EVOLTREE ELAB - An information system for forest genetics

François Ehrenmann, Stephan Gaubitzer, Dieter Kopecky, Schmidt Johanna, Silvia Fluch, Eva Maria Sehr, Antoine Kremer

▶ **To cite this version:**

François Ehrenmann, Stephan Gaubitzer, Dieter Kopecky, Schmidt Johanna, Silvia Fluch, et al.. EVOLTREE ELAB - An information system for forest genetics. Evolution of trees and forest communities, PG Edition, 175 p., 2016, 978-2-9519296-3-9. hal-01603666

HAL Id: hal-01603666
https://hal.science/hal-01603666

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

evoltree

A EUROPEAN RESEARCH GROUP

Evolution of Trees and Forest Communities
Ten years of the EVOLTREE network

CELEBRATING 10 YEARS

EDITED BY ANTOINE KREMER, STEPHANIE HAYES AND SANTIAGO C. GONZÁLEZ-MARTÍNEZ

evoltree

A EUROPEAN RESEARCH GROUP

linking

Genomics
Genetics
Ecology
Evolution

To find out more about EVOLTREE and how to get involved in its activities or make the use of its services, please consult the website:
www.evoltree.eu

AUTHORS

Dolores Abarca
Sally N. Aitken
Florian J. Alberto
Ricardo Alía
Catherine Bastien
Thomas Bataillon
Egbert Beuker
Gil Bohrer
Thomas Boivin
Aurore Bontemps
Roland Brandl
Katharina Bräutigam
Jon R. Bridle
Christian Brochmann
Carlo Calfapietra
Malcolm M. Campbell
Bastien Castagneyrol
Stephen Cavers
María-Teresa Cervera
Jean-Louis Churin
Cyrille Conord
François Courbet
Pierre-Emmanuel Courty
Hendrik Davi
Paolo De Angelis
Carmen Díaz-Sala
Marion Durand-Gillmann
François Ehrenmann
Bruno Fady
Mario Fernández Fraga
Silvia Fluch
Carl G. Fossdal
Alain Franc
Stephan Gaubitzer
Julie Gauzere
Sophie Gerber
Cindy Gidoin
Richard Gomulkiewicz
Santiago C. González-Martínez
Delphine Grivet
M. Ángeles Guevara
Felix Gugerli
Frédéric Guillaume
Jessica Gurevitch
José Gutiérrez Marcos
Heikki Hänninen
Stephanie Hayes
Hervé Jactel
Øystein Johnsen
Marie-Joe Karam
David F. Karnosky
Etienne K. Klein
Hans-Peter Koelewijn
Annegret Kohler
Dieter Kopecky
Jan Kowalczyk
Antoine Kremer
Anna Kuparinen
Jessy Labbé
Clément Lafon-Placette
Hadrien Lalagüe
Małgorzata Latałowa
François Le Tacon
François Lefèvre
Thomas Lenormand
Yunan Lin
Benoit Marçais
Francis Martin
Jan Matras
Stéphane Maury
Franco Miglietta
Marie Mirouze
Ran Nathan
Sylvie Oddou-Muratorio
Christophe Orazio
Martina Peter
Christian Pichot
Karin Pritsch
Kermit Ritland
Juan J. Robledo-Arnuncio
Alistair Rogers
Antje Rohde
Ophélie Ronce
Outi Savolainen
Johanna Schmidt
Hilke Schröder
Silvio Schueler
Federico Sebastiani
Eva Maria Sehr
Marinus J.M. Smulders
Christoph Sperisen
Michael Stierschneider
Steven H. Strauss
Nathaniel R. Street
Matthew James Tallis
Gail Taylor
Mari Mette Tollefsrud
Sara Torre
Inge van Halder
Giovanni G. Vendramin
Marc Villar
Kelly J. Vining
Willem O. van der Knaap
Otto Weichselbaum
Ross Whetten
Sam Yeaman
Mario Zabal-Aguirre
Jing Zhang
Birgit Ziegenhagen
Martin Ziehe

Photos: Fotolia, Bioversity, EFIATLANTIC, EVOLTREE.

Citation: Antoine Kremer, Stephanie Hayes and Santiago C. González-Martínez, Editors (2016). Evolution of Trees and Forest Communities: Ten Years of the EVOLTREE network. Document prepared by the EVOLTREE network. PG Edition - Bordeaux. 192 pages

ED. IEFC ISBN : 978-2-9519296-3-9

EVOLUTION OF TREES
as drivers of terrestrial biodiversity

CONTENTS

AUTHORS ..... Cover 2

FOREWORD 2

INTRODUCTION 3

EVOLTREE INFRASTRUCTURES, RESOURCES AND INITIATIVES 4

- Intensive Study Sites (ISS) 6
- Repository Centre 10
- eLab 15
- Training 20
- TreeType 23

RESEARCH HIGHLIGHTS: SCIENCE THAT MATTERS 26

- Potential for evolutionary responses to climate change - evidence from tree populations 27
- Epigenetic regulation of adaptive responses of forest tree species to the environment 45
- Large-scale longitudinal gradients of genetic diversity: a meta-analysis across six phyla in the Mediterranean basin 61
- Effect of poplar genotypes on mycorrhizal infection and secreted enzyme activities in mycorrhizal and non-mycorrhizal roots 75
- Molecular footprints of local adaptation in two Mediterranean conifers 88
- Community genetics in the time of next-generation molecular technologies 105
- Long-distance gene flow and adaptation of forest trees to rapid climate change 115
- Considering evolutionary processes in adaptive forestry 133
- The transcriptome of Populus in elevated CO_2 reveals increased anthocyanin biosynthesis during delayed autumnal senescence. 150
- Late Quaternary history of North Eurasian Norway spruce *(Picea abies)* and Siberian spruce *(Picea obovata)* inferred from macrofossils, pollen and cytoplasmic DNA variation 164

EVOLTREE PARTNERS 175

FOREWORD

EVOLTREE is ten years old this year (2016); but the idea of a European network in forest genetics dates well back before 2006. All current forest research involving genetics or genomics is transnational and it has become progressively apparent that better coordination of efforts and means is necessary if these issues are to be adequately addressed on a European level.

The forest genetics research community had felt the same need for a long time, but did not make the most of European Union funding opportunities at the beginning of the eighties.

It is unfortunate that pan-European initiatives, such as the comparison of provenances in plantations in the fifties and sixties, were not coordinated on a long-lasting scale in order to be able to answer current questions on assisted migration.

This historic reference highlights the necessity of coordinating research on forest genetics on a European level. The example of provenance tests is a past illustration of current needs in terms of means sharing in the field of genomics, the accessibility of as much biological and electronic resources as possible and the coordination of research efforts.

It is EVOLTREE's aim to rise up to this challenge. And since its creation ten years ago, this aim has been reaffirmed from year to year. We wish to celebrate the tenth anniversary by sharing the knowledge and results that we have accumulated during this period.

This book contains a selection of ten articles from the 165 listed on the Web of Science, which have been funded either by the EVOLTREE network of by other EVOLTREE initiatives, such as FORESTTRAC. It also describes the main physical and electronic infrastructures and other resources and initiatives, from which the forest genetics research community, in particular the network's present members can benefit.

ANTOINE KREMER
EVOLTREE Network Coordinator

INTRODUCTION

It all started in 2006, when twenty-five Universities and research institutes from fifteen European countries joined forces to set up EVOLTREE as a Network of Excellence. EVOLTREE aimed to link four major disciplines Ecology, Genetics, Genomics and Evolution to address global issues faced by European forests, such as environmental changes and the erosion of biodiversity.

Funded by the European Union within the 6th framework programme, it spent the next four years developing and setting up the necessary experimental and monitoring infrastructures and physical and electronic resources upon which long term research could be built.

These network infrastructures and resources were therefore well in place and up and running when EVOLTREE embarked on its next four year period in 2011. Twenty-three research groups from thirteen European countries agreed upon and signed a new consortium agreement. Now self-funding (based on financial contributions from some partners and "in-kind" contributions from others in the form of running relevant scientific training courses, for example), EVOLTREE was integrated into the European Forest Institute's (EFI) network.

Instrumental in EU projects (such as NOVELTREE, PROCOGEN, FORGER, NOVELTREE, TreesForFUTURE and GENTREE) EVOLTREE's research activities address topical issues such as the discovery of genes with economic and ecological relevance and the evaluation of their genetic diversity in natural tree populations and associated species, as well as the evolution, conservation, restoration, breeding and management of tree populations subject to environmental change and human interference.

EVOLTREE offered its resources to European projects, but also benefited from contributions of these projects by populating existing databases or upgrading existing infrastructures. During this period, EVOLTREE stimulated the organisation of workshops and summer schools on dedicated technical or broader issues related to EVOLTREE's interests.

It was in the latter part of the second period of the network (2011-2014) that the open-science research initiative "TreeType" was created for the widespread collection of data on simple (but very relevant) phenotypic traits for European trees. Such data can provide important insights into the balance between local adaptation and phenotypic plasticity in tree populations. Indeed, at a time when large DNA sequence data of individual trees are becoming available, the missing component is standardised phenotypic data.

The project is open to participation by anyone with the enthusiasm and skills to record the data for the trees of their choice and data will be made openly available. The recording website (accessible via the EVOLTREE website) was launched at the end of 2014 in time for its utilisation in the current four-year term 2015-2018.

The current four year term (2015-2018) will see TreeType taking off and becoming a valuable long term resource, alongside the EVOLTREE e-resources, the DNA repository centre, the ISS field network and the training courses. Furthermore, EVOLTREE intends to increase its involvement in EU projects; for example, it will be involved in FORESTING, a networking research infrastructure for forest ecosystem and resources research in Environmental and Earth Sciences.

EVOLTREE's ambition to widen and strengthen the network has been achieved by welcoming new partners from different parts of Europe to the consortium and by increasing synergies with EFI. The summer schools of the previous terms have been incorporated into wider training programmes to include all year-round workshops and EVOLTREE now endeavours to become more present in, and to organise its own, scientific events.

Therefore, at the age of ten, EVOLTREE is ready to take on new challenges and further its contribution to the field of genetics and genomics in Europe with fresh projects and ideas, and enthusiastic support and input from its partners and the forest genetics research community at large.

EVOLTREE infrastructures, resources and initiatives

INTENSIVE STUDY SITES

F. LEFÈVRE[1], C. PICHOT[1], E. BEUKER[2], J. KOWALCZYK[3], J. MATRAS[3], M. ZIEHE[4], M. VILLAR[5], M. PETER[6], F. GUGERLI[6], C. ORAZIO[7], R. CORDERO[7], R. ALIA[8], I. VAN HALDER[9], S. C. GONZÁLEZ-MARTÍNEZ[9], A. KREMER[9].

1 URFM, INRA, 84914 Avignon, France; 2 LUKE, 58450 Punkaharju, Finland; 3 Dept. of Silviculture and Genetics of Forest Trees, IBL, 05-090 Raszyn, Poland; 4 Dept. of Forest Genetics and Forest Tree Breeding, University of Goettingen, 37077 Goettingen, Germany; 5 AGPF, INRA, 45075 Orléans, France; 6 WSL Swiss Federal Research Institute, 8903 Birmensdorf, Switzerland; 7 EFIATLANTIC, 33610 Cestas, France; 8 CIFOR, INIA, 28040 Madrid, Spain; 9 BIOGECO, INRA, University of Bordeaux, 33612 Cestas, France.

The Intensive Study Sites (ISS) are large-scale ecosystem plots of a few thousands of hectares, where large samples of trees and associated species are progressively mapped, genotyped and phenotyped. The sites comprise entire portions of landscapes, where trees are present in different configurations from single trees to edges and woods.
More detailed description of the ISSs, their organisation and their information system are available on the EVOLTREE website.

Background and objectives of the ISS infrastructure

The recent revolution in high throughput technologies and methods in genetics has drastically changed the research perspectives and practices in ecology and evolutionary sciences. The availability of large genetic datasets, for a broad range of species far beyond the few so-called "model species", together with the development of appropriate analytical tools, has transformed the genomes of living forms into immeasurable sources of information on the ecological and evolutionary processes that shaped biodiversity, at different spatial and temporal scales.

A side effect of this revolution has been the emergence of common tools and methods shared by different branches of life sciences that had previously tended to diverge, such as genomics, population genetics, quantitative genetics, functional ecology and community ecology. All these interconnected disciplines were represented within EVOLTREE.

The EVOLTREE scientific community recognised the considerable advance in forest genetics and ecology based on these tools and methods. However, the projects that followed these different lines of research were conducted independently from each other, highlighting the need for shared experimental infrastructures. By using shared infrastructures it is possible to accumulate genetic information from trees and their associated species with environmental data and other information of interest collected in different projects. They also help to address the effects of ecological processes and human activities on forest systems on the relevant spatial scales.

Furthermore, the EVOLTREE scientists were convinced that innovative knowledge integration across disciplines often occurs *a posteriori* and does not always result from *a priori* planning. Thus, setting up adequate tools for sharing the information acquired by different research groups within the same experimental field sites seemed to be a priority.

The long-term EVOLTREE Intensive Study Sites (ISSs) infrastructure was therefore created with the following five main objectives:

1. To set up a European network of representative sites for long term research on the evolution of biodiversity in forest ecosystems at different hierarchical levels (from genes to phenotypes, from populations to communities) and with different management options.
2. To assess the spatial structure of biodiversity on various scales and at different hierarchical levels.
3. To monitor population dynamics in trees and their associated species, using demographic and genetic approaches, over different spatial scales.
4. To monitor the interaction between species (mainly trees, other plants, insects, and microorganisms).
5. To provide long-term and large-scale support for training, education and dissemination activities.

Each ISS has intrinsic value, but has also added value in comparison with other ISSs, as they were chosen to build a network of representative forests in Europe. The ISSs are used in two ways. First, due to the local heterogeneity and gradients, the impact of the environment or management practices on the dynamics of diversity within each type of ecosystem can be studied. Beyond the "natural" local heterogeneity, such as altitudinal gradients, some ISSs also host short-term and long-term experiments such as irrigation, reciprocal transplantations, or different sylvicultural treatments. Second, the drivers of diversity change in different ecological regions can be compared across the ISSs.

ISS integration in EVOLTREE research

Since the beginning of EVOLTREE, thirty-two international and national projects[1] have made use of the ISS infrastructure (Table 1). As explained above, the ISSs were selected - amongst other criteria - on the basis of research activities carried out prior to the EVOLTREE Network (not mentioned here). During a first phase (i.e., that benefitting from financial contributions from the European Commission), EVOLTREE was directly funding research activities carried out within the ISSs. Hence, international partnership was mandatory and priority was given to projects that involved multiple sites. Since 2011 (i.e., after EVOLTREE became a European Research Group without EC financial support), eighteen new projects funded by other sources[2] used the ISSs: national and mono-site projects, mainly based on local research groups, but also on international and multi-site projects.

These figures reflect a long-lasting interest in the ISS infrastructure and that the ISS network was not only useful to local groups, but also to others from a wide range of European research groups.

The research projects have made use of the ISSs in various ways connected to some of the infrastructure's five main objectives (listed above): simply as a reservoir of biological samples; as natural sites for observation and monitoring; and as appropriate ecological settings to establish experiments.

The research supported by the ISSs so far has addressed a broad range of scientific questions in various fields of evolutionary science and ecology represented within EVOLTREE; for example, the characterisation of genetic and genomic diversity in trees and associated species; the detection of genes involved in local adaptation and biotic interaction; the characterisation of local adaptation patterns and processes on various spatial and temporal scales; the assessment of the functional, demographic and genetic response to climate change on individual,

THE ISS INFORMATION SYSTEM

The ISS Information System is designed to make the data collected in each ISS available. It uses metadata that follow the ISO 19115/19139 standard and are compliant with the EU directive INSPIRE. It proposes different ways to search for metadata: via geographic location, data categories (e.g., maps, datasets, pictures) or keywords selected from a dedicated thesaurus developed by the ISS partners. As well as general information about the sites, the system holds references to publications and information about research activities, permanent plots and transects, and permanent samples of individual trees georeferenced and tagged in the forest (from which DNA is available in the Repository Center).

1• A detailed list of these projects is available on the EVOLTREE website, and more information is accessible through the ISS information system
2• EVOLTREE offers mobility grants to support activities in the ISSs

FIGURE 1

In order to restrict the research activity to a few sites, we initially selected seven ISSs. An eighth site was added recently (2015). In a bottom-up process, potential sites were proposed by the EVOLTREE community and the ISSs were selected based on two sets of criteria:

Selection criteria based on scientific characteristics

- Within-site diversity of tree species, associated species, communities, population structures.
- Within-site comparability, i.e., temporal or spatial heterogeneity: historical records, environmental variability, diversity of management, comparative experiments.
- Within-site research history and available datasets, e.g., ecological records, climatic data, genetic data, management data.
- Potential for species comparisons across ISSs: presence of tree species in common with other ISSs.
- Network coverage of the range of terrestrial forest ecosystems across Europe: boreal, temperate, alpine, Mediterranean, riparian, untouched and intensively managed forests.

Selection criteria based on infrastructure characteristics

- Long-term perspectives, e.g., ownership of the land, legal and/or protection status, relation to other networks, interest for local managers, education and dissemination.
- Technical facilities, e.g., access to the site and to the biological samples.
- Expertise of the local partner institution and relationship with the local managers.

The eight ISSs are, from North to South (Figure 1):

- **PUNKAHARJU**, Finland, Northern temperate and boreal forest (1,500 ha, includes Abies, Acer, Alnus, Betula, Larix, Picea, Pinus, Populus, Prunus, Quercus, Salix, Sorbus, Tilia)
- **BLIZYN,** Poland, Continental temperate untouched forest (17,000 ha, includes Abies, Acer, Alnus, Betula, Carpinus, Corylus, Crataegus, Fagus, Fraxinus, Larix, Picea, Pinus, Populus, Prunus, Quercus, Salix, Sorbus, Tilia, Ulmus)
- **SOLLING**, Germany, Continental temperate managed forest (25,000 ha, includes Betula, Fagus, Larix, Picea, Pinus, Quercus)
- **LOIRE,** France, riparian forest (96 ha, includes Acer, Alnus, Castanea, Corylus, Crataegus, Fraxinus, Populus, Prunus, Quercus, Salix, Tilia, Ulmus)
- **VALAIS,** Switzerland, montane to alpine forest (150,000 ha, includes Abies, Acer, Alnus, Betula, Carpinus, Corylus, Crataegus, Fagus, Fraxinus, Larix, Picea, Pinus, Populus, Prunus, Quercus, Salix, Sorbus, Tilia, Ulmus)
- **LANDES,** France, oceanic intensively managed forest (25,000 ha, includes Alnus, Betula, Castanea, Corylus, Crataegus, Fagus, Fraxinus, Pinus, Prunus, Quercus, Salix, Sorbus)
- **VENTOUX,** France, Mediterranean and South alpine forest (29,000 ha, includes Abies, Acer, Alnus, Betula, Castanea, Carpinus, Cedrus, Corylus, Crataegus, Fagus, Larix, Picea, Pinus, Quercus, Sorbus)
- **CALDERONA-ESPADÁN,** Spain, Mediterranean forest (49,000 ha, includes Celtis, Ceratonia, Juniperus, Olea, Pinus, Quercus, Salix)

population and community levels; and the modeling of adaptive processes and responses to management practices. The ISSs also support projects developing multidisciplinary approaches with other scientific fields, including environmental and social sciences.

What will the future bring?

The use of the ISSs as a research infrastructure has come of age. Beyond its role in supporting national research projects, the ISS infrastructure will continue to strengthen the long-lasting integration of research in the field of forest genetics and genomics on the pan-European level, e.g., via the H2020 research and innovation project GENTREE[3] "Optimising the management and sustainable use of forest genetic resources in Europe" (2016-2020). Moreover, projects linking the ISSs with other long-term research infrastructures on forest ecology are underway, thus providing wider integration of multi-disciplinary knowledge on forest ecosystems on a pan-European scale.

An innovative use of the infrastructure is planned in the participative project TreeType[4], in which citizens, researchers or forest managers can contribute to collecting phenotypic data on individual trees, aiming to characterise the phenotypic variation in the forest. These data will then be analysed by the scientists to provide information on the genetic basis of traits and create a database for future research. This initiative will also support training and dissemination activities by providing easy-to-use protocols and tools to study adaptation in the wild that can be used as educational materials.

One of the major challenges for the future will be to assemble and share the extensive knowledge, results and original data that have been generated in the ISSs throughout the years. This contribution to "Open Science" will be facilitated by the regularly updated ISS information system (see Box) leading to the publication of datasets in data papers for the benefit of scientific progress.

3• Project website not yet available, see http://cordis.europa.eu/project/rcn/200286_en.html
4• www.treetype.org/, see specific review in this volume

TABLE 1

Number of projects conducted in the ISS since the beginning of EVOLTREE

	Mono-site projects	Multi-site projects
EVOLTREE Phase I (2006-2010)		
International partnership	3	10
EVOLTREE Phase II (2011-current)		
International partnership	3	4
National partnership	11	0
TOTAL (2006-current)		
International partnership	6	14
National partnership	11	0

ISS COORDINATION

A Framework of Agreement defines the organisation of the ISSs, data and metadata supply and access policy. Each ISS has a local coordinator in charge of managing the ISS Information System and facilitating research activities, by interacting with local managers or helping with local logistics. In order to be able to access the infrastructure resources (data, samples, experiments), partners must accept the Agreement and contact the ISS coordinator before submitting the project.

All photos: EVOLTREE partners

THE EVOLTREE REPOSITORY CENTRE
A CENTRAL ACCESS POINT FOR REFERENCE MATERIAL AND DATA OF FOREST GENETIC RESOURCES

Michael Stierschneider[1], Stephan Gaubitzer[1], Johanna Schmidt[1], Otto Weichselbaum[1], Dieter Kopecky[1], Antoine Kremer[2], Silvia Fluch[1] and Eva Maria Sehr[1].

1 AIT Austrian Institute of Technology GmbH, Health & Environment Dept., Konrad-Lorenz-Street 24, 3430 Tulln, Austria
2 INRA, UMR1202 BIOGECO, Cestas F-33610, France

The Repository Centre is a centralised storage facility that hosts biological and genetic resources and corresponding metadata collected by EVOLTREE partners, to be used for research upon request by any interested lab.

It deals with resource management ranging from whole organisms (plant and animal material, bacteria, and fungi) to DNA samples of whole genomes, organellar and sub-genomic clones and BACs or genes (ESTs).

Goal and objectives

For many decades, the importance of reference material for various types of plant based research from ecology to phylogenetic systematics has been acknowledged; whenever results need to be compared, the availability of reference material plays a critical role. Thus collections of reference material for plant breeding, biotechnology and biodiversity analyses have been established all around the world. These have traditionally been based on living plant material managed *ex situ* in seed banks, botanic gardens or *in situ* in nature reserves.

With the advent of DNA based analytics, a new and very powerful tool was uncovered to unlock the potential of the stored biological material. After all, genomic DNA samples represent the entire genetic information of an organism, from various resistance traits to phenotypic parameters, as well as records of their inherited characteristics and ancestral roots. Thus, since the discovery of the double helical structure of the DNA in the early 1950s (Watson and Crick, 1953), molecular data based on DNA sequences have become increasingly important for a variety of biological disciplines, including systematics, ecology, evolution and population genetics, genetic diversity assessment and data generation used as a basis for nature conservation.

Nowadays, the extraction of genomic DNA is easily done, even in high-throughput, and once purified, DNA can be stored for long periods of time. As a result, DNA based collections have become increasingly important, with biobanking in the human sector being the most prominent one; but natural history museums have also started engaging in uncovering the value of their historic collections by extracting DNA from various dried samples which have been collected and preserved for centuries (Droege *et al.* 2014).

Photos: AIT

However, the key to success is standardisation. Standardised tools and techniques are increasingly applied in diverse initiatives, such as the 'International Barcode of Life' (iBOL) which entirely relies on DNA information generated from reference samples collected throughout the world, aims to reconstruct the phylogenetic relationships in the Tree of Life This project can only succeed if all the partners apply the same techniques and have access to the same reference DNA, as well as collected data.

In order to meet the demands of future high-impact research, reference material and larger sets of DNA samples, as well as corresponding data, have to be made widely accessible. Particularly in the area of ecological research, where findings are based on the fact that population genetic patterns are being compared across borders and over large geographical distances and gradients, it is essential that researchers have access to reference material and to the respective data generated from this material. Such biological material, data and tools

are needed in a standardised and freely accessible way in order to guarantee comparability of research results across Europe. When dealing with forest trees, sample collection is a laborious and time-consuming task, involving long trips to remote areas, as well as dangerous or difficult situations when accessing the material (e.g., tree climbing, shooting down twigs); therefore, the sharing of this material and generated data leads to more economical and time saving research.

In view of this, one objective of the EVOLTREE network was to build up a centralised and standardised storage and management facility, known as the Repository Centre. By storing research material at one physical site, the aim was to generate high impact in forest research in the disciplines of ecology, genetics, genomics and evolution, not only in Europe and during the course of the project, but also as an international reference site for forest genetic resources.

Description of the facility

The Repository Centre gathers together all the available and dispersed research material in one reliable site and provides open-access to a continuously growing data-set. This high quality material is available to and provided by EVOLTREE partners and researchers outside the network.

Due to the huge number of samples which were already available at the EVOLTREE partner sites, special equipment for storing, managing and tracing material, together with a database for storage of all the corresponding (meta-) data, were prerequisites for the physical installation of the Repository Centre. The installation process for creating flexible and highly reliable workflows for DNA extraction, quality control and long-term storage at the Repository Centre laboratory was initiated in 2006. The goal was to have genomic DNA of populations, e.g., from the EVOLTREE Intensive Study Sites (Lefèvre *et al.* page 6, this book) as well as gene bank collections extracted and stored following the same extraction procedures. ESTs had to be available in the form of single clones, as well as spotted on micro arrays, in order to conduct large scale expression profiling in natural populations of non-model species. The most important features were guaranteed sample integrity, standardised quality of material and data, and sustainable and easy access to the material. To fulfil these requirements, a modular -20°C/-80°C fully automated storage system with a capacity for 11,230 microtiter plates at -20°C (Universal Store US-450; Nexus Biosystems) and 1,000 microtiter plates at -80°C was installed (BioBank™; Thermo Scientific).

The storage system comes with online monitoring and logging of the temperature status and includes an internal alarm system via SMS in case of any temperature failures. The redundant refrigeration systems are assembled outside the building; each system being able to hold the set temperature without the second one.

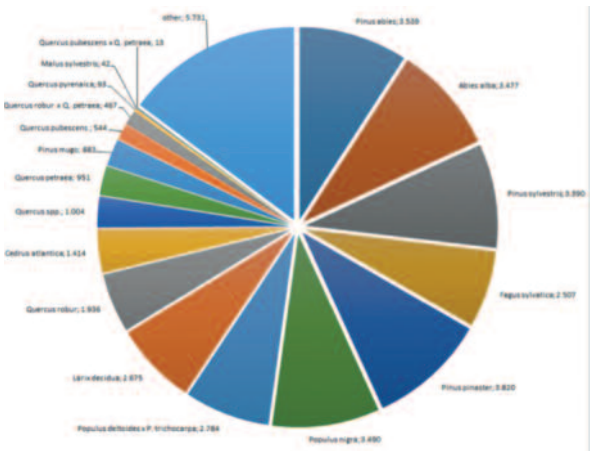
Barcoded, heat sealed microtiter plates are stored on blue barcoded trays (6 plates/tray) which are positioned randomly within a shelf system. A rail-based robotic arm stores or retrieves the plates on request and plates within the trays can be rearranged by a separate plate picking module under freezing conditions to prevent needless sample thawing. Plate picking and re-insertion is logged by a database system

Automation

To manage and distribute several thousand DNA resources, a high level of lab automation is required to guarantee failure-free sample and data handling, combined with high throughput and high quality. A newly introduced laboratory data management system, called *Material Administration and Preparation System* (MAPS™) serves as the backbone of the quality management system (Kopecky *et al.* 2012). It consists of a database which holds information about the lab processes and a web-based user interface for viewing and editing this information. Once the initial necessary Excel-based provider data has been imported into the system, the whole workflow being undertaken in the laboratory at the time (DNA extraction, electrophoresis images, PCR results, storage positions, DNA

FIGURE 1

Pie chart of the stored gDNA samples



quantity/quality, and results of downstream analysis), together with all needed supply information (e.g., barcodes, volumes, well IDs), is logged within the system and can be retrieved accordingly.

Genomic DNA (gDNA) extraction protocols have been optimised for semi-automated (whereby the pipetting steps are run automatically and the centrifuging steps manually), high-throughput processing for 96-well format, using liquid handling platforms like *Tecan* or *Hamilton*. Supported by these infrastructures, the Repository Centre is able to execute up to 960 gDNA extractions per day.

Protocols for a wide range of raw plant material like leaves, needles, buds, roots, cambium, and wood have been established. In order to guarantee access to the reference material over time, lyophilised plant material is stored at room temperature alongside 2 x 100 µl of gDNA extracted from this material at concentrations of approximately 50 ng/µl. One of these 100 µl is located in the working copy deposited in the automated storage facility, whereas the second copy is used as a backup plate stored in another section of the building. This is due to the requirement of risk mitigation – so that in the case of any sort of accident (e.g., fire, system break down, etc.) which could destroy the working copy of a sample, the backup copy is still available for use.

As soon as the gDNA extraction is finished, an aliquot (mostly 1 µl) of each sample is verified and documented via agarose gel electrophoresis. Analysis of the gel images is carried out automatically using proprietary image analysis software (Bajla *et al.* 2005). This software enables the analysis of 96 individual samples (loaded on one gel) in parallel. Pixel intensities are measured and compared to a pre-defined standardised mass ladder, whereby the gDNA yield is calculated. The corresponding picture as well as the estimated gDNA concentration and size are saved back to MAPS™.

In the case of failed samples, or samples showing degraded gDNA, the sample state is set to «failed» within MAPS™, and these samples will not be available for downstream analysis. This way, the quality and amount of gDNA is documented for each and every sample so that in the case of any problems occurring in downstream analyses, the data quality data of a sample can be retrieved and checked by the user directly via a central search portal.

At the request of a customer, single samples can be selected out of 96- or 384-well microtiter plates, a process known as cherry-picking. The requested samples are often distributed over a larger number of microtiter plates and their picking necessitates an error free workflow. To fulfil these requirements, worklists are generated by MAPS™ which are generated within the software and sent to the liquid handling platforms directly. The pipetting process monitored by barcode tracking starts according to the worklists and the generated log files provide information about possible sample manipulation at any time.

Capacity

At the time of writing, 27,073 gDNA samples (Figure 1), 485,893 cDNA samples (Figure 2), 202,752 BACs, and 26,329 source samples (Figure 3) from various organisms of forest genetic resources and allied species (e.g., moths, caterpillars, fungi) are managed in the Repository Centre (Table 1). Of the capacity of 11,230 microtiter plates, currently 9,569 plates in both formats, 96- and 384- well, are stored. Most of the gDNA and tissue samples are stored in duplicates.

FIGURE 2

Pie chart of the stored cDNA samples

Out of a total of 485,593 cDNA samples, 309,093 were generated in the framework of EVOLTREE.

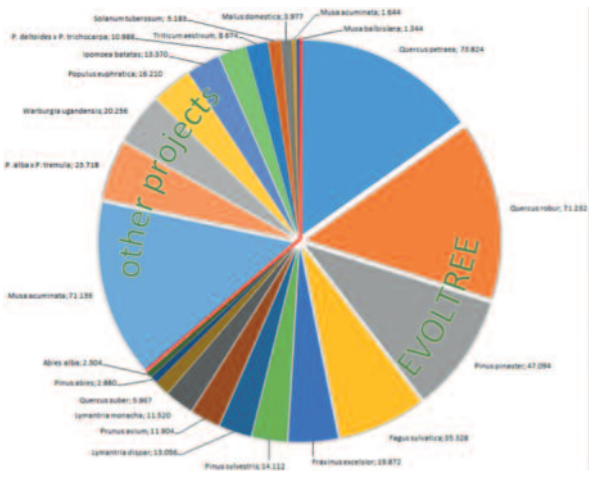
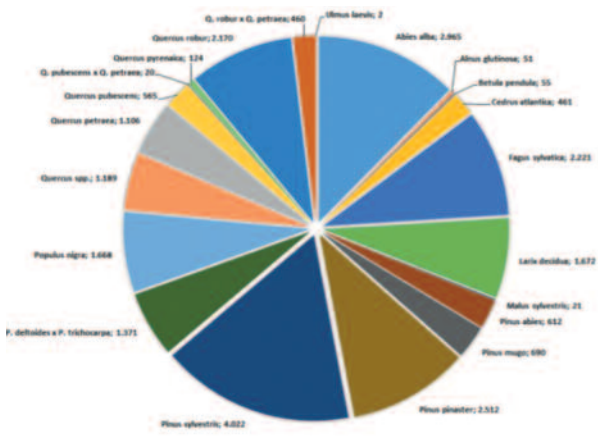


FIGURE 3

Pie chart of the stored tissue material



Data management and access

Databases are critical for efficient sample and data management, as well as for the efficient end utilisation of the DNA bank itself. To allow a quick and easy retrieval of a maximum amount of associated information (e.g., tissue type, collection conditions, date of collection, provider information, sample preparation, details on DNA quality and quantity), a tailor-made LIMS (Laboratory Information Management System) was developed in the framework of EVOLTREE to fulfil the demands of the Repository Centre. MAPS™ (Material Administration and Preparation System), a new concept of LIMS, meets these requirements with an innovative data model, as well as a modern service-oriented architecture, on top of state-of-the-art web technologies (Kopecky *et al.* 2012).

MAPS™ reflects all workflow steps in the laboratory and provides possibilities for recording these steps electronically. MAPS™ further communicates with the storage system via an application programming interface (API) based on the *Java Messaging Service* (JMS). This way the end user can access the organisation and reporting capabilities of the storage device seamlessly, without the need to work with two different applications in the laboratory. The user interface of MAPS™ has been implemented in a web-based way in *Java*. Access to the system therefore only requires a web-browser and no other specific software needs to be installed, so it is very easy to access MAPS™ from any computer available in the laboratory. The information in MAPS™ is recorded in a *PostgreSQL* database on a central server and regular backups of the data ensure its necessary integrity. Data access is only granted to laboratory workers and the administrator via user accounts with respective user rights. Should communication with external partners be required, MAPS™offers an interface based on the *Simple Object Access Protocol* (SOAP), which provides standardised services for querying information about samples stored within the LIMS, as well as for adding new samples to the system. These services can also be made available across organisational borders, so that customers are able to query information about the stored samples via a central search portal, known as the eLab (Ehrenmann *et al.*, page 15, this book). Due to security restrictions, however, the MAPS™ service, especially the storage system interface, is located in an internal network not reachable from the outside. The eLab services cannot directly access the MAPS™ databases, since they are located at the same institution (although in a different network). The MAPS™ database is therefore treated as an external database resource by the eLab that needs to be queried at regular intervals. The queried information is then inserted into the eLab search system, where it can be accessed by the users (Kopecky *et al.* 2010).

Current and future use

Besides its central role as a storage and retrieval system for biological material, DNA, and data in the framework of the EVOLTREE network serving forest ecosystem research, the Repository Centre became part of the Trees4Future (http://www.trees4future.eu/) project. This was an Integrative European Research Infrastructure project that integrated forest tree breeding infrastructures to improve and enhance gains in the area of European forest tree breeding. In this framework, the repository centre and its links to the EVOLTREE data collections serves as an integrative hub for European forest ecosystem research and European tree breeding efforts.

TABLE 1

Overview of the sample types and number of samples stored in the Repository Centre

Sample type	Nr. of samples stored
Source material	26,329
gDNA	27,073
cDNA	485,593
BAC	202,752

This way, the Repository Centre is currently the largest data provider to the Global Genome Biodiversity Network (GGBN) – a network of repositories of genomic sample collections aiming at allowing open access to reference material from botanical gardens as well as natural history museums. Data of about 53,402 samples, divided into 27,073 gDNA and 26,329 tissue samples are accessible via the GGBN portal (http://www.ggbn.org/) following the ABCD standard used by the BioCASe provider software (Holetschek *et al.* 2012). In addition to this, all georeferenced EVOLTREE repository data sets can be found on GBIF, the Global Biodiversity Information System (GBIF, http://www.gbif.org/).

Being part of these networks enhances visibility and contributes to the ever growing idea of open-access resources and data in order to further the development of DNA based forest ecosystem research. As can be seen by all these activities, and as also requested by the European Commission, open data and open material initiatives will help future research to better integrate research results and to enhance forest ecological understanding on a large scale. By integrating genetic, as well as environmental data, we will be able to generate forest systems modelling approaches that will allow a more general understanding of the complex ecosystem forest and will help to mitigate the impact of climate change on one of the most important socioeconomic factors in Europe.

REFERENCES

Bajla, I., Holländer, I., Fluch, S., Burg, K., Kollár, M., 2005. An alternative method for electrophoretic gel image analysis in the GelMaster software. Comput. Methods Programs Biomed. 77, 209–231. doi:10.1016/j.cmpb.2004.09.007

Droege, G., Barker, K., Astrin, J.J., Bartels, P., Butler, C., Cantrill, D., Coddington, J., Forest, F., Gemeinholzer, B., Hobern, D., Mackenzie-Dodds, J., Ó Tuama, É., Petersen, G., Sanjur, O., Schindel, D., Seberg, O., 2014. The Global Genome Biodiversity Network (GGBN) Data Portal. Nucleic Acids Res. 42, D607–D612. doi:10.1093/nar/gkt928

Holetschek, J., Dröge, G., Güntsch, A., Berendsohn, W.G., 2012. The ABCD of primary biodiversity data access. Plant Biosyst. - Int. J. Deal. Asp. Plant Biol. 146, 771–779. doi:10.1080/11263504.2012.740085

Kopecky, D., Schmidt, J., Fluch, S., 2010. Large-scale integration of distributed heterogeneous data sources within Evoltree, in: Proceedings of the EVOLTREE Conference on Forest Ecosystem Genomics and Adaptation. San Lorenzo de El Escorial, Madrid, Spain.

Kopecky, D., Weichselbaum, O., Fluch, S., 2012. MAPS – A service-oriented, customizable, multi-purpose LIMS. Nat. Methods Appl. Notes.

Watson, J.D., Crick, F.H., 1953. Molecular structure of nucleic acids. Nature 171, 737–738.

EVOLTREE ELAB - AN INFORMATION SYSTEM FOR FOREST GENETICS

F. EHRENMANN[1], S. GLAUBITZER[2], D. KOPECKY[2], J. SCHMIDT[2], S. FLUCH[2], E. MARIA SEHR[2], A. KREMER[1].

1 INRA, UMR1202 BIOGECO, Cestas F-33610, France; 2 AIT Austrian Institute of Technology GmbH, Health & Environment Dept., Konrad-Lorenz-Straße 24, 3430 Tulln, Austria.
* These authors contributed equally to the construction of the eLab and its components.

The EVOLTREE web portal acts as a platform for information and data storage, retrieval, exchange and communication. It was set up in the early years of the network (between 2007 and 2008) in order to fulfill one of EVOLTREE's objectives of maintaining and reinforcing electronic and physical resources, repositories and infrastructures.
It comprises what is known as the "electronic Lab" (eLab) which was designed as a centralised search engine for databases that are stored in different servers located in different institutions in Europe. The web portal can be accessed via the EVOLTREE website[1].

Background and objectives of the eLab

While some of the portal databases existed beforehand, most of the web applications corresponding to the databases were constructed in the first four years of the EVOLTREE network. These have been continuously updated and populated ever since by the member laboratories which host the databases and carry out this activity as part of their 'in-kind' contribution to the network.

The portal databases are connected through a standardised, HTTP transmittable interface (TAPIR - www.tdwg.org/activities/tapir/), so that queries can be made within the whole set of databases.

Given the number of tree species studied throughout Europe, it was decided to "virtually" subdivide the eLab into three major portals corresponding to the three major botanical forest tree families that are studied: the Quercus Portal (for species belonging to the Fagaceae family), the Pinus Portal (for the Pinaceae family), and the Populus Portal (for species belonging to Salicaceae). Depending on their field of interest, users can therefore enter the system and make queries via three channels:

- The individual database for queries targeting well-focused information.
- The eLab for an overall search across all the databases. Access via the eLab research engine is recommended if users do not know where - e.g., in which database - the information of interest is located.
- One of the family portals, for data corresponding to a particular species, or genera, of the Fagaceae, the Salicaceae, or the Pinaceae family. Databases concerning species not belonging to these families can be directly accessed via the eLab.

Photo: BioGeCo, INRA

The individual databases

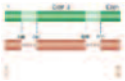
Passport, phenotypic, genetic and genomic data corresponding to different research units (genes, individuals, populations, species) were stored in separate databases, some of which existed before the launch of the network. At the beginning of EVOLTREE it was decided to keep the decentralized structure of the databases and to connect them via an interoperable interface in order to benefit from the already existing resources and the contributions of different partners.

Table 1 provides a summary of the content of the largest and most completed individual databases. All databases can be accessed individually and queries can be made internally without using the overall research engine of the eLab. A few databases offer some additional features and in some cases provide internal data analysis ; for example, genetic or QTL maps can be compared using Cmap. GD2 is dynamically linked with EUFGIS (http://portal.eufgis.org/), the georeferenced database of forest conservation units coordinated by EUFORGEN. The database connection is

1• The web portal can be accessed via the EVOLTREE website: www.evoltree.eu/index.php/e-recources/elab

TABLE 1

The main eLab databases and their content

Acronyms	Main content	Additionnal features	Access via eLab	Access via family portal
Map	Genetic and phenotypic records of trees belonging to mapping pedigrees.	Contains only data. Direct access to Cmap possible.	Yes	Yes, under the names of QuercusMap, PinusMap andPopulusMap.
Cmap	Position of markers and QTLs on genetic and QTL maps.	Comparison of different maps of different pedigrees.	Yes	Yes, under the the same name (Cmap) or all three families (Fagaceae, Pinaceae, Salicaceae).
Treepop TreePop	Genetic and phenotypic records of trees belonging to natural or unstructured populations.		Yes	Yes, under the the same name (Treepop) for all three families (Fagaceae, Pinaceae, Salicaceae).
Provenances	Passport, Genetic and phenotypic records of trees belonging to provenances established in provenance tests.	Also contains climatic data related to the provenance sources.	Yes	OakProvenance (exists only for oak).
GD2	Georeferenced data of allelic frequencies and diversity statistics in natural populations.	Is connected to the EUFGIS database.	Yes	Yes, under the same name (GD^2) for all three families (Fagaceae, Pinaceae, Salicaceae).
SSR	Sequences of microsatellites motifs and their flanking regions.		Yes	Only in the Quercus Portal.
SNP	Sequences of the contigcontaining the SNP and the two flanking regions.		Yes	No
Candidate genes	Sequences of candidate genes.		Yes	No
ESTs 	Expressed sequence tags of gene transcripts.		Yes	Yes, under the same name (EST) for all three families (Fagaceaea, Pinaceae, Salicaceae).

carried out using the TAPIR interface. The landscape of genetic diversity near conservation units can be drawn as a result of the connection between both databases, thus potentially helping to refine the setup of the conservation units.

Data access is controlled via user accounts and a hierarchy of roles is granted depending on user access rights. A high level of confidentiality is maintained using table fields with assigned values depending on the related user groups. Thus, data can be kept confidential and restricted to a particular group of users before publication.

Data input

In order to deal with the large amount of data created during the EVOLTREE project, it was decided to offer individual database-specific solutions for transferring data, using Excel or comma-separated text files (.csv). Most of the database-driven web applications enable EVOLTREE users and/or database administrators to manage data and use templates to import data of the same type (for example markers, populations or sequences).

Data output

EVOLTREE users can export data files (.csv) from the individual databases. Search menus are available in most of the database-driven web applications, with several criteria to avoid downloading issues for oversized datasets. The "download" buttons or sections are only available for registered users.

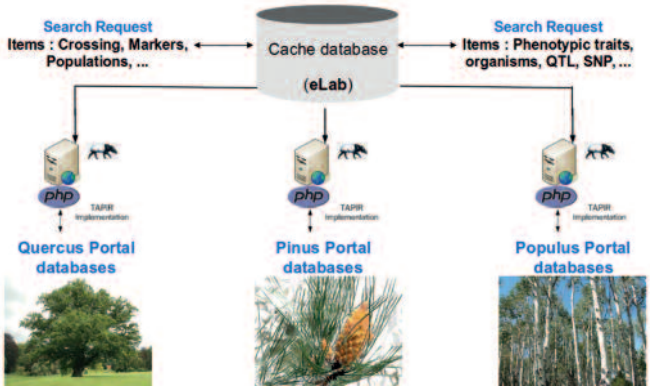
The eLab (electronic Lab)

The eLab gathers the data from all the individual EVOLTREE databases, and provides an interoperable web-interface for

FIGURE 1

General view of the eLab and its components

The eLab search interface consists of three parts:(1) a full-text search interface, (2) a guided search interface, and (3) a tree-view of the data. Within the full-text search interface, users can search for textual terms that occur somewhere in the cache database; whether the term defines a species name, a genus name, an annotation, a marker name or something else (e.g., a comment). The synonyms found during the standardisation process (see above) are also integrated into the full-text search, so that users can retrieve their own terms with their original names.



Data transfer is done over HTML using the TAPIR interface (http://www.tdwg.org/activities/tapir/). TAPIR is a XML-based protocol that can be used for information retrieval in distributed architectures. It is used to collect information from heterogeneous data sources in a pre-defined standardised format. TAPIR operates by harvesting all relevant information of the individual databases.

users to make queries against the data. Thus, the eLab functions as a clearinghouse mechanism enhancing the exchange of information and data throughout the different member labs of EVOLTREE. The relevant data parts of every database were defined when a new database was being integrated into the eLab.

The collected data is first stored locally and then transferred into a virtual cache database. This data collection is carried out at frequent intervals so that the latest information is always available in the cache. While transferring the data into the cache, the data is also merged into a standardised data format by using unique taxonomies; for example, different names for species (e.g., in English, German, French) are standardised so that only one name will be used in the cache database.

The implementation of a proper standardisation system played a major and important part in developing the eLab. The centralised search engine of the eLab only queries the cache database. Therefore, some specific information - only available in the individual databases and which was not considered to be relevant during the integration process - is not visible when using the eLab search engine. If users wish to access such additional information, their search will result in being redirected to the user interface of the corresponding individual database. During the redirection, the user information is encrypted when sent to the database (Figure 1).

Within the guided search interface, users can define more specific search queries. The existing data is presented in a web form and users can select their terms of interest (e.g., species, genus, institution, etc). It is also possible to refine queries further by selecting different pre-defined datatypes (e.g., genetic markers or population). The web-forms are updated dynamically when the selection changes. This way, users can, for example, search for all entries in the cache database that belong to a certain species, to a certain genus, or to a certain institution. The tree-view of the data represents a categorisation of the data available in the cache database. When transferring data into the cache, every data item is categorised according to pre-defined taxonomies. In the tree-view, users can browse through the hierarchical taxonomy and quickly find out how many, for example, data items for a certain sequence feature region exist. It is also possible to detect how many data items belong to a certain EVOLTREE partner.

In addition to the search tool, the eLab offers a reporting service. As they usually contain a large number of entries, the results are grouped together to get a better overview of the data. Each result entry is attributed a description to characterise it. If the user clicks on an entry, he will see all the information that is available for this entry in the cache database, which at this point may not be the "complete" data he is looking for. In order to view the "complete" data, the user can click on a second link and he will then be redirected to the external database the current result entry belongs to.

Photo: EFIATLANTIC

FIGURE 2

Main page of the Quercus Portal

The Quercus Portal comprises two sections:

A static section (left part of Figure 2) that provides general information regarding the biology, biogeography, phylogeny, botany and genetics of the botanical family and the different genera. The static page also comprises information about ongoing research and projects and links to their dedicated web pages.

A dynamic section that corresponds to the different databases related to the species or genera belonging to the Quercus family. They appear as different entry tabs in the headings of the webpage of the portal (upper part of Figure 2).

The family portals

To ease the queries of the user, the different databases were virtually subdivided into three families (Fagaceae, Pinaceae and Salicaceae); thus, users may directly enter one of the three portals (Quercus Portal, Pinus Portal or Populus Portal) and get direct access to the data they are looking for. As mentioned earlier, queries through the eLab retrieve the information stored in the cache database first and not the "complete" data stored in the individual database to which the user can be redirected; access via the portals is therefore much more rapid. As the different portals are designed in the same way, only the Quercus Portal is shown here, being the most complete at this stage (Figure 2).

The Quercus Portal has its own research engine (Global Search) which can be used to make queries across the databases hosted by the portal. An update of the current content of the different databases of the Quercus Portal is available in Table 2.

Current and future use of the eLab and the portals

A web analytic service has tracked and reported the EVOLTREE website traffic ever since the beginning of the network. From 2007 to 2015, 68,838 sessions were recorded by 37,576 users. On average, every time a person visited the EVOLTREE site (a single session), they looked at 4.45 pages for a total pages viewed of 307,000, and an average session duration of 2 minutes and 57 seconds.

While the main databases were constructed in the early years of EVOLTREE and the current portal structure was designed more recently, the main focus is now on the maintenance and regular updating of the databases. We anticipate, however, that very large data sets are still to come as a result of the development and applications of next generation sequencing (NGS) in population genomics of trees. Not all data collections corresponding to forest

FIGURE 3

Interoperability and data flow between information systems in the field of genomics and forestry

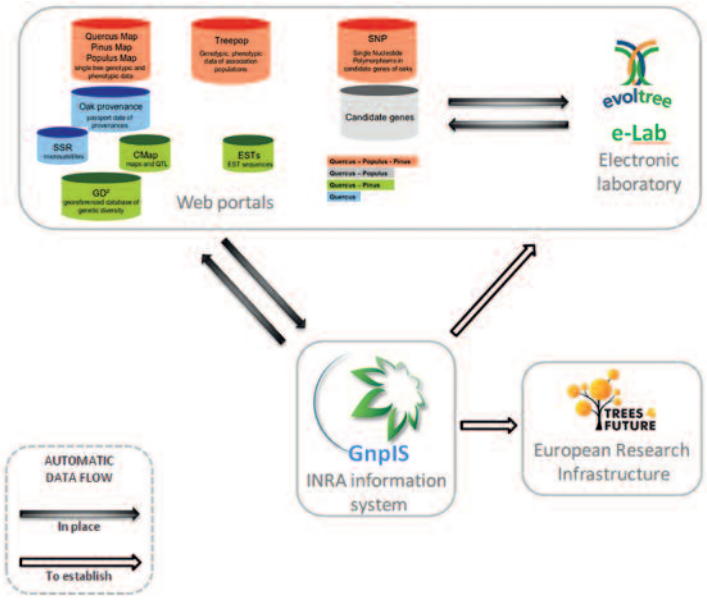


TABLE 2

Update of the content of Quercus Portal (March 31st 2016)

Databases	Taxons	Data types	Features
QuercusMap	*Q. robur, Q. petraea, Q. robur x Q. petraea*	Pedigrees	18
		Genotypes	11,000
		Traits	214
		Genotypic data	515,000
		Phenotypic data	335,000
Cmap	*Q. robur, Q. petraea, Q. robur x Q. petraea, C. sativa*	Geneticmap sets	24
		QTL map sets	13
		Maps	683
EST	*Q. robur, Q. petraea, Q. robur x Q. petraea*	Unigene sets	3
		Contigs OCV1	69,514
		Contigs OCV2	65,712
		Contigs OCV3	91,000
TreePop	*Q. robur, Q. petraea, Q. robur x Q. petraea*	ISS	4
		Association populations	7
		Genotypes	4,729
		Genotypic data	323,784
		Phenotypic data	83,813
GD²	*106 distinct species for Quercus genus*	Populations	4,017
		Trees	24,160
		Frequencymeasures	61,823
		Diversitymeasures	6,902
Oak provenance	*Q. robur, Q. petraea*	Provenances	419
		Provenance tests	60
		Seed lots	464
		Traits	1,874
		Phenotypic data	1,883,677
SSR	*Q. robur, Q. petraea*	Genetic markers	669
Candidate genes	*Q. robur, Q. petraea, Q. robur x Q. petraea*	Genes	648
		Traits	17
SNP	*Q. robur, Q. petraea*	SNP	7,576

trees can be hosted by international databases, such as GenBank, or dbSNP, and thus it is highly likely that in the future new databases will need to be constructed within the eLab.

In recent years, the eLab has also been connected to external data repositories related to either forestry or genomics. This is made possible by the use of a common set of exchange formats and of compatible protocols with the external repositories, similar to the TAPIR interface. Such interoperable protocols have now been installed with GnpIS (a multispecies integrative information system dedicated to genomic data of plants and fungi pests hosted and curated by INRA) and Trees4Future (an Integrative European Research Infrastructure in the field of Forestry) (Figure 3). Meetings, software demos and video conferences are organised to maintain the communication between collaborators and ensure a useful evolution of the information systems.

ACKNOWLEDGEMENTS

We are grateful to Catherine Bastien, Zamira Betancourt, Noémie Emeriau, Audrey Jacques-Gustave, Véronique Jorge, Thierry Labbé, Mélina Millox, Christophe Plomion, Frédéric Raspail, Richard Séverin, and Jean-Paul Soularue for their contribution during the construction of the databases and the eLab.

Photo: EFIATLANTIC

EVOLTREE TRAINING ACTIVITIES

F. GUGERLI[1], S. HAYES[2]

1 WSL Swiss Federal Research Institute, 8903 Birmensdorf, Switzerland; 2 EFIATLANTIC, 69, route d'Arcachon, 33612 Cestas, France.

One of EVOLTREE's four main Integration Activities (IAs) is dedicated to the training of students and young scientists.
For an integrative research consortium such as the Network of Excellence EVOLTREE, it is a prime responsibility to disseminate scientific knowledge and advances to the research community (e.g., through peer-reviewed articles) as well as to the public, but even more so directly to its constituting members; in particular to its students as the forthcoming generation of researchers.
One way in which EVOLTREE carries out such knowledge transfer is via the training opportunities created by its partners as part of their in-kind contribution to the network.

Transfer of knowledge within Evoltree – and beyond

During the initial, EC-supported phase (2006-2010), every partner was invited to suggest training courses that would fit the scope of Evoltree research in a broad sense. The organisation of a course was financially supported by the network's own funds so that expenses, e.g., for invited teachers, could be covered.

In EVOLTREE's second phase as a European Research Group (2011 onwards), it was agreed that partners could offer training courses as in-kind contribution, an alternative to direct monetary support of the network activities. As a consequence of this formal change, the types of training opportunities broadened so that excursions, workshops, existing courses from within the curriculum of a university, etc. could be integrated into the programme.

As a great benefit of the resources available within the network during both phases, all participants from the EVOLTREE partner institutions can be reimbursed for their expenses up to a pre-defined limit., these training courses are not only open to students from registered EVOLTREE partners, but also to interested students from outside the network (at their own costs). This way, it is possible to foster cross-disciplinary education and to establish or strengthen contacts between complementary fields of research.

Wide-ranging expertise for a variety of students

A multitude of disciplines are represented within the EVOLTREE partnership. Hence, training could greatly benefit from this broad range of expertise of leading scientists in their field within the consortium. To complement this competence in the training programme, EVOLTREE partners invited lecturers from a variety of

Photo: Julien Dumercq, LabEx COTE

AN INSIDER'S VIEW

Participating in an EVOLTREE training course has benefited many students over the past ten years. Students have appreciated the opportunities given by the network, be it the many topics explored within the various courses, or the depth and competence of the teaching in a particular course. This positive attitude is not only reflected by the often high numbers of participants, but also by respective feedback.

In their course feedback, participants have stated: "This course was beneficial for me and fulfilled my expectations. It was a good experience to learn interdisciplinary in approaching ecological problems. The necessity to combine natural, social and civil sciences to better understand the biodiversity loss and conservation was largely developed during the course." and "(...), a perfect place to forget about daily and mundane preoccupations and dedicate one's mind to the acquisition of new scientific skills". Such responses are great motivation for continuing our commitment to teaching our students and to dedicate time and resources towards these activities.

Photo: Patricia Gonzalez Diaz

disciplines and institutions for these training courses, which also contributed to the exchange of knowledge among the researchers involved.

An appealing outcome of the training programme was to see that not only young students of forest ecology, e.g., at PhD or post-graduate levels, took advantage of the opportunities to learn about new techniques, types of analyses, or concepts, but that established scientists also participated in the training events and could thus learn from – and at the same time actively contribute with their own background to – the training offered by their colleagues.

Over the years, the EVOLTREE training programme has accumulated an immense breadth of topics covered in the various courses (Table. 1). Students have been able to, for examples, learn about fundamental analytical tools in population genetics, take first steps towards effectively using the bioinformatic toolbox, debate about conceptual issues of the coalescent theory, obtain insights into and perform meta-analyses, attempt to detect genomic signatures of adaptation, or discuss ecological consequences of global change on forest ecosystems.

Outlook

The EVOLTREE community will continue to offer training opportunities that cover the entire breadth of EVOLTREE research and competence – and beyond. Benefiting from established courses or taking the opportunity for developing new teaching components, both researchers and their students of EVOLTREE partner organisations will be able to take part in the transfer of expertise and knowledge to the forthcoming generation of scientists in the fields of genetics, genomics, and ecology of forest ecosystems. These opportunities will also foster the integration of the European research laboratories taking part. Such personal contacts are fundamental and constitute a pre-requisite for continued integrative and interdisciplinary research.

TABLE 1

Overview of training opportunities during the Evoltree phase II as a European Research Group (2011–2015)

Year	Title/subject	Organising partner[1]
2011	Next-generation sequencing	U Udine
	Adaptation of forest management to climatic change	U West Hungary
	Functions of microbial communities in soils	Hemholtz; *TU Munich*
	Population genomics	U Oulu
	Evolutionary quantitative genetics in forest ecosystems	INRA Pierroton
2012	Genetic data analysis	CZ U Life Sciences; *N Carolina State U*
	Genome-wide association studies using mixed models	U Uppsala
	An interdisciplinary perspective on biodiversity and ecosystem services	*ALTER-Net*
	Ecophysiology techniques workshop	U Southampton
	Population genetic and genomic approaches	U Göttingen
2013	Estimating mating system and gene flow in plants	U Bygdosz
	NGS analysis for beginners	INIA; *U Valladolid*
	Global Ecology for Global Change	*LabEx COTE* (INRA Pierroton)
2014	Transfers and interactions between ecosystems	*LabEx COTE* (INRA Pierroton)
	NGS data analysis: from heaven to hell	U Udine
	Georeferenced genetic data and their evaluation	TU Zvolen
	Population structure and the architecture of quantitative traits	U Uppsala
2015	*Forest genetic monitoring*	*U Thessaloniki*
	Approximate Bayesian Computation	U Uppsala
	Ancestral graphs and SMC	U Uppsala
	Coalescent today	U Uppsala
	Ecology and society: biodiversity and global change	INRA Pierroton
	Global change and the evolutionary potential of forest trees	U Copenhagen
	NGS data for phylogenetics	U Marburg

1 Partner names in brackets indicate organisational link, but not full responsibility for course organisation; institutes in italics indicate non-Evoltree organisers.

FIGURE 1

Training courses combine expert lectures, hands-on computer work, guided disucssions, poster sessions, and excursions — in a creative and stimulating environment

Photos: RensingLab, C. Rosique

TREETYPE: COLLECTING TREE PHENOTYPES IN THE WILD

S. CAVERS[1], S.C. GONZÁLEZ-MARTÍNEZ[2].

1 Centre for Ecology and Hydrology, CEH Edinburgh, Bush Estate, Penicuik, Midlothian EH26 0QB, UK; 2 BIOGECO, INRA, University of Bordeaux, 33610 Cestas, France.

TreeType is a citizen science initiative of the EVOLTREE network to collect standardised phenotypic data in forest trees, from individual trees to whole forests. Data collection is centralised through a web application.

Additional information and detailed recording protocols are available on the TreeType website, www.treetype.org

Background and objectives

In an era when information from the genome is accessible as never before, the collection of standardised phenotypic data has become a limiting factor (Neale & Kremer, 2011). These data are essential if we are to establish the relationships between genes, phenotypes and natural selection in response to environmental drivers. For a wide variety of reasons, including breeding, climate change mitigation and conservation, we need to know how the genetic diversity observed in wild populations is translated, via the phenotype, into selectively relevant variation and how this reacts to the environment (in the broadest sense).

From individual genomes to populations and to species, most trees maintain very high levels of genetic diversity and, as a life history characteristic, have developed efficient mechanisms for dispersing that variation across spatially and temporally variable landscapes. Phenotypic data collected from experimental trial settings, whilst essential for establishing the genetic basis for traits, represents observation of unselected progeny and is always compromised by a lack of exposure to the home site environment, where genotypes may perform quite differently. To harness the power of new genomic techniques to understand selection and adaptation in tree species, we must now take observation of plant phenotypes into the wild.

Analysing quantitative phenotypic data from wild populations is challenging, but the application of appropriate methodological approaches can help to disentangle plasticity and local adaptation, in particular if phenotypic data from the same trees are collected at different time points or if molecular marker data is collected from the same trees (e.g., Castellanos *et al.* 2015). For example, Phillimore *et al.* (2012) showed that the slope of phenology on temperature through time will be due to mean plasticity plus any association between this trait values and temperature, that is, adaptive microevolution, and that these two components can be separated using spaciotemporal data collected in a citizen study similar to TreeType.

Another interesting use of phenotypic data collected in wild populations is the study of trade-offs among traits, which could either favour or prevent adaptation to the new

Photo: Fotolia

Photo: Fotolia

STRUCTURE AND FUNCTION

TreeType is structured as an openly-accessible, web-based data entry portal available at www.treetype.org, via which datasets for individual trees or in bulk via spreadsheets can be entered. A core set of species has been selected, to get the platform under way, but new users have the option to suggest and vote for new species to be included in the portal and these will then be added once a sufficient critical mass of interest is demonstrated. For each species, a 'leader' will be identified who will promote and guide the collection of data including optimising the protocols for that species. Across all species, however, the project will aim to collect data on the same general set of traits, as well as locally-specific environmental data.

INGREDIENTS FOR SUCCESS

The TreeType project was conceived using the 'citizen science' model, connecting a wide range of actors, from amateur naturalists to researchers, across the geographic ranges of target tree species and providing the infrastructure necessary for concerted data collection. Such approaches are now widespread and have been successfully deployed to gather data in many scientific fields, notably on phenology (e.g., Nature's Calendar, Track-a-tree).

When it works, citizen science has the great advantages of enabling data collection on a wide geographic scale at low cost. However, a number of critical factors must be taken into account to achieve success.

- Firstly, data quality depends on a careful selection of traits; they must be good indicators of the adaptive process in the wild, but at the same time easy to measure for the non-specialist.
- Secondly, a robust, simple platform for data collection is needed to promote participation.
- Finally, determined coordination, ensuring continual steering and promotion of the project is essential to see data collection through from project initiation to completion.

If these elements can be got right, then there is great potential to generate datasets of intrinsic value, as well as a resource on which to build future projects.

environments expected under impending climate change. Finally, the phenology and reproduction data collected by TreeType will increase our understanding of the interplay of demography and genetics for adaptation of forest trees to local environments.

Project construction

Following a period of consultation with experts within the EVOLTREE network, it was decided that measurements should be taken on a set of four basic traits as a minimum (diameter, number of fruits, average seed mass, and age) and where possible this should be extended with the addition of up to 8 additional characters.

The essence of the selection was to try to cover a range of general categories, of importance to adaptation in different ways, namely growth, reproduction, phenology and defence. In each, a simple trait was identified that could both be recorded reliably by specialists and non-specialists alike. Protocols for each identified trait were designed based on internationally agreed standards such as Perez-Harguindeguy *et al.* (2013). Recorders are asked to target a set of at least five trees growing from open-pollinated naturally dispersed seed, although it is made clear that any record of any tree is welcome. A basic set of environmental data for each tree is requested along with a photograph, when possible.

What will the future bring?

The coordinators of TreeType are following an open-access principle. Therefore, any data collected through the TreeType project will be made freely available, subject only to users contacting the TreeType managers to confirm their interest and intentions. Any use of the data for publications or reports will make a clear reference to the project and those TreeType contributors who have been directly involved in the collection of relevant data will have

the opportunity to be involved in collaborative work. Contributors are not automatically expected to be co-authors on any publications arising from the data, but they should be made aware of the data use, be offered an opportunity for involvement, and receive appropriate acknowledgement of any significant intellectual input. In addition to scientific publication, the data collected by TreeType will be regularly released to more general databases (e.g., the TRY database) and reported in data papers.

A phone application will be developed to facilitate data collections by citizens.

A concerted effort will be made to use the TreeType infrastructure to collect data for a set of key species, for which contributors will be actively sought. In parallel, the portal is essentially open to any enthusiastic researcher who feels it can fit their needs for data collection in their chosen species. We hope that initiating this project will enable and catalyse the establishment of some fundamental datasets which can be analysed in their own right, and which can form the basis of new projects and studies in the future, in particular those involving genome-wide genotyping.

ACKNOWLEDGEMENTS

the TreeType project was developed in collaboration and consultation with many colleagues across the EVOLTREE network. For their essential contributions, we thank Stephan Gaubitzer, Egbert Beuker, Celine Blanc-Jolivet, Oliver Brendel, José Climent, Thomas Källman, Christian Pichot, Tommi Suominen, Jill Thompson and the EVOLTREE Executive Committee.

REFERENCES

- Neale DB, Kremer A. 2011. Forest tree genomics: growing resources and applications. *Nature Reviews Genetics* 12: 111-122.
- Castellanos MC, González-Martínez SC, Pausas JG. 2015. Field heritability of a plant adaptation to fire in heterogeneous landscapes. *Molecular Ecology* 24: 5633-5642.
- Pérez-Harguindeguy N, Díaz S, Garnier E, *et al.* (32 more authors). 2013. New Handbook for standardized measurement of plant functional traits worldwide. *Australian Journal of Botany* 61: 167-234.
- Phillimore AB, Stålhandske S, Smithers RJ, Bernard R. 2012. Dissecting the contributions of plasticity and local adaptation to the phenology of a butterfly and its host plants. *American Naturalist* 180: 655-670.

RECORDED PHENOTYPIC TRAITS IN TREETYPE

1. **DBH (cm)**: Trunk diameter at 1.30 m, i.e., approximately at an adult's breast height.
2. **Height (m)**: Height from ground to tallest part of the crown.
3. **Crown size (m)**: Crown size is a measure of the footprint or plan area of the crown of a tree expressed as crown width.
4. **Stem form (class)**: Stem form is measured using a categorical classification based on straightness and verticality.
5. **Bark thickness (mm)**: Bark thickness is the thickness of the part of the stem that is external to the wood or xylem, including the vascular cambium.
6. **Number of fruits (units)**: Fruits counted from the ground using binoculars.
7. **Seed mass (g)**: Seed mass is the oven-dried mass of an average seed.
8. **First flowering date**: The date when the first male and/or female flower is observed.
9. **Bud flush date**: Bud flush or bud break is the first evidence of active growth resumption in the spring.
10. **SLA, Specific leaf area (cm^2/g)**: Specific leaf area (SLA) is the one-sided area of a fresh needle, divided by its oven-dry mass.
11. **Foliar damage (%)**: Leave damage regardless of the cause of the damage.
12. **Tree age (years)**: The age of the tree. Age determination is very important to compare trees growing in different environments.

Photo: S.C.Martinínez-González

Research highlights: science that matters

POTENTIAL FOR EVOLUTIONARY RESPONSES TO CLIMATE CHANGE - EVIDENCE FROM TREE POPULATIONS

Florian J. Alberto[1,2], Sally N. Aitken[3], Ricardo Alía[4], Santiago C. González-Martínez[4], Heikki Hänninen[5], Antoine Kremer[2], François Lefèvre[6], Thomas Lenormand[7], Sam Yeaman[3,8], Ross Whetten[9], Outi Savolainen[1]

[1] Department of Biology and Biocenter Oulu, University of Oulu, Oulu, FIN-90014 Finland
[2] INRA, UMR1202 Biodiversité Gènes et Communautés, Cestas, F-33610, France and Université de Bordeaux, UMR1202 Biodiversité Gènes et Communautés, Talence, F-33410, France
[3] Department of Forest and Conservation Sciences and Centre for Forest Conservation Genetics, University of British Columbia, Vancouver, BC V6T 1Z4, Canada
[4] Department of Forest Ecology and Genetics, INIA - Forest Research Centre, E-28040 Madrid, Spain
[5] Department of Biosciences, University of Helsinki, Helsinki, FIN-00014 Finland
[6] INRA, UR629 Ecologie des Forêts Méditerranéennes, URFM, Avignon, F-84914, France
[7] Centre d'Ecologie Fonctionnelle et Evolutive, CNRS, Université de Montpellier, UMR 5175, Montpellier F-34293, France
[8] Institute of Biology, Université de Neuchâtel, Neuchâtel, CH-2000, Switzerland
[9] Department of Forestry & Environmental Resources, NC State University Raleigh NC 27695-8008, USA.

Corresponding Author:
Outi Savolainen
Department of Biology and Biocenter Oulu
FIN-90014 University of Oulu
outi.savolainen@oulu.fi

Key words: quantitative genetics, adaptive traits, natural selection, local adaptation, phenotypic plasticity, provenance trials, conifers.

Tables S1, S2 and S3, along with the corresponding supporting information legends and references, can be downloaded here: http://onlinelibrary.wiley.com/wol1/doi/10.1111/gcb.12181/suppinfo

Published in *Global Change Biology* (2013) 19: 1645-1661, with doi:10.1111/gcb.12181

Evolutionary responses are required for tree populations to be able to track climate change. Results of two hundred years of common garden experiments show that most forest trees have evolved local adaptation, as evidenced by the adaptive differentiation of populations in quantitative traits, reflecting environmental conditions of population origins. Based on patterns of quantitative variation for 19 adaptation related traits studied in 59 tree species (mostly temperate and boreal species from the Northern hemisphere), we found that genetic differentiation between populations and clinal variation along environmental gradients were very common (respectively 90% and 78% of cases). Thus, responding to climate change will likely require that the quantitative traits of populations again match their environments.

We examine what kind of information is needed for evaluating the potential to respond, and what information is already available. We review the genetic models related to selection responses, and what is known currently about the genetic basis of the traits. We address special problems to be found at the range margins, and highlight the need for more modeling to understand specific issues at southern and northern margins. We need new common garden experiments, for less known species. For extensively studied species, new experiments are needed outside the current ranges. Improving genomic information will allow better prediction of responses. Competitive and other interactions within species and interactions between species deserve more consideration. Despite the long generation times, the strong background in quantitative genetics and growing genomic resources make forest trees useful species for climate change research. The greatest adaptive response is expected when populations are large, have high genetic variability, selection is strong, and there is ecological opportunity for establishment of improved genotypes.

Introduction

Populations can respond to environmental change through phenotypic plasticity, by moving to a new area corresponding to environmental conditions they are adapted to, by genetically adapting to the new conditions, or by combinations of these responses (Aitken *et al.* 2008). Most attention has been paid to range expansion or contraction (Chen *et al.* 2011, Parmesan, 2006), typically using models that assume the species are genetically homogenous. The potential for genetic responses has often been neglected, for instance in the IPCC reports (IPCC, 2001, IPCC, 2007), even if it is well known that evolutionary changes, i.e. genetic responses, have historically accompanied changes in climate (Davis & Shaw, 2001). Further, it is also now understood that the rate of adaptation required by climate change varies among geographic regions (Loarie *et al.* 2009). Modeling work on the potential of populations and species to

respond genetically to recent climate change is advancing (see Franks & Hoffmann, 2012, Hoffmann & Sgrò, 2011, Shaw & Etterson, 2012 for recent reviews). The immediate responses via phenotypic plasticity have also been considered in the context of climate change (Nicotra *et al.* 2010).

Here we examine the importance of and potential for genetic responses to climate change in forest tree populations. Trees are ecologically key species in many terrestrial ecosystems, including boreal and temperate forests in Europe and North America. Their response to climate change can substantively impact the global carbon cycle. Local adaptation (Kawecki & Ebert, 2004) is more common in trees than in some other plant species. Tree species are adapted, to the current climate, and they are thus potentially greatly influenced by the rapid changes in climate (Savolainen *et al.* 2007). The long generation times are a challenge for research, but trees also provide some advantages for these studies, as described below.

First, adaptation to climate change will depend on phenotypic traits relevant in the new environments, such as timing of growth and drought or cold tolerance. There is an extraordinary wealth of information on the quantitative genetics and population differentiation of trees for these traits, based on 250 years of forestry common garden experiments, known as provenance trials (Langlet, 1971, Morgenstern, 1996), and on extensive tree breeding experience.

Secondly, the demographic history since the last glacial maximum has been reconstructed for several tree species by combining phylogeographic and palynological approaches with coalescent-based studies of population demography (Cheddadi *et al.* 2006, Eckert *et al.* 2010, Heuertz *et al.* 2006, Magri *et al.* 2006, McLachlan *et al.* 2005, Parducci *et al.* 2012, Petit *et al.* 2002, Soltis *et al.* 2006). Rates of past adaptation of trees to climate changes can be inferred from these studies (Hendry & Kinnison, 1999). The increasing knowledge of the molecular basis of quantitative trait variation (see Neale & Kremer, 2011 for references) can improve predictive models (see e.g., Wilczek *et al.* 2010). This body of background information allows us to examine the potential for adaptation in natural conditions better than in many other organisms. For instance, in butterflies, studies of responses to climate change have relied nearly exclusively on examining molecular marker variation (Hill *et al.* 2011).

Trees have very long generation times, but they share population genetic characteristics with other outcrossing plants and animals with high levels of gene flow and large effective population sizes (Petit & Hampe, 2006). Trees are highly fecund, and may rapidly increase their population sizes. Because they are sessile, they generally have good tolerance of a range of environmental conditions and large plastic responses. There are ecologically and commercially important trees with large continuous distributions, such as *Picea abies*, *Pinus contorta*, and *Pinus sylvestris*, but also species with small, fragmented distributions more susceptible to genetic drift. The dispersal capacity of tree species will play a crucial role in their potential for adaptation. Hybridization between closely related tree species can also influence their adaptive capacity out of their current range, as it has been shown in other organisms (Hoffmann & Sgrò, 2011, Olson-Manning *et al.* 2012 and references therein).

The focus of this review is on predicting evolutionary responses, with as much evolutionary, genetic, and ecological realism as possible. We examine the models needed for prediction, starting with the simplest models of evolution in individual populations, and continuing to more complex and more realistic models involving multiple populations in heterogeneous environments. We discuss what data are needed for realistic prediction of genetic responses, what information is already available, and what additional information we need in terms of new models, new data, or new analyses of existing data (Lindner *et al.* 2010). Quantitative genetic models of evolutionary response deal with traits that will confer adaptation to future environments. While it is not easy to predict what traits will be most important in the future, it is reasonable to examine traits related to climate, such as the timing of growth and reproduction (Hänninen & Tanino, 2011, Rohde & Bhalerao, 2007) or cold and drought tolerance (Niinemets, 2010).

Evolution in one isolated population

A single population: the breeder's equation

According to the breeder's equation, the simplest model governing response to directional selection on a single trait, the response in a large population with no gene flow depends on the strength of selection, on the amount of genetic variation, and its ratio to total phenotypic variation (heritability; (see Falconer & Mackay, 1996). If there is no genetic variation, any change in phenotype in a novel environment inducing directional selection would be due to phenotypic plasticity alone. Forest tree populations harbor considerable genetic variability in many quantitative traits (Cornelius, 1994, Howe *et al.* 2003, Morgenstern, 1996) as well as at the DNA level (Savolainen & Pyhäjärvi, 2007). While tree breeders can control the intensity of selection and predict responses in breeding populations, it is much more difficult to make such predictions in the wild. Environmental variances will be higher, and heritabilities generally lower (Conner *et al.* 2003). Methods for estimating heritabilities in the wild are improving because of much better estimates of relatedness and improved methods (Ritland, 1996, Sillanpää, 2011), and will be of critical importance to understanding responses to climate change.

Assessing the strength of directional selection is a demanding task, as we do not even know exactly which traits are most important for fitness, and the longevity of

trees makes lifetime fitness estimates unattainable in a realistic timeframe. Estimates of directional selection are available for natural populations (Kingsolver & Diamond, 2011, Kingsolver *et al.* 2001), but studies on forest trees are lacking. Further, selection is likely to be variable across environments, years, and life stages. In natural populations the traits are also subject not just to directional selection but also to stabilizing and disruptive selection, not included in this simplest model. Thus, for most natural situations, the breeder's equation is far from the reality of populations responding to natural selection.

Temporal variation in selection

Two general classes of quantitative genetic models have been developed to study the risk of extinction in single populations: models with a sudden single step-change in the optimum phenotype (Gomulkiewicz & Holt, 1995, Gomulkiewicz *et al.* 2010, Gomulkiewicz & Houle, 2009, Pease *et al.* 1989), and models with a continuous change in the optimum phenotype (Björklund *et al.* 2009, Burger & Lynch, 1995, Chevin *et al.* 2010, Lynch & Lande, 1993). In single step-change models, extinction occurs as a consequence of decreasing population size due to selective deaths as the population adapts to the change in environment. In the continuous-change models, by contrast, extinction is assumed to occur when the pace of adaptation lags behind the rate of change in the optimum phenotype (see Aitken *et al.* 2008 for further discussion). There are several interesting ways in which these models could be extended to increase biological realism. Most of these models assume that the strength of selection does not vary with population density, which is unrealistic for most forest trees, as competition is likely greatly reduced at low densities (see Björklund *et al.* (2009) for a simulation model incorporating density dependent selection). Also, failing to account for changes in biotic interactions that may be associated with climatic change could cause models to under- or over-estimate extinction risks (Gilman *et al.* 2010). Climate change may result in the introduction of new pests, as for instance the mountain pine beetle (Robertson *et al.* 2009) or new pathogens (Netherer & Schopf, 2010), but also losses of current competitors, insects, or diseases caused for example by phenological shifts between trees and associated pests (van Asch & Visser, 2007).

While it is possible to parameterize some of these models to make quantitative predictions about extinction risk, the assumptions involved greatly limit the faith that should be placed in any such predictions (see Aitken *et al.* 2008 for further discussion). Rather, they seem most useful as heuristic tools to identify the most likely factors causing population extinction and to compare relative risk among species. Generally, these models find that the probability of extinction decreases for species with large population sizes, high fecundity, high heritabilities, and high amounts of standing genetic variation. While many forest trees present such characteristics, extra effort should be made to study species that are on the low end of the spectrum for any of these characteristics. Some examples of species that may be particularly vulnerable due to their small population sizes are *Pinus torreyana* in North America, or *Abies pinsapo* in Europe. More study is necessary to see whether such vulnerable species also have lower levels of standing variation.

Genetic basis of adaptive trait variation

The expected genetic responses in many models depend on the genetic architecture of the trait (e.g. Gomulkiewicz *et al.* 2010). While the traditional polygenic model of Fisher (Fisher, 1918, Fisher, 1930) is based on small effects at a very large number of loci, some models of selection predict larger effect sizes (Orr, 1998, Yeaman & Whitlock, 2011). Overall, quantitative trait locus (QTL) studies in forest trees have generally found large numbers of loci with relatively small effect sizes, compared to some crop plants (Barton & Keightley, 2002, Howe *et al.* 2003, Laurie *et al.* 2004). Association studies have further confirmed this view of moderate effect sizes (summarized in Fig. 2), e.g. in *Pinus taeda* (Cumbie *et al.* 2011, Quesada *et al.* 2010), in *Populus tremula* (Ingvarsson *et al.* 2008), *Picea sitchensis* (Holliday *et al.* 2010a), and *Pseudotsuga menziesii* (Eckert *et al.* 2009). These findings are consistent with the small effect sizes of flowering time and leaf trait variation loci in maize (Buckler *et al.* 2009, Tian *et al.* 2011), and human height (Hill *et al.* 2008). In contrast, Atwell *et al.* (2010) found large effect SNPs for many phenotypic traits of *Arabidopsis*. There may also be major effect loci for disease resistance, such as for rust disease caused by fungal pathogens in North American conifers (Kayihan *et al.* 2010). The associated loci may well differ between environments due to genotype by environment interactions (Jermstad *et al.* 2003) or because of different genetic basis in different areas (Goldstein & Holsinger, 1992, Hancock *et al.* 2011). In many conditions, the phenotypic differences between populations can be due to combined effects of several loci rather than differentiation at the level of individual loci (Kremer & Le Corre, 2012, Latta, 1998, LeCorre & Kremer, 2003).

Weak genetic correlations allow traits to respond to selection independently, whereas genetic correlations opposing the direction of selection will delay the response (Etterson & Shaw, 2001), and reinforcing correlations will accelerate it. Under stabilizing selection, responses are facilitated, if the selection is weak (Duputie *et al.* 2012). The underlying causes of genetic correlations are so far not known in trees.

Overall, the limited findings so far suggest that the response to strong selection on phenotypes will often be based on many loci with small effects, and fairly weak selection on individual loci, as has been also found in humans (Turchin *et al.* 2012). If larger effect loci are involved, response predictions could then use specific information on such loci. Alternatively, genomic selection methods could be used to build predictive models that do not need to identify the particular loci underlying adaptive

genetic responses (Grattapaglia & Resende, 2011, Holliday *et al.* 2012, Iwata *et al.* 2011, Resende *et al.* 2012).

We do not know whether most adaptations in trees are due to existing variation or new mutations. During interglacial periods, tree populations have repeatedly colonized northern areas and have rapidly adapted to those conditions, likely because the north-adapted variants may have remained in southern populations at lower frequencies (De Carvalho *et al.* 2010, Savolainen *et al.* 2011). Typically large effective population sizes in forest trees would have contributed to rapid fixation of adaptive variants. This supports an interpretation of evolution from standing rather than *de novo* variation.

Phenotypic plasticity and adaptation

Trees exhibit a high degree of phenotypic plasticity with respect to climatic variation. Phenological shifts of bud flush in response to recent increases in temperatures have been widely documented (Menzel & Fabian, 1999, Menzel *et al.* 2006, Parmesan, 2006). Arid years or an arid microsite may favor the development of deeper and denser root systems (Kozlowski & Pallardy, 2002). In such a context, adaptive plasticity can buffer the impact of changing conditions on population size (Lynch & Lande, 1993). However, these plastic changes may take time to develop (as in the root example above). In addition, more plasticity also means less intense selection, causing populations to genetically track changing optima more slowly. Recent models have shown that the decreased selection is more than compensated for by the increased phenotypic match allowed by plasticity (Chevin & Lande, 2010). In fact, the evolution of plasticity can provide populations with a transient and efficient response to large environmental changes (Lande, 2009).

Multiple-site provenance trials can be used to examine the plastic responses of populations in new environments. This can be quantified with response functions for individual populations, which describe the change in a trait as a function of transfer distance or change in environmental factors (Rehfeldt *et al.* 2002, Rehfeldt *et al.* 1999). Provenance trials have been planted in sites that vary with respect to many environmental variables, such as temperature or water availability (Kramer, 1995, Morgenstern, 1978, Rehfeldt *et al.* 2002, Rehfeldt *et al.* 1999, Reich & Oleksyn, 2008, Shutyaev & Giertych, 1997, Vitasse *et al.* 2010, Worrell *et al.* 2000). Transfers to the south have been used to predict responses to a warming climate (Beuker *et al.* 1998, Rehfeldt *et al.* 2002, Wang *et al.* 2006) even if the future conditions may be different (e.g. photoperiod). Further, these experiments take place in spaced plantings of seedlings, and thus ignore germination, establishment, and early intra- and inter-specific competitive effects. Response functions of individual populations have been developed for growth using very large datasets of multiple trials including more than a hundred populations available for *Pinus contorta* (Rehfeldt *et al.* 2001), *Pinus sylvestris* (Rehfeldt *et al.* 2002) and *Larix occidentalis* (Rehfeldt & Jaquish, 2010). Recently, Wang *et al.* (2010) developed a universal response function for Pinus contorta, which integrated populations and environment effects and can be used to predict the performance of any population in any climatic conditions. Incorporating provenance trial data on local adaptation and phenotypic plasticity in models predicting future distributions reduced dramatically the extinction risk in southern populations (Benito-Garzón *et al.* 2011, Morin & Thuiller, 2009). The plastic response of different traits (*e.g.* phenology in trees) to variation in climate is, however, often much more complex than in heuristic models of adaptation (see e.g. Caffarra *et al.* 2011, Hänninen & Tanino, 2011, Valladares *et al.* 2007).

Finally, epigenetic effects on phenotypic plasticity and inheritance of phenotypic variation need further investigation. Epigenetic variation can be partly inherited from one generation to the next while being still sensitive to environmental variation (Richards *et al.* 2010). Maternal epigenetic effects are known in *Arabidopsis* (Johannes *et al.* 2009), but so far their nature has not been studied much in trees (Bräutigam *et al.* 2013). Epigenetic effects can also occur during seed maturation. Temperature differences during embryogenesis caused differences in phenology in *Picea abies* (Skrøppa & Kohmann, 1997) and the molecular mechanisms involved are being studied (Yakovlev *et al.* 2010). They could have significant implications for the interpretation of provenance trial data, explaining some of the phenotypic variation among populations that is commonly interpreted as genetic variation.

Evolution in multiple populations

Geographic distribution and genetic structure

Natural populations of a species in a heterogeneous landscape may have very different patterns of distribution, which can influence its population genetic characteristics (Fig. 3) as reviewed by Charlesworth & Charlesworth (2010). The classical island model assumes populations of equal finite constant size, with equal migration rates between them (Wright, 1931). These assumptions can be relaxed, with variable migration rates and changing population sizes. Species can also be distributed in large continuous populations where parts of the range are connected by symmetric gene flow, as described in the isolation by distance model by Wright (1943). Populations located at range margins represent a special case, as they are at the edge of environmental gradients where carrying capacity may be limited. In such cases, there is more migration from the core populations to the margin than vice versa, resulting in asymmetric gene flow (Kirkpatrick & Barton, 1997).

Many economically important temperate and boreal species have large populations covering vast areas, but other tree species do not fit this distribution model.

TABLE 1

Distribution range and genetic estimates for the 27 European conifers

Species	Range	Distribution	Mean Q_{ST}[a]	Q_{ST} range[a]	F_{ST}	H_e	Reference
Abies nebrodensis	Sicilia	south small				*0.201*	Ducci *et al.* (1999)
Abies pinsapo	Andalusia	south small				*0.056*	Scaltsoyiannes *et al.* (1999)
Pinus nigra ssp dalmatica	South Croatia	south small			*0.091*	*0.292*	Nikolic & Tucic (1983)
Picea omorika	Croatia Serbia	south small			*0.261*	*0.067*	Ballian *et al.* (2006)
Pinus nigra ssp laricio	Corsica Calabria Sicilia	south small			*0.005*	*0.182*	Scaltsoyiannes *et al.* (1994)
Abies cephalonica	Balkans	south small	0.140	0.100 - 0.170	*0.048*	*0.221*	Fady & Conkle (1993)
Pinus peuce	Balkans	south small			*0.083*	*0.124*	Zhelev & Tsarska (2009)
Pinus brutia	Aegean Sea	south fragmented	0.040		*0.053*	*0.196*	Kara *et al.* (1997)
Pinus heldreichii	Balkans	south fragmented			*0.054*	*0.177*	Boscherini *et al.* (1994)
Abies borisii-regis	Balkans	south fragmented				*0.273*	Scaltsoyiannes *et al.* (1999)
Pinus nigra ssp pallasiana	Greece Serbia Bulgaria	south fragmented	0.028	0.020 - 0.040	*0.070*	*0.114*	Tolun *et al.* (1999)
Pinus nigra ssp salzmannii	East Spain South France	south fragmented				*0.216*	Scaltsoyiannes *et al.* (2009)
Pinus nigra ssp nigra	North Italy Croatia Greece	south large fragmented				*0.264*	Scaltsoyiannes *et al.* (2009)
Pinus pinaster	South West Europe	south large fragmented	0.616	0.441 - 0.791	*0.076*	*0.142*	Salvador *et al.* (2000)
Pinus pinea	South Europe	south large fragmented			*0.279*	*0.011*	Fallour *et al.* (1997)
Pinus halepensis	South Europe	south large fragmented	0.130			*0.040*	Schiller *et al.* (1986)
16 species with small or fragmented range			***0.192***		***0.082***[c]	*0.171*[c]	
Pinus cembra	Alps Romania	north large continuous	0.830		*0.040*	*0.081*	Belokon *et al.* (2005)
Pinus uncinata	Central West Europe	north large continuous			*0.006*	*0.260*	Lewandoski *et al.* (2000)
Larix decidua	Central Europe	north large continuous			*0.051*	*0.223*	Maier (1992)
Pinus sibirica	East Siberia	north large continuous			*0.027*	*0.278*	Goncharenko *et al.* (1992)
Pinus mugo	Central East Europe	north large continuous			*0.041*	*0.214*	Slavov and Zhelev (2004)
Abies alba	Central Europe	north large continuous	0.075	0.000 - 0.150		*0.252*	Ducci *et al.* (1999)
Abies sibirica	Siberia	north very large continuous			*0.102*	*0.083*	Semerikova & Semerikov (2006)
Larix sibirica	Siberia	north very large continuous			*0.082*	*0.159*	Semerikov *et al.* (1999)
Picea abies ssp obovata	Lapland Siberia	north very large continuous			*0.011*	*0.213*	Krutovskii & Bergmann (1995)
Picea abies ssp abies	North Central Europe	north very large continuous	0.417	0.106 - 0.727	*0.044*	*0.252*	Krutovskii & Bergmann (1995)
Pinus sylvestris	Whole Europe	north very large continuous	0.519	0.080 - 0.860	*0.033*	*0.286*	Goncharenko *et al.* (1994)
11 species with continuous range		***0.463***		***0.044***	*0.209*		

[a] Mean Q_{ST} and Q_{ST} range were calculated from estimates only for height increment, bud flush and bud set (for more details and references see Table S1). Q_{ST} estimates corresponds to the levels of population differentiation measured either as the proportion of phenotypic variation between populations (V_{pop}) or as the proportion of additive genetic variance between populations (Q_{ST}) in the provenance trials (for more details see Table S1).

[b] References of the studies using allozyme markers to assess F_{ST} and H_e. See supporting information references for full reference information.

[c] Pinus pinea, which has hardly any within-population variation (Vendramin *et al.* 2008), was not included in the calculation of mean F_{ST} and mean H_e.

We examined the population structure of European conifers in the *Pinaceae* (including pines, spruces and firs), a limited group of species with very good distributional and reasonable population genetics information. A compilation of the distributions of these 27 species (and sometimes subspecies; from (Jalas & Suominen, 1973), allowed us to classify them as having northern or central large, southern large or southern small or fragmented distributions (Table 1). Note that the classification is based on the current distributions, although some currently fragmented species may have had much larger distributions in the past (Soto *et al.* 2010). Species with a predominantly northern distribution, but also occurring in the south (e.g. *Pinus sylvestris*) were classified as northern species. Figure 3 shows examples of distributions of three species (*Picea omorika, Pinus pinaster and Pinus sylvestris*). There are 11 species with predominantly northern or central, large, continuous distributions, and four southern species with large but somewhat fragmented distributions. About half of the European conifers (12) have southern, small or fragmented distributions. Further, the southern margin of most species seems to consist of fragmented small populations, whereas in the north, the range margin is part of a continuous distribution for several species. This analysis shows that in many tree populations, the threats associated with climate change are accompanied by and likely exacerbated by the effects of fragmentation at southern range margins (see also Lynch, 1996). However, if there is still extensive gene flow among the fragments, the population structure should resemble that of a continuous population.

Consistent with the theoretical predictions, the European conifers with continuous distributions have higher genetic diversity (H_e) than the fragmented ones (Table 1). The widespread northern species such as Picea abies and Pinus sylvestris have low levels of genetic differentiation (F_{ST}) in their main range (Heuertz *et al.* 2006, Pyhäjärvi *et al.* 2007). Similar findings have been made in North America for species such as *Pseudotsuga menziesii* (Eckert *et al.* 2010), *Picea sitchensis*, *P. glauca* and *P. mariana* (Chen *et al.* 2009, Holliday *et al.* 2010a, Holliday *et al.* 2010b, Namroud *et al.* 2008). In contrast, the level of population differentiation is almost twice for the southern fragmented species compared to the northern widely-distributed ones (Table 1). Thus, the genetic data available are broadly consistent with the population structure classification based on species distribution and census size. However, current census size may ignore effects of past demographic history such as population size changes or hybridization, and we do not expect the current distributions to account for all variation in patterns of diversity.

Next we examine the patterns of quantitative genetic variation in tree species in general and in these European conifers in particular to evaluate the effects of selection for local adaptation. We reviewed the literature of provenance trials and found a total of 112 studies on 19 relevant traits related mostly to phenology, growth, cold or drought tolerance or other ecophysiological traits, among which were 36 studies on European conifers (Table S1). Among 59 tree species studied, most were native to Europe and North America (23 and 29 species respectively) while conifers were more studied than angiosperms (36 and 23 species respectively). Only three traits had been measured in a sufficiently large number of experiments (Table 2) to make general comparisons and draw general patterns. We focused on the patterns of genetic variation for height increment and for the timing of bud flush, at the beginning of the growing season in spring, and the timing of bud set, an indication of cessation of growth in fall. Over all studies, these three traits had comparable levels of genetic differentiation between populations (mean value equal to 0.249, 0.324 and 0.392 for bud flush, height increment and bud set, respectively; Table 2).

Quantitative variation in fragmented populations

In Europe, small and fragmented conifer populations occur mainly in the southern Mediterranean area. Provided population sizes are sufficiently large, species with greater differences among populations in local phenotypic optimum and higher levels of genetic variance would be expected to have higher equilibrium differentiation. Gene flow in contrast, would reduce differentiation (Hendry *et al.* 2001). Generally, if there is strong differential selection between populations, we would expect that the proportion of total genetic variance found between populations, Q_{ST}, should be higher than F_{ST} calculated from neutral markers with appropriate mutation rates (Edelaar *et al.* 2011, Leinonen *et al.* 2008).

In the limited set of provenance trials on European conifers, estimates of quantitative genetic differentiation among populations for species with small or fragmented range were low over all traits (mean Q_{ST} = 0.192, 5 species; Table 1). This average is about twice as high as the neutral F_{ST} (0.082; 9 species; Table 1). Even though sampling across an environmental gradient is clearly not concordant with the assumptions of the island model, data of this kind are frequently analyzed by comparing Q_{ST} and F_{ST} estimates for distinct samples from large and continuous populations. The average Q_{ST} estimate for large populations in northern areas is 0.463 while average F_{ST} is 0.044. Thus, in this small set of studies, the ratio of Q_{ST} to F_{ST} is much lower for species with small or fragmented range than that found in more widespread species. In small populations or fragments, selection for local adaptation is less efficient because of the effects of genetic drift on individual loci, and further, on the associations of alleles at different loci (Le Corre & Kremer, 2012). A review by Leimu & Fisher (2008) found that in plants only about 50% of all population pairs in reciprocal transplantations studies

showed evidence of local adaptation, *i.e.* each population at its native site had higher fitness than other populations introduced to that site. Local adaptation was much less likely in small than large populations. However, Q_{ST} values could also differ because the studies on species with limited distributions have sampled a smaller range of environmental variation than studies in species with large ranges, or because the scale of fragmentation does not match the scale of environmental variation. Reciprocal transplant experiments are needed to assess the level of local adaptation directly. In the large provenance trial data set over all 19 traits and 59 tree species, 264 of 294 analyses (around 90 %) showed significant differentiation across populations (Table S1), in most cases likely due to climatic selection.

There is also some evidence in the literature for local climatic adaptation in southern European fragmented populations, such as for water use efficiency in *Pinus halepensis* (Voltas *et al.* 2008). Further, some allelic variants at candidate loci for drought tolerance have also been found to be associated with environmental variables (Grivet *et al.* 2011). In some of these species, selection may have been strong enough for local adaptation to evolve. Clearly, more studies on the patterns of local adaptation are needed in the species with fragmented southern distributions. Forests at Mediterranean southern limits are threatened by faster changes in precipitations than in the northern range limit. If indeed their adaptive capacity is lower, this could make southern fragmented populations even more vulnerable.

It is also possible that these populations have evolved high adaptive phenotypic plasticity in response to environmental variability instead of genetic differentiation, either for some specific traits or across the genome (Nicotra *et al.* 2010). This could be likely if there is also a strong temporal component of environmental variation (Hedrick, 2006). In a changing climate, the responses due to phenotypic plasticity may maintain fitness despite climatic changes. More growth chamber or reciprocal transplant experiments will be needed to assess the response functions for these species.

Quantitative variation in continuous populations along environmental gradients

Species present in Central and Northern Europe generally have continuous distributions covering large areas encompassing much heterogeneity in abiotic and biotic environmental factors with large effective population sizes connected by extensive gene flow. If there is differential selection along environmental gradients, we expect to see patterns of clinal variation of traits (Barton, 1999). These patterns can be described by the slope of a regression along an environmental gradient. The proxies for environmental gradients most frequently used are latitude and altitude. For height increment, populations from warmer environments generally grew faster in the provenance trials (see Table S1) but quantitative estimates of the slopes were rarely available. Populations from cold

FIGURE 1

Mean silent nucleotide diversity per sites (π_{silent}) estimates for several tree species.
Average nucleotide diversity at silent sites (for more details and references see Table S2). Angiosperms appear in light grey and conifers in dark.

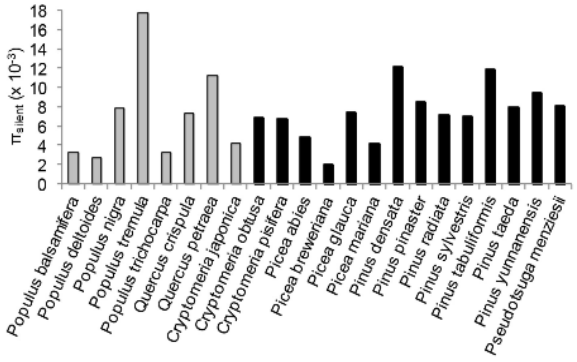
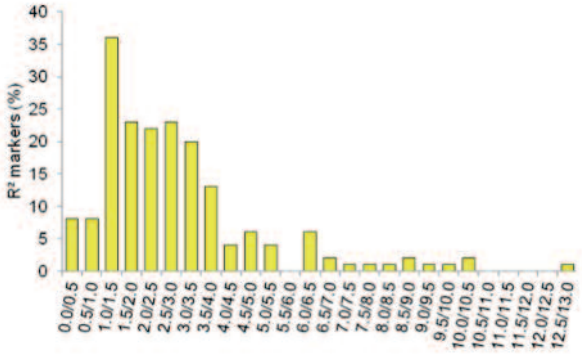


FIGURE 2

Distribution of allelic effect sizes in tree species

Distribution of the percentages of phenotypic variance explained by genotypic classes at SNP loci (R^2 marker) detected in significant associations with quantitative traits (for more details and references see Table S3).



environments cease growth earlier, as an adaptation to the approaching winter (see e.g., Savolainen *et al.* 2004). To compare slopes of clinal variation we focused on the two phenological traits, the timing of bud flush and the timing of bud set, and compared altitudinal and latitudinal clines. To summarize data across species and environments, we considered that one degree of latitude corresponds approximately to a temperature change of 0.6 °C, and correspondingly, 100 m of altitude (Jump *et al.* 2009). We show examples of an altitudinal cline in bud flush in Quercus petraea (Fig. 4a) and a latitudinal cline in bud set in *Pinus sylvestris* (Fig. 4b).

The results of the summary show that the two phenological traits differ in their patterns. For bud flush, both altitudinal and latitudinal clines showed similar slopes, but the direction of adaptation varied greatly among species (Table 3a). For example, high altitude populations

from the same transect flushed late in *Quercus petraea* (Fig. 4a), whereas in *Fagus sylvatica* they flushed early (Vitasse *et al.* 2009). This could reflect different compromises in the adaptive tradeoff between maximizing the growing season length and exposing new leaves to late frosts. Bud flush is triggered by the accumulation of cold (or chilling) sums followed by heat (or forcing) sums above a threshold temperature sum. These genetically determined critical temperature sums and thresholds may vary among species, and to a lesser extent among populations of the same species (Hänninen & Tanino, 2011). Bud flush in late successional species is also more influenced by photoperiod than in early successional species (Basler & Körner, 2012, Körner & Basler, 2010). Bud set showed steeper slopes for both gradients and in all species more northern or higher altitude populations had earlier bud set (Table 3b). These data suggest that differential selection on bud set is systematically stronger than on bud flush. Bud flush may display higher phenotypic plasticity as temperatures increase. In contrast, bud set is largely governed by photoperiods, and modulated by temperatures and drought, which results in a more predictable environmental signal from year to year and location to location (Böhlenius *et al.* 2006). In a warming climate, spring phenology can likely respond and advance without much genetic change, as has already been seen in many species (Gienapp *et al.* 2008), provided that the chilling requirement has been met. However, if chilling temperature requirements have not been met, in some cases bud flush may even be delayed (Hänninen & Tanino, 2011), as already seen recently in Tibet (Yu *et al.* 2010). In the fall, a change in bud set date is more likely to require a genetic change in photoperiodic responses. Some studies suggest that the heritability of bud flush is higher than for bud set (Howe *et al.* 2003), but estimates of the additive genetic component are rarely available in the literature. The latitudinal slopes were also much steeper than the altitudinal ones (4.91 and 2.35 days/°C, respectively). Sundblad & Andersson (1995) have suggested that along the altitudinal gradients there may be more gene flow so populations do not become as differentiated. The simple

TABLE 2

Genetic differentiation (Q_{ST}) estimates for the 19 quantitative traits studied in provenance trials.

Trait	Category	Q_{ST} estimates [a]			Qualitative estimation [b]	
		Mean Q_{ST}	Q_{ST} range	Nb [c]	trend	Nb [c]
Dark respiration	Ecophysiology			0	Moderate	2
Leaf mass per area	Ecophysiology	**0.022**	0.000-0.044	2	*Variable*	6
Net assimilation	Ecophysiology	**0.045**	0.015-0.075	2	*Variable*	8
Nitrogen leaf content	Ecophysiology	**0.042**	0.000-0.083	2	*Variable*	6
Photosynthetic capacity	Ecophysiology	**0.101**	0.000-0.201	2	*Variable*	1
Stomatal conductance	Ecophysiology	**0.061**	0.000-0.150	4	*Variable*	4
Stomatal density	Ecophysiology	**0.028**	0.000-0.056	2	Low	5
Water use efficiency (A/gs)	Ecophysiology	**0.075**		1	*Variable*	7
Water use efficiency ($\delta^{13}C$)	Ecophysiology			0	*Variable*	6
Fall frost hardiness	Frost hardiness	**0.581**	0.030-0.890	9	High	10
Spring frost hardiness	Frost hardiness	**0.126**	0.000-0.352	4	*Variable*	3
Winter frost hardiness	Frost hardiness	**0.170**	0.000-0.291	3		0
Growth rate per day	Growth	**0.284**	0.050-0.710	8	Moderate	3
Height increment	Growth	**0.324**	0.040-0.880	36	High	33
Root allocation	Growth	**0.340**	0.251-0.430	2	Moderate	4
Bud flush	Phenology	**0.249**	0.000-0.700	24	Moderate	37
Bud set	Phenology	**0.392**	0.040-0.904	16	High	16
Germination	Phenology	**0.521**	0.200-0.940	6	High	3
Senescence	Phenology	**0.108**	0.080-0.180	5	Low	3

[a] Q_{ST} estimates corresponds to the levels of population differentiation measured either as the proportion of phenotypic variation between populations (V_{pop}) or as the proportion of additive genetic variance between populations (Q_{ST}) in the provenance trials (for more details see Table S1).

[b] Qualitative estimation of genetic differentiation between populations corresponds to studies where no Q_{ST} estimate was available but significance of genetic differentiation was mentioned in the text.

[c] Nb: number of studies used to calculate mean Q_{ST} and Q_{ST} range, and the trend of population differentiation.

calibration factors we used also may not capture all aspects of the environment.

In the large set of provenance trial studies, clinal variation along environmental gradients was very common. In the 112 studies, 309 analyses of clinal variation of different quantitative traits, 243 (78 %) showed evidence of clinal variation with latitude, altitude, and sometimes longitude (Table S1).

Adaptation at range margins

An important hypothesis for species range limits is that gene flow constraints adaptation (Haldane, 1932, Mayr, 1963). Many models suggest that gene flow could limit adaptation, and even more so with asymmetrical gene flow towards small peripheral populations (see Lenormand, 2002 for review). In a model of species range involving local adaptation, a strong coupling between fitness and population size favors a feedback effect (a "migration meltdown") that acts to stabilize a range margin, as exemplified in the now well-known Kirkpatrick and Barton (1997) model. However, there is limited evidence to evaluate this model, and some issues that complicate the predictions. Some models assume that genetic variance is fixed (Kirkpatrick & Barton, 1997, Pease *et al.* 1989), while gene flow may also increase genetic variance and the response to selection (Barton, 2001, Polechova *et al.* 2009). Evidence in *Pinus contorta* suggests that gene flow between populations inhabiting heterogeneous environments can increase levels of standing genetic variation (Yeaman & Jarvis, 2006), but it is unclear whether this effect would be important in other species. In single locus models, gene flow can also decrease the genetic variance due to extinction of local low-frequency alleles and the different approaches are thus not fully reconciled (Lenormand, 2002). Genetic drift can also reduce genetic variance and thus adaptation in peripheral populations (Alleaume-Benharira *et al.* 2006, Bridle *et al.* 2010, Polechova *et al.* 2009), but gene flow may replenish genetic variation. Gene flow may even introduce better adapted genes than local ones, especially in a changing climate (Alleaume-Benharira *et al.* 2006).

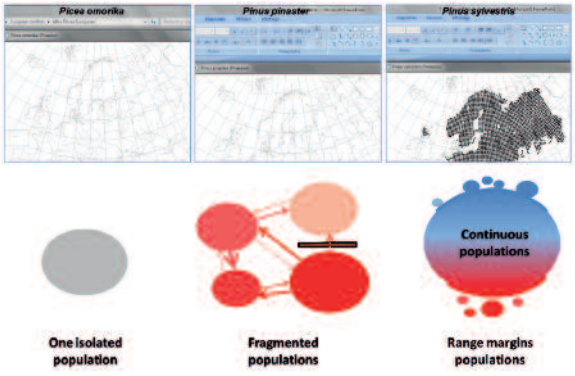
Some environments, in particular some polar or arid range margins, are intrinsically less favorable than others, and would sustain only very low population sizes even after a very long history of adaptation. Mainland-island models of local adaptation implicitly address this issue with population sizes, but spatially continuous models are still more informative. In particular, Nagylaki (1975) showed that extrinsic asymmetries in habitat quality strongly modified or could even compensate for asymmetries in selection across habitats. In other words, alleles showing a local advantage can be maintained despite having considerable antagonistic effects in other habitats, provided that the local habitat is of better quality (Nagylaki, 1978). Incorporating differences in carrying capacity in quantitative models could critically affect the potential for population adaptation (Bridle *et al.* 2010).

FIGURE 3

Schemes of the population models used to discuss evolutionary responses

The three different schematic models of population structure encountered in tree species illustrated by the different cases of *Picea omorika* (one limited population), *Pinus pinaster* (several fragmented populations) and *Pinus sylvestris* (large and continuous population). The colour of the circle indicates the environmental condition of the population which is either undefined (in grey) or following a temperature gradient from warm (in red) to cold (in blue).

The arrows represent gene flow connecting populations, with thickness indicating gene flow intensity. For the fragmented populations the brown line symbolizes a physical barrier to gene flow, such as a mountain.



The leading and the trailing edge of migrating tree distributions face quite different challenges due to the warming climate (Hampe & Petit, 2005). At the southern range edge (in northern hemisphere), the distributions are likely already limited by high temperatures or drought conditions, and associated biotic and abiotic stresses, whereas at the northern margin, many populations have been limited by the cold temperatures (Rehfeldt *et al.* 2002). For the southern margin, at least at low altitudes, the environment is clearly deteriorating. The risk of extinctions will come from the interplay of multiple factors. In particular, the reduction of water availability and a longer growing season with excessively warm temperatures (IPCC, 2007) could lead to massive diebacks of trees due to drought stress or carbon starvation (Bréda *et al.* 2006, Sabate *et al.* 2002) higher mortality due to reduced defense of trees against insects (Rouault *et al.* 2006), and more frequent forest fires (Mouillot & Field, 2005). Increased mortality due to heat and drought stress has already been observed in many locations globally (Allen *et al.* 2010). The impact of environmental change will be higher in small populations due to high demographic or environmental stochasticity (Hampe & Jump, 2011).

At the southern margin, there are no populations further south contributing genes conferring necessary adaptation, but gene flow from similar environments could still increase the variance within populations (Barton, 2001).

TABLE 3

Slopes of the linear regressions of a) bud flush and b) bud set along altitudinal and latitudinal gradients

a)

Gradient	Species	Pop [a]	Cline	Slopes	References
Altitudinal	*Abies amabilis*	5	High early	-1.18	Worrall (1983)
	Abies lasiocarpa	2	High early	-0.83	Worrall (1983)
	Fagus sylvatica	9	High early	-0.43	Vitasse *et al.* (2009)
	Fagus sylvatica	158	High early	-0.17	von Wuehlisch *et al.* (1995)
	Pseudotsuga menziesii	7	High early	-4.38	Acedevo-Rodriguez *et al.* (2006)
	Pseudotsuga menziesii	18	No cline	0.00	Rehfeldt (1978)
	Picea abies	23	No cline	-0.22	Chmura (2006)
	Picea abies	8	No cline	-0.03	Skroppa & Magnussen (1993)
	Abies alba	6	No cline	-0.20	Vitasse *et al.* (2009)
	Acer pseudoplatanus	7	No cline	-0.20	Vitasse *et al.* (2009)
	Fraxinus excelsior	9	Low early	1.90	Vitasse *et al.* (2009)
	Larix occidentalis	82	Low early	0.23	Rehfeldt (1982)
	Quercus petraea	10	Low early	1.15	Alberto *et al.* (2011)
	Quercus rubra	4	Low early	1.93	Mc Gee (1973)
	Total			**-0.17**	
Latitudinal	*Picea abies*	9	North early	-2.08	Sogaard *et al.* (2008)
	Picea glauca	63	No cline	0.43	Li *et al.* (1997a)
	Picea sitchensis	17	No cline	-0.08	Mimura & Aitken (Mimura & Aitken, 2007)
	Pinus strobus	66	No cline	-0.83	Li *et al.* (1997b)
	Populus balsamifera	4	No cline	0.10	Farmer (1993)
	Fagus sylvatica	158	South early	0.20	von Wuehlisch *et al.* (1995)
	Quercus petraea	16	South early	4.17	Deans & Harvey (1996)
	Tsuga heterophylla	8	South early	2.17	Hannerz *et al.* (1999)
	Total			**0.51**	

b)

Gradient	Species	Pop [a]	Cline	Slopes	References
Altitudinal	*Abies lasiocarpa*	5	High early	-3.33	Green (2005)
	Larix occidentalis	82	High early	-1.28	Rehfeldt (1982)
	Picea abies	23	High early	-9.07	Chmura (2006)
	Picea abies	8	High early	-2.63	Skroppa & Magnussen (1993)
	Picea glauca	5	High early	-1.00	Green (2005)
	Pinus contorta	5	High early	-1.67	Green (2005)
	Pinus contorta	173	High early	-0.22	Rehfeldt (1988)
	Pseudotsuga menziesii	7	No cline	0.37	Acedevo-Rodriguez *et al.* (2006)
	Total			**-2.35**	
Latitudinal	*Betula pendula*	7	North early	-4.63	Viherä-Aarnio *et al.* (2005)
	Picea glauca	63	North early	-3.83	Li *et al.* (1997a)
	Picea sitchensis	17	North early	-4.90	Mimura & Aitken (2007)
	Pinus strobus	66	North early	-3.33	Li *et al.* (1997b)
	Pinus sylvestris	4	North early	-5.00	Hurme *et al.* (1997)
	Pinus sylvestris	4	North early	-2.35	Notivol *et al.* (2007)
	Pinus sylvestris	2	North early	-6.83	Savolainen *et al.* (2004)
	Populus balsamifera	4	North early	-5.00	Farmer (1993)
	Populus tremula	12	North early	-8.33	Luquez *et al.* (2008)
	Total			**-4.91**	

Slopes of linear regressions are given for each study and expressed as days/°C (for details about the calculation see in the text and for references see Table S1). No cline indicates a non-significant regression.

[a] Number of populations in the provenance trial.

Experimental evidence of gene flow from like populations increasing fitness at warm range-edges exists for some plant species (e.g., Mimulus species Sexton *et al.* 2011), and long distance dispersal can be important in fragmented landscapes (Fayard *et al.* 2009, Klein *et al.* 2006, Kremer *et al.* 2012).

Until now, the severe climatic conditions at boreal northern range margins have constrained growth, pollen production, seed maturation and dispersal (Sarvas, 1962, Savolainen, 1996), as well as survival (Persson, 1998), and have limited expansion to the north (Chuine & Beaubien, 2001, Morin *et al.* 2007). In the northernmost areas, temperatures are expected to increase by about 4°C (Kattsov & Källen, 2005). Ecophysiologists have used the immediate plastic responses of trees to increased temperature to predict changes in species composition (Kellomäki & Kolström, 1992, Kellomäki *et al.* 2001). However, these predictions have not explicitly taken into account the possibilities of genetic response (Davis & Shaw, 2001, O'Neill *et al.* 2008). The warming in the north will improve survival, increase growth (Rehfeldt *et al.* 2002, Reich & Oleksyn, 2008), increase sexual reproduction (Andalo *et al.* 2005), and increase pollen production (Savolainen *et al.* 2011). Based on modeling studies, pollen and seed are predicted to be dispersed further than before (Kuparinen *et al.* 2009, Kuparinen *et al.* 2010). Production of mature filled seed will likely increase many fold (Kellomäki *et al.* 1997) and the warmer air and soil may result in improved germination and establishment. Northern range margin populations are already colonizing more northern and higher altitude areas (Chen *et al.* 2011, Juntunen *et al.* 2006, Kullman, 2002). The increased survival rates of existing, established trees may however reduce establishment opportunities for better adapted genotypes generated by gene flow and local selection (Kuparinen *et al.* 2010).

At altitudinal range limits, adaptation could be facilitated by the short geographical distance between populations, associated with low climate change velocity (Loarie *et al.* 2009). Gene flow from populations at low altitudes could help the populations at higher altitudes to adapt, as has already been observed, *e.g.* in oak phenological shifts *in situ* (Alberto *et al.* (2010). Both colonization of new areas at higher altitudes, if available, and local selection aided by gene flow may contribute to adaptation, as many altitudinal gradients show clinal genetic differentiation (see above).

Conclusions and suggestions for future research

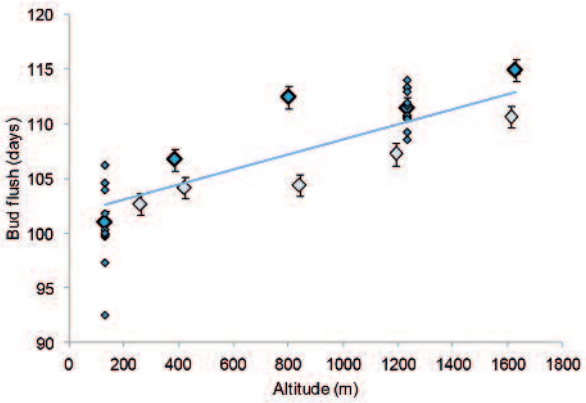
Forest trees are exceptionally well characterized with respect to adaptive quantitative variation, and with respect to responses to different climatic variables. The existing set of provenance trials can be used to extract even more information, for instance on the level of local adaptation, or even on the strength of selection, when the

FIGURE 4

Clines of phenological traits along environmental gradients

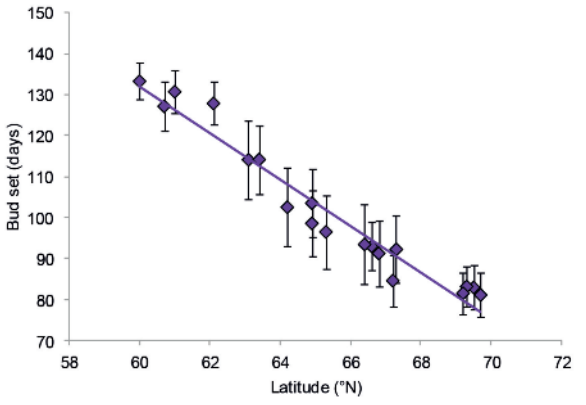
Timing of bud flush along an altitudinal gradient in *Quercus petraea*, based on data from Alberto *et al.* (2011).

The timing of bud flush is expressed as the number of days from 1st January to reach the fourth developmental stage of leaf unfolding. Means of populations (large diamonds) are plotted against the altitude of origin. Bars represent standard deviations of the populations. Means of maternal tree progenies (small diamonds) in populations located at 131 m and 1235 m of elevation illustrate high additive genetic variance within populations, slightly decreasing with increasing altitude. Dark colored points represent populations and maternal trees from Luz valley while light colored points represent populations from Ossau valley.



Timing of bud set along a latitudinal gradient in *Pinus sylvestris*, based on data from Mikola (1982).

The timing of bud set is measured as the number of days from the day of sowing. Means of populations (large diamonds) are plotted against latitude of origin. Bars represent standard deviations of the populations.



data sets are further analyzed. Long-term estimates of the strength of selection, in particular in natural conditions, would be very valuable for providing parameter range estimates for the prediction models. New reciprocal transplant experiments are needed for commercially less-important species, which may be most threatened, but which are under-represented in existing provenance trials. Furthermore, the present provenance trials ignore the likely important early fitness components of germination and establishment – these components also need to be studied (as is being done in herbaceous plants (Huang *et al.* 2010, Stanton-Geddes *et al.* 2012). The new experiments should include field sites at and beyond existing range margins. Experiments in controlled growth chambers can also help identify those abiotic aspects of temperature and moisture regimes to which populations are locally adapted, and to generate climatic regimes analogous to those predicted for the coming century.

The role of plasticity and its interaction with natural selection is just starting to be explored in the climate change context (Chevin *et al.* 2010) – provenance trials can also provide more information on these aspects. The extent and significance of adaptive phenotypic plasticity is still debated (Valladares *et al.* 2007), and experimental studies on range margins are still few (Angert, 2009, Stanton-Geddes *et al.* 2012). Wang's (2010) universal response function approach could be used as a mechanistic model to predict population responses.

Commercially less-important species are poorly represented in previously established common gardens, whether they have narrow or wide distributions. The species with smaller ranges are especially vulnerable. Are these species locally adapted to climate? Do these species have limited adaptive potential due to their historically small effective population sizes? While many important boreal and temperate species in the northern hemisphere (and some eucalypts or tropical acacias) have been extensively studied, there is much less information on subtropical or tropical species, which are outside the scope of this review. These species will also be affected by the changing climate, through both abiotic and many complex biotic factors.

Most of the studies on quantitative traits have been conducted in spaced, reasonably well-tended provenance trial experiments. Within or between species interactions, such as competition or diseases have largely been ignored. Many between-species interactions depend on the synchronous timing of events in the different species. Even before any evolutionary responses, phenotypic responses will affect such biotic interactions (Gilman *et al.* 2010, Yang & Rudolf, 2010). During the past decade phenological shifts have been already observed between trees and pest populations (Desprez-Loustau *et al.* 2010, Gordo & Sanz, 2010, van Asch *et al.* 2007, Visser & Holleman, 2001).

Much of the information on northern trees has been accumulated through decades of field experiments. Combining genomic tools with results from the quantitative and ecological approaches can significantly aid in predicting selection responses to climate change (for crop plants, see Morrell *et al.* 2012). Genomic studies will allow researchers to examine the geographical pattern of alleles conferring adaptation – are they globally occurring alleles with varying frequencies or very localized ones? Coupled with studies at the quantitative trait level, genomic surveys will aid in assessing the prospects for adaptation at the level of the population. Furthermore, the contribution of epigenetic and maternal effects to phenotypic variation needs to be assessed.

This review has pointed to several areas where management and breeding can possibly contribute to maintenance of populations. An evaluation of such options is beyond the scope of this review (see e.g. McLachlan *et al.* 2007).

In conclusion, the concordant patterns of current local adaptation among tree populations in numerous northern species in Europe and North America show that selection has repeatedly established such patterns. Populations facing the largest evolutionary challenges are at the range margins, but the northern and southern (or higher and lower latitude) margins face quite different limitations. Better data and models are thus necessary to evaluate accurately whether natural selection, and migration, may again allow evolutionary responses for populations to sufficiently match their new climates.

Acknowledgements

We acknowledge support from the project FoResTTraC FP7 – 244096 and thank Martin Lascoux and Matti Salmela for comments on the manuscript. Jouni Mikola kindly provided access to the original raw data related to Figure 4b. F.A. thanks the University Joseph Fourier of Grenoble for providing opportunity to work on the manuscript during the last stages of the redaction. SCG-M and RA were funded by RTA2010-00120-C02-01 and AdapTree (funded by Genome Canada and Genome BC) supported S.N.A. and S.Y. during the preparation and publication of this manuscript.

REFERENCES

Acevedo-Rodriguez R, Vargas-Hernandez JJ, Lopez-Upton J, Mendoza JV (2006) Effect of geographic origin and nutrition on shoot prenology of Mexican Douglas-Fir (Pseudotsuga sp.) seedlings. *Agrociencia*, **40**, 125-137.

Aitken SN, Yeaman S, Holliday JA, Wang TL, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, **1**, 95-111.

Alberto F, Bouffier L, Louvet JM, Lamy JB, Delzon S, Kremer A (2011) Adaptive responses for seed and leaf phenology in natural populations of sessile oak along an altitudinal gradient. *Journal of Evolutionary Biology,* **24**, 1442-1454.

Alberto F, Niort J, Derory J, Lepais O, Vitalis R, Galop D, Kremer A (2010) Population differentiation of sessile oak at the altitudinal front of migration in the French Pyrenees. *Molecular Ecology*, **19**, 2626-2639.

Alleaume-Benharira M, Pen IR, Ronce O (2006) Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *Journal of Evolutionary Biology*, **19**, 203-215.

Allen CD, Macalady AK, Chenchouni H *et al.* (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, **259**, 660-684.

Andalo C, Beaulieu J, Bousquet J (2005) The impact of climate change on growth of local white spruce populations in Quebec *Forest Ecology and Management*, **205**, 169-182.

Angert AL (2009) The niche, limits to species' distributions, and spatiotemporal variation in demography across the elevation ranges of two monkeyflowers. *Proceedings of the National Academy of Sciences*, **106**, 19693-19698.

Atwell S, Huang YS, Vilhjalmsson BJ *et al.* (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*, **465**, 627-631.

Barton N (2001) Adaptation at the edge of a species' range. In: *Integrating genetics and ecology in a spatial context*. (eds Silvertown J, Antonovics J) pp 365-392. London, Blackwell.

Barton NH (1999) Clines in polygenic traits. *Genetical Research*, **74**, 223-236.

Barton NH, Keightley PD (2002) Understanding quantitative genetic variation. *Nature Reviews Genetics*, **3**, 11-21.

Basler D, Körner C (2012) Photoperiod sensitivity of bud burst in 14 temperate forest tree species. *Agricultural and Forest Meteorology*, **165**, 73-81.

Benito-Garzón M, Alia R, Robson TM, Zavala MA (2011) Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography*, 766-778.

Beuker E, Valtonen E, Repo T (1998) Seasonal variation in the frost hardiness of Scots pine and Norway spruce in old provenance experiments in Finland. *Forest Ecology and Management*, **107**, 87-98.

Björklund M, Ranta E, Kaitala V, Bach LA, Lundberg P, Stenseth NC (2009) Quantitative trait evolution and environmental change. *PLoS ONE*, **4**, e4521.

Bréda N, Huc R, Granier A, Dreyer E (2006) Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences. *Annals of Forest Science*, **63**, 625-644.

Bridle JR, Polechova J, Kawata M, Butlin RK (2010) Why is adaptation prevented at ecological margins? New insights from individual-based simulations. *Ecology Letters*, **13**, 485-494.

Bräutigam K, Vining KJ, Lafon-Placette C *et al.* (2013) Epigenetic regulation of adaptive responses of forest tree species to the environment. *Ecology and Evolution*, **3**, n/a-n/a.

Buckler ES, Holland JB, Bradbury PJ *et al.* (2009) The Genetic Architecture of Maize Flowering Time. *Science*, **325**, 714-718.

Burger R, Lynch M (1995) Evolution and extinction in a changing environments - a quantitative genetic analysis. *Evolution*, **49**, 151-163.

Böhlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O (2006) CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science*, **312**, 1040-1043.

Caffarra A, Donelly A, Chuine I (2011) Modelling the timing of Betula pubescens budburst. II. Integrating the complex effects of photoperiod into process-based models of budburst. *Climate Research*, **46**, 159-170.

Charlesworth B, Charlesworth D (2010) *Elements of evolutionary genetics*, Greenwood Village, Co., Roberts & Co.

Cheddadi R, Vendramin GG, Litt T *et al.* (2006) Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecology and Biogeography*, **15**, 271-282.

Chen IC, Hill JK, Ohlemuller R, Roy DB, Thomas CD (2011) Rapid range shifts of species associated with high levels of climate warming. *Science*, **333**, 1024-1026.

Chen J, Källman T, Gyllenstrand N, Lascoux M (2009) New insights on the speciation history and nucleotide diversity of three boreal spruce species and a Tertiary relict. *Heredity*, **104**, 3-14.

Chevin L-M, Lande R (2010) When do adaptive plasticity and genetic evolution prevent extinction of a density-regulated population? *Evolution*, **64**, 1143-1150.

Chevin L-M, Lande R, Mace GM (2010) Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biology*, **8**, e1000357.

Chmura DJ (2006) Phenology differs among Norway spruce populations in relation to local variation in altitude of maternal stands in the Beskidy Mountains. *New Forests*, **32**, 21-31.

Chuine I, Beaubien EG (2001) Phenology is a major determinant of tree species range. *Ecology Letters*, **4**, 500-510.

Conner JK, Franks R, Stewart C (2003) Expression fo addtive genetic variances and covariances for wild radish floral traits: comparison between field and greenhouse environments *Evolution*, **57**, 487-495.

Cornelius J (1994) Heritabilities and additive genetic coefficients of variation in forest trees. *Canadian Journal of Forest*. *Research*, **24**, 372-379.

Cumbie WP, Eckert AJ, Wegrzyn JL, Whetten R, Neale DB, Goldfarb B (2011) Association genetics of carbon isotope discrimination, height, and foliar nitrogen in a natural population of *Pinus taeda* L. *Heredity*, **107**, 105-114.

Davis MB, Shaw RG (2001) Range shifts and adaptive responses to quaternary climate changes. *Science*, **292**, 673-679.

De Carvalho D, Ingvarsson PK, Joseph J *et al.* (2010) Admixture facilitates adaptation from standing variation in the European aspen (*Populus tremula* L.), a widespread forest tree. *Molecular Ecology*, **19**, 1638-1650.

Deans JD, Harvey FJ (1996) Frost hardiness of 16 European provenances of sessile oak growing in Scotland. *Forestry*, **69**, 5-11.

Desprez-Loustau ML, Vitasse Y, Delzon S, Capdevielle X, Marcais B, Kremer A (2010) Are plant pathogen populations adapted for encounter with their host? A case

study of phenological synchrony between oak and an obligate fungal parasite along an altitudinal gradient. *Journal of Evolutionary Biology*, **23**, 87-97.

Duputie A, Massol F, Chuine I, Kirkpatrick M, Ronce O (2012) How do genetic correlations affect species range shifts in a changing environment? *Ecology Letters*, **15**, 251-259.

Eckert AJ, Bower AD, Wegrzyn JL *et al.* (2009) Asssociation genetics of Coastal Douglas Fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-fardiness related traits. *Genetics*, **182**, 1289-1302.

Eckert AJ, van Heerwaarden J, Wegrzyn JL, Nelson CD, Ross-Ibarra J, Gonzalez-Martinez SC, Neale DB (2010) Patterns of population structure and environmental associations to aridity across the range of Loblolly Pine (*Pinus taeda* L., Pinaceae). *Genetics*, **185**, 969-982.

Edelaar P, Burraco P, Gomez-Mestre I (2011) Comparisons between Q(ST) and F(ST)-how wrong have we been? *Molecular Ecology*, **20**, 4830-4839.

Etterson JR, Shaw RG (2001) Constraint to adaptive evolution in response to global warming. *Science,* **294**, 151-154.

Falconer DS, Mackay TFC (1996) *Introduction to quantitative genetics*, Harlow, UK, Longman.

Farmer RE (1993) Latitudinal variation in height and phenology of balsam poplar. *Silvae Genetica*, **42**, 148-153.

Fayard J, Klein EK, Lefevre F (2009) Long distance dispersal and the fate of a gene from the colonization front. *Journal of Evolutionary Biology*, **22**, 2171-2182.

Fisher RA (1918) The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, **52**, 399-433.

Fisher RA (1930) *The genetical theory of natural selection*, Oxford, Oxford Univ. Press.

Franks SJ, Hoffmann AA (2012) Genetics of Climate Change Adaptation. *Annual Review of Genetics*, **46**, 185-208.

Gienapp P, Teplitsky C, Alho JS, Mills JA, Merilä J (2008) Climate change and evolution: disentangling environmental and genetic responses. *Molecular Ecology*, **17**, 167-178.

Gilman SE, Urban MC, Tewksbury J, Gilchrist GW, Holt RD (2010) A framework for community interactions under climate change. *Trends in Ecology & Evolution*, **25**, 325-331.

Goldstein DB, Holsinger KE (1992) Maintenance of polygenic variation in spatially structured populations: roles for local mating and genetic redundancy. *Evolution*, **46**, 412-429.

Gomulkiewicz R, Holt RD (1995) When does evolution by natural selection prevent extinction. *Evolution*, **49**, 201-207.

Gomulkiewicz R, Holt RD, Barfield M, Nuismer SL (2010) Genetics, adaptation, and invasion in harsh environments. *Evolutionary Applications*, **3**, 97-108.

Gomulkiewicz R, Houle D (2009) Demographic and genetic constraints on evolution. *American Naturalist*, **174**, 218-229.

Gordo O, Sanz JJ (2010) Impact of climate change on plant phenology in Mediterranean ecosystems. *Global Change Biology*, **16**, 1082-1106.

Grattapaglia D, Resende MDV (2011) Genomic selection in forest tree breeding. *Tree Genetics & Genomes*, **7**, 241-255.

Green DS (2005) Adaptive strategies in seedlings of three co-occurring, ecologically distinct northern coniferous tree species across an elevational gradient. *Can. J. For. Res.*, **35**, 910-917.

Grivet D, Sebastiani F, Alía R *et al.* (2011) Molecular footprints of local adaptation in two Mediterranean conifers. *Molecular Biology and Evolution*, **28**, 101-116.

Haldane JBS (1932) *The causes of evolution*, London, Longmans.

Hampe A, Jump AS (2011) Climate Relicts: Past, Present, Future. *Annual Review of Ecology, Evolution, and Systematics*, **42**, 313-333.

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461-467.

Hancock AM, Brachi B, Faure N *et al.* (2011) Adaptation to climate across the *Arabidopsis thaliana* genome. *Science*, **334**, 83-86.

Hannerz M, Aitken SN, King JN, Budge S (1999) Effects of genetic selection for growth on frost hardiness in western hemlock. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, **29**, 509-516.

Hedrick PW (2006) Genetic polymorphism in heterogeneous environments: the age of genomics. *Annual Review of Ecology Evolution and Systematics*, **37**, 67-93.

Hendry AP, Day T, Taylor EB (2001) Population mixing and the adaptive divergence of quantitative traits in discrete populations: A theoretical framework for empirical tests. *Evolution*, **55**, 459-466.

Hendry AP, Kinnison MT (1999) Perspective: The pace of modern life: Measuring rates of contemporary microevolution. *Evolution*, **53**, 1637-1653.

Heuertz M, De Paoli E, Källman T *et al.* (2006) Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce (*Picea abies* (L.) Karst). *Genetics*, **174**, 2095-2105.

Hill JK, Griffiths HM, Thomas CD (2011) Climate change and evolutionary adaptations at species' range margins. *Annual Review of Entomology*, **56**, 143-159.

Hill WG, Goddard ME, Visscher PM (2008) Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics*, **4**, e1000008.

Hoffmann AA, Sgrò CM (2011) Climate change and evolutionary adaptation. *Nature*, **470**, 479-485.

Holliday JA, Ritland K, Aitken SN (2010a) Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytologist*, **188**, 501-514.

Holliday JA, Wang T, Aitken S (2012) Predicting Adaptive Phenotypes From Multilocus Genotypes in Sitka Spruce (Picea sitchensis) Using Random Forest. *G3*: *Genes|Genomes|Genetics*, **2**, 1085-1093.

Holliday JA, Yuen M, Ritland K, Aitken SN (2010b) Postglacial history of a widespread conifer produces inverse clines in selective neutrality tests. *Molecular Ecology*, **19**, 3857-3864.

Howe GT, Aitken SN, Neale DB, Jermstad KD, Wheeler NC, Chen THH (2003) From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees. *Canadian Journal of Botany*, **81**, 1247-1266.

Huang X, Schmitt J, Dorn L *et al.* (2010) The earliest stages of adaptation in an experimental plant population: strong selection on QTLS for seed dormancy. *Molecular Ecology*, **19**, 1335-1351.

Hurme P, Repo T, Pääkkönen T, Savolainen O (1997) Climatic adaptation of bud set and frost hardiness in Scots pine (*Pinus sylvestris* L.). *Can. J. For. Res.*, **27**, 716-523.

Hänninen H, Tanino K (2011) Tree seasonality in a warming climate. *Trends in Plant Science*, **16**, 312-416.

Ingvarsson PK, Garcia MV, Luquez V, Hall D, Jansson S (2008) Nucleotide polymorphism and phenotypic associations within and around the phytochrome B2 locus in European aspen (*Populus tremula*, Salicaceae). *Genetics*, **178**, 2217-2226.

IPCC (2001) The Scientific Basis In: *Climate Change 2001*. (eds Houghton JT, Ding Y, Griggs DJ, Noguer M, Van Der Linden PJ, Dai X, Maskell K, Johnsson CA) pp 881. Cambridge, Cambridge University Press.

IPCC (2007) Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. In: *Climate Change 2007: Impacts, Adaptation and Vulnerability.* Canmbridge and New York, Cambridge University Press.

Iwata H, Hayashi T, Tsumura Y (2011) Prospects for genomic selection in conifer breeding: a simulation study of *Cryptomeria japonica. Tree Genetics Genomes*.

Jalas J, Suominen T (1973) Atlas Floreae Europaea. Coniferales.

Jermstad KD, Bassoni DL, Jech KS, Ritchie GA, Wheeler NC, Neale DB (2003) Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas Fir. III. Quantitative trait loci-by-environment interactions. *Genetics*, **165**, 1489-1506.

Johannes F, Porcher E, Teixeira FK *et al.* (2009) Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genetics*, **5**, e1000530.

Jump AS, Matyas C, Penuelas J (2009) The altitude-for-latitude disparity in the range retractions of woody species. *Trends in Ecology & Evolution*, **24**, 694-701.

Juntunen V, Neuvonen S, Sutinen R (2006) Männyn puurajan muutokset viimeisen 400 vuoden aikana ja metsänraja-puuraja vaihettumisvyöhykkeen ikärakenne. *Metlan työraportteja*, **25**, 25-32.

Kattsov VM, Källen E (2005) Future climate change: modeling and scenarios for the Arctic. In: *Arctic Climate Impact Assessment - scientific report.* pp 100-150. Cambridge, UK, Cambridge Univ. Press.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225-1241.

Kayihan GC, Nelson CD, Huber DA, Amerson HV, White TL, Davis JM (2010) Clonal evaluation for fusiform rust disease resistance: effects of pathogen virulence and disease escape. *Canadian Journal of Forest Research*, **40**, 1042-1050.

Kellomäki S, Kolström M (1992) Simulation of tree species composition and organic matter accumulation in Finnish boreal forests under changing climatic conditions. *Vegetatio*, **102**, 47-68.

Kellomäki S, Rouvinen I, Peltola H, Strandman H, Steinbrecher R (2001) Impact of global warming on the tree species composition of boreal forests in Finland and effects on emission of isoprenoids. *Global Change Biology*, **7**, 531-544.

Kellomäki S, Väisänen H, Kolström T (1997) Model computatons on the effects of elevating temperature and atmosphere CO2 on the regeneration of Scots pine at te timber line in Finland. *Climatic Change*, **37**, 683-708.

Kingsolver JG, Diamond SE (2011) Phenotypic selection in natural populations: what limits directional selection? *American Naturalist*, **177**, 346-357.

Kingsolver JG, Hoekstra HE, Hoekstra JM *et al.* (2001) The strength of phenotypic selection in natural populations. *American Naturalist*, **157**, 245-261.

Kirkpatrick M, Barton NH (1997) Evolution of a species' range. *American Naturalist*, **150**, 1-23.

Klein EK, Lavigne C, Gouyon PH (2006) Mixing of propagules from discrete sources at long distance: comparing a dispersal tail to an exponential. *BMC Ecology*, **6**.

Kozlowski T, Pallardy S (2002) Acclimation and adaptive responses of woody plants to environmental stresses. *The Botanical Review*, **68**, 270-334.

Kramer K (1995) Phenotypic plasticity of the phenology of 7 European tree species in relation to climate warming. *Plant Cell and Environment*, **18**, 93-104.

Kremer A, Le Corre V (2012) Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity*, **108**, 375-385.

Kremer A, Ronce O, Robledo-Arnuncio JJ *et al.* (2012) Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters*, **15**, 378-392.

Kullman L (2002) Rapid recent range-margin rise of tree and shrub species in the Swedish Scandes. *Journal of Ecology*, **90**, 68-77.

Kuparinen A, Katul G, Nathan R, Schurr FM (2009) Increases in air temperature can promote wind-driven dispersal and spread of plants. *Proceedings Royal Society B*, **276**, 3081-3087.

Kuparinen A, Savolainen O, Schurr FM (2010) Increased mortality can promote evolutionary adaptation of forest trees to climate change. *Forest Ecology and Management*, **259**, 1003-1008.

Körner C, Basler D (2010) Phenology Under Global Warming. *Science*, **327**, 1461-1462.

Lande R (2009) Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology*, **22**, 1435-1446.

Langlet O (1971) Two hundred years of genecology. *Taxon*, **20**, 653-721.

Latta RG (1998) Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. *American Naturalist*, **151**, 283-292.

Laurie CC, Chasalow SD, LeDeaux JR *et al.* (2004) The genetic architecture of response to long-term artificial selection for oil concentration in the maize kernel. *Genetics*, **168**, 2141-2155.

Le Corre V, Kremer A (2012) The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, **21**, 1548-1566.

LeCorre V, Kremer A (2003) Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics*, **164**, 1205-1219.

Leimu R, Fischer M (2008) A meta-analysis of local adaptation in plants. *PLoS ONE*, **3**, e4010.

Leinonen T, O'Hara RB, Cano JM, Merila J (2008) Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology*, **21**, 1-17.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends Ecol. Evol.*, **17**, 183-189.

Li P, Beaulieu J, Bousquet J (1997a) Genetic structure and patterns of genetic variation among populations in eastern white spruce. *Can. J. For. Res.*, **27**, 19-198.

Li P, Beaulieu J, Daoust G, Plourde A (1997b) Patterns of adaptive genetic variation in eastern white pine (*Pinus strobus*) from Quebec. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, **27**, 199-206.

Lindner M, Maroschek M, Netherer S *et al.* (2010) Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. *Forest Ecology and Management*, **259**, 698-709.

Loarie SR, Duffy PB, Hamilton H, Asner GP, Field CB, Ackerly DD (2009) The velocity of climate change. *Nature*, **462**, 1052-1055.

Luquez V, Hall D, Albrectsen BR, Karlsson J, Ingvarsson P, Jansson S (2008) Natural phenological variation in aspen (Populus tremula): the SwAsp collection. *Tree Genetics & Genomes*, **4**, 279-292.

Lynch M (1996) A quantitative-genetic perspective on conservation issues. In: *Conservation genetics. Case histories from nature*. (eds Avise JC, Hamrick JL) pp 471-501. New York, Chapman and Hall.

Lynch M, Lande R (1993) Evolution and extinction in response to environmental change. In: *Biotic interactions and global change*. (eds Kareiva PM, Kingsolver JG, Huey RB) pp 234-250. Sunderland, Mass., Sinauer.

Magri D, Vendramin GG, Comps B *et al.* (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, **171**, 199-221.

Mayr E (1963) *Animal species and evolution* Cambridge, Mass., Harvard Univ. Press.

McGee CE (1973) Is variation in budbreak of Red oak the result of heredity or environment? In: *Tree Improvement and Genetics Proceedings - Southern Forest Tree Improvement Conference*. pp. 185-189.

McLachlan JS, Clark JS, Manos PS (2005) Molecular indicators of tree migration capacity under rapid climate change. *Ecology*, **86**, 2088-2098.

McLachlan JS, Hellmann JJ, Schwartz MW (2007) A framework for debate of assisted migration in an era of climate change. *Conservation Biology*, 21, 297-302.

Menzel A, Fabian P (1999) Growing season extended in Europe. *Nature*, **397**, 659-659.

Menzel A, Sparks TH, Estrella N *et al.* (2006) European phenological response to climate change matches the warming pattern. *Global Change Biology*, **12**, 1969-1976.

Mikola J (1982) Bud-set phenology as an indicator of climatic adaptation of Scots pine in Finland. *Silva Fennica*, **16**, 178-184.

Mimura M, Aitken SN (2007) Adaptive gradients and isolation-by-distance with postglacial migration in in *Picea sitchensis*. *Heredity*, **98**, 224-232.

Morgenstern EK (1978) Range-wide variation of black spruce. *Canadian Journal of Forest Research*, **8**, 463-473.

Morgenstern EK (1996) *Geographic variation in forest trees: genetic basis and application of knowledge in silviculture*, Vancouver, B. C., UBC Press.

Morin X, Augspurger C, Chuine I (2007) Process-based modeling of species' distributions: What limits temperate tree species' range boundaries? *Ecology*, **88**, 2280-2291.

Morin X, Thuiller W (2009) Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology*, **90**, 1301-1313.

Morrell PL, Buckler ES, Ross-Ibarra J (2012) Crop genomics: advances and applications. *Nature Reviews Genetics*, **13**, 85-96.

Mouillot F, Field CB (2005) Fire history and the global carbon budget: a 1 degrees x 1 degrees fire history reconstruction for the 20th century. *Global Change Biology*, **11**, 398-420.

Nagylaki T (1975) Conditions for existence of clines. *Genetics*, **80**, 595-615.

Nagylaki T (1978) Clines with asymmetric migration *Genetics*, **88**, 813-827.

Namroud MC, Beaulieu J, Juge N, Laroche J, Bousquet J (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. *Molecular Ecology*, **17**, 3599-3613.

Neale DB, Kremer A (2011) Forest tree genomics: growing resources and applications. *Nature Reviews Genetics*, **12**, 111-122.

Netherer S, Schopf A (2010) Potential effects of climate change on insect herbivores in European forests-General aspects and the pine processionary moth as specific example. *Forest Ecology and Management*, **259**, 831-838.

Nicotra AB, Atkin OK, Bonser SP *et al.* (2010) Plant phenotypic plasticity in a changing climate. *Trends in Plant Science*, **15**, 684-692.

Niinemets U (2010) Responses of forest trees to single and multiple environmental stresses from seedlings to mature plants: Past stress history, stress interactions, tolerance and acclimation. *Forest Ecology and Management*, **260**, 1623-1639.

Notivol E, Garcia-Gil MR, Alia R, Savolainen O (2007) Genetic variation of growth rhythm traits in the limits of a latitudinal cline in Scots pine. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, **37**, 540-551.

O'Neill GA, Hamann A, Wang TL (2008) Accounting for population variation improves estimates of the impact of climate change on species' growth and distribution. *Journal of Applied Ecology*, **45**, 1040-1049.

Olson-Manning CF, Wagner MR, Mitchell-Olds T (2012) Adaptive evolution: evaluating empirical support for theoretical predictions. *Nature Reviews Genetics*, **13**, 867-877.

Orr HA (1998) The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution*, **52**, 935-949.

Parducci L, Jorgensen T, Tollefsrud MM *et al.* (2012) Glacial Survival of Boreal Trees in Northern Scandinavia. *Science*, **335**, 1083-1086.

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution and Systematics.*, **37**, 637-669.

Pease CM, Lande R, Bull JJ (1989) A model of population growth, dispersal and evolution in a changing environment. *Ecology*, **70**, 1657-1664.

Persson B (1998) Will climate change affect the optimal choice of *Pinus sylvestris* provenances. *Silva Fennica*, **32**, 121-128.

Petit RJ, Brewer S, Bordacs S *et al.* (2002) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, **156**, 49-74.

Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. *Annual Review of Ecology Evolution and Systematics*, **37**, 187-214.

Polechova J, Barton N, Marion G (2009) Species' range: adaptation in space and time. *American Naturalist*, **174**, E186-E204.

Pyhäjärvi T, García-Gil MR, Knürr T, Mikkonen M, Wachowiak W, Savolainen O (2007) Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. *Genetics*, **177**, 1713-1724.

Quesada T, Gopal V, Cumbie WP *et al.* (2010) Association mapping of quantitative disease resistance in a natural population of Loblolly Pine (*Pinus taeda L.*). *Genetics*, **186**, 677-686.

Rehfeldt GE (1978) Genetic differentiation of Douglas fir populations from the northern Rocky Mountains. *Ecology*, **59**, 1264-1270.

Rehfeldt GE (1982) Differentiation of *Larix occidentalis* populations from the northern Rocky Mountains. *Silvae Genetica*, 31, 13-19.

Rehfeldt GE (1988) Ecological genetics of *Pinus contorta* from the Rocky mountains (USA): a synthesis. *Silvae Genet.*, **37**, 131-135.

Rehfeldt GE, Jaquish BC (2010) Ecological impacts and management strategies for western larch in the face of climate-change. *Mitigation and Adaptation Strategies for Global Change*, **15**, 283-306.

Rehfeldt GE, Tchebakova NM, Parfenova YI, Wykoff WR, Kuzmina NA, Milyutin LI (2002) Intraspecific responses to climate in *Pinus sylvestris. Global Change Biology*, **8**, 912-929.

Rehfeldt GE, Wykoff WR, Ying CC (2001) Physiologic plasticity, evolution, and impacts of a changing climate on Pinus contorta. *Climatic Change*, **50**, 355-376.

Rehfeldt GE, Ying CC, Spittlehouse DL, Hamilton DA (1999) Genetic responses to climate change in *Pinus contorta*: niche breadth, climate change, and reforestation. *Ecological Monographs*, **69**, 375-407.

Reich PB, Oleksyn J (2008) Climate warming will reduce growth and survival of Scots pine except in the far north. *Ecology Letters*, **11**, 588-597.

Resende MFR, Munoz P, Acosta JJ *et al.* (2012) Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. *New Phytologist*, **193**, 617-624.

Richards CL, Bossdorf O, Verhoeven KJF (2010) Understanding natural epigenetic variation. *New Phytologist*, **187**, 562-564.

Ritland K (1996) Marker-based method for inferences about quantitative inheritance in natural populations. *Evolution*, **50**, 1062-1073.

Robertson C, Nelson TA, Jelinski DE, Wulder MA, Boots B (2009) Spatial-temporal analysis of species range expansion: the case of the mountain pine beetle, Dendroctonus ponderosae. *Journal of Biogeography*, **36**, 1446-1458.

Rohde A, Bhalerao RP (2007) Plant dormancy in the perennial context. *Trends Plant Sci.*, **15**, 217-223.

Rouault G, Candau JN, Lieutier F, Nageleisen LM, Martin JC, Warzee N (2006) Effects of drought and heat on forest insect populations in relation to the 2003 drought in Western Europe. *Annals of Forest Science*, **63**, 613-624.

Sabate S, Gracia CA, Sanchez A (2002) Likely effects of climate change on growth of Quercus ilex, Pinus halepensis, Pinus pinaster, Pinus sylvestris and Fagus sylvatica forests in the Mediterranean region. *Forest Ecology and Management*, **162**, 23-37.

Sarvas R (1962) Investigations on the flowering and seed crop of *Pinus sylvestris*. *Communicationes Instituti Forestalia Fennica*, **53.4**, 1-198.

Savolainen O (1996) Pines beyond the polar circle: adaptation to stress conditions. *Euphytica*, **92**, 139-145.

Savolainen O, Bokma F, Garcia-Gil R, Komulainen P, Repo T (2004) Genetic variation in cessation of growth and frost hardiness and consequences for adaptation of Pinus sylvestris to climatic changes. *Forest Ecology and Management*, **197**, 79-89.

Savolainen O, Kujala ST, Sokol C *et al.* (2011) Adaptive potential of northernmost tree populations to climate change, with emphasis on Scots pine (*Pinus sylvestris* L.). *Journal of Heredity*, **102**, 526-536.

Savolainen O, Pyhajarvi T, Knurr T (2007) Gene flow and local adaptation in trees. *Annual Review of Ecology Evolution and Systematics*, **38**, 595-619.

Savolainen O, Pyhäjärvi T (2007) Genomic diversity in forest trees. *Current Opinion in Plant Biology*, **10**, 162-167.

Sexton JP, Strauss SY, Rice KJ (2011) Gene flow increases fitness at the warm edge of a species' range. *Proceedings of the National Academy of Sciences*, **108**, 11704-11709.

Shaw RG, Etterson JR (2012) Rapid climate change and the rate of adaptation: insight from experimental quantitative genetics. *New Phytologist*, **195**, 752-765.

Shutyaev AM, Giertych M (1997) Height growth variation in a comprehensive Eurasian provenance experiment of *Pinus sylvestris* L. *Silvae Genetica*, **46**, 332-349.

Sillanpää MJ (2011) On statistical methods for estimating heritability in wild populations. *Molecular Ecology*, **20**, 1324-1332.

Skroppa T, Magnussen S (1993) Provenance variation in shoot growth components of Norway spruce. *Silvae Genetica*, **42**, 111-120.

Skrøppa T, Kohmann K (1997) Adaptation to local conditions after one generation in Norway spruce. *Forest Genetics*, **4**, 171-177.

Sogaard G, Johnsen O, Nilsen J, Junttila O (2008) Climatic control of bud burst in young seedlings of nine provenances of Norway spruce. *Tree Physiology*, **28**, 311-320.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261-4293.

Soto A, Robledo-Arnuncio JJ, González-Martínez SC, Smouse PE, Alía R (2010) Climatic niche and neutral genetic diversity of the six Iberian pine species: a retrospective and prospective view. *Molecular Ecology*, **19**, 1396-1409.

Stanton-Geddes J, Shaw RG, Tiffin P (2012) Interactions between Soil Habitat and Geographic Range Location Affect Plant Fitness. *PLoS ONE*, **7**, e36015.

Sundblad LG, Andersson B (1995) No difference in frost hardiness between high and low altitude Pinus sylvestris (L) offspring. *Scandinavian Journal of Forest Research*, **10**, 22-26.

Tian F, Bradbury PJ, Brown PJ *et al.* (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nature Genetics*, **43**, 159-162.

Turchin MC, Chiang CWK, Palmer CD, Sankararaman S, Reich D, Hirschhorn JN, Genetic Invest A (2012) Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nature Genetics*, **44**, 1015-1019.

Valladares F, Gianoli E, Gomez JM (2007) Ecological limits to plant phenotypic plasticity. *New Phytologist*, **176**, 749-763.

van Asch M, Tienderen PH, Holleman LJM, Visser ME (2007) Predicting adaptation of phenology in response to climate change, an insect herbivore example. *Global Change Biology*, **13**, 1596-1604.

van Asch M, Visser ME (2007) Phenology of forest caterpillars and their host trees: The importance of synchrony. *Annual Review of Entomology*, **52**, 37-55.

Wang T, Hamann A, Yanchuk A, O'Neill GA, Aitken SN (2006) Use of response functions in selecting lodgepole pine populations for future climates. *Global Change Biology*, **12**, 2404-2416.

Wang TL, O'Neill GA, Aitken SN (2010) Integrating environmental and genetic effects to predict responses of tree populations to climate. *Ecological Applications*, **20**, 153-163.

Vendramin GG, Fady B, Gonzalez-Martinez SC *et al.* (2008) Genetically depauperate but widespread: The case of an emblematic mediterranean pine. *Evolution*, **62**, 680-688.

Vihera-Aarnio A, Hakkinen R, Partanen J, Luomajoki A, Koski V (2005) Effects of seed origin and sowing time on timing of height growth cessation of Betula pendula seedlings. *Tree Physiology*, 25, 101-108.

Wilczek AM, Burghardt LT, Cobb AR, Cooper MD, Welch SM, Schmitt J (2010) Genetic and physiological bases for phenological responses to current and predicted climates. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **365**, 3129-3147.

Visser ME, Holleman LJM (2001) Warmer springs disrupt the synchrony of oak and winter moth phenology. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **268**, 289-294.

Vitasse Y, Bresson CC, Kremer A, Michalet R, Delzon S (2010) Quantifying phenological plasticity to temperature in two temperate tree species. *Functional Ecology*, **24**, 1211-1218.

Vitasse Y, Delzon S, Bresson CC, Michalet R, Kremer A (2009) Altitudinal differentiation in growth and phenology among populations of temperate-zone tree species growing in a common garden. *Canadian Journal of Forest Research*, **39**, 1259-1269.

Voltas J, Chambel M, Prada M, Ferrio J (2008) Climate-related variability in carbon and oxygen stable isotopes among populations of Aleppo pine grown in common-garden tests. *Trees-Structure and Function*, **22**, 759-769.

vonWuehlisch G, Krusche D, Muhs HJ (1995) Variation in temperature sum requirement for flushing of beech provenances. *Silvae Genetica,* **44**, 343-346.

Worrall J (1983) Temperature - bud-burst relationships in Amabilis and Subalpine fir provenance tests replicated at different elevations. *Silvae Genetica*, **32**, 203-209.

Worrell R, Cundall EP, Malcolm DC, Ennos RA (2000) Variation among seed sources of silver birch in Scotland. *Forestry*, **73**, 419-435.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97-159.

Wright S (1943) Isolation-by-distance. *Genetics*, **28**, 114-138.

Yakovlev IA, Fossdal CG, Johnsen O (2010) MicroRNAs, the epigenetic memory and climatic adaptation in Norway spruce. *New Phytologist*, **187**, 1154-1169.

Yang LH, Rudolf VHW (2010) Phenology, ontogeny and the effects of climate change on the timing of species interactions. *Ecology Letters*, **13**, 1-10.

Yeaman S, Jarvis A (2006) Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proceedings of the Royal Society B-Biological Sciences*, **273**, 1587-1593.

Yeaman S, Whitlock MC (2011) The genetic architecture of adaptation under migration-selection balance. *Evolution*, **65**, 1897-1911.

Yu H, Luedeling E, Xu J (2010) Winter and spring warming result in delayed spring phenology on the Tibetan Plateau *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 22151-22156.

EPIGENETIC REGULATION OF ADAPTIVE RESPONSES OF FOREST TREE SPECIES TO THE ENVIRONMENT

Bräutigam, Katharina[1]; Vining, Kelly J[2]; Lafon-Placette, Clément[3]; Fossdal, Carl G[4].; Mirouze, Marie[5]; Gutiérrez Marcos, José[6] ; Fluch, Silvia[7]; Fernández Fraga, Mario[8]; Guevara, M. Ángeles[9, 10]; Abarca, Dolores[11]; Johnsen, Øystein[12]; Maury, Stéphane[3]; Strauss, Steven H.[2]; Campbell, Malcolm M.[1, 13]; Rohde, Antje[14]; Díaz-Sala, Carmen[11] ; Cervera, María-Teresa[9, 10]

[1] **Katharina Bräutigam** Centre for the Analysis of Genome Evolution and Function, Department of Cell & Systems Biology, University of Toronto, Toronto, ON M5S 3B2, Canada.
[2] **Kelly J Vining** Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331-5752 USA.
[3] **Clément Lafon-Placette** University of Orléans, UFR-Faculté des Sciences, UPRES EA 1207 'Laboratoire de Biologie des Ligneux et des Grandes Cultures' (LBLGC), INRA, USC1328 'Arbres et Réponses aux Contraintes Hydrique et Environnementales' (ARCHE). Rue de Chartres, BP 6759, F-45067 Orléans, France.
[4] **Carl G. Fossdal** Department of Biology and Environment, Norwegian Forest and Landscape Institute, PO Box 115, N-1431, Aas, Norway.
[5] **Marie Mirouze** Epigenetic Regulations and Seed Development, Institut de Recherche pour le Développement, UMR232 ERL5300 CNRS-IRD, 911 Av. Agropolis, 34394 Montpellier, France.
[6] **José Gutiérrez Marcos** School of Life Sciences, University of Warwick, Wellesbourne, Warkwick, CV35 9EF, United Kingdom.
[7] **Silvia Fluch** Platform for Integrated Clone Management (PICME), Health & Environment Department, AIT Austrian Institute of Technology GmbH. Konrad-Lorenz-Straße 24, 3430 Tulln. Austria.
[8] **Mario Fernández Fraga** Cancer Epigenetics Laboratory, Institute of Oncology of Asturias (IUOPA HUCA), University of Oviedo, Spain.
[9, 10] **M. Ángeles Guevara**
[9] INIA. Forest Research Centre. Dpt. of Forest Ecology and Genetics, Forest Genomics and Ecophysiology group. Crta. La Coruña km 7,5 28040 Madrid. Spain.
[10] Mixed Unit of Forest Genomics and Ecophysiology, INIA/UPM, Madrid, Spain.
[11] **Dolores Abarca** Department of Life Sciences, University of Alcalá. Ctra. Madrid-Barcelona Km. 33,600. 28871, Alcalá de Henares, Madrid. Spain.
[12] **Øystein Johnsen** Department of Plant and Environmental Sciences, Norwegian University of Life Sciences, P. O. Box 5003, N-1432, Ås, Norway
[3] **Stéphane Maury** University of Orléans, UFR-Faculté des Sciences, UPRES EA 1207 'Laboratoire de Biologie des Ligneux et des Grandes Cultures' (LBLGC), INRA, USC1328 'Arbres et Réponses aux Contraintes Hydrique et Environnementales' (ARCHE). Rue de Chartres, BP 6759, F-45067 Orléans, France.
[2] **Steven H. Strauss** Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331-5752 USA.
[1, 13] **Malcolm M. Campbell**
[1] Centre for the Analysis of Genome Evolution and Function, Department of Cell & Systems Biology, University of Toronto, Toronto, ON M5S 3B2, Canada.
[13] Department of Biological Sciences, University of Toronto Scarborough, University of Toronto, 1265 Military Trail, Toronto, ON M1C 1A4, Canada.
[14] **Antje Rohde** Department Plant Growth & Development, Institute of Agriculture and Fisheries Research, Caritasstraat 21, 9090 Melle, Belgium.
[11] **Carmen Díaz-Sala** Department of Life Sciences, University of Alcalá. Ctra. Madrid-Barcelona Km. 33,600. 28871, Alcalá de Henares, Madrid. Spain.

Corresponding Author:
[9, 10] **María-Teresa Cervera**
[9] INIA. Forest Research Centre. Dpt. of Forest Ecology and Genetics, Forest Genomics and Ecophysiology group. Crta. La Coruña km 7,5. 28040 Madrid. Spain
[10] Mixed Unit of Forest Genomics and Ecophysiology, INIA/UPM, Madrid, Spain
E-mail: cervera@inia.es

Key words: Epigenetics, forest trees, adaptive response, environmental stress, phenotypic plasticity, epigenetic memory of stressful conditions.

Published in *Ecology and Evolution* (2013) 3: 399-415, with doi:10.1002/ece3.461

Epigenetic variation is likely to contribute to the phenotypic plasticity and adaptative capacity of plant species, and may be especially important for long-lived organisms with complex life cycles, including forest trees. Diverse environmental stresses and hybridisation / polyploidisation events can create reversible heritable epigenetic marks that can be transmitted to subsequent generations as a form of molecular "memory". Epigenetic changes might also contribute to the ability of plants to colonise or persist in variable environments. In this review we provide an overview of recent data on epigenetic mechanisms involved in developmental processes and responses to environmental cues in plant, with a focus on forest tree species. We consider the possible role of forest tree epigenetics as a new source of adaptive traits in plant breeding, biotechnology, and ecosystem conservation under rapid climate change.

Introduction

Epigenetics refers to the study of meiotically or mitotically heritable changes in gene function that do not result from changes in DNA sequence (Bonasio *et al.* 2010). At the molecular level epigenetic phenomena are mediated by reversible marks such as DNA methylation and histone modifications, and by small RNAs that can alter regulatory states of genes or genomic regions. DNA methylation in plants occurs at cytosines in all sequence contexts (CG, CHG, and CHH where H = A, T, or C), and well-studied post-translational modifications of histone proteins at specific amino acid residues include methylation (Krauss, 2008), acetylation, phosphorylation (Demidov *et al.* 2009), and ubiquitination (Kouzarides, 2007). Genome-wide epigenetic patterns, referred to as "epigenomes", are not static; rather, they can undergo precise changes. Epigenome modifications are involved in biological processes including genetic imprinting, transposon silencing, regulation of gene expression, and heterochromatin organization.

The influence of environmental factors on epigenetic marks, and on the resultant changes in gene expression and phenotype, has recently attracted considerable attention (Boyko & Kovalchuk, 2008; Groszmann *et al.* 2011a; Feil & Fraga, 2011; Mirouze & Paszkowski, 2011; Boyko & Kovalchuk, 2008). Knowledge of the regulatory mechanisms involved in adaptive epigenetic responses may help to guide management of genetic resources and plant breeding, especially in long-lived forest tree species

where changes in allele frequency are expected to occur very slowly. This review provides a brief overview of recent data on epigenetic mechanisms involved in developmental processes and responses to environmental cues in forest species, as well as the implications of forest tree epigenetics to adaptation as a possible new source of beneficial traits for plant breeding and conservation in ecosystems responding to climate change.

Factors driving epigenetic regulation in plants

Epigenetic regulation in plant development

Epigenetic regulation plays important roles in multiple aspects of plant development. Two distinct roles of this regulation can be distinguished, depending on whether they concern developmentally regulated genes or transposable elements (TEs). In developmentally regulated genes, epigenetic marks allow switches in gene expression in response to developmental transitions and/or environmental cues. After sexual reproduction, uniparental expression of parental alleles, imprinting, is associated with discrete differentially methylated regions (DMRs) that act in a genome context-independent manner (Gutierrez-Marcos *et al.* 2006; Jullien *et al.* 2006; Haun *et al.* 2007).

A well-characterised example of epigenetic control during postembryonic development is vernalisation, the phenomenon of cold temperature-induced competence to flower (Chouard, 1960; Schmitz & Amasino, 2007). In *Arabidopsis thaliana*, regulation of *FLOWERING LOCUS C (FLC)* gene expression during vernalisation illustrates how environmental cues are perceived and translated into epigenetic marks that affect plant development (Bastow, 2004; Heo, 2011; Kim & Sung, 2012).

The epigenetic marks of many loci involved in plant development are normally erased or reset at each generation following meiosis, thus preventing the establishment of new "epialleles" (alleles whose expression is conditioned by their epigenetic status). On the other hand, stable, *i.e.* heritable, epialleles can occur naturally and might confer specific phenotypes. Examples in plants include floral symmetry in *Linaria* that is influenced by DNA methylation levels at the *CYCLOIDEA* locus (Cubas *et al.* 1999) and absence of ripening in tomato, that is associated with hypermethylation at the *Colorless non-ripening* locus (Manning *et al.* 2006). Stable epialleles are potential targets for selection in evolutionary processes, or in applied plant breeding. More examples will contribute to a better understanding of their origin, their stability, and the role they might play in selection.

In contrast to the transient nature of many developmental epigenetic marks, those affecting TEs are more stable (Slotkin & Martienssen, 2007; Bourc'his & Voinnet, 2010; Lisch & Slotkin, 2011) and the mobility of TEs is observed when these marks are alleviated in mutants affected in the epigenetic machinery (Mirouze *et al.* 2009; Tsukahara *et al.* 2009; Ito *et al.* 2011). However, during development, transcription of activated TEs in hypomethylated gamete companion cells are thought to produce small RNAs that migrate into the germ cell and direct the silencing machinery to TEs. Hence, at each new generation, the "immune system" against transposons is perpetuated but also readjusted to prevent potential genome invasion of new mobile elements (Lisch & Slotkin, 2011). Given the abundance of TEs in tree genomes, they should be considered as potential sources of epigenetic variation potentially affecting regulation of nearby genes.

The importance of developmentally-related epigenetic modification has been underscored recently by its potential involvement in hybrid vigor. Hybrid vigor, also known as heterosis, describes the superior performance of hybrid progeny over their parents in traits like biomass and seed production or stress resistance. Various models explaining heterotic effects at single loci have been proposed, including dominance, overdominance, and pseudo-overdominance, while interactions between genes (epistasis) have been considered as well (Birchler *et al.* 2010). The molecular mechanisms causing non-additive gene expression in hybrids have been the focus of studies in rice and *A. thaliana*, and epigenetic regulation has recently been associated with heterosis (Ha *et al.* 2009; He, 2010; Groszmann *et al.* 2011a; Groszmann, 2011b). In hybrids, a number of short interfering RNAs (siRNAs) were found to accumulate to non-additive levels which in turn can lead to changes in DNA methylation and gene expression, thus contributing to hybrid vigor (Groszmann *et al.* 2011a, Groszmann *et al.* 2011b). Given the preponderance of hybrids in many plant taxa, including prominent tree genera like *Populus*, the putative involvement of epigenetics in heterosis is of great interest.

Epigenetic regulation in plant environmental responses

Various environmental signals and stresses can induce persistent changes in epigenetic modifications, thereby creating a flexible "memory" system for short or prolonged periods of time (Kvaalen & Johnsen, 2008; Chinnusamy & Zhu, 2009; Jablonka & Raz, 2009; Whittle *et al.* 2009). Environmental conditions have an impact on a number of different epigenetic marks and mechanisms, including DNA methylation and histone modifications, or on frequencies of homologous recombination and genomic rearrangements (Bond & Finnegan, 2007; Chinnusamy & Zhu, 2009; Feil & Fraga, 2011; Hauser *et al.* 2011; Mirouze & Paszkowski, 2011). For example, changes in genome-wide DNA methylation patterns in response to biotic and abiotic stress treatments (pathogen, herbivore, high salt, low nutrients) occur in asexually reproduced dandelions (*Taraxacum officinale*). Notably, altered DNA methylation patterns were transmitted to the non-stressed

progeny in this species and the potential role of stress-induced epigenetic inheritance in evolution has been discussed (Verhoeven *et al.* 2010). The involvement of a histone variant (H2A.Z) was found to mediate short-term adaptation to temperature change in *A. thaliana* (Kumar & Wigge, 2010), and cold stress-induced hypomethylation and transposition of a TE (Tam-3) has been observed in *Antirrhinum* (Hashida *et al.* 2006).

Epigenetic recombinant inbred lines (epiRILs) have emphasised the relationship between response to environmental conditions and epigenetic phenomena. In *A. thaliana*, epiRILs have nearly identical genomes but display diverse mosaic epigenomes with variant DNA methylation patterns (Johannes *et al.* 2009; Reinders *et al.* 2009). The range of pathogen sensitivity/resistance within one isogenic epiRIL population exhibited 60% of the range of pathogen responses observed in natural, genetically varying *A. thaliana* accessions (Reinders *et al.* 2009). In the context of environmental challenges, such epigenetic modifications may be thought of as relatively "plastic" yet heritable marks that allow for rapid responses and adaptations and, at the same time, might avoid excessive genetic diversification (Boyko & Kovalchuk, 2008; Lira-Medeiros *et al.* 2010).

Epigenetic control in forest tree species

Relationship between epigenetic and phenotypic plasticity

Forest trees are long-lived organisms with complex life cycles, which must contend with a variable environment over their long lifetimes (Rohde & Junttila, 2008). The long generation times impose limits on natural selection under rapidly changing climate conditions (Rehfeldt *et al.* 1999; Rehfeldt *et al.* 2002). Consequently, trees must be highly adaptable, displaying a wide range of phenotypes as a function of their environments, known as phenotypic plasticity (Nicotra *et al.* 2010). Phenotypic plasticity is likely to be of great importance for both individual trees and forest populations over near- and long-term timescales. Despite this, knowledge of the extent and underlying mechanisms of phenotypic plasticity in response to a variety of stress responses and developmental traits in trees is rudimentary (Rohde, 2009; Lira-Medeiros *et al.* 2010; Neale & Kremer, 2011).

In addition to the genetic component, epigenetic variation has been suggested to contribute to the phenotypic plasticity and adaptive potential of individuals and populations (Bossdorf *et al.* 2008; Jablonka & Raz, 2009; Herrera & Bazaga, 2010; Lira-Medeiros *et al.* 2010; Richards *et al.* 2012). Insight into epigenetic variation, and its relationship to phenotypic plasticity, will contribute to the understanding of adaptive plant responses, and might help to evaluate the risk of long-lived species to both short-term and long-term fluctuations in the environment. Moreover, understanding the interplay between epigenetics and adaptation should enhance the understanding of evolutionary trajectories, as natural selection also directly targets the proportion of phenotypic variation that is shaped by epigenetic phenomena (Bossdorf *et al.* 2008; Herrera & Bazaga, 2010).

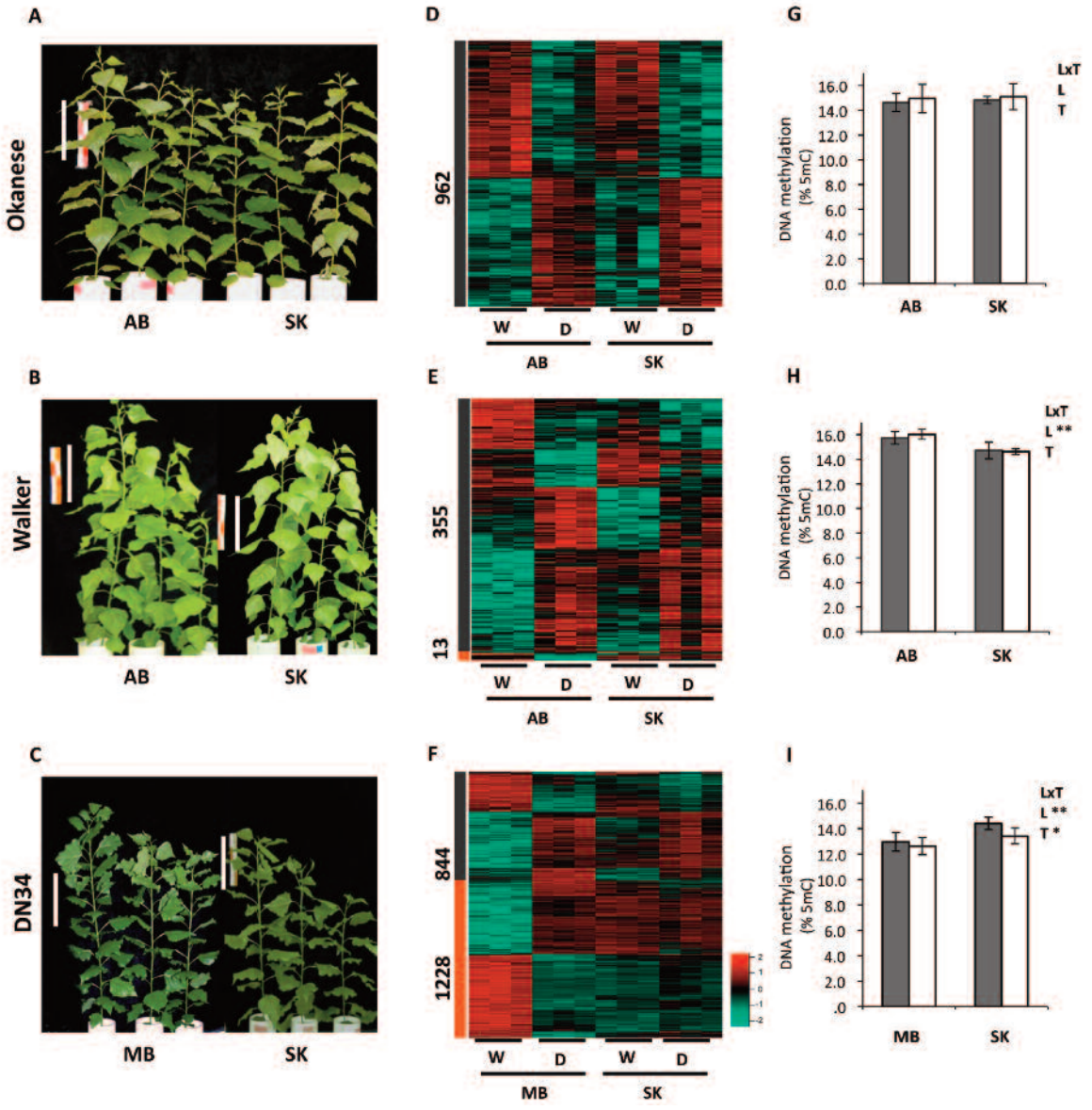
Epigenetic and phenotypic variation in natural populations, ecotypes and species

Despite the substantial impact that epigenetics might have in determining environmental compatibility, relatively few studies have investigated the extent of natural epigenetic variation and its relationship to phenotypic variation and adaptation potential (Cervera *et al.* 2002; Bossdorf *et al.* 2008; Jablonka & Raz, 2009; Marfil *et al.* 2009; Herrera & Bazaga, 2010; Lira-Medeiros *et al.* 2010; Paun *et al.* 2010). Among the different epigenetic mechanisms that are potentially involved in transgenerational inheritance and natural epigenetic variation, DNA methylation represents the most-studied modification (Akimoto *et al.* 2007; Jablonka & Raz, 2009; Herrera & Bazaga, 2010; Lira-Medeiros *et al.* 2010; Paun *et al.* 2010; Verhoeven *et al.* 2010). Two of the few published studies in higher plants that considered the interplay between genetic, epigenetic, as well as and phenotypic variation and environmental factors, focused on a perennial violet species (*Viola cazorlensis*) and orchids of the *Dactylorhiza majalis* complex (Herrera, 1990; Herrera & Bazaga, 2010; Paun *et al.* 2010). The studies detected coordinated genetic-epigenetic adaptive differentiation, indicating the involvement of epigenetic processes in adaptation and evolution by influencing primary phenotypic diversity.

In tree species, natural variation in epigenetic marks and the relation to phenotypic traits is still an under-explored area. Insight into the role of epigenetics in determining tree phenotype should identify key elements in the control of growth traits and contribute to the understanding of evolutionary capacity of tree species (Grattapaglia *et al.* 2009; Thumma *et al.* 2009; Lira-Medeiros *et al.* 2010). In keeping with this, evidence for the correlation between tree form and epigenetics is emerging. Trees of the white mangrove (*Laguncularia racemosa*) can occur naturally in contrasting habitats and can exhibit striking differences in morphological traits (Lira-Medeiros *et al.* 2010). Tree-like appearance was documented in a riverside habitat with abundant fresh water and nutrient supply whereas in a nearby salt marsh habitat mangrove plants were characterised by abnormal growth and shrub-like morphology. Notably, despite morphological dissimilarities, analysis of DNA nucleotide sequences and methylation patterns detected greater epigenetic than genetic variation within and between populations in contrasting environments, which indicates that epigenetic variation in natural populations plays an important role in long-term adaptation to different environments (Lira-Medeiros *et al.* 2010).

FIGURE 1

Clone history shapes drought responses in poplar hybrids. Transcriptome-level responses to water withholding are influenced by geographic origin for two of the three genotypes, and are paralleled by differences in total (genome-wide) DNA methylation. Ramets of hybrid poplar genotypes (A) Okanese, (B) Walker, and (C) DN34 with distinct histories were obtained from two different locations for each of the genotypes. The response to water deficit was assessed under common, controlled environmental conditions. Okanese (A, D, G); Walker (B, E, H); DN34 (C, F, I). Tree appearance (A-C). Transcriptome-level responses (D-F). Heat maps represent relative abundance of drought responsive transcripts at pre-dawn for Okanese (D), Walker (E), and Okanese (F) obtained from two locations each. The numbers indicated to the side of the heat map correspond to transcripts with significant treatment main effect only (gray) and with significant treatment: location interactions (orange bar) (BH adjusted, P<0.05). W, well watered samples; D, water-deficient samples. Global DNA methylation levels as percentage of 5mC under well-watered (shaded bars) and water-limited conditions (white bars) for the genotypes Okanese (G), Walker (H), and DN34 (I). L, location effect; T, treatment effect; and LxT, location: treatment interaction term (*P<0.05, **P<0.001, n=6, SD bars). Locations are abbreviated as follows AB, Alberta; SK, Saskatchewan; MB, Manitoba. Figure is adapted from Raj *et al.* 2011.



The lasting impact of previous environmental history on a tree's capacity to respond to a current environmental stimulus has recently been explored in *Populus* (Raj *et al.* 2011). Poplar trees are frequently propagated vegetatively through stem cuttings of branches containing dormant buds, generating genetically identical individuals or ramets of the same genotype. Clonally propagated poplar trees can be planted in different geographic locations, thus giving rise to populations of genetically identical ramets that are characterised by their own local environment and history. To study the lasting effect of clone history on current plant performance, cuttings of the same genotype were obtained from different geographic locations and grown under common environmental conditions, after which the transcriptome response to an important environmental stress, drought, was studied. Notably, differences in transcript abundance patterns in response to drought that were based on differences in geographic origin of clonally propagated trees were detected in two of the three investigated genotypes. These transcriptome-level patterns were paralleled by differences in genome-wide DNA methylation. Genotypes with the longest time since establishment and last common propagation showed the most pronounced location-specific patterns in transcriptome response and DNA methylation indicating a possible epigenomic basis for clone history-dependent transcriptome divergence (Fig.1). These findings underline the importance of epigenetic mechanisms related to the adaptation of long-lived species like poplar trees to the local environment (Raj *et al.* 2011).

The direct response of six hybrid poplar genotypes to water deficit revealed a relationship between epigenetic marks and the genotypic variability of phenotypic plasticity (Gourcilleau *et al.* 2010). Genotypic variation for both DNA methylation and traits related to biomass productivity was observed in hybrids (*Populus deltoids x P. nigra*), and a positive correlation was established among these variables in well-watered conditions (Fig.2). While poplar genotypes showed reduced growth in water deficit conditions, a significant genotype effect was observed for DNA methylation variations. This suggests that DNA methylation could participate in the fine-tuning of gene expression in poplar during water stress (Plomion *et al.* 2006; Bogeat-Triboulot *et al.* 2007; Bonhomme *et al.* 2009; Wilkins *et al.* 2009; Gourcilleau *et al.* 2010; Hamanishi & Campbell, 2011).

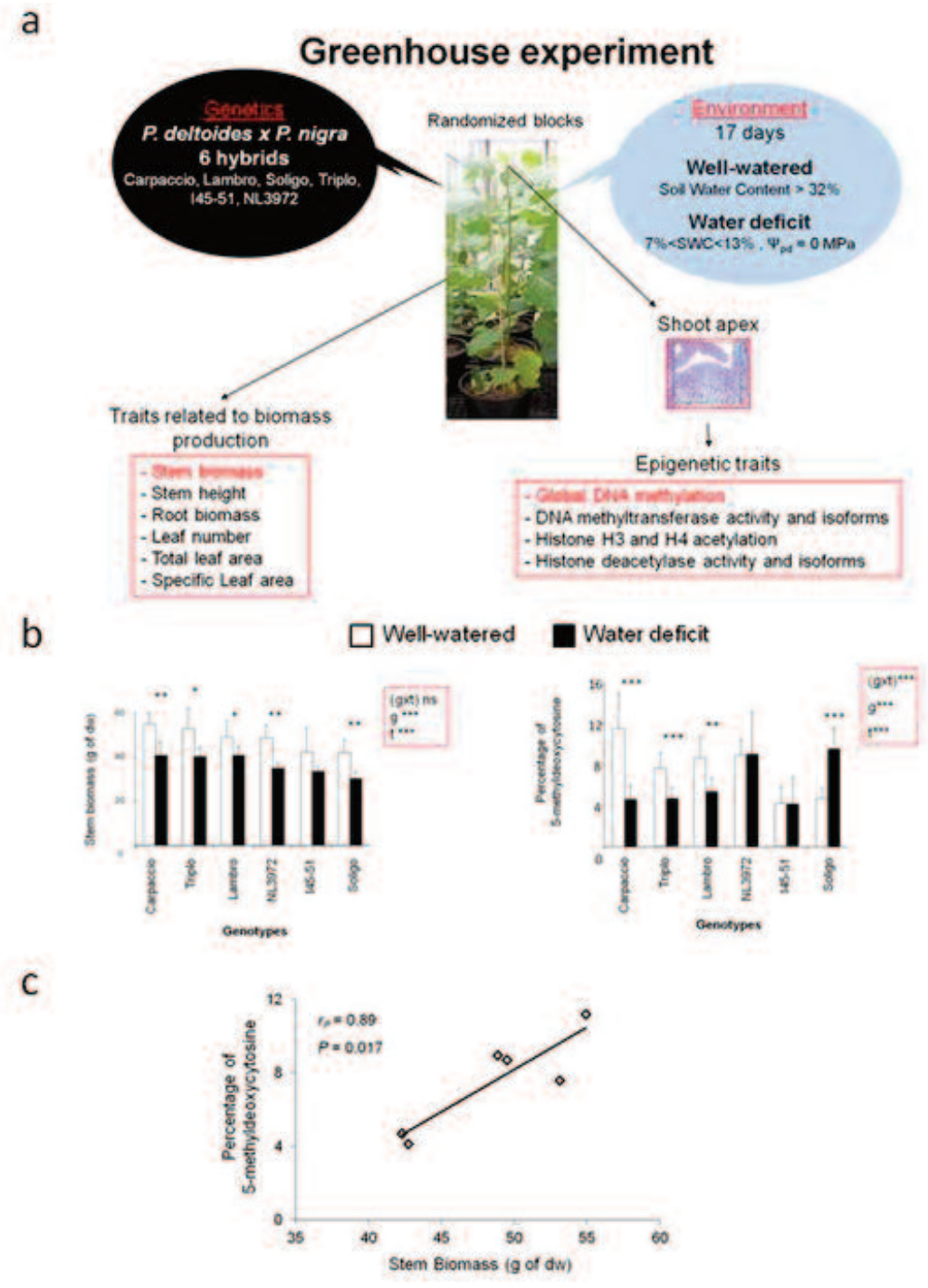
The potential link between natural epigenetic variation and phenotypic variability observed in trees is further supported by studies in ecotypes and individual populations of specific herbaceous plant species (Cervera *et al.* 2002; Marfil *et al.* 2009). Highly conserved DNA methylation patterns were detected within an *A. thaliana* ecotype (Ler) while clear DNA methylation differences existed between ecotypes that did not correlate with nucleotide sequence variation, but with their flowering time (Cervera *et al.* 2002). Furthermore, variation in the floral phenotype of individuals from a single natural population of a wild hybrid potato (*Solanum ruiz-lealii*) were found to correlate with distinct DNA methylation patterns but not with DNA sequence variation (Marfil *et al.* 2009).

Most studies assessing epigenetic variation in natural populations, ecotypes or species focused on the extent of epigenetic variability and paid less attention to the functional consequences. Indication for a functional link between a specific epigenetic mark at a specific position in the genome and variation in a quantitative trait was discovered by analysing polymorphisms in an association population and a full-sib family of eucalypt (*Eucalyptus nitens*; Thumma *et al.* 2009). Making use of the low linkage disequilibrium in populations of forest trees, variation in cellulose content was linked to polymorphisms within a gene potentially involved in cellulose synthesis and deposition (functional polymorphisms). The COBRA-like gene *En*COBL4A was strongly associated with a QTL region for cellulose content and fine mapping revealed a significant association with a SNP in exon 5. Notably, allelic expression imbalance was linked to allele-specific cytosine methylation upstream of this SNP in a full-sibling family. A heritable epigenetic polymorphism is thus likely to influence phenotypic variation in cellulose content; however, further functional analyses are required (Thumma *et al.* 2009). The findings suggest that epigenetic variations might contribute to quantitative trait variation (Thumma *et al.* 2009), and it has been suggested that this phenomenon might be common (Johannes *et al.* 2008; Reinders *et al.* 2009; Thumma *et al.* 2009; Long *et al.* 2011).

To date, some prominent, shared observations have emerged from the few studies of natural epigenetic variation and phenotypic plasticity. These studies established that a) epigenetic variation occurs in natural populations, ecotypes and species, b) this variation can correlate with naturally occurring phenotypic variation, and c) there is a potential role for epigenetic variation in adaptation and potentially in evolution. Despite the commonalities that have emerged from these studies, many questions remain unresolved. For example, is epigenetic variation in natural populations a wide-spread phenomenon? Moreover, the key molecular mechanisms involved and how they are regulated remain to be determined. Finally, it is unclear to what extent epialleles arise and how stable they are when considered in an evolutionary context (Bossdorf *et al.* 2008; Herrera & Bazaga, 2010; Lira-Medeiros *et al.* 2010; Paun *et al.* 2010). Answers to such questions might also contribute to a better understanding of the adaptive capability of long-lived forest trees that might help to assess their susceptibility to rapidly changing environments (Grattapaglia *et al.* 2009).

FIGURE 2

Relation between epigenetic marks and the genotypic variability of phenotypic plasticity under limited water availability or not in six poplar hybrids. a. Experimental design; **b.** Stem biomass and DNA methylation levels in the shoot apex (center of morphogenesis). For each graph, g indicates the genotype effect, t the treatment effect and (gxt) genotype by treatment effect. Means are accompanied by their standard errors SE (n = 6). Significant differences between well-watered and water deficit conditions are indicated by asterisk: * P ≤ 0.05, ** P ≤ 0.01 and *** P ≤ 0.001; c. Linear correlation (Pearson, r) between stem biomass and DNA methylation levels. Adapted from Gourcilleau *et al.* 2010.



Epigenetic plasticity in growth and development

During their relatively long lifespans, trees must make developmental adjustments while retaining flexibility to match and synchronise growth and development with prevailing environmental conditions. Epigenetic mechanisms are proposed to contribute to such flexible adjustments by generating transmittable and reversible marks that constitute temporary "memory" systems (Boyko & Kovalchuk, 2008; Kvaalen & Johnsen, 2008; Yakovlev *et al.* 2010; Jaskiewicz *et al.* 2011). Tissue-, organ-, and species-specific differences in DNA methylation levels are well known (Fraga *et al.* 2002a; Fraga *et al.* 2002b; Valledor *et al.* 2007; Monteuuis *et al.* 2009; Santamaria *et al.* 2009; Rodriguez Lopez *et al.* 2010; Valledor *et al.* 2010; Vining *et al.* 2012; Lafon-Placette *et al.* 2012). Changes in epigenetic marks were found to accompany morphological and physiological changes in trees in a wide variety of processes, including ageing, phase change, organ maturation, and bud set or burst (Fraga *et al.* 2002a; Fraga *et al.* 2002b; Santamaria *et al.* 2009; Valledor *et al.* 2010).

Bud dormancy is a vital adaptation to seasonal changes, and release and induction of bud dormancy are complex processes that largely determine length of the growth season, and thereby affect annual tree productivity. Regulation of bud burst integrates endogenous and exogenous signals such as hormone levels, day length, light quality and temperature (Santamaria *et al.* 2009) and involves substantial changes in gene expression and epigenetic modifications (Ruttink *et al.* 2007; Rohde, 2009; Santamaria *et al.* 2009). In apical buds of a chestnut *(Castanea sativa)*, a decrease in global DNA methylation level and concomitant increase in acetylation of histone 4 was observed during bud burst when conditions were favorable for active growth. The opposite pattern (i.e., DNA hypermethylation and lower histone acetylation levels), indicative of more repressive chromatin states, was detected during bud set when environmental conditions were less favorable for growth (Santamaria *et al.* 2009). The observed coordinated changes in DNA methylation and histone modifications are predicted to alter the control of gene expression to shape the processes of bud burst and bud set (Santamaria *et al.* 2009).

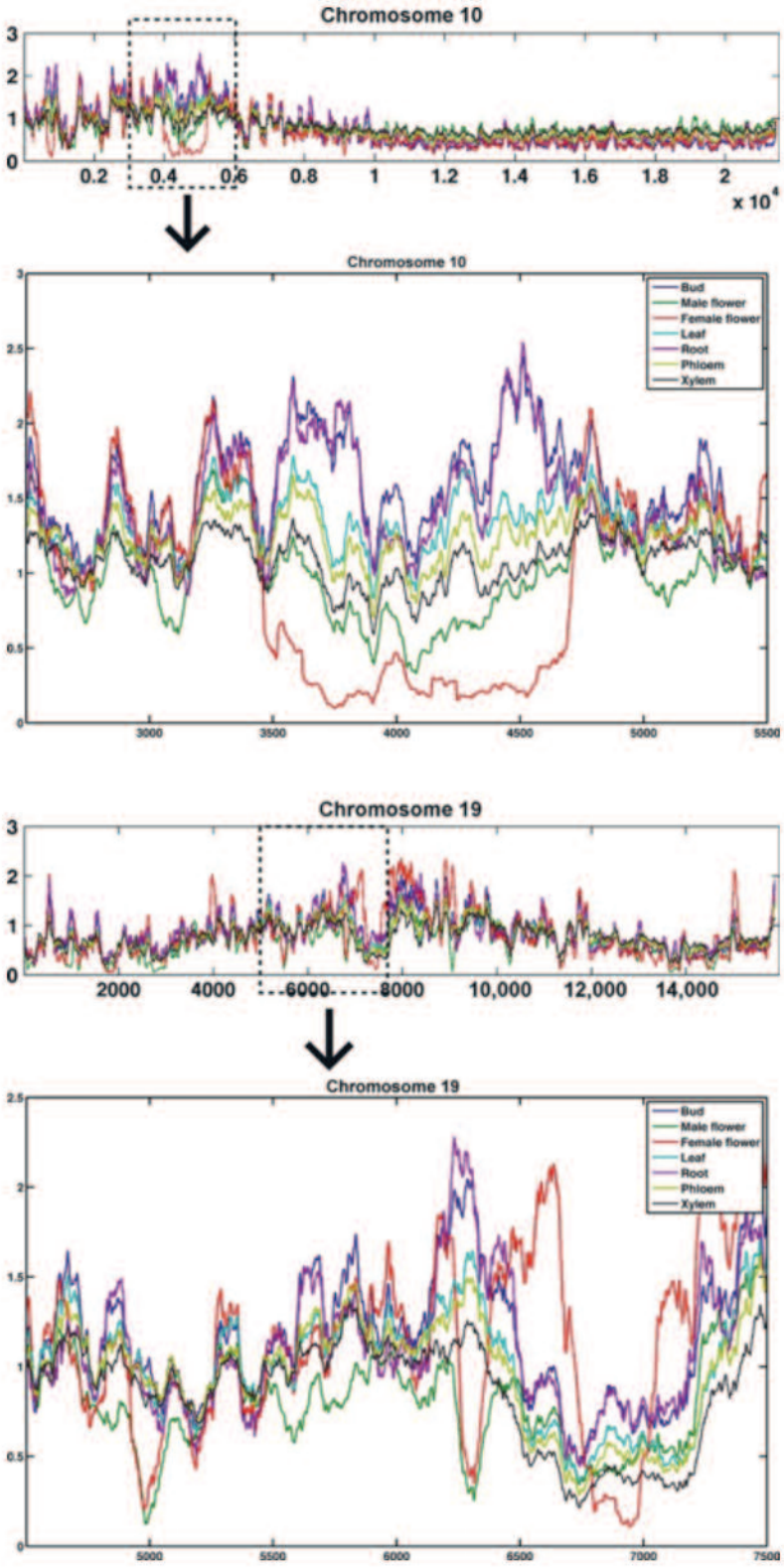
Ageing and maturation are characterised by altered patterns of cell differentiation and organ formation processes, and the potential role of DNA methylation in maturation has been studied in some tree species (Fraga *et al.* 2002a; Fraga *et al.* 2002b; Valledor *et al.* 2007; Monteuuis *et al.* 2009). For example, studies in radiata pine *(Pinus radiata)* support the involvement of DNA methylation in this process. Changes in global DNA methylation levels of up to 25% during maturation have been reported in this species (Fraga *et al.* 2002a; Fraga *et al.* 2002b). In juvenile plants without flowering capability, young needle tissue was characterised by a markedly lower extent of DNA methylation than corresponding tissues in adult trees with reproductive ability. Regarding histone modifications, decreased levels of euchromatin-associated marks, such as histone 4 acetylation and specific histone methylation (trimethylation of histone 3 on lysine 4 or H3K4me3) have been measured in mature needles when compared with juvenile ones (Valledor *et al.* 2010). Moreover, the observed increase in DNA methylation levels from juvenile to mature plants in meristematic tissue could be directly linked to phase change. Conversely, an increase in the degree of tree reinvigoration by serial grafting, measured by the recovery of morphogenetic competence, was accompanied by a decrease in global level of DNA methylation in meristematic tissue, thus pointing towards plasticity of DNA methylation marks during ageing and maturation. The degree of DNA methylation, as well as additional biochemical characteristics, were proposed to serve as suitable markers for ageing and reinvigoration in pine (Fraga *et al.* 2002a; Fraga *et al.* 2002b). However, differences between species and experimental systems might exist. In another conifer, *Larix laricina*, age-related changes in foliar traits were observed, while differences in DNA methylation levels between juvenile and mature scions could not be detected in DNA from whole needles (Greenwood *et al.* 1989).

In angiosperms, heteroblastic tree species like *Acacia mangium* with distinct leaf morphologies of juvenile and mature stages provide excellent systems to study ageing. Small but significant differences between microshoots with juvenile (pinnate) and mature (phyllode) morphology were observed in this acacia species when analysing global DNA methylation levels in physiologically active apical buds of *in vitro* grown plant material. Here, the degree of DNA methylation was higher in juvenile than in mature tissue, and might be influenced by *in vitro* culture conditions (see Epigenetic and phenotypic plasticity in artificial system) in addition to maturation-related processes (Baurens *et al.* 2004; Monteuuis *et al.* 2009). Taken together, the aforementioned studies establish a clear relationship between DNA methylation levels and maturation for some tissue types and species in woody plants (Fraga *et al.* 2002a; Fraga *et al.* 2002b; Baurens *et al.* 2004; Valledor *et al.* 2007; Monteuuis *et al.* 2009). Observed differences might be attributable to differences in taxonomy, tissue type (meristematic vs. differentiated) or experimental system (*in vitro*, field conditions) and might also reflect underlying mechanistic differences in the relationship between ageing and epigenetic marks (Fraga *et al.* 2002b; Monteuuis *et al.* 2009). Furthermore, the data indicate that DNA methylation patterns are not static and can exhibit remarkable dynamics and plasticity during development and seasonal changes (Fraga *et al.* 2002b; Valledor *et al.* 2007; Monteuuis *et al.* 2009).

Evidence for the remarkable dynamics and plasticity of epigenetic modification in tree species is growing.

FIGURE 3

DNA methylation profiles vary widely among chromosomes and among tissues at selected loci in *Populus trichocarpa*. Relative DNA methylation was determined using methylated DNA immunoprecipitation followed by Illumina sequencing (MeDIP-seq). The ratio of MeDIP-seq read counts in immunoprecipitated (IP) samples versus non-IP control is plotted in 1 kb windows for chromosomes 10 and 19, and areas of tissue-differential methylation are expanded below each chromosome. Figures from Vining *et al.* (2012).



Genome-level comparative analysis of cytosine methylation among differentiated poplar tissues revealed highly heterogeneous DNA methylation profiles among chromosomes, and a number of cases of tissue-specific methylation (Fig.3; Vining *et al.* 2012), many of them associated with gene bodies or promoters. Although a broadly similar chromosome methylation and gene expression profile was observed in poplar when compared to *A. thaliana* and other plant species, significant differences were also detected. For example, only in poplar was gene body (*i.e.*, the entire gene from the transcription start site to the end of the transcript) methylation associated with greater repression of gene expression than was promoter methylation. In addition, Vining *et al.* (2012) observed a distinctive pattern of transposon and gene body methylation for male catkins compared to other tissues, including female catkins. Recently, analysis of the methylome of open chromatin in poplar meristematic cells found that 74% of poplar gene models had gene body methylation, and its intensity, as well as cytosine context, varied depending on gene size, redundancy in the genome (presence of paralogs), and extent of tissue-specific gene expression (Lafon-Placette *et al.* 2012).

Plasticity in tree epigenetic modification has also been observed in conifer species, specifically as it relates to phenology. Phenology responses of seedlings that were produced in warm or cold years vary within the same stands (Kohmann & Johnsen, 1994). In Norway spruce, a temperature-dependent epigenetic "memory" from the time of embryo development, which thereafter influences the timing of bud phenology and gene expression, has been discovered (Skrøppa & Johnsen, 2000; Johnsen *et al.* 2005; Yakovlev *et al.* 2010). Colder-than-normal conditions during embryogenesis and seed development advance the timing, whereas temperatures above normal delay the onset of these adaptive processes, and the altered performance is long lasting in the progeny. This phenomenon was initially discovered when ecotypes from northern Norway were transferred to a southern seed orchard where they produced progenies with a phenology similar to that of southern ecotypes (Johnsen *et al.* 1996; Skrøppa & Johnsen, 2000). Notably, differences in day length and temperature applied during pollen formation did not affect the progeny performance. Differences in the female flowering environment did affect progeny performance. The temperature during zygotic embryogenesis and seed maturation shifted the developmental program of the seeds, resulting in significant phenotypic changes, with the effect lasting as long as over 20 years (Skrøppa & Johnsen, 2000; Skrøppa *et al.* 2010; Yakovlev *et al.* 2010). The traits that are affected include the timing of dehardening and bud burst in the spring; leader shoot growth cessation in the summer and bud set and cold acclimation in the autumn. All processes are thus advanced or delayed as influenced by the temperature during reproduction in progeny with identical genetic background. Similar effects have been observed in progeny from white spruce (*Picea glauca* × *Picea engelmannii*) crosses, Scots pine, Larix spp. and longleaf pine (Dormling & Johnsen, 1992; Greenwood & Hutchison, 1996; Stoehr *et al.* 1998; Webber *et al.* 2005) but there is lack of information regarding this phenomenon in angiosperm trees (Rohde & Junttila, 2008). In birch (*Betula pendula*), a small-scale study within a population revealed a close genetic relationship between trees that had established in a year of similar temperature (Kelly *et al.* 2003).

The importance of plastic epigenetic modification on phenology in conifer species extends beyond the individual to encompass the ecosystem. Epigenetic effects taking place during zygote development may create phenotypic diversity at the local community level, if temperature varies considerably amongst successive generations. This is particularly important as phenology traits are strongly genetically differentiated.

The molecular mechanism behind this striking epigenetic "memory" phenomenon is not yet clear, but transcriptional changes have been implicated (Johnsen *et al.* 2005; Yakovlev *et al.* 2010; Yakovlev *et al.* 2011). In progeny that differ epigenetically, transcriptional analysis revealed that seedlings from full-sib families produced at different embryogenesis temperatures under long and short day conditions differed. Suppressive subtracted cDNA libraries revealed considerable differences in their transcriptomes. MicroRNA pathway genes *DICER-LIKE1 (PaDCL1), DICER-LIKE 2 (PaDCL2)* and *SUPPRESSOR OF GENE SILENCING 3-LIKE (PaSGS3)*, as well as transposon-related genes, had altered transcript abundance in epigenetically-different progeny with phenotypic differences in bud burst and bud set (Yakovlev *et al.* 2011). Norway spruce contains a set of conserved miRNAs as well as a large proportion of novel non-conserved miRNAs involved in temperature-dependent epigenetic "memory". Most of the miRNAs were targeted to previously unknown genes, or genes with no known function. The expression of seven conserved and nine novel miRNAs showed significant differences in transcript levels in progenies showing distinct epigenetic difference in bud set, but not in the progeny from a non-responding family without differences in bud set, making them excellent candidate miRNAs. The altered transcript abundance of specific miRNAs suggests their putative participation in epigenetic regulation (Yakovlev *et al.* 2010). This epigenetic phenomenon is not only generated in controlled Norway spruce crosses, but such epitypes can also be produced by somatic embryogenesis (Kvaalen & Johnsen, 2008). Genetically identical plants generated at different temperatures by zygotic embryogenesis expressed a difference in timing of terminal bud formation that was equivalent to a 4-6° latitudinal ecotypic difference.

The "memory" effects acting on phenological traits lasted for more than 20 years after germination and affected long-term growth under field conditions (Skrøppa *et al.* 2007). Notably, there was absence of any genetic segregation distortion in the progeny, strongly supporting that this "memory", affecting the climatic adaptation in this species, is indeed an epigenetic phenomenon (Besnard *et al.* 2008). Thus distinct epitypes can be produced from the same genotype in Norway spruce, a process not well documented in other tree species so far. In view of rapid climate change, strategies to increase diversity for selection might be of prime importance for survival of species within their current geographic distribution, and therefore, this epigenetic "memory" mechanism is likely of evolutionary significance and has obvious practical implications.

Epigenetic and phenotypic plasticity in artificial systems

While epigenetic phenomena are clearly important for trees in a natural context, they also could be of great consequence during specific tree production processes integrated into the wood products chain. Long generation times and the out-crossing habit of a number of forest trees can make it difficult to rapidly propagate material and maintain valuable genotypes under natural conditions. Tissue culture can provide alternative means to keep desirable genotypes by vegetative propagation and to quickly produce commercial quantities of regenerants; therefore, micropropagation is widely used in forestry. It has been observed, however, that tissue culture can introduce variation in regenerated plants. This somaclonal variation, can result in subtle to drastic phenotypic variation and has been found to be attributable to genetic or epigenetic variations (*e.g.* reviewed in Kaeppler *et al.* 2000; Miguel & Marum, 2011). Somaclonal variation (heritable across mitotic and meiotic cell divisions) has been considered both beneficial and disadvantageous (Jaligot *et al.* 2000; Kaeppler *et al.* 2000; Schellenbaum *et al.* 2008), and a number of studies have focused on elucidating underlying mechanisms (Kaeppler *et al.* 2000; Rival *et al.* 2008; Schellenbaum *et al.* 2008; Rodriguez Lopez *et al.* 2010).

A well-studied example for somaclonal variants and their relation to epigenetic marks in a tree species is the *mantled* phenotype in somatic-embryo-derived oil palm (*Elaeis guineensis*). This phenotypic variant, found in about five percent of regenerants, is characterised by abnormal inflorescence development and has been associated with global DNA hypomethylation, but not to changes in genomic structure or nucleotide sequence (Jaligot *et al.* 2000; Rival *et al.* 2008). The exact mechanisms involved in generating somaclonal variants like the mantled phenotype remain largely unresolved. Ongoing studies of this phenomenon might help to better understand mechanisms of epigenetic responses to tissue-culture-induced stresses (Kaeppler *et al.* 2000; Rival *et al.* 2008).

It has also been observed that the ability to generate mature somatic embryos from cultured tissue can decrease as a culture ages and that somaclonal variation can increase with culture age (Phillips *et al.* 1994; Valledor *et al.* 2007; Krizova *et al.* 2009). In addition to other mechanisms, changes in DNA methylation were considered to contribute to the reduction of embryonic potential or organogenic potential in tissue culture and grafting procedures (Fraga *et al.* 2002b; Valledor *et al.* 2007). A detailed analysis of genetic and epigenetic variation in relation to callus age reports interesting plasticity in cocoa plants (*Theobroma cacao*) regenerated by somatic embryogenesis. Genetic variation was investigated using single sequence repeat (SSR) markers, and epigenetic variability was assessed by methylation-sensitive amplified polymorphism (MSAP), a method to detect genome-wide but anonymous DNA methylation patterns. Contrary to predictions, after an initial increase, a decrease in both genetic and epigenetic divergence between leaves of regenerants and the ortet plant was observed after the culture had reached an age of about 10 weeks (Rodriguez Lopez *et al.* 2010). One possible interpretation of the findings suggests a link between stable DNA methylation patterns and repression of *de novo* mutations during somatic embryogenesis (Rodriguez Lopez *et al.* 2010).

For many plant species, different physiological and developmental stages of diverse tissue explant types have been associated with distinct epigenetic characteristics, in particular DNA methylation (Fraga *et al.* 2002a; Fraga *et al.* 2002b; Monteuuis *et al.* 2009; Santamaría *et al.* 2009; Rodríguez López *et al.* 2010; Valledor *et al.* 2010). For example, some DNA methylation patterns and levels, characteristics of the source tissue used to start an *in vitro* culture, were retained in regenerants in acacia and cocoa (Monteuuis *et al.* 2009; Rodríguez López *et al.* 2010). This highlights the plasticity of DNA methylation marks under tissue culture conditions. Transitions from juvenile to adult phase are frequently accompanied by reduction or loss of morphogenetic ability in woody species (see Epigenetic plasticity in growth and development). Concomitant with maturation of pine needles, changes in epigenetic marks were measured when compared to immature needles. This finding could be in accordance with a less permissive and reprogrammable chromatin state and could account in part for the reduced organogenic capacity of explants from mature needles.

Generation of somaclonal genetic and epigenetic variants as well as plasticity in DNA methylation are widely documented outcomes of plant regeneration in tissue culture (Kaeppler *et al.* 2000; Marfil *et al.* 2009). Studying underlying mechanisms might be of relevance for basic research and applications in plant propagation such as the understanding of differentiation and dedifferentiation processes or the selection of appropriate *in vitro* culture conditions (Kaeppler *et al.* 2000; Marfil *et al.* 2009; Rodriguez Lopez *et al.* 2010).

Strategic and technical approaches to study epigenetic processes

Selection of appropriate systems

Different methods have been used in model plants to analyse epigenetic variation independently of genetic variation. These have included treatment with demethylating agents, analysis of natural epimutations, and study of DNA methylation-deficient mutants. Epigenetic recombinant inbred lines (epiRILs) have been developed in *A. thaliana* (Johannes *et al.* 2009; Reinders *et al.* 2009) using isogenic lines (wild types and mutant lines) differing only in the level and distribution of DNA methylation (see Epigenetic regulation in plant environmental responses). These lines represent a powerful tool to identify specific epigenomic regions that are associated with the observed phenotypic variation through epiQTL mapping approaches that are based on methylation sensitive markers. The epiQTL mapping approach requires the establishment of multiple plant generations, and may be difficult to apply to tree species that require a significant amount of time to reach sexual maturation.

To discern genetic and epigenetic effects, clonally propagated plants or systems that are characterized by reduced genetic variation, such as stone pine (*Pinus pinea*), represent ideal study subjects. To separate heritable from non-heritable epigenetic variation (resulting from developmental plasticity in response to different environments) it is necessary to study, when available, clonally propagated genotypes, the progeny of different natural populations or maternal families in a common environment, and to use the resemblance of epigenetic patterns among relatives as an indication of epigenetic inheritance (Bossdorf *et al.* 2008).

Technical approaches

A wide variety of techniques have been developed to study epigenetic patterns and modifications. Histone modifications can be analysed by chromatin immunoprecipation (ChIP) using antibodies that recognise specific histone modifications, followed by either microarray hybridisation (ChIP on chip) or by next generation sequencing (ChIP-Seq; Ku *et al.* 2011). DNA methylation at the genome level, the DNA methylome, can be investigated by methylated DNA immunoprecipitation (meDIP) or by bisulfite treatment of the DNA followed by hybridisation to a microarray, or by next generation sequencing (BS-Seq; Ku *et al.* 2011; Krueger *et al.* 2012; Cokus *et al.* 2008). Additionnally direct detection of methylated residues using DNA synthesis technologies based on variable polymerase kinetics depending on the chemical modification of the template nucleotide (*e.g.* 5-methylcytosine *vs.* cytosine) represents a novel method to directly detect DNA methylation (Flusberg *et al.* 2010).

Next generation sequencing technologies enable mapping of epigenetic modifications at single base resolution. The nature and large amount of data generated by such technologies will demand new approaches in data analysis techniques. Inference of the methylation status of bisulfite-treated DNA by BS-Seq can be challenging as the data obtained do not exactly match the reference sequence. Consequently, both DNA strands must be considered separately, and methylation at a specific site can be a percentage rather than a total presence or absence. Nevertheless, a number of tools have been developed to facilitate these analyses and are now available for application to tree epigenomes (Chen *et al.* 2010; Lim *et al.* 2010; Krueger *et al.* 2012).

Conclusion

Many questions remain about the mechanisms and roles of epigenetic processes in enabling rapid adaptation of plants to their environment, especially in forest trees. Recently, genome-wide studies of chromatin-bound proteins and epigenetic marks in *Drosophila melanogaster* and in *A. thaliana* have substantially revised our understanding of chromatin (Roudier *et al.* 2011; Van Steensel, 2011). The dogma of an uncompacted, transcriptionally-active euchromatin *versus* a compacted, silent heterochromatin is likely to be an oversimplification of the real chromatin architecture. It appears that chromatin might be composed of several types differing in their epigenetic marks as well as in their nuclear localisation and chromatin-associated proteins. These types could favor or prevent association with transcription factors, thus defining gene expression patterns. Whether these chromatin types exist in perennial species is not known, and the stability of these chromatin types in long-living organisms is to be established. Similarly, the maintenance of these types during clonal and *in vitro* culture propagation will give important clues about the effect of these biotechnologies on gene expression control.

It has been observed that the induction of alternative epigenetic states not only triggers the formation of new epialleles but also promotes the movement of DNA transposons and retroelements that are very abundant in plant genomes (Mirouze & Paszkowski, 2011). However, mechanisms counteracting accumulation of induced epialleles must also be in place, because otherwise we would be "constantly confronted with the inheritance of environmentally-induced phenotypic variation" (Richards, 2006). Additionally, in large genomes, such as those of conifer [with C estimates of DNA content ranging from 17 to 30 Gbp for pines and spruces of which more than 68% are attributed to repeated DNA (Rake *et al.* 1980; Ohri & Khoshoo, 1986)], cytosine methylation is implicated in genomic compartmentalisation, *i.e.*, non-coding highly repeated sequences get separated from low-copy

sequence and transcriptionally active regions. The differential methylation of genic and non-genic regions observed across plant taxa, may be involved in decreasing transcriptional 'noise' (Rabinowicz *et al.* 2005). In large genomes, epigenetic mechanisms might be more prominent, as a means to control the repetitive parts of the genome. This might render their entire genomes more amenable to epigenetic regulation.

From an economic and ecological point of view, it is important to integrate information on epigenetic control of environmental and developmental processes in both forest resources management and breeding. In quantitative genetic studies, estimates of genetic variance over the total phenotypic variance are typically used to assess the heritability of a trait. Akin to other genetic characters, variance in epigenetic characters will contribute to genetic variance and/or phenotypic variance, but might go undetected in some studies, or might be confounded with normal Mendelian-based quantitative inheritance (Kalisz & Purugganan, 2004). Epigenetic effects may thus inflate the true genetic variation in traits. As a consequence, the genetic clines observed for many phenology traits, even in common garden experiments, may reflect more local adaptation than DNA sequence-based genetic differences among populations.

Recent developments show that both energy efficiency and energy homeostasis, which are integral parts of yield, have an epigenetic component that can be directed and stabilised by artificial selection (i.e. selective breeding; De Block & Van Lijsebettens, 2011). These findings open new possibilities for engineering plant metabolism and improving complex traits. For example, in addition to the unintended genetic and epigenetic variation imparted by *in vitro*-manipulation, it may be considered and utilised as a means to amplify or release epigenetic variation of value to breeding programs Transgenic perturbation of epigenetic mechanisms might have similar effects; however, testing such effects using a transgenic approach with forest trees at a scale relevant to application and ecological variation are, at present, constrained by government regulations (Viswanath *et al.* 2012).

Genome perturbation, including epigenetic components, might be important for increasing the raw material for adaptive evolution under severe stress (Kalisz & Purugganan, 2004; Rapp & Wendel, 2005). Rapp & Wendel (2005) suggest that a population bottleneck, while reducing genetic diversity, might simultaneously create epigenetic novelty. In contrast to genetic alleles, epialleles might react more quickly to environmental change, be reversible, and persist for only a few generations (Kalisz & Purugganan, 2004). If a new epiallele were to cause a mild phenotype through alteration of gene expression, it might experience less strong selection than a loss-of-function sequence mutation (Kalisz & Purugganan, 2004) and thus enable rapid, yet fine-tuned, trait modifications. The significance of epialleles in wild populations will depend on their frequency and stability (Rohde & Junttila, 2008).

The analysis of the epigenetic processes in an ecological context, known as "ecological epigenetics" is set to transform our understanding of the way in which organisms function on the landscape. Forest trees offer excellent opportunities to examine some of the most compelling questions of ecological epigenetics (Bossdorf *et al.* 2008), particularly those related to the interplay between epigenetic variation and phenotypic variation in natural populations, and the role of epigenetic variation in evolutionary processes. Ecological epigenetics could readily address such questions by capitalising on the advantageous features of forest trees, including their long-lifespans, their dominance of many ecosystems, their wide geographical distribution, and their life histories, especially reproductive traits like clonal propagation. Analysis of the epigenetics of forest tree species will significantly improve our understanding of the mechanisms underlying natural phenotypic variation, and the responses of organisms to environmental change, and may thereby inform efforts to manage and breed tree species to help them cope with environmental stresses.

Acknowledgements

The preparation of this review was supported by funding from the FP7 UE projects FORESTTRAC (2440096-FP7) and ProCoGen (289841-FP7).

The authors are very grateful to Susana Ferrándiz for helping with the preparation of the manuscript.

Conflict of interest

All authors disclose any potential source of conflict of interest or relationship (financial or otherwise) that might be perceived as influencing an author's objectivity.

REFERENCES

Akimoto K, Katakami H, Kim HJ, Ogawa E, Sano CM, Wada Y, Sano H (2007) Epigenetic inheritance in rice plants. *Annals of Botany*, **100**, 205-217.

Bastow R, Mylne JS, Lister C, Lippman Z, Martienssen RA, Dean C (2004) Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* **427**, 164-167.

Baurens FC, Nicolleau J, Legavre T, Verdeil JL, Monteuuis O (2004) Genomic DNA methylation of juvenile and mature *Acacia mangium* micropropagated in vitro with reference to leaf morphology as a phase change marker. *Tree Physiology*, **24**, 401-407.

Besnard G, Acheré V, Jeandroz S, Johnsen Ø, Faivre Rampant P, Baumann R, Müller-Starck G, Skrøppa T, JM Favre (2008) Does maternal environmental condition during reproductive development induce genotypic selection in *Piceaabies*?. *Annals of Forest Science*, **65**, 109.

Birchler JA, Yao H, Chudalayandi S, Vaiman D, Vieti RA (2010) Heterosis. *The Plant Cell*, **22**, 2105-2112.

Bogeat-Triboulot MB, Brosché M, Renaut J, Jouve L, Le Thiec D, Fayyaz P, Vinocur B, Witters E, Laukens K, Teichmann T, Altman A, Hausman JF, Polle A, Kangasjärvi J, Dreyer E (2007) Gradual soil water depletion results in reversible changes of gene expression, protein profiles, ecophysiology, and growth performance in *Populus euphratica*, a poplar growing in arid regions. *Plant Physiology*, **143**, 876-892.

Bonasio R, Tu S, Reinberg D (2010) Molecular Signals of Epigenetic States. *Science*, **330**, 612-616.

Bond DM, Finnegan EJ (2007) Passing the message on: inheritance of epigenetic traits. *Trends in Plant Science*, **12**, 211-216.

Bonhomme L, Monclus R, Vincent D, Carpin S, Lomenech AM, Plomion C, Brignolas F, Morabito D (2009) Leaf proteome analysis of eight *Populus x euramericana* genotypes: genetic variation in drought response and in water-use efficiency involves photosynthesis-related proteins. *Proteomics*, **9**, 4121-4142.

Bossdorf O, Richards CL, Pigliucci M (2008) Epigenetics for ecologists. *Ecology, Letters*, **11**, 106-115.

Bourc'his D, Voinnet O (2010) A Small-RNA Perspective on Gametogenesis, Fertilization, and Early Zygotic Development. *Science*, **330**, 617-622.

Boyko A, Kovalchuk I (2008) Epigenetic control of plant stress response. *Environmental and Molecular Mutagenesis*, **49**, 61-72.

Cervera MT, Ruiz-Garcia L, Martínez-Zapater JM (2002) Analysis of DNA methylation in *Arabidopsis thaliana* based on methylation-sensitive AFLP markers. *Molecular Genetics and Genomics*, **268**, 543-552.

Chen PY, Cokus SJ, Pellegrini MBS (2010) Seeker: precise mapping for bisulfite sequencing. *BMC Bioinformatics*, **11**, 203.

Chinnusamy V, Zhu JK (2009) Epigenetic regulation of stress responses in plants. *Current Opinion in Plant Biology*, **12**, 133-139.

Chouard P (1960) Vernalization and its relations to dormancy. *Annual Review of Plant Physiology*, **11**, 191-238.

Cokus S J, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*, **452**, 215-219.

Cubas P, Vincent, C, Coen, E (1999) An epigenetic mutation responsible for natural variation in floral symmetry. *Nature*, **401**, 157-161.

De Block M, Van Lijsebettens M (2011) Energy efficiency and energy homeostasis as genetic and epigenetic components of plant performance and crop productivity *Current Opinion in Plant Biology*, **14**, 275-282.

Demidov D, Hesse S, Tewes A, Rutten T, Fuchs J, Ashtiyani RK, Lein S, Fischer A, Reuter G, Houben A (2009) Aurora1 phosphorylation activity on histone H3 and its cross-talk with other post-translational histone modifications in Arabidopsis. *The Plant Journal*, **59**, 221-230.

Dormling I, Johnsen Ø (1992) Effects of the parental environment on full-sib families of *Pinus sylvestris*. *Canadian Journal of Forest Research*, **22**, 88-100.

Feil, R, Fraga MF (2011) Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics*, **13**, 97-109.

Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares, EC, Tyson AC, Korlach J, Turner, SW (2010) Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature Methods*, **7**, 461-465.

Fraga MF, Canal MJ, Rodriguez R (2002a) Phase-change related epigenetic and physiological changes in *Pinus radiata* D. Don. *Planta*, **215**, 672-678.

Fraga MF, Rodriguez R, Canal MJ (2002b) Genomic DNA methylation-demethylation during aging and reinvigoration of *Pinus radiata*. *Tree Physiology*, **22**, 813-816.

Gourcilleau D, Bogeat-Triboulot MB, Le Thiec D, Lafon-Placette C, Delaunay A, El-Soud WA, Brignolas F, Maury S (2010) DNA methylation and histone acetylation: genotypic variations in hybrid poplars, impact of water deficit and relationships with productivity. *Annals of Forest Science*, **67**, 208

Grattapaglia D, Plomion C, Kirst M, Sederoff RR (2009) Genomics of growth traits in forest trees. *Current Opinion in Plant Biology*, **12**, 148-156.

Greenwood MS, Hopper CA, Hutchison KW (1989) Maturation in larch: I. Effect of age on shoot growth, foliar characteristics, and DNA methylation. *Plant Physiology*, **90**, 406-412.

Greenwood MS, Hutchison KW (1996) Genetic after effects of increased temperature in Larix. In: Hom J, Birdsey R, O'Brian K eds. *Proceedings of the 1995 Meeting of the Northern Global Change Program*. Radnor: USDA *Forest Service report* 56-62.

Groszmann, M, Greaves IK, Albert N, Fujimoto R, Helliwell CA, Dennis. ES, Peacock WJ (2011a) Epigenetics in plants-vernalisation and hybrid vigour. *Biochimica et Biophysica Acta*, **1809**, 427-437.

Groszmann M, Greaves IK, Albertyn ZI, Scofield GN, Peacock WJ, Dennis ES (2011b) Changes in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to hybrid vigor. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 2617-2622.

Gutierrez-Marcos JF, Costa LM, Dal Prà M, Scholten S, Kranz E, Perez P, Dickinson HG (2006) Epigenetic asymmetry of imprinted genes in plant gametes. *Nature Genetics*, **38**, 876-878.

Ha M, Lu J, Tian L, Ramachandran V, Kasschau KD, Capman EJ, Carrington JC, Chen X, Wang XJ, Chen ZJ (2009) Small RNAs serve as a genetic buffer against genomic shock in Arabidopsis interspecific hybrids and allopolyploids. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 17835-17840.

Hamanishi ET, Campbell MM (2011) Genome-wide responses to drought in forest trees. *Forestry*, **84**, 273-283.

Hashida S N, Uchiyama T, Martin C, Kishima Y, Sano Y, Mikami T (2006) The temperature-dependent change in methylation of the Antirrhinum transposon Tam3 is controlled by the activity of its transposase. *Plant Cell*, **18**, 104-118.

Haun WJ, Laoueille-Duprat S, O'Connell MJ, Spillane C, Grossniklaus U, Phillips AR, Kaeppler SM, Springer NM (2007) Genomic imprinting, methylation and molecular evolution of maize Enhancer of zeste (Mez) homologs. *The Plant Jorunal*, **49**, 325-337.

Hauser M, Aufsatz W, Jonak C, Luschnig C (2011) Transgenerational epigenetic inheritance in plants. *Biochimica et Biophysica Acta*, **1809**, 459-468.

He G, Zhu X, Elling AA, Chen L, Wang X, Guo L, Liang M, He H, Zhang H, Chen F, Qi Y, Chen R, Deng XW (2010) Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell*, **22**, 17-33.

Heo JB, Sung S (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*, **331**, 76-79.

Herrera CM (1990) The adaptedness of the floral phenotype in a relict endemic, hawkmoth-pollinated violet. 2. Patterns of variation among disjunct populations. *Biological Journal of the Linnean Society*, **40**, 275-291.

Herrera CM, Bazaga P (2010) Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. *New Phytologist*, **187**, 867-876.

Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J (2011) An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature*, **472**, 115-119.

Jablonka E, Raz G (2009) Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly Review of Biology*, **84**: 131-176.

Jaligot E, Rival A, Beule T, Dussert S, Verdeil JL (2000) Somaclonal variation in oil palm (*Elaeis guineensis Jacq.*): the DNA methylation hypothesis. *Plant Cell Reports*, **19**, 684-690.

Jaskiewicz M, Conrath U, Peterhansel C (2011) Chromatin modification acts as a memory for systemic acquired resistance in the plant stress response. *EMBO Reports*, **12**, 50-55.

Johannes F, Colot V, Jansen RC (2008) Epigenome dynamics: a quantitative genetics perspective. *Nature Reviews Genetics*, **9**, 883-890.

Johannes F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, Bulski A, Albuisson J, Heredia F, Audigier P, Bouchez D, Dillmann C, Guerche P, Hospital F, Colot V (2009) Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genetics*, **5**, e1000530

Johnsen Ø, Skrøppa T, Junttila O, Dæhlen OG (1996 Influence of the female flowering environment on autumn frost-hardiness of *Picea abies* progenies. *Theoretical and Applied Genetics*, **92**, 797-802.

Johnsen Ø, Fossdal CG, Nagy N, Molmann J, Dælen OG, Skrøppa T (2005) Climatic adaptation in *Picea abies* progenies is affected by the temperature during zygotic embryogenesis and seed maturation. *Plant, Cell & Environment*, **28**, 1090-1102.

Jullien PE, Kinoshita T, Ohad N, Berger F (2006) Maintenance of DNA methylation during the Arabidopsis life cycle is essential for parental imprinting. *The Plant Cell*, **18**, 1360-1372.

Kaeppler SM, Kaeppler HF, Rhee Y (2000). Epigenetic aspects of somaclonal variation in plants. *Plant Molecular Biology*, **43**, 179-188.

Kalisz S, Purugganan MD (2004) Epialleles via DNA methylation: consequences for plant evolution. *Trends in Ecology Evolution*, **19**, 309-314.

Kelly CK, Chase MW, De Bruijn A, Fay MF, Woodward FI (2003) Temperature-based population segregation in birch. *Ecology Letters*, **6**, 87-89.

Kim DH, Sung S (2012) Environmentally coordinated epigenetic silencing of FLC by protein and long noncoding RNA components. *Current Opinion in Plant Biology*, **15**, 51-56.

KohmannK, Johnsen Ø (1994) The timing of bud-set in seedlings of *Picea abies* from seed crops of a cool versus a warm summer. *Silvae Genetica*, **43**, 328-332.

Kouzarides T (2007) Chromatin modifications and their function. *Cell*, **128**, 693-705.

Krauss V (2008) Glimpses of evolution: heterochromatic histone H3K9 methyltransferases left its marks behind. *Genetica*, **133**, 93-106.

Krizova K, Fojtova M, DepickerA, Kovarik A (2009) Cell culture-induced gradual and frequent epigenetic reprogramming of invertedly repeated tobacco transgene epialleles. *Plant Physiology*, **149**, 1493-1504.

Krueger F, Kreck B, Franke A, Andrews SR (2012) DNA methylome analysis using short bisulfite sequencing data. *Nature Methods*, **9**, 145-151.

Ku CS, Naidoo N, Wu M, Soong R (2011) Studying the epigenome using next generation sequencing. *Journal of Medical Genetics*, **48**,721-30.

Kumar SV, Wigge PA (2010) H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. *Cell*, **140**, 136-47.

Kvaalen H, Johnsen Ø (2008) Timing of bud set in *Picea abies* is regulated by a memory of temperature during zygotic and somatic embryogenesis. *New Phytologist*, **177**, 49-59.

Lafon-Placette C, Faivre-Rampant P, Delaunay A, Street N, Brignolas F, Maury S (2012) Methylome of DNase I sensitive chromatin in *Populus trichocarpa* shoot apical meristematic cells: a simplified approach revealing characteristics of gene-body DNA methylation in open chromatin state. *New Phytologist*, in press. doi: 10.1111/nph.12026.

Lim SJ, Tan TW, Tong JC (2010) Computational epigenetics: the new scientific paradigm. *Bioinformation*, **4**, 331-337.

Lira-Medeiros CF, Parisod C, Fernandes RA, Mata CS, Cardoso MA, Ferreira PC (2010) Epigenetic variation in mangrove plants occurring in contrasting natural environment. *PLoS One*, **5**, e10326.

Lisch DR, Slotkin RK (2011) Strategies for silencing and escape: the ancient struggle between transposable elements and their hosts. *International review of cell and molecular biology*, **292**, 119-152.

Long Y, Xia W, Li R, Wang J, Shao M, Feng J, King GJ, Meng J (2011) Epigenetic QTL mapping in Brassica napus. *Genetics*, **189**, 1093-102.

Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nature Genetics*, **38**, 948-952.

Marfil CF, Camadro EL, Masuelli RW (2009) Phenotypic instability and epigenetic variability in a diploid potato of hybrid origin, *Solanum ruiz-lealii. BMC Plant Biology*, **9**, 21.

Miguel C, Marum L (2011) An epigenetic view of plant cells cultured in vitro: somaclonal variation and beyond. *Journal of Experimental Botany*, **62**, 3713-3725.

Mirouze M, Reinders J, Bucher E, Nishimura T, Schneeberger K, Ossowski S, Cao J, Weigel D, Paszkowski J, Mathieu O (2009) Selective epigenetic control of retrotransposition in *Arabidopsis*. *Nature*, **461**, 427-430.

Mirouze M, Paszkowski J (2011) Epigenetic contribution to stress adaptation in plants. *Current Opinion in Plant Biology*, **14**, 267-274.

Monteuuis O, Baurens FC, Goh DKS, Doulbeau S, Verdeil JL (2009) DNA methylation in *Acacia mangium* in vitro and ex-vitro buds, in relation to their within-shoot position, age and leaf morphology of the shoot. Silvae Genetica, **58**, 287-292.

Neale DB, Kremer A (2011) Forest tree genomics: growing resources and applications. *Nature Reviews Genetics*, **12**, 111-122.

Nicotra AB, Atkin OK, Bonser SP, Davidson AM, Finnegan EJ, Mathesius U, Poot P, Purugganan MD, Richards CL, Valladares F, van Kleunen M. (2010) Plant phenotypic plasticity in a changing climate. *Trends in Plant Science*, **15**, 684-692.

Ohri D, Khoshoo TN (1986) Genome size in gymnosperms. *Plant Systematics and Evolution*, **153**, 119-132.

Paun O, Bateman RM, Fay MF, Hedren M, Civeyrel L, Chase MW (2010) Stable epigenetic effects impact adaptation in allopolyploid orchids (*Dactylorhiza: Orchidaceae*). *Molecular Biology and Evolution*, **27**, 2465-2473.

Phillips RL, Kaeppler SM, Olhoft P (1994) Genetic instability of plant tissue cultures: breakdown of normal controls. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 5222-5226.

Plomion C, Lalanne C, Claverol S, Meddour H, Kohler A, Bogeat-Triboulot MB, Barre A, Le Provost G, Dumazet H, Jacob D, Bastien C, Dreyer E, de Daruvar A, Guehl JM, Schmitter JM, Martin F, Bonneu M (2006) Mapping the proteome of poplar and application to the discovery of drought-stress responsive proteins. *Proteomics*, **6**, 6509-6527.

Rabinowicz PD, Citek R, Budiman MA, Nunberg A, Bedell JA, Lakey N, O'Shaughnessy AL, Nascimento LU, McCombie WR, Martienssen RA (2005) Differential methylation of genes and repeats in land plants. *Genome Research*, **15**, 1431-1440.

Raj S, Bräutigam K, Hamanishi ET, Wilkins O, Schroeder W, Mansfield SD, Plant AL, Campbell MM (2011) Clone history shapes *Populus* drought responses. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 12521-12526.

Rake AV, Miksche JP, Hall RB, Hansen KM (1980) DNA reassociation kinetics of four Conifers. *Canadian Journal of Genetics and Cytology*, **22**, 69-79.

Rapp RA, Wendel JF (2005) Epigenetics and plant evolution. *New Phytologist*, **168**, 81-91.

Rehfeldt GE, Ying CC, Spittlehouse DL, Hamilton DA (1999) Genetic responses to climate in pinus contorta: niche breadth, climate change, and reforestation. *Ecological Monographs*, **69**, 375-407.

Rehfeldt GE, Tchebakova NM, Parfenova YI, Wykoff WR, Kuzmina NA, Milyutin LI (2002) Intraspecific responses to climate in *Pinus sylvestris. Global Change Biology*, **8**, 912-929.

Reinders J, Wulff BBH, Mirouze M, Mari-Ordonez A, Dapp M, Rozhon W, Bucher E, Theiler G, Paszkowski J. (2009) Compromised stability of DNA methylation and transposon immobilization in mosaic *Arabidopsis* epigenomes. *Genes & Development*, **23**, 939-950.

Richards EJ (2006) Inherited epigenetic variation - revisiting soft inheritance. *Nature Reviews Genetics*, **7**, 395-401.

Richards CL, Verhoeven KJF, Bossdorf O (2012) *Evolutionary significance of epigenetic variation*. In: Wendel JF, Greilhuber J, Dolezel J, Leitch IJ (eds.), Plant genome diversity Volume I. Springer, Vienna. pp 257-274.

Rival A, Jaligot E, Beule T, Finnegan EJ (2008) Isolation and expression analysis of genes encoding MET, CMT, and DRM methyltransferases in oil palm (*Elaeis guineensis Jacq.*) in relation to the 'mantled' somaclonal variation. *Journal of Experimental Botany*, **59**, 3271-3281.

Rodriguez Lopez CM, Wetten AC, Wilkinson MJ (2010) Progressive erosion of genetic and epigenetic variation in callus-derived cocoa (*Theobroma cacao*) plants. *New Phytologist*, **186**, 856-868.

Rohde A, Junttila O (2008) Remembrances of an embryo: long-term effects on phenology traits in spruce. *New Phytologist*, **177**, 2-5.

Rohde A (2009) Bud set - a landmark of the seasonal growth cycle in poplar. In: Gusta LV, Wisniewski ME, Tanino KK, *eds. Plant Cold Hardiness*. Cambridge, MA: CABI, 91-98.

Roudier F, Ahmed I, Bérard C, Sarazin A, Mary-Huard T, Cortijo S, Bouyer D, Caillieux E, Duvernois-Berthet E, Al-Shikhley L, Giraut L, Després B, Drevensek S, Barneche F, Dèrozier S, Brunaud V, Aubourg S, Schnittger A, Bowler C, Martin-Magniette ML, Robin S, Caboche M, Colot V (2011) Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *EMBO Journal*, **30**, 1928-1938.

Ruttink T, Arend M, Morreel K, Storme V, Rombauts S, Fromm J, Bhalerao RP, Boerjan W, Rohde A (2007) A molecular timetable for apical bud formation and dormancy induction in poplar. *The Plant Cell*, **19**, 2370-2390.

Santamaria ME, Hasbun R, Valera MJ (2009) Acetylated H4 histone and genomic DNA methylation patterns during bud set and bud burst in *Castanea sativa*. *Journal of Plant Physiology*, **166**, 1360-1369.

Schellenbaum P, Mohler V, Wenzel G, Walter B (2008) Variation in DNA methylation patterns of grapevine somaclones (*Vitis vinifera* L.). *BMC Plant Biology*, **8**, 78.

Schmitz RJ, Amasino RM (2007) Vernalization: a model for investigating epigenetics and eukaryotic gene regulation in plants. *Biochimica et Biophysica Acta*, **1769**, 269-275.

Schull GH (1946) Hybrid seed corn. *Science*, **103**, 547-550.

Skrøppa T, Johnsen Ø (2000) Patterns of adaptive genetic variation in forest tree species; the reproductive environment as an evolutionary force in *Picea abies*. In *Forest Genetics and Sustainability* (ed. C. Mátyás), pp. 49-58. *Kluwer Academic Publications*, Dordrecht, The Netherlands.

Skrøppa T, Kohmann K, Johnsen Ø, Steffenrem A, Edvardsen ØM (2007) Field performance and early test results of offspring from two Norway spruce seed orchards containing clones transferred to warmer climates. *Canadian Journal of Forest Research*, **37**, 515-522.

Skrøppa T, Tollefsrud M, Sperisen C, Johnsen Ø (2010) Rapid change in adaptive performance from one generation to the next in Picea abies-Central European trees in a Nordic environment. *Tree Genetics & Genomes*, **6**, 93-99.

Slotkin RK, Martienssen RA (2007) Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics*, **8**, 272-285.

Stoehr MU, L'Hirondelle SJ, Binder WD, Webber JE (1998) Parental environment after effects on germination, growth, and adaptive traits in selected spruce families. *Canadian Journal of Forest Research*, **28**, 418-426.

Thumma BR, Matheson BA, Zhang D, Meeske C, Meder R, Downes GM, Southerton SG (2009) Identification of a Cis-acting regulatory polymorphism in a Eucalypt COBRA-like gene affecting cellulose content. *Genetics*, **183**, 1153-1164.

Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T (2009). Bursts of retrotransposition reproduced in Arabidopsis. *Nature*, **461**, 423-426.

Valledor L, Hasbún R, Meijón M (2007) Involvement of DNA methylation in tree development and micropropagation. *The Plant Cell Tissue and Organ Culture*, **91**, 75-86.

Valledor L, Meijon M, Hasbun R, Jesus Canal M, Rodriguez R (2010) Variations in DNA methylation, acetylated histone H4, and methylated histone H3 during *Pinus radiata* needle maturation in relation to the loss of in vitro organogenic capability. *Journal of Plant Physiology,* **167**, 351-357.

Van Steensel B (2011) Chromatin: constructing the big picture. *The EMBO Journal*, **30**, 885-1895.

Verhoeven KJ, Jansen JJ, van Dijk PJ, Biere A (2010) Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytologist*, **185**, 1108-1118.

Vining KJ, Pomraning KR, Wilhelm LJ, Priest HD, Pellegrini M, Mockler TC, Freitag M, Strauss SH (2012) Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. *BMC Genomics*, **13**, 27.

Viswanath V, Albrechtsen BR, Strauss SH (2012) Global regulatory burden for field testing of genetically modified tress. *Tree Genetics & Genomes*, **8**, 221-226.

Webber J, Ott P, Owens J, Binder W (2005) Elevated temperature during reproductive development affects cone traits and progeny performance in *Picea glauca - engelmannii complex. Tree Physiology*, **25**, 1219-1227.

Whittle CA, Otto SP, Johnston MO, Krochko JE (2009) Adaptive epigenetic memory of ancestral temperature regime in *Arabidopsis thaliana. Botany*, **87**, 650-657.

Wilkins O, Waldron L, Nahal H, Provart NJ, Campbell MM (2009) Genotype and time of day shape the *Populus* drought response. *The Plant Journal*, **60**, 703-715.

Yakovlev IA, Fossdal CG, Johnsen Ø (2010) MicroRNAs, the epigenetic memory and climatic adaptation in Norway spruce. *New Phytologist*, **187**, 1154-1169.

Yakovlev IA, Asante DKA, Fossdal CG, Junttila O, Johnsen Ø (2011) Differential gene expression related to an epigenetic memory affecting climatic adaptation in Norway spruce. *Plant Science*, **180**, 132-139.

LARGE-SCALE LONGITUDINAL GRADIENTS OF GENETIC DIVERSITY: A META-ANALYSIS ACROSS SIX PHYLA IN THE MEDITERRANEAN BASIN

BIOGEOGRAPHY OF GENES IN THE MEDITERRANEAN

Cyrille CONORD[1,2]; Jessica GUREVITCH[3]; Bruno FADY[1]

[1] INRA, FR ECCOREV, UR629, Écologie des Forêts Méditerranéennes, Avignon, France.
[2] Université de Saint-Etienne, Jean Monnet, Laboratoire BVpam, EA2061, 23 rue du Dr Michelon, F-42000 Saint-Etienne, France.
[3] Department of Ecology & Evolution, Stony Brook University, NY, USA.

Corresponding Author:
Bruno FADY
INRA, UR629, Écologie des Forêts Méditerranéennes, Avignon, France, bruno.fady@avignon.inra.fr,

Biodiversity is the diversity of life at all scales, from genes to ecosystems. Predicting its patterns of variation across the globe is a fundamental issue in ecology and evolution. Diversity within species, i.e. genetic diversity, is of prime importance for understanding past and present evolutionary patterns, and highlighting areas where conservation might be a priority. Using published data on the genetic diversity of species whose populations occur in the Mediterranean basin, we calculated a coefficient of correlation between within-population genetic diversity indices and longitude. Using a meta-analysis framework, we estimated the role of biological, ecological, biogeographical and marker type factors on the strength and magnitude of this correlation in six phylla. Overall, genetic diversity increases from west to east in the Mediterranean basin. This correlation is significant for both animals and plants, but is not uniformly expressed for all groups. It is stronger in the southern than in the northern Mediterranean, in true Mediterranean plants than in plants found at higher elevations, in trees than in other plants and in bi-parentally and paternally than in maternally inherited DNA makers. Overall, this correlation between genetic diversity and longitude, and its patterns across biological and ecological traits, suggests the role of two non-mutually exclusive major processes that shaped the genetic diversity in the Mediterranean during and after the cold periods of the Pleistocene: east-west recolonization during the Holocene and population size contraction under local Last Glacial Maximum climate in resident western and low elevation Mediterranean populations.

Key words: Biodiversity; biogeography; past-climate; genetic diversity; recolonization; Holocene; longitude; meta-analysis; Arthropod; Mollusc; Chordata; Bryophyte; Pteridophyte; Spermaphyte; phylogeography; Pleistocene.

Published in *Ecology and Evolution* (2012) 2: 2600-2614, with doi:10.1002/ece3.350

Introduction

Biodiversity is the diversity of life at all scales, *i.e.* "the variability among living organisms from all sources including, *inter alia*, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems" (article 2 of the Convention on Biological Diversity1992). Although knowledge of the distribution of species is far from always being spatially accurate and detailed (Richardson & Whittaker 2010), species diversity and abundance is relatively well known for several taxonomic groups (*e.g.* mammals, birds, fishes, vascular plants) and particularly in the temperate regions of the world. This knowledge has helped shaping the delineation of hotspots of biological diversity (Myers *et al.* 2000) where conservation is most critical and is at the core of the field of conservation biogeography (Whittaker *et al.* 2005).

At a finer taxonomic scale, genetic diversity, diversity among individuals within species, yields valuable information for understanding how biodiversity changes within an evolutionary framework. Within the broad context of biogeography, genetic diversity has been assessed either through ecogeographic or through phylogeographic perspectives (Avise, 2000). The ecogeographic view focuses on patterns produced by contemporary natural selection, as for example, the genetic structure of Mediterranean pine stands exposed to wild fires (Aravanopoulos *et al.* 2004). Conversely, the phylogeographic approach focuses largely on historical evolutionary processes such as the balance between vicariance and dispersal to examine genetic differentiation both among and within populations. Measures of genetic differentiation among populations, mapped against major geographical barriers, have been used to derive the most likely Quaternary glacial refugia and Holocene recolonization routes of major temperate species (Taberlet *et al.* 1998, Hewitt 1999). Measures of genetic diversity within population (*GDpop*) have been used to refine how Holocene recolonization occurred for multiple species and vegetation types, *e.g.* from southern refugia in Europe and North America (Petit *et al.* 2003, Soltis *et al.* 2006).

Large-scale phylogeographic studies have identified latitudinal Holocene recolonization as well as complex

patterns of post-glaciation dispersal from refugia at landscape to regional scales as the major drivers of genetic diversity in the Northern Hemisphere (e.g., Petit *et al.* 2003, Brewer *et al.* 2002, Liepelt *et al.* 2009, for Europe). Phylogeographic studies at large scales have not often considered longitude as a potentially important ecological driver, even though refugia are distributed longitudinally in southern Europe (Stewart *et al.* 2010). In the Mediterranean basin, longitudinal trends are potentially an important factor in determining genetic diversity because of how the geography of southern Europe and the Mediterranean region is shaped. The Mediterranean Sea is a strong barrier to latitudinal movements of terrestrial species but also to longitudinal movements from one peninsula to the other in Europe, with potentially strong impacts in shaping contemporary biodiversity structures. Unveiling longitudinal patterns of diversity in the Mediterranean would be of great interest because, quoting from Atkinson *et al.* (2007), "*longitudinal processes represent the raw material on which later latitudinal processes work*" in Europe. The purpose of the present study is to examine whether longitudinal patterns of genetic diversity are important in the Mediterranean region and southern Europe across a large range of taxa. Such patterns may reveal a different perspective on post-glaciation colonization at large geographic and taxonomic scales.

In the same manner that a gene tree only depicts a very small part of the phylogenetic history of lineage, the population genetic structure of species can only represent a small slice of the history of a whole region or biome. Whether they have followed the ecogeographical or the phylogeographical approach, many empirical studies of population genetic structure provide *GDpop* estimates. Gathering and comparing the wealth of information contained in these empirical studies, using a proper statistical framework, represents an excellent opportunity to meet the challenge of testing processes determining genetic diversity patterns at regional scale or biome-wide, such as for the Mediterranean. We used meta-analysis to document and test the existence of cross-taxa longitudinal patterns of genetic diversity in the Mediterranean basin. No previous studies have examined genetic differentiation within and across populations at large geographic scales using the powerful statistical tools of meta-analysis.

Biogeographic genetic analyses have mostly focused on population structure and differentiation rather than on within-population diversity because genetic variation at neutral markers is not expected to respond to environmental effects (but see, *e.g.*, Petit *et al.* 2003). Strong spatial gradients of neutral genetic differentiation are thus only expected as a consequence of historical effects such as directional dispersal during range expansions from refugia during global warming periods, which leads to marked population structure (Petit *et al.* 2003). However, because demographic changes can impact *GDpop* (Young *et al.* 1996), strong gradients of *GDpop* can also be expected as an indirect response to clinal environmental effects, such as past climates.

GDpop is of fundamental importance in ecology and evolution because it is correlated with population demographic rates and, in numerous circumstances, with their potential for evolutionary adaptive change (Le Corre & Kremer 2003). *GDpop* is therefore important for identifying regions where evolutionary potential is either particularly low or high, thus providing insights for conservation strategies and planning (Schwartz *et al.* 2007). *GDpop* can be estimated in various ways, but the approaches fall within two general categories: "richness" (total amount of diversity, *e.g.* allelic richness, haplotypic richness) and "equitability" (the way diversity is distributed among samples, *e.g.* heterozygosity, Shannon's index, percentage of polymorphic loci). Whereas equitability measures are more sensitive to higher-frequency alleles and how they are distributed within populations, richness measures respond more to the presence and quantity of rare alleles. Thus, the two measures are needed concurrently to document and test for demographic events such as bottlenecks and expansions.

The Mediterranean basin is a hotspot of species diversity (Myers *et al.* 2000), and also a world region of unusually high *GDpop* (see Fady 2005 for conifers). Vascular plants are structured into regional hotspots of species diversity and endemism (Médail & Quézel 1997) often corresponding to glacial refugia (Médail & Diadéma 2009). The northern Mediterranean basin is made of south-north oriented peninsulas identified as independent Quaternary glacial refugia and starting points of Holocene recolonization for Europe (Hewitt 1999; Petit *et al.* 2003). The shoreline of the southern Mediterranean basin is more or less linear, without major peninsulas. Its western part, North Africa, is also recognized as a refugial zone (*e.g.* Cheddadi *et al.* 2009, Guzmán & Vargas 2009).

Two major causes can be hypothesized for longitudinal trends of *GDpop*, if such trends can be demonstrated, for natural populations in the Mediterranean (both in southern Europe and North Africa). First, longitudinal trends could result from genetic drift due to long distance dispersal and founder effects during Holocene recolonization from refugia (*e.g.*, from eastern Mediterranean refugia as in the tree *Pinus halepensis*, Grivet *et al.* 2009, or in the wasp *Andricus quercustozae*, Rokas *et al.* 2003, from western Mediterranean refugia as in the tree *Pinus sylvestris*, Soranzo *et al.* 2000). However, both uni-directional recolonization patterns appear less likely than multi-refugium recolonization patterns (Taberlet *et al.* 1998). The second cause for longitudinal trends in this region could be genetic drift due to decreasing effective population size, given the existence of a climate of increasing severity from east to west in the Mediterranean during the last glacial cycle, particularly the Last Glacial Maximum (LGM)

21 000 years before present (van Andel 2002; Wu *et al.* 2007). There is also evidence of climatic instability over the North-Atlantic Ocean leading to several extreme cooling events over the Iberian Peninsula during the last glacial cycle (Sánchez-Goñi *et al.* 2002). The potential effect of such past climate trends on *GDpop* was described for gallwasps (Atkinson *et al.* 2007) and trees (Fady & Conord 2010).

Looking at patterns across multiple levels of biodiversity provides a framework to understand processes beyond the idiosyncrasy of case studies. Here, using a meta-analytical framework, we tested the existence of a longitudinal trend of within-population genetic diversity in the Mediterranean basin at multiple taxonomic levels in the tree of life. For each population genetic study we retrieved from the literature, we calculated a correlation coefficient between longitudinal coordinates and genetic diversity. We examined data on populations in the Arthropods, Mollusks, Chordata, Bryophytes, Pteridophytes, and Spermaphytes. Previous studies of genetic diversity at large spatial scales have generally focused on far smaller taxonomic groups (*e.g.*, Riddle *et al.* 2000, vertebrates from Baja California, North America; Petit *et al.* 2003, trees from Europe; Kadereit *et al.* 2005, dicots from the Mediterranean basin).

We hypothesized that we would detect a west-east trend of increasing genetic diversity across all taxa if the demographic (and thus genetic) clinal imprint left by the climate of the last glacial cycle on resident populations in refugia was stronger than the imprints left by the incongruent Holocene recolonization patterns of different species from different refugia (Taberlet *et al.* 1998). In contrast, if recolonization from disparate refugia across multiple taxa is the dominant signal for current patterns of *GDpop*, we would not expect to find such a longitudinal imprint across taxa. Refugia have been identified in many different parts of the region. For example, Médail & Diadéma (2009) in their analysis of plant genetic patterns in the Mediterranean found that out of 52 refugia identified, 33 were in the western Mediterranean and 19 in the eastern Mediterranean (non-significantly different from an equal distribution in each zone). At species level, we expected different trends depending on where refugia were located. We expected that this trend would be weaker for studies using *GDpop* measures giving higher weight to frequent alleles because rare alleles are more likely to disappear with recolonization and bottleneck events than do frequent alleles.

We also predicted that, depending on their position in the tree of life, the different taxa would respond differently to longitude. We expected that their responses would depend on:

- their life-history traits (low versus high mobility);
- their bioclimatic requirements (particularly in plants, depending on their over-wintering abilities and their temperature requirements, and thus their sensitivity to local glacial climate); and
- their location within distribution areas (islands versus continents and southern versus northern Mediterranean, for which demographic effects and migration possibilities are different).

Material and methods

1. Gathering data from published sources

We collected published population genetic studies of terrestrial plant and animal species whose distributions were at least partially found within the Mediterranean basin, from endemic to widespread. For this, we searched the Web of Knowledge between for references published between Jan. 1980 and Oct. 2009. We used the following search expressions: 'population genetics', 'phylogeo*', 'genetic diversit*', 'Mediterr*' and individual Mediterranean country names. We also examined the references included in all retrieved publications for additional references.

2. Defining the geographic zone of interest and population sample size

From these papers, we selected all populations with *GDpop* estimates that were included within the Mediterranean basin. We used the delineation of the Mediterranean basin defined by Olson *et al.* (2001) which is the standard currently used by the World Wildlife Fund (WWF) to define the world eco-regions. We used a geographic information system (GIS) for selecting among published studies which population to allocate to that geographic envelope and to further qualify populations as continental *versus* insular and northern *versus* southern Mediterranean.

3. Constructing a database of *GDpop* estimates

We explain in details in Supplementary Material I how we constructed our database. The list of published studies used is referenced in Supplementary Material II.

4. Raw data analysis and calculation of effect-sizes

No published paper we retrieved had testing for a correlation between longitude and *GDpop* as its primary goal. We used the raw data from these studies to correlate the longitudinal position of each population with its *GDpop*. The statistics we used was the Pearson product-moment correlation coefficient.

The populations tested are not located on a strait longitudinal line but rather span a small latitudinal gradient (which reaches its maximum in each of the Mediterranean peninsulas). Also, not all organisms remained in the close vicinity of their glacial refugia after Holocene warming (receding edge populations may have moved farther away from refugia than rear edge populations or endemic species, Jump *et al.* 2009). Thus in order to account for the effect of latitude, we calculated a partial correlation coefficient which measured the degree of association of

GDpop with longitude while latitude was held constant. Partial correlation coefficients in meta-analysis represent the relationship between the independent and the dependent variable while controlling for other factors (Rosenthal & DiMatteo 2001, Keef & Roberts 2004). In our case, the controlling factor was always the same, latitude.

The partial correlation coefficients were transformed using Fisher's Z-transformation and used as effect-sizes: $Z = 0.5*ln(1+r/1-r)$, where r is the partial correlation coefficient. Finally, effect-sizes (Z-transformed partial correlation coefficients) were weighted by the inverse of their asymptotic variance (see Aloe & Becker 2009 for weighting partial coefficients in meta-analysis) which was calculated as $Vz = 1/(n-3)$, where n is the number of sampled populations in the source study. Finally, summary-effects (*Zr*) were computed as the weighted mean of the individual effect sizes using a fixed-effect meta-analysis model. We chose a fixed-effect model (Borenstein *et al.* 2009) because we were interested in testing primary factors affecting *GDpop* that acted in a similar way across all species and were expressed non-randomly across the whole Mediterranean basin.

Because data in our primary dataset are not all independent (several papers in our dataset address the same species), we tested the effect of non-independence on Type I error rates as well as on the precision of summary-effects (Hartung *et al.* 2008) using several methods (see Supplementary Material III). As the redundancy of our raw data affected neither the direction of the relationship between *GDpop* and longitude nor its significance, we decided to use the entire dataset in the following meta-analyses and to not perform any statistical treatment to reduce redundancy.

Data processing, effect-size computations as well as sensitivity analyses were done using R packages *plyr v0.1.9* (Wickham 2011), *MAc v1.1* (Del Re & Hoyt 2010) and custom scripts for sensitivity analyses (available upon request), under R v2.11 (http://www.r-project.org/), whereas the meta-analysis procedure was performed with MetaWin (Rosenberg *et al.* 2000).

Exploring moderators of the summary-effects

1. Moderators related to primary study design

Effect of the sampling geographical range in the primary studies

If we hypothesize a general and homogeneous link of *GDpop* with longitude (supposing that longitude is a proxy for the same phenomenon for all species/studies), then we may expect that the wider the range of population sampling across the Mediterranean basin in the primary studies was, the greater the probability of detecting a positive mean effect-size *Zr* would be. Also, *Zr* might be affected by the position of the range within the Mediterranean basin. We tested these relationships by regressing each *Zr* and their corresponding longitudinal span calculated as the absolute value of the difference in degrees of longitude between the easternmost and the westernmost populations and each *Zr* and their corresponding mean longitudinal coordinate.

Choice of genetic marker and *GDpop* metric

The choice of the genetic marker may impact the sign of the correlation between *GDpop* and geography because they may reflect different processes acting at different spatial and time scales (ecological vs evolutionary). They may also reflect different demographic histories via their different effective population size or sex-related transmission. Discrepancies have classically been found by phylogeographers between the nuclear and the mitochondrial DNA (Petit & Vendramin 2007). We thus tested marker type effects by categorizing them depending on their inheritance type (male, female or bi-parental inheritance) which may be related to an effect of dispersal ability. Because foundation events or distance to refugia may be imprinted differently on the different types of *GDpop* measures (see the *Fagus sylvatica* example in Comps *et al.* 2001), we tested metric type effects by categorizing *GDpop* measures as either 'equitability' or 'richness' measures (see introduction).

2. Biogeographic effect (north vs south, continents vs islands)

We tested biogeographic effects by categorizing the effect-sizes as either northern or southern Mediterranean, and as either from continents or islands.

3. Plant species biological attributes and ecological requirements

There may be a strong confounding effect between taxonomy and marker type in our general dataset. Specifically, cpDNA effects may be due to the type of DNA used or to traits specific to plants as this type of DNA is not present in animals. Thus, we used the part of our dataset restricted to plants to retest for marker type effects on overall trends and also to test for the imprint of biological attributes and ecological requirements on *GDpop* in the Mediterranean.

Along with demographic processes, many life history traits can be responsible more or less deterministically for spatial *GDpop* differences. For example, generation time is a life history trait that will moderate the imprint of past demographic events on the measured *GDpop*. As a first general and very coarse proxy of those traits, we used taxonomy and we characterized all species by their family, class and kingdom names.

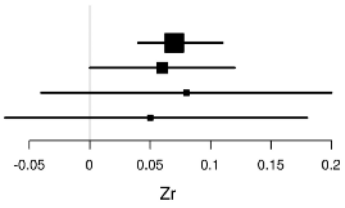
For each plant species, we also recorded several biological and ecological attributes using information from Quézel & Médail (2003), Pignatti (1982), Rameau *et al.* (2008) as well as the Telabotanica web database (*www.tela-botanica.org/* consulted in April 2009). We coded the

FIGURE 4

Square and error bars should be interpreted as indicated for Fig. 2a.

4a

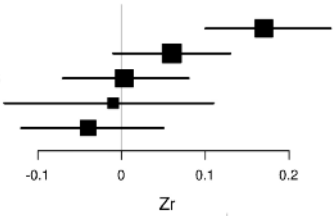
Moderator	Number of effect-sizes	Sign. of Q-test	Effect-size
nuclear	291	***	0.07
plastidial	64	***	0.06
mitochondrial	37	*	0.08
genomic	36	**	0.05



4a: Mean effect-sizes across the Mediterranean basin for marker type. The category "genomic" refers to unassigned marker types.

4b

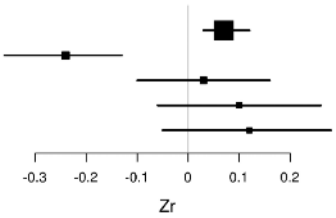
Me	53	***	0.17
Su	39	***	0.06
Eurytherms	64	***	0.0033
Th	27	***	-0.01
Mt	82	***	-0.04



4b: Mean effect-sizes (ordered from most positive to most negative) of Mediterranean basin plants for ecological requirements. The categories refer to altitudinal belts where plant species are predominantly found: 'Me' is meso-Mediterranean, 'Su' is supra-Mediterranean, 'Eurytherms' refers to plant species found across several altitudinal belts, 'Th' is thermo-Mediterranean and 'Mt' is mountain-Mediterranean (see Quézel & Médail 2003). Square and error bars should be interpreted as indicated for Fig. 2a. Nota Benne: effect-sizes for ecological requirements do not add up to the total number of effect-sizes in plants because raw data communicated by some authors were pooled at the genus level or because data included species for which we were not able to retrieve their ecological requirement.

4c

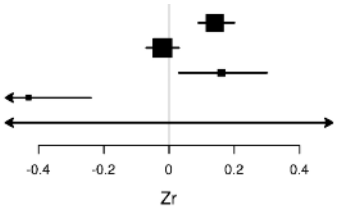
Phanerophytes	192	***	0.07
Chamaephytes	25	***	-0.24
Hemicryptophytes	23	.	0.03
Therophytes	16	***	0.10
Geophytes	13	NS	0.12



4c: Mean plant effect-sizes (arranged in decreasing effect-size frequency per category) across the Mediterranean basin for Raunkiaer biological types. Phanerophytes are woody plants with over-wintering buds situated over 50 cm from the ground, chamaephytes are low growing perennials (often woody plants) with wintering buds below 50 cm in height, hemicryptophytes are (often two-year cycle) perennials with ground-level wintering buds, geophytes are plants with bulbs or rhizomes (wintering buds below ground level) and therophytes are annuals (wintering organs as seeds). Nota Benne: Bryophytes were not assigned a Raunkiaer type (9 effect-sizes).

4d

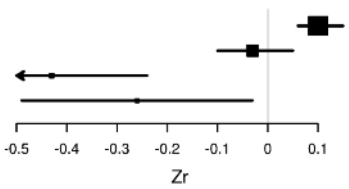
anemochorous	147	***	0.14
zoochorous	90	***	-0.02
barochorous	17	NS	0.16
Non assign.	12	NS	-0.43
hydrochorous	3	NS	-0.01



4d: Mean plant effect-sizes (arranged in decreasing effect-size frequency per category) across the Mediterranean basin for seed dispersal types. Anemochorous plants have wind-dispersed seeds, zoochorous plants animal-dispersed seeds, barochorous plants gravity-dispersed seeds and hydrochorous plants water-dispersed seeds.

4e

anemogamous	174	***	0.10
entomogamous	76	***	-0.03
Non assign.	12	NS	-0.43
hydrogamous	7	NS	-0.26



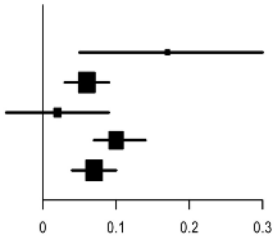
4e: Mean plant effect-sizes (arranged in decreasing effect-size frequency per category) across the Mediterranean basin for pollen dispersal types. Anemogamous plants have wind-dispersed pollen, entomogamous plants insect-dispersed pollen and hydrogamous plants water-dispersed pollen.

altitudinal thermo-climatic belts (from thermo- to oro-Mediterranean, see Fig. 4 for details) where each plant species was predominantly found. Bioclimatic requirements have been found to affect *GDpop* (Kadereit *et al.* 2005; Fady & Conord 2010; Soto *et al.* 2010). Species with higher temperature requirements may be more sensitive to demographic fluctuation in the Mediterranean as they will have been more strongly impacted by unfavorable cold climate during the cold cycles of the Quaternary. We also categorized each plant species according to their pollen and seed dispersal type (see Figs 4d and 4e for the detailed types within categories). Seed dispersal type may influence *GDpop* as a result of migration, as illustrated by the comparison of beech vs hornbeam in Europe. Beech with its animal-dispersed nuts conserved more genetic diversity when crossing mountain barriers than did hornbeam with its winged seeds (Coart *et al.* 2005). Finally, each species in our database was categorized according to its Raunkiaer life-form, which is based on the position of the plant's buds during the unfavorable season and it therefore may be a proxy of life history traits playing an important role in the survival of the species under harsh conditions at the LGM.

FIGURE 2

2a

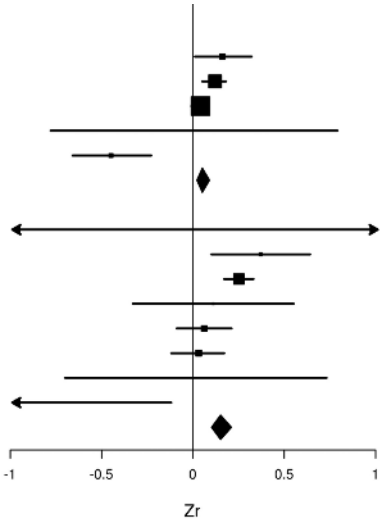
Dataset	Number of Effect-sizes	Signif. of Q-test	Effect-size
Southern Mediterranean	53	***	0.17
Northern Mediterranean	399	***	0.06
Mediterrean islands	108	***	0.02
Mediterranean continent	377	***	0.10
Mediterranean Basin	428	***	0.07



2a: Mean effect-sizes across the Mediterranean basin and for different biogeographical envelopes. Each square whose size is proportional to the number of effect-sizes, indicates mean values and each bar the 95% confidence interval around the mean. Bars not intercepting the Y-axis indicate a significant correlation between GDpop and longitude. When the mean is on the positive side of the X-axis, this correlation is positive, indicating that GDpop increases from west to east. Bars with arrows indicate that the 95% confidence interval falls outside the limits of the figure. Diamonds highlight mean values for plants and animals. All Q-tests are highly significant ($P < 0.0001$), indicating that significant differences between effect sizes remain within each biogeographical category. Nota Benne: northern / southern and continent / island effect-sizes do not add up to the total number of effect-sizes in the Mediterranean basin because raw data studies may include species with populations present north and south of the Mediterranean and both on continents and on islands.

2b

Moderator			
Monocotyledone	14	NS	0.15
Gymnosperm	122	***	0.11
Dicotyledone	121	***	0.03
Pteridophyte	3	NS	-0.01
Bryophyte	9	NS	-0.46
Plants	**269**	***	**0.04**
Crustacea	2	NS	0.56
Arachnida	10	NS	0.36
Insect	69	***	0.24
Reptile	8	NS	0.1
Mammal	33	*	0.05
Amphibian	29	***	0.02
Bird	5	*	0
Gasteropod	3	NS	-0.96
Animals	**159**	***	**0.14**



2b: Mean effect-sizes across the Mediterranean basin for 5 plant and 8 animal classes, and mean values for the plant and animal kingdoms. Classes are arranged in decreasing effect size order. Square and error bars should be interpreted as indicated for Fig. 2a

Results

1. Overall effect-sizes and role of range, taxonomy, and DNA markers

Overall, there was a positive and significant correlation between *GDpop* and longitude in the Mediterranean (Fig. 2a). Within-population genetic diversity decreases from east to west in the Mediterranean (Table 1; Supplementary Material IV). Out of the 428 effect sizes generated from 143 plant and animal species from 156 published studies in our meta-analysis, 54% showed a positive effect size. Considerable heterogeneity was found (*Q = 1180, df = 427, P < 0.0001*) leading to the tests of the categorical moderators reported below. The general longitudinal trend in *GDpop* was affected by sampling range and range type, by marker and metric types, by taxonomy (phylogeny), by biological traits and by ecological requirements.

Sampling range and range type

Both sampling range span and the mean longitudinal position of the studies were significantly but weakly correlated with *Zr (r = 0.0043, P < 0.0001 and r = 0.0086, P < 0.0001, respectively)*. Widely distributed species and species from the eastern part of the Mediterranean tended to have more significantly positive *Zr* than others. The *Zr* was five times higher for continents than for islands and almost three times higher for the southern than for the northern Mediterranean (Fig. 2a). However, *Zr* was positive and significant for all geographic envelopes except for Mediterranean islands. All tests based on Q statistics rejected the null hypothesis of homogeneity among effect-sizes within category, *i.e* that all studies shared a common effect size ($P < 0.0001$, Supplementary Material Table IV-1). This suggested that there was more variability among the effect-sizes of a category than expected by chance and justified the search and testing of moderating variables.

Studies in our dataset sampled a majority of northern Mediterranean populations (399 effect sizes in the north compared to 53 for the south). Differences in *Zr* between the northern and southern Mediterranean were not biased by high order taxonomic differences. The number of effect-sizes belonging to the 6 different phyla of the database (Arthropods, Mollusks, Chordata, Bryophytes, Pteridophytes, Spermaphytes) were not significantly different between the northern and southern Mediterranean (contingency Chi-square test = 12.05, df = 5, $P = 0.06$), although they were between continents and islands (contingency Chi-square test = 15.448, df = 5, $P = 0.01$).

Phylogeny and taxonomic group

The *Zr* computed for the animal and plant kingdoms were both positive and significant. *Zr* for animals was more than three times higher than that of plants. The two Q-tests rejected homogeneity of *Zr* within plants and animals (Fig. 2b), indicating that within each kingdom, finer level groups departed significantly from the positive trend. At a finer taxonomic level, 4 classes out of 13 had a summary-effect not intercepting the zero (Bryophytes, Gymnosperms, Arachnida and Insects, Fig. 2b). Heterogeneity tests were non-significant for all class levels with less than 10 effect-

TABLE 1

Description of the data set and summary statistics: Geographic range (including islands) from Olson *et al.* 2001.

Taxonomic group	Species nr.	Studies nr.	Effect-sizes nr.	mtDNA*	cpDNA*	cpDNA SSR*	nDNA*	nDNA Isozymes*	nDNA SSR*
Bryophyte	2	4	9	-	0	0	5	4	0
Dicotyledone	47	40	121	-	33	13	88	56	16
Gymnosperm	17	55	122	-	28	28	93	76	4
Monocotyledone	7	6	14	-	2	2	12	9	3
Pteridophyte	1	1	3	-	0	0	3	0	3
Total plants	*74*	*106*	*269*	*-*	*63*	*43*	*201*	*145*	*26*
Amphibian	14	7	29	6	-	-	23	23	0
Arachnida	6	4	10	0	-	-	10	10	0
Birds	3	3	5	5	-	-	0	0	0
Crustacea	1	1	2	2	-	-	0	0	0
Gasteropod	1	1	3	0	-	-	3	3	0
Insect	30	26	69	17	-	-	54	27	18
Mammal	10	11	33	8	-	-	25	12	12
Reptile	4	4	8	0	-	-	8	5	3
Total animals	*69*	*57*	*159*	*36*	*-*	*-*	*123*	*80*	*33*
Total	**143**	**163**	**428**	**36**	**63**	**43**	**324**	**225**	**59**

*: number of effect-sizes.

sizes except for birds (Fig. 1b), suggesting consistent responses among the members of these groups; however, the Q test is not very powerful and may fail to detect true heterogeneity, particularly in such small groups. The remaining groups were highly significantly heterogeneous. At the yet finer taxonomic level of the family, significant longitudinal *GDpop* structures could be observed in the Pinaceae, Cupressaceae, Poaceae and Asteraceae (positive summary-effects) and in the Lamiaceae, Nymphalidae and Pottiaceae (negative summary-effects, Fig. 3). As the taxonomic sampling was unbalanced, we ran the meta-analysis excluding successively the most represented groups, from the higher taxonomic level to the finer. Excluding the Phanerophytes (192 effect-sizes) had no effect on the *Zr* computed for the whole dataset and for the continental Mediterranean. However, *Zr* for the northern Mediterranean became non-significant whereas *Zr* for southern Mediterranean and for islands increased and even became significant for islands. Then, excluding gymnosperms (122 effect-sizes) confirmed the patterns observed for the phanerophytes exclusion except for the southern Mediterranean *Zr* which became non-significant. In animals, two families showed a positive trend (Buthidae and Tephritidae) whereas two other showed the opposite trend (Plethodontidae and Torymidae) (Fig. 3).

Markers type and metric type

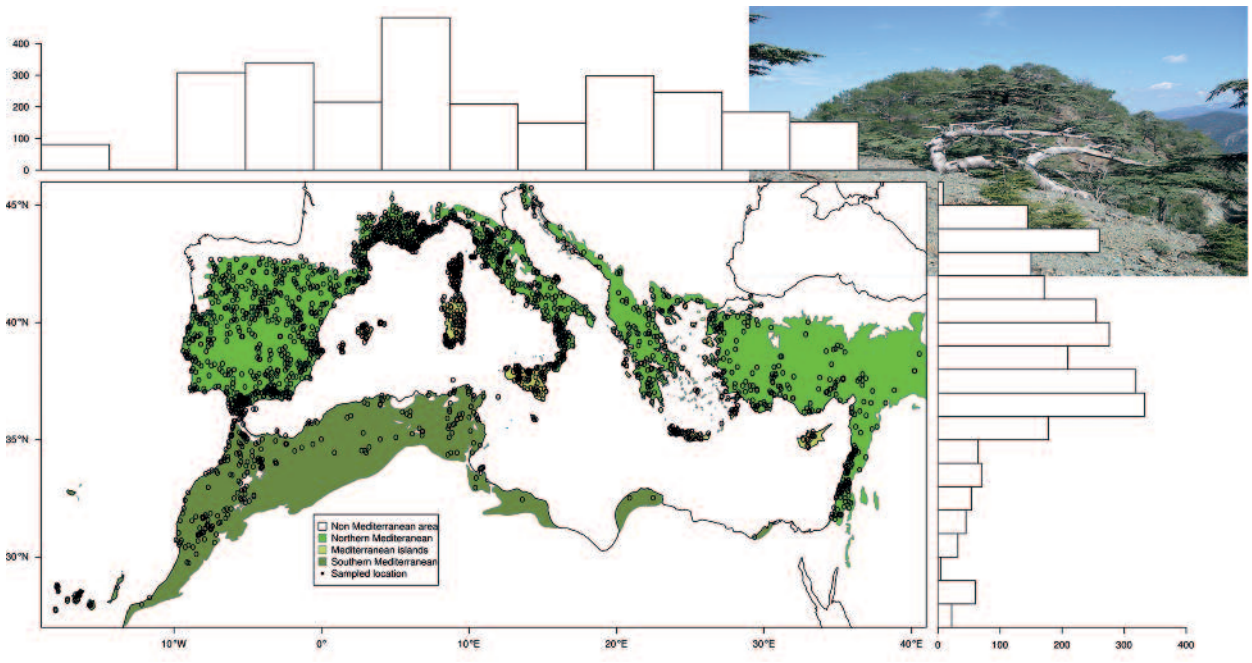
Nuclear and organelle markers showed a positive *Zr* of the same order of magnitude although mitochondrial *Zr* was not significantly different from zero (Fig. 4a). When assessing the effect of the inheritance of the genetic marker, we found that bi-parentally and paternally inherited markers yielded a significant *Zr*. On the contrary, *Zr* for maternally inherited markers (mitochondrial DNA in all species of our dataset and plastidial DNA in angiosperms) was positive but non-significant (not shown), indicating that *GDpop* for maternally inherited markers does not significantly increase with longitude. *Zr* for both 'equitability' and 'richness' indices were significant, positive and similar in magnitude. Excluding island populations from the dataset increased *Zr* values. *Zr* for 'equitability' was three times higher when restricting the dataset to the southern Mediterranean envelope whereas *Zr* for 'richness' stayed constant but became non-significant (not shown).

2. Effect sizes in plants and role of ecological requirements and biological traits

The summary-effects for bi-parentally and maternally inherited markers decreased and became (or remained in the case of maternally inherited markers) non-significant after excluding animals from the dataset (not shown). In

FIGURE 1

Map of the Mediterranean Basin showing the eco-climatic envelope as defined by Olson (2001) and the geographic partition (North vs South vs islands) tested in our study. Each black square represents a location sampled in the meta-analyzed raw studies. Histograms for latitudinal and longitudinal distributions of locations sampled in raw studies are given above and to the right of the map. The photographed tree in the upper right corner is a *Cedrus brevifolia* individual on a ridge in its natural habitat of the Trohodos mountains of Cyprus.



plants, positive and significant effects sizes were thus found for plastidial DNA and paternally-inherited DNA (gymnosperm plastidial DNA).

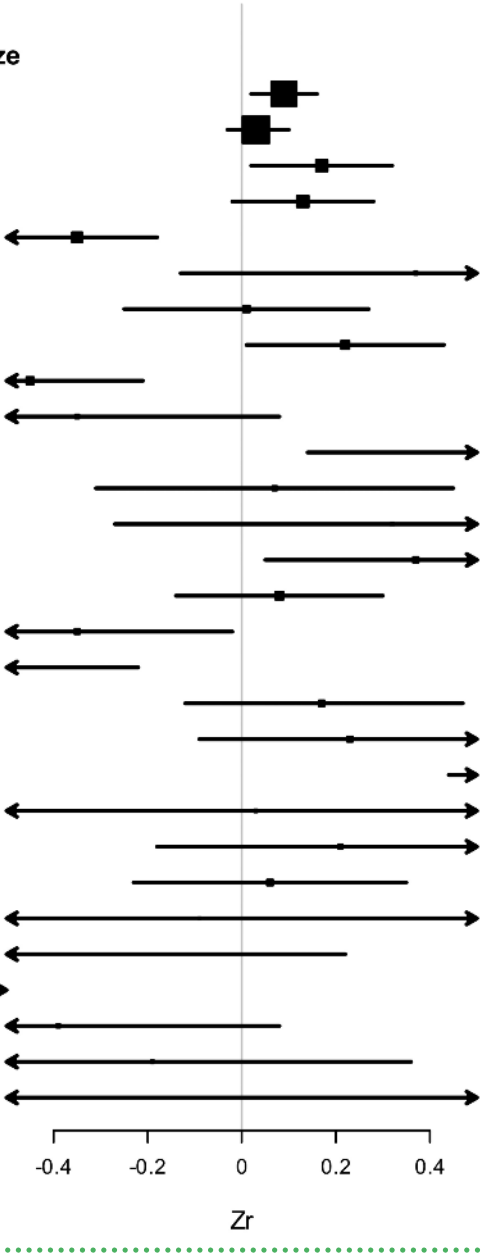
Ecological traits

We were able to assign 201 effect-sizes of plant species to a bioclimatic belt without ambiguity (Fig. 4b). The remaining species being found in two or more belts were labeled as 'eurytherm' species (64 effect-sizes). The true Mediterranean ecological group (meso-Mediterranean) had the highest positive *Zr* and was the only significant group. The supra-Mediterranean group had a positive but non-significant Zr. The category with the highest sample size (mountain-Mediterranean) had a negative and non-significant *Zr*, and the gymnosperms contributed predominantly to this group (70 of the 82 effect-sizes). The group with the highest requirements in terms of temperature (thermo-Mediterranean) had an almost null *Zr*. Heterogeneity was significant in all groups.

FIGURE 3

Mean effect-sizes across the Mediterranean basin for 12 plant and 17 animal families, arranged in decreasing effect size number per category. Out of 62 possible families represented in the dataset, only the 29 with a number of effect-sizes over 5 are represented here. Square and error bars should be interpreted as indicated for Fig. 2a.

Moderator	Number of effect-sizes	Sign. of Q-test	Effect-size
Pinaceae	105	***	0.09
Fagaceae	34	***	0.03
Cupressaceae	17	***	0.17
Oleaceae	14	***	0.13
Lamiaceae	12	*	-0.35
Fabaceae	9	NS	0.37
Ranunculaceae	8	**	0.01
Poaceae	7	NS	0.22
Pottiaceae	7	NS	-0.45
Myrtaceae	5	NS	-0.35
Asteraceae	5	*	0.77
Orchidaceae	5	NS	0.07
Brassicaceae	5	NS	0.32
Buthidae	8	NS	0.37
Scolytidae	8	NS	0.08
Nymphalidae	8	NS	-0.35
Plethodontidae	8	**	-0.72
Leporidae	8	*	0.17
Bacillaceae	6	.	0.23
Tephritidae	5	***	0.79
Lacertidae	5	NS	0.03
Bombinatoridae	5	NS	0.21
Ranidae	5	**	0.06
Canidae	5	NS	-0.09
Cervidae	5	NS	-1.02
Torymidae	5	***	-1.55
Vespertilionidae	5	NS	-0.39
Notodontidae	5	.	-0.19
Alytidae	5	.	-0.34



Biological traits

Phanerophytes (trees and shrubs) represented around 2/3 of the plants in our dataset and showed a positive and significant *Zr* (Fig. 4c). Among them, gymnosperms had a positive *Zr* (see above). Dicots among phanerophytes showed a contrasting pattern depending on the geographic envelope: at Mediterranean basin level, the effect was positive but non-significant; whereas it was strongly positive for the continental group (0.10 [0.04;0.16] 95% CI, *N* = 65) and strongly negative for islands (-0.56 [-0.71;-0.40] 95% CI, *N* = 17). Contrasting with the other moderator analyses, the southern group yielded a negative trend while the northern showed a positive one. Chamaephytes (low growing perennials with over-wintering buds below 50 cm) were the only group showing a negative (although non-significant) *Zr*.

Seed dispersal mode affected *Zr* values: Anemochorous (wind-dispersed, contributing to more than half of the effect sizes) and barochorous (gravity-dispersed) plants showed a significant and positive *Zr* while zoochorous (animal-dispersed) plants had a non-significant *Zr* (Fig. 4d). Pollen dispersal type also affected *Zr* values. Although calculated from a small sample size, the hydrogamous (water-dispersed pollen) plants showed a homogeneous negative trend (Fig. 4e) whereas entomogamous (insect-dispersed pollen) species had a non-significant *Zr* and anemogamous (wind-dispersed pollen) species had a positive and significant *Zr*.

Discussion

Organization of genetic diversity in Europe mostly follows latitudinal routes of recolonization dating from the Holocene (Petit *et al.* 2003). In the Mediterranean, although a latitudinal imprint exists, our analysis demonstrates the existence of an overall longitudinal imprint on genetic diversity. Using a meta-analysis on 143 plant and animal species, we found that overall within-population genetic diversity of plants and animals increases significantly from West to East in the Mediterranean basin, both in southern Europe and in North Africa, and for continental but not for island species. The longitudinal trend was not found in all taxonomic groups, however. This result broadens the evidence provided by Fady & Conord (2010) beyond tree species to include such taxonomic groups as arthropods, well-represented in the dataset, gymnosperms, and monocots. Other well-represented groups including dicots, mammals and amphibians do not follow the trend. Several poorly represented groups such as gastropods and bryophytes demonstrate a significant opposite trend, or no trend (reptiles and ferns).

What are the processes that shaped current genetic diversity longitudinally across the Mediterranean basin?

Trying to detect events / processes responsible for longitudinal imprints on genetic diversity in the Mediterranean may be challenging. The effects of older events such as divergence/diversification linked to vicariance during Pliocene (Blondel & Aronson, 1999; erection of geographical barriers) may coincide with that of younger events such as Last Glacial Maximum demographic bottlenecks, Holocene colonization events, admixture from secondary contact (Petit *et al.* 2003) or hybridization with closely related species (Papageorgiou *et al.* 2008).

However, it can be assumed that many current Mediterranean species have remained closer to their glacial refugia than their European counterparts and thus carry imprints (however attenuated) that commonly affected glacial refugia. A meta-analysis, using a fixed-effect model, precisely makes this assumption (Borenstein *et al.* 2009). It thus assumes that a common set of drivers affected *GDpop* across species along a longitudinal gradient in the Mediterranean basin. By analyzing the effect of moderators on summary-effect sizes, our goal was to discuss the most parsimonious explanations for such a trend.

Effect of past climate on resident populations or effect of recolonization on genetic diversity in the Mediterranean

Biological and life history traits:

We show that both plants and animals display a trend of increasing *GDpop* from west to east in the Mediterranean. We expected that, on average, the trend would be due more to migration effects (recolonization patterns) in mobile or high gene flow organisms, and more to local climatic effects in sessile or low gene flow organisms. With perhaps a difference in intensity between the two trends, that due to migration being less pronounced because of high gene flow and recurrent genetic exchange between populations blurring post recolonization foundation effects. Our data show that reality may be more complex because both highly mobile (insects) and more sessile (arachnids) animals showed trends similar in direction and magnitude. In arachnids, a group only comprising scorpions from the southern part of the Mediterranean basin in our dataset, the effect could be related to local past climate effects. The pattern shown by insects could reveal a link between their contemporary genetic structure and the mirrored structure of the plants they exploit. Recent studies have indeed illustrated the link between levels of diversity in keystone organisms such as trees and in their phytophagous associated organisms (Crutsinger *et al.* 2006, Whitham *et al.* 2006). The relationship has also been shown to hold true from local to region-wide scales (Bangert *et al.* 2006). Our work is certainly the first for trees and insects, indicating a clear common pattern of congruent *GDpop* at continental scale.

The absence of significant trend in mobile vertebrate groups such as birds, amphibians and mammals could reflect the coexistence of multiple Holocene recolonization routes from multiple refugia among species. For example,

in two species of rodents from the genus *Apodemus*, LGM survival had two very different outcomes, with *A. flavicollis* disappearing from the Iberian Peninsula whereas *A. sylvaticus* survived only there. The subsequent Holocene recolonization of Europe by these two currently sympatric species left two diverging imprints on genetic diversity (Michaux *et al.* 2005). In yet another species of rodent, the shrew *Crocidura russula*, refugial populations were located in North Africa from which western European Holocene populations derive (Cosson *et al.* 2005). As for the bank vole, *Clethrionomys glareolus*, most of its Holocene European range was recolonized from central European refugia although Mediterranean refugia existed (Deffontaine *et al.* 2005). In this group of mammals, the exception may be the rule in terms of LGM survival and Holocene recolonization, which is indicated by a non-significant trend of genetic diversity in our meta-analysis.

The only plant type (sensu Raunkiaer) displaying a negative trend in *GDpop* was the chameaphytes. Because of their ground level overwintering buds, chamaephytes are better suited to resist cold snowy winters than dry winters (Taulavuori *et al.* 2011). Snow may in fact be beneficial to their overwintering, thus potentially keeping larger populations during the LGM in the western than in the eastern Mediterranean.

The positive trend found in phanerophytes (*i.e.* woody plants having buds at least 50cm above the ground) overlaps with that of wind-dispersed species and may reflect both local demographic effects under *LGM* climate and long distance Holocene recolonization. Seed dispersal by gravity, however, does not allow for rapid and long-range dispersal. Thus, the positive trend found in gravity dispersed plant species matches the expectation of an effect of local *LGM* climate on genetic diversity. Conversely, animal-dispersed species depend on their dispersers' behavior for survival (Scofield *et al.* 2010). Their lack of trend in genetic diversity may mirror the diversity of vertebrate Holocene recolonization routes highlighted above for mammals.

Ecological requirements in plant taxa:
When assessing the impact of ecological (temperature) requirement on plants, we expected to find an increasing positive effect on *GDpop* from low to high elevation (from less to more cold tolerant) plant species. The rationale for this expectation was that the unfavorably cold *LGM* climate should affect more strongly population size in species with higher temperature requirements as they became trapped in reduced size habitats compared to those of lower temperature requirement species. In contrast, species adapted to colder climates such supra- and mountain Mediterranean species for example, benefiting from larger habitats during the *LGM*, should not have suffered demographic bottlenecks as did the more truly Mediterranean group. Thus, a west to east oriented clinal climate at the *LGM* should result in a stronger west to east clinal *GDpop* structure in true Mediterranean species ('Me' in Fig. 4b) than for species with other ecological requirements. The strength of the effects followed our expectations, with true Mediterranean plants showing a stronger cline of west-east increasing *GDpop* than Supra- and Mountain-Mediterranean plants ('Su' and 'Mt' in Fig. 4b). The group of plants with no precise thermal requirement showed no significant cline. The non-significant slightly negative summary-effect of the 'Th' group was more surprising. We expected this group of warm climate species to have been impacted even more strongly than other categories by cold climate at the *LGM* whether through strong bottlenecks or extirpation and subsequent migration / recolonization. Several explanations are possible. '*Th*' species may have survived locally without stronger loss in *GDpop* in the west as compared with the east, although we have no evidence to support this explanation. '*Th*' species could also have recolonized from glacial refugia not situated in the eastern Mediterranean, which was demonstrated for *Cistus ladanifer* (which came back into Spain from North Africa via the Strait of Gibraltar) by Guzmán & Vargas (2009). Finally, the thermo-Mediterranean belt is known to have endured severe human impact throughout the millennia, possibly obscuring clinal climatic and recolonization effects on *GDpop*. Interestingly, the two species generating most of the effect sizes in this group were species strongly impacted by humans (*Pinus pinea, Olea europaea*), both being valuable food crop.

Marker and metric types:
Studies using maternally-inherited markers are most of the time designed to detect phylogeographic signals and capture differentiation effects, e.g. those due to the imprints of contraction and recolonization to and from Pleistocene refugia. Studies using paternally-inherited (plastidial DNA in gymnosperms) or bi-parentally inherited markers are more often designed to detect local demographic signals (for example for conservation planning) resulting from current environmental drivers, in addition to phylogeographic signals. This confirms that the overall trend detect in our study is a within population demographic bottleneck effect, more weakly detected in maternally than paternally and bi-parentally inherited genomes. Phylogeographic studies have shown that the two different metric types we have used here, richness and equitability, can be negatively correlated along the distribution range of species (Comps *et al.* 2001; Petit *et al.* 2003). In Europe, from refugia to the newly colonized areas, heterozygosity increased (merging of recolonization routes that originated from different refugia) whereas allelic richness decreased (founder effects along the recolonization route) with local diversity peaks in suture zones. We found no such significant difference in our analysis. The two metric types had globally congruently positive and significant summary-effects. Taken together, the global positive effects we measured for marker and

metric types may indicate a stronger role of local climate over recolonization in shaping the genetic diversity of Mediterranean populations.

Biogeographic effects: South vs North and continents vs islands

The Mediterranean basin has a highly heterogeneous and fragmented geography (Blondel & Aronson 1999). Its different geographical compartments have likely experienced different past ecological conditions and evolutionary histories. The northern Mediterranean flora and fauna contain predominantly Nordic, Asian and local elements whereas the southern Mediterranean is predominantly made of Tropical and local elements (Quézel & Médail 2003). Because of its peninsulas, migration may have been more restricted in the northern Mediterranean than in the southern Mediterranean. Also, one might expect stronger demographic bottlenecks and stronger scale and size effects on islands than on the continent. Although the number of populations originating from the southern Mediterranean was eight times lower than in the north, its summary-effect was more strongly positive than that of northern Mediterranean populations. Also, in the southern Mediterranean, the equitability metric type was more strongly positive than the richness metric type (which was actually non-significant). Although the gradient was more restricted in its longitudinal span (mostly but not entirely limited to populations of the Maghreb) in the southern than in the northern Mediterranean, these results suggest that factors linked to local *LGM* climate may have more strongly affected genetic diversity in the Southern Mediterranean than those linked to recolonization.

The weakly positive summary-effect of Mediterranean islands is in sharp contrast with that of continents. It may reflect a Mediterranean 'insularity syndrome' globally independent of climatic factors and more likely linked to high endemism (Médail & Diadema 2009) and/or early human impact (Vigne *et al.* 2009).

Conclusions

Our population genetic diversity dataset covered an extensive range of animal and plant species. It also had the advantage of gathering data from species with strong economic importance (for example fruit trees and medicinal plants), strong ecological importance (for example forest trees) as well as endangered flagship species (for example butterflies and endemic plants). The taxonomic groups that were the most heavily sampled (gymnosperms for plants and arthropods for animals) showed a congruently positive summary-effect, *i.e.* an increasing genetic diversity from west to east. These abundant taxonomic groups are thus good models for detecting trends and patterns affecting biodiversity in general. The propensity of the genetic diversity of trees to be an excellent testimony of the imprint of ancient evolutionary and demographic processes (given their long generation time) has been mentioned for some time (Petit *et al.* 2003; Petit & Hampe 2006) and it seems that arthropods can be added to this category.

Longitude is a strong structural element of biodiversity at gene level in the Mediterranean, both in southern Europe and in North Africa. Continental / oceanic longitudinal type gradients also exist in many parts of the world. The generally overlooked role of longitude in shaping species ranges and genetic diversity deserves stronger focus (Stewart *et al.* 2010). Taken together, our results suggest that, on top of a genetic structure inherited from the existence of glacial refugia (which phylogeography is increasingly demonstrating as being very complex, Leppanen *et al.* 2011), local climate during the *LGM* durably affected the demography of resident populations in the Mediterranean, observable as a weak but highly significant longitudinal cline of genetic diversity.

For conserving and sustainably managing biodiversity, global or region-wide assessments are needed beyond the idiosyncrasy of single species or single taxonomic groups to detect trends and large scale patterns. Meta-analyses, by making it possible to compare already available data acquired within unrelated studies, provide an interesting framework for these assessments. Already successfully used in ecology to test theoretical predictions (*e.g.* Rapoport's law predicting an increase of species range with latitude, Ruggiero & Werenkraut 2007), we have shown that meta-analyses can also be powerful to test the determinants of large-scale biodiversity patterns.

Finally, our findings can now be compared with other measurements of past, current and expected biodiversity (species and functional traits for example) and their congruence tested (Devictor *et al.* 2010), provided that appropriate databases (species, ecosystems, past and present climate) exist or can be constructed at relevant scales. Data available for estimating biodiversity at gene level remain critically insufficient in the Mediterranean. For example the Mediterranean comprises approximately 200 mammal species and more than 300 bird species (Blondel & Aronson, 1999) whereas our dataset only included 10 mammal and 3 bird species! This remains a major challenge in the poorly politically structured Mediterranean, but also in other regions of the world where biodiversity is high and rapidly declining.

Acknowledgements

This research was made possible by financial support from the European Network of Excellence EVOLTREE (EVOLution of TREEs as drivers of terrestrial biodiversity) and from INRA Grant P-EFPA6 to Cyrille Conord. We would like to thank D. Betored and C. Pichot for help with database management and GIS analyses, and Drs F. Balfourier, D. Gömöry, C. Kerdelhué, R.J. Petit & G.G. Vendramin for kindly providing raw data.

REFERENCES

Aloe AM, Becker BJ (2009) Teacher verbal ability and school outcomes: Where is the evidence°? *Educational Researcher*, **38**, 612-624.

Aravanopoulos FA, Panetsos KP, Skaltsoyiannes A (2004) Genetic structure of Pinus brutia stands exposed to wild fires. *Plant Ecology*, **171**, 175-183.

Atkinson RJ, Rokas A, Stone GN (2007) Longitudinal patterns in species richness and genetic diversity in European oaks and oak gallwasps. In S. Weiss and N. Ferrand editors "*Phylogeography of southern European refugia*", pp. 127–151. Springer, Dordrecht, The Netherlands.

Avise JC (2000). *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA, USA. (447 pp.)

Bangert RK; Allan GJ, Turek RJ, Wimp GM, Meneses N, Martinsen GD, Keim P, Whitham TG (2006) From genes to geography: a genetic similarity rule for arthropod community structure at multiple geographic scales. *Molecular Ecology*, **15**, 4215-4228.

Blondel J, Aronson J (1999). *Biology and wildlife of the Mediterranean region*. Oxford University Press, New York.

Borenstein M, Hedges LV, Higgins JPT, Rothstein H, 2009. *Introduction to Meta-Analysis*. John Wiley and Sons.

Brewer S, Cheddadi R, de Beaulieu JL, Reille M (2002). The spread of deciduous Quercus throughout Europe since the last glacial period. *Forest Ecology & Management*, **156**, 27-48.

Crutsinger GM, Collins MD, Fordyce JA, Gompert Z, Nice CC, Sanders NJ (2006) Plant genotypic diversity predicts community structure and governs an ecosystem process. *Science*, **313**, 966-968

Cheddadi R, Fady B, François L, Hajar L, Suc JP, Huang K, Demarteau M, Vendramin GG (2009) Putative glacial refugia of *Cedrus atlantica* from Quaternary pollen records and modern genetic diversity. *Journal of Biogeography*, **36**, 1361–1371.

Coart E, Van Glabeke S, Petit RJ, Van Bockstaele E, Roldan-Ruiz I (2005) Range wide versus local patterns of genetic diversity in hornbeam (*Carpinus betulus L.*). *Conservation Genetics*, **6**, 259-273.

Comps B, Gomory D. Letouzey J, Thiebaut B, Petit RJ (2001). Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics*, **157**, 389-397

Convention of Biological Diversity (1992). United Nations, New York, USA. http://www.cbd.int/doc/legal/cbd-en.pdf

Cosson JF, Hutterer R, Libois R, *et al.* (2005). Phylogeographical footprints of the Strait of Gibraltar and Quaternary climatic fluctuations in the western Mediterranean: a case study with the greater white-toothed shrew, *Crocidura russula* (Mammalia: Soricidae). *Molecular Ecology*, **14**, 1151-1162.

Deffontaine V, Libois R, Kotlik P, Sommer R, Nieberding C, Paradis E, Searle JB, Michaux J (2005). Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole *(Clethrionomys glareolus). Molecular Ecology*, **14**, 1727-1739.

Del Re AC, Hoyt WT (2010). MAc: Meta-Analysis with correlations. R package version 1.1, available at http://cran.r-project.org/web/packages/MAc/index.html

Devictor V, Mouillot D, Meynard C, Jiguet F, Thuiller W, Mouquet N (2010) Spatial mismatch and congruence between taxonomic, phylogenetic and functional diversity: the need for integrative conservation strategies in a changing world. *Ecology Letters*, **13**, 1030-1040.

Fady B (2005) Is there really more biodiversity in Mediterranean forest ecosystems? *Taxon*, **54**, 905-910.

Fady B, Conord C (2010) Macroecological patterns of species and genetic diversity in vascular plants of the Mediterranean Basin. *Diversity & Distributions*, **16**, 53-64.

Grivet D, Sebastiani F, Gonzalez-Martinez SC, Vendramin GG. (2009) Patterns of polymorphism resulting from long-range colonization in the Mediterranean conifer Aleppo pine. *New Phytologist*, **184**, 1016-1028.

Guzmán B, Vargas P (2009) Long-distance colonization of the Western Mediterranean by *Cistus ladanifer* (Cistaceae) despite the absence of special dispersal mechanisms *Journal of Biogeography*, **36**, 954-968.

Hartung J, Knapp G, Sinha BK (2008) *Statistical meta-analysis with applications*. John Wiley, Sons.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87-112.

Jump AS, Matyas C, Penuelas J (2009) The altitude-for-latitude disparity in the range retractions of woody species. *Trends in Ecology and Evolution*, **24**, 694-701

Kadereit JW, Arafeh R, Somogyi G, Westberg E (2005) Terrestrial growth and marine dispersal? Comparative phylogeography of five coastal plant species at a European scale. *Taxon*, **54**, 861-876.

Keef SP, Roberts LA (2004) The meta-analysis of partial effect sizes. *British Journal of Mathematical and Statistical Psychology*, **57**, 97–129.

Le Corre V, Kremer A (2003) Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics*, **164**, 1205-1219.

Leppanen J, Vepsalainen K, Savolainen R (2011) Phylogeography of the ant *Myrmica rubra* and its inquiline social parasite. *Ecology and evolution*, **1**, 46-62.

Liepelt S, Cheddadi R, de Beaulieu JL, Fady B, Gömöry D, Hussendörfer E, Konnert M, Litt T, Longauer R, Terhürne-Berson R, Ziegenhagen B (2009) Postglacial range expansion and its genetic imprints in *Abies alba* (Mill.) - a synthesis from paleobotanic and genetic data. *Review of Palaeobotany and Palynology*, **153**, 139-149.

Médail F, Quézel P (1997) Hot-spots analysis for conservation of plant biodiversity in the Mediterranean basin. *Annals of the Missouri Botanical garden*, **84**, 112-127.

Médail F, Diadema K (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, **36**, 1333-1345.

Michaux JR, Libois R, Filippucci MG (2005). So close and so different: comparative phylogeography of two small mammal species, the Yellow-necked fieldmouse (*Apodemus flavicollis*) and the Woodmouse (*Apodemus sylvaticus*) in the Western Palearctic region. *Heredity*, **94**, 52-63.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853-858

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'Amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ. Allnutt TF, Ricketts TH. Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR (2001) Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, **51**, 933-938.

Papageorgiou AC, Vidalis A, Gailing O, Tsiripidis I, Hatziskakis S, Boutsios S, Galatsidas S, Finkeldey R (2008) Genetic variation of beech (*Fagus sylvatica L.*) in Rodopi (N.E. Greece). *European Journal of Forest Research*, **127**, 81-88.

Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Muller-Starck GM, Demesure-Musch B, Palme A, Martin JP, Rendell S, Vendramin GG (2003) Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563-1565.

Petit RJ, Hampe A (2006). Some evolutionary consequences of being a tree. *Annual Review of Ecology Evoluation and Systematics* **37**, 187-214.

Petit RJ, Vendramin GG (2007) Plant phylogeography based on organelle genes: an introduction. In S. Weiss and N. Ferrand editors "*Phylogeography of southern European refugia*", Springer, Dordrecht, The Netherlands, p 23-97.

Pignatti S (1982) *Flora d'Italia*. Edagricole, Bologna, Italy.

Quézel P, Médail F (2003) *Ecologie et biogéographie des forêts du bassin méditerranéen*. Elsevier, Paris.

Rameau JC, Mansion D, Dumé G, Gauberville C (2008) *Flore forestière française Tome 3*: Région méditerranéenne. IDF, Paris.

Richardson DM, Whittaker RJ (2010) Conservation biogeography - foundations, concepts and challenges. *Diversity & Distributions*, **16**, 313-320.

Riddle BR; Hafner DJ; Alexander LF, JR Jaeger (2000). Cryptic vicariance in the historical assembly of a Baja California Peninsular Desert biota. *Proceedings of the National Academy of Sciences*, **97**, 14438-14443.

Rokas A, Atkinson RJ, Webster L, Csoka G, Stone GN (2003) Out of Anatolia: longitudinal gradients in genetic diversity support an eastern origin for a circum-Mediterranean oak gallwasp *Andricus quercustozae*. *Molecular Ecology*, **12**, 2153-2174.

Rosenberg MS, Adams DC, J Gurevitch (2000) *MetaWin: statistical software for meta-analysis*. Version 2.0. Sinauer Associates, Sunderland, MA, USA.

Rosenthal R, DiMatteo MR, (2001) Meta-analysis: recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, **52**, 59-82.

Ruggiero A, Werenkraut V (2007) One-dimensional analyses of Rapoport's rule reviewed through meta-analysis. *Global Ecology, Biogeography*, **16**, 401-414.

Sánchez-Goñi MF, Cacho I, Turon JL, Guiot J, Sierro FJ, Peypouquet JP, Grimalt JO, Shackleton NJ (2002) Synchroneity between marine and terrestrial responses to millennial scale climatic variability during the last glacial period in the Mediterranean region. *Climate Dynamics*, **19**, 95–105.

Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution*, **22**, 11–16.

Scofield DG, Sork VL, Smouse PE (2010). Influence of acorn woodpecker social behaviour on transport of coast live oak (*Quercus agrifolia*) acorns in a southern California oak savanna. *Journal of Ecology*, **98**, 561-572.

Soltis DE, Morris AB, Lachlan JSM, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261-4293.

Soto A, Robledo-Arnuncio JJ, Gonzalez-Martinez SC, Smouse PE, Alia R (2010) Climatic niche and neutral genetic diversity of the six Iberian pine species: a retrospective and prospective view. *Molecular Ecology*, **19**, 1396-1409.

Soranzo N, Alía R, Provan J, Powell W (2000) Patterns of variation at a mitochondrial sequence-tagged-site locus provides new insights into the postglacial history of European Pinus sylvestris populations. *Molecular Ecology*, **9**, 1205-1211.

Stewart JR, Lister AM, Barnes I, Dalen L (2010). Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B*, **277**, 661–671.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453-464.

Taulavuori K, Bauer E, Taulavuori E (2011). Overwintering stress of *Vaccinium vitis-idaea* in the absence of snow cover. *Environmental and experimental botany*, **72**, 397-403.

Van Andel, TH (2002) The climate and landscape of middle part of Weichselian glaciation in Europe: the stage 3 project. *Quaternary Research*, **57**, 2-8.

Vigne JD, Zazzo A, Saliège JF, Poplin F, Guilaine J, Simmons A (2009) Pre-Neolithic wild boar management and introduction to Cyprus more than 11,400 years ago. *Proceedings of the National Academy of Sciences*, **106**, 16135-16138.

Whitham TG, Bailey JK, Schweitzer JA, Shuster SM, Bangert RK, Leroy CJ, Lonsdorf EV, Allan GJ, DiFazio SP, Potts BM, Fischer DG, Gehring CA, Lindroth RL, Marks JC, Hart SC, Wimp GM, Wooley SC (2006) A framework for community and ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics*, **7**, 510-523.

Whittaker RJ, Araujo MB, Paul J, Ladle RJ, Watson JEM, Willis KJ (2005) Conservation biogeography: assessment and prospect. *Diversity & Distributions*, **11**, 3–23.

Wickham H (2011) The Split-Apply-Combine Strategy for Data Analysis. *Journal of Statistical Software*, **40**, 1-29.

Wu HB, Guiot JL, Brewer S, Guo ZT (2007) Climatic changes in Eurasia and Africa at the last glacial maximum and mid-Holocene: reconstruction from pollen data using inverse vegetation modelling. *Climate Dynamics*, **29**, 211-229.

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in ecology and evolution*, **11**, 413-418.

EFFECT OF POPLAR GENOTYPES ON MYCORRHIZAL INFECTION AND SECRETED ENZYME ACTIVITIES IN MYCORRHIZAL AND NON-MYCORRHIZAL ROOTS

Courty PE[1,3*], Labbé J[1*], Kohler A[1], Marçais B[1], Bastien C[2], Churin JL[1], Garbaye J[1], Le Tacon F[1]

[1] UMR 1136 INRA-Nancy Université, Interactions Arbres / Microorganisms, INRA-Nancy, 54280 Champenoux, France.
[2] INRA Orléans, Unité Amélioration, Génétique et Physiologie forestières, Ardon, BP 20619, 45166 Olivet Cedex, France.
[3] Present address: University of Fribourg, Department of Biology, Rue Albert Gockel 3, CH-1700 Fribourg.

* These authors contributed equally to this work.

Corresponding Author:
Courty Pierre-Emmanuel
pierre-emmanuel.courty@unifr.ch

Poplar genotypes and secreted enzyme activities in mycorrhizal and non-mycorrhizal roots

The impact of ectomycorrhiza formation on the secretion of exoenzymes by the host plant and the symbiont is unknown. Thirty-eight F1 individuals from an interspecific *Populus deltoides* (Bartr.) x *Populus trichocarpa* (Torr. & A. Gray) controlled cross, were inoculated with the ectomycorrhizal fungus *Laccaria bicolor*. The colonization of poplar roots by *L. bicolor* dramatically modified their ability to secrete enzymes involved in organic matter breakdown or organic phosphorus mobilization, such as N-acetylhexosaminidase, glucuronidase, cellobiohydrolase, glucosidase, xylosidase, laccase and acid phosphatase. The expression of genes coding for laccase, N-acetylhexosaminidase and acid phosphatase was studied in mycorrhizal and non-mycorrhizal root tips. Depending on the genes, their expression was regulated upon symbiosis development. Moreover, it appears that poplar laccase or phosphatase are poorly contributing to ECM metabolic activity. Enzymes secreted by poplar roots were added or substituted to enzymes secreted by *L. bicolor*. The enzymatic activities expressed in mycorrhizal roots differed significantly between the two parents, while it did not differ in non-mycorrhizal roots. Significant differences were found between poplar genotypes for all enzymatic activities measured on ectomycorrhizas except for laccase activity. On the contrary, no significant differences were found between poplar genotypes for enzymatic activities of non-mycorrhizal root tips except for acid phosphatase activity. The level of enzymes secreted by the ectomycorrhizal root tips is under the genetic control of the host. Moreover, poplar heterosis was expressed through the enzymatic activities of the fungal partner.

Key words: Poplar, *Laccaria bicolor*, secreted enzymes, heterosis, heritability, host genetic control.

Published in *Journal of Experimental Botany* (2011) 62: 249-260, with doi:10.1093/jxb/erq274

Introduction

The fine roots of tree species in temperate and boreal forests are symbiotically associated with fungi, forming a composite organ called ectomycorrhiza (ECM) (Smith and Read, 2008). The establishment and the functioning of ECM lead to complex morphological and physiological changes in both the plant and the fungus (Martin and Nehls, 2009, Courty *et al*. 2010a). The ECM symbiosis has been described as a mutualistic association where the autotrophic plant supplies photosynthates to the heterotrophic fungus, which in turn supplies water and nutrients to the host (Smith and Read, 2008). Several studies also have shown that ectomycorrhizal fungi (ECMf) are able to produce extracellular enzymes, such as proteases, involved in the direct mobilization of nutrients from organic substrates (Courty *et al*. 2005, 2006, 2010b; Lindahl *et al.* 2005; Koide *et al.* 2008). In addition, a given species may contribute to significant functional variations through metabolic activities (Buée *et al.* 2007; Courty *et al.* 2010b).

The ecological fitness and the metabolic activity of ECMf depend on their genotypes, environmental factors (van der Heijden and Sanders 2002; Smith and Read, 2008), host plant genotypes (Barker *et al.* 2002; Linderman and Davis, 2004), and the interactions between all these factors (Khasa *et al.* 2002; Gehring *et al.* 2006; Karst *et al.* 2009). Recent studies also suggest that host plant genome may play a role in determining the dominant mycorrhizal type in dually colonized hosts (van der Heijden & Kuyper, 2001; Khasa *et al.* 2002). However, no studies have simultaneously examined the effect of host plant genotypes and the metabolic activity of one ECMf species in controlled conditions. In the *Laccaria bicolor*/poplar ECM symbiosis, Tagu *et al.* (2005) have

shown that the host genotype impacts on root colonization by the fungus. The heritability of mycorrhizal colonization of poplar was also studied (Tagu *et al.* 2001, 2005). However, the metabolic activity of one ECM fungal genotype colonizing different genotypes of the same host species was never studied. In our study, the use of poplar as host tree model was motivated by the availability of large genetic and genomic resources for this species (Brunner *et al.* 2004; Tuskan *et al.* 2006). Moreover, the study of heritability and variability of physiological parameters (i.e. water use efficiency, dry weight, Leaf Maximum Area) at family level were intensively studied in poplar (Marron *et al.* 2005; Dillen *et al.* 2007).

In this study, we selected as relevant functional traits, a standard set of seven enzymatic activities routinely used in field studies (Courty *et al.* 2005). The enzyme activity of a secreted laccase, an oxidative enzyme involved in the degradation of recalcitrant plant residues, such as lignin, five secreted glycosyl hydrolases (cellobiohydrolase, β-glucosidase, β-xylosidase, β-glucuronidase, N-acetylglucosaminidase) acting on polysaccharides and a phosphomonoesterase involved in the mobilisation of phosphorus from soil organic matter were assessed. *L. bicolor* has a low set of glycosyl hydrolases able to hydrolyse plant cell wall polysaccharides (Martin *et al.* 2008). However, its genome encodes several carbohydrate-active enzymes able to degrade bacterial, fungal and animal polysaccharides (Martin *et al.* 2008).

The impact of the host genotype on the ECM metabolic activity is unknown. Here, the responding functional trait in focus is the capacity to produce secreted or cell wall bound enzymes. The first objective was to determine whether the enzymatic activities expressed in mycorrhizal roots differed significantly between two parents, *P. deltoides* and *P. trichocarpa*, and different poplar hybrid genotypes (*P. deltoides X P. trichocarpa*). The second objective was to determine the effect of host genotypes on fungal traits by measuring the heritability of enzymatic activities in mycorrhizal and non-mycorrhizal root tips and by assessing for these traits a possible heterosis among the progeny.

Materials and Methods

Plant material, strain and culture conditions

Poplar material consisted of 38 F1 individuals from an interspecific *P. deltoides* (female clone from Illinois, no. 73028-62) and *P. trichocarpa* (male clone from Washington, no. 101-74) controlled cross (family 54B) (Tagu *et al.* 2001, 2005). We have tested the ability of the two parents and the 38 breeds to form mycorrhizas by inoculating them with *Laccaria bicolor* S238N (Di Battista *et al.* 1996; Tagu *et al.* 2001). The 38 F1 genotypes were chosen at random among the 336 genotypes used for the construction of a genetic map (Cervera *et al.* 2001; Jorge *et al.* 2005). The *L. bicolor* S238N fungal strain, coming from the INRA-Nancy collection of ECMf, was maintained on Pachlewski's. This model fungal strain was chosen for its ability to form ECMs with poplar and for the avaibility of genomic ressources (Tagu *et al.* 2001; Martin *et al.* 2008). The inoculum of *L. bicolor* S238N was prepared by aseptically growing the mycelium in a peat-vermiculite nutrient mix in glass jars for 2 months in the dark at 25 °C, and kept at 4° C during 2 months before use (Le Tacon and Bouchard, 1986).

Inoculation

Cuttings of one internode of each of the 38 poplar progenies and the two parents were rooted and individually inoculated at the same time, in 1-l pots containing a mixture of fungal inoculum (1:9 vol/vol) and calcinated attapulgite (Oil Dri US Special) during twelve weeks, in a greenhouse during spring with day-night temperatures of 28 and 15°C, respectively. Plants were watered during the whole experiment until measurements. From the second month, a low N, low P nutrient solution was applied weekly (Frey-Klett *et al.* 1997). In order to control environmental heterogeneity of the greenhouse, 8 replicates were done for each poplar genotype and were randomly distributed in 8 blocks. Each block contained one pot of each 38 progenies and the two parents.

Root colonization

Entire root systems (except roots present 1 cm depth from cal) were carefully washed under tap water and cut into approximately 1-cm pieces. For each root system, 100 randomly selected root tips were examined and assessed as mycorrhizal or non-mycorrhizal under a stereomicroscope (magnification x40) for calculating ECM percentages.

Chlorophyll content, leaf morphological measurement and dry weight

Before harvesting plants, chlorophylls *a* and *b* content was measured with a Minolta SPAD chlorophyll meter (Minolta Corp., Ramsey, N.J.). Three SPAD measurements were done on three leaves of each plant and then averaged (Monje and Bugbee, 1992). To convert SPAD measures into chlorophyll content, a standard curve was built by extracting chlorophylls with the dimethyl sulphoxide (DMSO) extraction technique (Monje and Bugbee, 1992; Richardson *et al.* 2002). Total leaf chlorophyll concentration (mg cm-2) of the extracts was calculated from this equation: $0.0202A_{645} + 0.00802A_{663}$. SPAD measurements were then converted to chlorophyll content using a third order polynomial equation: $-0.0064SPAD^3 + 0.5895SPAD^2 + 2.0891SPAD + 10.024$.

Once mycorrhizal infection had been determined, leaves, stems and roots were separated. The leaves were placed in plastic bags and kept at 4°C until leaf morphological measurements were completed. The leaf area (cm^2) of all leaves of each plantlet was measured by using a LI-COR 3100 (Li-Cor Inc., Lincoln, NE, USA). Then, leaves, stems and roots were dried at 70 °C for 1 week (Mettler, Toledo

balance). The leaf mass area (LMA) was calculated for each clone using the relationship between the area of each leaf and its corresponding dry weight.

Enzymatic activity profiling of ectomycorrhizal and non-mycorrhizal root tip

One mycorrhizal root tip and one non-mycorrhizal root tip were collected from each of the 320 cuttings in order to determine their potential enzymatic activities, using the high-throughput photometric and fluorimetric microplate assays described and detailed in Courty *et al.* (2005), and applied in previous studies (Buée *et al.* 2007; Courty *et al.* 2010b). As the variability of enzyme activities among ECM tips within a root system is low, one tip is sufficient to get a representative value (Courty *et al.* 2005). Each well of the 96-well micro-titration plate contained either one ectomycorrhizal root tip or one non-ectomycorrhizal root tip. Seven activities were successively measured on root tips: xylosidase (EC 3.2.1.37), glucuronidase (EC 3.2.1.31), cellobiohydrolase (EC 3.2.1.91), *N*-acetylglucosaminidase (EC 3.2.1.14), ß-glucosidase (EC 3.2.1.3), acid phosphatase (EC 3.1.3.2), and laccase (EC 1.10.3.2) activities. The enzymes activities were expressed as pmol mm^{-2} min^{-1} of developed surface area of root tips. The developed surface area of the root tips was measured after scanning and image analysis using the Mac/Win Rhizo software (Regent Instruments, Quebec City, Canada). They correspond to the activities of enzymes present on the surface of the roots or mycorrhiza mantles and released in the medium during the incubation.

Whole-genome expression oligoarray analyses

Genes coding for laccase, N-acetylglucosaminidase and acid phosphatase were known and characterized in the genome of *L. bicolor* and *P. trichocarpa*. As the genes involved in xylosidase, glucuronidase, cellobiohydrolase and ß-glucosidase activity were not characterized, we were not able to measure the corresponding transcript expression. Accumulation of predicted laccase (*Lac*), N-acetylglucosaminidase (*Nag*) and acid phosphatase (*Pap*) transcripts was detected in free-living mycelium of *L. bicolor* S238N, and in ectomycorrhizal and non-mycorrhizal root tips of poplar using NimbleGen *L. bicolor* whole-genome expression oligoarray v2 (Martin *et al.* 2008) and NimbleGen *P. trichocarpa* whole-genome expression oligoarray (Tuskan *et al.* 2006). Data are available at the GEO platform GPL2699. The *L. bicolor* 4-plex whole genome expression array contained 18,653 gene models with three oligonucleotide probes for each gene model. For 4,702 gene models, technical duplicates were included on the oligoarray (A. Kohler & F. Martin, unpublished results). Average expression levels were calculated for each gene from the independent probes and were used for further analysis. To estimate the signal background and the resulting threshold value for significant expression, the mean intensity of 2,032 random probes present on the microarray was calculated. Gene models with expression exceeding the threshold by three or more were considered to be transcribed. Raw array data were filtered for non-specific probes and renormalized using ARRAYSTAR software (DNASTAR). Three biological replicates were used. Therefore, the reported gene expression values corresponded to the mean intensity of hybridization signals obtained for the specific oligonucleotide probes. A student t-test with FDR (Benjamini-Hochberg) multiple testing corrections was applied on the data ($P<0.05$), using ARRAYSTAR sofware (DNASTAR).

Statistical analysis

The percentage of mycorrhizal colonization was transformed by arcsin $\sqrt{X/100}$ function prior to variance analysis (ANOVA). Xylosidase, glucuronidase, cellobiohydrolase, chitinase, ß-glucosidase, acid phosphatase and laccase activities, root, shoot and stem dry weight and LMA were also submitted to ANOVA. The following mixed linear model was applied on an individual basis to detect significant differences among the clones:

$Y_{ijk} = \mu + B_i + G_j + \varepsilon_{ijk}$

where μ is the overall mean, *B* is the block effect (fixed), *G* is the genotype effect (random), and ε is the random residual error.

Restricted maximum likelihood estimates of genetic, block and residual variance components (σ^2G, σ^2B and $\sigma^2\varepsilon$) were computed, and for each trait, individual broad sense heritability (h^2) was estimated as follows:

$h^2 = \sigma^2G / (\sigma^2G + \sigma^2\varepsilon / n)$ where n is the average number of replicates per genotype. Standard deviations (SD) were derived from classic estimation of SD for a ratio x/y where $x = \sigma^2G$ and $y = \sigma^2G + \sigma^2\varepsilon / n$.

All analyses were performed with the statistical programs JMP 5.0 (SAS Institute Inc., Cary, NC, USA) and R version 1.8.0 (R Development Core Team, 2006, www.R-project.org).

The genetic coefficient of variation (CV_G) was used (Cornelius, 1994) to compare the relative amounts of genetic variation of traits with different means:

$CVG = \sqrt{(\sigma^2G / \mu)}$

Relationships between the different traits were also analysed by Pearson linear correlations.

Developed projected area of mycorrhizal and mycorrhizal root tips were compared between genotypes by ANOVA.

Results

A total of 320 plants were harvested and studied in this experiment. Three dead plants were not used in the analysis. No significant block effect was found for any measured traits.

Poplar ecophysiological traits

Significant differences ($p<0.001$) were found between plant genotypes for all measured traits (chlorophyll content, leaf maximum area, stem and root dry weight).

Significant differences were found between the parents for all measured traits except for leaf maximum area.

Effect of poplar genotype on root colonization by L. bicolor

Twelve weeks after inoculation, progenies and parental clones were only colonized by *L. bicolor*. No other contaminant ectomycorrhizal fungi were found on roots. The two parental genotypes significantly differed in their mycorrhizal development, *P. trichocarpa* exhibiting a rate of colonization of 40 % ± 8, and P. *deltoides* a rate of 16 % ± 4. The percentages of colonization of the different genotypes (progenies) varied from 12 % ± 8 to 64 % ± 6, with an average of 31 % (Fig. 1). The variance analysis showed a significant genotype effect and no block effect. The developed projected area of mycorrhizal or non-mycorrhizal root tips was not significantly different between genotypes.

Enzymatic activity patterns of the parental and the hybrids root system

For each plant, the seven enzymatic activities were measured successively on one mycorrhizal and on one non-mycorrhizal root tip (40 poplar genotypes x 8 plant replicates x 2 root tips). Mycorrhizal root tips never lost their ability to secrete the seven enzymes in the conditions of the test, even if sometimes at a very low level (e.g. for xylosidase, glucuronidase, laccase).

The seven enzymatic activities expressed in non-mycorrhizal roots did not differ significantly between the two parents (Table 1a). The seven enzymatic activities measured on ECM root tips differed significantly between the two parents (Table 1a): five activities (xylosidase, cellobiohydrolase, ß-glucosidase, acid phosphatase, and laccase) had a higher level in *P. trichocarpa* and two (N-acetylhexosaminidase and glucuronidase) had a higher level in *P. deltoides*.

Enzyme activity patterns of mycorrhizal and non-mycorrhizal root tips of the parents and of their progeny were different (Figure 2). Significant differences were found between plant genotypes for all activities measured on ECM root tips except for laccase activity (Table 1b). No significant differences were found between plant genotypes for any activities measured on non-mycorrhizal root tips except for acid phosphatase activity (Table 1b).

Six of the enzymatic activities differed significantly between mycorrhizal and non-mycorrhizal roots, while no laccase activity could be detected in non-mycorrhizal roots (Table 1b). Compared to non-mycorrhizal root tips, N-acetylhexosaminidase activity was increased by more than 100 in mycorrhizas, while glucuronidase, cellobiohydrolase and glucosidase activities were multiplied by a factor ranging between 50 and 100, and xylosidase and acid phosphatase between 15 and 50 (Table 1b).

FIGURE 1

Percentage of root colonization of the different Poplar clones (n = 40). Genotypes are ranked in mean percentage of root colonization. Bars represent SE (n = 8). Grey corresponds to *Populus trichocarpa* (male) and black to *Populus deltoides* (female).

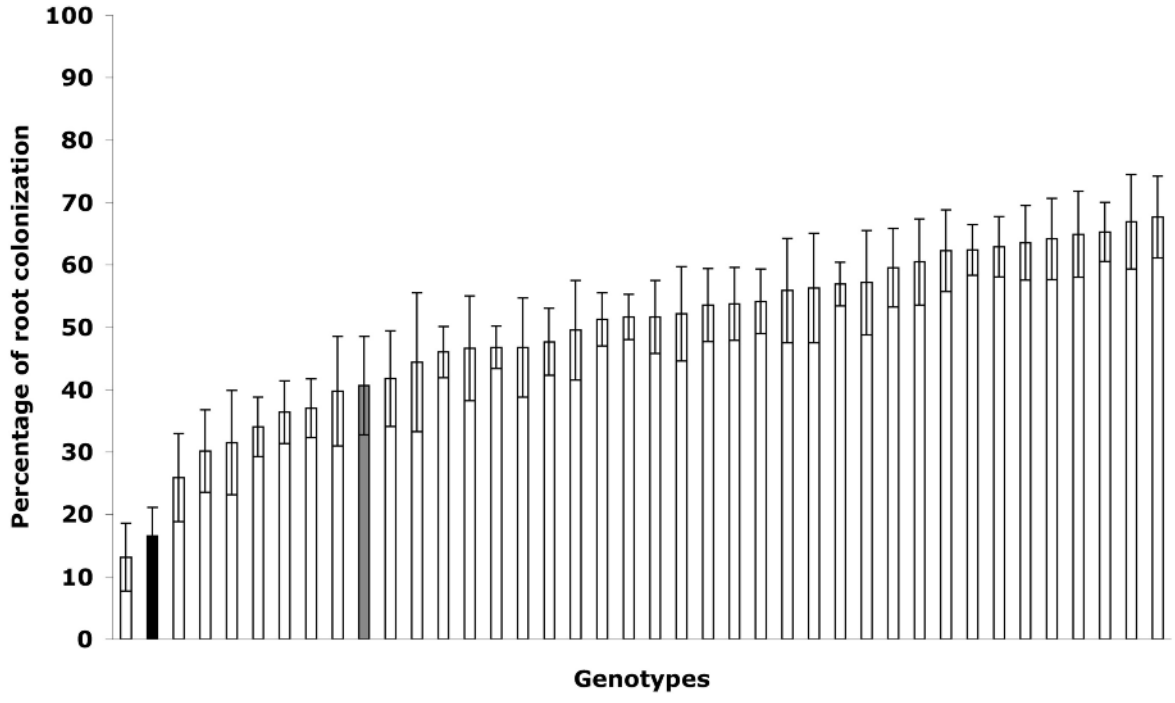
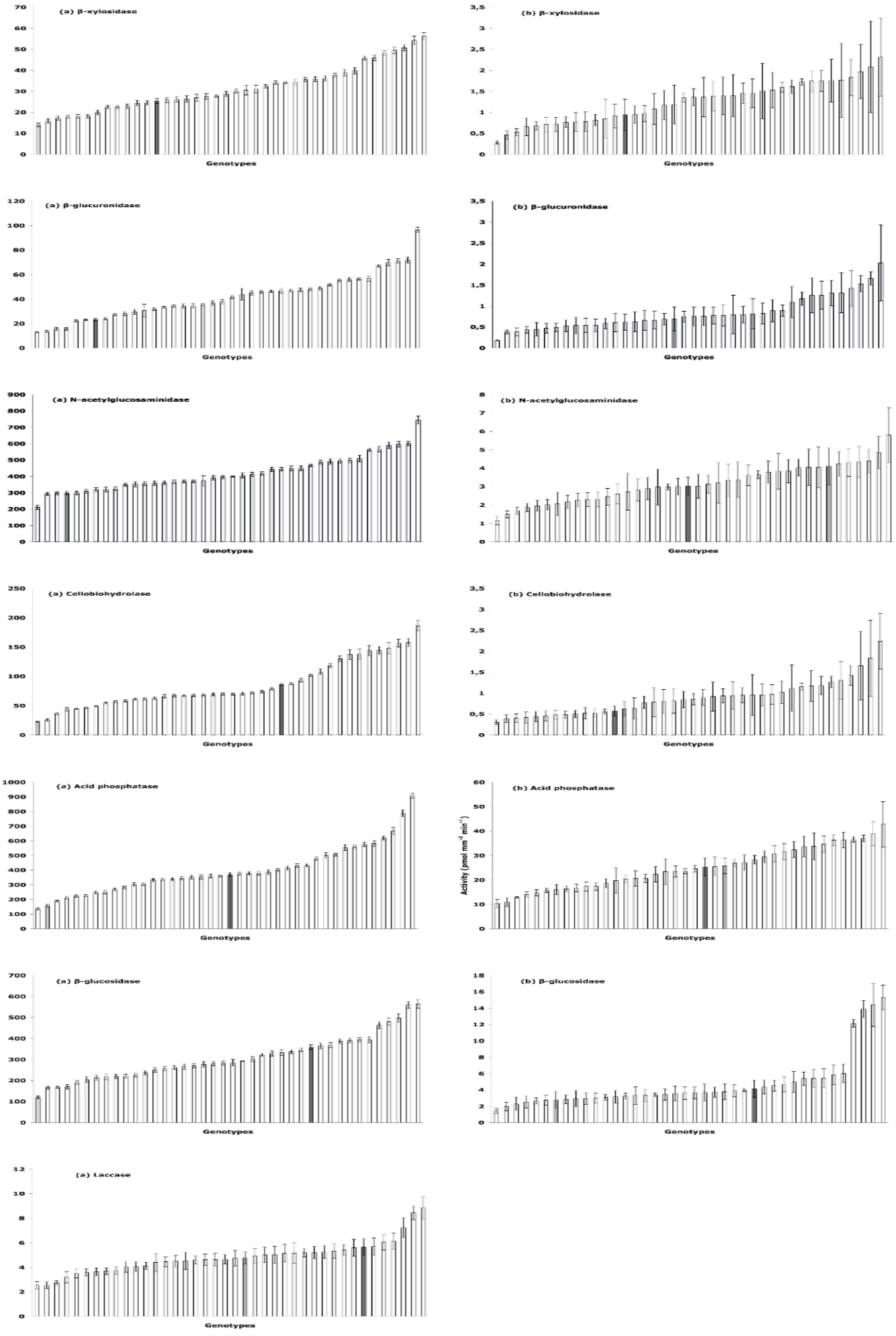


FIGURE 2


(a) β-xylosidase
(b) β-xylosidase
(a) β-glucuronidase
(b) β-glucuronidase
(a) N-acetylglucosaminidase
(b) N-acetylglucosaminidase
(a) Cellobiohydrolase
(b) Cellobiohydrolase
(a) Acid phosphatase
(b) Acid phosphatase
Activity (μmol mm⁻² min⁻¹)
(a) β-glucosidase
(b) β-glucosidase
(a) Laccase
Genotypes

Heritability

Heritability values (h^2) of plant phenotypic characters ranged from 0.21 to 0.48. The highest values of heritability were found for leaf maximum area (0.48 ± 0.01), chlorophyll content (0.45 ± 0.01), and stem and leaf dry weight (0.43 ± 0.01 and 0.50 ± 0.01, respectively). The lowest value was found for root dry weight (0.21 ± 0.01). A value of 0.45 ± 0.02 was found for the percentage of mycorrhizal colonization. Heritability values of enzymatic activities were similar for ectomycorrhizal and non-mycorrhizal root tips, except for laccase activity, which was not detected on non-mycorrhizal root tips (0.29 ± 0.01 in mycorrhizal root tips; 0 in non-mycorrhizal root tips). The highest heritabilities were found for N-acetylhexosaminidase (mycorrhizal root tips, 0.42 ± 0.01, non-mycorrhizal root tips 0.40 ± 0.01), acid phosphatase (mycorrhizal root tips, 0.41 ± 0.01; non-mycorrhizal root tips, 0.40 ± 0.01), glucosidase (mycorrhizal root tips 0.36 ± 0.01; non-mycorrhizal root tips, 0.34 ± 0.01) and cellobiohydrolase (mycorrhizal root tips, 0.33 ± 0.02; non-mycorrhizal root tips, 0.31 ± 0.02) activity. The lowest value was found for glucuronidase activity (mycorrhizal root tips 0.04 ± 0.01; non-mycorrhizal root tips, 0.04 ± 0.01). A medium value was found for xylosidase activity (mycorrhizal root tips 0.16 ± 0.01; non-mycorrhizal root tips, 0.19 ± 0.01).

Heterosis

For each trait, we have calculated the ratio between the average of the hybrids and the best parent (a) or between the average of the hybrids and the average of the two parents (b) (Table S1). The leaf maximum area exhibited a high positive heterosis (a = + 46, b = + 51), while the leaf

TABLE 1

Average enzymatic activities of ectomycorrhizal and non-mycorrhizal root tips of poplar. Enzyme activities are expressed as pmol.mm^{-2}.min^{-1} (Courty *et al.* 2005). Mean and SE are given for each activity. Significant differences (S, $p < 0.001$) are indicated by an asterisk (*). The effect of plant genotype on enzyme activities was assessed for mycorrhizal root tips and non-mycorrhizal root tips.

(a) Comparison of the parents: *P. trichocarpa* (Pt) and *P. deltoides* (Pd)

	Mycorrhizal root tips			Non-mycorrhizal root tips		
	Pt	Pd	S	Pt	Pd	S
Xylosidase	25.43 ± 1.67	14.04 ± 1.17	*	0.94 ± 0.52	0.46 ± 0.14	
Glucuronidase	23.29 ± 1.82	35.47 ± 1.47	*	0.69 ± 0.41	0.81 ± 0.51	
N-acetylhexosaminidase	297.83 ± 13.86	412.81 ± 17.72	*	3.02 ± 0.70	4.1 ± 1.42	
Cellobiohydrolase	86.51 ± 2.21	22.42 ± 1.61	*	0.57 ± 0.17	0.62 ± 0.24	
Glucosidase	358.94 ± 16.96	119.77 ± 9.77	*	4.13 ± 1.47	2.78 ± 1.45	
Acid phosphatase	368.46 ± 19.47	155.94 ± 12.84	*	25.33 ± 5.15	25.77 ± 4.47	
Laccase	5.64 ± 0.95	4.76 ± 0.72	*	0	0	

(b) Comparison of the 40 plant genotypes (the two parents and the 38 progenies). Significant differences in enzymatic activities between mycorrhizal and non-mycorrhizal root tips are also reported

	n	Mycorrhizal root tips		Non-mycorrhizal root tips		Mycorrhizal versus non-mycorrhizal	
		Values	S	Values	S	Ratio	S
Xylosidase	317	31.32 ± 1.76	*	1.23 ± 0.07		25.5	*
Glucuronidase	317	41.41 ± 2.90	*	0.82 ± 0.06		50.5	*
N-acetylhexosaminidase	317	418.30 ± 17.26	*	3.14 ± 0.16		133.2	*
Cellobiohydrolase	317	85.07 ± 6.48	*	0.87 ± 0.06		97.8	*
Glucosidase	317	306.78 ± 16.93	*	4.68 ± 0.52		65.5	*
Acid phosphatase	317	396.65 ± 26.42	*	24.81 ± 1.35	*	16	*
Laccase	317	4.8 ± 0.2		0			nd

dry weight exhibited a negative one (a = - 41, b = - 36). The percentage of mycorrhizal colonization also exhibited a positive heterosis (a = + 25, b = + 75). In mycorrhizal and non-mycorrhizal root tips, all the enzymatic activities displayed a positive heterosis at least for the b values, with the exception of laccase activity in mycorrhizal root tips (a = - 15, b = - 8) and acid phosphatase activity in non-mycorrhizal root tips (a = - 4, b = - 3).

Gene expression

Gene expression of laccase (*Lcc*), N-acetylglucosaminidase (*Nag*) and acid phosphatase (*Pap*) was assessed by whole-genome expression oligoarray analyses in poplar and *L. bicolor* (Table 2). In poplar, gene expression was compared between non-mycorrhizal and mycorrhizal roots. Thirty-two laccases (*Lcc1* to *Lcc32*) were detected in poplar. Three of them are mitochondrial and 27 have a signal peptide meaning that they belong to secreted pathways. On the 32 genes coding for laccases in poplar, the expression of 11 could be assessed. On the 11 expressed and also putatively expressed, one (*Lcc6*) was significantly up regulated in mycorrhizas and two were down regulated in mycorrhizas (*Lcc16* and *Lcc31*). Two N-acetylglucosaminidase (*Nag1* and *Nag2*) genes were expressed and a signal peptide was found for both of them. The expression of *Nag1* and *Nag2* was not modified by mycorrhizal establishment. Seven acid phosphatase genes (*Pap1* to *Pap7*) were expressed and a signal peptide was found for five of them (*Pap2, Pap3, Pap5, Pap6, Pap7*). Two (*Pap5* and *Pap7*) were significantly up expressed in mycorrhizas.

For *L. bicolor*, gene expression was compared between mycorrhizal roots of *P. trichocarpa* and *P. deltoides*, and mycelium growing in pure culture. On the 9 laccases previously described (Courty *et al.* 2009), one was mitochondrial (*Lcc5*) and 8 had a signal peptide meaning that they belong to secreted pathways. Seven of these laccase genes were expressed, while two were not (*Lcc2* and L*cc7*). *Lcc5* was only expressed in the free living mycelium. *Lcc9* was significantly down expressed in *P. deltoides* and *P. trichocarpa* mycorrhizas, while *Lcc6* was significantly down regulated only in *P. deltoides* mycorrhizas and *Lcc8* up regulated in *P. trichocarpa* mycorrhizas. The other laccases were not significantly regulated. The only acid phosphatase (*Pap1*) expressed displayed a peptide signal. Its expression was not significantly different between mycorrhizas and free-living mycelium. Among the two N-acetylglucosaminidase (*Nag1* and *Nag2*) expressed, only *Nag2* exhibited a signal peptide. The expression of *Nag2* was higher in the free-living mycelium than in *P. trichocarpa-L.bicolor* and *P.deltoides-L.bicolor* mycorrhizas.

Correlations between the different traits

No poplar trait was correlated with enzymatic activities of non-mycorrhizal root tips (Table 3). LMA, stem and roots dry weight were not correlated with any activities either from ectomycorrhizal and non-mycorrhizal root tips (Table 3). Chlorophyll content was significantly negatively correlated with xylosidase, cellobiohydrolase and glucosidase activities, three enzymes involved in cellulose and hemi-cellulose catabolism.

Enzymatic activities of ectomycorrhizal root tips were not correlated with those of non-mycorrhizal root tips. All the enzymatic activities of mycorrhizal root tips were correlated between them, except for laccase activity. Similarly, except for acid phosphatase, all the enzymatic activities of non-mycorrhizal root tips were correlated between them. Xylosidase activity of mycorrhizal roots was the only activity positively correlated with the percentage of mycorhizal infection (Table 3). Stem and root dry weights also were significantly correlated with the percentage of root colonization.

Discussion

Enzymatic activities of non-mycorrhizal root tips and mycorrhizas

The potential activities of enzymes involved in organic matter breakdown or organic phosphorus mobilization measured on poplar root tips colonized or not by *L. bicolor* were significantly different. Here, we found that the ectomycorrhizal complex adds or substitutes enzymes secreted from poplar roots. Compared to non-mycorrhizal root tips, N-acetylhexosaminidase activity was multiplied by more than 100 in mycorrhizas, while glucuronidase, cellobiohydrolase and glucosidase activities were multiplied between 50 and 100 and xylosidase and acid phosphatase between 15 and 50. Moreover, laccase activity could not be detected on non-mycorrhizal roots. By degrading organic compounds, including those from their own mycelia, and channelling nutrients directly to the host tree, ECMf have the capacity to shorten mineralization pathways in which free-living decomposers are involved. It is well known that ECMf are able to secrete enzymes, which allow the release of nutrients from soil organic matter (Cullings and Courty, 2009; Courty *et al.* 2010a). However, this study was not driven to understand the role of ECMf in the release of nutrients. Enzymes we measured should be considered as functional traits to study effects of soil or host tree parameters on ECMf. It is the first time that the breath of the modifications induced by the symbiotic association in the potential enzymatic secretion by the root system is measured.

Expression of genes involved in enzymatic activities

Although the complete sequences of the *P. trichocarpa* and *L. bicolor* genome are available, all genes putatively encoding the proteins responsible for the measured activities were not identified. The correspondence between enzymatic assay and gene expression could be done for three of them: laccase (*Lcc*), N-acetylglucosaminidase (*Nag*) and acid phosphatase (*Pho*). Two genes code for a secreted N-acetylglucosaminidase

TABLE 2

Quantification by exon expression array of the transcript levels of laccase (*lcc*), N-acetylglucosaminidase (*Nag*) and acid phosphatase (*Pap*) genes in different conditions. The length (bp) of signal peptide (Signal-P) was predicted with Signal P 3.0 (**http://www.cbs.dtu.dk/services/SignalP/**). The prediction of the subcellular location of the proteins was performed with TargetP 1.1 available on the webpage (http://www.cbs.dtu.dk/services/TargetP/); M=mitochondrial, S=secreted and -=unknown.
Three biological replicates were used for each treatment with NimbleGen oligoarrays (v.2.0; NG2). A Cyber-T test was performed on the mean for each transcript (*, P < 0.05).

(a) Poplar genes. The NimbleGen array analysis was carried out using *P. trichocarpa* root tips mycorrhizal or not by *L. bicolor*. Transcript levels in non-mycorrhizal root tips were used as the control values. NT: not transcribed. (-): gene not on the array or no reliable probe left.

	Protein ID	Signal-P	Target-P	Ratio
Lcc1	820390	24	S	1.0
Lcc2	557962	24	S	-
Lcc3	797646	22	S	NT
Lcc4	576931	23	S	-
Lcc5	762473	25	S	-
Lcc6	767563	28	S	4.3*
Lcc7	219290	29	S	1.4
Lcc8	759686		-	1.2
Lcc9	653089	24	S	-
Lcc10	819177	26	S	NT
Lcc11	797888	31	S	NT
Lcc12	579478	32	S	-
Lcc13	235935	32	S	-
Lcc14	831900	34	S	NT
Lcc15	235930	32	S	-
Lcc16	768177	26	S	0.4*
Lcc17	548008	32	S	-
Lcc18	783559	32	S	NT
Lcc19	822366	32	S	0.4
Lcc20	560853	23	S	NT
Lcc21	592533	23	S	-
Lcc22	832603	31	S	0.3
Lcc23	738903		M	-
Lcc24	574533	30	S	-
Lcc25	777748	23	S	0.8
Lcc26	738893		M	-
Lcc27	571858	33	S	-
Lcc28	574985	28	S	-
Lcc29	205176		M	1.6
Lcc30	569758	28	S	-
Lcc31	420672	17	S	0.4*
Lcc32	774519		-	0.2
Nag1	772972	27	S	1.4
Nag2	202916	25	S	1.5
Pap1	821155		-	1.4
Pap2	831269	28	S	0.6
Pap3	818768	27	S	0.5
Pap4	816041		-	1.4
Pap5	272725	22	S	2.1*
Pap6	259486	23	S	1.9
Pap7	825753	32	S	1.8*

(b) L. bicolor genes. The NimbleGen array analysis was carried out using P. trichocarpa (Pt) and P. deltoides (Pd) mycorrhizal by L. bicolor. Transcript levels in the mycelium grown in pure culture were used as the control values. NE: not expressed in mycorrhizas (signal under background). (-): gene not on the array or no reliable probe left.

Ratio Pd	Protein ID	Signal-P	Target-P	Ratio Pt
Lcc1	399743	17	S	0.1	0.2
Lcc2	399744	17	S	-	-
Lcc3	399745	17	S	4.4	2.0
Lcc4	399746	18	S	0.8	0.6
Lcc5	399747	19	M	NE	NE
Lcc6	399748	20	S	0.4	0.3*
Lcc7	399750	19	S	-	-
Lcc8	399749	22	S	2.0*	1.4
Lcc9	399751	16	S	0.3*	0.3*
Pap1	310810	21	S	0.6	0.9
Nag1	309753			0.6	0.6*
Nag2	182604	18	S	0.4*	0.4*

(*Nag*) involved in chitin catabolism in poplar and also two in *L. bicolor*. In this work, poplar *Nag1* and *Nag2* were expressed both in root tips and mycorrhizas but were not regulated by the symbiosis. On the contrary, the expression of the *L. bicolor Nag1* and *Nag2* were down regulated in mycorrhizas. Nevertheless, the activity of N-acetylglucosaminidase was multiplied by more than 130 times in the mycorrhizas compared to non-mycorrhizal root tips. We can make the assumption that (i) *L. bicolor* N-acetylglucosaminidases are secreted outside the mycelium of the sheath in mycorrhizas, while the Poplar N-acetylglucosaminidase are not secreted outside the root tips, (ii) *L. bicolor* N-acetylglucosaminidases can be involved in nitrogen mobilization from chitin by degrading its own mycelia and in defence against soil pathogenic fungi, (iii) *L. bicolor* can have the ability as *Trichoderma asperellum* (Ramot *et al.* 2004) to store a high amount of this enzyme in an active form and secrete it when mycelium is sensing the substrate.

In the *L. bicolor* genome, nine genes coding for laccases were characterized (Courty *et al.* 2009). In this experiment, six putatively secreted were expressed and three (*Lcc1, Lcc3, Lcc4*) were not significantly regulated by symbiosis. One (*Lcc8*) was significantly over expressed in mycorrhizas and two (*Lcc6* and *Lcc9*) under expressed. In *P. trichocarpa* genome, 32 genes are coding for laccases, 27 display a signal peptide meaning that they belong to secreted pathways and, in this work, 20 were expressed in root tips. Despite this large number of laccase genes which were expressed, no laccase activity was detected by the ABTS test on non-mycorrhizal root tips. This means that poplar laccases are not secreted in the rhizosphere. In *Arabidopsis thaliana*, only a few of the laccase genes were expressed in a pattern that could be considered consistent with a major role for these enzymes in lignin deposition (McCaig *et al.* 2005). Poplar laccases seems to be not cell wall bound, nor secreted outside the cells. They are probably involved in the polymerization of lignin precursors or in other functions.

Acid phosphatases, able to free phosphate groups from complex organic compounds, are widespread in living organisms. Both ECMf and plants secrete acid phosphatases in the rhizosphere. In most of the studies, mycorrhizas secrete more phosphatases than non-colonized roots (Colpaert *et al.* 1997; Conn and Dighton, 2000). Nevertheless, there are some exceptions (Cumming,

TABLE 3

Correlation matrix (Pearson Correlation r) between poplar traits, enzymatic activities and percentage of mycorrhizal colonization. Correlation is significant for $p < 0.01$ (light grey cells).
Abbreviations: %, percentage of mycorrhizal colonization; Chl, Chlorophyll (g/m²); DW, Dry Weight (g); LMA, Leaf maximum area (m²); M, mycorrhizal root tips; NM, non-mycorrhizal root tips; Pho, acid phosphatase; Nag, N-acetyl-glucosaminidase; Gls, β-glucosidase; Cel, cellobiohydrolase; Xyl, β-xylosidase; Lac, laccase; Glr, β-glucuronidase.

	%	Chl	DW Stem	DW Roots	LMA	M Xyl	M Glr	M Nag	M Cel	M Gls	M Pho	M Lac	NM Xyl	NM Glr	NM Nag	NM Cel	NM Gls	NM Pho
%	1	-0,09	0,20	0,20	0,06	0,22	0,14	0,13	0,07	0,13	0,04	0,05	0,05	0,13	0,02	-0,02	0,06	0,13
Chl		1	-0,12	-0,12	-0,01	-0,19	-0,01	-0,07	-0,18	-0,18	-0,05	0,01	-0,00	0,02	0,08	0,08	0,00	0,11
DW Stem			1	0,42	-0,28	-0,05	-0,04	0,04	-0,07	-0,13	-0,02	-0,01	-0,01	0,09	0,12	0,00	-0,02	0,01
DW Roots				1	-0,20	0,05	0,04	0,06	-0,01	0,02	-0,03	0,06	-0,07	0,00	-0,09	-0,09	-0,10	-0,02
LMA					1	0,14	0,09	0,14	0,06	0,10	0,12	0,13	-0,08	-0,13	-0,04	-0,07	-0,06	0,04
M Xyl						1	0,12	0,41	0,64	0,56	0,25	0,25	0,07	0,02	-0,03	0,11	-0,03	0,09
M Glr							1	0,27	0,18	0,15	0,30	0,15	0,01	0,07	-0,12	-0,13	-0,07	0,15
M Nag								1	0,35	0,33	0,47	0,31	-0,11	-0,08	-0,04	-0,01	-0,19	0,32
M Cel									1	0,69	0,22	0,16	0,06	-0,04	0,02	0,05	-0,06	0,05
M Gls										1	0,25	0,27	0,01	-0,05	-0,07	0,06	-0,05	0,09
M Pho											1	0,29	0,01	0,08	-0,01	0,02	0,01	0,36
M Lac												1	-0,12	-0,05	-0,14	-0,06	-0,10	0,21
NM Xyl													1	0,47	0,27	0,43	0,35	-0,03
NM Glr														1	0,25	0,39	0,29	0,04
NM Nag															1	0,24	0,31	-0,03
NM Cel																1	0,28	-0,03
NM Gls																	1	0,07
NM Pho																		1

1996). ECMf exhibit high phosphatase release in their environment, particularly under mineral phosphorus deficiency (Dighton, 1983; Nygren and Rosling, 2009). The *L. bicolor* genome comprises only one gene coding for a putative secreted acid phosphatase, while *P. trichocarpa* genome contains five. The phosphatase *Pap1* from L. bicolor was not regulated by symbiosis. *Pap5* and *Pap7* from poplar were significantly highly expressed under mycorrhizal conditions, whereas the three others (*Pap2, Pap3, Pap6*) were not significantly regulated.

Ezawa *et al.* (2005) have shown, on *Tagetes petala* in symbiosis with *Archaeospora leptoticha*, that the levels of transcripts of the *T. petala* acid phosphatase (*TpPAP1*) was increased eight times by *A. leptoticha* colonization. Our results support the hypothesis of Ezawa *et al.* (2005) on the fungal activation of the low-phosphate adaptation system of the plant partner and seem to show that the same mechanism of plant phosphatase activation exists both in arbuscular mycorrhizas and ectomycorhizas.

Another hypothesis could be involved in the explanation of the differences in enzyme secretion between non-mycorrhizal and mycorrhizal roots. Inside the mycorrhizas, the root tissues, being isolated from the external medium by the fungal sheath, probably poorly contribute to enzyme secretion. The ability to degrade cellulose, hemicelluloses and lignin is widespread among fungi and soil bacteria (i.e. *Streptomyces* sp., *Bacillus* sp., *Cellulomonas* sp., Lynd *et al.* 2002). However, it is assumed that most of cellulose degradation in soil is performed by fungi (de Boer *et al.* 2005). Even if it has been shown that laccase genes were present in bacteria (Kellner *et al.* 2008), bacterial lignin degradation appears to be negligible in terrestrial environments compared to fungal lignin degradation (Peng *et al.* 2002). This is supported by the fact that we did not find any laccase activity on non-mycorrhizal root tips. Among bacteria, *Collimonas* sp. display chitinolytic activities (de Boer *et al.* 2004). However, these bacteria are present under specific conditions, completely different from greenhouse experiments with artificial substrate. Thus, despite the fact that this experiment was performed in non-axenic conditions, we may assume that the secreted enzymes that we measured in mycorrhizas were mainly due to fungal activity.

Host genetic control of ECM enzyme secretion

The enzymatic activities expressed in mycorrhizal roots differed significantly between the two parents, while it did not differ in non-mycorrhizal roots. Significant differences were found between poplar genotypes for all enzymatic activities measured on ECMs except for laccase activity. On the contrary, no significant differences were found between poplar genotypes for enzymatic activities of non-mycorrhizal root tips except for acid phosphatase activity. Heritability values of enzymatic activities were similar for ectomycorrhizal and non-mycorrhizal root tips, except for glucuronidase in both types of roots and for laccase that was not detected on non-mycorrhizal root tips. It is remarkable to find a high heritability value among the poplar genotypes for the enzymatic secretions of mycorrhizal roots, which are mainly due to fungal activity. Several previous studies have demonstrated significant genetic variability within plants and/or fungal species for symbiotic capability in mycorrhizal interactions. Rosado *et al.* (1994) reported a high value of heritability for colonization of *Pinus ellilotii* by the ECMf *Pisolithus tinctorius*, and moderate heritability for the development of *P. tinctorius* extramatrical mycelium. *Eucalyptus grandis*, *E. globulus, E. marginata* and *Pinus muricata* varied greatly in their growth response to different *Pisolithus* and *Rhizopogon* genotypes, respectively (Tonkin *et al.* 1989; Burgess *et al.* 1994; Piculell *et al.* 2008). Tagu *et al.* (2003, 2005) have already shown that the ability of poplar to form ECMs is under its genetic control. Other studies with contrasting results have found that plant genotype can play a dominant role in controlling the associate soil microbial communities (Mari *et al.* 2003; Korkama *et al.* 2007). Short-term experiment have either shown variations in mycorrhizal colonization, in microbial and in mycorrhizal communities (Gehring *et al.* 2006; Barbour *et al.* 2009; Lojewski *et al.* 2009) or few differences in arbuscular fungal and bacterial communities (Bever *et al.* 1996; Madritch and Hunter, 2002). Here, the degree of fungal enzymatic secretion is modulated according to poplar genotype. An explanation could be that the host genotype controls the amount of fungal tissue in the mantle and that enzyme activity is determined by the amount of fungal tissue present on the root. However, we found no significant differences in the projected area of mycorrhized or non-mycorrhized root tips between genotypes. So, it means that the amount of mycelium in the root tip is similar whatever the genotype. These results suggest the potential for poplar genome to drive the microbial-plant interaction, to create environments to which ECM fungi can respond and that could be explained by the "extended phenotype" phenomenom (Schweitzer *et al.* 2008; Whitham *et al.* 2008). As defined by Whitham *et al.* (2003), the heritable genetic variation within individual species (poplar in our study) has community and ecosystem consequences. In addition, we found a high positive heterosis for the capacity of poplar to form mycorrhizas (h^2 a = + 45 %). We also found positive heterosis for characters such as leaf maximum area, dry weight of roots and for five of the seven enzymatic activities of mycorrhizal roots. Heterosis for poplar hybrids is a well-known phenomenon (Li and Wu, 1997; Marron *et al.* 2006). Heterosis is determined by non-mutually exclusive mechanisms, including genome-wide dominance complementation, locus-specific overdominance effects and epistasis, although the relative contribution of each of these mechanisms is still unclear (Lippman and Zamir, 2007). But it is also the first time that it is shown that plant heterosis could be expressed through the physiological activity of the fungal partner.

Conclusion

The genetic diversity in tree species can influence fluxes of nutrients as well as interactions with soil microorganisms. Assessing tree genotype x environment interactions is a major challenge in functional ecology. In this paper, our data linked and quantified the general relationships between poplar plant genetics, ECM fungal infection, and physiological parameters. In the association *L. bicolor*/poplar, variations in plant and fungal responses in these controlled conditions illustrate the broad plasticity of the interaction. In this study, the role of poplar genetics in determining both poplar growth characteristics and fungal activities has been highlighted.

Acknowledgements

Courty PE was supported by a grant of the French Ministry of Ecology and Sustainable Development and Labbé J by a scholarship of Région Lorraine/INRA; part of this research has been supported by the Biological Invasions program of the same Ministry. This project was also supported by the European Network of Excellence EVOLTREE. We thank Joseph Armento and Krista Plett for language corrections and helpful comments on the manuscript, and Francis Martin for valuable discussion.

REFERENCES

Barbour RC, Baker SC, O'Reilly-Wapstra JM, Harvest TM, Potts BM. 2009. A footprint of tree-genetics on the biota of the forest floor. *Oikos*, **118**, 1917-1923.

Barker SJ, Duplessis S, Tagu D. 2002. The application of genetic approaches for investigations of mycorrhizal symbioses. *Plant and Soil*, **244**, 85-95.

Bever JD, Morton J, Antonovics J, Schultz P. 1996. Host specificity and diversity of glomalean fungi: an experimental approach in an old-field community. *Journal of Ecology*, **84**, 71-82.

de Boer W, Folman LB, Summerbell RC, Boddy L. 2005. Living in a fungal world: impact of fungi on soil bacterial niche development. *FEMS Microbiology Review*, **29**, 795-811.

de Boer W, Leveau JHJ, Kowalchuk GA, Gunnewiek PJAK, Abeln ECA, Figge MJ, Sjollema K, Janse JD, van Veen JA. 2004. *Collimonas fungivorans* gen. nov., sp. nov., a chitinolytic soil bacterium with the ability to grow on living fungal hyphae. *International Journal of Systematic and Evolutionary Microbiology*, **54**, 857-864.

Brunner AM, Busov VB, Strauss SH. 2004. Poplar genome sequence: functional genomics in an ecologically dominant plant species. *TRENDS in Plant Science*, **9**, 49-56.

Buée M, Courty PE, Mignot D, Garbaye J. 2007. Soil niche effect on species diversity and catabolic activities in an ectomycorrhizal fungal community. *Soil Biology and Biochemistry*, **39**, 1947-1955.

Burgess T, Dell B, Malajczuk Thompson N. 1994. Variation in mycorrhizal development and growth stimulation by 20 *Pisolithus* isolates inoculated on to *Eucalyptus grandis* W. Hill ex Maiden. *New Phytologist*, **127**, 731-739.

Cervera MT, Storme V, Ivens B, Gusmao J, Liu BH, Hostyn V, Hostyn Van Slycken J, Van Montagu M, Boerjan W. 2001. Dense genetic linkagmaps of three Populus species (*Populus deltoides*, *P. nigra* and *P. trichocarpa* based on AFLP and microsatellite markers. *Genetics*, **158**, 787-809.

Colpaert JV, van Laere A, van Tichelen KK, van Assche JA. 1997. The use of inositol hexaphosphate as a phosphorus source by mycorrhizal and non-mycorrhizal scots pine (*Pinus sylvestris*). *Functional Ecology*, **11,** 407-415.

Conn C, Dighton J. 2000. Litter quality influences on decomposition, ectomycorrhizal community structure and mycorrhizal root surface acid phosphatase activity. *Soil Biology and Biochemistry*, **32**, 489-496.

Cornelius J. 1994. Heritabilities and additive genetic coefficients of variation in forest trees. *Canadian Journal of Forest Research*, **24**, 372-379.

Courty PE, Franc A, Garbaye J. 2010a. Temporal and functional pattern of secreted enzyme activities in an ectomycorrhizal community. *Soil Biology and Biochemistry, submitted*.

Courty PE, Buée M, Diedhiou AG, Frey-Klett P, Le Tacon F, Rineau F, Turpault MP, Uroz S, Garbaye J. 2010b. The role of ectomycorrhizal communities in forest ecosystem processes: new perspectives and emerging concepts. *Soil Biology and Biochemistry*, **42**, 679-698.

Courty PE, Hoegger P, Kilaru S, Kohler A, Buée M, Garbaye J, Martin F, Kües U. 2009. Phylogenetic analysis, genomic organization and expression analysis of multicopper oxidases in the ectomycorrhizal basidiomycete *Laccaria bicolor. New Phytologist*, **182**, 736-750.

Courty PE, Pouysegur R, Buée M, Garbaye J. 2006. Laccase and phosphatase activities of the dominant ectomycorrhizal types in a lowland oak forest. *Soil Biology and Biochemistry,* **38**, 1219-1222.

Courty PE, Pritsch K, Schloter M, Hartmann A, Garbaye J. 2005. Activity profiling of ectomycorrhiza communities in two forest soils using multiple enzymatic tests. *New Phytologist*, **167**, 309-319.

Cullings K, Courty PE. 2009. Saprotrophic capabilities as functional traits to study functional diversity and resilience of ectomycorrjizal communities. *Oecologia,* **161**, 661-664.

Cumming JR. 1996. Phosphate-limitation physiology in ectomycorrhizal pitch pine (*Pinus rigida*) seedlings. *Tree Physiology*, **16**, 977-983

Di Battista C, Selosse MA, Bouchard D, Stenström E, Le Tacon F. 1996. Variations in symbiotic efficiency, phenotypic characters and ploidy level among different isolates of the ectomycorrhizal basidiomycete *Laccaria bicolor* strain S238. *Mycological Research*, **100**, 1315-1324.

Dighton J. 1983. Phosphatase production by mycorrhizal fungi. *Plant and Soil*, **71**, 455-462.

Dillen SY, Marron N, Bastien C, Ricciotti L, Salani F, Sabatti M, Pinel MPC, Rae AM, Taylor G, Ceulemans R. 2007. Effects of environment and progeny on biomass estimations of five hybrid poplar families grown at three contrasting sites across Europe. *Forest Ecology and Management,* **252**, 12-23.

Ezawa T, Hayatsu M, Saito M. 2005. A new hypothesis on the strategy for acquisition of phosphorus in arbuscular mycorrhiza: up-regulation of secreted acid phosphatase gene in the host plant. *Molecular Plant-Microbe Interactions*, **18**, 1046-1053.

Frey-Klett P, Pierrat JC, Garbaye J. 1997. Location and survival of mycorrhiza helper *Pseudomonas fluorescens* during establishment of ectomycorrhizal symbiosis between *Laccaria bicolor* and Douglas fir. *Applied and Environmental Microbiology*, **63**, 139-144.

Gehring CA, Mueller R, Whitham TG. 2006. Environmental and genetic effects on the formation of ectomycorrhizal and arbuscular mycorrhizal associations in cottonwoods. *Oecologia*, **149**, 158-164.

Gehring CA, Whitham TG. 2002. *Mycorrhiza-herbivore interactions: population and community consequences*. In: van der Heijden MGA, Sanders IR (eds) Mycorrhizal ecology. Springer, Berlin Heidelberg New York, pp 295-320.

Jorge V, Dowkiw A, Faivre-Rampant P, Bastien C. 2005. Genetic architecture of qualitative and quantitative *Melampsora larici-populina* leaf rust resistance in hybrid poplar: genetic mapping and QTL detection. *New Phytologist*, **167**, 113-127.

Kellner H, Luis P, Zimdars D, Kiesel B, Buscot F. 2008. Diversity of laccase-like multicopper-oxidase genes in forest and grassland Cambisol soil samples. *Soil Biology and Biochemistry*, **40**, 638-648.

Khasa DP, Hambling B, Kernaghan G, Fung M, Ngimbi E. 2002. Genetic variability in salt tolerance of selected boreal woody seedlings. *Forest Ecology and Management*, **165**, 257-269.

Karst J, Jones MD, Turkington R. 2009. Ectomycorrhizal colonization and intraspecific variation in growth responses of lodgepole pine. *Plant Ecology*, **200**, 161-165.

Koide RT, Sharda JN, Herr JR, Malcolm GM. 2008. Ectomycorrhizal fungi and the biotrophy–saprotrophy continuum. *New Phytologist*, **178**, 230-233.

Korkama T, Fritze H, Pakkanen A, Pennanen T. 2007. Interactions between extraradical ectomycorrhizal mycelia, microbes associated with the mycelia and growth rate of Norway spruce (*Picea abies*) clones. *New Phytologist*, **173**, 798-807.

Le Tacon F, Bouchard D. 1986. Effects of different ectomycorrhizal fungi on growth of larch, Douglas fir, Scots pine and Norway spruce seedlings in fumigated nursery soil. *Oecologia Applicata*, **7**, 389-402.

Li B, Wu R. 1997. Heterosis and genotype×environment interactions of juvenile aspens in two contrasting sites. *Canadian Journal of Forest Research*, **27**, 1525-1537.

Lindahl BD, Finlay RD, Cairney JWG. 2005. *Enzymatic activities of mycelia in mycorrhizal fungal communities.* In: Dighton J, White JF, Oudemans P, eds. The fungal community: its organization and role in the ecosystem, 3rd edn. Boca Raton, FL, USA: CRC Press, 331-348.

Linderman RG, Davis EA. 2004. Varied response of marigold (Tagetes spp.) genotypes to inoculation with different arbuscular mycorrhizal fungi. *Scientia Horticulturae*, **99**, 67-78.

Lippman ZB, Zamir D. 2007. Heterosis: revisiting the magic. *Trends in Genetics*, **23**, 60-66.

Lojewski NR, Fischer DG, Bailey JK, Schweitzer JA, Whitham TG, Hart SC. 2009. Genetic basis of aboveground productivity in two native Populus species and their hybrids. *Tree Physiology*, **29**, 1133-1142.

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS. 2002. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiology Molecular Biology Review*, **66**, 506-577.

Madritch MD, Hunter MD. 2002. Phenotypic diversity influences ecosystem functioning in an oak sandhills community. *Ecology*, **83**, 2084-2090.

Mari S, Jonsson A, Finlay R, Ericsson T, Kähr M, Eriksson G. 2003. Genetic variation in nitrogen uptake and growth in mycorrhizal and nonmycorrhizal *Picea abies* (L.) Karst. seedlings. *Forest Science*, **49**, 258-267

Marron N, Bastien C, Sabatti M, Taylor G, Ceulemans R. 2006. Plasticity of growth and sylleptic branchiness in two poplar families grown at three sites across Europe. *Tree Physiology*, **26**, 935-946.

Marron N, Villar M, Dreyer E, Delay D, Boudouresque E, Petit JM, Delmotte FM, Guehl JM, Brignolas F. 2005. Diversity of leaf traits related to productivity in 31 *Populus deltoides-Populus nigra clones. Tree Physiology*, 25, 425-435.

Martin F, Nehls U. 2009. Harnessing ectomycorrhizal genomics for ecological insights. *Current Opinion in Plant Biology*, 12, 508-515.

Martin F, Aerts A, Ahrén D, et al. 2008. The genome sequence of the basidiomycete fungus *Laccaria bicolor* provides insights into the mycorrhizal symbiosis. *Nature*, **452**, 88-92.

Monje OA, Bugbee B. 1992. Inherent limitations of nondestructive chlorophyll meters: a comparison of two types of meters. *Hortscience*, **27**, 69-71.

McCaig B, Meagher RB, Dean JFD. 2005. Gene structure and molecular analysis of the laccase-like multicopper oxidase (LMCO) gene family in *Arabidopsis thaliana. Planta*, **221**, 619-636.

Nygren CMR, Rosling A. 2009. Localisation of phosphomonoesterase activity in ectomycorrhizal fungi grown on different phosphorus sources. *Mycorrhiza*, **19**, 197-204.

Peng X, Masai E, Kitayama H, Harada K, Katayama Y, Fukuda M. 2002. Characterization of the 5-carboxyvanillate decarboxilase gene an dits role in lignin-related biphenyl catabolism in *Sphingomonas paucimobilis* SYK-6. *Applied and Environmental Microbiology*, **68**, 4407-4415.

Piculell BJ, Hoeksema JD, Thompson JN. 2008. Interactions of biotic and abiotic environmental factors in an ectomycorrhizal symbiosis, and the potential for selection mosaics. *BMC Biology*, **6**, 23.

R Development Core Team. 2006. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ramot O, Viterbo A, Friesem D, Oppenheim A, Chet I. 2004. Regulation of two homodimer hexosaminidases in the mycoparasitic fungus *Trichoderma asperellum* by glucosamine. *Current Genetics*, **45**, 205-213.

Richardson AD, Duigan SP, Berlyn GP. 2002. An estimation of non-invasive methods to estimate foliar chlorophyll content. *New Phytologist*, **153**, 185-194.

Rosado SCS, Kropp BR, Piche Y. 1994. Genetics of ectomycorrhizal symbiosis. I. Host plant variability and heritability of ectomycorrhizal and root traits. *New Phytologist*, **126**, 105-110.

Schweitzer JA, Bailey JK, Fischer DG, LeRoy CJ, Lonsdorf EV, Whitham TG, Hart SC. 2008. Plant-soil-microorganism interactions: Heritable relationship between plant

genotype and associated soil microorganisms. *Ecology*, **89**, 773-781.

Smith SE, Read DJ. 2008. *Mycorrhizal Symbiosis* 3rd Ed. Academic Press, London.

Tagu D, Bastien C, Faivre-Rampant P, Garbaye J, Vion P, Villar M, Martin F. 2005. Genetic analysis of phenotypic variation for ectomycorrhiza formation in interspecific F1 poplar full-sib family. *Mycorrhiza*, **15**, 87-91.

Tagu D, Faivre Rampant P, Lapeyrie F, Frey-Klett P, Vion P, Villar M. 2001. Variation in the ability to form ectomycorrhizas in the F1 progeny of an interspecific poplar (*Populus* spp.) cross. *Mycorrhiza*, **10**, 237-240.

Tonkin CM, Malajczuk N, McComb JA. 1989. Ectomycorrhizal formation by micropropagated clones of *Eucalyptus marginata* inoculated with isolates of *Pisolithus tinctorius. New Phytologist,* **111**, 209-214.

Tuskan GA, DiFazio S, Jansson S, et al. 2006. The genome of black cottonwood, *Populus trichocarpa. Science,* 313, 1596-1604.

van der Heijden M, Sanders I. 2002. *Mycorrhizal ecological studies*. van der Heijden MGA, Sanders IR, eds. 2002. Berlin, Germany: Springer Verlag.

van der Heijden EW, Kuyper TW. 2001. Laboratory experiments imply the conditionality of mycorrhizal benefits for *Salix repens*: role of pH and nitrogen to phosphorus ratios. *Plant and Soil*, **228**, 275-290.

Whitham TG, Young WP, Martinsen GD, et al. 2003. Community and ecosystem genetics: a consequence of the extended phenotype. *Ecology*, **84**, 559-573.

Whitham TG, DiFazio SP, Schweitzer JA, Shuster SM, Allan GJ, Bailey JK, Woolbright SA. 2008. Extending genomics to natural communities and ecosystems. *Science*, **320**, 492-495.

MOLECULAR FOOTPRINTS OF LOCAL ADAPTATION IN TWO MEDITERRANEAN CONIFERS

Delphine Grivet [1]; Federico Sebastiani [2, 3]; Ricardo Alía [1,5]; Thomas Bataillon [4]; Sara Torre [3]; Mario Zabal-Aguirre [1]; Giovanni G. Vendramin [3]; Santiago C. González-Martínez [1,5]

[1] Department of Forest Ecology & Genetics, Center of Forest Research, CIFOR-INIA, Carretera de la Coruña km 7.5, E28040 Madrid, Spain.
[2] Department of Agricultural Biotechnology, Genexpress, University of Florence, Via della Lastruccia 14, I50019 Sesto Fiorentino (FI), Italy.
[3] Plant Genetics Institute, Division of Florence, National Research Council, via Madonna del Piano 10, I50019 Sesto Fiorentino (FI), Italy.
[4] BiRC – Bioinformatics Research Center, Aarhus University, C.F. Møllers Allé 8, DK8000 Aarhus C, Denmark
[5] Sustainable Forest Management Research Institute, University of Valladolid-INIA.

Corresponding Author:
Santiago C. González-Martínez
Department of Ecology & Genetics, Center of Forest Research, CIFOR-INIA, Carretera de la Coruña km 7.5, E28040 Madrid, Spain
e-mail: santiago@inia.es

This study combines neutrality tests and environmental correlations to identify non-neutral patterns of evolution in candidate genes related to drought stress in two closely-related Mediterranean conifers, *Pinus pinaster* Ait. and *Pinus halepensis* Mill. Based on previous studies, we selected 12 amplicons covering six candidate genes that were sequenced in a large sample spanning the full range of these two species. Neutrality tests relatively robust to demography (*DHEW* compound test and *ML-HKA* test) were used to detect selection events at different temporal scales. Environmental associations between variation at candidate genes and climatic variables were also examined. These combined approaches detected distinct genes that may be targeted by selection, most of them specific to only one of the two conifers, despite their recent divergence (< 10 Ma). An exception was *4-coumarate: CoA ligase* (*4cl*), a gene involved in the production of various important secondary products that appeared to play a role in local adaptation processes of both pines. Another remarkable result was that all significant environmental correlations involved temperature indices, highlighting the importance of this climatic factor as a selective driver on Mediterranean pines. The ability to detect natural selection at the DNA sequence level depends on the nature and the strength of the selection events, on the timescale at which they occurred and on the sensitivity of the methods to other evolutionary forces that can mimic selection (e.g. demography, population structure). Using complementary approaches can help to capture different aspects of the evolutionary processes that govern molecular variation at both intra- and interspecific levels.

Key words: Natural selection, neutrality tests, environmental associations, candidate genes, drought stress, *4cl*.

Published in *Molecular Biology and Evolution* (2011) 28: 101-116, with doi:10.1093/molbev/msq190

Tables S1, S2, S3, S4, S5, S6, S7 and S8, and Figures S1, S2 and S3, along with the corresponding supporting information legends and references, can be downloaded here: http://mbe.oxfordjournals.org/content/28/1/101/suppl/DC1

Introduction

Identifying candidate genes underlying genetic differences for adaptive traits can help to understand how species have adapted to their environment and to predict how they will respond to future climatic changes, which is a special concern in regions such as the Mediterranean Basin, where a substantial decrease in precipitation and a pronounced warming is expected in the near future (Petit, Hampe, and Cheddadi 2005; Giorgi and Lionello 2008). The ability to detect the footprints of natural selection in population DNA sequence samples depends on the nature and the strength of the selection events (Nielsen 2005), on the evolutionary scale at which they occur (Zhai, Nielsen, and Slatkin 2009) and on the sensitivity of the methods to other evolutionary forces that can mimic selection (e.g. demography, population structure; Biswas and Akey 2006).

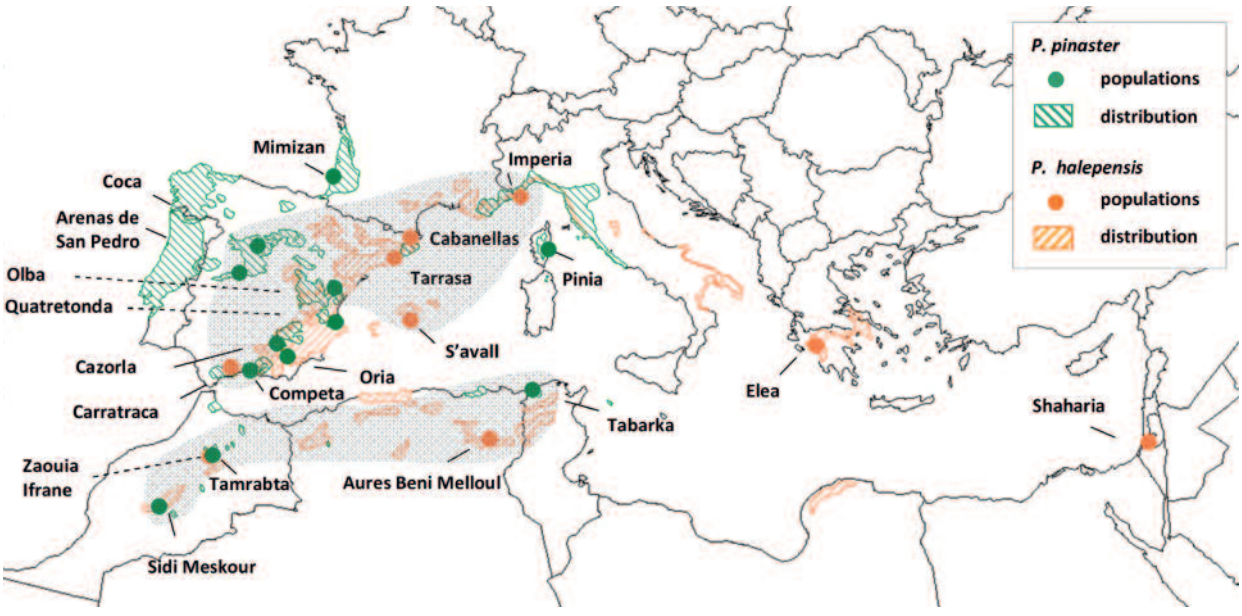
Positive selection, which drives the increase in frequency of advantageous mutations, is of particular interest because it underlies local adaptation. Recent or on-going positive selection events can be detected using methods based on polymorphism within species (Biswas and Akey 2006; Zhai, Nielsen, and Slatkin 2009). Recently, powerful methods based on site- and haplotype- frequency spectra that are relatively insensitive to demography or population structure have been developed (Zeng, Shi, and Wu 2007). Detecting selection events over a wider evolutionary scale (e.g. recurrent positive selection or ancient selective sweeps), can be efficiently done using methods contrasting within species polymorphism with among species divergence, as these methods are less sensitive to

assumptions regarding recombination or demography. The *HKA* (Hudson-Kreitman-Aguadé) neutrality test (Hudson, Kreitman, and Aguade 1987) and its extensions (e.g. Wright and Charlesworth 2004) are among the most powerful tests to detect positive selection (Zhai, Nielsen, and Slatkin 2009). In addition, the action of natural selection can be reflected in statistical associations between genetic and environmental data (Manel *et al.* 2003; Hancock *et al*. 2008; Coop *et al*. 2009). Several studies have shown that environmental heterogeneity influences the distribution of genetic diversity across plant populations (see Savolainen, Pyhäjärvi, and Knürr 2007, for forest trees; Nakazato, Bogonovich, and Moyle 2008; Montesinos *et al.* 2009). Detecting clinal genetic variation along environmental gradients can thus provide evidence for the action of selection (Gram and Sork 2001; Vasemagi and Primmer 2005; Parisod and Christin 2008). Approaches that detect adaptive clinal variation are very attractive because they can directly link environmental gradients with genotypes and phenotypes. However, detecting environmental associations depends on the spatial scale examined and the associations have to be controlled for selectively neutral processes that can also generate clines (e.g. Storz 2002; Hancock *et al*. 2008; Coop *et al*. 2009; Keller *et al.* 2009). Using environmental association approaches in conjunction with neutrality tests helps capturing different aspects of the evolutionary processes that govern molecular variation.

In the present study, we use the combined strategy outlined above to investigate patterns of polymorphism for a set of candidate genes related to drought stress in maritime pine (also known as cluster pine, *Pinus pinaster* Ait.), and in Aleppo pine (*Pinus halepensis* Mill.). *P. pinaster* is an economically important pine native to the occidental Mediterranean Basin and the European Atlantic front (Figure 1). This species is ecologically versatile, growing in a variety of substrates, in a wide range of elevations and under a range of Mediterranean and Atlantic climate regimes (semi-arid to very humid; Table 1). *P. halepensis*, a pine closely related to maritime pine, is one of the most abundant and widespread pine species of the Mediterranean Basin (Figure 1). This species, which also has a wide ecological breadth, is well adapted to dry conditions as well as high-intensity fire regimes (Tapias *et al.* 2004), which makes it a species of choice for afforestation in xeric Mediterranean areas. These two conifers, because they have extensive and partially overlapping distributions across the Mediterranean Basin, may possibly display adaptation to similar sets of contrasted environmental conditions, i.e. regional climate types and spatial variability. Among the key environmental factors to Pinus distribution in Mediterranean climates, temperature and rainfall have been shown to constitute good biological descriptors (Richardson and Bond 1991; Soto *et al.* 2010). Additionally, several common garden studies have documented differences in drought tolerance and/or gene expression among Mediterranean pine provenances (for *P. pinaster* see Nguyen and Lamant 1989; Costa *et al.* 1998; Chambel, Climent, and Alía 2007; Aranda *et al*. 2009; for *P. halepensis* see Atzmon, Moshe, and Schiller 2004; Sathyan, Newton, and Loopstra 2005;

FIGURE 1

Location of the 12 *P. pinaster* and the 9 *P. halepensis* populations and the distribution of the two species. Grey areas represent the two groups (Western Mediterranean and North African) defined by chloroplast SSRs (see text for details).



Voltas *et al*. 2008). Overall, these studies point to Mediterranean pines as an interesting system to study local adaptation mediated by abiotic stress response. In that context, studying allelic (SNP-allele and haplotype) frequency in drought-related candidate genes along temperature or precipitation clines can help to improve our understanding of pine adaptation to environmental heterogeneity.

Although *P. pinaster* and *P. halepensis* share various ecological requirements and may have responded to their environment in a similar manner, they present quite contrasted biogeographic and demographic histories that can also have a strong effect on the pattern of variation of the candidate genes examined. *Pinus pinaster* has a long-presence in the western Mediterranean Basin surviving the last glaciations in multiple refugia located in southeastern Spain, northern Africa and the Atlantic coast of Portugal (Bucci *et al*. 2007). In contrast, *P. halepensis* would have undergone long-range colonization of the western Mediterranean from ancient populations located in Greece and Turkey (Bucci *et al*. 1998; Morgante, Felice, and Vendramin 1998; Grivet *et al.* 2009). In addition, the dynamics of *P. halepensis* populations throughout the Mediterranean Basin highly depend on forest fires. This species-specific information on population dynamics has to be taken into account when looking at the molecular footprint of selection.

Based on previous population genomic studies identifying drought tolerance genes potentially under selection (see Eveno *et al.* 2008; Grivet *et al.* 2009) as well as on gene expression studies identifying genes affected by drought stress (Watkinson *et al.* 2003; Rani *et al.* 2009), we examine here patterns of nucleotide variation in an ortholog set of targeted candidate genes in *P. pinaster* and *P. halepensis*. We sample a broad geographic range of populations to: i) provide new evidence about positive selection acting on these genes, ii) identify the time scale at which selection events may have occurred, and iii) examine which environmental factors underlie molecular signatures of selection in both species. Our pluralistic approach provides insights on the adaptive strategy of two conifers that live under the same Mediterranean climate, but present distinct demographic, (re)colonization and life histories.

TABLE 1

Spatial coordinates and climatic variables for *P. pinaster* and *P. halepensis* populations

Population	Country	Spatial variables			Climatic variables							
		Altitude (m)	Latitude (degrees)	Longitude (degrees)	AMT (ºC)	TS	MTWM (ºC)	MTCM (ºC)	AP (mm)	PWM (mm)	PDM (mm)	PS
P. pinaster												
Arenas de San Pedro	Spain	733	40.194822	-5.116213	14.2	668.9	33.4	1.4	1318	199	12	60.8
Cazorla	Spain	1100	37.919675	-2.925765	11.5	650.2	30.6	-1.4	1257	179	11	58.3
Coca	Spain	800	41.254705	-4.497827	12.3	655.9	31.2	-0.6	454	55	15	30.0
Cómpeta	Spain	903	36.834265	-3.953989	14.0	512.1	27.5	4.7	899	132	5	64.1
Olba	Spain	1002	40.173309	-0.622966	12.4	600.7	28.3	0.6	509	63	24	31.5
Oria	Spain	1223	37.531165	-2.351138	13.1	633.1	30.6	0.4	357	46	5	46.2
Quatretonda	Spain	435	38.971645	-0.358844	15.3	553.7	30.3	3.8	777	120	9	52.1
Mimizan	France	19	44.134167	-1.303056	13.2	490.9	24.9	3.1	1235	149	62	24.3
Pinia	France	15	42.021083	9.464861	15.6	515.5	27.0	6.3	585	80	8	49.6
Tabarka	Tunisia	121	36.958397	8.703792	17.7	573.4	31.8	6.7	916	162	4	72.3
Tamrabta	Morocco	1758	33.600000	-5.016667	11.8	665.3	31.3	-3.2	721	96	8	54.4
Sidi Meskour	Morocco	1931	31.439375	-6.903864	11.4	711.3	32.5	-5.2	514	71	5	56.4
P. halepensis												
Cabanellas	Spain	210	42.235556	2.790000	14.8	575.5	27.4	3.9	713	97	35	26.0
Carratraca	Spain	650	36.841111	-4.834444	15.7	568.6	30.6	4.4	693	106	3	69.4
S'avall	Spain	10	39.287222	3.047778	16.8	525.9	28.7	6.8	566	92	5	54.7
Tarrasa	Spain	117	41.466667	2.100000	15.9	551.1	27.7	5.8	619	83	29	29.5
Imperia	Italy	109	43.900000	8.050000	15.1	550.5	27.5	4.6	804	107	18	38.5
Zaouia Ifrane	Morocco	1512	33.570000	-5.140000	11.5	668.7	31.2	-3.8	849	123	8	57.4
Aures Beni Melloul	Algeria	936	35.166667	6.833333	13.4	657.2	31.2	-0.1	375	53	10	36.1
Elea	Greece	155	37.766667	21.533333	16.7	592.0	30.8	5.4	808	150	6	77.1
Shaharia	Israel	236	31.600000	34.833333	20.0	496.4	31.8	8.0	397	100	0	116

AMT: annual mean temperature; TS: temperature seasonality (STD *100); MTWM: maximum temperature of the warmest month; MTCM: minimum temperature of the coldest month; AP: annual precipitation; PWM: precipitation of the wettest month; PDM: precipitation of the driest month; PS: precipitation seasonality (CV *100).

Materials and Methods

Study species

Pinus pinaster Ait.: Maritime pine populations have been assessed using various molecular markers: chloroplast and mitochondrial (e.g. Burban and Petit 2003; Bucci *et al.* 2007), as well as nuclear (e.g. Salvador *et al.* 2000; Eveno *et al.* 2008). In particular, chloroplast markers were able to identify different gene clusters related to the history of the species (Bucci *et al.* 2007). Because of its economic importance, various programs of genetic improvement have been developed in *Pinus pinaster*, in particular for the Atlantic provenances. Within this framework, several adaptive traits, such as growth, tolerance to drought and cold, and resistance to pests and diseases, have been the subject of genetic variability studies. Numerous approaches have been used to unravel the basis of quantitative traits in maritime pine: genetic (genetic cartography), physiologic (mechanisms implicated in traits), functional and structural genomic (candidate genes and proteins) and population genetic and genomic (genes under selection).

Pinus halepensis Mill.: Aleppo pine genetic variability has been studied with both chloroplast and nuclear markers (Bucci *et al.* 1998; Morgante, Felice, and Vendramin 1998; Grivet *et al.* 2009). A recent study revealed that the pattern of polymorphism observed in some candidate genes related to drought-tolerance in this species reflected long-range colonization and possibly natural selection during range expansion (Grivet *et al.* 2009). Various common garden studies have assessed Aleppo pine intraspecific variability in order to study the role of ecological factors in shaping adaptive strategies. Some experiments revealed adaptive variation to climate (total precipitation and dry season duration; Voltas *et al.* 2008) or reproductive features (Climent *et al.* 2008), highlighting thus the selective role of climate variables in determining population and family fitness in this species.

Sampling

Twelve populations of *P. pinaster* (77-122 individuals, depending on the gene; Table S1) and nine populations of *P. halepensis* (72-93 individuals; Table S1) were collected spanning the full range of each species (Figure 1, Table 1). Populations were selected considering not only spatial distribution but also environmental heterogeneity in both species, prioritizing populations that represent contrasted environments (Table 1). Our sampling covers also different soil (siliceous, calcareous) and Mediterranean forest types. Finally, we included representations of the five traditional varieties or landraces described in maritime pine (Resch 1974): *mesogeensis* (e.g. Olba), *atlantica* (e.g. Mimizan), *corteensis* (e.g. Pinia), *maghrebiana* (e.g. Tamrabta) and *renoui* (e.g. Tabarka), as well as all genecological groups normally considered in Aleppo pine: Western Europe, Eastern Europe and Northern Africa.

From each population, cones were collected from mother trees separated by at least 50 m without any phenotypic selection. Seeds from each mother tree were kept in individual paper bags and stored at 4°C in a dry environment till DNA extraction (see below).

TABLE 2

Gene diversity (all sites) for 12 amplicons from six putative candidate genes across the twelve *P. pinaster* populations and the nine *P. halepensis* populations

Amplicon	*P. pinaster*							*P. halepensis*						
	N	*L*	*S*	Θ_π	Θ_w	*K*	*He*	*N*	*L*	*S*	Θ_π	Θ_w	*K*	*He*
lp31-Pt														
a	122	456	12	6.26	4.89	6	0.717	79	353	3	2.48	1.72	4	0.460
b	111	560	7	3.91	2.37	3	0.446	74	488	7	3.93	2.94	4	0.464
lp33-Pp	97	449	9	4.29	3.89	9	0.798	90	375	1	0.17	0.53	2	0.065
dhn2-Pp														
a	120	472	12	6.06	4.74	11	0.807	89	448	2	0.20	0.88	3	0.088
b	85	596	21	8.42	7.03	11	0.811	93	346	4	0.55	2.26	5	0.105
dhn2-Ps														
a	92	743	27	8.10	7.13	14	0.899	na	na	na	na	na	na	na
b	80	513	15	5.92	5.90	14	0.855	92	457	11	1.93	4.73	6	0.204
dhn5-Ps	77	449	4	4.00	1.81	5	0.735	na	na	na	na	na	na	na
4cl-Pt														
a	91	551	13	5.67	4.64	4	0.678	88	461	10	5.88	4.30	13	0.805
b	92	256	7	3.56	5.37	5	0.291	79	688	21	10.67	6.18	13	0.793
c	87	543	4	0.69	1.46	4	0.341	78	528	3	2.47	1.15	5	0.697
d	na	na	na	na	na	na	na	72	452	3	2.80	1.37	4	0.484
Total	*1054*	*5588*	*131*	-	-	*86*	-	*834*	*4596*	*65*	-	-	*59*	
Mean	-	-	-	*5.17*	*4.48*	-	*0.671*	-	-	-	*3.11*	*2.61*	-	*0.416*

N: number of sequences; L: total analyzed length in base pair; S: polymorphic sites; Θ_π Tajima's nucleotide diversity × 10-3 (Tajima 1989) and Θ_w: Watterson's nucleotide diversity (Watterson 1975) per site (× 10-3); K: number of haplotypes; He: haplotype diversity; na: amplicons that failed to transfer across species.

Spatial, environmental and genetic data

Spatial, environmental and genetic data. Three spatial variables were recorded for each population: altitude, latitude and longitude (Table 1). Eight climatic variables were considered: annual mean temperature (AMT), temperature seasonality (expressed as the standard deviation across months multiplied by 100; TS), maximum temperature of the warmest month (MTWM), minimum temperature of the coldest month (MTCM), annual precipitation (AP), precipitation of wettest month (PWM), precipitation of driest month (PDM), and precipitation seasonality (coefficient of variation; PS). Climatic data for *P. pinaster*'s Iberian populations were obtained from a functional phytoclimatic model based on raw data from meteorological stations (Gonzalo 2007). Climatic data for *P. pinaster*'s non-Iberian populations as well as for all *P. halepensis*' populations were obtained from the WorldClim – Global Climate Data at 5 minute resolution (Hijmans *et al.* 2005) (Table 1). Graphical pairwise correlations between these 11 spatial and environmental variables are presented in Figure S1.

DNA extraction and candidate gene sequencing. Genomic DNA from *P. pinaster* (haploid) megagametophytes was extracted with a modified Dellaporta *et al.* (1983) protocol. DNA extraction for *P. halepensis* was carried out as reported in Grivet *et al.* (2009).

Candidate genes related to drought tolerance were originally identified on the basis of functional studies performed in *Pinus taeda* and other conifers, or derived from model species such as *Arabidopsis thaliana*, as described elsewhere (González-Martínez *et al.* 2006; Eveno *et al.* 2008; Grivet *et al.* 2009; Wachowiak, Balk, and Savolainen 2009). Altogether the candidate genes selected for this study belong to three well-known and relatively small multigene families: the ASR (*lp31-Pt and lp33-Pt*), dehydrin (*dhn2-Pp, dhn2-Ps and dhn5-Ps*) and *4-coumarate: CoA ligase* (*4cl-Pt*) families. The ASR family is named after the Abscisic acid (ABA), stress and ripening response (Frankel *et al.* 2006). These proteins are also induced by water deficit stress (WDS). The dehydrin gene, *dhn2-Pp*, previously described in maritime pine (Eveno *et al.* 2008), is not orthologous to the one described in Scots pine, i.e. *dhn2-Ps* (Wachowiak *et al.* 2009).

Previously published primer pairs of these putative candidate genes were transferred from *P. tadea* (Pt), *P. sylvestris* (Ps) or *P. pinaster* (Pp) to either *P. pinaster* or *P. halepensis* or both, and three new primer pairs, which overlap with pre-described amplicons, were also designed in order to extend the coverage of these target genes (Table S2). PCR conditions for the two pines studied here are given in Table S3. Outgroup sequences for each amplicon were obtained from GenBank: *P. taeda* for amplicons *lp31-Pt, lp33-Pt, dhn2-Pp* and *4cl-Pt*, and *P. sylvestris* for *dhn5-Ps* and one *dhn2-Ps* amplicon; or produced with our newly designed primers (*P. nigra* for the second *dhn2-Ps* amplicon) (Table S2). Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. HM481479-HM483364.

PCR products were purified and sequenced from both ends following standard protocols (for *P. pinaster*, see Eveno *et al.* 2008; for *P. halepensis*, see Grivet *et al.* 2009). Multiple sequence alignments and manual adjustments were done using SeqMan v7 (DNASTAR Lasergene software) and BIOEDIT. (http://www.mbio.ncsu.edu/BioEdit/page2.html).

Chloroplast microsatellites. To control for associations between candidate genes and environmental data that could be due to neutral processes, we included variation of chloroplast microsatellites (the only neutral genetic markers available for the two species at the wide range scale) as a covariate in the multivariate logistic regressions (see *Environmental associations* section below). We performed population-based Principal Component Analysis (PCA) on chloroplast markers, and kept the

FIGURE 2

Predicted and observed patterns of clinal variation at *4cl-Pt_c*. Sigmoid curves denote the clines in either allelic (SNP position given) or haplotypic (haplotypes represented by capital letters) frequencies under a logistic regression model. The logistic regression model uses a climatic variable (MTCM or TS) after controlling for neutral processes (i.e. using the PCs of neutral molecular markers as covariates) for *P. pinaster* and *P. halepensis*. MTCM: minimum temperature of the coldest month; TS: temperature seasonality.

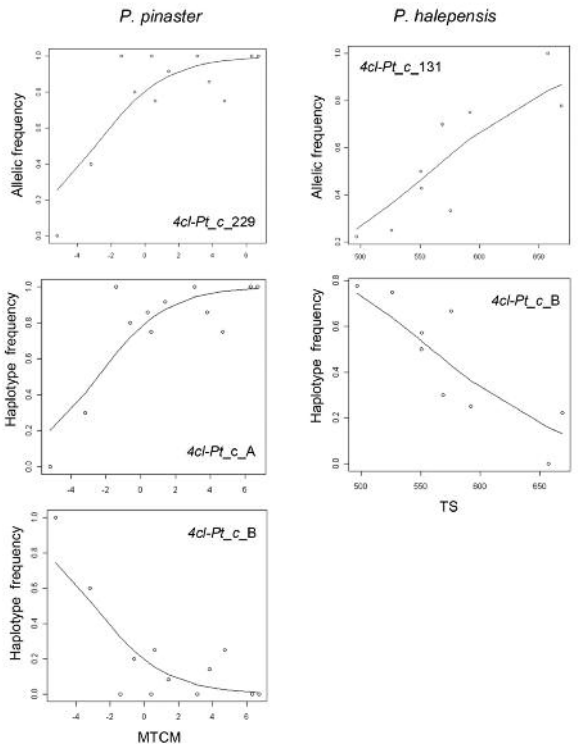


TABLE 3a

Polymorphism, divergence and neutrality tests (silent sites) in *P. pinaster* for the phylogeographical groups defined with chloroplast microsatellites (see text for details).

Amplicon	*S*	*D*	H_{norm}	*DHEW P-value*	*ML-HKA's K (P. taeda)*	K_s *(P. taeda)*	*ML-HKA's K (P. halepensis)*	K_s *(P. halepensis)*
Western Mediterranean [a]								
lp31-Pt								
a	8	2.172	-0.635	0.958	0.961	38.350	1.292	20.310
b	6	1.834	-1.380	0.913	0.698	44.250	0.720	24.360
lp33-Pp	6	1.133	0.911	0.984	0.696	60.060	1.839	15.990
dhn2-Pp								
a	7	0.269	-0.853	0.406	1.605	38.160	1.307	31.580
b	19	0.186	0.825	0.727	3.256 *	31.940	1.975	39.720
dhn2-Ps								
a	15	0.631	0.703	0.562	ptna	ptna	phna	phna
b	8	-0.126	0.232	0.209	ptna	ptna	d	d
dhn5-Ps	2	1.960						
	0.136	0.907	ptna	ptna	phna	phna		
4cl-Pt								
a	2	0.422	0.067	0.610	0.307	31.480	0.250	26.870
b	2	-1.041	-3.173 *	0.014 *	0.741	43.930	0.446	41.220
c	1	0.214	0.565	0.663	0.542	23.080	0.000	1.190
d	na	na	na	na	na	na	na	na
North Africa [b]								
lp31-Pt								
a	7	-0.276	-3.065 *	0.336	0.878	39.300	0.952	22.280
b	6	-2.046 *	-3.813 **	0.073	0.820	44.250	0.880	26.150
lp33-Pp	5	0.014	0.926	0.899	0.614	57.860	1.946	13.040
dhn2-Pp								
a	4	0.799	-0.554	0.615	0.868	36.610	0.792	28.620
b	13	1.095	0.361	0.718	2.036	32.780	1.362	33.600
dhn2-Ps								
a	6	0.190	-0.232	0.322	ptna	ptna	phna	phna
b	6	-1.116	-1.223	0.085	ptna	ptna	d	d
dhn5-Ps	2	0.688	-0.809	0.540	ptna	ptna	phna	phna
4cl-Pt								
a	8	2.797	0.411	0.995	1.279	39.160	0.442	28.880
b	3	2.266	0.090	0.982	1.303	48.190	1.017	34.730
c	2	0.201	0.387	0.494	1.039	25.240	0.000	3.410
d	na	na	na	na	na	na	na	na

D: Tajima's *D* (Tajima 1989); H_{norm}: Fay and Wu's normalized *H* (Zeng *et al.* 2006); *DHEW* test: compound test (Zeng *et al.* 2007); *ML-HKA's K*: Selection parameter of the Maximum-Likelihood multilocus Hudson-Kreitman-Aguadé neutrality tests (Wright and Charlesworth 2004); the species used as outgroup is given in parenthesis; *S*: number of segregating sites; K_s: average proportion of nucleotide differences between species per silent site $\times 10^{-3}$.

na: amplicons that failed to transfer across species; ptna: *P. taeda* outgroup sequence not available; phna: *P. halepensis* outgroup sequence not available; d: discarded because possible paralog.

P-values [*P* neutrality test (neutral) ≤ P neutrality test (observed)]: 0 < *** < 0.001 < ** < 0.01 < * < 0.05

[a] Western Mediterranean: Arenas de San Pedro, Cazorla, Coca, Cómpeta and Olba.

[b] North Africa: Tabarka, Tamrabta and Sidi Meskour.

Principal Components (PCs) that explained the majority of the overall inertia of the data. For *P. pinaster*, the three first PCs of the 16 most common haplotypes (determined on the basis of five chloroplast microsatellites) – accounting for 74% of the overall variation – were extracted from the lattice dataset (430 records X 16 variables) of Bucci *et al.* (2007). For populations that did not match any of the 430 locations in this lattice, the mean of the closest 3-4 locations was computed (Table S4). For *P. halepensis*, PCs were computed in R (R Development Core Team) for each of the nine populations based on the 22 haplotypes determined from three chloroplast microsatellites (our unpublished data) and the three first PCs, which explain 82% of the overall variation, were used for subsequent analysis (Table S4).

Chloroplast microsatellites were also used to determine groups with similar evolutionary history to conduct neutrality tests (Figure 1). A previous study based on chloroplast microsatellites pointed at eight gene pools in *P. pinaster* (Bucci *et al.* 2007). Within the sampling used in the present study, Mimizan (Continental French lineage) and Pinia (Northern Italy and Corsican lineages) represent single lineages. Tabarka (Tunisia), Tamrabta (Moroccan Middle Atlas) and Sidi Meskour (Moroccan High Atlas) are all part of a North African lineage that combines western and eastern African origins. The rest of *P. pinaster* populations sampled are part of a wide Western Mediterranean linage, except for Quatretonda and Oria that are considered marginal populations and may thus present a distinct evolutionary history (González-Martínez *et al.* 2007a; Eveno *et al.* 2008). In *P. halepensis*, eastern

TABLE 3b

Polymorphism, divergence and neutrality tests (silent sites) in *P. halepensis* for the phylogeographical groups defined with chloroplast microsatellites (see text for details)

Amplicon	*S*	*D*	H_{norm}	*DHEW P-value*	*ML-HKA's K (P. taeda)*	K_s *(P. taeda)*	*ML-HKA's K (P. halepensis)*	K_s *(P. halepensis)*
Western Mediterranean [a]								
lp31-Pt								
a	2	1.689	-1.714	0.860	0.597	34.590	0.452	20.960
b	4	0.535	-0.948	0.501	0.889	48.280	1.380	23.090
lp33-Pp	1	-0.629	-4.343 *	0.270	0.145 *	69.100	0.361	14.020
dhn2-Pp								
a	0	nps	nps	nps	nps	55.770	nps	44.020
b	0	nps	nps	nps	nps	57.180	nps	38.220
dhn2-Ps								
a	na	na	na	na	na	na	na	na
b	2	-0.595	0.317	0.382	ptna	ptna	d	d
dhn5-Ps	na	na	na	na	na	na	na	na
4cl-Pt								
a	6	0.752	-1.222	0.599	2.236	43.260	1.618	22.500
b	16	2.386	-0.320	0.985	2.279	52.000	1.242	30.660
c	2	1.701	-1.305	0.867	1.350	31.970	2.976	9.440
d	3	2.324	-0.981	0.965	0.674	54.780	ppna	ppna
North Africa [b]								
lp31-Pt								
a	2	1.818	-1.053	0.898	0.819	34.140	1.123	12.710
b	2	1.882	0.602	0.909	0.530	49.000	0.924	22.700
lp33-Pp	0	nps	nps	nps	nps	69.410	nps	9.550
dhn2-Pp								
a	1	-0.774	0.289	0.313	0.329	56.280	0.525	19.310
b	1	-1.147	0.160	0.189	0.215	57.320	0.205	33.600
dhn2-Ps								
a	na	na	na	na	na	na	na	na
b	0	nps	nps	nps	ptna	ptna	d	d
dhn5-Ps	na	na	na	na	na	na	na	na
4cl-Pt								
a	4	1.306	-0.201	0.791	2.072	44.030	1.815	22.450
b	14	2.316	0.044	0.981	3.251 *	50.760	1.951	31.290
c	2	1.077	0.221	0.834	1.665	27.720	3.309	5.200
d	3	-1.494	-6.584 ***	0.009 **	0.860	60.020	ppna	ppna

D: Tajima's *D* (Tajima 1989); H_{norm}: Fay and Wu's normalized *H* (Zeng *et al.* 2006); *DHEW* test: compound test (Zeng *et al.* 2007); *ML-HKA's K*: Selection parameter of the Maximum-Likelihood multilocus Hudson-Kreitman-Aguadé neutrality tests (Wright and Charlesworth 2004); the species used as outgroup is given in parenthesis; *S*: number of segregating sites; K_s: average proportion of nucleotide differences between species per silent site × 10-3.

na: amplicons that failed to transfer across species; ptna: *P. taeda* outgroup sequence not available; ppna: *P. pinaster* outgroup sequence not available; d: discarded because possible paralog; nps: monomorphic fragment.

P-values [*P* neutrality test (neutral) ≤ *P* neutrality test (observed)]: 0 < *** < 0.001 < ** < 0.01 < * < 0.05

[a] Western Mediterranean: Cabanellas, Carratraca, Imperia, S'avall and Tarrasa.

[b] North Africa: Aures Beni Melloul and Zaouia Ifrane.

Mediterranean populations (Elea in Greece and Shaharia in Israel) present a higher level of cpSSR diversity and substantial genetic differentiation from the rest and are considered single-population lineages (Grivet *et al.* 2009); Zaouia Ifrane (Morocco) and Aures Beni Melloul (Algeria) form part of the North African group, while all four Spanish populations and the Italian population of Imperia belong to the Western Mediterranean group. These regions (North Africa and Western Mediterranean) define relatively homogeneous zones of the Mediterranean Basin in terms of soil and climate (Barbéro *et al.* 1998).

Statistical analyses

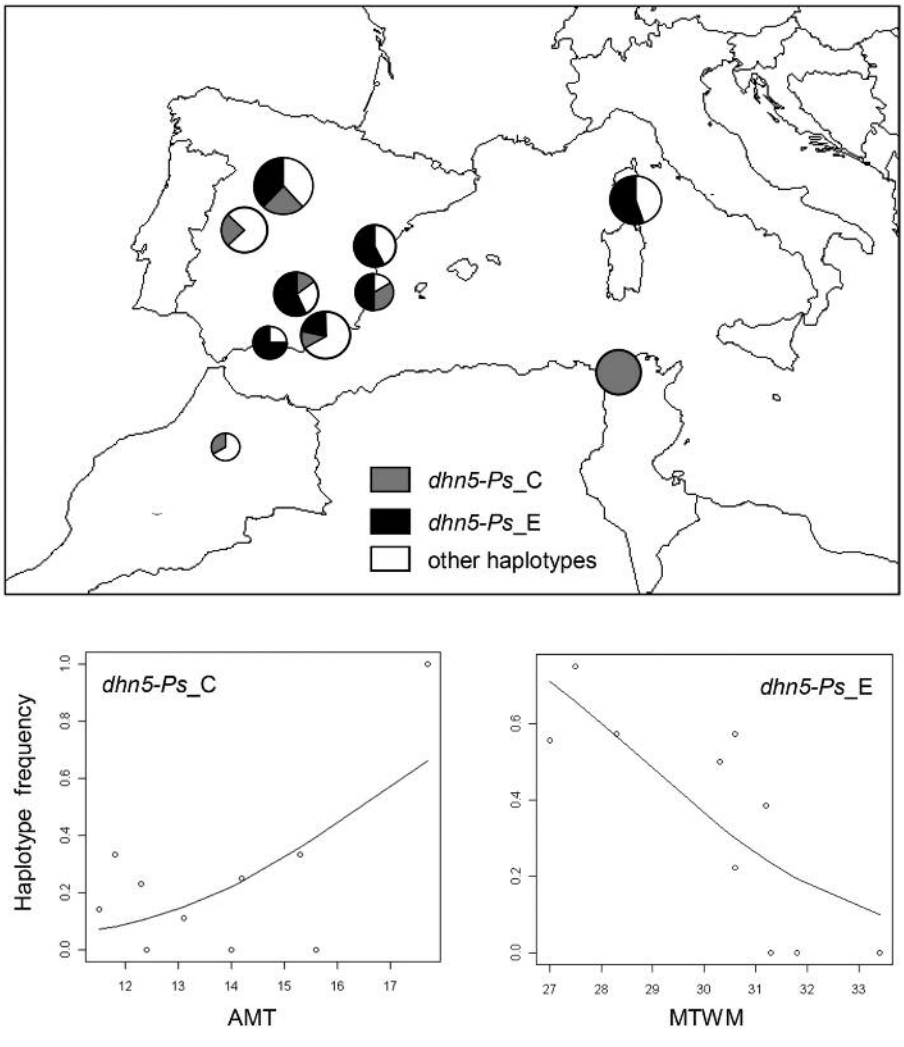
Gene diversity and divergence. Number of segregating sites (*S*), nucleotide diversity statistics ($\theta\pi$, Tajima 1989; θw, Watterson 1975), number of haplotypes (*K*), and haplotypic diversity (H_e) were computed for both North African and Western Mediterranean groups using scripts kindly provided by S.E. Ramos-Onsins (Department of Genetics, Faculty of Biology, University of Barcelona, Spain) and the program DnaSP v5 (Librado and Rozas 2009). Average divergence per site (K_{all}) for each geographical group was computed between each of the pines and the outgroup *Pinus taeda*, as well as between *P. pinaster* and *P. halepensis* themselves, using scripts also provided by S.E. Ramos-Onsins.

Neutrality tests. All neutrality tests were performed at the regional level (see group definition above) considering all sequenced gametes in each amplicon, using one sequence from another pine species as outgroup when appropriate, and considering all substitutions as well as only silent sites (except for the *MK* test where all sites must be used, see

below). First, we conducted neutrality tests based on within-species population genetic data with (to determine ancestral states) or without outgroup, including Tajima's D test (Tajima 1989), Fay and Wu's *H* normalized test (Zeng *et al.* 2006), both based on the site-frequency spectrum (SFS), Ewens-Watterson *EW* homozygosity test (Watterson 1978), and the *DHEW* neutrality test (Zeng, Shi, and Wu 2007) – a compound test that combines the properties of *D, H* normalized and *EW* tests – using scripts provided by K. Zeng (State Key Laboratory of Biocontrol and Key Laboratory of Gene Engineering of the Ministry of Education, Sun Yat-sen University, Guangzhou, China). Second, we used methods that compare patterns of polymorphism within species and divergence between species. The McDonald-Kreitman (*MK*) test, which compares the ratio of nonsynonymous and synonymous mutations between and within species (McDonald and Kreitman 1991), was computed using DnaSP v5 (Librado and Rozas 2009). *ML-HKA* (Maximum-Likelihood multilocus Hudson-Kreitman-Aguadé) neutrality tests were computed using the mlhka program (Wright and Charlesworth 2004). This test is an extension of the *HKA* test that allows to test specific locus or groups of loci against a group of neutrally evolving genes for their levels of polymorphism and divergence. First, each amplicon was compared to all the rest and the selection parameter *K* was computed. The *K* parameter measures the degree to which diversity is increased or decreased with respect to

FIGURE 3

Upper section: *dhn5-Ps* haplotype distribution showing the proportion of haplotypes C (in grey), E (in black) and all others (in white) in *P. pinaster*. The size of the pie is proportional to the sample size at each location. ***Lower section:*** expected (sigmoid curves) and observed covariation in haplotype frequencies of *dhn5-Ps*_C and *dhn5-Ps*_E and climatic variables. Expected curves were obtained under a logistic regression model including a single environmental variable (ATM or MTWM) after controlling for neutral processes (i.e. using the PCs of neutral molecular markers as covariates) for *P. pinaster.* AMT: annual mean temperature; MTWM: maximum temperature of the warmest month.



divergence. Second, for significant tests, the analysis was repeated considering only amplicons from different genes in order to avoid potential linkage. All runs consisted in 100,000 cycles of the Markov chain; long runs of 1,000,000 cycles were also computed for significant tests as well as different starting values for the divergence time parameter (*T*). However, no differences were found in the results and only basic runs are presented here. Standard likelihood ratio tests (LRTs) were used to compare among models considering one amplicon under selection at a time and the null model of no selection (degrees of freedom equal one). The *ML-HKA* test was performed using different outgroups to gain insights on evolutionary times at which positive selective wave(s) may have taken place: i) with *P. taeda* as outgroup for both *P. pinaster* and *P. halepensis* amplicons. In this case, the time frame would correspond to the split between the *Trifoliae* section (to which *P. taeda* belongs) and the *Pinus* section (to which the Mediterranean pines belong) in the Oligocene (~25 Ma; Gernandt *et al.* 2005; Gernandt *et al.* 2008). Three amplicons were discarded from these analyses because they had *P. sylvestris* or *P. nigra* as outgroups (two from *dhn2-Ps*, and one from *dhn5-Ps*); ii) with *P. halepensis* as outgroup for *P. pinaster* amplicons and *P. pinaster* as outgroup for *P. halepensis* amplicons. The time frame, in this case, would correspond to the diversification of Mediterranean pines (in the Miocene, ~10 Ma; Gernandt *et al.* 2008). One amplicon was discarded from this series of tests (*dhn2-Ps_b*; see Table S2 for amplicon nomenclature) due to bad alignment between species and, thus, the possible amplification of a paralogous fragment in *P. halepensis* (*dhn2-Ps_b* from *P. halepensis* had more similarity with the outgroup, *P. nigra*, than with its close relative *P. pinaster*).

Environmental associations. We tested if SNP-allele or haplotype frequencies at candidate loci correlated with climatic variables using logistic regression. We first carried out series of univariate logistic regressions to test for association between SNP/haplotype frequencies and environmental variables using the program SAM (Joost, Kalbermatten, and Bonin 2008). We considered a correlation as significant only when two likelihood ratio tests (G and Wald tests) rejected the null hypothesis of no association between the genetic and the environmental variables (at the 5% level). A strict Bonferroni correction was applied to correct for multiple testing of univariate models.

In a second step, whenever a SNP/haplotype was found significantly correlated with an environmental variable, we performed a multivariate logistic regression using neutral marker PCs (see above) as covariates (together with the SNP and the environmental variables) to control for neutral processes (e.g. postglacial migrations, geographical isolation) that may have also generated clines in the absence of any local adaptation. Multiple logistic regressions were performed in *R* (R Development Core Team) using the *glm* function (assuming that allelic counts at SNPs were binomially distributed) for each SNP/haplotype separately. Model selection was carried out using the *drop1* function and likelihood ratio tests (LRTs). Whenever the models compared were not nested (and LRTs were thus not appropriate), we used Akaike's Information Criterium (AIC). Tests for environmental associations were performed at the amplicon level (counting the SNPs located within overlapping regions only once) and including only populations with more than three gametes sampled (average number of gametes sampled per population: 7.98 in *P. pinaster* and 9.27 in *P. halepensis*, see Table S1).

Results

Gene diversity and divergence

Overall primer transfer rates across the two pine species were high (91.6% for *P. pinaster* and 83.3% for *P. halepensis*) and we obtained a total of 533 kbp and 380 kbp of sequence in *P. pinaster* and *P. halepensis*, respectively. All sequences represented putative orthologs (determined on the basis of similarities with available sequences in GenBank and construction of phylogenies including multiple gene members of the target families in different conifers) except probably for *dhn2-Ps_b* in *P. halepensis* (see Methods). Maritime pine showed a higher level of gene diversity than Aleppo pine (Table 2), and this result held when considering only the nine orthologous amplicons in the two species (data not shown). In total, 131 SNPs (24 non-synonymous) were detected in *P. pinaster* (4 to 27 per amplicon) and 65 SNPs (12 non-synonymous) in *P. halepensis* (1 to 21 per amplicon) (Table 2). There were only four shared polymorphisms (i.e. 2 % of polymorphic sites were shared) across species, with two of them (those within *4cl-Pt*) appearing only in one *P. pinaster* individual. Number of insertions/deletions was similar between the two species for the same locus, but these indels were not shared across species. None of them caused frame shifts. At the haplotype level, not only the number of haplotypes substantially differed between the two species (86 for *P. pinaster* vs. only 59 for *P. halepensis*) but also their frequency spectrum: *P. pinaster* had multiple common haplotypes along with some less frequent ones, while *P. halepensis* generally displayed one major haplotype together with some low-frequency ones (Figure S2). Finally, there was also a contrasted pattern of gene diversity between the regional groups defined by chloroplast microsatellites in *P. pinaster*: the Western Mediterranean group displayed higher nucleotide diversity than the North African group (average per population of 0.0039 vs. 0.0022). The SNP and haplotypes detected were used to conduct neutrality tests, as well as to detect associations with environmental variables.

Divergence at the DNA sequence level was examined using two outgroups: *P. taeda*, a New World pine, and each

of the Mediterranean pines as reference for the other (i.e. *P. halepensis* was used as outgroup for *P. pinaster* and *vice versa*). For both species, nucleotide divergence per site with *P. taeda* (K_{all} = 0.0397 for *P. pinaster*; K_{all} = 0.0498 for *P. halepensis*) was higher than with the other Mediterranean species (K_{all} = 0.0235 for *P. pinaster* vs. *P. halepensis*) except for *dhn2-Pp_b* in *P. pinaster* that displayed a slightly higher divergence between the Mediterranean pines than between each of them with *P. taeda.*

Neutrality tests

Neutrality tests rejected the null neutral model for distinct genes depending on the species and the geographic group (Table 3). The *D* and H_{norm} tests identified some genes departing significantly from expectation under the neutral model that are not detected by the more robust *DHEW* test: *lp31-Pt_a* and *lp31-Pt_b* in the North African group for *P. pinaster*; *lp33-Pp* in the Western Mediterranean group for *P. halepensis*. As shown below, these two tests are probably reflecting demographical events and not selective processes. More relevantly, two fragments of the *4cl-Pt* gene were identified as potential targets of selection by *DHEW* tests: *4cl-Pt_b* in the Western Mediterranean group for *P. pinaster* and *4cl-Pt_d* in the North African group for *P. halepensis*. Considering all polymorphic sites or only silent sites did not change the qualitative results of the test (Table 3 and Table S5).

None of the *MK* tests showed any departure from the null hypothesis of an equal ratio of nonsynonymous to synonymous variation within and between species (data not shown). The *ML-HKA* tests based on silent sites (with *P. taeda* as outgroup) revealed uncoupled patterns of polymorphism and divergence in only one amplicon in *P. pinaster* in the Western Mediterranean group (*dhn2-Pp_b*), while in *P. halepensis* one amplicon in this same group (*lp33-Pp)* and one in the North African group (*4cl-Pt_b*) were significant (Table 3). These results still stood when analyses were restricted to only one amplicon per gene (in order to avoid potential linkage between fragments from the same gene) and when fragments potentially under selection were not used as control loci (data not shown). When all substitution sites were included in the analyses, three more genes were significant for *P. halepensis* in the Western Mediterranean group (*dhn2-Ps_a, 4cl-Pt_a, 4cl-Pt_b*) (Table S5). In contrast, when Mediterranean pines were used as outgroup for each other, patterns of polymorphism and divergence either considering all sites (Table S5) or silent sites (Table 3) seemed to evolve neutrally and a similar trend of the maximum likelihood estimate of the selection parameter *(ML-HKA's K)* was observed in the two series of analyses.

Environmental associations

Environmental associations were examined at both the SNP and the haplotype level, and, overall, similar associations were detected. Nonetheless, a few correlations differed, with some found only at the haplotype level, which highlights how these two levels complement each other as haplotypes may reflect interactions among linked mutations. Within the two species, all significant associations detected between SNPs or haplotypes and climatic variables involved temperature indices as environmental variables.

As many as 23 significant associations were initially found in *P. pinaster* (non-corrected model as provided by SAM, see Methods). Only three associations remained after integrating neutral marker PCs as covariates, two of them with spatial variables (one with altitude, one with latitude) and one (*4cl-Pt_c*_229) with the minimum temperature of the coldest month (MTCM) (Figure 2; Table S6). At the haplotype level, initially 32 associations were found in maritime pine but only six remained in the corrected model: two (*4cl-Pt_c*_A *and 4cl-Pt_c*_B) with altitude and one (*dhn5-Ps*_C) with annual mean temperature (AMT), one (*dhn5-Ps_E*) with maximum temperature of the warmest month (MTWM), and two (*4cl-Pt_c*_A and *4cl-Pt_c*_B) with minimum temperature of the coldest month (MTCM) (Figure 2; Table S7). In the case of *P. halepensis*, only one association was found by the corrected models, both at the SNP (between *4cl-Pt_c*_131 and temperature seasonality, TS) and at the haplotype (*4cl-Pt_c*_B and TS) level (Figure 2; Tables S6 and S7).

In summary, association analysis identified two loci exhibiting significant correlations with temperature indices: *4cl-Pt_c* that is associated with MTCM in *P. pinaster* and with TS in *P. halepensis* at both the SNP and haplotype levels, and *dhn5-Ps* that is associated with AMT and MTWM at the haplotype level in *P. pinaster* only.

Discussion

In this study, we assessed the impact of natural selection on the same set of candidate genes related to drought tolerance in two widespread Mediterranean pine species. Our results revealed distinct selection patterns according to species, geographic regions and loci. Below, we discuss these findings in the light of the history of each species and the specificities of each of the methods used to reveal footprints of selection.

Neutrality tests

Neutrality tests examining selection events based on site- and haplotype-frequency spectrum identified distinct genes potentially targeted by selection. The *D* and/or H_{norm} tests detected three loci departing from neutrality that were not detected by the *DHEW* compound tests. These results may be explained by the sensitivity of the *D* and H_{norm} tests to demographic factors and different degrees of background selection (Zeng *et al.* 2006). Especially, Tajima's *D* is sensitive to background selection and population growth while the H_{norm} is more sensitive to population shrinkage and subdivision. In a previous

study in *P. halepensis* based on ten candidate genes, we have shown that this species has undergone historical bottlenecks and that the western populations of the species harbored some signatures of this demographic event (Grivet *et al.* 2009). In addition, based on the present data set and extensive coalescent simulation, we have found that observed values of *D* and H_{norm} for *P. pinaster* both in the Western Mediterranean and North African groups reject the standard neutral model and suggest past bottlenecks in this species too (see Figure S3). To test the robustness of the *DHEW* compound test, neutral coalescent simulations under realistic bottlenecks scenarios were further simulated in *P. pinaster* (1,000 coalescent simulations per scenario) and the significant values for this test were recorded. In all cases, the number of significant tests obtained was equal or lower than the number expected by chance (see Figure S3), highlighting the robustness of the *DHEW* test in the presence of bottleneck events and geographic population structure of magnitudes similar to the ones found in maritime and Aleppo pines. In addition to our simulations, the *DHEW* test has been shown to be robust to recombination and to have high sensitivity to detect positive selection (Zeng *et al.* 2007), as it combines powerful (and insensitive to recombination) haplotype-frequency spectrum tests (i.e the *EW* test) with SFS statistics, such as the *D* test that maintains power to detect positive selection across a wide period of time (Zeng *et al.* 2006; Zhai, Nielsen and Slatkin 2009). Only two on-going events of selection were detected by the *DHEW* test, one in *P. pinaster* and one in *P. halepensis*. Both cases involved the *4cl-Pt* gene (albeit in different geographical regions and distinct amplicons), suggesting a potential role of *4cl* in local adaption in both pine species (see below).

Among the tests examining polymorphism within species and divergence among species, the *MK* test did not detect any loci under selection. While this test is somewhat robust to demography, it seems less powerful in detecting positive selection than *HKA* type tests (Zhai, Nielsen, and Slatkin 2009). The *ML-HKA* tests for both silent and all sites, using *P. taeda* as outgroup, identified three amplicons with uncoupled levels of polymorphism and divergence in the two pines. Two genes (*dhn2-Pp_b* in *P. pinaster* in the Western Mediterranean group and *4cl-Pt_b* in *P. halepensis* in the North African group) showed high levels of diversity compared to divergence (selection parameter higher than one), a pattern compatible with balancing selection. One other gene (*lp33-Pp* in the Western Mediterranean group in *P. halepensis*) showed low diversity compared to divergence (selection parameter lower than one), which could reflect the transient reduction in variability occurring during a selective sweep. Three other genes showed uncoupled levels of polymorphism and divergence in *P. halepensis* when examining all sites, but not when examining silent sites alone. Two of these genes (*4cl-Pt_a* and *4cl-Pt_b* in the Western Mediterranean group) had an excess of nucleotide diversity, which could reflect recent balancing selection acting on non-synonymous sites, while for one gene (*dhn2-Pp_a* in the Western Mediterranean group) purifying selection reducing diversity in non-synonymous sites could explain the observed pattern. Evidence for extensive purifying selection in conifers comes from an approximately four-fold nucleotide diversity at synonymous compared to non-synonymous sites in most species (see Table 1 in González-Martínez *et al.* 2010), including maritime (dN/dS = 0.169) and Aleppo (dN/dS = 0.344) pines (this study). Although demographical events can result in a significant *ML-HKA* test, its multilocus nature combined with the *HKA* framework should produce a more robust test than those based on comparing different aspects of polymorphism at a single locus (Wright and Charlesworth 2004).

In order to gain power by increasing sample size while taking the specific history of each species into account, we performed the neutrality tests in groups of populations that present similar evolutionary history (i.e. North African and Western Mediterranean groups). However, grouping populations may bias the output of the SFS-based neutrality tests as it can increase the proportion of singletons, leading to an excess of low frequency variants and thus to more negative *D* (Städler *et al.* 2009). In our study, populations were grouped according to homogeneous gene pools (see Bucci *et al.* 2007 for *P. pinaster* and Grivet *et al.* 2009 for *P. halepensis*), the consequence of which should be that no substantial differences in the number of rare variants at the population and group levels are found. The extensive gene flow normally found in conifers with large and continuous distribution would also support this approach. Nevertheless, to check that our results were not biased because of population grouping, we performed neutrality tests at the population level too and found that values of *D* test at the group level were not more negative than at the single population level (data not shown). Moreover, results of the H_{norm} and *DHEW* tests should not be affected by the accumulation of singletons as the first is based on intermediary and high frequency variants while the second integrates all variants of the site frequency spectrum. Thus, our analyses appear reasonably robust to population grouping.

Selective events appear to have affected distinct genes in *P. pinaster* and in *P. halepensis* despite their overlapping environment and close phylogenetic relationship, a result that may be connected to the different histories of the two pines that have likely resulted in different selective pressures. In *P. pinaster*, selective events were detected only within the Western Mediterranean group, while in *P. halepensis* footprints of selection were identified both in the Western Mediterranean and North African groups. These results point to different geographical selection pressures that may have led to the process of regional

adaptation of the pines (see Barbéro *et al.* 1998 for a description of the different Mediterranean environmental zones; and Gómez and Zamora 2000; Nakazato, Bogonovich, and Moyle 2008; Montesinos *et al.* 2009 for some examples on adaptive variation across heterogeneous environments). It is also noticeable that *P. pinaster* had lower levels of nucleotide variation within the North African range. This fact can be attributed to population history and/or interpreted in the light of more extreme environmental conditions constraining population sizes and/or resulting in stronger selection in this range.

Environmental associations

Environmental associations identified two loci that were correlated with temperature, suggesting the importance of this climatic variable as selective agent (Saxe *et al.* 2001; Jump *et al.* 2006). One of these genes was common for both pines (*4cl-Pt* correlated with MTCM in *P. pinaster* and with TS in *P. halepensis*), while the other gene association was only significant in *P. pinaster* (*dhn5-Ps* correlated with AMT and MTWM). Some of the correlations between genetic and climatic data in *P. pinaster* were due basically to the extreme values of a few North African populations (Sidi Meskour and Tamrabta for *4cl-Pt*_c; Tabarka for *dhn5-Ps*_C; see Figures 2 and 3) and may not represent true adaptive responses to environmental gradients but local adaptation to particular environments or genetic drift due to population isolation (see references in Alleaume-Benharira, Pen, and Ronce 2006; Rosenblum, Hickerson, and Moritz 2007). In contrast, the other associations (*4cl-Pt*_c in *P. halepensis* and *dhn5-Ps*_E in *P. pinaster*) showed more robust patterns. For instance, dhn5-Ps_E tended to be absent from populations characterized by the highest MTWM (Arenas de San Pedro in central Spain, Tamrabta in Moroccan Middle Atlas and Tabarka in coastal Tunisia) regardless of their geographical location (Table 1 and Figure 3).

Environmental correlations with allelic variation can be spuriously inflated by neutral processes that may also generate genetic clines, such as population history or population genetic structure. This is particularly true for forest trees that may have followed a postglacial colonization pathway overlapping temperature and rainfall clines. Here, to control for confounding associations between candidate genes and environmental data, we included variation of chloroplast microsatellites as a covariate in the multivariate logistic regressions. In *P. pinaster*, the use of cpSSRs has allowed a more accurate description of genetic structure (Bucci *et al.* 2007) than previous studies based on biochemical markers, such as terpenes or allozymes (e.g. Baradat and Marpeau-Bezard 1988; Salvador *et al.* 2000). Recently, nuclear molecular data for a geographically limited set of populations (nuSSRs and SNPs) (Eveno *et al.* 2008; our unpublished results) have identified similar gene pools as Bucci *et al.* (2007). More extensive sampling with highly-polymorphic biparentally-inherited nuclear markers could, in principle, reveal a more complex spatial genetic pattern affecting the outcome of some of the multivariate logistic regressions reported in this study. However, because the available chloroplast data set we used strongly reflects the species' history, they are expected to 'correct' for the presence of overall neutral genetic gradients. Accordingly, including these markers as covariates identified a substantial fraction of the correlations initially retained as significant as false positives (83.6 % in *P. pinaster* and 85.7% in *P. halepensis*) (Tables S6 and S7).

The dehydrin gene family

Neutrality tests and environmental associations both point to dehydrins as potential targets of natural selection in *P. pinaster*. Some dehydrins were also suggested to be under selection in a study based on detection of outlier loci in this species (Eveno *et al.* 2008). Dehydrins displayed also non-neutral patterns of nucleotide diversity in *Pinus sylvestris* populations showing divergence for cold tolerance in Europe (Wachowiak *et al.* 2009) and were associated with carbon isotope discrimination (and, thus, potentially with drought tolerance) in *P. taeda* (González-Martínez *et al.* 2008). Altogether these studies suggest the involvement of dehydrins in the adaptive response of pines to abiotic stress.

Dehydrins are part of a relatively small multigene family of intracellular stabilizers that plays a major role in cell protection against desiccation. These proteins are produced in response to any type of stress that causes dehydration at the cellular level, such as cold, drought or salinity (Close 1997). Changes in dehydrin gene expression have been reported in response to drought and/or cold stress in many plants such as cowpea (Ismail, Hall, and Close 1999), barley (Suprunova *et al.* 2004), wheat (Lopez *et al.* 2001), apple (Wisniewski *et al.* 2008), tomato (Weiss and Egea-Cortines 2009) and blueberry (Panta, Rieger, and Rowland 2001); while transgenic experiments confirmed the role of dehydrin genes in enhancing tolerance to drought or freezing stress in plants (e.g. in Arabidopsis, Puhakainen *et al.* 2004). In conifers, at least eight dehydrin genes have been identified in *Pinus sylvestris* (Joosen *et al.* 2006) and *Picea abies* (Yakovlev *et al.* 2008), and dehydrin expression has been shown to increase under wounding, cold and drought stress (Richard *et al.* 2000; Watkinson *et al.* 2003). Thus, it is not surprising to find in our study significant associations between dehydrins and temperatures, as critical low temperatures can cause tissue injury while high temperatures accompany dehydration, both stimulating the accumulation of dehydrins (Lewitt 1980; Ingram and Bartels 1996; Rizhsky, Liang, and Mittler 2002). Although our results, and others in *P. pinaster*, suggest that selection may have acted on some of the dehydrins tested, further genetic association and functional studies are necessary to confirm the role of dehydrins in local adaptation of this Mediterranean pine.

The *4-coumarate: CoA ligase* (*4cl*) gene

There was converging evidence both from neutrality tests (*DHEW* in *P. pinaster* and both *DHEW* and *ML-HKA* in *P. halepensis*) and environmental associations that the *4-coumarate: CoA ligase* (*4cl*) gene may be under selection in the two Mediterranean pines studied. Preliminary results on phenotypical association corrected by neutral gradients in *P. pinaster* also point at *4cl-Pt*. Indeed, multiple regressions reveal significant associations between polymorphism in *4cl-Pt_c* at the SNP and haplotype levels and total height in populations growing in sites characterized by dry and intermediate humidity (Table S8). The *4cl* family has been extensively studied in plants, where it is encoded by four to five genes in the fully sequenced genomes of Arabidopsis, rice and poplar (reviewed in Souza *et al.* 2008). The *4cl* gene is involved, among other processes, in the production of basic enzymes of the phenylpropanoid metabolism that are important metabolites acting as protectants against biotic and abiotic stresses (Rani *et al.* 2009). The *4cl* gene also encodes key enzymes in the biosynthesis of lignin and several studies have demonstrated its involvement in plant growth (e.g. Yun *et al.* 2005; Wagner *et al.* 2009; Yun *et al.* 2009). Thus, we expect that changes in *4cl* function would have significant repercussions on tree physiology and morphology. Implication of 4cl in pine morphology and physiology has been shown in gene association studies (for *P. taeda* see González-Martínez *et al.* 2007b) as well as in gene suppression studies (for *P. radiata* see Wagner *et al.* 2009).

Neutrality tests versus Environmental correlations

The different approaches used in this study suggested different loci under selection within each of the two Mediterranean pine species. This fact has to be connected to the specificities of the statistical methods used, whose performance has been recently studied under various demographic and recombination scenarios (Zeng *et al.* 2007; Zeng, Shi, and Wu 2007; Ramírez-Soriano *et al.* 2008; Zhai, Nielsen, and Slatkin 2009). Within-species site and haplotype frequency spectrum-based methods are suitable to detect ongoing or recently fixed selective sweeps - *EW* tending to be more powerful around the time when a selected mutation reaches fixation (Zeng, Shi, and Wu 2007). In contrast, tests comparing levels of (within-species) polymorphism and (among-species) divergence are able to detect the cumulative effects of positive selection events over a wider evolutionary scale. Within this category, the *ML-HKA* test is suitable to quantify the amount of selection in the genome and, due to its multilocus nature, it is expected to be more robust to changes in population size such as population bottlenecks and expansion than traditional tests (Wright and Charlesworth 2004). Simulations have shown that the *HKA* test is relatively insensitive to change of divergence time as most of its power comes from the transient reduction in variability occurring during selective sweeps (Zhai, Nielsen, and Slatkin 2009). However, we only obtained significant tests when using the less closely-related outgroup (i.e. a New World pine, *P. taeda*). This points to either lower power when divergence time from the outgroup is low or, alternatively, to selection events that took place before the split of the two Mediterranean pines considered. In our study, none of the *EW* neutrality tests were significant (data not shown), while both the *DHEW* and the *ML-HKA* tests revealed distinct genes potentially under selection. As a consequence, potential selective events detected by the neutrality tests assessed in this study would correspond to either on-going or relatively old selective events (i.e. before the diversification of the Mediterranean pines in the Miocene, ~10 Ma; Gernandt *et al.* 2008).

Last generation tests, such as *DHEW* or *ML-HKA*, are powerful for detecting positive selection and are relatively insensitive to other evolutionary forces, but still do not integrate recombination rate, a factor that can produce substantial biases (Nielsen *et al.* 2007; Ramírez-Soriano *et al.* 2008), nor provide insights on selection drivers. Methods based on correlation between genetic and environmental data are appealing because they aim at understanding species-specific adaptations and processes that connect an organism to its environment by looking at allelic frequencies or the genetic structure of populations (Foll and Gaggiotti 2006; Joost, Kalbermatten, and Bonin 2008). These correlation methods are also more appealing than methods based on detection of outliers (e.g. Eveno *et al.* 2008) as they target particular selection drivers and provide a hypothesis testing framework. In addition, they are well suited for high-throughput genotyping even for non-model species (Namroud *et al.* 2008; Eckert *et al.* forthcoming). However, correlation methods are not without limitations, as it may be challenging to find the environmental factors that are relevant for each species adaptation and associations have to be controlled for historical and demographical processes, in particular for allelic clines produced by postglacial migrations in Europe and the Americas.

Together, neutrality tests and environmental association approaches complement each other by looking at different evolutionary scales and types of selection. In our study, they detected a relatively-high number of genes showing non-neutral patterns of evolution, a result that can be attributed to a selection of candidate genes based on earlier studies (González-Martínez *et al.* 2006; Eveno *et al.* 2008; Grivet *et al.* 2009). This fact supports a candidate gene approach with targeted genes, at least for organisms that have large genomes (e.g. conifers), which, so far, preclude dense genome-wide sampling.

Conclusion

Our pluralistic approach revealed a dynamic action of natural selection in space and time within maritime and Aleppo pines (Felsenstein 1976; see Vasemagi and Primmer 2005 for other examples). Selection events along with environmental associations were detected; some of these events differ between the two species reflecting individual histories (recolonization, demography, adaptation) while others are shared, which translates partly as a common history of these closely-related and partially sympatric Mediterranean pines. The simultaneous search for patterns of selection in two closely-related species can help understanding the evolutionary forces responsible for adaptive responses, and thus provides an effective way of assessing the degree of local adaptation, a key factor to integrate in future management and conservation strategies (Wright and Gaut 2005).

Acknowledgements

We are grateful to C. García-Barriga for her assistance in producing sequences of *P. pinaster* candidate genes, to C. Collada, M.T. Cervera, M.A. Guevara, G. Gill, G.R. Brown and D.B. Neale for providing some of the primers, as well as to A.I. de-Lucas and C. Ordoñez for supplying some of the *P. pinaster* seeds. We thank J. Gonzalo for kindly providing the climatic data from his functional phytoclimatic model, as well as S.E. Ramos-Onsins for his scripts to compute the diversity parameters and K. Zeng for his scripts to compute the *DHEW* neutrality test. Thanks are extended to M. Robson, M.A. Zavala and J. Climent for helpful discussion on the study. We are also thankful to A.J. Eckert for thoughtful comments on the manuscript, N. Takebayashi and two anonymous reviewers for constructive comments and suggestions, and P.C. Grant for revising the English. The EUFORGEN program (Bioversity International) provided the distribution maps used in Fig. 1.

Funding

This work was supported by the EU EVOLTREE Network of Excellence, the Collaborative Project on 'Conservation of Forest Genetic Resources' between the Spanish Ministry of Environment and INIA (AEG06-054) and Projects CGL2008-05289-C02-02 /BOS (VaMPiro) and EUI2008-03713 (BiodivERsA-LinkTree) from the Spanish Ministry of Science and Innovation, the later through the FP6 BiodivERsA Eranet. TB acknowledges support from the Danish Research Council (FNU) through a Steno fellowship.

LITERATURE CITED

Alleaume-Benharira M, Pen IR, Ronce O. 2006. Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. J Evol Biol. 19:203-215.

Aranda I, Alía R, Ortega U, Dantas AK, Majada J. 2009. Intra-specific variability in biomass partitioning and carbon isotopic discrimination under moderate drought stress in seedlings from *Pinus pinaster* populations. Tree Genet Genomes 6:169-178.

Atzmon N, Moshe Y, Schiller G. 2004. Ecophysiological response to severe drought in *Pinus halepensis* Mill. trees of two provenances. Plant Ecol. 171:15-22.

Baradat PH, Marpeau-Bezard A. 1988. Le pin maritime, *Pinus pinaster* Ait. Biologie et génétique des terpènes pour la connaissance et l'amélioration de l'espèce. PhD Thesis, University of Bordeaux I, France.

Barbéro M, Loisel R, Quézel P, Richardson DM, Romane F. 1998. Pines of the Mediterranean Basin. In Richardson DM, editor. Ecology and Biogeography of *Pinus*. Cambridge University Press, Cambridge. Pp. 153-170.

Biswas S, Akey JM. 2006. Genomic insights into positive selection. Trends Genet. 22:437-446.

Bucci G, Anzidei M, Madaghiele A, Vendramin GG. 1998. Detection of haplotypic variation and natural hybridization in halepensis-complex pine species using chloroplast simple sequence repeat (SSR) markers. Mol Ecol. 7:1633-1643.

Bucci G, González-Martínez SC, Le Provost G, Plomion C, Ribeiro MM, Sebastiani F, Alía R, Vendramin GG. 2007. Range-wide phylogeography and gene zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. Mol Ecol. 16:2137-2153.

Burban C, Petit RJ. 2003. Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. Mol Ecol. 12:1487-1495.

Chambel MR, Climent J, Alía R. 2007. Divergence among species and populations of Mediterranean pines in biomass allocation of seedlings grown under two watering regimes. Ann For Sci. 64:87-97.

Climent J, Prada MA, Calama R, Chambel MR, De Ron DS, Alía R. 2008. To grow or to seed: Ecotypic variation in reproductive allocation and cone production by young female Aleppo pine (*Pinus halepensis*, Pinaceae). Am J Bot. 95:833-842.

Close T. 1997. Dehydrins: a commonality in the response of plants to dehydration and low temperature. Physiol Plant. 100:291-296.

Coop G, Pickrell JK, Novembre J, Kudaravalli S, Li J, Absher D, Myers RM, Cavalli-Sforza LL, Feldman MW, Pritchard JK. 2009. The role of geography in human adaptation. PLoS Genetics 5(6): e1000500.

Costa P, Bahrman N, Frigerio JM, Kremer A, Plomion C. 1998. Water-deficit-responsive proteins in maritime pine. Plant Mol Biol. 38:587-596.

Dellaporta SL, Wood J, Hicks JB. 1983. A plant DNA minipreparation: version II. Plant Mol Biol Rep. 1:19-21.

Eckert AJ, van Heerwaarden J, Wegrzyn JL, Nelson CD, Ross-Ibarra J, González-Martínez SC, Neale DB. Forthcoming. Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). Genetics.

Eveno E, Collada C, Guevara MA, *et al.* (11 co-authors). 2008. Contrasting patterns of selection at *Pinus pinaster* Ait. drought stress candidate genes as revealed by genetic differentiation analyses. Mol Biol Evol. 25:417-437.

Felsenstein J. 1976. Theoretical population-genetics of variable selection and migration. Annu Rev Genet.10:253-280.

Foll M, Gaggiotti O. 2006. Identifying the environmental factors that determine the genetic structure of populations. Genetics 174:875-891.

Frankel N, Carrari F, Hasson E, Iusem ND. 2006. Evolutionary history of the Asr gene family. Gene 378:74-83.

Gernandt DS, Lopez GG, Garcia SO, Liston A. 2005. Phylogeny and classification of Pinus. Taxon 54:29-42.

Gernandt DS, Magallon S, Lopez GG, Flores OZ, Willyard A, Liston A. 2008. Use of simultaneous analyses to guide fossil-based calibrations of *Pinaceae* phylogeny. Int J Plant Sci.169:1086-1099.

Giorgi F, Lionello P. 2008. Climate change projections for the Mediterranean region. Glob Planet Change 63:90-104.

Gómez JM, Zamora R. 2000. Spatial variation in the selective scenarios of *Hormathophylla spinosa* (Cruciferae). Am Nat. 155:657-668.

González-Martínez SC, Dillon S, Garnier-Géré P, *et al.* (16 co-authors). Forthcoming. Patterns of nucleotide diversity and association mapping In Kole C, editor. Genomics of Conifers. Science Publishers, Inc., New Hampshire.

González-Martínez SC, Ersoz E, Brown GR, Wheeler NC, Neale DB. 2006. DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L. Genetics 172:1915-1926.

González-Martínez SC, Gómez A, Carrión JS, Agúndez D, Alía R, Gil L. 2007a. Spatial genetic structure of an explicit glacial refugium of maritime pine (*Pinus pinaster* Aiton) in southeastern Spain. In Weiss S, Ferrand N, editors. Phylogeography in southern European refugia: evolutionary perspectives on the origins of European biodiversity, London. Pp. 257-269.

González-Martínez SC, Huber D, Ersoz E, Davis JM, Neale DB. 2008. Association genetics in *Pinus taeda* L. II. Carbon isotope discrimination. Heredity 101:19-26.

González-Martínez SC, Wheeler NC, Ersoz E, Nelson CD, Neale DB. 2007b. Association genetics in *Pinus taeda* L. I. Wood property traits. Genetics 175:399-409.

Gonzalo J. 2007. Phytoclimatic analysis of the Spanish Peninsula. Update and geostatistical analysis. PhD Thesis, University of Valladolid, Spain.

Gram WK, Sork VL. 2001. Association between environmental and genetic heterogeneity in forest tree populations. Ecology 82:2012-2021.

Grivet D, Sebastiani F, González-Martínez SC, Vendramin GG. 2009. Patterns of polymorphism resulting from long-range colonization in the Mediterranean conifer Aleppo pine. New Phytol. 184:1016-1028.

Hancock AM, Witonsky DB, Gordon AS, Eshel G, Pritchard JK, Coop G, Di Rienzo A. 2008. Adaptations to climate in candidate genes for common metabolic disorders. PLoS Genetics 4(2): e32.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 25:1965-1978.

Hudson RR, Kreitman M, Aguade M. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics 116:153-159.

Ingram J, Bartels D. 1996. The molecular basis of dehydration tolerance in plants. Annu Rev Plant Physiol Plant Mol Biol. 47:377-403.

Ismail AM, Hall AE, Close TJ. 1999. Allelic variation of a dehydrin gene cosegregates with chilling tolerance during seedling emergence. Proc Natl Acad Sci USA. 96:13566-13570.

Joosen RVL, Lammers M, Balk PA, Bronnum P, Konings MCJM, Perks M, Stattin E, Van Wordragen MF, van der Geest AHM. 2006. Correlating gene expression to physiological parameters and environmental conditions during cold acclimation of *Pinus sylvestris*, identification of molecular markers using cDNA microarrays. Tree Physiol. 26:1297-1313.

Joost S, Kalbermatten M, Bonin A. 2008. Spatial analysis method (SAM): a software tool combining molecular and environmental data to identify candidate loci for selection. Mol Ecol Resources 8:957-960.

Jump AS, Hunt JM, Martínez-Izquierdo JA, Peñuelas J. 2006. Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. Mol Ecol. 15:3469-3480.

Keller SR, Sowell DR, Neiman M, Wolfe LM, Taylor DR. 2009. Adaptation and colonization history affect the evolution of clines in two introduced species. New Phytol. 183:678-690.

Kimura M, Ohta T. 1974. Some principles governing molecular evolution. Proc Natl Acad Sci USA. 71:2848-2852.

Lewitt J. 1980. Responses of plants to environmental stresses, Vol I. Chilling, freezing and high temperature stresses. Academic Press, New York, USA.

Librado P, Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451-1452.

Lopez CG, Banowetz G, Peterson CJ, Kronstad WE. 2001. Differential accumulation of a 24-kd dehydrin protein in wheat seedlings correlates with drought stress tolerance at grain filling. Hereditas 135:175-181.

Manel S, Schwartz MK, Luikart G, Taberlet P. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends Ecol Evol. 18:189-197.

McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the ADH locus in *Drosophila* Nature 351:652-654.

Montesinos A, Tonsor SJ, Alonso-Blanco C, Picó FX. 2009. Demographic and genetic patterns of variation among populations of *Arabidopsis thaliana* from contrasting native environments. PLoS One 4:e7213.

Morgante M, Felice N, Vendramin GG. 1998. Analysis of hypervariable chloroplast microsatellites in *Pinus halepensis* reveals a dramatic genetic bottleneck. In Karp A, Isaac PG, Ingram DS, editors. Molecular tools for screening biodiversity. Chapman & Hall, London, UK. Pp. 407-412.

Nakazato T, Bogonovich M, Moyle LC. 2008. Environmental factors predict adaptive phenotypic differentiation within and between two wild Andean tomatoes. Evolution 62:774-792.

Namroud MC, Beaulieu J, Juge N, Laroche J, Bousquet J. 2008. Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. Mol Ecol. 17:3599-3613.

Nei M. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.

Nguyen A, Lamant A. 1989. Variation in growth and osmotic regulation of roots of water-stressed maritime pine (*Pinus pinaster* Ait) provenances.Tree Physiol. 5:123-133.

Nielsen R. 2005. Molecular signatures of natural selection. Annu Rev Genet. 39:197-218.

Nielsen R, Hellmann I, Hubisz M, Bustamante C, Clark AG. 2007. Recent and ongoing selection in the human genome. Nat Rev Genet. 8:857-868.

Panta GR, Rieger MW, Rowland LJ. 2001. Effect of cold and drought stress on blueberry dehydrin accumulation. J Horticult Sci Biotechnol. 76:549-556.

Parisod C, Christin PA. 2008. Genome-wide association to fine-scale ecological heterogeneity within a continuous population of *Biscutella laevigata* (Brassicaceae). New Phytol. 178:436-447.

Petit RJ, Hampe A, Cheddadi R. 2005. Climate changes and tree phylogeography in the Mediterranean. Taxon 54:877-885.

Puhakainen T, Hess MW, Mäkelä P, Svensson J, Heino P, Palva ET. 2004. Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in *Arabidopsis*. Plant Mol Biol. 54:743-753.

Quézel P. 2000. Taxonomy and biogeography of Mediterranean pines. In Ne'eman G, Trabaud L, editors. Ecology, biogeography and management of *Pinus halepensis* and *P. brutia* forest ecosystems in the Mediterranean Bassin. Backhuys Publishers, Leiden, The Netherlands. Pp. 1-12.

Ramírez-Soriano A, Ramos-Onsins SE, Rozas J, Calafell F, Navarro A. 2008. Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. Genetics 179:555-567.

Rani A, Singh K, Sood P, Kumar S, Ahuja PS. 2009. p-Coumarate:CoA ligase as a key gene in the yield of catechins in tea [*Camellia sinensis* (L.) O. Kuntze]. Funct Integr Genomics 9: 271-275.

Resch T. 1974. Essai de distinction des races majeurs de *Pinus pinaster*. Annales de la Recherche Forestière au Maroc 14:91-102.

Richard S, Morency MJ, Drevet C, Jouanin L, Séguin A. 2000. Isolation and characterization of a dehydrin gene from white spruce induced upon wounding, drought and cold stresses. Plant Mol Biol. 43:1-10.

Richardson DM, Bond WJ. 1991. Determinants of plant-distribution - Evidence from pine invasions. Am Nat. 137:639-668.

Rizhsky L, Liang HJ, Mittler R. 2002. The combined effect of drought stress and heat shock on gene expression in tobacco. Plant Physiol. 130:1143-1151.

Rosenblum EB, Hickerson MJ, Moritz C. 2007. A multilocus perspective on colonization accompanied by selection and gene flow. Evolution 61:2971-2985.

Salvador L, Alía R, Agúndez D, Gil L. 2000. Genetic variation and migration pathways of maritime pine (*Pinus pinaster* Ait) in the Iberian peninsula. Theor Appl Genet. 100:89-95.

Sathyan P, Newton RJ, Loopstra CA. 2005. Genes induced by WDS are differentially expressed in two populations of Aleppo pine (*Pinus halepensis*). Tree Genet Genomes 1:166-173.

Savolainen O, Pyhäjärvi T, Knürr T. 2007. Gene flow and local adaptation in trees. Annu Rev Ecol Evol Syst. 38:595-619.

Saxe H, Cannell MGR, Johnsen B, Ryan MG, Vourlitis G. 2001. Tree and forest functioning in response to global warming. New Phytol. 149:369-399.

Soto A, Robledo-Arnuncio JJ, González-Martínez SC, Smouse PE, Alía R. 2010. Climatic niche and neutral genetic diversity of the six Iberian pine species: a retrospective and prospective view. Mol Ecol. 19:1396-1409.

Souza CD, Barbazuk B, Ralph SG, Bohlmann J, Hamberger B, Douglas CJ. 2008. Genome-wide analysis of a land plant-specific acyl: coenzymeA synthetase (ACS) gene family in *Arabidopsis*, poplar, rice and *Physcomitrella*. New Phytol. 179:987-1003.

Städler T, Haubold B, Merino C, Stephan W, Pfaffelhuber P. 2009. The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. Genetics 182:205-216.

Storz JF. 2002. Contrasting patterns of divergence in quantitative traits and neutral DNA markers: analysis of clinal variation. Mol Ecol. 11:2537-2551.

Suprunova T, Krugman T, Fahima T, Chen G, Shams I, Korol A, Nevo E. 2004. Differential expression of dehydrin genes in wild barley, *Hordeum spontaneum*, associated with resistance to water deficit. Plant Cell Environ. 27:1297-1308.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

Tapias R, Climent J, Pardos JA, Gil L. 2004. Life histories of Mediterranean pines. Plant Ecol. 171:53-68.

Vasemagi A, Primmer CR. 2005. Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. Mol Ecol. 14:3623-3642.

Voltas J, Chambel M, Prada M, Ferrio J. 2008. Climate-related variability in carbon and oxygen stable isotopes among populations of Aleppo pine grown in common-garden tests. Trees-Struct Funct. 22:759-769.

Wachowiak W, Balk PA, Savolainen O. 2009. Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvestris* L.). Tree Genet Genomes 5:117-132.

Wagner A, Donaldson L, Kim H, Phillips L, Flint H, Steward D, Torr K, Koch G, Schmitt U, Ralph J. 2009. Suppression of 4-Coumarate-CoA ligase in the coniferous gymnosperm *Pinus radiata*. Plant Physiol. 149:370-383.

Watkinson JI, Sioson AA, Vasquez-Robinet C, *et al.* (14 co-authors). 2003. Photosynthetic acclimation is reflected in specific patterns of gene expression in drought-stressed loblolly pine. Plant Physiol. 133: 1702-1716.

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. Theor Pop Biol. 7: 256-276.

Watterson GA. 1978. The homozygosity test of neutrality. Genetics 88:405-417.

Weiss J, Egea-Cortines M. 2009. Transcriptomic analysis of cold response in tomato fruits identifies dehydrin as a marker of cold stress. J Appl Genetics 50:311-319.

Wisniewski M, Bassett C, Norelli J, Macarisin D, Artlip T, Gasic K, Korban S. 2008. Expressed sequence tag analysis of the response of apple (*Malus* x *domestica* 'Royal Gala') to low temperature and water deficit. Physiol Plant. 133:298-317.

Wright SI, Charlesworth B. 2004. The HKA test revisited: a Maximum-Likelihood-Ratio test of the standard neutral model. Genetics 168:1071-1076.

Wright SI, Gaut BS. 2005. Molecular population genetics and the search for adaptive evolution in plants. Mol Biol Evol. 22:506-519.

Yakovlev IA, Asante DKA, Fossdal CG, Partanen J, Junttila O, Johnsen O. 2008. Dehydrins expression related to timing of bud burst in Norway spruce. Planta 228:459-472.

Yun MS, Chen WJ, Deng F, Yogo Y. 2005. Differential properties of 4-coumarate: CoA ligase related to growth suppression by chalcone in maize and rice. Plant Growth Regul. 46:169-176.

Yun MS, Chen WJ, Deng F, Yogo Y. 2009. Selective growth suppression of five annual plant species by chalcone and naringenin correlates with the total amount of 4-coumarate: coenzyme A ligase. Weed Biol Manag. 9:27-37.

Zeng K, Fu Y-X, Shi S, Wu C-I. 2006. Statistical tests for detecting positive selection by utilizing high-frequency variants. Genetics 174:1431-1439.

Zeng K, Mano SH, Shi SH, Wu CI. 2007. Comparisons of site- and haplotype-frequency methods for detecting positive selection. Mol Biol Evol. 24:1562-1574.

Zeng K, Shi S, Wu CI. 2007. Compound tests for the detection of hitchhiking under positive selection. Mol Biol Evol. 24:1898-1908.

Zhai WW, Nielsen R, Slatkin M. 2009. An investigation of the statistical power of neutrality tests based on comparative and population genetic data. Mol Biol Evol. 26:273-283.

COMMUNITY GENETICS IN THE TIME OF NEXT GENERATION MOLECULAR TECHNOLOGIES

FELIX GUGERLI[1], ROLAND BRANDL[2], BASTIEN CASTAGNEYROL[3], ALAIN FRANC[3], HERVÉ JACTEL[3], HANS-PETER KOELEWIJN[4], FRANCIS MARTIN[5], MARTINA PETER[1], KARIN PRITSCH[6], HILKE SCHRÖDER[7], MARINUS J.M. SMULDERS[8], ANTOINE KREMER[3], BIRGIT ZIEGENHAGEN[2] and EVOLTREE JERA[3] CONTRIBUTORS*

[1] WSL Swiss Federal Research Institute, 8903 Birmensdorf, Switzerland
[2] Fachbereich Biologie, Philipps-Universität Marburg, 35032 Marburg, Germany
[3] UMR1202 Biodiversity, Genes & Communities, INRA Pierroton, 33612 Cestas Cedex, France
[4] ALTERRA Centre for Ecosystem Studies, Wageningen UR, 6700 AA Wageningen, The Netherlands
[5] UMR "Interactions Arbres/Micro-Organismes", INRA Nancy, 54280 Champenoux, France
[6] Institute of Soil Ecology, Helmholtz Zentrum München, 85764 Neuherberg, Germany
[7] Institute for Forest Genetics, Johann Heinrich von Thuenen-Institute, 22927 Grosshansdorf, Germany
[8] Plant Research International, Wageningen UR, 6700 AA Wageningen, The Netherlands
* The EC-supported Network of Excellence Evoltree (http://www.evoltree.eu) formed a group of scientists involved in and actively contributing to Jointly Executed Research Activities on community genetics in forest ecosystems. In addition to the main authors, the group includes S. Augustin, M. Brändle, C. Burban, J. Burczyk, S. Cavers, I. Chybicki, C. Conord, E. Cremer, J. DeWoody, K. Donges, B. Fady, L. Karlinski, C. Kerdelhué, B. Kieliszewska-Rokicka, G. Kost, M. Kulczyk-Skrzeszewska, F. Lakatos, F. Lefèvre, S. Liepelt, S. Oddou-Muratorio, K.-H. Rexer, M. Rudawska, M. Schädler, G. Taylor, K. Tuba, M. Viger, F. Villani, M. Villar

Corresponding Author:
Felix Gugerli
WSL Swiss Federal Research Institute, 8903 Birmensdorf, Switzerland
E-mail: gugerli@wsl.ch)

Understanding the interactions of co-occurring species within and across trophic levels provides key information needed for understanding the ecological and evolutionary processes that underlie biological diversity. As genetics has only recently been integrated into the study of community-level interactions, the time is right for a critical evaluation of potential new, gene-based approaches to studying communities. Next generation molecular techniques, used in parallel with field-based observations and manipulative experiments across spatio-temporal gradients, are key to expanding our understanding of community-level processes. Here, we introduce a variety of "-omics" tools, with recent studies of plant–insect herbivores and of ectomycorrhizal systems providing detailed examples of how next generation approaches can revolutionize our understanding of interspecific interactions. We suggest ways that novel technologies may convert community genetics from a field that relies on correlative inference to one that reveals causal mechanisms of genetic co-variation and adaptations within communities.

Key words: ectomycorrhizal symbiosis, forest ecosystems, gene-to-gene interactions, herbivorous insects, population genomics, quantitative trait analysis.

Published in *Molecular Ecology* (2013) 22: 3198-3207, with doi:10.1111/mec.12300

Community genetics aims to understand how genetic variation within and among populations of host species affects the composition of associated organisms interacting with the host (Agrawal 2003; Whitham *et al.* 2006; Johnson & Stinchcombe 2007; Rowntree *et al.* 2011; Wymore *et al.* 2011). Empirical community genetics has been stimulated by pioneering work on poplars (*Populus* spp.), their genotype-based phenotypic variation, and associated communities (Whitham *et al.* 2006). However, community genetics has hitherto largely remained phenomenological, and the underlying genetic basis and processes involved in the interactions between host and associated organisms have not been studied in detail yet. Given the rapid development of molecular techniques (Rokas & Abbot 2009), it will soon be feasible to characterize the genomes of numerous members of a community. With whole-genome sequences or other types of -omics data at hand (Nadeau & Jiggins 2010), community genetics will be able to establish a solid genetic framework in which to understand the interplay between ecological and evolutionary processes (Rokas & Abbot 2009). Here, we sketch possible avenues along which research in community genetics may proceed, focussing in particular on how -omics may improve our understanding of the role of gene variants in species interactions. First, we argue for exploring spatio-temporal variation to investigate the fundamental ecological and evolutionary aspects of community genetics. Second, we describe how genomic, transcriptomic, proteomic, and metabolomic research can improve understanding of the interactions between trees as focal species and ectomycorrhizal fungi or herbivorous insects, the key players in forest ecosystems.

Community genetics in a spatio-temporal perspective

Let us consider populations of a focal species that start to diverge genetically. Genetic drift and/or selection may induce shifts in allele frequencies, leading to changes in the phenotypic traits mediating interactions with associated species that use the focal species as a host.

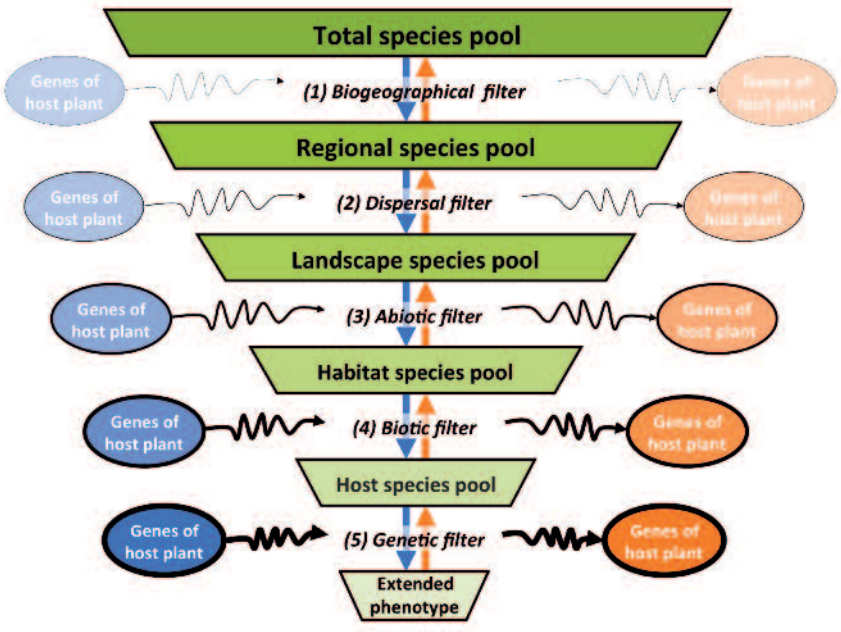
First, these genetic changes and changes in the associated traits may lead to shifts in the occurrence and abundance of species already associated with the host. Second, the new phenotypic traits of the focal species may allow new species from the regional species pool to colonize it. Finally, changes in the genetics of the host may induce evolutionary responses, including speciation events, in the associated organisms, which may feedback to evolutionary changes in the host.

If the above scenarios hold true, we expect the relatedness of host genotypes to co-vary with similarity among the communities of associated species (Bangert *et al.* 2006; Brändle & Brandl 2006). Within species, such patterns have received considerable attention under the concept of the "extended phenotype". This concept was introduced by Richard Dawkins (1982) to describe effects of genes on an individual's environment including other organisms. Whitham *et al.* (2003; 2005; 2006) adopted this concept and developed a framework for community and ecosystem genetics, which includes a feedback where an individual's phenotype is dependent on the interaction with other species.

Community assembly (Kraft *et al.* 2007; Emerson & Gillespie 2008) is shaped by successive filters, including regional species pool, habitat area and isolation (biogeographical filters), local environmental constraints

FIGURE 1

How host plant genes might shape assemblages of associated organisms (blue pathway on the left). Several ecological filters drive the structure of communities associated with one host plant. Among associated species co-occurring within a region and determined by evolutionary and biogeographical processes (1, Total species pool), local species assemblages depend on dispersal (2, Landscape species pool) and habitat filters (3, Habitat species pool). Dispersal filter refers to the ability of species to colonize the focal site. Habitat filters correspond to their capacity to develop and survive in a habitat given abiotic constraints. Biotic interactions with the host species contribute to the shaping of a host species pool (4, biotic filter). Finally, variation among host plant genotypes may further select different associated communities, shaping the extended phenotypes.
Genes of the focal host plant can interact with the four filters, as illustrated by the interaction between trees and associated insect herbivores: (1) There is evidence that pools of insect herbivore species of different tree families or genera are significantly different, probably owing to a long co-evolutionary process involving insect feeding traits and plant defence responses (Novotny *et al.* 2002); (2) insect herbivores use genetically controlled physical (e.g. shape, colour) and chemical cues (e.g. volatile organic compounds) provided by host plants to locate the plants; (3) trees can be seen as ecological engineers which can modify abiotic conditions that insects experience, e.g. wind, moisture, or light; (4) genes control plant phenotype and resistance traits that are deeply involved in interactions with insect herbivores (Schoonhoven *et al.* 2005); and (5) variants of host plant genes may ultimately induce quantitative changes in traits involved in plant–insect interactions with consequences for insect community structure (Crutsinger *et al.* 2008).
Presumed reciprocal effects, through which associated organisms feed back to the composition of host genes, are depicted by orange colors (right side).



(abiotic, biophysical filters) and biological interactions such as competition or predation (biotic filters; Fig. 1). The host genotype, interacting with the environment, may affect the structure of associated communities at several filtering steps by controlling phenotypic traits that allow associated organisms to locate, select and exploit resources of their host (Johnson & Agrawal 2005; Bailey *et al.* 2009) (Fig. 1). Thus, spatial variation in the composition of associated communities has a strong regional component.

Despite many reports demonstrating a correlation between genotypes of a focal species and the composition of associated communities, the fundamental ecological, genetic and evolutionary processes that generate this correlation remain poorly explored and require consideration in future studies. In this regard, three aspects deserve special attention: spatial variation, temporal variation, and gene-to-gene interactions.

First, space needs to be better integrated into study designs. As noted above, the assembly of species depends on the regional species pool, whose phylogenetic and functional structure imposes a constraint on the emerging local communities (Fig. 1). A group of genotypes of a focal species in natural or experimental populations is embedded in a landscape context that may include forest patches, arable land, urban environment or other habitats type, each of which has different species pools that might interact with the focal species. As the associated community influences the fitness of the focal species, the relative fitness of these genotypes will vary across sites, even if the abiotic conditions are similar. However, in single common garden experiments, genotypes of a focal species are exposed only to one particular species pool. Therefore, regional replicates of such experiments are necessary to estimate the stability of relationships between genotypes of the focal species and communities of associated species. Such replicates would enable us to distinguish between mainly spatial effects and those that can be attributed to the interaction between host genotypes and associated organisms. Alternatively, one might set up more complex common gardens including particular treatments, for example through fertilization or irrigation. Such an approach would allow tests of the effect of genotype x environment interactions on the assemblage of associated species for each local species pool. Furthermore, replicated common garden experiments would allow constructing reaction norms of different genotypes of the focal species. Do these genotypes respond differentially for their extended phenotypes to the changes of abiotic or biotic conditions across the testing sites? An initial step would be to identify the shape of the reaction norms (linear or quadratic) and then to estimate their variation among genotypes. Finally, the spatial context may also be dissected at the within-population level. For example, natural populations of trees usually exhibit strong spatial autocorrelation due to limited dispersal, which increases steadily over generations. On the other hand random spatial genetic structure is observed in recently planted forests. One would therefore expect very different spatial structures of extended phenotypes among these strongly contrasting cultural regimes.

Second, community genetics should consider temporal variation in species interactions, e.g. among seasons, among years along successional sequences, and other types of temporal gradients. Traits involved in plant–herbivore interactions are known to change during plant ontogeny (Boege & Marquis 2005; Holeski *et al.* 2009), which is why communities of insect herbivores – and herbivory pressure – on seedlings and mature individuals may differ (Le Corff & Marquis 1999; Basset 2001). Furthermore, although associated communities may change within and between years due to fluctuations in plant phenotypes, equally they may change due to differences in weather conditions. Thus, the phenotypic traits that are important for species interactions in a particular season or year may change within and between years, and drawing conclusions from short-term experiments may be misleading. Although such traits, and the underlying genes, are genuinely involved in community interactions, their relative importance compared to other genes may vary in time and can therefore only be established in long-term experiments. Hundreds of insect generations interact with a long-living host such as a tree during its lifetime, and each generation experiences different biophysical constraints and trophic interactions with other fungi, herbivores or predators. As a consequence, even though insect populations can adapt to individual host genotypes (Mopper *et al.* 2000), the strength and direction of these adaptations are likely to change over time (moving targets; Ruhnke *et al.* 2006).

Moreover, genetic processes underlie the formation of adaptive demes and co-evolution between host and associated organisms (Fig. 1). At present, the number and type of genes involved and the associated phenotypes of interacting species are largely unknown. Recent technological advances enable researchers to sequence whole genomes and to monitor gene expression of interacting species, offering the potential to identify the candidate genes mediating the interactions between focal and associated species. Such approaches will move community genetics from studying anonymous genotype/phenotype effects to studying gene-to-organism, gene-to-gene, and ultimately to genome-to-genome interactions. While current research has focused on the few "genome-enabled" species (Ekblom & Galindo 2011), the many ongoing whole-genome projects will widen the array of study systems applying genomics data in the near future (e.g. http://www.arthropodgenomes.org/wiki/i5K, http://1000.fungalgenomes.org/home/, http://pinegenome.org/pinerefseq/).

The following sections describe how the various types of -omics may stimulate community genetics. and how they enable the genetic component of variation in community composition to be addressed at the level of variants in adaptive genes and their differential expression.

An example of functional genomics based on complete genome sequences: ectomycorrhizal symbiosis

Ectomycorrhizae, the mutualistic symbiosis between tree roots and a cortege of soil fungal partners, are the most widespread and species-rich associations in temperate and boreal forests. Ectomycorrhizal fungi receive carbon from photosynthesis and, in turn, promote tree growth, enhance the survival of seedlings and increase the fitness of their plant partners under a wide range of environmental conditions. Despite the ecological significance of this mutualistic interaction, we have only started to explore its role for community ecology.

A breakthrough was the release of the first two full-genome drafts of mycorrhizal fungi, namely *Laccaria bicolor* (Basidiomycota) and *Tuber melanosporum*, the Périgord truffle (Ascomycota; Martin *et al.* 2008; Martin *et al.* 2010). Comparative genomics of the two mycorrhizal fungi indicated that they use different gene networks ('molecular toolkits') to establish symbiosis (Martin et *al. 2010*). There are vast differences between these two ectomycorrhizal genomes. *Laccaria bicolor* has a 65 Mb genome with more than 23 000 predicted proteins, which is the largest complement of genes known for any fungus, whereas *T. melanosporum* has the largest fungal genome so far with 125 Mb, but has only 7500 predicted genes, one of the smallest complement of proteins in any filamentous fungal genomes sequenced so far. Also, whereas the secretion of effector-like small secreted proteins seems to be crucial for the establishment of the symbiosis in *L. bicolor* (Plett *et al.* 2011), these so-called mycorrhiza-induced small secreted proteins (MiSSPs) are not present in the transcriptome of *T. melanosporum* symbiotic tissue (Martin *et al.* 2010). In spite of these differences,, some common features and some novelties emerged from the comparison with genomes of saprophytic and pathogenic fungi. Besides the loss of plant cell-wall degrading enzymes in ectomycorrhizae, an increase in the diversity and expression of nutrient transporters and signalling pathways (e.g. tyrosine kinases) in symbiotic tissues are hallmarks of mycorrhizal genomes (Martin *et al.* 2008; Kosti *et al.* 2010; Martin *et al.* 2010; Plett *et al.* 2011). These symbiosis-related genes are good candidates for gene expression studies of multi-species interactions in the field. On the tree side, it is not known how the host tree selects its symbiotic associates. Plant-encoded small secreted proteins may be required, as shown for nitrogen-fixing symbioses (Van de Velde *et al.* 2010). Genomic studies will probably be the only way to elucidate the mechanisms of interaction and to understand the effect of gene variants on this interplay. Therefore, we think that this system is an exciting model for community genetics in the -omics era.

Ectomycorrhizal fungi show a continuum of specialization to the host tree from strict specialists to generalists. Differences in the expansion of multigene families, in particular dynamic repertoires of genes encoding small secreted proteins and sugar-cleaving enzymes, might be responsible for the different host ranges of specialists, such as *T. melanosporum*, and generalists, such as *L. bicolor* (Martin *et al.* 2010). That is, the genome expansion observed in *L. bicolor* might be driven by selection of the symbiont to exploit diverse substrates provided by multiple potential hosts and by diverse soils. As more genomes of mycorrhizal fungi are sequenced (Martin *et al.* 2011), this hypothesis will become testable.

In addition to the genomics of host–symbiont interactions, studies of geographical patterns of co-evolution add to our knowledge of processes leading to reciprocal adaptation and specialization. There are only a handful of studies reporting the structure of geographic variation and patterns of co-evolution in mycorrhizal interactions, indicating that these patterns are geographically highly variable (Hoeksema 2010; Hoeksema *et al.* 2012). To date, mostly higher-level traits, such as intensity of mycorrhizal colonization or growth of host trees, have been studied. Several of these studies found significant genetic variation in either the host plant or the mycorrhizal fungus in its ecological effect on the other partner. For example, the relationship between the colonization intensity of the ectomycorrhizal fungus *Thelephora terrestris* and the growth of its host, Lodgepole pine (*Pinus contorta*), depends on the tree's genotype (Karst *et al.* 2009). In poplar, both the intensity of colonization and the amount of enzymes secreted by poplar root tips colonized by *L. bicolor* are under the genetic control of the host (Courty *et al.* 2011). Similar findings come from arbuscular mycorrhizal systems, where host identity has a strong effect on the fitness of different strains of *Glomus intraradices* (Ehinger *et al.* 2009).

An increasing body of evidence shows that subtle intraspecific differences in the genome of host plants determine the composition of interacting communities in mycorrhizal fungi (e.g. Korkama *et al.* 2006; Whitham *et al.* 2006; Sthultz *et al.* 2009; Karliński *et al.* 2010; Leski *et al.* 2010; Hoeksema *et al.* 2012). We have experimental evidence that such as intraspecific genetic variation in the host also affects the composition of interacting mycorrhizal populations (Hoeksema & Thompson 2007), but this has not yet been tested under natural conditions. To understand the links between structure and diversity of communities and ecosystem functioning, we need to know more about spatio-temporal patterns of genetic variation. There are indications that both interspecific (e.g. van der Heijden *et al.* 1998; Maherali & Klironomos 2007) and intraspecific (e.g. Johnson *et al.* 2012) diversity of

mycorrhizal fungi can regulate productivity and ecosystem functioning. We advocate studies of community and population diversity in forests and combining them with functional field studies, involving both partners of ectomycorrhizal symbioses. Numerous new techniques are emerging for gene expression studies, marker gene evaluation using comparative genomics, and enzyme activity profiling of whole ectomycorrhizal assemblages (Courty *et al.* 2010). The rapid development of high-throughput sequencing technologies facilitates the survey and comparison of whole microbial communities (Buée *et al.* 2009), although analysis, interpretation, and publication of data still needs to be optimized (Henrik Nilsson *et al.* 2012). Nevertheless, combined genotypic and functional studies are now feasible and may be expanded to natural and experimental gradients. Several reports indicate that soil microbe and mycorrhizal diversity differentially affect ecosystem functioning under different environmental conditions, e.g. nutrient status (van der Heijden *et al.* 2008). We also know that plant-associated microorganisms are an important factor influencing plant responses to climate change (Courty *et al.* 2010; Pickles *et al.* 2012). Combined genotypic and functional studies in diverse environments will help to understand current patterns and to predict changes and effects in the future.

Associations between genes and traits: potential of next generation approaches in community genetics

An essential part of future studies in community genetics will be to identify the genes that underlie the traits of hosts that affect associated organisms. For this, sequencing of the complete genome of a host species is not sufficient. Rather, it is essential to link the presence or action of particular variants of genes or genomic regions of a host plant to the presence or abundance of associated organisms or arrays of their genes. There are basically two strategies for this, namely QTL mapping and genome-wide association studies (GWAS). We briefly outline and illustrate below the pros and cons of these two approaches for community genetics.

An example of QTL mapping of community traits of poplar is a study aimed at identifying genomic regions associated with susceptibility to insects (DeWoody *et al.* unpubl.data). Parents and progeny of a poplar (*Populus trichocarpa* × *P. deltoides*) F2 mapping population were assessed for various categories of leaf damage, including chewers and skeletonizers. The damage levels significantly varied among offspring genotypes. Each category was treated as a quantitative trait in a QTL mapping approach and more than ten QTLs were detected. QTLs also varied seasonally, suggesting that the insect community responds to traits and the underlying genetic variation over time. This underlines the importance of considering temporal variation in studies of community genetics, as noted above.

Another example is a study on QTLs affecting ectomycorrhizal symbiosis in a *P. deltoides* × *P. trichocarpa* F1 population (Labbé *et al.* 2011). Four identified QTLs were associated with candidate genes, and differential transcript levels were assessed with the help of a whole-genome microarray. The transcripts with the highest overrepresentation were, based on their gene ontology, in the repress defense mechanisms and in pathogen resistance.

Relatively few mapping populations have been produced for long-lived tree species, due to the length of time needed to maintain and study them, and the high costs associated with it. As a single cross will not contain all alleles present in a large population of an outcrossing species, not all QTLs can be detected in a single cross, and most QTL interactions will go unnoticed. Hence, several populations are necessary, and producing them would be an important investment. Next to full-sib families it may be possible to use full or partial diallel designs with multiple parents, so that more alleles are included and many more allele combinations can be studied, similar to MAGIC populations (Kover *et al.* 2009) but without the need for selfing to multiply and maintain the population.

In the meantime, an elegant alternative for forest trees is to use existing progeny trials. Many of these have been established and often replicated at different locations, and phenotypic data are usually available for extensive periods of time. Many trials consist of half-sib families, in which the alleles from the mother segregate in the progeny. If only a limited number of fathers were involved, genotyping may even allow them to be split into a few interconnected full-sib families. Common garden experiments often include a sample of the diversity of an area. When these experiments are replicated at multiple sites, it may be possible to perform genome-wide association mapping with the advantage of multi-site / multi-year data.
An issue for community genetics, as mentioned above, is that the local species pool may be different between the locations of the trials. This can be tackled efficiently by replicating the populations and planting them in different locations. Replicated populations will also spread the risk of losing individual members of the populations.

After finding a QTL region based on the presence of an associated organism or, for example, damage caused by an insect species, the underlying mechanism can be unravelled, in this case by measuring the secondary compound composition of all progeny trees and locating such traits on the genetic map. Co-localization of a compound with a QTL would suggest that it was responsible for the effect on the insects and that a structural or regulatory gene involved in its synthesis is located in that genomic region. In some species, this can be tested by mutant analysis, but it is not practical with trees. Alternatively, one could analyse the naturally occurring genetic variation in a large set of unrelated trees

TABLE 1

From genes of focal species to traits of the extended phenotype – and back: questions and experimental considerations, related to (a) spatio-temporal variation, (b) the application of -omics approaches, and (c) reciprocal effects to stimulate future studies in community genetics

Theme	Questions	Experimental considerations
(a) Spatio-temporal variation	To what degree do regional species pools determine the composition of organisms associated to particular genotypes?	Assess naturally occurring spatial replicates of particular genotypes, e.g. agricultural, horticultural or silvicultural clones, and perform regionally replicated experiments using the same (set of) genotypes exposed to various regional species pools of potentially associated organisms.
	What is the relevance of phylogeographic structure in host species for the composition of associated communities?	Consider genetic structure and evolutionary lineages of the focal species.
	How do relationships between genotypes and associated organisms vary among seasons or among life stages?	Perform temporally replicated experiments or monitor natural communities across >1 year; establish long-term experiments with host plants from seedlings to mature adults.
	How does landscape configuration, e.g. differences in the relative abundance of, or connectivity among, particular habitat types, affect regional species pools and, thus, the communities of associated organisms in a focal species?	Include landscape characteristics when setting up experimental plots or assessing natural communities.
	To what degree does phenotypic plasticity shape extended phenotypes?	Set up common garden experiments along ecological gradients including reciprocal transplants to test for genotype-by-environment interactions and reaction norms.
(b) -omics approaches	Which QTL relate to particular groups of associated organisms?	Establish various fullsib families or diallel crosses to include a wide range of allele variants.
	What (classes of) compounds differ among host genotypes that are differentially affected by groups of associated organisms?	Genome/transcriptome sequencing of pools of host plants differing in their associated communities.
	Do traits affecting community composition of associated species rely on single or multiple genes, and how large is their allelic variation within host populations?	Identify genes directly involved in the interaction, e.g. through QTL mapping, and quantify the degree of polymorphism using high-throughput, reduced-representation sequencing.
	Does one gene of a focal species influence a single, a group of, or all associated species?	Use feeding (herbivores) or inoculation (ectomycorrhizae) experiments and perform co-expression profiling and subsequent protein annotation.
	How many such genes exist, given that a focal species may interact with hundreds of associated species?	Perform gene expression studies of focal species that are experimentally associated with different single species or groups of species of associated organisms.
(c) Reciprocal effects	How do different groups of associated species induce changes in the phenotypic traits (and the underlying allele frequencies) of the host?	Expose the same (set of) hosts to different (sets of) associated species and test for changes in traits and allele frequencies over time.
	What genes in host and associated species determine whether they interact as generalists or specialists?	Combine comparative genomics and expression profiling among generalists and specialists in both hosts and associated species.

with different combinations of compounds and conduct association tests (i.e. GWAS).

Genome-wide association studies assume that, in the absence of population substructure, markers that are physically linked to a gene associated with a phenotype of a trait can be distinguished from markers that are not linked, as the latter are assumed to occur randomly in individuals of the population regardless of the phenotype (Nordborg & Weigel 2008). There is no need to construct a mapping population as in QTL detection, but a reference genome or a dense genetic map in combination with sufficient linkage disequilibrium (LD) are required (Kim *et al.* 2007). LD appears to be limited in tree species (Ingvarsson 2005; Heuertz *et al.* 2006; Pyhäjärvi *et al.* 2007), which implies that high-density genetic marker arrays are needed for applying association mapping and that many more individuals need to be studied. For instance, Fournier-Level *et al.* (2009) tested target candidate genes and identified the functional variation responsible for the observed variation in anthocyanin variation in grape by association analysis. The very low LD often encountered in natural tree populations (Neale & Savolainen 2004) will assist in finding many of the possible combinations of compounds, thus increasing the power of the association study. A new approach, becoming feasible because of high-throughput sequencing technology, is to pool and sequence DNA from multiple individuals within a population with clearly distinct phenotypes or habitat conditions (Turner *et al.* 2010), and to identify those markers across the genome that display a large difference in allelic frequency between the pooled groups (Holderegger *et al.* 2008). The advantage of this 'population resequencing' approach, which vaguely resembles bulked segregant analysis (BSA), is that no mapping population or extensive LD is necessary; the drawback is that an annotated genome is still needed for reference. Since annotated genome sequences are increasingly becoming available, this will be less of a problem in the future. The approach can be readily extended to polygenic traits (Heard *et al.* 2010). A potential application to community genetics in trees would be to pool the DNA from trees that host a particular insect with DNA from those that do not, and compare the sequenced genomes of the two groups.

Next generation methods now enable genotyping-by-sequencing (Baird *et al.* 2008). In the context of segregating populations, restriction-site associated DNA (RAD) markers or transcriptome sequencing permit direct mapping-by-sequencing, thus skipping marker development altogether (Hartwig *et al.* 2012; Zhu *et al.* 2012). In QTL mapping this solves the problem of generating dense maps, so that the limiting factor for high resolution is the number of recombinations or the size of the segregating population. As forest trees have very small LD, the ability to generate high volumes of genomic data is a very promising development for GWAS.

Gene expression profiling, a complementary approach to association genomics as a strategy for functional genomics, is also being revolutionized by developments in next generation technologies. Gene expression profiling has been applied to study stress response in trees, for example following insect attack where transcript analyses by cDNA microarray profiles have been combined with 2-D protein and protein spectrometric analyses (Lippert *et al.* 2007). In this pioneering work on pines and pine weevils, the authors demonstrated that transcripts and their proteins were complimentary. Next generation sequencing of tagged cDNA ends now enables researchers to quantify the number of transcripts from different subsets of individuals (Xu *et al.* 2009). Given the availability of gene annotations, the transcripts will be associated with gene models and their regulators using publicly available databases. We expect that co-expression profiling will become feasible for populations as well as for individual ontogenetic stages of interacting species. Such an approach may also be scaled up from two-species interactions to multiple-species interactions, i.e. a true 'community transcriptome' approach.

Proteomic approaches allow for an efficient and simultaneous detection of the proteins in a sample. The proteome composition to some extent integrates fluctuations in expression over a period of time, thus potentially being robust with regard to sampling time in the field. The identification of peptides relies on either a large, high-quality RNA-seq dataset, a complete set of alleles from a multigene family, or the genome sequence. An example is the use of peptide identification (Q-TOF LC-MS^E) for fast screening of Bet v 1 isoforms in pollen of various birch species, as it was possible to determine both presence and relative abundances of individual isoforms (Schenk *et al.* 2009). For this, the mass spectra obtained from the pollen were compared with a set of predicted peaks based on a complete set of isoforms obtained by sequencing the genes. In species for which the genome sequence or a large amount of transcriptome data is available, this prediction becomes a relatively simple bioinformatics exercise.

Other -omics techniques, such as metabolomics, may be employed in similar experimental schemes. Recent advances have increased the sensitivity and throughput of metabolomics and proteomics assays ('next-gen biochem'). Now, one can directly map QTL controlling the metabolic profile of all offspring of a cross. For instance, untargeted GC-TOF-MS metabolite profiling allowed mapping of 100 mQTLs (Carreno-Quintero *et al.* 2012). The main drawbacks of metabolomics are the higher costs and the problem of interfering factors due to the different growing conditions of the trees included in the association analysis. Moreover the samples cannot be all taken at the same time. On the other hand, the compounds measured are also the ones that affect the interaction with associated insect species. So if genetic variation in multiple

genes affects the content of one important compound, the association of the compound with presence or absence of one or more insect species will be stronger than that of each of the underlying genes, and the association will also be more informative on the mechanism of the interaction. Even GWAS could be done in this way. In our example using a pool of trees including those that host a particular insect and those that do not, a comparison of compounds may be more straightforward than comparing DNA markers. In particular, if the insect is not always present on the same trees across years, the compounds present in each tree in each year could reveal a strong correlation, whereas the genes that enable the tree to produce the compounds would not.

If, as indicated above, a compound affects the presence of insect species, then one would expect, reciprocally, the presence of catabolites of the compound in insect species that tolerate the compound, when these insects are sampled on the trees that produce it. This can be used to experimentally validate the statistical associations between compounds in the tree and the presence of insect species or guilds, and for a starting point for understanding the mechanisms behind the interactions between trees and insects.

Perspectives

A suite of -omics approaches is available to pave the way for studying entire communities. Accordingly, we need to refine hypotheses and develop suitable study designs and statistical tools (Augustin *et al.* 2010; Ovaskainen *et al.* 2010), which will improve implementation once reduced costs make these tools applicable to large-scale sampling of community-level interactions (Table 1).

As outlined above, we see two main directions that should be followed in community genetics to substantiate inference on the interplay of genes, organisms, communities, and their respective environments. First, joint descriptive and experimental studies should include spatial and temporal gradients to account for environmental variation in these dimensions (Thompson 2005; Crutsinger *et al.* 2009; Tack *et al.* 2010). Second, researchers in community genetics should make better use of the exponentially increasing genomic information becoming available, which will requires solid expertise in bioinformatics. If this is achieved, gene-to-gene interactions can be explored in individual-based associations and at the level of entire communities and shift community genetics towards becoming community genomics.

Moreover, community genetics goes beyond the effects of genotypes in one species on the community of associated organisms. We also need to consider the reciprocal effects of how associated communities shape the genotypic composition of their hosts and of how the genotypes of associated species affect host communities (Fig. 1). There are virtually no studies available on this aspect of community interactions, which leaves a wide-open field of empirical research for the future. Exploring reciprocal interactions might help to extrapolate population genomics and quantitative genomics of focal species. We will then need to adopt a community-based understanding of selection and drift as well as to include G x G x E interactions into reaction norm calculations. However, elaborating on this subject goes beyond the scope of the present article.

In conclusion, we believe that the amalgamation of traditional population genetics, quantitative genetics and ecology, fostered by the advent of new genomic technologies, will revolutionise our perception of community and ecosystem processes and push community genetics into a new era.

Acknowledgments

Our sincere thanks go to Rolf Holderegger and the anonymous reviewers for helpful comments on earlier versions of the manuscript. Karen A. Brune and Stephen Cavers provided linguistic editing. This work was financed through the EC-supported Network of Excellence Evoltree (GOCE-016322).

REFERENCES

Agrawal AA (2003) Community genetics: new insights into community ecology by integrating population genetics. *Ecology*, **84**, 543–544.

Augustin S, Denux O, Castagneyrol B *et al.* (2010) Insect herbivory response to Populus nigra genetic diversity. Fifth International Symposium, Poplars and willows: from research models to multipurpose trees for a bio-based society. Orvieto, Sept. 20–25, 2010, p. 207.

Bailey JK, Hendry AP, Kinnison MT *et al.* (2009) From genes to ecosystems: an emerging synthesis of eco-evolutionary dynamics. *New Phytologist*, **184**, 746–749.

Baird NA, Etter PD, Atwood TS *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, **3**, e3376.

Bangert RK, Turek RJ, Rehill B, Wimp GM, Bailey JK, Whitham TG (2006) A genetic similarity rule determines arthropod community structure. *Molecular Ecology*, **15**, 1379–1391.

Basset Y (2001) Communities of insect herbivores foraging on saplings versus mature trees of *Pourouma bicolor* (Cecropiaceae) in Panama. *Oecologia*, **129**, 253–260.

Boege K, Marquis RJ (2005) Facing herbivory as you grow up: the ontogeny of resistance in plants. *Trends in Ecology & Evolution*, **20**, 441–448.

Brändle M, Brandl R (2006) Is the composition of phytophagous insects and parasitic fungi among trees predictable? *Oikos*, **113**, 296–304.

Buée M, Reich M, Murat C *et al.* (2009) 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytologist*, **184**, 449–456.

Carreno-Quintero N, Acharjee A, Maliepaard C *et al.* (2012) Untargeted metabolic quantitative trait loci analyses reveal a relationship between primary metabolism and potato tuber quality. *Plant Physiology*, **158**, 1306-1318.

Courty P-E, Buée M, Diedhiou AG *et al.* (2010) The role of ectomycorrhizal communities in forest ecosystem processes: New perspectives and emerging concepts. Soil *Biology & Biochemistry*, **42**, 679–698.

Courty PE, Labbé J, Kohler A *et al.* (2011) Effect of poplar genotypes on mycorrhizal infection and secreted enzyme activities in mycorrhizal and non-mycorrhizal roots. *Journal of Experimental Botany*, **62**, 249–260.

Crutsinger GM, Cadotte MW, Sanders NJ (2009) Plant genetics shapes inquiline community structure across spatial scales. *Ecology Letters,* **12**, 285–292.

Crutsinger GM, Reynolds WN, Classen AT, Sanders NJ (2008) Disparate effects of plant genotypic diversity on foliage and litter arthropod communities. *Oecologia*, **158**, 65–75.

Dawkins R (1982) The Extended Phenotype: *The Gene as the Unit of Selection*. Freeman, Oxford.

Ehinger M, Koch AM, Sanders IR (2009) Changes in arbuscular mycorrhizal fungal phenotypes and genotypes in response to plant species identity and phosphorus concentration. *New Phytologist*, **184**, 412–423.

Ekblom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, **107**, 1–15.

Emerson BC, Gillespie RG (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology & Evolution*, **23***,* 619–630.

Fournier-Level A, Cunff LL, Gomez C *et al.* (2009) Quantitative genetic bases of anthocyanin variation in grape (*Vitis vinifera* L. ssp. *sativa*) berry: A quantitative trait locus to quantitative trait nucleotide integrated study. *Genetics*, **183**, 1127–1139.

Hartwig B, Velikkakam James G, Konrad KK, Schneeberger K, Turck F (2012) Fast isogenic mapping-by-sequencing of ethyl methanesulfonate-induced mutant bulks. *Plant Physiology*, **160**, 591–600.

Heard E, Tishkoff S, Todd JA *et al.* (2010) Ten years of genetics and genomics: what have we achieved and where are we heading? *Nature Reviews Genetics*, **11**, 723–733.

Henrik Nilsson R, Tedersoo L, Lindahl BD *et al.* (2012) Towards standardization of the description and publication of next-generation sequencing datasets of fungal communities. *New Phytologist*, **191**, 314–318.

Heuertz M, de Paoli E, Källmann T *et al.* (2006) Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce [*Picea abies* (L.) Karst]. *Genetics*, **174**, 2095–2105.

Hoeksema JD (2010) Ongoing coevolution in mycorrhizal interactions. *New Phytologist*, **187**, 286–300.

Hoeksema JD, Hernandez JV, Rogers DL, Mendoza LL, Thompson JN (2012) Geographic divergence in a species-rich symbiosis: Interactions between Monterey pines and ectomycorrhizal fungi. *Ecology*, **93**, 2274-2285.

Hoeksema JD, Thompson JN (2007) Geographic structure in a widespread plant–mycorrhizal interaction: pines and false truffles. *Journal of Evolutionary Biology*, **20**, 1148–1163.

Holderegger R, Herrmann D, Poncet B *et al.* (2008) Land ahead: using genome scans to identify molecular markers of adaptive relevance. *Plant Ecology and Diversity*, **1**, 273–283.

Holeski LM, Kearsley MJC, Whitham TG (2009) Separating ontogenetic and environmental determination of resistance to herbivory in cottonwood. *Ecology*, **90**, 2969–2973.

Ingvarsson PK (2005) Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European aspen (*Populus tremula* L., Salicaceae). *Genetics*, **169**, 945–953.

Johnson D, Martin F, Cairney J, Anderson IC (2012) The importance of individuals: intraspecific diversity of mycorrhizal plants and fungi in ecosystems. *New Phytologist*, **194**, 614–628.

Johnson MTJ, Agrawal AA (2005) Plant genotype and environment interaction to shape a diverse arthropod community on evening primrose (*Oenothera biennis*). *Ecology*, **86**, 874–885.

Johnson MTJ, Stinchcombe JR (2007) An emerging synthesis between community ecology and evolutionary biology. *Trends in Ecology & Evolution*, **22**, 250–257.

Karliński L, Rudawska M, Kieliszewska-Rokicka B, Leski T (2010) Relationship between genotype and soil environment during colonization of poplar roots by mycorrhizal and endophytic fungi. *Mycorrhiza*, **20**, 315-324.

Karst J, Jones M, Turkington R (2009) Ectomycorrhizal colonization and intraspecific variation in growth responses of lodgepole pine. *Plant Ecology*, **200**, 161–165.

Kim S, Plagnol V, Hu TT *et al.* (2007) Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics*, **39**, 1151–1155.

Korkama T, Pakkanen A, Pennanen T (2006) Ectomycorrhizal community structure varies among Norway spruce (*Picea abies*) clones. *New Phytologist*, **171**, 815–824.

Kosti I, Mandel-Gutfreund Y, Glaser F, Horwitz B (2010) Comparative analysis of fungal protein kinases and associated domains. *BMC Genomics*, **11**, 133.

Kover PX, Valdar W, Trakalo J *et al.* (2009) A multiparent advanced generation nter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genetics*, **5**, e1000551.

Kraft NJB, Cornwell WK, Webb CO, Ackerly DD (2007) Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *American Naturalist*, **170**, 271–283.

Labbé J, Jorge V, Kohler A *et al.* (2011) Identification of quantitative trait loci affecting ectomycorrhizal symbiosis in an interspecific F1 poplar cross and differential expression of genes in ectomycorrhizas of the two parents: *Populus deltoides* and *Populus trichocarpa*. *Tree Genetics & Genomes*, **7**, 617–627.

Le Corff J, Marquis RJ (1999) Differences between understorey and canopy in herbivore community composition and leaf quality for two oak species in Missouri. *Ecological Entomology*, **24**, 46–58.

Leski T, Aučina A, Skridaila A, Pietras M, Riepšas E, Rudawska M (2010) Ectomycorrhizal community structure of different genotypes of Scots pine under forest nursery conditions. *Mycorrhiza*, **20**, 473–481.

Lippert D, Chowrira S, Ralph SG *et al.* (2007) Conifer defense against insects: Proteome analysis of Sitka spruce (*Picea sitchensis*) bark induced by mechanical wounding or feeding by white pine weevils (*Pissodes strobi*). *Proteomics*, **7**, 248–270.

Maherali H, Klironomos JN (2007) Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science*, **316**, 1746–1748.

Martin F, Aerts A, Ahren D *et al.* (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature*, **452**, 88–92.

Martin F, Cullen D, Hibbett D *et al.* (2011) Sequencing the fungal tree of life. *New Phytologist*, **190**, 818–821.

Martin F, Kohler A, Murat C *et al.* (2010) Perigord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature*, **464**, 1033–1038.

Mopper S, Stiling P, Landau K, Simberloff D, Van Zandt P (2000) Spatiotemporal variation in leafminer population structure and adaptation to individual oak trees. *Ecology*, **81**, 1577–1587.

Nadeau NJ, Jiggins CD (2010) A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in Genetics*, **26**, 484–492.

Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. *Trends in Plant Science*, **9**, 325–330.

Nordborg M, Weigel D (2008) Next-generation genetics in plants. *Nature*, **456**, 720–723.

Novotny V, Basset Y, Miller SE *et al.* (2002) Low host specificity of herbivorous insects in a tropical forest. *Nature*, **416**, 841–844.

Ovaskainen O, Nokso-Koivista J, Hottola J *et al.* (2010) Identifying wood-inhabiting fungi with 454 sequencing – what is the probability that BLAST gives the correct species? *Fungal Ecology*, **3**, 274–283.

Pickles BJ, Egger KN, Massicotte HB, Green DS (2012) Ectomycorrhizas and climate change. *Fungal Ecology*, **5**, 73–84.

Plett J, Kemppainen M, Kale S *et al.* (2011) A secreted effector protein of *Laccaria bicolor* is required for symbiosis development. *Current Biology*, **21**, 1197–1203.

Pyhäjärvi T, Rosario García-Gil M, Knürr T, Mikkonen M, Wachowiak W, Savolainen O (2007) Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. *Genetics*, **177**, 1713–1724.

Rokas A, Abbot P (2009) Harnessing genomics for evolutionary insights. *Trends in Ecology and Evolution*, **24**, 192–200.

Rowntree JK, Shuker DM, Preziosi RF (2011) Forward from the crossroads of ecology and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 1322–1328.

Ruhnke H, Schädler M, Matthies D, Klotz S, Brandl R (2006) Are sawflies adapted to individual host trees? A test of the adaptive deme formation hypothesis. *Evolutionary Ecology Research*, **8**, 1039–1048.

Schenk MF, Cordewener JHG, America AHP, van't Westende WPC, Smulders MJM, Gilissen LJWJ (2009) Characterization of PR-10 genes from eight *Betula* species and detection of Bet v 1 isoforms in birch pollen. *BMC Plant Biology*, **29**, 24.

Schoonhoven LM, Loon JJAv, Dicke M (2005) *Insect–Plant Biology*, 2nd edn. Oxford University Press, New York.

Sthultz CM, Whitham TG, Kennedy K, Deckert R, Gehring CA (2009) Genetically based susceptibility to herbivory influences the ectomycorrhizal fungal communities of a foundation tree species. *New Phytologist*, **184**, 657–667.

Tack AJM, Ovaskainen O, Pulkkinen P, Roslin T (2010) Spatial location dominates over host plant genotype in structuring an herbivore community. *Ecology*, **91**, 2660–2672.

Thompson JN (2005) *The geographic mosaic of coevolution*. University of Chicago Press, Chicago.

Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV (2010) Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature Genetics*, **42**, 260–263.

Van de Velde W, Zehirov G, Szatmari A *et al.* (2010) Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science*, **327**, 1122–1126.

van der Heijden MGA, Bardgett RD, Van Straalen NM (2008) The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters*, **11**, 296–310.

van der Heijden MGA, Klironomos JN, Ursic M *et al.* (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature*, **396**, 69–72.

Whitham TG, Bailey JK, Schweitzer JA *et al.* (2006) A framework for community and ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics*, **7**, 510–523.

Whitham TG, Lonsdorf EV, Schweitzer JA *et al.* (2005) "All effects of a gene on the world": Extended phenotypes, feedbacks, and multi-level selection. *Ecoscience*, **12**, 5–7.

Whitham TG, Young WP, Martinsen GD *et al.* (2003) Community and ecosystem genetics: a consequence of the extended phenotype. *Ecology*, **84**, 559–573.

Wymore AS, Keeley ATH, Yturralde KM, Schroer ML, Propper CR, Whitham TG (2011) Genes to ecosystems: exploring the frontiers of ecology with one of the smallest biological units. *New Phytologist*, **191**, 19–36.

Xu WJ, Wang ZX, Qiao ZD (2009) Modified PCR methods for 3' end amplification from serial analysis of gene expression (SAGE) tags. *FEBS Journal*, **276**, 2657-2668.

Zhu Y, Mang H-g, Sun Q, Qian J, Hipps A, Hua J (2012) Gene discovery using mutagen-induced polymorphisms and deep sequencing: application to plant disease resistance *Genetics*, **192**, 139–146.

Author contribution

All authors have conceptually contributed to the content of the present article through their active involvement and the many discussions held during the EvolTree Jointly Executed Research Activity 3 (JERA 3) on community ecology, which was managed by BZ. FG designed and organized the article preparation, FG, RB, BC, HJ, MP, MJMS, and BZ drafted the various sections, and all main authors participated in the writing of the final text.

LONG DISTANCE GENE FLOW AND ADAPTATION OF FOREST TREES TO RAPID CLIMATE CHANGE

Antoine Kremer[1,2†], Ophélie Ronce[3†], Juan J. Robledo-Arnuncio[4], Frédéric Guillaume[5], Gil Bohrer[6], Ran Nathan[7], Jon R. Bridle[8], Richard Gomulkiewicz[9], Etienne K. Klein[10], Kermit Ritland[11], Anna Kuparinen[12], Sophie Gerber[13,2], Silvio Schueler[14]

[1] INRA, UMR1202 Biodiversité Gènes et Communautés, Cestas, F-33610, France; email: antoine.kremer@pierroton.inra.fr
[2] Université de Bordeaux, UMR1202 Biodiversité Gènes et Communautés, Talence, F-33410, France;
[3] Université Montpellier 2 - CNRS, Institut des Sciences de l'Evolution, F-34095 Montpellier, Cedex 05, France; email: ophelie.ronce@univ-montp2.fr
[4] Department of Forest Ecology and Genetics, Forest Research Centre (CIFOR), INIA, 28040 Madrid, Spain; email: jjrobledo@gmail.com
[5] ETH, Department of Environmental Sciences, Universitätstrasse 16, 8092 Zürich, Switzerland; email: frederic.guillaume@env.ethz.ch
[6] Department of Civil, Environmental & Geodetic Engineering, Ohio State University, Columbus, OH 43210, USA; email: bohrer.17@osu.edu
[7] Movement Ecology Laboratory, Department of Ecology, Evolution and Behavior, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Jerusalem 91904, Israel; email: rnathan@cc.huji.ac.il
[8] School of Biological Sciences, University of Bristol, Bristol, BS8 IUG, UK; email: Jon.Bridle@bristol.ac.uk
[9] School of Biological Sciences and Department of Mathematics, Washington State University, Pullman, Washington 99164, USA; email: gomulki@wsu.edu
[10] INRA, UR Biostatistiques & Processus Spatiaux 546, F-84914 Avignon, France; email: Etienne.klein@avignon.inra.fr
[11] Department of Forest Sciences, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; email: kermit.ritland@ubc.ca
[12] Ecological Genetics Research Unit, Department of Biosciences, University of Helsinki, Helsinki FI-00014, Finland; e-mail: anna.kuparinen@helsinki.fi
[13] INRA, UMR1202 Biodiversité Gènes et Communautés, Cestas, F-33610, France; email: sophie.gerber@pierroton.inra.fr
[14] Federal Research and Training Centre for Forests, Natural Hazards and Landscape, Seckendorf-Gudent-Weg 8, 1131 Wien, Austria; email: silvio.schueler@bfw.gv.at
[†] Co-first authors

Corresponding Author:
Antoine Kremer, INRA
UMR1202 Biodiversité Gènes et Communautés 69 route d'Arcachon, Cestas, F-33610, France
Email: antoine.kremer@pierroton.inra.fr

Forest trees are keystone species, dominating many terrestrial ecosystems in regions where the most pronounced climate changes will occur. Climate change generates new adaptive challenges for trees. Their life history traits could either constrain or accelerate their adaptation. On one hand, long generation time can slow down evolutionary responses. On the other hand, long distance (LD) gene flow could compensate for their long generation time, facilitating evolutionary change in a shifting climate. We critically examine the latter hypothesis, by reviewing data and theory about the extent of gene flow in trees and its evolutionary consequences. Abundant evidence of LD effective dispersal indicates that genes may move within one generation over larger scales than the predicted shifts of tree habitat. Gene flow can have antagonistic effects on adaptation and persistence in the specific temporal and spatial frame of predicted climate change. Both theory and empirical data however suggest that the positive effects of LD gene flow in forest trees may dominate in many instances. The balance between the different effects of gene flow may however differ between the leading edge, the core and the rear of the distributions. Finally, we suggest future experimental and theoretical research areas for a better integration of trees dispersal biology and evolutionary quantitative genetics.

Key words: selection, adaptation, gene flow, climate change, forest trees.

Published in *Ecology Letters* (2012) 15: 378-392, with doi:10.1111/j.1461-0248.2012.01746.x

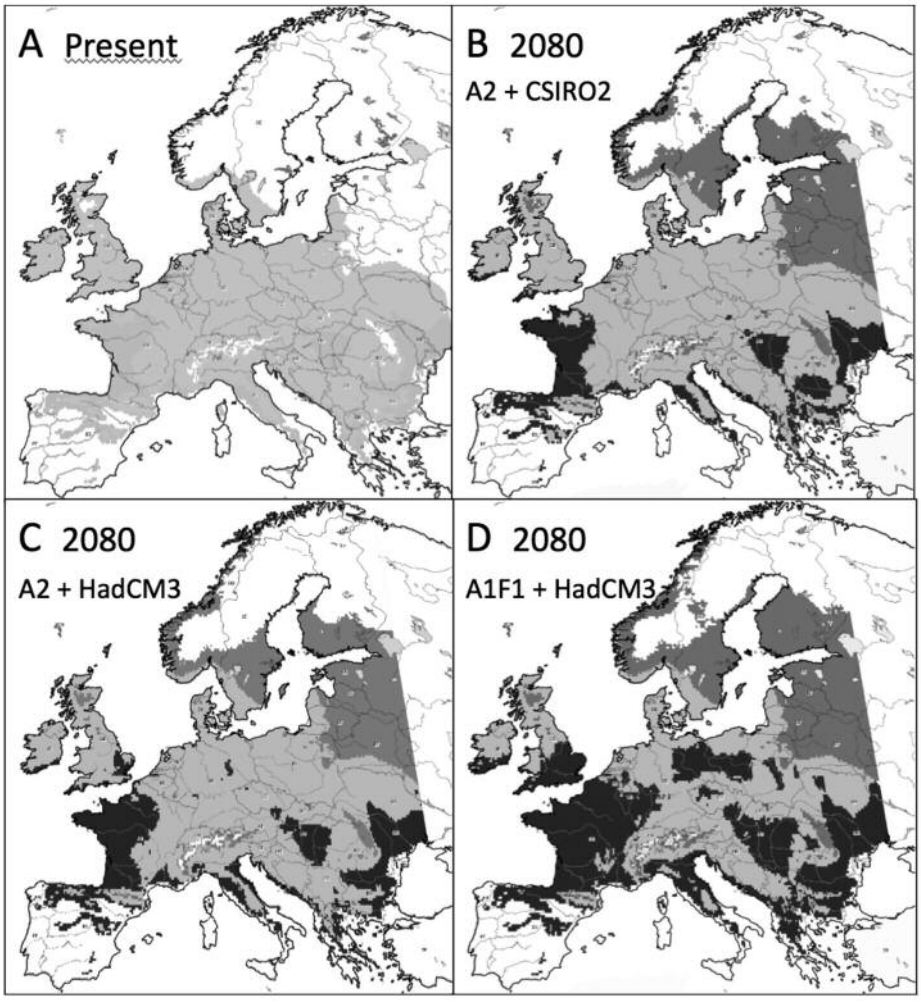
Introduction

While evidence of climate change and its impact on world's biota is steadily increasing, so are our concerns about the biological or human mediated capacities of species and populations to cope with these changes (Solomon *et al.* 2007). Such concerns are presumably more acute for sedentary and long lived organisms such as trees, which are less likely to track favourable conditions fast enough by migration. Furthermore, trees constitute a large ecologically and economically important functional group of woody plants that dominate many terrestrial ecosystems in regions where the most pronounced climate changes are projected to occur. The near-surface temperature is expected to shift northwards in mean rates of 110-430 m yr^{-1} during the 21st century for Mediterranean, temperate and boreal forests, the major forest biomes in mid and high latitudes (Loarie *et al.* 2009). Local estimates of this shift, in a scale of 1 km2, vary by three orders of magnitude (0.01 to 10 km yr^{-1}). Niche modelling under various climatic and greenhouse emission predictions suggests that *bioclimatic envelopes* (glossary) for forest trees will shift northwards in North America (Iverson *et al.* 2008) and north-eastwards in Europe (Thuillier 2003). The estimated shift distance varies from 300 to 800 km within one century depending on the climate or greenhouse gas emission scenario (Mc Kenney *et al.* 2007), with considerable variation across both models and species. The large *bioclimatic envelope* of many forest trees hides a collection of highly differentiated populations and genotypes with contrasting adaptation to local climate (Box 1). Shifts in *bioclimatic envelopes* are therefore likely to generate not only potential extinction and recolonization, but also large reorganization of genetic diversity within the species range if divergent or locally-adapted populations respond in different ways.

With climate change, environments change continuously and the optimal sets of adaptations maximizing fitness under local conditions may shift accordingly. The evolutionary responses of populations can then be pictured as a race where populations are tracking the moving optima both in time (Bürger & Krall, 2004) and in space (Pease *et al.* 1989; Polechova *et al.* 2009). Migration and adaptation are often perceived as alternative responses to these challenges (Aitken *et al.* 2008) because evolution allows populations to adapt to novel conditions without migrating, whereas migration lets populations track favourable conditions without evolving. Range shifts and adaptation can also occur simultaneously (e.g. Cwynar & MacDonald 1987). Seed dispersal allows colonization of new favourable habitat. However, both seeds and pollen dispersal in trees affect the spread of genetic variation within the range. Such gene flow affects adaptation by shaping the distribution of genetic variation both within and among populations (Lopez *et al.* 2008). Trees are characterized by their particular life history, combining long generation time (allowing divergent strategies at different life stages) and the capacity for long dispersal distances through pollen and seeds. Given the anticipated intensity and directionality of climatic change, do trees have the adaptive capacity to respond and how will gene flow affect that response? Valuable insight into

FIGURE 1

Predicted shifts of *bioclimatic envelopes* of sessile oak (*Quercus petraea*) in Europe (according to Thuiller 2003)



Predicted *bioclimatic envelopes* of sessile oak in 2080, assuming that correlations between present distribution (Panel A, light grey area) and climatic data are maintained. Climate of black areas would not be suited any more to sessile oak in 2080, while climate of dark grey areas would become favourable. Overall shifts of several hundred kilometres are foreseen which provide some hints on the scale of gene dispersal needed to track climate change. Predictions were made according to different IPCC models of greenhouse gas emissions (GG) and climatic changes (CC) (Solomon *et al.* 2007). Panel B: GG is A2 and CC is CSIRO2; Panel C: GG is A2 and CC is HadCM3; Panel D: GG is A1F1 and CC is HadCM3.

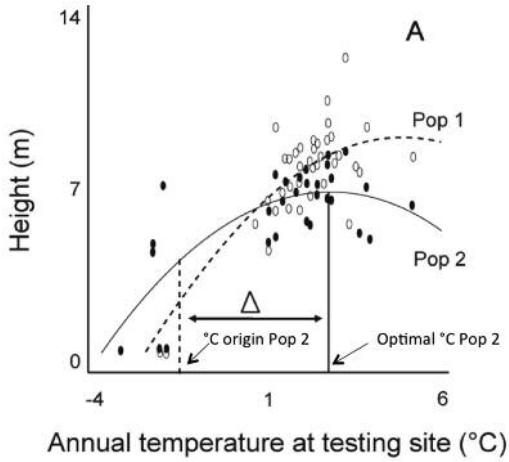
these issues is provided by the study of evolutionary changes in trees during the climate change that occurred following the last glaciations (Petit *et al.* 2008). Because of their economic importance, local adaptation in forest trees has very early on been the subject of intensive study at very large spatial scales. Decades-old common garden experiments of forest trees in the Northern hemisphere suggest that an interaction between divergent selection across contrasted environments and large pollen flow maintained enough diversity in local forest trees populations to support adaptability to past changing environments (Kremer *et al.* 2010). Whether interaction between gene flow and selection will be as efficient in the future remains unclear, as the predicted rates of environmental changes might exceed historical ones.

We here critically examine the hypothesis that long distance (LD) gene flow could compensate for the long generation time of trees, facilitating evolutionary change in a shifting climate, by reviewing both data and theory, about the extent of gene flow in forest trees and its evolutionary consequences. We first review the recent literature on long distance pollen and seed dispersal in trees and show that it can match the predicted climate change velocity within one generation. We then review the theoretical predictions and experimental evidence for the

BOX 1

Provenance tests and norms of reaction

Provenance tests are common garden experiments that gather usually very large number of populations planted by forest geneticists in multiple replicates over decades (Morgenstern 1996). They provide crucial information on the level of genetic variation within and between populations for fitness related traits, resulting from a balance of divergent selection across populations, gene flow, and random genetic drift. Extensive surveys of genetic diversity and variation have been conducted in these experiments and indicate that (1) extant populations harbour large levels of genetic variation (Hamrick *et al.* 1992) continuously replenished by extensive gene flow (2) *adaptive traits* exhibit high levels of population differentiation, despite gene flow, as a result of strong divergent selection (Savolainen *et al.* 2007) (3) clinal patterns of population differentiation along climatic or geographical variables are congruent across species suggesting systematic adaptive responses to directional selection, particularly for phenological traits; and (4) the extant distribution of between versus within population differentiation for fitness related traits has developed rapidly following post glacial recolonisation and is not the legacy of ancient population structure (Kremer *et al.* 2010). Furthermore, when replicated *provenance tests* were established, *reaction norms* of populations can be constructed that visualise their response across a wide range of environmental conditions (Rehfeldt *et al.* 1999; Rehfeldt *et al.* 2002). *Reaction norms* of fitness-related traits follow generally quadratic functions. Panel A illustrates the reactions norms of two populations of *Pinus contorta* (Rehfeldt *et al.* 1999) for height at age 20. Δ accounts for the difference between the climate of the site where the population stems from (dotted line) and the optimal climate corresponding to the site where the population exhibits the highest value for height (bold line). Such *reaction norms* show that the climatic tolerances of individual *provenances* are narrower than the whole species climatic envelope, that climatic optima for growth differ among local populations, correlating with their climate of origin, and that climatic tolerance only partially overlap between *provenances*. These tests also suggest that populations located at the extremities of the natural distribution inhabit climates that are suboptimal for their growth and development. For population 2 (coming from latitude 59.1°N), as illustrated in A, the optimal climate (2.5°C) is warmer than the climate of its geographic source (-2.5°C), while for southern populations the opposite pattern occurs. The overall picture is a negative correlation between Δ and the latitude of origin of the population as shown in panel B by the example in the case of *Pinus contorta* (Rehfeldt *et al.* 1999). This pattern is consistent with theoretical predictions that asymmetric gene flow from the core to peripheral populations increases maladaptation at the edges of the natural distribution (Kirkpatrick & Barton, 1997; Garcia-Ramos & Kirkpatrick 1997, see also text).



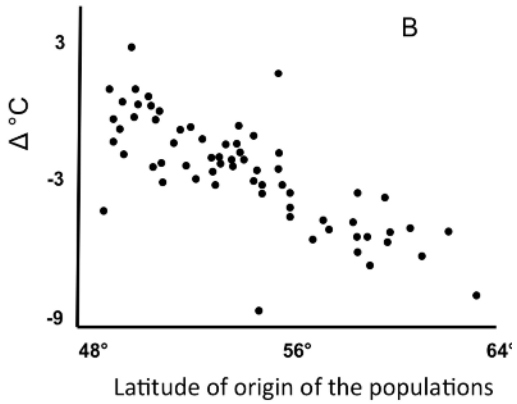


TABLE 1

Examples of observed LD pollen and seed dispersal in trees (more than 3 km for pollen and 1 km for seeds). The Table is arranged first by propagule type (pollen or seed), than by vector type (wind, insects, birds, bats, elephants, fish or generally vertebrates) and dispersal type (potential, viable or effective, defined in a footnote below), and finally alphabetically by species name.

Species	Dispersal system			Location	Method	Dispersal Distance		Reference
	Propagule	Vector	Type[a]			Maximum	Proportion≥ threshold[b]	
Betula spp.	Pollen	Wind	Potential	Central-North-Eastern Europe	Aerobiologic and phenological analysis	1000 km		Siljamo *et al.* 2008
Pinus banksiana and *Picea glauca*	Pollen	Wind	Potential	Canada	Aerobiologic analysis	3000 km		Campbell *et al.* 1999
Pinus sylvestris	Pollen	Wind	Viable	Northern Europe	Aerobiologic and phenological analysis	600 km		Varis *et al.* 2009
Pinus taeda	Pollen	Wind	Viable	Eastern North America	Aerobiologic analysis	40 km		Williams 2010
Cecropia obtusifolia	Pollen	Wind	Effective	Central America	Genetic paternity analysis		10% > 14 km	Kaufman *et al.* 1998
Fraxinus excelsior	Pollen	Wind	Effective	Scotland	Genetic parentage analysis		25-35% > 3 km	Bacles *et al.* 2006
Pinus sylvestris	Pollen	Wind	Effective	Spain	Genetic mixture analysis		4.3% > ~100 km	Robledo-Arnuncio 2011
Quercus robur	Pollen	Wind	Effective	Eastern Europe	Genetic parentage analysis		35 % > 80 km	Buschbom *et al.* 2011
Populus trichocarpa	Pollen	Wind	Effective	Western North America	Genetic paternity analysis		5% > ~5-10 km	Slavov *et al.* 2009
Ficus spp.	Pollen	Insects	Effective	Central America	Genetic parental reconstruction	14 km (isolated mother trees)		Nason *et al.* 1998
Ficus sycomorus	Pollen	Insects	Effective	Namibia	Genetic paternity analysis	165 km		Ahmed *et al.* 2009
Sorbus domestica	Pollen	Insects	Effective	Central Europe	Genetic paternity analysis		~1% > 12-16 km	Kamm *et al.* 2009
Swietenia humilis	Pollen	Insects	Effective	Central America	Genetic paternity analysis		40-80% ≥ 4 km (in small fragments)	White *et al.* 2002
Fraxinus excelsior	Seed	Wind	Effective	Scotland	Genetic parentage analysis	1.4 km	46-53% > 3 km	Bacles *et al.* 2006
Annona glabra	Seed	Birds	Potential	Australia	Empirically-based simulations of vector movements and seed passage time	5.2 km	1% > 4 km	Westcott *et al.* 2008
Xylopia hypolampra and 7 other *species*	Seed	Birds	Potential	Cameroon	Empirically-based simulations of vector movements and seed passage time	6.9 km		Holbrook & Smith 2000
Ficus carica and *Morus alba*	Seed	Bats	Potential	Israel	Empirically-based simulations of vector movements and seed passage time	20 km	17% > 1 km	Tsoar *et al.* 2011
Tamarindus indica	Seed	Elephants	Potential	Myanmar (Burma)	Empirically-based simulations of vector movements and seed passage time	5.4 km	50% > 1.2 km	Campos-Arceiz *et al.* 2008
Duroia duckei and 2 other species	Seed	Fish	Potential	Peru	Empirically-based simulations of vector movements and seed passage time	5.5 km	5% > 1.7 km	Anderson *et al.* 2011
Prunus mahaleb	Seed	Vertebrates	Potential	Spain	Genetic maternal analysis		33% >1500 m	Jordano *et al.* 2007
Sorbus domestica	Seed	Vertebrates	Effective	Central Europe	Genetic paternity analysis	12.2 km		Kamm *et al.* 2009

[a]Three types are distinguished: "potential" dispersal is the distance dispersed by a propagule (pollen or seeds) at any, commonly unknown, condition; "viable" is the same as "potential" but excluding non-viable propagules; cases of "effective" dispersal are the pollen that gave rise to seeds, or seeds that established, yielding seedlings, saplings or young/adult plants.

[b]The proportion (in %) of propagules dispersed to equal or greater distances than the specified threshold. The threshold distances were defined by the authors of each study, often arbitrarily or according to features of the study landscape and/or populations.

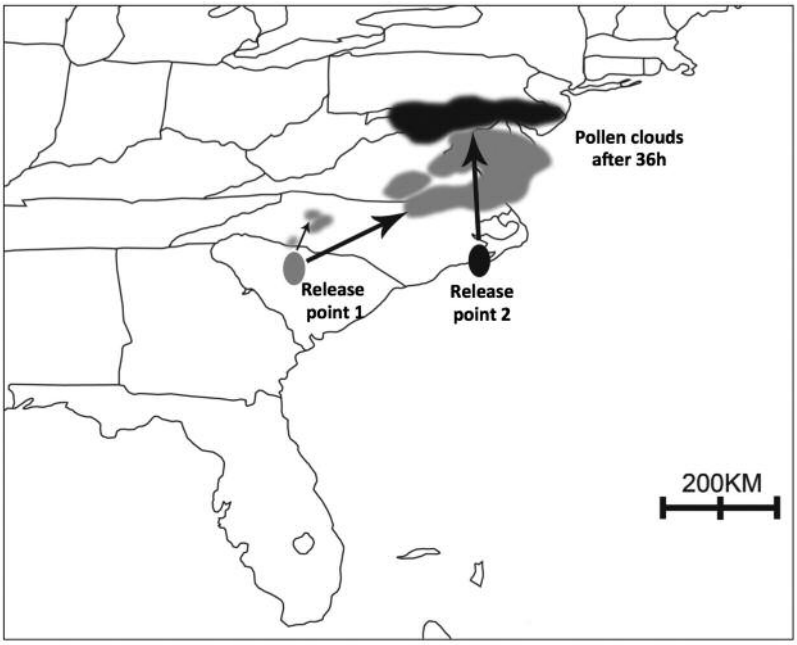
effects of gene flow on adaptation in trees. We conclude that the positive effects of gene flow may often dominate its negative effects, although regional variation may influence the balance of those effects. We finally elaborate on the theoretical and experimental approaches that should be implemented to improve our ability to predict the scale and distribution of gene flow effects on forest ecosystems in the context of climate change.

How far do seeds and pollen disperse in trees?

Gene flow in plants is mediated by both seed and pollen dispersal, which vary greatly among species (Ennos 1994). Seed and pollen dispersal however have distinct effects on the rate of demographic spread, and the rate at which genes move across the range of a species. The spatial scale of *effective* propagule *dispersal* (glossary) in trees depends on a variety of physical and biological processes that determine the amount and availability of pollen and seeds, their movement, their viability before and during movement, and the probability of successful pollination leading to viable seed and the seedling establishment rates. Different combinations of these components may yield *effective dispersal* (glossary) distances spanning from a few centimeters to thousands of kilometers (Nathan *et al.* 2008), generally following markedly leptokurtic patterns. Aerobiological studies show that airborne tree pollen (both viable and non-viable) has the potential to be transported in substantial amounts over hundreds to thousands of kilometers (Table 1, Fig. 2). However, documented dispersal distances of viable pollen (yet prior to successful fertilization) are about one order of magnitude shorter, up to 600 km (Table 1). Documented distances of effective pollen dispersal (when pollination led to successful mating) are of lower magnitude, up to 100 km (Table 1). Documented wind-driven effective seed dispersal is up to a few kilometers (Table 1), thus about two

FIGURE 2

Virtual long distance pollen dispersal of *Pinus taeda*



Virtual pollen release, using the Regional Atmospheric Modeling System (RAMS) and its Eulerian-Lagrangian particle transport module (HYPACT). This regional atmospheric simulation was forced with meteorological data from National Oceanic and Atmospheric Administration (NOAA) NCEP-DOE Reanalysis II data set. The experimental settings are described in Bohrerova et *al.* (2009). The figure shows a portion of the southeast United States, centered on eastern South Carolina. Pollen was arbitrarily released from two locations, in North Carolina outer banks (black point) and South Carolina (grey point), at a simulated afternoon on 27 March 2006, corresponding with the peak of pollen release at the Duke forest, NC. The dispersing pollen plumes (black for NC pollen, grey for SC pollen) are shown as "clouds." The wind was moderate, mainly toward the northeast. The figure shows a snapshot of the pollen plume at 6:00 AM, 36 hours after the release. The viability is resolved by the model as an additional property of the pollen. Mortality due to UV and vapour pressure deficit is calculated, with rates fitted to empirical equations, based on observations of a bench-scale experiment. The pollen in the image, 36 hours after release is about 40% viable.

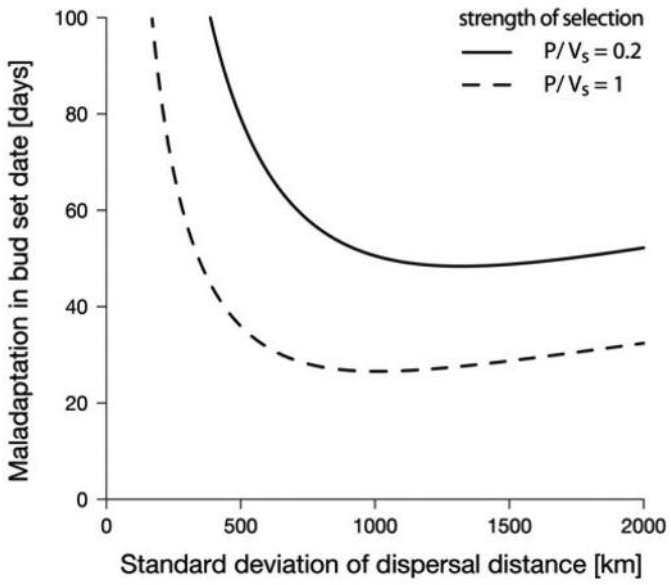
orders of magnitude shorter than effective pollen dispersal. Although animal-mediated seed dispersal can reach a scale of tens of kilometers, pollen dispersal distances are in general considerably longer than that of seeds, and especially in wind-driven dispersal systems.

Detecting mating events over hundreds of kilometers is experimentally very difficult and many biological processes take place (viability, phenology, pollen competition) between pollen release and fertilization. Similarly, quantifying the multiple demographic and dispersal parameters affecting seed dispersal and subsequent establishment is still hard to accomplish (but see Nathan *et al.* 2011). These experimental constraints and biological reality both explain the disparities between reported potential and *effective dispersal* distances. Their relative contributions remain unresolved and addressing this point will require innovation (see section 5.1). Interactions between physical and biological processes influencing long-distance dispersal (LDD) and the way these will be affected in a changing climate are particularly poorly understood. The local environment around the release point of the propagule and the conditions at which the propagule was released can have strong effects of either decreasing or increasing the rate and extent of LD propagule transport (Bohrer *et al.* 2008; Wright *et al.* 2008). Also important is the interaction between weather patterns that drive extreme LDD, and the conditions that affect survival during the dispersal event, particularly in pollen, where a migration event may span many hours and days (Schueler *et al.* 2005).

In that respect, predicting how climate change itself may modify patterns of gene flow in the future remains a formidable challenge. Projected changes such as warmer atmosphere (Kuparinen *et al.* 2009), or higher fecundity, earlier maturation, and changes (reduction or increase) in surface wind speed (Nathan *et al.* 2011) can be implemented into mechanistic models of plant spread in future environments. Although this approach facilitates assessing the upper and lower bounds of future gene flow and population spread, it is unlikely to provide accurate predictions for a particular species and system, due to the large uncertainty about key hard-to-measure parameters, such as the spatial patterns of phenological schedules for pollen and of post-dispersal survival of seeds. Moreover, genetic variation in traits affecting dispersal of both pollen and seeds is common in plants and, accordingly, plant dispersal traits have been observed to evolve fast in response to environmental change, especially in the context of range expansion (Darling *et al.* 2008). In particular, increased frequency of traits facilitating seed

FIGURE 3

Effect of dispersal distance on the evolutionary load



The average distance between population mean and optimal phenotype (i.e. maladaptation) is predicted from equation 2.2 in Box 2, as $\sqrt{2 \cdot V_S \cdot L}$. The model is illustrated with the example of the evolution of *bud set* date in the Sitka spruce as explained in Box 2. The standard deviation within populations for phenotypic expression of bud-set date is 17 days. *Heritability* was taken to be 0.5. No data is available on the strength of *stabilizing selection* V_S. Compiling the results of many studies (e.g. as shown in Johnson & Barton 2005), the median value of P/V_S is 0.2 (solid line). We also used a stronger value of *stabilizing selection* (P/V_S =1, dashed line). The optimal dispersal distance is smaller when selection is stronger (and thus the spatial fitness gradient steeper) but the *evolutionary load* is also much reduced (compare dashed and solid line).

dispersal have been found in recently founded populations, at the expanding edge of the range, while the converse trend was found in fragmented southern populations at the rear end (Riba *et al.* 2009). Identifying how global change (including climate change) will alter selection pressures acting on dispersal is therefore crucial in order to predict the extent of gene flow in future environments.

Despite gaps that prevent us from precisely predicting the extent of LD gene flow, the available data on *effective dispersal* and predicted habitat shifts still suggest that these two processes may operate over comparable scales for many tree species, but also that inter-specific variation in the rate and magnitude of LD gene flow may substantially affect the variation in the response of forest trees to climate change (Nathan *et al.* 2011, Table 1).

What are the possible effects of gene flow on adaptation?

We can conceptualize the potential effects of gene flow on adaptation through a simple model of species range evolution, where selection varies both in space and time (Pease *et al.* 1989; Polechova *et al.* 2009; Box 2). Local climate can be thought of as imposing specific selection pressures on a complex set of phenotypic traits (e.g.

BOX 2

A simple conceptual framework illustrating the antagonistic effects of gene flow

Two closely related theoretical models (Pease *et al.* 1989; Polechova *et al.* 2009) have explored the question of evolution in environments changing both in space and time, mimicking the effects of climate change in species with wide distributions. These models envision a species distributed along some linear environmental gradient, such as the Sitka spruce distributed along a large latitudinal gradient of temperatures (Mimura & Aitken 2007a). We use this empirical example to illustrate potentially realistic values of parameters for the model. We assume fitness depends quadratically on how well an individual tree is adapted to its local ecological conditions, i.e., how its phenotype matches the local optimum. By averaging over phenotypes in the local population, one may then write the mean fitness (here the Malthusian population exponential growth rate) in a given location as:

$$\bar{r} = r_0 - \frac{P}{2V_s} - \frac{(\bar{z}-\theta)^2}{2V_s}, \qquad (2.1)$$

where r_0 is the contribution to population growth of an individual with the optimal phenotype, V_s describes how well individuals that deviate from this optimal phenotype perform (and is thus inversely related to the strength of *stabilizing selection*), θ is the optimal phenotype in that location, $\bar{z}$ is the local mean phenotype in the population and P is the local phenotypic variance around this mean. This expression shows that the mean fitness in a variable population subject to stabilizing selection is reduced in two ways: (1) *Standing load* ($P/2V_S$): caused by phenotypic variation and present even when the mean phenotype matches the optimum; (2) *Evolutionary load* [$(\bar{z}-\theta)^2/2V_S$]: caused by departure of the mean phenotype from the local optimum (Lande & Shannon 1996). *Evolutionary loads* can be generated by selection that varies over space or time (see Bridle *et al.* 2009 for a review).

Pease *et al.* (1989) and Polechova *et al.* (2009) make specific predictions about how migration might affect the *evolutionary load* in a changing climate. Their models assume the optimum phenotype changes linearly through space, with slope *b*. This is similar to Sitka spruce, where *bud set* date increases linearly with the local mean annual temperature, which itself varies linearly with the distance to the Southern margin of the species range: assuming current bud set date corresponds to the optimum, this gives an estimate of *b* such that optimal *bud set* date increases by 13 days every °C, or by 3.24 days per 100 km (Mimura & Aitken 2007a; Aitken *et al.* 2008). Climate change can be approximated by this gradient of optimal phenotypes being constantly shifted through space at rate v. According to different climate models, mean annual temperature may increase by 3-5°C in the generation time of Sitka Spruce (Aitken *et al.* 2008), which gives an estimate of v as a shift of approximately 1000 to 2000 km per generation. Migration is modelled as a diffusion process, with average distance between parent and offspring σ. There are feedbacks between the evolution of the mean phenotype through time and space and that of the population density, mediated through gene flow and the local growth rates (Pease *et al.* 1989; Polechova *et al.* 2009).

Further assuming that genetic variation for the trait under selection is relatively weak and does not vary through space, Pease *et al.* (1989) predict that the loss of fitness at the scale of the range due to *evolutionary load* is approximately

$$L \approx \frac{1}{2}\left(\frac{\sigma b}{\sqrt{V_s}} + \frac{v^2}{\sigma^2} - \frac{G}{V_s}\right) \qquad (2.2)$$

where *G* is the genetic variance for the trait. Though this prediction might be crude in the case of forest trees with large within population *genetic variance*, it has heuristic value. Indeed, the first term in parentheses can be interpreted as the component of phenotypic mismatch due to spatial variability in the optimal phenotype and gene flow (*migration load*); this part of the load increases with dispersal distance (σ). The second term describes phenotypic mismatch due to the lagging response of the mean phenotype to temporal change in the local optimum (*lag load*). This part of the load decreases with dispersal distance because migration helps the species track its shifting optimum through space. The third term in (2.2) shows that the *evolutionary load* declines with genetic variance *G* because *response to selection* increases. Dispersal distance also affects the evolution of *genetic variance G* (Barton 2001; Polechova *et al.* 2009; Bridle *et al.* 2010), with positive effects on the *evolutionary load* (equation 2.2), but negative effects on the *standing load* $P/2V_S$ (equation 2.1)

phenology, frost hardiness, growth, seed size), and defining different optimal trait values through various trade-offs, depending on specific combinations of climatic conditions encountered within the range. As an illustration, in Sitka Spruce, trees originating from higher latitude with lower annual mean temperature cease growing earlier in the season than trees from lower latitude, when grown in common garden (Mimura & Aitken 2007a), suggesting different optimal *bud set* (glossary) dates along temperature gradients within the range. The simple conceptual model in Box 2 connects adaptation to demography by assuming that an individual's contribution to population growth declines as it departs from the locally optimal phenotype. This model suggests that gene flow has antagonistic effects on adaptation by modifying the various sources of *genetic load* (glossary) depressing population mean fitness, and thus population growth (see Bridle *et al.* 2009 for a review). We here review these effects by considering their alternative evolutionary consequences.

Gene flow constrains local adaptation

Because gene flow homogenises allele frequencies across space, high gene flow could constrain *adaptive divergence* (glossary) along environmental gradients (Garcia-Ramos & Kirkpatrick 1997; Bohrer *et al.* 2005; but see Barton 2001; Yeaman & Guillaume 2009; Bridle *et al.* 2010 for a revised consideration of the strength of such constraints). Some theoretical models predict in particular that gene flow from large central populations into small peripheral ones may swamp local adaptation in marginal areas, preventing range spread beyond some critical environmental limit (Kirkpatrick & Barton 1997; review in Bridle & Vines 2007). Gene flow then causes *phenotypic clines* (glossary) for adaptive traits to deviate from their optima (Box 2). Interestingly, the interaction of strong gene flow with selection on multiple traits could result in some *phenotypic clines* being flatter, and some steeper, than optimal because of genetic and selective interactions among traits (Guillaume 2011, Duputié *et al.* in revision). Furthermore, the constraining effects of migration on divergence are predicted to be more severe when divergence involves many loci of small effects rather than few major genes with large effects on the phenotype (Yeaman & Guillaume 2009). In forest trees, the former situation seems to be the most common (Neale & Kremer 2011).

Common garden experiments in forest trees (Box 1) suggest that genotypes can perform poorly when transferred to climates far from their location of origin. Maladaptation of LD migrants could thus reduce the mean fitness in forest tree populations, generating a *migration load* (glossary). Such *migration load* would be of concern if gene flow is extensive over long distances (see section 2) and if phenotypic mismatch of immigrants is mostly due to long lasting genetic effects (Aitken *et al.* 2008). In a Swedish population of *Pinus sylvestris*, Nilsson (1995) indeed found that offspring sired by naturally dispersing pollen had significantly slower growth and higher freezing resistance than expected if offspring were sired only by pollen produced locally. Natural pollination thus resulted in a phenotypic shift corresponding to that expected if most pollen originated from higher latitudes by 1 to 2 degrees (Nilsson 1995).

There is however little evidence that gene flow has strongly limited adaptation in forest trees in the past. Comparison of genetic differentiation at neutral molecular markers versus *adaptive traits* (glossary) repeatedly suggests that extensive gene flow (presumably mostly through pollen dispersal) has not prevented rapid adaptive divergence of extant populations (Savolainen *et al.* 2007; Kremer *et al.* 2010, for theoretical predictions see Kremer & Le Corre 2011). Populations under different climates may however have diverged while still being far from the locally optimum phenotype. Determining how much observed *phenotypic clines* deviate from what would be optimal under local conditions is however difficult to assess, and remains an open issue in evolutionary biology (Barton, 2001; Butlin *et al.* 2003). When *provenance tests* (glossary) have been replicated over a broad range of climatic conditions, *provenance* (glossary) *reaction norms* (glossary, Box 1) allow comparison of optimal and original climate for each population. Mismatches are not uncommon (Wang *et al.* 2010) and interestingly some studies (Rehfeldt *et al.* 1999; Rehfeldt *et al.* 2002) found more of them at the edge of distributions: e.g. populations of *Pinus contorta* from locations with extreme climates grow better in milder conditions, closer to the core, than in their original location. Such a pattern is consistent with the theoretical expectation that gene flow from the core increases maladaptation in marginal populations (Garcia-Ramos & Kirkpatrick 1997).

Gene flow enhances the response to selection

Natural selection operates by sieving from genetic variation found within populations. Local genetic diversity is therefore the fuel of evolutionary change. Forest tree populations harbour high diversity both for molecular markers and quantitative traits (Hamrick *et al.* 1992), with *heritabilities* (glossary) typically above 0.4 for wood characteristics or phenological traits such as bud set date (Cornelius, 1994). The maintenance of such high levels of quantitative variation for traits closely linked to fitness remains a paradox where strong *stabilizing selection* (glossary) is acting to reduce variation within populations (see Johnson & Barton 2005 for a review). Theoretical models predict that increases in *genetic variance* (glossary) due to gene flow could be substantial (Barton, 2001). Along climatic gradients, the increase in genetic variance within localities due to gene flow is predicted to be proportional to the change in mean *breeding value* (glossary) along the typical dispersal distance. For example, in Sitka Spruce, the *breeding value* for *bud set*

date varies by 3.24 days every 100 km while the within-population phenotypic standard deviation of *bud set* date is typically about 10-25 days (Mimura & Aitken 2007a, Aitken *et al.* 2008); LD gene flow over distances of about 100 km would then lead to *heritability* for *bud set* date greater than 0.4 even for relatively strong *stabilizing selection*. In addition to the mean dispersal distance, the whole shape of the *dispersal kernel* (glossary) is predicted to affect the spatial distribution of genetic variation and diversity of migrants (Travis *et al.* 2010; Fayard *et al.* 2009), especially in the context of range expansion.

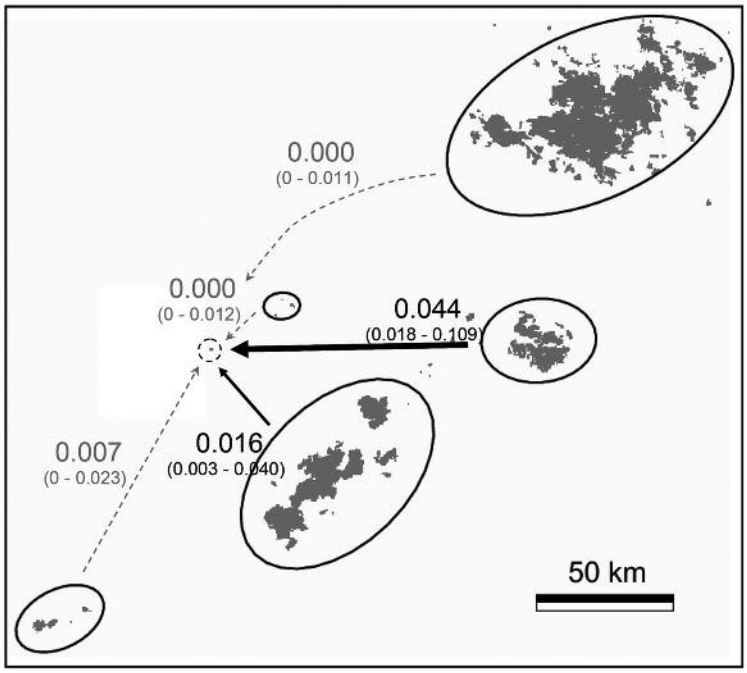
Extensive gene flow in trees is generally thought of as a major explanation for their high within-population diversity (Hamrick *et al.* 1992). Together with the strong selection acting at the juvenile stage experienced in trees, this may allow rapid adaptation to changing climate without large significant reductions in population mean fitness (for empirical examples of rapid genetic changes in forest trees see Jump *et al.* 2006). There is however little direct empirical demonstration of this. Using a mechanistic model of beech stand dynamics, Kramer *et al.* (2008) predicted little effect of pollen dispersal distance on the evolution of within-stand genetic diversity, but their model ignored the potentially large phenotypic divergence of immigrants (e.g. Nilsson 1995). If gene flow between differentiated populations is a persistent source of *genetic variation*, there should be strong correlations between genetic variance within populations and the amount of heterogeneity in the environment at the regional scale. In a study of 142 populations of lodgepole pine, Yeaman & Jarvis (2006) indeed found that the variance for growth among individuals from the same locality measured in *provenance tests* (glossary, Box 1) correlated with regional climatic heterogeneity.

Gene flow facilitates the tracking of environmental change

Assuming an initially locally adapted population, the new selection pressures induced by climate change will cause the optimal phenotype to deviate from the extant mean phenotype, thus creating a *lag load* (glossary, Box 2). Gene flow will reduce the *lag load* in two different ways: (i) by helping track the shift of the optimum phenotype through dispersal of pre-adapted genotypes found somewhere else in the range (see equation 2.2 in Box 2), and (ii) by augmenting the *response to selection* (glossary) through a general increase in genetic variation (see previous paragraph). Quite generally, dispersal should help adaptation in environments that are changing both in space and time (Blanquart & Gandon 2011). Beyond

FIGURE 4

Long distance effective pollen dispersal of *Pinus sylvestris* L.



Estimated effective pollen immigration rates into a *Pinus sylvestris* remnant (encircled with dashed line) from five long-distant populations (encircled with continuous lines) in Central Spain, obtained using maximum-likelihood genetic mixture analysis combined with Monte Carlo assessment of small parameter uncertainty. Continuous (resp. dashed) arrows indicate pollen immigration rates significantly (resp. non-significantly) different from zero. 95% confidence intervals between brackets (modified from Robledo-Arnuncio 2011 with permission from The Publisher).

theoretical predictions, there is little experimental evidence on the impact of gene flow on the *lag load* of tree populations engaged in an evolutionary race with a changing environment. Simulating the evolution of growth cessation date in pine and birch, Kuparinen *et al.* (2010) found that the large lag that developed after one century of climate warming (about 40 days behind the optimal date) was only moderately reduced (by 2 days) by the higher pollen and seed dispersal distances of birch along the climatic gradient.

Gene flow affects inbreeding levels

Gene flow may also affect genetic variation for fitness independently from the issue of adaptation to climate, by affecting genetic resemblance between mates. Early *inbreeding depression* (glossary) is widespread in largely outcrossing species such as trees and inbreeding was found more frequently in isolated marginal populations (Mimura & Aitken 2007b), which may depress their mean fitness and their ability to persist in a changing environment. In many plant species with small isolated populations, progeny formed by crosses between populations show higher fitness than that obtained by within population crosses (e.g. Willi & Fischer 2005). Such a pattern of *heterosis* (glossary) is however not expected to be generally very strong in trees, due to the combination of large population size and extensive gene flow (for theoretical predictions see Glémin *et al.* 2003; Lopez *et al.* 2009). Conversely, small amounts of gene flow between formerly isolated populations can also in theory seriously disrupt fitness due to negative interactions between genes having evolved separately (Edmands & Timmerman 2003), but contrasting results from artificial crosses between distant populations fail to provide solid evidence for this type of *outbreeding depression* (glossary) in trees (Harfouche *et al.* 2000; Goto et *al.* 2011). Because tree populations are seldom isolated from each other, *outbreeding depression* due to negative gene interactions is expected to be rare (Frankham *et al.* 2011).

How will the different effects of gene flow balance each other in the context of climate change?

We here suggest that the positive effects of gene flow may often dominate negative effects for forest trees confronted to climate change. This is in particular the case due to the specific life cycle of forest trees. However, this balance of effects is likely to be modulated by (i) the regional context (e.g. expanding edge or retracting part of the range), and (ii) the dispersal syndrome (relative strength of pollen versus seed dispersal).

Balance between antagonistic effects

Maladaptation in a changing climate is caused by mismatches between optimal and realized mean phenotypes, due to environments that are changing in time and space too fast for the population to adjust to these changes. Such mismatch depresses the mean fitness of populations, generating an *evolutionary load* (glossary, Box 2). Setting aside the effects of gene flow on *genetic variance*, the simple model in Box 2 predicts that there is an optimal level of migration that minimizes such *evolutionary load* under a shifting climate: when dispersal distance is short, the *lag load* decreases fast with increasing migration, which helps the population track the optimal climate (Fig. 3). If gene flow is too high, however, local adaptation is prevented (*migration load*) and maladaptation increases (albeit slowly) with increasing migration (Fig. 3). The optimal dispersal distance is higher if the environment changes more quickly in time in a given location, and if selection varies less sharply in space (Box 2). As an illustration, in the evolution of bud set date in Sitka Spruce (Mimura & Aitken 2007a; Aitken *et al.* 2008, see Box 2), the optimal migration distance is relatively large (immigrants should on average originate from locations with mean temperature differing by more than 3°C to the local site, i.e. more than 1000km). This suggests that, for a range of realistic dispersal distances, the positive tracking effect of dispersal should dominate its negative effects on local adaptation.

Once the effects of gene flow on the evolution of *genetic variance* are taken into account, the constraining effects of migration on adaptation in marginal populations is much weakened: very high gene flow seems instead to facilitate adaptation across a wide array of environmental conditions, but at the cost of a reduced fitness everywhere in the range, which could ultimately compromise species persistence (for theoretical predictions see Barton 2001; Polechova *et al.* 2009), and this effect on population fitness becomes greater when the stochastic effects of finite populations are included (Bridle *et al.* 2010). The very high fecundity, long life span and strong competition at the juvenile stage, which are characteristic of forest trees, could in principle permit very high *genetic load* causing massive mortality at the juvenile stage, without having much impact on adult density. Further exploration of connections between forest trees population dynamics and genetic diversity are needed to conclude when the demographic cost of adaptation compromises persistence.

Overall, models that integrate different antagonistic effects predict that intermediate levels of gene flow (e.g. between one and ten migrants per generation) suffice to replenish *genetic variance* eroded by drift and selection, and alleviate *inbreeding depression* without causing large *migration load*, thus maximize mean fitness in heterogeneous environments (Lopez *et al.* 2009; see also Blanquart & Gandon 2011 for the case of spatio-temporal variation). Empirical evidence that migration enhances fitness in marginal habitats of several plant species supports such predictions (Kawecki 2008). We lack similar direct evidence in forest trees. The admixture of genotypes of diverse geographical origin is increasingly thought of as

key to successful establishment of introduced populations, because of it increases the *genetic variance* necessary for adaptive responses (see Zheng & Ennos 1999 for an example in introduced pine populations). Experimental manipulation of gene flow in forest trees would provide valuable data to better understand its constraining or boosting effects on adaptation to local climate.

Regional variation

Predicted shifts in *bioclimatic envelopes* imply that current populations at the trailing and leading edges of the range will face different adaptive challenges. Southern margin populations will face climatic conditions that currently do not permit species growth. Will these challenged populations have the evolutionary potential (Gomulkiewicz & Houle 2009) to adapt before going extinct? Their persistence will depend on whether the evolutionary or demographic constraints preventing current establishment in warmer or drier climates will be relaxed enough to enable enlargement of species' fundamental niches. Conversely, at the northern margins, new areas will become suitable for growth, but the success of colonization may depend on the genetic make-up of new population founders.

Gene flow can have contrasted consequences for populations at trailing and leading edges of a shifting range (Hampe & Petit 2005): populations at the leading edge or in the central part of the distribution are likely to receive "pre-adapted" genes from more southern populations, and gene flow may facilitate their adaptation (Hu & He 2006). The opposite may be true for populations at the rear end that encounter an entirely novel environment. The flow of pre-adapted genes from central populations is then not possible, which may increase maladaptation and extinction probabilities in populations at the southern margins. However, both demographic and genetic rescue effects of dispersal from larger populations within the species' range may help those marginal populations to persist. The precise balance of the multifarious effects of gene flow remains to be explored in this context.

Balance between the effects of seed and pollen flow

Balance between the negative and positive effects of gene flow may also vary with the relative contribution of seed and pollen dispersal. Both pollen flow and seed flow contribute substantially to genetic diversity. On the one hand, pollen often disperses farther than seeds (see section 2 and Table 1) and in greater quantities. On the other hand, a single pollen grain carries half the number of alleles compared to a single seed, and only seeds can establish a new population in a remote habitat. Due to long generation times in trees, migrant seeds accumulate in a new population over years before the new generation reproduces, promoting high levels of diversity in recently founded populations (Austerlitz *et al.* 2000). Long distance gene flow mediated by pollen in marginal habitats is therefore conditional on the successful establishment of shorter distance migrating seeds. The movement of alleles by pollen necessarily involves combining with existing genetic variation, which explains why seed and pollen dispersal may have different consequences for population divergence, maintenance of within-population diversity, and mean fitness (Hu & Li 2003; Lopez *et al.* 2008). When selection varies sharply in space, pollen dispersal could, in particular, generate higher *migration loads* than equivalent seed dispersal (Lopez *et al.* 2008). This happens because selection is less efficient at removing badly adapted immigrant alleles when their deleterious effects are partly masked in hybrids (Lopez *et al.* 2008). Selection at the gametophytic stage may further affect these differences (Hu & Li 2003; Hu & He 2006).

Most models of adaptation and migration in a heterogeneous environment (e.g. Pease *et al.* 1989; Kirkpatrick & Barton 1997; Polechova *et al.* 2009, see Box 2; but see Butlin *et al.* 2003) consider a single dispersal parameter. With pollen and seed dispersal, demographic migration is partially uncoupled from gene flow. Hu & He (2006) predicted that pollen dispersal could slow down or accelerate range expansion in some homogeneous environments by interfering with the spread of beneficial or deleterious mutations. At retracting range margins, seed and pollen dispersal may play very different roles on adaptation: seed dispersal enhances the probability of adaptation in a sink habitat, while pollen dispersal generally compromises it (Aguilée *et al.* unpublished). Conversely, both pollen flow and seed flow could have positive effects at expanding range margins. In the presence of pollen limitation, long-distance pollen flow could moreover prevent extinction in marginal populations (Butlin *et al.* 2003).

Future research directions

Develop new methods to trace pollen and seeds

Experimental dispersal studies monitoring LD pollen and seed dispersal have often been limited in spatial scale due to (i) overlapping of the *pollen/seed shadows* (glossary) masking LDD, (ii) dilution effect (LDD is rare and requires high power to observe, let alone measure), and (iii) large numbers of putative sources (characterizing their positions and genotypes is time- and cost-intensive). Using highly polymorphic genetic markers like microsatellites greatly overcomes the first point, and the advent of next-generation sequencing will improve power and resolution, however it is still necessary to conceive new experimental designs dealing with points (ii)-(iii). We propose potential strategies here, mostly relying on a stronger interaction with mechanistic approaches.

Making use of meteorological data

For wind-mediated gene flow, available weather data could help determine the potential range of pollen and

seed dispersal within particular landscapes, regions, or continents. Such an analysis requires regional meteorological datasets, phenological observations over a wide region and sufficient understanding of the meteorological factors driving pollen and seed emission and spread. Products of regional and global weather reanalysis, such as the North American Regional Reanalysis (NARR) dataset and the European Centre for Medium Range Weather Forecast (ECMWF) data, offer useful observational and model-based information on wind, temperature, humidity, radiation and other meteorological data (Schueler *et al.* 2005). On-line interfaces for weather and radiation simulation tools can also be used to evaluate conditions across large dispersal ranges (Bohrerova *et al.* 2009). For pollen, phenological data are available from pollen monitoring networks (e.g. the European Aerobiology Network, EAN) or from phenological observations (e.g. the European Phenological Network). Model-driven weather reconstructions (Solomon *et al.* 2007) can provide estimates of dispersal potentials in past and future climates (Kuparinen *et al.* 2009; Nathan *et al.* 2011). Improved characterizations of wind dispersal mechanisms accounting for interactions between pollen/seeds and turbulent winds in relation to weather conditions can be combined to determine annual and multiannual wind-driven pollen and seed dispersal patterns throughout large geographic regions (Muñoz *et al.* 2004; Thompson & Katul 2008; Nathan *et al.* 2011).

To experimentally trace pollen or seed movement at the continental scale, a joint use of weather data, weather forecasting models and field observation of pollen/seed pools seems most promising. Large-scale spatial characterization of presence/absence of a species, phenology, and airflows were already used to identify temporal windows ideal for LDD and relate them to the actual presence of pollen grains in physical captors (Siljamo *et al.* 2008). A next step would be to measure the diversity of origins in the effective pollen pools through the genetic and/or phenotypic diversity of the seed produced (Nilsson 1995).

Taking advantage of adequate landscape configurations

Genetic assignment (glossary) methods linking pollen, seeds or seedlings to candidate parental populations could be used to evaluate the effective rate and range of contemporary gene flow among discrete populations (Manel *et al.* 2005). Focus could be placed initially on isolated populations or trees, particularly informative about LDD because they are less subject to dilution effects. A recent study using *genetic assignment* in such demographic setting has revealed effective pollen gene flow over 100-km distances in a wind-pollinated species (Fig. 4). For species with extremely low densities, even *parentage analysis* (glossary) may prove efficient in detecting LD gene flow (Ahmed *et al.* 2009). Female plants, male-sterile or self-incompatible isolated individuals might prove useful traps for investigating the composition of LD effective pollen clouds, and could be distributed at specific positions during the pollination period, e.g. using potted plants, flowering branches kept alive or flowering grafts. High precision aerial photographs and satellite images could be used to retrieve all potential sources at the regional scale and avoid biases due to ghost populations. An alternative solution, not requiring trap plants but not assessing *effective dispersal* directly, is to characterize the genetic content of the pollen pool by genotyping single pollen grains (Matsuki *et al.* 2007), sampled in volumetric traps from existing *aerobiology* (glossary) networks or placed at specific sites in a landscape. Note however that the atypical demographic conditions of isolated trees that facilitate LDD assessment may result in observed LDD patterns difficult to generalize (e.g. dilution effects in large populations), for which modelling approaches may be necessary.

Combining mechanistic and genetic models

Mechanistic and genetic tools for assessing dispersal have been developed and applied virtually independently, although they have complementary features and high potential for synergy, particularly for the analysis of LDD. For example, the ability of mechanistic approaches to assess dispersal across multiple scales complements the problematic extrapolation of genetic methods beyond the small scale in which individuals were sampled. Genetic methods, in turn, can provide data to validate mechanistically-derived kernels, and to add the required (and often hard to measure) component of post-dispersal establishment effects needed to assess *effective dispersal*. Mechanistically derived propagule transport probability functions over different distances could also be incorporated into the usual probabilistic (maximum-likelihood or Bayesian) migration rate estimation procedures based solely on genetic likelihoods, allowing jointly estimation of migration rates and mechanistic parameters that determine dispersal over long distances.

Developing the connectivity network

Current research on spatial patterns of dispersal and gene flow is dominated by the *dispersal kernel* concept, which bears significant disadvantages when applied to broad scales. Studies using *dispersal kernels* generally require sampling intensities that become unfeasible over long distances. Moreover, they often assume isotropy (i.e., the same *dispersal kernel* for all directions); although this assumption is unrealistic for many systems in which the dispersal vector moves in a directional manner, such as many seasonal winds, downward flow of rivers, and oriented movement of animals. Similarly, *genetic assignment* methods for migration rate estimation typically incorporate neither directional nor other kinds of spatial information. *Lagrangian dispersal* (glossary) simulations can account for dispersal anisotropy by incorporating turbulence patterns (Bohrer *et al.* 2008), and hourly, daily or seasonal variation in wind direction (Wright *et al.* 2008); this computationally-intensive

approach, however, is practically limited to relatively short-term small-scale applications. An alternative approach, *connectivity maps* (glossary), depicts dispersal probabilities between sites based on large-scale datasets and/or models available for the primary dispersal vector, for example, to assess wind connectivity of plants among islands in the southern oceans (Muñoz *et al.* 2004). In principle, this method could be adjusted to many plant species in a variety of spatial scales, if the patterns of movement of the dispersal vector can be estimated. Because some key vectors such as wind, inland water systems, ocean currents and migrating birds disperse many plant species, efforts to develop vector *connectivity maps* could advance the study of gene flow via pollen and seeds for a large number of species.

Implement experimental approaches to assess evolutionary changes in trees

Despite obvious biological constraints in trees, we recommend setting up experiments that would allow assessing evolutionary changes over a few generations. Such experiments would not only provide estimates of evolutionary rates, they would also offer the opportunity to test evolutionary hypotheses regarding responses to climate change. Existing *provenance tests* constitute in this respect a precious source of data, allowing the quantification of between and within sites genetic diversity for climate adaptation and the putative demographic impact of maladaptation (Box 1). Further exploitation of such data should be encouraged. Additional options can be foreseen:

Testing for the effect of gene flow on the changes of population means and genetic and phenotypic variance over one generation.
A straightforward design consists in conducting full sib control "hybrid" crosses between distant and close populations in comparison to "pure" within population crosses. Offspring should then be raised under controlled conditions mimicking different climatic scenarios. While more difficult to implement, because of potentially small sample sizes and unaccounted microenvironmental variation, an alternative "in situ" experiment consists in comparing "natural migrants" that have been identified by *parentage analysis* or *genetic assignment* methods to "local residents".

Testing the effects of the strength of selection over successive generations.
We suggest installing short generation tree populations (birch or willow) within open top chambers, and let the population reproduce under such conditions. Strength of selection can be set by manipulating conditions within the open top chambers. Foreign pollen can be supplemented at each generation to mimic gene flow.

Measuring the strength of selection at various filtering stages over the life cycle.
While filtering stages (i.e. stages with strong competition and selective mortality) are well known in trees especially at the young stage (from seeds to juvenile seedlings), the changes induced by selection at each stage have only rarely been assessed. We suggest to monitor population means, and genetic variances of relevant *adaptive traits* as well as allelic frequencies at genes of adaptive significance after each filtering stage.

Analysing adaptation in transferred populations.
Artificial transfers of populations have been done in the past in forest trees and some of them are well documented (Fallour-Rubio *et al.* 2009). They can provide alternative ways of tracking evolutionary changes at contemporary time scales. In some cases these transfers actually mimicked climate changes, as populations were moved from cooler to milder climates. Well known examples are transfers of North American tree species to Europe, whose introduced populations have differentiated in so-called land races (Northern red oak, Daubree & Kremer 1993), or large scale transfers of native trees within Europe. Transferred populations have usually been deployed over larger areas than *provenance tests* and the transferred material has been tested in a real forestry context, rather than in experimental plantations.

Develop integrative theoretical approaches

Extend evolutionary models of adaptation to climate change to the case of trees.
Analytical models –as the example shown in Box 2- provide conceptual insights into how gene flow, adaptation and biotic interactions shape species ranges in stable or changing environments (Pease *et al.* 1989; Kirkpatrick & Barton 1997; Barton 2001; Polechova *et al.* 2009; Price & Kirkpatrick 2009). Available models, however, rarely incorporate salient features of tree life cycles, such as distinct dispersal modes, overlapping generations, or *fat-tailed dispersal kernels* (glossary), which may profoundly affect their evolutionary responses to climate change. How the pace of adaptation in a changing environment depends on variation in fitness expressed before vs. after sexual maturity and on the correlation between juvenile vs. adult traits is for instance an important area for future research.
Exploring the effects of LD gene flow on adaptation also requires modelling dispersal as a more complex process than the simple homogeneous diffusion considered in the models summarized in Box 2. What is the evolutionary impact of rare long distance dispersal events well beyond the average dispersal distance? While the effect of *fat-tailed dispersal kernels* on rates of expansion (Thompson & Katul 2008) and neutral diversity (Travis *et al.* 2010; Fayard *et al.* 2009) have been explored, we lack similar theoretical investigation of their effects on adaptive diversity in the context of climate change. Answering this question would also help identifying critical features of pollen and seed dispersal distributions on which empirical estimates should focus.

The idea that climate change is equivalent to simple spatial shift of local climatic conditions is also a gross simplification. Rather, climatic change may results in new combinations of precipitation patterns, temperature, photoperiod and biotic conditions that occur nowhere within the current range, imposing entirely new selection pressures, and favouring the assembly of novel genotypes (Williams & Jackson 2007). Adaptation to climate change may thus require the production of new phenotypic combinations. Reaching such combinations means that natural selection is acting on multiple traits simultaneously. Evolutionary models have often been limited to single traits. Multivariate adaptive responses depend on the amount of *genetic correlation* (glossary) among traits, which may limit or accelerate adaptation to climate change (e.g., Etterson & Shaw 2001). Modelling those responses as a univariate rather than multivariate process, as done so far, might thus fail to provide an accurate picture of species' adaptive capacities. Although we have begun to incorporate multivariate evolution into models of migration-selection balance (see, Guillaume & Whitlock 2007; Guillaume 2011; Duputié *et al.* in review), empirical data are direly missing on patterns of *genetic correlations* among key ecological traits in trees and on the spatial and temporal variation of their joint selection pressures.

Efforts should be made to fill these gaps and help calibrate models with real data, to ultimately be able to merge evolutionary approaches with niche- and process-based ecological forecasting of climate induced range shifts.

Use integrative simulation platforms.
Trait-based, mechanistic models have recently been developed enabling predictions of species ranges under current and future non analogous climates (e.g., Morin *et al.* 2008). For instance the Phenofit model (Chuine & Beaubien 2001) predicts tree distributions based on existing phenological responses to local climate, drought and frost tolerance. Microevolutionary phenomena described above have only started to be incorporated in such ecological forecast models (Kearney *et al.* 2009; Kuparinen *et al.* 2010). There is therefore an urgent need to incorporate genetic and ecological concepts into integrated models to accurately predict the impact of environmental changes on species persistence over the next century and at the continental scale. Efforts should be dedicated to foster development of integrated computer simulation platforms with this aim. Individual-based, population and quantitative genetics simulation packages already exist (e.g. Nemo, Guillaume & Rougemont 2006; Metapop, e.g. Le Corre & Kremer 2003; Kremer & Le Corre, 2011) that could be extended to include the ecological and spatially explicit layers needed.

A key aspect of the modelling approach advocated here is the overlay of predictions from different processes; ecological niche and *bioclimatic envelope* modelling, variation of gene flow over geographical ranges, and evolutionary adaptation of local populations. The basal layer, the climatic layer, defines how changes in climatic conditions over the species' geographical range modify the localization of suitable habitats (Thuiller 2003). The second layer describes spatial variation of pollen and seed dispersal and should integrate information from the climatic layer to model the changes of seed and pollen movements caused by climate change through modification of the *dispersal kernels* (Kuparinen *et al.* 2009; Nathan *et al.* 2011), pollen viability (Bohrerova *et al.* 2009), or the timing of pollination and female receptiveness. The third layer integrates information from the two previous ones to predict how local populations adapt to their shifting conditions (e.g., Kuparinen *et al.* 2010). Information from the climatic layer will set the strength of selection acting on different *adaptive traits* by indicating how far from its local optimum a population might be. Information on gene flow from the second layer will indicate hybridization rates and fitness effects, depending on the geographical origin of the migrants (Savolainen *et al.* 2007; Lopez *et al.* 2008; Yeaman & Guillaume 2009). It will also indicate the potential for colonization of new habitats. Finally, the outcome of local adaptation can be interpreted in terms of growth and persistence of local populations and how this feeds back into predictions of the intensity of gene flow over larger geographical scales.

Conclusion

While much emphasis has been placed on the ability of tree populations to migrate fast enough in response to climate change, we have here examined the potential consequences of long distance gene flow on their adaptive response to climate change.
Many tree species have evolved dispersal syndromes enabling the effective flow of genetic information across distant populations inhabiting contrasting environments. We have argued how such exchanges, although potentially maladaptive in some evolutionary and demographic scenarios, may in the case of forest trees favour adaptation to changing climatic conditions, compensating for their long generation time.
Our understanding of the interaction between gene flow and local adaptation under realistic ecological, demographic and dispersal assumptions is however limited, and we have suggested potential theoretical and experimental avenues of research for the integration of dispersal biology, ecology and evolutionary quantitative genetics in a better predictive inferential framework.

Acknowledgements

This article is a synthesis of a workshop organized within the Coordination and Support Action (CSA) FORESTTRAC "Forest ecosystem genomics research: supporting transatlantic cooperation" which took place in

Barsac (France) from September 27th to October 1st 2010. FORESTTRAC was supported by the European Union under the FP7 Workprogramme Environment (FP7-244096). This is publication ISEM 2012-013 of the Institut des Sciences de l'Evolution Montpellier. This work was supported by the Agence Nationale de la Recherche, programme « 6th extinction », through the project EVORANGE (ANR-09-PEXT-011). FG was supported by a grant from the Swiss NSF (no. PZ00P3_121697). RN was supported by grants from the Israel Science Foundation (ISF-474 / 02, ISF-150 / 07 and ISF-FIRST-1316 / 05) and the Adelina and Massimo Della Pergola Chair of Life Sciences. JJRA was supported by a grant from the Spanish Ministry of Science and Innovation (CGL2009-09428).

REFERENCES

Ahmed, S., Compton, S.G., Butlin, R.K. & Gilmartin, P.M. (2009). Wind-borne insects mediate directional pollen transfer between desert fig trees 160 kilometers apart. *Proc. Natl. Acad. Sci.* U. S. A., 106, 20342-20347.

Aitken, S.N., Yeaman, S., Holliday, J.A., Wang, T. & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climatic changes outcomes for tree populations. *Evolutionary Applications*, 1, 95-111.

Anderson, J. T., Nuttle, T., Saldaña Rojas, J.S., Pendergast, T.H. & Flecker, A.S. (2011). Extremely long-distance seed dispersal by an overfished Amazonian frugivore. *Proceedings of the Royal Society B: Biological Sciences*:doi: 10.1098/rspb.2011.0155.

Austerlitz, F., Mariette, S., Machon, N., Gouyon, Ph., Godelle, B. (2000). Effects of colonization processes on genetic diversity: Differences between annual plants and tree species. *Genetics*, 154, 1309-1321.

Bacles, C.F.E., Lowe A.J. & Ennos, R.A. (2006). Effective seed dispersal across a fragmented landscape. *Science*, 311, 628-628.

Barton, N.H. (2001). Adaptation at the edge of a species' range. In: *Integrating Ecology and Evolution in a Spatial Context* (ed. J. Silvertown and J. Antonovics). Blackwell Sciences, Oxford, pp. 365-392.

Blanquart, F., Gandon, S. (2011). Evolution of migration in a periodically changing environment. *Am. Nat.*, 177, 188-201.

Bohrer, G., Nathan, R. & Volis, S. (2005). Effects of long-distance dispersal for metapopulation survival and genetic structure at ecological time and spatial scales. *J. Ecol.*, 93, 1029-1040.

Bohrer, G., Nathan, R., Katul, G.G., Walko, R.L. & Avissar, R. (2008). Effects of canopy heterogeneity, seed abscission, and inertia on wind-driven dispersal kernels of tree seeds. *J. Ecol.*, 96, 569-580.

Bohrerova, Z., Bohrer, G., Cho, K.D., Bolch, M.A. & Linden, K.G. (2009). Determining the viability response of *Pinus taeda* pollen under abiotic stresses typical to atmospheric conditions during long distance dispersal. *Ecol. Appl.*, 19, 656-667.

Bridle, J.R. & Vines, T.H. (2007). Limits to evolution at range margins: When and why does adaptation fail? *Trends Ecol.* Evol., 22,140-47.

Bridle, J.R., Polechova, J. & Vines, T.H. (2009). Limits to adaptation and patterns of diversity. In: *Speciation and patterns of diversity.* (eds Butlin, R.K., Bridle, J.R. & Schluter, D.). Cambridge University Press, pp77-102.

Bridle, J.R., Polechova, J., Kawata, M. & Butlin, R.K. (2010). Why is adaptation prevented at ecological margins? New insights from individual-based simulations. *Ecol. Lett.*, 13, 484-494.

Bürger, R. & Krall, C. (2004). Quantitative-genetic models and changing environments. In: *Evolutionary Conservation Biology* (eds. Ferriere, R., Dieckmann, U. & Couvet, D.). Cambridge University Press, pp. 171-187.

Buschbom, J., Yanbaev, Y. & Degen, B. (2011). Efficient long-distance gene flow into an isolated relict oak stand. *J. Hered.*, doi: 10.1093/jhered/esr023

Butlin, R.K., Bridle, J.R. & Kawata, M. (2003). Genetics and the boundaries of species´ distributions. In: *Macroecology: Concepts and Consequences* (ed. Blackburn, T. & Gaston, K.J.). Blackwell Science, pp. 274-295.

Campbell, I.D., McDonald, K., Flannigan, M.D. & Kringayark, J. (1999). Long-distance transport of pollen into the Arctic. *Nature*, 399, 29-30.

Campos-Arceiz, A., Larrinaga, A. R., Weerasinghe, U. R., Takatsuki, S., Pastorini, J., Leimgruber, P., *et al.* (2008). Behavior rather than diet mediates seasonal differences in seed dispersal by Asian elephants. *Ecology*, 89, 2684-2691.

Cornelius, J. (1994). Heritabilities and additive genetic coefficients of variation in forest trees. *Can. J. For. Res.*, 24, 372-379.

Cwynar, L.C. & Macdonald, G.M. (1987). Geographic variation of lodgpepole pine in relation to population history. *Amer. Nat.*, 129, 463-469

Darling E., Samis, K. E. & Eckert, C. G. (2008). Increased seed dispersal potential towards geographic range limits in a Pacific coast dune plant. *New Phytol.*, 178, 424-435.

Daubree J.B. & Kremer, A. (1993). Genetic and phenological differentiation between introduced and natural populations of *Quercus rubra L. Ann. For. Sci.* (Suppl 1), 271s-280s.

Edmands, S. & Timmerman, C. C. (2003). Modeling factors affecting the severity of outbreeding depression. *Conserv. Biol.*, 17, 883-892.

Ennos, R.A. (1994). Estimating the relative rates of pollen and seed migration among plant populations. *Heredity*, 72, 250-259.

Etterson, J. R. & Shaw, R. G. (2001). Constraint to adaptive evolution in response to global warming. *Science*, 294, 151-154.

Fallour-Rubio, D., Guibal, F., Klein, E.K., Bariteau, M. & Lefèvre, F. (2009). Rapidchanges in plasticity across generations within an expanding cedar forest. *J. Evol. Biol.*, 22, 553-563.

Fayard, J., Klein, E. K. & Lefèvre, F. (2009). Long distance dispersal and the fate of a gene from the colonization front. *J. Evol. Biol.*, 22, 2171-2182.

Frankham, R., Ballou, J.D., Eldridge, M.D.B., Lacy, R.C., Ralls, K., Dudash, M.R. & Fenster, C.B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology*, 25, 465-475.

Garcia-Ramos, G. & Kirkpatrick, M. (1997). Genetic models of adaptation and gene flow in peripheral populations. *Evolution*, 51, 21-28.

Glémin, S., Ronfort, J. & Bataillon, T. (2003). Patterns of inbreeding depression and architecture of the load in subdivided populations. *Genetics*, 165, 2193-2212.

Gomulkiewicz, R. & Houle, D. (2009). Demographic and genetic constraints on evolution. *Am. Nat.*, 174, E218-E229.

Goto, S., Lijima, H., Ogawa, H & Ohya, K. (2011). Outbreeding depression caused by intraspecific hybridization between local and non local genotypes in *Abies sachalinensis. Restor. Ecol.*, 19, 243-250.

Guillaume. F. & Rougemont, J. (2006). Nemo: an evolutionary and population genetics programming framework. *Bioinformatics*, 22, 2556-2557.

Guillaume, F. & Whitlock, M.C. (2007). Effects of migration on the genetic covariance matrix. *Evolution*, 61, 2398-2409.

Guillaume, F. (2011). Migration-induced phenotypic divergence: the migration-selection balance on correlated traits. *Evolution (In Press)* DOI: 10.1111/j.1558-5646.2011.01248.x

Hampe, A. & Petit, R.J. (2005). Conserving biodiversity under climate change: the rear edge matters. *Ecol. Lett.*, 8, 461-467.

Hamrick, J.L., Godt, M.J.W. & Sherman-Broyles, S.L. (1992). Factors influencing levels of genetic diversity in woody plant species. *New Forest.*, 6, 95-124.

Harfouche, A., Bahrman, N., Baradat, P., Guyon, J.P., Petit, R.J. & Kremer, A. (2000). Provenance hybridization in a diallel mating scheme of maritime pine *(Pinus pinaster* Ait*.)*. II. Heterosis. *Can. J. . For. Res.*, 30, 10-16.

Holbrook, K.M. & Smith, T.B. (2000). Seed dispersal and movement patterns in two species of *Ceratogymna hornbills* in a West African tropical lowland forest. *Oecologia*, 125, 249-257.

Hu, X.S. & Li, B. (2003). On migration load of seeds and pollen grains in a local population. *Heredity*, 90, 162-168.

Hu, X.S. & He, F.L. (2006). Seed and pollen flow in expanding a species' range. *J. Theor. Biol.*, 240, 662-672.

Iverson, L.R., Prasad, A.M., Matthews, S.N. & Peters, M. (2008). Estimating potential habitat for 134 eastern US tree species under six climate scenarios. *For. Ecol. Man.*, 254, 390-406.

Johnson, T. & Barton, N. (2005). Theoretical models of selection and mutation on quantitative traits. *Philos. Trans. R. Soc. B. Biol. Sci.*, 360, 1411-1425.

Jordano, P., Garcia, C., Godoy, J.A. & Garcia-Castano, J.L. (2007). Differential contribution of frugivores to complex seed dispersal patterns. *Proc..Natl. Acad. Sci. U S A.*, 104, 3278-3282.

Kamm, U., Rotach, P., Gugerli, F., Siroky, M., Edwards, P. & Holderegger, R. (2009). Frequent long-distance gene flow in a rare temperate forest tree (*Sorbus domestica*) at the landscape level. *Heredity*, 103, 476-482.

Kaufman, S.R., Smouse, P.E. & Alvarez-Buylla, E.R. (1998). Pollen-mediated gene flow and differential male reproductive success in a tropical pioneer tree, *Cecropia obtusifolia* Bertol. (Moraceae): a paternity analysis. *Heredity*, 81, 164-173.

Kawecki, T.J. (2008). Adaptation to marginal habitats. *Annu. Rev. Ecol. Evol. Syst.,* 39, 321-42.

Kearney, M., Porter, W.P., Williams, C., Ritchie, S. & Hoffmann, A.A. (2009). Integrating biophysical models and evolutionary theory to predict climatic impacts on species' ranges: the dengue mosquito *Aedes aegypti* in Australia. *Funct. Ecol.*, 23, 528-538.

Kelly, C.K., Chase, M.W., de Bruijn, A., Fay, M.F. & Woodward, F.I. (2003). Temperature-based population segregation in birch. *Ecol. Lett.*, 6, 87-89.

Kirkpatrick, M. & Barton N.H. (1997). Evolution of a species' range. *Am. Nat.*, 150, 1-23.

Kramer, K., Buiteveld, J., Forstreuter, M., Geburek, T., Leonardi, S., Menozzi, P., *et al.* (2008). Bridging the gap between ecophysiological and genetic knowledge to assess the adaptive potential of European beech. *Ecol. Modell.*, 216, 333-353.

Kremer, A., Le Corre, V., Petit, R.J. & Ducousso, A. (2010). Historical and contemporary dynamics of adaptive differentiation in European oaks. In: *Molecular approaches in natural resource Conservation* (ed. DeWoody, A., Bickham, J., Michler, C., Nichols K., Rhodes G. & Woeste K.). Cambridge University Press, pp. 101-117.

Kremer, A. & Le Corre, V. (2011). Decoupling of differentiation between traits and their underlying genes in response to divergent selection. Heredity doi:10.1038/hdy.2011.81

Kuparinen, A., Katul, G.G., Nathan, R. & Schurr, F.M. (2009). Increases in air temperature can promote wind-driven dispersal and spread of plants. *Proc. R. Soc. B Biol. Sci.*, 276, 3081-3087.

Kuparinen, A., Savolainen, O. & Schurr, F.M. (2010). Increased mortality can promote evolutionary adaptation of forest trees to climate change. *For. Ecol. Man.*, 259, 1003-1008.

Lande, R. & Shannon, S. (1996). The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, 50, 434-437.

Le Corre, V. & Kremer, A. (2003). Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics*, 164, 1205-1219.

Loarie S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B. & Ackerly, D.D. (2009). The velocity of climate change. *Nature*, 462, 1052-1056.

Lopez, S., Rousset, F., Shaw, F. H., Shaw, R. G. & Ronce, O. (2008). Migration load in plants: role of pollen and seed dispersal in heterogeneous landscapes. *J. Evol. Biol.*, 21 , 294-309.

Lopez, S., Rousset, F., Shaw, F. H., Shaw, R. G. & Ronce, O. (2009). Joint effects of inbreeding and local adaptation on the evolution of genetic load after fragmentation. *Conserv. Biol.*, 23,1618-1627.

Manel, S., Gaggiotti, O.E. & Waples, R.S. (2005). Assignment methods: matching biological questions with appropriate techniques. *Trends Ecol. Evol.*, 20, 136-142.

Matsuki, Y., Isagi, Y. & Suyama, Y. (2007). The determination of multiple microsatellite genotypes and DNA sequences from a single pollen grain. *Molecular Ecology Notes*, 7, 194-198.

Mc Kenney, D.W., Pedlar, J.H., Lawrence, K., Campbell, K. & Hutchinson, M.F. (2007). Potential impacts of climate change on the distribution of North American trees. *Bioscience*, 57, 939-948.

Mimura, M, & Aitken, S.N. (2007a). Adaptive gradients and isolation-by-distance with postglacial migration in *Picea sitchensis. Heredity*, 99, 224-232.

Mimura, M, & Aitken, S.N. (2007b). Increased selfing and decreased effective pollen donor number in peripheral relative to central populations in *Picea sitchensis* (Pinaceae). Am. J. Bot., 94, 991-998.

Morgenstern, E.K. (1996). Geographic variation in forest trees. University of British Columbia Press, Vancouver, Canada. 209 p.

Morin, X., Viner, D. & Chuine, I. (2008). Tree species range shifts at a continental scale: new predictive insights from a process-based model. *J. Ecol.*, 96, 784-794.

Muñoz, J., Felicísimo, A.M., Cabezas, F., Burgaz, A.R. & Martínez, I. (2004). Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science*, 304, 1144-1147.

Nason, J.D., Herre, E.A. & Hamrick, J.L. (1998). The breeding structure of a tropical keystone plant species. *Nature*, 391, 685-87.

Nathan, R., Schurr, F.M., Spiegel, O., Steinitz, O., Trakhtenbrot, A. & Tsoar, A. (2008). Mechanisms of long-distance seed dispersal. *Trends Ecol. Evol.*, 23, 638-647.

Nathan, R., Horvitz, N., He, Y., Kuparinen, A., Schurr, F.M. & Katul, G.G. (2011). Spread of North American wind-dispersed trees in future environments. *Ecol. Lett.*, 14, 211-219.

Neale, D.B. & Kremer, A. (2011). Forest tree genomics: growing resources and applications. *Nature Reviews Genetics*, 12, 111-122.

Nilsson, J.E. (1995). Genetic variation in the natural pollen cloud of *Pinus sylvestris* - A study based on progeny testing. *Scand. J. For. Res.*, 10, 140-148.

Pease, C.P., Lande, R. & Bull, J.J. (1989). A model of population growth, dispersal and evolution in a changing environment. *Ecology*, 70, 1657-64.

Petit, R.J., Hu, F.S. & Dick, C.K. (2008). Forests of the past: a window to future changes. *Science*, 320, 1450-1452.

Polechova, J., Barton, N. & Marion, G. (2009). Species' range: adaptation in space and time. *Am. Nat.*, 174, E186-204.

Price, T.D. & Kirkpatrick, M. (2009). Evolutionary stable range limits set by interspecific competition. Proceedings of the Royal Society B-Biological sciences, 276, 1429-1434.

Rehfeldt, G. E., Ying, C.C., Spittlehouse, D.L. & Hamilton, D.A. (1999). Genetic responses to climate in *Pinus contorta*: niche breadth, climate change, and reforestation. *Ecol. Monogr.*, 69, 375-407.

Rehfeldt, G.E, Tchebakova, N.M., Parfenova, Y.I., Wykoff, W.R., Kuzmina, N.A. & Milyutin, L.I. (2002). Intraspecific response to climate in *Pinus sylvestris. Glob. Change Biol.*, 8, 912-929.

Riba, M., Mayol, M., Giles, B.E., Ronce, O., Imbert, E., Van der Velde, M., Chauvet, S., Ericson, L. , Bijlsma, R., Vosman, B., Smulders, M.J.M. & Olivieri, I. (2009). Darwin's wind hypothesis: does it work for plant dispersal in fragmented habitats. *New Phytologist*, 183, 667-677.

Robledo-Arnuncio, J.J. (2011). Wind pollination over mesoscale distances: an investigation with Scots pine. *New Phytol.*, 190, 222-233.

Savolainen, O., Pyhäjärvi, T., & Knürr, T. (2007). Gene flow and local adaptation in trees. *Ann. Rev. Ecol. Evol. Syst.*, 38, 595-619.

Schueler, S., Schlünzen, K.H. & Scholz, F. (2005). Viability and sensitivity of oak pollen and its implications for pollen-mediated gene flow. *Trends Ecol. Evol.*, 19, 154-161.

Siljamo, P., Sofiev, M., Severova, E., Ranta, H., Kukkonen, J., Polevova, S., *et al.* (2008). Sources, impact and exchange of early-spring birch pollen in the Moscow region and Finland. *Aerobiologia*, 24, 211-230.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M, Averyt, K.B, *et al.* (2007). *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Thompson, S. & Katul, G.G. (2008). Plant propagation fronts and wind dispersal: An analytical model to upscale from seconds to decades using superstatistics *Am. Nat.*, 171, 468-479.

Thuiller, W. (2003). BIOMOD - optimizing predictions of species distributions and projecting potential future shifts under global change. *Glob. Change Biol.*, 9, 1353-1362.

Travis, J.M.J., Munkemuller, T. & Burton, O.J. (2010). Mutation surfing and the evolution of dispersal during range expansions. *J. Evol. Biol.*, 12, 2656-2667.

Tsoar, A., Shohami, D. & Nathan, R. (2011). A movement ecology approach to study seed dispersal and plant invasion: an overview and application of seed dispersal by fruit bats. Pages 103-119. *In:* Fifty Years of Invasion Ecology: The Legacy of Charles Elton (ed Richardson, D.M.), Wiley-Blackwell, Oxford, UK. doi: 10.1002/9781444329988.ch9

Varis, S., Pakkanen, A., Galofré, A. & Pulkkinen, P. (2009). The extent of south-north pollen transfer in Finnish Scots pine. *Silva Fenn.*, 43, 717-726.

Wang, T.L., O'Neill, G.A. & Aitken, S.N. (2010). Integrating environmental and genetic effects to predict responses of tree populations to climate. *Ecol. Appl.*, 20, 153-163.

Westcott, D.A., Setter, M., Bradford, M.G., McKeown, A. & Setter, S. (2008). Cassowary dispersal of the invasive pond apple in a tropical rainforest: the contribution of subordinate dispersal modes in invasion. *Diversity and Distributions*, 14, 432-439.

White, G.M., Boshier, D.H. & Powell, W. (2002). Increased pollen flow counteracts fragmentation in a tropical dry forest: An example from *Swietenia humilis* Zuccarini. *Proc. Natl. Acad. Sci. USA*, 99, 2038-42.

Willi, Y., & Fischer, M. (2005). Genetic rescue in interconnected populations of small and large size of the self-incompatible *Ranunculus reptans. Heredity*, 95, 437-443.

Williams, C.G. (2010). Long-distance pine pollen still germinates after meso-scale dispersal. *Am. J. Bot.*, 97, 846-855.

Williams J.W & Jackson, S.T. (2007). Novel climates, no-analog communities, and ecological surprises. *Front. Ecol. Environ.*, 5,475–82

Wright, S.J., Trakhtenbrot, A., Bohrer, G., Detto, M., Katul, G.G., Horvitz, N., *et al.* (2008). Understanding strategies for seed dispersal by wind under contrasting atmospheric conditions. *Proc. Nat. Acad. Sci. USA,* 105, 19084-19089.

Yeaman, S. & Jarvis, A. (2006). Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proc. R. Soc. B Biol. Sci.*, 273, 1587-1593.

Yeaman, S. & Guillaume, F. (2009). Predicting adaptation under migration load: the role of genetic skew. *Evolution*, 63, 2926-2938.

Zheng Y.Q. & Ennos R.A. (1999). Genetic variability and structure of natural and domesticated populations of Caribbean pine (*Pinus caribaea Morelet*). Theor. Appl. Genet. 98, 765-771.

GLOSSARY

Adaptive divergence: the differentiation among the mean phenotypes of populations subject to different selective pressures.

Adaptive trait: a phenotypic trait that enhances the fitness of an individual in a particular environment. Examples for trees are the time of flushing and bud set, and physiological traits determining water use efficiency.

Aerobiology: the study of airborne organisms and organic particles.

Bioclimatic envelope: the predicted potential geographic distribution of a species in a particular climatic scenario.

Breeding value: part of an individual's phenotype that can be transmitted to its offspring through the transmission of genetic material.

Bud set: formation of a terminal bud at the end of the vegetative period, the timing of which is heritable and determines cold tolerance in boreal and temperate trees.

Connectivity map: a map describing the cost or the probability of dispersal along possible trajectories linking a set of locations.

Dispersal kernel: a probability density function of dispersal distances or locations from a source point.

Effective dispersal: dispersal leading to successful establishment or reproduction.

Evolutionary load: the mean fitness loss in a population produced by the deviation of the mean phenotype from the local optimum due to varying selection in space and time.

Fat-tailed dispersal kernel: dispersal kernels with a slow probability decrease at long distances relative to a negative exponential.

Genetic assignment: the probabilistic ascertainment of the original population of an individual genotype.

Genetic correlation: non independent genetic variation for two phenotypic traits, which can be due in particular to the fact that the same genes affect variation of several traits.

Genetic load: the loss of mean fitness in a population due to the departure of individual phenotypes from the optimum in a given environment.

Genetic variance: part of the total phenotypic variance that is due to genetic differences between individuals.

Heritability: the proportion of phenotypic variation among the individuals of a population in a particular environment that is due to genetic variation.

Heterosis: the higher fitness of progeny obtained through crosses between populations rather than within the same population.

Inbreeding depression: reduced fitness of inbred individuals.

Lag load: the loss of mean fitness in a population due to the lagging response of the phenotypic mean to temporal changes in the optimum.

Lagrangian dispersal model: a mathematical description of the trajectories of individual dispersers.

Migration load: the contribution of immigrant genes to maladaptation.

Outbreeding depression: reduced fitness of individuals born to parents from different populations.

Parentage analysis: the probabilistic determination of the parents of an individual, frequently using genetic markers.

Phenotypic cline: a continuous change of a phenotypic trait along an environmental and/or geographical gradient.

Pollen/seed shadow: the density of pollen grains/seeds dispersed at different distances from an individual. It equals the product of the dispersal kernel by the individual's fecundity.

Provenance: the original geographic source of a population or group of individuals (used also to refer to such a population or group).

Provenance test: a common garden experiment, in one or more locations, where the genetic variation of different provenances is evaluated (see provenance).

Reaction norm: the set of phenotypes expressed by a particular genotype under a range of environments.

Response to selection: the difference between the mean phenotype of the offspring of a group of selected parents and the mean phenotype of the population before selection.

Stabilizing selection: selection that favours intermediate over extreme phenotypes.

Standing load: the loss of mean fitness in a population due to the phenotypic variance around the mean phenotype. Present even when the mean phenotype matches the optimum.

CONSIDERING EVOLUTIONARY PROCESSES IN ADAPTIVE FORESTRY

François Lefèvre*, Thomas Boivin, Aurore Bontemps, François Courbet, Hendrik Davi, Marion Durand-Gillmann, Bruno Fady, Julie Gauzere, Cindy Gidoin, Marie-Joe Karam, Hadrien Lalagüe, Sylvie Oddou-Muratorio, Christian Pichot.

Corresponding Author:
INRA, UR629 Ecologie des Forêts Méditerranéennes, URFM,
Domaine Saint Paul, Site Agroparc, 84914 Avignon Cedex 9, France.
* francois.lefevre@avignon.inra.fr

Context: Managing forests under climate change requires adaptation. The adaptive capacity of forest tree populations is huge but not limitless. Integrating evolutionary considerations into adaptive forestry practice will enhance the capacity of managed forests to respond to climate-driven changes.

Aims: Focusing on natural regeneration systems, we propose a general framework that can be used in various and complex local situations by forest managers, in combination with their own expertise, to integrate evolutionary considerations into decision making for the emergence of an evolution-oriented forestry.

Methods: We develop a simple process-based analytical grid, using few processes and parameters, to analyse the impact of forestry practice on the evolution and evolvability of tree populations.

Results: We review qualitative and, whenever possible, quantitative expectations on the intensity of evolutionary drivers in forest trees. Then, we review the effects of actual and potential forestry practice on the evolutionary processes. We illustrate the complexity of interactions in two study cases: the evolutionary consequences for forest trees of biotic interactions and of highly heterogeneous environment.

Conclusion: Evolution-oriented forestry may contribute adapting forests to climate change. It requires combining short-term and long-term objectives. We propose future lines of research and experimentation.

Key words: genetic resources ; silviculture ; adaptation ; climate change.

Annals of Forest Science (*in press, accepted 07/02/2013*)

Published in *Annals of Forest Sciences* (2014) 71: 723-739, with doi:10.1007/s13595-013-0272-1

Introduction

Ecosystem functioning depends on the adaptation of living organisms to their physicochemical environment. In particular, the maladaptation of trees to local conditions can provoke ecosystem dysfunctions such as forest dieback or failure of regeneration, and it can also affect biotic interactions between trees and associated species. Multi-site common-garden experiments, which allow modelling the reaction norm of current tree populations to climatic parameters, suggest that climate change will lead to a high risk of maladaptation of tree species, at least in some parts of the current distribution range (Savolainen *et al.* 2007; St Clair and Howe 2007). To maintain forest services under climate change, tree stands will have to respond within one to ten generations to (1) more frequent and more intense extreme climatic events, (2) changing mean climatic parameters and (3) other related changes such as parasite outbreaks (IPCC 2007). Consequently, ecological services of the forests will depend on the intensity and velocity of the evolution of tree populations in response to climate change (Rehfeldt *et al.* 2001). Thus, adaptation should be considered in a dynamic perspective, as a bouquet of evolutionary processes that change populations and communities to fit their environment. Among these processes, genetic adaptation, *i.e.* genetic change of a population responding to selection, can be rapid and contribute to the ecological success of species facing climate change including forest trees (Aitken *et al.* 2008; Hoffmann and Sgrò 2011). A recent review stressed the high potential of evolutionary response to climate change in trees (Alberto *et al.* 2013). However, evidence of lack of adaptation does also exist, *e.g.* niche limits and empty niches, including for tree species that have large population size and produce huge quantities of seeds (Bradshaw 1991).

During the last century, foresters have succeeded in adapting forest genetic resources to bioclimatic conditions very different from their native range, obtaining good survival, growth and reproduction in the new environments. Emblematic examples are the worldwide transfer of *Pinus radiata* (Yan *et al.* 2006) and the south to north translocation of *Picea abies* (Skrøppa *et al.* 2010). This adaptation was achieved in very few generations of trees, it proceeds from plasticity and/or evolution. For

each adaptive trait, the phenotypic plasticity and the capacity of evolution depend on the genetic content and the environment of the population, which can both evolve (Pigliucci 2008). Within each population, genetic changes of mean trait value, plasticity and evolvability result from the combination of random and selectively oriented processes that can be affected by forestry practice. Whether immediate response to selection can hamper future evolutions, *e.g.* due to erosion of the genetic diversity, remains an open question. Evidence from breeding experience shows that the genetic responsiveness of populations submitted to continuous selection can be maintained through time for some traits: evolvability was maintained over more than 100 generations of selection for protein and oil content in the Illinois maize breeding population (Moose *et al.* 2004). No such long-term empirical evidence is available for trees. However, local adaptation that commonly emerged in most tree species over the course of post-glacial recolonisation provides another illustration of achieved evolution (Savolainen *et al.* 2007). Noticeably, this local adaptation did not completely erode within-population genetic variation of adaptive traits (Mimura and Aitken 2007; Alberto *et al.* 2013). The long-term maintenance of evolvability also depends on the genetic architecture of the traits under selection, and in the case of polygenic inheritance, Kremer and Le Corre (2012) showed that evolutionary changes first result from the selection of the fittest combinations of gene alleles before it reduces the allelic diversity at individual gene loci.

However, adaptation is not limitless. Futuyma (2010) reviewed the factors that can limit adaptation from the short term to the phylogenic time scale. Focusing on an ecological rather than geological time scale, we can retain here seven constraints to evolutionary changes. Firstly, developmental constraints result from functional interactions among traits involved in the elaboration of the performance. We use here "performance" as a generic term, referring either to fitness components in an ecological perception or to forestry objectives like wood quantity or quality in an agronomic perspective, or to any combination of these traits. Secondly, genetic constraints result from the genetic architecture of traits, with complex epistatic interactions between several genes on one trait or pleiotropic effects of one single gene on several traits. Actually, forestry practices have little (but not null) impact on these first two limiting factors. Then, Futuyma (2010) identified four limiting factors of adaptation on which forestry practice may have direct or indirect impact: lack of genetic diversity, demographic stochasticity (counter-acting directional selection), random genetic drift and asymmetric gene flow (*e.g.* at niche limits). In addition to these, Kuparinen *et al.* (2010) identified another limiting factor potentially affected by forestry practice: low mortality.

Deciphering the factors that determine adaptation in the real forest, from the genes to the traits and from the traits to the performance, is complex. Each environment cannot be reduced to only one parameter, *e.g.* altitude combines temperature, soil, rainfall, biotic factors etc. Similarly, each performance, *e.g.* survival in stress conditions, can be achieved by different combinations of functional trait values. Finally, each value of a functional trait can be obtained by different combinations of gene alleles and interactions. As a consequence, one can hardly attribute a fixed intrinsic adaptive value to each physiological trait or to each gene allele. This complexity is also a chance for adaptation because it provides flexibility and there are multiple biological pathways to reach an ecological solution.

Forest management can enhance forest adaptation to climate change in three ways. Firstly, a full-control strategy consists in replacing the local population by a presumably better fit population. This is achieved through plantation of so-called forest reproductive material, which either comes from a breeding program or from a selected seed stand. This strategy allows for drastic stepwise evolutions, but it requires minimizing uncertainties about the ecological integration of the alien resource in the new site under future climates. Secondly, a driving strategy consists in guiding, *i.e.* supporting and accelerating, natural evolutionary processes using the local genetic resource, ecologically integrated within its current environment. This is achieved through natural regeneration. This strategy only produces progressive changes, limited by the evolutionary potential of the local resource, but it is flexible and relaxes the ecological uncertainty related to introduction of alien material. Thirdly, a combined strategy would follow the driving strategy after enrichment of the local resource with a certain amount of alien material in order to increase the evolutionary potential and to accelerate evolution while limiting the ecological uncertainty due to introduction. Since the first approach has already been treated elsewhere and deserves a complete treatment, *e.g.* see St Clair and Howe (2007) for a concrete experience in *Pseudotsuga menziesii*, we focus here on the second and third strategies. Three main questions emerge in this context: (1) How fast can tree populations respond to changes? (2) Will the populations keep their capacity to adapt to both continuous and unpredictable changes? (3) How can forestry practice affect, positively or negatively, the properties of adaptation and adaptability through time? Due to the complexity of evolutionary mechanisms interacting with highly diverse local conditions and climate change scenarios, the first two questions can only receive case-specific answers. Here, we call evolution-oriented forestry a particular form of adaptive forestry that integrates the enhancement of evolutionary processes among its possible objectives, and we propose a process-based approach to investigate the impact of silviculture on the evolution and evolvability of tree populations facing climate change.

In a first part, we describe the basic evolutionary mechanisms using a limited number of parameters in order to define a simple analytical grid. We show how these few parameters can help understand complex situations. The analytical grid also provides a mechanistic interpretation of the evolutionary constraints mentioned above. Secondly, we use this analytical grid to evaluate the potential effects of current forestry practice and to suggest other silvicultural options that could preserve as much as possible the objectives of forestry while driving the tree populations into faster evolutionary changes. Finally, we review the possible genetic tools available for monitoring adaptive changes and evolutionary processes and conclude with future perspectives for experimental management and research.

The basic mechanisms driving genetic changes as an analytical grid

Genetic diversity is continuously changing: each sexual reproduction event generates new and unique genotypic combinations, some of which are then eliminated by selection and random processes. The strength of selection and genetic drift can be efficiently approached with a limited number of parameters. A wide range of evolutionary scenarios can be obtained when considering the interactions between selection, genetic drift, gene flow and plasticity. Due to the short term considered here, we neglect the effect of mutation.

Single and multitrait response to selection

Selection is the elimination of the less fit, due to low reproduction or mortality. A selection pressure on one trait can change the mean of the population (directional selection), or its variance (stabilizing or disruptive selection), or both. Although it does not integrate all the biological processes that effectively operate, the model of quantitative genetics (Falconer 1960), which assumes that a large number of genes interact with the environment to determine the phenotypic variation of each trait, has long proven a remarkable predictive value throughout all the practical achievements in plant and animal breeding programmes since mid twentieth century. This model predicts the rate of change of trait mean per generation under direct directional selection as:

$R = i.h.\sigma_A$ or, expressed in phenotypic standard deviation units, $R' = R/\sigma_p = i.h^2$

where

i is the intensity of selection, i.e. mean differential between the adults that contribute and those that do not contribute to the next generation expressed in standard deviation units (i directly relates to the proportion of selected individuals)

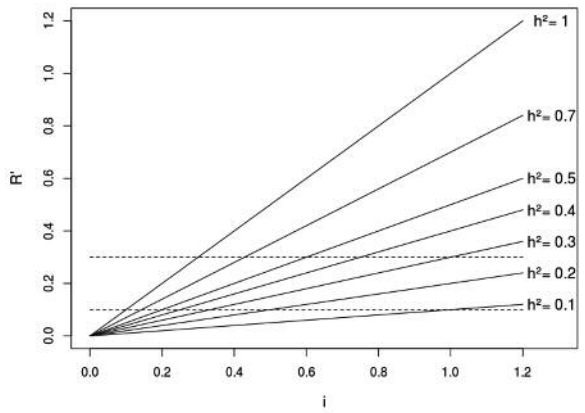
h is the square root of the narrow sense heritability, which is the ratio between the additive genetic variance and the phenotypic variance ($h^2=\sigma A^2/\sigma P^2$)

σA^2 is the additive genetic variance

σp^2 is the phenotypic variance

FIGURE 1

Expected rate of change per generation (R') for a single selected trait under direct selection, expressed in phenotypic standard deviation units, for different values of heritability (h^2) and different selection intensities (i). Assuming a Gaussian distribution of the trait, elimination of 25, 50 or 75% of the population corresponds to values of i = 0.42, 0.80 or 1.27, respectively. The range of empirical values of R' found in the literature by Gingerich (2009) lies between the dotted lines.



Natural selection does not operate directly on the traits but on the global fitness of the individuals. Therefore, the response of each trait to natural selection is a type of indirect selection, and the previous equation is slightly modified as: $R = \beta.\sigma A^2$ where β is the partial regression coefficient of the fitness on the trait, or selection gradient. This formulation extends to the multitrait case where the response on each trait integrates its correlations with other selected traits (Lande and Arnold 1983). Finally, selection is summarized with only four basic parameters: σP^2, σA^2, i (direct selection) or β (natural selection).

Based on a review of empirical studies in all kinds of organisms, Gingerich (2009) found relatively high rates of evolution (R') in the magnitude of 0.1 to 0.3 phenotypic standard deviation per generation, similar for long-term evolution and micro-evolution. Figure 1 shows the expected rate of change of a single trait under direct selection with different values of heritability and selection intensity. As explained in the next section, forest management can modify i and, therefore, proportionally change R'.

A trait will respond to selection if three conditions are simultaneously met (Endler 1986): (1) there is phenotypic variation within the population, (2) this variation is heritable and (3) this variation is correlated with the variation of fitness. Each term in the equation above is not fixed, but it varies depending both on the environment (Charmantier and Garant 2005) and the genetic background of the population (Carter *et al.* 2005). Natural selection in trees is a complex process because the

selection pressure varies between years and it can drastically change between life stages both in direction (*e.g.* shift in selective forces between a dense seedling patch under the canopy and the adult stage) and in intensity (*e.g.* mortality rate is much higher at juvenile stage while selection on reproductive success only occurs at adult stage). Thanks to this variability of the parameters, the response to selection does not systematically correlates with a total exhaustion of the genetic variation within populations: the partitioning of the genetic variance of adaptive traits into between- and within-population components (*Qst* approach) has revealed that a large part

BOX 1

Raising-up complexity (1): the interplay between biotic interactions and evolutionary processes in forest trees

Trees interact with diverse mutualistic and antagonistic insect species. Mutualistic insects may be essential for plant reproduction during pollination, while antagonistic insects may be damageable to plants by consuming and removing plant parts and by selectively feeding on their reproductive tissues (Crawley 1989). The strength of such negative effects depends on the timing, the type and the amount of damage, as well as the stage of the plant's life cycle at which the damage occurs (Marquis 1992). By directly affecting tree reproduction or survival, the demographic and evolutionary consequences of the feeding activities of seed-specialized and tree-killing insect species are thus likely to differ from most forms of herbivory which only result in partial removal of tissues from individual plants (Hulme 1998).

Since seed predation leads to the eradication of individuals in a population, it plays a crucial role in plant population dynamics with possible genetic drift effects when population size is limited and potentially acts as a selective force driving the evolution of particular plant traits such as flowering synchrony, flowering phenology, inflorescence characteristics, flower size, flower longevity and mast seeding (Janzen 1971; Brody 1997; Fenner *et al.* 2002; Cariveau *et al.* 2004; Rose *et al.* 2005; Strauss and Whittall 2006). Many tree species suffer from large seed losses due to pre-dispersal seed predation, which can have significant effects on recruitment and plant population growth rate (Maron and Crone 2006; Kolb *et al.* 2007). However, the effect of such parasites on the long-term fitness of their host plant appears controversial (Crawley 1989; Horvitz and Schemske 2002), mainly due to the lack of data addressing this issue, especially on perennial plants. There is still a crucial need for studies examining the genetic consequences of massive seed losses within a host tree population, especially in a context in which pre-dispersal seed predation shows significant variation between trees. Indeed, seed loss due to seed-specialized chalcid wasps may vary from less than 1% to 100% between trees (Roques 1981; Rappaport *et al.* 1993). At the tree population level, this raises the question of how such local variation in pre-dispersal seed predation may increase or, reversely, decrease the variance of effective seed set among trees, which influences the effective population size.

Evidence of insects directly acting as selective agents on forest trees is still lacking in the literature. Despite the extremely high tree mortality rates recorded during population outbreaks of the bark beetle Dendroctonus ponderosae Hopkins, many trees escape or survive bark beetle attacks, regardless of their vigour, age and/or size (Ott *et al.* 2011). Little is known about the heritability of tree traits involved in survival to bark beetle attacks such as resin acids (Baradat *et al.* 1978) and resin flow, viscosity and rate of crystallization (Nebeker *et al.* 1992), in the exception of monoterpene production, which has been shown to be under strong genetic control (Ott *et al.* 2011).

Understanding how abiotic and biotic disturbances and tree dynamics are interdependent is also crucial for predicting the overall impact of parasitism on tree evolution. Indeed, severe abiotic changes such as droughts and/or heat waves may affect trees and parasites, as well as their interactions (Jactel *et al.* 2012). Successive drought episodes can affect directly tree survival (Allen *et al.* 2010), or indirectly when higher temperatures and lower tree resistance trigger severe forest insect outbreaks (OFEFP 2005; Netherer and Schopf 2010; Durand-Gillmann *et al.* 2012). The interdependence between climate, biotic factors and tree dynamics remains complex to predict. Drought induced changes in tree nutritional quality (water, carbohydrates and nitrogen contents) or in tree defence mechanisms can limit the development and the damages of parasites (Rouault *et al.* 2006; Jactel *et al.* 2012; Forkner *et al.* 2004). Extreme droughts may even be directly involved in the collapse of herbivorous populations at wide scales (Yarnes and Boecklen 2005). But drought can also affect negatively tree physiology and decrease the effectiveness of tree resistance mechanisms to pathogens and parasites (McIntyre *et al.* 1996).

Fire ecology provides interesting additional examples of the complexity of integrating interdependencies between trees, biotic and abiotic factors. Bark beetle outbreaks and forest fires have indeed jointly increased in extent and severity during the last decades, raising concerns about their possible interactions (Parker *et al.* 2006; Simard *et al.* 2011). Bark beetle outbreaks may increase the probability and intensity of active crown fire because they create great quantities of dead and ladder fuels (Brown 1975; McCullough *et al.* 1998). However, Simard *et al.* (2011) suggest that active crown fire are less probable in the short-term after outbreaks due to insect-driven stand thinning, while the probability of passive crown fire does not change in the short term but greatly increases in the decades following an outbreak. Thus, bark beetles are likely to indirectly affect non-attacked trees through subsequent enhanced fire risks. This clearly illustrates the critical need to integrate the possible interplay between the abiotic environment, biotic interactions and trees dynamics when designing forest management strategies.

of the genetic variance of functional traits is maintained within population (Alberto *et al.* 2013, for a review), and an absence of within population genetic variance for a quantitative trait has exceptionally been reported in trees (Sáenz-Romero *et al.* 2006). Interestingly, the genetic architecture of a trait, *i.e.* the system of genes involved in the variation of the trait and their interactions with other traits, simultaneously determines long-term persistence of evolvability by a capacity to release cryptic variation in a new environmental or genetic context (Le Rouzic *et al.* 2007) as well as a potential limit to selection in case of detrimental genetic correlations (Walsh and Blows 2009). Functional constraints, resulting in environmental correlations between traits, can also limit the response to selection. Not included into this predictive model, epigenetic effects, i.e. environmentally determined heritable modification of gene expression, can also contribute to adaptation to sudden changes (Bossdorf *et al.* 2008; Skrøppa *et al.* 2010).

Random changes due to genetic drift and mating system

Genetic drift accounts for the reduction of genetic diversity that occurs in small populations, in absence of selection, mutation or migration, due to the variation of allele frequencies after random sampling from one generation to the next. Furthermore, small populations are prone to increased inbreeding due to the higher probability of mating between relatives. Inbreeding has a twofold effect: it reduces fitness whenever inbreeding depression is present, and it retains non-random association of gene alleles (linkage disequilibrium) at higher rate which represents a reduction in the diversity of genotypic combinations. Non-random mating system can also affect inbreeding: the mating system varies among individuals and populations, including selfing rate (most of tree species are not dioecious) and diversity of pollen donors, depending on the relative fecundity and spatial distribution of reproducing trees.

Under the assumptions of Wright-Fisher's model population (Wright 1931), the reduction of gene diversity and the increase of inbreeding are driven by one single parameter, population size. Using this model as a reference, the effective population size (*Ne*) of a real population of size *N* that has a per-generation rate of reduction of gene diversity (*DHe*) or increase of inbreeding (*DF*) is such that: $DHe = -1/2Ne$ or $DF = 1/2Ne$. It can be shown that, in the absence of dominance, *Ne* also measures the per-generation rate of reduction of additive variance: $D\sigma A^2 = -1/2Ne$. *Ne* is defined on the rate of change of gene diversity or inbreeding, not on the actual population size *N*. *Ne* is most often not directly estimable in natural populations (unless longitudinal estimates of gene diversity, *He or* σA^2, or inbreeding, *F*), but its changes can be predicted and decreasing *Ne* means intensifying the intensity of genetic drift. When the actual population only departs from the theoretical model by relaxing the assumption of Poissonian distribution of reproductive success, it can be shown that: $Ne = (4N-2)/(V+2)$ where *V* is the actual variance in reproductive success, *i.e.* *Ne* decreases substantially in proportion to this variance.

Tree populations are generally assumed to have large effective population size (Petit and Hampe 2006), in part because they are outcrossing and disperse their genes over long distances in particular through pollen (Ashley 2010); thus, they should not be too much affected by genetic drift. In their review, Schoen and Brown (1991) found *Ne* estimates for tree species in the range of other outbreeding plants: mean values around 3,000 for *Pseudotsuga menziesii* and >8,000 for *Pinus sylvestris* and *Picea abies*, with high variations among populations within each species. However, locally, seed and pollen contributions to reproduction are highly uneven among individuals (Burczyk *et al.* 2002; Krouchi *et al.* 2004; Oddou-Muratorio *et al.* 2005), and the great majority of the pollen disperses only in the close neighbourhood, which can greatly reduce the effective population size. In their review, Smouse and Sork (2004) found that the effective pollen pool size *Nep*, defined as the inverse of the probability that a female draws two offsprings from the same father, ranges from 2 to 200 in tree populations. *Nep* can be very small in some populations: fragmented populations of wind pollinated species of *Quercus humboldtii* (Fernandez-M and Sork 2005) and *Quercus alba* (Smouse *et al.* 2001) exhibit estimates of *Nep* around 6 and 8 respectively. By contrast, in continuous forest populations, several examples estimate high values of *Nep*: Robledo-Arnuncio *et al.* (2004) estimated a *Nep* >70 in a Spanish population of *Pinus sylvestris*. Fragmented populations, isolated populations and populations at low density have a higher risk of extinction due to the erosion of diversity by genetic drift (Goodell *et al.* 1997; Hardy *et al.* 2004; Robledo-Arnuncio *et al.* 2004; Aguilar *et al.* 2008).

Interactions between selection, drift, gene flow and phenotypic plasticity

A well-known interaction between drift and selection is the vortex of extinction (Gilpin and Soule 1986): when there is a genetic load in the population, a rapid decrease in population size leads to increased genetic drift and increased inbreeding, resulting in reduced mean fitness that further reduces population size, which over time will result in extinction in a geometric decline. However, there is no experimental evidence to our knowledge that this kind of extinction vortex ever occurred in trees. Alternatively, resistant genotypes that emerge in the population increase their contribution to the next generation and can restore population growth if not eliminated at random, a process known as evolutionary rescue (Gomulkiewicz and Holt 1995). Whether

populations can be rescued depends on population size, genetic diversity and the degree of maladaptation to the new environment.

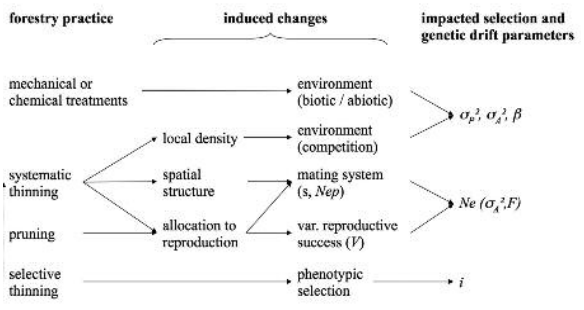
In the case when the environment changes both in space and time, gene flow can bring into the population pre-adapted genes (Pease *et al.* 1989; review by Kremer *et al.* 2012). Kuparinen *et al.* (2010) showed that pollen and seed dispersal at longer distance speed up the adaptation process. In tree populations, it is expected that pollen-mediated rather than seed mediated gene flow will contribute to this processes, with average pollination distances commonly being hundreds of meters (Ashley 2010), and maximum distance of 100km measured in *Pinus sylvestris* airborne transported pollen (Robledo-Arnuncio 2011). In this species, along a latitudinal gradient, Nilsson (1995) showed that long-distance pollen migration brings into the population new phenotypes with a phenology specific of other locations. Local individuals with extreme phenotypic values, in particular for flowering phenology, are keen to catch more long-distance pollen grains because they are better synchronized with source population of interest and also because they are less saturated by local pollen.

Abundant theoretical literature exists on the role of phenotypic plasticity in evolution; recent reviews of predictions on the interaction between selection and plasticity in the context of climate change can be found in Chevin *et al.* (2010, 2012), the second review also provides key references of empirical studies of plasticity in trees. Maladaptive plasticity is obviously detrimental to adaptation. When adaptive plasticity is not genetically variable, it is expected to slow down the genetic response to directional selection in each generation but also to allow the phenotypes to track the environmental change more closely. The sustainability of this process of adaptation depends on the fitness cost of plasticity. When adaptive plasticity varies genetically, *i.e.* there is GxE interaction and plasticity can evolve, the amount of genetic variance of the plastic trait depends on the environment: if the new environment increases the genetic variance, then plasticity tends to accelerate the genetic response to selection and plasticity is itself selected for (Lande 2009). Considering a steep spatial environmental gradient, where gene flow interacts with selection and plasticity, the evolution of plasticity is expected to allow the population to explore a larger range, and marginal habitats are expected to show higher plasticity (Chevin and Lande 2011).

A further degree of complexity arises when considering the interplay between abiotic environment, biotic interactions and evolutionary processes (Box 1).

FIGURE 2

Expected effects of forestry practice on evolutionary processes: schematic pathway from the forestry management in natural regeneration systems to the selection and genetic drift parameters. See text for explanation of the paths (σP^2 phenotypic variance, σA^2 additive genetic variance, β selection gradient for natural selection, i selection intensity for direct selection, s selfing rate, Nep effective pollen pool size, V variance of reproductive success, Ne effective population size, F inbreeding).



Potential effects of forestry practice on the rate of evolution of trees

A global framework to assess the evolutionary impact of silviculture

Considering the evolutionary impact of silviculture is thus an additional requisite to adaptive forestry. While facing climate change and uncertain future, forestry practice should simultaneously accelerate genetic adaptation by helping tree populations to track the known environmental changes and preserve the genetic diversity as a reservoir of future options to respond to the next unknown changes. This is a classical challenge in breeding programs: obtaining a rapid genetic gain while keeping the potential for long-term response to selection. Breeders solve this problem by combining selection and recombination. Genetic adaptation occurs if better performing genotypes emerge during reproduction and if the best performing genotypes spread in the population before extinction. Therefore, we can assign three objectives to evolution-oriented forestry practice: (1) increase the chance of emergence of "innovative" genetic combinations, (2) facilitate the spread of the best adapted genotypes and (3) preserve genetic diversity for long-term response to selection. Acting simultaneously on the demography and the environmental conditions, silviculture has an impact on the parameters of selection and genetic drift. Due to the multiplicity and complexity of the processes involved and to the huge diversity of biological and environmental features among forests, it is more reliable to predict the effects of silviculture on the parameters of evolution rather than on the final state of the genetic diversity. Thus, silviculture should aim to limit the intensity of genetic drift, increase the diversity of mating pairs, avoid counter-selection and maintain selection pressure in the right direction, reduce inbreeding etc. Here, we briefly consider the evolutionary benefits, or risks, associated to current or future forestry practice. A more quantitative prediction of

their balanced effects could rely on simulation studies using process-based models that explicitly integrate demographic, genetic and biophysical processes and allow to simulate interventions (Kramer *et al.* 2008).

Foreseen evolutionary impact of common forestry practice

A general consequence of silviculture on the drivers of evolution is the change of environmental conditions: competition and, possibly, other biotic or abiotic environmental factors. As mentioned above, the environmental conditions determines not only the phenotypic variance (σP^2) of the traits but also the genetic variance (σA^2), in case of GxE interactions, and the selection gradient that relates trait value to fitness (β) (Figure 2). Furthermore, increased canopy closure also affects pollen dispersal and mating system parameters: selfing rate (s), effective pollen pool size (*Nep*) and variance of reproductive success (*V*) (García *et al.* 2005; Milleron *et al.* 2012).

Beyond these general effects, thinning, whether systematic or selective, will affect the spatial clustering of genetically related reproducing trees and their allocation to reproduction and, therefore, the mating system (Figure 2). By removing neighbouring-related individuals, thinning may reduce spatial genetic structure and, consequently, inbreeding in the regeneration (Dounavi *et al.* 2002). Sagnard *et al.* (2011) showed that, when the density of seed trees is low (<16 trees.ha^{-1}), a clustered distribution of seed trees will produce less spatial genetic structure in the seedlings than a random or dispersed distribution. Robledo-Arnuncio *et al.* (2004), in *Pinus sylvestris*, and Restoux *et al.* (2008), in *Abies alba*, found at seed stage that low densities of reproductive trees simultaneously increase the probability of selfing (from <5% to >20%) and the diversity of the outcrossing pollen (*Nep*), which can be mechanistically explained by a higher rate of long distance pollen pollination (Klein *et al.* 2006). Therefore, lower seed trees density can result in higher genetic diversity at seedling stage, whenever pollen is not limiting and purge of inbreeding occurs early, but it can be a risk when pollen is limiting, which must be considered in the context of climate change.

When the local population size is limited, a genetic drift effect may result from the reduction of the number of reproducing trees, leading to a loss of rare alleles (among which currently deleterious genes). Such effect was observed in old-growth forests of *Pinus strobus* comparing pre- and post-harvest stands (>100 and <30 trees, respectively) (Buchert *et al.* 1997; Rajora *et al.* 2000), as well as in *Picea rubens* old growth forests (Mosseler *et al.* 2003). Konnert and Hussendörfer (2001) compared 16 even-aged and nine uneven-aged management systems in *Abies alba*, several forests for each group and also the two management systems within the same forest: they found a slightly higher number of rare alleles but smaller number of different gametic combinations in the uneven-aged forests. However, there are too few studies of this type to draw general conclusions on an eventual intrinsic difference between these two management systems regarding their effects on the genetic diversity. The drift effect is not *a priori* limited to the managed population but it may also affect secondary tree species as observed for *Abies amabilis, Tsuga heterophylla, Thuya plicata and Pinus monticola in Pseudotsuga menziesii* forests by El-Kassaby and Benowicz (2000).

In naturally regenerated stands, reproductive trees usually result from successive selective thinnings for their phenotypic value such as stem vigour, health consideration, lack of defects like thick branches or forked or twisted trunk, as well as for their spatial distribution in order to reduce competition. In *Fagus sylvatica*, selective thinning favouring the most vigorous trees was found to increase the heterozygosity by 4% to 9% (Lauber *et al.* 1997; Dounavi *et al.* 2002), even when selection occurs at a very early stage (Thiebaut *et al.* 1992). However, this effect was not detected in other studies on *Abies alba* (Hussendörfer and Konnert 2000) or *Pinus contorta* (McDonald *et al.* 2001). We must remind here that an increase in heterozygosity during maturation of forest stands is frequently observed and that natural selection of seedlings can start at a very early stage in overstocked young regeneration (Pichot *et al.* 2006). Selective thinning favouring the best growing trees may act as selection for competing ability: assuming a Gaussian distribution of tree height, the elimination of the 25% (resp. 50%) smallest individuals represents a selection intensity i = 0.42 (resp. 0.80) on this trait. Selective thinning may also integrate a direct selection on other traits chosen by the forester (Figure 2). We need to understand more clearly which functional traits are indirectly selected for and how these traits relate to future fitness in the context of changing climate. One main question that still needs to be addressed is how far silviculture will intensify selection for juvenile vigour and, if this is the case, how far juvenile vigour is genetically positively or negatively correlated with drought resistance.

The intensity of genetic drift, through the variance in reproductive success, and the mating system not only depends on the number and spatial distribution of reproductive trees but also on their allocation to reproduction. As indicated in Figure 2, this allocation is influenced by local stand density as well as other practices like pruning (Ayari *et al.* 2012). Thinning and pruning also affect the plasticity, which will interact with selection processes as previously mentioned: the question here is whether the plastic response induced by the silviculture is adaptive or maladaptive regarding future climate change.

Multiple forestry practices may also be combined and interact to modify the microenvironmental conditions of development, generally as to reduce stress and

competition in order to favour growth (Forrester *et al.* 2012). Thus, after an extreme drought event in 1976, stand decline was reduced in *Picea abies* stands that had previously been thinned in 1971 (Misson *et al.* 2003): compared to the control plot, heavy thinning had a more beneficial impact than moderate thinning (thinning from 36 to 14 $m^2.ha^{-1}$ or to 20 $m^2.ha^{-1}$ basal area, respectively). We are not aware of any study of the effects this could have on plasticity (*e.g.* reduced acclimation to future stress) and selection. As a case study, we present the analysis of evolution and potential effects of silviculture in the situation of highly heterogeneous environment within the forest (Box 2).

Evolutionary benefits and risks expected from some silvicultural recommendations related to climate change

New forestry practice is progressively implemented to reduce ecological and economic risks related to climate change (Legay and Mortier 2005; Yousefpour *et al.* 2012). From an evolutionary point of view, reducing environmental stress has a twofold effect. On the one hand, it reduces the damages and therefore contributes to increase the effective population size (*Ne*), which is a crucial issue when population size is already small or is expected to decrease drastically due to severe damages. But, on the other hand, it also slows down the genetic improvement in the next generation by reducing selection intensity (*i*) and it does not exploit potential adaptive plasticity (no acclimation to future stress), which is an important issue to consider in large populations. We briefly consider here some of these practices, or changes in practices, from the evolutionary point of view and illustrate their possible balanced effects. In all cases, whenever applying new practice, it is crucial to keep precise records of what is done, how and when, in order to facilitate future evaluation *ex post*, in particular after marked climatic events.

Shortening rotation reduces the probability of the risk, *e.g.* to extreme climatic events, but it can also increase the vulnerability to the risk if shorter rotations select for higher juvenile vigour and if juvenile vigour is genetically negatively correlated with stress resistance. Both of these conditions still have to be investigated. The answer will probably depend on the species, on the environment and on the management system considered.

Reducing the density of stands is envisaged to reduce the effective drought stress supported by the trees. However, this immediate positive effect may be partly balanced by a long-term detrimental effect on selection by inducing a maladaptive phenotypic response and by reducing selective mortality (see first section). This risk is reduced if time is left for sufficient natural selection to proceed before thinning. More generally, from the evolutionary point of view, interventions occurring at juvenile stage raise the question of age-age correlations. Apart from the

BOX 2

Raising-up complexity (2): evolution and silviculture in a highly spatially heterogeneous environment

In trees, the selection process is complex due to the long life-cycle and the high within-stand spatial environmental heterogeneity. Firstly, for long-lived and sessile organisms, different selection pressures may occur successively from the juvenile stage to the adult stage, e.g. selection for competing ability in a young dense regeneration vs. selection for stress resistance in the adult stage. Secondly, in heterogeneous environment, the phenotypic correlation between parents and offspring do not only depend on genetic control of the phenotype but also on the difference in environmental effects between the parents' and offsprings' sites. Finally, the environmental heterogeneity induces spatial variation in selection pressure, eventually leading to different selection pressure between parents and offsprings. Thus, evolution-oriented forestry should take environmental heterogeneity into account.

As a general objective, evolution-oriented silviculture should aim at favouring the mating success of the best growing trees located in the patches where the highest desired selection pressure occurs. In homogeneous conditions where the selection pressure is uniform, this objective is directly achieved by classical selective thinning. In heterogeneous conditions, this objective could be achieved if enough trees are selected for seeding within each patch where high selective pressure occurs, even though these trees may have lower growth than neighbours growing in more favourable conditions.

In such context, assessing and mapping the environmental heterogeneity among patches is essential to avoid confusion between micro-environmental and genetic effects on the performance of the trees. It is crucial not only to assess the individual performance but also the patch conditions with synthetic indicators, easy to use and independent of the competition, such as site index, species composition of vegetation etc. Local variations of site index can be assessed through spatial patterns of tree height; therefore it would be easier to identify and mark very early the 'good' phenotypes in highly selective areas, before a substantial reduction in number of trees happens due to other criteria.

Genetic improvement of the whole population will occur if these highly selected trees effectively contribute to the regeneration at stand level. If pollen or seed dispersal limits their effective contribution, considering here fecundity as included into the global dispersal process, it might be necessary to assist natural regeneration by local seed transfer from low to high selection patches within the stand.

temporal changes in environmental conditions, life stages differ in their physiology and development. During the complex and temporally changing selection process in trees, juvenile-adult genetic correlations contribute to determine how far a selection pressure (or release of selection) during the juvenile stage will genetically affect the adult population. This issue can hardly be addressed *in situ*. Partial answer, here again, comes from the breeding experience and early selection schemes. Studies on growth and wood density in different *Pinus* species revealed that genetic age-age correlation >0.8 is generally achieved from the age of 10-12 years (Hannrup and Ekberg 1998; Gwaze *et al.* 2000; Wu *et al.* 2007; Bouffier *et al.* 2008). It varies greatly with environmental conditions, and in *Pinus radiata*, this level of correlation can be reached as soon as 2-5 years in certain sites (Gwaze *et al.* 2000; Wu *et al.* 2007). Matheson *et al.* (2002) showed that genetic age-age correlations in *Pinus radiata* also vary with the genetic background with higher correlations in presence of inbreeding. Thus, silviculture may have an effect on age-age correlations through its effects on the environmental conditions and on the genetic background. Further investigations on age-age correlations of functional traits are deeply needed.

In the case of massive dieback, sanitary logging is necessary to reduce the spread of primary or secondary parasites, and it can also be necessary for fire prevention or for the protection of forest users. However, excessive elimination of surviving trees could result in the elimination of resistance to the pathogen (Burke 2011).

Evolution-oriented forestry, why not?

We imagined some specifically evolution-oriented forestry practices in the case of natural regeneration management system (Table 1). These interventions should not be directly considered as recommendations or guidelines as such; we rather propose them as case studies to illustrate innovative adaptive forestry that would take into account short- and long-term evolutionary potential, still to be associated with other clues. This is not an exhaustive list, and any combination of the proposed interventions can be envisaged.

In order to reduce the intensity of genetic drift (increase *Ne*) in small populations, silviculture may be oriented towards reduced variance of fecundity (*V*) between trees: reducing *V* will not only increase *Ne* within each annual seed production, as mentioned above, but it will also reduce the fluctuation of effective contributions across years and thus increase pluri-annual *Ne* estimates (Krouchi *et al.* 2004). This would be another objective assigned to thinning and pruning. It requires a balance between keeping sufficient number of seedling trees and sufficient spacing between them (optimisation should be made on a case-by-case basis, depending on dispersal capacities). Due to the tree x year interaction effect on the variation in fecundity generally observed in trees, a general recommendation would be to cumulate reproduction during several years. Actually, current practice may already be optimal for this purpose. A negative side effect is to slow down the elimination of detrimental alleles (Couvet and Ronfort 1994) and reduce the response to selection. A compromise between preserving the genetic diversity for the future (reduce genetic drift) and accelerating the immediate response to selection could be to equalize the mating success per patch, in particular when the environment is spatially heterogenous (Box 2).

In order to reshuffle the local genetic diversity and increase genetic recombination, silviculture could enhance local gene flow, either through artificial dispersal of local seeds or by assisting pollen dispersal. With the same objective, isolated seed trees should be considered with care: from one side, they may have a higher selfing rate but, on the other side, they can capture long distance pollen flow. If selfed seeds are eliminated at an early stage of development (*e.g.* empty seeds in some conifer species), the fertile seeds hamper a large genetic diversity. In the case of heterogeneous environment within the forest, areas for wood production and areas for evolution could be spatially dissociated while maintaining gene flow between these entities (Box 2).

In an environmental cline, typically an altitudinal cline, a strategy might be to accelerate the migration of the population towards more favourable areas. The velocity of migration depends on the effective dispersal, and effective dispersal is highly dependent on the local conditions for seedling establishment conditions (Amm *et al.* 2012). We can imagine to enhance seed germination and seedling growth by preparing the soil or controlling competition and predation at distance from the core of the population, in the direction wanted for migration. For zoochorous species, we can also imagine to attract seed dispersers along the wanted migration route (Oddou-Muratorio *et al.* 2004; García *et al.* 2009; Schleuning *et al.* 2011).

Genetic enrichment of the local genetic resource by the introduction of a limited amount of allochthonous material from a putatively pre-adapted origin, through seed or pollen introduction, could present a twofold benefit of introducing gene alleles of interest and increasing the global genetic diversity. To avoid gene swamping effect and reduction of the effective population size, it is essential to use a large genetic base of the introduced material (Lefèvre 2004). For long-lived organisms, it is also important to anticipate a possible trade-off between adaptation to long-term climatic trend and adaptation to current conditions and/or to annual fluctuations, such as vulnerability to late frost of early flushing genotypes.

As a complete utopy, we can imagine future access to intensive genotypic data on each adult tree. Inspired from the marker-assisted selection strategies used in plant and animal breeding, marker-assisted selective thinning could

combine the objectives of increasing the adaptive change for the target traits that are unambiguously identified while preserving the maximum diversity in the rest of the genome.

Genetic monitoring and study tools

In a recent publication, Hansen *et al.* (2012) reviewed the different tests and approaches for genetic monitoring of adaptive changes using phenotypic or molecular tools. Their focus was on the capacity to demonstrate the adaptive response and rule out alternative hypotheses that might explain the genetic change. In this section, we investigate how far recent genetic monitoring methods and tools can help to rationalize evolution-oriented forestry.

What's new in molecular and phenotypic tools?

A comprehensive review of the genetic markers and their use in trees was published by Prat *et al.* (2006). Almost all kinds of markers have been developed on one or several tree species, and they were mainly used to infer on the neutral genetic diversity and neutral processes (drift, mating system, dispersal). With the classical genome sequencing projects, which started in the 1990s for trees, a gap in terms of available tools had progressively appeared between a very limited number of model tree species and the other species. In the last 3 years, recent advances in DNA sequencing have revolutionized the field of genomics making it possible to generate a large amount of sequences and markers in time- and cost-effective way. Nowadays, thanks to the emergence and evolution of the so-called next generation sequencing techniques and

TABLE 1

Some examples of evolution-oriented forestry practice, including re-orientation of usual interventions (no supplementary cost) and additional interventions

Forestry practice	Expected benefits	Associated costs and risks
Ne-oriented regulation of the density and spatial distribution to equalize reproductive success between trees in small populations	- Reduce the variance in reproductive success to reduce genetic drift - Reduce spatial genetic structure in the seedlings and inbreeding in next generation	- No supplementary cost - Risk to slow down the elimination of detrimental genes, prefer equalization of mating success per patch (compatible with the next line)
In heterogeneous environment, dissociate areas of production and areas of evolution (selection patches in harsh areas) and allow gene flow between these entities	- Increase the reproductive contribution of the trees that have survived to drastic selection pressure	- Limited supplementary cost - Requires preliminary simulation studies to estimate benefits in various contexts (strength and spatial structure of the environmental heterogeneity)
Save the lone tree, which cumulates long distance dispersal (in allo-pollinated seeds) and can be adapted to marginal conditions; collect seeds for local assisted regeneration	- Diversify the mating pairs to favour the emergence of new genotypic combinations - Promote adaptation to marginal conditions	- Limited supplementary cost - Requires a protocol for assisted regeneration - Risk of inbreeding if self-pollinated seeds are not purged at a very early stage (e.g. seed abortion)
Assisted local seed dispersal (*e.g.* collecting, possibly over several years, mixing and replanting seeds within the stand) or pollen dispersal (*e.g.* air flow used in seed orchards)	- Enhance local gene flow to diversify the mating pairs and favour the emergence of new genotypic combinations - Reduce inbreeding	- Potentially significant supplementary cost - Requires preliminary studies to estimate benefits in various contexts (genetic diversity and spatial structure) - Requires a protocol for assisted regeneration
Enhance local migration capacity by favouring seed dispersal and germination at distance from the main stand	- Speed-up colonisation of locally favourable habitats in an environmental gradient	- Potentially significant supplementary cost
Genetic enrichment by introduction of a limited amount of seeds or pollen from presumably pre-adapted allochthonous origins	- Introduce pre-adapted genotypes - Increase local genetic diversity	- Potentially significant supplementary cost - Risk of gene swamping and reduction of effective population size (Ne) if local population is small and if introduced material has low genetic diversity - Risk of unforeseen local maladaptation
Marker-assisted selective thinning (futurist)	- Increase selection intensity on target major genes while retaining genetic diversity in the rest of the genome	- High supplementary cost - Requires accurate genetic knowledge and high-troughput genotyping capacities

related bioinformatics (Metzker 2009; Kircher and Kelso 2010), full-genome sequencing, reduced-representation sequencing and targeted sequencing are in progress not only for model species but also for non-model species where molecular monitoring is becoming worth considering from scratch. Evidence of genetic changes at molecular level in trees has recently emerged from the direct study of genome wide DNA polymorphisms: evidence of the correlations between genotype frequencies and environmental gradients (Eckert *et al.* 2010) or climate-related traits (Grivet *et al.* 2011), sometimes completed with functional information on the detected genes (Holliday *et al.* 2010). See also Alberto *et al.* (2013) for a recent review of single nucleotide polymorphisms associated to climate related traits in trees. However, as Rockman (2012) very wisely stated, the extraordinary potential of these approaches should not be misleading: due to the genetic and environmental sensitivity of the response to selection as previously discussed, we should not expect to find many single gene alleles having large, constant and uniform effect in all populations. If such nucleotides are detected, we should not reduce the genetic variation to them because most of the genetic diversity of interest for adaptation to climate change will remain cryptic. These approaches will nevertheless be very useful in providing genetic indicators of the selection pressure.

The phenotypic approach of adaptation has also evolved in two directions. Firstly, the physiologists have produced proxies of physiological functions that can be measured in large sample size (hundreds of individuals). The relation between the measured trait and the actual function is generally indirect and requires careful interpretation, *e.g.* carbon isotope discrimination or ring density used as proxies of the response to drought (Osório and Pereira 1994; Tene *et al.* 2011). It is probably worth reminding that splitting an integrated phenotypic trait into simpler functional components does not resolve the complexity, *e.g.* functional components do not necessarily have higher heritability or simpler genetic determinism than the integrated trait, because new interactions and regulations appear at finer scale. Secondly, methodologies combining phenotypic and genotypic information have been developed to estimate genetic parameters (variances and correlations) *in situ, i.e.* in the natural environment, at any life stage and without requiring controlled pedigrees (Ritland 1996). To estimate selection gradients (β) in trees, a common approach is to use performance traits like survival, growth or reproductive traits as proxies of fitness and study the impact of functional traits on these performance traits in controlled *ex situ* progeny tests. Using this approach in *Quercus suber*, Ramirez-Valiente *et al.* (2011) detected significant heritability but non-significant selection gradient for carbon isotope discrimination, contrasting with very low heritability and significant selection gradient for specific leaf area. An alternative approach of selection gradients through the assessment of actual reproductive success *in situ* was recently developed (Oddou-Muratorio *et al.* 2005; Burczyk *et al.* 2006; Klein *et al.* 2011). This method, based on the mixed-mating neighbourhood model, consists in estimating the reproductive success of individual adult trees using spatial genetic data of seedlings and their potential parents and then in relating this reproductive success to phenotypic traits. Bontemps (2012) used this method in a marginal population of *Fagus sylvatica*: in this case, the author found significant heritability and significant selection gradient for carbon isotope discrimination, contrasting with non-significant heritability and non-significant selection gradient for specific leaf area.

Possible monitoring

Various sets of state and pressure indicators have been proposed for the monitoring of the genetic diversity in forest trees, based on direct genetic assessment or indirect observations through the demography and ecology of the populations (Namkoong *et al.* 1996; Brown *et al.* 1997; Koski *et al.* 1997; Lefèvre and Kajba 2001; Aravanopoulos 2011). Table 2 briefly reviews the possible uses of molecular and phenotypic tools, and the requirements of these uses, for genetic monitoring of adaptation. We distinguish two main objectives for genetic monitoring: (1) quantification and characterization of the genetic diversity and (2) monitoring of the drivers of genetic changes.

The genes controlling the variance of most adaptive traits (quantitative trait loci, QTL) are expected to be numerous with small individual effect. This is confirmed by empirical results, even though molecular tools only detect a small fraction of these QTLs. The QTLs and their genetic effects vary depending on the environment and the genetic background. Therefore, defining the adaptive genetic diversity as the whole set of QTLs potentially affecting fitness components, there is no strict frontier between neutral and adaptive genetic diversity: a gene polymorphism that is neutral in one context may become adaptive in another environmental or genetic context and vice versa. Thus, considering climate change and its uncertainties, quantification and characterization of both currently adaptive genetic diversity and currently neutral genetic diversity are needed. Classically, the overall genetic diversity is assessed with neutral markers (Buchert *et al.* 1997). Quantification and characterization of the current adaptive genetic diversity is assessed using phenotypic and molecular tools. Using only the phenotypic tools, the assessment of the genetic matrix of variances and covariances of adaptive traits requires known pedigrees and *ex-situ* controlled experiments (preferable in trees for which deep pedigrees are not available in natural populations like in other organisms). Candidate gene polymorphisms, which include direct functional polymorphisms as well as linked markers, provide direct

or indirect information on the diversity of target genes. When markers and phenotypic tools are combined, several approaches can be conducted: QTL mapping in known pedigrees and controlled conditions, QTL association to traits in controlled conditions and *in-situ* studies of QTL association to the environment (Neale and Savolainen 2004).

All kinds of molecular tools of known heredity may provide information on the neutral drivers of genetic changes: genetic drift, mating system and dispersal. More interesting is to investigate the changes in these drivers, which generally requires transgenerational sampling. Genome-wide markers as well as candidate genes polymorphisms are now commonly used to detect signature of past selection events by testing a departure of the diversity pattern from the neutral expectation. Monitoring ongoing selection processes requires to relate directly the trait to the fitness or at least to a performance trait (selection gradient *b*). With phenotypic tools alone, trait to performance mapping can be assessed *ex-situ* with known pedigrees. When molecular and phenotypic tools are combined, *in-situ* selection gradient studies can be performed.

Conclusion

The adaptive capacity of tree populations is potentially huge, and silviculture can have a significant impact on the rate of phenotypic and genetic change per generation: the rate of change might probably be increased or reduced by a factor two depending on management interventions, which should not be neglected. The concept of evolution-oriented forestry that we introduced here does not pretend to allow for sufficient change in all cases. It should be considered as an option, with different associated benefits, risks and costs than those associated to the plantation strategy. Both strategies can also be combined. In any case, it is crucial to consider the potential evolutionary impact of silviculture when designing an adaptive forestry strategy. We proposed a simple framework to analyse and foresee the effects of forestry practice, and we identified a limited number of evolutionary processes and parameters that could be affected. Only few quantitative predictions can be made today, basically when evolutionary drivers are considered individually, and most expectations remain qualitative. Qualitative expectations can be used to draw research hypotheses. Quantitative predictions are needed to assess more precisely the cost effectiveness of forestry practice

TABLE 2

Objectives and requirements of genetic monitoring using molecular and phenotypic tools.

Monitoring objectives	Requirements		
	Molecular tools only	Phenotypic tools only	Combined approach
Quantification and characterization of the genetic diversity			
Quantify the global genetic diversity and characterize its organisation	Neutral markers	-	-
Decipher the genetic architecture of adaptive traits (QTLs, variances and correlations)	Known pedigrees, validated candidate genes polymorphisms	Known pedigrees, common garden experiments (*ex situ*)	QTL mapping or QTL association studies, *ex situ* experiments and *in situ* methods
Monitoring the drivers of genetic changes			
Monitor recent changes in demography and genetic drift intensity	Neutral markers, multi-generation sampling	-	-
Monitor the mating system and hybridization	Neutral markers, multi-cohort sampling	-	-
Characterize pollen and seed dispersal functions	Neutral markers, seed or/and seedling samples	Pollen and seed traps	Combined approach possible
Detect signatures of past selection (β)	Genome wide markers and candidate genes polymorphisms	-	-
Monitor current selection gradient (i,β)	Validated candidate genes polymorphisms, multi-generation sampling	Known pedigrees, path analysis traits to performance, *ex situ* experiments	Selection gradient studies, *in situ* methods

under various climate change scenarios. Quantitative comparison of evolution-oriented forestry with other baseline management options will require further process-based modelling and simulation studies for different forest types and species, different biotic and abiotic environments and different climate change scenarios. For a better understanding of the limits of the response to selection, we suggest to couple demogenetic models with biophysical models or host-parasite models.

One challenge for forestry decision-making under climate change is to reach a compromise between short-term and long-term objectives, *e.g.* speed-up the response to current selection pressure while preserving diversity and evolvability for uncertain future. A safe guideline is to favour natural selection for certainly adaptive traits, like drought resistance in the areas where more severe drought is expected, avoid random genetic erosion and increase genetic mixing. Over-selection for undue traits should also be avoided. As we have shown, each forestry practice has an effect on several evolutionary drivers (a geneticist would say pleiotropic effects) and interaction effects of different practices on a single evolutionary driver also exist (epistasis in genetic terms). To understand the global impact of forestry practice, in parallel to modelling approaches, long-term silvicultural options should be experimented. These new experiments will provide to the next generation of foresters very informative results, complementary to those obtained from comparative studies of existing situations.

Local decision should rely on a case by case approach taking into account each particular situation. Following quantitative genetics theory, we expect that in most cases genotypic diversity rather than allelic diversity constrains evolutionary changes. Can we obtain higher rate of change by driving recombination and selection within the current population? Do we need to introduce allochthonous pre-adapted genotypes to accelerate the emergence and spread of adequate allelic combinations? To address these questions and make a diagnostic, combined genophenotypic approaches *in situ* are very promising and should be further developed. Molecular tools initially developed for model species now become available for non-model species. Similarly, methodologies and knowledge about the genetic architecture of traits and phenotype construction (*e.g.* genetic and environmental correlations, age-age correlations) should be generalized to non-model species to help better understanding the gene to trait and gene to fitness mapping.

Acknowledgments

We particularly thank two anonymous reviewers and the editor for their relevant and useful comments on a previous version of the manuscript. The opinions expressed here benefitted from background experimental and modelling research funded by the EVOLTREE network of excellence (www.evoltree.eu), the ERA-Net Biodiversa LINKTREE project, the FRB VARIADAPT project, the ANR MACBI project, the Agropolis Foundation BIOFIS project as well as Région Provence Alpes Côte d'Azur and Département Santé des Forêts.

REFERENCES

Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J (2008) Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. Mol Ecol 17:5177–5188

Aitken SN, Yeaman S, Holliday JA, Wang T (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. Evol Appl 1:95-111

Alberto F, Aitken S, Alia A, Gonzáles-Martinez S, Hanninen H, Kremer A, Lefèvre F, Lenormand T, Yeaman S, Whetten R, Savolainen O (2013) Evolutionary response to climate change – evidence from tree populations. Global Change Biol (in press)

Allen CD, Macalady AK, Chenchouni H, Bachelet D, McDowell N, Vennetier M, Kitzberger T, Rigling A, Breshears D, Hogg EH(Ted), Gonzalez P, Fensham R, Zhang Z, Castro J, Demidova N, Lim JH, Allard G, Running SW, Semerci A, Cobb N (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. For Ecol Manag 259:660–684

Amm A, Pichot C, Dreyfus Ph, Davi H, Fady B (2012) Improving the estimation of landscape scale seed dispersal by integrating seedling recruitment. Ann For Sci 69:845-856

Aravanopoulos FA (2011) Genetic monitoring in natural perennial plant populations. Botany 89:75–81

Ashley M (2010) Plant parentage, pollination, and dispersal: how DNA microsatellites have altered the landscape. Crit Rev Plant Sci 29:148–161

Ayari A, Zubizarreta-Gerendiain A, Tome M, Tome J, Garchi S, Henchi B (2012) Stand, tree and crown variables affecting cone crop and seed yield of Aleppo pine forests in different bioclimatic regions of Tunisia. Forest Syst 21:128-140

Baradat P, Marpeau A, Bernard-Dagan C (1978) Variation of terpenes within and between populations of maritime pine. In: Rudin D (ed) Biochemical genetics of forest trees, Swedish University of Agricultural Sciences, Umea

Bontemps A (2012) Potentiel évolutif d'une population de hêtre commun sur le Mont Ventoux. Ph.D. dissertation University Paul Cézanne, Aix-Marseille III, Aix en Provence.

Bossdorf O, Richards CL, Pigliucci M (2008) Epigenetics for ecologists. Ecol Lett 11:106–115

Bouffier L, Charlot C, Raffin A, Rozenberg P, Kremer A (2008) Can wood density be efficiently selected at early stage in maritime pine (*Pinus pinaster Ait.*)? Ann For Sci 65:106p2-106p8

Bradshaw AD (1991) Genostasis and the limits to evolution. Philos T R Soc B 333:289–305

Brody AK (1997) Effects of pollinators, herbivores, and seed predators on flowering phenology. Ecology 78: 1624-1631

Brown JK (1975) Fire cycles and community dynamics in lodgepole pine forests. In: Baumgartner DM (ed) Symposium Proceedings: Management of lodgepole pine ecosystems. Washington State University Cooperative Extension Service, Pullman, pp430-456

Brown A, Young A, Burdon J, Christidis L, Clarke G, Coates D, Sherwin W (1997) Genetic indicators for state of the environment reporting, Australia: State of the Environment Technical Paper Series (Environmental indicators). Department of the Environment, Sport and Territories, Canberra

Buchert GP, Rajora OP, Hood JV, Dancik BP (1997) Effects of harvesting on genetic diversity in old-growth eastern white pine in Ontario, Canada. Conserv Biol 11:747-758

Burczyk J, Adams WT, Moran GF, Griffins AR (2002) Complex patterns of mating revealed in a *Eucalyptus regnans* seed orchard using allozyme markers and the neighbourhood model. Mol Ecol 11:2379-2391

Burczyk J, Adams WT, Birkes DS, Chybicki IJ (2006) Using genetic markers to directly estimate gene flow and reproductive success parameters in plants on the basis of naturally regenerated seedlings. Genetics 173:363–372

Burke KL (2011) The effects of logging and disease on American chestnut. For Ecol Manag 261:1027-1033

Cariveau D, Rebecca EI, Brody AK, Garcia-Mayeya LS, Ohe A (2004) Direct and indirect effects of pollinators and seed predators to selection on plant and floral traits. Oikos 104: 15-26

Carter AJR, Hermisson J, Hansen TF (2005) The role of epistatic gene interactions in the response to selection and the evolution of evolvability. Theor Popul Biol 68:179–196

Charmantier A, Garant D (2005) Environmental quality and evolutionary potential: lessons from wild populations. P Roy Soc B - Biol Sci 272:1415–1425

Chevin L-M, Collins S, Lefèvre F (2012) Phenotypic plasticity and evolutionary demographic responses to climate change: taking theory out to the field. Funct Ecol doi: 10.1111/j.1365-2435.2012.02043.x

Chevin LM, Lande R (2011) Adaptation to marginal habitats by evolution of increased phenotypic plasticity. J Evolution Biol, 24:1462–1476

Chevin L-M, Lande R, Mace GM (2010) Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. PLoS Biology 8:e1000357.

Couvet D, Ronfort J (1994) Mutation load depending on variance in reproductive success and mating system. In: Loeschcke V, Tomiuk J, Jain SK (eds.) Conservation genetics. Birkhauser Basel, pp. 55–68

Crawley MJ (1989). Insect herbivores and plant population dynamics. Annu Rev Entomol 34: 531–564

Dounavi K, Steiner W, Maurer W (2002) Effects of different silvicultural treatments on the genetic structure of European beech populations (*Fagus sylvatica* L.). Manag For Ecosyst 4:81-90

Durand-Gillmann M, Cailleret M, Boivin T, Nageleisen LM, Davi H (2012) Individual vulnerability of Silver fir (*Abies alba* Mill.) to parasitism by two contrasting biotic agents: mistletoe (*Viscum album* L. ssp. *abietis*) and bark beetles (*Coleoptera: Curculionidae: Scolytinae*) during a decline process. Ann For Sci doi:10.1007/s13595-012-0251-y

Eckert AJ, Bower AD, Gonzalez-Martinez SC, Wegrzyn J, Coop G, Neale DB (2010) Back to nature: ecological genomics of loblolly pine (*Pinus taeda, Pinaceae*). Mol Ecol 19:3789–3805

El-Kassaby YA, Benowicz A (2000) Effects of commercial thinning on genetic, plant species and structural diversity in second growth Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) stands. Forest Genetics 7:193-203

Endler JA (1986) Natural selection in the wild. Monographs in population biology 21. Princeton University Press, Princeton

Falconer DS (1960) Introduction to quantitative genetics. Oliver and Boyd, Edinburgh

Fenner M, Cresswell JE, Hurley RA, Baldwin T (2002) Relationship between capitulum size and predispersal seed predation by insect larvae in common *Asteraceae*. Oecologia 130: 72-77

Fernandez-M JF, Sork VL (2005) Mating patterns of a subdivided population of the Andean oak (*Quercus humboldtii* Bonpl., *Fagaceae*). J Hered 96:635–643

Forkner RE, Marquis RJ, Lill JT (2004) Feeny revisited: condensed tannins as anti-herbivore defences in leaf-chewing. Herbivore communities of *Quercus*. Ecol Entomol 29: 174–187

Forrester DI, Collopy JJ, Beadle, Christopher L, Warren CR, Baker TG (2012). Effect of thinning, pruning and nitrogen fertiliser application on transpiration, photosynthesis and water-use efficiency in a young *Eucalyptus nitens* plantation. For Ecol Manag 266:286-300

Futuyma DJ (2010) Evolutionary constraint and ecological consequences. Evolution 64:1865–1884

García C, Arroyo JM, Godoy JA, Jordano P (2005) Mating patterns, pollen dispersal, and the ecological maternal neighbourhood in a *Prunus mahaleb* L. population. Mol Ecol 14:1821-1830

García C, Jordano P, Arroyo JM, Godoy JA (2009) Maternal genetic correlations in the seed rain: effects of frugivore activity in heterogeneous landscapes. J Ecol 97:1424-1435

Gilpin ME, Soule ME (1986) Minimum viable populations: processes of species extinction. In: Soule ME (ed) Conservation biology: the science of scarcity and diversity. Sinauer Associates, Sunderland, pp. 19-34

Gingerich PD (2009) Rates of evolution. Annu Rev Ecol Evol S 40:657–675

Gomulkiewicz R Holt RD (1995) When does evolution by natural selection prevent extinction? Evolution 49:201-207

Goodell K, Elam DE, Nason JD, Ellstrand NC (1997) Gene flow among small populations of a self-incompatible plant: an interaction between demography and genetics, Am J Bot 84:1362-1371

Grivet D, Sebastiani F, Alía R, Bataillon T, Torre S, Zabal-Aguirre M, Vendramin GG, González-Martínez SC (2011) Molecular footprints of local adaptation in two Mediterranean conifers. Mol Biol Evol 28:101-116

Gwaze DP, bridgewater FE, Byram TD, Woolliams JA, Williams CG (2000) Predicting age-age correlations in tree-breeding programs: a case study in *Pinus taeda* L. Theor Appl Genet 100:199-206

Hannrup B, Ekberg (1998) Age-age correlations for tracheid length and wood density in *Pinus sylvestris*. Can J For Res 28:1373-1379

Hansen MM, Olivieri I, Waller DM, Nielsen EE, the GeM Working Group (2012) Monitoring adaptive genetic responses to environmental change. Mol Ecol 21:1311–1329

Hardy O, Gonzales-Martinez SC, Colas B, Fréville H, Mignot A, Olivieri I (2004) Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa (Asteraceae)*. II. Correlated paternity within and among sibships. Genetics 168:1601–1614

Hoffmann AA, Sgrò CM (2011) Climate change and evolutionary adaptation. Nature 470:479-485

Holliday JA, Ritland K, Aitken SN (2010) Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). New Phytol 188:501-514

Horvitz CC, Schemske DW (2002) Effects of plant size, leaf herbivory, local competition and fruit production on survival, growth and future reproduction of a neotropical herb. J Ecol 90: 279–290

Hulme PE (1998) Post-dispersal seed predation: consequences for plant demography and evolution. Perspect Plant Ecol 1: 32-46

Hussendörfer E, Konnert M (2000) Impact of forest management on genetic variation of Silver fir and European beech populations. For Landsc Res 75:187-204

IPCC (2007) Climate change 2007: AR4 synthesis report. An assessment of the Intergovernmental Panel on Climate Change, adopted at IPCC Plenary XXVII, Valencia, Spain, 12-17/11/2007

Jactel H, Petit RJ, Desprez-Loustau M-L, Delzon S, Piou D, Battisti A, Koricheva J (2012) Drought effects on damage by forest insects and pathogens: a meta-analysis. Global Change Biol 195: 267-276

Janzen DH (1971) Seed predation by animals. Annu Rev Ecol S 2: 465-492

Kircher M, Kelso J (2010) High-throughput DNA sequencing? Concepts and limitations. BioEssays 32:524-536

Klein EK, Carpentier FH, Oddou-Muratorio S (2011) Estimating the variance of male fecundity from genotypes of progeny arrays: evaluation of the Bayesian forward approach. Methods Ecol Evol 2:349–361

Klein EK, Lavigne C, Gouyon PH (2006) Mixing of propagules from discrete sources at long distance: comparing a dispersal tail to an exponential. BMC Ecology, 6:3

Kolb A, Ehrlén J, Eriksson O (2007) Ecological and evolutionary consequences of spatial and temporal variation in pre-dispersal seed predation. Perspect Plant Ecol 9:79–100

Konnert M, Hussendörfer E (2001) Genetic variation of silver fir (*Abies alba*) in unevenaged forests (Plenter forest) in comparison with evenaged forests (Altersklassenwald). Forest Sci 70:307-320

Koski V, Skrøppa T, Paule L, Wolf H, Turok J (1997) Technical guidelines for genetic conservation of Norway spruce (*Picea abies* (L.) Karst.). International Plant Genetic Resources Institute, Rome

Kramer K, Buiteveld J, Forstreuter M, Geburek T, Leonardi S, Menozzi P, Povillon F, Schelhaas MJ, Teissier du Cros E, Vendramin GG, van der Werf DC (2008). Bridging the gap between ecophysiological and genetic knowledge to assess the adaptive potential of European beech. Ecol Model 216:333–353

Kremer A, Le Corre V (2012) Decoupling of differentiation between traits and their underlying genes in response to divergent selection. Heredity 108:375-385

Kremer A, Ronce O, Robledo-Arnuncio JJ, Guillaume F, Bohrer G, Nathan R, Bridle JR, Gomulkiewicz R, Klein EK, Ritland K, Kuparinen A, Gerber S, Schueler S (2012) Long-distance gene flow and adaptation of forest trees to rapid climate change. Ecol Lett 15:378-392

Krouchi F , Derridj A, Lefèvre F (2004) Year and tree effect on reproductive organisation of *Cedrus atlantica* in a natural forest. For Ecol Manag 197:181-189

Kuparinen A, Savolainen O, Schurr FM (2010) Increased mortality can promote evolutionary adaptation of forest trees to climate change. For Ecol Manag 259:1003–1008

Lande R (2009) Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. J Evolution Biol 22:1435–1446

Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. Evolution 37:1210–1226

Lauber U, Rotach P, Hussendörfer E (1997) The influence of silvicultural interventions on the genetic structure of a young beech stand (*Fagus sylvatica* L.). Schweiz Z Forstwes 148:847-862

Lefèvre F (2004) Human impacts on forest genetic resources in the temperate zone: an updated review. For Ecol Manag 197:257-271

Lefèvre F, Kajba D (2001) Indicators for monitoring genetic diversity. In: Lefèvre F, Barsoum N, Heinze B, Kajba D, Rotach P, de Vries S, Turok J (eds) *In situ* conservation of *Populus nigra*. International Plant Genetic Resources Institute, Rome, pp 36-46

Legay M, Mortier F (2005) La forêt face au changement climatique: adapter la gestion forestière - Prise en compte dans les documents d'orientation de la gestion forestière. Synthèse de l'atelier ONF/INRA du 20 octobre 2005. Office National des Forêts, Paris. Les Dossiers Forestiers n°16

Le Rouzic A, Siegel PB, Carlborg O (2007) Phenotypic evolution from genetic polymorphisms in a radial network architecture. BMC Biol 5:50 doi:10.1186/1741-7007-5-50

Maron JL, Crone E (2006) Herbivory: effects on plant abundance, distribution and population growth. P Roy Soc B 273: 2575–2584

Marquis RJ (1992) The selective impact of herbivores. In: Fritz RS, Simms EL (eds) Plant resistance to herbivores and pathogens. University of Chicago Press, Chicago, pp301–325

Matheson AC, Wu HX, Spencer DJ, Raymond CA, Griffin AR (2002) Inbreeding in *Pinus radiata* III. The effect of inbreeding on age-age correlation and early selection efficiency. Silvae Genet 51:115-122

McCullough DG, Werner RA, Neumann D (1998) Fire and insects in northern and boreal ecosystems of North America. Annu Rev Entomol 43: 107-127

McDonald SE, Thomas BR, Cherniawsky DM, Purdy BG (2001) Managing genetic resources of lodgepole pine in west-central Alberta: patterns of isozyme variation in natural populations and effects of forest management. For Ecol Manag 152:45-58

McIntyre GA, Jacobi WR, Ramaley AW (1996) Factors affecting Cytospora canker occurrence on aspen. J Arboric 22: 229–233

Metzker ML (2009) Sequencing technologies - the next generation. Nat Rev Genet 11:31-46

Milleron M, Lopez de Heredia U, Lorenzo Z, Perea R, Dounavi A, Alonso J, Gil L, Nanos N (2012) Effect of canopy closure on pollen dispersal in a wind-pollinated species (*Fagus sylvatica* L.) Plant Ecol 213:1715-1728

Mimura M, Aitken SN (2007) Adaptive gradients and isolation-by-distance with postglacial migration in *Picea sitchensis*. Heredity 99:224–232

Misson L, Nicault A, Guiot J (2003) Effects of different thinning intensities on drought response in Norway spruce (*Picea abies* (L.) Karst.) For Ecol Manage 183:47-60

Moose SP, Dudley JW, Rocheford TR (2004) Maize selection passes the century mark: a unique resource for 21st century genomics. Trends Plant Sci 9:358-364

Mosseler A, Major JE, Rajora OP (2003) Old-growth red spruce forests as reservoirs of genetic diversity and reproductive fitness. Theor Appl Genet 106:151-161

Namkoong G, Boyle T, Gregorius HR, Joly H, Savolainen O, Ratnam W, Young A (1996) Testing criteria and indicators for assessing the sustainability of forest management: genetic criteria and indicators. Center for International Forestry Research CIFOR, Bogor. Working Paper N°10

Neale D, Savolainen O (2004) Association genetics of complex traits in conifers. Trends Plant Sci 9:325-330

Nebeker TE, Hodges JD, Blanche CA, Honea CR, Tisdale RA (1992) Variation in the constitutive defensive system of Loblolly pine in relation to bark beetle attack. Forest Sci 38: 457-466

Netherer S, Schopf A (2010) Potential effects of climate change on insect herbivores – general aspects and a specific example (Pine processionary moth, Thaumetopoea pityocampa). For Ecol Manag 259:831–838

Nilsson JE (1995) Genetic variation in the natural pollen cloud of *Pinus sylvestris*: a study based on progeny testing. Scand J Forest Res 10:140-148

Oddou-Muratorio S, Demesure-Musch B, Pélissier R, Gouyon PH (2004) Impacts of gene flow and logging history on the local genetic structure of a scattered tree species, *Sorbus torminalis* L. Crantz. Mol Ecol 13:3689–3702

Oddou-Muratorio S, Klein EK, Austerlitz F (2005) Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. II. Pollen dispersal and heterogeneity in mating success inferred from parent–offspring analysis. Mol Ecol 14:4441-4452

OFEFP, Rapport forestier (2005) Faits et chiffres sur l'état de la forêt suisse. Office fédéral de l'environnement, des forêts et du paysage, Berne

Osório J, Pereira JS (1994) Genotypic differences in water-use efficiency and ^{13}C discrimination in *Eucalyptus globulus*. Tree Physiol 14:871-882

Ott DS, Yanchuk AD, Huber DPW, Wallin KF (2011) Genetic variation of lodgepole pine, *Pinus contorta* var. *latifolia*, chemical and physical defenses that affect mountain pine beetle, *Dendroctonus ponderosae*, attack and tree mortality. J Chem Ecol 37: 1002-1012

Parker TJ, Clancy KM, Mathiasen RL (2006) Interactions among fire, insects and pathogens in coniferous forests of the interior western United States and Canada. Agri For Entomol 8:167-189

Pease CM, Lande R, Bull JJ (1989) A model of population growth, dispersal and evolution in a changing environment. Ecology 70:1657-1664

Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. Annu Rev Ecol Evol S 37:187–214.

Pichot C, Bastien C, Courbet F, Demesure-Musch B, Dreyfus P, Fady B, Frascaria-Lacoste N, Gerber S, Lefèvre F, Morand-Prieur ME, Oddou S, Teissier du Cros E, Valadon A (2006). Déterminants et conséquences de la qualité génétique des graines et semis lors de la phase initiale de régénération naturelle des peuplements forestiers. Les Actes du BRG n°6, Burreau des Ressources Génétiques, Paris, pp 277-297

Pigliucci M (2008) Is evolvability evolvable? Nat Rev Genet 9:75–82

Prat D, Faivre-Rampant P, Prado E (2006) Analyse du génome et gestion des ressources génétiques forestières. INRA ed Quae, Paris

Rajora OP, Rahman MH, Buchert GP, Dancik BP (2000) Microsatellite DNA analysis of genetic effects of harvesting in old-growth eastern white pine (*Pinus strobus*) in Ontario, Canada. Mol Ecol 9:339-348

Ramirez-Valiente JA, Valladares F, Huertas AD, Granados S, Aranda I (2011) Factors affecting cork oak growth under dry conditions: local adaptation and contrasting additive genetic variance within populations. Tree Genet Genomes 7:285-295

Rappaport NG, Mori S, Roques A (1993) Estimating effect of *Megastigmus spermotrophus (Hymenoptera: Torymidae)* on Douglas-fir seed production: the new paradigm. J Econ Entomol 86:845-849

Rehfeldt GE, Wykoff WR, Ying CC (2001) Physiologic plasticity, evolution, and impacts of a changing climate on. Climatic Change 50:355–376

Restoux G, Silva ED, Sagnard F, Torre F, Klein E, Fady B (2008) Life at the margin: the mating system of Mediterranean conifers. Web Ecology 8: 94–102

Ritland K (1996) A marker-based method for inferences about quantitative inheritance in natural population. Evolution 50:1062-1073

Robledo-Arnuncio JJ (2011) Wind pollination over mesoscale distances: an investigation with Scots pine. New Phytol 190:222-233

Robledo-Arnuncio JJ, Alía R, Gil L (2004) Increased selfing and correlated paternity in a small population of a predominantly outcrossing conifer, *Pinus sylvestris*. Mol Ecol 13:2567–2577

Rockman MV (2012) The QTN program and the alleles that matter for evolution: all that's gold does not glitter. Evolution 66:1–17

Roques A (1981) Biologie et répartition de *Megastigmus spermotrophus* Wachtl. (*Hymenoptera: Chalcidoidea Torymidae*) et des autres insectes liés aux cônes dans les peuplements forestiers et vergers à graines français de sapin de Douglas *Pseudotsuga menziesii* [Mirb.] Franco. Acta Oecol 2:161-180

Rose KE, Louda SM, Rees M (2005) Demographic and evolutionary impacts of native and invasive insect herbivores on *Cirsium canescens*. Ecology 86: 453–465

Rouault G, Candau JN, Lieutier F, Nageleisen L-M, Martin JC, Warzée N (2006) Effects of drought and heat on forest

insect populations in relation to the 2003 drought in Western Europe. Ann For Sci 63: 613-624

Sáenz-Romero C, Guzmán-Reyna R, Rehfeldt G (2006) Altitudinal genetic variation among *Pinus oocarpa* populations in Michoacan, Mexico. Implications for seed zoning, conservation, tree breeding and global warming. For Ecol Manag 229:340-350

Sagnard F, Oddou-Muratorio S, Pichot C, Vendramin GG, Fady B (2011) Effects of seed dispersal, adult tree and seedling density on the spatial genetic structure of regeneration at fine temporal and spatial scales. Tree Genet Genomes 7:37-48

Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene flow and local adaptation in trees. Annu Rev Ecol Evol S 38:595–619

Schleuning M, Farwig N, Peters MK, Bergsdorf T, Bleher B, Brandl R, Dalitz H, Fischer G, Freund W, Gikungu MW, Hagen M, Garcia FH, Kagezi GH, Kaib M, Kraemer M, Lung T, Naumann CM, Schaab G, Templin M, Uster D, Wagele JW, Bohning-Gaese K (2011) Forest fragmentation and selective logging have inconsistent effects on multiple animal-mediated ecosystem processes in a tropical forest. PloS ONE 6:e27785

Schoen D, Brown HD (1991) Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. P Natl Acad Sci USA 88:4494-4497

Simard M, Romme WH, Griffin JM, Turner MG (2011) Do mountain pine beetle outbreaks change the probability of active crown fire in lodgepole pine forests? Ecol Monogr 81: 3-24

Skrøppa T, Tollefsrud M, Sperisen C, Johnsen Ø (2010). Rapid change in adaptive performance from one generation to the next in *Picea abies* - central European trees in a Nordic environment. Tree Genet Genomes 6:93–99

Smouse P, Dryer RJ, Westfall RD, Sork VL (2001) Two-generation analysis of pollen flow across a landscape. I. Male gamete heterogeneity among females. Evolution 55:260-271

Smouse P, Sork VL (2004) Measuring pollen flow in forest trees: an exposition of alternative approaches, For Ecol Manag 197:21–38

St Clair JB, Howe GT (2007) Genetic maladaptation of coastal Douglas-fir seedlings to future climates. Global Change Biol 13:1441–1454

Strauss SY, Whittall JB (2006) Non-pollinator agents of selection on floral traits. In: Harder LD, Barrett SCH (eds.), Ecology and evolution of flowers. Oxford University Press, Oxford, pp. 120-138

Tene A, Tobin B, Dyckmans J, Ray D, Black K, Nieuwenhuis M (2011) Assessment of tree response to drought: validation of a methodology to identify and test proxies for monitoring past environmental changes in trees. Tree Physiol 31:309-322

Thiebaut B, Comps B, Leroux A (1992) Relation hauteur-génotype dans une régénération naturelle de hêtre (*Fagus sylvatica*) équienne et agée de 18 ans. Ann Sci For 49:321-335

Walsh B, Blows MW (2009) Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. Annu Rev Ecol Evol S 40: 41–59

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97-158

Wu HX, Powell MB, Yang JL, Ivkovic M, McRae TA (2007) Efficiency of early selection for rotation-aged wood quality traits in radiata pine. Ann For Sci 64:1-9

Yan H, Bi HQ, Li R, Eldridge R, Wu Z, Li Y, Simpson J (2006) Assessing climatic suitability of *Pinus radiata* (D. Don) for summer rainfall environment of southwest China. For Ecol Manag 234:199-208

Yarnes C, Boecklen WJ (2005) Abiotic factors promote plant heterogeneity and influence herbivore performance and mortality in Gambel's Oak (*Quercus gambelii,* Nutt.). Entomol Exp Appl 114:87–95

Yousefpour R, Jacobsen JB, Thorsen BJ, Meilby H, Hanewinkel M, Oehler K (2012) A review of decision-making approaches to handle uncertainty and risk in adaptive forest management under climate change. Ann For Sci 69:1-15

THE TRANSCIPTOME OF *POPULUS* IN ELEVATED CO_2 REVEALS INCREASED ANTHOCYANIN BIOSYNTHESIS DURING DELAYED AUTUMNAL SENESCENCE

M.J. Tallis[1], Y Lin[1], A. Rogers[2], J. Zhang[1,3], N.R. Street[1], F. Miglietta[4], D.F. Karnosky[5], P. De Angelis[6] and G. Taylor[1*]

[1] School of Biological Science, Bassett Crescent East, University of Southampton, SO16 7PX.
[2] Environmental Sciences Department, Brookhaven National Laboratory, Upton NY 11973, USA.
[3] College of Life Sciences, Peking University, Beijing, 100871, P.R.China
[4] Institute of Biometeorology-CNR, Via Caproni, 8 50145 Firenze, Italy.
[5] School of Forest Resources & Environmental Science, Michigan Technological University, 1400 Townsend Drive, Houghton, Michigan 49931-1295 USA.
[6] Department of Forest Environment and Resources (DISAFRI), University of Tuscia, Via S. Camillo De Lellis, 01100 Viterbo, Italy.

Present addresses:
NR Street: *Umeå Plant Science Centre, Department of Plant Physiology, University of Umeå, SE-901 87 Umeå, Sweden.*

Corresponding authors:
Professor Gail Taylor
School of Biological Sciences, Bassett Crescent East,
University of Southampton, SO16 7PX
g.taylor@soton.ac.uk

Key words: Populus, autumnal senescence, cDNA microarray, anthocyanin biosynthesis, elevated CO2, LDOX.

Published in *New Phytologist* (2010) 186: 415-428, with doi:10.1111/j.1469-8137.2010.03184.x

Summary

• Over recent decades the observed delay in autumnal senescence has been linked to rising temperatures. Here we suggest that increasing atmospheric CO_2 alone may partly account for extensions of the growing season and for the first time, through transcriptome analysis, identify gene expression changes associated with this delayed senescence.

• Using a plantation of *Populus x euramericana* grown in elevated [CO_2] (e[CO_2]) using Free-air CO_2 Enrichment (FACE) technology, we investigated the molecular and biochemical basis underlying this response. Leaf biosynthetic pathways influenced by e[CO_2] during senescence were identified using a *Populus* cDNA microarray and an analysis of expression changes for genes representing multiple biochemical pathways, RT-qPCR, and leaf biochemical assays.

• Both pathways for secondary metabolism and for the glycolytic pathway were significantly up-regulated by e[CO_2] during senescence. Within these pathways the two most significantly up-regulated ESTs in e[CO_2] represented transcripts for and *LDOX* and *DFR,* which regulate anthocyanin biosynthesis, with a normalised expression (e[CO_2] / a[CO_2]) of 39.6 and 19.3 respectively.

• We propose that in e[CO_2] there is an association between increased autumnal sugar accumulation and changes in gene expression in genes determining anthocyanin biosynthesis. This prolongs leaf longevity during natural autumnal senescence through improved stress tolerance.

Introduction

Evidence from phenological records suggests that recent global warming is leading to longer growing seasons. An analysis of over 1700 species showed significant shifts in plant phenology (Parmesan & Yohe, 2003) including extension of the growing season (Menzel & Fabian, 1999; Menzel *et al.* 2006; Myneni *et al.* 1997 and Zhou *et al.* 2001) which has been attributed to rising air temperature (Menzel & Fabian, 1999). On average over the past 35 years autumnal senescence has been delayed across Europe by 1.3 days decade^{-1} (Menzel *et al.* 2006). However, whilst a strong correlation exists between atmospheric warming and an earlier spring phenophase, the correlation between warming and a later on-set of the autumn phenophase is very weak (Menzel *et al.* 2006). Understanding this process is important since changing phenology can alter bio-geochemical cycling and albedo both feeding back on climate change (Peñuelas, Rutishauser, and Filella, 2009). For example, an extended autumn has been reported to increase carbon storage in the boreal zone of Northern latitude forests (Lucht *et al.* 2002) and in the aspen boreal forests of North America (Chen *et al.* 1999)

Over the time period of the Menzel *et al.* (2006) study (1971 -2000) atmospheric carbon dioxide has increased by 44 µmol mol-1 (13.5 %). We have shown previously that elevated atmospheric CO_2 (e[CO_2]) delays autumnal senescence in a forest canopy exposed for six years to e[CO_2] in a free air CO_2 exposure (FACE) ecosystem experiment. At the canopy level the decline in greenness (NDVI) and leaf area index (LAI) were both significantly reduced by e[CO_2]. Also significantly reduced was the decline in leaf chlorophyll indicating delayed senescence in these trees (Taylor *et al.* 2008). However, these finding are controversial since rising [CO_2] has been shown to shorten (Sigurdsson, 2001), extend (Li *et al.* 2000; Rae *et al.* 2005; Taylor *et al.* 2008) or have no effect (Herrick & Thomas, 2003) on forest senescence.

Natural autumnal senescence is regulated by day length, temperature, light, nitrogen supply and water supply and by plant carbon-nitrogen and source-sink balance (Wingler *et al.* 2006). The timing of which can be regarded as the result of a trade-off between the conflicting

requirements for optimizing the nitrogen and carbon status of the plant (Keskitalo *et al.* 2005). The strength of the plants sink for photosynthate can positively influence photosynthetic responses to e[CO_2] (Bryant, Taylor and Frehner, 1999; Ainsworth *et al.* 2004) and also reduce the rate of senescence (Wingler *et al.* 2004; Kaschuk *et al.* 2009). Recent studies using girdled sugar maple trees have shown that sugar accumulation in leaves resulted in the formation of anthocyanins (Murakami, Schaberg and Shane, 2008) and leaves with increased anthocyanin content were associated with a delayed senescence (Schaberg *et al.* 2008). In poplar over expressing an Arabidopsis sucrose phosphate synthase gene resulted in increased leaf sucrose content between August and throughout senescence which was associated with a delayed senescence (Park *et al.* 2009). These data indicate the complex interactions between the plants developmental state, source sink balance and rate of senescence. Nevertheless, the initiation and sequence of events during senescence are well conserved. The stimulus for the process of autumnal senescence in *Populus* is a shortening of the photoperiod initiating bud-set, at least for high-latitude trees (Böhlenius *et al.* 2006; Fracheboud *et al.* 2009; Keskitalo *et al.* 2005; Olsen *et al.* 1997) which is considered an adaptive trait related to plant fitness (Ingvarsson *et al.* 2006). Following this initial stimulus a well characterised sequence of cellular events has been reported (Keskitalo *et al.* 2005) from chloroplast breakdown, carotenoid and soluble sugar loss, to anthocynain production, a massive 80 % nitrogen remobilisation and leaf abscission. During this process in poplar 166 genes were classed as the most up-regulated revealing a shift from gene expression associated with anabolism to that of catabolism and an increased role of the mitochondria for energy generation as photosynthesis breaks down (Andersson *et al.* 2004).

The aim of this research was to understand how exposure to increased atmospheric CO_2 disrupts the process of autumnal senescence and to identify key changes in metabolism and gene expression associated with delayed senescence. We conducted our investigation at the POP/EUROFACE (Free Air CO_2 Enrichment) experiment (Miglietta *et al.* 2001) where trees had been grown for six years, since planting to canopy closure, in a fully open air environment at either ambient [CO_2] (a[CO_2]) or e[CO_2] (550 µmol mol-1). Conducting this study at a FACE site allowed us to eliminate the potentially confounding problems of sink, N and water limitation that are common in experiments using other CO_2 fumigation techniques (McLeod & Long, 1999) which are known to influence the rate of senescence. The highly productive, fast growing trees reported to be non-resource limited (Liberloo *et al.* 2009) at the POP/EUROFACE experiment provided the ideal model system in which to investigate the changes in natural autumnal senescence of a forest canopy growing in e[CO_2]. Furthermore *Populus* is now recognised as a model tree genus (Taylor, 2002; Tuskan *et al.* 2006; Jansson & Douglas, 2007) enabling genomic resources to be deployed to answer questions of ecological and evolutionary significance on plant response and adaptation to climate change.

Materials and Methods

The POP/EUROFACE site

The PopFACE experiment (9 ha) was situated on a nutrient rich, clay soil in Tuscania, Italy (42° 22′ N, 11° 48′ E; altitude 150 m a.s.l.; www.unitus.it/euroface). Three species of *Populus, P. alba* L. (clone 2AS-11), *P. nigra* L. (clone Jean Pourtet), and *P.* x *euramericana* (Dode) Guinier (clone I-214) were grown in the experiment within six experimental plots. The whole site was assigned to three blocks each containing two 314 m^2 octagonal plots assigned two treatments of [CO_2] (control or ambient CO_2, a[CO_2] and eleveated CO_2 e[CO_2] of 550 µmol mol^{-1}). Complete design characteristics of PopFACE are available elsewhere (Miglietta, *et al.* 2001; Scarascia-Mugnozza *et al.* 2006). During the period of this study trees had been planted for six years, coppiced after three years and a closed canopy was evident. Canopy characterisation and climatic data during this study have been described in detail (Taylor *et al.* 2008) and only the ambient nitrogen sub-plots were used in this study, the same treatments as Taylor *et al.* (2008). A strong chlorosis of the canopies in plots five and six was evident during this study so these plots were discounted from any further analysis as described in Liberloo *et al.* (2007). Day time CO_2 enrichment was provided from bud burst until bud-set except during this study when CO_2 enrichment was continued throughout. The e[CO_2] measured at 1 min intervals was within a 20% deviation from the target concentration of 550 ppm for 94% of the time during the first three year rotation, and for 78% of the time during the second rotation (Liberloo *et al.* 2009). The leaf sampling regime is described in the Supporting Information and all sampled leaves were instantly placed in foil, added to a weighted bag and dropped from the canopy to be placed in liquid N2, from removal until placing in liquid N2 was ~ 10 s.

Canopy level spectral reflectance

Canopy reflectance was measured with a field portable spectroradiometer (GER 3700) (GER, Buffalo, NY, USA. Mod. 3700), and a chlorophyll specific NDVI (Gitelson and Merzlyak, 1994; Gamon and Surfus, 1999) was calculated as described in Taylor *et al.* (2008). Further details are given in Supporting Information.

Microarray Hybridisation

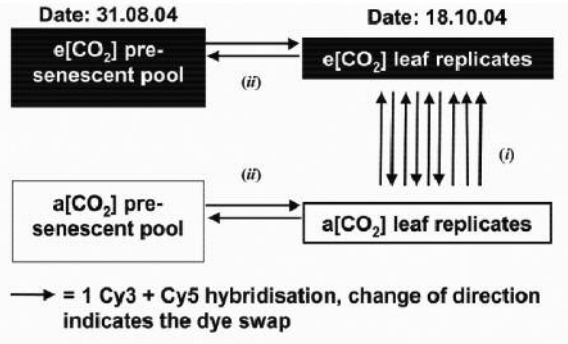
Total RNA was extracted using the protocol of Chang *et al.* (1993) with the following modifications, 2.67 % β-Mercapto-ethanol was added to the CTAB extraction buffer instead of spermidine. After the overnight 4 ºC LiCl precipitation and re-suspension in SSTE two additional

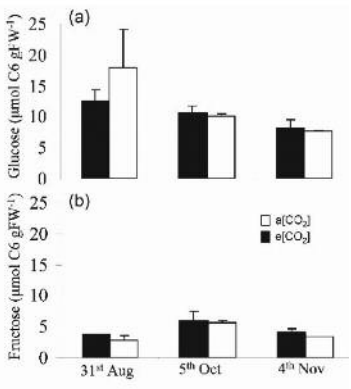
CHISAM (chloroform:isoamylalcohol, 24:1) extractions were performed. The RNA pellet was re-suspended in 20 µl of DEPC treated H_2O. RNA quantity was checked (NanoDrop® ND-1000 Spectrophotometer, Wilmington, DE, USA.) and RNA quality checked (Agilent 2100 bioanaylser, Agilent Technologies, USA). 100 µg of total RNA was denatured (65 º for 10 minutes) with 2 µl of anchored oligo$(dt)_{20}$ primer then chilled on ice (max. 1 minute). A reverse transcription master mix was prepared. This consisted of 6 µl 5x -RT-buffer (first strand buffer), 3 µl of 10mM DTT, 1 µl of 50x dNTP mix (a mix of dA, dC and dG, and aa-dUTP and dTTP in a ratio of 4:1 aa-dUTP:dTTP) (Amersham UK; except aa-dUTP, Sigma UK), 1 µl RNase inhibitor and 2 µl of Superscript™ reverse transcriptase. The master mix was made in bulk such that 13 µl was added to each oligo $(dt)_{20}$ primed RNA sample for overnight cDNA synthesis at 48 ºC. After overnight synthesis the reverse transcription reaction was inhibited by addition of 10 µl 0.5 M EDTA and any remaining RNA degraded by the addition of 10 µl 1 M NaOH and heating at 65 ºC for 15 minutes. The remaining cDNA was then neutralised with 50 µl of 1M HEPES (pH 7.5). The cDNA purification was carried out according to the manufacturer's instructions (Qiagen PCR purification kit, Qiagen, Crawley, UK) with the following exceptions. A phosphate-ethanol wash buffer (PWB: pH 8.0, 5 mM KPO4) was used instead of buffer PE and two PWB steps were included. cDNA was then eluted via two elutions each with 30 µl of 0.1 M $NaHCO_3$ (pH 9.0), 1 µl of cDNA (60 µl total) was then taken for spectrometric quantification. The purified cDNA (59 µl total) was taken and 35 µl 100 mM sodium acetate (pH 5.2) added. Under minimal light, purified cDNA was added to an aliquot of CyDye™ ester (Amersham, Buckinghamshire, UK). Cy3 and Cy5 were added to control and treatment respectively, and for nearly 50 % of the samples this orientation was reversed to account for any dye binding bias. The samples were gently agitated and then left in the dark at room temperature for 2 hours. Following a dye coupled cDNA purification step (Qiagen protocol except an additional buffer PE wash step was included and two repeated elution steps were carried out) the labelled samples were randomly paired between control and treated samples. The total elute containing 200 µl of Cy3-and-Cy5 coupled cDNA was concentrated down to 25 µl in a spin concentrator (Eppendorf Concentrator 5301, Eppendorf, Cambridge, UK). The dye-labelled cDNA target (25 µl) was denatured by the addition of 50 µl formamide, 25 µl of hybridisation buffer (Amersham, UK) was added and the sample heated at 95 ºC for 1 minute and then chilled on ice.

Microarrays slides were purchased from PICME (www.picme.at). The 26,915 ESTs spotted on glass slides were produced by INRA-Nancy (Rinaldi *et al.* 2007), INRA-Orleans (Dejardin *et al.* 2004), and University of Helsinki (Brosche *et al.* 2005) within the framework of the LIGNOME and ESTABLISH programme respectively. This was estimated to represent approximately 10, 000 predicted gene models in the *P. trichocarpa* genome sequence (Rinaldi *et al.* 2007). Full MIAME-compliant details of the array content and production can be found at www.picme.at. An overview of the experimental design is illustrated in Figure 1. In approach (i), a direct comparisons between replicate senescent samples (18th October) exposed to either treated (elevated CO_2 e[CO_2]) or control ((ambient CO_2 (a[CO_2])) was conducted. Of the 12 leaves sampled per treatment on 18 October, RNA of sufficient quality was obtained from nine leaves per treatment. These nine samples were randomly paired so that each CO_2 treatment was represented in each pair. In approach (ii), comparisons between pre-senescent (31st August) and senescent material were undertaken using a common pre-senescent reference pool. For the progress of senescence in a[CO_2] the a[CO_2] reference pool was used and this was designated a[CO_2] senescence and in e[CO_2] the e[CO_2] reference pool was used and

FIGURE 1

A schematic representation of the microarray hybridisation design used. *(i)* Nine replicate hybridisations are shown for probes obtained from leaves sampled on 18th October 2004 with each CO_2 treatment represented in each hybridisation. *(ii)* Two hybridisations are shown for senescence within each CO_2 treatment. The pre-senescent probe for each treatment was created by pooling probes from seven replicate leaves sampled from within the respective CO_2 environment.





designated e[CO_2] senescence. Microarray hybridisations were carried out to directly assess the transcriptome changes occurring during senescence in a[CO_2] and in e[CO_2]. RNA was extracted individually from pre-senescent (31 August 2004) leaves and then pooled. The pool for e[CO_2] consisted of seven samples (four from plot one and three from plot four). The pool for a[CO_2] consisted of seven samples (four from plot two and three from plot three). In each pool 100 µg of total RNA was established using 14.3 µg of total RNA from each replicate leaf.

Pre-hybridisation, hybridisation and scanning of the PICME microarray slides are described in Street *et al.* (2009). The transcript profile data were also analysed according to Street *et al.*(2009). Linear models with *B* statistics implemented in the LIMMA (Smyth, 2004: http://bioinf.wehi.edu.au/limma/) package for the statistical software 'R' (http://www.r-project.org) were used to identify ESTs that may be differentially expressed (Diaz *et al.* 2003). The model contained only one specified factor for treatment either (i) e[CO_2] for the 18th October direct comparison between exposure to a[CO_2] and e[CO_2] or (ii) expression on 18th October for the comparison between the pre-senescent reference pools and the senescent material. Calculated B and *P* values are adjusted for multiple testing with a false discover rate (FDR) of 0.05 (Benjamini & Hochberg, 1995) considered a very conservative statistical analysis for FACE experiments of field acclimated material (Leakey *et al.* 2009a). As the influence of senescence on transcript abundance can be very large for some transcripts (Andersson *et al.* 2004) this conservative statistical analysis was considered appropriate. ESTs considered as significantly differentially expressed in e[CO_2] compared to a[CO_2] late in senescence (18 October 2004) and those differentially expressed between pre-senescent and senescent tissues were those with a B value of ≥ 3. A B value of zero equals a 50:50 probability of differential expression where as a B value of 3 represents approximately 95% certainty of differential expression (exp[3] / (1+exp[3])) = 0.95, or 95 %). We used a B value of ≥ 3 and a two-fold change in mean normalized expression as our threshold for declaring an EST as significantly differentially expressed. Sequence annotation was obtained using the tblastx algorithm run by the DOE Joint Genome Institute (JGI, http://genome.jgi-psf.org). All microarray data generated have been deposited in the Gene Expression Omnibus (GEO) database as Series GSE15874.

Analysis of expression changes for genes representing multiple biochemical pathways

Microarray data was further analysed by MapMan (version 2.2.0, downloaded from (http://gabi.rzpd.de/projects/MapMan/). Mapman software displays large data sets onto diagrams of metabolic pathways or other processes (Thimm *et al.* 2004). *Arabidopsis thaliana* ortholog ids were obtained using the gene model ids of each EST sequence on the PICME array and the ortholog extractor function in PopGenIE (Sjödin *et al.* 2009). The mean Log_2 ratio for all the *Populus* ESTs representing a single ortholog Arabidopsis gene model were used in the Mapman pathway analysis. The Wilcoxon Rank Sum Test was used within MapMan to identify any functional group of genes that exhibit a different behaviour in terms of expression profile compared to all the other remaining functional groups. Data were Benjamini Hochberg corrected in MapMan and $P \leq 0.05$ was considered the cut-off for identifying functional groups considered to have a different behaviour in terms of expression profiles. The pathway diagrams for anthocyanin biosynthesis were based on that described for Arabidopsis (Solfanelli *et al.* 2006) and additional information for Populus were obtained from Tsai *et al.* (2006).

Real-time quantitative PCR

The selection and validation of the internal reference gene is described in Supporting Information. ESTs to be validated were searched by their EST name in the PICME database and the EST information was extracted from NCBI using Accession Number. Gene model information was downloaded from JGI. Real-time qPCR primers were designed using Beacon Design 5.0 (PREMIER Biosoft International) and the following criteria: Tm of 55-60°C and PCR amplicon lengths of 115 to 160 bp, yielding primer sequences with lengths of 19 to 22 nucleotides and GC contents of 45% to 55%. Primers were also designed to amplify close to the 3' end of the transcripts or EST, and at least one primer of a pair was designed to cover an exon-exon junction if possible. All primers used in this study were synthesized and desalted by Sigma-Genosys.

The protocols for cDNA synthesis and SYBR Green qPCR are described in supplementary materials and methods. The primer pairs used and mean ct values of the reference gene are also reported in supplementary materials and methods.

Leaf Biochemistry

Anthocyanin content

Frozen leaves were ground and 50 mg from each sample was used for analysis according to the method of Martin *et al.* (2002). Data were calculated from the mean of three technical repeats for eight replicate leaves per treatment (four per plot).

Soluble carbohydrates and starch

Extraction and measurement of glucose, fructose, sucrose and starch content has been described previously (Rogers *et al.* 2006). Glucose, fructose and sucrose were extracted from frozen ground material using sequential incubations in ethanol. Starch was extracted from the residual material and converted to glucose. Glucose, fructose, sucrose and the glucose resulting from the starch degradation were then assayed using a continuous enzymatic substrate assay.

Statistical analysis of leaf biochemistry and spectral reflectance

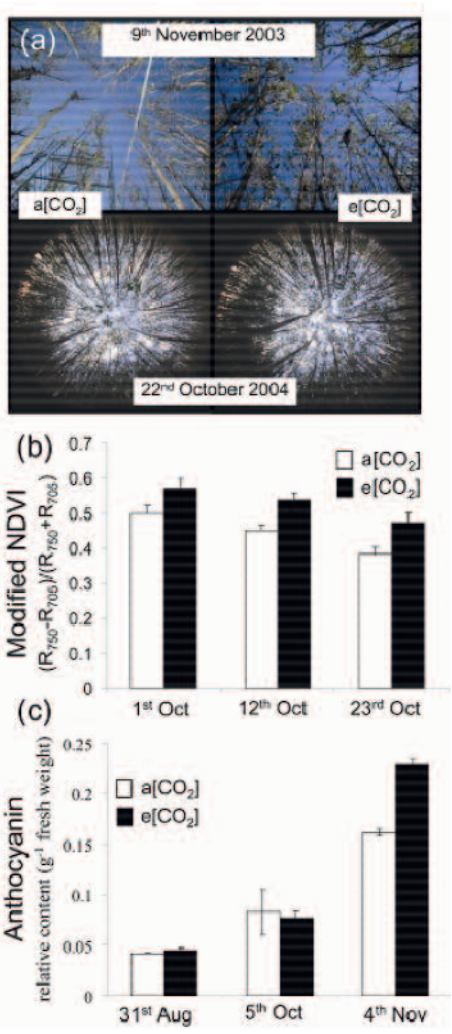
Responses of leaf biochemistry during the progression of autumnal senescence were subjected to general linear model ANOVA (using Minitab® 15.1.0.0. Minitab Inc., Philadelphia), with the model: Response = Block' | CO_2 and leaf was considered the unit of replication. No significant Block * CO_2 interactions were detected. Responses of canopy spectral reflectance during the progression of autumnal senescence were analysed by a two-way ANOVA with repeated measures in SPSS (SPSS 16.0 for windows) with the model: Response = time | CO_2

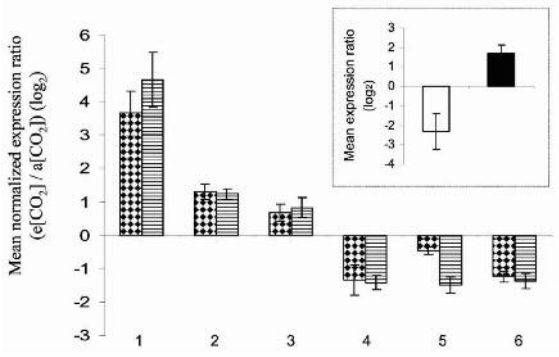
Results and Discussion

The trees used in this study exhibit both a net increase in photosynthesis (Tricker *et al.* 2005; Liberloo *et al.* 2007) increased biomass production (Liberloo *et al.* 2009) a significantly reduced leaf N (on a leaf mass basis) (Liberloo *et al.* 2007) and were considered not to be resource limited (Liberloo *et al.* 2009). We have previously established that the field grown poplar trees in this study, exhibit delayed autumnal senescence in response to growth at e[CO_2] (Taylor *et al.* 2008). Figure 2a shows digital photography examples of one a[CO_2] and one e[CO_2] plot for two years and a detailed analysis of the LAI for all experimental plot canopies is given in Taylor *et al.* (2008) where a significant delay in the decline of LAI was reported in e[CO_2]. A canopy modified NDVI (a chlorophyll specific NDVI) (Gitelson and Merzlyak, 1994; Gamon and Surfus, 1999) used to estimate changes in whole canopy chlorophyll content during October 2004 is reported in Figure 2b. Modified NDVI significantly declined with time ($F_{2,6}$ = 115.2 $P \leq 0.001$, as expected during senescence). At all time points modified NDVI was significantly greater in e[CO_2] ($F_{1,3}$ = 104.1 $p \leq 0.01$) indicating that during this time the canopy contained more chlorophyll in e[CO_2] than in a[CO_2] this is in agreement with extracted leaf chlorophyll content reported in Taylor *et al.* (2008) there was no significant interaction with CO_2 treatment and time. A more detailed report addressing canopy changes during senescence in e[CO_2] is given in Taylor *et al.* 2008. This study aims to provide the first snap-shot of what may be occurring within the metabolism of the plants during senescence in e[CO_2]. In order to do this the focus was initially on genes that exhibited a statistically significant change in regulation in e[CO_2] during senescence as evidenced from the microarray EST data. This data identified a significant up-regulation of key transcripts in anthocyanin biosynthesis. Additional to this individual EST data, functional groups of genes were examined which together exhibited a significant change in regulation. These data from an e[CO_2], a[CO_2] direct comparison late in senescence (18th Oct) are presented in Figure 3. Changes in the leaf transcritome between non-senescent and senescent leaf material in either e[CO_2] or a[CO_2] were

FIGURE 2

The influence of e[CO_2] on canopy chlorophyll content and leaf anthocyanin content during senescence. (a) Digital images of an a[CO_2] and an e[CO_2] plot during senescence and on two separate years. (b) The chlorophyll specific NDVI (modified NDVI) measured over the e[CO_2] (black bars) and a[CO_2] (open bars) canopies at three time-points through October 2004 mean (+/- SE; *n* = 4). (c) Extractable leaf anthocyanin content in leaves senescing in either e[CO2] (black bars) or a[CO_2] (open bars), mean (+/- SE; *n* = 8).





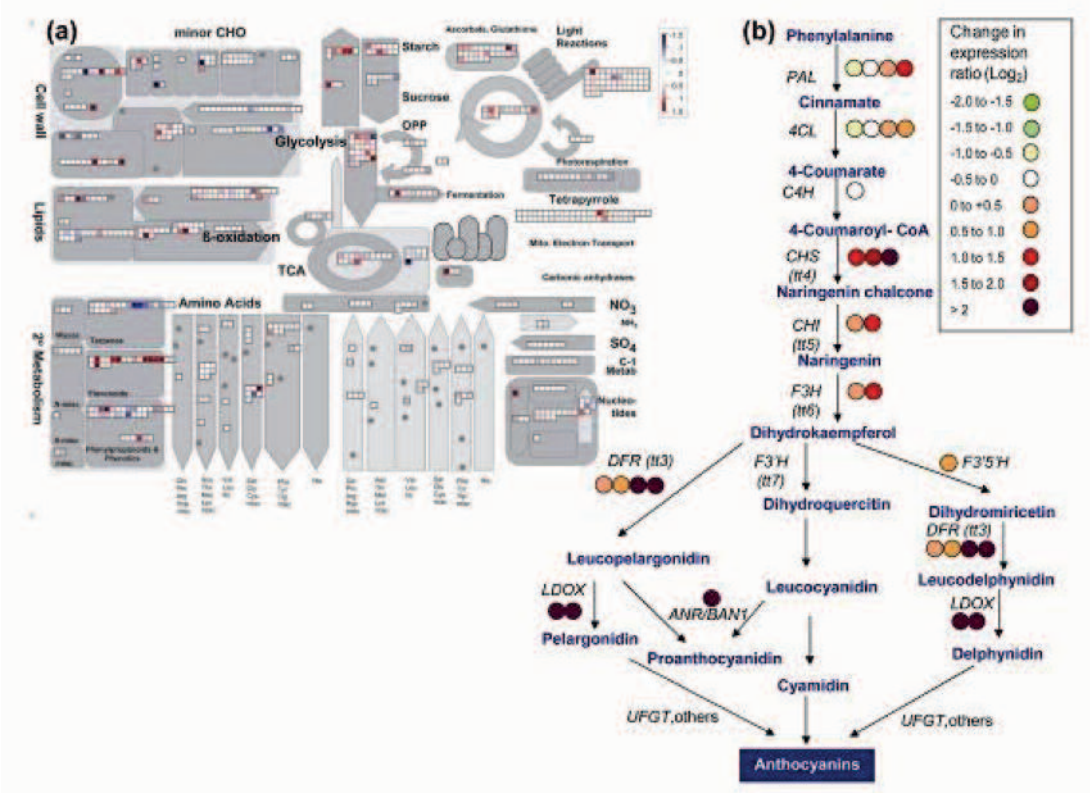
also examined and a change in leaf metabolism was identified. These data representing the changes between non-senescent and senescent material are presented in Figures 4 and 5. Leaf anthocyanin content and soluble and insoluble carbohydrate contents were also measured to support the data identified from the microarrays.

Identifying the genes most significantly influenced by e[CO2] late in autumnal senescence

The effect of senescence in e[CO_2] on transcript abundance, represented by ~ 23, 000 ESTs accounting for approximately 10,000 predicted gene models in the P. trichocarpa genome sequence (Rinaldi *et al.* 2007) was tested using the PICME microarray (http://www.picme.at/). Differential gene expression between e[CO_2] and a[CO_2] grown leaves was assessed on 18th October during the latter phase of senescence. Canopy Leaf Area Index (LAI) at this time had declined by 43 % and 37 % from 31st August in a[CO_2] and e[CO_2] treatments respectively (data not shown). Using nine replicate microarray hybridisations each of ~ 23,000 ESTs then by random chance ~ 1150 ESTs would be classed as significantly differentially expressed within each microarray ($P \leq 0.05$) using a traditional ANOVA framework. In order to account for this a robust statistical analysis was applied to identify the most consistently and significantly differentially expressed ESTs during the heterogeneous process of canopy senescence. Bayesian statistics were used with a FDR of 0.05 considered by some as very stringent for such FACE experiments (Leaky

FIGURE 3

Differential gene expression involved in (a) metabolism and (b) anthocyanin biosynthesis from replicate microarray data for probes prepared from leaf samples taken late in senescence (18th October 2004). (a) The differences in transcript abundance for Arabidopsis othologs of the *Populus* gene models. Each coloured square represents the mean Log_2 expression data for the Arabidopsis ortholog calculated from the *Populus* EST expression data. The resulting file was loaded into the MapMan Image Annotator module to generate the metabolism overview map. The logarithmic colour scale bar ranging from - 1.5 (dark blue, representing a three-fold down-regulated gene in e[CO_2]) to + 1.5 (dark red, representing a three-fold up-regulated gene in e[CO_2]) is given.. (b) Log_2 transcript abundance data for Populus gene models involved in anthocyanin biosynthesis, from the nine replicate microarray hybridizations as a function of (e[CO_2] / a[CO_2]) expression. Genes coding for enzymes in this pathway were identified using the *Populus* EST sequence data and annotation and obtaining the Arabidopsis ortholog gene model. Coloured circles represent *Populus* gene models predicted to code for enzymes involved in each metabolic step. Pathway diagrams were constructed using Solfanelli *et al.* (2006) and Tsai *et al.* (2006). The expression data and annotations for Figure 3b are given as Supplementary Table S3.



et al. 2009a) for example, in the study by Taylor *et al.* (2005) zero ESTs would have been classed as significantly differentially expressed at this level of significance. Sixty six ESTs were classed as significantly differentially expressed using a Bayesian Log odds (B-stat) cut-off value of ≥ 3. Of these 66 ESTs, 15 were significantly up-regulated in e[CO_2] 13 of which were also ≥ two-fold up-regulated and these are given as supplementary information (Table S1) and 51 were significantly down-regulated, 38 of which were also ≥ two-fold down-regulated and these are given as supplementary information (Table S2).The two most significantly differentially expressed ESTs showing the greatest increase in abundance in the e[CO_2] treatment were annotated *leucoanthocyanidin dioxgenase* (*LDOX*, clone id R71B12) and *dihydroflavonol reductase* (DFR, clone id RSH03D11) exhibiting a normalised change in transcript abundance (e[CO_2] / a[CO_2]) of 39.6 (5.3 log_2) and 19.3 (4.3 log_2) respectively (Table S1). The gene models for the two Arabidopsis orthologs of the above were identified (*LDOX* gene model at4g22880 and *DFR* gene model at5g42800) and the mean normalised increase in transcript abundance (e[CO_2] / a[CO_2]) for all (not only those classed as significant) of the *Populus* ESTs showing homology to these Arabidopsis orthologs were 16.7 (4.1 log_2) and 30.2 (4.9 log_2) respectively. These two transcripts code for enzymes in the anthocyanin biosynthetic pathway of *Populus* (Tsai *et al.* 2006). The influence of this change in gene expression on leaf anthocyanin content was investigated. Anthocyanin content was measured on three occasions in August, October and November. Irrespective of CO_2 treatment, leaf anthocyanin content increased over time from late August to early-November, as expected during senescence (Keskitalo *et al.* 2005). Although not significant, anthocyanin content had increased by 413 % in e[CO_2] compared to a 342 % increase in a [CO_2] between 31st August and 4th November resulting in a 23 % increase in the anthocyanin content by 4th November in e[CO_2] compared with a[CO_2] (Figure 2c).

Identifying the pathways most significantly influenced by e[CO_2] late in autumnal senescence.

Although a relationship between mRNA and protein levels can be inferred here for anthocyanin biosynthesis this may often not be the case (Feder and Welser, 2005). Therefore, by grouping genes into functional categories Andersson *et al.* (2004) considered that mean values should represent a good approximation of the relative effort that plants are making to synthesize the proteins of the respective categories. Using the pathway analysis software Mapman (http://gabi.rzpd.de/projects/MapMan/) the metabolism overview map was explored against the Arabidopsis TAIR 8 database. Functional groups of genes which together exhibit a statistically significantly different behaviour in terms of expression profiles compared to all the other remaining functional groups were identified. The mean expression data were calculated for each EST passing the quality controls from the late senescent replicate microarray hybridisations. Of the 13,241 ESTs with expression data, 12,491 exhibited homology with the Arabidopsis genome, and unique Arabidopsis ortholog gene models numbered 4,712. The mean Log_2 expression data for each Arabidopsis ortholog was imported into MapMan and the functional groups within the metabolism class were analysed (Figure 3a). The statistically differentially regulated functional groups (BINs) were BIN 16.8 secondary metabolism of flavoniod biosynthesis *P* ≤0.02 (mean normalised expression ratio, e/a = 4.56 (2.19 log_2,)) and BIN 4 glycolysis P ≤0.04 (mean normalised expression ratio, e/a = 2.39 (1.26 log_2)). This approach provided further support for up-regulation of secondary metabolism leading to anthocyanin biosynthesis in e[CO_2] compared to a[CO_2]-grown leaves. Absent from the individual EST statistical analysis, an up-regulation of the glycolytic pathway was also identified.

In *Populus* the flavoniod biosynthetic pathway of secondary metabolism leading to anthocyanin biosynthesis contains increased gene copy numbers for many enzymes when compared with Arabidopsis (Tsai *et al.* 2006). Therefore when analysing this pathway further the mean expression data for each *Populus* gene model were used. Figure 3b shows that transcript abundance for enzymes catalysing the biosynthetic pathway from phenylalanine to anthocyanin were generally increased in e[CO_2]. Supplementary Information Table S3 gives the expression data for each EST, the *Populus* gene models and the Arabidopsis orthologs used in this figure. Although post-transcriptional processes play an important role in regulating metabolism, the greater transcript abundance for nearly the entire pathway, not just a few enzymes, provides strong evidence for a transcriptionally driven mechanism responding to e[CO_2] and stimulating anthocyanin biosynthesis. Anthocyanin pigments have a multi-faceted protective role in leaves, including protection from UV damage, pathogens, photoinhibiton, photo-oxidative stress and scavenging free radicals (Gould, 2004). The increased anthocyanin biosynthesis seen here is consistent with the idea that induction of stress responsive pathways can extend the viability of senescing cells (Buchanan-Wollaston *et al.* 2005). Diaz *et al.* (2006) suggest that anthocyanin may influence Arabidopsis leaf lifespan by protecting from photo-oxidative stress and Schaberg *et al.* (2008) identified a delay in abscission layer formation during autumnal senescence in sugar maple leaves containing increased anthocyanin.

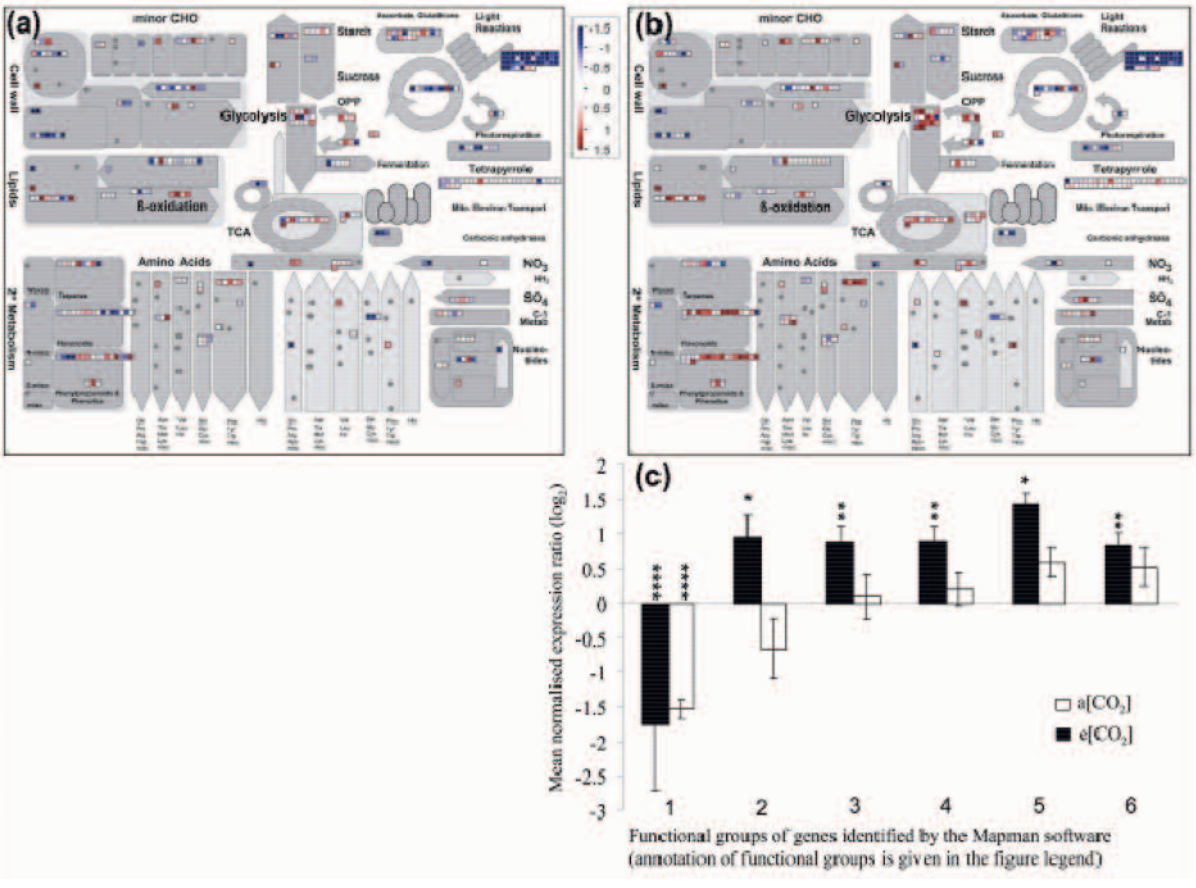
Identifying the genes most significantly influenced by e[CO_2] during the progression of autumnal senescence.

To determine whether this shift in metabolism late in senescence was in response to delayed senescence, the

changes in leaf transcript abundance and selected metabolites during the progress of autumnal senescence were examined. Differential gene expression between the pre-senescent (31 August) and the senescent canopies (18 October) was investigated using RNA from a pre-senescent reference pool, similar to the approach of Andersson *et al.* (2004). During senescence in a[CO_2] normalised transcript abundance was calculated as (18 Oct / 31 Aug) and 921 ESTs representing 148 unique *Populus* gene models were significantly differentially expressed (≥ 3 Bvalue and ≥ 2-fold change) whilst in the e[CO_2] treated canopies those ESTs significantly differentially expressed numbered 494 representing 49 unique *Populus* gene models. These differences are interesting as they suggest that the senescence transcriptional programme was disrupted by the e[CO_2] treatment. Using a different microarray platform, Andersson *et al.* (2004) identified 201 ESTs as the most up-regulated during senescence in *Populus tremula* and these represent 166 unique *Populus* gene models, an amount similar to that identified in

FIGURE 4

Mapman metabolism overview maps showing changes in transcript abundance during senescence in both e[CO_2] and a[CO_2] as a function of expression data (18 Oct / 31 Aug) within each CO_2 environment for Arabidopsis othologs of the *Populus* ESTs. The mean normalised expression ratio (log_2) is also given for those functional groups that exhibited a significantly different behaviour in terms of expression profiles compared to all the other remaining groups. Figures (a) and (b) represent gene expression changes during senescence in a[CO_2] and e[CO_2] respectively. The logarithmic colour scale bar ranges from - 1.5 (dark blue, representing a three-fold down-regulated gene on 18th Oct) to + 1.5 (dark red, representing a three-fold up-regulated gene on 18th Oct) is given.. The normalized expression values for all ESTs classed as significantly differentially expressed through the process of senescence in e[CO_2] are reported in Supplementary Tables 4 and 5 and those during senescence in a[CO_2] in Supplementary Tables 6 and 7. (c) Ratio for all functional groups classed as significantly differentially expressed (Benjamini Hochberg corrected Wilcoxon Rank Sum test in MapMan) during senescence in e[CO_2] in terms of expression profiles compared to all the other remaining groups. The expression ratio for the same group is given for the a[CO_2] and *P* values are reported as * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; **** $P \leq 0.0001$. Functional groups are annotated in MapMan as: (1) PS. light reactions, (2) Secondary metabolism. Flavonoids, (3) Secondary metabolism. Phenylpropanoids, (4) Glycolysis, (5) Amino acid metabolism. Synthesis. aromatic aa. Chorismate, (6) TCA cycle.



a[CO_2]. The ESTs exhibiting a two-fold change in normalized transcript abundance and a significant Bvalue and unique to senescence in each CO_2 environment are reported as supplementary tables (Tables S4 to S7). For example Table S4 gives the details of all 75 ESTs significantly up-regulated during senescence in e[CO_2] and Table S6 the 274 ESTs in a[CO_2] while 161 up-regulated ESTs were common to both environments. Table S5 reports the 67 down-regulated ESTs in e[CO_2] and table S7 the 295 EST down-regulated in a[CO_2] while 191 down-regulated ESTs were common to both environments. Although the Andersson *et al.* (2004) study and the data reported here used different microarray platforms of the significantly up-regulated ESTs during senescence in e[CO_2] only seven had commonality with Andersson's up-regulated list, whilst in a[CO_2] this numbered 15, data for these ESTs are reported in table S8. Table S4 reports the ESTs significantly up-regulated during senescence in e[CO_2] and ESTs representing four Arabidopsis ortholog gene models in the anthocyanin biosynthetic pathway were among those most abundant. These represented the Arabidopsis ortholog genemodels; at4g22880 (*LDOX*), at5g42800 (*DFR*), at2g37040 (*Phenylalanine ammonia lyase, PAL1*) and at5g05270 (*chalcone-flavanone isomerise, CHI*). The mean normalised increases in transcript abundance in e[CO_2] were 20.26 (4.3 $\log_2$), 12.75 (3.7 $\log_2$), 8.36 (3.1 $\log_2$)and 6.31 (2.7 $\log_2$) respectively for each of these gene models. Of these gene models both those expressing the products LDOX and DFR were among those at least four-fold up-regulated and *CHI* was between two – four fold up-regulated during the development of bud dormancy in *Populus* (Ruttink *et al.* 2007).

Identifying the pathways most significantly influenced by e[CO_2] during the progression of senescence.

The MapMan software was used to identify functional groups of genes which together exhibited a significantly different behaviour in expression during the progression of senescence in either a[CO_2] (18.10.04 / 31.08.04) or e[CO_2] (18.10.04 / 31.08.04). To identify changes in transcript abundance between senescence in a[CO_2] and senescence in e[CO_2] the sequences for all expression data from ESTs existing in both CO_2 environments and passing quality control were taken. These numbered 7404 ESTs, 7171 showed homology to the Arabidopsis genome and of these sequences, 2148 were classed as unique Arabidopsis ortholog gene models. Figure 4a represents functional groups within the metabolism overview in a[CO_2] and Figure 4b in e[CO_2]. In both CO_2 environments these figures show a significant down-regulation of genes involved in the photosynthetic regulation of the light reactions, as would be expected during senescence. It is notable that the mean down-regulation of this functional group is greater in e[CO_2] than in a[CO_2] (Figure 4c). It is possible that anthocyanin accumulation results in a stabilisation of photosynthetic proteins and pigments in e[CO_2] and thus delayed functional senescence despite induction of the senescence programme at the transcriptional level. During senescence in e[CO_2] the phenylpropanoid and flavonoid biosynthetic pathways were significantly up-regulated (18th Oct / 31st Aug) 1.84 (0.88 $\log_2$) and 1.93 (0.95 $\log_2$) respectively while during senescence in a[CO_2] the flavonoid biosynthetic pathway was down-regulated 0.63 (-0.67 $\log_2$) (Figure 4c). Studies of developing and mature soybean have shown an e[CO_2] induced increase in transcripts associated with glycolysis (Ainsworth *et al.* 2006; Leakey *et al.* 2009b). Data here indicates that both the glycolytic pathway and the TCA cycle were significantly up-regulated during senescence in e[CO_2] as were genes for enzymes of the shikimate pathway leading to chorismate biosynthesis, a pre-cursor for the aromatic amino acids such as phenylalanine (Figure 4c). These data provide support for the up-regulation of flavonoid biosynthesis and glycolysis obtained from a snap-shot late in senescence. Together these data indicate a shift in metabolism between senescence in a[CO_2] and senescence in e[CO_2] which appears to coincide with an up-regulation of glycolysis and secondary metabolism. As carried out for the late senescence samples the whole anthocyanin biosynthetic pathway was analysed. The change in transcript abundance was calculated from EST expression data as a function of (senescence in e[CO_2] / senescence in a[CO_2]) and the mean expression data for *Populus* gene model were reported (Figure 5a). The pathway analysis (Figure 5a) and data from the EST statistical analysis (Table S4) all indicate an active up-regulation of the anthocyanin pathway during the progress of senescence in e[CO_2] compared with that in a[$CO_2$2]. This supports the transcript data obtained from a direct comparison between CO_2 growth environments late in senescence (Figure 3b) and the leaf anthocyanin content data (Figure 2c).

Leaf carbohydrate contents

As photosynthate production declines during the active process of senescence, energy is generated by mitochondrial respiration through processes such as beta-oxidation (Andersson *et al.* 2004). During senescence in e[CO_2] it could be postulated that metabolism through glycolysis was still sufficient for energy generation, and the products of this metabolism were used in flavonoid biosynthesis. If this were the case then substrate for glycolysis should be present in the leaves of e[CO_2] exposed leaves and beta-oxidation could be expected to be up-regulated in the a[CO_2] leaves. Although, not classed as a significantly up-regulated functional group, those genes comprising BIN 11.9.4.2: (lipid metabolism.lipid degradation.beta-oxidation) exhibited an increase in mean normalised expression during senescence in a[CO_2] compared with e[CO_2], of 4.20 (2.07 $\log_2$) and 3.07 (1.62 $\log_2$) respectively (Figure 4a and b). Sucrose and starch content were increased in e[CO_2] this was only significant at two time points and only for sucrose (Figures 5 b and

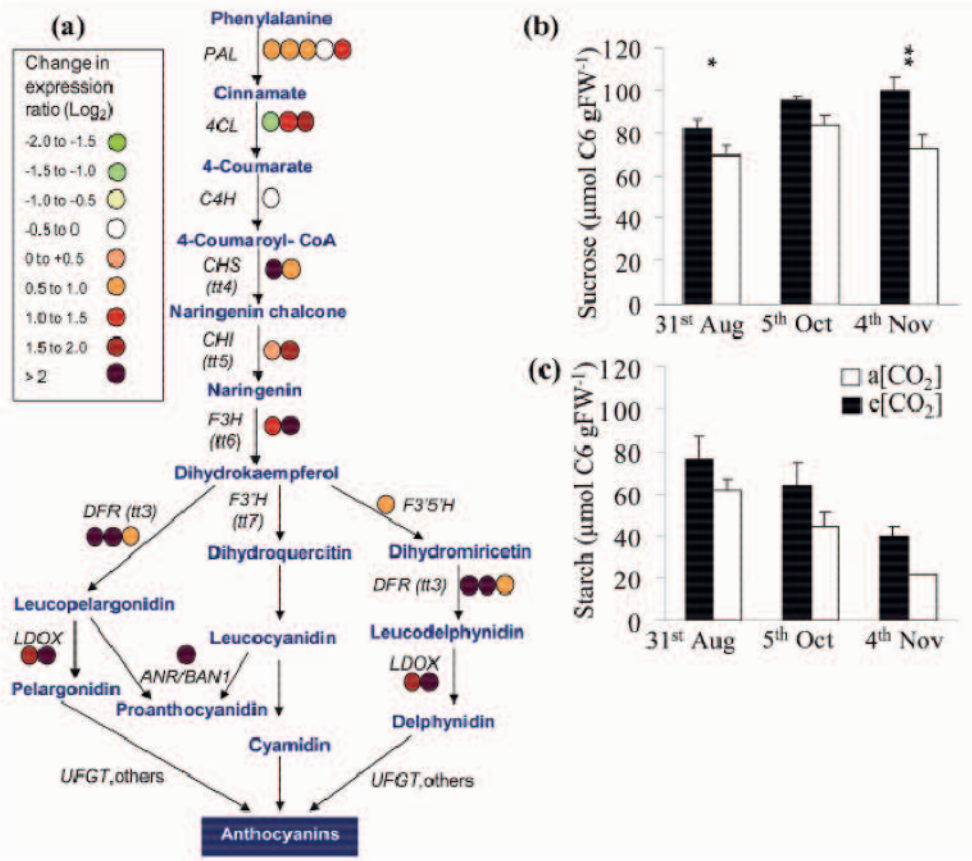
c). A gradual increase in sucrose content was apparent from August to November in e[CO_2] (Figure 5b). This is in contrast to leaves sampled in the growing season prior to the on-set of senescence during a preceding year where in e[CO_2] no accumulation of sucrose or hexoses were observed (Davey *et al.* 2006). Similarly the measured starch content was higher in leaves grown in e[CO_2] (Figure 5c) consistent with previous measurements on these trees (Davey *et al.* 2006). During senescence starch content declined in all leaves, suggesting that catabolism contributed to the energy requirements of the leaf (Figure 5c). Glucose and fructose contents are reported as supplementary Figure 1 (Figure S1)..During senescence the glucose content of leaves decreased and no significant difference was reported between treatments for either glucose or fructose.

Additional microarray confirmation using real-time quantitative PCR (RT-qPCR)

Transcript abundance data derived from the microarray hybridisations for the pathways reported here is consistent with the findings from leaf biochemical assays of anthocyanin and carbohydrate contents. Real-time quantitative PCR (RT-qPCR) was used to further assess the reliability of the microarray data and particularly to observe the CO_2 influences on *LDOX* transcript abundance. Primers were designed against the sequences of three genes up-regulated and three down-regulated in e[CO_2] on the 18th October. The expression of these genes was quantified using RT-qPCR and normailised with *PDF1* as a reference gene. Supplementary Figure 2 (Figure S2) shows good confirmation between the microarray data

FIGURE 5

The influence of e[CO_2] on gene expression within the anthocyanin biosynthetic pathway and leaf sucrose and starch content during senescence (between 31st August and 18th Oct) . (a) Log_2 expression data for *Populus* gene model from two replicate hybridizations for each CO2 treatment calculated as ((e[CO_2] 18 Oct / 31 Aug) / (a[CO2] 18 Oct/ 31 Aug)). Dark red circles represent a Log_2 value > 2 (a > four-fold increase in transcript abundance) for that gene during senescence in e[CO_2] compared with senescence in a[CO_2]. Green circles represent a Log2 value between -1.5 and -2 (a three to four-fold decrease in transcript abundance). The pathway diagram was constructed using Solfanelli *et al.* (2006) and Tsai *et al.* (2006). The expression data and annotations for figure 5a are given as Supplementary Table S9.(b) and (c) Data for leaf sucrose and starch content respectively in both a[CO_2] and e[CO_2], n = 8, *P* values are reported as * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$ where no *P* value was reported data were not significant.



and those from RT-qPCR. The relative expression of *LDOX* between the late growing season and late senescence samples was also examined using RT-qPCR. A clear up-regulation of *LDOX* during senescence in e[CO_2] can be seen, with the opposite occurring in a[CO_2] (Supporting Information (Fig S2 inset).

Summary

Once autumnal senescence in *Populus* is initiated by a change in photoperiod (Olsen *et al.* 1997; Keskitalo *et al.* 2005; Böhlenius *et al.* 2006) the balance between reactive oxygen species (ROS) production and ROS scavenging can determine the rate of senescence (McKersie *et al.* 1988; Buchanan-Wollaston *et al.* 2003; Gepstein *et al.* 2003). We have shown that soluble carbohydrates and starch were increased by exposure to e[CO_2] at a concentration predicted for 2050, and that this may act as a signal to stimulate the synthesis of anthocyanin. This supports findings in several contrasting plant species and organs and is supported by work on the mutant *pho3* (Lloyd & Zakhleniuk, 2004). In this sucrose-export mutant (*pho3*) LDOX and DFR were up-regulated by 190 and 31 times respectively, again linking leaf sucrose content and anthocyanin biosynthesis. Direct sugar induction of anthocyanin biosynthesis in Arabidopsis has also been reported (Teng *et al.* 2005; Solfanelli *et al.* 2006). Furthermore, the association between e[CO_2] and the partitioning of carbon to the synthesis of secondary metabolites was evident in tobacco plants exposed to an e[CO_2] of 1000 ppm (Matros *et al.* 2006). The Matros *et al.* (2006) study provides evidence for a direct link between e[CO_2], an increased leaf C:N ratio and an increased activity of Phenylalanine ammonia lyase (*PAL*), a key enzyme catalyzing the first committed step in the biosynthesis of phenylpropanoids, with a concomitant increase of secondary metabolites. Long-lived trees, such as *Populus*, have evolved strategies for defence, dormancy and wood formation that may not be well represented in the genomes of annuals such as Arabidopsis. Enzymes involved in the flavonoid biosynthetic pathway leading to anthocyanin production are coded by multiple copy genes in *Populus* and generally single copy genes in Arabidopsis for example, *LDOX* and *DFR* (Tsai *et al.* 2006). This study further highlights the importance of using Populus as a model to study natural autumnal senescence (Jansson & Douglas, 2007) and the open-field environment as essential to gain a mechanistic understanding of how trees may respond in the natural environment (Taylor *et al.* 2005; Sjödin *et al.* 2008).

In conclusion, we have identified an association between a delayed autumnal senescence in e[CO_2] a change in leaf carbohydrate status, gene expression profiles and anthocyanin content. It is possible that this may be a secondary response to other factors. For example, canopy temperature is often increased during growth in e[CO_2] at FACE sites, as is evident for soy bean (Long *et al.* 2006). Nevertheless the data reported here begin to identify processes by which climate change can influence plant phenology. Using transcript profiling we identified a number of genes that changed in expression during senescence in an atmosphere enriched with CO_2 and the most conspicuous of these were genes involved in the biosynthetic pathway of anthocyanin; they were strongly and significantly induced in e[CO_2] resulting in an increased leaf anthocyanin content. The carbohydrate content of senescing leaves was also increased in e[CO_2] and we propose that this provided both a signal and a fuel for anthocyanin biosynthesis. Additional leaf anthocyanin then provides increased protection to the senescing leaves in e[CO_2] and extends the senescent phase. This CO_2 stimulated shift in metabolism is consistent with the growth/differentiation balance hypothesis extended by Herms and Matteson (1992) and observed in forest studies (Harding *et al.* 2005; Mattson *et al.* 2005; Cseke *et al.* 2009). Evidence from this study allows us to postulate that excess carbon in e[CO_2] is available to be partitioned to carbon rich secondary metabolites. These provide a protective role in senescing leaves, so extending leaf longevity, as the carbon demand from growth declines in the autumn. We have begun to identify the genetic mechanisms for adaptation to future CO_2, but the long-term consequences of such changes for forest ecosystem function and micro-evolutionary adaptation remain uncertain.

Acknowledgements

This research was supported by the European Commission through the Directorate General Research within the Fifth Framework for Research – Quality of Life and Management of the Living Resources Programme, contract number QLK5-CT-2002-00953 (POPYOMICS), coordinated by the University of Southampton. The PopFACE experiment within the EUROFACE infrastructure was supported by the EC through its Environment R and D programme within the Fourth Framework as a research contract ENV4-CT97-0657 (PopFACE) and within the Fifth Framework as contract EVR1-CT-2002-40027 (EUROFACE) coordinated by Giuseppe Scarascia-Mugnozza at the University of Viterbo. This research was also supported by the Office of Science (BER), US Department of Energy, Grant No. DE-FG02-04ER63792, 'POPGENICS' awarded to GT and DFK. AR was supported by the US Department of Energy Office of Science grant no. DEFG02-04ER63792 and by contract no. DE-AC02-98CH10886 to Brookhaven National Laboratory. Research in the laboratory of GT was also supported by the DEFRA project BEGIN and the Seventh Framework Network of Excellence, EVOLTREE. The authors thank B Gielen, C Calfapietra, T Oro, GJJ Clarkson, M Pecchiari and C Steynen for help in the field, and LE Graham and J Tucker for help with the microarray hybridisations. A Zaldei is thanked for support of the infrastructure at the POPFACE field site.

REFERENCES

Ainsworth EA, Rogers A, Nelson R and Long SP. 2004. Testing the "source-sink" hypothesis of down-regulation of photosynthesis in elevated [CO_2] in the field with single gene substitutions in *Glycine max. Agricultural and Forest Meteorology* **122**: 85-94

Ainsworth EA, Rogers A, Vodkin LO, Walter A & Schurr U. 2006. The effects of elevated CO_2 concentration on soybean gene expression. An analysis of growing and mature leaves. *Plant Physiology* **142**: 135-147.

Andersson A, Keskitalo J, Sjodin A, Bhalerao R, Sterky F, Wissel K, Tandre K, Aspeborg H, Moyle R, Ohmiya Y *et al.* 2004. A transcriptional timetable of autumn senescence. *Genome Biology* **5**: art. no.-R24.

Benjamini Y and Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) **57**: 289-300.

Bryant JB, Taylor G and Frehner M. 1998. Photosynthetic acclimation in chalk grassland herbs exposed to elevated CO_2 in a FACE system. *Plant Cell and Environment* **21**: 159-168.

Böhlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH and Nilsson O. 2006. CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* **312**: 1040-1043.

Brosché M, Vinocur B, Alatalo ER, Lamminmaki A, Teichmann T, Ottow EA, Djilianov D, Afif D, Bogeat-Triboulot MB, Altman A, Polle A, Dreyer E, Rudd S, Lars P, Auvinen P and Kangasjarvi J. 2005. Gene expression and metabolite profiling of *Populus euphratica* growing in the Negev desert. *Genome Biology* **6**: art. no.-**R101**

Buchanan-Wollaston V, Earl S, Harrison E, Mathas E, Navabpour S, Page T and Pink D. 2003. The molecular analysis of leaf senescence - a genomics approach. *Plant Biotechnology Journal* **1**: 3-22.

Buchanan-Wollaston V, Page T, Harrison E, Breeze E, Lim PO, Nam HG, Lin J-F, Wu S-H, Swidzinski J, Ishizaki K *et al.* 2005. Comparartive transcriptome analysis reveals significant differences in gene expression and signalling pathways between developmental and dark/starvation-induced senescence in Arabidopsis. *The Plant Journal* **42**: 567-585.

Chang S, Puryear J and Cairney J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* **11**: 113-116.

Chen WJ, Black TA, Yang PC, Barr AG, Neumann HH, Nesic Z, Blanken PD, Novak MD, Eley J, Ketler RJ *et al.* 1999. Effects of climatic variability on the annual carbon sequestration by a boreal aspen forest. *Global Change Biology* **5**: 41-53.

Cseke LJ, Tsai C-J, Rogers A, Nelson MP, White HL, Karnosky DF, Podila GK. 2009. Transcriptomic comparison in the leaves of two aspen genotypes having similar carbon assimilation rates but different allocation patterns under elevated CO_2. *New Phytologist.* In Press.

Davey PA, Olcer H, Zakhleniuk O, Bernacchi CJ, Calfapietra C, Long SP and Raines CA. 2006. Can fast-growing plantation trees escape biochemical down-regulation of photosynthesis when grown throughout their complete production cycle in the open air under elevated carbon dioxide? *Plant, Cell and Environment* **29**: 1235 - 1244.

Déjardin A, Leplé J-C, Lesage-Descauses M-C, Costa G and Pilate G. 2004. Expressed Sequence Tags from poplar wood tissues - A comparative analysis from multiple libraries. *Plant Biology*, **6**: 55 - 64

Diaz E, Yang YH, Ferreira T, Loh KC, Okazaki Y, Hayashizaki Y, Tessier-Lavingne M, Speed TP, Ngai J. 2003. Analysis of gene expression in the developing mouse retina. *Proceedings of the National Academy of Sciences, USA* **100**: 5491–5496.

Diaz C, Saliba-Colombani V, Loudet O, Belluomo P, Moreau L, Daniel-Vedele F, Morot-Gaudry J-F and Masclaux-Daubresse C. 2006. Leaf yellowing and anthocyanin accumulation are two genetically independent strategies in response to nitrogen limitation in Arabidopsis thaliana. *Plant Cell Physiol*. **47**: 74-83.

Feder ME and Walser JC. 2005. The biological limitations of transcriptomics in elucidating stress and stress responses. *Journal of Evolutionary Biology.* **18**: 901-910

Fracheboud Y, Luquez V, Björkén L, Sjödin A, Tuominen H, and Jansson S. 2009. The control of autumn senescence in European Aspen. *Plant Physiology* **149**:1982-1991

Gamon JA and Surfus JS. 1999. Assessing leaf pigment content and activity with a reflectometer. *New Phytologist* **143**: 105-117

Gepstein S, Sabehi G, Carp M-J, Hajouj T, Nesher MFO, Yariv I, Dor C and Bassani M. 2003. Large-scale identification of leaf senescence-associated genes. *The Plant Journal* **36**: 629-642.

Gitelson AA and Merzlyak MN. 1994. Spectral reflectance changes associated with autumn senescence of *Aesculus-Hippocastanum* L and *Acer-Platanoides* L leaves - spectral features and relation to chlorophyll estimation. *Journal of Plant Physiology* **143**: 286-292

Gould KS. 2004. Natures Swiss army knife: the diverse protective roles of anthocyanins in leaves. *Journal of Biomedicine and Biotechnology* **5**: 314-320.

Harding SA, Jiang H, Jeong ML, Casado FL, Lin H-W and Tsai C-J. 2005. Functional genomics analysis of foliar condensed tannin and phenolic glycoside regulation in natural cottonwood hybrids. *Tree Physiology* **25**: 1475-1486.

Herms DA and Mattson WJ. 1992. The dilemma of plants: to grow or defend. *The Quarterly Review of Biology* **67**: 283-335.

Herrick JD and Thomas RB. .2003. Leaf senescence and late-season net photosynthesis of sun and shade leaves of overstory sweetgum (*Liquidambar styraciflua*) grown in elevated and ambient carbon dioxide concentrations. *Tree Physiology* **23**: 109-118.

Ingvarsson PK, Garcia MV, Hall D, Luquez V, Jansson S. 2006. Clinal variation in *phyB2*, a candidate gene for day-lengh induced growth cessation and bud set, across a latitudinal gradient in European aspen (*Populus tremula*). *Genetics* **172**:1845–1853

Jansson S and Douglas C. 2007. Populus: a model system for plant biology. Annual *Review of Plant Biology*. **58**: 435-458.

Keskitalo J, Bergquist G, Gardeström and Jansson S. 2005. A cellular timetable of autumnal senescence. *Plant Physiology* **139**: 1635-1648.

Kaschuk G, Hungria M, Leffelaar PA, Giller KE and Kuyper TW. 2009. Differences in photosynthetic behaviour and leaf senescence of soybean (*Glycine max* [L.] Merrill) dependent on N2 fixation or nitrate supply. *Plant Biology doi:10.1111/j.1438-8677.2009.00211.x.*

Leakey ADB, Ainsworth EA, Bernard SM, Cody Markelz RJ, Ort DR, Placella SA, Rogers A, Smith MD, Sudderth EA, Weston DJ *et al.* 2009a. Gene expression profiling: opening the black box of plant ecosystem responses to global change. *Global Change Biology* **15**: 1201–1213

Leakey ADB, Xu F, Gillespie KM, McGrath JM, Ainsworth EA and Ort DR. 2009b. Genomic basis for stimulated respiration by plants growing under elevated carbon dioxide. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 3597-3602.

Li, J-H, Dijkstra P, Hymus GJ, Wheeler RM, Piastuch WC, Hinkle CR, and Drake BG. 2000. Leaf senescence of Quercus myrtifolia as affected by long-term CO_2 enrichment in its native environment. *Global Change Biology* **6**: 727-733.

Liberloo M, Tulva I, Raïm O and Ceulemans R. 2007. Photosynthetic stimulation under long-term CO_2 enrichment and fertilization is sustained across a closed *Populus* canopy profile (EUROFACE). *New Phytologist* **173**: 537-549.

Liberloo M, Lukac M, Calfapietra C, Hoosbeek MR, Gielen B, Miglietta F, Scarascia-Mugnozza GE and Ceulemans R. 2009. Coppicing shifts CO_2 stimulation of poplar productivity to above-ground pools: a synthesis of leaf to stand level results from the POP/EUROFACE experiment. *New Phytologist* **182**: 331–346

Lloyd JC and Zakhleniuk OV. 2004. Responses of primary and secondary metabolism to sugar accumulation revealed by microarray expression analysis of the Arabidopsis mutant, *pho3. Journal of Experimental Botany* **55**: 1221 - 1230.

Long SP, Ainsworth EA, Leakey ADB, Nösberger J and Ort DR. 2006. Food for thought: lower-than-expected crop yield stimulation with rising CO_2 concentrations. *Science* **312**: 1918-1921

Lucht W, Prentice CI, Myneni RB, Sitch S, Friedlingstein P, Cramer W, Bousquet P, Buermann W and Smith B. 2002. Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science* **296**: 1687-1688.

Martin T, Oswald O, and Graham IA. 2002. Arabidopsis seedling growth, storage lipid mobilization, and photosynthetic gene expression are regulated by carbon: nitrogen availability. *Plant Physiology* **128**: 472-481.

Matros A, Amme S, Kettig B, Buck-Sorlin GH, Sonnewald U and Mock H-P. 2006. Growth at elevated CO_2 concentrations leads to modified profiles of secondary metabolites in tobacco cv. SamsunNN and to increased resistance against infection with potato virus Y. *Plant, Cell and Environment* **29**: 126-137.

Mattson WJ, Julkunen-Tiitto R and Herms DA. 2005. CO_2 enrichment and carbon partitioning to phenolics: do plant responses accord better with the protein competition or the growth-differentiation balance models? *OIKOS* **111**: 337-347.

McLeod AR and Long SP. 1999. Free-air carbon dioxide enrichment (FACE) in global change research: A review. In: *Advances in Ecological Research* **28**: 1-56.

McKersie BD, Senaratna T, Walker MA, Kendall EJ and Hetherington PR. 1988. Deterioration of membranes during aging in plants: evidence for free radical mediation. In LD Noodén and AC Leopold eds, *Senescence and aging in plants.* Academic Press Inc. London.

Menzel A, and Fabian P. 1999. Growing season extended in Europe. *Nature* **397**: 659-659.

Menzel, A, Sparks TH, Estrella N Koch E, Aasa A, Ahas R, Alm-kübler K, Bissolli P, Braslavská O, Briede A *et al.* 2006. European phenological response to climate change matches the warming pattern. *Global Change Biology* **12**: 1-8.

Miglietta F, Peressotti A, Vaccari FP, Zaldei A, DeAngelis P and Scarascia-Mugnozza G. 2001. Free-air CO_2 enrichment (FACE) of a poplar plantation: the POPFACE fumigation system. *New Phytologist* **150**: 465-476.

Murakami PF, Schaberg PG and Shane JB. 2008. Stem girdling manipulates leaf sugar concentrations and anthocyanin expression in sugar maple trees during autumn. *Tree Physiology* **28**: 1467–1473

Myneni RB, Keeling CD, Tucker CJ, Asrar G and Nemani RR. 1997. Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* **386**: 698-702.

Olsen JE, Junttila O, Nilsen J, Eriksson ME, Martinussen I, Olsson O, Sandberg G and Moritz T. 1997. Ecotpic expression of oat phytochrome A in hybrid aspen changes critical daylength for growth and prevents cold acclimatization. *The Plant Journal* **12**: 1339 - 1350.

Park J-Y, Canam T, Kang K-Y, Unda F and Mansfield SD. 2009. Sucrose phosphate synthase expression influences poplar phenology. *Tree Physiology* **29**: 937–946

Parmesan C and Yohe G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**: 37-42.

Peñuelas J, Rutishauser T, and Filella I. 2009. Phenology feedbacks on climate change. *Science* **324**: 887-888

Rae A, Ferris R, Tallis MJ and Taylor G. 2006. Elucidating genomic regions determining enhanced leaf growth and delayed senescence in elevated CO_2. *Plant, Cell and Environment* **29**: 1730-1741.

Rinaldi C, Kohler A, Frey P, Duchaussoy F, Ningre N, Couloux A, Wincker P, Thiec DL, Fluch S, Martin F *et al.* 2007. Transcript profiling of poplar leaves upon infection with compatible and incompatible strains of the foliar rust *Melampsora larici-populina. Plant Physiology* **144**: 347-366.

Ruttink T, Arend M, Morreel K, Storme V, Rombauts S, Fromm J, Bhalerao RP, Boerjan W, and Rohdea A. 2007. A molecular timetable for apical bud formation and dormancy induction in Poplar. *The Plant Cell* **19**: 2370-2390.

Rogers A, Gibon Y, Stitt M, Morgan PB, Bernacchi CJ, Ort DR, Long SP. 2006. Increased carbon availability at elevated carbon dioxide concentration improves nitrogen assimilation in a legume. *Plant, Cell & Environment*. **29**: 1651-1658.

Scarascia-Mugnozza G, Calfapietra C, Ceulemans R, Gielen B, Cotrufo MF, De Angelis P, Godbold D, Hoosbeek MR, Kull O, Lukac M *et al.* 2006. Responses to elevated [CO_2] of a short rotation, multispecies poplar plantation: the POPFACE / EUROFACE experiment. In: Nösberger J, Long SP, Norby RJ, Stitt M, Hendrey GR and Blum H, eds. *Ecological Studies*, Vol. 187. *Managed ecosystems and CO_2*

Case studies, processes and perspectives. Springer-Verlag, Berlin Heidelberg, 173-195.

Schaberg PG, Murakami PF, Turner MR, Heitz HK and Hawley GJ. 2008. Association of red coloration with senescence of sugar maple leaves in autumn. *Trees* **22**: 573–578

Sigurdsson BD. 2001. Elevated CO_2 and nutrient status modified leaf phenology and growth rhythm of young *Populus trichocarpa* trees in a 3-year field study. *Trees-Structure and Function* **15**: 403-413.

Sjödin A, Wissel K, Bylesjö M, Trygg J and Jansson S. 2008. Global expression profiling in leaves of free-growing aspen. *BMC Plant Biology* **8**:61

Sjödin A, Street NR, Sandberg G, Gustafsson P, Jansson S. 2009. The *Populus* Genome Integrative Explorer (PopGenIE): a new resource for exploring the *Populus* genome. *New Phytologist* **182**: 1013–1025

Smyth G K. 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* **3**: No. 1, Article 3

Solfanelli C, Poggi A, Loreti E, Alpi A and Perata P. 2006. Sucrose-specific induction of the anthocyanin biosynthetic pathway in Arabidopsis. *Plant Physiology* **140**: 637-646.

Street NR, Tucker J, Brosché M, Kangasjärvi J, Broadmeadow M, Taylor G. 2009. A genetical genomics approach to understand ozone sensitivity in *Populus*. *Environmental pollution*, In press.

Taylor G. 2002. *Populus*. Arabidopsis for forestry. Do we need a model tree? *Annals of Botany* **90**: 681-689.

Taylor G, Street NR, Tricker PJ, Sjodin A, Graham L, Skogstrom O, Calfapietra C, Scarascia-Mugnozza G and Jansson S. 2005. The transcriptome of *Populus* in elevated CO_2. *New Phytologist* **167**: 143-154.

Taylor G, Tallis MJ, Giardina CP, Percy KE, Miglietta F, Gupta PS, Gioli B, Calfapietra C, Gielen B, Kubiske ME *et al.* 2008. Future atmospheric CO_2 leads to delayed autumnal senescence in *Populus* over two continents. *Global Change Biology* **14**: 264-275.

Teng S, Keurentjes J, Bentsink L, Koornneef M and Smeekens S. 2005. Sucrose-specific induction of anthocyanin biosynthesis in Arabidopsis requires the *MYB75/PAP1* gene. *Plant Physiology* **139**: 1840-1852.

Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller L, Rhee SYMS. 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant Journal* **37**: 914–939

Tricker PJ, Trewin H, Kull O, Clarkson GJJ, Eensalu E, Tallis MJ, Colella A, Doncaster CP, Sabatti M and Taylor G .2005. Stomatal conductance and not stomatal density determines the long-term reduction in leaf transpiration of poplar in elevated CO_2. *Oecologia* **143**: 652-660

Tsai C-J, Harding SA, Tschaplinski TJ, Lindroth RL and Yuan Y. 2006. Genome-wide analysis of the structural genes regulating defense phenylpropanoid metabolism in Populus. *New Phytologist* **172**: 47-62.

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A *et al.* 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**: 1596–1604.

Wingler A, Mares M and Pourtau N. 2004. Spatial patterns and metabolic regulation of photosynthetic parameters during leaf senescence. *New Phytologist* **161**: 781-789.

Wingler A, Purdy S, MacLean JA and Pourtau N. 2006. The role of sugars in integrating environmental signals during the regulation of leaf senescence. *Journal of Experimental Botany* **57**: 391–399.

Zhou L, Tucker C J, Kaufmann RK, Slayback D, Shabanov NV and Myneni RB. 2001. Variations in northern vegetation activity infrerred from satellite data of vegetation index during 1981 to 1999. *Journal of Geophysical Research* **106**: 20069-20083.

LATE QUATERNARY HISTORY OF NORTH EURASIAN NORWAY SPRUCE (*PICEA ABIES*) AND SIBERIAN SPRUCE (*PICEA OBOVATA*) INFERRED FROM MACROFOSSILS, POLLEN AND CYTOPLASMIC DNA VARIATION

Mari Mette Tollefsrud[1], Małgorzata Latałowa[2], W. O. van der Knaap[3], Christian Brochmann[4] and Christoph Sperisen[5]

[1] Norwegian Institute of Bioeconomy Research (NIBIO), NO-1431 Ås, Norway,
[2] Laboratory of Palaeoecology and Archaeobotany, Department of Plant Ecology, University of Gdańsk, PL-80-308 Gdańsk, Poland,
[3] Institute of Plant Sciences and Oeschger Centre for Climate Change Research, University of Bern, CH-3013 Bern, Switzerland,
[4] National Centre for Biosystematics, Natural History Museum, University of Oslo, Blindern, NO-0318 Oslo, Norway,
[5] Swiss Federal Research Institute for Forest, Snow and Landscape Research (WSL), CH-8903 Birmensdorf, Switzerland.

Corresponding authors:
Christoph Sperisen
Swiss Federal Research Institute for Forest, Snow and Landscape Research (WSL), Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland
E-mail: sperisen@wsl.ch

Aim We used combined palaeobotanical and genetic data to assess whether Norway spruce (*Picea abies*) and Siberian spruce (*Picea obovata*), two major components of the Eurasian boreal forests, occupied separate glacial refugia, and to test previous hypotheses on their distinction, geographical delimitation and introgression.

Location The range of Norway spruce in northern Europe and Siberian spruce in northern Asia.

Methods Pollen data and recently compiled macrofossil records were summarized for the Last Glacial Maximum (LGM), late glacial and Holocene. Genetic variation was assessed in 50 populations using one maternally (mitochondrial *nad1*) and one paternally (chloroplast *trn*T–*trn*L) inherited marker and analysed using spatial analyses of molecular variance (SAMOVA).

Results Macrofossils showed that spruce was present in both northern Europe and Siberia at the LGM. Congruent macrofossil and pollen data from the late glacial suggested widespread expansions of spruce in the East European Plain, West Siberian Plain, southern Siberian mountains and the Baikal region. Colonization was largely completed during the early Holocene, except in the formerly glaciated area of northern Europe. Both DNA markers distinguished two highly differentiated groups that correspond to Norway spruce and Siberian spruce and coincide spatially with separate LGM spruce occurrences. The division of the mtDNA variation was geographically well defined and occurred to the east of the Ural Mountains along the Ob River, whereas the cpDNA variation showed widespread admixture. Genetic diversity of both DNA markers was higher in western than in eastern populations.

Main conclusions North Eurasian Norway spruce and Siberian spruce are genetically distinct and occupied separate LGM refugia, Norway spruce on the East European Plain and Siberian spruce in southern Siberia, where they were already widespread during the late glacial. They came into contact in the basin of the Ob River and probably hybridized. The lower genetic diversity in the eastern populations may indicate that Siberian spruce suffered more from past climatic fluctuations than Norway spruce.

Key words: DNA markers, fossil records, glacial refugia, introgression, *Picea abies, Picea obovata*, post-glacial colonization.

Published in Journal of Biogeography (2015) 42: 1431-1442, with doi:10.1111/jbi.12484

Appendices S1, S2 and S3 and their respective references can be downloaded here: http://onlinelibrary.wiley.com/wol1/doi/10.1111/jbi.12484/suppinfo

Introduction

The boreal forest or taiga of Eurasia, with fir (*Abies*), larch (*Larix*), pine (*Pinus*) and spruce (*Picea*) as the most abundant components, is one of the largest continuous forest areas worldwide. Like most Northern Hemisphere biota, the taiga experienced repeated range contractions and expansions driven by the Quaternary climate fluctuations and the associated formation and decay of ice sheets and permafrost (Frenzel, 1968). However, little is known about the distribution of taiga taxa during glacial maxima. Based on fossil pollen and macrofossils, it was assumed that the taiga was confined to small areas located south of its current limits during the Last Glacial Maximum (LGM) (Frenzel, 1968; Tarasov *et al.* 2000). This view has recently been challenged by new macrofossil data, demonstrating that several tree taxa, including birch (*Betula*), larch, pine and spruce, persisted locally at high latitudes during the LGM in areas south and east of the Scandinavian ice sheet (Binney *et al.* 2009). A more northerly and wider LGM distribution of taiga taxa is also supported by genetic data. For example, a detailed survey

of molecular variation in Siberian larch (*Larix sibirica*) suggested the presence of several isolated refugia in the southern Siberian mountains, the foothills of the Sayan Mountains and northern Siberia (Semerikov *et al.* 2013). In Norway spruce [*Picea abies* (L.) H. Karst.], genetic data have corroborated the fossil-based inference that populations survived on the East European Plain (e.g. Tollefsrud *et al.* 2008) and suggested a refugium in Scandinavia (Parducci *et al.* 2012a).

In this study, we further examine the history of Norway spruce, and present data for it and the closely related Siberian spruce (*Picea obovata* Ledeb.), two major components of the taiga. Their combined range extends almost continuously from Norway in the west to the coast of the Sea of Okhotsk in the east (Fig. 1; Schmidt-Vogt, 1977). Despite their ecological importance, the delimitation and taxonomic rank of Norway spruce and Siberian spruce are not clear. Some authors recognize two separate species (Pravdin, 1975; Popov, 2003). Others consider them either as two closely related subspecies (Tutin *et al.* 1993) or geographical varieties (Schmidt-Vogt, 1974a), a suggestion which is supported by allozyme data: based on a limited number of populations, very little allozyme differentiation was detected between the two taxa, consistent with a rank of subspecies or variety (Krutovskii & Bergmann, 1995). Phylogenetic analyses using molecular markers have corroborated the close relationship (Ran *et al.* 2006) and indicated that the Norway spruce of northern Europe is more closely related to Siberian spruce than to the Norway spruce of central and south-eastern Europe (Lockwood *et al.* 2013).

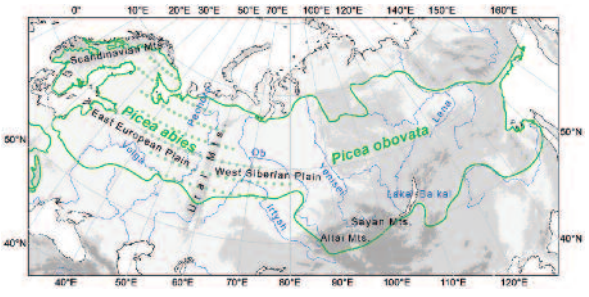
At the morphological level, Norway spruce and Siberian spruce can only be distinguished by the shape of their cone-scales (Schmidt-Vogt, 1974a). The cone-scales of Norway spruce are slightly pointed, whereas those of Siberian spruce are rounded, although a range of intermediate forms occur. Popov (2003) described a smooth longitudinal gradient of cone-scale morphology, with distinct Norway spruce populations found in the extreme west of northern Europe and distinct Siberian spruce populations east of the Ural Mountains (Fig. 1). It has been suggested that this morphological pattern is a result of introgression, when the two taxa came into contact after the last glaciation (Pravdin, 1975; Popov, 2003). Introgression is indeed supported by allozyme data, which revealed higher levels of genetic diversity in putative hybrid populations than in populations with 'pure' cone-scale shapes (Krutovskii & Bergmann, 1995). Schmidt-Vogt (1974a), on the other hand, argued that the morphological gradient reflects adaptation to cold climate rather than introgression, because the frequency of rounded forms also increases with latitude and elevation.

The glacial and post-glacial history of Norway spruce has been intensively studied using fossil pollen (e.g. Giesecke & Bennett, 2004; Latałowa & van der Knaap, 2006), macrofossils (Terhürne-Berson, 2005), genetic data (e.g. Lagercrantz & Ryman, 1990; Vendramin *et al.* 2000; Sperisen *et al.* 2001; Collignon *et al.* 2002; Heuertz *et al.* 2006) and pollen and genetic data combined (Tollefsrud *et al.* 2008). These studies all agree that the Norway spruce of northern Europe and that of central and south-eastern Europe originate from separate glacial refugia. In the north, detailed pollen data (Giesecke & Bennett, 2004; Latałowa & van der Knaap, 2006), mitochondrial DNA (mtDNA) (Tollefsrud *et al.* 2008) and nuclear microsatellites (Tollefsrud *et al.* 2009) have consistently revealed the presence of a major refugium on the East European Plain. Inferences from macrofossils, stomata and pollen indicated that the refugium extended as far north as the Pechora region (Väliranta *et al.* 2011). The presence of a separate refugium in northern Scandinavia was inferred from ancient DNA and the current distribution of a mitotype that appears to be unique to Scandinavia (Parducci *et al.* 2012a,b). Population survival in northern Scandinavia was also suggested based on macrofossils (e.g. Kullman, 2008). The existence of refugial populations in Scandinavia has, however, been questioned (e.g. Birks *et al.* 2012; Vorren *et al.* 2013).

No comparable genetic and palaeobotanical data exist for spruce in Siberia. The available palaeobotanical data suggest that Siberian spruce survived in the lowlands of southern Siberia (Tarasov *et al.* 2000; Binney *et al.* 2009) and the Baikal region (Blyakharchuk *et al.* 2004; Bezrukova *et al.* 2005). These studies did not, however, consider cone-scale morphology, and it remains unclear whether Norway spruce and Siberian spruce also shared refugia, as was proposed based on allozyme data (Krutovskii & Bergmann, 1995).

FIGURE 1

Geographical distribution of Norway spruce (*Picea abies*) and Siberian spruce (*Picea obovata*) in northern Eurasia after Schmidt-Vogt (1974b; northern Europe and easternmost Siberia) and Bezrukova *et al.* (2005; Siberia). The current combined distribution of the two taxa is indicated by the green outline. The green dotted area marks the introgression zone suggested by Popov (2003, 2010). The shading illustrates topography: darker is higher altitude.



Here, we summarize macrofossil and pollen data and assess cytoplasmic DNA variation in spruce populations sampled across boreal Eurasia to test whether Norway spruce and Siberian spruce occupied separate or shared glacial refugia and to clarify whether they are today genetically distinct and fully allopatric or connected by a zone of introgression. Cytoplasmic DNA variation was assessed in one mtDNA and one chloroplast DNA (cpDNA) marker. As in most conifers, mtDNA is maternally inherited in Picea (Sutton *et al.* 1991; Grivet *et al.* 1999) and thus dispersed by seeds only, whereas cpDNA is paternally inherited (Sutton *et al.* 1991) and dispersed by both pollen and seeds. Hence, their combined use provides opportunities for studying the geographical distribution of both maternal and paternal lineages. We hypothesize that Norway spruce and Siberian spruce underwent largely independent glacial histories. We therefore make the following predictions. (1) There are two major genetic groups, corresponding to the morphologically defined Norway spruce and Siberian spruce. (2) One genetic group coincides spatially with LGM spruce occurrences in northern Europe, and the other coincides with occurrences in Siberia. (3) The two genetic groups form a zone of contact along or in the vicinity of the Ural Mountains. (4) In the area of contact, the two genetic groups are admixed, reflecting introgression.

Materials and methods

Macrofossil and pollen data

Macrofossil and pollen data play complementary roles in palaeoecology. Macrofossils provide the most direct evidence for the past presence of a plant taxon in a particular area (Birks & Birks, 2000), but macrofossil sites are scarce in many regions. Owing to long-distance transport, pollen data alone have limited value in demonstrating the presence of a species, particularly when the pollen abundance of a given taxon is low. Pollen sites are, however, often numerous and may thus support other types of evidence. Pollen is also valuable for assessing the dynamics of expansion, because it can be quantified (e.g. Latałowa & van der Knaap, 2006). To obtain the best resolution, we used both types of records.

Macrofossil records were mainly derived from the Northern Eurasian Macrofossil Database (Binney *et al.* 2009). Although many records refer to either Norway spruce or Siberian spruce, the distinction between the taxa must be regarded as tentative because only very few records of cones exist in the database. Consequently, macrofossil records were used solely as evidence for the presence of spruce (*Picea* sp.). For records younger than 12 ^{14}C kyr BP, calibrated years were taken from the table CALIBRATED_AGE BP in the database. Older radiocarbon dates were recalibrated using OXCAL 4.2, based on the IntCal13 and Marine13 calibration curves (Reimer *et al.* 2013). Additional macrofossil and stomata records were included from recent studies or studies that are not integrated into the database. The compiled records are given in Appendix S1 of Supporting Information.

Pollen records were derived from the European Pollen Database (http://www.europeanpollendatabase.net/; see Appendix S2) and from the literature. A subset of pollen sites was taken from Latałowa & van der Knaap (2006). In the present study, we report ages when spruce pollen reached the 2% threshold for the first time, as we regard this as an acceptable marker for early population expansion. The possibilities and restrictions of this threshold are discussed in Latałowa & van der Knaap (2006).

Macrofossils and 2% pollen thresholds dated to 12–11 cal. kyr BP were carefully inspected to determine whether they belong to the late glacial or the Holocene. For plotting on maps, the palaeodata were divided into five time-slices: (1) > 27 cal. kyr BP, prior to the LGM; (2) 27 to > 18 cal. kyr BP, the local LGM when the Scandinavian and the Barents-Kara ice sheet were at their southern and south-western maximum, respectively (Clark *et al.* 2009); (3) 18 to > 11.7 cal. kyr BP, the late glacial, including the Younger Dryas, when the climate warmed and the ice retreated (Svendsen *et al.* 2004); the Holocene, divided into (4) 11.7 to > 6 cal. kyr BP and (5) 6 to 0 cal. kyr BP. In regions where sites are geographically clustered, only the oldest sites are shown on the maps. Ages are consistently expressed in calibrated years before present (cal. yr BP).

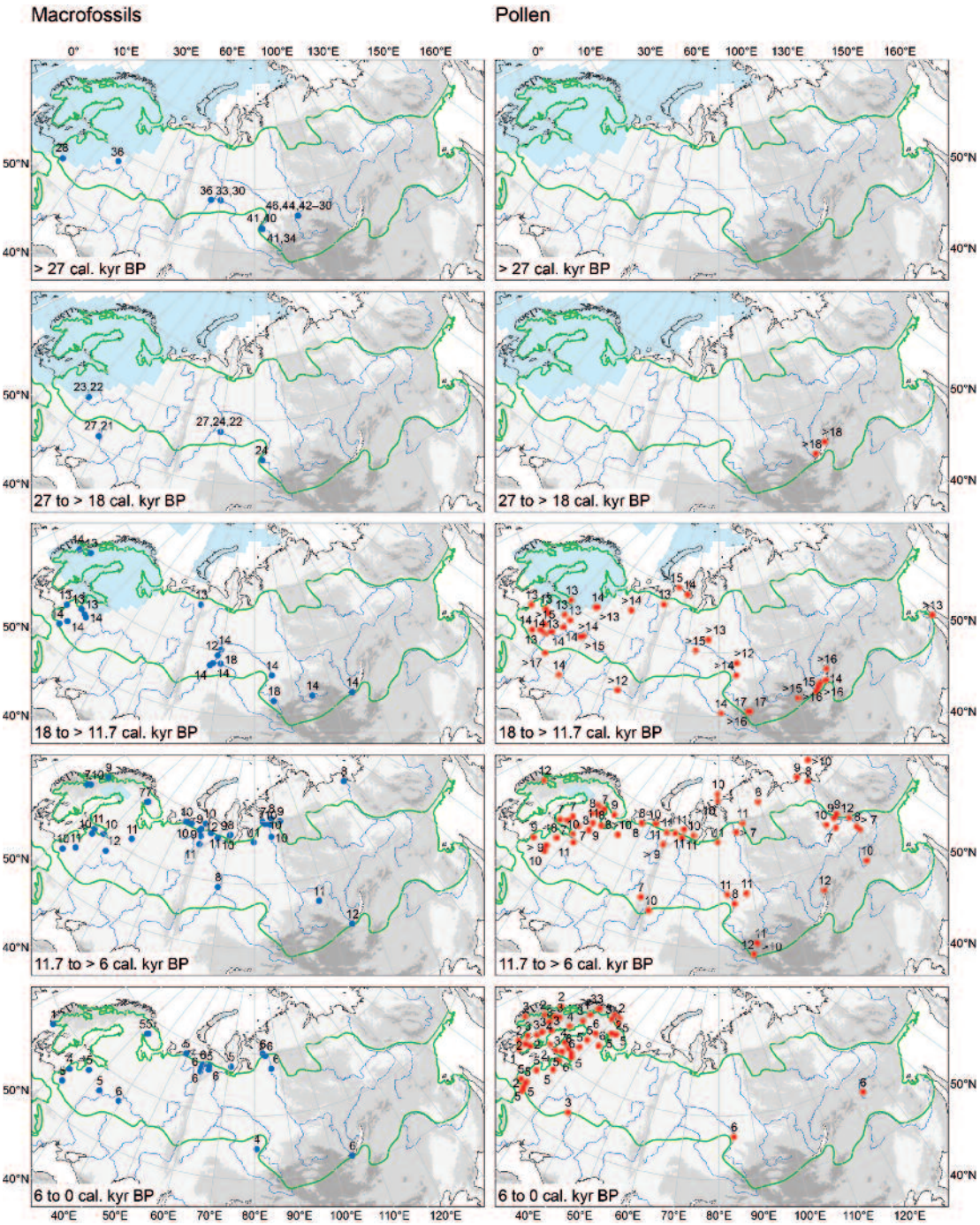
Genetic analyses

The genetic analyses were based on 551 individuals from 50 populations distributed throughout the boreal range of Norway spruce and the range of Siberian spruce (Appendix S3). All the sampled populations represented putatively natural forests. DNA was extracted from frozen needles using the DNeasy 96 Plant Kit or the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). A subset of the Norway spruce DNAs (32 populations) and mtDNA data was taken from two previous studies (Tollefsrud *et al.* 2008, 2009). The mtDNA data of Siberian spruce and the cpDNA data are new to this study.

Variation in mtDNA was assessed using the second intron of the *nad1* gene, which includes two variable minisatellites of 32 and 34 bp (Sperisen *et al.* 2001). The polymorphic fragment was amplified, sized and scored according to Tollefsrud *et al.* (2008). Sequencing was performed for 15 samples collected to the west of the Ural Mountains, and 17 samples collected to the east. Additional sequences were taken from Tollefsrud *et al.* (2008) to obtain sequences for all the detected size variants. In a preliminary cpDNA screen, six intergenic spacer (IGS) regions were sequenced in eight samples collected on either side of the Ural Mountains. The IGS regions surveyed were *trn*H–*psb*A (for primer sequences, see Sang *et al.* 1997; Tate & Simpson, 2003), *trn*S–*trn*G (Hamilton, 1999),

FIGURE 2

Spruce macrofossils and pollen records from northern Eurasia in five time-slices (cal. kyr BP is calibrated thousand years before present): > 27 cal. kyr BP is the period prior to the LGM; 27 to > 18 cal. kyr BP represents the local LGM when the Scandinavian and the Barents-Kara ice sheet were at their southern and south-western maximum, respectively (Clark *et al.* 2009); 18 to > 11.7 cal. kyr BP is the late glacial including the Younger Dryas; 11.7 to > 6 cal. kyr BP is the first part of the Holocene; and 6 to 0 cal. kyr BP is the last part of the Holocene. Macrofossil sites: numbers given in millennia cal. kyr BP indicate the age of spruce macrofossils. In the maps of the lateglacial and Holocene, only the oldest ages are shown. Pollen sites: numbers given in millennia cal. kyr BP indicate when spruce pollen reached the 2% threshold for the first time. Ice sheets are shown in blue with approximate extent according to time period: extent at 21 cal. kyr BP in the > 27 cal. kyr BP and 27 to > 18 cal. kyr BP maps; at 15 cal. kyr BP in the 18 to > 11.7 cal. kyr BP maps; and at 9 cal. kyr BP in the 11.7 to > 6 cal. kyr BP maps (Peltier, 1994). The current combined distribution of Norway spruce and Siberian spruce is indicated by the green outline.



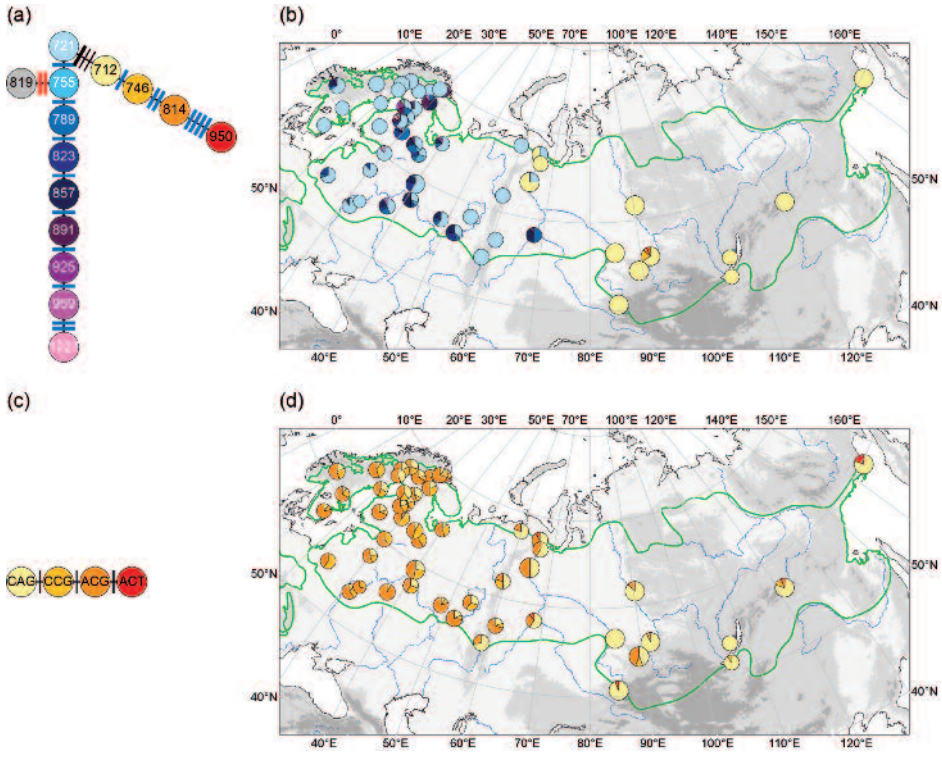
*trn*S–*trnf*M (Demesure *et al.* 1995) and *trn*T–*trn*F (Taberlet *et al.* 1991). Variation was detected in three regions, among which the most polymorphic (*trn*T–*trn*F) was genotyped in all samples. After amplification of the entire *trn*T–*trn*F region with primers 'a' and 'f' (Taberlet *et al.* 1991), a 593-bp fragment that included the *trn*T–*trn*L spacer and the *trn*L intron was sequenced with a newly designed forward primer (5'-GGAGGATAATAACATTGCAT-3') and the reverse primer 'd' (Taberlet *et al.* 1991). Mitochondrial and chloroplast sequences were assembled and refined manually with LASERGENE 7.7 (DNAStar, Madison, WI, USA) or with AUTOASSEMBLER (Applied Biosystems, Foster City, CA, USA). BIOEDIT (Hall, 1999) was used to align the sequences.

To establish the phylogenetic relationships among mitotypes and chlorotypes, a minimum spanning network was computed using ARLEQUIN 3.11 (Excoffier *et al.* 2005) and drawn in HAPSTAR (Teacher & Griffiths, 2011). In the alignment of the mitochondrial sequences, the two minisatellites were coded as separate multistate ordered characters (0–2 for the 32-bp minisatellite and 0–9 for the 34-bp minisatellite).

We applied spatial analyses of molecular variance (SAMOVA; using SAMOVA 1.0) to investigate the spatial patterns of genetic subdivision across the ranges of the two taxa and across the two markers (Dupanloup *et al.* 2002). This approach uses an iterative procedure to delineate contiguous groups of populations that are maximally differentiated. In the mitochondrial marker, each repeat copy of the two minisatellites was coded as a single character (base for presence and gap for absence). The analysis was carried out for *k* groups, where *k* ranged from 2 to 10. The grouping was considered to be optimal when the differentiation among groups (F_{CT}) reached a plateau and before single populations began to be delimited. Within-population gene diversity (H_S), total gene diversity (H_T) and genetic differentiation among populations (G_{ST}) were estimated according to Pons & Petit (1995). Because

FIGURE 3

Variation in the mitochondrial *nad1* gene and the chloroplast *trn*T–*trn*L fragment in spruce populations from northern Eurasia. The current combined distribution of Norway spruce and Siberian spruce is indicated by the green outline. (a) Minimum spanning network of the 14 mitotypes identified in the *nad1* minisatellite region. The mitotypes are indicated according their size in base pairs. Characters separating the mitotypes are shown as black bars (single-base substitutions) and boxes (indels); black box, 9-bp indel; blue boxes, indels of the 34-bp repeat; red boxes, indels of the 32-bp repeat. (b) Geographical distribution of the mitotypes. Their frequencies are represented in circle charts; chart size is proportional to the number of trees analysed, ranging from 7 to 16. (c) Minimum spanning network of the four chlorotypes identified in the *trn*T–*trn*L fragment. Chlorotypes are indicated with their polymorphic bases. Black bars indicate base substitutions. (d) Geographical distribution of chlorotypes.



genetic differentiation depends on the level of variation, the standardized genetic differentiation was calculated as $G'_{ST} = G_{ST}(1 + H_S) / (1 - H_S)$ following Hedrick (2005), thus allowing a comparison between GST of the mitochondrial and chloroplast markers, which exhibited very different levels of variation. These indices were calculated over all populations, and for groups of populations delineated by SAMOVA.

Results

Macrofossil and pollen data

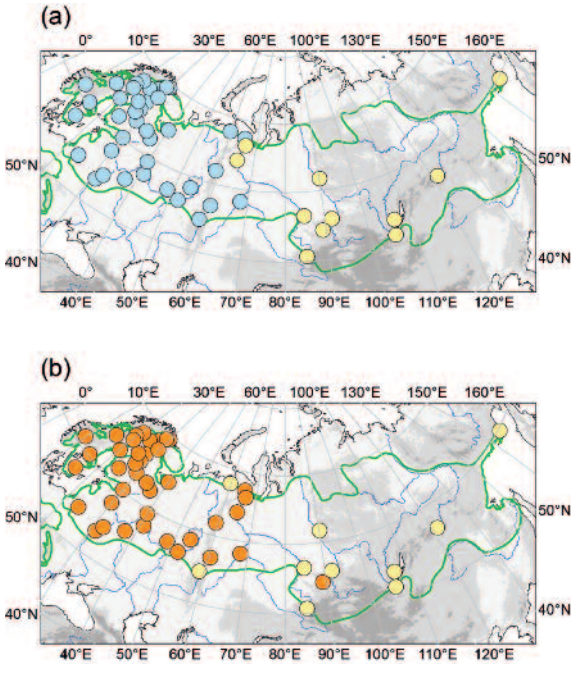
Macrofossils dated to the period before the LGM (> 27 cal. kyr BP) and to the LGM (27 to > 18 cal. kyr BP) were scarce, but indicated full-glacial presence of spruce in both northern Europe and Siberia (Fig. 2). Macrofossils dated to the LGM are reported at four sites, two on the East European Plain and two on the West Siberian Plain. The two sites on the West Siberian Plain were identical to the pre-LGM sites (Appendix S1), suggesting the long-term persistence of spruce in this area. Two additional occurrences, based on pollen dated to earlier than 18 cal. kyr BP, are described for the Baikal region.

Fossil records of the late glacial (18 to > 11.7 cal. kyr BP) were numerous and widespread, particularly the pollen records (Fig. 2). On the East European Plain, pollen records suggest that spruce established rapidly across large areas, including the Baltic countries, the northern basin of the Volga River and the area west of the northern Ural Mountains. Most pollen sites reached the 2% threshold between 15 and 13 cal. kyr BP, including those situated close to the northern Ural Mountains. In southern Siberia, pollen and macrofossils consistently showed early spruce occurrences on the West Siberian Plain, Altai and Sayan mountains and the Baikal region. The records from these areas are generally older than those from the East European Plain, ranging from 18 to 13 cal. kyr BP.

In northern Europe, range expansion as recorded at the > 2% pollen threshold was very slow up to the millennium 9–8 cal. kyr BP. Northward expansion towards Finland was fast, particularly from 7 to 5 cal. kyr BP, and spruce reached its modern northern limits in Finland around 2 cal. kyr BP. Spruce expanded westwards towards the Baltic Sea and colonized Scandinavia from 5 cal. kyr BP to the present millennium. The timing of northward colonization in Siberia is less clear because of a scarcity of data, particularly for central Siberia. Fossil pollen records indicate that spruce had already reached its current northern limits in Siberia during the early Holocene (11–8 cal. kyr BP), and also revealed spruce occurrences north of these limits.

FIGURE 4

Population structure of northern Eurasian spruce populations as revealed by spatial analysis of molecular variance (SAMOVA). Populations delineated to a particular group are presented in the same colour: (a) mtDNA groups; (b) cpDNA groups.



Genetic data

Sizing and sequencing of the mitochondrial *nad1* marker revealed 14 mitotypes in the 50 populations surveyed (Fig. 3, Appendix S3). Three mitotypes were not previously reported. The size variation was mainly a result of copy-number variation in the two minisatellites. Two nucleotide substitutions and a 9-bp indel defined one western and one eastern lineage (Fig. 3a). The four mitotypes of the eastern lineage were restricted to Siberia and two populations in the northern Ural Mountains. The 10 mitotypes of the western lineage were mainly found in northern and north-eastern Europe. Mitotype 721 bp and the mitotypes 789, 857 and 891 bp were also detected in two separate populations situated in the western part of the West Siberian Plain (Fig. 3b). One of these populations and one population from the northern Ural Mountains included mitotypes of both lineages.

The chloroplast *trn*T–*trn*L marker included three polymorphic sites, together defining four chlorotypes (CAG, CCG, ACT and ACG). Only the first character was parsimony-informative (Fig. 3c). The rare chlorotype ACT was only found in Siberia and in one population from the northern Ural Mountains. The three other chlorotypes occurred across northern Eurasia, but their frequencies varied considerably between eastern and western regions (Fig. 3d). For example, chlorotype CAG was abundant in Siberia, but rare in most of northern Europe, whereas chlorotype CCG was rare in Siberia but abundant in northern Europe.

Spatial analysis of molecular variance (SAMOVA) delineated two highly differentiated groups in both markers (mtDNA, F_{CT} = 0.69; cpDNA, F_{CT} = 0.44), one

western and one eastern (Fig. 4). In the case of the mtDNA, a slightly higher F_{CT} was obtained for k = 3 (F_{CT} = 0.73), but because the third group comprised two neighbouring populations in northernmost Fennoscandia, k = 2 was retained to capture the large-scale geographical structure across northern Eurasia. The two marker groups were largely congruent, with their division located to the east of the Ural Mountains. The division of the mtDNA variation was well defined and centred along the Ob River. This grouping roughly corresponds to the suggested ranges of Norway spruce and Siberian spruce, and coincides with LGM occurrences of spruce in northern Europe and Siberia. The western group also coincides with LGM populations on the West Siberian Plain.

On the northern Eurasian scale, population differentiation was higher for mtDNA (G_{ST} = 0.569) than for cpDNA (G_{ST} = 0.194), including when the level of variation was taken into account (Table 1). For assessing the genetic diversity within groups, we used the SAMOVA groups delineated by mtDNA, because these groups are geographically most structured. The within-population gene diversity was clearly higher in the western than in the eastern group, both in mtDNA (west, H_S = 0.332; east, H_S = 0.086) and cpDNA (west, H_S = 0.580; east, H_S = 0.281). In the western group, gene diversity was slightly higher when populations from formerly glaciated areas were excluded (Table 1).

Discussion

Genetic groups

Using cytoplasmic DNA markers, we identified two highly differentiated genetic groups, one western and one eastern, demonstrating for the first time that northern European Norway spruce and Siberian spruce should be considered two well-defined taxa. This finding resolves the long-term controversy on their distinction, which was primarily based on their high morphological similarity, the vast geographical cline in cone-scale shape (e.g. Schmidt-Vogt, 1974a; Popov, 2003) and their vague differentiation at allozyme loci (Krutovskii & Bergmann, 1995). The distinct east–west division of the mtDNA variation along the Ob River is consistent with the morphological data of Popov (2003). He described the most typical Siberian spruce cone-scales east of a line stretching from the Pechora River south-eastwards to the Ob River; in the area to the west, the scales gradually change to the shape of Norway spruce. Interestingly, the Ob River has also been found to constitute the border between Russian larch (*Larix sukaczewii*) and Siberian larch (Semerikov *et al.* 2007, 2013) as well as between several boreal animal species (Fedorov *et al.* 2008), suggesting that the same historical factors have been important to these taxa.

Glacial refugia

The combined palaeobotanical and genetic data strongly suggest that Norway spruce and Siberian spruce occupied separate LGM refugia – Norway spruce on the East European Plain and Siberian spruce in southern Siberia. Macrofossil records dated to the LGM and the widespread distribution of late glacial records on the East European Plain indicate that Norway spruce survived the LGM across vast areas, possibly including the area to the west of the northern Ural Mountains (cf. Väliranta *et al.* 2011). Consistent with a large refugium, current populations of the East European Plain were generally characterized by high mtDNA and cpDNA diversity, a pattern also reported for nuclear microsatellites (Tollefsrud *et al.* 2009). In contrast, populations in northern Scandinavia showed much lower diversity, both in mtDNA and microsatellites, reflecting the immigration history of spruce (Tollefsrud *et al.* 2009). The refugial population on the East European Plain may even have extended into the West Siberian Plain, where mitotypes of Norway spruce spatially coincide with LGM macrofossils (Figs 2 & 3).

Previous palaeobotanical studies, mainly based on pollen, have indicated that Siberian spruce persisted during the LGM in the southern part of the West Siberian Plain (e.g. Tarasov *et al.* 2000) and in the Baikal region, where large LGM populations were suggested (e.g. Bezrukova *et al.* 2005). The data of the present study lend further support to these suggestions and indicate that, during the late glacial, Siberian spruce was present in large parts of southern Siberia, namely the West Siberian Plain, Altai and Sayan mountains and the Baikal region. Most records from these areas are situated in the basins of the rivers Irtysh, Ob and Yenisei, whereas many records from the Holocene are located on interfluves. River floodplains may have been locally sheltered and/or provided the moisture required for tree growth (cf. Binney *et al.* 2009).

Post-glacial expansion

Based on pollen data, spruce started to expand on the East European Plain between 15 and 13 cal. kyr BP. No northward trend of decreasing age was detectable, suggesting that fossil pollen largely reflects local expansions. This supports the view of Väliranta *et al.* (2011) that expansion in the north-eastern region of European Russia is a result of local population expansion rather than migration from the south. In a former study, we hypothesized based on fossil cones with rounded scales that late-glacial expansions of spruce on the East European Plain involved Siberian spruce as the more cold-tolerant taxon, whereas Norway spruce was favoured during the Holocene expansion (Latałowa & van der Knaap, 2006). The genetic data of the present study do not support this hypothesis, but historical presence of Siberian spruce cannot be excluded. Following the argument of Schmidt-Vogt (1974a) that rounded cone-scales reflect adaptation to cold climate, the East European Plain may have harboured different Norway spruce ecotypes, some of them more cold-tolerant than others, and the cold-tolerant ecotypes may even have originated from introgression between Norway spruce and Siberian spruce during interglacials.

In Siberia, the pollen data show that the earliest expansions of spruce took place in the Baikal region and in the Altai and Sayan mountains (Fig. 2). Population expansion probably started in these areas concomitantly with the onset of climate warming after the LGM. The late-glacial expansion in westernmost Siberia possibly included both Norway spruce and Siberian spruce as suggested by mtDNA. The pollen data indicate that Siberian spruce reached the coast of the Kara and Laptev seas during the early Holocene. Many of these northernmost populations must, however, have declined again, because the northern limits of the modern Siberian spruce range are now further south (MacDonald *et al.* 2000; Binney *et al.* 2009). Whether northern populations expanded from southern or more northerly refugia cannot be inferred from our data. A survey of nuclear microsatellites and functional genes in six Siberian spruce populations located along the Yenisei River between 56° and 67° N latitude did not reveal any population structure, except for photoperiodic and circadian-clock genes which showed significant clinal variation and/or evidence of local selection (Chen *et al.* 2014). This pattern could indicate that Siberian spruce expanded from numerous source populations, including populations at high latitudes, but a wave-like front of population expansion from the south involving many individuals cannot be excluded.

Contact and introgression between Norway spruce and Siberian spruce

Unlike the mitotypes, the chlorotypes showed widespread admixture. Three of the four chlorotypes were widely distributed and were shared between the two taxa. Haplotype sharing can be a result of introgression and/or retention of ancestral polymorphisms. Generally, introgression is predicted to be less common for genes that experience high rates of gene flow than for those that experience little gene flow (Currat *et al.* 2008). The rationale is that high rates of gene flow help to dilute migrant alleles and thus to preserve the species' integrity. In our study, we observed a clear difference in overall population differentiation in the two markers (Table 1). As in other conifers (e.g. Du *et al.* 2009; Polezhaeva *et al.* 2010), population differentiation was lower in cpDNA than in mtDNA (Table 1), indicating greater gene flow in the former. Accordingly, it seems unlikely that the widespread distribution of the three chlorotypes solely reflects introgression. Rather, the retention of ancestral polymorphism, which is frequently observed between closely related *Picea* species (Li *et al.* 2010), seems to be the main cause for their widespread distribution. The rare chlorotype ACT, which was restricted to Siberian populations and one population in the northern Ural Mountains, may represent a more recent mutation.

In the north-western part of the West Siberian Plain, two populations exhibited mitotypes of both mitochondrial lineages, which strongly suggests that Norway spruce and Siberian spruce are indeed in contact today. The palaeobotanical data show that spruce was already present in this region during the late glacial, suggesting that the two taxa came into contact soon after the LGM. Based on morphological data, the north-western part of the West Siberian Plain is characterized by Siberian spruce (Popov, 2003). It is therefore not unlikely that mtDNA of Siberian spruce was replaced by that of Norway spruce as a consequence of recurrent hybridization. Replacement of mtDNA is also conceivable further west, in the Ural Mountains and the Pechora region (cf. Fig. 1). However, because the east–west division of the cpDNA variation was largely congruent with that of mtDNA, the zone of introgression may be narrower than suggested by morphological criteria (e.g. Pravdin, 1975; Popov, 2003).

TABLE 1

Genetic diversity within populations (H_S), total genetic diversity (H_T) and population differentiation estimated as G_{ST} and standardized G'_{ST} (Hedrick, 2005) for mtDNA and cpDNA of spruce populations of northern Eurasia, both globally and for the SAMOVA groups delineated by mtDNA. The western group consists of populations from the East European Plain, Fennoscandia, the Ural Mountains and a population on the West Siberian Plain. Population-genetic parameters are also given for the western group, excluding populations from formerly glaciated areas. The eastern group consists of populations from Siberia. Standard errors are given in parentheses; n.c., not calculated because of low variation among populations.

Grouping of populations	Mitochondrial DNA				Chloroplast DNA			
	H_S	H_T	G_{ST}	G'_{ST}	H_S	H_T	G_{ST}	G'_{ST}
Northern Eurasia (n = 50)	0.278 (0.047)	0.644 (0.049)	0.569 (0.064)	1.001	0.514 (0.027)	0.638 (0.009)	0.194 (0.044)	0.605
Western group (n = 39)	0.332 (0.056)	0.492 (0.071)	0.326 (0.067)	0.649	0.580 (0.019)	0.606 (0.017)	0.043 (0.030)	0.163
Western group without glaciated areas (n = 14)	0.354 (0.094)	0.493 (0.110)	0.281 (0.073)	0.589	0.534 (0.044)	0.626 (0.026)	0.141 (0.052)	0.464
Eastern group (n = 11)	0.086 (0.058)	0.105 (0.067)	0.180 n.c.	0.214	0.281 (0.066)	0.347 (0.080)	0.192 n.c.	0.342

Contrasting patterns of genetic diversity

The genetic diversity in both mtDNA and cpDNA was found to be clearly higher in western than in eastern populations, a pattern that also has been observed in allozymes, albeit less pronouncedly (Krutovskii & Bergmann, 1995). In the case of mtDNA, all but one of the populations of the eastern mitochondrial lineage were fixed for the 712-bp mitotype. The three other mitotypes, each containing several copies of the 34-bp minisatellite, were found in a single population located in the northern foothills of the southern Siberian mountains. Assuming stepwise copy-number changes in the minisatellite, it seems likely that ancestral populations carried more mitotypes.

The higher genetic diversity in Norway spruce than in Siberian spruce could have two causes – less pronounced fluctuations in population size of Norway spruce due to more stable and suitable climatic conditions during the Quaternary and/or gene flow from Siberian spruce. It seems unlikely that the clear difference in diversity is solely a result of introgression. Diversities of mtDNA and cpDNA are both quite high in the western part of the East European Plain, an area that has not been suggested to be influenced by introgression (cf. Fig. 1). Glaciological evidence indicates that the climate of the East European Plain was rather humid during glacial maxima, whereas that of Siberia was very dry, particularly during periods when the Scandinavian ice sheet was large, such as during the LGM, absorbing much of the moisture spread from the Atlantic Ocean (Stauch & Gualtieri, 2008; Krinner *et al.* 2011). A cold and extremely dry LGM climate in Siberia and much milder conditions on the East European Plain are also supported by palaeobotanical data (Hubberten *et al.* 2004) and palaeovegetation maps (Allen *et al.* 2010). Siberian spruce may thus have repeatedly experienced reductions in population size and consequently in genetic diversity, whereas Norway spruce was less affected. A similar pattern of reduced diversity into the continent was observed in *Abies* in the Russian Far East, where a more stable climate along the coast was suggested to have prevented major demographic fluctuations (Semerikova *et al.* 2011). It seems possible that glaciations also played a role, particularly during the Saalian stage, when glaciers in Siberia were more extensive (Stauch & Gualtieri, 2008). Indeed, climatic events pre-dating the LGM were suggested to have influenced the genetic structure of Siberian larch (Semerikov *et al.* 2013) and several Eurasian boreal animal species (Fedorov *et al.* 2008).

Conclusions

The combined palaeobotanical and genetic data allowed us to draw the following major conclusions. (1) North Eurasian Norway spruce and Siberian spruce are genetically distinct and occupied separate LGM refugia. Norway spruce persisted during the LGM on the East European Plain, where it was already widely distributed during the late glacial, suggesting widespread LGM tree occurrences. Siberian spruce persisted in southern Siberia, but may also have survived locally at high latitudes, facilitating rapid northward expansion. (2) The two taxa came into contact in the area of the Ob River, where they probably hybridized. (3) The higher cytoplasmic DNA diversity in Norway spruce than in Siberian spruce may be a consequence of more favourable past climatic conditions for Norway spruce and/or introgression from Siberian spruce.

Our survey of cytoplasmic DNA variation did not provide evidence for population subdivision either on the East European Plain or in Siberia. Multilocus surveys with nuclear markers will be necessary to provide further insights into the structure of refugial populations. Nuclear markers could also help to delineate the introgression zone and to elucidate patterns of past and current gene flow in this zone. Furthermore, variation in functional genes could be used to determine whether the gradient in cone-scale morphology is associated with introgression and/or climatic adaptation.

Acknowledgements

We thank Paolo Cherubini, Thomas Geburek, Maxim Kapralov, Daniel Nievergelt, Andreas Rigling and Fritz Schweingruber for providing spruce samples from Russia. We are grateful to the staff at the ABI lab at the University of Oslo for producing excellent DNA sequences. We also thank Heather Binney for providing valuable information about the structure of the Eurasian Macrofossil Database and Arne Steffenrem, who assisted with extraction of data from the database. Marcelina Zimny is acknowledged for her help in preparing tables and references of palaeobotanical data, György Sipos for critically reading and commenting on the manuscript, Donatella Magri and an anonymous referee for their helpful comments. Finally, we thank the data contributors of the EMD and EPD. This study was supported by the Norwegian Institute of Bioeconomy Research, the Research Council of Norway (Strategic Institute Programme 501500: DNA markers for characterization of genetic variation in Norwegian Forest trees), and the EU network EVOLTREE; the work done by M.L. was based on statutory funds in the Faculty of Biology, University of Gdańsk (DS/530-L145-D486-14).

REFERENCES

Allen, J.R.M., Hickler, T., Singarayer, J.S., Sykes, M.T., Valdes, P.J. & Huntley, B. (2010) Last glacial vegetation of northern Eurasia. *Quaternary Science Reviews*, **29**, 2604–2618.

Bezrukova, E.V., Abzaeva, A.A., Letunova, P.P., Kulagina, N.V., Vershinin, K.E., Belov, A.V., Orlova, L.A., Danko, L.V. & Krapivina, S.M. (2005) Post-glacial history of Siberian spruce (*Picea obovata*) in the Lake Baikal area and the significance of this species as a paleo-environmental indicator. *Quaternary International*, **136,** 47–57.

Binney, H.A., Willis, K.J., Edwards, M.E., Bhagwat, S.A., Anderson, P.M., Andreev, A.A., Blaauw, M., Damblon, F., Haesaerts, P., Kienast, F., Kremenetski, K.V., Krivonogov, S.K., Lozhkin, A.V., MacDonald, G.M., Novenko, E.Y., Oksanen, P., Sapelko, T.V., Väliranta, M. & Vazhenina, L. (2009) The distribution of late-Quaternary woody taxa in northern Eurasia: evidence from a new macrofossil database. *Quaternary Science Reviews*, **28**, 2445–2464.

Birks, H.H. & Birks, H.J.B. (2000) Future uses of pollen analysis must include plant macrofossils. *Journal of Biogeography*, **27**, 31–35.

Birks, H.H., Giesecke, T., Hewitt, G.M., Tzedakis, P.C., Bakke, J. & Birks, H.J.B. (2012) Comment on "Glacial survival of boreal trees in northern Scandinavia". *Science*, **338**, 742–743.

Blyakharchuk, T.A., Wright, H.E., Borodavko, P.S., van der Knaap, W.O. & Ammann, B. (2004) Late Glacial and Holocene vegetational changes on the Ulagan high-mountain plateau, Altai Mountains, southern Siberia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **209**, 259–279.

Chen, J., Tsuda, Y., Stocks, M., Källman, T., Xu, N-N., Kärkkäinen, K., Huotari, T., Semerikov, V.L., Vendramin, G.G. & Lascoux, M. (2014) Clinal variation at phenology-related genes in spruce: parallel evolution in FTL2 and Gigantea? *Genetics*, **197**, 1025–1038.

Clark, P.U., Dyke, A.S., Shakun, J.D., Carlson, A.E., Clark, J., Wohlfarth, B., Mitrovica, J.X., Hostetler, S.W. & McCabe, A.M. (2009) The Last Glacial Maximum. *Science*, **325**, 710–714.

Collignon, A.-M., Van de Sype, H. & Favre, J.-M. (2002) Geographical variation in random amplified polymorphic DNA and quantitative traits in Norway spruce. *Canadian Journal of Forest Research*, **32**, 266–282.

Currat, M., Ruedi, M., Petit, R.J. & Excoffier, L. (2008) The hidden side of invasions: massive introgression by local genes. *Evolution*, **62**, 1908–1920.

Demesure, B., Sodzi, N. & Petit, R.J. (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology*, 4, 129–131.

Du, F-K., Petit, R.J. & Liu, J-Q. (2009) More introgression with less gene flow: chloroplast vs. mitochondrial DNA in the Picea asperata complex in China, and comparison with other conifers. *Molecular Ecology*, **18**, 1396–1407.

Dupanloup, I., Schneider, S. & Excoffier, L. (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**, 2571–2581.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Fedorov, V.B., Goropashnaya, A.V., Boeskorov, G.G. & Cook, J.A. (2008) Comparative phylogeography and demographic history of the wood lemming (*Myopus schisticolor*): implications for late Quaternary history of the taiga species in Eurasia. *Molecular Ecology*, **17**, 598–610.

Frenzel, B. (1968) The Pleistocene vegetation of northern Eurasia. *Science*, **161**, 637–649.

Giesecke, T. & Bennett, K.D. (2004) The Holocene spread of Picea abies (L.) Karst. in Fennoscandia and adjacent areas. *Journal of Biogeography*, **31**, 1523–1548.

Grivet, D., Jeandroz, S. & Favre, J.-M. (1999) *Nad1* b/c intron polymorphism reveals maternal inheritance of the mitochondrial genome in Picea abies. *Theoretical and Applied Genetics*, **99**, 346–349.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Hamilton, M.B. (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology*, **8**, 521–523.

Hedrick, P.W. (2005) A standardized genetic differentiation measure. *Evolution,* **59**, 1633–1638.

Heuertz, M., De Paoli, E., Källman, T., Larsson, H., Jurman, I., Morgante, M., Lascoux, M. & Gyllenstrand, N. (2006) Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce [*Picea abies* (L.) Karst]. *Genetics*, **174**, 2095–2105.

Hubberten, H.W., Andreev, A., Astakhov, V.I. *et al.* (2004) The periglacial climate and environment in northern Eurasia during the Last Glaciation. *Quaternary Science Reviews*, **23**, 1333–1357.

Krinner, G., Diekmann, B., Colleoni, F. & Stauch, G. (2011) Global, regional and local scale factors determining glaciation extent in Eastern Siberia over the last 140,000 years. *Quaternary Science Reviews*, **30**, 821–831.

Krutovskii, K.V. & Bergmann, F. (1995) Introgressive hybridization and phylogenetic relationships between Norway, *Picea abies* (L.) Karst., and Siberian, *P. obovata* Ledeb., spruce species studied by isozyme loci. *Heredity*, **74**, 464–480.

Kullman, L. (2008) Early postglacial appearance of tree species in northern Scandinavia: review and perspective. *Quaternary Science Reviews*, **27**, 2467–2472.

Lagercrantz, U. & Ryman, N. (1990) Genetic structure of Norway spruce (*Picea abies*): concordance of morphological and allozymic variation. *Evolution*, **44**, 38–53.

Latałowa, M. & van der Knaap, W.O. (2006) Late Quaternary expansion of Norway spruce *Picea abies* (L.) Karst. in Europe according to pollen data. *Quaternary Science Reviews*, **25**, 2780–2805.

Li, Y., Stocks, M., Hemmilä, S., Källman, T., Zhu, H-T., Zhou, Y-F., Chen, J., Liu, J-Q. & Lascoux, M. (2010) Demographic histories of four spruce (*Picea*) species of the Qinghai-Tibetan Plateau and neighboring areas inferred from multiple nuclear loci. *Molecular Biology and Evolution*, **27,** 1001–1014.

Lockwood, J.D., Aleksić, J.M., Zou, J-B., Wang, J., Liu, J-Q. & Renner, S.S. (2013) A new phylogeny for the genus *Picea* from plastid, mitochondrial, and nuclear sequences. *Molecular Phylogenetics and Evolution*, **69**, 717-727.

MacDonald, G.M., Velichko, A.A., Kremenetski, C.V., Borisova, O.K., Goleva, A.A., Andreev, A.A., Cwynar, L.C., Riding, R.T., Forman, S.L., Edwards, T.W.D., Aravena, R., Hammarlund, D., Szeicz, J.M. & Gataullin, V.N. (2000) Holocene treeline history and climate change across northern Eurasia. *Quaternary Research*, **53**, 302–311.

Parducci, L., Jørgensen, T., Tollefsrud, M.M. *et al.* (2012a) Glacial survival of boreal trees in northern Scandinavia. *Science*, **335**, 1083–1086.

Parducci, L., Edwards, M.E., Bennett, K.D., Alm, T., Elverland, E., Tollefsrud, M.M., Jørgensen, T., Houmark-Nielsen, M., Larsen, N.K., Kiær, K.H., Fontana S.L., Alsos I.G. & Willerslev,

E. (2012b) Response to comment on “Glacial survival of boreal trees in northern Scandinavia”. *Science*, **338**, 742.

Peltier, W.R. (1994) Ice age paleotopography. *Science*, **265**, 195–201.

Polezhaeva, M.A., Lascoux, M. & Semerikov, V.L. (2010) Cytoplasmic DNA variation and biogeography of *Larix* Mill. in northeast Asia. *Molecular Ecology*, **19**, 1239–1252.

Pons, O. & Petit, R.J. (1995) Estimation, variance and optimal sampling of gene diversity. I. Haploid locus. *Theoretical and Applied Genetics*, **90**, 462–470.

Popov, P.P. (2003) Structure and differentiation of spruce populations in eastern Europe and western Siberia. *Russian Journal of Ecology*, **34**, 27–33.

Popov, P.P. (2010) Form structure and geographic differentiation of spruce populations in northwestern Russia. *Russian Journal of Ecology*, **41**, 336–343.

Pravdin, L.F. (1975) *European spruce and Siberian spruce in the USSR* (in Russian). Nauka, Moskow.

Ran, J-H., Wei, X-X. & Wang, X-Q. (2006) Molecular phylogeny and biogeography of *Picea* (Pinaceae): implications for phylogeographical studies using cytoplasmic haplotypes. *Molecular Phylogenetics and Evolution*, **41**, 405–419.

Reimer, P.J., Bard, E., Bayliss, A. *et al.* (2013) IntCal13 and Marine13 radiocarbon age calibration curves, 0–50,000 years cal BP. *Radiocarbon*, **55**, 1869–1887.

Sang, T., Crawford, D.J. & Stuessy, T.F. (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany*, **84**, 1120–1136.

Schmidt-Vogt, H. (1974a) Die systematische Stellung der gemeinen Fichte (*Picea abies* [L.] Karst.) und der sibirischen Fichte (*Picea obovata* Ledeb.) in der Gattung *Picea*. Ein Beitrag zur Einführung von Erkenntnissen auf dem Gebiet forstlicher Genökologie in die Systematik der Baumarten. *Allgemeine Forst- und Jagdzeitung*, **145**, 45–60.

Schmidt-Vogt, H. (1974b) Das natürliche Verbreitungsgebiet der Fichte (*Picea abies* [L.] Karst.) in Eurasien. *Allgemeine Forst- und Jagdzeitung*, **145**, 185–197.

Schmidt-Vogt, H. (1977) *Die Fichte*. Verlag Paul Parey, Hamburg.

Semerikov, V.L., Iroshnikov, A.I. & Lascoux, M. (2007) Mitochondrial DNA variation pattern and postglacial history of the Siberian larch (*Larix sibirica* Ledeb.). *Russian Journal of Ecology*, **38**, 147–154.

Semerikov, V.L., Semerikova, S.A., Polezhaeva, M.A., Kosintsev, P.A. & Lascoux, M. (2013) Southern montane populations did not contribute to the recolonization of West Siberian Plain by Siberian larch (*Larix sibirica*): a range-wide analysis of cytoplasmic markers. *Molecular Ecology*, **22**, 4958–4971.

Semerikova, S.A., Semerikov, V.L. & Lascoux, M. (2011) Post-glacial history and introgression in *Abies* (Pinaceae) species of the Russian Far East inferred from both nuclear and cytoplasmic markers. *Journal of Biogeography*, **38**, 326–340.

Sperisen, C., Büchler, U., Gugerli, F., Mátyás, G., Geburek, T. & Vendramin, G.G. (2001) Tandem repeats in plant mitochondrial genomes: application to the analysis of population differentiation in the conifer Norway spruce. *Molecular Ecology*, **10**, 257–263.

Stauch, G. & Gualtieri, L. (2008) Late Quaternary glaciations in northeastern Russia. *Journal of Quaternary Science*, **23**, 545–558.

Sutton, B.C.S., Flanagan, D.J., Gawley, J.R., Newton, C.H., Lester, D.T. & Elkassaby, Y.A. (1991) Inheritance of chloroplast and mitochondrial DNA in *Picea* and composition of hybrids from introgression zones. *Theoretical and Applied Genetics*, **82**, 242–248.

Svendsen, J.I., Alexanderson, H., Astakhov, V.I. *et al.* (2004) Late Quaternary ice sheet history of northern Eurasia. *Quaternary Science Reviews*, **23**, 1229–1271.

Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, **17**, 1105–1109.

Tarasov, P.E., Volkova, V.S., Webb, T., Guiot, J., Andreev, A.A., Bezusko, L.G., Bezusko, T.V., Bykova, G.V., Dorofeyuk, N.I., Kvavadze, E.V., Osipova, I.M., Panova, N.K. & Sevastyanov, D.V. (2000) Last glacial maximum biomes reconstructed from pollen and plant macrofossil data from northern Eurasia. *Journal of Biogeography*, **27**, 609–620.

Tate, J.A. & Simpson, B.B. (2003) Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany*, **28**, 723–737.

Teacher, A.G.F. & Griffiths, D.J. (2011) HapStar: automated haplotype network layout and visualization. *Molecular Ecology Resources*, **11**, 151–153.

Terhürne-Berson, R. (2005) *Changing distribution patterns of selected conifers in the Quaternary of Europe caused by climatic variations*. PhD Thesis, University of Bonn, Bonn.

Tollefsrud, M.M., Kissling, R., Gugerli, F., Johnsen, Ø., Skrøppa, T., Cheddadi, R., van der Knaap, W.O., Latałowa, M., Terhürne-Berson, R., Litt, T., Geburek, T., Brochmann, C. & Sperisen, C. (2008) Genetic consequences of glacial survival and postglacial colonization in Norway spruce: combined analysis of mitochondrial DNA and fossil pollen. *Molecular Ecology*, **17**, 4134–4150.

Tollefsrud, M.M., Sønstebø, J.H., Brochmann, C., Johnsen, Ø., Skrøppa, T. & Vendramin, G.G. (2009) Combined analysis of nuclear and mitochondrial markers provide new insight into the genetic structure of North European *Picea abies*. *Heredity*, **102**, 549–562.

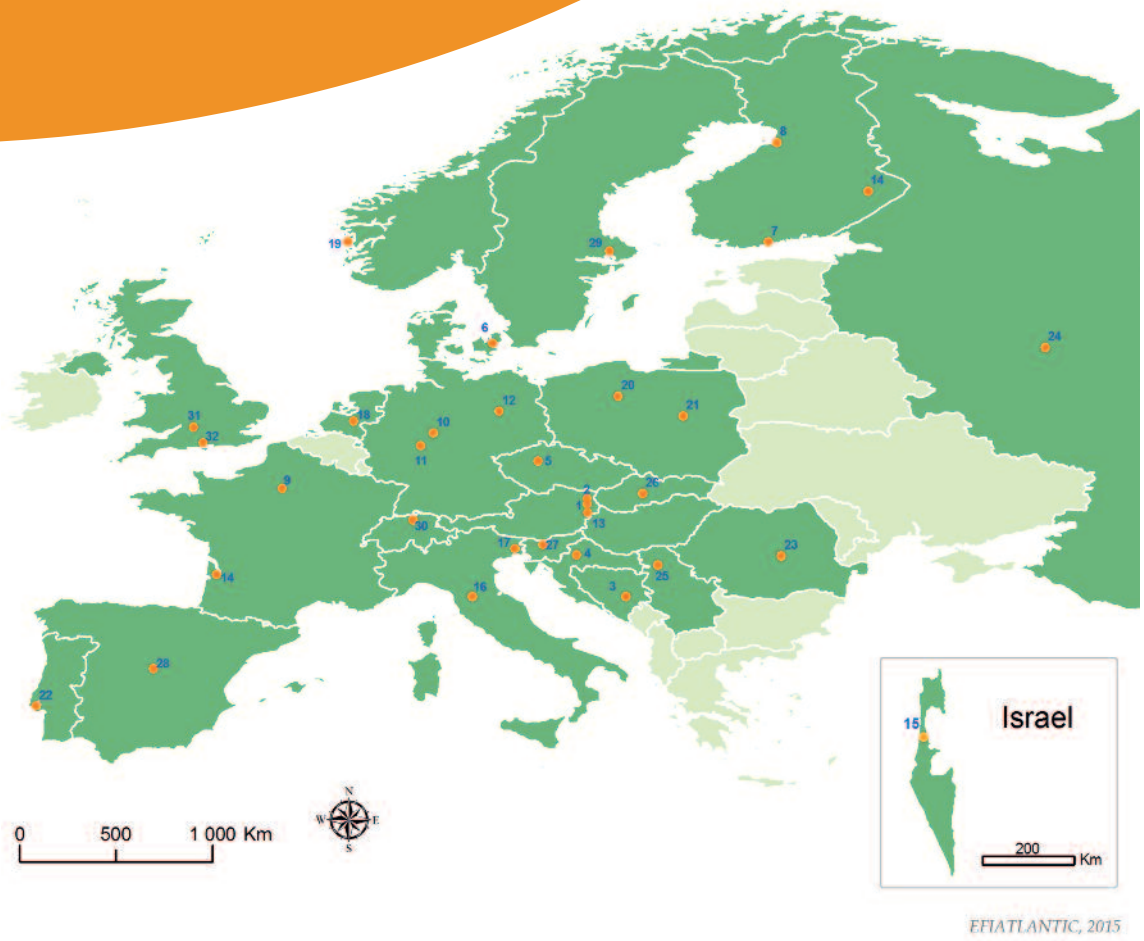
Tutin, T.G., Burges, N.A., Chater, A.O., Edmonds, J.R., Heywood, D.H., Moore, D.M., Valentine, D.H., Walters, S.M. & Webb, D.A. (1993) *Flora Europaea*, 2nd edn. Cambridge University Press, Cambridge, UK.

Väliranta, M., Kaakinen, A., Kuhry, P., Kultti, S., Salonen, J.S. & Seppä, H. (2011) Scattered late-glacial and early Holocene tree populations as dispersal nuclei for forest development in north-eastern European Russia. *Journal of Biogeography*, **38**, 922–932.

Vendramin, G.G., Anzidei, M., Madaghiele, A., Sperisen, C. & Bucci, G. (2000) Chloroplast microsatellite analysis reveals the presence of population subdivision in Norway spruce (*Picea abies* K.). *Genome*, **43**, 68–78.

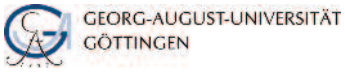
Vorren, T.O., Vorren, K.-D., Aasheim, O., Dahlgren, K.I.T., Forwick, M. & Hassel, K. (2013) Palaeoenvironment in northern Norway between 22.2 and 14.5 cal. ka BP. *Boreas*, **42**, 876–895.

EVOLTREE PARTNERS



In green: core partners *(i.e., partners who provide financial contribution to the network)*
In blue: associate partners *(i.e., partners who provide in-kind contribution to the network)*

1. Austrian Institute of Technology GmbH (AIT)
2. Bundesforschungszentrum für Wald Institut für Waldgenetik (BWF), Austria
3. University of Sarejevo Faculty of Forestry (UNSA), Bosnia and Herzegovina
4. Croatian Forest Research Institute (CFRI)
5. Czech University of Life Sciences (CULS)
6. University of Copenhagen (UCPH), Denmark
7. Natural Resources Institute Finland (LUKE)
8. University of Oulu (UOULU), Finland
9. National Institute for Agricultural Research (INRA), France
10. Georg-August-University Göttingen (UGOE), Germany
11. Phillips University of Marburg (UNIMAR), Germany
12. Johann Heinrich von Thünen Institut, Germany
13. University of West Hungary Faculty of Forestry (UWH)
14. European Forest Institute (EFI), International
15. Institute of Plant Sciences Agriculture Research Organisation (ARO), Israel
16. National Research Council Plant Genetics Institute (CNR-IBBR), Italy
17. Institute for Applied Genomics (IGA-UNIUD), Italy
18. Wageningen UR (WUR), The Netherlands
19. Norwegian Institute of Bioeconomy Research (NIBIO)
20. University of Bydgoszcz (UKW), Poland
21. The Forest Research Institute (IBL), Poland
22. Universidade de Lisboa Instituto Superior de Agronomia (ISA), Portugal
23. Universitatea Transilvania Brasov (UNITBV), Romania
24. Russian Academy of Sciences Vavilov Institute of General Genetics (VIGG)
25. Institute of Lowland Forestry and Environment (ILFE), Serbia
26. Technical University of Zvolen (TUZVO), Slovakia
27. Slovenian Forestry Institute (SFI)
28. National Institute for Agriculture, Food Research and Technology (INIA), Spain
29. Uppsala University (UU), Norway
30. Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Switzerland
31. Centre for Ecology and Hydrology (NERC), UK
32. University of Southampton (SOTON), UK



The EVOLTREE coordination team would like to thank all EVOLTREE participants, past and present, who have supported the network throughout the last 10 years by contributing to its successful work and activities. We are also grateful to the European Commission for funding during the period as Network of Excellence (2006-2010) under FP6 and to all institutional EVOLTREE members for self-funding after that period.

EFI

THE EVOLTREE NETWORK
IS A EUROPEAN RESEARCH GROUP
FACILITATED BY THE EUROPEAN
FOREST INSTITUTE (EFI)

EVOLTREE/EFIATLANTIC
SITE DE RECHERCHE FORÊT-BOIS / 69 ROUTE D'ARCACHON / 33612 CESTAS / FRANCE