



HAL
open science

Limitation of gene flow by distance in the common yellow jasmine (*Chrysojasminum fruticans*, Oleaceae): implications for the study of its mating strategies

Aurore Puyoou, Hervé Gryta, Anne-Laure Fuchs, Pierrick Blanchard, Laure Civeyrel, Pierre-Olivier Cheptou, Mathilde Dufaÿ, Julia Dupin, Patricia Jargeat, Emilie Lecompte, et al.

► **To cite this version:**

Aurore Puyoou, Hervé Gryta, Anne-Laure Fuchs, Pierrick Blanchard, Laure Civeyrel, et al.. Limitation of gene flow by distance in the common yellow jasmine (*Chrysojasminum fruticans*, Oleaceae): implications for the study of its mating strategies. *Botanical Journal of the Linnean Society*, inPress, 10.1093/botlinnean/boad035 . hal-04133393

HAL Id: hal-04133393

<https://hal.science/hal-04133393>

Submitted on 19 Jun 2023

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.



**Limitation of gene flow by distance in the common yellow
jasmine (*Chrysojasminum fruticosum*, Oleaceae): implications
for the study of its mating strategies**

Journal:	<i>Botanical Journal of the Linnean Society</i>
Manuscript ID	BOTJLS-Jan-2023-4211-ART.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	14-Apr-2023
Complete List of Authors:	<p>Puyou, Aurore; Université Paul Sabatier - Complexe Scientifique de Rangueil, UMR5174, Laboratoire Evolution et Diversité Biologique Gryta, Hervé; Université Toulouse III Paul Sabatier, UMR5174, Laboratoire Evolution et Diversité Biologique Fuchs, Anne-Laure; Université Toulouse III Paul Sabatier, UMR5174, Laboratoire Evolution et Diversité Biologique Blanchard, Pierrick; Université Toulouse III Paul Sabatier, UMR5174, Laboratoire Evolution et Diversité Biologique Cheptou, Pierre-Olivier; CEFE, CNRS Civeyrel, Laure; Université Toulouse III Paul Sabatier, UMR5174, Laboratoire Evolution et Diversité Biologique Dufay, Mathilde; Université Montpellier Faculté des Sciences de Montpellier, UMR5175, CEFE (Centre d'Ecologie Fonctionnelle et Evolutive) Dupin, Julia; Université Paul Sabatier, Laboratoire Evolution et Diversité Biologique, UMR 5174 Jargeat, Patricia; Université Toulouse III Paul Sabatier, UMR5174, Laboratoire Evolution et Diversité Biologique Lecompte, Emilie; Université Toulouse III Paul Sabatier, UMR5174, Evolution & Diversité Biologique Besnard, Guillaume; CNRS, Université Paul Sabatier, EDB, Laboratoire Evolution & Diversité Biologique</p>
Keywords:	cpDNA < Genetics, gene flow < Genetics, genetic differentiation < Genetics, genetic structure < Genetics, microsatellites < Genetics, mitochondria < Genetics, self-incompatibility < Genetics, Mediterranean < Geography

SCHOLARONE™
Manuscripts

1
2
3 **Limitation of gene flow by distance in the common yellow jasmine**
4
5
6 **(*Chrysojasminum fruticans*, Oleaceae): implications for the study of its**
7
8
9 **mating strategies**
10
11
12
13
14
15

16 Running title: Population genetics of a yellow jasmine
17
18
19
20

21 AUTHORS:
22
23
24

25 Aurore Puyou¹, Hervé Gryta¹, Anne-Laure Fuchs¹, Pierrick Blanchard¹, Pierre-
26
27 Olivier Cheptou², Laure Civeyrel¹, Mathilde Dufay², Julia Dupin¹, Patricia
28
29 Jargeat¹, Emilie Lecompte¹, Guillaume Besnard^{1*}
30
31
32
33
34
35
36

37 ¹*Laboratoire Evolution et Diversité Biologique (EDB), UMR 5174, CNRS, IRD, Université*
38
39 *Toulouse Paul Sabatier, 31062 Toulouse, France*
40
41

42 ²*CEFE (Centre d'Ecologie Fonctionnelle et Evolutive), UMR 5175, CNRS, Université de*
43
44 *Montpellier, Université Paul Valéry, EPHE, IRD, 34293 Montpellier, France*
45
46
47
48
49

50
51 * Corresponding author: Guillaume Besnard, guillaume.besnard@univ-tlse3.fr
52
53
54
55
56
57
58
59
60

ABSTRACT

The common yellow jasmine (*Chrysojasminum fruticans*; Oleaceae) is a distylous shrub occurring in the wild in southwestern Europe and the Mediterranean Basin. Little is known about the genetics of its populations and such information would be necessary to investigate its spread and mating strategies. Here, the organization of its genetic diversity was investigated among and between 13 populations from southern France, including a 35-years old experimental plot ('CEFE', CNRS Montpellier). Markers (microsatellites and indels) were developed to screen polymorphisms in nuclear, chloroplast, and mitochondrial genomes. Low linkage disequilibrium was observed between chloroplast and mitochondrial haplotypes likely resulting from paternal leaks in their inheritance as reported in other species of tribe Jasmineae. Yet, analyses of 36 progenies issued from parents with distinct chloroplast and/or mitochondrial DNA haplotypes only revealed a maternal contribution. Natural populations of *C. fruticans* are moderately to highly differentiated at the regional scale with a strong isolation-by-distance pattern detected on nuclear data, indicating limited gene flow. An isolated site ('Moulis'), located on the marginal distribution area, was remarkably genetically depauperate and highly differentiated from other populations. Further studies on the variation of mating strategies in *C. fruticans* should consider populations with contrasting patterns of genetic diversity. The artificial 'CEFE' population also offers opportunities for experiments in a closed system.

ADDITIONAL KEYWORDS: Chloroplast genome – Heterostyly – IBD – Jasmineae – Microsatellite – Mitochondrial genome – Organelle paternal leakage – Population genetics – Self-incompatibility.

INTRODUCTION

Many ecological and evolutionary processes directly depend on how genetic diversity is structured within and among natural populations. For instance, levels of genetic variation and patterns of gene flow among populations determine both the opportunity for selection and the geographical scale at which selection can act (Slatkin, 1987; Sork, 2016). In plants, this mainly depends on dispersal strategies of pollen and seeds, but also on cross-compatibilities between individuals (Barrett, 2002; Charlesworth, 2006). Beyond the acknowledged usefulness of population genetics descriptors for applied purposes, such as conservation biology, describing the distribution of genetic diversity in a species also constitutes the first step of many basic research studies, in particular those that investigate the ecology and evolution of reproductive systems.

The common yellow jasmine [*Chrysojasminum fruticans* (L.) Banfi; Oleaceae] is a distylous shrub, which is the only native European species of Jasmineae, the sister lineage of the olive tribe or Oleae (Wallander and Albert, 2000; Dupin *et al.*, 2020). In various members of Oleae, a homomorphic self-incompatibility (SI) system with only two groups (di-allelic SI: DSI) has been discovered. This unusual DSI, compared to highly diversified SI (see for instance Bush *et al.*, 2010), asks the question of its maintenance as well as its potential role in the evolution of separate sexes in Oleaceae (Vassiliadis *et al.*, 2000; Saumitou-Laprade *et al.*, 2010; Vernet *et al.*, 2016; Billiard *et al.*, 2015). In addition, a breakdown of the same DSI system has been documented in some individuals of one Oleae species (de Cauwer *et al.*, 2021). No information of the genetic structure is currently available in the common yellow jasmine, and more generally in any species of Jasmineae. Such data has both practical and theoretical importance: (i) studying levels of variation will help in defining molecular markers and identifying populations of potential interest; (ii) measurements of heterozygosity should provide some indirect insight on the mating system and on possible breakdown of the DSI; and

1
2
3 (iii) patterns of gene flow through both seed and pollen may help us to infer geographical scale
4
5 at which evolutionary processes take place.
6
7

8
9 Our study species (*C. fruticans*) occurs in the wild in southwestern Europe and the
10 Mediterranean Basin (Dommée, Thompson & Cristini, 1992; Olesen *et al.*, 2003; Banfi, 2014).
11
12 It is generally found on dry and sunny soils of hillsides, often rocky with a slight preference for
13 limestone, leading to a heterogeneous and fragmented habitat. The species is a small
14
15 rhizomatous semi-evergreen shrub with slightly fragrant yellow flowers, which appear from
16
17 April to June, and small black berries (0.5 cm in diameter) in the autumn. Individuals of *C.*
18
19 *fruticans* either have “pin” flowers (long style and low anthers = longistylous) or “thrum”
20
21 flowers (short style and high anthers = brevistylous), and reciprocal crosses mostly occur
22
23 between the two morphs (Dommée, Thompson & Cristini, 1992; Gutiérrez, Gutiérrez & Medrano,
24
25 1998). Pollination and seed dispersal are respectively mediated by numerous insects of different
26
27 orders (Thompson, 2001) and by birds (Puech, 1986). A generalist pollination system that may
28
29 impede specialization of the plant to particular pollinators has been notably described by
30
31 Thompson (2001). On the other hand, no study has specifically documented fruit dispersal of
32
33 *C. fruticans*, but its frequent occurrence in abandoned cultivated areas and firebreaks suggests
34
35 an efficient spread at the local scale (e.g. Debussche, Escarré & Lepart, 1982). Although
36
37 pollinated by an insect community and its fruits spread by birds, the dispersal of the common
38
39 yellow jasmine should be limited by geographic distance due to its occurrence in a fragmented
40
41 habitat. A strong structure of its genetic diversity among populations is therefore expected, but
42
43 gene flow measured on the organellar genomes or on the nuclear genome should not be
44
45 impacted in the same way due to their different transmission modes (Petit *et al.*, 2005).
46
47
48
49
50
51
52
53
54

55 In this study, we thus aim at describing the population genetics of *Chrysojasminum*
56
57 *fruticans* in southern France, in particular the isolation by distance (IBD) and the genetic
58
59 differentiation between populations. Patterns of genetic diversity on different genomic
60

1
2
3 compartments were investigated and compared to assess the relative contribution of pollen and
4 seeds in the global gene flow. First, genotyping methods (microsatellites and indels) were
5 developed to screen polymorphisms in nuclear, chloroplast, and mitochondrial genomes, which
6 should show contrasting mutation rates (Wolfe *et al.*, 1987) and could therefore also give
7 complementary information about the impact of demographic processes operating at different
8 historical scales on the species genetic variability and structure. The inheritance of both
9 cytoplasmic organelles was beforehand tested in a few progenies in order to ensure that their
10 polymorphisms can be used to study seed dispersal. A marked genetic differentiation between
11 populations was detected, especially on nuclear data for which a strong IBD was observed
12 confirming a limitation of gene flow, especially for the pollen. The implications of our results
13 for investigating mating strategies in *C. fruticans* were finally briefly discussed.

31 32 MATERIAL AND METHODS

33 34 PLANT SAMPLING AND DNA EXTRACTION

35
36 Leaves of a total of 196 individuals ('ramets') from 13 French populations and one Spanish
37 population were collected and preserved in silica gel (Table 1). To avoid sampling clones, we
38 collected ramets that were at least 3 m apart. Floral morph and GPS coordinates of individuals
39 (or populations) were also recorded (Table 1; Supporting Information, Table S1). Although
40 style length varies in both morphs (Thompson & Dommée, 2000), the distinction between
41 brevistylous (B) and longistylous (L) individuals has never been ambiguous on open flowers.
42 In addition, 15 herbarium specimens from the Museum National d'Histoire Naturelle de Paris
43 (MNHN-P) were sampled (Supporting Information, Table S2), representing *Chrysojasminum*
44 *fruticans* populations throughout its range (from Morocco-Portugal in the West to Iran-
45 Turkmenistan in the East). Nine of those specimens were collected more than a century ago.

1
2
3 Those specimens were used as outgroups to our study and also allowed us to test the utility of
4 our markers over the whole distribution area of *C. fruticans*. All populations sampled in this
5 study are assumed to have a natural origin, with the exception of the one from the “plateforme
6 des terrains d’expérience du LabEx CeMEB” (‘CEFE’, Centre d’Étude Fonctionnelle et
7 Evolutive de Montpellier), known to have an artificial origin since it was planted ca. 35 years
8 ago by John D. Thompson and Bertrand Dommée, with individuals from different localities
9 near Montpellier.

10
11 The DNA of each sample was extracted with the BioSprint DNA Plant Kit (QIAGEN).
12 About 20 mg of material was placed in a 2-mL tube with three tungsten beads and frozen at -
13 80°C for one hour before being crushed using TissueLyser II (QIAGEN). Then, after one hour
14 of incubation at 65°C in 500 µL of 2X CTAB buffer, DNA extraction was performed following
15 the supplier's protocol. Finally, the DNA was resuspended in 130 µL of TE buffer and stored at
16 4°C.

17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

GENOPYING PROCEDURE

To genetically characterize all the sampled *C. fruticans* individuals, we developed 29 nuclear DNA loci (including 23 microsatellites and six indels) and 15 cytoplasmic loci [five mitochondrial DNA (mtDNA) loci (four indels and one microsatellite) and ten chloroplast DNA (cpDNA) loci (two indels and eight microsatellites)] (Supporting Information, Tables S3 to S5). We used shotgun genome sequencing data (*i.e.* HiSeq data on individual 'Celleneuve L11') to random search and assemble microsatellite or indel regions as described in Salmons *et al.* (2020). All assembled nuclear sequences are deposited in GenBank. In contrast, the cytoplasmic DNA loci were developed from length polymorphisms revealed on plastomes and mitogenomes of individuals collected in France (GenBank no MH559274, and OQ124187 to OQ124196).

1
2
3 The primers used to amplify the 44 loci are given in the Supporting Information (Tables S3 to
4
5 S5). All primers were defined to have the same annealing temperature (T_m).
6
7

8
9 Some markers were amplified alone (simplex), while others were amplified in multiplex
10 (either two markers for nuclear loci, or three or four for cytoplasmic loci). The PCR conditions
11 are those reported by Besnard *et al.* (2011). These amplifications were performed in a total
12 volume of 25 μ L following the method of Schuelke (2000), which allows the integration of
13 fluorescence on fragments amplified with a dye-labeled M13 primer. Three different
14 fluorochromes were used depending on the markers: 6-FAM, ATTO565 and HEX. The PCR
15 multiplexes and the fluorochrome used for each locus are given in the Supporting Information
16 (Tables S3 to S5). The same PCR program was used for all markers: denaturation of the DNA
17 at 95°C for 2 min followed by 25 cycles with a 30 sec denaturation at 95°C, followed by an
18 annealing at 55°C and then a 90 sec elongation at 72°C. Ten additional cycles with a lowered
19 annealing temperature of 51.5°C allowing integration of the dye-labeled M13 primer into
20 amplified fragments were added and were followed by a final elongation of 20 min at 72°C.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

37 After PCR amplification, the markers were multiplexed (several markers of different
38 sizes were mixed together) in formamide with ROX-500 (Applied Biosystems) used as internal
39 size standard. The amplified fragments at each locus were separated by chromatography on an
40 ABI 3730 capillary sequencer (Applied Biosystems) by GenoScreen (Lille). The
41 chromatograms were then read on GENEIOUS v.9.1.2 (Kearse *et al.*, 2012). To limit reading
42 errors, all loci were read twice independently by AP and GB.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

DEFINING CYTOPLASMIC HAPLOTYPES AND NETWORK RECONSTRUCTION

For chloroplastic and mitochondrial genomes, polymorphisms were revealed at multiple loci (see above). Plastid and mitochondrial DNA profiles (chlorotypes and mitotypes) were defined by the combination of alleles from these loci. Relationships between chlorotypes on one hand, and between mitotypes on the other hand, were established by constructing a reduced-median haplotype network with NETWORK v.10.2 (Bandelt, Forster & Röhl, 1999). All polymorphisms observed were coded as in Gryta *et al.* (2017) [*i.e.* number of repeats for microsatellites, or binary (0/1) for indels].

ON THE INHERITANCE OF ORGANELLES IN *C. FRUTICANS*

As potential for paternal leakage of chloroplasts and mitochondria has been previously reported in several *Jasminum* species (Sodmergen *et al.*, 1998; Zhang, Liu & Sodmergen, 2003), it was necessary to check the inheritance of organelles in *C. fruticans*. As a first source of evidence, we measured the linkage disequilibrium (LD) between chlorotypes and mitotypes. In the case of a strict maternal inheritance, we expect a strong LD (as already shown in beet, *Silene*, and olive; Desplanque *et al.*, 2000; Olson & McCauley, 2000; Besnard *et al.*, 2002). In contrast, the presence of a few paternal leakages may reduce LD between haplotypes of the two organelle genomes. The LD was estimated with Lewontin's normalized D' (Lewontin, 1964, 1988), and its significance was tested with Fisher's exact tests using ARLEQUIN (Excoffier & Lischer, 2010). The D' index was calculated (i) separately on each sampled population where polymorphism was observed on both organellar genomes, and (ii) on all French individuals considered as a single population. Then, we also verified the inheritance of organelles in a few progenies (*i.e.* seeds resulting from open pollination) for which parents display different cpDNA and/or mtDNA haplotypes. In the 'CEFE' and 'Moulis' populations, we previously identified 36 progenies issued from such crosses (*i.e.* 33 with a different chlorotype, and 32

1
2
3 with a different mitotype; Supporting Information, Table S6). Genotyping of the chloroplastic
4 and mitochondrial haplotypes in these progenies allowed us to determine from which parent
5 cytoplasmic organelles were inherited.
6
7
8
9

10 POPULATION GENETICS ANALYSES

11
12
13
14 To avoid bias in population genetics analyses due to repetitions of the same individual, clones
15 (genets) needed to be excluded. In a population, plants sampled a few meters apart may
16 correspond to ramets from the same stump, and the probability of finding the same genotype at
17
18
19
20
21 29 multi-allelic nuclear loci is theoretically very low (e.g. Baali-Cherif & Besnard, 2005). Thus,
22
23 when several close individuals shared the same genotype or a very similar genotype (*i.e.*
24 differing by a mutation step on a long allele), we considered them as a clone. In this case, only
25 one ramet was kept for analyses. In order to represent the genetic proximity between all
26 analyzed individuals, Bruvo *et al.* (2004) genetic distances were first calculated on the nuclear
27 microsatellite data between each pair of genotypes with the R package *Poppr* (Kamvar, Tabima
28 & Grünwald, 2014) and a dendrogram was reconstructed according to the unweighted pair
29 group method with arithmetic mean (UPGMA). For further analyses, we only considered
30 populations with at least seven genotypes.
31
32
33
34
35
36
37
38
39
40
41

42 *Populations' genetic diversity*

43
44
45 Genetic diversity patterns were described for each population. First, observed (H_O) and
46 expected (H_E) heterozygosities under the Hardy-Weinberg ideal population model (*i.e.*
47 population of infinite size, panmictic reproduction, non-overlapping generations), and the Weir
48 & Cockerham Fixation Index (F_{IS} , Weir & Cockerham, 1984) were calculated using the
49 GENETIX software (Belkhir *et al.*, 2004) for each nuclear locus. Allelic richness (A_R) was
50 calculated with the R package HIERFSTAT (Goudet, 2005) to compare the number of alleles
51 between populations. The presence of null alleles was tested for each population with a 95%-
52
53
54
55
56
57
58
59
60

1
2
3 confidence interval using MICRO-CHECKER v.2.2.3 (van Oosterhout *et al.*, 2004). As all loci
4 were successfully amplified, the null allele frequency was estimated according to Brookfield
5 (1996). For cytoplasmic data (cpDNA and mtDNA), haplotypic diversity was calculated from
6 the formula $D = 1 - \sum p_i^2$ with p_i the frequency of haplotype i and was multiplied by $n / (n - 1)$
7 to account for differences in sample sizes, as described in Gryta *et al.* (2017). This parameter
8 was estimated independently for chloroplast and mitochondrial data (D_{cp} and D_{mt}), and for their
9 combination (D_{cp+mt}).

10 11 12 13 14 15 16 17 18 19 20 *Genetic structure of populations*

21
22
23 To measure genetic differentiation between populations, the fixation index F_{ST} of Weir &
24 Cockerham (1984) was calculated between each pair of populations using GENETIX. These
25 calculations were performed separately for nuclear data (bi-paternally transmitted, and thus
26 dispersed by pollen and seed) and cytoplasmic data (assumed to be maternally transmitted, and
27 dispersed by seed). However, for cytoplasmic data, the main mode of transmission of plastids
28 and mitochondria remains uncertain in Jasmineae (Sodmergen *et al.*, 1998; Zhang, Liu &
29 Sodmergen, 2003). Occasional paternal transmission of one or both cytoplasmic genomes could
30 impact on differentiation patterns between populations. Because of this uncertainty, we
31 estimated the F_{ST} of the cytoplasmic data using four different approaches: (i) only on haplotypes
32 of the chloroplast genome, (ii) only on haplotypes of the mitochondrial genome, (iii)
33 considering the chloroplast and mitochondrial genomes as two independent cytoplasmic loci,
34 and (iv) combining cpDNA and mtDNA information to define cytoplasmic haplotypes. Patterns
35 of isolation by distance (IBD) were then compared between each approach. The geographical
36 distribution of cpDNA and mtDNA haplotypes across populations was also mapped using
37 QGIS v.3.0.3 (QGIS Development Team, 2018).
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 In order to determine the number of genetic clusters actually present in our sample and
4
5 to assess the isolation between populations, a Bayesian approach was then used to assign
6
7 individuals to genetic clusters on the basis of their genotypes. STRUCTURE v.2.3.4 (Pritchard,
8
9 Stephens & Donnelly, 2000) was used on the nuclear data set. For this analysis, the Spanish
10
11 population ('La Aliseda') was excluded due to its great distance from all the other studied
12
13 populations and the absence of intermediate populations between them. The artificial 'CEFE'
14
15 population was also excluded in order not to bias the analyses (e.g. due to admixture between
16
17 individuals). This analysis was carried out under the admixture model (mixing of gametes
18
19 possible at reproduction) for a duration of 200,000 generations after a burn-in period of 50,000
20
21 generations. The analysis was performed for a number of clusters (K) ranging from 1 to 14 with
22
23 ten iterations for each value of K . The most likely number of clusters was determined using the
24
25 *ad hoc* measure of ΔK performed with STRUCTURE HARVESTER (Earl & vonHoldt, 2012). In
26
27 order to determine the putative origin of the artificial CEFE population, we then assigned each
28
29 individual from this location to the ten natural French populations (with at least ten genets)
30
31 using GENECLASS2 (Piry *et al.*, 2004). The standard criterion described by Rannala & Mountain
32
33 (1997) was used to compute the probabilities with Bayesian statistics. The population
34
35 assignment simulation algorithm described by Paetkau *et al.* (2004) was used in order to
36
37 simulate 10,000 genotypes for each population. An arbitrary threshold probability value of 0.05
38
39 or greater was then applied to identify the possible origin of CEFE genotypes among natural
40
41 populations.
42
43
44
45
46
47
48
49

50 *Limitation of gene flow by distance and relative contribution of pollen and seeds*

51
52

53 As IBD was expected, Mantel (1967) tests between geographical distances and F_{ST} values
54
55 between pairs of populations were performed in southern France on the ten natural populations
56
57 with at least ten genotypes. These tests were done with 999 permutations using the R package
58
59 *ade4* (Chessel, Dufour & Thioulouse, 2004) for each genomic compartment in order to compare
60

1
2
3 patterns according to the origin of the data (nuclear *vs.* cytoplasmic). In addition, the relative
4 contribution of pollen and seeds in the global gene flow (R) was investigated following the
5 method of Ennos (1994). This approach allows an indirect estimation of pollen and seed flows
6 from genetic differentiation estimates (F_{ST}) on both bi-parentally inherited nuclear markers
7 (F_{STn}) and maternally inherited cytoplasmic markers (F_{STc}). The R estimation is based on a
8 model of island migration (Wright, 1965), and can be used under several assumptions: (i) a very
9 low mutation rate compared to migration rate, (ii) a constant and equal size of populations, and
10 (iii) a constant and equal migration rate between populations, the last two points representing a
11 drift/migration equilibrium. This equilibrium turns out to be rapidly reached in nature (Slatkin,
12 1993). However, when the species has a long generation time and/or if perturbations are quite
13 recent, it might be necessary to check these assumptions. In addition, the precision of estimates
14 of genetic differentiation used to infer gene flow depends on the sampling of populations (Petit
15 *et al.*, 2005; Besnard *et al.*, 2021). Although the R estimate can help comparing the relative
16 contribution of pollen and seeds in global gene flow among taxa (Ennos, 1994; Petit *et al.*,
17 2005), it thus still needs to be interpreted with caution.
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

42 RESULTS

43
44
45 All loci were successfully amplified in all *Chrysojasminum fruticans* individuals freshly
46 collected in southern France and Spain (whole dataset provided in the Supporting Information,
47 Table S1), and all revealed polymorphisms. A few genets with several ramets were identified
48 in populations 'La Aliseda', 'Amélie-les-Bains', 'Prades-le-Lez', and 'La Turbie'. These genets
49 can cover a relatively large surface (up to about 90 m² for an individual of the Spanish
50 population 'La Aliseda'). Only one ramet per genet was subsequently kept for our analyses (for
51 a total of 186 genets). Among the 179 genets sampled in France, 89 were recorded as
52
53
54
55
56
57
58
59
60

1
2
3 brevistylous and 90 as longistylous (Table 1). In addition, most of our markers yielded
4
5 amplifications with herbarium samples, pointing out their high potential to analyze genetic
6
7 variation in old specimens collected more than 100 years ago. Cytoplasmic markers revealed
8
9 the presence of seven chlorotypes and six mitotypes in French populations, most of them being
10
11 shared with the few analyzed individuals from the Southern Iberian Peninsula to the Eastern
12
13 Maghreb (e.g. all mitotypes, and chlorotypes 1, 3, 4 and 7; Supporting Information, Table S1).
14
15 By combining cpDNA and mtDNA information, 20 cytoplasmic haplotypes were identified in
16
17 southern France (Table 3), where five of the six mitotypes were associated with more than one
18
19 chlorotype, and five of the seven chlorotypes were associated with more than one mitotype.
20
21 Consequently, the LD between haplotypes of the two organellar genomes, measured according
22
23 to Lewontin (1964), was relatively low ($D' = 0.368$, p -value < 0.001) compared to other plants
24
25 where such data are available [e.g. $D' > 0.95$ in *Olea europaea* (Besnard *et al.*, 2002), *Silene*
26
27 *vulgaris* (Olson & McCauley, 2000), and *Beta vulgaris* (Desplanque *et al.*, 2000)]. However,
28
29 relatively strong and significant LD between organellar genomes was found at the population
30
31 scale (*i.e.* 'La Aliseda' and 'CEFE'; $D' > 0.88$, p -value < 0.005 ; Supporting Information, Table
32
33 S7). Overall, a non-strict association between haplotypes of both cytoplasmic genomes
34
35 indicates that organelles are not always inherited together (implying paternal leaks). Yet,
36
37 analyses of all 36 progenies only revealed a maternal inheritance of the two organellar genomes
38
39 (Supporting Information, Table S6), suggesting that paternal leaks may be occasional.
40
41 Consequently, the structure of organellar DNA diversity should be mostly influenced by seed
42
43 dispersal. In addition, all 36 crosses involved a pair of brevistylous and longistylous individuals
44
45 (24 B x L and 12 L x B; Supporting Information, Table S6).
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

STRUCTURE OF THE NUCLEAR GENETIC DIVERSITY ON THE WHOLE SAMPLE

The genetic proximity between all individuals of *C. fruticans* is represented by a dendrogram (Supporting Information, Fig. S1). All individuals from French populations were clustered together. Individuals from the same population were generally found in the same group suggesting a strong structure of populations in the South of France. Individuals from the Spanish population 'La Aliseda' formed the closest cluster to the French populations. Individuals from the rest of the Mediterranean were more distant from French and Spanish populations, except for individuals from Tunisia and Portugal. This analysis also showed that individuals from Iran and Turkmenistan are very similar to each other and genetically the most distant from all other individuals analyzed.

GENETIC DIVERSITY OF POPULATIONS ANALYSES

Most nuclear loci revealed polymorphism within populations, even the di-allelic indel markers (Supporting Information, Table S8). Nevertheless, a few loci were not variable in some populations, especially in 'Moulis' (eight loci) and 'La Aliseda' (six loci). Among populations, most loci showed no excess of homozygosity or significant deviation from expected heterozygosities at the Hardy-Weinberg equilibrium, suggesting that the species reproduces mainly via outcrossing. For all populations, the global F_{IS} value was thus not significantly different from 0. The MICRO-CHECKER analysis detected probable presence of null allele in only five populations for one locus each (Cf-TA-1 for 'Bages', Cf-ATT2 for 'CEFE', Cf-GAA19 for 'Moulis', Cf-GAA22 for 'St-Maurice-Navacelles', and Cf-GAA4 for 'Puycelis', with respectively the followings Brookfield estimates of null allele frequencies: 0.019, 0.215, 0.093, 0.079, and 0.154). Considering the multiple tests performed (*i.e.* 29 loci x 12 populations), such observations may be fortuitous and all loci should have a very low frequency of null alleles. Finally, the average expected heterozygosity and allelic richness (calculated for seven diploid

1
2
3 individuals) showed that 'Moulis' ($H_E = 0.35$; $A_R = 2.36$) is genetically impoverished compared
4
5 to all other populations ($0.41 < H_E < 0.56$; $2.84 < A_R < 3.89$; Table 2). The highest level of
6
7 nuclear diversity was revealed in 'St-Maurice-Navacelles'.
8
9

10
11 Most populations showed polymorphisms on both chloroplast and mitochondrial
12
13 genomes (Fig. 2; Supporting Information, Table S8), with a maximum of five mitotypes in 'St-
14
15 Maurice-Navacelles' and four chlorotypes in 'CEFE'. The combination of both genomes
16
17 allowed us to detect from two to eight cytoplasmic haplotypes per population, with a haplotypic
18
19 diversity (D_{cp+mt}) varying from 0.19 ('Celleneuve') to 0.91 ('St-Maurice-Navacelles').
20
21 Conversely to the nuclear genome, the 'Moulis' population did not show a particularly reduced
22
23 diversity of cytoplasmic haplotypes ($D_{cp+mt} = 0.50$), although its chlorotype diversity was
24
25 among the lowest ($D_{cp} = 0.11$; Table 2).
26
27
28
29

30 POPULATION STRUCTURE

31
32
33 The F_{ST} values calculated from the nuclear data supported a significant genetic differentiation
34
35 between most pairs of populations (Supporting Information, Table S9). The dendrogram
36
37 reconstructed from these F_{ST} (Supporting Information, Fig. S2) showed that geographically
38
39 close populations were also genetically close, except in the case of 'Moulis', which was found
40
41 here genetically more differentiated from the other French populations than the Spanish
42
43 population was. Also, the artificial population of 'CEFE' (Montpellier) whose individuals
44
45 originate from neighboring populations was found to be close to the other populations sampled
46
47 near Montpellier ('Celleneuve' and 'Prades-le-Lez').
48
49
50

51
52 An analysis performed with STRUCTURE on nuclear data from ten natural French
53
54 populations (with at least ten genotypes) revealed the presence of several genetic clusters. Peaks
55
56 in the graphical representation of ΔK (Fig. 1A) allowed the identification of population
57
58 structuring at two levels: $K = 2$ and 10. For $K = 2$, 'Moulis' was distinguished from other
59
60

1
2
3 populations (Fig. 1B), whereas the ten populations defined *a priori* were relatively well
4 supported at $K = 10$ (Fig. 1C). The assignment of ‘CEFE’ individuals to these natural
5
6 supported at $K = 10$ (Fig. 1C). The assignment of ‘CEFE’ individuals to these natural
7
8 populations confirmed the putative origin of this artificial population from collects around
9
10 Montpellier, since successful assignments (27% of genets) only concerned the three locations
11
12 collected in Hérault ('St-Maurice-de-Navacelles', 'Celleneuve', and 'Prades-le-Lez'; Supporting
13
14 Information, Table S10).
15

16
17
18 Genetic relationships between cytoplasmic haplotypes assessed by reduced-median
19
20 network analyses (Fig. 2) showed three slightly differentiated lineages of chlorotypes (*i.e.* 1-2,
21
22 3-6, and 7) and no clear genetic structure for mitotypes. The geographical distribution of these
23
24 chlorotypes and mitotypes (Fig. 2) did not show any clear pattern of geographical structuring.
25
26 However, F_{ST} values calculated from these cytoplasmic data were relatively high and most of
27
28 them were significantly different from 0 (Supporting Information, Table S8).
29
30

31 32 ISOLATION BY DISTANCE AND CONTRIBUTION OF POLLEN AND SEEDS TO GLOBAL GENE FLOW 33

34
35 Strong IBD was observed for the nuclear data (Mantel test, $r = 0.528$, p -value = 0.004; Fig. 3A).
36
37 In contrast, the cytoplasmic data did not show any isolation pattern by distance (Fig. 3B),
38
39 regardless of how the F_{ST} values were calculated (*i.e.* based on cpDNA, mtDNA, or when
40
41 combining information of both organellar genomes; Supporting Information, Fig. S3). Based
42
43 on the estimates of global genetic differentiation between populations for nuclear markers
44
45 ($F_{STn} = 0.17$, $F_{IS} = 0.17$) and cytoplasmic genomes ($F_{STc} = 0.50$), the Ennos's index ($R = 3.618$)
46
47 showed that the contribution of the pollen to global gene flow is somewhat greater than that of
48
49 seeds. However, this index was relatively low compared to many other plants (median $R = ca.$
50
51 17; Petit *et al.*, 2005).
52
53
54
55
56
57
58
59
60

DISCUSSION

This work represents the first population genetics study of *Chrysojasminum fruticans*. We mainly sampled populations in southern France and recorded floral morphs on all individuals. Genetic variation was then detected in the three genome compartments allowing revealing contrasting levels of diversity among populations as well as a strong genetic structure resulting from limited gene flow. It also brought information on the relative contribution of seeds and pollen in the global gene flow. Our results will have various methodological implications when developing genetic studies on this species, especially to study its reproductive and dispersal strategies.

First, our sampling in southern France revealed that both floral morphs have a balanced frequency among all populations (89 B vs. 90 L), and the few analyzed progenies (36) also confirmed that crosses should mainly occur between these two morphs (Dommée, Thompson & Cristini, 1992; Gutiérrez, Gutiérrez & Medrano, 1998). These preliminary observations are consistent with a diallelic self-incompatibility system but a wider sampling is still necessary, especially to measure how mating patterns vary among populations and floral morphs (for instance, to test if cross-compatibility or selfing rate vary among morphs in relation with stigma height or any other traits; Thompson & Dommée, 1993, 2000).

Then, a set of markers was developed for investigating variation of the three genomic compartments in *C. fruticans*. Such tools may have applications in different fields from population genetics and phylogeography to parentage analyses. Here, chloroplast and mitochondrial markers showed great potential for investigating the inheritance of both organellar genomes by either investigating the linkage disequilibrium between chloroplast and mitochondrial haplotypes, or directly by checking their inheritance in progenies. The low linkage disequilibrium on the whole sample supports a non-strict maternal inheritance in

1
2
3 *C. fruticans* as already documented in other Jasmineae (Sodmergen *et al.*, 1998; Zhang, Liu &
4 Sodmergen, 2003), but our preliminary progeny analyses indicate that paternal leakage should
5 be rare. Consequently, we used organellar DNA polymorphism for investigating seed dispersal.
6
7 Furthermore, nuclear microsatellite markers may be useful in assisting the identification of
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Furthermore, nuclear microsatellite markers may be useful in assisting the identification of
genets in populations and in determining fine population structure to further study demographic
processes at a local scale. Here, they revealed shared genotypes by close ramets indicating the
presence of clones in the investigated populations due to vegetative growth. The shoots of a
single stump can cover up to several tens of m² (as already observed in olive trees; Baali-Cherif
& Besnard, 2005), indicating *a priori* a long persistence of certain individuals. Based on our
observations, avoiding multiple sampling of the same genotype in population genetic studies
may therefore require collecting ramets at least ten meters apart.

Once the clones were removed, our population analysis showed a strong structure of the
genetic diversity in *C. fruticans* (*i.e.* high differentiation between pairs of populations) on both
nuclear and cytoplasmic data. In addition, strong IBD is observed on the nuclear data, but not
on the cytoplasmic data (Fig. 3). The presence of a very strong IBD is common in insect-
pollinated species with limited pollen and seed dispersal (e.g. Jump *et al.*, 2009). In *C. fruticans*,
the pollen is sticky and should be mostly spread by various insects (Thompson, 2001) at the
local scale. In contrast, fleshy fruits can be spread by birds (Puech, 1986), and the longest
dispersal events may thus occur at a larger distance for seeds than for pollen, especially in a
very fragmented landscape (e.g. García-Verdugo *et al.*, 2010; Trapnell *et al.*, 2022). The relative
contribution of pollen and seeds to global gene flow reflects this expectation, with an *R* value
inferior to 5 as frequently observed in most insect-pollinated and fleshy fruited species (Oddou-
Muratorio *et al.*, 2001; Petit *et al.*, 2005). Furthermore, random long-distance dispersal by seeds
through the landscape would lead to a patchy structure of organellar lineages in space (Petit *et al.*, 1997), and could explain the absence of IBD on both maternally inherited genomes. In

1
2
3 addition, the plants mainly produced fruits in autumn (Puech, 1986) when migratory birds are
4 travelling southwards, and a latitudinal pattern could thus be expected. It is noticeable that
5 organellar haplotypes observed in southern France are shared with populations from the
6 Southern Iberian Peninsula to the Eastern Maghreb, suggesting that seeds can indeed travel long
7 distances with birds and could have even crossed the Mediterranean Sea, as already shown in
8 another Mediterranean shrub, *Pistacia lentiscus* L. (Martinez-Lopez *et al.*, 2020). Nonetheless,
9 it is possible that cytoplasmic haplotypes identified here have diverged a long time ago, and
10 more variable loci would reveal some differentiation between such distant populations. It is
11 thus presently too speculative to interpret our diversity patterns in the context of Mediterranean
12 micro-refugia (*i.e.* south-eastern Pyrenees, southern Cevennes, Mont Ventoux, eastern
13 Provence, Alpes Maritimes; Médail & Diadema, 2009), and a denser and wider sampling of the
14 studied area would also be required for this purpose.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

31 Finally, in a context of gene flow limitation, small marginal populations are expected to
32 diverge rapidly. In our study, the 'Moulis' population is highly differentiated from the others
33 and is genetically impoverished compared to other populations, supporting the fact that it is
34 highly isolated. This population is relatively large (several hundred thousand individuals over
35 several hundred ha), and its low diversity may be explained by a relatively recent expansion
36 from a few founders. However, the necessity that recent populations were initially founded from
37 at least two compatible individuals could be a constraint, and self-compatible individuals may
38 be temporarily favored on the front of colonization to ensure reproduction (Baker, 1955; Pannell
39 & Barrett, 1998). Although no significant inbreeding index was observed in all populations
40 indicating that the species is mainly out-crossing, preliminary paternity analyses on seeds
41 collected in 'Moulis' from two mother plants in two following years showed a high rate of self-
42 fertilization (ca. 70%) on a brevistylous individual (unpublished data), suggesting that this
43 population is indeed not completely allogamous as similarly reported in some populations of
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 *Ligustrum vulgare* L. (de Cauwer *et al.*, 2021). Thus, effective selfing events could have also
4
5 contributed to genetic diversity erosion of the 'Moullis' population.
6
7
8
9

10 11 12 CONCLUSION AND PERSPECTIVES 13 14

15 Our results have several implications for further studies on the variation of mating strategies in
16
17 *Chrysojasminum fruticans*. Variable levels of genetic diversity among unconnected populations
18
19 imply contrasting demographic histories with possibly varying loads of deleterious mutations
20
21 (e.g. Liu *et al.*, 2022). In a distylous species, if the selfing rate is higher in one morph than in
22
23 the other, questions arise about the allocation of resources and their evolutionary consequences
24
25 on the reproductive strategies (e.g. possible transition to cryptic dioecy). For this reason, it
26
27 seems relevant to study mating of the common yellow jasmine in various contexts. One could
28
29 first compare the rates of self-fertilization (by paternity analysis on seeds) on the two floral
30
31 morphs in several populations, and in case of marked differences, study the associated
32
33 inbreeding depression. If one morph is much more self-pollinating in some populations, it
34
35 would also be very relevant to follow the investment of each morph in reproduction and
36
37 vegetative growth (Dommée, Thompson & Cristini, 1992).
38
39
40
41
42

43 Finally, the artificial population on the experimental field of the 'CEFE' Montpellier
44
45 (with approximately 200-300 mature individuals in a hedge of ca. 250 m²) may also offer some
46
47 opportunities for studying mating in a semi-natural context. Our genetic data supported that
48
49 individuals have mostly originated from different localities near Montpellier. Its high
50
51 accessibility may allow long-term follow up of individuals (e.g. measure of their investment in
52
53 sexual functions over years). Due to the limitation of gene flow by distance, this population can
54
55 be considered as strongly isolated from natural stands (at several kilometers). It will be fully
56
57 characterized with our markers in order to investigate pollen and seed dispersal, as well as
58
59
60

1
2
3 mating between morphs (Besnard *et al.*, 2020). Microsatellite analyses should allow
4
5 investigating pollen gene flow via paternity analyses of seeds, while recruitment over last
6
7 decades could be investigated via reconstructing pedigrees (e.g. Bacles, Lowe & Ennos, 2006;
8
9 Bittencourt & Sebben, 2007). Finally, cytoplasmic variation in the population could be useful
10
11 to investigate the inheritance of organelles and estimate the rate of paternal leaks in known
12
13 crosses.
14
15
16
17
18
19

20 ACKNOWLEDGEMENTS

21
22
23 We thank Myriam Gaudeul (MNHN-P), Alex Baumel, Elodie Chapuis, Nicolas Juillet, Gérard
24
25 Largier (CBN-PMP), Elwen Le Bras, Frédéric Médail, James Molina (CBN-Med), and John
26
27 Thompson, for their help in plant sampling. The 'CEFE' population is managed by the Platform
28
29 “Terrains d'Expériences” of the LabEx CeMEB (ANR-10-LABX-04-01). We also thank two
30
31 anonymous referees for their constructive comments.
32
33
34
35
36
37

38 FUNDING

39
40 This research was funded by the Region Occitanie of France, through the BIODIVOC
41
42 programme (grant REPOL). J.D. and G.B. received support from the H2020 Marie Skłodowska-
43
44 Curie Actions (grant 842234-FRUITFUL), and A-L.F., P-O.C. and G.B. from the European
45
46 Union's Horizon 2020 project GEN4OLIVE (H2020-SFS-2020-1; G.A. No. 101000427). P.B.,
47
48 L.C., P.J., H.G. and G.B. are members of the EDB laboratory, which is supported by the
49
50 Laboratoire d'Excellence (LABEX) CEBA (grant ANR-10-LABX-25-01) and LABEX TULIP
51
52 (grant ANR-10-LABX-0041), both managed by the Agence Nationale de la Recherche in
53
54 France.
55
56
57
58
59
60

DATA AVAILABILITY

Genotype data generated in this study are provided in the Supporting Information (Table S1). Thirteen herbarium specimens were collected in Amélie-les-Bains, Bages, CEFE, Moulis, Moussoulens, and Puycelsi, and deposited at the Montpellier herbarium (MPU; HG-2023-001 to -005, and GB-2023-001 to -008).

REFERENCES

- Baali-Cherif D, Besnard G. 2005.** High genetic diversity and clonal growth in relict populations of *Olea europaea* subsp. *laperrinei* (Oleaceae) from Hoggar, Algeria. *Annals of Botany* **96**: 823–830.
- Bacles CFE, Lowe AJ, Ennos RA. 2006.** Effective seed dispersal across a fragmented landscape. *Science* **311**: 628–628.
- Baker HG. 1955.** Self-compatibility and establishment after "long-distance" dispersal. *Evolution* **9**: 347–348.
- Bandelt HJ, Forster P, Röhl A. 1999.** Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**: 37–48.
- Banfi E. 2014.** *Chrysojasminum*, a new genus for *Jasminum* sect. *Alternifolia* (Oleaceae, Jasmineae). *Natural History Sciences* **1**: 3–6.
- Barrett SCH. 2002.** The evolution of plant sexual diversity. *Nature Reviews Genetics* **3**: 274–284.
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. 2004.** GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).
- Besnard G, Cheptou PO, Debbaoui M, Lafont P, Huguény B, Dupin J, Baali-Cherif D. 2020.** Paternity tests support a diallelic self-incompatibility system in a wild olive (*Olea europaea* subsp. *laperrinei*, Oleaceae). *Ecology and Evolution* **10**: 1876–1888.
- Besnard G, Gorrilliot O, Raimondeau P, Génot B, El Bakkali A, Anthelme F, Baali-Cherif D. 2021.** Contrasting genetic footprints among Saharan olive populations: potential causes and conservation implications. *Plants* **10**: 1207.

- 1
2
3 **Besnard G, Hernández P, Khadari B, Dorado G, Savolainen V. 2011.** Genomic profiling of
4 plastid DNA variation in the Mediterranean olive tree. *BMC Plant Biology* **11**: 80.
- 5
6 **Besnard G, Khadari B, Baradat P, Bervillé A. 2002.** Combination of chloroplast and
7 mitochondrial DNA polymorphisms to study cytoplasm genetic differentiation in the
8 olive complex (*Olea europaea* L.). *Theoretical and Applied Genetics* **105**: 139–144.
- 9
10
11 **Billiard S, Husse L, Lepercq P, Godé C, Bourceaux A, Lepart J, Vernet P, Saumitou-**
12 **Laprade P. 2015.** Selfish male-determining element favors the transition from
13 hermaphroditism to androdioecy. *Evolution* **69**: 683–693.
- 14
15
16 **Bittencourt JVM, Sebben AM. 2007.** Patterns of pollen and seed dispersal in a small,
17 fragmented population of the wind pollinated tree *Araucaria angustifolia* in southern
18 Brazil. *Heredity* **99**: 580–591.
- 19
20
21 **Brookfield JFY. 1996.** A simple new method for estimating null allele frequency from
22 heterozygote deficiency. *Molecular Ecology* **5**: 453–455.
- 23
24
25 **Bruvo R, Michiels NK, D’Souza TG, Schulenburg H. 2004.** A simple method for the
26 calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular*
27 *Ecology* **13**: 2101–2106.
- 28
29
30 **Busch JW, Joly S, Schoen DJ. 2010.** Does mate limitation in self-incompatible species
31 promote the evolution of selfing? The case of *Leavenworthia alabamica*. *Evolution* **64**:
32 1657–1670.
- 33
34
35 **Charlesworth D. 2006.** Evolution of plant breeding systems. *Current Biology* **16**: R726–R735.
- 36
37 **Chessel D, Dufour A, Thioulouse J. 2004.** The ade4 package - I: One-table methods. *R News*
38 **4**: 5–10.
- 39
40
41 **de Cauwer I, Vernet P, Billiard S, Godé C, Bourceaux A, Ponitzki C, Saumitou-Laprade**
42 **P. 2021.** Widespread coexistence of self-compatible and self-incompatible phenotypes in
43 a diallelic self-incompatibility system in *Ligustrum vulgare* (Oleaceae). *Heredity* **127**:
44 384–392.
- 45
46
47 **Debussche M, Escarré J, Lepart J. 1982.** Ornithochory and plant succession in Mediterranean
48 abandoned orchards. *Vegetatio* **48**: 255–266.
- 49
50
51 **Desplanque B, Viard F, Bernard J, Forcioli D, Saumitou-Laprade P, Cuguen J, Van Dijk**
52 **H. 2000.** The linkage disequilibrium between chloroplast DNA and mitochondrial DNA
53 haplotypes in *Beta vulgaris* ssp. *maritima* (L.): the usefulness of both genomes for
54 population genetic studies. *Molecular Ecology* **9**: 141–154.
- 55
56
57
58
59
60

- 1
2
3 **Dommée B, Thompson JD, Cristini F. 1992.** Distylie chez *Jasminum fruticans* L. : Hypothèse
4 de la pollinisation optimale basée sur les variations de l'écologie intraflorale. *Bulletin de*
5 *la Société Botanique de France* **139**: 223–234.
6
7
8 **Dupin J, Raimondeau P, Hong-Wa C, Manzi S, Gaudeul M, Besnard G. 2020.** Resolving
9 the phylogeny of the olive family (Oleaceae): confronting information from organellar
10 and nuclear genomes. *Genes* **11**: 1508.
11
12 **Earl DA, vonHoldt BM. 2012.** STRUCTURE HARVESTER: A website and program for visualizing
13 STRUCTURE output and implementing the Evanno method. *Conservation Genetics*
14 *Resources* **4**: 359–361.
15
16 **Ennos RA. 1994.** Estimating the relative rates of pollen and seed migration among plant-
17 populations. *Heredity* **72**: 250–259.
18
19 **Excoffier L, Lischer HEL. 2010.** ARLEQUIN suite ver 3.5: A new series of programs to perform
20 population genetics analyses under Linux and Windows. *Molecular Ecology Resources*
21 **10**: 564–567.
22
23 **García-Verdugo C, Forrest AD, Fay MF, Vargas P. 2010.** The relevance of gene flow in
24 metapopulation dynamics of an oceanic island endemic, *Olea europaea* subsp.
25 *guanchica*. *Evolution* **64**: 3525–3536.
26
27 **Goudet J. 2005.** HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics.
28 *Molecular Ecology Notes* **5**: 184–186.
29
30 **Gryta H, Van de Paer C, Manzi S, Holota H, Roy M, Besnard G. 2017.** Genome skimming
31 and plastid microsatellite profiling of alder trees (*Alnus* spp., Betulaceae): Phylogenetic
32 and phylogeographical prospects. *Tree Genetics & Genomes* **13**: 118.
33
34 **Gutián J, Gutián P, Medrano M. 1998.** Floral biology of the distylous Mediterranean shrub
35 *Jasminum fruticans* (Oleaceae). *Nordic Journal of Botany* **18**: 195–201.
36
37 **Jump AS, Rico L, Lloret F, Peñuelas J. 2009.** Microspatial population genetic structure of
38 the Mediterranean shrub *Fumana thymifolia*. *Plant Biology* **11**: 152–160.
39
40 **Kamvar ZN, Tabima JF, Grünwald NJ. 2014.** Poppr: An R package for genetic analysis of
41 populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**: e281.
42
43 **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper**
44 **A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012.**
45 GENEIOUS Basic: An integrated and extendable desktop software platform for the
46 organization and analysis of sequence data. *Bioinformatics* **28**: 1647–1649.
47
48 **Lewontin RC. 1964.** The interaction of selection and linkage. I. General considerations;
49 heterotic models. *Genetics* **49**: 49–67.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 **Lewontin RC. 1988.** On measures of gametic disequilibrium. *Genetics* **120**: 849–852.
- 4
5 **Liu S, Zhang L, Sang Y, Lai Q, Zhang X, Jia C, Long Z, Wu J, Ma T, Mao K, Street NR,**
6 **Ingvarsson PK, Liu J, Wang J. 2022.** Demographic history and natural selection shape
7 patterns of deleterious mutation load and barriers to introgression across *Populus* genome.
8 *Molecular Biology and Evolution* **39**: msac008.
- 9
10
11 **Mantel N. 1967.** The detection of disease clustering and a generalized regression approach.
12 *Cancer Research* **27**: 209–220.
- 13
14
15 **Martínez-López V, García C, Zapata V, Robledano F, De la Rúa P. 2020.** Intercontinental
16 long-distance seed dispersal across the Mediterranean Basin explains population genetic
17 structure of a bird-dispersed shrub. *Molecular Ecology* **29**: 1408–1420.
- 18
19
20 **Médail F, Diadema K. 2009.** Glacial refugia influence plant diversity patterns in the
21 Mediterranean Basin. *Journal of Biogeography* **36**: 1333–1345.
- 22
23
24 **Oddou-Muratorio S, Petit RJ, Le Guerroué B, Guesnet D, Demesure B. 2001.** Pollen-
25 versus seed-mediated gene flow in a scattered forest tree species. *Evolution* **55**: 1123–
26 1135.
- 27
28
29 **Olesen JM, Dupont YL, Ehlers BK, Valido A, Hansen DM. 2003.** Heterostyly in the
30 Canarian endemic *Jasminum odoratissimum* (Oleaceae). *Nordic Journal of Botany* **23**:
31 537–539.
- 32
33
34 **Olson MS, McCauley DE. 2000.** Linkage disequilibrium and phylogenetic congruence
35 between chloroplast and mitochondrial haplotypes in *Silene vulgaris*. *Proceedings of the*
36 *Royal Society of London, Series B* **267**: 1801–1808.
- 37
38
39 **Pannell JR, Barrett SCH. 1998.** Baker's law revisited: reproductive assurance in a
40 metapopulation. *Evolution* **52**: 657–668.
- 41
42
43 **Paetkau D, Slade R, Burden M, Estoup A. 2004.** Genetic assignment methods for the direct,
44 real-time estimation of migration rate: a simulation-based exploration of accuracy and
45 power. *Molecular Ecology* **13**: 55–65.
- 46
47
48 **Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG. 2005.** Comparative
49 organization of chloroplast, mitochondrial and nuclear diversity in plant populations.
50 *Molecular Ecology* **14**: 689–701.
- 51
52
53 **Petit RJ, Pineau E, Demesure B, Bacilieri R, Ducouso A, Kremer A. 1997.** Chloroplast
54 DNA footprints of postglacial recolonization by oaks. *Proceedings of the National*
55 *Academy of Sciences of the United States of America* **94**: 9996–10001.
- 56
57
58
59
60

- 1
2
3 **Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A. 2004.** GENECLASS2:
4 A software for genetic assignment and first-generation migrant detection. *Journal of*
5 *Heredity* **95**: 536–539.
6
7
8 **Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure using
9 multilocus genotype data. *Genetics* **155**: 945–959.
10
11 **Puech S. 1986.** Production des diaspores et potentialités de germination chez quelques espèces
12 à fruits charnus, ornithochores, dans le sud-est de la France. *Ecologia Mediterranea* **12**:
13 143–158.
14
15
16 **QGIS Development Team. 2018.** QGIS Geographic Information System. Open Source
17 Geospatial Foundation Project. <https://qgis.org/>
18
19
20 **Rannala B, Mountain JL. 1997.** Detecting immigration by using multilocus
21 genotypes. *Proceedings of the National Academy of Sciences of the United States of*
22 *America* **94**: 9197–9201.
23
24
25 **Salmona J, Olofsson JK, Hong-Wa C, Razanatosoa J, Rakotonasolo F, Ralimanana H,**
26 **Randriamboavonjy T, Suescun U, Vorontsova MS, Besnard G. 2020.** Late Miocene
27 origin and recent population collapse of the savanna Malagasy olive (*Noronhia lowryi*).
28 *Biological Journal of the Linnean Society* **129**: 227–243.
29
30
31 **Saumitou-Laprade P, Vernet P, Vassiliadis C, Hoareau Y, de Magny G, Dommée B,**
32 **Lepart J. 2010.** A self-incompatibility system explains high male frequencies in an
33 androdioecious plant. *Science* **327**: 1648–1650.
34
35
36 **Schuelke M. 2000.** An economic method for the fluorescent labeling of PCR fragments. *Nature*
37 *Biotechnology* **18**: 233–234.
38
39
40 **Slatkin M. 1987.** Gene flow and the geographic structure of natural populations. *Science* **236**:
41 787–792.
42
43
44 **Slatkin M. 1993.** Isolation by distance in equilibrium and non-equilibrium populations.
45 *Evolution* **47**: 264–279.
46
47
48 **Sodmergen, Bai HH, He JX, Kuroiwa H, Kawano S, Kuroiwa T. 1998.** Potential for
49 biparental cytoplasmic inheritance in *Jasminum officinale* and *Jasminum nudiflorum*.
50 *Sexual Plant Reproduction* **11**: 107–112.
51
52
53 **Sork VL. 2016.** Gene flow and natural selection shape spatial patterns of genes in tree
54 populations: implications for evolutionary processes and applications. *Evolutionary*
55 *Applications* **9**: 291–310.
56
57
58 **Thompson JD, Dommée B. 1993.** Sequential variation in the components of reproductive
59 success in the distylous *Jasminum fruticans* (Oleaceae). *Oecologia* **94**: 480–487.
60

- 1
2
3 **Thompson JD, Dommée B. 2000.** Morph-specific patterns of variation in stigma height in
4 natural populations of distylous *Jasminum fruticans*. *New Phytologist* **148**: 303–314.
- 5
6 **Thompson JD. 2001.** How do visitation patterns vary among pollinators in relation to floral
7 display and floral design in a generalist pollination system? *Oecologia* **126**: 386–394.
- 8
9
10 **Trapnell DW, Smallwood PA, Dixon KW, Philipps RD. 2022.** Are small populations larger
11 than they seem? Genetic insights into patchily distributed populations of *Drakaea*
12 *glyptodon* (Orchidaceae). *Botanical Journal of the Linnean Society* **198**: 99–116.
- 13
14
15 **Van Oosterhout C, Hutchinson W, Willis D, Shipley P. 2004.** MICRO-CHECKER: software for
16 identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology*
17 *Notes* **4**: 535–538.
- 18
19
20 **Vassiliadis C, Valero M, Saumitou-Laprade P, Godelle B. 2000.** A model for the evolution
21 of high frequencies of males in an androdioecious plant based on a cross-compatibility
22 advantage of males. *Heredity* **85**: 413–422.
- 23
24
25 **Vernet P, Lepercq P, Billiard S, Bourceaux A, Lepart J, Dommée B, Saumitou-Laprade**
26 **P. 2016.** Evidence for the long-term maintenance of a rare self-incompatibility system in
27 Oleaceae. *New Phytologist* **210**: 1408–1417.
- 28
29
30 **Wallander E, Albert VA. 2000.** Phylogeny and classification of Oleaceae based on *rps16* and
31 *trnL-F* sequence data. *American Journal of Botany* **87**: 1827–1841.
- 32
33
34 **Weir BS, Cockerham CC. 1984.** Estimating *F*-statistics for the analysis of population
35 structure. *Evolution* **38**: 1358–1370.
- 36
37
38 **Wolfe KH, Li WH, Sharp PM. 1987.** Rates of nucleotide substitution vary greatly among
39 plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National*
40 *Academy of Sciences of the United States of America* **84**: 9054–9058.
- 41
42
43 **Wright S. 1965.** The interpretation of population structure by *F*-statistics with special regard
44 to systems of mating. *Evolution* **19**: 395–420.
- 45
46
47 **Zhang Q, Liu Y, Sodmergen. 2003.** Examination of the cytoplasmic DNA in male
48 reproductive cells to determine the potential for cytoplasmic inheritance in 295
49 Angiosperm species. *Plant and Cell Physiology* **44**: 941–951.
- 50
51
52
53
54
55
56
57
58
59
60

Table 1. List of sampled populations, with their location and geographic coordinates. All populations except one ('La Aliseda', Spain) are located in France. Nr and Ng refer to the number of collected 'ramets' and of recorded genotypes (after excluding clones), respectively. For each French locality, the number of brevistylous (B) and longistylous (L) genotypes is given in bracket (for more details see Supporting Information, Table S1). Lat. = Latitude; Long. = Longitude.

Locality	Department (or province)	Lat., Long.	Nr	Ng (B/L)
La Aliseda	Jaén	38.335, -3.569	10	7
Amélie-les-Bains	Pyrénées Orientales	42.478, 2.666	9	5 (3/2)
Bages	Aude	43.132, 2.985	12	12 (8/4)
Bidon	Ardèche	44.370, 4.532	10	10 (5/5)
CEFE (Montpellier)	Hérault	43.638, 3.862	15	15 (6/9)
Celleneuve	Hérault	43.611, 3.821	21	21 (10/11)
Eguilles	Bouches du Rhône	43.601, 5.2669	10	10 (5/5)
Moulis	Ariège	42.979, 1.084	37	37 (18/19)
Moussoulens	Aude	43.286, 2.249	15	15 (8/7)
Prades-le-Lez	Hérault	43.693, 3.861	11	10 (5/5)
Puycelsi	Tarn	44.015, 1.676	14	14 (8/6)
Sainte-Baume	Bouches du Rhône	43.327, 5.657	10	10 (4/6)
St-Maurice-Navacelles	Hérault	43.891, 3.510	17	17 (8/9)
La Turbie	Alpes Maritimes	43.747, 7.411	5	3 (1/2)

Table 2. Summary of genetic diversity parameters estimated within populations with at least seven individuals. Expected and observed heterozygosities were calculated using GENETIX (Belkhir *et al.*, 2004), the allelic richness (A_R) with the R package HIERFSTAT (Goudet, 2005), and the haplotypic diversities (D) according to Gryta *et al.* (2017). Parameters estimated per locus are given in the Supporting Information (Table S8). A_R = Allelic richness, H_E = Expected heterozygosity under Hardy-Weinberg, H_O = Observed heterozygosity, D_{cp} = Chloroplast haplotype diversity, D_{mt} = Mitochondrial haplotype diversity, D_{cp+mt} = Cytoplasmic haplotype diversity.

Population	Nuclear DNA			Cytoplasmic DNA		
	A_R	H_E	H_O	D_{cp}	D_{mt}	D_{cp+mt}
La Aliseda	3.17	0.42	0.45	0.76	0.76	0.76
Bages	3.38	0.50	0.49	0.53	0.44	0.65
Celleneuve	3.32	0.47	0.47	0.10	0.10	0.19
CEFE	3.78	0.55	0.57	0.70	0.53	0.70
Prades-le-Lez	3.21	0.47	0.51	0.69	0.00	0.69
St-Maurice-de-Navacelles	3.89	0.56	0.58	0.63	0.76	0.91
Moussoulens	3.32	0.50	0.54	0.75	0.75	0.83
Puycelsi	3.36	0.50	0.46	0.44	0.54	0.74
Bidon	3.29	0.49	0.58	0.36	0.00	0.36
Sainte-Baume	2.84	0.41	0.46	0.20	0.20	0.20
Eguilles	3.19	0.47	0.52	0.38	0.20	0.38
Moulis	2.36	0.35	0.37	0.11	0.47	0.50

Table 3. Association between mitotypes and chlorotypes on all *C. fruticans* individuals collected in southern France (179 individuals). Occurrence number of each chlorotype-mitotype combination is reported. Based on these data, the LD between haplotypes of the two organellar genomes, measured according to Lewontin (1964), was relatively low, but still highly significant ($D' = 0.368$, p -value < 0.001).

Mitotype no	Chlorotype no							Total
	1	2	3	4	5	6	7	
1	36		26	3	1	8		74
2	6		1	3				10
3	2		4	3			1	10
4	37	12	10	15			2	76
5				1				1
6				5		3		8
Total	81	12	41	30	1	11	3	179

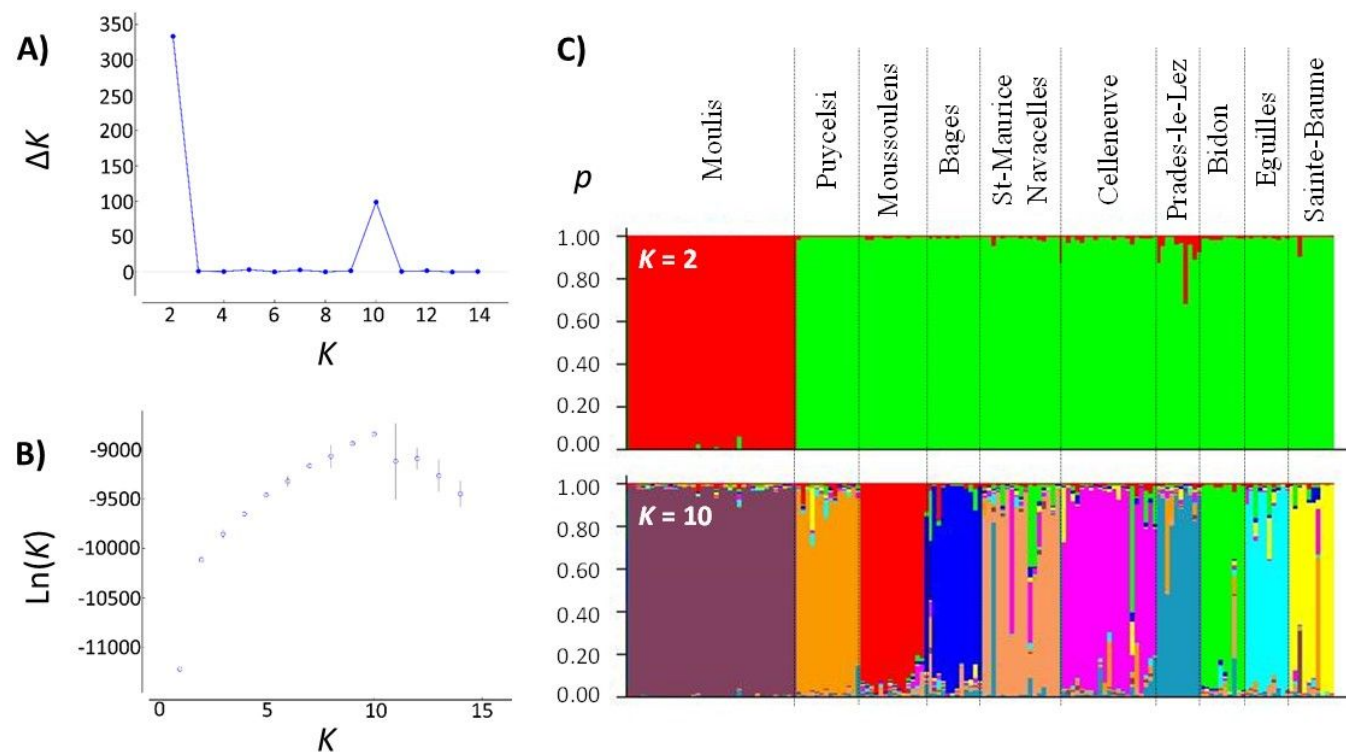


Figure 1. Results of the STRUCTURE analysis on ten French populations. A Bayesian approach was used to assign individuals to populations on the basis of their genotypes with STRUCTURE v.2.3.4 (Pritchard, Stephens & Donnelly, 2000). The analysis was performed for a number of clusters (K) ranging from 1 to 14 with ten iterations for each value of K . The most likely numbers of clusters were determined using the *ad hoc* measure of ΔK performed with STRUCTURE HARVESTER (Earl & vonHoldt, 2012). **(A)** Absolute values of the second-order rate of change of the likelihood distribution divided by the s.d. of the likelihoods (ΔK); **(B)** Mean log likelihood [$\ln(K)$ +/- s.d.] averaged over the ten iterations for each K value; **(C)** Barplots of the population structure for $K = 2$ and 10. Each genetic cluster is represented by a color. Each vertical bar corresponds to an individual with their percentage assignment (p) to the clusters. For $K = 2$, the population of 'Moulis' (in red) differs from the nine others (in green). At $K = 10$, the clusters correspond fairly well to the populations defined *a priori*.

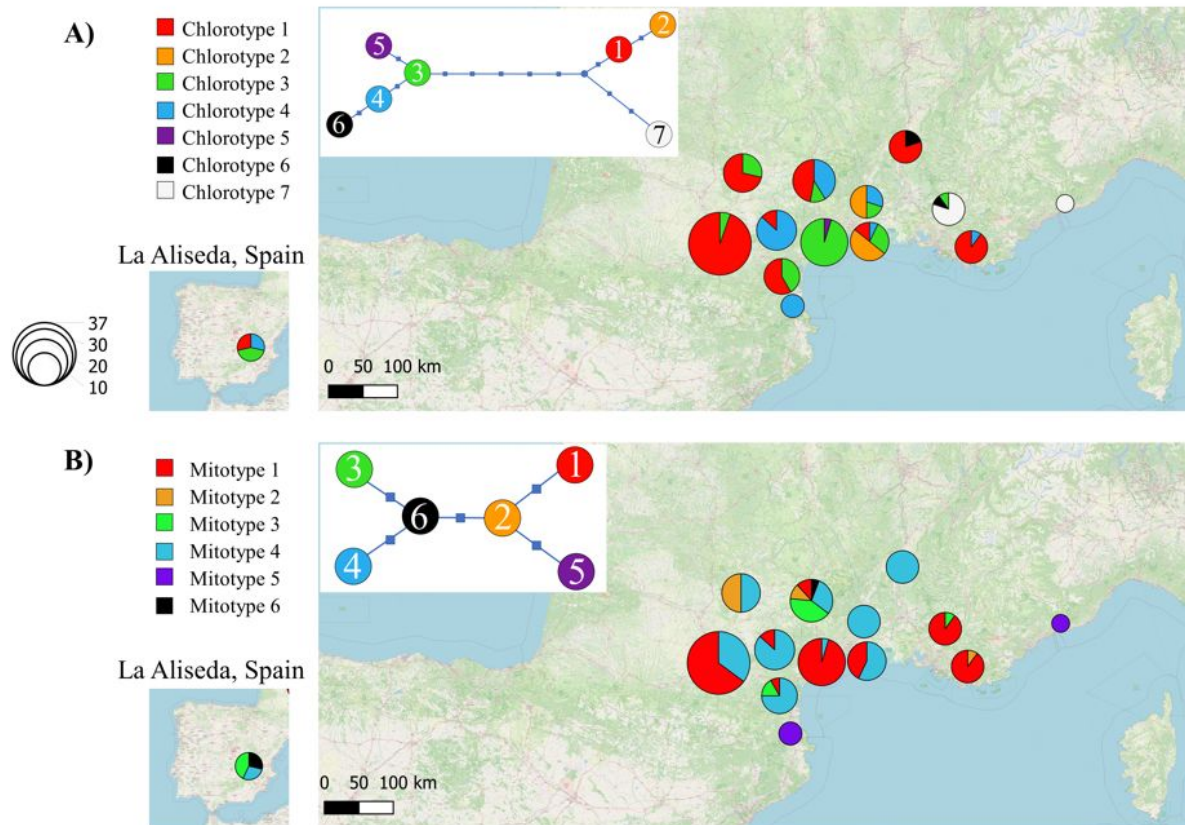


Figure 2. Distribution of (A) chloroplast and (B) mitochondrial haplotypes in southwestern France. The haplotypes found in a Spanish population ('La Aliseda', Jaén Province) are also indicated. In the left-hand, top corner of each map, the reduced-median haplotype network (reconstructed with NETWORK v.10.2; Bandelt, Forster & Röhl, 1999) is given for each organellar genome.

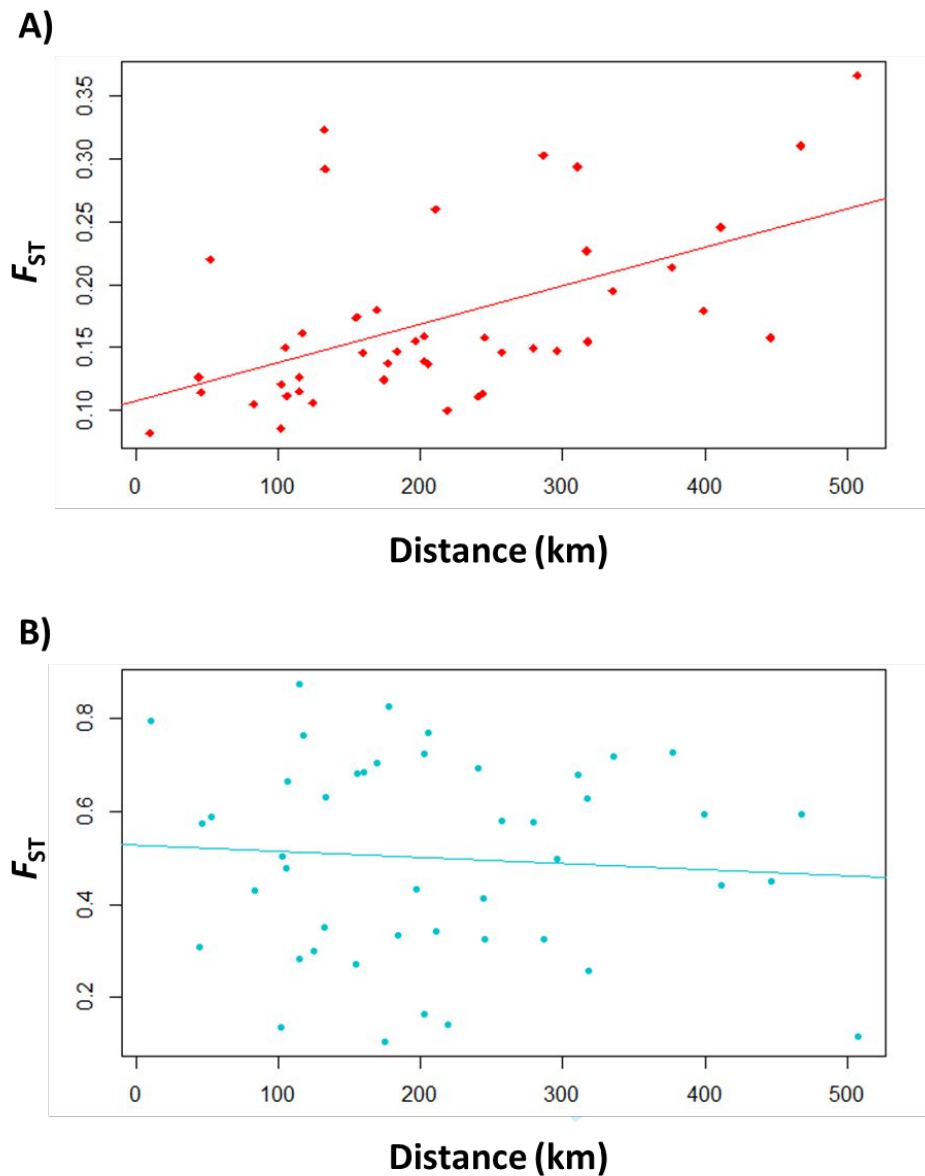


Figure 3. Genetic differentiation between populations as a function of geographical distance for (A) nuclear and (B) cytoplasmic data. Tests of Mantel (1967) were done with 999 permutations using the R package *ade4* (Chessel, Dufour & Thioulouse, 2004). Genetic distances (F_{ST}) and geographical distances between population pairs show a significant relationship (Mantel test, $r = 0.528$, p -value = 0.003) for nuclear data, but not for cytoplasmic data ($r = -0.052$, p -value > 0.05).

Supporting information

Limitation of gene flow by distance in the common yellow jasmine (*Chrysojasminum fruticans*, Oleaceae): implications for the study of its mating strategies

Aurore Puyooou *et al.*

SI Appendix includes:

Table S1. Plant sampling and complete genotyping dataset - (see excel file)

Table S2. List of herbarium specimens of the MNHN-P included in our study

Table S3. Characteristics of the 23 nuclear microsatellite loci developed in the present study

Table S4. Characteristics of the six nuclear indel loci developed in the present study

Table S5. Characteristics of chloroplastic and mitochondrial loci developed in the present study

Table S6. Genotyping of 36 progenies issued from parents with distinct chloroplast and/or mitochondrial DNA haplotypes in the 'CEFE' and 'Moulis' populations

Table S7. Measure of the LD between organellar genomes at the population scale

Table S8. Genetic diversity parameters estimated on both nuclear and cytoplasmic loci within populations with at least seven individuals

Table S9. F_{ST} values calculated between each pair of populations

Table S10. Results of the assignment test comparing nuclear SSR genotypes of individuals of the CEFE collection to nuclear SSR genotypes of natural populations

Figure S1. Dendrogram representing genetic similarity between all individuals based on genetic distances of Bruvo *et al.* (2004) and reconstructed with the unweighted average linkage clustering criterion.

Figure S2. Dendrogram representing genetic similarity between populations based on pairwise F_{ST} estimated from nuclear data and reconstructed on the unweighted average linkage clustering criterion

Figure S3. Cytoplasmic genetic differentiation between populations as a function of geographical distance according to different F_{ST} estimation methods

References

Table S2. List of herbarium specimens of the MNHN-P (Muséum National d'Histoire Naturelle de Paris) included in our study. For each specimen, the collector name (with a collector number when available) and the herbarium inventory number (in brackets) are given. Samples can be viewed at the [MNHN website](#). The geographic origin (locality and country) and date of collection are also indicated (“00” means that day or month are not indicated on the label).

MNHN-P specimen	Locality, country	Date of collection
J. Lewalle 9060 (P03426484)	Oued Cherrad, Morocco	1979-05-01
J.M. Losa Quintana 17304 (P03883854)	Granada, Spain	1994-05-29
E. Bourgeau (P04074654)	Algarve, Portugal	1853-05-30
E. Reverchon 1182 (P04074663)	Granada, Spain	1900-06-00
A. Faure (P01194578)	Oran, Algeria	1940-03-17
A. Letourneux (P03426428)	Ghardimaou, Tunisia	1884-03-31
J. Bourcart (P04074655)	Pishkupije, Albania	1923-06-00
V. Stribny (P04074662)	Philippopolin, Bulgaria	1893-06-00
B. Balansa (P03869118)	Smyrne, Turkey	1854-05-05
A.H. Brotherus 787 (P03869098)	Gori, Georgia	1881-07-00
R. Gombault 4458 (P03869122)	Zahlé, Lebanon	1930-06-09
A. Callier 252 (P03869093)	Soudak, Crimea	1896-00-00
R. Gombault 4456 (P03869103)	Radjou, Syria	1934-04-15
G. Remaudière 202 (P04254353)	South Sari, Iran	1966-00-00
P. Sintenis 1838 (P03869104)	Kisil-Arwat, Turkmenistan	1901-05-25

Table S3. Characteristics of the 23 nuclear microsatellite (SSR) loci developed in the present study. Name, putative homolog in olive [with SSR gene location: coding sequence (cds), intron, or untranscribed region (UTR)], GenBank accession nos, SSR motif (number of repeats based on the sequenced allele in *C. fruticans*), forward and reverse primers, fluorochrome used (Flu), PCR multiplex (Mx; locus amplified in simplex when no indication), allele size range, and number of alleles revealed (Na). * a M13 tail (TGTAACGACGGCCAGT) was added on the 5' extremity of each forward primer. ** Two loci are amplified with GAA20 primers (allele intensity of locus B is lower; allele size range indicates they may occasionally overlap).

Locus ID	Putative homolog in olive (encoded protein) – marker location in gene	GenBank no	SSR Motif	Forward primer*	Reverse primer	Flu	Mx	Allele size	Na
Cf-ATT-1	LOC111367380 (Kinesin prot.) – 3' UTR	OQ124178	(ATT) ₉₋₁₀	TGTGGACTCATTTTATTCTC	TACACTCCTTAGACGGTGGA	HEX	-	130-158	7
Cf-ATT-2	No gene homology	OQ124179	(AAT) ₉	TGTGTCACGCTGGTTTTTC	CCCGATGTGTTTTGGATACA	HEX	-	188-203	5
Cf-ATT-3	LOC111367396 (F-box prot.) – 3' UTR	OQ127180	(AAT) ₁₁₋₁₂	CCCACATTTGACTCCTCATC	GAATACGATTATAGAAGTGTGTC	HEX	-	143-160	8
Cf-GAA-1	LOC111372306 (Translation initiation factor) – cds	OQ230738	(GAA) ₁₅	TGCGTTTGACAAAGGATGCT	CAGATGGATTAACATATTTCCAC	FAM	Dx-I	154-193	7
Cf-GAA-2	LOC111405212 (Unknown prot.) – 3' UTR	OQ230739	(GAA) ₉₋₁₀	GAAGTGAACAATCAGACACC	CCTCAATGACTTGCCTTTGGT	AT550	-	204-213	4
Cf-GAA-3	No gene homology	OQ230740	(GAA) ₁₃₋₁₄	TCAGTAGTATGTGCTAATTGATG	AAGGTTCTACCAATGGTATCT	FAM	Dx-II	205-232	9
Cf-GAA-4	No gene homology	OQ124164	(GAA) ₁₃₋₁₅	CGGTGAAAGGGTAGCAGATA	ATAGGAATGCTCGAAGTATCCT	FAM	Dx-II	247-307	19
Cf-GAA-5	LOC111387121 (Beta-1,3-galactosyl transferase) – 5' UTR	OQ124165	(GAA) ₆	CACCAAGTCGACCAACTGT	GCCTGTCTTCTACTCTTCTC	HEX	Dx-III	161-170	6
Cf-GAA-6	No gene homology	OQ124166	(GAA) ₇₋₁₇	GCTTCCTCRAACCAATCATG	GAACAYTGATGTTGCGCTTGG	AT565	-	104-161	11
Cf-GAA-7	LOC111411240 (ATP-dependent 6-phosphofructo kinase) – 5' UTR	OQ124167	(GAA) ₁₃₋₁₄	ACCTCAGTTTCAGTTGAAAGCT	CTTTCCGAATTTGAAATTTGGTAG	HEX	Dx-III	235-250	5
Cf-GAA-9	LOC111395981 (Unknown prot.) - intron	OQ124168	(GAA) ₁₇₋₁₈ (GAT) ₆ GAA(GAT) ₇	ATTTGGTATGGTCTCTGTTG	TTTCTCAAAATCAGCCATTTTCGA	FAM	-	250-289	13
Cf-GAA-11	LOC111378787 (Fasciclin arabino galactan prot.) – cds?	OQ124169	(AAG) ₁₁	TGTGCTTGAAACAGCTCCT	AATGCCATGGATTTCCAGATCT	FAM	Dx-I	117-132	6
Cf-GAA-14	No gene homology	OQ124170	(GAA) ₂₅₋₂₆	AATAACGACCAACACCGATCAA	GGAGTTTTTCACCTATTCTCCT	FAM	-	101-182	23
Cf-GAA-16	No gene homology	OQ124112	(AAG) ₁₈₋₂₀	AAGCCCCAATCCATTATCTAAC	AAGCTAATATCCACAATCAAAGC	FAM	-	162-243	23
Cf-GAA-17	LOC111374342 (60S ribosomal prot.) – 5' UTR	OQ124172	(GAA) ₁₈₋₂₃	CACATGCTTGAGTTTGACAAG	CTGGGAGACAATTGTTTCC	FAM	-	206-257	14
Cf-GAA-18	LOC111402076 (Histone acetyl transferase) – cds	OQ124173	(GAA) ₂₁₋₂₃	AAGACGGACACCAATGCACT	AGCTTTGCGGCAAGTAAGGA	AT565	-	154-214	16
Cf-GAA-19	No gene homology	OQ124174	(TAA) ₁₂₋₁₃ (GAA) ₂₇₋₃₀	ATAAGATGCATATGCATCTACCTA	ATCGCTAAAAGGTCTTCTTCTCCT	AT565	-	191-284	25
Cf-GAA-20	No gene homology	OQ124175	(GAA) ₂₄₋₂₅	GACCTGAGATTCCAGATCGA	GACAAAGTCTGGAAGCTTCG	HEX	-	112-151	15
-(B)**	-	-	-	-	-	HEX	-	94-115	8
Cf-GAA-21	No gene homology	OQ124176	(GAA) ₁₁₋₂₆ (TAA) ₈₋₉	TTCTTTGTTGATGCAACAAATC	AACCAAGAAACTTGTGTGGACT	HEX	-	204-357	36
Cf-GAA-22	No gene homology	OQ124177	(GAA) ₂₆₋₃₂	TGCTCTAGAGCTATAGTTG	AAGCAAGATATTGAGGAATTCAAG	AT565	-	131-230	32
Cf-ATA-1	No gene homology	OQ808969	(ATA) ₁₂₋₁₃	CTTGTTGTTACGACAGTGCT	TAACTTCTAACCAACTCCATAGT	FAM	-	223-241	7
Cf-TA-1	LOC111402738 (Transcription factor) – 5' UTR	OQ808968	(TA) ₁₁₋₁₂	GGACTTCTTTCTTTAGTTGTATG	AAACCAAGATGGAATATTACCTCA	AT565	-	213-223	3

Table S4. Characteristics of the six nuclear indel loci developed in the present study. Name, putative homolog in olive [with indel gene location: coding sequence (cds), intron, or untranscribed region (UTR)], GenBank accession nos, polymorphic motif, forward and reverse primers, fluorochrome used (Flu), PCR multiplex (Mx; locus amplified in simplex when no indication), allele size range, and number of alleles revealed (Na). * a M13 tail (TGTAACGACGGCCAGT) was added on the 5' extremity of each forward primer.

Locus ID	Putative homolog in olive (encoded protein) - marker location in gene	GenBank no	Polymorphic motif	Forward primer*	Reverse primer	Flu	Mx	Size range	Na
Cf-ind-1	LOC111410094 (Fructose-bisphosphate aldolase) – intron	OQ124181	Indel 25 bp	GACCCTCATCAAGTATTTCCA	CATTAGGACAGTCGTAACATGT	FAM	Dx-IV	136-161	2
Cf-ind-2	LOC111367381 (Chaperone-like protein) – 5' UTR	OQ124182	(AT) _{5,6} + indel 8 bp	GGAATCTATCATCAAACACTATAT C	CTAATCATTTTCTTGATTGCGGT	HEX	-	274-285	6
Cf-ind-3	LOC111392924 (Trans-resveratrol di-O-methyl transferase) – Intron	OQ124183	Indel 2 bp	CAGGGATCTTCCAAATCCGA	AGCTAGCGTGCATATCTTAGT	AT565	Dx-V	121-123	2
Cf-ind-4	LOC111406052 (Pollen allergen Ole e 1) – 5'UTR	OQ124184	Indel 7 bp	CTTCCCGAACACGATACTC	ATTGAGAGACTGGACGAACG	AT565	Dx-V	168-175	2
Cf-ind-5	LOC111367387 (GATA transcription factor) - 3' part	OQ124185	Indel 20 bp	ATGGCATGGGTTTGAGTAGC	CTTTTAAAGTCTAGCTAGAGAGA	FAM	-	101-121	2
Cf-ind-6	LOC111367389 (Cytochrome P450) – intron	OQ124186	Indels 6+37 bp	AGCTCACTGTGCTCTCT	AGCTTCGCAATAGTATTGGAC	FAM	Dx-IV	150-200	3

Table S5. Characteristics of chloroplastic and mitochondrial loci developed in the present study. Name, indel length or microsatellite motif (number of repeats based on the sequenced allele in *C. fruticans*), forward and reverse primers, fluorochrome used (Flu), PCR multiplex (Mx), allele size range, and number of alleles revealed (Na). * a M13 tail (TGTAACGACGGCCAGT) was added on the 5' extremity of each forward primer.

Locus ID	GenBank no	Polymorphic motif	Forward primer*	Reverse primer	Flu	Mx	Size range	Na
- Plastome								
	MH559274							
CP1		(T) ₁₃₋₁₄	cp17-For (olive; Besnard <i>et al.</i> , 2011)	CGATCAGGGTCTTTTTGTATCAC	HEX	cpII	176-177	2
CP2		(C) ₉₋₁₂ WWW(G) ₁₃₋₈	GGAGTTTGGCTCTATTAAGTAG	ATGATAGTGTTCGAGTCTATC	HEX	cpI	116-120	3
CP3		(A) ₈₋₉	GGATTTATCGACATCAAGTTCG	CCGTACAAGGAATCCGTAG	HEX	cpII	166-167	2
CP4		(TAATAATA) ₁₂ + (A) ₈ (TA) ₁₀	CAGATGTATCAGAAGCTTTTCGAT	TTGGTCGACTAGATATCGTACA	HEX	cpIII	349-357	2
CP5		(A) ₈₋₁₀ TTC(A) ₈₋₁₀	GCATGATGACAAGAATTCTATCT	ATAGGAACGAAATAAACTCGAAGA	HEX	cpI	126-129	4
CP6		(T) ₉ +(T) ₈₋₁₀	cp45-For (olive; Besnard <i>et al.</i> , 2011)	GTTTTCAAACCGAAGGGCTC	HEX	cpII	136-137	2
CP7		(A) ₁₁₋₁₂	cp46-For (olive; Besnard <i>et al.</i> , 2011)	cp46-Rev (olive; Besnard <i>et al.</i> , 2011)	HEX	cpI	108-109	2
CP8		(A) ₈₋₁₀	CTTGAGCCACACCTACATTG	GAATCTATGGAGGGCCAT	HEX	cpIV	90-91	2
CP9		(T) ₁₄₋₁₅	GGAGAAAAAGAGGCATTACCT	AATCACGTGTCTTTCTCGTAAG	HEX	cpIV	110-111	2
CP10		Indel 8 bp	TTGGAACAAATTACTCTAATGCGT	CGGAGGATGATTCCAAGGA	HEX	cpIV	236-245	2
- Mitogenome								
	OQ124187-96							
MT1		Indel 8 bp	TGCATAAGGACTGTAAGCT	CGATAGCAGACATCTCCATC	HEX	mtI	301-309	2
MT2		(A) ₈₋₁₀	CTACCTGCTTCTACGGAAGA	AGGGAACAAGCAACGGATGA	HEX	mtI	316-317	2
MT3		Indel 8 bp	GTATAGTGGACACGAGCA	TCTGTACTCCTCTCATGC	HEX	mtI	327-335	2
MT4		Indel 8 bp	GTAATATCCTATTGGTCAAATCA	CGTATACAGGATTTTCAGTATTAGA	HEX	mtII	231239	2
MT5		Indel 8 bp	TAGTAATAGGCTCAACTACAAGT	ACCTTCTATTTCGGCCAAGT	HEX	mtII	217-225	2

Table S6. Genotyping of 36 progenies issued from parents with distinct chloroplast and/or mitochondrial DNA haplotypes in the 'CEFE' and 'Moulis' populations. In these two locations, a dense sampling was performed to characterize the genotype of most putative parents in the population (data available on request). Floral morph (L = Longistylous; B = Brevistylous) of each parent is also recorded. Father of each progeny was identified from nuclear microsatellite data with CERVUS (Marshall et al., 1998). Mitochondrial and chloroplast haplotypes were characterized on all parents and progenies, and highlighted in colour when different from both parents. A maternal contribution was systematically revealed in all progenies for both genomes. MD = Missing data. All crosses were between L and B (12), or B and L (24).

Table with columns for Individual, Floral morph (GAA20_B, GAA20, GAA21, GAA13, GAA22, GAA19, GAA16, GAA17, GAA1, GAA11, GAA3, GAA4, GAA2, GAA9, GAA5, GAA7, GAA14, AT3, GAA6, AT1, indel1, indel6, TA-1, ATA-1), mtDNA haplotype (MT1-MT5), cpDNA haplotype (CP1-CP10), and a final column with numerical values. Rows list various progenies and their parents (Mother/Father) with associated genotypes and haplotypes.

Table S6, end.

Individual		GAA20_B	GAA20	GAA21	GAA18	GAA22	GAA19	GAA16	GAA17	GAA1	GAA11	GAA3	GAA4	GAA2	GAA9	GAA5	GAA7	GAA14	ATT3	GAA6	ATT1	indel1	indel6	TA-1	ATA-1	MT1	MT2	MT3	MT4	MT5	mDNA haplotype										cpDNA haplotype									
Mother	CEFE_S4	B	106:106	115:121	210:288	196:202	173:185	212:221	162:168	236:251	169:172	126:126	214:220	253:268	204:204	274:280	163:163	241:244	152:155	151:151	104:104	156:156	161:161	150:194	221:221	229:235	301	317	335	239	225	4	177	120	166	357	127	137	109	90	110	245	2	1						
Father	CEFE_136	L	106:106	118:121	210:210	175:184	170:200	209:215	168:198	236:251	169:169	126:126	211:220	286:301	213:213	274:277	163:170	241:241	134:137	151:157	104:104	156:156	MD	150:150	221:221	226:235	309	316	327	239	225	1	177	119	166	357	127	137	109	90	110	245	1	1						
Progeny (2021)	CEFE_S4s3A	MD	106:106	118:121	MD	175:202	MD	212:215	162:198	236:251	169:172	MD	MD	268:286	204:213	MD	MD	241:241	134:152	151:157	MD	MD	MD	MD	MD	229:235	301	317	335	239	225	4	15	177	120	166	357	127	137	109	90	110	245	2	17					
Progeny (2021)	CEFE_S4s3B	MD	115:121	MD	175:202	MD	MD	162:198	251:251	169:172	MD	MD	268:286	204:213	MD	MD	241:244	134:152	151:157	MD	156:156	MD	MD	MD	235:235	301	317	335	239	225	4	16	177	120	166	357	127	137	109	90	110	245	2	18						
Progeny (2021)	CEFE_S4s6	MD	115:121	MD	175:196	MD	MD	168:198	236:251	169:169	MD	MD	MD	MD	MD	MD	241:241	137:155	151:157	MD	MD	MD	MD	MD	MD	301	317	335	239	225	4	17	177	120	166	357	127	137	109	90	110	245	2	19						
Progeny (2021)	CEFE_S4s8	MD	106:106	121:121	MD	184:202	MD	209:212	162:168	251:251	169:172	MD	MD	MD	MD	MD	MD	241:241	134:152	151:151	MD	MD	MD	MD	MD	301	317	335	239	225	4	18	177	120	166	357	127	137	109	90	110	245	2	20						
Mother	CEFE_S4	B	106:106	115:121	210:288	196:202	173:185	212:221	162:168	236:251	169:172	126:126	214:220	253:268	204:204	274:280	163:163	241:244	152:155	151:151	104:104	156:156	161:161	150:194	221:221	229:235	301	317	335	239	225	4	177	120	166	357	127	137	109	90	110	245	2	1						
Father	CEFE_109	L	103:109	118:121	210:210	175:196	170:200	209:221	168:198	236:236	169:169	126:126	211:214	253:301	204:204	274:277	170:170	241:244	134:152	148:151	104:104	136:156	MD	150:150	221:221	226:232	309	316	327	239	225	1	177	119	166	357	127	137	109	90	110	245	1	1						
Progeny (2021)	CEFE_S4s4	MD	106:106	115:118	210:210	175:202	MD	209:221	162:198	236:251	169:172	MD	MD	204:204	MD	MD	241:244	134:155	151:151	MD	MD	MD	MD	MD	MD	232:235	301	317	335	239	225	4	19	177	120	166	357	127	137	109	90	110	245	2	21					
Mother	CEFE_S4	B	106:106	115:121	210:288	196:202	173:185	212:221	162:168	236:251	169:172	126:126	214:220	253:268	204:204	274:280	163:163	241:244	152:155	151:151	104:104	156:156	161:161	150:194	221:221	229:235	301	317	335	239	225	4	177	120	166	357	127	137	109	90	110	245	2	1						
Father	CEFE_L6	L	103:109	121:139	288:288	190:202	167:200	221:257	162:207	215:254	163:169	126:126	214:214	253:253	204:204	277:277	163:163	235:241	149:152	148:154	104:143	156:156	136:161	150:150	221:221	226:235	309	316	327	239	225	1	176	119	167	349	128	136	108	90	110	245	3	1						
Progeny (2021)	CEFE_S4s5	MD	103:106	121:121	MD	190:202	MD	212:257	162:207	215:236	169:172	MD	MD	MD	MD	MD	235:241	149:152	148:151	MD	MD	MD	MD	MD	MD	301	317	335	239	225	4	20	177	120	166	357	127	137	109	90	110	245	2	22						
Mother	CEFE_S6	B	106:106	118:136	210:291	199:202	179:203	221:236	162:162	236:251	169:169	126:126	214:214	253:253	204:204	277:280	170:170	241:244	134:152	151:151	104:104	156:156	136:161	150:194	213:221	229:235	301	317	335	239	225	4	177	120	166	357	127	137	109	90	110	245	2	1						
Father	CEFE_L5	L	106:106	115:142	204:294	190:199	188:227	218:257	198:198	215:239	166:169	126:126	211:214	268:280	204:204	265:277	163:170	235:241	140:140	154:157	104:104	136:136	161:161	150:194	213:221	229:229	309	316	327	239	225	1	176	119	167	349	128	136	108	90	110	245	3	1						
Progeny (2021)	CEFE_S6s1A	MD	106:106	115:136	MD	199:199	MD	162:198	215:236	169:169	MD	MD	MD	MD	MD	MD	241:244	140:152	151:154	MD	MD	MD	MD	MD	MD	301	317	335	239	225	4	21	177	120	166	357	127	137	109	90	110	245	2	23						
Progeny (2021)	CEFE_S6s2	MD	115:136	MD	199:199	MD	218:221	162:198	215:236	166:169	MD	MD	MD	MD	MD	MD	241:244	140:152	157:157	MD	MD	MD	MD	MD	MD	301	317	335	239	225	4	22	177	120	166	357	127	137	109	90	110	245	2	24						
Progeny (2021)	CEFE_S6s3	MD	136:142	MD	199:202	MD	MD	162:198	215:236	166:166	MD	MD	MD	MD	MD	MD	235:241	140:152	151:154	MD	MD	MD	MD	MD	MD	301	317	335	239	225	4	23	177	120	166	357	127	137	109	90	110	245	2	25						
Progeny (2021)	CEFE_S6s6	MD	106:106	136:142	MD	199:199	MD	162:198	236:239	169:169	MD	MD	MD	MD	MD	MD	241:244	134:140	151:154	MD	MD	MD	MD	MD	MD	301	317	335	239	225	4	24	177	120	166	357	127	137	109	90	110	245	2	26						
Mother	CEFE_S6	B	106:106	118:136	210:291	199:202	179:203	221:236	162:162	236:251	169:169	126:126	214:214	253:253	204:204	277:280	170:170	241:244	134:152	151:151	104:104	156:156	136:161	150:194	213:221	229:235	301	317	335	239	225	4	177	120	166	357	127	137	109	90	110	245	2	1						
Father	CEFE_136	L	106:106	118:121	210:210	175:184	170:200	209:215	168:198	236:251	169:169	126:126	211:220	286:301	213:213	274:277	163:170	241:241	134:137	151:157	104:104	156:156	MD	150:150	221:221	226:235	309	316	327	239	225	1	177	119	166	357	127	137	109	90	110	245	1	1						
Progeny (2021)	CEFE_S6s7	MD	106:106	121:136	MD	175:199	MD	170:203	209:236	162:198	236:251	MD	MD	253:301	204:213	MD	MD	241:241	134:134	151:157	MD	MD	MD	MD	MD	226:229	301	317	335	239	225	4	25	177	120	166	357	127	137	109	90	110	245	2	27					
Mother	CEFE_S6	B	106:106	118:136	210:291	199:202	179:203	221:236	162:162	236:251	169:169	126:126	214:214	253:253	204:204	277:280	170:170	241:244	134:152	151:151	104:104	156:156	136:161	150:194	213:221	229:235	301	317	335	239	225	4	177	120	166	357	127	137	109	90	110	245	2	1						
Father	CEFE_149	L	106:106	115:142	288:288	190:199	173:206	215:236	162:162	251:254	169:172	126:126	220:220	253:268	204:204	274:277	163:163	241:241	134:152	151:157	104:104	156:156	MD	150:150	221:221	232:235	309	316	327	239	225	1	176	119	167	349	128	136	108	90	110	245	3	1						
Progeny (2021)	CEFE_S6s1B	MD	118:142	MD	199:202	MD	MD	162:162	251:254	169:172	MD	MD	MD	MD	MD	MD	241:241	134:134	151:151	MD	MD	MD	MD	MD	MD	301	317	335	239	225	4	26	177	120	166	357	127	137	109	90	110	245	2	28						
Progeny (2021)	CEFE_S6s8	MD	106:106	136:142	MD	199:202	MD	162:162	236:254	169:172	126:126	MD	MD	MD	MD	MD	MD	241:241	134:152	151:157	MD	MD	MD	MD	MD	301	317	335	239	225	4	27	177	120	166	357	127	137	109	90	110	245	2	29						
Mother	CEFE_119	B	106:106	121:142	210:288	190:202	179:194	215:236	162:162	236:251	169:172	126:126	211:217	268:268	204:204	259:280	163:170	238:241	152:155	151:151	104:104	136:156	136:161	150:150	213:221	229:232	301	317	335	239	225	4	177	1																

Table S7. Measure of the LD (D' ; Lewontin, 1964) between organellar genomes at the population scale. Significance (p -value) of the D' estimate was measured with a Fisher's test. This test was only done on populations where polymorphism was observed on both organellar genomes (see Table S1). Ng = Number of genotypes.

Population	Ng	D'	p-value
La Aliseda	7	1.000	0.005
Bages	12	1.000	0.048
CEFE	15	0.875	<0.001
Celleneuve	21	1.000	1.000
Eguilles	10	1.000	0.212
Moulis	37	1.000	0.118
Moussoulens	12	1.000	1.000
St-Maurice-de-Navacelles	17	0.558	0.181
Puycelsi	15	0.500	0.073
Sainte-Baume	10	1.000	0.109

Table S9. F_{ST} values calculated between each pair of populations. The parameters were calculated using GENETIX (Belkhir *et al.*, 2004). Values above the diagonal correspond to values obtained with nuclear loci, values below to those obtained with cytoplasmic data (considering chloroplast and mitochondrial data as two different loci). ^{ns} Values not significantly different from 0 at the 5% threshold, estimated on a total of 1000 permutations. All other values are significantly different from 0.

Population	Pop. Code	ALB	BAGE	BIDO	CEFE	CELL	EGUI	MOUL	MOUS	SMN	PLL	PUYC	BAUM	TURB	ALIS
Amelie-Les-Bains	ALB		0,220	0,212	0,194	0,244	0,204	0,433	0,265	0,169	0,269	0,230	0,322	0,260	0,318
Bages	BAGE	0,672		0,100	0,078	0,112	0,146	0,260	0,105	0,086	0,115	0,124	0,147	0,208	0,239
Bidon	BIDO	0,880	0,143 ^{ns}		0,084	0,126	0,161	0,246	0,149	0,106	0,150	0,155	0,180	0,209	0,298
CEFE	CEFE	0,563	0,179	0,384		0,047	0,107	0,258	0,120	0,070	0,071	0,106	0,105	0,149	0,258
Celleneuve	CELL	0,923	0,664	0,874	0,454		0,146	0,294	0,138	0,114	0,082	0,111	0,137	0,194	0,306
Eguilles	EGUI	0,808	0,579	0,764	0,394	0,684		0,311	0,195	0,155	0,174	0,179	0,220	0,238	0,324
Moulis	MOUL	0,763	0,343	0,442	0,387	0,680	0,593		0,292	0,303	0,227	0,323	0,366	0,445	0,419
Moussoulens	MOUS	0,688	0,431	0,576	0,384	0,825	0,718	0,632		0,174	0,147	0,121	0,214	0,292	0,232
St-Maurice-Navacelles	SMN	0,416	0,136	0,299	0,210	0,575	0,434	0,327	0,272		0,126	0,139	0,158	0,108	0,209
Prades-le-Lez	PLL	0,734	0,284	0,478	0,149	0,795	0,680	0,629	0,333	0,307		0,113	0,159	0,255	0,311
Puycelsi	PUYC	0,659	0,106 ^{ns}	0,259	0,284	0,693	0,593	0,350	0,503	0,165	0,413		0,158	0,223	0,242
Sainte-Baume	BAUM	0,858	0,499	0,705	0,409	0,770	0,589	0,117 ^{ns}	0,728	0,327	0,724	0,451		0,265	0,341
La Turbie	TURB	1,000	0,628	0,859	0,530	0,916	0,639	0,749	0,795	0,472	0,736	0,619	0,842		0,310
La Aliseda	ALIS	0,497	0,107 ^{ns}	0,410	0,177	0,612	0,481	0,464	0,383	-0,022 ^{ns}	0,308	0,207	0,478	0,485	

Table S10. Results of the assignment test comparing nuclear SSR genotypes of individuals of the CEFE collection to nuclear SSR genotypes of natural populations (with at least ten sampled individuals). We first considered the 15 genotypes initially used in our population genetic study, and then, to increase the power of this analysis, we considered 125 genotypes of the CEFE collection (data not shown). The analysis was performed with GENECLASS2 (Piry *et al.*, 2004); in both cases, 27% of individuals (4/15, 34/125) were assigned to at least one putative population (all in Hérault, at maximum 30 km from Montpellier). This result is consistent with our expectations, since the population was initially funded with individuals from different localities around Montpellier, such as Causse-de-la-Selle near St-Maurice-de-Navacelles (J.D. Thompson, pers. com.). An arbitrary threshold probability value of 0.05 or greater was applied to determine the origin of CEFE individuals.

Population	Percent of CEFE individuals assigned (15 genets)	Percent of CEFE individuals assigned (125 genets)
Bages	0	0
Bidon	0	0
Celleneuve	13.3	5.6
Eguilles	0	0
Moulis	0	0
Moussoulens	0	0
St-Maurice-de-Navacelles	20.0	24.0
Puycelsi	0	0
Prades-le-Lez	0	2.4
Sainte-Baume	0	0

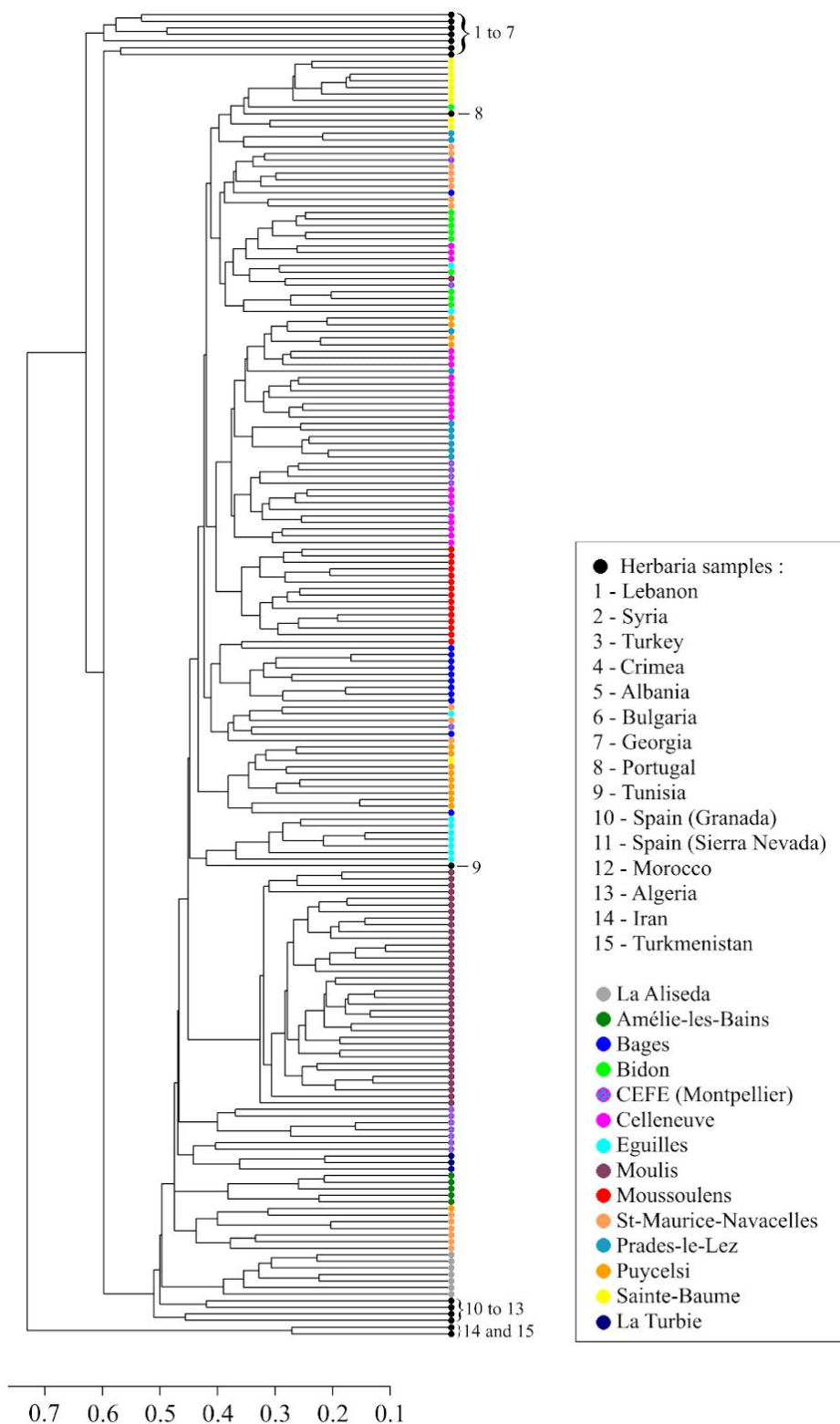
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure S1. Dendrogram representing genetic similarity between all individuals based on genetic distances of Bruvo *et al.* (2004) and reconstructed with the unweighted pair group method with arithmetic mean (UPGMA) using *Poppr* (Kamvar *et al.*, 2014). French individuals are grouped in the same cluster with the individuals from Tunisia and Portugal. Individuals from the same population also tend to cluster together (e.g. 'Moulis', 'La Aliseda').

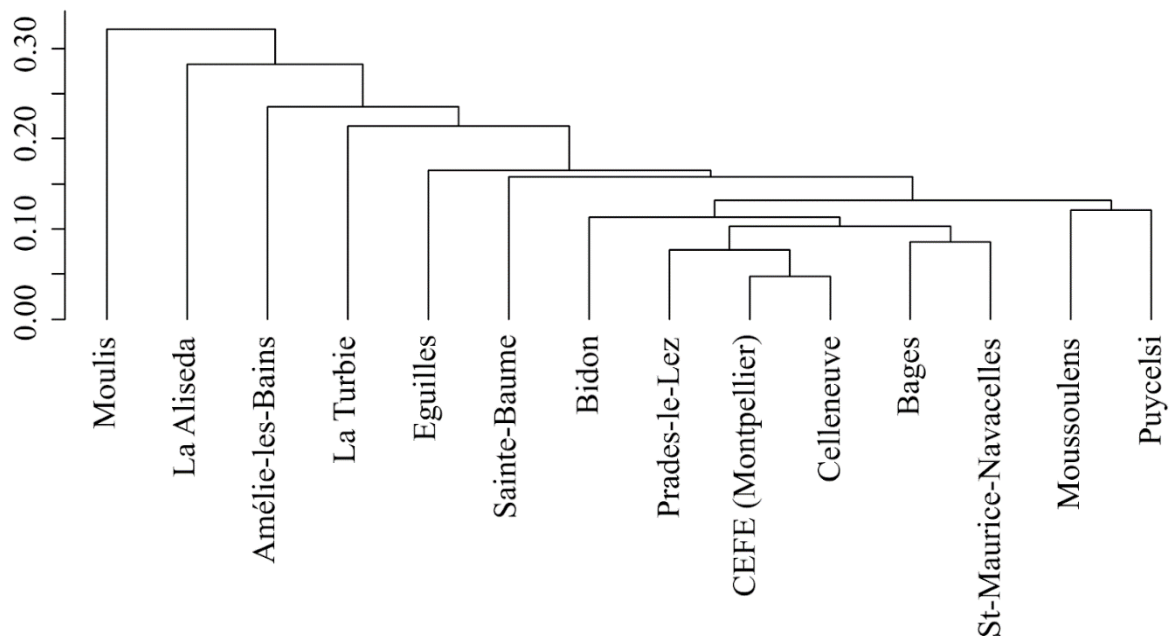


Figure S2. Dendrogram representing genetic similarity between populations based on pairwise F_{ST} estimated from nuclear data and reconstructed on the unweighted pair group method with arithmetic mean (UPGMA) with *Poppr* (Kamvar *et al.*, 2014). With the exception of the 'Moulis' population, geographically close populations tend to cluster together in this dendrogram (e.g. 'Prades-le-Lez', 'CEFE' and 'Celleneuve').

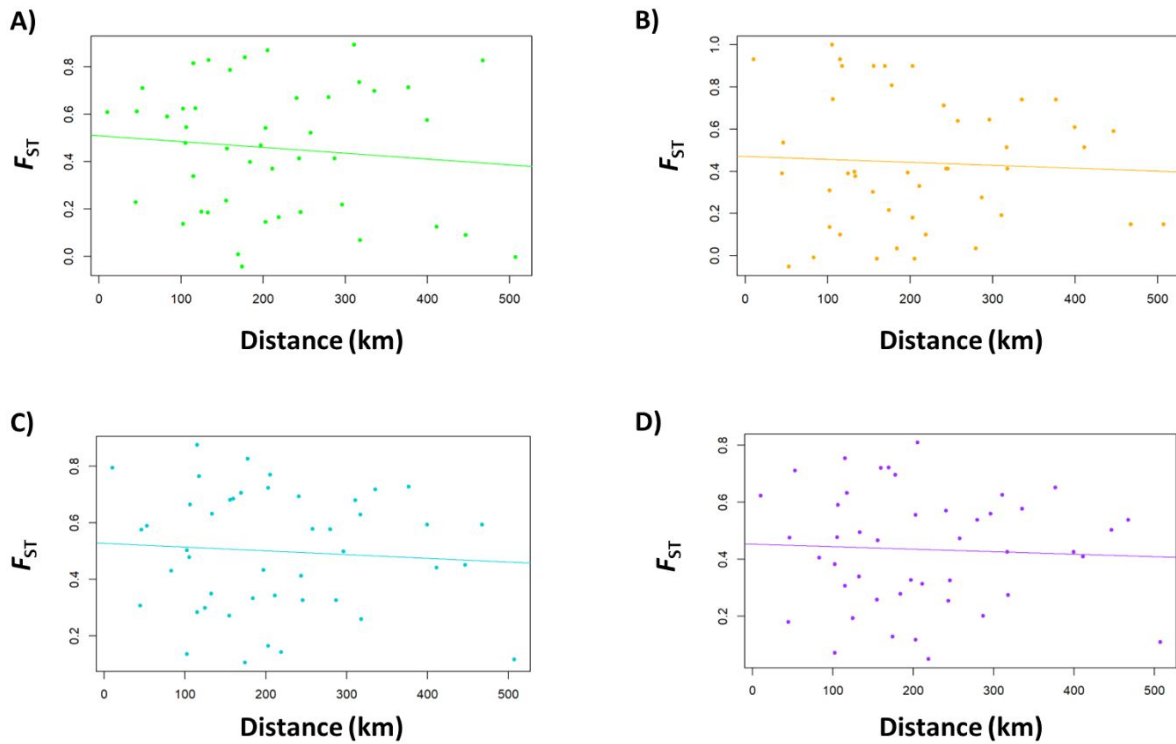


Figure S3. Cytoplasmic genetic differentiation between populations as a function of geographical distance according to different F_{ST} estimation methods: **(A)** with chloroplast data alone, **(B)** with mitochondrial data alone, **(C)** with chloroplast and mitochondrial data considered as two independent cytoplasmic loci, and **(D)** by combining chloroplast and mitochondrial information as a single locus for defining cytoplasmic haplotypes. Tests of Mantel (1967) were done with 999 permutations using the R package *ade4* (Chessel *et al.*, 2004). Genetic and geographical distances between population pairs do not show any significant relationship (Mantel test, $r = 0.005, -0.110, -0.009,$ and $-0.064,$ respectively for A, B, C and D; p -value > 0.05) regardless of how the F_{ST} values are estimated.

References

- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. 2004.** GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).
- Besnard G, Hernández P, Khadari B, Dorado G, Savolainen V. 2011.** Genomic profiling of plastid DNA variation in the Mediterranean olive tree. *BMC Plant Biology* **11**: 80.
- Bruvo R, Michiels NK, D'souza TG, Schulenburg H. 2004.** A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology* **13**: 2101–2106.
- Chessel D, Dufour A, Thioulouse J. 2004.** The ade4 package - I: One-table methods. *R News* **4**: 5–10.
- Goudet J. 2005.** HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. *Molecular Ecology Notes* **5**: 184–186.
- Gryta H, Van de Paer C, Manzi S, Holota H, Roy M, Besnard G. 2017.** Genome skimming and plastid microsatellite profiling of alder trees (*Alnus* spp., Betulaceae): Phylogenetic and phylogeographical prospects. *Tree Genetics & Genomes* **13**: 118.
- Kamvar ZN, Tabima JF, Grünwald NJ. 2014.** *Poppr*: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**: e281.
- Lewontin RC. 1964.** The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* **49**: 49–67.
- Mantel N. 1967.** The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**: 209–220.
- Marshall TC, Slate J, Kruuk LE, Pemberton JM. 1998.** Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* **7**: 639–655.
- Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A. 2004.** GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**: 536–539.
- Van Oosterhout C, Hutchinson W, Willis D, Shipley P. 2004.** MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**: 535–538.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

Table with columns for Species Name, Date, Locality, and various numerical data points. Includes entries for Moulii_115, Moulii_144, Moulii_161, etc., and a large section for St_Maurice_Navelles_L1 through L11. The table contains multiple rows of data for each species, with some cells highlighted in yellow.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

P04074663_Spain	37.95000	-2.450	00/06/1900	-	169:172	204:213	214:214	259:262	163:165	104:161	235:241	277:280	123:123	107:125	162:162	236:239	154:190	224:230	103:103	136:136	210:216	149:149	156:156	188:190	151:154	235:235	0	136:161	277:285	121:121	168:175	121:121	150:150	301	317	335	239	225	4	177	119	166	357	127	137	109	90	110	245	1
P04074654_Portugal	37.12000	-8.010	30/05/1853	-	163:163	204:204	214:214	253:253	163:163	104:104	241:241	268:274	123:126	152:152	198:201	215:215	184:196	212:221	106:106	130:133	210:210	140:164	156:158	190:190	151:151	229:232	0	136:161	279:279	121:121	168:175	121:121	150:150	301	317	335	239	225	4	177	119	166	349	127	137	109	91	111	245	7
P03426484_Morocco	33.81500	-7.109	5/1/1979	-	166:172	204:213	211:214	262:271	165:170	143:143	235:241	262:265	126:126	122:152	168:168	245:257	154:181	227:251	106:112	115:115	210:228	137:140	133:156	188:188	151:151	229:232	0	136:161	275:279	123:123	175:175	101:121	194:194	309	317	335	239	225	3	177	119	166	349	127	137	109	90	110	245	8
P01194578_Algeria	35.70000	-0.700	3/17/1940	-	163:163	204:213	214:217	253:253	163:163	104:104	235:241	265:271	123:126	134:149	162:162	242:251	181:181	215:215	109:112	115:115	213:216	137:188	133:156	188:188	151:157	232:232	0	136:136	275:279	121:121	175:175	121:121	150:194	309	316	335	231	225	5	177	119	166	349	127	137	109	91	111	245	7
P03426428_Tunisia	36.43800	8.44300	31/03/1884	-	166:166	204:204	211:214	265:268	163:170	104:143	238:241	277:277	123:126	137:143	162:189	242:245	190:193	236:239	103:106	124:136	116:231	137:161	133:133	188:188	151:157	235:235	0	136:136	275:277	121:121	175:175	0	150:194	309	316	335	239	225	2	177	119	166	349	127	137	109	91	111	245	7
P04074655_Albania	40.89500	20.78000	00/00/1923	-	163:193	204:210	211:214	256:268	170:170	137:137	235:241	259:259	123:126	116:131	204:207	230:230	154:154	218:221	103:106	130:130	228:240	140:140	133:133	191:191	143:143	0	0	136:136	275:275	121:121	168:168	0	150:150	309	317	335	239	225	3	177	119	167	349	127	137	108	90	110	245	10
P04074662_Bulgaria	42.05000	24.70000	00/06/1893	-	164:184	204:204	208:223	268:280	163:170	107:111	241:241	268:280	111:126	128:179	219:237	215:245	181:187	212:223	106:106	127:130	210:219	131:137	156:156	188:188	148:148	229:238	0	136:136	277:277	121:121	168:168	121:121	194:194	309	317	335	239	225	3	177	119	167	349	127	137	108	90	110	245	11
P03869116_Turkey	38.37000	27.00000	05/06/1854	-	163:169	204:204	226:226	0	163:170	104:137	235:235	268:268	120:123	110:128	215:222	224:248	154:181	209:242	106:109	112:118	213:249	131:143	133:133	0	151:157	0	0	136:136	275:275	123:123	168:168	121:121	150:150	309	317	335	239	225	3	177	119	167	349	127	136	108	90	110	245	9
P03869093_Crimea	44.84000	34.96000	06/04/1896	-	169:193	207:210	226:232	262:277	161:163	104:104	235:235	268:271	123:126	128:134	204:213	248:248	154:154	248:257	106:109	127:127	255:258	146:146	133:133	190:191	147:149	223:235	0	136:136	274:276	121:121	168:175	121:121	150:150	309	317	335	239	225	3	177	116	167	349	127	137	108	90	110	245	12
P03869098_Georgia	41.92000	44.11000	00/07/1881	-	166:193	204:207	211:220	277:277	166:170	104:122	241:241	268:268	123:123	128:149	210:228	239:257	154:181	254:269	106:106	115:142	240:288	152:155	156:156	190:190	149:149	223:223	0	136:136	275:275	121:121	175:175	101:121	150:150	309	317	335	239	225	3	177	116	167	349	127	137	108	90	110	245	12
P03869122_Lebanon	33.87000	35.88000	6/9/1930	-	163:163	207:213	212:211	256:268	161:163	107:119	241:241	268:274	123:123	101:140	216:219	215:236	154:154	224:245	106:106	115:115	216:240	131:131	130:133	190:191	154:154	226:238	0	136:161	275:277	121:121	168:168	0	150:200	309	317	335	239	225	3	177	119	167	349	126	137	108	90	110	245	13
P03869103_Syria	36.72500	36.65500	4/15/1934	-	163:163	204:213	217:226	268:268	161:163	143:143	241:244	250:271	123:123	101:134	191:192	215:248	181:181	218:239	109:109	130:130	270:291	143:152	133:136	191:191	149:154	226:241	0	136:161	275:285	123:123	175:175	121:121	150:200	309	317	335	239	225	3	177	116	167	349	127	137	108	90	110	245	12
P03869104_Turkmenistan	38.91000	56.21000	5/25/1901	-	0	210:210	214:226	256:256	167:170	110:110	250:250	259:271	126:126	116:116	207:243	206:206	154:166	227:227	115:115	136:136	0	140:140	133:133	203:203	149:149	223:223	0	161:161	274:274	123:123	175:175	0	0	309	317	335	239	225	3	176	119	167	349	127	136	108	90	110	245	8
P04254353_Iran	36.45000	53.10000	00/00/1966	-	0	210:210	217:226	256:256	167:167	110:110	250:250	259:271	117:126	116:116	213:225	206:209	166:166	212:212	115:115	112:112	0	137:140	133:136	194:203	149:149	223:223	0	161:161	274:274	123:123	175:175	0	0	309	317	335	239	225	3	176	119	167	349	127	136	108	90	110	245	8