



HAL
open science

XEML Interactive designer

Benjamin Dartigues, Kenny Biliau, Cécile Cabasson, Yves Gibon, Joachim Kopka, A. Erban, Jan Hummel, Macha Nikolski, Dirk Walther

► **To cite this version:**

Benjamin Dartigues, Kenny Biliau, Cécile Cabasson, Yves Gibon, Joachim Kopka, et al.. XEML Interactive designer: Machine Readable Metadata describing Genotypes, Environmental Conditions and Sampling Strategy. Bordeaux Computational Biology and Bioinformatics, Nov 2014, Bordeaux, France. , 1 p., 2014. hal-01203290

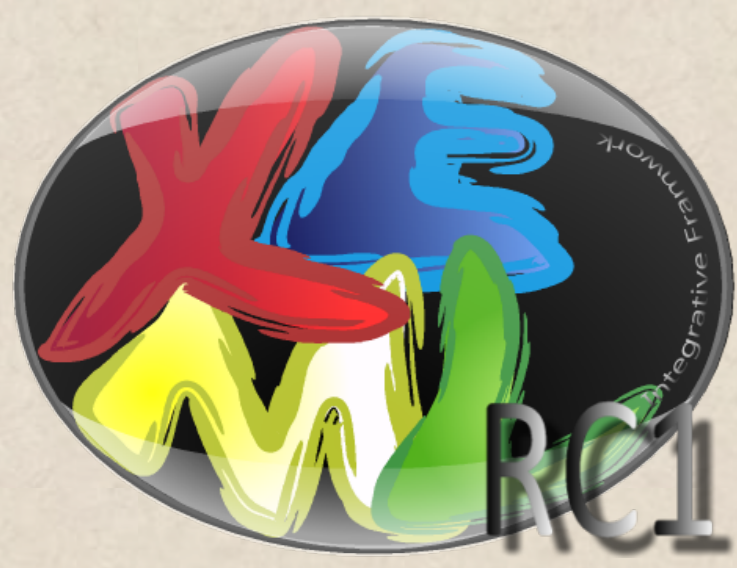
HAL Id: hal-01203290

<https://hal.science/hal-01203290>

Submitted on 22 Sep 2015

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.

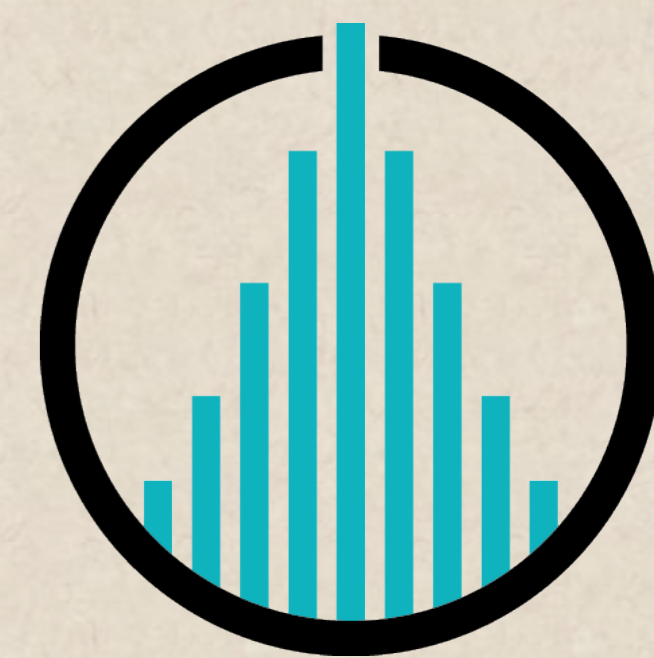


XEML interactive Designer

Benjamin Dartiques², Kenny Billiau¹, Cabasson C³, Gibon Y³, Erban A¹, Kopka J¹, Jan Hummel¹, Macha Nikolski² & Dirk Walther¹

¹ Max-Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, D-14476 Golm, Germany
² Bordeaux Bioinformatics Center, Université Bordeaux, rue Léo Saignat 146, 33076 Bordeaux, France
³ UMR 1332 Biologie du fruit et Pathologie, Centre INRA Bordeaux 71, avenue Edouard Bourlaux, CS20032 33882 Villenave d'Ornon Cedex, France

Contact: {billiau|walther}@mpimp-golm.mpg.de/ bdartiques@gmail.com



MACHINE READABLE METADATA DESCRIBING GENOTYPES, ENVIRONMENTAL CONDITIONS AND SAMPLING STRATEGY

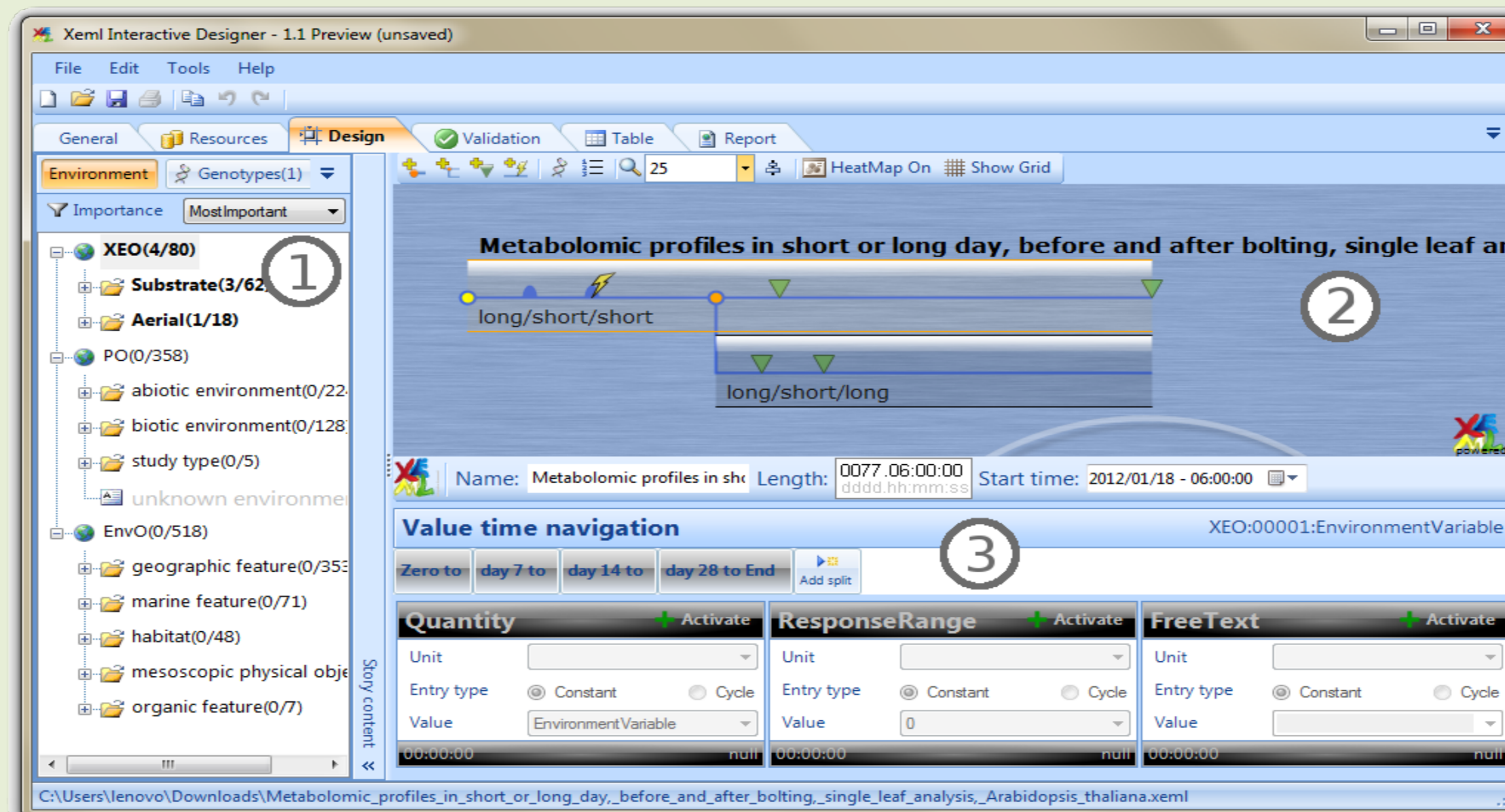
Data mining depends on the ability to access machine readable metadata that describe genotypes, environmental conditions, and sampling times and strategy. XEML interactive Designer provides an interactive graphical interface in which complex experiments can be designed, and conjointly generates machine-readable metadata files.

It uses a new eXtensible Mark-up Language (XML)-derived dialect termed XEML. The Designer includes a new ontology for environmental conditions, called XEML Environment Ontology. However, to provide versatility, it is designed to be generic and accepts other commonly used ontology formats (OBO, OWL).

1 Environmental parameters are determined with the help of preloaded ontologies. Environmental, developmental and structural terms can be found to describe your experiment.

2 Define the general setup of your experiment, plot treatments into different stories, choose observation points where you want to take measures and select environmental parameters from predefined ontologies to describe and document your experiment.

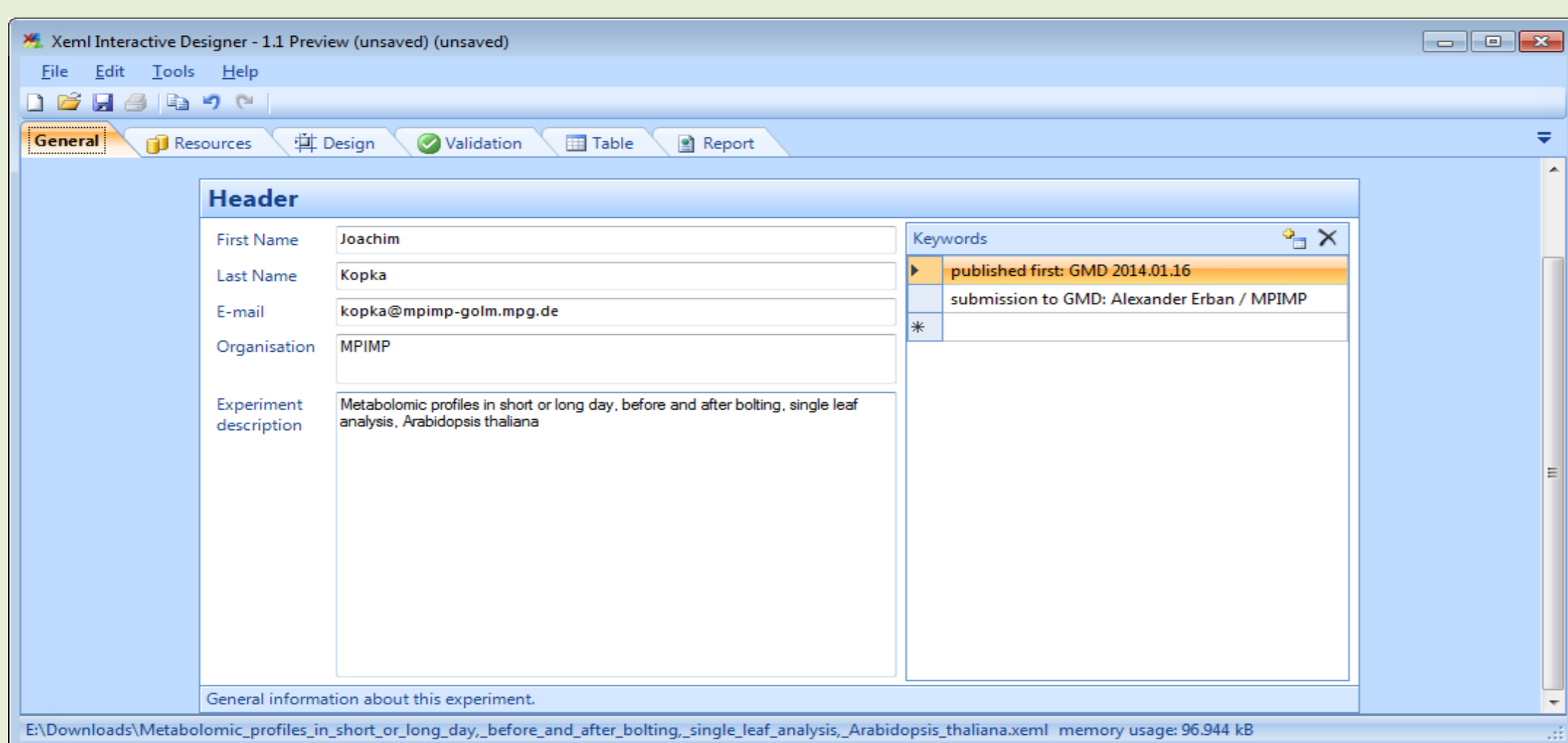
3 Enter values for each environmental parameter you want to control on any time point within your storyboard. Values for parameters can be defined under different contexts. This is possible to enter intensities, and to indicate the level of optimality or stressfulness.



XEML interactive Designer is compatible with the open source ISA metadata tracking tools to help to manage an increasingly diverse set of life science, environmental and biomedical experiments that employing one or a combination of technologies.

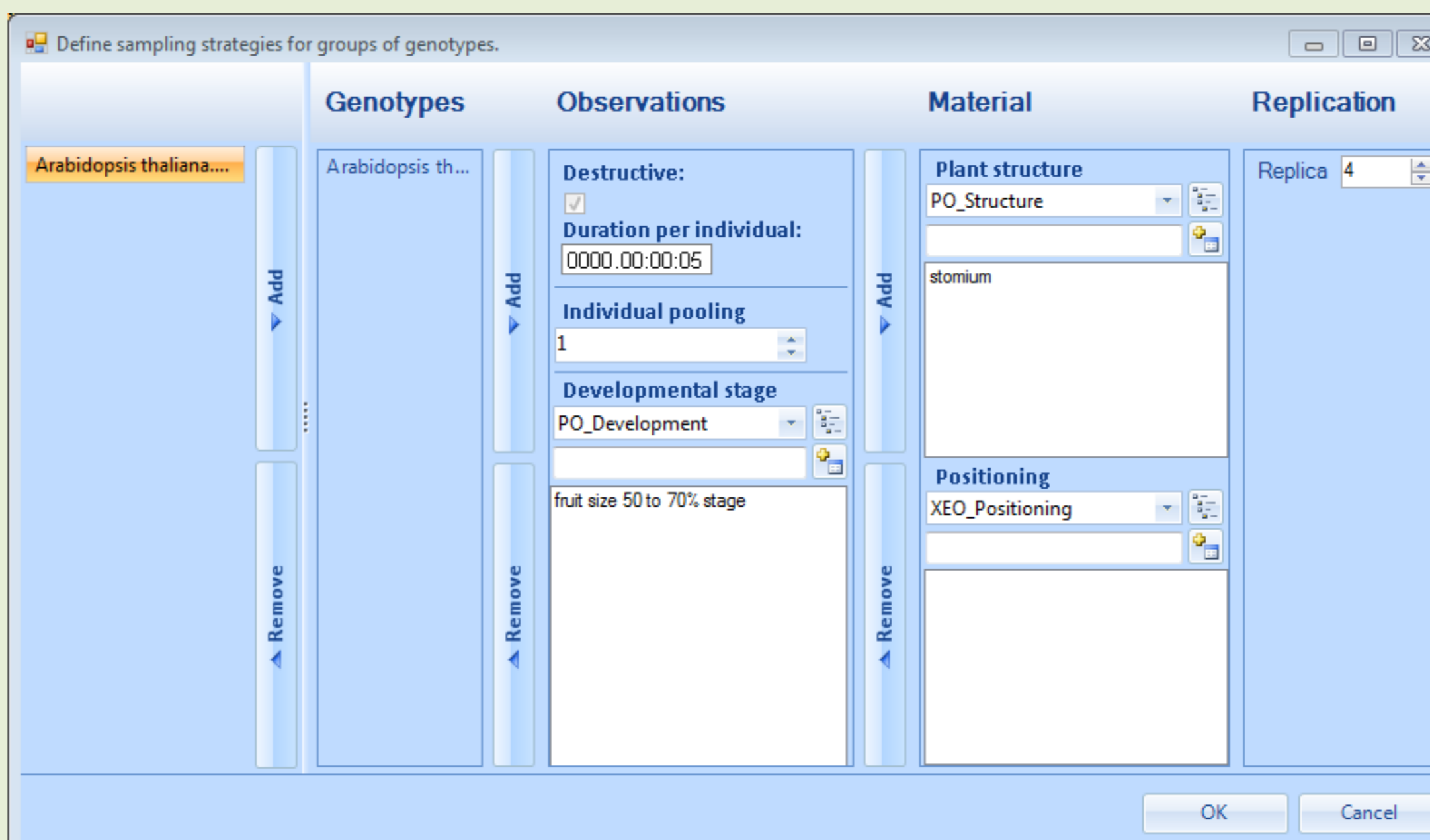
4 Define the germplasm of your individual pools by entering species, accession, mutant information and transgenic modification. Since a germplasm ontology has not been implemented yet, you can make detailed annotation to any entry.

Define the germplasm of your individual pools by entering species, accession, mutant information and transgenic modification. Since a germplasm ontology has not been implemented yet, you can make detailed annotation to any entry.



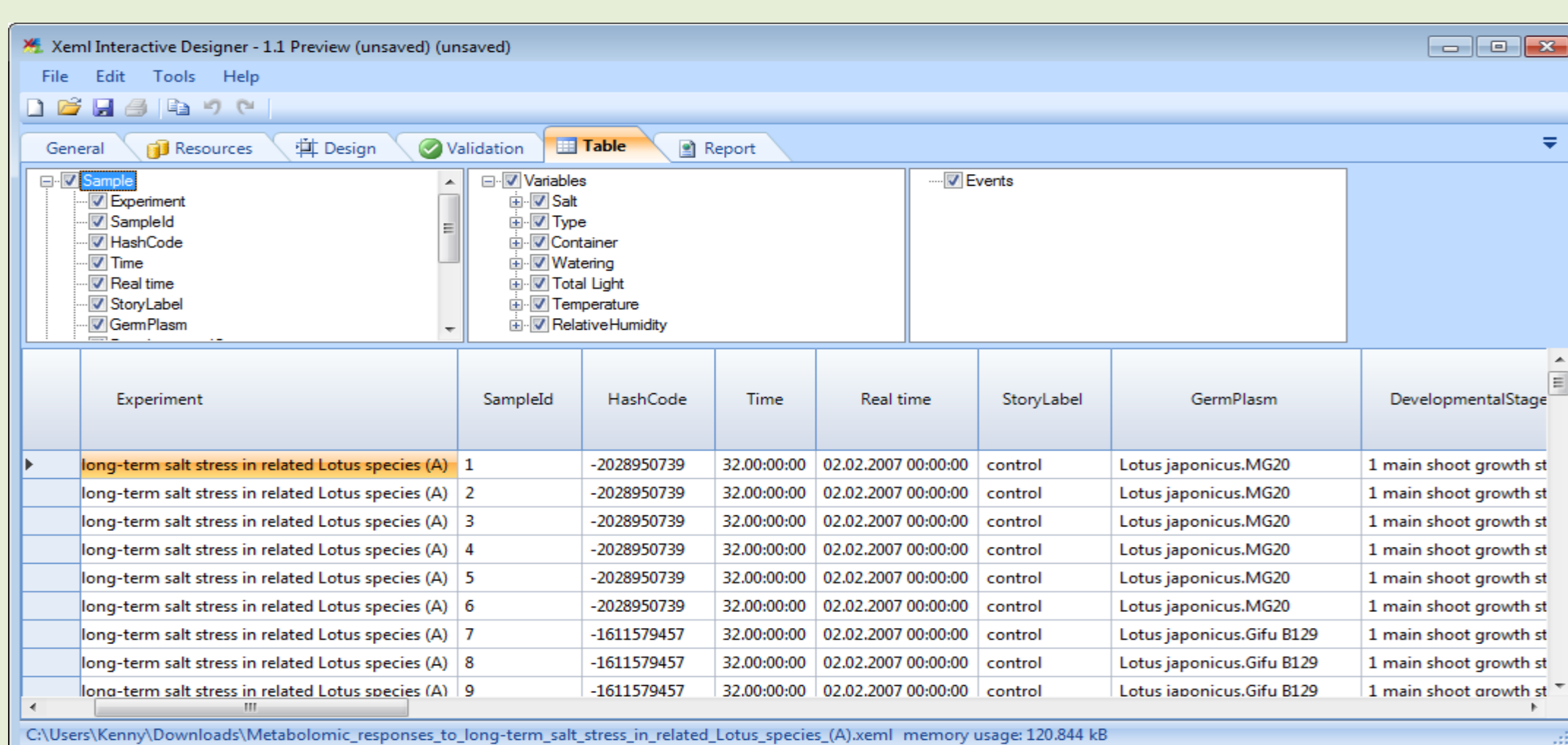
5

The sampling strategy for multiple genotypes in terms of the number of individuals, plant part and replicates, can be defined for each of the genotypes under investigation.



6

As a final step you can produce a table as an overview of all entered design information. This table can be copied to the clipboard and used for further processing steps in programs like Excel.



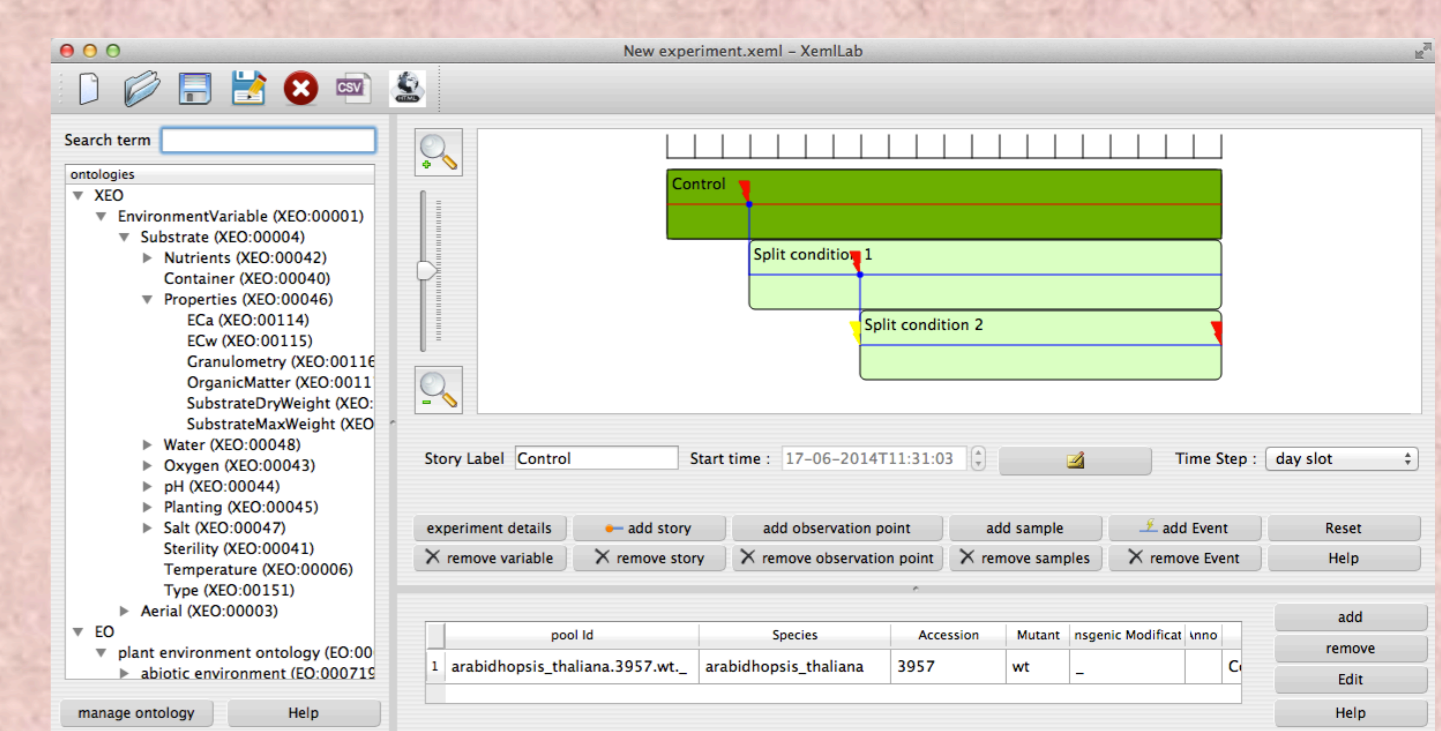
FEATURES



- Formulate experimental setup
- Include environmental parameters
- Individual germplasm information
- Import/Export XEML
- Export as ISATab
- Integration of and possibility to import ontologies
- Sample data mapping to heterogeneous databases and general annotations
- Define a "Response Range"

WHAT'S NEXT?

Now a new version of XEML Designer is available for all platforms (Windows, Mac and Linux). This new version offers all functionalities previously described plus new ones facilitating description of experiments.

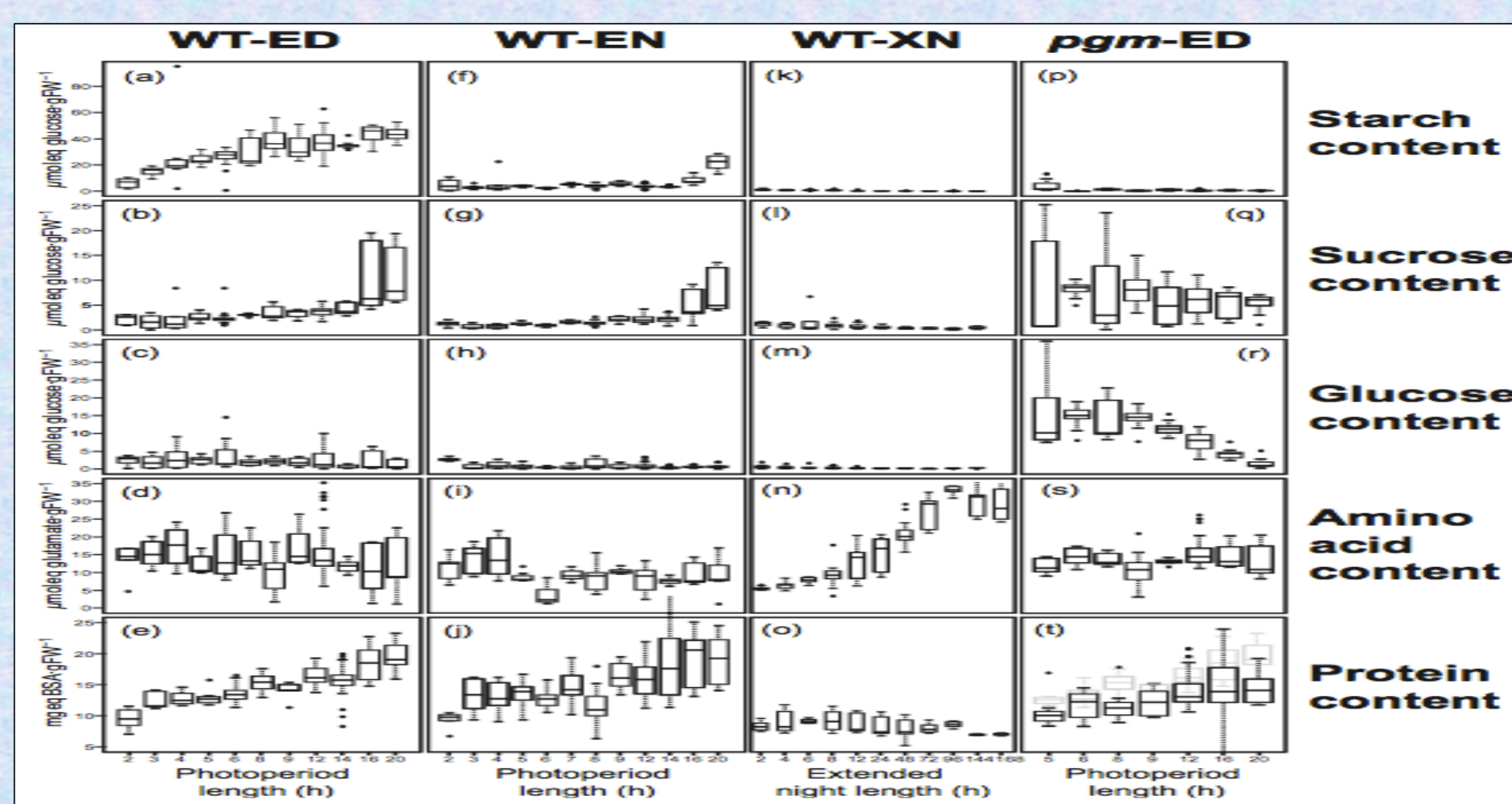


- Import/export of data from tabular files (csv, tsv)
 - Plot the environmental data for intuitive display of correlations
 - More coming soon
- This way, users can add their experimental or environmental data and easily visualize them

EXPERIMENTAL DATA COUPLED WITH ENVIRONMENTAL METADATA

Box plots representing variations in starch (a, f, k, p), sucrose (b, g, l, q), glucose (c, h, m, r), total amino acids (d, i, n, s) and protein content (e, j, o, t) in rosette leaves of wild-type *Arabidopsis thaliana* grown under various day and regimes, harvested at the end of the photoperiod (a–e), at the end of the night (f–j), or following transfer into continuous darkness (k–o),

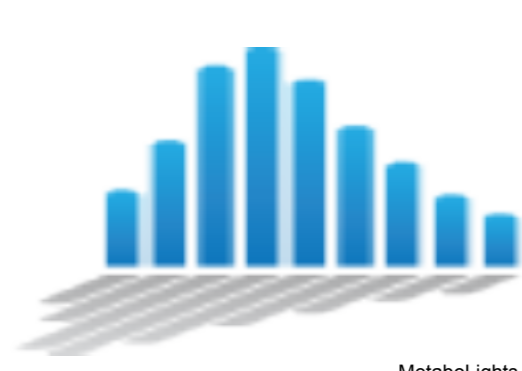
and in rosette leaves of starchless *pgm* mutants grown under various day and night regimes. The light grey data shown in panel T is the corresponding data for wild-type Col-0 (extracted from panel e) to facilitate comparison with the values in *pgm*. The data were collected from a total of 23 separate experiments performed between 2002 and 2007.



XEML interactive Designer generates metadata files that can be crossed with experimental files in order to perform Data mining analysis. The figure above summarizes an experiment performed to extract information about the response of metabolism to changes in the carbon status.



This project is funded through European Commission COSMOS Grant EC312941



- MetaboLights: <http://www.ebi.ac.uk/metabolights/>
- ISATools: <http://isa-tools.org/>
- Codeplex (C#): <http://XEML.codeplex.com/>
- GitHub (C++): <https://github.com/cbib/XEML-Lab>
- Cosmos: <http://www.cosmos-fp7.eu/>



cgfb
BIOINFORMATIQUE

Max-Planck-Institut für Molekulare Pflanzenphysiologie

