Genetic and population analysis

BioMercator V3: an upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms

Olivier Sosnowski, Alain Charcosset and Johann Joets*

INRA, UMR 0320/UMR 8120 Génétique Végétale, F-91190 Gif-sur-Yvette, France

Associate Editor: Jeffrey Barrett

ABSTRACT

Summary: Compilation of genetic maps combined to quantitative trait loci (QTL) meta-analysis has proven to be a powerful approach contributing to the identification of candidate genes underlying quantitative traits. BioMercator was the first software offering a complete set of algorithms and visualization tool covering all steps required to perform QTL meta-analysis. Despite several limitations, the software is still widely used. We developed a new version proposing additional up to date methods and improving graphical representation and exploration of large datasets.

Availability and implementation: BioMercator V3 is implemented in JAVA and freely available (http://moulon.inra.fr/biomercator)

Contact: joets@moulon.inra.fr

1 INTRODUCTION

Integration of multiple independent QTL mapping experiments is of major interest to unravel genetic factors underlying complex traits. As a first step, a consensus map is built from independent QTL maps and, if available, the organism reference map. QTLs are then projected onto the consensus map and subjected to meta-analysis (Goffinet and Gerber, 2000). As a result, meta-QTLs and mapped genes lie on the same consensus map, greatly simplifying the inventory of candidate genes. In addition, the confidence interval (CI) of a meta-QTL is often shorter than CI of corresponding QTLs, decreasing the number of candidate genes to consider or the extent of fine mapping to conduct. In a survey of flowering-time traits in maize, 62 meta-QTLs were identified (Chardon *et al.*, 2004) including one corresponding to VGT1, which was subsequently cloned in the predicted region by fine mapping (Salvi *et al.*, 2007). As map compilation and QTL meta-analysis are not based on genotyping raw data or trait measure, they can be easily achieved from maps available from the literature or databases.

The BioMercator software (Arcade *et al.*, 2004) offers a complete set of analysis and visualization tools dedicated to these approaches and is widely used (Blackman *et al.*, 2011; Chang *et al.*, 2011; Quraishi *et al.*, 2011) but suffers several limitations. We propose here a major upgrade of the software to overcome algorithm and graphical user interface limitations; map compilation can now be achieved in a single step whatever the number of maps, there is no limitation to the number of meta-QTL per chromosome and the code

was rewritten to support high-density maps with no limitation of the number of loci.

2 COMPILATION OF GENETIC MAPS

The previous versions of BioMercator proposed an iterative map compilation method, which required at least $n-1$ steps to integrate n maps. The process could be fastidious for a large number of maps and in addition did not ensure optimality of consensus map. A one-step procedure to compile n maps based on weighted least squares and described in Veyrieras *et al.* (2007) is now available in BioMercator. To assist user in examining the extent of shared loci between maps, we have implemented a graphical comparative map viewer, which highlights inverted loci.

3 QTL META-ANALYSIS

The meta-analysis method implemented in BioMercator 2.1 suffered several limitations. Only models from 1 to 4 meta-QTLs per chromosome were considered, possibly forcing user to iterate meta-analysis process several times to fully cover a chromosome. In addition, classification of QTL into meta-QTL was discrete despite in some cases it would have been more realistic to take into account the uncertainty in assignation of a given QTL to two or more meta-QTL. Methods and code (Veyrieras *et al.*, 2007) addressing these issues have been included into BioMercator V3; among the new functionalities are full chromosome meta-analysis and probabilistic clusterization of QTLs. QTLs from related traits can be jointly subjected to a single meta-analysis by grouping corresponding traits into meta-traits. User may define meta-traits with any combination of traits looking relevant to him.

As an alternative or complement to meta-analysis, Chardon *et al.* (2004) have introduced the 'QTL overview' analysis, which summarizes QTL information by estimating the probability of identification of a QTL along the consensus map. BioMercator V3 automatically computes these probabilities, which are represented as a curve along the chromosomes.

4 DATA REPRESENTATION

The datasets to be handled by BioMercator, if not carefully represented, can quickly overload the display as experienced in the previous version. New compact graphical representations have been developed; (i) a chromosome cascading zoom replaces the simple scale change of the previous version; this zoom allows the user to enlarge a region of chromosome as deep as he needs, while keeping an overview of the whole chromosome map. (ii) When compiling several maps into a reference one, QTLs may stack in

---

*To whom correspondence should be addressed.
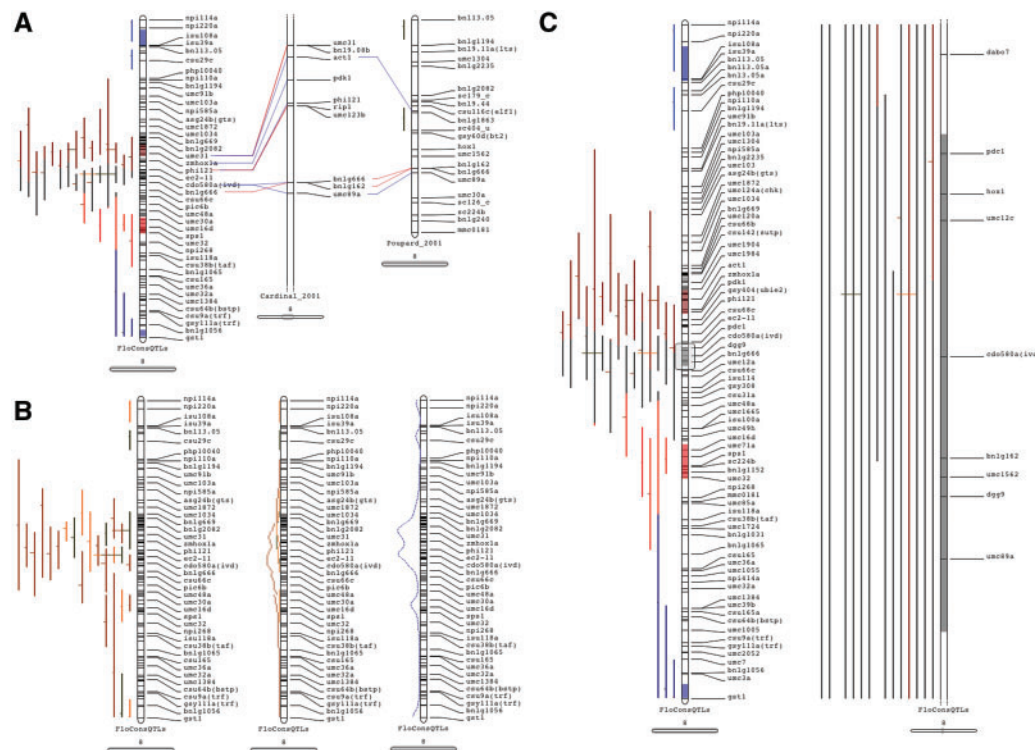
Fig. 1. Genetic maps graphical representation. (**A**) Dynamic maps comparison; shared loci are linked by blue line if collinear and red line otherwise. (**B**) Full representation of QTLs (left map), QTL density curves (one by trait) and QTL overview curve (all traits, right map). (**C**) Meta-analysis output; meta-QTL are drawn within chromosome and QTL-CI is colored according to the probability of its belonging to meta-QTLs; an enlarged region is displayed on the right along the full chromosome. The dataset used is maize maps collected by Chardon *et al.* (2004). Traits are silking date (brown), days to pollen shed (black) and plant height (orange)

some regions and occupy a really large space, impairing whole map representation (Fig. 1). To address such cases, a new QTL track summarizing QTL data along the chromosomes has been developed. Basically, this track is a curve depicting QTL density that may be weighted or not by R2. (iii) For whole map representation, each chromosome can be enlarged independently at the user convenience in order to focus on regions of interest. By default, for a lighter display, only a randomly selected set of loci is represented at whole chromosome scale. By zooming in progressively, more loci are displayed. However, an option forcing the display of all loci is available. Datasets as well as analysis output may be exported in several formats (tabulation-delimited text file, xml). Either whole genome map or single chromosome can be exported in JPEG, PNG or the useful SVG vectorial format allowing image editing in dedicated third-party software.

5 FUTURE DEVELOPMENTS

Genetic maps and genome sequence are now integrated for many species. It would be a great enhancement for end-user if the structural and functional annotations of regions underlying QTL and meta-QTL could be reached directly from BioMercator.

REFERENCES

Arcade,A. *et al.* (2004) Biomercator: integrating genetic maps and Qtl towards discovery of candidate genes. *Bioinformatics*, **20**, 2324–2326.

Blackman,B.K. *et al.* (2011) Contributions of flowering time genes to sunflower domestication and improvement. *Genetics*, **187**, 271–287.

Chang,W. *et al.* (2011) Qtl underlying resistance to two Hg types of *Heterodera glycines* found in soybean cultivar 'L-10'. *BMC Genomics*, **12**, 233.

Chardon,F. *et al.* (2004) Genetic architecture of flowering time in maize as inferred from quantitative trait loci meta-analysis and synteny conservation with the rice genome. *Genetics*, **168**, 2169–2185.

Goffinet,B. and Gerber,S. (2000) Quantitative trait loci: a meta-analysis. *Genetics*, **155**, 463–473.

Quraishi,U.M. *et al.* (2011) Cross-genome map based dissection of a nitrogen use efficiency ortho-metaQtl in bread wheat unravels concerted cereal genome evolution. *Plant J.*, **65**, 745–756.

Salvi,S. *et al.* (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc. Natl. Acad. Sci. USA.*, **104**, 11376–11381.

Veyrieras,J.B. *et al.* (2007) MetaQtl: a package of new computational methods for the meta-analysis of Qtl mapping experiments. *BMC Bioinformatics*, **8**, 49.