The Genetic Map Comparator: a user-friendly application to display and compare genetic maps

Yan Holtz[1], Jacques David[1], Vincent Ranwez[1*]

[1]*Montpellier SupAgro, UMR AGAP, 34060 Montpellier, France*

*To whom correspondence should be addressed.

Abstract

Motivation: Marker-assisted selection strongly relies on genetic maps to accelerate breeding programs. High-density maps are now available for numerous species. Dedicated tools are required to compare several high-density maps on the basis of their key characteristics, while pinpointing their differences and similarities.

Results: We developed the Genetic Map Comparator—a web-based application for easy comparison of different maps according to their key statistics and the relative positions of common markers.

Availability: The genetic map comparator is available online at: www.agap-sunshine.inra.fr/genmapcomp. The source code is available for free download on GitHub under the under the CeCILL general public license: https://github.com/holtzy/GenMap-Comparator.

Contact: Holtz@supagro.fr; Ranwez@supagro.fr

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Drawing and displaying genetic maps is a common task for anyone working with genetic markers. Multiple maps that share some markers have to be dealt with when several segregating populations are studied. These maps are compared in order to highlight their overall relative strengths and weaknesses (e.g. via marker distributions or map lengths), or local marker inconsistencies. Genetic maps are usually compared on the basis of some summary statistics (e.g. average gap size between successive markers or total map length) and dedicated graphical representations (e.g. marker positions along chromosomes). These comparative data have to be updated after each map recalculation according to the different purposes (e.g. focusing on a given chromosome or a subset of available maps), so an effective user-friendly tool is essential to facilitate the genetic map comparisons.

A few tools like MapChart (Voorrips 2002) are already available for graphical presentation of genetic maps. They are usually focused on charting quantitative trait loci (QTLs) along a single map, but also offer a few map comparison possibilities. However, their main target is not map comparisons and they are not tailored for high resolution genetic maps that are currently produced thanks to rapid advances in biomolecular tools. Cmap (Fang et al. 2003) and Cmap-3D (Duran et al. 2003) are designed for map comparisons but are presently unsuitable for dealing with high density maps. Scientists and breeders are now frequently faced with the challenge of having to compare several genetic maps, each bearing thousands of markers. Initially developed for low density microsatellite or DArT markers, etc., current tools produce overly cluttered charts that do not enable proper data comparisons.

The Genetic Map Comparator takes advantage of R shiny (RStudio Inc 2014) and Plotly (https://plot.ly) to jointly draw and compare multiple genetic maps. These interactive tools allow users to more efficiently explore the data. Among other features, the Genetic Map Comparator allows users to select a subset of chromosomes, zoom on loci, hover over interesting markers to get their names, and perform inter-chromosomal analyses to detect markers assigned to very different positions in different maps. Moreover, it also provides some key descriptive statistics concerning the selected data, which are updated based on user data selection and hence always synchronized with the charted data. This feature avoids the tedious (and error prone) task of juggling between a map visualization tool and statistical software to analyze the data.

The Genetic Map Comparator is thus a valuable tool to help scientists and breeders explore and compare dense genetic maps and produce informative graphic summaries of their work, which is crucial for efficient scientific communications.

2 Technology and installation

The Genetic Map Comparator is an R shiny (RStudio Inc 2014) application using Plotly. It can be installed locally for private/custom uses (https://github.com/holtzy/GenMap-Comparator); the installation procedure is quite straightforward on most operating systems as it depends only on R (R Core Team 2013) and a handful of R packages. The simplest way to use the Genetic Map Comparator is via our online version: www.agap-sunshine.inra.fr/genmapcomp which does not require installation. Two input formats, compatible with the leading software OneMap, MapMaker or Carthagene (Ahfock et al. 2014), are accepted for uploading user maps.

3 Key functionalities

The Genetic Map Comparator site includes several web pages (Fig. 1) that are accessible through tabbed browsing buttons. On the "**Home"** page, personal maps may be uploaded or one of the provided datasets consisting of several maps and thousands of markers may be selected. The "**Summary Statistics"** page offers general information such as number of markers, total map length, number of unique positions and inter-marker distances. Statistics are provided for the whole map and for each individual chromosome. The "**Compare position"** page displays a graph of markers along the chosen chromosome on the selected maps. Selected maps are displayed side by side and common markers are linked by purple lines, highlighting colinearity and suspicious markers. The map order can be modified by the user. When focusing on only two maps, an alternative display is proposed in the "**Inter-chromosomal analyses"** page. This view displays a scatter plot with genetic positions of two selected maps only. This enables users to simultaneously compare the number and position of markers of all chromosomes of the two selected maps and hence to detect chromosomal assignment discrepancies that may be caused by chromosomal recombination or by segmental or chromosomal duplications. The "**Raw data"** page displays a table with one row per uploaded marker. This table can be sorted according to marker names or positions, it can be filtered to focus on a subset of maps or chromosomes and it provides searching facilities based on marker names. The last "**Help"** page offers a step-by-step guide for optimal use of the Genetic Map Comparator.

4 Case study

The Genetic Map Comparator project was initiated to support a breeding project on durum wheat geared towards identifying QTLs for virus resistance. Two recombinant inbred line segregating populations sharing a common parent were genotyped using specific allelic capture (Holtz, Y., et al. 2016). The two individual maps and their consensus map were built, with 3729, 6886 and 8568 SNP markers, respectively. A physical map with the putative position of markers was also available. The Genetic Map Comparator was found to be a really efficient tool for quickly and easily exploring these map data. It enabled a seamless shift from the genetic map building to the visualization step, thus enhancing the genetic map building efficiency. Moreover, it streamlines collaboration by providing interactive access to (up to date) data to all project members without requiring any prerequisite (bio-)informatics skills. Several of the figures needed for publication were synthesized by exporting charts from the Genetic Map Comparator.
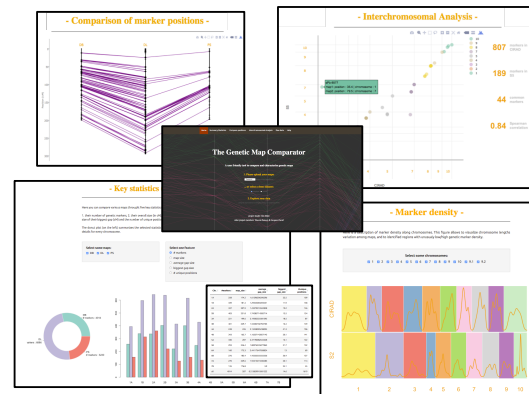

Fig. 1: Screenshots from the Genetic Map Comparator. A set of genetic maps may be uploaded from the Home page (center) and then compared using: the "Compare position" (top left), the "Interchromosomal analyses" (top right) and the "Summary statistics" (bottom) pages.

Funding

Acknowledgment

References

Voorrips, R.E (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered*, **93**(1):77-78.

Fang, Z., et al (2003). cMap: the comparative genetic map viewer. *Bioinformatics* 19(3):416-417.

Duran, C., et al (2010). CMap3D: a 3D visualization tool for comparative genetic maps. *Bioinformatics* **26**(2):273-274.

RStudio Inc (2014). Shiny: Web Application Framework for R. (http://cran.r-project.org/package=shiny)

R Core Team. R (2013). A Language and Environment for Statistical Computing. (http://www.R-project.org/).

Ahfock, D., et al. (2014). Characterizing uncertainty in high-density maps from multiparental populations. *Genetics* **198**(1):117-128.

Holtz, Y., et al (2016). Genotyping by Sequencing Using Specific Allelic Capture to Build a High-Density Genetic Map of Durum Wheat. *PLoS One* **11**(5):e0154609.