*Bioinformatics*, YYYY, 0–0

doi: 10.1093/bioinformatics/xxxxx

Advance Access Publication Date: DD Month YYYY

Application Note

|  |
| --- |
| Subject Section  The Genetic Map Comparator: a user friendly application to visualize and compare genetic maps  Yan Holtz1, Jacques David1, Vincent Ranwez1\*  1*1Montpellier SupAgro, UMR AGAP, 34060 Montpellier, France*  \*To whom correspondence should be addressed.  Associate Editor: XXXXXXX  Received on XXXXX; revised on XXXXX; accepted on XXXXX  Abstract  **Motivation:** Genomic selection strongly relies on genetic map to accelerate breeding programs. Thanks to high throughput sequencing high density maps are now available for numerous species. Comparing several maps require dedicated tools able to summarize the key characteristics of high density maps and to pinpoint their differences and similarities.  **Results:** We developed the Genetic Map Comparator, a web based application allowing to easily compare different maps by depicting their key statistics and the relative positions of their common markers.  **Availability:** The genetic map comparator is available online: [www.agap-sunshine.fr/genmapcomp](http://www.agap-sunshine.fr/genmapcomp). Source code is freely available on GitHub under the under the CeCILL general public license: <https://github.com/holtzy/GenMap-Comparator>.  **Contact:** [Holtz@supagro.fr](mailto:Holtz@supagro.fr); [Ranwez@supagro.fr](mailto:Ranwez@supagro.fr)  **Supplementary information**: Supplementary data are available at *Bioinformatics* online. |

# Introduction

Drawing and visualizing genetic maps is a common task for whoever works with genetic markers. When several populations and crossing are studied one have to deal with multiple maps that share some markers. The comparison of those maps allows to assert their relative strengths and weaknesses or to pinpoint recombination events among parent varieties of the studied populations. Genetic maps are usually compared through values of some key statistics (e.g. average gap size between successive markers or total map length) and dedicated graphical representations (e.g. marker positions along chromosome). As those comparative data have to be update after each map recalculation and are considered using different point of views (e.g. focusing on a given chromosome or a subset of available maps) a user-friendly and effective tool is strongly needed to ease those genetic map comparisons.

A few tools like MapChart [1] are already available for genetic maps visualization. They are most often focused on the visualization of Quantitative Trait Loci (QTLs) along a single map, but also propose a few map comparison possibilities. However, their main target is not the map comparisons and they are not well adapted to the high resolution genetics maps that are nowadays produced thanks to the rapid advances in bio-molecular tools. Cmap [2] and Cmap-3D [3] are dedicated to map comparisons but are also unfit to nowadays high density maps. Scientists and breeders are now frequently facing the challenge to compare several genetic maps each carrying several hundreds of marker. With current tools, initially develop for low density markers such as microsatellite or darts, it leads to messed-up visualizations that does not permit to understand data properly.

The Genetic Map Comparator uses emerging interactive data visualization tools such as R shiny [4] and Plotly [XX] to represent and compare genetic maps. Using reactive tools allows the user to more efficiently explore the data. The Genetic Map Comparator allows, among other things, selecting a subset of chromosomes, zooming on figures, hovering on interesting markers to get their names, performing inter-chromosomal analyses to detect markers assigned to different chromosomes in different maps. Moreover, the Genetic Map Comparator also provides some key statistics concerning the selected data that are updated based on user data selection and hence always synchronized with the visualized data. This avoid the tedious (and error prone) task of juggling between a map visualization tool and a statistical software to analyze the data.

The Genetic Map Comparator should thus significantly help scientists and breeders to explore and compare dense genetic maps and to produce informative graphic summaries of their works, which are crucial for efficient scientific communications.

# Technology and installation

The Genetic Map Comparator is an R shiny [4] application using Plotly [XX]. It can be installed locally for private/custom utilization (lien GitHubXXX); the installation procedure is quite straightforward on most operating systems as it depends only on R (ref R) and a handful of R packages. The simplest way to use the Genetic Map Comparator is to rely on our online version: [www.agap-sunshine.fr/genmapcomp](http://www.agap-sunshine.fr/genmapcomp) for which no installation whatsoever is needed. Two input formats are accepted for uploading user dataset that correspond to the output format of today genetic map building software such as OneMap, MapMaker or Carthagene.. ~~.~~

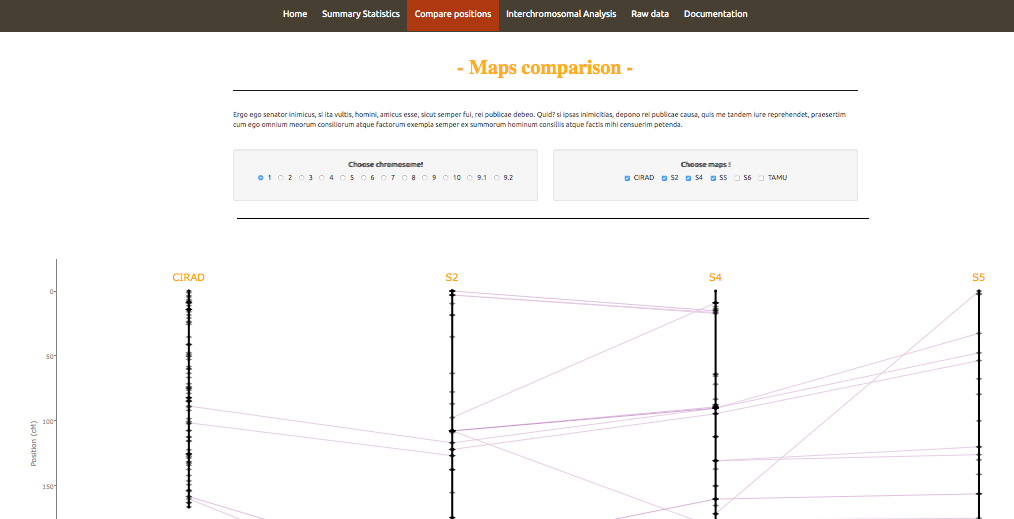
# Key functionalities

The Genetic Map Comparator site is made of several web pages that are accessible through buttons mimicking tabbed browsing. The **Home page** allows to upload a personal map dataset or to select one of the three provided example datasets taken from publication related to durum wheat [5] and sorghum [6]. Those example datasets contain several maps made of hundreds of markers and confirm the responsiveness of the Gene Map Comparator on large dataset and illustrate the relevance of its functionalities. The “**Summary Statistics” page** provides 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. Spreadsheet of key statistics are available for download in .csv format for further analysis or publication purpose. The **Compare position** page displays a graphic representation of markers along the chosen chromosome of the selected maps (Fig 1). More precisely, for each selected map, a vertical black line represents the chosen chromosome. Dots along each lines represents genetic marker positions along this chromosome/map and information concerning each marker is available by hovering the dot. Purple lines link every common marker between two adjacent maps (map order can be modified by the user). When focusing on only two maps an alternative view is proposed in the “**inter-chromosomal analyses”** page. This view displays a scatter plot with genetic positions of two selected maps only. This enables 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. The **raw data** page displays a table with one row per upload 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 provides step by step guiding for optimal use of the Genetic Map Comparator .

# Case study

The Genetic Map Comparator project has been initiated to support a breeding project on durum wheat aiming at identifying QTLs for virus resistance (ref XXX). Using three parent genitors, two populations were breed and genotyped. Two individual maps and a consensus map were built, with respectively xx, xx and xx SNP markers). A physical map with the putative position of markers was also available. To investigate the consistency of markers assignments and positions, an important number of comparisons were needed hence the need for this tool.

The Genetic Map Comparator revealed to be a really efficient to explore map data quickly and easily. It permitted to pass fluently from the genetic map building to the visualization step, thus improving the genetic map building efficiency. Moreover, it smoothens collaboration by providing interactive access to (up to date) data to all project members without the need of any (bio-)informatics skill pre-requisite. Several of the published figures were synthetized exporting charts from the tool (e.g. Fig XXX and Fig YYY in XXX).



**Fig. 1**: Parallel comparison of 4 genetic maps (via the **Compare position** sheet of the application). Each vertical black line represents the selected chromosome in one of the 4 populations, black dots along those lines indicate marker positions and purple lines link markers shared by 2 adjacent chromosomes. Numbers on the left side specifies positions in cM. Names of markers are available on hover.

Funding

This work has been supported by …

*Conflict of Interest:* none declared.

References

1. Voorrips RE. MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered. 2002;93: 77–78. doi:10.1093/jhered/93.1.77

2. Fang Z, Polacco M, Chen S, Schroeder S, Hancock D, Sanchez H, et al. cMap: The comparative genetic map viewer. Bioinformatics. 2003;19: 416–417. doi:10.1093/bioinformatics/btg012

3. Duran C, Boskovic Z, Imelfort M, Batley J, Hamilton NA, Edwards D. CMap3D: A 3D visualization tool for comparative genetic maps. Bioinformatics. 2010;26: 273–274. doi:10.1093/bioinformatics/btp646

4. Chang W, Cheng J, Allaire JJ, Xie Y, McPherson J. shiny: Web Application Framework for R [Internet]. 2015. Available: http://cran.r-project.org/package=shiny

5. Maccaferri M, Ricci A, Salvi S, Milner SG, Noli E, Martelli PL, et al. A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. Plant Biotechnol J. 2014; 1–16. doi:10.1111/pbi.12288

6. Mace ES, Rami J-F, Bouchet S, Klein PE, Klein RR, Kilian A, et al. A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers. BMC Plant Biol. 2009;9: 1–14. doi:10.1186/1471-2229-9-13