

# VIEWpoly: a visualization tool to integrate and explore results of polyploid genetic analysis

Cristiane Hayumi Taniguti<sup>1</sup>, Gabriel de Siqueira Gesteira<sup>2</sup>, Jeekin Lau<sup>1</sup>, Guilherme da Silva Pereira<sup>3</sup>, Zhao-Bang Zeng<sup>2</sup>, David Byrne<sup>1</sup>, Oscar Riera-Lizarazu<sup>1</sup>, and Marcelo Mollinari<sup>2</sup>

<sup>1</sup> Department of Horticultural Sciences, Texas A&M University, College Station, TX, USA <sup>2</sup> Bioinformatics Research Center, Department of Horticultural Sciences, North Carolina State University, Raleigh, NC, USA <sup>3</sup> Department of Agronomy, Federal University of Viçosa, Brazil

DOI: [10.21105/joss.04242](https://doi.org/10.21105/joss.04242)

## Software

- [Review](#)
- [Repository](#)
- [Archive](#)

Editor: [Mikkel Meyer Andersen](#)

## Reviewers:

- [@cpalmer718](#)
- [@jkanche](#)
- [@raivivek](#)

Submitted: 21 February 2022

Published: 09 June 2022

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

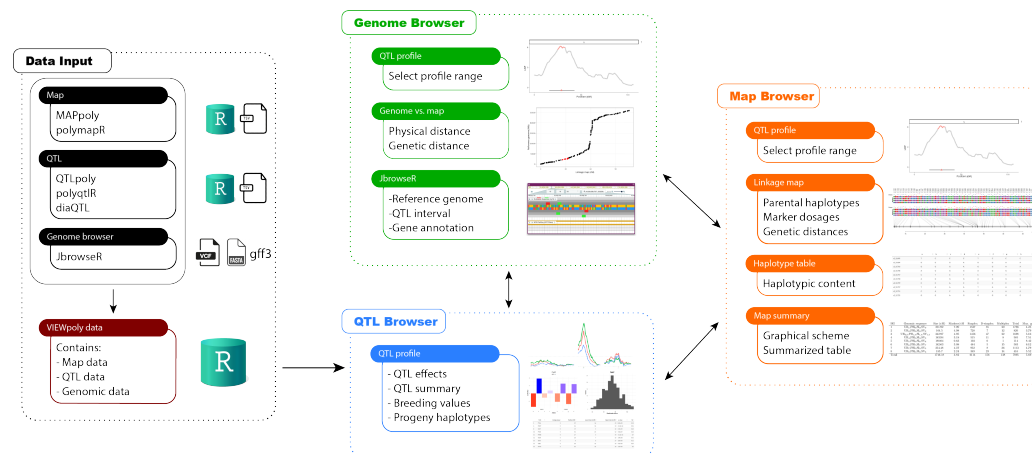
## Summary

Autopolyploid species comprise a significant part of the agronomic market. Some examples are potato, sweet potato, roses, and blueberries. They all present more than two copies of their entire genome, which increase the complexity of genetic models. Advancements in computational tools for linkage and quantitative trait loci (QTL) analysis in autopolyploids have allowed the in-depth exploration and understanding of their genetics. Despite providing advanced methods, the results of these tools may be challenging to interpret and connect to other sources of genomic information and consequently apply them in practical breeding applications. VIEWpoly is an R package for visualizing, exploring, and integrating results from different polyploid computational tools using an interactive graphical user interface. The package allows users to explore linkage and QTL analysis results and integrates these with genomic information in a genome browser, facilitating the search for candidate genes. In addition, it provides features to download comprehensive tables and graphics for further analysis and presentation. VIEWpoly is freely available as an R package at <https://CRAN.R-project.org/package=viewpoly>.

## Statement of need

In recent years, genetic analysis in autopolyploid crops has been gaining attention due to the flourishing of high-throughput genotyping technologies, which deliver massive amounts of DNA sequences. This technological advance prompted the development of many computational tools to perform an in-depth analysis of these complex genomes. Programs such as TetraPoloidSNPmap ([Hackett et al., 2017](#)), polymapR ([Bourke et al., 2018](#)), and MAPpoly ([Mollinari et al., 2020](#); [Mollinari & Garcia, 2019](#)) use dosage-based markers to build genetic maps and obtain the linkage phases in full-sib mapping populations. TetraOrigin ([Zheng et al., 2016](#)), PolyOrigin ([Zheng et al., 2021](#)), MAPpoly ([Mollinari & Garcia, 2019](#)), and polyqtLR ([Bourke et al., 2021](#)) provide the probabilistic haplotype inheritance profiles of individuals in mapping populations in terms of their respective parental haplotypes. These probabilities are used to estimate the position and genetic effects of quantitative trait loci (QTL). Using this information, researchers can benefit from the inferred genome-wide genotype and phenotype relationship to make informed breeding decisions. Software programs for performing such analysis in autopolyploids include GWASpoly ([Rosyara et al., 2016](#)), QTLpoly ([Silva Pereira et al., 2020](#)), polyqtLR ([Bourke et al., 2021](#)), and diaQTL ([Amadeu et al., 2021](#)). Despite the increasing availability of computational and analytical tools focused on autopolyploid species, the need for programming skills to comprehensively organize, integrate, display results of these programs still imposes an obstacle to exploring and using this information in practical situations. We present

VIEWpoly, a user-friendly R package that allows easy integration, visualization, and exploration of upstream genetic and genomic analysis focused on polyploid organisms. VIEWpoly is a self-contained interactive R package built as Shiny modules using *golem* (Fay et al., 2022) framework. It includes documentation, data examples, and comprehensive video tutorials to guide users in easy-to-follow steps to extract information from polyploid computational tools in a seamless manner. It can be deployed locally or on servers.



**Figure 1:** VIEWpoly conceptual flowchart highlighting the inputs and the three modules available: VIEWqtl, VIEWgenome, and VIEWmap.

## Software implementation

VIEWpoly comprises three main modules arranged in an intuitive interface that guides the user through the analysis (Figure 1). In its current version (V 0.2.0), the package supports polymapR (Bourke et al., 2018) and MAPpoly (Mollinari et al., 2020; Mollinari & Garcia, 2019) results to display genetic maps. The QTL analysis results can be imported from QTLpoly (Silva Pereira et al., 2020), polyqtlR (Bourke et al., 2021), and diaQTL (Amadeu et al., 2021) packages. Maps and QTL input files can be provided in the RData format or standard text formats (CSV or TSV). VIEWpoly is also compatible with genome-related files supported by JBrowseR, including FASTA, BED, GFF3, BAM, CRAM, VCF, Wiggle, and bigWig files (Hershberg et al., 2021). Once the data is uploaded, the user can export the information in a single compressed file to expedite the upload process in further analysis.

### VIEWqtl: QTL Browser

The VIEWqtl module is designed to explore information from the QTL analysis. The first graphic contains the QTL profile, where users can display single or multiple linkage groups and phenotypes. The QTL peak positions are represented by triangles at the bottom of the graphic and can be interactively selected to explore the genetic effects of parental haplotypes. It is possible to compare pleiotropic and linked QTL effects simultaneously and select individuals carrying alleles with specific effects. In this module, users can also find a table with a summary of selected QTL characteristics, and progeny breeding values will be displayed if results from QTLpoly are available. The information shown in graphics and tables may change according to the programs used in upstream analysis. For example, when QTLpoly is used, the QTL profile will show the negative logarithm of the P values ( $LOP = -\log_{10} P$ ) associated with the score statistics used to test for QTL significance. When polyqtlR is used, the logarithm of the odds (LOD) ratio is shown instead. If diaQTL is used, the significance curve can be evaluated using the Deviance Information Criterion (DIC). Other results derived from QTL analysis, such as heritabilities, allele effects, breeding values, and confidence intervals, are

available and shown according to upstream method used. All information is displayed in tables that can be downloaded in CSV, EXCEL, and PDF formats, and interactive graphics in static versions can be exported in PNG, TIFF, JPEG, and PDF formats.

### **VIEWgenome: Genome Browser**

This module lets users explore the relationship between QTL regions and a genome sequence with its features such as gene annotations, transcripts, and domains. After selecting a range in the genetic map, the JBrowseR interface ([Hershberg et al., 2021](#)) displays all tracks identified in the uploaded files. The display will be set to meet the corresponding genomic region according to the selected map region. The region can be queried and explored while all related information is automatically updated. Users can select the genomic sequence content and all annotated features for any given genomic range, including QTL peaks and their supporting intervals. A scatterplot of the genetic versus the physical positions highlights the relationship between the genetic map, reference genome, QTL positions and intervals, and recombination rate along the chromosomes.

### **VIEWmap: Map Browser**

In the VIEWmap module, users can interactively select a genomic range based on a QTL profile and explore specific regions of the map. Given a selected region, it is possible to interactively explore the estimated haplotypes and inspect their marker positions, allele dosages, genetic distances, and sequence base contents for parental homologs. Users can optionally generate a table with the haplotypes and download them in CSV, EXCEL, and PDF formats. The module highlights all linkage groups in a graphical representation with their respective marker positions and estimated recombination fractions, allowing for a direct comparison between homology groups. The module also shows a summarized table with information about the genetic maps, including their estimated sizes, densities, and genetic and genomic composition. Both figures and tables can be downloaded by users in standard static formats, such as CSV, EXCEL, PNG, TIFF, JPEG, and PDF.

## **Conclusions and perspectives**

VIEWpoly facilitates the integration, exploitation, and use of phenotypic, genetic, and genomic information provided by upstream analytical tools, leveraging the adoption of genomic-based tools for polyploid species. Its easy-to-use interface assists the decision-making process in practical applications, such as marker-assisted selection in breeding programs. Due to its modular nature, the package can be expanded and updated to incorporate further developments in diploid and polyploid genetic software. We expect to provide continuous integration and exploitation of novel features and tools, such as the analysis of multi-parental and multi-generation populations.

## **Acknowledgments**

We acknowledge Tessa Hochhaus for voicing-over the tutorial video.

## **Funding**

This work was supported by the USDA, National Institute of Food and Agriculture (NIFA), Specialty Crop Research Initiative (SCRI) project “Tools for Genomics-Assisted Breeding in Polyploids: Development of a Community Resource” (Award No. 2020-51181-32156), and by the Bill & Melinda Gates Foundation under the Genetic Advances and Innovative Seed Systems for Sweetpotato project (SweetGAINS) (grant number OPP1213329).

## References

- Amadeu, R. R., Muñoz, P. R., Zheng, C., & Endelman, J. B. (2021). QTL mapping in outbred tetraploid (and diploid) diallel populations. *Genetics*, 219(3). <https://doi.org/10.1093/genetics/iyab124>
- Bourke, P. M., Geest, G. van, Voorrips, R. E., Jansen, J., Kranenburg, T., Shahin, A., Visser, R. G. F., Arens, P., Smulders, M. J. M., & Maliapaard, C. (2018). polypmapR—linkage analysis and genetic map construction from F1 populations of outcrossing polyploids. *Bioinformatics*, 34(20), 3496–3502. <https://doi.org/10.1093/bioinformatics/bty371>
- Bourke, P. M., Voorrips, R. E., Hackett, C. A., Geest, G. van, Willemsen, J. H., Arens, P., Smulders, M. J. M., Visser, R. G. F., & Maliapaard, C. (2021). Detecting quantitative trait loci and exploring chromosomal pairing in autopolyploids using polyqtlR. *Bioinformatics*, 37(21), 3822–3829. <https://doi.org/10.1093/bioinformatics/btab574>
- Fay, C., Guyader, V., Rochette, S., & Girard, C. (2022). *Golem: A framework for robust shiny applications*. <https://github.com/ThinkR-open/golem>
- Hackett, C. A., Boskamp, B., Vogogias, A., Preedy, K. F., & Milne, I. (2017). Tetraploid-SNPMap: Software for Linkage Analysis and QTL Mapping in Autotetraploid Populations Using SNP Dosage Data. *Journal of Heredity*, 108(4), 438–442. <https://doi.org/10.1093/jhered/esx022>
- Hershberg, E. A., Stevens, G., Diesh, C., Xie, P., De Jesus Martinez, T., Buels, R., Stein, L., & Holmes, I. (2021). JBrowseR: an R interface to the JBrowse 2 genome browser. *Bioinformatics*, 37(21), 3914–3915. <https://doi.org/10.1093/bioinformatics/btab459>
- Mollinari, M., & Garcia, A. A. F. (2019). Linkage Analysis and Haplotype Phasing in Experimental Autopolyploid Populations with High Ploidy Level Using Hidden Markov Models. *G3 Genes/Genomes/Genetics*, 9(10), 3297–3314. <https://doi.org/10.1534/g3.119.400378>
- Mollinari, M., Olukolu, B. A., Pereira, G. da S., Khan, A., Gemenet, D., Yencho, G. C., & Zeng, Z.-B. (2020). Unraveling the Hexaploid Sweetpotato Inheritance Using Ultra-Dense Multilocus Mapping. *G3 Genes/Genomes/Genetics*, 10(1), 281–292. <https://doi.org/10.1534/g3.119.400620>
- Rosyara, U. R., De Jong, W. S., Douches, D. S., & Endelman, J. B. (2016). Software for genome-wide association studies in autopolyploids and its application to potato. *The Plant Genome*, 9(2), plantgenome2015.08.0073. <https://doi.org/10.3835/plantgenome2015.08.0073>
- Silva Pereira, G. da, Gemenet, D. C., Mollinari, M., Olukolu, B. A., Wood, J. C., Diaz, F., Mosquera, V., Gruneberg, W. J., Khan, A., Buell, C. R., Yencho, G. C., & Zeng, Z.-B. (2020). Multiple QTL Mapping in Autopolyploids: A Random-Effect Model Approach with Application in a Hexaploid Sweetpotato Full-Sib Population. *Genetics*, 215(3), 579–595. <https://doi.org/10.1534/genetics.120.303080>
- Zheng, C., Amadeu, R. R., Munoz, P. R., & Endelman, J. B. (2021). Haplotype reconstruction in connected tetraploid F1 populations. *Genetics*, 219(2). <https://doi.org/10.1093/genetics/iyab106>
- Zheng, C., Voorrips, R. E., Jansen, J., Hackett, C. A., Ho, J., & Bink, M. C. A. M. (2016). Probabilistic Multilocus Haplotype Reconstruction in Outcrossing Tetraploids. *Genetics*, 203(1), 119–131. <https://doi.org/10.1534/genetics.115.185579>