

1 **ShinyAIM: Shiny-based Application of Interactive Manhattan Plots for Longitudinal**
2 **Genome-Wide Association Studies**

3 Waseem Hussain^{1,2}, Malachy Campbell^{1,2}, Harkamal Walia², and Gota Morota¹

4 ¹Department of Animal Science, University of Nebraska-Lincoln

5 ²Department of Agronomy and Horticulture, University of Nebraska-Lincoln

6

7 Corresponding author:

8 Waseem Hussain

9 Department of Animal Science, Department of Agronomy and Horticulture

10 University of Nebraska-Lincoln

11 Lincoln, Nebraska 68583

12 Email: waseem.hussain@unl.edu

13

14 **Abstract**

15 Due to advancements in sensor-based, non-destructive phenotyping platforms, researchers are
16 increasingly collecting data with higher temporal resolution. These phenotypes collected over
17 several time points are cataloged as longitudinal traits and used for genome-wide association
18 studies (GWAS). Longitudinal GWAS typically yield a large number of output files, posing a
19 significant challenge for data interpretation and visualization. Efficient, dynamic, and integrative
20 data visualization tools are essential for the interpretation of longitudinal GWAS results for
21 biologists but are not widely available to the community. We have developed a flexible and user-
22 friendly Shiny-based online application, ShinyAIM, to dynamically view and interpret temporal
23 GWAS results. The main features of the application include (i) interactive Manhattan plots for
24 single time points, (ii) grid plot to view Manhattan plots for all time points simultaneously, (iii)
25 dynamic scatter plots for p-value-filtered selected markers to investigate co-localized genomic
26 regions across time points, (iv) and interactive phenotypic data visualization to capture variation
27 and trends in phenotypes. The application is written entirely in the R language and can be used
28 with limited programming experience. ShinyAIM is deployed online as a Shiny web server
29 application at <https://chikudaisei.shinyapps.io/shinyaim/>, enabling easy access for users without
30 installation.

31 **Keywords**

32 ShinyAIM, Longitudinal traits, GWAS, Interactive visualization

33

34 **Introduction**

35 Owing to the availability of high-throughput phenotyping platforms, there is growing interest in
36 the quantitative genetics of longitudinally measured traits, i.e., traits that are measured over
37 multiple time points by advanced imaging systems (Araus and Kefauver 2018; Araus et al.
38 2018). For example, the application of GWAS to responses to abiotic stresses, such as drought,
39 salinity, and temperature tolerance, measured at temporal resolution may provide insights into
40 the mechanisms underlying plant physiological processes measured throughout the duration of
41 stress or development (Busemeyer et al., 2013; Moore et al., 2013; Topp et al., 2013; Slovak et
42 al., 2014; Wu \square rschum et al., 2014; Yang et al., 2014; Bac-Molenaar et al., 2015; Campbell et al
43 2015; Campbell, Walia, and Morota 2018).

44 Data visualization is a fundamental aspect of big data analysis in genetics. Manhattan plots are
45 standard tools used to visualize GWAS results and to identify the genomic regions associated
46 with a given phenotype. However, the static nature of these plots limits the information that can
47 be displayed and extracted. Further, the number of Manhattan plots that can be viewed at once
48 time is limited, making comparisons across phenotypes tedious. The situation becomes more
49 challenging in the case of longitudinal GWAS, which are performed across multiple time points,
50 with each time point treated as an independent phenotype. Furthermore, it is difficult to share
51 GWAS outputs in an easy and convenient way, requiring novel applications for dynamic data
52 visualization and sharing. Many browsers have been built to visualize GWAS outputs (e.g.,
53 Khramtsova and Stranger 2017; Cuellar-Partida, Renteria, and MacGregor 2015; Juliusdottir et
54 al. 2018; Ziegler, Hartsock, and Baxter 2015). However, none of these are specifically tailored to
55 longitudinal traits. Further, none of the applications offer features for the dynamic visualization
56 of Manhattan plots online and for comparisons across timepoints simultaneously.

57 To this end, we developed a Shiny-based application, ShinyAIM, for visualizing and interpreting
58 longitudinal GWAS data in an interactive way. The application is distinct from previously
59 developed GWAS browsers because it is specifically designed for longitudinal traits, allowing
60 the simultaneous visualization of all time points or phenotypes and comparisons of top
61 associated markers across time points. The interactive and integrative data visualization features

62 embedded in the application offer a new resource for users to readily extract extensive
63 information from temporal GWAS results.

64 **Overview of ShinyAIM**

65 **Implementation**

66 ShinyAIM is entirely written in the R language (R Core Team 2018) with the underlying R code
67 encapsulated by the shiny R package (Chang et al. 2018), which is a web application framework
68 for R offering an interactive graphical user interface. Shiny has been making inroads into plant
69 breeding and quantitative genetics for research and teaching purposes, such as Be-Breeder
70 (Fritsche-Neto and Matias, 2016) and ShinyGPAS (Morota 2017). ShinyAIM leverages the
71 cumulative utility of the R packages manhattanly (Sahir 2016) and plotly (Sievert et al. 2017) to
72 create a cohesive web browser-based application. The ShinyAIM application does not require
73 any working knowledge of R and is intuitively operated by mouse clicks. ShinyAIM is hosted by
74 a Shiny web server (<https://chikudaisei.shinyapps.io/shinyaim/>) for online use or can be run
75 locally within RStudio by downloading the source code from the GitHub repository
76 (<https://github.com/whussain2/ShinyAIM>). The ShinyAIM application is open source and is
77 distributed under Artistic License 2.0. The user guidelines, such as input data formats and data
78 upload instructions, are provided in the main tab labeled ‘Information.’ Links to sample files and
79 a video demonstration are also available in the Information tab.

80 **Main features and functionality**

81 The application has four main features to explore GWAS results: (i) interactive Manhattan plots
82 for single time points, (ii) Manhattan grid plot to compare results across all time points
83 simultaneously, (iii) dynamic views of p-value-filtered top associated markers in a scatter plot to
84 identify co-localized markers over time, and (iv) visualization of phenotypic data used for
85 GWAS (Figure 1). These features are supported by user-defined data filtering criteria in
86 ShinyAIM to smoothly navigate the application. Each feature is briefly described in the
87 following sections.

88 **Interactive Manhattan Plots**

89 In the Interactive Manhattan Plots panel, users can interactively view the Manhattan plot for each
90 time point (Figure 1A). After the correct file format is selected and the file is uploaded, the
91 available time points will be automatically updated in the 'Choose Time Point or Phenotypes.'
92 An interactive Manhattan plot is automatically generated on the right hand panel after selecting a
93 target time point. Users can move the mouse over the points in the plot to display detailed
94 information, including the marker name, position, chromosome location, and $-\log_{10}$ p-value.
95 Furthermore, it is possible to zoom in on potential candidate regions to obtain additional detail.
96 ShinyAIM offers the flexibility to choose the significance level by moving the slider input bar. In
97 addition, users have a choice to display a list of markers arranged in decreasing order of p-values
98 in the table below the Manhattan plot panel. The display also includes marker information in the
99 input data file. The slider input bar controls the number of markers shown in the table.

100 **Manhattan Grid Plot**

101 Manhattan Grid Plot tab allows users to visualize the Manhattan plots combined for all time
102 points and can be used to explore how GWAS peaks change over time to facilitate data
103 interpretation (Figure 1B). The significance threshold for markers can be modified by moving
104 the slider input bar. Moreover, ShinyAIM enables users to choose the number of columns and
105 rows in the grid plot by moving the slider input bar 'Select the Number of Columns in Grid Plot.'

106 **Comparison of Associated Markers**

107 Users are able to dynamically view only the top associated markers in a scatter plot (Figure 1C).
108 This feature is implemented in ShinyAIM to enable users to focus only the topmost associated
109 markers and compare these markers across time points to identify co-localized regions. Users can
110 select the number of markers displayed in a scatter plot by filtering the markers based on p-
111 values. This is achieved by directly typing or selecting the option 'Select Top Markers Based on
112 p-value.' The scatter plot is interactive and users can move the mouse over a point to display
113 information, including the time point, chromosome name, position of the marker, name of the
114 marker, and $-\log_{10}$ p-value (Figure 1C).

115 **Phenotypic Data Visualization**

116 Phenotypic data visualization helps users view phenotypes used for GWAS in the forms of
117 dynamic histograms and density plots (Figure 1D). The trends and variability in phenotypic
118 values at each time point can be visualized using box plots. All plot types are interactive and
119 users can move the mouse over a particular point to obtain detailed information.

120 **Conclusion**

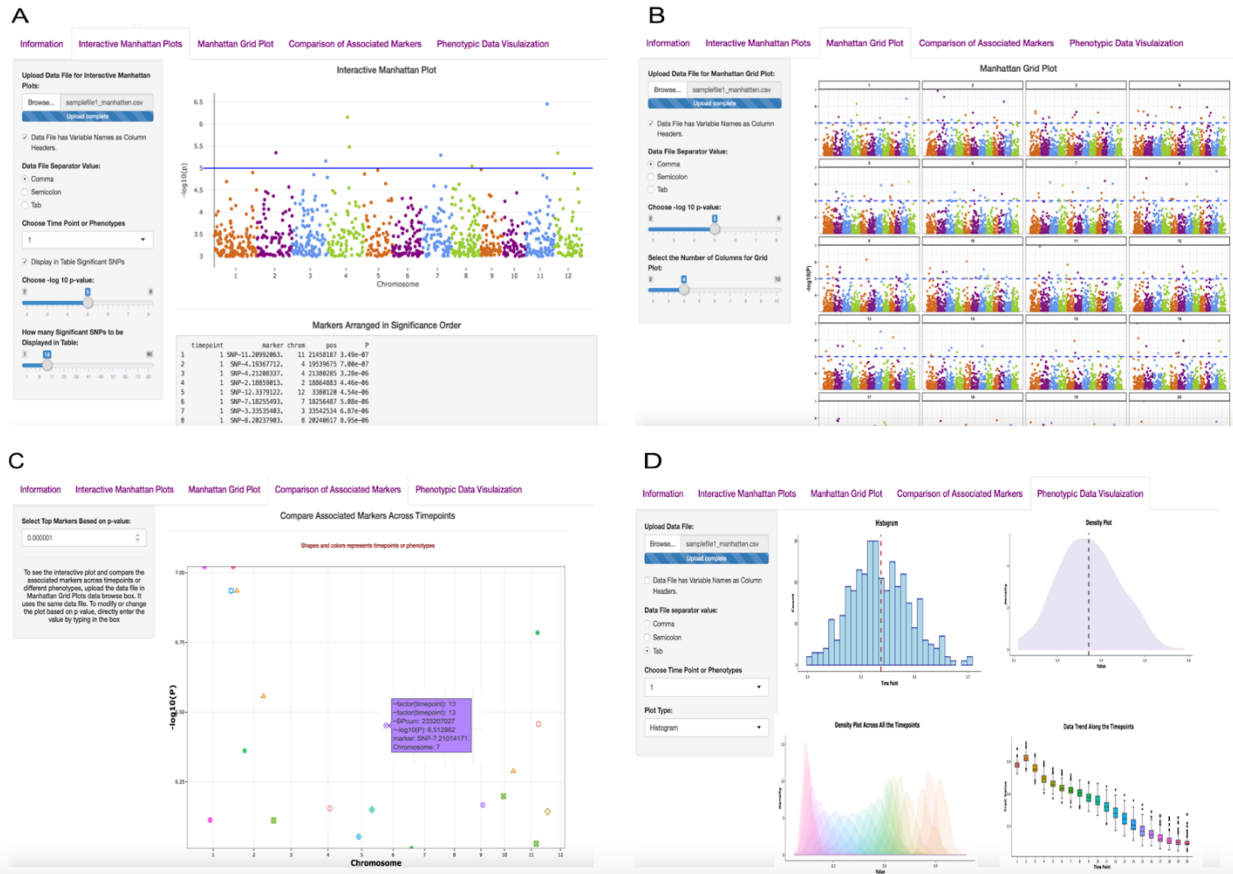
121 We have developed a user-friendly integrative Shiny-based application to dynamically visualize
122 and interpret longitudinal GWAS results, providing an easy-to-use online tool to the community.

123 **Availability**

124 The source code for the ShinyAIM application is freely available at
125 <https://github.com/whussain2/ShinyAIM> licensed under Artistic License 2.0. ShinyAIM can be
126 launched on any system that has RStudio installed or available online at the Shiny web server
127 <https://chikudaisei.shinyapps.io/shinyaim/>.

128

129 Figures



130

131 **Figure 1:** Main interface of the ShinyAIM application. Screenshots of panels for the main tabs
132 are shown. (A) The Interactive Manhattan Plots tab allows users to display interactive Manhattan
133 plots for a selected time point. Users have the flexibility to choose the significance level and can
134 display the top associated markers in tabular format. (B) The Manhattan Grid Plot tab allows
135 users to visualize Manhattan plots for all time points simultaneously. Users have the flexibility to
136 choose the significance level and the number of columns in the grid plot. (C) The Comparison of
137 Associated Markers tab allows users to filter markers based on p-values, display a scatter plot for
138 comparisons across all time points, and search for co-localized markers. (D) The Phenotypic
139 Data Visualization tab generates histogram and density plots and summarizes trends in temporal
140 phenotypic data in the form of box plots.

141

142 **Acknowledgments**

143 This work was funded and supported by National Science Foundation Grant No.1736192.

144 **Conflict of interest**

145 The authors declare there are no competing interests.

146 **References**

147 Araus, J. L., & Kefauver, S. C. (2018). Breeding to adapt agriculture to climate change:

148 Affordable phenotyping solutions. *Current Opinion in Plant Biology* (In Press).

149 <https://doi.org/10.1016/j.pbi.2018.05.003>

150 Araus, J.L., Kefauver, S. C., Zaman-Allah, M., Olsen, M. S., & Cairns, J. E. (2018). Translating
151 high-throughput phenotyping into genetic gain. *Trends in Plant Science*, 23(5), 451–466.

152 <https://doi.org/10.1016/j.tplants.2018.02.001>

153 Bac-Molenaar, J. A., Vreugdenhil, D., Granier, C., & Keurentjes, J. J. B. (2015). Genome-wide
154 association mapping of growth dynamics detects time-specific and general quantitative trait loci.

155 *Journal of Experimental Botany*, 66(18), 5567-5580. <http://dx.doi.org/10.1093/jxb/erv176>

156 Bhatnagar, S. (2016). Manhattanly: Interactive Q-Q and Manhattan Plots Using “Plotly.js.”

157 <https://CRAN.R-project.org/package=manhattanly>.

158 Busemeyer, L., Ruckelshausen, A., Möller, K., Melchinger, A. E., Alheit, K. V., Maurer, H. P.,

159 Hahn, V., Weissmann, E. A., Reif, J. C., & Wu \square rschum, T. (2013). Precision phenotyping of

160 biomass accumulation in triticale reveals temporal genetic patterns of regulation. *Scientific*

161 *Reports*, 3, 2442. <https://www.nature.com/articles/srep02442>

162 Campbell, M. T., Walia, H., & Morota, G. (2018). Utilizing random regression models for

163 genomic prediction of a longitudinal trait derived from high-throughput phenotyping. BioRxiv

164 319897 [Preprint]. <https://doi.org/10.1101/319897>

- 165 Chang, Winston., Cheng, J., & Allaire, et al. (2018) Shiny: Web application framework for R.
166 <https://CRAN.R-project.org/package=shiny>.
- 167 Cuellar-Partida, G., Renteria, M. E., & MacGregor, S. (2015). LocusTrack: Integrated
168 Visualization of GWAS Results and Genomic Annotation. *Source Code for Biology and*
169 *Medicine*, 10, 1. <https://doi.org/10.1186/s13029-015-0032-8>
- 170 Fritsche-Neto, Roberto., & Matias, F. I. (2016). Be-Breeder - Learning: A new tool for teaching
171 and learning plant breeding principles. *Crop Breeding and Applied Biotechnology*, 16(3), 240–
172 245. <http://dx.doi.org/10.1590/1984-70332016v16n3n36>.
- 173 Juliusdottir, T., Banasik, K., Robertson, N. R., Mott, R., & McCarthy, M. I. (2018). Toppar: An
174 Interactive Browser for Viewing Association Study Results. *Bioinformatics*, 34(11), 1922–1924.
175 <https://doi.org/10.1093/bioinformatics/btx840>
- 176 Khramtsova, E. A., & Stranger, B. E. (2017). Assocplots: A Python package for static and
177 interactive visualization of multiple-group GWAS results. *Bioinformatics*, 33(3), 432–434.
178 <https://doi.org/10.1093/bioinformatics/btw641>
- 179 Moore, C. R., Johnson, L. S., Kwak, I. Y., Livny, M., Broman, K. W., & Spalding, E. P.
180 (2013). High-throughput computer vision introduces the time axis to a
181 quantitative trait map of a plant growth response. *Genetics*, 195, 1077–
182 1086. <https://doi.org/10.1534/genetics.113.153346>
- 183 Morota, G. (2017). ShinyGPAS: Interactive Genomic Prediction Accuracy Simulator based on
184 deterministic formulas. *Genetics, Selection, Evolution* 49: 91. [https://doi.org/10.1186/s12711-](https://doi.org/10.1186/s12711-017-0368-4)
185 [017-0368-4](https://doi.org/10.1186/s12711-017-0368-4)
- 186 R Core Team. (2018). R: The R Project for Statistical Computing. Vienna, Austria.
187 <https://www.r-project.org/>.
- 188 Sievert, C., Parmer, C., Hocking, et al. (2017). Plotly: Create Interactive Web Graphics via
189 “Plotly.Js.” <https://CRAN.R-project.org/package=plotly>

- 190 Slovak, R., Göschl, C., Su, X., Shimotani, K., Shiina, T., & Busch, W. (2014). A
191 scalable open-source pipeline for large-scale root phenotyping of Arabidopsis.
192 *Plant Cell*, 26, 2390–2403. <https://doi.org/10.1105/tpc.114.124032>
- 193 Topp, C. N., Iyer-Pascuzzi, A. S., Anderson, J. T., Lee, C. R., Zurek, P. R., Symonova, O.,
194 Zheng, Y., Bucksch, A., Mileyko, Y., Galkovskyi, T., et al. (2013). 3D phenotyping and
195 quantitative trait locus mapping identify core regions of the rice genome controlling root
196 architecture. *Proceedings of National Academy of Science USA*, 110: E1695–E1704.
197 <https://doi.org/10.1073/pnas.1304354110>
- 198 Wu□rschum, T., Liu, W., Busemeyer, L., Tucker, M. R., Reif, J.C., Weissmann, E. A., Hahn,
199 V., Ruckelshausen, A., & Maurer, H. P. (2014). Mapping dynamic QTL for plant height in
200 triticale. *BMC Genetics*, 15, 59. <https://doi.org/10.1186/1471-2156-15-59>
- 201 Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., Fang, W., Feng, H., Xie, W., Lian,
202 X., et al. (2014). Combining high-throughput phenotyping and genome-wide association studies
203 to reveal natural genetic variation in rice. *Nature Communications*, 5, 5087.
204 <https://www.nature.com/articles/ncomms6087>
- 205 Ziegler, G. R., Ryan, H. H., & Baxter, I. (2015). Zbrowse: An interactive GWAS results
206 browser. *PeerJ Computer Science* 1, e3. <https://peerj.com/articles/cs-3/>