*Bioinformatics, 33(18), 2023, 2933–2935*

doi: 10.1093/*bioinformatics*/ab345

Advance Access Publication Date: 21 March 2023

Applications Note

|  |
| --- |
| Genome Analysis  **Unlocking the Secrets of Allelic Variation: Insights from Allele Frenzy, an Interactive application**  Eleni Theofania Skorda1,\*  1Department of Biology, Box 118, 221 00, Lund University, Sweden.  \*To whom correspondence should be addressed.  Abstract  **Motivation:** A dynamic and interactive tool for investigating genetic information and comprehending the connections between different alleles and phenotypes is provided by the Allele Frenzy program. This application enables users to visualize the impacts of genetic variants on phenotype expression, examine allele frequencies in many or few populations, investigate allele frequencies and depict the effects of natural selection. We also discuss the future of this application for enhancing genetic research and education, as well as directions for further development and improvement. All in all, Allele Frenzy offers a user-friendly environment and entertaining approach for anyone interested in population genetics, genomics, and evolution.  **Contact:** ethskorda@gmail.com  **Supplementary information:** <https://github.com/elenitskorda/AlleleFrenzy> |

# Introduction

SNPs (single nucleotide polymorphisms) are a generic variation in the genome and are used in order to detect genomic regions associated with diseases and traits. Their high abundance and variability made the scientific audience explore them correlated to genetic association studies and to identify genes related to several diseases and the development of medicine. Also, in the population genetics field, SNPs could be further examined allowing scientists to detect migrations, gene flow, and evolutionary history (Broeckel et al. 2004).

Thus, it is important to investigate genetic variation to understand the heritability of complex traits and the underlying genetic basis of diseases(Broeckel et al. 2004). To address this challenge, we have developed an interactive web application called "Allele Frenzy" that enables users to visualize and explore genetic variation data. Allele Frenzy gives the possibility to users to explore and analyze genomic variation data in a user-friendly and interactive way. The application offers several plotting options, including allele frequency spectrum, genotype counts, and allele frequency by generation.

Compared to existing tools such as Genomepop (Rodriguez, 2008) Allele Frenzy offers several advantages. Firstly, the app gives the opportunity to researchers and scientists without strong programming skills to visualize and analyze their genetic data. Secondly, the tool is interactive, enabling users to explore their data in real-time, which makes it not complicated and facilitates the patterns and relationships within the data. Thirdly, the application includes a scale for each input, and thus making the visualization easier for big data and large datasets.

The aim of this paper is the creation of the tool Allele Frenzy and to establish its significance. We further demonstrate how this tool may be used to recognize population structure, visualize patterns and trends of genetic variation. We believe that Allele Frenzy will be a useful resource for academics looking at genetic variety, especially those who have modest programming knowledge.

# Methods

* 1. **Data**

For the construction of the "Allele Frenzy" app we used R version (4.2.2). The app was built using the shiny package (Chang et al., 2021) version (1.7.4) and deployed using the shinyapps.io platform. First, to calculate the minor allele frequencies (MAF) for different generations, the user inserts the input parameters. Later, we include the genetic drift, by adding a normally distributed random variable to the allele frequencies using the rnorm() function, with the mean equal to zero and the standard deviation set to the genetic drift. For generating mixed populations, we calculated the weighted aSverage of the allele frequencies using the sample() function, same as the calculation of minor allele frequency for all SNPs. The results were saved to a data frame and then appended to a list. Finally, we combined the data frames into one data frame using the do.call() and rbind() functions.

* 1. **Simulation**

After configuring the input parameters and the equivalent equations, we simulate the minor allele frequency for each SNP at each generation by using ggplot2 package version (3.4.1) It is significant to note that the input data used within the application toll did not contain any real-world in formations and data.

# Features

The application developed for this study is a web-based tool for simulating genetic drift in multiple populations over time. The user provides input parameters such as Figure 1:

1. Initial allele frequencies
2. Number of populations
3. Number of generations
4. Genetic drift.

Graphical user interface, chart, scatter chart

Description automatically generatedThe application then uses these parameters to simulate changes in allele frequencies over time and generates a plot of the minor allele frequency for each SNP over the specified number of generations selected. One of the main features of the application is its user-friendly interface, which allows users to easily adjust input settings and visualize the output. The web application has the benefit of a standalone program as it can be accessed from any device with an internet connection, making it more accessible to a wider audience and can be simply updated and maintained by the developers without requiring users to download and install new versions of the software. However, one potential drawback is that it may require a reliable internet connection and might run slower than a standalone program.

Figure 1: Demonstration of Allele Frenzy application tool. On the left side of the window, the user adjusts the input parameters, and the plot are adjusted accordingly to the right side with the generation change on the y axis and the minor allele frequency on the x axis.

# Discussion

In this study, we have developed a web application to model the evolution of MAF under different generations in mixed populations subject to genetic drift. The software creates plot displaying the MAF over generations on user defined settings for the number of populations, number of SNPs, starting allele frequencies, and genetic drift rate. The application has a user-friendly interface that makes scientists investigate the effects of various parameters on allele frequency changes in mixed populations.

One key benefit of our web application is its accessibility. The application is hosted on a server, and thus users can access it from any device with an internet connection, avoiding installation of any software or dealing with compatibility difficulties. This raises the application accessibility to researchers who may lack coding or computational approach knowledge. Also, the interactive aspect of the online interface makes it for users easy to explore different parameter choices and visualize the outcomes.

However, the tool has certain acknowledged flaws and restrictions. For instance, the simulation makes the assumption that all populations have the same starting allele frequencies, which may not always reflect biology in the real world. Additionally, the application’s scalability may be constrained by how computationally intensive large parameter settings are.

There are several actions which need to take place for further development of Allele Frenzy app. One of these could be the incorporation of more intricate models that take different population structures into account. The Allele Frenzy app should also enable users to upload selected form of input files in order to investigate real time scenarios and clinical cases. Last but not list, the generated code could be improved for scalable of larger configurations, for instance by adding more input parameters. All things considered; we anticipate that our web application will be a helpful tool for researchers looking to explore the dynamics of allele frequency changes in mixed populations under genetic drift.

Acknowledgements

We would like to express our gratitude to the editor and two anonymous reviewers for their insightful comments and suggestions that greatly improved the quality of our manuscript. We also acknowledge Lund University for providing financial support for this research project".

References

Broeckel U, Schork NJ. Identifying genes and genetic variation underlying human diseases and complex phenotypes via recombination mapping. J Physiol. 2004 Jan 1;554(Pt 1):40-5. doi: 10.1113/jphysiol.2003.051128. PMID: 14678489; PMCID: PMC1664744.

Carvajal-Rodríguez, A. GENOMEPOP: A program to simulate genomes in populations. BMC Bioinformatics 9, 223 (2008). <https://doi.org/10.1186/1471-2105-9-223>