Springer Plus
a SpringerOpen Journal

SOFTWARE   **Open Access**

STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface

Ramesh Krishnan Ramasamy[1†], Sumathy Ramasamy[2†], Bharat Bushan Bindroo[1,2] and V Girish Naik[1*]

Abstract

Background: Understanding structure of the population is one of the major objective of many genetic studies. The program STRUCTURE is commonly used to infer population structure using multi-locus genotype data. However, a tool with graphical-user interface is currently not available to visualize STRUCTURE bar plots.

Results: We introduce STRUCTURE PLOT, a program for drawing STRUCTURE bar plots. The program generates publication ready, aesthetic STRUCTURE bar plots by using individual Q matrix from STRUCTURE or CLUMPP output. The program is very simple to use and includes variety of options like sorting bar by original order or by K, and selection of colors from R colors or RColorBrewer palette. Individual or population labels can be printed below or above the plot in any angle. Size of the graph and label can be defined, and option is provided to save plot in variety of picture formats in user defined resolution.

Conclusion: The program is implemented as a web application for online users and also as a standalone shiny application. Web application is compatible to majority of leading web browsers and standalone version can be launched using a simple R command. The program can be freely accessed at http://btismysore.in/strplot.

Keywords: DISTRUCT; CLUMPP; STRUCTURE HARVESTER; R program; Shiny

Introduction

The program STRUCTURE (Pritchard et al. 2000) is one of the widely used genetic analysis software over the last decade to infer population structure and gene flow by using multi-locus genotypic data. STRUCTURE implements model based method to assign each individual to one assumed population (K) or more than one population, if it is an admixture. Estimating suitable assumed population size for the dataset is usually based on maximum likelihood (LnPD) value inferred from STRUCTURE run or delta K (Δk) value

(Evanno et al. 2005). STRUCTURE HARVESTER (Earl 2012) program was designed to carry out downstream processing of STRUCTURE results to calculate Evanno's Δk value and prepares input file for CLUMPP program (Jakobsson and Rosenberg 2007). CLUMPP program permutes replicated runs of STRUCTURE software to find a close match among iterated runs. CLUMPP results are used to generate bar graphs using DISTRUCT (Rosenberg 2004) program. DISTRUCT is a standalone program without Graphical User Interface (GUI), and the drawing parameters needs to be defined in a separate file. DISTRUCT returns the result file in postscript format and users have to install third-party software to convert postscript to the graphical format. To overcome these limitations, we introduce a user-friendly program – 'STRUCTURE

* Correspondence: vgirishnaik@yahoo.com
†Equal contributors
[1]Molecular Biology Laboratory-1, Host Plant Improvement, Central Sericultural Research and Training Institute, Srirampura, Manandavadi Road, Mysore 570008, India
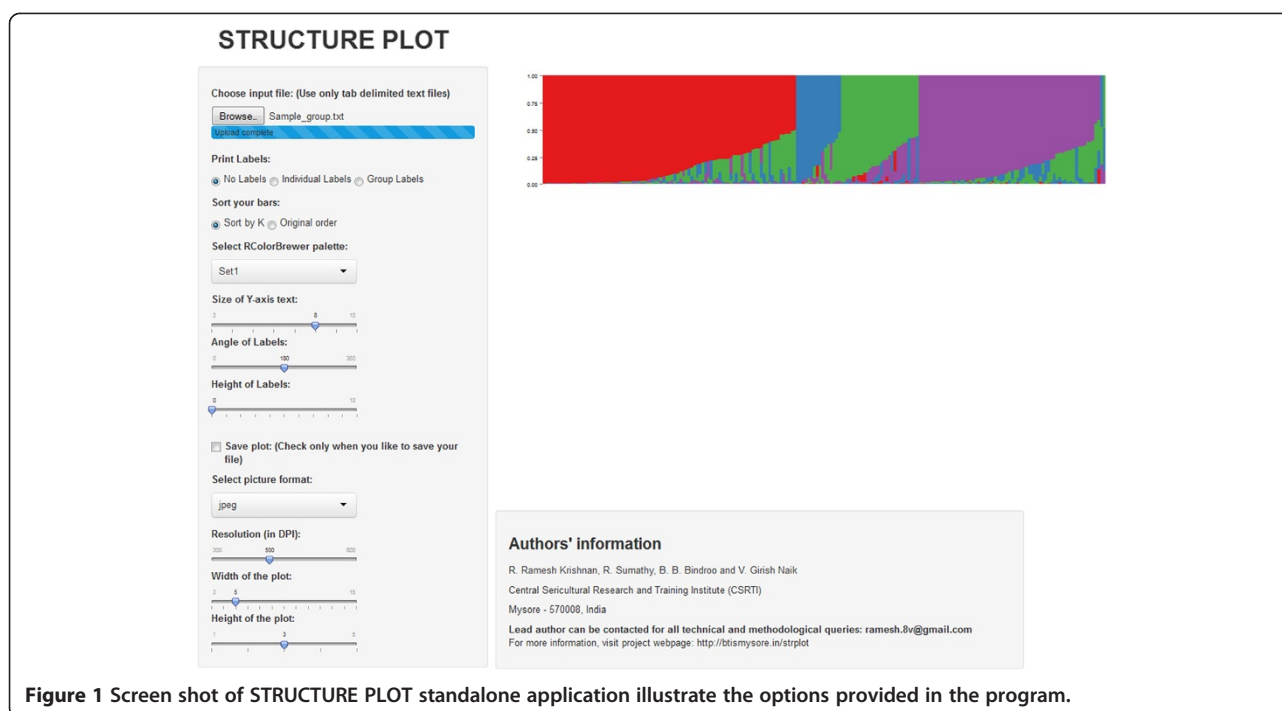Full list of author information is available at the end of the article

Figure 1 Screen shot of STRUCTURE PLOT standalone application illustrate the options provided in the program.

PLOT' to draw elegant bar plots with graphical user interface.

Functionality description

The program is implemented as a web application with user friendly interface also as a shiny (http://www.rstudio.com/shiny/) standalone application. The program accepts individual Q matrix as input file and plot settings can be adjusted using a graphical interface (Figure 1). R is used as scripting language, and ggplot2 package (Wickham 2009) is used for plotting bar graphs. Any suitable colour palette can be selected from RColorBrewer (Neuwirth 2007) or custom colours can be passed to each K. Users can simply copy the individual Q matrix from STRUCTURE or CLUMPP output and used as input file for STRUCTURE PLOT. For a demo, a sample dataset with 202 diverse coconut genotypes sampled by Krishnan et al. (2014) was provided in the webpage. Bars can be drawn in the original order, or it can be sorted by Q. Bar plots are displayed on the same page instead of separate result page. Therefore, the user can play around with settings and visualise the changes in the plot. The program is capable of printing individual or population labels below or above the plot in any angle. Shiny application is built in reactive programming environment, therefore, results can be viewed instantly while adjusting plotting parameters. Options are provided to change the dimensions of the graph and size of the axis labels. Plots can be downloaded in a variety of picture formats in user defined resolution. STRUCTURE PLOT web application can be accessed from any computer with internet connectivity and compatible to majority of the leading web browsers. Shiny application can be launched from R console by using a simple command, and all the parameters can be defined using a graphical interface. We are intended to provide a variety of plotting options in coming days.

Author details
[1]Molecular Biology Laboratory-1, Host Plant Improvement, Central Sericultural Research and Training Institute, Srirampura, Manandavadi Road, Mysore 570008, India. [2]Bioinformatics Centre, Central Sericultural Research and Training Institute, Srirampura, Manandavadi Road, Mysore 570008, India.

References

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4(2):359–361

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14(8):2611–2620

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23(14):1801–1806

Krishnan RR, Sumathy R, Ramesh SR, Bindroo BB, Naik VG (2014) SimEli: Similarity Elimination method for sampling distant entries in development of core collections. Crop Sci doi:10.2135/cropsci2013.09.0600

Neuwirth E (2007) RColorBrewer: ColorBrewer palettes. R package version 1(5):

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 4(1):137–138

Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer, New York