AFCMHeatMap: A Shiny web Tool for HeatMap Generation of genetic expression datasets using R packages

Mohammad M. Tarek[1], Ayman E. Shafei[2], Mahmoud A. Ali[3]

[1]Bioinformatics Department, Armed Forces College of Medicine (AFCM) Cairo, Egypt.

[2]Biomedical sciences Department, Armed Forces College of Medicine (AFCM) Cairo, Egypt.

[3]Microbiology Department, Armed Forces College of Medicine (AFCM) Cairo, Egypt.

Abstract

Generating heatmaps of genetic datasets is a 2D graphical visualization of data where the individual expression values contained in a matrix are represented as colors. Herein, we describe AFCMHeatMap a shiny web App that integrates quantitative interaction of genomics data and results from microarrays or RNA-Seq to highlight expression levels of various genetic datasets with a *.CSV input file. The application also facilitates downloading heatmaps as a supplementary material for user's publications. Written in R using Shiny framework, it is a user-friendly framework for interactive expression data visualization that can be easily deployed without any restrictions to any operating system used by any online user.

Availability: It is an open-source and available at https://github.com/MTarekM/AFCMHeatMapGeneratingTool

Corresponding Author: Mohammad M. Tarek[1]

Contact: mohammadtareq459@gmail.com

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2961v2 | CC BY 4.0 Open Access | rec: 26 Dec 2017, publ: 26 Dec 2017

Introduction

Heatmaps are 2D graphical false-color image demonstrations of genetic expression information with a predefined color pattern to visualize values in a form of data matrix.[1] Heatmaps visualize genetic expression levels within different samples obtained whether from high-throughput sequencing tools like RNA-seq or DNA microarrays.[2] Overexpressed genes are marked in a different way than genes that are downregulated, therefore it provides an instantaneous visualization of genetic expression levels within different types of cancer samples.[3] Heatmaps are also being used to differentiate which clusters of genes are being upregulated or downregulated in variety of cancer samples as compared to control samples as well as profiling expression data of cancer samples to evaluate the differential expression patterns of genes across multiple tumors.[4] It is so helpful to have an open-source application that provides an easy user-interface which provides users with the capability to generate heatmaps for their genetic data. Many Software Applications depending on different programming languages have provided several platforms for generating static non-interactive heatmaps without interactive parameters and overloaded with unfriendly user-interfaces. Recently, R language has been enriched with different packages such as shinyheatmap[5] which is a newly developed shiny platform that could facilitate the process of developing heatmap generation tools. Therefore, we tried to design an application with easy user-interface that helps users to input their data in a CSV file format embedded with the genetic dataset to be further used for implementing and drawing heatmaps. The Application

Unlike static non-interactive heatmap generation tools, could help users to interact with the produced heatmaps and permit flexible visualization of regions of interest on the heatmap.

Input

The Application uses a .CSV file containing the gene expression matrix, data manipulation and adjustment of matrix rows and columns was done by calling functions of dplyr[6] R package The user could select the non-control groups for normalization process prior to visualization processing. Statistical cutoffs could also be changed according to user preferences of p-value, FDR or both.

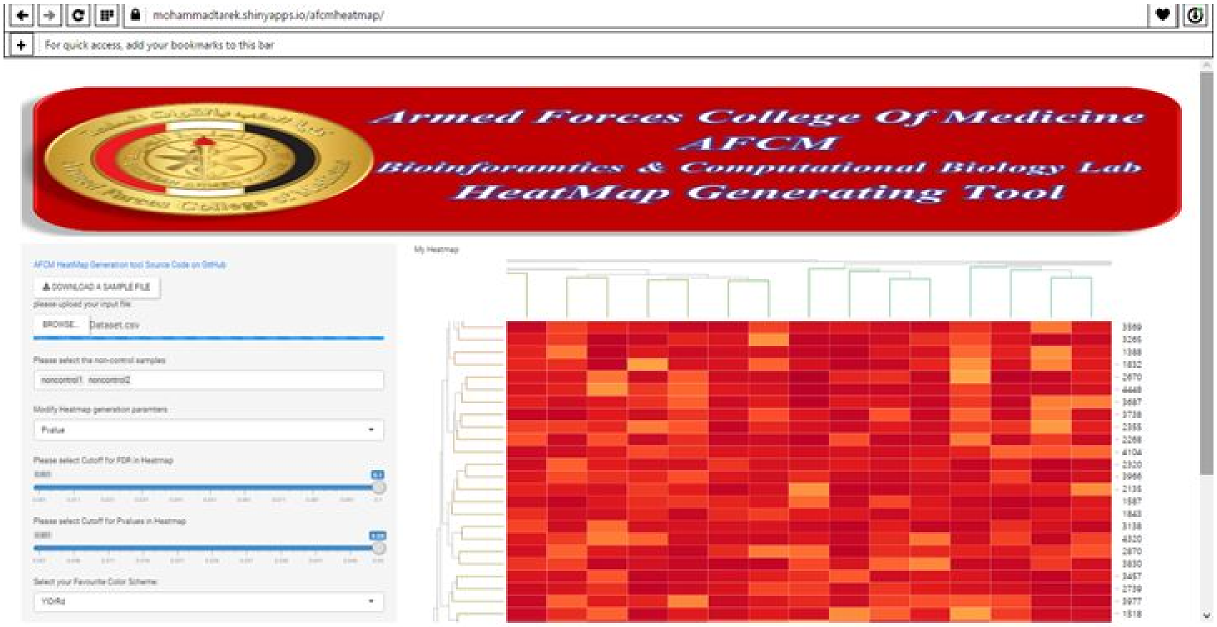
Output

The output file is an HTML file containing the heatmap where the user can download that file for supplementing publications. For submitting another file, the user should just refresh the server without any Limitations to submissions number.

Colors may be edited on the user demand using color scheme parameters provided through RColorBrewer[7] package, which represents a vector of variety of colors, or a function which is of highly scaled data points as input and gives back colors. The User can perform clustering on rows or columns, can interactively apply log2 transform for his heatmap and modify heatmap parameters based on P-values, False Discovery rates FDR or both which is done through calling functions of edgeR[8] package. Deploying edgeR package functions for data analysis was customized to allow higher limit for analysed results focusing on data normalization and data filtration cutoffs.

Implementation

The Application is written in code based on R standard packages using Shiny framework.[9] Our Application implements the growing utility of R packages shiny. D3heatmap[10] is a novel R package which provides interactive capabilities for heatmaps plotting using htmlwidgets[11] & d3.js. D3heatmap is written in a way to have an easy-deployed feature set as well as API for any bioinformatician has previously used heatmap.2 function to build static non-interactive heatmaps. So the user could specify clustering, dendrograms, and variety of scaling preferences in a very friendly way. d3heatmap also comprises variety of features like displaying the values of rows or columns with pointer, Clicking on rows or columns to visualize expression values easily. The package also allows clustering ad zooming through the heatmap allowing users to inspect specific values on large heatmaps.The deployed functions in this tool were inspired by the usage of the same packages on MicroScope[12] shiny App that facilitated deploying d3heatmap and edgeR functions to provide user-friendly heatmap generation. AFCMstats Function that was deployed on the server file of AFCMHeatMap was inspired by the MicroScope[13] SlimStats function with customization to limit visualization spaces and give more space to user specific zooming.



please upload your input file: Dataset.csv

Please select the non-control samples: noncontrol1 noncontrol2

Modify Heatmap generation parameters: Pvalue

Please select Cutoff for FDR in Heatmap

Please select Cutoff for Pvalues in Heatmap

Select your Favourite Color Scheme: YlOrRd

Figure-1 A screenshot of the AFCMHeatMap tool showing the user online view

Design

The HTML framework for the app was created in R console using HTMLwidgets package.[11] Heatmap image extension was subjected to image processing using png[12] R package. For making the application more user-friendly a bootstrap theme was applied using shinythemes[13] R package.

Conclusions

Heatmaps generation within the R shiny App AFCMHeatMap, without any pre required programming skills of the user provides a helpful tool that targets wide population of scientists and students studying genetic expression to provide them with easier ways of graphical visualization and analysis of expression data to investigate genetic datasets using an interactive heatmap generation platform.

Acknowledgement

The Authors are very thankful to the authors of the paper entitled “MicroScope: ChIP-seq and RNA-seq software analysis suite for gene expression heatmaps”,[11] for providing their source code freely available on github,[15] in which their developed functions have helped so much to use the required packages for this Shiny App.

Availability and requirements

- **App name:** AFCM HeatMap Generation tool
- **App home page:** https://mohammadtarek.shinyapps.io/AFCMHeatMap/
- **Operating system:** Only an Internet Connection is required.
- **App Programming platform:** R, Shiny
- **License:** GNU General Public License version 3.0 (GPL-3.0)

Source Code Availability

All source code files were made available on Github at: https://github.com/MTarekM/AFCMHeatMapGeneratingTool

Competing interests: **The authors show no competing interests.**

References

1. Bailey T, Krajewski P, Ladunga I, Lefebvre C, Li Q, Liu T, Madrigal P, Taslim C, Zhang J. Practical Guidelines for the Comprehensive Analysis of ChIP-seq Data. PLoS Comput Biol. 2013; 9(11):e1003326.
2. Soneson C, Delorenzi M. A comparison of methods for differential expression analysis of RNA-seq data. BMC Bioinforma. 2013; 14:91.
3. Škuta C, Bartůněk P, Svozil D. InCHlib — interactive cluster heatmap for web applications. J Cheminformatics. 2014; 6(44):1–9.
4. Turkay C, Lex A, Streit M, Pfister H, Hauser H. Characterizing cancer subtypes using dual analysis in Caleydo StratomeX. IEEE Comput Graph Appl. 2014; 34(2):38–47.
5. Khomtchouk BB, Hennessy JR, Wahlestedt C (2017) shinyheatmap: Ultra fast low memory heatmap web interface for big data genomics. PLoS ONE 12(5): e0176334.
6. Wickham H, Francois R. RStudio. dplyr: A Grammar of Data Manipulation. 2015. R package version 0.4.3.
7. Neuwirth E. RColorBrewer: ColorBrewer Palettes. 2014. R package version 1.1-2.
8. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26:139–40.
9. Chang W, Cheng J, Allaire JJ, Xie Y, McPherson J, RStudio, et al. shiny: Web Application Framework for R. 2015. R package version 0.12.2.
10. Cheng J, Galili T, RStudio Inc., Bostock M, Palmer J. d3heatmap: Interactive Heat Maps Using ‘htmlwidgets’ and ‘D3.js’. 2015. R package version 0.6.1.

11. Vaidyanathan R, Xie Y, Allaire JJ, Cheng J, Russell K. RStudio. htmlwidgets: HTML Widgets for R. 2015. R package version 0.5.
12. *Khomtchouk BB, Hennessy JR, Wahlestedt C. Microscope: ChIP-seq and RNA-seq software analysis suite for gene expression heatmaps. BMC Bioinformatics. 2016; 17(1):390.*
13. Urbanek S. png: Read and write PNG images 2013. R package version 0.1-7
14. Chang W., RStudio , Park T., Dziedzic L., Willis N., Google Corporation, et al. shinythemes: Themes for Shiny 2016. R package version 1.1.1
15. https://github.com/Bohdan-Khomtchouk/Microscope

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2961v2 | CC BY 4.0 Open Access | rec: 26 Dec 2017, publ: 26 Dec 2017