

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

Mohammad M.Tarek¹, Ayman E.Shafei², Mahmoud A.Ali³

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

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

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

Abstract: Generating heatmaps of genes 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 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 integrative genetic analyses that can be easily deployed without any restrictions to any operating system used by any user just an internet connection is required.

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

Contact: mohamadtareq459@gmail.com

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 RNAseq. or DNA microarrays. ^[2] Overexpressed genes are marked in a different way than genes that are down regulated, 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 cluster of genes are being upregulated or downregulated in variety of cancer samples as compared to control samples as well as expression profiling of cancer samples to evaluate the differential expression patterns of genes across multiple tumors^[4]. It's so helpful to have an open-source application that provides easy user interface that allows the user to generate heatmaps for his genetic data. Many Software Applications depending on different programming languages have provided several platforms for generating static heatmaps without interactive parameters and overloaded with unfriendly user interfaces. Therefore we tried to design an application with easy user interface that allow it's user to input his data-file as a CSV file embedded with the genetic dataset to be further used for implementing and drawing heatmaps. The Application Unlike static heatmap generation tools, should allow the user to interact with the produced heatmap and permit flexible visualization of area of interest.

Input:

The Application uses a .CSV file containing the gene expression matrix, then the user should select the non-control samples.

Output:

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

Implementation

the Application is written in code based on R (R Core Team, 2016) standard packages using Shiny framework (Chang et al., 2016).^{[5][6]} Our Application implements the growing utility of R packages shiny ^[6], htmlwidgets ^[7], RColorBrewer ^[8], dplyr ^[9], [d3heatmap](#) ^[10] and edgeR ^[11]. D3heatmap is a novel R package which provides interactive capabilities for heat maps plotting using [htmlwidgets](#) & [d3.js](#). d3heatmap is written in a way to have an easy-deployed feature set as well as API for any bioinformatician who has used either [heatmap.2](#) package to create static heatmaps. So the user could specify clustering, dendrograms, and variety of scaling preferences in a very friendly way. d3heatmap also comprises variety of features such as Displaying the values of rows or columns with pointer, Clicking on rows or columns to highlight data and Dragging a quad over the heatmap for zooming in.

Colors may be edited on the user demand using color scheme parameters. Using [RColorBrewer](#) 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.

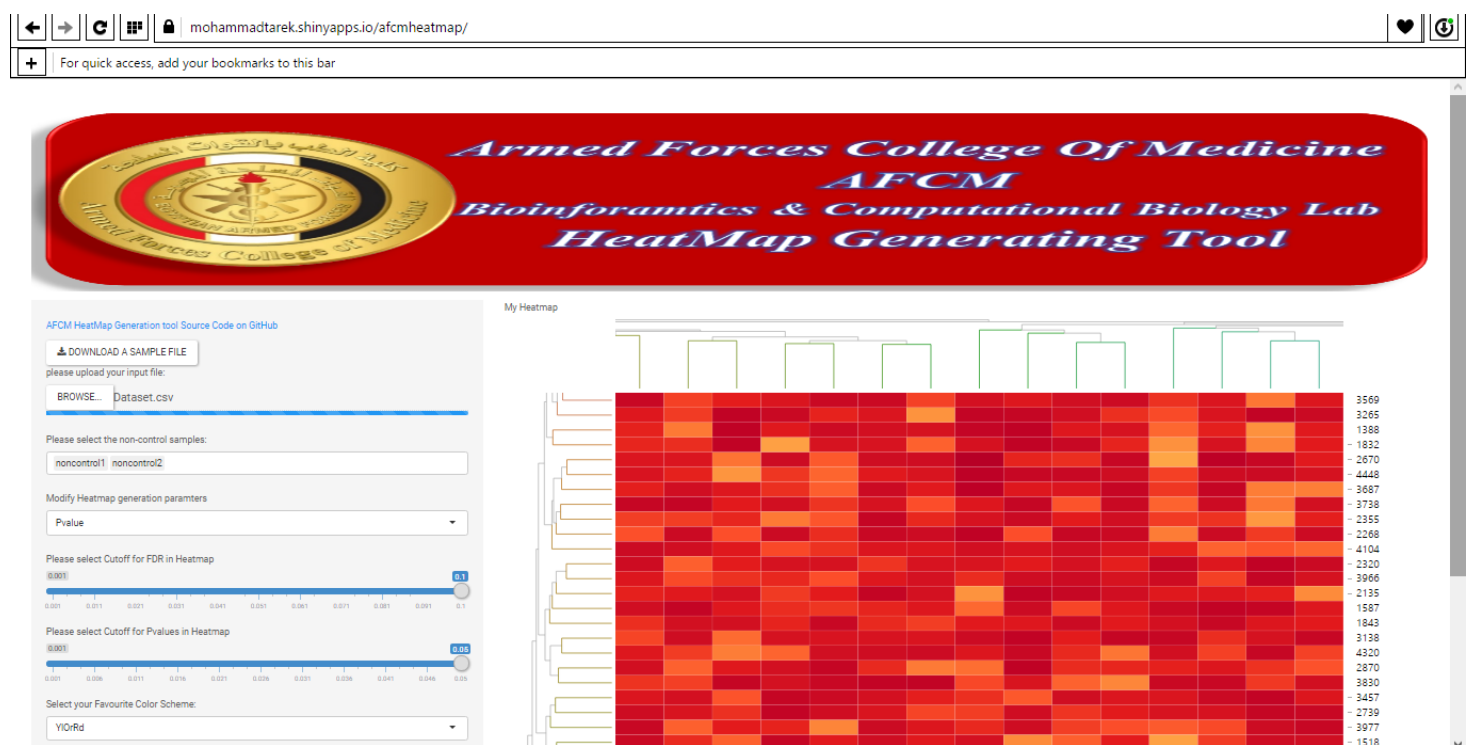


Figure-1 A screenshot of the AFCMHeatMap tool

The Application is designed with huge amount of gene expression data in heatmaps, where there is high probability for overlapping between gene expression values. Application provides the user with ability to interact through the pointer over any specific gene to visualize expression magnitude, gene name, and column ID. Our Application also permits user to repeatedly focus zooming in to any branch or cluster of the heatmap to investigate up to a single gene data. The Specifying heatmap statistical parameters influences the contents of the heatmap visualization itself, because rigid cutoffs will eventually result in less genes displayed, so any further analysis is performed on the entire input dataset. Experiment-specific data normalization procedures are restricted to the preference of the user, depending on whether the user needs to visualize expression differences within different genes or detect differences within samples by specifying non-control samples.

Conclusions

Heatmaps generation within the R shiny App, without any prerequired programming skills of the user. Our shiny tool targets wide population of scientists and students studying genetic expression to provide them the easier visualization and analysis of expression data to investigate genetic datasets using an interactive heatmap generation platform to graphically visualize them.

Availability and requirements

- **Project name:** AFCM HeatMap Generation tool
- **Project home page:**
<https://mohammadtarek.shinyapps.io/afcmheatmap/>
- **Operating system:** Only requires an Internet Connection
- **Programming language:** R, Shiny
- **License:** GNU General Public License version 3.0 (GPL-3.0)
- **Prerequired programming skills:** No

Availability of Source Code files:

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

Competing interests: The authors declare 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. Sonesson 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. 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.
6. Chang W, Cheng J, Allaire JJ, Xie Y, McPherson J, RStudio, jQuery Foundation, jQuery contributors, jQuery UI contributors, Otto M, Thornton J, Bootstrap contributors, Twitter Inc., Farkas A, Jehl S, Petre S, Rowls A, Gandy D, Reavis B, Kowal KM, es5-shim contributors, Ineshin D, Samhuri S, SpryMedia Limited, Fraser J, Gruber J, Sagalaev I, R Core Team. shiny: Web Application Framework for R. 2015. R package version 0.12.2.
7. Vaidyanathan R, Xie Y, Allaire JJ, Cheng J, Russell K. RStudio. htmlwidgets: HTML Widgets for R. 2015. R package version 0.5.
8. Neuwirth E. RColorBrewer: ColorBrewer Palettes. 2014. R package version 1.1-2.
9. Wickham H, Francois R. RStudio. dplyr: A Grammar of Data Manipulation. 2015. R package version 0.4.3.
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. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010; 26:139–40.