Author/s:
Mullan, KA;Bramberger, LM;Munday, PR;Goncalves, G;Revote, J;Mifsud, NA;Illing, PT;Anderson, A;Kwan, P;Purcell, AW;Li, C

Title:
ggVolcanoR: A Shiny app for customizable visualization of differential expression datasets

Date:
2021-01-01

Citation:
Mullan, K. A., Bramberger, L. M., Munday, P. R., Goncalves, G., Revote, J., Mifsud, N. A., Illing, P. T., Anderson, A., Kwan, P., Purcell, A. W. & Li, C. (2021). ggVolcanoR: A Shiny app for customizable visualization of differential expression datasets. Computational and Structural Biotechnology Journal, 19, pp.5735-5740. https://doi.org/10.1016/j.csbj.2021.10.020.

Persistent Link:
https://hdl.handle.net/11343/296633

Computational and Structural Biotechnology Journal 19 (2021) 5735–5740

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Communications

ggVolcanoR: A Shiny app for customizable visualization of differential expression datasets

Kerry A. Mullan [a,*], Liesl M. Bramberger [a], Prithvi Raj Munday [a], Gabriel Goncalves [a], Jerico Revote [b], Nicole A. Mifsud [a], Patricia T. Illing [a], Alison Anderson [c,d], Patrick Kwan [c,d,e], Anthony W. Purcell [a], Chen Li [a,*]

[a] Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Melbourne, VIC 3800, Australia
[b] Monash eResearch Centre, Monash University, Melbourne, VIC 3800, Australia
[c] Department of Neuroscience, Central Clinical School, Monash University, Melbourne, Victoria, Australia
[d] Departments of Medicine and Neurology, University of Melbourne, Royal Melbourne Hospital, Melbourne, Victoria, Australia
[e] Department of Neurology, Alfred Health, Melbourne, Victoria, Australia

ARTICLE INFO

Article history:
Received 11 August 2021
Received in revised form 12 October 2021
Accepted 12 October 2021
Available online 13 October 2021

Keywords:
Volcano plot
Correlation plot
Heatmap
Upset plot
Differential expression data
Transcriptomics
Proteomics
Data visualization

ABSTRACT

Volcano and other analytical plots (e.g., correlation plots, upset plots, and heatmaps) serve as important data visualization methods for transcriptomic and proteomic analyses. Customizable generation of these plots is fundamentally important for a better understanding of dysregulated expression data and is therefore instrumental for the ensuing pathway analysis and biomarker identification. Here, we present an R-based Shiny application, termed ggVolcanoR, to allow for customizable generation and visualization of volcano plots, correlation plots, upset plots, and heatmaps for differential expression datasets, via a user-friendly interactive interface in both local executable version and web-based application without requiring programming expertise. Compared to currently existing packages, ggVolcanoR offers more practical options to optimize the generation of publication-quality volcano and other analytical plots for analyzing and comparing dysregulated genes/proteins across multiple differential expression datasets. In addition, ggVolcanoR provides an option to download the customized list of the filtered dysregulated expression data, which can be directly used as input for downstream pathway analysis. The source code of ggVolcanoR is available at https://github.com/KerryAM-R/ggVolcanoR and the webserver of ggVolcanoR 1.0 has been deployed and is freely available for academic purposes at https://ggvolcanor.erc.monash.edu/.

1. Introduction

Investigation of gene expression perturbations plays a critical role in understanding complex diseases [1–4]. With the improved accessibility, transcriptomic and proteomic techniques have been widely applied in biomedical studies to advance molecular profiling in a variety of disease samples. These experiments consequently generate a large volume of data describing dysregulated genes and proteins [5–8]. Visualizing data from differential expression datasets can be challenging, due to their complexity. In addition, comparing differential expression datasets, for example, proteomic and transcriptomic data, serves as an important approach to identify critical functions and pathways [5,9–13]. Conventionally, these transcriptomic and proteomic datasets are usually presented as scatter plots, commonly known as volcano plots. To facilitate such analysis, there have been a number of tools developed for processing transcriptomic and multi-omic datasets, however limitations exist. Some tools lack fully functional webservers or need to be manually setup (e.g., EnhancedVolcano [14], ShinyOmics [15], and WIlsON [16]), which restricts their use. While VolcaNoseR [17] has many customizable features in the R shiny format, it lacks multiple dataset comparisons. In addition, EnhancedVolcano [14] is specific to RNA-seq data that can be processed by the 'DESeq2' R package [18]. Therefore, there is an urgent need to create a computational tool with the following important features: (1) easy-to-use for non-coders, (2) fully functional webserver, (3) highly customizable analytical plots, (4) straightforward comparison of two or more datasets utilizing either correlation analysis and/or heatmap/upset plots, and (5) not be restricted to a certain differential expression pipeline.

* Corresponding authors.
E-mail addresses: Kerry.Mullan@monash.edu (K.A. Mullan), Chen.Li@monash.edu (C. Li).

https://doi.org/10.1016/j.csbj.2021.10.020

Here we present our Shiny app, termed 'ggVolcanoR', for publication-quality figure generation utilizing transcriptomic and proteomic datasets for non-coders. In comparison to currently well-established tools for analyzing differential expression and omic datasets [19–23], ggVolcanoR specifically focuses on facilitating the generation of highly customizable and publication-ready plots and the comparison of multiple transcriptomic and/or proteomic differential expression datasets. The plot generation can be customized with a variety of user-parameter guided features including the selection of font, image quality, and labels. Additionally, there is an interactive table with links provided by ggVolcanoR to well-established third-party biological databases for the gene/protein mapping in uploaded datasets. We anticipate that ggVolcanoR can facilitate the visualization and comparison of differential expression datasets (i.e. transcriptomic and proteomic), thereby informing the discovery of potential biomarkers in complex diseases.

2. Material and methods

ggVolcanoR is an R-based Shiny application constructed using various third-party R packages, including tidyverse [24], ggplot2 [25], ggrepel [26], shiny [27], shinyBS [28], gridExtra [29], DT [30], plyr [31], dplyr [32], and re-shape2 [33], colourpicker [34], ComplexHeatmap [35] and circlize [36]. Two versions of 'ggVolcanoR' are available for users to run in their local systems and on our webserver, respectively. The Shiny R platform was deployed on the webserver to host the web application of ggVolcanoR. The webserver is managed by the Apache framework as reverse proxy which prevents the underlying R application from being accessed directly by the users and secures the webserver with SSL encryption. The webserver of ggVolcanoR resides on the Nectar (the National eResearch Collaboration Tools and Resources project) service equipped with 4-core CPU and 16 GBs of RAM and managed by the Monash eResearch Centre. To showcase the functionality and usability of ggVolcanoR, we used the publicly available datasets by Goncalves *et al.* [5], which contain both transcriptomic and proteomic results of differential expressed genes and proteins. The datasets were generated from a triple negative breast cancer (TNBC) cell line to determine the effect of Interferon γ (IFNγ) treatment. These datasets are available in our 'test-data folder' on GitHub (i.e., 'Proteomic data.csv' as input).

3. Results

3.1. Overall features of ggVolcanoR

Several major types of plots can be generated in a customizable manner by ggVolcanoR, including volcano plots, correlation plots, heatmaps, and upset plots. The volcano plot is mainly for single-group analysis, where only one differential expression dataset (either transcriptomics or proteomics) is used. A highly customized volcano plot and two interactive tables providing the information of the dataset will be generated. The correlation plot, on the other hand, is generated to demonstrate the agreement between two differential expression datasets (either transcriptomics or proteomics). While the heatmap and upset plot are generated to compare the $\log_2$ fold change (logFC) of two or more differential expression datasets. For volcano and correlation plots in ggVolcanoR, the input file is simple to organize, consisting of a list containing gene/transcript/protein IDs, logFC and *p*-value (Pvalue) from the differential expression data (either .txt or .csv format). To generate heatmaps and upset plots, two additional columns, 'group' and 'group.direction' need to be added. The 'group' column specifies the experimental/condition information (e.g., proteomics or transcriptomics) and the 'group.direction' column specifies the experimental information and the direction of fold change (e.g., 'proteomics.up'). Additionally, for both volcano and correlation plots, a variety of options are provided in ggVolcanoR for users to generate publication-quality plots, such as font/color choices, and adjustable Pvalue and logFC cut-offs. There is also an option for users to download the pre-set styles with a variety of parameters. All plots can be exported in PDF or PNG formats with customizable size and resolution (for PNG only).

3.2. Using ggVolcanoR to generate publication-quality volcano plots

A collection of parameters, such as data selection, font, data point/label, can be configured using the different panels demonstrated in Fig. 1**A**. The outputs in the 'Volcano plot' tab are organized in four subtabs, including 'Volcano plot' (Fig. 1**B**), 'Volcano plot (selected colours)' (Fig. 1**C**), 'Table with links' (Fig. 1**D**), and 'Summary table' (Fig. 1**E**). There are 5 customizable styles (i.e., 'default', 'all.datapoints', 'up.ID', 'down.ID', and 'selected.ID') that will alter the outputs in both 'Volcano plot' and 'Volcano plot (selected colours)' tabs. For example, one might be interested in the top 30 proteins dysregulated after IFNγ treatment, as shown in the 'Volcano plot' tab (Fig. 1**B**). ggVolcanoR also allows users to investigate a selection of IDs of interest by providing 'selected.ID'. Lastly, users can color and highlight specific genes of interest under the 'Volcano plot (selected colours)' tab using the color palettes provided (Fig. 1**C**). Under the 'Table with links' tab (Fig. 1**D**), the collection of dysregulated genes in the generated volcano plot are listed in an interactive table with highlighted logFC and Pvalue, together with the links to third-party gene and protein databases, including GeneCards [37], the Human Protein Atlas (http://www.proteinatlas.org) [38] and UniProt [39]. Users can select if their IDs are either gene symbol, Ensembl [40] IDs or UniProt accessions for linking purposes. Users can also define if they are investigating a human or non-human species (limited to UniProt database). The 'Summary table' tab (Fig. 1**E**) offers the summary statistics regarding the numbers of down- and up-regulated genes. We have also added a label function for creating a file for the heatmap/upset plots. After downloading the 'Filtered table', users can upload this list of significant dysregulated expression IDs to g:Profiler [41] for pathway analysis.

3.3. Correlating two differential expression datasets using ggVolcanoR

As technology costs reduce, there is now capacity to obtain multi-omics datasets from immunopeptidome, proteome and transcriptome sources to interrogate disease complexity. This generates a need to facilitate comparisons not only of replicated data sets but also of distinct data types as performed by Goncalves *et al.* [5], where proteomic and transcriptomic datasets were compared. The 'Correlation plot' tab has the capacity to compare relationships between the logFC of two differential expression datasets, e.g., transcriptomic and proteomic datasets (Fig. 2). The graph features can be modified based on the choices in the side panel (Fig. 2**A**) or above the graph (Fig. 2**B**). Users can choose to label the most significant IDs based on the logFC or Pvalue of either dataset. If desired, the linear model regression line with the 95% confidence interval can be added (Fig. 2**B**). The 'Correlation Pearson statistics' tab includes the Pearson's correlation for the overall dataset as well as the overlap of the significant IDs for the positive correlation or negative correlation based on the logFC direction (Fig. 2**C**). The 'Correlation table' (Fig. 2**D**) displays the combined two datasets, with the option to export the IDs that are significant in both datasets. The 'Bar graph' (Fig. 2**E**) demonstrates the logFC values and users can visualize these values for: 'all', 'up', 'down', 'same' or 'opposite'. There is also capacity to display a user defined

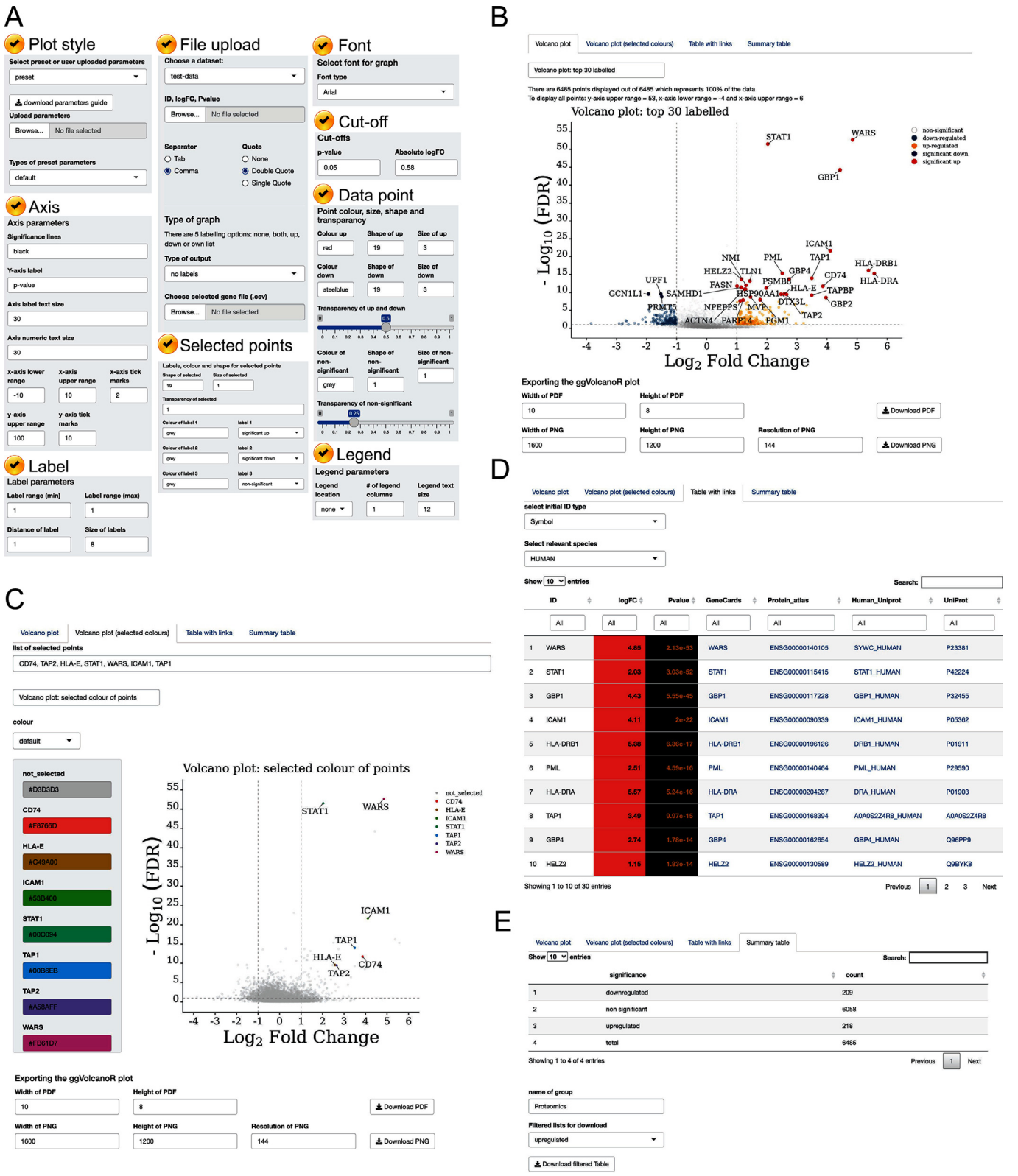

Fig. 1. Using ggVolcanoR to generate volcano plots. (A) Nine panels for data uploading and parameter configuration; (B) an example of the generated volcano plot using the dataset by Goncalves *et al.* [5]; (C) an example demonstrated seven selected genes of interest in the volcano plot; (D) the 'Table with links' tab for plotted dysregulated genes; and (E) the statistical information of different types of genes in the 'Summary table' tab.

list of IDs of interest (type of output: "own list"; which is uploaded using selected gene file) regardless of significance threshold. All 'Bar graph' features are modifiable at the top of the graph (Fig. 2**E**). Here the datasets by Goncalves *et al.* [5] have been used to showcase the function of ggVolcanoR for comparing two differential expression datasets. This analysis identified 6,488 IDs were common to the proteomic and transcriptomic datasets with a Pearson's correlation of 0.275 (*p*-value = 1.16e-112).

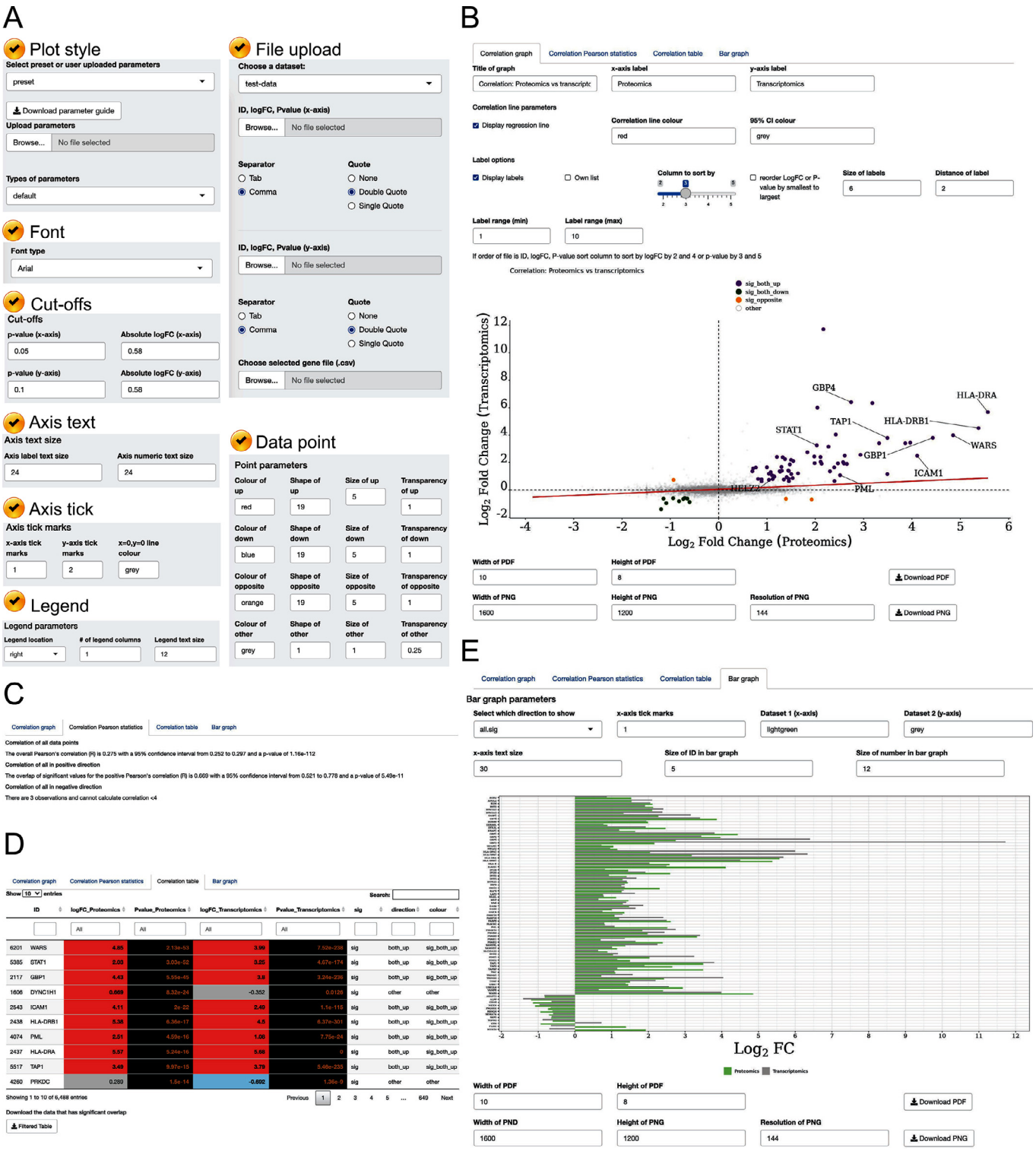

Fig. 2. Using ggVolcanoR to analyze and compare dysregulated genes and proteins across transcriptomic and proteomic datasets. (A) Eight panels for parameter configuration to generate the correlation plot; (B) the correlation plot generated using the transcriptomic and proteomic datasets by Goncalves *et al.* [5]; (C) the statistics of the Pearson's correlation; (D) the 'Correlation table' tab representing the overlap of the two datasets; and (E) the bar chart illustrating the agreement of gene/protein dysregulation in the 'Correlation table' tab across the transcriptomic and proteomic.

3.4. Comparing multiple differential expression datasets using heatmaps and upset plots

Heatmaps and upset plots can also be generated by ggVolcanoR to compare dysregulated genes/proteins across multiple datasets and conditions. The heatmap (Fig. 3**A**) and upset plots (Fig. 3**B**) can compare the logFC of two or more datasets. There is no limit to the number of groups for the heatmap, but the upset plot is limited to 31 groups due to the capacity of the package. Users can create the file for this section by downloading the filtered file from the 'Volcano plot' →'summary table' tab and copying these into one file. Here we showcase the functionality using Goncalves *et al.* [5]. For example, the upset plot has clearly demonstrated that nine genes/proteins (green rectangle in Fig. 3**B**) were down regulated

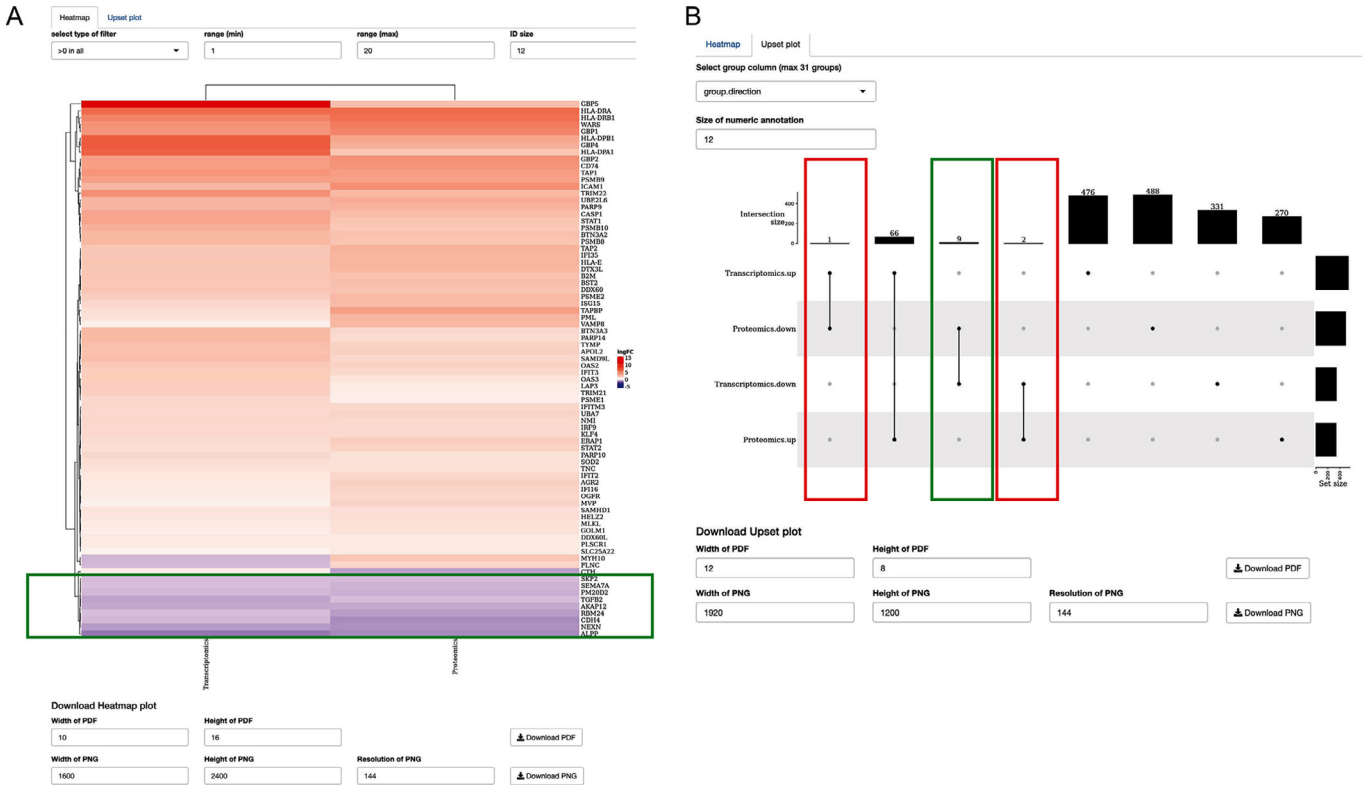

Fig. 3. Using heatmap and upset plot to compare dysregulated genes in both transcriptomic and proteomic datasets. (A) A heatmap by ggVolcanoR demonstrating dysregulated proteins/genes; and (B) an upset plot showing agreement of dysregulated proteins/genes across the transcriptomic and proteomic datasets.

both in proteomic and transcriptomic experiments. The results can also be reflected in the heatmap generated by ggVolcanoR (Fig. 3**A**). In contrast, three genes/proteins (red rectangles in Fig. 3**B**) in total show inconsistent regulation across proteomic and transcriptomic experiments.

4. Discussion

ggVolcanoR, an R-based package using the Shiny user interface, was developed for customizable generation of high-quality volcano plots, correlation graphs, heatmaps and upset plots using differential expression datasets. Unlike previous tools [14,17], ggVolcanoR can perform an in-depth comparison of two datasets via correlation analysis, which is essential for replication comparisons or comparing two differential expression datasets. Users can also use individual filters for each dataset, as different cut-offs may be required (e.g., false discovery rate = 0.1 and q-value = 0.05) and different logFC values. The 'Correlation plot' tab has the capacity to compare relationships between the logFC of two differential expression datasets, e.g., transcriptomic and/or proteomic datasets (Fig. 2). Note, it is important that the nomenclature for gene/protein identifiers is consistent across the compared datasets to avoid misalignment/duplication of IDs. Furthermore, the volcano and correlation plots also have the added features of preset customization, where users only need to make minimal adjustments or can upload their own styles for future use. ggVolcanoR also provides a heatmap and upset plot to compare multiple differential expression datasets based on the logFC rather than the pre-processed raw data which was presented in ShinyOmics [15].

In addition to multiple analysis plots, there have been other improvements to the presentation of the volcano plot. For instance, users can download both PDF and PNG files of the generated plots. Users can also select the type of font which may be needed for specific journals or publication types. The size, transparency, color, and shape of the data points (e.g., selected, up-regulated, down-regulated or non-significant) can be further customized in ggVolcanoR, which is limited in VolcaNoseR [17]. ggVolcanoR axis labels give users the freedom to specify the type of p-value test used, such as FDR for RNA-seq analysis or q-value for proteomics, which cannot be customized in other tools like OmicLoupe [42]. Unlike other tools, ggVolcanoR generates a table with links to third party databases for gene mapping, providing additional utility and connectivity of the analyzed data. One of the limitations is the 'table with links' function may not always match available databases which could be due to different aliases from our specified databases. Another advantage of ggVolcanoR is that users are able to download the filtered list based on the user defined filters that can be used directly for downstream pathway analysis in tools such as g:Profiler [41] or use this for the heatmap/upset plot sections.

In this paper, we have highlighted how ggVolcanoR, as both a stand-alone tool and a webserver version, will assist researchers in the analysis of dysregulated genes and proteins in broad research areas with minimal training and coding expertise required to generate informative figures. We anticipate that ggVolcanoR will serve as a routinely used application for visualizing and analyzing differential expression omic datasets in laboratories, thereby contributing to the identification of significant biomarkers.

Availability

The webserver of ggVolcanoR is freely available for academic purposes at https://ggvolcanor.erc.monash.edu/. The source code of ggVolcanoR and the testing data used to demonstrate the usability and functionalities of ggVolcanoR in this study are accessible at https://github.com/KerryAM-R/ggVolcanoR.

Authors contribution

K.A.M. and C.L. conceived and designed the project. K.A.M. designed and implemented the local version of ggVolcanoR. P.R.

M. implemented the webserver of ggVolcanoR, which is managed by J.R. L.M.B., G.G., N.A.M., P.T.I., A.A., P.K., and A.W.P. provided critical suggestions and comments on the functionality and usability of the application. K.A.M. and C.L. drafted the manuscript, which has been revised and approved by all the other co-authors.

CRediT authorship contribution statement

Kerry A. Mullan: Conceptualization, Methodology, Software, Validation, Writing – original draft. **Liesl M. Bramberger:** Validation, Writing – review & editing. **Prithvi Raj Munday:** Resources, Validation, Writing – review & editing. **Gabriel Goncalves:** Validation, Writing – review & editing. **Jerico Revote:** Resources, Validation, Writing – review & editing. **Nicole A. Mifsud:** Validation, Writing – review & editing. **Patricia T. Illing:** Validation, Writing – review & editing. **Alison Anderson:** Validation, Writing – review & editing. **Patrick Kwan:** Validation, Writing – review & editing. **Anthony W. Purcell:** Validation, Writing – review & editing, Funding acquisition. **Chen Li:** Conceptualization, Methodology, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work has been financially supported by a National Health and Medicine Research Council of Australia (NHMRC) Project Grant to A.W.P. (1122099). K.A.M. is supported by an Australian Government Research Training Program (RTP) Scholarship. C.L. is currently supported by an NHMRC CJ Martin Early Career Research Fellowship (1143366). P.K. is supported by a Medical Research Future Fund Practitioner Fellowship (MRF1136427). A.W.P. is supported by an NHMRC Principal Research Fellowship (1137739).

References

[1] Kim YA, Wuchty S, Przytycka TM. Identifying causal genes and dysregulated pathways in complex diseases. PLoS Comput Biol 2011;7:e1001095.
[2] Schadt EE. Molecular networks as sensors and drivers of common human diseases. Nature 2009;461:218–23.
[3] Venkat S, Alahmari AA, Feigin ME. Drivers of Gene Expression Dysregulation in Pancreatic Cancer. Trends Cancer 2021;7:594–605.
[4] Wang Y, Chakravarty P, Ranes M, Kelly G, Brooks PJ, Neilan E, et al. Dysregulation of gene expression as a cause of Cockayne syndrome neurological disease. Proc Natl Acad Sci U S A 2014;111:14454–9.
[5] Goncalves G, Mullan KA, Duscharla D, Ayala R, Croft NP, Faridi P, et al. IFNgamma Modulates the Immunopeptidome of Triple Negative Breast Cancer Cells by Enhancing and Diversifying Antigen Processing and Presentation. Front Immunol 2021;12:645770.
[6] Gupta S, Ellis SE, Ashar FN, Moes A, Bader JS, Zhan J, et al. Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. Nat Commun 2014;5:5748.
[7] Hecker M. Blood transcriptome profiling captures dysregulated pathways and response to treatment in neuroimmunological disease. EBioMedicine 2019;49:2–3.
[8] Lapek Jr JD, Greninger P, Morris R, Amzallag A, Pruteanu-Malinici I, Benes CH, et al. Detection of dysregulated protein-association networks by high-throughput proteomics predicts cancer vulnerabilities. Nat Biotechnol 2017;35:983–9.
[9] Hegde PS, White IR, Debouck C. Interplay of transcriptomics and proteomics. Curr Opin Biotechnol 2003;14:647–51.
[10] Pavelka N, Fournier ML, Swanson SK, Pelizzola M, Ricciardi-Castagnoli P, Florens L, et al. Statistical similarities between transcriptomics and quantitative shotgun proteomics data. Mol Cell Proteomics 2008;7:631–44.
[11] Schenk S, Bannister SC, Sedlazeck FJ, Anrather D, Minh BQ, Bileck A, et al. Combined transcriptome and proteome profiling reveals specific molecular brain signatures for sex, maturation and circalunar clock phase. Elife 2019;8.
[12] Wang D, Eraslan B, Wieland T, Hallstrom B, Hopf T, Zolg DP, et al. A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. Mol Syst Biol 2019;15:e8503.
[13] Zhang W, Ambikan AT, Sperk M, van Domselaar R, Nowak P, Noyan K, et al. Transcriptomics and Targeted Proteomics Analysis to Gain Insights Into the Immune-control Mechanisms of HIV-1 Infected Elite Controllers. EBioMedicine 2018;27:40–50.
[14] Blighe K, Rana S, Lewis M (2021), 'EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling'.
[15] Surujon D, van Opijnen T. ShinyOmics: collaborative exploration of omics-data. BMC Bioinf 2020;21:22.
[16] Schultheis H, Kuenne C, Preussner J, Wiegandt R, Fust A, Bentsen M, et al. WIIsON: Web-based Interactive Omics VisualizatioN. Bioinformatics 2019;35:1055–7.
[17] Goedhart J, Luijsterburg MS. VolcaNoseR is a web app for creating, exploring, labeling and sharing volcano plots. Sci Rep 2020;10:20560.
[18] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
[19] Harshbarger J, Kratz A, Carninci P. DEIVA: a web application for interactive visual analysis of differential gene expression profiles. BMC Genomics 2017;18:47.
[20] Kallio MA, Tuimala JT, Hupponen T, Klemela P, Gentile M, Scheinin I, et al. Chipster: user-friendly analysis software for microarray and other high-throughput data. BMC Genomics 2011;12:507.
[21] Li Y, Andrade J. DEApp: an interactive web interface for differential expression analysis of next generation sequence data. Source Code Biol Med 2017;12:2.
[22] Wang YE, Kutnetsov L, Partensky A, Farid J, Quackenbush J. WebMeV: A Cloud Platform for Analyzing and Visualizing Cancer Genomic Data. Cancer Res 2017;77:e11–4.
[23] Younesy H, Moller T, Lorincz MC, Karimi MM, Jones SJ. VisRseq: R-based visual framework for analysis of sequencing data. BMC Bioinf 2015;16(Suppl 11):S2.
[24] Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the tidyverse. Journal of Open Source Software;4.
[25] Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2016.
[26] Slowikowski K, Schep A, Hughes S, Dang TK, Lukauskas S, Irisson J-O, et al. (2021), 'ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2''.
[27] Chang W, Cheng J, Allaire JJ, Sievert C, Schloerke B, Xie Y, et al. (2021), 'shiny: Web Application Framework for R'.
[28] Bailey E (2015), 'shinyBS: Twitter Bootstrap Components for Shiny'.
[29] Auguie B, Antonov A. gridExtra: Miscellaneous Functions for "Grid. Graphics"; 2017.
[30] Xie Y, Cheng J, Tan X, Allaire JJ, Girlich M, Ellis GF, et al. 'DT: A Wrapper of the JavaScript Library 'DataTables''.
[31] Wickham H. The Split-Apply-Combine Strategy for Data Analysis. J Stat Softw 2011;40:29.
[32] Wickham H, François R, Henry L, Müller K, RStudio (2021), 'dplyr: A Grammar of Data Manipulation'.
[33] Wickham H. Reshaping Data with the reshape Package. J Stat Softw 2007;21:20.
[34] Attali D. Colourpicker: A colour picker tool for shiny and for selecting colours in plots. R package version 2017;1.
[35] Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016;32:2847–9.
[36] Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize implements and enhances circular visualization in R. Bioinformatics 2014;30:2811–2.
[37] Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Curr Protoc Bioinformatics 2016;54:1 30 1-1 3.
[38] Thul PJ, Akesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, et al. A subcellular map of the human proteome. Science 2017;356.
[39] UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res 2021;49:D480–9.
[40] Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, et al. Ensembl 2021. Nucleic Acids Res 2021;49:D884–91.
[41] Reimand J, Kull M, Peterson H, Hansen J, Vilo J. g:Profiler–a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Res 2007;35:W193–200.
[42] Willforss J, Siino V, Levander F. OmicLoupe: facilitating biological discovery by interactive exploration of multiple omic datasets and statistical comparisons. BMC Bioinf 2021;22:107.