iMeta WILEY

ggVennDiagram: Intuitive Venn diagram software extended

To the Editor,

Venn diagrams are widely used to illustrate relationships among multiple datasets. As one of the most popular data visualization platforms, the R programming language stands out with numerous packages, including VennDiagram, venn, and RVenn, designed for plotting Venn diagrams [1–3]. Building upon the foundations laid by these existing tools, we developed ggVennDiagram [4]. Due to the support for the grammar of graphics, precise region/subset filling, and some other easy-to-use features, the ggVennDiagram package has emerged as one of the most favored tools in this domain.

Over the past 2 years, ggVennDiagram has undergone continuous refinement, benefiting from a number of functional optimizations, and earned more than 100 literature citations. In this study, we release a milestone version 1.5 that includes the following new features.

EASIER FOR NEW INSTALLATION THAN BEFORE

Before version 1.5, ggVennDiagram had a large package dependency tree, which contains more than 90 packages. Among them, the sf package emerged as the heaviest package. It not only has the largest package size but also has several system requirements, which are often not installed by most users. It was a primary cause for installation failures, as highlighted in GitHub Issues (https://github.com/gaospecial/ggVennDiagram/issues). However, sf is essential for shape generation in ggVennDiagram. Removing it directly is not feasible. Taking this into consideration, we relocated the shape generation functions to a new package, namely shapeMageR. Furthermore, several other dependencies, including RVenn, purrr, magrittr, and plotly were also removed (see Table S1). Consequently, the total file size of ggVennDiagram, together with its package dependencies, has been substantially reduced from 210 MB (1.1.0) to 36 MB in the current version, as been analyzed by the pak package manager [5] (see File S1). This makes the installation of ggVennDiagram much easier than before.

NATIVE SUPPORT OF UPSET PLOT

Venn Diagrams and Upset plot are two types of visualizing methods that are generally used under the same scenario. Plotting Upset plot is supported by the UpsetR package in the R platform [6]. However, we believe it would be more convenient to combine Upset method together with the Venn Diagram. Therefore, we introduced native support for Upset plots in ggVennDiagram. The implementation of Upset plot is inspired by the aplot package [7], which enables automatic axis alignments between subplots. We first create subplots of Upset plot, and subsequently combine the three components together. Figure 1 shows the Venn diagram and Upset plot generated by ggVennDiagram. Please note that the shapes for this five set diagram, as well as those for six and seven sets, are imported from the original package "venn," authored by Adrian Dușa [3].

FULL FUNCTIONAL VENN CALCULATOR

The S4 Venn class were implemented to store structured sets data such as the set members and set names. A series of methods were implemented for enabling set calculation of the Venn object. As shown in Figure 1C, we can get the overlapping, different and specific members of one or more sets by using these methods. Thus, these functions provide a Venn Calculator for further in-depth analysis of sets data.

(A)

(B)

(C)

```
# Suppose we have a four-set Venn object
venn_y # names of sets are A/B/C/D
#> An object of class 'Venn':
#>     Slots: sets, names;
#>     No. Sets: 4    SetNames: A, B, C, D.

# Get the overlapping members of two or more sets
overlap(venn_y, 1:2) # members in both the first two sets
overlap(venn_y) # members in all the sets

# Get the different members between sets and set unions
discern(venn_y, 1)    # members in set 1, but not in all the resting sets
discern(venn_y, c("A","B"), 3) # members in set A & B, but not in set 3

# Get the specific members in one or more sets
discern_overlap(venn_y, 1)    # specific items in set 1
discern_overlap(venn_y, 1:2)    # specific items in set 1 and set 2
```
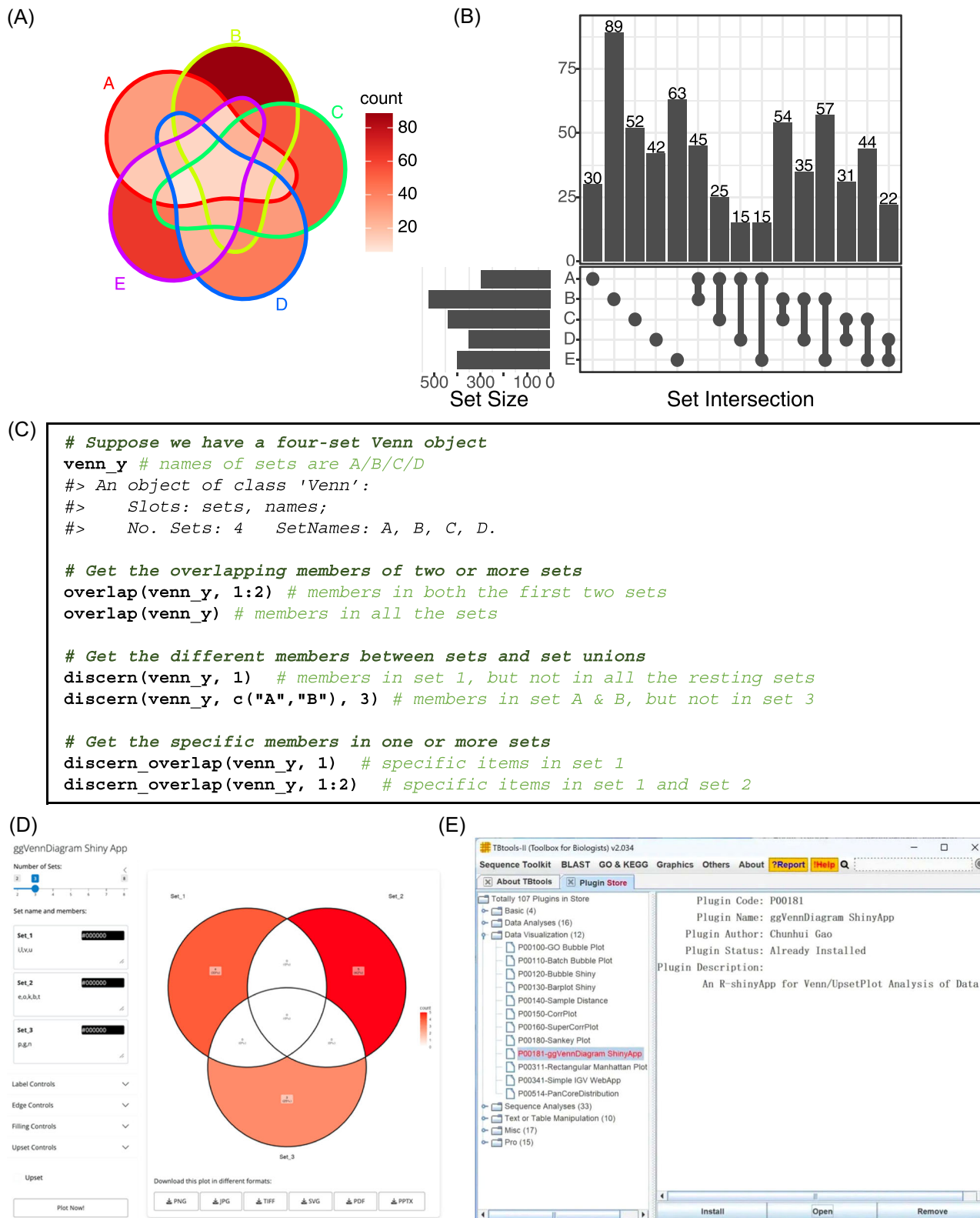
(D)

(E)

FIGURE 1 Features of ggVennDiagram. Venn diagram (A) and upset plot (B) with five sets input. (A) The numbers of subset members can be distinguished by light and deep red color fillings. (B) Top plot shows the size of intersections; left plot shows the size of five sets; main plot shows what these intersections come from. Intersections and sets were sorted alphabetically. (C) Example of Venn calculator. User can easily access the subsets with Venn object and generic methods. A reproducible example is provided in File S1. Screenshot of the ggVennDiagram Shiny App (D) and ggVennDiagram plugin in TBtools (E).

OFFICIAL SHINY APP AND TBTOOLS PLUGIN

Shiny is a web application framework that allows developer to create interactive web-based data visualization tools. The ggVennDiagram Shiny app was developed and deployed to *shinyapps.io*, a cloud-based service provided by Posit/RStudio. Most of the parameters required for plotting a Venn Diagram are provided on the sidebar. Specifically, the exported vector-format figures (svg, pdf, and pptx) can be further edited in place with handy tools, such as Adobe Illustrator, Microsoft PowerPoint, and so on (Figure 1D). TBtools is a bioinformatics software developed by Dr. Chengjie Chen [8]. It serves as a comprehensive toolset for completing multiple popular bioinformatics tasks. The plugin store provided in TBtools-II makes it possible to integrate the Shiny app (Figure 1E) into the platform [9]. Those new features produced an interactive user interface, allowing light users who have limited coding abilities to use and generate high-quality plots with ggVennDiagram.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in GitHub at https://github.com/gaospecial/ggVennDiagram. The ggVennDiagram R package is open source and freely available on CRAN (https://cran.r-project.org/package=ggVennDiagram) and GitHub (https://github.com/gaospecial/ggVennDiagram). The ggVennDiagram Shiny app can be accessed at Shinyapps.io (https://bio-spring.shinyapps.io/ggVennDiagram). The TBtools plugin can be accessed through its plugin store. Supplementary materials (figures, tables, scripts, graphical abstract, slides, videos, Chinese translated version, and update materials) may be found in the online DOI or iMeta Science http://www.imeta.science/.

Chun-Hui Gao[1,2,3]
Chengjie Chen[4]
Turgut Akyol[5]
Adrian Dusa[6]
Guangchuang Yu[7]
Bin Cao[3]
Peng Cai[1,2,8]

[1]*National Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan, China*
[2]*College of Resources and Environment, Huazhong Agricultural University, Wuhan, China*
[3]*School of Civil and Environmental Engineering and Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore, Singapore*
[4]*State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, College of Horticulture, South China Agricultural University, Guangzhou, China*
[5]*Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark*
[6]*Department of Sociology, University of Bucharest, Bucharest, Romania*
[7]*Department of Bioinformatics, School of Basic Medical Sciences, Southern Medical University, Guangzhou, China*
[8]*Hubei Key Laboratory of Soil Environment and Pollution Remediation, Huazhong Agricultural University, Wuhan, China*

Correspondence

Peng Cai, National Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan 430070, China.
Email: cp@mail.hzau.edu.cn

ORCID

Chun-Hui Gao http://orcid.org/0000-0002-1445-7939
Chengjie Chen http://orcid.org/0000-0001-5964-604X
Turgut Akyol http://orcid.org/0000-0003-0897-7716
Adrian Dusa http://orcid.org/0000-0002-3525-9253
Guangchuang Yu http://orcid.org/0000-0002-6485-8781
Bin Cao http://orcid.org/0000-0002-9462-496X
Peng Cai http://orcid.org/0000-0002-3103-1542

REFERENCES

1. Chen, Hanbo, and Paul C. Boutros. 2011. "VennDiagram: A Package for the Generation of Highly-Customizable Venn and Euler Diagrams in R." *BMC Bioinformatics* 12(1): 35. https://doi.org/10.1186/1471-2105-12-35
2. Akyol, Turgut Yigit. 2019. "RVenn: Set Operations for Many Sets." Version 1.1.0. https://CRAN.R-project.org/package=RVenn
3. Dusa, Adrian. 2024. "Venn: Draw Venn Diagrams." Version 1.12. https://CRAN.R-project.org/package=venn
4. Gao, Chun-Hui, Guangchuang Yu, and Peng Cai. 2021. "ggVennDiagram: An Intuitive, Easy-to-Use, and Highly Customizable R Package to Generate Venn Diagram." *Frontiers in Genetics* 12: 1598. https://doi.org/10.3389/fgene.2021.706907
5. Csárdi, Gábor, Jim Hester, Posit Software, P. B. C. Winston Chang, Mango Solutions, Hadley Wickham, et al. 2023. "Pak: Another Approach to Package Installation." Version 0.7.1. https://CRAN.R-project.org/package=pak
6. Conway, Jake R., Alexander Lex, and Nils Gehlenborg. 2017. "UpSetR: An R Package for the Visualization of Intersecting Sets and Their Properties." *Bioinformatics* 33(18): 2938–40. https://doi.org/10.1093/bioinformatics/btx364
7. Yu, Guangchuang, Shuangbin Xu, and Thomas Hackl. 2023. "Aplot: Decorate a 'ggplot' With Associated Information." Version 0.2.2. https://CRAN.R-project.org/package=aplot
8. Chen, Chengjie, Hao Chen, Yi Zhang, Hannah R. Thomas, Margaret H. Frank, Yehua He, and Rui Xia. 2020. "TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data." *Molecular Plant* 13(8): 1194–202. https://doi.org/10.1016/j.molp.2020.06.009
9. Chen, Chengjie, Ya Wu, Jiawei Li, Xiao Wang, Zaohai Zeng, Jing Xu, Yuanlong Liu, et al. 2023. "TBtools-II: A 'One for All, All for One' Bioinformatics Platform for Biological Big-Data Mining." *Molecular Plant* 16(11): 1733–42. https://doi.org/10.1016/j.molp.2023.09.010

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Data S1: Analysis of ggVennDiagram dependency.

Data S2: Example for Venn Calculator.

Table S1. Direct dependencies of ggVennDiagram version 1.1.