

ggpubfigs: colorblind friendly color palettes and ggplot2 graphic system extensions for publication-quality scientific figures

Jacob L. Steenwyk* & Antonis Rokas*

Department of Biological Sciences, Vanderbilt University, Nashville, TN 37235, United States of America

*Correspondence: jacob.steenwyk@vanderbilt.edu and antonis.rokas@vanderbilt.edu

Abstract

Clear and effective figures are central to successfully communicating scientific data. Here, we present ggpubfigs, an R package with colorblind friendly color palettes and extensions of the ggplot2 graphic system, which helps make publication-quality scientific figures from quantitative data. ggpubfigs is an open-source and user-friendly tool that is available from <https://github.com/JLSteenwyk/ggpubfigs>.

Announcement

Scientific figures are graphical representations of scientific data (Rougier et al. 2014). Several tools have been developed to generate scientific figures in numerous computer programming languages, including seaborn (Waskom 2021) and Matplotlib (Matplotlib Org. 2019) in the Python programming language, and lattice (Sarkar 2017) and ggplot2 (Wickham 2009) in the R programming language. These and other data visualization pieces of software have empowered researchers with the ability to generate scientific figures from diverse sources of quantitative data. As a result, methods and standards for effective scientific figures—which we define as accurate, clear, and precise representations of scientific data—have been a topic of rigorous debate that is in part influenced by field, audience, and data type (Moere & Purchase 2011; Lau & Vande Moere 2007; Bertini et al. 2011).

Although certain rules of effective scientific figures are context-dependent and subject to change, some rules are broadly applicable to several disciplines, including in microbiology and the life sciences. These include two rules from Rougier et al.'s article titled *Ten Simple Rules for Better Figures*: “Do Not Trust the Defaults” and “Use Color Effectively” (Rougier et al. 2014). For the first rule, the authors suggest that default plotting parameters (e.g., font size, ticks, etc.) are sufficient to make a scientific figure but insufficient to make the “best” scientific figure; for the second, the authors suggest that color is an important component of human vision and therefore is equally important when making scientific figures. Effective color use can also make scientific figures more accessible. For example, 8% and 0.4% of European Caucasian men and women, respectively, are red-green color deficient (Birch 2012). Thus, effective figure making is also a matter of inclusion.

To facilitate generating effective figure making, we present ggpubfigs, an R package with colorblind friendly color palettes and ggplot2 extensions that facilitates the generation of publication-quality scientific figures for quantitative data (<https://github.com/JLSteenwyk/ggpubfigs>). More specifically, ggpubfigs contains six color palettes that are colorblind friendly and aim to increase accessibility of scientific figures and eight “themes,” which modify 21 parameters of a default ggplot2 figure. To demonstrate how ggpubfigs can improve scientific figures in R, we compare default ggplot2 settings (Figure 1A) to those modified using extensions or colorblind friendly color palettes available in ggpubfigs (Figure 1B-G). Users can create additional modifications to a scientific figure according to their specific needs.

Color palettes can be accessed using the “friendly_pal()” function. For example, friendly_pal(“contrast_three”) will provide users access to an object of class “palette” that contains the hex codes for the “contrast_three” color palette. Color palettes can be converted into a colormap of N colors—which may be useful for plotting data as a heatmap—using the following command: friendly_pal(“contrast_three”, N , type=“continuous”). Themes that modify ggplot2 plots can be appended to the end of ggplot2 plotting command. For example, to use the “simple theme” in ggpubfigs, theme_simple(), an object of class gg theme, should be appended to the end of a ggplot2 plotting command.

We anticipate ggpubfigs will assist researchers in generating effective scientific figures that are accessible to broad audiences including those that are colorblind.

Data Availability

ggpubfigs is freely available under the MIT license and is available for download on GitHub (<https://github.com/JLSteenwyk/ggpubfigs>). The GitHub repository comes complete with installation instructions and tutorials. Installing ggpubfigs is simple and only requires executing one command. Tutorials detail how to use color palettes for qualitative and continuous and discrete quantitative data as well as utilizing ggplot2 theme extensions.

Acknowledgements

J.L.S. and A.R. were supported by the Howard Hughes Medical Institute through the James H. Gilliam Fellowships for Advanced Study program. A.R. was supported by the National Science Foundation (DEB-1442113 and DEB-2110404), the National Institutes of Health/National Institute of Allergy and Infectious Diseases (R56 AI146096), the Guggenheim Foundation, and the Burroughs Wellcome Fund.

References

- Bertini E, Tatu A, Keim D. 2011. Quality Metrics in High-Dimensional Data Visualization: An Overview and Systematization. *IEEE Trans. Vis. Comput. Graph.* 17:2203–2212. doi: 10.1109/TVCG.2011.229.
- Birch J. 2012. Worldwide prevalence of red-green color deficiency. *J. Opt. Soc. Am. A.* 29:313. doi: 10.1364/JOSAA.29.000313.
- Lau A, Vande Moere A. 2007. Towards a Model of Information Aesthetics in Information Visualization. In: 2007 11th International Conference Information Visualization (IV '07). IEEE pp. 87–92. doi: 10.1109/IV.2007.114.
- Matplotlib Org. 2019. Matplotlib: Python plotting. Matplotlib.
- Moere A Vande, Purchase H. 2011. On the role of design in information visualization. *Inf. Vis.* 10:356–371. doi: 10.1177/1473871611415996.
- Rougier NP, Droettboom M, Bourne PE. 2014. Ten Simple Rules for Better Figures. *PLoS Comput. Biol.* 10:e1003833. doi: 10.1371/journal.pcbi.1003833.
- Sarkar D. 2017. Package ‘lattice’: Trellis Graphics for R. R Doc.
- Waskom M. 2021. seaborn: statistical data visualization. *J. Open Source Softw.* 6:3021. doi: 10.21105/joss.03021.
- Wickham H. 2009. *ggplot2*. Springer New York: New York, NY doi: 10.1007/978-0-387-98141-3.

Figure and figure legend

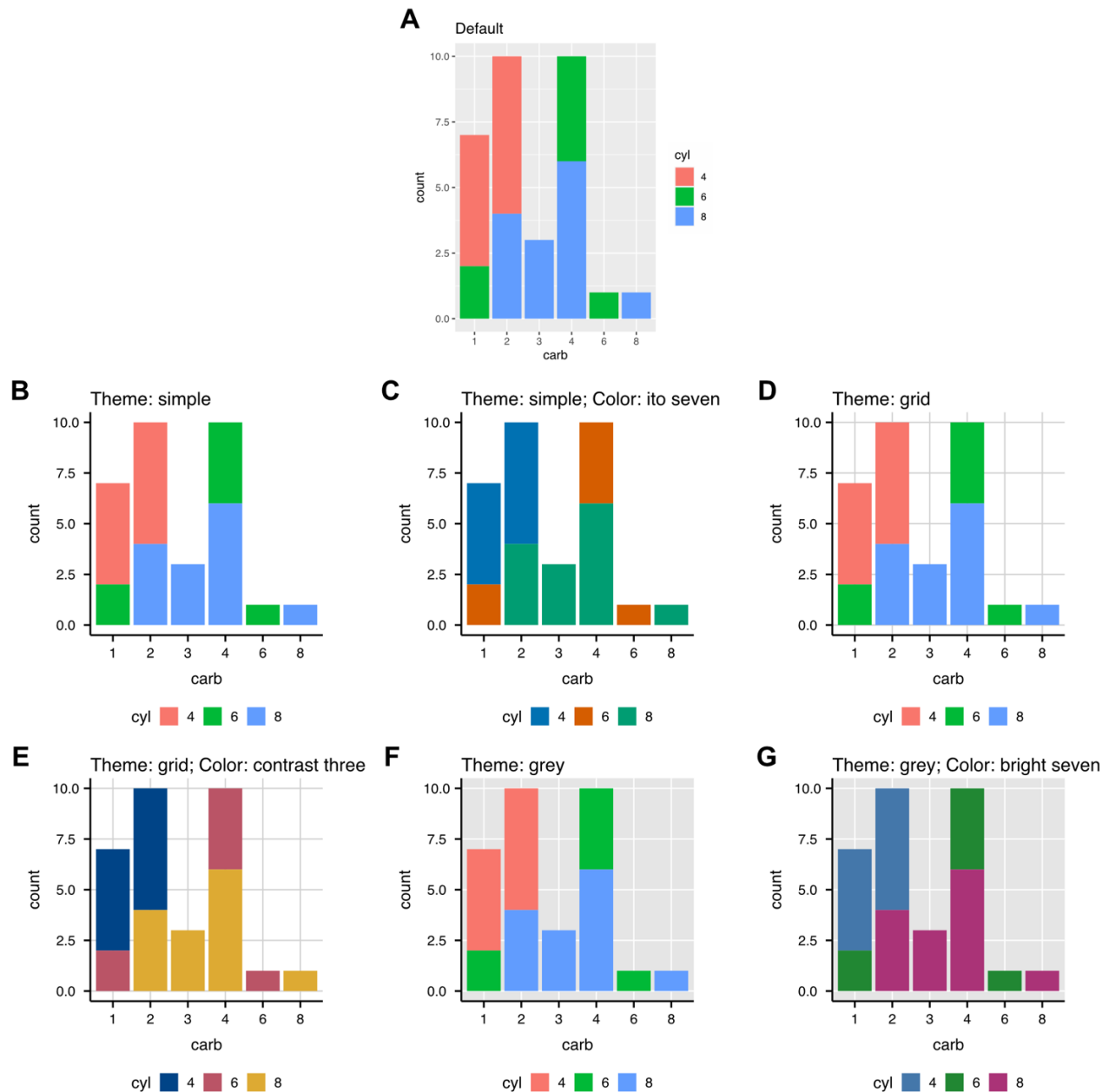


Figure 1. Examples of ggplot2 extensions and color palettes available in ggpubfigs. (A) Default scientific visualization made using ggplot2. Modified scientific figures using (B) the 'simple theme,' (C) the 'simple theme' and the 'ito seven' color palette, (D) the 'grid theme,' (E) the 'grid theme' and the 'contrast three' color palette, (F) the 'grey theme,' and (G) the 'grey theme' and the 'bright seven' color palette. Data are from the mtcars data frame available through the datasets package.