



# *microshades*: An R Package for Improving Color Accessibility and Organization of Microbiome Data

Erin M. Dahl,<sup>a</sup> Emory Neer,<sup>a</sup> Kate R. Bowie,<sup>b</sup> Eric T. Leung,<sup>a</sup>  Lisa Karstens<sup>a</sup>

<sup>a</sup>Department of Medical Informatics and Clinical Epidemiology, Division of Bioinformatics and Computational Biomedicine, Oregon Health & Science University, Portland, Oregon, USA

<sup>b</sup>Knight Cancer Institute, Cancer Early Detection Advanced Research Center, Oregon Health & Science University, Portland, Oregon, USA

**ABSTRACT** When creating figures, it is important to consider that individuals with color vision deficiency (CVD) may not perceive all colors. While there are several CVD-friendly color palettes, they are often insufficient for working with microbiome data. Here, we introduce *microshades*, an R package for creating CVD-accessible microbiome figures.

Color vision deficiency (CVD), commonly known as colorblindness, affects 1 in 12 men and 1 in 200 women, approximately 300 million people worldwide (1). Individuals with CVD do not experience complete loss of color vision but have reduced ability to distinguish between different colors. There are three common types of CVD, i.e., deuteranopia, protanopia, and tritanopia. Individuals with deuteranopia (red-green colorblindness) have difficulty distinguishing between shades of red, green, and yellow. Individuals with protanopia (red colorblindness) have diminished ability to distinguish between colors containing red, whereas those with tritanopia have difficulty distinguishing between blues and yellows.

Despite the large numbers of individuals who experience CVD, many scientific figures rely on color to convey information. To ensure that colors in figures are accessible to all, researchers can use colorblind-friendly color palettes and evaluate accessibility with colorblind simulators. There are several CVD-friendly color palettes available (2–5); however, they are typically limited to 8 to 15 distinct colors or may not be accessible for all forms of CVD. Even with these resources, it can be challenging to apply these schemes to scientific figures. For example, for visualization of microbiome data it is common to represent

**TABLE 1** Colors in the palettes available in *microshades*

microshades_cvd_palettes color hex values					
micro_cvd_purple	#EFB6D6	#E794C1	#CC79A7	#A1527F	#7D3560
micro_cvd_turquoise	#A3E4D7	#48C9B0	#43BA8F	#009E73	#148F77
micro_cvd_blue	#E7F4FF	#BCE1FF	#7DCCFF	#56B4E9	#098BD9
micro_cvd_orange	#FFD5AF	#FCB076	#F09163	#C17754	#9D654C
micro_cvd_green	#DDFFA0	#BDEC6F	#97CE2F	#6D9F06	#4E7705
micro_cvd_gray	#F5F5F5	#D6D6D6	#B7B7B7	#8B8B8B	#616161
microshades_palettes color hex values					
micro_purple	#DADAEB	#BCBDDC	#9E9AC8	#807DBA	#6A51A3
micro_blue	#EFF3FF	#C6DBEF	#9ECAE1	#6BAED6	#4292C6
micro_orange	#FEEDA0	#FEC44F	#FDAE6B	#FE9929	#FF7F00
micro_green	#C7E9C0	#A1D99B	#74C476	#41AB5D	#238B45
micro_brown	#D8C7BE	#CAA995	#B78560	#9E5C00	#7D3200
micro_gray	#D9D9D9	#BDBDBD	#969696	#737373	#525252

**Editor** Irene L. G. Newton, Indiana University, Bloomington

**Copyright** © 2022 Dahl et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

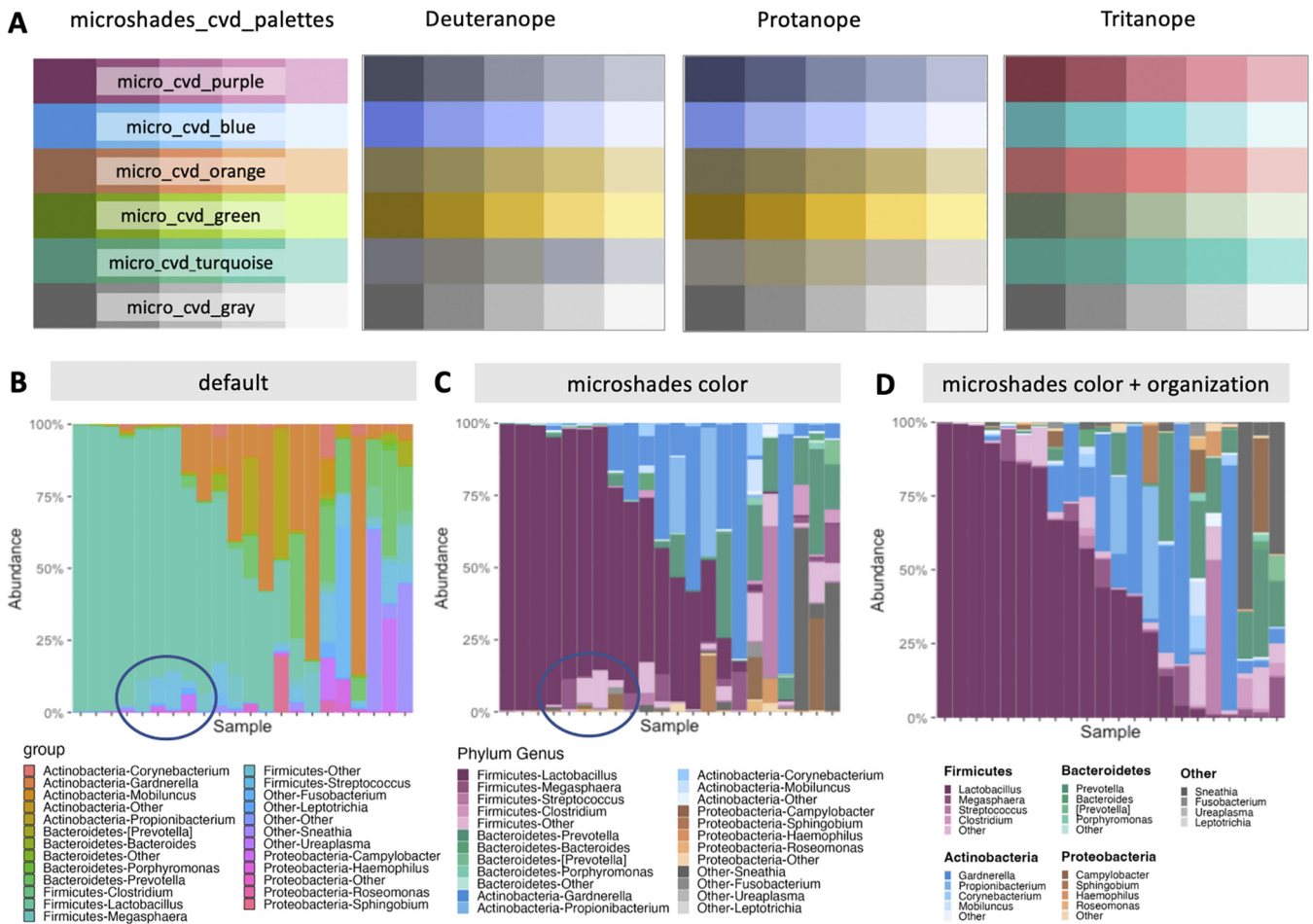
Address correspondence to Lisa Karstens, karstens@ohsu.edu.

The authors declare no conflict of interest.

**Received** 5 August 2022

**Accepted** 12 October 2022

**Published** 31 October 2022



**FIG 1** (A) The *microshades* universal CVD-friendly palette, *microshades\_cvd\_palettes*, with CVD emulations for deuteranope, protanope, and tritanope accessibility. The *cvdemulator* (<http://hclwizard.org:3000/cvdemulator>) was used. (B to D) Example of application to a stacked barplot representing microbiome data (7). In each plot, columns represent samples and the colors represent different bacteria within each sample. The blue circles in B and C highlight the visual improvement that results from using the *microshades* colors. (B) In the default plot, colors are determined by the default ggplot2 color scheme. (C) The *microshades* color plot contains the same data and in the same sample order, with the *microshades\_cvd\_palettes* applied. Colors are organized to represent the phylum-genus classification of bacteria found in the samples. (D) The *microshades* color plus organization plot uses the *microshades\_cvd\_palettes*, *microshades* organizational functions, and *microshades* custom legend.

tens to hundreds of bacterial taxa in one figure, for which the currently available CVD color palettes are insufficient.

To overcome this limitation, we developed the *microshades* R package, which provides custom color shading palettes to improve CVD accessibility and data organization. The *microshades* package includes two color palettes, namely, *microshades\_cvd\_palettes* and *microshades\_palettes* (Table 1). To construct these palettes, hue (type of color), chroma (colorfulness), and luminance (brightness) were adjusted for optimal visual distinction and CVD accessibility. Each color palette contains six hues with five sequential variations of chroma and luminance per hue, for a total of 30 available colors per palette. All shades have been tested with a CVD simulator, *cvdemulator* (3), for deuteranope, protanope, and tritanope accessibility (Fig. 1A).

In addition to color palettes, the *microshades* package contains functions to aid in the complex data visualization common in microbiome studies (Fig. 1B to D), including functions to group data by taxonomic ranking. For example, hues correspond to a high-order taxonomic group (e.g., phylum) and shades of each hue can represent subgroups of the taxonomic group (e.g., genus). Subgroup shading is determined by abundance in the data set, with darker shades indicating the most abundant subgroups, and less abundant subgroups are collapsed into an “other” category. There are also functions to aid in data organization, such as vertical and horizontal sorting of the data and restructuring of the plot legends. Contribution plots can also be created, providing greater insight by displaying boxplots, median barplots, or mean barplots for individual taxa.

In summary, the *microshades* R package is a visualization tool for microbiome researchers. The package contains two CVD-accessible palettes, along with several organization features. The *microshades* package can be used in conjunction with common microbiome R packages, such as phyloseq (6), to enhance microbiome data visualization.

**Data availability.** *microshades*, along with microbiome and nonmicrobiome examples using *microshades*, is available for download at <https://karstenslab.github.io/microshades>.

### ACKNOWLEDGMENTS

This work was supported by the NIH under NIDDK award K01 DK116706 (L.K.), NLM award T15LM007088 (E.M.D. and E.T.L.), and the NIH Common Fund and Office of Scientific Workforce Diversity under three linked awards (grants RL5GM118963, TL4GM118965, and UL1GM118964) administered by the NIGMS (E.N. and L.K.).

The content of the manuscript is solely the authors' responsibility and does not represent the official views of the NIH or any other funding agency.

### REFERENCES

1. Simunovic MP. 2010. Colour vision deficiency. *Eye (Lond)* 24:747–755. <https://doi.org/10.1038/eye.2009.251>.
2. Okabe M, Ito K. 2008. Color Universal Design (CUD): how to make figures and presentations that are friendly to colorblind people. <https://jfly.uni-koeln.de/color>.
3. Zeileis A, Fisher JC, Hornik K, Ihaka R, McWhite CD, Murrell P, Stauffer R, Wilke CO. 2020. colorspace: a toolbox for manipulating and assessing colors and palettes. *J Stat Softw* <https://doi.org/10.18637/jss.v096.i01>.
4. Krzywinski M. 2022. Designing for colorblindness. <http://mkweb.bcgsc.ca/colorblind/palettes.mhtml>.
5. Garnier S, Ross N, Rudis B, Filipovic-Pierucci A, Galili T, Greenwell B, Sievert C, Harris DJ, Chen JJ. 2021. Viridis: colorblind-friendly color maps for R. <https://doi.org/10.5281/zenodo.4679424>.
6. McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. <https://doi.org/10.1371/journal.pone.0061217>.
7. Stansfield J, Smirnova E, Zhao N, Fettweis J, Waldron L, Dozmorov M. 2021. HMP2Data: 16s rRNA sequencing data from the Human Microbiome Project 2. <https://bioconductor.org/packages/release/data/experiment/manuals/HMP2Data/man/HMP2Data.pdf>.