Introduction

Data visualization is essential for the biological mining of the vast amount of information generated by high-throughput technologies. One of the most popular plotting technique in genomics is the heatmap. It is used to display a single quantitative information from a two dimensional table (Wilkinson and Friendly, 2009). For instance, in transcriptomics, rows and columns usually represent genes and samples, respectively, and boxes are colour-coded according to expression signals. Heatmaps however appear intrinsically limited to provide comprehensive representation of the diversity of information now available from single-cell technologies. Accordingly, alternative methods such as dot (or spot) plots are increasingly used.

A dot plot is a modified heatmap where each box in the grid is replaced by a dot. In addition to quantitative values that are dis-played via a colour gradient, the dot size is also used to represent another quantitative information, e.g. the fraction of cells express-ing a given gene within a single-cell RNA-sequencing (scRNA-seq) dataset (Lukassen et al., 2018; Habib et al., 2017; Ordovas-Montanes et al., 2018; Wu et al., 2018).

In recent years, multiple tools have been developed to make dot plots like the rain plot method (Henglin et al., 2019) or the corrplot function of the feature-expression heatmap method (Benno Haar-man et al., 2015; Wei and Simko, 2017). Several programs dedi-cated to scRNAseq analysis (Seurat, scClustViz or cell phonedb) also provide a dot plot function (Innes and Bader, 2019; Stuart et al., 2019; Efremova et al., 2019).

Implementation

FlexDotPlot is implemented in R and takes advantage of several publicly available R packages for data visualization, manipulation and analysis (Supplementary Table 2). It requires a simple data frame as input: the first two columns contain the two factors to spread along the x and y axes (e.g. genes and cell populations for scRNA-seq datasets), followed by the corresponding continuous and/or discrete data to be displayed (Supplementary Fig. S1A). From this input, dot plots can be produced with a single command line or interactively with a Shiny application.

Features highlights

FlexDotPlot consists of a single function to generate dot plots with several easy-to-tune parameters allowing users to specify which and how information has to be represented (Supplementary Fig. S1B). In addition to the traditional size and colour features, users can display two additional information by adding some text or by using dot shapes (Fig. 1). One major improvement of FlexDotPlot relies on the possibility to represent a defined factor by using variable dot shapes, which is not possible with other existing dot plot representation methods (Supplementary Table 1). This characteristic is really suitable to represent percentages as it is classically done in scRNA-seq related dot plots (Supplementary Fig. S2).

Conclusion

FlexDotPlot is an easy to use tool for generating highly customizable dot plot representations. It uses standardized input and output formats (data frame and ggplot2 object respectively), which also makes the combination of FlexDotPlot and other R pipelines possible. A set of vignettes are included in the package to illustrate the dot plot construction procedures and their associated parameters.

References

Benno Haarman,B.C.M. et al. (2015) Feature-expression heat maps - A new visual method to explore complex associations between two variable sets. J. Biomed. Inform., 53, 156–161.

Bi,K. et al. (2021) Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma. Cancer Cell, 39, 649-661.e5.

Efremova,M. et al. (2019) CellPhoneDB v2.0: Inferring cell-cell communication from combined expression of multi-subunit receptor-ligand complexes. bioRxiv, 680926.

Habib,N. et al. (2017) Massively parallel single-nucleus RNA-seq with DroNc-seq. Nat. Methods, 14, 955–958.

A

Gene	Cell type	Pct.exp (ADT)	Average expression (ADT)	Canonical marker ?	Pct.exp (RNA)
FCGR3A (CD16)	CD16+ Monocytes	90	1.88	yes	87
PTPRC (CD45RA)	CD16+ Monocytes	42	-0.6	no	82
FCGR3A (CD16)	NK cells	86	1.77	no	65
PTPRC (CD45RA)	NK cells	100	2.42	no	62

size colour shape text

B