Intro to R for Biologists
Session 5
Data Visualisation

Irina & Rao

03/02/2021

Hilary 2021

INTRO TO R FOR BIOLOGISTS

- **Data Visualisation**
  - Plotting continued from Session 4
  - Statistics
  - Problem set from Session 4
  - Heatmaps
  - Colours
  - Q & A

Basic statistics

- Correlations – `cor(), cor.test()`
- T-test (parametric):
  - Unpaired 2-group t-test

`t.test(y~x)` # where y is numeric and x is a binary factor

`t.test(y1,y2)` # where y1 and y2 are numeric

  - Paired 2-group t-test

`t.test(y1,y2, paired=TRUE)` # where y1 & y2 are numeric

- Non-parametric tests:
  - Unpaired 2-group Mann-Whitney U Test

`wilcox.test(y~A)` # where y is numeric and A is A binary factor

`wilcox.test(y,x)` # where y and x are numeric

  - Paired 2-group Wilcoxon Signed Rank Test

`wilcox.test(y1,y2, paired=TRUE)` # where y1 and y2 are numeric

- More than 2 groups - analysis of variance (ANOVA, parametric)

`aov(), anova()`

Adding statistics to the plots

```
library(ggpubr)
```

- Perform the test

```
compare_means(formula, data, method = "wilcox.test", paired = FALSE, group.by = NULL, ref.group = NULL, ...)
```

*method = "t.test", "anova"

- Or just add the significance levels to the plot

```
stat_compare_means(mapping = NULL, comparisons = NULL, hide.ns = FALSE, label = NULL, label.x = NULL, label.y = NULL, ...)
```

Adding statistics to the plots - alternatively

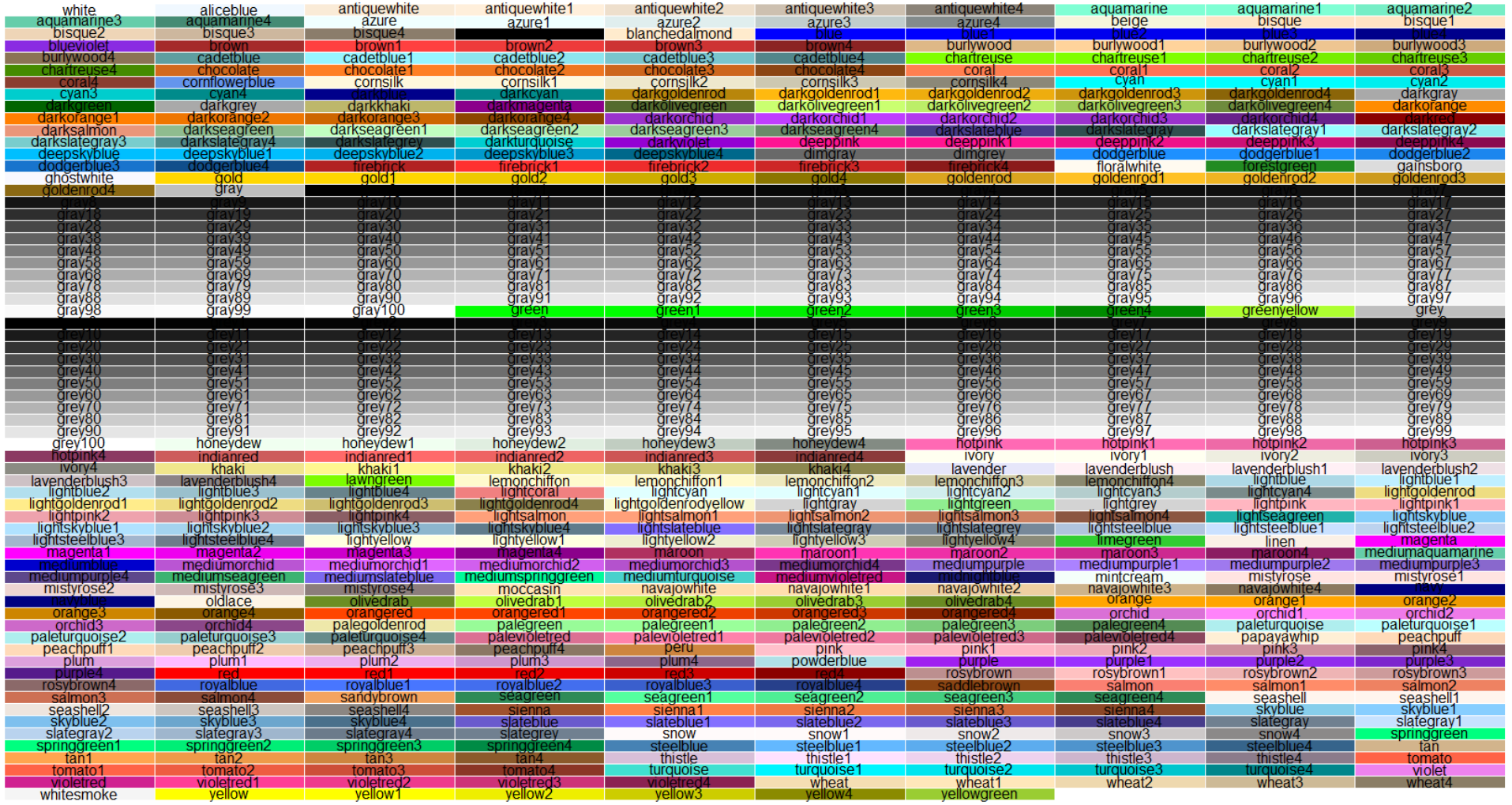
```
library(ggsignif)
stat_signif(mapping = NULL, data = NULL,
position = "identity", na.rm = FALSE,
show.legend = NA, inherit.aes =
TRUE, comparisons = NULL, test =
"wilcox.test", test.args = NULL, annotations
= NULL, map_signif_level = FALSE ...)

map_signif_level = c("***"=0.001, "**"=0.01,
"*"=0.05)
```

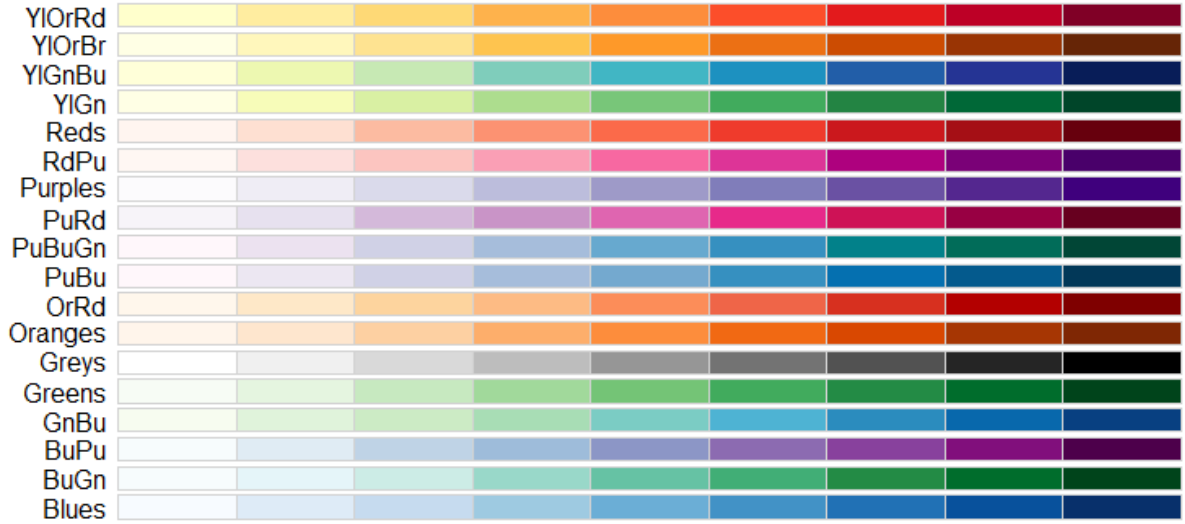
Let’s explore practically

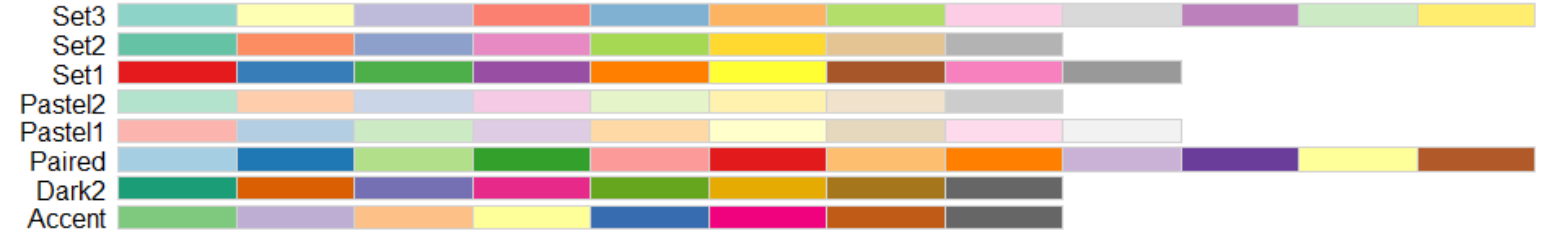
Colours in R

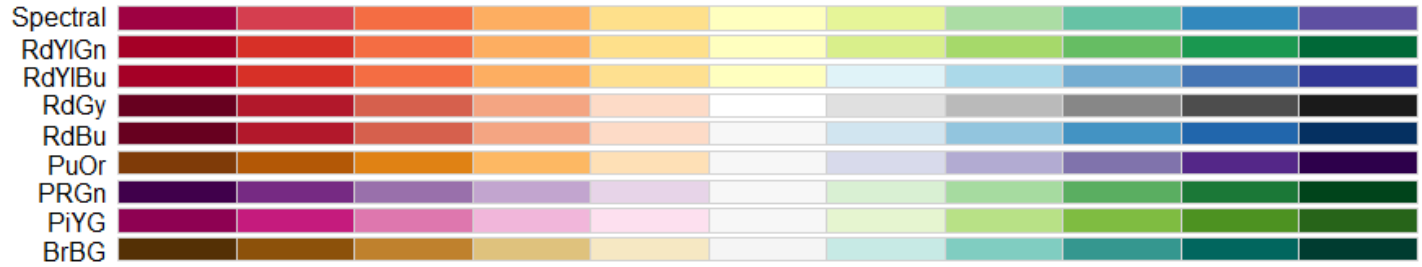
- Represented as named colours (e.g. “red”, “mediumspringgreen”) or hexadecimal code (e.g. “#FF0000FF”, “#00FA9AFF”)



Colours in R



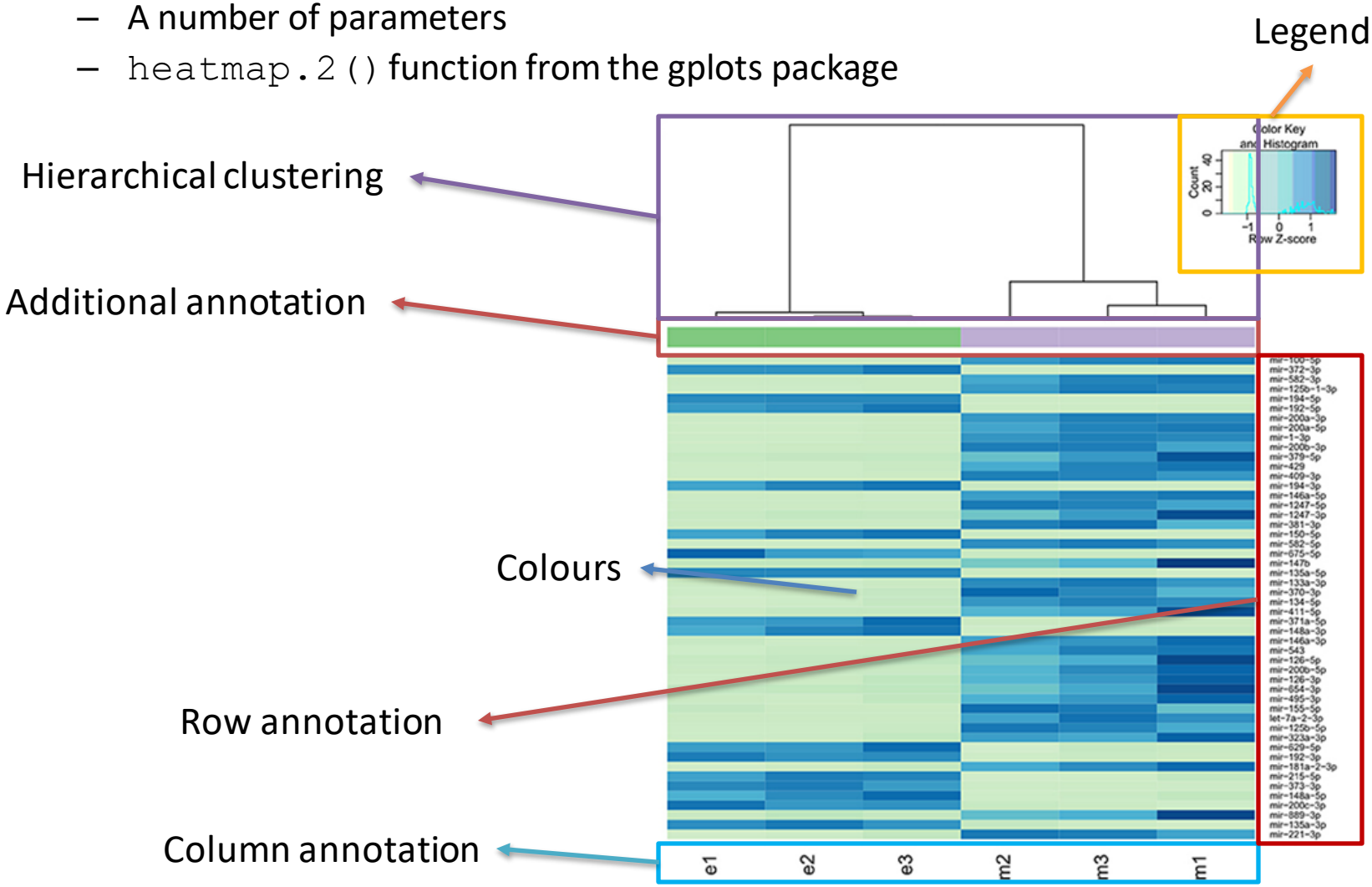




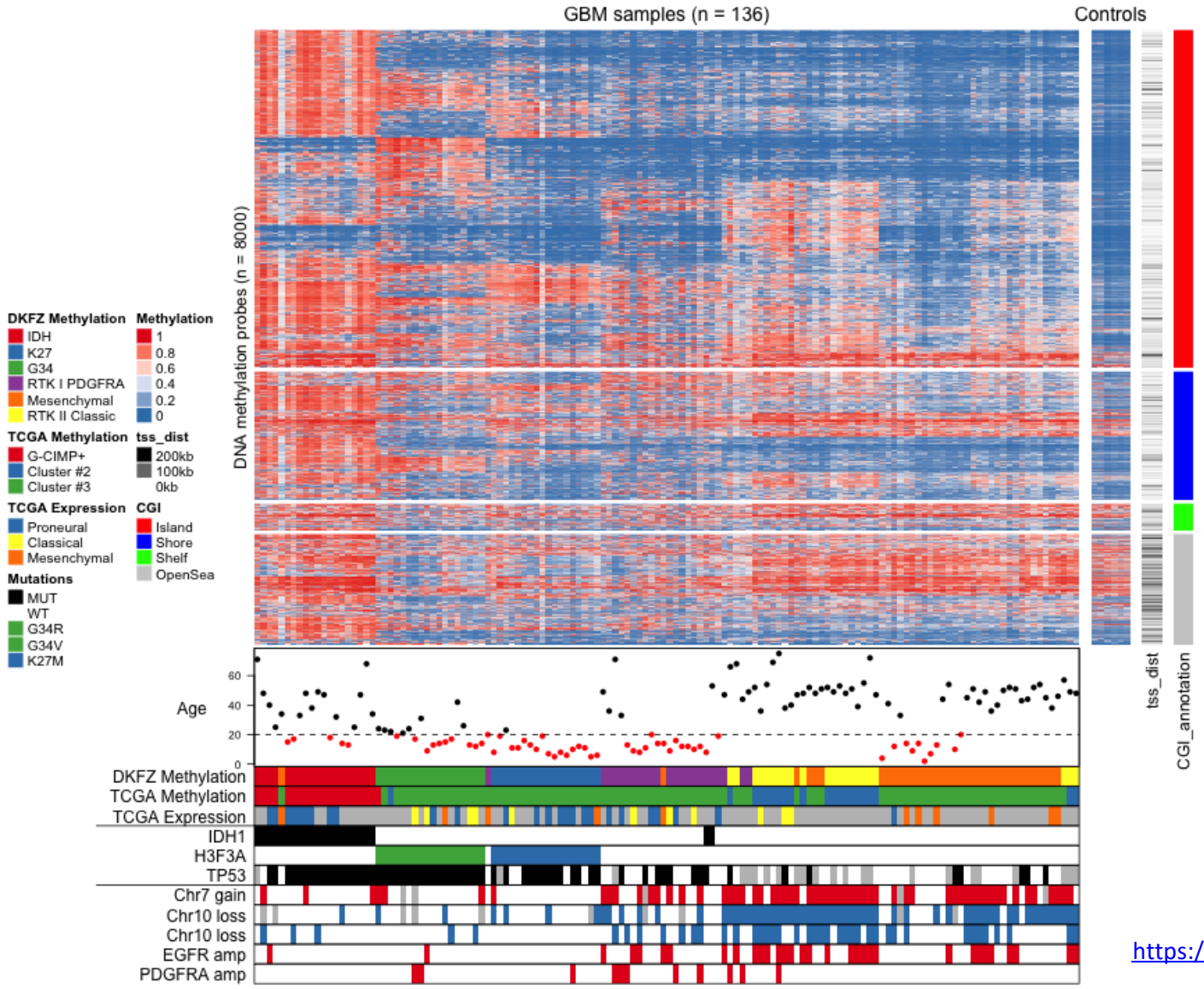
- Palettes – a set of colours that can be used together
- RColorBrewer package
  - Sequential
  - Qualitative
  - Diverging
- https://github.com/EmilHvitfeldt/r-color-palettes for a huge list of palettes and the packages they come in

Heatmaps

- Useful for displaying multidimensional data
  - A number of parameters
  - `heatmap.2()` function from the gplots package



A lot of information can be conveyed!

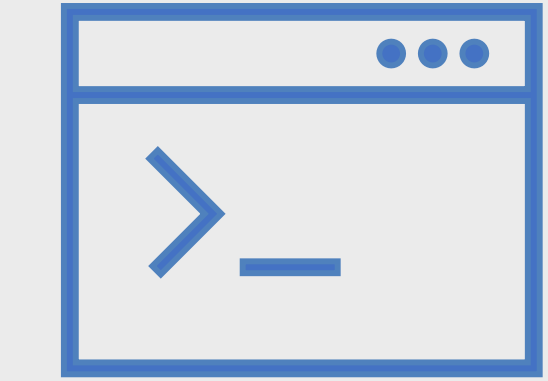


- A more complicated example with the Complex Heatmap package

https://github.com/jokergoo/ComplexHeatmap

Let’s explore practically

Let’s explore practically



Address the tasks in breakout rooms!

Useful references

- https://github.com/EmilHvitfeldt/r-color-palettes - a huge list of R palettes
- https://colorbrewer2.org/ - RColorBrewer package based on this tool by Cynthia Brewer
- https://jokergoo.github.io/ComplexHeatmap-reference/book/ - R package to produce Complex Heatmaps