R in Biology
Examples of genomic data visualisation using R

Valentine Murigneux
R-Ladies Brisbane Meetup, 19th August 2020

About me and R

What do I do?
- Bioinformatician at the University of Queensland
  QCIF Facility for Advanced Bioinformatics (QFAB) and UQ Genome Innovation Hub
- Research in genomics, data analysis

How do I use R?
- Bioinformatics/biostatistics packages
- Projects reports using R Markdown
- Data visualisation

Examples of genomic data visualisation using R

Useful packages in cancer genomics

- GenVisR: Genomic Visualizations in R
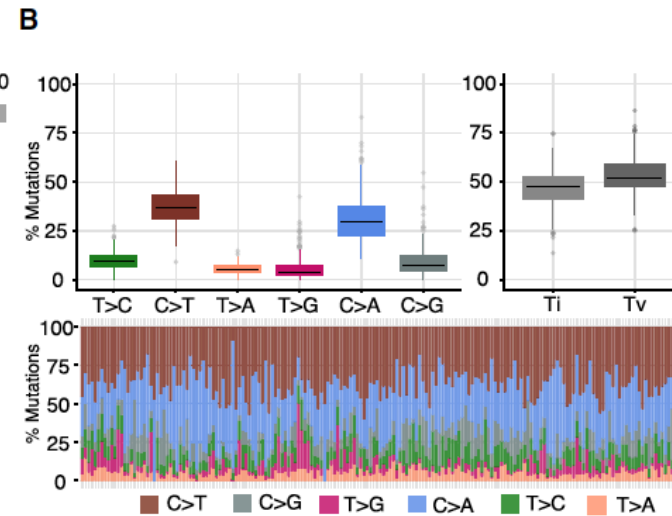- Maftools: Summarize, Analyze and Visualize Mutation Annotation Format Files
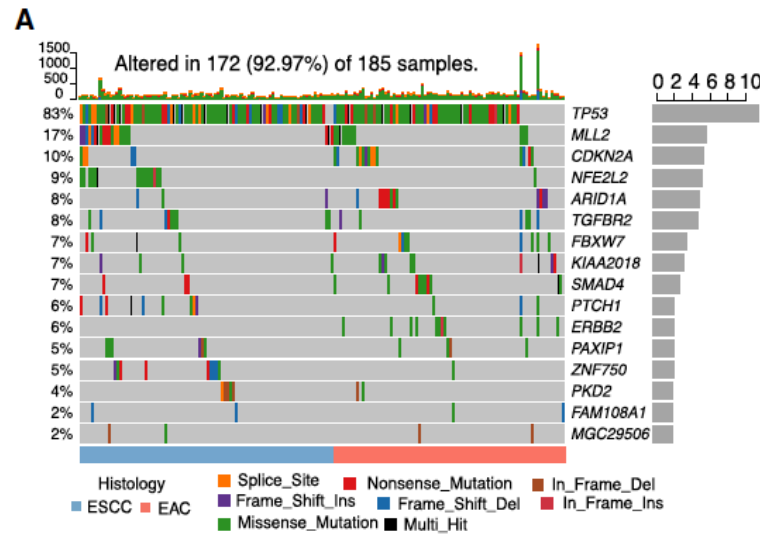
General packages

- Heatmaps: pheatmap
- Visualisation of set intersections: Venn diagrams, UpSetR

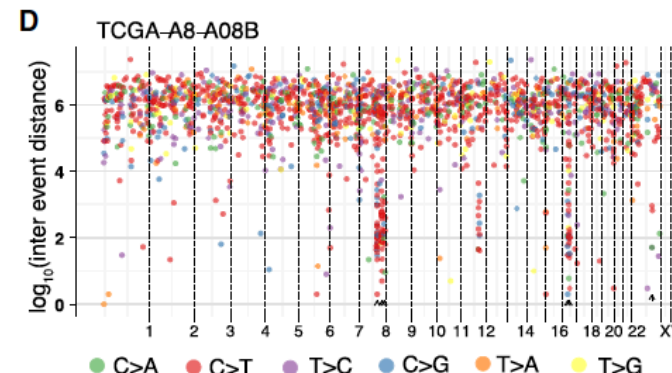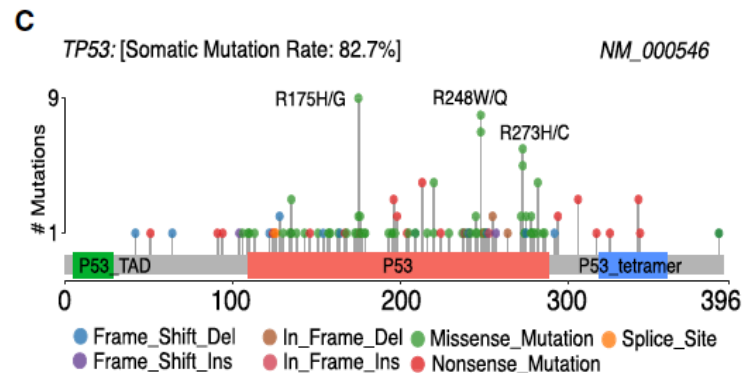# Maftools: Summarize, Analyze and Visualize Mutation Annotation Files
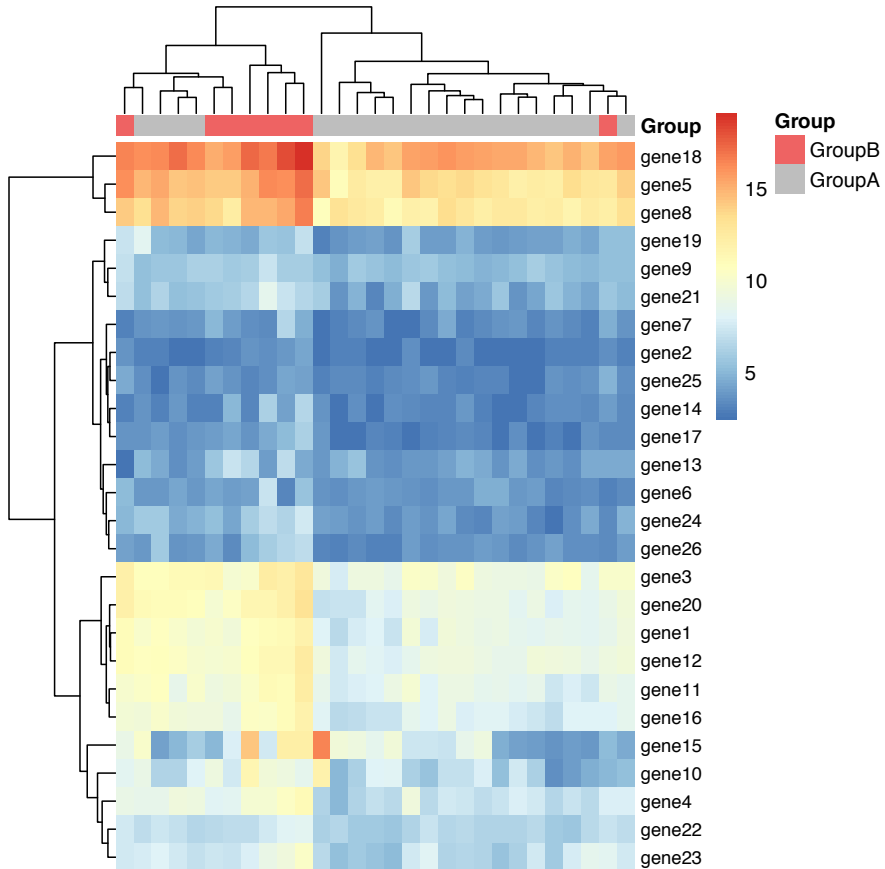

Oncoplot
(waterfall plot)

Mutation classification

Lollipop plot
for amino
acid changes

Rainfall plot

Mayakonda A, Lin D, Assenov Y, Plass C, Koeffler PH (2018). *Genome Research*. doi: 10.1101/gr.239244.118.
https://www.bioconductor.org/packages/release/bioc/html/maftools.html

pheatmap: Pretty Heatmaps

Features
- Numeric matrix object as input
- Hierarchical clustering of rows and columns
- Value scaling
- **Add annotations to rows and columns (via a data frame)**
- Cut the heatmap to pieces

https://cran.r-project.org/web/packages/pheatmap/index.html
Tutorial: https://towardsdatascience.com/pheatmap-draws-pretty-heatmaps-483dab9a3cc

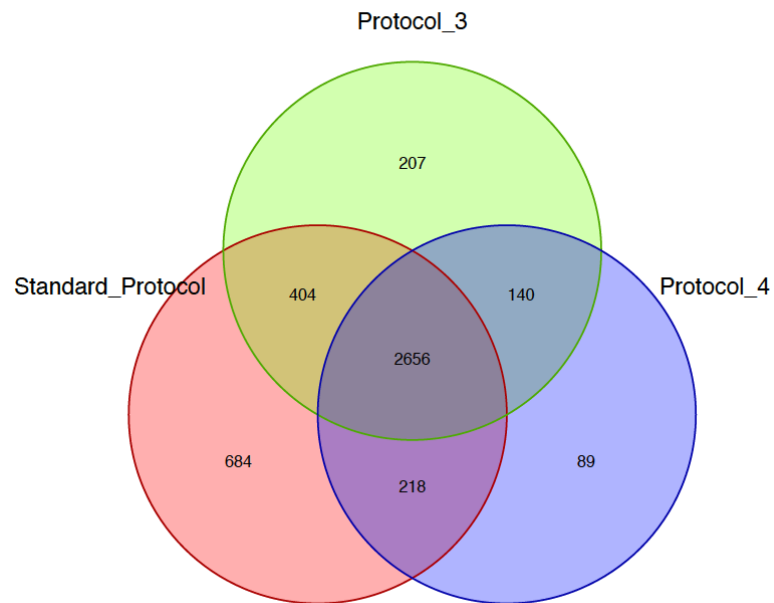# Visualisation of set intersections: Venn diagrams, UpSet plots

Example: comparing lists of expressed genes obtained from 5 different protocols

- **Venn package**: Draw Venn diagrams
- **Vennerable package**: Venn and Euler area-proportional diagrams
- **UpSetR package**: A More Scalable Alternative to Venn and Euler Diagrams for Visualizing Intersecting Sets
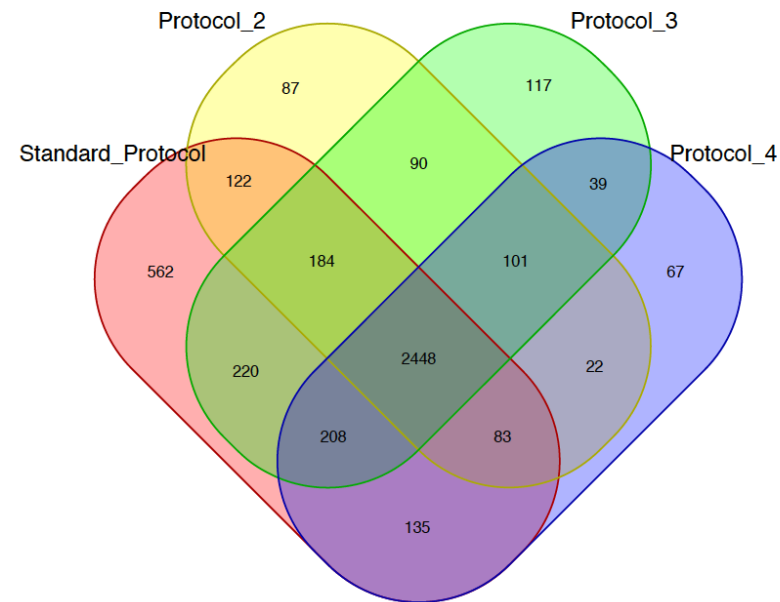
Venn package: Draw Venn diagrams (up to 7 sets)

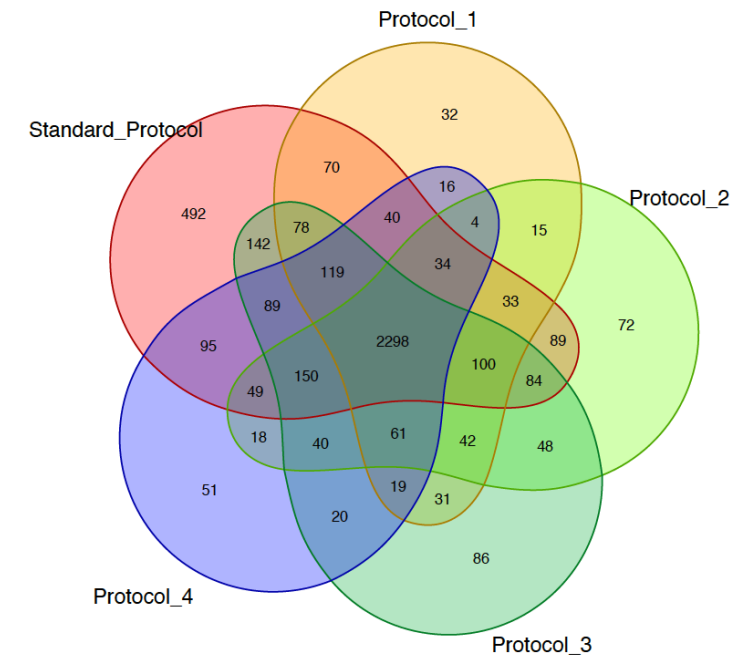Example: comparing lists of expressed genes obtained from 5 different protocols
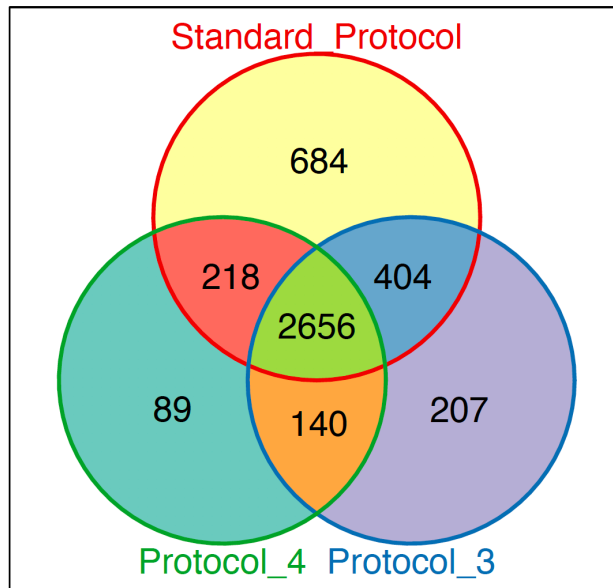

3 subsets

4 subsets

5 subsets

https://www.rdocumentation.org/packages/venn/versions/1.9/topics/venn

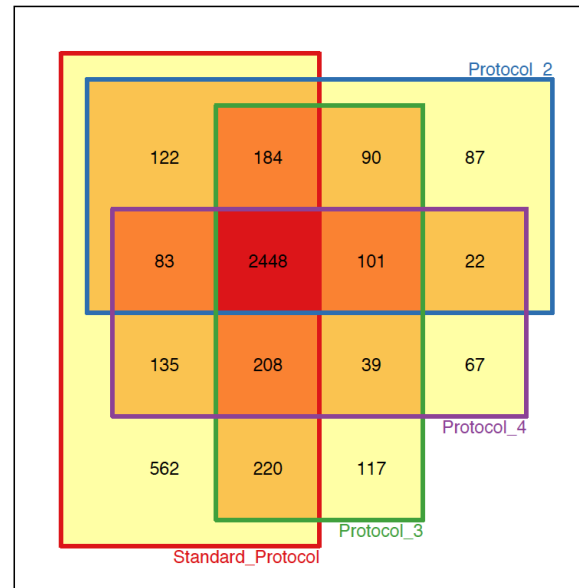# Vennerable package: Venn and Euler area-proportional diagrams

Example: comparing lists of expressed genes obtained from 5 different protocols
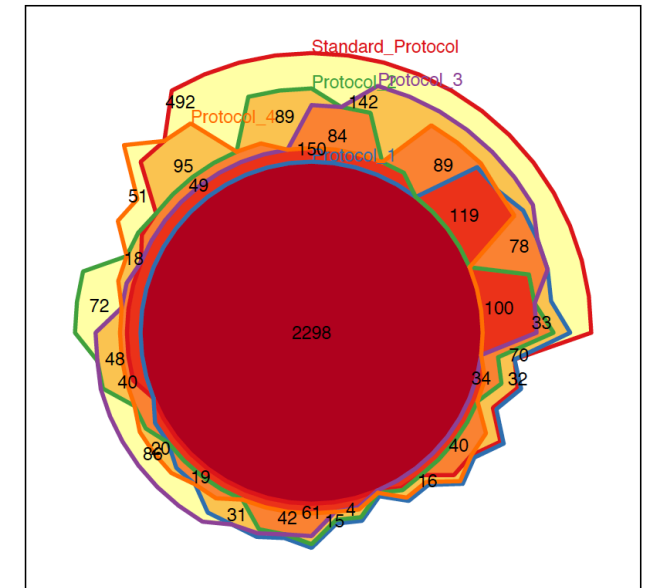

3 subsets

plot(V3, doWeight=FALSE)
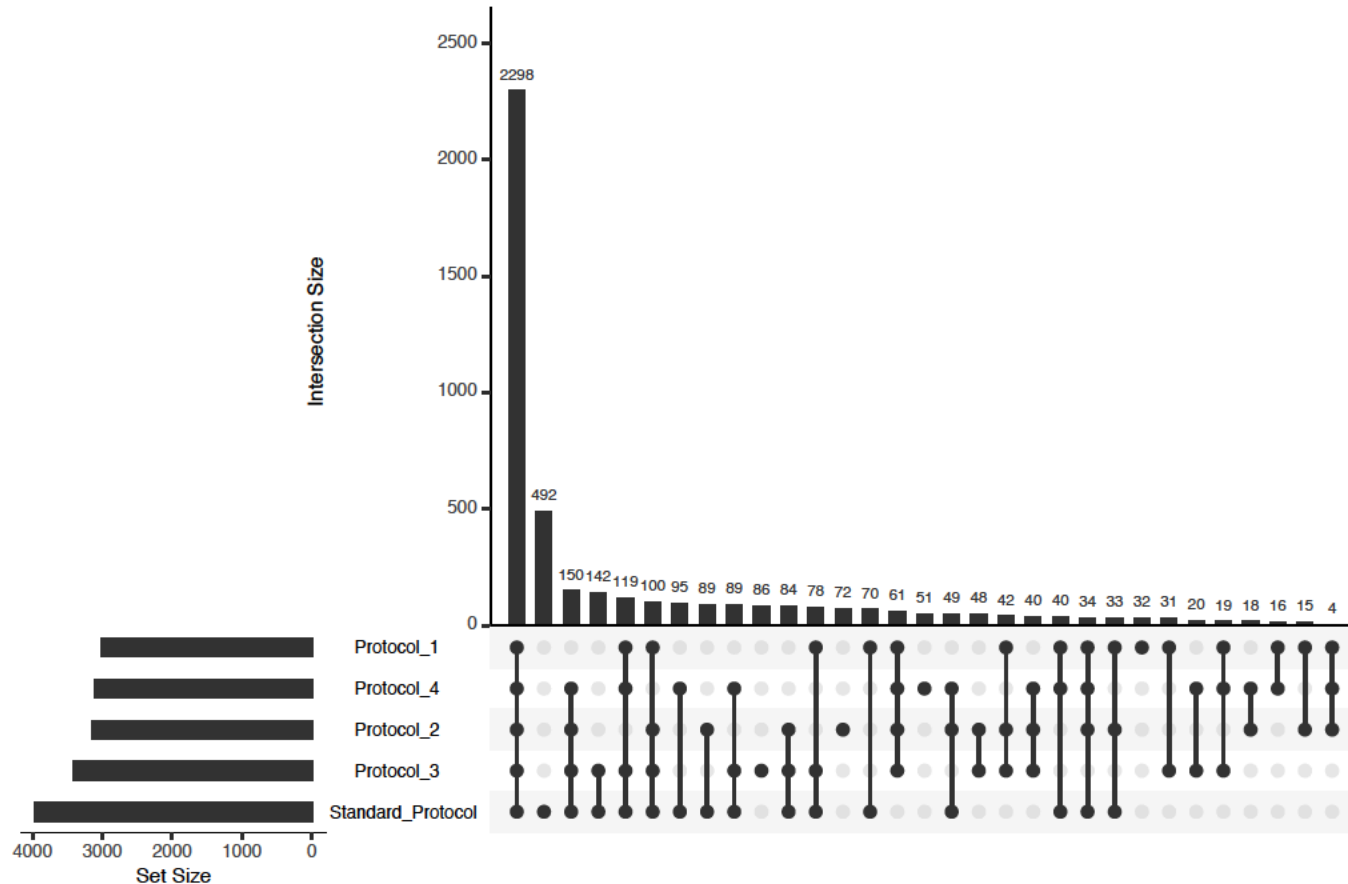
4 subsets

plot(V4, doWeight=FALSE)

5 subsets

plot(V5, doWeight=TRUE)

Use doEuler parameter to hide empty intersections

https://rdrr.io/rforge/Vennerable/

UpSetR package: A More Scalable Alternative to Venn Diagrams for Visualizing Intersecting Sets

Example: comparing lists of expressed genes obtained from 5 different protocols

R code

```
library(UpSetR)

# input data (list of named vectors)
listInput<-list(
        Standard_Protocol=SP,
        Protocol_1=NP1,
        Protocol_2=NP2,
        Protocol_3=NP3,
        Protocol_4=NP4
        )

# plot the graphe
upset(fromList(listInput), order.by = "freq")
```

https://github.com/hms-dbmi/UpSetR
Jake R Conway, Alexander Lex, Nils Gehlenborg, Bioinformatics, 2017

UpSetR : An R implementation of the UpSet set visualization technique

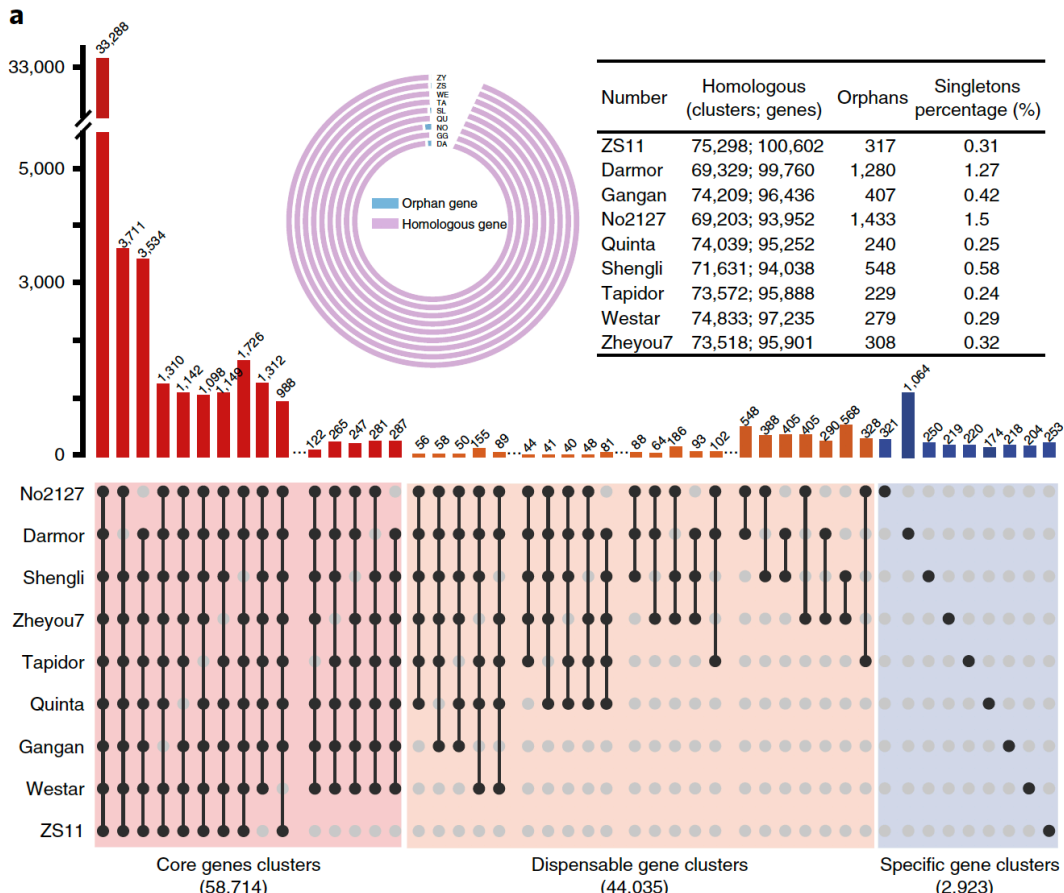**Useful resources**

- UpSet: http://caleydo.org/tools/upset/
- UpSetR, R version: https://github.com/hms-dbmi/UpSetR
- UpSetR Shiny Version: https://gehlenborglab.shinyapps.io/upsetr/
- UpSet Web Version: https://vcg.github.io/upset/
- pyUpSet, Python version: https://github.com/ImSoErgodic/py-upset

References

- UpSet: Lex, Gehlenborg, et al., IEEE Transactions on Visualization and Computer Graphics, 2014. https://doi.org/10.1109/TVCG.2014.2346248
- UpSetR: Jake R Conway, Alexander Lex, Nils Gehlenborg, Bioinformatics, 2017. https://doi.org/10.1093/bioinformatics/btx364
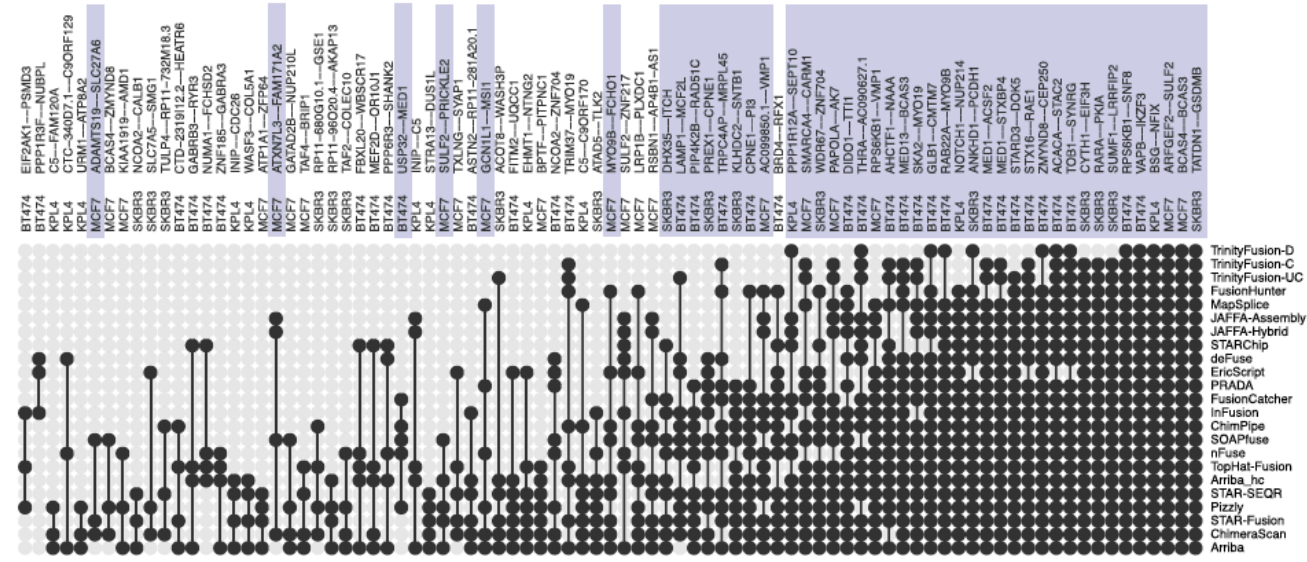- Points of view: Sets and intersections. Alexander Lex, Nils Gehlenborg. Nature Methods, vol. 11, no. 8, pp. 779, 2014. https://doi.org/10.1038/nmeth.3033

UpSet visualisation examples

Jenjaroenpun et al. Nucleic Acids Research, 2020

Song et al. Nature Plants, 2020

Haas et al. Genome Biology (2019) 20:213

Thank you !