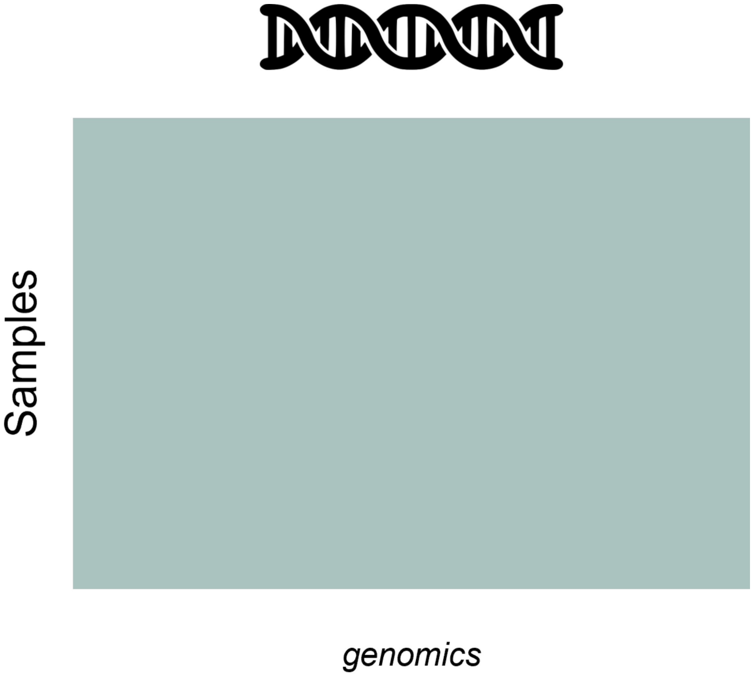
Integrating and visualising multi-omics data with the `{moiraine}` R package

Olivia Angelin-Bonnet
The New Zealand Institute for Plant and Food Research Ltd

3 September 2024

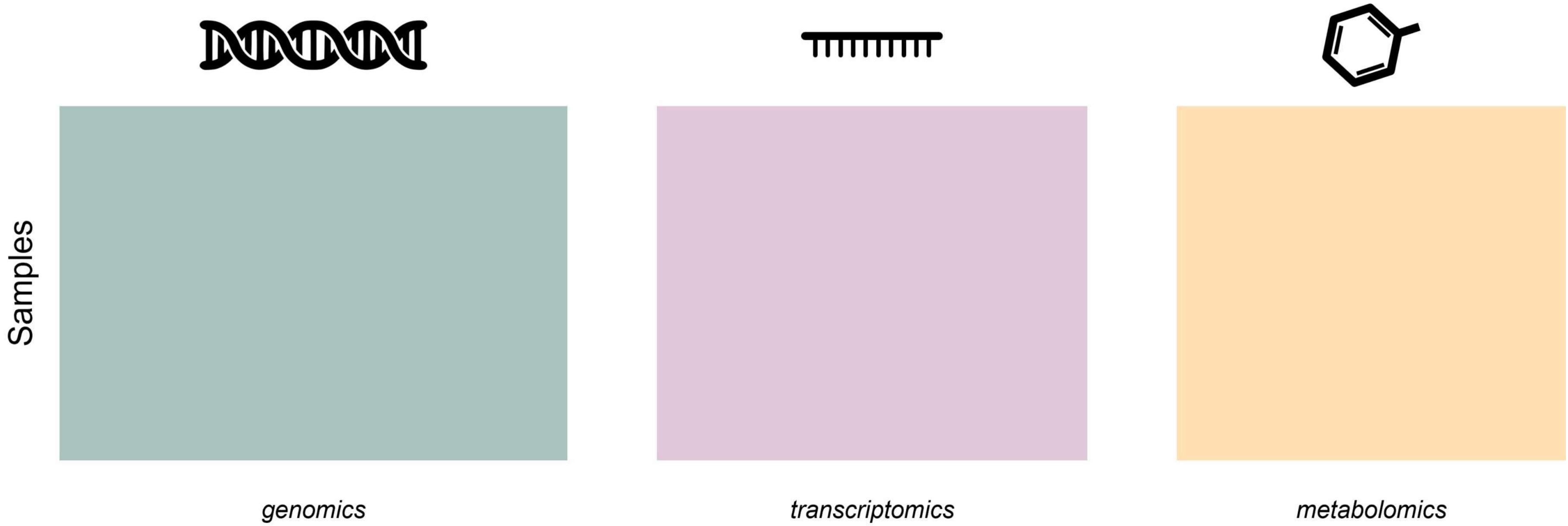
Multi-omics data

Omics data: High-throughput measurement of molecular layer across biological samples

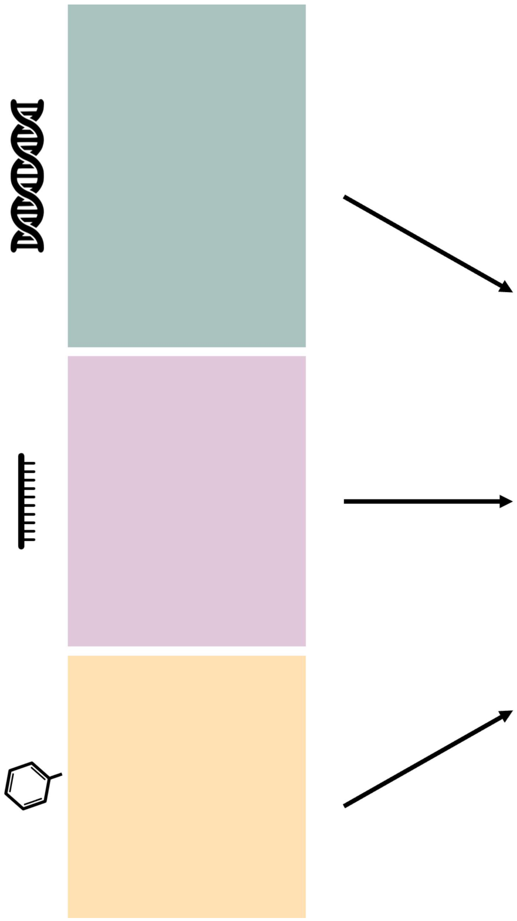


Multi-omics data

Multi-omics data: Different omics layers measured on the same biological samples



Multi-omics integration



Multi-omics integration methods:

- DIABLO – `{mixOmics}`
- MOFA – `{MOFA2}`
- MCIA – `{omicade4}`
- omeSOM
- MFA – `{FactoMineR}`
- …

Technical challenges

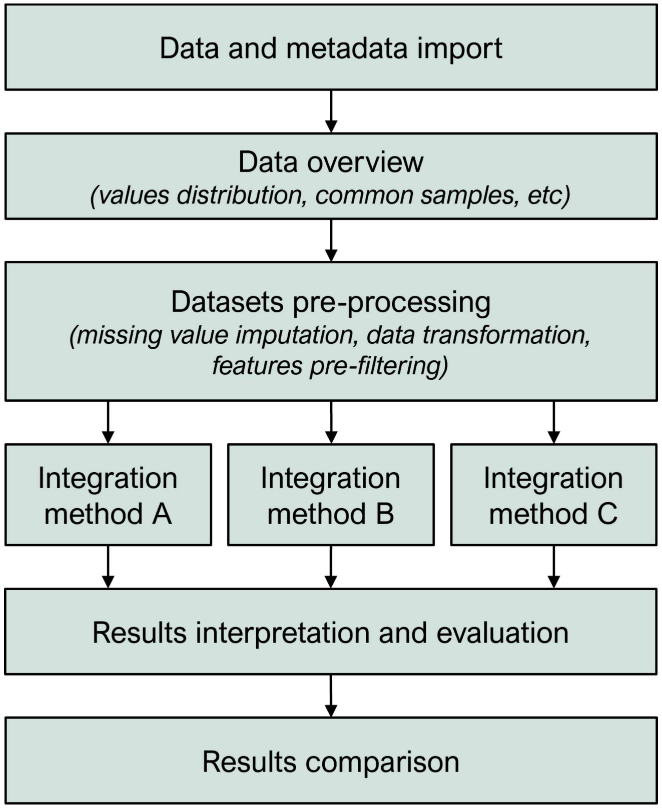
Different requirements for input data format

Different formatting of results

Unequal visualisation capabilities

The `{moiraine}` R package

- Construction of reproducible multi-omics data integration pipelines
- Comparison of results from several integration methods



https://github.com/Plant-Food-Research-Open/moiraine/

Three key features of `{moiraine}`

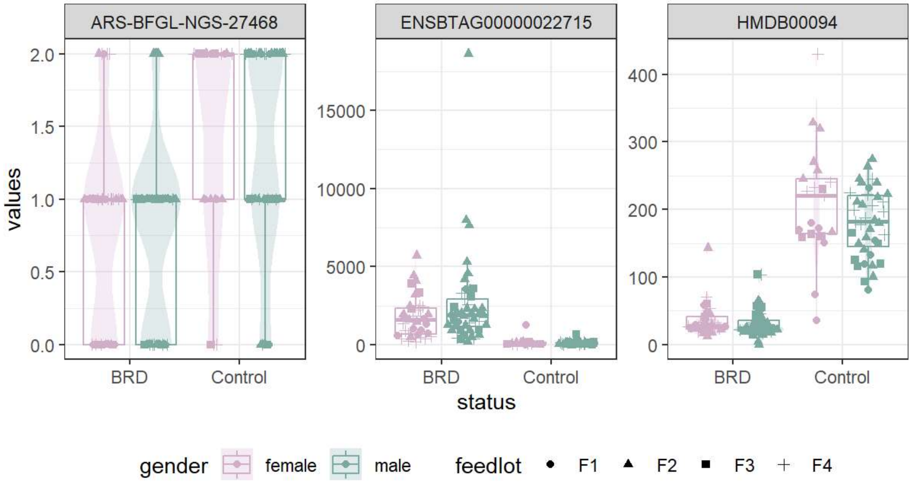
Storing data with `{MultiDataSet}`

Using the `{MultiDataSet}` package to store omics datasets and related metadata in one R object:

```r
mo_set
```

```
Object of class 'MultiDataSet'
 . assayData: 3 elements
    . snps: 23036 features, 139 samples
    . rnaseq: 20335 features, 143 samples
    . metabolome: 55 features, 139 samples
 . featureData:
    . snps: 23036 rows, 13 cols
    . rnaseq: 20335 rows, 15 cols
    . metabolome: 55 rows, 16 cols
 . rowRanges:
    . snps: YES
    . rnaseq: YES
    . metabolome: NO
 . phenoData:
    . snps: 139 samples, 10 cols
    . rnaseq: 143 samples, 10 cols
    . metabolome: 139 samples, 10 cols
```

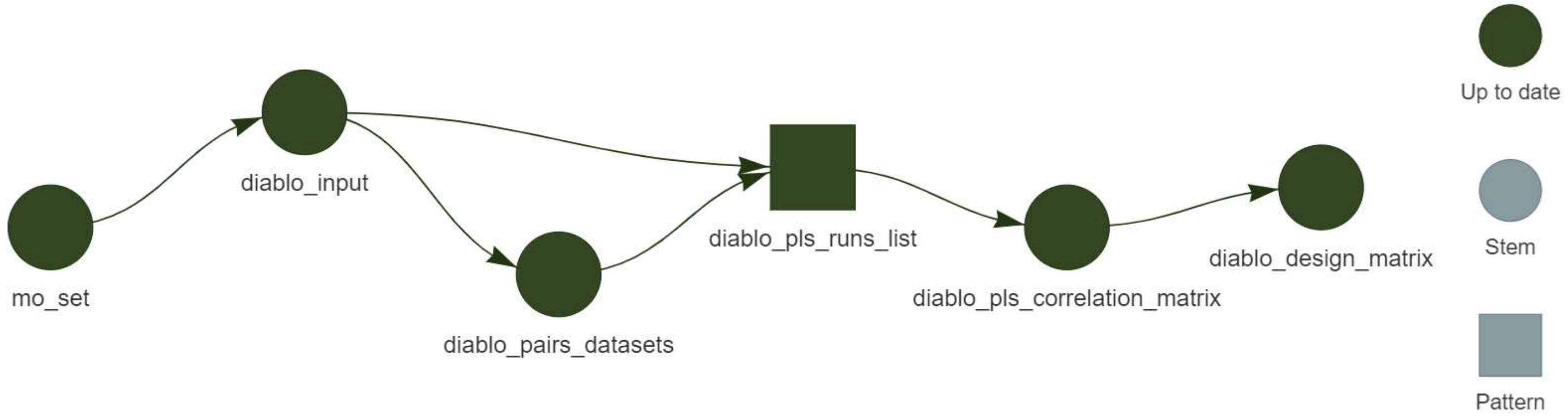
```r
plot_data_covariate(
  mo_set,
  covariate = "status",
  features = c("ARS-BFGL-NGS-27468",
               "ENSBTAG00000022715",
               "HMDB00094"),
  colour_by = "gender",
  shape_by = "feedlot"
)
```



Reproducible pipelines with `{targets}`

Uses the `{targets}` package to build reproducible analysis pipelines

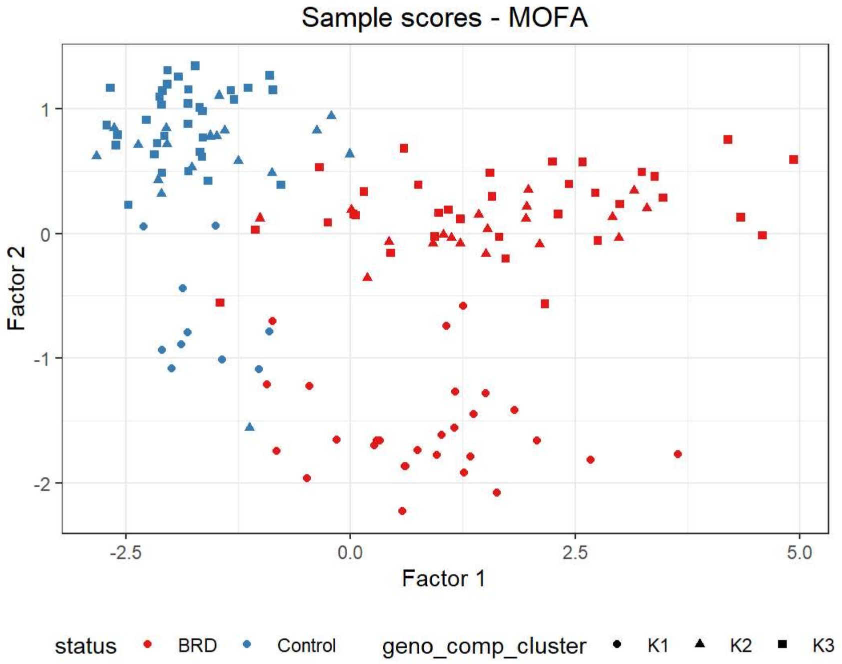
e.g. automates multi-step processes such as estimation of a parameter of interest:



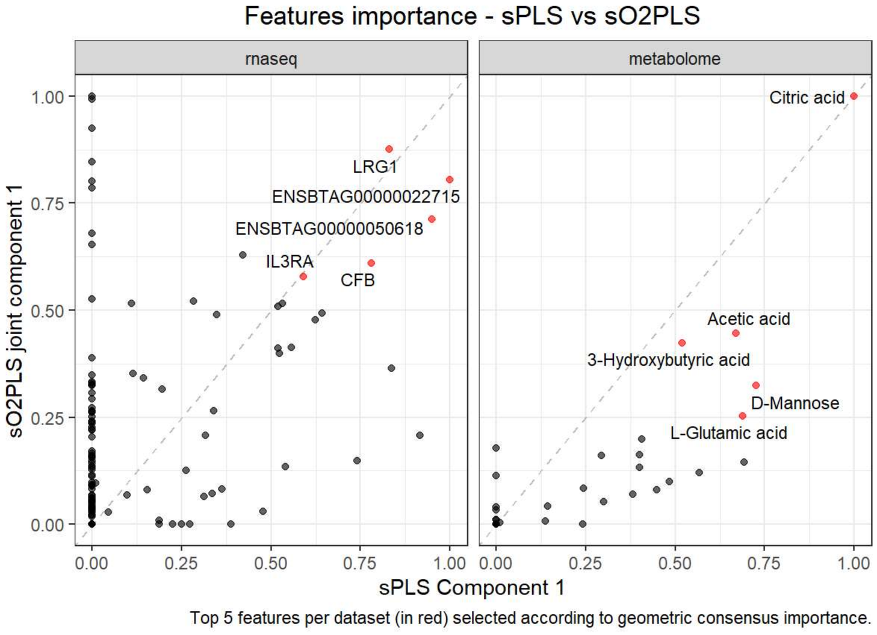
Consistent formatting of integration output

Converts methods’ output into S3 object for storing dimension reduction results

Results interpretation:



Methods comparison:



Top 5 features per dataset (in red) selected according to geometric consensus importance.

Thank you for your attention!

olivia.angelin-bonnet@plantandfood.co.nz

Acknowledgments

- Susan Thomson
- Roy Storey
- Lindy Guo
- Nigel Joyce
- Blake List
- Leonardo Salgado

Presentation disclaimer

Presentation for
Australasian Applied Statistics Conference 2024, Rottnest Island, Western Australia, 3–5 September 2024

Publication data:
Angelin-Bonnet O. September 2024. Integrating and visualising multi-omics data with the `{moiraine}` R package. A Plant & Food Research PowerPoint presentation. SPTS No. 26038.

Presentation prepared by:
Olivia Angelin-Bonnet
Scientist, Data Science
September 2024

Presentation approved by:
Mark Wohlers
Science Group Leader, Data Science
September 2024

For more information contact:
Olivia Angelin-Bonnet
DDI: +64 6 355 6156
Email: Olivia.Angelin-Bonnet@plantandfood.co.nz

DISCLAIMER

The New Zealand Institute for Plant and Food Research Limited does not give any prediction, warranty or assurance in relation to the accuracy of or fitness for any particular use or application of, any information or scientific or other result contained in this presentation. Neither The New Zealand Institute for Plant and Food Research Limited nor any of its employees, students, contractors, subcontractors or agents shall be liable for any cost (including legal costs), claim, liability, loss, damage, injury or the like, which may be suffered or incurred as a direct or indirect result of the reliance by any person on any information contained in this presentation.

COPYRIGHT

© COPYRIGHT (2024) The New Zealand Institute for Plant and Food Research Limited. All Rights Reserved. No part of this report may be reproduced, stored in a retrieval system, transmitted, reported, or copied in any form or by any means electronic, mechanical or otherwise, without the prior written permission of The New Zealand Institute for Plant and Food Research Limited. Information contained in this report is confidential and is not to be disclosed in any form to any party without the prior approval in writing of The New Zealand Institute for Plant and Food Research Limited. To request permission, write to: The Science Publication Office, The New Zealand Institute for Plant and Food Research Limited – Postal Address: Private Bag 92169, Victoria Street West, Auckland 1142, New Zealand; Email: SPO-Team@plantandfood.co.nz.

This report has been prepared by The New Zealand Institute for Plant and Food Research Limited (Plant & Food Research).
Head Office: 120 Mt Albert Road, Sandringham, Auckland 1025, New Zealand, Tel: +64 9 925 7000, Fax: +64 9 925 7001.
www.plantandfood.co.nz