



# InMoose: the Integrated Open Source Python Package for Multi-omic Analyses

Poster 6287

M. Colange<sup>1</sup>, G. Appé<sup>1</sup>, L. Meunier<sup>1</sup>, S. Weill<sup>1</sup>, A. Nordor<sup>1</sup>, A. Behdenna<sup>1</sup>. <sup>1</sup>Epigene Labs, Paris, France

AACR

American Association  
for Cancer Research

## Bridging the Reproducibility Gap between R and Python

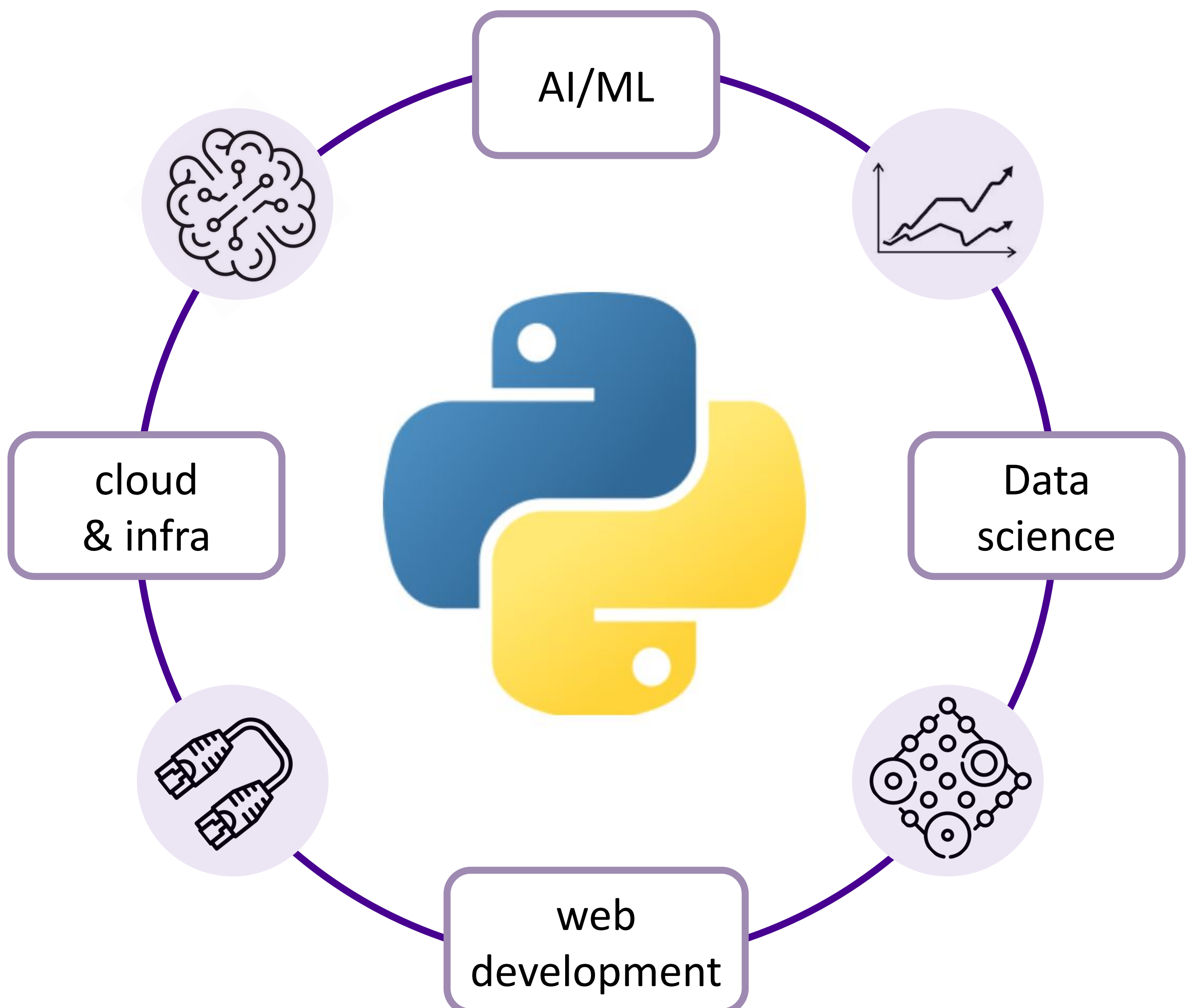
### Approach

#### Toward a Unified Python Ecosystem for Omics

- As a general-purpose language, **Python** is **versatile, accessible and performant**.
- Python is gaining momentum in bioinformatics, and proves crucial to connect omic data with ML/AI tools or with user-friendly user interfaces.

InMoose provides **drop-in replacements** of state-of-the-art R tools, packaged in a **single toolbox package**.

- Open source**
- Aims to foster a larger collaborative effort to build and grow a consistent state-of-the-art **Python ecosystem** for cancer bioinformatics.

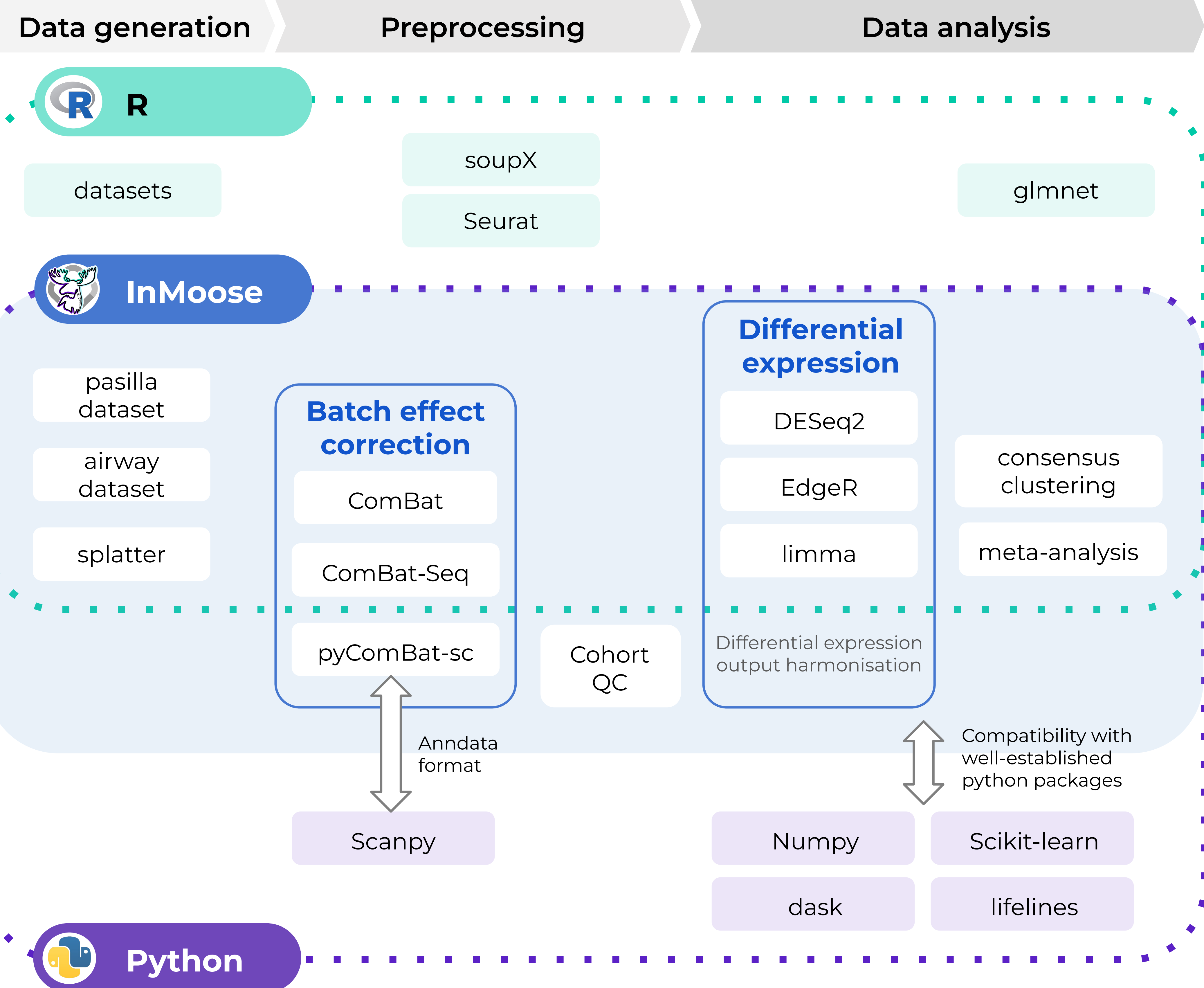


### References

- Behdenna A, Colange M, Haziza J, Cerna A, Appé G, Azencott CA, et al. pyComBat, a Python tool for batch effects correction in high-throughput molecular data using empirical Bayes methods. BMC Bioinformatics. 2023 Dec 7;24(1):459.
- Colange M, Appé G, Meunier L, Weill S, Nordor A, Behdenna A. Differential Expression Analysis with InMoose, the Integrated Multi-Omic Open-Source Environment in Python. bioRxiv; 2024. p. 2024.11.14.623578.
- Colange M, Appé G, Meunier L, Weill S, Johnson WE, Nordor A, et al. Bulk Transcriptomic Analysis with InMoose, the Integrated Multi-Omic Open-Source Environment in Python. bioRxiv; 2024. p. 2024.11.29.625982.
- Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007 Jan 1;8(1):118–27.
- Zhang Y, Parmigiani G, Johnson WE. ComBat-seq: batch effect adjustment for RNA-seq count data. NAR Genomics Bioinforma. 2020 Sep;2(3):lqaa078.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014 Dec 5;15(12):550.
- Chen Y, Lun ATL, Smyth GK. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. F1000Research; 2016.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20;43(7):e47–e47.
- Zappia L, Phipson B, Oshlack A. Splatter: simulation of single-cell RNA sequencing data. Genome Biol. 2017 Sep 12;18(1):174.

### Integrated Multi Omic Open Source Environment

Overlap representation between R and Python implementations within the open-source InMoose framework



- Harmonized data structures across InMoose modules
  - Unified output format for differential expression, enabling direct comparisons and direct meta-analysis, within or across technologies
- Seamless integration with other Python tools through standardized, best-practice data formats
- Additional modules extend the pipeline, including cohort-level quality control and batch effect correction for single cell datasets

### Expanding InMoose

Already adopted by part of the bioinformatics community

- 47,000+** downloads on PyPI
- Public GitHub repository** with active usage and community contributions
- pyComBat **cited in 22 peer-reviewed publications** (via CrossRef)
- 2 additional papers** on InMoose currently under peer-review

European  
Innovation  
Council



Co-funded by  
the European Union

### Contact

Akpéli Nordor, PharmD, PhD ([akpeli@epigenelabs.com](mailto:akpeli@epigenelabs.com)).  
The authors have no conflict of interest to declare.

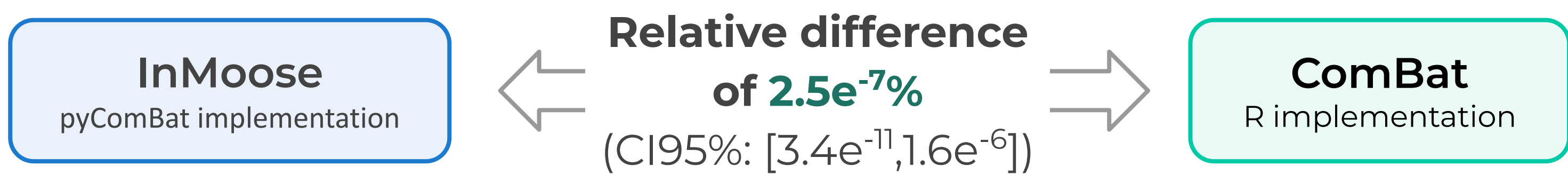
CHECK US  
OUT!



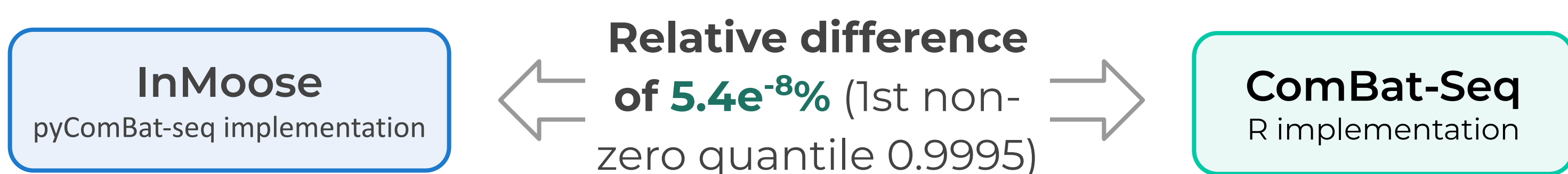
### Reproducibility

Enables reproducible R-to-Python migration

#### Microarray batch effect correction results comparison



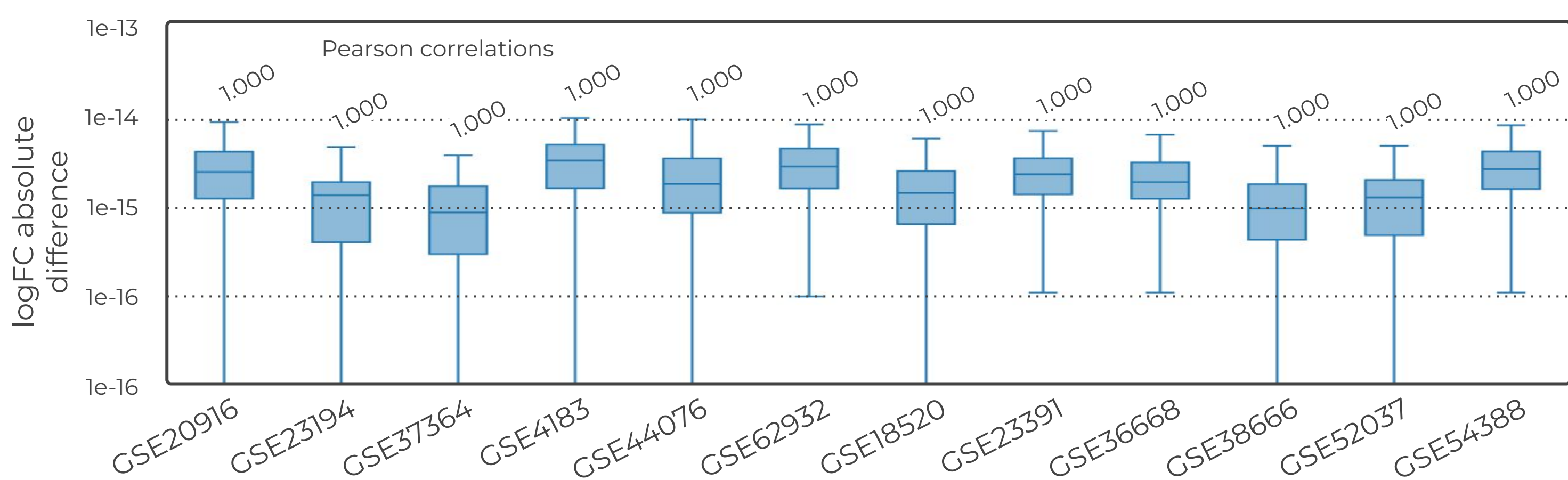
#### RNAseq batch effect correction results comparison



**InMoose replicates** the results of ComBat and ComBat-Seq more than **4x faster** than the original R implementation.

#### Microarray differential expression analysis results comparison

##### C Results difference InMoose vs Limma

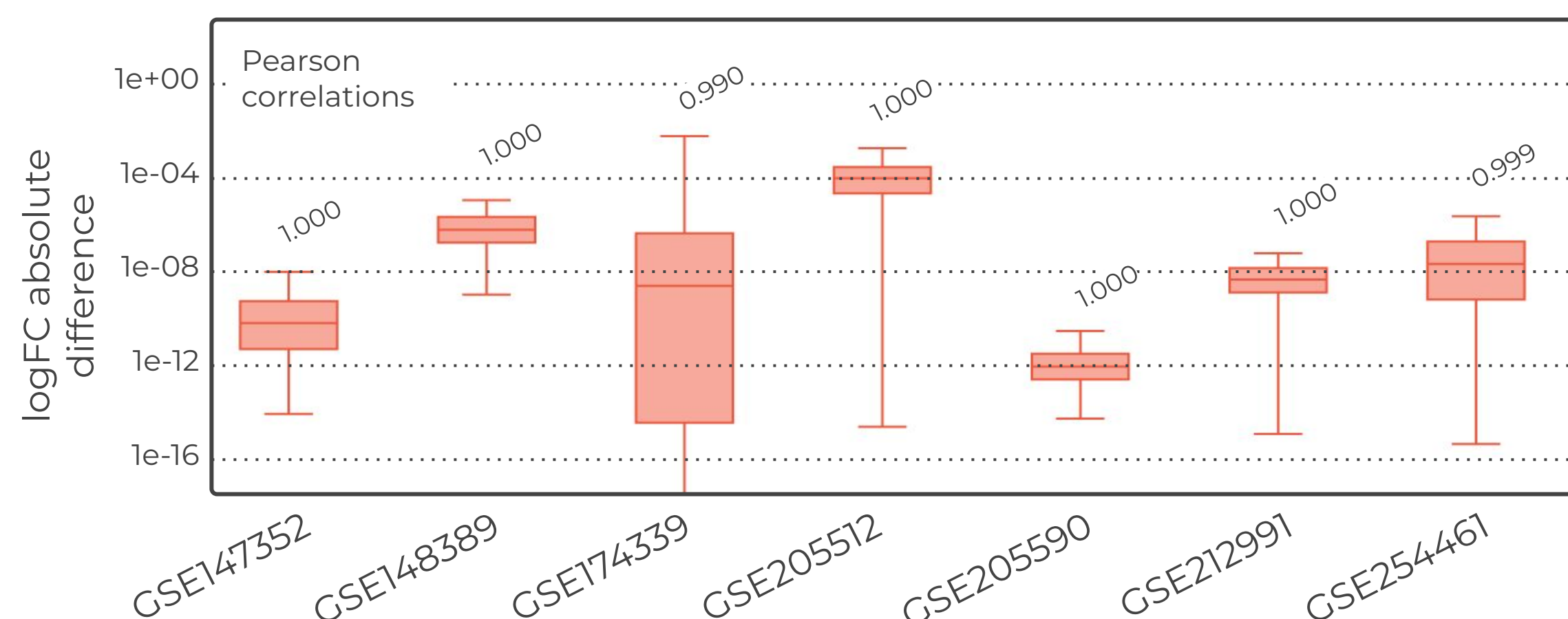


→ logFC **perfectly correlated** between InMoose and limma

Figure 2: Comparison of InMoose with limma, EdgeR, DESeq2. Absolute differences of log-fold-change between inMoose and C. limma D. DESeq2 and E. EdgeR. Boxes represent the median, and first and third quartiles, whiskers represent the 2.5% and 97.5% quantiles and number above the box are Pearson correlations between the results

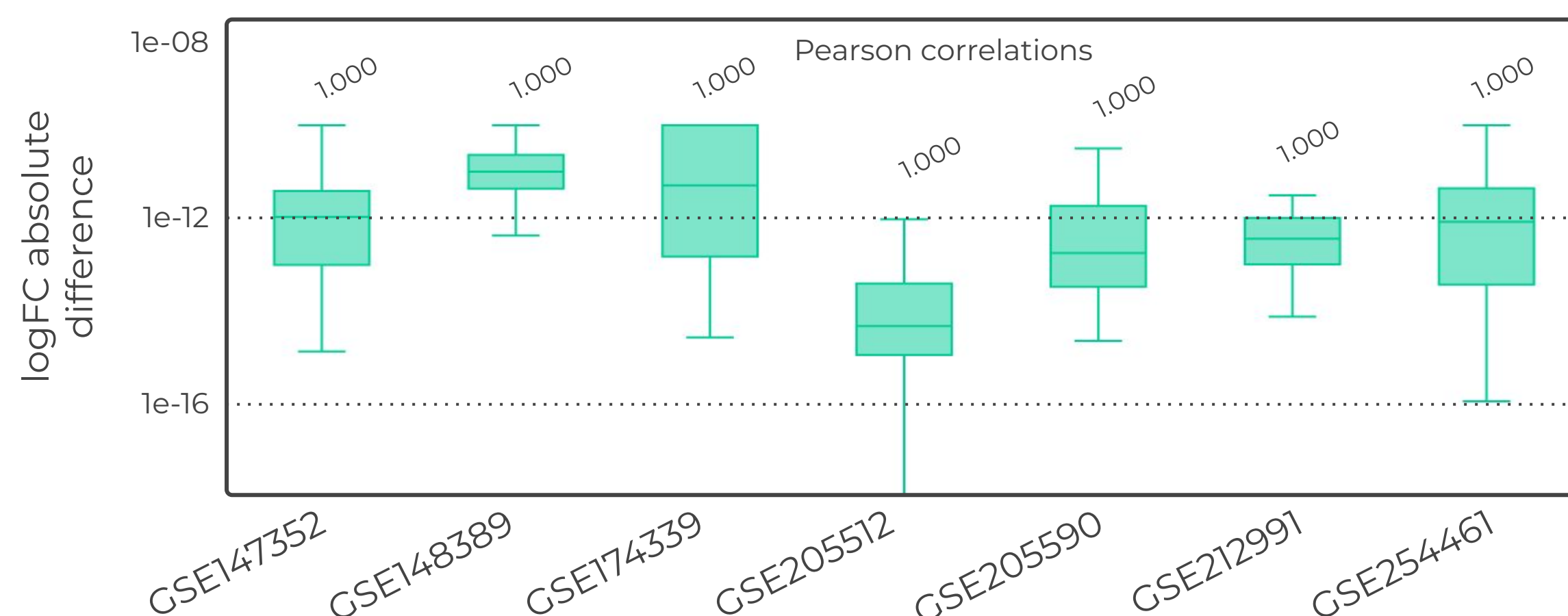
#### RNAseq differential expression analysis results comparison

##### D Results difference InMoose vs DESeq2



→ logFC **highly correlated** between InMoose and DESeq2

##### E Results difference InMoose vs EdgeR



→ logFC **perfectly correlated** between InMoose and EdgeR

#### Future implementation of other key R-based bioinformatics tools

- Potential next candidate: *glmnet*, a package for regularized regression (Lasso, Ridge, Elastic Net) widely used in outcome prediction, feature selection, and high-dimensional omics modeling

#### Your input matters

- Which tools do you need in Python?
- Which R packages should we prioritize next?



Check it out  
on PyPI

