



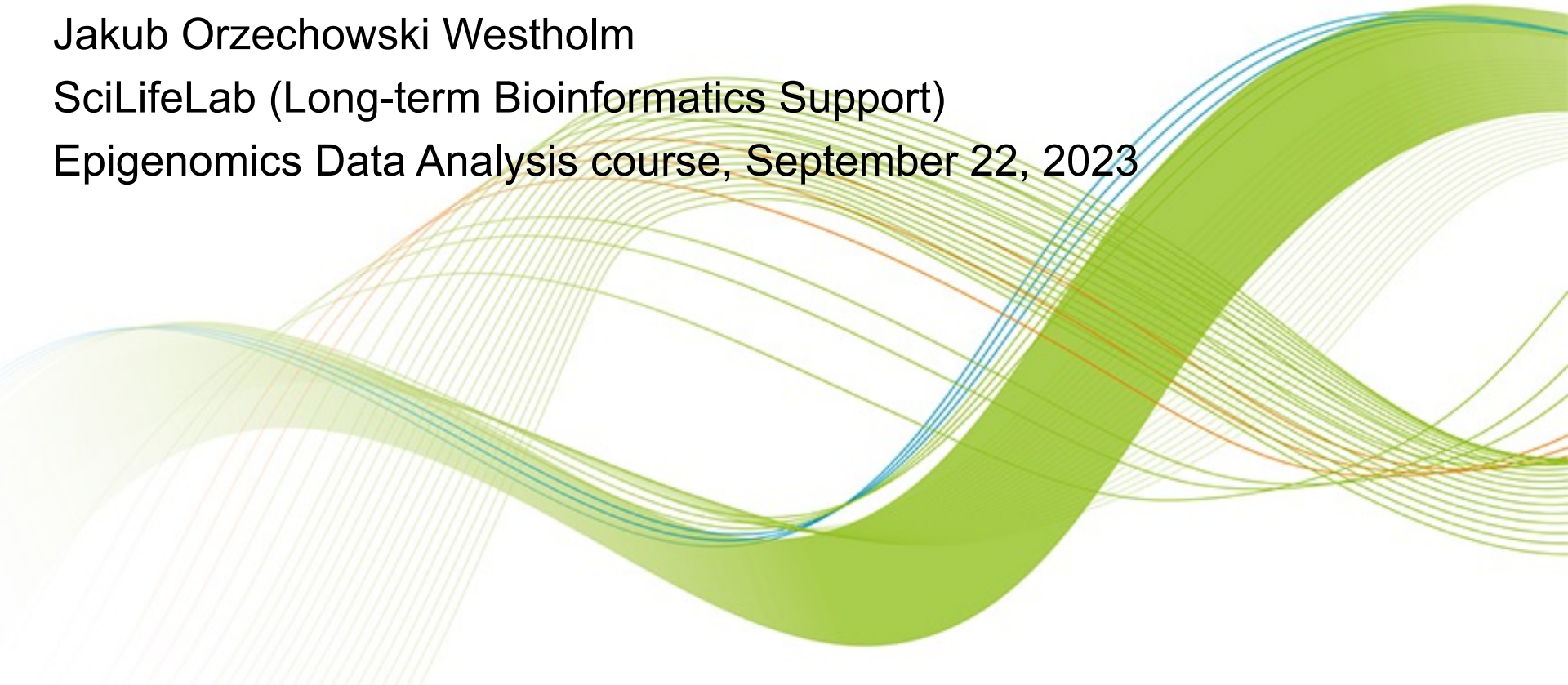
SciLifeLab

Single cell and data integration exercises

Jakub Orzechowski Westholm

SciLifeLab (Long-term Bioinformatics Support)

Epigenomics Data Analysis course, September 22, 2023



- **Single cell ATAC-seq**
 - Learn to process single cell ATAC-seq data:
 - QC, filter cells
 - Normalization, clustering, visualization
 - Motif analysis
 - Find differentially accessible regions
 - Integration with RNA-seq data
 - Quite advanced exercise, may take a while
 - Some steps take a while (around 10 mins) to run



-
- **Genomic overlaps**
 - Learn to:
 - Manipulate genomic ranges
 - Compute overlaps between data sets
 - Check if those overlaps as statistically significant
 - Quite basic exercise

-
- **Unsupervised data integration**
 - Learn to:
 - Use MOFA to integrate single cell ATAC-seq, RNA-seq and bisulphite-seq.
 - Optional exercise, if you have time and interest

- All these are R exercises
- The single cell exercise can be run in three ways
 - On Uppmax using modules.
Recommended. No setting up, everything is already installed. Should work without problems, but you need an Uppmax account.
 - On uppmax, using containers. Good for reproducibility. Also a good option if Rstudio is slow using ssh
 - On your laptop. You need to install R and all packages. Also you need around 13Gb to run the single cell exercise.
- [Read instructions](#), then ask us!





**KEEP
CALM
AND
GOOD
LUCK**