VIB-UGENT CENTER FOR INFLAMMATION RESEARCH

How to: cytometry data preprocessing and quality control

Sarah Bonte

SaeysLab Data mining and Modelling for Biomedicine

GHENT UNIVERSITY

EuroFlow

ISAC INTERNATIONAL SOCIETY FOR ADVANCEMENT OF CYTOMETRY

MB&C course 2026 Leuven
6 February 2026

Automated computational analysis pipeline

Automated computational analysis pipeline

Automated computational analysis pipeline

Tools:

FLOWJO

OMIQ

FCS 7 Express Cytometry

Infinicyt™

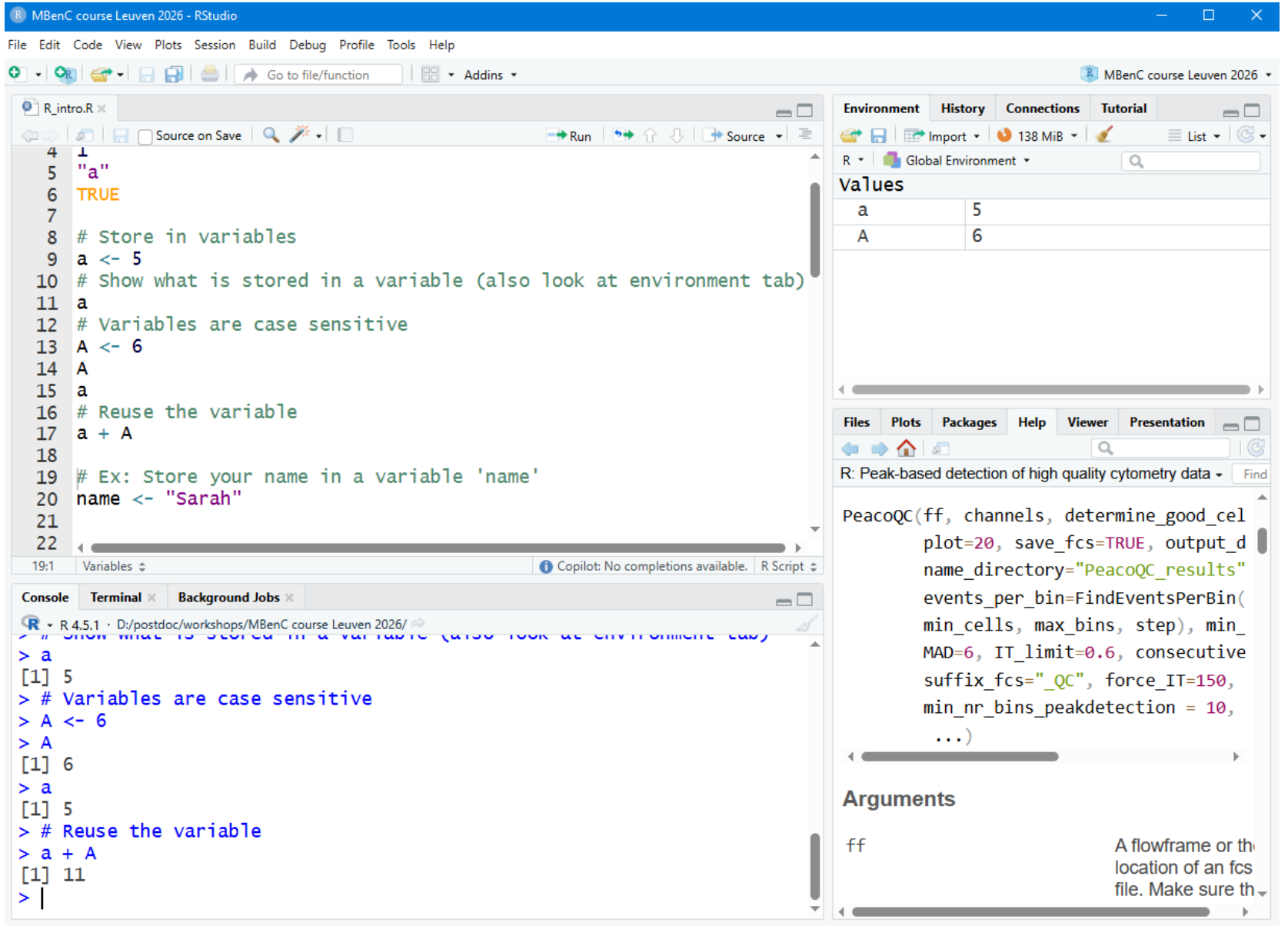
R

R basics

R studio

script:
store code

console:
execute code

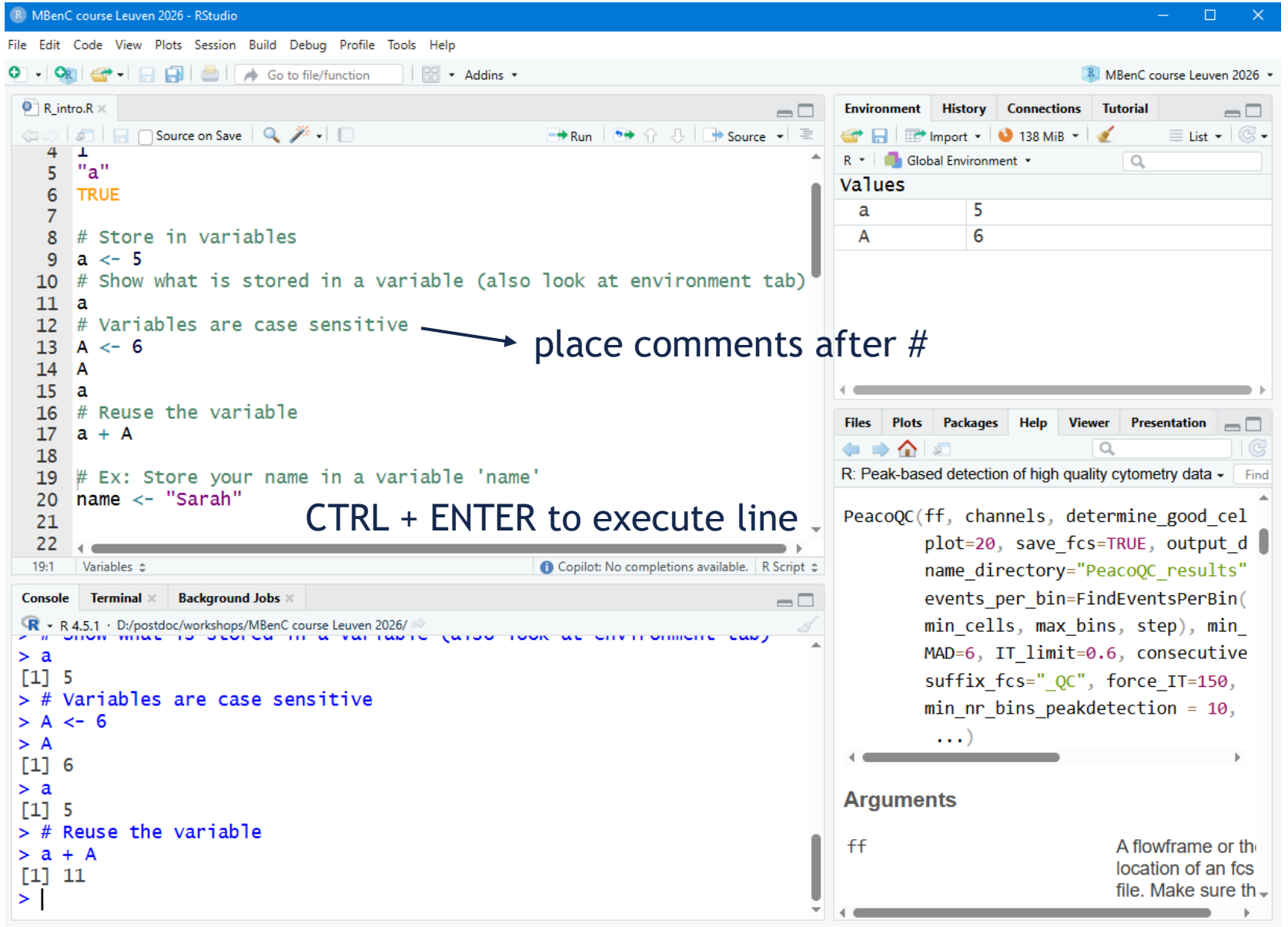


environment:
collection of
user-defined
objects

help:
how to use
functions

R studio

script:
store code



R basics

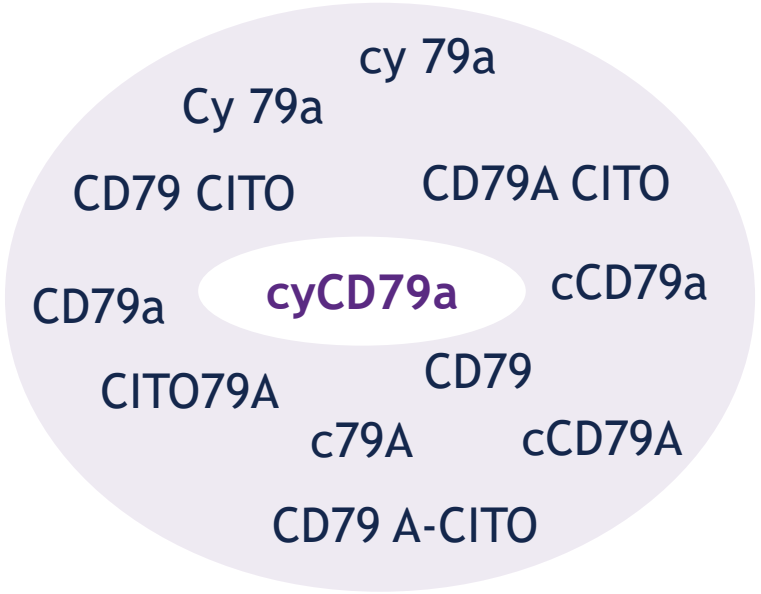
variable, function, for loop, …

Data acquisition

Where it all starts: data acquisition

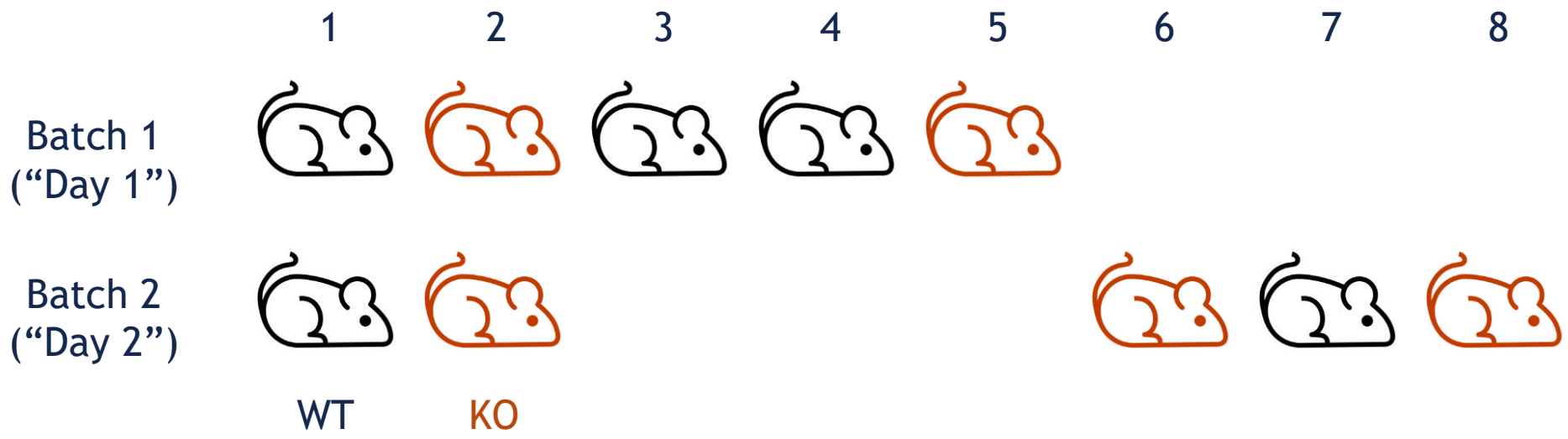
TIP: standardization

- marker names
- channel names
- machine settings
- single stain controls
- …



Dataset introduction

- Spleens from 8 mice
  - 4 WT
  - 4 KO (IKK2 fl/fl CD11c-cre Tg/+, expected autoimmune phenotype)
- 22-color panel
- Experimental batch effect: YG laser power adjusted



many thanks to Victor Bosteels and the VIB Flow core

What's in an FCS file?

.FCS file

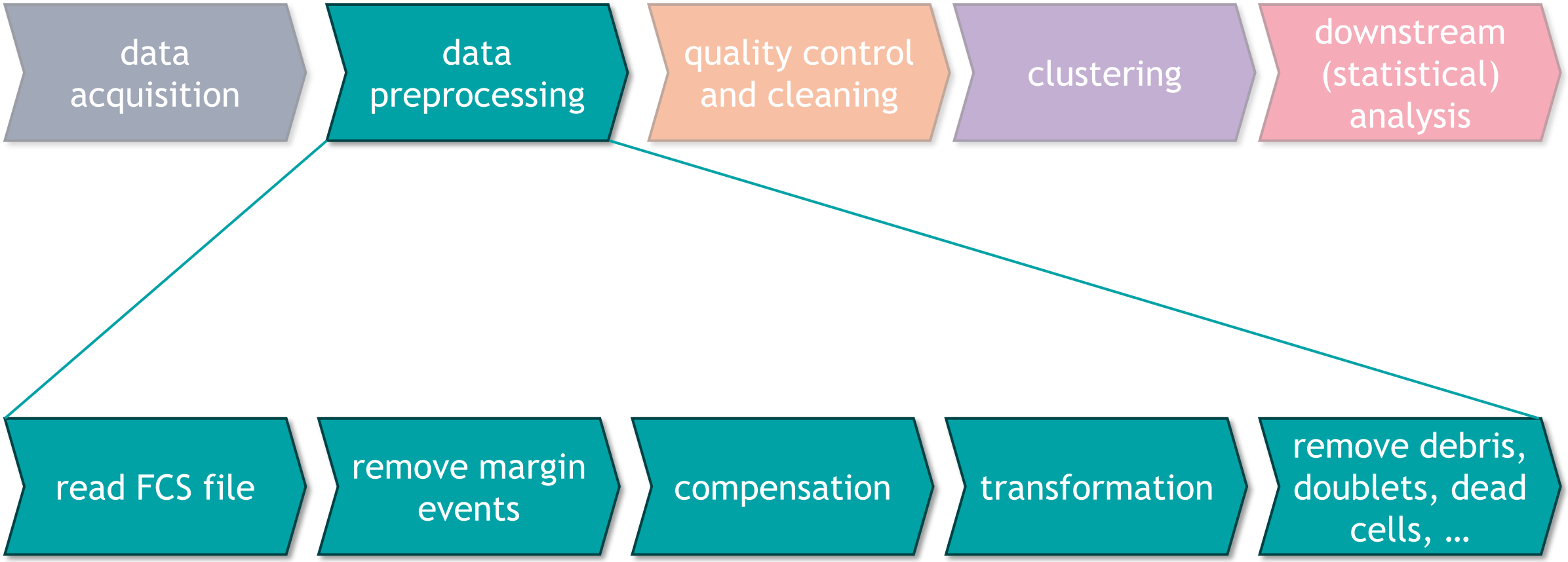
install and load libraries

set and create directories

load in FCS file

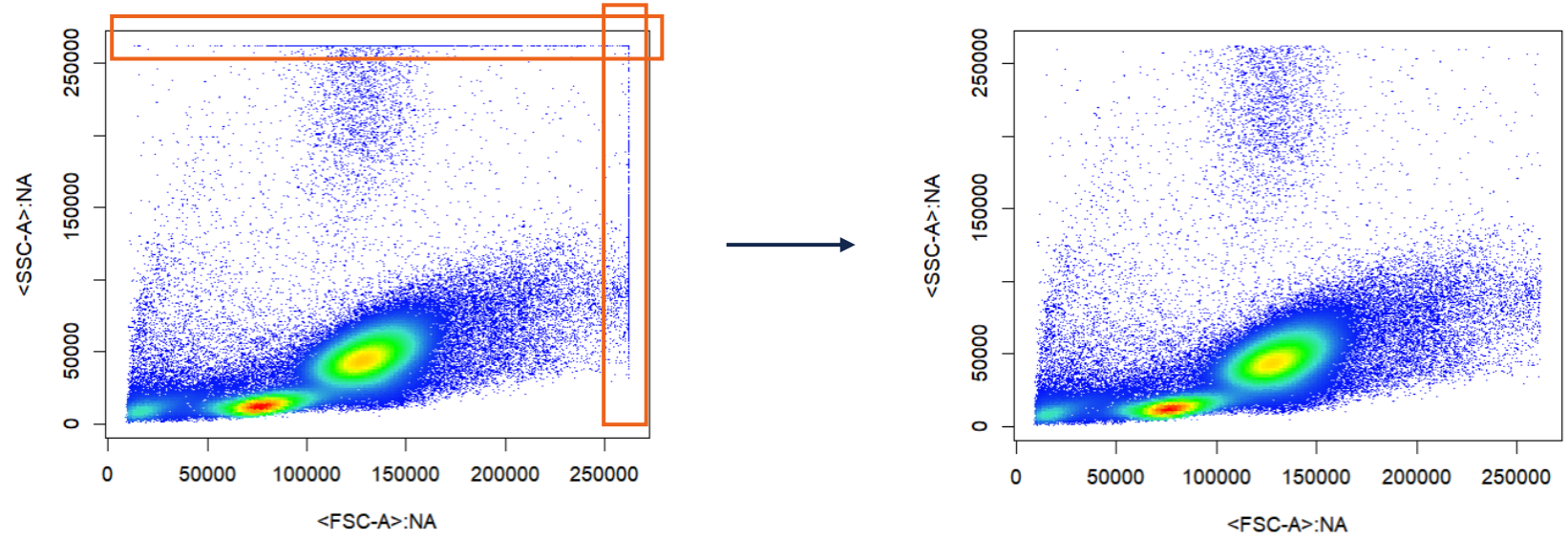
Data preprocessing

Automated computational analysis pipeline



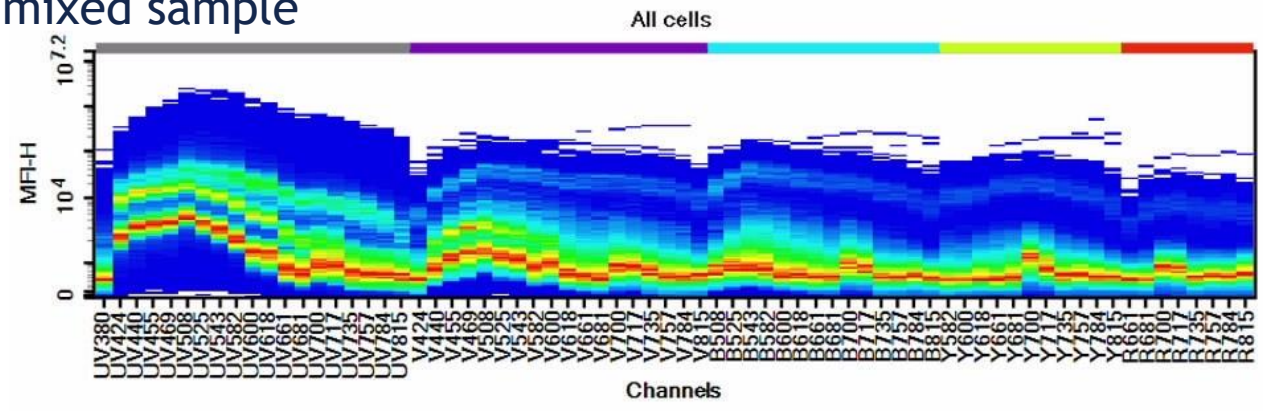
Removal of margin events

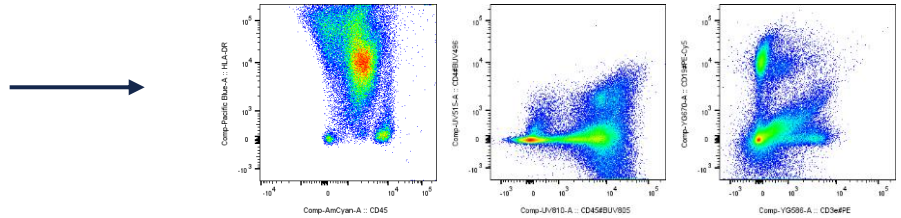
Events that are outside, or at the borders of, the detectable range of the cytometer should be removed

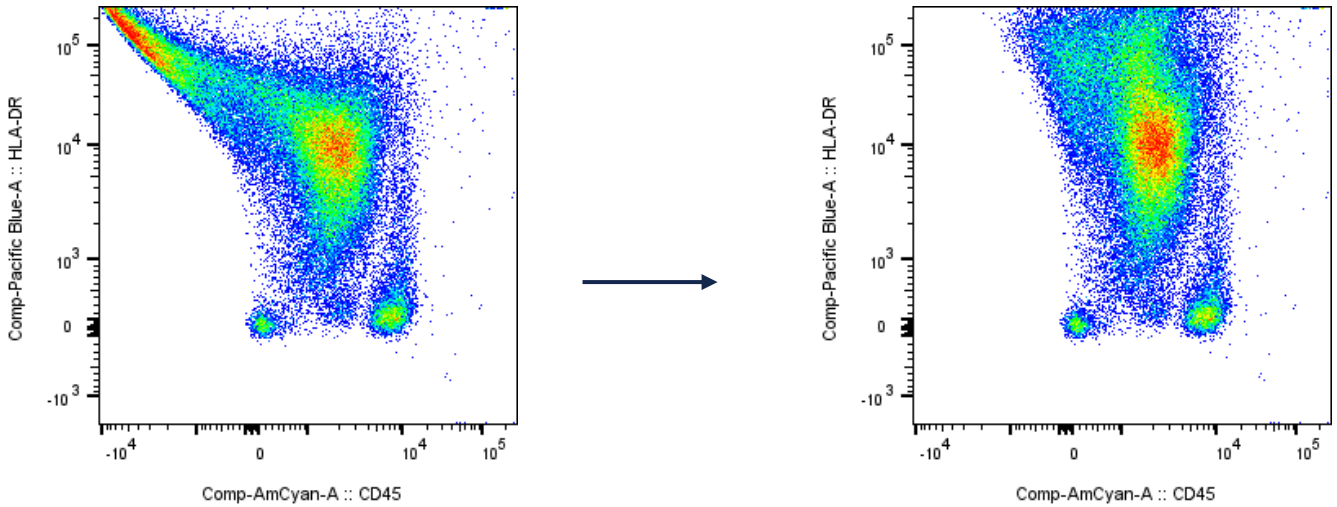


Unmixing and/or compensation

mixed sample







e.g. overcompensation

Unmixing and/or compensation

- Mass cytometry

  No compensation needed,
  unless for small impurities

- Conventional flow cytometry

  Spillover/compensation matrix
  Use either the one stored in the FCS file, or an external one
  Compensation = apply the compensation matrix to the FCS file

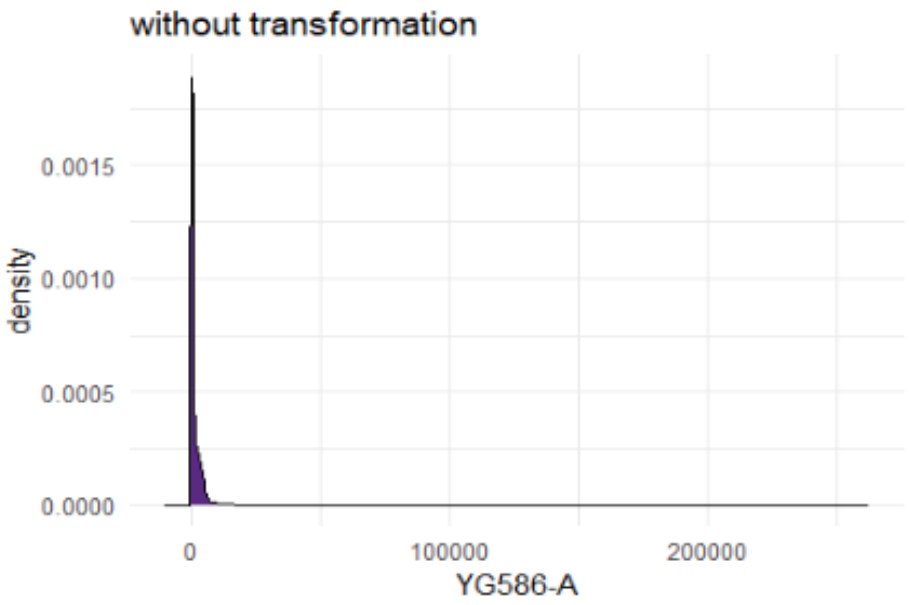
- Spectral flow cytometry

  Unmixing
  Unmixed file requires no further compensation (in theory)

Instrument/vendor specific!
e.g. Cytek adds spillover matrix in the FCS file when you make modifications in the software

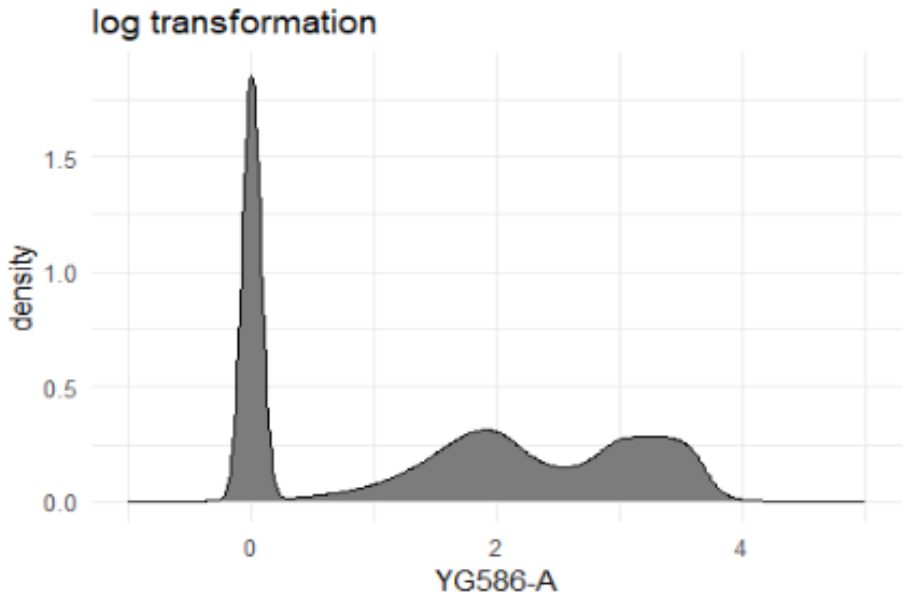
Transformation (data scaling)

- Raw signal intensity values span a large data range (several orders of magnitude)



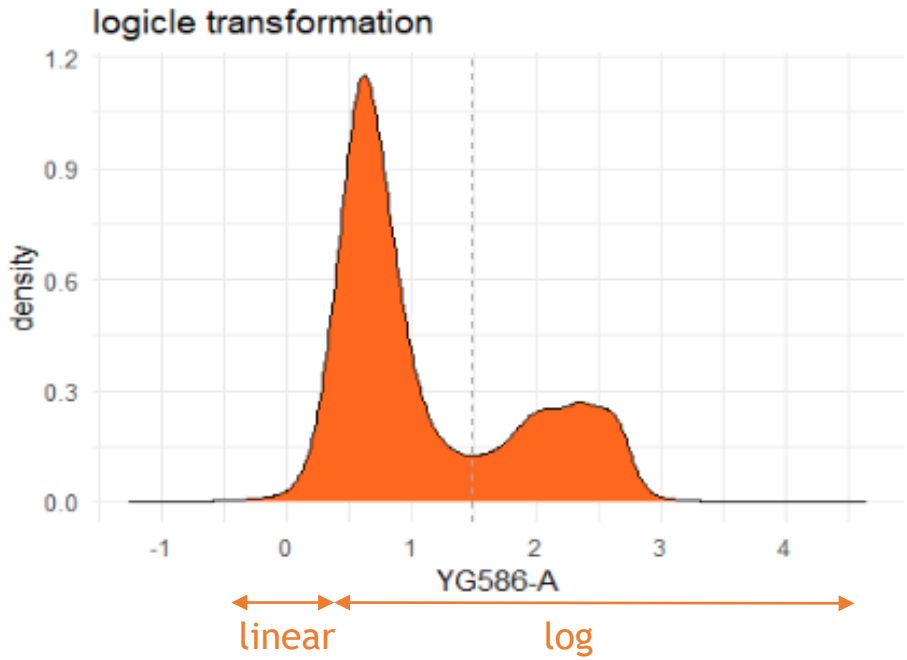
Transformation (data scaling)

- Log transformation works well for high fluorescence intensity values
- However, due to compensation/unmixing, values can be below zero → not possible to log transform



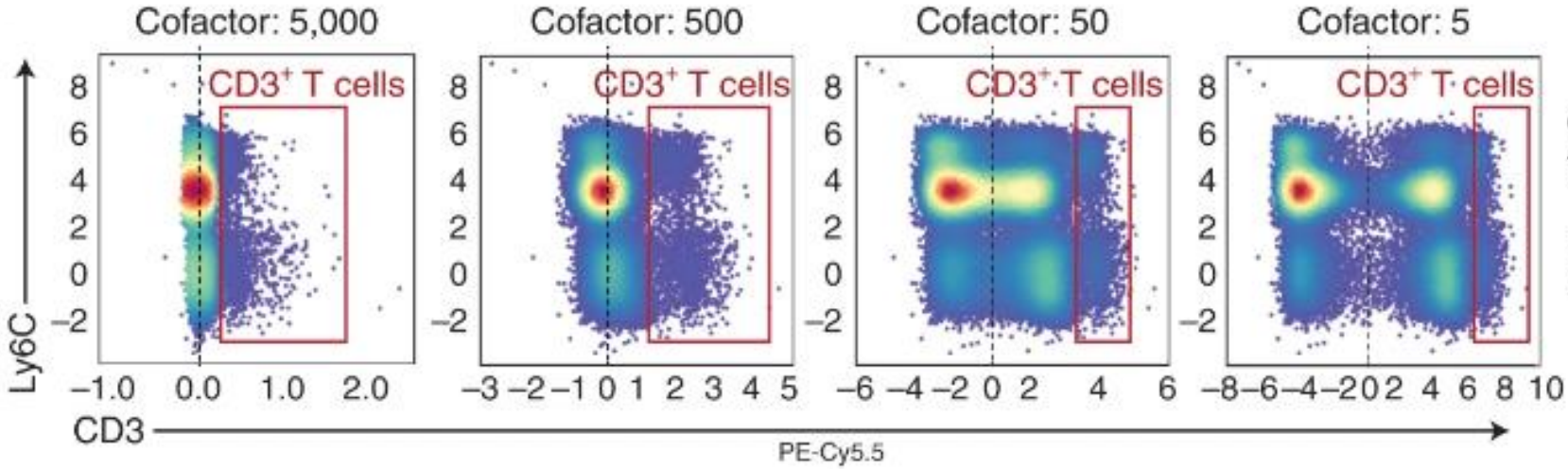
Transformation (data scaling)

- Hybrid scaling, e.g. logicle or arcsinh
- Linear part around zero, rest of the scale behaves logarithmically



Transformation (data scaling)

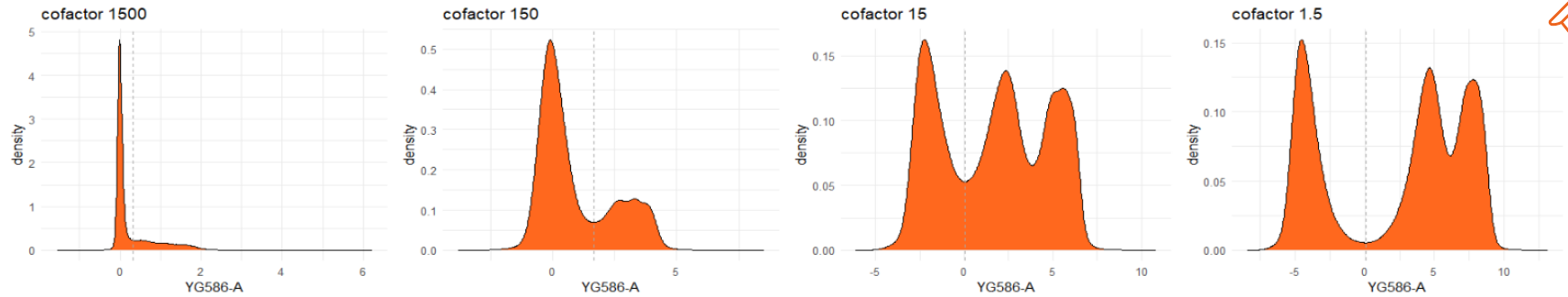
- Hybrid scaling, e.g. logicle or arcsinh
- Width of the linear part is an important parameter choice: arcsinh “cofactor” ~ logicle “width basis”



Cofactor for arcsinh transformation:

FITC: 250
PE-Cy5.5: 5 to 5,000

FLOWJO



negative population is split into two populations centered around zero

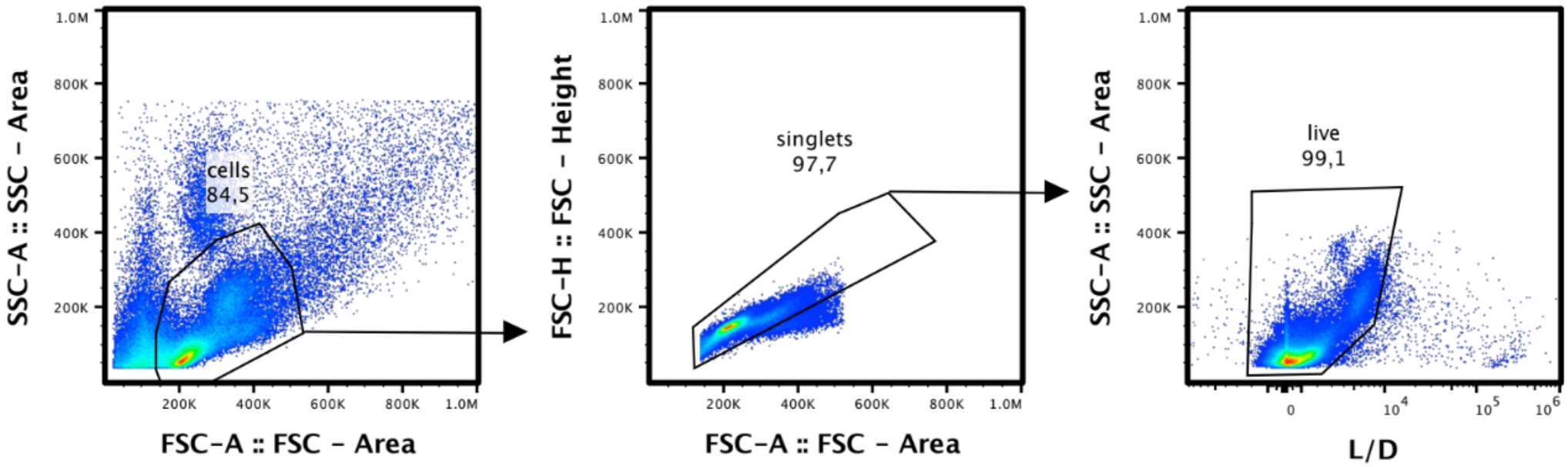
Liechti et al 2021 Nature Immunology

Transformation (data scaling)

- Typical cofactor values
  - 5 for mass cytometry
  - 120 – 1000 for conventional flow cytometry
  - 5000 – 8000 for spectral flow cytometry
    - Cytek: 6000
    - BD: 8000

- **Check and optimize for each individual marker and instrument!**

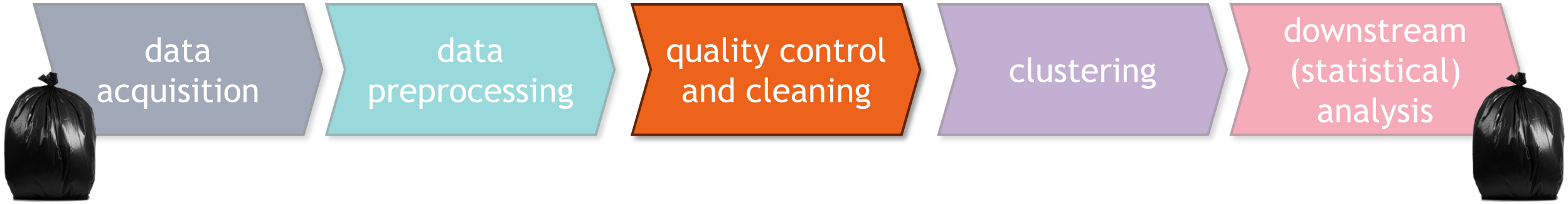
Removal of debris, dead cells, doublets, …



Quality control

Automated computational analysis pipeline

Automated computational analysis pipeline



“garbage in, garbage out”

Quality control at two levels

per file

What can happen? *e.g.*
- clogs during acquisition
- changes in flow rate
- tube runs dry

What to do?
Check the signal consistency over time

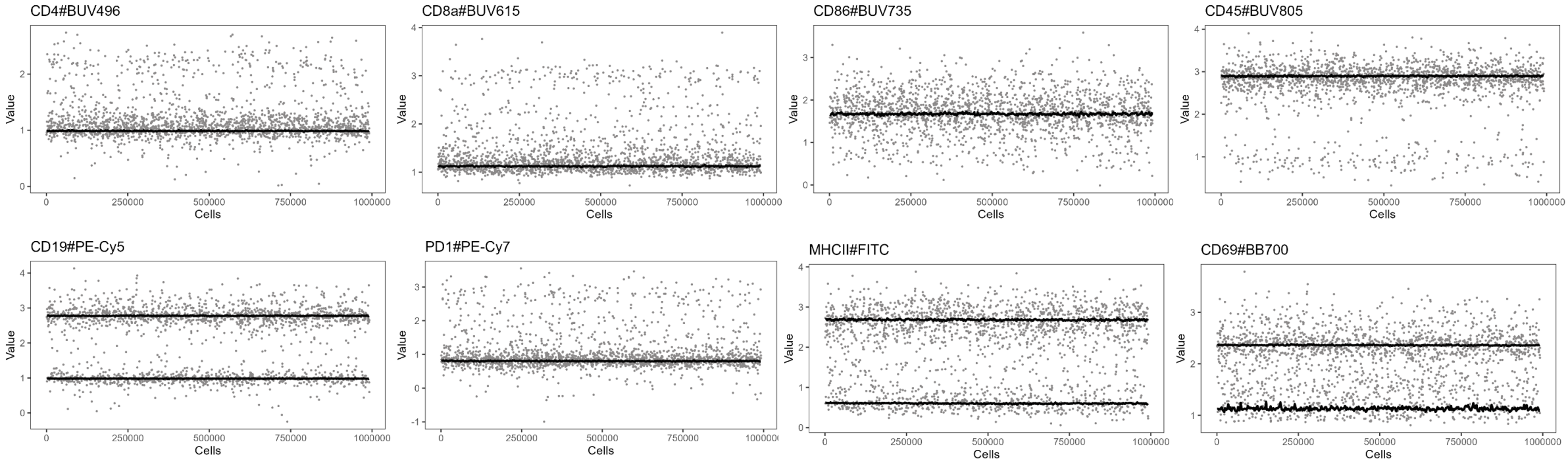
between files

What can happen? *e.g.*
- new antibody batch
- instrument maintenance

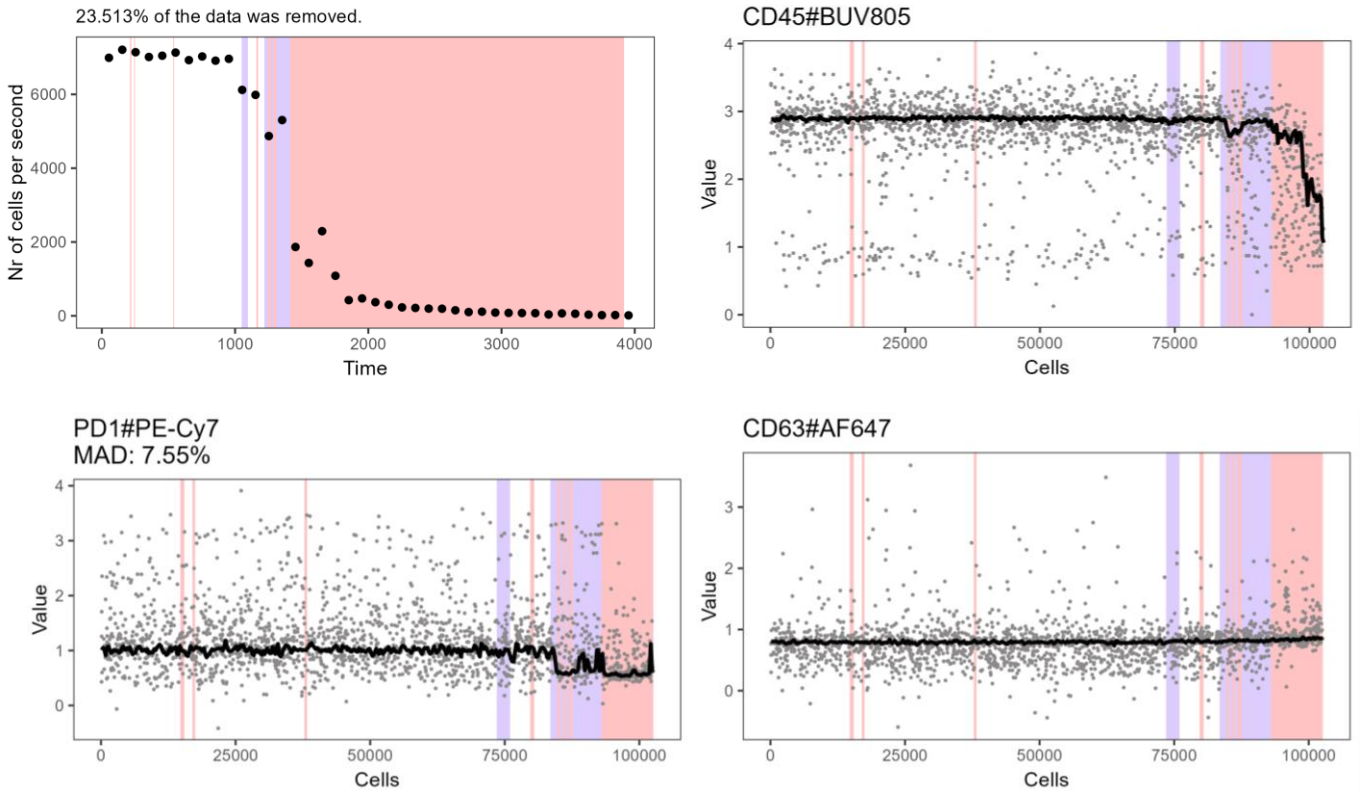
What to do?
Check for batch effects between samples

Quality control of individual files

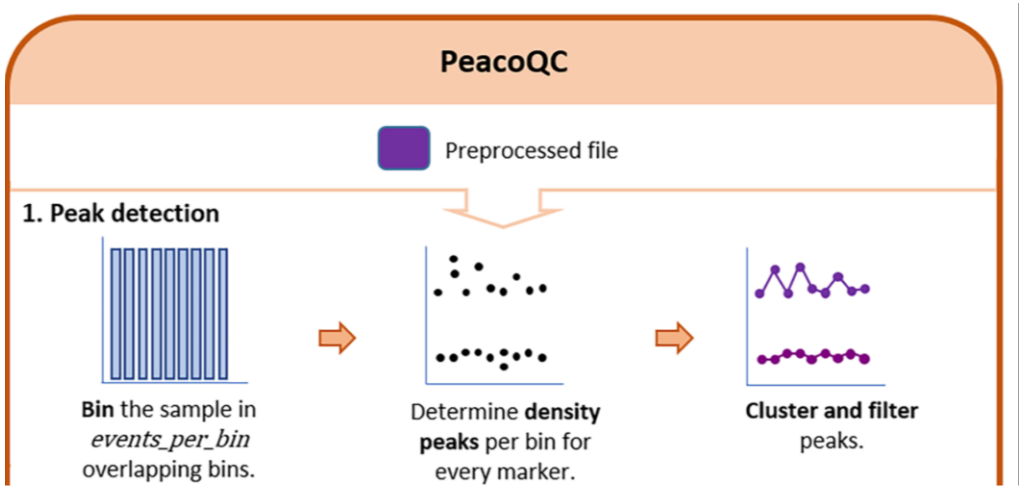
We expect a stable signal over time, as cells go through the system in a random order



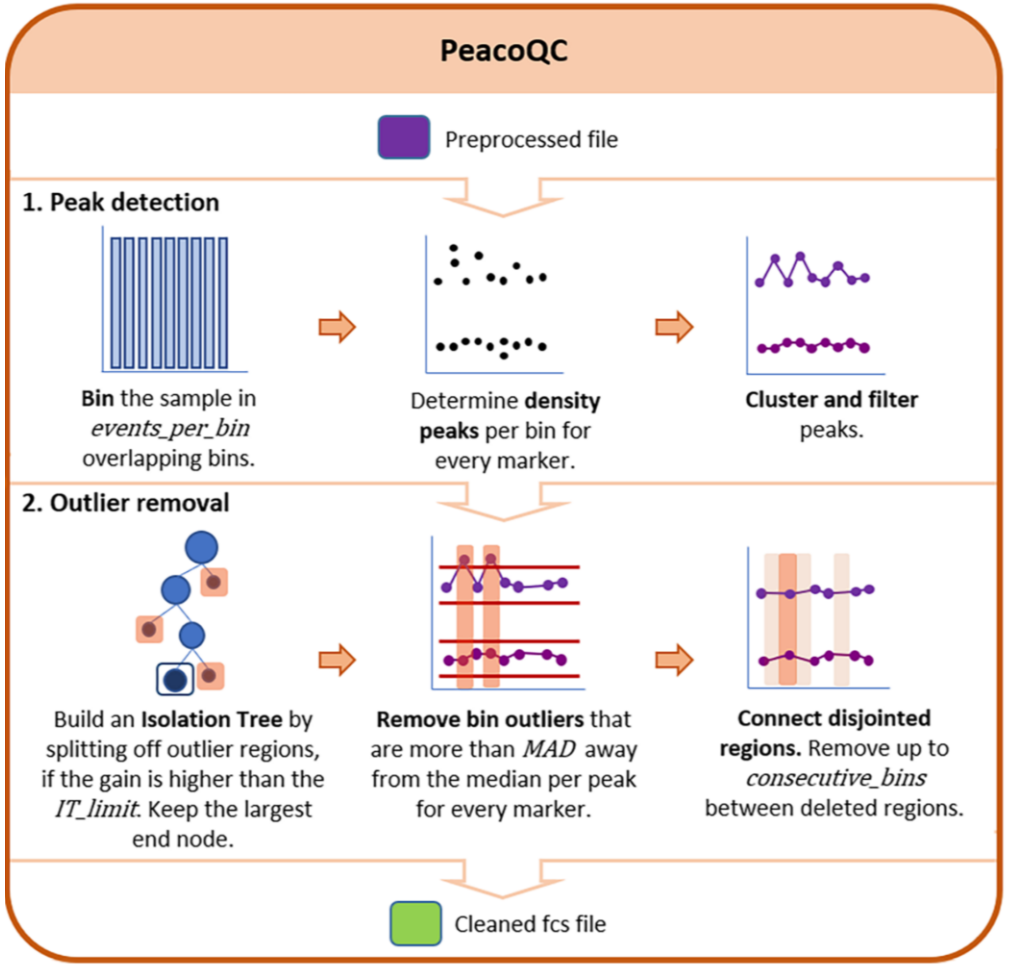
Quality control of individual files



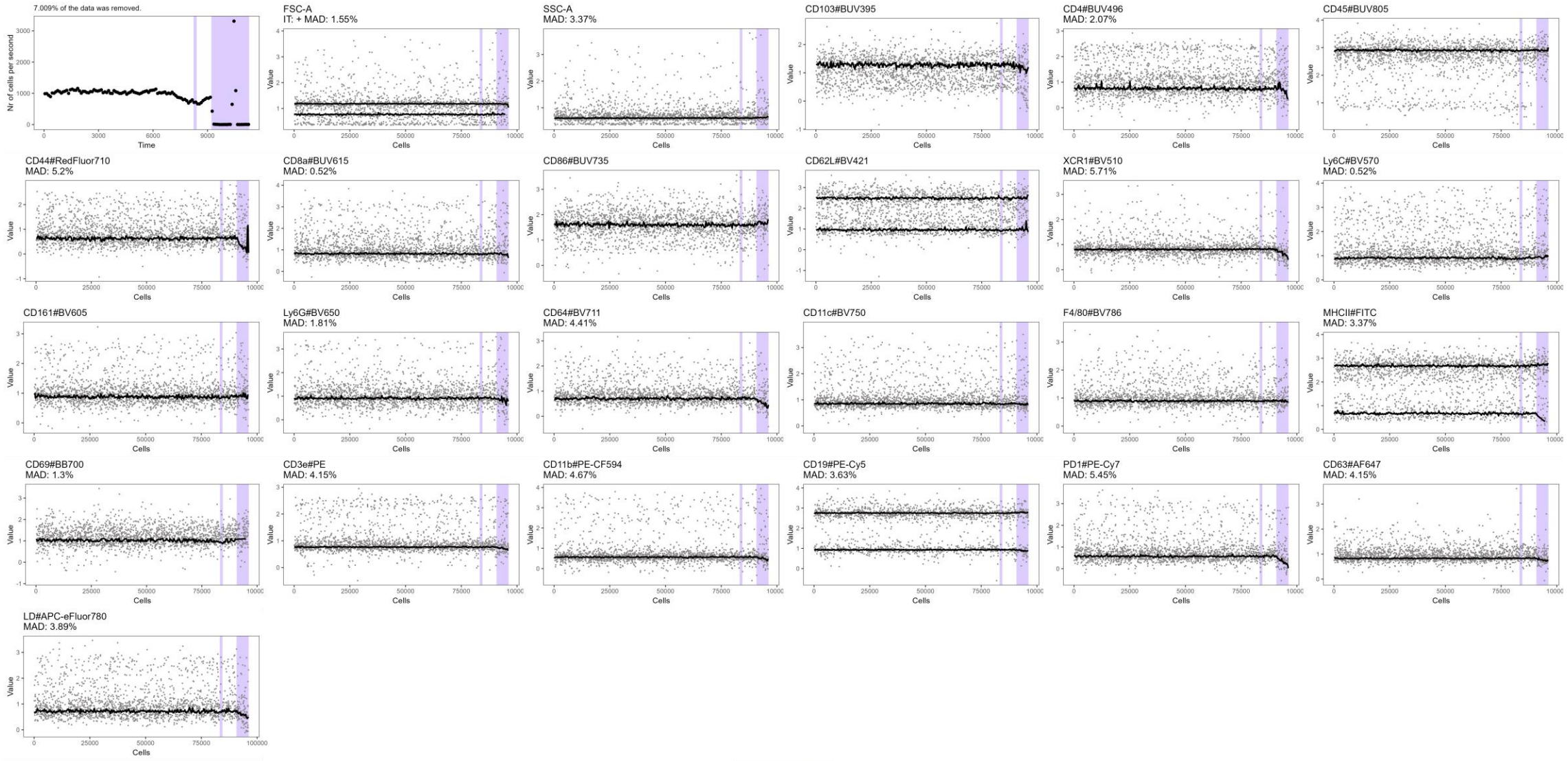
PeacoQC


PeacoQC
Preprocessed file
1. Peak detection
Bin the sample in events_per_bin overlapping bins.
Determine density peaks per bin for every marker.
Cluster and filter peaks.

PeacoQC


PeacoQC
Preprocessed file
1. Peak detection
Bin the sample in events_per_bin overlapping bins.
Determine density peaks per bin for every marker.
Cluster and filter peaks.
2. Outlier removal
Build an Isolation Tree by splitting off outlier regions, if the gain is higher than the IT_limit. Keep the largest end node.
Remove bin outliers that are more than MAD away from the median per peak for every marker.
Connect disjointed regions. Remove up to consecutive_bins between deleted regions.
Cleaned fcs file

PeacoQC

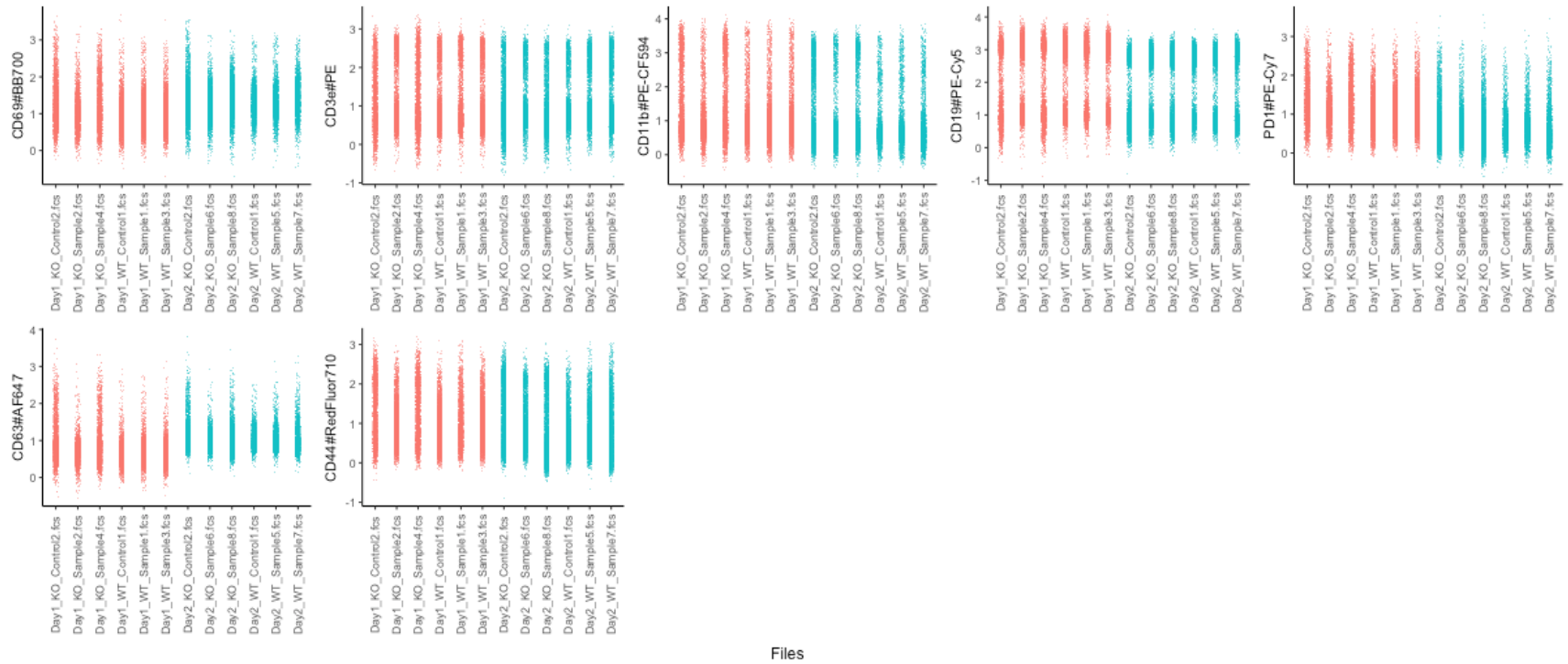

7.009% of the data was removed.
Nr of cells per second
Time
FSC-A
IT: + MAD: 1.55%
SSC-A
MAD: 3.37%
CD103#BUV395
CD4#BUV496
MAD: 2.07%
CD45#BUV805
CD44#RedFluor710
MAD: 5.2%
CD8a#BUV615
MAD: 0.52%
CD86#BUV735
CD62L#BV421
XCR1#BV510
MAD: 5.71%
Ly6C#BV570
MAD: 0.52%
CD161#BV605
Ly6G#BV650
MAD: 1.81%
CD64#BV711
MAD: 4.41%
CD11c#BV750
F4/80#BV786
MHCII#FITC
MAD: 3.37%
CD69#BB700
MAD: 1.3%
CD3e#PE
MAD: 4.15%
CD11b#PE-CF594
MAD: 4.67%
CD19#PE-Cy5
MAD: 3.63%
PD1#PE-Cy7
MAD: 5.45%
CD63#AF647
MAD: 4.15%
LD#APC-eFluor780
MAD: 3.89%
Value
Cells
Good Values
MAD

Preprocessing + quality control

Batch effects

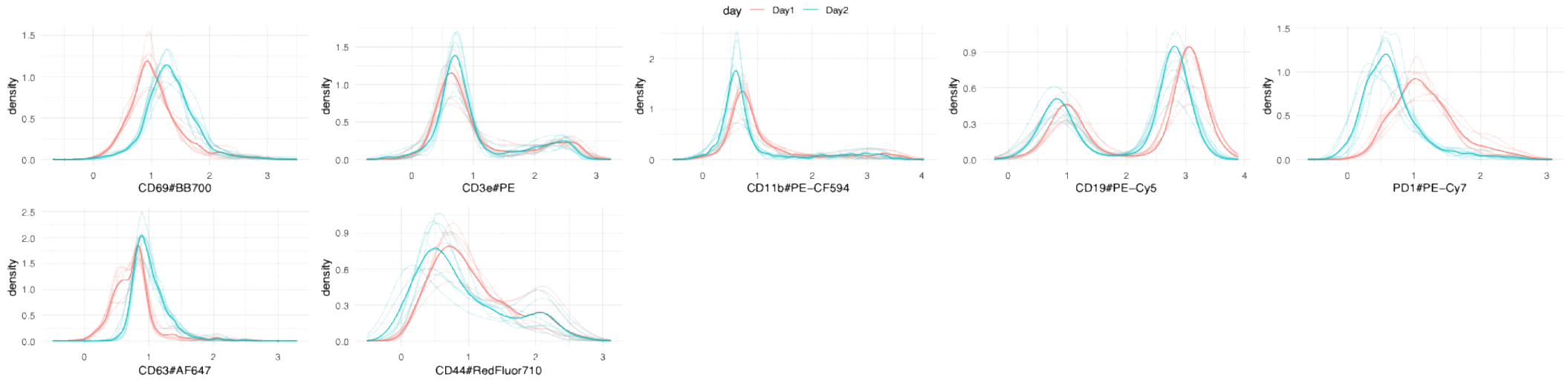
Between-files QC: check for batch effects

- Visual assessment: file scatter plots



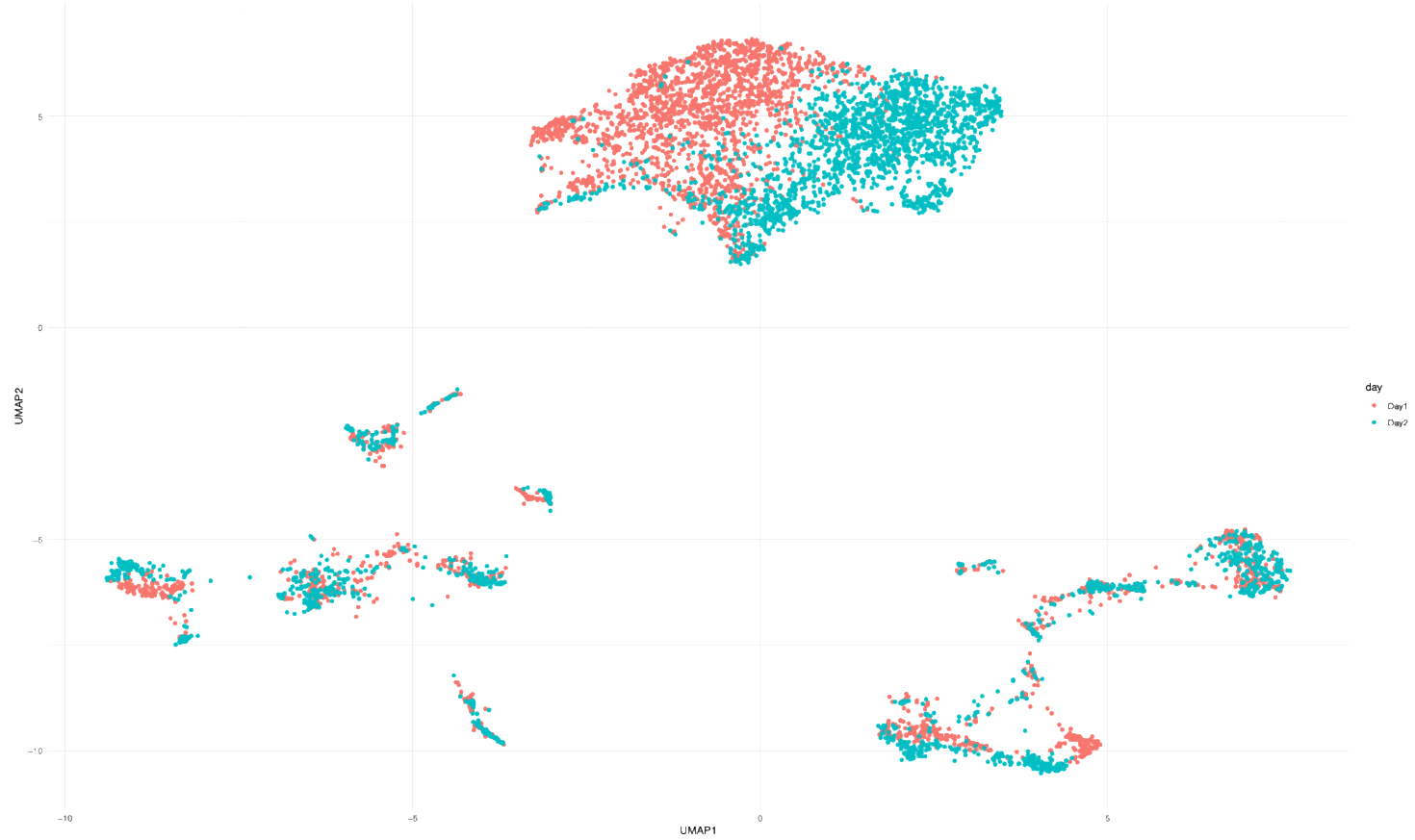
Between-files QC: check for batch effects

- Visual assessment: density plots



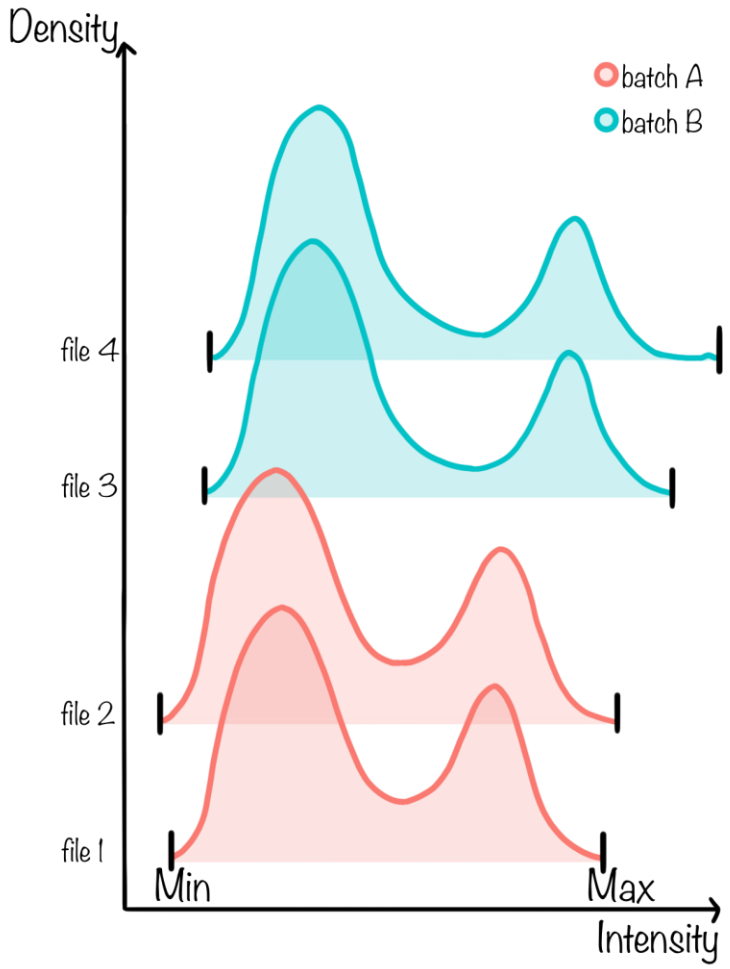
Between-files QC: check for batch effects

- Visual assessment: UMAP plot



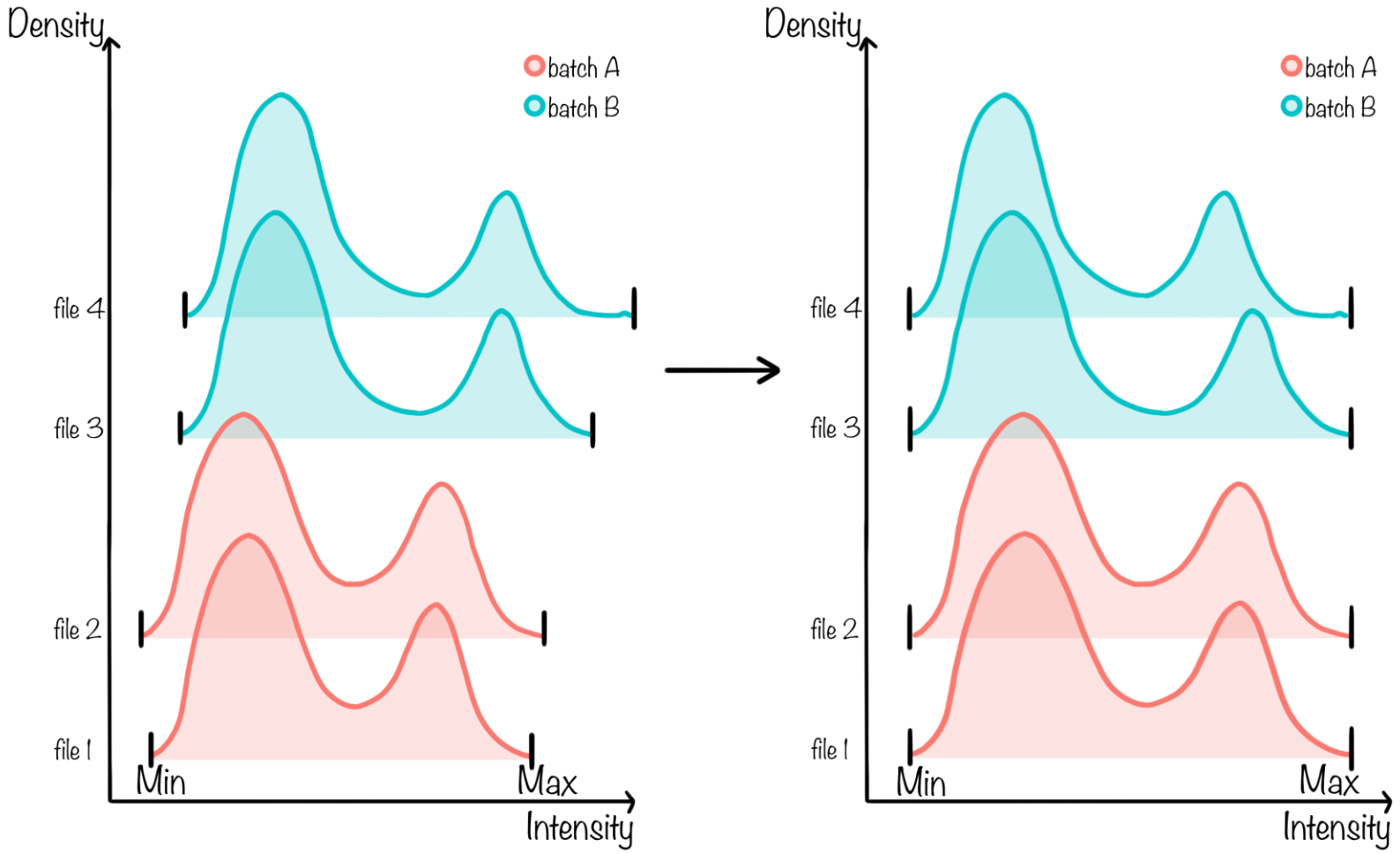
Batch effects: visualizations

Batch effect correction: MinMax normalization



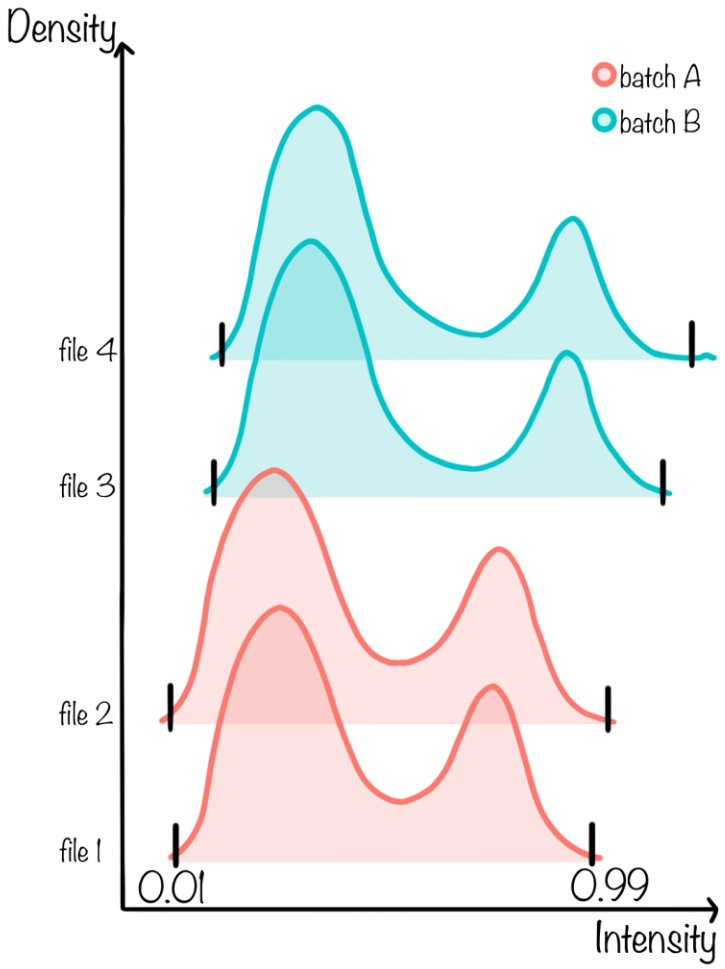
Figures: Katrien Quintelier

Batch effect correction: MinMax normalization



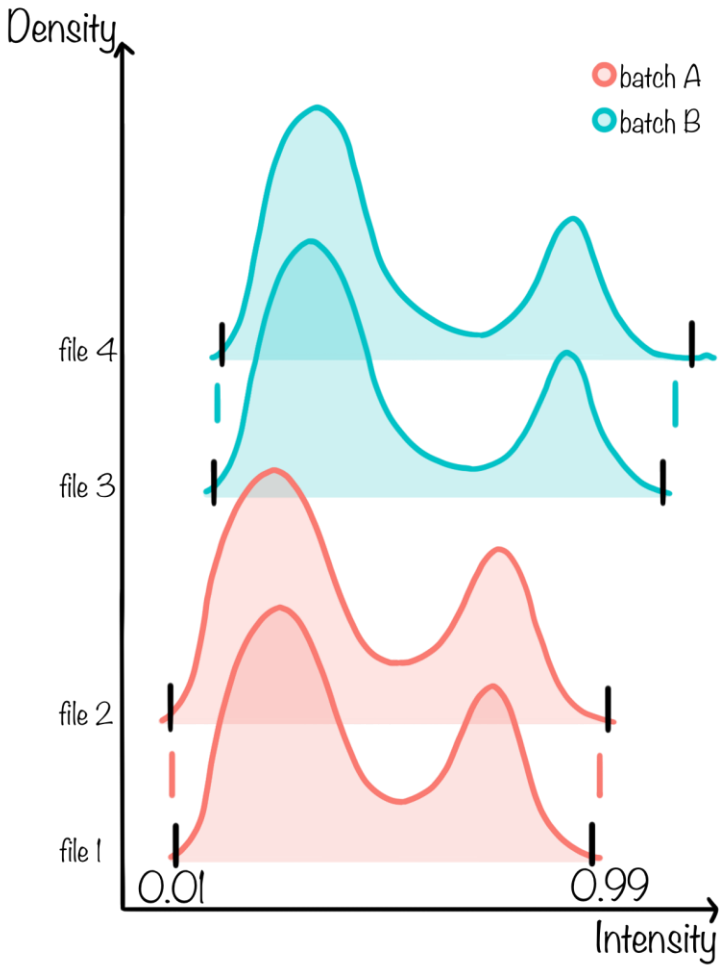
Figures: Katrien Quintelier

Batch effect correction: Quantile normalization



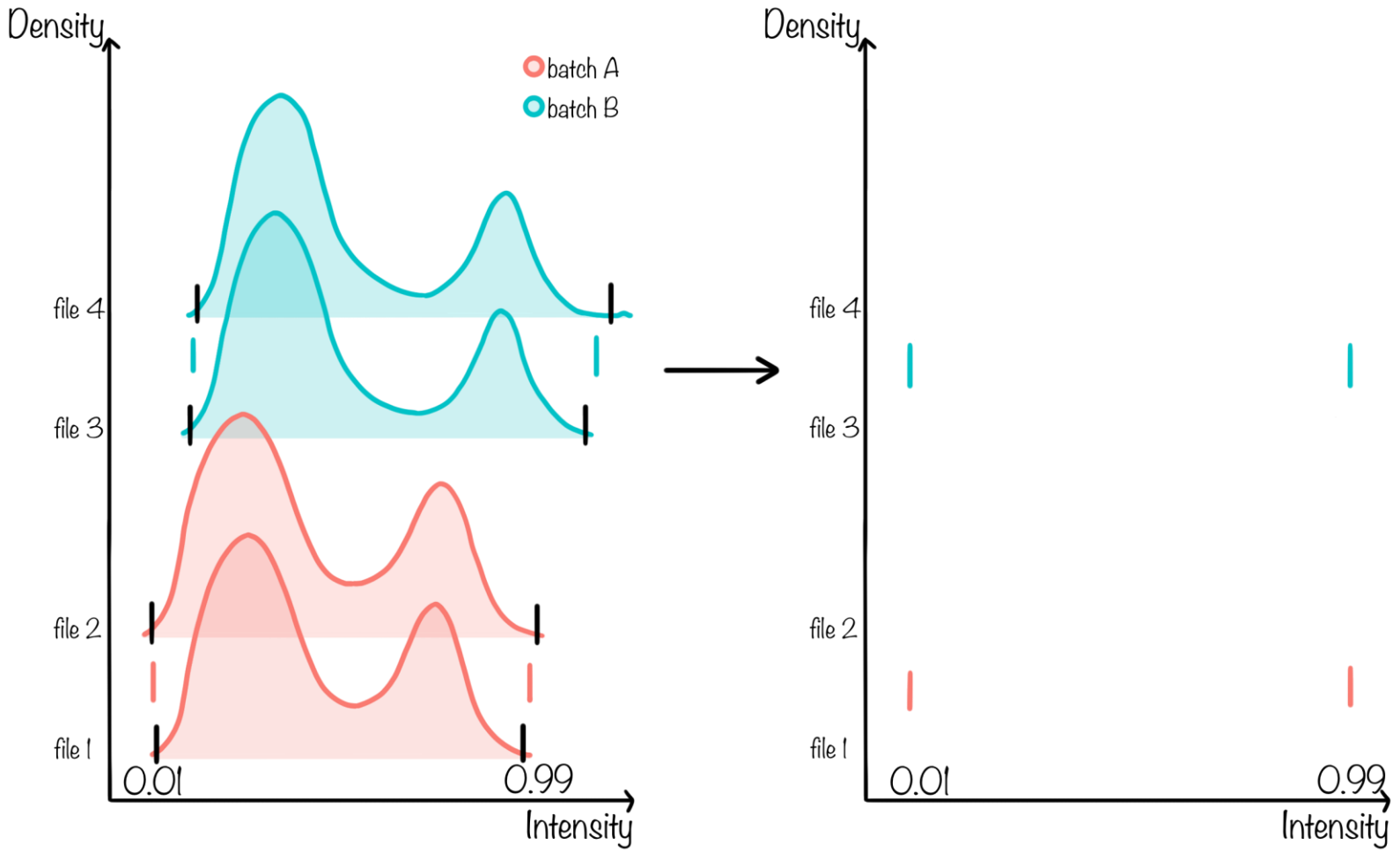
Figures: Katrien Quintelier

Batch effect correction: Quantile normalization



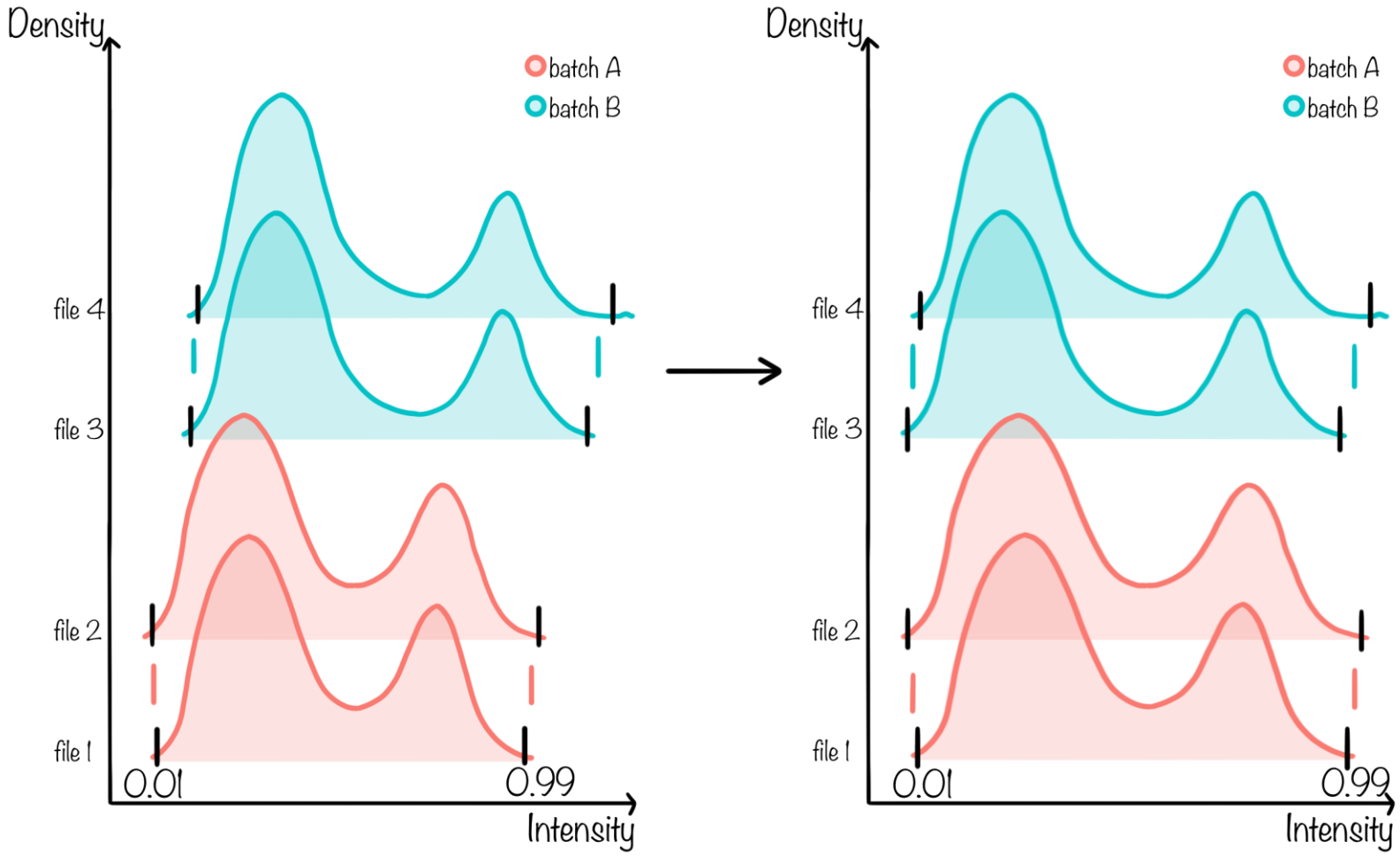
Figures: Katrien Quintelier

Batch effect correction: Quantile normalization



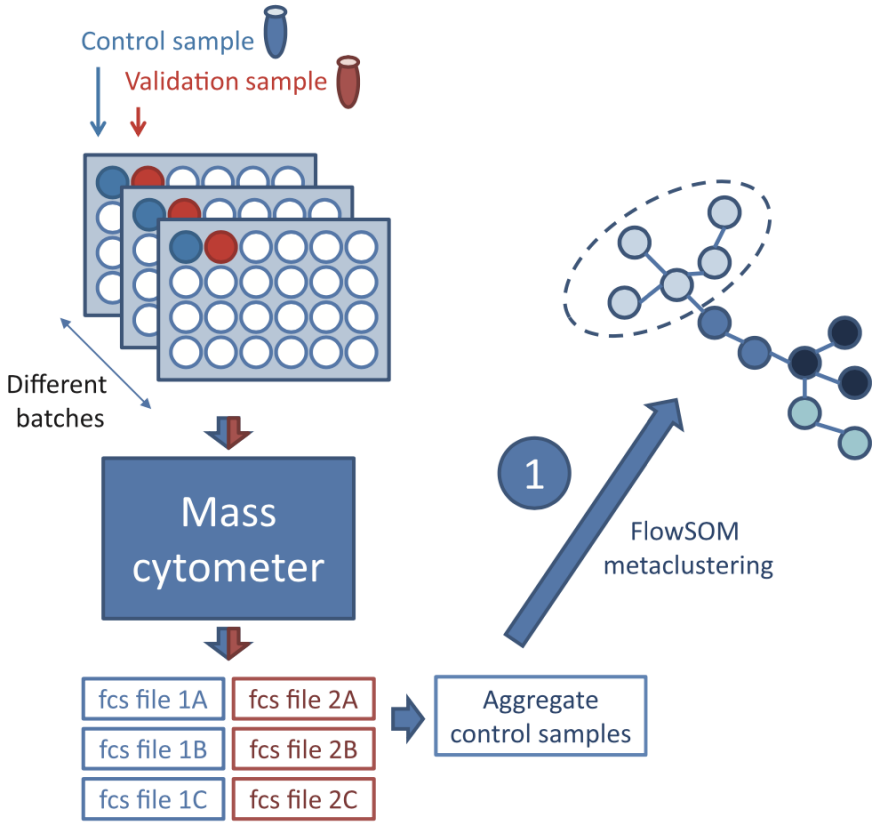
Figures: Katrien Quintelier

Batch effect correction: Quantile normalization



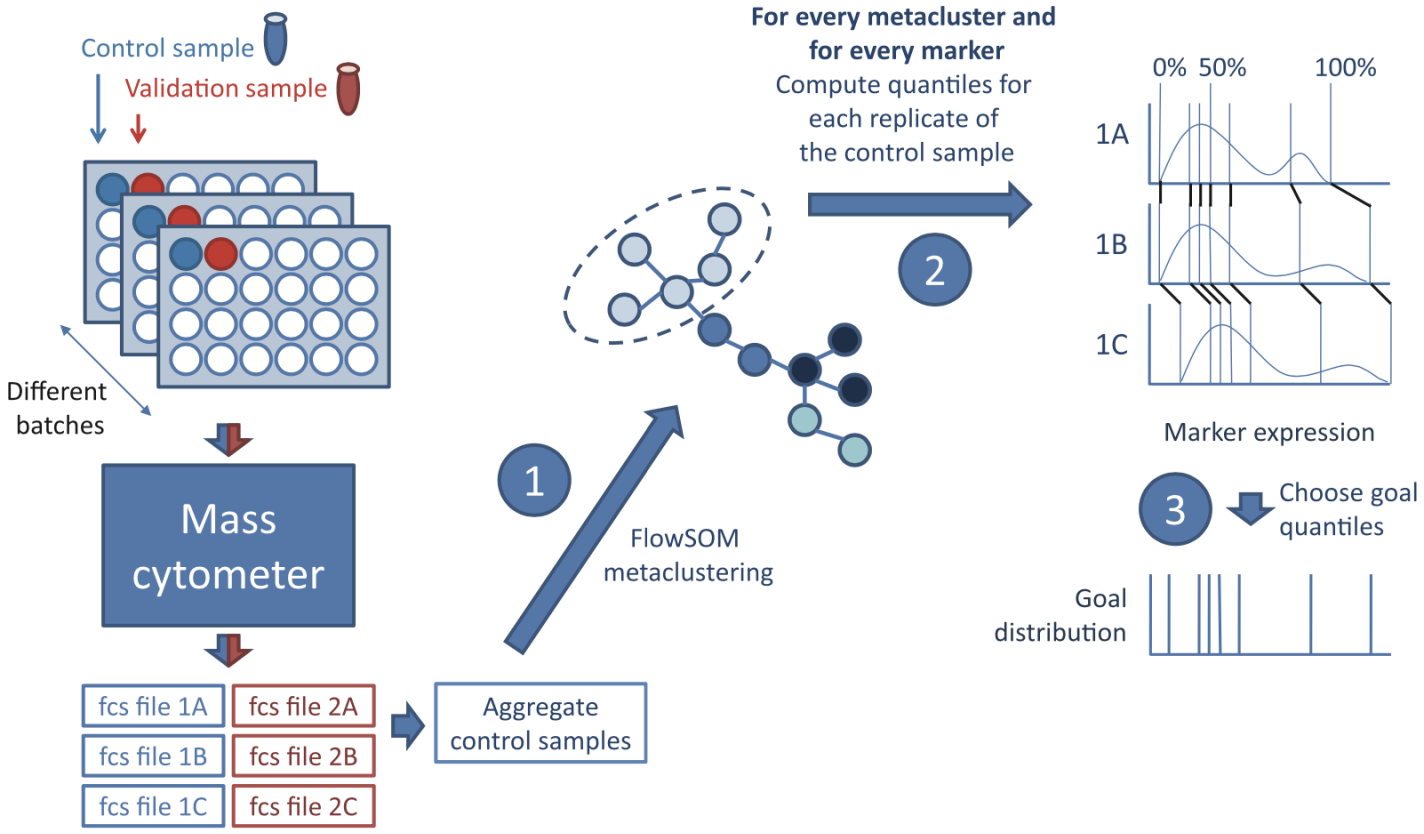
Figures: Katrien Quintelier

Batch effect correction: CytoNorm



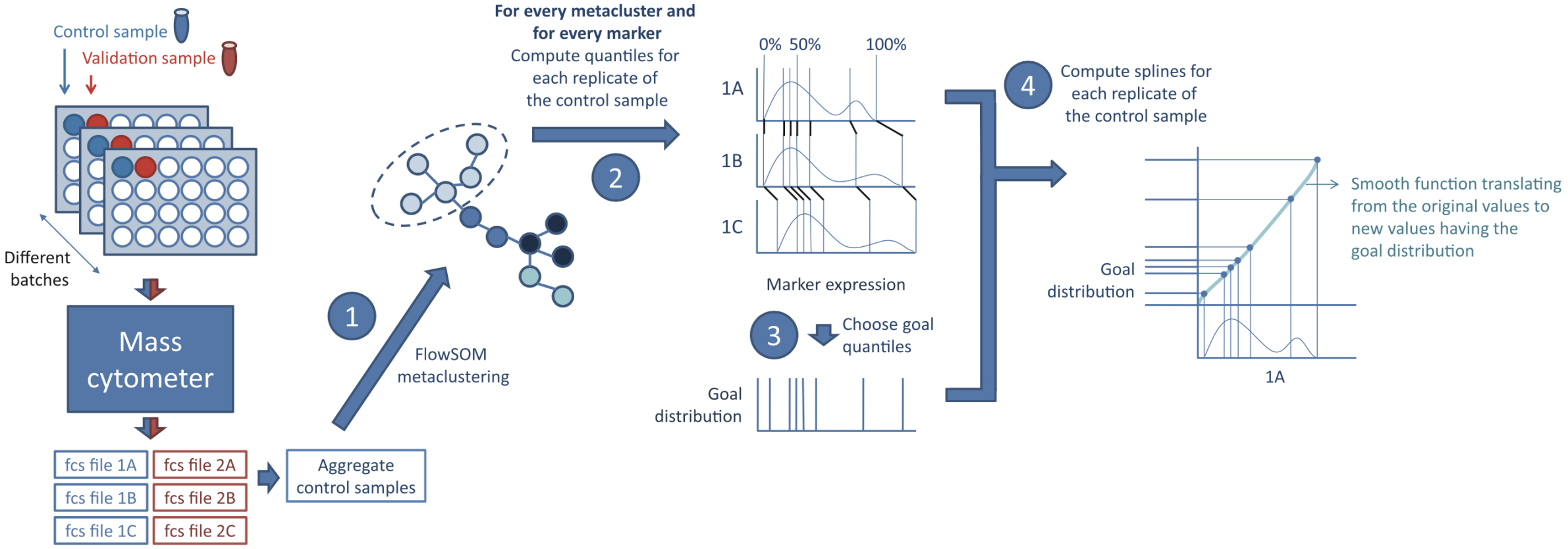
Van Gassen S et al. (2019). CytoNorm: A Normalization Algorithm for Cytometry Data. Cytometry part A, 97(3):268-278.
Quintelier K et al. (2025). CytoNorm 2.0: A flexible normalization framework for cytometry data without requiring controls. Cytometry part A, 107(2):69-87.

Batch effect correction: CytoNorm



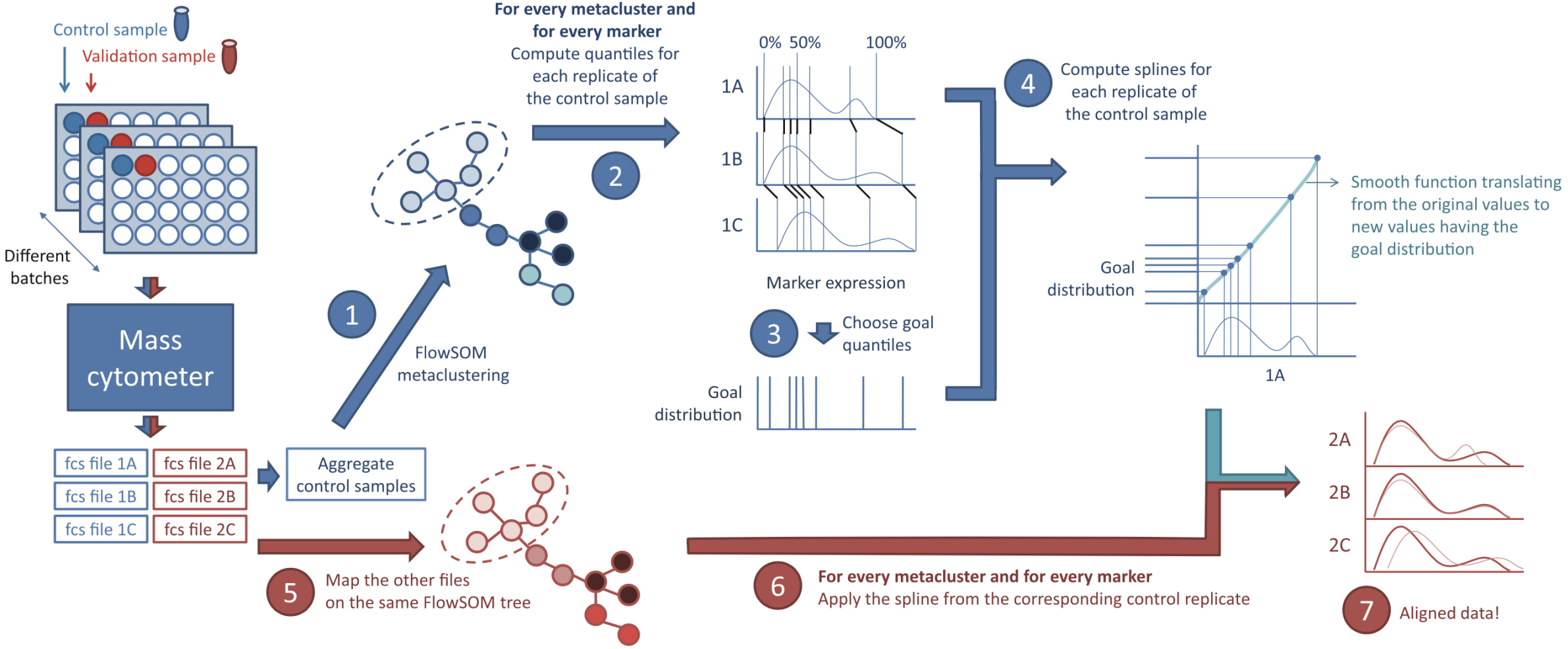
Van Gassen S et al. (2019). CytoNorm: A Normalization Algorithm for Cytometry Data. Cytometry part A, 97(3):268-278.
Quintelier K et al. (2025). CytoNorm 2.0: A flexible normalization framework for cytometry data without requiring controls. Cytometry part A, 107(2):69-87.

Batch effect correction: CytoNorm



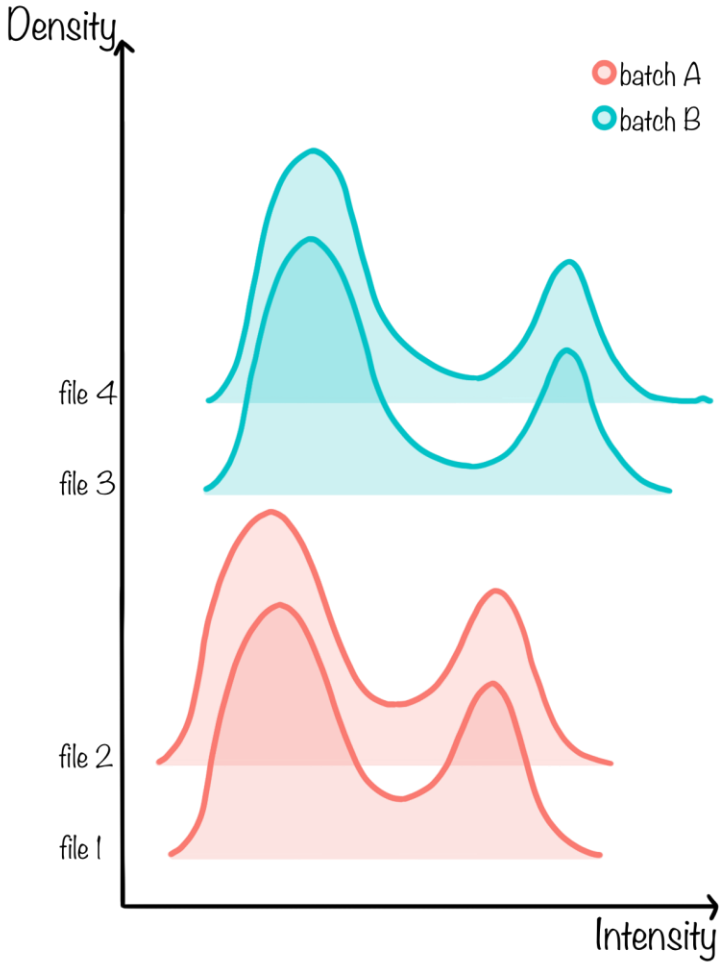
Van Gassen S et al. (2019). CytoNorm: A Normalization Algorithm for Cytometry Data. Cytometry part A, 97(3):268-278.
Quintelier K et al. (2025). CytoNorm 2.0: A flexible normalization framework for cytometry data without requiring controls. Cytometry part A, 107(2):69-87.

Batch effect correction: CytoNorm



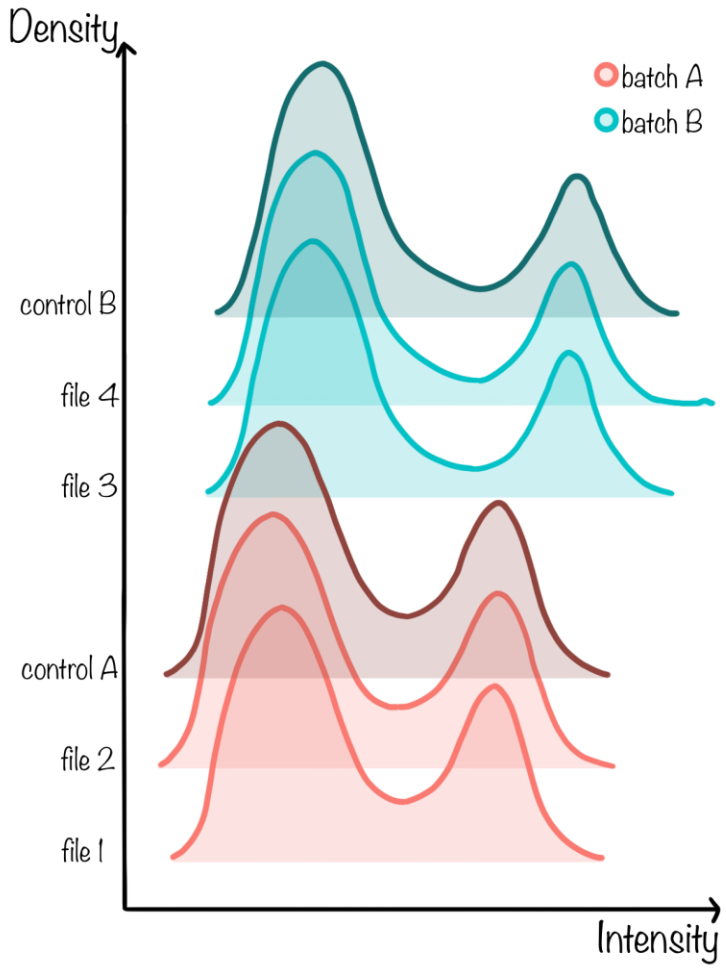
Van Gassen S et al. (2019). CytoNorm: A Normalization Algorithm for Cytometry Data. Cytometry part A, 97(3):268-278.
Quintelier K et al. (2025). CytoNorm 2.0: A flexible normalization framework for cytometry data without requiring controls. Cytometry part A, 107(2):69-87.

Batch effect correction: CytoNorm



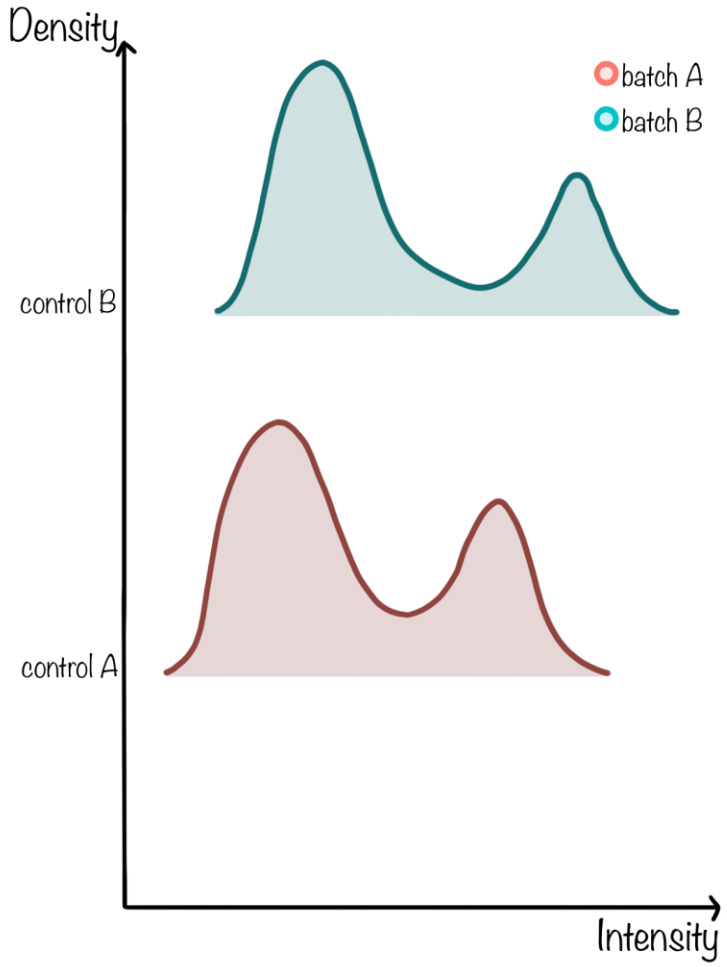
Figures: Katrien Quintelier

Batch effect correction: CytoNorm



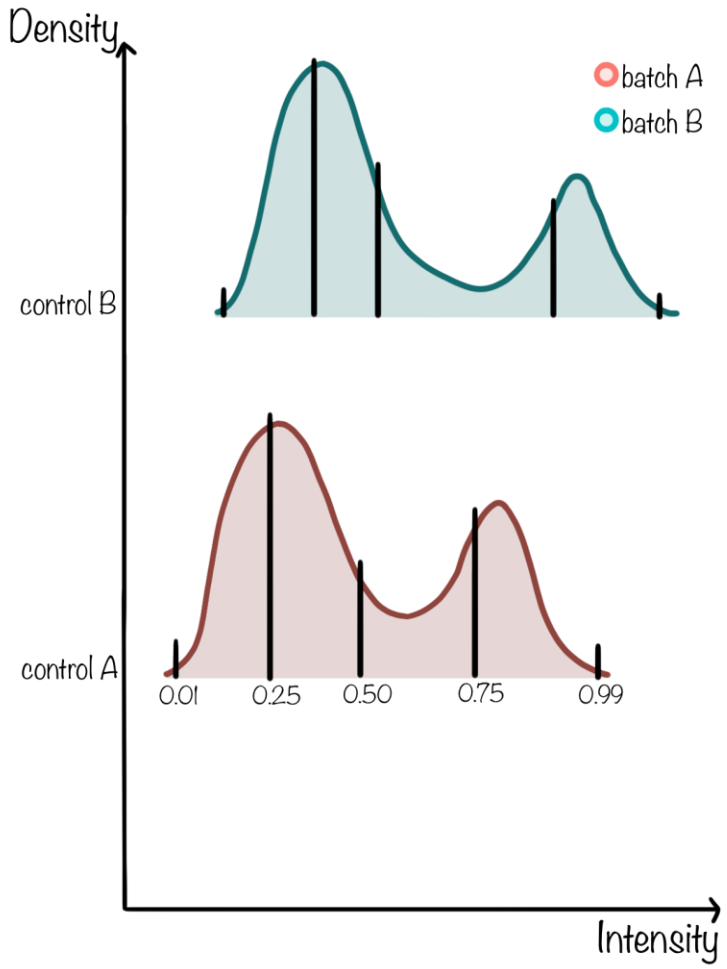
Figures: Katrien Quintelier

Batch effect correction: CytoNorm



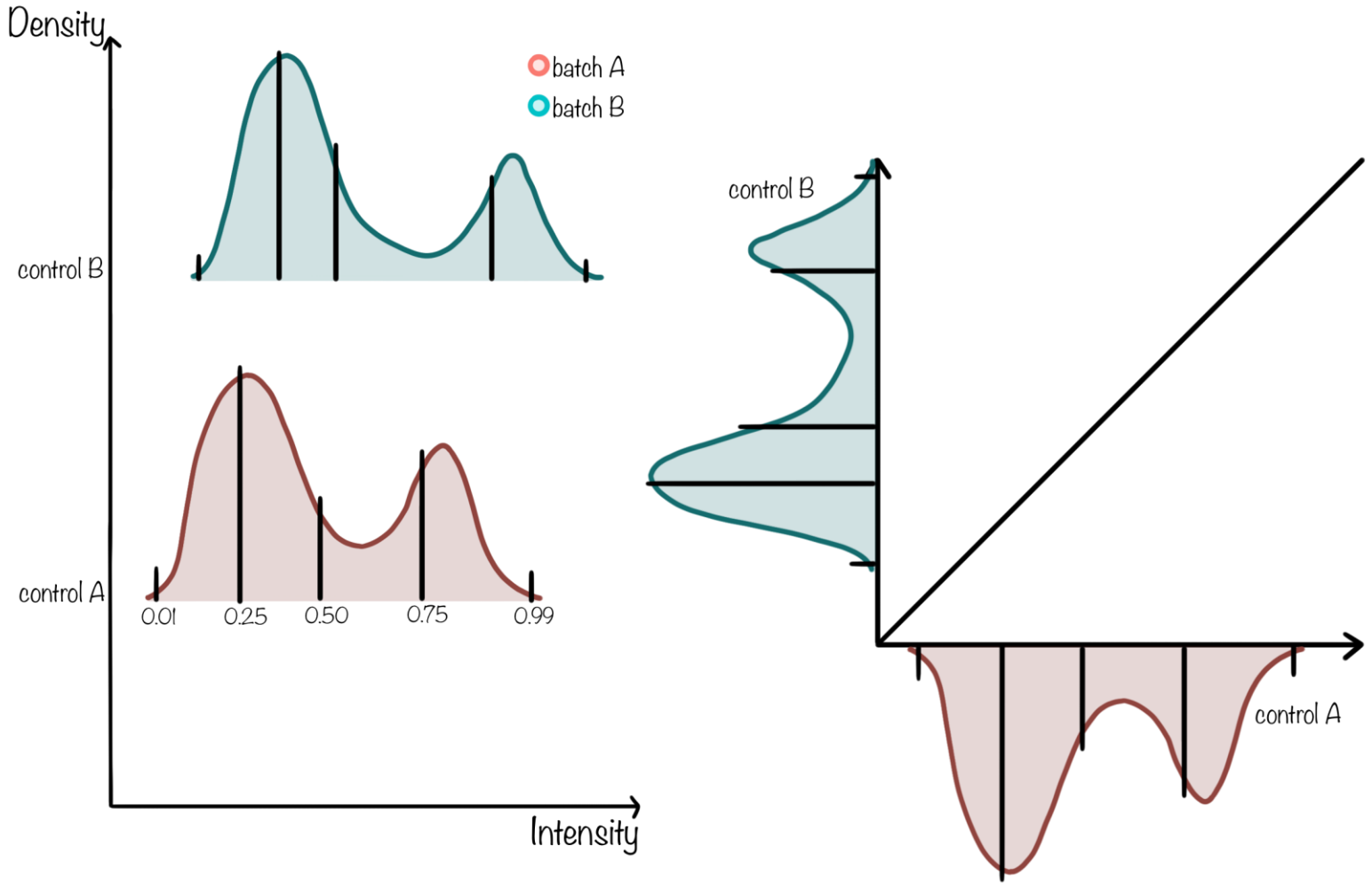
Figures: Katrien Quintelier

Batch effect correction: CytoNorm



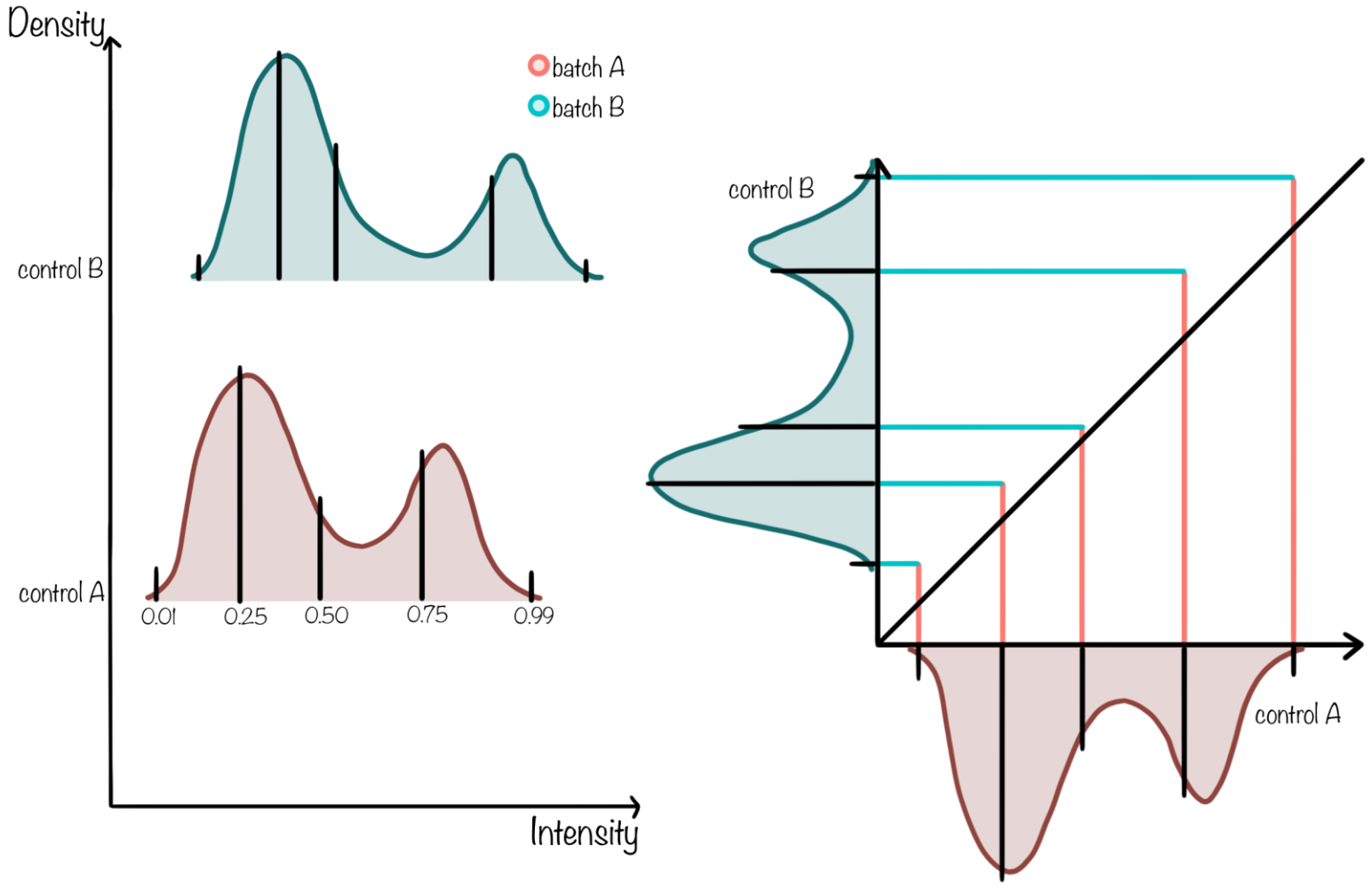
Figures: Katrien Quintelier

Batch effect correction: CytoNorm



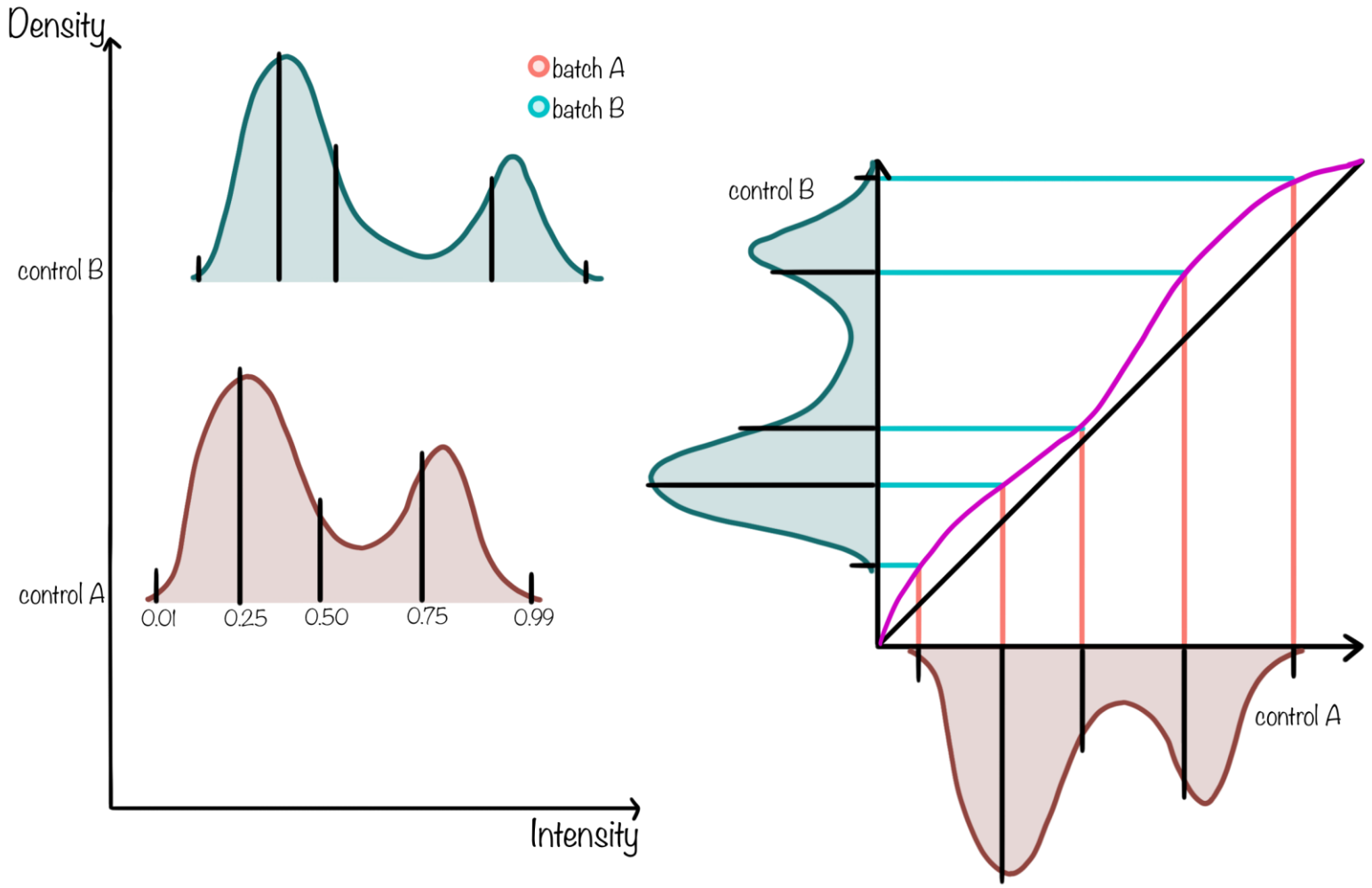
Figures: Katrien Quintelier

Batch effect correction: CytoNorm



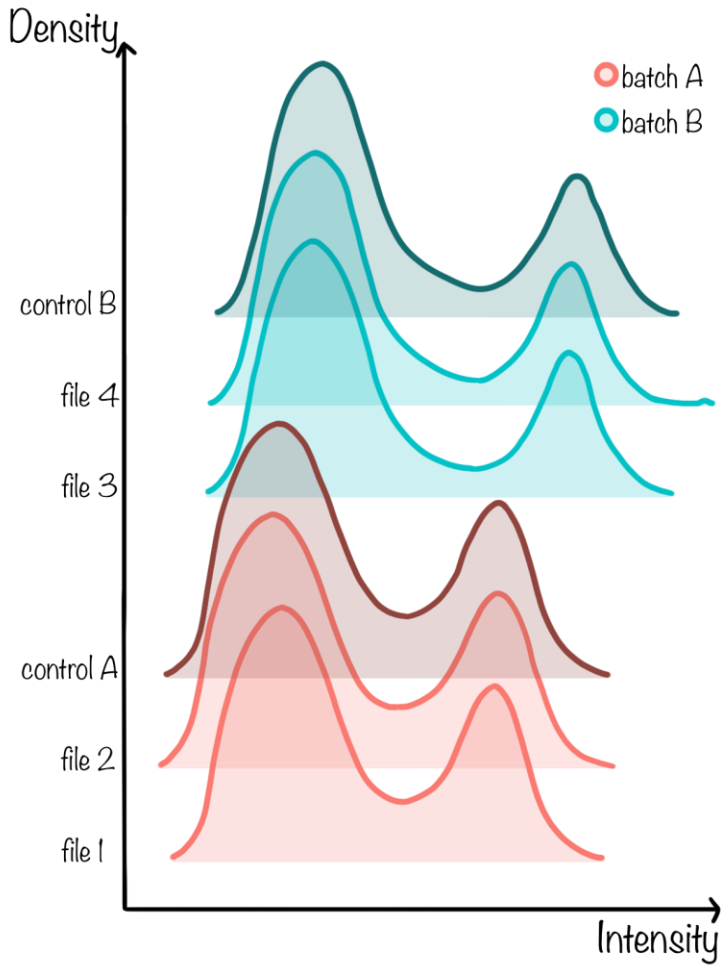
Figures: Katrien Quintelier

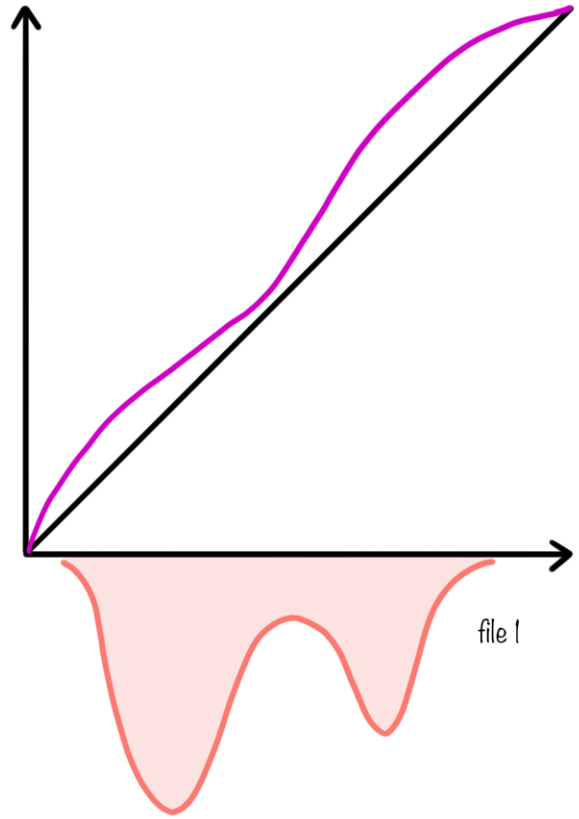
Batch effect correction: CytoNorm



Figures: Katrien Quintelier

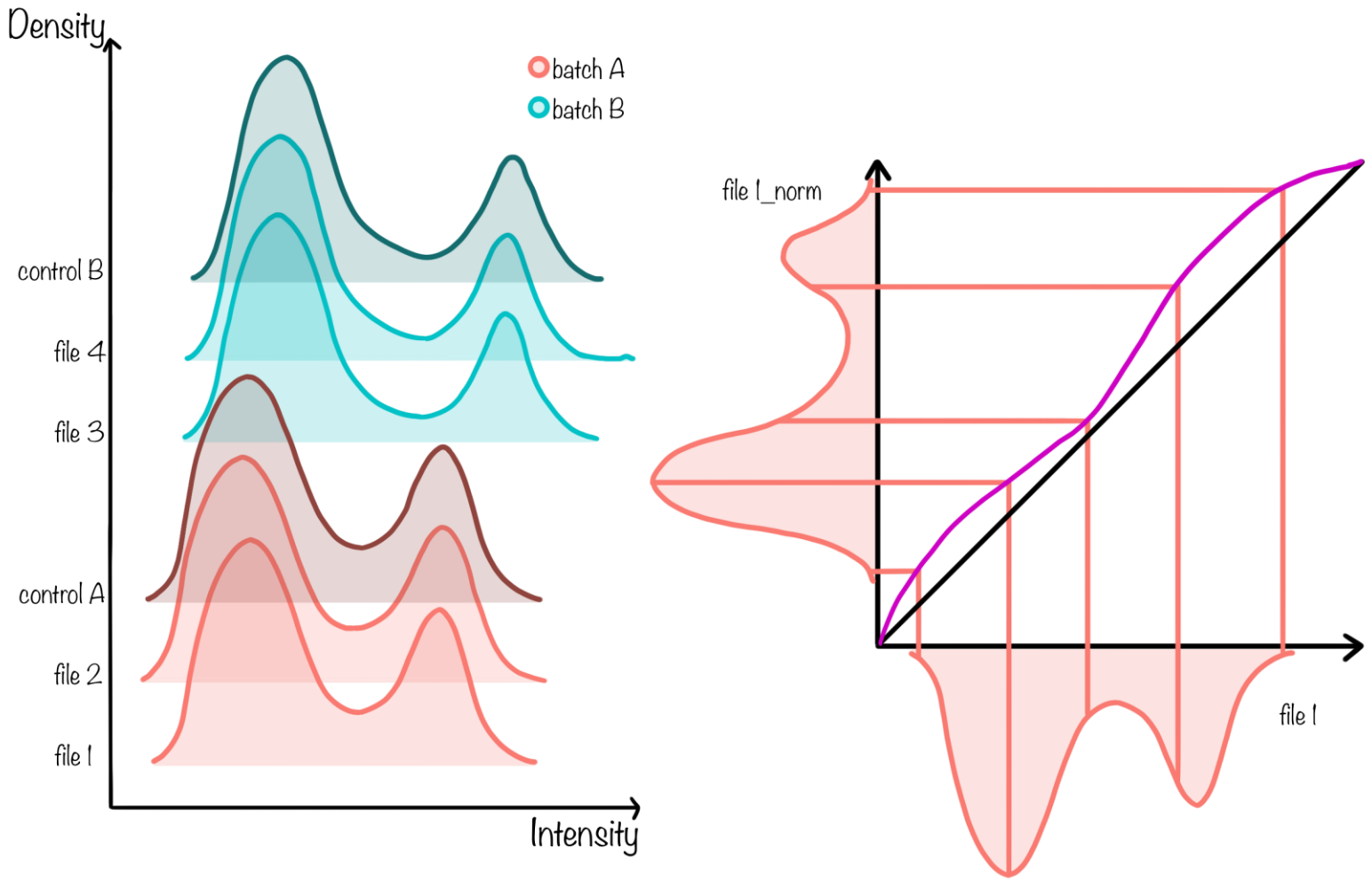
Batch effect correction: CytoNorm





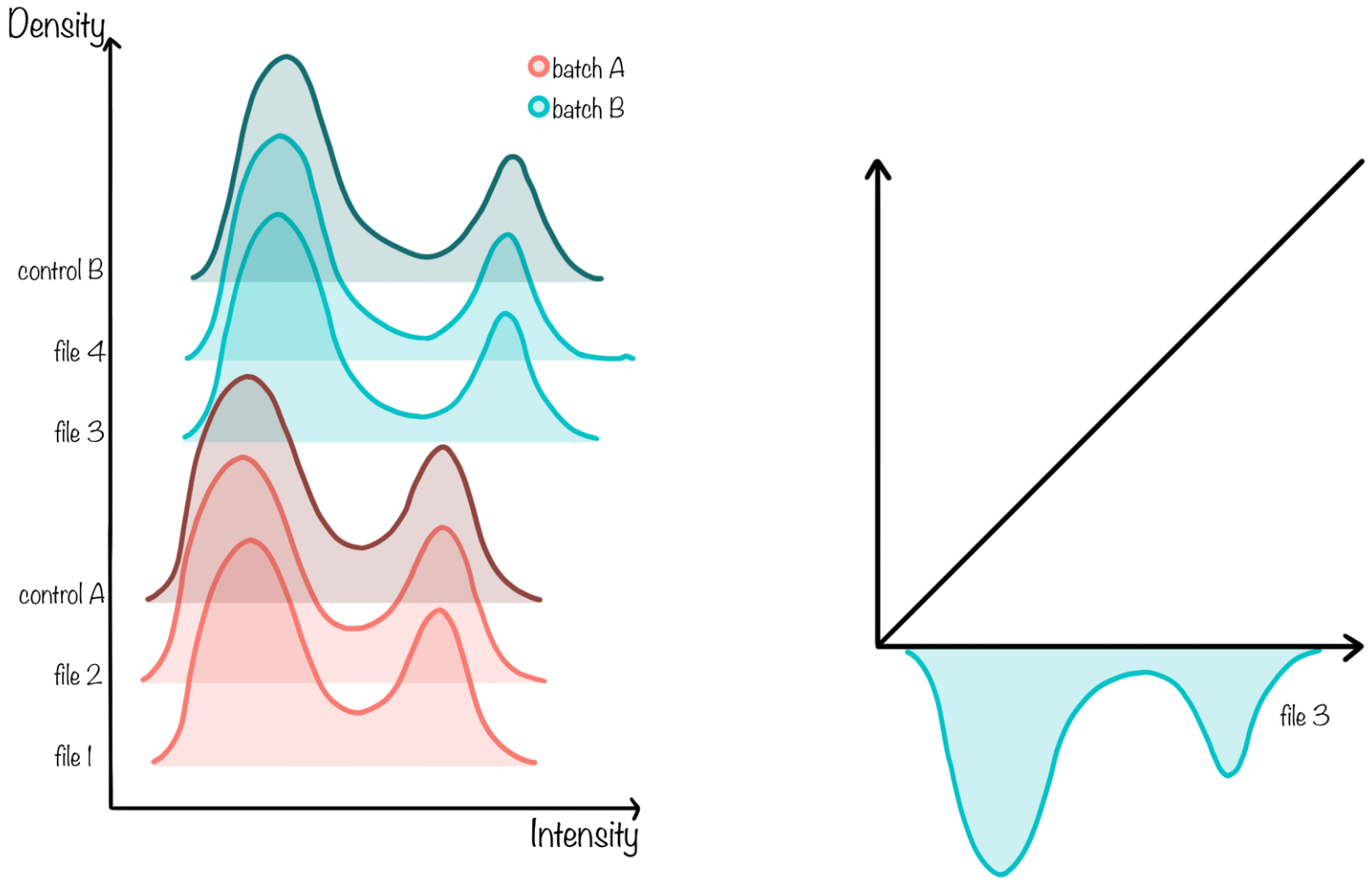
Figures: Katrien Quintelier

Batch effect correction: CytoNorm



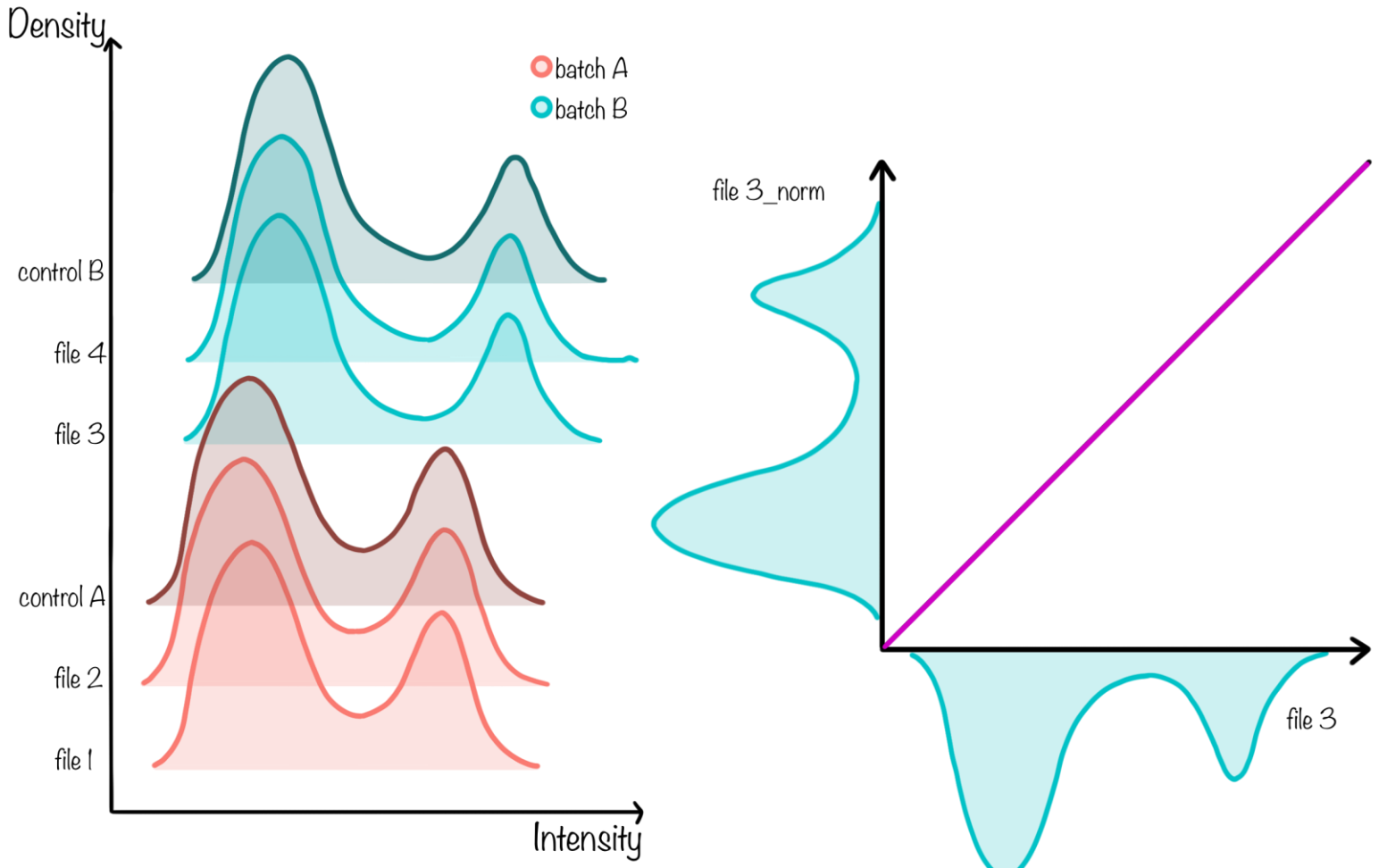
Figures: Katrien Quintelier

Batch effect correction: CytoNorm



Figures: Katrien Quintelier

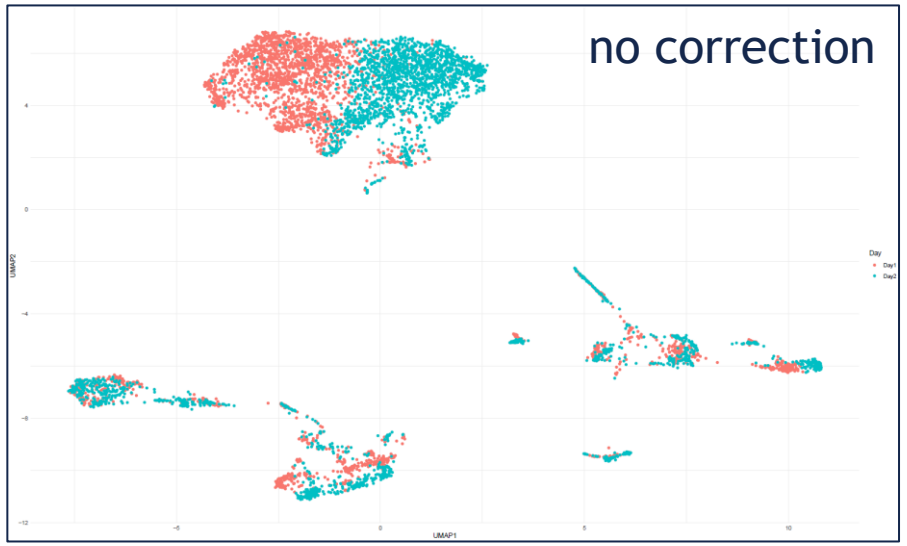
Batch effect correction: CytoNorm

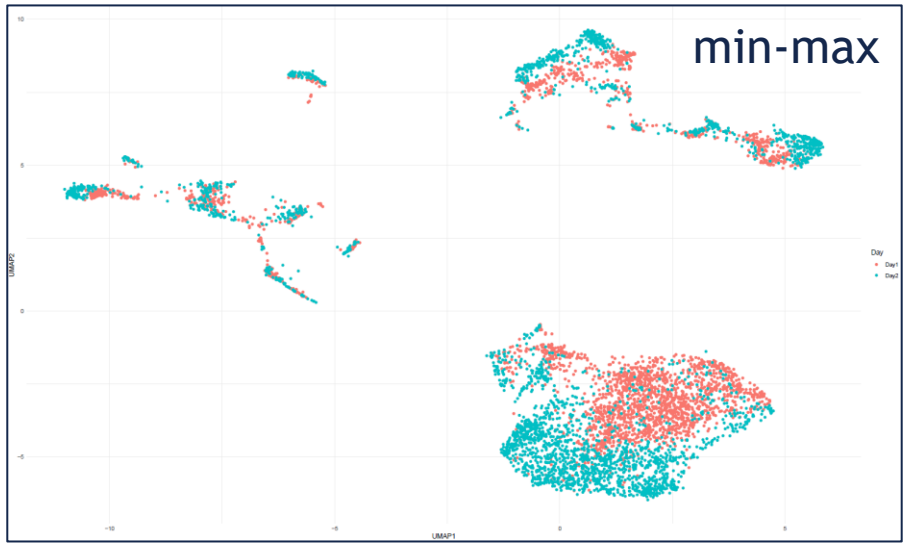


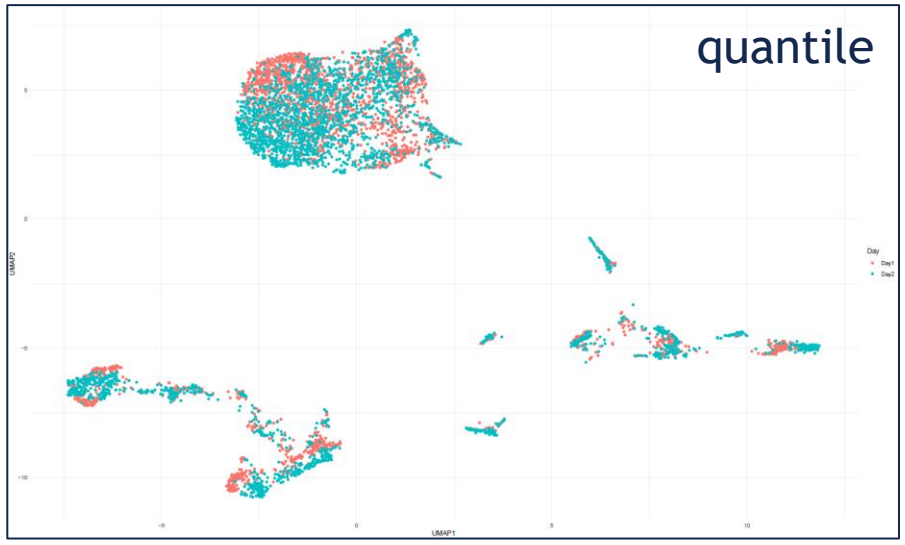
Figures: Katrien Quintelier

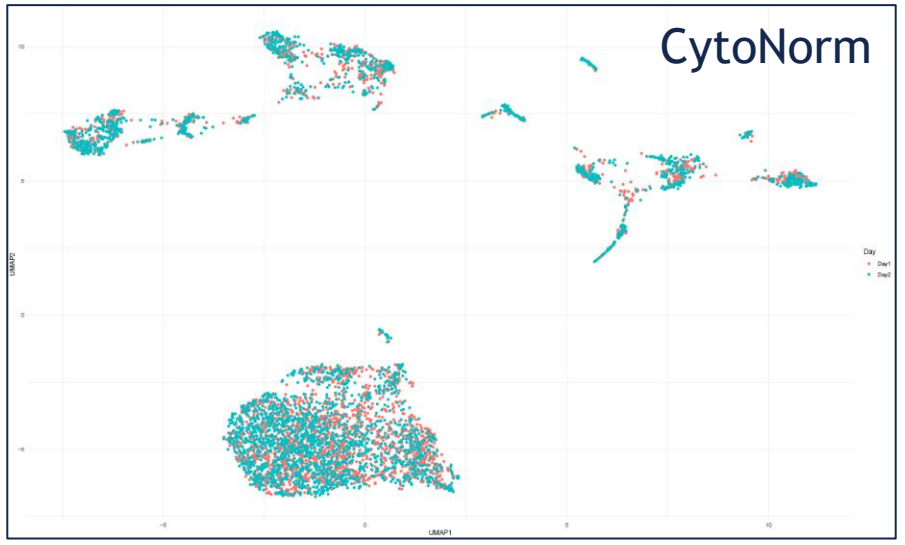
Batch effect correction

Overview batch effect correction results









What’s next?

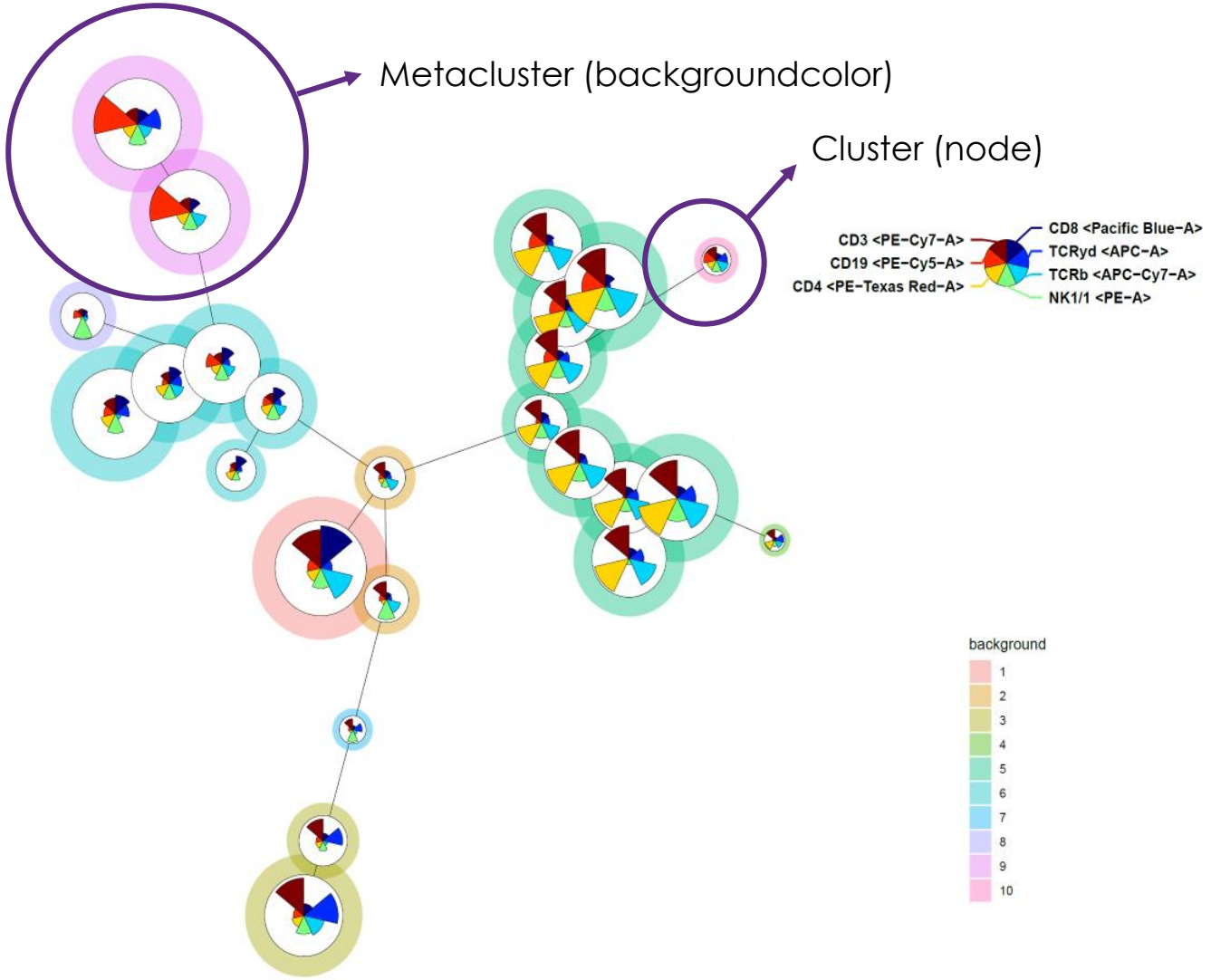
Automated computational analysis pipeline

Clustering

Cluster cells with
similar marker expression profiles

e.g. FlowSOM

metaclusters ≈ cell populations



Downstream analysis

FlowSOM output

- Features: counts, percentages, MFIs, percentage positive
- Levels: cluster, metacluster

Use this output for

- Comparing abundances or population MFIs between outcomes/groups of interest
- Correlation analyses
- Predictive modeling
- …

Resources

Resources

CytoBytes: short, practical tutorials - covering tools, techniques and best practices in cytometry data analysis

https://www.youtube.com/@CytoBytes

Computational cytometry (FlowSOM) protocol

Quintelier, K., Couckuyt, A., Emmaneel, A. et al. (2021). Analyzing high-dimensional cytometry data using FlowSOM. Nat Protoc 16, 3775–3801. doi: 10.1038/s41596-021-00550-0

Protocol for spectral data acquisition and analysis (incl. scaling and batch effect correction)

Ferrer-Font, L., Kraker, G., Hally, K. E., & Price, K. M. (2023). Ensuring full spectrum flow cytometry data quality for high dimensional data analysis. Current Protocols, 3, e657. doi: 10.1002/cpz1.657

General overview of high dimensional data analysis

Liechti, T., Weber, L.M., Ashhurst, T.M. et al. (2021). An updated guide for the perplexed: cytometry in the high-dimensional era. Nat Immunol 22, 1190–1197. doi: 10.1038/s41590-021-01006-z

R code from today's workshop

https://github.com/saeyslab/ComputationalCytometry

SarahM.Bonte@UGent.be

SaeysLab
Data mining and Modelling for Biomedicine

VIB

UNIVERSITEIT GENT