

- Detection p-value >.01 (in one or more samples)
- Probes on the sex chromosomes
- Probes affected by common SNPs
- Cross-reactive probes and polymorphic probes (Chen et al.)
- Cross-hybridizing and polymorphic probes on EPIC (McCartney et al.)
- → BMIQ.quantileN\_filtered
- → quantileN\_filtered & Betas\_quantileN\_filtered & Ms\_quantileN\_filtered

## Remove batch effects with Combat:

- NA probes without variance (M-values)
- PCA on the M values → PCobj
- determine which batches show strongest association with PCs
- Remove batches with combat (→M\_combat\_1plate/etc..., last latest number)
- → Betas\_combated & Betas\_combated\_ExprSet

## Remove batch effects with Combat:

- NA probes without variance (M-values)
- PCA on the M values → PCobj
- determine which batches show strongest association with PCs
- Remove batches with combat (→M\_combat\_1plate/etc..., last latest number)
- → BMIQ.quantileN\_combated& BMIQ.quantileN\_combated\_ExprSet

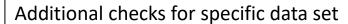
This is added for cases where all probes are needed!



- Apply MixedUpMapper to check mismatches between genotypes and methylation
- → exclude mismatches from Betas\_combated & Betas\_combated\_ExprSet & BMIQ.quantileN
- → if exclusions: Betas\_combated\_mixc & Betas\_combated\_ExprSet\_mixc & BMIQ.quantileN\_combated\_mixc



→ Betas\_gapped



- Cord blood: contamination maternal blood
- → Contaminated IDs excluded in: **Betas\_combated\_CC**, Betas\_combated\_ExprSet\_CC, BMIQ.quantileN\_combated\_CC, BMIQ.quantileN\_combated\_ExprSet\_CC, pd\_cleaned\_CC



- methylation\_betas\_XX in subfolder finalData
- And a sample-cleaned RGSet named RGSet\_cleaned\_final

## estimate cell type proportions

- reference-free: RefFreeEWAS
- reference-based: e.g. for cord blood Salas et al. or Bakulski et al.