scater pre-processing and quality control workflow

From raw RNA-seq reads to a clean, tidy dataset ready for downstream analysis

Raw RNA-seq Reads
[Fastq format]

Summarised feature expression values
[e.g. produced by bioinformatics core]

1. Obtain RNA-seq expression data

runKallisto/
readKallisto
runSalmon/readSalmon

newSCESet

SCESet
[Container: S4 class inheriting Bioconductor's ExpressionSet]

Object that contains assay data, phenotype data, feature data, and more, for single-cell analysis

calculateQCMetrics
QC methods

plot

plotQC

plotPCA

plotTSNE

plotMDS

plotDiffusionMap

plotReducedDim

plotExpression

plotPhenoData

plotFeatureData

plotMetadata

Plotting methods

2. QC and filter features

3. QC and filter cells

Filtered SCESet

normaliseExprs
normalise
Normalisation methods

4. Simple normalisation

5. QC of explanatory variables

getBMFeatureAnnos

summariseExprsAcross
Features

Miscellaneous methods

Tidy filtered and normalised SCESet

(6. Further normalisation)

Downstream modelling and statistical analysis