

# Package ‘BASiCS’

November 25, 2024

**Type** Package

**Title** Bayesian Analysis of Single-Cell Sequencing data

**Version** 2.19.0

**Date** 2024-03-25

**Description** Single-cell mRNA sequencing can uncover novel cell-to-cell heterogeneity in gene expression levels in seemingly homogeneous populations of cells. However, these experiments are prone to high levels of technical noise, creating new challenges for identifying genes that show genuine heterogeneous expression within the population of cells under study. BASiCS (Bayesian Analysis of Single-Cell Sequencing data) is an integrated Bayesian hierarchical model to perform statistical analyses of single-cell RNA sequencing datasets in the context of supervised experiments (where the groups of cells of interest are known a priori, e.g. experimental conditions or cell types). BASiCS performs built-in data normalisation (global scaling) and technical noise quantification (based on spike-in genes). BASiCS provides an intuitive detection criterion for highly (or lowly) variable genes within a single group of cells. Additionally, BASiCS can compare gene expression patterns between two or more pre-specified groups of cells. Unlike traditional differential expression tools, BASiCS quantifies changes in expression that lie beyond comparisons of means, also allowing the study of changes in cell-to-cell heterogeneity. The latter can be quantified via a biological over-dispersion parameter that measures the excess of variability that is observed with respect to Poisson sampling noise, after normalisation and technical noise removal. Due to the strong mean/over-dispersion confounding that is typically observed for scRNA-seq datasets, BASiCS also tests for changes in residual over-dispersion, defined by residual values with respect to a global mean/over-dispersion trend.

**License** GPL-3

**Depends** R ( $\geq 4.2$ ), SingleCellExperiment

**Imports** Biobase, BiocGenerics, coda, cowplot, ggExtra, ggplot2, graphics, grDevices, MASS, methods, Rcpp ( $\geq 0.11.3$ ), S4Vectors, scran, scuttle, stats, stats4, SummarizedExperiment, viridis, utils, Matrix ( $\geq 1.5.0$ ), matrixStats, assertthat, reshape2, BiocParallel, posterior, hexbin

**Suggests** BiocStyle, knitr, rmarkdown, testthat, scRNAseq, magick

**LinkingTo** Rcpp, RcppArmadillo

**VignetteBuilder** knitr

**biocViews** ImmunoOncology, Normalization, Sequencing, RNASeq, Software,  
GeneExpression, Transcriptomics, SingleCell,  
DifferentialExpression, Bayesian, CellBiology, ImmunoOncology

**SystemRequirements** C++11

**NeedsCompilation** yes

**URL** <https://github.com/catavallejos/BASiCS>

**BugReports** <https://github.com/catavallejos/BASiCS/issues>

**RoxygenNote** 7.3.1

**Encoding** UTF-8

**LazyData** false

**Config/testthat/edition** 3

**Collate** 'AllClasses.R' 'AllGenerics.R' 'BASiCS\_CalculateERCC.R'  
'BASiCS\_CorrectOffset.R' 'BASiCS\_DenoisedCounts.R'  
'BASiCS\_DenoisedRates.R' 'BASiCS\_DetectHVG\_LVG.R'  
'BASiCS\_DiagHist.R' 'BASiCS\_DiagPlot.R'  
'BASiCS\_DivideAndConquer.R' 'BASiCS\_Draw.R'  
'BASiCS\_EffectiveSize.R' 'BASiCS\_Filter.R' 'BASiCS\_LoadChain.R'  
'BASiCS\_MCMC.R' 'BASiCS\_MockSCE.R' 'BASiCS\_Package.R'  
'BASiCS\_PlotDE.R' 'BASiCS\_PlotOffset.R' 'BASiCS\_PlotVG.R'  
'BASiCS\_PlotVarianceDecomp.R' 'BASiCS\_PriorParam.R'  
'BASiCS\_ShowFit.R' 'BASiCS\_Sim.R' 'BASiCS\_TestDE.R'  
'BASiCS\_VarThresholdSearchHVG\_LVG.R' 'BASiCS\_VarianceDecomp.R'  
'HiddenBASiCS\_Sim.R' 'HiddenHeaderBASiCS\_Sim.R'  
'HiddenHeaderTest\_DE.R' 'HiddenVarDecomp.R' 'utils\_Misc.R'  
'Methods.R' 'RcppExports.R' 'data.R' 'makeExampleBASiCS\_Data.R'  
'newBASiCS\_Chain.R' 'newBASiCS\_Data.R' 'utils\_Data.R'  
'utils\_DivideAndConquer.R' 'utils\_MCMC.R' 'utils\_Store.R'  
'utils\_Tests.R' 'utils\_VG.R' 'welcome.R'

**git\_url** <https://git.bioconductor.org/packages/BASiCS>

**git\_branch** devel

**git\_last\_commit** 055feb8

**git\_last\_commit\_date** 2024-10-29

**Repository** Bioconductor 3.21

**Date/Publication** 2024-11-25

**Author** Catalina Vallejos [aut, cre] (ORCID:  
<<https://orcid.org/0000-0003-3638-1960>>),  
Nils Eling [aut],  
Alan O'Callaghan [aut],  
Sylvia Richardson [ctb],  
John Marioni [ctb]

**Maintainer** Catalina Vallejos <catalina.vallejos@igmm.ed.ac.uk>

## Contents

.generateSubsets . . . . .	4
as.data.frame,BASiCS_ResultsDE-method . . . . .	5
BASiCS-defunct . . . . .	6
BASiCS_CalculateERCC . . . . .	6
BASiCS_Chain . . . . .	7
BASiCS_Chain-methods . . . . .	8
BASiCS_CorrectOffset . . . . .	9
BASiCS_DenoisedCounts . . . . .	10
BASiCS_DenoisedRates . . . . .	11
BASiCS_DetectVG . . . . .	12
BASiCS_DiagHist . . . . .	14
BASiCS_DiagPlot . . . . .	16
BASiCS_DivideAndConquer . . . . .	17
BASiCS_Draw . . . . .	19
BASiCS_EffectiveSize . . . . .	20
BASiCS_Filter . . . . .	21
BASiCS_LoadChain . . . . .	22
BASiCS_MCMC . . . . .	23
BASiCS_MockSCE . . . . .	27
BASiCS_PlotDE . . . . .	28
BASiCS_PlotOffset . . . . .	30
BASiCS_PlotVarianceDecomp . . . . .	31
BASiCS_PlotVG . . . . .	32
BASiCS_PriorParam . . . . .	32
BASiCS_Result . . . . .	34
BASiCS_ResultDE . . . . .	35
BASiCS_ResultsDE . . . . .	35
BASiCS_ResultVG . . . . .	36
BASiCS_ShowFit . . . . .	36
BASiCS_Sim . . . . .	38
BASiCS_Summary . . . . .	39
BASiCS_Summary-methods . . . . .	40
BASiCS_TestDE . . . . .	41
BASiCS_VarianceDecomp . . . . .	44
BASiCS_VarThresholdSearchHVG . . . . .	46
ChainRNA . . . . .	47
ChainRNAReg . . . . .	48
ChainSC . . . . .	48
ChainSCReg . . . . .	49
dim . . . . .	49
dimnames . . . . .	50
displayChainBASiCS-BASiCS_Chain-method . . . . .	50
displaySummaryBASiCS-BASiCS_Summary-method . . . . .	51
format,BASiCS_ResultsDE-method . . . . .	52

makeExampleBASiCS_Data . . . . .	53
newBASiCS_Chain . . . . .	54
newBASiCS_Data . . . . .	56
plot-BASiCS_Chain-method . . . . .	57
plot-BASiCS_Summary-method . . . . .	58
rowData,BASiCS_ResultsDE-method . . . . .	59
show,BASiCS_ResultDE-method . . . . .	60
show,BASiCS_ResultsDE-method . . . . .	61
show,BASiCS_ResultVG-method . . . . .	61
subset . . . . .	62
Summary . . . . .	63
[,BASiCS_ResultsDE,ANY,ANY,ANY-method . . . . .	64

## Index 65

---

.generateSubsets	<i>Generate balanced subsets for divide and conquer BASiCS</i>
------------------	--

---

### Description

Partitions data based on either cells or genes. Attempts to find a partitioning scheme which is "balanced" for either total reads per cell across all genes (partitioning by gene) or total expression per gene across all cells (partitioning by gene). When partitioning by cell, at least 20 cells must be in each partition or BASiCS\_MCMC will fail. If this partitioning fails, it will continue recursively up to a maximum number of iterations (20 by default).

### Usage

```
.generateSubsets(
  Data,
  NSubsets,
  SubsetBy = c("cell", "gene"),
  Alpha = 0.05,
  WithSpikes = FALSE,
  MaxDepth = 20,
  .Depth = 1
)
```

### Arguments

Data	a SingleCellExperiment object
NSubsets	Integer specifying the number of batches into which to divide Data for divide and conquer inference.
SubsetBy	Partition by "cell" or by "gene".
Alpha	p-value threshold for ANOVA testing of "balance"
WithSpikes	Similar to argument for BASiCS_MCMC - do the Data contain spikes?
MaxDepth	Maximum number of recursive
.Depth	Internal parameter. Do not set.

**Value**

A list of SingleCellExperiment objects

---

as.data.frame,BASiCS\_ResultsDE-method

*Converting BASiCS results objects to data.frames*

---

**Description**

Converting BASiCS results objects to data.frames

**Usage**

```
## S4 method for signature 'BASiCS_ResultsDE'
as.data.frame(x, Parameter, Filter = TRUE, ProbThreshold = NULL)

## S4 method for signature 'BASiCS_ResultDE'
as.data.frame(x, Filter = TRUE, ProbThreshold = NULL)

## S4 method for signature 'BASiCS_ResultVG'
as.data.frame(x, Filter = TRUE, ProbThreshold = NULL)
```

**Arguments**

x	An object of class <a href="#">BASiCS_ResultVG</a> , <a href="#">BASiCS_ResultDE</a> , or <a href="#">BASiCS_ResultsDE</a> .
Parameter	For <a href="#">BASiCS_ResultsDE</a> objects only. Character scalar specifying which table of results to output. Available options are "Mean", (mu, mean expression), "Disp" (delta, overdispersion) and "ResDisp" (epsilon, residual overdispersion).
Filter	Logical scalar. If TRUE, output only entries corresponding to genes that pass the decision rule used in the probabilistic test.
ProbThreshold	Only used if Filter=TRUE. Numeric scalar specifying the probability threshold to be used when filtering genes. Default is to use the threshold used in the original decision rule when the test was performed.

**Value**

A data.frame of test results.

---

BASiCS-defunct      *Defunct functions in package 'BASiCS'*

---

### Description

The functions listed here are no longer part of BASiCS.

### Details

## Removed

- BASiCS\_D\_TestDE has been replaced by BASiCS\_TestDE.

### usage

## Removed

- BASiCS\_D\_TestDE()

### Author(s)

Catalina A. Vallejos <cnvallej@uc.cl>

### See Also

- [BASiCS\\_TestDE](#)

---

BASiCS\_CalculateERCC      *Convert concentration in moles per microlitre to molecule counts*

---

### Description

Convert concentration in moles per microlitre to molecule counts

### Usage

```
BASiCS_CalculateERCC(Mix, DilutionFactor, VolumePerCell)
```

### Arguments

Mix	The name of the spike-in mix to use.
DilutionFactor	The dilution factor applied to the spike-in mixture. e.g., 1 microlitre per 50ml would be a 1/50000 DilutionFactor.
VolumePerCell	The volume of spike-in mixture added to each well, or to each cell.

### Value

The molecule counts per well, or per cell, based on the input parameters.

BASiCS\_Chain

*The BASiCS\_Chain class***Description**

Container of an MCMC sample of the BASiCS' model parameters as generated by the function [BASiCS\\_MCMC](#).

**Slots**

**parameters** List of matrices containing MCMC chains for each model parameter. Depending on the mode in which BASiCS was run, the following parameters can appear in the list:

**mu** MCMC chain for gene-specific mean expression parameters  $\mu_i$ , biological genes only (matrix with `q.bio` columns, all elements must be positive numbers)

**delta** MCMC chain for gene-specific biological over-dispersion parameters  $\delta_i$ , biological genes only (matrix with `q.bio` columns, all elements must be positive numbers)

**phi** MCMC chain for cell-specific mRNA content normalisation parameters  $\phi_j$  (matrix with `n` columns, all elements must be positive numbers and the sum of its elements must be equal to `n`) This parameter is only used when spike-in genes are available.

**s** MCMC chain for cell-specific technical normalisation parameters  $s_j$  (matrix with `n` columns, all elements must be positive numbers)

**nu** MCMC chain for cell-specific random effects  $\nu_j$  (matrix with `n` columns, all elements must be positive numbers)

**theta** MCMC chain for technical over-dispersion parameter(s)  $\theta$  (matrix, all elements must be positive, each column represents 1 batch)

**beta** Only relevant for regression BASiCS model (Eling et al, 2017). MCMC chain for regression coefficients (matrix with `k` columns, where `k` represent the number of chosen basis functions + 2)

**sigma2** Only relevant for regression BASiCS model (Eling et al, 2017). MCMC chain for the residual variance (matrix with one column, `sigma2` represents a global parameter)

**epsilon** Only relevant for regression BASiCS model (Eling et al, 2017). MCMC chain for the gene-specific residual over-dispersion parameter (matrix with `q` columns)

**RefFreq** Only relevant for no-spikes BASiCS model (Eling et al, 2017). For each biological gene, this vector displays the proportion of times for which each gene was used as a reference (within the MCMC algorithm), when using the stochastic reference choice described in (Eling et al, 2017). This information has been kept as it is useful for the developers of this library. However, we do not expect users to need it.

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

## Examples

```
# A BASiCS_Chain object created by the BASiCS_MCMC function.
Data <- makeExampleBASiCS_Data()

# To run the model without regression
Chain <- BASiCS_MCMC(Data, N = 100, Thin = 2, Burn = 2, Regression = FALSE)

# To run the model using the regression model
ChainReg <- BASiCS_MCMC(Data, N = 100, Thin = 2, Burn = 2, Regression = TRUE)
```

---

BASiCS\_Chain-methods *'show' method for BASiCS\_Chain objects*

---

## Description

'show' method for [BASiCS\\_Chain](#) objects.

'updateObject' method for [BASiCS\\_Chain](#) objects. It is used to convert outdated [BASiCS\\_Chain](#) objects into a version that is compatible with the Bioconductor release of BASiCS. Do not use this method if [BASiCS\\_Chain](#) already contains a parameters slot.

## Usage

```
## S4 method for signature 'BASiCS_Chain'
show(object)

## S4 method for signature 'BASiCS_Chain'
updateObject(object, ..., verbose = FALSE)
```

## Arguments

object	A <a href="#">BASiCS_Chain</a> object.
...	Additional arguments of <a href="#">updateObject</a> generic method. Not used within BASiCS.
verbose	Additional argument of <a href="#">updateObject</a> generic method. Not used within BASiCS.

## Value

Prints a summary of the properties of object.

Returns an updated [BASiCS\\_Chain](#) object that contains all model parameters in a single slot object (list).

## Author(s)

Catalina A. Vallejos <cnvallej@uc.cl>  
Nils Eling <eling@ebi.ac.uk>



**Examples**

```
Data <- makeExampleBASiCS_Data()
Chain <- BASiCS_MCMC(Data, N = 50, Thin = 2, Burn = 2, Regression = FALSE)

# Not run
# New_Chain <- updateObject(Old_Chain)
```

---

BASiCS\_CorrectOffset *Remove global mean expression offset*

---

**Description**

Remove global offset in mean expression between two BASiCS\_Chain objects.

**Usage**

```
BASiCS_CorrectOffset(Chain, ChainRef, min.mean = 1)
```

**Arguments**

Chain	a 'BASiCS_MCMC' object to which the offset correction should be applied (with respect to 'ChainRef').
ChainRef	a 'BASiCS_MCMC' object to be used as the reference in the offset correction procedure.
min.mean	Minimum mean expression threshold required for inclusion in offset calculation. Similar to 'min.mean' in 'scrans::computeSumFactors'.

**Value**

A list whose first element is an offset corrected version of 'Chain' (using 'ChainRef' as a reference), whose second element is the point estimate for the offset and whose third element contains iteration-specific offsets.

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

Alan O'Callaghan

## Examples

```
# Loading two 'BASiCS_Chain' objects (obtained using 'BASiCS_MCMC')
data(ChainSC)
data(ChainRNA)

A <- BASiCS_CorrectOffset(ChainSC, ChainRNA)

# Offset corrected versions for ChainSC (with respect to ChainRNA).
A$Chain
A$Offset
```

---

BASiCS\_DenoisedCounts *Calculates denoised expression counts*

---

## Description

Calculates denoised expression counts by removing cell-specific technical variation. The latter includes global-scaling normalisation and therefore no further normalisation is required.

## Usage

```
BASiCS_DenoisedCounts(Data, Chain, WithSpikes = TRUE)
```

## Arguments

Data	An object of class <a href="#">SingleCellExperiment</a>
Chain	An object of class <a href="#">BASiCS_Chain</a>
WithSpikes	A logical scalar specifying whether denoised spike-in genes should be generated as part of the output value. This only applies when the <a href="#">BASiCS_Chain</a> object was generated with the setting <code>WithSpikes=TRUE</code> .

## Details

See vignette `browseVignettes("BASiCS")`

## Value

A matrix of denoised expression counts. In line with global scaling normalisation strategies, these are defined as  $X_{ij}/(\phi_j\nu_j)$  for biological genes and  $X_{ij}/(\nu_j)$  for spike-in genes. For this calculation  $\phi_j$   $\nu_j$  are estimated by their corresponding posterior medians. If spike-ins are not used,  $\phi_j$  is set equal to 1.

## Author(s)

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

**See Also**[BASiCS\\_Chain](#)**Examples**

```
Data <- makeExampleBASiCS_Data(WithSpikes = TRUE)
## The N and Burn parameters used here are optimised for speed
## and should not be used in regular use.
## For more useful parameters,
## see the vignette (\code{browseVignettes("BASiCS")})
Chain <- BASiCS_MCMC(Data, N = 1000, Thin = 10, Burn = 500,
                    Regression = FALSE, PrintProgress = FALSE)

DC <- BASiCS_DenoisedCounts(Data, Chain)
```

---

BASiCS\_DenoisedRates *Calculates denoised expression rates*

---

**Description**

Calculates normalised and denoised expression rates, by removing the effect of technical variation.

**Usage**

```
BASiCS_DenoisedRates(Data, Chain, Propensities = FALSE)
```

**Arguments**

Data	an object of class <a href="#">SingleCellExperiment</a>
Chain	an object of class <a href="#">BASiCS_Chain</a>
Propensities	If TRUE, returns underlying expression propensities $\rho_{ij}$ . Otherwise, denoised rates $\mu_i \rho_{ij}$ are returned. Default: Propensities = FALSE.

**Details**

See vignette

**Value**

A matrix of denoised expression rates (biological genes only)

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

**See Also**[BASiCS\\_Chain](#)**Examples**

```

Data <- makeExampleBASiCS_Data(WithSpikes = TRUE)
## The N and Burn parameters used here are optimised for speed
## and should not be used in regular use.
## For more useful parameters,
## see the vignette (\code{browseVignettes("BASiCS")})
Chain <- BASiCS_MCMC(Data, N = 1000, Thin = 10, Burn = 500,
                    Regression = FALSE, PrintProgress = FALSE)

DR <- BASiCS_DenoisedRates(Data, Chain)

```

---

 BASiCS\_DetectVG

*Detection method for highly (HVG) and lowly (LVG) variable genes*


---

**Description**

Functions to detect highly and lowly variable genes. If the BASiCS\_Chain object was generated using the regression approach, BASiCS finds the top highly variable genes based on the posteriors of the epsilon parameters. Otherwise, the old approach is used, which initially performs a variance decomposition.

**Usage**

```

BASiCS_DetectVG(
  Chain,
  Task = c("HVG", "LVG"),
  PercentileThreshold = NULL,
  VarThreshold = NULL,
  ProbThreshold = 2/3,
  EpsilonThreshold = NULL,
  EFDR = 0.1,
  OrderVariable = c("Prob", "GeneIndex", "GeneName"),
  Plot = FALSE,
  MinESS = 100,
  ...
)

BASiCS_DetectLVG(Chain, ...)

BASiCS_DetectHVG(Chain, ...)

```

**Arguments**

Chain	an object of class <a href="#">BASiCS_Chain</a>
Task	Search for highly variable genes (Task="HVG") or lowly variable genes (Task="LVG").
PercentileThreshold	Threshold to detect a percentile of variable genes (must be a positive value, between 0 and 1). Default: PercentileThreshold = NULL.
VarThreshold	Variance contribution threshold (must be a positive value, between 0 and 1). This is only used when the BASiCS non-regression model was used to generate the Chain object. Default: VarThreshold = NULL.
ProbThreshold	Optional parameter. Posterior probability threshold (must be a positive value, between 0 and 1). If EFDR = NULL, the posterior probability threshold for the test will be set to ProbThreshold.
EpsilonThreshold	Threshold for residual overdispersion above which
EFDR	Target for expected false discovery rate related to HVG/LVG detection. If EFDR = NULL, EFDR calibration is not performed and the posterior probability threshold is set equal to ProbThreshold. Default EFDR = 0.10.
OrderVariable	Ordering variable for output. Possible values: 'GeneIndex', 'GeneName' and 'Prob'. Default ProbThreshold = 'Prob'
Plot	If Plot = TRUE error control and expression versus HVG/LVG probability plots are generated.
MinESS	The minimum effective sample size for a gene to be included in the HVG or LVG tests. This helps to remove genes with poor mixing from detection of HVGs/LVGs. Default is 100. If set to NA, genes are not checked for effective sample size the tests are performed.
...	Graphical parameters (see <a href="#">par</a> ).

**Details**

In some cases, the EFDR calibration step may fail to find probability threshold that controls the EFDR at the chosen level. In cases like

See vignette

**Value**

An object of class [BASiCS\\_ResultVG](#).

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

**References**

Vallejos, Marioni and Richardson (2015). PLoS Computational Biology.

**See Also**[BASiCS\\_Chain](#)**Examples**

```

# Loads short example chain (non-regression implementation)
data(ChainSC)

# Highly and lowly variable genes detection (within a single group of cells)
DetectHVG <- BASiCS_DetectHVG(ChainSC, VarThreshold = 0.60,
                             EFDR = 0.10, Plot = TRUE)
DetectLVG <- BASiCS_DetectLVG(ChainSC, VarThreshold = 0.40,
                             EFDR = 0.10, Plot = TRUE)

# Loads short example chain (regression implementation)
data(ChainSCReg)

# Highly and lowly variable genes detection (within a single group of cells)
DetectHVG <- BASiCS_DetectHVG(ChainSCReg, PercentileThreshold = 0.90,
                             EFDR = 0.10, Plot = TRUE)
DetectLVG <- BASiCS_DetectLVG(ChainSCReg, PercentileThreshold = 0.10,
                             EFDR = 0.10, Plot = TRUE)

## Highly and lowly variable genes detection based on residual overdispersion
## threshold
DetectHVG <- BASiCS_DetectHVG(ChainSCReg, EpsilonThreshold = log(2), Plot = TRUE)
DetectLVG <- BASiCS_DetectLVG(ChainSCReg, EpsilonThreshold = -log(2), Plot = TRUE)

```

---

BASiCS\_DiagHist

*Create diagnostic plots of MCMC parameters*


---

**Description**

Plot a histogram of effective sample size or Geweke's diagnostic z-statistic. See [effectiveSize](#) and [geweke.diag](#) for more details.

**Usage**

```

BASiCS_DiagHist(
  object,
  Parameter = NULL,
  Measure = c("ess", "geweke.diag", "rhat"),
  VLine = TRUE,
  na.rm = TRUE
)

BASiCS_diagHist(...)

```

**Arguments**

object	an object of class <a href="#">BASiCS_Summary</a>
Parameter	Optional name of a chain parameter to restrict the histogram; if not supplied, all parameters will be assessed. Default Parameter = NULL.
Measure	Character scalar specifying the diagnostic measure to plot. Current options are effective sample size, the Geweke diagnostic criterion, and the <a href="#">rhat</a> diagnostic.
VLine	Numeric scalar indicating a threshold value to be displayed as a dashed line on the plot. Alternatively, can be set to FALSE to disable line drawing, or TRUE to use the default thresholds.
na.rm	Logical scalar indicating whether NA values should be removed before calculating effective sample size.
...	Unused.

**Value**

A ggplot object.

**Author(s)**

Alan O'Callaghan

**References**

Geweke, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *\_Bayesian Statistics 4\_* (ed JM Bernardo, JO Berger, AP Dawid and AFM Smith). Clarendon Press, Oxford, UK.

**See Also**

[BASiCS\\_Chain](#)

**Examples**

```
# Built-in example chain
data(ChainSC)

# See effective sample size distribution across all parameters
BASiCS_DiagHist(ChainSC)
# For mu only
BASiCS_DiagHist(ChainSC, Parameter = "mu")
```

---

BASiCS\_DiagPlot      *Create diagnostic plots of MCMC parameters*

---

## Description

Plot parameter values and effective sample size. See [effectiveSize](#) for more details on this diagnostic measure.

## Usage

```
BASiCS_DiagPlot(
  object,
  Parameter = "mu",
  Measure = c("ess", "geweke.diag", "rhat"),
  x = NULL,
  y = NULL,
  LogX = isTRUE(x %in% c("mu", "delta")),
  LogY = isTRUE(y %in% c("mu", "delta")),
  Smooth = TRUE,
  HLine = TRUE,
  na.rm = TRUE
)
```

```
BASiCS_diagPlot(...)
```

## Arguments

object	an object of class <a href="#">BASiCS_Summary</a>
Parameter	Name of the parameter to be plotted. Default Parameter = 'mu'
Measure	Character scalar specifying the diagnostic measure to plot. Current options are effective sample size, the Geweke diagnostic criterion, and the <a href="#">rhat</a> diagnostic.
x, y	Optional MCMC parameter values to be plotted on the x or y axis, respectively. If neither is supplied, Parameter will be plotted on the x axis and effective sample size will be plotted on the y axis as a density plot.
LogX, LogY	A logical value indicating whether to use a log10 transformation for the x or y axis, respectively.
Smooth	A logical value indicating whether to use smoothing (specifically hexagonal binning using <a href="#">geom_hex</a> ).
HLine	Numeric scalar or vector indicating threshold value(s) to be displayed as a dashed line on the plot when DrawHLine = TRUE. Alternatively, can be set to FALSE to disable line drawing, or TRUE to use the default thresholds.
na.rm	Logical value indicating whether NA values should be removed before calculating effective sample size.
...	Unused.



**Value**

A ggplot object.

**Author(s)**

Alan O'Callaghan

**See Also**

[BASiCS\\_Chain](#)

**Examples**

```
# Built-in example chain
data(ChainSC)

# Point estimates versus effective sample size
BASiCS_DiagPlot(ChainSC, Parameter = "mu")
# Effective sample size as colour, mu as x, delta as y.
BASiCS_DiagPlot(ChainSC, x = "mu", y = "delta")

# Point estimates versus Geweke diagnostic
BASiCS_DiagPlot(ChainSC, Parameter = "mu", Measure = "geweke.diag")
```

---

BASiCS\_DivideAndConquer

*Run divide and conquer MCMC with BASiCS*

---

**Description**

Performs MCMC inference on batches of data. Data is divided into NSubsets batches, and BASiCS\_MCMC is run on each batch separately.

**Usage**

```
BASiCS_DivideAndConquer(
  Data,
  NSubsets = 5,
  SubsetBy = c("cell", "gene"),
  Alpha = 0.05,
  WithSpikes,
  Regression,
  BPPARAM = BiocParallel::bpparam(),
  PriorParam = BASiCS_PriorParam(Data, PriorMu = "EmpiricalBayes"),
  RunName,
  StoreChains,
  StoreDir,
```

```

    Start,
    ...
  )

```

### Arguments

Data	SingleCellExperimnt object
NSubsets	The number of batches to create and perform MCMC inference with.
SubsetBy	A character value specifying whether batches should consist of a subset of the cells in Data (when SubsetBy="cell") or a subset of the genes in Data (when SubsetBy="gene").
Alpha	A numeric value specifying the statistical significance level used to determine whether the average library size or average count are significantly different between batches.
WithSpikes, Regression, PriorParam	See <a href="#">BASiCS_MCMC</a> .
BPPARAM	A <a href="#">BiocParallelParam</a> instance.
RunName, StoreChains, StoreDir, Start	Unused. If used when calling this function, they are likely to result in undefined behaviour.
...	Passed to <a href="#">BASiCS_MCMC</a> . All arguments required by <a href="#">BASiCS_MCMC</a> must be supplied here, for example N, Thin, Burn.

### Details

Subsets are chosen such that the average library size (when partitioning by cells) or average count (when partitioning by genes) is not significantly different between batches, at a significance level Alpha.

### Value

A list of [BASiCS\\_Chain](#) objects.

### References

Simple, Scalable and Accurate Posterior Interval Estimation Cheng Li and Sanvesh Srivastava and David B. Dunson arXiv (2016)

### Examples

```

bp <- BiocParallel::SnowParam()
Data <- BASiCS_MockSCE()
BASiCS_DivideAndConquer(
  Data,
  NSubsets = 2,
  SubsetBy = "gene",
  N = 8,
  Thin = 2,
  Burn = 4,

```

```

    WithSpikes = TRUE,
    Regression = TRUE,
    BPPARAM = bp
  )

```

---

BASiCS_Draw	<i>Generate a draw from the posterior of BASiCS using the generative model.</i>
-------------	---

---

### Description

BASiCS\_Draw creates a simulated dataset from the posterior of a fitted model implemented in BASiCS.

### Usage

```

BASiCS_Draw(
  Chain,
  BatchInfo = gsub(".*_Batch([0-9a-zA-Z])", "\\1", colnames(Chain@parameters[["nu"]])),
  N = sample(nrow(Chain@parameters[["nu"]]), 1)
)

```

### Arguments

Chain	An object of class <a href="#">BASiCS_Chain</a> .
BatchInfo	Vector of batch information from the <code>SingleCellExperiment</code> object used as input to <code>BASiCS_MCMC</code> .
N	The integer index for the draw to be used to sample from the posterior predictive distribution. If not supplied, a random value is chosen.

### Value

An object of class [SingleCellExperiment](#), including synthetic data generated by the model implemented in BASiCS.

### Author(s)

Alan O'Callaghan

### References

Vallejos, Marioni and Richardson (2015). PLoS Computational Biology.

**Examples**

```
data(ChainSC)
BASiCS_Draw(ChainSC)

data(ChainSC)
BASiCS_Draw(ChainSC)
```

---

BASiCS\_EffectiveSize *Calculate effective sample size for BASiCS\_Chain parameters*

---

**Description**

A function to calculate effective sample size [BASiCS\\_Chain](#) objects.

**Usage**

```
BASiCS_EffectiveSize(object, Parameter, na.rm = TRUE)

BASiCS_effectiveSize(...)
```

**Arguments**

object	an object of class <a href="#">BASiCS_Chain</a> .
Parameter	The parameter to use to calculate effective sample size. Possible values: 'mu', 'delta', 'phi', 's', 'nu', 'theta', 'beta', 'sigma2' and 'epsilon'.
na.rm	Remove NA values before calculating effective sample size. Only relevant when Parameter = "epsilon" (genes with very low expression are excluding when inferring the mean/over-dispersion trend. Default: na.rm = TRUE.
...	Unused.

**Value**

A vector with effective sample sizes for all the elements of Parameter

**Examples**

```
data(ChainSC)
BASiCS_EffectiveSize(ChainSC, Parameter = "mu")
```

---

BASiCS\_Filter                      *Filter for input datasets*

---

### Description

BASiCS\_Filter indicates which transcripts and cells pass a pre-defined inclusion criteria. The output of this function used to generate a [SingleCellExperiment](#) object required to run BASiCS. For more systematic tools for quality control, please refer to the scater Bioconductor package.

### Usage

```
BASiCS_Filter(
  Counts,
  Tech = rep(FALSE, nrow(Counts)),
  SpikeInput = NULL,
  BatchInfo = NULL,
  MinTotalCountsPerCell = 2,
  MinTotalCountsPerGene = 2,
  MinCellsWithExpression = 2,
  MinAvCountsPerCellsWithExpression = 2
)
```

### Arguments

Counts	Matrix of dimensions $q$ times $n$ whose elements corresponds to the simulated expression counts. First $q$ .bio rows correspond to biological genes. Last $q$ - $q$ .bio rows correspond to technical spike-in genes.
Tech	Logical vector of length $q$ . If Tech = FALSE the gene is biological; otherwise the gene is spike-in.
SpikeInput	Vector of length $q$ - $q$ .bio whose elements indicate the simulated input concentrations for the spike-in genes.
BatchInfo	Vector of length $n$ whose elements indicate batch information. Not required if a single batch is present on the data. Default: BatchInfo = NULL.
MinTotalCountsPerCell	Minimum value of total expression counts required per cell (biological and technical). Default: MinTotalCountsPerCell = 2.
MinTotalCountsPerGene	Minimum value of total expression counts required per transcript (biological and technical). Default: MinTotalCountsPerGene = 2.
MinCellsWithExpression	Minimum number of cells where expression must be detected (positive count). Criteria applied to each transcript. Default: MinCellsWithExpression = 2.
MinAvCountsPerCellsWithExpression	Minimum average number of counts per cells where expression is detected. Criteria applied to each transcript. Default value: MinAvCountsPerCellsWithExpression = 2.

**Value**

A list of 2 elements

Counts Filtered matrix of expression counts

Tech Filtered vector of spike-in indicators

SpikeInput Filtered vector of spike-in genes input molecules

BatchInfo Filtered vector of the 'BatchInfo' argument

IncludeGenes Inclusion indicators for transcripts

IncludeCells Inclusion indicators for cells

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

**Examples**

```
set.seed(1)
Counts <- matrix(rpois(50*10, 2), ncol = 10)
rownames(Counts) <- c(paste0('Gene', 1:40), paste0('Spike', 1:10))
Tech <- c(rep(FALSE,40),rep(TRUE,10))
set.seed(2)
SpikeInput <- rgamma(10,1,1)
SpikeInfo <- data.frame('SpikeID' = paste0('Spike', 1:10),
                      'SpikeInput' = SpikeInput)

Filter <- BASiCS_Filter(Counts, Tech, SpikeInput,
                      MinTotalCountsPerCell = 2,
                      MinTotalCountsPerGene = 2,
                      MinCellsWithExpression = 2,
                      MinAvCountsPerCellsWithExpression = 2)
SpikeInfoFilter <- SpikeInfo[SpikeInfo$SpikeID %in% rownames(Filter$Counts),]
```

---

BASiCS_LoadChain	<i>Loads pre-computed MCMC chains generated by the <a href="#">BASiCS_MCMC</a> function</i>
------------------	---

---

**Description**

Loads pre-computed MCMC chains generated by the [BASiCS\\_MCMC](#) function, creating a [BASiCS\\_Chain](#) object

**Usage**

```
BASiCS_LoadChain(RunName = "", StoreDir = getwd(), StoreUpdatedChain = FALSE)
```

**Arguments**

RunName	String used to index '.Rds' file containing the MCMC chain (produced by the <a href="#">BASiCS_MCMC</a> function, with StoreChains = TRUE)
StoreDir	Directory where '.Rds' file is stored. Default: StoreDir = getwd()
StoreUpdatedChain	Only required when the input files contain an outdated version of a <a href="#">BASiCS_Chain</a> object. If StoreUpdatedChain = TRUE, an updated object is saved (this overwrites original input file, if it was an '.Rds' file).

**Value**

An object of class [BASiCS\\_Chain](#).

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

**See Also**

[BASiCS\\_Chain](#)

**Examples**

```
Data <- makeExampleBASiCS_Data()
Chain <- BASiCS_MCMC(
  Data,
  N = 50,
  Thin = 5,
  Burn = 5,
  Regression = FALSE,
  StoreChains = TRUE,
  StoreDir = tempdir(),
  RunName = "Test"
)
ChainLoad <- BASiCS_LoadChain(RunName = "Test", StoreDir = tempdir())
```

---

BASiCS\_MCMC

*BASiCS MCMC sampler*

---

**Description**

MCMC sampler to perform Bayesian inference for single-cell mRNA sequencing datasets using the model described in Vallejos et al (2015).

**Usage**

```

BASiCS_MCMC(
  Data,
  N,
  Thin,
  Burn,
  Regression,
  WithSpikes = TRUE,
  PriorParam = BASiCS_PriorParam(Data, PriorMu = "EmpiricalBayes"),
  FixNu = FALSE,
  SubsetBy = c("none", "gene", "cell"),
  NSubsets = 1,
  CombineMethod = c("pie", "consensus"),
  Weighting = c("naive", "n_weight", "inverse_variance"),
  Threads = getOption("Ncpus", default = 1L),
  BPPARAM = BiocParallel::bpparam(),
  ...
)

```

**Arguments**

Data	A <a href="#">SingleCellExperiment</a> object. If <code>WithSpikes = TRUE</code> , this MUST be formatted to include the spike-ins and/or batch information (see vignette).
N	Total number of iterations for the MCMC sampler. Use $N \geq \max(4, \text{Thin})$ , N being a multiple of Thin.
Thin	Thinning period for the MCMC sampler. Use $\text{Thin} \geq 2$ .
Burn	Burn-in period for the MCMC sampler. Use $\text{Burn} \geq 1$ , $\text{Burn} < N$ , Burn being a multiple of Thin.
Regression	If <code>Regression = TRUE</code> , BASiCS exploits a joint prior formulation for mean and over-dispersion parameters to estimate a measure of residual over-dispersion is not confounded by mean expression. Recommended setting is <code>Regression = TRUE</code> .
WithSpikes	If <code>WithSpikes = TRUE</code> , BASiCS will use reads from added spike-ins to estimate technical variability. If <code>WithSpikes = FALSE</code> , BASiCS depends on replicated experiments (batches) to estimate technical variability. In this case, please supply the <code>BatchInfo</code> vector in <code>colData(Data)</code> . Default: <code>WithSpikes = TRUE</code> .
PriorParam	List of prior parameters for BASiCS_MCMC. Should be created using <a href="#">BASiCS_PriorParam</a> .
FixNu	Should the scaling normalisation factor <code>nu</code> be fixed to the starting value when <code>WithSpikes=FALSE</code> ? These are set to scan scaling normalisation factors.
SubsetBy	Character value specifying whether a divide and conquer inference strategy should be used. When this is set to "gene", inference is performed on batches of genes separately, and when it is set to "cell", inference is performed on batches of cells separately. Posterior distributions are combined using posterior interval estimation (see Li et al., 2016).
NSubsets	If <code>SubsetBy="gene"</code> or <code>SubsetBy="cell"</code> , <code>NSubsets</code> specifies the number of batches to create and perform divide and conquer inference with.



CombineMethod	The method used to combine subposteriors if SubsetBy is set to "gene" or "cell". Options are "pie" corresponding to posterior interval estimation (see Li et al., 2016) or "consensus" (see Scott et al., 2016). Both of these methods use a form of weighted average to combine subposterior draws into the final posterior.
Weighting	The weighting method used in the weighted average chosen using CombineMethod. Available options are "naive" (unweighted), "n_weight" (weights are chosen based on the size of each partition) and "inverse_variance" (subposteriors are weighted based on the inverse of the variance of the subposterior for each parameter).
Threads	Integer specifying the number of threads to be used to parallelise parameter updates. Default value is the globally set "Ncpus" option, or 1 if this option is not set.
BPPARAM	A <a href="#">BiocParallelParam</a> instance, used for divide and conquer inference.
...	Optional parameters.
	AR Optimal acceptance rate for adaptive Metropolis Hastings updates. It must be a positive number between 0 and 1. Default (and recommended): AR = 0.44.
	StopAdapt Iteration at which adaptive proposals are not longer adapted. Use StopAdapt>=1. Default: StopAdapt = Burn.
	StoreChains If StoreChains = TRUE, the generated BASiCS_Chain object is stored as a '.Rds' file (RunName argument used to index the file name). Default: StoreChains = FALSE.
	StoreAdapt If StoreAdapt = TRUE, trajectory of adaptive proposal variances (in log-scale) for all parameters is stored as a list in a '.Rds' file (RunName argument used to index file name). Default: StoreAdapt = FALSE.
	StoreDir Directory where output files are stored. Only required if StoreChains = TRUE and/or StoreAdapt = TRUE. Default: StoreDir = getwd().
	RunName String used to index '.Rds' files storing chains and/or adaptive proposal variances.
	PrintProgress If PrintProgress = FALSE, console-based progress report is suppressed.
	Start Starting values for the MCMC sampler. We do not advise to use this argument. Default options have been tuned to facilitate convergence. If changed, it must be a list containing the following elements: mu0, delta0, phi0, s0, nu0, theta0, ls.mu0, ls.delta0, ls.phi0, ls.nu0 and ls.theta0
	GeneExponent/CellExponent Exponents applied to the prior for MCMC updates. Intended for use only when performing divide & conquer MCMC strategies.

### Value

An object of class [BASiCS\\_Chain](#).

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

**References**

Vallejos, Marioni and Richardson (2015). PLoS Computational Biology.

Vallejos, Richardson and Marioni (2016). Genome Biology.

Eling et al (2018). Cell Systems

Simple, Scalable and Accurate Posterior Interval Estimation Cheng Li and Sanvesh Srivastava and David B. Dunson arXiv (2016)

Bayes and Big Data: The Consensus Monte Carlo Algorithm Steven L. Scott, Alexander W. Blocker, Fernando V. Bonassi, Hugh A. Chipman, Edward I. George and Robert E. McCulloch International Journal of Management Science and Engineering Management (2016)

**Examples**

```
# Built-in simulated dataset
set.seed(1)
Data <- makeExampleBASiCS_Data()
# To analyse real data, please refer to the instructions in:
# https://github.com/catavallejos/BASiCS/wiki/2.-Input-preparation

# Only a short run of the MCMC algorithm for illustration purposes
# Longer runs might be required to reach convergence
Chain <- BASiCS_MCMC(Data, N = 50, Thin = 2, Burn = 10, Regression = FALSE,
  PrintProgress = FALSE, WithSpikes = TRUE)

# To run the regression version of BASiCS, use:
Chain <- BASiCS_MCMC(Data, N = 50, Thin = 2, Burn = 10, Regression = TRUE,
  PrintProgress = FALSE, WithSpikes = TRUE)

# To run the non-spike version BASiCS requires the data to contain at least
# 2 batches:
set.seed(2)
Data <- makeExampleBASiCS_Data(WithBatch = TRUE)
Chain <- BASiCS_MCMC(Data, N = 50, Thin = 2, Burn = 10, Regression = TRUE,
  PrintProgress = FALSE, WithSpikes = FALSE)

# For illustration purposes we load a built-in 'BASiCS_Chain' object
# (obtained using the 'BASiCS_MCMC' function)
data(ChainSC)

# `displayChainBASiCS` can be used to extract information from this output.
# For example:
head(displayChainBASiCS(ChainSC, Param = 'mu'))

# Traceplot (examples only)
plot(ChainSC, Param = 'mu', Gene = 1)
```

```

plot(ChainSC, Param = 'phi', Cell = 1)
plot(ChainSC, Param = 'theta', Batch = 1)

# Calculating posterior medians and 95% HPD intervals
ChainSummary <- Summary(ChainSC)

# `displaySummaryBASiCS` can be used to extract information from this output
# For example:
head(displaySummaryBASiCS(ChainSummary, Param = 'mu'))

# Graphical display of posterior medians and 95% HPD intervals
# For example:
plot(ChainSummary, Param = 'mu', main = 'All genes')
plot(ChainSummary, Param = 'mu', Genes = 1:10, main = 'First 10 genes')
plot(ChainSummary, Param = 'phi', main = 'All cells')
plot(ChainSummary, Param = 'phi', Cells = 1:5, main = 'First 5 cells')
plot(ChainSummary, Param = 'theta')

# To contrast posterior medians of cell-specific parameters
# For example:
par(mfrow = c(1,2))
plot(ChainSummary, Param = 'phi', Param2 = 's', SmoothPlot = FALSE)
# Recommended for large numbers of cells
plot(ChainSummary, Param = 'phi', Param2 = 's', SmoothPlot = TRUE)

# To contrast posterior medians of gene-specific parameters
par(mfrow = c(1,2))
plot(ChainSummary, Param = 'mu', Param2 = 'delta', log = 'x',
     SmoothPlot = FALSE)
# Recommended
plot(ChainSummary, Param = 'mu', Param2 = 'delta', log = 'x',
     SmoothPlot = TRUE)

# To obtain denoised rates / counts, see:
# help(BASiCS_DenoisedRates)
# and
# help(BASiCS_DenoisedCounts)

# For examples of differential analyses between 2 populations of cells see:
# help(BASiCS_TestDE)

```

---

BASiCS\_MockSCE

*Create a mock [SingleCellExperiment](#) object.*


---

## Description

Creates a [SingleCellExperiment](#) object of Poisson-distributed approximating a homogeneous cell population.

**Usage**

```
BASiCS_MockSCE(
  NGenes = 100,
  NCells = 100,
  NSpikes = 20,
  WithBatch = TRUE,
  MeanMu = 1
)
```

**Arguments**

NGenes	Integer value specifying the number of genes that will be present in the output.
NCells	Integer value specifying the number of cells that will be present in the output.
NSpikes	Integer value specifying the number of spike-in genes that will be present in the output.
WithBatch	Logical value specifying whether a dummy BatchInfo is included in the output.
MeanMu	The log mean used to generate per-gene mean expression levels.

**Value**

A [SingleCellExperiment](#) object.

**Examples**

```
BASiCS_MockSCE()
```

---

BASiCS\_PlotDE

*Produce plots assessing differential expression results*

---

**Description**

Produce plots assessing differential expression results

**Usage**

```
BASiCS_PlotDE(object, ...)

## S4 method for signature 'BASiCS_ResultsDE'
BASiCS_PlotDE(
  object,
  Plots = c("MA", "Volcano", "Grid"),
  Parameters = intersect(c("Mean", "Disp", "ResDisp"), names(object@Results)),
  MuX = TRUE,
  ...
)
```

```

## S4 method for signature 'BASiCS_ResultDE'
BASiCS_PlotDE(
  object,
  Plots = c("Grid", "MA", "Volcano"),
  Mu = NULL,
  TransLogit = FALSE
)

## S4 method for signature 'missing'
BASiCS_PlotDE(
  GroupLabel1,
  GroupLabel2,
  ProbThresholds = seq(0.5, 0.9995, by = 0.00025),
  Epsilon,
  EFDR,
  Table,
  Measure,
  EFDRgrid,
  EFNRgrid,
  ProbThreshold,
  Mu,
  TransLogit = FALSE,
  Plots = c("Grid", "MA", "Volcano")
)

```

### Arguments

object	A <a href="#">BASiCS_ResultsDE</a> or <a href="#">BASiCS_ResultDE</a> object.
...	Passed to methods.
Plots	Plots plot to produce? Options: "MA", "Volcano", "Grid".
Parameters	Character vector specifying the parameter(s) to produce plots for, Available options are "Mean", (mu, mean expression), "Disp" (delta, overdispersion) and "ResDisp" (epsilon, residual overdispersion).
MuX	Use Mu (mean expression across both chains) as the X-axis for all MA plots? Default: TRUE.
Mu, GroupLabel1, GroupLabel2, ProbThresholds, Epsilon, EFDR, Table, Measure, EFDRgrid, EFNRgrid, ProbThreshold	Internal arguments.
TransLogit	Logical scalar controlling whether a logit transform is applied to the posterior probability in the y-axis of volcano plots. As <code>logit(0)</code> and <code>logit(1)</code> are undefined, we clip these values near the range of the data excluding 0 and 1.

### Value

A plot (possibly several combined using [plot\\_grid](#)).

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

Alan O'Callaghan

**Examples**

```
data(ChainSC)
data(ChainRNA)

Test <- BASiCS_TestDE(Chain1 = ChainSC, Chain2 = ChainRNA,
                     GroupLabel1 = 'SC', GroupLabel2 = 'P&S',
                     EpsilonM = log2(1.5), EpsilonD = log2(1.5),
                     OffSet = TRUE)

BASiCS_PlotDE(Test)
```

---

BASiCS\_PlotOffset      *Visualise global offset in mean expression between two chains.*

---

**Description**

Visualise global offset in mean expression between two BASiCS\_Chain objects.

**Usage**

```
BASiCS_PlotOffset(
  Chain1,
  Chain2,
  Type = c("offset estimate", "before-after", "MAPlot"),
  GroupLabel1 = "Group 1",
  GroupLabel2 = "Group 2"
)
```

**Arguments**

Chain1, Chain2    **BASiCS\_Chain** objects to be plotted.

Type              The type of plot generated. "offset estimate" produces a boxplot of the offset alongside an estimate of the global offset. "before-after" produces MA plots of Mean expression against  $\log_2(\text{fold-change})$  before and after offset correction. "MA plot" produces an MA plot of Mean expression against  $\log_2(\text{fold-change})$ .

GroupLabel1, GroupLabel2    Labels for Chain1 and Chain2 in the resulting plot(s).

**Value**

Plot objects.

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

Alan O'Callaghan

**Examples**

```
# Loading two 'BASiCS_Chain' objects (obtained using 'BASiCS_MCMC')
data("ChainSC")
data("ChainRNA")

BASiCS_PlotOffset(ChainSC, ChainRNA)
```

---

```
BASiCS_PlotVarianceDecomp
```

*Plot variance decomposition results.*

---

**Description**

Plot variance decomposition results.

**Usage**

```
BASiCS_PlotVarianceDecomp(
  Decomp,
  beside = FALSE,
  nBatch = ((ncol(Decomp) - 2)/3) - 1,
  main = "Overall variance decomposition",
  xlabs = if (nBatch == 1) "Overall" else c("Overall", paste("Batch", seq_len(nBatch))),
  ylab = "% of variance"
)
```

**Arguments**

Decomp	The output of <a href="#">BASiCS_VarianceDecomp</a> .
beside	If TRUE, bars are placed beside each other. If FALSE, bars are stacked.
nBatch	Number of batches.
main	Plot title.
xlabs	x-axis labels. Defaults to "Batch 1", "Batch 2", etc.
ylab	y axis label.

**Value**

A ggplot object.

---

BASiCS\_PlotVG      *Plots of HVG/LVG search.*

---

**Description**

Plots of HVG/LVG search.

**Usage**

```
BASiCS_PlotVG(object, Plot = c("Grid", "VG"), ...)
```

**Arguments**

object	<a href="#">BASiCS_ResultVG</a> object.
Plot	Character scalar specifying the type of plot to be made. Options are "Grid" and "VG".
...	Optional graphical parameters passed to <code>.VGPlot</code> (internal function).

**Value**

A plot.

**Examples**

```
data(ChainSC)

# Highly and lowly variable genes detection (within a single group of cells)
DetectHVG <- BASiCS_DetectHVG(ChainSC, VarThreshold = 0.60,
                             EFDR = 0.10, Plot = TRUE)

BASiCS_PlotVG(DetectHVG)
```

---

BASiCS\_PriorParam      *Prior parameters for BASiCS\_MCMC*

---

**Description**

This is a convenience function to allow partial specification of prior parameters, and to ensure default parameters are consistent across usage within the package.



**Usage**

```

BASiCS_PriorParam(
  Data,
  k = 12,
  mu.mu = NULL,
  s2.mu = 0.5,
  s2.delta = 0.5,
  a.delta = 1,
  b.delta = 1,
  p.phi = rep(1, times = ncol(Data)),
  a.s = 1,
  b.s = 1,
  a.theta = 1,
  b.theta = 1,
  RBFMinMax = TRUE,
  FixLocations = !is.null(RBFLocations) | !is.na(MinGenesPerRBF),
  RBFLocations = NULL,
  MinGenesPerRBF = NA,
  variance = 1.2,
  m = numeric(k),
  V = diag(k),
  a.sigma2 = 2,
  b.sigma2 = 2,
  eta = 5,
  PriorMu = c("default", "EmpiricalBayes"),
  PriorDelta = c("log-normal", "gamma"),
  StochasticRef = TRUE,
  ConstrainProp = 0.2,
  GeneExponent = 1,
  CellExponent = 1
)

```

**Arguments**

Data	<a href="#">SingleCellExperiment</a> object (required).
k	Number of regression terms, including k - 2 Gaussian radial basis functions (GRBFs).
mu.mu, s2.mu	Mean and variance parameters for lognormal prior on mu.
s2.delta	Variance parameter for lognormal prior on delta when PriorDelta="lognormal".
a.delta, b.delta	Parameters for gamma prior on delta when PriorDelta="gamma".
p.phi	Parameter for dirichlet prior on phi.
a.s, b.s	Parameters for gamma prior on s.
a.theta, b.theta	Parameters for gamma prior on theta.
RBFMinMax	Should GRBFs be placed at the minimum and maximum of log(mu)?

FixLocations	Should RBFLocations be fixed throughout MCMC, or adaptive during burn-in? By default this is FALSE, but it is set to TRUE if RBFLocations or MinGenesPerRBF are specified.
RBFLocations	Numeric vector specifying locations of GRBFs in units of $\log(\mu)$ .
MinGenesPerRBF	Numeric scalar specifying the minimum number of genes for GRBFs to be retained. If fewer than MinGenesPerRBF genes have values of $\log(\mu)$ within the range of an RBF, it is removed. The range covered by each RBF is defined as centre of the RBF plus or minus half the distance between RBFs.
variance	Variance of the GRBFs.
m, V	Mean and (co)variance priors for regression coefficients.
a.sigma2, b.sigma2	Priors for inverse gamma prior on regression scale.
eta	Degrees of freedom for t distribution of regression errors.
PriorMu	Indicates if the original prior (PriorMu = 'default') or an empirical Bayes approach (PriorMu = 'EmpiricalBayes') will be assigned to gene-specific mean expression parameters.
PriorDelta	Scalar character specifying the prior type to use for delta overdispersion parameter. Options are "log-normal" (recommended) and "gamma" (used in Vallejos et al. (2015)).
StochasticRef	Logical scalar specifying whether the reference gene for the no-spikes version should be chosen randomly at MCMC iterations.
ConstrainProp	Proportion of genes to be considered as reference genes if StochasticRef=TRUE.
GeneExponent, CellExponent	Exponents for gene and cell-specific parameters. These should not be outside of divide and conquer MCMC applications.

### Value

A list containing the prior hyper-parameters that are required to run the algorithm implemented in [BASiCS\\_MCMC](#).

### Examples

```
BASiCS_PriorParam(makeExampleBASiCS_Data(), k = 12)
```

---

BASiCS\_Result

*The BASiCS\_Result class*

---

### Description

Container of results for a single test (HVG/LVG/DE). This should be an abstract class (but this is R so no) and shouldn't be directly instantiated. Defines a very small amount of common behaviour for [BASiCS\\_ResultDE](#) and [BASiCS\\_ResultVG](#).

**Slots**

Table Tabular results for each gene.

Name The name of the test performed (typically "Mean", "Disp" or "ResDisp")

ProbThreshold Posterior probability threshold used in differential test.

EFDR, EFNR Expected false discovery and expected false negative rates for differential test.

Extra Additional objects for class flexibility.

---

BASiCS\_ResultDE      *The BASiCS\_ResultDE class*

---

**Description**

Container of results for a single differential test.

**Slots**

Table Tabular results for each gene.

Name The name of the test performed (typically "Mean", "Disp" or "ResDisp")

GroupLabel1, GroupLabel2 Group labels.

ProbThreshold Posterior probability threshold used in differential test.

EFDR, EFNR Expected false discovery and expected false negative rates for differential test.

EFDRgrid, EFNRgrid Grid of EFDR and EFNR values calculated before thresholds were fixed.

Epsilon Minimum fold change or difference threshold.

Extra objects for class flexibility.

---

BASiCS\_ResultsDE      *The BASiCS\_ResultsDE class*

---

**Description**

Results of BASiCS\_TestDE

**Slots**

Results [BASiCS\\_ResultDE](#) objects

Chain1, Chain2 [BASiCS\\_Chain](#) objects.

GroupLabel1, GroupLabel2 Labels for Chain1 and Chain2

Offset Ratio between median of chains

RowData Annotation for genes

Extras Slot for extra information to be added later

---

BASiCS\_ResultVG      *The BASiCS\_ResultVG class*

---

### Description

Container of results for a single HVG/LVG test.

### Slots

Method Character value detailing whether the test performed using a threshold directly on epsilon values (Method="Epsilon"), variance decomposition (Method="Variance") or percentiles of epsilon (Method="Percentile").

RowData Optional [DataFrame](#) containing additional information about genes used in the test.

EFDRgrid,EFNRgrid Grid of EFDR and EFNR values calculated before thresholds were fixed.

Threshold Threshold used to calculate tail posterior probabilities for the HVG or LVG decision rule.

ProbThresholds Probability thresholds used to calculate EFDRgrid and EFNRgrid.

ProbThreshold Posterior probability threshold used in the HVG/LVG decision rule.

---

BASiCS\_ShowFit      *Plotting the trend after Bayesian regression*

---

### Description

Plotting the trend after Bayesian regression using a [BASiCS\\_Chain](#) object

### Usage

```
BASiCS_ShowFit(
  object,
  xlab = "log(mu)",
  ylab = "log(delta)",
  pch = 16,
  smooth = TRUE,
  variance = 1.2,
  colour = "dark blue",
  markExcludedGenes = TRUE,
  GenesSel = NULL,
  colourGenesSel = "dark red",
  Uncertainty = TRUE
)
```

**Arguments**

object	an object of class <code>BASiCS_Chain</code>
xlab	As in <code>par</code> .
ylab	As in <code>par</code> .
pch	As in <code>par</code> . Default value <code>pch = 16</code> .
smooth	Logical to indicate whether the <code>smoothScatter</code> function is used to plot the scatter plot. Default value <code>smooth = TRUE</code> .
variance	Variance used to build GRBFs for regression. Default value <code>variance = 1.2</code>
colour	colour used to denote genes within the scatterplot. Only used when <code>smooth = TRUE</code> . Default value <code>colour = "dark blue"</code> .
markExcludedGenes	Whether or not lowly expressed genes that were excluded from the regression fit are included in the scatterplot. Default value <code>markExcludedGenes = TRUE</code> .
GenesSel	Vector of gene names to be highlighted in the scatterplot. Only used when <code>smooth = TRUE</code> . Default value <code>GenesSel = NULL</code> .
colourGenesSel	colour used to denote the genes listed in <code>GenesSel</code> within the scatterplot. Default value <code>colourGenesSel = "dark red"</code> .
Uncertainty	logical indicator. If true, statistical uncertainty around the regression fit is shown in the plot.

**Value**

A `ggplot2` object

**Author(s)**

Nils Eling <eling@ebi.ac.uk>

Catalina Vallejos <cnvallej@uc.cl>

**References**

Eling et al (2018). Cell Systems <https://doi.org/10.1016/j.cels.2018.06.011>

**Examples**

```
data(ChainRNAReg)
BASiCS_ShowFit(ChainRNAReg)
```

---

BASiCS_Sim	<i>Generates synthetic data according to the model implemented in BASiCS</i>
------------	--

---

### Description

BASiCS\_Sim creates a simulated dataset from the model implemented in BASiCS.

### Usage

```
BASiCS_Sim(Mu, Mu_spikes = NULL, Delta, Phi = NULL, S, Theta, BatchInfo = NULL)
```

### Arguments

Mu	Gene-specific mean expression parameters $\mu_i$ for all biological genes (vector of length <code>q.bio</code> , all elements must be positive numbers)
Mu_spikes	$\mu_i$ for all technical genes defined as true input molecules (vector of length <code>q-q.bio</code> , all elements must be positive numbers). If <code>Mu_spikes = NULL</code> , the generated data will not contain spike-ins. If <code>Phi = NULL</code> , <code>Mu_spikes</code> will be ignored. Default: <code>Mu_spikes = NULL</code> .
Delta	Gene-specific biological over-dispersion parameters $\delta_i$ , biological genes only (vector of length <code>q.bio</code> , all elements must be positive numbers)
Phi	Cell-specific mRNA content normalising parameters $\phi_j$ (vector of length <code>n</code> , all elements must be positive numbers and the sum of its elements must be equal to <code>n</code> ). <code>Phi</code> must be set equal to <code>NULL</code> when generating data without spike-ins. If <code>Mu_spikes = NULL</code> , <code>Phi</code> will be ignored. Default: <code>Phi = NULL</code>
S	Cell-specific technical normalising parameters $s_j$ (vector of length <code>n</code> , all elements must be positive numbers)
Theta	Technical variability parameter $\theta$ (must be positive). <code>Theta</code> can be a scalar (single batch of samples), or a vector (multiple batches of samples). If a value for <code>BatchInfo</code> is provided, the length of <code>Theta</code> must match the number of unique values in <code>BatchInfo</code> .
BatchInfo	Vector detailing which batch each cell should be simulated from. If spike-ins are not in use, the number of unique values contained in <code>BatchInfo</code> must be larger than 1 (i.e. multiple batches are present).

### Value

An object of class [SingleCellExperiment](#), including synthetic data generated by the model implemented in BASiCS.

### Author(s)

Catalina A. Vallejos <cnvallej@uc.cl>, Nils Eling

## References

Vallejos, Marioni and Richardson (2015). PLoS Computational Biology.

## Examples

```
# Simulated parameter values for 10 genes
# (7 biological and 3 spike-in) measured in 5 cells
Mu <- c(8.36, 10.65, 4.88, 6.29, 21.72, 12.93, 30.19)
Mu_spikes <- c(1010.72, 7.90, 31.59)
Delta <- c(1.29, 0.88, 1.51, 1.49, 0.54, 0.40, 0.85)
Phi <- c(1.00, 1.06, 1.09, 1.05, 0.80)
S <- c(0.38, 0.40, 0.38, 0.39, 0.34)
Theta <- 0.39

# Data with spike-ins, single batch
Data <- BASiCS_Sim(Mu, Mu_spikes, Delta, Phi, S, Theta)
head(SingleCellExperiment::counts(Data))
dim(SingleCellExperiment::counts(Data))
altExp(Data)
rowData(altExp(Data))

# Data with spike-ins, multiple batches
BatchInfo <- c(1,1,1,2,2)
Theta2 <- rep(Theta, times = 2)
Data <- BASiCS_Sim(Mu, Mu_spikes, Delta, Phi, S, Theta2, BatchInfo)

# Data without spike-ins, multiple batches
Data <- BASiCS_Sim(Mu, Mu_spikes = NULL, Delta,
                  Phi = NULL, S, Theta2, BatchInfo)
```

---

BASiCS\_Summary

*The BASiCS\_Summary class*

---

## Description

Container of a summary of a [BASiCS\\_Chain](#) object. In each element of the parameters slot, first column contains posterior medians; second and third columns respectively contain the lower and upper limits of an high posterior density interval (for a given probability).

## Slots

**parameters** List of parameters in which each entry contains a matrix: first column contains posterior medians, second column contains the lower limits of an high posterior density interval and third column contains the upper limits of high posterior density intervals.

**mu** Posterior medians (1st column), lower (2nd column) and upper (3rd column) limits of gene-specific mean expression parameters  $\mu_i$ .

**delta** Posterior medians (1st column), lower (2nd column) and upper (3rd column) limits of gene-specific biological over-dispersion parameters  $\delta_i$ , biological genes only

- phi** Posterior medians (1st column), lower (2nd column) and upper (3rd column) limits of cell-specific mRNA content normalisation parameters  $\phi_j$
- s** Posterior medians (1st column), lower (2nd column) and upper (3rd column) limits of cell-specific technical normalisation parameters  $s[j]$
- nu** Posterior medians (1st column), lower (2nd column) and upper (3rd column) limits of cell-specific random effects  $\nu_j$
- theta** Posterior median (1st column), lower (2nd column) and upper (3rd column) limits of technical over-dispersion parameter(s)  $\theta$  (each row represents one batch)
- beta** Posterior median (first column), lower (second column) and upper (third column) limits of regression coefficients  $\beta$
- sigma2** Posterior median (first column), lower (second column) and upper (third column) limits of residual variance  $\sigma^2$
- epsilon** Posterior median (first column), lower (second column) and upper (third column) limits of gene-specific residual over-dispersion parameter  $\epsilon$

## Examples

```
# A BASiCS_Summary object created by the Summary method.
Data <- makeExampleBASiCS_Data()
Chain <- BASiCS_MCMC(Data, N = 100, Thin = 2, Burn = 2, Regression = FALSE)
ChainSummary <- Summary(Chain)
```

---

BASiCS\_Summary-methods

*'show' method for BASiCS\_Summary objects*

---

## Description

'show' method for [BASiCS\\_Summary](#) objects.

## Usage

```
## S4 method for signature 'BASiCS_Summary'
show(object)
```

## Arguments

object            A [BASiCS\\_Summary](#) object.

## Value

Prints a summary of the properties of object.



**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

**Examples**

```
data(ChainSC)
show(ChainSC)
```

---

BASiCS\_TestDE

*Detection of genes with changes in expression*

---

**Description**

Function to assess changes in expression between two groups of cells (mean and over-dispersion)

**Usage**

```
BASiCS_TestDE(
  Chain1,
  Chain2,
  EpsilonM = log2(1.5),
  EpsilonD = log2(1.5),
  EpsilonR = log2(1.5)/log2(exp(1)),
  ProbThresholdM = 2/3,
  ProbThresholdD = 2/3,
  ProbThresholdR = 2/3,
  OrderVariable = c("GeneIndex", "GeneName", "Mu"),
  GroupLabel1 = "Group1",
  GroupLabel2 = "Group2",
  Plot = TRUE,
  PlotOffset = TRUE,
  PlotOffsetType = c("offset estimate", "before-after", "MA plot"),
  Offset = TRUE,
  EFDR_M = 0.05,
  EFDR_D = 0.05,
  EFDR_R = 0.05,
  GenesSelect = rep(TRUE, ncol(Chain1@parameters[["mu"]])),
  min.mean = 1,
  MinESS = 100,
  ...
)
```

**Arguments**

Chain1	an object of class <a href="#">BASiCS_Chain</a> containing parameter estimates for the first group of cells
Chain2	an object of class <a href="#">BASiCS_Chain</a> containing parameter estimates for the second group of cells
EpsilonM	Minimum fold change tolerance threshold for detecting changes in overall expression (must be a positive real number). Default value: $\text{EpsilonM} = \log_2(1.5)$ (i.e. 50% increase).
EpsilonD	Minimum fold change tolerance threshold for detecting changes in biological over-dispersion (must be a positive real number). Default value: $\text{EpsilonD} = \log_2(1.5)$ (i.e. 50% increase).
EpsilonR	Minimum distance threshold for detecting changes in residual over-dispersion (must be a positive real number). Default value: $\text{EpsilonR} = \log_2(1.5) / \log_2(\exp(1))$ (i.e. 50% increase).
ProbThresholdM	Optional parameter. Probability threshold for detecting changes in overall expression (must be a positive value, between 0 and 1). If $\text{EFDR}_M = \text{NULL}$ , the posterior probability threshold for the differential mean expression test will be set to $\text{ProbThresholdM}$ . If a value for $\text{EFDR}_M$ is provided, the posterior probability threshold is chosen to achieve an EFDR equal to $\text{EFDR}_M$ and $\text{ProbThresholdM}$ defines a minimum probability threshold for this calibration (this avoids low values of $\text{ProbThresholdM}$ to be chosen by the EFDR calibration. Default value $\text{ProbThresholdM} = 2/3$ , i.e. the probability of observing a $\log_2$ -FC above $\text{EpsilonM}$ must be at least twice the probability of observing the complementary event ( $\log_2$ -FC below $\text{EpsilonM}$ ).
ProbThresholdD	Optional parameter. Probability threshold for detecting changes in cell-to-cell biological over-dispersion (must be a positive value, between 0 and 1). Same usage as $\text{ProbThresholdM}$ , depending on the value provided for $\text{EFDR}_D$ . Default value $\text{ProbThresholdD} = 2/3$ .
ProbThresholdR	Optional parameter. Probability threshold for detecting changes in residual over-dispersion (must be a positive value, between 0 and 1). Same usage as $\text{ProbThresholdM}$ , depending on the value provided for $\text{EFDR}_R$ . Default value $\text{ProbThresholdR} = 2/3$ .
OrderVariable	Ordering variable for output. Possible values: 'GeneIndex' (default), 'GeneName' and 'Mu' (mean expression).
GroupLabel1	Label assigned to reference group. Default: $\text{GroupLabel1} = \text{'Group1'}$
GroupLabel2	Label assigned to reference group. Default: $\text{GroupLabel2} = \text{'Group2'}$
Plot	If $\text{Plot} = \text{TRUE}$ , MA and volcano plots are generated.
PlotOffset	If $\text{Plot} = \text{TRUE}$ , the offset effect is visualised.
PlotOffsetType	See argument <code>Type</code> in <a href="#">BASiCS_PlotOffset</a> for more information.
Offset	Optional argument to remove a fix offset effect (if not previously removed from the MCMC chains). Default: $\text{Offset} = \text{TRUE}$ .
EFDR_M	Target for expected false discovery rate related to the comparison of means. If $\text{EFDR}_M = \text{NULL}$ , EFDR calibration is not performed and the posterior probability threshold is set equal to $\text{ProbThresholdM}$ . Default $\text{EFDR}_M = 0.05$ .

EFDR_D	Target for expected false discovery rate related to the comparison of dispersions. If EFDR_D = NULL, EFDR calibration is not performed and the posterior probability threshold is set equal to ProbThresholdD. Default EFDR_D = 0.05.
EFDR_R	Target for expected false discovery rate related to the comparison of residual over-dispersions. If EFDR_R = NULL, EFDR calibration is not performed and the posterior probability threshold is set equal to ProbThresholdR. Default EFDR_D = 0.05.
GenesSelect	Optional argument to provide a user-defined list of genes to be considered for the comparison. Default: GenesSelect = rep(TRUE, nGene) When used, this argument must be a vector of TRUE (include gene) / FALSE (exclude gene) indicator, with the same length as the number of intrinsic genes and following the same order as how genes are displayed in the table of counts. This argument is necessary in order to have a meaningful EFDR calibration when the user decides to exclude some genes from the comparison.
min.mean	Minimum mean expression threshold required for inclusion in offset calculation. Similar to 'min.mean' in 'scrans::computeSumFactors'. This parameter is only relevant with 'Offset = TRUE'.
MinESS	The minimum effective sample size for a gene to be included in the tests for differential expression. This helps to remove genes with poor mixing from differential expression tests. Default is 100. If set to NA, genes are not checked for effective sample size before differential expression tests are performed.
...	Optional parameters.

**Value**

BASiCS\_TestDE returns an object of class [BASiCS\\_ResultsDE](#)

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

**Examples**

```
# Loading two 'BASiCS_Chain' objects (obtained using 'BASiCS_MCMC')
data(ChainSC)
data(ChainRNA)

Test <- BASiCS_TestDE(
  Chain1 = ChainSC, Chain2 = ChainRNA,
  GroupLabel1 = "SC", GroupLabel2 = "P&S",
  EpsilonM = log2(1.5), EpsilonD = log2(1.5),
  Offset = TRUE
)

# Results for the differential mean test
head(as.data.frame(Test, Parameter = "Mean"))

# Results for the differential over-dispersion test
```

```

# This only includes genes marked as 'NoDiff' in Test$TableMean
head(as.data.frame(Test, Parameter = "Disp"))

# For testing differences in residual over-dispersion, two chains obtained
# via 'BASiCS_MCMC(Data, N, Thin, Burn, Regression=TRUE)' need to be provided
data(ChainSReg)
data(ChainRNAREg)

Test <- BASiCS_TestDE(
  Chain1 = ChainSReg, Chain2 = ChainRNAREg,
  GroupLabel1 = 'SC', GroupLabel2 = 'P&S',
  EpsilonM = log2(1.5), EpsilonD = log2(1.5),
  EpsilonR = log2(1.5)/log2(exp(1)),
  OffSet = TRUE
)

## Plotting the results of these tests
BASiCS_PlotDE(Test)

```

---

BASiCS\_VarianceDecomp *Decomposition of gene expression variability according to BASiCS*

---

## Description

Function to decompose total variability of gene expression into biological and technical components.

## Usage

```

BASiCS_VarianceDecomp(
  Chain,
  OrderVariable = c("BioVarGlobal", "GeneName", "TechVarGlobal", "ShotNoiseGlobal"),
  Plot = TRUE,
  main = "Overall variance decomposition",
  ylab = "% of variance",
  beside = FALSE,
  palette = "Set1",
  legend = c("Biological", "Technical", "Shot noise"),
  names.arg = if (nBatch == 1) "Overall" else c("Overall", paste("Batch",
    seq_len(nBatch)))
)

```

## Arguments

**Chain** an object of class `BASiCS_Chain`

**OrderVariable** Ordering variable for output. Possible values: 'GeneName', 'BioVarGlobal', 'TechVarGlobal' and 'ShotNoiseGlobal'. Default: `OrderVariable = "BioVarGlobal"`.

Plot	If TRUE, a barplot of the variance decomposition (global and by batches, if any) is generated. Default: Plot = TRUE.
main	Plot title.
ylab	y axis label.
beside	If TRUE, bars are placed beside each other. If FALSE, bars are stacked.
palette	Palette to be passed to <a href="#">scale_fill_brewer</a> to create a discrete colour mapping.
legend	Labels for variance components.
names.arg	X axis labels.

### Details

See vignette

### Value

A [data.frame](#) whose first 4 columns correspond to

GeneName Gene name (as indicated by user)

BioVarGlobal Percentage of variance explained by a biological component (overall across all cells)

TechVarGlobal Percentage of variance explained by the technical component (overall across all cells)

ShotNoiseGlobal Percentage of variance explained by the shot noise component (baseline Poisson noise, overall across all cells)

If more than 1 batch of cells are being analysed, the remaining columns contain the corresponding variance decomposition calculated within each batch.

### Author(s)

Catalina A. Vallejos <cnvallej@uc.cl>

### References

Vallejos, Marioni and Richardson (2015). PLoS Computational Biology.

### See Also

[BASiCS\\_Chain](#)

### Examples

```
# For illustration purposes we load a built-in 'BASiCS_Chain' object
# (obtained using the 'BASiCS_MCMC' function)
data(ChainSC)

VD <- BASiCS_VarianceDecomp(ChainSC)
```

---

BASiCS\_VarThresholdSearchHVG

*Detection method for highly and lowly variable genes using a grid of variance contribution thresholds*

---

### Description

Detection method for highly and lowly variable genes using a grid of variance contribution thresholds. Only used when HVG/LVG are found based on the variance decomposition.

### Usage

```
BASiCS_VarThresholdSearchVG(
  Chain,
  Task = c("HVG", "LVG"),
  VarThresholdsGrid,
  EFDR = 0.1,
  Progress = TRUE
)
```

```
BASiCS_VarThresholdSearchHVG(...)
```

```
BASiCS_VarThresholdSearchLVG(...)
```

### Arguments

Chain	an object of class <a href="#">BASiCS_Chain</a>
Task	See ?BASiCS_DetectVG.
VarThresholdsGrid	Grid of values for the variance contribution threshold (they must be contained in (0,1))
EFDR	Target for expected false discovery rate related to HVG/LVG detection. Default: EFDR = 0.10.
Progress	If Progress = TRUE, partial output is printed in the console. Default: Progress = TRUE.
...	Passed to methods.

### Details

See vignette

### Value

**BASiCS\_VarThresholdSearchHVG** A table displaying the results of highly variable genes detection for different variance contribution thresholds.

**BASiCS\_VarThresholdSearchLVG** A table displaying the results of lowly variable genes detection for different variance contribution thresholds.

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

**References**

Vallejos, Marioni and Richardson (2015). PLoS Computational Biology.

**See Also**

[BASiCS\\_Chain](#)

**Examples**

```
data(ChainSC)

BASiCS_VarThresholdSearchHVG(ChainSC,
                             VarThresholdsGrid = seq(0.55,0.65,by=0.01),
                             EFDR = 0.10)
BASiCS_VarThresholdSearchLVG(ChainSC,
                             VarThresholdsGrid = seq(0.35,0.45,by=0.01),
                             EFDR = 0.10)
```

---

ChainRNA

*Extract from the chain obtained for the Grun et al (2014) data: pool-and-split samples*

---

**Description**

Small extract (75 MCMC iterations, 350 randomly selected genes) from the chain obtained for the pool-and-split samples (this corresponds to the RNA 2i samples in Grun et al, 2014).

**Usage**

```
ChainRNA
```

**Format**

An object of class [BASiCS\\_Chain](#) containing 75 MCMC iterations.

**References**

Grun, Kester and van Oudenaarden (2014). Nature Methods.

---

ChainRNAREg	<i>Extract from the chain obtained for the Grun et al (2014) data: pool-and-split samples (regression model)</i>
-------------	--

---

**Description**

Small extract (75 MCMC iterations, 350 randomly selected genes) from the chain obtained for the pool-and-split samples (this corresponds to the RNA 2i samples in Grun et al, 2014).

**Usage**

ChainRNAREg

**Format**

An object of class [BASiCS\\_Chain](#) containing 75 MCMC iterations.

**References**

Grun, Kester and van Oudenaarden (2014). Nature Methods.

---

ChainSC	<i>Extract from the chain obtained for the Grun et al (2014) data: single-cell samples</i>
---------	--

---

**Description**

Small extract (75 MCMC iterations, 350 randomly selected genes) from the chain obtained for the pool-and-split samples (this corresponds to the SC 2i samples in Grun et al, 2014).

**Usage**

ChainSC

**Format**

An object of class [BASiCS\\_Chain](#) containing 75 MCMC iterations.

**References**

Grun, Kester and van Oudenaarden (2014). Nature Methods.



---

ChainSCReg	<i>Extract from the chain obtained for the Grun et al (2014) data: single-cell samples (regression model)</i>
------------	---

---

**Description**

Small extract (75 MCMC iterations, 350 randomly selected genes) from the chain obtained for the pool-and-split samples (this corresponds to the SC 2i samples in Grun et al, 2014).

**Usage**

```
ChainSCReg
```

**Format**

An object of class `BASiCS_Chain` containing 75 MCMC iterations.

**References**

Grun, Kester and van Oudenaarden (2014). Nature Methods.

---

dim	<i>'dim' method for BASiCS_Chain objects</i>
-----	--

---

**Description**

Returns the dimensions (genes x cells) of a `BASiCS_Chain`

**Usage**

```
## S4 method for signature 'BASiCS_Chain'  
dim(x)
```

**Arguments**

x                    A `BASiCS_Chain` object.

**Value**

An vector of dimensions

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

**Examples**

```
data(ChainSC)
dimnames(ChainSC)
```

---

dimnames	<i>'dimnames' method for BASiCS_Chain objects</i>
----------	---

---

**Description**

Returns the dimension names (genes x cells) of a [BASiCS\\_Chain](#)

**Usage**

```
## S4 method for signature 'BASiCS_Chain'
dimnames(x)
```

**Arguments**

x                    A [BASiCS\\_Chain](#) object.

**Value**

A list of two elements: (1) a vector of gene names and (2) a vector of cell names.

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

**Examples**

```
data(ChainSC)
dimnames(ChainSC)
```

---

displayChainBASiCS-BASiCS_Chain-method	<i>Accessors for the slots of a BASiCS_Chain object</i>
--	---

---

**Description**

Accessors for the slots of a [BASiCS\\_Chain](#)

**Usage**

```
## S4 method for signature 'BASiCS_Chain'
displayChainBASiCS(object, Parameter = "mu")
```

**Arguments**

object            an object of class [BASiCS\\_Chain](#)  
Parameter        Name of the slot to be used for the accessed. Possible values: 'mu', 'delta', 'phi', 's', 'nu', 'theta', 'beta', 'sigma2' and 'epsilon'.

**Value**

The requested slot of a [BASiCS\\_Chain](#) object

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

**See Also**

[BASiCS\\_Chain](#)

**Examples**

```
help(BASiCS_MCMC)
```

---

displaySummaryBASiCS-BASiCS\_Summary-method

*Accessors for the slots of a [BASiCS\\_Summary](#) object*

---

**Description**

Accessors for the slots of a [BASiCS\\_Summary](#) object

**Usage**

```
## S4 method for signature 'BASiCS_Summary'  
displaySummaryBASiCS(object, Parameter = "mu")
```

**Arguments**

object            an object of class [BASiCS\\_Summary](#)  
Parameter        Name of the slot to be used for the accessed. Possible values: 'mu', 'delta', 'phi', 's', 'nu', 'theta', 'beta', 'sigma2' and 'epsilon'.

**Value**

The requested slot of a [BASiCS\\_Summary](#) object

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>  
 Nils Eling <eling@ebi.ac.uk>

**See Also**

[BASiCS\\_Summary](#)

**Examples**

```
help(BASiCS_MCMC)
```

---

```
format,BASiCS_ResultsDE-method
```

*Methods for formatting [BASiCS\\_Result](#) and [BASiCS\\_ResultsDE](#) objects.*

---

**Description**

Methods for formatting [BASiCS\\_Result](#) and [BASiCS\\_ResultsDE](#) objects.

**Usage**

```
## S4 method for signature 'BASiCS_ResultsDE'
format(x, Parameter, Filter = TRUE, ProbThreshold = NULL, ...)
```

```
## S4 method for signature 'BASiCS_ResultDE'
format(x, Filter = TRUE, ProbThreshold = NULL, ...)
```

```
## S4 method for signature 'BASiCS_ResultVG'
format(x, Filter = TRUE, ProbThreshold = NULL, ...)
```

**Arguments**

x	Object being subsetted.
Parameter	Character scalar indicating which of the <a href="#">BASiCS_Result</a> should be formatted.
Filter	Logical scalar indicating whether results should be filtered based on differential expression or HVG/LVG status if ProbThreshold=NULL, or a probability threshold if ProbThreshold=NULL
ProbThreshold	Probability threshold to be used if Filter=TRUE
...	Passed to <a href="#">format</a> .

**Value**

A data.frame.

---

`makeExampleBASiCS_Data`

*Create a synthetic `SingleCellExperiment` example object with the format required for `BASiCS`*

---

## Description

A synthetic `SingleCellExperiment` object is generated by simulating a dataset from the model underlying `BASiCS`. This is used to illustrate `BASiCS` in some of the package and vignette examples.

## Usage

```
makeExampleBASiCS_Data(WithBatch = FALSE, WithSpikes = TRUE)
```

## Arguments

<code>WithBatch</code>	If TRUE, 2 batches are generated (each of them containing 15 cells). Default: <code>WithBatch = FALSE</code> .
<code>WithSpikes</code>	If TRUE, the simulated dataset contains 20 spike-in genes. If <code>WithSpikes = FALSE</code> , <code>WithBatch</code> is automatically set to TRUE. Default: <code>WithSpikes = TRUE</code>

## Details

Note: In `BASiCS` versions < 1.5.22, `makeExampleBASiCS_Data` used a fixed seed within the function. This has been removed to comply with Bioconductor policies. If a reproducible example is required, please use `set.seed` prior to calling `makeExampleBASiCS_Data`.

## Value

An object of class `SingleCellExperiment`, with synthetic data simulated from the model implemented in `BASiCS`. If `WithSpikes = TRUE`, it contains 70 genes (50 biological and 20 spike-in) and 30 cells. Alternatively, it contains 50 biological genes and 30 cells.

## Author(s)

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

## Examples

```
Data <- makeExampleBASiCS_Data()
is(Data, 'SingleCellExperiment')
```

---

newBASiCS\_Chain      *Creates a BASiCS\_Chain object from pre-computed MCMC chains*

---

### Description

BASiCS\_Chain creates a [BASiCS\\_Chain](#) object from pre-computed MCMC chains.

### Usage

```
newBASiCS_Chain(parameters)
```

### Arguments

parameters      List of matrices containing MCMC chains for each model parameter.

**mu**      MCMC chain for gene-specific mean expression parameters  $\mu_i$ , biological genes only (matrix with q.bio columns, all elements must be positive numbers)

**delta**      MCMC chain for gene-specific biological over-dispersion parameters  $\delta_i$ , biological genes only (matrix with q.bio columns, all elements must be positive numbers)

**phi**      MCMC chain for cell-specific mRNA content normalisation parameters  $\phi_j$  (matrix with n columns, all elements must be positive numbers and the sum of its elements must be equal to n). This parameter is only used when spike-in genes are available.

**s**      MCMC chain for cell-specific technical normalisation parameters  $s_j$  (matrix with n columns, all elements must be positive numbers)

**nu**      MCMC chain for cell-specific random effects  $\nu_j$  (matrix with n columns, all elements must be positive numbers)

**theta**      MCMC chain for technical over-dispersion parameter(s)  $\theta$  (matrix, all elements must be positive, each column represents 1 batch)

beta      Only used for regression model. MCMC chain for regression coefficients (matrix with k columns, where k represent the number of chosen basis functions + 2)

sigma2      Only used for regression model. MCMC chain for the residual variance (matrix with one column, sigma2 represents a global parameter)

epsilon      Only used for regression model. MCMC chain for the gene specific residual over-dispersion parameter (mean corrected variability) (matrix with q columns)

### Value

An object of class [BASiCS\\_Chain](#).

### Author(s)

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>



---

newBASiCS_Data	<i>Creates a <code>SingleCellExperiment</code> object from a matrix of expression counts and experimental information about spike-in genes</i>
----------------	--

---

### Description

newBASiCS\_Data creates a [SingleCellExperiment](#) object from a matrix of expression counts and experimental information about spike-in genes.

### Usage

```
newBASiCS_Data(
  Counts,
  Tech = rep(FALSE, nrow(Counts)),
  SpikeInfo = NULL,
  BatchInfo = NULL,
  SpikeType = "ERCC"
)
```

### Arguments

Counts	Matrix of dimensions q times n whose elements contain the expression counts to be analysed (including biological and technical spike-in genes). Gene names must be stored as <code>rownames(Counts)</code> .
Tech	Logical vector of length q. If <code>Tech = FALSE</code> the gene is biological; otherwise the gene is spike-in. Default value: <code>Tech = rep(FALSE, nrow(Counts))</code> .
SpikeInfo	<code>data.frame</code> whose first and second columns contain the gene names assigned to the spike-in genes (they must match the ones in <code>rownames(Counts)</code> ) and the associated input number of molecules, respectively. If <code>SpikeInfo = NULL</code> , only the horizontal integration implementation (no spikes) can be run. Default value: <code>SpikeInfo = NULL</code> .
BatchInfo	Vector of length n whose elements indicate batch information. Not required if a single batch is present on the data. Default value: <code>BatchInfo = NULL</code> .
SpikeType	Character to indicate what type of spike-ins are in use. Default value: <code>SpikeType = "ERCC"</code> (parameter is no longer used).

### Value

An object of class [SingleCellExperiment](#).

### Author(s)

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>



**See Also**[SingleCellExperiment](#)

---

`plot-BASiCS_Chain-method`*'plot' method for BASiCS\_Chain objects*

---

**Description**`'plot'` method for [BASiCS\\_Chain](#) objects**Usage**

```
## S4 method for signature 'BASiCS_Chain,ANY'
plot(
  x,
  Parameter = "mu",
  Gene = NULL,
  Cell = NULL,
  Batch = 1,
  RegressionTerm = NULL,
  ...
)
```

**Arguments**

<code>x</code>	A <a href="#">BASiCS_Chain</a> object.
<code>Parameter</code>	Name of the slot to be used for the plot. Possible values: 'mu', 'delta', 'phi', 's', 'nu', 'theta', 'beta', 'sigma2' and 'epsilon'.
<code>Gene</code>	Specifies which gene is requested. Required only if <code>Parameter = 'mu'</code> or <code>'delta'</code>
<code>Cell</code>	Specifies which cell is requested. Required only if <code>Parameter = 'phi'</code> , <code>'s'</code> or <code>'nu'</code>
<code>Batch</code>	Specifies which batch is requested. Required only if <code>Parameter = 'theta'</code>
<code>RegressionTerm</code>	Specifies which regression coefficient is requested. Required only if <code>Parameter = 'beta'</code>
<code>...</code>	Unused.

**Value**

A plot object

**Author(s)**

Catalina A. Vallejos &lt;cnvallej@uc.cl&gt;

Nils Eling &lt;eling@ebi.ac.uk&gt;

**Examples**

```
help(BASiCS_MCMC)
```

---

```
plot-BASiCS_Summary-method
      'plot' method for BASiCS_Summary objects
```

---

**Description**

'plot' method for [BASiCS\\_Summary](#) objects

**Usage**

```
## S4 method for signature 'BASiCS_Summary,ANY'
plot(
  x,
  Param = "mu",
  Param2 = NULL,
  Genes = NULL,
  Cells = NULL,
  Batches = NULL,
  RegressionTerms = NULL,
  xlab = "",
  ylab = "",
  xlim = "",
  ylim = NULL,
  pch = 16,
  col = "blue",
  bty = "n",
  SmoothPlot = TRUE,
  ...
)
```

**Arguments**

x	A <a href="#">BASiCS_Summary</a> object.
Param	Name of the slot to be used for the plot. Possible values: 'mu', 'delta', 'phi', 's', 'nu', 'theta', 'beta', 'sigma2' and 'epsilon'.
Param2	Name of the second slot to be used for the plot. Possible values: 'mu', 'delta', 'epsilon', 'phi', 's' and 'nu' (combinations between gene-specific and cell-specific parameters are not admitted).
Genes	Specifies which genes are requested. Required only if Param = 'mu', 'delta' or 'epsilon'.
Cells	Specifies which cells are requested. Required only if Param = 'phi', 's' or 'nu'

Batches	Specifies which batches are requested. Required only if Param = 'theta'
RegressionTerms	Specifies which regression coefficients are requested. Required only if Param = 'beta'
xlab	As in <a href="#">par</a> .
ylab	As in <a href="#">par</a> .
xlim	As in <a href="#">par</a> .
ylim	As in <a href="#">par</a> .
pch	As in <a href="#">par</a> .
col	As in <a href="#">par</a> .
bty	As in <a href="#">par</a> .
SmoothPlot	Logical parameter. If TRUE, transparency will be added to the color of the dots.
...	Other graphical parameters (see <a href="#">par</a> ).

**Value**

A plot object

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

**Examples**

```
help(BASiCS_MCMC)
```

---

rowData,BASiCS\_ResultsDE-method

*rowData* getter and setter for [BASiCS\\_ResultsDE](#) and [BASiCS\\_ResultVG](#) objects.

---

**Description**

rowData getter and setter for [BASiCS\\_ResultsDE](#) and [BASiCS\\_ResultVG](#) objects.

**Usage**

```
## S4 method for signature 'BASiCS_ResultsDE'  
rowData(x)  
  
## S4 replacement method for signature 'BASiCS_ResultsDE'  
rowData(x) <- value  
  
## S4 method for signature 'BASiCS_ResultVG'  
rowData(x)  
  
## S4 replacement method for signature 'BASiCS_ResultVG'  
rowData(x) <- value
```

**Arguments**

x                    [BASiCS\\_ResultVG](#) or [BASiCS\\_ResultsDE](#) object.  
value                New rowData value for setter method.

**Value**

For the getter, a [DFrame](#). For setter, the modified x.

---

show,BASiCS\_ResultDE-method

*Accessors for the slots of a [BASiCS\\_ResultDE](#) object*

---

**Description**

Accessors for the slots of a [BASiCS\\_ResultDE](#) object

**Usage**

```
## S4 method for signature 'BASiCS_ResultDE'  
show(object)
```

**Arguments**

object                an object of class [BASiCS\\_ResultDE](#)

**Value**

Prints a summary of the properties of object.

**See Also**

[show](#)

### Examples

```
help(BASiCS_MCMC)
```

---

show,BASiCS\_ResultsDE-method

*Accessors for the slots of a [BASiCS\\_ResultsDE](#) object*

---

### Description

Accessors for the slots of a [BASiCS\\_ResultsDE](#) object

### Usage

```
## S4 method for signature 'BASiCS_ResultsDE'  
show(object)
```

### Arguments

object            an object of class [BASiCS\\_ResultsDE](#)

### Value

Prints a summary of the properties of object.

### See Also

[show](#)

### Examples

```
help(BASiCS_MCMC)
```

---

show,BASiCS\_ResultVG-method

*Accessors for the slots of a [BASiCS\\_ResultVG](#) object*

---

### Description

Accessors for the slots of a [BASiCS\\_ResultVG](#) object

### Usage

```
## S4 method for signature 'BASiCS_ResultVG'  
show(object)
```

**Arguments**

object            an object of class [BASiCS\\_ResultsDE](#)

**Value**

Prints a summary of the properties of object.

**See Also**

[show](#)

**Examples**

```
help(BASiCS_MCMC)
```

---

subset            *A 'subset' method for 'BASiCS\_Chain' objects*

---

**Description**

This can be used to extract a subset of a 'BASiCS\_Chain' object. The subset can contain specific genes, cells or MCMC iterations

**Usage**

```
## S4 method for signature 'BASiCS_Chain'
subset(x, Genes = NULL, Cells = NULL, Iterations = NULL)
```

**Arguments**

x                A [BASiCS\\_Chain](#) object.

Genes, Cells    A vector of characters, logical values, or numbers, indicating which cells or genes will be extracted.

Iterations      Numeric vector of positive integers indicating which MCMC iterations will be extracted. The maximum value in Iterations must be less or equal than the total number of iterations contained in the original [BASiCS\\_Chain](#) object.

**Value**

An object of class [BASiCS\\_Chain](#).

**Author(s)**

Catalina A. Vallejos <cnvallej@uc.cl>

## Examples

```
data(ChainSC)

# Extracts 3 first genes
ChainSC1 <- subset(ChainSC, Genes = rownames(ChainSC)[1:3])
# Extracts 3 first cells
ChainSC2 <- subset(ChainSC, Cells = colnames(ChainSC)[1:3])
# Extracts 10 first iterations
ChainSC3 <- subset(ChainSC, Iterations = 1:10)
```

---

Summary

*'Summary' method for BASiCS\_Chain objects*

---

## Description

For each of the BASiCS parameters (see Vallejos et al 2015), Summary returns the corresponding posterior medians and limits of the high posterior density interval (probability equal to prob)

## Usage

```
## S4 method for signature 'BASiCS_Chain'
Summary(x, ..., prob = 0.95, na.rm = FALSE)
```

## Arguments

x	A <a href="#">BASiCS_Chain</a> object.
...	Unused, only included for consistency with the generic.
prob	prob argument for <a href="#">HPDinterval</a> function.
na.rm	Unused, only included for consistency with the generic.

## Value

An object of class [BASiCS\\_Summary](#).

## Author(s)

Catalina A. Vallejos <cnvallej@uc.cl>

Nils Eling <eling@ebi.ac.uk>

## Examples

```
data(ChainSC)
SummarySC <- Summary(ChainSC)
```

---

```
[,BASiCS_ResultsDE,ANY,ANY,ANY-method
```

*Methods for subsetting [BASiCS\\_Result](#) and [BASiCS\\_ResultsDE](#) objects.*

---

## Description

Methods for subsetting [BASiCS\\_Result](#) and [BASiCS\\_ResultsDE](#) objects.

## Usage

```
## S4 method for signature 'BASiCS_ResultsDE,ANY,ANY,ANY'
x[i, j, drop = FALSE]
```

```
## S4 method for signature 'BASiCS_ResultsDE,ANY,ANY'
x[[i]]
```

```
## S4 method for signature 'BASiCS_Result,ANY,ANY,ANY'
x[i, j, drop = FALSE]
```

## Arguments

x	Object being subsetted.
i	See ?`[, ?`[[`
j, drop	Ignored.

## Value

An object of the same class as x.



# Index

- \* **datasets**
  - ChainRNA, [47](#)
  - ChainRNAREg, [48](#)
  - ChainSC, [48](#)
  - ChainSCReg, [49](#)
  - .generateSubsets, [4](#)
  - [,BASiCS\_Result,ANY,ANY,ANY-method
    - ([,BASiCS\_ResultsDE,ANY,ANY,ANY-method), [64](#)
  - [,BASiCS\_ResultsDE,ANY,ANY,ANY-method,
    - [64](#)
  - [[,BASiCS\_ResultsDE,ANY,ANY-method
    - ([,BASiCS\_ResultsDE,ANY,ANY,ANY-method), [64](#)
- as.data.frame,BASiCS\_ResultDE-method
  - (as.data.frame,BASiCS\_ResultsDE-method), [5](#)
- as.data.frame,BASiCS\_ResultsDE-method,
  - [5](#)
- as.data.frame,BASiCS\_ResultVG-method
  - (as.data.frame,BASiCS\_ResultsDE-method), [5](#)
  
- BASiCS-defunct, [6](#)
- BASiCS\_CalculateERCC, [6](#)
- BASiCS\_Chain, [7, 8, 10–15, 17–20, 22, 23, 25, 30, 35–37, 39, 42, 44–51, 54, 55, 57, 62, 63](#)
- BASiCS\_Chain-class (BASiCS\_Chain), [7](#)
- BASiCS\_Chain-methods, [8](#)
- BASiCS\_CorrectOffset, [9](#)
- BASiCS\_D\_TestDE (BASiCS-defunct), [6](#)
- BASiCS\_DenoisedCounts, [10](#)
- BASiCS\_DenoisedRates, [11](#)
- BASiCS\_DetectHVG (BASiCS\_DetectVG), [12](#)
- BASiCS\_DetectHVG\_LVG (BASiCS\_DetectVG), [12](#)
- BASiCS\_DetectHVG\_LVG (BASiCS\_DetectVG), [12](#)
  
- BASiCS\_DetectLVG (BASiCS\_DetectVG), [12](#)
- BASiCS\_DetectVG, [12](#)
- BASiCS\_DiagHist, [14](#)
- BASiCS\_diagHist (BASiCS\_DiagHist), [14](#)
- BASiCS\_DiagPlot, [16](#)
- BASiCS\_diagPlot (BASiCS\_DiagPlot), [16](#)
- BASiCS\_DivideAndConquer, [17](#)
- BASiCS\_Draw, [19](#)
- BASiCS\_EffectiveSize, [20](#)
- BASiCS\_effectiveSize
  - (BASiCS\_EffectiveSize), [20](#)
- BASiCS\_Filter, [21](#)
- BASiCS\_LoadChain, [22](#)
- BASiCS\_MCMC, [7, 18, 22, 23, 23, 34](#)
- BASiCS\_MockSCE, [27](#)
- BASiCS\_PlotDE, [28](#)
- BASiCS\_PlotDE,BASiCS\_ResultDE-method
  - (BASiCS\_PlotDE), [28](#)
- BASiCS\_PlotDE,BASiCS\_ResultsDE-method
  - (BASiCS\_PlotDE), [28](#)
- BASiCS\_PlotDE,missing-method
  - (BASiCS\_PlotDE), [28](#)
- BASiCS\_PlotOffset, [30, 42](#)
- BASiCS\_PlotVarianceDecomp, [31](#)
- BASiCS\_PlotVG, [32](#)
- BASiCS\_PriorParam, [24, 32](#)
- BASiCS\_Result, [34, 52, 64](#)
- BASiCS\_Result-class (BASiCS\_Result), [34](#)
- BASiCS\_ResultDE, [5, 29, 34, 35, 35, 60](#)
- BASiCS\_ResultDE-class
  - (BASiCS\_ResultDE), [35](#)
- BASiCS\_ResultsDE, [5, 29, 35, 43, 52, 59–62, 64](#)
- BASiCS\_ResultsDE-class
  - (BASiCS\_ResultsDE), [35](#)
- BASiCS\_ResultVG, [5, 13, 32, 34, 36, 59–61](#)
- BASiCS\_ResultVG-class
  - (BASiCS\_ResultVG), [36](#)
- BASiCS\_ShowFit, [36](#)

- BASiCS\_showFit (BASiCS\_ShowFit), 36
- BASiCS\_Sim, 38
- BASiCS\_Summary, 15, 16, 39, 40, 51, 52, 58, 63
- BASiCS\_Summary-class (BASiCS\_Summary), 39
- BASiCS\_Summary-methods, 40
- BASiCS\_TestDE, 6, 41
- BASiCS\_VarianceDecomp, 31, 44
- BASiCS\_VarThresholdSearchHVG, 46
- BASiCS\_VarThresholdSearchHVG\_LVG (BASiCS\_VarThresholdSearchHVG), 46
- BASiCS\_VarThresholdSearchLVG (BASiCS\_VarThresholdSearchHVG), 46
- BASiCS\_VarThresholdSearchVG (BASiCS\_VarThresholdSearchHVG), 46
- BiocParallelParam, 18, 25
  
- ChainRNA, 47
- ChainRNAREg, 48
- ChainSC, 48
- ChainSCReg, 49
  
- data.frame, 45
- DataFrame, 36
- DFrame, 60
- dim, 49
- dim, BASiCS\_Chain-method (dim), 49
- dimnames, 50
- dimnames, BASiCS\_Chain-method (dimnames), 50
- displayChainBASiCS (displayChainBASiCS-BASiCS\_Chain-method), 50
- displayChainBASiCS, BASiCS\_Chain-method (displayChainBASiCS-BASiCS\_Chain-method), 50
- displayChainBASiCS-BASiCS\_Chain-method, 50
- displaySummaryBASiCS (displaySummaryBASiCS-BASiCS\_Summary-method), 51
- displaySummaryBASiCS, BASiCS\_Summary-method (displaySummaryBASiCS-BASiCS\_Summary-method), 51
- displaySummaryBASiCS-BASiCS\_Summary-method, 51
  
- effectiveSize, 14, 16
- format, 52
- format, BASiCS\_ResultDE-method (format, BASiCS\_ResultsDE-method), 52
- format, BASiCS\_ResultsDE-method, 52
- format, BASiCS\_ResultVG-method (format, BASiCS\_ResultsDE-method), 52
  
- geom\_hex, 16
- geweke.diag, 14
  
- HPDinterval, 63
  
- makeExampleBASiCS\_Data, 53
  
- newBASiCS\_Chain, 54
- newBASiCS\_Data, 56
  
- par, 13, 37, 59
- plot, BASiCS\_Chain, ANY-method (plot-BASiCS\_Chain-method), 57
- plot, BASiCS\_Chain-method (plot-BASiCS\_Chain-method), 57
- plot, BASiCS\_Summary, ANY-method (plot-BASiCS\_Summary-method), 58
- plot, BASiCS\_Summary-method, (plot-BASiCS\_Summary-method), 58
- plot-BASiCS\_Chain-method, 57
- plot-BASiCS\_Summary-method, 58
- plot\_grid, 29
- rhat, 15, 16
- rowData, BASiCS\_ResultsDE-method, 59
- rowData, BASiCS\_ResultVG-method (rowData, BASiCS\_ResultsDE-method), 59
- rowData<-, BASiCS\_ResultsDE-method (rowData, BASiCS\_ResultsDE-method),
- rowData<-, BASiCS\_ResultVG-method (rowData, BASiCS\_ResultsDE-method),
- scale\_fill\_brewer, 45
- show, 60–62

show,BASiCS\_Chain-method  
    (BASiCS\_Chain-methods), 8  
show,BASiCS\_ResultDE-method, 60  
show,BASiCS\_ResultsDE-method, 61  
show,BASiCS\_ResultVG-method, 61  
show,BASiCS\_Summary-method  
    (BASiCS\_Summary-methods), 40  
SingleCellExperiment, 10, 11, 19, 21, 24,  
    27, 28, 33, 38, 53, 56, 57  
subset, 62  
subset,BASiCS\_Chain-method (subset), 62  
Summary, 63  
Summary,BASiCS\_Chain-method (Summary),  
    63  
  
updateObject, 8  
updateObject,BASiCS\_Chain-method  
    (BASiCS\_Chain-methods), 8