

Package ‘mixhvg’

September 15, 2024

Type Package

Title Mixture of Multiple Highly Variable Feature Selection Methods

Version 1.0.1

Maintainer Ruzhang Zhao <ruzhangzhao@gmail.com>

Description Highly variable gene selection methods, including popular public available methods, and also the mixture of multiple highly variable gene selection methods, <<https://github.com/RuzhangZhao/mixhvg>>. Reference: <[doi:10.1101/2024.08.25.608519](https://doi.org/10.1101/2024.08.25.608519)>.

Imports scran, Seurat, Matrix, methods, SingleCellExperiment, scuttle

License GPL-3

Encoding UTF-8

RoxygenNote 7.2.1

NeedsCompilation no

Author Ruzhang Zhao [aut, cre] (<<https://orcid.org/0000-0002-3420-5771>>),
Hongkai Ji [aut]

Repository CRAN

Date/Publication 2024-09-15 16:30:02 UTC

Contents

FindVariableFeaturesMix	1
Index	4

FindVariableFeaturesMix
FindVariableFeaturesMix

Description

FindVariableFeaturesMix

Usage

```
FindVariableFeaturesMix(
  object,
  method.names = c("scrn", "scrn_pos", "seuratv1"),
  nfeatures = 2000,
  loess.span = 0.3,
  clip.max = "auto",
  num.bin = 20,
  binning.method = "equal_width",
  verbose = FALSE
)
```

Arguments

<code>object</code>	An object, SeuratObject and matrix(including sparse matrix) are both acceptable
<code>method.names</code>	The following methods can be directly used for highly variable feature selection. The mixture of methods take a vector of method list, e.g. <code>c("scrn", "scrn_pos", "seuratv1")</code> , which is also default. scrn: Use mean-variance curve adjustment on lognormalized count matrix, which is <code>scrn ModelGeneVar</code> . mv_ct: Use mean-variance curve adjustment on count matrix, inherited from <code>scrn ModelGeneVar</code> . mv_nc: Use mean-variance curve adjustment on normalized count matrix, inherited from <code>scrn ModelGeneVar</code> . mv_lognc: The same as <code>scrn</code> . mv_PFlogPF: Use mean-variance curve adjustment on PFlog1pPF matrix, inherited from <code>scrn ModelGeneVar</code> . scrn_pos: Use <code>scrn poisson</code> version, <code>modelGeneVarByPoisson</code> . seuratv3: Use logmean-logvariance curve adjustment on count matrix, which is <code>vst</code> , Seurat FindVariableFeatures Function(https://satijalab.org/seurat/reference/findvariablefeatures). logmv_ct: The same as <code>seuratv3</code> . logmv_nc: Use logmean-logvariance curve adjustment on normalized count matrix, inherited from <code>seuratv3(vst)</code> . logmv_lognc: Use logmean-logvariance curve adjustment on lognormalized count matrix, inherited from <code>seuratv3(vst)</code> . logmv_PFlogPF: Use logmean-logvariance curve adjustment on PFlog1pPF matrix, inherited from <code>seuratv3(vst)</code> . seuratv1: Use dispersion on lognormalized count matrix, which is dispersion (<code>disp</code>), Seurat FindVariableFeatures Function(https://satijalab.org/seurat/reference/findvariablefeatures). disp_lognc: The same as <code>seuratv1</code> . disp_PFlogPF: Use dispersion on PFlog1pPF matrix, inherited from <code>seuratv1(disp)</code> . mean_max_ct: Highly Expressed Features with respect to count matrix. mean_max_nc: Highly Expressed Features with respect to normalized count matrix.

	mean_max_lognc: Highly Expressed Features with respect to lognormalized count matrix
nfeatures	Number of features to select as top variable features.
loess.span	(Only work for logmv based methods like <code>seuratv3</code>). Loess span parameter used when fitting the variance-mean relationship
clip.max	(Only work for logmv based methods like <code>seuratv3</code>). After standardization values larger than <code>clip.max</code> will be set to <code>clip.max</code> ; default is 'auto' which sets this value to the square root of the number of cells
num.bin	(Only work for logmv or dispersion based methods)Total number of bins to use in the scaled analysis (default is 20)
binning.method	Specifies how the bins should be computed. Available methods are: equal_width: each bin is of equal width along the x-axis[default]. equal_frequency: each bin contains an equal number of features (can increase statistical power to detect overdispersed features at high expression values, at the cost of reduced resolution along the x-axis).
verbose	Whether to show progress bar for calculations. Default is FALSE.

Details

The function inherits from `FindVariableFeatures` function of Seurat Package. Refer to <https://github.com/RuzhangZhao/mixhvg> for user manual.

Value

object: If the input is `SeuratObject`, the return is also `SeuratObject`; if the input is `matrix`(including sparse matrix), the return is the highly variable feature names.

Examples

```
if(0){
  simple_matrix<-matrix(1:2e4,nrow=4000,ncol=5)
  rownames(simple_matrix)<-1:nrow(simple_matrix)
  colnames(simple_matrix)<-1:ncol(simple_matrix)
  simple_matrix_HVG<-FindVariableFeaturesMix(simple_matrix)
}
```

Index

FindVariableFeaturesMix, 1