R documentation

of 'deconvDigitalDLSorter.Rd'

September 3, 2020

`deconvDigitalDLSorter` *Deconvolute bulk gene expression samples (bulk RNA-Seq) using a pre-trained DigitalDLSorter model.*

Description

Deconvolute bulk gene expression samples (RNA-Seq) quantifying the proportion of cell types present in a bulk sample. See in Details the available models. This method uses a pre-trained Deep Neural Network model to enumerate and quantify the cell types present in bulk RNA-Seq samples. For the moment, the available models allow to deconvolute the immune infiltration breast cancer (Chung et al., 2017) at two levels: specific cell types (`'breast.chung.specific'`) and generic cell types (`'breast.chung.generic'`). See `breast.chung.generic` and `breast.chung.specific` documentation for details.

Usage

```
deconvDigitalDLSorter(
  data,
  model = "breast.generic",
  batch.size = 128,
  normalize = TRUE,
  simplify.set = NULL,
  simplify.majority = NULL,
  verbose = TRUE
)
```

Arguments

`data`	A `matrix` or a `data.frame` with bulk gene expression of samples. Rows must be genes in symbol notation and columns must be samples.
`model`	Pre-trained DNN model to use for deconvoluting process. For the moment, the available models are for RNA-Seq samples from breast cancer (`'breast.chung.generic'` and `'breast.chung.specific'`) environment.
`batch.size`	Number of samples loadad in-memory each time of deconvolution process. If unspecified, `batch.size` will default to 128.

`normalize` Normalize data before deconvolution. `TRUE` by default.

`simplify.set` List specifying which cell types should be compressed into a new label whose name will be the list name item. See examples for details.

`simplify.majority` List specifying which cell types should be compressed into the cell type with greater proportions in each sample. Unlike `simplify.set`, it allows to maintain the complexity of the results while compressing the information, because it is not created a new label.

`verbose` Show informative messages during the execution.

Details

This function is oriented for users that only want to use the method for deconvoluting their bulk RNA-Seq samples. For users that are building their own model from scRNA-seq, see `deconvDigitalDLSorterObj` The former works with base classes, while the last uses `DigitalDLSorter` objects.

For situations where there are cell types exclusive to each other because it does not make sense that they appear together, see arguments `simplify.set` and `simplify.majority`.

Value

A `data.frame` with samples (i) as rows and cell types (j) as columns. Each entry represents the predicted proportion of j cell type in i sample.

References

Chung, W., Eum, H. H., Lee, H. O., Lee, K. M., Lee, H. B., Kim, K. T., et al. (2017). Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. Nat. Commun. 8 (1), 15081. doi: 10.1038/ncomms15081.

See Also

`deconvDigitalDLSorterObj`

Examples

```
results1 <- deconvDigitalDLSorter(
  data = TCGA.breast.small,
  model = "breast.chung.specific",
  normalize = TRUE
)

## simplify arguments
simplify <- list(Tumor = c("ER+", "HER2+", "ER+/HER2+", "TNBC"),
                 Bcells = c("Bmem", "BGC"))

## in this case,  the item names from list will be the new labels
results2 <- deconvDigitalDLSorter(
  TCGA.breast.small,
  model = "breast.chung.specific",
  normalize = TRUE,
  simplify.set = simplify)

## in this case, the cell type with greatest proportion will be the new label
## the rest of proportion cell types will be added to the greatest
```

```
results3 <- deconvDigitalDLSorter(
  TCGA.breast.small,
  model = "breast.chung.specific",
  normalize = TRUE,
  simplify.majority = simplify)
```

Index