Package 'MIND'

March 31, 2019

Type Package

Title Using Multiple Measurements of Tissue to Estimate Subject- And Cell-Type-Specific Gene Expression via Deconvolution

Version 0.2.0

Author Jiebiao Wang

Maintainer Jiebiao Wang <randel.wang@gmail.com>

Description A method to glean more insights from bulk gene expression. It borrows information across multiple measurements of the same tissue per subject, such as multiple regions of the brain, using an empirical Bayes approach to estimate subject- and cell-type-specific gene expression via deconvolution.

Depends nnls, doParallel, foreach

License GPL

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

URL https://github.com/randel/MIND

BugReports https://github.com/randel/MIND/issues

R topics documented:

`est_frac` *Estimating cell type fractions using non-negative least squares (NNLS)*

Description

It calls the nnls package to estimate cell type fractions using a signature matrix and bulk expression data.

Usage

```
est_frac(sig, bulk)
```

Arguments

sig	signature matrix (marker gene x cell type).
bulk	bulk gene expression data that need to be deconvolved (gene x tissue sample).

Value

A matrix containing the estimated cell type fractions (tissue sample x cell type). Row sums have been normalized to be 1.

---

example *A data example*

Description

A data list for demonstration.

Value

A list containing

X	bulk gene expression (gene x subject x measure).
W	subject-specific cell type fraction (subject x measure x cell type).

Examples

```
data(example)
```

---

mind *The Multi-measure INdividual Deconvolution (MIND) algorithm*

Description

It calculates the empirical Bayes estimates of subject- and cell-type-specific gene expression, via a computationally efficient EM algorithm.

Usage

```
mind(X, W, maxIter = 100, tol = 0.001, verbose = F, ncore = 4)
```

Arguments

X	bulk gene expression (gene x subject x measure).
W	subject-specific cell type fraction (subject x measure x cell type).
maxIter	maximum number of iterations for the EM algorithm.
tol	tolerance level of absolute relative change of the log-likelihood to stop the EM algorithm.
verbose	logical, to print the detailed information for each iteration: iter (the iteration number), logLike_change, sigma2_e, mean(diag(Sigma_c))).
ncore	number of cores to run in parallel

Value

A list containing the output of the EM deconvolution algorithm

A	the deconvolved cell-type-specific gene expression (gene x cell type x subject).
mu	the estimated profile matrix (gene x cell type).
iter	the number of iterations used in the EM algorithm.
Sigma_c	the covariance matrix for the deconvolved cell-type-specific expression (cell type x cell type).
sigma2_e	the error variance.
loglike	the log-likelihood for each EM iteration.
var_A	the posterior covariance matrix for A (vectorized covariance matrix by subject).

References

Jiebiao Wang, Bernie Devlin, Kathryn Roeder. Using multiple measurements of tissue to estimate subject- and cell-type-specific gene expression. Submitted.

Examples

```
data(example)

deconv = mind(X = example$X, W = example$W, ncore = 2)
```

Index