Package 'dyebiasexamples'

January 23, 2025

Version 1.46.0

Date 2 March 2016

Title Example data for the dyebias package, which implements the GASSCO method.

Author Philip Lijnzaad and Thanasis Margaritis

Description Data for the dyebias package, consisting of 4 self-self hybrizations of self-spotted yeast slides, as well as data from Array Express accession E-MTAB-32

Maintainer Philip Lijnzaad <plijnzaad@gmail.com>

License GPL-3

Depends R (>= 1.4.1), marray, GEOquery

Suggests dyebias, convert, Biobase

URL http://www.holstegelab.nl/publications/margaritis_lijnzaad

biocViews ExperimentData, SAGEData, CGHData, MicroarrayData, TwoChannelData, ArrayExpress

git_url https://git.bioconductor.org/packages/dyebiasexamples

git_branch RELEASE_3_20

git_last_commit dd94026

git_last_commit_date 2024-10-29

Repository Bioconductor 3.20

Date/Publication 2025-01-23

Contents

data.raw 2
dyebias.geo2marray 3
dyebias.umcu.proper.estimators 4

Index **6**

`data.raw` *Example data for the dyebias package*

Description

The `dyebias`-package, described in Margaritis et al. (2009) can be used to get rid of dye bias in two-colour microarrays. The `data.raw` and `data.norm` objects are used in its examples.

The objects represent four hybridizations of identical mRNA, with increasing Cy3 and Cy5 labeling percentages (identical per slide) and differently spiked-in external controls to judge the process of dyebias correction.

Usage

```
data(data.raw)
data(data.norm)
```

Format

The data uses the `marray`-package by Dudoit and Yang (2002). `data.raw` is a `marrayRaw` object, `data.norm` is a `marrayNorm` object derived from it by print-tip LOESS normalization. Neither is dyebias-corrected yet.

Details

The column `R.group` of `maInfo(maTargets(data.norm))` shows the details. Eg., `4%_2EC` indicates that the labeling (of both channels) was at 4%, and the external controls were spiked in at a concentration twice that of the green channel. See Margaritis et~al. (2009) for details.

Note

The Tuteja data is also included in this package under the (inst)/doc directory, as this data is not proper rda, tab or csv data. For details, refer to the original publication and/or the `dyebias` vignette.

Author(s)

Philip Lijnzaad

Source

All accession numbers below refer to ArrayExpress (http://www.ebi.ac.uk/microarray).

This two-colour microarrray data was obtained from identical mRNA extracts (protocol P-UMCU-37), spiked with external controls, dUTP-labeled with Cy3 and Cy5 (protocol P-UMCU-38). This was hybridized (protocol P-UMCU-39) onto self-spotted slides containing 70-mer oligonucleotides (2 replicates per oligo, Operon "Array-Ready", and including 2838 control features; protocol P-UMCU-34). Scanning was done with an Agilent G2565AA scanner (protocol P-UMCU-40) and images were quantified with BioDiscovery's ImaGene 7.x (protocol P-UMCU-42)

References

Margaritis, T., Lijnzaad, P., van~Leenen, D., Bouwmeester, D., Kemmeren, P., van~Hooff, S.R and Holstege, F.C.P. (2009). Adaptable gene-specific dye bias correction for two-channel DNA microarrays. *Molecular Systems Biology, submitted*

Dudoit, S. and Yang, Y.H. (2002) Bioconductor R packages for exploratory analysis and normalization of cDNA microarray data. In: Parmigiani, G., Garrett, E.S. , Irizarry, R.A., and Zeger, S.L. (eds.) *The Analysis of Gene Expression Data: Methods and Software*, New~York: Springer

Examples

```
data(data.raw)
data(data.norm)
```

`dyebias.geo2marray` *convenience function to convert GEO objects to marray objects*

Description

convenience function to convert GEO objects to marray objects

Arguments

`gse`	GSE data set
`slide.ids`	Return only the slides with these ids. If NULL, return all.
`type`	what to extract; must be either "norm" or "raw".
`gene.selector`	function(table) acting on Table(GPL) giving back an index with the rows considered to be genes.
`reporterid.name`	column containing the reporter.id, in Table(gpl).
`cy3.name`	The column name containing the factor value for the Cy3 (green) channel
`cy5.name`	The column name containing the factor value for the Cy5 (red) channel
`R.name`	column name for extracting the R data from Table(gsm)
`G.name`	column name for extracting the G data from Table(gsm)
`M.name`	column name for extracting the M data from Table(gsm)
`A.name`	column name for extracting the A data from Table(gsm)
`Rf.name`	column name for extracting the Rf data from Table(gsm)
`Gf.name`	column name for extracting the Gf data from Table(gsm)
`Rb.name`	column name for extracting the Rb data from Table(gsm)
`Gb.name`	column name for extracting the Gb data from Table(gsm)

Details

The `XYZ.name` mechanism is the same as that used in `read.marrayRaw`; i.e. you specify the name of the column that contains the desired data.

Value

A full-fledged marrayRaw (if `type` was "raw") or marrayNorm (if `type` was "norm") is returned.

Note

At some point, this functionality should be merged into the convert package.

Author(s)

Philip Lijnzaad

References

Davis, S. and Meltzer, P.S (2007). GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* 23, 1846–1847 (doi:10.1093/bioinformatics/btm254).

Dudoit, S. and Yang, Y.H. (2002) Bioconductor R packages for exploratory analysis and normalization of cDNA microarray data. In: Parmigiani, G., Garrett, E.S. , Irizarry, R.A., and Zeger, S.L. (eds.) *The Analysis of Gene Expression Data: Methods and Software*, New~York: Springer

Chen,S., de~Vries, M.A. and Bell, S.P. (2007) *Genes Dev.* 21, 2897–2907 "Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading" (doi:10.1101/gad.1596807)

Examples

```
## Not run:
  ## Running this example takes too much time; if you want that, see the
  ## second example in the vignette

## End(Not run)
```

`dyebias.umcu.proper.estimators`

Determine which spots should not be ruled out as slide bias estimators

---

Description

Some spots (reporters/probes) should not be used when estimating the slide bias. Typical examples are mitochondrial genes and spots known to cross-hybridize. This function finds the ones that are OK to use.

Arguments

`reporter.info`	A data.frame, one row per spot, with (at least) columns `reporterId` (e.g. gene id or oligo id) and any of the following characteristics: `reporterGroup`, `chromosomeName`, `bioSeqType`, `crosshybRank` and `reporterSequence`. They are used to get rid of reporters that are not suitable when estimating the slide bias.
`verbose`	Logical speficying whether to be verbose or not

Details

This function is particular to the slides and database set-up at the Holstege lab, but might serve as inspiration.

Value

Returns and index vector that can be used as the `estimator.subset`-argument to `dyebias.application.subset`.

Author(s)

Philip Lijnzaad <p.lijnzaad@umcutrecht.nl>

References

Margaritis, T., Lijnzaad, P., van~Leenen, D., Bouwmeester, D., Kemmeren, P., van~Hooff, S.R and Holstege, F.C.P. (2009) Adaptable gene-specific dye bias correction for two-channel DNA microarrays. *Molecular Systems Biology, submitted*

See Also

`dyebias.apply.correction`

Examples

```
### choose the estimators and which spots to correct:
estimator.subset <- dyebias.umcu.proper.estimators(maInfo(maGnames(data.norm)))

summary(estimator.subset)

### do the correction
## Not run:
   correction <- dyebias.apply.correction(data.norm=data.norm,
                                          iGSDBs = iGSDBs.estimated,
                                          estimator.subset=estimator.subset,
                                          application.subset = TRUE,
                                          verbose=TRUE)

## End(Not run)
```

Index

∗ **datasets**
data.raw, 2
∗ **misc**
dyebias.geo2marray, 3
dyebias.umcu.proper.estimators, 4

data.norm *(*data.raw*)*, 2
data.raw, 2
dyebias.apply.correction, *5*
dyebias.geo2marray, 3
dyebias.umcu.proper.estimators, 4

read.marrayRaw, *3*