Package 'normalize450K'

December 29, 2015

Type Package

Title Preprocessing of Illumina Infinium 450K data

Version 0.99.3

Date 2015-08-31

Author Jonathan Alexander Heiss

Maintainer Jonathan Alexander Heiss <jonathan.heiss@posteo.de>

Description Precise measurements are important for epigenome-
wide studies investigating DNA methylation in whole blood samples, where effect sizes are
expected to be small in magnitude. The 450K platform is often affected by batch ef-
fects and proper preprocessing is recommended. This
package provides functions to read and normalize 450K '.idat' files. The normalization cor-
rects for dye bias and biases related to signal
intensity and methylation of probes using local regression. No adjust-
ment for probe type bias is performed to avoid the trade-off of
precision for accuracy of beta-values.

Depends R (>= 3.3), Biobase, illuminaio

Suggests minfiData, data.table

License BSD_2_clause + file LICENSE

biocViews Normalization, DNAMethylation, Microarray, TwoChannel,
Preprocessing, MethylationArray

NeedsCompilation no

R topics documented:

`read_and_normalize450K` *Normalization of 450K data by LOESS method*

Description

Read 450K '.idat' files and compute raw or normalized beta-values.

Usage

```
read450K(idat_files)
normalize450K(intensities)
dont_normalize450K(intensities)
```

Arguments

`idat_files`	a character vector containing the paths to the .idat files stripped from the '_Grn.idat' suffix with one entry for each sample (.idat files for green and red intensities have to be in the same folder).
`intensities`	List object containing raw signal intensities. Result of calling `read450K`.

Details

Function `read450K` reads .idat files and returns a list object containing raw signal intensities. `dont_normalize450K` returns an ExpressionSet containing beta-values without normalization. `normalize450K` performs dye bias correction using the extension controls probes followed by normalization by local regression (Heiss and Brenner, 2015) and returns an ExpressionSet containing beta-values, too.

Value

For `read450K` a list containing the methylated, unmethylated and control signal intensities. For `dont_normalize450K` and `normalize450K` an ExpressionSet containing beta-values, rows corresponding to CpG sites (named) and columns to samples (in the same order as 'idat_files').

Note

A benchmark comparing the performance of this method with other normalization approaches is provided in the vignette.

Author(s)

Jonathan A. Heiss

References

Heiss, Jonathan A., and Hermann Brenner. *Between-array normalization for 450K data*. Frontiers in genetics 6 (2015). doi:10.3389/fgene.2015.00092

Examples

```
library(minfiData) ## this package includes some .idat files
library(data.table)

path <- system.file("extdata",package="minfiData")
samples = fread(file.path(path, 'SampleSheet.csv'),integer64='character')

samples[,file:=file.path(path,Sentrix_ID,paste0(Sentrix_ID,'_',Sentrix_Position))]
## samples$file is a character vector containing the location of the
## .idat files, but without the suffixes "_Red.idat" or "_Grn.idat"

raw = read450K(samples$file)
none = dont_normalize450K(raw) ## no normalization
norm = normalize450K(raw)
```

Index