MAQCsubset

March 23, 2012

`MAQCsubset` *Experimental Data Package: MAQCsubset*

Description

selected data from the MAQC project (Nature Biotechnology, Sept. 2006)

Usage

```
data(afxsubRMAES)
data(afxsubRMA)
data(afxsub)
```

Format

The format is: An `ExpressionSetObject` with covariates:

- `site`: from cel
- `samp`: rna src/mixture code
- `repl`: replicate

Note

`afxsubRMA` is an `exprSet` (deprecated) and `afxsub` is an `AffyBatch`. `afxsubRMAES` is a proper ExpressionSet instance.

ilmMAQCsubR is the result of applying lumiR to the files in the vicinity of GEO GSM122901 with filename suffixes matching those of the sampleNames in the set.

Examples

```
data(afxsubRMAES)
```

`gehMAQCsubDef` *Excerpt from GE Codelink array contributions to MAQC*

Description

Excerpt from GE Codelink contributions to MAQC

Usage

```
data(gehSubRaw)
data(gehMAQCsubDef)
```

Details

`gehSubRaw` is a `codelink::Codelink` instance based on reading the raw GEO files: "GSM123122_GEH_1_A1. "GSM123123_GEH_1_A2.TXT" "GSM123127_GEH_1_B1.TXT" "GSM123128_GEH_1_B2.TXT" "GSM123132_GEH_1_C1.TXT" "GSM123133_GEH_1_C2.TXT" "GSM123137_GEH_1_D1.TXT" "GSM123138_GEH_1_D2.TXT"

`gehMAQCsubDef` is an `ExpressionSet` instance based on default background correction and normalization of the `codelink` package. The original feature names include duplicates; these were made unique by `make.names` with `unique=TRUE`.

Author(s)

Vince Carey <stvjc@channing.harvard.edu>

Examples

```
data(gehMAQCsubDef)
gehMAQCsubDef
```

`gnfCerebHi` *Gene lists for hi or low abundance in cerebellum according to Novartis GNF symatlas*

Description

Data frames with gene lists for hi or low abundance in cerebellum according to Novartis GNF symatlas

Usage

```
data(gnfCerebHi)
data(gnfCerebLow)
```

Details

The symatlas.gnf.org database was searched using the gcrma version of human gene atlas for genes having expression in cerebellum at least 3 times (or at most 1/3 times) median expression over all organs surveyed. The resulting gene lists were intersected with genes present on GE codelink (version used in MAQCsubset) and hgu95a.

Value

data.frame instances with columns providing gene name, affy probe set identifier, codelink probe identifier, illuminaHumanv1 identifier.

Author(s)

Vince Carey <stvjc@channing.harvard.edu>

Examples

```
data(gnfCerebHi)
gnfCerebHi[1:3,]
```

`proboStruct-class` *Class "proboStruct"*

Description

structure for managing proboscis plot data

Objects from the Class

Objects can be created by calls of the form `new("proboStruct", ...)`.

Slots

`.Data`: Object of class `"list"` ~~

`call`: Object of class `"call"` ~~

Extends

Class `"list"`, from data part. Class `"vector"`, by class "list", distance 2. Class `AssayData-class`, by class "list", distance 2.

Methods

plot

Note

The proboscis plot shows how the probability of self-consistent monotone titration (SCMT) varies with the spiked difference in concentrations of two mRNA preparations in an MAQC dataset.

Author(s)

V Carey <stvjc@channing.harvard.edu>

References

For Figure 2 of Shippy et al., Using RNA sample titrations... (Nat Biotech, 24(9):1123-1131, Sep 2006)

Examples

```
data(afxsubRMAES)
NN1 = proboscis(afxsubRMAES)
plot(NN1)
showClass("proboStruct")
```

`proboscis` *Produce a plot similar to Figure 2 of the Shippy MAQC paper (PMID 16964226).*

Description

Produce a plot similar to Figure 2 of the Shippy MAQC paper (PMID 16964226).

Usage

```
proboscis(es, site=1, ABp=0.001, CDp=0.01, mmrad=100)
```

Arguments

`es`	`ExpressionSet` instance with MAQC assay results
`site`	numeric code – site to be assessed
`ABp`	ABp – p-value threshold to declare concentration of gene in sample A to be different from ehe concentration in sample B
`CDp`	CDp – p-value threshold to declare concentration of gene in sample C to be different from the concentration in sample D
`mmrad`	numeric radius of the moving mean used to smooth the proportions differentially expressed

Details

Figure 2 of the Shippy paper consists of a collection of plots of estimated probabilities of self-consistent monotone titration – briefly, samples are such that A has 100% USRNA, B has 100% Ambion brain, C has 75% USRNA+25% brain, D has 25% USRNA, 75% brain. Self-consistent monotone titration holds for gene g if microarray measures for that gene satisfy A > C > D > B or B > C > D > A. The estimated probability functions look like a creature sticking its nose over a wall, thus the name of this function.

Value

an instance of `proboStruct`, for which a plot and lines method are available.

Author(s)

Vince Carey <stvjc@channing.harvard.edu>

References

PMID 16964226

Examples

```
data(afxsubRMAES)
NN2 = proboscis(afxsubRMAES, site=2)
plot(NN2)
```

Index