oneChannelGUI

April 20, 2011

`annotateNGSeset` *Associate annotation to NGS data loaded in oneChannelGUI*

Description

This function associates annotation to NGS data loaded in oneChannelGUI, using the information retrieved by genome annotation from ENSEMBL by ChIPpeakAnn package

Usage

```
annotateNGSeset()
```

Author(s)

Raffaele A Calogero

`AptMidas` *Graphical interface to APT midas*

Description

This is a graphical interface to the midas program for detection of alternative splicing detection implemented in the Affymetrix APT tools

Usage

```
AptMidas()
```

Note

For more information see Affymetrix Alternative Transcript Analysis Methods for Exon Arrays whitepaper. Before using MiDAS is strongly recomanded to filter out gene level probe sets with low intensity values to avoid searching alternative splicing for probe sets which are not expressed. This can be done using filtering method implemented in oneChannelGUI which define a background intensity threshold on the basis of the intron exon signals of a set of housekeeping genes present in the exon arrays. However it is also possible to use a filter based on the dabg p-value calculated using Affymetrix APT tools. This function will also calculate Splice Index

Author(s)

Raffaele A Calogero

See Also

erankProdAltSpl

`bayseqInterface` *graphical interface to baySeq library.*

Description

BaySeq provides statistical routines for determining differential expression in digital gene expression data. The routines can be applied equally to two groups experiments only.

Usage

```
bayseqInterface()
```

Details

This interfase reorganizes NGS data loaded in oneChannelGUI to be analysed by baySeq

Author(s)

Raffaele A Calogero

`biomartFilter` *Filtering only gene-level probe sets with multiple ensembl transcripts*

Description

This function allows to filter exon array data to selected only those gene-level probe sets associated to multiple mRNAs annotated in ensembl data base

Usage

```
biomartFilter()
```

Author(s)

Raffaele A Calogero

bofa — *Bovine microRNA data set*

Description

microRNA data set fasta produced using miRbase B taurus precursors reformated to have only miR name as descriptor of the fasta file

Usage

```
bofa
```

Format

Fasta file

buildingLocalAnnotation — *Updates local gene-level annotation data for gene and exon arrays using the netaffx database*

Description

Internal oneChannelGUI Gene-level annotation data can be upgraded using this function, which queries netaffx database. The exon-level annotations are upadated using UCSC golden path data. annotation RDA files are saved in the data subdir of oneChannelGUI dir. Windows users need to exchange the older copies present in Rdata.zip, simply dragging them in the zip file.

Usage

```
buildingLocalAnnotation()
```

Author(s)

Raffaele A Calogero

chrLength *Chromosomes lenghts*

Description

A list containing chromosome lengths for hg19, mm9 and rn4 genome releases

Usage

```
chrLength
```

Format

A list with three observations: chromosome lengths for hg19, mm9 and rn4 genome releases

Author(s)

Raffaele A Calogero

colExtract *Extract a column from a tab delimited file with header*

Description

This function allows to extract a specific column from a tab delimited file generated by oneChannelGUI. The file should contain an header. This function is useful to extract probe set ids to be used fro ven diagram representations

Usage

```
colExtract()
```

Author(s)

Raffaele A Calogero

`combineGeoMSF` *This function allows to combine GEO Matrix Series Files belonging to the same experiment.*

Description

The function combines in a unique ExpressionSet the data derived from multiple Matrix Series Files belongig to a GEO experiment contaning more than 255 arrays.

Usage

```
combineGeoMSF()
```

Note

see oneChannelGUI vignette for more info

Author(s)

Raffaele A Calogero

`combining2eSet` *This function allows to combine two matrices extracted from oneChannelGUI.*

Description

The function combines two matrices in one using column binding. This function is useful to combine two matrices generated for meV visualization.

Usage

```
combining2eSet()
```

Note

see oneChannelGUI vignette for more info

Author(s)

Raffaele A Calogero

`consistentFilters` *This function allows filtering using the combination of multiple paramenters, e.g. MiDAS p-values and Rank Product p-values*

Description

This filter can be used to moderate multiple tests errors. E.g. finding the intersection between MiDAS p-values and Rank Product p-values user will remove some of the false positive produced by the two methods. A filter on the size of delta Splice Index associated to MiDAS p-values filter will will allow to remove statistical significant splicing events which are characterized by a very limited variation.

Usage

```
consistentFilters()
```

Note

This fuction needs the presence of Splice Index data, MiDAS p-values and RP p-values. It works for two groups only

Author(s)

Raffaele A Calogero

See Also

erankProdAltSpl, AptMidas

`cosieHscore` *Parameters precalculated for the Affymetrix core human annotation subgroup from FMI Bioinformatics group and needed to correct SI*

Description

The description of how these data were generated is described in Gaidatzis et al. Nucleic Acids Research, 2009, 1

Usage

```
cosieHscore
```

Format

A matrix with 7 observations: status, geneTrainFrom, geneTrainTo, a, b, c, d

References

www.fmi.ch/groups/gbioinfo/cosie/cosie.html

cosieHsfull — *Parameters precalculated for the Affymetrix full human annotation subgroup from FMI Bioinformatics group and needed to correct SI*

Description

The description of how these data were generated is described in Gaidatzis et al. Nucleic Acids Research, 2009, 1. Authors recommend the use of the core set since the full set is not as well characterized.

Usage

```
cosieHsfull
```

Format

A matrix with 7 observations: status, geneTrainFrom, geneTrainTo, a, b, c, d

References

www.fmi.ch/groups/gbioinfo/cosie/cosie.html

cosieMmcore — *Parameters precalculated for the Affymetrix core mouse annotation subgroup from FMI Bioinformatics group and needed to correct SI*

Description

The description of how these data were generated is described in Gaidatzis et al. Nucleic Acids Research, 2009, 1

Usage

```
cosieMmcore
```

Format

A matrix with 7 observations: status, geneTrainFrom, geneTrainTo, a, b, c, d

References

www.fmi.ch/groups/gbioinfo/cosie/cosie.html

`cosieMmfull` *Parameters precalculated for the Affymetrix full mouse annotation subgroup from FMI Bioinformatics group and needed to correct SI*

Description

The description of how these data were generated is described in Gaidatzis et al. Nucleic Acids Research, 2009, 1. Authors recommend the use of the core set since the full set is not as well characterized.

Usage

```
cosieMmfull
```

Format

A matrix with 7 observations: status, geneTrainFrom, geneTrainTo, a, b, c, d

References

www.fmi.ch/groups/gbioinfo/cosie/cosie.html

`cosieWrapper` *A wrapper to cosie, Corrected Splicing Indices for Exon arrays, that for any given set of new exon array experiments corrects for the observed bias and improves the detection of alternative splicing*

Description

COSIE is a function that for a given set of exon arrays corrects for the observed bias and improves the detection of alternative splicing. It adjusts splicing indices for exons, especially for those that belong to differentially expressed genes. For this adjustment, COSIE uses parameters that are specific for each probeset which were trained from a large number of published exon arrays. The downside of this approach is that such parameters cannot be estimated for all probesets on the microarray. Based on our training set, COSIE corrects 95.1 percent of the probesets. Separate parameter files are provided for both the full and core sets, including all probesets that are linked to transcripts. We recommend the use of the core set that was also used in the cited study below. The full set is not as well characterized.

Usage

```
cosieWrapper()
```

Author(s)

Gaidatzis et al. Nucleic Acids Research, 2009, 1

See Also

cosieHscore, cosieMmcore, cosieHsfull, cosieMmfull

`createGeoTarget` *Creating a affylmGUI Target starting from a GEO Matrix series file*

Description

The function extracts from GEO series matrix file all the information to create a Target file, that can be used to load the GEO series matrix file into oneChannelGUI.

Usage

```
createGeoTarget()
```

Note

see oneChannelGUI vignette for Target file description

Author(s)

Raffaele A Calogero

`crosshybFilter` *Removing from exon array gene/exon level probe sets characterized by cross hybridization with other transcripts*

Description

XHYB field is mainly an indicator of weak assignment between a transcript cluster and the assigned mRNA, suggesting a potential crosshyb, CRSSHYB is a measure of the promiscuity of the probes within a probe set among transcribed sequences.

1. 1 unique. All probes in the probe set perfectly match only one sequence in the putatively transcribed array design content. The vast majority of probe sets are unique.
2. 2 similar. All the probes in the probe set perfectly match more than one sequence in the putatively transcribed array design content.
3. 3 mixed. The probes in the probe set either perfectly match or partially match more than one sequence in the putatively transcribed array design content.

XHYB and CRSSHYB are used to remove probe sets characterized by multiple hybridization of exon probes

Usage

```
crosshybFilter()
```

Author(s)

Raffaele A Calogero

`ctrtrtHeatmap` *Creating heat map for maSigPro data with only one treatment condition*

Description

This function needs a maSigPro target econmpassing only one treatment condition. This function is ideal to visualize the expression data for putative mRNA targets upon miRNA expression perturbation with antagomir or mimics. Expression data reorganized as the average expression differences between treatment and untreated condition over the time course.

Usage

```
ctrtrtHeatmap()
```

Author(s)

Raffaele A Calogero

`dfMAplot` *MA and Volcano plots from data present in a limma derived topTable*

Description

MA and Volcano plots can be generated starting from limma results summarized in a topTable. Specific subsets of the topTable defined by p-value below an user-defined threshold and/or log2 fold changes over an user-defined threshold can be saved. The subset of data can be saved as a tab delimited file

Usage

```
dfMAplot(table1)
```

Arguments

`table1`	topTable data.frame generate by affylmGUI

Note

To know more about topTable see limma help

Author(s)

Raffaele A Calogero

edgerInterface — *graphical interface to edgeR library.*

Description

edgeR provides statistical routines for determining differential expression in digital gene expression data. The routines can be applied equally to SAGE, CAGE, Illumina, Solexa, 454 or ABI SOLiD experiments.

Usage

```
edgerInterface()
```

Details

This interfase reorganizes NGS data loaded in oneChannelGUI to be analysed by edgeR

Author(s)

Raffaele A Calogero

edgeRNorm — *This function scale-normalize NGS raw data as described by Robinson and Oshlack Genome Biology 2010, 11:R25*

Description

This function is scale-normalize NGS data using the normalization function provided in edgeR.

Usage

```
edgeRNorm()
```

Author(s)

Raffaele A Calogero

`EG2probeset` *This function allows to link GeneBank and Entrez Gene ids to gene-level probe set ids*

Description

This function allows to link oneChannelGUI embedded Affymetrix annotated accession numbers to gene-level probe set ids. Usnig the ACC EG are linked using the Bioconductor human, mouse or rat LLMappings annotation library

Usage

```
EG2probeset()
```

Author(s)

Raffaele A Calogero

`erankProdAltSplFilter` *Filtering Rank Product results for the detection of alternative splicing events*

Description

This is a graphical interface to filter data on the basis of p-value generated by rank product analysis applied for the detection of alternative splicing

Usage

```
erankProdAltSpl()
```

Author(s)

Raffaele A Calogero

See Also

erankProdAltSpl, AptMidas

`erankProdAltSpl` *Implementation of the Rank Product method for the detection of alternative splicing events*

Description

This is a graphical interface to the RP function from RankProd package applied to detection of alternative splicing

Usage

```
erankProdAltSpl()
```

Details

Before using this method it is strongly suggested to perform a filter on the basis of DABG p-values using the filtering function available in the filtering menu. DABG values can be calculated if exon array probe set data are generated using the oneChannelGUI graphical implementation to APT tools. Affymetrix suggests to calculate probe set intensity at gene level using iterPlier and at exon level using plier. Subsequently SpliceIndex need to be calculated using the function available in the exon menu. Finally the Rank Product method could be applied exon by exon. For more details on the method see RankProd package. Selection of putative alternative splicing could be done using the filtering function available in the filtering menu of oneChannelGUI

Note

IMPORTANT we are still evaluating the efficacy of this method for detection of alternative splicing events. Use it being concious of this!

Author(s)

Raffaele A Calogero

See Also

inspecting.splice, spliceIndex

`eSet4meV` *This function allows to reformat an eSet to be loaded in meV visualization software*

Description

The function allows to reformat an eSet to be loaded in meV visualization software. The eSet extracted from onechannelGUI exchanging array names, used for column names, with target Names.

Usage

```
eSet4meV()
```

Note

see oneChannelGUI vignette for more info

Author(s)

Raffaele A Calogero

---

exonContrasts — *Defining t-test regularized p-values*

Description

This function constructs the contrasts as affylmGUI but applied to exon-level. It also perfoms Bayes regularization. Raw p-values are plotted to see if BH or BY type I error correction can be applied. Corrected p-values are saved and used for extraction of alternative spliced exons.

Usage

```
exonContrasts()
```

Details

It is important to note that if multiple contrasts should be considered after the calculation of each of them it is essential to extract the alternative spliced exons with exonTopTableExtract, because everytime exonContrasts is run it overwrites the previous results.

Author(s)

Raffaele A Calogero

---

exonScaffold — *stand alone function to generate a scaffold containing only exon location from ENSEMBL*

Description

The scaffold contains chr stand start and end position of exon. Ensembl geneID and exonID together with exon rank are also provided

Usage

```
exonScaffold(genome)
```

Arguments

genome	The genome code for H sapiens is hg19, for M musculus is mm9 and for R norvegicus is rn4

Details

The scaffold is used to map reads data into regions only associated to exons

Author(s)

Raffaele A Calogero

---

exonsSpecific2as *Defining the exons associated to the various alternative isoforms*

Description

This function uses the output derived from the function mapping2ensembl and produces a list of 1 and 0 for each of the alternative trasncripts associated to a specific Entrez Gene. This funciton is useful to define which splicing events are not associated to exons conserved over all the possible isoforms

Usage

```
exonsSpecific2as()
```

Author(s)

Raffaele A Calogero

---

exonTopTableExtract *Extracts data on the basis of a defined t-test regularized p-value*

Description

This function filters the data produced by exonContrasts to extract a list of alternative spliced exons that are saved in a file and they can be used for further analysis, i.e. extracting only variant exons. The function also filter the data present in the onechannelGUI project bot at gene and exon-level.

Usage

```
exonTopTableExtract()
```

Details

It is important to note that if multiple contrasts should be considered after the calculation of each of them it is essential to extract the alternative spliced exons with exonTopTableExtract, because everytime exonContrasts is run it overwrites the previous results.

Author(s)

Raffaele A Calogero

extractAffyids — *Extracting probe ids associated to a specific Gene Ontology term*

Description

It is possible to identify the affy ids associated to a specific GO term using the extractAffyids function.

Usage

```
extractAffyids()
```

Details

The function asks to the user to select a file containing probe set ids separated by carriage return. The file should contain only one column and no header. The user is also asked to select a specific GO term. The probe sets associated to the specific GO term will be annotated ans saved in a HTML file.

Note

For the annotation the annotation library associated to the raw data loaded in the affylmGUI environment is used.

Author(s)

Raffaele A. Calogero

See Also

GOenrichment, plotGO

extractmirTargets — *Extract miRNA targets showing a opposite regulation with respect to miRNA*

Description

This function extract from a RmiR output the subset of genes showin a fold change variation inversely correlated to that of the miRNA selected by the user.

Usage

```
extractmirTargets()
```

Details

The output as the same structure of an RmiR output

Author(s)

Raffaele A Calogero

filteringmiRtargets

Subsetting an expression set using a list of gene which are putative targets for a miRNA

Description

This function subsets the normalized expression set present in the affylmGUI environment on the basis of a list of probe set ids extracted on the basis of the predicted gene targets for a microRNA. Extraction is based on predicted targets for miRNA in human.

Usage

```
filteringmiRtargets()
```

Details

The function asks to the user to select a file containing probe set ids separated by carriage return. The file should contain only one column and no header.

Note

In transcriptional studies focusing on genes characterized by specific feature (i.e. transcription factor elements in promoters) the best filtering approach is selecting only those genes linked to the interesting biological feature.

Author(s)

Raffaele A Calogero

See Also

IPAlistFilter, iqrFilter, intensityFilter

filteringTable

Filtering a tab delimited file

Description

This function allows to filter a tab delimited file using a vector of data present in an other file. The two files should have an header and the column name to be used for the filtering should be equal in both files

Usage

```
filteringTable()
```

Author(s)

Raffaele A Calogero

geneExonLibs — *Download the Library files for gene and exon analysis*

Description

Affymetrix Gene/Exon library files are necessary to APT tools to calculate probe set summaries. The versions downloaded from www.bioinformatica.unito.it, with htis function, contain all informations needed to analyse gene exon arrays.

Usage

```
geneExonLibs()
```

Author(s)

Raffaele A Calogero

genomePlot — *This function plots average intensity signals for two group experiment over the genes and transcripts structure mapped on ENSEMBL to identify a specific splicing event*

Description

This function plots average intensity signals over the genes and transcripts structure mapped on ENSEMBL to identify a specific splicing event

Usage

```
genomePlot()
```

Author(s)

Raffaele A Calogero

See Also

variantSI, variantExons, makeBED15

geoVSbioc — *linking GEO platforms to available BioC annotations libraries*

Description

This data file gives the linke between GEO platforms and BioC annotation libraries. If the GEO BioC link exists the Bioconductor annotation lib is directly loaded in the annotation fild of the SespressionSet

Usage

```
geoVSbioc
```

Format

A data frame with 4 observations: GEOAcc, Organisms, Title, BiocAnLIb

References

GEO and Bioconductor

getNGSannotation — *Retrieving genome annotation from ENSEMBL*

Description

This function retrieves genome annotation from ENSEMBL using ChIPpeakAnn package

Usage

```
getNGSannotation()
```

Author(s)

Raffaele A Calogero

`GOenrichment` *Searching for Gene Ontology enriched terms within a set of differentially expressed genes*

Description

In Bioconductor is available a library called GOstats, which allows the calculation of enriched GO terms within a set of differentially expressed probe sets. This is a graphical implementation of a function allowing the extraction of GO enriched term in a sub set of differentially expressed probe sets. To know more about it see GOstat library

Usage

```
GOenrichment()
```

Details

The function asks to the user to select a file containing probe set ids separated by carriage return. The file should contain only one column and no header. The set of enriched terms are plotted in red over the graph of all GO term associated to the differentially expressed genes. GO enriched terms can be also saved in a tab delimited file.

Author(s)

Raffaele A Calogero

References

Robert Gentleman GOstat package

See Also

extractAffyids, plotGO

`hsfa` *Human microRNA data set*

Description

microRNA data set fasta produced using miRbase H sapiens precursors reformated to have only miR name as descriptor of the fasta file

Usage

```
hsfa
```

Format

Fasta file

huex.annotation *Annotation data for CORE subset of human exon array 1.0 ST*

Description

These annotation data are derived from Affymetrix annotation file huex.annotation

Usage

```
huex.annotation
```

Format

A data frame with 5 observations: PROBESETID, SYMBOL, DESCRIPTION, CYTOBAND, ACC

References

Affymetrix web site

huex.variantexons *table linking exon-level probe set ids to variant exons*

Description

The table is made integrating exon-level netaffx annotaiton with UCSC derived annotation. Specifically exon-level rpobesets are mapped on variant exons, i.e. those that are specific of only a subset of all isoforms associated to a gene. This table is used to filter differentially expressed exons to select only those associated to variant isoforms

Usage

```
huex.variantexons
```

Format

A data frame with 12 observations: affyname, affystart, affyend, affywidth, affystrand, vspname, vspstart, vspend, vspwidth,vspstrand,chr,genome

References

Affymetrix web site and UCSC database

inspecting.one.splice.index

Plotting on the profiles of splice indexes for a transcript cluster ID

Description

This function plots the splice index profiles for one trnascript cluster ID

Usage

```
inspecting.one.splice.index()
```

Author(s)

Raffaele A Calogero

See Also

spliceIndex

inspecting.splice.index

Plotting on a pdf file the profiles of splice indexes

Description

This function prints in a pdf file the splice index profiles of the available genes

Usage

```
inspecting.splice.index()
```

Author(s)

Raffaele A Calogero

See Also

spliceIndex

intensityFilter — *intensity filtering with a mouse click*

Description

This function removes all probe sets in which a certain percentage of experiments is below a user defined intensity threshold.

Usage

```
intensityFilter()
```

Details

The aim of non specific filtering is to remove the genes that are unlikely to carry information about the phenotypes under investigation. This filtering remove genes that do not have a centain level of, user defined, intensities in a set of, user defined, experiments.

Note

Factor analysis will be limited by the problem of having fewer samples than genes. Therefore, preselecting a smaller set of genes is definetively helpful.

Author(s)

Raffaele A Calogero

See Also

iqrFilter, listFilter, IPAlistFilter

IPAlistFilter — *Filtering an expression set using a set of Entrez genes extracted from Ingenuity Pathways analysis (IPA)*

Description

It is possible to sub set an expression set loaded in the affylmGUI environment starting form a list of Entrez genes derived by IPA search tool.

Usage

```
IPAlistFilter()
```

Details

The function asks to the user to select a file containing Entrez genes separated by carriage return. The file should contain only one column and no header.

Author(s)

Raffaele A Calogero

See Also

iqrFilter, listFilter, intensityFilter

---

iqrFilter *Interquantile filtering with a mouse click*

Description

This function implements the interquantile filtering proposed by Heydebreck in 2004

Usage

```
iqrFilter()
```

Details

The aim of non specific filtering is to remove the genes that are unlikely to carry information about the phenotypes under investigation. This filtering remove genes that show little changes within the experimental points.

Note

Factor analysis will be limited by the problem of having fewer samples than genes. Therefore, preselecting a smaller set of genes is definetively helpful.

Author(s)

Raffaele A Calogero

References

Heydebreck et al. Bioconductor project Papers 2004

See Also

IPAlistFilter, listFilter, intensityFilter

`limma2paired` *graphical interface apply linear model to two sample groups with batch effect.*

Description

This function allows to run Bayes regularized t-test on two groups with batch effects. Derived from the mailing list question: Paired samples and origin question, Limma 2006-05-11 13:45:41 GMT

Usage

```
limma2paired()
```

Details

This function allows to run Bayes regularized t-test on two groups with batch effects.

Author(s)

Raffaele A Calogero

`limmaExons` *graphical interface to limma for alternative splicing detection*

Description

Applying the limma model fitting to eoxn-level data. Same implemetnaiton of AffylmGUI but applied to exon-level data. The first indicatin of alternative splicing detection using limma was proposed by Shah and Pallas in BMC Bioinformatics. 2009 Jan 20;10:26

Usage

```
limmaExons()
```

Details

The funciton fit the limma linear model to exon-level data

Author(s)

Raffaele A Calogero

listFilter — *Subsetting an expression set using a list of Affymetrix ids*

Description

This function subsets the normalized expression set present in the affylmGUI environment on the basis of a list of probe set ids passed via flat file.

Usage

```
listFilter()
```

Details

The function asks to the user to select a file containing probe set ids separated by carriage return. The file should contain only one column and no header.

Note

In transcriptional studies focusing on genes characterized by specific feature (i.e. transcription factor elements in promoters) the best filtering approach is selecting only those genes linked to the interesting biological feature.

Author(s)

Raffaele A Calogero

See Also

IPAlistFilter, iqrFilter, intensityFilter

makeBED15 — *This function creates files in BED15 format to be loaded on the UCSC browser*

Description

This function creates files in BED15 format to be loaded on UCSC genome browser. The functin uses the data derived by variantSI filter on the basis of chromosome annotation

Usage

```
makeBED15()
```

Author(s)

Raffaele A Calogero

See Also

variantSI, variantExons, plotVariantSI

`mapping2ensembl` *Associating e-level probe sets to entrez gene exonic structure*

Description

This function associates the statistical and expression data produced by a oneChannleGUI exon-level analysis to the exonic structure of Entrez Gene ID. This function uses biomaRt to retrieve the sequence of EG exons. RRE database is instead used to retrieve the exon-level target sequences. Any exon-level probe set id to be associated to the EG exonic sequence need to be a perfct matching substring of the exon. In the otehr case no exon is associated to the probe set

Usage

```
mapping2ensembl()
```

Author(s)

Raffaele A Calogero

`mapping2exon` *This function maps on exon-level Probe Selection Region (PSR) starting for the file produced by function oneChannelGUI: Mapping exon level probe sets to Reference Sequences*

Description

This function retrieve from RRE the PSR sequences associated to the exon-level probe sets and all exons associated to the gene associated to PRS. Subsequently identify the exon where PSR maps and procuces a fasta file were are located exon-level PSR and target exon. The mapping is done using the countPattern function of the Biostrings package. Up to three mismathces are allowed in PSR mapping on exonic sequence.

Usage

```
mapping2exon()
```

Author(s)

Raffaele A Calogero

`mapping2RefSeq` *This function maps on NCBI Reference sequences spliced exons detected by the function oneChannelGUI: Inspecting splice indexes*

Description

This function retrieve from RRE the PSR sequences associated to the exon-level probe sets using blastn detects the best refseq associated to any of the exon-level probe sets retrieve from org.XX.eg.db the EG associated to any of the detected refseq and retrieves all the refseqs associated to the EG. Subsequently check if PSR maps on all the refseqs associated to the eg (conserved exon) or only some of them (isoform specific exon)

Usage

```
mapping2RefSeq()
```

Author(s)

Raffaele A Calogero

`masigpro.edesign` *The function creates an edesign object needed to run maSigPro*

Description

The function creates an edesign object needed to run maSigPro. To know more about edesign object see maSigPro help. This function uses a specific configuration of Target column of the affylmGUI target file. To know more about target file see affylmGUI help. Each row of the column named Target, in the affylmGUI target file, describes the array on the basis of the experimental design. Each element needed for the construction of edesign is separated from the others by an underscore. The first three elements of the row are fixed and represent Time Replicate Control all separated by an underscore: `Time_Replicate_Control`. All the other elements refer to various experimental conditions. Considering two different conditions to be evaluated each row is made of 5 elements: `Time_Replicate_Control_cond1_cond2` all separated by an underscore. Having an experiment made of 9 arrays, with two time points, 0h and 24h, in triplicate, and two different experimental conditions to be evaluated, the affylmGUI target file will look like:

Name	FileName	Target
mC1	M1.CEL	`0_1_1_0_0`
mC2	M4.CEL	`0_1_1_0_0`
mC3	M7.CEL	`0_1_1_0_0`
mE1	M3.CEL	`24_2_0_1_0`
mE2	M6.CEL	`24_2_0_1_0`
mE3	M9.CEL	`24_2_0_1_0`
mI1	M2.CEL	`24_3_0_0_1`
mI2	M5.CEL	`24_3_0_0_1`
mI3	M8.CEL	`24_3_0_0_1`

Usage

```
masigpro.edesign()
```

Author(s)

Raffaele A Calogero

See Also

masigpro, masigpro.view

---

masigpro — *The function executes maSigPro analysis*

Description

The function creates: 1. Create a regression matrix for the full regression model (make.design.matrix function). 2. Computes the p-value associated to the F-Statistic of the model, which is used to select significant genes (p.vector function). 3. Applies a variable selection procedure to find significant variables for each gene (T.fit function). This will ultimatelly be used to find which are the profile differences between experimental groups. 4. Finally, it generates lists of significant genes according to R-squared of the models (get.siggenes function). To know more about the various steps see maSigPro help.

Usage

```
masigpro()
```

Author(s)

Raffaele A Calogero

See Also

masigpro.edesign, masigpro.view

---

masigpro.view — *The function allows the visualization of maSigPro results*

Description

The function is a graphical implementation of the maSigPro PlotGroups function. To know more about it see maSigPro help.

Usage

```
masigpro.view()
```

Author(s)

Raffaele A Calogero

See Also

masigpro.edesign, masigpro

---

metaArrayIC — *Graphical interface to metaArray Integrative Correlation function*

Description

The integrative correlation analysis (Parmigiani et al., 2004) is a convenient tool to monitor the interstudy concordance of within-study correlations of gene expression. The gene-specific reproducibility score takes the correlation between each gene and all other genes within individual study and calculate the average correlation of these correlations across all pairs of studies.

Usage

```
metaArrayIC()
```

Author(s)

Raffaele A Calogero

References

MergeMaid package and metaArray Package

---

metaArrayMerge — *Tool to create a merge object for metaArray package*

Description

This function will create an ExpressionSet from a study starting from a tab delimite file and a target file this ExpressionSet will be merged with the NormalizedAffyData if they contain the same number of row and rownames in the same order. Data generated with this function could be analyzed using metaArrayIC function.

Usage

```
metaArrayIC()
```

Author(s)

Raffaele A Calogero

See Also

mataArrayIC

`ML.edesign` *The function creates an data frame containing the parameters useful for class prediction*

Description

This function uses a specific configuration of Target column of the affylmGUI target file. To know more about target file see affylmGUI help. Each row of the column named Target, in the affylmGUI target file, describes the array on the basis of the experimental design. Each element needed for the construction of the data frame is separated from the others by an underscore. All the other elements refer to experimental conditions or clinical parameters. The absence of a parameter NEEDS to be described in the Target file by NA Considering two different conditions to be evaluated each row is made of 5 elements: `Time_Replicate_Control_cond1_cond2` all separated by an underscore. Having an experiment made of 9 arrays with 4 different experimental parameters the affylmGUI target file will look like:

Name	FileName	Target
mC1	M1.CEL	`0_1_pos_0_NA`
mC2	M4.CEL	`0_1_pos_0_yes`
mC3	M7.CEL	`0_1_neg_0_no`
mE1	M3.CEL	`24_2_neg_1_NA`
mE2	M6.CEL	`24_2_NA_1_yes`
mE3	M9.CEL	`24_2_neg_1_yes`
mI1	M2.CEL	`12_3_0_pos_yes`
mI2	M5.CEL	`12_3_0_pos_no`
mI3	M8.CEL	`12_3_0_pos_no`

Usage

```
ML.edesign()
```

Author(s)

Raffaele A Calogero

`mmfa` *Mouse microRNA data set*

Description

microRNA data set fasta produced using miRbase M musculus precursors reformated to have only miR name as descriptor of the fasta file

Usage

```
mmfa
```

Format

Fasta file

moex.annotation *Annotation data for CORE subset of mouse exon array 1.0 ST*

Description

These annotation data are derived from Affymetrix annotation file moex10stv1na23mm8.

Usage

```
moex.annotation
```

Format

A data frame with 5 observations: PROBESETID, SYMBOL, DESCRIPTION, CYTOBAND, ACC

References

Affymetrix web site

moex.variantexons *table linking exon-level probe set ids to variant exons*

Description

The table is made integrating exon-level netaffx annotaiton with UCSC derived annotation. Specifically exon-level rpobesets are mapped on variant exons, i.e. those that are specific of only a subset of all isoforms associated to a gene. This table is used to filter differentially expressed exons to select only those associated to variant isoforms

Usage

```
moex.variantexons
```

Format

A data frame with 12 observations: affyname, affystart, affyend, affywidth, affystrand, vspname, vspstart, vspend, vspwidth,vspstrand,chr,genome

References

Affymetrix web site and UCSC database

`mRNAmiRCor` *Filtering by mean of correlation between expression data and miRNA expression changes*

Description

This function extract the subset of gene ids which are associated to the changes of expression on putative miRNA targets upon perturbation of miR expression.

Usage

```
mRNAmiRCor()
```

Details

To run this analysis a time course experiment suitable for maSigPro analysis is needed. Expression data need to have at least 3 time points for samples transfected with a scrambled and with an Antagomir or a Mimic. The expression changes at the same time points used for microarray data need to be measured also for miRNA. miRNA expression data need to be saved as table delimited file with two columns: column Name, which has to be organized exactly as the target file Name column,and a column named Value with expression of miRNA, e.g. deltaCt. Target file for mRNA data need to be organized as for maSigPro time course analysis. Output is a three column (gene level id, miRNA-mRNA correlation, mRNA untreated - mRNA treated correlation) tab delimited file with gene IDs and correlation coefficient between mRNA profile and miRNA profile upon treatment with antagomir, i.e. mRNA expression should rise due to the reduction of miRNA expression, or mimics, i.e. mRNA expression should be reduced due to the over-expression of the specific miRNA.

Author(s)

Raffaele A Calogero

See Also

IPAlistFilter, iqrFilter, intensityFilter

`myExpresso` *Running the affy expresso function with the widget*

Description

Various probe set intensity summary and normalization can be customized using the expresso function.

Usage

```
myExpresso()
```

Details

This function run expresso with the graphical interface for parameters selection. It is important to note that expresso is more slow than the C coded rma)

Author(s)

Raffaele A Calogero

---

ncHs.data — *Subset of genomic regions with ncRNA of Homo sapiens*

Description

Two IRanges object referring to the plus and minus strand of genomic region encoding for ncRNAs. The ncRNAs taken in consideration are: miRNA, Mitocondrial rRNA, Mitocondrial tRNA, rRNA, snoRNA, snRNA

Usage

```
ncHs.data
```

Format

A list with two observations: IRanges object referring to the plus and minus strand of genomic region encoding ofr ncRNAs

Author(s)

Raffaele A Calogero

---

ncMm.data — *Subset of genomic regions with ncRNA of Mus musculus*

Description

Two IRanges object referring to the plus and minus strand of genomic region encoding ncRNAs. The ncRNAs taken in consideration are: miRNA, Mitocondrial rRNA, Mitocondrial tRNA, rRNA, snoRNA, snRNA

Usage

```
ncMm.data
```

Format

A list with two observations: IRanges object referring to the plus and minus strand of genomic regions encoding ncRNAs

Author(s)

Raffaele A Calogero

ncRn.data — *Subset of genomic regions with ncRNA in Rattus norvegicus*

Description

Two IRanges object referring to the plus and minus strand of genomic region encoding ncRNAs. The ncRNAs taken in consideration are: miRNA, Mitocondrial rRNA, Mitocondrial tRNA, rRNA, snoRNA, snRNA

Usage

```
ncRn.data
```

Format

A list with two observations: IRanges object referring to the plus and minus strand of genomic region encoding ncRNAs

Author(s)

Raffaele A Calogero

ncScaffold — *stand alone function to generate a scaffold containing only ncRNA location from ENSEMBL*

Description

The scaffold contains chr stand start and end position of ncRNA. ENSEMBL retrieved spicies are miRNA,Mt_rRNA,Mt_tRNA,rRNA,snoRNA,snRNA

Usage

```
ncScaffold(genome, fasta)
```

Arguments

genome	The genome code for H sapiens is hg19, for M musculus is mm9 and for R norvegicus is rn4
fasta	If fasta is TRUE a fasta file with the sequences retrieved by ncScaffold will be produced

Details

The scaffold is used to map reads data into regions only associated to ncRNAs

Author(s)

Raffaele A Calogero

NGSreformat — *reformating output of NGS primary tools software*

Description

This function allows to reorganize the output files produced by SHRIMP, MicroRazerS in two files with the extension: bed and logos used by oneChannelGUI

Usage

```
NGSreformat()
```

Author(s)

Raffaele A Calogero

normBoxplot — *Box plot of the arrays data available in NormalizeAffyData slot*

Description

Box plot visualization of normalized array data

Usage

```
normBoxplot()
```

Author(s)

Raffaele A Calogero

ocPlotHist — *Gene/Exon level density plots*

Description

This function runs a modified version of the plotHist of the affycoretools to be used to check density distribution plots for gene and exon expression data generated by expression console.

Usage

```
ocPlotHist()
```

Author(s)

Raffaele A Calogero

See Also

ocPlotPCA

ocPlotPCA *Gene/Exon level density plots*

Description

This function runs a modified version of the plotPCA of the affycoretools to be used to check density distribution plots for gene and exon expression data.

Usage

```
ocPlotPCA()
```

Author(s)

Raffaele A Calogero

See Also

ocPlotHist

oneChannelGUI-package *Set of functions extending the capability of affylmGUI package*

Description

This package is directed to Bioconductor beginners that have little or no experience in writing R code. The package implements, as simple functions accessible over the affylmGUI graphical interface, some code useful for QC, data filtering, data output manipulation and identification of GO enriched classes.

Details

Package:	oneChannelGUI
Type:	Package
Version:	1.0
Date:	2006-12-16
License:	GPL version 2 or newer

Author(s)

Author: Raffaele A Calogero Maintainer: Raffaele A Calogero <raffaele.calogeor@unito.it>

Examples

```
#  library(oneChannelGUI)
```

```
## To start the oneChannelGUI with the modifications
#oneChannelGUI()
```

`oneChannelGUI` *Starting oneChannelGUI package*

Description

Starting oneChannelGUI package. oneChannelGUI contains a set of functions extending the capabilities of affylmGUI package

Usage

```
AboutextendedaffylmGUI()
AboutaffylmGUI()
oneChannelGUI()
oneChannelGUIHelp()
maSigProHelp()
siggenesHelp()
oneChannelGUIHelp()
initialize.extAffylmGUI()
OpenExonandTargetsfiles()
GOstatsHelp()
NewLimmaFile()
OpenLimmaFile()
OpenALimmaFile(FileName)
OpenAFile(FileName)
OpenExonFile()
OpenLargeFile()
changeMenu()
oneChannelGUI.start()
libraryFilesDir()
whichKindOfArray()
intronicBg()
ExportNormalizedExpressionValues1()
ExportNormalizedExpressionValues()
ExportfeatureNames()
SaveAsLimmaFile()
addAnnLib()
OpenCDFandTargetsfiles()
ComputeLinearModelFit()
GetNormalizationMethod()
NormalizeNow()
ComputeContrasts()
midasFilter()
dabgFilter()
largedatasetNorm()
RankProdHelp()
affyPLMHelp()
genefilterHelp()
pamrHelp()
```

```
pdmclassHelp()
sizepowerHelp()
ssizeHelp()
OpenAGeoFile(FileName)
OpenGeoFile()
OpenGeoFiles()
OpenNgsfiles()
recoverUnfiltered()
delete.ML()
affyPlotMA()
changes()
aptFolder()
deleteLocalData()
SetED()
chooseEDir()
.annotation(eset)
.myfindOverlaps(x,y)
log2Conversion()
.myrk(x,df)
exportFASTA()
OpenmiRanalyserfiles()
OpenmiRProfiles()
OpenmiRExpress()
OpenmiRShrimp()
ngsperlDownload()
meVDownload()
extPckInfo()
startmeV()
```

Arguments

`FileName`	Internal argument not to be set by the user
`eset`	Internal argument not to be set by the user
`x`	Internal argument not to be set by the user
`y`	Internal argument not to be set by the user
`df`	Internal argument not to be set by the user

Details

This function launches a modify version of the Graphical User Interface by James Wettenhall for the limma package by Gordon Smyth. The GUI uses Tk widgets (via the R TclTk interface by Peter Dalgaard) in order to provide a simple interface to various tools for quality control and statistical analysis of Affymetrix gene chips.

Author(s)

Raffaele A Calogero

Examples

```
#  library(oneChannelGUI)
## To start the affylmGUI with the modifications
#oneChannelGUI()
```

OpenBeadStudioFiles

Read BeadStudio expression data file

Description

Read BeadStudio expression data file

Usage

```
OpenBeadStudioFiles()
```

Details

Reads an Illumina intnesity data file produced by BeadStudio. Using BeadStudio version 'One" the file will have a 'gene profile.csv' extension and using version "Two" the file will have a .txt extension. See package vignette for more information. Multiple filenames can be specified as a vector and the data are then combined into one output file. This function should only really be used for custom analysis as the beadAnalysis() function provides easier, flexible use.

Author(s)

Derived from readBead by Gareth Elvidge <gareth.elvidge@well.ox.ac.uk>

OpenLargefiles

This function loads large data set made from tab delimited files

Description

The function creates and expressionSet starting from al file containing the expresison data in a tab delimited format. This file is loaded together with the description of the clinical paramenters present in Target This function uses a specific configuration of Target column of the affylmGUI target file. To know more about target file see affylmGUI help. Each row of the column named Target, in the affylmGUI target file, describes the clinical parameters. Each clinical parameter is separated from the others by an underscore. The affylmGUI target file will look like:

Name	FileName	Target
mC1	M1.CEL	pos_yes_1_NA_0
mC2	M4.CEL	pos_no_2_NA_0
mC3	M7.CEL	neg_no_3_pos_0
mE1	M3.CEL	neg_yes_3_neg_0
mE2	M6.CEL	neg_no_NA_1_0
mE3	M9.CEL	neg_yes_3_pos_0
mI1	M2.CEL	pos_no_2_neg_1
mI2	M5.CEL	pos_yes_2_pos_1
mI3	M8.CEL	pos_no_2_pos_1

Usage

```
OpenLargefiles()
```

Author(s)

Raffaele A Calogero

---

plierToZero — *Setting to 0 low log2 intensity values generated with plier*

Description

The calculation of log2 of probe set intensity by mean of plier generates a set of intensities very low this function will set to 0 all the log2 intensities below 1 produced by iter-plier or plier alghoritm

Usage

```
plierToZero()
```

Author(s)

Raffaele A Calogero

---

plotGO — *Plotting parents of a GO term with few mouse click*

Description

To know more on the parents of a specific GO term you can use the plotGO function

Usage

```
plotGO()
```

Details

A GO term to be investigated for its parents has to be placed in the graphical window.

Author(s)

Raffaele A Calogero

See Also

GOenrichment, extractAffyids

`PlotOptionsv1` *A modified version of the function used in affyPLM library*

Description

As default the plots are generated on the R GUI to reduce RAM consumption.

Usage

```
PlotOptionsv1()
```

Author(s)

Raffaele A Calogero

`plotVariantSI` *This function plots on UCSC genome browser data derived by variantSI filter*

Description

This function plots on UCSC genome browser data derived by variantSI filter on the basis of chromosome annotation

Usage

```
plotVariantSI()
```

Author(s)

Raffaele A Calogero

See Also

variantSI, variantExons, makeBED15

`qcMDS` *This function plots the sample relations based on multidimensional scaling using the plotMDS.dge edgeR function*

Description

This function is a variation on the usual multdimensional scaling (or principle coordinate) plot, in that a distance measure particularly appropriate for the digital gene expression (DGE) context is used. The distance between each pair of samples (columns) is the square root of the common dispersion for the top genes which best distinguish that pair of samples. These top genes are selected according to the tagwise dispersion of all the samples.

Usage

```
qcMDS()
```

Author(s)

Raffaele A Calogero

`raex.annotation` *Annotation data for CORE subset of rat exon array 1.0 ST*

Description

These annotation data are derived from Affymetrix annotation file raex.annotation

Usage

```
raex.annotation
```

Format

A data frame with 5 observations: PROBESETID, SYMBOL, DESCRIPTION, CYTOBAND, ACC

References

Affymetrix web site

raex.variantexons *table linking exon-level probe set ids to variant exons*

Description

The table is made integrating exon-level netaffx annotaiton with UCSC derived annotation. Specifically exon-level rpobesets are mapped on variant exons, i.e. those that are specific of only a subset of all isoforms associated to a gene. This table is used to filter differentially expressed exons to select only those associated to variant isoforms

Usage

```
raex.variantexons
```

Format

A data frame with 12 observations: affyname, affystart, affyend, affywidth, affystrand, vspname, vspstart, vspend, vspwidth,vspstrand,chr,genome

References

Affymetrix web site and UCSC database

rankingConversion *This function transforms intensity data in normalized ranks*

Description

This function transforms intensity data in normalized ranks, i.e. high intensity genes will have a value near to 0 as instead low intensity genes a normalized rank near to 1.

Usage

```
rankingConversion()
```

Author(s)

Raffaele A Calogero

`rankProd` *graphical interface to rank product method implemented in RankProd Bioconductor library.*

Description

To know more about rank product method see RankProd help.

Usage

```
rankProd()
```

Details

The target file for the RankProd implementation contain the origin of the data as a number separated by an under score from the corresponding covariate. If all data are from the same origin the origin definition is not needed. Therefore target will contain only the covariates.

Name	FileName	Target
mC1	M1.CEL	`CTRL_1`
mC2	M4.CEL	`CTRL_1`
mC3	M7.CEL	`CTRL_2`
mE1	M3.CEL	`CTRL_2`
mE2	M6.CEL	`CTRL_2`
mE3	M9.CEL	`TRT_1`
mI1	M2.CEL	`TRT_1`
mI2	M5.CEL	`TRT_2`
mI3	M8.CEL	`TRT_2`

Author(s)

Raffaele A Calogero

`rawBoxplotPN` *Plotting raw log2 intensities from controls*

Description

This function produces a box plot of the log2 raw intensities, extracted directly from CEL files, for positive and negative controls presente in XXXX.control.ps Affymetrix library file. Positive and negative controls are made of housekeeping exon and introns. It gives an idea of signal behaviour before data normalization both in the high and low intensity range

Usage

```
rawBoxplotPN()
```

Author(s)

Raffaele A Calogero

rawpCheck — *Raw p-value distribution from limma analysis by a mouse click*

Description

This function allow to visualize the histogram of raw p-value distribution generated by limma analysis.

Usage

```
rawpCheck()
```

Details

The histogram of raw p-value distribution will show if raw p-values are uniform in the non significant range and therefore the BH correction can be applied.

Note

BH is the most used method for the correction of type I errors in microarray analysis. However, it has some limitation due to the initial hypotheses: The gene expressions are independent from each other. The raw distribution of p values should be uniform in the non significant range.

Author(s)

Raffaele A Calogero

References

To know more see limma package help

Reads2logos — *Extracting info on the counts associated to a differentially expressed ncRNA gene*

Description

This function allow to plot the raw counts associated to samples or experimental groups associated to a ncRNA found differentially expressed. Furthermore, it saves the count matrix for the mapped positions on the differentially expressed ncRNA

Usage

```
Reads2logos()
```

Author(s)

Raffaele A Calogero

`refiningPeaks` *This function refines the structure of the genomics peaks, checking for the presence nearby peaks that cna be combined*

Description

This function allows the user to define a threshold for peaks merging. if threshold is set to 0 non merging is provided.

Usage

```
refiningPeaks()
```

Author(s)

Raffaele A Calogero

`reformatGdl` *This function reorganizes single NGS data in a matrix to be used for statistical analysis.*

Description

This function reorganizes the raw NGS data saved in a GenomeDataList object in a matrix. Matrix rows are peaks containing multiple reads. For peak detection, user has to indicate the amount of extension of each read, e.g. 200 nt in chipseq experiments, and the number of reads due by random events, e.g. 8.

Usage

```
reformatGdl()
```

Author(s)

Raffaele A Calogero

`refseqDownload` *Retrieving Reference Sequences from NCBI ftp*

Description

This function retieves refeence sequences from NCBI ftp. RefSeq are used for mapping exon-level probe sets to refseq specific isoforms.

Usage

```
refseqDownload()
```

Author(s)

Raffaele A Calogero

`retrieveMirTargets` *graphical interface to RmiR.Hs.miRNA library.*

Description

This function retrieves putative targets from RmiR.Hs.miRNA given a specific microRNA.

Usage

```
retrieveMirTargets()
```

Details

This interfase provides a graphical GUI to RmiR.Hs.miRNA.

Author(s)

Raffaele A Calogero

`retrievePSRseq1gid` *This function, given a gene-level probeset id, retrieves Probe selection Regions, PRS, from RRE db*

Description

This function retrieves from RRE the PSR sequences associated to exon-level probe sets

Usage

```
retrievePSRseq1gid()
```

Author(s)

Raffaele A Calogero

`retrievePSRseq` *This function, given a file containing exon-level probesets retrieves Probe selection Regions, PRS, from RRE db*

Description

This function retrieve from RRE the PSR sequences associated to the exon-level probe sets

Usage

```
retrievePSRseq()
```

Author(s)

Raffaele A Calogero

reviqrFilter — *Reverse interquantile filtering with a mouse click*

Description

This function implements a reverse version of the interquantile filtering proposed by Heydebreck in 2004 to select low variance genelevel probe set. To be used to remove putative differentially expressed genes that will make more difficult the detection of alternative splicing events.

Usage

```
reviqrFilter()
```

Details

This function can be used in a analysis focused to the detection of alternative splicing events. The aim of this non specific filtering is to remove the genes that are likely to carry information about the phenotypes under investigation at gene level. This filtering remove genes that show strong changes within the experimental points at the gene level.

Author(s)

Raffaele A Calogero

References

Heydebreck et al. Bioconducotor project Papers 2004

See Also

dabgFilter, crosshybFilter

RmiRInterface — *graphical interface to RmiR library.*

Description

RmiR package is to couple microRNA and gene expression data (coming from the same RNA). Input data are obtainable with limma interface (gene-level differential expression) in oneChannelGUI and baySeq interface (NGS microRNA differential expression).

Usage

```
RmiRInterface()
```

Details

This interfase provides a graphical GUI to RmiR package. The output structure is described in RmiR package.

Author(s)

Raffaele A Calogero

---

rnfa — *Rat microRNA data set*

Description

microRNA data set fasta produced using miRbase R norvegicus precursors reformated to have only miR name as descriptor of the fasta file

Usage

```
rnfa
```

Format

Fasta file

---

runningJetta — *graphical interface to MADS/jetta R library.*

Description

MADS, which stands for Microarray Analysis of Differential Splicing, is a tool to identify differential alternative splicing by exon array. The principle of MADS is to increase the precision of exon-level and gene-level expression estimates by correcting, as much as possible, noise in observed probe intensities due to background and cross-hybridization. MADS incorporates a series of novel algorithms motivated by the probe-rich design of exon-tiling arrays, such as background correction, iterative probe selection and removal of sequence-specific cross-hybridization to off-target transcripts. MADS was published in RNA,2008,14(8): 1470-1479. Junction and Exon array Toolkit for Transcriptome Analysis (JETTA) is compacted version of MADS.

Usage

```
runningJetta()
```

Details

Expression indexes are calculated as the order of Background Correction, Normalization and Summarization. In the Summarization step, background corrected and normalized probe intensities of a meta probeset are summarized to expression of the meta probeset. Meta probesets can be defined as gene/transcript clust/exon level.

Background Correction JETTA estimates background signal using background probes and subtracts it from the probe intensity. If the probe intensity is less than the estimated background signal, the background subtracted signal is truncated to 1. Estimation of background signal is based on several models: Median GC: median of background probe signal of the same GC counts MAT: linear model of probe sequence with 80 parameters. see Kapur et al, 2007

Normalization Normalization of JETTA is done for core probes defined in probeset annotation file. If the PSA file is not specified, it considers all probes in the MPS files as core probes. Median scaling: scaling each array so that its median is 100 Quantile: all probes of the same signal quantile have the same signal

Summarization LiWong model: multiplication model of expression and probe effect, see Li and Wong, 2001 Probe selection: select probes based on cross-array correlation of signal. see Xing et al, 2006 Median-polish

Alternative Splicing Detection Detecting alternatively expressed PSR/Exon between two sample groups based on background corrected and normalized probe intensities. It has several criteria to filter out transcript clusts and probes from the analysis. TC expression level: excluding low-expressed transcript clusts TC expression fold change: excluding transcript clusts which have big fold change between two groups Extreme probe signal: excluding probes of which signal is extremely high Cross-hybridized probes: excluding cross hybridized probes, currently pre-calculated results are needed

Author(s)

jseok@stanford.edu

`sample.size.evaluation1`
The function executes functions from the sizepower packahe

Description

This function represents a tool for helping users to understand the trade off between sample size and statistical power. To know more about see ssize help.

Usage

```
sample.size.evaluation1()
```

Details

see sizepower help

Author(s)

Raffaele A Calogero

sample.size.evaluation

The function executes and plots results from ssize and delta fulction from the ssize packahe

Description

This function represents avisual tool for helping users to understand the trade off between sample size and statistical power. To know more about see ssize help.

Usage

```
sample.size.evaluation()
```

Details

Both ssize and delta outputs are calculated using the BH type I error correction instead of the Bonferroni used as default in the ssize package. Furthermore, instead using the control group variance, this implementation uses the common variance described in Wei et al. BMC Genomics. 2004, 5:87

Main assumptions: A microarray experiment is set up to compare gene expressions between one treatment group and one control group. Microarray data has been normalized and transformed so that the data for each gene is sufficiently close to a normal distribution that a standard 2-sample pooled-variance t-test will reliably detect differentially expressed genes.

Author(s)

Raffaele A Calogero

showDataset

Grabbing info about the available expression set

Description

The size of the normalized expression set can change upon filtering. This function show info about the exact size of the data set.

Usage

```
showDataset()
```

Author(s)

Raffaele A Calogero

`showTopTable` *Modification of the function immplemented in affylmGUI to generate a topTable*

Description

Modification of the function immplemented in affylmGUI to generate a topTable. To know more about topTable see limma package help

Usage

```
showTopTable(...,export=FALSE)
```

Arguments

`export`	defining the possibility to export data
`...`	Arguments to be passed to methods

Author(s)

Raffaele A Calogero

`siggenes` *The function executes SAM analysis implemented in siggenes bioconductor library*

Description

To know more about SAM in Bioconductor see siggenes help.

Usage

```
siggenes()
```

Author(s)

Raffaele A Calogero

simFilter — *This function allows filtering on the basis of the average splice index mean or min difference between two groups*

Description

Filtering out gene/exon level probe sets associated to average splice index mean or min difference between two groups lower than user defined value

Usage

```
simFilter()
```

Note

This fuction needs the presence of Splice Index.

Author(s)

Raffaele A Calogero

See Also

simFilter

spliceIndex — *This function coverts the exon intenisty data in a slice index*

Description

Exons intensities are divided for the expression of the corresponding gene, as descrived by Clark et al. Science 2002 May 3;296(5569):907-10.

Usage

```
spliceIndex()
```

Details

The function is not yet optimized, therefore it could take quite a long time to compute spliceIndex if more than 1000 genes are used.

Author(s)

Raffaele A Calogero

See Also

inspecting.splice.index

`standAloneAddingAnnotation`
Attach to a data frame containing gene-level data derived from Affymetrix exon arrays the annotations derived by netaffx

Description

Standalone oneChannelGUI function attaches gene-level annotation to a data frame.

Usage

```
standAloneAddingAnnotation(annotationdf, df.tobe.annotated, ids.column)
```

Arguments

`annotationdf` An annotation data frame generated with standAloneBuildingLocalAnnotation function

`df.tobe.annotated` A data frame containing a gene-level data of any type. It is mandatory that onne of the column contains gene-level ids

`ids.column` the column of the df.tobe.annotated containing gene-level ids

Value

A data frame.

Author(s)

Raffaele A Calogero

`standAloneBuildingLocalAnnotation`
Creates a data frame with gene-level annotation data for exon arrays using the netaffx database

Description

Standalone oneChannelGUI function to create gene-level annotation data using netaffx database.

Usage

```
standAloneBuildingLocalAnnotation(libDirLocation = getwd(), netaffxUser = "my
```

Arguments

`libDirLocation` Folder where to save the annotation object

`netaffxUser` The email registered to Affymetrix netaffx web site

`netaffxUserPw` The password to access to netaffx

`whichAnnotation` Which annotation table should be used

Value

Location of the annotation data frame.

Author(s)

Raffaele A Calogero

`targetWidget` *Widget to create a target file to load .CEL files*

Description

Widget to create a target file to load .CEL files to be used with NewLimmaFile function.

Usage

```
targetWidget()
```

Author(s)

Raffaele A Calogero

`templA` *Generating a template A to be uploaded in Ingenuity Pathways analysis (IPA)*

Description

A template A file to be used in Ingenuity can be generated starting from a topTable containing the full array data.

Usage

```
templA()
```

Note

Template A file will contain a column with the gene ID, a column with fold change, a column with true p-value and a column with p-values for discriminating between the set of differentially expressed probe sets and the background. This coulumn is needed to allow IPA to identify the set of enriched functional classes associated to the differentially expressed probe sets.

Author(s)

Raffaele A Calogero

See Also

IPAlistFilter

trainTest — *Creating a training set and a test set by a mouse click*

Description

This function allows the creation of a training set and a test set to be used for classification purposes.

Usage

```
trainTest()
```

Details

User will be asked to assign names to the available classification parameters. User will be asked to select the number associated to one of the available classification parameters. The training set will be made, using the selected classification parameter and it will be made of 2/3 of the original data set. The test set will be the remaining 1/3.

Author(s)

Raffaele A Calogero

updateLibs — *This function allows to update the present installation of Bioconductor*

Description

The function allows the updating of local installation of Bioconductor. It might be quite long depending on the internet connection speed.

Usage

```
updateLibs()
```

Author(s)

Raffaele A Calogero

`variantExons` *This function is used to generate a table containing exon-level probe set data linked to variant exons*

Description

Internal oneChannelGUI annotation data linking exon-level probesets to variant exons, i.e. those exons that ar specific for a subgroup of the isoforms associated to a specific gene, can be upgraded using this function. Annotation RDA files need to be saved in the data subdir of oneChannelGUI dir. Windows users need to exchange the older copies present in Rdata.zip, simply dragging them in the zip file.

Usage

```
variantExons()
```

Author(s)

Raffaele A Calogero

`variantSI` *This function allows filtering on the basis of variant exons*

Description

The function interesects a list of alternative splice exon-level probe sets detected by oneChannelGUI analysis and intersects it to the list of exon-level probe sets associated to variant exons, i.e. these exons that are associated only to a subset of all isoforms associated to a gene. The table of variant Exon is stored in RRE database and is retrieved by the function updating the UCSC tables linking probe set ids with variant exons, located in the General Tools menu. The variantSI function also attaches to the variant exons table the deltaSI associated exon-level probe set.

Usage

```
variantSI()
```

Note

This fuction needs the presence of Splice Index

Author(s)

Raffaele A Calogero

See Also

variantExons, plotVariantSI

`VennDiagram` *Venn diagrams using two or three lists*

Description

Venn diagrams can be generated using probe sets ids or Entrez gene ids saved in flat files.

Usage

```
VennDiagram()
```

Details

The function asks to the user to select two/three files containing probe set ids or EGs separated by carriage return. Each file should contain only one column and no header.

Author(s)

Raffaele A Calogero

Index

∗Topic **IO**
oneChannelGUI, 38
∗Topic **datasets**
bofa, 3
chrLength, 4
cosieHscore, 6
cosieHsfull, 7
cosieMmcore, 7
cosieMmfull, 8
geoVSbioc, 19
hsfa, 20
huex.annotation, 21
huex.variantexons, 21
mmfa, 31
moex.annotation, 32
moex.variantexons, 32
ncHs.data, 34
ncMm.data, 34
ncRn.data, 35
raex.annotation, 43
raex.variantexons, 44
rnfa, 50
∗Topic **hplot**
dfMAplot, 10
normBoxplot, 36
plotGO, 41
rawpCheck, 46
VennDiagram, 59
∗Topic **manip**
annotateNGSeset, 1
AptMidas, 1
biomartFilter, 2
colExtract, 4
consistentFilters, 6
cosieWrapper, 8
edgeRNorm, 11
EG2probeset, 12
erankProdAltSpl, 13
erankProdAltSplFilter, 12
exonsSpecific2as, 15
filteringTable, 17
genomePlot, 18
getNGSannotation, 19
makeBED15, 26
mapping2ensembl, 27
mapping2exon, 27
mapping2RefSeq, 28
metaArrayIC, 30
metaArrayMerge, 30
NGSreformat, 36
plierToZero, 41
plotVariantSI, 42
qcMDS, 43
rankingConversion, 44
rawBoxplotPN, 45
Reads2logos, 46
refiningPeaks, 47
reformatGdl, 47
refseqDownload, 47
retrievePSRseq, 48
retrievePSRseq1gid, 48
simFilter, 54
variantSI, 58
∗Topic **methods**
buildingLocalAnnotation, 3
combineGeoMSF, 5
combining2eSet, 5
createGeoTarget, 9
crosshybFilter, 9
dfMAplot, 10
eSet4meV, 13
extractAffyids, 16
filteringmiRtargets, 17
geneExonLibs, 18
GOenrichment, 20
inspecting.one.splice.index, 22
inspecting.splice.index, 22
intensityFilter, 23
IPAlistFilter, 23
iqrFilter, 24
listFilter, 26
mRNAmiRCor, 33
OpenBeadStudioFiles, 40
PlotOptionsv1, 42
reviqrFilter, 49

showDataset, 52
showTopTable, 53
spliceIndex, 54
standAloneAddingAnnotation, 55
standAloneBuildingLocalAnnotation, 55
templA, 56
updateLibs, 57
variantExons, 58
*Topic **package**
oneChannelGUI-package, 37
*Topic **utilities**
bayseqInterface, 2
ctrtrtHeatmap, 10
edgerInterface, 11
exonContrasts, 14
exonScaffold, 14
exonTopTableExtract, 15
extractmirTargets, 16
limma2paired, 25
limmaExons, 25
masigpro, 29
masigpro.edesign, 28
masigpro.view, 29
ML.edesign, 31
myExpresso, 33
ncScaffold, 35
ocPlotHist, 36
ocPlotPCA, 37
OpenLargefiles, 40
rankProd, 45
retrieveMirTargets, 48
RmiRInterface, 49
runningJetta, 50
sample.size.evaluation, 52
sample.size.evaluation1, 51
siggenes, 53
targetWidget, 56
trainTest, 57
.annotation *(oneChannelGUI)*, 38
.myfindOverlaps *(oneChannelGUI)*, 38
.myrk *(oneChannelGUI)*, 38

AboutaffylmGUI *(oneChannelGUI)*, 38
AboutextendedaffylmGUI *(oneChannelGUI)*, 38
addAnnLib *(oneChannelGUI)*, 38
affyPLMHelp *(oneChannelGUI)*, 38
affyPlotMA *(oneChannelGUI)*, 38
annotateNGSeset, 1
aptFolder *(oneChannelGUI)*, 38
AptMidas, 1

bayseqInterface, 2
biomartFilter, 2
bofa, 3
buildingLocalAnnotation, 3

changeMenu *(oneChannelGUI)*, 38
changes *(oneChannelGUI)*, 38
chooseEDir *(oneChannelGUI)*, 38
chrLength, 4
colExtract, 4
combineGeoMSF, 5
combining2eSet, 5
ComputeContrasts *(oneChannelGUI)*, 38
ComputeLinearModelFit *(oneChannelGUI)*, 38
consistentFilters, 6
cosieHscore, 6
cosieHsfull, 7
cosieMmcore, 7
cosieMmfull, 8
cosieWrapper, 8
createGeoTarget, 9
crosshybFilter, 9
ctrtrtHeatmap, 10

dabgFilter *(oneChannelGUI)*, 38
delete.ML *(oneChannelGUI)*, 38
deleteLocalData *(oneChannelGUI)*, 38
dfMAplot, 10

edgerInterface, 11
edgeRNorm, 11
EG2probeset, 12
erankProdAltSpl, 13
erankProdAltSplFilter, 12
eSet4meV, 13
exonContrasts, 14
exonScaffold, 14
exonsSpecific2as, 15
exonTopTableExtract, 15
exportFASTA *(oneChannelGUI)*, 38
ExportfeatureNames *(oneChannelGUI)*, 38
ExportNormalizedExpressionValues *(oneChannelGUI)*, 38
ExportNormalizedExpressionValues1 *(oneChannelGUI)*, 38
extPckInfo *(oneChannelGUI)*, 38
extractAffyids, 16

extractmirTargets, 16

filteringmiRtargets, 17
filteringTable, 17

geneExonLibs, 18
genefilterHelp *(oneChannelGUI)*, 38
genomePlot, 18
geoVSbioc, 19
getNGSannotation, 19
GetNormalizationMethod *(oneChannelGUI)*, 38
GOenrichment, 20
GOstatsHelp *(oneChannelGUI)*, 38

hsfa, 20
huex.annotation, 21
huex.variantexons, 21

initialize.extAffylmGUI *(oneChannelGUI)*, 38
inspecting.one.splice.index, 22
inspecting.splice.index, 22
intensityFilter, 23
intronicBg *(oneChannelGUI)*, 38
IPAlistFilter, 23
iqrFilter, 24

largedatasetNorm *(oneChannelGUI)*, 38
libraryFilesDir *(oneChannelGUI)*, 38
limma2paired, 25
limmaExons, 25
listFilter, 26
log2Conversion *(oneChannelGUI)*, 38

makeBED15, 26
mapping2ensembl, 27
mapping2exon, 27
mapping2RefSeq, 28
masigpro, 29
masigpro.edesign, 28
masigpro.view, 29
maSigProHelp *(oneChannelGUI)*, 38
metaArrayIC, 30
metaArrayMerge, 30
meVDownload *(oneChannelGUI)*, 38
midasFilter *(oneChannelGUI)*, 38
ML.edesign, 31
mmfa, 31
moex.annotation, 32
moex.variantexons, 32
mRNAmiRCor, 33

myExpresso, 33

ncHs.data, 34
ncMm.data, 34
ncRn.data, 35
ncScaffold, 35
NewLimmaFile *(oneChannelGUI)*, 38
ngsperlDownload *(oneChannelGUI)*, 38
NGSreformat, 36
NormalizeNow *(oneChannelGUI)*, 38
normBoxplot, 36

ocPlotHist, 36
ocPlotPCA, 37
oneChannelGUI, 38
oneChannelGUI *(oneChannelGUI-package)*, 37
oneChannelGUI-package, 37
oneChannelGUI.start *(oneChannelGUI)*, 38
oneChannelGUIHelp *(oneChannelGUI)*, 38
OpenAFile *(oneChannelGUI)*, 38
OpenAGeoFile *(oneChannelGUI)*, 38
OpenALimmaFile *(oneChannelGUI)*, 38
OpenBeadStudioFiles, 40
OpenCDFandTargetsfiles *(oneChannelGUI)*, 38
OpenExonandTargetsfiles *(oneChannelGUI)*, 38
OpenExonFile *(oneChannelGUI)*, 38
OpenExonProbeFile *(oneChannelGUI)*, 38
OpenGeoFile *(oneChannelGUI)*, 38
OpenGeoFiles *(oneChannelGUI)*, 38
OpenLargeFile *(oneChannelGUI)*, 38
OpenLargefiles, 40
OpenLimmaFile *(oneChannelGUI)*, 38
OpenmiRanalyserfiles *(oneChannelGUI)*, 38
OpenmiRExpress *(oneChannelGUI)*, 38
OpenmiRProfiles *(oneChannelGUI)*, 38
OpenmiRShrimp *(oneChannelGUI)*, 38
OpenNgsfiles *(oneChannelGUI)*, 38

pamrHelp *(oneChannelGUI)*, 38
pdmclassHelp *(oneChannelGUI)*, 38
plierToZero, 41
plotGO, 41
PlotOptionsv1, 42
plotVariantSI, 42

qcMDS, 43

raex.annotation, 43
raex.variantexons, 44
rankingConversion, 44
rankProd, 45
RankProdHelp *(oneChannelGUI)*, 38
rawBoxplotPN, 45
rawpCheck, 46
Reads2logos, 46
recoverUnfiltered
(oneChannelGUI), 38
refiningPeaks, 47
reformatGdl, 47
refseqDownload, 47
retrieveMirTargets, 48
retrievePSRseq, 48
retrievePSRseq1gid, 48
reviqrFilter, 49
RmiRInterface, 49
rnfa, 50
runningJetta, 50

sample.size.evaluation, 52
sample.size.evaluation1, 51
SaveAsLimmaFile *(oneChannelGUI)*,
38
SetED *(oneChannelGUI)*, 38
showDataset, 52
showTopTable, 53
siggenes, 53
siggenesHelp *(oneChannelGUI)*, 38
simFilter, 54
sizepowerHelp *(oneChannelGUI)*, 38
spliceIndex, 54
ssizeHelp *(oneChannelGUI)*, 38
standAloneAddingAnnotation, 55
standAloneBuildingLocalAnnotation,
55
startmeV *(oneChannelGUI)*, 38

targetWidget, 56
templA, 56
trainTest, 57

updateLibs, 57

variantExons, 58
variantSI, 58
VennDiagram, 59

whichKindOfArray *(oneChannelGUI)*,
38