Package ‘miRNAtap’

March 22, 2017

Type Package

Title miRNAtap: microRNA Targets - Aggregated Predictions

Version 1.8.0

Date 2016-10-03

Author Maciej Pajak, T. Ian Simpson

Maintainer Maciej Pajak <m.pajak@sms.ed.ac.uk>

Description The package facilitates implementation of workflows requiring miRNA predictions, it allows to integrate ranked miRNA target predictions from multiple sources available online and aggregate them with various methods which improves quality of predictions above any of the single sources. Currently predictions are available for Homo sapiens, Mus musculus and Rattus norvegicus (the last one through homology translation).

License GPL-2

Depends R (>= 3.3.0), AnnotationDbi

Imports DBI, RSQLite, stringr, sqldf, plyr, methods

Suggests topGO, org.Hs.eg.db, miRNAtap.db, testthat

biocViews Software, Classification, Microarray, Sequencing, miRNA

Roxygen list(wrap = FALSE)

RoxygenNote 5.0.1

NeedsCompilation no

R topics documented:

aggregateRanks ......................................................... 2
getPredictedTargets .................................................... 3
getTargetsFromSource .................................................. 4
MirnaDb-class .......................................................... 5
miRNAtap ................................................................. 6
translate ................................................................. 7

Index 8
aggregateRanks

Aggregated ranks from multiple sources with various methods

Description

This function performs aggregation phase of target prediction for `getPredictedTargets`. Consensus ranking is derived from multiple individual rankings. Available methods include minimum, maximum and geometric mean with further tuning parameters which promote true positives at the top of the final ranking.

Usage

```r
aggregateRanks(ranks, n_valid_srcs, min_src, method = "geom", promote = TRUE)
```

Arguments

- `ranks`: data.frame with ordered scores
- `n_valid_srcs`: number of valid sources in the dataset
- `min_src`: minimum acceptable number of sources
- `method`: "min", "max", or "geom", default "geom"
- `promote`: add weights to improve accuracy of the method, default `TRUE`

Value

data.frame object with ranks per source and aggregate ranks

Author(s)

Maciej Pajak <m.pajak@sms.ed.ac.uk>

Examples

```r
data = data.frame(GeneID=c("15364", "56520", "57781", "58180", "18035"),
                  source1scores=c(0.9,0.5,0.3,NA,NA),
                  source2scores=c(0.7,NA,0.8,0.6,0.5),
                  source3scores=c(0.5,NA,0.3,0.1,0.2))
data # dataframe with scores
aggregateRanks(data, n_valid_srcs=3, min_src=2, method="geom")
# note how gene 56520 is eliminated as it appeared in fewer than 2 sources
```
getPredictedTargets  

Description

This method performs aggregation of target lists from multiple sources. Aggregated list is more accurate than any list from a single source. Multiple aggregation methods are available. Direct target data from five sources for Human and Mouse is supplied through miRNAtap.db package, for Rat targets are derived through homology translations whenever direct ones are not available.

Usage

getPredictedTargets(mirna, sources = c("pictar", "diana", "targetscan", "miranda", "mirdb"), species = "mmu", min_src = 2, method = "geom", promote = TRUE, synonyms = TRUE, both_strands = FALSE, ...)

Arguments

mirna  miRNA in a standard format
sources a list of sources to use for aggregation, default is all five sources, i.e. c('pictar','diana','targetscan','miranda','mirdb')
species species in a standard three-letter acronym, 'mmu' and 'hsa' available as direct targets, 'rno' as homology translations, default 'mmu'
min_src minimum number of sources required for a target to be considered, default 2
method method of aggregation - choose from 'min', 'max', and 'geom': 'min' is a minimum of ranks, 'max' is a maximum of ranks, and default 'geom' is based on geometric mean of the ranks which proves to be the most accurate method.
promote add weights to improve accuracy of the method, default TRUE
synonyms when searching for -3p miRNA automatically also searches for miRNA with the same name but ending with * (some databases list -3p miRNA this way) and other way around, similarly for -5p miRNA, default TRUE
both_strands overrides synonyms and searches for targets of both -5p and -3p strands together
... any optional arguments

Details

Tuning min_src parameter is an easy way of prioritising precision at the top of the list (high values) or total recall (low values). For the five default input sources, recommended values are 2, 3, or 4.

Value

data.frame object where row names are entrez IDs of target genes, ranks from individual sources and aggregated rank are shown in columns. If no targets are found in any of the sources NULL and a warning are returned.

Author(s)

Maciej Pajak <m.pajak@sms.ed.ac.uk>
getTargetsFromSource

Get target list from a single source

Description
This function queries precompiled annotation SQLite database which contains miRNA - target gene associations with their respective scores.

Usage
getTargetsFromSource(mirna, species = "mmu", source = "diana", synonyms = TRUE, both_strands = FALSE)

Arguments

- **mirna**: miRNA in a standard format
- **species**: species in a standard three-letter acronym, default 'mmu'
- **source**: a source target prediction algorithm table to query, default 'diana', other possible values are 'miranda', 'targetscan', and 'pictar'.
- **synonyms**: when searching for -3p miRNA automatically also searches for miRNA with the same name but ending with * (some databases list -3p miRNA this way) and other way around, similarly for -5p miRNA, default TRUE
- **both_strands**: overrides synonyms and searches for targets of both -5p and -3p strands together

Value
data.frame object with entrez IDs of target genes and their scores, if there are no targets found for a given miRNA in a given table then an empty
Author(s)

Maciej Pajak <m.pajak@sms.ed.ac.uk>

References


Examples

```r
targets <- getTargetsFromSource('let-7a', species='hsa', source='targetscan')
head(targets)
# top of the list of human targets of let-7a from TargetScan only
```

Description

Object of `MirnaDb` class holds the SQLite database connection, and extends `AnnotationDb` class from `AnnotationDbi` package. `columns`, `keys`, `keytypes` and `select` methods allow access to database tables and retrieval of miRNA target information.

`select` is the most important method, allows querying the database for predictions from a specific source and species for a given miRNA.

Usage

```r
columns(x)
keytypes(x)
keys(x, keytype, ...)
select(x, keys, columns, keytype, ...)
```

```r
## S4 method for signature 'MirnaDb'
columns(x)
```

```r
## S4 method for signature 'MirnaDb'
keytypes(x)
```

```r
## S4 method for signature 'MirnaDb'
keys(x, keytype, ...)
```

```r
## S4 method for signature 'MirnaDb'
select(x, keys, columns, keytype, ...)
```
Arguments

- **x**: the `MirnaDb` object
- **keytype**: the keytype that matches the keys used; the table in which the search should be performed.
- **keys**: the key to select records for from the database - miRNA name; all possible keys (miRNAs) are returned by using the `keys` method.
- **columns**: in this case same as keytype, the table in which the search should be performed, this value specifies the source of predictions as well as species; as with keys, all possible columns are returned by using the `columns` method.

Value

string vectors, for select a data.frame with target genes and scores

Author(s)

Maciej Pajak <m.pajak@sms.ed.ac.uk>, Ian Simpson

Examples

```r
#first load the annotations
require(miRNAtap.db)
#see all available tables
keytypes(miRNAtap.db)
```

miRNAtap: microRNA Targets - Aggregated Predictions.

Description

It is a package with tools to facilitate implementation of workflows requiring miRNA prediction through access to multiple prediction results (DIANA, Targetscan, PicTar, Miranda, and miRDB) and their aggregation. Three aggregation methods are available: minimum, maximum and geometric mean, additional parameters provide further tuning of the results. Predictions are available for Homo sapiens, Mus musculus and Rattus norvegicus (the last one through homology translation).

Author(s)

Maciej Pajak <m.pajak@sms.ed.ac.uk>, Ian Simpson

Examples

```r
#direct targets in mouse aggregated from all sources:
targets_mouse <- getPredictedTargets('let-7a', species='mmu', method='geom')
#homology-translated targets in rat aggregated from all sources
targets_rat <- getPredictedTargets('let-7a', species='rno', method='geom')
```
**translate**

*Homology transfer for miRNAtap*

**Description**

This function maps gene entrez ID between species using homology information from Homologene.

**Usage**

```r
translate(entrezes, from = "mmu", to = "rno", ...)
```

**Arguments**

- `entrezes`: data.frame with entrez Gene IDs and their scores
- `from`: origin species, default 'mmu', Mus musculus
- `to`: target species, default
- `...`: any optional arguments

**Value**

data.frame object with orthologous genes’ entrez IDs and corresponding scores

**Author(s)**

Maciej Pajak <m.pajak@sms.ed.ac.uk>

**Examples**

```r
mouse_genes <- data.frame(GeneID = 
  c("15364", "56520", "57781", "58180", "18035", "239857"))
translate(mouse_genes, from='mmu', to='rno')
```
Index

.MirnaDb (MirnaDb-class), 5
aggregateRanks, 2

columns (MirnaDb-class), 5
columns, MirnaDb-method (MirnaDb-class), 5

getPredictedTargets, 2, 3
getTargetsFromSource, 4

keys (MirnaDb-class), 5
keys, MirnaDb-method (MirnaDb-class), 5
keytypes (MirnaDb-class), 5
keytypes, MirnaDb-method (MirnaDb-class), 5

MirnaDb (MirnaDb-class), 5
MirnaDb-class, 5

miRNAtap, 6
miRNAtap-package (miRNAtap), 6

select (MirnaDb-class), 5
select, MirnaDb-method (MirnaDb-class), 5

translate, 7