

Package ‘ChIPXpressData’

October 3, 2024

Type Package

Title ChIPXpress Pre-built Databases

Version 1.42.0

Date 2012-07-24

Author George Wu

Maintainer George Wu <gewu@jhsp.h.edu>

Description Contains pre-built mouse (GPL1261) and human (GPL570) database of gene expression profiles to be used for ChIPXpress ranking.

License GPL (>=2)

Depends bigmemory

biocViews Homo_sapiens_Data, Mus_musculus_Data, GEO

git_url <https://git.bioconductor.org/packages/ChIPXpressData>

git_branch RELEASE_3_19

git_last_commit afd6ff1

git_last_commit_date 2024-04-30

Repository Bioconductor 3.19

Date/Publication 2024-10-03

Contents

| | |
|----------------------------------|----------|
| ChIPXpressData-package | 2 |
| DB_GPL1261.bigmemory | 2 |
| DB_GPL570.bigmemory | 3 |
| GPL1261mean | 4 |
| GPL1261var | 5 |
| GPL570mean | 6 |
| GPL570var | 7 |
| Index | 8 |

ChIPXpressData-package

ChIPXpress Gene Expression Databases

Description

Pre-built databases of gene expression profiles for ChIPXpress analysis in big.matrix format. DB_GPL1261 contains mouse data. DB_GPL570 contains human data.

Details

Package: ChIPXpressData
Type: Package
Version: 1.00.0
Date: 2012-07-24
License: GPL 2.0

Author(s)

George Wu Maintainer: George Wu <gewu@jhsph.edu>

References

McCall M.N., Bolstad B.M., and Irizarry R.A. (2010) Frozen robust multiarray analysis (fRMA). *Biostatistics* 11, 242-253.

Barrett T., et al. (2007) NCBI GEO: mining tens of millions of expression profiles - database and tools update. *Nucl. Acids Res.* 35, D760-D765.

DB_GPL1261.bigmemory *Database of gene expression profiles from the Affymetrix Mouse 430 2.0 Array (GPL1261) in big.matrix format*

Description

The data set contains 9634 mouse profiles downloaded from NCBI GEO, processed using fRMA, and normalized. It is in big.matrix format.

Format

The format is: Formal class 'big.matrix' [package "bigmemory"] with 1 slots ..@ address:<externalptr>

Details

The database is formatted as a `big.matrix` for more efficient loading into memory. It is stored in `DB_GPL1261.bigmemory` and the corresponding description file is `DB_GPL1261.bigmemory.desc`. To utilize the `big.matrix` format, it requires the package `bigmemory` to be loaded. See the `bigmemory` package for more information.

The database contains 20757 rows and 9643 columns. Each row represents the expression vector for each gene and each column represents the gene expression measurements for a sample from NCBI GEO obtained using the GPL1261 platform. Each gene will match uniquely to a single probe ID; only the probe with the highest variance in the compendium apriori to normalization are retained as the representative measurement for each gene.

Source

www.ncbi.nlm.nih.gov/geo/

References

McCall M.N., Bolstad B.M., and Irizarry R.A. (2010) Frozen robust multiarray analysis (fRMA). *Biostatistics* 11, 242-253.

Barrett T., et al. (2007) NCBI GEO: mining tens of millions of expression profiles - database and tools update. *Nucl. Acids Res.* 35, D760-D765.

Examples

```
## Load the GPL1261 database
library(bigmemory)
path <- system.file("extdata", package="ChIPXpressData")
DB_GPL1261 <- attach.big.matrix("DB_GPL1261.bigmemory.desc", path=path)
## DB_GPL1261 is then ready for input into the ChIPXpress function.

## To see info about the database matrix
describe(DB_GPL1261)
```

| | |
|---------------------|---|
| DB_GPL570.bigmemory | <i>Database of gene expression profiles from the Affymetrix Human U133 Plus 2.0 array (GPL570) in big.matrix format</i> |
|---------------------|---|

Description

The data set contains 18257 human profiles downloaded from NCBI GEO, processed using fRMA, and normalized. It is in `big.matrix` format.

Format

The format is: Formal class 'big.matrix' [package "bigmemory"] with 1 slots ..@ address:<externalptr>

Details

The database is formatted as a `big.matrix` for more efficient loading into memory. It is stored in `DB_GPL570.bigmemory` and the corresponding description file is `DB_GPL570.bigmemory.desc`. To utilize the `big.matrix` format, it requires the package `bigmemory` to be loaded. See the `bigmemory` package for more information.

The database contains 19798 rows and 18257 columns. Each row represents the expression vector for each gene and each column represents the gene expression measurements for a sample from NCBI GEO obtained using the GPL570 platform. Each gene will match uniquely to a single probe ID; only the probe with the highest variance in the compendium apriori to normalization are retained as the representative measurement for each gene.

Source

www.ncbi.nlm.nih.gov/geo/

References

McCall M.N., Bolstad B.M., and Irizarry R.A. (2010) Frozen robust multiarray analysis (fRMA). *Biostatistics* 11, 242-253.

Barrett T., et al. (2007) NCBI GEO: mining tens of millions of expression profiles - database and tools update. *Nucl. Acids Res.* 35, D760-D765.

Examples

```
## Load the GPL570 database
library(bigmemory)
path <- system.file("extdata", package="ChIPXpressData")
DB_GPL570 <- attach.big.matrix("DB_GPL570.bigmemory.desc", path=path)
## DB_GPL570 is then ready for input into the ChIPXpress function.

## To see info about the database matrix
describe(DB_GPL570)
```

GPL1261mean

Mean of each probeset in the GPL1261 database

Description

Mean across all samples for each probe set in the GPL1261 compendium prior to standardization

Usage

```
data(GPL1261mean)
```

Format

The format is: Named num [1:20757] 10.2 10.88 7.77 8.87 10.92 ... - attr(*, "names")= chr [1:20757] "11972" "57437" "100678" "13481" ...

Details

Used to check for low expression probesets

Source

www.ncbi.nlm.nih.gov/geo/

References

Barrett T., et al. (2007) NCBI GEO: mining tens of millions of expression profiles - database and tools update. *Nucl. Acids Res.* **35**, D760-D765.

Examples

```
data(GPL1261mean)
```

GPL1261var

Variance of each probeset in the GPL1261 database

Description

Variance across all samples for each probe set in the GPL1261 compendium prior to standardization

Usage

```
data(GPL1261var)
```

Format

The format is: Named num [1:20757] 0.773 0.509 1.817 0.326 0.568 ... - attr(*, "names")= chr [1:20757] "11972" "57437" "100678" "13481" ...

Details

Used to check for low variance probesets

Source

www.ncbi.nlm.nih.gov/geo/

References

Barrett T., et al. (2007) NCBI GEO: mining tens of millions of expression profiles - database and tools update. *Nucl. Acids Res.* **35**, D760-D765.

Examples

```
data(GPL1261var)
```

GPL570mean

Mean of each probeset in the GPL570 database

Description

Mean across all samples for each probe set in the GPL570 compendium prior to standardization

Usage

```
data(GPL570mean)
```

Format

The format is: Named num [1:19944] -9.25e-16 -6.07e-16 5.80e-16 -3.01e-16 4.10e-16 ... - attr(*, "names")= chr [1:19944] "112597" "203102" "91937" "266675" ...

Details

Used to check for low expression probesets

Source

www.ncbi.nlm.nih.gov/geo/

References

Barrett T., et al. (2007) NCBI GEO: mining tens of millions of expression profiles - database and tools update. *Nucl. Acids Res.* **35**, D760-D765.

Examples

```
data(GPL570mean)
```

`GPL570var`*Variance of each probeset in the GPL570 database*

Description

Variance across all samples for each probeset in the GPL570 compendium prior to standardization

Usage

```
data(GPL570var)
```

Format

The format is: Named num [1:19944] 0.177 0.0892 0.5552 0.4545 1.0399 ... - attr(*, "names")= chr [1:19944] "112597" "203102" "91937" "266675" ...

Details

Used to check for low variance probesets

Source

www.ncbi.nlm.nih.gov/geo/

References

Barrett T., et al. (2007) NCBI GEO: mining tens of millions of expression profiles - database and tools update. *Nucl. Acids Res.* **35**, D760-D765.

Examples

```
data(GPL570var)
```

Index

- * **datasets,GPL1261,database**
 - DB_GPL1261.bigmemory, [2](#)
 - * **datasets,GPL570,database**
 - DB_GPL570.bigmemory, [3](#)
 - * **datasets**
 - GPL1261mean, [4](#)
 - GPL1261var, [5](#)
 - GPL570mean, [6](#)
 - GPL570var, [7](#)
 - * **package, database, ChIPXpress**
 - ChIPXpressData-package, [2](#)
- ChIPXpressData
(ChIPXpressData-package), [2](#)
- ChIPXpressData-package, [2](#)
- DB_GPL1261.bigmemory, [2](#)
- DB_GPL570.bigmemory, [3](#)
- GPL1261mean, [4](#)
- GPL1261var, [5](#)
- GPL570mean, [6](#)
- GPL570var, [7](#)