Package ‘Affyhgu133aExpr’

April 17, 2025

Type Package

Title Affymetrix Human hgu133a Array (GPL96) Expression Data Package

Version 1.47.0

Date 2014-4-9

Author Zhicheng Ji, Hongkai Ji

Maintainer Zhicheng Ji <`zji4@jhu.edu`>

Description Contains pre-built human (GPL96) database of gene expression profiles. The gene expression data was downloaded from NCBI GEO, preprocessed and normalized consistently. The biological context of each sample was recorded and manually verified based on the sample description in GEO.

License GPL (>=2)

Depends R (>= 2.10)

biocViews Homo_sapiens_Data, GEO

git_url https://git.bioconductor.org/packages/Affyhgu133aExpr

git_branch devel

git_last_commit e2ea0d3

git_last_commit_date 2025-04-15

Repository Bioconductor 3.22

Date/Publication 2025-04-17

Contents

`Affyhgu133aExpr-package` *Affymetrix Human hgu133a Array (GPL96) Expression Data Package*

Description

Contains gene expression profiles from Affymetrix Human hgu133a Array (GPL96). This package is primarily designed for GSCA (Gene Set Context Analysis). All gene expression data are downloaded from NCBI GEO. Gene expression data were preprocessed and normalized consistently using fRMA. Gene expression values are further standardized using gene expression barcode approach. The biological context of each sample was recorded and manually verified based on the sample description in GEO. Gene expression profiles are stored as hdf5 format.

Details

Package:	Affyhgu133aExpr
Type:	Package
Version:	1.0.0
Date:	2014-4-9
License:	GPL 2.0

Author(s)

Author: Zhicheng Ji, Hongkai Ji Maintainer: Zhicheng Ji <zji4@jhu.edu>

References

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

McCall, M. N., Uppal, K., Jaffee, H. A., Zilliox, M. J., & Irizarry, R. A. (2011). The Gene Expression Barcode: leveraging public data repositories to begin cataloging the human and murine transcriptomes. Nucleic acids research, 39(suppl 1), D1011-D1015.

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

`Affyhgu133aExpr` *Data of human gene expression profiles from the Affymetrix Human hgu133a Array (GPL96).*

Description

The data set contains 11778 human profiles on 12495 genes downloaded from NCBI GEO. Gene expression data were preprocessed and normalized consistently using fRMA. Gene expression values are further standardized using gene expression barcode approach. The biological context of each sample was recorded and manually verified based on the sample description in GEO. The gene expression value matrix is stored in hdf5 format using rhdf5 package.

Details

This data package contains expression values of 12495 genes and 11778 samples measurements 11778 samples measurements from NCBI GEO obtained using the GPL96 platform. Gene expression data were preprocessed and normalized consistently using fRMA. Gene expression values were further standardized using gene expression barcode approach. Probeset with the largest coefficient of variaction from all probesets that corresponds to the same gene is retained, so that each gene uniquely matches to one row in the database. The biological context of each sample was also recorded and manually verified based on the sample description in GEO. To enhance the reading speed of the dataset, the gene expression value matrix is stored in a hdf5 format using rhdf5 package. The rows of the matrix represents samples and the columns of the matrix represent genes. Notice that all values are 1000 times the actual value so that the values can be stored as integers to minimize file size and reading time. The package is specifically designed to be manipulated by GSCA package so users are not expected to read the expression values by themselves. The sample id, sample type, and experiment id for each sample in the gene expression compendium are also included in this data package.

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.

McCall, M. N., Uppal, K., Jaffee, H. A., Zilliox, M. J., & Irizarry, R. A. (2011). The Gene Expression Barcode: leveraging public data repositories to begin cataloging the human and murine transcriptomes. Nucleic acids research, 39(suppl 1), D1011-D1015.

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

`Affyhgu133aExprtab` *Reference table for Affyhgu133aExpr gene expression compendium*

Description

Contains the sample id, sample type, and experiment id for each sample in the Affymetrix Human hgu133a Array (GPL96) gene expression compendium.

Usage

```
data(Affyhgu133aExprtab)
```

Format

A data frame with 11778 observations on the following 3 variables.

`SampleID`	a character vector
`ExperimentID`	a character vector
`SampleType`	a character vector

Details

SampleID is a GSM ID that NCBI GEO uses to as a sample identifier. ExperimentID is a GEO ID that NCBI GEO uses to identify an experiment. SampleType denotes the cell type or tissue and whether the sample is given a specific treatment or in a specific condition.

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.

McCall, M. N., Uppal, K., Jaffee, H. A., Zilliox, M. J., & Irizarry, R. A. (2011). The Gene Expression Barcode: leveraging public data repositories to begin cataloging the human and murine transcriptomes. Nucleic acids research, 39(suppl 1), D1011-D1015.

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 reference table
data(Affyhgu133aExprtab)
str(Affyhgu133aExprtab)
```

Index