

# Package ‘Affymoe430Expr’

February 28, 2014

**Type** Package

**Title** Affymetrix Mouse 430 2.0 Array (GPL1261) Expression Data Package

**Version** 0.99.0

**Date** 2014-2-6

**Author** Zhicheng Ji, Hongkai Ji

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

**Description** Contains pre-built mouse (GPL1261) databases of gene expression profiles. The gene expression data was downloaded from NCBI GEO and preprocessed and normalized consistently using fRMA. 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** AffymetrixChIP, Mus\_musculus

## R topics documented:

Affymoe430Expr-package	2
Affymoe430Expr	2
Affymoe430Exprtab	4

<b>Index</b>	<b>6</b>
--------------	----------

Affymoe430Expr-package

*Affymetrix Mouse 430 2.0 Array (GPL1261) Expression Data Package*

---

**Description**

Contains gene expression profiles from Affymetrix Mouse 430 2.0 Array (GPL1261). 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. The biological context of each sample was recorded and manually verified based on the sample description in GEO. Gene expression profiles are stored as vectors in separate .rda files (with xz compression).

**Details**

Package: Affymoe430Expr

Type: Package

Version: 0.99.0

Date: 2014-2-6

License: GPL 2.0

**Author(s)**

Author: Zhicheng Ji, Hongkai Ji Maintainer: Zhicheng Ji &lt;zji4@jhu.edu&gt;

**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.

---

Affymoe430Expr*Data of human gene expression profiles from the Affymetrix Mouse 430 2.0 Array (GPL1261).*

---

**Description**

The data set contains 9643 mouse profiles on 20877 genes downloaded from NCBI GEO. Gene expression data were preprocessed and normalized consistently using fRMA. The biological context of each sample was recorded and manually verified based on the sample description in GEO.

Expression profiles of the 20877 genes are stored as vectors of length 9643 in separate .rda files (with xz compression).

## Details

This data package contains 20877 .rda files with xz compression. Each .rda file contains the gene expression vector of length 9643 for the gene ID represented by its file name, and the vector consists of 9643 samples measurements from NCBI GEO obtained using the GPL1261 platform. After the data is processed and normalized consistently with frma, all probesets that corresponds to the same gene are averaged, such 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. The sample name, 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/](http://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

```
## GSCA will automatically read in the necessary gene expression data according to given input.
## Thus, normally users do not need to read in the data themselves.

## To manually read in gene expression data, for example Entrez GeneID 11287:
geneid <- 11287
path <- system.file("extdata",package="Affymoe430Expr")
load(paste0(path,"/",geneid,".rda"))
## The expression vector named "Expr" is read in

## To read in gene expression data of multiple Entrez GeneID:
## This function represents select() interface in most Bioconductor annotation packages
moe430search <- function(geneid) {
  exprdat <- matrix(0,nrow=length(geneid),ncol=9643)
  path <- system.file("extdata",package="Affymoe430Expr")
  for (i in 1:length(geneid)) {
    load(paste0(path,"/",geneid[i],".rda"))
    exprdat[i,] <- Expr
  }
  return(exprdat)
}
## Run the function on desired genes, for example Entrez GeneID 11287,11298,11302,11303
Exprgeneset <- moe430search(c(11287,11298,11302,11303))
str(Exprgeneset)
```

```
## Which returns a matrix of 4 rows and 9643 columns containing the expression profiles of the 4 genes

## To check all Entrez GeneID included in the data package:
path <- system.file("extdata",package="Affymoe430Expr")
sub(".rda","",list.files(path))

## To check the reference table which contains
## sample name, sample id, sample type, and experiment id for each sample
data(Affymoe430Exprtab)
str(Affymoe430Exprtab)

## Add sample name to expression data
colnames(Exprgeneset) <- Affymoe430Exprtab$SampleName
```

---

Affymoe430Exprtab      *Reference table for Affymoe430 gene expression compendium*

---

### Description

Contains the sample name, sample id, sample type, and experiment id for each sample in the Affymetrix Mouse 430 2.0 Array (GPL1261) gene expression compendium.

### Usage

```
data(Affymoe430Exprtab)
```

### Format

A data frame with 9643 observations on the following 4 variables.

SampleName a character vector

SampleID a character vector

ExperimentID a character vector

SampleType a character vector

### Details

The sample name is a .CEL.gz file name for the downloaded source file. 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/](http://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 reference table
data(Affymoe430Exprtab)
str(Affymoe430Exprtab)
```

# Index

\*Topic **datasets, Affymoe430Exprtab**

Affymoe430Exprtab, [4](#)

\*Topic

**datasets,GPL1261,database,Affymoe430Expr**

Affymoe430Expr, [2](#)

\*Topic **package, database,  
Affymoe430Expr**

Affymoe430Expr-package, [2](#)

Affymoe430Expr, [2](#)

Affymoe430Expr-package, [2](#)

Affymoe430Exprtab, [4](#)