

# Package ‘cnvGSAdata’

October 18, 2017

**Title** Data used in the vignette of the cnvGSA package

**Version** 1.12.0

**Author** Joseph Lugo <joseph.r.lugo@gmail.com>

**Description** This package contains the data used in the vignette of the cnvGSA package.

**Maintainer** Joseph Lugo <joseph.r.lugo@gmail.com>

**License** LGPL

**biocViews** ExperimentData, Genome, CopyNumberVariationData

**Depends** R (>= 2.10), cnvGSA

**NeedsCompilation** no

## R topics documented:

cnvGSAdata-package . . . . .	1
cnvGSA_input_example.RData . . . . .	2
cnvGSA_output_example.RData . . . . .	2
cnv_AGP_demo.txt . . . . .	3
enrGeneric_AGP_demo.txt . . . . .	4
enrGMT_AGP_demo.gmt . . . . .	4
gene_ID_demo.txt . . . . .	5
gs_data_example.RData . . . . .	5
kl_gene_AGP_demo.txt . . . . .	6
kl_loci_AGP_demo.txt . . . . .	6
ph_AGP_demo.txt . . . . .	7
<b>Index</b>	<b>8</b>

---

cnvGSAdata-package     *Data used in the examples and vignette of the cnvGSA package*

---

## Description

This package contains data used in the examples and vignette of the cnvGSA package.

**Details**

Raw input files: [cnv\\_AGP\\_demo.txt](#) - CNV data [enrGeneric\\_AGP\\_demo.txt](#) - enrichment generic data [enrGMT\\_AGP\\_demo.gmt](#) - enrichment GMT data [kl\\_gene\\_AGP\\_demo.txt](#) - genes of interest data [kl\\_loci\\_AGP\\_demo.txt](#) - known loci data [ph\\_AGP\\_demo.txt](#) - phenotype/covariate data [gene\\_ID\\_demo.txt](#) - gene ID file

Pre-built input: [cnvGSA\\_input\\_example.RData](#)

Saved output: [cnvGSA\\_output\\_example.RData](#)

Pre-built Gene-set data: [gs\\_data\\_example.RData](#)

**Author(s)**

Joseph Lugo <[joseph.r.lugo@gmail.com](mailto:joseph.r.lugo@gmail.com)>

---

cnvGSA\_input\_example.RData  
*cnvGSA example input*

---

**Description**

Pre-built input used in the full workflow example in the `cnvGSA` vignette.

**Usage**

```
data("cnvGSA_input_example")
```

**Details**

The dataset contains `cnvGSA.in`, a single object of class `CnvGSAInput` as defined in the `cnvGSA` package. When processed by `cnvGSAlogRegTest()` – the main function in that package – it produces as its output an object of class `CnvGSAOutput` (such as the one stored in `cnvGSA_output_example.RData`).

---

cnvGSA\_output\_example.RData  
*cnvGSA example output*

---

**Description**

Saved output from the full workflow example in the `cnvGSA` vignette.

**Usage**

```
data("cnvGSA_output_example")
```

## Details

The dataset contains `cnvGSA.out`, a single object of class `CnvGSAOutput` as defined in the `cnvGSA` package. `CnvGSAOutput` is a simple S4 class containing a slot for each data structure output by `cnvGSAlogRegTest` (the main function in the package):

```
res.ls
gsTables.ls
gsData.ls
phData.ls
config.df
```

`res.ls` contains the output from the regression tests, `gsTables.ls` contains the the gene-set tables, `gsData.ls` contains the gene-set data needed for other scripts, `phData.ls` contains the phenotype/covariate data and `config.df` contains the config data frame that allows the other scripts to read in the params.

See the `cnvGSA` vignette for complete details and discussion.

---

cnv\_AGP\_demo.txt

*Rare CNV data from Pinto et al. 2014 ASD study*

---

## Description

The file contains rare CNV data similar to that used in the Pinto et al. 2014 ASD study.

## Details

The data is stored in a similar format as the Genome Variation Format <http://www.sequenceontology.org/resources/gvf.html>. See the user manual for more information.

## Source

Pinto, D et al. Convergence of Genes and Cellular Pathways Dysregulated in Autism Spectrum Disorders. *Am J Hum Genet.* 2014 May 1; 94(5): 677–694.

## Examples

```
cnvFile <- system.file( "extdata", "cnv_AGP_demo.txt", package="cnvGSAdata" )
cnv.df <- read.table( cnvFile, header = TRUE, sep = "\t", quote = "\"", stringsAsFactors = FALSE)
```

---

```
enrGeneric_AGP_demo.txt
```

*Generic file for Enrichment Map with data from Pinto et al. 2014 ASD study*

---

### Description

The file contains gene-sets with their FDR and p-values similar to that used in the Pinto et al. 2014 ASD study.

### Details

The data is stored in a generic file format which can be seen here [http://www.baderlab.org/Software/EnrichmentMap/UserManual#Generic\\_results\\_files](http://www.baderlab.org/Software/EnrichmentMap/UserManual#Generic_results_files)

### Source

Pinto, D et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*. 2010 Jul 15; 466(7304): 368–72.

### Examples

```
enrFile      <- system.file( "extdata", "enrGeneric_AGP_demo.txt", package="cnvGSadata" )
enrGeneric.df <- read.table( enrFile, header = TRUE, sep = "\t", quote = "\"", stringsAsFactors = FALSE)
```

---

```
enrGMT_AGP_demo.gmt
```

*Gene-set data from Pinto et al. 2014 ASD study*

---

### Description

The file contains gene-set data similar to that used in the Pinto et al. 2014 ASD study.

### Details

The data is stored in the Gene Matrix Transposed format [http://www.baderlab.org/Software/EnrichmentMap/UserManual#Gene\\_sets\\_file\\_.28GMT\\_file.29](http://www.baderlab.org/Software/EnrichmentMap/UserManual#Gene_sets_file_.28GMT_file.29).

### Source

Pinto, D et al. Convergence of Genes and Cellular Pathways Dysregulated in Autism Spectrum Disorders. *Am J Hum Genet*. 2014 May 1; 94(5): 677–694.

### Examples

```
gmtFile <- system.file( "extdata", "enrGMT_AGP_demo.gmt", package="cnvGSadata" )
no_col  <- max(count.fields(gmtFile, sep = "\t"))
gmt.df  <- read.table(gmtFile, sep="\t", fill=TRUE, col.names=1:no_col)
```

---

gene_ID_demo.txt	<i>Gene System data</i>
------------------	-------------------------

---

**Description**

The file contains Entrez gene ID's

**Details**

The data is stored in a format that look like (geneID) -tab- (Symbol) -tab- (Name).

**Examples**

```
geneIDFile <- system.file( "extdata", "gene_ID_demo.txt", package="cnvGSAdata" )
geneID.df <- read.table (geneIDFile, header = TRUE, sep = "\t", quote = "\"", stringsAsFactors = FALSE)
```

---

gs_data_example.RData	<i>gsData example output</i>
-----------------------	------------------------------

---

**Description**

Saved gsData described in the cnvGSA vignette.

**Usage**

```
data("gs_data_example")
```

**Details**

The dataset contains gsData. It contains the gene-sets analyzed in the cnvGSA vignette as well as the gene set names.

gs\_all.ls

gsid2name.chv

gs\_all.ls contains the gene-sets used in the analysis and gsid2name.chv contains the gene-set names.

See the cnvGSA vignette for complete details and discussion.

---

kl\_gene\_AGP\_demo.txt *Genes of interest used in the Pinto et al. 2014 ASD study*

---

**Description**

Contains the genes of interest that were also used in the Pinto et al. 2014 ASD study.

**Details**

This file is used to provide the genes of interest that will be looked at in the testing.

**Source**

Pinto, D et al. Convergence of Genes and Cellular Pathways Dysregulated in Autism Spectrum Disorders. *Am J Hum Genet.* 2014 May 1; 94(5): 677–694.

**Examples**

```
klGeneFile <- system.file( "extdata", "kl_gene_AGP_demo.txt", package="cnvGSadata" )
kl_gene.df <- read.table( klGeneFile, header = TRUE, sep = "\t", quote = "\"", stringsAsFactors = FALSE)
```

---

kl\_loci\_AGP\_demo.txt *Known loci of genes in the Pinto et al. 2014 ASD study*

---

**Description**

Contains the known loci of genes as in the Pinto et al. 2014 ASD study.

**Details**

This file is used to provide the loci of certain genes and this will affect the results from the test.

**Source**

Pinto, D et al. Convergence of Genes and Cellular Pathways Dysregulated in Autism Spectrum Disorders. *Am J Hum Genet.* 2014 May 1; 94(5): 677–694.

**Examples**

```
klLociFile <- system.file( "extdata", "kl_loci_AGP_demo.txt", package="cnvGSadata" )
kl_loci.df <- read.table( klLociFile, header = TRUE, sep = "\t", quote = "\"", stringsAsFactors = FALSE)
```

---

ph\_AGP\_demo.txt

*Covariate/phenotype data from Pinto et al. 2014 ASD study*

---

### **Description**

The file contains covariate/phenotype data similar to that used in the Pinto et al. 2014 ASD study.

### **Details**

The data is stored in a format that is a subset of the the CNV data. It includes all the covariates and phenotypes that the user wants.

### **Source**

Pinto, D et al. Convergence of Genes and Cellular Pathways Dysregulated in Autism Spectrum Disorders. *Am J Hum Genet.* 2014 May 1; 94(5): 677–694.

### **Examples**

```
phFile <- system.file( "extdata", "ph_AGP_demo.txt", package="cnvGSAdata" )
ph.df <- read.table (phFile, header = TRUE, sep = "\t", quote = "\"", stringsAsFactors = FALSE)
```

# Index

cnv\_AGP\_demo.txt, [2, 3](#)  
cnvGSA.in (cnvGSA\_input\_example.RData),  
[2](#)  
cnvGSA.out  
(cnvGSA\_output\_example.RData),  
[2](#)  
cnvGSA\_input\_example.RData, [2, 2](#)  
cnvGSA\_output\_example.RData, [2, 2](#)  
cnvGSAdata (cnvGSAdata-package), [1](#)  
cnvGSAdata-package, [1](#)  
  
enrGeneric\_AGP\_demo.txt, [2, 4](#)  
enrGMT\_AGP\_demo.gmt, [2, 4](#)  
  
gene\_ID\_demo.txt, [2, 5](#)  
gs\_data\_example.RData, [2, 5](#)  
gsData (gs\_data\_example.RData), [5](#)  
  
kl\_gene\_AGP\_demo.txt, [2, 6](#)  
kl\_loci\_AGP\_demo.txt, [2, 6](#)  
  
ph\_AGP\_demo.txt, [2, 7](#)