

# Package ‘coGPS’

February 28, 2025

**Type** Package

**Title** cancer outlier Gene Profile Sets

**Version** 1.51.0

**Date** 2011-10-20

**Author** Yingying Wei, Michael Ochs

**Maintainer** Yingying Wei <ywei@jhsph.edu>

**Description** Gene Set Enrichment Analysis of P-value based statistics  
for outlier gene detection in dataset merged from multiple  
studies

**Depends** R (>= 2.13.0)

**Suggests** limma

**Imports** graphics, grDevices

**License** GPL-2

**LazyLoad** yes

**biocViews** Microarray, DifferentialExpression

**git\_url** <https://git.bioconductor.org/packages/coGPS>

**git\_branch** devel

**git\_last\_commit** 2f7fc38

**git\_last\_commit\_date** 2024-10-29

**Repository** Bioconductor 3.21

**Date/Publication** 2025-02-28

## Contents

|                                   |   |
|-----------------------------------|---|
| coGPS-package . . . . .           | 2 |
| coGPS internal . . . . .          | 2 |
| PatientSpecificGeneList . . . . . | 2 |
| PCOPA . . . . .                   | 4 |
| permCOPA . . . . .                | 5 |
| PlotTopPCOPA . . . . .            | 7 |
| SampleData . . . . .              | 8 |



**Arguments**

|                         |  |
|-------------------------|--|
| <code>exprslist</code>  | Each element of <i>exprslist</i> is a list with the first element being <i>exprs</i> and the second element being <i>classlab</i> . Each row of <i>exprs</i> represents one gene and each column represents one sample. <i>classlab</i> is a zero-one vector indicating the status of samples. We use 0 for the baseline group, usually the normal group, and 1 for the comparison group, usually the tumor group. |
| <code>alpha</code>      | Significance level for P-value.  |
| <code>side</code>       | A vector specifying the definition of P-value in each of the study, which could be either <i>up</i> , <i>down</i> , or <i>twosided</i> .   |
| <code>type</code>       | A vector specifying whether the outlier pattern is <i>subtype</i> or <i>uniform</i> .  |
| <code>TopGeneNum</code> | a number specifying the top number of outlier genes scored by PCOPA to be included in the generation of individual outlier gene list for each patient.   |

**Value**

|                                    |   |
|------------------------------------|---|
| <code>outliergene_bypatient</code> | a list whose length equals the number of tumor samples (patients). each element of the list is a list of length equaling to the length of <i>exprslist</i> , in other words the number of studies(or data type), showing the outlier gene for each patient in each study (or data type) |
|------------------------------------|---|

**Author(s)**

Yingying Wei

**References**

Wei, Y., Hennessey, P., Gaykalova, D., Califano, J.A., Ochs, M.F., (2011) Cancer Outlier Gene Profile Sets Elucidate Pathways in Head and Neck Squamous Cell Carcinoma.

**Examples**

```
#read in data
data(Exon_exprs_matched)
data(Methy_exprs_matched)
data(CNV_exprs_matched)
data(Exon_classlab_matched)
data(Methy_classlab_matched)
data(CNV_classlab_matched)
head(Exon_exprs_matched)

#exprslist[[i]]$exprs should be in matrix format
Exon_exprs<-as.matrix(Exon_exprs_matched)
Methy_exprs<-as.matrix(Methy_exprs_matched)
CNV_exprs<-as.matrix(CNV_exprs_matched)

#exprslist[[i]]$classlab should be in vector format
Exon_classlab<-unlist(Exon_classlab_matched)
Methy_classlab<-unlist(Methy_classlab_matched)
```

```

CNV_classlab<-unlist(CNV_classlab_matched)

#make an exprslist consisting 3 studies
trylist<-list()
trylist[[1]]<-list(exprs=Exon_exprs,classlab=Exon_classlab)
trylist[[2]]<-list(exprs=Methy_exprs,classlab=Methy_classlab)
trylist[[3]]<-list(exprs=CNV_exprs,classlab=CNV_classlab)

#generate an outlier gene list for each patient restricted to the top PCOPA scored genes
IndividualList7<-PatientSpecificGeneList(trylist,0.05,side=c("up","down","up"),type="subtype",TopGeneNum=100)

```

---

PCOPA

*P-value based outlier gene detection*


---

### Description

Calculate P-value based statistics for outlier gene detection in dataset merged from multiple studies and give out outlier gene list for each patient.

### Usage

```
PCOPA(exprslist, alpha, side, type)
```

### Arguments

|                        |  |
|------------------------|--|
| <code>exprslist</code> | Each element of <i>exprslist</i> is a list with the first element being <i>exprs</i> and the second element being <i>classlab</i> . Each row of <i>exprs</i> represents one gene and each column represents one sample. <i>classlab</i> is a zero-one vector indicating the status of samples. We use 0 for the baseline group, usually the normal group, and 1 for the comparison group, usually the tumor group. |
| <code>alpha</code>     | Significance level for P-value.  |
| <code>side</code>      | A vector specifying the definition of P-value in each of the study, which could be either <i>up</i> , <i>down</i> , or <i>twosided</i> .   |
| <code>type</code>      | A vector specifying whether the outlier pattern is <i>subtype</i> or <i>uniform</i> .  |

### Value

|                                    |   |
|------------------------------------|---|
| <code>PCOPAstatistics</code>       | the P-value based outlier gene detection statistics   |
| <code>outliergene_bypatient</code> | a list whose length equals the number of tumor samples (patients). each element of the list is a list of length equaling to the length of <i>exprslist</i> , in other words the number of studies(or data type), showing the outlier gene for each patient in each study (or data type) |

### Author(s)

Yingying Wei

## References

Wei, Y., Hennessey, P., Gaykalova, D., Califano, J.A., Ochs, M.F., (2011) Cancer Outlier Gene Profile Sets Elucidate Pathways in Head and Neck Squamous Cell Carcinoma.

## Examples

```
#read in data
data(Exon_exprs_matched)
data(Methy_exprs_matched)
data(CNV_exprs_matched)
data(Exon_classlab_matched)
data(Methy_classlab_matched)
data(CNV_classlab_matched)
head(Exon_exprs_matched)

#exprslist[[i]]$exprs should be in matrix format
Exon_exprs<-as.matrix(Exon_exprs_matched)
Methy_exprs<-as.matrix(Methy_exprs_matched)
CNV_exprs<-as.matrix(CNV_exprs_matched)

#exprslist[[i]]$classlab should be in vector format
Exon_classlab<-unlist(Exon_classlab_matched)
Methy_classlab<-unlist(Methy_classlab_matched)
CNV_classlab<-unlist(CNV_classlab_matched)

#make an exprslist consisting 3 studies
trylist<-list()
trylist[[1]]<-list(exprs=Exon_exprs,classlab=Exon_classlab)
trylist[[2]]<-list(exprs=Methy_exprs,classlab=Methy_classlab)
trylist[[3]]<-list(exprs=CNV_exprs,classlab=CNV_classlab)

#calculate P-value based statistics for outlier gene detection and output the outlier gene list for each patient
a7<-PCOPA(trylist,0.05,side=c("up","down","up"),type="subtype")
```

---

permCOPA

*Calculate PCOPA value for permutations*

---

## Description

Run permutations by randomly shuffling the sample class labels and calculate a vector of PCOPA values for each permutation.

## Usage

```
permCOPA(exprslist, alpha=0.05, side, type, perms=100)
```

**Arguments**

|                        |  |
|------------------------|--|
| <code>exprslist</code> | Each element of <i>exprslist</i> is a list with the first element being <i>exprs</i> and the second element being <i>classlab</i> . Each row of <i>exprs</i> represents one gene and each column represents one sample. <i>classlab</i> is a zero-one vector indicating the status of samples. We use 0 for the baseline group, usually the normal group, and 1 for the comparison group, usually the tumor group. |
| <code>alpha</code>     | Significance level for P-value.  |
| <code>side</code>      | A vector specifying the definition of P-value in each of the study, which could be either <i>up</i> , <i>down</i> , or <i>twosided</i> .   |
| <code>type</code>      | A vector specifying whether the outlier pattern is <i>subtype</i> or <i>uniform</i> .  |
| <code>perms</code>     | Number of permutations to run.   |

**Value**

|                         |   |
|-------------------------|---|
| <code>permResult</code> | A matrix where each row correspond to a gene and each column correspond to one permutation. |
|-------------------------|---|

**Author(s)**

Michael Ochs

**References**

Wei, Y., Hennessey, P., Gaykalova, D., Califano, J.A., Ochs, M.F., (2011) Cancer Outlier Gene Profile Sets Elucidate Pathways in Head and Neck Squamous Cell Carcinoma.

**Examples**

```
#read in data
data(Exon_exprs_matched)
data(Methy_exprs_matched)
data(CNV_exprs_matched)
data(Exon_classlab_matched)
data(Methy_classlab_matched)
data(CNV_classlab_matched)
head(Exon_exprs_matched)

#exprslist[[i]]$exprs should be in matrix format
Exon_exprs<-as.matrix(Exon_exprs_matched)
Methy_exprs<-as.matrix(Methy_exprs_matched)
CNV_exprs<-as.matrix(CNV_exprs_matched)

#exprslist[[i]]$classlab should be in vector format
Exon_classlab<-unlist(Exon_classlab_matched)
Methy_classlab<-unlist(Methy_classlab_matched)
CNV_classlab<-unlist(CNV_classlab_matched)

#make an exprslist consisting 3 studies
trylist<-list()
```

```

trylist[[1]]<-list(exprs=Exon_exprs,classlab=Exon_classlab)
trylist[[2]]<-list(exprs=Methy_exprs,classlab=Methy_classlab)
trylist[[3]]<-list(exprs=CNV_exprs,classlab=CNV_classlab)

#run 2 permutations
perma7<-permCOPA(trylist,0.05,side=c("up","down","up"),type="subtype",perms=2)

```

---

PlotTopPCOPA

*Plot expression patterns of top ranked genes.*


---

### Description

It first sorts the expression value  $exprslist[[i]]$exprs[j,]$  among the baseline samples (e.g. normal ones) and comparison group (e.g. tumor ones) separately for selected gene  $j$ , and then plot the sorted expression values. The first argument *exprslist* should be the same one as for *PCOPA*; the second argument *PCOPAResult* should be an output of *PCOPA*; the third argument *topcut* determines how far we would go down the top ranked list; and the last argument *typelist* is a vector specifying the titles for each graph corresponds to a specific study.

### Usage

```
PlotTopPCOPA(exprslist, PCOPAResult, topcut, typelist)
```

### Arguments

|                    |  |
|--------------------|--|
| <i>exprslist</i>   | Each element of <i>exprslist</i> is a list with the first element being <i>exprs</i> and the second element being <i>classlab</i> . Each row of <i>exprs</i> represents one gene and each column represents one sample. <i>classlab</i> is a zero-one vector indicating the status of samples. We use 0 for the baseline group, usually the normal group, and 1 for the comparison group, usually the tumor group. |
| <i>PCOPAResult</i> | Output of <i>PCOPA</i> .   |
| <i>topcut</i>      | Cutoff of top ranked gene list.  |
| <i>typelist</i>    | A vector specifying the titles for each graph corresponds to a specific study.   |

### Author(s)

Michael Ochs, Yingying Wei

### Examples

```

#read in data
data(Exon_exprs_matched)
data(Methy_exprs_matched)
data(CNV_exprs_matched)
data(Exon_classlab_matched)
data(Methy_classlab_matched)
data(CNV_classlab_matched)
head(Exon_exprs_matched)

```

```

#exprslist[[i]]$exprs should be in matrix format
Exon_exprs<-as.matrix(Exon_exprs_matched)
Methy_exprs<-as.matrix(Methy_exprs_matched)
CNV_exprs<-as.matrix(CNV_exprs_matched)

#exprslist[[i]]$classlab should be in vector format
Exon_classlab<-unlist(Exon_classlab_matched)
Methy_classlab<-unlist(Methy_classlab_matched)
CNV_classlab<-unlist(CNV_classlab_matched)

#make an exprslist consisting 3 studies
trylist<-list()
trylist[[1]]<-list(exprs=Exon_exprs,classlab=Exon_classlab)
trylist[[2]]<-list(exprs=Methy_exprs,classlab=Methy_classlab)
trylist[[3]]<-list(exprs=CNV_exprs,classlab=CNV_classlab)

#calculate P-value based statistics for outlier gene detection and output the outlier gene list for each patient
a7<-PCOPA(trylist,0.05,side=c("up","down","up"),type="subtype")

#plot expression patterns of top ranked genes.
PlotTopPCOPA(trylist,a7,topcut=1,typelist=c("Exon","Methy","CNV"))

```

---

SampleData

*Sample Data for coGPS*

---

## Description

Here we present an example of coGPS analysis.

## Arguments

Exon\_exprs\_matched  
Expression data for 44 tumors and 25 normals. Each row indicates a gene with row name showing gene name and each column indicates a sample with column name showing sample name.

Exon\_class\_matched  
A length 69 vector showing status of corresponding exon samples, 0 for normals and 1 for tumors.

Methy\_exprs\_matched  
Methylation data for 44 tumors and 25 normals.

Methy\_class\_matched  
A length 69 vector showing status of corresponding methylation samples, 0 for normals and 1 for tumors.

CNV\_exprs\_matched  
Copy number data for 44 tumors and 25 normals.

CNV\_class\_matched  
A length 69 vector showing status of corresponding copy number samples, 0 for normals and 1 for tumors.

Hs.gmt1.c1  
Broad Institute C1 Positional Gene Sets.



**Details**

In this application, the columns of each data type are matched. In other words, the first columns of Exon\_exprs\_matched, Methy\_exprs\_matched and CNV\_exprs\_matched correspond to the same patient. And hence the Exon\_class\_matched, Methy\_class\_matched and CNV\_class\_matched are identical. However, suppose in applications that we are not concerned with the outlier gene list for each patient, we can leave with the samples (columns) unmatched.

# Index

## \* **Microarray, Bioinformatics, DifferentialExpression**

coGPS-package, [2](#)

CNV\_classlab\_matched (SampleData), [8](#)

CNV\_exprs\_matched (SampleData), [8](#)

coGPS (coGPS-package), [2](#)

coGPS internal, [2](#)

coGPS-package, [2](#)

Exon\_classlab\_matched (SampleData), [8](#)

Exon\_exprs\_matched (SampleData), [8](#)

Hs.gmt1.c1 (SampleData), [8](#)

Methy\_classlab\_matched (SampleData), [8](#)

Methy\_exprs\_matched (SampleData), [8](#)

PatientSpecificGeneList, [2](#)

PCOPA, [4](#)

permCOPA, [5](#)

plotCOPA (coGPS internal), [2](#)

PlotTopPCOPA, [7](#)

SampleData, [8](#)