

# Package ‘gramm4R’

April 15, 2020

**Type** Package

**Title** Generalized correlation analysis and model construction strategy  
for metabolome and microbiome

**Version** 1.0.0

**Author** Mengci Li, Dandan Liang, Tianlu Chen and Wei Jia

**Maintainer** Tianlu Chen <chentianlu@sjtu.edu.cn>

**Description** Generalized Correlation Analysis for Metabolome and Microbiome (GRaMM), for inter-correlation pairs discovery among metabolome and microbiome.

**License** GPL-2

**Encoding** UTF-8

**biocViews** GraphAndNetwork, Microbiome

**Depends** R (>= 3.6.0)

**LazyData** true

**Imports** basicTrendline, investr, minerva, psych, grDevices, graphics,  
stats, DelayedArray, SummarizedExperiment, DMwR, phyloseq

**VignetteBuilder** knitr

**Suggests** knitr, rmarkdown

**git\_url** <https://git.bioconductor.org/packages/gramm4R>

**git\_branch** RELEASE\_3\_10

**git\_last\_commit** f0c9bc5

**git\_last\_commit\_date** 2019-10-29

**Date/Publication** 2020-04-14

## R topics documented:

covariates . . . . .	2
Gramm . . . . .	2
metabolites . . . . .	4
microbes . . . . .	4
naiveGramm . . . . .	5
nlfittGramm . . . . .	6
preGramm . . . . .	7
<b>Index</b>	<b>9</b>

---

covariates	<i>Test data of covariates</i>
------------	--------------------------------

---

### Description

The test data for examples at function [naiveGramm](#), [Gramm](#), [nlfitGramm](#)

### Usage

```
data("covariates")
```

### Format

The format is: Formal class 'SummarizedExperiment' [package "SummarizedExperiment"] with 5 slots ..@ colData : Formal class 'DataFrame' [package "S4Vectors"] with 6 slots ... ..@ rownames : chr [1:42] "B1" "B2" "B3" "B4" ... ..@ nrows : int 42 .. ..@ listData : Named list() .. ..@ elementType : chr "ANY" .. ..@ elementMetadata: NULL .. ..@ metadata : list() ..@ assays : Formal class 'SimpleAssays' [package "SummarizedExperiment"] with 1 slot .. ..@ data: Formal class 'SimpleList' [package "S4Vectors"] with 4 slots .. ..@ listData : List of 1 .. ..@ ..\$ counts: int [1:3, 1:42] 25 1 90 26 1 70 27 1 68 28 ... ..@ ..\$ : chr [1:42] "B1" "B2" "B3" "B4" ... ..@ ..@ elementType : chr "ANY" .. ..@ elementMetadata: NULL .. ..@ metadata : list() ..@ NAMES : NULL ..@ elementMetadata: Formal class 'DataFrame' [package "S4Vectors"] with 6 slots .. ..@ rownames : NULL .. ..@ nrows : int 3 .. ..@ listData : List of 1 .. ..@ ..\$ X: Factor w/ 3 levels "age","gender",,..: 1 2 3 .. ..@ elementType : chr "ANY" .. ..@ elementMetadata: NULL .. ..@ metadata : list() ..@ metadata : list()

### Examples

```
data(covariates)
## maybe str(covariates) ; plot(covariates) ...
```

---

Gramm	<i>Get the association between metabolites and microbes</i>
-------	---

---

### Description

The entire strategy to get the association between metabolites and microbes, using linear and non-linear methods, and plot the regression figures.

### Usage

```
Gramm(A,B,C,metaNor,rarefaction,r,alpha)
```

**Arguments**

A	A SummarizedExperiment object contains data of metabolome, where rows represent features of metabolites and columns represent samples.
B	A SummarizedExperiment object contains data of microbiome, where rows represent features of microbes and columns represent samples.
C	A SummarizedExperiment object contains data of covariates, where rows represent features of covariates and columns represent samples.
metaNor	Should metabolome data normalized? Using normalization when your metabolites are qualitative; and no normalization when the metabolites are quantitative. Default:TRUE.
rarefaction	Resample an OTU table such that all samples have the same library size. Here refers to a repeated sampling procedure to assess species richness, first proposed in 1968 by Howard Sanders.(see wikipedia for more detail.) Default:FALSE.
r	The linear regression coefficients threshold for using nonlinear method. Default: 0.5.
alpha	The linear regression p-value threshold for using nonlinear method.Default: 0.05.

**Value**

pretreatment	The result of pretreatment
correlation	The result of correlation, see <a href="#">naiveGramm</a> for detail

A file named "R value top 10 pairs.pdf" will be created automatically (correlation coefficient top 10 pairs).

**Author(s)**

Mengci Li, Dandan Liang, Tianlu Chen and Wei Jia

**References**

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., Egozcue, J. J., Microbiome Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* 2017, 8 (2224). Chambers, J. M. (1992) Linear models. Chapter 4 of *Statistical Models in S* eds J. M. Chambers and T. J. Hastie, Wadsworth & Brooks/Cole. D. Reshef, Y. Reshef, H. Finucane, S. Grossman, G. McVean, P. Turnbaugh, E. Lander, M. Mitzenmacher, P. Sabeti. (2011) Detecting novel associations in large datasets. *Science* 334, 6062. D. Albanese, M. Filosi, R. Visintainer, S. Riccadonna, G. Jurman, C. Furlanello. *minerva and minepy: a C engine for the MINE suite and its R, Python and MATLAB wrappers.* *Bioinformatics* (2013) 29(3): 407-408.

**See Also**

[preGramm](#) for pretreatment;[nlfitGramm](#) for nonlinear fitting;[naiveGramm](#) for naive correlation method.

**Examples**

```
data("metabolites");data("microbes");data("covariates")
Gramm(metabolites,microbes,covariates)
```

---

metabolites

*Test data of metabolome*


---

### Description

The test data for examples at function [naiveGramm](#), [Gramm](#), [nlfitGramm](#)

### Usage

```
data("metabolites")
```

### Format

The format is: Formal class 'SummarizedExperiment' [package "SummarizedExperiment"] with 5 slots ..@ colData :Formal class 'DataFrame' [package "S4Vectors"] with 6 slots .. ..@ rownames : chr [1:42] "B1" "B2" "B3" "B4" ... .. ..@ nrows : int 42 .. ..@ listData : Named list() .. ..@ elementType : chr "ANY" .. ..@ elementMetadata: NULL .. ..@ metadata : list() ..@ assays :Formal class 'SimpleAssays' [package "SummarizedExperiment"] with 1 slot .. ..@ data:Formal class 'SimpleList' [package "S4Vectors"] with 4 slots .. .. ..@ listData :List of 1 .. .. .. .. ..\$ counts: num [1:40, 1:42] 8.75 3.31 5.09 1.49 6.2 8.81 2.37 1.76 4.47 2.18 ... .. .. .. ..\$ attr(\*, "dimnames")=List of 2 .. .. .. .. ..\$ : NULL .. .. .. .. ..\$ : chr [1:42] "B1" "B2" "B3" "B4" ... .. .. .. ..@ elementType : chr "ANY" .. .. .. ..@ elementMetadata: NULL .. .. .. ..@ metadata : list() ..@ NAMES : NULL ..@ elementMetadata:Formal class 'DataFrame' [package "S4Vectors"] with 6 slots .. .. ..@ rownames : NULL .. .. ..@ nrows : int 40 .. .. ..@ listData :List of 1 .. .. .. ..\$ X: Factor w/ 40 levels "Decadienylcarnitine",,..: 16 35 31 25 3 4 1 8 9 6 ... .. .. ..@ elementType : chr "ANY" .. .. ..@ elementMetadata: NULL .. .. ..@ metadata : list() ..@ metadata : list()

### Examples

```
data(metabolites)
## maybe str(metabolites) ; plot(metabolites) ...
```

---

microbes

*Test data of microbiome*


---

### Description

The test data for examples at function [naiveGramm](#), [Gramm](#), [nlfitGramm](#)

### Usage

```
data("microbes")
```

**Format**

The format is: Formal class 'SummarizedExperiment' [package "SummarizedExperiment"] with 5 slots ..@ colData :Formal class 'DataFrame' [package "S4Vectors"] with 6 slots .. ..@ rownames : chr [1:42] "B1" "B2" "B3" "B4" ... ..@ nrows : int 42 .. ..@ listData : Named list() .. ..@ elementType : chr "ANY" .. ..@ elementMetadata: NULL .. ..@ metadata : list() ..@ assays :Formal class 'SimpleAssays' [package "SummarizedExperiment"] with 1 slot .. ..@ data:Formal class 'SimpleList' [package "S4Vectors"] with 4 slots .. .. ..@ listData :List of 1 .. .. ..@ counts: num [1:85, 1:42] 1 1 13 2 4 8 11 24 6 1 ... .. ..@ attr(\*, "dimnames")=List of 2 .. .. ..@ elementType : chr "ANY" .. ..@ elementMetadata: NULL .. ..@ metadata : list() ..@ NAMES : NULL ..@ elementMetadata:Formal class 'DataFrame' [package "S4Vectors"] with 6 slots .. ..@ rownames : NULL .. ..@ nrows : int 85 .. ..@ listData :List of 1 .. .. ..@ X: Factor w/ 84 levels "[Ruminococcus] spp",...: 74 24 1 12 20 21 29 70 65 20 ... ..@ elementType : chr "ANY" .. ..@ elementMetadata: NULL .. ..@ metadata : list() ..@ metadata : list()

**Examples**

```
data(microbes)
## maybe str(microbes) ; plot(microbes) ...
```

---

naiveGramm

*Correlation methods of Gramm*


---

**Description**

A function to test the association among metabolites and microbes. Using linear (linear regression) or nonlinear (Maximal Information Coefficient) methods.

**Usage**

```
naiveGramm(x,y,z,r,alpha)
```

**Arguments**

x	A SummarizedExperiment object contains data of preprocessed data, where rows represent features of metabolites and columns represent samples.
y	A SummarizedExperiment object contains data of microbiome preprocessed data, where rows represent features of microbes and columns represent samples.
z	An optional SummarizedExperiment object containing the confounders.
r	The linear regression coefficients threshold for using nonlinear method. Default: 0.5.
alpha	The linear regression p-value threshold for using nonlinear method. Default: 0.05.

**Value**

r	Correlation coefficients.
p	Correlation p values.
type	Correlation methods for computing (linear or nonlinear).

**Author(s)**

Mengci Li, Dandan Liang, Tianlu Chen and Wei Jia

**References**

Chambers, J. M. (1992) Linear models. Chapter 4 of Statistical Models in S eds J. M. Chambers and T. J. Hastie, Wadsworth & Brooks/Cole. D. Reshef, Y. Reshef, H. Finucane, S. Grossman, G. McVean, P. Turnbaugh, E. Lander, M. Mitzenmacher, P. Sabeti. (2011) Detecting novel associations in large datasets. Science 334, 6062. D. Albanese, M. Filosi, R. Visintainer, S. Riccadonna, G. Jurman, C. Furlanello. minerva and minepy: a C engine for the MINE suite and its R, Python and MATLAB wrappers. Bioinformatics (2013) 29(3): 407-408.

**See Also**

[preGramm](#) for pretreatment; [nlfitGramm](#) for nonlinear fitting; [Gramm](#): the whole strategy of this method.

**Examples**

```
data("metabolites")
data("microbes")
data("covariates")
naiveGramm(metabolites,microbes,covariates)
```

---

nlfitGramm

---

*Plot the nonlinear fitting of input data*


---

**Description**

Plot, draw regression line and confidence interval, and show regression equation, R-square and P-value. The function includes the following models in the latest version: "line2P" (formula as:  $y=a*x+b$ ), "line3P" ( $y=a*x^2+b*x+c$ ), "log2P" ( $y=a*\ln(x)+b$ ), "exp2P" ( $y=a*\exp(b*x)$ ), "exp3P" ( $y=a*\exp(b*x)+c$ ), "power2P" ( $y=a*x^b$ ), "power3P" ( $y=a*x^b+c$ ), and "S" ( $y=a/(1+\exp((x-b)/c))$ ). Of each correlation pairs the figure who has the highest R-square among these models will be saved in a "pdf" file at working directory.

**Usage**

```
nlfitGramm(X,Y)
```

**Arguments**

X	Metabolome preprocessed data (SummarizedExperiment object).
Y	Microbiome preprocessed data (SummarizedExperiment object).

**Details**

nothing

**Value**

A file named "curve fit.pdf" will be created automatically.

**Author(s)**

Mengci Li, Dandan Liang, Tianlu Chen and Wei Jia

**References**

Bates, D. M., and Watts, D. G. (2007) Nonlinear Regression Analysis and its Applications. Wiley.  
 Greenwell B. M., and Schubert-Kabban, C. M. (2014) investr: An R Package for Inverse Estimation. The R Journal, 6(1), 90-100.

**See Also**

[naiveGramm](#) for naive correlation method; [preGramm](#) for pretreatment; [Gramm](#): the whole strategy of this method.

**Examples**

```
data("metabolites")
data("microbes")
nlfitGramm(metabolites,microbes)
```

---

```
preGramm
```

---

*Preprocess the metabolome data and microbiome data*

---

**Description**

Preprocess the input data of metabolome and microbiome. Missing values may be imputed and filled (KNN method). Metabolome data and microbiomedata may be normalized and transformed by logarithm transformation and centered log-ratio (CLR) algorithm.

**Usage**

```
preGramm(A,B,metaNor = TRUE,rarefaction = FALSE)
```

**Arguments**

A	The metabolome data under pretreatment (SummarizedExperiment object).
B	The microbiome data under pretreatment (SummarizedExperiment object).
metaNor	Should metabolome data normalized? Using normalization when your metabolites are qualitative; and no normalization when the metabolites are quantitative. Default:TRUE.
rarefaction	Resample an OTU table such that all samples have the same library size. Here refers to a repeated sampling procedure to assess species richness, first proposed in 1968 by Howard Sanders.(see wikipedia for more detail.) Default:FALSE.

**Value**

x	Metabolome preprocessed data
y	Microbiome preprocessed data

**Author(s)**

Mengci Li, Dandan Liang, Tianlu Chen and Wei Jia

**References**

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., Egozcue, J. J., Microbiome Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* 2017, 8 (2224).

**See Also**

[naiveGramm](#) for naive correlation method; [nlfitGramm](#) for nonlinear fitting; [Gramm](#): the whole strategy of this method.

**Examples**

```
data("metabolites")
data("microbes")
preGramm(metabolites,microbes)
```



# Index

## \*Topic **Nonlinear Fitting**

nlfitGramm, 6

## \*Topic **datasets**

covariates, 2

metabolites, 4

microbes, 4

covariates, 2

Gramm, 2, 2, 4, 6–8

metabolites, 4

microbes, 4

naiveGramm, 2–4, 5, 7, 8

nlfitGramm, 2–4, 6, 6, 8

preGramm, 3, 6, 7, 7