

Package ‘PhageCocktail’

July 14, 2021

Type Package

Title Design of the Best Phage Cocktail

Version 1.0.0

Description There are 4 possible methods: ``ExhaustiveSearch``; ``ExhaustivePhi``; ``ClusteringSearch``; and ``ClusteringPhi``.
``ExhaustiveSearch``--> gives you the best phage cocktail from a phage-bacteria infection network. It checks different phage cocktail sizes from 1 to 7 and only stops before if it lyses all bacteria. Other option is when users have decided not to obtain a phage cocktail size higher than a limit value.
``ExhaustivePhi``--> firstly, it finds Phi out. Phi is a formula indicating the necessary phage cocktail size. Phi needs nestedness temperature and fill, which are internally calculated. This function will only look for the best combination (phage cocktail) with a Phi size.
``ClusteringSearch``--> firstly, an agglomerative hierarchical clustering using Ward's algorithm is calculated for phages. They will be clustered according to bacteria lysed by them. PhageCocktail() chooses how many clusters are needed in order to select 1 phage per cluster. Using the phages selected during the clustering, it checks different phage cocktail sizes from 1 to 7 and only stops before if it lyses all bacteria. Other option is when users have decided not to obtain a phage cocktail size higher than a limit value.
``ClusteringPhi``--> firstly, an agglomerative hierarchical clustering using Ward's algorithm is calculated for phages. They will be clustered according to bacteria lysed by them. PhageCocktail() chooses how many clusters are needed in order to select 1 phage per cluster. Once the function has one phage per cluster, it calculates Phi. If the number of clusters is less than Phi number, it will be changed to obtain, as minimum, this quantity of candidates (phages). Then, it calculates the best combination of Phi phages using those selected during the clustering with Ward algorithm.

Imports readxl, stringr, factoextra, bipartite, smerc, RJSONIO

License GPL-3

Encoding UTF-8

RoxygenNote 7.1.1.9001

NeedsCompilation no

Author Maria Victoria Diaz-Galian [aut, cre]
 (<<https://orcid.org/0000-0003-2543-6081>>),
 Miguel A. Vega-Rodriguez [aut]
 (<<https://orcid.org/0000-0002-3003-758X>>),
 Felipe Molina [aut] (<<https://orcid.org/0000-0002-9167-7097>>)

Maintainer Maria Victoria Diaz-Galian <mvdiazgalian@unex.es>

Repository CRAN

Date/Publication 2021-07-14 08:30:05 UTC

R topics documented:

PhageCocktail	2
Index	6

PhageCocktail	<i>Design of the Best Phage Cocktail</i>
---------------	--

Description

It gives you the best phage cocktail which lyses the maximum number of bacteria.

There are 4 possible methods within the function PhageCocktail(): "ExhaustiveSearch"; "ExhaustivePhi"; "ClusteringSearch"; and "ClusteringPhi".

"*ExhaustiveSearch*" gives you the best phage cocktail from a phage bacteria infection network. It checks different phage cocktail sizes from 1 to 7 and only stops before if it lyses all bacteria. Other option is when users have decided not to obtain a phage cocktail size higher than a limit value.

"*ExhaustivePhi*" firstly, the function PhageCocktail() finds Phi out. Phi is a formula indicating the necessary phage cocktail size [Molina et al. (2021)<[doi:10.3389/fmicb.2021.564532](https://doi.org/10.3389/fmicb.2021.564532)>]. Phi needs nestedness temperature and fill [Molina et al. (2021)<[doi:10.3389/fmicb.2021.564532](https://doi.org/10.3389/fmicb.2021.564532)>][Rodríguez Gironés et al. (2006)<[doi:10.1111/j.1365-2699.2006.01444.x](https://doi.org/10.1111/j.1365-2699.2006.01444.x)>], which are internally calculated. This method will only look for the best combination (phage cocktail) with a Phi size.

"*ClusteringSearch*" firstly, an agglomerative hierarchical clustering using Ward's algorithm is calculated for phages [Strauss et al. (2017)<[doi:10.1371/journal.pone.0168288](https://doi.org/10.1371/journal.pone.0168288)>]. They will be clustered according to bacteria lysed by them. 'PhageCocktail' chooses how many clusters are needed in order to select 1 phage per cluster [Kodinariya et al. (2013)]. Using the phages selected during the clustering, it checks different phage cocktail sizes from 1 to 7 and only stops before if it lyses all bacteria. Other option is when users have decided not to obtain a phage cocktail size higher than a limit value.

"*ClusteringPhi*" firstly, an agglomerative hierarchical clustering using Ward's algorithm is calculated for phages [Strauss et al. (2017)<[doi:10.1371/journal.pone.0168288](https://doi.org/10.1371/journal.pone.0168288)>]. They will be clustered according to bacteria lysed by them. 'PhageCocktail' chooses how many clusters are needed in order to select 1 phage per cluster [Kodinariya et al. (2013)]. Once the function (PhageCocktail()) has one phage per cluster, it calculates Phi [Molina et al. (2021)<[doi:10.3389/fmicb.2021.564532](https://doi.org/10.3389/fmicb.2021.564532)>]. If

the number of clusters is less than Phi number, it will be changed to obtain, as minimum, this quantity of candidates (phages). Then, it calculates the best combination of Phi phages using those selected during the clustering with Ward algorithm.

Usage

```
## Default method:
PhageCocktail(pathway,FUN="ClusteringSearch",file.output=FALSE,
file.name="output.txt",limit=7)
```

Arguments

pathway	string with the input file pathway. The input file must be in .xlsx format and must contain one PhageBacteria Infection Network (PBIN). This PBIN must be in matrix format, that is, the input file must have only one sheet with a matrix, whose rows are the bacteria and whose columns are the phages. It is important to say that the first row must be devoted to the names/codes of the phages. In the same way, the first column must be devoted to the names/codes of the bacteria. Only the values 0 and 1 are considered to reflect the infection relationship between involved phages and bacteria (0 = does not lyse and 1 = lyses). In this way, the infection relationship can be represented by the names of phages and bacteria, and by the position of the value (0 or 1) in the matrix.
FUN	string indicating the function/method to execute: " <i>ExhaustiveSearch</i> " ; " <i>ExhaustivePhi</i> "; " <i>ClusteringSearch</i> " (default option) ; " <i>ClusteringPhi</i> ".
file.output	logical. If <i>FALSE</i> (default option), the program will not create an output file. If <i>TRUE</i> , the program will write a file with the results.
file.name	string with the filename of the output file. Default option is "output.txt"
limit	Positive integer number indicating maximum phage cocktail size desired. The range of options is: (1:7). Default option is 7. It is desirable to use limit when the number of phages and/or bacteria is extremely high.

Value

\$file	string indicating the input file used.
\$method	string indicating the function chosen.
\$maxCocktailSize	integer. Maximum phage cocktail size.
\$allBacterialLysed	logical. If <i>TRUE</i> indicates that the set of phages lyses all bacteria. If <i>FALSE</i> , not all bacteria are lysed by the phage cocktail.
\$Result	dataframe composed of the information of each phage cocktail size (phage cocktail and number of bacteria lysed) and the list of lysed bacteria. An example ("ClusteringPhi"): "Result": "Size3": ["SajbenNagyP16", "SajbenNagyP3", "SajbenNagyP10", "18"], "BacterialLysed":

```
"Size3": [ "SajbenNagyB1", "SajbenNagyB2", "SajbenNagyB3", "SajbenNagyB4",
"SajbenNagyB5", "SajbenNagyB6", "SajbenNagyB10", "SajbenNagyB11", "Sajben
NagyB13", "SajbenNagyB15", "SajbenNagyB16", "SajbenNagyB18", "Sajben
NagyB19", "SajbenNagyB23", "SajbenNagyB24", "SajbenNagyB30", "Sajben
NagyB31", "SajbenNagyB32" ]
```

As it is an output of executing "ClusteringPhi", there is only one phage cocktail. The meaning is that the best Phage Cocktail of three phages is composed by "Sajben-NagyP16", "Sajben-NagyP3" and "Sajben-NagyP10" lysing 18 bacteria shown in the list called "BacteriaLysed".

```
$non_usefulPhages
```

vector of phages removed because they do not lyse any bacterium. If all of them lyse some bacterium, this message will appear: "All the phages are useful (all of them lyse some bacterium)"

```
$non_lysedBacteria
```

vector of bacteria not lysed by any phage. If there is no one, this message will appear: "All the bacteria are lysed by some phage"

Author(s)

Díaz-Galián, María Victoria <mvdiagalían@unex.es>

Vega-Rodríguez, Miguel A. <mavega@unex.es>

Molina, Felipe <fmolina@unex.es>

University of Extremadura, Spain

References

Molina, F., Simancas, A., Ramírez, M., Tabla, R., Roa, I., & Rebollo, J. E. (2021). A new pipeline for designing phage cocktails based on phage-bacteria infection networks. *Frontiers in Microbiology*, 12, 149. doi: [10.3389/fmicb.2021.564532](https://doi.org/10.3389/fmicb.2021.564532)

Strauss, T., & von Maltitz, M. J. (2017). Generalising Ward's method for use with Manhattan distances. *PloS ONE*, 12(1), e0168288. doi: [10.1371/journal.pone.0168288](https://doi.org/10.1371/journal.pone.0168288)

Rodríguez Gironés, M. A., & Santamaría, L. (2006). A new algorithm to calculate the nestedness temperature of presence-absence matrices. *Journal of Biogeography*, 33(5), 924-935. doi: [10.1111/j.13652699.2006.01444.x](https://doi.org/10.1111/j.13652699.2006.01444.x)

Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of cluster in K-means clustering. *International Journal of Advance Research in Computer Science and Management Studies*, 1(6), 90-95.

Examples

```
library(readxl)
x<-system.file("extdata", "Sajben-Nagy.xlsx", package="PhageCocktail", mustWork=TRUE)
```

```
PhageCocktail(x, "ExhaustiveSearch")
```

```
PhageCocktail(x, "ExhaustivePhi")
```

```
#If the number of phages and/or bacteria is extremely high, it is  
#recommended to use limit option starting by 1 for "ExhaustiveSearch"  
#avoiding that the result takes too much time. If process took a short  
#period of time and it was desirable to test with a higher number of  
#phages, it could be done.
```

```
PhageCocktail(x,"ExhaustiveSearch",limit=1)
```

Index

PhageCocktail, [2](#)