

Package ‘RchyOptimyx’

September 24, 2012

Type Package

Title Optimized Cellular Hierarchies for Flow Cytometry

Version 1.0.0

Date 2012-03-19

Author Adrin Jalali, Nima Aghaeepour

Maintainer Adrin Jalali <adrin.jalali@gmail.com>, Nima Aghaeepour <naghaeep@bccrc.ca>

Description Constructs a hierarchy of cells using flow cytometry for maximization of an external variable (e.g., a clinical outcome or a cytokine response).

Imports Rgraphviz, sfsmisc, graphics, methods, graph, grDevices

Depends R (>= 2.10)

Suggests flowCore, flowType

biocViews FlowCytometry

License Artistic-2.0

LazyLoad yes

R topics documented:

RchyOptimyx-package	2
HIVData	3
merge	4
plot	4
RchyOptimyx	6
summary-methods	7
Index	8

RchyOptimyx-package *Gating Hierarchy Optimization for Flow Cytometry*

Description

RchyOptimyx is a computational tool which determines the minimal sets of markers that can identify a target population within a given purity standard. This can enable panel redesign using smaller subsets of markers, with consequent savings on reagents, potential for adaptation to older instruments, or the addition of other markers of interest to the panel design.

Details

Package:	RchyOptimyx
Type:	Package
Version:	0.99.1
Date:	2011-10-12
License:	Artistic-2.0
LazyLoad:	yes

See the RchyOptimyx function.

Author(s)

Nima Aghaeepour <<naghaeep@bccrc.ca>> and Adrin Jalali <<ajalali@bccrc.ca>>

References

D. Eppstein, Finding the k shortest paths, SIAM J. Comput., 1998 - Citeseer.

See Also

RchyOptimyx

Examples

```
data(HIVData)

res<-RchyOptimyx(Signs, -log10(LogRankPvals), '2111012110', 5,FALSE)
plot(res, phenotypeScores=-log10(LogRankPvals), ylab='-log10(Pvalue)')

res<-RchyOptimyx(Signs, OverlapScores, '1201100221', 5,FALSE)
plot(res, phenotypeScores=OverlapScores, ylab='Purity')
```

HIVData

HIVData

Description

A dataset of two sets of scores (particularly, correlation with protection against HIV and overlap with the Naive T-cell population) assigned to immunophenotypes measured by flow cytometry. 10 markers were measured: KI-67, CD28, CD45RO, CD8, CD4, CD57, CCR5, CD27, CCR7, and CD127.

Usage

```
data(HIVData)
```

Details

This dataset consists of a matrix (`Signs`) and 2 numeric vectors (`LogRankPvals` and `OverlapScores`). The `Signs` matrix consists of 10 columns (one per measured marker) and $3^{10} - 1 = 59048$ rows (one per immunophenotype). See Aghaeepour et.al. or the `flowType` package for more details. For every phenotype (row), the entity corresponding to a given marker (column) can be 0, 1, and 2 for negative, neutral, and positive respectively. The rownames and column names are set respectively.

`LogRankPvals` is a vector of log-rank test P-values to determine the correlation between HIV's progression and each of the measured immunophenotypes in 466 HIV positive patients (lower values represent a stronger correlation). For more details see the Aghaeepour et.al. manuscript below. The names of the vector match the names of the `Signs` matrix.

Ganesan et. al. define Naive T-cells as CD28+CD45RO-CD57-CCR5-CD27+ CCR7+ within the CD3+CD14- compartment. The `OverlapScores` vector has the proportion of Naive T-cells (as defined above) to the total number of cells in any given immunophenotype (a higher value represents a larger overlap). The names of the vector match the names of the `Signs` matrix.

Author(s)

Nima Aghaeepour <<naghaeep@bccrc.ca>> and Adrin Jalali <<ajalali@bccrc.ca>>

References

Nima Aghaeepour, Pratip K. Chattopadhyay, Anuradha Ganesan, Kieran O'Neill, Habil Zare, Adrin Jalali, Holger H. Hoos, Mario Roederer, and Ryan R. Brinkman. Early Immunologic Correlates of HIV Protection can be Identified from Computational Analysis of Complex Multivariate T-cell Flow Cytometry Assays. *Bioinformatics*, pending on minor revisions, 2012.

Anuradha Ganesan, Pratip K Chattopadhyay, Tess M. Brodie, Jing Qin, Wenjuan Gu, John R. Mascola, Nelson L. Michael, Dean A. Follmann, and Mario Roederer. Immunologic and virologic events in early HIV infection predict subsequent rate of progression. *Journal of Infectious Diseases*, 2010.

merge	<i>Merges two OptimizedHierarchy objects.</i>
-------	---

Description

Merges two OptimizedHierarchy objects.

Usage

```
## S4 method for signature 'OptimizedHierarchy,OptimizedHierarchy'
merge(x, y)
```

Arguments

x	First OptimizedHierarchy.
y	Second OptimizedHierarchy.

Value

OptH	An OptimizedHierarchy object
------	------------------------------

Author(s)

Nima Aghaeepour <<naghaeep@bccrc.ca>> and Adrin Jalali <<ajalali@bccrc.ca>>

Examples

```
data(HIVData)

res1<-RchyOptimyx(Signs, -log10(LogRankPvals), paste(Signs['KI-67+CD4-CCR5+CD127-',], collapse=''), 4,FALSE)
res2<-RchyOptimyx(Signs, -log10(LogRankPvals), paste(Signs['CD45RO-CD8+CD57+CCR5-CD27+CCR7-CD127-',], collapse=''), 4,FALSE)
res3<-RchyOptimyx(Signs, -log10(LogRankPvals), paste(Signs['CD28-CD45RO+CD4-CD57-CD27-CD127-',], collapse=''), 4,FALSE)

res=merge(res1, res2)
res=merge(res,res3)

plot(res, phenotypeScores=-log10(LogRankPvals), ylab='-log10(Pvalue)')
```

plot	<i>Plots an optimized cellular hierarchy</i>
------	--

Description

Plots an optimized cellular hierarchy

Usage

```
## S4 method for signature 'OptimizedHierarchy,ANY'
plot(x, phenotypeScores,
     uniformColors = FALSE, ylab = NULL, xlab = NULL, colors = c("blue",
     "cyan", "yellow", "red"), edgeWeights = TRUE, edgeLabels = TRUE,
     nodeLabels = TRUE, min.score = NA, max.score = NA, cell.proportions =
     NULL, min.proportion = NA, max.proportion = NA, proportion.colors =
     c("black", "white"), node.lwd = 5, root.name = "All Cells", legend.size=1.25)
```

Arguments

x	An OptimizedHierarchy object as generated by RchyOptimyx.
phenotypeScores	The score vector used for determining the colors of the nodes.
uniformColors	Boolean variable. If TRUE, the graph will not have any colors.
ylab	y-axis label of the density plot on the right.
xlab	x-axis label of the density plot of the phenotype cell proportion in the bottom.
colors	Color vector indicating colors to be used in the right pallet.
edgeWeights	Boolean value indicating whether edgeWeights should be plotted or not.
edgeLabels	Boolean value indicating whether edgeLabels should be plotted or not.
nodeLabels	Boolean value indicating whether nodeLabels should be plotted or not.
min.score	Double value indicating the minimum value of scores. Default value will use minimum node scores of the given OptimizedHierarchy object.
max.score	Double value indicating the maximum value of scores. Default value will use maximum node scores of the given OptimizedHierarchy object.
cell.proportions	The score vector used to determine the color of node borders to illustrate phenotype cell proportions.
min.proportion	Double value indicating the minimum value of the cell proportions. Default value will use minimum cell proportion of the given OptimizedHierarchy object.
max.proportion	Double value indicating the minimum value of the cell proportions. Default value will use minimum cell proportion of the given OptimizedHierarchy object.
proportion.colors	Color vector indicating colors to be used in the bottom pallet.
node.lwd	Integer value indicating node border width.
root.name	Character value indicating the root node's name.
legend.size	Size of the color legend in inches.

Author(s)

Nima Aghaeepour <<naghaeep@bccrc.ca>> and Adrin Jalali <<ajalali@bccrc.ca>>

See Also

[RchyOptimyx-methods](#)

Description

RchyOptimyx is a computational tool which determines the minimal sets of markers that can identify a target population within a given purity standard. This can enable panel redesign using smaller subsets of markers, with consequent savings on reagents, potential for adaptation to older instruments, or the addition of other markers of interest to the panel design.

Usage

```
## S4 method for signature 'ANY,numeric,character,numeric,logical'
RchyOptimyx(phenotypeSigns, phenotypeScores, startPhenotype, pathCount=1, trimPaths=FALSE, trim.
```

Arguments

phenotypeSigns A NXM matrix where N is the number of markers and M is the number of phenotypes (3^N-1). For every phenotype (row), the entity corresponding to a given marker (column) can be 0, 1, and 2 for negative, neutral, and positive respectively. The rownames must have the phenotype names and column names must have the marker names. See the provided vignette (<inst/doc/RchyOptimyx.pdf>) for more details and examples.

phenotypeScores A vector of the scores assigned to every phenotype. The optimization algorithm will try to maximize this score. The vector name must be identical to phenotypeSigns' rownames. See the provided vignette (<inst/doc/RchyOptimyx.pdf>) for more details and examples.

startPhenotype The terminal phenotype which includes every marker that must be considered in the analysis. This variable is a vector of the length of the number of markers. Every element of the vector can be 0, 1, or 2 for negative, neutral, or positive respectively (see details).

pathCount The number of paths that must be generated.

trimPaths A boolean indicating whether the paths should be trimmed or not. If set to TRUE, the paths will be terminated as soon as adding `trim.tolerance` extra number of edges to the path doesn't result in an improvement in the score.

trim.tolerance An integer indicating the number of levels before the path is trimmed if the score is decreasing by adding extra edges.

trim.level An integer indicating from which level the tree is going to be trimmed. This parameter has no effect if set to zero.

Details

Every marker can be positive, negative, or neutral. Neutral markers are not included in the measurement of the respective phenotype.

A path is trimmed in one of the following conditions: 1. If the `trimPaths` parameter is TRUE, the path is trimmed as soon as it decreases for equal or more levels that `trim.tolerance` parameter. If the path is decreasing while reaching the `trim.level`, the path is trimmed from the point that it started to decrease. 2. If the path reaches the `trim.level`.

Value

OptH An OptimizedHierarchy object

Author(s)

Nima Aghaeepour <<naghaeep@bccrc.ca>> and Adrin Jalali <<ajalali@bccrc.ca>>

References

D. Eppstein, Finding the k shortest paths, SIAM J. Comput., 1998 - Citeseer.

Nima Aghaeepour, Pratip K. Chattopadhyay, Anuradha Ganesan, Kieran O'Neill, Habil Zare, Holger H. Hoos, Mario Roederer, and Ryan R. Brinkman. Early Immunologic Correlates of HIV Protection can be Identified from Computational Analysis of Complex Multivariate T-cell Flow Cytometry Assays. Bioinformatics, Feb. 2012.

Examples

```
data(HIVData)

res<-RchyOptimyx(Signs, -log10(LogRankPvals), '2111012110', 5,FALSE)
plot(res, phenotypeScores=-log10(LogRankPvals), ylab='-log10(Pvalue)')

res<-RchyOptimyx(Signs, OverlapScores, '1201100221', 5,FALSE)
plot(res, phenotypeScores=OverlapScores, ylab='Purity')
```

summary-methods *~~ Methods for Function summary ~~*

Description

~~ Methods for function summary ~~

Methods

signature(object = "OptimizedHierarchy") Prints basic characteristics of an OptimizedHierarchy object.

Author(s)

Nima Aghaeepour <<naghaeep@bccrc.ca>> and Adrin Jalali <<ajalali@bccrc.ca>>

See Also

[RchyOptimyx-methods](#)

Index

*Topic **FlowCytData**

merge, 4

RchyOptimyx, 6

RchyOptimyx-package, 2

*Topic **package**

RchyOptimyx-package, 2

*Topic **print**

plot, 4

summary-methods, 7

HIVData, 3

LogRankPvals (HIVData), 3

merge, 4

merge, OptimizedHierarchy, OptimizedHierarchy-method
(merge), 4

merge-methods (merge), 4

OverlapScores (HIVData), 3

plot, 4

plot, OptimizedHierarchy, ANY-method
(plot), 4

plot, OptimizedHierarchy-method (plot), 4

RchyOptimyx, 6

RchyOptimyx, ANY, numeric, character, numeric, logical-method
(RchyOptimyx), 6

RchyOptimyx-methods (RchyOptimyx), 6

RchyOptimyx-package, 2

Signs (HIVData), 3

summary, OptimizedHierarchy-method
(summary-methods), 7

summary-methods, 7