

Package ‘rds’

May 29, 2024

Type Package
Title Root diversity score (RDS)
Version 0.1.2
Date 2024-05-11
Author Johannes Reiter, Martin Blohmer, Alexander Gorelick
Maintainer Alexander Gorelick <alexander_gorelick@hms.harvard.edu>
Description R implementation of root diversity score (RDS) calculation (Reiter et al, Nature Genetics, 2020). Translated from the original python by Martin Blohmer. Packaged into an R library and maintained by Alexander Gorelick. Please cite: Reiter et al., Nature Genetics, 2020. <https://doi.org/10.1038/s41588-020-0633-2>
License MIT
RoxygenNote 7.3.1

R topics documented:

calculate_no_common_trees	1
calculate_no_trees	2
calculate_rds	2
cluster_probability	3
min_cluster_probability	3
multinomial	4
no_trees_min_cluster	4
no_trees_with_clusters	5
rds	5
Index	6

calculate_no_common_trees
<i>calculate_no_common_trees</i>

Description

Calculate the number of monophyletic trees given by Theorem 3 in the SI of Reiter et al. 2020 Nat Gen.

Usage

```
calculate_no_common_trees(k, m)
```

Arguments

k	number of remaining samples
m	number of metastases samples

Value

number of monophyletic trees

calculate_no_trees	<i>calculate_no_trees</i>
--------------------	---------------------------

Description

Calculate the total number of distinct trees given by Edwards and Cavalli-Sforza, 1963

Usage

```
calculate_no_trees(n)
```

Arguments

n	number of samples
---	-------------------

Value

number of distinct trees

calculate_rds	<i>calculate_rds</i>
---------------	----------------------

Description

Calculates the root diversity scores as the probability that a tree where l metastases form a cluster (have one root) exists by chance as defined by Equation (2).

Usage

```
calculate_rds(k, m, l, use_gmp = T)
```

Arguments

k	number of primary tumor region samples
m	total number of metastases samples
l	number of metastases that actually cluster together
use_gmp	Use gmp R package to calculate RDS for large trees. Required for trees with more around 30 tips (default=T)

Value

probability that a cluster of size l arises by chance

cluster_probability	<i>cluster_probability</i>
---------------------	----------------------------

Description

Calculate the probability that a tree where all m metastases form a cluster (have one root) exists by chance

Usage

cluster_probability(k, m)

Arguments

- k number of remaining cancer samples of the same phylogeny
- m number of metastases samples

Value

probability of cluster by chance

min_cluster_probability	<i>min_cluster_probability</i>
-------------------------	--------------------------------

Description

Calculates the probability that a tree where all m metastases form a cluster (have one root) exists by chance

Usage

min_cluster_probability(k, m, l, use_gmp)

Arguments

- k number of primary tumor region samples
- m total number of metastases samples
- l number of metastases that actually cluster together

Value

probability that a cluster of size l arises by chance

multibinomial

multibinomial

Description

Calculate the product of the binomials

Usage

```
multibinomial(m, l, cluster_numbers)
```

Arguments

m	number of remaining cancer samples
l	size of observed metastases cluster
cluster_numbers	vector of possible coexisting clusters of size l for a given phylogeny

Value

product of binomials

no_trees_min_cluster

no_trees_min_cluster

Description

Calculates the number of trees where m metastases form a cluster of size at least l. Implements Formula (S1) from SI.

Usage

```
no_trees_min_cluster(k, m, l)
```

Arguments

k	number of primary tumor region samples
m	number of metastases samples
l	minimal cluster size

Value

number of distinct tree topologies where l out of m metastases form a cluster

```
no_trees_with_clusters
no_trees_with_clusters
```

Description

Calculates the number of trees where m1, m2, ..., mC, metastases form a cluster (have one root)

Usage

```
no_trees_with_clusters(k, clusters)
```

Arguments

k number of remaining cancer samples of the same phylogeny
clusters vector of number of metastases samples per C clusters

Value

number of trees where the given clusters have independent origin

```
rds                      rds
```

Description

Main function to run RDS calculations on a cancer phylogeny (phylo object). Must use Naxerova Lab Normal/Primary/Metastasis tip label conventions.

Usage

```
rds(phy, drop_na = T, use_gmp = T)
```

Arguments

phy A phylo-class cancer phylogeny
drop_na Omit rows in the output data.frame with NA RDS values (default=T)
use_gmp Use gmp R package to calculate RDS for large trees. Required for trees with more around 30 tips (default=T)

Value

data.frame where rows are metastasis sample-types and columns are k, l, m, and RDS values.

Index

calculate_no_common_trees, [1](#)
calculate_no_trees, [2](#)
calculate_rds, [2](#)
cluster_probability, [3](#)

min_cluster_probability, [3](#)
multibinomial, [4](#)

no_trees_min_cluster, [4](#)
no_trees_with_clusters, [5](#)

rds, [5](#)