

Package ‘CongreveLamsdell2016’

January 23, 2019

Version 1.0.0

Date 2019-01-17

Title Distance metrics for trees generated by Congreve and Lamsdell (2016)

Description Includes the 100 datasets simulated by Congreve and Lamsdell (2016) <doi:10.1111/pala.12236>, and analyses of the partition and quartet distance of reconstructed trees from the generative tree, as analysed by Smith (2019).

URL <https://github.com/ms609/CongreveLamsdell2016>

BugReports <https://github.com/ms609/Quartet/issues>

Copyright Data from Congreve & Lamsdell (2016) released under a CC0 license <doi:10.5061/dryad.7dq0j/1>.

License GPL (>= 2)

Encoding UTF-8

Language en-GB

Depends R (>= 3.4.0)

Imports Ternary

Suggests ape,
bookdown,
knitr,
phangorn,
Quartet,
rmarkdown,
TreeSearch (> 0.2.0),
usethis

LazyData true

ByteCompile true

VignetteBuilder knitr

RoxygenNote 6.1.1

Roxygen list(markdown = TRUE)

R topics documented:

clCI	2
clColours	2
clInitializeTernaryQuarts	3
clPhyDat	4
clPlotQuartets	4
clReferenceTree	6
clResults	7
Index	10

clCI	<i>Consistency indices</i>
------	----------------------------

Description

Consistency indices of Congreve & Lamsdell datasets.

Usage

clCI

Format

An object of class `numeric` of length 100.

clColours	<i>Default colours for analyses.</i>
-----------	--------------------------------------

Description

Default colours for analyses.

Usage

clColours

Format

An object of class `character` of length 8.

`clInitializeTernaryQuarts`*Initialize ternary plots for quartet plotting*

Description

Sets up a blank ternary plot ready for analytical results to be added.

Usage

```
clInitializeTernaryQuarts(zoom = 1, padding = 0.1, gridLines = 10,  
  fontSize = 1, gridCol = "#EBEBEB", backgroundCol = "#FBFBFD",  
  xLim = c(0, 1/zoom) - 0.01, yLim = c(0.5 - (1/zoom), 0.5),  
  isometric = TRUE)
```

```
clInitializeTernarySplits(fontSize = 1, xLim = NULL, yLim = NULL,  
  gridCol = "#EBEBEB", backgroundCol = "#FBFBFD", padding = 0.1,  
  isometric = TRUE)
```

Arguments

<code>zoom</code>	Level of magnification (times), used to adjust ticks and scale.
<code>padding</code>	Padding, passed to TernaryPlot.
<code>gridLines</code>	Number of grid lines, passed to TernaryPlot as <code>grid.lines</code> .
<code>fontSize</code>	Font size, passed to TernaryPlot as <code>lab.cex</code> .
<code>gridCol</code>	Colour, passed to TernaryPlot as <code>grid.col</code> .
<code>backgroundCol</code>	Background colour, passed to TernaryPlot as <code>col</code> .
<code>xLim, yLim</code>	x and y limits, passed to TernaryPlot as <code>xlim, ylim</code> .
<code>isometric</code>	Logical specifying whether plot should be isometric, passed to TernaryPlot as <code>isometric</code> .

Functions

- `clInitializeTernarySplits`: Initialize ternary plots for partition plotting.

clPhyDat	<i>100 simulated data matrices</i>
----------	------------------------------------

Description

Contains the 100 simulated matrices generated by Congreve & Lamsdell (2016) using a heterogeneous Markov-k model, generated from the [clReferenceTree](#) topology, with all branches sharing an equal length.

Usage

```
clPhyDat

clMatrices
```

Format

- `clPhyDat`: A list with 100 entries, each comprising a `phyDat` object of 55 characters for 22 taxa.
- `clMatrices`: A list with 100 entries, each comprising a list of character tokens for each simulated character, as read from raw nexus files using `ape::read.nexus.data`. The four dummy 'characters' have been removed.

Source

<http://datadryad.org/resource/doi:10.5061/dryad.7dq0j>

References

- Congreve, C. R. & Lamsdell, J. C. (2016). Implied weighting and its utility in palaeontological datasets: a study using modelled phylogenetic matrices. *Palaeontology* 59(3), 447–465. doi: [10.1111/pala.12236](https://doi.org/10.1111/pala.12236).
- Congreve, C. R. & Lamsdell, J. C. (2016). Data from: Implied weighting and its utility in palaeontological datasets: a study using modelled phylogenetic matrices. Dryad Digital Repository. doi: [10.5061/dryad.7dq0j](https://doi.org/10.5061/dryad.7dq0j).

clPlotQuartets	<i>Plot results</i>
----------------	---------------------

Description

Plots the results of the analyses of the Congreve & Lamsdell (2016) datasets.

Usage

```

cIPlotQuartets(dataset, tree, cex = 1.1, pch = 2,
  col = CongreveLamsdell2016::cIColours, ...)

cIPlotAverageQuartets(dataset, cex = 1.1, pch = 2,
  col = CongreveLamsdell2016::cIColours, ...)

cIPlotTheseAverageQuartets(dataset, cex = 1.1, pch = 2,
  col = "black", ...)

cIPlotSplits(dataset, tree, cex = 1.1, pch = 2,
  col = CongreveLamsdell2016::cIColours, ...)

cIPlotTheseAverageSplits(dataset, cex = 1.1, pch = 2, col = "black",
  ...)

cIPlotTheseBestAverageSplits(dataset, cex = 1.1, pch = 2,
  col = "black", ...)

cIPlotAverageSplits(dataset, cex = 1.1, pch = 2,
  col = CongreveLamsdell2016::cIColours, ...)

cIPlotBestAverageSplits(dataset, cex = 1.1, pch = 2,
  col = CongreveLamsdell2016::cIColours, ...)

```

Arguments

dataset	Dataset to plot, for example cIBootGcQuartets .
tree	Integer specifying which tree to plot.
cex, pch, ...	Graphical parameters to pass to JoinTheDots .
col	Named vector specifying colours to use to plot each analysis, named to match <code>names(dataset)</code> .

Value

Returns [invisible](#).

Functions

- `cIPlotAverageQuartets`: Plots average across all 100 trees.
- `cIPlotTheseAverageQuartets`: Plot average for single dataset across all 100 trees.
- `cIPlotSplits`: Splits equivalent of `cIPlotQuartets`.
- `cIPlotTheseAverageSplits`: Splits equivalent of `cIPlotTheseAverageQuartets`.
- `cIPlotTheseBestAverageSplits`: Splits equivalent of `cIPlotTheseBestAverageQuartets`.
- `cIPlotAverageSplits`: Splits equivalent of `cIPlotAverageQuartets`.
- `cIPlotBestAverageSplits`: Splits equivalent of `cIPlotAverageQuartets`.

Author(s)

Martin R. Smith

clReferenceTree

Tree topology for matrix simulation

Description

The tree topology used to generate the matrices in [clMatrices](#) Congreve & Lamsdell (2016).

Usage

clReferenceTree

Format

A single phylogenetic tree saved as an object of class phylo.

Source

Congreve & Lamsdell (2016).

References

- Congreve, C. R. & Lamsdell, J. C. (2016). Implied weighting and its utility in palaeontological datasets: a study using modelled phylogenetic matrices. *Palaeontology* 59(3), 447–465. doi: [10.1111/pala.12236](#).
- Congreve, C. R. & Lamsdell, J. C. (2016). Data from: Implied weighting and its utility in palaeontological datasets: a study using modelled phylogenetic matrices. Dryad Digital Repository. doi: [10.5061/dryad.7dq0j](#).

Examples

```
library(ape) # Contains tree plotting functions
data(clReferenceTree)
plot(clReferenceTree)
```

`clResults`*Congreve and Lamsdell tree distances*

Description

Distance of CL trees from generative tree.

Usage`clBremQuartets``clBremPartitions``clMkvPartitions``clMkvQuartets``clBootFreqPartitions``clBootFreqQuartets``clJackFreqPartitions``clJackFreqQuartets``clBootGcPartitions``clBootGcQuartets``clJackGcPartitions``clJackGcQuartets`**Format**

An object of class `list` of length 7.

Details

For each of the 100 matrices generated by Congreve & Lamsdell (2016), I conducted phylogenetic analysis under different methods:

- Mkv: using the Markov K model in MrBayes;
- eq: using equal weights in TNT;
- k1, k2, k3, k5, kX: using implied weights in TNT, with the concavity constant (k) set to 1, 2, 3, 5, or 10;

- kC: by taking the strict consensus of all trees recovered by implied weights parsimony analysis under the k values 2, 3, 5 and 10 (but not 1).

For each analysis, I recorded the strict consensus of all optimal trees, and also the consensus of trees that were suboptimal by a specified degree.

I then calculated, of the total number of quartets or partitions that were resolved in the reference tree, how many were the same or *different* in the tree that resulted from the phylogenetic analysis, and how many were not resolved in this tree ($r2$).

The data object contains a list whose elements are named after the methods, as listed above.

Each list entry is a three-dimensional array, whose dimensions are:

1. The suboptimality of the tree. Different measures of node support are employed:
 - Mkv: Posterior probabilities, at 2.5% intervals (50%, 52.5%, ... 97.5%, 100%).
 - Brem: Bremer supports: the consensus of all trees that are (equal weights) 0, 1, ..., 19, 20 steps less optimal than the optimal tree (implied weights: the consensus of all trees that are $0.73^{(19:0)}$ less optimal than the optimal tree).
 - Boot: Bootstrap supports (symmetric resampling, $p = 0.33$).
 - Jack: Jackknife supports ($p = 0.36$).

Boot and Jack results are reported both as the frequency of splits among replicates, and using the gc (Groups Present / Contradicted) measure (Goloboff *et al.* 2003); frequency columns correspond to 100%, 97.5%, 95% ... 0% support; gc columns correspond to 100%, 95%, ... 0% present, 5%, 10%, ... 100% contradicted.
2. Counts of the condition of each quartet or partition:
 - Q: The total number of quartets defined on 22 taxa.
 - N: The total number of partitions present, counting each tree separately.
 - P1: The number of partitions in tree 1 (the reconstructed tree).
 - P2: The number of partitions in tree 2 (the generative tree).
 - s: The number of quartets or partitions resolved identically in each tree. * d: The number of quartets resolved differently in each tree.
 - d1: The number of partitions resolved in tree 1, but contradicted by tree 2.
 - d2: The number of partitions resolved in tree 2, but contradicted by tree 1.
 - r1: The number of partitions or quartets resolved in tree 1 that are neither present in nor contradicted by tree 2.
 - r2: The number of partitions or quartets resolved in tree 2 that are neither present in nor contradicted by tree 1.
 - u: The number of quartets that are not resolved in either tree.
3. The number of the matrix, from 1 to 100.

Source

Congreve, C. R. & Lamsdell, J. C. (2016). Implied weighting and its utility in palaeontological datasets: a study using modelled phylogenetic matrices. *Palaeontology* 59(3), 447–465. doi: [10.1111/pala.12236](https://doi.org/10.1111/pala.12236).

References

Goloboff, P. A., J. S. Farris, M. Källersjö, B. Oxelman, M. J. Ramírez, and C. A. Szumik. 2003. Improvements to resampling measures of group support. *Cladistics* 19, 324–332. doi: [10.1016/S07483007\(03\)000604](https://doi.org/10.1016/S07483007(03)000604).

See Also

[clMatrices](#), [clReferenceTree](#).

Index

*Topic **datasets**

- clCI, [2](#)
- clColours, [2](#)
- clPhyDat, [4](#)
- clReferenceTree, [6](#)
- clResults, [7](#)

- clBootFreqPartitions (clResults), [7](#)
- clBootFreqQuartets (clResults), [7](#)
- clBootGcPartitions (clResults), [7](#)
- clBootGcQuartets, [5](#)
- clBootGcQuartets (clResults), [7](#)
- clBremPartitions (clResults), [7](#)
- clBremQuartets (clResults), [7](#)
- clCI, [2](#)
- clColours, [2](#)
- clInitializeTernaryQuarts, [3](#)
- clInitializeTernarySplits
 (clInitializeTernaryQuarts), [3](#)
- clJackFreqPartitions (clResults), [7](#)
- clJackFreqQuartets (clResults), [7](#)
- clJackGcPartitions (clResults), [7](#)
- clJackGcQuartets (clResults), [7](#)
- clMatrices, [6](#), [9](#)
- clMatrices (clPhyDat), [4](#)
- clMkvPartitions (clResults), [7](#)
- clMkvQuartets (clResults), [7](#)
- clPhyDat, [4](#)
- clPlotAverageQuartets (clPlotQuartets),
 [4](#)
- clPlotAverageSplits (clPlotQuartets), [4](#)
- clPlotBestAverageSplits
 (clPlotQuartets), [4](#)
- clPlotQuartets, [4](#)
- clPlotSplits (clPlotQuartets), [4](#)
- clPlotTheseAverageQuartets
 (clPlotQuartets), [4](#)
- clPlotTheseAverageSplits
 (clPlotQuartets), [4](#)

- clPlotTheseBestAverageSplits
 (clPlotQuartets), [4](#)
- clReferenceTree, [4](#), [6](#), [9](#)
- clResults, [7](#)
- invisible, [5](#)
- JoinTheDots, [5](#)