Package 'maxaltall'

July 22, 2025

Type Package

Title 'FASTA' ML and 'altall' Sequences from IQ-TREE .state Files

Version 0.1.0

Maintainer Philip D. Kiser <pkiser@uci.edu>

Description Takes a .state file generated by IQ-TREE as an input and, for each ancestral node present in the file, generates a FASTA-formatted maximum likelihood (ML) sequence as well as an 'AltAll' sequence in which uncertain sites, determined by the two parameters thres_1 and thres_2, have the maximum likelihood state swapped with the next most likely state as described in Geeta N. Eick, Jamie T. Bridgham, Douglas P. Anderson, Michael J. Harms, and Joseph W. Thornton (2017), ``Robustness of Reconstructed Ancestral Protein Functions to Statistical Uncertainty'' <doi:10.1093/molbev/msw223>.

License GPL (>= 3)

Imports data.table, dplyr, magrittr, tidyr

Encoding UTF-8

RoxygenNote 7.3.2

NeedsCompilation no

Author Alec S. Chu [aut],
Philip D. Kiser [aut, cre]

Repository CRAN

Date/Publication 2024-09-25 09:30:02 UTC

Contents

---

max_altall *'FASTA' ML and 'altall' Sequences from IQ-TREE .state Files*

Description

Takes a .state file generated by IQ-TREE as an input and, for each ancestral node present in the file, generates a FASTA-formatted maximum likelihood (ML) sequence as well as an 'AltAll' sequence in which uncertain sites, determined by the two parameters thres_1 and thres_2, have the maximum likelihood state swapped with the next most likely state as described in Geeta N. Eick, Jamie T. Bridgham, Douglas P. Anderson, Michael J. Harms, and Joseph W. Thornton (2017), "Robustness of Reconstructed Ancestral Protein Functions to Statistical Uncertainty" <doi:10.1093/molbev/msw223>.

Usage

```
max_altall(file, type, thres_1, thres_2, export, export_dir)
```

Arguments

file	This argument specifies the IQ-TREE .state file containing the data the function will use.
type	This argument is either 'aa' for amino acid states or 'nuc' for nucleotide states.
thres_1	This argument specifies the probability threshold at which the most probable state will be considered ambiguous enough for possible substitution (as determined by thres_2) by the second most probable state. Permissible values obey the following inequalities:
	$0.05 <= thres\_1 <= 1$ for amino acid states $0.25 <= thres\_1 <= 1$ for nucleotide states
	The default value is 0.8 (see Eick et al).
thres_2	This argument specifies the probability threshold at which the second most probable state will be substituted for the most probable state. Permissible values obey the following inequalities:
	$thres\_2 <= thres\_1$, $thres\_2 <= (1 - thres\_1)$, $thres\_2 >= (1 - thres\_1)/19$ for amino acid states $thres\_2 <= thres\_1$, $thres\_2 <= (1 - thres\_1)$, $thres\_2 >= (1 - thres\_1)/3$ for nucleotide states
	The default value is 0.2 (see Eick et al).
export	This argument is either "TRUE" or "FALSE." A "TRUE" input will cause the sequences to be saved to a new text file called 'node_sequences_all.txt' in a directory specified by the export_dir argument described below. A "FALSE" input will cause the sequences to be printed to the R console. The default value is "FALSE"
export_dir	This argument specifies the directory where the new text file containing the deposited FASTA sequences will be located. The default directory is tempdir().

Details

An IQ-TREE .state file contains posterior probabilities for each possible character state at each position of an amino acid or nucleotide alignment and for each ancestral node of the phylogenic tree used for the ancestral state reconstruction calculation. The purpose of this R script is to extract the maximum likelihood sequence and a user-defined 'AltAll' sequence and output the sequences in FASTA format for facile downstream use. The 'AltAll' sequence concept is described by Eick et al.

Value

Returns a new text file called node_sequences_all.txt with the ML and altall sequences for all nodes present in the input .state file.

Author(s)

Alec Chu and Philip D. Kiser at the University of California, Irvine

References

Paper describing IQ-TREE 1) B.Q. Minh, H.A. Schmidt, O. Chernomor, D. Schrempf, M.D. Woodhams, A. von Haeseler, R. Lanfear (2020) IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol., 37:1530-1534. https://doi.org/10.1093/molbev/msaa015

Online IQ-TREE command reference 2) http://www.iqtree.org/doc/Command-Reference

Paper describing the concept of an AltAll sequence 3) Geeta N. Eick, Jamie T. Bridgham, Douglas P. Anderson, Michael J. Harms, Joseph W. Thornton, Robustness of Reconstructed Ancestral Protein Functions to Statistical Uncertainty, Molecular Biology and Evolution, Volume 34, Issue 2, February 2017, Pages 247–261, https://doi.org/10.1093/molbev/msw223

Examples

```
example_aa.state <- system.file("extdata", "example_aa.state", package = "maxaltall")
max_altall(example_aa.state, "aa", 0.8, 0.2, "TRUE", tempdir())


example_nuc.state <- system.file("extdata", "example_nuc.state", package = "maxaltall")
max_altall(example_nuc.state, "nuc", 0.8, 0.2, "TRUE", tempdir())
```

Index