Infer_ST_MDC_UR Command

Description

Infer the species tree from unrooted gene trees using MDC criterion. The input gene trees must be specified in the Rich Newick Format. Gene trees must be unrooted. The generated output trees will also be generated in the rich newick format.

Usage

```
Infer_ST_MDC_UR (gene_tree_ident1 [, gene_tree_ident2...]) [-e proportion] [-x] [-b threshold] [-a taxa map] [-ur] [-t time] [result output file]
```

gene_tree_ident1 [, *gene_tree_ident2*...]	Comma delimited list of gene tree identifiers. See details.	mandatory
-e *proportion*	Get optimal and sub-optimal trees.	optional
-x	Use all clusters in generation.	optional
-b *threshold*	Specifies bootstrap threshold. Edges in the gene trees that have support lower than *threshold* will be contracted.	optional
-a *taxa map*	Gene tree / species tree taxa association.	optional
-ur	Allow non-binary species tree generation.	optional
-t *time*	Limit search time to *time* minutes.	optional
result output file	Optional file destination for command output.	optional

By default, the method returns the optimal tree. But the option -e allows the users to get the optimal tree and a set of sub-optimal trees. If the optimal tree has *n* extra lineages, all the sub-optimal trees that have extra lineages less than *(1+proportion/100)*n* will be returned with the optimal tree.

By default, the method uses clusters induced from gene trees to infer species tree. However, the option -x allows users to specify using all possible clusters to infer species tree.

If input gene trees have bootstrap values a threshold can be set with the `-b` option.

By default, the method will always return a binary species tree. But users can use option `-ur` to allow non-binary species tree. If the gene trees are not binary and the degree of resolution are low, it is recommended to use this option. Otherwise, the program will do some exhaustive search for a binary species tree. In this case, users can also use option `-t` to limit the search time. The time is in the unit of minutes.

By default, it is assumed that only one individual is sampled per species in gene trees. However, the option `-a` allows multiple alleles to be sampled.

Examples

```
#NEXUS

BEGIN NETWORKS;

Network g1 = ((((a:5,b:5):4,c:9):3,d:12):3,e:15);
Network g2 = ((a:6,b:6):11,((c:12,e:12):2,d:14):3);
Network g3 = ((a:8,c:8):7,((b:14,e:14):1,d:15));

END;


BEGIN PHYLONET;

Infer_ST_MDC_UR (g1, g2, g3);

END
```

```
#NEXUS

BEGIN NETWORKS;

Network g1 = ((((a1::.5,b1::.5)::.5,c::.5)::.5,d::.5)::.5,e::.5)::.5;
Network g2 = ((a2::.5,b2::.5)::.5,((c::.5,e::.5)::.5,d::.5)::.5)::.5;
Network g3 = ((a::.5,c::.5)::.5,((b::.5,e::.5)::.5,d::.5)::.5)::.5;

END;


BEGIN PHYLONET;

InferST_MDC_UR (g1, g2, g3) -b .5 -e .2 -x -ur -t 1 -a <z:a1,a2,a; y:b1,b2,b; c:c; d:d; e:e>;

END;
```

Command Refernces

- Y. Yu, T. Warnow, and L. Nakhleh. Algorithms for MDC-based multi-locus phylogeny inference. *The 15th Annual International Conference on Research in Computational Molecular Biology (RECOMB)*, pages 531--545, 2011. LNBI 6577.

See Also

- List of PhyloNet Commands