Intoducción a la filoinformática – pan-genómica y filogenómica. IBBM-CONYCET, UNLP Argentina. Julio 2018

https://github.com/vinuesa/get_phylomarkers

Benchmark analyses of the phylogenetic performance of FT vs IQ-TREE

Consensus core-genome (n=239) computed with GET_HOMOLOGUES

Distribution of SH-alrt branch support values of gene-trees found by the FThigh and IQThigh Searches.

Wilcoxon signed-rank test p < 2.2e-16

Distribution of consensus values from majority-rule consensus trees computed from the gene trees passing all the filters, as a function Of search-type.

Kruskal-Wallis p < 0.027

Computation time required FT and IQT when run under "default" (FTdef, IQTdef) and thorough (FThigh, IQThigh) search modes (s. type) on the 239 consensus Clusters

Kruskal-Wallis p < 2.2e-16

Association plot summarizing the results of multi-way Chi-Square analyses of the InL score ranks (1–4, meanin best to worst) of the 105 top-scoring ML gene-trees passing the kdetrees filter in the IQThigh run (Table 2) for each search-type

p < 2.2e-16

Distribution of the edge-lengths of species-trees computed from the concatenated top-scoring markers, as a function of Search-type

Kruskal-Wallis p < 1e-7

Flowchart showing how the GET_PHYLOMARKERS package integrates with GET_HOMOLOGUES

Vinuesa et al., 2018

ML species tree

- 118 genomes
- Top 52(/231) markers
- GTR+ASC+F+R7
- ~8% misclassified genomes 13/169 ~8% unclassified genomes
- reclassified 27 RefSeq genome sequences!
- The Smc is split into 5 major lineages including S. maltophilia s. lato
- 13 clades (species?) are recognized within the latter, which are consistent with a 95.9% cgANIb cutoff

Vinuesa et al., 2018

Intoducción a la filoinformática – pan-genómica y filogenómica. IBBM-CONYCET, UNLP Argentina. Julio 2018

Unsupervised learning methods to find groups within the Smc and <i>S. maltophilia s. lato</i> lineages based on cgANDb (= 100 - cgANlb)	
up to 13 potential species resolved within Smsl at cqANDb = 4.1 (cqANlb = 95.9)	
	Vinuesa et al., 2018

ML pan-genome tree	
 118 genomes 29,623 markers GTR2+F0+R4 	
Unsupervised learning methods to find species-like clusters in the Smc based on the pan-genome matrix	

Vinuesa et al., 2018