If you use these data, please cite:

Jackson, AP et al. 2012. A cell-surface phylome for African trypanosomes, *manuscript submitted.*

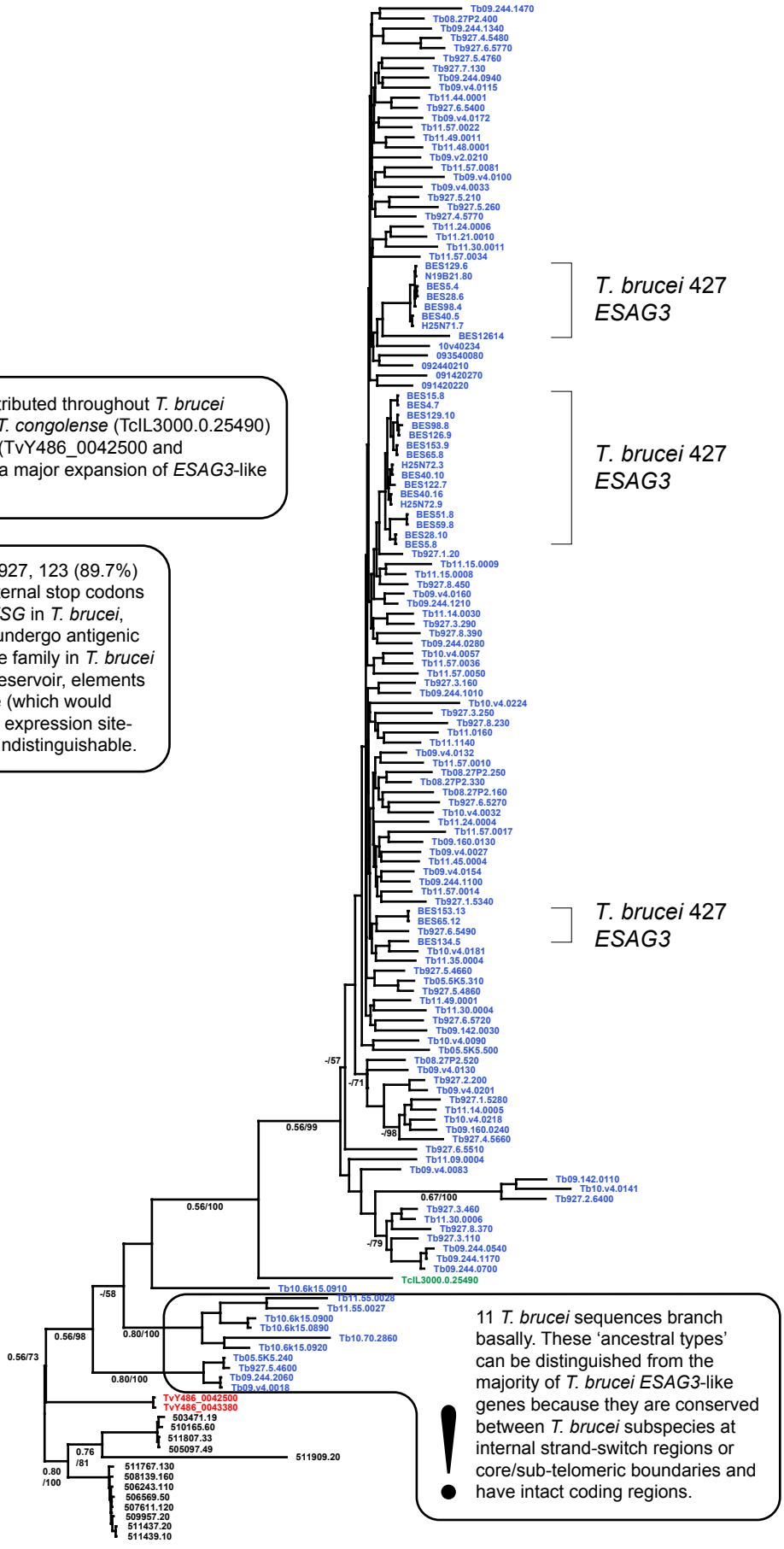
Fam53:
Expression site-associated gene 3-like genes

Fam53 includes 137 *ESAG3*-like genes distributed throughout *T. brucei* sub-telomeric regions, a single homolog in *T. congolense* (TcIL3000.0.25490) and two highly similar homologs in *T. vivax* (TvY486_0042500 and TvY486_0043380). Hence, there has been a major expansion of *ESAG3*-like genes has occurred uniquely in *T. brucei*.

Of 137 sub-telomeric *ESAG3*-like genes in *T. brucei* 927, 123 (89.7%) are fragmentary or predicted pseudogenes due to internal stop codons or frameshifts. This is reminiscent of sub-telomeric *VSG* in *T. brucei*, 72% of which are pseudogenes. *ESAG3* might also undergo antigenic variation like *VSG*. In which case, the expanded gene family in *T. brucei* would represent a transcriptionally-silent, structural reservoir, elements of which periodically transpose to the expression site (which would account for the polyphyly of *ESAG3* and the fact that expression site-associated and sub-telomeric genes are structurally indistinguishable.

Key:
Tb927. *T. brucei*
TvY486_ *T. vivax*
TcIL3000. *T. congolense*
500000.1 *T. cruzi*

0.1 sub/site



NOTES:

Fam53 includes *ESAG3* sequences from *T. brucei* 427 bloodstream expression site sequences as well as homologous *ESAG3*-like genes from *T. brucei* 927 core chromosomes, and other species. Note that most *T. brucei* homologs are predicted pseudogenes.

The maximum likelihood phylogram was estimated from a multiple nucleotide sequence alignment of 882 characters, using PHYML with an LG model with rate heterogeneity. The tree is rooted with *T. cruzi* genes. Selected nodes are supported by posterior probabilities and non-parametric bootstraps generated from a maximum likelihood analysis under a GTR+Γ model.