

If you use these data, please cite:

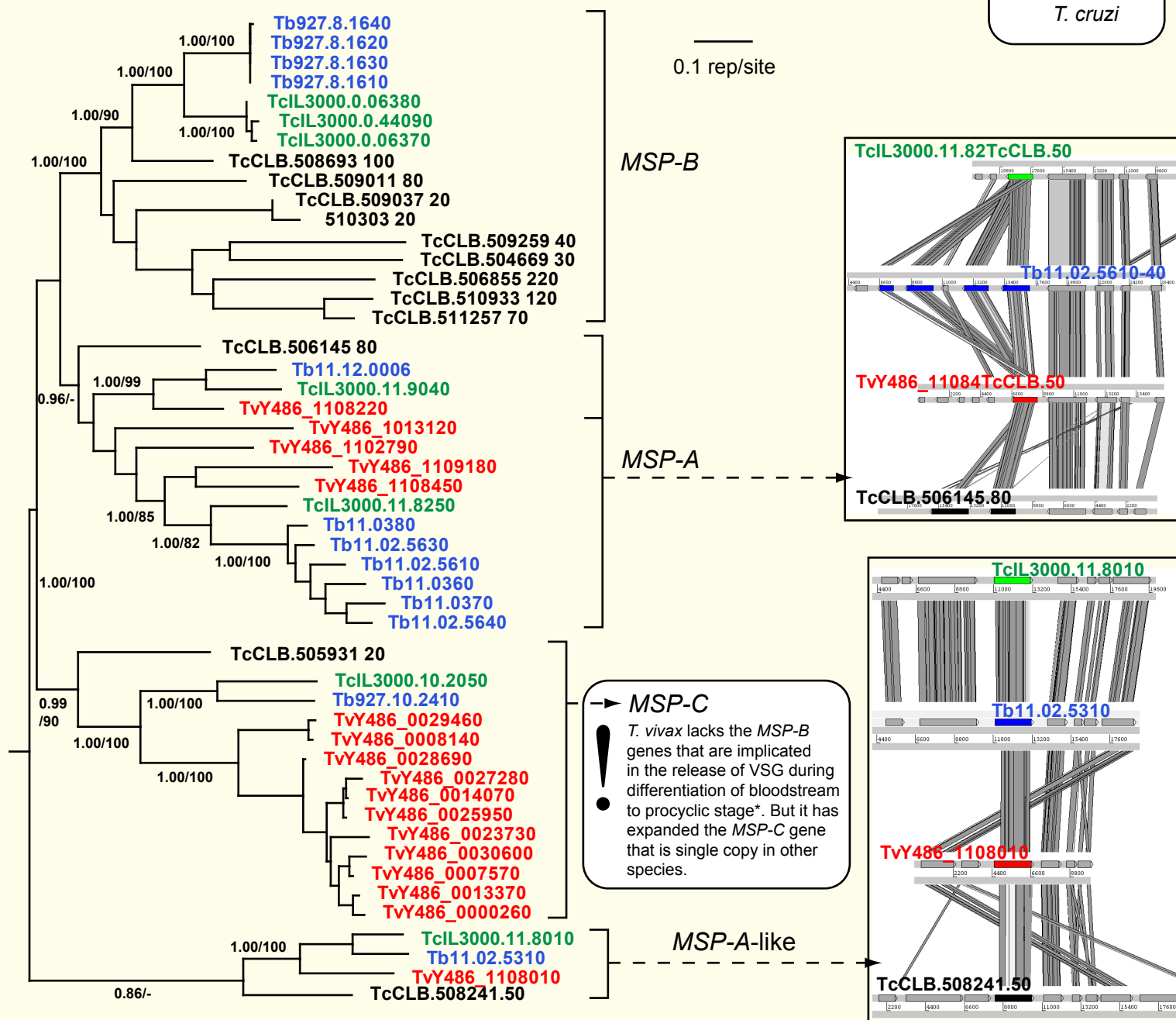
Jackson, AP et al. 2012. A cell surface phylome for African trypanosomes. *Manuscript submitted.*

# Fam46: Major surface protease



*MSP* genes from *Trypanosoma* fall into four clades. (i) *MSP-A* is arranged as a polymorphic tandem gene array in *T. brucei*, with single genes conserved at the same position in other species. (ii) *MSP-B* is arranged in a polymorphic tandem gene array in *T. brucei* and *T. congolense*, and is present at multiple species-specific positions in *T. cruzi*. *MSP-B* is absent from *T. vivax*. (iii) *MSP-C* is a single-copy gene positionally conserved in *T. brucei* and *T. congolense* but multiplied substantially in *T. vivax* at species-specific positions. (iv) A fourth locus annotated as '*MSP-A-like*' is a single copy gene conserved in position in all species. These genes branch basally within the phylogeny and are most closely related to *MSP* homologs in *Leishmania*.

Key: **Tb927.**  
*T. brucei*  
**TvY486\_**  
*T. vivax*  
**TcIL3000.**  
*T. congolense*  
**TcCLB.**  
*T. cruzi*



**NOTES:** Fam46 comprises predicted major surface protease genes (i.e. *gp63*) from various loci throughout genomes from all species. All members encode 9-11 transmembrane helices; several members have predicted signal peptides. The Bayesian phylogram was estimated from a multiple protein sequence alignment of 485 characters, using MrBayes under default settings. The tree is rooted with orthologs from the divergent *MSP-A-like* locus (TcCLB.508241.50). Selected nodes are supported by posterior probability values and non-parametric bootstraps generated from a maximum likelihood analysis using an LG model with rate heterogeneity. \* Gruszyński et al. 2003. *J Biol Chem.* **278**:24665-72.