If you use these data, please cite:

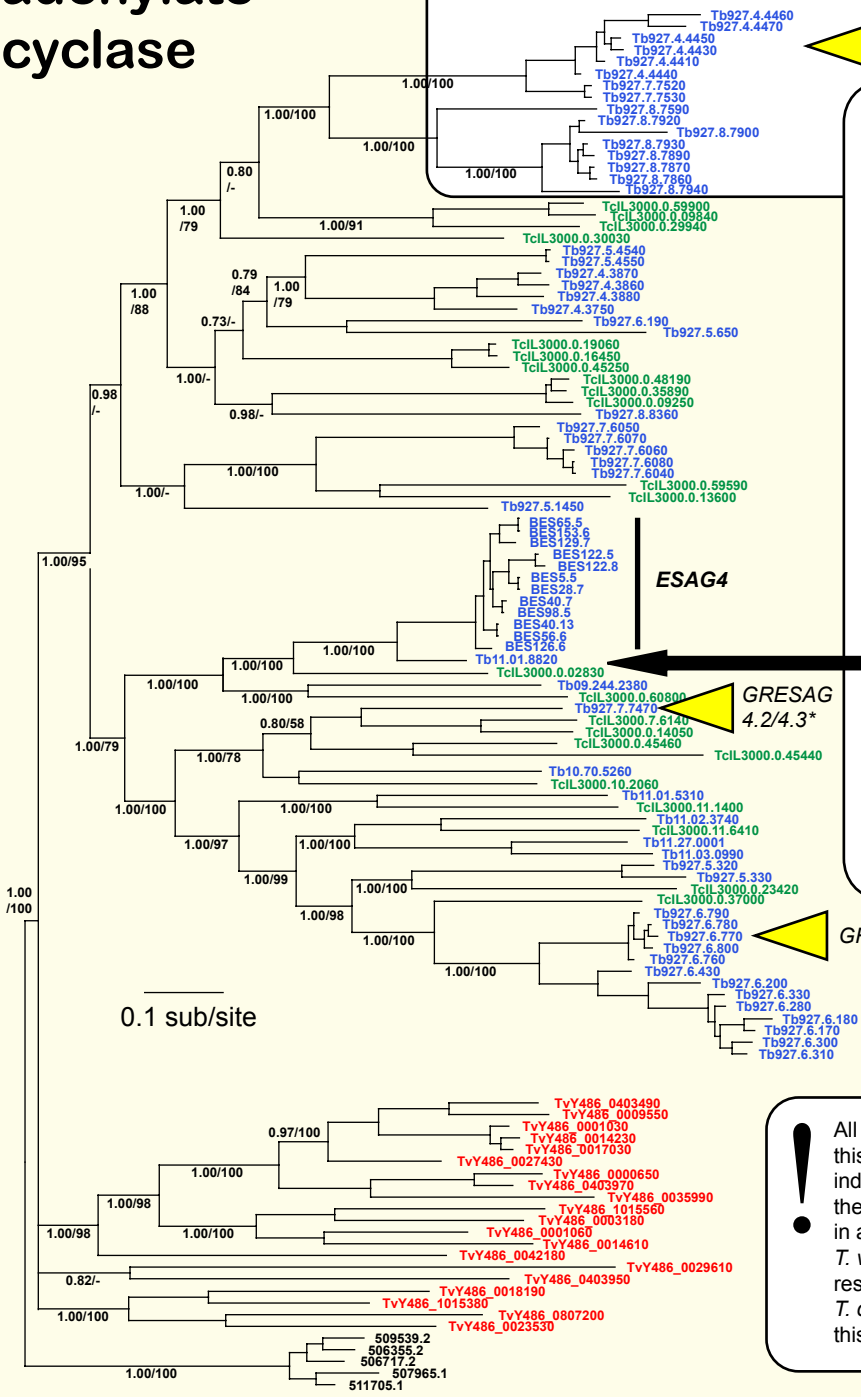
Jackson, AP et al. 2012. A cell-surface phylome for African trypanosomes, *manuscript submitted.*

Fam51: **Receptor-type adenylate cyclase**

These *T. brucei*-specific genes have originated through a segmental duplication affecting chromosomes 4 and 8. This event duplicated a large tandem gene array of adenylate cyclase genes, resulting in an array on each chromosome. For more information, go to:
http://www.genedb.org/Page/48dup_image

Key:

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



ESAG4 is just one lineage of receptor-type adenylate cyclase gene, which are otherwise common features in trypanosomatids. *ESAG4* is one of multiple expansions of adenylate cyclases in *T. brucei* (see above). Otherwise, most lineages are present in *T. congolense* and *T. brucei* and these are positionally conserved.

T. congolense has no orthologs to *ESAG4*. The most closely related core chromosomal locus to *ESAG4* is Tb11.01.8820, a locus at the sub-telomeric boundary of chromosome 11. This gene does have a *T. congolense* ortholog (TcIL3000.0.02830), which is found in the same genomic position. TcIL3000.0.02830 is identical to the single adenylate cyclase sequence cloned by Alexandre *et al*. (1996) from a *T. congolense* cDNA library (Z67964); so it may be highly expressed among *T. congolense* adenylate cyclases.

Therefore, we can place the likely origin of *ESAG4* at a conserved sub-telomeric locus on chromosome 11. Comparison of Tb11.01.8820 and *ESAG4* sequences shows that this remodelling has primarily affected the catalytic and intracellular domains. 245 amino acid differences are distributed preferentially towards the C-terminal, with 69% occurring after the putative transmembrane helix (a portion accounting for only 35% of total characters). Furthermore, of 54 sites where Tb11.01.8820 and TcIL3000.11.16970 are conserved, but *ESAG4* is derived (i.e. unambiguous *ESAG4* apomorphies), 41 occur in the intracellular domain.

All *T. vivax* genes are monophyletic. A literal interpretation of this topology would indicate that diversification has occurred independently in *T. brucei/T. congolense* and *T. vivax*. Given that these genes encode diverse families with core physiological roles in all trypanosomatids, this is unlikely. Rather the monophyly of *T. vivax* and *T. cruzi* genes suggests a high rate of gene turnover resulting in the loss of orthology over time. In *T. brucei* and *T. congolense*, where species-specific expansions have begun this process, of which *ESAG4* is a part.

NOTES: **Fam51** includes the expression-site associated adenylate cyclase *ESAG4* from *T. brucei* 427, as well as related genes from *T. brucei* 927 ('*GRESAG4*') and other species. Copy number of tandemly arrayed adenylate cyclase genes is imprecise.* Alexandre *et al. Mol Biochem Parasitol.* **43**, 279-288 (1990). ** Naula *et al. Mol. Biochem. Parasitol.* **112**, 19-28 (2001). Published *GRESAG4* are indicated by yellow arrows.

The Bayesian phylogram was estimated from a multiple protein sequence alignment of 247 characters, using MrBayes under default settings. The tree is mid-point rooted with *T. cruzi* sequences (only a fraction of total *T. cruzi* sequences shown). Selected nodes are supported by posterior probabilities and non-parametric bootstraps from a maximum likelihood analysis under an LG model with rate heterogeneity.