

**Comparative genomics of the three model trypanosomatids:
Trypanosoma brucei, *Trypanosoma cruzi* and *Leishmania***

Najib M. El-Sayed¹, Peter Myler, Gaëlle Blandin, Neil Hall & the TriTryp consortium

¹ Department of Parasite Genomics, The Institute for Genomic Research (TIGR); Department of Microbiology and Tropical Medicine, School of Medicine and Health Sciences, The George Washington University Medical Center.

The TriTryp sequencing consortium has completed the first draft genome sequences of the three reference Trypanosomatids – *Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania major*. While these three pathogens share many general characteristics, each organism presents distinct aspects, including transmission by a different insect, unique life cycle features, different target tissues and immune evasion mechanisms, as well as different disease characteristics in their mammalian host, all of which are reflected by differences in their genome sequences.

In an effort to characterize the core TriTryp proteome as well as genes that are species-specific, we have identified the orthology relationships between individual genes of the three genomes. Clusters of orthologous genes (COGs) were constructed and used to compare gene content as well as genome architecture. Our results show that regions encoding the TriTryp core proteome present a remarkably high degree of synteny (conservation of gene order) while in *Trypanosoma*, non-syntenic regions have been expanded and harbor many large species-specific gene families, majority of which appear to be surface antigen families. The different numbers of such species-specific genes largely reflect the different strategies of immune evasion used in each organism. Examination of large-scale synteny also sheds some light on chromosome evolution history in Trypanosomatids and emphasizes the role of multigene families, mobile elements or structural RNAs in shaping the genomic landscape of the present day parasites. We will also report on our analysis of the composition and organization of protein domains encoded by each genome, and an analysis of the rates of evolution for selected gene families.