Episomal gene amplification in natural populations of *Leishmania donovani*

An Mannaert[1], Hideo Imamura[1], Manu Vanaerschot[1], Tim Downing[2], Matt Berriman[2], Shyam Sundar[3], Suman Rijal[4], Jean-Claude Dujardin[1]

[1]Institute of Tropical Medicine, Antwerpen, Belgium; [2]Wellcome Trust Sanger Institute, Hinxton, UK; [3]Baranas Hindu University, Varanasi, India; [4]BP Koirala Institute of Health Sciences, Dharan, Nepal

Leishmania donovani is the major cause of visceral leishmaniasis, which affects the life of millions, especially in the Indian subcontinent. Recent whole-genome sequencing of Indian and Nepalese clinical isolates of *L. donovani* showed a high level of sequence conservation, but a large amount of gene copy number variation. This structural variation includes massive aneuploidy, with different chromosome copy numbers in the different lines, and expansion or contraction of tandem repeats. In addition, several genes were amplified as extra-chromosomal, circular episomes. The presence of two episomes was predicted in a population of 17 drug resistant and drug sensitive clinical *L. donovani* lines by analysis of whole-genome sequences and experimentally verified. One of the episomes is the H-locus containing the ABC transporter gene MRPA, which is already known to appear in experimental conditions such as induction of drug resistance. The other episome is new and carries a MAP kinase homolog (MPK1) and an acid phosphatase gene. To assess the occurrence of these episomes in the field, we expanded the analysis to >100 clinical lines and found that both episomes are present in most lines and that their copy number is linked with treatment failure of visceral leishmaniasis in the corresponding patients. In a second step, we tested *in vitro* and *in vivo* stability of these episomes, as the MRPA-containing episome is known to appear in laboratory strains with induced drug resistance, but quickly disappears upon release of drug pressure. To determine whether both episomes are stable throughout the life cycle and whether their gene expression correlates with DNA copy number, the copy number and expression levels of the MRPA and MPK1 genes in the logarithmic and stationary growth phase of promastigotes and in amastigotes were quantified by RT-PCR. We selected five field isolates of *L. donovani* with differential episome copy number and drug (SSG) tolerance as determined a priori by read depth analysis of whole-genome sequences (two strains with high episome copy numbers, two strains with lower copy numbers and one strain lacking both episomes). Our results, combined with phenotypic data from the parasites provide the basis to further explore the function of these episomes. Moreover, they highlight the importance of the choice of the life stage of *Leishmania* for analyses of structural variation in genome studies, as copy numbers – not only of episomes, but also of genomic DNA fragments and chromosomes – can fluctuate over the life cycle.