Episomal gene amplification in natural populations of *Leishmania donovani*

An Mannaert[1], Hideo Imamura[1], Manu Vanaerschot[1], Tim Downing[2], Matt Berriman[3], Shyam Sundar[4], Suman Rijal[5], Jean-Claude Dujardin[1]

[1]Institute of Tropical Medicine, Antwerpen, Belgium; [2]National University of Ireland, Galway, Ireland; [3]Wellcome Trust Sanger Institute, Hinxton, UK; [4]Baranas Hindu University, Varanasi, India; [5]BP Koirala Institute of Health Sciences, Dharan, Nepal

Leishmania donovani is the major cause of visceral leishmaniasis, a disease afflicting millions worldwide, 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 and evidence 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 are 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 genomic read coverage variability 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 newly discovered and carries a MAP kinase homolog (MPK1) and an acid phosphatase gene (Downing et al, 2011).

To assess the occurrence of these episomes in the field, we expanded the analysis to >100 clinical lines and determined the copy number of both episomes by quantitative PCR. In a second step, we tested *in vitro* and *in vivo* stability of these episomes, because 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 if both episomes are stable throughout the life cycle of *Leishmania* and if 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 (fly stage) and in amastigotes (human stage) were quantified by RT-PCR. We selected five field isolates of *L. donovani* with differential episome copy number, as determined a priori by read depth analysis of whole-genome sequences, and drug (sodium stibogluconate) tolerance.

We found that the two episomes are present in 94% of the lines and that their copy number is linked with treatment failure of visceral leishmaniasis in the corresponding patients. The episomes are stable during promastigotes growth stages, but amastigotes of drug resistant strains appear to have higher copy numbers. The number of transcripts of the episomal genes correlates with the genomic copy numbers, which is not necessarily straightforward in *Leishmania*, since it partially relies on mRNA degradation as a mechanism to regulate gene expression instead of transcriptional control.

Our results provide a basis to further explore the function of these episomes and highlight the importance of the choice of the host or vector life stage of *Leishmania* for analyses of structural variation in genome studies, as copy numbers of episomes, genomic DNA fragments and chromosomes can fluctuate over the life cycle.

Keywords: *Leishmania donovani*, episome, structural variation