

## In the Shadow of the Genome: A Challenging Journey to Diversity in *Leishmania donovani*

Hideo Imamura<sup>1</sup>, An Mannaert<sup>1</sup>, Tim Downing<sup>2</sup>, Matthew Berriman<sup>2</sup>, Jean-Claude Dujardin<sup>1</sup>

<sup>1</sup>Institute of Tropical Medicine, Antwerp, Belgium;

<sup>2</sup>Wellcome Trust Sanger Institute, Hinxton, UK

<http://www.itg.be/>

Leishmaniasis is a disease complex caused by protozoan parasites of the genus *Leishmania*, which are transmitted by sandflies. It affects 350 million people worldwide, but the most severe form, visceral leishmaniasis (VL), is most prevalent in India, Nepal, Bangladesh, Sudan and Brazil. In the Indian subcontinent, VL is caused by *Leishmania donovani*, and efficient treatment is highly challenged by the emergence of drug resistance. We are running two projects – Kaladrug (2) and GeMInI (3) – to characterize the genetic and phenotypic diversity of *L. donovani* in India and Nepal using comparative genomics and metabolomics in order to identify genetic and metabolic signatures associated with drug resistance. Both omics represent two extremes, from genotype to phenotype. The ultimate goal is to identify the factors that lead to the different clinical phenotypes (cure versus treatment failure). Therefore, different strains have been and are being subjected to whole-genome sequencing and metabolic profiling. Sequencing, genome assembly and comparative analyses are performed in collaboration with the Parasite Genomics group of the Wellcome Trust Sanger Institute. One phenotypically well-characterized *L. donovani* strain was chosen as a reference and a *de novo* draft genome sequence was generated with 454 and Illumina sequencing technologies, resulting in a 35 Mbp genome distributed over 36 chromosomes. Approximately 50 additional strains with known in vitro drug susceptibility from VL patients with differential response to treatment are sequenced and analyzed to identify natural variation. Despite high sequence conservation and thus a limited number of single nucleotide polymorphisms (SNPs), a substantial amount of structural variation has been observed among the different strains (1). The chromosome ploidy is highly variable between different strains, such that all strains examined so far have a different chromosome content, and contractions and expansions of tandem gene repeats appear to occur frequently. In addition, we observed extra-chromosomal amplification of three gene loci. These structural polymorphisms can result in a change in the gene dosage and can have an effect on the metabolome and thus the phenotype of the parasite. Preliminary genome analyses identified a number of SNPs and structural changes that may contribute to the resistant phenotype, and the first metabolome analyses of the same samples revealed a significant difference in metabolites between drug susceptible and drug resistant strains. The amount of information generated by next-generation sequencing and other technologies is growing, as is the challenge to process and interpret the increasing amount of data. The major task here is the integration of the information coming from both ends of the omics chain in order to understand how complex biological traits, such as drug resistance, are acquired.

### References

1. Downing, Imamura et al. Whole genome sequencing of multiple *Leishmania donovani* clinical isolates provides insights into population structure and mechanisms of drug resistance. *Genome Research*, in press.

### Relevant Web sites

2. <http://www.leishrisk.net/kaladrug>
3. <http://www.leishrisk.net/gemini>