

Tb035-SNP typing reveals similarity in *Mycobacterium tuberculosis* genetic diversity in Portugal and Northeast Brazil

Lopes J S^{1*}, Marques I¹, Soares P¹, Guimaraes H¹, Costa J¹, Miranda A², Alves A², Macedo R³, Antunes A F³, Duarte T A⁵, Barbosa T⁴, Nery J S⁴, Boechat N⁵, Pereira S M⁶, Barreto M L⁶, Pereira-Leal J¹, Gomes M G M¹, Penha-Goncalves C¹

1-Instituto Gulbenkian de Ciencia, Oeiras, Portugal, 2-Instituto Nacional de Saude Dr. Ricardo Jorge, Porto, Portugal, 3-Direccao Geral da Saude, Lisboa, Portugal, 4-Fundacao Oswaldo Cruz, Salvador, Brazil, 5-Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, 6-Universidade Federal da Bahia, Salvador, Brazil

Introduction: Human tuberculosis is an infectious disease primarily caused by *Mycobacterium tuberculosis* (Mtb). At the moment, the gold standard of epidemiological genotyping of Mtb is Spoligotyping and MIRU-VNTR. These markers, however, have limited use for phylogenetics and strain identification since their fast substitution rate leads to convergent evolution. Recent studies have introduced the use of single nucleotide polymorphisms (SNP) data for identification of Mtb strain groups. In this work we use SNP typing techniques to characterize and compare Mtb populations from Portugal and Brazil. We complement this analysis by discussing the benefits and caveats of this technique. **Material and Methods:** We defined as set of 79 previously described SNPs to characterize Mtb samples. Using this set we analyze 1915 samples collected in Portugal. From these, 111 were further characterized using Spoligotypes. The set of SNPs was also used to analyze 141 samples from Brazil. All of these were further characterized using Spoligo types. **Results:** The most frequent bacterial lineage in both countries is lineage 4 corresponding to around 95% of the analyzed samples. Lineage 1 is also present in both populations equally, but in considerably low frequencies (around 1%). Contrastingly, Mtb bacteria of lineage 2 are only present in the Portuguese population with a frequency of almost 3%. Within lineage 4, the most frequent strain group in both Portugal and Brazil are LAM, followed by Haarlem and then X. Contrarily to these groups, strain group T shows a very different prevalence between Portugal and Brazil, with a frequency of 7% and less than 1.5%, respectively. A comparison between strain identification using Spoligotypes and SNPs show that the former marker misidentifies more than 11% of the samples and are unable to identify almost 1%. Furthermore, SNP typing, as opposed to Spoligotypes can provide phylogenetic relationships between the strain group. By close examination of the genotypes of our samples we seem to observe signs of often disregarded recombination events. **Main Conclusions:** Overall, the use of SNP typing reveals striking similarities between Mtb populations from Portugal and Brazil. We also show that this technique can be useful to reveal signs of recombination events in Mtb. **E-mail:** j.sollari.lopes@gmail.com