Review of: "Whole-genome single nucleotide variant phylogenetic analysis of Mycobacterium tuberculosis Lineage 1 in endemic regions of Asia and Africa"

ro Zenteno-Cuevas

Potential competing interests: The author(s) declared that no potential competing interests exist.

In this article Netikul et al. made a review and reorganized the sublineages present in Lineage 1 (L1) of Mycobacterium tuberculosis (Mtb), using 1764 genomes, mostly from public repositories. The large number of isolates included in this study confers an important value to the results generated. Undoubtedly, the results of the study are of great value in the area, because they address the L1, little characterized, but widely dispersed in regions such as India, Oceania, Asia and Africa. In addition to the phylogeographic analysis of the sublineages found, a correlation of the information of these sublineages with the spoligotyping patterns inferred from the genome, the analysis of the SNPS found in genes associated with drug resistance, and finally an analysis of genetically clustered genotypes were performed.

The phylogeny analysis of the L1 sublineages showed the presence of three major and 24 minor sublineages. One of the most interesting contributions of the study is that it provides a redefinition of several of the minor sublineages, this definition was achieved thanks to the SNPS used for this analysis, as well as, the large number of genomes that were used. Undoubtedly, this reclassification will be an important reference for subsequent studies involving phylogenomic analysis with this lineage.

The correlation analysis between genomically defined sublineages and the spoligotyping patterns showed the absence of certain genetic regions that would explain the spoligotyping patterns and their location in specific sublineages. An additional data generated in the combined analysis with spoligotyping was the limited value of the DVR62 marker and its possible discard as a reference marker in future studies. On the other hand, a better phylogenetic relationship between certain spoligotypes was achieved, which could facilitate the development of a better grouping in future studies, considering that this technique remains as an important tool to genotypically characterize tuberculosis in many of the regions considered in the study.

The phylogeographic analysis was extremely interesting, since it shows in detail how certain lineages differentiate, disperse, and show a dominant dispersion in certain regions or countries. The level of detail that can be appreciated confirms the value of genomic analysis to determine the geographic distribution of the sublineages that occur in a region and the evolution of Mtb according to the social groups in which it occurs.

Perhaps the analysis of mutations associated with drug resistance was the most limited, since it showed very general information on the percentages of resistance of the total number of isolates analyzed, as well

as the SNPS that were characterized. For rifampicin resistance alone, 31 alleles in the rpoB gene were described, but no detailed description of these polymorphisms or the amino acids with the greatest number of changes is given. This aspect is taken up in the discussion as a limitation of the study considering that the genomes recovered were obtained in the context of other types of studies.

The low number of genetic clustering observed is striking, which would indicate a limited rate of contagious; however, this does not agree with the existing epidemiological information on tuberculosis in the regions studied, which are characterized by frequent outbreaks of infection. Unfortunately, the authors did not make further analysis of this information, and only refer in the discussion to the need to increase this type of analysis.

Perhaps the main conclusion of the study is based on the redefinition of the sublineages present in the L1, and locates a geographical point, island Southeast Asia, as a possible site of interest to study the evolutionary history of the L1 lineage. Finally, it confirms the usefulness of WGS for in-depth studies of phylogenomics, phylogeographic, epidemiological and drug resistance characterization in tuberculosis.