The importance of genetic and phenotypic variance of *Mycobacterium tuberculosis* in tuberculosis transmission

The understanding of the mechanisms behind the spread of tuberculosis is crucial in combating this disease and essential for developing effective global control policies by the public health sector. It is known that the transmissibility of tuberculosis is influenced by a range of complex and interrelated factors, which significantly complicates the design and analysis of studies on its spread. These factors include environmental conditions such as temperature and humidity, socioeconomic factors like population density or quick and efficient access to healthcare, as well as host-related factors affecting immune status, such as coexisting diseases (e.g., diabetes) and malnutrition. Important pathogen characteristics, such as its infectiousness or ability to induce a pro-inflammatory response, as well as its genetic background (originating from a particular phylogenetic lineage), are also crucial. However, pathogen-related factors influencing tuberculosis transmissibility have not yet been fully identified and remain the subject of intensive research, due to conflicting results from various studies.

This study aimed to determine the genotypic and phenotypic variability of *Mycobacterium tuberculosis* strains with high and low transmissibility in Poland.

In the first stage of this project, high or low transmission strains were selected from the collection of the National Tuberculosis and Lung Diseases Research Institute in Warsaw based on their spoligotype. A total of 347 candidates for high-transmissibility strains and 44 candidates for low-transmissibility strains were selected. Whole-genome sequencing (WGS) of these strains was then performed, allowing for association analysis and assignment to appropriate clusters based on single nucleotide polymorphism (SNP) differences. Initially, 23 clusters of high-transmissibility strains were identified; however, this number was reduced to 13 due to difficulties in conducting *in vitro* studies. Strains that did not group into clusters were classified as low transmission. The criteria for strain selection are discussed in detail in Chapter 4: Methods. In the next step, selected high transmission strains were matched with low transmission strains, enabling further comparative phenotypic analyses. The strain pairs originated from the same phylogenetic lines and sublines and could not differ by more than a specified number of SNPs. All analyzed strains belonged to phylogenetic lineage 4, except for one strain from lineage

2. Statistical analysis of the variants between the high and low transmission strains strains did not show significant differences between these two groups.

In subsequent stages of the study, phenotypic research focused on identifying differences in growth kinetics between high or low transmission strains. Initially, cultures were conducted under standard laboratory conditions in liquid 7H9 media with OADC supplement, and for control, the strains were also plated on solid 7H10 media with OADC to assess their viability. The results of this analysis did not show differences in growth kinetics between the two groups of strains. Next, an experiment was conducted under limited oxygen conditions, mimicking the *in vitro* environment that the bacteria may encounter in the host's lungs. The strains were cultured for 30 days in an anaerobic medium and then reactivated in a liquid medium to assess their ability to grow after a dormancy period. This analysis also did not reveal significant differences between the high and low transmission strains.

In the next stage, recombinant high-transmissibility strains were obtained using the plasmid pMV306*attP*$_{_ercc3_Mtb}$-*gfp*, containing the *gfp* reporter gene, encoding the green fluorescent protein under the control of a strong promoter from the *ercc3* gene. This allowed for the efficient identification of recombinants using fluorescence microscopy. A total of 9 recombinant high transmission strains were obtained. The plasmid also contained a gene for kanamycin resistance, which allowed for the selection of strains on the 7H10 media supplemented with OADC and kanamycin. Competition assays were then conducted to determine the level of relative fitness by co-culturing high-transmissibility (recombinant) and low transmission strains, as well as absolute fitness assays, where high or low transmission strains were co-cultured with the control strain *M. tuberculosis ΔkatG*. The analysis of these experiments did not reveal statistically significant differences between the high and low-transmissibility strains.

Next, the level of intracellular uptake of selected *M. tuberculosis* strains by macrophages was assessed. This analysis showed statistically significant differences between the studied groups of strains – high transmission strains were taken up to a greater extent than low transmission strains. An experiment on the survival of bacteria inside macrophages was also conducted. The survival analysis did not reveal statistically significant differences between the high and low-transmission strains.

The final stage of the study involved transcriptome analysis of the high and low transmission strains. Bacterial cultures were grown in rich liquid 7H9 media supplemented with OADC, and in a second variant, in only liquid 7H9 media with the addition of 0.01% cholesterol and 0.01% tyloxapol. No specific shared genes with altered expression were observed for either high or low transmission strains.

The obtained results demonstrate that studying transmissibility is a complex process. No statistically significant genotypic or phenotypic differences were observed between the studied groups of strains.