論文の内容の要旨

論文題目　Search for susceptibility gene(s) to tuberculosis in the candidate region of Chromosome 20
(20 番染色体上候補領域における結核の感受性遺伝子の探索)

氏名　中内　彩香

Tuberculosis (TB) is one of the three major world-wide infectious diseases along with acquired immune deficiency syndrome and malaria. Even though a decline in incidence rates was observed in recent years, TB still remains a challenging issue for the world's health, especially due to co-infection with human immunodeficiency virus and the existence of multidrug-resistant TB as well as extensively drug-resistant TB.

It is estimated that approximately one-third of the world's population is infected with the pathogen, *Mycobacterium tuberculosis (M. tb)*, but only 5-10 percent of infected persons will progress to develop the clinical disease. Early studies concluded that TB is a complex disease that both genetic and environmental factors contribute to its development. To gain insight into host genetic factors, a number of genome-wide linkage studies (GWLS), candidate gene studies and genome-wide association studies were carried out. However, not much consistence in the results has been observed among different populations and more genetic studies are needed.

A previous GWLS in Thais identified chromosome 20p13-12.3 as a candidate region for young onset of TB. The present study aims to find any unreported susceptibility genes to young TB within a 1Mbp target region around the top GWLS marker. For this purpose, next-generation sequencing (NGS) was performed on the region in 13 young patients from Thai multi-case families to screen for variants. NGS is a recent novel technology which enables massive and rapid sequencing at comparatively lower cost than conventional methods.

After applying the stringent criteria such as having a threshold 20x coverage in two software to detect variants and consolidating duplicated variants among the sequenced samples as one count, 1,878 variants were detected. Among them, five functionally interesting single nucleotide polymorphisms (SNPs) were selected as candidates for further case-control association analysis: 3 non-synonymous SNPs, 1 SNP that locates in DNaseI hypersensitivity site and in regions with high acetylation of histone H3 lysine 27, as well as 1 SNP in 3' untranslated region (UTR) of a gene whose proxy SNP was reported to have strong association with anemia and thrombocytopenia induced by pegylated interferon and ribavirin therapy for Japanese patients with chronic hepatitis C.

For the case-control association analysis of the SNPs, 665 unrelated case and 777 unrelated control samples from Thai population were used. Among the 5 SNPs tested, a non-synonymous SNP in exon region, SNP A, and a SNP in 3'UTR, SNP B, of *ITPA* showed association (p=0.015; OR=0.71, and p=4.4E-3; OR=0.67, respectively) in young Thai patients (< 45 years old).

To investigate the gene region, three additional tagSNPs were selected for case-control association analysis and none of them surpassed significance level of SNP A and SNP B. In order to improve statistical power to see if the two SNPs are possibly significant, genotyping with an additional sample set, which included 545 unrelated cases and 407 unrelated controls was performed. After the analysis, SNP A and SNP B showed lower p-values (p=1.3E-03; OR=0.72, and p=5.1E-5; OR=0.66, respectively) in young patients. In the studied young Thai population, the two SNPs showed high linkage disequilibrium (r^2=0.88). No effective SNP interaction was concluded in the gene region according to the result of haplotype analysis.

In silico expression quantitative trait loci (eQTL) analysis was carried out in order to assess association between SNP B in 3'UTR of *ITPA* gene and the gene's expression. The eQTL on *ITPA* gene showed a significant association with SNP B ($p < 0.01$) within a 2 Mbp region around the SNP. In addition, a significant correlation between mRNA level of *ITPA* and the genotypes of SNP B was observed (permutated p-value=0.0045), displaying higher expression in accordance with the number of minor allele, which was observed to be protective.

The *ITPA* gene encodes the enzyme inosine triphosphate pyrophosphatase (ITPase), which functions to catalyze the hydrolysis of inosine triphosphate (ITP) to inosine monophosphate (IMP) and pyrophosphate. Although the role of ITPase in humans is not well-defined, it has been speculated that there might be a potential role for *ITPA* in immunity. Furthermore, a search in the UCSC genome browser revealed that expression of ITPA is high in immune cells. Considering these factors, it may be speculated that the minor allele of SNP B may serve to increase the stability of the mRNA transcript and increase expression of the gene, which in turn may lead to better functioning of the cells involved in the immune response against *M. tb*.

In conclusion, this study, to my knowledge, is the first attempt in using NGS to gain insight into host genetic factors associated with TB and being the first to report a significant association of the *ITPA* gene with young onset TB. The study also demonstrates the effectiveness of NGS in searching for susceptibility genes in common diseases. To confirm these findings, further genetic and functional studies are needed.