［課程―２］

審査の結果の要旨

氏名 ウオン ジン ハオ
Wong Jing Hao

TB is a serious global health issue and factors such as multidrug-resistant (MDR) and extensively drug-resistant (XDR) TB make treatment and control of the disease difficult. Both environmental and genetic factors contribute to the development of TB, however, numerous genetic studies conducted across a variety of populations have showed little consistency, and there may be as yet unidentified host genetic factors for TB susceptibility. Furthermore, there has not been a genome-wide study conducted regarding host genetic factors affecting susceptibility to drug-resistant TB.

The current study consists of two parts. The first part was a search for as yet unidentified host genetic factors affecting susceptibility to Tuberculosis (TB) in a Thai population, by conducting variant screening using next-generation sequencing (NGS) methods in an approximately 1Mb candidate region on chromosome (chr.) 5q31.1. This was followed by variant filtering and prioritization of candidate variants for testing by genotyping in a larger sample set and conducting case-control association analysis. The results of the first part are summarized as follows:

1. The candidate region on chr. 5q31.1 was defined based on previous family-based genome-wide linkage and SNP association studies conducted in Thai TB-affected families. The candidate region was captured and sequenced using NGS in 18 Thai TB-affected individuals. A total of 904 variants were detected from the sequencing reads and after filtering, 5 candidates were prioritized according to their predicted functional annotations. Among the candidates, one was a missense SNP (rs28552) located in the *CSF2* gene, two were 3'-Untranslated Region (UTR) SNPs (rs6873269 and rs12153431) in the *IRF1* gene, and another two were 3'-UTR SNPs (rs17510339 and rs17690923) in the *KIF3A* gene.

2. These candidate variants were confirmed by Sanger direct sequencing before being genotyped in a larger unrelated case and control sample set consisting of 663 Thai TB cases and 774 controls, using the Taqman genotyping assay. Case-control association analysis showed that the SNP rs17690923, located in the *KIF3A* gene, had moderate evidence of association with TB susceptibility in the genotypic model ($p = 0.012$) and the dominant model ($p = 0.029$, OR = 1.43, 95% CI = 1.04 – 1.96).

The second part of the study was a genome-wide association study (GWAS) conducted in 160 drug resistant TB cases and 192 drug-sensitive controls obtained from the Indonesian islands of Java and Madura. Genotyping was conducted using the Illumina Human OmniExpressExome-8 v1.2 BeadChip array.

Data quality control and case-control association analysis was first conducted, followed by regional SNP genotype imputation and *HLA* alleles imputation. A conditional logistic regression analysis was also conducted in the *HLA* region on chr. 6 using SNPs showing moderate evidence of association ($p = 5.0 \times 10^{-05}$) in the *BTNL2* gene as covariates. Gene-based association analysis was also conducted using the genotyping data and the results of the gene-based association analysis was used to conduct gene pathway, gene process networks, and gene ontology enrichment analysis. The results of the second part are summarized as follows:

1. The association analysis identified ten SNPs that showed suggestive evidence of association ($p \leq 1.0 \times 10^{-05}$) with drug-resistant TB. The most significant association was seen in the SNP

rs4859461 (p-value = 2.83 x 10^{-07}, OR = 2.33), located within the *SHROOM3* gene. A number of genes were also observed to contain a number of SNPs showing suggestive and moderate ($p \leq 1.0 \times 10^{-04}$) evidence of association, including the *LAMA2*, *BTNL2*, *GRM8*, *ADD2*, *ASIC2* and *NELL1* genes. The genes identified may have functions related to kidney function and subsequent drug clearance from blood plasma, granuloma formation or maintenance, as well as also potentially being involved in apoptosis and necrotization of the granuloma core, possibly leading to lowered drug penetration and exposure and increasing the risk and chance of developing drug-resistant TB.

2. Regional SNP genotype imputation identified an imputed SNP, rs111789237, located near the *HLA-B* gene was seen to have a lower p-value ($p_{imputation}$ =2.04 x 10^{-06}) compared to the most significant SNP seen in the GWAS for that region, rs9264916 (p = 1.93 x 10^{-05}). Validation of the SNPs showing suggestive associations as well as rs111789237 was conducted using the Taqman genotyping assay, with high concordance seen between the two platforms (~98%).

3. The results of *HLA* alleles imputation showed that a number of class I alleles in *HLA-B* and *HLA-C*, as well as class II alleles in *HLA-BRB1, HLA-DQA1,* and *HLA-DQB1* showed association with drug-resistant TB. After correcting for multiple testing, the *HLA-B*18:01* and *HLA-C*06:02* alleles remained significantly associated (p = 0.003, OR = 2.42 and 0.13, respectively).

4. Conditional logistic regression was conducted using 3 SNPs (rs10947262, rs10947261 and rs28362680) showing moderate evidence of association in the *BTNL2* gene. A SNP, rs3104404, located between the *HLA-DQA2* and *HLA-DQB1* genes, was seen to have a more significant p-value (p = 1.29 x 10^{-04}) compared to its original GWAS results (p = 1.53 x 10^{-01}). However, the result was not more significant than the ones seen in *BTNL2*, and the main association signals for the region was determined to be located in *BTNL2*.

5. Results of the gene-based association analysis showed that genes such as *SHROOM3*, *ADD2*, and *GRM8*, which were seen to contain multiple SNPs of suggestive and moderate association in the GWAS, were also identified to be significantly associated. Genes with $p \leq 0.01$ (n = 364) in the gene-based association analysis were used for the subsequent enrichment analysis by the Metacore software. Gene pathways and gene ontologies relating to inflammation, as well as immune response and regulation were seen to be significantly enriched in the dataset. Additionally, immune response and inflammation signalling gene process networks were seen to be enriched in the dataset as well.

This study found a new, as yet unidentified variant in the 3'-UTR of the *KIF3A* gene to show moderate association with TB in Thais. The present study was also the first report of a genome-wide study for host genetic susceptibility factors for drug-resistant TB and reported a number of genetic loci and genes with suggestive evidence of association. Additionally, new *HLA* alleles were also identified to be associated with susceptibility to drug-resistant TB in the current study. Furthermore, gene pathways, gene process networks and gene ontologies related to inflammation and immune response and signalling were found to be enriched for drug-resistant TB. The study contributes a better understanding of host genetic susceptibility factors for the development and pathogenesis of TB and drug-resistant TB. For these reasons, we consider the candidate worthy of a Ph.D degree.