Up-regulated long noncoding RNA AC007128.1 and its genetic polymorphisms associated with Tuberculosis susceptibility

Hong Yan¹,²#, Guoye Liu³#, Yuan Liang⁴,⁵#, Wei Wu⁵,⁶, Rui Xia¹, Lin Jiao⁷, Han Shen⁸, Zhijun Jia⁹, Qian Wang¹, Zhiqiang Wang³, Yi Kong¹, Binwu Ying⁷, Hualiang Wang¹⁰, Chengbin Wang¹

¹Department of Clinical Laboratory Medicine, the First Medical Center, Chinese PLA General Hospital, Beijing, China; ²Laboratory Medicine Center, the Second Affiliated Hospital, Nanjing Medical University, Nanjing, China; ³Department of Laboratory Medicine, Affiliated Brain Hospital of Nanjing Medical University (Chest Branch), Nanjing, China; ⁴The Affiliated Cancer Hospital & Hepatobiliary Center, First Affiliated Hospital, Nanjing Medical University, Nanjing, China; ⁵Department of Bioinformatics, Nanjing Medical University, Nanjing, China; ⁶Research Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, China; ⁷Department of Laboratory Medicine, West China Hospital, Sichuan University, Chengdu, China; ⁸Department of Clinical Laboratory, the Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, China; ⁹Department of Nuclear Medicine, the Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, China; ¹⁰Department of Molecular Biology, Shanghai Centre for Clinical Laboratory, Shanghai, China

Contributions: (I) Conception and design: C Wang, H Wang; (II) Administrative support: C Wang; (III) Provision of study materials or patients: H Yan, G Liu, R Xia, H Shen, Z Jia, Q Wang, Z Wang, Y Kong; (IV) Collection and assembly of data: Y Liang, W Wu, L Jiao, B Ying; (V) Data analysis and interpretation: H Yan, G Liu, Y Liang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

These authors contributed equally to this work.

Correspondence to: Chengbin Wang. Department of Clinical Laboratory Medicine, the First Medical Center, Chinese PLA General Hospital, Beijing 100853, China. Email: wangchbin301@163.com; Hualiang Wang. Department of Molecular Biology, Shanghai Centre for Clinical Laboratory, 528 Hongshan Road, Pudong, Shanghai 200126, China. Email: wanghualiang@sccl.org.cn; Binwu Ying. Department of Laboratory Medicine, West China Hospital, Sichuan University, Chengdu 610041, China. Email: docbwy@126.com.

Background: Tuberculosis (TB) remains a major public health problem. Long non-coding RNAs (lncRNAs) are important regulators of gene expression. In this study, we explored the association between the expression of lncRNA AC007128.1 and TB susceptibility.

Methods: Three single-nucleotide polymorphisms (SNPs) (rs12333784, rs6463794, and rs720964) of lncRNA AC007128.1 were selected using the 1000 Genomes Project database and offline software Haploview V4.2, and were genotyped by a customized 2x48-Plex SNPscan™ Kit.

Results: We identified two differentially expressed lncRNA including AC007128.1 and AP001065.3 in comparisons of expression profiles between ATB vs. LTBI, LTBI vs. HC, and AC007128.1 expression was specifically and significantly up-regulated in TB patients by verification of external data. Gene Ontology functional enrichment analysis and co-expression network showed up-regulated mRNA was mainly involved in negative regulation of the G protein-coupled receptor (GPCR) signaling pathway, and FPR1 and CYP27B1 were involved in the co-expression of AC007128.1. Using the 1000 Genomes Project, software Haploview V4.2, and SNP genotype, we screened out SNP rs12333784 which locus at 7p21.3 in AC007128.1 associated with TB susceptibility. The G carrier of rs12333784 was then finally verified to be significantly associated with pulmonary TB (PTB) and extrapulmonary tuberculosis (EPTB) susceptibility (pBonferroni =0.03878), and a similar but more significant effect was observed under the dominant model analysis (pBonferroni =0.013, OR =1.349, 95% CI, 1.065–1.709). In addition, the GG + GA genotype of SNP rs12333784 was significantly correlated with higher glucose (GLU) (P=0.03), higher gamma-glutamyl transferase (GGT) (P=0.05), and higher erythrocyte sedimentation rate (ESR) (P=0.05).

Conclusions: Our findings show lncRNA AC007128.1 can be regarded as biomarkers discriminating between ATB and LTBI and may also be a diagnostic biomarker for LBTI. These findings may aid clinical decision making in the management of TB.

Keywords: AC007128.1; tuberculosis infection; biomarker; genetic susceptibility; single nucleotide polymorphisms (SNPs)
Introduction

Tuberculosis (TB) remains a major challenge to global public health. According to the WHO Global Tuberculosis 2020 Report there were 10 million new TB cases worldwide in 2019, of which pulmonary TB (PTB) accounted for more than 80% (1). Approximately 1.4 million people died of TB in 2019, making it the major cause of mortality among chronic infectious diseases. In addition, new cases of TB in China in 2019 accounted for about 8.4% of those worldwide, ranking it third to only India and Indonesia in new cases, indicating that the need to explore potential biomarkers for the diagnosis of TB is urgent. The complications of tuberculosis include hemoptysis, pneumothorax, bronchiectasis, secondary lung infection, and cardiopulmonary failure. The diagnostic characteristics of hemoptysis are the presence of exudation and cavitiation or bronchial tuberculosis and local tuberculosis that cause bronchial deformation, distortion and dilation. The diagnostic characteristics of pneumothorax include the lesion or cavity breaking into the thoracic cavity, and the thoracic cavity exuding more fluid which can form liquid pneumothorax and empyema. The bronchiectasis is characterized by pulmonary tuberculosis lesions destroying the bronchial wall and surrounding tissues of the bronchus. Bronchial tuberculosis itself can also cause bronchial deformation and expansion. The secondary lung infection is associated with airway obstruction caused by bronchiectasis, atelectasis, and bronchial tuberculosis. The diagnostic characteristics of cardiopulmonary failure are that the treatment of pulmonary tuberculosis is ineffective, and chronic disease damages the lung tissue, causing acute respiratory failure and long-term hypoxia.

Deficiencies exist in current laboratory TB diagnosis technology. Although traditional identification based on Mtb culture is the current “gold standard” for clinical TB diagnosis, it is time-consuming and the delay in results may cause further spread and a delay in timely treatment (2). As a form of immunological testing, T-SPOT.TB is an enzyme-linked immunospot assay designed to use specific T cells present in peripheral blood mononuclear cells (PBMC) of TB patients to secrete γ-interferon after stimulation by an antigen. Although T-SPOT.TB greatly shortens the testing time, it is associated with a high-risk of false positives as it may detect infections and inflammatory reactions caused by other diseases and does not distinguish between latent TB infection (LTBI), active TB infection (ATB), and past TB infection. The WHO estimates that as many as one-third of the global population have LTBI with continuous bacterial survival and bacterial immune control, but no symptoms of ATB. However, there is no direct test to diagnose LTBI, resulting in a reliance on the pure protein derivatives (PPD) test and IGRA assay for auxiliary diagnosis (3), and as these tests have poor prognostic value, there is an urgent need to discover new biomarkers for the timely clinical diagnosis of TB.

The host immune response has become a significant topic for research into the diagnosis of TB. After TB infection, the host immune system initiates a series of responses and host-specific TB diagnostic markers may be found based on this. Long non-coding RNA (lncRNA) are defined as transcripts with lengths exceeding 200 nucleotides that are not translated into protein (4). An increasing number of studies have shown that lncRNAs play crucial roles in the development of many different types of tumors, including XIAP-AS1 in gastric cancer (5), HOXA-AS3 in human glioma (6), OIP5-AS1 in multiple myeloma (7), MIAT in breast cancer (8), and CRNDE in colon cancer (9). In addition, some evidence suggests lncRNAs participate in host inflammatory responses. Atianand et al. reported that LncRNA-EPS, as a repressor of inflammatory responses, is precisely regulated in macrophages to control the expression of immune response genes (IRGs) (10). Zhang et al. found that MARCKS or ROCKI, induced by multiple TLR stimuli and interacted with APEX1, acted as a master regulator of inflammatory responses (11). Genetic variants of lncRNAs were also associated with TB susceptibility and clinical phenotypes of active TB (12-14). Until now, the pathogenesis and roles of the host lncRNAs in TB were unclear, and further understanding of the pathogenesis of TB in humans will help find novel approaches to prevention and therapy.

AC007128.1 belongs to a lncRNA transcribed from chromosomal region 7p21.3. Studies have found that AC007128.1 promoted migration and invasion in esophageal cancer cell lines and was up-regulated in both...
esophageal cancer tissues and cells, resulting in a poor prognosis (15), and a linkage analysis study based on a full microsatellite genomes scan showed that chromosomal region 7p21-22 was significantly associated with susceptibility to active TB (P=0.0002) (16). Therefore, we speculated that significantly up-regulated AC007128.1 may correlate with TB development.

Genetic predisposition plays an important role in defense of TB (17), and the association between polymorphisms in host immune-related genes and TB susceptibility has been identified (18,19). However, the exact molecular regulatory mechanism of TB remains unclear. Genetic polymorphisms associated with PTB, including OXTR polymorphisms in the endocrine system (rs2254298, rs237911, and rs2228485) were found to be significantly associated with an increased risk of PTB (20) and maternal KCNN3 polymorphisms in vascular angiogenesis (rs1218585, rs1218584, and rs883319) increased the infection risk of PTB (21). In addition, a previous analysis of innate immunity and inflammation gene polymorphisms found that polymorphisms in TNFα promoter and fetal catechol-o-methyltransferase (COMT) genes (rs4818) each conferred an increased risk of PTB (22). However, there are no reports of research on correlation analysis between gene polymorphism and EPTB and PTB.

Our study revealed that AC007128.1 was a potential specific biomarker for TB and was significantly likely to participate in immune responses to TB through the G protein-coupled receptor (GPCR) signaling pathway. We aimed to investigate the association between three single nucleotide polymorphisms (SNPs) within AC007218.1 genes and susceptibility for TB in the western Chinese population to provide useful information for understanding the role of the AC007128.1 gene variant in the clinical diagnosis and prevention of PTB and EPTB. We present the following article in accordance with the MDAR reporting checklist (available at https://dx.doi.org/10.21037/atm-21-2724).

Methods

Study subjects

Our retrospective study recruited 900 TB patients and 1,534 healthy controls. All subjects were recruited from the West China Hospital of Sichuan University from Jan 2014 to Feb 2016 as described previously (23,24).

GEO dataset reanalysis

We used the AnnoProbe (V0.1.0) R package to reannotate microarray probes from the GSE98461 dataset (25). We then filtered lncRNAs for the analysis and performed the Wilcoxon rank sum test to determine significantly differently expressed lncRNAs with log2 (fold change) ≥1 and P value <0.05.

Function enrichment analysis

Through correlation analysis, mRNAs with significant positive correlation (cor ≥0.5 and P value <0.05) with AC007218.1 were screened in groups with a TB and healthy control group. This saw 760 mRNAs obtained in the TB group, 455 mRNAs in the healthy control group, and a total of 23 mRNAs obtained by the two groups. The top eight significant enrichment pathways (ranked by the P value) were screened through Gene Ontology functional enrichment analysis.

SNPs selection

The genetic variation information of lncRNA was obtained from VCF to PED Converter on Ensembl (http://asia.ensembl.org/Homo_sapiens/Tools/VcftoPed), then imported into Haploview software. The marker SNPs represented by the Han Chinese in Beijing (CHB) minor allele frequency (MAF) ≥0.2 were screened. Genotype distributions of SNPs within the lncRNA gene did not deviate from Hardy-Weinberg equilibrium (HWE) (P>0.05 for all loci).

Genotyping

We used QIAamp® DNA blood Mini Kit (Qiagen, Germany) to isolate genomic DNA from peripheral blood samples and used modified multiple linkage analysis (Genesky biotechnology Inc., Shanghai, China) to genotype the candidate SNPs. The primers for multiplex PCR amplification and probes for all three SNPs linkage reactions were then added. The 0.5 μL ligation product was fractionated and loaded into ABI 3730xl and the original data were analyzed by GeneMapper v4.1 software. Different fluorescent-labeled allele-specific oligonucleotide probes were used to distinguish the specific alleles of each
SNP and differentiate the extension length of different SNPs at the 3’ end. In addition, we introduced ddH$_2$O as a negative control reaction to monitor the quality of genotyping. Genotyping was blinded for all investigators, and we randomly selected 10% of genotypes with a 100% coincidence rate. The study conformed to the provisions of the Declaration of Helsinki (as revised in 2013). The study was approved by institutional ethics committee of Clinical Trial and Biomedical Ethics Committee of West China Hospital of Sichuan University [No. (198) trial 2014]. Individual consent for this retrospective analysis was waived.

**Statistical analysis**

Chi-square test and Mann-Whitney U test were used to analyze categorical and continuous variables, respectively, and HWE was used to evaluate all SNPs in the control group. The relationship between candidate SNPs and TB infection was determined by the distribution of alleles and genotype frequencies, and genetic models (both dominant and recessive). The Plink v1.07 command-line program was used for unconditional logistic regression analysis, odds ratio (ORs) and 95% confidence interval (CIs) were used for estimation, and age and gender were adjusted for correction. The Bonferroni method was used to further calibrate many tests by using the “--adjust” option in the Plink program. The statistical software was SPSS version 19.0 (IBM, Chicago, USA) and all statistical tests were bilateral (P<0.05).

**Results**

**Differential expression of lncRNAs in TB**

To discover the lncRNA that were differentially expressed in TB, differentially expressed genes (Kruskal-Wallis Test P<0.05) were re-detected in comparisons of ATB patients vs. LTBI patients vs. healthy controls from the GEO dataset GSE98461. Hierarchical cluster demonstrated the expression patterns of lncRNA (Figure 1A). In addition, AC007128.1 (Kruskal-Wallis; P=0.012) and AP001065.3 (Kruskal-Wallis; P=0.018) showed significant differences in comparisons of expression profiles between ATB vs. LTBI and LTBI vs. healthy controls (Kruskal-Wallis Test P<0.05, Fold Change TB/LTBI >1.2, and Fold Change LTBI/HC >1.2) (Figure 1B). This suggested AC007128. and AP001065.3 could be regarded as biomarkers which could discriminate between ATB and LTBI and may also be a diagnostic biomarker for LBTI.

**AC007128.1 specifically expressed in TB patients**

To verify whether the above two lncRNAs in the heat map were specifically expressed in patients with TB but not with other diseases, we analyzed GSE107231 with non-alcoholic fatty liver disease (NAFLD) datasets, GSE94519 with rheumatoid arthritis (RA) datasets, GSE163980 with type 2 diabetes (T2D) datasets, GSE145227 with pediatric sepsis datasets, GSE123932 with hepatitis B (HBV) datasets, GSE169256 with coronary heart disease (CHD), GSE102541 with acute cerebral infarction (ACI), and GSE165934 with asthma, all of which have AC007128.1. The results showed that while AC007128.1 expression was significantly up-regulated in TB compared to LTBI patients as well as healthy controls, the expression of AC007128.1 in patients with metabolic diseases (NAFLD, RA, and T2D), bacterial infectious (pediatric sepsis), and viral infectious disease (HBV) were down-regulated in comparison to healthy controls (Figure 2A,B,C). Compared with healthy controls, AC007128.1 was up-regulated in cardiovascular diseases (CHD and ACI) and asthma with no statistical significance (Figure 2D). Given the specific expression of AC007128.1 in TB patients, this suggested AC007128.1 plays an important role in the pathogenesis of TB and may be a potential TB biomarker. Since AP001065.3 cannot be annotated in other datasets, no further data analysis was performed.

**Biological function of AC007128.1 in TB patients**

We further explored whether AC007128.1 was involved in the regulation of TB by taking the mRNAs in each group and those present in both groups by correlation analysis (Figure 3A). Gene Ontology functional enrichment analysis using the 23 mRNAs that had both TB and healthy control expression showed that the upregulated mRNAs were mainly involved in negative regulation of the GPCR signaling pathway (Figure 3B). It also revealed that up-regulated mRNA may block the G protein-coupled receptor signaling pathway that elicits the host immune response and inflammation, subsequently suppressing the immune response of TB patients.

To understand the AC007128.1 expression profile and assess its relationships with protein-coding genes, a co-expression network was constructed and hub nodes were analyzed. Networks of AC007128.1 and 16 mRNA which
were strongly co-expressed in the network were assembled, and the topological characteristics of the co-expression network were analyzed with degree >80 and topological coefficient >0.04 (Figure 3C). As reported in previous research, among these mRNAs, FPR1 was detected to be important in neutrophil function-related pathways (26) and has been found to play a key role in human immune responses and progression to ATB (27), consistent with our data analysis. CYP27B1 genes encode for the 1α-hydroxylase enzyme that is involved in the activation of 25-hydroxyvitamin D3 (25-OH-D3) to 1,25-(OH)2D3, and genetic variants of these genes may alter Vitamin D levels and render TB susceptibility (28). This implies FPR1 and CYP27B1 also have crucial roles in TB through certain key genes.

Figure 1 Differential expression of lncRNAs by reannotating microarray probes from GSE98461 dataset. (A) Unsupervised cluster analysis of differentially expressed lncRNAs in pair-wise comparisons. (B) AC007128.1 and AP001065.3 showed significant differences in comparisons of expression profiles between TB vs. LTBI and LTBI vs. HCs. TB, tuberculosis; ATB, active tuberculosis; LTBI, latent tuberculosis infection; HCs, healthy controls.
LncRNA AC007128.1 polymorphism is associated with the susceptibility of PTB and EPTB

Both PTB and EPTB infection are difficult to diagnose. To explore whether tagSNPs of IncRNA AC007128.1 polymorphism were associated with TB risk, we selected tagSNPs using data from the 1000 Genomes Project database and offline software Haploview V4.2. The data were then genotyped by a customized 2×48-Plex SNPscan™ Kit among a total of 900 TB patients and 1,534 healthy individuals in a western Chinese Han population. Among the 900 TB patients, 657 (73%) were PTB, 93 (10.3%) were EPTB, and 150 (16.7%) were PTB and EPTB (Figure 4A,B). We found that rs12333784, rs6463794, and rs720964 (3 SNPs) were successfully genotyped in the case-control population which was consistent with the HWE test results (P>0.05) (Figure 4C). Statistical analysis showed that subjects carrying the rs12333784 (A > G) mutant G allele had an increased susceptibility to PTB and EPTB compared with A allele carriers (P=0.013, or 1.349, 95% CI, 1.065–1.709). Furthermore, this relationship remained significant via Bonferroni correction with P=0.039 (Table 1), and after adjusting for age and gender, the genotype distributions of rs12333784 were significantly associated with PTB and EPTB both before and after Bonferroni correction (Table 1). We also conducted a genetic model analysis where rs12333784 also showed a significantly increased association with PTB and EPTB susceptibility (OR =1.702, 95% CI, 1.196–2.424, P=0.01 after Bonferroni correction) in the recessive model (Table 2). Therefore, the data revealed that the rs12333784 of AC007128.1 was associated with the risk of PTB and EPTB.

To further explore whether rs12333784 was associated with the clinical features of PTB and EPTB, we analyzed data between rs12333784 genotypes and laboratory indices in the PTB and EPTB group. This revealed patients with the rs12333784 GG + GA genotype (in the dominant model) were significantly correlated with higher glucose
(GLU) (P=0.03), gamma-glutamyl transferase (GGT) (P=0.05), and erythrocyte sedimentation rate (ESR) (P=0.05) (Figure 4D,E,F, Table 3) and indicates the rs12333784 GG + GA genotype may be involved in the immune response of PTB and EPTB patients.

**Discussion**

TB remains one of the top 10 causes of morbidity and mortality worldwide (1). A major challenge in the management of TB is the existence of a large population infected with LTBI, and the lifetime risk of reactivation for those with documented LTBI is estimated to be 5–10%. The effective treatment of LTBI is essential to prevent progression to ATB and may reduce the risk of the latter by at least 60%. The Tuberculin Skin Test (TST) and Interferon-Gamma Release Assays (IGRAs) are currently the main tests used for the diagnosis of LTBI, and positive results in either may indicate an immune response to *Mtb*. However, these tests have limitations as they cannot distinguish between LTBI with viable microorganisms and ATB groups (1). Our study explored lncRNA as a biomarker to discriminate between ATB and LTBI groups and the application of screening methods for lncRNA specifically expressed on TB to resolve the difficulty of false positives in the clinical diagnostic process. In addition, we explored the characteristics of susceptible populations at the molecular level and used these to better understand TB in the Chinese Han population.

Previous research reported that lncRNAs are the largest group of non-coding RNAs produced from the genome and mediate their functions through interactions with proteins, RNA, DNA, or a combination of these. lncRNAs can clearly function through a variety of mechanisms to regulate gene expression both at the transcriptional and post-transcriptional levels (29). The bulk of lncRNA studies
to date have focused on cancer (30). Recent years, data on the association of polymorphism in lncRNAs with cancer susceptibility suggested that SNPs in lncRNAs hold great potential as prognostic biomarkers for cancer. It has been confirmed that the rs1041279 SNP in the promoter region of HULC could enhance hepatocellular carcinoma risk without significantly altering the expression level of the lncRNA (31). Data indicated that the rs3787016 SNP in POLR2E might increase the susceptibility to prostate cancer (32). In addition, three HOTAIR SNPs, rs12826786, rs874945 and rs1899663, were identified as being associated with increased neuroblastoma risk (33). The SNP rs10845671 on lncRNA RP11-392P7.6 is associated with an increased risk of colorectal cancer (34). Moreover, More and more tumors including lung cancer, breast cancer, pancreatic cancer, etc., have been found to be associated with lncRNA SNP. In-depth research on it may provide the possibility to find suitable markers for the diagnosis of malignant tumors.

While they are now branching out to include

Figure 4 Clinical characteristics and clinical indicators in relation to polymorphisms of AC007128.1 in PTB & EPTB patients vs. HCs. (A) The age of PTB & EPTB patients was significantly younger than HCs. (B) Distribution of sex between PTB & EPTB patients and HCs, showed more males than females in both groups. (C) Map showing the genomic context of 100 candidate SNPs in AC007128.1. (D) GLU in relation to the rs12333784 genotype; (E) GGT in relation to the rs12333784 genotype. (F) ESR in relation to the rs12333784 genotype. Genotypes were stratified based on the dominant model. PTB & EPTB, both pulmonary tuberculosis and extrapulmonary tuberculosis; HCs, healthy controls; SNP, single nucleotide polymorphism; GLU, glucose; GGT, gamma-glutamyl transferase; ESR, erythrocyte sedimentation rate. *, P<0.05; **, P<0.01.
Table 1 Alleles and genotypes of AC007128.1 in Chinese Han patients with PTB & EPTB

| SNPs   | Allele | Case, n (%) | Control, n (%) | OR (95% CI) | P     | P**     | Genotype | Case, n (%) | Control, n (%) | P*     | P**     |
|--------|--------|-------------|----------------|-------------|-------|---------|----------|-------------|----------------|-------|---------|
|        |        |             |                |             |       |         |          |             |                |       |         |
| rs12333784 | G     | 168 (56.76) | 1,495 (48.73)  | 1.349       | 0.01293 | 0.03878 | GG       | 52 (35.14)  | 364 (23.73)    | 0.01433 | 0.04298 |
| A > G  | A      | 128 (43.24) | 1,573 (51.27)  |             |       |         | AG       | 64 (43.24)  | 767 (50.00)    |       |         |
|        | A      | 128 (43.24) | 1,573 (51.27)  |             |       |         | AA       | 32 (21.62)  | 403 (26.27)    |       |         |
| rs6463794 | A    | 79 (26.69)  | 731 (23.80)    | 1.174       | 0.2382 | 0.5958  | AA       | 11 (7.43)   | 83 (5.40)      | 0.2315  | 0.5956  |
| C > A  | C      | 217 (73.31) | 2,341 (76.20)  |             |       |         | CA       | 57 (38.51)  | 565 (36.78)    |       |         |
|        | C      | 217 (73.31) | 2,341 (76.20)  |             |       |         | CC       | 80 (54.05)  | 888 (57.81)    |       |         |
| rs720964 | A    | 115 (38.85) | 1,112 (36.27)  | 1.171       | 0.1986 | 0.7147  | AA       | 24 (16.22)  | 191 (12.46)    | 0.1985  | 0.6946  |
| G > A  | G      | 181 (61.15) | 1,954 (63.73)  |             |       |         | GA       | 67 (45.27)  | 730 (47.62)    |       |         |
|        | G      | 181 (61.15) | 1,954 (63.73)  |             |       |         | GG       | 57 (38.51)  | 612 (39.92)    |       |         |

P*, P value after adjusting for gender and age; P**, P or P* value after Bonferroni correction. SNP, single nucleotide polymorphism; OR, odds ratio; PTB, pulmonary tuberculosis; EPTB, extrapulmonary tuberculosis.

Table 2 Analysis of the AC007128.1 genotype relevant to PTB & EPTB risk in Chinese Han population

| SNPs   | Dominant model | Recessive model |
|--------|----------------|-----------------|
|        | OR (95% CI)    | P*     | P**     | OR (95% CI)      | P*     | P**     |
| rs12333784 A > G | 1.244 (0.8359–1.852) | 0.2787 | 0.8361  | 1.702 (1.196–2.424) | 0.003442 | 0.01033 |
| rs6463794 C > A  | 1.189 (0.8521–1.66)  | 0.3053 | 0.916   | 1.351 (0.704–2.594) | 0.1139 | 0.3416  |
| rs720964 G > A   | 1.125 (0.7983–1.586) | 0.4939 | 1       | 1.45 (0.9262–2.271) | 0.356  | 1       |

P*, P value after adjusting for gender and age; P**, P or P* value after Bonferroni correction. SNP, single nucleotide polymorphism; OR, odds ratio; PTB, pulmonary tuberculosis; EPTB, extrapulmonary tuberculosis.

Table 3 Clinical indicators in relation to rs12333784 polymorphisms of AC007128.1 in PTB & EPTB patients

| Variables | Total (n=149) | rs12333784 GG + GA (n=116) | rs12333784 AA (n=33) | P value |
|-----------|--------------|------------------------------|---------------------|---------|
| Age       | 31.00 (22.00–47.00) | 30.50 (22.00–47.00) | 32.00 (24.00–50.00) | 0.71    |
| Sex       | 149 (100%)   | 116 (100%)                  | 33 (100%)           | 0.7     |
| Male      | 79 (53%)     | 60 (52%)                    | 19 (58%)            |         |
| Female    | 70 (47%)     | 56 (48%)                    | 14 (42%)            |         |
| RBC       | 4.26 (3.90–4.59) | 4.22 (3.86–4.61) | 4.32 (4.00–4.56) | 0.69    |
| WBC       | 6.44 (5.28–8.33) | 6.48 (5.062–8.335) | 6.41 (5.72–8.23) | 0.57    |
| N         | 5.07 (3.66–6.96) | 5.15 (3.53–6.76) | 4.73 (3.78–7.25) | 0.63    |
| L         | 1.05 (0.66–1.43) | 0.95 (0.63–1.35) | 1.23 (0.83–1.72) | 0.96    |
| M         | 0.49 (0.34–0.64) | 0.48 (0.34–0.64) | 0.55 (0.35–0.65) | 0.77    |
| PLT       | 259.00 (188.00–307.50) | 260.00 (187.20–309.00) | 258.00 (197.00–301.00) | 0.47    |
| HCT       | 0.37 (0.32–0.40) | 0.37 (0.32–0.40) | 0.36 (0.28–0.42) | 0.52    |

Table 3 (continued)
Table 3 (continued)

| Variables | Total (n=149) | rs12333784 GG + GA (n=116) | rs12333784 AA (n=33) | P value |
|-----------|---------------|----------------------------|----------------------|---------|
| Hb        | 119.00 (108.00–131.50) | 118.50 (108.00–131.00) | 124.00 (108.00–135.00) | 0.64    |
| TP        | 69.70 (63.45–74.80) | 69.60 (63.40–74.80) | 70.35 (65.10–75.40) | 0.58    |
| ALB       | 37.90 (33.80–42.10) | 37.55 (33.70–42.17) | 39.40 (35.20–42.05) | 0.76    |
| GLB       | 30.55 (27.10–35.02) | 30.45 (27.02–35.35) | 31.30 (28.25–34.75) | 0.60    |
| ALP       | 80.00 (64.00–104.00) | 81.00 (65.50–104.00) | 79.50 (63.25–101.25) | 0.38    |
| ALT       | 17.00 (11.00–27.00) | 17.00 (11.00–27.75) | 17.00 (12.00–22.00) | 0.27    |
| AST       | 22.00 (15.00–31.00) | 22.00 (15.00–31.75) | 25.00 (17.50–28.00) | 0.71    |
| GGT       | 35.00 (22.00–66.50) | 37.00 (25.50–69.50) | 29.50 (16.00–54.75) | 0.05    |
| TG        | 1.06 (0.77–1.52) | 1.07 (0.75–1.62) | 1.04 (0.89–1.30) | 0.43    |
| CHOL      | 3.95 (3.01–4.73) | 3.95 (3.03–4.72) | 3.95 (2.94–4.71) | 0.36    |
| HDL-C     | 1.07 (0.69–1.34) | 0.99 (0.74–1.36) | 1.09 (0.67–1.29) | 0.57    |
| LDL-C     | 2.25 (1.64–2.75) | 2.25 (1.68–2.74) | 2.35 (1.57–2.86) | 0.40    |
| TBIL      | 8.80 (6.30–13.70) | 9.25 (6.23–14.63) | 8.40 (6.50–10.70) | 0.20    |
| DBIL      | 3.60 (2.50–6.30) | 3.65 (2.43–6.68) | 3.60 (2.85–5.00) | 0.57    |
| IBIL      | 5.00 (3.40–6.80) | 5.00 (3.33–7.12) | 4.80 (3.55–5.90) | 0.35    |
| GLU       | 5.13 (4.73–5.65) | 5.22 (4.74–5.82) | 4.96 (4.73–5.24) | 0.03    |
| UREA      | 3.90 (2.88–5.00) | 3.74 (2.80–4.90) | 4.87 (3.25–5.21) | 0.93    |
| CREA      | 55.0 (46.00–69.00) | 53.50 (45.00–66.90) | 60.30 (50.00–70.00) | 0.88    |
| URIC      | 292.00 (197.00–446.00) | 286.50 (199.20–397.50) | 314.00 (189.00–481.50) | 0.74    |
| Cys-C     | 0.92 (0.79–1.06) | 0.90 (0.77–1.04) | 0.97 (0.85–1.09) | 0.88    |
| CRP       | 12.75 (2.78–31.83) | 13.40 (2.70–32.55) | 10.30 (3.00–31.10) | 0.35    |
| ESR       | 31.00 (18.00–56.00) | 34.00 (18.50–59.50) | 21.50 (14.00–50.50) | 0.05    |
| TBA       | 34.00 (35%) | 28.00 (35%) | 6.00 (35%) | 1.00    |
| Smear     | 20.00 (18%) | 17.00 (19%) | 3.00 (14%) | 0.76    |
| TB culture | 18.00 (26%) | 14.00 (26%) | 4.00 (27%) | 1.00    |
| Liver damage | 24.00 (35%) | 20.00 (37%) | 4.00 (27%) | 0.59    |
| Fever     | 70.00 (53%) | 53.00 (51%) | 17.00 (58%) | 0.53    |
| Loss      | 61.00 (47%) | 49.00 (48%) | 12.00 (41%) | 0.67    |
| night sweat | 40.00 (31%) | 32.00 (32%) | 8.00 (30%) | 1.00    |
| Poor appetite | 61.00 (47%) | 50.00 (49%) | 11.00 (41%) | 0.52    |
| Fatigue   | 34.00 (26%) | 30.00 (29%) | 4.00 (14%) | 0.15    |

PTB, pulmonary tuberculosis; EPTB, extrapulmonary tuberculosis; RBC, erythrocyte/red blood cell; WBC, white blood cell; N, neutrophil; L, lymphocyte; M, monocytes; PLT, platelets; HCT, red blood cell specific volume; HB, hemoglobin; TP, total protein; ALB, albumin; GLB, globulin; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, γ-glutamyl transpeptidase; TG, triglyceride; CHOL, cholesterol; HDL-C, high density liprotein cholesterol; LDL-C, low density lipoprotein cholesterol; TBIL, total bilirubin; DBIL, direct bilirubin; IBIL, indirect bilirubin; GLU, glucose; CREA, creatinine; URIC, uric acid; Cys-C, cystatin C; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; TBA, total bile acid.
questions about the regulatory impact of these pervasive long noncoding gene species. Therefore, we aimed at determining the disease association between IncRNA and TB. Pan et al. found that expression of the resistin (RETN) and kallikrein 1 (KLK1) genes showed the greatest difference between TB and LTBI or HC groups (25). We filtered IncRNAs for reannotating microarray probes data and surprisingly discovered that IncRNA AC007128.1 showed significant differences in comparisons of expression profiles between ATB vs. LTBI and LTBI vs. healthy controls. While AC007128.1 has been previously reported as correlating with the overall survival of patients with esophageal cancer (15), we were the first to show its expression was significantly different between ATB and LTBI, and that it is expected to become a key molecule in their differential diagnosis.

We then analyzed the interaction relationship of AC007128.1 and differentially expressed mRNAs using co-expression analysis, and found 16 mRNAs in the prospective cohort, four of which have been found to play key roles in human immune responses or progression to ATB. Human FPR1 is a high affinity receptor for N-formyl-methionyl peptides (fMLP), which are powerful neutrophil chemotactic factors, and binding of fMLP to the receptor stimulates intracellular calcium mobilization and superoxide anion release. This response is mediated via a G-protein that activates a phosphatidylinositol-calcium second messenger system (35-37). FPR1 could serve as a receptor for TAF4 and mediate its effects on chemotaxing macrophages, promoting phagocytosis and increasing ROS release (38). Both Ai and Chen et al. found FPR1 separated ATB from LTBI patients in a Chinese Han population (27,39). CYP27B1 encodes for the 1α-hydroxylase enzyme that is involved in the activation of 25-hydroxy Vitamin D3 (25-OH-D3) to 1,25-(OH)2D3, and genetic variants of these genes may alter Vitamin D levels and render TB susceptibility (28). Wang et al. found that CYP27B1 in the metabolic pathway of Vitamin D is related to the risk and prognosis of TB (40). In addition, SAMSN1, a negative regulator of B-cell activation, was involved in B cell related pathways and played a key role in the immune response (41). Further, CD276 may participate in the regulation of T-cell-mediated immune responses, and the isoform 2 of CD276 has been shown to enhance the induction of cytotoxic T-cells and selectively stimulate interferon gamma production in the presence of T-cell receptor signaling (42). Our analysis shows these co-expression mRNAs are proven to be critical elements in the development of TB. Using the Gene Ontology functional enrichment analysis of 23 mRNAs simultaneously expressed in TB and HC in dataset GSE98461, we successfully found upregulated mRNAs were mainly involved in negative regulation of the GPCR signaling pathway. A previous study suggested a variety of GPCRs were involved in immune and inflammatory responses, and the stimuli involved were usually activated by some cytokines and inflammatory mediators in the immune or inflammatory process. In turn, this activated downstream effector molecules and transcription factors and eventually affected gene expression and biological functions. Our data also revealed FPR1, which co-expressed with AC007128.1, was involved in neutrophils function-related pathways by mediating via a G-protein that activated a phosphatidylinositol-calcium second messenger system. This suggested that AC007128.1 may participate in the up-regulation of mRNA, which blocks the G protein coupled receptor signaling pathway that elicits the host immune response and inflammation, and subsequently suppresses the immune response of TB patients.

Genome-wide association studies (GWAS) have revolutionized the study of complex diseases by allowing the quantitative disease-association of thousands of genetic loci (43) studies include evaluations of single-nucleotide polymorphisms (SNPs) or deletions, and determination of their association with disease phenotypes. Previous studies have shown that 90% of disease-related SNPs were located in non-coding regions of the genome (44). Zhang et al. found that IncRNA LIN00305 had five SNPs that were associated with atherosclerosis and all were located within an intrinsic region (45) SNP rs917997 is associated with celiac disease and is located within IncRNA Lnc13 (46). This suggests that SNPs located in the IncRNA region are associated with inflammatory diseases. While further research is required to elucidate the relationship between SNPs in IncRNAs and TB susceptibility, we discovered that SNP rs12333784 in AC007128.1, which locates at 7p21.3, showed a significant difference in PTB and EPTB and healthy controls. One hypothesis to be proposed is that SNPs in the gene responsible for immune responses to Mtb infection might change its coded proteins’ molecular function and consequently influence this dissemination process (47) and another concerned the linkage imbalance with other undetermined susceptible loci in different populations (15).

We found that rs12333784 was related to higher GLU, GGT, and ESR concentrations. Patients carrying
rs12333784 GG + GA genotypes had a higher inflammatory level and host immunity to PTB and EPTB infection, causing an increase in the ESR of patients with PTB and EPTB. In addition, patients with rs12333784 GG + GA genotypes showed high levels of GGT, which may be because this genotype is likely to cause severe liver damage. At the same time, the GLU level also increased, which may be due to the high sugar content in lung tissue, which contributes to the growth of Mtb and potentially, diabetes. Previous studies have demonstrated that rs9297758 in CASC8 is significantly associated with neutrophil counts among TB patients (23). Moreover, another study showed that rs218916 and rs160441 within lncRNA RP11-37B2.1 were significantly related to thrombocytopenia in anti-TB treatment (13). Taken together, our results suggested that rs12333784 is important in multiple facets of PTB and EPTB susceptibility.

To our knowledge, this is the first study to analyze the role of AC007128.1 and its genetic polymorphisms in the gene susceptibility of latent and active TB infection status. Our results provide a new direction for further study on the mechanism of TB infection and provide the groundwork for future studies on potential biomarkers or the development of novel therapeutic targets for a variety of TB. However, further clinical research is required to verify this.

Conclusions

In this study, we found AC007128.1 specifically differentially expressed in comparisons of expression profiles between ATB vs. LTBI and LTBI vs. healthy controls. Up-regulated mRNAs were mainly involved in negative regulation of the GPCR signaling pathway and AC007128.1 co-expressed with FPR1 and CYP27B1, participating in TB development. We found 12333784 were associated with PTB and EPTB susceptibility and might affect human immune responses in PTB and EPTB. Our research may provide potential biomarkers for discriminating between ATB and LTBI groups, and might assist clinical research.

Acknowledgments

Funding: Sponsorship for this study was funded by the Medical Science and Technology Development Projects of Nanjing (YKK18167), Medical Key Science and Technology Development Projects of Nanjing (ZKKX18016).

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at https://dx.doi.org/10.21037/atm-21-2724

Data Sharing Statement: Available at https://dx.doi.org/10.21037/atm-21-2724

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/atm-21-2724). All authors report that sponsorship for this study was funded by the Medical Science and Technology Development Projects of Nanjing (YKK18167), Medical Key Science and Technology Development Projects of Nanjing (ZKX18016). The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study conformed to the provisions of the Declaration of Helsinki (as revised in 2013). The study was approved by institutional ethics committee of Clinical Trial and Biomedical Ethics Committee of West China Hospital of Sichuan University [No. (198) trial 2014]. Individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Chakaya J, Khan M, Ntoumi F, et al. Global Tuberculosis Report 2020 - Reflections on the Global TB burden, treatment and prevention efforts. Int J Infect Dis 2021. doi: 10.1016/j.ijid.2021.02.107.
2. Woods GL. Susceptibility testing for mycobacteria. Clin Infect Dis 2000;31:1209-15.
3. Trajman A, Steffen RE, Menzies D. Interferon-Gamma Release Assays versus Tuberculin Skin Testing for the Diagnosis of Latent Tuberculosis Infection: An Overview of the Evidence. Pulm Med 2013;2013:601737.

4. Perkel JM. Visiting “noncodarnia”. Biotechniques 2013;54:301, 303-4.

5. Cai J, Wang D, Bai ZG, et al. The long noncoding RNA XIAP-AS1 promotes XIAP transcription by XIAP-AS1 interacting with Sp1 in gastric cancer cells. PLoS One 2017;12:e0182433.

6. Wu F, Zhang C, Cai J, et al. Upregulation of long noncoding RNA HOXA-AS3 promotes tumor progression and predicts poor prognosis in glioma. Oncotarget 2017;8:53110-23.

7. Yang N, Chen J, Zhang H, et al. LncRNA OIP5-AS1 loss-induced microRNA-410 accumulation regulates cell proliferation and apoptosis by targeting KLF10 via activating PTEN/PI3K/AKT pathway in multiple myeloma. Cell Death Dis 2017;8:e2975.

8. Luan T, Zhang X, Wang S, et al. Long non-coding RNA MIAT promotes breast cancer progression and functions as ceRNA to regulate DUSP7 expression by sponging miR-155-5p. Oncotarget 2017;8:76153-64.

9. Han P, Li JW, Zhang BM, et al. The lncRNA CRNDE promotes colorectal cancer cell proliferation and chemoresistance via miR-181a-5p-mediated regulation of Wnt/β-catenin signaling. Mol Cancer 2017;16:9.

10. Atianand MK, Hu W, Satpathy AT, et al. A Long Noncoding RNA lincRNA-EPS Acts as a Transcriptional Brake to Restrain Inflammation. Cell 2016;165:1672-85.

11. Mann PC, Cooper ME, Ryckman KK, et al. Polymorphisms in the fetal progesterone receptor and a calcium-activated potassium channel isofrom are associated with preterm birth in an Argentinian population. J Perinatol 2013;33:51-6.

12. Schepetkin IA, Khlebnikov AI, Kirpotina LN, et al. Antagonism of human formyl peptide receptor 1 with natural compounds and their synthetic derivatives. Int Immunopharmacol 2016;37:43-58.

13. Song J, Liu T, Zhao Z, et al. Genetic polymorphisms of long noncoding RNA RP11-37B2.1 associate with susceptibility of tuberculosis and adverse events of antituberculosis drugs in west China. J Clin Lab Anal 2019;33:e22880.

14. Zhang X, Liang Z, Zhang Y, et al. Comprehensive analysis of long non-coding RNAs expression pattern in the pathogenesis of pulmonary tuberculosis. Genomics 2020;112:1970-7.

15. Liu H, Zhang Q, Lou Q, et al. Differential Analysis of IncRNA, miRNA and mRNA Expression Profiles and the Prognostic Value of IncRNA in Esophageal Cancer. Pathol Oncol Res 2020;26:1029-39.

16. Stein CM, Zalwango S, Malone LL, et al. Genome scan of M. tuberculosis infection and disease in Ugandans. PLoS One 2008;3:e4094.

17. Alcaïs A, Fieschi C, Abel L, et al. Tuberculosis in children and adults: two distinct genetic diseases. J Exp Med 2005;202:1617-21.

18. Eskandari-Nasab E, Moghadampour M, Tahmasebi A, et al. Interleukin-17 A and F gene polymorphisms affect the risk of tuberculosis: An updated meta-analysis. Indian J Tuberc 2018;65:200-7.

19. Salie M, Daya M, Lucas LA, et al. Association of toll-like receptors with susceptibility to tuberculosis suggests sex-specific effects of TLR8 polymorphisms. Infect Genet Evol 2015;34:221-9.

20. Mann PC, Cooper ME, Ryckman KK, et al. Clinical relevance of LINC00152 and its variants in western Chinese tuberculosis patients. Oncotarget 2017;8:115456-68.

21. Li J, Wu L, Guo W, et al. Clinical relevance of LINC00152 and its variants in western Chinese tuberculosis patients. Oncotarget 2017;8:115456-68.

22. Schepetkin IA, Khlebnikov AI, Kirpotina LN, et al. Antagonism of human formyl peptide receptor 1 with natural compounds and their synthetic derivatives. Int Immunopharmacol 2016;37:43-58.

23. Ai JW, Zhang H, Zhou Z, et al. Gene expression pattern analysis using dual-color RT-MLPA and integrative genome-wide association studies of eQTL for tuberculosis susceptibility. Respir Res 2021;22:23.

24. DeLuca HF. Evolution of our understanding of vitamin D. Nutr Rev 2008;66:S73-87.
29. Atianand MK, Fitzgerald KA. Long non-coding RNAs and control of gene expression in the immune system. Trends Mol Med 2014;20:623-31.
30. Chan JJ, Tay Y. Noncoding RNA:RNA Regulatory Networks in Cancer. Int J Mol Sci 2018;19:1310.
31. Panzitt K, Tschernatsch MM, Guelly C, et al. Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. Gastroenterology 2007;132:330-42.
32. Huang S, Cui H, Lou Z, et al. Association of rs3787016 in Long Non-coding RNAs POLR2E and rs2910164 in MiRNA-146a with Prostate Cancer: A Systematic Review and Meta-analysis. Iran J Public Health 2018;47:623-32.
33. Yang X, He J, Chang Y, et al. HOTAIR gene polymorphisms contribute to increased neuroblastoma susceptibility in Chinese children. Cancer 2018;124:2599-606.
34. Jin M, Gu S, Ye D, et al. Association between genetic variants in the promoter region of a novel antisense long noncoding RNA RP11-392P7.6 and colorectal cancer risk. Environ Mol Mutagen 2017;58:434-42.
35. Postna B, Poppelier MJ, van Galen JC, et al. Chemotaxis inhibitory protein of Staphylococcus aureus binds specifically to the C5a and formylated peptide receptor. J Immunol 2004;172:6994-7001.
36. Murphy PM, McDermott D. Functional expression of the human formyl peptide receptor in Xenopus oocytes requires a complementary human factor. J Biol Chem 1991;266:12560-7.
37. Maestes DC, Potter RM, Prossnitz ER. Differential phosphorylation paradigms dictate desensitization and internalization of the N-formyl peptide receptor. J Biol Chem 1999;274:29791-5.
38. Wang W, Li T, Wang X, et al. FAM19A4 is a novel cytokine ligand of formyl peptide receptor 1 (FPR1) and is able to promote the migration and phagocytosis of macrophages. Cell Mol Immunol 2015;12:615-24.
39. Chen YC, Chang YP, Hsiao CC, et al. Blood M2a monocyte polarization and increased formyl peptide receptor 1 expression are associated with progression from latent tuberculosis infection to active pulmonary tuberculosis disease. Int J Infect Dis 2020;101:210-9.
40. Wang M, Kong W, He B, et al. Vitamin D and the promoter methylation of its metabolic pathway genes in association with the risk and prognosis of tuberculosis. Clin Epigenetics 2018;10:118.
41. Zhu YX, Benn S, Li ZH, et al. The SH3-SAM adaptor HACS1 is up-regulated in B cell activation signaling cascades. J Exp Med 2004;200:737-47.
42. Chapoval AI, Ni J, Lau JS, et al. B7-H3: a costimulatory molecule for T cell activation and IFN-gamma production. Nat Immunol 2001;2:269-74.
43. Bush WS, Moore JH. Chapter 11: Genome-wide association studies. PLoS Comput Biol 2012;8:e1002822.
44. Kumar V, Westra HJ, Karjalainen J, et al. Human disease-associated genetic variation impacts large intergenic non-coding RNA expression. PLoS Genet 2013;9:e1003201.
45. Zhang DD, Wang WT, Xiong J, et al. Long noncoding RNA LINC00305 promotes inflammation by activating the AHRR-NF-κB pathway in human monocytes. Sci Rep 2017;7:46204.
46. Atianand MK, Caffrey DR, Fitzgerald KA. Immunobiology of Long Noncoding RNAs. Annu Rev Immunol 2017;35:177-98.
47. Caws M, Thwaites G, Dunstan S, et al. The influence of host and bacterial genotype on the development of disseminated disease with Mycobacterium tuberculosis. PLoS Pathog 2008;4:e1000034.

(English Language Editor B. Draper)