

Article

Comparative Genetic Association Analysis of Human Genetic Susceptibility to Pulmonary and Lymph Node Tuberculosis

Abhimanyu^{1,2*}, Mridula Bose^{1*}, Astha Giri¹ and Mandira Varma-Basil¹

¹ Department of Microbiology, Vallabhbhai Patel Chest Institute, University of Delhi, Delhi, India, 110007

² The Global tuberculosis program, Baylor College of Medicine, and William T Shearer Center for Immunobiology, Texas Children's Hospital, Houston Texas, USA, 77004

* Correspondence: Abhimanyu@bcm.edu; bosmridula@gmail.com

Abstract: Background: Tuberculosis (TB) manifests itself primarily in the lungs as pulmonary disease (PTB) and sometimes disseminates to other organs to cause extra-pulmonary TB, such as lymph node TB (LNTB). This study aimed to investigate the role of host genetic polymorphism in immunity related genes to find a genetic basis for such differences. Methods: Sixty-three, Single nucleotide polymorphisms (SNPs) in twenty-three, TB-immunity related genes including eleven innate immunity (*SLCA11*, *VDR*, *TLR2*, *TLR4*, *TLR8*, *IRGM*, *P2RX7*, *LTA4H*, *SP110*, *DCSIGN* and *NOS2A*) and twelve cytokine (*TNFA*, *IFNG*, *IL2*, *IL12*, *IL18*, *IL1B*, *IL10*, *IL6*, *IL4*, *IL1RA*, *IL8* and *TNFB*) genes were investigated to find genetic associations in both PTB and LNTB as compared to healthy community controls. The serum cytokine levels were correlated for association with the genotypes. Results: PTB and LNTB showed differential genetic associations. The genetic variants in the cytokine genes (*IFNG*, *IL12*, *IL4*, *TNFB* and *IL1RA* and *TLR2,4* associated with PTB susceptibility and cytokine levels but not LNTB ($p < 0.05$). Similarly, genetic variants in *LTA4H*, *P2RX7*, *DCSIGN* and *SP110* showed susceptibility to LNTB and not PTB. Pathway analysis showed abundance of cytokine related variants for PTB and apoptosis related variants for LNTB. Conclusions: PTB and LNTB outcomes of TB infection have a genetic component and should be considered for any future susceptibility and functional studies.

Keywords: pulmonary tuberculosis; lymph node tuberculosis; extra-pulmonary tuberculosis; single nucleotide polymorphisms; cytokine; innate immunity; genetic association; genotype; serum

1. Introduction

Tuberculosis (TB), a major health hazard worldwide, is characterized by different clinical manifestations including localized infection in the lungs or pulmonary TB (PTB) and various forms of extra-pulmonary (EPTB). PTB accounts for 80% of all forms of TB [1], while EPTB constitute about 15-20% of all immunocompetent TB cases and 50% in cases infected with Human Immunodeficiency Virus (HIV) [2]. The most common form of EPTB is tubercular lymphadenitis (LNTB) with 50% of the cases involving the peripheral lymph nodes [3]. The basis of the variability of disease manifestation by the same infectious organism is not fully understood. It is not well understood as to why some individuals have EPTB disease which can infect other sites such as lymph nodes, while most persons have localized infection in the lungs.

The propensity for such different manifestations can be attributed to environmental exposures, pathogen virulence traits and host genetics of immune response. It is not really understood which of the aforementioned factors is the most important. India being an endemic country for tuberculosis with highest number of incident TB cases in 2021 [1], the prominent role of environmental exposures would most likely not be a driving factor in this population. As for pathogen virulence traits, there is association between infectivity of *Mycobacterium tuberculosis* (*Mtb*) strain and extra-pulmonary infections [4]. Pathogen variance can differ in disease presentations and *Mtb* strains in EPTB show mutations as in the *pncA* gene [5], or insertions and deletions in the phospholipase-C gene D (*plcD*) gene

[6], which is not seen strains isolated from PTB. To, that angle, in north India the most commonly circulating strain is CAS- 1, Delhi and the Beijing strains [7], which have not been shown to be associated with either PTB or EPTB alone. Although *Mtb* strains show genetic diversity, in various manifestations of TB like PTB or meningeal TB, the association is not lineage-specific [8], pointing to the decisive nature of host-genetics in such a scenario.

Both host and pathogen genotypes can be important as *Mtb* lineage specific association with disseminated disease of tubercular meningitis was identified to be associated with the *TLR2* polymorphism. [9]. The question remains as to why individuals infected with genotypically same, or different circulating strains of *Mtb* show different immune response. As far as the host genetics is concerned, ethnic specificity of the immune response in TB has been demonstrated [10].

So, to address the question of existing variability in the important immune response genes and their effect on governing this differential manifestation of TB, we designed a comparative study between PTB and LNTB. Host genetic association studies on different manifestations of tuberculosis are few (11-21). Workers have either focused on tubercular meningitis (TBM) [11, 12], pleural TB ,mixed cases with any form of EPTB together [13]. Some of the common EPTB studied gene variants include *TLR2* [11], *IL10*, *IFNG*[13], *VDR* [14, 15], *P2RX7*[16, 17], *SP110* [18, 19], *IL1B*, *IL1RA*[15, 20] and a pilot scale genome-wide association study [21].

In this study, we have considered multiple genes in a candidate gene approach in twenty three genes which have been shown to be important in tuberculosis immunopathogenesis [22] and genotyped 63 human genetic variants encompassing both innate and adaptive branch of tuberculosis in order to investigate both the arms of immunity.

We found that polymorphisms in the cytokine genes were associated for susceptibility to PTB. These cytokine genetic variants showed no or weaker association with LNTB. The cytokine genetic variants also showed high degree of gene-gene interaction among them accentuating their importance in governing susceptibility to PTB in north Indians. Interestingly, the variants in *SP110* [23] and *P2RX7* [24], which control macrophage apoptosis during tuberculosis infection , were associated with the risk of developing LNTB.

Overall, the study described here demonstrates that differences in host genetics is associated with different manifestations of tuberculosis. The study contributes to the emerging knowledge of key players in host-pathogen interaction in tuberculosis. We show that host genotype is a key determinant of the outcome of host-pathogen interaction and thus manifestation of tuberculosis as pulmonary and LNTB form.

2. Materials and Methods

2.1. Study Population

Venous blood Samples were collected from TB hospitals in and around New Delhi between 2009-2011. TB cases were 15 years or older culture confirmed or clinically diagnosed PTB cases with sputum smear microscopy for acid-fast bacilli (AFB) and chest X – ray data. Individuals included were not on any anti-tubercular therapy. For LNTB, patients were carefully selected to only have peripheral lymph node tuberculosis. Fine needle aspiration cytology (FNAC) was used to make histological confirmation of granulomatous structure and the FNAC was stained for AFB. Either histologically confirmed or AFB positive patients were considered for the LNTB group of the study. Patients with mixed EPTB and PTB infection were excluded from the study. All the enrolled patients were HIV negative. HIV negative community controls were enrolled from in an around New Delhi. The controls were confirmed to have never been diagnosed with TB and had no family history of TB. Genotypes for additional, ethnicity matched controls (n =135), were added for the validation phase analysis and were obtained from the Indian genome variation consortium database [25].

All patients and volunteers were informed about the study and an informed written consent was obtained from all the study participants. This study was approved the Institutional ethics committee of Vallabhbhai Patel Chest Institute, University of Delhi.

2.2. SNP selection

For the discovery panel, thirty-nine SNPs from twelve cytokine genes were selected. The SNPs mostly were in the intronic, exonic and 3'UTR regions. To avoid selection of non-polymorphic loci, as there was no data available on the Indian population at the time of the study initiation; we relied on the HapMap database (www.hapmap.org). SNPs were selected based on the following criteria: 1. Reported frequency >10% in at least 3 world population in HapMap database; 2. Reported frequency >20% in at least 2 world populations in HapMap database; 3. Average heterozygosity, which is a measure of genetic diversity at population scale was considered from dbSNP (<http://www.ncbi.nlm.nih.gov/projects/SNP>). It indicates the average proportion of individuals which are heterozygous in dbSNP from all the SNP data submitted to it, and this reduces selection on non-polymorphic loci. We have successfully implemented this strategy previously identifying novel genetic associations for TB [26]. New associations identified from the discovery panel in the study, were intersected previously identified strong association with PTB [26] and used for the validation panel.

2.3. SNP Genotyping

Genomic DNA was extracted using QIAamp DNA kit (Qiagen, Hilden, Germany). The concentrations of the DNA samples were determined by Nanodrop using Nanoquant™ plate of Infinite® Pro 200 system (Tecan, Männedorf Switzerland), checked for purity on a 1% agarose gel and stored at -20°C until further analyses. All the cytokine SNPs were genotyped using the matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (Sequenom Inc., USA). Assays for all SNPs were designed using Spectro DESIGNER software (Sequenom Inc., USA) and genotyped using the iPLEX assays (www.sequenom.com/iplex) as described previously [26].

The innate panel SNPs were typed using tetra-primer Amplification refractory mutation system (ARMS) PCR following the method of Ye et al [27] with modifications optimized for our SNP panel.

Briefly, primers for each selected polymorphism were designed using primer design software available at http://cedar.genetics.soton.ac.uk/public_html/primer.html. The software has been optimized to include two deliberate mismatches in the inner primer sets at 3' termini and -2 bases from 3' terminal to aid in the allele specificity. Each PCR reaction was carried out in a total volume of 10 µl containing 30 ng of template DNA, 10 pmol of each outer and inner primers, 200 mM of dNTPs, appropriate concentration of MgCl₂ (2.5 -3.5 mM), 20 mM Tris-Cl pH 8.4, 50mM KCl, 0.05% (v/v) W1 and 0.5 Units of thermostable Taq polymerase. The PCR conditions included 2 min at 95°C, 1 min annealing (annealing temperatures different according to primers) and 1 min extension (72°C), and additional two minutes extension at 72°C at the end of 35 cycles. Representative images of all genotyped polymorphisms are available in Figure S1 A-O.

To confirm that ARMS PCR detection of genotypes matches up with Sequenom genotyping, we selected a SNP (rs3212220) from *IL12*. This SNP was which was been genotyped on the Sequenom platform as well as ARMS-PCR was carried out for all the individuals of the control group and full concordance was observed (Table S1). *LTA4H* gene SNP rs17525495 was typed using allelic discrimination assay (catalog number: 4351379, Applied Biosystems) using manufacturer's instructions.

2.4. ELISA for serum cytokine measurement

Serum collected from the abovementioned cohort was quantified for the level of circulating cytokines by using Enzyme-linked Immunosorbent assay (ELISA) using cytokine

kits following the manufacturer's instructions. The unknown values were extrapolated from a standard curve within the linear range.

2.4. Statistical analysis

Hardy-Weinberg equilibrium was assessed in cases and controls for all tested variants to ensure that the samples were within allelic population equilibrium by using Haploview v 4.2 (<http://www.broad.mit.edu/mpg/haploview/>). A stringent cut off offered by the Haploview v 4.2 was used to further analysis which was used as a filtering criterion which included the following parameters: Minimum genotype = 75% and minimum minor allele frequency 0.0010) and HWE controls $p > 0.05$. The samples and SNPs failing this test were not selected for further analysis. PLINK v 1.07 (<http://pngu.mgh.harvard.edu/purcell/plink/>) was used to correct for multiple comparison, using Bonferroni methods, P-value after correction considered significant in the validation panel. Haplotype block generation was performed using the algorithm by Gabriel et al, 2002 [28] implemented in the Haploview software (<http://www.broad.mit.edu/mpg/haploview/>) which was also used for initial association testing. Genetic association testing was done using a 2 X 2 contingency table. Odds ratio, two tailed p- value was calculated for alleles using GraphPad Prism (version 5.00 for Windows, Graph Pad Software, San Diego California USA, www.graphpad.com). A two-tailed p value < 0.05 was considered statistically significant. The Odds ratio was confirmed by PLINK v 1.07, using a general model with fisher's exact test options. Multidimensional scaling using pairwise identity-by-state distances which was inferred based on genotypes of the 34 SNPs of the cytokine was carried out in PLINK v 9.0. Raw Distances were plotted on a 3D plot in R using the 'rgl' package.

Gene-gene interaction analysis we applied semi-exhaustive testing for pairwise interaction using PLINK v 1.07. The --fast epistasis along with --case-only option was used for this purpose. This has been hailed as a powerful approach by some workers. It provides a logistic regression test for interaction. This analysis exploits the fact that under certain conditions an interaction term in logistic regression equation corresponds to dependency or correlation between relative predictor variables within the population of cases. It uses an allelic model for both main effects and interactions and genotypes are not correlated.

2.5. SNP targeted Pathway analysis

We performed a SNP targeted pathway analysis using the PANOGA protocol [29] as shown by Bakir-Gungor et al [30]. The PANOGA protocol uses the association information in terms of p-value and creates files that can be used as input files in Cytoscape [31] application of the JActiveModules [32], which takes the genes containing the SNPs information and extrapolates it to the human whole human protein-protein interaction network and derives network and sub-network based on the input genes in the query. The JActiveModules output consisting of networks is then used as an input in ClueGo app [33] in which one can look for gene annotations from various sources including the KEGG, WikiPathways [34], GO database for immunological, biological, molecular and other networks that has been used to visualize the pathways to the gene interaction scale. We chose the WikiPathways to visualize the genes in the results.

3. Results

3.1. The study population was devoid of population stratification

False-positives associations can arise as a result of population stratification [35]. To investigate any hint of population-substructure, the self-reported ethnicity of each subject and his/her parents was carefully considered. To rule out population stratification a Multidimensional scaling (MDS) analysis was carried out on the genotyping data from the groups, which generated a compact cluster, without separating, indicating that population of patients (PTB and LNTB) and control subjects (HC) were homogenous with no substructures (Figure 1).

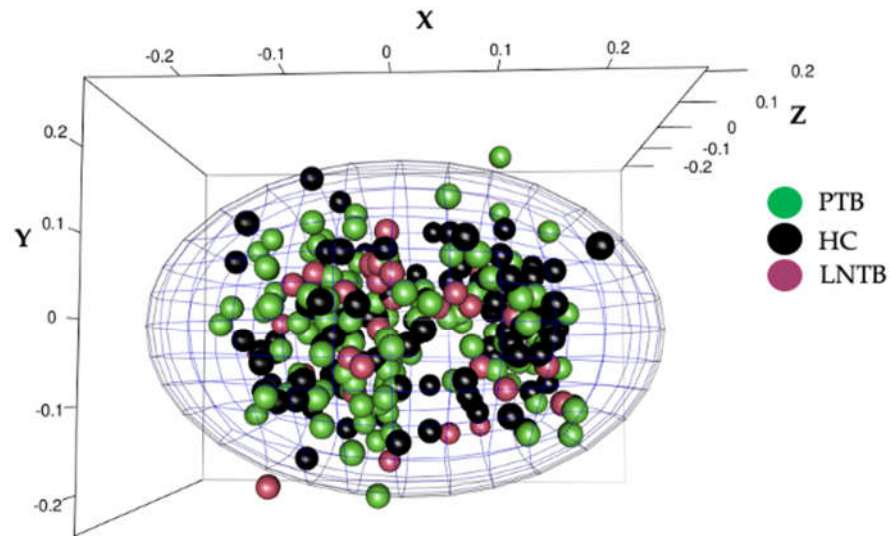


Figure 1. The study groups show homogeneity forming a tight cluster showing no stratification in the samples. Figure depicts a three-dimensional plot for checking population stratification among the study groups i.e., PTB (green spheres), LNTB (Purple Spheres) and HC (green Spheres). Raw Hamming distances as multidimensional scaling (MDS) co-ordinates are plotted on X, Y and Z-axes, to visualize genetic distance between the study groups.

3.2. Cytokine genetic variants show significant allelic and haplotypic association in PTB and not in LNTB

For analysis of the genetic association of the cytokine variants, we initially constituted a discovery panel of 39 SNPs listed in Table S2. We genotyped these cytokine polymorphisms in 110 PTB cases, 78 HC and 35 LNTB cases. A Combined (PTB+ LNTB) comparison (Table 1) was followed by a separate PTB and LNTB comparison for association (Table 1), with the aim to identify the SNPs- linked to differential susceptibility to PTB and LNTB. 26 SNPs for PTB, 23 SNPs for LNTB and 25 SNPs combined, passed the filtering criteria, and were analyzed for allelic association are listed in Table S3. The significant and borderline significant associations are enlisted in Table 1.

Table 1. Comparative allelic association statistics of cytokine variants in the discovery group between PTB and LNTB.

Gene	SNP (rsID)	Risk allele#	Case, Control (n = 78) allele frequencies			P-Value [@]				
			PTB	(n=110)	LNTB	(n=35)	Combined (n =145)	PTB	LNTB	Combined
<i>TNFB</i>	rs746868	C/C/C	0.650, 0.543		0.589, 0.543		0.637, 0.543	0.0442^a	0.5545	0.0647
<i>IL10</i>	rs1878672	C/G/C	0.937, 0.812		0.250, 0.188		0.894, 0.812	0.0007^b	0.3487	0.0271^f
<i>IL1B</i>	rs1143643	G/A/G	0.935, 0.817		0.286, 0.183		0.887, 0.817	0.001^c	0.1174	0.0635
<i>IL1RA</i>	rs419598	C/T/C	0.418, 0.264		0.796, 0.736		0.369, 0.264	0.0079^d	0.401	0.054
<i>IL6</i>	rs1548216	G/C/C	0.985, 0.978		0.089, 0.022		0.031, 0.022	0.6447	0.032^e	0.5846

#: The minor allele; @: The association P-Values, significant P-Values and borderline P-Values are shown in bold.

Odds Ratio (95% CI): ^a1.6 (1.0-2.4); ^b 3.4 (1.6-7.2); ^c 3.2 (1.6-6.6); ^d2 (1.2-3.4); ^e 4.4 (1.0-19); ^f 2.0 (1.1-3.6)

When considering PTB and LNTB cases together for analysis (Combined) we could identify only one variant from *IL10* gene rs1878672 of significance while certain others showed trend for association such as rs746868 of *TNFB*, rs1143643 *IL1B* and rs3882891 of *IL1RA*. From the PTB only analysis we could identify only one variant from *IL10* gene rs1878672 of significance with C allele showing 3-fold risk of developing PTB. While certain others which showed trend for association in all TB group such as rs746868 of *TNFB* (1.6-fold risk), rs1143643 of *IL1B* (3-fold risk) showed association with PTB group, indicating that disease type has a bearing on the susceptibility to TB. From the LNTB only analysis we could identify only variant from *IL6* gene, rs1548216 of significance with C allele showing a 4.4-fold risk of developing LNTB.

Table 2. Comparative Haplotypic association statistics of cytokine variants in the discovery group between PTB and LNTB.

Gene	Block	Haplotype			Case, Control Frequencies			P-Value		
		PTB	LNTB	Combined	PTB	LNTB	Combined	PTB	LNTB	Combined
<i>IL18</i>	Block 1 ^a									
	TG	0.907	0.884	0.905	0.926, 0.880	0.893, 0.880	0.919, 0.880	0.153	0.8036	0.211
	CC	0.093	0.116	0.095	0.074, 0.120	0.107, 0.120	0.081, 0.120	0.153	0.8036	0.211
<i>TNFB</i>	Block 2 ^b									
	GC/GCC	0.609	0.281	0.601	0.650, 0.549	0.313, 0.268	0.629, 0.548	0.0568	0.5305	0.1146
	TG/GTC	0.383	0.269	0.387	0.343, 0.441	0.241, 0.280	0.358, 0.442	0.0637	0.5736	0.098
	TTC		0.011			0.036, 0.001			0.0303^c	
	TTG		0.433			0.411, 0.442			0.6911	

Multiallelic haplotype blocks ^aFor *IL18* Block 1 consisted of rs5744256, rs1834481; ^b For *TNFB* Block 2 consisted of rs2239704, rs746868 for PTB and Combined and rs2239704, rs909253 and rs746868 for LNTB. The rsIDs are indicated were in the same order in the haplotype blocks. Biallelic combination of haplotype showed two combinations for PTB and Combined and four combinations for LNTB.

^c Odds Ratio (95% CI) = 4.4 (1.0-19).

Analysis of the gene structure in the combined analysis revealed two haplotype blocks formed by SNPs in *TNFB* and *IL18* (Table 2). The haplotypic frequency among case and controls did not differ significantly for both PTB and combined groups (Table 2). While the combined and PTB analysis showed two combinations each for *TNFB* and *IL8*, LNTB showed four combinations for *TNFB*, the haplotypic frequency of one of which, the TTC showing 5-fold risk of developing LNTB (Table 2). No multiple corrections were carried out at this stage in the analysis. The aim was not to prematurely discard SNPs and select them for further validation in a larger sample size.

After, discovering that cytokine gene polymorphism associations with PTB and to a lesser extent in LNTB, we wanted to independently validate, these findings before making a conclusion. We selected, 15 SNPs for the validation panel from seven cytokine genes (Table S2). Seven of these genes and SNPs were selected association in the discovery panel

(Table 1, Table S3) and the eight SNPs found from our previous study on 25 SNPs (Table S2) from six cytokine genes [26], which achieved a replication sample size of 160 PTB cases, 50 LNTB cases and 265 controls, giving the validation panel a 91% power of study to detect an odds ratio of 1.8 and above for these five cytokine SNPs [36].

Upon analysis of 15 cytokine SNPs after applying Bonferroni's correction for multiple testing, (Table 3), for PTB we found association for *IFNG* at rs1861493, *IL4* at rs2853694 and *IL12* at rs3212220. This replicated our previous findings about cytokine gene polymorphisms increasing the risk for PTB [26]. In contrast, for LNTB two variants rs2070874 of *IL4* and rs2853694 of *IL12*, showed significance but these associations were lost after correction for multiple testing. (Table 3). Interestingly, out of the 7 SNPs from the discovery panel, which was significantly associated in (Table 1 and 2), only rs3024498 of *IL10* gene achieved a borderline significance ($p = 0.07$) (Table S4a) and the rest was not replicated in the validation cohort. This could be related to a lower replication sample size of the validation panel. Therefore, these SNPs need larger sample size for validation as with good numbers we could validate our previous cytokine gene association findings [26]. This analysis overall, showed that having polymorphic cytokine genes increases the risk of PTB but not LNTB, indicating that genetic polymorphisms play a role in TB manifestation as pulmonary or extra-pulmonary form.

Table 3. Comparative allelic association of the cytokine gene variants in the validation panel between PTB and LNTB.

Gene	SNP (rsID) ^a	Risk allele	Case, Control Frequencies			Odds Ratio (CI) ^b			P-Value ^c			P-Bonferroni ^d		
			PTB	LNTB	Combined	PTB	LNTB	Combined	PTB	LNTB	Combined	PTB	LNTB	Combined
<i>IFNG</i>	rs1861493	A	0.935, 0.858	0.148, 0.142	0.912, 0.858	2.4 (1.3 - 4.2)	1.1 (0.62 - 1.8)	1.7 (1.1 - 2.7)	0.002	0.853	0.0197	0.013	1	0.138
<i>IL1RA</i>	rs4252019	C	0.974, 0.934	0.090, 0.066	0.962, 0.934	2.6 (1.1-6.0)	1.4 (0.72 - 2.7)	1.8 (0.93 - 3.4)	0.018	0.314	0.078	0.159	1	0.71
<i>IL4</i>	rs2070874	T	0.306, 0.239	0.847, 0.761	0.280, 0.239	1.4 (1.0 - 2.0)	1.7 (1.1 - 2.9)	1.2 (0.90 - 1.7)	0.048	0.027	0.194	0.303	0.2436	1
	rs2853694	A	0.624, 0.508	0.620, 0.508	0.608, 0.508	1.6 (1.2 - 2.2)	1.6 (1.1 - 2.3)	1.5 (1.1 - 2.0)	0.002	0.018	0.004	0.017	0.1624	0.032
<i>IL12</i>	rs3212220	T	0.409, 0.297	0.355, 0.297	0.394, 0.297	1.6 (1.2 - 2.2)	1.3 (0.8 - 1.9)	1.5 (1.2 - 2.1)	0.002	0.186	0.003	0.017	1	0.031

SNP, single nucleotide polymorphism; ^adb SNP, the SNP database (<http://www.ncbi.nlm.nih.gov/projects/SNP>); Two tailed $P < 0.05$ was considered significant. ^b Confidence Interval; ^c P-value from a chi-square test; ^d P-value after applying Bonferroni correction for multiple testing.

3.3. Gene-gene epistatic interaction analysis reveals a higher risk for cytokine genes majorly in PTB and not LNTB

After determining that cytokine gene polymorphisms contributed to increased risk for PTB susceptibility, we applied semi-exhaustive epistatic testing for pairwise interaction among the significantly associated SNPs from a previous panel [26] and the current cytokine gene validation panel to understand their genetic interaction. Thirteen significant interactions were identified and are enlisted in Table 4. Interestingly, the *IL4* locus showed interaction in LNTB as well, highlighting a critical role for this SNP in the north Indian population (Table 4). This approach also identified some SNPs which were not associated in the single locus analysis. *IL1RA* emerged as the gene having a significant interaction with *IL12*, *IL4* and *TNFB* genetic variants. Most of the *IL1RA* interactions were protective with odds ratio < 1 . Only one interaction between its own SNP was showing an eight-fold risk ($p = 3.066E-05$). This interaction could be important in defining the genetic susceptibility to TB. The other important player was *IL4* which showed interaction with variants of *IL12*, *IL1RA* and *TNFB*. Interestingly, all the interactions of *IL4*, an anti-inflammatory cytokine with other proinflammatory cytokines such as *IL12* and *TNFB* were

showing a very high risk (18 –fold risk) with very highly significant p-values. These genetic interactions enabled us to test the hypothesis that the disease outcome in tuberculosis can be due to interaction of the cytokine gene polymorphism. Also, many of the loci identified here were not significant in single variant association analysis. This analysis confirmed that cytokine gene polymorphisms affect the outcome of PTB more than LNTB, adding evidence to support the role of genetic polymorphisms in differential disease manifestation in TB.

Table 4. Comparison of Epistatic interactions in cytokine genes between PTB and LNTB.

CHR1	Gene	SNP1	CHR2	Gene	SNP2	OR_INT	STAT	P-Value	TB form
2	<i>IL1RA</i>	rs1794068	2	<i>IL1RA</i>	rs3213448	0.07834	21.47	3.60E-06	PTB
2	<i>IL1RA</i>	rs1794068	2	<i>IL1RA</i>	rs3181052	0.08759	22.68	1.92E-06	PTB
2	<i>IL1RA</i>	rs1794068	5	<i>IL12</i>	rs3213119	0.04328	24.7	6.69E-07	PTB
2	<i>IL1RA</i>	rs1794068	5	<i>IL12</i>	rs3213096	0.06435	25.68	4.03E-07	PTB
2	<i>IL1RA</i>	rs1794068	6	<i>TNFB</i>	rs3093542	0.05699	26.41	2.76E-07	PTB
2	<i>IL1RA</i>	rs315951	2	<i>IL1RA</i>	rs9005	8.265	17.38	3.07E-05	PTB
2	<i>IL1RA</i>	rs315951	5	<i>IL4</i>	rs2243266	0.1028	17.66	2.64E-05	PTB
2	<i>IL1RA</i>	rs315951	5	<i>IL12</i>	rs3213119	0.008753	31.27	2.25E-08	PTB
2	<i>IL1RA</i>	rs315951	5	<i>IL12</i>	rs3213096	0.007691	32.86	9.92E-09	PTB
2	<i>IL1RA</i>	rs315951	6	<i>TNFB</i>	rs3093542	0.008263	32.32	1.31E-08	PTB
5	<i>IL4</i>	rs2070874	5	<i>IL12</i>	rs3213119	18.18	29.85	4.66E-08	PTB
5	<i>IL4</i>	rs2070874	5	<i>IL12</i>	rs3213096	18.33	30.69	3.03E-08	PTB
5	<i>IL4</i>	rs2070874	6	<i>TNFB</i>	rs3093542	18.86	31.75	1.76E-08	PTB
5	<i>IL4*</i>	rs2070874	5	<i>IL12</i>	rs3213096	16.71	37.64	8.52E-10	PTB
5	<i>IL4*</i>	rs2070874	5	<i>IL12</i>	rs730690	5.056	16.02	6.26E-05	PTB
5	<i>IL4*</i>	rs2070874	5	<i>IL4</i>	rs2243266	5.007	23.61	1.18E-06	LNTB

Abbreviations: CHR1: Chromosome of 1st SNP, SNP1: Identifier of 1st SNP; OR_INT: Odds ratio for interaction; STAT: Chi-square,1-df (degree of freedom); P-Value: Asymptomatic p-value, *the interaction for the validation panel.

3.4. Lack of major association of cytokine levels with genotypes in LNTB

We have previously shown that cytokine levels affected by their genotypes, and individuals with a certain genotype secrete more or less of cytokines in their serum in people with PTB [37]. Since, we observed such stark differences in the association of cytokine gene polymorphism in PTB and LNTB, we carried out a similar analysis for the LNTB samples in this study. Overall, LNTB showed higher levels of the cytokine as compared to the healthy controls (Figure 4A). Out of 34 SNPs tested, none of the cytokine genotypes except for *IL8* at rs3882891 showed any significant difference in cytokine levels as governed by their genotype (Figure 4B), lending credibility to a major role of cytokine gene polymorphism in PTB but not LNTB.

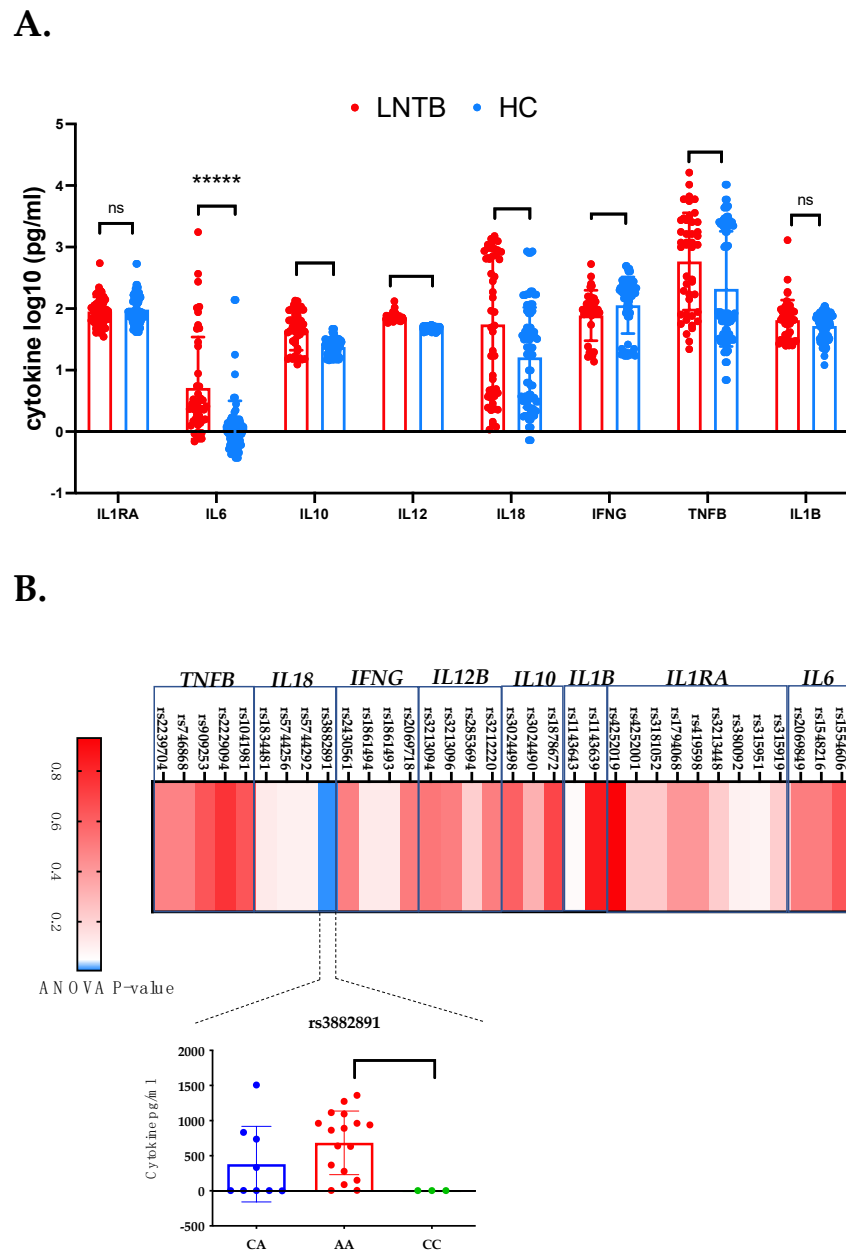


Figure 2. Cytokine levels have limited correlation with genotypes in LNTB. A. Increased cytokine levels in LNTB (red) vs HC (Blue). The groups were compared using a multiple Mann-Whitney test. **** $P < 0.00001$, *** $P < 0.0001$, ** $P < 0.001$, * $P < 0.05$. B. Heatmap of overall ANOVA P-Values of 34 SNPs in the cytokine genes with the corresponding cytokine levels. Shades of Red depict non-significant P-values and shades of blue significant p-value. Only significant variant, *IL18* at rs3882891 is shown with the genotypes with their respective levels, where AA genotype individuals are highest *IL18* producers and compared to CA and CC ($P < 0.05$) genotypes.

3.5. Innate immunity related genes are majorly associated with LNTB and not PTB

Innate immunity forms the first line of defense and multiple of innate immune genes have been implicated in susceptibility to PTB in various populations of the world where TB is endemic [22]. Since we observed such stark differences in cytokine gene polymorphisms, we hypothesized that a number of these gene polymorphisms would be PTB or LNTB specific. Widely studied polymorphisms were selected for the study, as it would offer us a great comparative insight with other world populations for TB susceptibility. The allelic association of the innate genes is listed in Table 5.

P2RX7 gene showed a 7-fold risk for: -762 T/C (rs2393799) C allele for the development of LNTB this association was marginally associated with risk of developing PTB. For rs37511431 we didn't detect any association. Out of three studied variants of the *VDR* gene variants i.e., *FokI* (rs2228570), *TaqI* (rs731236), *BsmI* (rs1544410), rs1544410 was found to be not polymorphic (presence of only one allele detected), rest of the two polymorphisms were not found to be associated with TB (both PTB and LNTB) risk or protection in this population. We did not find any association between either PTB or LNTB and *IRGM* genetic variant rs9637876 (Table 5). The results indicated that *NRAMP1 / SLC11A1* gene polymorphic variants may not be associated with the susceptibility to TB in north Indians. In fact, we could detect the presence of only a single genotype in all cases and controls; a CC genotype for rs3731865 and a heterozygous AG genotype for rs17235409. No haplotypes were observed. G allele of the *TLR2* genetic variant rs6265786 (Arg677Trp) of showed a high risk for PTB but not for LNTB. Similar results obtained for rs4986790 of *TLR4* gene where the G allele of shows a 2-fold risk of development of PTB. For *DCSIGN (CD209)* in PTB cases, A allele of rs4804803 was overrepresented in cases as compared to healthy controls showed a very significant association posing a 4-fold risk for developing PTB in north Indians and a 1.9-fold risk in LNTB cases (Table 5). None of the two tested *NOS2A* variants i.e., rs2274894 and rs7215373 showed any association either in PTB or LNTB, although the variant rs7215373 showed marginal association for both PTB and LNTB. Of the three studied 3 polymorphisms in the *LTA4H* gene rs1978331, rs2660898, and rs17525495, none of the variant showed association in either PTB or LNTB. *LTA4H* gene polymorphisms have been shown to provide heterozygous protection, meaning having a heterozygous genotype is protective from TB [12]. When we compared the heterozygous genotypes vs the homozygotes as proposed by Tobin et al, we observed, that out of 3 typed variants, rs1978331 have a protective association (odds ratio < 1) in combined and LNTB but not in PTB (Table 6). Similar, odds were observed between the haplotype of rs1978331-rs2660898, where when both the SNPs are heterozygous, they are borderline protective for LNTB and not PTB (Table S4). We have previously shown that, *SP110* gene polymorphisms were associated with risk of LNTB and not PTB in this population [19]. To continue exploring this gene in an independent cohort, we genotyped, *SP110* variants rs6436915, rs1346311, rs7580900. As shown previously and none of these showed any allelic associations in PTB (Table 5). Due to limited independent samples for LNTB, so these were not genotyped.

Table 5. Comparative Allelic association between various innate gene polymorphisms between PTB and LNTB.

Gene	db SNP ^a rsID/ Another name	Gene Location	Risk allele	Case, Control Frequencies		Chi Square		P-Value ^b		Odds Ratio (CI) ^c	
				PTB	LNTB	PTB	LNTB	PTB	LNTB	PTB	LNTB
P2RX7	rs2393799/ -762T/C	Exon	C	0.707, 0.611	0.917, 0.611	3.892	19.23	0.048	1.15E-05	1.5 (1- 2.4)	7 (2.6- 18.4)
	rs3751143/ 1513 A/C	Exon	C	0.574, 0.525	0.574, 0.525	0.816	0.079	0.3664	0.7791	1.2 (0.79-1.7)	1.1 (0.5-2.2)
VDR	rs2228570/ <i>FokI</i>	Exon	G	0.505, 0.472	0.552, 0.472	0.436	1.151	0.5	0.288	ND	ND
	rs731236/ <i>TaqI</i>	Exon	C	0.893, 0.861	0.895, 0.525	0.849	0.298	0.357	0.5849	ND	ND
	rs1544410/ <i>BsmI</i>	Exon		Only GG Genotype observed							
IRGM	rs9637876/ -261T/C	Alu se- quence	C	0.306, 0.255	0.272, 0.255	1.53	1.2	0.2161	0.515	ND	ND
TLR2	rs6265786/Arg 299 Trp	intron	G	1.000, 0.958	0.979, 0.958	9.53	0.471	0.002	0.4925	20.92 (1.2-369.3)	2.1 (0.25 - 17.25)
TLR4	rs498670/ Asp 299 Gly	Intron	G	0.223, 0.125	0.05, 0.125	6.09	2.622	0.0135	0.1054	2.1 (1.2 - 2.5)	0.11 (0.10 - 1.3)
NOS2A	rs7215373	Intron	T	0.481, 0.405	0.514, 0.400	2.732	2.914	0.0983	0.0878	1.3 (0.94-1.95)	1.6 (0.94-2.65)
	rs2274894	3'UTR	T/G	0.213, 0.171	0.05, 0.125	1.348	0.45	0.2456	0.5022	1.3 (0.83-2.06)	1.3 (0.63-2.6)
SLC11A1/ NRAMP1	rs3731865 17235409	Exon Intron		Only CC genotype observed Only AG genotype observed							
SP110	rs6436915	Intron	T	0.421, 0.347	0.417, 0.350	2.766	1.038	0.0963	0.3082	ND	ND
	rs1346311	Intron	T	0.105, 0.104	0.959, 0.900	0.001	2.516	0.9701	0.1127	ND	ND
	rs7580900	Intron	T	0.531, 0.473	0.554, 0.472	1.613	1.471	0.204	0.2252	ND	ND
DCSIGN (CD209)	rs4804803/ -336 A/G	Promoter	A	0.745, 0.380	0.534, 0.380	55.47	4.52	9.48E-14	3.34E-02	4.0 (2.6 - 6.0)	1.9 (1.0 -3.4)

SNP, single nucleotide polymorphism; ^adbSNP, the SNP database (<http://www.ncbi.nlm.nih.gov/projects/SNP>); Two tailed P<0.05 was considered significant. ^cp value from a chi-square test, significant comparisons indicated in bold, ^cConfidence Interval

Table 6. Allelic association for the *LTA4H* gene with and without the heterozygosity model.

SNP	Groups	N	Genotype Frequencies			Allele frequencies		Without heterozygosity model		Heterozygosity model (01 vs 00+11)			
			00	01	11	0	1	OR(CI) ^a	P-Value ^a	OR (CI) ^b	P-Value	Padj ^c	
rs1978331	All TB	185	0.26	0.44	0.30	0.48	0.52	1.02 (0.73-1.41)	0.9209	0.55 (0.34 -0.87)	0.013	0.039	
	T=0,C=1	PTB	135	0.25	0.46	0.29	0.48	0.52	1 (0.70 - 1.43)	0.9738	0.60 (0.36 - 0.99)	0.076	0.168
	LNTB	50	0.3	0.38	0.32	0.49	0.51	1.04 (0.65 -1.66)	0.8533	0.43 (0.22 - 0.85)	0.018	0.037	
	Control	120	0.18	0.59	0.23	0.47	0.52						
rs2660898			TT	TG	GG								
	T=0,C=1	All TB	185	0.17	0.41	0.42	0.38	0.63	1.21 (0.85 -1.7)	0.2761	0.76(0.48-1.2)	0.283	0.849
	PTB	135	0.15	0.42	0.44	0.36	0.64	1.1 (0.76 - 1.61)	0.5822	0.77(0.46-1.3)	0.309	0.927	
	LNTB	50	0.22	0.4	0.36	0.43	0.57	1.5(0.93-2.44)	0.093	0.75(0.38-1.5)	0.496	1.488	
	Control	120	0.09	0.48	0.43	0.33	0.67						
rs17525495			TT	TC	CC								
	All TB	185	0.06	0.24	0.69	0.19	0.82	1.33(0.83-2.1)	0.238	0.95(0.53-1.7)	0.8827	2.6481	
	T=0, C=1	PTB	135	0.05	0.25	0.65	0.19	0.81	1.31(0.78-2.2)	0.298	1.1(0.60-2.1)	0.753	2.259
	LNTB	50	0.08	0.18	0.62	0.19	0.80	1.38(0.72-2.65)	0.329	0.82(0.35-1.9)	0.831	2.493	
	Control	120	0.03	0.24	0.73	0.15	0.85						

^aOdds ratio for the minor allele are shown. ^aP-Value from a chi-square test. ^aConfidence Interval, ^bFor each SNP, odds ratios (OR) calculated for heterozygosity (01) versus homozygosity (00+11) for cases versus controls. ^cP-Values adjusted by Bonferroni correction for multiple tests. Significant comparisons are indicated in bold.

Since, we observed a uniform *TLR* gene polymorphism risk for both PTB and LNTB, we also genotyped four *TLR8* gene polymorphisms, as their genetic association has been shown to be important for outcome of TB. Uniquely, *TLR8* is located on the X chromosome, so the males as they have only one copy of the X chromosome, would be hemizygous. A stratified by sex analysis revealed a higher risk for males as expected carrying A allele for rs3788935 (17-fold risk), rs3761624 (4-fold risk). A risk for female population was also detected for rs3761624 which was lost after multiple corrections testing. If A allele is risk factor for males as they carry only one copy a corresponding homozygous phenotype can be a risk factor for females too as depicted by rs3761624. The sample size of LNTB group (n = 50) was limited for a stratified analysis by sex, so it was not carried out for LNTB (Table 6). Interestingly, in a case vs control analysis not stratified by sex *TLR8* PTB showed an increased risk for 3 (rs3788935, rs3761624, rs3764880) and LNTB 2 (rs3761624, rs3764879) among the four variants typed (Table S5).

Table 7. Sex-specific associations for *TLR8* gene variants in PTB.

db SNP ^a rsID	Risk allele	Males (Case=75; Control=70)			Female (Case=50; Control=54)				
		Case, Control Frequencies	P-Value	P Bonferroni	Odds Ratio (CI)	Case, Control Frequencies	P-Value	P Bonferroni	Odds Ratio (CI)
rs3788935	A	0.1102, 0.01	0.0072	0.0288	17 (0.94 - 300)	0.160, 0.106	0.1673	0.6692	1.6 (0.82-3.2)
rs3761624	A	0.2, 0.05	0.012	0.048	4.3 (1.2-16)	0.356, 0.244	0.0423	0.1692	1.7 (1.0-2.9)
rs3764880	A	0.107, 0.029	0.1012	0.4048	3.9 (0.81-19)	0.229, 0.149	0.0557	0.2228	1.7 (0.98-2.9)
rs3764879	C	0.558, 0.482	0.5934	2.3736	1.3 (0.61-2.6)	0.489, 0.440	0.3981	1.5924	1.2 (0.77-1.9)

3.6. Pathway analysis reveals an apoptotic axis for LNTB and a cytokine axis for PTB

SNP association and their respective p-values from the study was used as an input to identify associated modules from a protein-protein interaction network, which was used to identify the associated pathways and the results obtained were subjected to a gene-ontology annotation using WikiPathways [29, 30]. All the SNPs that are significantly associated in this study and previous studies on this population [19, 26] were considered. An abundance for cytokine pathways was seen for PTB (Figure 3 A), while abundance of apoptosis modules was seen for LNTB (Figure 3 B). This pathway level difference is in line with the genetic association findings presented above, showing these differential genetic association could contribute to differential pathway activation and hence activate the immune response distinctly. This adds another level of evidence of differential host genotype being responsible for different manifestation of PTB and LNTB.

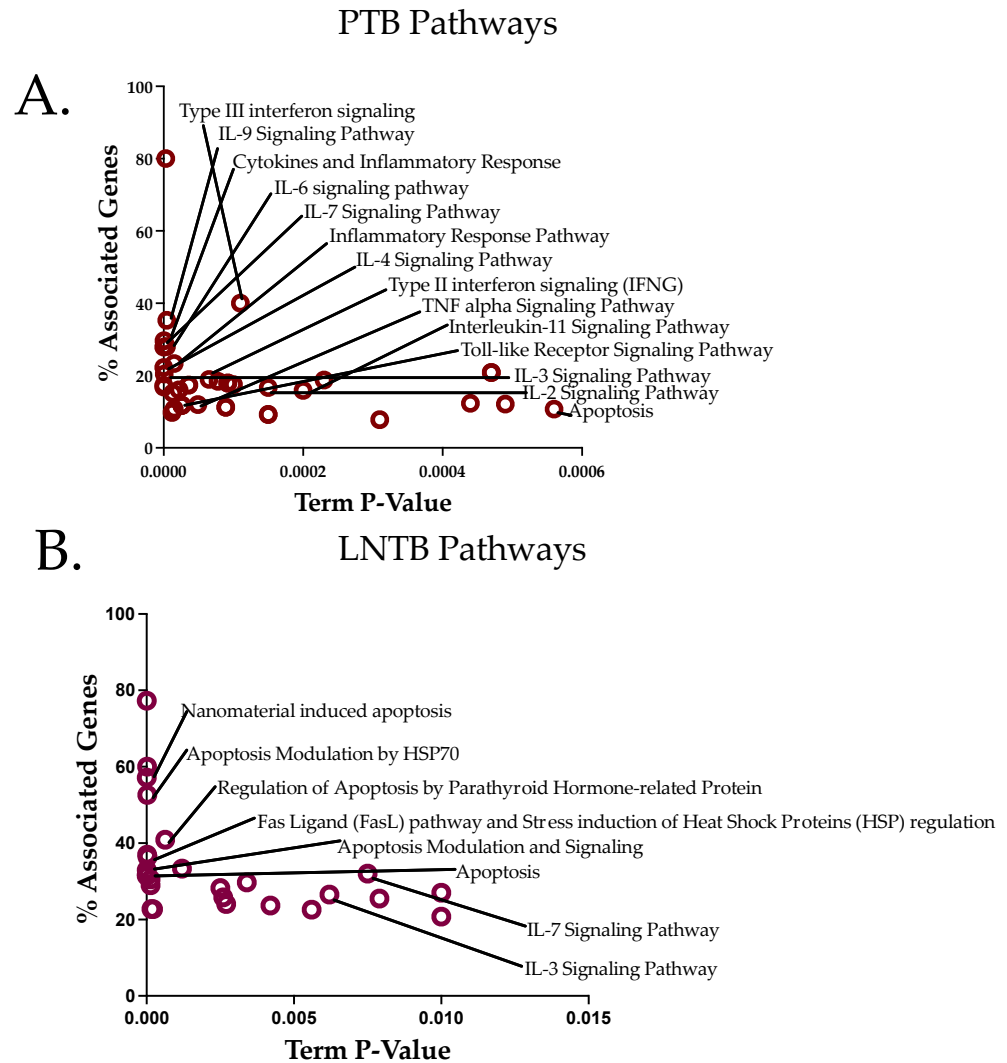


Figure 3. PTB and LNTB risk associated SNP related genes enrich in different pathways. Pathways enriched in PTB and LNTB through a pathway analysis based on the SNPs with significant P-values. The modules obtained from JActivemodules were used for enrichment in ClueGO App in Cytoscape. Two metrics, %Associated genes (proportion of genes in the pathway from the network of all genes) and Term P-value (enrichment P-value) is plotted here to show the abundance of the pathway with all the input genes and the P-value for the GO term obtained from WikiPathways.

4. Discussion

Human genetic diversity is hugely impacted by co-evolving pathogens such as *Mtb* [38]. Candidate gene studies using the case – control design provides one of the most direct means of identifying human genetic variants that currently impact on susceptibility to infectious disease. Such information would help better the understanding of disease pathogenesis, disease resistance at an individual level, that could inform targeted intervention strategies based on their genotype as has been successfully implemented [12].

Several studies have shown that TB susceptibility has a genetic component (summarized in [22, 39], but comparative studies on genetic susceptibility to different forms of TB are limited [21, 40]. Such studies can provide insight into the role genetic polymorphisms in different manifestations of TB. Even rarer are studies on genetic susceptibility to EPTB. Some of the studied forms of EPTB have been involving multiple sites [41], LNTB [19, 42, 43], TB meningitis [43, 44], intestinal TB [45], bone [43] and pleural [43].

In this study, we aimed to do a comparative study between PTB and LNTB to investigate differential genetic associations between PTB and LNTB. We tested several candidate gene polymorphisms, never investigated before, as associated with differential TB susceptibility (Table S1). In addition, we validated susceptibility loci previously identified in other populations [22, 39] and our previous studies [19, 26]. This is important as ethnic validation of commonly reported genetic variants in different populations is desirable. In total, 63 polymorphisms across 23 genes were selected and genotyped from both the innate and adaptive immune branches of immunity to TB in the north Indian population and their allele frequencies compared and linkage disequilibrium (LD) and haplotypes investigated. Thus, we have employed a comprehensive coverage of SNPs and genes to compare the genetic susceptibility differences between PTB and LNTB.

In our study, genetic variants in the cytokine were validated to be significant risk factors PTB (Table 3). We also showed that a significant gene – gene interaction among cytokine SNPs may further accentuate the importance of the identified SNPs in governing the genetic susceptibility to PTB. Interestingly, we didn't find significant cytokine gene polymorphisms associate with LNTB. The important difference was lack of a major association between cytokine SNPs and serum cytokine levels in LNTB, which has been shown to be associated with PTB in multiple studies [37, 46-48]. The difference between such association clearly shows that distinct genetic association associates with PTB and LNTB. Highly enriched cytokine pathways in PTB and limited in LNTB (Figure 3) add strength to this argument.

Interestingly the innate immunity genes, *P2RX7* [49] and *DCSIGN* [50-52], which are critical for immune response to *Mtb*, were risk factors for both PTB and LNTB, as expected. *P2RX7* has been very widely studied as risk factor for both PTB and EPTB. Macrophages from patients with loss of function homozygous allele for rs3751143, could not kill *Mtb* in vitro in EPTB [16]. We didn't observe any association with this variant in our study. Interestingly, another functional variant rs2393799, showed an increased risk for both PTB and LNTB, but the risk was much higher for LNTB (7-fold as compared to 1.5-fold for PTB). *P2RX7* is known to have a role in apoptosis of *Mtb* infected macrophage [24]. Similar theme was seen for *SP110* gene for which we have previously identified a risk for rs1427294 of in LNTB but not pulmonary TB [19]. Recently, this gene has been shown to inhibit apoptosis of infected macrophages, thereby resisting *Mtb* infection [23]. This in conjunction of identifying more apoptotic pathways do suggest that the apoptotic axis may be important in LNTB. The other genetic variants of importance in the *LTA4H* gene showed heterozygous protection in LNTB and not PTB (Table 6). So, it emerges from the current study that variations in the innate immune variants have a closer relation to development of LNTB, while the cytokine genetic variants have little influence and associations in LNTB. Similarly, among the pattern recognition receptors, *TLR2* and *TLR4* showed risks for PTB and not LNTB, while *TLR8* showed risk for both PTB and LNTB and more risk for males in PTB. This adds to the theme of differential association among PTB and LNTB, (*TLR2,4* while a conserved risk among critical genes susceptibility (*TLR8*). *TLRs* variants have been shown to be critical risk factors for TB[53-56] but have not been studied for LNTB. Although, limited by sample size there appears to be differences in TLR gene polymorphisms in PTB and LNTB.

Similar to our study a few other studies have shown a selective genetic association with EPTB, for example, like our study (Table 6) Yang et al show that *LTA4H* gene polymorphism rs1978331 and rs2540474, are only risk factor with EPTB and not PTB in Han Chinese population [43]. Similarly, a GWAS could identify 4 loci that were only associated with EPTB and not PTB [21]. These studies support the differential nature of genetic polymorphism in EPTB, which is distinct from PTB. Similar studies are warranted for validation in a larger sample size and in multiple populations to test whether genetic polymorphism can associate with various forms of tuberculosis. The limitation of the study is that for certain polymorphisms we could not achieve a good sample size and thus the results need to be validated in a larger sample size.

5. Conclusions

Our study contributes to the growing knowledge that PTB and EPTB manifestations have a genetic basis, and in our study the differences are more polymorphic cytokine genes in PTB and more polymorphic apoptosis/innate genes in LNTB.

Supplementary Materials: Figure S1A-O: Representative ARMS-PCR gel images for the typed variants; Table S1: Concordance ARMS-PCR and Sequenom Genotyping for *IL12* at rs3212220; Table S2a: List of 39 Cytokine SNPs for the discovery Panel; Table S2b: List of 15 Cytokine SNPs for the Validation Panel; Table S3: Allelic association for the Discovery Panel; Table S4a: Validation results of the Discovery Panel; Table S4 b: Haplotypic association following the heterozygosity model for *LTA4H* polymorphisms; Table S5: Comparative allelic association among PTB and LNTB for *TLR8* gene.

Author Contributions: Conceptualization, A and MB.; methodology, A, MB.; software, A; formal analysis, A.; investigation, A, AG.; resources, MB, MVB.; data curation, A.; writing—original draft preparation, A; writing—review and editing, A, MB, MVB.; visualization, A.; supervision, MB, MVB; funding acquisition, MB. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Council for Scientific and Industrial Research; Government of India funded the Research and Fellowship for Dr. Abhimanyu via Grant Number: 60/80 EMR-II. The research was also supported by Indian council of medical research (ICMR) grant No. 5/8/5/14/2012-ECD-1 for the work on *SP110* and a Senior Research fellowship to Dr. Abhimanyu. The authors also acknowledge an Emeritus Professor Fellowship to Dr. Bose from ICMR.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of V P Chest Institute, University of Delhi, Delhi, India.

Informed Consent Statement: Informed written consent was obtained from all subjects involved in the study.

Data Availability Statement: Data for the study has been summarized in multiple tables for the study. If needed, the data can be obtained from the first and corresponding authors upon reasonable request.

Acknowledgments: The authors would like to acknowledge the help of Dr. Shyama Jain of Lok Nayak TB hospital, Drs. Nilang Tale, Dr. J N Banavaliker of Rajan Babu Institute of pulmonary medicine for facilitating study enrolment. Contribution of Dr. Ashima Jain is acknowledged for enrolment of LNTB cases. We would like to acknowledge Dr. Pankaj Jha and the Indian Genome Variation consortium for making genotype data for additional controls available for the validation cohort analyses.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Some of the results were submitted to the funders as a part of the final project report. This work was part of doctoral thesis of Dr. Abhimanyu, Th 19865, University of Delhi. This work has been presented as poster at the Indian Society for Human Genetics conference by Dr Abhimanyu and at the Indian association of medical microbiologist meeting by Dr A Giri.

References

1. Organisation WH. Global tuberculosis Report 2017. Geneva; 2017.
2. Sharma SK, Mohan A. Extrapulmonary tuberculosis. Indian J Med Res. 2004;120(4):316-53.
3. Polesky A, Grove W, Bhatia G. Peripheral tuberculous lymphadenitis: epidemiology, diagnosis, treatment, and outcome. Medicine (Baltimore). 2005;84(6):350-62.
4. Garcia de Viedma D LG, Cardona PJ, Rodriguez NA, Gordillo S, Serrano MJ, Bouza E. Association between the infectivity of Mycobacterium tuberculosis strains and their efficiency for extrarespiratory infection. J Infect Dis. 2005;12(195):2059-65.

5. Lee J, Yun YJ, Kqueen CY, Lee JH, Kim HY, Kim YR, et al. pncA Mutations in the Specimens from Extrapulmonary Tuberculosis. *Tuberc Respir Dis (Seoul)*. 2012;72(6):475-80.
6. Yang Z, Yang D, Kong Y, Zhang L, Marrs CF, Foxman B, et al. Clinical Relevance of Mycobacterium tuberculosis plcD Gene Mutations. *American Journal of Respiratory and Critical Care Medicine*. 2005;171(12):1436-42.
7. Varma-Basill M, Nair D. Molecular epidemiology of tuberculosis: Opportunities & challenges in disease control. *Indian Journal of Medical Research*. 2017;146(1).
8. Ruesen C, Chaidir L, van Laarhoven A, Dian S, Ganiem AR, Nebenzahl-Guimaraes H, et al. Large-scale genomic analysis shows association between homoplastic genetic variation in Mycobacterium tuberculosis genes and meningeal or pulmonary tuberculosis. *BMC Genomics*. 2018;19(1).
9. Caws M, Thwaites G, Dunstan S, Hawn TR, Lan NT, Thuong NT, et al. The influence of host and bacterial genotype on the development of disseminated disease with Mycobacterium tuberculosis. *PLoS Pathog*. 2008;4(3):e1000034.
10. Coussens AK, Wilkinson RJ, Nikolayevskyy V, Elkington PT, Hanifa Y, Islam K, et al. Ethnic variation in inflammatory profile in tuberculosis. *PLoS Pathog*. 2013;9(7):e1003468.
11. Thuong NT, Hawn TR, Thwaites GE, Chau TT, Lan NT, Quy HT, et al. A polymorphism in human TLR2 is associated with increased susceptibility to tuberculous meningitis. *Genes Immun*. 2007;8(5):422-8.
12. Tobin DM, Vary JC, Jr., Ray JP, Walsh GS, Dunstan SJ, Bang ND, et al. The It4h locus modulates susceptibility to mycobacterial infection in zebrafish and humans. *Cell*. 2010;140(5):717-30.
13. Hasan Z, Jamil B, Khan J, Ali R, Khan MA, Nasir N, et al. Relationship between circulating levels of IFN-gamma, IL-10, CXCL9 and CCL2 in pulmonary and extrapulmonary tuberculosis is dependent on disease severity. *Scand J Immunol*. 2009;69(3):259-67.
14. Wilkinson RJ, Llewelyn M, Toossi Z, Patel P, Pasvol G, Lalvani A, et al. Influence of vitamin D deficiency and vitamin D receptor polymorphisms on tuberculosis among Gujarati Asians in west London: a case-control study. *Lancet*. 2000;355(9204):618-21.
15. Motsinger-Reif AA, Antas PR, Oki NO, Levy S, Holland SM, Sterling TR. Polymorphisms in IL-1beta, vitamin D receptor Fok1, and Toll-like receptor 2 are associated with extrapulmonary tuberculosis. *BMC Med Genet*. 2010;11:37.
16. Fernando SL, Saunders BM, Sluyter R, Skarratt KK, Goldberg H, Marks GB, et al. A Polymorphism in the P2X7 Gene Increases Susceptibility to Extrapulmonary Tuberculosis. *American Journal of Respiratory and Critical Care Medicine*. 2007;175(4):360-6.
17. Shandera W, Webster A. The extrapulmonary dissemination of tuberculosis: A meta-analysis. *International Journal of Mycobacteriology*. 2014;3(1).
18. Tailleux L, Fox GJ, Sy DN, Nhung NV, Yu B, Ellis MK, et al. Polymorphisms of SP110 Are Associated with both Pulmonary and Extra-Pulmonary Tuberculosis among the Vietnamese. *PLoS ONE*. 2014;9(7).
19. Abhimanyu, Jha P, Jain A, Arora K, Bose M. Genetic association study suggests a role for SP110 variants in lymph node tuberculosis but not pulmonary tuberculosis in north Indians. *Human Immunology*. 2011;72(7):576-80.
20. Wilkinson RJ, Patel P, Llewelyn M, Hirsch CS, Pasvol G, Snounou G, et al. Influence of polymorphism in the genes for the interleukin (IL)-1 receptor antagonist and IL-1beta on tuberculosis. *J Exp Med*. 1999;189(12):1863-74.
21. Oki NO, Motsinger-Reif AA, Antas PR, Levy S, Holland SM, Sterling TR. Novel human genetic variants associated with extrapulmonary tuberculosis: a pilot genome wide association study. *BMC Res Notes*. 2011;4:28.
22. Möller M, Hoal EG. Current findings, challenges and novel approaches in human genetic susceptibility to tuberculosis. *Tuberculosis*. 2010;90(2):71-83.
23. Wu Y, Guo Z, Liu F, Yao K, Gao M, Luo Y, et al. Sp110 enhances macrophage resistance to Mycobacterium tuberculosis via inducing endoplasmic reticulum stress and inhibiting anti-apoptotic factors. *Oncotarget*. 2017;8(38):64050-65.
24. Wiley JS, Sluyter R, Gu BJ, Stokes L, Fuller SJ. The human P2X7 receptor and its role in innate immunity. *Tissue Antigens*. 2011;78(5):321-32.
25. Indian Genome Variation C. The Indian Genome Variation database (IGVdb): a project overview. *Hum Genet*. 2005;118(1):1-11.

26. Abhimanyu, Bose M, Jha P, IGVC. Footprints of genetic susceptibility to pulmonary tuberculosis: Cytokine gene variants in north Indians. *Indian Journal of Medical Research*. 2012;135(5):763-70.
27. Ye S, Dhillon S, Ke X, Collins AR, Day IN. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res*. 2001;29(17):E88-8.
28. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. *Science*. 2002;296(5576):2225-9.
29. Bakir-Gungor B, Egemen E, Sezerman OU. PANOGA: a web server for identification of SNP-targeted pathways from genome-wide association study data. *Bioinformatics*. 2014;30(9):1287-9.
30. Burcu Bakir-Gungor OUS. Identification of SNP Targeted Pathways From Genome-wide Association Study (GWAS) Data. *Protocol Exchange*. 2012.
31. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-504.
32. Ideker T, Ozier O, Schwikowski B, Siegel AF. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*. 2002;18 Suppl 1:S233-40.
33. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*. 2009;25(8):1091-3.
34. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, et al. WikiPathways: connecting communities. *Nucleic Acids Res*. 2021;49(D1):D613-D21.
35. Abhimanyu, Bose M. Reporting genetic association studies: the roadblocks and guiding rules for robust results. *Lung*. 2012;190(6):587-8; author reply 8.
36. Dupont W, Plummer WJ. PS: Power and Sample Size Calculation 2014 [Available from: <http://biostat.mc.vanderbilt.edu/wiki/Main/PowerSampleSize>].
37. Abhimanyu, Mangangcha IR, Jha P, Arora K, Mukerji M, Banavaliker JN, et al. Differential serum cytokine levels are associated with cytokine gene polymorphisms in north Indians with active pulmonary tuberculosis. *Infect Genet Evol*. 2011;11(5):1015-22.
38. Brites D, Gagneux S. Co-evolution of Mycobacterium tuberculosis and Homo sapiens. *Immunol Rev*. 2015;264(1):6-24.
39. Yim JJ, Selvaraj P. Genetic susceptibility in tuberculosis. *Respirology*. 2010;15(2):241-56.
40. Souza De Lima D, Bomfim CCB, Leal VNC, Reis EC, Soares JLS, Fernandes FP, et al. Combining Host Genetics and Functional Analysis to Depict Inflammation Contribution in Tuberculosis Susceptibility and Outcome in Endemic Areas. *Front Immunol*. 2020;11:550624.
41. Fernando SL, Saunders BM, Sluyter R, Skarratt KK, Goldberg H, Marks GB, et al. A polymorphism in the P2X7 gene increases susceptibility to extrapulmonary tuberculosis. *Am J Respir Crit Care Med*. 2007;175(4):360-6.
42. Abhimanyu, Bose M, Komal, Varma-Basil M. Lack of association between IL17A and IL17F polymorphisms and related serum levels in north Indians with tuberculosis. *Gene*. 2013;529(1):195-8.
43. Yang J, Chen J, Yue J, Liu L, Han M, Wang H. Relationship between human LTA4H polymorphisms and extra-pulmonary tuberculosis in an ethnic Han Chinese population in Eastern China. *Tuberculosis (Edinb)*. 2014;94(6):657-63.
44. Tobin DM, Roca FJ, Oh SF, McFarland R, Vickery TW, Ray JP, et al. Host genotype-specific therapies can optimize the inflammatory response to mycobacterial infections. *Cell*. 2012;148(3):434-46.
45. Yu ZQ, Wang WF, Zhu CZ, Zhang KH, Chen XC, Chen JY. The Polymorphism rs17525495 of LTA4H Is Associated with Susceptibility of Crohn's Disease instead of Intestinal Tuberculosis in a Chinese Han Population. *Biomed Res Int*. 2019;2019:9537050.
46. Joshi L, Ponnana M, Sivangala R, Chelluri LK, Nallari P, Penmetsa S, et al. Evaluation of TNF-alpha, IL-10 and IL-6 Cytokine Production and Their Correlation with Genotype Variants amongst Tuberculosis Patients and Their Household Contacts. *PLoS One*. 2015;10(9):e0137727.

47. Peresi E, Oliveira LR, da Silva WL, da Costa EA, Araujo JP, Jr., Ayres JA, et al. Cytokine Polymorphisms, Their Influence and Levels in Brazilian Patients with Pulmonary Tuberculosis during Antituberculosis Treatment. *Tuberc Res Treat.* 2013;2013:285094.
48. Singh AK, Yadav AB, Garg R, Misra A. Single nucleotide polymorphic macrophage cytokine regulation by *Mycobacterium tuberculosis* and drug treatment. *Pharmacogenomics.* 2014;15(4):497-508.
49. Soare AY, Freeman TL, Min AK, Malik HS, Osota EO, Swartz TH. P2RX7 at the Host-Pathogen Interface of Infectious Diseases. *Microbiol Mol Biol Rev.* 2021;85(1).
50. Lugo-Villarino G, Troegeler A, Balboa L, Lastrucci C, Duval C, Mercier I, et al. The C-Type Lectin Receptor DC-SIGN Has an Anti-Inflammatory Role in Human M(IL-4) Macrophages in Response to *Mycobacterium tuberculosis*. *Front Immunol.* 2018;9:1123.
51. Tailleux L, Pham-Thi N, Bergeron-Lafaurie A, Herrmann JL, Charles P, Schwartz O, et al. DC-SIGN induction in alveolar macrophages defines privileged target host cells for mycobacteria in patients with tuberculosis. *PLoS Med.* 2005;2(12):e381.
52. Tailleux L, Schwartz O, Herrmann JL, Pivert E, Jackson M, Amara A, et al. DC-SIGN is the major *Mycobacterium tuberculosis* receptor on human dendritic cells. *J Exp Med.* 2003;197(1):121-7.
53. Wang JJ, Xia X, Tang SD, Wang J, Deng XZ, Zhang Y, et al. Meta-analysis on the associations of TLR2 gene polymorphisms with pulmonary tuberculosis susceptibility among Asian populations. *PLoS One.* 2013;8(10):e75090.
54. Varshney D, Singh S, Sinha E, Mohanty KK, Kumar S, Kumar Barik S, et al. Systematic review and meta-analysis of human Toll-like receptors genetic polymorphisms for susceptibility to tuberculosis infection. *Cytokine.* 2022;152:155791.
55. Zhou Y, Zhang M. Associations between genetic polymorphisms of TLRs and susceptibility to tuberculosis: A meta-analysis. *Innate Immun.* 2020;26(2):75-83.
56. Davila S, Hibberd ML, Hari Dass R, Wong HE, Sahiratmadja E, Bonnard C, et al. Genetic association and expression studies indicate a role of toll-like receptor 8 in pulmonary tuberculosis. *PLoS Genet.* 2008;4(10):e1000218.