Supplementary Material: Metabolite Selection Process

The metabolite selection process followed a sequential filtering strategy. Initially, a total of 5,011 molecular features were detected in positive ion mode. After quality control filtering (removing features with >30% missing values), 3,847 features remained. Using the OPLS-DA model with criteria of VIP > 1 and $P < 0.05$, 371 differentially expressed metabolites were identified. Subsequent filtering by fold change ($FC \geq 2$ or ≤ 0.5) reduced the number to 198 significant differential metabolites. Among these, 43 metabolites were successfully annotated using MS/MS fragmentation matching with public databases (HMDB, METLIN). Finally, ROC curve analysis identified four metabolites (leucylleucine, tryptophyl-phenylalanine, biliverdin, and lysoPE(18:1(11Z)/0:0) with AUC > 0.8 as the key candidate biomarkers.

Supplementary Table 1. The AUC of 43 biomarkers between LTBI and non-LTBI groups.

Compound Name	Group	AUC
Tryptophyl-Phenylalanine	down	0.975202
Leucylleucine	down	0.981263
Biliverdin	up	0.8555
LysoPE (18:1(11Z)/0:0)	down	0.8405
N-(1-Deoxy-1-fructosyl)phenylalanine	down	0.7992
Proline betaine	Down	0.773
3-Hydroxyocta-2,4-dienoylcarnitine	Down	0.7652
N2-Acetylornithine	Down	0.7644
Carboxyethyl-hydroxychroman	Down	0.762
Lumirubin	up	0.7599
gamma-Glutamylvaline	Down	0.7578
Lysine	Down	0.7506
alpha-Terpineol butanoate	up	0.7482
Chenodeoxycholic acid glycine conjugate	down	0.7433
Fructosylvaline	down	0.7268
N-Hexadecanoylpyrrolidine	up	0.721
Citpressine I	down	0.703
Hypoxanthine	down	0.7002
N-Methyl-L-proline	down	0.6931
Bilirubin	up	0.6917
L-Arginine	down	0.688
Caffeine	down	0.6851
Hydroxyprolyl-Proline	down	0.6771
2-methyl-1,3-thiazolidine-2,4-dicarboxylic acid	up	0.6768
N-(1-Deoxy-1-fructosyl)leucine	down	0.6595
N-Docosahexaenoyl GABA	down	0.6493
Muramic acid	down	0.6385

4-Trimethylammoniobutanoic acid	down	0.6302
Glutamylglutamine	down	0.6141
2-Hydroxypurine	down	0.6125
Oleamide	up	0.5955
2-Hydroxypurine	down	0.5875
Deoxycholic acid glycine conjugate	down	0.5391
Tetracosahexaenoic acid	down	0.466
LysoPE(18:0/0:0)	up	0.4464
LysoPC(14:0/0:0)	up	0.4298
LysoPC(15:0/0:0)	up	0.4271
LysoPC(17:0/0:0)	up	0.4208
Biliverdin	up	0.4155
LysoPE(16:0/0:0)	up	0.4153
LysoPC(20:2(11Z,14Z)/0:0)	up	0.4081
Aspartylhydroxyproline	up	0.3505
(4E,7E,10E,13E)-Hexadeca-4,7,10,13-tetraenoic acid	up	0.3401

Abbreviations: AUC, area under the receiving operating characteristic curve.

Supplementary Table 2. AUC, sensitivity, and specificity for ROC curves calculated at optimal cutoff of Leucyl leucine, Tryptophyl-Phenylalanine, Biliverdin, LysoPE(18:1(11Z)/0:0).

Compound name	AUC	Sensitivity	Specificity	P-value	Fold Change
Leucylleucine	0.98	0.96	0.94	3.28E-13	32.9447
Tryptophyl-Phenylalanine	0.98	0.96	0.97	3.38E-34	15.6357
Biliverdin	0.86	0.91	0.70	6.74E-14	0.0915827
LysoPE(18:1(11Z)/0:0)	0.84	0.69	0.85	6.74E-15	2.09266

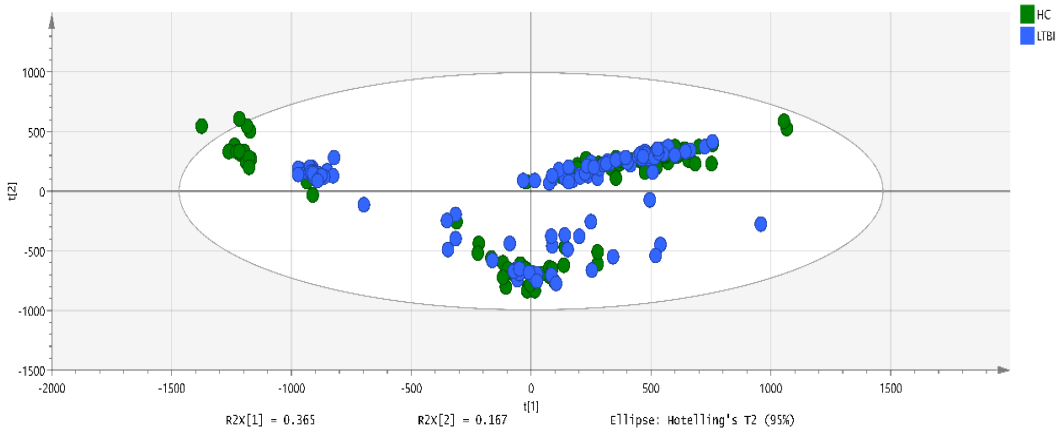
Abbreviations: AUC, area under the receiving operating characteristic curve.

Supplementary Table 3. Cross-validation results of the random forest model combining four key metabolites

Metric	Value	95% CI
AUC	0.9998	0.9996–1.0000
Sensitivity	0.998	-
Specificity	0.986	-
Best mtry parameter	1	-

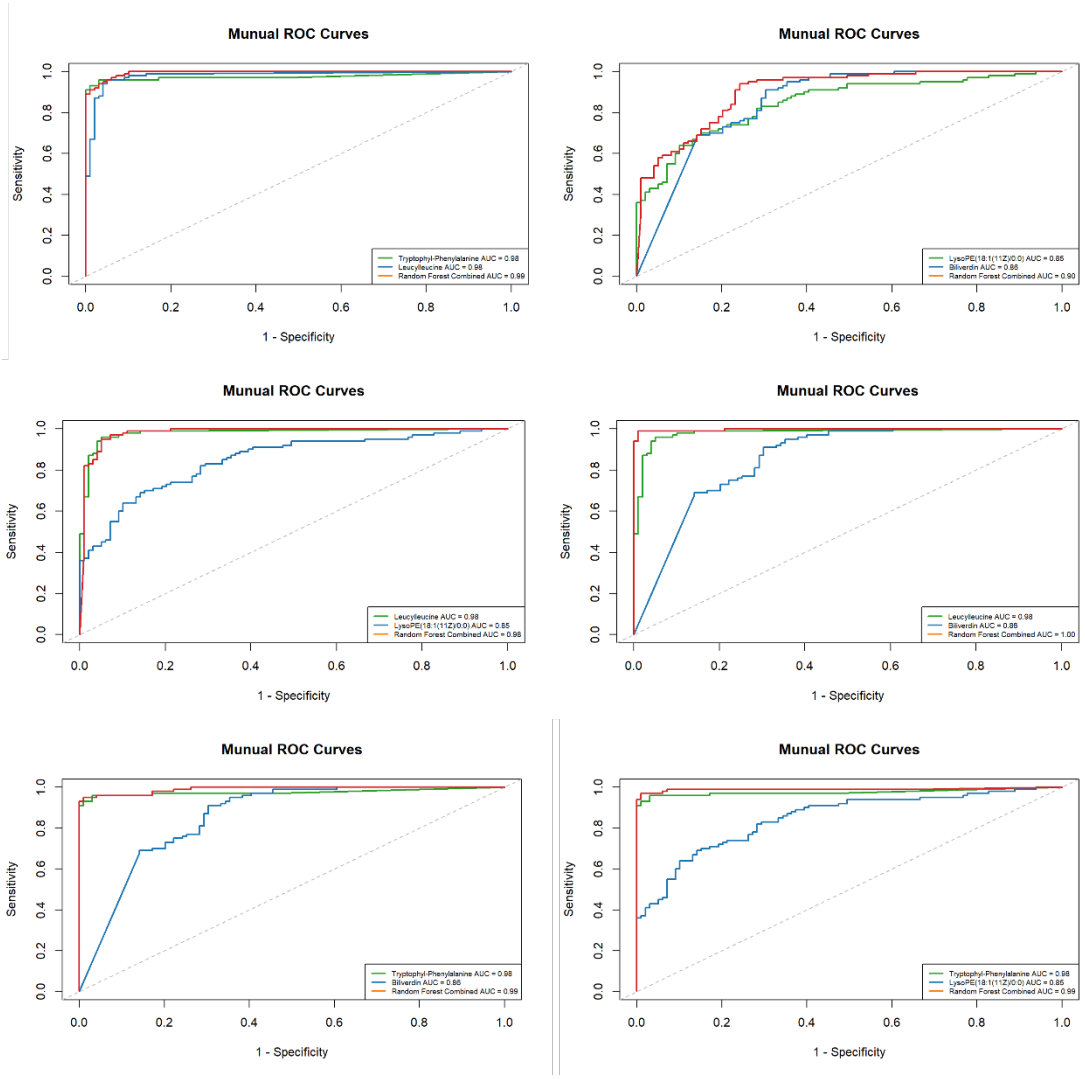
Supplementary Table 4. Complete regression results including covariate effects

Metabolite	Variable	β	SE	t-value	P-value
Leucylleucine	group (LTBI vs non-LTBI)	4.148	0.201	20.60	7.20×10^{-50}
	Age	0.003	0.005	0.68	0.500
	Sex (Female vs Male)	-0.263	0.248	-1.06	0.290
	Smoking (Former vs Never)	-0.115	0.298	-0.52	0.604
	Smoking (Current vs Never)	0.144	0.283	0.51	0.612
Tryptophyl-phenylalanine	group (LTBI vs non-LTBI)	3.449	0.292	11.79	2.51×10^{-24}
	Age	0.013	0.007	1.82	0.070
	Sex (Female vs Male)	-0.515	0.360	-1.43	0.155
	Smoking (Former vs Never)	-0.206	0.432	-0.48	0.634
	Smoking (Current vs Never)	-0.014	0.411	-0.04	0.972
Biliverdin	group (LTBI vs non-LTBI)	-3.410	0.395	-8.63	2.66×10^{-15}
	Age	-0.007	0.009	-0.70	0.482
	Sex (Female vs Male)	0.137	0.487	0.28	0.778
	Smoking (Former vs Never)	-0.182	0.584	-0.31	0.756
	Smoking (Current vs Never)	1.362	0.555	2.45	0.015
LysoPE(18:1(11Z)/0:0)	group (LTBI vs non-LTBI)	0.654	0.117	5.58	8.58×10^{-8}
	Age	0.007	0.003	2.32	0.022
	Sex (Female vs Male)	0.010	0.144	0.07	0.943
	Smoking (Former vs Never)	-0.050	0.173	-0.29	0.772
	Smoking (Current vs Never)	0.182	0.165	1.10	0.271

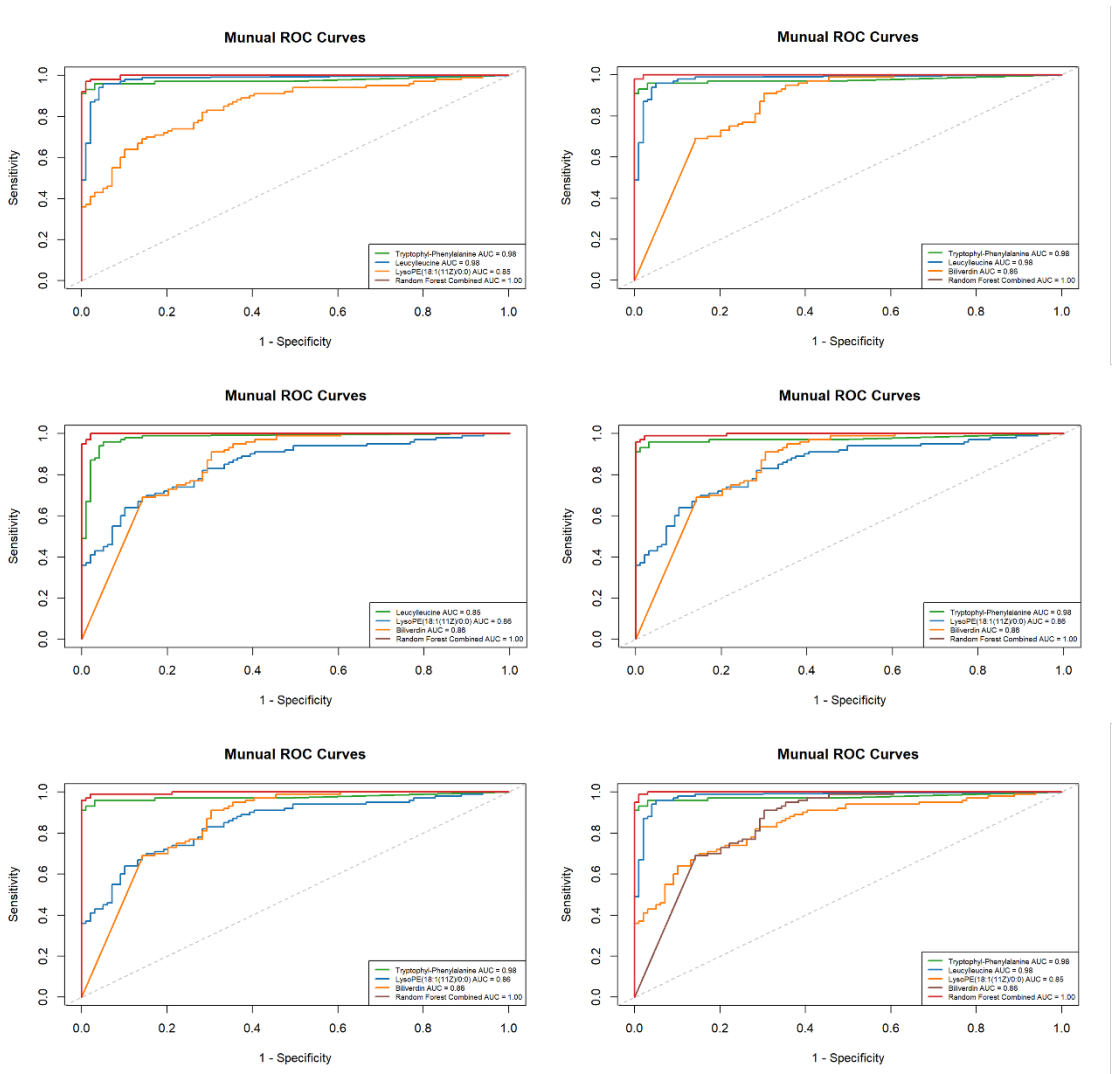


Supplementary Figure 1. Scatter plot of PCA model of the LTBI and non-LTBI groups.

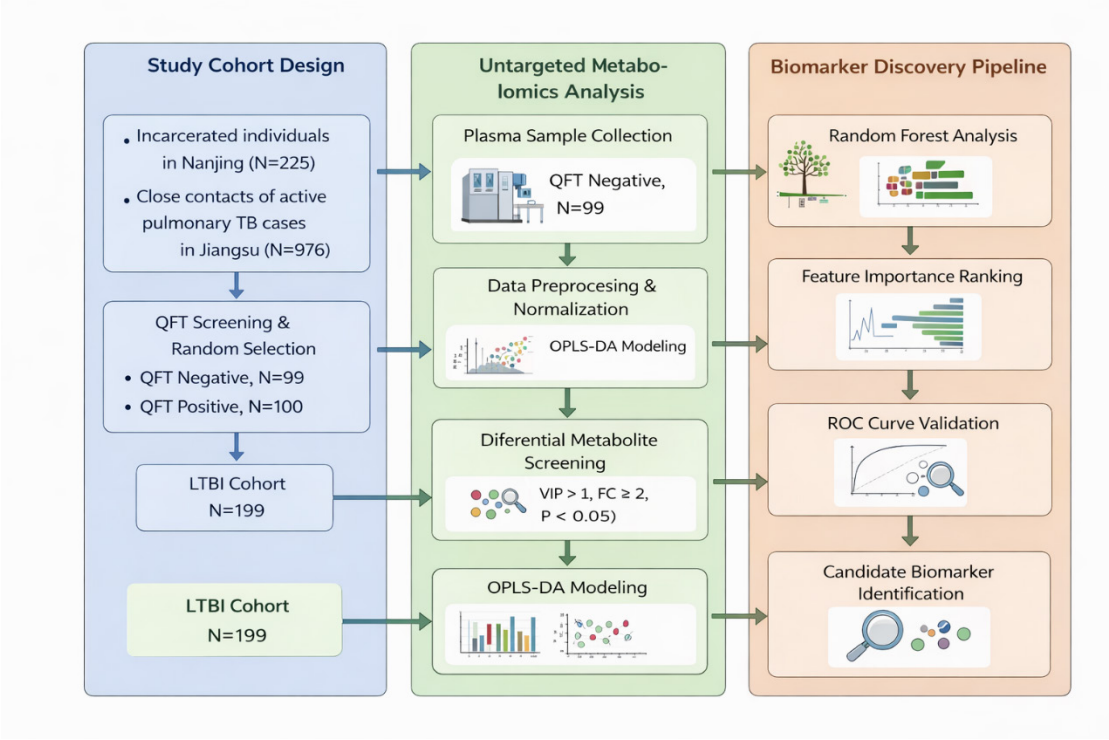
Abbreviations: PCA, Principal component analysis; Orthogonal partial least squares-discriminant analysis; LTBI, latent tuberculosis infection.



Supplementary Figure 2. Combined ROC curves of pairwise combinations of four metabolites (leucylleucine, tryptophyl-phenylalanine, biliverdin, and lysoPE(18:1(11Z)/0:0) for distinguishing LTBI from non-LTBI.



Supplementary Figure 3. Combined ROC curves of triple combinations of four metabolites (leucylleucine, tryptophyl-phenylalanine, biliverdin, and lysoPE(18:1(11Z)/0:0) for distinguishing LTBI from non-LTBI.



Supplementary Figure 7 Schematic workflow of metabolomic analysis and biomarker discovery.