

Comparative whole-genome sequence analysis of *Mycobacterium tuberculosis* isolated from tuberculous meningitis and pulmonary tuberculosis patients

Kiatichai Faksri^{1,2,*}, Eryu Xia³, Rick Twee-Hee Ong³, Jun Hao Tan⁴, Ditthawat Nonghanphithak^{1,2}, Nampueng Makhao⁵, Nongnard Thamnongdee⁵, Arirat Thanormchat⁵, Arisa Phurattanakornkul⁵, Somcharn Rattanarangsee⁵, Chate Ratanajaraya⁵, Prapat Suriyaphol⁶, Therdsak Prammananan⁷, Yik-Ying Teo^{3,8,9,10,11}, Angkana Chaiprasert^{5*}

¹Department of Microbiology Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand.

²Research and Diagnostic Center for Emerging Infectious Diseases (RCEID), Khon Kaen University, Khon Kaen 40002, Thailand.

³Saw Swee Hock School of Public Health, National University of Singapore 119077, Singapore.

⁴Duke-NUS Medical School, National University of Singapore 119077, Singapore.

⁵ Drug Resistant Tuberculosis Laboratory, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.

⁶Bioinformatics and Data Management for Research Unit, Office for Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.

⁷National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Ministry of Science and Technology, Pathum Thani 12120, Thailand.

⁸NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore 119077, Singapore.

⁹Genome Institute of Singapore 138672, Singapore.

¹⁰Department of Statistics and Applied Probability, National University of Singapore 119077, Singapore.

¹¹Life Sciences Institute, National University of Singapore 119077, Singapore.

Correspondence and requests for materials should be addressed to K.F (email: kiatichai@kku.ac.th) and A.C (email: angkana.cha@mahidol.ac.th)

Positions (SNPs)	Genes	Gene function	Note
Common missense SNPs specifically found in <i>Mtb</i> causing TBM compared to PulTB			
12,210 (A>G)	<i>Rv0103c (ctpB)</i>	Cation-transporting ATPase	
225,323 (T>C)	<i>Rv0193c</i>	Unknown	
475,178 (T>C)	<i>Rv0395</i>	Unknown	
502,589 (C>G)	<i>Rv0417 (thiG)</i>	Thiamine biosynthesis	
598,475 (G>A)	<i>Rv0507 (mmpL2)</i>	Unknown (Fatty acid transport)	
623,472 (A>G)	<i>Rv0532 (PE_PGRS6)</i>	Unknown	LIV
623,508 (C>G)	<i>Rv0532 (PE_PGRS6)</i>	Unknown	LIV
648,002 (T>G)	<i>Rv0556</i>	Unknown	
775,639 (T>C)	<i>Rv0676c (mmpL5)</i>	Unknown (Fatty acid transport)	
1,552,547 (G>A)	<i>Rv1378c (Rv1378c)</i>	Unknown	
1,885,772 (G>A)	<i>Rv1662 (pks8)</i>	Polyketide synthesis	
1,944,402 (T>C)	<i>Rv1716</i>	Unknown	
2,057,774 (A>T)	<i>Rv1815</i>	Unknown	
2,143,328 (G>C)	<i>Rv1895</i>	Unknown (cellular metabolism)	
2,269,780 (T>C)	<i>Rv2024c</i>	Unknown	
2,270,102 (A>G)	<i>Rv2024c</i>	Unknown	
2,362,041 (C>A)	<i>Rv2101 (helZ)</i>	Has helicase activity	
2,386,389 (G>A)	<i>Rv2125 (Rv2125)</i>	Unknown	
2,415,656 (G>C)	<i>Rv2155c (murD)</i>	Peptidoglycan biosynthesis	
2,910,461 (G>T)	<i>Rv2584c (apt)</i>	Purine salvage	
2,911,293 (C>G)	<i>Rv2585c</i>	Unknown	
3,296,843 (A>G)	<i>Rv2947c (pks15)</i>	Polyketide synthesis	
3,591,063 (T>C)	<i>Rv3213c</i>	Unknown (partitioning regulation)	
3,736,628 (T>G)	<i>Rv3343c (PPE54)</i>	Unknown	
3,746,409 (A>G)	<i>Rv3347c (PPE55)</i>	Unknown	
3,826,684 (C>T)	<i>Rv3408 (vapC47)</i>	Unknown	
3,959,418 (C>T)	<i>Rv3522 (ltp4)</i>	Unknown (lipid metabolism)	
4,302,036 (T>C)	<i>Rv3827c</i>	IS1537 transposition	
Other LIVs specifically found in <i>Mtb</i> causing TBM compared to PulTB			
71,336 (G>C)	<i>Rv0064A (vapB1)</i>	Unknown	Upstream gene variant
132,417 (C>G)	<i>Rv0107c (ctpI)</i>	Cation-transporting ATPase	Upstream gene variant
837,033 (A>G)	<i>Rv0728c-Rvnt09 (serA2-thrV)</i>	L-serine biosynthesis/ tRNA	Intergenic region
1,636,826 (C>A)	<i>Rv1439c-1562c (Rv1439c-treZ)</i>	Unknown/ trehalose biosynthesis	Intergenic region

Supplementary Table S1. Annotation of missense SNPs and lineage independent variants commonly found in TBM isolates of *M. tuberculosis* compared to PulTB. Note: LIV= lineage independent variant.

No.	Positions (SNPs)	Size (bp)	Affected genes	Functions
1	368,305 - 368,332	28	<i>Rv0304c (PPE5)</i>	Unknown
2	369,563 - 369,583	21	<i>Rv0304c (PPE5)</i>	Unknown
3	372,008 - 372,117	110	<i>Rv0304c (PPE5)</i>	Unknown
4	598,532 - 598,870	339	<i>Rv0507 (mmpL2)</i>	Unknown (fatty acid transport?)
5	848,706 - 849,164	459	<i>Rv0755c (PPE12)</i>	Unknown
6	1,021,831 - 1,021,873	43	<i>Rv0916c (PE7)- Rv0917 (betP)</i> intergenic region	Unknown/ glycine betaine uptake
7	1,340,761 - 1,340,842	83	<i>Rv1197 (esxK)</i>	Unknown
8	1,985,087 - 1,985,110	24	<i>Rv1754c</i>	Unknown
9	2,169,553 - 2,169,578	26	<i>Rv1918c (PPE35)</i>	Unknown
10	2,616,893 - 2,617,134	242	<i>Rv2339 (mmpL9)</i>	Unknown (fatty acid transport?)
11	3,047,645 - 3,047,709	65	<i>Rv2734 and Rv2735c</i>	Unknown/ Unknown
12	3,482,733 - 3,482,800	68	<i>Rv3115 – Rv3116 (moeB)</i> intergenic region	IS1081 insertion/ molybdopterin metabolism
13	3,527,891 - 3,528,048	158	<i>Rv3159c (PPE53)</i>	Unknown
14	3,798,712 - 3,798,742	31	<i>Rv3383c (idsB) – Rv3384c (vapC46)</i> intergenic region	Membrane lipids biosynthesis / Unknown

Supplementary Table S2. Characteristic of large indels identified from genomic positional comparison between *M. tuberculosis* isolates from TBM and PulTB cases.