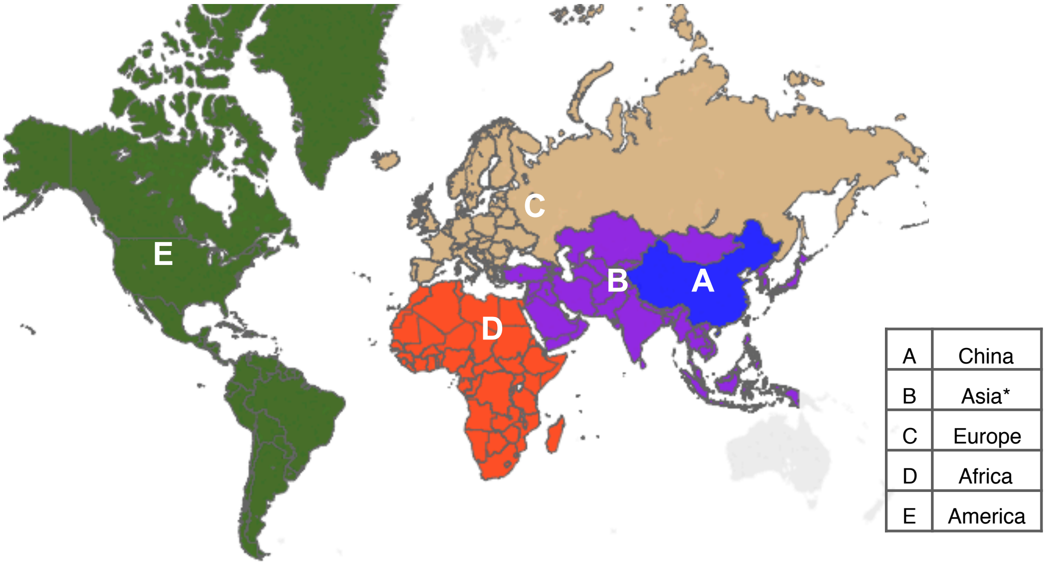
Supplementary Appendix

Supplementary Figures

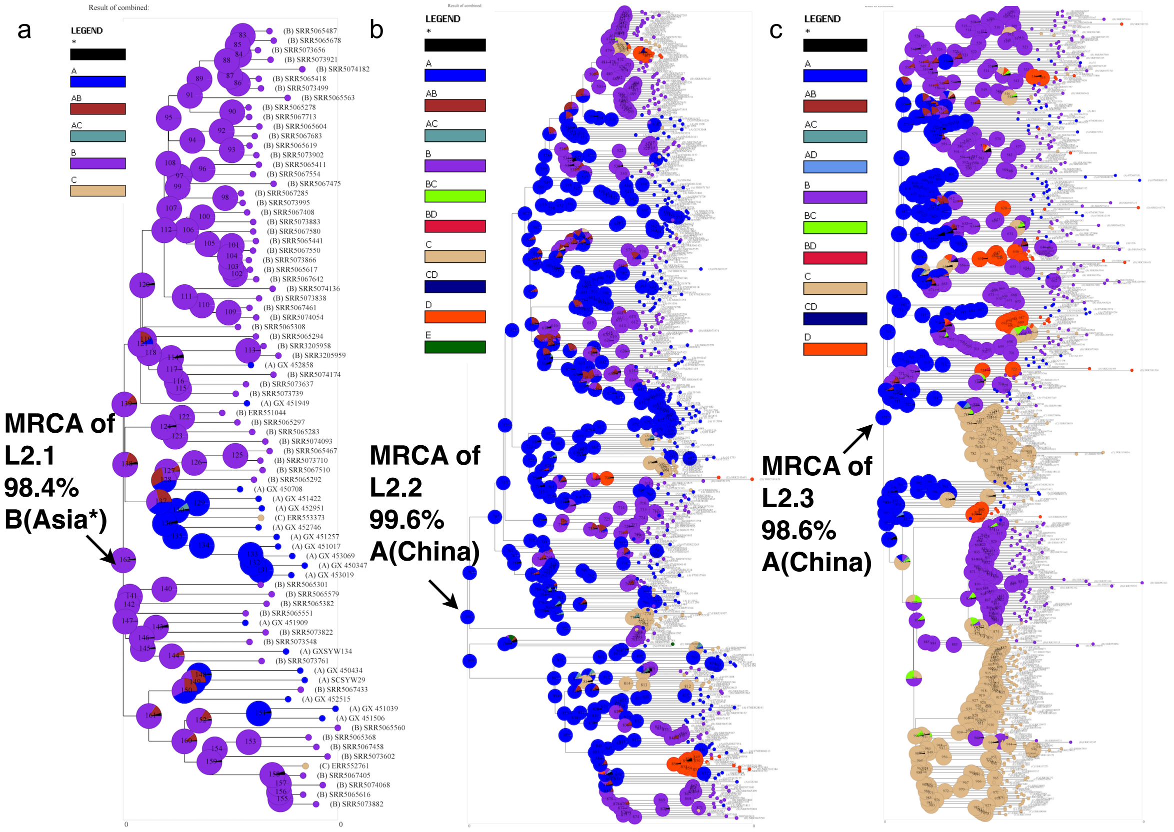
Supplementary Tables

Supplementary Text

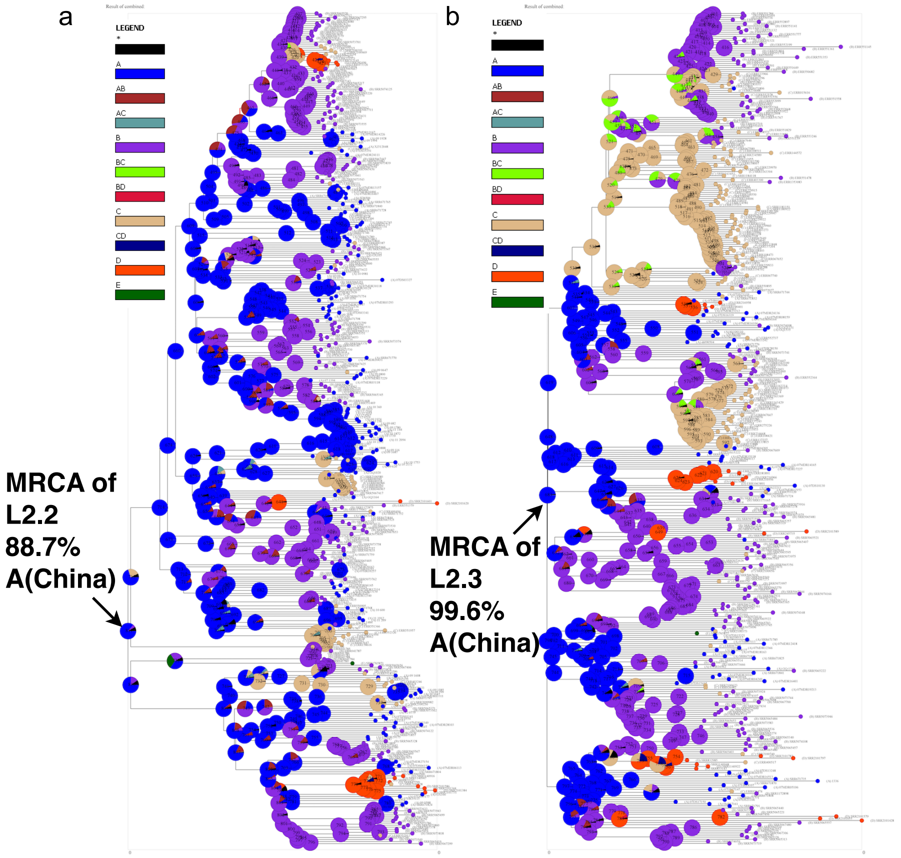


A	China
B	Asia*
C	Europe
D	Africa
E	America

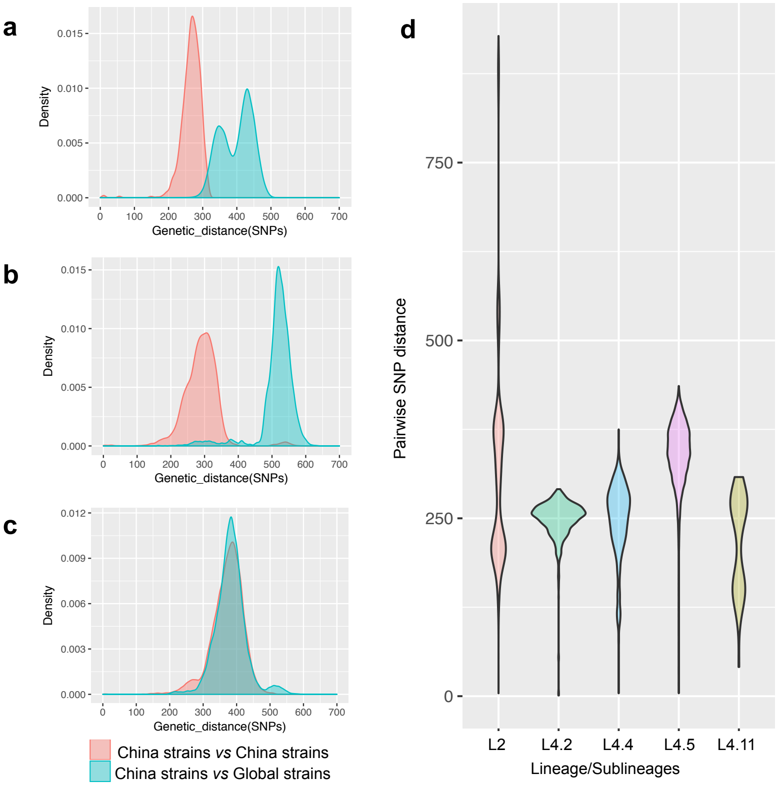
Supplementary Figure 1. The divisions of global geographic areas for the phylogeographic analysis. To estimate the origins of those MTBC sublineages detected in China, we divided the world map into five broad geographic areas and used them as a proxy for the most likely origin of each strain. We set China as an independent area from Asia because we want to test whether the indigenous genotypes were from single origins. Asia* refers to Asia region excluding China. We didn't subdivide America continent into different regions because only a few isolates with whole genome sequence were isolated there.



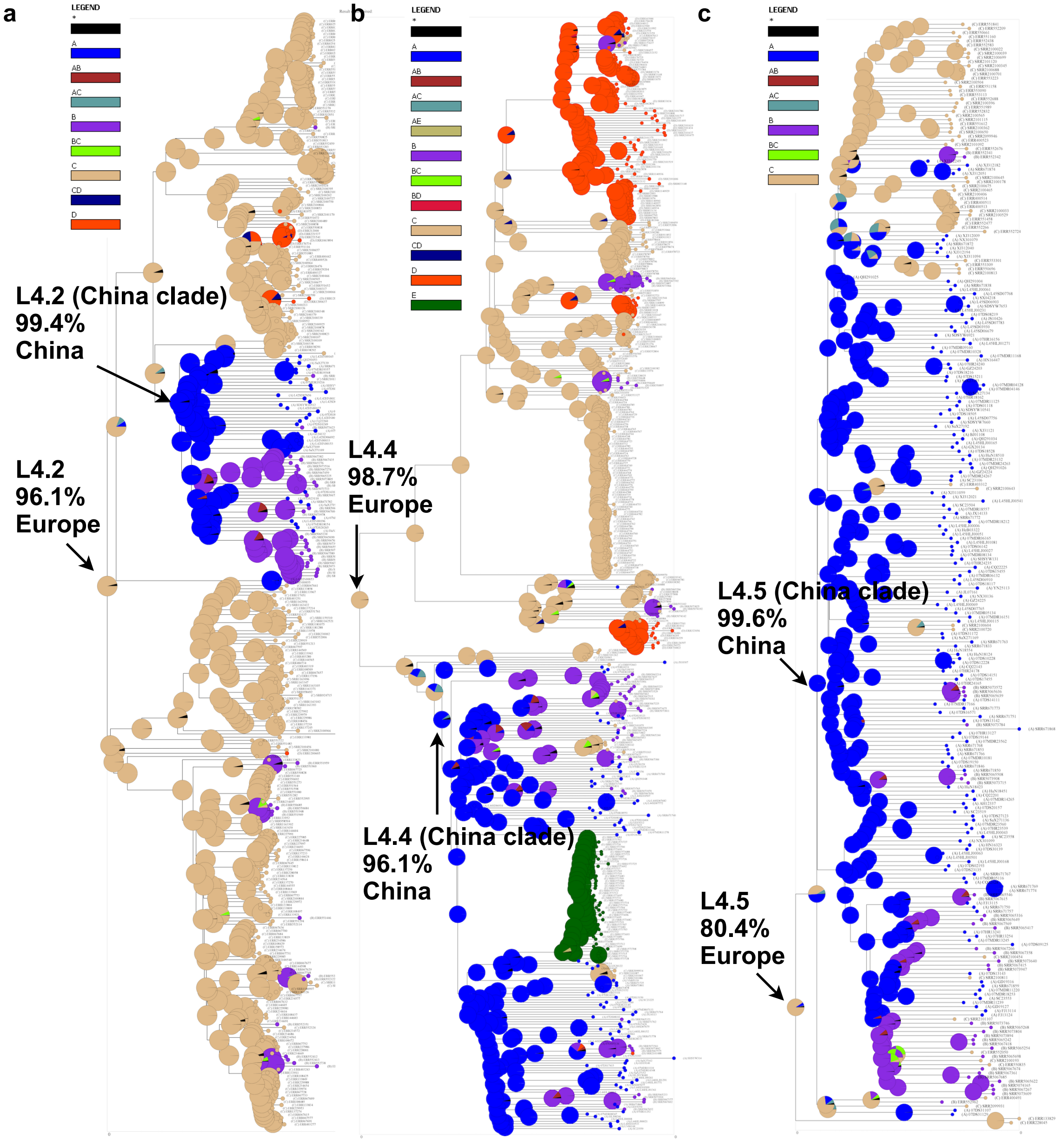
Supplementary Figure 2. Phylogeographic estimation for the geographic states of the MRCAs of L2.1, L2.2 and L2.3 sublineages. The results from Bayesian Binary MCMC (BBM) method showing the ancestor states of L2.1 (**a**), L2.2 **(b)** and L2.3 **(c)**. Pie-charts reflect probability of the respective area. Areas are colored as in the legends allowing for single or combined distributions. The probabilities for each sublineage's MRCA are indicated in the corresponding node. As RASP software only allow a maximum of 500 samples for analysis, we randomly sampled 500 isolates from 3,449 L2.3 isolates of global origin while all isolates of L2.1 or L2.2 were used for this analysis. The results from Statistical Dispersal-Vicariance analysis (S-DIVA) method showed very high consistency with BBM method and were not shown here.



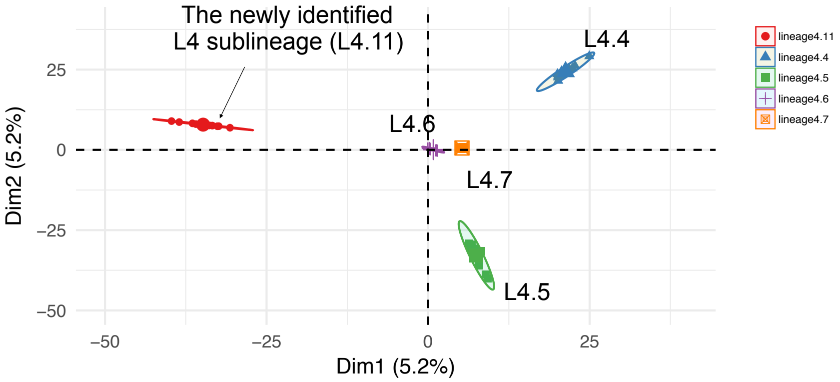
Supplementary Figure 3. Phylogeographic estimation for L2.2 and L2.3 after randomly reduced the isolate number from China. The results from Bayesian Binary MCMC (BBM) method showing the ancestor states L2.2 **(a)** and L2.3 **(b)**. For the analysis of L2.2 sublineage, the isolates numbers from China, Germany, Guatemala, India, Japan, Malawi, Mali, Portugal, Russia, South Africa, UK, Uzbekistan and Vietnam were 120, 11, 1, 8, 7, 4, 1, 2, 6, 10, 27, 3 and 161 respectively. For the analysis of L2.3 sublineage, the isolates numbers from Afgan, Belarus, China, Germany, Guatemala, India, Malawi, Portugal, Russia, South_Africa, Thailand, UK, Uzbekistan and Vietnam were 3, 22, 77, 6, 1, 3, 12, 1, 84, 10, 1, 12, 57 and 111 respectively.



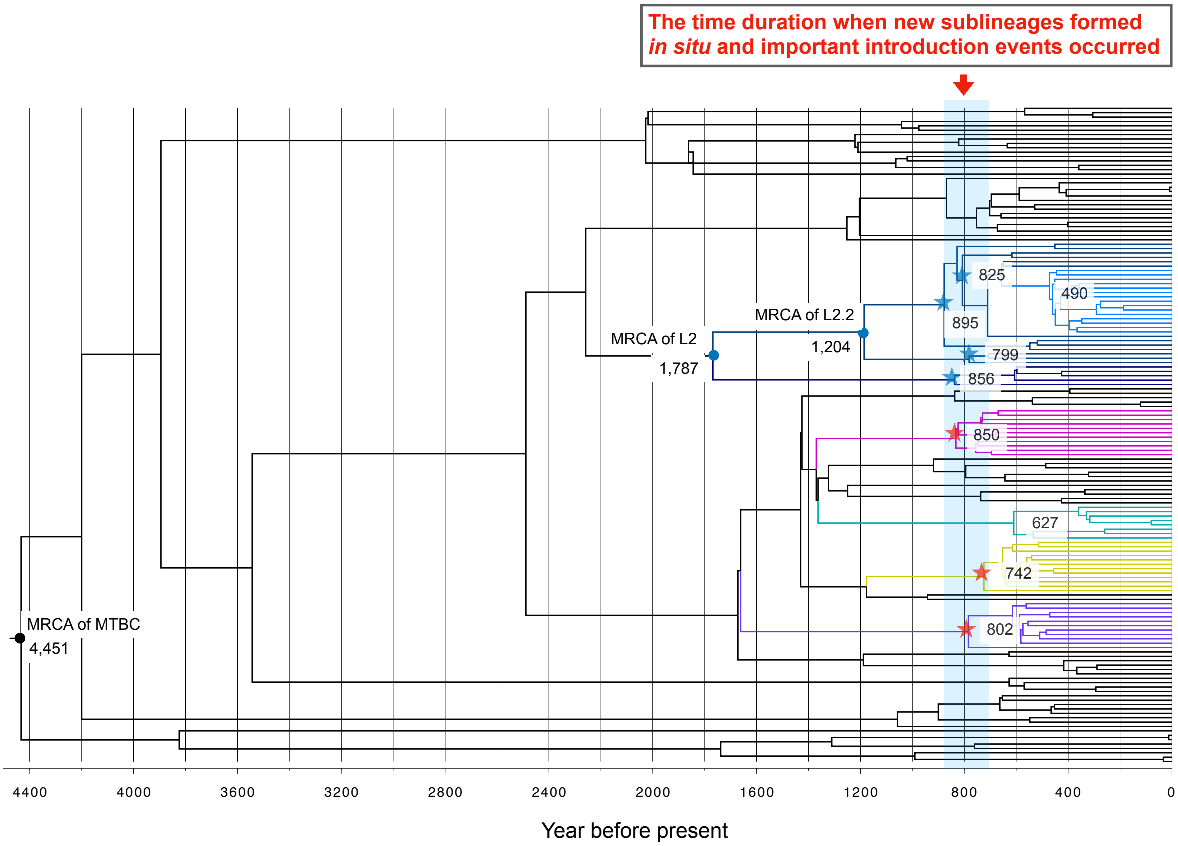
Supplementary Figure 4. Pairwise SNP distance distributions of indigenous genotypes. "China strains vs China strains" shows the pairwise SNP distance of MTBC strains from China of the corresponding sublineage. "China strains vs Globe strains" shows the pairwise SNP distance distribution of each China strain and each strain Global strain in the corresponding sublineage. These distributions indicate L4.2 (**a**) and L4.4 (**b**) strains from China formed deeply separated populations to those from global collections. However, L4.5 (**c**) strains in China could not be separated from global L4.5 strains, which was because the most of the genetic diversity in L4.5 were sampled in China (except two early diverged strains sampled from Russia) and most of globally sampled L4.5 strains were descendants from China L4.5 clade. **d**) The violin plot shows the pairwise SNP distance distributions between the major genotypes observed in China.



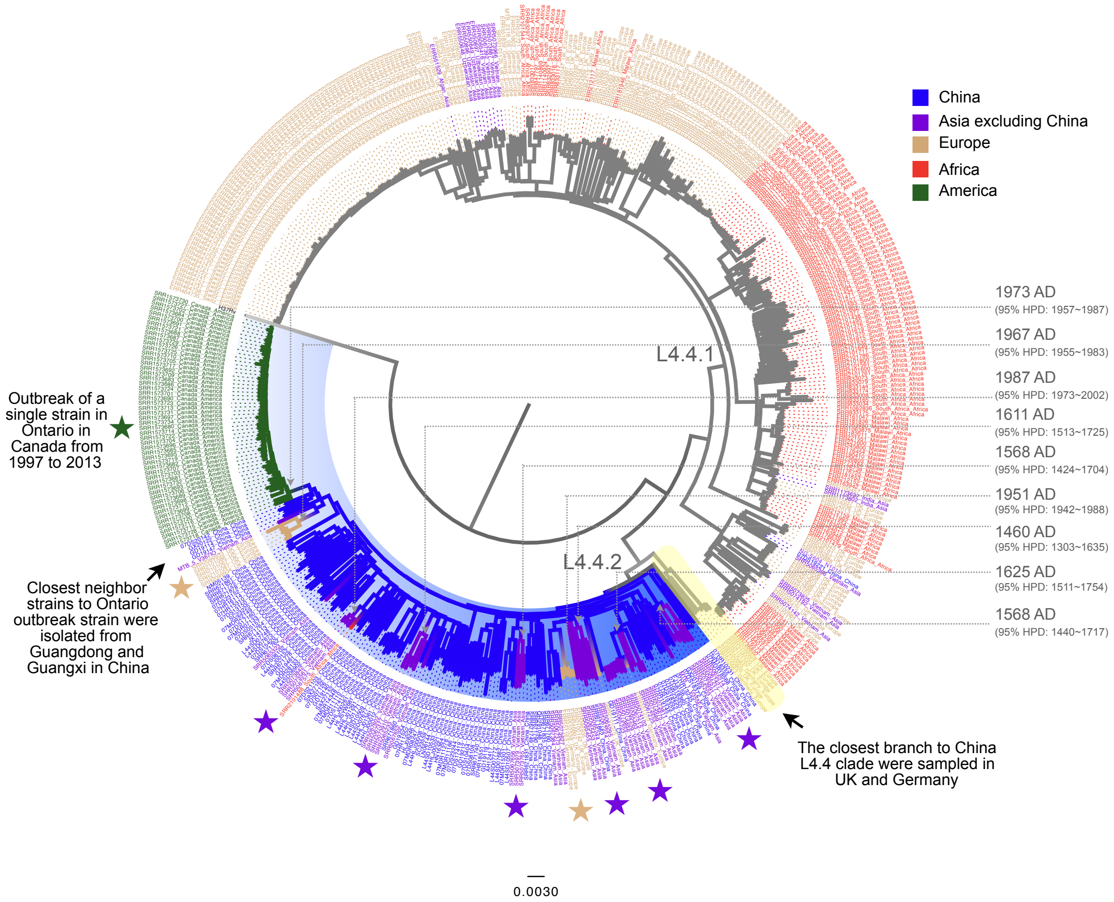
Supplementary Figure 5. Ancestral state estimation for L4.2, L4.4 and L4.5 sublineages. The geographic areas divisions refer Supplementary Figure 1. The probabilities for each sublineage's MRCA and the China specific clades are indicated in the corresponding node.



Supplementary Figure 6. Principal-component analysis of five lineage 4 sublineages. Principal-component analysis was conducted for the newly identified sublineage L4.11 and its four closest neighbor sublineages using the SNP matrices derived from whole-genome analysis. The first two principal-component axes (Dim1 and Dim2) are shown. This analysis demonstrated discriminative separation between L4.11 and the neighbor sublineages.



Supplementary Figure 7. Bayesian based dating results from MTBC-6 model. Numbers on the nodes are the diverged time (before present) corresponding to the median height of the node.

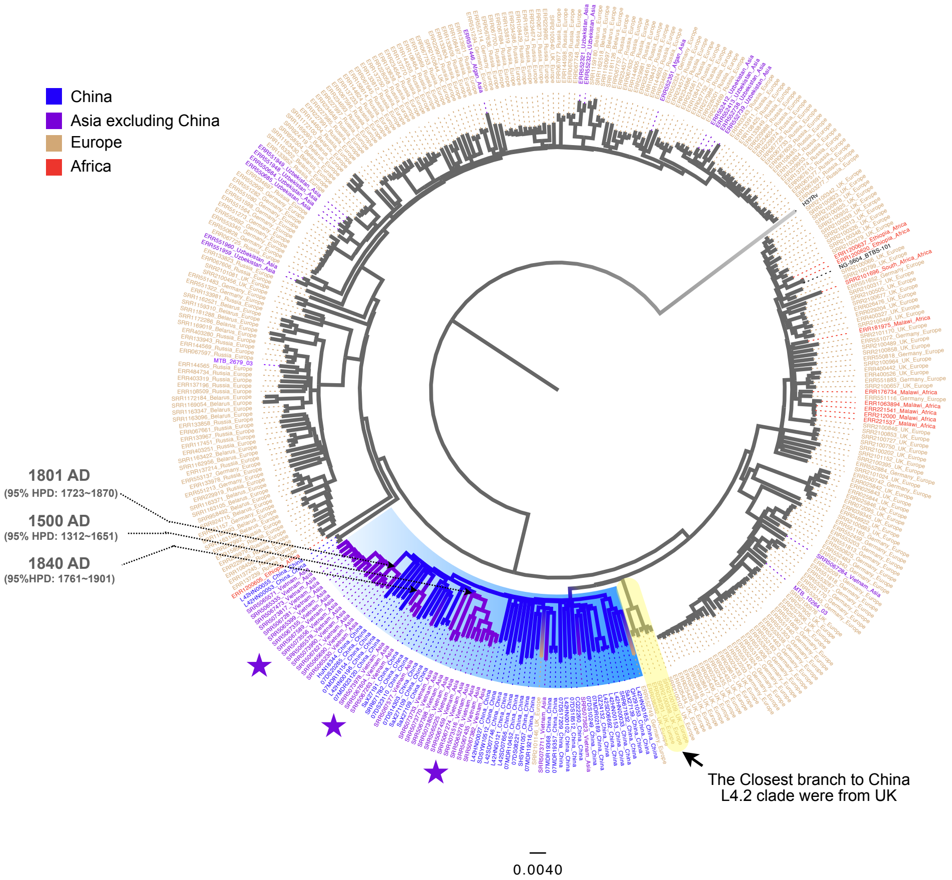


Outbreak of a single strain in Ontario in Canada from 1997 to 2013

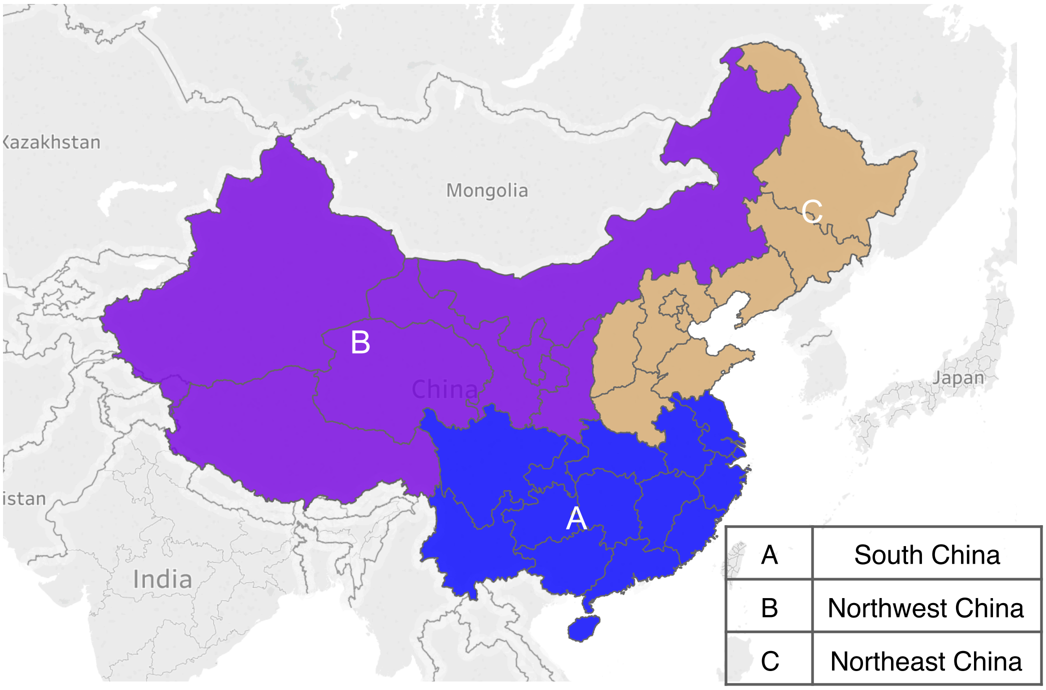
Closest neighbor strains to Ontario outbreak strain were isolated from Guangdong and Guangxi in China

The closest branch to China L4.4 clade were sampled in UK and Germany

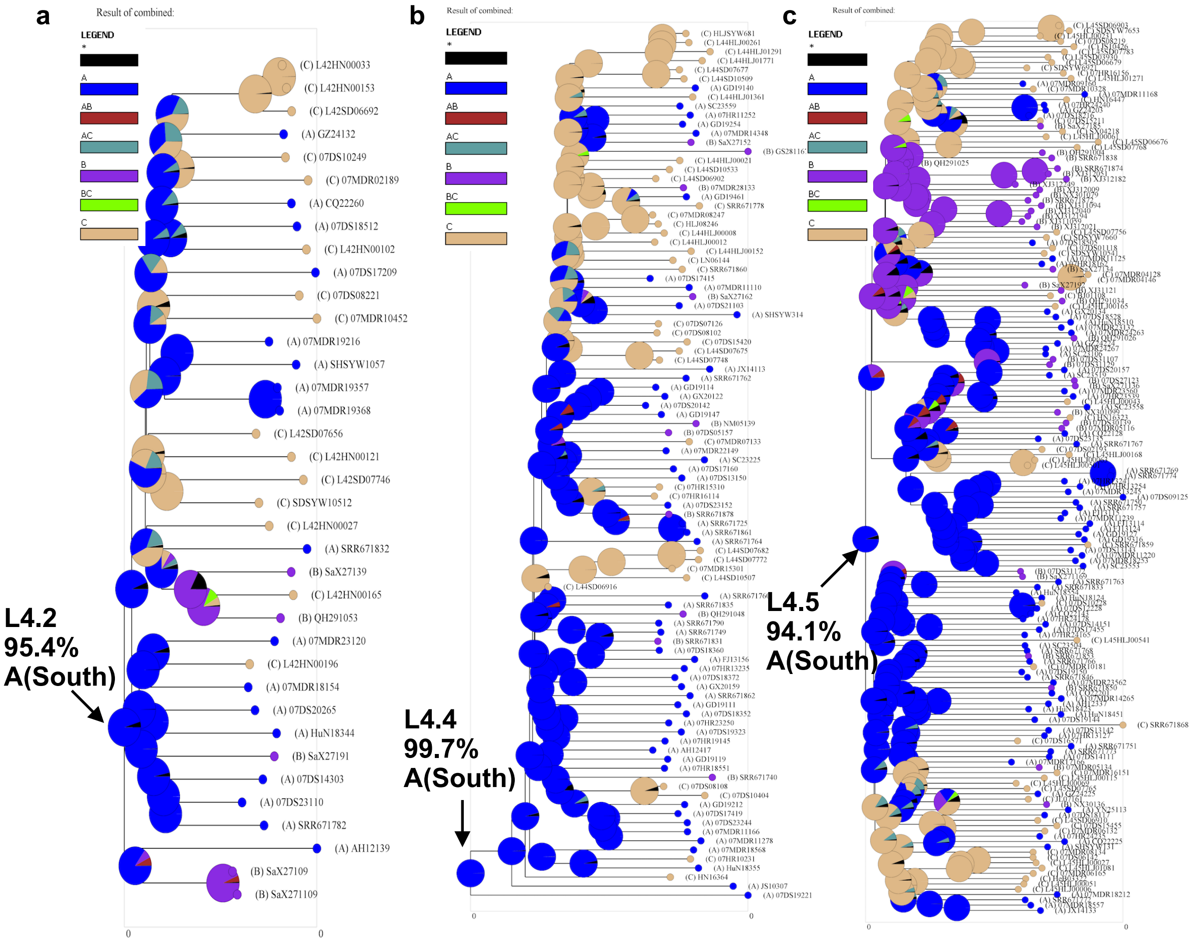
Supplementary Figure 8. Maximum likelihood phylogeny of sublineage L4.4. Different leaf colors refer to different geographic origin of those isolates. The China L4.4 clade was highlighted with blue background and the branches were colored according to the geographic their geographic attributions. The major country transition events were marked with a star symbol, and transition time of each event was estimated with BEAST under MTBC-6 model. The closest brach that diverged before China clade was sampled from UK and Germany. The outbreak strain in Ontario turned out to be a descendant of China L4.4 clade, and the most recent common ancestor of that outbreak strain was estimated to emerge around 1973 AD. The closest neighbor strains to that clade were sampled from Guangdong and Guangxi provinces(South China), suggesting recent human migration brought that strain to Ontario.



Supplementary Figure 9. Maximum likelihood phylogeny of sublineage L4.2. Different leaf colors refer to different present geographic distribution of those isolates. The China L4.2 clade was highlighted with blue background and the branches were colored according to the geographic their geographic attributions. The major country transition events were marked with a star symbol, and transition time of each event was estimated with BEAST under MTBC-6 model. The closest brach that diverged before China clade was sampled from UK.



Supplementary Figure 10. China geographic areas divisions for the phylogeographic analysis. For estimation of L4 strains' origin in China, we divided the map into three broad areas refer different origin hypothesis. "A" represent south region and include provinces of Anhui, Chongqing, Fujian, Guangdong, Guangxi, Guizhou, Hainan, Hubei, Hunan, Jiangxi, Jiangsu, Shanghai, Sichuan, Yunan, Zhejiang. "B" represent northwest region and includes provinces of Gansu, Inner Mongolia, Ningxia, Qinghai, Shaanxi, Tibet and Xinjiang. "C" represents northeast region including Beijing, Hebei, Heilongjiang, Henan, Jilin, Liaoning, Shandong, Shanxi and Tianjin.



Supplementary Figure 11. Estimation of the geographic origins of L4 sublineages in China. (a), (b) and (c) show the phylogeographic reconstructions of ancestral states of China's L4.2, L4.4 and L4.5 strains respectively.

Supplementary Table 1. Collected spoligotyping records of 16,621 MTBC isolates from China.

Provinces	Reference	Total	L1	L3	L2	L4	Beijing	T1	T2	T3	T4	T5	H1	H3	H4	U	S	X1	H37Rv	LAM1	LAM3	LAM	LAM7	LAM8	LAM9	LAM10	CAS	CAS-Delhi	MANU2	BCG	New
Anhui	1	236			152	17	152	9	4	2						2													19		48
Beijing	2	1585		1	1300	122	1300	69	24	10	2			3	1	12	1											1	76	2	84
Beijing	3	260			211	37	211	27	7	1						2															12
Beijing	1	57			47	8	47	6		2																					2
Chongqing	4	208			156	25	156	17	4					2			1								1				7		20
Chongqing	5	297			198	77	198	53	13	1	1			4		3									1	1			2		20
Chongqing	1	127		2	67	37	67	21	9	5				1		1											2				21
Fujian	1	100		4	30	27	30	8	6					4		9											4		3		36
Guangdong	1	235			75	63	75	26	16	1	1			9		10													9		88
Gansu	6	467		1	409	32	409	20	3	2					5	2												1	3		22
Gansu	1	55			47	6	47	1	1						3	1															2
Guangxi	1	108		1	34	24	34	20	3							1											1		12		37
Guizhou	1	149		1	89	39	89	20	5	7		1	3			3											1		5		15
Hainan	1	13			7	1	7	1																					2		3
Hebei	7	1017			920	73	920	39	15	5	1	1		7		1				1					3				7		17
Hebei	1	188			155	9	155	4	3	2																			7	1	16
Heilongjiang	8	200			189	11	189																								
Heilongjiang	1	118		2	96	9	96	3	3			2				1											2		2		9
Henan	9	443			387	52	387	50									1					1							4		0
Henan	1	263		1	213	26	213	10	8	1	1	1				1	1												12		11
Hubei	10	105			86	19	86																								
Hubei	1	236		1	168	27	168	16	4	3		1				3											1		26		14
Hunan	11	171		1	126	38	126	34								4												1	1		5
Hunan	1	319		7	155	104	155	73	15	8	1	1				2											7		9		44
Jiangsu	12	287			233	38	233	33								1	4												3		13
Jiangsu	13	497			403	64	403	42	15	1		3					2									1			20		10

Jiangsu	1	197		3	134	32	134	20	9	1		1															3		10		18
Jiangsu, Zhejiang	14	209	1		163	45	163	32								5						8									
Jilin	1	64		1	49	8	49	2	2	1						3											1				6
Jiangxi	1	172		2	111	23	111	18	2	1				1		1											2		6		30
Liaoning	1	63		1	48	4	48	1	2							1													7		3
Inner Mongolia	15	372			318	23	318	13	4			1		3							1								1		30
Inner Mongolia	1	35			31	2	31	1								1															2
Ningxia	1	63		1	52	7	52	4	1	1						1											1		0		3
Qinghai	1	41		2	34	4	34	1	3																			2	1		
Shaanxi	16	89			79	10	79	4						1			1	1	1						1	1					
Shaanxi	1	53			36	9	36	4	3	2																			1		7
Shandong Jiangsu	17	1222	8	8	967	178	967																								61
Shandong	1	235		1	184	31	184	9	11	1						10											1		2		17
Shanxi	1	117			82	4	82	3	1																				17		14
Shanghai	1	62		1	50	4	50	3			1																1		4		3
Sichuan	18	265			143	96	143	80						9		7															26
Sichuan	19	213			150	42	150	39						3															9		12
Sichuan	20	306			212	85	212	57						18																	9
Sichuan	1	305		4	138	87	138	70	13					2		2											4		5		71
Taiwan	21	421	120		176	104	176																								21
Taiwan	22	334	91	0	116	110	116																								17
Tianjin	1	66		0	57	3	57	3																					5		1
Tibet	23	576		18	522	27	522	25		1		1															18				9
Tibet	1	17		1	12	2	12	2																					1		1
Xinjiang	1	133		35	56	23	56	12	1					1	8														17		2
Yunnan	1	55		3	27	10	27	6		1	1	1		1															6		9
Zhejiang	24	274			251	20	251	12	4							4													3		
Zhejiang	1	135		2	64	23	64	15	4	1				1		2													9		37
13 Provinces	25	2346	7	12	1738	494	1738	219	74	23	12	3	11	74	10	38	8	2	2		1		2	2	10	3	7	5	38		57
Whole country	26	440	0	1	349	75	349																								15
Total		16621	227	118	12302	2570	12302	1257	292	84	21	17	14	144	27	134	19	3	3	1	2	9	2	2	16	6	56	10	371	3	1030

Reference:

1. Pang, Y. et al. Spoligotyping and drug resistance analysis of Mycobacterium tuberculosis strains from national survey in China. PLoS One 7, e32976 (2012).
2. Liu, Y. et al. Genotypic diversity analysis of Mycobacterium tuberculosis strains collected from Beijing in 2009, using spoligotyping and VNTR typing. PLoS One 9, e106787 (2014).
3. Lu, B. et al. Genetic diversity of Mycobacterium tuberculosis isolates from Beijing, China assessed by Spoligotyping, LSPs and VNTR profiles. BMC Infect Dis 12, 372 (2012).
4. Zhang, D., An, J., Wang, Y. & Pang, Y. Genetic diversity of multidrug-resistant tuberculosis in a resource-limited region of China. Int J Infect Dis 29, 7-11 (2014).
5. Zhang, D. et al. Molecular typing and drug susceptibility of Mycobacterium tuberculosis isolates from Chongqing Municipality, China. Infect Genet Evol 13, 310-6 (2013).
6. Liu, J. et al. First insight into the genotypic diversity of clinical Mycobacterium tuberculosis isolates from Gansu Province, China. PLoS One 9, e99357 (2014).
7. Li, Y. et al. Characterization of Mycobacterium tuberculosis isolates from Hebei, China: genotypes and drug susceptibility phenotypes. BMC Infect Dis 16, 107 (2016).
8. Wang, J. et al. Genotypes and characteristics of clustering and drug susceptibility of Mycobacterium tuberculosis isolates collected in Heilongjiang Province, China. J Clin Microbiol 49, 1354-62 (2011).
9. Zhao, Y. et al. Molecular typing of Mycobacterium tuberculosis isolates circulating in Henan, central China. Exp Ther Med 4, 949-953 (2012).
10. Han, H. et al. Utility of mycobacterial interspersed repetitive unit typing for differentiating Mycobacterium tuberculosis isolates in Wuhan, China. J Med Microbiol 56, 1219-23 (2007).
11. Zhao, L.L. et al. Prevalence and molecular characteristics of drug-resistant Mycobacterium tuberculosis in Hunan, China. Antimicrob Agents Chemother 58, 3475-80 (2014).
12. Pang, H. et al. Molecular characterization and drug-resistance of Mycobacterium tuberculosis strains in Xuzhou, China. Biomed Environ Sci 27, 960-4 (2014).
13. Lu, W. et al. Genotypes of Mycobacterium tuberculosis isolates in rural China: using MIRU-VNTR and spoligotyping methods. Scand J Infect Dis 46, 98-106 (2014).
14. Hu, Y. et al. Transmission pattern of drug-resistant tuberculosis and its implication for tuberculosis control in eastern rural China. PLoS One 6, e19548 (2011).
15. Yu, Q. et al. Genetic diversity of Mycobacterium tuberculosis isolates from Inner Mongolia, China. PLoS One 8, e57660 (2013).
16. Zhou, A. et al. Molecular characterization of isoniazid-resistant Mycobacterium tuberculosis isolates from Xi'an, China. Microb Drug Resist 17, 275-81 (2011).
17. Hu, Y. et al. Comparison of the socio-demographic and clinical features of pulmonary TB patients infected with sub-lineages within the W-Beijing and non-Beijing Mycobacterium tuberculosis. Tuberculosis (Edinb) 97, 18-25 (2016).
18. Liu, H.C. et al. Molecular Typing Characteristic and Drug Susceptibility Analysis of Mycobacterium tuberculosis Isolates from Zigong, China. Biomed Res Int 2016, 6790985 (2016).
19. Tang, K. et al. Characterization of rifampin-resistant isolates of Mycobacterium tuberculosis from Sichuan in China. Tuberculosis (Edinb) 93, 89-95 (2013).
20. Zhao, Y. et al. The population structure of drug-resistant Mycobacterium tuberculosis clinical isolates from Sichuan in China. Infect Genet Evol 12, 718-24 (2012).
21. Chen, Y.Y. et al. Molecular epidemiology of tuberculosis in Kaohsiung City located at southern Taiwan, 2000-2008. PLoS One 10, e0117061 (2015).
22. Chen, Y.Y. et al. Distinct modes of transmission of tuberculosis in aboriginal and non-aboriginal populations in Taiwan. PLoS One 9, e112633 (2014).
23. Dong, H. et al. Genetic diversity of Mycobacterium tuberculosis isolates from Tibetans in Tibet, China. PLoS One 7, e33904 (2012).
24. Xia, Q. et al. Phenotypic and genotypic characterization of pyrazinamide resistance among multidrug-resistant Mycobacterium tuberculosis isolates in Zhejiang, China. Antimicrob Agents Chemother 59, 1690-5 (2015).
25. Dong, H. et al. Spoligotypes of Mycobacterium tuberculosis from different Provinces of China. J Clin Microbiol 48, 4102-6 (2010).
26. Jiao, W. et al. A country-wide study of spoligotype and drug resistance characteristics of Mycobacterium tuberculosis isolates from children in China. PLoS One 8, e84315 (2013).

Supplementary Table 2-1. SNP typing results of 3,227 MTBC isolates from 70 random sites.

Province	Site No.	No. of Isolates	L2.1	L2.2	L2.3	L4.1	L4.2	L4.3	L4.4.1	L4.4.2	L4.5	L4.8	L4.11	L1	L3
Anhui	4	174		27	135		4			3	5				
Beijing	1	52		10	34		2			0	6				
Chongqing	2	84		13	44		9			6	11		1		
Fujian	2	77		13	41	1	2			7	13				
Gansu	1	82		27	51		0			1	2				1
Guangdong	4	175		26	97		10		1	26	14			1	
Guangxi	2	62	5	21	18		5			8	5				
Guizhou	2	85		30	36		2			5	12				
Hainan	1	13		6	5		0			1	1				
Hebei	3	143		21	110		0			3	9				
Heilongjiang	2	86		12	65		1			5	3				
Henan	5	263		57	176		6			11	12				1
Hubei	4	158		29	111		2			5	11				
Hunan	5	208		50	86		14			28	30				
Inner Mongolia	1	64		9	51		0			2	2				
Jiangsu	4	201		35	129		9			15	13				
Jiangxi	3	142		13	106		7			6	10				
Jilin	1	53		5	41		0			5	2				
Liaoning	1	59		9	47		0			1	2				
Ningxia	1	90		48	31		0			2	8	1			
Qinghai	1	63		29	24		1			2	6				1
Shaanxi	1	97		34	39		0			9	13				2
Shandong	4	169		19	133		3			7	7				
Shanghai	1	57		14	36		0			3	4				
Shanxi	2	89		10	68		3			0	8				
Sichuan	5	168	2	40	59		12			14	41				
Tianjin	1	64		2	59		1			1	1				
Tibet	1	9		6	2		0			1	0				
Xinjiang	2	115		33	37		0	2		0	22				21
Yunnan	1	36		13	11		2			0	10				
Zhejiang	2	89	1	26	49		4			4	5				
Whole country	70	3227	8	687	1931	1	99	2	1	181	288	1	1	1	26

Supplementary Table 2-2. SNP typing results of 1,351 MTBC isolates from 6 county sites (population sites).

Province	County Name	Longitude	Latitude	No. of Isolates	L2.1	L2.2	L2.3	L4.5	L4.2	L4.4	L4.8	L4.1	L4.11	L1
Guangxi	Baise	23.00	108.00	174	3	49	55	16	16	34		1		
Heilongjiang	Wuchang	44.00	126.01	183		12	146	15		10				
Henan	Weishi	34.41	114.17	194		44	130	8	9	3				
Sichuan	Wusheng	30.00	106.00	214	2	47	65	44	17	33	1		5	
Shandong	Feixian	35.26	117.97	201		25	134	19	4	17		1		1
Shanghai	Songshi	31.00	121.00	385		80	233	24	22	24			1	1
All				1351	5	257	763	126	68	121	1	2	6	2

The fates of 32 isolates that showed ambiguous typing results:

1) Six isolates belonged to a newly described sublineage. There were typed as L4 strains, but could not be assigned to a previously defined sublineage. As described in the Results, we found that these isolates formed a phylogenetic clade deeply separated from its sister clades/sublineages. We defined this clade as a new L4 sublineage (L4.11) based on principle component analysis.

2) Eleven isolates belonged to rare sublineages that our Real-time PCR probe detecting method was not designed to detect. To reduce cost, we designed probes for the common sublineages identified in our pilot typing results (L2.3, L2.2, L4.5, L4.4, and L4.2). Isolates that did not fit into these sublineages were sent for whole genome sequencing. This is described in the Methods section (line 448-453). Among the eleven isolates in this group, two were identified as L2.1, three as L4.1, one as L4.3, two as L4.8 and three as L1.

3) Three isolates were non-tuberculosis mycobacteria (NTM). These isolates failed in one probe detecting assay but were successful in the other probes' detections. The whole-genome sequencing results showed that these isolates were NTM such as Mycobacterium kansassii (2 isolates), Mycobacterium intracellulare (1 isolate).

4) Twelve isolates had poor DNA quality. These isolates failed Real-time PCR probe detection assays and also failed in WGS library construction ten isolates) or only generated very small numbers of sequencing reads (two isolates).

Samples with signals of mixed infection were not sent for whole-genome sequencing as this situation can easily be identified by checking double peaks from a single probe. In these situations, we recorded the typing result from the major melting peak.

Supplementary Table 3. Geographic and sequencing information of 306 whole-genome sequenced isolates in this study.

Strain ID	Lineage	Mapped reads number	Genome Coverage	Average Depth	Province
GXSYW247	LINEAGE4	5113618	0.9968	103.58	Guangxi
07HR13233	LINEAGE4	6860847	0.9967	148.53	Fujian
SDSYW7641	LINEAGE4	5039782	0.9949	102.91	Shandong
L42HN00053	LINEAGE4	3462503	0.9991	113.52	Henan
L42HN00055	LINEAGE4	7657175	0.9989	253.73	Henan
QH29124	LINEAGE4	3398448	0.9982	98.11	Qinghai
07DS08221	LINEAGE4	3630439	0.9963	52.28	Heilongjiang
07DS10249	LINEAGE4	7934890	0.9948	172.21	Jiangsu
07DS14303	LINEAGE4	4193573	0.9876	59.27	Jiangxi
07DS17209	LINEAGE4	6586784	0.9888	93.66	Hubei
07DS18512	LINEAGE4	5056754	0.9952	72.43	Hunan
07DS20265	LINEAGE4	9222464	0.9967	191.93	Guangxi
07DS23110	LINEAGE4	7021082	0.9974	151.67	Sichuan
07MDR02189	LINEAGE4	2771190	0.9805	43.65	Tianjin
07MDR10452	LINEAGE4	2011774	0.9928	30.17	Jiangsu
07MDR18154	LINEAGE4	2884045	0.9963	40.30	Hunan
07MDR19216	LINEAGE4	2485760	0.9935	33.34	Guangdong
07MDR19357	LINEAGE4	3717574	0.9938	51.99	Guangdong
07MDR19368	LINEAGE4	4924194	0.9958	68.83	Guangdong
07MDR23120	LINEAGE4	6710399	0.9930	94.59	Sichuan
AH12139	LINEAGE4	4893382	0.9979	138.90	Anhui
CQ22260	LINEAGE4	4468519	0.9928	142.12	Chongqing
GZ24132	LINEAGE4	5158399	0.9964	162.73	Guizhou
HuN18344	LINEAGE4	4784054	0.9947	145.92	Hunan
L42HN00027	LINEAGE4	10645299	0.9986	352.80	Henan
L42HN00033	LINEAGE4	7280770	0.9990	239.29	Henan
L42HN00102	LINEAGE4	1565626	0.9985	50.85	Henan
L42HN00121	LINEAGE4	7551346	0.9987	248.74	Henan
L42HN00153	LINEAGE4	8292383	0.9988	272.26	Henan
L42HN00165	LINEAGE4	2899035	0.9988	94.01	Henan
L42HN00196	LINEAGE4	7201007	0.9990	235.73	Henan
L42SD06692	LINEAGE4	5698344	0.9987	186.51	Shandong
L42SD07656	LINEAGE4	5364454	0.9987	175.62	Shandong
L42SD07746	LINEAGE4	4203978	0.9989	137.90	Shandong
QH291053	LINEAGE4	2620110	0.9957	84.58	Qinghai
SaX27109	LINEAGE4	4125985	0.9977	132.55	Shaanxi
SaX271109	LINEAGE4	3241980	0.9982	103.69	Shaanxi
SaX27139	LINEAGE4	3837212	0.9974	123.17	Shaanxi
SaX27191	LINEAGE4	3696062	0.9965	116.72	Shaanxi
SDSYW10512	LINEAGE4	5217800	0.9960	105.72	Shandong
SHSYW1057	LINEAGE4	3889891	0.9950	79.08	Shanghai
XJ312027	LINEAGE4	4958130	0.9970	161.47	Xinjiang
07DS19221	LINEAGE4	6762318	0.9960	99.67	Guangdong
07DS05157	LINEAGE4	6392532	0.9897	92.45	Inner Mongolia
07DS07126	LINEAGE4	6507780	0.9976	140.51	Jilin
07DS08102	LINEAGE4	8153051	0.9970	175.07	Heilongjiang
07DS08108	LINEAGE4	4449229	0.9972	63.10	Heilongjiang
07DS10404	LINEAGE4	5226047	0.9786	75.50	Jiangsu
07DS13150	LINEAGE4	5899445	0.9954	123.08	Fujian
07DS15420	LINEAGE4	4438677	0.9879	62.86	Shandong
07DS17160	LINEAGE4	5631448	0.9954	79.80	Hubei
07DS17415	LINEAGE4	2972003	0.9946	46.12	Hubei
07DS17419	LINEAGE4	7890291	0.9992	169.81	Hubei

07DS18352	LINEAGE4	9727693	0.9968	210.22	Hunan
07DS18360	LINEAGE4	5233570	0.9939	74.36	Hunan
07DS18372	LINEAGE4	4937205	0.9935	72.76	Hunan
07DS19323	LINEAGE4	7300849	0.9871	116.33	Guangdong
07DS20142	LINEAGE4	6294413	0.9937	90.99	Guangxi
07DS21103	LINEAGE4	2261216	0.9911	33.63	Hainan
07DS23152	LINEAGE4	7864495	0.9989	168.51	Sichuan
07DS23244	LINEAGE4	7744523	0.9980	161.12	Sichuan
07HR10231	LINEAGE4	8732458	0.9963	188.95	Jiangsu
07HR11252	LINEAGE4	9066695	0.9993	195.53	Zhejiang
07HR13235	LINEAGE4	6884319	0.9952	149.36	Fujian
07HR15310	LINEAGE4	6541338	0.9969	139.74	Shandong
07HR16114	LINEAGE4	9884345	0.9993	213.47	Henan
07HR18551	LINEAGE4	6367008	0.9959	138.06	Hunan
07HR19145	LINEAGE4	6761980	0.9985	144.70	Guangdong
07HR23250	LINEAGE4	10632890	0.9997	227.53	Sichuan
07MDR07133	LINEAGE4	6627455	0.9986	103.77	Jilin
07MDR08247	LINEAGE4	3823904	0.9991	57.32	Heilongjiang
07MDR11110	LINEAGE4	3508146	0.9908	52.87	Zhejiang
07MDR11166	LINEAGE4	3471420	0.9878	51.75	Zhejiang
07MDR11278	LINEAGE4	2989808	0.9908	45.27	Zhejiang
07MDR14348	LINEAGE4	5313152	0.9876	79.73	Jiangxi
07MDR15301	LINEAGE4	4089826	0.9946	60.25	Shandong
07MDR18568	LINEAGE4	2515974	0.9927	34.66	Hunan
07MDR22149	LINEAGE4	4454920	0.9930	64.83	Chongqing
07MDR28133	LINEAGE4	6262862	0.9971	89.66	Gansu
AH12417	LINEAGE4	5045332	0.9967	157.97	Anhui
FJ13156	LINEAGE4	5116127	0.9975	149.34	Fujian
GD19111	LINEAGE4	4573882	0.9961	143.83	Guangdong
GD19114	LINEAGE4	4327072	0.9968	135.43	Guangdong
GD19119	LINEAGE4	2327561	0.9972	72.04	Guangdong
GD19140	LINEAGE4	4689544	0.9957	149.01	Guangdong
GD19147	LINEAGE4	4873097	0.9956	152.98	Guangdong
GD19212	LINEAGE4	7421544	0.9973	236.08	Guangdong
GD19254	LINEAGE4	3974276	0.9959	126.24	Guangdong
GD19461	LINEAGE4	3638998	0.9965	111.95	Guangdong
GS281167	LINEAGE4	460961	0.9975	14.52	Gansu
GX20122	LINEAGE4	5507056	0.9961	174.41	Guangxi
GX20159	LINEAGE4	4197019	0.9985	126.82	Guangxi
HLJ08246	LINEAGE4	5626712	0.9967	173.91	Heilongjiang
HLJSYW681	LINEAGE4	5722652	0.9978	117.03	Heilongjiang
HN16364	LINEAGE4	3420189	0.9987	102.12	Henan
HuN18355	LINEAGE4	4639210	0.9958	144.80	Hunan
JS10307	LINEAGE4	3556302	0.9989	97.47	Jiangsu
JX14113	LINEAGE4	3723514	0.9975	111.07	Jiangxi
L44HLJ00008	LINEAGE4	4946363	0.9997	161.59	Heilongjiang
L44HLJ00012	LINEAGE4	6925093	0.9997	228.21	Heilongjiang
L44HLJ00021	LINEAGE4	4647278	0.9999	152.50	Heilongjiang
L44HLJ00152	LINEAGE4	4429540	0.9990	144.01	Heilongjiang
L44HLJ00261	LINEAGE4	4053732	0.9995	132.98	Heilongjiang
L44HLJ01291	LINEAGE4	6341900	0.9996	206.28	Heilongjiang
L44HLJ01361	LINEAGE4	6599592	0.9992	215.58	Heilongjiang
L44HLJ01771	LINEAGE4	2121794	0.9997	68.15	Heilongjiang
L44SD06902	LINEAGE4	6404323	0.9997	211.25	Shandong
L44SD07675	LINEAGE4	5912449	0.9998	193.78	Shandong
L44SD07677	LINEAGE4	5456985	0.9988	178.77	Shandong

L44SD07682	LINEAGE4	6024187	0.9997	198.30	Shandong
L44SD07748	LINEAGE4	6776230	0.9997	222.99	Shandong
L44SD07772	LINEAGE4	5420804	0.9997	178.28	Shandong
L44SD10507	LINEAGE4	11986079	0.9999	393.93	Shandong
L44SD10509	LINEAGE4	11856175	0.9986	392.03	Shandong
L44SD10533	LINEAGE4	10111261	0.9998	333.92	Shandong
LN06144	LINEAGE4	5193348	0.9969	158.48	Liaoning
NM05139	LINEAGE4	4270008	0.9980	124.65	Inner Mongolia
QH291048	LINEAGE4	2958935	0.9992	95.96	Qinghai
SaX27152	LINEAGE4	3876884	0.9964	117.25	Shaanxi
SaX27162	LINEAGE4	3152959	0.9971	100.68	Shaanxi
SC23225	LINEAGE4	2257746	0.9979	67.42	Sichuan
SC23559	LINEAGE4	5650036	0.9979	175.73	Sichuan
SHSYW314	LINEAGE4	5825326	0.9971	119.33	Shanghai
07DS01118	LINEAGE4	2468965	0.9938	36.28	Beijing
07DS02193	LINEAGE4	5707208	0.9940	82.07	Tianjin
07DS06142	LINEAGE4	5093503	0.9897	73.99	Liaoning
07DS08219	LINEAGE4	3326905	0.9992	47.38	Heilongjiang
07DS09125	LINEAGE4	3845756	0.9668	57.54	Shanghai
07DS10228	LINEAGE4	5225853	0.9867	73.37	Jiangsu
07DS12228	LINEAGE4	7882355	0.9950	169.26	Anhui
07DS13142	LINEAGE4	3281153	0.9929	45.66	Fujian
07DS13143	LINEAGE4	6741593	0.9974	145.24	Fujian
07DS14111	LINEAGE4	4429661	0.9888	62.46	Jiangxi
07DS14151	LINEAGE4	3082021	0.9893	43.95	Jiangxi
07DS15211	LINEAGE4	3669734	0.9878	52.31	Shandong
07DS15455	LINEAGE4	4058486	0.9943	58.64	Shandong
07DS16571	LINEAGE4	6024358	0.9964	85.64	Henan
07DS17455	LINEAGE4	3284037	0.9961	46.93	Hubei
07DS18117	LINEAGE4	3038644	0.9938	43.18	Hunan
07DS18216	LINEAGE4	5790127	0.9966	81.55	Hunan
07DS18505	LINEAGE4	7474367	0.9972	159.01	Hunan
07DS18528	LINEAGE4	6232699	0.9963	88.85	Hunan
07DS19144	LINEAGE4	7749568	0.9969	161.34	Guangdong
07DS19150	LINEAGE4	8784553	0.9992	187.87	Guangdong
07DS20157	LINEAGE4	4350433	0.9931	61.72	Guangxi
07DS23135	LINEAGE4	7682808	0.9991	158.98	Sichuan
07DS27123	LINEAGE4	8631295	0.9954	185.10	Shaanxi
07DS30139	LINEAGE4	8289529	0.9961	179.52	Ningxia
07DS31107	LINEAGE4	11325838	0.9968	236.34	Xinjiang
07DS31129	LINEAGE4	8454763	0.9969	182.82	Xinjiang
07DS31172	LINEAGE4	10869306	0.9964	156.87	Xinjiang
07HR13127	LINEAGE4	7521560	0.9979	161.32	Fujian
07HR13241	LINEAGE4	5688232	0.9968	122.05	Fujian
07HR13254	LINEAGE4	8362889	0.9987	178.81	Fujian
07HR16156	LINEAGE4	7331670	0.9969	157.74	Henan
07HR18162	LINEAGE4	7795069	0.9974	168.49	Hunan
07HR23539	LINEAGE4	10393384	0.9954	224.92	Sichuan
07HR24165	LINEAGE4	9573228	0.9980	207.31	Guizhou
07HR24178	LINEAGE4	7260680	0.9992	156.65	Guizhou
07HR24235	LINEAGE4	8390063	0.9971	181.83	Guizhou
07HR24240	LINEAGE4	7680962	0.9962	166.31	Guizhou
07MDR04128	LINEAGE4	3004168	0.9840	46.80	Shanxi
07MDR04146	LINEAGE4	3873635	0.9857	62.73	Shanxi
07MDR05116	LINEAGE4	2091343	0.9946	31.74	Inner Mongolia
07MDR05134	LINEAGE4	6567896	0.9853	105.30	Inner Mongolia

07MDR06132	LINEAGE4	5044129	0.9973	79.60	Liaoning
07MDR06165	LINEAGE4	2808908	0.9958	42.44	Liaoning
07MDR08134	LINEAGE4	3451075	0.9878	52.88	Heilongjiang
07MDR09160	LINEAGE4	2283069	0.9956	34.39	Shanghai
07MDR10181	LINEAGE4	7809592	0.9954	116.33	Jiangsu
07MDR10328	LINEAGE4	1968695	0.9976	29.58	Jiangsu
07MDR11125	LINEAGE4	4619162	0.9829	70.70	Zhejiang
07MDR11168	LINEAGE4	9029761	0.9906	134.43	Zhejiang
07MDR11220	LINEAGE4	2737591	0.9923	41.24	Zhejiang
07MDR11239	LINEAGE4	2799336	0.9899	43.58	Zhejiang
07MDR13245	LINEAGE4	4832604	0.9917	58.70	Fujian
07MDR14265	LINEAGE4	7016985	0.9841	84.19	Jiangxi
07MDR16151	LINEAGE4	3794297	0.9885	60.73	Henan
07MDR17166	LINEAGE4	2311568	0.9929	31.53	Hubei
07MDR18212	LINEAGE4	4668878	0.9974	65.51	Hunan
07MDR18253	LINEAGE4	4878943	0.9949	71.07	Hunan
07MDR18557	LINEAGE4	3453099	0.9955	45.69	Hunan
07MDR23132	LINEAGE4	6551663	0.9977	92.74	Sichuan
07MDR23560	LINEAGE4	2675300	0.9911	41.40	Sichuan
07MDR23562	LINEAGE4	6140169	0.9882	85.67	Sichuan
07MDR24263	LINEAGE4	3528159	0.9949	48.72	Guizhou
07MDR24267	LINEAGE4	4477545	0.9902	63.70	Guizhou
AH12337	LINEAGE4	4668098	0.9929	147.06	Anhui
BJ01108	LINEAGE4	4969847	0.9960	148.82	Beijing
CQ22128	LINEAGE4	4879467	0.9930	155.79	Chongqing
CQ22143	LINEAGE4	5383610	0.9945	170.97	Chongqing
CQ22201	LINEAGE4	5502432	0.9951	172.78	Chongqing
CQ22225	LINEAGE4	2347328	0.9947	71.82	Chongqing
FJ13114	LINEAGE4	4661433	0.9976	140.02	Fujian
FJ13115	LINEAGE4	5667590	0.9954	173.42	Fujian
FJ13124	LINEAGE4	3954904	0.9931	126.43	Fujian
GD19127	LINEAGE4	7056366	0.9965	218.10	Guangdong
GD19316	LINEAGE4	4961801	0.9954	154.76	Guangdong
GX20134	LINEAGE4	4747034	0.9949	149.10	Guangxi
GZ24203	LINEAGE4	5022607	0.9935	158.31	Guizhou
GZ24224	LINEAGE4	4504584	0.9940	141.24	Guizhou
GZ24225	LINEAGE4	5188081	0.9940	163.91	Guizhou
HeB03322	LINEAGE4	5066512	0.9962	155.70	Hebei
HN16323	LINEAGE4	4703515	0.9972	146.03	Henan
HN16447	LINEAGE4	3450046	0.9970	104.50	Henan
HuN18124	LINEAGE4	3537389	0.9961	107.16	Hunan
HuN18423	LINEAGE4	4742483	0.9970	146.54	Hunan
HuN18451	LINEAGE4	5547221	0.9969	169.45	Hunan
HuN18510	LINEAGE4	4070791	0.9987	124.00	Hunan
HuN18554	LINEAGE4	3779086	0.9960	116.68	Hunan
JL07161	LINEAGE4	4748511	0.9961	145.55	Jilin
JS10426	LINEAGE4	3670962	0.9975	111.49	Jiangsu
JX14133	LINEAGE4	5229965	0.9971	160.95	Jiangxi
L45HLJ00006	LINEAGE4	4030234	0.9999	131.89	Heilongjiang
L45HLJ00027	LINEAGE4	8288846	0.9991	274.24	Heilongjiang
L45HLJ00043	LINEAGE4	7252907	0.9986	238.78	Heilongjiang
L45HLJ00051	LINEAGE4	14655899	0.9999	485.99	Heilongjiang
L45HLJ00061	LINEAGE4	4608846	0.9997	150.73	Heilongjiang
L45HLJ00063	LINEAGE4	6154221	0.9993	202.05	Heilongjiang
L45HLJ00069	LINEAGE4	5229735	0.9994	170.66	Heilongjiang
L45HLJ00115	LINEAGE4	5358089	0.9994	175.02	Heilongjiang

L45HLJ00165	LINEAGE4	3247111	0.9995	104.80	Heilongjiang
L45HLJ00168	LINEAGE4	3127754	0.9989	100.27	Heilongjiang
L45HLJ00231	LINEAGE4	6806127	0.9998	224.86	Heilongjiang
L45HLJ00501	LINEAGE4	10479503	0.9989	346.38	Heilongjiang
L45HLJ00541	LINEAGE4	3405242	0.9994	110.64	Heilongjiang
L45HLJ01081	LINEAGE4	4283268	0.9993	139.63	Heilongjiang
L45HLJ01271	LINEAGE4	4499978	0.9999	145.99	Heilongjiang
L45SD03930	LINEAGE4	8834649	0.9997	291.51	Shandong
L45SD06676	LINEAGE4	421037	0.9991	13.10	Shandong
L45SD06679	LINEAGE4	2576641	0.9995	83.55	Shandong
L45SD06903	LINEAGE4	8886105	0.9987	294.60	Shandong
L45SD06910	LINEAGE4	4997948	0.9986	163.67	Shandong
L45SD07756	LINEAGE4	3939955	0.9996	128.34	Shandong
L45SD07765	LINEAGE4	6024351	0.9991	197.63	Shandong
L45SD07768	LINEAGE4	2687010	0.9999	87.53	Shandong
L45SD07783	LINEAGE4	1624010	0.9993	51.62	Shandong
NX301079	LINEAGE4	3228752	0.9982	104.12	Ningxia
NX301099	LINEAGE4	1012965	0.9958	32.56	Ningxia
NX30136	LINEAGE4	4307911	0.9991	136.11	Ningxia
QH291004	LINEAGE4	3943322	0.9983	126.64	Qinghai
QH291026	LINEAGE4	2761201	0.9970	89.63	Qinghai
QH291034	LINEAGE4	3361775	0.9997	109.30	Qinghai
SaX271136	LINEAGE4	3263278	0.9962	106.01	Shaanxi
SaX271169	LINEAGE4	4518418	0.9965	146.84	Shaanxi
SaX27134	LINEAGE4	3311743	0.9991	105.04	Shaanxi
SaX27185	LINEAGE4	3900606	0.9986	124.05	Shaanxi
SaX27192	LINEAGE4	4094860	0.9979	131.05	Shaanxi
SC23106	LINEAGE4	5300260	0.9932	169.26	Sichuan
SC23504	LINEAGE4	5478925	0.9953	169.55	Sichuan
SC23519	LINEAGE4	5146755	0.9945	160.80	Sichuan
SC23553	LINEAGE4	4761992	0.9965	147.92	Sichuan
SC23558	LINEAGE4	5020115	0.9965	151.66	Sichuan
SDSYW10541	LINEAGE4	5608956	0.9994	113.37	Shandong
SDSYW6921	LINEAGE4	4212037	0.9965	85.60	Shandong
SDSYW7653	LINEAGE4	4738719	0.9960	96.33	Shandong
SDSYW7660	LINEAGE4	2080761	0.9966	42.05	Shandong
SHSYW131	LINEAGE4	5832992	0.9968	118.08	Shanghai
SX04218	LINEAGE4	6294702	0.9970	189.78	Shanxi
XJ311059	LINEAGE4	4902367	0.9983	159.01	Xinjiang
XJ311094	LINEAGE4	2749030	0.9974	88.37	Xinjiang
XJ31121	LINEAGE4	2912579	0.9980	87.65	Xinjiang
XJ312009	LINEAGE4	4687730	0.9985	151.96	Xinjiang
XJ312021	LINEAGE4	2776555	0.9982	88.94	Xinjiang
XJ312040	LINEAGE4	2540666	0.9966	81.55	Xinjiang
XJ312051	LINEAGE4	3179880	0.9978	103.07	Xinjiang
XJ312054	LINEAGE4	949674	0.9971	30.18	Xinjiang
XJ312182	LINEAGE4	1937195	0.9973	62.24	Xinjiang
XJ312194	LINEAGE4	3105087	0.9968	100.61	Xinjiang
XJ312210	LINEAGE4	992734	0.9968	32.21	Xinjiang
XJ312249	LINEAGE4	1823396	0.9980	58.68	Xinjiang
YN25113	LINEAGE4	5435002	0.9952	162.94	Yunan
07DS30138	LINEAGE4	3527003	0.9966	50.63	Ningxia
SCSYW2010330	LINEAGE4	4795038	0.9969	97.15	Sichuan
07DS22168	LINEAGE4	7152646	0.9947	152.40	Chongqing
SCSYW2010140	LINEAGE4	2780141	0.9894	57.12	Sichuan
SCSYW24	LINEAGE4	5504953	0.9959	111.52	Sichuan

SCSYW53	LINEAGE4	6408272	0.9950	129.93	Sichuan
SCSYW92	LINEAGE4	5306101	0.9934	107.44	Sichuan
SHSYW1610	LINEAGE4	5158862	0.9950	104.61	Shanghai
07DS31117	LINEAGE3	7912865	0.9968	164.83	Xinjiang
07DS31130	LINEAGE3	8176306	0.9965	170.18	Xinjiang
07DS31169	LINEAGE3	7144530	0.9974	152.94	Xinjiang
07DS31265	LINEAGE3	7085531	0.9976	147.49	Xinjiang
07MDR16424	LINEAGE3	5301190	0.9947	83.69	Henan
07MDR29131	LINEAGE3	3171998	0.9966	44.68	Qinghai
07MDR31213	LINEAGE3	2219793	0.9985	31.57	Xinjiang
07MDR31251	LINEAGE3	5507020	0.9891	81.52	Xinjiang
XJ311018	LINEAGE3	2595348	0.9972	83.46	Xinjiang
XJ311093	LINEAGE3	2452392	0.9986	78.84	Xinjiang
XJ312004	LINEAGE3	4125012	0.9990	133.90	Xinjiang
XJ312005	LINEAGE3	3708057	0.9981	119.88	Xinjiang
XJ312006	LINEAGE3	3036957	0.9975	97.88	Xinjiang
XJ312007	LINEAGE3	1008497	0.9964	32.78	Xinjiang
XJ312011	LINEAGE3	2214607	0.9980	71.09	Xinjiang
XJ312029	LINEAGE3	5028231	0.9989	163.57	Xinjiang
XJ312059	LINEAGE3	2428021	0.9969	78.01	Xinjiang
XJ312081	LINEAGE3	2885554	0.9962	93.98	Xinjiang
XJ312181	LINEAGE3	1340690	0.9987	42.77	Xinjiang
XJ312207	LINEAGE3	2042888	0.9989	65.26	Xinjiang
XJ312217	LINEAGE3	784016	0.9983	24.91	Xinjiang
XJ312221	LINEAGE3	1468810	0.9971	47.23	Xinjiang
XJ312240	LINEAGE3	2147255	0.9978	69.37	Xinjiang
SDSYW6693	LINEAGE1	4577321	0.9928	93.51	Shandong
07DS19453	LINEAGE1	7618472	0.9892	110.34	Guangdong
SHSYW1603	LINEAGE1	4398325	0.9954	89.07	Shanghai

Supplementary Table 4. The source information of 15,319 MTBC genomes representing the global diversity of MTBC.

Continent	County	Source	Isolate No.	Reference
Asia	China	Zhang H. 2013	161	Zhang H, Li D, Zhao L, et al. Genome sequencing of 161 Mycobacterium tuberculosis isolates from China identifies genes and intergenic regions associated with drug resistance. *Nature genetics* 2013.
Asia	China	Luo T. 2015	85	Luo T, Comas I, Luo D, et al. Southern East Asian origin and coexpansion of Mycobacterium tuberculosis Beijing family with Han Chinese. Proceedings of the National Academy of Sciences of the United States of America 2015; 112(26): 8136-41.
Asia	China	Luo T. 2014	26	Luo T, Yang C, Peng Y, et al. Whole-genome sequencing to detect recent transmission of Mycobacterium tuberculosis in settings with a high burden of tuberculosis. *Tuberculosis* 2014; **94**(4): 434-40.
Asia	Vietnam	Kathryn EH. 2017	1635	Genomic analysis of Mycobacterium tuberculosis reveals complex etiology of tuberculosis in Vietnam including frequent introduction and transmission of Beijing lineage and positive selection for EsxW Beijing variant.
Asia	India	Manson AL. 2017	605	Manson AL, Abeel T, Galagan JE, et al. Mycobacterium tuberculosis Whole Genome Sequences From Southern India Suggest Novel Resistance Mechanisms and the Need for Region-Specific Diagnostics. *Clin Infect Dis* 2017; **64**(11): 1494-501.
Asia	Uzbekistan	Walker TM. 2015	531	Walker TM, Kohl TA, Omar SV, et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. *The Lancet infectious diseases* 2015.
Asia	Thailand	Coker OO.2016	3	Coker OO, Chaiprasert A, Ngamphiw C, et al. Genetic signatures of Mycobacterium tuberculosis Nonthaburi genotype revealed by whole genome analysis of isolates from tuberculous meningitis patients in Thailand. *PeerJ* 2016; **4**: e1905.
Asia	Thailand	Faksri K. 2016	7	Faksri K, Tan JH, Disratthakit A, et al. Whole-Genome Sequencing Analysis of Serially Isolated Multi-Drug and Extensively Drug Resistant Mycobacterium tuberculosis from Thai Patients. *PloS one* 2016; **11**(8): e0160992.
Asia	Thailand	Regmi SM. 2015	4	Regmi SM, Chaiprasert A, Kulawonganunchai S, et al. Whole genome sequence analysis of multidrug-resistant Mycobacterium tuberculosis Beijing isolates from an outbreak in Thailand. *Mol Genet Genomics* 2015; **290**(5): 1933-41.
Asia	Afghan	Eldholm V. 2016	85	Eldholm V, Pettersson JH, Brynildsrud OB, et al. Armed conflict and population displacement as drivers of the evolution and dispersal of Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America* 2016.
Asia	Myanmar	Aung HL. 2016	14	Aung HL, Tun T, Moradigaravand D, et al. Whole-genome sequencing of multidrug-resistant Mycobacterium tuberculosis isolates from Myanmar. *J Glob Antimicrob Resist* 2016; **6**: 113-7.
Asia	Japan	Sekizuka. 2015	7	Sekizuka T, Yamashita A, Murase Y, et al. TGS-TB: Total Genotyping Solution for Mycobacterium tuberculosis Using Short-Read Whole-Genome Sequencing. *PloS one* 2015; **10**(11): e0142951.
Europe	Russia	Casali N. 2014	1292	Casali N, Nikolayevskyy V, Balabanova Y, et al. Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nature genetics* 2014; **46**(3): 279-86.
Europe	Russia	Casali N. 2012	58	Casali N, Nikolayevskyy V, Balabanova Y, et al. Microevolution of extensively drug-resistant tuberculosis in Russia. *Genome research* 2012; **22**(4): 735-45.
Europe	UK	Satta G. 2016	16	Satta G, Witney AA, Shorten RJ, Karlikowska M, Lipman M, McHugh TD. Genetic variation in Mycobacterium tuberculosis isolates from a London outbreak associated with isoniazid resistance. *BMC Med* 2016; **14**(1): 117.
Europe	UK	Walker TM. 2015	1966	Walker TM, Kohl TA, Omar SV, et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. *The Lancet infectious diseases* 2015.
Europe	UK	Walker TM. 2014	364	Walker TM, Lalor MK, Broda A, et al. Assessment of Mycobacterium tuberculosis transmission in Oxfordshire, UK, 2007-12, with whole pathogen genome sequences: an observational study. *The Lancet Respiratory medicine* 2014; **2**(4): 285-92.
Europe	UK	Walker TM. 2013	390	Walker TM, Ip CL, Harrell RH, et al. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. *The Lancet infectious diseases* 2013; **13**(2): 137-46.
Europe	UK	Koser CU. 2013	1	Koser CU, Bryant JM, Becq J, et al. Whole-genome sequencing for rapid susceptibility testing of M. tuberculosis. *The New England journal of medicine* 2013; **369**(3): 290-2.
Europe	UK	Torok ME. 2013	2	Torok ME, Reuter S, Bryant J, et al. Rapid whole-genome sequencing for investigation of a suspected tuberculosis outbreak. *J Clin Microbiol* 2013; **51**(2): 611-4.
Europe	UK	Walker TM. 2017	2	Walker TM, Crook DW, Peto TE, Conlon CP. Whole-genome sequencing identifies nosocomial transmission of extra-pulmonary M. tuberculosis. *QJM* 2016; **109**(12): 819-20.
Europe	Netherlands	Bryant JM. 2013	213	Bryant JM, Schurch AC, van Deutekom H, et al. Inferring patient to patient transmission of Mycobacterium tuberculosis from whole genome sequencing data. *BMC Infect Dis* 2013; **13**: 110.
Europe	Netherlands	Nebenzahl GH. 2016	0	Nebenzahl-Guimaraes H, van Laarhoven A, Farhat MR, et al. Transmissible Mycobacterium tuberculosis Strains Share Genetic Markers and Immune Phenotypes. *Am J Respir Crit Care Med* 2016.
Europe	Switzerland	Stucki D. 2015	74	Stucki D, Ballif M, Bodmer T, et al. Tracking a tuberculosis outbreak over 21 years: strain-specific single-nucleotide polymorphism typing combined with targeted whole-genome sequencing. *The Journal of infectious diseases* 2015; **211**(8): 1306-16.
Europe	Germany	Kohl TA. 2014	58	Kohl TA, Diel R, Harmsen D, et al. Whole-genome-based Mycobacterium tuberculosis surveillance: a standardized, portable, and expandable approach. *J Clin Microbiol* 2014; **52**(7): 2479-86.
Europe	Germany	Walker TM. 2015	846	Walker TM, Kohl TA, Omar SV, et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. *The Lancet infectious diseases* 2015.
Europe	Greenland	Bjorn-Mortensen K. 2016	275	Bjorn-Mortensen K, Soborg B, Koch A, et al. Tracing Mycobacterium tuberculosis transmission by whole genome sequencing in a high incidence setting: a retrospective population-based study in East Greenland. *Scientific reports* 2016; **6**: 33180.
Europe	Beralus	Wollenberg KR. 2016	598	Wollenberg KR, Desjardins CA, Zalutskaya A, et al. Whole genome sequencing of Mycobacterium tuberculosis provides insight into the evolution and genetic composition of drug-resistant tuberculosis in Belarus. *J Clin Microbiol* 2016.
Europe	Portugal	Phelan J. 2016	11	Phelan J, O'Sullivan DM, Machado D, et al. The variability and reproducibility of whole genome sequencing technology for detecting resistance to anti-tuberculous drugs. *Genome Med* 2016; **8**(1): 132.
Europe	Portugal	Perdigao J. 2014	84	Perdigao J, Silva H, Machado D, et al. Unraveling Mycobacterium tuberculosis genomic diversity and evolution in Lisbon, Portugal, a highly drug resistant setting. *BMC Genomics* 2014; **15**: 991.
Europe	Finland	Korhonen V. 2016	18	Korhonen V, Smit PW, Haanpera M, et al. Whole genome analysis of Mycobacterium tuberculosis isolates from recurrent episodes of tuberculosis, Finland, 1995-2013. *Clin Microbiol Infect* 2016; **22**(6): 549-54.
Africa	Malawi	Guerra-Assuncao JA. 2015	2267	Guerra-Assuncao JA, Crampin AC, Houben RM, et al. Large-scale whole genome sequencing of M. tuberculosis provides insights into transmission in a high prevalence area. *eLife* 2015; **4**.
Africa	Malawi	Guerra-Assuncao JA. 2015	353	Guerra-Assuncao JA, Houben RM, Crampin AC, et al. Recurrence due to relapse or reinfection with Mycobacterium tuberculosis: a whole-genome sequencing approach in a large, population-based cohort with a high HIV infection prevalence and active follow-up. *The Journal of infectious diseases* 2015; **211**(7): 1154-63.
Africa	South Africa	Black PA. 2015	14	Black PA, de Vos M, Louw GE, et al. Whole genome sequencing reveals genomic heterogeneity and antibiotic purification in Mycobacterium tuberculosis isolates. *BMC Genomics* 2015; **16**: 857.

Africa	South Africa	Witney AA. 2017	36	Witney AA, Bateson AL, Jindani A, et al. Use of whole-genome sequencing to distinguish relapse from reinfection in a completed tuberculosis clinical trial. *BMC Med* 2017; **15**(1): 71.
Africa	South Africa	Walker TM. 2015	504	Walker TM, Kohl TA, Omar SV, et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. *The Lancet infectious diseases* 2015.
Africa	South Africa	Cohen KA.2015	473	Cohen KA, Abeel T, Manson McGuire A, et al. Evolution of Extensively Drug-Resistant Tuberculosis over Four Decades: Whole Genome Sequencing and Dating Analysis of Mycobacterium tuberculosis Isolates from KwaZulu-Natal. *PLoS Med* 2015; **12**(9): e1001880.
Africa	Mali	Winglee K. 2016	356	Winglee K, Manson McGuire A, Maiga M, et al. Whole Genome Sequencing of Mycobacterium africanum Strains from Mali Provides Insights into the Mechanisms of Geographic Restriction. *PLoS Negl Trop Dis* 2016; **10**(1): e0004332.
Africa	Siera Lion	Walker TM. 2015	79	Walker TM, Kohl TA, Omar SV, et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. *The Lancet infectious diseases* 2015.
Africa	Djibouti	Blouin Y. 2012	38	Blouin Y, Hauck Y, Soler C, et al. Significance of the identification in the Horn of Africa of an exceptionally deep branching Mycobacterium tuberculosis clade. *PloS one* 2012; **7**(12): e52841.
Africa	Uganda	Clark TG. 2013	51	Clark TG, Mallard K, Coll F, et al. Elucidating emergence and transmission of multidrug-resistant tuberculosis in treatment experienced patients by whole genome sequencing. *PloS one* 2013; **8**(12): e83012.
Africa	Ethiopia	Comas I. 2015	66	Comas I, Hailu E, Kiros T, et al. Population Genomics of Mycobacterium tuberculosis in Ethiopia Contradicts the Virgin Soil Hypothesis for Human Tuberculosis in Sub-Saharan Africa. *Curr Biol* 2015; **25**(24): 3260-6.
America	America	Shea J. 2017	608	Shea J, Halse TA, Lapierre P, et al. Comprehensive Whole-Genome Sequencing and Reporting of Drug Resistance Profiles on Clinical Cases of Mycobacterium tuberculosis in New York State. *J Clin Microbiol* 2017; **55**(6): 1871-82.
America	America	Brown TS. 2016	34	Brown TS, Narechania A, Walker JR, et al. Genomic epidemiology of Lineage 4 Mycobacterium tuberculosis subpopulations in New York City and New Jersey, 1999-2009. *BMC Genomics* 2016; **17**(1): 947.
America	Canada	Lee RS. 2015	169	Lee RS, Radomski N, Proulx JF, et al. Population genomics of Mycobacterium tuberculosis in the Inuit. *Proceedings of the National Academy of Sciences of the United States of America* 2015; **112**(44): 13609-14.
America	Canada	Mehaffy C. 2014	60	Mehaffy C, Guthrie JL, Alexander DC, Stuart R, Rea E, Jamieson FB. Marked microevolution of a unique Mycobacterium tuberculosis strain in 17 years of ongoing transmission in a high risk population. *PloS one* 2014; **9**(11): e112928.
America	Canada	Gardy JL. 2011	36	Gardy JL, Johnston JC, Ho Sui SJ, et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *The New England journal of medicine* 2011; **364**(8): 730-9.
America	Haiti	Ocheretina O.2015	7	Ocheretina O, Shen L, Escuyer VE, et al. Whole Genome Sequencing Investigation of a Tuberculosis Outbreak in Port-au-Prince, Haiti Caused by a Strain with a "Low-Level" rpoB Mutation L511P - Insights into a Mechanism of Resistance Escalation. *PloS one* 2015; **10**(6): e0129207
America	Guatemala	Saelens JW. 2015	6	Saelens JW, Lau-Bonilla D, Moller A, et al. Whole genome sequencing identifies circulating Beijing-lineage Mycobacterium tuberculosis strains in Guatemala and an associated urban outbreak. *Tuberculosis* 2015; **95**(6): 810-6.
Australia	Multiple	Gurjav U. 2016	30	Gurjav U, Outhred AC, Jelfs P, et al. Whole Genome Sequencing Demonstrates Limited Transmission within Identified Mycobacterium tuberculosis Clusters in New South Wales, Australia. *PloS one* 2016; **11**(10): e0163612.
Multiple	Multiple	Comas I. 2013	171	Comas I, Coscolla M, Luo T, et al. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. *Nature genetics* 2013; **45**(10): 1176-82.
Multiple	Multiple	Phelan J. 2016	127	Phelan J, Coll F, McNerney R, et al. Mycobacterium tuberculosis whole genome sequencing and protein structure modelling provides insights into anti-tuberculosis drug resistance. *BMC Med* 2016; **14**: 31.
Multiple	Multiple	Pankhurst LJ. 2016	356	Pankhurst LJ, Del Ojo Elias C, Votintseva AA, et al. Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study. *The Lancet Respiratory medicine* 2016; **4**(1): 49-58.
Multiple	Multiple	Bryant JM. 2013	47	Bryant JM, Harris SR, Parkhill J, et al. Whole-genome sequencing to establish relapse or re-infection with Mycobacterium tuberculosis: a retrospective observational study. *The Lancet Respiratory medicine* 2013; **1**(10): 786-92.

Supplementary Table 5. Comparision of cluster rates between sublineages using different sets of VNTR loci.

VNTR-9 loci

	GX	HLJ	HN	SC	SD	SH	All Sites	OR [95%CI]	P value
L2.3	47%(27/58)	60%(87/146)	48%(62/130)	46%(30/65)	62%(83/134)	61%(142/233)	56%(431/766)	1	-
L2.2	31%(15/49)	25%(3/12)	39%(17/44)	19%(9/47)	16%(4/25)	35%(28/80)	30%(76/257)	0.326 [0.238-0.446]	<0.0001
L4.2	31%(5/16)	#DIV/0!	22%(2/9)	24%(4/17)	0%(0/5)	27%(6/22)	25%(17/69)	0.254 [0.135-0.457]	<0.0001
L4.4	24%(8/34)	60%(6/10)	0%(0/3)	18%(6/33)	12%(2/17)	8%(2/24)	20%(24/121)	0.192 [0.115-0.312]	<0.0001
L4.5	27%(4/15)	0%(0/15)	0%(0/8)	34%(15/44)	26%(5/19)	8%(2/25)	21%(26/126)	0.202 [0.123-0.323]	<0.0001

VNTR-9+3 loci

	GX	HLJ	HN	SC	SD	SH	All Sites	OR [95%CI]	P value
L2.3	21%(12/56)	39%(57/144)	18%(24/114)	6%(4/65)	20%(27/128)	34%(80/232)	28%(204/739)	1	-
L2.2	8%(4/48)	33%(4/12)	23%(10/40)	4%(2/47)	16%(4/22)	25%(20/79)	18%(44/248)	0.566 [0.384-0.822]	0.0019
L4.2	0%(0/16)	0%(/0)	22%(2/9)	12%(2/17)	0%(0/5)	36%(8/22)	17%(12/69)	0.552 [0.264-1.068]	0.0668
L4.4	12%(4/29)	50%(5/10)	0%(0/3)	12%(4/33)	12%(2/16)	8%(2/23)	15%(17/114)	0.460 [0.251-0.799]	0.0040
L4.5	0%(0/15)	0%(0/15)	0%(0/8)	14%(6/44)	21%(4/19)	8%(2/25)	10%(12/126)	0.276 [0.136-0.516]	<0.0001

Reference for the VNTR typing schemes:

Luo, T. et al. Development of a hierarchical variable-number tandem repeat typing scheme for Mycobacterium tuberculosis in China. PLoS One 9, e89726 (2014).

Supplementary Table 6. SNP typing results of global MTBC isolates based on whole genome sequences.

	Belarus	Vietnam	Uzbekistan	UK	Serria Leone	South Africa	Mali	Malawi	Inuit	India	Greenland	Germany	Russia	Ethiopia	China
lineage1.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
lineage1.1.1	0	46	0	1	3	0	3	0	0	0	0	5	2	3	0
lineage1.1.1.1	0	318	0	3	0	0	0	73	0	0	0	8	1	0	0
lineage1.1.2	0	0	0	111	0	0	0	182	0	238	0	12	2	0	0
lineage1.1.3	0	0	0	42	0	1	0	0	0	12	0	4	1	0	0
lineage1.2.1	0	23	0	44	0	0	0	98	0	6	0	18	0	0	1
lineage1.2.2	0	0	0	59	0	8	0	0	0	62	0	3	0	0	0
lineage2.1	0	58	0	0	0	0	0	0	0	0	0	2	0	0	16
lineage2.2	0	0	0	0	0	0	0	108	0	0	0	0	0	0	0
lineage2.2.1	373	795	446	104	2	87	5	0	0	47	0	56	738	2	892
lineage2.2.1.1	0	155	0	11	0	9	0	0	0	8	0	4	1	0	72
lineage2.2.1.2	0	5	0	0	0	0	0	0	1	4	0	0	1	0	0
lineage2.2.2	0	39	0	4	0	4	0	201	0	0	0	3	1	0	53
lineage3.1.1	0	0	0	108	0	9	0	5	0	0	0	5	1	0	0
lineage3.1.2	0	0	0	41	0	0	0	0	0	21	0	3	0	0	0
lineage3.1.2.1	0	0	0	26	0	0	0	0	0	18	0	2	0	0	0
lineage3.1.2.2	0	0	0	17	0	0	0	1	0	0	0	0	0	0	0
lineage4.1	4	0	0	10	10	5	169	0	0	0	0	12	17	10	0
lineage4.1.1	0	0	0	2	0	10	0	2	0	0	0	1	0	0	0
lineage4.1.1.1	0	0	0	70	1	14	0	0	0	2	0	3	0	1	0
lineage4.1.1.2	0	0	0	1	0	3	0	55	1	0	0	1	0	0	0
lineage4.1.1.3	0	0	0	153	4	40	3	82	0	2	0	4	0	4	1
lineage4.1.2	0	15	0	16	0	16	4	24	154	0	0	23	28	0	0
lineage4.1.2.1	26	42	11	86	10	33	20	0	0	4	0	205	44	10	0
lineage4.2	1	0	0	0	0	0	0	0	0	0	0	0	2	0	5
lineage4.2.1	35	0	12	5	0	0	0	6	0	0	0	19	108	0	1
lineage4.2.2	0	31	0	46	0	1	0	0	0	0	0	9	0	0	26
lineage4.2.2.1	0	1	0	1	0	0	0	10	0	0	0	24	0	0	0
lineage4.3	53	1	3	12	1	2	2	0	0	0	0	4	3	1	0
lineage4.3.1	0	2	0	9	0	0	1	0	0	0	0	8	0	0	0
lineage4.3.2	0	1	0	5	0	16	0	12	0	0	0	5	0	0	0
lineage4.3.2.1	0	0	0	6	0	63	0	36	0	0	0	1	0	0	0
lineage4.3.3	54	3	19	25	8	15	0	0	0	0	1	24	88	8	0
lineage4.3.4	0	0	0	3	0	6	0	82	0	0	0	0	0	0	0
lineage4.3.4.1	0	0	0	13	0	3	0	15	0	0	0	11	4	0	1
lineage4.3.4.2	0	0	0	10	3	11	0	952	0	0	0	10	4	3	0
lineage4.3.4.2.1	0	0	0	27	0	23	0	0	0	0	0	1	0	0	0
lineage4.4	0	0	0	5	0	0	0	0	0	0	0	1	0	0	0
lineage4.4.1	0	0	0	1	0	3	0	27	0	0	2	0	0	0	0
lineage4.4.1.1	0	4	4	16	1	43	0	11	0	2	27	18	5	1	0
lineage4.4.1.2	0	4	0	7	0	0	0	0	0	0	0	2	0	0	1
lineage4.4.2	0	35	0	7	0	1	0	0	0	0	0	4	0	0	101
lineage4.5	0	33	2	41	0	0	0	0	0	0	0	26	1	0	160
lineage4.6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
lineage4.6.1	0	0	0	2	0	0	0	15	0	0	0	0	0	0	0
lineage4.6.1.1	0	0	0	4	0	1	0	13	0	0	0	3	0	0	0
lineage4.6.1.2	0	0	0	7	0	0	0	1	0	0	0	1	0	0	0

lineage4.6.2	0	0	0	6	0	0	0	1	0	0	0	2	0	0	0
lineage4.6.2.1	0	0	0	3	0	0	0	3	0	0	0	0	0	0	0
lineage4.6.2.2	0	3	0	23	3	0	23	10	0	0	0	26	1	3	0
lineage4.7	8	1	0	10	0	5	4	42	0	0	0	34	3	0	0
lineage4.8	32	16	22	166	11	27	0	0	11	2	231	145	102	11	2
lineage4.9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LINEAGEBOV	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LINEAGEBOV_AFRI	0	0	0	0	18	0	0		0	0	0	0	0	18	0
All	586	1631	519	1369	75	459	234	2067	167	428	261	752	1158	75	1332

Supplementary Table 7. Primer and probe sequences for SNP genotyping

Probe/Primer	SNP	Probe/Primer sequence	5’-label	3’-label
Probe_L45	2789341 A-C	TGCCAGCGATCTTAGATTC	FAM	BQ1
F_L45		GTCCTTTGGCGCTAGAGCTT		
R_L45		GAAGGCGGCTACAAAGAAGG		
Probe_L44	3966059 G-C	TGCCACCGTACCCAGCG	ROX	BQ2
F_L44		GACGGTTGTCATTGCCTGT		
R_L44		CGTGGACTACCACGACTTGA		
Probe_L42	2181026 G-C	CACATCGCAGCAGACGGGCACGACCT	FAM	BQ1
F_L42		ATCCAACATGCTGGTCACCT		
R_L42		AGAGAAGTTGGCGGATGAGA		
Probe_L22	2752122 C-G	CTGGTGACGTGGTCACCAAC	FAM	BQ1
F_L22		TCAAAGCCCACGGCATCC		
R_L22		AACGTTGAAGAACGCCCGTA		
Probe_L23	1286766 G-C	AAGAACTGGGACTCGAGGT	ROX	BQ2
F_L23		CGGTAAGGTCGCCGCC		
R_L23		TTCAACGCAATATCGTCGCC		
Probe-L2-L4	2752122 G-C	CTGGTGACGTGGTCACCAAC	FAM	BQ1
F-L2-L4		TCAAAGCCCACGGCATCC		
R-L2-L4		AACGTTGAAGAACGCCCGTA		

Supplementary Discussion

Reconstructions of the global migratory history of the MTBC indicate that the pathogen spread *via* diverse human movements within the past several thousand years[1]. One might expect that the Chinese epidemic arose in the context of continuous introductions of new MTBC genotypes as the pandemic bloomed. However, although we identified more than ten different MTBC lineages/sublineages in our sample of 4,578 isolates, the population was dominated by the four indigenous genotypes that made up ~99.0% of the Chinese population of MTBC. In a previous analysis of the global population, relative rates of migration to and from East Asia were low[1]. Together with the results presented here, this suggests that distinct historical phenomena shaped the spread of TB in China. We did not observe the expansion of any indigenous or foreign strains during the last century, contradicting the assumption that the current TB epidemic in China is due to re-emergence of native or introduced populations of pathogens. On the contrary, we inferred a precipitous decline of N_e in MTBC population in the second half of the last century.

Descriptions of TB in Chinese historical texts predate the oldest introduction that we inferred for extant MTBC in China. This inconsistency could result from several scenarios: similarity in the clinical presentation of TB and other diseases, and/or clinical TB arising from a population of bacteria that are distinct from currently extant MTBC. A similar situation has been observed in the Americas, where the current population of MTBC clearly derives from European colonial migrations, but an abundance of morphological and DNA evidence indicated pre-Columbian TB[2,3]. The recent discovery of 1,000-year-old MTBC genomes from Peruvian human skeletons suggests that pre-Columbian TB in this region could have been caused by *M. pinnipedii*[4]. Ancient MTBC genomes from China would similarly clarify the early history of TB in this country.

Substitution rate model: MTBC-6 vs. MTBC-70 The substitution/mutation rate of MTBC has been debated many times[4-10]. There are two competing substitution rate models for MTBC. One is MTBC-70, which is based on the “out of Africa” hypothesis and assumption of codivergence between MTBC and its human host[6]; the other is MTBC-6, which was calibrated by ancient MTBC DNA in two studies[4,5]. In this study, we used the MTBC-6 model for dating because 1) the substitution rate of MTBC-6 model was calibrated by ancient MTBC DNA and a preponderance of recent literature favors this model [1,11-13]; 2) results from MTBC-6 model fit well with multiple aspects of the demographic history of China, e.g., the dispersal of L4.5 *via* the overland Silk Road, the synchronous growth of MTBC and human population, the decrease of MTBC N_e in the second half of last century. Evaluating the plausibility of the two competing models is not the goal of this project and our major findings are independent of the model we chose, as both models point to historical origins and expansions determining the current TB epidemic in China. Formal consensus on dating and correlations between TB history and historical events will probably have to wait for the analysis of more samples from ancient MTBC DNA to emerge[14].

1. O'Neill, M.B. *et al.* Lineage specific histories of Mycobacterium tuberculosis dispersal in Africa and Eurasia. *bioRxiv* (2017).
2. Salo, W.L., Aufderheide, A.C., Buikstra, J. & Holcomb, T.A. Identification of Mycobacterium tuberculosis DNA in a pre-Columbian Peruvian mummy. *Proceedings of the National Academy of Sciences* **91**, 2091-2094 (1994).
3. Roberts, C.A. & Buikstra, J.E. *The bioarchaeology of tuberculosis: a global perspective on a re-emerging disease*, (University Press of Florida, 2003).
4. Bos, K.I. *et al.* Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature* **514**, 494-7 (2014).
5. Kay, G.L. *et al.* Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe. *Nat Commun* **6**, 6717 (2015).
6. Comas, I. *et al.* Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. *Nat Genet* **45**, 1176-82 (2013).
7. Merker, M. *et al.* Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage. *Nat Genet* **47**, 242-9 (2015).
8. Ford, C.B. *et al.* Use of whole genome sequencing to estimate the mutation rate of Mycobacterium tuberculosis during latent infection. *Nat Genet* **43**, 482-6 (2011).
9. Ford, C.B. *et al.* Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. *Nat Genet* **45**, 784-90 (2013).
10. Eldholm, V. *et al.* Armed conflict and population displacement as drivers of the evolution and dispersal of Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **113**, 13881-13886 (2016).
11. Comas, I. *et al.* Population Genomics of Mycobacterium tuberculosis in Ethiopia Contradicts the Virgin Soil Hypothesis for Human Tuberculosis in Sub-Saharan Africa. *Curr Biol* **25**, 3260-6 (2015).
12. Mokrousov, I. *et al.* Emerging peak on the phylogeographic landscape of Mycobacterium tuberculosis in West Asia: Definitely smoke, likely fire. *Mol Phylogenet Evol* **116**, 202-212 (2017).
13. Eldholm, V. *et al.* Armed conflict and population displacement as drivers of the evolution and dispersal of Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* (2016).
14. Rood, J. Tracking Tuberculosis Over Time. (The Scientist, 2015).