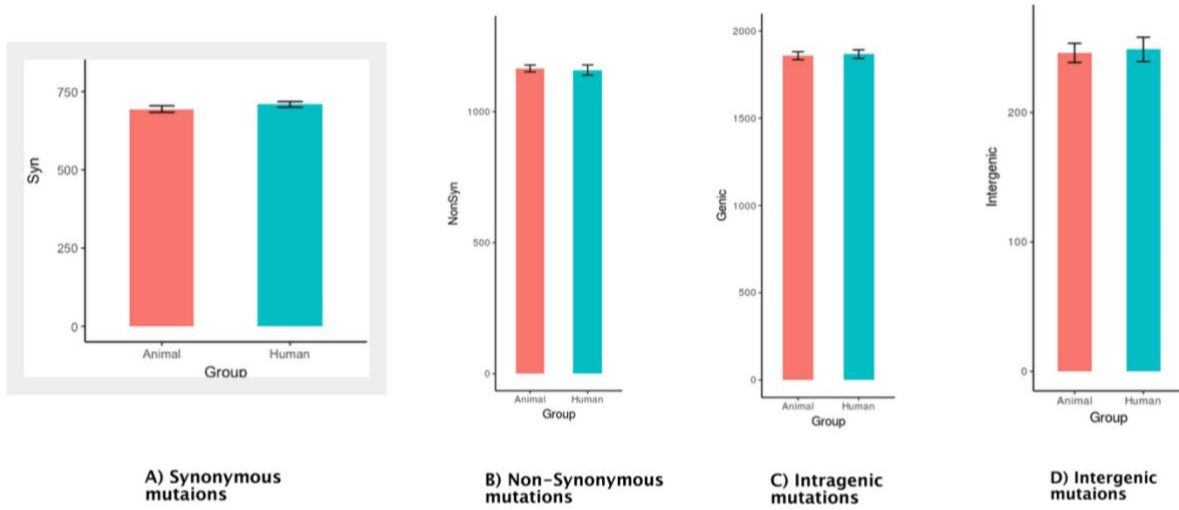


Isolation and comparative genomics of *Mycobacterium tuberculosis* isolates from cattle and their attendants in SouthIndia

Kannan Palaniyandi^{1,*,+}, Narender Kumar^{2,+}, Maroudam Veerasamy³, Ahmed Kabir Refaya¹, Chandrakumar Dolla¹, Subramanyam Balaji¹, Dhanaraj Baskaran¹, Kannan Thiruvengadam¹, Ananthi Rajendran¹, Sujatha Narayanan¹, Dhinakar Raj³, Soumya Swaminathan⁴, and Sharon J. Peacock^{2,5}

Supplementary Figure 1: Genetic variations identified among the genomes and plotted as groups:
Animal; isolates obtained from cattle and Human; isolates obtained from farm-workers.



Supplementary Table 1. IS6110 insertion sites identified among the genomes.

The ISMapper tool was used to identify the insertion sites of the *IS6110* insertion element. “X”

and “Y” represent the left and right side of the insertion site identified in each genome.

Strain id	Source	X	Y	Left gene (Rv ID)	Right gene (Rv ID)	Genic/Intergenic
AH03	Human	3120177	3120212	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3120540	3121857	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3119408	3119223	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
AH29	Human	3119223	3119408	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3120169	3120212	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3122380	3122521	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3123466	3123353	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
AH69	Human	3120540	3121856	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
AH85	Human	3119223	3119408	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3120177	3120212	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3123466	3123354	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
AH90	Human	3119223	3119408	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3120177	3120212	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3123466	3123350	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
AH91	Human	3120513	3121874	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
KH126	Animal	1895487	1895382	<i>Rv1667c</i>	<i>Rv1667c</i>	genic
		2041598	2041665	<i>ppe28</i> (Rv1800)	<i>ppe29</i> (Rv1801)	intergenic

		2368569	2368526	<i>ppe36</i> (Rv2108)	<i>prcA</i> (Rv2109c)	intergenic
		3120521	3121898	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		4056795	4056802	<i>espA</i> (Rv3616c)	<i>ephA</i> (Rv3617)	intergenic
KH127	Animal	888788	888796	oxidoreductase (Rv0794c)	insertion sequence element IS1547 transposase (Rv0797)	intergenic
		889020	890443	oxidoreductase (Rv0794c)	insertion sequence element IS1547 transposase (Rv0797)	intergenic
		1895413	1895394	Rv1667c	Rv1667c	genic
		1995677	1995767	hypothetical protein (Rv1761c)	hypothetical protein (Rv1761c)	genic
		1999316	1999263	hypothetical protein (Rv1761c)	hypothetical protein (Rv1766)	intergenic
		3120499	3121922	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3711993	3711772	transposase fusion protein (Rv3327)	transposase fusion protein (Rv3327)	genic
		4056783	4056802	<i>espA</i> (Rv3616c)	<i>ephA</i> (Rv3617)	intergenic
KH128	Animal	888783	888786	oxidoreductase (Rv0794c)	insertion sequence element IS1547 transposase (Rv0797)	intergenic
		1995761	1995764	hypothetical protein (Rv1761c)	hypothetical protein (Rv1761c)	genic
		1999271	1999275	hypothetical protein (Rv1761c)	hypothetical protein (Rv1766)	intergenic
		3119408	3119223	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3119551	3119449	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		3120212	3120177	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		4056799	4056803	<i>espA</i> (Rv3616c)	<i>ephA</i> (Rv3617)	Intergenic
KH143	Animal	3120538	3121866	hypothetical protein (Rv2813)	CRISPR-associated endoribonuclease <i>cas2</i> (Rv2816c)	intergenic
		4197857	4197790	hypothetical protein (Rv3749c)	hypothetical protein (Rv3749c)	genic
		1889840	1889683	<i>pks9</i> (Rv1664)	<i>pks9</i> (Rv1664)	genic

Supplementary Table 2. Sequence data metrics for the 10 *M. tuberculosis* isolates used in the study.

Strain Id	Source	Strain name (NCBI)	NCBI Bioproject id	Reads (x10⁶)	Contigs (>500b)	rRNA	tRNA	CDS	N50	Rd analysis	Lineage
AH03	Human	NIRTAH01	PRJNA512047	5.2	81	3	45	4120	129915	Rd239	L1
AH29	Human	NIRTAH02	PRJNA512047	10.2	91	3	45	4138	110988	Rd239	L1
AH69	Human	NIRTAH03	PRJNA512047	5.6	82	3	45	4118	129914	Rd239	L1
AH85	Human	NIRTAH04	PRJNA512047	16.7	89	3	45	4136	129915	Rd239	L1
AH90	Human	NIRTAH05	PRJNA512047	16.1	88	3	45	4127	125613	Rd239	L1
AH91	Human	NIRTAH06	PRJNA512047	1.3	96	3	45	4140	125590	Rd239	L1
KH126	Animal	NIRTAH07	PRJNA512047	1.2	101	3	45	4158	106348	Rd239	L1
KH127	Animal	NIRTAH08	PRJNA512047	1.1	87	3	45	4138	129915	Rd239	L1
KH128	Animal	NIRTAH09	PRJNA512047	5.0	87	3	45	4127	129915	Rd239	L1
KH143	Animal	NIRTAH10	PRJNA512047	5.6	119	3	45	4168	94243	Rd239	L1

