Mutational and phylogenetic analyses of the mycobacterial *mbt* gene cluster

Sivagami sundaram Chavadi[1], Karen L. Stirrett[2¥], Uthamaphani R. Edupuganti[1], Olivia Vergnolle[1], Gigani Sadhanandan[2‡], Emily Marchiano[2†], Che Martin[3], Wei-Gang Qiu[3], Clifford E. Soll[4], and Luis E. N. Quadri[1]*

[1]Department of Biology, Brooklyn College-City University of New York, Brooklyn, NY 11210, USA; [2]Department of Microbiology and Immunology, Weill Medical College of Cornell University, New York, NY 10021, USA; [3]Department Biological Sciences, Hunter College-City University of New York, New York, NY 10065, USA; [4]Department of Chemistry, Hunter College-City University of New York, New York, NY 10065, USA.

* Corresponding author. Mailing address: Luis E. N. Quadri, Ph.D., Biology Department, Brooklyn College, 2900 Bedford Ave., Brooklyn, NY 11210. Phone: 718-951-5000, Ext. 6254. Fax: 718-951-4659. E-Mail: LQuadri@brooklyn.cuny.edu

¥ Present address: Southeastern Community College, 4564 Chadbourn Hwy., P.O. Box 151, Whiteville, NC 28472.

‡ Present address: New York Methodist Hospital, CAHE, 1401 Kings Hwy., Brooklyn, NY 11229.

† Present address: New Jersey Medical School, UMDNJ, 185 South Orange Ave., Newark, NJ 07103.

Supplemental Information

Table S1. Plasmids

Plasmid	**Characteristics**	**Source or Reference**
pCR2.1Topo	Cloning vector, kanamycin resistance and ampicillin resistance genes	Invitrogen
pCP0	*E. coli*-mycobacteria shuttle vector for *hsp60* promoter-driven gene expression in mycobacteria, kanamycin resistance gene	(1)
p2NIL	Vector containing kanamycin resistance gene and OriE	(2)
pGOAL19	Vector containing hygromycin resistance gene, *sacB*-*lacZ Pac*I cassette, and OriE	(2)
pCP0-*mbtA*	pCP0 expressing *mbtA* (MSMEG_4516)	This study
pCP0-*mbtB*	pCP0 expressing *mbtB* (MSMEG_4515)	This study
pCP0-*mbtC*	pCP0 expressing *mbtC* (MSMEG_4513)	This study
pCP0-*mbtD*	pCP0 expressing *mbtD* (MSMEG_4512)	This study
pCP0-*mbtE*	pCP0 expressing *mbtE* (MSMEG_4511)	This study
pCP0-*mbtF*	pCP0 expressing *mbtF* (MSMEG_4510)	This study
pCP0-*mbtG*	pCP0 expressing *mbtG* (MSMEG_4509)	This study
pCP0-*mbtH*	pCP0 expressing *mbtH* (MSMEG_4508)	This study
pCP0-*mbtT*	pCP0 expressing *mbtT* (MSMEG_4514)	This study
p2NIL-GOALc-ΔmbtAc	suicide delivery vector carrying a *mbtA* deletion cassette (ΔmbtAc)	This study
p2NIL-GOALc-ΔmbtBc	suicide delivery vector carrying a *mbtB* deletion cassette (ΔmbtBc)	This study
p2NIL-GOALc-ΔmbtCc	suicide delivery vector carrying a *mbtC* deletion cassette (ΔmbtCc)	This study
p2NIL-GOALc-ΔmbtDc	suicide delivery vector carrying a *mbtD* deletion cassette (ΔmbtDc)	This study
p2NIL-GOALc-ΔmbtEc	suicide delivery vector carrying a *mbtE* deletion cassette (ΔmbtEc)	This study
p2NIL-GOALc-ΔmbtFc	suicide delivery vector carrying a *mbtF* deletion cassette (ΔmbtFc)	This study
p2NIL-GOALc-ΔmbtGc	suicide delivery vector carrying a *mbtG* deletion cassette (ΔmbtGc)	This study
p2NIL-GOALc-ΔmbtHc	suicide delivery vector carrying a *mbtH* deletion cassette (ΔmbtHc)	This study
p2NIL-GOALc-ΔmbtTc	suicide delivery vector carrying a *mbtT* deletion cassette (ΔmbtTc)	This study

Table S2. PCR primers and amplicon information used for mutant screening and verification

Mutant probed	Primer name	Primer sequence (5'-3')	Amplicon (bp)	
			Mutant strain	Wild-type strain
Δ*mbtA*	mbtAOR	GGGTGTGGCGGTCCGCAAATCCTCGACGGTGACCGGAAATTCGTT	1332	3001
	mbtAOF	CGGTGGCATAGACAGACAGGGCTGTCTATTGCCGTTTTCGGGTTA		
	mbtAF	*GAATTC*GAGGTCCAGCCGTGACTCTGACCA	0	1708
	mbtAR	*AAGCTT*TGAGGCGGACCTACCCGCCGAGCT		
Δ*mbtB*	mbtBOF	GGATGGGACCTCACGAGTTCGCGTGCAAGGTCGCGGTGGTCGAAA	1203	4679
	mbtBOR	CGATCGTCACGCCCTGATCGCACGAGTACCGGTTGAACGCCTCGT		
	mbtB1F	*AAGCTT*GTGTTCGGGAGGCCATATGGGCGA	75	3552
	mbtB2R	*GCTAGC*AGACCTCGGTGCCCGTCACGATGT		
Δ*mbtC*	mbtCOF	CTTCGAGGCGGGCGACTGGACACGGCTCGGGCCGCTGTGTCTGTT	1050	2319
	mbtCOR	GACGCCGTGCACGAACTGGGCGCCCTGGATCTGTACCTGTGAGAA		
	mbtCF	CTGCAGCGTCGCGGAGCTGGTCAATGAGCT	68	1337
	mbtCR	AAGCTTTCGGTCATCACGCGGCCTCCGGGA		
Δ*mbtD*	mbtDOF	ACGAGGACTTCTTCGAACTGGGCGGAGATTCCGTGCTCGCGACCG	1148	4115
	mbtDOR	AACCGTCGTCCCATGCGATGTGGTGTGCGACGAGCAGCAGTACGT		
	mbtD2F	CGAAGTGTCCGGACACCACCGCATCGCGGT	0	1582
	mbtD2R	GCTAGCATCGGCAATGTCAGTCACGCGAGA		
Δ*mbtE*	mbtEOF	CGCGAACCGTCTGCTCGACGTCATGGCGGAACGGCTTCGCTCCGA	1176	6606
	mbtEOR	TCATGATCGCGCCCGGCGAGCCAGCCGATGTAATCACGGTACGGT		
	mbtE2F	CGGATGGCCACCGGCGAGGCCTTCGACACA	0	2120
	mbtE2R	ACAGCACGTGGTGCTCAGCGGTGGTCAGCA		
Δ*mbtF*	mbtFOF	GCTGTGGATGCGTCGGCTGAAGCCGCGAACGGCGTTGCCGCACAA	1354	5779
	mbtFOR	CTGCGCGATGGACAACACACGCGGGTTGCCGGGAAGGATCGACCT		
	mbtF1F	*AAGCTT*CAGAGCCAAGACGGGAAACGATGA	78	4503
	mbtF3R	*GCTAGC*TCGTCTCGCTCACTCTCTCGCTCA		
Δ*mbtG*	mbtGOF	ACGCGGACAGCCGACACCCGCGCCGCTCTTGGCGCTGGAAACCCA	1247	2503
	mbtGOR	CCGCGCCGTGAGCAGGTGAAAGCAACCGGTGGAAAATTCGCCGAT		
	mbtGF	GAATTCTCGAGTGAGCGAGACGGTGAGCGA	100	1357
	mbtGR	AAGCTTGCCGTTGTCGTCATCGAACGGATT		
Δ*mbtH*	mbtHOF	CAGTCCTATCTCGTCGACATCGGCCAGTTCGCGGAGTGGAT	2033	2267
	mbtHOR	GAGCTCTGGTCCACGCGCCCGGTGTCGGTGATCA		
	mbtHF	CTGCAGAACGATCAGGAGAACCGATGAGCA	71	248
	mbtHR	AAGCTTCTGAACGGGCCTATCGAACGGCCT		
Δ*mbtT*	mbtTOF	CCGCGTCAAGCTCAGCGGCTACCGCATCGAACTGGGCGAGGTCGA	1400	2090
	mbtTOR	GGCGGCAGCGGCCGTCGTCGGACAGTGCGTGCTGTTTGGAGAATT		
	mbtTF	*CTGCAG*CTCGACACCGCCGCGACATCGTGA	99	787
	mbtTR	*AAGCTT*ACCGGGTCCTCACCAGAGCTCATT		

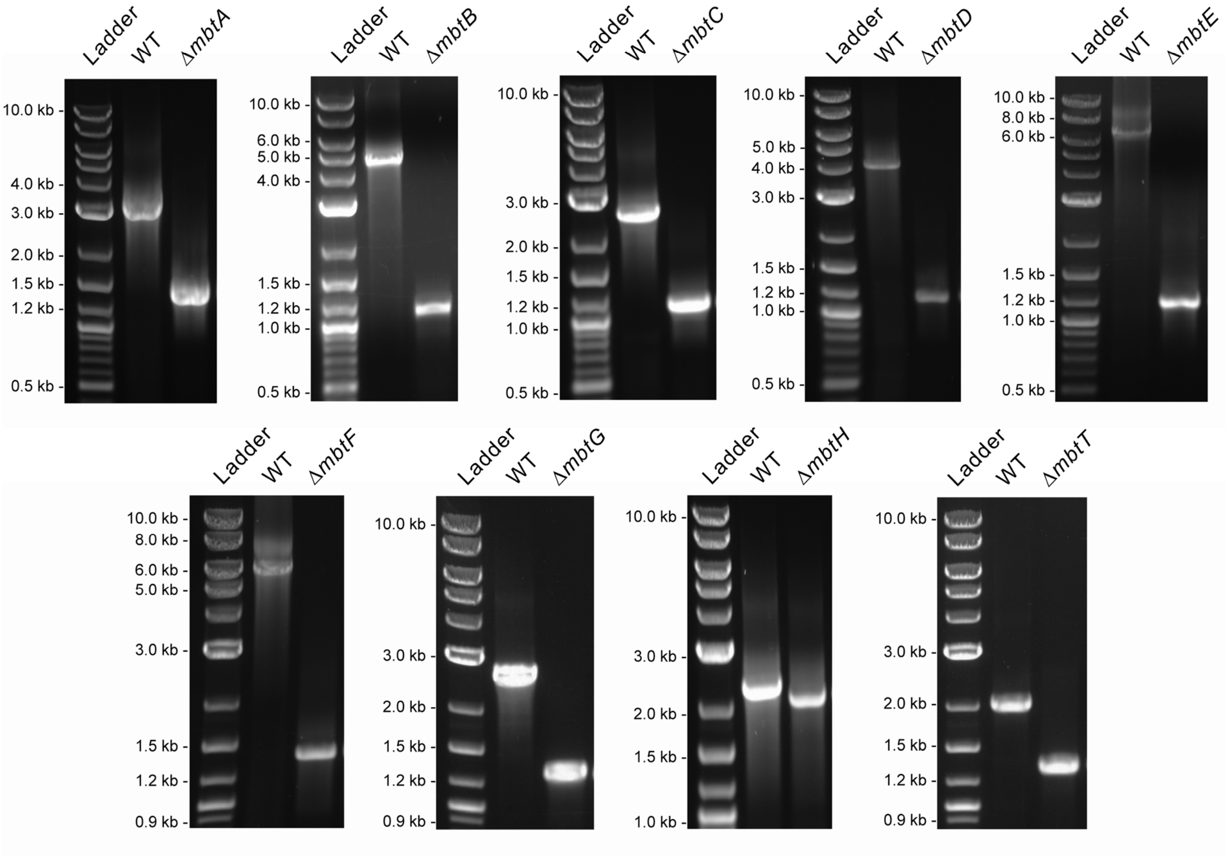

Figure S1. Agarose gel electrophoresis showing PCR-based confirmation of gene-specific deletions in Δ*mbt* mutants. PCR primers and amplicon information used for mutant screening and verification is presented in Table S2.

Table S3. Amplicon information and PCR primers

Amplicon name (bp)	**Primer name**	**Primer sequence (5’-3’)**	**Characteristics**
ΔmbtAc 5’ arm (681)	mbtAIF	GCTGAGGCGGACCTACACGGCTGGACCTCGGTGGGTAGGCGAAAT	Cloned using natural *Hin*dIII and *Not*I site from pCR2.1 Topo into p2NIL
	mbtAOR	GGGTGTGGCGGTCCGCAAATCCTCGACGGTGACCGGAAATTCGTT	
ΔmbtAc 3’ arm (679)	mbtAIR	CGAGGTCCAGCCGTGTAGGTCCGCCTCAGCGCCGGCCGAGTTCGA	
	mbtAOF	CGGTGGCATAGACAGACAGGGCTGTCTATTGCCGTTTTCGGGTTA	
ΔmbtAc full-length (1332)	mbtAOR	GGGTGTGGCGGTCCGCAAATCCTCGACGGTGACCGGAAATTCGTT	
	mbtAOF	CGGTGGCATAGACAGACAGGGCTGTCTATTGCCGTTTTCGGGTTA	
ΔmbtBc 5’ arm (621)	mbtBOF	GGATGGGACCTCACGAGTTCGCGTGCAAGGTCGCGGTGGTCGAAA	Cloned using natural *Hin*dIII and *Not*I site from pCR2.1 Topo into p2NIL
	mbtBIR	GGTGTCGAGTGCGGACATATGGCCTCCCGAACACGATTAGCACAG	
ΔmbtBc 3’ arm (612)	mbtBIF	TCGGGAGGCCATATGTCCGCACTCGACACCGCCGCGACATCGTGA	
	mbtBOR	CGATCGTCACGCCCTGATCGCACGAGTACCGGTTGAACGCCTCGT	
ΔmbtBc full-length (1203)	mbtBOF	GGATGGGACCTCACGAGTTCGCGTGCAAGGTCGCGGTGGTCGAAA	
	mbtBOR	CGATCGTCACGCCCTGATCGCACGAGTACCGGTTGAACGCCTCGT	
ΔmbtCc 5’ arm (1022)	mbtCOFNew	ACGAGGACTTCTTCGAACTGGGCGGAGATTCCGTGCTCGCGACCG	Cloned using natural *Kpn*I and *Not*I site from pCR2.1 Topo into p2NIL.
	mbtCIR	CGCGGCCTCCGGGATCTCACCAGAGCTCATTGACCAGCTCCGCGA	
ΔmbtCc 3’ arm (1031)	mbtCIF	ATGAGCTCTGGTGAGATCCCGGAGGCCGCGTGATGACCGACGGAA	
	mbtCORNew	TGACGACGTCGGCCGTGGCCGAGCCGATGAACTGCACGGGAGTCT	
ΔmbtCc full-length (2021)	mbtCOFNew	ACGAGGACTTCTTCGAACTGGGCGGAGATTCCGTGCTCGCGACCG	
	mbtCORNew	TGACGACGTCGGCCGTGGCCGAGCCGATGAACTGCACGGGAGTCT	
ΔmbtDc 5’ arm (598)	mbtDOF	CGCGGCGATGTTCGTCCTGCAACGCCGGTCCGCCGCAGTGCGCGA	Cloned using natural *Hin*dIII and *Not*I site from pCR2.1 Topo into p2NIL.
	mbtDIR	GAACGTTTCCTTTTCGGTCATCACGCGGCCTCCGGGATCGCGACC	
ΔmbtDc 3’ arm (594)	mbtDIF	GGCCGCGTGATGACCGAAAAGGAAACGTTCTCGCGTGACTGACAT	
	mbtDOR	AACCGTCGTCCCATGCGATGTGGTGTGCGACGAGCAGCAGTACGT	
ΔmbtDc full-length (1148)	mbtDOF	ACGAGGACTTCTTCGAACTGGGCGGAGATTCCGTGCTCGCGACCG	
	mbtDOR	AACCGTCGTCCCATGCGATGTGGTGTGCGACGAGCAGCAGTACGT	
ΔmbtEc 5’ arm (599)	mbtEOF	CGCGAACCGTCTGCTCGACGTCATGGCGGAACGGCTTCGCTCCGA	Cloned using natural *Hin*dIII and *Not*I site from pCR2.1 Topo into p2NIL.
	mbtEIR	TCGTTTCCCGTCTTGAATGTCAGTCACGCGAGAACGTTTCCTTTT	
ΔmbtEc 3’ arm (607)	mbtEIF	CGCGTGACTGACATTCAAGACGGGAAACGATGACCGCCACCGGAA	
	mbtEOR	TCATGATCGCGCCCGGCGAGCCAGCCGATGTAATCACGGTACGGT	
ΔmbtEc full-length (1176)	mbtEOF	CGCGAACCGTCTGCTCGACGTCATGGCGGAACGGCTTCGCTCCGA	
	mbtEOR	TCATGATCGCGCCCGGCGAGCCAGCCGATGTAATCACGGTACGGT	
ΔmbtFc 5’ arm (770)	mbtFOF	GCTGTGGATGCGTCGGCTGAAGCCGCGAACGGCGTTGCCGCACAA	Cloned using natural *Hin*dIII and *Not*I site from pCR2.1 Topo into p2NIL.
	mbtFIR	CCGTCTCGCTCACTCGGCGGTCATCGTTTCCCGTCTTGGCTCTGA	
ΔmbtFc 3’ arm (614)	mbtFIF	GAAACGATGACCGCCGAGTGAGCGAGACGGTGAGCGAGAGAGTGA	
	mbtFOR	CTGCGCGATGGACAACACACGCGGGTTGCCGGGAAGGATCGACCT	
ΔmbtFc full-length (1354)	mbtFOF	GCTGTGGATGCGTCGGCTGAAGCCGCGAACGGCGTTGCCGCACAA	
	mbtFOR	CTGCGCGATGGACAACACACGCGGGTTGCCGGGAAGGATCGACCT	
ΔmbtGc 5’ arm (594)	mbtGOF	ACGCGGACAGCCGACACCCGCGCCGCTCTTGGCGCTGGAAACCCA	
	mbtGIR	TCTGCTGATCGTTCGCTCGCTCACTCTCTCGCTCACCGTCTCGCT	
ΔmbtGc 3’ arm (683)	mbtGIF	GAGAGAGTGAGCGAGCGAACGATCAGGAGAACCGATGAGCACCAA	
	mbtGOR	CCGCGCCGTGAGCAGGTGAAAGCAACCGGTGGAAAATTCGCCGAT	

ΔmbtGc full-length (1247)	mbtGOF	ACGCGGACAGCCGACACCCGCGCCGCTCTTGGCGCTGGAAACCCA	
	mbtGOR	CCGCGCCGTGAGCAGGTGAAAGCAACCGGTGGAAAATTCGCCGAT	
ΔmbtHc 5’ arm (1027)	mbtHOF	CAGTCCTATCTCGTCGACATCGGCCAGTTCGCGGAGTGGAT	Cloned using natural *Bam*HI and *Sac*I site from pCR2.1 Topo into p2NIL.
	mbtHIR	CTATCGAACGGCCTCGAAGGTGCCGTTGTCGTCATCGAACGGATTGG	
ΔmbtHc 3’ arm (1036)	mbtHIF	GACAACGGCACCTTCGAGGCCGTTCGATAGGCCCGTTCAGGC	
	mbtHOR	GAGCTCTGGTCCACGCGCCCGGTGTCGGTGATCA	
ΔmbtHc full-length (2033)	mbtHOF	CAGTCCTATCTCGTCGACATCGGCCAGTTCGCGGAGTGGAT	
	mbtHOR	GAGCTCTGGTCCACGCGCCCGGTGTCGGTGATCA	
ΔmbtTc 5’ arm (711)	mbtTOF	CCGCGTCAAGCTCAGCGGCTACCGCATCGAACTGGGCGAGGTCGA	Cloned using natural *Hin*dIII and *Not*I site from pCR2.1 Topo into p2NIL.
	mbtTIR	GACCAGCTCCGCGACTTCGAGCGAGACCTCGGTGCCCGTCACGAT	
ΔmbtTc 3’ arm (719)	mbtTIF	GAGGTCTGCCTCGAAGTCGCGGAGCTGGTCAATGAGCTCTGGTGA	
	mbtTOR	GGCGGCAGCGGCCGTCGTCGGACAGTGCGTGCTGTTTGGAGAATT	
ΔmbtTc full-length (1400)	mbtTOF	CCGCGTCAAGCTCAGCGGCTACCGCATCGAACTGGGCGAGGTCGA	
	mbtTOR	GGCGGCAGCGGCCGTCGTCGGACAGTGCGTGCTGTTTGGAGAATT	
RBS-*mbtA* (1708)	mbtAF	*GAATTC*GAGGTCCAGCCGTGACTCTGACCA	*Eco*RI
	mbtAR	*AAGCTT*TGAGGCGGACCTACCCGCCGAGCT	*Hin*dIII
RBS-*mbtC* (1337)	mbtCF	*CTGCAG*CGTCGCGGAGCTGGTCAATGAGCT	*Pst*I
	mbtCR	*AAGCTT*TCGGTCATCACGCGGCCTCCGGGA	*Hin*dIII
RBS-*mbtG* (1357)	mbtGF	*GAATTC*TCGAGTGAGCGAGACGGTGAGCGA	*Eco*RI
	mbtGR	*AAGCTT*GCCGTTGTCGTCATCGAACGGATT	*Hin*dIII
RBS-*mbtH* (248)	mbtHF	*CTGCAG*AACGATCAGGAGAACCGATGAGCA	*Pst*I
	mbtHR	*AAGCTT*CTGAACGGGCCTATCGAACGGCCT	*Hin*dIII
RBS-*mbtT* (787)	mbtTF	*CTGCAG*CTCGACACCGCCGCGACATCGTGA	*Pst*I
	mbtTR	*AAGCTT*ACCGGGTCCTCACCAGAGCTCATT	*Hin*dIII
RBS-*mbtB* B1 (1845)	mbtB1F	*AAGCTT*GTGTTCGGGAGGCCATATGGGCGA	*Hin*dIII
	mbtB1R	CCTTTGGGTCCCATCACCGCGACGCTCTCA	Natural *Ppu*MI Site in the gene
RBS-*mbtB* B2 (1705)	mbtB2F	GGTGAGAGCGTCGCGGTGATGGGACCCAAA	
	mbtB2R	*GCTAGC*AGACCTCGGTGCCCGTCACGATGT	*Nhe*I
RBS-*mbtD* D1 (1449)	mbtD1F	*AAGCTT*GCGATCCCGGAGGCCGCGTGATGA	*Hin*dIII
	mbtD1R	CGATGCGGTGGTG*TCCGGA*CACTTCGGGCA	Natural *Bsp*EI
RBS-*mbtD* D2 (1574)	mbtD2F	CGAAGTG*TCCGGA*CACCACCGCATCGCGGT	
	mbtD2R	*GCTAGC*ATCGGCAATGTCAGTCACGCGAGA	*Nhe*I
RBS-*mbtE* E1 (1458)	mbtE1F	*AAGCTT*CGAAAAGGAAACGTTCTCGCGTGA	*Hin*dIII
	mbtE1R	GGGTGTGTCGA*AGGCCT*CGCCGGTGGCCAT	Natural *Stu*I
RBS-*mbtE* E2 (2125)	mbtE2F	CGGATGGCCACCGGCG*AGGCCT*TCGACACA	
	mbtE2R	ACAGCACGTGGT*GCTCAG*CGGTGGTCAGCA	Natural *Blp*I
RBS-*mbtE* E3 (1906)	mbtE3F	TGCTGACCACC*GCTGAG*CACCACGTGCTGT	
	mbtE3R	*GCTAGC*CCGGTGGCGGTCATCGTTTCCCGT	*Nhe*I
RBS-*mbtF* F1 (1267)	mbtF1F	*AAGCTT*CAGAGCCAAGACGGGAAACGATGA	*Hin*dIII
	mbtF1R	TCGAGCATTCC*GAATGC*GCCGTCGAGTGTT	Natural *Bsm*I
RBS-*mbtF* F2 (1890)	mbtF2F	AACACTCGACGGCGCATTGG*GAATGC*TCGA	
	mbtF2R	AAT*CCTCAGG*AGTGTGGTCGACGGCAGCGT	Natural *Bsu*36I
RBS-*mbtF* F3 (1313)	mbtF3F	ACGCTGCCGTCGACCACACT*CCTGAGG*ATT	
	mbtF3R	*GCTAGC*TCGTCTCGCTCACTCTCTCGCTCA	*Nhe*I

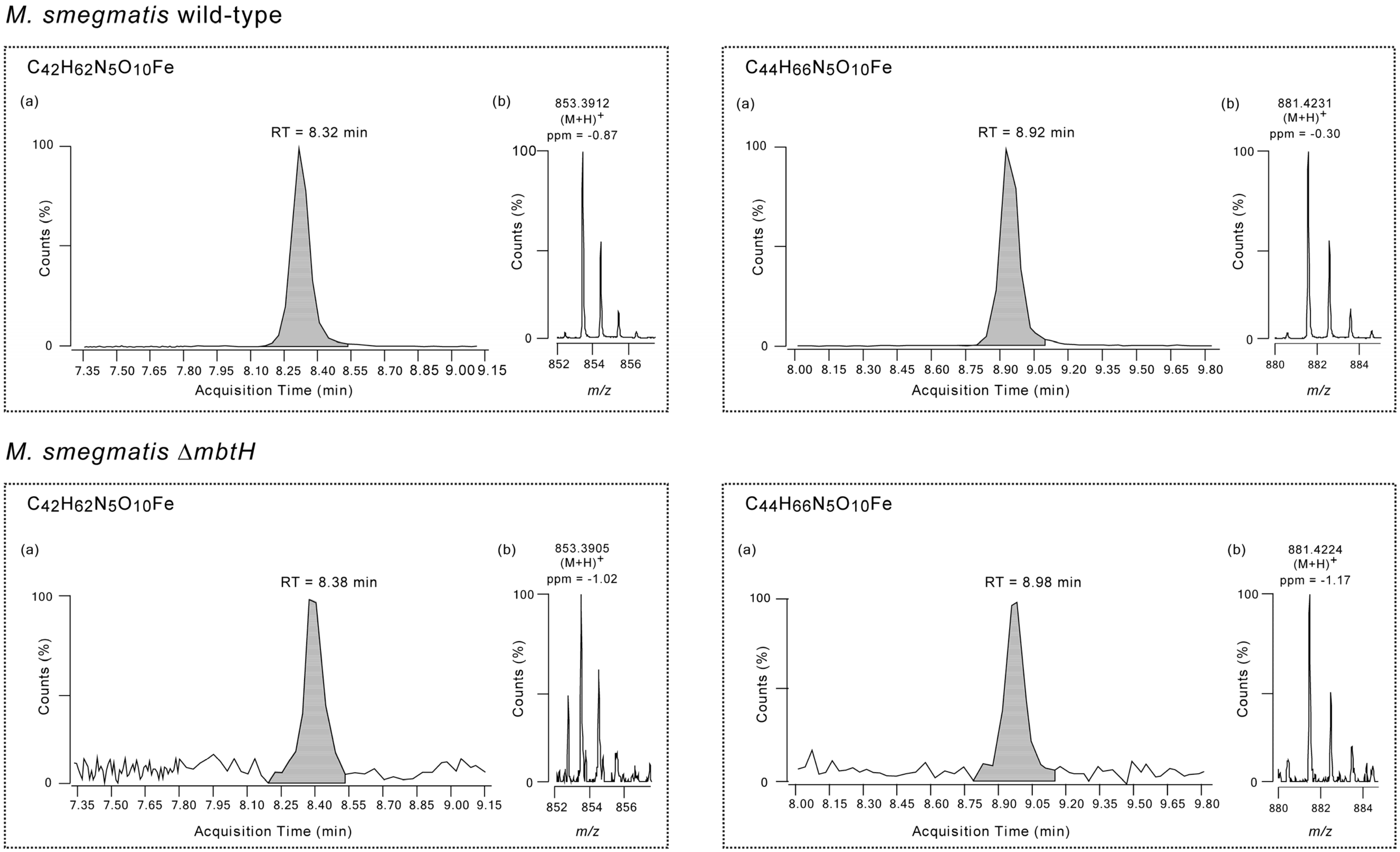

Figure S2. Extracted ion chromatograms (a) and mass spectra (b) supporting the presence of mycobactins in the *mbtH* knock-out mutant. The data from *M. smegmatis* wild-type + pCP0 and *M. smegmatis* Δ*mbtH* + pCP0 samples are consistent with the presence of two representative mycobactin species differing in the chain length of the acyl substituent on the hydroxylysine residue. The ion peak arrays in the left and right spectra match those predicted for ferrimycobactins of elemental composition $C_{42}H_{62}N_5O_{10}Fe$ and $C_{44}H_{66}N_5O_{10}Fe$, respectively, with less than 1.5 ppm error. Only the *m*/*z* of the most prominent (M+H)$^+$ ion peak in the isotopic distribution is indicated in each spectrum. The retention time (RT) is indicated in the chromatograms. Ferrimycobactin-derived ions were not detected by LC-MS in samples from *M. smegmatis* Δ*mbtC* + pCP0 (not shown), a representative mycobactin deficient mutant (Fig. 3). Mass spectral data was collected on an Agilent Technologies G6520A high resolution Q-TOF mass spectrometer attached to an Agilent Technologies 1200 Capillary HPLC system. Samples were ionized by electrospray ionization in positive mode. Chromatography was performed on an Thermo Hypersil Gold C8 column (2.1 × 50 mm, 3 µm, 175 Å, Part No. 25203-052130) using water with 0.1% formic acid and 5 µM ammonium formate (Solvent A) and methanol with 0.1% formic acid and 5 µM ammonium formate (Solvent B) at a flow rate of 400 µL/min. The HPLC gradient was as follows: 50%-100% B (0-10 min) and 100% B (10-15 min). Total analysis time was 15 min. The HPLC eluent was diverted to waste for the first 1.5 minutes. The temperature of the column was held at 50°C for the entire analysis. Mass spectrometer parameters were as follows: fragmentor = 165 V; drying gas temperature = 300°C; drying gas flow = 10 L/min; nebulizer pressure = 25 psi; capillary voltage = 3,500 V. Data was collected with the instrument set to low mass range (100-1,700 *m*/*z*) extended dynamic range conditions (2 GHz mode) and data was stored as both centroid and profile spectra with a threshold of 10 counts for MS mode. MS spectra were collected over a range of 100-1,700 *m*/*z* at 2 spectra/sec. The reference masses used were purine with M+H$^+$ ion at 121.0508 *m*/*z* and HP-922 with M+H$^+$ ion at 922.009798 *m*/*z*, were infused into the spray chamber using Agilent's calibrant delivery system. The instrument was controlled with Agilent MassHunter Workstation Acquisition Software B.03.01 and data was analyzed using Agilent MassHunter Workstation Qualitative Analysis Software B.04.00.

Table S4. Sequence accession numbers of 16S rRNA genes and sequence information of *mbt* genes and their orthologs

Bacterium	**16SrDNA**	**A**^{a}	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**T**
M. abscessus ATCC 19977	NC_010397	YP_001702982/ MAB_2247c	YP_001702859/ MAB_2124	YP_001702855/ MAB_2120c	YP_001702854/ MAB_2119c	YP_001702857/ MAB_2122	YP_001702858/ MAB_2123	YP_001702984/ MAB_2249	YP_001702985/ MAB_2250	YP_001702980/ MAB_2245	NP	YP_001702856/ MAB_2121c
M. avium 104	NC_008595	YP_881227/ MAV_2008	YP_881228/ MAV_2009	YP_881230/ MAV_2011	YP_881231/ MAV_2012	YP_881232/ MAV_2013	YP_881233/ MAV_2014	YP_881234/ MAV_2015	YP_882427/ MAV_3245	YP_881016/ MAV_1792	YP_881022/ MAV_1798	YP_881229/ MAV_2010
M. avium paratuberculosis K-10	NC_002944	NP_961112/ MAP2178/ *mbtA*	NP_961111/ MAP2177c/ *mbtB*	NP_961109/ MAP2175c/ *mbtC*	NP_961108/ MAP2174c/ *mbtD*	NP_961107/ MAP2173c/ *mbtE*	NP_961105/ MAP2171c/ *mbtF*	NP_961104/ MAP2170c/ *mbtG*	NP_961103/ MAP2169c/ *mbtH_3*	NP_961139/ MAP2205c/ *trpE2*	NP_961131 MAP2197/ *lipK*	NP_961110/ MAP2176c/
M. bovis BCG 1173P2	NC_008769	CAL72386/ BCG_2398/ *mbtA*	YP_978486/ BCG_2397c/ *mbtB*	YP_978485/ BCG_2396c/ *mbtC*	YP_978484/ BCG_2395c/ *mbtD*	YP_978483/ BCG_2394c/ *mbtE*	YP_978482/ BCG_2393c/ *mbtF*	YP_978481/ BCG_2392c/ *mbtG*	YP_978480/ BCG_2391c	YP_978489/ BCG_2400c/ *mbtI*	YP_978488/ BCG_2399	B-T
M. bovis subsp. bovis AF2122/97	BX248338	*	*	*	*	*	*	*	*	*	*	*
M. gilvum PYR-GCK	NC_009338	YP_001133956/ Mflv_2691	YP_001133957/ Mflv_2692	YP_001133959/ Mflv_2694	YP_001133960/ Mflv_2695	YP_001133961/ Mflv_2696	YP_001133962/ Mflv_2697	YP_001133963/ Mflv_2698	YP_001133964/ Mflv_2699	YP_001134864/ Mflv_3602	YP_001135862/ Mflv_4606	YP_001133958/ Mflv_2693
M. marinum M	NC_010612	YP_001851969/ MMAR_3698/ *mbtA*	YP_001851960/ MMAR_3689/ *mbtB*	YP_001851965/ MMAR_3694/ *mbtC*	YP_001851966/ MMAR_3695/ *mbtD*	YP_001851962/ MMAR_3691/ *mbtE*	YP_001851961/ MMAR_3690/ *mbtF*	YP_001851968/ MMAR_3697/ *mbtG*	YP_001851967/ MMAR_3696 *mbtH_1*	YP_001851977/ MMAR_3706/ *mbtI*	YP_001851970/ MMAR_3699/ *mbtJ*	YP_001851964/ MMAR_3693
M. smegmatis MC2 155	NC_008596	YP_888788/ MSMEG_4516	YP_888787/ MSMEG_4515	YP_888785/ MSMEG_4513	YP_888784/ MSMEG_4512	YP_888783/ MSMEG_4511	YP_888782/ MSMEG_4510	YP_888781/ MSMEG_4509	YP_888780/ MSMEG_4508	YP_888796/ MSMEG_4524	YP_889517/ MSMEG_5271	YP_888786/ MSMEG_4514
M. sp. JLS	NC_009077	YP_001071751/ Mjls_3482	YP_001071750/ Mjls_3481	YP_001071748/ Mjls_3479	YP_001071747/ Mjls_3478	YP_001071746/ Mjls_3477	YP_001071745/ Mjls_3476	YP_001071744/ Mjls_3475	YP_001071743/ Mjls_3474	YP_001071339/ Mjls_3069	YP_001071472/ Mjls_3202	YP_001071749/ Mjls_3480
M. sp. KMS	NC_008705	YP_939518/ Mkms_3534	YP_939517/ Mkms_3533	YP_939515/ Mkms_3531	YP_939514/ Mkms_3530	YP_939513/ Mkms_3529	YP_939512/ Mkms_3528	YP_939511/ Mkms_3527	YP_939510/ Mkms_3526	YP_939096/ Mkms_3112	YP_939239/ Mkms_3255	YP_939516/ Mkms_3532
M. sp. MCS	NC_008146	YP_640634/ Mmcs_3471	YP_640633/ Mmcs_3470	YP_640631/ Mmcs_3468	YP_640630/ Mmcs_3467	YP_640629/ Mmcs_3466	YP_640628/ Mmcs_3465	YP_640627/ Mmcs_3464	YP_640626/ Mmcs_3463	YP_640216/ Mmcs_3053	YP_640356/ Mmcs_3193	YP_640632/ Mmcs_3469
M. tuberculosis CDC1551	NC_002755	*	*	*	*	*	*	*	*	*	*	*
M. tuberculosis H37Ra	NC_009525	*	*	*	*	*	*	*	*	*	*	*
M. tuberculosis H37Rv	BX842576	NP_216900/ Rv2384/ *mbtA*	NP_216899/ Rv2383c/ *mbtB*	NP_216898/ Rv2382c/ *mbtC*	NP_216897/ Rv2381c/ *mbtD*	NP_216896/ Rv2380c/ *mbtE*	NP_216895/ Rv2379c/ *mbtF*	NP_216894/ Rv2378c/ *mbtG*	NP_216893/ Rv2377c/ *mbtH*	YP_177877/ Rv2386c/ *mbtI*	YP_177876/ Rv2385/ *mbtJ*	B-T
M. tuberculosis F11	NC_009565	*	*	*	*	*	*	*	*	*	*	*
M. ulcerans Agy99	NC_008611	YP_907245/ MUL_3641/ *mbtA*	YP_907236/ MUL_3632/ *mbtB*	YP_907241/ MUL_3637/ *mbtC*	YP_907242/ MUL_3638/ *mbtD*	YP_907238/ MUL_3634/ *mbtE*	YP_907237/ MUL_3633/ *mbtF*	YP_907244/ MUL_3640/ *mbtG*	YP_907243/ MUL_3639/ *mbtH_1*	YP_907251/ MUL_3648/ *mbtI*	YP_907246/ MUL_3642/ *mbtJ*	YP_907240/ MUL_3636/
M. vanbaalenii PYR-1	NC_008726	YP_954637/ Mvan_3850	YP_954636/ Mvan_3849	YP_954634/ Mvan_3847	YP_954633/ Mvan_3846	YP_954632/ Mvan_3845	YP_954631/ Mvan_3844	YP_954630/ Mvan_3843	YP_954629/ Mvan_3842	YP_953627/ Mvan_2814	YP_955456/ Mvan_4675	YP_954635/ Mvan_3848
N. farcinica IFM 10152	NC_006361	YP_116829/ nfa6200	YP_116977/ nfa7680/ ntbF	YP_116973/ nfa7640/ nbtB	YP_116974/ nfa7650/ nbtC	YP_116975/ nfa7660/ nbtD	YP_116976/ nfa7670/ nbtE	YP_116970/ nfa7610/ nbtG	NP	YP_116828/ nfa6190	YP_116414/ nfa2080	YP_116972/ nfa7630

[a], letters in column headings correspond to the letter nomenclature of the *mbt* genes of *M. tuberculosis*. Sequence information is given in the format: accession number / locus tag / gene name (when available). The star (*) indicates that the sequence is identical to that in the strain of the same species for which sequence information is given. Only one representative strain per species was included in our parsimonious reconstruction analysis when the sequences were identical. B-T indicates a gene fusion that results in the fusion of MbtT to the C-terminus of MbtB. NP indicates that the gene is not present.

Reference

(1) **Onwueme, K. C., J. A. Ferreras, J. Buglino, C. D. Lima, and L. E. Quadri.** 2004. Mycobacterial polyketide-associated proteins are acyltransferases: Proof of principle with *Mycobacterium tuberculosis* PapA5. Proc. Natl. Acad. Sci. U. S. A. **101:**4608-4613.

(2) **Parish, T., and N. G. Stoker.** 2000. Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis tlyA plcABC* mutant by gene replacement. Microbiology **146:**1969-1975.