

Supplementary Table S1. Set of randomly selected clinical isolates* used to calculate discriminatory index of PCR-based genotyping methods

Name or number	Place of isolation	Time of isolation	hsp65 type (code)†	LSP-MVR type‡	8-locus MIRU type§	rep-PCR type**
BR80	São Paulo, Brazil	1998	CCAGCGA (1)	1111-5282	2522-1228	unique
W120	Southern California	1989	CCAGCGA (1)	2221-4282	2422-1228	unique
W214	Southern California	1991	CCAGCGA (1)	1221-4282	2422-1228	B
W272	Southern California	1991	CCGGCGG (2)	2112-5483	0543-1228	not done
W41	Southern California	1990	CCAGCGA (1)	1212-4282	2422-1228	X
W61	Southern California	1990	CCAGCGA (1)	2211-3383	3333-1228	unique
9033	Montreal, Canada	1999-2003	GGGGGGG (16)	211n-1n81	n1p1-nn(>12)8	not done
NET11 (NLA009600829)	Netherlands	1996	CGGCGGG (3)	1n11-5282	2522-1228	not done
NET17 (NLA009600751)	Netherlands	1996	CCAGCGA (1)	1121-3282	2322-1228	C
NET19 (NLA009601234)	Netherlands	1996	CCAGCGA (1)	1121-2282	2222-1228	T
NET20 (NLA009601375)	Netherlands	1996	CGGCGGG (3)	1111-4282	2422-1228	unique
NET38 (NLA009601391)	Netherlands	1996	CCAGCGA (1)	1111-2282	2222-1228	E
NET39 (NLA009601393)	Netherlands	1996	CCAGCGA (1)	1111-2282	2222-1228	E
NET56 (NLA009601899)	Netherlands	1996	CCAGCGA (1)	1121-3282	2322-1228	C
NET61 (NLA009601392)	Netherlands	1996	CCAGCGA (1)	1111-2282	2222-1228	F
NET63 (NLA009600536)	Netherlands	1996	CGGCGGG (3)	1111-4282	2422-1228	F
NET67 (NLA009600205)	Netherlands	1995	CGGCGGG (3)	1111-4282	2422-1228	I
NET71 (NLA009600531)	Netherlands	1996	CCAGCGA (1)	1111-2282	2222-1228	F
NET76 (NLA009601555)	Netherlands	1996	CCAGCGA (1)	1111-2282	2222-1228	F
HMC10	Seattle, U.S.A.	1999-2001	CCAGCGA (1)	1221-2282	2222-1228	unique
HMC11	Seattle, U.S.A.	1999-2001	CCAGCGA (1)	1n12-4282	2422-1218	not done
HMC19	Seattle, U.S.A.	1999-2001	CCAGCGA (1)	1121-3282	2322-1228	P
HMC20	Seattle, U.S.A.	1999-2001	CCGGCGG (2)	2212-3383	3333-n298	unique
102	Southern California	1983	CCGGCGG (2)	2212-5483	1543-1228	unique
105	Southern California	1983	CCAGCGA (1)	1221-4282	2422-1228	unique
107	Southern California	1983	CCGGCGG (2)	2112-3383	3333-1298	unique
108	Southern California	1983	CCGGCGG (2)	2112-3383	3333-1298	J
110	Southern California	1983	CCGGCGG (2)	2212-5483	0543-1228	unique

*isolates in this set were previously described in Horan *et al* 2006 (7)

†hsp65 type (code) = Concatenation of bases present at hsp65 633, 1128, 1218, 1269, 1272, 1488 and 1536 and hsp65 code #s correspond to those presented in Turenne *et al* 2006 (15) except 16 and 17, which correspond to sequences not previously reported (GenBankAccession #1198075 and 1198086)

‡LSP-MVR type = large-sequence polymorphism/MIRU/VNTR analysis; digits of type represent loci HSDR, DA2/LSP2, DA7/LSP7, DEL11/LSPP5, X3, 25, 32, 47 (7); "1" is used to indicate the absence of the large-sequence polymorphism, which corresponds to product lengths of 343bp (for HSDR), 834bp (for DA2/LSP2), 329bp (for DA7/LSP7) and 728bp (for DEL11/LSPP5). "2" is used to indicate the presence of the large-sequence polymorphism, which corresponds to product lengths of 480bp (for HSDR), 954bp (for DA2/LSP2), 181bp (for DA7/LSP7) and 794bp (for DEL11/LSPP5) (7); digits for MIRU/VNTR type represent the number of repeats inferred from the product length at that locus, based on product sizes and repeat numbers from Thibault *et al* (14) and personal correspondence with V. Thibault: 196bp when 2 repeats of 53bp are present [X3], 350bp when 3 repeats of 58bp are present [25], 298bp when 8 repeats of 18bp are present [32] and 217bp when 3 repeats of 35bp are present [47].

§ 8-locus MIRU-VNTR = mycobacterial interspersed repetitive unit – variable-number tandem repeat, using loci: 292, X3, 25, 47, 3, 7, 10, 32; n = no bands observed

**rep-PCR type = repetitive-sequence-based PCR type; results for isolates in this analysis were combined with results for all MAC isolates in the SBRI rep-PCR library using the Diversilab software, and letter type designations were assigned to branches of the dendrogram with at least 92% average similarity

Supplementary Table S2. Clinical strains used in geographical analysis

Name or number	Place of isolation	Time of isolation	hsp65 type (code)*	LSP-MVR type†	rep-PCR type‡	PFGE pattern§	Citation(s)
1061-1-BC-2 o	Boston, U.S.A.	1990s	not done	2112-3282	not done	not done	(16-18)
AR14	Little Rock, U.S.A.	1999	not done	not done	unique	not done	(7,9)
BR27	São Paulo, Brazil	1998	not done	2111-2483	not done	not done	(7,10)
BR34	São Paulo, Brazil	1998	not done	2212-3383	unique	not done	(7,10)
BR37	São Paulo, Brazil	1998	not done	not done	unique	not done	(7,10)
BR51	São Paulo, Brazil	1998	not done	2112-1683	S	not done	(7,10)
BR54	São Paulo, Brazil	1998	not done	2112-1682	S	not done	(7,10)
BR62BR	São Paulo, Brazil	1998	not done	2112-1183	not done	not done	(7,10)
BR62TL	São Paulo, Brazil	1998	not done	not done	H	not done	(7,10)
BR65	São Paulo, Brazil	1998	not done	not done	H	not done	(7,10)
BR67	São Paulo, Brazil	1998	not done	not done	unique	not done	(7,10)
BR69	São Paulo, Brazil	1998	not done	1121-5282	unique	not done	(7,10)
BR75	São Paulo, Brazil	1998	not done	2112-3383	unique	not done	(7,10)
BR80	São Paulo, Brazil	1998	CCAGCGA (1)	1111-5282	unique	not done	(7,10)
5026-1-F-3	Hanover, U.S.A.	1990s	not done	2112-5283	not done	not done	(16-18)
5029-3-SK-1	Hanover, U.S.A.	1990s	not done	1121-4282	not done	not done	(16-18)
5047-1-F-1 o	Hanover, U.S.A.	1990s	not done	2211-2383	not done	not done	(16-18)
5060-1-R-3	Hanover, U.S.A.	1990s	not done	2212-3283	not done	not done	(16-18)
CW12	Southern California	1994	not done	1211-2483	not done	not done	(1,2,7)
CW22	Southern California	1994	not done	1211-4282	not done	4	(1,2,7)
CW26	Southern California	1994	CCAGCGA (1)	not done	C	not done	(1,2,7)
CW4	Southern California	1993	not done	2211-2282	unique	not done	(1,2,7)
CW5	Southern California	1994	not done	1211-2483	not interpretable	2	(1,2,7)
CW68	Southern California	1995	not done	1211-2483	not done	3	(1,2,7)
W10	Southern California	1990	not done	1121-4282	D	not done	(1,2,7)
W17	Southern California	1990	not done	not done	F	not done	(1,2,7)
W112	Southern California	1989	not done	2211-2283	unique	1	(1,2,7)
W120	Southern California	1989	CCAGCGA (1)	2221-4282	unique	not done	(1,2,7)
W123	Southern California	1989	not done	not done	unique	not done	(1,2,7)
W124	Southern California	1989	not done	1211-4282	not done	4	(1,2,7)
W126	Southern California	1989	not done	1221-3282	B	8	(1,2,7)
W129	Southern California	1989	not done	1121-3282	P	9	(1,2,7)
W166	Southern California	1989	not done	1211-2483	not done	1	(1,2,7)
W167	Southern California	1989	not done	1112-4282	unique	not done	(1,2,7)
W182	Southern California	1990	not done	1211-2483	unique	1	(1,2,7)
W187	Southern California	1991	not done	2221-3482	unique	not done	(1,2,7)
W198	Southern California	1991	not done	1121-2282	unique	not done	(1,2,7)
W203	Southern California	1991	not done	2211-2283	unique	1	(1,2,7)
W205	Southern California	1991	not done	not done	D	not done	(1,2,7)
W214	Southern California	1991	CCAGCGA (1)	1221-4282	B	7	(1,2,7)
W216-8	Southern California	1991	not done	1221-4282	X	5	(1,2,7)
W272	Southern California	1991	CCGGCGG (2)	2112-5483	not done	not done	(1,2,7)
W349	Southern California	1993	not done	1211-2283	not done	not done	(1,2,7)
W350-8	Southern California	1993	not done	not done	unique	not done	(1,2,7)
W355	Southern California	1993	CCGGCGG (2)	2112-3383	unique	13	(1,2,7)
W357	Southern California	1993	not done	not done	N	not done	(1,2,7)
W393	Southern California	1993	not done	1211-4282	not done	6	(1,2,7)
W394	Southern California	1993	not done	2112-3383	not done	not done	(1,2,7)
W41	Southern California	1990	CCAGCGA (1)	1212-4282	X	not done	(1,2,7)
W60	Southern California	1990	not done	not done	D	not done	(1,2,7)
W61	Southern California	1990	CCAGCGA (1)	2211-3383	unique	not done	(1,2,7)
W80	Southern California	1990	not done	not done	T	not done	(1,2,7)
W88	Southern California	1990	not done	not done	F	not done	(1,2,7)
395	Montreal, Canada	1999-2003	not done	2112-3383	unique	not done	(7,11)
976	Montreal, Canada	1999-2003	not done	2112-5383	unique	not done	(7,11)
2049	Montreal, Canada	1999-2003	not done	2112-5383	not done	not done	(7,11)
03240	Montreal, Canada	1999-2003	not done	2112-3383	not interpretable	12	(7,11)

4840	Montreal, Canada	1999-2003	not done	2112-5383	O	not done	(7,11)
4904	Montreal, Canada	1999-2003	CCGGCGG (2)	2112-3383	W	10	(7,11)
8422	Montreal, Canada	1999-2003	CCGGCGG (2)	2112-5383	N	not done	(7,11)
13373	Montreal, Canada	1999-2003	GGGGGCG(17)	211n-1n8n	not done	not done	(7,11)
28132	Montreal, Canada	1999-2003	CGGCGGG (3)	1111-5282	I	12	(7,11)
53198	Montreal, Canada	1999-2003	not done	2112-5383	N	not done	(7,11)
57413	Montreal, Canada	1999-2003	not done	not done	N	not done	(7,11)
62490	Montreal, Canada	1999-2003	not done	2112-5383	unique	not done	(7,11)
63787	Montreal, Canada	1999-2003	not done	2112-5383	V	not done	(7,11)
74301	Montreal, Canada	1999-2003	not done	2112-5383	not done	not done	(7,11)
76102	Montreal, Canada	1999-2003	not done	2112-5383	O	not done	(7,11)
82307	Montreal, Canada	1999-2003	not done	2112-5383	N	not done	(7,11)
98838	Montreal, Canada	199-2003	not done	not done	V	not done	(7,11)
NET1 (NLA009601988)	Netherlands	1996	not done	2211-2283	not done	not done	(7,8)
NET10 (NLA009600097)	Netherlands	1995	not done	1111-4282	I	not done	(7,8)
NET12 (NLA009601343)	Netherlands	1996	not done	1211-2483	not done	not done	(7,8)
NET13 (NLA009601477)	Netherlands	1996	not done	1211-4282	F	not done	(7,8)
NET14 (NLA009600753)	Netherlands	1996	not done	1111-2282	E	not done	(7,8)
NET15 (NLA009600902)	Netherlands	1996	not done	1111-2282	E	not done	(7,8)
NET16 (NLA009600217)	Netherlands	1995	not done	2211-2283	not done	not done	(7,8)
NET17 (NLA009600751)	Netherlands	1996	CCAGCGA (1)	1121-3282	C	not done	(7,8)
NET19 (NLA009601234)	Netherlands	1996	CCAGCGA (1)	1121-2282	T	not done	(7,8)
NET2 (NLA009600730)	Netherlands	1996	not done	2111-2283	unique	not done	(7,8)
NET20 (NLA009601375)	Netherlands	1996	CGGCGGG (3)	1111-4282	unique	not done	(7,8)
NET21 (NLA009601442)	Netherlands	1996	not done	1111-2282	E	not done	(7,8)
NET22 (NLA009601499)	Netherlands	1996	CCAGCGA (1)	1221-3282	C	not done	(7,8)
NET23 (NLA009601554)	Netherlands	1996	CCGGCGG (2)	2112-3383	W	not done	(7,8)
NET24 (NLA009601696)	Netherlands	1996	not done	not done	I	not done	(7,8)
NET26 (NLA009600192)	Netherlands	1996	not done	1111-4282	E	not done	(7,8)
NET27 (NLA009600433)	Netherlands	1996	not done	1111-4282	I	not done	(7,8)
NET28 (NLA009601594)	Netherlands	1996	not done	2112-3383	not interpretable	not done	(7,8)
NET29 (NLA009600281)	Netherlands	1995	not done	1111-4282	E	not done	(7,8)
NET30 (NLA009600514)	Netherlands	1996	not done	1111-4282	I	not done	(7,8)
NET31 (NLA009600877)	Netherlands	1996	not done	1111-4282	E	not done	(7,8)
NET32 (NLA009601344)	Netherlands	1996	not done	1211-4282	not done	not done	(7,8)
NET33 (NLA009601739)	Netherlands	1996	not done	2211-2283	not done	not done	(7,8)
NET38 (NLA009601391)	Netherlands	1996	CCAGCGA (1)	1111-2282	E	not done	(7,8)
NET39 (NLA009601393)	Netherlands	1996	CCAGCGA (1)	1111-2282	E	not done	(7,8)
NET4 (NLA009600239)	Netherlands	1996	not done	1111-2282	F	not done	(7,8)
NET40 (NLA009601542)	Netherlands	1996	not done	1111-2282	E	not done	(7,8)
NET41 (NLA009601734)	Netherlands	1996	not done	1111-2282	E	not done	(7,8)
NET42 (NLA009601158)	Netherlands	1996	CGGCGGG (3)	1111-4282	I	not done	(7,8)
NET43 (NLA009601794)	Netherlands	1996	not done	1111-4282	I	not done	(7,8)
NET46 (NLA009600403)	Netherlands	1996	not done	1111-2282	E	not done	(7,8)
NET47 (NLA009600841)	Netherlands	1996	CGGCGGG (3)	1111-5282	I	not done	(7,8)
NET48 (NLA009600842)	Netherlands	1996	not done	1111-4282	I	not done	(7,8)
NET49 (NLA009601198)	Netherlands	1996	not done	not done	E	not done	(7,8)
NET52 (NLA009601544)	Netherlands	1995	not done	1211-2483	not done	not done	(7,8)
NET53 (NLA009601738)	Netherlands	1996	not done	1112-2282	F	not done	(7,8)
NET54 (NLA009601816)	Netherlands	1995	not done	not done	I	not done	(7,8)
NET55 (NLA009601888)	Netherlands	1996	not done	1111-2282	E	not done	(7,8)
NET56 (NLA009601899)	Netherlands	1996	CCAGCGA (1)	1121-3282	C	not done	(7,8)
NET57 (NLA009601549)	Netherlands	1996	not done	1111-2282	F	not done	(7,8)
NET58 (NLA009600206)	Netherlands	1995	not done	1111-4282	I	not done	(7,8)
NET59 (NLA009600436)	Netherlands	1996	not done	1111-4282	I	not done	(7,8)
NET6 (NLA009600846)	Netherlands	1996	not done	1111-7282	not done	not done	(7,8)
NET61 (NLA009601392)	Netherlands	1996	CCAGCGA (1)	1111-2282	F	not done	(7,8)
NET62 (NLA009601543)	Netherlands	1996	not done	1111-2282	F	not done	(7,8)
NET63 (NLA009600536)	Netherlands	1996	CGGCGGG (3)	1111-4282	F	not done	(7,8)
NET67 (NLA009600205)	Netherlands	1995	CGGCGGG (3)	1111-4282	I	not done	(7,8)

NET68 (NLA009602012)	Netherlands	1995	not done	1111-5282	I	not done	(7,8)
NET70 (NLA009600511)	Netherlands	1996	not done	1111-3282	F	not done	(7,8)
NET71 (NLA009600531)	Netherlands	1996	CCAGCGA (1)	1111-2282	F	not done	(7,8)
NET75 (NLA009600098)	Netherlands	1995	not done	1111-4282	I	not done	(7,8)
NET76 (NLA009601555)	Netherlands	1996	CCAGCGA (1)	1111-2282	F	not done	(7,8)
NET8 (NLA009601245)	Netherlands	1994	not done	1111-4282	I	not done	(7,8)
NET80 (NLA009600224)	Netherlands	1995	not done	2112-5382	not done	not done	(7,8)
HMC02	Seattle, U.S.A.	1999-2001	CCAGCGA (1)	2112-3383	J	not interpretable	(2,3,7,9)
HMC05	Seattle, U.S.A.	1999-2001	CCAGCGA (1)	1221-2282	B	not done	(2,3,7,9)
HMC07	Seattle, U.S.A.	1999-2001	not done	not done	B	not done	(2,3,7,9)
HMC08	Seattle, U.S.A.	1999-2001	not done	1221-4282	B	7	(2,3,7,9)
HMC10	Seattle, U.S.A.	1999-2001	CCAGCGA (1)	1221-2282	unique	not done	(2,3,7,9)
HMC19	Seattle, U.S.A.	1999-2001	CCAGCGA (1)	1121-3282	P	not done	(2,3,7,9)
HMC20	Seattle, U.S.A.	1999-2001	CCGGCGG (2)	2212-3383	unique	not done	(2,3,7,9)
HMC22	Seattle, U.S.A.	1999-2001	not done	2211-3283	not done	not done	(2,3,7,9)
HMC24	Seattle, U.S.A.	1999-2001	CCAGCGA (1)	1221-4282	B	4	(2,3,7,9)
HMC31	Seattle, U.S.A.	1999-2001	not done	2112-1483	not done	not done	(2,3,7,9)
HMC33	Seattle, U.S.A.	1999-2001	not done	2112-3383	not interpretable	not done	(2,3,7,9)
HMC34	Seattle, U.S.A.	1999-2001	not done	not done	unique	not done	(2,3,7,9)
HMC36	Seattle, U.S.A.	1999-2001	not done	2112-2483	A	not done	(2,3,7,9)
HMC45	Seattle, U.S.A.	1999-2001	not done	2112-3383	not interpretable	not done	(2,3,7,9)
HMC46	Seattle, U.S.A.	1999-2001	not done	2112-1483	not done	not done	(2,3,7,9)
101	Southern California	1983	CCAGCGA (1)	1221-5292	B	not done	(2,7)
102	Southern California	1983	CCGGCGG (2)	2212-5483	unique	not done	(2,7)
103	Southern California	1983	not done	not done	B	not done	(2,7)
104	Southern California	1983	CCAGCGA (1)	1221-5292	B	not done	(2,7)
105	Southern California	1983	CCAGCGA (1)	1221-4282	unique	8	(2,7)
107	Southern California	1983	CCGGCGG (2)	2112-3383	unique	not done	(2,7)
108	Southern California	1983	CCGGCGG (2)	2112-3383	J	10	(2,7)
109	Southern California	1983	not done	not done	unique	not done	(2,7)
110	Southern California	1983	CCGGCGG (2)	2212-5483	unique	not done	(2,7)
111	Southern California	1983	not done	not done	unique	not done	(2,7)
113	Southern California	1983	not done	2212-5283	not done	not done	(2,7)
501	Southern California	1985	not done	not done	unique	not done	(2,7)
502	Southern California	1985	CCAGCGA (1)	1221-3282	B	not done	(2,7)
503	Southern California	1985	CCAGCGA (1)	1221-5292	B	not done	(2,7)
504	Southern California	1985	not done	not done	unique	not done	(2,7)
505	Southern California	1985	not done	not done	unique	not done	(2,7)

*hsp65 type (code) = Concatenation of bases present at hsp65 633, 1128, 1218, 1269, 1272, 1488 and 1536 and hsp65 code #s correspond to those presented in Turenne *et al* 2006 (15).

†LSP-MVR type = large-sequence polymorphism/MIRU/VNTR analysis; digits of type represent loci HSDR, DA2/LSP2, DA7/LSP7, DEL11/LSP5, X3, 25, 32, 47 (7,14); "1" is used to indicate the absence of the large-sequence polymorphism, which corresponds to product lengths of 343bp (for HSDR), 834bp (for DA2/LSP2), 329bp (for DA7/LSP7) and 728bp (for DEL11/LSP5). "2" is used to indicate the presence of the large-sequence polymorphism, which corresponds to product lengths of 480bp (for HSDR), 954bp (for DA2/LSP2), 181bp (for DA7/LSP7) and 794bp (for DEL11/LSP5) (7); digits for MIRU/VNTR type represent the number of repeats inferred from the product length at that locus, based on product sizes and repeat numbers from Thibault *et al* (14) and personal correspondence with V. Thibault: 196bp when 2 repeats of 53bp are present [X3], 350bp when 3 repeats of 58bp are present [25], 298bp when 8 repeats of 18bp are present [32] and 217bp when 3 repeats of 35bp are present [47]

‡rep-PCR type = repetitive-sequence-based PCR type; results for isolates in this analysis were combined with results for all MAC isolates in the SBRI rep-PCR library using the Diversilab software, and letter type designations were assigned to branches of the dendrogram with at least 92% average similarity

§PFGE pattern = pulsed-field electrophoresis results that are indistinguishable by the criteria of Tenover *et al* 1995 (13).

Supplementary Table S3. Environmental strains used in LSP-MVR analysis*

Name or number	Place of isolation	Time of isolation	LSP-MVR type	Citation(s)
E1508-3	Boston, U.S.A.	1990s	2112-5383	(18)
E5502-1	Hanover, U.S.A.	1990s	2122-3383	(18)
Crud-3	Denver, U.S.A.	2000s	1122-3383	(5)
Hot-2	Denver, U.S.A.	2000s	2122-3382	(5)
CA2	Southern California	1997	1121-2282	(6,19)
CA3	Southern California	1997	2112-3383	(6,19)
CA4	Southern California	1997	2112-3383	(6,19)
CA5	Southern California	1997	2112-3383	(6,19)
CA7	Southern California	1997	2112-3383	(6,19)
D10CC5tlsp	Southern California	1991	2112-3383	(6,19)
D10CN4W	Southern California	1991	2112-3383	(6,19)
D20BN1N	Southern California	1992	2112-3383	(6,19)
D20DN8N	Southern California	1992	2112-3383	(6,19)
D41BC8N	Southern California	1992	2112-3383	(6,19)
D41CO1	Southern California	1992	2112-3383	(6,19)
D41EC8N	Southern California	1992	2112-3383	(6,19)
D41EO8N	Southern California	1992	2112-3383	(6,19)
D55AC8N	Southern California	1992	2112-3383	(6,19)
D55AO1	Southern California	1992	2112-3383	(6,19)
F100	Southern California	1997	1211-4282	(6,19)
F103B	Southern California	1997	1121-4282	(6,19)
F72B	Southern California	1997	1121-4282	(6,19)
F72C	Southern California	1997	1121-4282	(6,19)
F7A	Southern California	1997	1121-4282	(6,19)
F86	Southern California	1997	1121-4282	(6,19)
H10BO1	Southern California	1992	2112-3383	(6,19)
H14FC8N	Southern California	1993	2212-3383	(6,19)
HO4AC5	Southern California	1991	2112-3383	(6,19)
HO5DN5	Southern California	1991	2112-3383	(6,19)
HO8FC8N	Southern California	1992	2212-3383	(6,19)
W292	Southern California	1992	2112-3383	(6,19)
P50073	U.S. West Coast City	2001	2112-3483	(6)
P50075	U.S. West Coast City	2001	2112-3383	(6)
P50079	U.S. West Coast City	2001	2112-3383	(6)
P50091	U.S. West Coast City	2001	2212-3383	(6)
P50099	U.S. West Coast City	2001	2112-3383	(6)
P50103	U.S. West Coast City	2001	2112-3383	(6)
P50112	U.S. West Coast City	2001	2212-3383	(6)
P50135	U.S. West Coast City	2001	2112-3383	(6)
P50139	U.S. West Coast City	2001	2112-3383	(6)
P50171	U.S. West Coast City	2002	2212-3383	(6)
P50181	U.S. West Coast City	2002	2112-3383	(6)
P50183	U.S. West Coast City	2002	2112-3383	(6)
P50187	U.S. West Coast City	2002	2212-3383	(6)
P50195	U.S. West Coast City	2002	2112-3383	(6)
P50199	U.S. West Coast City	2002	2112-3383	(6)
P50207	U.S. West Coast City	2002	2212-3383	(6)
P50211	U.S. West Coast City	2002	2112-3383	(6)
A001-2-1-2	U.S. Potting Soil	2000s	1121-4282	(4)
A001-2-5-1	U.S. Potting Soil	2000s	1121-4282	(4)
A002-1-1-1	U.S. Potting Soil	2000s	1121-4282	(4)
A019-C-4-1	U.S. Potting Soil	2000s	2112-3273	(4)
W978	Westchester, U.S.A.	1980s	1112-5282	(12)

* LSP-MVR type = large-sequence polymorphism/MIRU/VNTR analysis; digits of type represent loci HSDR, DA2/LSP2, DA7/LSP7, DEL11/LSP5, X3, 25, 32, 47 (7,14); "1" is used to indicate the absence of the large-sequence polymorphism, which corresponds to product lengths of 343bp (for HSDR), 834bp (for DA2/LSP2), 329bp (for DA7/LSP7) and 728bp (for DEL11/LSP5). "2" is used to indicate the presence of the large-sequence polymorphism, which corresponds to product lengths of 480bp (for HSDR), 954bp (for DA2/LSP2), 181bp (for DA7/LSP7) and 794bp (for DEL11/LSP5) (7); digits for MIRU/VNTR type represent the number of repeats inferred from the product length at that locus, based on product sizes and repeat numbers from Thibault et al (14) and personal correspondence with V. Thibault: 196bp when 2 repeats of 53bp are present [X3], 350bp when 3 repeats of 58bp are present [25], 298bp when 8 repeats of 18bp are present [32] and 217bp when 3 repeats of 35bp are present [47].

Supplemental Information S4. Mycobacterial genomic DNA preparation protocol modified from ArchivePure DNA Cell/Tissue Kit from 5'

1. Pre-heat hybridization oven or waterbath to 55°C.
2. Transfer mobio garnet beads to 2 mL screw-top tube.
3. Add 800 µL lysis solution to each tube.
4. Add appx 10 µL loop of cells (or 1-5 OD) to each tube.
5. Bead-beat on vortex adapter for 30 min @ high power.
6. Spin at max for 1 min.
7. Transfer lysate supernatant to a new 2 mL screw-top tube.
8. Add 15 µL of Proteinase K (20 mg/ml), mix by inverting 25 times.
9. Incubate lysate at 55°C 30 min or overnight.
10. Let cool 10min, then add 3 µL Rnase A Solution (20 mg/mL) to the cell lysate. Incubate at 37°C for 30-60 min.
11. Cool sample to room temperature (~10 min), then add 0.3 mL Protein Precipitation Solution to the lysate.
12. Vortex vigorously at high speed for ~20 seconds to mix, then cool on wet ice for 5-30 minutes.
13. Spin at max for 5 minutes to pellet protein precipitate.
14. Pipette supe (go for purity over yield) to a clean 2mL (snap-top) tube
15. Precipitate (ppt) DNA by adding 800 µL isopropanol.
16. Gently invert 50 times, and let sit at room temp 5-30 min to ppt DNA; a white stringy ppt may emerge (DNA).
17. Spin at max for 5 min to pellet precipitated DNA.
18. Wash w/ 1ml 70% ethanol, vortex, then spin at max speed for 1 min. You should still see a pellet.
19. Suction or gently dump off ethanol. Pellet may be loose, take care not to lose it.
20. Dry by leaving upside down 1 hour, or by low-mid temperature speedvac for 20 minutes.
21. Make sure all ethanol has evaporated.
22. Resuspend in 100 µl TE. Vortexing/heat/more TE may be necessary if yield is high.
23. Analyze by spectrophotometry to determine purity and yield of gDNA.

Supplementary Table S5. LSP-MVR* distribution among clinical isolates at seven geographical locations, including GDCs*, local clinical clusters and strains unique among clinical isolates

Group†	LSP-MVR Genotype	Sao Paulo, Brazil	Southern California	Seattle, U.S.A.	Montreal, Canada	Hanover, U.S.A.	Boston, U.S.A.	Netherlands	Total
GDC1	2112-3383	1	4	3	3			2	13
GDC2	1111-5282	1			1			2	4
GDC3	1121-3282		1	1				2	4
GDC4	1211-2483		5					2	7
GDC5	1211-4282		3					2	5
GDC6	1221-3282		2					1	3
GDC7	1221-4282		3	2					5
GDC8	2211-2283		2					3	5
GDC9	2212-3383	1		1					2
GDC10	1121-2282		1					1	2
GDC11	1121-4282		1			1			2
GDC12	2112-5383				10	1			11
Local 1	1111-2282							15	15
Local 2	1111-4282							16	16
Local 3	1221-2282			2					2
Local 4	1221-5292		3						3
Local 5	2112-1483			2					2
Local 6	2212-5483		2						2
Unique	1111-3282							1	1
	1111-7282							1	1
	1112-2282							1	1
	1112-4282		1						1
	1121-5282	1							1
	1211-2283		1						1
	1212-4282		1						1
	2111-2283							1	1
	2111-2483	1							1
	2112-1183	1							1
	2112-1682	1							1
	2112-1683	1							1
	2112-2483			1					1
	2112-3282						1		1
	2112-5382							1	1
	2112-5483		1						1
	2211-2282		1						1
	2211-2383					1			1
	2211-3283			1					1
	2211-3383		1						1
	2212-3283					1			1

	2212-5283		1					1	
	2221-3482		1					1	
	2221-4282		1					1	
Total typed	-	8	36	13	14	4	1	51	127
In GDCs	-	3 (38%)	22 (61%)	7 (54%)	14 (100%)	2 (50%)	0	15 (29%)	63 (50%)
In Local Clusters	-	0	5 (14%)	4 (31%)	0	0	0	31 (61%)	40 (31%)
Unique	-	5 (62%)	9 (25%)	2 (15%)	0	2 (50%)	1	5 (10%)	24 (19%)

*LSP-MVR = large-sequence polymorphism/MIRU-VNTR analysis; GDCs = geographically dispersed clinical types

† GDCs were defined by the same LSP-MVR type being found for at least one clinical isolate at each of two or more geographical locations. Local clinical clusters were defined by finding the same LSP-MVR type for two or more clinical isolates at a single geographic location, but not being found among clinical isolates from other geographic locations. Unique types were found for a single isolate in the entire set of clinical genotypes. The presence or absence of the LSP-MVR type among environmental isolates (at the same or a different location) did not affect designation as GDC, local clinical cluster, or unique among clinical isolates.

GDCs were defined by the same LSP-MVR type being found for at least one clinical isolate at each of two or more geographical locations. Local clinical clusters were defined by finding the same LSP-MVR type for two or more clinical isolates at a single geographic location, but not being found among clinical isolates from other geographic locations. Unique types were found for a single isolate in the entire set of clinical genotypes. The presence or absence of the LSP-MVR type among environmental isolates (at the same or a different location) did not affect designation as GDC, local clinical cluster, or unique among clinical isolates.

Supplementary Table S6. 8-locus MIRU-VNTR types of some isolates sharing GDC* genotypes

Name	LSP-MVR type †	Isolate ID †	Site	8-locus MIRU-VNTR type †
GDC1	2112-3383	4904	Montreal	3333-12p8
		HMC02	Seattle	3333-12(≥9)8
		107	Southern California	3333-1298
		108	Southern California	3333-1298
		W355	Southern California	n333-12p8
GDC2	1111-5282	28132	Montreal	2522-1218
		BR80	Sao Paolo	2522-1228
GDC3	1121-3282	HMC19	Seattle	2322-1228
		NET17	The Netherlands	2322-1228
		NET56	The Netherlands	2322-1228
		W129	Southern California	2322-1218
GDC7	1221-4282	W214	Southern California	2422-1228
		HMC08	Seattle	2422-1218
		HMC24	Seattle	2422-1218
		W216-8	Southern California	2422-1218
		105	Southern California	2422-1228

* GDCs = geographically dispersed clinical strains, GDCs were defined by the same LSP-MVR type being found for at least one clinical isolate at each of two or more geographical locations

† LSP-MVR = large-sequence polymorphism – MIRU/VNTR analysis; MIRU-VNTR = mycobacterial interspersed repetitive unit – variable-number tandem repeat; LSP-MVR and 8-locus MIRU-VNTR types are given for all members of LSP-MVR-defined GDCs for which genomic DNA was randomly selected for method evaluation or archived cultures from more than one location were available for PFGE analysis; identity by 8-locus MIR of isolates from distant geographic locations could only be tested for .

Supplementary Table S7. LSP-MVR* distribution among environmental isolates from at least six sites.

Group†	LSP-MVR†	Southern California	U.S. West Coast Metropolitan Area	U.S. Potting Soil‡	Denver, U.S.A.	Westchester, U.S.A.	Hanover, U.S.A.	Boston, U.S.A.	Total
GDC1	2112-3383	18	15						33
GDC5	1211-4282	1							1
GDC9	2212-3383	2							2
GDC10	1121-2282	1							1
GDC11	1121-4282	5		3					8
GDC12	2112-5383							1	1
Unique	1112-5282					1			1
	1122-3382				1				1
	2112-3273			1					1
	2112-3483		1						1
	2122-3382				1				1
	2122-3383						1		1
Total typed	-	27	16	4	2	1	1	1	52
In GDCs	-	27 (100%)	15 (94%)	3 (75%)	0	0	0	1 (100%)	46 (88%)

*LSP-MVR = large-sequence polymorphism – MIRU/VNTR analysis

† HSDR, DA2/LSP2, DA7/LSP7, DEL11/LSP5, X3, 25, 32, 47 (7,14); “1” is used to indicate the absence of the large-sequence polymorphism, which corresponds to product lengths of 343bp (for HSDR), 834bp (for DA2/LSP2), 329bp (for DA7/LSP7) and 728bp (for DEL11/LSP5); “2” is used to indicate the presence of the large-sequence polymorphism, which corresponds to product lengths of 480bp (for HSDR), 954bp (for DA2/LSP2), 181bp (for DA7/LSP7) and 794bp (for DEL11/LSP5) (7); digits for MVR type represent the number of repeats inferred from the product length at that locus, based on product sizes and repeat numbers from Thibault *et al* 2007 (14) and personal correspondence with V. Thibault: 196bp when 2 repeats of 53bp are present [X3], 350bp when 3 repeats of 58bp are present [25], 298bp when 8 repeats of 18bp are present [32] and 217bp when 3 repeats of 35bp are present [47].

‡ For the U.S. Potting Soil collection, individuals with nontuberculous mycobacterial lung disease seen at a reference center were asked to send samples of potting soil from their homes for analysis; these individuals lived throughout the U.S.

Supplementary information reference list

1. **Aronson, T., A. Holtzman, N. Glover, M. Boian, S. Froman, O. G. W. Berlin, H. Hill, and G. Stelma.** 1999. Comparison of large restriction fragments of *Mycobacterium avium* isolates recovered from AIDS and non-AIDS patients with those of isolates from potable water. *J Clin Microbiol* **37**:1008-1012.
2. **Cangelosi GA, Freeman R, Lewis KN, Livingston-Rosanoff D, Milan SJ, and Goldberg S.** 2004. Evaluation of a high throughput rep-PCR system for DNA fingerprinting of *Mycobacterium tuberculosis* and *Mycobacterium avium* complex. *J Clin Microbiol* **42**:2685-2693.
3. **Cangelosi, G. A., C. O. Palermo, J. P. Laurent, A. M. Hamlin, and W. H. Brabant.** 1999. Colony morphotypes on Congo red agar segregate along species and drug susceptibility lines in the *Mycobacterium avium*--*intracellulare* complex. *Microbiology* **145**:1317-1324.
4. **De Groote, M. A., N. R. Pace, K. Fulton, and J. O. Falkinham, III.** 2006. Relationships between *Mycobacterium* Isolates from Patients with Pulmonary *Mycobacterial* Infection and Potting Soils. *Appl.Environ.Microbiol.* **72**:7602-7606.
5. **Falkinham, J. O., M. D. Iseman, P. de Haas, and D. van Soolingen.** 2008. *Mycobacterium avium* in a shower linked to pulmonary disease. *Journal of Water & Health* **6**:209-213. doi:Article.
6. **Hilborn, E. D., T. C. Covert, M. A. Yakus, S. I. Harris, S. F. Donnelly, E. W. Rice, S. Toney, S. A. Bailey, and G. N. Stelma, Jr.** 2006. Persistence of Nontuberculous *Mycobacteria* in a Drinking Water System after Addition of Filtration Treatment. *Appl.Environ.Microbiol.* **72**:5864-5869.
7. **Horan, K. L., R. Freeman, K. Weigel, M. Semret, S. Pfaller, T. C. Covert, D. van Soolingen, S. C. Leao, M. A. Behr, and G. A. Cangelosi.** 2006. Isolation of the Genome Sequence Strain *Mycobacterium avium* 104 from Multiple Patients over a 17-Year Period. *J.Clin.Microbiol.* **44**:783-789.
8. **Komijn, R. E., P. E. áW. de Haas, M. M. Schneider, T. Eger, J. H. Nieuwenhuijs, R. J. van den Hoek, D. Bakker, F. G. Zijd Erveld, and D. van Soolingen.** 1999. Prevalence of *Mycobacterium avium* in Slaughter Pigs in The Netherlands and Comparison of IS1245 Restriction Fragment Length Polymorphism Patterns of Porcine and Human Isolates. *J.Clin.Microbiol.* **37**:1254-1259.
9. **Mukherjee S, Petrofsky M, Yaraei K, Bermudez LE, and Cangelosi GA.** 2001. The white morphotype of *Mycobacterium avium-intracellulare* is common in infected humans and virulent in infection models. *J Infect Dis* **184**:1480-1484.
10. **Oliveira, R. S., M. P. Sircili, E. M. D. Oliveira, S. C. Balian, J. S. Ferreira-Neto, and S. C. Le+úo.** 2003. Identification of *Mycobacterium avium* Genotypes with Distinctive Traits by Combination of IS1245-Based Restriction Fragment Length Polymorphism and Restriction Analysis of hsp65. *J.Clin.Microbiol.* **41**:44-49.
11. **Semret M, Zhai G, Mostowy S, Alexander D, Cangelosi GA, Cousins D, van Soolingen D, and Behr M.** 2004. Extensive Genomic Polymorphism within *Mycobacterium avium*. *J.Bacteriol.* **186**:6332-6334.
12. **Stormer RS and Falkinham JO 3rd.** 1989. Differences in antimicrobial susceptibility of pigmented and unpigmented colonial variants of *Mycobacterium avium*. *J Clin Microbiol* **27**:2459-2456.

13. **Tenover, F. C., R. D. Arbeit, R. V. Goering, P. A. Mickelsen, B. E. Murray, D. H. Persing, and B. Swaminathan.** 1995. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J.Clin.Microbiol.* **33**:2233-2239.
14. **Thibault, V. C., M. Grayon, M. L. Boschioli, C. Hubbans, P. Overduin, K. Stevenson, M. C. Gutierrez, P. Supply, and F. Biet.** 2007. New variable number tandem repeat markers for typing *M. avium* subsp. *paratuberculosis* and *M. avium* strains: comparison with IS900 RFLP and IS1245 RFLP typing. *J.Clin.Microbiol.JCM.*
15. **Turenne, C. Y., M. Semret, D. V. Cousins, D. M. Collins, and M. A. Behr.** 2006. Sequencing of *hsp65* Distinguishes among Subsets of the *Mycobacterium avium* Complex. *J.Clin.Microbiol.* **44**:433-440.
16. **Von Reyn, C. F., R. D. Arbeit, C. R. Horsburgh, M. A. Ristola, R. D. Waddell, S. M. Tvaroha, M. Samore, L. R. Hirschhorn, J. Lumio, A. D. Lein, and M. R. Grove.** 2002. Sources of Disseminated *Mycobacterium avium* Infection in AIDS. *Journal of Infection* **44**:166-170.
17. **Von Reyn, C. F., J. N. Maslow, T. W. Barber, J. O. I. Falkinham, and R. D. Arbeit.** 1994. Persistent colonisation of potable water as a source of *Mycobacterium avium* infection in AIDS. *Lancet* **343**:1137-1141.
18. **Von Reyn, C. F., R. D. Waddell, T. Eaton, R. D. Arbeit, and et al.** 1993. Isolation of *Mycobacterium avium* complex from water in the United States, Finland, Zaire, and Kenya. *J Clin Microbiol* **31**:3227-3230.
19. **Yoder, S., C. Argueta, A. Holtzman, T. Aronson, O. G. Berlin, P. Tomasek, N. Glover, S. Froman, and G. Stelma, Jr.** 1999. PCR Comparison of *Mycobacterium avium* Isolates Obtained from Patients and Foods. *Appl.Environ.Microbiol.* **65**:2650-2653.