Mycobacterium riyadhense clinical isolates provide insights into ancestry and adaptive evolution in tuberculosis

Qingtian Guan[1], Musa Garbati[2,&], Sara Mfarrej[1], Talal AlMutairi[2], Alicia Smyth[3&&], Albel Singh[4&&], Shamsudeen Fagbo[5&&], John A. Browne[3], Muhammad Amin urRahman[2], Alya Alruwaili[2], Anwar Hoosen[2], Chie Nakajima[6,7], Yasuhiko Suzuki[6,7], Apoorva Bhatt[4], Stephen V. Gordon[3,6], Faisal AlAsmari[2], A. Pain[1,6*]

[1]Pathogen Genomics Laboratory, BESE Division, King Abdullah University of Science and Technology (KAUST), Thuwal-Jeddah, Kingdom of Saudi Arabia; [2]King Fahad Medical City (KFMC), Riyadh, Saudi Arabia; [3]UCD School of Veterinary Medicine, University College Dublin, Dublin, D04 W6F6, Ireland; [4]Institute of Microbiology and Infection, School of Biosciences, University of Birmingham, Edgbaston, Birmingham, UK;[5]One Health Unit, Executive Directorate for Surveillance and Response, National Center for Disease Prevention and Control, Riyadh, Saudi Arabia; [6]Global Institution for Collaborative Research and Education, Hokkaido University, Kita 20 Nishi 10, Kita-ku,

17 *Sapporo, Japan;*[7]*Research Center for Zoonosis Control, Hokkaido University, Kita 20*

18 *Nishi 10, Kita-ku, Sapporo, Japan.*

19

20

21 Key words: *Mycobacterium riyadhense*; Evolution of *Mycobacterium tuberculosis*

22 complex; Phylogeny; Virulence factors; Diagnostic marker.

23 [&]Current address: Infectious Diseases Unit, Department of Medicine, University of

24 Maiduguri, Nigeria

25 [&&]These authors contributed equally to the work

26 *****For correspondence

27 Arnab Pain (arnab.pain@kaust.edu.sa)

28

29

30

31

32

33

34

Abstract

Mycobacterium riyadhense is a novel non-tuberculous mycobacterium (NTM) that is increasingly isolated from clinical cases with tuberculosis (TB)-like symptoms in humans in various parts of the world. To elucidate the genomic hallmarks that define the relationship between *M. riyadhense* and other mycobacterial species, including members of the *Mycobacterium tuberculosis* complex (MTBC), eight clinical isolates of *M. riyadhense* were analysed. *M. riyadhense* shares a large number of conserved orthologues with the MTBC and contains linear and circular plasmids carrying type IV and type VII secretion systems. We provide evidence that *M. riyadhense* is likely the 'missing link' in the evolution of *M. tuberculosis*, sharing a common ancestor with members of the MTBC that has evolved through a process of genome reduction, expansion of toxin/antitoxin systems, and ultimately host adaptation. We conclude that *M. riyadhense* is an emerging mycobacterial pathogen that can serve as an experimental model to study the evolution and pathogenesis of tubercle bacilli.

Introduction

Non-tuberculous mycobacteria (NTMs), including *Mycobacterium riyadhense* (MR), are ubiquitous, naturally occurring environmental bacteria commonly found in water and soil (Falkinham, 2013)(J. Van Ingen et al., 2009). A wide range of animal and environmental

55 sources (aquaria, swimming pools) act as reservoirs for NTMs, and several human
56 disease outbreaks caused by exposure to environmental NTMs have been described (Ding
57 et al., 2005)(Panwalker and Fuhse, 1986)(Carbonne et al., 2009)(Singh et al., 2018). With
58 the ability to cause infections in both immunocompromised (Garbati and Hakawi, 2014)
59 and immunocompetent (Choi et al., 2012) individuals, *M. riyadhense* has positioned
60 itself as a clinically important pulmonary pathogen since its discovery in 2009 (van Ingen
61 et al., 2009) as a cause of pulmonary infection. The clinical and radiologic characteristics
62 of pulmonary infection caused by *M. riyadhense* are indistinguishable from those caused
63 by *Mycobacterium tuberculosis*, the most important human pathogen of the *M.*
64 *tuberculosis* complex (MTBC), which cause TB in humans and animals (Choi et al.,
65 2012)(van Ingen et al., 2009).

66 Similar to *M. tuberculosis*, *M. riyadhense* grows at 37°C and requires 2~3 weeks (van
67 Ingen et al., 2009) to form visible colonies on agar media. However, unlike *M.*
68 *tuberculosis*, which is a common worldwide pathogen transmitted directly from human to
69 human with no known environmental reservoirs (King et al., 2017), *M. riyadhense*
70 infections are rare and are transmitted to patients via contact with contaminated water
71 (King et al., 2017) and soil (Narendrula-Kotha and Nkongolo, 2017), with no evidence of
72 human-to-human transmission yet reported. Infections with *M. riyadhense* have been
73 reported in Asia and Europe including, Bahrain, South Korea, France, Italy and Germany
74 (Godreuil et al., 2012)(Van der Werf et al., 2014), (Van der Werf et al., 2014), (Choi et

75 al., 2012), although most of the recent cases originated in patients from Saudi Arabia
76 (Althawadi et al., 2017). Indeed, the very first case of acute *M. riyadhense* infection was
77 initially misdiagnosed as a case of *M. tuberculosis* infection in a Saudi hospital using
78 commercially available diagnostic tests (Tortoli et al., 2010).
79
80 It is postulated that *M. tuberculosis* evolved from a free-living environmental ancestor
81 into an obligate pathogen (Murty DS, 2014)(Jang et al., 2008). Our current knowledge
82 indicates that *Mycobacterium canettii* is the most closely related obligate pathogenic
83 species to the MTBC (Supply et al., 2013)(Van Soolingen et al., 1997). Infections with *M.*
84 *canettii* are rare and found solely in people from the Horn of Africa, with no
85 environmental reservoir described to date (Blouin et al., 2012).
86
87 Previous phylogenetic studies have suggested that *Mycobacterium kansasii* (Wang et al.,
88 2015), *Mycobacterium marinum* (Stinear et al., 2008), *Mycobacterium lacus* (Tortoli et
89 al., 2017), *Mycobacterium decipiens* (Brown-Elliott et al., 2018), *Mycobacterium*
90 *shinjukuense* (Saito et al., 2011) and *M. riyadhense* (Tortoli et al., 2017)(Fedrizzi et al.,
91 2017) are closely related to the free-living ancestor of the MTBC based on a single
92 marker gene (e.g., *hsp65*, 16S). In this evolutionary context, a comprehensive
93 understanding of the genomic hallmarks that drive the close phylogenomic relationship of
94 *M. riyadhense* with the MTBC and NTMs is of particular scientific interest.

95

96 In this study, we report the characterization of eight clinical isolates of *M. riyadhense*,
97 compared their genomes to members of the MTBC, and provided comprehensive insight
98 into the adaptive changes associated with the evolution of the MTBC from environmental
99 bacteria. We examined the lipid profiles of both rough and smooth variants of *M.*
100 *riyadhense*, compared them to those of other related mycobacteria, and analysed the
101 transcriptional response of immunity-related host genes in a murine macrophage infection
102 model with *M. riyadhense*, *M. kansasii* and *M. bovis* BCG.

103

104 Using genome information, we developed a simple PCR-based diagnostic test for the
105 rapid and accurate identification of *M. riyadhense* to minimize the risk of misdiagnosis in
106 a clinical setting.

107

108 Our analyses provide a comprehensive description of the hallmarks of *M. riyadhense* that
109 make it one of the closest known environmental relatives of the MTBC and that can serve
110 to illuminate studies into the evolution and pathogenesis in the MTBC.

111

112 **Methods and Materials**

113

114 **Clinical reports and bacterial strains**

115 Eight *M. riyadhense* strains were collected in Riyadh, Saudi Arabia, between June 2011
116 and March 2016 (Table 1) from patients with a positive culture for *M. riyadhense* isolated
117 from the microbiology laboratory at the King Fahad Medical City (KFMC) in Riyadh,
118 Saudi Arabia. The patient and sample data collected included demographic and clinical
119 characteristics, age, sex, clinical features at presentation, presence of comorbidities,
120 including HIV co-infection, antimicrobial susceptibility, initial and modified therapy,
121 where applicable; and treatment outcome. Once an isolate was suspected to be an NTM,
122 the samples were sent to the reference laboratory for full identification and antimicrobial
123 susceptibility testing. Radiographic and pathologic data were also collected.
124

125 **Culturing, DNA isolation and sequencing of bacteria**

126 The *M. riyadhense* strains were grown on Lowenstein Jensen (LJ) slants at 37 °C for two
127 weeks, DNA was extracted using a phenol-chloroform protocol (Belisle and Sonnenberg,
128 1998), and the quality was measured by Qubit. Twenty micrograms of
129 high-molecular-weight (HMW) DNA from the eight *M. riyadhense* strains was
130 sequenced using a PacBio RSII sequencer (Pacific Biosciences, Menlo Park, USA) with a
131 10 kb library. An NEBNext Ultra II DNA library preparation kit (New England BioLabs,
132 Massachusetts USA) was used to prepare libraries according to the manufacturer's
133 instructions, and sequences from each library were generated for all *M. riyadhense* strains
134 using the Illumina HiSeq 4000 platform (Illumina, San Diego, United States).

135 **Genome assembly and annotation**

136 The Illumina short reads were trimmed, and low-quality reads were removed by
137 Trimmomatic (Bolger et al., 2014). Eight consensus genomes based on each strain were
138 assembled into contigs with the PacBio long reads using the Canu assembler (Koren et al.,
139 2017). After assembly, the draft genome was then corrected with short Illumina reads
140 using Pilon (Walker et al., 2014) software. The circularity of assemblies was checked by
141 Gepard (Krumsiek et al., 2007), and assemblies were annotated by Prokka (Seemann,
142 2014). A circular map of the chromosome was compared with that of *M. tuberculosis*
143 H37Rv and visualized with BRIG (Alikhan et al., 2011). The genome of the
144 *M. riyadhense* MR226 strain was used as a high-quality reference in this study.

145

146 **Comparison of chromosomal and plasmid gene contents of *M. riyadhense* to those of**
147 **various *Mycobacterium* species**

148 DNA sequences of 152 mycobacterial species and 77 mycobacterial plasmids were
149 obtained from the NCBI genome database and independently annotated by Prokka
150 (Seemann, 2014). The predicted protein sequences from the chromosome and each of the
151 two plasmids (pMRLP01 and pMR01) of the *M. riyadhense* MR226 strain were then
152 compared with the annotated genes from the rest of the mycobacterial species using
153 orthoMCL (Li et al., 2003) with a 50% identity cutoff and an inflation parameter of 1.5.
154 The obtained orthologues were visualized with the heatmap package in R (Ihaka and

155 Gentleman, 1996). A focused OrthoMCL comparison was performed between (1) *M.*
156 *riyadhense*, *M. marinum*, *M. kansasii*, *M. szulgai* and *M. tuberculosis* and (2) *M.*
157 *riyadhense* and five species from the MTBC, namely, *M. tuberculosis*, *M. bovis*, *M.*
158 *canettii*, *M. mungi*, and *M. africanum*.
159

160 **SNP calling and phylogeny based on SNPs**

161 The corrected Illumina reads were mapped using BWA (Li and Durbin, 2009) on to the
162 *M. riyadhense* MR226 genome assembly. Picard tool (Broad Institute, 2016) was used to
163 clean SAM files, fix mate-pair information and mark duplicates. SNPs were called for
164 two iterations and filtered (QD<2.0, FS>60.0, SOQ>4.0, ReadPosRankSum<-8.0) with
165 Genome Analysis Toolkit (GATK) (Alkan et al., 2011). An alignment file was generated
166 by SVAMP (Naeem et al., 2014), and phylogeny was generated by RaxML (Stamatakis
167 et al., 2005) with the TVM model.
168

169 **Phylogeny of *M. riyadhense***

170 The AMPHORA2 (Wu and Scott, 2012) pipeline was used to identify protein sequences
171 from 31 conserved genes in the pangenome datasets and the *Mycobacterium* genomes
172 available at NCBI or JGI as at 1st January 2018. A total of 152 species were selected after
173 applying a filter of at least 10 marker genes to be detected in each of the assemblies.
174 *Nocardia abscessus* was used as the out-group. 12 genes, which have only one copy in

175 each of the species (i.e., genes *rplA*, *rplB*, *rplF*, *rplM*, *rplN*, *rplP*, *rplS*, *rplT*, *rpmA*, *rpsJ*,
176 *rpsM*, and *rpsS*), were concatenated, aligned and trimmed. The phylogenetic tree was
177 then constructed using RaxML with the Dayhoff model, which is calculated with a script
178 provided in AMPHORA2 (Wu and Scott, 2012).
179
180 A whole-genome phylogenetic tree was constructed with the MTBC species (*M.*
181 *tuberculosis*, *M. bovis*, *M. canettii*, *M. mungi*, *Mycobacterium orygis*, *M. africanum*) and
182 with *M. kansasii*, *M. marinum*, *M. shinjukuense*, *Mycobacterium leprae*,
183 *Mycobacterium smegmatis*, *Mycobacterium parmense*, *Mycobacterium avium* and
184 *Mycobacterium abscessus*. The one-to-one orthologues of each species were obtained
185 using OrthoMCL and concatenated, then aligned with Muscle (Edgar, 2004) and trimmed
186 with TrimAL (Capella-Gutiérrez et al., 2009). The concatenated sequences were
187 composed of 906 genes encoding 296,124 amino acids and were used to build a
188 phylogenetic tree with the LG+G+F model, which was selected by ModelGenerator. The
189 phylogenetic tree was generated by RaxML (Stamatakis et al., 2005).
190
191 **Toxin/antitoxin systems, *mce*/*mce*-associated genes and ESX systems in *M.***
192 ***riyadhense* and other mycobacteria**
193 One hundred fifty-eight toxin/antitoxin (T/A) proteins belonging to the VapBC, RelEF,
194 HigBA, MazEF, ParDE and UCAT families were downloaded from the NCBI protein

195 database. We identified the *M. tuberculosis* T/A orthologues from all of the 152 species
196 by OrthoMCL (Li et al., 2003), and the orthologue groups were also examined by Blast+
197 2.4.0. The same pipeline was also applied for the MCE family, PhoPR, PE/PPE,
198 PE-PGRS, and ESX1-5.
199
200 **Infection of the RAW 246.7 cell line with *M. riyadhense*, *M. kansasii* and *M. bovis***
201 **BCG Denmark**
202 The murine macrophage RAW264.7 cell line obtained from the American Type Culture
203 Collection (ATCC, Manassas, USA) was cultured in Dulbecco's modified Eagle's
204 medium (DMEM) (ThermoFisher Scientific, Waltham, USA) supplemented with 10%
205 FCS, streptomycin, and penicillin. *M. riyadhense*, *M. kansasii* (subtype I), and *M. bovis*
206 BCG Denmark strains were grown in Middlebrook 7H9 liquid medium after
207 single-colony isolation from LJ slants or 7H10 agar. 7H9 was supplemented with 10%
208 albumin, dextrose and catalase (ADC), while 7H10 was supplemented with oleic acid,
209 albumin, dextrose and catalase (OADC) in addition to 0.2% glycerol. Ready to use LJ
210 slants were provided by Saudi Prepared Media Laboratory (SPML, Riyadh, Saudi
211 Arabia).
212
213 Before the infection, all bacterial cultures were centrifuged at 1,000xg for 10 minutes.
214 The supernatant was discarded, and 10-15 3 mm glass beads were added to the pellet and

215 then vortexed for 1 minute to break up clumps. Six ml of DMEM was added to the pellet
216 and left to rest for 5 minutes. The upper 5 ml of the suspension was then removed to a
217 fresh 15 ml falcon tube, which was then centrifuged for an additional 3 minutes at 200xg
218 to remove remaining bacterial clumps. The supernatant was then taken and passaged
219 using a 26G hypodermal syringe approximately 15 times to further break up any clumps
220 of bacteria. The optical density of the culture was measured again before the infection
221 experiment.
222
223 RAW264.7 cells were seeded at 2 x 10^5 cells per well in 24-well flat-bottom tissue
224 culture plates 24 hours prior, reaching 5 x 10^5 cells per well. The DMEM over the cells
225 was removed, and the cells were washed once with phosphate-buffered saline (PBS). The
226 cells were infected by applying 1 ml of DMEM containing the prepared mycobacteria
227 from the former steps at an appropriate concentration to reach an MOI of 5:1, with
228 DMEM alone used for the control wells. The culture plates were then returned to the
229 incubator at 37°C with 5% CO_2 for 3 hours to allow for bacterial uptake by the
230 RAW264.7 cells. The supernatant was removed after 3 hours, and the infected cells were
231 washed with PBS to remove extracellular bacteria. Subsequently, the cells were
232 incubated in fresh DMEM with 10% FCS for 24 hours and 48 hours. For harvesting, 400
233 μl of TRIzol (ThermoFisher Scientific, Waltham, USA) was added to the wells at each
234 time point, and the adherent cells were scraped out and stored at -80°C for RNA

235 extraction. Each bacterial infection was performed in triplicate, in addition to the
236 non-infected controls. RNA was isolated from the samples using the Direct-zol™ RNA
237 Miniprep kit (Zymo Research, Irvine, USA) according to the manufacturer's instructions.
238
239 An Agilent RNA 6000 Nano kit was used to check the quality and quantity of the total
240 RNA. The NanoString murine nCounter Myeloid Innate Immunity Gene Expression
241 Panel (NanoString Technologies, USA) was used to assess transcript abundance across
242 infections and timepoints using the nCounter MAX Analysis System (NanoString
243 Technologies, Seattle, USA). The counts obtained were normalised using the nSolver™
244 Advanced Analysis plugin (NanoString Technologies, Seattle, USA) using the geNorm
245 algorithm, and differential gene (DE) expression was analysed using multivariate linear
246 regression in nSolver™ software, with 0.05 as the p-value cutoff.
247
248 **Thin layer chromatography (TLC) analysis of lipooligosaccharides (LOSs) in *M.***
249 ***riyadhense*, *M. kansasii* and *M. marinum***
250 For TLC analysis, mycobacterial strains were grown at 30°C (*M. marinum*) or 37°C
251 (*M. smegmatis*, *M. kansasii*, *M. riyadhense*) on LJ slants, and after sufficient incubation,
252 grown cells were collected and washed once with PBS. Apolar and polar lipids were
253 extracted from the cell pellets using methods described by Dobson et al. (G. Dobson et al.,
254 1985). Polar lipids were analysed by 2D-TLC using solvent system E, which is designed

255 to separate phospholipids and LOSs (G. Dobson et al., 1985). Glycolipids were visualized
256 by charring following staining with either molybdophophoric acid (MPA) or
257 alpha-napthol (for glycolipids).
258
259 **Diagnostic PCR markers for *M. riyadhense***
260 To develop diagnostic markers for *M. riyadhense* for potential use in clinical diagnosis as
261 well as environmental and animal studies, unique regions within the *M. riyadhense*
262 reference genome compared to that of 152 other mycobacterial species were detected
263 using Shustring (Haubold et al., 2005). These regions were also examined by OrthoMCL
264 (Li et al., 2003) and Blastn (Camacho et al., 2009). *mr_00036*, *mr_00263*, *mr_00606*, and
265 *mr_01005* were selected as the amplification targets. Two primers for each gene were
266 designed in this study:
267 MRDP-MR_00036-F (5'-TTCGTTGTCGGTTTCGTCGC-3') and MRDP-MR_00036-R
268 (5'-GCGTCAGCTCCACCGAAAAC-3');
269 MRDP-MR_00263-F (5'-CCACCGCTGTTGGCGA-3') and MRDP-MR_00263-R
270 (5'-TTCGTCCCGTTGATCCCGTT-3');
271 MRDP-MR_00606-F (5'-AACCTGCCCGATACGCACTT-3') and
272 MRDP-MR_00606-R (5'-ACTGTTCCTCCGTGGGGTTG-3');
273 MRDP-MR_01005-F (5'-GACTGTGGGGTAACGGTGGA-3') and
274 MRDP-MR_01005-R (5'-CCGGTGATGTCGCCTACTCC-3').

275 PCR was performed in a 25 µl reaction volume with 12.5 µl of GoTaq ® Green Master
276 Mix (Promega, USA), 1 µl of 100 ng/µl gDNA, 1 µl with 10 nmol of forward and reverse
277 primers, 3 µl of dimethyl sulfoxide (DMSO) and 19 µl of nuclease-free water. The PCR
278 mixture was denatured for 5 minutes at 94°C; followed by 35 cycles of amplification
279 involving a denaturation step at 94°C for 30 seconds, a primer annealing step at 59°C for
280 45 seconds, and a primer extension step at 72°C for 45 seconds; and a final extension step
281 at 72°C for 7 minutes. The ITS-F/mycom-2 primer set (Park et al., 2000), which is a
282 *Mycobacterium* genus-specific primer set, was used as a control, with amplification
283 conditions as described previously (Park et al., 2000). The products were electrophoresed
284 in a 2% agarose gel for 60 minutes and visualized.
285
286 **Results**
287
288 **Clinical manisfestation and culture characteristics of *M. riyadhense* isolates**
289 Between April 2011 and March 2017, eight clinical cases of infection with *M. riyadhense*
290 were recorded in male patients aged from 8-82 years (Table 1). In addition to HIV/AIDS,
291 most of the patients had multiple comorbidities, such as pulmonary and/or systemic
292 hypertension, malignancies, and diabetes mellitus (DM), with the lung being the major
293 site of disease (Table 1).
294

295 Bronchoalveolar lavage, endotracheal, sputum and lymph node biopsy specimens grew *M.*
296 *riyadhense* colonies when cultured in solid and liquid media using LJ agar and
297 *Mycobacteria* growth indicator tube (MGIT) broth, respectively, albeit with varying
298 times to positivity. Although no consistent findings existed for chest imaging, three out of
299 the eight patients presented with upper lobe consolidation, cavitation, ground-glass
300 opacities, 'tree-in-bud' appearance, hilar lymphadenopathy and pleural effusion (Figure
301 supplementary 1).
302
303 The isolates were subjected to susceptibility testing for antibiotics commonly used to
304 treat both typical and atypical infections. Isolates showed 100% susceptibility to rifampin
305 (RIF), rifabutin (RFB), ethambutol (EMB), clarithromycin (CLR), linezolid (LZD),
306 amikacin (AMK), moxifloxacin (MOXI), and trimethoprim-sulfamethoxazole
307 (TMP-SMX). Three out of the eight isolates were resistant to ciprofloxacin (CIP).
308
309 **Assembly and annotation of the *M. riyadhense* MR226 genome**
310 The comparison of different assemblies of all sequenced *M. riyadhense* strains is listed in
311 Supplementary Table 1. We obtained chromosomes of all eight isolates in single
312 contiguous sequences for genome comparison at a high resolution. The *M. riyadhense*
313 MR226 strain contains a 6,243,587 bp chromosome, a linear plasmid (pMRLP01) of
314 550,247 bp and a circular plasmid (pMR01) of 94,344 bp. The circular nature of the

315 chromosome and the pMR01 plasmid were demonstrated through Gepard (Krumsiek et
316 al., 2007).
317 As expected of a free-living opportunistic pathogen, the chromosome of *M. riyadhense* is
318 significantly larger than the chromosomes of the MTBC (Table 2). The number of genes
319 unique in a species showed that the members of the MTBC have a considerably lower
320 percentage of unique genes than *M. riyadhense* and other closely related NTMs (Table 2).
321
322 The comparison of the annotated protein-coding genes from the *M. riyadhense* MR226
323 strain to the genome assemblies of 152 *Mycobacterium* species (including 77 known
324 mycobacterial plasmids) shows that *M. riyadhense* forms a tight cluster with the MTBC
325 and some NTMs (Figure supplementary 2A, highlighted in box a). A total of 335 genes
326 were identified as unique from the orthologue group comparison, being present in only all
327 sequenced *M. riyadhense* isolates, including MR226, and the vast majority of these genes
328 belong to the PE/PPE family, which are thought to be involved in antigen variation and
329 are widely spread across the slow-growing species within the *Mycobacterium* genus.
330
331 Linear plasmids were first described in 1989 in maize (which has a linear
332 mitochondrion)(Leon et al., 1989) and have also been found in *Actinomycetales*,
333 including *Streptomyces* (Kinashi, 2011), *Rhodococcus* (Crespi et al., 1992) and
334 *Mycobacterium* species, such as *Mycobacterium xenopi*, *Mycobacterium branderi* and

335 *Mycobacterium celatum*. They are often accompanied by a circular plasmid in the same
336 host (Picardeau and Vincent, 1997). The linear plasmid pMRLP01 in *M. riyadhense*
337 contains a pair of partitioning genes (*parA*/*parB*) that are involved in active segregation
338 and thus stabilize the inheritance of the plasmid (Surtees and Funnell, 2003). The latter
339 are known to contribute to genome evolution by active DNA transfer and exchange
340 (Zrimec and Lapanje, 2018). As is often the case for both circular and linear large
341 plasmids, a relatively higher proportion of pMRLP01 genes (51%) have no known
342 function compared to those of the main chromosome (26%). This reinforces the notion
343 that plasmids are an important route by which new genes are introduced into the genome
344 in *Mycobacteria*. Of the 443 predicted protein-coding genes of pMRLP01, 118 have at
345 least one orthologue on the main chromosome. Furthermore, we observed several
346 protein-coding genes in pMRLP01 that have orthologues in the genomes of
347 *Mycobacterium tusciae* JS617, *Mycobacterium aromaticivorans* JCM 16368,
348 *Mycobacterium llatzerense*, *Mycobacterium obuense*, *Mycobacterium novocastrense* and
349 *Mycobacterium holsaticum* (Figure supplementary 2B). This finding indicates that
350 pMRLP01 can be readily exchanged with other environmental mycobacterial species.
351
352 It is well known that plasmids are important "vehicles" for the exchange of genetic
353 material between bacteria or between chromosomes and extra-chromosomal plasmids. In
354 this study, we further identified a circular plasmid termed pMR01 in *M. riyadhense*

355 (Figure supplementary 2C, Figure supplementary 3). When compared with the circular
356 plasmids of other species, such as pRAW in *M. marinum* (Ummels et al., 2014),
357 pMAH135 (Uchiya et al., 2015) and pMA100 (da Silva Rabello et al., 2012) of *M. avium*,
358 pMyong1 from *Mycobacterium yongonense* (Kim et al., 2013), pMK12478 (Veyrier et al.,
359 2009) from *M. kansasii* and several plasmids from *Mycobacterium chimaera* (van Ingen
360 et al., 2017), a high similarity was observed. These plasmids all harbour both a TS4 and
361 TS7, which are necessary for conjugation (Ummels et al., 2014)(Morgado et al., 2017),
362 and facilitate the exchange of genetic material between different species of slow-growing
363 mycobacteria (Ummels et al., 2014). We therefore speculate that pMR01 is a novel
364 conjugative plasmid.
365
366 In mycobacteria, five type VII secretion systems have been described, named ESX-1 to
367 ESX-5 (Houben et al., 2014),(Abdallah et al., 2007). An *esx-5* locus on pMR01, which
368 shows high similarity to the ESX-5 loci on pMK12478, pRAW and pMAH135, is
369 markedly different from the ESX-5 system found on the main *M .riyadhense*
370 chromosome. ESX-5 is involved in the secretion of PE and PPE proteins in
371 *M. tuberculosis* and is involved in modulating the host immune responses to maintain a
372 persistent infection (Weerdenburg et al., 2012). The potential transmissibility of pMR01
373 and other pMR01-like plasmids may therefore play a role in the evolution of the ESX
374 systems in mycobacteria.

375

376 The progressive alignments (Figure supplementary 4A) of the assembled chromosomes
377 and plasmids (Figure supplementary 4B) of each *M. riyadhense* strain show that the
378 chromosomes are relatively conserved; however, the linear plasmids present in all eight
379 sequenced isolates are diverse from both structural and similarity perspectives, while the
380 pRAW-like plasmids are present in only the MR226, MR193 and MR222 strains.

381

382 The SNP-based phylogeny of the sequenced *M. riyadhense* isolates based on 43,136
383 polymorphic sites (Supplementary Table 2, Figure supplementary 5) indicates the
384 presence of two different clades of *M. riyadhense* among the clinical isolates sequenced
385 in this study. The nucleotide diversity between the MR222 clade is greater than the
386 diversity between *M. tuberculosis* strains (Jia et al., 2017), while it is smaller than that
387 seen between *M. canettii* strains (Supply et al., 2013), and the variation between the
388 MR226 clade is comparable to the SNP variation in *M. tuberculosis* strains.

389

390 **Regions of difference (RDs) in *M. riyadhense***

391 The RDs were originally described as genomic regions present in virulent *M. bovis* and *M.*
392 *tuberculosis* but absent from the *M. bovis* BCG genome (Mahairas et al., 1996). RD loci
393 were subsequently described across the MTBC (Gordon et al., 1999) and contain
394 functions believed to contribute to pathogenicity (Jakko Van Ingen et al., 2009)(Kozak et

395 al., 2011)(Ru et al., 2017) and evolution of MTBC species (Brosch et al., 2002). *M.*
396 *riyadhense* was found to harbour most of the RD loci (RD1, RD3-R11, R13-RD16) that
397 are also intact in *M. tuberculosis,* while 2 of the RDs show unique deletions, $RD2^{riyadh}$
398 (Figure supplementary 6A) and $RD12^{riyadh}$ (Figure supplementary 6B).
399
400 RD2 was originally described as deleted in BCG vaccine strains. Subsequently, it was
401 shown that disruption of RD2 in *M. tuberculosis* leads to decreased proliferation *in vivo*
402 and impaired modulation of the innate immune response (Kozak et al., 2011). The RD2
403 locus also has a deletion from the *M. riyadhense* genome. It is a larger deletion than the
404 originally described $RD2^{BCG}$ as $RD2^{riyadh}$ contains 29 genes (*rv1971~rv2000*, location
405 2,216,498~2,246,766); eight genes within this locus (*rv1978*, *rv1979c*, *rv1980c*, *rv1981c*,
406 *rv1983*, *rv1984*, *rv1987*, *rv1988*) have orthologues elsewhere in the *M. riyadhense*
407 genome (*mr_05764*, *mr_05852*, *mr_02310*, *mr_02993*, *mr_00486*, *mr_02995*, *mr_02325*,
408 *mr_02349*, *mr_01747*), suggesting possible functional redundancy.
409
410 The RD12 locus shows deletions across MTBC members, including *M. bovis*,
411 *Mycobacterium caprae* and *M. orygis* (Alexander et al., 2010), but it is present in other
412 MTBC members. *M. canettii* isolates (except group B (M Cristina et al., 2005)) also show
413 an independent deletion at the RD12 locus named $RD12^{can}$ (3,479,430~3,491,866,
414 *rv3111~rv3126*), which is distinct from $RD12^{bovis}$ (3,484,740~3,487,515,

415 *rv3117~rv3121*). We identified another unique deletion at the RD12 locus in *M.*
416 *riyadhense*, designated RD12^{riyadh}, which encompasses a larger region than RD^{can} and
417 RD12^{bovis}, encompassing *rv3108-rv3127* (3,477,171~3,492,150) (Figure supplementary
418 6B). It is intriguing that multiple mycobacteria show independent deletion events at the
419 RD2 and RD12 loci, suggesting selective forces play a role in this variation.
420

421 **Comparative phylogeny of *M. riyadhense* with other Mycobacteria**

422 The phylogenetic tree shows that the slow-growers and rapid-growers are separated into
423 two different clades and that fast-growers evolved from slow-growers (Figure 1, the
424 detailed figure is shown in Figure supplementary 7). The overall topology of our tree is
425 remarkably similar to that of previously published phylogenetic trees (Tortoli et al., 2017).
426 In the tree, *M. riyadhense* is located within the same clade as the obligate and
427 opportunistic causal organisms of mycobacterial diseases in humans that include the
428 MTBC, *M. marinum*, *M. kansasii*, *M. leprae* and related host-restricted mycobacteria
429 with reduced genomes and decreased survivability in the environment. Notably,
430 *M. shinjukuense*, *M. lacus* and *M. riyadhense* form a sub-clade that is phylogenetically
431 close to the MTBC.
432

433 The *PE*/*PPE* family *mce* and *mce*-associated genes are known to be important for host
434 adaptation (Delogu and Brennan, 2001) and pathogenicity (Isom et al., 2017). The

435 *PE*/*PPE* family genes are enriched in the MTBC members but also in *M. riyadhense*
436 MR226 (278) and other pathogenic species, such as *M. kansasii* (228) and
437 *Mycobacterium ulcerans* (200). The number of *mce*- or *mce*-associated genes has not
438 significantly changed across mycobacterial genomes, indicating that this group of genes
439 is under an evolutionary constraint and plays functional roles bridging both
440 environmental and obligate pathogen lifestyles. Our results agree with previous finding
441 that during their evolution, the ESX systems were derived from the ancestor ESX-4, as
442 shown in Figure supplementary 7 at the root node, and then ESX-3, ESX-1, ESX-5 and
443 ESX-2 evolved by horizontal transfer (Newton-Foot et al., 2016).
444
445 A comparative phylogenetic map based on 1,301 conserved proteins (Figure
446 supplementary 8) reveals this downsizing of the genome and the dynamic changes in
447 genome components. Certain functional categories of genes are relatively enriched during
448 the evolution of MTBC, including protein metabolism, regulation and cell signalling, cell
449 division and cell cycle.
450
451 *M. riyadhense* shares a larger number of orthologues (3,122) with *M. tuberculosis* than
452 with *M. kansasii* (2,978), *M. marinum* (2,962) and *M. szulgai* (2,724) among the
453 environmental mycobacteria that are closely related to the MTBC (Figure 2A). A total of
454 134 orthologues are exclusively shared between *M. riyadhense* and *M. tuberculosis*,

455 while the number of orthologues exclusively shared between *M. tuberculosis* and *M.*
456 *kansasii* (30), *M. marinum* (48) and *M. szulgai* (18) is less (Figure 2A). It is notable that
457 genes from the phage-derived regions of RD3 and RD11 are shared exclusively between
458 *M. riyadhense* and *M. tuberculosis*.
459
460 The comparative analysis of the orthologue groups of *M. riyadhense* and the MTBC is
461 revealing. Firstly, 385 orthologue groups present across all MTBC species are absent
462 from *M. riyadhense*. Secondly, 221 orthologues uniquely present in *M. riyadhense* are not
463 found amongst the MTBC (Figure 2B). This latter group of *M. riyadhense* unique
464 orthologues likely depict the constraints imposed by the free-living biology of *M.*
465 *riyadhense* which necessitates maintaining a broad functional repertoire to secure
466 environmental survival. In contrast, the obligate MTBC species have lost genes involved
467 in environmental survival but gained a large number of genes required for survival in the
468 *in vivo* environment.
469
470 A hallmark of *M. tuberculosis* infection is the ability to survive long-term in host
471 granulomas and develop a latent stage of infection. The molecular mechanisms and
472 cellular components that are involved in the persistence of *M. tuberculosis* are still poorly
473 understood, but several T/A systems have been implicated in the pathogenicity of *M.*
474 *tuberculosis* (Slayden et al., 2018). T/A systems were first found on plasmids or

475 plasmid-derived chromosomal loci where they promoted plasmid maintenance in
476 bacterial populations (Cooper and Heinemann, 2000), but when compared to other
477 mycobacteria, the MTBC are remarkable for the extensive expansion of T/A systems. We
478 thus compared the 79 pairs of T/A systems (belonging to the HigAB, MazEF, ParDE,
479 RelEF, VapBC and UCAT families) in *M. tuberculosis* with the T/A pairs found in other
480 members of the MTBC and NTMs. Based on the presence of 49 out of the 79 T/A
481 orthologue pairs (Figure 3), *M. riyadhense* appears more closely related to the MTBC
482 than to other mycobacteria, including *M. lacus, M. shinjukuense* and *M. decipiens*. It can
483 be hypothesized that the shared component of T/A pairs play a functional role in the
484 pathogenicity or persistence of *M. riyadhense* infection in a way similar to that described
485 for the importance of T/A pairs in *M. tuberculosis in vivo* biology.
486
487 ***M. riyadhense* strains produce a distinct pattern of LOSs**
488 The LOSs are an important class of glycolipids that have previously been linked to
489 diverse mycobacterial phenotypes, such as colony morphology, secretion of PE/PPE
490 family proteins, and the pathogenicity of *M. marinum* (Van Der Woude et al., 2012).
491 Notably, the presence or absence of LOS from *M. canettii* causes a transition from
492 smooth to rough colony variants, respectively, with rough variants showing increased
493 virulence; MTBC strains lack LOS, and it has been suggested that the removal of LOS
494 was a key event in the evolution of the MTBC species towards their current obligate
495 pathogen status (Boritsch et al., 2016). It is noteworthy that we observed both smooth

496 (MR210, MR22, MR226, MR244, MR246, and MR1023) (Figure 4A) and rough
497 (MR193 and MR206) (Figure 4B) morphologies in *M. riyadhense* strains. We therefore
498 sought to first examine whether the genetic machinery for the production of LOS is
499 present in the *M. riyadhense* genome and then to follow up on the genome-level
500 predictions with lipid analyses of rough and smooth variants using TLC.
501
502 We observed that the *wecE* and *galE6* LOS genes are absent from the *M. riyadhense*
503 genome (Table 3, Supplementary Table 3). These genes are linked with the removal of
504 LOS II* and the production of LOS IV (Van Der Woude et al., 2012). Thus, their absence
505 is likely to cause an accumulation of LOS II* and the lack of fully formed LOS IV, which
506 have previously been shown to increase the pathogenicity of *M. marinum* (Van Der
507 Woude et al., 2012). Furthermore, both the *pks5* and *pap* genes in the LOS locus are
508 intact in *M. riyadhense*, as is the case in *M. canettii*, but not in *M. tuberculosis,* where the
509 former is truncated and the latter deleted (Boritsch et al., 2016). This finding indicates
510 that *pks5* recombination and *pap* deletion occurred in a common ancestor of the MTBC
511 after its differentiation from both *M. riyadhense* and *M. canettii.* Remarkably, the *M.*
512 *riyadhense* LOS gene locus layout is dissimilar to that in *M. canettii*, *M. tuberculosis*, *M.*
513 *kansasii* and *M. marinum*: indeed, exclusive rearrangements of this locus in *M.*
514 *riyadhense* were observed (Figure 4C).
515

516 To correlate rough vs. smooth colony morphology with LOS production, we extracted
517 polar lipids from the strains and analysed them by 2D-TLC using solvent system E (G.
518 Dobson et al., 1985), which is designed to separate phospholipids and LOSs. Charring of
519 the TLC plates with alpha-napthol revealed glycolipids, including the accumulation of a
520 species that migrated at a position similar to that of LOS III. This lipid was seen in only
521 smooth strains; species with migration patterns similar to LOS I and LOS II were
522 observed, while LOS IV was absent. This result was not unexpected because all *M.*
523 *riyadhense* strains lack a functional *wecA* gene, which is required for the extension of
524 LOS II to LOS IV (Figure 4). Additionally, the relative levels of the predominant LOS
525 species in *M. riyadhense* seem to be quite high when compared to those seen in other
526 LOS-producing mycobacteria (Figure 4 (D)(E)). Conversely, the rough strains did not
527 produce any glycolipids that migrated in the positions corresponding to LOSs (Figure 4
528 (F)(H)). Overall, *M. riyadhense* exhibits a LOS production phenotype distinct from that
529 of other LOS-producing mycobacteria.
530
531 ***PE-PGRS33* locus and type VII secretion system of *M. riyadhense***
532 The *pe-pgrs33* (*rv1818c*) locus encodes the exported protein PE_PGRS33 and plays an
533 important role in the pathogenesis of *M. tuberculosis* (Cohen et al., 2014). A previous
534 study (Boritsch et al., 2014) showed that *pe-pgrs33* is present in all MTBC members but
535 not in *M. canettii*, which implies a specific *pe_pgrs33* insertion event in the ancestor of

536 MTBC strains. Genome comparison of *M. riyadhense* with *M. tuberculosis*, *M. kansasii*,
537 *M. marinum* and *M. canettii* provides additional evidence that *M. riyadhense* is the
538 missing link of the *pe-pgrs33* deletion/insertion event. Our phylogeny strongly suggests
539 that the deletion of *pe-pgrs33* from *M. kansasii* and *M. marinum* occurred before the
540 divergence of environmental mycobacteria and the smooth tubercle bacilli (STB)/MTBC
541 clade (Figure supplementary 9).
542
543 All 5 ESX systems (ESX1-ESX5) were found in the *M. riyadhense* genome (Figure
544 supplementary 10) but with minor modifications. Hence, the *eccA* and *eccB* genes are
545 absent from the ESX-2 system, while the *espACD* operon, which is essential for secretion
546 of virulence factors via ESX1, is also missing in *M. riyadhense*. The overall gene
547 arrangement of the ESX1-ESX5 loci is very similar in both *M. riyadhense* and
548 *M. tuberculosis* (Houben et al., 2014). This conserved synteny reinforces the previous
549 results that phylogenetically *M. riyadhense* may represent an ancestral state to MTBC. As
550 noted before, the pMR01 plasmid also contains an extra ESX-5 locus, which could
551 indicate a role for this plasmid in mediating pathogenicity (Figure supplementary 4).
552
553 **Transcriptional response of murine macrophage cells upon *M. riyadhense* infection**
554 Our genomic analysis of *M. riyadhense* revealed a range of genes and potential systems
555 that could play a role in host-pathogen interactions. We therefore sought to assess the

556 initial interaction of *M. riyadhense* with macrophages, using the RAW264.7 cell line of
557 murine origin as our experimental model. As comparator strains in our analysis, we
558 performed parallel infections with *M. kansasii*, an opportunistic pathogen that also
559 contains an orthologous ESX-1 system (Figure 1, Figure supplementary 10), and *M. bovis*
560 BCG, the live TB vaccine that is attenuated through deletion of the ESX-1 system. These
561 comparisons allowed us to explore whether *M. riyadhense* triggered innate immune
562 responses were more similar to those triggered by *M. kansasii* or BCG or were
563 intermediate between the two.
564
565 To analyse the innate immune responses of the macrophages in the intracellular presence
566 or absence of these mycobacterial isolates, the transcriptional response was analysed
567 using a 754 probe NanoString Murine Myeloid Innate Immunity panel V2 (Geiss et al.,
568 2008) at 3, 24 and 48 hours post infection. These analyses revealed an expected
569 commonality in the responses to infection with all mycobacteria, such as induction of
570 proinflammatory genes through TLR signalling (e.g., upregulation of *IL1B*, *TNFA*, *CCL4*,
571 *PTGS2* and *CXCL2*, albeit to different absolute levels, Figure 5, Figure supplementary
572 11A and 11B). Distinct responses triggered by BCG infection compared to *M. riyadhense*
573 and *M. kansasii* included, for example, upregulation of *MARCO* by BCG infection
574 (Figure supplementary 11B); *MARCO* is involved in pathogen uptake via trehalose
575 dimycolate, a lipid that is known to show variation in structure between the MTBC and *M.*

576 *kansasii* (Bowdish et al., 2009)(Fujita et al., 2005). *CCL24* and *CXCL14*, which are
577 involved in the attraction of immune cells to the site of infection (Stavrum et al., 2011),
578 were upregulated to higher levels with *M. riyadhense* and *M. kansasii* than with BCG at
579 24 and 48 hours.
580 Overall, our analysis of the macrophage transcriptional profiles showed that the response
581 to infection with *M. riyadhense* and *M. kansasii* triggered more similar macrophage
582 transcriptional responses in comparison to the response produced by BCG infection
583 (Figure supplementary 11).
584
585 **Developing a rapid PCR-based diagnostic marker for *M. riyadhense***
586 Due to the issues previously encountered in diagnosing *M. riyadhense* infections (Choi et
587 al., 2012)(van Ingen et al., 2009), correct and prompt identification of cases upon
588 presentation at healthcare units is of paramount importance. We therefore sought to
589 translate our knowledge on the genome sequences into a PCR diagnostic test that could
590 be used in a clinical microbiology setting to distinguish *M. riyadhense* from other
591 mycobacteria, including the members of the MTBC.
592
593 By identifying unique K-mers ranging in size from 11 bp to 4,209 bp (Figure
594 supplementary 12) in the assembled genome compared to the genomes of 152 other
595 mycobacterial species, a primer set was developed targeting the *mr_00036*, *mr_00263*,

596 *mr_00606* and *mr_01005* genes. The MRDP primer pair MRDP-F/MRDP-R amplified a
597 single product of 372 bp from each of the eight isolates of *M. riyadhense* (Figure 6) but
598 not from other mycobacterial species, including *M. tuberculosis*, *M. bovis*, *M. kansasii*, *M.*
599 *marinum*, *M. szulgai*, *M. avium* and *Mycobacterium angelicum*. This result shows that the
600 MRDP-F/MRDP-R primers are highly specific to *M. riyadhense* and form the basis for a
601 simple diagnostic PCR that can inform appropriate treatment protocols.
602
603 **Discussion**
604
605 *M. riyadhense* has become a clinically relevant NTM species globally (Choi et al.,
606 2012)(Katoch, 2004). Contrary to prior publications on *M. riyadhense* have been based
607 primarily on clinical case reports, here we present the largest and most comprehensive
608 genomic study undertaken to date on clinical *M. riyadhense* isolates. Our comparative
609 analysis of *M. riyadhense* genomes with the MTBC and a large collection of NTMs
610 provide unequivocal evidence that *M. riyadhense* is one of the closest known
611 environmental mycobacteria species to the MTBC and forms a phyloclade with *M. lacus*,
612 *M. decipiens* and *M. shinjukuense*.
613
614 Our analyses of *M. riyadhense* reveal that expansion of T/A pairs, modification of
615 secretion systems, alterations in cell wall lipids, and plasmid-mediated horizontal

616 acquisition of new functionality all played key roles in the evolution of the MTBC. Our
617 work adds new insight into the evolution of the MTBC from free-living environmental
618 bacteria to their obligate pathogen biology.
619
620 The eight new *M. riyadhense* strains sequenced in this study originated from pulmonary
621 infections (with some having additional extra-pulmonary involvement), which fulfilled
622 the American Thoracic Society/ Infectious Diseases Society of America (ATS/IDSA)
623 criteria for NTM infection (Griffith et al., 2007).
624
625 Our study shows that *M. riyadhense* shares a larger number of orthologues with *M.*
626 *tuberculosis* than *M. kansasii* and *M. marinum*, notably in the T/A gene family (Figure
627 supplementary 6). Forty-nine pairs of T/A gene orthologues were found in *M. riyadhense*,
628 far greater than the number of orthologues observed in any of the other NTMs. The
629 expansion of T/A genes among the MTBC offers additional proof that gene acquisition
630 through lateral gene transfer played a key role in the development of pathogenicity; this
631 agrees with previous findings (Stinear et al., 2008)(Becq et al., 2007).
632
633 *M. riyadhense* strains appeared as both smooth and rough colony forms when grown on
634 solid LJ media. We observed the presence of LOS I and LOS II and the absence of LOS
635 IV. These biochemical observations agree with the genomic prediction that *M.*

636 *riyadhense* strains lack a functional *wecA* gene, which is required for the extension of
637 LOS II to LOS IV. *In silico* RD analysis *M. riyadhense* revealed independent RD2^{riaydh}
638 and RD12^{riyadh} deletions when compared to *M. canettii* and other members of the MTBC.
639 Frequent and independent deletion events in these specific genomic regions across these
640 groups of closely related mycobacteria suggest a role for natural selection that might be
641 important for pathogenicity.
642
643 All the 5 ESX type VII secretion systems present in *M. tuberculosis*, and known to be
644 involved in virulence and pathogenicity, were found in *M. riyadhense* with very similar
645 gene arrangement. An additional ESX-5 system was also found on the circular plasmid
646 pMR01. pMR01-like plasmids have been shown in many different NTMs, including *M.*
647 *kansasii*, *M. marinum*, *M. chimaera* and *M. avium*, indicating that these groups of
648 plasmids containing both TS4 and TS7 are conjugative plasmids. The extensive presence
649 of these plasmids may also explain the origin of the type VII secretion system (Ummels
650 et al., 2014). ESAT-6 and CFP-10, which are secreted through ESX-1, are critical to
651 phagosome perturbation and manipulation of host macrophages induced by *M.*
652 *tuberculosis* (Jamwal et al., 2016). Linked to this, our transcriptome analysis revealed
653 that *M. riyadhense* and *M. kansasii* trigger similar overall macrophage responses after
654 infection when compared to *M. bovis* BCG (as a representative of the MTBC). These data

655 reinforce the phylogenetic relationship of *M. riyadhense* with the MTBC species and its
656 transitional status between an opportunistic and an obligate pathogen.
657
658 The clinical presentation of our cases was mainly indistinguishable from disease caused
659 by *M. tuberculosis*, as reported earlier (Garbati and Hakawi, 2014)(Choi et al., 2012), but
660 with a negative *M. tuberculosis* PCR. Due to the relatively recent emergence of *M.*
661 *riyadhense* as an important clinical pathogen, and coupled with its misdiagnosis as *M.*
662 *tuberculosis* by commercially available kits, we developed an accurate set of diagnostic
663 markers based on the genomic datasets generated in this study. The primer sets accurately
664 detect *M. riyadhense* in a mixed cocktail of closely related mycobacteria: it can serve as
665 part of an accurate and fast diagnostic protocol in clinical settings thus reducing the need
666 for strict isolation, laborious contact tracing as well as the inappropriate use of TB
667 antimicrobials. . These primers can be used in a global epidemiological survey of cases of
668 *M. riyadhense* infections in humans, animals and the environment. This may thus provide
669 a more complete picture of the epidemiology of *M. riyadhense* following a broad One
670 Health approach. It would, for example, be of clinical interest to see if *M. riyadhense*
671 infections occur in Africa and South America, for which no reports are available, or if *M.*
672 *riyadhense* is uniquely endemic in the Arabian Peninsula. Indeed, a limitation of this and
673 previous similar studies is that all isolates available for genomic studies are clinical. *M.*
674 *riyadhense* is an environmental pathogen that is most likely be found in animals sharing

675 the same niche with humans that were infected from the environment. As the natural
676 environmental reservoir of *M. riyadhense* remains unknown, we believe that systematic
677 screening of relevant environmental samples with the MRDP established in this study
678 may help to establish the natural habitat of this bacterium.
679
680 In conclusion, our study provides unprecedented insights into the ancestry and adaptive
681 evolution of the MTBC relative to extant NTM species and places *M. riyadhense* as one
682 of the closest environmental relatives to the MTBC. Our work provides compelling data
683 supporting the use of *M. riyadhense* as a novel mycobacterium model for the study of
684 virulence evolution and pathogenesis in MTBC.
685
686 **References**
687
688 Abdallah A, Pittius G, Nicolaas C, DiGiuseppe Champion P, Cox J, Luirink J,
689 Vandenbroucke-Grauls C, Appelmelk B, Bitter W. 2007. Type VII secretion -
690 Mycobacteria show the way. *Nat Rev Microbiol* **5**:883–891.
691 doi:10.1038/nrmicro1773
692 Alexander KA, Laver PN, Michel AL, Williams M, van Helden PD, Warren RM, van
693 Pittius NCG. 2010. Novel mycobacterium tuberculosis complex pathogen, M.
694 Mungi. *Emerg Infect Dis* **16**:1296–1299. doi:10.3201/eid1608.100314
695 Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. 2011. BLAST Ring Image
696 Generator (BRIG): Simple prokaryote genome comparisons. *BMC Genomics* **12**.
697 doi:10.1186/1471-2164-12-402

698 Alkan C, Coe BP, Eichler EE. 2011. GATK toolkit. *Nat Rev Genet* **12**:363–76.
699 doi:10.1038/nrg2958
700 Althawadi S, Johani S, Fernandez GM, Al-ghafli H, Enani RMA, Abdul K, Medical A.
701 2017. Mycobacterium riyadhense **23**:2015–2017.
702 Becq J, Gutierrez MC, Rosas-Magallanes V, Rauzier J, Gicquel B, Neyrolles O,
703 Deschavanne P. 2007. Contribution of horizontally acquired genomic islands to the
704 evolution of the tubercle bacilli. *Mol Biol Evol* **24**:1861–1871.
705 doi:10.1093/molbev/msm111
706 Belisle JT, Sonnenberg MG. 1998. Isolation of genomic DNA from mycobacteria.
707 *Methods Mol Biol* **101**:31–44. doi:10.1385/0-89603-471-2:31
708 Blouin Y, Hauck Y, Soler C, Fabre M, Vong R, Dehan C, Cazajous G, Massoure P-L,
709 Kraemer P, Jenkins A, Garnotel E, Pourcel C, Vergnaud G. 2012. Significance of
710 the Identification in the Horn of Africa of an Exceptionally Deep Branching
711 Mycobacterium tuberculosis Clade. *PLoS One* **7**:e52841.
712 doi:10.1371/journal.pone.0052841
713 Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina
714 sequence data. *Bioinformatics* **30**:2114–2120. doi:10.1093/bioinformatics/btu170
715 Boritsch EC, Frigui W, Cascioferro A, Malaga W, Etienne G, Laval F, Pawlik A, Le
716 Chevalier F, Orgeur M, Ma L, Bouchier C, Stinear TP, Supply P, Majlessi L, Daffé
717 M, Guilhot C, Brosch R. 2016. Pks5-recombination-mediated surface remodelling in
718 Mycobacterium tuberculosis emergence. *Nat Microbiol* **1**.
719 doi:10.1038/nmicrobiol.2015.19
720 Boritsch EC, Supply P, Honoré N, Seeman T, Stinear TP, Brosch R. 2014. A glimpse into
721 the past and predictions for the future: The molecular evolution of the tuberculosis
722 agent. *Mol Microbiol*. doi:10.1111/mmi.12720
723 Bowdish DME, Sakamoto K, Kim MJ, Kroos M, Mukhopadhyay S, Leifer CA,
724 Tryggvason K, Gordon S, Russell DG. 2009. MARCO, TLR2, and CD14 are
725 required for macrophage cytokine responses to mycobacterial trehalose dimycolate
726 and Mycobacterium tuberculosis. *PLoS Pathog* **5**. doi:10.1371/journal.ppat.1000474

727 Broad Institute. 2016. Picard tools. *https://broadinstitute.github.io/picard/*.
728 https://broadinstitute.github.io/picard/%5Cnhttp://broadinstitute.github.io/picard/
729 Brosch R, Gordon S V., Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T,
730 Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, van
731 Soolingen D, Cole ST. 2002. A new evolutionary scenario for the Mycobacterium
732 tuberculosis complex. *Proc Natl Acad Sci* **99**:3684–3689.
733 doi:10.1073/pnas.052548299
734 Brown-Elliott BA, Simmer PJ, Trovato A, Hyle EP, Droz S, Buckwalter SP, Borroni E,
735 Branda JA, Iana E, Mariottini A, Nelson J, Matteelli A, Toney NC, Scarparo C, de
736 Man TJB, Vasireddy R, Gandhi RT, Wengenack NL, Cirillo DM, Wallace RJ,
737 Tortoli E. 2018. Mycobacterium decipiens sp. nov., a new species closely related to
738 the Mycobacterium tuberculosis complex. *Int J Syst Evol Microbiol* **68**:3557–3562.
739 doi:10.1099/ijsem.0.003031
740 Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL.
741 2009. BLAST+: Architecture and applications. *BMC Bioinformatics* **10**.
742 doi:10.1186/1471-2105-10-421
743 Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: A tool for automated
744 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**:1972–
745 1973. doi:10.1093/bioinformatics/btp348
746 Carbonne A, Brossier F, Arnaud I, Bougmiza I, Caumes E, Meningaud JP, Dubrou S,
747 Jarlier V, Cambau E, Astagneau P. 2009. Outbreak of nontuberculous mycobacterial
748 subcutaneous infections related to multiple mesotherapy injections. *J Clin Microbiol*
749 **47**:1961–1964. doi:10.1128/JCM.00196-09
750 Choi JI, Lim JH, Kim SR, Lee SH, Park JS, Seo KW, Jeon JB, Jeong J. 2012. Lung
751 infection caused by Mycobacterium riyadhense confused with Mycobacterium
752 tuberculosis: The first case in Korea. *Ann Lab Med* **32**:298–303.
753 doi:10.3343/alm.2012.32.4.298
754 Cohen I, Parada C, Acosta-Gío E, Espitia C. 2014. The PGRS domain from PE_PGRS33
755 of Mycobacterium tuberculosis is target of humoral immune response in mice and
756 humans. *Front Immunol* **5**:1–9. doi:10.3389/fimmu.2014.00236

757 Cooper TF, Heinemann JA. 2000. Postsegregational killing does not increase plasmid
758 stability but acts to mediate the exclusion of competing plasmids. *Proc Natl Acad*
759 *Sci* **97**:12643–12648. doi:10.1073/pnas.220077897
760 Crespi M, Messens E, Caplan AB, van Montagu M, Desomer J. 1992. Fasciation
761 induction by the phytopathogen Rhodococcus fascians depends upon a linear
762 plasmid encoding a cytokinin synthase gene. *EMBO J* **11**:795–804.
763 da Silva Rabello MC, Matsumoto CK, de Almeida LGP, Menendez MC, de Oliveira RS,
764 Silva RM, Garcia MJ, Leão SC. 2012. First description of natural and experimental
765 conjugation between mycobacteria mediated by a linear plasmid. *PLoS One* **7**.
766 doi:10.1371/journal.pone.0029884
767 Delogu G, Brennan MJ. 2001. Comparative Immune Response to PE and PE _ PGRS
768 Antigens of Mycobacterium tuberculosis Comparative Immune Response to PE and
769 PE _ PGRS Antigens of Mycobacterium tuberculosis. *Infect Immun* **69**:5606–5611.
770 doi:10.1128/IAI.69.9.5606
771 Ding L-W, Lai C-C, Lee L-N, Huang L-M, Hsueh P-R. 2005. Lymphadenitis caused by
772 non-tuberculous mycobacteria in a university hospital in Taiwan: predominance of
773 rapidly growing mycobacteria and high recurrence rate. *J Formos Med Assoc*
774 **104**:897–904.
775 Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time
776 and space complexity. *BMC Bioinformatics* **5**:113. doi:10.1186/1471-2105-5-113
777 Falkinham JO. 2013. Ecology of nontuberculous mycobacteria-where do human
778 infections come from? *Semin Respir Crit Care Med* **34**:95–102.
779 doi:10.1055/s-0033-1333568
780 Fedrizzi T, Meehan CJ, Grottola A, Giacobazzi E, Fregni Serpini G, Tagliazucchi S,
781 Fabio A, Bettua C, Bertorelli R, De Sanctis V, Rumpianesi F, Pecorari M, Jousson O,
782 Tortoli E, Segata N. 2017. Genomic characterization of Nontuberculous
783 Mycobacteria. *Sci Rep* **7**. doi:10.1038/srep45258
784 Fujita Y, Naka T, McNeil MR, Yano I. 2005. Intact molecular characterization of cord
785 factor (trehalose 6,6′-dimycolate) from nine species of mycobacteria by

786 MALDI-TOF mass spectrometry. *Microbiology* **151**:3403–3416.
787 doi:10.1099/mic.0.28158-0
788 G.Dobson., D.E.Minnikin., S.M.Minnikin., M.Parlett., M.Goodfellow., M.Ridell.,
789 M.Magnusson. 1985. Systematic analysis of complex mycobacterial lipids.Chemical
790 Methods in Bacterial Systematics. Academic Press, London, United Kingdom. pp.
791 237–265.
792 Garbati, Hakawi AM. 2014. Mycobacterium Riyadhense Lung Infection in a Patient with
793 HIV/AIDS, Sub-Saharan African Journal of Medicine.
794 Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, Fell HP, Ferree
795 S, George RD, Grogan T, James JJ, Maysuria M, Mitton JD, Oliveri P, Osborn JL,
796 Peng T, Ratcliffe AL, Webster PJ, Davidson EH, Hood L. 2008. Direct multiplexed
797 measurement of gene expression with color-coded probe pairs. *Nat Biotechnol*
798 **26**:317–325. doi:10.1038/nbt1385
799 Godreuil S, Marchandin H, Michon AL, Ponsada M, Chyderiotis G, Brisou P, Bhat A,
800 Panteix G. 2012. Mycobacterium riyadhense pulmonary infection, France and
801 Bahrain. *Emerg Infect Dis*. doi:10.3201/eid1801.110751
802 Gordon S V., Brosch R, Billault A, Garnier T, Eiglmeier K, Cole ST. 1999. Identification
803 of variable regions in the genomes of tubercle bacilli using bacterial artificial
804 chromosome arrays. *Mol Microbiol* **32**:643–655.
805 doi:10.1046/j.1365-2958.1999.01383.x
806 Griffith DE, Aksamit T, Brown-Elliott B a., Catanzaro A, Daley C, Gordin F, Holland
807 SM, Horsburgh R, Huitt G, Iademarco MF, Iseman M, Olivier K, Ruoss S, von
808 Reyn CF, Wallace RJ, Winthrop K. 2007. An Official ATS/IDSA Statement:
809 Diagnosis, Treatment, and Prevention of Nontuberculous Mycobacterial Diseases.
810 *Am J Respir Crit Care Med* **175**:367–416. doi:10.1164/rccm.200604-571ST
811 Haubold B, Pierstorff N, Möller F, Wiehe T. 2005. Genome comparison without
812 alignment using shortest unique substrings. *BMC Bioinformatics* **6**.
813 doi:10.1186/1471-2105-6-123

814 Houben ENG, Korotkov K V., Bitter W. 2014. Take five - Type VII secretion systems of
815 Mycobacteria. *Biochim Biophys Acta - Mol Cell Res* **1843**:1707–1716.
816 doi:10.1016/j.bbamcr.2013.11.003
817 Ihaka R, Gentleman R. 1996. R: A Language for Data Analysis and Graphics. *J Comput*
818 *Graph Stat* **5**:299–314. doi:10.1080/10618600.1996.10474713
819 Isom GL, Davies NJ, Chong ZS, Bryant JA, Jamshad M, Sharif M, Cunningham AF,
820 Knowles TJ, Chng SS, Cole JA, Henderson IR. 2017. MCE domain proteins:
821 Conserved inner membrane lipid-binding proteins required for outer membrane
822 homeostasis. *Sci Rep* **7**. doi:10.1038/s41598-017-09111-6
823 Jamwal S V., Mehrotra P, Singh A, Siddiqui Z, Basu A, Rao KVS. 2016. Mycobacterial
824 escape from macrophage phagosomes to the cytoplasm represents an alternate
825 adaptation mechanism. *Sci Rep* **6**. doi:10.1038/srep23089
826 Jang J, Becq J, Gicquel B, Deschavanne P, Neyrolles O. 2008. Horizontally acquired
827 genomic islands in the tubercle bacilli. *Trends Microbiol* **16**:303–308.
828 doi:10.1016/j.tim.2008.04.005
829 Jia X, Yang L, Dong M, Chen S, Lv L, Cao D, Fu J, Yang T, Zhang J, Zhang X, Shang Y,
830 Wang G, Sheng Y, Huang H, Chen F. 2017. The Bioinformatics Analysis of
831 Comparative Genomics of Mycobacterium tuberculosis Complex (MTBC) Provides
832 Insight into Dissimilarities between Intraspecific Groups Differing in Host
833 Association, Virulence, and Epitope Diversity. *Front Cell Infect Microbiol* **7**.
834 doi:10.3389/fcimb.2017.00088
835 Katoch VM. 2004. Infections due to non-tuberculous mycobacteria (NTM). *Indian J Med*
836 *Res*. doi:10.1016/j.phrs.2009.12.004
837 Kim B-J, Kim B-R, Lee S-Y, Seok S-H, Kook Y-H, Kim B-J. 2013. Whole-Genome
838 Sequence of a Novel Species, Mycobacterium yongonense DSM 45126 T. *Genome*
839 *Announc* **1**:604–13. doi:10.1128/genomeA.00604-13
840 Kinashi H. 2011. Giant linear plasmids in Streptomyces: A treasure trove of antibiotic
841 biosynthetic clusters. *J Antibiot (Tokyo)*. doi:10.1038/ja.2010.146
842 King HC, Khera-Butler T, James P, Oakley BB, Erenso G, Aseffa A, Knight R,
843 Wellington EM, Courtenay O. 2017. Environmental reservoirs of pathogenic

844 mycobacteria across the Ethiopian biogeographical landscape. *PLoS One* **12**.
845 doi:10.1371/journal.pone.0173811
846 Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu:
847 Scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat
848 separation. *Genome Res* **27**:722–736. doi:10.1101/gr.215087.116
849 Kozak RA, Alexander DC, Liao R, Sherman DR, Behr MA. 2011. Region of difference 2
850 contributes to virulence of mycobacterium tuberculosis. *Infect Immun* **79**:59–66.
851 doi:10.1128/IAI.00824-10
852 Krumsiek J, Arnold R, Rattei T. 2007. Gepard: A rapid and sensitive tool for creating
853 dotplots on genome scale. *Bioinformatics* **23**:1026–1028.
854 doi:10.1093/bioinformatics/btm039
855 Kumar S, Warren S, Cotechini T, Medler T, Bailey C, Beechem JM, Beechem JM,
856 Coussens LM. 2017. Abstract 3706: Validation of human and mouse myeloid panels
857 on the NanoString® nCounter® Platform. pp. 3706–3706.
858 doi:10.1158/1538-7445.am2017-3706
859 Leon P, Virginia W, Patricia B. 1989. Molecular analysis of the linear 2.3 kb plasmid of
860 maize mitochondria: Apparent capture of tRNA genes. *Nucleic Acids Res* **17**:4089–
861 4099. doi:10.1093/nar/17.11.4089
862 Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler
863 transform. *Bioinformatics* **25**:1754–1760. doi:10.1093/bioinformatics/btp324
864 Li L, Stoeckert CJ, Roos DS. 2003. OrthoMCL: Identification of ortholog groups for
865 eukaryotic genomes. *Genome Res* **13**:2178–2189. doi:10.1101/gr.1224503
866 M Cristina G, Brisse S, Brosch R, Fabre M, Omaïs B, Marmiesse M, Supply P, Vincent
867 V. 2005. Ancient origin and gene mosaicism of the progenitor of Mycobacterium
868 tuberculosis. *PLoS Pathog* **1**:0055–0061. doi:10.1371/journal.ppat.0010005
869 Mahairas GG, Sabo PJ, Hickey MJ, Singh DC, Stover CK. 1996. Molecular analysis of
870 genetic differences between Mycobacterium bovis BCG and virulent M. bovis. *J*
871 *Bacteriol* **178**:1274–1282. doi:10.1128/jb.178.5.1274-1282.1996
872 Morgado SM, Marín MA, Freitas FS, Fonseca EL, Vicente ACP. 2017. Complete
873 plasmid sequence carrying type IV-like and type VII secretion systems from an

874 atypical mycobacteria strain. *Mem Inst Oswaldo Cruz* **112**:514–516.
875 doi:10.1590/0074-02760160546
876 Murty DS PK. 2014. A Brief review on Ecology and Evolution of Mycobacteria.
877 *Mycobact Dis* **4**. doi:10.4172/2161-1068.1000172
878 Naeem R, Hidayah L, Preston MD, Clark TG, Pain A. 2014. SVAMP: Sequence
879 variation analysis, maps and phylogeny. *Bioinformatics* **30**:2227–2229.
880 doi:10.1093/bioinformatics/btu176
881 Narendrula-Kotha R, Nkongolo KK. 2017. Microbial response to soil liming of damaged
882 ecosystems revealed by pyrosequencingand phospholipid fatty acid analyses. *PLoS*
883 *One* **12**. doi:10.1371/journal.pone.0168497
884 Newton-Foot M, Warren RM, Sampson SL, Van Helden PD, Gey Van Pittius NC. 2016.
885 The plasmid-mediated evolution of the mycobacterial ESX (Type VII) secretion
886 systems. *BMC Evol Biol* **16**. doi:10.1186/s12862-016-0631-2
887 Panwalker AP, Fuhse E. 1986. Nosocomial Mycobacterium gordonae pseudoinfection
888 from contaminated ice machines. *Infect Control* **7**:67–70.
889 doi:10.1017/S0195941700063918
890 Park H, Jang H, Kim C, Chung B, Chang CL, Soon Kew Park, Song S. 2000. Detection
891 and identification of mycobacteria by amplification of the internal transcribed spacer
892 regions with genus- and species-specific PCR primers. *J Clin Microbiol* **38**:4080–
893 4085.
894 Picardeau M, Vincent V. 1997. Characterization of large linear plasmids in mycobacteria.
895 *J Bacteriol* **179**:2753–2756. doi:10.1128/jb.179.8.2753-2756.1997
896 Ru H, Liu X, Lin C, Yang J, Chen F, Sun R, Zhang L, Liu J. 2017. The Impact of
897 Genome Region of Difference 4 (RD4) on Mycobacterial Virulence and BCG
898 Efficacy. *Front Cell Infect Microbiol* **7**. doi:10.3389/fcimb.2017.00239
899 Saito H, Iwamoto T, Ohkusu K, Otsuka Y, Akiyama Y, Sato S, Taguchi O, Sueyasu Y,
900 Kawabe Y, Fujimoto H, Ezaki T, Butler R. 2011. Mycobacterium shinjukuense sp.
901 nov., a slowly growing, non-chromogenic species isolated from human clinical
902 specimens. *Int J Syst Evol Microbiol* **61**:1927–1932. doi:10.1099/ijs.0.025478-0

903 Seemann T. 2014. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*
904 **30**:2068–2069. doi:10.1093/bioinformatics/btu153
905 Singh J, O'Donnell K, Ashouri N, Adler-Shohet FC, Nieves D, Tran MT, Arrieta A, Tran
906 L, Cheung M, Zahn M. 2018. 926. Outbreak of Invasive Nontuberculous
907 Mycobacterium (NTM) Infections Associated With a Pediatric Dental Practice .
908 *Open Forum Infect Dis* **5**:S29–S29. doi:10.1093/ofid/ofy209.067
909 Slayden RA, Dawson CC, Cummings JE. 2018. Toxin-antitoxin systems and regulatory
910 mechanisms in Mycobacterium tuberculosis. *Pathog Dis*.
911 doi:10.1093/femspd/fty039
912 Stamatakis A, Ludwig T, Meier H. 2005. RAxML-III: A fast program for maximum
913 likelihood-based inference of large phylogenetic trees. *Bioinformatics* **21**:456–463.
914 doi:10.1093/bioinformatics/bti191
915 Stavrum R, Stavrum AK, Valvatne H, Riley LW, Ulvestad E, Jonassen I, Aßmus J,
916 Doherty TM, Grewal HMS. 2011. Modulation of transcriptional and inflammatory
917 responses in murine macrophages by the Mycobacterium tuberculosis Mammalian
918 cell entry (Mce) 1 complex. *PLoS One* **6**. doi:10.1371/journal.pone.0026295
919 Stinear TP, Seemann T, Harrison PF, Jenkin G a, Davies JK, Johnson PDR, Abdellah Z,
920 Arrowsmith C, Chillingworth T, Churcher C, Clarke K, Cronin A, Davis P,
921 Goodhead I, Holroyd N, Jagels K, Lord A, Moule S, Mungall K, Norbertczak H,
922 Quail M a, Rabbinowitsch E, Walker D, White B, Whitehead S, Small PLC, Brosch
923 R, Ramakrishnan L, Fischbach M a, Parkhill J, Cole ST. 2008. Insights from the
924 complete genome sequence of Mycobacterium marinum on the evolution of
925 Mycobacterium tuberculosis. *Genome Res* **18**:729–41. doi:10.1101/gr.075069.107
926 Supply P, Marceau M, Mangenot S, Roche D, Rouanet C, Khanna V, Majlessi L,
927 Criscuolo A, Tap J, Pawlik A, Fiette L, Orgeur M, Fabre M, Parmentier C, Frigui W,
928 Simeone R, Boritsch EC, Debrie AS, Willery E, Walker D, Quail MA, Ma L,
929 Bouchier C, Salvignol G, Sayes F, Cascioferro A, Seemann T, Barbe V, Locht C,
930 Gutierrez MC, Leclerc C, Bentley SD, Stinear TP, Brisse S, Médigue C, Parkhill J,
931 Cruveiller S, Brosch R. 2013. Genomic analysis of smooth tubercle bacilli provides

932 insights into ancestry and pathoadaptation of Mycobacterium tuberculosis. *Nat*
933 *Genet* **45**:172–179. doi:10.1038/ng.2517
934 Surtees JA, Funnell BE. 2003. Plasmid and Chromosome Traffic Control: How ParA and
935 ParB Drive Partition. *Curr Top Dev Biol* **56**:145–180.
936 doi:10.1016/S0070-2153(03)01010-X
937 Tortoli E, Fedrizzi T, Meehan CJ, Trovato A, Grottola A, Giacobazzi E, Serpini GF,
938 Tagliazucchi S, Fabio A, Bettua C, Bertorelli R, Frascaro F, De Sanctis V, Pecorari
939 M, Jousson O, Segata N, Cirillo DM. 2017. The new phylogeny of the genus
940 Mycobacterium: The old and the news. *Infect Genet Evol* **56**:19–25.
941 doi:10.1016/j.meegid.2017.10.013
942 Tortoli E, Pecorari M, Fabio G, Messinò M, Fabio A. 2010. Commercial DNA probes for
943 mycobacteria incorrectly identify a number of less frequently encountered species. *J*
944 *Clin Microbiol*. doi:10.1128/JCM.01536-09
945 Uchiya KI, Takahashi H, Nakagawa T, Yagi T, Moriyama M, Inagaki T, Ichikawa K,
946 Nikai T, Ogawa K. 2015. Characterization of a novel plasmid, pMAH135, from
947 Mycobacterium avium subsp. hominissuis. *PLoS One* **10**.
948 doi:10.1371/journal.pone.0117797
949 Ummels R, Abdallah AM, Kuiper V, Aâjoud A, Sparrius M, Naeem R, Spaink HP, Bitter
950 W, Pain A, Bitter W. 2014. Identification of a novel conjugative plasmid in
951 mycobacteria that requires both type IV and type VII secretion. *MBio* **5**.
952 doi:10.1128/mBio.01744-14
953 Van der Werf MJ, K??dm??n C, Katalini??-Jankovi?? V, Kummik T, Soini H, Richter E,
954 Papaventsis D, Tortoli E, Perrin M, van Soolingen D, ??olnir-Dov?? M, ??stergaard
955 Thomsen V. 2014. Inventory study of non-tuberculous mycobacteria in the
956 European Union. *BMC Infect Dis* **14**. doi:10.1186/1471-2334-14-62
957 Van Der Woude AD, Sarkar D, Bhatt A, Sparrius M, Raadsen SA, Boon L, Geurtsen J,
958 Van Der Sar AM, Luirink J, Houben ENG, Besra GS, Bitter W. 2012. Unexpected
959 link between lipooligosaccharide biosynthesis and surface protein release in
960 Mycobacterium marinum. *J Biol Chem* **287**:20417–20429.
961 doi:10.1074/jbc.M111.336461

962 van Ingen J, Al-Haijoj SAM, Boeree M, Al-Rabiah F, Enaimi M, de Zwaan R, Tortoli E,
963 Dekhuijzen R, van Soolingen D. 2009. Mycobacterium riyadhense sp. nov., a
964 non-tuberculous species identified as Mycobacterium tuberculosis complex by a
965 commercial line-probe assay. *Int J Syst Evol Microbiol* **59**:1049–1053.
966 doi:10.1099/ijs.0.005629-0
967 Van Ingen J, Bendien SA, De Lange WCM, Hoefsloot W, Dekhuijzen PNR, Boeree MJ,
968 Van Soolingen D. 2009. Clinical relevance of non-tuberculous mycobacteria
969 isolated in the Nijmegen-Arnhem region, The Netherlands. *Thorax* **64**:502–506.
970 doi:10.1136/thx.2008.110957
971 Van Ingen J, De Zwaan R, Dekhuijzen R, Boeree M, Van Soolingen D. 2009. Region of
972 difference 1 in nontuberculous Mycobacterium species adds a phylogenetic and
973 taxonomical character. *J Bacteriol* **191**:5865–5867. doi:10.1128/JB.00683-09
974 van Ingen J, Kohl TA, Kranzer K, Hasse B, Keller PM, Katarzyna Szafrańska A,
975 Hillemann D, Chand M, Schreiber PW, Sommerstein R, Berger C, Genoni M,
976 Rüegg C, Troillet N, Widmer AF, Becker SL, Herrmann M, Eckmanns T, Haller S,
977 Höller C, Debast SB, Wolfhagen MJ, Hopman J, Kluytmans J, Langelaar M,
978 Notermans DW, ten Oever J, van den Barselaar P, Vonk ABA, Vos MC, Ahmed N,
979 Brown T, Crook D, Lamagni T, Phin N, Smith EG, Zambon M, Serr A, Götting T,
980 Ebner W, Thürmer A, Utpatel C, Spröer C, Bunk B, Nübel U, Bloemberg G V.,
981 Böttger EC, Niemann S, Wagner D, Sax H. 2017. Global outbreak of severe
982 Mycobacterium chimaera disease after cardiac surgery: a molecular epidemiological
983 study. *Lancet Infect Dis* **17**:1033–1041. doi:10.1016/S1473-3099(17)30324-9
984 Van Soolingen D, Hoogenboezem T, De Haas PEW, Hermans PWM, Koedam MA,
985 Teppema KS, Brennan PJ, Besra GS, Portaels F, Top J, Schouls LM, Van Embden
986 JDA. 1997. A Novel Pathogenic Taxon of the Mycobacterium tuberculosis Complex,
987 Canetti: Characterization of an Exceptional Isolate from Africa. *Int J Syst Bacteriol*
988 **47**:1236–1245. doi:10.1099/00207713-47-4-1236
989 Veyrier F, Pletzer D, Turenne C, Behr MA. 2009. Phylogenetic detection of horizontal
990 gene transfer during the step-wise genesis of Mycobacterium tuberculosis. *BMC*
991 *Evol Biol* **9**. doi:10.1186/1471-2148-9-196

992 Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q,
993 Wortman J, Young SK, Earl AM. 2014. Pilon: An integrated tool for comprehensive
994 microbial variant detection and genome assembly improvement. *PLoS One* **9**.
995 doi:10.1371/journal.pone.0112963
996 Wang J, McIntosh F, Radomski N, Dewar K, Simeone R, Enninga J, Brosch R, Rocha EP,
997 Veyrier FJ, Behr MA. 2015. Insights on the emergence of Mycobacterium
998 tuberculosis from the analysis of Mycobacterium kansasii. *Genome Biol Evol* **7**:856–
999 870. doi:10.1093/gbe/evv035
1000 Weerdenburg EM, Abdallah AM, Mitra S, De Punder K, Van der Wel NN, Bird S,
1001 Appelmelk BJ, Bitter W, Van der Sar AM. 2012. ESX-5-deficient Mycobacterium
1002 marinum is hypervirulent in adult zebrafish. *Cell Microbiol* **14**:728–739.
1003 doi:10.1111/j.1462-5822.2012.01755.x
1004 Wu M, Scott AJ. 2012. Phylogenomic analysis of bacterial and archaeal sequences with
1005 AMPHORA2. *Bioinformatics* **28**:1033–1034. doi:10.1093/bioinformatics/bts079
1006 Zrimec J, Lapanje A. 2018. DNA structure at the plasmid origin-of-Transfer indicates its
1007 potential transfer range. *Sci Rep* **8**. doi:10.1038/s41598-018-20157-y
1008

1009

1010 **Data Availability**

1011 The *M. riyadhense* dataset is available at European Nucleotide Archive (ENA) under the
1012 study accession no. PRJEB32162. The assemblies are available with DOI
1013 10.5281/zenodo.2873972.

1014

1015 **Funding**

1016 Work in AP's laboratory is supported by the KAUST faculty baseline fund
1017 (BAS/1/1020-01- 01).

1018

1019 **Conflicts of Interest**

1020 The authors have no conflicts of interest to declare.

1021

1022 **Acknowledgement**

1023 We would like to thank members of the Biological Core Lab (BCL) of King Abdullah

1024 University of Science and Technology for sequencing the bacteria on the Illumina HiSeq

1025 4000 and PacBio RSII platforms. The research protocol was approved by the Institutional

1026 Review Board of King Fahad Medical City (Riyadh, Saudi Arabia; #16-345) and the

1027 Institutional Biosafety and Bioethics Committee of King Abdullah University of Science

1028 and Technology (Jeddah, Saudi Arabia; #18IBEC23).

1029

1030 **Author Contributions Statement**

1031 AP conceived the comparative genomics part of the study, obtained the funding and

1032 supervised the work; MG and AH conceived the clinical part of the study; MG and FA

1033 wrote the clinical part of the text; FA, MG, TA, SF, AH, MAR, AR, and TS collected the

1034 microbiological and clinical information; and AP, QG, SG, AB, and FA designed the

1035 experiments. QG performed the data analysis and prepared the initial draft of the

1036 manuscript, followed by edits from AP, SG, MG, SF and AB. QG, SM, ASm and JB

1037 performed the transcriptome experiment. ASi and AB performed the TLC analysis. CN

1038 and YS provided materials for the diagnostic markers and intellectual advice. All authors

have commented on various sections of the manuscript, which were finally curated and incorporated in the final version by QG and AP.

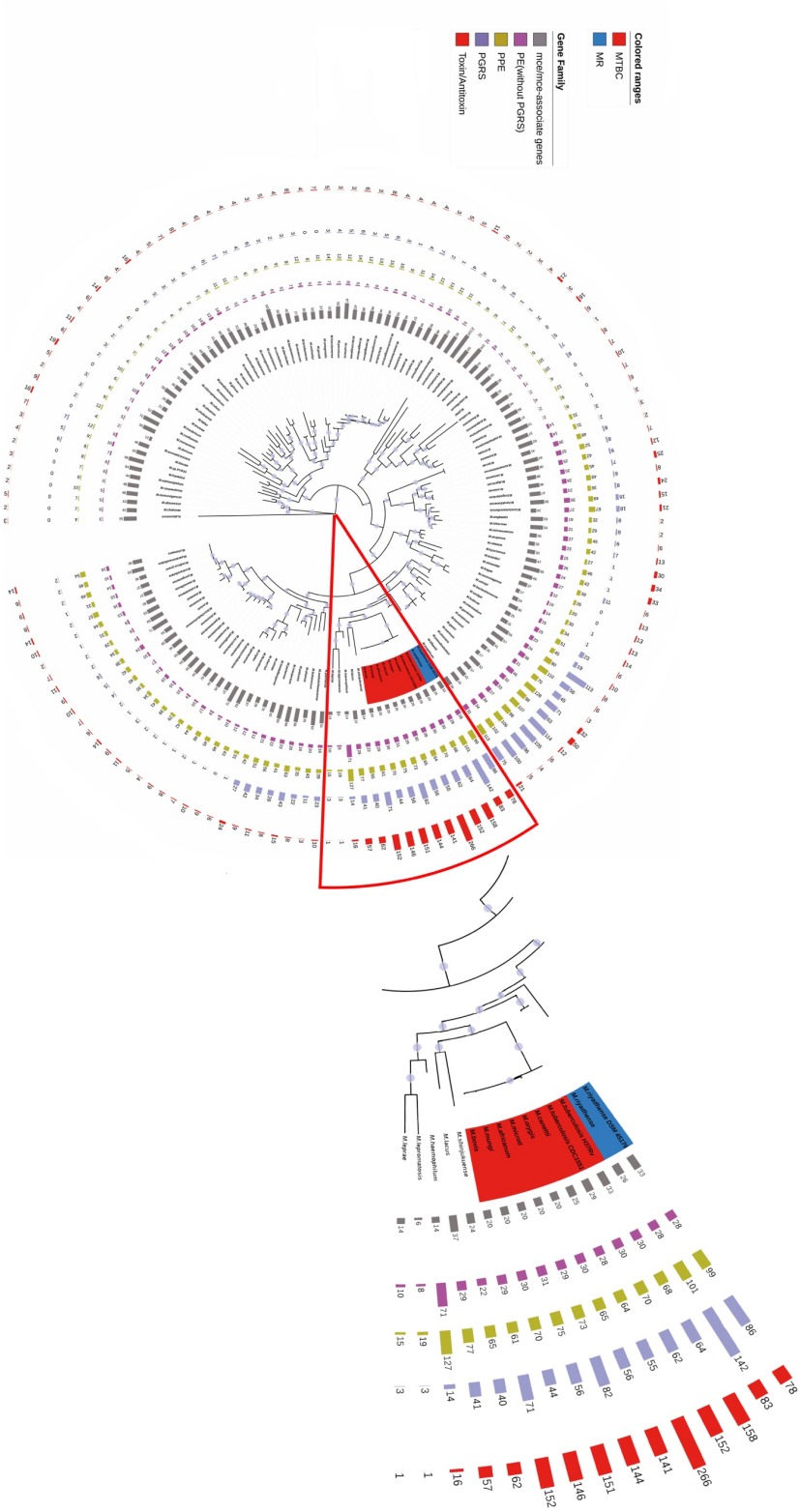

Colored ranges
MTBC
MR
Gene Family
mce/mce-associate genes
PE(without PGRS)
PPE
PGRS
Toxin/Antitoxin

Figure.1

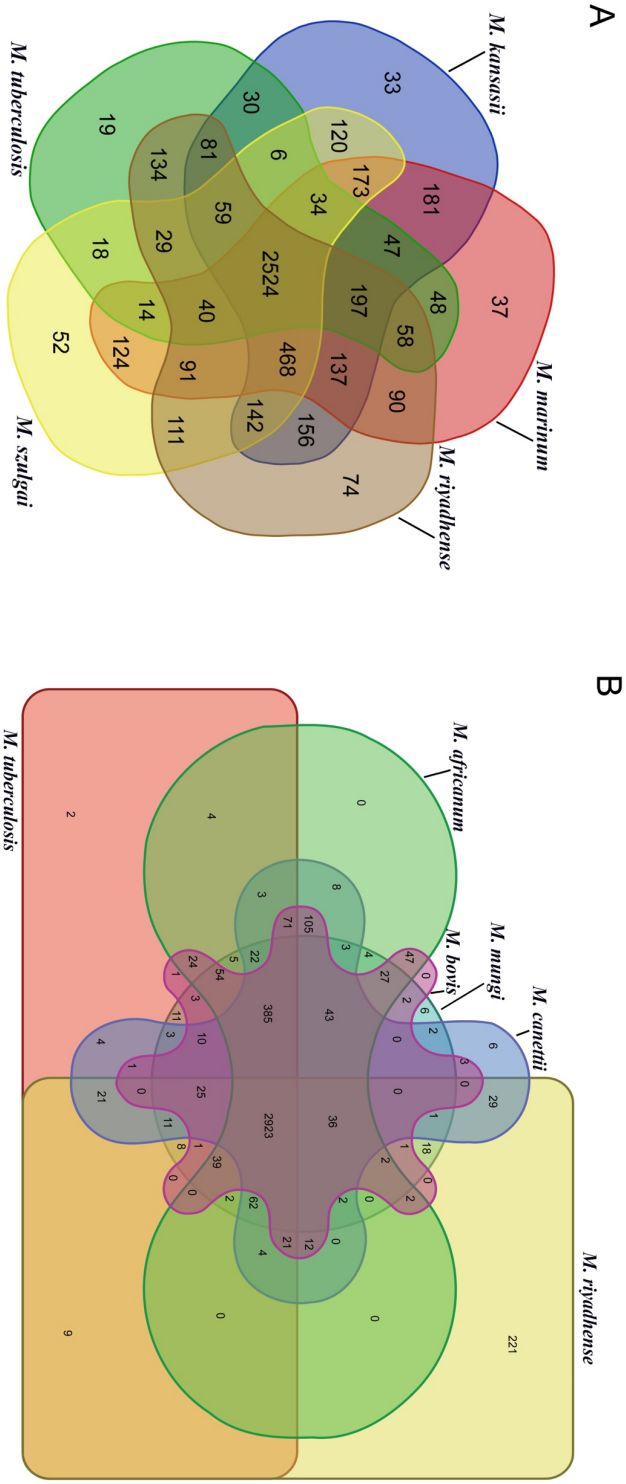

A
M. kansasii
M. marinum
M. riyadhense
M. szulgai
M. tuberculosis
33
30
19
120
6
81
134
173
181
34
59
29
18
47
2524
197
48
37
14
40
58
52
124
468
137
91
90
142
156
111
74
B
M. africanum
M. mungi
M. bovis
M. canettii
M. riyadhense
M. tuberculosis
0
4
2
8
3
105
71
3
4
47
0
27
2
6
2
6
3
0
29
0
0
1
18
1
2
0
2
0
2
0
12
21
4
0
0
0
221
9
22
5
54
24
1
3
11
3
10
4
1
0
25
21
11
8
1
39
0
0
2
62
385
43
2923
36

Figure.2

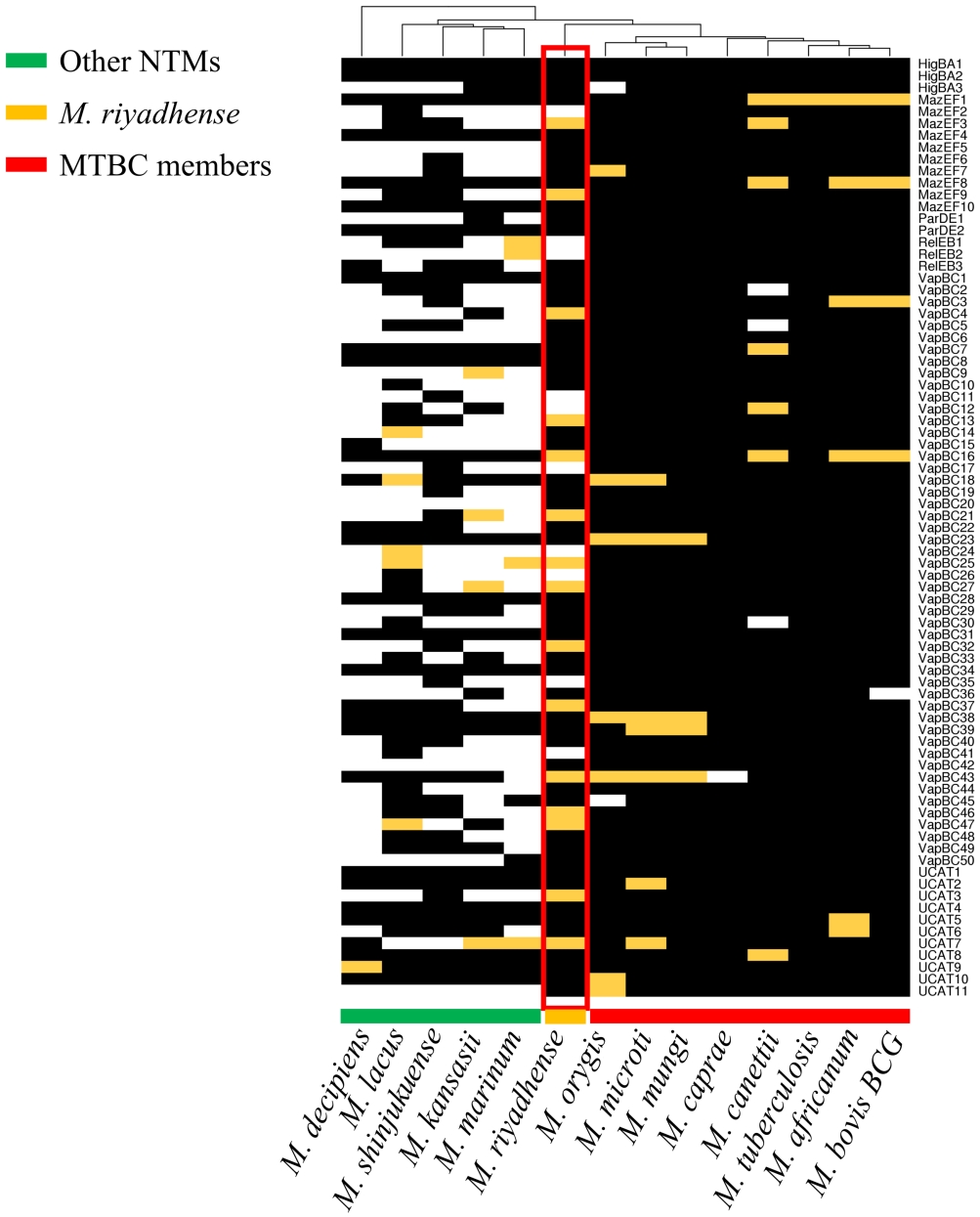

Other NTMs
M. riyadhense
MTBC members
HigBA1
HigBA2
HigBA3
MazEF1
MazEF2
MazEF3
MazEF4
MazEF5
MazEF6
MazEF7
MazEF8
MazEF9
MazEF10
ParDE1
ParDE2
RelEB1
RelEB2
RelEB3
VapBC1
VapBC2
VapBC3
VapBC4
VapBC5
VapBC6
VapBC7
VapBC8
VapBC9
VapBC10
VapBC11
VapBC12
VapBC13
VapBC14
VapBC15
VapBC16
VapBC17
VapBC18
VapBC19
VapBC20
VapBC21
VapBC22
VapBC23
VapBC24
VapBC25
VapBC26
VapBC27
VapBC28
VapBC29
VapBC30
VapBC31
VapBC32
VapBC33
VapBC34
VapBC35
VapBC36
VapBC37
VapBC38
VapBC39
VapBC40
VapBC41
VapBC42
VapBC43
VapBC44
VapBC45
VapBC46
VapBC47
VapBC48
VapBC49
VapBC50
UCAT1
UCAT2
UCAT3
UCAT4
UCAT5
UCAT6
UCAT7
UCAT8
UCAT9
UCAT10
UCAT11
M. decipiens
M. lacus
M. shinjukuense
M. kansasii
M. marinum
M. riyadhense
M. orygis
M. microti
M. mungi
M. caprae
M. canettii
M. tuberculosis
M. africanum
M. bovis BCG

Figure.3

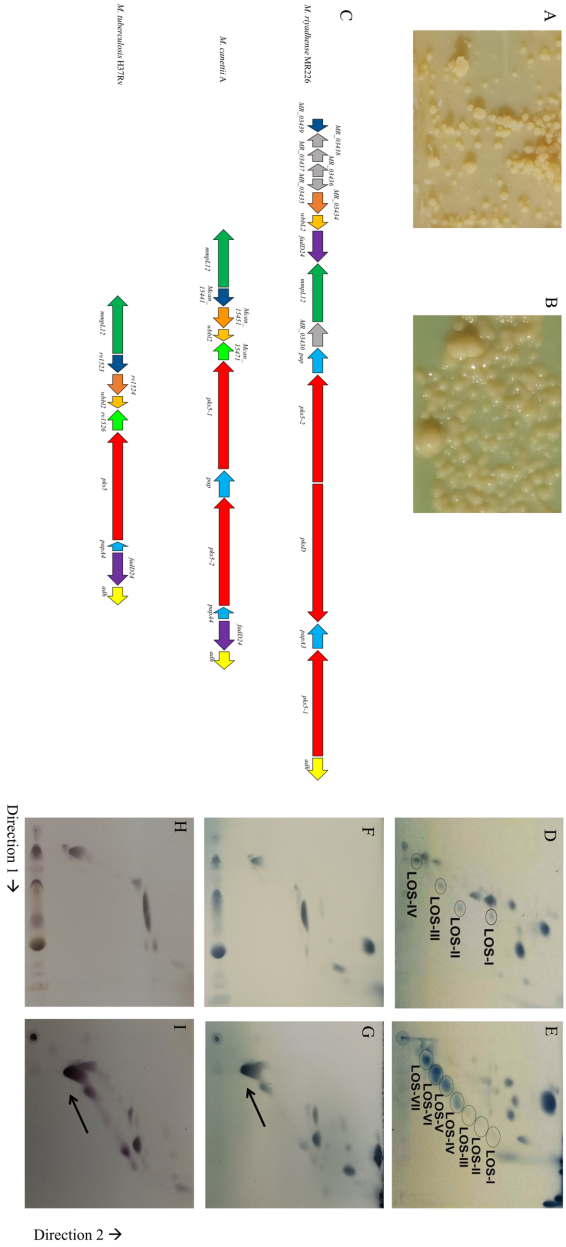

A
B
C
M. rhyadhense MR226
M. canettii A
M. tuberculosis H37Rv
D
E
F
G
H
I
LOS-I
LOS-II
LOS-III
LOS-IV
LOS-V
LOS-VI
LOS-VII
Direction 1 →
Direction 2 →

Figure.4

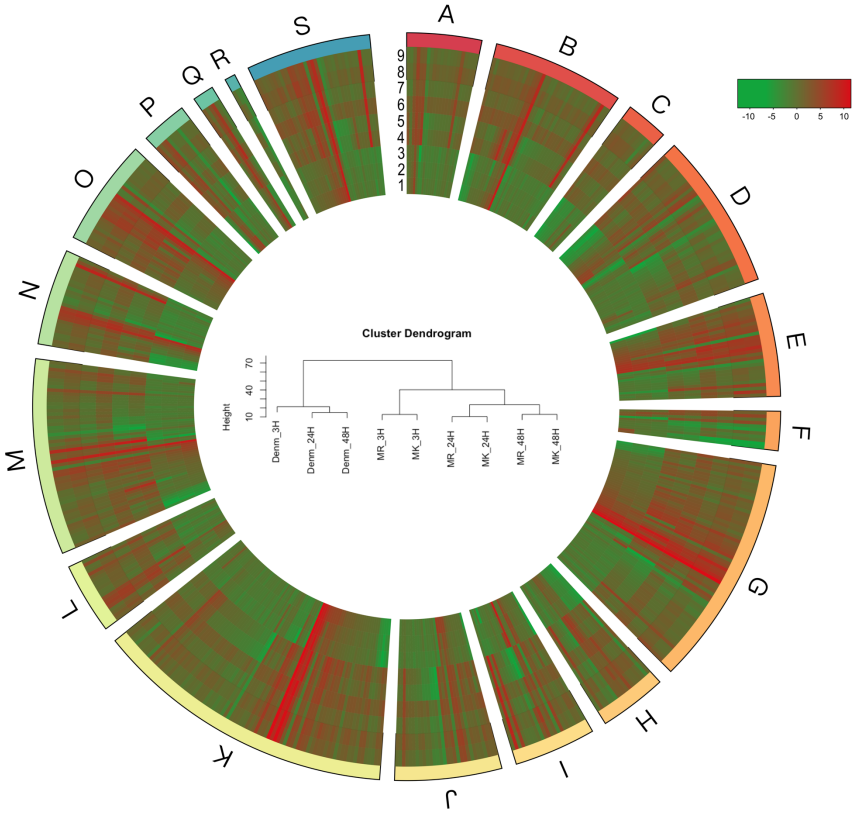
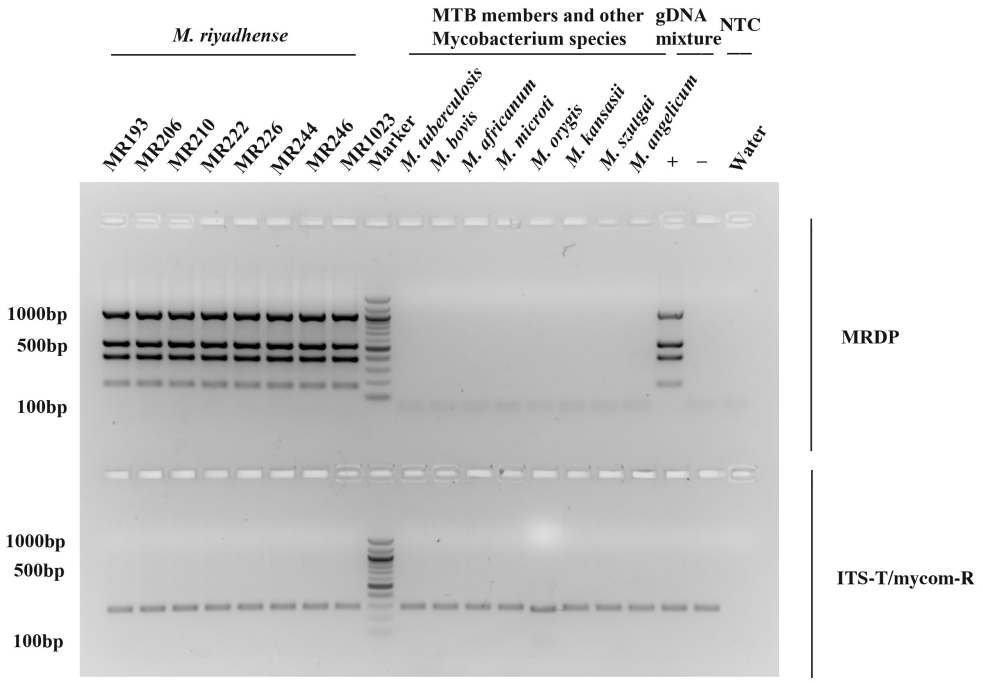

Cluster Dendrogram
Height
10
40
70
Denm_3H
Denm_24H
Denm_48H
MR_3H
MK_3H
MR_24H
MK_24H
MR_48H
MK_48H
-10
-5
0
5
10
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
1
2
3
4
5
6
7
8
9

Figure.5

1085



1086

1087

1088

1089

1090

1091

1092

1093

1094 Figure.6

1095 Captions for the figures:

1096 Figure.1: A hypothesis for relationships of *M. riyadhense* and MTBC among
1097 Mycobacterium genus. 12 marker genes from the chromosome used for
1098 generating the phylogeny are selected by Amphora2, Dayhoff model
1099 selected by Modelgenerator are applied to generated this phylogeny using
1100 RaxML. The circle on each branch represent the bootstrap values. Part of the
1101 phylogenetic tree was highlighted for illustration purpose.

1102

1103 Figure.2: Genome content comparison. Venn diagrams represents the
1104 overlap of gene orthologs between the (A) *M. riyadhense*, *M. tuberculosis*
1105 H37Rv, *M. marinum* M, *M. kansasii* 12478 and *M. szulgai* and (B) *M.*
1106 *riyadhense* and five species within MTBC.

1107

1108 Figure.3: The presence (black) and absence(while) of *M. tuberculosis*
1109 H37Rv toxin/antitoxin orthologs in *M. riyadhense*, *M. marinum*, *M. kansasii*,
1110 *M. shinjukuense*, *M. lacus*, *M. decipiens* and MTBC species. The orange
1111 color represent only toxin or antitoxin protein ortholog has been found in
1112 that pair of toxin/antitoxin system. T/A systems within *M. riyadhense* are
1113 highlighted in the red box.

1114

1115 Figure.4: LOS systems in *M. riyadhense* and other related mycobacteria and
1116 2D-TLC analysis of polar lipids extracted from selected *M. riyadhense*
1117 strains. (A) Genetic locus map of the LOS biosynthesis gene cluster from *M.*
1118 *riyadhense*, *M. canettii* A and *M. tuberculosis* H37Rv (drawn to scale). The
1119 arrows showing the direction of transcription and the genes are colored
1120 according to the orthologous relationships. (B) Rough-dry colony
1121 morphology (MR193) and (C) smooth morphology (MR226) of *M.*
1122 *riyadhense*. Polar lipids from two known LOS producers, *M. marinum* (D)
1123 and *M. kansasii* (E) are included to show the migration pattern of LOS
1124 species in System E. (F) 2D-TLC analysis of polar lipids extracted from
1125 select *M. riyadhense* rough strain or smooth (G) strain. A separate staining
1126 with alpha-napthol also confirmed that this was a glycolipid species from the
1127 same (H) rough and (I) smooth strain. (D), (E), (F) and (G) were charred
1128 after staining with MPA, while (H) and (I) were charred after staining with
1129 alpha napthol. LOS III from *M. riyadhense* is indicated by a solid arrow.
1130

1131 Figure. 5 This circular heatmap demonstrates the expression values from
1132 RAW264.7 infections with *M. riyadhense* (MR), *M. kansasii* (MK) and *M.*
1133 *bovis* BCG Denmark (Denm) over 3, 24 and 48 hrs. The gene sets of each
1134 clusters are: A: Angiogenesis; B: Antigen Presentation; C: Cell Cycle and

1135 Apoptosis; D: Cell Migration and Adhesion; E: Chemokine signaling; F:
1136 Complement Activation; G: Cytokine Signaling; H: Differentiation and
1137 Maintenance of Myeloid Cells; I: ECM remodeling; J: Fc Receptor
1138 Signaling; K: Growth Factor Signaling; L: Interferon Signaling; M:
1139 Lymphocyte activation; N: Metabolism; O: Pathogen Response; P: T-cell
1140 Activation and Checkpoint Signaling; Q: TH1 Activation; R: TH2
1141 Activation; S: TLR signaling. The cycles 1-9 represent RAW264.7
1142 transcriptome response upon infections of 1: Denm-3h; 2: Denm-24h; 3:
1143 Denm-48h; 4: MR-3h; 5: MK-3h; 6: MR-24h; 7: MK-24h; 8: MR-48h; 9:
1144 MK-48h.
1145
1146 Figure.6: Agarose gel (2%) electrophoresis analysis of a PCR diagnostic test
1147 for *M. riyadhense*.
1148 Lane M: DNA Marker, Lane 1-8: Varies *M. riyadhense* strains (From left to
1149 right: MR193, MR206, MR210, MR222, MR226, MR244, MR246,
1150 MR1023), Lane 9-16: Varies *Mycobacterium* species (From left to right: *M.*
1151 *tuberculosis*, *M. bovis*, *M. africanum*, *M. microti*, *M. oryis*, *M. kansasii*, *M.*
1152 *szulgai* and *M. angelicum*) template, Template cocktail, mycobacterium
1153 species (*M. tuberculosis*, *M. bovis*, *M. kansasii*, *M. marinum*, *M. szulgai*, *M.*
1154 *avium* and *M. angelicum*) with (Lane 17, +) and without (Lane 18,-) *M.*

1155 *riyadhense* MR226 gDNA template. Lane 19: Non-template control (NTC).

1156 Upper Panel: MRDP (*M. riyadhense* diagnostic marker) set. Lower Panel B:

1157 *Mycobacterium* genus specific primer ITS-T and mycom-R amplified with

1158 mycobacterial gDNA.

1159

1160

1161

1162

1163

1164

1165

1166

1167

1168

1169

1170

1171

1172

1173

1174

1175

1176

1177

Tables:

Table.1 Clinical characteristics of the studied population

	Age (Years)	Gender	HIV status	Other co-morbid conditions	Site(s) of the infection	Smooth/Rough	Auramine-rhodamine stain	Ziehl–Neelsen stain	Therapy		Radiographic Location of lesion (Chest CT)	Baseline		Treatment Outcome
									Initial medication/duration	Continuous medication/duration		CD4 count ($X10^3$/L)	Viral load (Copies/mL)	
MR193	82	M	Unknown	Smoking; DM; HTN; heart disease	Pulmonary	R	4+	2+	NA	NA	Left upper lobe consolidation	NA	N/A	Died
MR206	32	M	Negative	Eisenmenger's syndrome; pulmonary HTN	Pulmonary	R	NA	1+	IMP/2M CLR/2M EMB/2M INH/1M PYZ/1M	CLR/4M EMB/4M	Bilateral air spice consolidation and nodular opacity	NA	N/A	Survived
MR210	66	M	Positive	PCP co-infection	Pulmonary	S	NA	Negtive	RIF/1M EMB/18M CLR/18M	NA	Bilateral ground glass opacities	0.7	415,927	Survived
MR222	37	M	Positive	Lymphoblastic lymphoma	Pulmonary; LN; Abdominal	S	NA	Negtive	CLR/1M EMB/1M IMP/1M	EMB/2M IMP/2M	Multiple lung lesions with central cavitation	0	838	Died
MR226	8	M	Negative	Free	Pulmonary; LA	S	NA	Negtive	NA	NA	Multiple bilateral nodular infiltrates, more in the left upper lobe	NA	N/A	Survived
MR244	28	M	Positive	Pulmonary disease; hematologic disease	Pulmonary; LN; Abdominal; Peritoneal;	S	NA	2+	CLR/1M RIF+INH/1M EMB/1M	MXF/4M	Multiple bilateral lung nodules; cavitation left lower lobe	0.02	399,652	Survived
MR246	17	M	Negative	Pulmonary disease; chronic diarrhea	Pulmonary; Mediastinal LN	S	NA	Negative	RIF/17M CLR/21M EMB/17M	NA	Left upper lobe consolidation	NA	N/A	Survived
MR1023	47	M	Negative	DM; subclinical hypothyroidism	Pulmonary	S	NA	Negative	NA	NA	Bilateral nodular infiltrate; cavity,left upper lobe	NA	N/A	Survived

Table.2 Comparison of general features of genome in Mycobacteria

Features	*M. tuberculosis*	*M. africanum*	*M. bovis*	*M. canettii*	*M. riyadhense*	*M. kansasii*	*M. marinum*
Genome Size (M)	4.41	4.39	4.26	4.48	Chromosome: 6.25 Plasmid pMRLP01: 0.55 Plasmid pMR01: 0.09	Chromosome: 6.43 Plasmid pMK12478: 0.14	Chromosome: 6.34 Plasmid pRAW: 0.11
GC content (%)	65.6	65.6	65.6	65.6	Chromosome: 65.35 Plasmid pMRLP01: 67.36 Plasmid pMR01: 65.69	Chromosome: 66.2 Plasmid pMK12478: 65.8	Chromosome: 65.8 Plasmid pRAW: 64.1
CDS	3,906	4,045	4,042	4,064	Chromosome: 5,490 Plasmid pMRLP01: 443 Plasmid pMR01: 88	Chromosome: 5,631 Plasmid pMK12478: 152	Chromosome: 5,149 Plasmid pRAW: 94
Gene Density(bp)	1101	1085	1032	1044	Chromosome: 1126 Plasmid pMRLP01: 1239 Plasmid pMR01: 1072	Chromosome: 1142 Plasmid pMK12478: 954	Chromosome: 1176 Plasmid pRAW: 1166
Conserved protein with assigned function	2855 (73%)	3078 (76%)	3072 (76%)	3125 (77%)	Chromosome: 4079 (74%) Plasmid pMRLP01: 217 (49%) Plasmid pMR01: 22 (25%)	Chromosome: 4468 (79%) Plasmid pMK12478:74 (49%)	Chromosome: 4091 (79%) Plasmid pRAW: 38 (40%)
Unique CDS	62 (1.59%)	25 (0.62%)	40 (0.99%)	83 (2.04%)	Chromosome:1059(19.29%) Plasmid pMRLP01:293 (66.14%) Plasmid pMR01:32 (36.36%)	Chromosome: 1,080 (19.18%) Plasmid pMK12478: 61 (40.13%)	Chromosome: 993 (19.29%) Plasmid pMK12478: 42 (44.68%)
tRNA	45	45	47	45	Chromosome: 51 Plasmid pMRLP01:1	Chromosome: 46 Plasmid: NA	Chromosome: 47 Plasmid: NA
rRNA	3	3	3	3	Chromosome: 3 Plasmid: NA	Chromosome: 3 Plasmid: NA	Chromosome: 3 Plasmid: NA
Transposase /integrase	53	68	56	98	Chromosome:110 Plasmid pMR01:1 Plasmid pMRLP01: 72	Chromosome:79 Plasmid pMK12478:3	Chromosome:41 Plasmid pRAW:4
Reference	1	2	3	4	MR226 This Study	5	6

1181

1182

1183

Table.3 Presence and absence of the *M. marinum* orthologs related to LOS synthesis in *M. riyadhense* and other species.

	M. smegmatis	*M. kansasii*	*M. canettii*	*M. riyadhense*	*M. tuberculoisis* H37Rv
MMAR_1008	P	P	P	P	P
MMAR_2307	A	P	A	P	A
MMAR_2313	A	P	P	P	P
MMAR_2319	A	A	A	A	A
MMAR_2320 (*wecE*)	A	P	P	A	A
MMAR_2327	A	P	P	P	P
MMAR_2332	A	A	A	A	A
MMAR_2336 (*galE*)	A	P	A	A	A
MMAR_2340 (*pks*)	P	P	P	P	T
MMAR_2341	P	P	P	P	P
MMAR_2343(*pap*)	P	P	P	P	A

1184

1185

1186

1187

1188

1189

1190

1191

1192

1193

1194

1195

Supplementary Files:

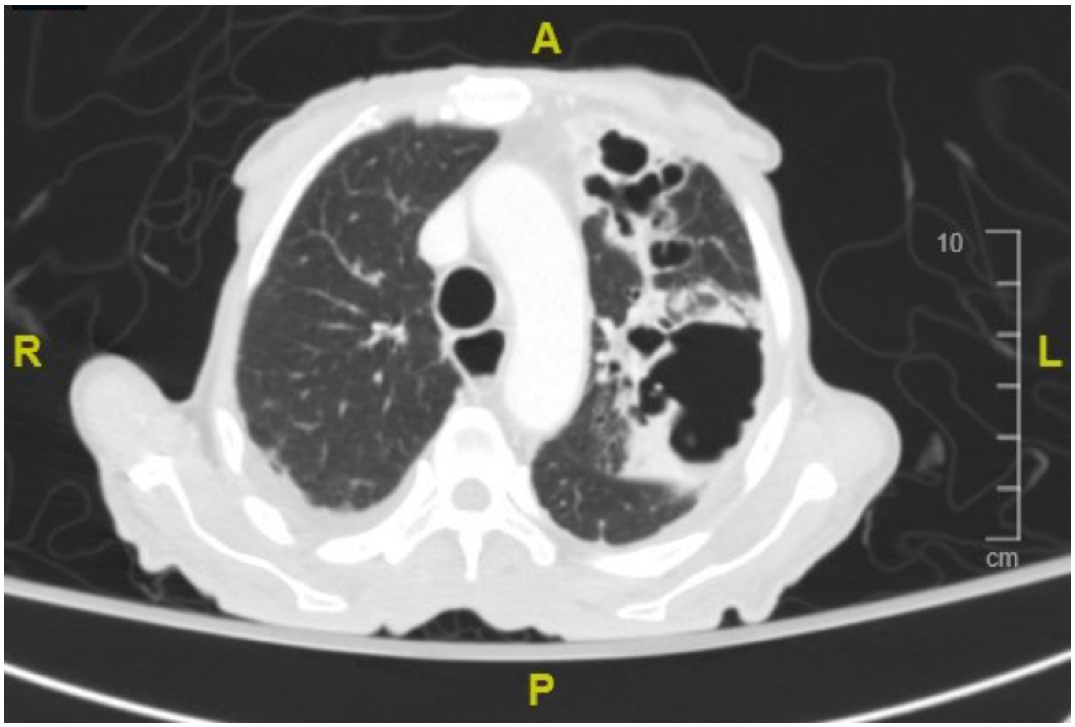


Figure supplementary 1

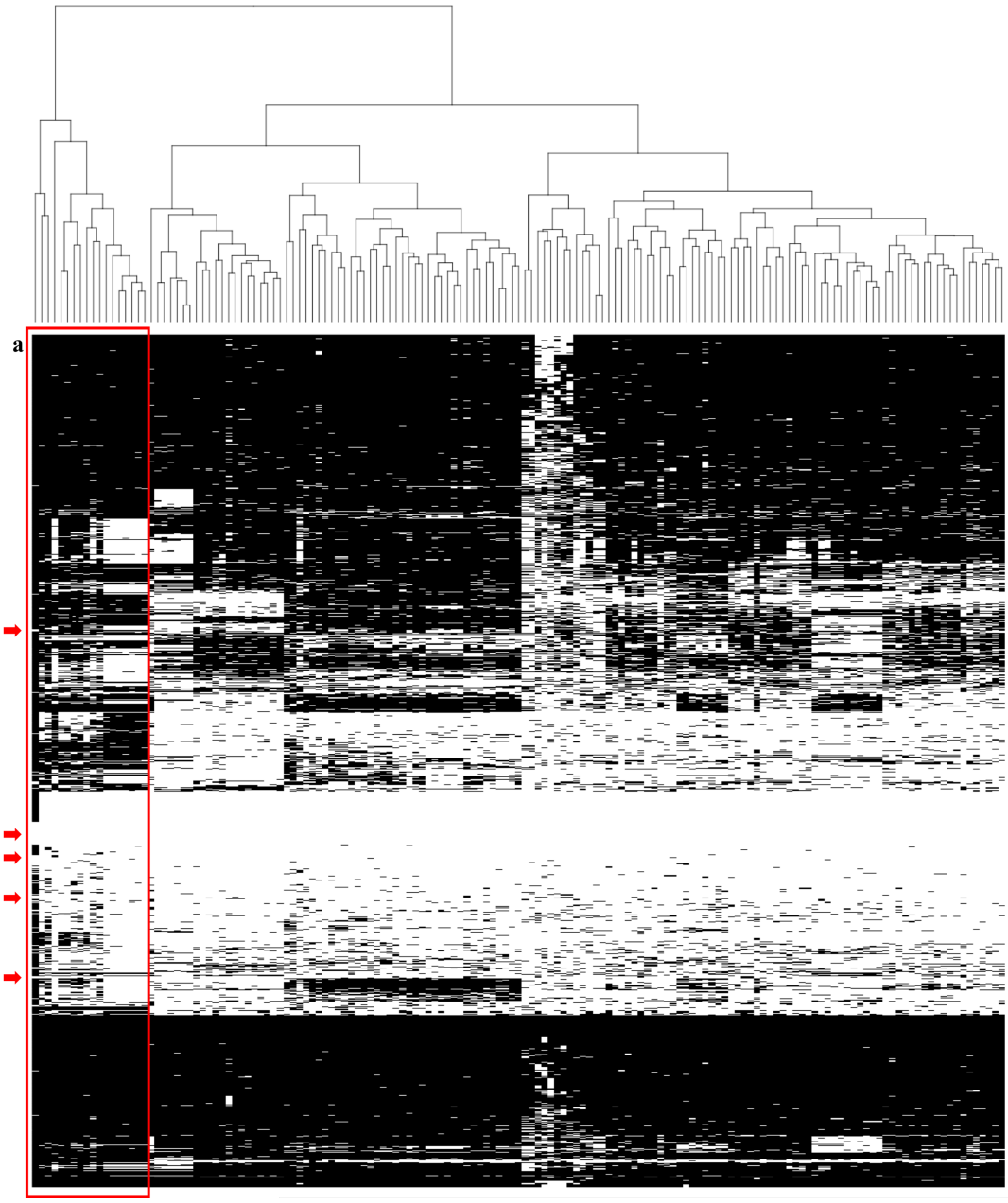


Figure supplementary 2A

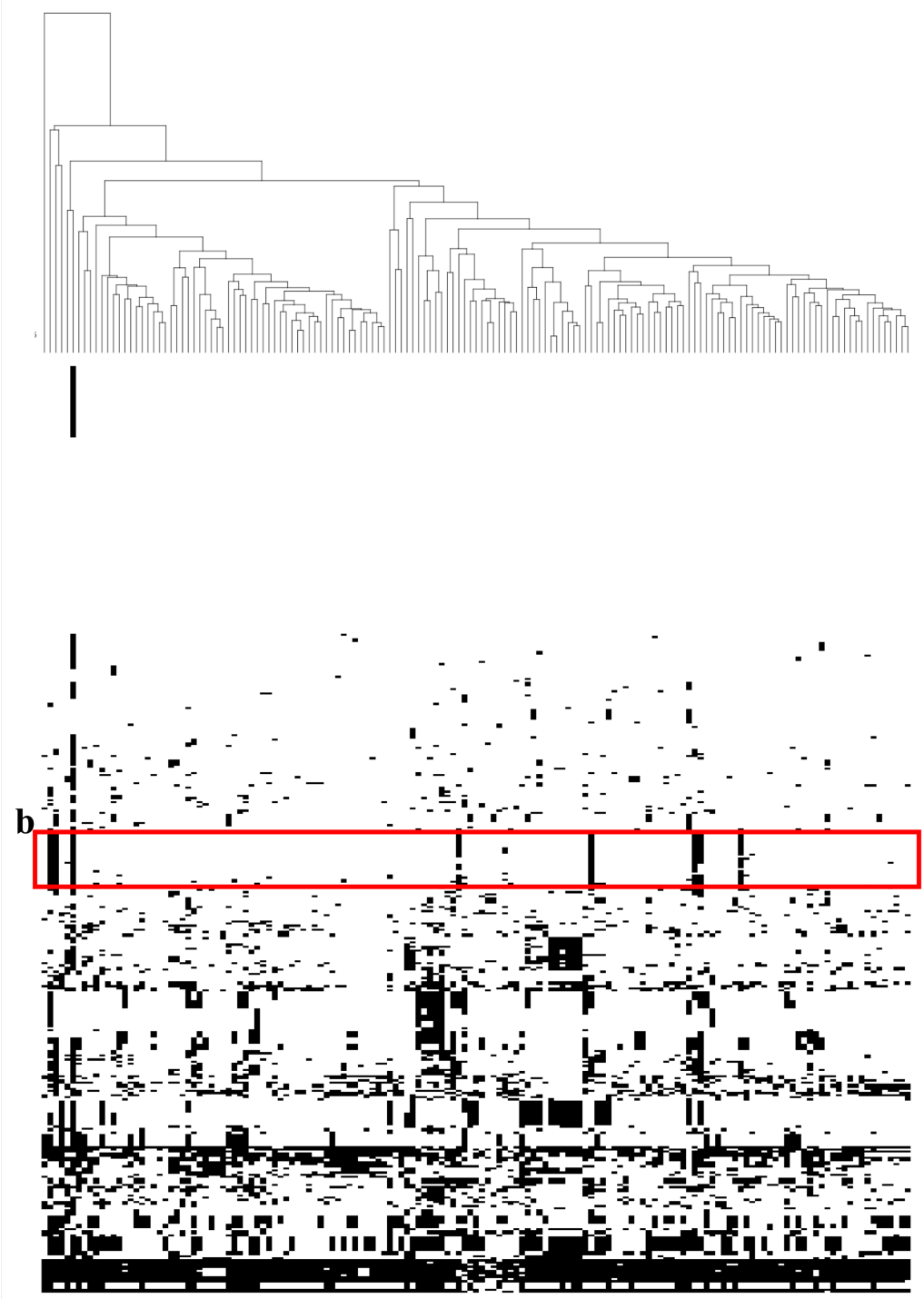


Figure supplementary 2B

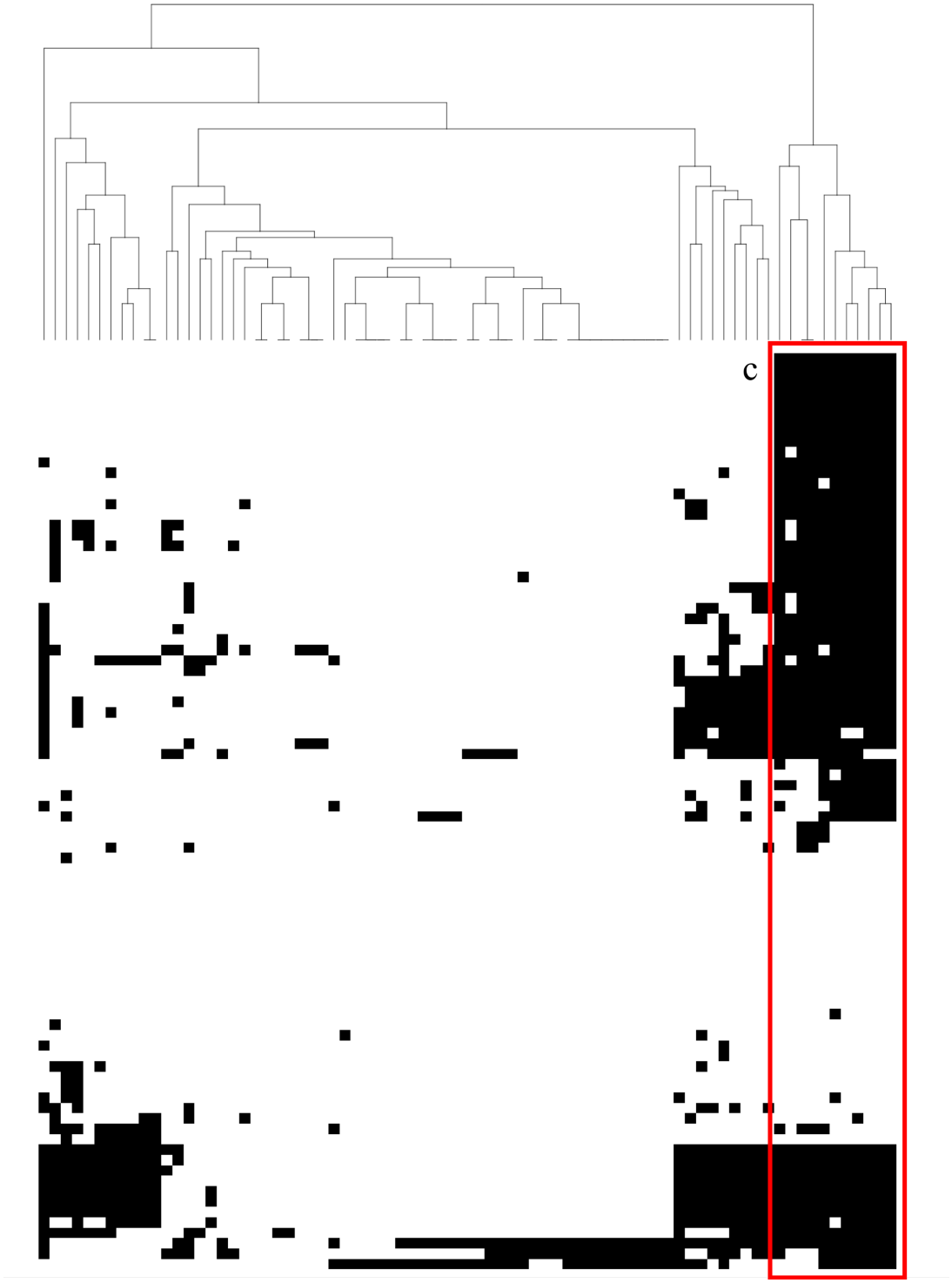


Figure supplementary 2C

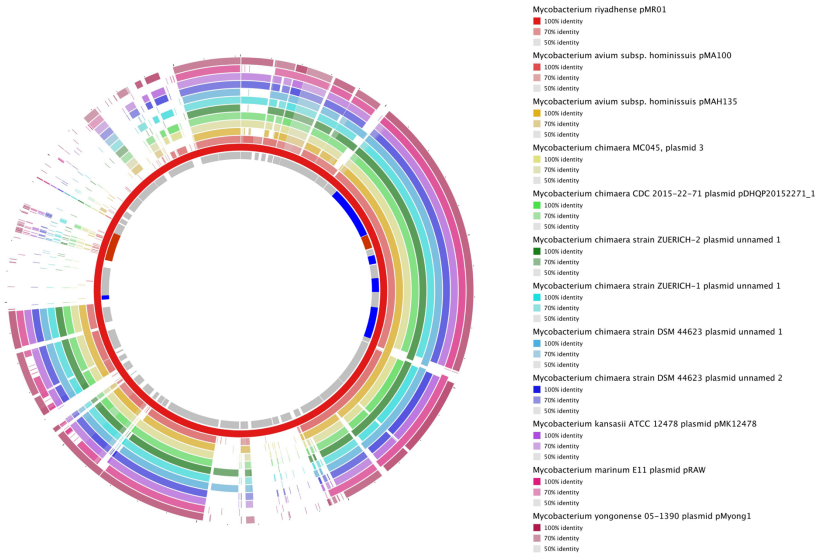


Figure supplementary 3

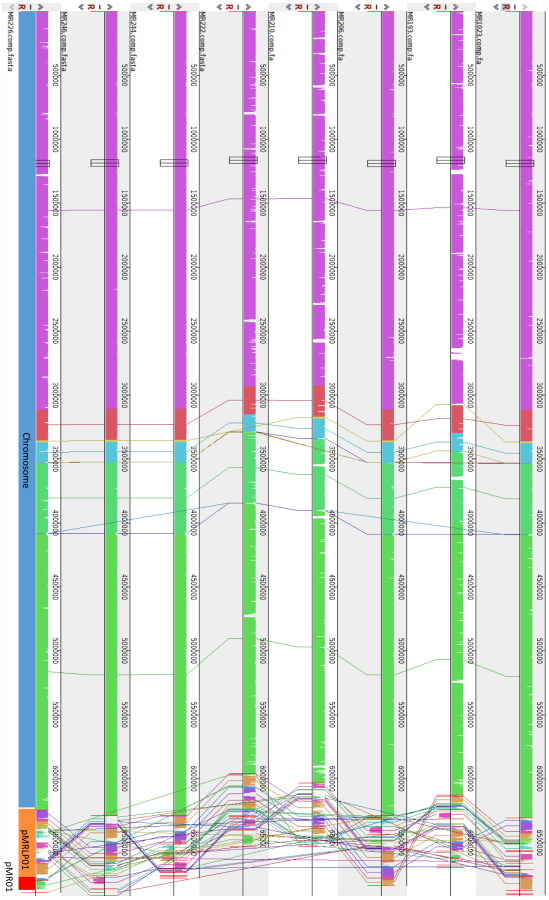


Figure supplementary 4A

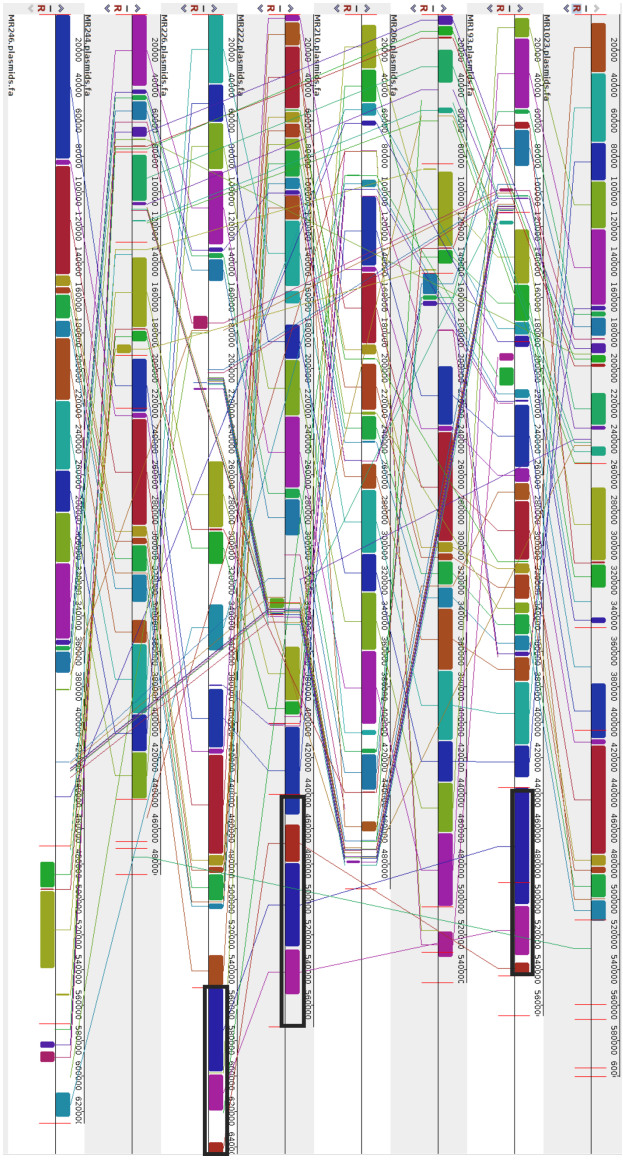


Figure supplementary 4B

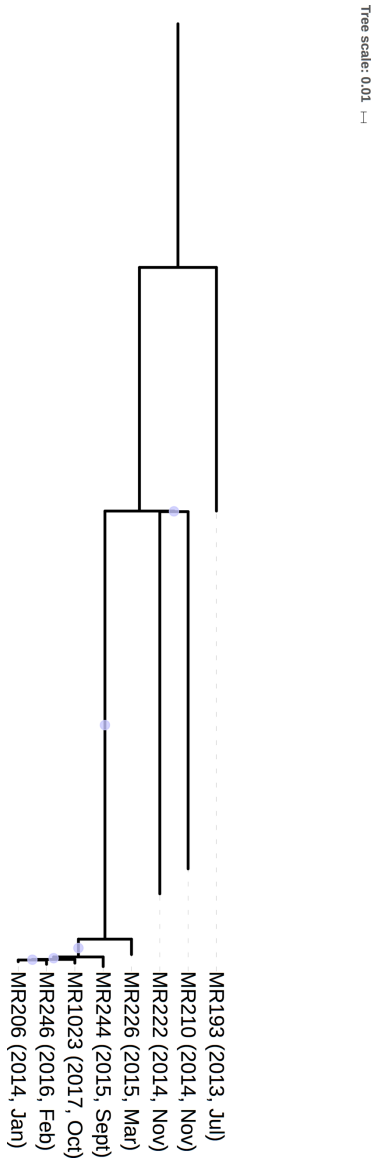


Figure supplementary 5

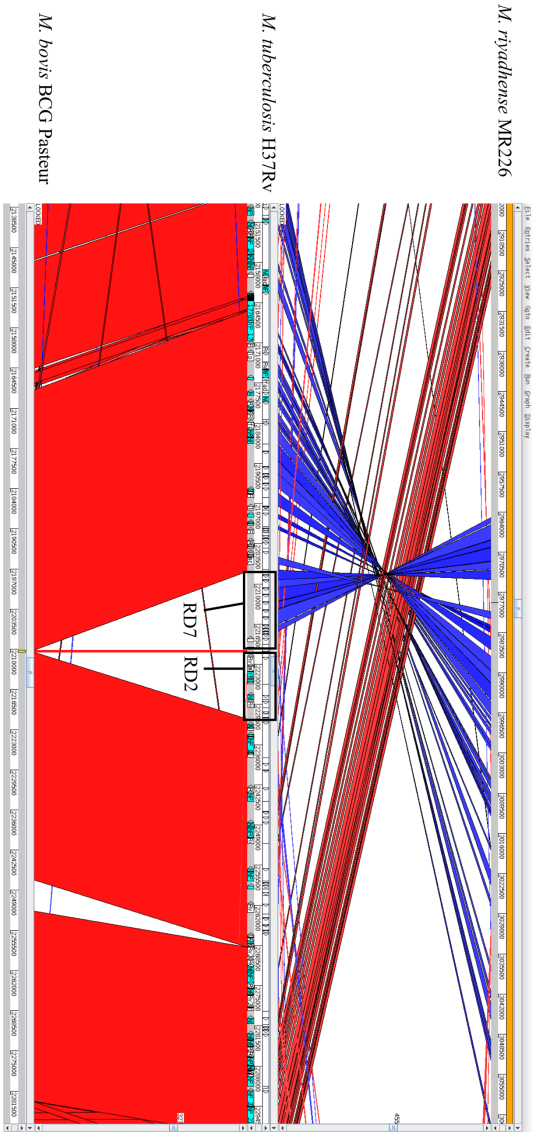


Figure supplementary 6A

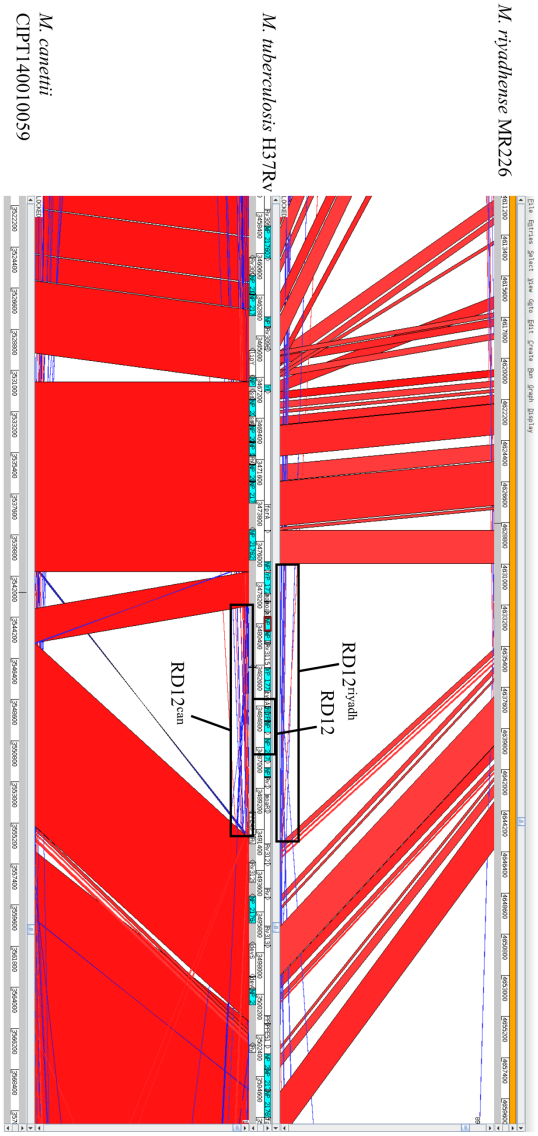


Figure supplementary 6B

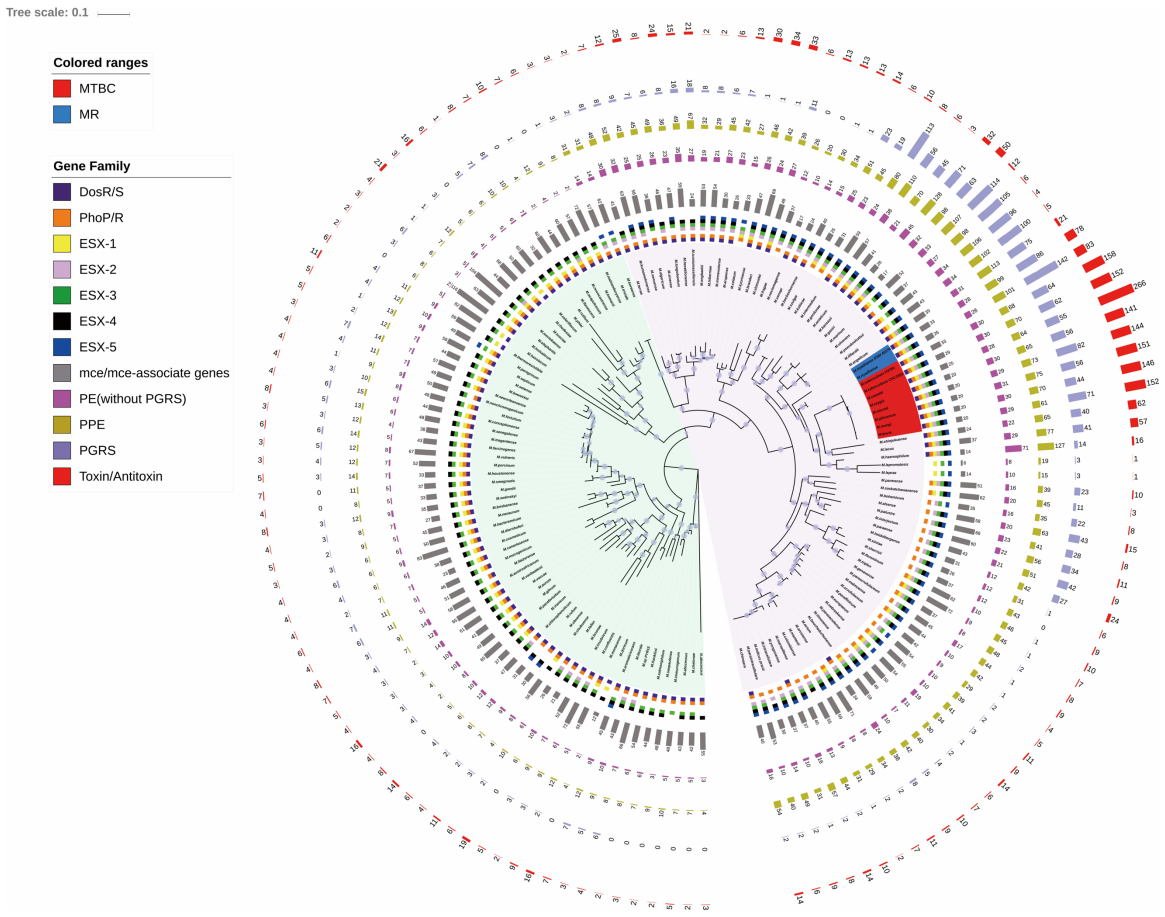


Figure supplementary 7

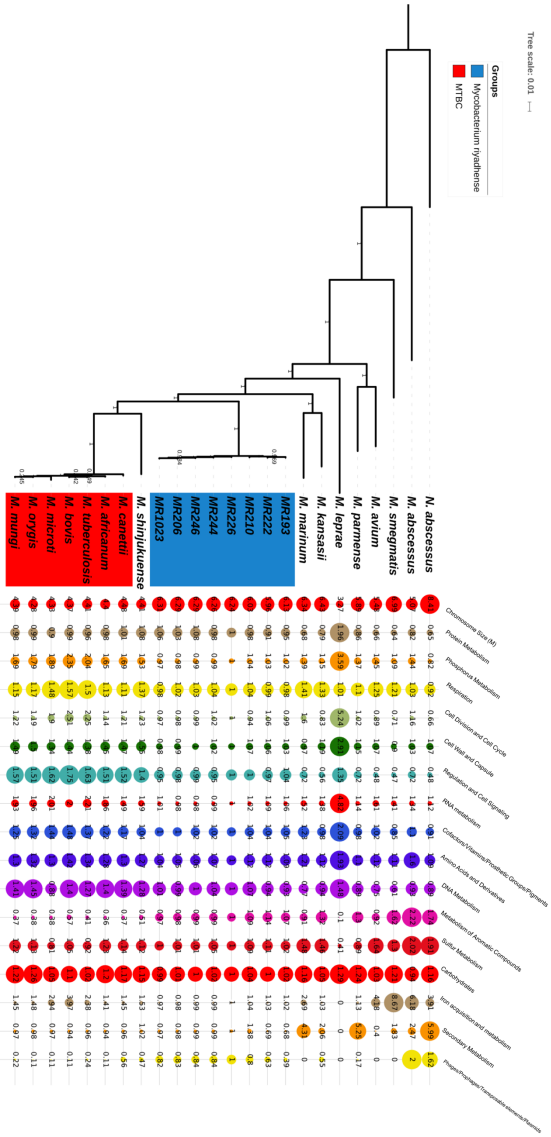


Figure supplementary 8

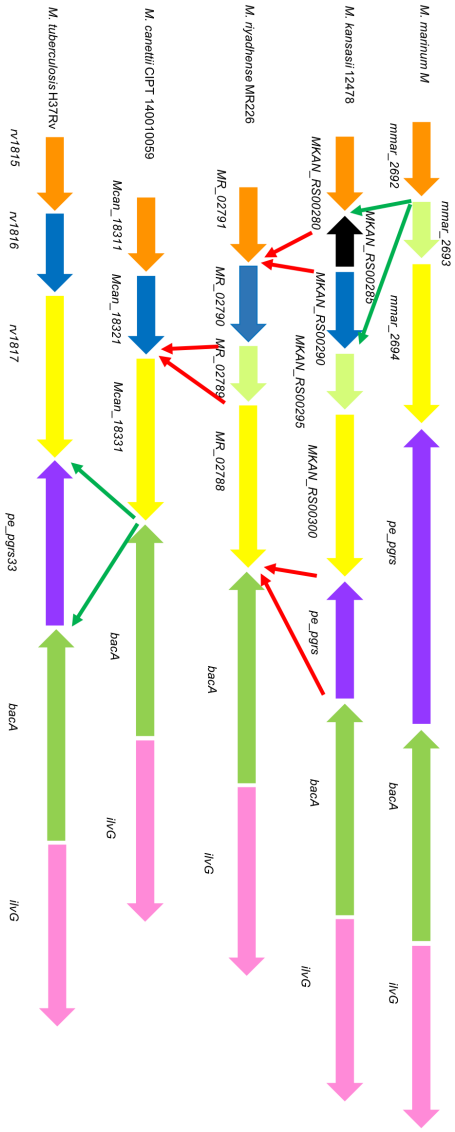


Figure supplementary 9

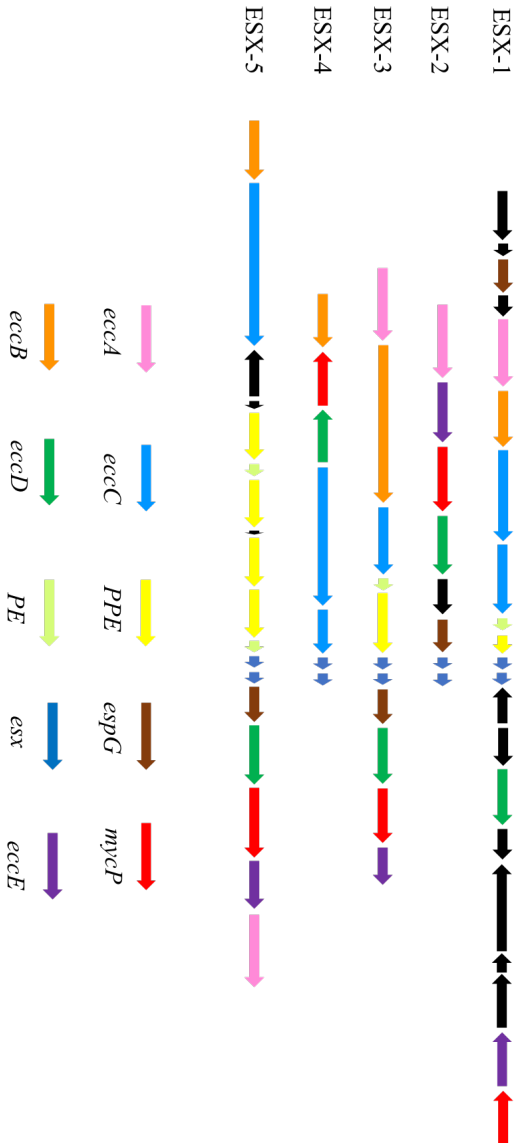


Figure supplementary 10

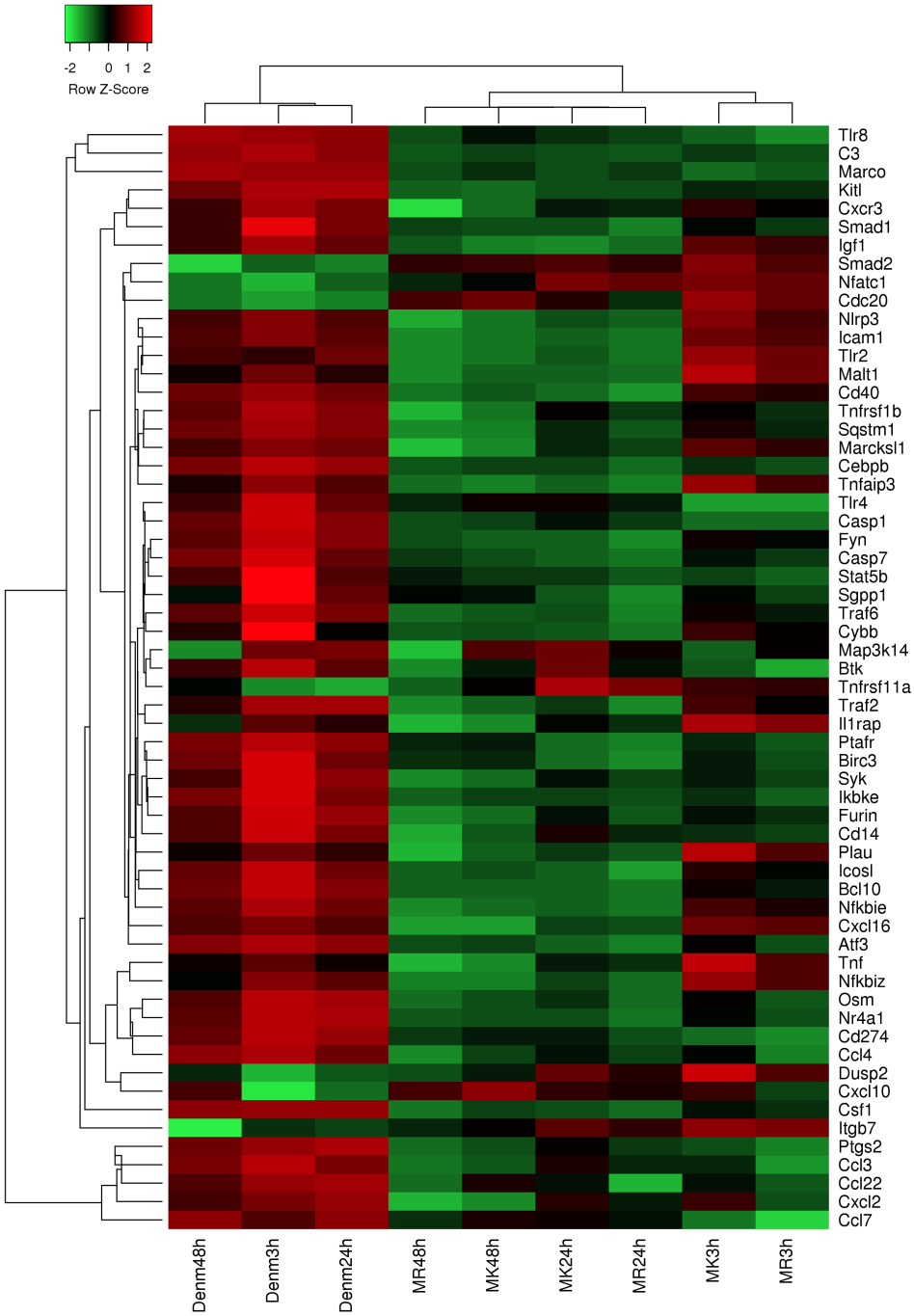
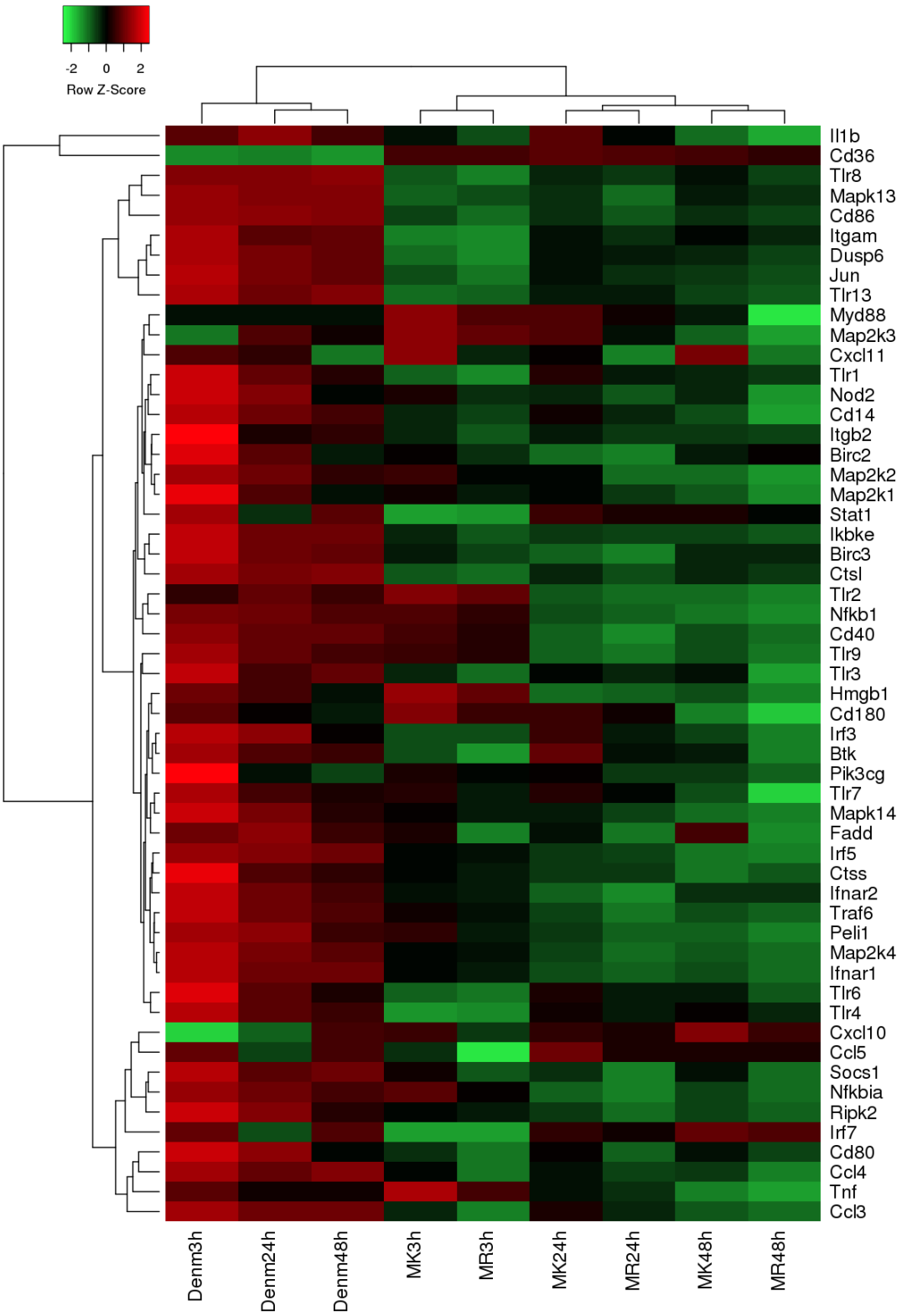


Figure supplementary 11A



1341

1342 Figure supplementary 11B

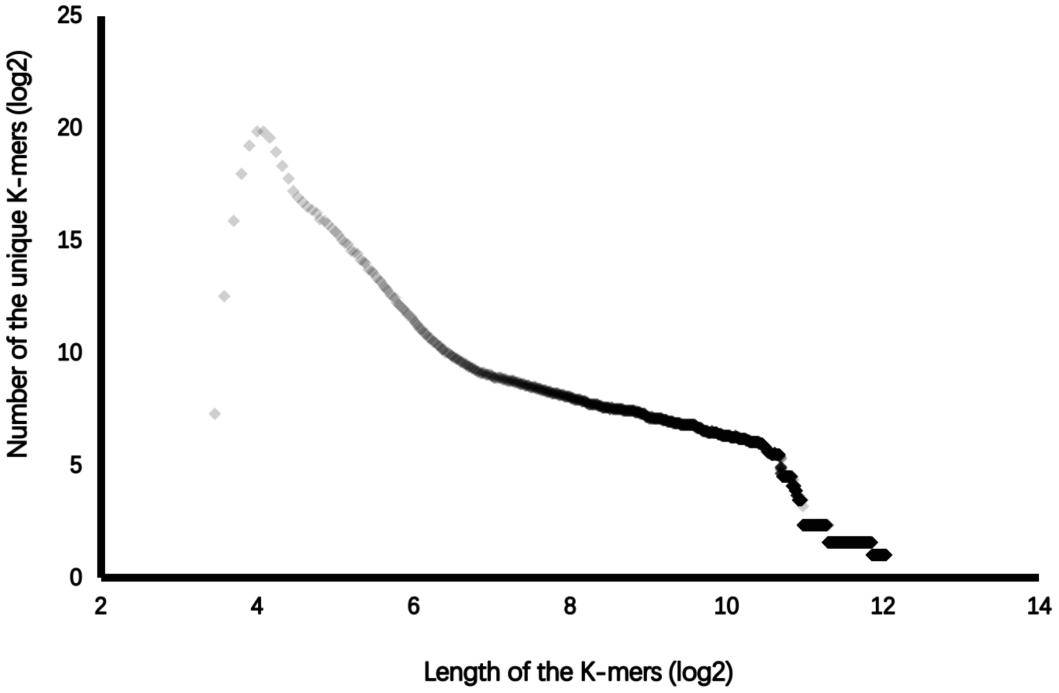


Figure supplementary 12

Supplementary Table 1. Comparison of *M. riyadhense* strains' assemblies

	MR1023	MR193	MR206	MR210	MR222	MR226	MR244	MR246	Type strain assembly GCA_002101845.1
Assembly Size (bp)	6,906,827	6,695,517	6,835,855	6,528,955	6,533,138	6,888,178	6,775,970	6,916,580	6,269,850
Chromosome size (bp)	6,306,178	6,129,600	6,288,531	6,034,715	5,960,707	6,243,587	6,258,360	6,289,824	6,269,850
Contigs	9	7	6	2	7	3	9	4	263
Gaps in chromosome	0	0	0	0	0	0	0	0	262
pMRLP01	P	P	P	P	P	P	P	P	A
pMR01	P	A	A	A	P	P	P	P	A
Transposase number in pMRL01	91	43	81	56	42	67	87	87	73
CDS	6185	5901	6124	5800	5852	6077	6030	6060	5,168
Chromosomal CDS	5630	5392	5634	5420	5353	5545	5637	5546	5,168

1360

Supplementary Table 2. The SNPs analysis of different samples using MR226 as the reference

	MR193	MR206	MR210	MR222	MR244	MR246	MR1023
Non-coding SNPS	2,550	160	2,478	2,727	160	192	165
Synonymous SNPs	16,574	1,107	15,757	16,959	1,250	1,138	1,120
Non-synonymous SNPs	3,446	227	3,161	3,287	266	250	232

1361

Supplementary Table 3. *M. marinum* LOS locus genes orthologs shared with *M. ziziulhense* strains

	MR1023	MR193	MR206	MR210	MR222	MR226	MR244	MR246
MMAR_1008	MR1023_04688	MR193_04455	MR206_04685	MR210_04486	MR222_04444	MR_04617	MR244_04622	MR246_04625
MMAR_2302	MR1023_03305	MR193_03158	MR206_03308	MR210_03179	MR222_03113	MR_03260	MR244_03260	MR246_03263
MMAR_2303	MR1023_03304	MR193_03157	MR206_03307	MR210_03178	MR222_03112	MR_03259	MR244_03259	MR246_03262
MMAR_2304	MR1023_03303	MR193_03156	MR206_03306	MR210_03177	MR222_03111	MR_03258	MR244_03258	MR246_03261
MMAR_2305	MR1023_03302	MR193_03155	MR206_03305	MR210_03176	MR222_03110	MR_03257	MR244_03257	MR246_03260
MMAR_2306	MR1023_03301	MR193_03154	MR206_03304	MR210_03175	MR222_03109	MR_03256	MR244_03256	MR246_03259
MMAR_2307	MR1023_03295	MR193_03148	MR206_03298	MR210_03167	MR222_03103	MR_03250	MR244_03250	MR246_03253
MMAR_2308								
MMAR_2309	MR1023_03299	MR193_03152	MR206_03302	MR210_03173	MR222_03107	MR_03254	MR244_03254	MR246_03257
MMAR_2310	MR1023_03298	MR193_03151	MR206_03301	MR210_03172	MR222_03106	MR_03253	MR244_03253	MR246_03256
MMAR_2311	MR1023_03297	MR193_03150	MR206_03300	MR210_03171	MR222_03105	MR_03252	MR244_03252	MR246_03255
MMAR_2312				MR210_03166				
MMAR_2313 (*ilvx4*)		MR193_03128						
MMAR_2314	MR1023_00361	MR193_03129	MR206_00366	MR210_00381	MR222_00356	MR_00367	MR244_00362	MR246_00362
MMAR_2315								
MMAR_2316	MR1023_03023	MR193_02856	MR206_03023	MR210_01388,MR210_02888	MR222_02834	MR_02982	MR244_02976	MR246_02979
MMAR_2317	MR1023_05374	MR193_05138	MR206_05362	MR210_05150	MR222_05107	MR_05289	MR244_05292	MR246_05296
MMAR_2318		MR193_03130						
MMAR_2319								
MMAR_2320 (*mecE*)								
MMAR_2321		MR193_03145						
MMAR_2322		MR193_03146						
MMAR_2323								
MMAR_2324								
MMAR_2325		MR193_03147						
MMAR_2326								
MMAR_2327	MR1023_03300	MR193_03153	MR206_03303	MR210_03174	MR222_03108	MR_03255	MR244_03255	MR246_03258
MMAR_2328								
MMAR_2329	MR1023_03443	MR193_03296	MR206_03447	MR210_03317	MR222_03253	MR_03400	MR244_03395	MR246_03398
MMAR_2330	MR1023_00875,MR1023_02733	MR193_00837	MR206_00875,MR206_02735	MR210_00864,MR210_02597	MR222_00847	MR_00870,MR_02697	MR244_00866,MR244_	MR246_00867,MR246_02692
MMAR_2331								
MMAR_2332	MR1023_03970	MR193_03792	MR206_03973	MR210_03161	MR222_03768	MR_03910	MR244_03914	MR246_03918
MMAR_2333	MR1023_03293	MR193_03123	MR206_03296	MR210_03160	MR222_03101	MR_03248	MR244_03248	MR246_03251
MMAR_2334				MR210_03159				
MMAR_2335		MR193_04282		MR210_03158				
MMAR_2336 (*galE6*)								
MMAR_2337	MR1023_03292	MR193_03122	MR206_03295	MR210_03156	MR222_03100	MR_03247	MR244_03247	MR246_03250
MMAR_2338				MR210_03153				
MMAR_2339	MR1023_04010	MR193_03832	MR206_04012	MR210_03873	MR222_03809	MR_03949	MR244_03953	MR246_03957
MMAR_2340 (*pkc5*)	MR1023_03273	MR193_03102	MR206_03274	MR210_03130	MR222_03081	MR_03228	MR244_03228	MR246_03231
MMAR_2341 (*dxlX25*)	MR1023_03277	MR193_03106	MR206_03912	MR210_03134	MR222_03085	MR_03232	MR244_03232	MR246_03235
MMAR_2342	MR1023_03276	MR193_03105	MR206_03277	MR210_03133	MR222_03084	MR_03231	MR244_03231	MR246_03234
MMAR_2343	MR1023_03271	MR193_03100	MR206_03272	MR210_03128	MR222_03079	MR_03226	MR244_03226	MR246_03229
MMAR_2344	MR1023_03270	MR193_03099	MR206_03271	MR210_03127	MR222_03078	MR_03225	MR244_03225	MR246_03228
MMAR_2345	MR1023_03268	MR193_03097	MR206_03269	MR210_03125	MR222_03076	MR_03223	MR244_03223	MR246_03226
MMAR_2346	MR1023_03266	MR193_03095	MR206_03267	MR210_03123	MR222_03073	MR_03221	MR244_03221	MR246_03224
MMAR_2347								
MMAR_2348	MR1023_03264	MR193_03093	MR206_03265	MR210_03121	MR222_03071	MR_03219	MR244_03219	MR246_03222
MMAR_2349	MR1023_03278,MR1023_03285	MR193_03107,MR193_03	MR206_03281,MR206_03288	MR210_03135,MR210_03142	MR222_03086,MR222,MR_03233,MR_03240		MR244_03233,MR244_	MR246_03236,MR246_03243
MMAR_2350	MR1023_02109	MR193_02040	MR206_02111	MR210_02028		MR_02086	MR244_02076	MR246_02077
MMAR_2351								
MMAR_2352	MR1023_01970	MR193_01915	MR206_01970	MR210_01904	MR222_01910	MR_01948	MR244_01936	MR246_01937
MMAR_2353	MR1023_03275,MR1023_03279	MR193_03104,MR193_03	MR206_03276,MR206_03282	MR210_03132,MR210_03136	MR222_03083,MR222,MR_03230,MR_03234		MR244_03230,MR244_	MR246_03233,MR246_03237
MMAR_2354	MR1023_02623	MR193_02508	MR206_02626	MR210_02487	MR222_02487	MR_02589	MR244_02580	MR246_02583
MMAR_2403								
MMAR_2404								
MMAR_2405 (*cpkhB*)								
MMAR_2406								
MMAR_2355 (*psps43*)								
MMAR_4419								
MMAR_5170	MR1023_00312	MR193_00322	MR206_00317	MR210_00331	MR222_00307	MR_00318	MR244_00313	MR246_00313

1362