Worldwide spreading of pMAB625 plasmid encoding MMPL family transporter in multidrug-resistant clinical isolates of *Mycobacterium abscessus* by horizontal transfer

Author names:

Kensuke Ohse[1], Atsushi Yoshida[2,3,4], Keisuke Kamada[3,5], Hironobu Kitazawa[6,7], Yusuke Ito[8], Takayo Shoji[9], Kenichiro Watanabe[10], Hiroshi Koganemaru[2], Ken Kikuchi[3] and Masashi Toyoda[1]

Affiliations:

[1] Research Team for aging science (Vascular Medicine), Tokyo Metropolitan Institute for Geriatrics and Gerontology (TMIG), Tokyo, Japan

[2] Department of Infectious Diseases, TMIG, Tokyo, Japan

[3] Department of Infectious Diseases, Tokyo Women's Medical University, Tokyo, Japan

[4] Department of Genetic Diagnosis and Laboratory Medicine, Dokkyo Medical University, Tochigi, Japan

[5] Department of Mycobacterium Reference and Research, The Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association, Tokyo, Japan

[6] Department of Pediatrics, Nagoya University Graduate School of Medicine, Aichi, Japan

[7] Department of Hematology and Oncology, Children's Medical Center, Japanese Red Cross Aichi Medical Center Nagoya First Hospital, Aichi, Japan

[8] Department of Pediatric Infectious Diseases, Hyogo Prefectural Amagasaki General Medical Center, Hyogo, Japan

[9] Department of Infectious Diseases, Shizuoka Children's Hospital, Shizuoka, Japan

[10] Department of Hematology and Oncology, Shizuoka Children's Hospital, Shizuoka, Japan

Corresponding author and email address: Masashi Toyoda, mtoyoda@tmig.or.jp

Keywords: *Mycobacterium abscessus*, plasmid, horizontal gene transfer, MMPL family transporter

Repositories: AP028613, AP028614, AP028615, and AP028616 for Isolate57625; AP028617, AP028618, AP028619, and AP028620 for Isolate57629; AP028621, AP028622, AP028623, and AP028624 for Isolate57630; AP028625, AP028626, AP028627, and AP028628 for Isolate57596; AP028629, AP028630, AP028631, and AP028632 for Isolate57626.

Abstract

In recent years, the number of infections caused by non-tuberculous mycobacteria has increased worldwide, and diseases caused by *Mycobacterium abscessus* are often difficult to treat due to their multidrug resistance. Research into the mechanisms of multidrug resistance in *M. abscessus* has focused on genetic mutations but not on genetic mobile elements such as plasmids. We performed a comparative analysis of clinical isolates of *M. abscessus* collected from the same patient at different time points, indicating the decreased susceptibility to carbapenem antibiotics and growth rate of the isolates compared to the type strain. Whole genome sequencing analysis revealed that isolates had three plasmids (pMAB625-1, pMAB625-2 and pMAB625-3) not found in the type strain. pMAB625-1 and pMAB625-2 encoded genes such as the ESX secretion system, and pMAB625-3 encoded the MMPL family transporter. We investigated the distribution of plasmids previously identified in *M. abscessus,* including pMAB625 plasmids, *in silico* using the whole genome sequencing data from 462 clinical isolates worldwide. The data from 2.2% and 11.5% isolates were mapped to pMAB625-2 and pMAB625-3, respectively. pMAB625-3 was detected in the largest number of isolates of the plasmids analysed. Furthermore, phylogenetic tree analysis and the comparative analysis for mutation detection revealed that these plasmids transferred beyond isolates, regions and

48 subspecies and acquired unique mutations. These results indicate that horizontal transfer of
49 pMAB625 plasmids increases the genomic diversity of the multidrug-resistant clinical isolates
50 of *M. abscessus* worldwide.
51

52 Abbreviations

53 ABS, *M. abscessus* subsp. *abscessus*; AMR, antimicrobial resistance; ANI, average nucleotide
54 identity; CDS, coding sequence; CLSI, Clinical and Laboratory Standards Institute; DEG,
55 differentially expressed gene; FDR, false discovery rate; IPM, imipenem; MAC, *Mycobacterium*
56 *avium* complex; MAS, *M. abscessus* subsp. *massiliense*; MEPM, meropenem; MIC, minimum
57 inhibitory concentration; MMPL, mycobacterial membrane protein large; NTM, non-
58 tuberculous mycobacteria; ORF, open reading frame; PBP, penicillin-binding protein; SNV,
59 single nucleotide variant; SRA, sequence reads archive; SV, structural variant; TA, toxin-
60 antitoxin; WGS, whole genome sequencing
61

62 Impact statement

63 Genomic mutations and the acquisition of genes such as drug efflux transporters are known
64 to cause drug resistance in bacteria. Of these, the horizontal transfer of plasmids and
65 bacteriophages plays an important role in the acquisition of genes. The involvement of
66 plasmids in bacterial virulence is known in gram-negative bacteria. However, it has not been
67 reported in gram-positive bacteria, such as *Mycobacterium abscessus,* except in a few cases. In
68 this study, we performed a comparative genomic analysis of multidrug-resistant clinical
69 isolates of *M. abscessus* using long-read sequencing, indicating plasmids in mediating
70 antimicrobial resistance. Specifically, our study identified conserved plasmids across different
71 clinical isolates beyond regions and subspecies. In addition, our study indicated that plasmid

distribution was strictly differentiated between *M. abscessus* and other non-tuberculous mycobacteria except for very few cases and that there is an incompatibility of plasmid present in non-tuberculous mycobacteria. These findings provide a novel perspective on the acquisition of genomic diversity in *M. abscessus*.

Data Summary

All raw sequence data generated in this study have been deposited in the DDBJ database at https://www.ddbj.nig.ac.jp/ under the BioProject accession number PRJDB16220. The complete genome sequences of the isolates are available in the DDBJ database. They can be accessed using the accession numbers indicated in the repository section for the chromosome and three plasmids in each isolate. The analysed data of RNA-seq have been deposited in the Genomic Expression Archive at https://www.ddbj.nig.ac.jp/gea/ under the accession number E-GEAD-827. Supplementary figures and tables providing details of the used data and results were available on the Figshare data repository platform at https://doi.org/10.6084/m9.figshare.27186183.

Introduction

In recent years, the threat of antimicrobial resistance (AMR) has been increasing worldwide and AMR could be responsible for 1.91 million deaths per year by 2050 [1]. Hence, research into the mechanisms and routes of transmission of AMR is becoming increasingly important to develop effective countermeasures against antimicrobial-resistant bacteria. The global incidence of non-tuberculous mycobacterial infections has been on the rise, presenting a significant challenge due to their increasing prevalence and innate drug resistance, particularly in *Mycobacterium abscessus* infections [2]. *M. abscessus* is not limited to pulmonary infections

96 but also serves as the causative agent for various extrapulmonary infections, including
97 bloodstream, skin, soft tissue, surgical site, and peritoneal dialysis-related infections [3].
98 Furthermore, it demonstrates resistance to a wide range of antimicrobial agents.
99
100 The resistance mechanism has been thoroughly investigated, primarily concerning macrolide
101 antibiotics, which are regarded as essential drugs, and the involvement of the *erm* and *rrl*
102 genes has been elucidated [4–6]. The Clinical and Laboratory Standards Institute (CLSI) also
103 recommends an analysis of the *erm* gene in microbiological laboratories [7]. Among *M.*
104 *abscessus* strains with a non-truncated *erm*(41) gene, the most common variant is T28 sequevar,
105 which has been observed to exhibit inducible resistance to macrolides. However, it has been
106 pointed that *M. abscessus* subsp. *abscessus* (ABS) can develop inducible resistance even without
107 T28 [8], and although whole genome sequencing (WGS) is useful for predicting resistance, they
108 are not sufficient [9, 10].
109
110 We conducted WGS of *M. abscessus* subsp. *massiliense* (MAS), which developed intravascular
111 infections and subsequently became resistant to clarithromycin during treatment. Despite
112 having a non-functional *erm* gene and no *rrl* mutation, we were unable to identify any
113 mutations in the genome that could explain macrolide resistance in this strain with sufficient
114 evidence [11]. This highlights the existence of strains for which the mechanism of AMR remains
115 unexplained by WGS.
116
117 Regarding other antimicrobial agents, apart from resistance genes for amikacin and
118 fluoroquinolones, there is limited information available about the correlation between
119 genomic variations and resistance, and such information is mostly confined to a few

120 antimicrobial agents. Therefore, despite the comprehensive genomic data provided by WGS,
121 identifying the genes responsible for resistance, especially in cases of multi-drug resistance,
122 remains challenging. As a result, the utility of WGS in predicting resistance remains
123 inconclusive. In recent years, it was reported that plasmids mediate the virulence in non-
124 tuberculous mycobacteria (NTM) [12, 13]. Access to long-read sequencers has become easier
125 and complete genome sequences, including plasmids, are accumulating in the public database.
126 Therefore, in case that cannot be explained the cause of bacterial pathogenicity by analysis of
127 chromosomes alone, plasmids should be investigated as the cause of pathogenicity in NTM.
128 Therefore, the plasmid epidemiology in NTM is effective in understanding the causes of NTM
129 pathogenicity.
130
131 In this study, we observed elevated minimum inhibitory concentrations (MICs) for carbapenem
132 antibiotics in a case of disseminated infection caused by ABS T28 sequevar during
133 antimicrobial therapy. Several causes of the decreased susceptibility to carbapenem antibiotics
134 have been reported, including the acquisition of the carbapenemase [14], efflux pump
135 overexpression [15], and mutations in PBPs [16]. However, the mechanism of the increased
136 resistance to carbapenem antibiotics in *M. abscessus* is unclear. We then performed *de novo*
137 genome assembly of the strains isolated sequentially from this patient, including chromosome
138 as well as plasmid, by WGS using short- and long-read sequencers. We present the results,
139 which suggest the involvement of plasmids in the mechanism of AMR. Taking these results, we
140 reanalysed the read data obtained by WGS from the clinical isolates of *M. abscessus* worldwide
141 using a short-read sequencer deposited in the public database. We found that the data from
142 2.2% and 11.5% of phylogenetically different isolates were mapped to the two plasmid

143 sequences, respectively, suggesting that these plasmids have been horizontally transferred
144 among clinical isolates of *M. abscessus* worldwide.
145

146 Methods

147 Clinical isolates and type strain

148 Ethical approval for this study was obtained from the Ethics Committee of Dokkyo Medical
149 University (approval number 28045). Regarding Isolate57625, Isolate57629, Isolate57630,
150 Isolate57596, and Isolate57626, specimens were obtained from a patient with osteomyelitis at
151 the five time points in six months indicated in Table 1. Each strain was isolated from the
152 specimen at Shizuoka Children's Hospital for diagnosis. The other clinical isolates used for PCR
153 or WGS were stored at Tokyo Metropolitan Institute for Geriatrics and Gerontology or Tokyo
154 Women's Medical University or Dokkyo Medical University [17]. The type strain of ABS, ATCC
155 19977, was purchased from ATCC (Manassas, VA, US). The type strain of MAS, JCM 15300, and
156 the type strain of BOL, JCM 15297, were provided by Japan Collection of Microorganisms,
157 RIKEN BRC which is participating in the National BioResource Project of the MEXT, Japan.
158

159 Growth curve

160 The type strain and the five clinical isolates were inoculated onto trypticase soy agar with 5%
161 sheep blood (Beckton Dickinson, Franklin Lakes, NJ, US) and incubated at 30 °C for five days.
162 Next, some colonies were inoculated into BBL™ Mueller Hinton II Broth (Cation-Adjusted)
163 (Beckton Dickinson) and incubated overnight at 30°C. These culture media were measured for
164 absorbance (OD_{600}), from which they were prepared to 0.1 by medium. Seventy microliter of
165 the adjusted culture solution was inoculated into 7 mL of BBL™ Mueller Hinton II Broth
166 (Cation-Adjusted) and incubated at 30 °C for eight days with agitation (200 rpm). OD_{600} was

167 measured on days one to eight after inoculation. Statistical analysis was performed using R
168 ver. 4.3.1.
169

170 Antimicrobial susceptibility testing

171 The type strain and the clinical isolates were subcultured on 2% OGAWA medium (Kyokuto
172 Pharmaceutical Industrial, Tokyo, Japan) at 30 °C for five days. Antimicrobial susceptibility
173 testing was performed using BrothMIC RGM (Kyokuto Pharmaceutical Industrial) following
174 CLSI M24 3rd ed [7]. MICs were determined as the minimum concentration at which bacteria
175 had not grown on day four (all antibiotics) and day 14 (clarithromycin). The test was repeated
176 three times and all MICs were confirmed to be the same between the experiments, except for
177 imipenem (IPM) for the isolates. For IPM, the MICs of the type strain were confirmed to be the
178 same between all three experiments, while for the MICs of Isolate57625 and Isolate57626, they
179 were confirmed to be the same between two out of three experiments, as indicated in Table 1.
180

181 Mycobacterial DNA isolation and whole-genome sequencing

182 The type strain and the clinical isolates of *M. abscessus* were stored in frozen stock containing
183 skim milk and 5% (v/v) glycerol. Each strain was inoculated from frozen stocks into
184 Middlebrook 7H9 Medium (Sigma-Aldrich, St. Louis, MO, US) containing 10% (v/v) oleic acid-
185 albumin-dextrose-catalase and 0.2% (v/v) glycerol and cultured at 30 °C for six days with
186 agitation (200 rpm). Bacterial genomic DNA for short-read sequencing was extracted as
187 previously reported [18]. The quality of the extracted DNA was confirmed by agarose gel
188 electrophoresis, and the concentration was measured using Qubit™ BR DNA Quantitation Kit
189 (Thermo Fisher Scientific, Waltham, MA, US). Genomic DNA was fragmented, and the
190 sequencing library was prepared using the Ion Xpress™ Plus Fragment Library Kit and the Ion

191 Xpress™ Barcode Adapters 1-96 kit (Thermo Fisher Scientific) according to the manufacturer's
192 instructions. The quantity of sequencing library was assessed using the Ion Library TaqMan™
193 Quantitation Kit (Thermo Fisher Scientific). Emulsion PCR for sequence template synthesis was
194 performed with the Ion PI™ Hi-Q™ OT2 200 Kit v2 using the Ion One Touch™ II System and
195 the Ion One Touch™ II ES (Thermo Fisher Scientific). Sequencing was performed using the Ion
196 PI™ Chip kit v3 and the Ion Proton™ System (Thermo Fisher Scientific). Data were collected
197 using the Torrent Suite v5.0.5 software (Thermo Fisher Scientific).
198
199 High-molecular-weight genomic DNA for long-read sequencing was extracted as follows. The
200 bacterial pellet was collected by centrifugation and homogenised using EZ-beads (Promega,
201 Madison, WI, US) and Shakeman 6 (Bio Medical Science, Tokyo, Japan). The homogenised
202 pellet was lysed with Lysozyme (Sigma-Aldrich, L4919) at 37 °C for one hour. Next, DNA was
203 extracted using Wizard® HMW DNA Extraction Kit (Promega) following the manufacturer's
204 instructions. Fragmented DNA in the extracted DNA was removed using the Short Read
205 Eliminator Kit (PacBio, Menlo Park, CA, US). The sequencing library for long-read sequencing
206 was prepared using the Ligation Sequencing Kit (SQK-LSK109, Oxford Nanopore Technologies,
207 Oxford, UK) and the Native Barcoding Expansion 1-12 (EXP-NBD104, Oxford Nanopore
208 Technologies). In brief, the barcode sequences for multiplex sequencing were attached to the
209 ends of the HMW genomic DNA. In addition, the sequencing adapters were ligated to the end
210 of HMW genomic DNA. The prepared DNA was loaded into a flow cell (FLO-MIN106, Oxford
211 Nanopore Technologies) attached to MinION Mk1B (Oxford Nanopore Technologies).
212 Sequencing data were acquired using MinKNOW ver.21.06.0 (Oxford Nanopore Technologies),
213 and base calling was performed using Guppy ver. 5.0.7 (Oxford Nanopore Technologies) with
214 high accuracy mode using the National Institute of Genomics supercomputer system (NIG,

215 Shizuoka, Japan). Data quality was confirmed using NanoPlot ver. 1.38.0 [19] and adapter
216 sequences were trimmed using Porechop ver. 0.2.4 [20]. Raw reads shorter than 1000 bp
217 (Isolate57625, Isolate57629, Isolate57630, Isolate57596) or 5000 bp (Isolate57626) were
218 removed using Filtlong ver. 0.2.1 (https://github.com/rrwick/Filtlong). Filtered reads were
219 corrected using Fmlrc ver. 1.0.0 [21] with the data of short-read sequencing acquired by the
220 Ion Proton™ Sequencer. *De novo* assembly was performed using Flye ver. 2.9-b1768 [22] with
221 corrected long-reads. Assembled genome sequences were polished using Pilon ver. 1.24.0 [23],
222 evaluated using QUAST ver. 5.0.2 [24], and visualised using Bandage ver. 0.8.1 [25]. Genome
223 annotation was performed using DFAST v.1.5.0 [26, 27].
224

225 Plasmid sequence analysis

226 Known plasmid sequences were downloaded from PLSDB [28] by the taxonomic search for
227 "*abscessus*". Average nucleotide identity (ANI) was calculated using FastANI ver.1.34 [29] and
228 structural variants (SVs) were analysed using Mauve ver. 2.4.0 [30].
229
230 Homology search was performed by blast+ [31] using the amino acid sequence as the query
231 or by Abricate software (http://github.com/tseemann/abricate) using the nucleotide
232 sequences as the query. NCBI AMRFinderPlus [32], CARD [33], Resfinder [34], ARG-ANNOT [35],
233 VFDB [36], PlasmidFinder [37], EcOH [38], and MEGARes 2.00 [39] were used as the database
234 for the search of AMR genes. Mycobrowser [40] was used as the database for the homology
235 search of the mycobacterium genes.
236

237 SNVs and SVs analysis

238 Short-read data acquired by the Ion Proton™ Sequencer were mapped using BWA-MEM ver.
239 0.7.17-r1188 [41] against the reference sequence (CU458896.1, CU458745.1). Output BAM files
240 were sorted using samtools ver.1.12 [42]. Valiant calling was performed using Freebayes ver.
241 1.3.5 [43], and variants were annotated using SnpEff ver. 5.2 [44] with annotated genome
242 sequence. The positions of single nucleotide variants (SNVs) were visualised using IGV (Broad
243 Institute, Cambridge, MA, US). Assembled genome sequences were aligned using Mauve ver.
244 2.4.0 [30] for SVs analysis. Mapping the short- and long-read to the assembled genome
245 sequence using minimap2 ver. 2.21-r1071 [45] was performed for the validation of assembly.
246
247 PCR for the validation of SVs was performed using KOD One® PCR Master Mix (TOYOBO,
248 Osaka, Japan) and the primers indicated in Table S1. Amplified DNA was confirmed using
249 agarose gel electrophoresis following staining with SYBR™ Safe Gel Stain (Thermo Fisher
250 Scientific).
251

252 RNA-seq analysis

253 Total RNA was extracted as follows: The bacterial pellet was collected by centrifugation and
254 suspended in UltraPure™ DNase/RNase-free Distilled Water (Thermo Fisher Scientific) and
255 homogenised using EZ-beads (Promega) and Shakeman 6 (Bio Medical Science). SUPERase-
256 In™ RNase Inhibitor (Thermo Fisher Scientific) was added to the homogenised pellet. Next, the
257 homogenised pellet was lysed by Lysozyme (Sigma-Aldrich) at 37 °C for an hour. Total RNA
258 was purified by mirVana™ miRNA Isolation Kit (Thermo Fisher Scientific). The size and quality
259 of total RNA were assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa
260 Clara, CA, US) and the RNA 6000 Nano Kit. rRNA was depleted using the NEBNext® rRNA
261 Depletion Kit (Bacteria) (NEB, Ipswich, MA, US) from 500 ng of total RNA. The whole

262 transcriptome library was prepared using the Ion Total RNA-Seq Kit v2 (Thermo Fisher
263 Scientific) as follows: First, the rRNA-depleted RNA was fragmented by RNase III at 37 ˚C for
264 10 minutes and purified using the Magnetic Beads Cleanup Module. Next, adapter sequences
265 were ligated and reverse transcription was performed. The whole transcriptome library was
266 amplified and barcoded using the IonXpress™ RNA-Seq Barcode 1-16 Kit (Thermo Fisher
267 Scientific). The quantity of the amplified whole transcriptome library was assessed using the
268 Ion Library TaqMan™ Quantitation Kit (Thermo Fisher Scientific). Emulsion PCR for sequence
269 template synthesis was performed using the Ion PI™ Hi-Q™ OT2 200 Kit v2 with the Ion One
270 Touch™ II System and the Ion One Touch™ II ES (Thermo Fisher Scientific). Sequencing was
271 performed using the Ion Proton™ System with Ion PI™ Chip kit v3 (Thermo Fisher Scientific).
272 Data was collected using the Torrent Suite ver. 5.0.5 software (Thermo Fisher Scientific).
273 Sequence reads were mapped using Bowtie ver. 1.0.0 [46] against the assembled genome
274 sequence of Isolate57625. Mapped reads were counted using HTseq ver. 1.99.2 [47]. Merged
275 transcript count data were normalized using TCC ver. 1.26.0 [48] and differentially expressed
276 genes (DEG) were detected under false discovery rate (FDR) < 0.01.
277

278 Phylogenetic tree analysis

279 The genome sequences of the various clinical isolates shown in Fig. 2c for phylogenetic tree
280 analysis were obtained from the NCBI database and are listed in Table S2. Core-genome SNVs
281 were detected using Snippy ver. 4.6.0 (https://github.com/tseemann/snippy) with each
282 assembled genome sequence and the type strain (ATCC 19977, CU458896.1) as the reference.
283 The concatenated core-genome SNVs were also aligned using the snippy-core script. The
284 phylogenetic tree was constructed using IQ-TREE ver.2 [49] using the maximum likelihood
285 criterion and rendered using iTOL ver.6.7 [50]. Phylogenetic tree analysis of pMAB625

286 sequences were performed by the same method. For the phylogenetic tree analysis regarding
287 the components of the ESX secretion system, the amino acid sequences of ESX loci (*eccB*, *eccC*,
288 *eccD*, *mycP*, and *eccE*) were concatenated and used as input for NGphylogeny.fr [51] using the
289 PhyML+SMS inference method.
290

291 PCR for plasmid detection and copy number analysis

292 PCR was performed using KOD One® PCR Master Mix (TOYOBO) and the primers indicated in
293 Table S1. Amplified DNA was confirmed using agarose gel electrophoresis. The plasmid copy
294 number was analysed by qPCR using PowerUp™ SYBR™ Green Master Mix (Thermo Fisher
295 Scientific). qPCR was performed using the QuantStudio™ 5 Real-Time PCR System (Thermo
296 Fisher Scientific). Five nanogram of genomic DNA was used for a 20 µL reaction volume of
297 qPCR. The plasmid copy number is shown as a relative value normalised using the copy number
298 of *HrpA*, which is present in one copy on the chromosome.
299

300 Mapping the read data of various clinical isolates to the 301 genome sequence of Isolate57625

302 Raw read data of *M. abscessus* clinical isolate in Japan, Taiwan [52], China [53], the US, the
303 United Kingdom (UK) and Spain were obtained from the NCBI sequence reads archive (SRA).
304 Raw read data were trimmed using fastp ver. 0.23.4 [54] and reduced using seqkit ver.2.4.0 [55]
305 to 200 Mb for the Ion Proton™ single-end data and 100 Mb for the Illumina pair-end data to
306 equalise the depth of coverage among isolates. Reads were mapped using BWA-MEM ver.
307 0.7.17-r1188 against the chromosome and plasmid sequences of Isolate57625 and other
308 plasmids which were hit using PLSDB by the taxonomic searching of "*abscessus*" with removing
309 the redundant plasmid sequences (ANI (%)>97, without large SV (>8 kbp)). The depth of

310 coverage was calculated using Qualimap ver. 2.2.2-dev [56]. The covered genome region was
311 calculated using bedtools ver. 2.30.0 [57]. We considered the isolate whose depth of coverage
312 was higher than 10x and the covered genome region was higher than 75% as the plasmid-
313 harbouring isolate. The 152 sequences for the analysis of the distribution of plasmids identified
314 in other NTM and previously reported [58] were listed in Table S2. For the phylogenetic tree
315 analysis, the aligned assumed chromosome sequences using Snippy-core script were used as
316 input for Gubbins ver.3.3.1 [59]. The phylogenetic tree was rendered using ggtree ver 3.10.1
317 [60]. Statistical analysis of the number of clinical isolates retaining plasmid was performed
318 using Fisher's exact test with Holm's method for correction of multiple comparisons using
319 RVAideMemoire ver. 0.9-83-7 (https://github.com/cran/RVAideMemoire).
320

321 Results

322 The comparison of growth rate and antimicrobial 323 susceptibility of the isolates with the type strain

324 We sequentially isolated five MAB strains from one patient who had osteomyelitis over six
325 months (Table 1). We first compared the growth rate and antimicrobial susceptibility of these
326 clinical isolates with the type strain (ATCC 19977). The growth rates of Isolate57625 and
327 Isolate57626 were significantly slower than the type strain (Fig. S1). Antimicrobial susceptibility
328 testing showed that the MIC of Isolate57626 against IPM, which binds the penicillin-binding
329 proteins (PBPs) and inhibits the bacterial cell wall synthesis, was higher than that of the type
330 strain (Table 1). The MICs of Isolate57625 and Isolate57626 against meropenem (MEPM), which
331 has the same mechanism of action as IPM and improved stability, were higher than that of the
332 type strain (Table 1). However, there were no differences in MIC between ATCC 19977 and two

333 isolates against the other antibiotics. These results indicated the decreased susceptibility of
334 the isolates to carbapenem antibiotics compared to the type strain.
335

336 WGS of the clinical isolates

337 To explore the factors determining the difference in growth rate and antimicrobial
338 susceptibility between the type strain and the clinical isolates, we performed WGS using short-
339 read sequencer and Nanopore long-read sequencer. We analysed the *de novo* assembled
340 chromosome sequences and found that two large insertions (13 kbp and 8 kbp) and a deletion
341 (16 kbp) existed in the sequence of Isolate57625 (Fig. S2a, b). The coding sequences of phage
342 integrase and transposase were present near the three large SVs (Fig. S3), indicating that these
343 regions were prophage. We also analysed the SNVs in the chromosome (Table S3), however,
344 we could not identify the cause of the phenotypic differences. Phylogenetic analysis showed
345 that the isolates were genetically close to the type strain (Fig. S2c). Furthermore, to analyse the
346 difference in the chromosome sequence between five clinical isolates, we analysed SVs and
347 SNVs between the isolates (Fig. S4). This analysis suggested one mutation located at the
348 putative *PbpA*, which was associated with the synthesis of the bacterial cell wall and the target
349 of carbapenem antibiotics, may be responsible for the reduced susceptibility of Isolate57626
350 to IPM compared to Isolate57625.
351

352 Sequence analysis of pMAB625 plasmids

353 We could not identify the cause of phenotypic differences of the isolates compared with the
354 type strain through the analysis of chromosomes. However, WGS analysis revealed that the
355 clinical isolates have three plasmids not found in the type strain (named pMAB625-1,
356 pMAB625-2 and pMAB625-3, respectively) (Table 2). Thus, we next focused on the pMAB625

357 plasmids and compared the pMAB625 plasmid sequences with other plasmids registered to
358 PLSDB [28] previously (Table 3). A taxonomic keyword search for "*abscessus*" returned 20
359 plasmids. To investigate the homology between these plasmids, we calculated ANI between
360 them and pMAB625. No plasmid had a high ANI for pMAB625-1, however, the ANI between
361 pMAB625-2 and pGD21-2 was 97.4% with some SVs. Furthermore, the ANI for pMAB625-3
362 against pGD69A-1 and pGD42-1 was 100%. These results indicated that pMAB625-2 and
363 pMAB625-3 were retained in the other clinical isolates of *M. abscessus*. Next, we performed
364 qPCR using primers specific to each plasmid to verify the copy numbers of pMAB625 plasmids
365 in the clinical isolates. This analysis showed that the copy number of pMAB625-1 and
366 pMAB625-2 was less than one per bacterium, and that of pMAB625-3 was more than five per
367 bacterium in the isolates (Fig. S5). Therefore, the clinical isolates were considered to be a
368 heterogeneous population in terms of plasmid possession.
369
370 We then investigated the function of the genes encoded on pMAB625 plasmids using
371 homology search. Plasmids encode various factors that influence host characteristics.
372 Therefore, we listed the lost and acquired genes in Isolate57625 compared to the type strain
373 to investigate the factors affecting antimicrobial susceptibility (Table S4). Of note, 480
374 candidate open reading frames (ORFs) were acquired and 37 candidate ORFs were lost in
375 Isolate57625 compared to the type strain. We then searched the AMR genes from several
376 databases using Blast+ based on the amino acid sequences of the acquired candidate ORFs.
377 This analysis showed that the coding sequence (CDS) named Isolate57625_54550 had a
378 similarity to *qacA*, a subunit of the qac multidrug efflux pump [61], (identity, 31 %; *e*-value,
379 1.69e-10; coverage, 16.1%) and that the CDS named Isolate57625_54800 had a similarity to

380 *MexL*, a specific repressor of *MexJK* multidrug efflux system [62] (identity, 33%; *e*-value, 5.54e-
381 4; coverage, 34.4%). However, the identity and coverage were not sufficiently high.
382
383 To gain further insight into the effect of gene acquisition, we searched the homology of the
384 acquired candidate ORFs using the mycobacterial gene database of Mycobrowser [40]. This
385 analysis detected 37 mycobacterial proteins with amino acid homology to candidate ORFs of
386 the isolates. (identity > 30 %, *e*-value < 1e-50) (Table S5). Notably, genes associated with the
387 ESX secretion system were highly enriched (10/37, 27%). ESX secretion systems are encoded
388 in various bacteria on their chromosomes and plasmids. It has been reported that the ESX
389 secretion system (also known as the type VII secretion system) is involved in bacterial virulence
390 through several mechanisms, including evasion of the host immune system [63, 64], and
391 biofilm formation [65]. We further searched whether other ORFs of the components of the ESX
392 secretion system were encoded near the loci of the detected ORFs of the components of the
393 ESX secretion system using blast+. We detected several components of the ESX secretion
394 system in pMAB625-1 and pMAB625-2 (Fig. 1a). Core genes that construct membrane pores
395 (*eccB, eccC, eccD, mycP,* and *eccE*) were present in these gene clusters. The type strain and
396 clinical isolates have ESX-3 and ESX-4 loci on their chromosomes; therefore, the isolates
397 acquired two additional ESX loci by the plasmid acquisition. To predict the function of these
398 genes, we performed phylogenetic tree analysis using known amino acid sequences of the ESX
399 secretion system components in chromosomes and plasmids. It was previously reported that
400 plasmid-encoded ESX loci can be classified into 4 major clusters [66]. Importantly, our analysis
401 revealed that plasmid-encoded ESX loci were classified into 5 major clusters (ESX-P cluster 1-
402 5) (Fig. 1b). Notably, the newly classified ESX-P cluster 5 contained the ESX loci of pMAB625-
403 1 and pMAB625-2. As shown in the results of plasmid homology search (Table 3), the ESX loci

404 of pGD21-2 and pMAB625-2 were genetically close. The ESX loci of pMAB625-1 were
405 genetically close to that of pMFLV01, which was harboured by *Mycobacterium gilvum* PYR GCK.
406 It was previously reported that pMFLV01 was enriched in biofilm formed in household water
407 purifiers by metagenomic analysis [67]. Biofilm formation is one of the drug resistance
408 mechanisms in mycobacterium and acts as a barrier against antibiotics. We also detected the
409 toxin-antitoxin (TA) system components, YefM/YoeB and VapB5/VapC5 in pMAB625-1. The TA
410 system is present in various bacteria and archaea and encoded on their chromosomes and
411 plasmids. It contributes to plasmid persistence, inhibiting bacteriophage propagation, and
412 increases antimicrobial tolerance by inducing dormancy [68]. These results indicated that the
413 clinical isolates acquired factors that could contribute to AMR through plasmid transfer.
414

415 Verification of the gene expression from pMAB625 plasmids

416 We found that the various genes associated with bacterial virulence were encoded on the
417 pMAB625 plasmid. However, it was unclear whether genes encoded on these plasmids were
418 expressed in the isolates. To address this question, we analysed the global gene expression in
419 the type strain and the isolates using RNA-seq analysis (Fig. 1c and Table S6). In the clinical
420 isolates, 393 genes were high expressed and 29 genes were low expressed compared to the
421 type strain. Of these differentially expressed genes, 87.2% of genes (368 high expressed genes)
422 were acquired genes by SVs or plasmid and 12.8% of genes (25 high expressed genes and 29
423 low expressed genes) were present on the chromosomes of both the isolates and the type
424 strain (Fig 1c). Notably, the expression levels of the genes associated with the ESX secretion
425 system, which were encoded on the pMAB625-1 and pMAB625-2, were 26- to 19016-fold
426 higher in the isolates than in the type strain (Fig. 1c). In addition, the genes associated with the
427 TA system and the mycobacterial membrane protein large (MMPL) family transporter, which

428 were also encoded on the pMAB625-1 and pMAB625-3, respectively, were highly expressed in
429 the isolates compared to the type strain. These results indicated that the numerous genes
430 associated with bacterial virulence encoded on the plasmids were more highly expressed in
431 the clinical isolates than in the type strain, despite the low number of copies per bacterium
432 regarding the pMAB625-1 and pMAB625-2.
433

434 Investigation of plasmid possession in the various clinical 435 isolates worldwide

436 So far, we have found that the isolates carried three plasmids that the type strain did not, that
437 these plasmids encoded the genes that could contribute to the AMR, and that these genes
438 were also highly expressed in the isolates. However, we had analysed the strains sequentially
439 isolated from only one patient and it was uncertain whether the plasmid acquisition occurred
440 extensively in clinical isolates of mycobacterium. We therefore investigated whether other
441 clinical isolates of mycobacterium carry the pMAB625 plasmids. We first performed PCR using
442 the genomic DNA of the clinical isolates collected in Japan [17] with primers specific to each
443 pMAB625 plasmid. We detected pMAB625-3 in two isolates other than Isolate57625, although
444 we did not detect pMAB625-1 and pMAB625-2 in any isolates other than Isolate57625 (Fig.
445 2a). Both the isolates in which pMAB625-3 was detected were *M. abscessus*. We then verified
446 the plasmid possession by mapping the read data from the short-read sequencer to the
447 pMAB625 plasmid sequences. We used the read data of *M. abscessus* clinical isolates in Japan
448 [52] obtained from the NCBI database in addition to our original data. Out of 125 clinical
449 isolates, pMAB625-2 was detected in four isolates (3.2%) and pMAB625-3 was detected in 16
450 isolates (12.8%) (Fig. 2b). These isolates were collected from the four prefectures in Japan
451 (Tokyo, Shizuoka, Aomori, and Okinawa) (Fig. 2c). We also mapped the read data of the other

452 mycobacterium species (three of *Mycobacterium iranicum*, 20 of *Mycobacterium fortuitum*, and
453 25 of *Mycobacterium chelonae*), however, no pMAB625 plasmid was detected in these isolates
454 (data not shown).
455
456 Moreover, to investigate whether clinical isolates in other areas of East Asia possess pMAB625
457 plasmids, we mapped the available read data about the clinical isolates from Taiwan [52] and
458 China [53]. pMAB625-2 was detected in the isolates from China (2/69, 2.9%) and pMAB625-3
459 was detected in the isolates from Taiwan (8/98, 8.2%), and China (5/69, 7.2%) (Fig. S6). Two
460 plasmids were detected in ABS and MAS, but not in *M. abscessus* subsp. *bolletii* (BOL). In
461 addition, we analysed the clinical isolates in Spain, the UK, and the US to determine whether
462 these isolates carried the pMAB625 plasmids (Table 4). pMAB625-2 plasmid was detected in
463 isolates from the UK (2/30, 6.7%), and the US (2/110, 1.8%). pMAB625-3 plasmid was detected
464 in the isolates from Spain (9/30, 30%), the UK (1/30, 3.3%), and the US (14/110, 12.7%). Overall,
465 pMAB625-1 was detected in one isolate (1/462, 0.2%), pMAB625-2 was detected in 10 isolates
466 (10/462, 2.2%), and pMAB625-3 was detected in 53 isolates (53/462, 11.5%).
467
468 As previous analyses have shown that the pMAB625 plasmid is retained in many clinical
469 isolates, we investigated to what extent other plasmids are retained in the clinical isolates. First,
470 we searched the plasmid sequence in PLSDB as the taxonomic name for "*abscessus*" and got
471 20 plasmids sequences (Table 3). We removed 3 plasmids (pGD42-1, pGD69A-1 and
472 FDAARGOS_1640_plasmid) because they were identical to other plasmids with no SV. We
473 selected 20 plasmids shown in Table 3 as "TRUE" for "USE for Reference" together with three
474 pMAB625 plasmids. Then, we mapped the read data of the clinical isolates to these plasmid
475 sequences. Except for the pMAB625 plasmid, three plasmids (pGD69A-2, pGD22-2, and

476 pGD42-2) were detected in more than 10 isolates (Table 4). Seven plasmids (pGD21-1, pGD21-
477 2, pGD22-1, pGD25-3, pJCM30620_2, BRA100 and 50594_plasmid1) were not detected in our
478 dataset. Statistical analysis showed regional differences in the frequency of appearance of
479 these plasmids (Table 4). pGD69A-2 was detected more frequently in the US (9/110, 8.2%) than
480 in Japan (0/125, p=0.014). pGD22-2 was detected more frequently in Spain (3/40, 13.3%) than
481 in Japan (0/125, p=0.018) and Taiwan (0/98, p=0.036). pGD42-2 was detected more frequently
482 in Spain (4/30, 13.3%) than in Japan (0/125, p=0.017). pGD42-2 was also detected more
483 frequently in the US (12/110, 10.9%) than in Japan (0/125, p=0.001), China (0/69, p=0.045), and
484 Taiwan (1/98, 1.0%, p=0.041). pGD25-1 was detected more frequently in the UK (4/30, 13.3%)
485 than in Japan (0/125, p=0.018) and in Taiwan (0/98, p=0.036). Moreover, we performed a
486 phylogenetic tree analysis of the chromosomes to determine whether these plasmids are
487 horizontally transferred or whether the same strain harbouring the plasmids are spreading (Fig.
488 3). The phylogenetic tree analysis showed that pGD22-2 and pGD42-2 were only detected in
489 the isolates belonging to Clade756. In contrast, pMAB625-3 was detected mainly in the isolates
490 belonging to the four clades (Clade595, 674, 791 and 813). pGD25-1 and pGD69A-2 were
491 detected mainly in the isolates belonging to Clade678 and Clade813. pMAB625-2 and pGD19
492 were also detected in the isolates belonging to several clades. These results indicated that
493 some plasmids, such as pMAB625-3, were spread across clades and others were restricted to
494 the specific clade.
495
496 To analyse the genetic relations between plasmids, we performed a phylogenetic tree analysis
497 of 53 sequences of pMAB625-3 (Fig. 4a). Forty-two sequences belonged to the same clade of
498 the sequence of Isolate57625. However, 11 sequences belonged to the different clades. In
499 particular, four pMAB625-3 sequences of the isolates belonging to Clade791 and three

500 sequences of the isolates belonging to Clade674 of the phylogenetic tree of host bacterial
501 chromosome were classified to the clades genetically distant from the pMAB625-3 sequence
502 of Isolate57625. In addition, the verification of the mapping results showed that pMAB625-3
503 sequences in the six isolates (DRR317373, Japan; SRR18969469, China; SRR7800518, US;
504 SRR7800703, US; SRR7800409, US; SRR7800511, US) were heterogeneous because the
505 unmutated reads and mutated ones were mixed (Fig. 4b). Thus, we ran freebays with parameter
506 "--ploidy 2" to detect SNVs in these mixed reads. Interestingly, the variant hotspot region was
507 located at the CDS of the MMPL family transporter, which was the most reliably highly
508 expressed gene in Isolate57625 and the other four isolates in the RNA-seq analysis. The variant
509 pattern in the MMPL family transporter was classified into three types (Fig. 4b). Four sequences
510 (DRR317373, Japan; SRR7800518, US; SRR7800703, US; SRR7800409 US) were classified as
511 Type A and one sequence each was classified as Type B (SRR18969469, China) and Type C
512 (SRR7800511, US), the other 47 sequences had no mutation. There were three missense
513 mutations and 14 or 15 silent mutations in Type A, 20 missense mutations and 43 silent
514 mutations in Type B, and 47 missense mutations and 126 silent mutations in Type C (Table S7).
515 Surprisingly, despite many mutations, no mutation was responsible for the frameshift and
516 nonsense mutations. It was therefore speculated that the mutant MMPL family transporters
517 were being translated into full-length proteins in these isolates. Thus, these data indicated that
518 pMAB625-3 was spread across strains and regions, each acquiring unique mutations and some
519 clinical isolates retaining pMAB625-3 expressed mutated MMPL family tranporters.
520
521 Then, we classified pMAB625-2 sequences in terms of SVs. The sequences of pMAB625-2 were
522 classified into four types (Fig. 5). Assuming that Type I, which was the sequence of pMAB625-
523 2 in Isolate57625, was the base sequence, Type II and III had the large SVs at Region B, and

524 Type II but not Type III had at Region C. In contrast, Type IV had large SVs at Region A and C,
525 but not at Region B. Type I, Type III, and Type IV were detected in ABS, and Type II was detected
526 in MAS, except for ERR3198514. In addition, Type III was detected only in the isolates collected
527 in China, and Type IV was detected only in the isolates collected in the US. All pMAB625-2
528 sequences retained the ESX locus. These results indicated that pMAB625-2 had been
529 horizontally transferred between subspecies and had acquired the SVs unique to each
530 subspecies and region.
531
532 Finally, we investigated the distribution in our dataset regarding the plasmids identified in
533 other NTMs except for *M. abscessus* to verify whether plasmids transferred between *M.*
534 *abscessus* and other NTM species. Recently, Wetzstein et al. reported clinical and genomic
535 features in Europe about *Mycobacterium avium* complex (MAC) [58], the most frequently
536 detected NTM globally [69, 70]. In the report, they investigated the distribution of the known
537 152 plasmids registered in PLSDB and showed that plasmids identified in *M. abscessus*,
538 including pGD69B-1 (identical to pGD69A-1) and pGD42-1, which were identical to pMAB625-
539 3, were not detected in their dataset. Therefore, we investigated the distribution of these 152
540 plasmids in our dataset using our methods in this study. As a result, we detected the same
541 plasmids indicated in Table 3 but did not other plasmids identified in other NTM except for
542 pMN23 (accession number: NC_010604.1) and pMFLV03 (accession number: NC_009341.1) in
543 two and one isolates, respectively (Table S8). pMN23 was almost identical to pMAB23 (ANI:
544 99.9%), so pMN23 was detected redundantly in ATCC 19977 and SRR7800408, which had
545 pMAB23. Interestingly, pMFLV03 was detected in ERR3198408 (covered genome region: 78.7%,
546 depth of coverage: 73.1x). pMN23 had originally identified in *Mycobacterium marinum* [71]
547 and pMFLV03 had identified in *M. gilvum* PYR-GCK (according to the GenBank registry

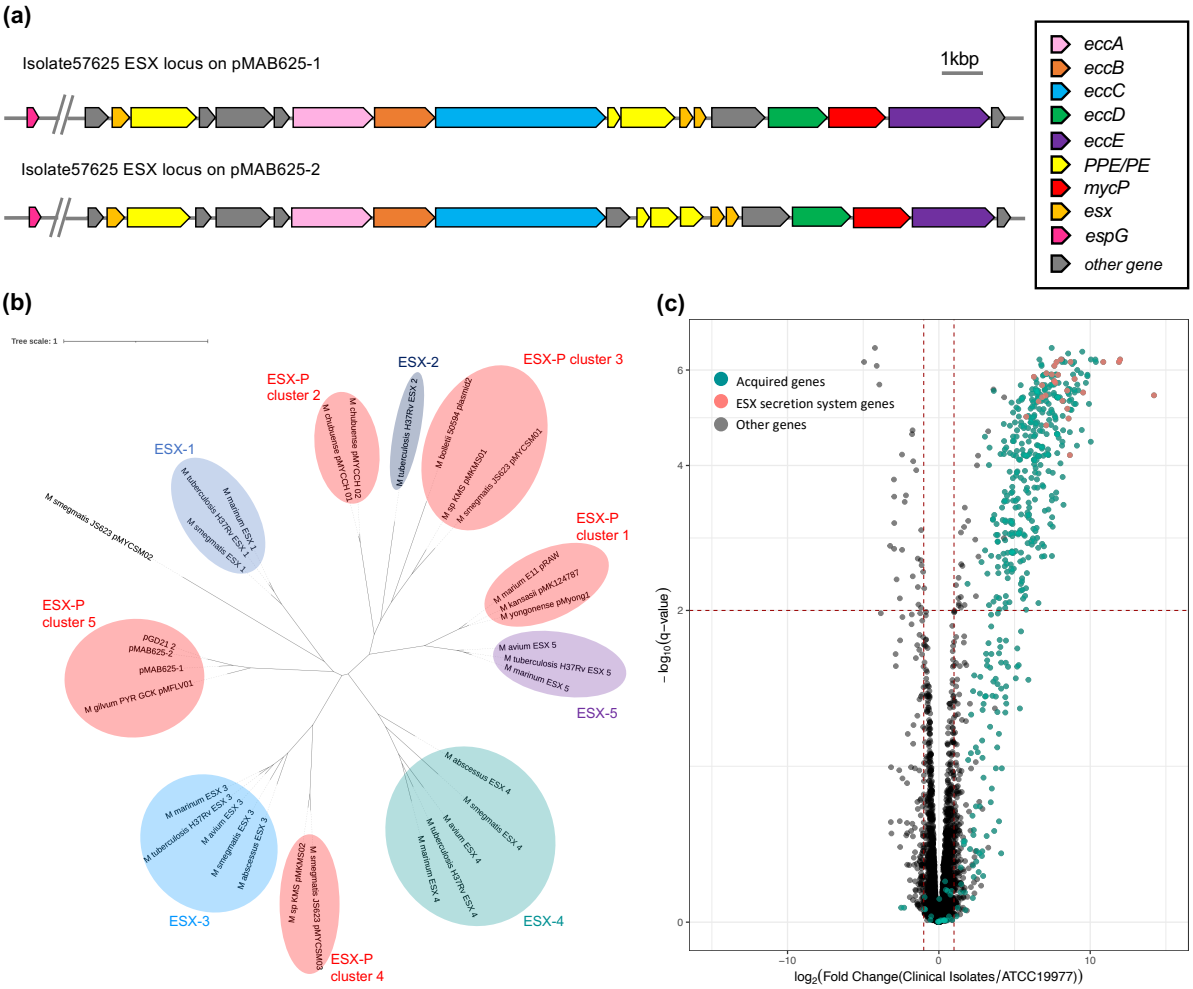
548 information), indicating that these plasmids could be present across NTM species. These

549 results suggested that plasmid distribution is strictly different from *M. abscessus* and other

550 NTMs except for very few cases.

551

552 # Figures and tables



553

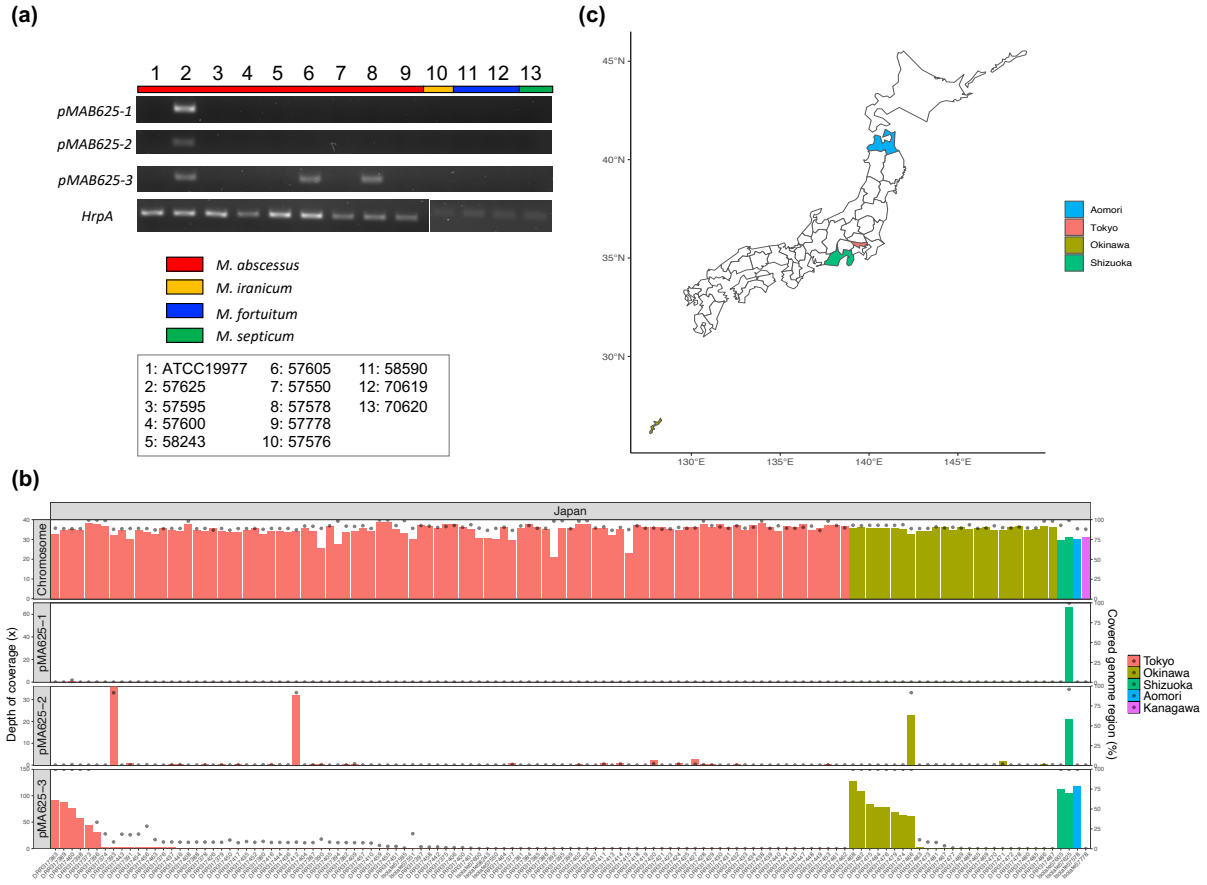
554 Fig. 1. Phylogenetic and gene expression analysis of the genes related to the ESX secretion

555 system. (a) The loci of the ESX secretion system are encoded on pMAB625-1 and pMAB625-2.

556 CDSs of the ESX secretion system were searched using Blast+. The core components that

557 construct membrane pore (*eccB, eccC, eccD, mycP,* and *eccE*) are contained in this operon. (b)

558 Phylogenetic tree analysis of various ESX loci encoded on chromosomes and plasmids. The
559 amino acid sequence of the genes that construct the membrane pore were concatenated and
560 the phylogenetic tree was created by NGphylogeny.fr [51] using the PhyML+SMS inference
561 method. The classification of each cluster was followed as previously reported [66]. (c) Volcano
562 plot of RNA-seq analysis of the type strain and the isolates. The green dots represent the data
563 of genes that the isolates acquired by SVs and plasmid acquisition. The red dots represent the
564 data of ESX secretion system genes encoded on the plasmids. The dashed line on the X-axis
565 represents the 2- and 0.5-fold change. The dashed line on the Y-axis represents the *q*-value of
566 0.01.



567

568 Fig. 2. The distribution of pMAB625 plasmid in the clinical isolates in Japan. (a) PCR analysis of

569 the possession of pMAB625 plasmid in the clinical isolates which our group isolated in Japan.
570 PCR was performed using the primers specific to each plasmid and *HrpA*. The colour bar
571 represents the species of each isolate. The primer sets of *HrpA* were different depending on
572 the species. (b) Mapping the read data of the clinical isolates in Japan to Isolate57625 genome
573 sequence. The bar plot represents the depth of coverage to chromosome and each plasmid
574 and the dot plot represents the percentage of covered genome region. (c) Geographical map
575 of Japan. The prefectures in which the pMAB625 plasmids were detected are represented by
576 colour.

577

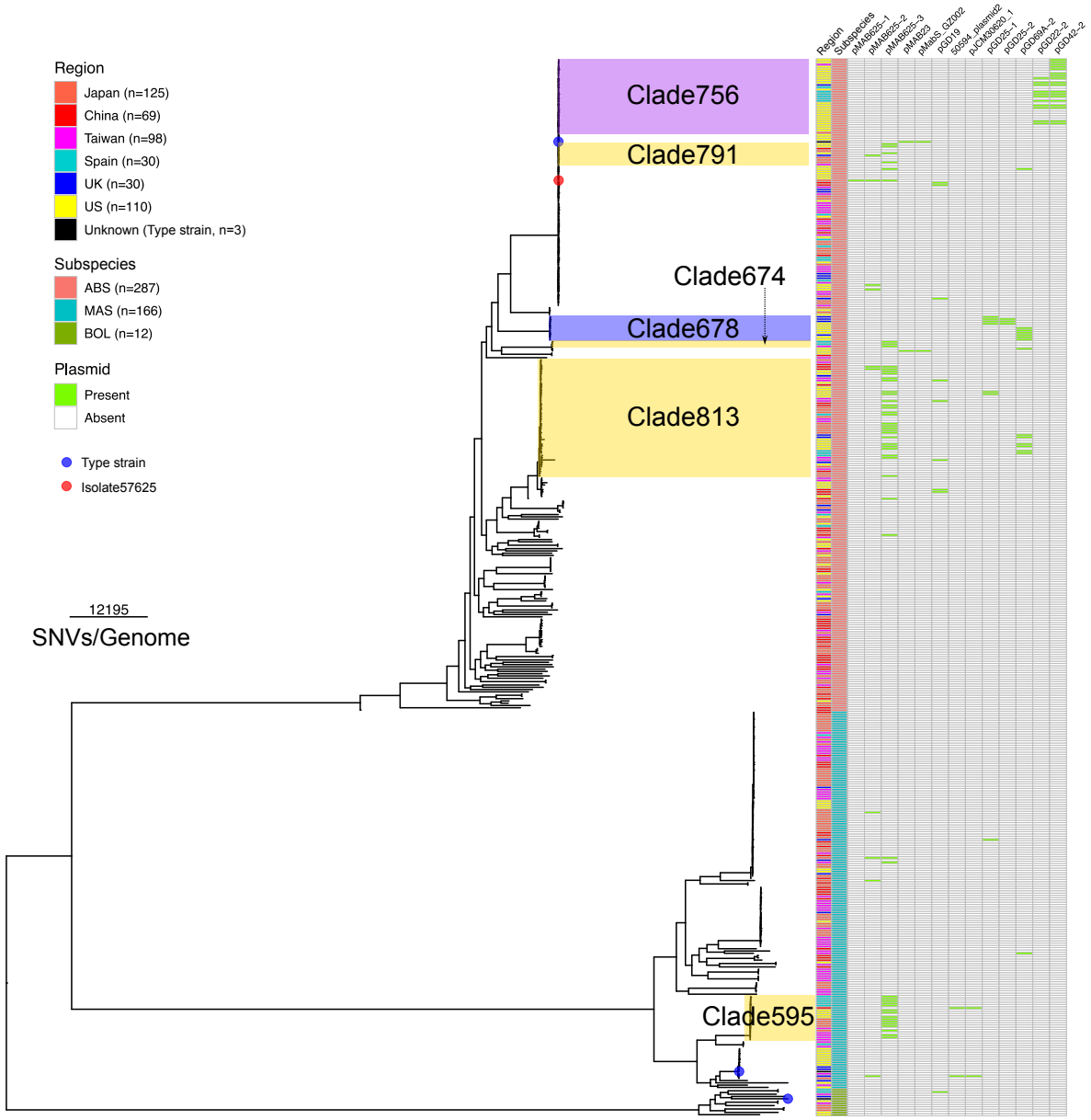


Fig. 3. The relation between the phylogenetic distance of host bacterial chromosomes and plasmid distribution in clinical isolates worldwide. The phylogenetic tree was created by Gubbins [59] using chromosome sequences and visualised by ggtree [60]. The read data of the clinical isolates collected worldwide were mapped to pMAB625 and known plasmid sequences to verify the plasmid distribution. Tree-scale represents the number of SNVs per genome.

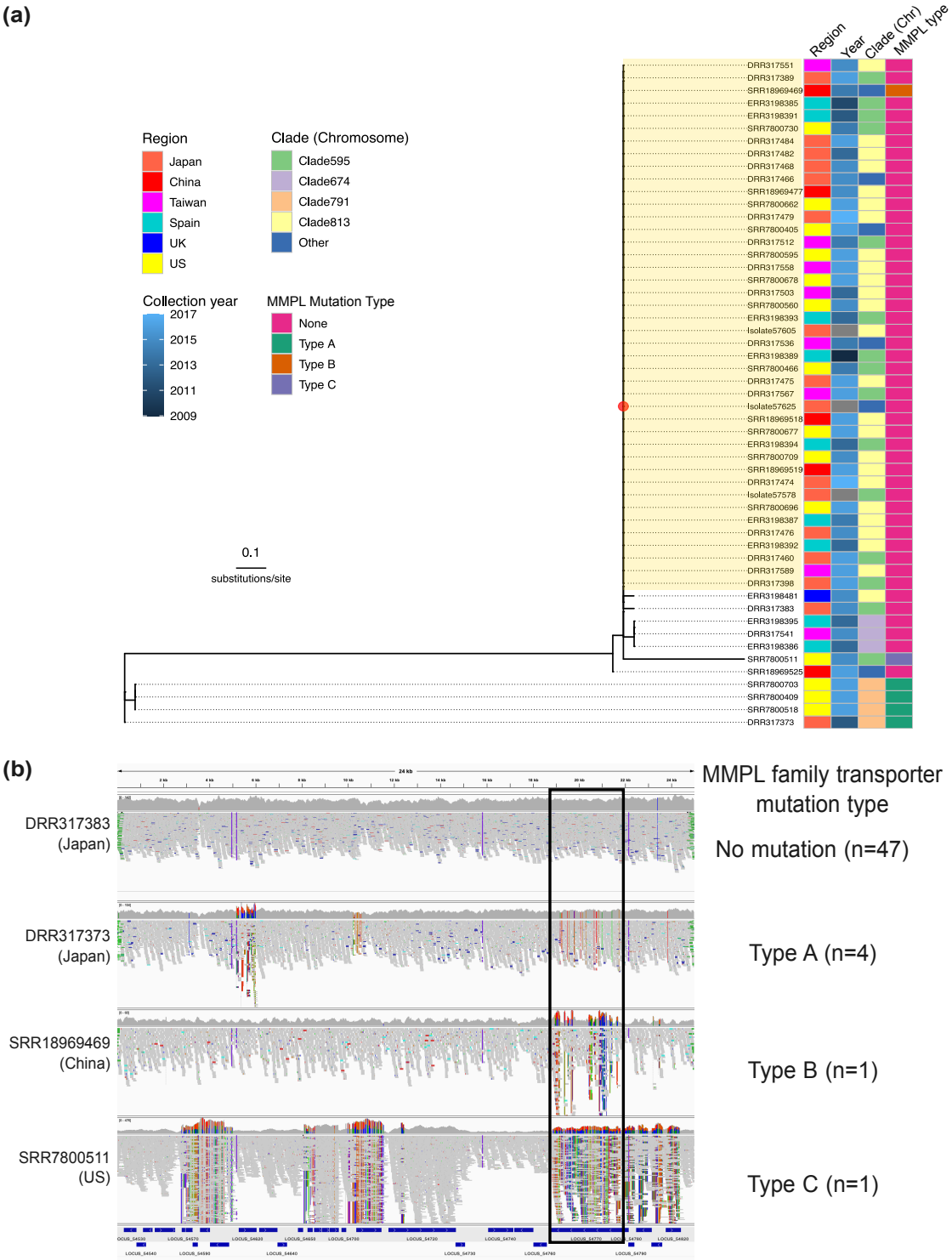
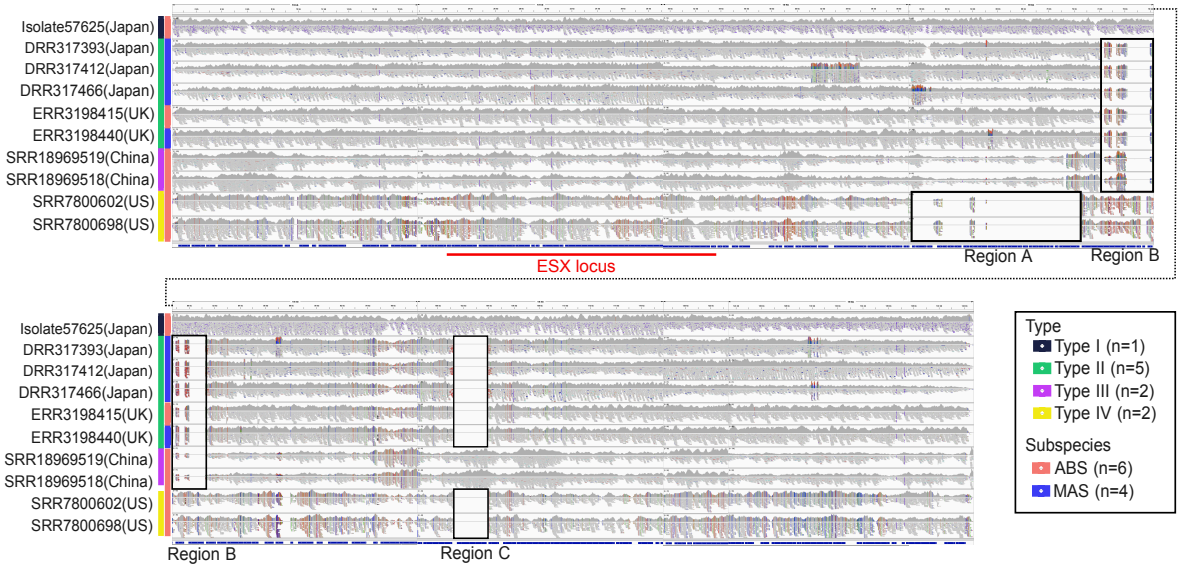


Fig. 4 Phylogenetic tree and SNVs analysis of pMAB625-3. (a) Phylogenetic tree of the sequences of pMAB625-3 detected in the 53 isolates. The phylogenetic tree was created using the maximum-likelihood method using iQ-TREE2 based on the core-SNVs data. The sequences belonging to the clade covered by yellow colour were identical, although the sequence of

590 SRR18969469 was heterogeneous. The classification of "Clade (Chromosome)" is based on the
591 clade in the phylogenetic tree indicated in Fig. 3. (b) The classification of the pMAB625
592 sequences based on the SNVs. The mapping results of the reads data from each isolate to
593 pMAB625-3 were also visualised by IGV. The mutation hotspot was located at the MMPL family
594 transporter (surrounded by a black line square). The mutation types of the MMPL transporters
595 were classified into three types, and all mutations at this region were missense or silent.



596

597 Fig. 5. Comparison of the pMAB625-2 sequences. The mapping results of the reads data from
598 each isolate to pMAB625-2 were visualised using IGV. Each sequence was classified into 4 types
599 by the pattern of structural variants (surrounded by black line squares).

600

601

602

603

604

605

606 Table 1. The isolation time point of the clinical isolates and antimicrobial susceptibility of the

607 type strain and the isolates

Isolates	Source	Period from the first isolation (month)	MIC* (μg/mL)														
			AMK	TOB	IPM	FRPM	LVFX	MFLX	AZM	CAM 4day	CAM 14day	ST	DOXY	MEPM	LZD	CLF	STFX
ATCC 19977	-	-	8	8	16	>64	32	>8	64	2	>64	>152/8	>16	64	32	1	2
Isolate57625	catheter	0	8	8	16†	>64	16	>8	64	2	>64	>152/8	>16	>64	32	1	2
Isolate57629	catheter	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Isolate57630	bone biopsy	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Isolate57596	heel	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Isolate57626	malleolus	6	8	8	>64 ‡	>64	32	>8	64	2	>64	>152/8	>16	>64	32	1	2

608 *MICs were determined as the minimum concentration at which the bacteria did not grow. No

609 antimicrobial susceptibility testing was conducted on the isolates other than Isolate57625 and

610 Isolate57626, representing hyphens in the table. The test was repeated three times and all

611 MICs were confirmed to be the same between the experiments with the exception of imipenem

612 for the isolates. AMK, Amikacin; TOB, Tobramycin; IPM, Imipenem; FRPM, Faropenem; LVFX,

613 Levofloxacin; MFLX, Moxifloxacin; AZM, Azithromycin; CAM, Clarithromycin; ST,

614 Trimethoprim/sulfamethoxazole; DOXY, Doxycycline; MEPM, Meropenem; LNZ, Linezolid; CLF,

615 Clofazimine; STFX, Sitafloxacin

616 †For Isolate57625, the MICs against IPM in the three experiments were 64, 16, and 16 μg/mL,

617 respectively.

618 ‡For Isolate57626, the MICs against IPM in the three experimenst were >64, 16, and >64 μg/mL,

619 respectively.

620

621

622

623 Table 2. Genome assembly and the annotation statics of the Isolate57625

Contig	Form	Length(bp)	GC contents (%)	Number of CDS	Number of rRNA	Number of tRNA	Number of tmRNA
Chromosome	circular	5,072,346	64.1	5,034	3	48	1
pMAB625-1	linear	202,398	64.7	247	0	29	0
pMAB625-2	linear	142,689	62.2	171	0	0	0
pMAB625-3	circular	24,988	65.6	30	0	0	0

624

625 Table 3. Homology search result between plasmids identified in *M. abscessus*

Plasmid	Accession no.	Topology	Length (bp)	Homologous plasmid (ANI(%)>99)	Homologous plasmid (97< ANI(%) <99)	Use for "Reference"
pMAB625-1	AP028614.1	linear	202,398	-	-	TRUE
pMAB625-2	AP028615.1	linear	142,689	-	pGD21-2†	TRUE
pMAB625-3	AP028616.1	circular	24,988	pGD69A-1, pGD42-1	-	TRUE
pMAB23	NC_010394.1	circular	23,319	pMabS_GZ002*, FDAARGOS_1604_plasmid	-	TRUE
pMabS_GZ002	NZ_CP034180.1	linear	15,203	pMAB23, FDAARGOS_1604_plasmid	-	TRUE
pGD19	NZ_CP063329.1	circular	18,605	-	-	TRUE
pGD21-1	NZ_CP065285.1	circular	112,633	-	-	TRUE
pGD21-2	NZ_CP065286.1	linear	155,609	-	pMAB625-2†	TRUE
pGD22-1	NZ_CP063325.1	circular	19,694	-	-	TRUE
pGD22-2	NZ_CP063326.1	circular	18,117	pGD42-2*	-	TRUE
pGD25-1	NZ_CP063321.1	circular	31,413	-	-	TRUE
pGD25-2	NZ_CP063322.1	circular	27,424	-	-	TRUE
pGD25-3	NZ_CP063323.1	circular	23,599	-	-	TRUE
pGD42-1	NZ_CP065281.1	circular	24,993	pMAB625-3, pGD69A-1	-	FALSE
pGD42-2	NZ_CP065282.1	circular	9,547	pGD22-2	pGD69A-2	TRUE
pGD69A-1	NZ_CP065270.1	circular	25,000	pMAB625-3, pGD42-1	-	FALSE
pGD69A-2	NZ_CP065271.1	circular	18,611	-	pGD42-2*	TRUE
pJCM30620_1	NZ_AP022622.1	linear	119,389	50594_plasmid2*	-	TRUE
pJCM30620_2	NZ_AP022623.1	linear	38,763	50594_plasmid1†	-	TRUE
BRA100	NC_017908.2	circular	56,265	-	-	TRUE
FDAARGOS_1604_plasmid	NZ_CP085920.1	circular	23,319	pMAB23, pMabS_GZ002*	-	FALSE
50594_plasmid1	NC_021278.1	circular	172,814	pJCM30620_2*	-	TRUE
50594_plasmid2	NC_021279.1	circular	97,240	pJCM30620_1†	-	TRUE

626 * One Structural Variant (SV) exists in the sequence.

627 † More than two SVs exist in the sequence.

628

629

630 Table 4. pMAB625 and known plasmids distribution in the clinical isolates worldwide

	Number (%) of isolates retaining plasmid							
Region	Unknown (Type strain)	Japan	China	Taiwan	US	Spain	UK	total*
Analyzed strains	3	125	69	98	110	30	30	462
pMAB625-1	0	1 (0.8)	0	0	0	0	0	1 (0.2)
pMAB625-2	0	4 (3.2)	2 (2.9)	0	2 (1.8)	0	2 (6.7)	10 (2.2)
pMAB625-3	0	16 (12.8)	5 (7.2)	8 (8.2)	14 (12.7)	9 (30.0)	1 (3.3)	53 (11.5)
pMAB23	1 (33.3)	0	0	0	1 (0.9)	0	0	2 (0.4)
pMabS_GZ002	1 (33.3)	0	0	0	1 (0.9)	0	0	2 (0.4)
pGD69A-2	0	0	0	1 (1.0)	9 (8.2)†	2 (6.7)	3 (10.0)	15(3.2)
pGD19	0	0	4 (5.8)	3 (3.1)	0	1 (3.3)	1 (3.3)	9(1.9)
pJCM30620_1	0	0	1 (1.4)	0	0	0	1 (3.3)	2(0.4)
50594_plasmid2	0	0	1 (1.4)	0	0	0	1 (3.3)	2(0.4)
pGD22-2	0	0	0	0	6 (5.5)	4 (13.3)‡	1 (3.3)	11(2.4)
pGD42-2	0	0	0	1 (1.0)	12 (10.9)§	4 (13.3)†	1 (3.3)	18(3.9)
pGD25-1	0	0	0	0	3 (2.7)	0	4 (13.3)‡	7(1.5)
pGD25-2	0	0	0	0	1 (0.9)	0	2 (6.7)	3(0.6)

631 Fisher's exact tests were performed to verify regional differences in the frequency of
632 appearance of each plasmid. Corrections for the multiple comparisons were performed using
633 Holm's method.
634 *Total number excludes the number of the type strains.
635 †$p<0.05$, compared with Japan; ‡$p<0.05$, compared with Japan and Taiwan; §$p<0.05$, compared
636 with Japan, China and Taiwan.
637

638 # Discussion

639 Plasmids are mobile genetic elements that can move between different bacterial strains and
640 alter host characteristics, such as virulence, and antimicrobial susceptibility. However, it
641 remains unclear whether plasmids confer AMR in *M. abscessus*. In this study, we showed that
642 the multidrug-resistant clinical isolates of ABS isolated from a patient with osteomyelitis had
643 three plasmids (two of three were linear, and one was circular) that the type strain did not
644 harbour. In addition, several factors associated with virulence and antimicrobial susceptibility
645 were encoded on the plasmids. Analysis of the clinical isolates identified worldwide indicated

646 that pMAB625 plasmids were horizontally transferred, and these plasmid transfers could be
647 the driving force of the genomic diversity among the clinical isolates of *M. abscessus*.
648
649 A recent study also showed that prophages and plasmids in various clinical isolates
650 contributed to diversity [72]. They reported that most strains have less than two types of
651 plasmids and that strains with more than three types of plasmids were rare. The type of plasmid
652 could contribute to the genomic diversity of the host bacterium, and our study indicated the
653 diverse distribution of plasmids in *M. abscessus*. Furthermore, our analysis also revealed that
654 pMAB625-3 which were detected in the isolates belonging to the Clade791, which were
655 genetically close to Isolate57625, had many SNVs (Fig. 4b). Given the rate of replication, the
656 mutation rate of high-copy plasmids is expected to be faster than that of chromosomes.
657 Therefore, the genomic diversity could accelerate faster in high-copy plasmids than in
658 chromosomes and more easily modify features of the host bacterial genome by horizontal
659 transfer. In addition, the copy number analysis indicated that Isolate57625 was a
660 heterogeneous population with some remaining pMAB625-1 and pMAB625-2 but others not
661 (Fig. S5). Retaining plasmids is detrimental to bacterial growth, however, if genes for AMR are
662 encoded on the plasmid, it is advantageous for survival during antimicrobial exposure.
663 Therefore, a certain number of bacteria that retain the plasmid may be present in the
664 population as a survival strategy against antimicrobial agents while growing as a population.
665
666 We identified various factors associated with multidrug resistance on the pMAB625 plasmids.
667 The components of the ESX secretion system, the MMPL family transporter, and the
668 components of the TA system were encoded on the plasmids, and these genes were highly
669 expressed in the clinical isolates compared to the type strain. The ESX secretion system is

670 involved in the export of the various pathogenic proteins and metal ions associated with
671 proliferation, survival in the cytosol of macrophage, and conjugation. It has also been reported
672 that ESX-1 components are required for sliding motility and biofilm formation in *M. avium* [65].
673 Therefore, given the close genetic distance between ESX-1 and ESX-P cluster 5, which include
674 the components of the ESX secretion system on the pMAB625-1 and pMAB625-2 (Fig. 1b), the
675 components of ESX-P cluster 5 may be involved in biofilm formation. In contrast, MMPL family
676 transporters can transport antibiotics across the cell membrane. It has been reported that the
677 increased expression of MMPL5 is associated with decreased susceptibility to bedaquiline, and
678 clofazimine in *M. tuberculosis* [73, 74], bedaquiline, and clofazimine in *M. intracellulare* [75],
679 and thiactazone derivative in *M. abscessus* [76]. The function of the MMPL family transporter
680 encoded on the pMAB625-3 is unknown; however, the presence of mutational hotspots (Fig.
681 4b) indicates that this MMPL transporter may be compatible with a variety of substrates,
682 including antimicrobial agents. Therefore, our results suggest that Isolate57625 enhanced
683 defense mechanisms against antibiotics through the acquisition of the ESX secretion system
684 and transporters, accompanied by changes in biofilm barriers, and transport of antibiotics. TA
685 systems are known as the plasmid maintenance systems; however, they can transform the
686 bacterial phenotype into dormancy by inhibiting vital cellular processes, such as proton-
687 motive force and translation and contribute to antimicrobial torelance [77, 78]. Thus, the
688 acquisition of TA systems by the plasmid may be the cause of the decreased growth rate of
689 Isolate57625 and Isolate57626 compared to the type strain (Fig. S1). These results support that
690 pMAB625 plasmids appeared to be the main cause of the low susceptibility to carbapenem
691 antibiotics and intractability to treatment; however, further studies will be essential to
692 determine whether these plasmids are the main cause of AMR.
693

694 As important findings, we found that 2.2% (10/462) and 11.5% (53/462) of the clinical isolates
695 worldwide had pMAB625-2, and pMAB625-3, respectively (Table 4). This result provides two
696 significant insights. The first is an epidemiological insight into the spread of plasmids in *M.*
697 *abscessus* beyond regions and strains. Because *M. abscessus* has not been known to cause
698 human-to-human transmission, this global spread of pMAB625 plasmids may be mainly due
699 to increased global logistical activity and human mobility with associated interference of
700 human and microbial habitats [79]. The habitat of *M. abscessus* is soil and water systems, and
701 it is also possible that increased flooding due to recent climate change may have increased
702 the chance of contact between the different strains. Analysis of plasmid distribution revealed
703 that some plasmids (pMAB625-2, pMAB625-3, pGD19, pGD25-1 and pGD69A-2) have been
704 widely spread between strains, regions and subspecies, while others (pGD22-2 and pGD42-2)
705 have been restricted to specific strains and regions (Fig. 3). Therefore, it is speculated that
706 regarding pGD22-2 and pGD42-2, the identical isolates retaining these plasmids had been
707 spread within or between the regions, whereas, regarding pMAB625-2, pMAB625-3, pGD19,
708 pGD25-1 and p69A-2, plasmids themselves had been spread between different strains,
709 subspecies and region through horizontal transfer. It is unknown what factors determine these
710 differences in plasmid distribution. It is possible that the factors encoded on the plasmid or
711 the periods from the emergence were different among the plasmids. In addition, we also
712 showed the plasmid distribution was different from *M. abscessus* and MAC (Table S8), indicating
713 that incompatibility also exists in the plasmids present in NTM, strictly distinguishing *M.*
714 *abscessus* from other NTMs except for very few cases, such as pMN23 and pMFLV03. The
715 second is the potential of plasmids as pharmacological targets. The mechanisms of action of
716 current antimicrobials inhibit specific cellular processes; therefore, protein mutation at the
717 point of action or the acquisition of a multidrug transporter can lead to the emergence of

718 resistant strains. Plasmids such as pMAB625-3, which are highly prevalent in multidrug-
719 resistant bacteria, can be promising targets for the therapy using gene-editing methods such
720 as CRISPR-Cas13a because these methods do not necessarily target genes or proteins that are
721 targeted by conventional antibiotics. The CRISPR-Cas13a system can recognise mRNA
722 expressed not only from the loci on the chromosome but also on plasmids and induce non-
723 specific RNA degradation and cell death [80]. In addition, the gene-editing method can also
724 target multiple regions of locus. Thus, the multiple, low-mutation loci encoded on plasmids,
725 which are only found in antimicrobial-resistant bacteria, can be pharmaceutical targets that
726 are robust to genetic mutations using the CRISPR-Cas13a system.
727
728 In summary, we have identified the plasmids that have been horizontally transferred among
729 the multidrug-resistant clinical isolates of *M. abscessus* worldwide. These findings provide a
730 novel perspective on the acquisition of genomic diversity in *M. abscessus*. Future studies will
731 be required to verify whether the pMAB625 plasmids induce multidrug resistance in *M.*
732 *abscessus* by plasmid transfection. These studies will lead to the development of robust
733 therapies against multidrug resistance caused by genetic mutations, and a rapid diagnostic
734 method to prevent antibiotics abuse, and an understanding of the mechanisms of multidrug
735 resistance in *M. abscessus*.
736

737 Author statements

738 **Author contributions:**

739 Kensuke Ohse: Conceptualization, Formal analysis, Methodology, Data curation, Investigation,
740 Software, Validation, Visualization, Writing—original draft, Writing—review & editing. Atushi
741 Yoshida: Conceptualization, Methodology, Funding acquisition, Writing—original draft,

Writing—review & editing. Kamada Keisuke: Investigation. Hironobu Kitazawa, Yusuke Ito, Takayo Shoji, and Kenichiro Watanabe: Patient care and Resources. Hiroshi Koganemaru: Methodology, Patient care, and Resources. Ken Kikuchi: Investigation and Funding acquisition. Masashi Toyoda: Project administration, Supervision, Writing—review & editing.

Conflicts of interest:

The authors declare no conflict of interest.

Funding information:

This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP19K08938.

Ethical approval:

Ethical approval for this study was obtained from the Ethics Committee of Dokkyo Medical University (approval number 28045).

Consent for publication:

This manuscript does not contain personal details of individuals that could lead to their identification.

Acknowledgements:

The authors thank the staff of the microbiology laboratory of Dokkyo Medical University Hospital for performing identification and drug susceptibility tests, and Dr. Yoshishige Masuda and Dr. Takashi Inamatsu at TMIG for supervising in clinical practice. Computations were partially performed on the NIG supercomputer at ROIS National Institute of Genetics.

References

764 1. **Naghavi M, Vollset SE, Ikuta KS, Swetschinski LR, Gray AP, *et al.*** Global
765 burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with
766 forecasts to 2050. *The Lancet* 2024;404:1199–1226.
767 2. **Johansen MD, Herrmann J-L, Kremer L**. Non-tuberculous mycobacteria and
768 the rise of Mycobacterium abscessus. *Nat Rev Microbiol* 2020;18:392–407.
769 3. **Cristancho-Rojas C, Varley CD, Lara SC, Kherabi Y, Henkle E, *et al.***
770 Epidemiology of *Mycobacterium abscessus*. *Clin Microbiol Infect* 2024;30:712–717.
771 4. **Rubio M, March F, Garrigó M, Moreno C, Español M, *et al.*** Inducible and
772 Acquired Clarithromycin Resistance in the Mycobacterium abscessus Complex. *PLoS*
773 *ONE* 2015;10:e0140166.
774 5. **Huh HJ, Kim S-Y, Jhun BW, Shin SJ, Koh W-J**. Recent advances in molecular
775 diagnostics and understanding mechanisms of drug resistance in nontuberculous
776 mycobacterial diseases. *Infect Genet Evol* 2019;72:169–182.
777 6. **López-Roa P, Esteban J, Muñoz-Egea M-C**. Updated Review on the
778 Mechanisms of Pathogenicity in Mycobacterium abscessus, a Rapidly Growing
779 Emerging Pathogen. *Microorganisms*;11. Epub ahead of print January 2023. DOI:
780 10.3390/microorganisms11010090.
781 7. **Clinical & Laboratory Standards Institute**. *Susceptibility Testing of*
782 *Mycobacteria, Nocardia spp., and Other Aerobic Actinomycetes, M24Ed3rd*, CLSI Puplication,
783 Pennsylvania
784 8. **Jong B-E, Wu T-S, Chen N-Y, Yang C-H, Shu C-C, *et al.*** Impact on
785 Macrolide Resistance of Genetic Diversity of Mycobacterium abscessus Species.
786 *Microbiol Spectr*;10:e02749-22.

787 9. **Solanki P, Lipman M, McHugh TD, Satta G**. Whole genome sequencing and
788 prediction of antimicrobial susceptibilities in non-tuberculous mycobacteria. *Front*
789 *Microbiol* 2022;13:1044515.
790 10. **Lipworth S, Hough N, Buchanan R, Smith EG, Robinson E, *et al.*** Improved
791 Performance Predicting Clarithromycin Resistance in Mycobacterium abscessus on an
792 Independent Data Set. *Antimicrob Agents Chemother* 2019;63:e00400-19.
793 11. **Akiyama Y, Iwamoto N, Kamada K, Yoshida A, Osugi A, *et al.*** Aortic
794 endograft infection by Mycobacterium abscessus subsp. massiliense with acquired
795 clarithromycin resistance: a case report. *BMC Infect Dis* 2023;23:694.
796 12. **Moriyama M, Ogawa K, Nakagawa T, Nikai T, Uchiya K**. ASSOCIATION
797 BETWEEN A pMAH135 PLASMID AND THE PROGRESSION OF PULMONARY DISEASE
798 CAUSED BY MYCOBACTERIUM AVIUM.
799 13. **van Ingen J, Kohl TA, Kranzer K, Hasse B, Keller PM, *et al.*** Global outbreak
800 of severe *Mycobacterium chimaera* disease after cardiac surgery: a molecular
801 epidemiological study. *Lancet Infect Dis* 2017;17:1033–1041.
802 14. **Queenan AM, Bush K**. Carbapenemases: the Versatile β-Lactamases. *Clin*
803 *Microbiol Rev* 2007;20:440–458.
804 15. **Zhang Y, Li Z, He X, Ding F, Wu W, *et al.*** Overproduction of efflux pumps
805 caused reduced susceptibility to carbapenem under consecutive imipenem-selected
806 stress in Acinetobacter baumannii. *Infect Drug Resist* 2017;11:457.
807 16. **Ranjitkar S, Reck F, Ke X, Zhu Q, McEnroe G, *et al.*** Identification of
808 Mutations in the mrdA Gene Encoding PBP2 That Reduce Carbapenem and
809 Diazabicyclooctane Susceptibility of Escherichia coli Clinical Isolates with Mutations in
810 ftsI (PBP3) and Which Carry blaNDM-1. *mSphere* 2019;4:e00074-19.

811 17. **Kamada K, Yoshida A, Iguchi S, Arai Y, Uzawa Y, *et al.*** Nationwide
812 surveillance of antimicrobial susceptibility of 509 rapidly growing mycobacteria
813 strains isolated from clinical specimens in Japan. *Sci Rep* 2021;11:12208.
814 18. **Bouso JM, Planet PJ**. Complete nontuberculous mycobacteria whole
815 genomes using an optimized DNA extraction protocol for long-read sequencing.
816 *BMC Genomics* 2019;20:793.
817 19. **De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C**.
818 NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics*
819 2018;34:2666–2669.
820 20. **Wick RR, Judd LM, Gorrie CL, Holt KE**. Completing bacterial genome
821 assemblies with multiplex MinION sequencing. *Microb Genomics* 2017;3:e000132.
822 21. **Wang JR, Holt J, McMillan L, Jones CD**. FMLRC: Hybrid long read error
823 correction using an FM-index. *BMC Bioinformatics* 2018;19:50.
824 22. **Kolmogorov M, Yuan J, Lin Y, Pevzner PA**. Assembly of long, error-prone
825 reads using repeat graphs. *Nat Biotechnol* 2019;37:540–546.
826 23. **Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, *et al.*** Pilon: An
827 Integrated Tool for Comprehensive Microbial Variant Detection and Genome
828 Assembly Improvement. *PLOS ONE* 2014;9:e112963.
829 24. **Gurevich A, Saveliev V, Vyahhi N, Tesler G**. QUAST: quality assessment tool
830 for genome assemblies. *Bioinformatics* 2013;29:1072–1075.
831 25. **Wick RR, Schultz MB, Zobel J, Holt KE**. Bandage: interactive visualization of
832 de novo genome assemblies. *Bioinformatics* 2015;31:3350–3352.

833 26. **Tanizawa Y, Fujisawa T, Nakamura Y**. DFAST: a flexible prokaryotic genome
834 annotation pipeline for faster genome publication. *Bioinformatics* 2018;34:1037–
835 1039.
836 27. **Tanizawa Y, Fujisawa T, Kaminuma E, Nakamura Y, Arita M**. DFAST and
837 DAGA: web-based integrated genome annotation tools and resources. *Biosci*
838 *Microbiota Food Health* 2016;35:173–184.
839 28. **Schmartz GP, Hartung A, Hirsch P, Kern F, Fehlmann T, *et al.*** PLSDB:
840 advancing a comprehensive database of bacterial plasmids. *Nucleic Acids Res*
841 2022;50:D273–D278.
842 29. **Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S**. High
843 throughput ANI analysis of 90K prokaryotic genomes reveals clear species
844 boundaries. *Nat Commun* 2018;9:5114.
845 30. **Darling ACE, Mau B, Blattner FR, Perna NT**. Mauve: Multiple Alignment of
846 Conserved Genomic Sequence With Rearrangements. *Genome Res* 2004;14:1394–
847 1403.
848 31. **Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, *et al.***
849 BLAST+: architecture and applications. *BMC Bioinformatics* 2009;10:421.
850 32. **Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, *et al.*** Validating
851 the AMRFinder Tool and Resistance Gene Database by Using Antimicrobial
852 Resistance Genotype-Phenotype Correlations in a Collection of Isolates. *Antimicrob*
853 *Agents Chemother* 2019;63:e00483-19.
854 33. **Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, *et al.*** CARD 2017:
855 expansion and model-centric curation of the comprehensive antibiotic resistance
856 database. *Nucleic Acids Res* 2017;45:D566–D573.

857 34. **Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, *et al.***
858 Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother*
859 2012;67:2640–2644.
860 35. **Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, *et al.***
861 ARG-ANNOT, a New Bioinformatic Tool To Discover Antibiotic Resistance Genes in
862 Bacterial Genomes. *Antimicrob Agents Chemother* 2014;58:212–220.
863 36. **Chen L, Zheng D, Liu B, Yang J, Jin Q**. VFDB 2016: hierarchical and refined
864 dataset for big data analysis—10 years on. *Nucleic Acids Res* 2016;44:D694–D697.
865 37. **Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, *et***
866 ***al.*** In Silico Detection and Typing of Plasmids using PlasmidFinder and Plasmid
867 Multilocus Sequence Typing. *Antimicrob Agents Chemother* 2014;58:3895–3903.
868 38. **Ingle DJ, Valcanis M, Kuzevski A, Tauschek M, Inouye M, *et al.*** In silico
869 serotyping of E. coli from short read data identifies limited novel O-loci but extensive
870 diversity of O:H serotype combinations within and between pathogenic lineages.
871 *Microb Genomics* 2016;2:e000064.
872 39. **Doster E, Lakin SM, Dean CJ, Wolfe C, Young JG, *et al.*** MEGARes 2.0: a
873 database for classification of antimicrobial drug, biocide and metal resistance
874 determinants in metagenomic sequence data. *Nucleic Acids Res* 2020;48:D561–D569.
875 40. **Kapopoulou A, Lew JM, Cole ST**. The MycoBrowser portal: a comprehensive
876 and manually annotated resource for mycobacterial genomes. *Tuberc Edinb Scotl*
877 2011;91:8–13.
878 41. **Li H**. Aligning sequence reads, clone sequences and assembly contigs with
879 BWA-MEM. Epub ahead of print 26 May 2013. DOI: 10.48550/arXiv.1303.3997.

880 42. **Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, *et al.*** Twelve years of
881 SAMtools and BCFtools. *GigaScience* 2021;10:giab008.
882 43. **Garrison E, Marth G**. Haplotype-based variant detection from short-read
883 sequencing. Epub ahead of print 17 July 2012. DOI: 10.48550/arXiv.1207.3907.
884 44. **Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, *et al.*** A program for
885 annotating and predicting the effects of single nucleotide polymorphisms, SnpEff:
886 SNPs in the genome of Drosophila melanogaster strain w ^{1118}; iso-2; iso-3. *Fly*
887 *(Austin)* 2012;6:80–92.
888 45. **Li H**. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*
889 2018;34:3094–3100.
890 46. **Kim D, Langmead B, Salzberg SL**. HISAT: a fast spliced aligner with low
891 memory requirements. *Nat Methods* 2015;12:357–360.
892 47. **Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, *et al.***
893 StringTie enables improved reconstruction of a transcriptome from RNA-seq reads.
894 *Nat Biotechnol* 2015;33:290–295.
895 48. **Sun J, Nishiyama T, Shimizu K, Kadota K**. TCC: an R package for comparing
896 tag count data with robust normalization strategies. *BMC Bioinformatics* 2013;14:219.
897 49. **Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, *et al.***
898 IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the
899 Genomic Era. *Mol Biol Evol* 2020;37:1530–1534.
900 50. **Letunic I, Bork P**. Interactive Tree Of Life (iTOL) v5: an online tool for
901 phylogenetic tree display and annotation. *Nucleic Acids Res* 2021;49:W293–W296.

902 51. **Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F, *et al.***
903 NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic*
904 *Acids Res* 2019;47:W260–W265.
905 52. **Yoshida M, Chien J-Y, Morimoto K, Kinjo T, Aono A, *et al.*** Molecular
906 Epidemiological Characteristics of Mycobacterium abscessus Complex Derived from
907 Non-Cystic Fibrosis Patients in Japan and Taiwan. *Microbiol Spectr* 2022;10:e00571-
908 22.
909 53. **Jin P, Dai J, Guo Y, Wang X, Lu J, *et al.*** Genomic Analysis of Mycobacterium
910 abscessus Complex Isolates from Patients with Pulmonary Infection in China.
911 *Microbiol Spectr* 2022;10:e00118-22.
912 54. **Chen S**. Ultrafast one-pass FASTQ data preprocessing, quality control, and
913 deduplication using fastp. *iMeta* 2023;2:e107.
914 55. **Shen W, Le S, Li Y, Hu F**. SeqKit: A Cross-Platform and Ultrafast Toolkit for
915 FASTA/Q File Manipulation. *PLOS ONE* 2016;11:e0163962.
916 56. **García-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Götz S, *et al.***
917 Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics*
918 2012;28:2678–2679.
919 57. **Quinlan AR, Hall IM**. BEDTools: a flexible suite of utilities for comparing
920 genomic features. *Bioinformatics* 2010;26:841–842.
921 58. **Wetzstein N, Diricks M, Anton TB, Andres S, Kuhns M, *et al.*** Clinical and
922 genomic features of Mycobacterium avium complex: a multi-national European
923 study. *Genome Med* 2024;16:86.

924 59. **Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, *et al.*** Rapid
925 phylogenetic analysis of large samples of recombinant bacterial whole genome
926 sequences using Gubbins. *Nucleic Acids Res* 2015;43:e15.
927 60. **Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y**. ggtree: an r package for
928 visualization and annotation of phylogenetic trees with their covariates and other
929 associated data. *Methods Ecol Evol* 2017;8:28–36.
930 61. **Paulsen IT, Brown MH, Littlejohn TG, Mitchell BA, Skurray RA**. Multidrug
931 resistance proteins QacA and QacB from Staphylococcus aureus: membrane
932 topology and identification of residues involved in substrate specificity. *Proc Natl*
933 *Acad Sci* 1996;93:3630–3635.
934 62. **Chuanchuen R, Gaynor JB, Karkhoff-Schweizer R, Schweizer HP**. Molecular
935 Characterization of MexL, the Transcriptional Repressor of the mexJK Multidrug Efflux
936 Operon in Pseudomonas aeruginosa. *Antimicrob Agents Chemother* 2005;49:1844–
937 1851.
938 63. **Osman MM, Shanahan JK, Chu F, Takaki KK, Pinckert ML, *et al.*** The C
939 terminus of the mycobacterium ESX-1 secretion system substrate ESAT-6 is required
940 for phagosomal membrane damage and virulence. *Proc Natl Acad Sci*
941 2022;119:e2122161119.
942 64. **Laencina L, Dubois V, Le Moigne V, Viljoen A, Majlessi L, *et al.***
943 Identification of genes required for Mycobacterium abscessus growth in vivo with a
944 prominent role of the ESX-4 locus. *Proc Natl Acad Sci* 2018;115:E1002–E1011.
945 65. **Lai L-Y, Lin T-L, Chen Y-Y, Hsieh P-F, Wang J-T**. Role of the Mycobacterium
946 marinum ESX-1 Secretion System in Sliding Motility and Biofilm Formation. *Front*
947 *Microbiol* 2018;9:1160.

948 66. **Dumas E, Christina Boritsch E, Vandenbogaert M, Rodríguez de la Vega**
949 **RC, Thiberge J-M, *et al.*** Mycobacterial Pan-Genome Analysis Suggests Important
950 Role of Plasmids in the Radiation of Type VII Secretion Systems. *Genome Biol Evol*
951 2016;8:387–402.
952 67. **Zhou Z, Xu L, Zhu L, Liu Y, Shuai X, *et al.*** Metagenomic analysis of
953 microbiota and antibiotic resistome in household activated carbon drinking water
954 purifiers. *Environ Int* 2021;148:106394.
955 68. **Harms A, Brodersen DE, Mitarai N, Gerdes K**. Toxins, Targets, and Triggers:
956 An Overview of Toxin-Antitoxin Biology. *Mol Cell* 2018;70:768–784.
957 69. **Namkoong H, Kurashima A, Morimoto K, Hoshino Y, Hasegawa N, *et al.***
958 Epidemiology of Pulmonary Nontuberculous Mycobacterial Disease, Japan. *Emerg*
959 *Infect Dis* 2016;22:1116.
960 70. **Hoefsloot W, Ingen J van, Andrejak C, Ängeby K, Bauriaud R, *et al.*** The
961 geographic diversity of nontuberculous mycobacteria isolated from pulmonary
962 samples: an NTM-NET collaborative study. *Eur Respir J* 2013;42:1604–1613.
963 71. **Stinear TP, Seemann T, Harrison PF, Jenkin GA, Davies JK, *et al.*** Insights
964 from the complete genome sequence of Mycobacterium marinum on the evolution
965 of Mycobacterium tuberculosis. *Genome Res* 2008;18:729–741.
966 72. **Dedrick RM, Aull HG, Jacobs-Sera D, Garlena RA, Russell DA, *et al.*** The
967 Prophage and Plasmid Mobilome as a Likely Driver of Mycobacterium abscessus
968 Diversity. *mBio* 2021;12:e03441-20.
969 73. **Andries K, Villellas C, Coeck N, Thys K, Gevers T, *et al.*** Acquired Resistance
970 of Mycobacterium tuberculosis to Bedaquiline. *PLOS ONE* 2014;9:e102135.

971 74. **Hartkoorn RC, Uplekar S, Cole ST**. Cross-Resistance between Clofazimine
972 and Bedaquiline through Upregulation of MmpL5 in Mycobacterium tuberculosis.
973 *Antimicrob Agents Chemother* 2014;58:2979–2981.
974 75. **Alexander DC, Vasireddy R, Vasireddy S, Philley JV, Brown-Elliott BA, *et***
975 ***al.*** Emergence of mmpT5 Variants during Bedaquiline Treatment of Mycobacterium
976 intracellulare Lung Disease. *J Clin Microbiol* 2017;55:574–584.
977 76. **Halloum I, Viljoen A, Khanna V, Craig D, Bouchier C, *et al.*** Resistance to
978 Thiacetazone Derivatives Active against Mycobacterium abscessus Involves
979 Mutations in the MmpL5 Transcriptional Repressor MAB_4384. *Antimicrob Agents*
980 *Chemother* 2017;61:e02509-16.
981 77. **Sonika S, Singh S, Mishra S, Verma S**. Toxin-antitoxin systems in bacterial
982 pathogenesis. *Heliyon* 2023;9:e14220.
983 78. **Harms A, Maisonneuve E, Gerdes K**. Mechanisms of bacterial persistence
984 during stress and antibiotic exposure. *Science* 2016;354:aaf4268.
985 79. **Castañeda-Barba S, Top EM, Stalder T**. Plasmids, a molecular cornerstone of
986 antimicrobial resistance in the One Health era. *Nat Rev Microbiol* 2024;22:18–32.
987 80. **Kiga K, Tan X-E, Ibarra-Chávez R, Watanabe S, Aiba Y, *et al.*** Development
988 of CRISPR-Cas13a-based antimicrobials capable of sequence-specific killing of target
989 bacteria. *Nat Commun* 2020;11:2934.
990