Metabolic phenotype of clinical and environmental *Mycobacterium avium* subsp. *hominissuis* isolates

Andrea Sanchini [1], **Flavia Dematheis** [2], **Torsten Semmler** [3], **Astrid Lewin** Corresp. [1]

[1] Division 16, Mycotic and Parasitic Agents and Mycobacteria, Robert Koch Institute, Berlin, Germany

[2] Institute of Microbiology and Epizootics, Free University Berlin, Berlin, Germany

[3] NG 1 Microbial Genomics, Robert Koch Institute, Berlin, Germany

Corresponding Author: Astrid Lewin
Email address: LewinA@rki.de

Background. *Mycobacterium avium* subsp. *hominissuis* (MAH) is an emerging opportunistic human pathogen. It can cause pulmonary infections, lymphadenitis and disseminated infections in immuno-compromised patients. In addition, MAH is widespread in the environment, since it has been isolated from water, soil or dust. In the recent years, knowledge on MAH at molecular level substantially increased. On the contrary, knowledge of the MAH metabolic phenotypes remains limited.

Methods. In this study for the first time we analyzed the metabolic substrate utilization of ten MAH isolates, five from clinical and five from environmental source. We used the BIOLOG Phenotype Microarray™ technology. This technology permits the rapid and global analysis of metabolic phenotypes.

Results. The ten MAH isolates tested showed different metabolic patterns pointing to high intra-species diversity. Our MAH isolates preferred to use fatty acids such as Tween, caproic, butyric and propionic acid as a carbon source, and L-cysteine as a nitrogen source. Environmental MAH isolates resulted to be more metabolically active than clinical isolates, since the former metabolized more strongly butyric acid (p = 0.0209) and propionic acid (p = 0.00307).

Discussion. Our study provides new insight into the metabolism of MAH. Understanding how bacteria utilize substrates during infection might help the developing of strategies to fight such infections.

Article Title

Metabolic phenotype of clinical and environmental *Mycobacterium avium* subsp. *hominissuis* isolates

Authors

Andrea Sanchini[1], Flavia Dematheis[2], Torsten Semmler[3], Astrid Lewin[1,*]

Author affiliations

[1] Division 16, Mycotic and Parasitic Agents and Mycobacteria, Robert Koch Institute, Berlin, Germany

[2] Institute of Microbiology and Epizootics, Free University Berlin, Berlin, Germany

[3] NG 1 Microbial Genomics, Robert Koch Institute, Berlin, Germany

Corresponding author

Dr Astrid Lewin

Division 16, Mycotic and Parasitic Agents and Mycobacteria, Robert Koch Institute, Seestraße 10, 13353, Berlin, Germany.

Telephone: +4930187542112

E-mail: LewinA@rki.de

Abstract

Background. *Mycobacterium avium* subsp. *hominissuis* (MAH) is an emerging opportunistic human pathogen. It can cause pulmonary infections, lymphadenitis and disseminated infections in immuno-compromised patients. In addition, MAH is widespread in the environment, since it has been isolated from water, soil or dust. In the recent years, knowledge on MAH at molecular level substantially increased. On the contrary, knowledge of the MAH metabolic phenotypes remains limited.

Methods. In this study for the first time we analyzed the metabolic substrate utilization of ten MAH isolates, five from clinical and five from environmental source. We used the BIOLOG Phenotype Microarray™ technology. This technology permits the rapid and global analysis of metabolic phenotypes.

Results. The ten MAH isolates tested showed different metabolic patterns pointing to high intra-species diversity. Our MAH isolates preferred to use fatty acids such as Tween, caproic, butyric and propionic acid as a carbon source, and L-cysteine as a nitrogen source. Environmental MAH isolates resulted to be more metabolically active than clinical isolates, since the former metabolized more strongly butyric acid ($p = 0.0209$) and propionic acid ($p = 0.00307$).

Discussion. Our study provides new insight into the metabolism of MAH. Understanding how bacteria utilize substrates during infection might help the developing of strategies to fight such infections.

40

41

Introduction

Mycobacterium avium subsp. *hominissuis* (MAH) is clinically one of the most relevant non-tuberculous mycobacteria (Tortoli 2014). MAH is an opportunistic human pathogen causing pulmonary infections, lymphadenitis in small children and disseminated infections (Despierres et al. 2012; Rindi & Garzelli 2014). It is of increasing public health relevance, with reports of MAH infections increasing worldwide (Hoefsloot et al. 2013). Moreover, MAH is widespread in the environment (Falkinham 2013; Lahiri et al. 2014). In recent years, there have been substantial advances in the analysis of bacteria at the molecular level. Indeed, several whole genome sequences are now available for many mycobacterial species, including MAH (Bannantine et al. 2014; Kim et al. 2012; Uchiya et al. 2013; Wynne et al. 2010). In contrast, there has been little concomitant advance in knowledge at the phenotypic level. Phenotype analysis deserves greater attention, as it is the phenotype that selection pressure acts upon to confer evolutionary advantages to the bacterial species (Plata et al. 2015). In order to address this knowledge gap for bacterial phenotypes, BIOLOG Inc. developed the Phenotype MicroArray™ (PM) (BIOLOG, Hayward CA), a high throughput method for the rapid and global analysis of microbial metabolic phenotypes (Bochner 2003; Bochner 2009; Bochner et al. 2001; Bochner et al. 2008). The PM technology consists of several commercially available 96-well plates in which every well has a different substrate, allowing nearly 2000 different microbial metabolic phenotypes to be tested (Bochner 2003; Bochner 2009; Bochner et al. 2001; Bochner et al. 2008). PM technology has been applied to several microorganisms, including mycobacteria (Baloni et al. 2014; Bochner et al. 2008; Borglin et al. 2012; Chen et al. 2012; Gupta et al. 2015; Johnson et al. 2008; Khatri et al. 2013; Lofthouse et al. 2013; Mackie et al. 2014; Mishra & Daniels 2013; Nai et al. 2013; Omsland et al. 2009; Tohsato & Mori 2008). One possible

application of PM is the detection of phenotype changes due to gene knock-out. For example, Chen and co-authors showed that a *leuD* mutant of *M. avium* subsp *paratuberculosis* lost the ability to use several carbon, nitrogen, sulfur and phosphorous substrates (Chen et al. 2012). Other researchers showed that the use of 12 carbon substrates differentiated *M. tuberculosis* from *M. bovis* (Khatri et al. 2013; Lofthouse et al. 2013).

To date, there have been no reports of PM data for the emerging human pathogen MAH. In this study we tested clinical and environmental isolates of MAH using the PM technology. Our aim was to investigate if the PM technology was applicable to MAH isolates, to describe the metabolic substrates utilized by MAH isolates and to identify any metabolic differences between clinical and environmental MAH isolates.

Materials and Methods

Bacterial isolates and BIOLOG phenotype microarray

We analyzed five clinical and five environmental MAH isolates (Table 1).

We performed the BIOLOG Phenotype Microarray ™ (BIOLOG, Hayward, CA) according to the manufacturer´s recommendations. The technology is based on the measurement of bacterial respiration, which produces NADH (Bochner et al. 2001). If bacteria are able to metabolize a specific substrate, electrons from NADH reduce a tetrazolium dye in an irreversible reaction generating a purple color in the PM plate wells. This color change is measured and recorded every 15 minutes by the reporter instrument OmniLog™ (BIOLOG, Hayward, CA), generating a kinetic response curve for each well (Bochner 2003; Bochner 2009).

The ten MAH isolates were tested with the 96-wells plates PM1 to PM4, containing 190 carbon (PM1 and PM2), 95 nitrogen (PM3), 59 phosphorous (PM4) and 35 sulfur (PM4) substrates. The PM plates 1, 2 and 3 include one negative control well, in which bacteria are tested without any substrate. The PM4 plate includes two negative control wells, one for the phosphorus and one for the sulfur substrates. All isolates were tested three times. Briefly, we cultivated each MAH isolate in 30 ml of 7H10 Middlebrook medium supplemented with 10% modified ADC-enrichment (2% of glucose, 5% of BSA, 0,85% of NaCl) until an $OD_{600\ nm}$ of 0.3-0.6 was achieved (mid-logarithmic phase of growth). The use of liquid cultures in place of agar reduces bacterial clumping. Bacterial cultures were harvested by centrifugation for 10 minutes at 4000g and pellets were re-suspended in 10 ml of distilled water. Bacterial cells were starved for one night in water at room temperature to minimize false positive reactions due to nutrient accumulation in MAH cells and to ensure the use of the substrates provided by the PM plates.

98 The following day the cells were centrifuged and re-suspended using a sterile stick in tubes
99 containing 10 ml of GN/GP-IF-0a (BIOLOG inoculating fluid), 120 µl of 100× BIOLOG Redox
100 Dye Mix G and 1 ml of the appropriate additive (Table 2), until 85% transmittance was reached
101 as measured using the turbidimeter provided by BIOLOG. In order to reduce bacterial clumping,
102 the sterile stick used for inoculation was ground against the wall of the tube. A volume of 100 µl
103 of this final suspension was added to each of the 96 wells of the PM plates. The PM plates were
104 then sealed to avoid drying and incubated at 37°C in the OmniLog® (BIOLOG, Hayward, CA)
105 incubator reader for 8 days.

106 As recommended by BIOLOG, we tested plates PM1 to PM4 using the same assay protocol but
107 without addition of bacteria in order to identify wells with abiotic dye reduction, which can
108 generate false positive results.

109 **Analysis of BIOLOG phenotype microarray data**

110 The raw kinetic data were exported as CSV files using OmniLog PM file Management/kinetic
111 Analysis module (Bochner 2003; Khatri et al. 2013). Differences in the metabolization of the
112 different substrates by the ten MAH isolates were investigated by analyzing the maximum height
113 of the bacterial respiration curves (parameter A) using the R-package opm (Vaas et al. 2013). To
114 allow comparisons across plates processed in different experimental runs, the A parameters were
115 normalized by subtracting the well mean of the negative control (Vaas et al. 2013). Furthermore,
116 the A parameters of the triplicates were combined by calculating the mean and discretized into
117 “positive”, “moderate” and “negative” metabolization using the method “discrete” within the R-
118 opm package. Substrates differentiating the isolates from each other were visualized as a
119 heatmap generated using the R-packages heatmap.plus with the Euclidean algorithm. The

heatmap displays the utilization of each substrate with a color key: yellow for strong positive metabolization, green for moderate metabolization and blue for no metabolization.

Analysis of metabolic pathways

The metabolic pathways of butyric and propionic acids were identified from the KEGG pathway database of the reference strain *M. avium* 104 (Kanehisa et al. 2016). All genes of the pathways were screened in the genomes of the ten MAH using Geneious version 8 (Kearse et al. 2012).

Statistical analyses

We generated two groups, one with data from all clinical isolates and the other with data from all environmental isolates. Statistical differences between clinical and environmental isolates in the metabolization of butyric acid and propionic acid were evaluated by means of 95% family-wise comparison of group means (Tukey contrast test) of the parameter A on specific wells using the function "opm_mcp" within the opm R-package. A p value less than 0.05 was considered to be statistically significant.

Whole genome sequencing of MAH isolates

Genomic DNAs were extracted from the MAH isolates as described previously (Lewin et al. 2003). Whole genome sequencing (WGS) was performed using Illumina MiSeq 300 bp paired-end sequencing, yielding a coverage that exceeded 100x. The NGS QC tool kit was used to assess the quality of the data reads, which was set as reads with a minimum of 70 % of bases having a phred score greater than 20 (Patel & Jain 2012). De novo assembly of the resulting reads into multiple contigs was performed using CLC Genomics Workbench 8.0 (CLC bio, Aarhus, Denmark) and contigs annotation was done using RAST (Aziz et al. 2008).

141 **Determination of the maximum common genome and clustering analysis**

142 We determined the maximum common genome (MCG), comprising those genes present in all of
143 the ten MAH genomes, as reported previously (von Mentzer et al. 2014). All these genes were
144 then extracted from all genomes, concatenated and aligned. The resulting alignment was used to
145 generate a clustering tree using RAxML 8.1 (Stamatakis 2014).

Results

Substrate utilization of the ten MAH isolates

We tested the capability of our ten MAH isolates to metabolize 379 different substrates. In total, 334/379 (88.1 %) substrates were negative for all of the isolates (see Supplemental Table S1). A total of 23/379 (6.1 %) substrates caused abiotic reactions and were excluded from further analysis. A list of false-positive substrates is shown in the Supplemental Table S2. The kinetic curves corresponding to the control plates PM1 to PM4 tested without bacteria are presented in the Supplemental Fig. S1

Only two carbon substrates, the fatty acid derivatives Tween 20 and Tween 40 were strongly positive for all of the ten MAH isolates. The kinetic curves for these substrates reached 250 Omnilog units, amongst the highest values recorded in our analysis (see Supplemental Fig. S2 for all kinetic curves of the ten MAH isolates). The opm analysis revealed that a total of 20/379 (5.3 %) substrates were metabolized differently among the MAH isolates (Table 3). We therefore carried out further analysis using only these substrates. The majority of these 20 substrates were carbon substrates, 15/20 (75.0 %), followed by 3 nitrogen and 2 phosphorous substrates. The heatmap in Fig. 1 shows the utilization of these 20 substrates among the ten MAH isolates. The isolates are grouped according to their substrate utilization. Isolates utilizing similar substrates appear to cluster together.

Two major clusters, each composed of five isolates, could be observed. One was rich in environmental isolates (4/5) and the other was rich in clinical isolates (4/5). The substrates predominantly contributing to this clustering were butyric acid and propionic acid and indeed,

the Tukey's test revealed that environmental isolates metabolized more strongly butyric acid (p = 0.0209) and propionic acid (p = 0.00307) than clinical isolates with statistical significance.

Regarding the analysis of the metabolic pathways of these two substrates, in the reference strain MAH 104, 147 genes are known to be involved in the butyric acid and propionic acid pathways (Kanehisa et al. 2016). Based the WGS of the ten isolates we analysed the distribution of these genes in our ten isolates. However, no difference was observed in gene distribution among clinical and environmental MAH isolates (data not shown), pointing to differences in gene regulation and not in gene content.

Clustering analysis by whole genome sequencing

The WGS of the two reference strains MAH 104 and MAH TH135 were already in the GenBank database and we submitted the remaining genomes at DDBJ/EMBL/GenBank under the BioProject Number PRJNA299461. The MCG, the maximum number of genes shared by all ten MAH isolates was 1,658, the alignment of which spanned 1.378 Mbp. The clustering analysis of the ten MAH isolates is shown in Fig. 2. By comparing the genetic clustering obtained by WGS with the phenotypic clustering obtained through BIOLOG PM we observed slight differences. For examples, the isolates MAH E-96-2 and MAH E-82-7, which share identical metabolic profiles, were genetically more distant from each other. Interestingly, at the genetic level there was no obvious clustering between the group of clinical and the group of environmental isolates.

Discussion

This study represents the first phenotypic analysis of a collection of clinical and environmental MAH isolates using the Biolog PM technology. We showed that the PM technology works well and can be performed with MAH isolates. Strong positive reactions with several substrates were observed with kinetic curves exceeding 200 Omnilog dye units. Although some substrates were metabolized only moderately by our MAH isolates (green in Fig. 1), this might be due to the fact that the use of such substrates by bacteria has a time lag.

The ten MAH isolates showed different metabolic patterns pointing to high intra-species diversity. Only two out of the ten isolates had identical heatmap profiles (MAH E-96-2 and MAH E-82-7).

Our study showed that MAH isolates prefer to metabolize fatty acids as a carbon source. Indeed, the Tween substrates were strongly metabolized by all MAH isolates tested. This is in agreement with prior studies, showing that Tween substrates were widely used by different mycobacterial species (Baloni et al. 2014; Chen et al. 2012; Hayashi et al. 2010; Khatri et al. 2013; Lofthouse et al. 2013; Wang et al. 2011). It has been reported that mycobacteria hydrolyze Tween 80 to generate the fatty acid oleic acid, which can enter the Tricarboxylic acid (TCA) cycle or can be used as a substrate for energy production (Lofthouse et al. 2013; Vandal et al. 2009). Other fatty acids used by the majority of our MAH isolates are represented by two short fatty acids, caproic acid and butyric acid (Kanehisa & Goto 2000; Kanehisa et al. 2016; Khatri et al. 2013). Caproic acid and its derivatives are involved in several mycobacterial pathways such as the degradation of aromatic compounds, oxocarboxylic acid metabolism or lysine degradation (Kanehisa & Goto 2000; Kanehisa et al. 2016). The butyric acid is the final product of butanoate metabolism.

Propionic acid is another fatty acid used by our MAH isolates and this represents the terminal product of propanoate metabolism (Kanehisa & Goto 2000; Kanehisa et al. 2016). The nitrogen source L-cysteine, used by six of our MAH isolates, is the final product of cysteine metabolism and is involved in the biosynthesis of other amino acids such as methionine and histidine (Baloni et al. 2014; Kanehisa & Goto 2000; Kanehisa et al. 2016).

The question of whether bacteria of the same species originating from either clinical or environmental sources differ from each other is still a matter of discussion. Li and co-authors showed that comparative genome analysis clearly distinguished clinical and environmental *Vibrio parahaemolyticus* isolates from each other (Li et al. 2014). In contrast, other researchers have reported no difference between clinical and environmental *Pseudomonas aeruginosa* isolates with regard to virulence and metabolic properties (Alonso et al. 1999; Vives-Florez & Garnica 2006). Although our study did not reveal any clear distinction between clinical or environmental MAH isolates at the level of the whole genome, we observed differences between clinical and environmental isolates with regard to substrate utilization. The most intriguing difference is that the two fatty acids butyric acid and propionic acid are metabolized more by the environmental than by clinical isolates. Such a higher metabolic activity might be advantageous for survival in an environment presenting a wider range of nutritional conditions than the host cells alone. Further studies testing a larger number of isolates from different origins might clarify this. In addition, it has been showed that in bacteria the fatty acids have a role in adaptation to different environmental conditions (de Sarrau et al. 2012; de Sarrau et al. 2013; Diomande et al. 2015).

Conclusions

Our study contributes to the understanding of the emerging pathogen MAH at the phenotypic and metabolic level. Understanding how bacteria utilize their own or host-derived substrates during infection might help the development of strategies to fight such infections. We encourage phenotypic testing of microbial isolates from different ecological niches to identify key substrates or pathways that can be used as targets for drug development or for selective growth media development.

Acknowledgments

We would like to thank Barry Bochner (President of BIOLOG) for his invaluable input to this study and Brian Weinrick (Albert Einstein College of Medicine, New York City) for his support in developing the BIOLOG laboratory protocol. Elvira Richter (National Reference Center for Mycobacteria, Borstel, Germany) and Roland Schulze-Röbbecke (University Hospital Düsseldorf) provided a number of MAH isolates, Carsten Schwarz (Christiane Herzog Zentrum, Charité, Berlin) provided respiratory samples from cystic fibrosis patients and Kei-ichi Uchiya provided the reference strain MAH TH135. We thank Katharina Schaufler (Free University Berlin) for her support with the BIOLOG data analysis and Inga Eichorn (Free University Berlin) for her help with the whole genome sequencing data. We thank Steve Norley (Robert Koch Institute, Berlin) for the English revision of the manuscript.

References

Alonso A, Rojo F, and Martinez JL. 1999. Environmental and clinical isolates of Pseudomonas aeruginosa show pathogenic and biodegradative properties irrespective of their origin. *Environ Microbiol* 1:421-430.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, and Zagnitko O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. 10.1186/1471-2164-9-75

Baloni P, Padiadpu J, Singh A, Gupta KR, and Chandra N. 2014. Identifying feasible metabolic routes in Mycobacterium smegmatis and possible alterations under diverse nutrient conditions. *BMC Microbiol* 14:276. 10.1186/s12866-014-0276-5

Bannantine JP, Bayles DO, Robbe-Austerman S, Burrell AM, and Stabel JR. 2014. Draft Genome Sequence of a Mycobacterium avium Complex Isolate from a Broadbill Bird. *Genome Announc* 2. 10.1128/genomeA.01268-13

Bochner BR. 2003. New technologies to assess genotype-phenotype relationships. *Nat Rev Genet* 4:309-314. 10.1038/nrg1046

Bochner BR. 2009. Global phenotypic characterization of bacteria. *FEMS Microbiol Rev* 33:191-205. 10.1111/j.1574-6976.2008.00149.x

Bochner BR, Gadzinski P, and Panomitros E. 2001. Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. *Genome Res* 11:1246-1255. 10.1101/gr.186501

Bochner BR, Giovannetti L, and Viti C. 2008. Important discoveries from analysing bacterial phenotypes. *Mol Microbiol* 70:274-280. 10.1111/j.1365-2958.2008.06383.x

Borglin S, Joyner D, DeAngelis KM, Khudyakov J, D'Haeseleer P, Joachimiak MP, and Hazen T. 2012. Application of phenotypic microarrays to environmental microbiology. *Curr Opin Biotechnol* 23:41-48. 10.1016/j.copbio.2011.12.006

Chen JW, Scaria J, and Chang YF. 2012. Phenotypic and transcriptomic response of auxotrophic Mycobacterium avium subsp. paratuberculosis leuD mutant under environmental stress. *PLoS One* 7:e37884. 10.1371/journal.pone.0037884

de Sarrau B, Clavel T, Clerte C, Carlin F, Ginies C, and Nguyen-The C. 2012. Influence of anaerobiosis and low temperature on Bacillus cereus growth, metabolism, and membrane properties. *Appl Environ Microbiol* 78:1715-1723. 10.1128/AEM.06410-11

de Sarrau B, Clavel T, Zwickel N, Despres J, Dupont S, Beney L, Tourdot-Marechal R, and Nguyen-The C. 2013. Unsaturated fatty acids from food and in the growth medium improve growth of Bacillus cereus under cold and anaerobic conditions. *Food Microbiol* 36:113-122. 10.1016/j.fm.2013.04.008

Despierres L, Cohen-Bacrie S, Richet H, and Drancourt M. 2012. Diversity of Mycobacterium avium subsp. hominissuis mycobacteria causing lymphadenitis, France. *Eur J Clin Microbiol Infect Dis* 31:1373-1379. 10.1007/s10096-011-1452-2

Diomande SE, Nguyen-The C, Guinebretiere MH, Broussolle V, and Brillard J. 2015. Role of fatty acids in Bacillus environmental adaptation. *Front Microbiol* 6:813. 10.3389/fmicb.2015.00813

Falkinham JO, 3rd. 2013. Ecology of nontuberculous mycobacteria--where do human infections come from? *Semin Respir Crit Care Med* 34:95-102. 10.1055/s-0033-1333568

295 Gupta KR, Kasetty S, and Chatterji D. 2015. Novel functions of (p)ppGpp and Cyclic di-GMP in
296 mycobacterial physiology revealed by phenotype microarray analysis of wild-type and
297 isogenic strains of Mycobacterium smegmatis. *Appl Environ Microbiol* 81:2571-2578.
298 10.1128/AEM.03999-14
299 Hayashi D, Takii T, Mukai T, Makino M, Yasuda E, Horita Y, Yamamoto R, Fujiwara A, Kanai
300 K, Kondo M, Kawarazaki A, Yano I, Yamamoto S, and Onozaki K. 2010. Biochemical
301 characteristics among Mycobacterium bovis BCG substrains. *FEMS Microbiol Lett*
302 306:103-109. 10.1111/j.1574-6968.2010.01947.x
303 Hoefsloot W, van Ingen J, Andrejak C, Angeby K, Bauriaud R, Bemer P, Beylis N, Boeree MJ,
304 Cacho J, Chihota V, Chimara E, Churchyard G, Cias R, Daza R, Daley CL, Dekhuijzen
305 PN, Domingo D, Drobniewski F, Esteban J, Fauville-Dufaux M, Folkvardsen DB,
306 Gibbons N, Gomez-Mampaso E, Gonzalez R, Hoffmann H, Hsueh PR, Indra A, Jagielski
307 T, Jamieson F, Jankovic M, Jong E, Keane J, Koh WJ, Lange B, Leao S, Macedo R,
308 Mannsaker T, Marras TK, Maugein J, Milburn HJ, Mlinko T, Morcillo N, Morimoto K,
309 Papaventsis D, Palenque E, Paez-Pena M, Piersimoni C, Polanova M, Rastogi N, Richter
310 E, Ruiz-Serrano MJ, Silva A, da Silva MP, Simsek H, van Soolingen D, Szabo N,
311 Thomson R, Tortola Fernandez T, Tortoli E, Totten SE, Tyrrell G, Vasankari T, Villar M,
312 Walkiewicz R, Winthrop KL, Wagner D, and Nontuberculous Mycobacteria Network
313 European Trials G. 2013. The geographic diversity of nontuberculous mycobacteria
314 isolated from pulmonary samples: an NTM-NET collaborative study. *Eur Respir J*
315 42:1604-1613. 10.1183/09031936.00149212
316 Johnson DA, Tetu SG, Phillippy K, Chen J, Ren Q, and Paulsen IT. 2008. High-throughput
317 phenotypic characterization of Pseudomonas aeruginosa membrane transport genes. *PLoS*
318 *Genet* 4:e1000211. 10.1371/journal.pgen.1000211
319 Kanehisa M, and Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic*
320 *Acids Res* 28:27-30.
321 Kanehisa M, Sato Y, Kawashima M, Furumichi M, and Tanabe M. 2016. KEGG as a reference
322 resource for gene and protein annotation. *Nucleic Acids Res* 44:D457-462.
323 10.1093/nar/gkv1070
324 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A,
325 Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, and Drummond A. 2012.
326 Geneious Basic: an integrated and extendable desktop software platform for the
327 organization and analysis of sequence data. *Bioinformatics* 28:1647-1649.
328 10.1093/bioinformatics/bts199
329 Khatri B, Fielder M, Jones G, Newell W, Abu-Oun M, and Wheeler PR. 2013. High throughput
330 phenotypic analysis of Mycobacterium tuberculosis and Mycobacterium bovis strains'
331 metabolism using biolog phenotype microarrays. *PLoS One* 8:e52673.
332 10.1371/journal.pone.0052673
333 Kim BJ, Choi BS, Lim JS, Choi IY, Lee JH, Chun J, Kook YH, and Kim BJ. 2012. Complete
334 genome sequence of Mycobacterium intracellulare strain ATCC 13950(T). *J Bacteriol*
335 194:2750. 10.1128/JB.00295-12
336 Lahiri A, Kneisel J, Kloster I, Kamal E, and Lewin A. 2014. Abundance of Mycobacterium
337 avium ssp. hominissuis in soil and dust in Germany - implications for the infection route.
338 *Lett Appl Microbiol* 59:65-70. 10.1111/lam.12243

Lewin A, Freytag B, Meister B, Sharbati-Tehrani S, Schafer H, and Appel B. 2003. Use of a quantitative TaqMan-PCR for the fast quantification of mycobacteria in broth culture, eukaryotic cell culture and tissue. *J Vet Med B Infect Dis Vet Public Health* 50:505-509.

Li L, Wong HC, Nong W, Cheung MK, Law PT, Kam KM, and Kwan HS. 2014. Comparative genomic analysis of clinical and environmental strains provides insight into the pathogenicity and evolution of Vibrio parahaemolyticus. *BMC Genomics* 15:1135. 10.1186/1471-2164-15-1135

Lofthouse EK, Wheeler PR, Beste DJ, Khatri BL, Wu H, Mendum TA, Kierzek AM, and McFadden J. 2013. Systems-based approaches to probing metabolic variation within the Mycobacterium tuberculosis complex. *PLoS One* 8:e75913. 10.1371/journal.pone.0075913

Mackie AM, Hassan KA, Paulsen IT, and Tetu SG. 2014. Biolog Phenotype Microarrays for phenotypic characterization of microbial cells. *Methods Mol Biol* 1096:123-130. 10.1007/978-1-62703-712-9_10

Mishra MN, and Daniels L. 2013. Characterization of the MSMEG_2631 gene (mmp) encoding a multidrug and toxic compound extrusion (MATE) family protein in Mycobacterium smegmatis and exploration of its polyspecific nature using biolog phenotype microarray. *J Bacteriol* 195:1610-1621. 10.1128/JB.01724-12

Nai C, Wong HY, Pannenbecker A, Broughton WJ, Benoit I, de Vries RP, Gueidan C, and Gorbushina AA. 2013. Nutritional physiology of a rock-inhabiting, model microcolonial fungus from an ancestral lineage of the Chaetothyriales (Ascomycetes). *Fungal Genet Biol* 56:54-66. 10.1016/j.fgb.2013.04.001

Nelson, D., Michael M. Cox (2004) *Lehninger Principles of Biochemistry fourth edition.* : W. H. Freeman.

Omsland A, Cockrell DC, Howe D, Fischer ER, Virtaneva K, Sturdevant DE, Porcella SF, and Heinzen RA. 2009. Host cell-free growth of the Q fever bacterium Coxiella burnetii. *Proc Natl Acad Sci U S A* 106:4430-4434. 10.1073/pnas.0812074106

Patel RK, and Jain M. 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* 7:e30619. 10.1371/journal.pone.0030619

Plata G, Henry CS, and Vitkup D. 2015. Long-term phenotypic evolution of bacteria. *Nature* 517:369-372. 10.1038/nature13827

Rindi L, and Garzelli C. 2014. Genetic diversity and phylogeny of Mycobacterium avium. *Infect Genet Evol* 21:375-383. 10.1016/j.meegid.2013.12.007

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312-1313. 10.1093/bioinformatics/btu033

Tohsato Y, and Mori H. 2008. Phenotype profiling of single gene deletion mutants of E. coli using Biolog technology. *Genome Inform* 21:42-52.

Tortoli E. 2014. Microbiological features and clinical relevance of new species of the genus Mycobacterium. *Clin Microbiol Rev* 27:727-752. 10.1128/CMR.00035-14

Uchiya K, Takahashi H, Yagi T, Moriyama M, Inagaki T, Ichikawa K, Nakagawa T, Nikai T, and Ogawa K. 2013. Comparative genome analysis of Mycobacterium avium revealed genetic diversity in strains that cause pulmonary and disseminated disease. *PLoS One* 8:e71831. 10.1371/journal.pone.0071831

Vaas LA, Sikorski J, Hofner B, Fiebig A, Buddruhs N, Klenk HP, and Goker M. 2013. opm: an R package for analysing OmniLog(R) phenotype microarray data. *Bioinformatics* 29:1823-1824. 10.1093/bioinformatics/btt291

Vandal OH, Nathan CF, and Ehrt S. 2009. Acid resistance in Mycobacterium tuberculosis. *J Bacteriol* 191:4714-4721. 10.1128/JB.00305-09

Vives-Florez M, and Garnica D. 2006. Comparison of virulence between clinical and environmental Pseudomonas aeruginosa isolates. *Int Microbiol* 9:247-252.

von Mentzer A, Connor TR, Wieler LH, Semmler T, Iguchi A, Thomson NR, Rasko DA, Joffre E, Corander J, Pickard D, Wiklund G, Svennerholm AM, Sjoling A, and Dougan G. 2014. Identification of enterotoxigenic Escherichia coli (ETEC) clades with long-term global distribution. *Nat Genet* 46:1321-1326. 10.1038/ng.3145

Wang C, Mahrous EA, Lee RE, Vestling MM, and Takayama K. 2011. Novel Polyoxyethylene-Containing Glycolipids Are Synthesized in Corynebacterium matruchotii and Mycobacterium smegmatis Cultured in the Presence of Tween 80. *J Lipids* 2011:676535. 10.1155/2011/676535

Wynne JW, Seemann T, Bulach DM, Coutts SA, Talaat AM, and Michalski WP. 2010. Resequencing the Mycobacterium avium subsp. paratuberculosis K10 genome: improved annotation and revised genome sequence. *J Bacteriol* 192:6319-6320. 10.1128/JB.00972-10

Table 1(on next page)

Characteristics of the ten MAH isolates analyzed in this study

1 **Table 1** Characteristics of the ten MAH isolates analyzed in this study.

MAH Isolate name	Year of isolation	Source	Provider or reference	Accession of whole genome sequence
P-10091-06	2006	Clinical - Child with lymphadenitis	NRC for Mycobacteria, Borstel, Germany	LNAV00000000
2721	2004	Clinical - Child with lymphadenitis	NRC for Mycobacteria, Borstel, Germany	AWXJ00000000
P-9-13	2013	Clinical - Adult pulmonary infection	Charité Hospital, Berlin, Germany	LNBB00000000
104	1983	Clinical - Adult pulmonary infection	Reference strain, USA	CP000479
TH135	2013	Clinical - Adult pulmonary infection	Reference strain, Japan	AP012555
E-128	2010	Environmental - Soil	Friedrich Löffler Institute, Jena, Germany	LVCS00000000
E-96-2	2010	Environmental - Soil	This study	LMVW00000000
E- 82-7	2010	Environmental - Dust	This study	LNAF00000000
27-1	2010	Environmental - Dust	This study	AWXK00000000
E-2514	na	Environmental - Water	University of Düsseldorf, Germany	LNBJ00000000

2 MAH: *Mycobacterium avium* subsp. *hominissuis*; NRC: National reference center; na: not available

3

Table 2(on next page)

Additives used for each PM plates

As additive are usually provided nutrient that are absent to the PM minimal media, but present in a standard MAH growth conditions. We used additives to make a complete minimal medium but omitted anything that could act as a source of the substrates of interest (for example, we did not include nitrate additives in the nitrogen source plates).

Table 2 Additives used for each PM plates. As additive are usually provided nutrient that are absent to the PM minimal media, but present in a standard MAH growth conditions. We used additives to make a complete minimal medium but omitted anything that could act as a source of the substrates of interest (for example, we did not include nitrate additives in the nitrogen source plates).

	Additive a	**Additive b**
PM plate usage	**PM1, PM2, PM4**	**PM3**
Ingredients	24mM MgCl2 12mM CaCl2 0,0012% ZnSO4 0,06% ferric ammonium citrate 1,2% NH4Cl 0,01% tween 80	24mM MgCl2 12mM CaCl2 0,0012% ZnSO4 0,01% tween 80

Table 3(on next page)

The 20 substrates differentiating the ten MAH isolates analyzed in this study.

1 **Table 3** The 20 substrates differentiating the ten MAH isolates analyzed in this study.

PM Plate	Substrate and well number	Pathway involved	Reference
PM1 Carbon	Acetic acid – C08	Pyruvate metabolism	(Baloni *et al*. 2014; Nai *et al*. 2013)
	Acetoacetic acid – G07	Pyruvate metabolism	(Baloni *et al*. 2014; Nai *et al*. 2013)
	Methyl pyruvate – G10	Pyruvate metabolism	(Baloni *et al*. 2014; Nai *et al*. 2013)
	Mono−methyl Succinate – G09	Tricarboxylic acid cycle	(Baloni *et al*. 2014; Nai *et al*. 2013)
	Propionic acid – F07	Propanoate metabolism, Nicotinate and nicotinamide metabolism, Degradation of aromatic compounds	(Baloni, *et al*. 2014; Kanehisa and Goto 2000; Kanehisa *et al*. 2016; Nai *et al*. 2013)
	D−psicose – H05	Glycolysis and branches	(Baloni *et al*. 2014; Nai *et al*. 2013)
	Pyruvic acid – H08	Pyruvate metabolism	(Baloni *et al*. 2014; Nai *et al*. 2013)
	Tween 80 – E05	Fatty acid metabolism	(Baloni *et al*. 2014; Nai *et al*. 2013)
PM2 Carbon	L−alaninamide – G02	Amino acid metabolism	(Nai *et al*. 2013)
	Butyric acid – D12	Butanoate metabolism	(Baloni, *et al*. 2014; Kanehisa and Goto 2000; Kanehisa *et al*. 2016; Nai *et al*. 2013)
	Caproic acid – E02	Carboxylic acid metabolism	(Nai *et al*. 2013)
	L−histidine – G06	Amino acid metabolism	(Nai *et al*. 2013)
	γ−hydroxy−butyric acid – E09	Succinate metabolism	(Breitkreuz *et al*. 2003; Nai *et al*. 2013)
	β−methyl−D−galactoside – C07	Galactose Metabolism	(Nai *et al*. 2013)
	Sebacic acid – F08	Carboxylic acid metabolism	(Nai *et al*. 2013)
PM3 Nitrogen	D,L−α−amino−caprylic acid – G10	Amino acid metabolism	(Baloni *et al*. 2014)
	L−cysteine – A11	Amino acid metabolism	(Baloni *et al*. 2014)
	D−galactosamine – E09	Amino-sugar pathway	(Baloni *et al*. 2014)
PM4 Phosphorous and sulphur	Carbamyl phosphate – B05	Urea cycle and Pyrimidine synthesis	(Nelson 2004)
	Sodium pyrophosphate – A03	Phosphoric acid synthesis	(Nelson 2004)

2

3

4

5

Figure 1(on next page)

Heatmap showing the 20 substrates that were differently metabolized by the ten MAH isolates analyzed in this study.

The color key scale for each substrate is based on dye reduction quantified by Omnilog units. A yellow color indicates strong positive substrate metabolization, a green color moderate metabolization and a blue color indicates no substrate metabolization. Regarding the MAH isolates, environmental isolates are marked in orange, while clinical isolates are marked in blue.

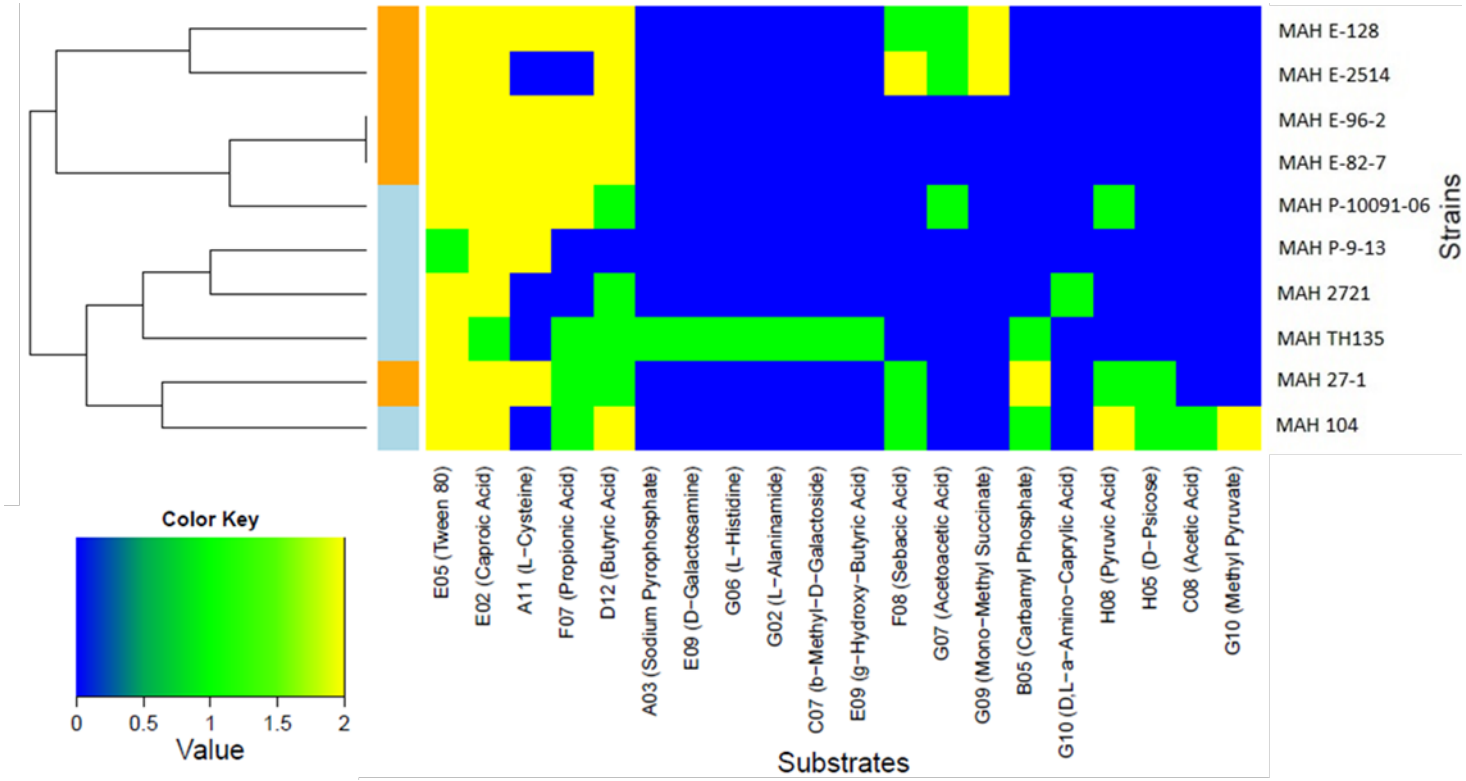
MAH E-128
MAH E-2514
MAH E-96-2
MAH E-82-7
MAH P-10091-06
MAH P-9-13
MAH 2721
MAH TH135
MAH 27-1
MAH 104
Strains
E05 (Tween 80)
E02 (Caproic Acid)
A11 (L−Cysteine)
F07 (Propionic Acid)
D12 (Butyric Acid)
A03 (Sodium Pyrophosphate)
E09 (D−Galactosamine)
G06 (L−Histidine)
G02 (L−Alaninamide)
C07 (b−Methyl−D−Galactoside)
E09 (g−Hydroxy−Butyric Acid)
F08 (Sebacic Acid)
G07 (Acetoacetic Acid)
G09 (Mono−Methyl Succinate)
B05 (Carbamyl Phosphate)
G10 (D,L−a−Amino−Caprylic Acid)
H08 (Pyruvic Acid)
H05 (D−Psicose)
C08 (Acetic Acid)
G10 (Methyl Pyruvate)
Substrates
Color Key
0
0.5
1
1.5
2
Value

Figure 1.

Figure 2(on next page)

Clustering of the 10 MAH isolates.

The tree was generated using RAxML 8.1. The alignment comprised 1,658 genes constituting the maximum common genome of our ten MAH isolates. Two reference strains were also included (MAH 104 and MAH TH135). The genome sequence of *M. avium* subsp. *paratuberculosis* K10 (Accession Number: AE016958) was used as outgroup. Isolate origin is also represented by blue for clinical origin and orange for environmental origin. The percentage of trees in which the associated taxa clustered together is shown adjacent to the branches.

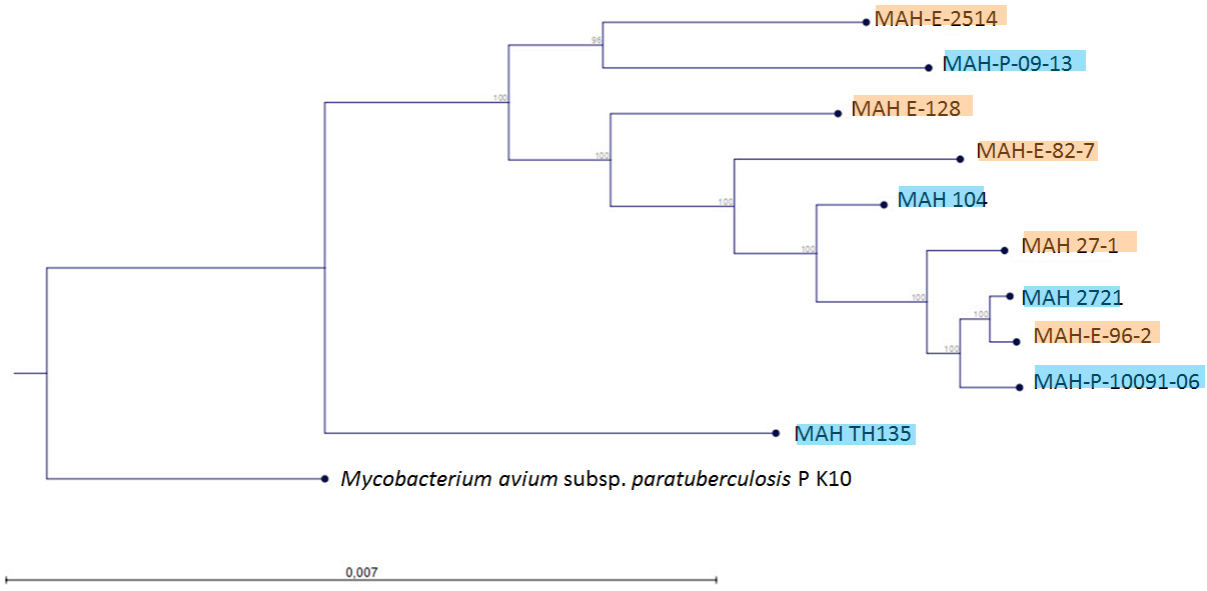
MAH-E-2514
MAH-P-09-13
MAH E-128
MAH-E-82-7
MAH 104
MAH 27-1
MAH 2721
MAH-E-96-2
MAH-P-10091-06
MAH TH135
Mycobacterium avium subsp. paratuberculosis P K10
0,007

Figure 2.