Functional genomics reveals strain-specific genetic requirements conferring hypoxic growth in *Mycobacterium intracellulare*

Yoshitaka Tateishi[1,2*], Yuriko Ozeki[1], Akihito Nishiyama[1], Yuta Morishige[3], Yusuke Minato[4,5], Anthony D. Baughn[5], Sohkichi Matsumoto[1,6,7,8]

1. Department of Bacteriology, Graduate School of Medicine and Dental Sciences, Niigata University, 1-757, Asahimachi-dori, Chuo-ku, Niigata 951-8510, Japan.
2. Department of Microbiology, Fukushima Medical University, 1, Hikariga-oka, Fukushima City, Fukushima, 960-1295, Japan.
3. Department of Mycobacterium Reference and Research, Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association, Kiyose, Tokyo, 2048533, Japan.
4. Department of Microbiology, Fujita Health University School of Medicine, Toyoake, Aichi, Japan.
5. Department of Microbiology and Immunology, University of Minnesota Medical School, 689 23rd Avenue S.E. Microbiology Research Facility, Minneapolis, 55455, MN, USA.
6. Laboratory of Tuberculosis, Institute of Tropical Disease, Universitas Airlangga,

Kampus C JI. Mulyorejo, Surabaya, 60113, Indonesia

7. Department of Bacteriology, Osaka Metropolitan University Graduate School of Medicine, 1-4-3 Asahi-machi, Abeno-ku, Osaka 545-8586, Japan

8. Division of Research Aids, Hokkaido University Institute for Vaccine Research & Development, Kita 20, Nishi 10, Kita-Ku, Sapporo, Hokkaido 001-0020, Japan

e-mail: y-tateishi@med.niigata-u.ac.jp, y-tate25@fmu.ac.jp

Abstract

Mycobacterium intracellulare is a major etiological agent of the recently expanding *Mycobacterium avium–intracellulare* complex pulmonary disease (MAC-PD). Therapeutic regimens that include a combination of macrolides and antituberculous drugs have been implemented with limited success. To identify novel targets for drug development that accommodate the genomic diversity of *M. avium–intracellulare*, we subjected eight clinical MAC-PD isolates and the type strain ATCC13950 to genome-wide profiling to comprehensively identify universally essential functions by transposon sequencing (TnSeq). Among these strains, we identified 131 shared essential or growth-defect-associated genes by TnSeq. Unlike the type strain, the clinical strains showed increased requirements for genes involved in gluconeogenesis and the type VII

secretion system under standard growth conditions, the same genes required for hypoxic pellicle-type biofilm formation in ATCC13950. Consistent with the central role of hypoxia in the evolution of *M. intracellulare*, the clinical MAC-PD strains showed more rapid adaptation to hypoxic growth than the type strain. Importantly, the increased requirements of hypoxic fitness genes were confirmed in a mouse lung infection model. These findings confirm the concordant genetic requirements under hypoxic conditions *in vitro* and hypoxia-related conditions *in vivo*, and highlight the importance of using clinical strains and host-relevant growth conditions to identify high-value targets for drug development.

Introduction

In contrast to the ongoing decline in the rate of tuberculosis, *Mycobacterium avium–intracellulare* complex pulmonary disease (MAC-PD) is increasing in many parts of the world[1,2]. The standard therapy for MAC-PD involves multidrug chemotherapy that includes clarithromycin and several antituberculous drugs. Unfortunately, treatment failure is common, resulting in relapse, disease progression, and death[3]. Therefore, it is critical to investigate the bacterial physiological characteristics that underlie the ability of MAC-PD strains to establish and maintain chronic drug-recalcitrant infections.

Recent advances in comparative genomics have revealed major differences in the genomic features of reference type strains and circulating pathogenic clinical isolates. Extensive genomic diversity has been recognized across clinical isolates of nontuberculous mycobacteria (NTM). *Mycobacterium avium* has been divided into four subspecies including subspecies, *avium*, *hominissuis*, *silvaticum*, and *paratuberculosis* [4]. Whereas *M. intracellulare* has been divided into two subgroups, typical *M. intracellulare* (TMI) and *M. paraintracellulare–M. indicus pranii* (MP-MIP), with currently two additional distinguished subspecies, *yongonense*, and *chimaera*[5,6]. Despite the increasing number of sequenced MAC-PD strains, the molecular genetic mechanisms that underlie their virulence and pathogenesis are still poorly understood.

The burgeoning field of functional genomics has allowed the rapid assessment of molecular aspects of organisms on a genome-wide scale. Transposon (Tn) sequencing (TnSeq) is a functional genomic approach that combines saturation-level transposon-insertion mutagenesis with next-generation sequencing (NGS) to comprehensively evaluate the fitness costs associated with gene inactivation in bacteria. Over the last decade, TnSeq has been applied to numerous bacterial species using various approaches[7]. Whereas initial studies applied TnSeq to reference strains, recent approaches have included the comparative assessment of representative clinical strains

to better understand the fundamental aspects of pathogen biology[8,9]. We have previously used TnSeq to globally characterize the genes that are essential for growth versus those that are essential for hypoxic pellicle formation in the *M. intracellulare* type strain ATCC13950[10]. However, given the recent expansion of MAC-PD and the well-recognized high levels of genomic diversity in this complex of bacteria, data obtained from the study of a decades-old reference strain is not expected to represent the diversity of extant clinical isolates. ATCC13950 was originally isolated from the abdominal lymph node of the 34-month-old female[11]. Even though the etiological bacterial species are the same, the clinical manifestation of infant lymphadenitis differs markedly from MAC-PD, which often occurs in middle-aged and elderly female patients with no predisposing immunological disorder. So far, there are a total of two reports on TnSeq that compare genetic requirements between clinical mycobacterial strains and the type strains in *M. tuberculosis*[8] and *M. abscessus*[9]. They reported the diversity of genetic requirements among strains including the type strain such as *M. tuberculosis* H37Rv and *M. abscessus* ATCC19977. Therefore, it is essential to investigate a collection of strains that captures the diversity of recent clinical MAC-PD strains in order to define key molecular aspects of their pathogenesis and virulence.

In this study, we used TnSeq to analyze a set of recently isolated clinical *M.*

intracellulare strains with diverse genotypes. We integrated the gene essentiality data for these strains and the type strain to identify the common essential and growth-defect-associated genes that represent the genomic diversity of this group of pathogens. We also identified the greater requirements of genes involved in gluconeogenesis, the type VII secretion system, and cysteine desulfurase in the clinical MAC-PD strains than in the type strain. Furthermore, the requirement for these genes was confirmed in a mouse lung infection model. The profiles of genetic requirements in clinical MAC-PD strains suggest the mechanism of hypoxic adaptation in NTM in terms of promising drug targets, especially for strains that cause clinical MAC-PD.

Results

Common essential and growth-defect-associated genes representing the genomic diversity of *M. intracellulare* strains

To obtain comprehensive information on the genome-wide gene essentiality of *M. intracellulare*, we included nine representative strains of *M. intracellulare* with diverse genotypes for TnSeq *in vitro* in this study (Fig. 1a, b, Supplementary Fig. 1, Supplementary Table 1). We used TRANSIT[12] to obtain 2–6 million reads of Tn insertions. The average number of Tn insertion reads was > 50 at Tn-inserted TA sites,

which confirmed the TnSeq data (Supplementary Table 2). The numbers of essential genes in the clinical MAC-PD strains, calculated with the hidden Markov model (HMM; a transition probability model), tended to be smaller than that in ATCC13950 (Supplementary Table 3). In total, 131 genes were identified as essential or growth-defect-associated with the HMM analysis across all *M. intracellulare* strains, including the eight MAC-PD clinical strains and ATCC13950 (Fig. 2a, Supplementary Table 4). The genes identified as universal essential or growth-defect-associated were involved in fundamental functions, such as glyoxylate metabolism, purine/pyrimidine metabolism, amino acid synthesis, lipid biosynthesis, arabinogalactan/peptidoglycan biosynthesis, porphyrin metabolism, DNA replication, transcription, amino acid tRNA ligases, ribosomal proteins, general secretion proteins, components of iron–sulfur cluster assembly, some ABC transporters, and the type VII secretion system. The genes identified corresponded to the genes that encode conventional and under-development drug targets, such as those encoding gyrase, arabinosyltransferase, enoyl-(acyl-carrier-protein) reductase, alanine racemase, DNA-directed RNA polymerase, malate synthase[13] and F_0F_1 ATP synthase (Fig. 2b). These data suggest that the set of universal essential or growth-defect-associated genes identified with the HMM analysis of *M. intracellulare* strains with diverse genotypes has potential utility

as antibiotic drug targets.

The sharing of strain-dependent and accessory essential and growth-defect-associated genes with genes required for hypoxic pellicle formation in the type strain ATCC13950

By comparative genomic analysis, we have revealed 3,153 core genes among 55 *M. intracellulare* strains including those used in this study[5] (Supplementary Fig. 1). Of these core genes, 439 genes were classified as strain-dependent essential or growth-defect-associated with the HMM analysis (Supplementary Table 5). And among 5,824 accessory genes, 222 genes were classified as essential or growth-defect-associated with the HMM analysis. The proportion of these strain-dependent and accessory essential or growth-defect-associated genes was comparable to the similar analysis of *Streptococcus pneumoniae*[14]. Of note, 17 (3.8%) of strain-dependent essential or growth-defect-associated genes and 14 (6.3%) of accessory essential or growth-defect-associated genes were also included in the 175 genes required for hypoxic pellicle formation in ATCC13950 reported in our previous study[10] (Supplementary Table 6). Gene set enrichment analysis (GSEA)[15] showed that 9.1% (16 out of 175) genes were hit as core enrichment (Supplementary Table 6). Of

them, 4 genes were hit commonly as genes showing increased genetic requirements analyzed by resampling plus HMM analyses including genes of phosphoenolpyruvate carboxykinase *pckA* (OCU_RS48660), type VII secretion-associated serine protease *mycP5* (OCU_RS38275), type VII secretion protein eccC5 (OCU_RS38345) and glycine cleavage system amino-methyltransferase *gcvT* (OCU_RS35955).

Partial overlap of the genes showing increased genetic requirements in clinical MAC-PD strains with those required for hypoxic pellicle formation in the type strain ATCC13950

The profiles of genetic requirements in each bacterial strain reflect the adaptation to the environment in which each strain lives. When the strains are placed in a special environment, they can adapt to the situation by altering the profiles of genetic requirements, resulting in the remodeling of metabolic pathways. The sharing of strain-dependent and accessory essential or growth-defect-associated genes with genes required for hypoxic pellicle formation in ATCC13950 prompts us to consider that the profiles of genetic requirements in clinical MAC-PD strains may be associated with the genes required for hypoxic pellicle formation in ATCC13950. To clarify this, we compared the genetic requirements of the clinical MAC-PD strains with that in

ATCC13950 by integrating the HMM analysis with the resampling analysis, a method of calculating the number of significantly altered Tn insertion reads (a gene-based permutation model). With this integrating analysis, 121 genes in the clinical MAC-PD strains were shown to have significantly fewer Tn insertions than ATCC13950 (Fig. 3, Supplementary Tables 7, 8). Of these, nine genes were identified in 3–5 clinical MAC-PD strains. These included genes encoding phosphatases (OCU_RS30185, OCU_RS35795), cysteine desulfurase (*csd* [OCU_RS40305]), methyl transferase (OCU_RS41540), a hypothetical protein (OCU_RS47290), type VII secretion components (*mycP5* [OCU_RS38275], *eccC5* [OCU_RS38345]), a glycine cleavage system aminomethyltransferase (*gcvT* [OCU_RS35955]), and phosphoenolpyruvate carboxykinase (*pckA* [OCU_RS48660]) (Figs 3, 4). The genes involved in gluconeogenesis (such as *pckA*), the glycine cleavage system, and the type VII secretion system were also identified as required for hypoxic pellicle formation in ATCC13950 in our previous study[9]. The Tn insertion reads were significantly reduced in the pathways of gluconeogenesis, including in fructose-1,6-bisphosphatase (*glpX*), phosphoenolpyruvate carboxykinase (*pckA*), and pyruvate carboxylase (*pca*), but were significantly increased in those of glycolysis, including in pyruvate kinase (*pykA*) and pyruvate dehydrogenase (*aceE*, *lpdC*), suggesting that carbohydrate metabolism was

metabolically remodeled in the clinical MAC-PD strains, which mimics the metabolism observed during hypoxic pellicle formation in ATCC13950 (Fig. 5).

Minor role of gene duplication on reduced genetic requirements in clinical MAC-PD strains

On the other hand, 162 genes were detected with significantly more Tn insertions in the clinical MAC-PD strains than in ATCC13950 (Fig. 3). Of these, 12 genes were identified in more than seven strains. The increased Tn insertion reads in the genes encoding superoxide dismutase (OCU_RS25945) and pyruvate dehydrogenase (*aceE* [OCU_RS35720]) in the clinical MAC-PD strains suggest the inessentiality of the detoxification of oxygen radicals and glycolysis metabolism in these strains (Figs 3–5). These genes also included those encoding amidohydrolases (OCU_RS34400, OCU_RS34425, OCU_RS34445), some transcription factors (OCU_RS28440, OCU_RS30230), a CoA transferase (OCU_RS34390), and some predicted transporters (OCU_RS28435, OCU_RS39950, OCU_RS39955, OCU_RS44685). Furthermore, the increase in Tn insertion reads in predicted multidrug-resistant genes, such as those encoding the DrrAB-ABC transporter complex (OCU_RS39950, OCU_RS39955)[16] and a putative lipoprotein LpqB (OCU_RS44685), encoded by the two-component system

mtrAB operon[17-18], in the clinical MAC-PD strains were found.

In order to investigate the effect of gene duplication on the change of genetic requirements between strains, we conducted blast search for the 162 genes showing increased Tn insertion reads in the clinical MAC-PD strains compared to ATCC13950. We found that M019 has duplicate genes of OCU_RS44705 coding adenosylhomocysteinase (LOCUS_42940: ahcY_1, LOCUS_21000: ahcY_2). However, there were no duplicate genes found in the remaining 161 genes. These data suggest that gene duplication has minor effects on the change of genetic requirements between strains. Rather, sequence differences and accessory genes may play a key role in determining the difference of genetic requirements.

Identification of genes in the clinical MAC-PD strains required for mouse lung infection

Hypoxia is the major environmental condition in natural water[19], biofilms[20], and infections in humans and animals characterized by caseous granuloma lesions[21]. The impact of hypoxia on mycobacteria under various ecological circumstances implies that the genes required for pathogenesis of MAC-PD may be in some degrees, overlapped with the genes with increased requirements in the clinical MAC-PD strains compared to

ATCC13950, and also with the genes required for hypoxic pellicle formation in ATCC13950. To identify genes required for *in vivo* infection of clinical MAC-PD strains, we performed TnSeq in mice infected intratracheally with the Tn mutant library strains of M.i.198 or M.i.27, which we had demonstrated to be virulent in mice (Fig. 1b, Supplementary Fig. 3)[22]. (It is impossible to perform TnSeq in mouse lungs infected with ATCC13950 because ATCC13950 is eliminated within 4 weeks of infection[22].) The time course of the changes in the bacterial burden showed a pattern similar to those of the wild-type strains M.i.198 and M.i.27, respectively, except that it was not possible to harvest sufficient colonies (as few as 10^4/mouse) in the few mice infected with the M.i.27 Tn mutant strain in Week 8 and Week 16 (Supplementary Fig. 3b, Supplementary Table 9). With a resampling analysis, 629 and 512 genes were shown to be required for *in vivo* infection with transposon mutant library strains M.i.27 and M.i.198, respectively, at least at one time point between Day 1 and Week 16 of infection (Fig. 6a, Supplementary Fig. 3c, d, Supplementary Tables 10, 11). Of these genes, 72 (11.9%) and 50 (9.76%) in M.i.27 and M.i.198, respectively, were also required for hypoxic pellicle formation in ATCC13950. Forty-one genes also required for hypoxic pellicle formation in ATCC13950 were identified in both M.i.27 and M.i.198, and included genes associated with succinate production (multifunctional 2-oxoglutarate

metabolism enzyme) and the glyoxylate cycle (isocitrate lyase), cysteine and methionine synthesis (thiosulfate sulfurtransferase [CysA], 5-methyltetrahydrofolate-homocysteine methyltransferase [MetH]), the serine/threonine-protein kinase PknG system, proteasome subunits (PrcA, PrcB), DNA repair UvrC, secretion protein EspR, cell-wall synthesis proteins (such as phosphatidylinositol [polyprenol-phosphate-mannose-dependent alpha-[1-2]-phosphatidylinositol pentamannoside mannosyltransferase], peptidoglycan (penicillin-binding protein A pbpA), and L,D-transpeptidase 2), putative membrane proteins and transporters (such as the MmpL family, transport accessory protein MmpS3, transport protein MmpL11, vitamin B_{12} transport ATP-binding protein BacA, ABC transporter substrate-binding protein, and ABC transporter permease), and mammalian cell entry (Mce) proteins (OCU_RS48855–48875) (Supplementary Table 12).

At all timepoints from Day1 to Week 16, there were 74 and 99 genes were shown to be required for lung infection with M.i.27 and M.i.198, respectively. Of them, 21 (28.3%) and 12 (12.1%) genes belonged to the genes involved in hypoxic pellicle formation in the type strain. At timepoints throughout Week 4 to Week 16 except for Day1, there were 121 and 172 genes were shown to be required for lung infection with M.i.27 and M.i.198, respectively. Of them, 21 (23.1%) and 30 (18.0%) genes belonged

to genes involved in hypoxic pellicle formation in the type strain. These hypoxic pellicle-associated genes detected both in M.i.27 and M.i.198 encoded methionine synthesis (MetH), acyl-CoA dehydrogenase, isocitrate lyase, MMPL family transporter (MmpL11) (from Day1 to Week 16), additionally multifunctional oxoglutarate decarboxylase/dehydrogenase, proteasome subunits (PrcA, PrcB), ABC transporter ATP-binding protein/permease, lipase chaperone (from Week 4 to Week 16) (new Supplementary Tables 13, 14).

The genes only identified in M.i.27 that were also required for hypoxic pellicle formation in ATCC13950 included genes associated with gluconeogenesis (*pckA*, *glpX*), the glycine cleavage system (*gcvT*), the type VII secretion system (*eccD5, espG5, eccC5* [OCU_RS38280-OCU_RS38285-OCU_RS38345], *eccC3* [OCU_RS48145]), and D-inositol 3-phosphate glycosyltransferase (Supplementary Table 12). These three systems, gluconeogenesis, the glycine cleavage system, and the type VII secretion system, also showed increased genetic requirements in 3–5 clinical MAC-PD strains than in ATCC13950 (Figs 3, 6). Although M.i.27 was excluded as a relevant strain showing increased genetic requirements of these three metabolic systems, the increase in genetic requirements for infection *in vivo* suggests a positive relationship between hypoxic growth *in vitro* and bacterial growth *in vivo*. In contrast, M.i.198 already

showed increased genetic requirements in these three metabolic systems under aerobic conditions *in vitro*, so the genes associated with these three systems were not identified as having fewer Tn insertion reads during infection *in vivo*. Genes only identified in M.i.198 and required for hypoxic pellicle formation in ATCC13950 included genes associated with succinate synthesis (2-oxoglutarate oxidoreductase subunit KorA) or encoding the HTH-type transcriptional repressor KstR and menaquinone reductase (Supplementary Table 12). As increased genetic requirements in 3–5 clinical strains (excluding M.i.198) compared with ATCC13950, an increased requirement for the RNA methyltransferase OCU_RS41540 was identified in M.i.198 during infection *in vivo*. These *in vivo* TnSeq data suggest that, despite the differences in the profiles of the genes required for *in vivo* infection between strains, increase of genetic requirements for hypoxic growth in part contribute to the pathogenesis *in vivo*.

We have performed a statistical enrichment analysis of gene sets by GSEA[15] to investigate whether the genes required for growth of M.i.27 and M.i.198 in mouse lungs overlap the gene sets required for hypoxic pellicle formation in ATCC13950 as well as the gene sets showing increased genetic requirements observed in the clinical MAC-PD strains listed in Fig 3. Of these gene sets including a total of 180 genes, 54% (98 genes) and 40% (73 genes) were listed as enriched in genes required for mouse infection in

M.i.27 and M.i.198, respectively ($P < 0.01$ in M.i.27 and $P = 0.016$ in M.i.198) (Supplementary Fig. 4, Supplementary Table 15). The enriched genes both in M.i.27 and M.i.198 included genes encoding fructose -1,6-bisphosphatase (*glpX*), a type VII secretion protein (*eccC5*), glycine cleavage system aminomethyltransferase (*gcvT*), some phosphatase (OCU_RS30185) and a hypothetical protein (OCU_RS47290) (Fig. 6). The enriched genes only in M.i.27 included genes of phosphoenolpyruvate carboxylinase (*pckA*), cysteine desulfurase (*csd*) and the ESX-5 type VII secretion component serine protease (*mycP5*). The enriched genes only in M.i.198 included genes encoding some phosphatase (OCU_RS35975) and class I SAM-dependent RNA methyltransferase (OCU_RS41540). Similarly, 40-50% of the genes required for ATCC13950's hypoxic pellicle formation and strain-dependent/accessory essential genes are significantly enriched individually with the genes required for *in vivo* bacterial growth (Supplementary Figs. 5, 6, Supplementary Tables 16, 17). These data suggest the sharing of the hypoxia-adaptation pathways between *in vitro* hypoxic pellicle formation and *in vivo* bacterial growth.

Effects of knockdown of universal essential or growth-defect-associate genes in clinical MAC-PD strains

To facilitate the functional analysis of the TnSeq-hit genes, it is necessary to construct the genetically-engineered strains. Still now, there are very few reports of gene engineering of MAC strains and none of them has been accompanied by the follow-up studies[23-28]. Moreover, no attempts have been reported to construct the knockdown strains in MAC strains. With an intention to evaluate the effect of suppressing TnSeq-hit genes on bacterial growth, we used the pRH2052/pRH2521 clustered regularly interspaced short palindromic repeats interference (CRISPR-i) plasmids to construct knockdown strains where the expression of the target gene is suppressed[29-31]. Of the nine *M. intracellulare* strains analyzed in this study, seven could be transformed with the vectors for single guide RNA (sgRNA) expression and dCas9 expression. However, the other two strains, M001 and MOTT64, could not be transformed with these CRISPR-i vectors. To confirm the gene essentiality detected with the HMM analysis, we evaluated the consequent growth inhibition in the knockdown strains of representative universal essential or growth-defect-associated genes, including *glcB*, *inhA*, *gyrB*, and *embB*. The knockdown strains showed growth suppression with comparative growth rates of knockdown strains to vector control strains that do not express single sgRNA reaching 10^{-1} to 10^{-4} (Fig. 7a).

Differential effects of knockdown of accessory/strain-dependent essential or growth-defect-associated genes among clinical MAC-PD strains

As representative accessory and strain-dependent essential or growth-defect-associated genes, we evaluated the 9 genes showing reduced Tn insertion reads in the 3–5 clinical MAC-PD strains as listed in Fig. 3, as well as *glpX*, a gene of gluconeogenesis where four MAC-PD strains showed reduced Tn insertion reads in a resampling analysis (Fig. 5b). The growth was suppressed in the *pckA*-knockdown strains in M.i.198 and M003, in the *glpX*-knockdown strain in M003, and in the cysteine desulfurase gene (*csd*)-knockdown strains in M.i.198 and M003, with comparative growth rates of knockdown strains to vector control strains reaching 10^{-2} to 10^{-4} (Fig. 7b). The growth tended to be suppressed in the type VII secretion protein gene (*eccC5*)-knockdown strain in M003 and in the type VII secretion-associated serine protease mycosin gene (*mycP5*)-knockdown strain in M003, with comparative growth rates of knockdown strains to vector control strains reaching less than 10^{-1}. By contrast, the growth was not suppressed in these knockdown strains in ATCC13950, with comparative growth rates of knockdown strains to vector control strains not reaching 10^{-1}, which was consistent to the inessentiality of these genes under aerobic conditions in ATCC13950. As for M.i.198 and M003, the consistency between TnSeq data and

CRISPR-i data was supported by a correlation between the growth suppression of the knockdown strains and the fitness costs evaluated by a resampling analysis (Supplementary Table 18).

In contrast to the cases of M.i.198 and M003, the growth of knockdown strains of the accessory and strain-dependent essential genes was not effectively suppressed in the other MAC-PD strains despite being positively detected in TnSeq (Fig. 7b, Supplementary Fig. 7). Although the gene expression levels were not suppressed to < 1% as reported in *M. tuberculosis* (Mtb)[32], the suppression of gene expression was confirmed to approximately 20-70% in the knockdown strains of these accessory and strain-dependent essential or growth-defect-associate genes, as well as universal essential or growth-defect-associated genes such as *glcB*, *inhA*, *gyrB*, and *embB* (Supplementary Fig. 8). The possible reason may be explained by the recently revealed bypass mechanism of gene essentiality in strain-dependent/accessory essential or growth-defect-associated genes, where gene essentiality can be bypassed by several mechanisms including the composition of the accessory genome and pathway rewiring[14].

To be summarized, we could validate the universal essential or growth-defect-associated genes by the CRISPR-i system in the *M. intracellulare* strains

used in this study. And, although not successful with all strains, we could validate the accessory and strain-dependent essential or growth-defect-associated genes, especially genes of gluconeogenesis, type VII secretion components and iron-sulfur cluster assembly in several MAC-PD strains like M.i.198 and M003.

Preferential hypoxic adaptation of clinical MAC-PD strains evaluated with bacterial growth kinetics

The metabolic remodeling, such as the increased genetic requirements of gluconeogenesis and the type VII secretion system, and the overlap of the genes required for mouse lung infection and those required for hypoxic pellicle formation involved by these metabolic pathways suggest the preferential adaptation of the clinical MAC-PD strains to hypoxic conditions. To confirm this, we investigated the growth kinetics of each strain under aerobic and hypoxic conditions. In terms of growth kinetics, most of the clinical MAC-PD strains (except M001) entered logarithmic (log) phase earlier than ATCC13950 under 5% oxygen (Fig. 8a, 8b, Supplementary Table 19). However, under aerobic conditions, the clinical MAC-PD strains entered log phase at a similar time to ATCC13950. The growth rate at mid-log point (midpoint) was significantly reduced under hypoxic conditions in a total of 5 clinical MAC-PD strains

(M.i.198, M.i.27, M003, M019 and M021) compared to aerobic conditions (Fig. 8a, 8b, Supplementary Table 19). The growth rate at midpoint under hypoxic conditions was significantly lower in these 5 clinical MAC-PD strains than in ATCC13950. The early entry to log phase followed by long-term logarithmic growth (slow growth rate at midpoint) suggests the capacity for these 5 clinical MAC-PD strains to continue replication long time under hypoxic conditions. On the other hand, the rest 3 clinical MAC-PD strains (M018, M001 and MOTT64) did not show significant change in the growth rate between aerobic and hypoxic conditions, suggesting that there are different levels of capacity in maintaining long-term replication under hypoxia among clinical MAC-PD strains. In ATCC13950, the entry to log phase was significantly delayed under 5% oxygen compared to aerobic conditions, and the growth rate at midpoint was significantly increased under hypoxic conditions compared to aerobic conditions in ATCC13950. Such long-term lag phase followed by short-term log phase suggests lower capacity for ATCC13950 to continue replication under hypoxic conditions compared to clinical MAC-PD strains.

The pattern of hypoxic adaptation not simply determined by genotypes

Since ATCC13950 and M018 are phylogenetically neighbors within the same TMI

group each other, there may be a claim that the growth phenotypes of a longer lag phase compared to other strains under hypoxic conditions might be determined simply by the genotypes. To examine this, we performed hypoxic growth assay for additional two strains (M005, M016) phylogenetically more closely related to ATCC13950 (Supplementary Fig. 1, Supplementary Table 1). The result showed different growth pattern under hypoxia; M005 reached midpoint later than ATCC13950, by contrast M016 reached midpoint three-quarters earlier than ATCC13950 (Supplementary Fig. 9). These data suggest that genotypes seem to have some impact on bacterial growth pattern under hypoxia; however, since there was a significant difference in the timing of hypoxic adaptation among ATCC13950 and its neighbor strains, bacterial growth pattern under hypoxia seems to be determined by multiple factors other than genotypes such as strain-specific profiles of genetic requirements and unproven regulatory systems.

In summary, the data of more rapid adaptation to hypoxia followed by reduced growth rate once adapted indicate the greater metabolic advantage on continuous logarithmic bacterial growth under hypoxic conditions in the clinical MAC-PD strains than in ATCC13950. Because the profiles of genetic requirements reflect the adaptation to the environment in which bacteria habits, it is reasonable to assume that the increase

of genetic requirements in hypoxia-related genes such as gluconeogenesis (*pckA*, *glpX*), type VII secretion system (*mycP5*, *eccC5*) and cysteine desulfurase (*csd*) play an important role on the growth under hypoxia-relevant conditions *in vivo*.

Discussion

In this study, using TnSeq of nine genetically diverse *M. intracellulare* strains, we have identified 131 shared essential and growth-defect-associated genes with an HMM analysis. The genes identified are promising drug targets, representing the genomic diversity of the clinical MAC-PD strains and reflecting the actual clinical situation. The strategy of narrowing the essential genes of bacteria is a useful approach for refining the database of drug targets. We have also demonstrated an overview of the differences in genetic requirements for hypoxic growth between the ATCC type strain and the clinical *M. intracellulare* strains responsible for MAC-PD, with the evidence of earlier entry to log phase followed by slower growth rate observed in most of these clinical strains compared to ATCC13950. This is the first study to demonstrate the positive relationship between the profiles of increased genetic requirements for hypoxic metabolic remodeling and the phenotypes advantageous for bacterial survival under hypoxic conditions.

In this study, 121 genes were detected as significantly reduced Tn insertion in the clinical MAC-PD strains compared to ATCC13950, which was higher than the case of Mtb showing 10–50 genes with altered Tn insertion reads detected by a resampling analysis[8]. This may reflect the genomic diversity of *M. intracellulare* compared with that of Mtb. The proportion of accessory genes is larger in NTM than in Mtb[33]. In a clinical trial of phage therapy for *M. abscessus* infection, the number of accessory genes was shown to decrease with the loss of virulence in *M. abscessus* during the course of therapy[34], suggesting some role for accessory genes for its clinical pathogenesis. Our data of the overlap of the accessory essential or growth-defect-associated genes such as *pckA* and *mycP5* (Supplementary Table 7) with genes required for mouse lung infection (Fig. 6b, Supplementary Fig. 3) supports the idea that accessory genes (non-core genes) may encode functions required for bacterial survival, especially under nonoptimal conditions, such as hypoxia and during the infection of human and animal bodies.

Our TnSeq data are consistent with metabolomic data that have shown the importance of the gluconeogenic carbon flow of TCA cycle intermediates for bacterial growth under hypoxic conditions in terms of the essentiality of PckA and the reverse TCA cycle[35–38]. Gluconeogenesis is essential for bacterial growth when fatty acids are used as a major nutrient[37]. PckA catalyzes the conversion of the metabolites derived

from the TCA cycle to the metabolites of gluconeogenesis[36]. The reaction catalyzed by PckA is essential for bacterial growth when the available nutrients are predominantly fatty acids[39,40]. The main nutrients for mycobacteria inside hypoxic granulomas in an infected human or mouse body are lipids, such as fatty acids and cholesterol, rather than carbohydrates[37,41]. In the present study, *M. intracellulare* grew under 5% oxygen, whereas Mtb stops growing under 1% oxygen and enters a dormant state[36]. The depletion of phosphoenolpyruvate in the *pckA* deletion mutant of Mtb reduced the phosphoenolpyruvate–carbon flux essential for bacterial growth and drug sensitivity[36]. The present study supports the idea that PckA is essential for the microaerobic growth of mycobacteria. Our TnSeq data imply the preferential adaptation of the clinical MAC-PD strains for circumstances *in vivo*, which are characterized by hypoxia.

On the other hand, such metabolic shift towards gluconeogenesis and reverse TCA cycle can also be observed under reactive nitrogen species and *in vivo* infection as well. This suggests that the metabolic shift observed under hypoxia is a general stress response rather than being specific to hypoxia due to the nature of the provided/available carbon source[42]. Several factors other than hypoxia can be estimated to be related to the change of the genetic requirements in the clinical MAC-PD strains such as host-related factors and general stress conditions (i.e. nutrition of abundant fatty

acids and less carbohydrates, low pH circumstances)[37,39-41]. Except for our data of the profile of genes required for hypoxic pellicle formation in *M. intracellulare*[10], no fundamental genome-wide data are available on the change of the genetic requirements in response to each stress factor in non-tuberculous mycobacteria. Further studies are needed to discriminate which factors are specifically involved in the increased requirements of each gene. Although of interest, it is beyond the scope of the current study.

As in the *glpX*-knockdown strains, *csd* knockdown caused growth suppression in several MAC-PD strains. This type II Csd is involved in the biosynthesis of the iron–sulfur clusters of the SUF machinery. Compared with the type I cysteine desulfurase IscS, the SUF machinery is active under oxidative-stress and iron-limiting conditions[50]. Representatives of iron–sulfur cluster proteins include components of the electron-transport system, such as NADH dehydrogenase I (Nuo) and succinate dehydrogenase (Sdh), which are components of respiratory complexes I and II, respectively. Under hypoxic conditions, the non-proton-translocating type II NADH dehydrogenase Ndh is upregulated, and cytochrome *bd* oxidase (CydAB) and the proton-pumping type I NADH dehydrogenase (Nuo) are downregulated[42]. In the reverse TCA cycle on the same reaction pathway as Sdh, fumarate reductase (Fnr) is also an

487 iron–sulfur cluster protein that produces succinate, and is a proposed alternative electron
488 acceptor under hypoxic conditions[38,51]. The increased essentiality of the *csd* gene
489 observed in the clinical MAC-PD strains seems to reflect the switching of the
490 electron-transport system to less-energy-efficient respiration to support preferential
491 hypoxic survival, including *in vivo*.

492 About 10% of the genes required for mouse infection are also required for hypoxic
493 pellicle formation. The genes showing increased genetic requirements in the clinical
494 MAC-PD strains, such as *pckA* and genes encoding the ESX-5 type VII secretion
495 system, including *eccC5* and *mycP5*, also showed increased genetic requirements in
496 mouse infections. In a TnSeq study of mice infected intravenously with *M. avium*, the
497 genes of the ESX-5 type VII secretion system were required for infection of the liver
498 and spleen[43]. The ESX-5 type VII secretion system is involved in the transport of fatty
499 acids and the virulence of mycobacteria[44,45]. These similarities in the profiles of genetic
500 requirements prompt us to propose some role of biofilm formation in mycobacterial
501 pathogenesis in animals and humans. In the present study, we shed light on the impact
502 of hypoxia on biofilm formation, the acquisition of virulence, and the *in vivo*
503 pathogenesis of NTM in terms of global genetic requirements. Although the
504 morphologies of the pellicle and macroscopic lesions (including cavitation and nodular

505 bronchiectasis) are not similar, monitoring the expression of the genes identified in this
506 study may provide clues to the mechanism of biofilm formation *in vivo*.
507 The data of validation of the universal essential or growth-defect associated genes
508 by the CRISPR-i system was overall acceptable in the clinically MAC-PD strains whose
509 knockdown strains could be constructed (Fig. 7a). However, we have encountered the
510 difficulties of validating the TnSeq-hit genes by CRISPR-i experiment, especially
511 accessory and strain-dependent essential or growth-defect-associated genes in the
512 MAC-PD strains other than M.i.198 or M003 (Fig. 7b, Supplementary Fig. 7). We found
513 the lower efficiency of the suppression of gene expression by CRISPR-i in *M.*
514 *intracellulre* than in Mtb that is reported to be suppressed to 6% under 10 ng/ml
515 anhydrotetracycline (aTc) and by < 1% under 100 ng/ml aTc in the knockdown strains,
516 compared to strains expressing scrambled single guide RNAs (sgRNA)[32]. However,
517 because the levels of knockdown efficiency were comparable between
518 strain-dependent/accessory essential genes and universally essential genes
519 (Supplementary Fig. 8), the low efficiency of knockdown does not seem to explain the
520 reason of the discrepancy between TnSeq and CRISPR-i results. Although it is difficult
521 to find the clear definitive reason for this phenomenon only from the current study, the
522 possible reason may be speculated by the bypass mechanism of gene essentiality which

is characteristically observed in strain-dependent/accessory essential or growth-defect-associated genes. According to the publication by Rosconi[14] reporting the 'forced-evolution experiments' of 36 clinical *Streptococcus pneumoniae* strains., gene essentiality can be bypassed by several mechanisms including the composition of the accessory genome and pathway rewiring, They recovered successfully knockout mutants from transformation experiments in strain-dependent/accessory essential genes such as cytidine monophosphate kinase, a folate pathway enzyme formate tetrahydrofolate ligase and an undecaprenyl phosphate-biosynthesis pathway enzyme farnesyl-diphosphate synthase. The bypassing of gene essentiality could be suggested by observing suppressor mutations and synthetic lethality in knockout strains. By contrast, universal essential genes were reported to fulfill the three categories including high levels of conservation within and often across species, limited genetic diversity, and high and stable expression levels. Consequently, universal essential genes were estimated to be rigid, largely immutable key components to an organism's survival. The knockout recovery of the universal essential genes was reported to fail as shown by no colonies or only appearance of merodiploids. Taking into consideration such bypass mechanism of gene essentiality, the inconsistent effect of gene silencing of strain-dependent/accessory essential genes on bacterial growth may reflect some kinds

of pathway rewiring that helps the bacteria grow under suppression of the target gene expression. Nevertheless, our study has opened the avenue for functional analysis of NTM that leads to the development of novel drugs targeting universal and accessory/strain-dependent essential metabolic pathways, especially by making use of M.i.198 and M003 as optimal NTM strains for gene manipulation.

In this study, we revealed the pattern of hypoxic adaptation characteristic to the clinical MAC-PD strains by the growth curve analysis: an early entry to log phase followed by a reduced growth rate (Fig. 8). This pattern implicates the continuous impact on the infected hosts during logarithmic bacterial growth, which may be involved in the persistent and steadily progressive illness for years in humans[3]. The advantage of slow growth in mycobacteria may be inferred by the role of a histone-like nucleoid protein, mycobacterial DNA-binding protein 1 (MDP1), also designated HupB[46-48]. The expression of MDP1 is enhanced in the stationary phase, resulting in reduced growth speed and long-term survival with increased resistance to reactive oxygen species (ROS)[46-48]. The reduced growth rate after entry to log phase may allow long-term survival of mycobacteria under ROS-rich circumstances inside host cells. The pattern of hypoxic adaptation as observed in the clinical MAC-PD strains may enhance pathogenesis by extending the period of the repeated process of bacterial invasion and

replication in infected macrophages.

Contrary to our expectation, not all clinical strains showed rapid adaptation from aerobic to hypoxic conditions relative to ATCC13950. The two strains belonging to the MP-MIP subgroup, MOTT64 and M001 showed similar time at midpoint under aerobic conditions. However, the time at midpoint was significantly different between MOTT64 and M001 under hypoxia, the latter showing great delay of timing of entry to logarithmic phase. In contrast to the majority of the clinical strains that showed slow growth at midpoint under hypoxia, neither strain showed such phenomenon. These data suggest that some, but not a major proportion of, clinical *M. intracellulare* strains do not have an advantageous phenotype for continuous logarithmic growth under hypoxia. In Mtb, the adaptation to hypoxia is considered to be necessary for clinical pathogenesis because the conditions *in vivo*, such as inside macrophages or granulomas, are hypoxic[42,49]. Our inability to construct knockdown strains in M001 and MOTT64 prevented us from clarifying the factors that discriminate against the pattern of hypoxic adaptation. Although the implication in clinical situations has not been proven, strains without slow growth under hypoxia may have different (possibly strain-specific) mechanisms of hypoxic adaptation corresponding to their growth phenotypes under hypoxia. Nevertheless, because we have shown that most of the *M. intracellulare* strains

with different genotypes showed preferential adaptation to hypoxia in this study, our results largely reflect the clinical situation of the etiological agents of MAC-PD. A possible strategy for identifying factors that confer clinical pathogenesis will be to focus on the major clinical strains that show preferential hypoxic adaptation.

We tried to use inhibitors of phosphoenolpyruvate carboxykinase to confirm the increased essentiality of the *pckA* gene. However, we detected no positive effect because millimolar 3-mercaptopicolinic acid is required to show the effect of inhibition, even in eukaryotic cells, which have no cell wall[52]. Moreover, 3-alkyl-1,8-dibenzylxanthine[53] also only effectively inhibits PckA in eukaryotic cells. However, PckA has been proposed as a drug target for Mtb[54], so our TnSeq data seem to be consistent with studies directed towards antimycobacterial drug discovery.

We found the difference in capacity of transformation between strains belonging to the same genomic *M. paraintracellulare-M. indicus pranii* (MP-MIP) subgroup. Although the direct mechanism determining the competency for foreign DNA has not been elucidated in *M. intracellulare* and other NTM species, several studies on general bacteria suggest the difficulties of introducing foreign DNA into clinical strains compared to the laboratory strains. As suggested in *Staphylococcus aureus*[55], some clinical strains develop elimination systems of foreign nucleic acids such as a type

III-like restriction endonuclease. As suggested in gran-negative bacteria[56], there may be some difference in cell surface structures between strains, resulting in the necessity of polymyxin B nonapeptide targeting cell membrane for transforming clinical strains. The efficiency of eliminating foreign DNA may be attributed to various kinds of strain-specific factors including restriction endonuclease, natural CRISPR-interference system and cell wall structures rather than a simple genotype factor.

There may be a claim that the saturation of Tn insertion in most of each replicate were 50-60% in our libraries. However, our method of constructing Tn mutant libraries is in accordance with the previous studies of “high-density” transposon libraries[9] because the saturation of the Tn mutant libraries by combining replicates are 62-79% as follows: ATCC13950: 67.6%, M001: 72.9%, M003: 63.0%, M018: 62.4%, M019: 74.5%, M.i.27: 76.6%, M.i.198: 68.0%, MOTT64: 77.6%, M021: 79.9%. That is, we predicted gene essentiality from the Tn mutant libraries with 62-79% saturation in each strain. However, there is non-permissive sequence pattern in mariner transposon mutagenesis and using more than 10 replicates of Tn mutant libraries is reported to be a more accurate method[57,58]. The issue of high probability of judging nonessential region as mistakenly as “essential” are mainly found in small ORFs (2-9 TA sits) rather than larger ORFs (TA sites > = 10))[57]. There are 4136 (6.43% of total ORFs) nonpermissive

613 TA sites in ATCC13950, which is less than in *M. tuberculosis* H37Rv (9% of total
614 ORFs)[57] and in *M. abscessus* ATCC19977 (8.1% of total ORFs)[58]. As for small ORFs
615 (2-9 TA sites), there are nonpermissive TA sites in 41 genes (ORFs) of common
616 essential or growth-defect- associated (1.35% of a total of 3021 larger ORFs in
617 ATCC13950). With respect to the covering of non-permissive sites, our TnSeq data may
618 need to be improved to classify quite accurately the gene essentiality particularly on the
619 nonstructural small regulatory genes including small RNAs, although the study by
620 DeJesus shows the comparable number of essential genes in large genes possessing
621 more than 10 TA sites between 2 and 14 TnSeq datasets, most of which seem to be
622 structural genes[57].
623 We have provided the data suggesting the preferential hypoxic adaptation in clinical
624 strains compared to the ATCC type strain by the growth assay of individual strains. To
625 strengthen our claim, several experiments are suggested including mixed culture
626 experiments of clinical and reference strains under hypoxia. However, co-culture
627 conditions introduce additional variables, including inter-strain competition or synergy,
628 which can obscure the specific contributions of hypoxic adaptation in each strain.
629 Therefore, we took the current approach using monoculture growth curves under
630 defined oxygen conditions, which offers a clearer interpretation of strain-specific

hypoxic responses. Furthermore, one of the limitations of this study is the lack of validation of TnSeq results with individual gene knockouts. Contrary to the case of Mtb, the technique of constructing knockout mutants of slow-growing NTM including *M. intracellulare* has not been established long time. We have just recently succeeded in constructing the vector plasmids for making knockout mutants of *M intracellulare*[59]. Growth assay of individual knockout strains of genes showing increased genetic requirements such as *pckA*, *glpX*, *csd*, *eccC5* and *mycP5* in the clinical strains is suggested to provide the direct involvement of these genes on the preferential hypoxic adaptation in clinical strains. We have a future plan to construct knockout mutants of these genes to confirm the involvement of these genes on preferential hypoxic adaptation.

In summary, we have identified 131 universal essential and growth-defect-associated genes that cover the genomic diversity of *M. intracellulare* strains responsible for MAC-PD, which narrowed down the drug targets against MAC-PD. Furthermore, we have demonstrated preferential adaptation to hypoxia in clinical NTM strains compared with the type strain in terms of the increased genetic requirements for hypoxic growth, such as those encoding phosphoenolpyruvate carboxykinase, the ESX-5 type VII secretion system, and cysteine desulfurase. These

genes were shown to have a significant impact on infection *in vivo* by increasing their genetic requirements. Our data highlight the importance of using clinical strains and hypoxic conditions in experimental research, and provide fundamental resources for developing novel drugs directed towards nontuberculous mycobacterial strains that cause severe clinical diseases.

Methods

Bacterial strains and phages used to prepare Tn mutants

A total of eleven strains were used in this study, including nine clinical strains isolated from patients with MAC-PD in Japan (M.i.198, M.i.27, M018, M003, M019, M021, M001, M005, M016)[22], *M. intracellulare* genovar *paraintracellulare* type strain MOTT64 isolated from MAC-PD in South Korea[60]and the *M. intracellulare* type strain ATCC13950 (Fig. 1, Supplementary Table 1). All eleven clinical strains from MAC-PD patients in Japan were isolated from sputum[5,22]. Sputum samples were treated by the standard method for clinical isolation of mycobacteria with 0.5% (w/v) N-acetyl-L-cysteine and 2% (w/v) sodium hydroxide and plated on 7H10/OADC agar. Single colonies were picked up for use in experiments as isolated strains. Of these strains, ATCC13950, M.i.198, M.i.27, M018, M005 and M016 belong to the typical *M.*

667 *intracellulare* (TMI) genotype and M001, M003, M019, M021 and MOTT64 belong to
668 the *M. paraintracellulare-M. indicus pranii* (MP-MIP) genotype (Fig. 1, Supplementary
669 Table 1). The Tn mutant libraries were constructed using glucose as a carbon source
670 under aerobic conditions according to the transduction method used for ATCC13950[10],
671 except that a higher concentration of kanamycin (100 μg/ml) was used to select the
672 transposon mutants than was used for ATCC13950 (25 μg/ml). In brief, a hundred
673 milliliter of the bacteria were cultured aerobically in Middlebrook 7H9 medium (Difco
674 Laboratories, Detroit, MI, USA) containing 0.2% glycerol, 0.1% Tween 80 (MP
675 Biochemicals, Illkrich, France), and 10% Middlebrook albumin–dextrose–catalase
676 (7H9/ADC/Tween 80) until mid-log phase (optical density at a wavelength of 600 nm
677 [OD_{600}] of approximately 0.4–0.6). After harvesting the bacteria, 10 mL of high-titer
678 mycobacteriophage phAE180[61] was transduced into the bacteria 1-day at 37°C for
679 inserting Tn and plated on Middlebrook 7H10 solid medium supplemented with 10%
680 oleic acid-albumin-dextrose-catalase (Becton Dickinson and Co., Sparks, MD)
681 (7H10/OADC) agar plates containing kanamycin. After 2 or 3 week-cultivation at 37°C
682 aerobically, the colonies were harvested *en masse* and used as Tn mutant libraries. The
683 kanamycin-resistant colonies (> 1×10^5) were harvested and evenly resuspended in
684 7H9/ADC medium containing 20% glycerol, aliquoted, and stored at −80 °C until

further use.

Next-generation sequencing and TnSeq analysis

DNA libraries were prepared as reported previously[62]. The TnSeq libraries were sequenced with the NextSeq 500 System, 150-bp single-end run (Illumina, San Diego, CA, USA). The sequence reads were analyzed as previously described[10]. Genetic requirements were calculated, and gene essentiality was predicted by resampling and HMM analyses using TRANSIT, respectively[12]. First, an HMM analysis was conducted to calculate the fluctuation of the number of Tn insertion reads individually for each strain. HMM method can categorize the gene essentiality throughout the genome including “Essential”, “Growth-defect”, “Non-essential” and “Growth-advantage”. “Essential” genes are defined as no insertions in all or most of their TA sites. “Non-essential” genes are defined as regions that have usual read counts. “Growth-defect” genes are defined as regions that have unusually low read counts. “Growth-advantage” genes are defined as regions that have unusually high low read counts. Universal essential and growth-defect-associated genes were identified as the genes classified as “Essential” and “Growth-defect” with an HMM analysis in all 9 subject strains (Supplementary Table 4). Next, a resampling analysis was conducted to calculate the difference of the number of Tn insertion reads between each clinical strain

703 and ATCC13950 (Supplementary Table 8). By comparing the data obtained by HMM
704 and resampling analyses, we summarized the difference of genetic requirements
705 between the clinical MAC-PD strains and ATCC13950. To compare the genetic
706 requirements of the clinical MAC-PD strains and ATCC13950, the annotation for the
707 clinical MAC-PD strains was adapted from that of ATCC13950 by adjusting the START
708 and END coordinates of each open reading frame (ORF) in the clinical MAC-PD strains
709 according to their alignment with the corresponding ORFs of ATCC13950. The number
710 of Tn insertion in our datasets varied between 1.3 to 5.8 million among strains. To
711 reduce the variation in the Tn insertion across strains, we adopt a non-linear
712 normalization method, Beta-Geometric correction (BGC). BGC normalizes the datasets
713 to fit an "ideal" geometric distribution with a variable probability parameter ρ[12], and
714 BGC improves resampling by reducing the skew. On TRANSIT software, we set the
715 replicate option as Sum to combine read counts. And we normalized the datasets by
716 BGC and performed resampling analysis by using normalized datasets to compare the
717 genetic requirements between strains. The function of TnSeq-hit genes was categorized
718 according to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.

719 **Animals**

720 Female C57BL/6JJcl mice were purchased from CLEA Japan (Tokyo, Japan).

721 Six-week-old mice were used for the infection experiment (Supplementary Fig. 3). All
722 the mice were kept under specific-pathogen-free conditions in the animal facility of
723 Niigata University Graduate School of Medicine, according to the institutional
724 guidelines for animal experiments. The mice were housed (four per cage) under a
725 12-h/12-h light/dark cycle. Food and water were available *ad libitum*. Kanamycin was
726 added to the drinking water to a final concentration of 100 μg/ml to maintain the mice's
727 resistance to the Tn mutant strains. Euthanasia was conducted with an intraperitoneal
728 injection of a combined anesthetic (15 μg of medetomidine, 80 μg of midazolam, and
729 100 μg of butorphanol), followed by cervical dislocation.

730 **Mouse TnSeq experiment**

731 The protocol was based on our previous mouse lung infection experiment[22]. Twenty
732 mice each were used for the Tn mutant strains M.i.198 and M.i.27. The mice were
733 injected intratracheally with 3.25×10^8 and 5.93×10^7 CFUs of the Tn mutant strains,
734 respectively. The infected lungs were expected to be harvested from five mice at day 1,
735 week 4, week 8, and week 16 after infection. However, the lungs were harvested from
736 only four mice infected with the M.i.27 Tn mutant strains in week 8 because we
737 experienced difficulties with sampling, and the lungs were harvested from only three,
738 one, and two mice infected with M.i.198 Tn mutant strains in week 8, week 12, and

week 16, respectively, because morbidity was high in the mice infected with M.i.198. The harvested lungs were homogenized in 4.5 ml of distilled water with the gentleMACS™ Tissue Dissociator (Miltenyi Biotec, Bergish Gladbach, Germany) and the homogenates were plated on 7H10/OADC agar plates and cultured for 3 weeks. The colonies were harvested *en masse* for TnSeq.

Gene set enrichment analysis

To investigate the overlap of the gene sets required for growth in mouse lungs with those required for hypoxic pellicle formation observed in ATCC13950 and those showing increased genetic requirements observed in the clinical MAC-PD strains, a statistical enrichment of gene sets was performed by GSEA 4.3.2[15]. Parameters set for GSEA were as follows: permutations = 1,000, permutation type: gene_set, enrichment statistic: weighted, metric for ranking genes: Signal2Noise, max size: 500, min size: 15.

Construction and growth assay of knockdown strains

Knockdown strains were constructed with the CRISPR-dCas9-mediated RNA interference system, as reported previously[29]. Gene-specific 20-nt sequences were selected as protospacers for incorporation into the sgRNAs (Supplementary Table 20). pRH2502 (a vector expressing an inactivated version of the *cas9* DNA sequence) and pRH2521 (an sgRNA-expressing vector whose expression is regulated by the

757 TetR-regulated smyc promotor P_{myc1}tetO) were introduced into the bacteria to obtain
758 colonies of the knockdown strains. For selection, 25 μg/ml kanamycin and 50 μg/ml
759 hygromycin were used for ATCC13950, and 100 μg/ml kanamycin and 50 μg/ml
760 hygromycin were used for the clinical strains. The growth assay of the knockdown
761 strains was performed in 7H9/ADC broth containing the above-mentioned
762 concentrations of kanamycin and hygromycin under aerobic conditions at 37 °C with
763 humidification (APM-30D incubator, Astec, Tokyo, Japan). During the assay, the
764 cultures were supplemented with anhydrotetracycline (aTc) to a final concentration of
765 50 ng/ml (as a concentration that is not toxic to ATCC13950 and M.i.27) or 200 ng/ml
766 (for the other clinical strains). aTc was added repeatedly every 48 h to maintain the
767 induction of dCas9 and sgRNAs in experiments that extended beyond 48 h[29]. An aliquot
768 (4 μl) of the culture and its diluents were streaked onto 7H10/OADC agar plates
769 containing kanamycin and hygromycin, and the CFUs were counted. Growth inhibition
770 was evaluated as the ratio of the CFUs of knockdown strains compared with those of the
771 vector control strains lacking the guide RNA sequence. To confirm the downregulation
772 of gene expression in the knockdown strains, quantitative reverse transcription PCR
773 (qRT-PCR) was performed. Bacteria were cultured in 10 ml of 7H9/ADC/Tween 80
774 containing kanamycin and hygromycin until mid-log phase (OD_{600} approximately 0.4–

775 0.6). The cultures were supplemented with aTc every 48 h, as in the growth assay for the
776 knockdown strains. One day after the second addition of aTc, RNA was extracted by
777 bead-beating six times at 5000 rpm for 30 s with cooling (Micro Smash MS 100R Cell
778 Disruptor; Tomy Seiko, Tokyo, Japan) and cDNA was synthesized with the Direct-zol™
779 RNA Miniprep (Zymo Research, Orange, CA, USA) and ReverTra Ace qPCR RT
780 Master Mix with gDNA Remover (Toyobo, Osaka, Japan). qRT-PCR was performed
781 with the CFX Connect Real-time PCR Detection System (Bio-Rad, Hercules, CA,
782 USA), according to the manufacturer's instructions. The data were normalized to the
783 expression of *sigA*.

784 **Comparison of growth curves under aerobic and hypoxic conditions**

785 First, we precultured the bacteria in 7H9/ADC broth until the bacteria reached early
786 log-phase (OD_{600} approximately 0.1–0.3). Next, the bacterial cultures were diluted with
787 new 7H9/ADC broth (final OD_{600} 0.003) for the aerobic and hypoxic growth assay.
788 Bacterial cells were grown statically in 96-well plates containing 250 μl of 7H9/ADC
789 broth in each well under aerobic conditions or 5% oxygen at 37 °C with humidification
790 (APM-30D incubator, Astec, Tokyo, Japan). The culture was sampled every 1–3 days.
791 At sampling, the culture was mixed by pipetting 50 times. An aliquot (4 μl) of the
792 culture and its diluents were streaked onto 7H10/OADC agar and the CFUs counted.

The growth curve was fitted to the logistic 4P model with the JMP software (SAS Institute Inc., Cary, NC, USA) to calculate the inflection point (midpoint) and growth rate at midpoint in the log growth phase. Entry into the log phase was evaluated from the timing of the inflection point on the growth curve.

Statistical analyses

The cut-off of fold-change (log_2FC) to compare genetic requirements by TnSeq resampling analysis was set as adjusted *P*-values less than 0.05. Pearson's correlation test was performed between the fitness costs (log_2FC calculated by the resampling assay), efficiency of knockdown (assessed by qRT-PCR) and growth suppression in the knockdown strains. Statistical analyses of the growth curves were performed using a 4-parameter logistic model (logistic 4P model), Wilcoxon two-sample test, and a one-way analysis of variance (ANOVA) followed by Steel's multiple comparisons test using the JMP software program (SAS Institute Inc., Cary, NC, USA). Data are presented as the mean ± SD. Statistical significance was set as *P* values less than 0.05.

[References]

1. Adjemian, J., Olivier, K.N., Seitz, A.E., Holland, S.M. & Prevots, D.R. Prevalence of nontuberculous mycobacterial lung disease in U.S. Medicare beneficiaries. *Am J Respir Crit Care Med* **185**, 881-886 (2012).

2. Namkoong, H. *et al.* Epidemiology of pulmonary nontuberculous mycobacterial disease, Japan. *Emerg Infect Dis* **22**, 1116-1117 (2016).

3. Daley, C.L. *et al.* Treatment of nontuberculous mycobacterial pulmonary disease: an official ATS/ERS/ESCMID/IDSA clinical practice guideline. *Eur Respir J* **56**, 2000535 (2020).

4. Uchiya, K.I. *et al.* Comparative genome analyses of *Mycobacterium avium* reveal genomic features of its subspecies and strains that cause progression of pulmonary disease. *Sci Rep* **7**, 39750 (2017).

5. Tateishi, Y. *et al.* Comparative genomic analysis of *Mycobacterium intracellulare*: implications for clinical taxonomic classification in pulmonary *Mycobacterium avium-intracellulare* complex disease. *BMC Microbiol* **21**, 103 (2021).

6. Tortoli, E. *et al.* Genome-based taxonomic revision detects a number of synonymous taxa in the genus *Mycobacterium*. *Infect Genet Evol* **75**, 103983 (2019).

7. van Opijnen, T. & Camilli, A. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. *Nat Rev Microbiol* **11**, 435-442 (2013).

8. Carey, A.F. *et al.* TnSeq of *Mycobacterium tuberculosis* clinical isolates reveals strain-specific antibiotic liabilities. *PLoS Pathog* **14**, e1006939 (2018).

9. Akusobi, C. *et al.*. Transposon-sequencing across multiple *Mycobacterium abscessus* isolates reveals significant functional genomic diversity among strains. *mBio* **6**, e0337624 (2025).

10. Tateishi, Y. *et al.* Genome-wide identification of essential genes in *Mycobacterium intracellulare* by transposon sequencing - Implication for metabolic remodeling. *Sci Rep* **10**, 5449 (2020).

11. Cuttino, J.T. & Mc, C.A. Pure granulomatous nocardiosis, a new fungus disease distinguished by intracellular parasitism; a description of a new disease in man due to a hitherto undescribed organism, *Nocardia intracellularis*, n. sp., including a study of the biologic and pathogenic properties of this species. *Am J Pathol* **25**, 1-47 (1949).

12. DeJesus, M.A., Ambadipudi, C., Baker, R., Sassetti, C. & Ioerger, T.R. TRANSIT--A software tool for Himar1 TnSeq analysis. *PLoS Comput Biol* **11**,

e1004401 (2015).

13. Krieger, I.V. *et al.* Structure-guided discovery of phenyl-diketo acids as potent inhibitors of *M. tuberculosis* malate synthase. *Chem Biol* **19**, 1556-1567 (2012).

14. Rosconi, F. *et al.* A bacterial pan-genome makes gene essentiality strain-dependent and evolvable. *Nat Microbiol* **7**, 1580–1592 (2022).

15. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102,** 15545–15550 (2005).

16. Coll, F. *et al.* Genome-wide analysis of multi- and extensively drug-resistant *Mycobacterium tuberculosis*. *Nat Genet* **50**, 307-316 (2018).

17. Nguyen, H.T., Wolff, K.A., Cartabuke, R.H., Ogwang, S. & Nguyen, L. A lipoprotein modulates activity of the MtrAB two-component system to provide intrinsic multidrug resistance, cytokinetic control and cell wall homeostasis in *Mycobacterium*. *Mol Microbiol* **76**, 348-364 (2010).

18. Li, S. *et al.* CRISPRi chemical genetics and comparative genomics identify genes mediating drug potency in *Mycobacterium tuberculosis*. *Nat Microbiol* **7**, 766–779 (2022).

19. Kirschner, R.A., Jr., Parker, B.C. & Falkinham, J.O., 3rd. Epidemiology of

infection by nontuberculous mycobacteria. *Mycobacterium avium, Mycobacterium intracellulare,* and *Mycobacterium scrofulaceum* in acid, brown-water swamps of the southeastern United States and their association with environmental variables. *Am Rev Respir Dis* **145**, 271-275 (1992).

20. Werner, E. *et al.* Stratified growth in *Pseudomonas aeruginosa* biofilms. *Appl Environ Microbiol* **70**, 6188-6196 (2004).

21. Via, L.E. *et al.* Tuberculous granulomas are hypoxic in guinea pigs, rabbits, and nonhuman primates. *Infect Immun* **76**, 2333-2340 (2008).

22. Tateishi, Y. *et al.* Virulence of *Mycobacterium intracellulare* clinical strains in a mouse model of lung infection - role of neutrophilic inflammation in disease severity. *BMC Microbiol* **23**, 94 (2023).

23. Kryzywinska, E., Bhatnagar, S., Sweet, L., Chatterjee, D. & Schorey, J.S. *Mycobacterium avium* 104 deleted of the methyltransferase D gene by allelic replacement lacks serotype-specific glycopeptidolipids and shows attenuated virulence in mice. *Mol Microbiol* **56**, 1262-1273 (2005).

24. Irani, V.R., Lee, S.-H., Eckstein, T.M., Inamine, J.M., Belisle, J.T. & Maslow, J.N. Utilization of a *ts-sacB* selection system for the generation of a *Mycobacterium avium* serovar-8 specific glycopeptidolipid allelic exchange

mutant. A*nnals Clin Microbiol Antimicrob* **3**, 18 **(**2004).

25. Marklund, B.I., Speert, D.P. & Stokes, R.W. Gene replacement through homologous recombination in *Mycobacterium intracellulare*. *J Bacteriol* **177**, 6100-6105 (1995).

26. Marklund, B.I., Mahenthiralingam, E. & Stokes, R.W. Site-directed mutagenesis and virulence assessment of the *katG* gene of *Mycobacterium intracellulare*. *Mol Microbiol* **29**, 999-1008 (1998).

27. Mahenthiralingam, E., Marklund, B.I., Brooks, L.A., Smith, D.A., Bancroft, G.J. & Stokes, R.W. Site-directed mutagenesis of the 19-kilodalton lipoprotein antigen reveals No essential role for the protein in the growth and virulence of *Mycobacterium intracellulare*. *Infect Immun*. **66**, 3626-3634 (1998).

28. Hinds, J. *et al.* Enhanced gene replacement in mycobacteria. *Microbiology* **145**, 519-527 (1999).

29. Singh, A.K. *et al.* Investigating essential gene function in *Mycobacterium tuberculosis* using an efficient CRISPR interference system. *Nucleic Acids Res* **44**, e143 (2016).

30. Choudhary, E., Thakur, P., Pareek, M. & Agarwal, N. Gene silencing by CRISPR interference in mycobacteria. *Nat Commun* **6**, 6267 (2015).

31. Rock, J.M. *et al.* Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. *Nat Microbiol* **2,** 16274 (2017).

32. McNeil, M.B. & Cook, G.M. Utilization of CRISPR interference to validate MmpL3 as a drug target in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **63,** e00629-19 (2019).

33. Yang, T. *et al.* Pan-genomic study of *Mycobacterium tuberculosis* reflecting the primary/secondary genes, generality/individuality, and the interconversion through copy number variations. *Front Microbiol* **9**, 1886 (2018).

34. Nick, J.A. *et al.* Host and pathogen response to bacteriophage engineered against *Mycobacterium abscessus* lung infection. *Cell* **185**, 1860-1874.e12 (2022).

35. Eoh, H. *et al.* Metabolic anticipation in *Mycobacterium tuberculosis*. *Nat Microbiol* **2**, 17084 (2017).

36. Lim, J. *et al.* Phosphoenolpyruvate depletion mediates both growth arrest and drug tolerance of *Mycobacterium tuberculosis* in hypoxia. *Proc Natl Acad Sci U S A* **118,** e2105800118 (2021).

37. Marrero, J., Rhee, K.Y., Schnappinger, D., Pethe, K. & Ehrt, S. Gluconeogenic carbon flow of tricarboxylic acid cycle intermediates is critical for *Mycobacterium tuberculosis* to establish and maintain infection. *Proc Natl Acad*

Sci U S A **107**, 9819-9824 (2010).

38. Watanabe, S. *et al.* Fumarate reductase activity maintains an energized membrane in anaerobic *Mycobacterium tuberculosis*. *PLoS Pathog* **7**, e1002287 (2011).

39. Gouzy, A., Healy, C., Black, K.A., Rhee, K.Y. & Ehrt, S. Growth of *Mycobacterium tuberculosis* at acidic pH depends on lipid assimilation and is accompanied by reduced GAPDH activity. *Proc Natl Acad Sci U S A* **118,** e2024571118 (2021).

40. Quinonez, C.G. *et al.* The role of fatty acid metabolism in drug tolerance of *Mycobacterium tuberculosis*. *mBio* **13**, e0355921 (2022).

41. Pandey, A.K. & Sassetti, C.M. Mycobacterial persistence requires the utilization of host cholesterol. *Proc Natl Acad Sci U S A* **105**, 4376-4380 (2008).

42. Prosser, G. *et al.* The bacillary and macrophage response to hypoxia in tuberculosis and the consequences for T cell antigen recognition. *Microbes Infect* **19**, 177-192 (2017).

43. Dragset, M.S. *et al.* Global assessment of *Mycobacterium avium* subsp. *hominissuis* genetic requirement for growth and virulence. *mSystems* **4,** e00402-19 (2019).

44. Ates, L.S. *et al.* Essential role of the ESX-5 secretion system in outer membrane permeability of pathogenic mycobacteria. *PLoS Genet* **11**, e1005190 (2015).

45. Ates, L.S. *et al.* The ESX-5 system of pathogenic mycobacteria is nvolved in capsule integrity and virulence through its substrate PPE10. *PLoS Pathog* **12**, e1005696 (2016).

46. Savitskaya, A. *et al.* C-terminal intrinsically disordered region-dependent organization of the mycobacterial genome by a histonelike protein. *Sci Rep* 8, 1–15 (2018).

47. Enany, S. *et al.* Mycobacterial DNA-binding protein 1 is critical for long term survival of *Mycobacterium smegmatis* and simultaneously coordinates cellular functions. *Sci Rep* 7, 1–11 (2017).

48. Shaban, A.K. *et al.* Mycobacterial DNA-binding protein 1 is critical for BCG survival in stressful environments and simultaneously regulates gene expression. *Sci Rep* 13, 14157 (2023).

49. Rustad, T.R., Sherrid, A.M., Minch, K.J. & Sherman, D.R. Hypoxia: a window into *Mycobacterium tuberculosis* latency. *Cell Microbiol* **11**, 1151-1159 (2009).

50. Takahashi, Y. & Tokumoto, U. A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids. *J Biol Chem* **277**,

28380-28383 (2002).

51. Cecchini, G., Schröder, I., Gunsalus, R.P. & Maklashina, E. Succinate dehydrogenase and fumarate reductase from *Escherichia coli*. *Biochim Biophys Acta* **1553**, 140-157 (2002).

52. Brearley, M.C., Daniel, Z., Loughna, P.T., Parr, T. & Brameld, J.M. The phosphoenolpyruvate carboxykinase (PEPCK) inhibitor, 3-mercaptopicolinic acid (3-MPA), induces myogenic differentiation in C2C12 cells. *Sci Rep* **10**, 22177 (2020).

53. Montal, E.D. *et al.* Inhibition of phosphoenolpyruvate carboxykinase blocks lactate utilization and impairs tumor growth in colorectal cancer. *Cancer Metab* **7**, 8 (2019).

54. McLeod, M.J., Krismanich, A.P., Assoud, A., Dmitrienko, G.I. & Holyoak, T. Characterization of 3-[(carboxymethyl)thio]picolinic acid: a novel inhibitor of phosphoenolpyruvate carboxykinase. *Biochemistry* **58**, 3918-3926 (2019).

55. Corvaglia, A.R., François, P., Hernandez, D., Perron, K., Linder, P. & Schrenzel, J. A type III-like restriction endonuclease functions as a major barrier to horizontal gene transfer in clinical *Staphylococcus aureus* strains. *Proc Natl Acad Sci U S A* **107,** 11954-11958 (2010).

56. Qin, J., Hong, Y., Pullela, K., Morona, R., Henderson, I.R. & Totsika, M. A method for increasing electroporation competence of Gram-negative clinical isolates by polymyxin B nonapeptide. *Sci Rep* **12,**:11629 (2022).

57. DeJesus, M.A. *et al.* Comprehensive essentiality analysis of the *Mycobacterium tuberculosis* genome via saturating transposon mutagenesis. *mBio* **8,** e02133-16 (2017).

58. Rifat, D., Chen L., Kreiswirth, B.N. & Nuermberger, E.L. Genome-wide essentiality analysis of *Mycobacterium abscessus* by aaturated rransposon mutagenesis and deep sequencing. *mBio* **12,** e0104921 (2021).

59 Tateishi, Y., Nishiyama, A., Ozeki, Y. & Matsumoto, S. Construction of knockout mutants in *Mycobacterium intracellulare* ATCC13950 strain using a thermosensitive plasmid containing negative selection marker *rpsL*$^{+}$. *Microbiol Immunol* **68,** 339-347 (2024).

60. Kim, B.J. *et al.* Complete genome sequence of *Mycobacterium intracellulare* clinical strain MOTT-64, belonging to the INT1 genotype. *J Bacteriol* **194**, 3268 (2012).

61. Kriakov, J., Lee, S. & Jacobs, W.R., Jr. Identification of a regulated alkaline phosphatase, a cell surface-associated lipoprotein, in *Mycobacterium smegmatis*.

J Bacteriol **185**, 4983-4991 (2003).

62. Minato, Y. *et al*. Genomewide assessment of *Mycobacterium tuberculosis* conditionally essential metabolic pathways. *mSystems* **4,** e00070-19 (2019).

63. Shannon, P. *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).

Acknowledgments

The authors thank Dr. Torin Weisbrod and Dr. William R. Jacobs, Jr. (Albert Einstein College of Medicine, Bronx, NY, USA) for kindly providing mycobacterial phage phAE180, and Dr. Norikazu Hara and Dr. Takeshi Ikeuchi (Department of Molecular Genetics, Bioresource Science, Brain Center for Bioresources, Brain Research Institute, Niigata University) for next-generation sequencing with the NextSeq 500 System. The authors thank Mr. Naomi Yoshihara and Ms. Chisa Kumagai for constructing Tn mutant library strains and hypoxic growth experiments, respectively. This work was supported by Grants-in-Aid for Scientific Research (grant numbers 20KK0216, 21K19637, and 22H03115 to Yoshitaka Tateishi) from the Ministry of Health, Labour, and Welfare, of Japan This work was also supported by grants from the Japanese Ministry of Education, Culture, Sports, Science, and Technology, the Ministry of Health, and the Research

Program on Emerging Infectious Disease from the Japan Agency for Medical Research and Development (23fk0108673h0501, AMED-CREST 25gm1610009h0004, and JP223fa627005).

Author contributions

Y.T. conceived, designed and performed the experiments. Y.O., A.N., and S.M. contributed reagents or materials. Y. Morishige., Y. Minato, and A.D.B. contributed analytical tools. Y. Morishige., Y. Minato, A.D.B., and S.M. provided critical suggestions throughout the preparation of the manuscript. Y.T. wrote the manuscript. All authors read and approved the final manuscript.

Availability of data and materials

The datasets used or analyzed in this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

All experimental protocols were approved by the Institutional Review Board of Niigata

University (approval numbers: 2019-0020 for the use of human tissue samples and SA00506 for experiments on live vertebrates). All experiments on humans and human samples were conducted in accordance with relevant guidelines and regulations. Informed consent was obtained from all patients. All experiments on animals were conducted in accordance with the relevant guidelines and regulations, and the study complied with the ARRIVE guidelines.

Consent for publication

According to the protocols approved by the Institutional Review Board of Niigata University mentioned above, consent for publication was obtained from the patients enrolled in this study.

Competing interests

The authors declare that they have no competing interests.

Fig. 1 Phylogenetic tree of the *M. intracellulare* strains and strategy of the TnSeq analyses in this study. a Phylogenetic tree of the *M. intracellulare* strains used in this study. The tree was generated based on average nucleotide identity (ANI) with the neighbor-joining method. TMI: typical *M. intracellulare*; MP-MIP: *M. paraintracellulare–M. indicus pranii*. The detail of phylogenetic tree is shown in Supplementary Fig. 1. **b** Strategy of the procedure of TnSeq analyses.

Fig. 2 Identification of the essential and growth-defect-associated genes across genetically diverse nine ***M. intracellulare* strains used in this study. a** Functional categories of 131 genes identified as universal essential or growth-defect-associated with an HMM analysis. The genes were categorized according to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. **b** Universal essential or growth-defect-associated genes corresponding to the genes of existing antituberculous drug targets. **c** The number of the essential or growth-defect-associated genes in the accessory genes, and the number of strain-dependent essential or growth-defect-associated genes in the core genes.

Fig. 3 Detection of genes showing increased or reduced genetic requirements in the

1056 **clinical *M. intracellulare* strains.** Left panel shows the genes identified as having
1057 fewer Tn insertion reads than the type strain ATCC13950. The fold changes in the
1058 number of Tn insertion reads calculated by a resampling analysis are represented by the
1059 color scale. Red squares indicate genes required for hypoxic pellicle formation in
1060 ATCC13950[10]. *Genes identified with a combination of resampling and HMM analyses.
1061 †Genes identified only with resampling analysis. Right panel shows the genes identified
1062 as having more Tn insertion reads than the type strain ATCC13950. NH: no homolog
1063 with ATCC13950, NS: no significant increase in Tn insertion reads by a resampling
1064 analysis (adjusted $P < 0.05$).
1065
1066 **Fig. 4 Overview of the differences in genetic requirements between the clinical *M.*
1067 *intracellulare* strains and ATCC13950, drawn with Cytoscape[63].** Each central node
1068 represents the functional category of genes, assigned with a KEGG pathway analysis.
1069 Each peripheral node represents the genes showing significant changes in the number of
1070 Tn insertion reads. The size of each node represents the number of strains identified.
1071 Each edge represents a strain identified as showing significant changes in the number of
1072 Tn insertion reads. The thickness of each edge represents the adjusted *P* value. The
1073 color scale of each edge represents the fold change in the number of Tn insertion reads

calculated by a resampling analysis. The graphical organization was referenced to the previous publication by Carey, *et al*[8].

Fig. 5 Preferential genetic requirements for gluconeogenesis in the clinical *M. intracellulare* strains inferred from the TnSeq results. a Carbohydrate pathway showing the changes in Tn insertion reads. Genes identified with the resampling analysis are framed by squares. The essentiality of gluconeogenesis-related genes (*pckA*, *glpX*) was higher and that of glycolysis-related genes (*aceE*, *lpdC*) was lower in the clinical *M. intracellulare* strains compared with those in ATCC13950. **b** Data on genetic requirements calculated with a resampling analysis in the clinical *M. intracellulare* strains compared with ATCC13950. Red indicates a reduction in Tn insertion reads in the clinical *M. intracellulare* strains. Blue indicates an increase of Tn insertion reads in the clinical *M. intracellulare* strains. Asterisks indicate statistical significance of $\log_2$ fold changes ($\log_2$FCs). The graphical organization was referenced to the previous publication by Carey, *et al*[8].

Fig. 6 Detection of genes required for mouse lung infection in the clinical *M. intracellulare* strains. a Consistency between the genes required for infection in mouse

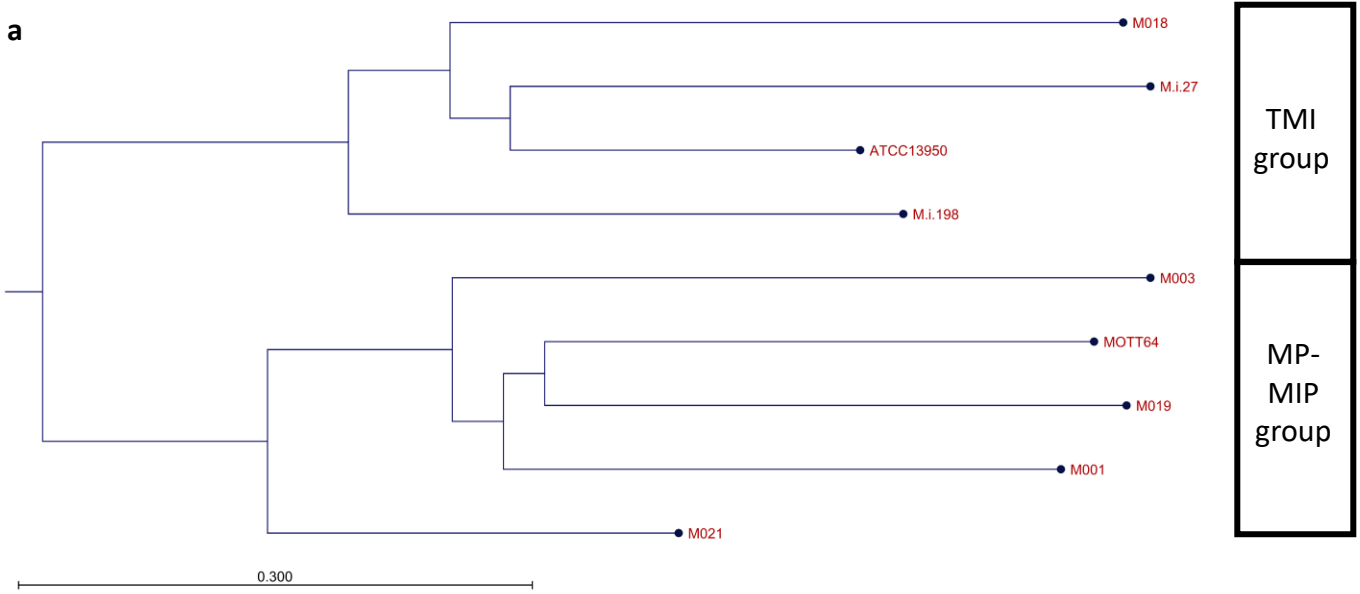
lungs and the 175 genes required for hypoxic pellicle formation in ATCC13950. **b** Changes in the requirements for genes with increased genetic requirements observed in the clinical *M. intracellulare* strains after the infection of mouse lungs. Highlighted genes are also required for hypoxic pellicle formation by ATCC13950. Genetic requirements of the strain itself: genetic requirements compared with that of ATCC13950. Genetic requirements in mouse lungs: changes in the genetic requirements in the infected mouse lungs compared with those before infection. Gene set enrichment analysis: The genes listed as core enrichment are shown as "Yes" and the genes not listed as core enrichment are shown as "No."

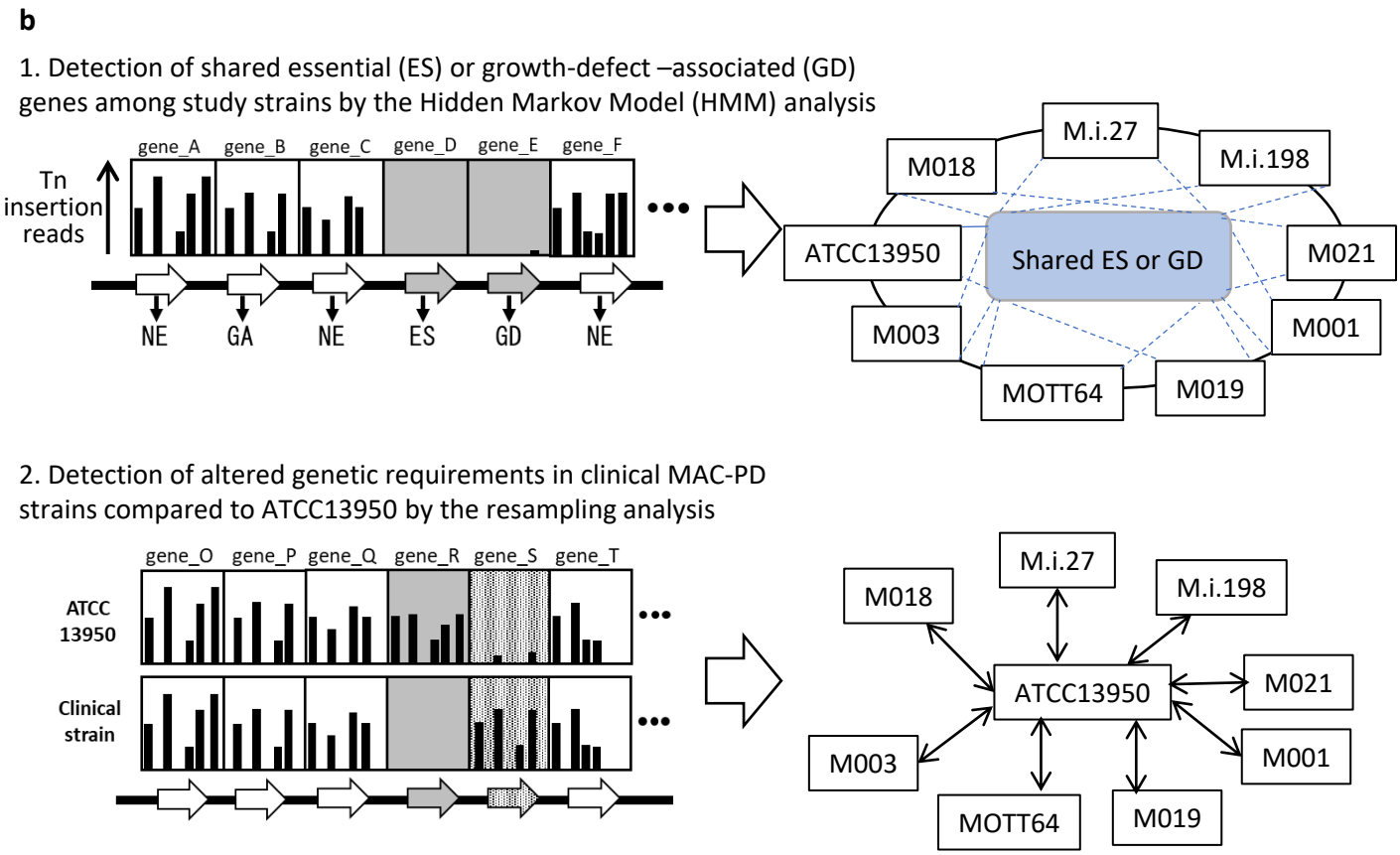
Fig. 7 Evaluation of the effect of the suppression of TnSeq-hit genes on the bacterial growth using the CRISPR-i system. Open bar: day 3; closed bar: day 7. **a** Comparative growth rates of the knockdown strains relative to those of the vector control strains in the representative universal essential or growth-defect-associated genes: *glcB, inhA*, *gyrB* and *embB*. **b** Comparative growth rates of the knockdown strains relative to those of the vector control strains in the representative accessory and strain-dependent essential or growth-defect-associated genes: *pckA*, *glpX*, *csd* and ESX-5 type VII secretion components. Data are shown as the means ± SD of triplicate

experiments. Data from one experiment representative of two independent experiments (N = 2) are shown.

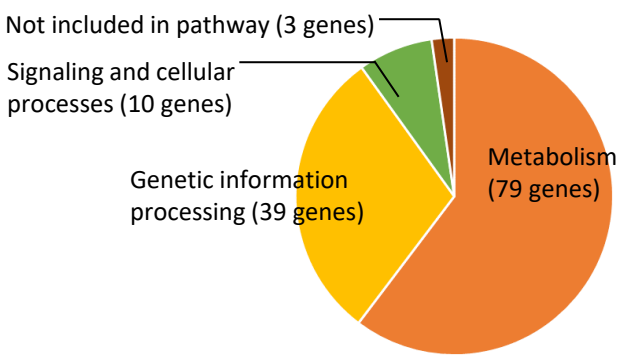
Fig. 8 Comparison of the timing of entry into logarithmic growth and logarithmic growth rate by the clinical *M. intracellulare* strains and ATCC13950. a Representative data on the growth curves of each strain under aerobic and 5% oxygen conditions. The assay was performed three times in 96-well plates containing 250 μl of broth medium per well, in triplicate. Data are represented as CFUs in 4 μl sample at each timepoint. Data on the growth curves are the means of three biological replicates from one experiment. Data from one experiment representative of three independent experiments (N = 3) are shown. **b** Time at the inflection point (midpoint) on the sigmoid growth curve. *significantly earlier than an aerobic culture of ATCC13950; #significantly earlier than a hypoxic culture of ATCC13950; †significantly later than an aerobic culture of ATCC13950; ‡significantly later than a hypoxic culture of ATCC13950. **c** Growth rate at midpoint of the growth curve in each strain. #significantly slower than a hypoxic culture of ATCC13950; †significantly slower than an aerobic culture of the corresponding strain; ‡significantly faster than an aerobic culture of ATCC13950. Open bars: aerobic; closed bars: 5% O_2. Data are shown as the

means ± SD of triplicate experiments. Data from one experiment representative of three independent experiments (N = 3) are shown.


a
M018
M.i.27
ATCC13950
M.i.198
M003
MOTT64
M019
M001
M021
0.300
TMI group
MP-MIP group


b
1. Detection of shared essential (ES) or growth-defect –associated (GD) genes among study strains by the Hidden Markov Model (HMM) analysis
gene_A
gene_B
gene_C
gene_D
gene_E
gene_F
Tn insertion reads
NE
GA
NE
ES
GD
NE
M.i.27
M018
M.i.198
ATCC13950
Shared ES or GD
M021
M003
M001
MOTT64
M019
2. Detection of altered genetic requirements in clinical MAC-PD strains compared to ATCC13950 by the resampling analysis
gene_O
gene_P
gene_Q
gene_R
gene_S
gene_T
ATCC 13950
Clinical strain
M.i.27
M018
M.i.198
ATCC13950
M021
M003
M001
MOTT64
M019
3. Detection of genes required for mouse lung infection
Intratracheal injection
M.i.27
M.i.198
Day0
Day1
Week4
Week8
Week16
TnSeq
TnSeq
TnSeq
TnSeq

a



Metabolism: glyoxylate, purine/pyrimidine, amino acid, lipid, arabinogalactan/peptidoglycan, porphyrin

Genetic information processing: DNA replication, transcription, amino acid tRNA ligases, ribosomal proteins, general Sec proteins

Signaling and cellular processes: ABC transporter, type VII secretion system, iron-sulfur cluster assembly

Not Included in Pathway: hypothetical proteins

b

Gene locus	gene	function	target drug
OCU_RS25020, OCU_RS25025	*gyrB, gyrA*	DNA topoisomerase subunit B, intein-containing DNA gyrase subunit A	fluoroquinolones
OCU_RS26150	*embB*	arabinosyltransferase	ethambutol
OCU_RS40185	*inhA*	enoyl[acyl-carrier-protein] reductase FabI	isoniazid/ethionamid
OCU_RS45570	*alr*	alanine racemase	cycloserine
OCU_RS46225	*rpoC*	DNA-directed RNA polymerase subunit beta'	rifampicin
OCU_RS38100	*glcB*	malate synthase G	malate synthase inhibitors
OCU_RS32840, OCU_RS32845, OCU_RS32850	*atpA, atpG, atpB*	F_0F_1 ATP synthase subunit alpha, gamma, beta	bedaquiline

c

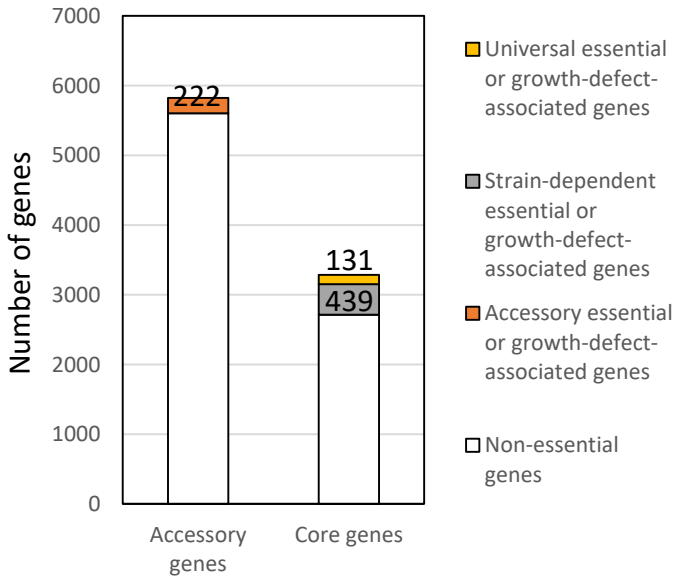
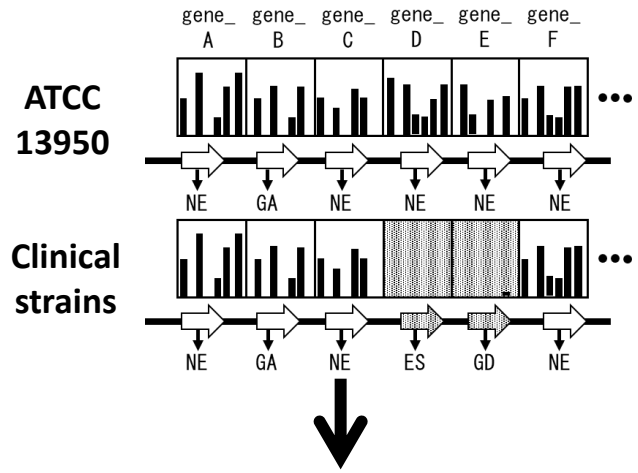
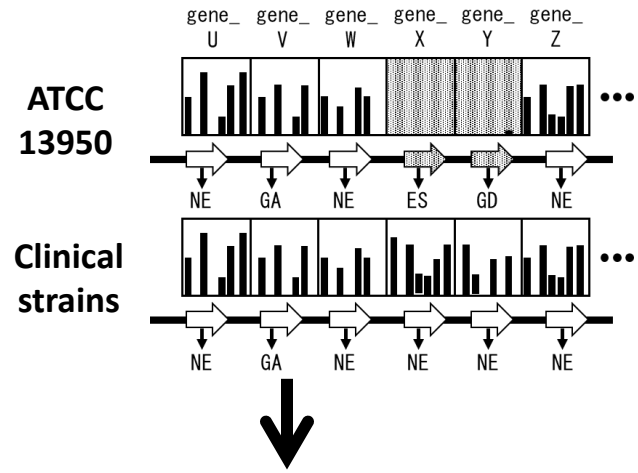


Fig. 2



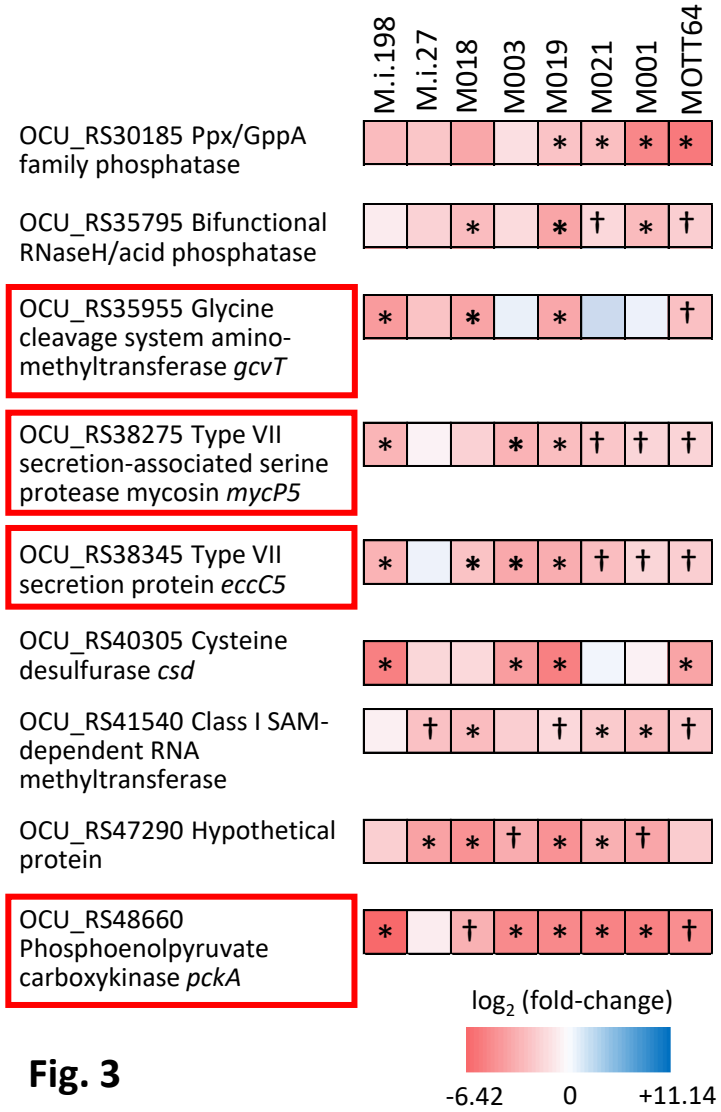


Tn decrease in clinical strains

A total of 121 genes were detected as significantly reduced Tn in clinical MAC-PD strains compared to ATCC13950. Of those, 9 genes were commonly hit in 3 to 5 clinical strains.

Tn increase in clinical strains

A total of 162 genes were detected as significantly increased Tn insertion compared to ATCC13950. Of those, 12 genes were commonly hit in more than 7 clinical MAC-PD strains.



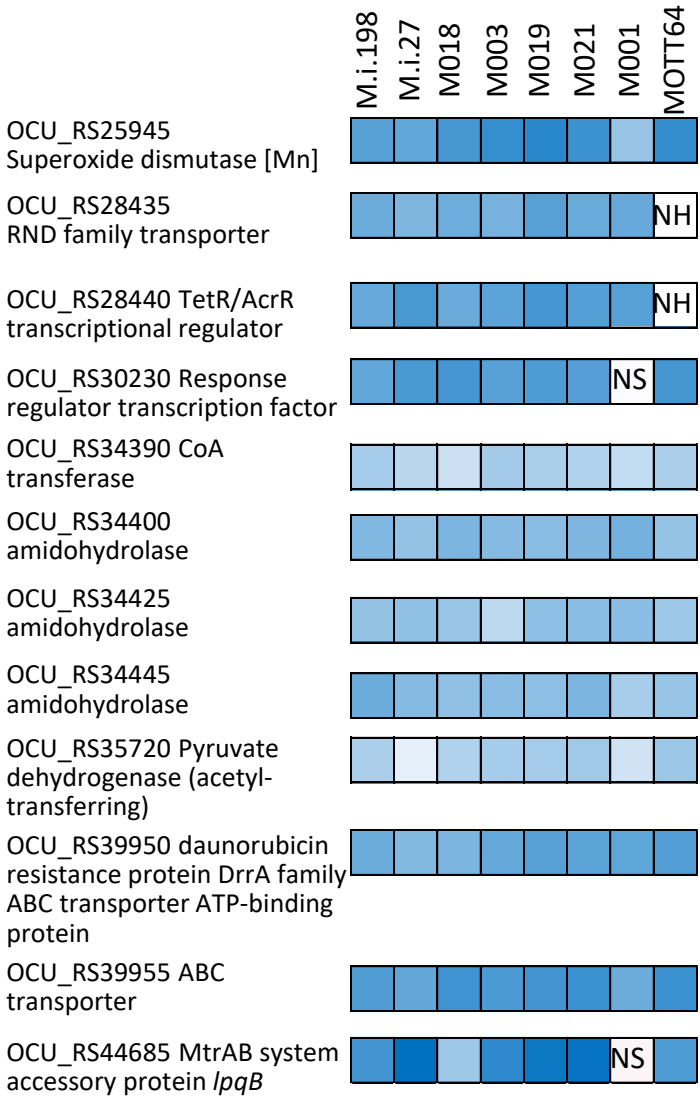


Fig. 3

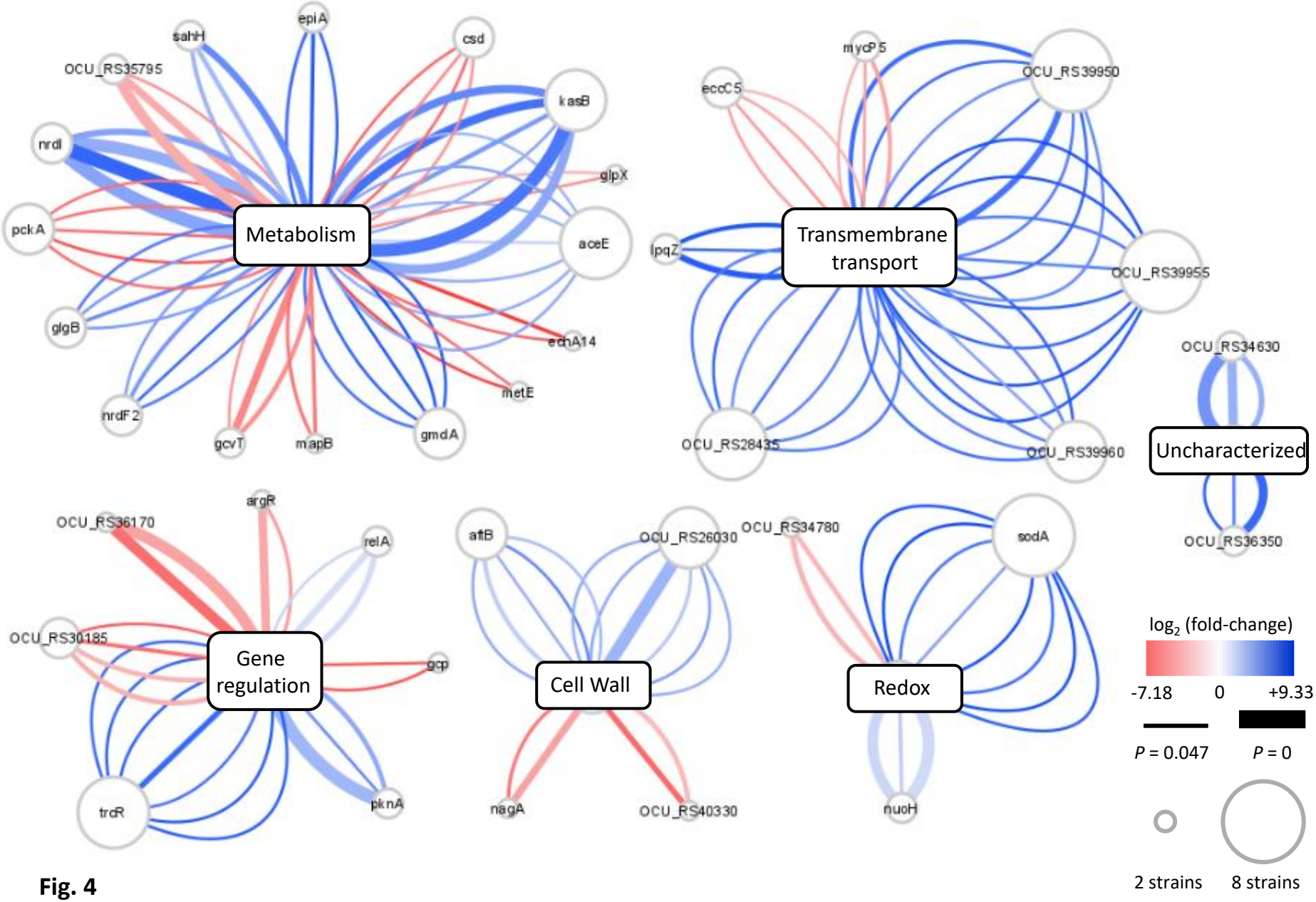

Metabolism
epiA
sahH
csd
OCU_RS35795
kasB
nrdI
glpX
pckA
aceE
glgB
echA14
metE
nrdF2
gcvT
mapB
gmdA
Transmembrane transport
mycP5
OCU_RS39950
eccC5
lpqZ
OCU_RS39955
OCU_RS28435
OCU_RS39960
Uncharacterized
OCU_RS34630
OCU_RS36350
Gene regulation
argR
OCU_RS36170
relA
OCU_RS30185
gcp
trcR
pknA
Cell Wall
aftB
OCU_RS26030
nagA
OCU_RS40330
Redox
OCU_RS34780
sodA
nuoH
log2 (fold-change)
-7.18
0
+9.33
P = 0.047
P = 0
2 strains
8 strains

Fig. 4

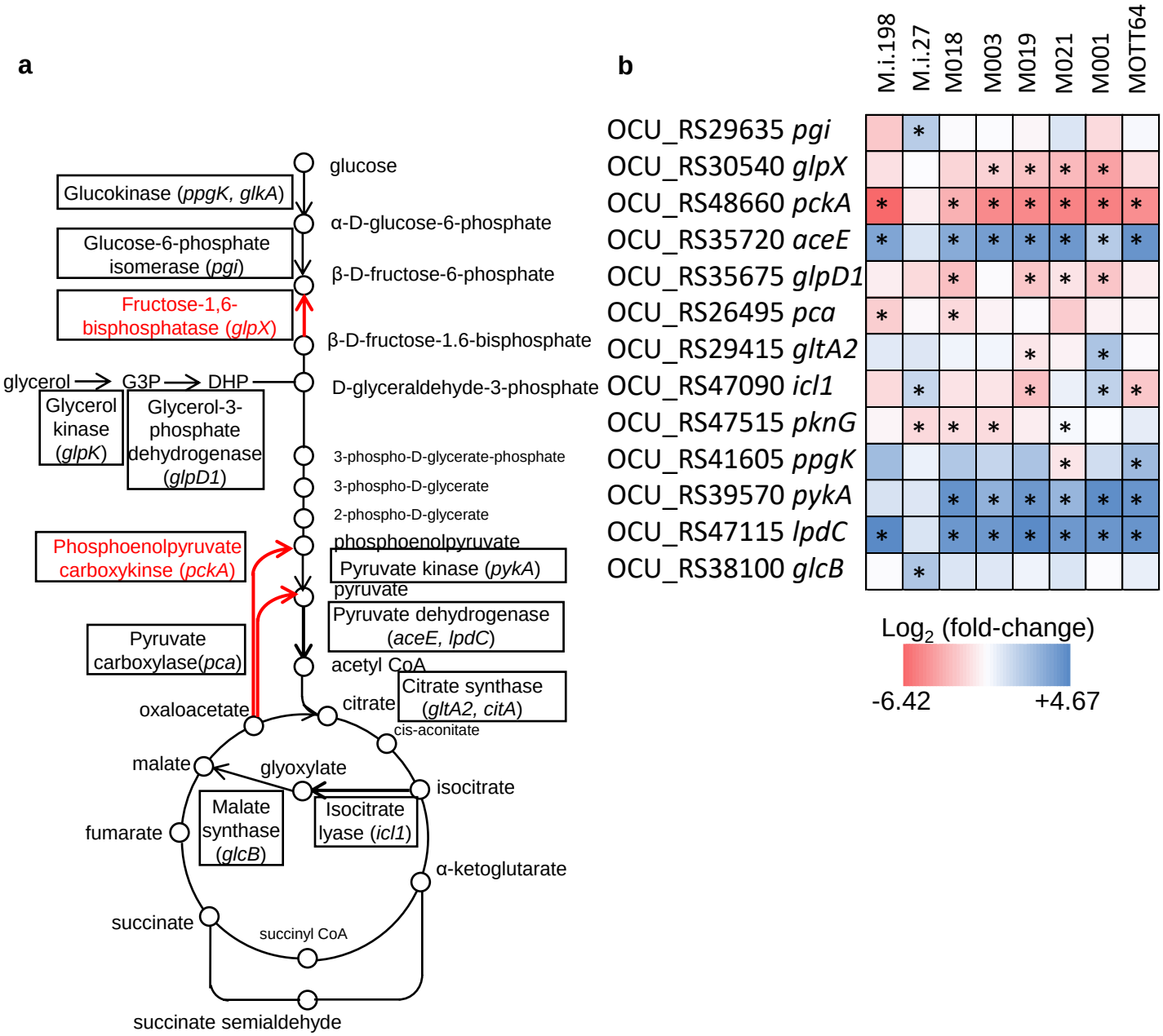
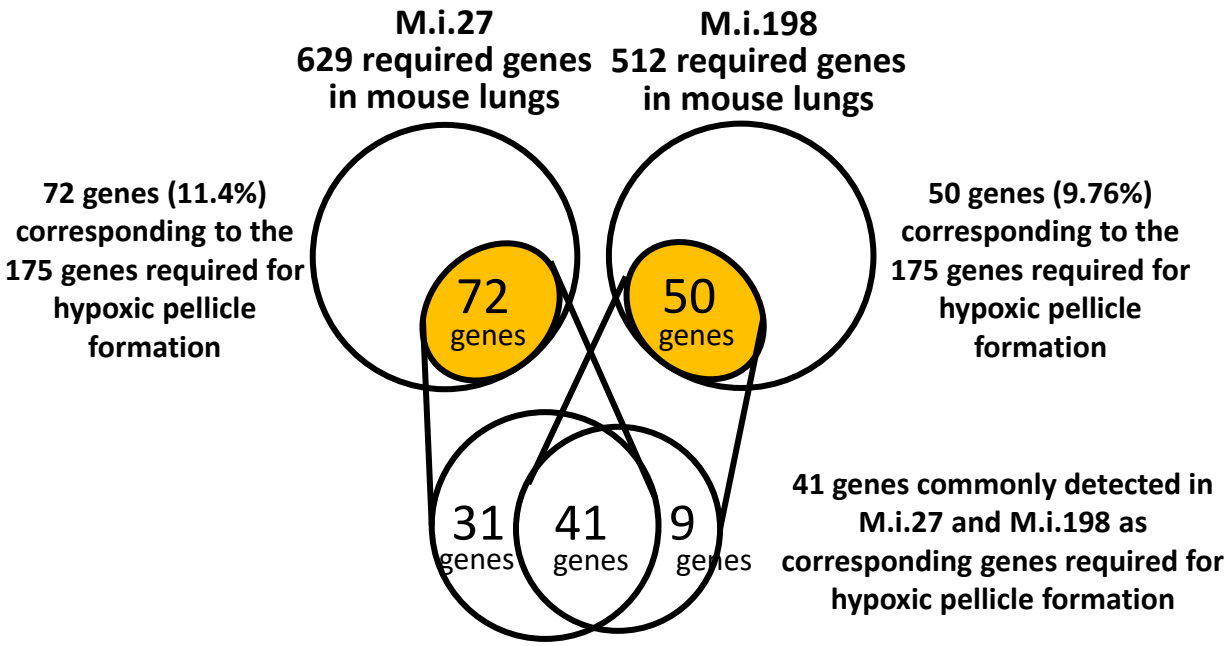

a
glucose
Glucokinase (ppgK, glkA)
α-D-glucose-6-phosphate
Glucose-6-phosphate isomerase (pgi)
β-D-fructose-6-phosphate
Fructose-1,6-bisphosphatase (glpX)
β-D-fructose-1.6-bisphosphate
glycerol
G3P
DHP
D-glyceraldehyde-3-phosphate
Glycerol kinase (glpK)
Glycerol-3-phosphate dehydrogenase (glpD1)
3-phospho-D-glycerate-phosphate
3-phospho-D-glycerate
2-phospho-D-glycerate
phosphoenolpyruvate
Phosphoenolpyruvate carboxykinse (pckA)
Pyruvate kinase (pykA)
pyruvate
Pyruvate dehydrogenase (aceE, lpdC)
Pyruvate carboxylase(pca)
acetyl CoA
Citrate synthase (gltA2, citA)
oxaloacetate
citrate
cis-aconitate
malate
glyoxylate
isocitrate
Malate synthase (glcB)
Isocitrate lyase (icl1)
fumarate
α-ketoglutarate
succinate
succinyl CoA
succinate semialdehyde
b
M.i.198
M.i.27
M018
M003
M019
M021
M001
MOTT64
OCU_RS29635 pgi
OCU_RS30540 glpX
OCU_RS48660 pckA
OCU_RS35720 aceE
OCU_RS35675 glpD1
OCU_RS26495 pca
OCU_RS29415 gltA2
OCU_RS47090 icl1
OCU_RS47515 pknG
OCU_RS41605 ppgK
OCU_RS39570 pykA
OCU_RS47115 lpdC
OCU_RS38100 glcB
Log2 (fold-change)
-6.42
+4.67

Fig 5.

a



b

Locus_tag	Gene annotation	M.i.27 Genetic require-ments of strain itself	M.i.27 Genetic require-ments in mouse lungs	M.i.27 Gene set enrichi-ment analysis	M.i.198 Genetic require-ments of strain itself	M.i.198 Genetic require-ments in mouse lungs	M.i.198 Gene set enrichi-ment analysis
OCU_RS48660	Phoshoenolpyruvate carboxykinase PckA	NC	UP	Yes	UP	NC	No
OCU_RS30540	Fructose-1,6-bisphosphatase GlpX	NC	UP	Yes	NC	NC	Yes
OCU_RS40305	Cysteine desulfurase Csd	NC	UP	Yes	UP	NC	No
OCU_RS38345	TypeVII secretion protein EccC5	NC	UP	Yes	UP	NC	Yes
OCU_RS38275	TypeVII serine protease MycP5	NC	UP	Yes	UP	NC	No
OCU_RS35955	Glycine cleavage system aminomethyltransferase GcvT	NC	UP	Yes	UP	NC	Yes
OCU_RS30185	Ppx/GppA family phosphatase	NC	NC	Yes	NC	NC	Yes
OCU_RS35795	Bifunctional RNaseH/acid phosphatase	NC	NC	No	NC	NC	Yes
OCU_RS41540	Class I SAM-dependent RNA methyltransferase	UP	NC	No	NC	UP	Yes
OCU_RS47290	Hypothetical protein	UP	NC	Yes	NC	NC	Yes

Fig. 6

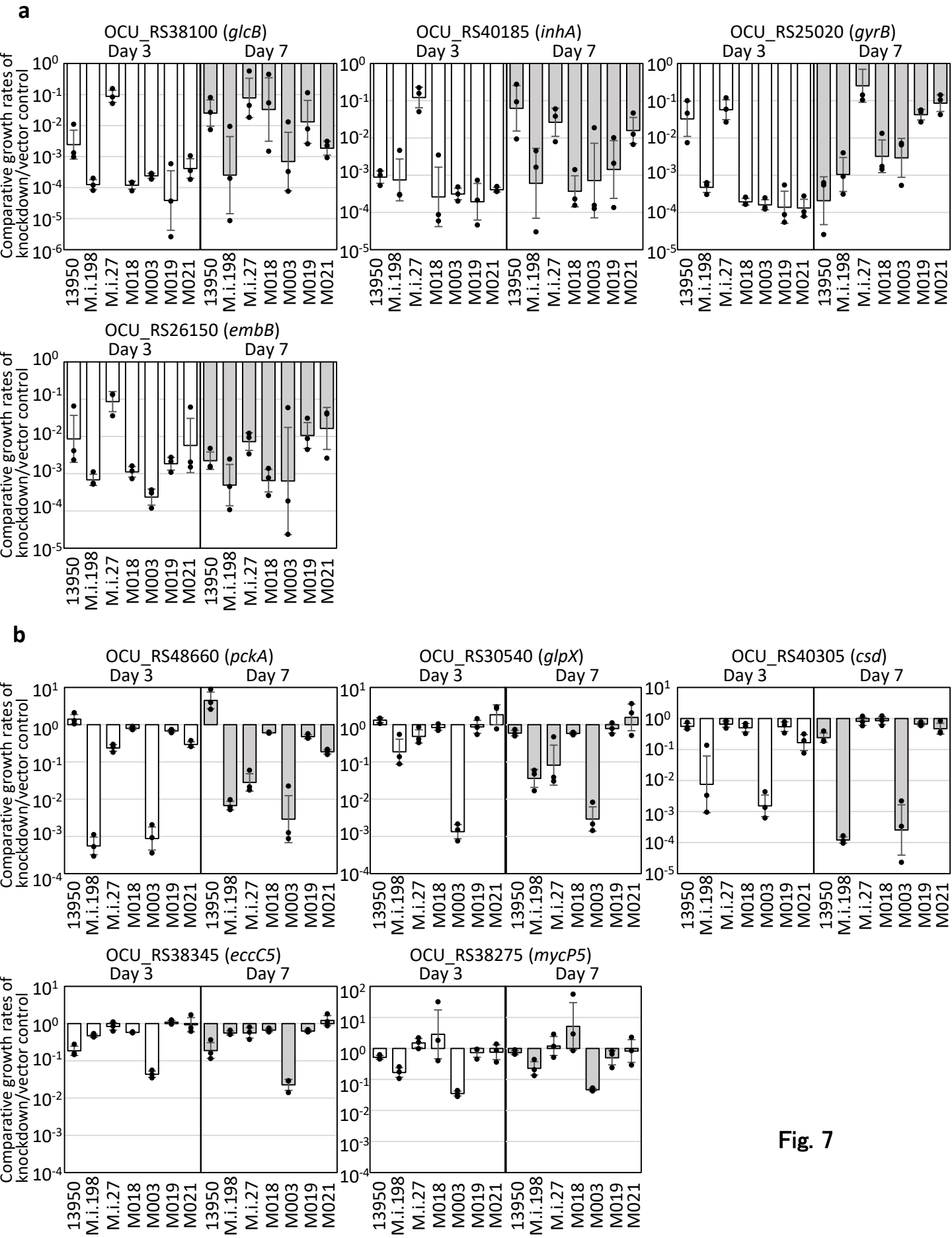
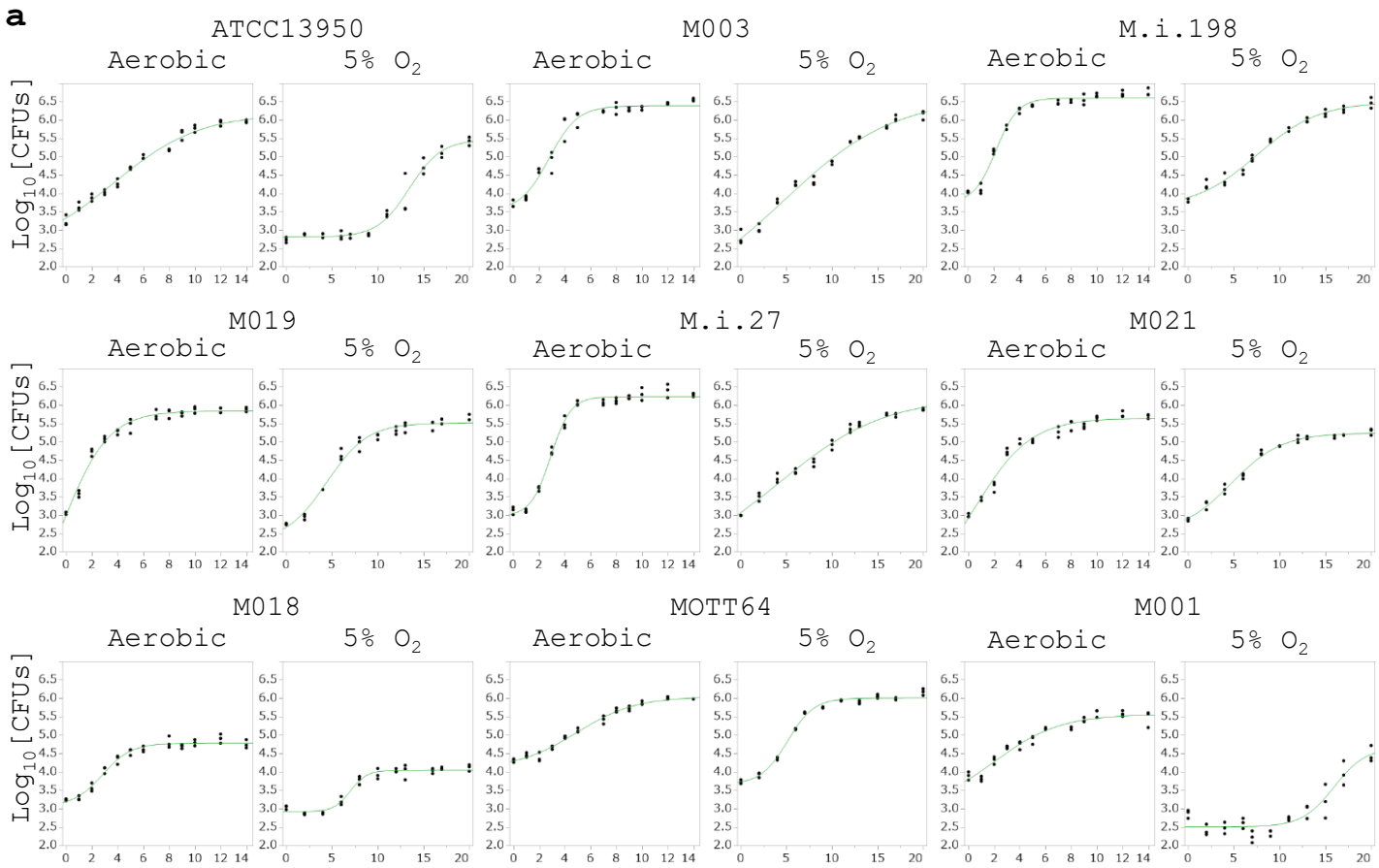

a
OCU_RS38100 (glcB)
OCU_RS40185 (inhA)
OCU_RS25020 (gyrB)
OCU_RS26150 (embB)
Day 3
Day 7
Comparative growth rates of knockdown/vector control
13950
M.i.198
M.i.27
M018
M003
M019
M021
b
OCU_RS48660 (pckA)
OCU_RS30540 (glpX)
OCU_RS40305 (csd)
OCU_RS38345 (eccC5)
OCU_RS38275 (mycP5)

Fig. 7


a
ATCC13950
M003
M.i.198
M019
M.i.27
M021
M018
MOTT64
M001
Aerobic
5% O2
Log10 [CFUs]

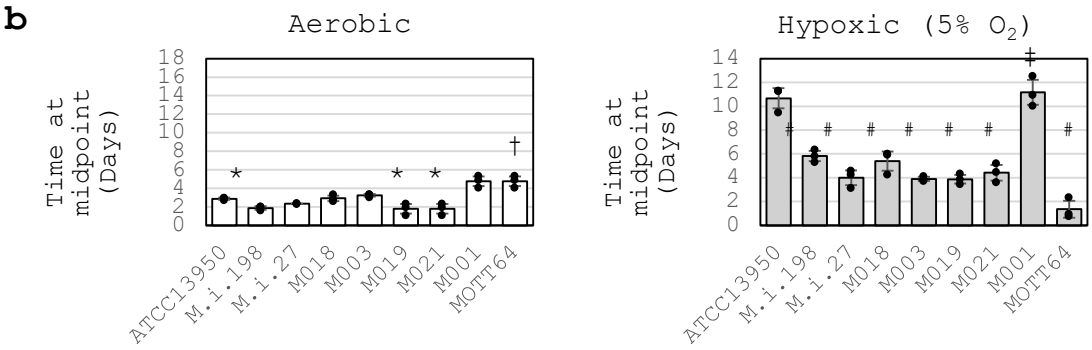
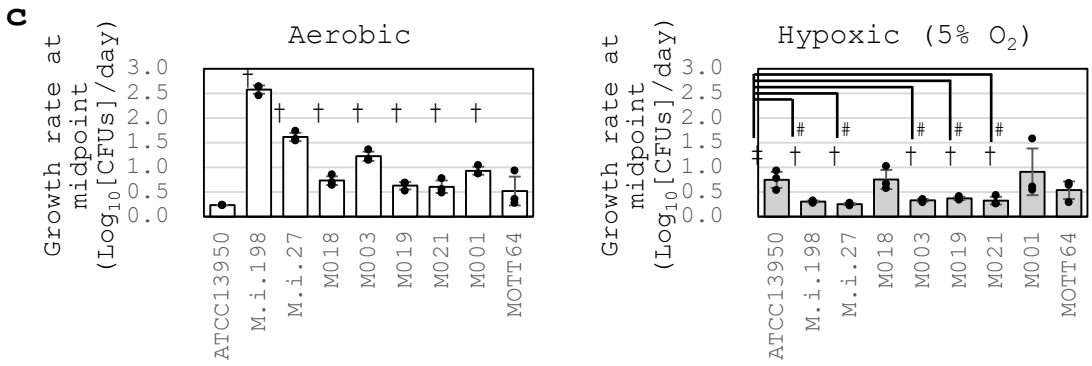

b
Aerobic
Hypoxic (5% O2)
Time at midpoint (Days)
ATCC13950
M.i.198
M.i.27
M018
M003
M019
M021
M001
MOTT64

Fig. 8


c
Aerobic
Hypoxic (5% O2)
Growth rate at midpoint (Log10[CFUs]/day)
ATCC13950
M.i.198
M.i.27
M018
M003
M019
M021
M001
MOTT64