Genome-wide association study in patients with pulmonary *Mycobacterium avium* complex disease

Ho Namkoong1,2,3,25,26, Yosuke Omae4,5,25, Takanori Asakura1,6, Makoto Ishii1, Shoji Suzuki1, Kozo Morimoto7, Yosuke Kawai4,5, Katsura Emoto8, Andrew J. Oler9, Eva P. Szymanski2, Mitsunori Yoshida6, Shuichi Matsuda7, Kazuma Yagi1, Isano Hase10, Tomoyasu Nishimura11, Yuka Sasaki7, Takahiro Asami1, Tetsuya Shiomi12, Hiroaki Matsubara13, Hisato Shimada14, Junko Hamamoto1, Byung Woo Jhun15, Su-Young Kim15, Hee Jae Huh16, Hong-Hee Won17, Manabu Ato6, Kenjiro Kosaki18, Tomoko Betsuyaku1, Koichi Fukunaga1, Atsuyuki Kurasima7, Hervé Tettelin19,20, Hideki Yanai21, Surakameth Mahasirimongkol22, Kenneth N. Olivier23, Yoshihiko Hoshino6, Won-Jung Koh15, Steven M. Holland2, Katsushi Tokunaga4,5,26, and Naoki Hasegawa24,26, the Nontuberculous Mycobacteriosis and Bronchiectasis – Japan Research Consortium (NTM-JRC)

1Division of Pulmonary Medicine, Dept of Medicine, Keio University School of Medicine, Tokyo, Japan. 2Laboratory of Clinical Immunology and Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH), Bethesda, MD, USA. 3JSPS Overseas Research Fellow, Japan Society for the Promotion of Science, Tokyo, Japan. 4Dept of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. 5Genome Medical Science Project (Toyama), National Center for Global Health and Medicine, Tokyo, Japan. 6Dept of Mycobacteriology, Leprosy Research Center, National Institute of Infectious Diseases, Tokyo, Japan. 7Respiratory Disease Center, Fukuiji Hospital, Japan Anti-Tuberculosis Association, Tokyo, Japan. 8Dept of Pathology, Keio University School of Medicine, Tokyo, Japan. 9Bioinformatics and Computational Biosciences Branch, Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH), Bethesda, MD, USA. 10Dept of Respiratory Medicine, National Hospital Organization Utsunomiya Hospital, Tochigi, Japan. 11Keio University Health Center, Tokyo, Japan. 12Dept of Pulmonary Medicine, Keiyu Hospital, Kanagawa, Japan. 13Dept of Pulmonary Medicine, Fussa Hospital, Tokyo, Japan. 14Dept of Pulmonary Medicine, Kawasaki Municipal Ida Hospital, Kanagawa, Japan. 15Division of Pulmonary and Critical Care Medicine, Dept of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea. 16Dept of Laboratory Medicine and Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea. 17Samsung Advanced Institute for Health Sciences and Technology (SAIHST), Sungkyunkwan University, Samsung Medical Center, Seoul, South Korea. 18Center for Medical Genetics, Keio University School of Medicine, Tokyo, Japan. 19Dept of Microbiology and Immunology, School of Medicine, University of Maryland, Bethesda, MD, USA. 20Institute for Genome Sciences, School of Medicine, University of Maryland, Bethesda, MD, USA. 21Dept of Clinical Laboratory, Fukuiji Hospital, Japan Anti-Tuberculosis Association, Tokyo, Japan. 22Medical Genetics Center, Medical Life Sciences Institute, Dept of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand. 23Pulmonary Branch, National Heart, Lung, and Blood Institute, National Institutes of Health (NIH), Bethesda, MD, USA. 24Dept of Infectious Diseases, Keio University School of Medicine, Tokyo, Japan. 25H. Namkoong and Y. Omae are co-first authors. 26H. Namkoong, N. Hasegawa and K. Tokunaga contributed equally to this article as lead authors and supervised the work.

Shareable abstract (@ERSpublications)
The first GWAS of pulmonary MAC disease in Japanese, Korean and European patients. SNPs in the CHP2 region were associated with the disease risk. *CHP2* may play an important role in host susceptibility to pulmonary MAC disease. https://bit.ly/39iCIio

polymorphisms (SNPs) in another Japanese cohort. For verification by Korean and European ancestry, we performed SNP genotyping.

**Results** The GWAS discovery set included 475 pulmonary MAC cases and 417 controls. Both GWAS and replication analysis of 591 pulmonary MAC cases and 718 controls revealed the strongest association with chromosome 16p21, particularly with rs109592 ($p=1.64\times10^{-13}$, OR 0.54), which is in an intronic region of the calcineurin-like EF-hand protein 2 (CHP2). Expression quantitative trait loci analysis demonstrated an association with lung CHP2 expression. CHP2 was expressed in the lung tissue in pulmonary MAC disease. This SNP was associated with the nodular bronchiectasis subtype. Additionally, this SNP was significantly associated with the disease in patients of Korean ($p=2.18\times10^{-12}$, OR 0.54) and European ($p=5.12\times10^{-03}$, OR 0.63) ancestry.

**Conclusions** We identified rs109592 in the CHP2 locus as a susceptibility marker for pulmonary MAC disease.

**Introduction** Pulmonary nontuberculous mycobacterial (NTM) diseases are a major global public health issue [1]. Studies have revealed a steady increase in the incidence and prevalence of pulmonary NTM diseases in North America, Europe and Asia [2, 3]. The reasons are unclear, but may be related to increased awareness and improved diagnostic techniques.

*Mycobacterium avium* complex (MAC) is the most common cause of pulmonary NTM diseases, which are generally chronic and progressive. Eradicating this pathogen using standard antimicrobial chemotherapy is difficult and some patients develop respiratory failure, which is ultimately fatal [4]. To improve the treatment of pulmonary NTM disease, a fundamental understanding of its pathogenesis is essential.

While NTM pathogens exist ubiquitously in the environment including in the soil and drinking water, few exposed people develop pulmonary disease [5]. Although pulmonary NTM diseases are associated with underlying disorders such as bronchiectasis, COPD and primary ciliary dyskinesia, most patients with pulmonary NTM lack any recognised coexisting illness or apparent immunodeficiency [6–8]. Most patients are immunocompetent post-menopausal women with a low body mass index [9]. Familial clustering has been reported with a higher incidence in Asians than Caucasians or African Americans, which suggests population differences in susceptibility [2, 10]. These findings indicate that it is likely that genetic predispositions exist for pulmonary NTM disease.

Mendelian susceptibility to mycobacterial disease (MSMD) has been investigated previously [11, 12]. Impairment of the interleukin-12/interferon-γ axis in the T-helper 1 response plays a central role in MSMD. Additionally, anti-interferon-γ autoantibodies have been examined as a cause of adult-onset acquired immunodeficiency-associated mycobacterial disease [13]. However, the clinical phenotypes associated with MSMD and anti-interferon-γ autoantibodies are typically disseminated infections, which are quite different from the pulmonary NTM disease phenotype. Accordingly, the pathogeneses of pulmonary NTM disease and MSMD are probably distinct.

Previous genetic studies of pulmonary NTM diseases involved whole-exome sequencing, linkage analysis and candidate gene analysis [14–18]. However, these studies employed relatively small sample sizes and therefore require additional verification. Pulmonary NTM diseases are likely to be influenced both by environmental and polygenic factors. Genome-wide association studies (GWAS) are a powerful tool for investigating host susceptible genes, and thus generating functional insights into disease pathogenesis [19].

In this study, we performed the first GWAS study in Japanese patients with pulmonary MAC, followed by replication in independent Japanese, Korean and European cohorts.

**Methods**

For further details on the applied methods, please refer to the supplementary material.

**Sample and clinical data collection**

The study design and workflow for the GWAS of pulmonary MAC disease are summarised in figure 1.
For subsequent GWAS analysis, the pulmonary MAC patient subgroup was extracted from the pulmonary NTM patient group. 419 controls were independently recruited from the Tokyo area, as described in our previous GWAS report [20]. The validation cohort of 591 patients with pulmonary MAC disease was collected from other collaborating institutions near Tokyo and compared to 718 healthy Japanese adult volunteers recruited from the Pharma SNP Consortium and Health Science Research Resources Bank of Japan Health Sciences Foundation. With a sample size of 1066 cases and 1135 controls in our Japanese cohort, we calculated that we had >80% power (88.6%) to detect a genotype relative risk of 1.6 at a minor allele frequency (MAF) of 0.2 [21].

The Korean cohort included samples from patients with pulmonary MAC from the database of the NTM Registry of Samsung Medical Center (Seoul, South Korea). As a Korean control group, we used single-nucleotide polymorphism (SNP) data from the Korean Reference Genome Database.

The European cohort included patients with pulmonary MAC disease recruited from 2001 to 2018 at the National Institutes of Health Clinical Center (Bethesda, MD, USA) [14]. Among the European population in the 1000 Genomes Project phase 3, we used TSI (Toscani in Italia, n=107) and CEU (Utah residents with northern and western European Ancestry, n=99) populations as a reference, because principal component analysis of these patients revealed that most patients were identical to those of the TSI and CEU.
We collected demographic data to evaluate sex, age of onset and chest high-resolution computed tomography images from accessible cases in Japan and Korea (n=1175). Images were evaluated by two independent pulmonologists who were blinded to the clinical data. Discrepancies were resolved by consensus review. Using these images, we classified the subjects into the nodular bronchiectasis (NB) and non-NB (fibrocavitary (FC), mixed and unclassified) subgroups.

**Replication study**

DigiTag2 assay or TaqMan SNP genotyping assays (Applied Biosystems, Foster City, CA, USA) were used to genotype the candidate SNPs identified in the GWAS discovery set in the Japanese population. For the replication study, we selected SNPs with info score 1.000 and genotyped by the Affymetrix Axiom Japonica array and in linkage disequilibrium ($r^2>0.7$) with SNPs having a suggestive evidence of association ($p$-value<$3\times10^{-6}$). For the replicated chromosomal region, additional SNPs with expression quantitative trait locus (eQTL) evidence in RegulomeDB v 1.1 (category 1) and moderately suggestive association ($p$-value<$5\times10^{-3}$) were further validated to detect primary association in the locus [22]. In total, six candidate SNPs were genotyped successfully with a call rate of >99% in 591 MAC cases and 718 healthy controls.

**Results**

**GWAS discovery stage**

We genotyped 659503 variants in patients with MAC and healthy control individuals and 622273 autosomal variants passed quality control filtering. Study samples that clustered with the HapMap JPT samples following principal component analysis were selected (supplementary figure S1). 475 patients with pulmonary MAC and 417 healthy control individuals were included in the GWAS discovery stage. After applying whole-genome imputation using a phased reference panel of 2049 Japanese individuals from a prospective general population cohort study performed by the Tohoku Medical Megabank Organization, 5876247 autosomal variants were included in the GWAS [23, 24]. The genomic inflation factor was 1.033, suggesting low genomic inflation due to population stratification or cryptic relatedness among samples (supplementary figure S2). Although no SNP passed the standard genome-wide significance threshold ($p<5\times10^{-8}$), the strongest association was observed on rs194788 ($p=1.75\times10^{-7}$, OR 0.54, 95% CI 0.44–0.69) at chromosome 16p21 (supplementary table S1). The most significant SNP, rs194788, was in complete linkage disequilibrium ($r^2=1.000$) with genotyped SNP rs11646605 with info score 1.000 ($p=2.67\times10^{-7}$, OR 0.55, 95% CI 0.44–0.69) (figure 2 and supplementary table S1). 18 SNPs showed suggestive evidence of association ($p<3\times10^{-6}$) at chromosome 16p21, and these SNPs are in the intergenic region between the calcineurin-like EF-hand protein 2 (CHP2) gene and protein kinase C β (PRKCB) gene (figure 3).

Loci previously reported to be associated with susceptibility to leprosy and tuberculosis showed no significant associations in this study, suggesting that the genetic risk factors for pulmonary MAC infection differ from those for other mycobacterial species (supplementary table S2).

**Replication study**

Besides the most significant association at chromosome 16p21, 18 SNPs from three independent loci showed suggestive evidence of associations (supplementary table S1). To confirm these preliminary findings, four candidate variants from four independent loci with info scores of 1.000 and which might tag other imputed SNPs were genotyped in 591 patients with MAC and 718 healthy individuals (supplementary table S3). These subjects were recruited independently from among Japanese GWAS participants. For the chromosome 16p21 locus replication, two imputed SNPs with a $p$-value=$5\times10^{-3}$ and eQTL evidence in RegulomeDB (category 1) were further validated to detect any primary association at this locus. Three SNPs at chromosome 16p21 showed associations ($p<0.05$) in this replication dataset, and they remained significant after Bonferroni correction (supplementary table S3). An intrinsic CHP2 SNP (rs1095952) reached the GWAS significance threshold after combining the discovery and replication datasets ($p=1.64\times10^{-13}$, OR=0.54; table 1). The significant association of rs11646605 was lost when conditional analysis was performed with rs1095952 ($p=0.118$) (table 1). The association of chromosome 16p21 observed at the discovery stage also disappeared after conditional analysis with rs1095952 (supplementary figure S3). These results suggest that the association between this region and MAC susceptibility is in linkage disequilibrium with rs1095952. Candidate SNPs from the other three loci showed no association with the replication dataset (supplementary table S3).
Intronic SNPs may modulate the expression of neighbouring genes. To assess the functional effects of rs109592, we conducted eQTL analysis. We found that the rs109592 genotype significantly affected \textit{CHP2} expression in the lung ($p=2.1\times10^{-5}$, normalised effect size (NES) 0.28) (figure 4). Additionally, \textit{PRKCB} expression showed significant eQTL associations with rs109592 in several tissues including artery-tibial ($p=7.8\times10^{-14}$, NES 0.33), but not in lung ($p=0.033$). An intronic \textit{PRKCB} SNP, rs7200798, showed more significant eQTL association in artery-tibial and whole blood than rs109592 ($p=1.7\times10^{-17}$, NES −0.31 and $p=5.2\times10^{-20}$, NES 0.14, respectively), but showed only a moderate association with MAC infection in our GWAS (table 1). Colocalisation analysis did not detect any significant overlap of specific eQTL signals with our GWAS association on chromosome 16p21 from the discovery dataset.

**Immunostaining of resected lung tissue from patients with pulmonary MAC infection**

Immunostaining showed that CHP2 was positively associated with fibroblasts surrounding granulomas in all stained cases (figure 5 and supplementary figure S4). In addition, CHP2 was positively related to a reticular pattern in the germinal centre of tertiary lymphoid structures (figure 5).

**Replication study in other populations (Korean and European ancestry)**

Next, we investigated the association of this SNP in a Korean cohort (MAC n=722; controls n=1722) and a European cohort (MAC n=276; controls n=206). This SNP showed a significant association in both the Korean population ($p=2.18\times10^{-12}$, OR 0.54, 95% CI 0.45–0.64; table 2) and the European population ($p=5.12\times10^{-3}$, OR 0.63, 95% CI 0.46–0.87; table 2). These results demonstrate that rs109592 is a shared risk variant between East Asian and European populations.

**Genetic association with clinical characteristics**

As indicated in figure 6, no significant associations were observed between the rs109592 genotype and age of onset (one-way ANOVA, $p=0.9166$) or sex (Chi-squared, $p=0.5199$). Radiological findings indicated that the NB type accounted for 83.2%, 75.6% and 60.7% of individuals with the rs109592 CC, CT and TT genotypes, respectively (figure 6). This indicates that the CC genotype was significantly associated with the NB type (Chi-squared, $p=0.0008$).
Discussion

The present study of 1066 patients with pulmonary MAC disease and 1135 controls identified an association with SNPs around CH2P in Japanese patients. We demonstrated the association of the minor T allele of rs109592 (MAF 11.8% in MAC patients) with protection from pulmonary MAC disease (OR 0.54). eQTL analysis indicated that this protective allele was associated with increased CH2P expression in the lung and increased expression of PRKCB in several tissues. In addition, the expression of CH2P was pathologically confirmed in resected lung tissue from patients with pulmonary MAC infection. Furthermore, this SNP was also found to be associated with disease risk in Korean and European populations. Taken together, these findings suggest that these genes may influence susceptibility to pulmonary MAC disease. This is the first GWAS to analyse patients with pulmonary MAC or pulmonary NTM.

Previous studies of candidate polymorphisms and host susceptibility to pulmonary NTM diseases did not apply genome-wide approaches nor involve large sample sizes [15–17, 19]. The present study did not reproduce the genetic findings reported by these studies. Other previous investigations of host susceptibility employed exome sequencing approaches and parametric linkage analysis and identified MPEG1, TTK and MST1R as individual but not population candidates [18, 25, 26]. While exome sequencing can identify low-frequency or rare coding variants, GWAS can identify associations between more common genetic variants and complex traits [27]. The present study identified genome-wide associations, indicating that host genetic factors contribute to pulmonary NTM disease.

We identified a disease risk associated with SNPs in the CH2P region. CH2P is a cofactor for plasma membrane sodium ion (Na⁺)/hydrogen ion (H+) exchangers (NHEs), which regulate ion homeostasis by catalysing the electroneutral counter-transport of Na⁺ and H⁺ [28]. Of the nine isoforms (NHE1–9), NHE1 is most widely investigated. NHE1 shows high ion transport activity under acidic conditions, which decreases as pH increases and is nearly absent at a pH of 7.5. NHE1 activity is increased upon interaction.
<table>
<thead>
<tr>
<th>SNP</th>
<th>Chromosome</th>
<th>Position (hg19)</th>
<th>Nearest gene (location)</th>
<th>RegulomeDB score (v 1.1)</th>
<th>Data</th>
<th>Minor allele</th>
<th>Major allele</th>
<th>Discovery stage (475 MAC versus 417 controls)</th>
<th>Replication stage (591 MAC versus 718 controls)</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs109592</td>
<td>16</td>
<td>23768710</td>
<td>CHP2 (intron)</td>
<td>1f</td>
<td>Imputed</td>
<td>T</td>
<td>C</td>
<td>0.127 (0.44–0.74) 1.39×10⁻⁵ 0.112 (0.41–0.63) 1.55×10⁻⁹</td>
<td>0.54 (0.45–0.63) 1.64×10⁻¹³</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>23793439</td>
<td>CHP2 (23kbp 3')</td>
<td>6</td>
<td>Genotyped</td>
<td>T</td>
<td>G</td>
<td>0.164 (0.44–0.69) 2.67×10⁻⁷ 0.157 (0.48–0.72) 1.39×10⁻⁷</td>
<td>0.58 (0.50–0.67) 3.62×10⁻¹³</td>
<td>0.118</td>
</tr>
<tr>
<td>rs7200798</td>
<td>16</td>
<td>23853860</td>
<td>PRKCB (intron)</td>
<td>1f</td>
<td>Imputed</td>
<td>G</td>
<td>A</td>
<td>0.298 (0.60–0.90) 2.75×10⁻³ 0.293 (0.67–0.93) 4.77×10⁻³</td>
<td>0.77 (0.68–0.87) 5.74×10⁻⁵</td>
<td>0.800</td>
</tr>
</tbody>
</table>

MAF: minor allele frequency.
with CHP via an intracellular CHP-binding domain. CHP has at least three isoforms (CHP1–3), and the binding affinity of CHP2 for NHE1 is reportedly five-fold more potent than that of CHP1 [29]. The critical role of ion transport in pulmonary NTM susceptibility is exemplified by the robust association of mutations in CFTR with active disease. Importantly, this association is true even for heterozygous CFTR mutation carriers, suggesting that even relatively modest changes in normal CFTR function have important late-onset effects on pulmonary NTM disease [14].

Interestingly, multi-tissue eQTL analysis identified a significant association between the rs109592 genotype and lung CHP2 expression levels in the GTEx database. This suggests that increased expression of CHP2 protects against the development of pulmonary MAC.

After finding that rs109592 within the CHP2 locus was associated with the development of pulmonary MAC infection, we further analysed the lung tissue to identify the cellular location of CHP2 by staining resected lung tissues from patients with pulmonary MAC. The results were positive in fibroblasts surrounding the granuloma and stroma in the germinal centre. Pathological findings indicated stroma in the germinal centre, suggesting a follicular dendritic cell network. Immunopathogenesis involving CHP2-expressing cells should be evaluated further.

PRKCB also displayed eQTL effects in several tissues. PRKCB is activated by calcium and the second messenger diacylglycerol [30]. This molecule is involved in B-cell signalling, apoptosis, insulin signalling, endothelial cell proliferation and autophagy, and reports of its involvement in other immune cells are increasing. This indicates that genes other than CHP2 are candidate causal genes, including PRKCB.

The present study showed that the rs109592 CC genotype was more common in patients with the NB type than the FC type (figure 6). The FC type is more common in older males with underlying diseases such as COPD and previous pulmonary tuberculosis. The NB type is more common in females with no history of smoking. The global increase in pulmonary NTM disease arises mainly from an increase in the NB type. Previous studies reported that the NB type showed a higher redevelopment rate than the FC type after standard antimicrobial treatment, although patients with the FC type had a poorer long-term prognosis [31, 32]. Interestingly, different mycobacterial strains were present at the pre- and post-treatment stages in most patients with relapsed NB-type, indicating that redevelopment of MAC in these patients reflected their intrinsic vulnerability to infection rather than a recurrence of a latent pathogen. Based on these prior

**FIGURE 4** Expression quantitative trait loci (eQTL) analysis of rs109592 and rs7200798. The relationships between the rs109592 and rs7200798 alleles and expression of the indicated genes are shown for the indicated tissues, lung, whole blood and artery-tibial. The risk allele (C) of rs109592 decreased the endogenous expression of CHP2 in the lungs, while rs109592 also showed eQTL with PRKCB in the artery tibial tissue. Data were extracted from the GTEx database (release v8).
studies, the NB subtype may be more influenced by host genetic factors, while the FC subtype may be more influenced by other factors such as smoking or previous pulmonary tuberculosis. This finding is partially consistent with those of a previous study reporting that the NB type was more influenced by host susceptibility [31].

We demonstrated that this SNP is associated with disease risk not only in the Japanese population, but also in Korean and European populations. The sharing of risk variants among different populations suggests that this SNP is robust for the development of pulmonary MAC disease. Further international collaborative GWAS using larger sample size may reveal additional candidate genes related to NTM diseases, including genes whose MAF is lower than those in this study.

There are several limitations to this study. First, our study focused on MAC and excluded other NTM pathogens. Because other pulmonary NTM diseases may display different clinical characteristics, the governing genetic factors may be distinct from those influencing the risk for pulmonary MAC infections. *Mycobacterium abscessus* is a rapid-growing mycobacterium that often causes refractory lung disease.

![FIGURE 5](Image)

**FIGURE 5** Immunostaining of resected lung tissue from pulmonary *Mycobacterium avium* complex patients. a and b) CHP2 was positive with fibroblasts surrounding the granuloma; c and d) CHP2 was positive and showed a reticular pattern in the germinal centre of the tertiary lymphoid structures. Scale bars a and c) 500 μm, b and d) 100 μm.

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Associations with rs109592 in Korean and European patients with <em>Mycobacterium avium</em> complex (MAC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAF (cases)</td>
<td>MAF (controls)</td>
</tr>
<tr>
<td>Korean (722 MAC versus 1722 controls)</td>
<td>0.127</td>
</tr>
<tr>
<td>European (276 MAC versus 206 controls)</td>
<td>0.161</td>
</tr>
</tbody>
</table>

MAF: minor allele frequency.

https://doi.org/10.1183/13993003.02269-2019
Pulmonary Mycobacterium kansasii, which is more common in males and smokers, can be treated with classic anti-tuberculous medicines. When we included the other NTM populations in our analysis, the candidate SNPs displayed the same statistical tendencies (supplementary figure S5 and table S4). The use of a larger sample size may enable stratified analysis of NTM species. Second, a single causal variant at chromosome 16p21 locus remains unclear in our GWAS. Although rs109592 showed the most significant association with pulmonary MAC disease, after combining the discovery and replication datasets, rs109592 was not included in the 95% credible variant set to explain the association at chromosome 16p21 locus under the fine-mapping of the discovery dataset. Related to this causal variant ambiguity, our GWAS and eQTL signals for CHP2 and PRKCB in GTEx project data did not support a shared causal variant in the colocalisation analysis. These are possibly due to the weakness of our GWAS, in that no SNP showed a highly significant genome-wide association (p<5×10^{-8}) in our discovery dataset or error using a single causal variant assumption. Future genome-wide analyses with increased sample sizes are necessary to conclude that the causal variant is really at chromosome 16p21 and that CHP2 and/or other surrounding genes regulate susceptibility to pulmonary MAC disease. The associations of CHP2 expression levels in resected lung tissues based on rs109592 genotype are intriguing. However, we were unable to analyse this difference due to inadequate sample size. We plan to analyse it in our next project by increasing sample size.

In conclusion, we performed GWAS for pulmonary MAC susceptibility, which identified rs109592 at the CHP2 locus as a susceptibility marker. The minor T allele of rs109592 was associated with protection from pulmonary MAC disease and increased expression of CHP2 in lung by eQTL analysis. These findings may provide novel insight into pulmonary NTM disease and may help form the basis for future collaborative research into the pathogenesis of this condition.

Acknowledgements: We thank all the participants in this study. We thank Sachika Wada, Shoko Takahashi and Miyuki Yamamoto (Keio University, Tokyo, Japan), and Kayoko Yamada, Kayoko Kato and Natsumi Baba (University of Tokyo, Tokyo, Japan) for providing technical assistance.

Availability of data and materials: All raw data are available by request to the corresponding author.


Conflict of interest: H. Namkoong reports grants from JSPS (JSPS KAKENHI grant number 15H06590), AMED (2019 US-Japan Collaborative Proposal 28863) and NIAID/CBER Intramural Tuberculosis Research Initiative (FY19 Start-up support from CAN 8020315), during the conduct of the study. Y. Omae reports grants from JSPS KAKENHI (18K15041), during the conduct of the study. T. Asakura reports grants from JSPS, outside the submitted work.
M. Ishi reports grants from JSPS KAKENHI (18H02821, 18K19566), outside the submitted work. S. Suzuki reports grants from JSPS, outside the submitted work. K. Morimoto reports grants from AMED (2019 US-Japan Collaborative Proposal 28863), during the conduct of the study; personal fees for consultancy from INSmed, outside the submitted work. K. Kawai has nothing to disclose. K. Emoto has nothing to disclose. A.J. Oler has nothing to disclose. E.P. Seymanski has nothing to disclose. M. Yoshida reports grants from JSPS (grant-in-aid for young scientists (B), JSPS KAKENHI grant number 17K16066), during the conduct of the study. S. Matsuda has nothing to disclose. K. Yagi has nothing to disclose. I. Hase has nothing to disclose. T. Nishimura reports grants from JSPS KAKENHI (16K09942), outside the submitted work. Y. Sasaki reports grants from JSPS (JSPS KAKENHI grant number 15H056590), during the conduct of the study. T. Asami has nothing to disclose. T. Shiomi has nothing to disclose. H. Matsubara has nothing to disclose. H. Shimada has nothing to disclose. J. Hamamoto reports grants from JSPS KAKENHI (19K08610), outside the submitted work. B.W. Jhun reports grants from AMED (2019 US-Japan Collaborative Proposal 28863), during the conduct of the study. S.-Y. Kim has nothing to disclose. H.-J. Huh has nothing to disclose. H.-H. Won has nothing to disclose. M. Ato reports grants from Japan Agency for Medical Research and Development, during the conduct of the study. K. Kosaki reports grants from Japan Agency for Medical Research and Development (AMED), outside the submitted work. T. Betsuyaku reports grants from JSPS KAKENHI and Japan Agency for Medical Research and Development (AMED), outside the submitted work. K. Fukunaga reports grants from JSPS KAKENHI and Japan Agency for Medical Research and Development (AMED), outside the submitted work. A. Kurashima reports personal fees for consultancy from Insmed, during the conduct of the study. H. Tettelin reports grants from NIAID (contract number HHSN27220090009C/JSPS), during the conduct of the study. H. Yanai has nothing to disclose. S. Mahasirimongkol has nothing to disclose. K.N. Olivier reports non-financial support from AIT Therapeutics, non-financial support from Insmed, outside the submitted work. Y. Hoshino reports grants from The Japan Agency for Medical Research and Development/Japan International Cooperation Agency (AMED), during the conduct of the study (JP18K0108043 and JP18K0108075). W.-J. Koh reports grants from National Research Foundation of Korea (grant number NRF-2018R1A2A1A05018309), during the conduct of the study. S.M. Holland has nothing to disclose. K. Tokunaga reports grants from Japan Agency for Medical Research and Development (grant-in-aid for research JP17km0405205h0002), during the conduct of the study. N. Hasegawa reports grants from Japan Agency for Medical Research and Development (AMED), during the submitted work; grants from JSPS KAKENHI, personal fees for consultancy from INSmed, outside the submitted work.

Support statement: This work was supported by Japan Agency for Medical Research and Development (AMED) (project numbers JP17FK0108116H0401, JP17KM0405205H0002, JP18FK0108043, JP18FK0108075, 18km0405205h0003, 19km0405205h9904, 19k0205012h0004, US–Japan Cooperative Medical Sciences Program Collaborative Awards 2019), Keio Gijuku Academic Development Funds, grant-in-aid for research activity start-up (JSPS KAKENHI) (grant number 15H06590), grant-in-aid for young scientists (B) (JSPS KAKENHI) (grant numbers 15K19039, 17K16066), grant-in-aid for scientific research (B) (grant numbers 19H03704, 15H05271, 24406010), and National Research Foundation of Korea (NRF) funded by the Korea government (MSIT) (grant number NRF-2018R1A2A1A05018309). This work was supported in part by federal funds to the Institute for Genome Sciences from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under contract number HHSN27220090009C/JSPS. This work was supported by NIAID/CBER Intramural Tuberculosis Research Initiative and by the intramural programs of the National Institute of Allergy and Infectious Diseases and the National Heart, Lung, and Blood Institute, National Institutes of Health. Funding information for this article has been deposited with the Crossref Funder Registry.

References