Research Articles

Unique genomic sequences in novel *Mycobacterium avium* subsp. *hominissuis* lineage enabled fine scale tracing the transmission route during pig movement

Tetsuya Komatsu[a¶], Kenji Ohya[b,c,d¶*a], Atsushi Ota[e], Yukiko Nishiuchi[f], Hirokazu Yano[g], Kayoko Matsuo[h], Justice Opare Odoi[c], Shota Suganuma[b*b], Kotaro Sawai[b,i], Akemi Hasebe[j], Tetsuo Asai[c,d], Tokuma Yanai[b,c,k], Hideto Fukushi[b,c], Takayuki Wada[l], Shiomi Yoshida[m], Toshihiro Ito[n], Kentaro Arikawa[o], Mikihiko Kawai[p], Manabu Ato[q], Anthony D. Baughn[r], Tomotada Iwamoto[o], Fumito Maruyama[f,s,t#]

[a]Aichi Prefectural Chuo Livestock Hygiene Service Center, Okazaki, Aichi, Japan.

[b]Faculty of Applied Biological Sciences, Gifu University, Gifu, Gifu, Japan.

[c]United Graduate School of Veterinary Sciences, Gifu University, Gifu, Gifu, Japan.

[d]Education and Research Center for Food Animal Health, Gifu University (GeFAH), Gifu, Gifu, Japan.

[e]Data Science Center, Division of Biological Science, Nara Institute of Science and Technology, Ikoma, Nara, Japan.

[f]Office of Academic Research and Industry-Government Collaboration, Hiroshima University, Higashi-Hiroshima, Hiroshima, Japan.

[g]Graduate School of Life Sciences, Tohoku University, Sendai, Miyagi, Japan.

[h]Kumamoto Prefectural Aso Public Health Center, Aso, Kumamoto, Japan.

[l]Division of Transboundary Animal Disease Research, National Institute of Animal Health, National Agriculture Research Organization, Tsukuba, Ibaraki, Japan.

[j]Toyama Prefectural Meat Inspection Center, Imizu, Toyama, Japan.

[k]Hiwa Natural History Museum, Shobara, Hiroshima, Japan.

[l]Graduate School of Human Life Science, Osaka City University, Osaka, Osaka, Japan.

[m]Clinical Research Center, National Hospital Organization Kinki-Chuo Chest Medical Center, Sakai, Osaka, Japan.

[n]Laboratory of Proteome Research, Proteome Research Center, National Institutes of Biomedical Innovation, Health and Nutrition, Ibaraki, Osaka, Japan.

[o]Department of Infectious Diseases, Kobe Institute of Health, Kobe, Hyogo, Japan

[p]Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Kyoto, Japan

[q]Department of Mycobacteriology, Leprosy Research Center, National Institute of Infectious Diseases, Higashimurayama, Tokyo, Japan

[r]Department of Microbiology and Immunology, University of Minnesota Medical School, Minneapolis, Minnesota, USA

[s]Project Research Center for Holobiome and Built Environment (CHOBE), Hiroshima University, Higashi-Hiroshima, Hiroshima, Japan

[t]Scientific and Technological Bioresource Nucleus, Universidad de La Frontera, Temuco, Chile.

[*a]Current Address: National Institute of Health Sciences, Kawasaki, Kanagawa, Japan.

[*b]Current Address: Central Research Institute for Feed and Livestock of Zen-noh, Tsukuba, Ibaraki, Japan.

[#]Corresponding author
fumito@hiroshima-u.ac.jp (FM)

[¶]These authors contributed equally to this work.

The order of authors: We determined the order of co-authors according to the magnitude of their contribution to the study. Of these, the two authors who made particular contributions to the design of this study, analyses, interpretation and discussion of the results, and preparation of the paper were designated as the first authors according to the magnitude of their contributions.

Running title: NOVEL MAH LINEAGE CAN TRACE THE TRANSMISSION ROUTE

Abstract: 245 words, Importance: 150 words, Manuscript: 4,162 words.

Abstract

Mycobacterium avium subsp. *hominissuis* (MAH) is one of the most prevalent mycobacteria causing non-tuberculous mycobacterial (NTM) disease in human and animals. Of note, MAH is a major cause of mycobacterial granulomatous mesenteric lymphadenitis outbreaks in pig populations. To determine the precise source of infection of MAH in a pig farm and to clarify the epidemiological relationship among pig and human MAH lineages, we collected 50 MAH isolates from pigs reared in Japan and

determined draft genome sequences of 30 isolates. A variable number of tandem repeat (VNTR) analysis revealed that most pig MAH isolates in Japan were closely related to North American, European and Russian human isolates but not to those from East Asian human. Historical recombination analysis revealed that most pig isolates could be classified into SC2/4 and SC3. To our surprise, four isolates belonged to a new lineage (SC5) in the global MAH population. Members of SC5 had few footprints of inter-lineage recombination in the genome, and carried 80 unique genes, most of which were located on lineage specific-genomic islands. As its unique genetic features, we were able to trace the putative transmission route via their host pigs. Together, these results reveal the diversity in the pig MAH population and clarify the possibility of species-specificity in MAH in addition to local adaptation. Moreover, our study also clarified the difference of source of infection and the vector of transmission between humans and pigs and provides a clear countermeasure to prevent MAH spreading in pigs.

Importance

Mycobacterium avium subsp. *hominissuis* (MAH) is a dominant cause of non-tuberculous mycobacterial (NTM) disease in human and animals. In this study, in order to clarify the epidemiological relatedness between MAH isolates from human and pigs, and the MAH transmission route among pigs, we conducted molecular and genomic-based epidemiological research. The pig isolates were genetically distant from those of humans in Japan, however, they were closely related to human isolates from North America, Europe and Russia, suggesting a proportion of pig MAH isolates are imported from abroad. Importantly, we identified a novel lineage of MAH that was distributed in a

specific area in Japan. Epidemiological survey and its unique genomic features allowed us to trace putative transmission routes via pig movement. Our study provides critical information for understanding the diversity of MAH population and the possibility of species-specificity in MAH. Our study will be useful for countermeasures of pig NTM disease.

Introduction

Mycobacterium avium complex (MAC) is the most common causative agent of non-tuberculous mycobacterial disease (NTM) in humans in Asia and Europe (1, 2). The frequency of pulmonary MAC infection is increasing worldwide (3). *Mycobacterium avium* subsp. *hominissuis* (MAH) is the most prevalent subspecies of MAC clinical isolates in Asia and Europe (1, 2). MAH causes pulmonary disease (4) and sometimes causes systemic infection in immunocompromised people, such as those infected with human immunodeficiency virus (HIV) (5).

MAH also causes lymphadenitis most often involving mesenteric lymph nodes in pigs (6). MAH sometimes causes systemic infection in pigs (7). Pigs infected with MAH usually have no clinical signs and the lesions are incidentally found during slaughter (8), bringing great financial losses in pig farm management in the event of an outbreak (1).

It is thought that environmental materials in pig farms such as sawdust (9) and peat (10) are sources of MAH infection in pigs, suggesting that the environmental materials are the most important factor in the spread of MAH (11). In a previous study, use of IS1245 restriction fragment length polymorphism (RFLP) analysis indicated that the pig flow and the mother sow are important sources of MAH infection (12). However,

the MAH infection by pig flow and sow to piglet transmission in that study were anecdotal events within a single farm and it is possible that employees and mobile farm equipment could also serve as sources of MAH exposure. In addition, RFLP analysis is inferior to a variable number of tandem repeat (VNTR) analysis (13) for MAH typing and for more precise determination of source of MAH infection, molecular and genomic epidemiological study using the isolates from several farms and environments is essential.

Pigs are thought to be a potential source of exposure for human MAH infection (14, 15). Recently, several molecular epidemiological studies of MAH involving VNTR analysis have been described from around the world (6, 16, 17, 18, 19, 20). Some reports showed epidemiological relationships of MAH strains between humans and pigs (6, 20, 21, 22). From these studies, it was suggested that MAH would have region-specific genetic diversity and distribution. In East Asian countries, human MAH strains isolated from patients with lung disease were not identical to MAH isolates of pigs (6, 20). On the contrary, MAH strains of pigs were identical to those of humans from Europe (18, 22). However, MAH strains of pigs used in these studies were isolated from limited regions (Kinki - Western region of Honshu Island, Hokkaido, Okinawa), that harbor about 9.8% of farmed pigs in Japan. VNTR analysis of MAH isolates from pigs in Kyushu area in Japan, where 31.2% of pigs in Japan are reared, was also reported (23). However, this study was limited due to use of only 8 loci for VNTR, which was thought to be less useful than 16 loci used in other analyses (13). To circumvent these shortcomings, MAH isolates from additional regions are needed to provide a more robust view of MAH distribution.

Phylogenetic and population structure analysis based on MAH genomes has recently revealed that MAH can be divided into six major clades, including EastAsia1 (EA1), EastAsia2 (EA2) and SC1 - 4, suggesting that MAH adaptation is geographically restricted around the world (24, 25). For example, EA1 and EA2 lineages were mainly distributed in humans in Japan and Korea although SC1 - 4 lineages were distributed in Europe and USA (24, 25). Although these studies utilized many MAH genomes, only a few pig isolates were included. Therefore, it is still unknown what lineages are distributed in pigs in Japan, and highlights a major knowledge gap in understanding likely zoonotic sources of MAH infection in humans.

To understand the diversity of MAH and its potential risk for humans, it is essential to characterize the distribution of MAH lineages in farmed pigs in Japan and their genetic relationship with human MAH isolates. In addition, it is also critical to conduct detailed epidemiological study at a genomic level for the determination of transmission route of MAH for pigs. In this study, we collected 50 MAH of pig isolates from Tokai and Hokuriku area, which produce about 10% of pigs in Japan. Initially, we determined the genetic relatedness of human and pig isolates via VNTR and the genomic features of major and minor MAH lineages distributed in pig population in Japan. Next, we characterized genomic and epidemiological features of a novel MAH lineage distributed in a limited geographic area. The unique genes of these isolates may be useful to trace MAH transmission routes via pig movement.

Results

Insertion Sequence (IS) and SNP analysis of *hsp65*.

To investigate molecular features of pig MAH, we compared *hsp65* gene sequences and IS patterns with those of some reference strains (Table S1). Among MAH isolates, IS*1245* was absent in 14/50 isolates (28.0%). No isolates had IS*900*, IS*901* and IS*902*. Four isolates (GM10, GM12, GM21 and GM32), which formed a distinct new lineage as described in the following section, had a conserved SNP pattern in the *hsp65* region that was unique to this lineage (Figure S1). These isolates had two characteristic SNPs, G486A and G828T. GM16 also had a characteristic SNP, G441A.

Mycobacterial Interspersed Repetitive Unit (MIRU) and *Mycobacterium avium* Tandem Repeat (MATR) VNTR.

Next, to investigate molecular epidemiological relationship between human and pig MAH strains isolated from the world and Japan, VNTR analysis was performed and two types of MST trees were generated using 7 loci for comparison with domestic MAH isolates and 13 loci for those of global isolates.

VNTR based on 7 loci. MST based on 7 loci is shown in Figure 1. From VNTR analysis, MAH isolates were classified into two groups. Group 1 contained mainly the human isolates, although almost all isolates derived from pigs in Japan were classified in Group 2. Some isolates in this study were classified in Group 1, however, no isolate was clustered with human isolates. The pig isolates also formed a cluster with environmental isolates in Group 2. The Isolates of the new lineage described above formed a distinct subgroup with Group 2.

VNTR based on 13 loci. MST based on 13 loci is shown in Figure 2. Isolates were divided into six clonal complexes. Most isolates of human origin from Japan and Korea were classified into complex 1. Complex 2 was composed of MAH isolates from

USA, Europe, Russia and some human derived isolates from Korea and pig isolates from Japan. Complex 4 and 5 were mainly composed of pig isolates from Japan, although, complex 3 contained human and pig isolates from Japan. The isolates of the newly described lineage were classified into complex 6 and were genetically far from any adjacent node.

Phylogenetic tree analysis.

To clarify the phylogenetic features of pig MAH at the genome level, phylogenetic analysis was performed using draft genome sequences of pig isolates, reference strains described by Yano et al (25) and *Mycobacterium avium* subsp. *paratuberculosis* (MAP) strain K-10 as an outgroup. Phylogenetic trees generated by the neighbor joining (NJ) tree method and maximum likelihood (ML) tree method via CSIphylogeny (26) and IQ-TREE (27) are shown in Figure 3, Figure S2a and Figure S2b, respectively. As a result of topological evaluation of both phylogenetic trees via IQ-TREE, phylogenetic tree of CSI phylogeny considered to be more appropriate than that of IQ-TREE. From the phylogenetic tree based on draft genome sequences of 30 MAH isolates, 24 isolates were classified into 3 main lineages (SC2: 13, SC3: 1, SC4: 10), and 4 isolates were classified into a new lineage (SC5) (Figure 3). This new lineage was the earliest branch from the MAP (Figure 3, Figure S2a, Figure S2b). The isolates GM16 and OCU484 were unclassified and placed between EA2 and SC3 (Figure 3, Figure S2a), however, they were classified in EA2 by IQ-TREE (Figure S2b).

SNP analysis for *cinA* and *sugA* genes.

SNP analysis for *cinA* and *sugA* genes, which are considered as epidemiological markers to distinguish the six major lineages of MAH (25), are shown in Figure 4 and

Figure 5. The new lineage shared a unique SNP pattern in both genes compared with the other lineage (Figure 4 and Figure 5). The isolate GM5 had a different SNP pattern from SC4 in *cinA* (Figure 4). GM16 had the same SNP patterns as EA1 in *cinA* (Figure 4). GM16 and OCU484 had unique SNP patterns in *sugA*, respectively (Figure 5). Interestingly, SNP in the *sugA* gene of the isolates of lineage SC2 were divided into two patterns, SC2a and SC2b. SNP patterns of SC2a were of a new type and had a characteristic SNP, C25T. The patterns of SC2b were the same as *Mycobacterium avium* subsp *avium* (MAA) (Figure 5).

Bayesian analysis of population structure (BAPS) analysis and recombination analysis for lineage classification.

In order to predict precise MAH lineage, we conducted BAPS analysis and recombination analysis. From population structure analysis using fastBAPs (28), MAH isolates in this study were divided into 4 lineages, SC5, SC1, SC2 and large SC (Figure S3). The occurrence of inter-lineage recombination events assessed by fastGEAR (29) is shown in Figure 6. The strains and isolates were divided into 5 lineages (SC5, EastAsia, SC2/4, and SC3). Each of the pig isolates was classified into three lineages (SC5: 4, SC2/4: 25, SC3: 1). SC5 had few recombination events with any other lineages. Similarly, some isolates of lineage SC2/4, which were classified into SC2a by SNP analysis, also had few occurrences of recombination among lineages.

Pan and core genome analysis.

To search the unique genes of new lineage MAH, pan and core genome analysis were conducted using Roary software (30). From results of pan and core genome analysis, core and accessory gene content were 2,314 and 12,115, respectively. Six

core genes were absent only in the SC5 and 97 genes were characteristic in this lineage (Table S2). From Basic Local Alignment Search Tool (BLAST) analysis, 80 genes of the SC5 had less than 95% coverage and similarity with the other MAH isolates. Most of the genes unique to the SC5 were clustered and were located on some genomic islands (GIs), which were not previously reported (31) (Table S2). Clustering analysis based on accessory genes showed that SC5 made one cluster (Figure S4).

Detection of mobile elements (Plasmid, phage and GI) in MAH genome.

In order to search for mobile elements of pig isolates, plasmid, phage region and GI were identified from draft genome sequences via online programs (32-36). No plasmids were detected in any isolates via Plasmidfinder (32), however, *avium*MD30_plasmid_pMD2-like sequence was detected in 9/30 isolates of draft genome sequences via JSpecies (33). Interestingly, all isolates of the new lineage had these plasmid-like sequences. Intact and questionable phage regions were also not detected, although incomplete sequences of phage were detected in all isolates (34-35). GIs were detected in all isolates (36), however, there was diversity in the number of GIs (Table 1).

Farm relationship where the novel lineage MAH was isolated.

GM10, GM12, GM21 and GM32 were isolated from pigs reared in three farms located in Gifu, Shiga and Aichi Prefecture (Table S1). From interviewing farm owners, these 3 farms were related epidemiologically (Figure 7). Aichi B and Gifu A farms had introduced boars and sows from Aichi Y farm. Shiga A farm had introduced fattening pigs from Aichi X farm, whose sows were introduced from Aichi Y farm.

Discussion

In the present study, we revealed two critical points, first was the genetic relationship among MAH isolates of pigs, humans and the environment in Japan, and the second was the identification of a new MAH lineage and its genomic and epidemiological features.

Genetic relationship of pig isolates in Japan

From VNTR analysis using both 7 and13 loci, most isolates in this study were grouped with those from pigs and the environment in Japan and humans in Europe, Russia and USA (Figure 2: Complex 2, 4), supporting previous studies (6, 16, 18, 20). In Japan, sow farms in various regions supply sows to each region. Nevertheless, pig MAH isolates from various regional origins are closely genetically related, except for the new lineage. This observation indicates that genetically close MAHs are distributed among Japanese pigs.

In addition, by phylogenetic tree analysis based on draft genome sequences, 30 isolates were classified into SC2 (13 isolates), SC4 (10), SC3 (1), unclassified (2) and a SC5, as a new lineage (4) (24, 25). These results are consistent with previous reports (24, 25), but it is worth noting that SC3 was isolated from pigs in Japan for the first time (Table S3). Recombination analysis clarified two previously unclassified isolates GM16 and OCU484 that were classified into SC2/4. GM16 and OCU484 had many recombination events with EastAsia. In addition, these isolates were classified into EA1 and SC4 by SNP analysis of *cinA* gene, respectively. Frequent recombination events and multiple SNP in recombination-cold regions may contribute to the result of these isolates being unclassified in genome-based phylogenetic tree analysis.

On the other hand, SNP analysis reveals there is diversity in pig isolates in Japan at the level of a single gene, such as *sugA*. Our SC2/4 isolates were divided into 5 patterns by phylogenetic analysis of *sugA*, SC4 (10 isolates), SC2a (7), SC2b (6) and two unclassified (2). These isolates had been isolated from multiple farms across the prefectures. Focusing on SC2a, these isolates had a unique SNP pattern in their core gene and had few recombination events with the other lineages. This indicates that SC2a has evolved in a species- or region-specific manner in Japan.

In lineage SC4, most isolates had the same SNP pattern as the reference strain and had relatively frequent recombination events among lineages, consistent with a past study (25). SC4 was isolated from human and the environment in Europe, USA and Russia (25) (Table S3) and Japanese pig isolates are genetically close to those of environmental isolates in these areas (16). According to a recent annual report from the Animal Quarantine Service in Japan, most live pigs that are imported to Japan originate from Canada (43.7%), America (33.4%) and Denmark (12.9%) in 2019 (https://www.maff.go.jp/aqs/tokei/toukeinen.html). These observations suggest that SC4 may be transmitted from overseas countries by a pig movement. Our study also reveals that phylogenetic analysis at the level of genome is useful to correctly understand the ecology of human, pig and environmental MAH. It is also though that simple phylogenetic classification using *cinA* and *sugA* genes would be useful for epidemiological survey, even for MAH in pigs.

Genomic and epidemiological features of the novel lineage

Four isolates (GM10, GM12, GM21, GM32) had the same characteristic SNP patterns in the *hsp65*, *cinA* and *sugA* genes and possessed related IS patterns. In

addition, these isolates also had the same VNTR profiles and were present in a very unique place in MST trees. These results including genome based phylogenetic tree, BAPS and recombination analysis indicated that this cluster was the novel lineage (SC5). SC5 had 80 unique genes in their genomes, which were not observed in the other lineages (Table S2). These genes were clustered and most of them were on what is predicted to be a GI. In addition, these genes contained mobile element associated genes, such as IS*481* family transposase IS*3514* and putative prophage phiRv2 integrase, although any plasmids and prophage regions were not detected via finder programs. SC5 had few recombination events with the other lineages. These results suggest that this new lineage can be partly characterized by mobile elements rather than recombination events (24, 25). These unique genes can serve as candidate markers for molecular epidemiological study of SC5.

It was unknown why this new lineage was isolated to the Tokai area in Japan. In general, the *Mycobacterium* genus is mainly present in the environment, such as in water sources, soil and dust (37, 38), however, our results of VNTR revealed that SC5 is genetically distinct from the other MAH isolates including environmental, human and pig isolates from Japan, East Asia, Europe, Russia and USA. Based on this observation, it is unlikely that SC5 was introduced from these regions. It is reported that MAH is isolated from wild boars in Europe (39), however, MAH is not isolated from wild animals in the same area of this study (40). Epidemiological information from farms harboring infected pigs revealed that the source of SC5 was likely introduced by pig movement. Specifically, Aichi B and Gifu A farms had directly introduced pigs from the same farm (Aichi Y), and Shiga A farm had contact through Aichi X farm, which obtained pigs from

Aichi Y farm. There was nothing in common about using animal housing materials, such as sawdust, among farms. Further, the other farms including in the Hokuriku area in this study had not introduced pigs from the Tokai area. These facts suggest that SC5 is spreading via pig movement. Based on our study, the transfer of bacteria with the movement of infected pigs seems to be the most plausible model. This hypothesis is also supported by the observation that there were few inter-lineage recombination events in SC5, because it is thought that acquisition of genetic variation through recombination (24) or plasmid transfer (41) occurs more frequently in the environment than in living hosts. Our results also indicate that SC5 has species-specificity for pigs in addition to geographical local adaptation, suggesting the possibility of adaptation for specific environment including living host and discovery of additional lineages. Our study highlights that the elimination of subclinically infected pigs with MAH can prevent MAH invasion farms and outbreaks.

In conclusion, we have clarified the relationship among MAH isolates of pigs and humans in Japan and the world using draft genome sequences and VNTR analysis. We have also identified a new lineage of MAH distributed only in the Tokai area in Japan and based on its genomic features, we were able to identify the route of infection for this MAH as a pig movement. This one-health approach revealed some novel aspects of dynamics of MAH, which can be useful for implementing countermeasures of pig and human NTM disease.

Materials and methods

Bacterial isolation, Insertion Sequence and *hsp65* profile and SNP analysis. Bacterial isolation and identification of MAH isolates of 30 isolates (GM5, GM10, GM12, GM16, GM17, GM21, GM32, GM44, OCU467 – OCU486 and Toy194 – Toy195) were referred in our past study (42). 20 isolates used for VNTR analysis (gifu-1 – gifu-92, GM6, GM11, GM23 – GM25) were isolates from mesenteric lymph nodes, feces, sawdust and farm soil (Table S1). Isolation and identification of these isolates including determination of IS profiles and *hsp* code were conducted by our study (42). All of them were isolated from pigs or environment in Tokai area (Gifu, Aichi and Shiga Prefecture) and Hokuriku area (Toyama and Ishikawa Prefecture), whose geographical relationship was showed in Figure 8. The sequences of *hsp65* gene were referred from repository related to our past study (43) with genes retrieved from genomes of representative lineages (OCU404: MahEastAsia1, TH135: MahEastAsia2, A5: SC1, 104: SC2, 100: SC3, OCU491: SC4) and were aligned with Molecular Evolutionary Genetics Analysis (MEGA) and were used for SNP analysis.

MIRU and MATR VNTR. Primer sets for 15 loci of MATR (MATR1-9, 11-16) and 7 loci selected from MIRU loci (32, 292, X3, 25, 7, 10, 47) were used to get VNTR profiles, as previously reported (20). Three loci were overlapped between MATR-VNTR (MATR-2, MATR-3 MATR-9) and MIRU-VNTR (MIRU-292, MIRU-X3, MIRU-10). From all data, we selected 13 MATR (MATR1-8, 11-13, 15-16) and 7 MIRU profiles for VNTR analysis to compare with variable data of past reports: Japanese patients and Japanese bathroom of healthy volunteers (6, 16, 18, 20), Korea patients (18, 44), European patients (18, 45), Japanese pig (6, 16, 18, 20, 23) and the Japanese natural environment (16, 20). The data profiles 50 MAH isolates in this study were available in Table S4. Detail number of VNTR

profiles using in this study were showed in Table S5. MSTs based on 7 loci and 13 loci using above data set and data in this study were generated using Bionumerics version 7.0 (Applied Maths, Sint-Martens-Latem, Belgium) for clustering analysis. Clonal complex in MST of 13 loci was defined as a different group when there were two or more differences in the VNTR profile of adjacent nodes according to the past study with one modify, that a clonal complex was composed when it had over five genotypes (15).

Draft genome sequence. Draft genome sequences of 30 isolates were obtained according to our past study (42).

Genome phylogenetic tree and SNP analysis of *cinA* and *sugA* gene. Genome phylogenetic tree analysis based on 30 MAH isolates in this study and 20 reference strains were performed via online tool CSI Phylogeny version 1.4 (https://cge.cbs.dtu.dk/services/CSIPhylogeny/) with default parameters (26) and IQ-TREE (27). Obtained newick file was visualized by MEGA according to our past study (42). Alignment file generated by CSI Phylogeny was used for phylogenetic analysis by NJ method with p-distance method via MEGA. Topological evaluation of obtained newick file was conducted by IQ-TREE (27). All of the reference genome sequences were retrieved from National Center for Biotechnology Information (NCBI) genome database. The sequence of two marker genes, *cinA* (product name: cytochrome P-450) and *sugA* (product name: trehalose ABC transporter) of 30 MAH located in recombination cold region, which are offered to be able to distinguish major six MAH lineages (25), were retrieved from draft genome sequences. These sequences in this study and the reference strains shown in the past report (25) were aligned and used for

the phylogenetic tree analysis by ML method via MEGA according to our past study (42).

BAPS and recombination analysis for determination of MAH lineage. MAH lineage in this study was predicted by using fastBAPS software (28) and fastGEAR (29), which was able to infer the MAH lineages, considering genetic linkage (25, 28). Used genome sequences were the same as using for followed phylogenetic tree analysis except for MAP. Filtered polymorphic sites in vcf file generated by CSI Phylogeny were used for fastBAPS (28). In addition, these sites were combined with intervening reference genome sequences TH135, and were organized in an alignment file (25) and then used as an input for fastGEAR (29). The lineage is generated by a result of hierarchical clustering (29). We adopted the result of fastGEAR as the lineage classification in this study because it was the most consistent with previous reports (24, 25).

Pan and core genome analysis. The pan and core genome analysis was conducted using Roary software (30). Used genomes sequences were the same as BAPS and recombination analysis. Draft genome sequences in this study and references were re-annotated to get gff files by Prokka v1.14.5 software (46). Genes which were present or absent only in new lineage (GM10, GM12, GM21, GM32) were searched by Roary software. BLAST analysis was performed with the genes present only in GM10, GM12, GM21 and GM32 and marked the genes which had high coverage and similarity (over 95%) with genes of MAH in NCBI gene database. We also researched whether these genes were located on GI and/or phage regions or not in next paragraph. Hierarchical clustering analysis based on the present or absent accessory genes in MAH were also

performed by Cluster 3.0 and visualized by TreeView 3.0 (https://zenodo.org/record/1303402#.X4cME-ZxeUk).

Detection of mobile element (Plasmid, phage and GI) in MAH genome. Plasmid was identified from draft genome sequences by PlasmidFinder version 2.1 (https://cge.cbs.dtu.dk/services/PlasmidFinder/) with following setting (threshold for minimum % identity: 50%, minimum % coverage: 20%) (32). In addition, we retrieved MAH plasmid sequences which were more than 2 kbp and used them for pairwise genome comparisons with draft genome sequences in this study by MUMmer program via JSpecies (33). Prophage sequences were identified by PHASTER (PHAge Search Tool Enhanced Release) (https://phaster.ca) (34, 35). GIs were predicted by IslandViewer version 4 (36).

Epidemiological information of farms. Epidemiological information of each farm condition and pig movement was obtained via interviewing farm owners.

Data availability.

The draft genome sequence data used in this study have been submitted to the NCBI and all of accession number was available in our past study (42).

Acknowledgments

We thank the member of Gifu Prefectural Chuo Livestock Hygiene Service Center for sampling. We thank the Data Integration and Analysis Facility at the National Institute for Basic Biology for providing some of the computational resources. All authors reported no potential conflicts of interest. This work was supported by a grant from the Japan Agency for Medical Research and Development (AMED) (17fk0108116h040,

21fk0108129h0502), the Japan Racing Association (JRA) Livestock Industry Promotion Project (H28-29_239, H29-30_7) of the JRA, a grant for Meat and Meat Products (H28-130, H30-60) managed by the Ito Foundation for Research in design study, collection, analysis; and was supported by grants from the Japan Society for the Promotion of Science (JSPS) KAKENHI (JP26304039, JP18K19674, 16H05501, 16H01782, 20H00562) and the Joint Research Program of the Research Center for Zoonosis Control in Hokkaido University. JOO is a recipient of a Japanese government (Ministry of Education, Culture, Sports, Science and Technology) scholarship. All funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

References

1. Prevots DR, Marras TK. 2015. Epidemiology of human pulmonary infection with nontuberculous mycobacteria: a review. Clin Chest Med 36:13-34.
2. Namkoong H, Kurashima A, Morimoto K, Hoshino Y, Hasegawa N, Ato M, Mitarai S. 2016. Epidemiology of pulmonary nontuberculous mycobacterial disease, Japan. Emerg Infect Dis 22:1116-1117.
3. Daley CL. 2017. *Mycobacterium avium* Complex Disease. Microbiol Spectr 5.
4. Tran QT, Han XY. 2014. Subspecies identification and significance of 257 clinical strains of *Mycobacterium avium*. J Clin Microbiol 52:1201-1206.
5. Uchiya KI, Asahi S, Futamura K, Hamaura H, Nakagawa T, Nikai T, Ogawa K. 2018. Antibiotic susceptibility and genotyping of *Mycobacterium avium* strains that cause pulmonary and disseminated infection. Antimicrob Agents Chemother 62:e02035-17.

6. Adachi T, Ichikawa K, Inagaki T, Moriyama M, Nakagawa T, Ogawa K, Hasegawa Y, Yagi T. 2016. Molecular typing and genetic characterization of *Mycobacterium avium* subsp. *hominissuis* isolates from humans and swine in Japan. J Med Microbiol 65:1289-1295.
7. Agdestein A, Johansen TB, Kolbjørnsen Ø, Jørgensen A, Djønne B, Olsen I. 2012. A comparative study of *Mycobacterium avium* subsp. avium and *Mycobacterium avium* subsp. *hominissuis* in experimentally infected pigs. BMC Vet Res 8:11.
8. Thorel MF, Huchzermeyer HF, Michel AL. 2001. *Mycobacterium avium* and *Mycobacterium intracellulare* infection in mammals. Rev Sci Tech 20:204-218.
9. Matlova L, Dvorska L, Palecek K, Maurenc L, Bartos M, Pavlik I. 2004. Impact of sawdust and wood shavings in bedding on pig tuberculous lesions in lymph nodes, and IS1245 RFLP analysis of *Mycobacterium avium* subsp. *hominissuis* of serotypes 6 and 8 isolated from pigs and environment. Vet Microbiol 102:227-236.
10. Agdestein A, Johansen TB, Polaček V, Lium B, Holstad G, Vidanović D, Aleksić-Kovačević S, Jørgensen A, Žultauskas J, Nilsen SF, Djønne B. 2011. Investigation of an outbreak of mycobacteriosis in pigs. BMC Vet Res 7:63.
11. Gardner IA, Hird DW. 1989. Environmental source of mycobacteriosis in a California swine herd. Can J Vet Res 53:33-37.
12. Hibiya K, Kazumi Y, Nishiuchi Y, Sugawara I, Miyagi K, Oda Y, Oda E, Fujita J. 2010. Descriptive analysis of the prevalence and the molecular epidemiology of *Mycobacterium avium* complex-infected pigs that were slaughtered on the main island of Okinawa. Comp Immunol Microbiol Infect Dis 33:401-421.

13. Inagaki T, Nishimori K, Yagi T, Ichikawa K, Moriyama M, Nakagawa T, Shibayama T, Uchiya K, Nikai T, Ogawa K. 2009. Comparison of a variable-number tandem-repeat (VNTR) method for typing *Mycobacterium avium* with mycobacterial interspersed repetitive-unit-VNTR and IS*1245* restriction fragment length polymorphism typing. J Clin Microbiol 47:2156-2164.
14. Klanicova B, Slana I, Vondruskova H, Kaevska M, Pavlik I. 2011. Real-time quantitative PCR detection of *Mycobacterium avium* subspecies in meat products. J Food Prot 74:636-640.
15. Agdestein A, Olsen I, Jørgensen A, Djønne B, Johansen TB. 2014. Novel insights into transmission routes of *Mycobacterium avium* in pigs and possible implications for human health. Vet Res 45:46.
16. Arikawa K, Ichijo T, Nakajima S, Nishiuchi Y, Yano H, Tamaru A, Yoshida S, Maruyama F, Ota A, Nasu M, Starkova DA, Mokrousov I, Narvskaya OV, Iwamoto T. 2019. Genetic relatedness of *Mycobacterium avium* subsp. *hominissuis* isolates from bathrooms of healthy volunteers, rivers, and soils in Japan with human clinical isolates from different geographical areas. Infect Genet Evol. 74:103923.
17. Genua F, Menichini M, Lari N, Rindi L. 2018. Genotyping and clarithromycin susceptibility testing of *Mycobacterium avium* subsp. *hominissuis* isolated in Tuscany, Italy. Infect Genet Evol 65:144-149.
18. Ichikawa K, van Ingen J, Koh WJ, Wagner D, Salfinger M, Inagaki T, Uchiya KI, Nakagawa T, Ogawa K, Yamada K, Yagi T. 2015. Genetic diversity of clinical *Mycobacterium avium* subsp. *hominissuis* and *Mycobacterium intracellulare* isolates

causing pulmonary diseases recovered from different geographical regions. Infect Genet Evol 36:250-255.

19. Imperiale BR, Moyano RD, DI Giulio AB, Romero MA, Alvarado Pinedo MF, Santangelo MP, ravería GE, Morcillo NS, Romano MI. 2017. Genetic diversity of *Mycobacterium avium* complex strains isolated in Argentina by MIRU-VNTR. Epidemiol Infect 145:1382-1391.

20. Iwamoto T, Nakajima C, Nishiuchi Y, Kato T, Yoshida S, Nakanishi N, Tamaru A, Tamura Y, Suzuki Y, Nasu M. 2012. Genetic diversity of *Mycobacterium avium* subsp. *hominissuis* strains isolated from humans, pigs, and human living environment. Infect Genet Evol 12:846-852.

21. Kalvisa A, Tsirogiannis C, Silamikelis I, Skenders G, Broka L, Zirnitis A, Jansone I, Ranka R. 2016. MIRU-VNTR genotype diversity and indications of homoplasy in *M. avium* strains isolated from humans and slaughter pigs in Latvia. Infect Genet Evol 43:15-21.

22. Leão C, Canto A, Machado D, Sanches IS, Couto I, Viveiros M, Inácio J, Botelho A. 2014. Relatedness of *Mycobacterium avium* subspecies *hominissuis* clinical isolates of human and porcine origins assessed by MLVA. Vet Microbiol 173:92-100.

23. Subangkit M, Yamamoto T, Ishida M, Nomura A, Yasiki N, Sudaryatma PE, Goto Y, Okabayashi T. 2019. Genotyping of swine *Mycobacterium avium* subsp. *hominissuis* isolates from Kyushu, Japan. J Vet Med Sci 81:1074-1079.

24. Yano H, Iwamoto T, Nishiuchi Y, Nakajima C, Starkova DA, Mokrousov I, Narvskaya O, Yoshida S, Arikawa K, Nakanishi N, Osaki K, Nakagawa I, Ato M, Suzuki Y,

Maruyama F. 2017. Population structure and local adaptation of MAC lung disease agent *Mycobacterium avium* subsp. *hominissuis*. Genome Biol Evol 9:2403-2417.

25. Yano H, Suzuki H, Maruyama F, Iwamoto T. 2019. The recombination-cold region as an epidemiological marker of recombinogenic opportunistic pathogen *Mycobacterium avium*. BMC Genomics 20:752.

26. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. 2014. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. PLoS One 11:e104984.

27. Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res 44:W232-W235.

28. Tonkin-Hill G, Lees JA, Bentley SD, Frost SDW, Corander J. 2019. Fast hierarchical Bayesian analysis of population structure. Nucleic Acids Res 47:5539-5549.

29. Mostowy R, Croucher NJ, Andam CP, Corander J, Hanage WP, Marttinen P. 2017. Efficient inference of recent and ancestral recombination within bacterial populations. Mol Biol Evol 34:1167-1182.

30. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691-3693.

31. Sanchini A, Semmler T, Mao L, Kumar N, Dematheis F, Tandon K, Peddireddy V, Ahmed N, Lewin A. 2016. A hypervariable genomic island identified in clinical and environmental *Mycobacterium avium* subsp. *hominissuis* isolates from Germany. Int J Med Microbiol 306:495-503.

32. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 58:3895-3903.
33. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929-931.
34. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16-W21.
35. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. Nucleic Acids Res 39:W347-352.
36. Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Winsor GL, Brinkman FSL. 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res 45:W30-W35.
37. Lahiri A, Kneisel J, Kloster I, Kamal E, Lewin A. 2014. Abundance of *Mycobacterium avium* ssp. *hominissuis* in soil and dust in Germany - implications for the infection route. Lett Appl Microbiol 59:65-70.
38. Makovcova J, Slany M, Babak V, Slana I, Kralik P. 2014. The water environment as a source of potentially pathogenic mycobacteria. J Water Health 12: 254-263.
39. Ghielmetti G, Hilbe M, Friedel U, Menegatti C, Bacciarini L, Stephan R, Bloemberg G. 2021. Mycobacterial infections in wild boars (*Sus scrofa*) from Southern

Switzerland: diagnostic improvements, epidemiological situation and zoonotic potential. Transbound Emerg Dis 68:573-586.

40. Odoi JO, Ohya K, Moribe J, Takashima Y, Sawai K, Taguchi K, Fukushi H, Wada T, Yoshida S, Asai T. 2020. Isolation and antimicrobial susceptibilities of nontuberculous Mycobacteria from wildlife in Japan. J Wildl Dis 56:851-862.

41. Shoulah SA, Oschmann AM, Selim A, Semmler T, Schwarz C, Kamal E, Hamouda F, Galila E, Bitter W, Lewin A. 2018. Environmental *Mycobacterium avium* subsp. *hominissuis* have a higher probability to act as a recipient in conjugation than clinical strains. Plasmid 95:28-35.

42. Komatsu T, Ohya K, Ota A, Nishiuchi Y, Hirokazu Y, Matsuo K, Opare Odoi, Shota Suganuma, Kotaro Sawai, Akemi Hasebe, Tetsuo Asai, Tokuma Yanai, Hideto Fukushi, Takayuki Wada, Shiomi Yoshida, Toshihiro Ito, Kentaro Arikawa, Mikihiko Kawai, Manabu Ato, Anthony D Baughn, Tomotada Iwamoto, Fumito Maruyama. 2021. Genomic features of *Mycobacterium avium* subsp. *hominissuis* isolated from pigs in Japan. Gigabyte 1.

43. Komatsu T, Ohya K, Ota A, Nishiuchi Y, Hirokazu Y, Matsuo K, Opare Odoi, Shota Suganuma, Kotaro Sawai, Akemi Hasebe, Tetsuo Asai, Tokuma Yanai, Hideto Fukushi, Takayuki Wada, Shiomi Yoshida, Toshihiro Ito, Kentaro Arikawa, Mikihiko Kawai, Manabu Ato, Anthony D Baughn, Tomotada Iwamoto, Fumito Maruyama. 2021. Supporting data for "Genomic features of *Mycobacterium avium* subsp. *hominissuis* isolated from pigs in Japan." GigaScience Database http://dx.doi.org/10.5524/100934.

44. Iwamoto T, Arikawa K, Nakajima C, Nakanishi N, Nishiuchi Y, Yoshida S, Tamaru A, Tamura Y, Hoshino Y, Yoo H, Park YK, Saito H, Suzuki Y. 2014. Intra-subspecies sequence variability of the MACPPE12 gene in *Mycobacterium avium* subsp. *hominissuis*. Infect Genet Evol 21:479-483.
45. Starkova DA, Otten TF, Mokrousov IV, Viazovaia AA, Vishnevskiĭ BI, Narvskaia OV. 2013. Genotypic characteristics of *Mycobacterium avium* subsp. *hominissuis* strains. Genetika 49:1048-1054.
46. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-2069.

Supporting Information Legends

Table S1. Information of MAH isolates in this study.

Table S2. Uniquely present or absent genes in new lineage (SC5).

Table S3. Geographic and host features of MAH lineages.

Table S4. VNTR profiles of 50 MAH isolates in this study.

Table S5. Information of MAH strains or isolates used for VNTR analysis in this study.

Figure S1. SNP position of *hsp65* gene. Characteristic SNPs were detected and shared in the new lineage (indicated by gray background). *1: Turenne et al., 2006 J Clin Microbiol. *2: Iwamoto et al., 2012 Infect Genet Evol (15).

Figure S2. Phylogenetic tree based on draft genome sequences generated by CSIphylogeny and IQ-TREE. a) Maximum likelihood tree was generated by CSI phylogeny version 1.4 and was visualized via MEGA 7.0. The area and Prefecture of isolates in this study were indicated by the color dots. Lineage classification except for new lineage was based on Yano et al., 2019. BMC Genomics. **b) Phylogenetic tree based on draft genome sequences generated by IQ-TREE.** Maximum likelihood tree was generated by IQ-TREE and visualized via MEGA 7.0. The Area and Prefecture of isolates in this study were indicated by the color dots. Lineage classification except for new lineage was based on Yano et al., 2019. BMC Genomics.

Figure S3. Lineage classification by fastBAPS. Lineage classification was predicted by fastBAPS (Tonkin-Hil et al., 2019 Nucleic Acids Res). Bootstrap values were calculated by 1,000 replications. The scale bar indicates genetic distance of each strain. Lineage classification in the right side except for new lineage was based on Yano et al., 2019. BMC Genomics. Neighbor joining tree was generated by via MEGA 7.0 using alignment file obtained from CSI phylogeny version 1.4.

Figure S4. Clustering analysis based on the presence or absence of accessory genes in MAH. Phylogenetic tree based on the presence or absence of accessory genes in MAH was constructed by Cluster 3.0 and visualized by TreeView 3.0.

Figure Legends

Figure 1. Minimum spanning tree (MST) based on 7 loci MIRU-VNTR genotyping of MAH isolates in Japan. Circles indicate different VNTR profiles. The size of each circle depended on the number of isolates sharing the same profiles. New lineages

indicated by an arrow were clustered and consisted of isolates GM10, GM11, GM12, GM21, GM23, GM24, GM25, GM32.

Figure 2. Minimum spanning tree (MST) based on 13 loci MATR-VNTR genotyping of MAH isolates from various geographic regions and sources. Circles indicate different VNTR profiles. The size of each circle depended on the number of isolates sharing the same profiles. Clonal complexes were highlighted in gray. Each clonal complex consisted of over five isolates.

Figure 3. Phylogenetic tree based on draft genomes. Neighbor joining tree was generated by MEGA 6.0 using alignment file obtained from CSI phylogeny version 1.4. Bootstrap values were calculated by 1,000 replications. The scale bar indicates genetic distance of each strain. Lineage classification except for new lineage was based on Yano et al (21).

Figure 4. Allelic variants in *cinA* gene of pig MAH. A) Maximum likelihood tree generated by MEGA 7.0. Bootstrap values were calculated by 1,000 replications. The scale bar indicates genetic distance of each strain. Lineage classification except for new lineage was based on Yano et al (25). B) SNP cites in pig MAH. New lineage shared the same SNP patterns, which were highlighted by gray. Some variations were detected in SC4. Reference SNP patterns were referred from Yano et al (25).

Figure 5. Allelic variants in *sugA* gene of pig MAH. A) Maximum likelihood tree generated by MEGA 7.0. Bootstrap values were calculated by 1,000 replications. The scale bar indicates genetic distance of each strain. Lineage classification except for new lineage was based on Yano et al (25). B) SNP cites in pig MAH. New lineage shared

the same SNP patterns, which were highlighted by gray. Two type of SNP patterns were detected in SC2. Reference SNP patterns were referred from Yano et al (25).

Figure 6. Determination of MAH lineages. Left: Visualization of recent horizontal acquisition events generated by fastGEAR. Color of chromosomal regions expresses donor lineage of the regions. Right: Neighbor joining tree was generated by via MEGA 7.0 using alignment file obtained from CSI phylogeny version 1.4.

Figure 7. Relationship among farms where new lineage MAH was isolated. Aichi Y farm bred and supplied pigs to Aichi X, Aichi B and Gifu A farm, where these farms had the novel lineage (GM21, GM10 and GM32). Aichi X farm whose sows were introduced from Aichi Y farm supplied pigs to Shiga A farm where GM12 was isolated. Finally, pigs of Gifu A, Aichi B and Shiga A farm were slaughtered in the same slaughtered plant (Gifu meat inspection center).

Figure 8. Location of each regions and prefectures in Japan. Pigs from Tokai area (southern Gifu, Shiga and Aichi Prefecture) are slaughtered in Gifu meat inspection center. Pigs from Hokuriku area (northern Gifu, Toyama and Ishikawa Prefecture) are slaughtered in Toyama meat inspection center.

Table 1. Predicted genomic islands and a number of genes on the genomic island.

Isolates	Number of genomic islands	Number of genes in genomic island
GM5	43	464
GM10	37	454
GM12	35	485
GM16	51	565
GM17	43	391
GM21	36	434
GM32	34	424
GM44	45	638
OCU467	7	89
OCU468	9	253
OCU469	28	321
OCU470	7	217
OCU471	35	311
OCU472	8	180
OCU473	38	489
OCU474	41	459
OCU475	9	164
OCU476	10	179
OCU477	35	376
OCU478	8	122
OCU479	11	295
OCU480	37	462
OCU481	8	111
OCU482	7	93
OCU483	11	155
OCU484	32	451
OCU485	8	112
OCU486	40	379
Toy194	44	686
Toy195	50	624

671

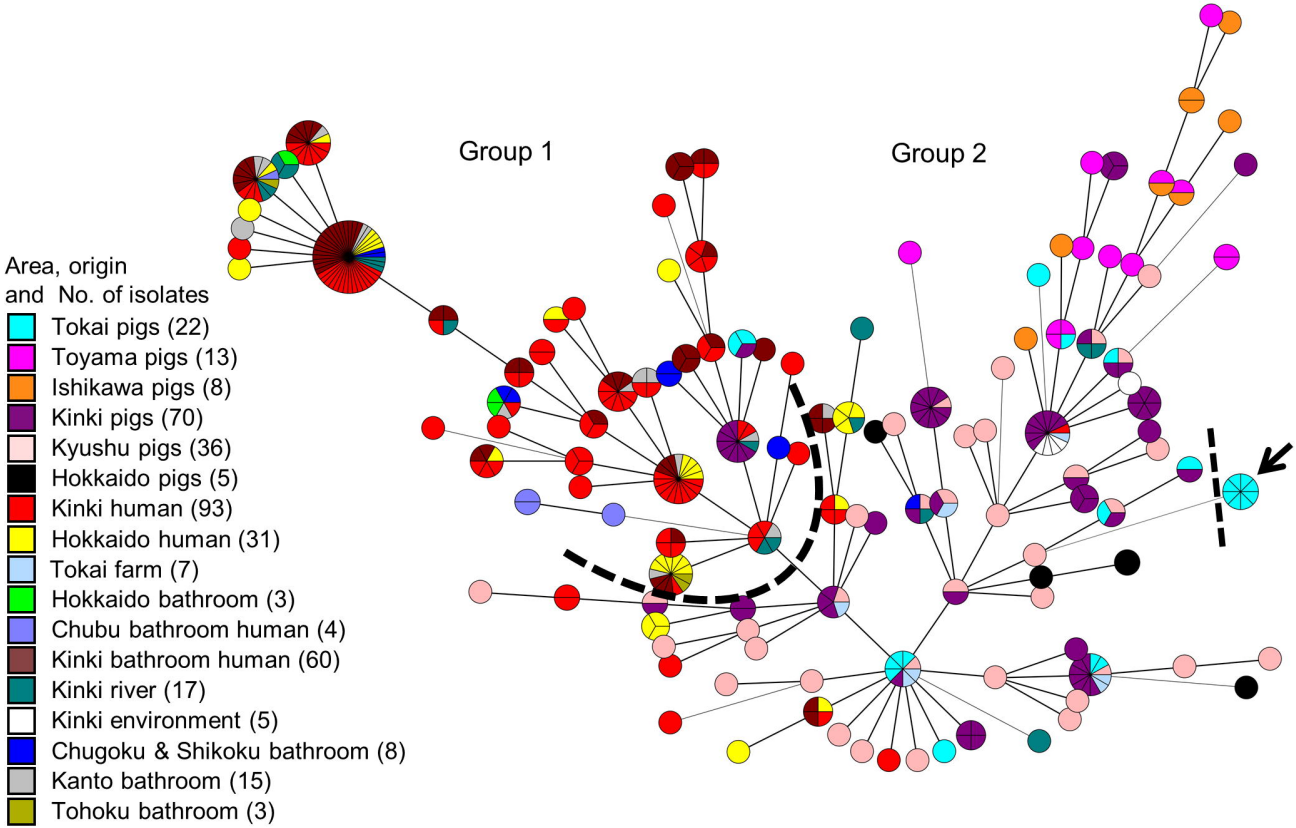
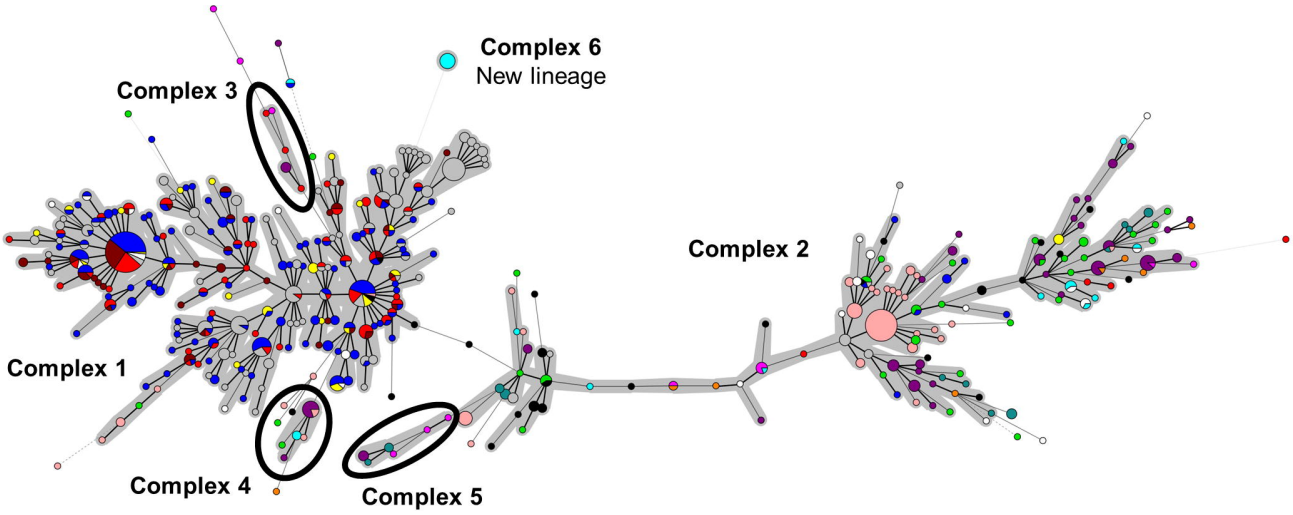


Fig 1. Minimum spanning tree (MST) based on 7 loci MIRU-VNTR genotyping of MAH isolates in Japan. Circles indicate different VNTR profiles. The size of each circle depended on the number of isolates sharing the same profiles. New lineages indicated by an arrow were clustered and consisted of isolates GM10, GM11, GM12, GM21, GM23, GM24, GM25, GM32.



Area, origin and No. of isolates

- Japan Gifu pigs (22)
- Japan Toyama pigs (13)
- Japan Ishikawa pigs (8)
- Japan Kinki pigs (70)
- Japan Okinawa pigs (19)
- Japan Chubu human (191)
- Japan Kinki human (93)
- Japan Hokkaido human (31)
- Japan bathroom human (60)
- Japan environment (29)
- Korea human (175)
- Russia human (90)
- Europe human (37)
- USA human (32)

Fig 2. Minimum spanning tree (MST) based on 13 loci MATR-VNTR genotyping of MAH isolates from various geographic regions and sources.

Circles indicate different VNTR profiles. The size of each circle depended on the number of isolates sharing the same profiles. Clonal complexes were highlighted in gray. Each clonal complex was consisted of over five isolates.

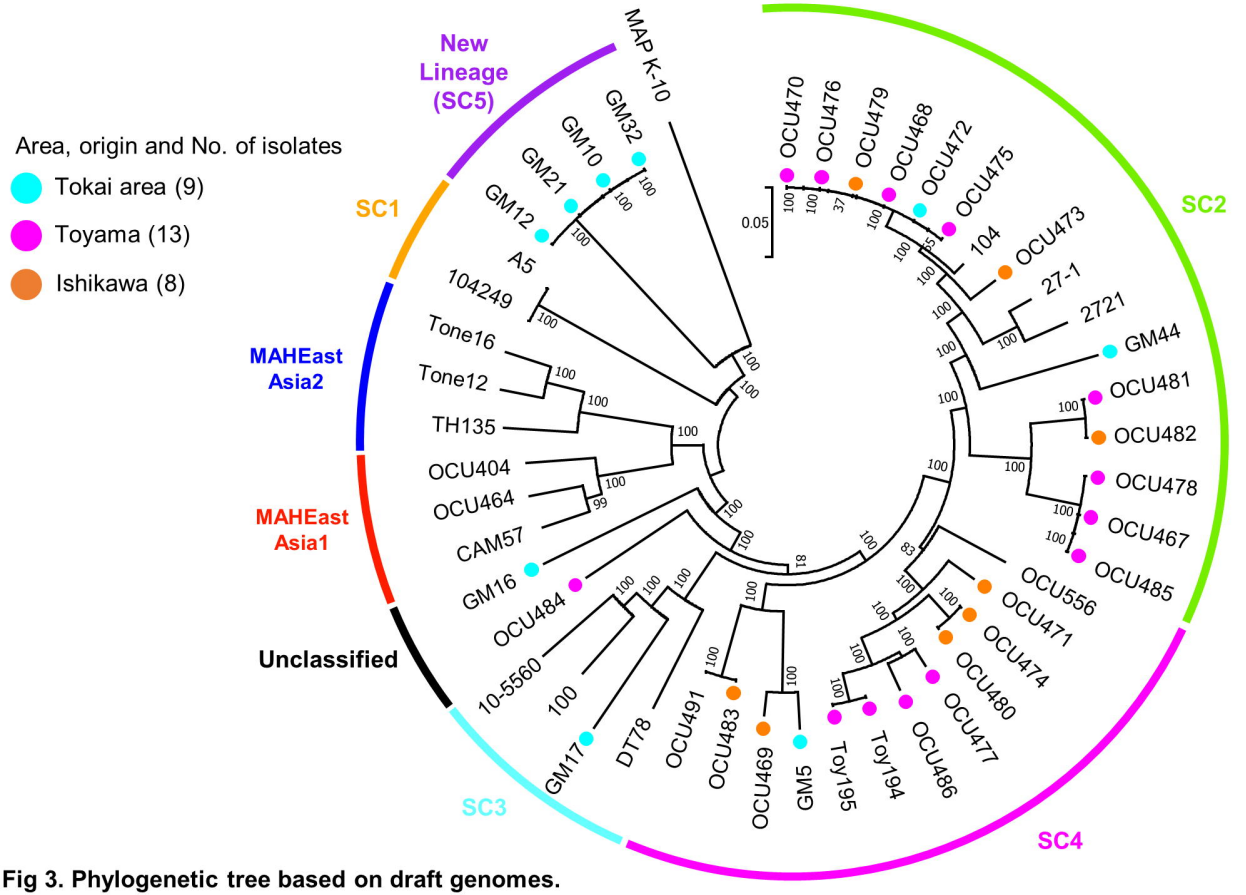
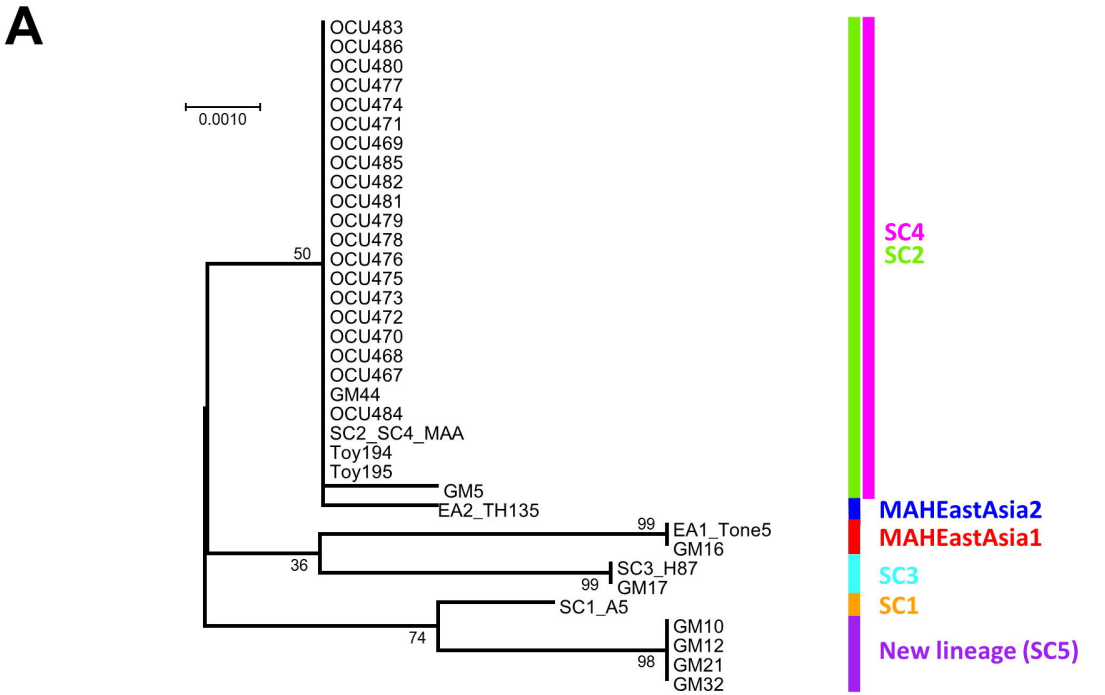


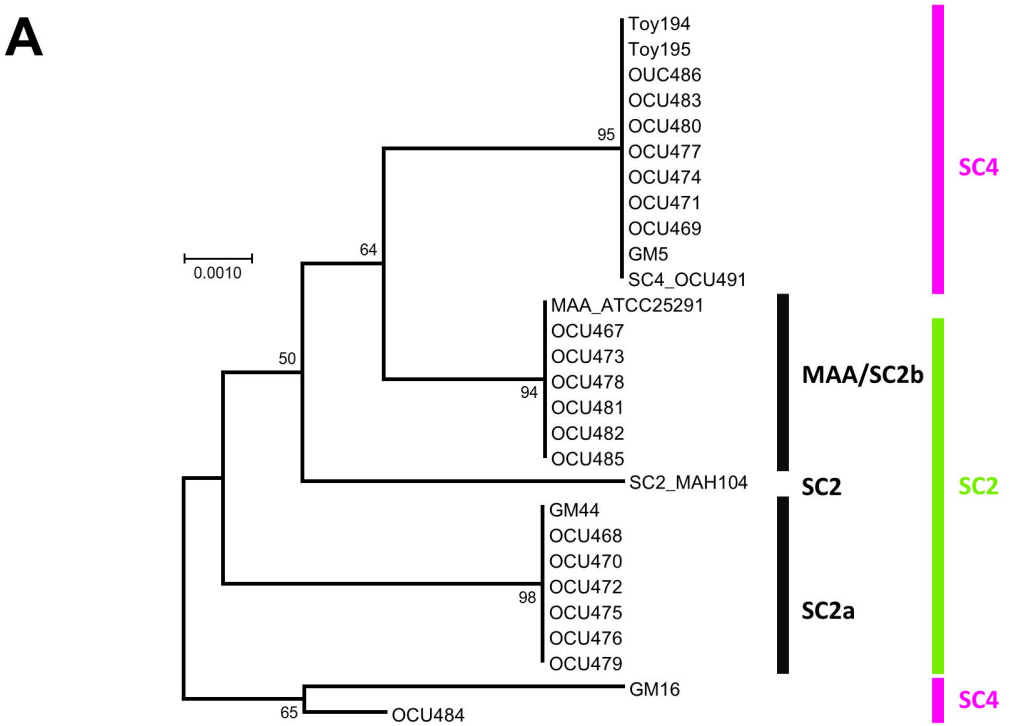
Fig 3. Phylogenetic tree based on draft genomes.
Neighbor joining tree was generated by MEGA 6.0 using alignment file obtained from CSI phylogeny version 1.4. Bootstrap values were calculated by 1,000 replications. The scale bar indicates genetic distance of each strain. Lineage classification except for new lineage was based on Yano et al (21).



B

		SNP position	36	72	157	198	231	240	267	288	351	552	564	570	705	708	816	888	897	900	918	1077	1140	1239	1269	1317
Reference		EA1 (Tone5)	C	G	T	T	C	T	C	G	T	C	C	A	C	A	T	C	T	C	C	G	G	G	A	A
		EA2 (TH135)	C	G	T	T	C	C	C	A	C	C	G	G	C	A	T	C	C	G	G	G	G	G	G	G
		SC1 (A5)	C	T	T	C	T	C	G	G	C	C	G	G	C	A	T	C	C	G	C	C	C	A	A	G
		SC2/SC4/MAA	C	G	T	T	C	C	C	A	C	C	G	G	C	A	T	C	C	G	G	G	G	A	A	G
		SC3 (H87)	C	G	T	T	C	T	C	G	C	T	G	G	G	G	T	T	C	G	C	G	G	G	G	G
		MAS (ATCC49884)	C	G	T	T	C	C	C	A	C	C	G	G	C	A	C	C	C	G	G	G	G	A	A	G
		MAP (K10)	-	G	C	C	C	C	G	G	T	C	G	G	C	A	T	C	C	G	C	C	C	A	A	G
This study	New lineage (SC5)	GM10	T	G	T	C	C	C	G	G	T	C	G	G	C	A	T	C	T	C	C	C	C	A	A	G
		GM12	T	G	T	C	C	C	G	G	T	C	G	G	C	A	T	C	T	C	C	C	C	A	A	G
		GM21	T	G	T	C	C	C	G	G	T	C	G	G	C	A	T	C	T	C	C	C	C	A	A	G
		GM32	T	G	T	C	C	C	G	G	T	C	G	G	C	A	T	C	T	C	C	C	C	A	A	G
	SC2	GM44, OCU467 OCU468, OCU470 OCU472, OCU473 OCU475, OCU476 OCU478, OCU479 OCU481, OCU482 OCU485	C	G	T	T	C	C	C	A	C	C	G	G	C	A	T	C	C	G	G	G	G	A	A	G
	SC3	GM17	C	G	T	T	C	T	C	G	C	T	G	G	G	G	T	T	C	G	C	G	G	G	G	G
	SC4	GM5	C	G	T	T	C	C	C	A	C	C	C	A	C	A	T	C	C	G	G	G	G	A	A	G
		GM16	C	G	T	T	C	T	C	G	T	C	C	A	C	A	T	C	T	C	C	G	G	G	A	A
		OCU469, OCU471 OCU474, OCU477 OCU480, OCU483 OCU484, OCU486 Toy194, Toy195	C	G	T	T	C	C	C	A	C	C	G	G	C	A	T	C	C	G	G	G	G	A	A	G

Figure 4. Allelic variants in *cinA* gene of pig MAH. A) Maximum likelihood tree generated by MEGA 7.0. Bootstrap values were calculated by 1,000 replications. The scale bar indicates genetic distance of each strain. Lineage classification except for new lineage was based on Yano et al (25). B) SNP cites in pig MAH. New lineage shared the same SNP patterns, which were highlighted by gray. Some variations were detected in SC4. Reference SNP patterns were referred from Yano et al (25).



B

			SNP position	25	71	165	168	298	336	387	393	432	441	495	592	654	664	669	705	789	852	855	868
Reference			SC2 (Mah104)	C	T	T	C	A	G	G	G	C	C	G	C	C	T	A	C	A	A	C	G
			SC4 (OCU491)	C	C	C	T	G	G	G	A	C	C	C	C	C	T	A	C	G	C	C	G
			MAA (ATCC25291)	C	C	T	C	G	G	G	G	A	C	C	C	A	T	A	C	G	C	C	G
This study	New lineage (SC5)		GM10	C	C	C	T	G	C	G	A	C	C	C	C	A	T	A	C	G	C	C	G
			GM12	C	C	C	T	G	C	G	A	C	C	C	C	A	T	A	C	G	C	C	G
			GM21	C	C	C	T	G	C	G	A	C	C	C	C	A	T	A	C	G	C	C	G
			GM32	C	C	C	T	G	C	G	A	C	C	C	C	A	T	A	C	G	C	C	G
	SC2	SC2a	GM44, OCU468 OCU470, OCU472 OCU475, OCU476 OCU479	T	C	C	C	G	G	G	G	C	C	G	T	C	C	G	C	G	C	C	G
		SC2b	OCU467, OCU473 OCU478, OCU481 OCU482, OCU485	C	C	T	C	G	G	G	G	A	C	C	C	A	T	A	C	G	C	C	G
	SC3		GM17	C	T	T	C	G	G	G	G	C	C	C	C	C	T	A	T	G	C	C	G
	SC4		GM5, OCU469 OCU471, OCU474 OCU477, OCU480 OCU483, OCU486 Toy194, Toy195	C	C	C	T	G	G	G	A	C	C	C	C	C	T	A	C	G	C	C	G
			GM16	C	C	T	C	G	G	A	G	C	G	G	T	C	T	G	T	G	A	G	G
			OCU484	C	C	T	C	G	G	A	G	C	G	G	T	C	T	A	C	G	C	C	C

Figure 5. Allelic variants in *sugA* gene of pig MAH. A) Maximum likelihood tree generated by MEGA 7.0. Bootstrap values were calculated by 1,000 replications. The scale bar indicates genetic distance of each strain. Lineage classification except for new lineage was based on Yano et al (25). B) SNP cites in pig MAH. New lineage shared the same SNP patterns, which were highlighted by gray. Two type of SNP patterns were detected in SC2. Reference SNP patterns were referred from Yano et al (25).

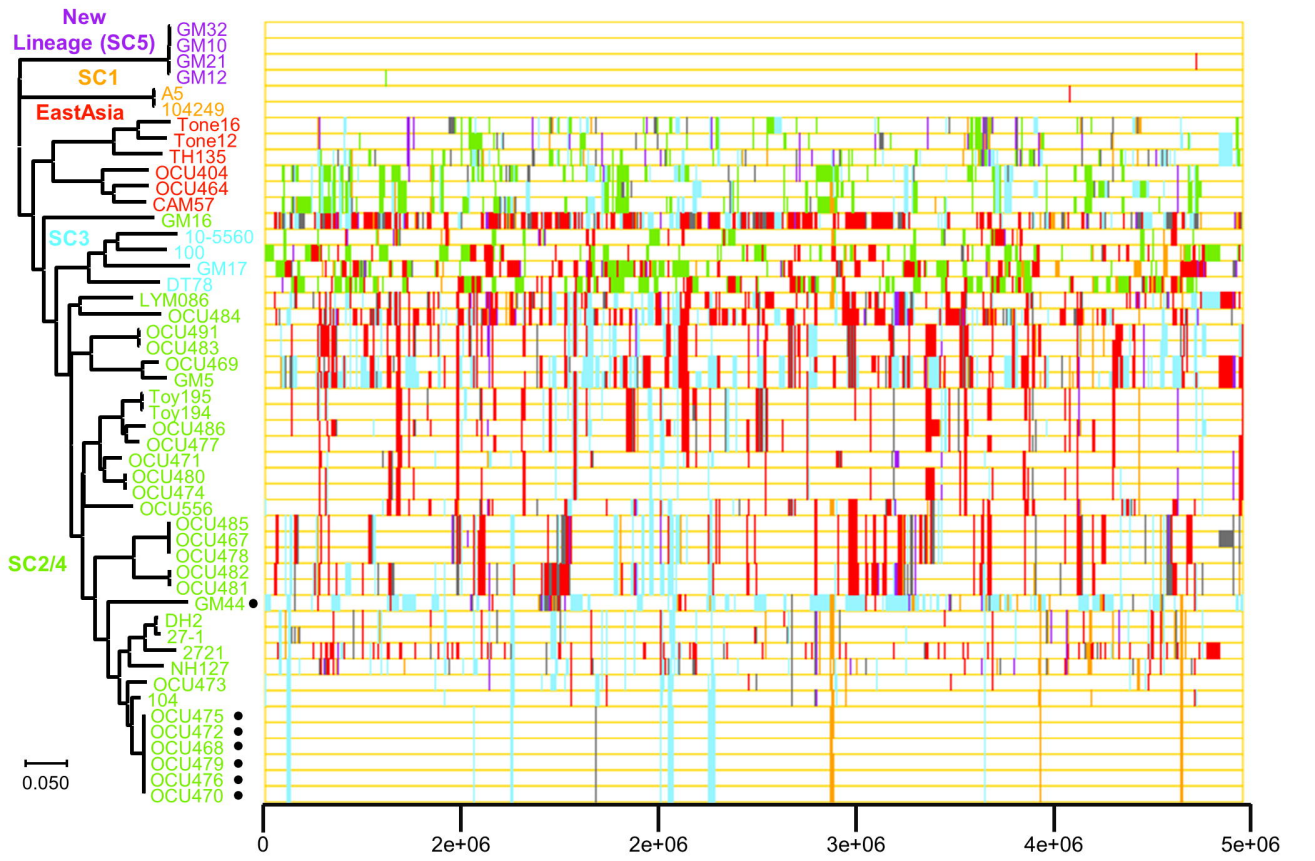


Fig 6. Determination of MAH lineages.

Left: Visualization of recent horizontal acquisition events generated by fastGEAR. Color of chromosomal regions expresses donor lineage of the regions. The isolates with dots were classified into SC2a by SNP analysis. Right: Neighbor joining tree was generated by via MEGA 7.0 using alignment file obtained from CSI phylogeny version 1.4.

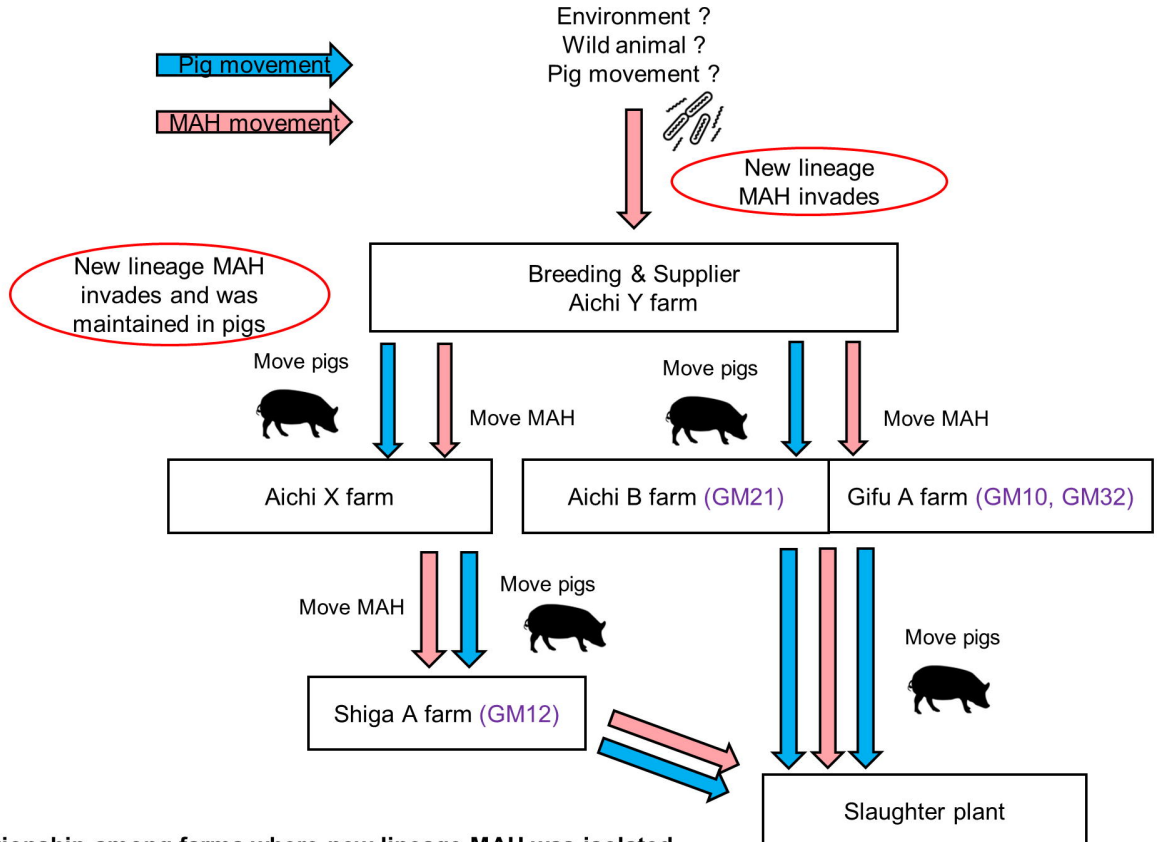


Fig 7. Relationship among farms where new lineage MAH was isolated.
Aichi Y farm bred and supplied pigs to Aichi X, Aichi B and Gifu A farm, where these farms had the novel lineage (GM21, GM10 and GM32). Aichi X farm whose sows were introduced from Aichi Y farm supplied pigs to Shiga A farm where GM12 was isolated. Finally, pigs of Gifu A, Aichi B and Shiga A farm were slaughtered in the same slaughtered plant (Gifu meat inspection center).

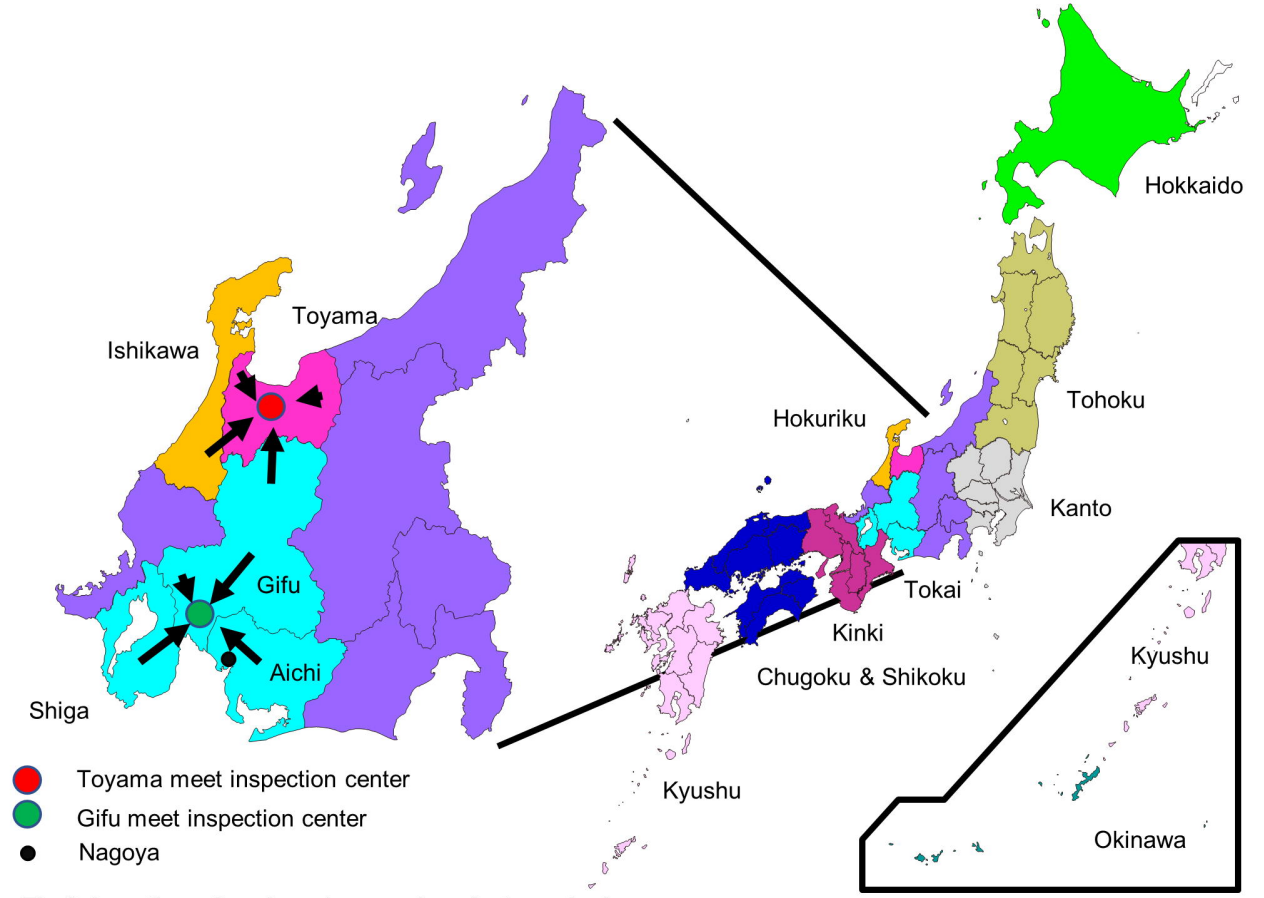


Fig 8. Location of each regions and prefectures in Japan.
Pigs from Tokai area (southern Gifu, Shiga and Aichi Prefecture) are slaughtered in Gifu meat inspection center. Pigs from Hokuriku area (northern Gifu, Toyama and Ishikawa Prefecture) are slaughtered in Toyama meat inspection center.