International Journal of Infectious Diseases 71 (2018) 82–88

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Genome sequence of *Mycobacterium yongonense* RT 955-2015 isolate from a patient misdiagnosed with multidrug-resistant tuberculosis: First clinical detection in Tanzania

Nicholaus Peter Mnyambwa[a,b,*], Dong-Jin Kim[a], Esther Ngadaya[b], Jongsik Chun[c,d], Sung-min Ha[c,d], Pammla Petrucka[a,e], Kennedy Kwasi Addo[f], Rudovick R. Kazwala[g], Sayoki G. Mfinanga[b]

[a] School of Life Science and Bioengineering, Nelson Mandela African Institution of Science and Technology, Arusha, Tanzania
[b] National Institute for Medical Research Muhimbili Medical Research Centre, Dar es Salaam, Tanzania
[c] School of Biological Sciences, Seoul National University, Seoul, Republic of Korea
[d] ChunLab, Inc., Seocho-gu, Seoul, Republic of Korea
[e] University of Saskatchewan, Saskatoon, Canada
[f] Bacteriology Department, Noguchi Memorial Institute for Medical Research (NMIMR), Accra, Ghana
[g] Faculty of Veterinary Medicine, Sokoine University of Agriculture, Morogoro, Tanzania

ARTICLE INFO

Article history:
Received 8 March 2018
Received in revised form 18 April 2018
Accepted 18 April 2018
Corresponding Editor: Eskild Petersen, Aarhus, Denmark

Keywords:
Mycobacterium yongonense
Mycobacterium avium complex (MAC)
Multidrug-resistant tuberculosis (MDR-TB)
Non-tuberculous mycobacteria (NTM)
Tanzania

ABSTRACT

Background: Mycobacterium yongonense is a recently described novel species belonging to *Mycobacterium avium* complex, which is the most prevalent aetiology of non-tuberculous mycobacteria associated with pulmonary infections, and poses tuberculosis diagnostic challenges in high-burden, resource-constrained settings.
Methods: Whole genome shotgun sequencing and comparative microbial genomic analyses were used to characterize the isolate from a patient diagnosed with multidrug-resistant tuberculosis (MDR-TB) after relapse.
Results: The genome sequence of the first case of *M. yongonense* (*M. yongonense* RT 955-2015) in Tanzania is presented. Sequence analysis revealed that the RT 955-2015 strain had a high similarity to *M. yongonense* 05-1390(T) (98.74%) and *Mycobacterium chimaera* DSM 44623(T) (98%). Its *16S rRNA* showed similarity to *Mycobacterium paraintracellulare* KCTC 290849(T) (100%), *Mycobacterium intracellulare* ATCC 13950(T) (100%), *M. chimaera* DSM 44623(T) (99.9%), and *M. yongonense* 05-1390(T) (98%). The strain exhibited a substantially different *rpoB* sequence to that of *M. yongonense* 05-1390 (95.16%), but closely related to that of *M. chimaera* DSM 44623(T) (99.86%), *M. intracellulare* ATCC 13950(T), (99.53%), and *M. paraintracellulare* KCTC 290849(T) (99.53%).
Conclusions: In light of the OrthoANI algorithm and phylogenetic analysis, it was concluded that the isolate was *M. yongonense* Type II genotype, which is an indication that the patient was misdiagnosed with TB/MDR-TB and received inappropriate treatment.

Introduction

Mycobacterium yongonense is a recently reported species from human specimens in South Korea (Kim et al., 2013a,c). Its *16S rRNA* gene shows a high degree of similarity to members of the *Mycobacterium avium* complex (MAC), namely *Mycobacterium marseillense* (100%), *Mycobacterium chimaera* (99.9%), and *Mycobacterium intracellulare* (99.8%), and it has a distinct *rpoB* gene sequence that resembles that of *Mycobacterium parascrofulaceum* (Kim et al., 2013a). Other members of MAC include *Mycobacterium timonense*, *Mycobacterium bouchedurhonense* (Ben et al., 2009), *Mycobacterium vulneris*, and *Mycobacterium colombiense* (Tortoli et al., 2004), as well as *Mycobacterium arosiense* (van Ingen et al., 2009). MAC is the most common group of non-tuberculous mycobacteria (NTM), which are frequently associated with pulmonary infections in humans, particularly in immune-compromised persons (Gordin et al., 1997). A complete genome

* Corresponding author at: School of Life Science and Bioengineering, Nelson Mandela African Institution of Science and Technology, Arusha, Tanzania.
E-mail address: lodnicho@gmail.com (N.P. Mnyambwa).

https://doi.org/10.1016/j.ijid.2018.04.796

sequence of *M. yongonense* isolated in Korea has been published previously, with an estimated circular DNA of 5.5 Mb and 5222 protein-coding genes (CDS) (Kim et al., 2013c). Subsequently, two clinical cases of *M. yongonense* in Italian patients with lung infections have been reported (Tortoli et al., 2013). Both strains had *rpoB* sequences showing high similarity to that of *M. intracellulare* (99.4%), but not to *M. parascrofulaceum*.

Kim et al. have suggested two distinct genotypes of *M. yongonense* based on the *rpoB* sequence, specifically *M. yongonense* Type I genotype with the *rpoB* gene acquired through horizontal gene transfer (HGT) from *M. parascrofulaceum* and *M. yongonense* Type II genotype with the *M. intracellulare rpoB* gene (Kim et al., 2016). More recently the same authors have provided additional evidence that the entire *M. yongonense* Type I *rpoBC* operon resulted from a distantly related species of *M. parascrofulaceum* (Kim et al., 2017). The authors also revealed that members of *M. yongonense* Type I genotype harbour a unique DNA mismatch repair gene *MutS4* family, that potentially serves as a putative driving force for the suggested HGT between the *M. parascrofulaceum* and *M. yongonense* Type I genomes through homologous recombination events (Kim et al., 2017). The occurrence of NTM in Tanzania has been reported frequently (Hoza et al., 2016a,b; Mfinanga et al., 2014; Kilale et al., 2016; Mnyambwa et al., 2017a), but characterization of the isolates has relied on suboptimal typing methods that have potentially hindered accurate species identification.

In 2008, the World Health Organization (WHO) endorsed the molecular tests known as line probe assays (LPA) and, more recently, the GeneXpert assay for the rapid detection of clinically resistant TB. MTBDR*plus* is one of the LPA that has been endorsed for rapid detection (within 24 h) of multidrug-resistant tuberculosis (MDR-TB) (World Health Organization, 2013). MTBDR*plus* can detect drug resistance-conferring mutations for rifampicin and isoniazid. The MDR-TB treatment guidelines in Tanzania recommend the use of MTBDR*plus* and a patient should be placed on MDR-TB treatment if there is proof of rifampicin resistance by LPA, GeneXpert, or conventional culture-based drug-susceptibility testing (DST) proportional methods (Ministry of Health, 2012).

The whole genome sequence of *M. yongonense* isolate RT 955-2015–the first case of infection to be reported in Tanzania – is described herein. The organism was isolated from a sputum specimen of a patient who was diagnosed with MDR-TB by GenoType MTBDR*plus* (Hain Lifescience GmbH, Nehren, Germany) after relapse.

Materials and methods

Culture and drug susceptibility testing

The GenoType MTBDR*plus* was performed in accordance with manufacturer's instructions. Culture and DST were performed in a contained biosafety level 3 laboratory at the Central TB Reference Laboratory (CTRL) in Dar es Salaam, Tanzania. The bacterial isolates were inoculated onto Löwenstein–Jensen (LJ) medium slants and incubated at 35–37 °C, yielding growth in the sixth week. All drug-sensitivity assays were performed in triplicate and susceptibility to a particular drug was determined by the standard proportion method, as recommended by the WHO (World Health Organization, 2009).

Genomic DNA extraction

The genomic DNA was extracted from heat-killed *Mycobacterium tuberculosis* by the cetyltrimethylammonium bromide (CTAB) method, as described in detail elsewhere (van Soolingen et al., 1991). In summary, bacterial isolates were grown on LJ medium slants at CTRL. Using a sterile loop, viable colonies were lifted from the LJ slants and emulsified into tubes containing 1 × Tris–ethylenediaminetetraacetic acid (EDTA) buffer (TE), pH 8.0, then heat-killed in a water bath at 80 °C for 20 min. This was followed by the addition of 10 mg/ml lysozyme in each tube and incubation at 37 °C overnight. On the following day, the DNA was extracted with chloroform–isoamyl alcohol (24:1) and the pellets of genomic DNA were rehydrated in 80 μl TE and left overnight at 4 °C. The gDNA was then temporarily stored at −20 °C before shipping to Kilimanjaro Clinical Research Institute (KCRI), Moshi, Tanzania for the sequencing procedure.

Whole genome sequencing

Dual-indexed (24) DNA libraries were constructed by shearing genomic DNA of approximately 300 bp using the NexteraXT DNA sample preparation kit according to the manufacturer's instructions (Illumina Inc., San Diego, CA, USA). Genomic DNA libraries were sequenced using a paired-end 2 × 250 bp protocol (MiSeq Illumina) at KCRI.

Genome assembly, annotation, and identification

The pre-processed sequencing reads were assembled by means of SPAdes 3.9.1, using default parameters to generate contigs and scaffolds. Scaffold-level assembly was evaluated using QUAST with the reference genome (*M. yongonense* 05-1390(T)) to generate summary statistics, such as N50, maximum scaffold length, G + C content, and others. The assembled scaffold sequences were ordered using ABACAS (Assefa et al., 2009) with *M. yongonense* 05-1390 as a reference genome. Gene prediction was performed using Prodigal and annotation was conducted using a homology search against the Clusters of Orthologous Groups (COG), eggNOG, SEED systems, Swiss-Prot, and KEGG databases. A species-level identification was performed using OrthoANI and *16S rRNA* with 98% and 97% as the threshold value, respectively (Chun's Bioinformatics Lab (ChunLab)).

Gene content (presence/absence) data were used to generate a heatmap (ChunLab). The construction of the heatmap utilizes tetranucleotides of CDS through the whole genome (pan-genome). The Jaccard coefficient was used to calculate similarities between genomes based on presence/absence data. Sequence alignment and syntenic assessment were performed using Mugsy, and an up-to-date bacterial core genome (UBCG) approximately maximum likelihood phylogenetic tree was reconstructed using the FastTree program. The phylogenetic tree was derived from the alignment of the pseudomolecule with reference sequences. The *M. tuberculosis* H37Rv was included as a reference genome representing members of the *Mycobacterium tuberculosis* complex (MTBC), while *M. intracellulare* ATCC 13950 and *M. parascrofulaceum* provided additional evidence for the classification of the isolate into a specific genotype (Type I or Type II genotype). Additionally, BLAST analysis against the NCBI RefSeq database was performed.

Multilocus sequence typing

Multilocus sequence typing (MLST) sequences were analyzed using 10 public *M. yongonense* genomes from the EzBioCloud database (ChunLab). The Kruskal algorithm, based on the allelic distance matrix of the shared loci, was used to construct a minimum spanning tree. Datasets used and/or analyzed in this study are freely available at https://www.bioiplug.com/genome/YongonenseRT9552015Set.

Quality control

The quality and quantity of gDNA were confirmed using a Qubit 2.0 fluorometer (Thermal Fisher Scientific, Waltham, MA USA).

FastQC and K-mers spectra analysis by K-mer Analysis Toolkit (KAT 2.3.2) were used to assess the quality of the raw sequence data. Trimmomatic version 0.36 and FastX-ToolKit (Hannon Lab) were used to remove Illumina adapter sequences and filter out low quality sequenced bases and reads. The de novo genome assembly was validated against a complete genome sequence of *M. yongonense* 05-1390. Genome contamination was checked using the ContEst16s tool and coverage was checked based on bacterial core genes.

Results

Demographic and clinical characteristics of the patient

The patient was a 60-year-old, HIV-negative male living in Dar es Salaam, Tanzania, who was diagnosed with TB relapse by smear examination and mycobacterial culture (2015). Additionally, drug resistance-conferring mutations for rifampicin and isoniazid were detected using a rapid GenoType MTBDR*plus* VER 2.0 test. Subsequently, the conventional phenotypic culture-based DST (proportion method) on LJ medium was performed and the isolates were found to be susceptible to all first-line anti-TB drugs. Based on the MTBDR*plus* test and medical history, the patient was classified as having MDR-TB and was hence placed on the MDR-TB treatment regimen. The patient converted to sputum smear- and culture-negative status in the third month of MDR-TB treatment, thereby being considered cured on completion of treatment.

Sequence analysis

MiSeq sequencing generated 400 000 reads, which passed quality checks with an average of approximately 28 × depth coverage. No evidence of genome contamination was found. The de novo assembly generated 202 scaffolds containing 402 contigs. The findings revealed a genome size of 5 528 170 bp with 99.61% degree of genome coverage. The *M. yongonense* RT 955-2015 genome has 5331 open reading frames (ORFs), 46 tRNA genes, four rRNA genes, and one tmRNA. The mean (standard deviation) length of the CDS was 913.8 (652.3) bp and the median length of the CDS was 807 bp. The mean (standard deviation) length of the Intergenic region (IGR) was 137.7 (194.3) bp. The genome of *M. yongonense* RT 955-2015 has a G + C content of 67.96%. A comparison of predicted ORFs of *M. yongonense* RT 955-2015 and that of *M. yongonense* 05-1390(T) showed that the two genomes share 4763 ORFs with an average genome identity of 98.74%. A large number of unshared genes among *M. yongonense* genomes was observed. Comparative functional analysis of genes showed that about one-third of genes were assigned unknown function and the majority of remaining genes were assigned to transcription (6.7%), lipid transport and metabolism (6.4%), energy conversion and production (6.0%), secondary metabolite biosynthesis, transport and catabolism (5.5%), replication, and amino acid transport and metabolism (see Figure 1 and Supplementary material, Table S1 in the online version, at DOI:10.1016/j.ijid.2018.04.796).

Species identification

Genome similarity analysis of the RT 955-2015 isolate showed the highest similarity to *M. yongonense* 05-1390, while both *16S rRNA* and *rpoB* showed high similarity to *M. intracellulare* ATCC 13950 and its related members (*M. paraintracellulare* KCTC 290849 and *M. chimaera* DSM 44623) (Table 1). Gene content analysis showing the evolutionary relationship of the isolate suggested that the RT 955-2015 isolate shares a recent ancestry with *M. yongonense* species. The clustering analysis of gene encoding protein (CDS) content information as expressed by the presence (blue) and absence (red) of genes is shown in Figure 2A. The RT 955-2015 strain was found to be closely related to *M. yongonense* species (I and II) and quite different from *M. tuberculosis* H37Rv in

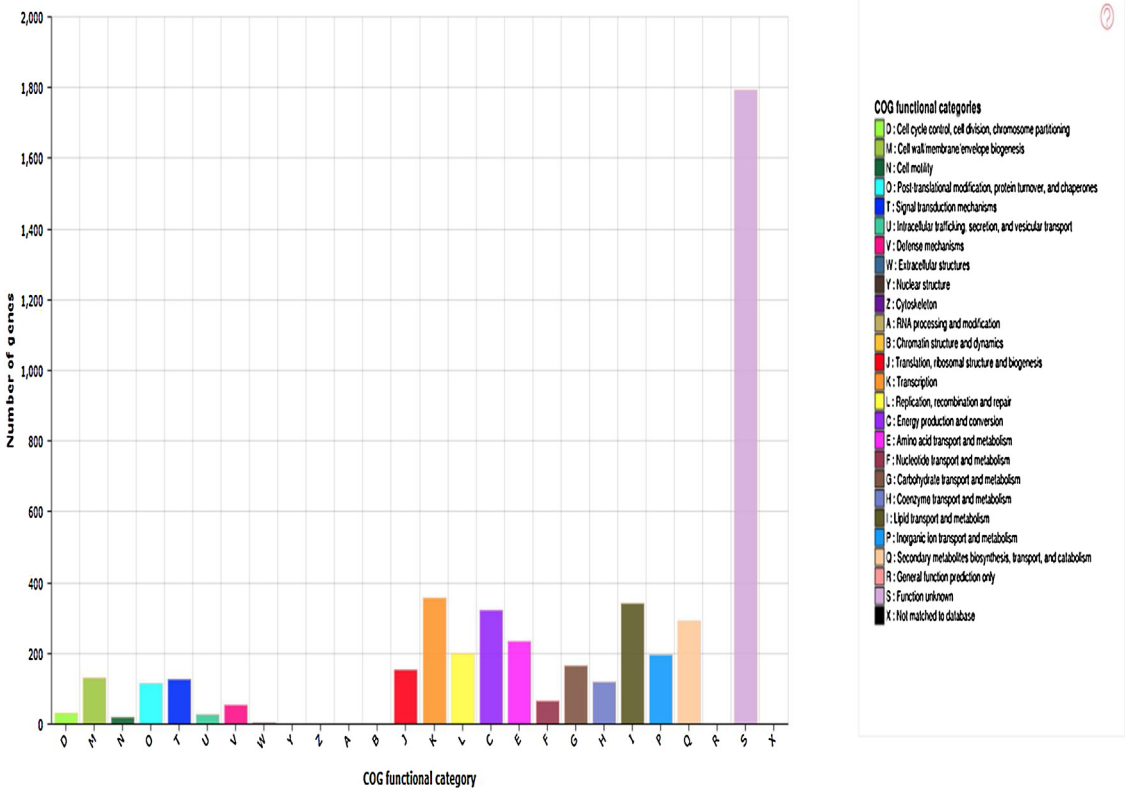

Figure 1. Distribution of CDS based on function prediction.

Table 1
Similarity values generated by the OrthoANI showing the similarity of isolate RT 955-2015 to the closest members of *Mycobacterium avium* complex.

Taxonomy name	Genome identity (%)	*16S rRNA* similarity (%)	*hsp65* similarity (%)	*rpoB* similarity (%)	ITS1 similarity (%)
M. yongonense 05-1390(T)	98.74	99.8	99.12	95.16	98.23
M. chimaera DSM 44623(T)	98.00	99.9	99.26	99.86	94.00
M. paraintracellulare KCTC 290849(T)	97.50	100	99.26	99.53	99.29
M. intracellulare ATCC 13950(T)	97.50	100	99.45	99.53	99.29
M. marseillense DSM 43457(T)	92.10	99.8	98.22	97.24	98.27
M. parascrofulaceum ATCC BAA-614(T)	83.83	97.89	94.46	95.07	89.36

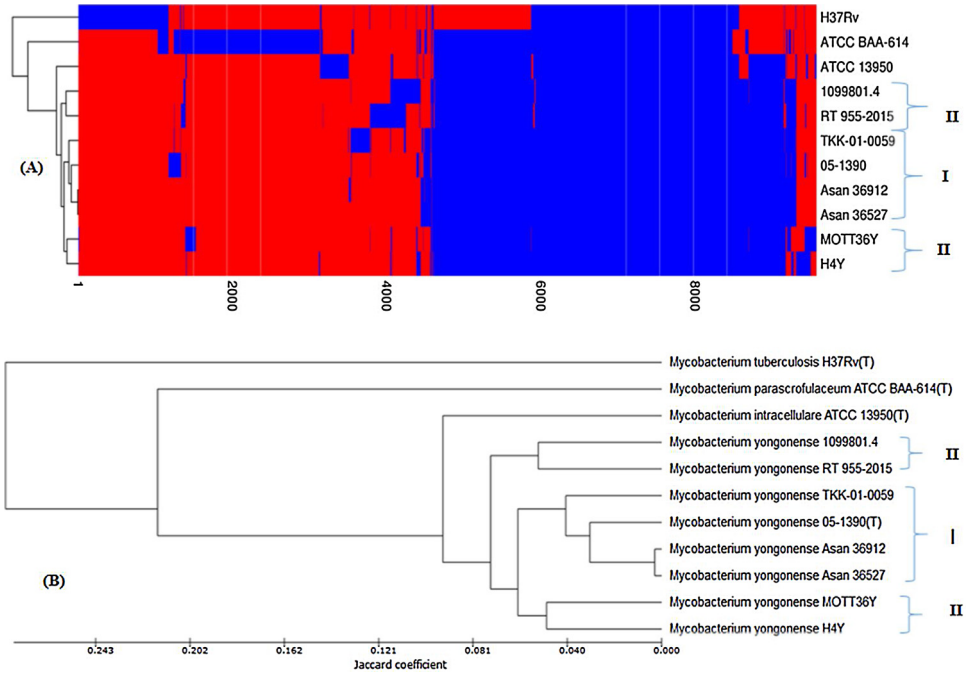

Figure 2. Phylogeny based on the presence/absence of genes. (A) Heatmap showing clustering analysis of the gene content among genomes. The pairwise orthologous relationship is checked through reciprocal UBLAST hits. (B) UPGMA clustering of the strains based on tetranucleotide compositions.

the composition of tetranucleotides. Figure 2B shows the same dendrogram without detailed gene content information. The UBCG-based maximum likelihood phylogenetic tree also supports the finding that the RT 955-2015 strain is a member of *M. yongonense* (Figure 3). These analyses also reflect the previous classification of *M. yongonense* species into two distinct genotypes (Type I genotype and Type II genotype). BLAST analysis against the NCBI RefSeq database of all hits and top hits suggested that the isolate RT 955-2015 was *M. yongonense*.

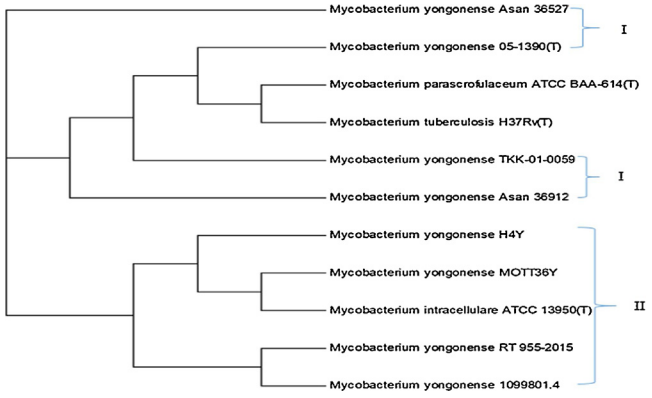

Figure 3. UBCG-based maximum likelihood phylogenetic tree showing the evolutionary relationship of *Mycobacterium yongonense* RT-955-2015.

Mutations

A number of single nucleotide polymorphism (SNPs) were detected in six genes, namely *rpoB*, *rpoC*, *kasA*, *katG*, *inhA*, and *pnacA*. More polymorphisms (including non-synonymous mutations) were common in the *rboB* gene (Supplementary material, Table S2 in the online version, at DOI:10.1016/j.ijid.2018.04.796).

MLST of M. yongonense

A total of 4019 core genes were retrieved from 10 public *M. yongonense* genomes and the RT 955-2015 isolate for strain typing, and a minimum spanning tree was generated (Figure 4). Of the 10 genomes retrieved from public databases, nine were reclassified as *M. yongonense* in the EzBioCloud database. The profile was then used to search the core genes in each genome including the RT 955-2015 strain. Coloured circles in the tree represent the country where a strain was isolated and the distance between the nodes indicates the genetic distance of the isolates. The shorter the distance the closer the strains are; thus, the RT 955-2015 strain seems to be very close to the *M. yongonense* 1099801.4 strain.

Discussion

This study presents the whole genome sequence of a clinical isolate that was confirmed as *M. yongonense* RT 955-2015–the first

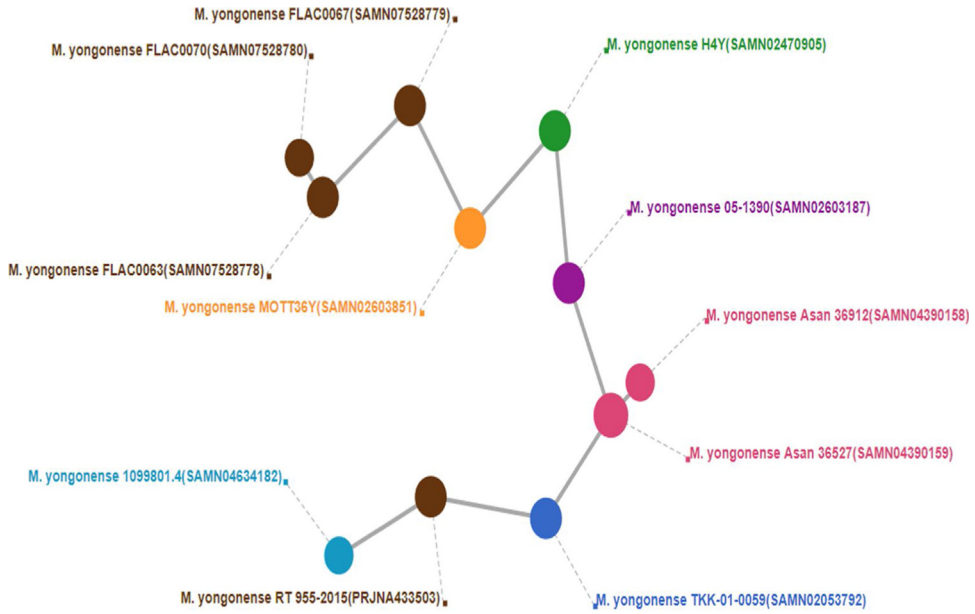

Figure 4. Minimum spanning tree based on MLST showing genetic distances among 11 strains of *Mycobacterium yongonense* species. The BioSample accession code of each strain is provided in parenthesis.

case detected in Tanzania. The study findings are based on an OrthoANI algorithm, gene content (presence/absence), and complete genome sequence-based phylogenetic analyses, which further suggested that the *M. yongonense* RT 955-2015 belongs to the Type II genotype. Of note, the heatmap and UPGMA dendrogram, based on the gene presence/absence, showed a similar strain clustering pattern as the UBCG-based maximum likelihood phylogenetic tree. Isolate identification based on similarity analysis of the *16S rRNA* gene sequence alone was not possible due to its high degree of similarity (>99%) with other members of MAC (Table 1). Previous findings have demonstrated that the *16S rRNA* gene lacks resolution power above 98% similarity (Lee et al., 2016). As there was no evidence of genome contamination, the involvement *M. yongonense* RT 955-2015 as the causative agent of the disease and that the patient was misdiagnosed with TB/MDR-TB seems incontrovertible.

The substantial difference between the *M. yongonense* RT 955-2015 *rpoB* sequence and that of *M. parascrofulaceum* and *M. yongonense* 05-1390, while showing high similarity to *M. intracellulare* and its closely related species (*M. chimaera* DSM 44623, *M. paraintracellulare* KCTC 290849), is additional evidence that the RT 955-2015 strain belongs to the Type II genotype group. Kim et al. used genome-based phylogenetic analysis to classify *M. yongonense* species into two distinct genotypes (Type I genotype and Type II genotype) (Kim et al., 2016), and demonstrated that Type I genotype has laterally acquired *rpoB* from *M. parascrofulaceum* while Type II genotype possess the *M. intracellulare rpoB* gene (Kim et al., 2016, 2017). Kim et al. classified *M. yongonense* 05-1390, *M. yongonense* Asan 36912, and *M. yongonense* Asan 36527 as members of Type I genotype, while *M. yongonense* MOTT36Y (INT-5) and *M. yongonense* H4Y (INT-5) were classified as Type II genotype (Kim et al., 2016). The present study findings for these strains are consistent with their classification. Furthermore, the study findings suggest that *M. yongonense* TKK-01-0059 (formerly known as *Mycobacterium sp* TKK-01-0059) from South Africa, belongs to *M. yongonense* Type I genotype. The pronounced difference in a number of unshared genes among the *M. yongonense* genomes highlights the possibility of extensive HGT from outside the cell in these genomes. Evidence of HGT among NTM members has been described previously (Kim et al., 2016, 2017; Fedrizzi et al., 2017; Kim et al., 2013b).

MLST demonstrated that the *M. yongonense* RT 955-2015 strain was closely related to *M. yongonense* 1099801.4, which was formerly classified as a member of *M. intracellulare*, sampled in Mozambique in 2012. The strain is placed in the same group (Type II genotype) with *M. yongonense* RT 955-2015 based on pan-genome and phylogenetic analysis.

In this study, *M. yongonense* TKK-01-0059 (formerly *Mycobacterium sp* TKK-01-0059) from South Africa (2009) was also placed in the Type I genotype. A similar classification has been proposed by Kim et al., based on the presence of *MutS4* and supported by *hsp65* and genome sequence-based phylogenetic analyses (Kim et al., 2017). The first three stains in the top of the spanning tree were isolated in the USA and the remaining five isolates were isolated from South Korea. The clustering of these isolates in the minimum spanning tree based on the three geographical regions where they were isolated (Africa, Korea, and the USA) may infer an important evolutionary relationship among the strains.

M. yongonense is a recently described species (Kim et al., 2013a, c) of NTM belonging to the MAC members, which are associated with pulmonary infections in humans. The steady increase in health complexities and mortality rate caused by several NTM species have recently been recognized (Yeung et al., 2016) and are raising a public health concern. The epidemiology and clinical implications of *M. yongonense* remain a mystery, as only a few cases have been described. This can be attributed to the use of suboptimal diagnostic tests, particularly in developing countries like Tanzania. The detection of acid-fast bacilli by smear examination of NTM can result in the misdiagnosis of TB cases (Hoza et al., 2016a,b; Mfinanga et al., 2014; Kilale et al., 2016; Mnyambwa et al., 2017a), which normally do not respond to anti-TB drugs, thereby leading to classification as MDR-TB (Lu et al., 2016; Shahraki et al., 2015). Most of the MAC members are naturally resistant to anti-TB drugs (Shahraki et al., 2015; Mpagama et al., 2013), which leads to the unnecessary administration of toxic second-line anti-TB treatments, thereby increasing the risk of the evolution of drug resistance. Thus, proper tests are needed for an effective healthcare system, providing critical

information for the correct medical decisions and management of a particular health condition. Such diagnostic tests should also help to address epidemiological concerns regarding the *M. yongonense* species and other underreported members of NTM. SNP-based phylogenetic analysis targeting five genes (*dnaA*, *deaD*, *argH*, *recF*, and *hsp65*) identified potential diagnostic biomarkers specific for *M. yongonense* (Kim et al., 2016).

It is estimated that between 7% and 12% of the adult population has been previously infected with MAC, but the rate of the disease varies significantly by geographic region (Gordin et al., 1997; Benson, 1994). MAC disease is typically prevalent in patients with compromised immune systems (Gordin et al., 1997). A high incidence rate, between 20% and 40%, has been reported in patients with severely low CD4 cell counts associated with HIV infection, especially when there is an ineffective antiretroviral treatment or chemoprophylaxis (Gordin et al., 1997). MAC infections have been reported to cause serious disease in HIV patients globally (Gordin et al., 1997; Kwak et al., 2017), and evidence of such mycobacterial infections has been reported in Tanzania (Crump et al., 2009).

Although the GenoType MTBDR*plus* assay is designed to detect common resistance mutations of the *rpoB* gene for rifampicin and *katG* and *inhA* for isoniazid in MTBC strains (Lacoma et al., 2008), sequence analysis of *M. yongonense* RT 955-2015 revealed a catalogue of non-synonymous mutations in the same genes and in three other genes, namely *rpoC*, *kasA*, and *pnacA*. Pyrazinamide resistance-conferring mutations are known to occur in the *pnacA* gene, while mutations in the *rpoC* gene are associated with resistance to rifampicin in MTBC (Mnyambwa et al., 2017b). Although evidence of genetic markers that are known to cause drug resistance in MTB strains are provided here, it remains a mystery whether these genetic signatures play the same role in *M. yongonense*.

In conclusion, sequence analysis suggests that the isolate RT 955-2015 is *M. yongonense* Type II genotype, a member of MAC, and this is the first detected case in Tanzania. These findings strongly affirm the hypothesis that the condition was misdiagnosed as TB/MDR-TB and that the patient received inappropriate treatments. This finding highlights the need for accurate diagnostic assays to ensure the effective use of drugs and to minimize the risk of the development of drug resistance and worsening of the patient's health condition.

This whole genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number **PSQD00000000** and assigned the name *Mycobacterium yongonense* RT 955-2015. The version described in this paper is version **PSQD01000000**.

Acknowledgements

The authors extend their sincere appreciation to Happiness Kumburu (KCRI) and Daphne Mtunga and Salim Bossy (CTRL) for their support during the laboratory work, and to Jean-Baka Domelevo and Dedan Githae (BeCA-ILRI) and Edson Ishengoma (Mkwawa University College of Education) for their technical support during the sequence analysis.

Funding

This work was mainly supported by the Government of Tanzania through the Nelson Mandela African Institution of Science and Technology and was partially supported by the DELTAS Africa Initiative (Afrique One-ASPIRE/DEL-15-008). Afrique One-ASPIRE is funded by a consortium of donors including the African Academy of Sciences (AAS)Alliance for Accelerating Excellence in Science in Africa (AESA), the New Partnership for Africa's Development Planning and Coordinating (NEPAD) Agency, the Wellcome Trust (107753/A/15/Z), and the UK government. The sponsors had no involvement in the study design, data collection, or in the interpretation of the data, and the views expressed in this publication are those of the authors and not necessarily those of Afrique One-ASPIRE or the Government of Tanzania.

Ethical considerations

This study was part of a research study aimed at investigating the genetic basis of drug resistance patterns in *Mycobacterium tuberculosis* complex in clinical isolates from tuberculosis patients in Tanzania, which was approved by the National Institute for Medical Research, Tanzania.

Availability of data and material

All important datasets used and/or analyzed in this study are available at https://www.bioiplug.com/genome/YongonenseRT9552015Set.

Conflict of interest

The authors declare that there is no conflict of interest.

Author contributions

NPM, DJK, SM, AKK, and RK conceived the study idea and co-designed the protocol. DJK, SM, RK, AKK, and EN co-supervised the study. NPM performed a bacterial culture, prepared DNA isolates, and performed genome sequencing. NPM, SH, and JC analyzed the sequence data and drafted the paper. SGM, ESN, RK, DJK, and PP revised the paper. All authors approved the final version for publication.

References

Assefa S, Keane TM, Otto TD, Newbold C, Berriman M. ABACAS: algorithm-based automatic contiguation of assembled sequences. Bioinformatics 2009;25:1968–9, doi:http://dx.doi.org/10.1093/bioinformatics/btp347.

Ben Salah I, Cayrou C, Raoult D, Drancourt M. *Mycobacterium marseillense* sp. nov., *Mycobacterium timonense* sp. nov. and *Mycobacterium bouchedurhonense* sp. nov., members of the *Mycobacterium avium* complex. Int J Syst Evol Microbiol 2009;59:2803–8, doi:http://dx.doi.org/10.1099/ijs.0.010637-0.

Benson CA. Disease due to the *Mycobacterium avium* complex in patients with AIDS: epidemiology and clinical syndrome. Clin Infect Dis 1994;18(Suppl. 3):S218–22.

Crump JA, van Ingen J, Morrissey AB, Boeree MJ, Mavura DR, Swai B, et al. Invasive disease caused by nontuberculous mycobacteria, Tanzania. Emerg Infect Dis 2009;15:53–5, doi:http://dx.doi.org/10.3201/eid1501.081093.

Fedrizzi T, Meehan CJ, Grottola A, Giacobazzi E, Fregni Serpini G, Tagliazucchi S, et al. Genomic characterization of Nontuberculous Mycobacteria. Sci Rep 2017;7:45258, doi:http://dx.doi.org/10.1038/srep45258.

Gordin FM, Cohn DL, Sullam PM, Schoenfelder JR, Wynne BA, Horsburgh CR. Early manifestations of disseminated *Mycobacterium avium* complex disease: a prospective evaluation. J Infect Dis 1997;176:126–32.

Hoza AS, Lupindu AM, Mfinanga SGM, Moser I, König B. The role of nontuberculous mycobacteria in the diagnosis, management and quantifying risks of tuberculosis in Tanga, Tanzania. Tanzan J Health Res 2016a;18:.

Hoza AS, Mfinanga SGM, Rodloff AC, Moser I, König B. Increased isolation of nontuberculous mycobacteria among TB suspects in Northeastern, Tanzania: public health and diagnostic implications for control programmes. BMC Res Notes 2016b;9:109, doi:http://dx.doi.org/10.1186/s13104-016-1928-3.

Kilale AM, Ngadaya E, Muhumuza J, Kagaruki GB, Lema YL, Ngowi BJ, et al. Who has mycobacterial disease? A cross sectional study in agropastoral communities in Tanzania. PLoS One 2016;11:e0153711, doi:http://dx.doi.org/10.1371/journal.pone.0153711.

Kim B-J, Math RK, Jeon CO, Yu H-K, Park Y-G, Kook Y-H, et al. *Mycobacterium yongonense* sp. nov., a slow-growing non-chromogenic species closely related to *Mycobacterium intracellulare*. Int J Syst Evol Microbiol 2013a;63:192–9, doi:http://dx.doi.org/10.1099/ijs.0.037465-0.

Kim B-J, Hong S-H, Kook Y-H, Kim B-J. Molecular evidence of lateral gene transfer in rpoB gene of *Mycobacterium yongonense* strains via multilocus sequence analysis. PLoS One 2013b;8:e51846, doi:http://dx.doi.org/10.1371/journal.pone.0051846.

Kim B-J, Kim B-R, Lee S-Y, Seok S-H, Kook Y-H, Kim B-J. Whole-genome sequence of a novel species, *Mycobacterium yongonense* DSM 45126T. Genome Announc 2013c;1:, doi:http://dx.doi.org/10.1128/genomeA.00604-13.

Kim B-J, Kim B-R, Lee S-Y, Kim G-N, Kook Y-H, Kim B-J. Molecular taxonomic evidence for two distinct genotypes of *Mycobacterium yongonense* via genome-based phylogenetic analysis. PLoS One 2016;11:e0152703, doi:http://dx.doi.org/10.1371/journal.pone.0152703.

Kim B-J, Kim B-R, Kook Y-H, Kim B-J. Role of the DNA mismatch repair GeneMutS4in driving the evolution of *Mycobacterium yongonense* Type I via homologous recombination. Front Microbiol 2017;8:2578, doi:http://dx.doi.org/10.3389/fmicb.2017.02578.

Kwak N, Park J, Kim E, Lee C-H, Han SK, Yim J-J. Treatment outcomes of *Mycobacterium avium* complex lung disease: a systematic review and meta-analysis. Clin Infect Dis 2017;65:1077–84, doi:http://dx.doi.org/10.1093/cid/cix517.

Lacoma A, Garcia-Sierra N, Prat C, Ruiz-Manzano J, Haba L, Rosés S, et al. GenoType MTBDRplus assay for molecular detection of rifampin and isoniazid resistance in *Mycobacterium tuberculosis* strains and clinical samples. J Clin Microbiol 2008;46:3660–7, doi:http://dx.doi.org/10.1128/JCM.00618-08.

Lee I, Ouk Kim Y, Park S-C, Chun J. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2016;66:1100–3, doi:http://dx.doi.org/10.1099/ijsem.0.000760.

Lu W, Feng Y, Wang J, Zhu L. Evaluation of MTBDR plus and MTBDR sl in detecting drug-resistant tuberculosis in a Chinese population. Dis Markers 2016;2016:, doi:http://dx.doi.org/10.1155/2016/2064765.

Mfinanga SG, Warren RM, Kazwala R, Ngadaya E, Kahwa A, Kazimoto T, et al. Genetic profile of *Mycobacterium tuberculosis* and treatment outcomes in human pulmonary tuberculosis in Tanzania. Tanzan J Health Res 2014;16:, doi:http://dx.doi.org/10.4314/thrb.v16i2.1.

Mnyambwa PN, Kim D-J, Ngadaya ES, Kazwala R, Petrucka P, Mfinanga SG. Assessment of sputum smear-positive but culture-negative results among newly diagnosed pulmonary tuberculosis patients in Tanzania. Int J Gen Med 2017a;199–205.

Mnyambwa NP, Kim D-J, Ngadaya ES, Kazwala R, Petrucka P, Mfinanga SG. Clinical implication of novel drug resistance-conferring mutations in resistant tuberculosis. Eur J Clin Microbiol Infect Dis 2017b;36:2021–8, doi:http://dx.doi.org/10.1007/s10096-017-3027-3.

Mpagama SG, Heysell SK, Ndusilo ND, Kumburu HH, Lekule IA, Kisonga RM, et al. Diagnosis and interim treatment outcomes from the first cohort of multidrug-resistant tuberculosis patients in Tanzania. PLoS One 2013;8:e62034, doi:http://dx.doi.org/10.1371/journal.pone.0062034.

Shahraki AH, Heidarieh P, Bostanabad SZ, Khosravi AD, Hashemzadeh M, Khandan S, et al. "Multidrug-resistant tuberculosis" may be nontuberculous mycobacteria. Eur J Intern Med 2015;26:279–84, doi:http://dx.doi.org/10.1016/j.ejim.2015.03.001.

Ministry of Health and Social Welfare. National Tuberculosis and Leprosy Program. Operational guidelines for the management of drug resistant TB in Tanzania. 1st ed. 2012.

Tortoli E, Rindi L, Garcia MJ, Chiaradonna P, Dei R, Garzelli C, et al. Proposal to elevate the genetic variant MAC-A, included in the *Mycobacterium avium* complex, to species rank as *Mycobacterium chimaera* sp. nov. Int J Syst Evol Microbiol 2004;54:1277–85, doi:http://dx.doi.org/10.1099/ijs.0.02777-0.

Tortoli E, Mariottini A, Pierotti P, Simonetti TM, Rossolini GM. *Mycobacterium yongonense* in pulmonary disease, Italy. Emerg Infect Dis 2013;19:1902–4, doi:http://dx.doi.org/10.3201/eid1911.130911.

van Ingen J, Boeree MJ, Kosters K, Wieland A, Tortoli E, Dekhuijzen PNR, et al. Proposal to elevate *Mycobacterium avium* complex ITS sequevar MAC-Q to *Mycobacterium vulneris* sp. nov. Int J Syst Evol Microbiol 2009;59:2277–82, doi:http://dx.doi.org/10.1099/ijs.0.008854-0.

van Soolingen D, Hermans PW, de Haas PE, Soll DR, van Embden JD. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. J Clin Microbiol 1991;29:2578–86.

World Health Organization. Guidelines for surveillance of drug resistance in tuberculosis. 4th ed. 2009.

World Health Organization. The use of molecular line probe assay for detection of resistance to second-line anti-tuberculosis drugs: Epert group meeting report Geneva. 2013.

Yeung MW, Khoo E, Brode SK, Jamieson FB, Kamiya H, Kwong JC, et al. Health-related quality of life, comorbidities and mortality in pulmonary nontuberculous mycobacterial infections: A systematic review. Respirology 2016;21:1015–25, doi:http://dx.doi.org/10.1111/resp.12767.