# Genomic Characterization of Non-Tuberculous Mycobacteria Using Whole Genome Sequencing

**Nadeem Gul Dar[1], Dr. Geeta Gupta[1], Dr. Nazia Khanum[2], Dr. Shagufta shahi[3], Dr. Zafar Nowshad Wani[3], Dr. Sachin Kishore[4], Dr. Shabir Ahmad Lone[5]**

[1]*Microbiology department, Santosh Medical College, Ghaziabad, UP, India-201009.*

[2]*Microbiology department, SVS medical college, Mahbubnagar, Telangana, India-509001.*

[3]*IRL (Intermediate reference laboratory, Chest Disease Hospital, Srinagar, Jammu and Kashmir, India-190001.*

[4]*Microbiology department, Autonomous Sate Medical College, Mirzapur, UP, India-231001.*

[5]*Microbiology department, Sher-i-Kashmir Institute of Medical Science, Srinagar, Jammu and Kashmir, India-190011.*

***Corresponding Author:** Dr. Sachin Kishore

Microbiology department, Autonomous Sate Medical College, Mirzapur, UP, India-231001.

**KEYWORDS**

- Whole genome sequencing
- Nontuberculous mycobacteria (NTM)
- Tuberculosis (TB)
- Mycobacterium tuberculosis complex (MTBC)

**ABSTRACT:**

**Introduction:**The study aims to identify and characterize nontuberculous mycobacteria at the molecular level in patients suspected of having tuberculosis, as standardized identification criteria are recommended by the American Thoracic Society for quicker clinical and laboratory procedures.

**Method:** 772 samples were collected from 386 patients (2 each) in selected districts of Jammu and Kashmir. In addition to phenotypic, molecular methods were also used to detect the number, species/ sub-species and AMR gene/SNP's detection of Non tuberculosis mycobacteria. Records of patients were collected for clinical information, such as symptoms and radiological findings.

**Results:**Out of 772 samples, 180 (23.31%) were positive for acid-fast bacteria, with 164 (21.24%) and 16 (2.07%) identified as M. tuberculosis complex and NTM respectively. Mycobacterium abscessus and M. intracellular isolates were the most frequent, identified by restriction endonuclease assay (REA-PCR) and whole genome sequencing. Furthermore, whole genome sequencing helps in detection of AMR gene and SNP's. Common symptoms included cough, fever, shortness of breath, weight loss, sputum production, appetite loss, night sweating, and thrombocytosis.

**Conclusion:**A study of 386 patients revealed that most were over 40 years old, with a higher rate of non-tuberculous mycobacteria (NTM) infections in males. Common symptoms included cough, fever, shortness of breath, weight loss, and sputum production. Positive cultures showed 21.24% MTB complexes and 2.07% NTM growth. Although biochemical testing is time-consuming, the study found five distinct mycobacterial species with 100% concordance across REA-PCR, whole genome sequencing, and biochemical tests. Whole genome sequencing not only helped find NTM at the species level but also at the sub-species level. It also aids in the discovery of AMR genes and SNPs.

## INTRODUCTION

Tuberculosis, an ancient disease, has been present in humans since prehistory. It may have first appeared around 150 million years ago[1]. Despite the first humans leaving Africa 1.7 million years ago[2], they likely brought TB with them. Written accounts date back as long as 2300 years in China and 3300 years in India[3,4].

Mycobacterium tuberculosis was once the only Mycobacterium infection in humans, causing significant social impact. Other mycobacterium species, known as anonymous or atypical mycobacteria, mycobacteria other than tuberculosis (MOTT), and non-tuberculous mycobacteria (NTM), are more common and have thicker, lipid-rich cell walls.

Mycobacteria, resistant to hydrophilic nutrients, heavy metals, antibiotics, and disinfectants, are found in various environments. Non-tuberculous mycobacteria (NTM) spread disease at varying rates, and host factors are now considered more significant in the pathophysiology of NTM infections, despite environmental variables being suspected.

Host-organism interaction can occur due to environmental exposure or host components[5,6]. Soil and water sources have high concentrations of NTM, which contributes to biofilm growth and antibiotic resistance. NTM's hydrophobicity allows preferential aerosolization from water, and many species can withstand high temperatures and low pH[7,8].

There is limited evidence that NTM is spread from human to human, unlike leprosy and tuberculosis. Since NTM infections cannot widely spread, they are not reported; yet, there are few surveillance data available. It has been demonstrated that over 140 distinct mycobacteria species are hazardous to humans. Among NTM species that cause sickness in humans, Mycobacterium avium complex (MAC) and Mycobacterium kansasii are the most common[9,10].

From the early 20th century, reports of non-tuberculous mycobacteria from clinical specimens were infrequent. However, in the 1950s, a new concept emerged, focusing on mycobacterial infection and the relationship between infection with organisms other than the tubercle bacillus and mild tuberculin responses. Two "novel" mycobacterial infections were described, establishing mycobacteria other than tunberculous bacilli as serious human illnesses.

NTM infections are not reportable in India, making it difficult to determine prevalence due to clinical, radiographic, and microbiological criteria. Factors such as immune-suppressive medications, immune deficiency disorders, and chronic structural lung illnesses contribute to the increase in NTM infections[11,12].

NTM infection can manifest in various areas, including the lung, skin, soft tissues, lymphadenitis, empyema, eye, central nervous system, and genitourinary infections. Initially believed to be untransferable, genetic evidence suggests human-to-human transmission[13]. Treatment typically involves antibiotics, with surgery for non-responsive patients, posing a high risk of complications[14].

Non-tuberculous mycobacterial pulmonary disease (NTM-PD) is a common comorbidity in patients with underlying respiratory diseases like bronchiectasis, cystic fibrosis, and chronic obstructive pulmonary disease[15]. Common NTM species causing NTM-PD include Mycobacterium avium complex, Mycobacterium kansasii, Mycobacterium xenopi, Mycobacterium abscessus, and Mycobacterium malmoense[16]. The discussion of pulmonary NTM disease in relation to tuberculosis is motivated by two factors: NTM-related lung illness shares symptoms with TB, and NTM isolates may become resistant to first-line anti-TB drugs[17].

Unfortunately, it can be challenging to diagnose NTM infections, and patients commonly receive a delayed or inaccurate diagnosis, which can have detrimental long-term effects. The purpose of this research was to identify and characterize non-tuberculous mycobacteria at the molecular level in people who may have tuberculosis.
Material and methods:

A study analyzed 772 clinical samples of 386 patients suspected of having tuberculosis between June 2021 and August 2024 in Jammu and Kashmir. The majority were male. Standardization from ATS/IDSA and American Thoracic Society was used to identify NTM isolates[18]. Patients with at least two positive cultures were included.

NTM identification using phenotypic and genotypic testing
The study involved decontaminating sputum specimens with N-acetyl-L-cysteine, staining them with Ziehl-Neelsen, and preparing them for Acid Fast Bacilli (AFB) smear microscopy. Cultured in Lowenstein-Jensen medium[19], they were incubated for eight weeks[20]. To differentiate between MTB and NTM[21], SD Bioline TB Ag was used to expose AFB-positive growth to the MPT64 antigen. PCR and biochemical testing were used to identify the species, all based on Centers for Disease Control (CDC) procedures for the isolation of NTM strains.
Immunochromatographic assay (ICA):
The SD Bio-Line MPT64 TB test was used for rapid immunochromatographic assay (ICA) to differentiate between MTB and NTM21. 200 µL of extraction buffer was used to emulsify three or four colonies of mycobacterial strains culture, adding 100 µL to each well, and visually assessing the results using color development after 15 minutes of incubation.
Whole genome sequencing

DNA extraction: DNA was extracted by using saline and bead beating. Bead beating is an effective method of cell lysis used to disrupt virtually any biological sample by rapidly agitating samples with a lysing matrix, sometimes referred to as grinding media or beads, in a bead beater. while saline solution maintains isotonic conditions, stabilize cells, solubilize components after lysis, and aid in removing proteins and contaminants during nucleic acid extraction.

DNA Quality check: The extracted DNA was quantified using Qubit dsDNA High sensitivity Assay kit (Invitrogen, Cat# Q32854) (Qubit Fluorometer 3):

Qubit fluorescence technology: Qubit fluorometers and assay kits are designed to measure the intensity of the signal from fluorescent dyes bound to specific biological molecules. These optimized dyes bind selectively to DNA, RNA, or protein and only emit a fluorescent signal when bound to the target. Qubit fluorometers use specialized curve-fitting algorithms to develop a calibration curve using standard samples with a known concentration. An unknown sample concentration of DNA, RNA, or protein is calculated by comparing the relative fluorescence units (RFUs) of the sample to the RFUs of the standards used in calibration. The detection limits of the measurements are specific to each assay.

Formats: standard assay and 1X assay • Invitrogen™ Qubit™ standard assays require same-day mixing of the buffer with the reagent to create the working solution prior to preparing standards and samples for quantification • Invitrogen™ Qubit™ 1X assays eliminate the step of preparing the working solution – The Invitrogen™ Qubit™ 1X dsDNA HS Assay Kit provides the same dynamic range and limit of detection as the standard assay, while the Invitrogen™ Qubit™ 1X dsDNA BR Assay Kit has a wider dynamic range than the standard assay, achieving 4,000 ng/μL in the extended range – This format offers a simplified workflow while reducing the tubes in the kit, therefore reducing the amount of plastic used – Simply add your sample or standard to the premixed solution, incubate, and read your results

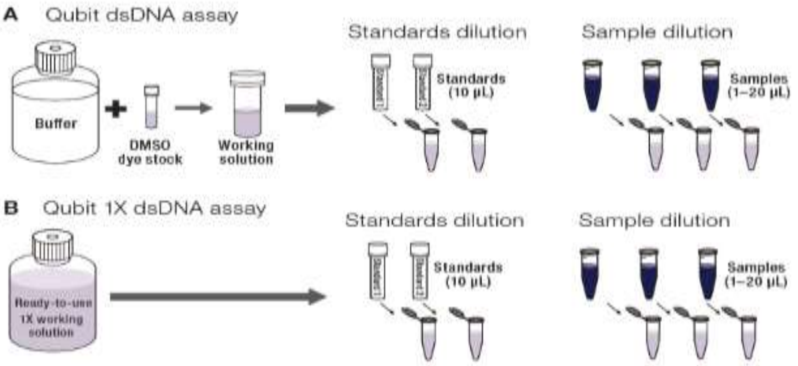



Fig 1: Workflow comparison for the (A) Qubit dsDNA and (B) Qubit 1X dsDNA assays. Standard Qubit dsDNA High Sensitivity (HS) and Qubit dsDNA Broad Range (BR) assay kits include a fluorogenic dye, buffer, and dsDNA standards. Prior to each assay, a fresh aqueous working solution needs to be prepared by diluting the dye stock in the provided buffer in a 1:200 ratios. Qubit 1X dsDNA assay kits eliminate this step by providing a ready-to-use working solution.

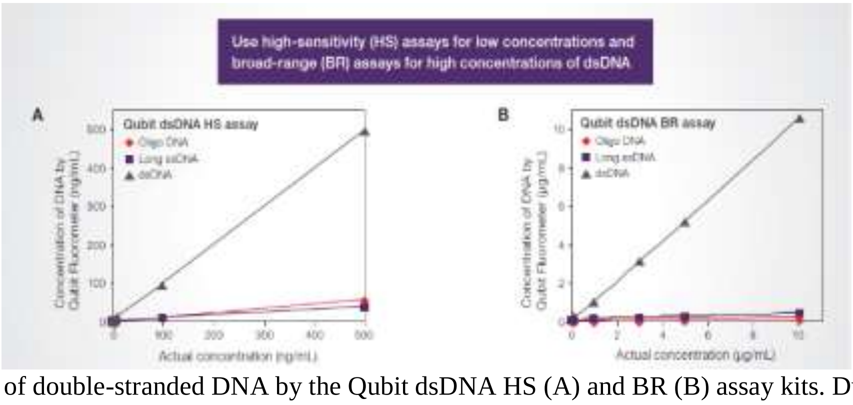



Fig 2: Detection of double-stranded DNA by the Qubit dsDNA HS (A) and BR (B) assay kits. Duplicate samples of long ssDNA, oligo DNA, or lambda dsDNA at concentrations of 0.5 to 500 ng/mL in the assay tube were quantified using the Qubit dsDNA HS assay, and at concentrations of 0.01 to 10 μg/mL in the assay tube using the Qubit dsDNA BR assay according to kit protocols.

Library Preparation: The DNA sample is further taken for DNA library preparation using Twist Library Preparation Enzymatic Fragmentation (EF) Kit (TWIST Biosciences, Cat. No: 104175).

The EF Kit includes the reagents required for end-repair, dA-tailing, adapter ligation, and library amplification. This Kit also incorporates the enzymes for fragmentation of gDNA samples and allows for tunable sample sizes. Following core library construction, either full-length or Twist Universal Adapaters can be used to suit your application needs

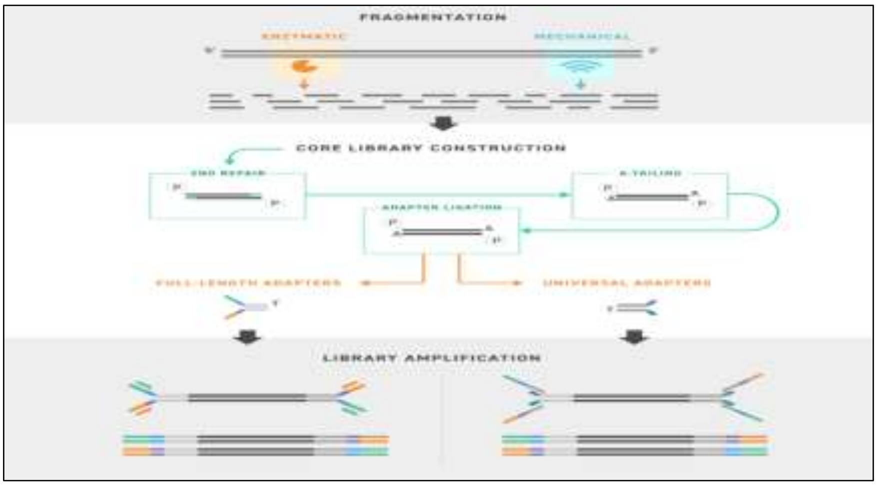


Figure 1: Library preparation workflow

gDNA SAMPLES:
Use the Thermo Fisher Scientific Qubit dsDNA Broad Range Quantitation Assay to accurately quantify input purified gDNA.
Measuring DNA concentration by absorbance at 260 nm is not recommended.
Input DNA should be suspended in Molecular Biology Grade Water, 10 ml Tris-HCl pH 8.0, or Buffer EB.
It is important to remove all cations and chelators from the starting gDNA sample. The presence of cations and chelators may affect the initial fragmentation reaction.
The recommended DNA input is 50 ng of high quality gDNA.
Reagents are compatible with mass input of 1 ng to 500 ng, but may require optimization of the following steps in library preparation to achieve optimal performance:
Incubation Time for Fragmentation:
Program the thermal cycler with the following conditions. Use the Incubation Time table below to select conditions for fragmentation to achieve the desired insert size. Set the temperature of the heated lid to 105°C. Start the program to pre-chill the thermal cycler.
NOTE: Fragmentation temperature (Step 2) does not impact library performance

INCUBATION TIME*

| DESIRED INSERT SIZE (BP) | @30°C | @37°C |
|---|---|---|
| 145–175 | — | 30 min |
| 180–220 | — | 20 min |
| 250–300 | — | 10 min |
| 275–350 | 15 min | — |
| 350–425 | 10 min | — |

*20 min at 37ºC is recommended for Twist target enrichment applications utilizing 50 ng of high quality gDNA. Additional conditions in the Step 2 Incubation Time table were also derived using high quality gDNA and should be optimized for each sample type/application. See Appendix C for additional guidance.
Amount of Twist Universal Adapter:
Add 5µl Twist Universal Adapters into each sample well or tube containing the dA-tailed DNA fragments from Step 1. Mix gently by pipetting and keep on ice.
Incubation Time for Ligation Reaction:
Incubate the ligation reaction at 20°C for 15 minutes in the thermal cycler, then move the samples to the bench top. Proceed to the Purify step.
IMPORTANT: Turn off the heated lid or set to minimum temperature.
NOTE: While the thermal cycler program is running, prepare the reagents for Step 3: PCR Amplify Using Twist UDI Primers, Purify, and Perform QC
PCR cycles for Amplification:

Program a thermal cycler with the following conditions. Set the temperature of the heated lid to 105°C.

| STEP | | TEMPERATURE | TIME | NUMBER CYCLES |
|---|---|---|---|---|
| 1 | Initialization | 98°C | 45 seconds | 1 |
| 2 | Denaturation | 98°C | 15 seconds | 6-8* |
| | Annealing | 60°C | 30 seconds | |
| | Extension | 72°C | 30 seconds | |
| 3 | Final Extension | 72°C | 1 minute | 1 |
| 4 | Final Hold | 4°C | HOLD | – |

*6–8 cycles are recommended for Twist target enrichment workflows when starting with 50 ng high quality gDNA.
Library QC: The prepared library is checked for fragment distribution by 5300 Fragment analyzer system (Agilent)
The 5300 Fragment Analyzer systems use parallel capillary electrophoresis technology for precise RNA and DNA analysis, offering a range of kits and the ability to process 12, 48, or 96 samples simultaneously, enhancing lab productivity. The 5300 Fragment Analyzer System automates nucleic acid analysis through key steps: sample preparation (extract, quantify, dilute, and mix with loading buffer and ladder), instrument loading (samples, gel matrix, marker solution, and electrophoresis cassette), software setup (method file selection and plate layout configuration via ProSize Software), running the analysis (capillary electrophoresis with results displayed as electropherograms or gel images), and data analysis (evaluation of fragment size, RNA integrity, and library quality, with results exported as needed).
Sequencing: The library is sequenced in Novaseq X Plus Illumina platform. The read chemistry for Novaseq X plus is 2X 150bp. Data generated for NTM species is 1.2gb/sample.

Bioinformatics analysis:


Figure 2: Bioinformatics analysis workflow

Read quality check.
Initially, we will check the following parameters from the sample fastq file –
Base quality score distribution, sequence quality score distribution, average base content per read, GC distribution in the reads, PCR amplification issue, overrepresented sequences and adapters.
Based on the quality report, the fastq files will be trimmed to retain high-quality sequences and the low-quality sequence reads will be excluded from the analysis using the fastq-mcf tool.

Denovo Metagenome assembly
Unaligned reads after contamination removal step was de novo assembled using Spades (version. 3.11.1). Assembly was performed with k-mer size 55 using de-Bruijn graph method. The assembled scaffolds were taken for further downstream analysis.

ORF prediction and annotation
The Primarily assembled genome will be used for ORF prediction and ORF annotation using Prodigal.

Blast
The NCBI blast was performed for the primarily assembled genome to predict the closest reference genome of each sample.

Reference-guided assembly
The genome fasta of the predominant species (predicted from above step) was downloaded from NCBI database and reference-guided assembly was performed. Consensus genome fasta was generated using samtools and Bcftools.

Variant Prediction and annotation
The Reference aligned reads were used to predict variants using GATK and the variants were annotated using the snpEff.

AMR prediction
AMR (Anti-microbial resistance) prediction was performed using the RGI tool (version 6.0.0) and database used was Comprehensive Antibiotic Resistance Database (CARD) https://card.mcmaster.ca/, a biological database that collects and organizes reference information on antimicrobial resistance genes, proteins and phenotypes. The cutoff used in the RGI tool were perfect (100% identity), strict (>=95% identity) and loose (< 95% identity).

**RESULTS**
The study collected demographic data from 386 patients, categorized into three age groups based on ten-year intervals. The highest age group was those over 40 years, with 178 cases. The 30-40year age group had 116 cases, and the 19-29year age group had 92 cases.

Table 3: Age group wise distribution of suspected cases of pulmonary tuberculosis

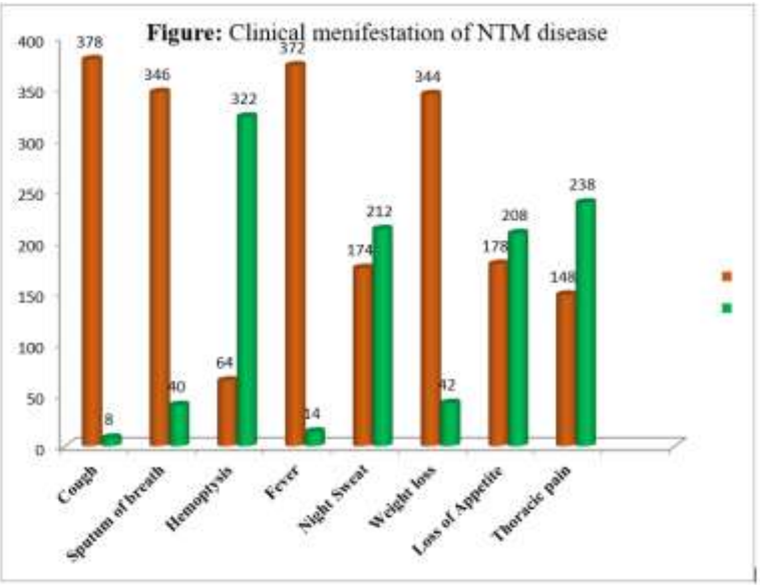
| Age groups | No. of cases | Percentage |
|---|---|---|
| 19-29 | 92 | 23.83 |
| 30-40 | 116 | 30.05 |
| >40 | 178 | 46.11 |
| Total | 386 | 100 |

The study revealed a higher prevalence of NTM infections among males, with 204 males and 182 females identified out of 386 patients.

Table 2: Gender wise distribution of suspected cases of pulmonary tuberculosis

| **Gender** | **No. of cases** | **Percentage** |
|---|---|---|
| Male | 204 | 52.85 |
| Female | 182 | 47.15 |
| Total | 386 | 100 |

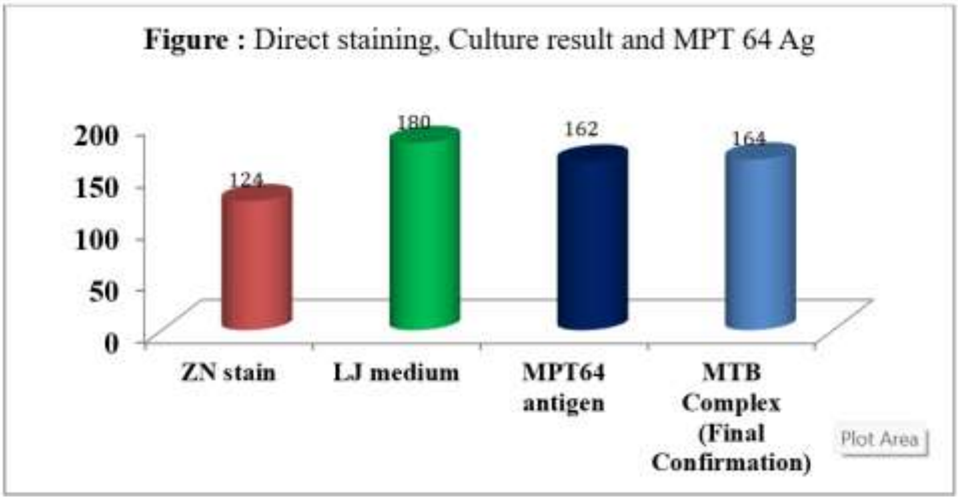
Mycobacterium infection symptoms include cough, fever, shortness of breath, weight loss, sputum production, appetite loss, night sweat, Thoracic pain, and hemoptysis.




The study involved collecting 772 sputum samples using a modified Pertroff's method. After decontamination, the samples were inoculated onto Lowenstein-Jensen medium slants and incubated at 37°C for up to four weeks. 124 samples tested positive for ZN staining, and 180 samples contained positive mycobacteria. The MPT 64 antigen-based immunological test was used to analyze the presence of MTB complexes or non-tuberculous mycobacteria (NTM) growth. Biochemical tests, including niacin, nitrate, and heat-resistant catalase, confirmed the positive growth findings. The results showed that 2.07% of the samples were NTM growth, while 21.24% were identified as M. tuberculosis complexes.

Table 5: Identification of Mycobacterium species by various methods

| Test method | Test Positive |
|---|---|
| Direct ZN staining | 124/772 (62 out of 386 patients) |
| Culture on LJ media | 180/772 (90 out of 386 patients) |
| MPT64 antigen | 162/180 (81 out of 90 patients) |
| MTB Complex | 164/180 (82 out of 90 patients) |
| NTM | 16/180 (8 out of 90 patients) |



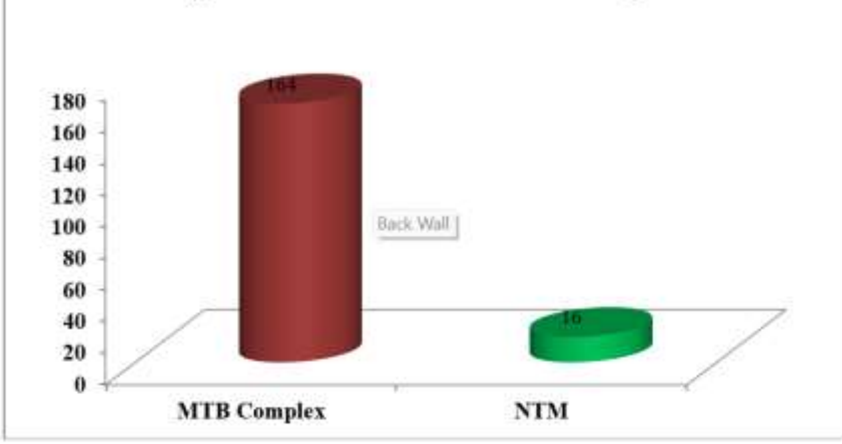



**Figure:** Detection of NTM and MTB complex

Lung fibrosis, observed in 88.86% of cases, is the most prevalent clinical manifestation, followed by cervical lymphadenopathy at 68.65%, nodularity at 61.40%, and unilateral lung involvement at 60.88%.

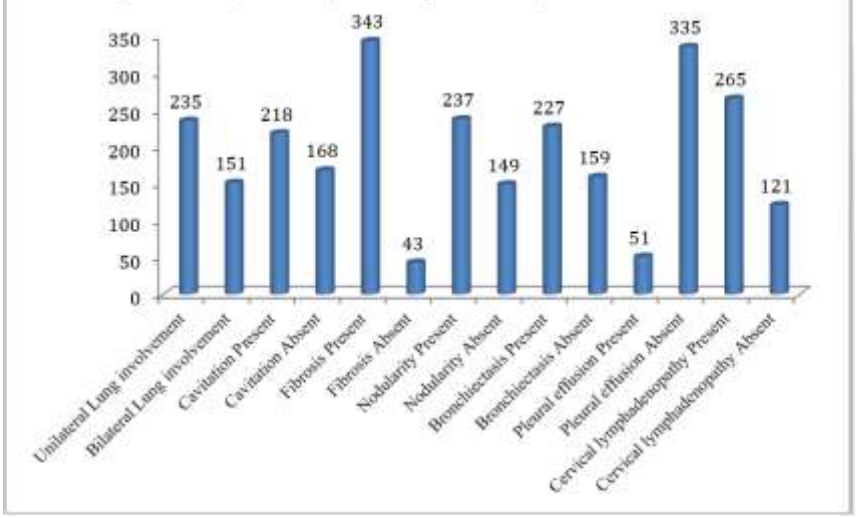



**Figure:** X-ray findings of lung with suspected TB infections

**Molecular Diagnosis of NTM**

Whole Genome Sequencing:
Among the eight samples analyzed, three were identified as M. abscessus, and two were categorized as Mycobacterium sp. Furthermore, sub species level identification, three were identified as Mycobacteroidesabscessus subsp. abscessus strain T9277, Mycobacterium intracellulare subsp. chimaera strain FLAC0070, Mycobacterium intracellulare subsp. chimaera strain 852, Mycobacterium chelonae, Mycobacterium fortuitum, Mycobacterium kansasii, highlighted in Table 1. The Phred quality score (Q30), was consistently high across all samples, ranging from 91% to 93%. This indicates that more than 90% of bases had an error probability of less than 1 in 1,000, ensuring reliable downstream analysis. High Phred scores are a cornerstone for high-confidence variant calling and AMR gene detection, particularly in clinical settings where accuracy is paramount. The GC content of the samples ranged from 62.85% to 66.065%, consistent with the known genomic characteristics of mycobacterial species. The read quality scores were uniformly high for all samples, reflecting the overall integrity of the sequencing process. Data sizes varied, from 1.05 GB to 2.5 GB.

Alignment of sequencing reads to reference genomes is a critical step for species identification and detection of genetic variations. Samples exhibited alignment percentages ranging from 86.44% to 94.19%. This highlights the availability of well- characterized reference genomes for M. abscessus, facilitating accurate strain identification. The coverage (≥30x) ranged from 86.19% to 95.42%. High coverage ensures the reliability of SNP detection and AMR gene identification; as low coverage can lead to false negatives or misinterpretation of variants. Sequencing depth varied from 83.45x to 185.88x, with deeper coverage observed in Mycobacterium sp. isolates. This depth allows for high-confidence identification of low-frequency variants, which are often clinically significant in heterogeneous infections or mixed populations.

Antimicrobial Resistance Gene Identification and Mechanisms

AMR genes were detected in five samples, reflecting the significant resistance potential of both Mycobacteroides and Mycobacterium species. In the M. abscessus group, AMR determinants included:

1. vanY Gene in the vanB Cluster: Present in samples number 1, 2, and 3, this gene confers resistance to glycopeptide antibiotics, such as vancomycin. The presence of glycopeptide resistance genes underscores the challenge of treating infections caused by M. abscessus.
2. Beta-Lactamase Genes: Identified in all M. abscessus isolates, these genes encode enzymes that hydrolyze beta-lactam antibiotics, including cephalosporins and penems. This highlights the limited efficacy of these drug classes against M. abscessus.
3. Erm(41) Gene: Found in sample number 2 and 3, this gene encodes a 23S rRNA methyltransferase, conferring macrolide resistance. This is clinically significant, as macrolides are commonly used to treat M. abscessus infections.
4. A2059G Mutation in 23S rRNA: This SNP, present in sample number 2 and 3, alters the antibiotic binding site, further contributing to macrolide resistance. In contrast, Mycobacterium sp. isolates (sample number 4 and 5) exhibited distinct AMR profiles.
5. mtrA Gene: This efflux pump gene mediates resistance by actively transporting antibiotics out of the bacterial cell. It confers resistance to macrolides and phosphonic acid antibiotics, reducing drug efficacy.
6. C1170 Mutation: This SNP in mtrA enhances its function, providing additional resistance.

The presence of diverse AMR mechanisms, including target alteration, antibiotic inactivation, and efflux pumps, underscores the genomic adaptability of these pathogens and the importance of comprehensive resistance profiling for effective treatment planning.

Sequencing enabled precise strain differentiation, particularly within the M. abscessus group. For example, sample number 2 was identified as M. abscessus subsp. abscessus strain GD54, while 3 was classified as M. abscessus strain FLAC049. Such differentiation is critical in clinical settings, as different strains may exhibit varying resistance profiles and virulence factors.

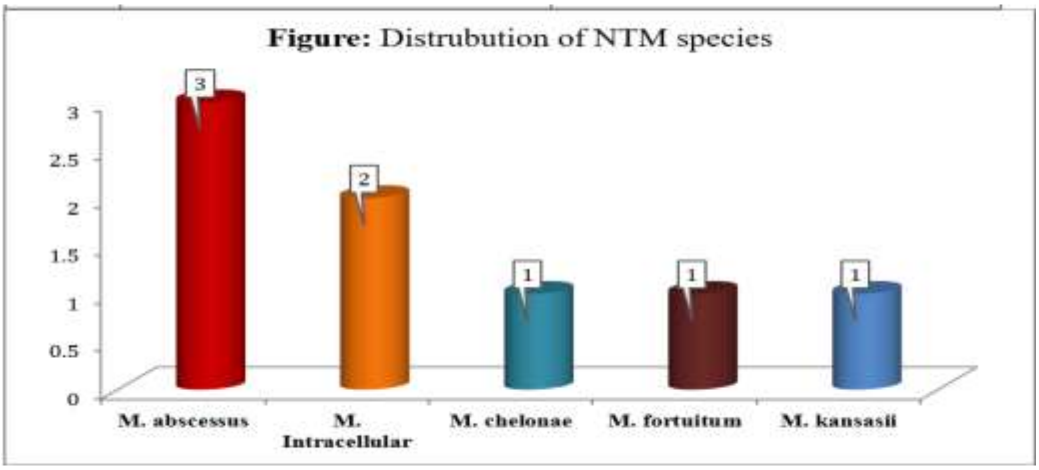
Comparative Analysis of the Samples

The comparative analysis of the eight samples revealed distinct genomic and AMR characteristics:

Sequencing Quality and Coverage: While all samples exhibited high Phred scores, Mycobacterium sp. isolates (sample number 4 and 5) had larger data sizes, higher coverage, and greater sequencing depth compared to M. abscessus. This reflects the larger genome size and complexity of Mycobacterium sp., necessitating deeper sequencing for comprehensive analysis.

AMR Profiles: M. abscessus isolates shared common resistance determinants, such as vanY and beta-lactamase genes, while Mycobacterium sp. isolates exhibited distinct efflux pump-mediated resistance. The identification of specific SNPs, such as A2059G and C1170, further differentiated the samples.

| Sr. No. | Organism identified | Sub Species level identification |
|---|---|---|
| 1 | Mycobacteroidesabscessus subsp. abscessus strain GD25 | N.A. |
| 2 | Mycobacteroidesabscessus subsp. abscessus strain GD54 | N.A. |
| 3 | Mycobacteroidesabscessus strain FLAC049 | Mycobacteroidesabscessus subsp. abscessus strain T9277 |
| 4 | Mycobacterium sp. 5-140-3-1 | Mycobacterium intracellulare subsp. chimaera strain FLAC0070 |
| 5 | Mycobacterium sp. 5-140-3-1 | Mycobacterium intracellulare subsp. chimaera strain 852 |
| 6 | Mycobacterium chelonae | N.A. |
| 7 | Mycobacterium fortuitum | Mycobacterium fortuitum subsp. fortuitum |
| 8 | Mycobacterium kansasii | N.A. |




**DISCUSSION**

The advent of whole genome sequencing (WGS) has significantly transformed our understanding of non-tuberculous mycobacteria (NTM), particularly in the context of antimicrobial resistance (AMR) and strain differentiation. The current study presents a comprehensive analysis of Non-Tuberculous Mycobacteria (NTM) through whole genome sequencing, providing valuable insights into the genomic characteristics and antimicrobial resistance (AMR) profiles of various mycobacterial species. The findings not only underscore the utility of WGS in clinical microbiology but also the importance of advanced genomic techniques in enhancing our understanding of these pathogens, which are increasingly recognized for their clinical significance and resistance challenges.

Genomic Insights and Sequencing Quality

The high-quality sequencing results, characterized by high Phred quality scores (Q30) consistently above 90% across all samples (91% to 93%), affirm the reliability of the data generated. Such high scores are essential for ensuring accurate downstream analyses, particularly in variant calling and AMR gene detection, which are critical in clinical microbiology settings where treatment decisions hinge on precise pathogen characterization[22]. The observed GC content, ranging from 62.85% to 66.065%, aligns with established genomic profiles of mycobacterial species [22, 23] reinforcing the validity of our sequencing approach and the integrity of the extracted DNA.

In terms of alignment and coverage, our results demonstrated alignment percentages ranging from 86.44% to 94.19%, with coverage exceeding 30x in all samples. This high coverage is critical for the reliable detection of single nucleotide polymorphisms (SNPs) and AMR genes, as low coverage can lead to false negatives or misinterpretation of variants [24]. It demonstrate the robustness of the reference genomes available for M. abscessus, facilitating accurate strain identification. This is particularly relevant given the clinical implications of distinguishing between closely related strains, which may exhibit divergent resistance profiles and virulence factors[24]. The depth of coverage, exceeding 30x, is crucial for the reliable detection of single nucleotide polymorphisms (SNPs) and other genetic variations[25]. The sequencing depth observed (83.45x to 185.88x) provides a strong foundation for identifying low-frequency variants, which are particularly relevant in the context of heterogeneous infections.

Antimicrobial Resistance Profiles

The identification of diverse AMR mechanisms within the M. abscessus group is of particular concern. Presence of diverse AMR genes among the isolates, highlights the complex resistance mechanisms employed by these pathogens. The presence of the vanY gene, conferring resistance to glycopeptides, alongside beta-lactamase genes, highlights the increasing difficulty in treating infections caused by M. abscessus, which is known for its multidrug-resistant nature[26]. This aligns with findings from previous studies that have documented the alarming rise of multidrug-resistant mycobacteria and the associated clinical challenges[27]. The identification of Erm(41) and the A2059G mutation in 23S rRNA further further illustrates the genetic adaptability of these organisms, contributing to macrolide resistance—a critical concern given that macrolides are frequently used in therapeutic regimens for NTM infections[28,29] This emphasizes the complexity of resistance mechanisms that can evolve within these pathogens, necessitating ongoing surveillance and tailored therapeutic strategies.

In contrast, the distinct AMR profiles observed in Mycobacterium sp. isolates, particularly the efflux pump-mediated resistance, underscore the genomic adaptability of these organisms. This adaptability is critical in the context of rising AMR, as it allows for the rapid acquisition and dissemination of resistance traits, complicating treatment efforts and highlighting the need for alternative therapeutic strategies[30,31]. The identification of the mtrA gene and its C1170 mutation demonstrates the multifaceted nature of AMR in mycobacteria, where mechanisms can vary significantly even among closely related species. This finding aligns with previous studies that have documented the role of efflux pumps in conferring resistance to a broad spectrum of antibiotics[30,28].

Comparative Analysis and Clinical Implications

The implications of these findings are profound, particularly in light of the rising incidence of NTM infections and the associated treatment challenges. The resistance profiles exhibited by the M. abscessus isolates suggest a limited arsenal of effective therapeutic options, necessitating the exploration of alternative treatments and the development of novel antibiotics. The comparative analysis of the samples reveals significant genomic and AMR characteristics that can inform clinical practices. The shared resistance determinants among M. abscessus isolates indicate a pressing need for the development of new therapeutic agents and treatment regimens that can effectively target these resistant strains. The distinct resistance mechanisms exhibited by Mycobacterium sp. isolates further suggest that treatment strategies need to be tailored not only to the species but also to the specific resistance profiles identified and tailoring treatment regimens to individual patient needs.

The transformative role of sequencing in clinical microbiology is evident in this study. By providing high-resolution genomic data, sequencing facilitates the precise identification of pathogens, differentiation of closely related strains, and comprehensive profiling of AMR mechanisms. This information is invaluable for guiding antibiotic selection, monitoring the emergence of resistance, and designing targeted treatment regimens. Moreover, the identification of resistance genes and mutations contributes to the global understanding of AMR, informing public health strategies aimed at combating this growing challenge.

Moreover, the capability of WGS to facilitate precise strain differentiation is invaluable in clinical microbiology. The identification of specific strains, such as M. abscessus subsp. abscessus strain GD54 and strain FLAC049, is essential for understanding the epidemiology of NTM infections and their associated resistance profiles. This level of detail can inform infection control measures and public health strategies aimed at curbing the spread of resistant strains[32].

Future Directions

As sequencing technologies continue to evolve, their application in clinical settings is likely to expand, offering new opportunities for precision medicine and improved patient outcomes. Future studies should focus on integrating genomic data with clinical outcomes to better understand the implications of AMR profiles on treatment efficacy. The integration of WGS into routine clinical practice could enhance our ability to monitor the emergence of AMR and inform public health strategies aimed at mitigating the impact of resistant infections. Future research should focus on exploring the role of mobile genetic elements in the dissemination of resistance genes, the longitudinal tracking of resistance patterns and the potential for horizontal gene transfer among mycobacterial populations, which could further complicate treatment efforts, will be crucial for developing strategies to mitigate the spread of AMR among mycobacterial populations

## CONCLUSION

A study involving 386 patients found that the majority were over 40 years old, with a higher rate of non-tuberculous mycobacteria (NTM) infections observed in males. Common symptoms of mycobacterium infection included cough, fever, shortness of breath, weight loss, sputum production, loss of appetite, night sweats, thoracic pain, and hemoptysis. Sputum samples were collected and processed for culture, with negative cultures confirmed by smear examination. Positive cultures were further evaluated for mycobacterial growth using immunochromatography and biochemical tests. The results showed that 21.24% of positive cultures were MTB complexes, while 2.07% were identified as NTM growth. This study highlights the power of whole genome sequencing to unravel the genomic and resistance landscapes of non-tuberculous mycobacteria. The insights gained not only enhance our understanding of NTM but also underscore the urgent need for innovative approaches to address the challenges posed by antimicrobial resistance in clinical practice. It will be pivotal in shaping future approaches to the management of NTM infections and improving patient outcomes

**Funding:**

No specific grant was obtained for this research from governmental, private, or nonprofit funding organizations.

**Ethics statement**

This study was approved by the Institutional Ethics Committee, Santosh Medical College Ghaziabad, UP, India (Letter No: SU/2019/1531[5])

**Conflict of interest:**

The authors declare that there is no conflict of interest.

## REFERENCES

1. Daniel T M. History of tuberculosi. Respiratory Med 2006;100:1862-1870.

2. Zimmerman MR. Pulmonary and osseous tuberculosis in an Egyptian mummy. Bull NY Acad Med 1979;55:604–8
3. Morse D, Brothwell DR, Ucko PJ. Tuberculosis in ancient Egypt. Am Rev Respir Dis 1964;90:5224–541.
4. Immerman MR. Pulmonary and osseous tuberculosis in an Egyptian mummy. Bull NY Acad Med 1979;55:604–8.
5. Nerlich AG, Haas CJ, Zink A, Szeimies U, Hagedorn HG. Molecular evidence for tuberculosis in an ancient Egyptian mummy. Lancet 1997;350:1404.
6. Crube´zy E´, Ludes B, Poveda J-D, Clayton J, Crouau-Roy B, Montagnon D. Identification of Mycobacterium DNA in an Egyptian Pott‘s disease of 5400 years old. C R Acad Sci Paris (Sciences de la vie) 1998;321:941–51.
7. Daniel VS, Daniel TM. Old testament biblical references to tuberculosis. Clin Infect Dis 1999;29:1557–8.
8. Daniel TM. The early history of tuberculosis in central East Africa: insights from the clinical records of the first twenty years of Mengo Hospital and review of the relevant literature. Int J Tuberc Lung Dis 1998;2:1–7.
9. Gibbons A. Modern men trace ancestry to African migrants. Science 2001;292:1051– 2.
10. Brown L. The story of clinical pulmonary tuberculosis. Baltimore, MD: Williams & Wilkins Company; 1941. 198
11. Winthrop KL. Pulmonary Disease Due to Nontuberculous mycobacteria: an Epidemiologist's View. Future Microbiology. 2010 Mar 8;5(3):343–5.
12. Stout JE, Koh WJ, Yew WW. Update on pulmonary disease due to non-tuberculous mycobacteria. International Journal of Infectious Diseases. 2016 Apr;45:123–34.
13. Bryant JM, Grogono DM, Rodriguez-Rincon D, Everall I, Brown KP, Moreno P, et al. Emergence and spread of a human-transmissible multidrug-resistant nontuberculous mycobacterium. Science [Internet]. 2016 Nov 11 [cited 2020 Apr 20];354(6313):751–7. Available from: https://science.sciencemag.org/content/sci/354/6313/751.full.pdf
14. Hyung Koo Kang, Hye Yun Park, Kim D, Jeong BH, Jeon K, Jong Ho Cho, et al. Treatment outcomes of adjuvant resectional surgery for nontuberculous mycobacterial lung disease. BMC Infectious Diseases. 2015 Feb 19;15(1).
15. The Voice of the Patient Non-Tuberculous Mycobacterial (NTM) Lung Infection [Internet]. 2016 Apr. Available from: https://www.fda.gov/media/96932/download
16. Griffith DE, Aksamit T, Brown-Elliott BA, Catanzaro A, Daley C, Gordin F, et al. An Official ATS/IDSA Statement: Diagnosis, Treatment, and Prevention of Nontuberculous Mycobacterial Diseases. American Journal of Respiratory and Critical Care Medicine. 2007 Feb 15;175(4):367–416.
17. Payam Tabarsi, Parvaneh Baghaei, Parissa Farnia, Mansouri N, Ehsan Chitsaz, Fatemeh Sheikholeslam, et al. Nontuberculous Mycobacteria Among Patients Who are Suspected for Multidrug-Resistant Tuberculosis—Need for Earlier Identification of Nontuberculosis Mycobacteria. The American Journal of the Medical Sciences. 2009 Mar 1;337(3):182–4.
18. Aksamit T, Brown-Elliott B, Catanzaro A, Daley C, Gordin F, Holland S, Horsburgh R, Huitt G, Iademarco M, Iseman M (2007) ATS Mycobacterial diseases subcommittee american thoracic society, infectious disease society of america an ofcial ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. Am J Respir Crit Care Med 175(4):367416.
19. Kent PT. (1985). Public health mycobacteriology: a guide for the level III laboratory: US Department of Health and Human Services, Public Health Service, Centers.
20. Mackie &Mccartney, J.G Collie, A.G.Fraser, B.P. Marmion, A.Simmons. Practical Medical Microbiology. Eigth edition. Page: Mycobacterium tubercle bacilli.Page: 412.
21. Ribon W. Biochemical Isolation and Identification of Mycobacteria. In Jimenez-Lopez JC. (editor) Biochemical Testing: InTech 2012. doi: 10.5772/34309.
22. Bouso JM, Planet PJ. Complete nontuberculous mycobacteria whole genomes using an optimized DNA extraction protocol for long-read sequencing. BMC Genomics. 2019 Oct 30;20(1):793. doi: 10.1186/s12864-019-6134-y. PMID: 31666009; PMCID: PMC6822416.
23. Fedrizzi T, Meehan CJ, Grottola A, Giacobazzi E, FregniSerpini G, Tagliazucchi S, Fabio A, Bettua C, Bertorelli R, De Sanctis V, Rumpianesi F, Pecorari M, Jousson O, Tortoli E, Segata N. Genomic characterization of Nontuberculous Mycobacteria. Sci Rep. 2017 Mar 27;7:45258. doi: 10.1038/srep45258. PMID: 28345639; PMCID: PMC5366915.
24. Yoon, JK., Kim, T.S., Kim, JI. et al. Whole genome sequencing of Nontuberculous Mycobacterium (NTM) isolates from sputum specimens of co-habiting patients with NTM pulmonary disease and NTM isolates from their environment. BMC Genomics 21, 322 (2020). https://doi.org/10.1186/s12864-020-6738-2
25. Khieu V, Ananta P, Kaewprasert O, Laohaviroj M, Namwat W, Faksri K. Whole-Genome Sequencing Analysis to Identify Infection with Multiple Species of Nontuberculous Mycobacteria. Pathogens. 2021 Jul 11;10(7):879. doi: 10.3390/pathogens10070879. PMID: 34358029; PMCID: PMC8308675.

26. Zeng D, Debabov D, Hartsell TL, Cano RJ, Adams S, Schuyler JA, McMillan R, Pace JL. Approved Glycopeptide Antibacterial Drugs: Mechanism of Action and Resistance. Cold Spring Harb Perspect Med. 2016 Dec 1;6(12):a026989. doi: 10.1101/cshperspect.a026989. PMID: 27663982; PMCID: PMC5131748.
27. Chin KL, Anibarro L, Chang ZY, Palasuberniam P, Mustapha ZA, Sarmiento ME, Acosta A. Impacts of MDR/XDR-TB on the global tuberculosis epidemic: Challenges and opportunities. Curr Res Microb Sci. 2024 Oct 19;7:100295. doi: 10.1016/j.crmicr.2024.100295. PMID: 39512261; PMCID: PMC11541417.
28. Solanki P, Lipman M, McHugh TD and Satta G (2022) Whole genome sequencing and prediction of antimicrobial susceptibilities in non-tuberculous mycobacteria. Front. Microbiol. 13:1044515. doi: 10.3389/fmicb.2022.1044515
29. L. Elaine Epperson, Rebecca M. Davidson, Sara M. Kammlade, Nabeeh A. Hasan, Sophie E. Nick, Iara M.P. Machado, Valerie H. Rodriguez, Aaron Appleman, Niels K. Helstrom, Michael Strong, Evaluation of the GenoType NTM-DR line probe assay for nontuberculous mycobacteria using whole genome sequences as reference standard, Diagnostic Microbiology and Infectious Disease, Volume 110, Issue 4, 2024, 116526, ISSN 0732-8893, https://doi.org/10.1016/j.diagmicrobio.2024.116526.
30. Dohál, M.; Porvazník, I.; Soloviˇc, I.; Mokrý, J. Whole Genome Sequencing in the Management of Non-Tuberculous Mycobacterial Infections. Microorganisms 2021, 9, 2237. https://doi.org/10.3390/ microorganisms9112237
31. Saxena S, Spaink HP, Forn-Cuní G. Drug Resistance in Nontuberculous Mycobacteria: Mechanisms and Models. Biology (Basel). 2021 Jan 29;10(2):96. doi: 10.3390/biology10020096. PMID: 33573039; PMCID: PMC7911849.
32. Jauneikaite E, Baker KS, Nunn JG, Midega JT, Hsu LY, Singh SR, Halpin AL, Hopkins KL, Price JR, Srikantiah P, Egyir B, Okeke IN, Holt KE, Peacock SJ, Feasey NA; SEDRIC Genomics Surveillance Working Group. Genomics for antimicrobial resistance surveillance to support infection prevention and control in health-care facilities. Lancet Microbe. 2023 Dec;4(12):e1040-e1046. doi: 10.1016/S2666-5247(23)00282-3. Epub 2023 Nov 14. PMID