



## Identification of a Novel Nucleic Acid Target for the Rapid and Specific Detection of *Mycobacterium Simiae* Using Comparative Genomic Analysis

Reza Kamali Kakhki <sup>1†</sup>, Mohammad Abavisani <sup>2†</sup>, Kiarash Ghazvini <sup>3</sup>, Hosna Zare <sup>2,4</sup>,  
Noshin Hojatpanah <sup>2,3</sup>, Jamal Falahi <sup>5</sup> and Alireza Neshani <sup>2,4\*</sup>

1. Mashhad Gene Azma Inc., Mashhad, Iran

2. Student Research Committee, Mashhad University of Medical Sciences, Mashhad, Iran

3. Department of Microbiology and Virology, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

4. Department of Laboratory Sciences, School of Paramedical Sciences, Mashhad University of Medical Sciences, Mashhad, Iran

5. Department of Laboratory Sciences, School of Paramedical Sciences, Torbat Heydariyeh University of Medical Sciences, Torbat Heydariyeh, Iran

† The first and the second authors have had equal contribution to this manuscript

### Abstract

**Background:** *Mycobacterium simiae* (*M. simiae*) is a non-tuberculous mycobacterium (NTM) that closely resembles *Mycobacterium tuberculosis* (*M. tuberculosis*) in clinical and biochemical characteristics, notably its niacin-positive phenotype. This similarity frequently leads to misdiagnosis and inappropriate treatment with first-line anti-tuberculosis drugs, to which *M. simiae* is frequently resistant. Current diagnostic methods are expensive or need complex equipment, highlighting the urgent need for a rapid, specific, and accessible molecular target to identify *M. simiae* accurately.

**Methods:** In this study, a modified genome comparison method was applied to the complete reference genome of *M. simiae* (AP022568.1) in order to identify a putative species-specific nucleotide sequence. A conventional PCR assay was designed to amplify a 168-bp fragment within this target, designated MST601 (*M. simiae* Target, 601 bp). The analytical sensitivity [Limit of Detection (LOD)] was determined using serial dilutions of genomic DNA. The pilot evaluation of the assay was assessed using 10 well-characterized clinical isolates of *M. simiae*.

**Results:** The MST601-PCR assay demonstrated high analytical sensitivity, with a limit of detection of ~10 fg (~2 genome equivalents per reaction) of *M. simiae* genomic DNA. No cross-reactivity among the tested species was observed with any of the 10 non-target mycobacterial species tested. The assay successfully amplified the target sequence from all 10 clinical isolates. Sequencing of the amplicons revealed ≥99% identity to reference *M. simiae* strains in the GenBank database, validating the assay's accuracy.

**Conclusion:** A species-specific nucleic acid target, MST601, facilitating the rapid and accurate detection of *M. simiae* via conventional PCR was presented. This assay provides a low-cost and accessible option for diagnostic laboratories.

**Keywords:** Molecular diagnostics, MST601, *Mycobacterium simiae*, Novel target, PCR, Sensitivity, Specificity

**To cite this article:** Kamali Kakhki R, Abavisani M, Ghazvini K, Zare H, Hojatpanah N, Falahi J, et al. Identification of a Novel Nucleic Acid Target for the Rapid and Specific Detection of *Mycobacterium Simiae* Using Comparative Genomic Analysis. Avicenna J Med Biotech 2026;18(2):157-166.

\* Corresponding author:  
Alireza Neshani, Ph.D.,  
Department of Laboratory  
Sciences, School of Paramedical  
Sciences, Mashhad University of  
Medical Sciences, Mashhad, Iran  
Tel: +98 9395100919,  
513 8829278  
E-mail: Neshania@mums.ac.ir,  
alireza.neshani@gmail.com  
Received: 29 Sept 2025  
Accepted: 31 Jan 2026

### Introduction

Non-Tuberculous Mycobacteria (NTM) constitute a notable category of environmental micro-organisms

that have garnered heightened interest in clinical microbiology owing to their potential to induce opportu-

istic infections, especially in immunocompromised patients and individuals with pre-existing pulmonary disorders<sup>1</sup>. Within the over 200 species classified in the NTM group, *Mycobacterium simiae* (*M. simiae*) is distinguished as a distinct pathogen due to its slow growth, non-pigmented colonies, and particularly its niacin-positive feature, a property it shares exclusively with *Mycobacterium tuberculosis* (*M. tuberculosis*; MTB). The biochemical resemblance frequently results in diagnostic ambiguity between *M. simiae* and MTB, complicating clinical management and therapy choices<sup>2</sup>.

*M. simiae* is an emerging pathogen predominantly linked to pulmonary infections, particularly in individuals with pre-existing lung conditions such as Chronic Obstructive Pulmonary Disease (COPD), bronchiectasis, or cystic fibrosis<sup>3,4</sup>. Historically perceived as geographically confined to areas like Cuba, Gaza, and the southern United States, current reports reveal its occurrence in several more places worldwide, including the Middle East<sup>2,5</sup>. A meta-analysis study indicated that between 1993 and 2012, the prevalence of NTM infections among Tuberculosis (TB) suspects in Iran was approximately 10.2%, with 43.3% of these cases attributed to *M. simiae*<sup>6</sup>. During the same period, the World Health Organization (WHO) reported an incidence rate of TB in Iran ranging from 18 to 54 cases per 100,000 population<sup>7,8</sup>. The clinical signs of *M. simiae* infections closely resemble those of TB, rendering precise diagnosis essential yet challenging<sup>9</sup>.

Accurate identification of *M. simiae* is further complicated by the potential for misdiagnosis with other NTM, particularly those sharing similar growth characteristics or biochemical profiles. For instance, *M. simiae* may be misidentified as members of the *Mycobacterium avium* Complex (MAC) or other slow-growing NTM due to overlapping clinical presentations and radiological features<sup>10</sup>. This risk is especially critical in regions where multiple NTM species coexist and in cases of co-infection, where the presence of more than one mycobacterial species can lead to diagnostic ambiguity and suboptimal treatment strategies. Failure to distinguish *M. simiae* from other NTM in such scenarios may result in the inappropriate use of macrolide-based regimens, which are ineffective against *M. simiae* due to its frequent resistance<sup>11,12</sup>. Therefore, a highly specific diagnostic tool is essential not only to differentiate *M. simiae* from *M. tuberculosis* but also to accurately identify it in the context of mixed or co-occurring NTM infections. The absence of a widely accepted, species-specific genomic target for *M. simiae* suitable for simple PCR-based detection has limited the development of efficient and uncomplicated molecular diagnostic methods like Polymerase Chain Reaction (PCR) or Loop-mediated isothermal amplification (LAMP)<sup>5</sup>.

Conventional diagnostic approaches reliant on culture techniques and phenotypic attributes, such as col-

ony morphology, growth rate, pigment synthesis, and biochemical assays, are not only laborious, but also challenging to interpret and frequently unreliable<sup>13</sup>. Despite the development of various advanced molecular techniques for the detection of *M. simiae*, including Restriction Fragment Length Polymorphism (RFLP)<sup>14</sup>, real-time PCR<sup>15</sup>, Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS)<sup>16</sup>, line probe assays<sup>13,17</sup>, sequencing<sup>15</sup>, and electrochemical label-free DNA Nanobiosensors<sup>5</sup>, these methods are considerably more costly than traditional PCR-based assays and necessitate highly skilled personnel. Importantly, most of these assays either target conserved loci (e.g., *16S rRNA*, *hsp65*, *rpoB*) that do not provide unequivocal discrimination of *M. simiae* from closely related NTM, or depend on specialized, high-cost platforms that are not universally accessible. Although conventional PCR is a well-established and routine technique in molecular microbiology, its widespread availability, low cost, and modest equipment requirements make it particularly suitable for laboratories in low- and middle-income countries and regional TB reference centers. Therefore, identifying a robust, species-specific genomic target that can be detected using conventional PCR has substantial practical value for the accurate diagnosis of *M. simiae*.

In light of these challenges, there is an urgent need for cost-effective, accessible, and highly precise diagnostic tools to distinguish *M. simiae* from other mycobacterial infections. The lack of a specific nucleic acid diagnostic target for *M. simiae* intensifies this problem, necessitating the development of novel markers suitable for straightforward, cost-effective PCR-based assays. To address this gap, a comparative genomic analysis, a method previously established by Neshani *et al*<sup>18,19</sup> was applied, to identify a nucleotide sequence that appears species-specific to *M. simiae* based on comparative genomic analysis. The primary novelty of this study lies in the identification and validation of this new target and its translation into a simple assay that can be implemented using conventional PCR, a platform that is already routinely available in most clinical and reference laboratories. The specific objectives of this study were: (1) to identify a candidate genomic region through *in silico* genome comparison; (2) to design and optimize a conventional PCR assay for the detection of this target; and (3) to evaluate the analytical sensitivity, analytical specificity, and performance of the PCR assay on reference strains and well-characterized clinical isolates.

## Materials and Methods

This study utilized comparative genomic analysis to discover a species-specific genomic target in *M. simiae*, which has a genome size of around 5.7 million base pairs<sup>18</sup>. To reduce analytical errors and guarantee precision, the complete target-identification workflow was

allocated among four researchers, with each step executed independently and cross-validated. This organization ensured that the genome-wide comparative analysis was performed in a systematic and reproducible manner, rather than as an ad hoc or single BLAST search.

#### Identifying specific target

The entire genome sequences of *M. simiae* were compared with the genomic data accessible in the NCBI database to ascertain the most precise and dependable target sequence. The workflow was executed as follows: Initially, all accessible *M. simiae* genomic sequences in the NCBI database were evaluated, and a reference quality sequence, ideally from the RefSeq database, was chosen as a starting point for subsequent research (Table 1). The *M. simiae* genome with accession number AP022568.1 was used as the reference sequence for this work.

This choice was predicated on its categorization within the NCBI RefSeq database, which offers fully assembled, high-quality genomes accompanied by standardized and meticulously reviewed annotations by the NCBI database. In comparison to other sequences classified as WGS (Whole Genome Shotgun) or GenBank entries—some of which may be incomplete or of inferior quality—the AP022568.1 sequence provides enhanced reliability and precision. Consequently, it was considered the most appropriate genomic reference for executing accurate comparative analyses of *M. simiae*.

The genomic sequence of *M. simiae* was obtained from the NCBI database and divided into about 5000 bp segments, yielding a total of 1157 pieces. Each fragment underwent independent analysis using BLASTn against the NCBI Nucleotide collection (nt) via the online BLASTn interface (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) between March and May 2024. The BLAST findings were thoroughly analyzed to identify the most distinctive segments unique to *M. simiae*. Two primary factors were utilized in this selection process: all four selected *M. simiae* strains enumerated in table 1 must appear in the BLAST results with both a sequence identity and query coverage of at least 90%. No additional bacterial species should exhibit a query coverage of over 10% in the BLAST output. As BLAST results are inherently organized by query coverage in descending order, non-target organ-

isms exhibiting greater resemblance can be readily discerned towards the conclusion of the list. The most specific genomic area was chosen based on these studies. A subsegment of this region exhibiting 100% specificity was selected and designated as MST601 (*Mycobacterium simiae* Target, 601 bp) (Figure 1). This approach represents a systematic, genome-wide screening rather than the selection of a single candidate gene. By fragmenting the complete reference genome into 1157 overlapping segments and testing each fragment against the entire NCBI nucleotide collection, the aim was to minimize bias with focusing only on commonly used loci (e.g. 16S rRNA, hsp65, rpoB). The stringent dual thresholds ( $\geq 90\%$  identity and coverage for all available *M. simiae* strains and  $\leq 10\%$  query coverage for any non-*M. simiae* species) were chosen to reduce the risk of cross-reactivity even in the presence of incomplete or misannotated genomes. This genome-wide, threshold-based strategy allowed the unbiased identification of MST601 as a highly specific marker for *M. simiae*.

#### In silico specificity study of the MST601 sequence

An *in silico* alignment study was conducted to assess the species specificity of the MST601 sequence. The whole nucleotide sequence of MST601 was individually compared with the genomes of phylogenetically related mycobacterial species. Reference genome sequences were acquired from the NCBI database. The MST601 sequence for each selected species was aligned using the "NCBI BLAST 2 Sequences (BLASTn)" tool. Accessible online at: ([https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&USE\\_DEFAULTS=on&BLAST\\_SPEC=blast2seq](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&USE_DEFAULTS=on&BLAST_SPEC=blast2seq)). Alignment parameters including query coverage, percentage identity, and match count were documented. This method facilitated a focused evaluation of probable homology between MST601 and non-target *Mycobacterium* genomes.

The selected panel of species included both phylogenetically related mycobacteria and clinically relevant non-tuberculous mycobacteria that are commonly encountered in pulmonary infections. The reference genomes and their accession numbers are listed in table 2. This second-level, targeted *in silico* validation complements the initial genome-wide screening and provides an additional layer of assurance that MST601 does not share substantial homology with non-target

Table 1. Summary of the genome sequence records for *M. simiae* accessible in the NCBI database

Accession number	Database source	Genome assembly level	Annotation quality	Reliability
AP022568.1	RefSeq	Complete Genome	High (Standardized)	Most reliable, complete, and well-annotated reference sequence
HG315953.1	GenBank	Unknown/Incomplete	Moderate	Submitted by local lab, possibly incomplete
NZ_AP022568.1	WGS (GenBank)	Scaffold/Contig	Low to Moderate	Likely an older or partial version of AP022568.1
NZ_HG315953.1	WGS (GenBank)	Scaffold/Contig	Low	Part of a WGS project, likely fragmented

The dependability and comprehensiveness of the entries exhibit considerable variation, with AP022568.1 serving as the fullest and most credible reference.

## Identification of a Novel Nucleic Acid Target for the Detection of *M. Simiae*

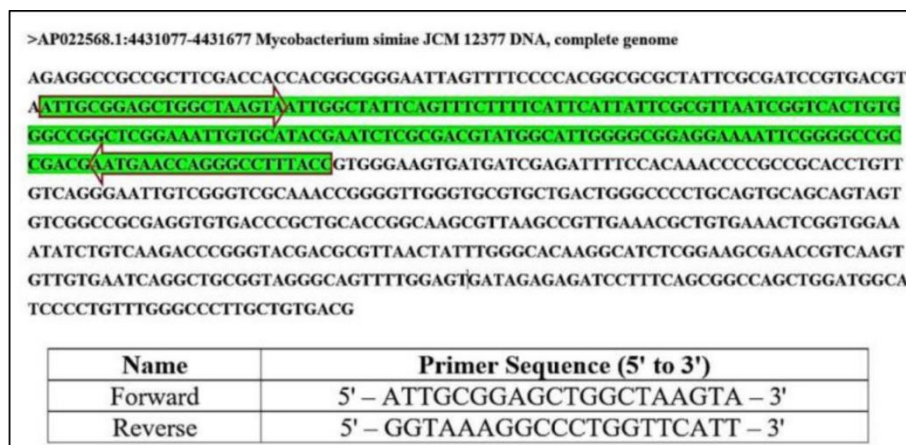


Figure 1. The particular target sequence of *M. simiae*, referred to as MST601, encompassing 601 base pairs. The locations and sequences of the forward and reverse primers are specified (Tm: 59.17 and 57.77, respectively; GC%: 50.00 for both).

*Mycobacterium* species.

### Primer design and PCR protocol

The design and analysis of primers targeting the *M. simiae* specific genomic region MST601 were conducted using Oligo7 and the OligoAnalyzer™ Tool version 3.1 (<https://eu.idtdna.com/calc/analyzer>) (Integrated DNA Technologies, Coralville, IA, USA). The primer combination is engineered to amplify a 168 bp fragment, ensuring enhanced specificity and efficiency in PCR detection. The primer sequences and amplicon information are presented in figure 1.

The PCR was conducted in a total volume of 25  $\mu$ l, comprising 1  $\mu$ l of sample DNA, 2.5  $\mu$ l of 10 $\times$ PCR buffer (100 mM Tris-HCl [pH=8.3], 500 mM KCl), 1.5  $\mu$ l of 25 mM MgCl<sub>2</sub>, 0.5  $\mu$ l of 200  $\mu$ M of each of the four dNTPs, 1  $\mu$ l of each 10  $\mu$ M forward and reverse primers, PCR-grade water, and 0.625 U of Taq DNA polymerase. The DNA amplification was conducted using a Thermal Cycler (Techne TC-3000X, Stone, Staffordshire, UK). The procedure commenced with an initial denaturation at 95°C for 5 min, followed by 40 cycles consisting of (i) denaturation at 95°C for 15 s, (ii) annealing at 61°C for 20 s, (iii) extension at 72°C for 30 s, and concluded with a final extension at 72°C for 5 min.

Then, 5  $\mu$ l of PCR product was electrophoresed on a 1.5% agarose gel prepared in 1 $\times$ TAE buffer and containing DNA Green Viewer (Parstous, Tehran, Iran). Gels were run at 90 V for 45 min using an Owl EasyCast B1 electrophoresis system (Thermo Fisher Scientific, Waltham, MA, USA), alongside a GeneRuler 100 bp DNA Ladder (Thermo Fisher Scientific) as a size marker. DNA bands were visualized under ultraviolet illumination using a Spectroline UV transilluminator (Spectroline, New York, NY, USA). The detection of a 168 bp fragment signified a favorable outcome. Ten percent of positive specimens were randomly selected, and sequencing was conducted on their amplicons. To ensure assay reliability, each PCR run included a No-Template Control (NTC) with nuclease-free

water instead of DNA template, and an extraction blank processed alongside samples during DNA isolation. Both controls consistently yielded negative results, confirming the absence of contamination. Given that all tested samples were purified genomic DNA from cultured isolates (not crude clinical specimens), the risk of PCR inhibition was minimal, and thus an Internal Amplification Control (IAC) was not included in this validation study.

### Analytical sensitivity (detection limit)

Genomic DNA of *M. simiae* (verified isolate, stored in the microbial culture collection of the Department of Laboratory Sciences, Mashhad University of Medical Sciences) was extracted from cultured colonies using the commercial mycobacterial DNA extraction kit (Karmania Pars Gene, Iran; Cat. No. KPG-DNAMTB), which is based on silica-membrane technology for efficient recovery of mycobacterial DNA. The protocol was performed according to the manufacturer's instructions. The concentration of DNA was measured utilizing a NanoDrop ND-1000 Spectrophotometer (NanoDrop, Fisher Thermo, Wilmington, DE, USA). Serial dilutions of genomic DNA were created in nuclease-free distilled water to obtain concentrations of 100 pg, 10 pg, 1 pg, 100 fg, 10 fg, 5 fg, and 1 fg/ $\mu$ l, with 1  $\mu$ l of each dilution utilized as a template in PCR experiments. The sensitivity of MST601-PCR was assessed by identifying the minimum concentration that produced a visible amplicon on agarose gel electrophoresis. Each dilution was evaluated on three distinct days to determine repeatability. Assuming a genome size of approximately 5.7 Mb for *M. simiae*<sup>18</sup> and an average molecular weight of 660 Da per base pair of double-stranded DNA, 10 fg of genomic DNA per reaction corresponds to roughly two genome equivalents.

### Analytical specificity

To assess the analytical specificity of MST601-PCR, genomic DNA from *M. simiae* and other closely related mycobacterial species, including *M. tuberculosis*, *Mycobacterium bovis* (*M. bovis*) BCG, *Mycobacte-*

Table 2. Sequence alignment results of MST601 against various mycobacterial species

Group	Species	Accession number	Query coverage (%)	Identity (%)	Length of matched region (bp)
Mycobacterium species	<i>M. simiae</i>	AP022568.1	100	100	601
	<i>M. simiae</i>	HG315953.1	100	100	601
	<i>M. simiae</i>	NZ_AP022568.1	100	100	601
	<i>M. simiae</i>	NZ_HG315953.1	100	100	601
	<i>M. tuberculosis</i>	NC_000962.3	0	0	0
	<i>M. bovis</i>	NC_002945.4	4	88.89	24
	<i>M. africanum</i>	NC_015758.1	0	0	0
	<i>M. leprae</i>	NC_002677.1	0	0	0
	<i>M. avium</i>	NC_008595.1	0	0	0
	<i>M. intracellulare</i>	NC_016946.1	0	0	0
	<i>M. kansasii</i>	NC_022663.1	4	92	23
	<i>M. abscessus</i>	NC_010397.1	0	0	0
	<i>M. fortuitum</i>	NZ_CP060409.1	5	90	27
	<i>M. chelonae</i>	CP007220.1	0	0	0
	<i>M. marinum</i>	NC_010612.1	0	0	0
	<i>M. ulcerans</i>	NC_008611.1	0	0	0
	<i>M. xenopi</i>	NZ_AP022314.1	0	0	0
	<i>M. terrae</i>	NZ_AP022564.1	4	92	23
	<i>M. smegmatis</i>	NC_008596.1	0	0	0
	<i>M. mucogenicum</i>	NZ_CP062008.1	0	0	0
<i>M. immunogenum</i>	CP011530.1	5	86.67	26	
<i>M. goodnae</i>	NZ_CP070973.1	0	0	0	
<i>M. lentiflavum</i>	NZ_CTEE01000001.1	0	0	0	
<i>M. triplex</i>	NZ_HG964446.1	0	0	0	
<i>M. genavense</i>	NZ_JAGZ01000002.1	0	0	0	
<i>M. sherrisii</i>	taxid:243061	0	0	0	
Closely related genera	<i>Nocardia asteroides</i>	NZ_CP089212.1	0	0	0
	<i>Rhodococcus equi</i>	NZ_CP157276.1	0	0	0
	<i>Corynebacterium diphtheriae</i>	NC_002935.2	0	0	0
	<i>Tsukamurella paurometabola</i>	NZ_CP085954.1	0	0	0
Common respiratory bacteria	<i>Gordonia bronchialis</i>	NC_013441.1	0	0	0
	<i>Streptococcus pneumoniae</i>	NZ_CP113114.1	0	0	0
	<i>Haemophilus influenzae</i>	NZ_CP135913.1	0	0	0
Fungi	<i>Pseudomonas aeruginosa</i>	NZ_CP084484.1	0	0	0
	<i>Aspergillus fumigatus</i>	taxid:746128	0	0	0
Human	<i>Candida albicans</i>	taxid:5476	0	0	0
	<i>Homo sapiens</i>	GRCh38 (GCF_000001405.40)	0	0	0

Query coverage is the proportion of the query sequence aligned with the reference genome, whereas identity signifies the percentage of identical nucleotides inside the aligned region.

*rium smegmatis* (*M. smegmatis*), *Mycobacterium kansasii* (*M. kansasii*), *Mycobacterium avium* (*M. avium*), *Mycobacterium chelonae* (*M. chelonae*), *Mycobacterium fortuitum* (*M. fortuitum*), *Mycobacterium ulcerans* (*M. ulcerans*), *Mycobacterium intracellulare* (*M. intracellulare*), and *Mycobacterium marinum* (*M. marinum*) was utilized as templates in PCR assays (verified isolate, preserved in the microbial culture collection of the Department of Laboratory Sciences, Mashhad University of Medical Sciences). A quantity of 1 ng of genomic DNA per reaction was used for all

tested species to ensure standardized conditions during specificity evaluation.

The PCRs were conducted under the same circumstances as those employed for sensitivity testing. Specificity was deemed adequate when amplification occurred solely in samples containing *M. simiae* DNA, with no amplification observed for any of the non-target mycobacterial species tested.

#### Assay performance on DNA from clinical isolates

Genomic DNA was taken from a collection of 10 verified clinical isolates of *M. simiae*, which had been

obtained from sputum samples of patients suspected for TB, and previously cultured and confirmed *via* ITS sequencing. The isolates were stored in the microbial culture collection of the Department of Laboratory Sciences at Mashhad University of Medical Sciences. All DNA samples performed MST601-PCR following the previously outlined methodology. Amplification products were examined by agarose gel electrophoresis under the same conditions described in section 2.3.

To test the specificity of MST601-PCR, amplicons from all ten *M. simiae* clinical isolates were sequenced utilizing forward and reverse primers. Sequencing reactions were conducted using an automated DNA sequencer. Amplicon sequencing was performed by the BGI Group (Beijing Genomics Institute, Shenzhen, China). The produced sequences were examined with the BLASTN program (<https://blast.ncbi.nlm.nih.gov/>) for comparison to the reference sequences in the GenBank database.

## Results

### Target finding

The 601 bp fragment encompassing nucleotides 4,431,077 to 4,431,677 of *M. simiae* JCM 12377 (AP022568.1) was recognized as the most specific sequence (Figure 1) (GenBank: accession number: PV878584). The BLASTN search indicated that within the NCBI registered complete genomes (nucleotide collection database), this sequence could specifically identify all strains of *M. simiae*.

### In silico specificity analysis of the MST601 sequence

The alignment analysis revealed that the MST601 sequence was solely found in *M. simiae* and exhibited no substantial similarity to the genomes of other mycobacterial species analyzed (Table 2). No comprehensive or highly conserved matches were identified among the non-target species, thereby affirming the specificity of MST601 for *M. simiae*. A limited number of non-simiae species exhibited partial alignments; however, these correspondences constituted less than 5% of the overall sequence length and demonstrated identities below 93%, signifying negligible cross-reactivity.

### Analytical sensitivity (detection limit)

The analytical sensitivity of MST601-PCR was assessed using serial dilutions of *M. simiae* genomic DNA. Clear and consistent amplification was observed for DNA concentrations ranging from 100 pg to 10 fg/ $\mu$ l, as indicated by the presence of a distinct band of the expected size on agarose gel electrophoresis (Figure 2). At the lowest concentrations tested (5 fg and 1 fg), no visible amplicon was detected under UV transillumination, establishing the detection limit of the assay at approximately 10 fg of genomic DNA per reaction (equivalent to approximately 1.5 copies of *M. simiae* genomic DNA per  $\mu$ l). All dilutions were tested in triplicate across three independent runs, demonstrating the

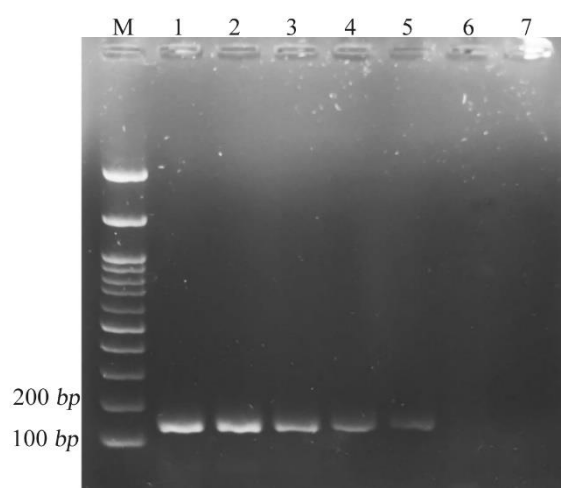


Figure 2. Agarose gel electrophoresis of MST601-PCR products using serial dilutions of *M. simiae* genomic DNA. M: 100 bp DNA marker, Lane 1: 100 pg, Lane 2: 10 pg, Lane 3: 1 pg, Lane 4: 100 fg, Lane 5: 10 fg, Lane 6: 5 fg, and Lane 7: 1 fg. A specific amplicon of 168 bp was detected up to the 10 fg (approximately equivalent to 1.5 copies of *M. simiae* genomic DNA per  $\mu$ l) with no amplification observed at 5 fg and 1 fg.

high reproducibility of the assay on different experimental days.

### Analytical specificity

The analytical specificity of MST601-PCR was evaluated using genomic DNA from *M. simiae* and closely related mycobacterial species, including *M. tuberculosis*, *M. bovis* BCG, *M. smegmatis*, *M. kansasii*, *M. avium*, *M. chelonae*, *M. fortuitum*, *M. ulcerans*, *M. intracellulare*, and *M. marinum*. Amplification occurred solely in samples containing *M. simiae* DNA, yielding a distinct 168 bp band on agarose gel electrophoresis (Figure 3). No amplicon was identified in any of the non-simiae mycobacterial species tested, demonstrating the high specificity of the MST601-PCR assay for *M. simiae*.

### Assay performance on DNA from clinical isolates

In a pilot clinical evaluation, MST601-PCR effectively amplified the target sequence from all ten verified clinical isolates of *M. simiae*. Electrophoretic analysis demonstrated a specific 168 bp amplicon in every sample, with no amplification seen in negative controls (Figure 4). The results validate the assay's excellent sensitivity in detecting *M. simiae* DNA among a range of clinical isolates. Analysis of all 10 amplicons demonstrated  $\geq 99\%$  identity with *M. simiae* reference strains in GenBank, hence validating the specificity of the MST601-PCR test. It is important to note that this assessment evaluates analytical performance on purified DNA from isolates and should not be interpreted as clinical sensitivity, which would require testing of direct patient specimens.

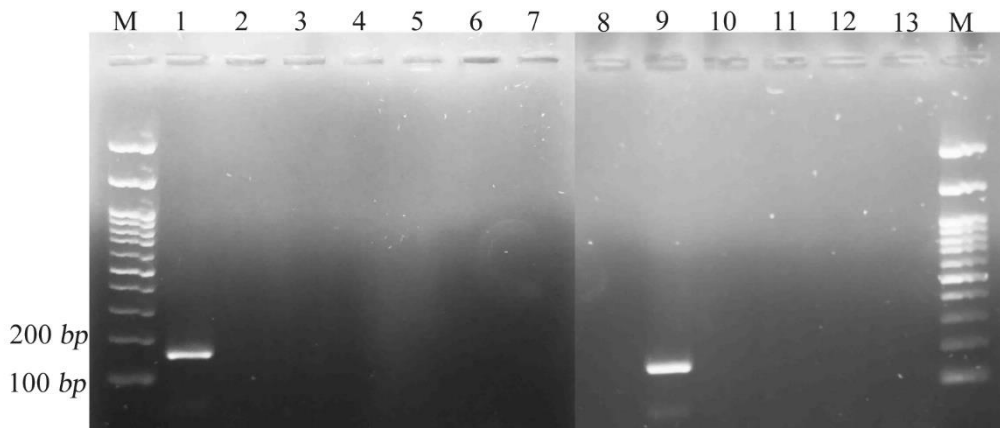


Figure 3. Specificity of MST601-PCR assessed using genomic DNA from various mycobacterial species. M: 100 bp DNA marker, Lane 1: positive control, Lane 2: *M. bovis*, Lane 3: *M. smegmatis*, Lane 4: *M. kansasii*, Lane 5: *M. avium*, Lane 6: *M. chelonae*, Lane 7: *M. intracellulare*, Lane 8: *M. fortuitum*, Lane 9: *M. simiae*, Lane 10: *M. ulcerans*, Lane 11: *M. marinum*, Lane 12: *M. tuberculosis*, and Lane 13: negative control. The absence of bands in non-simiae species confirms the high specificity of the assay, with no cross-reactivity observed.

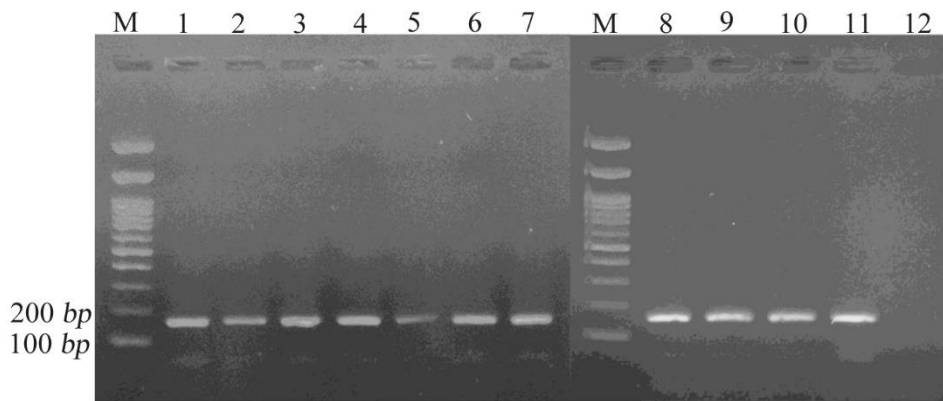


Figure 4. Agarose gel electrophoresis showing amplification of the 168 bp MST601 target in ten clinical isolates of *M. simiae*. Lane M: 100 bp DNA marker; Lane 1: positive control; Lanes 2-11: *M. simiae* isolates; and Lane 12: negative control.

### Discussion

The precise and prompt identification of *M. simiae* is essential for optimal clinical care, especially because of its phenotypic and biochemical resemblances to *M. tuberculosis*<sup>20</sup>. This study identifies MST601 as the nucleotide sequence identified for *M. simiae*, enabling the rapid and specific detection of this pathogen via a conventional PCR assay. This strategy resolves a significant diagnostic deficiency in contemporary laboratory operations, where protracted culture techniques and expensive molecular assays frequently postpone diagnosis and complicate treatment choices. In this study, aim was not to introduce a new amplification technology. Instead, the main contribution was the identification of MST601 as a highly specific genomic marker for *M. simiae* and its implementation in a conventional PCR format. While conventional PCR is a routine method in molecular microbiology, its broad availability, minimal equipment demands, and low cost make it a practical backbone for diagnostics in many

clinical laboratories, particularly in resource-limited settings. By coupling this widely accessible platform with a newly defined species-specific target, the present work addresses a critical gap in the molecular diagnosis of *M. simiae*.

Although *M. avium* and *M. intracellulare* are the most common NTM species globally, *M. simiae* demonstrates notable regional prevalence<sup>21,22</sup>. In Iran, it accounts for up to 43% of NTM isolates, emphasizing the need for region-specific diagnostic tools<sup>6</sup>. The MST601-PCR assay is particularly valuable in such endemic areas where access to advanced molecular diagnostics is limited.

Genome comparison methods have identified highly specific genetic sequences for molecular diagnostics, improving the sensitivity and specificity of PCR-based tests for *M. tuberculosis* complex and other species<sup>23,24</sup>. The MST601 area was picked from the reference genome AP022568.1 using comparative genomic analysis and *in silico* alignment. This sequence exhibited great specificity for *M. simiae*, showing no significant

relationship to other mycobacterial species or closely related genera. The BLASTN study verified that MST601 was detected in all accessible *M. simiae* genomes, exhibiting little (<5%) query coverage with non-target species, suggesting a minor likelihood of cross-reactivity. This discovery corresponds with prior attempts to pinpoint distinctive molecular markers in mycobacteria, including those employed in line probe tests and MALDI-TOF MS identification techniques<sup>13,25</sup>. In contrast to existing methodologies, the method of this study provides a cost-efficient and accessible alternative that does not necessitate specialized equipment, rendering it appropriate for use in routine diagnostic laboratories, particularly in resource-constrained environments.

The MST601-PCR assay demonstrated a detection limit of 10 fg of *M. simiae* genomic DNA per reaction. Based on the estimated genome size of ~5.7 Mb and a haploid genome mass of ~6.25 fg, this corresponds to the detection of approximately 1.5 to 2 genomic copies, indicating near single-cell sensitivity. This sensitivity level is analogous to previously documented advanced molecular diagnostic methods, although it does not require costly chemicals or a specialist apparatus<sup>5,13,15,25,26</sup>. The capacity to identify minimal quantities of DNA guarantees the assay's use in clinical specimens with a restricted bacterial presence, such as sputum or bronchoalveolar lavage fluid<sup>27</sup>. From a methodological perspective, the comparative-genomic strategy used in this study is conceptually straightforward, relying on BLAST-based sequence comparison. However, its implementation for *M. simiae* required a comprehensive and systematic workflow. The complete reference genome was fragmented into 1157 segments and each fragment was analysed against the entire NCBI nucleotide collection, applying strict criteria for inclusion of all available *M. simiae* strains and exclusion of non-target species. This genome-wide, threshold-based screening, followed by targeted *in silico* verification against phylogenetically related mycobacteria, goes beyond a simple single BLAST query and enabled the unbiased identification of MST601 as a new species-specific marker for *M. simiae*.

The analytical specificity testing further validated the assay's robustness. Among nine mycobacterial species tested, including closely related pathogens such as *M. tuberculosis*, *M. kansasii*, and *M. chelonae*, only *M. simiae* produced a positive result. These data highlight the significant discriminatory capability of MST601 for *M. simiae*, mitigating the likelihood of false-positive identification resulting from cross-reactivity.

The pilot evaluation of the MST601-PCR assay was evaluated in a pilot study on ten confirmed *M. simiae* isolates, all of which had been previously confirmed as *M. simiae* by ITS sequencing, with MST601-PCR amplicon sequencing further confirming ≥99% identity to reference strains in GenBank. These results provide a strong proof-of-concept for the assay's diagnostic po-

tential and its applicability in epidemiological research and clinical screening. However, larger-scale, multi-center studies are warranted to further validate its sensitivity and specificity across diverse clinical and geographical settings.

Conventional diagnostic methods, including phenotypic assays and culture-based identification, are both time intensive and susceptible to misidentification owing to the overlapping traits among NTM species<sup>28</sup>. Molecular techniques such as sequencing<sup>29</sup> and MALDI-TOF MS<sup>30</sup> provide enhanced precision, but need greater technical proficiency and financial resources. Real-time PCR and LAMP techniques, while sensitive and specific, frequently exceed the capabilities of smaller or under-resourced laboratories<sup>31</sup>. The MST601-PCR test presents several significant advantages, rendering it an invaluable instrument for the detection of *M. simiae*. It exhibits great specificity, only amplifying *M. simiae* DNA without cross-reactivity to other mycobacterial species. The technique demonstrates sensitivity at clinically significant levels, facilitating dependable detection even in samples with minimal bacterial burden. Moreover, its simplicity of execution with a conventional PCR apparatus guarantees extensive applicability in regular diagnostic laboratories. MST601-PCR is a low-cost and technically simple technique that obviates the necessity for costly reagents or real-time PCR equipment. The assay offers a rapid turnaround time, yielding data within hours, thus facilitating prompt diagnosis and clinical decision-making.

Although the MST601-PCR assay demonstrates encouraging outcomes, several limitations must be recognized. First, the panel of clinical isolates included in this study was confined to ten *M. simiae* isolates, which limits the precision of our estimates of diagnostic performance. Second, MST601-PCR was not assessed directly on patient-derived clinical specimens (e.g., respiratory samples); future studies should evaluate the assay under routine diagnostic conditions using primary samples. Third, the analytical specificity assessment was performed on a restricted panel of nine non-simiae mycobacterial species and therefore does not cover the full diversity of NTM; expanded specificity testing against a broader range of clinically relevant and environmental mycobacteria is warranted. Finally, the analytical sensitivity was evaluated with three replicate reactions per dilution, which is insufficient to calculate a formal 95% limit of detection using probit analysis. Future work should include larger replicate numbers, statistical LOD estimation, and validation in multicenter settings to strengthen the generalizability of present findings.

The utilization of MST601 in a quantitative real-time PCR (qPCR) format may further improve sensitivity and enable quantification of the *M. simiae* load in clinical specimens. In addition, adaptation of MST601 into multiplex qPCR panels for the concurrent identifi-

cation and differentiation of *M. simiae* from other common NTM species would increase its utility in routine diagnostics. Future work should include multi-center clinical trials that compare MST601-based assays with existing molecular methods and evaluate their performance as part of standardized diagnostic panels for NTM differentiation. A further strength of the proposed MST601-PCR assay is its reliance on conventional PCR and agarose gel electrophoresis, which are available in most microbiology and TB reference laboratories. This design choice facilitates rapid adoption of the assay without additional investment in specialized equipment, while also allowing straightforward translation of the MST601 target into more advanced qPCR or multiplex platforms where resources permit.

### Conclusion

This study presents a unique molecular target, MST601, for the specific and sensitive detection of *M. simiae* via a conventional PCR experiment. The results of the present study indicate that MST601-PCR has high analytical precision and exceptional specificity, with no cross-reactivity among the tested species. The assay's simplicity and cost-effectiveness render it an invaluable resource for ordinary diagnostic laboratories, especially in areas where *M. simiae* is common and sophisticated diagnostic facilities are scarce.

### Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT by OpenAI and Quillbot in order to improve the grammar, clarity, and overall English language quality of the manuscript. After using these tools, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

### Ethics approval

This study was conducted as part of a research project approved by the Research Ethics Committees of the School of Paramedical Sciences and the School of Health, with approval ID: IR.MUMS.FHMPM.REC.1403.142; approval date: 28 September 2024.

### Acknowledgement

This study is related to the project NO 4030855 from the Student Research Committee, Mashhad University of Medical Sciences, Mashhad, Iran. We also appreciate the "Student Research Committee" and "Research & Technology Chancellor" of Mashhad University of Medical Sciences for their financial support of this study. We also appreciate the generous cooperation of Mashhad Gene Azma Inc.

Funding: This work was supported by the Student Research Committee, Mashhad University of Medical Sciences, Mashhad, Iran (Project No. 4030855), through the provision of laboratory facilities and insti-

tutional resources. No external financial grant was received for this study.

### Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### References

1. Chai J, Han X, Mei Q, Liu T, Walline JH, Xu J, et al. Clinical characteristics and mortality of non-tuberculous mycobacterial infection in immunocompromised vs. immunocompetent hosts. *Front Med (Lausanne)* 2022; 9:884446.
2. Lotfi H, Sankian M, Meshkat Z, Khalifeh Soltani A, Aryan E. Mycobacterium simiae pulmonary infection: a case series and literature review. *Respirol Case Rep* 2021;9(3):e00719.
3. Wyrostkiewicz D, Opoka L, Filipczak D, Jankowska E, Skorupa W, Augustynowicz-Kopec E, et al. Non-tuberculous mycobacterial lung disease in the patients with cystic fibrosis—a challenging diagnostic problem. *Diagnostics (Basel)* 2022;12(7):1514.
4. Coolen-Allou N, Touron T, Belmonte O, Gazaille V, Andre M, Allyn J, et al. Clinical, Radiological, and Microbiological Characteristics of Mycobacterium simiae Infection in 97 Patients. *Antimicrob Agents Chemother* 2018;62(7):10.1128/aac.00395-18.
5. Zare H, Meshkat Z, Hatamluyi B, Rezayi M, Ghazvini K, Derakhshan M, et al. The first diagnostic test for specific detection of Mycobacterium simiae using an electrochemical label-free DNA nanobiosensor. *Talanta* 2022; 238(Pt 2): 123049.
6. Nasiri MJ, Dabiri H, Darban-Sarokhalil D, Hashemi Shahraki A. Prevalence of non-tuberculosis mycobacterial infections among tuberculosis suspects in Iran: systematic review and meta-analysis. *PLoS One* 2015;10 (6):e0129073.
7. Data WHO. Tuberculosis incidence (per 100 000 population) Geneva: WHO; 2024 [Available from: <https://data.who.int/indicators/i/13B4226/C288D13>].
8. Suresh P, Kumar A, Biswas R, Vijayakumar D, Thulasid-haran S, Anjaneyan G, et al. Epidemiology of nontuberculous mycobacterial infection in tuberculosis suspects. *Am J Trop Med Hyg* 2021;105(5):1335-8.
9. Jabbour JF, Hamieh A, Sharara SL, Kanj SS. Mycobacterium simiae: Harmless colonizer or deadly pathogen? *PLOS Pathog* 2020;16(4):e1008418.
10. Leão SC, Briones MR, Sircili MP, Balian SC, Mores N, Ferreira-Neto JS. Identification of two novel Mycobacterium avium allelic variants in pig and human isolates from Brazil by PCR-restriction enzyme analysis. *J Clin Microbiol* 1999; 37(8):2592-7.
11. Bachofner JA, Ikenberg K, Schulthess B, Nemeth J. Disseminated Mycobacterium simiae and Mycobacterium avium infection causing an immune reconstitution inflammatory syndrome in a female patient with HIV infection. *BMJ Case Rep* 2021;14(5):e241037.

12. Lee JS, Lee JH, Yoon SH, Kim TS, Seong MW, Han SK, et al. Implication of species change of Nontuberculous Myco-bacteria during or after treatment. *BMC Pulm Med* 2017;17 (1):213.
13. Kamali Kakhki R, Aryan E, Meshkat Z, Sankian M. Development of a Cost-Effective Line Probe Assay for Rapid Detection and Differentiation of Mycobacterium Species: A Pilot Study. *Rep Biochem Mol Biol* 2020;8 (4):383-93.
14. Cerit SKE, Tarhan G, Ceyhan I. Investigation of species distribution of nontuberculosis mycobacteria isolated from sputum samples in patients with suspected pulmonary tuberculosis. *The Int J Mycobacteriol* 2022;11 (2):145-9.
15. Hashemzadeh M, Hasanvand M, Montazeri EA. Analysis of relative genes expression and mutation of *pstB* and *efpA* efflux pumps in Mycobacterium *simiae* isolates from suspected tuberculosis patients by using quantitative Real-time PCR. *BMC Microbiol* 2025;25(1):144.
16. Rodriguez-Temporal D, Alcaide F, Mareković I, O'connor JA, Gorton R, van Ingen J, et al. Multicentre study on the reproducibility of MALDI-TOF MS for nontuberculous mycobacteria identification. *Sci Rep* 2022;12(1):1237.
17. Epperson LE, Davidson RM, Kammlade SM, Hasan NA, Nick SE, Machado IM, et al. Evaluation of the Geno-Type NTM-DR line probe assay for nontuberculous mycobacteria using whole genome sequences as reference standard. *Diagn Microbiol Infect Dis* 2024;110 (4):116526.
18. Neshani A, Kamali Kakhki R, Sankian M, Zare H, Hooshyar Chichaklu A, Sayyadi M, et al. Modified genome comparison method: a new approach for identification of specific targets in molecular diagnostic tests using Mycobacterium tuber-culosis complex as an example. *BMC Infect Dis* 2018;18:517.
19. Khoshbakht R, Zare H, Kakhki RK, Neshani A, Arfa-atabar M. A New Specific DNA Target Sequence for Identification of Staphylococcus epidermidis using Modified Comparative Genomic Analysis. *Avicenna J Med Biotechnol* 2022;14 (3):216-22.
20. Nqwata L, Pasipanodya JG, Black M, Feldman C. Non-Tuberculous Mycobacteria: Single Center Analyses of Risk Factors, Management and Mortality Outcomes of Adults with HIV. *Diagnostics (Basel)* 2024;14(23):2682.
21. Xu N, Li L, Wu S. Epidemiology and laboratory detection of non-tuberculous mycobacteria. *Heliyon* 2024;10 (15):e35311.
22. Jabbour JF, Hamieh A, Sharara SL, Kanj SS. Myco-bacterium *simiae*: Harmless colonizer or deadly pathogen? *PLoS Pathog* 2020;16(4):e1008418.
23. Neshani A, Kamali Kakhki R, Sankian M, Zare H, Hooshyar Chichaklu A, Sayyadi M, et al. Modified genome comparison method: a new approach for identification of specific targets in molecular diagnostic tests using Mycobacterium tuber-culosis complex as an example. *BMC Infect Dis* 2018;18 (1):517.
24. Kakhki RK, Neshani A, Sankian M, Ghazvini K, Hooshyar A, Sayadi M. The short-chain dehydroge-nases/reductases (SDR) gene: a new specific target for rapid detection of Mycobacterium tuberculosis complex by modified com-parative genomic analysis. *Infect Genet Evol* 2019;70:158-64.
25. Park JS, Choi SH, Hwang SM, Hong YJ, Kim TS, Park KU, et al. The impact of protein extraction protocols on the performance of currently available MALDI-TOF mass spectrometry for identification of mycobacterial clinical isolates cultured in liquid media. *Clin Chim Acta* 2016;460:190-5.
26. Dezhkhi H, Farnia P, Haddadi A, Farnia P, Velayati AA. Characterization of clinical isolates of Mycobacterium *simiae* using drug susceptibility tests and molecular analyses. *Curr Microbiol* 2021;78(6):2324-31.
27. Jiang L, Han L, Zhong Y, Zhang M, Li J, Rao G, et al. High utility of bronchoalveolar lavage fluid meta-genomic next-generation sequencing approach for etiologi-cal diagnosis of pneumonia. *BMC Infect Dis* 2024;24 (1):1232.
28. Maboni G, Prakash N, Moreira MAS. Review of methods for detection and characterization of non-tuber-culous mycobacteria in aquatic organisms *J Vet Diagn Invest* 2024; 36(3):299-311.
29. Zhang X, Chen H, Lin Y, Yang M, Zhao H, Hu J, et al. Diagnosis of non-tuberculous mycobacterial pulmonary disease by metagenomic next-generation sequencing on bronchoalveolar lavage fluid. *Infect Drug Resist* 2023; 16:4137-45.
30. Yao L, Gui X, Wu X, Yang J, Fang Y, Sun Q, et al. Rapid identification of nontuberculous mycobacterium species from respiratory specimens using nucleotide MALDI-TOF MS. *Microorganisms* 2023;11(8):1975.
31. Anand AR, Biswas J. TB or NTM: Can a new multiplex PCR assay be the answer? *EBioMedicine* 2021;71: 103552.