

RESEARCH ARTICLE

Open Access



# Modified genome comparison method: a new approach for identification of specific targets in molecular diagnostic tests using *Mycobacterium tuberculosis* complex as an example

Alireza Neshani<sup>1,2</sup>, Reza Kamali Kakhki<sup>1</sup>, Mojtaba Sankian<sup>3</sup>, Hosna Zare<sup>1</sup>, Amin Hooshyar Chichaklu<sup>1</sup>, Mahsa Sayyadi<sup>1</sup> and Kiarash Ghazvini<sup>1,4\*</sup> 

## Abstract

**Background:** The first step of designing any genome-based molecular diagnostic test is to find a specific target sequence. The modified genome comparison method is one of the easiest and most comprehensive ways to achieve this goal. In this study, we aimed to explain this method with the example of *Mycobacterium tuberculosis* complex and investigate its efficacy in a diagnostic test.

**Methods:** A specific target was identified using modified genome comparison method and an in-house PCR test was designed. To determine the analytical sensitivity and specificity, 10 standard specimens were used. Also, 230 specimens were used to determine the clinical sensitivity and specificity.

**Results:** The identity and query cover of our new diagnostic target (5KST) were  $\geq 90\%$  with *M. tuberculosis* complex. The 5KST-PCR sensitivity was 100% for smear-positive, culture-positive and 85.7% for smear-negative, culture-positive specimens. All of 100 smear-negative, culture-negative specimens were negative in 5KST-PCR (100% clinical specificity). Analytical sensitivity of 5KST-PCR was approximately 1 copy of genomic DNA per microliter.

**Conclusions:** Modified genome comparison method is a confident way to find specific targets for use in diagnostic tests. Accordingly, the 5KST-PCR designed in this study has high sensitivity and specificity and can be replaced for conventional TB PCR tests.

**Keywords:** *Mycobacterium tuberculosis* complex, PCR, Genome comparison method, Specific target, 5KST

## Background

One of the methods to find a specific target to use in diagnostic tests is genome comparison method. In this method, the full genome of the organism is compared with close organisms and the most specific sequences can be identified and used for the design of nucleic acid-based diagnostic tests [1]. Unfortunately, this

method causes confusion due to the lack of transparency in the protocol, and use of the unfamiliar software. In this study, we applied some changes to the previous method such as defining an obvious pathway using online and available software and aimed to employ this method to find a new specific target for *Mycobacterium tuberculosis* complex. It is clear that this new approach to detect specific diagnostic targets is also appropriate in other microorganisms, and it makes the first step of designing any genome-based diagnostic test (finding a specific target) easier.

Tuberculosis (TB) has been among the 10 leading causes of human death since 2000 [2]. It causes more

\* Correspondence: [Ghazvinik@mums.ac.ir](mailto:Ghazvinik@mums.ac.ir); [kiarash\\_ghazvini@yahoo.com](mailto:kiarash_ghazvini@yahoo.com)

<sup>1</sup>Antimicrobial Resistance Research Center, Mashhad University of Medical Sciences, Mashhad, Iran

<sup>4</sup>Department of Microbiology and Virology, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

Full list of author information is available at the end of the article



than 1.3 million deaths each year, which is even higher than road traffic accidents and HIV/AIDS [2].

WHO has planned for a 90% reduction in TB deaths by 2030. To achieve this, both parts of diagnosis and treatment must be strengthened. TB is a treatable infection, provided it is diagnosed promptly. The importance of the diagnostic part becomes clearer by knowing this fact that timely diagnosis and treatment from 2000 to 2016, saved the lives of over 53 million people with TB [2].

So far, many laboratory tests have been designed to diagnose TB, but some of these tests despite their high levels of performance, have not been used in high TB burden countries due to the costs [3]. The most common methods for TB diagnosis in low-income countries include acid-fast bacilli microscopy, culture and PCR [4].

PCR is the most important molecular diagnostic method used to detect TB during past two decades, and it has maintained its position despite the introduction of newer methods such as *loop*-mediated isothermal amplification (LAMP). Over the past years, various target sequences were used for TB PCR such as *IS6110*, *mpb64*, *devR*, *hsp65*, 38 KDa (*pstS1*), 30 KDa (*fbpB*), *esat6*, *cfp10*, 16S rDNA gene (*rrs*) and *rpoB* [5–9]. Diagnostic tests based on these targets have high sensitivity and specificity. Nevertheless, sensitivity of the test is greatly reduced on clinical specimens [10, 11]. This reduction of sensitivity can be attributed to the presence of inhibitors in clinical specimens. According to the results of some articles, there are unknown substances in the respiratory specimens (especially sputum) that have been shown to cause 10 to 26% false-negative results in PCR test and thereby reduce the sensitivity [12–14]. Other items such as poorly designed primers [15], or partial loss of target DNA during purification [16] are also effective in the reduction of PCR sensitivity, especially on clinical specimens.

Each of the targets used to date has limitations. For example, about two widely used targets (*IS6110*, *mpb64*), it has been shown that in some parts of Southeast Asia, there are strains of *M. tuberculosis* which lack the *IS6110* sequence [17], and the nonspecifically positive result of *mpb64*-PCR was also seen in *M. scrofulaceum* [18].

Therefore, we still need to find a specific and long enough sequence that would allow us to design the best possible diagnostic primers with minimal mispairing, hairpin, and dimer, to enhance the PCR sensitivity.

In this study, by modifying the genome comparison method, we identified a 5 Kbp sequence which is specific for *M. tuberculosis* complex (MTBC) and named it 5KST (5 Kbp Specific Target). It was then used to design a sensitive and highly specific PCR test. Next, the efficacy of this target was evaluated on clinical specimens (true positives and negatives) and pure genomic DNA.

## Methods

Employment of this method for the genome of *Mycobacterium tuberculosis* with about 4.4 million bp, required 45–50 h of continuous work. To reduce the errors, all dedicated work of finding the target was divided between 4 people and each part was performed separately. With this approach, 11–12 h were needed for them to investigate their fragments, which were included in three working days (4 h of continuous work per day).

### Finding specific target by modified genome comparison method

At this stage, the complete genome sequences of MTBC members were compared with available genomic sequences on nucleotide collection database (a part of NCBI) and the most specific sequence was selected. It should be noted that this instruction can be accomplished for any other microorganisms. The steps are summarized as follows:

1. First, available genomic sequences of MTBC members on NCBI database were determined and one case, preferably RefSeq, was considered as the reference for each species (Table 1).
2. Then, one of these seven items should be selected as the initial basis. For this study, the sequence of *M. tuberculosis* H37Rv (NC\_000962.3) was considered as the basis and its genomic sequence was downloaded from the NCBI and cut to about 5000 bp fragments, providing about 882 fragments.
3. Each fragment was compared separately with other NCBI available genomic sequences by blastn (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). The required time for each fragment was about 30 s.
4. Then, the blast search results were evaluated, and appropriate fragments were selected. Two criteria were considered for the result evaluation: 1. Presence of all 7 species of MTBC (Table 1) in Blast search results, with both identity and query cover  $\geq 90\%$ . To simplify, we used Ctrl+F of windows operating system. 2. No bacteria other than MTBC

**Table 1** MTBC members with available complete genome on NCBI database (preferably RefSeq)

Species	Accession number
<i>Mycobacterium tuberculosis</i> H37Rv	NC_000962.3
<i>Mycobacterium africanum</i> GM041182	NC_015758.1
<i>Mycobacterium bovis</i> AF2122/97	NC_002945.4
<i>Mycobacterium bovis</i> BCG Pasteur 1173P2	NC_008769.1
<i>Mycobacterium canettii</i> CIPT 140010059	NC_015848.1
<i>Mycobacterium caprae</i> strain Allgaeu	NZ_CP016401.1
<i>Mycobacterium microti</i> strain 12	CP010333.1



**Table 2** The number of randomly studied strains and their geographical regions, to prove the 5KST conservation

Location	Number of strains
India	53
Sweden	36
Iran	8
Uganda and South Korea	40
Moldova	16
South Africa(KwaZulu-Natal)	56
Africa	51
Mali	27
Romania	9
USA	8
Total	304

*tuberculosis* purified DNA, and the negative control contained no DNA.

The DNA amplification was performed by thermocycler (Atlas G Japan). Initial denaturation at 95 °C for 5 min, was proceeded by 40 cycles of (i) denaturation at 95 °C for 15 s, (ii) annealing at 60 °C for 20 s, (iii) extension at 72 °C for 30 s, followed by a final extension at 72 °C for 5 min.

Then, 5 µL of PCR product was electrophoresed on 2% agarose gel containing DNA green viewer and DNA bands were visualized under ultraviolet light from UV transilluminator (UVTEC). The presence of a 128 bp fragment indicated the positive result. 10% of positive specimens were selected randomly and sequencing was performed on their amplicons.

#### Analytical sensitivity (detection limit)

Purified genomic DNA of *M. tuberculosis* H37Rv was extracted from colonies by Dick Van Soolingen method [21]. DNA concentration was determined by spectrophotometer (Thermo Scientific). Then, serial dilution of genomic DNA was prepared using distilled water (10 pg, 1 pg, 100 fg, 10 fg, 5 fg, 1 fg) and 1 µL was used as template. This process was performed for three times, on three different days.

#### Analytical specificity

To determine the analytical specificity for 5KST-PCR, genomic DNA of two members of MTBC (*M. tuberculosis*

H37Rv and *M. bovis* BCG) and four Non-tuberculosis mycobacteria (*M. smegmatis*, *M. chelonae*, *M. simiae*, *M. fortuitum*) and four non-Mycobacterium bacteria (*Corynebacterium diphtheriae*, *Escherichia coli*, *Staphylococcus aureus* and *Streptococcus pneumoniae*) were used which purchased from Pasteur Institute (Tehran, Iran). The amount of 10 ng of genomic DNA was used in the reaction for each species.

#### Evaluation of clinical specimens

##### Specimen collection

A total of 100 smear-positive and culture-positive, and 100 smear-negative and culture-negative sputum specimens were obtained from tuberculosis laboratory of Qaem hospital in Mashhad. The ethical approval for performing this study was obtained from the Ethics Committee of the Mashhad University of Medical Sciences (Ethics code of IR.mums.fm.rec.1397.143). To confirm TB, *mpb64*-PCR and *IS6110*-PCR were performed on colonies of each specimen [22].

##### Specimen processing

To homogenize and concentrate the specimens, all clinical specimens examined in this study was processed using modified Petroff's method and cultured on LJ medium [23]. In brief, 5 ml of sputum was homogenized for 15 min in a shaker using an equal volume of 4% NaOH. After centrifugation at 6000 rpm for 15 min, the sediment was neutralized with 20 ml of sterile distilled water. The samples were again centrifuged. The supernatant was discarded and the precipitate was dissolved in the little residual liquid in the bottom of falcon to be used for the next steps.

##### Providing smear negative, culture-positive specimens

A total of 30 smear-positive and culture-positive processed samples were diluted with the processed negative specimen and two-fold serial dilutions were prepared (1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 1/128). Then, all of the dilutions were subjected to ziehl-neelsen staining and smear-negative dilutions were cultured on Lowenstein Jensen medium (LJ). The lowest concentrations with a positive culture were considered as smear-negative, culture-positive specimens.

##### DNA extraction

Autoclave method with some modification was used to extract DNA [24]. The processed specimens were transferred to 1.5 ml microtubes. Then, microtube lids were sealed and the autoclave process was performed (121°C for 5 min). No other substance was added to the specimens. In order to remove the debris, microtubes were centrifuged at 12000 rpm for 5 min. Then, the supernatant was used as the template for 5KST-PCR test.

**Table 3** Primers and the amplicon sequence of 5KST-PCR

Forward	5'-TTGCTGAACCTGACCTGCCCGTA
Reverse	5'-GCGTCTCTGCCTCCTCCGAT
Amplicon	5'TTGCTGAACCTGACCTGCCCGTAGCCACGGGTTCGGCTGC CGCCGAGGTAGTCGAGTTCGAGCAACTTCAGGCCGCGCGCA TGGCGTTGAAGTCCCTCGATGATCTCATCGGAGGAAGGCAGA GACGC

### Investigating the inhibitory effect of clinical specimens on 5KST-PCR

To study the inhibitory effect of clinical specimens, 8 true negative processed sputum samples were pooled together. Then the pure genomic DNA provided in the previous step, was added to 50  $\mu\text{l}$  of this pooled specimen up to the final concentration of 100  $\text{pg}/\mu\text{l}$  of DNA. Then, two groups of six dilutions were prepared by the remained pooled specimen. The first group included 10  $\text{pg}$ , 1  $\text{pg}$ , 100  $\text{fg}$ , 10  $\text{fg}$ , 5  $\text{fg}$ , and 1  $\text{fg}$  dilutions and the second group included 5  $\text{pg}$ , 0.5  $\text{pg}$ , 50  $\text{fg}$ , 5  $\text{fg}$ , 2.5  $\text{fg}$ , and 0.5  $\text{fg}$  dilutions. Then, 1  $\mu\text{l}$  of DNA for the first group and 2  $\mu\text{l}$  of DNA for the second group were used for the 5KST-PCR test. So, the amount of DNA used in both groups was equal and the only difference was the amount of clinical specimen used in the reactions.

## Results

### Target finding

The 5000 bp fragment containing nucleotides 3,127,000 to 3,132,000 of *M. tuberculosis* H37Rv genome (NC\_000962.3) was identified as the most specific possible sequence (Fig. 1). The blastn search showed that among the NCBI-registered complete genomes (nucleotide collection database), this sequence could specifically detect 237 strains of *M. tuberculosis* (Additional file 1: Table S1), 12 strains of *M. bovis* BCG, 8 strains of *M. bovis*, 3 strains of *M. africanum*, 2 strains of *M. canettii*, 1 strain of *M. caprae* and 1 strain of *M. microti*.

### Results of bioinformatics control

Two hundred thirty-seven strains of *M. tuberculosis* (with complete genome registered in nucleotide collection database) could be identified by *rpoB* sequence, which was quite identical to the strains which could be identified by 5KST-PCR (Additional file 1: Table S1).

Furthermore, the study of 5KST presence in 304 strains of *M. tuberculosis* related to the TB-ARC project also showed that this sequence is present in the all 304 strains and is considered as a conserved sequence (Additional file 1: Table S2).

### Analytical sensitivity (detection limit)

Analytical sensitivity is the lowest DNA concentration that a test can detect. Accordingly, the 5KST-PCR analytical sensitivity was 5  $\text{fg}$  approximately equivalent to 1 copy of *M. tuberculosis* genomic DNA per  $\mu\text{l}$  (Fig. 2).

### Analytical specificity

Both species of *M. tuberculosis* and *M. bovis* BCG which studied in this study were identified by 5KST-PCR, but no amplicon was produced for studied Non-tuberculosis mycobacteria (*M. smegmatis*, *M. chelonae*, *M. simiae*, *M. fortuitum*) and non-Mycobacterium bacteria

(*Corynebacterium diphtheriae*, *Escherichia coli*, *Staphylococcus aureus* and *Streptococcus pneumoniae*) (Table 4).

### Clinical specimen's result

Before beginning the study, all the specimens were cultured on the LJ medium. To select the true positive specimens, DNA extraction of the colonies was performed by simple boiling method and positive specimens were confirmed by *mpb64*-PCR and *IS6110*-PCR.

Only those specimens were included in the study as the true positives, which had clinical symptoms of the patient, and also simultaneous positive results of acid-fast bacilli smear, culture, and both PCR tests on the colonies. In addition, only those negative specimens were included in the study as true negatives, which had negative result of acid-fast bacilli smear and the culture, with the final clinical diagnosis of non-tuberculosis infection.

### Clinical sensitivity and specificity

All 100 smear-positive, culture-positive specimens were positive for 5KST-PCR test (100% sensitivity) and 25 of 30 smear-negative, culture-positive samples were positive by the test (85.7% sensitivity). All 100 smear-negative and culture-negative specimens, were negative in 5KST-PCR test (100% clinical specificity). The 5KST-PCR results on clinical specimens are provided in Table 5.

### Inhibitory effect of clinical specimens on 5KST-PCR

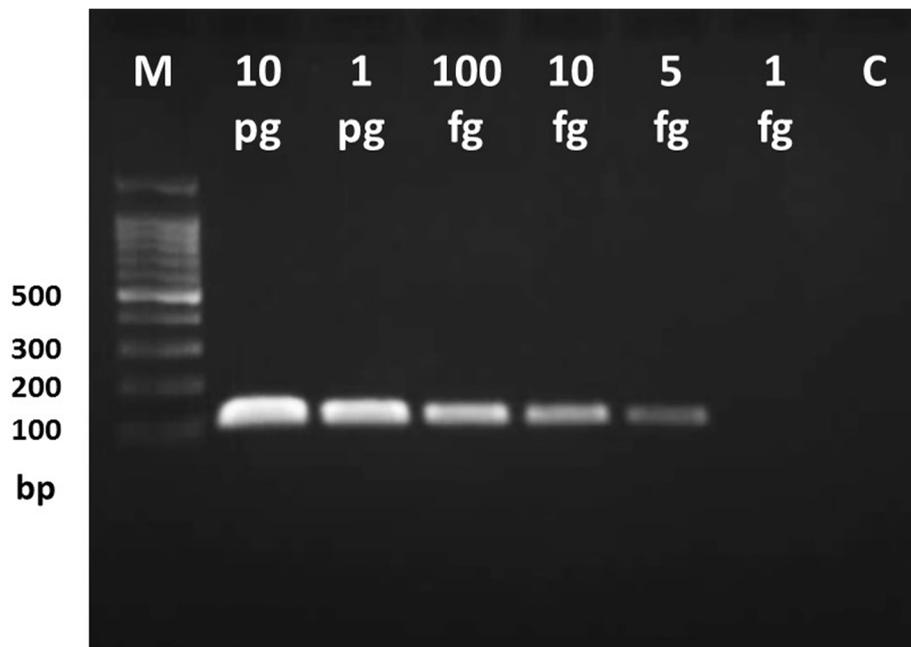
The results showed the sensitivity of 5  $\text{fg}/\mu\text{l}$  and 10  $\text{fg}/\mu\text{l}$  when using 1  $\mu\text{l}$  and 2  $\mu\text{l}$  of processed clinical specimen in our test, respectively. (Additional file 1: Figure S1).

## Discussion

One of the most important parts of designing a nucleic acid amplification test (NAAT) such as PCR, is having a completely specific and long enough target. Specificity of the target would eliminate the false positive results. The long sequence would help the researcher to design efficient and sensitive primers. In this study, genome comparison method with few modifications was used to achieve completely specific and long sequence (1). In this method, using bioinformatic facilities provided by the NCBI database, genomic sequences of 7 members of MTBC were compared with complete genomes available on the database (up to 2018.08.20) and a 5000 bp sequence with high specificity was identified.

Based on in silico studies, the 5KST can specifically detect the most of NCBI-registered complete genomes of MTBC members including: 237 strains of *M. tuberculosis*, 12 strains of *M. bovis* BCG, 8 strains of *M. bovis*, 3 strains of *M. africanum*, 2 strains of *M. canettii*, 1 strain of *M. caprae* and 1 strain of *M. microti*.

Furthermore, the investigation of 5KST presence in 304 strains of *M. tuberculosis* from different parts of the



**Fig. 2** Gel electrophoresis of 5KST-PCR products. It shows target amplicons at different concentrations of *M. tuberculosis* H37Rv genomic DNA as template. Lane M shows 100 bp DNA ladder and successive lanes show amplicons using 10 pg, 1 pg, 100 fg, 10 fg, 5 fg, 1 fg of *M. tuberculosis* genomic DNA. Lane C shows negative control

world (TB-ARC project) showed that this sequence is conserved among all these strains (Additional file 1: Table S2). It should be noted that few of these 304 strains had the query cover of less than 100% (about 81–85%). After evaluating the sequences of these strains, we found that the query cover of less than 100% was not due to the less similarity to our sequence, but also it resulted from the gaps of the sequence during the sequencing process that recorded NNN's instead of G, C, A, T bases.

Tuberculosis is a disease which has caused millions of deaths so far. Despite the advances in human science, it still puts millions of people at the risk of death every year. Therefore, extensive studies are being conducted in

many parts of the world to control it in prevention, diagnosis and treatment parts [2]. The aim of the present study was to improve the diagnostic part of TB. If the diagnosis is timely, TB can be controlled and treated with minimal side effects when the disease is not advanced [25].

Three tests that are routinely used in most TB laboratories include acid-fast bacilli microscopy, culture and PCR [26]. PCR is fast, sensitive and highly specific. In most studies on clinical specimens, it showed higher sensitivity than acid-fast bacilli microscopy, but lower or equal sensitivity than culture. PCR acts more specific than the two other tests [26–28].

The most important factors that reduce the sensitivity of the PCR test on clinical specimens include 1. Presence of inhibitors in clinical specimens 2. Partial loss of DNA during purification 3. Chemical inhibitors residue from the purification process 4. Poorly designed and inefficient primers [16, 28, 29].

In this study, to prevent the negative effects of inhibitors in clinical specimens, only 1 µL of the specimen was used in 25 µL of reaction. The results of this study, previous experiences of our team in molecular

**Table 4** The results for analytical specificity of 5KST-PCR

Groups	Bacterial Species	5KST-PCR result
<i>M. tuberculosis</i> complex	<i>M. tuberculosis</i> H37Rv	+
	<i>M. bovis</i> BCG	+
Non-tuberculosis mycobacteria	<i>M. smegmatis</i>	-
	<i>M. chelonae</i>	-
	<i>M. simiae</i>	-
	<i>M. fortuitum</i>	-
Non-Mycobacterium bacteria	<i>Corynebacterium diphtheriae</i>	-
	<i>Escherichia coli</i>	-
	<i>Staphylococcus aureus</i>	-
	<i>Streptococcus pneumoniae</i>	-

**Table 5** 5KST-PCR results on clinical specimens

	Clinical sensitivity (Smear +, culture +)	Clinical sensitivity (Smear -, culture +)	Clinical specificity (Smear -, culture -)
5KST-PCR result	100%	85.7%	100%

laboratory, as well as reviewing the studies of other researchers, showed that the sensitivity of PCR test is lower when using clinical specimens (either natural or artificial) than pure DNA in the reaction. We don't know exactly which inhibitor exists, but generally the substances of clinical specimens, especially sputum, cause the reduction in the sensitivity of test. As JE Clarridge et al. (1993), in a large study to evaluate the PCR test on clinical specimens, showed that the substances of clinical specimens, especially sputum, could cause up to 20% false-negative results and reduce the sensitivity [14]. In another study on sputum specimens, FS Nolte et al. reported that the sputum could produce 10 to 17% false-negative [12]. In a study on 76 respiratory specimens in Turkey, the inhibition rate of up to 26% was observed in the real-time PCR test and caused false-negative results [13]. Also in our study, the use of  $\geq 2$   $\mu$ l clinical specimen in the 25  $\mu$ l 5KST-PCR reaction worsened the sensitivity from 5 fg to 10 fg.

Also, to prevent DNA loss during purification, this step was removed and only autoclave extraction with few modifications was performed. Furthermore, since the 5KST sequence is long enough, we were able to design efficient primers with the best probable condition of not producing hairpin and dimer structures.

Our studies with blastn showed that all the important targets that have been used so far, have short length or contain long nonspecific regions. Unlike other targets, the 5KST target in addition to being long enough (5000 bp), does not have any statistically significant relationship to NTM bacteria. Some of the targets that have been used more than others or had better detection limit include: *rpoB*, *IS6110*, *devR*, *mpb64*, *sdaA*.

*IS6110* has been used as one of the most widely used diagnostic targets for TB [8]. This 1361 bp sequence, is repeated multiple times in the genome of *M. tuberculosis* and thus is regarded as a sensitive diagnostic target. Up to now, various PCR tests are designed based on this sequence, and the reported detection limit usually equals 2 copy of *M. tuberculosis* genomic DNA per  $\mu$ l [18, 30–32]. Various studies showed that the PCR tests based on this target had a clinical sensitivity of 63–98% and clinical specificity of 82.1–100%. [18, 33–36] This sequence has limitations despite the high sensitivity. For example, some strains of *M. tuberculosis* have been found which lack the *IS6110* sequence [17]. Furthermore, our in silico analysis with blastn showed that large fragments of this sequence have similarities with other NTM bacteria (*M. rutilum*, *M. smegmatis*, *M. chimaera*) and some Nocardia species such as *N. brasiliensis*. The 5KST sequence has only one copy in the genome of MTBC, nevertheless, it acts completely specific. In addition, unlike the *IS6110*, the 5KST is quite long which allows to design highly effective primers with minimal hairpin and

dimer structures. Although the result of poorly designed primers may not appear in the test sensitivity of pure genomic DNA, it would quietly affect the clinical specimens and reduce the sensitivity [15].

Another diagnostic target sequence which many studies are based on, is *mpb64* [28, 37, 38]. This sequence has 687 bp length [39]. Our in silico study by blastn showed that it also contains large regions with high similarity to NTM species such as *M. kansasii*, *M. ulcerans*, and *M. hemophilum*. These nonspecific regions as well as the short length of sequence, make it very difficult to design efficient primers. As in most cases, the sensitivity reported for *mpb64* primers is lower than the *IS6110* primers [35]. An analytical sensitivity that eventually reported for 5KST-PCR was 1 copy per  $\mu$ l, which according to previous studies is better than *mpb64* -PCR (20 copies per  $\mu$ l) [18]. Also, *mpb64*-PCR had 88–91% specificity and 48–91% sensitivity on clinical specimens [18, 34, 35, 37].

*rpoB* diagnostic target has also been used to detect rifampicin-resistant TB in several studies [40]. Sensitivity and specificity of the PCR assays based on this target gene have been reported in two clinical trials about 93.3–95.8 and 100% respectively [41, 42]. However, our in silico analysis showed that this sequence has nonspecific similarities with some other mycobacteria. As the analytical specificity of this target in another study showed that the common primers designed for this sequence can also detect *M. chelonae*, *M. kansasii*, *M. scrofulaceum*, *M. smegmatis* and *M. szulgai* nonspecifically [18].

The *devR* sequence is another diagnostic target. In various studies, the detection limit of 200–500 copy per  $\mu$ l has been reported. This sequence has similarities with *M. kansasii* and may therefore result in false positives. In a clinical study of intraocular TB, same specificity and lower sensitivity compared to *mpb64* were reported [18, 43, 44].

In a comparative study between common diagnostic target sequences of TB, the sensitivity of *mpb64* was highest (84% in confirmed cases and 77.5% in clinically suspected cases). The clinical sensitivity of other targets was as follows: *mpb64* > *IS6110* > *hsp65* > 38KDa (*pstS1*) > 30KDa (*fbpB*) > *esat6* > *cfp10* > *devR* [5].

*sdaA* is another target sequence with few studies. This 1383 bp sequence encodes a protein called serine dehydratase. Although the detection limit equivalent to 1 copy per  $\mu$ l has been reported for this target [18], but our blastn results showed the presence of many nonspecific regions. This sequence has 70–80% similarity to some non-tuberculosis mycobacteria (*M. ulcerans*, *M. marinum*, *M. smegmatis*, and *M. fortuitum*) and many non-Mycobacterium bacteria such as Rhodococcus and Nocardia.

## Conclusions

In this study, we succeeded to introduce highly specific novel target using modified genome comparison method. We also designed highly specific primers, as they were able to detect even 1 copy of *M. tuberculosis* genomic DNA per  $\mu\text{l}$ . We think that the unique sensitivity of these primers is due to our long enough target which allows the software to give us the best possible primers.

Our recommendations for future studies include: 1. Using modified genome comparison method to identify diagnostic targets for other pathogens. 2. 5KST-PCR analysis on more and various clinical specimens. 3. Simultaneous comparison of 5KST-PCR with other PCR tests based on other targets to determine its efficiency. 4. Discovering the 5KST genes and their functions 5. Use of 5KST target in other diagnostic tests such as LAMP.

## Additional file

**Additional file 1: Table S1.** Investigating the presence of *rhoB* and 5KST diagnostic targets in *M. tuberculosis* complete genomes registered in nucleotide collection database (a part of NCBI). **Table S2.** Investigating the presence of 5KST sequence in strains collected from different parts of the world (TB Antibiotic Resistance Catalog project: [https://olive.broadinstitute.org/projects/tb\\_arc/tree](https://olive.broadinstitute.org/projects/tb_arc/tree)). **Figure S1.** Gel electrophoresis of 5KST-PCR products. It shows target amplicons at different concentrations of *M. tuberculosis* H37Rv genomic DNA spiked in processed clinical specimen as the template. (DOCX 250 kb)

## Abbreviations

LAMP: loop-mediated isothermal amplification;; MTBC: Mycobacterium tuberculosis complex; NCBI: National Center for Biotechnology Information; NTM: Nontuberculous Mycobacteria; PCR: Polymerase Chain Reaction; TB-ARC: Tuberculosis Antibiotic Resistance Catalog

## Availability of data and materials

All data generated or analysed during this study are included in this published article [and its Additional file 1].

## Authors' contributions

AN, KG, RKK, MS contributed in the study design and target finding. AN designed the primers. AN, RKK, HZ, AH, Ma.S collected and evaluated the clinical specimens. AN, HZ wrote the manuscript. All authors approved the final version of the paper.

## Ethics approval and consent to participate

The ethical approval for performing this study was obtained from the Ethics Committee of the Mashhad University of Medical Sciences (Ethics code of IR.mums.fm.rec.1397.143) and written informed consent was obtained from participants.

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Antimicrobial Resistance Research Center, Mashhad University of Medical Sciences, Mashhad, Iran. <sup>2</sup>Student Research Committee, Mashhad University of Medical Sciences, Mashhad, Iran. <sup>3</sup>Immunology Research Center, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.

<sup>4</sup>Department of Microbiology and Virology, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.

Received: 26 June 2018 Accepted: 26 September 2018

Published online: 12 October 2018

## References

- ZHU DS, Zhou M, FAN YL, SHI XM. Identification of new target sequences for PCR detection of *Vibrio parahaemolyticus* by genome comparison. *J Rapid Methods Autom Microbiol.* 2009;17(1):67–79.
- The top 10 causes of death [Internet]. World Health Organization. 2017. Available from: <http://www.who.int/mediacentre/factsheets/fs310/en/>.
- Bojang AL, Mendy FS, Tientcheu LD, Otu J, Antonio M, Kampmann B, et al. Comparison of TB-LAMP, GeneXpert MTB/RIF and culture for diagnosis of pulmonary tuberculosis in the Gambia. *J Inf Secur.* 2016;72(3):332–7.
- Kivihya-Ndugga L, van Cleeff M, Juma E, Kimwoni J, Githui W, Oskam L, et al. Comparison of PCR with the routine procedure for diagnosis of tuberculosis in a population with high Prevalences of tuberculosis and human immunodeficiency virus. *J Med Microbiol.* 2004;42(3):1012–5.
- Raj A, Singh N, Gupta KB, Chaudhary D, Yadav A, Chaudhary A, et al. Comparative evaluation of several gene targets for designing a multiplex-PCR for an early diagnosis of Extrapulmonary tuberculosis. *Yonsei Med J.* 2016;57(1):88–96.
- Mehta PK, Raj A, Singh N, Khuller GK. Diagnosis of extrapulmonary tuberculosis by PCR. *FEMS Immunol Med Microbiol.* 2012;66(1):20–36.
- Sharma K, Gupta N, Sharma A, Singh G, Gupta PK, Rajwanshi A, et al. Multiplex polymerase chain reaction using insertion sequence 6110 (IS6110) and mycobacterial protein fraction from BCG of Rm 0.64 in electrophoresis target genes for diagnosis of tuberculous lymphadenitis. *Indian J Med Microbiol.* 2013;31(1):24–8.
- Eisenach KD, Donald Cave M, Bates JH, Crawford JT. Polymerase chain reaction amplification of a repetitive DNA sequence specific for *Mycobacterium tuberculosis*. *J Infect Dis.* 1990;161(5):977–81.
- Haldar S, Bose M, Chakrabarti P, Dagainawala HF, Harinath BC, Kashyap RS, et al. Improved laboratory diagnosis of tuberculosis—the Indian experience. *Tuberculosis (Edinburgh, Scotland).* 2011;91(5):414–26.
- Besetti J. An introduction to PCR inhibitors. *J Microbiol Methods.* 2007; 28:159–67.
- Wilson IG. Inhibition and facilitation of nucleic acid amplification. *Appl Environ Microbiol.* 1997;63(10):3741.
- Noite FS, Metchock B, McGowan J, Edwards A, Okwumabua O, Thurmond C, et al. Direct detection of *Mycobacterium tuberculosis* in sputum by polymerase chain reaction and DNA hybridization. *J Clin Microbiol.* 1993; 31(7):1777–82.
- Döşkaya M, Caner A, Değirmenci A, Wengenack NL, Yolasiğmaz A, Turgay N, et al. Degree and frequency of inhibition in a routine real-time PCR detecting pneumocystis jirovecii for the diagnosis of pneumocystis pneumonia in Turkey. *J Med Microbiol.* 2011;60(7):937–44.
- Jr C, Shawar RM, Shinnick TM, Pliakytis BB. Large-scale use of polymerase chain reaction for detection of *Mycobacterium tuberculosis* in a routine mycobacteriology laboratory. *J Clin Microbiol.* 1993;31(8):2049–56.
- Chou Q, Russell M, Birch DE, Raymond J, Bloch W. Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nucleic Acids Res.* 1992;20(7):1717–23.
- Afghani B, Stutman HR. Polymerase chain reaction for diagnosis ofM. Tuberculosis:comparison of simple boiling and a conventional method for DNA extraction. *Biochem Mol Med.* 1996;57(1):14–8.
- Lok KH, Benjamin WH Jr, Kimerling ME, Pruitt V, Lathan M, Razeq J, et al. Molecular differentiation of *Mycobacterium tuberculosis* strains without IS6110 insertions. *Emerg Infect Dis.* 2002;8(11):1310.
- Nimesh M, Joon D, Pathak AK, Saluja D. Comparative study of diagnostic accuracy of established PCR assays and in-house developed sdaA PCR method for detection of *Mycobacterium tuberculosis* in symptomatic patients with pulmonary tuberculosis. *J Inf Secur.* 2013;67(5):399–407.
- Manson AL, Cohen KA, Abeel T, Desjardins CA, Armstrong DT, Barry CE III, et al. Genomic analysis of globally diverse *Mycobacterium tuberculosis* strains provides insights into the emergence and spread of multidrug resistance. *Nat Genet.* 2017;49(3):395.
- Rychlik W. OLIGO 7 primer analysis software. *Methods Mol Biol.* 2007:35–59.
- Parish T, Stoker NG. *Mycobacterium tuberculosis* protocols: Springer Science & Business Media; 2001.

22. Aziz MM, Khan AY, Hasan KN, Azad Khan AK, Hassan MS. Comparison between IS6110 and MPB64 primers for the diagnosis of *Mycobacterium tuberculosis* in Bangladesh by polymerase chain reaction (PCR). *Bangladesh Med Res Council Bull.* 2004;30(3):87–94.
23. Tripathi K, Tripathi PC, Nema S, Shrivastava AK, Dwiwedi K, Dhanvijay AK. Modified Petroff's method: an excellent simplified decontamination technique in comparison with Petroff's method. *Int J Recent Trends Sci Tech.* 2014;10:461–4.
24. Simmon KE, Steadman DD, Durkin S, Baldwin A, Jeffrey WH, Sheridan P, et al. Autoclave method for rapid preparation of bacterial PCR-template DNA. *J Microbiol Methods.* 2004;56(2):143–9.
25. Lönnroth K, Migliori GB, Abubakar I, D'Ambrosio L, De Vries G, Diel R, et al. Towards tuberculosis elimination: an action framework for low-incidence countries. *Eur Respir J.* 2015;45(4):928–52.
26. Aslanzadeh J, de la Viuda M, Fille M, Smith WB, Namdari H. Comparison of culture and acid-fast bacilli stain to PCR for detection of *Mycobacterium tuberculosis* in clinical samples. *Mol Cell Probes.* 1998;12(4):207–11.
27. Organization WH. Definitions and reporting framework for tuberculosis–2013 revision. 2013.
28. Narotam S, Veena S, Ch NS, Raj SP, Kushwaha RS, Shivani S, et al. Conventional PCR usage for the detection of *Mycobacterium Tuberculosis* complex in Cerebrospinal Fluid by MPB64-Target PCR. *Int J Drug Dev & Res.* 2012;4(4):206–210.
29. Amita J, Vandana T, Guleria R, Verma R. Qualitative evaluation of mycobacterial DNA extraction protocols for polymerase chain reaction. *Mol Biol Today.* 2002;3(2):43–9.
30. Sankar S, Kuppanan S, Balakrishnan B, Nandagopal B. Analysis of sequence diversity among IS6110 sequence of *Mycobacterium tuberculosis*: possible implications for PCR based detection. *Bioinformatics.* 2011;6(7):283–5.
31. Thierry D, Cave M, Eisenach K, Crawford J, Bates J, Gicquel B, et al. IS6110, an IS-like element of *Mycobacterium tuberculosis* complex. *Nucleic Acids Res.* 1990;18(1):188.
32. Tumwasorn S, Kwanlertjit S, Mokmued S, Charoenlap P. Comparison of DNA targets for amplification by polymerase chain reaction for detection of *Mycobacterium tuberculosis* in sputum. *J Med Assoc Thai.* 1996;79(Suppl 1):S113–8.
33. Haldar S, Sharma N, Gupta V, Tyagi JS. Efficient diagnosis of tuberculous meningitis by detection of *Mycobacterium tuberculosis* DNA in cerebrospinal fluid filtrates using PCR. *J Med Microbiol.* 2009;58(5):616–24.
34. Rafi W, Venkataswamy MM, Ravi V, Chandramuki A. Rapid diagnosis of tuberculous meningitis: a comparative evaluation of in-house PCR assays involving three mycobacterial DNA sequences, IS6110, MPB-64 and 65 kDa antigen. *J Neurol Sci.* 2007;252(2):163–8.
35. Asthana AK, Madan M. Study of target gene IS 6110 and MPB 64 in diagnosis of pulmonary tuberculosis. *Int J Curr Microbiol Appl Sci.* 2015; 4(8):856–63.
36. Haldar S, Chakravorty S, Bhalla M, De Majumdar S, Tyagi JS. Simplified detection of *Mycobacterium tuberculosis* in sputum using smear microscopy and PCR with molecular beacons. *J Med Microbiol.* 2007; 56(10):1356–62.
37. Martins LC, Paschoal IA, Von Nowakowski A, Silva SA, Costa FF, Ward LS. Nested-PCR using MPB64 fragment improves the diagnosis of pleural and meningeal tuberculosis. *Rev Soc Bras Med Trop.* 2000;33(3):253–7.
38. Aryal R, Sah AK, Paudel DS, Joshi B, Lekhak SP, Rajbhandari R, et al. Polymerase chain reaction using the MPB64 fragment for detection of *Mycobacterium tuberculosis* complex DNA in suspected TB cases. *Int J Med Sci Public Heal.* 2014;3:1.
39. *Mycobacterium tuberculosis* H37Rv complete genome [Internet]. The National Center for Biotechnology Information 2015. Available from: <https://www.ncbi.nlm.nih.gov/nuccore/AL123456.3?from=2223343&to=2224029>.
40. Kim B-J, Hong S-K, Lee K-H, Yun Y-J, Kim E-C, Park Y-G, et al. Differential identification of *Mycobacterium tuberculosis* complex and nontuberculous mycobacteria by duplex PCR assay using the RNA polymerase gene (*rpoB*). *J Med Microbiol.* 2004;42(3):1308–12.
41. Yang S, Zhong M, Zhang Y, Wang Y. Rapid detection of *rpoB* and *katG* genes from the sputum of multidrug-resistant *Mycobacterium tuberculosis* by polymerase chain reaction (PCR)-direct sequencing analysis. *Afr J Microbiol Res.* 2011;5(26):4519–23.
42. Li J, Xin J, Zhang L, Jiang L, Cao H, Li L. Rapid detection of *rpoB* mutations in rifampin resistant *M. tuberculosis* from sputum samples by denaturing gradient gel electrophoresis. *Int J Med Sci.* 2012;9(2):148.
43. Kataria P, Kumar A, Bansal R, Sharma A, Gupta V, Gupta A, et al. devR PCR for the diagnosis of intraocular tuberculosis. *Ocul Immunol Inflamm.* 2015; 23(1):47–52.
44. Dasgupta N, Kapur V, Singh K, Das T, Sachdeva S, Jyothisri K, et al. Characterization of a two-component system, devR-devS, of *Mycobacterium tuberculosis*. *Tuber Lung Dis.* 2000;80(3):141–59.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

