

RNomic identification and evaluation of npcTB_6715, a non-protein-coding RNA gene as a potential biomarker for the detection of *Mycobacterium tuberculosis*

Priyatharisi Kanniappan ^{a, b}, Siti Aminah Ahmed ^{a, *} , Ganeswri Rajasekaram ^b, Citartan Marimuthu ^a, Ewe Seng Ch'ng ^a, Li Pin Lee ^a, Carsten A. Raabe ^{c, d, e}, Timofey S. Rozhdestvensky ^f, Thean Hock Tang ^a

^a Advanced Medical & Dental Institute (AMDI), Universiti Sains Malaysia, Kepala Batas, Penang, Malaysia

^b Department of Pathology, Johor Bahru General Hospital, Johor, Malaysia

^c Institute of Experimental Pathology (ZMBE), University of Muenster, Münster, Germany

^d Institute of Evolutionary and Medical Genomics, Brandenburg Medical School (MHB), Neuruppin, Germany

^e Institute of Medical Biochemistry, Centre for Molecular Biology of Inflammation (ZMBE), University of Muenster, Münster, Germany

^f Department of Medicine (TRAM), University Hospital of Muenster, Münster, Germany

Received: October 25, 2016; Accepted: February 1, 2017

Abstract

Technological advances in RNA biology greatly improved transcriptome profiling during the last two decades. Besides the discovery of many small RNAs (sRNA) that are involved in the physiological and pathophysiological regulation of various cellular circuits, it becomes evident that the corresponding RNA genes might also serve as potential biomarkers to monitor the progression of disease and treatment. sRNA gene candidate npcTB_6715 was previously identified *via* experimental RNomic (unpublished data), and we report its application as potential biomarker for the detection of *Mycobacterium tuberculosis* (MTB) in patient samples. For proof of principle, we developed a multiplex PCR assay and report its validation with 500 clinical cultures, positive for *Mycobacteria*. The analysis revealed 98.9% sensitivity, 96.1% specificity, positive and negative predictive values of 98.6% and 96.8%, respectively. These results underscore the diagnostic value of the sRNA gene as diagnostic marker for the specific detection of MTB in clinical samples. Its successful application and the general ease of PCR-based detection compared to standard bacterial culture techniques might be the first step towards 'point-of-care' diagnostics of *Mycobacteria*. To the best of our knowledge, this is the first time for the design of diagnostic applications based on sRNA genes, in *Mycobacteria*.

Keywords: small RNA (sRNA) • *Mycobacterium tuberculosis* • multiplex PCR • diagnosis • RNomic

Introduction

npcRNAs (non-protein-coding RNAs) do not encode significant open reading frames and are not translated into proteins [1]. However, these RNAs serve important regulatory roles, often in complex with proteins [2]. npcRNAs are further classified according to RNA processing patterns [3] or size criteria [4].

sRNAs in bacteria represent small npcRNAs with sizes typically ranging between 10 and 500 nucleotides (nts) [2, 5]. sRNAs often encode conserved RNA secondary structure or antisense elements to allow specific interactions with nucleic acid or protein targets [6],

respectively. Notably, the conservation of either element does not necessarily entail appreciable primary sequence preservation. In agreement, deep sequencing uncovered copious instances of species-specific non-protein-coding transcripts [5].

Mycobacterium tuberculosis is the causative agent of tuberculosis (TB). It was responsible for 9.6 million infections in 2014 [7] and an estimated number of 1.5 million deaths worldwide [4]. The analysis of global surveillance data revealed that tuberculosis infections are most widespread in Asia (58%) [7]. Effective control of tuberculosis outbreaks and appropriate treatment of diseased individuals demands the timely identification of infected patients. The detection of MTB in patient samples *via* PCR is less time consuming, cheaper and provides most accurate results [8, 9]. In addition, this approach requires

*Correspondence to: Siti Aminah AHMED
E-mail: asiti2000@usm.my

doi: 10.1111/jcmm.13148

less sophisticated laboratory equipment than the original, more demanding cell culture-based methods.

The *IS6110* is an insertion element specific for the mycobacterial complex and is represented by up to 25 copies per genome [10]. So far, the *IS6110* insertion sequence has been extensively exploited as target for multiplex PCR-based analyses of MTB [11–13]. But, of late there are reports of extreme instances where genomes of some MTB strains only contain a single copy of *IS6110* [14], while others even lost the insertion sequence entirely [15, 16]. The recent emergence of *IS6110*-negative strains particularly in Southeast Asia [17, 18], warrant for alternative targets that can complement the use of *IS6110* in a multiplex-based assay.

Despite of the growing evidence that sRNAs can serve as biological markers for many human diseases [19–21], the application of sRNA genes as targets for diagnosis of infectious agents has so far not been scrutinized in greater detail [22]. With small size cDNA libraries and intensive blast analysis, we identified sRNA gene candidates that were specific to MTB complex. Depending on sequence composition, length and G/C content, these candidates potentially represent suitable targets for the design of multiplex PCR assays in combination with *IS6110*.

To this end, we describe the experimental validation of the *npctB_6715* sRNA gene as complementing target to *IS6110* in a multiplex PCR assay for the detection of MTB. To reduce false negatives, amplification controls were designed and routinely included in the mPCR (multiplex) assay. The PCR test was validated with 500 culture positive clinical samples, and to rank its diagnostic potential, we compared the results with a standard commercial kit for TB diagnosis (Genotype Mycobacterium CM Hains line probe assay).

Materials and methods

Bacterial isolates and genomic DNA extraction

All bacterial strains in this study were part of the culture stock collection of the Department of Microbiology & Parasitology, School of Medical Sciences and Infectomic Cluster, Advanced Medical and Dental Institute, Universiti Sains Malaysia. DNA extraction for MTB and non-tuberculosis mycobacteria (NTM), respectively, was carried out as described [23]. In brief, the DNA extraction is based on chemical and enzymatic lysis of the bacterial cells followed by a chloroform-isoamyl alcohol extraction. Basically, MTB cells were collected into 500 μ l Tris-EDTA buffer, pH 8.0. The cell suspension was then heated for 20 min. at 80°C. Lysozyme was added (final concentration at 1 mg/ml), followed by incubation at 37°C for 2 hrs. After the incubation, 10% sodium dodecyl sulfate (final concentration 1.1%) and proteinase K (final concentration at 0.2 mg/ml; Promega Inc. Madison, Wisconsin, USA) were added. Tubes were gently vortexed gently and incubated at 65°C for additional 20 min. Next, a mixture of *N*-acetyl-*N,N,N*-trimethyl ammonium bromide (CTAB; final concentration of 40 mM) and NaCl (final concentration of 0.1 M) was added, followed immediately by the addition of NaCl alone (final concentration 0.6 M). The tubes were again vortexed before further incubation at 65°C for 10 min. Equal volume (800 μ l) of chloroform-isoamyl alcohol (24:1) was added to each tube, vortexed and then centrifuged at 12,000 \times *g* for 5 min. The genomic

DNA in the aqueous phase was precipitated with equal volume of isopropanol and washed with 80% ice-cold ethanol. The DNA pellet was air-dried briefly and resuspended in 20 μ l double-distilled water.

For cDNA library construction, MTB H37Rv was grown in Middlebrook 7H9 broth (Difco, Beckton Dickinson and Company, Sparks, MD 21152, USA) supplemented with 10% Middlebrook ADC enrichment (Becton, Dickinson and Company, Sparks, MD 21152 USA). Bacterial cultures were incubated at 37°C in a 250 ml screw cap Erlenmeyer flask and shaken manually twice per day. Cells were harvested during the lag, mid-log, stationary and late stationary phases. The cell pellets were stored at –80°C or directly processed for total RNA extraction.

For all other bacteria (Table S2, excluding MTB and NTM), we followed an in-house designed DNA extraction protocol. In brief, 1.5 ml of bacterial culture was centrifuged at 5000 \times *g*, the supernatant was discarded and tubes were vortexed to disperse bacterial pellets, the cells were suspended in 200 μ l Solution I (20% sucrose, 50 mM Tris-HCl [pH 8.0], 1% SDS, 0.2 M NaOH, 25 mM EDTA [pH 8.0], 0.1 M NaCl).

The suspension was subsequently neutralized with 200 μ l Solution II (3 M NaOAc, pH 5.2). After centrifugation at 13,000 \times *g*, the genomic DNA was precipitated with 100% ice-cold ethanol, pelleted at 13,000 \times *g*, washed with 70% ice-cold ethanol, air-dried and re-suspended in 50 μ l autoclaved ddH₂O. The concentration of the genomic DNA was measured *via* the NanoDrop™ 2000 system and stored at –20°C until used for PCR.

Total RNA isolation, cDNA library construction and sequencing

The RNA isolation method is based on single-step guanidine isothiocyanate/phenol/chloroform extraction [24]. Cell pellets (100 mg) were resuspended in 1 ml of TRIzol reagent (Gibco BRL; Life Technologies, Applied Biosystems Inc., Foster City, California, USA). The suspension was then transferred to pre-chilled 2 ml screw cap tubes containing 0.1 mm diameter zirconia/silica beads (Biospec Products Bartlesville, OK 74005, USA). Cells were homogenized with the FastPrep®-24 (MP Biomedicals Irvine, California, USA) by subjecting to 3 \times 1 min. pulses at 4.5 m/s with 1-minute rests on ice during each interval. The supernatant was transferred into 1.5-ml microcentrifuge tubes, and total RNA isolation was conducted according to the manufacturer's protocol (Gibco BRL, Life Technologies). To 1 ml of lysate was added: 0.1 ml of 2 M sodium acetate, pH 4.0 and 1 ml water-saturated phenol. The mixture was then mixed thoroughly by inversion and 0.2 ml chloroform-isoamyl alcohol (49:1) were added. The resulting biphasic system was vigorously shaken by hand for 10 sec. The sample was cooled on ice for 15 min. and centrifuged 12,000 \times *g* for 20 min. at 4°C. The upper aqueous phase was then carefully transferred to a clean tube, and 1 ml isopropanol was added for RNA precipitation for at least 1 hr at –20°C. The resulting RNA pellet was collected by centrifugation for 10 min. at 10,000 \times *g* at 4°C and washed with 1 ml of 75% ethanol. Finally, the RNA pellet was air-dried for 5–10 min. at room temperature and solubilized in 50 μ l autoclaved ddH₂O. Total RNA isolated was then treated with DNase (Ambion; Applied Biosystems, Foster City, California USA).

For analysis of sRNA transcription in MTB, we generated small size cDNA libraries, which are enriched for candidate sRNAs. Combined total RNA (200 μ g) representing the four different growth stages, lag (5 days), mid-log (14 days), early stationary (28 days) and late stationary (50 days), were size fractionated (20–350 nts) on 8% denaturing polyacrylamide gels (7 M urea, 1 \times TBE buffer). Size fractionation of total RNA starting material *via* 8% polyacrylamide gel-electrophoresis enabled the

selection of RNAs with size ranges of 20–350 nts. Passive elution was conducted in 0.3 M sodium acetate, NaOAc (pH 5.3) overnight at 4°C. After passive elution, the resulting starting material was 3' C-tailed and linked to DNA adapters at 5' termini according to Raabe *et al.*, [25]. The procedure enabled subsequent steps of reverse transcription and single stranded PCR amplification. The PCR product was then double digested with *SaI* (Roche Applied Science, Mannheim, Germany) and *NotI* (Roche Applied Science, Mannheim, Germany) restriction endonucleases and ligated (T4 DNA ligase, Roche Applied Science, Mannheim, Germany) into the pSPORT1 cloning vector (Invitrogen Life Technologies, Applied Biosystems Inc., Foster City, California, USA). Finally, 5400 colonies were selected at random and sequenced [26]. With the help of Seqman, all cDNAs were assembled into 3030 contigs, which ultimately represent all sRNA candidates of our survey. Contigs with sizes below 15 nts were excluded from downstream analysis.

Bio-computational screens for MTB complex specific sRNA genes and primer design

5'-adapter and C-tail sequences were trimmed from the cDNAs using custom UNIX scripts. cDNAs that were shorter than 15 nts were excluded from further analysis. The resulting cDNAs were assembled into contigs using SeqMan Pro 8.1.5 (LaserGene, DNASTAR, Madison, WI, USA) [27] with parameters as described by Raabe *et al.*, [25]. The assembled contigs were manually analysed and mapped against the MTB H37Rv genome (Refseq: NC_000962) using Basic Local Alignment Search Tool (BlastN) [28]. The Blast search was conducted with E-values of 10^{-5} and a minimal word size of 11 for the identification of sRNA genes specific for MTB complex. The NCBI Primer design software (www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi) was used to design suitable PCR primers for amplification of sRNA gene, IS6110 and the control plasmid pL250 (Table S1).

mPCR assay

The mPCR assay was conducted with the MyCycler™ (Bio-Rad Laboratories, Hercules, California, USA) in a total reaction volumes of 20 µl, each PCR was supplemented with 0.25 µM of forward and reverse primer for IS6110, npcRNA_6715 and plasmid DNA pL250, 200 µM dNTPs (Fermentas, ThermoFisher Scientific, Waltham, Massachusetts, USA), 1.5 mM MgCl₂ (Fermentas ThermoFisher Scientific, Waltham, Massachusetts, USA) and 2 U *Taq* DNA polymerase (Biotools, Madrid Spain) in 1× PCR buffer (10 mM Tris-HCl, [pH 8.3], 50 mM KCl). One ng genomic DNA served as template for PCR amplification and 10 ng plasmid DNA pL250 were employed as amplification internal control (AIC) to rule out false negatives caused by potential PCR inhibitors.

The PCR parameters: initial denaturation at 95°C for 1 min., 30 amplification cycles (i.e. 30 sec. denaturation at 95°C, 30 sec. annealing at 66°C, 30 sec. extension at 72°C), with final elongation step of 2 min. at 72°C. The PCR conditions were optimized with varying concentrations of primers, MgCl₂, dNTPs and AIC. In addition, annealing temperatures were optimized by gradient PCR. The specificity of the mPCR assay was determined with 10 NTM species, 7 respiratory tract-associated pathogens and 7 normal flora bacteria (i.e. bacteria that would represent non-diseased physiological conditions) (Table S2). The analytical sensitivity of the mPCR assay was determined based on the lowest amount of DNA template per reaction that yields a PCR band of the expected size in agarose gel electrophoresis.

Gel electrophoresis

All PCR products were analysed by gel electrophoresis. Twenty microlitres of each amplified sample was loaded on a 3.5% agarose gel and electrophoresed in 1× TAE buffer at 60 V for 1 hr. Gels were stained with ethidium bromide and visualized using a gel-imaging system (Bio-Rad, USA). PCR product from the samples was randomly chosen and sequenced [26].

Line probe assay (LPA)

The Line probe assay (GenoType Mycobacterium CM, Hain Lifesciences, Nehren, Germany) is a molecular test for the identification and differentiation of the MTB complex and common NTM [29]. The assay amplifies the 23S rRNA gene, followed by reverse hybridization to specific oligonucleotide probes for different mycobacterial species immobilized on membrane strip. Hybridization events were detected *via* colorimetric changes mediated by the enzymatic reaction of the streptavidin-conjugated alkaline phosphatase against the substrate.

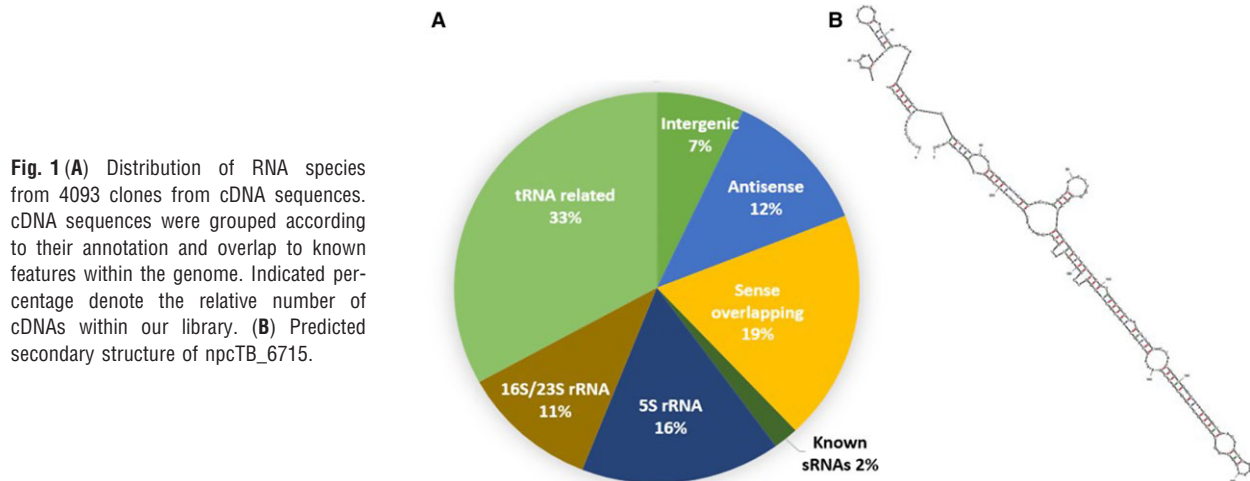
The identification of individual mycobacterial species relies on the detection of specific banding patterns. LPA analysis was performed according to the manufacturer's recommendations. In brief, PCR amplifications were conducted in 50 µl reaction volume, comprising of 35 µl of primer-nucleotide mix, 5 µl of 10× amplification buffer containing 2.5 mM MgCl₂ and 1.25 U of hot start *Taq* polymerase. Five µl genomic DNA (15 ng/µl) and distilled water were added to adjust the final reaction volumes. The amplification products were denatured in a denaturing reagent (provided with the kit) at room temperature, followed by hybridization at 45°C for 30 min. in a shaking water bath (Memmert, Germany). After two washing steps, streptavidin conjugated-alkaline phosphatase and the substrate were added for the colorimetric reaction. To control for cross-contamination, sterile double-distilled water in lieu of DNA template was added to the negative control included for each run. (For details regarding buffer compositions: <http://www.hain-lifescience.de/en/products/microbiology/mycobacteria/genotype-mycobacterium-cmas.html>).

Results and discussion

Analysis of cDNA library

About 19% of all sRNAs within our surveys were derived from protein-coding mRNAs. Here, the majority of candidate contigs partially overlapped 5' and 3' termini of annotated CDS regions. Notably, we also uncovered small antisense transcripts, which reside on the reverse strand to known CDS sequences in 12% of all sRNA candidates (Fig. 1A). The vast majority of all cDNAs represented tRNA and ribosomal RNA fragments (60%).

Cis-regulatory elements in bacteria often harbour extended RNA secondary structures, which convey increased metabolic stability [30]. The corresponding transcripts are of small size and hence are frequently detectable in small RNA libraries. Whether these are actual sRNAs or rather represent simple by-product of the cellular turnover remains matter of debate [31]. We included even these contigs in our survey for the identification of genomic regions, which would be



suitable to design multiplex PCR assays. The resulting contigs were analysed *via* blast to identify candidates that are specific for MTB. Finally out of 3030 cDNA contigs, 80% of the isolated sRNA gene candidates are found to be specific to MTB genome (data not shown).

IS6110, en masse with *npcTB_6715*, as the diagnostic markers for the mPCR assay

Multiplex PCRs to detect pathogens within patient samples demand species-specific primer design for identification of the disease-causing agent and also to discriminate against other yet closely related species [32, 33]. Here, we evaluate sRNA genes as potential targets for PCR-based amplification for diagnosis of MTB. The mPCR assay also compensates for drawbacks typically associated with bacterial cell culture and microscopy for the detection of MTB. The inclusion of additional targets provides increased selectivity to PCR-based techniques.

Phylogenetic analysis *via* blast revealed that the *npcTB_6715* candidate is detectable only within the mycobacterial complex (Supp. Dataset 1-Alignment HitTable). The RNA overlaps partially the 5' leader and CDS of hypothetical protein LH57_10405 (606 nts) (GenBank: AIR14655.1) (Fig. S1). The examination of RNA secondary structures *via* mfold [34] identified extended RNA stem loop structures. Given its genomic localization, the candidate might be leftover of *cis*-regulatory circuits to control AIR14655.1 gene expression. However, its actual molecular functions are not known. The candidate was selected for PCR assays based on its sequence features, which allowed simple PCR design.

For long, IS6110 has been target of choice because the multi-copy sequence is specific for the mycobacterial complex [10]. Recently however, various reports emphasized the emergence of IS6110-negative MTB strains [15, 16]. The addition of the sRNA gene as marker would aid in the detection of IS6110-negative variants. So far, different regions of the IS6110 were scrutinized for primer design and most often they relied on nucleotide positions 762–883 [35]. Here, we selected PCR primers from position 127–326 nts of the IS6110, which has not been explored for the design of diagnostic assays.

Optimization of the mPCR assay

Multiplex PCR assays demand extensive optimization to achieve robust results. The annealing temperature, primer design, MgCl₂ and dNTPs concentrations impact on the efficiency of PCR amplification. Analysis with gradient PCR revealed that for our assay primer annealing at 65.2°C is optimal and yields strongest amplification for all three targets (Fig. S2A–C).

Primer concentration

The parallel application of two primer pairs might cause unexpected interactions between both PCRs. We titrated primer concentrations in serial dilutions; all reactions were programmed with 1 ng of *M. tuberculosis* genomic DNA. Finally, 0.1 μM of IS6110 and 0.5 μM of *npcTB_6715* primers yielded balanced amplifications (Fig. S3A, red box).

The AIC internal control plasmid to identify PCR inhibitors

Depending on specific protocols for genomic DNA extraction sample contamination by PCR inhibitors must be assumed [36]. We designed the 'Amplification Internal Control' (AIC) plasmid as internal control, which enables the detection of false negative results. Certainly, AIC concentrations have to be titrated extensively for identification of thresholds that ensure the amplification of control plasmids but minimize the inhibition of PCRs for other targets. A wide range of pL250 plasmid DNA concentrations (1000 ng to 0.1 fg per reaction) was subjected to PCR (Seegene, USA). The results revealed that concentrations of up to 10 pg AIC per reaction allow both the detection of effective control and target gene amplification (Fig. S3B, red box).

The titration of optimal dNTPs and MgCl₂ concentrations

To determine the most effective dNTP concentration, serial dNTP dilutions ranging from 200 to 900 μM were analysed. The strongest amplification was observed with 0.4 mM dNTPs, as evident by gel electrophoresis (Fig. S4A). Similarly, we analysed the

optimal MgCl₂ concentration with serial dilutions ranging from 1.0 to 2.8 mM. Finally, 1.8 mM MgCl₂ proved best and yielded strongest amplification of both IS6110 and npcTB_6715 (Fig. S4B), respectively.

Analytical specificity of the mPCR

The specificity of the assay was validated with 10 closely related NTM strains as template. Individual strains were selected according to their prevalence within Asia to account for potential cross-reaction with common targets [37]. In addition, we included seven pathogens of the lower respiratory tract and seven microorganisms common to non-diseased individuals. We only observed successful amplifications for MTB strains (MTB H37Rv and MTB H37Ra) (Fig. 2A–C). Under these test conditions, the results revealed 100% specificity. Thus, the assay specifically detects MTB under test conditions.

Analytical sensitivity of the mPCR

The sensitivity of our mPCR assay was analysed with serial dilutions of MTB genomic DNA samples (100 ng to 1 pg). The actual detection threshold was about 100 pg genomic DNA per

20 µl total reaction, which is roughly equivalent to 10³ bacteria (one bacterium contains 10 fg genomic DNA) (Fig. 3, red box) [38]. Apparently, with this sensitivity, the mPCR assay is about 10 times more sensitive than smear microscopy, which requires minimally 10,000 AFB/ml sample for the detection of MTB [39].

Validation of the mPCR with culture positive samples

For this mPCR assay, samples that contain either the IS6110 (200 nts) or npcTB_6715 (167 nts) amplicon deliver positive signals. To exclude false negatives, AIC (250 nts) supplement is mandatory. For validation and to test the mPCR with more diverse patient material, we utilized genomic DNA derived from 500 culture positive samples and compared our results with those of the 'Mycobacterium CM line probe assay' (LPA). The LPA test is based on reverse hybridization of PCR products to their complementary probes and allows the detection of different mycobacteria on (sub) species level.

With the above collection as template, the LPA analysis identified 371 MTB and 129 NTM samples. Among the latter, 124 samples represented various NTM isolates (41 *Mycobacterium abscessus*, 27 *Mycobacterium intracellulare*, 23 *Mycobacterium kansasii*, 11

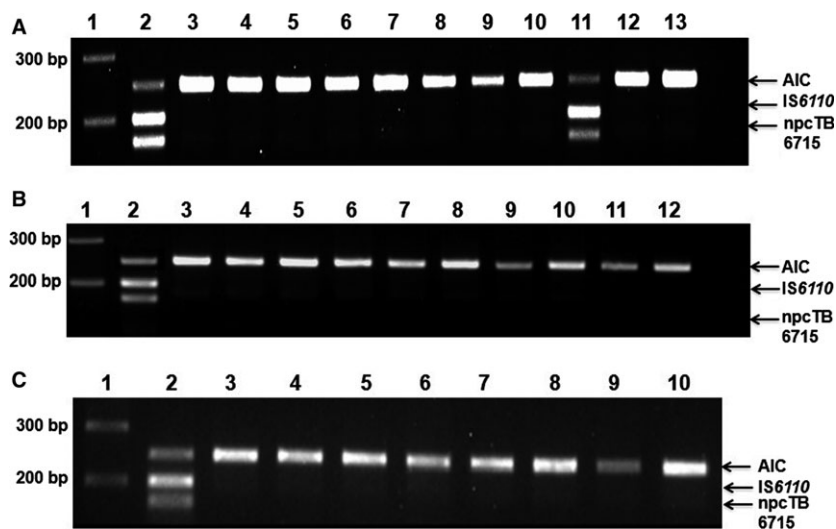


Fig. 2 Specificity of the mPCR assay analysed with 4% agarose gel-electrophoresis. The template DNA was 10 ng of MTB H37Rv genomic DNA per reaction. Primers for IS6110 and npcTB_6715 were used at 0.1 µM and 0.5 µM, respectively. (A) MTB complex and NTM, Lane 1: 100 bp DNA ladder, Lane 2: MTB H37Rv, Lane 3: *M. avium*, Lane 4: *M. abscessus*, Lane 5: *M. fortuitum*, Lane 6: *M. fortuitum*, Lane 7: *M. goodii*, Lane 8: *M. scrofaeum*, Lane 9: *M. intracellulare*, Lane 10: *M. kansasii*, Lane 11: MTB H37Ra, Lane 12: *M. malmoense*, Lane 13: Negative control (B) NTM and lower respiratory pathogens, Lane 1: 100 bp DNA ladder, Lane 2: MTB H37Rv, Lane 3: *M. marinum*, Lane 4: *M. chelonae*, Lane 5: *Staphylococcus aureus*, Lane 6: *Streptococcus pneumoniae*, Lane 7: *Klebsiella pneumoniae*, Lane 8: *Pseudomonas aeruginosa*, Lane 9: *Haemophilus influenzae*, Lane 10: *Haemophilus parainfluenzae*, Lane 11: *Moraxella catarrhalis*, Lane 12: Negative control (C) normal flora bacterial strains, Lane 1: 100 bp DNA ladder, Lane 2: MTB H37Rv, Lane 3: *Escherichia coli*, Lane 4: *Bifidobacterium bifidum*, Lane 5: *Corynebacteria sp.*, Lane 6: *Lactobacillus sp.*, Lane 7: *Neisseria meningitidis*, Lane 8: *Staphylococcus epidermidis*, Lane 9: *Coagulase Negative Streptococcus*, Lane 10: Negative control.

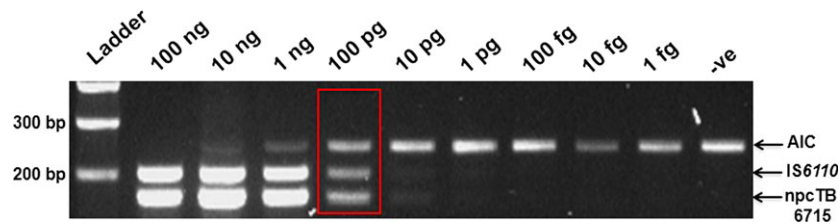


Fig. 3 Sensitivity of the mPCR assay using different amounts of template genomic DNA of MTB H37Rv ranging from 100 ng to 1 pg per reaction analysed with 3% agarose gel-electrophoresis. Primers for *IS6110* and *npcTB_6715* were used at 0.1 μ M and 0.5 μ M, respectively. Ladder: 100 bp DNA ladder, -ve: Negative control.

Table 1 Performance of mPCR (*npcTB_6715* and *IS6110*) compared to that of the Line Probe assay (LPA) for the 500 culture positive samples

(a)			
Assay sensitivity	98.9%		
Assay specificity	96.1%		
Positive predictive value	98.6%		
Negative predictive value	96.8%		
(b)	mPCR positive	mPCR negative	Total
LPA positive	367	4	371
LPA negative	5	124	129
Total	372	128	500

Mycobacterium fortuitum, 7 *Mycobacterium avium*, 4 *Mycobacterium gordonae*, 4 *Mycobacterium scrofulaceum*, 3 *Mycobacterium chelonae*, 1 *Mycobacterium xenopi*, 1 *Mycobacterium ulcerance*, 1 *Mycobacterium marinum*, 1 *Mycobacterium malmoeense* isolates); and finally 5 samples were classified as non-mycobacterial species. These results were utilized as reference point to assess the validity of the mPCR design. Among the 500 samples analysed, 372 (74.4%) isolates were mPCR positive for MTB. A representative gel picture of the mPCR result was shown (Fig. S5). Interestingly, with the mPCR we did not uncover any IS-negative isolate. Finally, the remaining 128 (25.6%) isolates were mPCR negative for MTB, as only the AIC displayed substantial PCR signals (which also indicates the absence of PCR inhibitors) (Table 1 [a]).

Based on the LPA assay, five samples were found to be false positives and four samples were false negatives. Although in general the cause of false negative results might be the presence of PCR inhibitors [36], this is highly unlikely due to the successful amplification of the AIC for all 4 samples. Nevertheless, the cause of false negatives may be linked to the loss of MTB DNA during the extraction procedure. On the other hand, the false positives might be caused by contaminations of the amplicons or the stock culture of the NTM/MTB, or carry-over during the PCR step.

Our mPCR assay reached sensitivity and specificity levels of 98.9% and 96.1%, respectively (Table 1 [b]). These values were comparable to those obtained by other assays based on PCR for the detection of MTB. Sensitivity and specificity of PCR assays vary widely, ranging from 63 to 100% and from 62 to 100%, respectively [40–45]. The differences in PCR sensitivities certainly depend on the PCR target [46]. The high sensitivities of our assay substantiated the diagnostic potential of using the sRNA gene with *IS6110* as diagnostic biomarkers for MTB. The PPV and NPV for this mPCR assay are 98.6% and 96.8% respectively (Table 1 [b]). These results indicate that the mPCR assay is useful for the rapid identification of MTB in cultures.

Conclusion

Collectively, our findings demonstrate that the multiplex PCR, amplifying *npcTB_6715* sRNA gene and *IS6110*, has potential to detect MTB complex especially in culture positive isolates. The detection limit of the mPCR assay was 100 pg of genomic DNA per reaction. Our novel sRNA-based mPCR assay may have clinical importance, whereby it can be used to differentiate between viable and non-viable TB. The sensitivity and specificity of the mPCR assay were 98.4% and 96.1%, respectively, which underscore the potential of this assay for specific diagnosis of MTB. Moreover, the PPV and NPV values of 98.6% and 95.4% were comparable to that of the LPA. The mPCR assay is also cheaper and easier to handle because it involves only one amplification cycle compared to the 3 steps of LPA. Thus, the clinical potential of the *npcTB_6715* was corroborated that it can be used alone or in concert with any target gene for the gene-based detection of MTB.

Acknowledgements

We thank the Department of Microbiology & Parasitology, School of Medical Sciences, USMCK, Kelantan for bacterial samples. The project was supported by a grant from Universiti Sains Malaysia (research grant #1001/CIPPT/811319 to TTH). PK was supported by Hadiah Latihan Persekutuan from the Ministry of Education. LPL received a MyPhD scholarship from the Ministry of Education. CAR is a member of the phi Club of the Münster Alliance for Infection Research. The views expressed in this report are those of the authors and not necessarily those of the funders. The funders had no role in study design, data collection and analysis, or preparation of the manuscript.

Author contribution

PK and LPL performed the experiments. TSR and THT conceived and designed the experiments. CAR, TSR and THT analysed the data. GR and ESC performed the clinical strain and supervised. PK, CM, SAA, CAR and THT wrote the study. PK, GR, CM, SAA, ESC, LPL, CAR, TSR and THT performed the final approval of the version to be submitted.

Conflict of interest

The authors confirm that there are no conflict of interests.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Fig. S1 Depicted is the genomic blast organization of npcTB_6715 including flanking genes

Fig. S2 Optimization of annealing temperature for IS6110, npcTB_6715 and pL250 primers. Gradient PCR were done for individual primer pair with the annealing temperature ranging from 62°C to 70°C, run on a 2% agarose gel. **(a) IS6110 (b) npcTB_6715 (c) amplification internal control (AIC)**. Ladder: 100 bp DNA ladder.

Fig. S3 (A) Optimization of primer concentrations for IS6110 and npcTB_6715 targets in mPCR, run on a 3% agarose gel.

The template DNA was 1 ng of *M. tuberculosis* H37Rv genomic DNA per reaction. In each lane, the primer concentrations were indicated respectively for IS6110 and npcTB_6715 (μM). Ladder: 100 bp DNA ladder, **(B)** Optimization of the amounts of AIC in the mPCR assay per reaction, run on a 3% agarose gel. Different amounts of AIC ranging from 1000 ng to 0.1 fg, 10 ng of *M. tuberculosis* H37Rv genomic DNA, 0.1 μM of IS6110 primers, and 0.5 μM of npcTB_6715 primers were used in the reaction. Ladder: 100 bp DNA ladder

Fig. S4 (A) Agarose gel electrophoresis of multiplex PCR using different concentrations of dNTPs ranging from 200 to 900 μM , run on a 3% agarose gel. The template DNA was 10 ng of MTB H37Rv genomic DNA per reaction. Primers for IS6110 and npcTB_6715 were used at 0.1 μM and 0.5 μM , respectively. Ladder: 100 bp DNA ladder. **(B)** Optimization of the amount of MgCl_2 ranging from 1.0 to 2.8 mM, run on a 3% agarose gel. The template DNA used was 10 ng of MTB H37Rv genomic DNA per reaction, with 0.1 μM of IS6110 primers and 0.5 μM of npcTB_6715 primers. Ladder: 100 bp DNA ladder

Fig. S5 Representative gel picture of multiplex PCR products derived from culture positive samples, analyzed with 4% agarose gel-electrophoresis. M: 100 bp DNA ladder. Lane 1: Negative control, Lane 2: Positive control, Lane 3-17: 15 culture positive samples.

Table S1 Primers used in the mPCR

Table S2 Bacterial strains for specificity testing

References

1. Brosius J, Tiedge H. RNomenclature. *RNA Biol*. 2004; 1: 81–83. Epub 2004 Jul 9.
2. Li W, Ying X, Lu Q, *et al*. Predicting sRNAs and their targets in bacteria. *Genomics Proteomics Bioinformatics*. 2012; 10: 276–84
3. Pundhir S, Gorodkin J. Differential and coherent processing patterns from small RNAs. *Sci Rep*. 2015; 5: 12062.
4. Erhard F, Zimmer R. Classification of ncRNAs using position and size information in deep sequencing data. *Bioinformatics*. 2010; 26: i426–i432.
5. Raabe CA, Tang TH, Brosius J, *et al*. Biases in small RNA deep sequencing data. *Nucleic Acids Res*. 2014; 42: 1414–26.
6. Gottesman S, Storz G. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb Perspect Biol*. 2011; 3: a003798.
7. WHO 2015. Global tuberculosis report 2015. http://www.who.int/tb/publications/global_report/en/
8. Moore DF, Guzman JA, Mikhail LT. Reduction in turnaround time for laboratory diagnosis of pulmonary tuberculosis by routine use of a nucleic acid amplification test. *Diag Microbiol Infect Dis*. 2005; 52: 247–54.
9. Kusum S, Aman S, Pallab R, *et al*. Multiplex PCR for rapid diagnosis of tuberculous meningitis. *J Neurol*. 2011; 258: 1781–87.
10. Thierry D, Brisson-Noël A, Vincent-Lévy-Frébault V, *et al*. Characterization of a *Mycobacterium tuberculosis* insertion sequence, IS6110, and its application in diagnosis. *J Clin Microbiol*. 1990; 28: 2668–73.
11. Tang TH, Ahmed SA, Musa M, *et al*. Rapid detection of *Mycobacterium tuberculosis* in clinical samples by multiplex polymerase chain reaction (mPCR). *World J Microbiol Biotechnol*. 2013; 29: 2389–95.
12. Sharma K, Sinha SK, Sharma A, *et al*. Multiplex PCR for rapid diagnosis of gastrointestinal tuberculosis. *J Glob Infect Dis*. 2013; 5: 49–53.
13. Sharma K, Appannanavar SB, Modi M, *et al*. Role of multiplex polymerase chain reaction using IS6110 and protein b for the diagnosis of extra-pulmonary tuberculosis: North India. *Indian J Pathol Microbiol*. 2015; 58: 27–30.
14. Dale JW, Mat Nor R, Ramayah S, *et al*. Molecular epidemiology of tuberculosis in Malaysia. *J Clin Microbiol*. 1999; 37: 1265–68.
15. Boruń M, Sajduda A, Pawłowska I, *et al*. Detection of *Mycobacterium tuberculosis* in clinical samples using insertion sequences IS6110 and IS990. *Tuberculosis*. 2001; 81: 271–78.
16. Radhakrishnan I, Manju YK, Kumar RA, *et al*. Implications of low frequency of IS6110 in fingerprinting field isolates of *Mycobacterium tuberculosis* from Kerala. *India. J Clin Microbiol*. 2001; 39: 1683.

17. **Huyen MN, Tiemersma EW, Kremer K, et al.** Characterisation of *Mycobacterium tuberculosis* isolates lacking IS6110 in Viet Nam. *Int J Tuberc Lung Dis.* 2013; 17: 1479–85.
18. **Lok KH, Benjamin WH Jr, Kimerling ME, et al.** Molecular differentiation of *Mycobacterium tuberculosis* strains without IS6110 insertions. *Emerg Infect Dis.* 2002; 8: 1310–13.
19. **Calin GA, Liu CG, Sevignani C, et al.** MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc Natl Acad Sci USA.* 2004; 101: 11755–60.
20. **Esteller M.** Non-coding RNAs in human disease. *Nat Rev Genet.* 2011; 12: 861–74.
21. **Meng X, Wu J, Pan C, et al.** Genetic and epigenetic down-regulation of microRNA-212 promotes colorectal tumor metastasis via dysregulation of MnSOD. *Gastroenterol.* 2013; 145: 426–36.
22. **Ravichantar N, Ahmed SA, Hoe CH, et al.** Non-protein coding RNA genes as the novel diagnostic markers for the discrimination of *Salmonella* species using PCR. *PLoS One.* 2015; 10: e0118668.
23. **Somerville W, Thibert L, Schwartzman K, et al.** Extraction of *Mycobacterium tuberculosis* DNA: a question of containment. *J Clin Microbiol.* 2005; 43: 2996–97.
24. **Chomczynski P, Sacchi N.** The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat Protoc.* 2006; 1: 581–85.
25. **Raabe CA, Sanchez CP, Randau G, et al.** A global view of the nonprotein-coding transcriptome in *Plasmodium falciparum*. *Nucleic Acids Res.* 2010; 38: 608–17.
26. **Sanger F, Nicklen S, Coulson AR.** DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad.* 1977; 74: 5463–67.
27. **Burland TG.** DNASTAR's Lasergene sequence analysis software. *Methods Mol Biol.* 2000; 132: 71–91.
28. **Altschul SF, Gish W, Miller W, et al.** Basic local alignment search tool. *J Mol Biol.* 1990; 215: 403–10.
29. **Lee AS, Jelfs P, Sintchenko V, et al.** Identification of non-tuberculous mycobacteria: utility of the GenoType *Mycobacterium* CM/AS assay compare with HPLC and 16S rRNA gene sequencing. *J Med Microbiol.* 2009; 58: 900–4.
30. **Georg J, Hess WR.** cis-Antisense RNA, another level of gene regulation in bacteria. *Microbiol Mol Biol Rev.* 2011; 75: 286–300.
31. **Brosius J, Raabe CA.** What is an RNA? A top layer for RNA classification. *RNA Biol.* 2016; 13: 140–44.
32. **Fournier P-E, Dubourg G, Raoult D.** Clinical detection and characterization of bacterial pathogens in the genomics era. *Genome Med.* 2014; 6: 114.
33. **Neogi SB, Chowdhury N, Asakura M, et al.** A highly sensitive and specific multiplex PCR assay for simultaneous detection of *Vibrio cholera*, *Vibrio parahaemolyticus* and *Vibrio vulnificus*. *Lett Appl Microb.* 2010; 51: 293–300.
34. **Zuker M.** Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003; 31: 3406–15.
35. **Sankar S, Kuppanan S, Balakrishnan B, et al.** Analysis of sequence diversity among IS6110 sequence of *Mycobacterium tuberculosis*: possible implications for PCR based detection. *Bioinformatics.* 2011; 6: 283–85.
36. **Huggett JF, Novak T, Garson JA, et al.** Differential susceptibility of PCR reactions to inhibitors: an important and unrecognised phenomenon. *BMC Res Notes.* 2008; 1: 70.
37. **Simons S, van Ingen J, Hsueh P-R, et al.** Nontuberculous mycobacteria in respiratory tract infections, eastern Asia. *Emerg Infect Dis.* 2011; 17: 343–49.
38. **De Wit D, Steyn L, Shoemaker S, et al.** Direct detection of *Mycobacterium tuberculosis* in clinical specimens by DNA amplification. *J Clin Microbiol.* 1990; 28: 2437–41.
39. **Parsons LM, Somoskövi A, Gutierrez C, et al.** Laboratory diagnosis of tuberculosis in resource-poor countries: challenges and opportunities. *Clin Microbiol Rev.* 2011; 24: 314–50.
40. **Abe C, Hirano K, Wada M, et al.** Detection of *Mycobacterium tuberculosis* in clinical specimens by polymerase chain reaction and Gen-Probe Amplified *Mycobacterium Tuberculosis* Direct Test. *J Clin Microbiol.* 1993; 31: 3270–74.
41. **Beige J, Lokies J, Schaberg T, et al.** Clinical evaluation of a *Mycobacterium tuberculosis* PCR assay. *J Clin Microbiol.* 1995; 33: 90–95.
42. **Gopinath K, Singh S.** Multiplex PCR assay for simultaneous detection and differentiation of *Mycobacterium tuberculosis*, *Mycobacterium avium* complexes and other *Mycobacterium* species directly from clinical specimens. *J Appl Microbiol.* 2009; 107: 425–35.
43. **Kim JH, Kim YJ, Ki C-S, et al.** Evaluation of Cobas TaqMan MTB PCR for Detection of *Mycobacterium tuberculosis*. *J Clin Microbiol.* 2011; 49: 173–76.
44. **Miller MB, Popowitch EB, Backlund MG, et al.** Performance of Xpert MTB/RIF RUO assay and IS6110 real-time PCR for *Mycobacterium tuberculosis* detection in clinical samples. *J Clin Microbiol.* 2011; 49: 3458–62.
45. **Rimek D, Tyagi S, Kappe R.** Performance of an IS6110-based PCR assay and the COBAS AMPLICOR MTB PCR system for detection of *Mycobacterium tuberculosis* complex DNA in human lymph node samples. *J Clin Microbiol.* 2002; 40: 3089–92.
46. **Sachadyn P, Kur J.** The construction and use of a PCR internal control. *Mol Cell Probes.* 1998; 12: 259–62.