Highly sensitive sequence specific qPCR detection of *Mycobacterium tuberculosis* complex in respiratory specimens

Jennifer L. Reed a, Zachary J. Walker a, Debby Basu a, Veronica Allen b, Mark P. Nicol b, David M. Kelso a, Sally M. McFall a, *

a Center for Innovation in Global Health Technologies (CIGHT), Department of Biomedical Engineering, Northwestern University, Evanston, IL 60208, USA
b Division of Medical Microbiology, Department of Pathology, University of Cape Town and National Health Laboratory Service, Cape Town, South Africa

**A R T I C L E   I N F O**

**Article history:**
Received 24 July 2016
Received in revised form 25 August 2016
Accepted 3 September 2016

**Keywords:**
*Mycobacterium tuberculosis*
Quantitative PCR
Sequence specific capture
Diagnosis
Sputum

**S U M M A R Y**

Nucleic acid amplification tests for *Mycobacterium tuberculosis* (MTB) detection from sputum are highly sensitive and specific with smear microscopy positive specimens, but their sensitivity with smear-negative/culture-positive specimens is much lower; therefore, these tests cannot rule out a tuberculosis diagnosis. Co-extraction of PCR inhibitors may be a cause of decreased test sensitivity. Here the design and early validation of a MTB screening assay with sample preparation and qPCR methods designed to specifically address this diagnostic gap is reported. First, human genomic DNA is identified as a significant qPCR inhibitor. To circumvent this problem, a novel, streamlined sample preparation method utilizing detergent and proteolysis to thin the sputum and DNA sequence specific MTB DNA isolation was developed. Additionally, a multiplexed qPCR assay targeting two MTB complex-specific loci: the potentially multi-copy IS6110 and the single-copy *senX3-regX3*, combined with the *cotE* gene from *Bacillus atrophaeus* spores amplified as a process control was developed. The limit of detection of the test was estimated to be 20 cfu/ml which is significantly lower than the Xpert® MTB/RIF assay. In a preliminary field study of 60 de-identified blinded spuata, a test sensitivity of 96% and specificity of 100% was observed when compared to the Xpert® MTB/RIF assay.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

**1. Introduction**

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (MTB) complex species, is the deadliest infectious disease globally with an estimated 9.6 million people falling ill, and 1.5 million people dying from TB in 2014 [1]. HIV-1 positive populations are particularly susceptible to TB and are 29 times more likely to die from active TB disease than HIV-1 negative populations. HIV-TB coinfection contributes substantially to TB-associated deaths worldwide, as 1.2 million (12%) individuals who developed TB were HIV positive, and 0.4 million co-infected patients died in 2014, accounting for 27% of the estimated 1.5 million deaths from TB [1]. Accurate and timely diagnosis is the first step in providing care to patients and preventing transmission.

Mycobacterial culture has long been the conventional gold standard test for TB diagnosis in high resource settings. It has high sensitivity (limit of detection ~10 cfu/ml) [2] but the time-to-result is lengthy (ranging from 2 to 8 weeks) [3] and the sample preparation is technically challenging, prone to contamination and requires a BSL-3 laboratory facility. In most low resource settings, bacterial culture is unavailable, leaving sputum smear microscopy as the only direct bacteriological test available [4]. The limit of detection (LOD) of the unconcentrated smear test is approximately 10,000 cfu/ml, and it has poor specificity in settings where non-tuberculosis mycobacteria are commonly isolated [3]. Smear microscopy’s poor clinical performance is particularly troublesome in settings with high HIV-1 incidence. The HIV-1 epidemic has led to a disproportionate increase in the reported rate of smear-negative TB patients, and co-infection with HIV-1 changes the presentation of smear-negative TB to more rapidly progressive disease with a high mortality rate [5,6].

Nucleic acid amplification tests (NAT) have the potential to provide highly sensitive detection of paucibacillary forms of tuberculosis with rapid turn-around-time; however, the commercially available NATs are not yet meeting this need. To date, the analytical sensitivity of the most widely adopted NAT, the Xpert®...
MTB/RIF Assay (Cepheid; Sunnyvale, CA) is reported to be 131 cfu/ml sputum (95% CI) which is only slightly less sensitive than culture (10–100 cfu/ml), and the qPCR assay LOD is < 4.5 genomes per reaction [7]. The clinical sensitivity with smear-positive specimens is 98% when compared to culture; however, clinical sensitivity with smear-negative specimens is only 67% [8]. This sensitivity can be improved to greater than 90% by testing the same patient 3 times [7]. Suboptimal sensitivity is likely to lead to reduced test impact, as clinicians will continue to use empiric treatment in test-negative patients [9] or alternatively may miss the diagnosis of TB in patients with paucibacillary TB, such as children and people living with HIV.

The performance of MTB NATs is greatly impacted by the sample preparation, mycobacterial lysis, and DNA extraction methods utilized [10,11]. Leung et al. [12] reported up to a 12.5-fold difference in DNA yield simply due to different cell lysis and DNA extraction protocols. Additionally, biological factors in sputum can cause PCR inhibition with a reported 5-fold decrease in analytical sensitivity of an MTB-specific PCR assay when comparing extraction from buffer to sputum [13]. Indeed, sputum appeared to be one of the more difficult sample type for detecting MTB using the Xpert® MTB/RIF test because pulmonary specimens had a greater proportion of inhibited specimen. Clinicians are aware of this non-target nucleic acid process control of the than non-pulmonary sample types (6% vs. 1%) [14]. In a large study of 741 patients, 15.5% of the Xpert® MTB/RIF positive specimens were inaccurately quantified due to PCR inhibition observed by substantial delays in the internal control values [15], and in another smaller study, ~20% of false negatives were reported to be caused by the presence of Taq inhibitors from sputum [16].

The sensitivity gap between culture and NAT performance implies that an analysis of the different sample preparation and testing methods may reveal strategies that can improve NAT performance to approach that of culture.

MTB NAT sample preparation can be performed by concentrating and purifying intact MTB from sputum, or by directly lysing MTB bacteria in sputum followed by extraction of DNA. MTB in liquefied sputum can be concentrated through centrifugation followed by resuspension of the pellet in buffer, or MTB can be purified through filtration and washing as in the Xpert® MTB/RIF Assay [17]. However, logistical obstacles often thwart successful execution of these established sample processing methods. The sample volume of sputum is generally limited because the high concentration of PCR inhibitors and nucleases found in the specimen can cause assay failure. Sputum itself is a viscous, heterogeneous mixture that contains high levels of nucleic acid, and concentrating by filtration limits the test volume because the filter will readily clog. Tests that rely on centrifugation to concentrate the bacilli from the liquefied sputum also neglect the free MTB DNA present in the supernatant [18], which may reduce the overall sensitivity of the assay.

Here a MTB screening assay is described with a novel sample preparation method that achieves an LOD of 20 cfu/ml which is more sensitive than the current NATs and approaches that of culture. Human genomic DNA was identified as the key qPCR inhibitor present in sputum, and a sample preparation method was developed that substantially eliminates it from the amplification reaction. A multiplexed qPCR assay was also developed that targets two M. tuberculosis complex (MTBC)-specific loci to increase sensitivity and minimize the likelihood of false negatives due to target gene mutation or deletion. It includes the potentially multi-copy insertion sequence IS6110 [19], and the highly conserved single-copy backup target senX3-regX3, a mycobacterial two-component regulatory operon critical for MTB virulence [20]. Amplification of the cotJC gene in Bacillus atrophaeus spores, which are added to the raw sputum, serves as a control for adequate processing of the target bacteria and to monitor qPCR inhibitors. The effectiveness of the sequence specific sample preparation method was verified with 60 sputum specimens collected from symptomatic TB patients displaying 96% sensitivity and 100% specificity compared to the GeneXpert® MTB/RIF Assay.

2. Materials and methods

2.1. Bacterial strains, genomic and plasmid DNA and sample sources

Two MTBC strains, M. tuberculosis H37Ra and Mycobacterium bovis BCG were acquired from the American Type Culture Collection (ATCC; Manassas, VA) and used as positive controls for assay development. Mycobacterial samples were sonicated to break up cell clumps prior to contriving sputum specimens as described by Helb et al. [7]. B. atrophaeus spores (MesaLabs; Lakewood, CO) were used as the process control. M. tuberculosis H37Rv TMC 303 genomic DNA (ATCC; Manassas, VA) was used as the template for DNA-specific capture and real-time PCR assays. Genomic DNA from 6 Mycobacterium species: Mycobacterium gordonae, Mycobacterium intracellulare, Mycobacterium terrae, Mycobacterium malmoense, Mycobacterium celatum, and Mycobacterium abscessus (ATCC; Manassas, VA) was used as the template for the specificity panel. The β-globulin plasmid used to generate a standard curve of the human single-copy β-globulin gene. HBB [21], was constructed by cloning 476 bp of nucleotide 76–552 of the HBB sequence (sequenceC: gblKP309822.1) into the EcoRI site of pGEM(R)-(T-Easy Vector (Promega; Madison, WI). Residual sputum specimens were obtained from TriCore Reference Laboratories (Albuquerque, NM) in frozen 1 ml aliquots.

2.2. Quantitative PCR

PCR master mix for a 25 μl total reaction volume (15 μl master mix and 10 μl of eluted DNA) consisted of: 0.2 mg/ml bovine serum albumin (BSA; Life Technologies Corporation; Grand Island, NY), 0.2% tween-20, 150 mM trehalose (Sigma; St. Louis, MO), 10% glycerol, 62.5 mM bicine pH 8 (Affymetrix; Santa Clara, CA), 135 mM potassium acetate pH 7.5 (Affymetrix; Santa Clara, CA), 1.5 mM manganese chloride (Sigma; St. Louis, MO), 0.325 mM each dNTP (Life Technologies Corporation; Grand Island, NY), 3.75 U R M S 205 DNA polymerase (Roche Molecular Systems, Inc., Branchburg, NJ), and sequence-specific oligonucleotides (Table 1). For the IS6110 single-plex assays, 200 nM IS6110 F7, 200 nM IS6110 R10, and 300 nM IS610 probe were used. For the β-globulin assays, 100 nM bgF, 100 nM bgR and 150 nM β-globulin probe were used. For the multiplexed MTB assay, 200 nM IS6110 F7, 200 nM IS6110 R10, 300 nM IS6110 probe, 300 nM senX3-regX3 F3, 200 nM senX3-regX3 R2, 200 nM senX3-regX3 probe 8, 100 nM cotJC F2, 100 nM cotJC R2 and 100 nM cotJC probe3. Freeze-dried qPCR master mix was used for the DNase I experiment and for the field testing study using clinical specimens. The freeze-dried master mix composition is the same as the liquid qPCR master mix except that the amount of qPCR enzyme stabilizers were modified as follows: 2.5 mg/ml BSA, 0.03% tween-20 and 138 mM trehalose, and the bicine buffer, potassium acetate and manganese were not freeze dried rather they were added as the resuspension buffer. Oligonucleotides were supplied by IDT (Corvallis, IA) with the exception of the cotJC probe3 which was supplied by Biosearch Technologies (Petaluma, CA).

Amplification was performed in a 5-plex Qiagen (Hilden, Germany) Rotor-Gene Q thermocycler, with the following cycling conditions: 1.95 °C 2:00, 2.95 °C 0:15, 3: 60 °C 0:45, 4. repeat steps
eluates and comparing the quantification cycles (Cqs) of the diluted vs. the undiluted samples. Assuming a 100% efficient qPCR assay, the Cqs should be 3.3 cycles apart. A semi-quantitative inhibition rating (IR) was established based upon dCq, the difference between the expected Cq and the actual Cq. A dCq less than 1 was represented by a score of IR 0 (uninhibited); dCq between 1 and 3 was represented by a score of IR 1 (moderately inhibited); and dCq greater than 3 or reaction failure was represented by a score of IR 2 (severely inhibited).

2.5. Sputum thinning and extraction method comparison

Three different sputum thinning and decontamination methods were compared with 16 individual sputum specimens (Figure 3A): N-acetyl-L-cysteine-sodium hydroxide (NALC), guanidinium hydrochloride (GuHCl) which is the first step of the Dynabeads® SILANE Genomic DNA Kit (Life Technologies, Carlsbad, CA), and sodium dodecyl sulfate with proteinase K (SDS/PK). The NALC-NaOH and GuHCl treated samples were subsequently extracted using the Dynabeads® SILANE Genomic DNA Kit, and the SDS/PK treated samples were extracted using DNA sequence specific capture. One million cfu of M. bovis BCG was added to each test.

Human DNA was quantified using a standard curve of 100 to 1,000,000 copies of β-globulin plasmid DNA. MTB yield was quantified via a standard curve of 100 to 100,000 copies of IS6110 M. bovis BCG genomic DNA. Samples were run at 10 μl eluate, 10 μl of 10⁻¹ dilution and 10 μl of 10⁻² dilution in duplicate to insure that the Cq would be on the standard curve and to use the dilutions to determine if amplification was inhibited. For the GuHCl samples, the β-globulin gene was quantified using the 10⁻² dilution and determined the inhibition by comparing the Cqs of the 10⁻¹ and 10⁻² dilutions. For the NALC treated samples, the β-globulin gene was quantified using the 10⁻¹ dilutions and determined inhibition by comparing the Cqs of the neat samples vs. the 10⁻¹ dilutions. For the SDS/PK treated samples, the β-globulin gene was quantified using the neat eluates and determined inhibition by comparing the Cqs of the neat samples vs. the 10⁻¹ dilutions.

2.4. DNase I treatment

A cocktail of sputum specimens was prepared by vortexing 15 ml residual sputum specimens with approximately 7.5 g acid-washed 5 mm glass beads (Sigma; St. Louis, MO) for 5 min. Five aliquots of 350 μl each were extracted using the Dynabeads® SILANE Genomic DNA Kit (Life Technologies, Carlsbad, CA) as per manufacturer instructions, and the elutions were pooled. 1.5 μl of the pooled eluate was treated with 1 μl Turbo DNase I (Life Technologies, Carlsbad, CA), 1.5 μl Turbo DNase I buffer, and 1 μl molecular grade water in a total of 15 μl and incubated for 30 min at 37 °C followed by 20 min at 95 °C to inactivate the DNase I. The controls for this study included heat inactivated DNase I treated eluate, untreated eluate, elution buffer and elution buffer plus DNase I buffer. For the heat inactivated DNase I control, Turbo DNase I was first inactivated for 20 min at 95 °C, and then combined with the eluted DNA and Turbo DNase I buffer as described above. Fifteen microliters of the untreated eluate, elution buffer and elution buffer plus DNase I buffer were added directly to the PCR reaction. Five thousand copies of H37Rv MTB genomic DNA was amplified PCR using the MTB IS6110 assay.

2.5.2. Dried reagent tubes

Two tubes per reaction were prepared with dried reagents: one for sputum thinning and the second for specific capture probe (oligo) binding. The sputum thinning buffer was prepared by combining 50 μl of 20% SDS, 30 μl of 1 M Tris, pH 8.0; 20 μl of 0.5 M
Figure 1. Development of qPCR assay for sample prep development and DNase I treatment of samples removes qPCR inhibition of IS6110 assay. (A) IS6110 qPCR assay. Standard curve of 500,000; 50,000; 5000; 500; 50, and 5 MTB H37Rv genomic DNA copies in triplicate. The equation of the line is $y = -3.7x + 35.5$; $R^2 = 0.997$. qPCR efficiency = 89%. (B) Means diamonds plot of DNase I treated samples. Eluates from bulk sputum extractions were spiked into control qPCR assay to measure PCR inhibition. The center line of the diamond is the mean. The tips of the diamond are the 95% confidence interval and the lines near the tips indicate significant differences if they don't overlap with other diamonds. $n = 6$ DNase I treated eluates, 3 DNase I buffer controls, 3 Tris buffer controls, and 6 untreated eluates. Heat-killed DNase I data not shown. 4 of 6 replicates failed to amplify.
EDTA, pH 8.0 per reaction for a final concentration in 1 ml of 1% SDS, 30 mM Tris pH 8.0 and 10 mM EDTA, pH 8.0. One hundred microliters of this SDS solution was pipetted on to the side of sterile 1.5 ml conical-bottom tubes; the tubes were kept in a horizontal position and dried in a 55 °C oven overnight. The oligo binding buffer was prepared by combining 50 μl of 5 M NaCl, 10 μl of 1 M Tris, pH 8.0, 2 μl of 0.5 M EDTA, pH 8.0 and 0.5 μl of 10% Tween-20 per reaction for a final concentration of 250 mM NaCl, 10 mM Tris, pH 8.0, 1 mM EDTA and 0.005% Tween-20. Sixty-two and one half microliters of oligo binding buffer was pipetted on to the side of a 1.5 ml sterile conical tube and kept in a horizontal position while dried in a 55 °C oven overnight. After drying, the tubes were capped and stored at room temperature in aluminum moisture-barrier pouches (Ted Pella; Redding, CA), with silica gel desiccant (Mcmaster Carr; Elmhurst, IL) to maintain dryness and a humidity indicator card (Static Control Components; San Diego, CA) to monitor moisture.

A cocktail of biotin-labeled oligonucleotides was prepared in advance which 5 μl contained a total of 5 or 5.5 pmol of probe: for the singleplex IS6110 assay, 2.5 pmol of IS6110 capture probes 1 and 2 were used, and for the multiplexed assay, 1.25 pmol each of IS6110 capture probes 1 and 2, 1.25 pmol each of senX3-regX3 capture probes 3 and 4, and 0.25 pmol each of cotfC capture probe 1 and capture probe 2 were used. All capture oligonucleotides contained a 5' end biotin moiety, included a spacer of 5 adenine residues prior to the specific sequences described above, and were HPLC-purified. Probes were obtained from Integrated DNA Technologies (IDT; Coralville, IA), were diluted in 10 mM Tris pH 8, and stored at −20 °C until time of use.

2.5.3. Sputum liquefaction, decontamination and lysis

To initiate sputum thinning, 950 μl sputum, 50 μl of Proteinase K solution containing 30 U Proteinase K, and 1 μl 1 M CaCl₂ were added to a 1.5 ml tube containing dried thinning reagent. The samples were heated to 55 °C for 8 min with mixing at 1000 rpm (Benchmark Scientific, Inc; South Plainfield, NJ). The temperature was then ramped to 95 °C, and the specimens were incubated for 10 min with 1000 rpm mixing.

2.5.4. Specific capture based MTB DNA extraction

Samples were removed from the thermal mixer, spun briefly, and the entire volume was transferred to the tube containing dried oligo binding buffer. Five microliters of biotin-labeled oligonucleotide cocktail mix (5 pmol in total) was added, and the samples were incubated at 60 °C and 1000 rpm for 20 min. Twenty microliters of pre-washed Dynal Streptavidin M–270 paramagnetic particles (PMPs) (Thermo Fisher Scientific; Waltham, MA) were added and samples were incubated for 10 min at room temperature with end-over-end rotation. The liquid was collected to the bottom of the tube by a short centrifugation step and the PMPs were pelleted on a magnetic rack (Dynabeads™) for 2 min. Supernatants were aspirated, the PMP pellet was resuspended in 1 ml wash buffer (10 mM Tris, pH 8.0 & 0.01% Tween 20), and the entire volume was transferred to a new 1.5 ml tube. Again, PMPs were collected on a magnetic rack, followed by supernatant aspiration and resuspension of PMPs in 1 ml wash buffer. Following this second wash, PMPs were collected on a magnetic rack and the supernatant was aspirated, taking care to remove all remaining wash buffer volume. The pellet was carefully resuspended in 10 μl elution buffer (10 mM Tris, pH 8.0 0.01% Tween 20, and 10% glycerol) and heated at 75 °C with 1500 rpm mixing, for 3 min. It is important that PMPs are resuspended in the elution buffer during the heated elution step. The liquid was collected at the bottom of the tube by a brief spin, and the samples were placed on the magnetic rack to collect PMPs. The 10 μl elution was carefully transferred to clean tube ready for qPCR.
2.5.5. Bulk DNA capture

Prior to bulk DNA extraction using the Dynabeads® SILANE Genomic DNA Kit, the sputum specimens were either thinned by NALC treatment or by the addition of GuHCl, as specified in the kit's instructions. Fresh NALC reagent was prepared by adding equal volumes of 4% NaOH and 2.9% sodium citrate with 0.5 g NALC per 100 ml NaOH/citrate solution [23]. To each sputum specimen, an equal volume of the NALC reagent was added and mixed by vortexing, then incubated for 15 min at room temperature with intermittent shaking. The samples were neutralized by the addition of phosphate buffer [pH 6.8] at a ratio of 1:1.5 (sample: phosphate buffer) and mixed by vortexing. The pellet was collected via centrifugation for 15 min at 3000 × g, resuspended with 350 μl phosphate buffer and immediately processed using the Dynabeads® SILANE Genomic DNA Kit according to the manufacturer's instructions. The GuHCl thinned sputa were immediately processed using the Dynabeads® SILANE Genomic DNA Kit according to the manufacturer's instructions.

2.6. Laboratory validation study

The Xpert® MTB/RIF Assay was used as the clinical standard since South Africa has adopted Xpert® MTB/RIF Assay (Cepheid; Sunnyvale, CA) testing on sputum samples as the standard first-line test for diagnosis of MTB, with MTB culture only performed on selected samples. Sixty de-identified blinded sputum specimens were tested; 27 were Xpert® MTB/RIF positive and 33 were Xpert® MTB/RIF negative. One of the Xpert negative specimens was invalid (no Cq in cotJC test).

2.7. Statistical methods

JMP® software (SDS Institute; Cary, NC) was used to calculate differences between means via one-way analysis of variance for independent samples, as well as to generate means diamonds plots. P-values of less than 0.01 were considered significantly different. Pearson correlation tests were used to generate R² values to evaluate correlations between sputum-specific factors and MTB assay inhibition.

3. Results

3.1. IS6110 qPCR assay

The MTBC-specific potentially multi-copy insertion sequence IS6110 [19] was selected as the genetic target for a qPCR assay to determine the efficacy of MTB DNA extraction protocols from sputum. The primers and probes were designed with 100%
sequence homology to all complete MTBC species sequences available in the NCBI database. The assay was demonstrated to be linear across at least 6 logs of concentration from 500,000 to 5 copies (Figure 1A) and could detect as few as 1 MTB H37Rv genome 6 out of 6 times (data not shown).

3.2. **MTB extraction from sputum**

Multiple investigators have reported that silica-based extraction methods such as the Boom method [24] yield MTB DNA with less PCR inhibition than alternative methods [12,25,26]. The Dynabeads® SILANE Genomic DNA Kit (Life Technologies, Carlsbad, CA) protocol is designed to extract genomic DNA from 350 µl whole blood using 50 µl paramagnetic particles (PMP), but input volume can be adjusted to suit specific experimental needs. In order to evaluate bulk DNA extraction, initial studies were performed with 350 µl contrived TB-positive sputum specimens with the intention of ultimately scaling up the sample volume to 1 ml. However, in extracting MTB DNA using this system, the qPCR results differed markedly between independent sputum samples. By amplifying 10-fold dilutions of the eluted DNA and quantifying the yield using a standard curve derived from serial dilutions of MTB H37Rv genomic DNA, it was determined that the MTB DNA yield from the Dynabeads® extraction was similar between samples, but qPCR inhibition varied dramatically (data not shown). In order to identify the potential source of inhibition, 7 sputum samples were evaluated for: appearance (bloody, mucopurulent, etc.), density, viscosity, protein levels and human genomic DNA levels and compared these results to qPCR inhibition. Of these variables, only human genomic DNA (gDNA) content correlated with inhibition (R² = 0.985, Pearson correlation test). Three of the 7 samples displayed qPCR inhibition of greater than 1 Cq, and these 3 samples also had high levels of co-extracted human gDNA as measured by copy number of the single copy gene for β-globulin (data not shown). The other four sputa tested had much lower β-globulin copy number in the eluted sample and no evidence of qPCR inhibition.

Co-extraction of high amounts of human gDNA was confirmed to be the source of inhibition by treating sputum-derived DNA eluates with Turbo DNase I for 1 h at 37 °C, heat-inactivating the enzyme and then adding the eluates or equal volume of controls to qPCR assays containing 5000 genomic copies of H37Rv genomic DNA. Relief of inhibition was observed in the specimens that were treated with active DNase I (Figure 1B). The Tris buffer control had an average Cq of 22.0 ± 0.1, and the untreated eluate from sputum extraction had an average Cq of 26.0 ± 0.3. This four Cq delay would lead to more than a 10-fold under-quantification of the MTB bacterial load. Treatment of eluate with DNase I allowed a mean recovery of 5.4 Cq. DNase I activity was required for this relief of inhibition because heat-inactivated DNase I did not improve amplification (data not shown). In fact, 4 of the 6 replicates incubated with inactivated DNase I failed to amplify while the remaining two samples had CqS of 24.6 and 28.5, similar to the untreated eluate. Eluates from 3 additional sputum extractions gave analogous results with DNase I treatment (data not shown) indicating that human genomic DNA in sputum is a frequent and potent qPCR inhibitor.

3.3. **Development of sequence specific capture method**

To limit the co-extraction of human DNA from sputum specimens as observed with bulk DNA extraction, a protocol to selectively purify mycobacterial DNA was developed using sequence specific capture. The protocol is divided into two phases (Figure 2): 1) sputum thinning and MTB lysis (steps 1–3) and 2) MTB DNA capture (steps 4–6). The sputum is thinned to a pipettable consistency and simultaneously sterilized with minimum operator steps. The process involves treating with Proteinase K (PK) and SDS with a subsequent incubation at 95 °C for 10 min. The SDS acts as an emulsifier breaking up mucoid complexes by unfolding proteins, and PK facilitates sputum thinning by digesting the unfolded proteins. Heating to 95 °C decontaminates the specimen by killing MTB and other flora present, while simultaneously denaturing the PK and initiating the specific capture reaction through melting the double stranded DNA in the specimen. The thinned sputum is then transferred to a tube containing dried NaCl and MTB-specific biotinylated oligonucleotides (capture probes) where a two-step capture process using streptavidin-coated paramagnetic particles (PMP) is performed. The high affinity of biotin and streptavidin allows for the biotinylated oligo-DNA complex to be isolated from the thinned sputum solution for washing. The eluted DNA is subsequently amplified and detected using real-time PCR.

3.4. **Sequence specific capture versus silica bulk DNA capture**

Sixteen MTB-negative sputum specimens with sufficient volume (≥1.8 ml) were selected for the study. MTB-spiked sputum specimen were thinned in one of 3 ways: 1) NALC, 2) GuHCl, or 3) SDS/PK, and 350 µl of the thinned sputa were extracted by either Dynabeads® SILANE Genomic DNA Kit (NALC & GuHCl) or sequence specific capture (SDS/PK) (Figure 3A). Human gDNA yield measured in β-globulin copies (Figure 3B), MTB DNA yield measured in IS6110 copies (Figure 3C), and qPCR inhibition rating (IR) was determined by comparing the amplification of the straight eluate versus the 1:10 dilution of the eluate. By using the SDS/PK thinning technique with MTB specific capture, the average amount of human genomic DNA co-extracted was reduced 2.5 to 3.7 logs. Mean β-globulin copy number co-extracted from 350 µl of sputum was 6.2 ± 0.7 log for NALC treatment, 7.4 ± 1.4 log for GuHCl treatment, and 3.7 ± 0.7 log for SDS/PK treatment. The IR was zero for all 16 of the NALC-treated or SDS/PK-treated samples indicating no qPCR inhibition. However, 75% (12/16) of the GuHCl treated specimens exhibited qPCR inhibition with the amount of β-globulin co-extracted corresponding to qPCR inhibition (Figure 4). The mean β-globulin log copy yield of GuHCl treated samples with an IR of 0 (5.9 ± 1.6) was significantly different (p < 0.01) from IR1 (7.8 ± 0.5) and IR2 (8.1 ± 0.4), but the mean yields of IR1 and IR2 were not significantly different.

The yield of MTB DNA was measured using 10⁻¹ dilutions of the eluate to bypass the effect of qPCR inhibition discussed above. There were no significant differences between mean MTB yield of the GuHCl treated specimens (4.7 ± 1.0 log copies), the NALC treated specimens (4.1 ± 0.6 log copies), or the SDS/PK treated specimens (4.1 ± 0.8 log copies). It was estimated that at least 1 ml of sputum must be extracted and amplified in order to achieve equivalent sensitivity to culture (10 cfu/ml) [2]. Because only 10% of the eluate from a 350 µl specimen was added, roughly 30-fold less specimen volume was used in this experiment than the target 1 ml. If the average genomic DNA yield were multiplied by a factor of 30, the average yield of β-globulin copies would increase to 7.7 logs for NALC, 8.9 logs for GuHCl, and 5.2 logs for SDS/PK thinning with specific capture extraction. The amount of co-extracted β-globulin DNA in the NALC treated specimens would then be equivalent to the average amount of DNA found in the inhibited GuHCl specimens, suggesting that 1 ml NALC-treated samples would also likely be inhibited. The SDS/PK treated specimens average gDNA yield for 1 ml of sputum would still be below the likely inhibited level. Due to the specific capture based DNA extraction resulting in equivalent IS6110 copies (Figure 3B), but
significantly lower human gDNA (Figure 3C), it was selected as the preferred nucleic acid extraction strategy for the MTB diagnostic assay.

3.5. Multiplexed qPCR assay

Since IS6110 is not present in some clinical strains [27], the IS6110 assay was multiplexed with primers and probes specific to an intergenic region in the two-component regulatory operon senX3-regX3 [20]. An amplicon of the B. atrophaeus cotJC gene that encodes a spore coat composition polypeptide [28] was also included to monitor the test for extraction, qPCR inhibitors, and amplification efficiency. The B. atrophaeus spores were added to the sputum prior to sample processing and were thinned and lysed along with the MTB. Capture probes were designed for each specific target, tested in isolation and then in combination for optimal specific MTB DNA extraction (Figure 5). Contrived sputum specimens of 5000 cfu of MTB H37Ra and 500 B. atrophaeus spores in triplicate were extracted using the specific capture MTB DNA extraction method, and the multiplexed PCR assay was performed to confirm capture of multiple targets and assay performance. The IS6110 amplicon had an average Cq of 24.8 ± 0.4, the senX3-regX3 had an average Cq of 28.7 ± 0.5 and the cotJC amplicon had an average Cq of 32.1 ± 0.6 (Figure 5).

To estimate the limit of detection, 1 ml sputum specimens were contrived containing 100, 50 or 20 cfu/ml, and six replicates of each concentration were assayed. All 6 of the 50 and 100 cfu/ml specimens were detected by both the IS6110 and senX3-regX3 assays. Five of the 20 cfu/ml specimens were detected by the IS6110 assay, and none of these samples were detected by the senX3-regX3 assay. Therefore, the LOD of the combined assay was estimated to be 20 cfu/ml. One of the 20 cfu/ml samples was excluded from analysis because the Cqs in both MTB assays were anomalously low. MTB has a tendency to clump, and despite the fact that the bacterial preparation was sonicated to break up clumps before it was serially diluted for contriving the sputum specimens; this sample may have had a clump of bacteria. Six different Mycobacterium species were also tested for potential cross reactivity. Genomic DNA extracted from M. gordonae, M. intracellulare, M. terrae, M. malmoense, M. celatum, and M. abscessus was added to qPCR reactions at 50,000 copies/PCR reaction, and no amplification was observed for any of these species.

3.6. Laboratory validation study of sputum processing and work flow

To determine if the SDS/PK sputum pre-treatment combined with sequence specific capture is effective in processing clinical specimens, a small-scale laboratory validation study was performed at the routine TB diagnostic laboratory at Groote Schuur Hospital, Cape Town, South Africa. A convenient sample of 60 blinded routinely collected specimens that had 1 ml of sputum left over after GeneXpert® testing were assayed using the multiplexed MTB assay protocol described here; 27 Xpert® MTB/RIF positive specimens and 33 negative specimens were analyzed. All 60 distinct sputum samples thinned to a pipetable liquid in the sputum thinning step. Of the 27 positive specimens, three samples were negative in senX3-regX3 assay and one of these same samples was also negative in the IS6110 assay yielding a sensitivity of 88.9% (24/27) and 96.3% (26/27), respectively (Table 2). One of the Xpert® MTB/RIF positive specimens and 33 negative specimens were analyzed. All 60 distinct sputum samples thinned to a pipetable liquid in the sputum thinning step. Of the 27 positive specimens, three samples were negative in senX3-regX3 assay and one of these same samples was also negative in the IS6110 assay yielding a sensitivity of 88.9% (24/27) and 96.3% (26/27), respectively (Table 2). One of the Xpert® MTB/RIF negative specimens failed to amplify the cotJC gene and was therefore considered an invalid result. The sample that was a false negative, having failed in both the IS6110 and senX3-regX3 assays, was reported as a very low positive in the Xpert® MTB/RIF assay. The 32 negative specimens with valid test results were also negative in the IS6110 and senX3-regX3 assays, yielding 100% specificity in both assays.
A major contributor to MTBC-specific qPCR assay inhibition. Furthermore, high amounts of human genomic DNA could have a secondary technical effect of making the PMPs sticky, leading to greater co-extraction of inhibitors from either the biological specimen or DNA extraction solvents such as guanidinium or alcohols. Guanidinium hydrochloride treated sputum yielded the highest levels of co-extracted human DNA among the three DNA extraction methods tested, as reflected by the single-copy human gene β-globulin, as well as strongest assay inhibition.

In this study, one ml sputum specimen was introduced into the tube containing dried sputum thinning reagent for sputum liquefaction and decontamination. In this regard, the sample prep strategy mimics the processing of sputum for mycobacterial culture growth, but the cumbersome sample neutralization and centrifugation steps have been eliminated. This strategy also addresses the biohazard risk associated with inadequate management of infectious samples and the limitations of sample preparation methods to enhance the overall sensitivity of the assay. Sputum collection and manipulation from putative TB patients puts healthcare workers at risk for infection [30,31], and, for NAT testing in settings without biosafety precautions, it is necessary to completely sterilize the sample. This MTB screening assay is able to use a larger volume of sputum (1 ml) which is fully homogenized by treatment with SDS and proteinase K thus ameliorating sampling error. Although it was not explicitly demonstrated in this study, the 95 °C incubation step reduces the biosafety hazard to the test operator by fully lysing the bacteria and melting the double stranded DNA to prepare for the subsequent sequence specific capture step. This method is not reliant on a BSL-3 facility thereby expanding its potential to be applied in limited-resource settings or even at the patient point-of-care, nor is it reliant on centrifugation, therefore improving sensitivity by using both intact and lysed mycobacterial DNA [18] that would otherwise be discarded in supernatants of centrifugation-based extraction methods. By using paramagnetic particles (PMP), the MTB DNA can be extracted from a larger specimen volume, and the sputum-derived inhibitors can be removed with simple wash steps. This sputum processing protocol is currently being adapted
so that it may be performed directly in the sputum collection cup, in order to minimize operator handling and potential for cross-contamination.

It was recently reported that in patients who have been previously treated for TB, 1 in 7 Xpert® MTB/RIF-positive results were false-positive relative to mycobacterial culture in a high incidence setting [32]. False positives are correlated with a previous infection that occurred in the last 3 years [33]. The test described here is designed to lyse all the bacteria before an aliquot is removed to reduce sampling error and promote safety of the operator. Therefore, testing the specificity of this assay with specimens from recently recovered TB patients must be carefully evaluated. The test was designed to have a high negative predictive value so that it can be used to rule out a TB diagnosis. Additional algorithms designed to determine if a patient has been treated for TB in the past 3 years could be used to increase the specificity of the test in previously infected patients [32,33].

This study also lends merit to the use of specific capture based DNA extraction in other applications of clinical diagnostics and in other clinical sample types where target DNA may be underrepresented relative to a large pool of human DNA or other potential assay inhibitors. Compared to the two other common sputum processing and DNA extraction methods discussed here (NALC or GuHCl treatment with bulk DNA extraction), while specific capture based DNA extraction resulted in an equivalent level of MTB-specific IS6110 copies, there was a significant reduction (2.5–3.7 log copies β-globulin DNA) in co-eluted human genomic DNA. When tested on a panel of TB positive and negative clinical spuata, showed a sensitivity of 88.9–96.3% for senX2-regX3 and IS6110, respectively relative to Xpert® MTB/RIF, with 100% specificity in both targets. The assay now requires further validation using samples that test culture-positive, Xpert® MTB/RIF negative to determine whether the clinical accuracy exceeds that of Xpert® MTB/RIF.

The WHO Millennium Development Goal of stopping the increase of TB incidence and halving the TB-associated mortality by 2015 was achieved globally, but not in Africa [1,34]. The next goal set by WHO is to eliminate TB by 2050. This will require a reduction of TB of 16% per year compared to the current rate of 1% per year. In order to achieve these goals, major advances in diagnostics, vaccines and social policy are required [35]. Increasing diagnostic sensitivity is capable of dramatically improving detection in highly vulnerable populations such as children and people living with HIV. Specifically, this will require targeted development of diagnostic tests designed to address the challenges of MTB including the intrinsic properties of difficult clinical sample types that may limit diagnostic assay performance, and barriers of implementation in settings with high HIV and TB incidence.

Acknowledgements

This work was funded by the Northwestern Global Health Foundation and Quidel Corporation. The study sponsors played no role in the design of any of the studies, in the collection, analysis and interpretation of data; in the writing of the manuscript; nor in the decision to submit the manuscript for publication.

Funding: None.

Competing interests: None declared.

Ethical approval: Not required.

References


