A novel multiplex real-time PCR diagnostics assay for the identification and
differentiation of *Mycobacterium tuberculosis*, *Mycobacterium canettii* and
the *Mycobacterium tuberculosis* complex

Kate Reddington¹, ³†, Justin O’Grady¹, ³†, Siobhan Dorai-Raj¹, ³, Majella Maher³,
Dick van Soolingen², Thomas Barry¹, ³*

**Running Title:** Multiplex *Mycobacterium* real-time PCR assay

Microbiology, School of Natural Sciences, National University of Ireland, Galway,
Ireland ¹; National Tuberculosis Reference Laboratory, National Institute for Public
Health and the Environment, Bilthoven, The Netherlands ²; Molecular Diagnostics
Research Group, NCBES, National University of Ireland, Galway, Ireland ³

* Corresponding Author. Mailing Address: Microbiology, School of Natural Sciences,
National University of Ireland, Galway, University Road, Co. Galway, Ireland.
Telephone: +35391493189, Fax: +35391494598,
Email: thomas.barry@nuigalway.ie

† Both authors contributed equally to this work

¹ Present address: Department of Infection, Windeyer Institute of Medical Sciences,
University College London, London W1T 4JF
Abstract

Tuberculosis (TB) in humans is caused by members of the Mycobacterium tuberculosis complex (MTC). Rapid detection of the MTC is necessary for the timely initiation of antibiotic treatment, while differentiation between members of the complex may be important to guide the appropriate antibiotic treatment and provide epidemiological information.

In this study, a multiplex real-time PCR diagnostics assay using novel molecular targets was designed to identify the MTC while simultaneously differentiating between M. tuberculosis and M. canettii. The lepA gene was targeted for the detection of members of the MTC, the wbbL1 gene was used for the differentiation of M. tuberculosis and M. canettii from the remainder of the complex and a unique region of the M. canettii genome, a possible novel “region of difference” (RD), was targeted for the specific identification of M. canettii.

The multiplex real-time PCR assay was tested using 125 bacterial strains (64 MTC isolates, 44 non-tuberculosis mycobacteria (NTM) and 17 other bacteria). The assay was determined to be 100% specific for the mycobacteria tested. Limits of detection of 2.2, 2.17 and 0.73 cell equivalents were determined for M. tuberculosis/M. canettii, the MTC and M. canettii, respectively, using Probit regression analysis.

Further validation of this diagnostics assay, using clinical samples, should demonstrate its potential for the rapid, accurate and sensitive diagnosis of TB caused by M. tuberculosis, M. canettii and the other members of the MTC.
1.0 Introduction

Tuberculosis (TB) is the leading cause of death worldwide from an infectious agent (13), with the WHO estimating that one third of the global population are infected with TB. In a global report from the WHO (2009), it was estimated that there was 9.27 million cases of TB in 2007, with 2 million associated deaths. TB in humans is caused by members of the *Mycobacterium tuberculosis* complex (MTC). The eight closely related species of the MTC have a wide range of natural hosts including humans hosts (*M. tuberculosis*, *M. africanum*, *M. canettii*), bovine hosts (*M. bovis*), caprine hosts (*M. caprae*), rodent hosts (*M. microti*) and pinniped hosts (*M. pinnipedii*), along with the attenuated *M. bovis* strain BCG (*Bacillus Calmette-Guérin*), the commonly used vaccine strain. While there are a number of natural hosts, each species of the MTC has been implicated in human infection (6, 20).

Traditionally, diagnosis of TB relies on smear microscopy and culture techniques in combination with a battery of biochemical tests which are time consuming, labour intensive and often yield unreliable results (18). Nucleic Acid Diagnostics (NAD) techniques, in particular real-time PCR, offer a rapid, reliable and highly sensitive alternative diagnostic tool for many infectious agents (25, 42). Advances in real-time PCR such as the availability of multiple fluorophores, along with the development of non-fluorescent quenchers has facilitated multiplexing, allowing for the simultaneous detection and differentiation of multiple targets, along with internal controls, in one reaction (3).

While significant advances have been made in the diagnosis of TB using NAD techniques (19), the differentiation of members of the MTC to the species level is not routinely performed. Commercially available real-time PCR kits for the diagnosis of TB generally identify the MTC but not individual species. The high
degree of nucleotide sequence homology between members of the complex makes species differentiation challenging (31). Comparative genomics revealed that *M. tuberculosis* and *M. bovis* genomes are 99.95% similar at the nucleotide level (14), with whole genome DNA microarrays identifying 16 regions of difference (RD 1-16) (4). These RD’s represent regions of the genome deleted in *M. bovis BCG* which are present in *M. tuberculosis* and have been used for the differentiation of members of the MTC. One RD commonly targeted for the specific detection of *M. tuberculosis* is RD9 (31), however, this RD is also present in *M. canettii* (7). There is currently no real-time PCR test which can diagnose TB, while differentiating between *M. tuberculosis* and *M. canettii*, as the causative agent of infection.

*M. tuberculosis* is the most important human pathogen in the MTC and is thought to be responsible for 95% of human cases of TB, yet rarely causes disease in other mammals (1, 7, 9). While drug resistant strains of *M. tuberculosis* are emerging, it is considered sensitive to anti-tuberculosis drugs such as Pyrazinamide (PZA), a first line antibiotic that reduces patient treatment time from 9 months to 6 months (27, 35). However, *M. canettii* which has been reported to cause TB in humans, is intrinsically resistant to PZA, therefore, the ability to differentiate it from *M. tuberculosis* is important for indicating the therapeutic regimen necessary for patient treatment (34).

*M. canettii* is considered to be the most phenotypically distinct member of the MTC and is considered the species from which other members of the complex may have evolved (6). *M. canettii* is phenotypically characterised by its smooth glossy white colonies, however a small number of these colonies have been shown to revert to rough colony variants when individual colonies are replated (37). Smooth colonies are uncharacteristic of the MTC and are due to the presence of large amounts of
lipooligosaccharides in the *M. canettii* cell wall (30). Like *M. tuberculosis*, *M. canettii* contains all the RD’s with the exception of RD12 *canetti* (RD12can) which has been targeted for the specific detection of *M. canettii* in a complex conventional PCR methodology (18).

While infection with *M. canettii* is thought to be rare and confined to eastern African countries, there is a lack of rapid diagnostic tests available to differentiate between it and *M. tuberculosis*. Cases of human TB caused by *M. canettii* have now been reported in Europe and America (12). In addition, recent reports have suggested that the number of true cases of TB caused by *M. canettii* may in fact be underrepresented (15, 34). Therefore, an ability to differentiate *M. tuberculosis* and *M. canettii* is not only important from a patient treatment perspective but will also provide important epidemiological information for clinicians.

We report the design, development and testing of a multiplex real-time PCR assay using novel nucleic acid diagnostics targets to detect the MTC while simultaneously detecting and differentiating between *M. tuberculosis* and *M. canettii* in one reaction.

2.0 Methods

2.1 Diagnostics target identification

The diagnostics target genes used in this study were identified using a number of approaches. In order to identify a target for collective detection of the MTC, a number of housekeeping genes, which are highly conserved throughout the *Mycobacterium* genus, were evaluated. To identify novel targets for the detection of *M. tuberculosis*, approximately 3000 genes were evaluated based on regions deleted in other members of the MTC but present in *M. tuberculosis* or alternatively, present in other members
of the MTC but deleted in *M. tuberculosis*. These potential target regions were identified using the Mycobacterial Genome Divergence Database (MGDD) (http://mirna.jnu.ac.in/mgdd/), which allowed for identification of insertions, deletions and single nucleotide polymorphisms between *M. tuberculosis*, *M. bovis* and *M. bovis BCG*. Potential target regions were also identified using the web based version of the Artemis comparison tool, WebACT (http://www.webact.org/WebACT/home).

Nucleotide sequence information was retrieved from the *M. africanum* and *M. microti* genomes (currently being sequenced by the Welcome Trust Sanger Institute) to determine, *in silico*, if the candidate diagnostics target nucleotide sequences for *M. tuberculosis* identification were specific. For the remaining members of the complex, namely, *M. canettii*, *M. caprae* and *M. pinnipedii*, for which no nucleotide sequence information was available at the time of this study, the specificity of potential targets for these species was determined empirically and further validated through sequencing.

As the diagnostics target used in this study for the specific detection of *M. tuberculosis*, was experimentally determined to also detect *M. canettii*, a novel diagnostics target specific for *M. canettii* was also required. When nucleotide sequence information became publicly available on the Sanger website for *M. canettii*, a number of large regions of nucleotide sequence were evaluated using BLAST for putative species specific nucleotide sequence diagnostics motifs.

For each putative diagnostics target identified, alignments were carried out using clustalW multiple sequence alignment programme (http://www.ebi.ac.uk/Tools/clustalw2/index.html), from which real-time PCR primers and probes were designed (Table 1).
2.2 Bacterial strains, culture media and growth conditions

Sixty four MTC isolates (26 M. tuberculosis, 11 M. bovis, 7 M. bovis BCG, 5 M. canettii, 2 M. caprae, 5 M. africanum, 5 M. microti and 3 M. pinnipedii), 44 non tuberculosis mycobacteria (NTM) and 17 other bacterial species were used in this study (supplementary Tables 1 and 2). Of the 64 MTC isolates, 36 strains, previously characterised by a variety of methods as described in the literature (12, 22, 23, 26, 30, 36-38), were provided by Dr. Dick van Soolingen (RIVM, Bilthoven, Netherlands). All other MTC strains, provided by Professor Mario Vaneechoutte (University of Ghent, Ghent, Belgium) were clinical isolates collected over a ten year period from reference laboratories in Belgium and the Netherlands. These isolates were characterised based on custom techniques available at the time. Twenty eight NTM were purchased from the German Collection of Microorganisms and Cell Cultures (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, DSMZ) and grown on Middlebrook agar/broth at either 30 °C or 37 °C. “Fast growing” mycobacteria were cultured for 3-6 days, and “slow growing” mycobacteria were incubated for six weeks, or until sufficient growth was visible. All media were purchased from BD Biosciences (Oxford, United Kingdom). For the 16 remaining NTM, in addition to 4 of the 5 Norcardia strains used in this study, DNA was supplied by Professor Vaneechoutte, which had been characterised using techniques previously described (10, 40, 41)

2.3 DNA isolation and quantification

Genomic DNA from NTM and 2 M. bovis BCG cultures was isolated from 1 ml of culture (Middlebrook 7H9 broth, Becton Dickinson), using a modified procedure combining mechanical lysis (IDI lysiskit, GeneOhm, Quebec, Canada) and
purification using a DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany).

Briefly, 1 ml of culture was centrifuged in a bench-top centrifuge (Microcentrifuge 5415, Eppendorf) at 13,000 rpm for 3 min. The supernatant was discarded and the pellet resuspended in 250 µl GeneOhm sample buffer. The suspension was transferred to a GeneOhm lysis tube and bead beaten (Mini-Bead-Beater-16™, Stratech, UK) for 3 min. After bead-beating, 200 µl was transferred to a sterile microcentrifuge tube and steps 3-8 of the “purification of total DNA from animal tissue” procedure in the Qiagen DNeasy Blood and Tissue kit were followed according to the manufacturer’s instructions. Total genomic DNA samples provided by the RIVM were extracted using methods previously described (39), while genomic DNA from Professor Vaneechoutte was extracted as described in a study by De Baere et al. (10). For all other bacterial species tested for, DNA was provided from stocks held within this laboratory. DNA concentrations for all NTM and members of the MTC used in this study were determined using the PicoGreen dsDNA Quantitation Kit (Molecular Probes, Eugene, Oregon, USA) and the TBS-380 mini-fluorometer (Invitrogen Corporation, California, USA). All DNA samples were stored at -20 °C before use.

2.4 Conventional and real-time PCR primers and hydrolysis probe design

Oligonucleotide primers and hydrolysis probes were designed in accordance with general recommendations and guidelines (11, 33), following alignments of each of the nucleic acid diagnostics target genes identified in this study. All primers and probes (Table 1) used in this study were supplied by MWG-BIOTECH AG (Essenberg, Germany). The primers used in this study were designed to have a melting temperature (Tm) of 58-61 °C, with all probes designed to have a Tm of 4-7 °C degrees higher. Hydrolysis probes were designed to be specific for each target
following published design guidelines (11). These parameters were adhered to during the design of monoplex assays so that the assays could be easily multiplexed after specificity and sensitivity testing was complete.

For the MTC and IAC assay, PCR primers MTC_IAC Fw and MTC_IAC Rv (Table 1), were designed to amplify a 155 bp fragment of the lepA gene for all members of the MTC and *M. smegmatis*. The MTC_IAC Fw primer was located at positions 618-634 bp and the MTC_IAC Rv primer located at positions 754-772 bp of the *M. tuberculosis* H37RV lepA gene. For the *M. tuberculosis* and *M. canettii* assay, *wbbl1*_Fw and *wbbl1*_Rv were designed to amplify a 146 bp fragment of the *wbbl1* gene. The *wbbl1*_Fw primer was located at positions 15-34 bp and the *wbbl1*_Rv primer located at positions 141-159 bp of the *M. tuberculosis* H37RV *wbbl1* gene.

Finally, the *M. canettii* specific assay was designed to amplify a 128 bp fragment of a 2869 bp region of the genome identified in this study as specific to *M. canettii* (an *M. canettii* specific assay could potentially be designed anywhere in this region). This 2869 bp region of the genome has been mapped to the *M. tuberculosis* H37Rv genome and is inserted between Rv0150c (hypothetical protein) and Rv0151c (gene name *PE1*, a PE family protein) at position 177,445 bp on the *M. tuberculosis* H37Rv genome.

All real-time PCR assays were initially tested for in a monoplex format, evaluating their specificity and sensitivity, using probes labelled with FAM and Black Hole Quencher 1 (BHQ1). After the monoplex real-time PCR assays were optimised, three of the four assay probes were labelled with different fluorescent dyes to allow for multiplex real-time PCR. The MTC probe was labelled with HEX and BHQ1, the *M. canettii* specific probe with ROX and BHQ2 and the internal amplification control (IAC) probe with Cy5 and BHQ2.
2.5 *Conventional PCR*

Conventional PCR was performed using the sequencing primers outlined in Table 1 on the iCycler iQ thermal cycler (Bio-Rad Laboratories Inc., California, USA). All reactions were carried out in a final volume of 50 µl, containing 5 µl 10X buffer (15mM MgCl$_2$), forward and reverse primers (0.2 mM final conc.), 2 µl Taq DNA polymerase (1 U/ µl, Roche Diagnostics, Basel, Switzerland), 1 µl dNTP mix (10 mM: deoxynucleoside triphosphate set (Roche Diagnostics), 2 µl of template DNA, 38 µl Nuclease free water (Applied Biosystems/Ambion, Texas, USA). The cycling parameters consisted of initial denaturation at 95 °C for 5 mins, followed by 35 cycles of denaturation at 95 °C (1 min), amplification at 55 °C (1 min), and extension at 72 °C (1 min), followed by a final elongation at 72 °C for 10 min.

2.6 *Sequencing*

Nucleotide sequence data for real-time PCR assay design was generated in this study or was obtained from either the National Centre for Biotechnology Information (NCBI), or the Sanger website (where partial nucleotide sequences for *M. canettii*, *M. africanum* and *M. microti* were available). The primers used for the real-time PCR assays were also used in conventional PCR to generate nucleotide sequence information for each of the assays developed. In addition, sequencing primers were designed to span the full 2869 bp *M. canettii* specific nucleotide sequence identified, to evaluate if this region is conserved in all *M. canettii* strains tested in this study and to identify potential diagnostics target nucleotide sequences for probe design. PCR products were generated according to section 2.5, followed by purification using the High Pure PCR product purification kit (Roche Diagnostics).
The purified PCR products were sequenced externally (Sequiserve, Vaterstetten, Germany).

2.7 Development of an IAC for real-time PCR

An IAC targeting the M. smegmatis lepA gene was developed for the multiplex real-time PCR. The IAC and MTC targets were both amplified using the same primer set, however the IAC probe, targeting M. smegmatis, was labelled with Cy5, while the MTC probe was labelled with HEX (Table 1). Titrations of MTC and IAC DNA were performed to determine the optimum concentration of IAC target per reaction such that it was always detected without impacting on detection of the primary MTC target (17). An IAC concentration of 500 genome equivalents per reaction allowed for the positive detection of the IAC at low concentrations or in the absence of primary target.

The lepA gene was chosen as the target for the IAC because sufficient sequence heterogeneity exists between the M. smegmatis and MTC lepA gene sequences for the design of independent, specific probes. There was also enough sequence homology, flanking these probe regions, to design one set of primers to amplify both MTC and IAC targets. This resulted in three rather than four primer pairs in the multiplex PCR, reducing assay complexity.

2.8 Real-time PCR

Monoplex real-time PCR was performed on the LightCycler 2.0 Instrument (Roche Diagnostics) using the LightCycler® TaqMan® Master kit (Roche Diagnostics). A final volume of 20 µl was used in each reaction, containing 5X master mix, forward and reverse primers (0.5 mM final conc.), FAM labelled probe (0.2 mM final conc.),...
template DNA (2 µl) and the volume adjusted to 20 µl with the addition of nuclease free dH$_2$O. The cycling parameters consisted of incubation for 10 min at 95 °C to activate enzymes and denature DNA followed by 50 cycles of 95 °C for 10 s and 60 °C for 30 s, followed by a cooling step at 40 °C for 10 s. The temperature transition rate for all cycling steps was 20 °C/s.

Multiplex real-time PCR reactions were carried out on the LightCycler 480 using LightCycler® 480 Probes Master kit (Roche Diagnostics). A final volume of 40 µl was used for each multiplex reaction. The optimised master mix contained 2X LightCycler 480 Probes Master (6.4 mM MgCl$_2$), forward and reverse primer (0.5mM final conc.), FAM labelled probe (0.4 µM final conc.), HEX, ROX and CY5 labelled probes (0.2 µM final conc.), template DNA (MTC 2 µl, IAC 2 µl, NTM 10 µl) and the volume adjusted to 40 µl with the addition of nuclease free dH$_2$O. The internal control DNA was diluted to contain 500 genome equivalents per 2 µl and the NTM contained ~10$^4$ genome equivalents per 10 µl.

The cycling parameters used were the same as those used on the LightCycler 2.0. The temperature transition rate, referred to as the ramp rate on the LightCycler 480 was 4.4 °C/s while heating and 2.2 °C/s while cooling. Prior to experimental analysis on the LightCycler 480, a colour compensation file was generated using the technical note outlined in the Advanced Software Functionalities of the operator manual (2).

2.9 Nucleotide sequence accession numbers

Partial *wbbl1* gene nucleotide sequences generated in this study for the five *M. canettii* strains (supplementary data Table 1) were deposited in GenBank with the following accession numbers: HQ625205 through to HQ625209. Nucleotide
sequences generated in this study for the proposed novel RD$_{canetti}^{1}$ were also deposited to Genbank with the following accession numbers: HQ625200 through to HQ625204.

Results 3.0

3.1 Diagnostics targets identification

The diagnostics nucleotide sequence target identified in this study for detection of the MTC in combination with detection of the IAC was lepA (Rv2404c). LepA is an elongation factor required for accurate and efficient protein synthesis capable of inducing back-translocation of mistranslocated tRNA’s. The lepA gene is present in all bacteria sequenced to date and codes for one of the most conserved proteins in bacteria (55-68 % amino acid sequence similarity between bacteria), with a homologue (Guf1) found in higher organisms (32).

For the specific detection of *M. tuberculosis* and *M. canettii*, a *wbbI* gene nucleotide sequence target (Rv3265c) was identified. The *wbbI* gene encodes rhamnosyl transferase, which inserts rhamnose into the cell wall and is thought to be essential for mycobacterial viability (24). Nucleotide sequence analysis of members of the MTC revealed a 12 base pair region of the *wbbI* gene present only in *M. tuberculosis* and *M. canettii*, which has been deleted in all other members of the MTC.

The *M. canettii* specific diagnostics nucleotide sequence target identified in this study is a region of the genome that appears to be deleted in all other members of the MTC. This 2869 bp region was discovered while mapping RD$_{12}^{can}$ to the unfinished genome sequence of *M. canettii* available on the Sanger website. This region appears to be a novel RD specific for *M. canettii*. To date, the *M. canettii*
genome is not annotated, therefore, the function of the gene/s in this region are unknown. The *M. canettii* diagnostics target region identified was BLAST analysed and revealed similarity to a putative ATP binding protein gene in *Nocardioides* species (Query coverage 52%, max indent 73%). This region was sequenced using the primers listed in Table 1 for the *M. canettii* strains used in this study. Sequence analysis revealed 100% similarity between the 5 strains and the sequence available on the Sanger website.

### 3.2 Assay design and development

While the guidelines for primer and probe design were adhered to as closely as possible, the high GC content (60-65%) of the *Mycobacterium* species did have an impact on assay design. The *wbbl1* specific probe was based on a region present in *M. tuberculosis* and *M. canettii* and deleted in other MTC, that is very G/C rich making probe design difficult. This probe was labelled with FAM and double the standard probe concentration (0.4µM/reaction) was used to improve the endpoint fluorescence, sensitivity and robustness of the assay.

### 3.3 Internal amplification control

In order for a result to be considered valid using the multiplex real-time PCR assay developed in this study, a positive signal must be obtained in at least one of the four detection channels on the LightCycler 480. If none of the assay targets or the IAC are detected, the result is considered invalid and must be repeated (17, 28). In this study, *M. smegmatis* DNA is spiked into the PCR master mix to act as an internal control target. Equally *M. smegmatis* cells could also be used as a process control, when spiked into patient samples before total genomic DNA extraction. This could then act
as a control for both DNA extraction and PCR efficiency when patient samples are tested for.

### 3.4 Specificity of the diagnostic assays

The specificity of each real-time PCR assay was confirmed both in monoplex and multiplex formats using the specificity panel listed in supplementary Tables 1 and 2. The *wbb1* assay was specific for the detection of the 26 *M. tuberculosis* and the 5 *M. canettii* strains. Figure 1 A shows the detection of 6 *M. tuberculosis* strains (circles) and 5 *M. canettii* strains (triangles). The remaining members of the MTC, the NTM and other bacteria were not detected. The 64 MTC strains were all detected in the MTC assay (Figure 1 B) and 44 NTM and 17 other bacteria were not detected. The *M. canettii* assay was specific for the 5 *M. canettii* isolates (Figure 1 C) and did not cross-react with the specificity panel. The specificity of the IAC assay was tested using the MTC panel and was specific for *M. smegmatis* DNA. As the MTC and IAC assays are competitive, the IAC is only detected at low concentrations or in the absence of primary target. In Figure 1 B, the *M. africanum* 3 sample tested in the assay is at a lower concentration than the other members of the MTC, hence a stronger amplification signal is observed in the Cy 5 channel (IAC) with this sample (Figure 1 D, *M. africanum* 3 represented by diamonds). A simple interpretation of the results that can be obtained from the multiplex real-time PCR developed is outlined in Table 2.

### 3.5 Sensitivity of the assays

The limit of detection (LOD) of each assay was evaluated in a monoplex real-time PCR format. Genomic DNA was quantified and serial dilutions were prepared from
200,000 to 2 genome equivalents based on the genome size of *M. canettii* (4,525,000 bp), with equates to approximately 4.9 fg DNA per cell. *M. canettii* was chosen for sensitivity testing of the multiplex assay as it is detected in each of the primary target assays (*M. tuberculosis*/*M. canettii*, MTC and *M. canettii* specific).

In a monoplex format, the dilution series was run in duplicate and a sensitivity of 2-20 *M. canettii* genome equivalents was determined for each assay. In multiplex format the lower limit of detection was established using probit regression analysis. In this analysis, 12 replicates of each of 20, 15, 12, 10, 7.5, 4, 2, 0.2 genome equivalents were evaluated. LOD’s of 2.17, 2.20, 0.73 genome equivalents for the *M. canettii*/*M. tuberculosis*, MTC and *M. canettii* specific assays respectively were determined with 95% probability. The IAC, at a concentration of 500 genome equivalents per reaction, was detected in all samples tested.

### 4.0 Discussion

In 1998 the ‘Stop TB’ initiative was established with the ultimate goal of obtaining a world free of TB. Within this initiative a core group, the New Diagnostics Working Group (NDWG), was founded in an effort to establish new diagnostics for TB. This working group, along with a number of collaborators, such as FIND (Foundation for Innovative New Diagnostics), have launched a new web resource ‘Evidence-based Tuberculosis Diagnosis’ which highlights the importance of new diagnostics for the rapid and cost effective detection of TB (29)

At present, literature describing TB NAD assays, in addition to commercially available TB NAD kits (5, 8, 16), are limited in their scope and capacity to differentiate the MTC to the species level. For example, the GenoType MTBC kit (Hain Lifescience GmbH, Nehren, Germany), which claims to differentiate the MTC,
does not in fact differentiate between *M. tuberculosis* and *M. canettii* nor between *M. africanum* and *M. pinnipedii* (21). Recently a study by Pinsky and Banaei (2008), which describes a real-time PCR for differentiation of the MTC, fails to differentiate *M. tuberculosis* and *M. canettii*, and is also unable to differentiate between *M. africanum*, *M. microti*, *M. pinnipedii* and *M. caprae*. A capability to identify the specific MTC species causing infection is important for determining the appropriate therapeutic regimen for the patient (34).

The multiplex real-time PCR assay described in this study uses novel nucleic acid diagnostics targets for the identification of the MTC, *M. tuberculosis* and *M. canettii*. The first novel target identified was a *lepA* gene nucleotide sequence. This target was used to detect the MTC and to develop the IAC for the real-time PCR diagnostics assay. The second novel molecular target identified and evaluated in this study was a *wbbl1* gene nucleotide sequence target which enables the simultaneous detection of *M. tuberculosis* and *M. canettii*, a target with the same diagnostics potential as the widely used RD9 region for *M. tuberculosis* identification. The third novel diagnostics target nucleotide sequence identified, allowed for the differentiation of *M. canettii* from *M. tuberculosis*. As this nucleotide sequence is present in the 5 *M. canettii* strains used in this study, and is 100 % homologous, we propose this to be a novel RD (RD<sub>canettii</sub>1).

An *M. canettii* RD has previously been described by Huard et al. (2003) which represents a region of the genome flanking RD 12 which is deleted in *M. canettii* but present in *M. tuberculosis*. The study uses conventional PCR for differentiation of the MTC including *M. tuberculosis* and *M. canettii*, based on PCR product size. The method requires time consuming multiple reactions and produces results that require detailed interpretation. In this current study, we have identified a
putative new RD which is present in *M. canettii* but deleted in *M. tuberculosis* and all other members of the MTC. As this region is only present in *M. canettii*, the interpretation of results is less complex.

The multiplex real-time PCR developed in this study is the first description of a hydrolysis probe based diagnostic tool capable of rapid detection of the MTC, combined with the detection and differentiation of *M. tuberculosis* and *M. canettii*, using novel targets. This rapid, specific and sensitive multiplex real-time PCR assay produces a diagnostic result in less than one hour after DNA extraction. These novel genetic markers will be further optimised and validated against a more extensive range of clinical isolates in addition to clinical samples in the future.

While this current study has focused on identifying novel nucleotide sequence diagnostics targets and the development of a multiplex real-time PCR for differentiation of *M. tuberculosis* and *M. canettii*, work has begun in this laboratory on developing a series of nucleic acid based diagnostics assays with the ability to differentiate all species of the MTC. For example, there is a requirement to differentiate infection caused by both *M. bovis* and *M. bovis BCG* from the remainder of the MTC as they, like *M. canettii*, are intrinsically resistant to PZA (35).

Ultimately, there will be a requirement to transfer this set of assays to a platform with a capability of detecting and differentiating the MTC in a single test diagnostics format. For example, the design of a microarray utilising these novel MTC diagnostics targets, in combination with diagnostics targets for the determination of MTC drug resistance, could be significantly advantageous. This rapid diagnostics approach would provide the clinician with important information as to the optimal therapeutic regimen required for an infected patient, while at the same time providing the clinical laboratory with unambiguous MTC epidemiological data.
Acknowledgements

We thank Dr. Kristen Kremer (RIVM) for supplying genomic DNA’s for some MTC strains used in this study. We also thank Enterprise Ireland for partial funding of this project. Finally, we wish to acknowledge the National University of Ireland, Galway who have funded KR with a postgraduate scholarship.
5.0 Bibliography


25. Malhotra-Kumar, S., K. Haccuria, M. Michiels, M. Ieven, C. Poyart, W. Hryniewicz, H. Goossens, and on behalf of the MOSAR WP2 Study


<table>
<thead>
<tr>
<th>Name</th>
<th>Function</th>
<th>Sequence 5’→3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTC_IAC Fw</td>
<td>Forward Sequencing primer, forward MTC and internal control real-time PCR assay primer</td>
<td>AGACCGTGCGGATCTTG</td>
</tr>
<tr>
<td>MTC_IAC Rv</td>
<td>Reverse Sequencing primer, Reverse MTC and internal control real-time PCR assay primer</td>
<td>CATGGAGATCACCGTG</td>
</tr>
<tr>
<td>MTC Probe</td>
<td>MTC probe</td>
<td>HEX-ACGGATTTGTCACCGGATT-BHQ1</td>
</tr>
<tr>
<td>IAC probe</td>
<td>Internal control probe</td>
<td>CY5-ACGACCTTCTCGGAAACGT-BHQ2</td>
</tr>
<tr>
<td>wbbl1_Fw</td>
<td>Forward sequencing primer, Forward real-time PCR assay primer</td>
<td>TACCAGCTTCAGTTCCGT</td>
</tr>
<tr>
<td>wbbl1_Rv</td>
<td>Reverse sequencing primer, Reverse real-time PCR assay primer</td>
<td>GCACCTATATCTTCTTACCGG</td>
</tr>
<tr>
<td>wbbl1 probe</td>
<td>wbbl1 probe</td>
<td>FAM-ATGGTGCGCAGTTCACTGC-BHQ1</td>
</tr>
<tr>
<td>M. canetti sp Fw</td>
<td>Forward M. canetti specific primer</td>
<td>ATGTGTTTTGACTAGCAGTTC</td>
</tr>
<tr>
<td>M. canetti sp Rv</td>
<td>Reverse M. canetti specific primer</td>
<td>GATGGCAGTGTCTTACCAA</td>
</tr>
<tr>
<td>M. canetti sp probe</td>
<td>M. canetti specific probe</td>
<td>ROX-TGAGAGGTGTGGCAGC-AA-BHQ2</td>
</tr>
<tr>
<td>M. canetti seq 1.a</td>
<td>Forward sequencing primer 1</td>
<td>TGTCGGCGCCACAGT</td>
</tr>
<tr>
<td>M. canetti seq 1.b</td>
<td>Reverse sequencing primer 1</td>
<td>GAAGTCACATCTTGGC</td>
</tr>
<tr>
<td>M. canetti seq 2.a</td>
<td>Forward sequencing primer 2</td>
<td>TGTCGGCGCCACAGT</td>
</tr>
<tr>
<td>M. canetti seq 2.b</td>
<td>Reverse sequencing primer 2</td>
<td>ATCGTGCAGTGCGCCA</td>
</tr>
<tr>
<td>M. canetti seq 3.a</td>
<td>Forward sequencing primer 3</td>
<td>GCACGTATTGGTTGACGCA</td>
</tr>
<tr>
<td>M. canetti seq 3.b</td>
<td>Reverse sequencing primer 3</td>
<td>TCCACGCTTGCGCTT</td>
</tr>
<tr>
<td>M. canetti seq 4.a</td>
<td>Forward sequencing primer 4</td>
<td>TGATGGCGCTGCTCAAGC</td>
</tr>
<tr>
<td>M. canetti seq 4.b</td>
<td>Reverse sequencing primer 4</td>
<td>TGTCACAGGGCATGCGGGAAC</td>
</tr>
</tbody>
</table>
Table 2: Result interpretation table

<table>
<thead>
<tr>
<th>Diagnostics result profile</th>
<th>FAM (WbbI assay)</th>
<th>HEX (MTC assay)</th>
<th>ROX (M. canettii RD assay)</th>
<th>Cy5 (IAC*)</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ ve + ve + ve +/-ve</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M. canettii present.</td>
</tr>
<tr>
<td>+ ve + ve -ve +/-ve</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M. tuberculosis present.</td>
</tr>
<tr>
<td>-ve + ve -ve +/-ve</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MTC present (other than M. canettii or M. tuberculosis).</td>
</tr>
<tr>
<td>-ve -ve -ve + ve</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Not a member of the MTC.</td>
</tr>
<tr>
<td>-ve -ve -ve -ve</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Result invalid and test must be repeated.</td>
</tr>
</tbody>
</table>

* The IAC is only detected in the presence of low concentrations of primary target or absence of primary target.
**Figure Legend**

Figure 1 A. Real-time amplification curves for *M. tuberculosis* (circle) and *M. canettii* (triangle) using *wbbL1* gene in FAM channel (438-533). Figure 1 B Amplification curves for all MTC using *lepA* gene in HEX channel (523-568), with the non-template control highlighted with circles through line. Figure 1 C Amplification curves for *M. canettii* specific assay in ROX channel (558-610), with the 5 *M. canettii* strains depicted with triangles. Figure 1 D Amplification curves for IAC in Cy5 channel (615-670) with *M. africanum 3* highlighted with boxes through amplification curve. The Reduction of the IAC signal is due to competition from high concentrations of primary target, resulting in competition between the MTC and IAC assays.