Simultaneous mycobacterial speciation and determination of tuberculosis drug resistance by PCR/ESI-MS.

Christian Massire*, Cristina Agasino Ivy, Robert Lovari, Natalia Kurepina, Haijing Li, Lawrence B. Blyn, Steven A. Hofstadler, George Khechinashvili, Charles W. Stratton, Rangarajan Sampath, Yi-Wei Tang, David J. Ecker and Barry N. Kreiswirth

1 Ibis Biosciences, Inc., a subsidiary of Abbott Molecular, Inc., Carlsbad, CA.
2 Public Health Research Institute Tuberculosis Center, University of Medicine and Dentistry of New Jersey, Newark, NJ.
3 Department of Pathology, Vanderbilt University Medical Center, Nashville, TN.
4 Department of Medicine, Vanderbilt University Medical Center, Nashville, TN.
5 Georgian Foundation against Tuberculosis and Lung Diseases, Tbilisi, Georgia.

*Corresponding author.

Mailing address: Ibis Biosciences, Inc., 2251 Faraday Ave., Carlsbad, CA 92008.

E-mail: cmassire@ibisbio.com. Phone: (760) 476-3200. Fax: (760) 476-3301.

Running title: Mycobacterium profiling by PCR/ESI-MS
Abstract

*Mycobacterium tuberculosis* that is resistant to both isoniazid (INH) and rifampin (RIF) is spreading. It has become a public health problem in part because the standard culture methods used to determine the appropriate treatment regimen for patients often take months following the presumptive diagnosis of tuberculosis. Furthermore, the misidentification of non-tuberculosis mycobacteria (NTM) in patients presumably suffering from tuberculosis results in additional human and healthcare costs. The mechanisms of resistance for several drugs used to treat *Mycobacterium tuberculosis* are well understood and therefore should be amenable to determination by rapid molecular methods. We describe here the use of PCR followed by electrospray ionization mass spectrometry (PCR/ESI-MS) in an assay that simultaneously determines INH and RIF resistance in *Mycobacterium tuberculosis* and identifies and speciates NTMs. The assay panel includes 16 primer pairs in 8 multiplexed reactions and was validated using a collection of 1340 DNA samples from cultured specimens collected in the New York City area, the republic of Georgia and South Africa. Compared with phenotypic data, the PCR/ESI-MS assay had 89.3% sensitivity and 95.8% specificity in the determination of INH resistance, and 96.3% sensitivity and 98.6% specificity in the determination of RIF resistance. Using a set of 264 previously characterized liquid culture specimens, the PCR/ESI-MS method had 97.0% sensitivity and 99.9% specificity for determination of NTM identity. The assay also provides information on ethambutol, fluoroquinolone and diarylquinoline resistance and lineage-specific polymorphisms to yield highly discriminative digital signatures potentially suitable for epidemiology tracking.
The spread of multi-drug resistant (MDR) *Mycobacterium tuberculosis* (MTB) resistant to both isoniazid (INH) and rifampin (RIF) is an increasingly worrisome health concern. Resistance testing requires 3-4 weeks in the best of laboratories using the state-of-the-art culture systems, due to the extremely slow growth rate of MTB. At the same time, the spread of non-tuberculosis mycobacteria (NTM) infections is also emerging as an issue that must be addressed (21, 39). The proper characterization of NTMs is critical as some species are known to be naturally resistant to one or more antitubercular agents and, under current testing paradigms, the presence of an NTM may be suspected only after the failure of a regular tuberculosis treatment (11). Molecular methods have been reported that directly identify and differentiate mycobacterial isolates from acid-fast bacillus culture broth (23, 34, 38, 40, 41). Li *et al.* reported the use of broad-range PCR amplification followed by suspension array analysis to identify and differentiate 17 commonly encountered mycobacterial species, in the clinical setting (25). However, the detection of resistance mutations and NTM characterization remain distinct tasks requiring multiple assays (24, 30).

In the present study, we evaluated the use of an assay that employs PCR followed by amplicon characterization using electrospray ionization mass spectroscopy (PCR/ESI-MS) for the speciation and drug-resistance characterization of mycobacteria, as illustrated in Figure 1. PCR/ESI-MS is a rapid, high-throughput method for identification, characterization and quantification of microorganisms (bacteria, fungi, and viruses) present in cultured specimens or patient samples (6, 7, 9). Several factors converge to make PCR/ESI-MS an ideal tool for MTB genotyping, NTM characterization, and characterization of MDR. First, a PCR/ESI-MS assay
provides extensive flexibility in the choice of custom-made primer pairs that specifically
interrogate the desired markers (i.e., the determination of RIF resistance or the characterization
of a specific lineage). Second, the process is automated and allows the batch analysis of up to
240 isolates per day, with the first isolate being analyzed under four hours (9). Third, the digital
nature of PCR/ESI-MS results allows portability between different facilities. And finally, the
type of information that needs to be queried for MTB resistance profiling fits perfectly within the
specifications of the methodology. In general, the main limitations of PCR/ESI-MS technology
are found in the usable amplicon size (preferably less than 160 nucleotides long) and in the
limited information content when compared to traditional sequencing, as the linking order of the
bases is not determined using PCR/ESI-MS. Neither limitation is significant in the case of MTB
profiling as the targeted mutations are sparse and can be accurately identified using base
composition alone (8). The genes carrying mutations that confer drug resistance in MTB lack
secondary (incidental) synonymous mutations, an observation that has been consistently verified
by numerous genotyping studies (27). The virtual absence of background genomic variation
ensures that any observed deviation from the expected mass represents an expressed mutation
that is significant for the determination of drug resistance. These characteristics allowed us to
assemble a PCR/ESI-MS assay using 16 primer pairs in eight multiplex reactions that target the
main markers associated with RIF and INH resistance and allow speciation of NTMs. The assay
was tested using a large collection of 1340 isolates with various resistance profiles and a set of
264 previously characterized NTMs. This study was performed using the Ibis T5000 biosensor,
which allowed the simultaneous analysis of 12 samples on a 96-well microtiter plate. Sample
extraction, PCR amplification, and T5000 analysis were typically performed by batches of up to
fifteen plates. Time to first answer was typically in the order of 6 to 8 hours, with additional results becoming available by increments of one hour per plate.

Materials and Methods

Isolates tested

Four sets of DNA samples isolated from clinical MTB isolates were used in this study. First, 45 reference isolates were collected from the Public Health Research Institute (PHRI) for the initial primer pair testing and panel assembly. These isolates had been extensively characterized by IS6110 restriction fragment length polymorphism (RFLP), spoligotyping, mycobacterial interspersed repetitive unit (MIRU) typing and principal genetic group analysis (27). Susceptibilities to INH, RIF and EMB were also previously determined using direct culture-based growth inhibition methods (29), and the presence of the corresponding mutations within the katG, inhA, rpoB and gyrA genes was confirmed by direct sequencing.

Second, a total of 1340 MTB DNA samples was gathered to evaluate the drug resistance determination capabilities of the PCR/ESI-MS assay. This evaluation set was divided into three subsets: The first subset of 962 isolates originated primarily from the greater New York area, with 911 isolates provided by an ongoing collaboration between the PHRI TB Center, the New York City Bureau of Laboratories and the New Jersey Department of Health and Senior Services (4). An additional 51 isolates were obtained from the Russian Federation and were characterized at the PHRI. The second subset of 188 MTB isolates was collected in the Republic of Georgia (10) and the third subset included 190 isolates from South Africa. Statistical analyses of these three subsets were performed separately since the drug susceptibility testing was performed at different locations.
Third, a panel of 47 mycobacterial reference strains was collected from the ATCC to provide reference NTM signatures. This set represents 33 species: *M. agri*, *M. asiaticum*, *M. aurum*, *M. avium*, *M. branderi*, *M. chelonae*, *M. engbaekii*, *M. farcinogenes*, *M. flavescens*, *M. fortuitum*, *M. gallinarum*, *M. gastri*, *M. gordoniae*, *M. haemophilum*, *M. intracellulare*, *M. kansasii*, *M. lactis*, *M. lufu*, *M. malmoense*, *M. marinum*, *M. neoaurum*, *M. phlei*, *M. rhodesiae*, *M. scrofulaceum*, *M. senegalense*, *M. sherrisii*, *M. simiae*, *M. szulgai*, *M. terrae*, *M. tokaiense*, *M. ulcerans*, *M. vaccae* and *M. xenopi*.

Finally, to test the speciation capabilities of the assay, a total of 264 culture broth specimens that had tested positive for acid-fast bacilli using the BACTEC MGIT 960 (BD Diagnostic Systems, Franklin Lakes, NJ) were collected at the Vanderbilt University Medical Center and the Veteran Affairs Tennessee Valley Health Care system (25).

**Sample culture and drug susceptibility testing**

The strains are sub-cultured on Löwenstein-Jensen slants (29). Drug susceptibility testing (DST) has performed in New York City Department of Health Laboratory as described (29) with the indirect proportion method with Middlebrook 7H10 medium for the following anti-TB drugs (at concentrations): isoniazid (1.0 µg/mL), rifampin (1.0 µg/mL), streptomycin (10.0 µg/mL), ethambutol (5.0 µg/mL).

**Genome preparation and PCR**

Genomic material from cultured samples was prepared using the DNeasy Tissue Kit (Qiagen, Valencia, CA) according to the manufacturer’s protocols. All PCR reactions were assembled in 40 µL reactions in the 96-well microtiter plate format using a Packard MPII liquid handling robotic platform and an Eppendorf Mastercycler pro (Eppendorf, Hauppauge, NY). The primers
The PCR reaction mix consisted of 1 unit of Immolase (Bioline, Taunton, MA), 1x buffer II, 1.5 mM MgCl₂, 0.4 M betaine, 800 µM dNTP mix, and 250 nM of each primer. The following PCR conditions were used to amplify the sequences used for PCR/ESI-MS analysis: 95°C for 10 minutes followed by 8 cycles of 95°C for 30 seconds, 48°C for 30 seconds, and 72°C for 30 seconds, with the 48°C annealing temperature increasing 0.9°C each cycle. The PCR was then continued for 37 additional cycles of 95°C for 15 seconds, 56°C for 20 seconds, and 72°C for 20 seconds.

**PCR/ESI-MS gene targets and selection of primers**

General methods for PCR/ESI-MS have been described previously using the first PCR/ESI-MS instrument, the Ibis T5000 biosensor (6, 12), and remain fully applicable with the currently available instrument marketed as Plex-ID (9). The choice of the genes used for PCR/ESI-MS analysis was primarily dictated by the location of the main mutations associated with resistance to INH and RIF (e.g., katG codon 315) and to rifampin (e.g., rpoB codons 516, 526 and 531). Whenever possible, primer pairs were designed to exclusively amplify the codons of interest. Other primer pairs were designed to amplify a longer portion of a gene in regions where rarer mutations are known to occur at different locations to allow query of multiple sites simultaneously. An example is primer pair BCT4366, designed to yield amplicons that encompass rpoB codons 505 to 516 of the rifampin-resistance determining region. This primer pair not only detects point mutations in individual codons (e.g., L511P or D516G) but also more complex insertion or deletion patterns of one or more codons (18). Such mutations are often not surveyed by alternate molecular typing methods focusing on point mutations. An internal positive control, containing a target for one of the two primer pairs per well, was designed, tested...
and added at 150 molecules per well of the assay plates. This calibrant provided a distinct base composition signature as a positive control for the PCR.

**Mass spectrometry and base composition analysis**

Following amplification, 15 µL aliquots of each PCR reaction were desalted and purified using a weak anion exchange protocol as described elsewhere (20). Accurate mass (±1 ppm), high-resolution (M/dM > 100, 000 FWHM) mass spectra were acquired for each sample using the PCR/ESI-MS protocols described previously (19). For each sample, approximately 1.5 µL of analyte solution was consumed during the 74-second spectral acquisition. Raw mass spectra were post-calibrated with an internal mass standard and deconvolved to monoisotopic molecular masses. Unambiguous base compositions were derived from the exact mass measurements of the complementary single-stranded oligonucleotides. Quantitative results were obtained by comparing the peak heights with the internal PCR calibration standard (present in every PCR well at 150 molecules).
Results

Assay panel

We assembled a 16-primer pair multiplexed assay panel; nucleic acid isolated from each specimen was amplified in eight reactions each containing two primer pairs. This allowed the analysis of 12 samples on the same 96-well plate. We tested over 80 primer pairs with dilution-to-extinction experiments using purified DNA from the reference strain H37Rv and from strains selected to represent the variety of mutations expected within the surveyed loci. Information on the 16 primers pairs selected for final configuration is reported in Table 1. The assay primarily focuses on the determination of resistance to INH and RIF (five and six primer pairs, respectively). The panel was completed by the incorporation of primer pairs that target mutations associated with ethambutol, fluoroquinolone and diarylquinoline resistance; all are chiefly associated with mutations in single hotspots (27, 37). In addition, two primer pairs were included for the determination of the Principal Genetic Group (PGG) classification (36) and provide a broad phylogenetic classification.

Linearity and dynamic range

All the results described in this work are from examination of DNA extracted from isolated colonies or broths from MGIT culture. However, the assay has potential for use where quantitative measurements are important. Therefore, we measured the linearity and dynamic range of the PCR/ESI-MS assay in serial limiting dilution experiments (Figure 2). PCR/ESI-MS measurements are calibrated by spiking, during the manufacturing process, a precisely known amount of synthetic DNA into each PCR reaction (19). This synthetic DNA has the same
sequence as the target organism with the exception of a small deletion in the amplicon region to distinguish it from amplicons generated from the target. In this assay the calibration standard was present at 150 molecules in each PCR reaction. Target amplicon was quantitated by a comparison of the mass spectral peak heights from amplicons generated from the target organism and calibration standard. As shown in Figure 2, the assay was linear and has a dynamic quantitative response over a range of at least 300-fold.

**Determination of reference base composition signatures for MTB and NTM species**

We then established reference base composition signatures for a large panel of mycobacterial species. A collection of 92 isolates were tested in the PCR/ESI-MS assay: 23 isolates from diverse lineages of *M. tuberculosis*, 22 isolates representing other species within the MTB complex (*M. canetti*, *M. africanum*, *M. bovis* and *M. microti*) and 47 isolates representing 33 distinct NTM species. Typical signatures associated with a number of clinically relevant mycobacterial species are reported in Table 2.

Base composition signatures for the pan-susceptible MTB isolates revealed their PGG classification (Table 2, top three rows). Signatures for the other species within the MTB complex were indistinguishable from the one for MTB (PGG1), except for the four *M. africanum* strains which had G to T mutations within the *atpE* locus (Table 2). This mutation (Rv1305_0207s) is one of the eleven synonymous (silent) single nucleotide polymorphisms (sSNP) that characterize the *M. africanum* type 1a, a.k.a. MTB lineage 6 (5, 17). In contrast, all NTM isolates outside of the MTB complex yielded distinct signatures, though only nine out of the sixteen primer pairs included in the assay actually contribute to speciation: The two primer pairs that target the *inhA* and *ahpC* promoter sequences do not yield amplicons outside of the MTB complex, whereas the
five primer pairs that result in amplification of only one or two codon regions yielded the same signature for NTMs and MTB. The nine primer pairs that target stretches of 7 to 40 nucleotides segregate Mycobacteria from each other (Table 2). Six of these loci provide the bulk of the resolving power: BCT4236 (*inhA* 91-96), BCT4237 (*rpoB* 142-154), BCT4366 (*rpoB* 505-516), BCT3697 (*rpoB* 562-572), BCT3555 (*gyrA* 90-96) and BCT4364 (*atpE* 58-69). The last three loci, BCT4235 (*inhA* 21), BCT3551 (*embB* 306) and BCT3828 (*rpoB* 531-538) mainly provide signatures specific for *M. kansasii, M. simiae, M. szulgai* and *M. gordonae*. All species tested were clearly differentiated from *M. tuberculosis* in at least five loci. Furthermore, each species tested so far displayed unique base composition signatures in at least two distinct loci. Thus, the PCR/ESI-MS assay has potential for the characterization of NTM species.

Characterization of mycobacterial species from liquid culture broth

Once base composition signatures for these reference samples were determined, we analyzed 264 DNA samples that had been characterized previously using the BACTEC MGIT 960 culture system (25). The correlation between the PCR/ESI-MS identification and the reference identification is shown in Table 3. Among the 50 MTB positive specimens, 49 showed typical 16-loci MTB complex signatures when analyzed by PCR/ESI-MS and could be further segregated into PGG1, PGG2 and PGG3 (11, 31 and 7 isolates, respectively) based on our data. Only one resistance mutation was detected, an A to T change in BCT3908 indicating an *rpoB* H526L mutant (CAC → CTC). One sample yielded a highly unusual signature with only one locus compatible with MTB; as the signature did not match those of any of the NTM isolates tested, the identity of this isolate remains unclear.
One hundred fifty-five of the specimens were previously characterized as belonging to the *M. avium/M. intracellulare* (MAI) complex (25). Most of these were further characterized as either *M. avium* or *M. intracellulare* by PCR/ESI-MS (74 and 76 isolates, respectively). The species assignment was based on the retrieval of at least five out of the six MAI-specific base counts (Table 2). Seven of these isolates showed unexpected mutations in one locus, resulting in the identification of two new *M. avium* and five *M. intracellulare* signature variants in addition to the consensus signatures already established for these species. Further studies will be needed to determine if these mutations correlate with variations in drug susceptibility. In two of the 155 MAI isolates, both *M. avium* and *M. intracellulare* base counts were detected in the three loci where these species differ, revealing the simultaneous occurrence of both species within the same sample. Conversely, three isolates yielded partial matches to both species, allowing a partial identification to the MAI complex but not a definitive species assignment – these may represent evolutionarily intermediate strains, recombinants, or strains in which homoplasy has generated identical mutations in multiple lineages.

For the remaining 59 samples, reference and PCR/ESI-MS characterizations were overall in good agreement. However, the species of six isolates (Table 3, bottom row) could not be determined by PCR/ESI-MS as the base composition signatures for these isolates contained multiple novel base counts and/or inconsistent or partial matches to base composition signatures of known species. The diagnostic performance of the PCR/ESI-MS assay for *Mycobacterium* speciation based on these 268 AFB-positive MGIT broth culture specimens is reported in Table 4.
Determination of drug resistance profiles

A total of 1340 DNA samples from MTB isolates were tested to evaluate the utility of the PCR/ESI-MS assay in determination of drug resistance. Phenotypic testing was performed previously for each of these isolates by the absolute-concentration method on solid agar medium, allowing a direct evaluation of the reliability of the PCR/ESI-MS assay. Four isolates were characterized as M. avium by PCR/ESI-MS, whereas strain mixtures were identified in nine isolates through the recognition of multiple base composition signatures in at least two loci. The PCR/ESI-MS characterization of each subset of isolates is shown in Figure 3.

Mutations conferring resistance to INH were found in 745 of the 835 isolates previously shown to be INH resistant for an overall sensitivity of 89.2% (Table 5). Interestingly, the INH mutations in regions targeted by our primer pairs were more often observed in the isolates with the MDR phenotype (339/356 or 95.2%) than in those with the mono-INH-resistant isolates (406/479 or 84.8%). As expected, the most common mutations for INH resistance were found in the two main loci, katG S315T (ACC codon: 492 isolates or 58.9%) and inhA promoter C-15T (151 isolates or 18.1%). The signature katG S315N/T was also found in 89 isolates (10.7%); it corresponds to either the single mutant S315N (AAC) or the double mutant S315T (ACA) found in the W MDR strain. Both mutations have the same mass signature and are therefore indistinguishable by PCR/ESI-MS. Four instances of rarer katG mutations (S315T, S315R or S315I) and 29 other mutations within the inhA operon were also detected. Mutations were simultaneously found in both katG 315 and inhA promoter in 32 of the 835 INH-resistant isolates (3.8%). Three secondary loci were also evaluated for INH resistance in the PCR/ESI-MS assay. Within the ahpC promoter four distinct types of mutations were observed, but in only 16 of 37 instances were they seen in the absence of mutations within the inhA promoter or katG.
Similarly, mutations within the \textit{inhA} reading frame were observed alone in only eight of 20 instances (five I21V, six I21T and nine I94T). The inclusion of the \textit{ahpC} promoter locus and the two \textit{inhA} loci therefore allow modest gains of 2\% and 1\%, respectively, in the overall level of INH susceptibility. Finally, mutations were found in 21 isolates with INH-susceptible phenotypes, including seven instances of \textit{katG} S315T and nine instances of \textit{inhA} C-15T. These findings highlight the still incomplete understanding of the molecular mechanisms underlying INH resistance.

For RIF, the correlation between phenotypes and PCR/ESI-MS profiles was higher than for INH with only 14 missed detections among 393 RIF-resistant isolates (96.4\% sensitivity) and 13 unexpected mutations among 933 RIF-susceptible isolates (98.5\% specificity). The data are summarized in Table 6. Most mutations were found within the so-called Rifampin-Resistance Determining Region or RRDR, particularly in codons 531 (207 instances, mainly S513L), 526 (115 instances) and 516 (58 instances). These findings are in agreement with the commonly held view that RIF resistance mutations are confined within the \textit{rpoB} gene (3). Mutations conferring resistance to rifampin have also been characterized in secondary \textit{rpoB} hotspots (13, 15, 26). We found ten RIF-resistant isolates with the mutation V146F and five isolates possessed the I572F mutation. These two mutations outside of the RRDR accounted for 3.8\% of the instances of RIF resistance. Neither mutation was present in the RIF-susceptible isolates. L511P, L533P or D516Y mutations were found in seven RIF-susceptible isolates; these mutations are usually associated with low-level RIF resistance (13). The mutations H526Y, H526N or S531L, commonly associated with RIF resistance, were observed in six isolates with a RIF-susceptible phenotype.
The PCR/ESI-MS assay also includes three markers for ethambutol, fluoroquinolone and diarylquinoline resistance (2, 27, 31, 37). Phenotypes for EMB resistance were available for 755 isolates (Table 7). Mutations at \(embB\) codon 306 were observed in 92 of the 125 EMB-resistant isolates (73.6% sensitivity) and in 17 of 630 EMB-susceptible isolates (97.3% specificity). The mutations identified were M306V (69 isolates), M306L (4 isolates) and the three M306I variants (21, 13 and 2 isolates with ATA, ATC and ATT codons, respectively). The M306I mutants were noticeably overrepresented in the isolates with the EMB-susceptible phenotype (9/17, 53%) compared to the ones with EMB-resistant phenotypes (27/92, 29%). No phenotypes were available for fluoroquinolone resistance, but at least eight distinct mutations were detected within the quinolone-resistance determining region of 49 isolates, including five isolates that were both INH- and RIF-susceptible. Finally, analysis of amplicons from the primer pair targeting the \(atpE\) locus indicated diarylquinoline resistance in two isolates and allowed the identification of the \(M. africanum\) lineage in seven isolates.

**Discussion**

In the present study we investigate the use of the PCR/ESI-MS technology for the molecular characterization of \(M. tuberculosis\) and other mycobacteria species. The PCR/ESI-MS mycobacterium assay provides, on the time frame of a few hours and with a throughput of 300 specimens per day, the rapid speciation of mycobacteria present in the sample (MTB vs. NTM, the latter being identified at the species level), the characterization of drug resistance to the primary front line drugs (INH and RIF) if MTB is present, the characterization of ethambutol and fluoroquinolone resistance, and the partial characterization of lineages within the \(M. tuberculosis\) complex itself. A unique feature of this assay is the simultaneous targeting of specific mutations.
conferring drug resistance and lineage markers using complementary priming strategies. In addition to the traditional MDR markers that were captured individually, primer pairs targeting the \textit{atpE}, the \textit{rpoB} and the \textit{inhA} loci were designed to capture rare resistance mutations that may occur in MDR MTB. In addition, these primer pairs targeting housekeeping genes cover stretches of DNA long enough to capture species-specific variations, thus rendering the characterization of NTM feasible with the same assay used for MTB drug resistance testing. A prime example is the \textit{atpE} primer pair that amplifies a region associated with mutations that provide resistance to diarylquinoline (2, 31) but is also useful for NTM characterization.

Although there was excellent correlation between PCR/ESI-MS data and previously determined rifampin resistance of the analyzed isolates (96.4% sensitivity and 98.6% specificity overall), the PCR/ESI-MS assay did not perform as well for the determination of INH resistance (89.2% sensitivity and 95.8% specificity). The sensitivity and specificity of the PCR/ESI-MS assay, however, compare favorably with other molecular methods for resistance determination (28). In our analysis of the PCR/ESI-MS assay, we characterized 1340 isolates collected on three continents to maximize strain diversity. We also included an unusually large contingent of isolates with INH-only resistant phenotypes (479 INH-only versus 356 MDR). The PCR/ESI-MS assay interrogates five regions commonly associated with INH resistance. These mutations were less often present in the isolates resistant only to INH (406/479, 84.8%) than in the MDR isolates (339/356, 95.2%). Demonstrating the diversity due to geography, in the Georgian isolates, these common INH-resistance mutations were found in 18 of 26 (69%) INH-only isolates versus 12 of 13 (92%) MDR isolates, hence the significantly lower sensitivity achieved for INH testing with the Georgian isolates (Table 5). The isolates where INH mutations were expected but not observed were evenly distributed between the three principal genetic groups with no apparent
sampling bias, and further investigation will be required to identify the molecular basis for INH resistance in these isolates. Other discrepancies, however, may rather find an explanation in the questionable portability of phenotyping methods (22). For example, we found *embB* M306I mutants primarily in samples with an EMB-susceptible phenotype. Previous studies showed that the marginal increase in EMB MIC levels resulting from the incorporation of a M306I only brings the EMB MIC close to the critical concentration and therefore could be inadvertently defined as EMB susceptible (33, 37).

Since the PCR/ESI-MS technology uses aggregate base composition signatures and not sequencing, potentially ambiguous results might occur if two mutations that result in no mass change (e.g., G→C and C→G) are simultaneously present within the same amplicon (8). This concern guided the primer pair design and panel assembly for the MTB assay, and a subsequent survey of the literature revealed that such a situation would arise in at least one instance encountered in past studies: *gyrA* D94H+T95S (GAC→CAC + ACC→AGC). This ambiguity was alleviated by the inclusion of a primer pair that unambiguously resolves codon *gyrA* 95 alone (BCT3556). In this study, four isolates harboured both mutations, and each occurrence was identified. Mutations at this position characterize the sub-lineage that includes strain H37Rv, and are the only known example of lineage-specific mutation appearing within a resistance-determining region (14).

The other lineage-specific mutation used for Principal Genetic Group (PGG) typing, *katG* L463R, was recently confirmed by multi-locus analysis to be congruent with the whole Euro-American lineage (5). Although the PGG classification does not distinguish older and less frequent lineages, it does provide valuable insights into the global phylogeny of sample collections and helps delineate clonal outbreaks. The characterization of these lineage-specific
mutations within the \textit{katG} and \textit{gyrA} loci provides a basic framework upon which mutation profiles can be mapped and interpreted (Figure 3). For example, a cluster of 36 PGG1 isolates from New York shared the same string of resistance mutations (\textit{katG} S315T(ACA), \textit{rpoB} H526Y, \textit{embB} M306V). This combination was seen only in the New York PGG1 isolates and, based upon IS6110 and spoligotyping analysis, indicated the W strain. Similarly, another cluster of six PGG3 South African isolates was defined by simultaneous mutations of \textit{katG} S315N(AAC), \textit{rpoB} H526Y and \textit{embB} M306V. Three distinct \textit{gyrA} mutations were also retrieved within this cluster, hinting that extensively drug-resistant strains may be common within this particular lineage. The inclusion of the \textit{atpE} marker in the PCR/ESI-MS assay allowed characterization of one of the two MTB lineages commonly known as \textit{M. africanum} (17). The future inclusion of additional markers would similarly provide unambiguous characterization of the remaining \textit{M. tuberculosis} lineages, providing valuable information for epidemiology tracking. To our knowledge, there is only one study that has previously attempted to speciate mycobacteria and assess drug resistance with the same assay (35). This reverse line blot hybridization assay, however, uses a panel of 40 specific probes to target the main mutations associated with RIF, INH and streptomycin resistance, and a parallel set of sixteen probes for NTM characterization. This assay achieved sensitivity levels comparable with our own; the set of isolates evaluated appears to be less diverse than those evaluated here, as there was an unusually high prevalence of \textit{rpoB} S531L and \textit{katG} S315T mutations detected in the Shenai \textit{et al.} study.

Most of the initial testing and validation work performed in this study involved isolates cultured on Löwenstein-Jensen medium obtained through existing collaborations. Current laboratory techniques for identification of mycobacteria to the species level involve growth in broth and then nucleic acid probing or subculture to solid media for phenotypic tests, which can
take weeks. Use of DNA probes (AccuProbe; GenProbe, San Diego, CA), which are specific for 
*M. tuberculosis* complex (MTB), *M. avium-intracellulare* complex (MAI), *M. kansasii*, and *M. 
gordonae*, on samples taken directly from BACTEC TB broth culture has dramatically reduced 
the time to identification and differentiation of mycobacterial infections (1, 32). Recent work 
(16) has demonstrated that the characterization of resistance mutations directly from sputum 
samples is achievable, and validation of the PCR/ESI-MS assay using direct sputum specimens is 
planned.

The present study was primarily performed using the Ibis T5000 biosensor, a prototype of the 
next generation of PCR/ESI-MS instrument commercialized by Abbott under the Plex-ID brand 
(9). Since the completion of this study, the MTB assay has been further validated for use with the 
Plex-ID instrument. While the core of the technology and the assays remains the same, the new 
instrument platform allows significant cost cuts and a faster turnover with an average analysis 
time of 30s per well, half of what was needed with the T5000.

Finally, it should be reminded that, in order to be used as a clinical diagnostic tool, the 
technology must be FDA-approved in the U.S. and CE-Marked in Europe. That process has been 
initiated but approval is still pending. Consequently, the assay is for now available for Research 
Use Only.

**Acknowledgements**

The authors thank Dr. Dorothy Fallows and Dr. Gilla Kaplan from the PHRI Center for 
providing the South African isolates, Dr. Kristin Sannes-Lowery, Jenna Cromwell and Maria 
Tobar-Mosquera for their help during sample analysis. We also thank Dr Jacqueline Wyatt and 
Dr. David Metzgar for editorial assistance. This study was supported in part by grants # R44
AI078694-03 from the Center for Disease Control and GEB2-2605-TB-04 from the U.S. Civilian Research and Development Foundation.
References


Development of a simple and low-cost real-time PCR method for the identification of
107:1433–1439.


Figure 1. Basic flow chart of the PCR/ESI-MS assay

Resistance profiling:
INH: katG 5315T/G/N/R/...
    inhA promoter C-15T/T-8A/...
    ahpC promoter G->A/...
    inhA 121 wt/T/V/...
RIF:  rpoB 5531L/Y/W/...
    rpoB H526D/G/N/R/Y/...
    rpoB D516Y/F/H/N/...
    rpoB P564 R/A/S/...
    rpoB V146F/...
EMB: embB M306V/I/...
FQ:  gyrA D94A/G/N/V, A90

Principal Genetic Group:
katG L463wt / L463R
    gyrA 595wt / 595T

Speciation by name:
M. avium
M. intracellulare
M. kansasii
M. simiae
M. szulgai
M. gordonae
M. chelonae/M. abscessus
M. fortuitum
Figure 2. Linearity and dynamic range of the PCR/ESI-MS assay. Synthetic controls including the intended target sequence for each primer pair were designed and known quantities were analyzed by PCR/ESI-MS (x-axis). The measured molecular count detected by ESI-MS is reported on the y-axis for each of four primer pairs using the same logarithmic scale as was used for the x-axis. Each data point represents the average detection read for the same primer pair across five distinct plates with the minimal and maximal values represented by vertical bars. Data are not represented for lower input values as the amplification ceased to be reproducible below a theoretical input of 8 molecules per well. The average response of the 16 primer pairs is indicated by the brown solid line.
Figure 3. Characterization by PCR-ESI-MS of the 1340 isolates used for the validation of resistance profiling. Isolates were first segregated into three distinct doughnut charts according to their origin. Within each chart, isolates are then sorted and color-coded in accordance with their phylogenetic characterization (light, medium and dark green slices for PGG 1, 2 and 3, respectively; blue for *M. africanum*; grey for strain mixtures or undetermined lineages). Mutations for INH, RIF, EMB and FQ resistance are indicated by coloring the corresponding sectors, with the three most common mutations reported in yellow, light and dark orange (see color key below). All other mutations or multiple detections are indicated with red sectors. Brackets indicate two distinct clusters of multi-resistant strains sharing largely congruent PCR/ESI-MS signatures.
<table>
<thead>
<tr>
<th>Row</th>
<th>Locus</th>
<th>Target</th>
<th>Forward and Reverse primer sequence</th>
<th>Full amplicon length</th>
<th>Sequence length between primers</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>BCT3633*</td>
<td>*rpoB</td>
<td>TCCAGCCAGCTGAGCCAACTTC</td>
<td>47</td>
<td>6</td>
<td>RIF resistance; does not discriminate NTM from MTB</td>
</tr>
<tr>
<td>B</td>
<td>BCT4235</td>
<td>inhA 19, 23</td>
<td>TTGGTACGGGAACTACCCGGAAGCCATGCTGTTGAGCGGGTTGTTGCTG</td>
<td>60</td>
<td>17</td>
<td>INH resistance; speciates NTM</td>
</tr>
<tr>
<td>B</td>
<td>BCT3552*</td>
<td>inhA promoter</td>
<td>TGCTGGCCAGAGTACCGATTCGTATGCTG</td>
<td>75</td>
<td>32</td>
<td>INH resistance; does not prime NTM</td>
</tr>
<tr>
<td>B</td>
<td>BCT3908</td>
<td>*rpoB 526</td>
<td>TGGCCCTGCTGAGGTCGAC</td>
<td>40</td>
<td>3</td>
<td>RIF resistance; does not discriminate NTM from MTB</td>
</tr>
<tr>
<td>C</td>
<td>BCT4234*</td>
<td>ahpC promoter</td>
<td>TGGGATGCCGATAATATGGTGGTAGAACTACCCGGAAGCCATGCTG</td>
<td>101</td>
<td>49</td>
<td>INH resistance; does not prime NTM</td>
</tr>
<tr>
<td>C</td>
<td>BCT4364</td>
<td>atpE 58, 69</td>
<td>TCACCGTTCTTGATCGACCCGGAAGCCATGCTG</td>
<td>79</td>
<td>36</td>
<td>Diarylquinoline resistance; speciates NTM</td>
</tr>
<tr>
<td>D</td>
<td>BCT3551*</td>
<td>embB 306</td>
<td>TGACGGCTACACCTGCCCACGAT</td>
<td>43</td>
<td>7</td>
<td>Ethambutol resistance; speciates NTM</td>
</tr>
<tr>
<td>D</td>
<td>BCT4236</td>
<td>inhA 91, 96</td>
<td>TGGCCACATGCTGAGGGAAGCCATGCTG</td>
<td>55</td>
<td>20</td>
<td>INH resistance; speciates NTM</td>
</tr>
<tr>
<td>E</td>
<td>BCT3553*</td>
<td>katG 315</td>
<td>TGGTAGGGACGGGATCACCACCTGTGATCACCCGGAAGCCATGCTG</td>
<td>44</td>
<td>3</td>
<td>INH resistance; does not discriminate NTM from MTB</td>
</tr>
<tr>
<td>E</td>
<td>BCT3463</td>
<td>*rpoB 505, 516</td>
<td>TGCCAGGCTACACCTGCCCACGAT</td>
<td>72</td>
<td>32</td>
<td>RIF resistance; speciates NTM</td>
</tr>
<tr>
<td>F</td>
<td>BCT3554*</td>
<td>katG 463</td>
<td>TGCCAGCCTTACAGGACCCACGAT</td>
<td>48</td>
<td>3</td>
<td>PGGs determination; does not discriminate NTM from MTB</td>
</tr>
<tr>
<td>F</td>
<td>BCT3828</td>
<td>*rpoB 631, 539</td>
<td>TGACGGCTACACCTGCCCACGAT</td>
<td>66</td>
<td>28</td>
<td>RIF resistance; speciates NTM</td>
</tr>
<tr>
<td>G</td>
<td>BCT3555*</td>
<td>gyrA 90, 95</td>
<td>TGGCAGGCTACACCTGCCCACGAT</td>
<td>51</td>
<td>18</td>
<td>PGG determination; speciates NTM</td>
</tr>
<tr>
<td>G</td>
<td>BCT4237</td>
<td>*rpoB 142, 154</td>
<td>TGCCAGGCTACACCTGCCCACGAT</td>
<td>86</td>
<td>40</td>
<td>RIF resistance; speciates NTM</td>
</tr>
<tr>
<td>H</td>
<td>BCT3556*</td>
<td>gyrA 95</td>
<td>TGCCAGGCTACACCTGCCCACGAT</td>
<td>40</td>
<td>3</td>
<td>Fluoroquinolone resistance; does not discriminate NTM from MTB</td>
</tr>
<tr>
<td>H</td>
<td>BCT3697</td>
<td>*rpoB 562, 572</td>
<td>TGCCAGGCTACACCTGCCCACGAT</td>
<td>72</td>
<td>35</td>
<td>RIF resistance; speciates NTM</td>
</tr>
</tbody>
</table>

* Indicates the primer pair for which an internal positive control was designed and included in the assay.
<table>
<thead>
<tr>
<th>Species</th>
<th>BCT4235</th>
<th>BCT4236</th>
<th>BCT3551</th>
<th>BCT3554</th>
<th>PCR/ESI-MS locus and target gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTB (PGG 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A13G15C18T14</td>
</tr>
<tr>
<td>MTB (PGG 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A13G15C18T14</td>
</tr>
<tr>
<td>M. tuberculosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A13G15C18T14</td>
</tr>
<tr>
<td>M. bovis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A13G15C18T14</td>
</tr>
<tr>
<td>M. canetti</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A13G15C18T14</td>
</tr>
<tr>
<td>M. gordonae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A11G20C13T11 A8G14C15T6 A16G28C24T16</td>
</tr>
<tr>
<td>M. chelonae / M. abscessus</td>
<td>A12G17C15T12</td>
<td></td>
<td></td>
<td></td>
<td>A15G25C26T20 A17G20C22T13 A12G24C24T12</td>
</tr>
</tbody>
</table>

Table 2: Expected PCR/ESI-MS signatures for members of the MTB complex and the most common NTMs tested in this study.

M. fortuitum, Mycobacterium tuberculosis; PGG, Principal Genetic Group. Only ten loci (out of sixteen) are represented here; the six loci not shown do not provide lineage-specific signatures.
### Table 3: Correlation between the PCR/ESI-MS and reference identification for the collection of 268 NTM samples.

<table>
<thead>
<tr>
<th>PCR/ESI-MS Identification</th>
<th>Reference Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTB complex (3 PGG variants)</td>
<td>MTB</td>
</tr>
<tr>
<td>MTB complex (3 PGG variants)</td>
<td>49</td>
</tr>
<tr>
<td>MAI complex (3 variants)</td>
<td>76</td>
</tr>
<tr>
<td>M. avium &amp; M. intracellulare (6 variants)</td>
<td>2</td>
</tr>
<tr>
<td>MAI complex sp.</td>
<td>3</td>
</tr>
<tr>
<td>M. chelonae/M. abscessus complex</td>
<td>14</td>
</tr>
<tr>
<td>M. fortuitum complex (3 variants)</td>
<td>7</td>
</tr>
<tr>
<td>M. gordonae (3 variants)</td>
<td>9</td>
</tr>
<tr>
<td>M. kansasii</td>
<td>16</td>
</tr>
<tr>
<td>M. simiae</td>
<td>1</td>
</tr>
<tr>
<td>M. szulgai (2 variants)</td>
<td>3</td>
</tr>
<tr>
<td>M. intracellulare &amp; M. gordonae</td>
<td>1</td>
</tr>
<tr>
<td>unique signatures</td>
<td>1</td>
</tr>
<tr>
<td>All isolates</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 4: Sensitivity, specificity and predictive values of the PCR/ESI-MS mycobacterium assay for identification and speciation of 264 mycobacterial isolates directly from AFB-positive MGIT broth culture media.

<table>
<thead>
<tr>
<th>Mycobacterial species, complex or group by PCR/ESI-MS</th>
<th>Number of isolates with PCR/ESI-MS result</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTB complex</td>
<td>Correct positive ID Incorrect positive ID Incorrect negative ID Correct negative ID</td>
<td>98.0% 100.0% 100.0% 99.5%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. avium /M. intracellulare</td>
<td>157 0 1 106 100.0% 99.1% 99.4% 100.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. chelonae /M. abscessus group</td>
<td>15 2 0 249 88.2% 100.0% 100.0% 99.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. fortuitum group</td>
<td>7 0 0 259 100.0% 100.0% 100.0% 100.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. gordonae</td>
<td>10 2 0 256 83.3% 100.0% 100.0% 99.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. kansasii</td>
<td>16 3 0 249 84.2% 100.0% 100.0% 98.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. simiae</td>
<td>1 0 0 267 100.0% 100.0% 100.0% 100.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. szulgai</td>
<td>3 0 0 265 100.0% 100.0% 100.0% 100.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>258 8 1 1869 97.0% 99.9% 99.6% 99.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5: Determination of INH resistance (1340 isolates).

<table>
<thead>
<tr>
<th>INH</th>
<th>Number of isolates with result (reference / PCR/ESI-MS)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>R/R</td>
<td>R/S</td>
<td>S/R</td>
<td>S/S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York (962 isolates)</td>
<td>598</td>
<td>67</td>
<td>16</td>
<td>281</td>
<td>89.9%</td>
</tr>
<tr>
<td>Georgia (188 isolates)</td>
<td>30</td>
<td>9</td>
<td>5</td>
<td>144</td>
<td>76.9%</td>
</tr>
<tr>
<td>South Africa (190 isolates)</td>
<td>117</td>
<td>14</td>
<td>0</td>
<td>59</td>
<td>89.3%</td>
</tr>
<tr>
<td>All 1340 isolates</td>
<td>745</td>
<td>90</td>
<td>21</td>
<td>484</td>
<td>89.2%</td>
</tr>
</tbody>
</table>

R, resistant; S, susceptible; PPV, Positive Predictive Value; NPV, Negative Predictive Value.
Table 6: Determination of RIF resistance (1340 isolates).

<table>
<thead>
<tr>
<th>RIF</th>
<th>Number of isolates with result (reference / PCR/ESI-MS)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R/R R/S S/R S/S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York (962 isolates)</td>
<td>238 9 8 707</td>
<td>96.4%</td>
<td>98.9%</td>
<td>96.7%</td>
<td>98.7%</td>
</tr>
<tr>
<td>Georgia (188 isolates)</td>
<td>16 0 4 168</td>
<td>100%</td>
<td>97.7%</td>
<td>80.0%</td>
<td>100%</td>
</tr>
<tr>
<td>South Africa (190 isolates)</td>
<td>125 5 1 59</td>
<td>96.2%</td>
<td>98.3%</td>
<td>99.2%</td>
<td>92.2%</td>
</tr>
<tr>
<td>All 1340 isolates</td>
<td>379 14 13 934</td>
<td>96.4%</td>
<td>98.6%</td>
<td>96.7%</td>
<td>98.5%</td>
</tr>
</tbody>
</table>

R, resistant; S, susceptible; PPV, Positive Predictive Value; NPV, Negative Predictive Value.
Table 7: Determination of EMB resistance (755 isolates).

<table>
<thead>
<tr>
<th>EMB</th>
<th>Number of isolates with result (reference / PCR/ESI-MS)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R/R  R/S  S/R  S/S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York (567 isolates)</td>
<td>87   32   9   439</td>
<td>73.1%</td>
<td>98.0%</td>
<td>90.6%</td>
<td>93.2%</td>
</tr>
<tr>
<td>Georgia (188 isolates)</td>
<td>5    1    8   174</td>
<td>83.3%</td>
<td>95.6%</td>
<td>38.5%</td>
<td>99.4%</td>
</tr>
<tr>
<td>All 749 isolates</td>
<td>92   33   17  613</td>
<td>73.6%</td>
<td>97.3%</td>
<td>84.4%</td>
<td>94.9%</td>
</tr>
</tbody>
</table>

R, resistant; S, susceptible; PPV, Positive Predictive Value; NPV, Negative Predictive Value.