

RESEARCH ARTICLE

Molecular drug susceptibility testing and strain typing of tuberculosis by DNA hybridization

Hillary N. Wood¹[✉], Tom Venken²[✉], Hanny Willems²[✉], An Jacobs², Ana Júlia Reis³, Pedro Eduardo Almeida da Silva³, Susanne Homolka^{4,5}, Stefan Niemann^{4,5}, Kyle H. Rohde^{1*}[✉], Jef Hooyberghs^{2,6}

1 Division of Immunity and Pathogenesis, Burnett School of Biomedical Sciences, College of Medicine, University of Central Florida, Orlando, Florida, United States of America, **2** Flemish Institute for Technological Research, VITO, Mol, Belgium, **3** Laboratory of Tuberculosis, Faculty of Medicine, Universidade Federal do Rio Grande- FURG, Rio Grande so Sul, RS, Brazil, **4** Molecular and Experimental Mycobacteriology, Research Center Borstel, Borstel, Germany, **5** German Center for Infection Research, Borstel, Germany, **6** Theoretical Physics, Hasselt University, Diepenbeek, Belgium

 These authors contributed equally to this work.

[✉] Current address: MRIGlobal, Palm Bay, Florida, United States of America

[✉] Current address: VIB Center for Cancer Biology, Leuven, Belgium

[✉] Current address: Laboratory for Translational Genetics, Department of Human Genetics, KU Leuven, Leuven, Belgium

* kyle.rohde@ucf.edu



 OPEN ACCESS

Citation: Wood HN, Venken T, Willems H, Jacobs A, Reis AJ, Almeida da Silva PE, et al. (2019) Molecular drug susceptibility testing and strain typing of tuberculosis by DNA hybridization. PLoS ONE 14(2): e0212064. <https://doi.org/10.1371/journal.pone.0212064>

Editor: HASNAIN SEYED EHTESHAM, Jamia Hamdard, INDIA

Received: October 8, 2018

Accepted: January 25, 2019

Published: February 7, 2019

Copyright: © 2019 Wood et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Microarray data are available from the ArrayExpress database (accession number E-MTAB-7610). All other relevant data are within the manuscript and its Supporting Information files.

Funding: The authors would like to acknowledge funding for this project provided by startup funds from UCF to KHR. Jef Hooyberghs acknowledges funding of a travel grant by FWO (K226314N). The funders had no role in study design, data collection

Abstract

In *Mycobacterium tuberculosis* (*Mtb*) the detection of single nucleotide polymorphisms (SNPs) is of high importance both for diagnostics, since drug resistance is primarily caused by the acquisition of SNPs in multiple drug targets, and for epidemiological studies in which strain typing is performed by SNP identification. To provide the necessary coverage of clinically relevant resistance profiles and strain types, nucleic acid-based measurement techniques must be able to detect a large number of potential SNPs. Since the *Mtb* problem is pressing in many resource-poor countries, requiring low-cost point-of-care biosensors, this is a non-trivial technological challenge. This paper presents a proof-of-concept in which we chose simple DNA-DNA hybridization as a sensing principle since this can be transferred to existing low-cost hardware platforms, and we pushed the multiplex boundaries of it. With a custom designed probe set and a physicochemical-driven data analysis it was possible to simultaneously detect the presence of SNPs associated with first- and second-line drug resistance and *Mtb* strain typing. We have demonstrated its use for the identification of drug resistance and strain type from a panel of phylogenetically diverse clinical strains. Furthermore, reliable detection of the presence of a minority population (<5%) of drug-resistant *Mtb* was possible.

Introduction

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*), remains a major global health problem and is responsible for 10.4 million new infections and 1.8 million deaths annually [1]. The problem is exacerbated by the emergence of drug-resistant MTBC strains. According to

and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

the World Health Organization in 2014, 3.3% of new infections and 20% of previously treated cases were multidrug resistant TB (MDR-TB). MDR-TB strains are resistant to the first-line anti-TB drugs rifampin (RIF) and isoniazid (INH). Furthermore, 9.7% of MDR strains were extensively drug resistant (XDR) based on additional resistance to fluoroquinolones (FQ) and an injectable drug (amikacin, capreomycin, or kanamycin) [1]. Non-molecular diagnostic tools for *Mtb* such as sputum smear microscopy (SSM) and culture have inherent limitations that warrant the development of improved tools for the rapid, sensitive, and accurate detection and drug-resistance profiling of *Mtb*. SSM detects <50% of TB cases with even lower detection rates in children and HIV-positive patients and does not provide information on drug resistance [2, 3]. Culture-based testing is time consuming (requiring 2–6 weeks) and is limited by the requirement for elevated biosafety precautions, trained personal, and risk of contamination [4]. To help eradicate this disease it is important therefore to develop a diagnostic tool that can quickly identify *Mtb* and detect resistance to first- and second-line drugs.

Molecular diagnostic tools enable rapid identification, genotypic drug susceptibility testing (DST) and strain typing of bacterial strains. In clinical *Mtb* strains, the main source of drug resistance is the acquisition of single nucleotide polymorphisms (SNPs) that prevent the interaction of drugs with their altered protein target, alter pro-drug activation, or cause upregulation of drug targets [5]. In addition to the detection of drug resistance, SNPs are used in molecular epidemiology studies for strain typing and are considered the most valid markers for phylogenetic classification of clinical strains of the *Mycobacterium tuberculosis* complex (MTBC) [6, 7]. Defined sets of SNPs for differentiating species and strains within the MTBC have been published [8–11]. Several molecular diagnostic tools such as Cepheid's GeneXpert [12, 13], line probe assays (LiPA) [14] and whole genome sequencing (WGS) technologies are available, but the need for high multiplexing with a simple low-cost robust solution for resource-poor countries is currently still unmet.

The goal of this project was to develop a proof-of-concept, multi-purpose DNA hybridization-based method capable of MTBC detection, genomic DST, and strain typing through the detection of SNPs. We designed a 15-loci multiplex PCR, followed by amplicon detection with a custom, high-density microarray (Fig 1) targeting the most common mutations associated with first- and second-line drug resistance and the strain typing SNPs established by Homolka *et al.* [8]. In principle any hybridization-based hardware platform could have been used for this study, but in the current proof-of-principle stage the maturity of the microarray technology was an advantage. What is crucial about the presented method is that it relies on an established thermodynamic framework [15–19] using unique probe sets which contain mismatch mutations to improve the dynamic range of the detection. These probe sets enable the specific identification of targeted mutations even when present at ~1% relative to wild-type.

Materials and methods

Mycobacterial strains used in this study

DNA for multiplex PCR was isolated from drug-susceptible wild-type (WT) *Mtb* strain CDC1551 (BEI Resources, ATCC) or clinical isolates containing a variety of SNPs. Clinical isolates used in this study (Table C in S1 File) were obtained from archived stocks at the Mycobacteria Laboratory of the Federal University of Rio Grande Brazil or the National Reference Center for Mycobacteria, Borstel, Germany. DNA was extracted using the standardized CTAB (cetyltrimethylammonium bromide-NaCl) method and stored at -20°C [20].

Multiplex PCR (mPCR)

The multiplex PCR assay was designed to produce 15 amplicons (230–1018 base pairs) from genes associated with detection of *Mtb* (16S rRNA and 23S rRNA), drug resistance (*rpoB*,

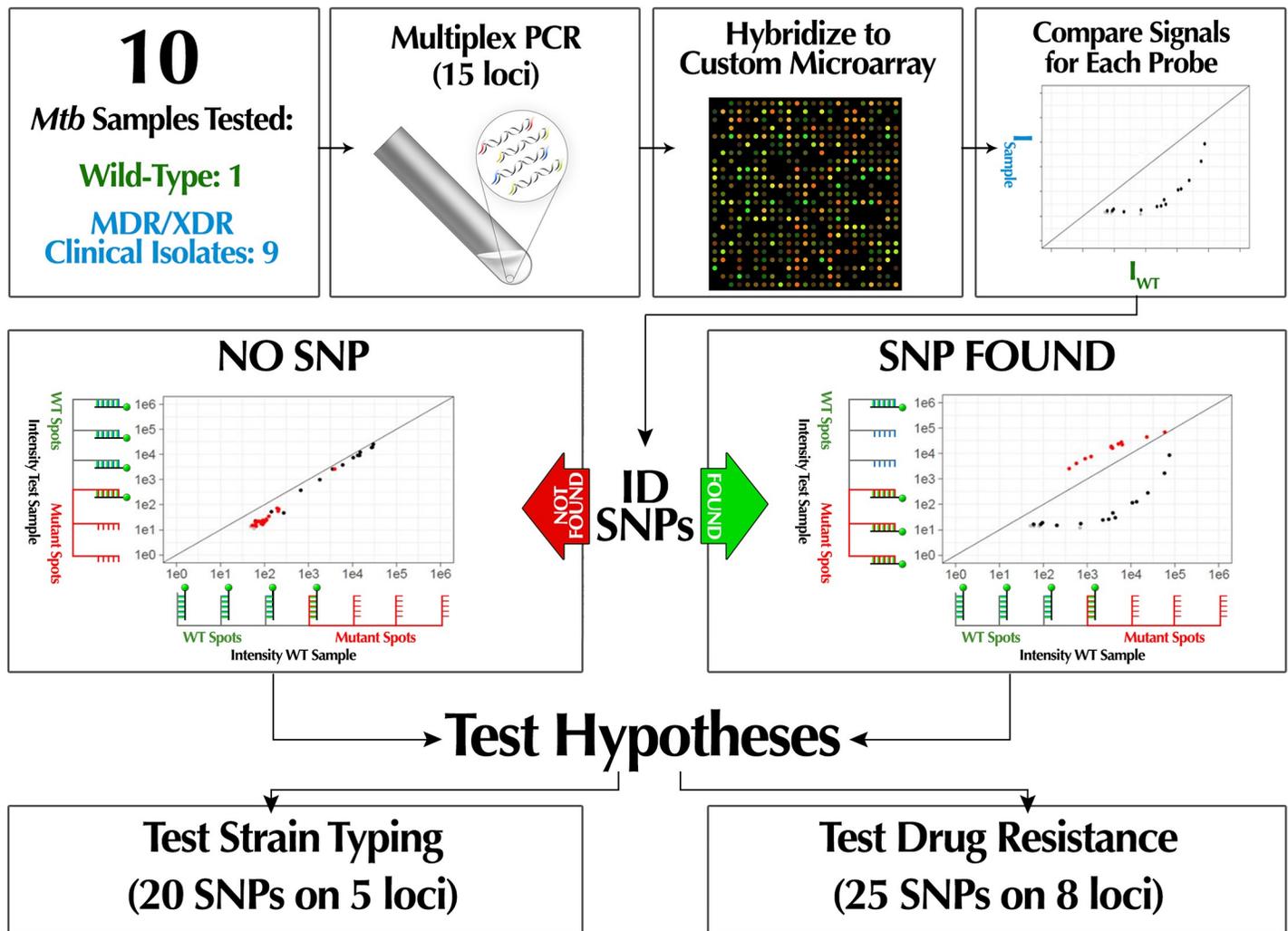


Fig 1. Scheme depicting experimental workflow and data interpretation. A 15-loci multiplex PCR developed in this study serves to generate sufficient analyte DNA from the targeted genes relevant to drug resistance and strain typing. Conceptual diagrams overlaid onto example dot plots of microarray data illustrate how hybridization signals from each probe set can distinguish between two alleles differing by a single polymorphism. Comparison with results from wild-type control samples and alternative mutant alleles allows determination of the specific nucleotide at the tested position.

<https://doi.org/10.1371/journal.pone.0212064.g001>

katG, *inhA*, *gyrA*, *embB*, *pncA*, *rpsL*, and *rrs*), and lineage specific polymorphisms (*Rv0129c*, *Rv1009*, *Rv1811*, *Rv2926*, and *Rv0557*) [8] (**Table A in S1 File**). The most common mutations associated with drug resistance were identified using the tuberculosis drug resistance mutation database (<https://tbdreamdb.ki.se/Info/>). The SNPs associated with strain typing were defined by Homolka *et al.* [8]. Nucleotide sequences of *Mtb* CDC1551 were used to design primers for amplicons encompassing the polymorphisms associated with drug resistance and strain typing. Reverse primers contained a 5' phosphate to allow for digestion of the antisense strand by lambda exonuclease and an adapter sequence (5'-Cy3-aaaaactggcgtcatagctgttctgtgta-3') for hybridization of microarray probes to mPCR products. mPCR samples contained 3 units of Phusion DNA Polymerase (NEB), 147 nM– 1.176 μM of each primer (**Table A in S1 File**), 200 μM of each NTP, 1x Phusion HF Buffer, 20 ng of *Mtb* DNA, and water to a final reaction volume of 25 μL. The samples were PCR amplified using the following cycling conditions: initial denaturation 98 °C for 30 sec, followed by 50 cycles of denaturation at 98 °C for 10 sec,

annealing at 65°C for 40 sec, extension at 72°C for 1 min, and final extension at 72°C for 2 min. Following mPCR, samples were purified using QIAgen PCR purification kit following manufacturer protocol. The dsDNA samples were then lyophilized before shipping to VITO.

Gel validation & densitometry

Multiplex PCR samples were run on a 2% agarose gel at 120 V, 10°C for 12.75 hr followed by staining in GelRed (Biotium, Inc., Fremont, CA) for 2 hr. Densitometry was completed using UN-SCAN-IT gel version 7.1 for relative quantification of each product in the multiplex.

Microarray experimental design

The microarray was designed to enable the specific detection of SNPs associated with *Mtb* drug resistance and strain type. Although amplicons for *16S* and *23S* rRNA were included in the mPCR, we did not include probe sets for these loci on this prototype array because definitive identification of *Mtb* involves distinguishing multiple mutations in hypervariable regions rather than SNPs [21]. Although we envision this as a second-tier assay to provide comprehensive DST and strain typing information on *Mtb* samples detected by existing methods, we could enable species-specific detection of *Mtb* by adding probe sets targeting MTBC specific regions of *16S/23S* rRNA in subsequent versions. Probe sets were designed according to parameters previously established by the Hooyberghs laboratory [16, 18] to enable accurate identification of SNPs. Each probe set is designed to test a specific hypothesis—the presence or absence of a SNP at a defined nucleotide (nt) position. More information on probe set design can be found in the **Supplemental Information (Table B in S1 File)**.

Microarray experiments

Lyophilized mPCR products were resuspended in elution buffer from the QIAgen PCR purification kit. Next, 2 µg of purified mPCR products were treated with 10 U lambda exonuclease (Fermentas, St. Leon-Rot, Germany) in 1 µL reaction buffer and 7 µL water for 30 minutes at 37°C and 10 minutes at 80°C. The concentration of the resulting ssDNA was measured on a NanoDrop spectrophotometer. For microarray hybridization measurements we used the commercially available Agilent microarray platform and followed a standard protocol with Agilent products. Hybridization mixtures were prepared by combining a Cy3-labeled adapter (5'-Cy3-aaaaactggcgtcatagctgttctgtgtga-3') diluted in nuclease-free water to a final concentration of 0.05 µM with 500 ng ssDNA, 5 µl 10x blocking agent and 25 µl 2x GEx hybridization buffer HI-RPM. Each hybridization mixture (40 µl) containing ssDNA mPCR products of a single test sample was centrifuged at 13,000 rpm for 1 min and then added to one well on the 8×15K custom Agilent slide. Hybridization occurred in an Agilent oven at 65°C for 17 h with rotor setting 10 and the washing was performed according to the manufacturer's instructions. The slides were scanned on an Agilent scanner (G2565BA) at 5-µm resolution and further processed using Agilent Feature Extraction Software (GE1 v5 95 Feb07) that performs automatic gridding, fluorescence intensity measurement, background subtraction and quality checks.

Analysis of mixed-strain multiplex PCR samples using microarray

To determine the ability of a custom microarray to identify a mixed infection in which mutations are often present in low abundance compared to WT, we simulated this scenario using our mPCR assay to amplify DNA from two strains: *Mtb* CDC1551 (WT) and the MDR-TB clinical isolate 08–1074 (MT) (refer to **Table C in S1 File**). We completed a mPCR with 4 concentrations of MT DNA (1 ng, 250 pg, 62.5 pg, and 15.6 pg) in the presence of excess WT

DNA (20 ng) which is equivalent to approximately 5%, 1.25%, 0.31% and 0.08% MT DNA, respectively. The samples were then processed and applied to the microarray as described above.

Data analysis

A detailed description of the statistical analysis used in this study is described in **Supplemental Information (Figure B in S1 File)**. The data was analyzed by comparing the fluorescence intensities of the WT reference sample (x-axis) versus the test sample (y-axis) on a scatterplot. The scatterplots allow for easy visualization of the presence or absence of mutation, but results are only considered statistically significant if the data passes three statistical tests. As described by Willems et al. [22], a scatterplot contains branches representing hybridization based on at least two probe sets: the reference branch (or WT, black dots) and the mutant branch (the targeted SNP, red dots). In the case of drug resistance, we tested for all possible SNPs; therefore, two additional branches are formed and are referred to as side branches (alternative SNPs, blue dots). Dots below the detection limit were colored grey in the scatterplots. When the test sample is WT all branches will be on the diagonal line. Alternatively, when the test sample contains a SNP the branches separate; the mutant branch (corresponding to the targeted SNP) will have the highest intensity, the reference branch will have the lowest, and the side branches are intermediate. Furthermore, we designed schemes (Figs 2C and 3D) to summarize the results for drug resistance and strain typing. Only clinically relevant mutations are included in the drug resistance scheme. To validate genotypes determined by microarray for all strains, results were compared with whole genome sequencing data where available (strains 03/8864, 03/9532,

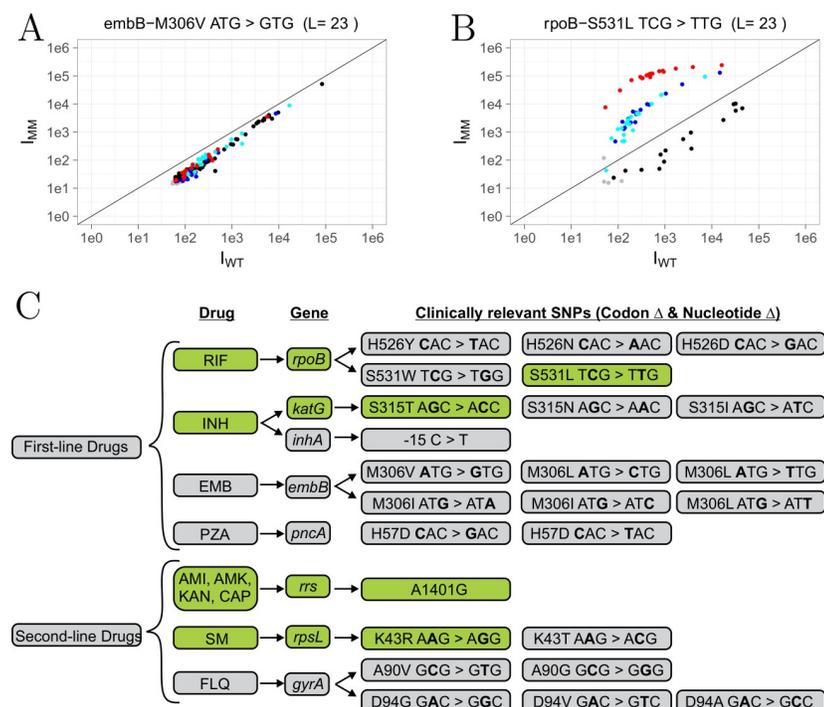


Fig 2. Microarray drug resistance results of XDR-TB strain (08-1186). A) Example of wild-type microarray result. B) Example of microarray result showing detection of mutation in *rpoB* codon 531. Fluorescence intensities of the wild-type sample (x-axis, I_{WT}) are plotted versus the test sample (y-axis, I_{MM} , 08-1186). C) Drug resistance scheme with detected mutations in green. L, probe length (nt).

<https://doi.org/10.1371/journal.pone.0212064.g002>

03/4850) or targeting sequencing (Eurofins) of PCR amplicons of individual loci from the remaining strains.

Results

Design and validation of 15 loci multiplex PCR (mPCR)

The mPCR was designed to amplify fragments of 15 genes (Table A in S1 File) associated with the specific detection of *Mtb*, genotypic DST and strain typing of *Mtb*. After optimization of the cycling conditions and primer concentrations (Table A in S1 File), we achieved <4-fold variation in relative concentration of PCR products as determined by gel quantification software. Our mPCR assay reliably amplified all 15 specific products at the correct size and with few nonspecific products as shown in Figure A in S1 File. Nonspecific products (Figure A in S1 File, white arrows) were present in the mPCR, but did not interfere with the accurate and specific detection of 45 mutations (MT).

Microarray results

Genotypic DST. To evaluate the utility of the customized microarray to detect MDR and XDR-TB, we analyzed DNA from 9 clinical isolates with diverse phenotypic and genotypic profiles. Complete results are summarized in Table C in S1 File and supplementary figures. Selected results are presented here to illustrate the data analysis process. Fig 2 shows the drug-resistance SNP profiling results for an XDR-TB strain (08–1186). The scatterplot shown in Fig 2A illustrates the pattern observed when a strain is wild-type at codon 306 of *embB* where an ATG>GTG SNP would indicate ethambutol resistance. In this case, the fluorescence intensity of both wild-type reference (I_{WT}) and the test sample signal (I_{MM}) lie on a diagonal line indicating that no mutation was found. In contrast, Fig 2B demonstrates the presence of a mutation in *rpoB* codon 531 (TCG > TTG), which is responsible for rifampicin resistance. The most deviating side branch (Fig 2B, red dots), the mutant branch, corresponds to the probes complementary to the targeted *rpoB* SNP and therefore has the highest signal intensity compared to WT. Analysis of scatterplots for each targeted mutation (25 SNPs in total) was completed as described in Materials and Methods and the resulting scheme is shown in Fig 2C. In all cases, sequence analysis confirmed the genotype assigned based on microarray results (Table C in S1 File).

Strain typing. To classify *Mtb* isolates based on phylogenetic lineage we designed probe sets based on SNPs in five genes described by Homolka *et al.* as shown in Fig 3D (scheme, adopted from Homolka *et al.* [8]). The 19 SNPs tested in each sample were previously shown to be superior to standard MIRU-VNTR typing for defining deep phylogenetics groups with high confidence [8]. Here we demonstrate the strain typing results of strain 08–1186 of the Beijing genotype. Fig 3A–3C illustrates the scatterplots of each green box highlighted in Fig 3D. The microarray result confirms the presence of two strain-specific SNPs, a non Euro-American SNP (*Rv0557*, nt 321, T>C—Fig 3A) and Beijing SNP (*Rv2629*, nt 191, A>C—Fig 3C). No other mutations were found as indicated by grey boxes in the scheme. Fig 3B shows that no mutation was present in the *Rv0557* gene at position nt 221 as an illustration of a negative result. The strain types for all other strains are summarized in Table B in S1 File and corresponding schemes and scatterplots are found in Supplemental Information (S1 File, pg 12–28). All strain types were confirmed by sequencing.

Probe length comparison. In this proof-of-concept study, each hypothesis (presence/absence of a SNP at a defined location) was tested using multiple probe lengths (between 21–25 nt) to optimize both the microarray signal sensitivity and dynamic range. In some cases, for example the detection of AGC > ACC in *katG* codon 315 (Fig 4A and 4B), the dynamic range

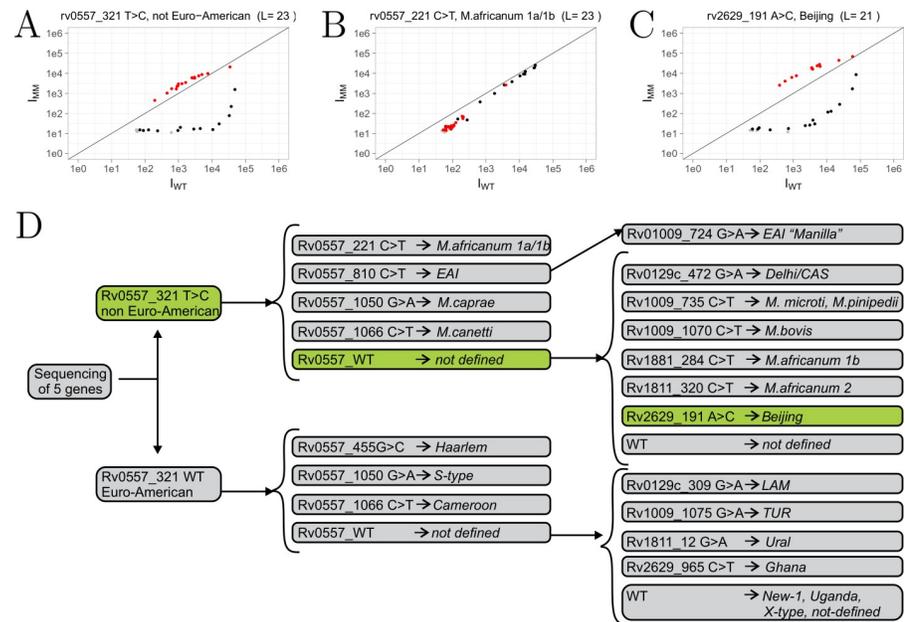


Fig 3. Microarray strain typing results of Beijing strain (08–1186). Fluorescence intensities of the wild-type sample (x-axis, I_{WT}) plotted versus the test sample (y-axis, I_{MM}). **A)** Microarray result for probe set targeting *Rv0557* nt 321. **B)** Example where no mutation was found in *Rv0557* at nt 221. **C)** Example for *Rv2629* nt 191 A>C confirming Beijing strain type. **D)** Strain typing scheme with detected mutations in green. L, probe length (nt).

<https://doi.org/10.1371/journal.pone.0212064.g003>

was altered when using short (23nt) vs long (25nt) probes without affecting SNP identification. The smaller dynamic range observed with probe length of 25 nt arises from a stronger increase in hybridization free energies between the longer probes and the WT sample compared to the MT sample (Fig 4B). However, in a few cases (<1%) probe length affected the ability to confidently determine the presence or absence of a SNP. One of these scenarios is illustrated in Fig 4C and 4D. In this example, the test sample (strain 08–1186, Beijing genotype) does not contain the *M. canetti* specific SNP (C > T) at position 1066 in *Rv0557*. Despite Test 2 providing the correct result (False = no SNP), Test1 suggests the presence of a mutation using a 25 nt probe ($p = 0.0025$). In this case, the 23 nt probe allowed us to confirm the absence of a SNP by a significant negative result ($p = 0.9952$). These results demonstrate how probe length can be used to optimize the detection of specific SNPs, if needed. Furthermore, we completed a global analysis by evaluating the effect of probe length for each hypothesis tested and we conclude that the use of two probe lengths for each hypothesis tested is redundant in the majority of cases (>99%).

Detection of mixed infections. Given the clinical relevance of heteroresistance [23–27], we assessed the utility of our microarray to detect drug resistant *Mtb* in a mixed infection in which only a minority subpopulation of bacilli harbor resistance-conferring mutations. We performed an mPCR with samples containing a mixture of template DNA from WT (CDC1551) and 5%, 1.25%, 0.31% or 0.08% of an XDR-TB strain (strain 08–1074, Beijing genotype). This concentration range was chosen to determine the relative sensitivity of the microarray assay for the detection of resistant alleles in the context of a mixed infection. The use of an XDR-strain from a different strain-type background allowed us to evaluate detection of mutations with several probes sets in a single experiment. Based on analysis of probe sets targeting *rpoB* codon 531, we were able to detect drug resistance mutations even in the presence of ~80-fold excess WT DNA (1.25% MT concentration, Fig 5). The capability of our

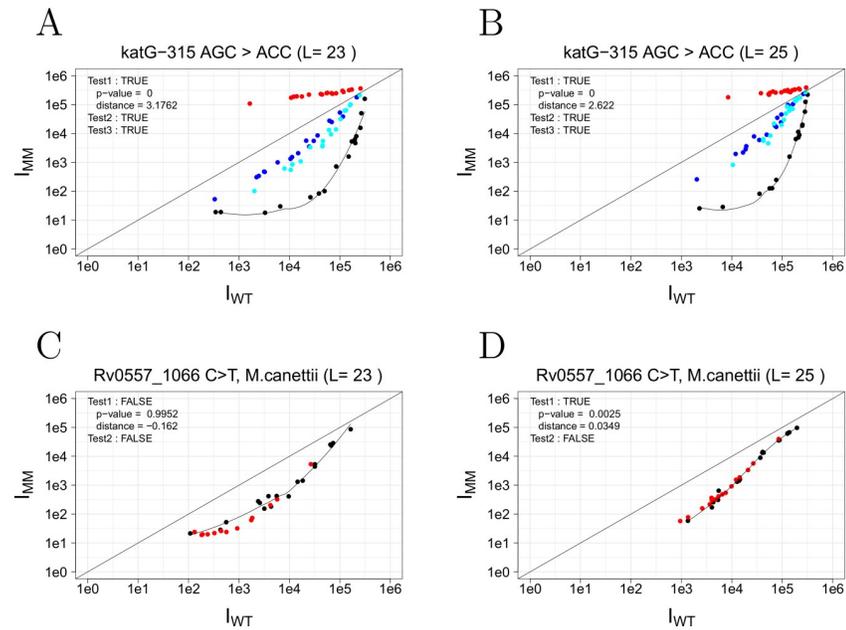


Fig 4. Probe length comparison of strain 08-1186. Fluorescence intensities of the wild-type sample (x-axis, I_{WT}) plotted versus the test sample (y-axis, I_{MM}). Probe set targeting the SNP (AGC > ACC) in *katG* codon 315 using a probe length (L) of **A**) 23 nt or **B**) 25 nt. Probe set targeting C > T SNP in *Rv0557* at position 1066 using **C**) 23 nt or **D**) 25 nt probe length.

<https://doi.org/10.1371/journal.pone.0212064.g004>

hybridization platform to detect the presence of low abundance mutations could be particularly valuable for early identification of drug resistance, before resistant clones come to dominate the infection. **Table 1** summarizes the results of the probe lengths comparison for the probe sets targeting mutations present in the XDR-TB strain tested in this experiment. The assay reliably detected the presence of mutations as low as 1.25% with some probe sets (**Table 1**). The inability to detect the *katG* 315 mutant allele when present at 5% of the population is consistent sequence context dependence of hybridization we have reported previously [22].

Discussion

In this study, we demonstrated the use of a 15-locus mPCR coupled with a custom, high-density microarray for robust detection of SNPs clinically relevant to TB diagnosis. This proof-of concept assay, capable of identifying 45 distinct alleles, was designed to enable simultaneous genotypic DST and strain typing of *Mtb*. Based on best estimates of clinical prevalence of drug resistance mutations, this prototype assay would detect >95% of rifampicin resistance, 70–90% of isoniazid resistance, and > 80% of strains resistance to second-line fluoroquinolones, as well as common mutations associated with resistance to additional clinically utilized therapeutics [28–31]. We tested this platform using *Mtb* clinical isolates from diverse lineages, including MDR and XDR strains, containing a variety of SNPs and successfully identified the presence/absence of SNPs in all cases.

Furthermore, we addressed the clinically relevant problem of detecting mixed *Mtb* infections, often referred to as heteroresistance. Heteroresistance can occur early during treatment as *Mtb* acquires drug resistance-conferring SNPs or upon reinfection with a drug-resistant strain [32]. Current methods are unable to detect low levels of drug-resistant bacilli which results in a delay of appropriate treatment. The WHO endorsed GeneXpert platform for

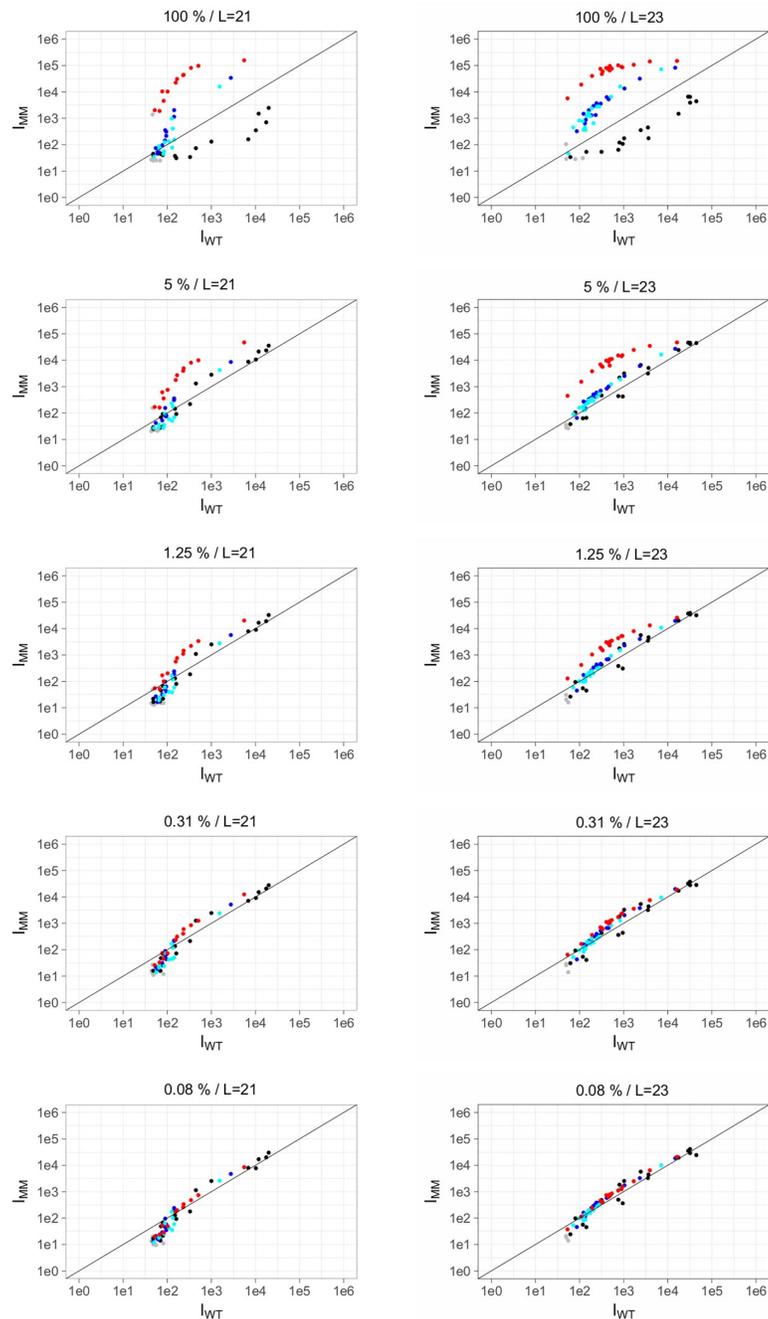


Fig 5. Detection of heteroresistance using probe sets targeting *rpoB* codon 531, SNP TCG > TTG (S531L). Probe lengths (L) of 21 or 23 nt long were used. A wild-type sample (CDC1511) and a mutant sample (strain 08–1074, Beijing genotype) were combined using the following mutant concentrations: 100%, 5%, 1.25%, 0.31% and 0.08%.

<https://doi.org/10.1371/journal.pone.0212064.g005>

detecting rifampin resistance, which relies on loss of signal in the presence of a MT, can only detect drug-resistant alleles when present in ~90% of the population [33]. According to a recent report, the new GeneXpert Ultra assay exhibited enhanced sensitivity for detection of heteroresistance for select *rpoB* mutations ranged from 10–40% mutant allele [34]. Performance for detection of other alleles from a mixed sample was not tested. In contrast, our

Table 1. Summary of drug resistance results for probe sets targeting differences between the WT and MT strain (08–1074).

% 08–1074	<i>rpoB</i> codon 531 TCG > TTG		<i>katG</i> codon 315 AGC > ACC		<i>rpsL</i> codon 43 AAG > AGG		<i>embB</i> codon 306 ATG > ATA	
	L = 21	L = 23	L = 23	L = 25	L = 23	L = 25	L = 23	L = 25
5	+	+	-	-	+	+	+	-
1.25	+	-	-	-	+	-	-	-
0.31	-	-	-	-	-	-	-	-
0.08	-	-	-	-	-	-	-	-

The plus (+) sign indicates the ability to identify the presence of the mutation with statistical significance. The minus (-) sign indicates the inability to detect the mutation by failure of any of the 3 statistical tests. L, probe length (nt).

<https://doi.org/10.1371/journal.pone.0212064.t001>

hybridization approach enables the specific detection of MTs by utilizing probe sets specific to SNPs of interest enabling detection of MTs present at ~1.25% relative to WT (Table 1). This is a significant advantage for the early detection and treatment of drug-resistant *Mtb*. Similarly, this platform could also be designed to detect mixed infection involving multiple pathogens in addition to drug resistant variants of *Mtb*.

In the future, we can improve our assay by including probe sets for ~50 additional clinically relevant SNPs. In this proof-of-concept design, we included probe sets of two lengths for each targeted MT to enable robust identification of SNPs. We found in the majority of cases (>99%) the use of two probe sets was not required to identify the presence or absence of a SNP. Furthermore, our microarray included 5 technical replicates and 12–17 probes for each targeted MT enabling a high dynamic range and ability to differentiate SNPs. These features allow for the specific identification of the MT present and the ability to identify MTs present in a minor population of target analyte. The importance of multiple probes for detection of a SNP is highlighted by the ability of our platform to specifically identify mutations despite the presence of polymorphisms present in the same or nearby codons (Table C in S1 File, A, in blue). This can occur due to strain dependent genetic variants unrelated to drug resistance or unknown SNPs conferring resistance not targeted by probes that could interfere with hybridization based detection [35, 36]. In this scenario, there is a global loss of intensity with all probe sets but the relative intensity difference between the branches remains the same enabling the identification of a SNP. This is a major advantage of our platform compared to current hybridization based platforms for the detection of drug resistance. However, if this level of sensitivity is not required, the number of probe sequence variations for each set can be decreased to allow for the detection of even more clinically relevant SNPs.

Conclusions

We have demonstrated the use of this methodology as a promising alternative to current molecular methods for DST and strain typing for *Mtb*. We envision positioning of a diagnostic tool for high-resolution SNP analysis of DST and strain typing based on this proof-of-concept platform as a second-line assay following initial detection of *Mtb*. Miniaturization of this hybridization based platform as a lab-on-chip product could be amenable for front-line use. The principles of the current fluorescence-based approach can in principle be transferred to other methods such as impedimetric biosensors or surface plasmon resonance interfaces [37, 38]. The data analysis is robust and can be automated, making the methodology simpler compared to next-generation sequencing approaches. Other groups have utilized microarray-based approaches, but are limited in their coverage of clinically relevant SNPs and use fewer probes per SNP [39, 40]. A large number of probes are not always necessary for each SNP, as

this depends on the dynamic range in each case. However, this method ensures that also minor mutations (from a thermodynamic point of view) can be detected accurately. The methodology presented here for the detection of SNPs offers a novel alternative for the detection of *Mtb*, strain type, and drug susceptibility with the potential to provide next day results. Recently, our hybridization-based approach for the detection of SNPs was applied to the simultaneous detection of a limited number of clinically relevant KRAS oncogene point mutations [22]. This study further validates and extends our platform for the robust detection of clinically relevant SNPs as a valuable diagnostic tool for a variety of priority health concerns ranging from cancer to infectious disease.

Supporting information

S1 File. Table A. Multiplex PCR primers. Table B. Example of probe set design. Table C. Summary of drug resistance and strain typing results for all strains tested. Figure A. Agarose gel validation of fifteen loci multiplex PCR. Lane M is a 2 log DNA ladder (NEB) and lane 1 is the mPCR profile of WT *Mtb* DNA (CDC1551 DNA). The white arrows indicate non-specific PCR products. This gel is representative of all mPCR products used in this study. **Figure B. Illustration of statistical tests. A)** Example where the minimal distance (red arrow) of the mutant branch (red dots) from the LOWESS curve (black curve) is higher than the p75 quantile distances of the reference branch (black dots) with the LOWESS curve. The black arrows represent the distances of each spot of the reference branch from the LOWESS curve, from which the p75 quantile is subsequently calculated. **B)** Same as in A, but here the minimal distance of the mutant branch from the LOWESS curve is too small and Test 2 is rejected. **C)** Example where the median distance (red arrow) of the mutant branch (red dots) from the LOWESS curve (black curve) is higher than the median distance (blue or cyan arrow) of each side branch (blue and cyan dots) from the LOWESS curve. **D)** Same as C, but here the median distance of the mutant branch is smaller than the median distances of the side branches, thus Test 3 is rejected. (DOCX)

Acknowledgments

The authors are grateful to Dr. Travis Jewett for technical support, Dr. Inge Nelissen for advice, and Ryan Dickerson of the UCF Educational Technology Department for assistance with preparation of graphics.

Author Contributions

Conceptualization: Hillary N. Wood, Tom Venken, Kyle H. Rohde, Jef Hooyberghs.

Data curation: Hillary N. Wood, Tom Venken, Hanny Willems, An Jacobs, Ana Júlia Reis, Jef Hooyberghs.

Formal analysis: Hillary N. Wood, Tom Venken, Hanny Willems, An Jacobs, Jef Hooyberghs.

Funding acquisition: Jef Hooyberghs.

Investigation: Hillary N. Wood, Tom Venken, Hanny Willems, An Jacobs.

Methodology: Hillary N. Wood, Tom Venken, Hanny Willems.

Project administration: Pedro Eduardo Almeida da Silva, Susanne Homolka, Stefan Niemann, Kyle H. Rohde, Jef Hooyberghs.

Resources: Ana Júlia Reis, Pedro Eduardo Almeida da Silva, Susanne Homolka, Stefan Niemann, Kyle H. Rohde, Jef Hooyberghs.

Supervision: Pedro Eduardo Almeida da Silva, Kyle H. Rohde.

Visualization: Hillary N. Wood, Tom Venken, Hanny Willems, An Jacobs.

Writing – original draft: Hillary N. Wood, Tom Venken, Kyle H. Rohde, Jef Hooyberghs.

Writing – review & editing: Hillary N. Wood, Tom Venken, Hanny Willems, An Jacobs, Ana Júlia Reis, Pedro Eduardo Almeida da Silva, Susanne Homolka, Stefan Niemann, Kyle H. Rohde, Jef Hooyberghs.

References

1. Organization WH. Global Tuberculosis Report 2015: World Health Organization; 2015 [cited 2015]. Available from: <http://www.who.int/iris/handle/10665/191102>.
2. Shingadia D, Novelli V. Diagnosis and treatment of tuberculosis in children. *Lancet Infect Dis*. 2003; 3(10):624–32. PMID: 14522261.
3. Steingart KR, Ramsay A, Pai M. Optimizing sputum smear microscopy for the diagnosis of pulmonary tuberculosis. *Expert Rev Anti Infect Ther*. 2007; 5(3):327–31. <https://doi.org/10.1586/14787210.5.3.327> PMID: 17547496.
4. Dunlap NE, Harris RH, Benjamin WH Jr., Harden JW, Hafner D. Laboratory contamination of *Mycobacterium tuberculosis* cultures. *Am J Respir Crit Care Med*. 1995; 152(5 Pt 1):1702–4. <https://doi.org/10.1164/ajrccm.152.5.7582316> PMID: 7582316.
5. Dookie N, Rambaran S, Padayatchi N, Mahomed S, Naidoo K. Evolution of drug resistance in *Mycobacterium tuberculosis*: a review on the molecular determinants of resistance and implications for personalized care. *J Antimicrob Chemother*. 2018. <https://doi.org/10.1093/jac/dkx506> PMID: 29360989; PubMed Central PMCID: PMC5909630.
6. Comas I, Homolka S, Niemann S, Gagneux S. Genotyping of genetically monomorphic bacteria: DNA sequencing in *mycobacterium tuberculosis* highlights the limitations of current methodologies. *PLoS One*. 2009; 4(11):e7815. Epub 2009/11/17. <https://doi.org/10.1371/journal.pone.0007815> PMID: 19915672; PubMed Central PMCID: PMC2772813.
7. Stucki D, Ballif M, Egger M, Furrer H, Altpeter E, Battegay M, et al. Standard Genotyping Overestimates Transmission of *Mycobacterium tuberculosis* among Immigrants in a Low-Incidence Country. *J Clin Microbiol*. 2016; 54(7):1862–70. <https://doi.org/10.1128/JCM.00126-16> PMID: 27194683; PubMed Central PMCID: PMC4922098.
8. Homolka S, Projahn M, Feuerriegel S, Ubben T, Diel R, Nubel U, et al. High resolution discrimination of clinical *Mycobacterium tuberculosis* complex strains based on single nucleotide polymorphisms. *PLoS One*. 2012; 7(7):e39855. Epub 2012/07/07. PONE-D-12-02200 [pii]. <https://doi.org/10.1371/journal.pone.0039855> PMID: 22768315.
9. Comas I, Gagneux S. The past and future of tuberculosis research. *PLoS Pathog*. 2009; 5(10):e1000600. Epub 2009/10/27. <https://doi.org/10.1371/journal.ppat.1000600> PMID: 19855821; PubMed Central PMCID: PMC2745564.
10. Abadia E, Zhang J, dos Vultos T, Ritacco V, Kremer K, Aktas E, et al. Resolving lineage assignment on *Mycobacterium tuberculosis* clinical isolates classified by spoligotyping with a new high-throughput 3R SNPs based method. *Infect Genet Evol*. 2010; 10(7):1066–74. <https://doi.org/10.1016/j.meegid.2010.07.006> PMID: 20624486.
11. Dos Vultos T, Mestre O, Rauzier J, Golec M, Rastogi N, Rasolofo V, et al. Evolution and diversity of clonal bacteria: the paradigm of *Mycobacterium tuberculosis*. *PLoS One*. 2008; 3(2):e1538. Epub 2008/02/07. <https://doi.org/10.1371/journal.pone.0001538> PMID: 18253486; PubMed Central PMCID: PMC2211405.
12. El-Hajj HH, Marras SA, Tyagi S, Kramer FR, Alland D. Detection of rifampin resistance in *Mycobacterium tuberculosis* in a single tube with molecular beacons. *J Clin Microbiol*. 2001; 39(11):4131–7. <https://doi.org/10.1128/JCM.39.11.4131-4137.2001> PMID: 11682541; PubMed Central PMCID: PMC5909630.
13. Helb D, Jones M, Story E, Boehme C, Wallace E, Ho K, et al. Rapid detection of *Mycobacterium tuberculosis* and rifampin resistance by use of on-demand, near-patient technology. *J Clin Microbiol*. 2010; 48(1):229–37. <https://doi.org/10.1128/JCM.01463-09> PMID: 19864480; PubMed Central PMCID: PMC2812290.

14. Makinen J, Marttila HJ, Marjamaki M, Viljanen MK, Soini H. Comparison of two commercially available DNA line probe assays for detection of multidrug-resistant *Mycobacterium tuberculosis*. *J Clin Microbiol*. 2006; 44(2):350–2. <https://doi.org/10.1128/JCM.44.2.350-352.2006> PMID: 16455882; PubMed Central PMCID: PMCPMC1392655.
15. Hooyberghs J, Carlon E. Hybridisation thermodynamic parameters allow accurate detection of point mutations with DNA microarrays. *Biosens Bioelectron*. 2010; 26(4):1692–5. <https://doi.org/10.1016/j.bios.2010.07.039> PMID: 20719495.
16. Hooyberghs J, Van Hummelen P, Carlon E. The effects of mismatches on hybridization in DNA microarrays: determination of nearest neighbor parameters. *Nucleic Acids Res*. 2009; 37(7):e53. <https://doi.org/10.1093/nar/gkp109> PMID: 19270064; PubMed Central PMCID: PMCPMC2673445.
17. Hadiwikarta WW, Carlon E, Hooyberghs J. Dynamic range extension of hybridization sensors. *Biosens Bioelectron*. 2015; 64:411–5. Epub 2014/10/04. <https://doi.org/10.1016/j.bios.2014.09.043> PMID: 25280340.
18. Hadiwikarta WW, Walter JC, Hooyberghs J, Carlon E. Probing hybridization parameters from microarray experiments: nearest-neighbor model and beyond. *Nucleic Acids Res*. 2012; 40(18):e138. <https://doi.org/10.1093/nar/gks475> PMID: 22661582; PubMed Central PMCID: PMCPMC3467032.
19. Hooyberghs J, Baiesi M, Ferrantini A, Carlon E. Breakdown of thermodynamic equilibrium for DNA hybridization in microarrays. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2010; 81(1 Pt 1):012901. <https://doi.org/10.1103/PhysRevE.81.012901> PMID: 20365418.
20. van Soolingen D, de Haas PE, Hermans PW, van Embden JD. DNA fingerprinting of *Mycobacterium tuberculosis*. *Methods Enzymol*. 1994; 235:196–205. Epub 1994/01/01. PMID: 8057895.
21. Chakravorty S, Helb D, Burday M, Connell N, Alland D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J Microbiol Methods*. 2007; 69(2):330–9. <https://doi.org/10.1016/j.mimet.2007.02.005> PMID: 17391789; PubMed Central PMCID: PMCPMC2562909.
22. Willems H, Jacobs A, Hadiwikarta WW, Venken T, Valkenburg D, Van Roy N, et al. Thermodynamic framework to assess low abundance DNA mutation detection by hybridization. *PLoS One*. 2017; 12(5): e0177384. Epub 2017/05/26. <https://doi.org/10.1371/journal.pone.0177384> PMID: 28542229; PubMed Central PMCID: PMCPMC5444680.
23. Zhang D, Gomez JE, Chien JY, Haseley N, Desjardins CA, Earl AM, et al. Genomic analysis of the evolution of fluoroquinolone resistance in *Mycobacterium tuberculosis* prior to tuberculosis diagnosis. *Antimicrob Agents Chemother*. 2016. <https://doi.org/10.1128/AAC.00664-16> PMID: 27572408.
24. Cohen T, van Helden PD, Wilson D, Colijn C, McLaughlin MM, Abubakar I, et al. Mixed-strain *Mycobacterium tuberculosis* infections and the implications for tuberculosis treatment and control. *Clin Microbiol Rev*. 2012; 25(4):708–19. <https://doi.org/10.1128/CMR.00021-12> PMID: 23034327; PubMed Central PMCID: PMCPMC3485752.
25. Folkvardsen DB, Thomsen VO, Rigouts L, Rasmussen EM, Bang D, Bernaerts G, et al. Rifampin heteroresistance in *Mycobacterium tuberculosis* cultures as detected by phenotypic and genotypic drug susceptibility test methods. *J Clin Microbiol*. 2013; 51(12):4220–2. <https://doi.org/10.1128/JCM.01602-13> PMID: 24068005; PubMed Central PMCID: PMCPMC3838044.
26. Zhang X, Zhao B, Liu L, Zhu Y, Zhao Y, Jin Q. Subpopulation analysis of heteroresistance to fluoroquinolone in *Mycobacterium tuberculosis* isolates from Beijing, China. *J Clin Microbiol*. 2012; 50(4):1471–4. <https://doi.org/10.1128/JCM.05793-11> PMID: 22259208; PubMed Central PMCID: PMCPMC3318552.
27. Tolani MP, D'Souza D T, Mistry NF. Drug resistance mutations and heteroresistance detected using the GenoType MTBDRplus assay and their implication for treatment outcomes in patients from Mumbai, India. *BMC Infect Dis*. 2012; 12:9. <https://doi.org/10.1186/1471-2334-12-9> PMID: 22260344; PubMed Central PMCID: PMCPMC3283446.
28. Ahmad N, Javaid A, Sulaiman SA, Ming LC, Ahmad I, Khan AH. Resistance patterns, prevalence, and predictors of fluoroquinolones resistance in multidrug resistant tuberculosis patients. *Braz J Infect Dis*. 2016; 20(1):41–7. <https://doi.org/10.1016/j.bjid.2015.09.011> PMID: 26626164.
29. Avalos E, Catanzaro D, Catanzaro A, Ganiats T, Brodine S, Alcaraz J, et al. Frequency and geographic distribution of *gyrA* and *gyrB* mutations associated with fluoroquinolone resistance in clinical *Mycobacterium tuberculosis* isolates: a systematic review. *PLoS One*. 2015; 10(3):e0120470. <https://doi.org/10.1371/journal.pone.0120470> PMID: 25816236; PubMed Central PMCID: PMCPMC4376704.
30. Bobadilla-del-Valle M, Ponce-de-Leon A, Arenas-Huertero C, Vargas-Alarcon G, Kato-Maeda M, Small PM, et al. *rpoB* Gene mutations in rifampin-resistant *Mycobacterium tuberculosis* identified by polymerase chain reaction single-stranded conformational polymorphism. *Emerg Infect Dis*. 2001; 7(6):1010–3. <https://doi.org/10.3201/eid0706.010615> PMID: 11747730; PubMed Central PMCID: PMCPMC2631921.
31. Lorenzo D, Mousa SA. Mechanisms of drug resistance in *Mycobacterium tuberculosis* and current status of rapid molecular diagnostic testing. *Acta Trop*. 2011; 119(1):5–10. Epub 2011/04/26. S0001-706X

- (11)00078-7 [pii] 10.1016/j.actatropica.2011.04.008. <https://doi.org/10.1016/j.actatropica.2011.04.008> PMID: 21515239.
32. Small PM, Shafer RW, Hopewell PC, Singh SP, Murphy MJ, Desmond E, et al. Exogenous reinfection with multidrug-resistant *Mycobacterium tuberculosis* in patients with advanced HIV infection. *N Engl J Med*. 1993; 328(16):1137–44. <https://doi.org/10.1056/NEJM199304223281601> PMID: 8096066.
 33. Cho SN, Brennan PJ. Tuberculosis: diagnostics. *Tuberculosis (Edinb)*. 2007; 87 Suppl 1:S14–7. <https://doi.org/10.1016/j.tube.2007.05.001> PMID: 17584529.
 34. Chakravorty S, Simmons AM, Rowneki M, Parmar H, Cao Y, Ryan J, et al. The New Xpert MTB/RIF Ultra: Improving Detection of *Mycobacterium tuberculosis* and Resistance to Rifampin in an Assay Suitable for Point-of-Care Testing. *MBio*. 2017; 8(4). <https://doi.org/10.1128/mBio.00812-17> PMID: 28851844; PubMed Central PMCID: PMC5574709.
 35. Mathys V, van de Vyvere M, de Droogh E, Soetaert K, Groenen G. False-positive rifampicin resistance on Xpert(R) MTB/RIF caused by a silent mutation in the *rpoB* gene. *Int J Tuberc Lung Dis*. 2014; 18(10):1255–7. Epub 2014/09/14. <https://doi.org/10.5588/ijtld.14.0297> PMID: 25216843.
 36. Feuerriegel S, Koser CU, Niemann S. Phylogenetic polymorphisms in antibiotic resistance genes of the *Mycobacterium tuberculosis* complex. *J Antimicrob Chemother*. 2014; 69(5):1205–10. Epub 2014/01/25. <https://doi.org/10.1093/jac/dkt535> PMID: 24458512.
 37. van Grinsven B, Vanden Bon N, Grieten L, Murib M, Janssens SD, Haenen K, et al. Rapid assessment of the stability of DNA duplexes by impedimetric real-time monitoring of chemically induced denaturation. *Lab Chip*. 2011; 11(9):1656–63. Epub 2011/03/31. <https://doi.org/10.1039/c1lc20027e> PMID: 21448492.
 38. Zagorodko O, Spadavecchia J, Serrano AY, Larroulet I, Pesquera A, Zurutuza A, et al. Highly sensitive detection of DNA hybridization on commercialized graphene-coated surface plasmon resonance interfaces. *Anal Chem*. 2014; 86(22):11211–6. Epub 2014/10/24. <https://doi.org/10.1021/ac502705n> PMID: 25341125.
 39. Linger Y, Kukhtin A, Golova J, Perov A, Lambarqui A, Bryant L, et al. Simplified microarray system for simultaneously detecting rifampin, isoniazid, ethambutol, and streptomycin resistance markers in *Mycobacterium tuberculosis*. *J Clin Microbiol*. 2014; 52(6):2100–7. <https://doi.org/10.1128/JCM.00238-14> PMID: 24719444; PubMed Central PMCID: PMC4042730.
 40. Tang P, Wang X, Shen X, Shi M, Zhu X, Yu X, et al. Use of DNA microarray chips for the rapid detection of *Mycobacterium tuberculosis* resistance to rifampicin and isoniazid. *Experimental and Therapeutic Medicine*. 2017; 13(5):2332–8. PMC5443298. <https://doi.org/10.3892/etm.2017.4250> PMID: 28565846