Detecting Rifampin Resistance in *Mycobacterium tuberculosis* Using Different Technologies

Aman Melles

Supervisor: Pontus Juréen

Rifampin is one of the most important drugs in the treatment of tuberculosis. Its widespread use makes it important to detect rifampin resistance in *Mycobacterium tuberculosis* effectively, as increasingly resistant strains of TB continue to spread around the world. In conjunction with this study, I have used two different sequencing technologies to detect rifampin resistance in *M. tuberculosis* by sequencing an 81-bp region of the *rpoB* gene, where previous research has shown that an overwhelming majority of rifampin-resistant strains of tuberculosis have mutations that are positively correlated with rifampin resistance. I used the conventional chain-termination method (commonly known as the Sanger method) to examine the genotypes of 16 different strains of *M. tuberculosis*. The second method employed was pyrosequencing. This relatively new technology is distinct from the Sanger method, as it is dependent on detecting the release of pyrophosphate during the incorporation of nucleotides into DNA.

Introduction

Tuberculosis (shortened TB or MTB) is a common infectious disease that results in many deaths annually. It is estimated that a third of the global population has been infected with TB at some point in their lives (3). In the struggle to fight this disease, the most commonly used drugs are rifampin, a bactericidal antibiotic, and isoniazid (11). The strains of *M. tuberculosis* that are resistant to rifampin and isoniazid are commonly referred to as MDR (multidrug-resistant) strains. The worldwide spread of these MDR strains threatens to make this highly penetrative infectious disease incurable (7).

Recent advances in cutting the costs of sequencing technology have facilitated whole-genome sequencing on a wider scale than was practically possible in the past. This has resulted in a greater understanding of the correlation between genotype and phenotype in MDR strains. Whole-genome sequencing has been vital to the localization of compensatory mutations in MDR strains, which negate the association between resistance mutations and reduced fitness, giving rise to high fitness strains of *M. tuberculosis* that are also multidrug-resistant (5).

The *rpoB* gene is found in many species of bacteria. It codes for the β subunit of RNA polymerase, which catalyzes RNA synthesis. The RNA polymerase enzyme is targeted by rifampin, which kills the bacteria by binding to the β subunit of RNA polymerase (4). Previous studies have shown that an overwhelming majority (96%) of *M. tuberculosis* strains with resistance to rifampin have a mutation in an 81-bp region of the *rpoB* gene that alters the sequence of a 27-amino-acid (507 through 533) region of the β subunit of RNA polymerase (10).

The aim of this study is to sequence this 81-bp region in 16 strains of *M. tuberculosis,* in order to test the reliability of predicting rifampin resistance in contemporary strains by detecting mutations in this region. This is done by two different sequencing techniques.

Material and Methods

The 16 strains of *M. tuberculosis* that were sequenced as part of this study (Table S1) are from the strain collection at the Swedish Institute for Infectious Disease Control, originally having been sent from different parts of the world. The original DNA concentrations of the samples were detected using the NanoDrop spectrophotometer (Table S2) and diluted to 5 ng/µl.

Pyrosequencing

Pyrosequencing is a relatively new technology that was commercialized by Pyrosequencing AB in Uppsala, Sweden. This novel technology is based on the detection of the pyrophosphate (PPi), which accompanies the nucleotide incorporation that is catalyzed by DNA polymerase. Sequencing is made possible by the subsequent conversion of PPi to ATP by ATP sulfurylase. This ATP is picked up by luciferase, which drives the conversion of the luciferin substrate to oxyluciferin. This process emits light that is easily detectable and proportional to the number of incorporated nucleotides. Because this principle does not enable any differentiation between nucleotides, this process is divided into four sections during DNA sequencing. In the sequence, the corresponding dNTP is added sequentially for each nucleotide (2).

The desired region of *rpoB* was amplified in a Gene-Amp PCR system 2400. Two sequencing primers, F1 and F9, were used to cover the 81-bp region. Each 50 µl reaction mixture consisted of 5 µl of 10X PCR buffer (Perkin Elmer), 1 U Ampli*Taq* Gold, 1 mM Mg^{2+}, 200 nM of each primer, 100 µM of each deoxynucleoside triphosphate, and 10 ng template, as done by Jureen et al. (9). The reaction started at 95 °C, followed by 30 cycles of 95 °C for 45 s, 53 °C for 45 s and 72 °C for 45 s. Finally, it was elongated at 72 °C for 7 minutes.

The template was prepared for sequencing using the PSQ 96 SQA reagent kit 1X96 manual (obtained from Biotage AB). All resulting pyrograms were analyzed manually, and mutations were detected by contrasting the sequenced strains to the wild-type H37Rv (ATCC 25618) master sequence.

Sanger sequencing

The second method used in this study was the conventional Sanger method, also known as the chain-termination method. The results were analyzed using the CLC Main Workbench 6.5 software. Each allele that was a designated "conflict" in the program was manually interpreted.

Results

In order to locate mutations in the analyzed strains, the genotypes were aligned with the H37Rv wild-type master sequence. In the results obtained by chain-termination (Table S3), one of the strains (BTB 05-596) was omitted due to low quality readings. Similarly, the pyrosequencing results (Table S4) lack results for the second half of the 81-bp region (covered by the F9 primer) for the BTB 04-416 and BTB 05-580 strains. Aside from this minor, albeit unfortunate, lack of interpretable data, the mutations detected by pyrosequencing and the Sanger method are almost perfectly concordant. The sole discrepancy is found in the single mutation for BTB 05-595, where the results obtained by chain-termination show the mutated codon 531 as TYG (Y is a mixture of C and T; C is the wild-type allele), whereas pyrosequencing identified it as a fully mutated TTG (leucine) codon.

When comparing the genotypic results for both pyrosequencing (Table S4) and Sanger sequencing (Table S3) with the phenotypic data (Table S5), 90% (9/10) of the mutated strains are rifampin-resistant. All mutated strains have no more than one mutation, which is in accordance with the basic principle that each strain is in need of just a single resistance-associated mutation, as any additional mutations would be disadvantageous due to reduced fitness (5).

In the genotypic data obtained by Sanger sequencing (Fig. S1), mutated base pairs are distinguished by the "conflict" label. All

manual edits have been marked by lower-case letters. The dual Sanger method PCR products, which were analyzed on a 1.5% agarose gel (Fig. S2 and S3), include the 16 strains (a 17th strain was also included as a backup), a positive control (H37Rv) and a negative control (H_2O), in addition to the reference DNA ladders on each end. The pyrosequencing PCR product (Fig. S4) has the same composition.

Discussion

Sequencing mutations for rifampin resistance in the studied 81-bp region continues to have high prediction rates for rifampin resistance. This highlights the continued importance of being able to sequence for mutations in rifampin-resistant strains of *M. tuberculosis* reliably.

The widespread distribution of MDR strains of *M. tuberculosis* has made a swift identification of rifampin resistance all the more important. It also has implications for the importance of the reliability and affordability of different sequencing technologies. The results obtained in this study may help shed light on the contrasts between different sequencing technologies and their accuracy for detecting rifampin resistance.

There are several advantages associated with detecting rifampin resistance by pyrosequencing, including reduced costs (1), as well as an optimal sequencing length for sequencing the relevant 81-bp region of the *rpoB* gene. This explains the popularity of this technology in previous work on the detection rifampin resistance (6, 8, 9, 12, 13). In addition, pyrosequencing appears to produce results that are fairly accurate for achieving this purpose, considering the near perfect concordance between the results of both methods in this study. The disadvantage of the shorter read length that is associated with pyrosequencing did not negatively affect the results because of the small size of the studied region in the *rpoB* gene.

However, it is important to note that the correlation between mutations in the studied 81-bp region of *rpoB* is not perfect, and a single mutation (GAC516TAC) in BTB 05-598, originally from Russia, was found in a rifampin-susceptible strain in the results obtained by Sanger sequencing and pyrosequencing.

Acknowledgments

My utmost thanks go to Pontus Juréen, PhD, for his guidance, time, and eagerness to teach. I would also like to thank Andreas Bråve, Head of the Unit of Highly Pathogenic Microorganisms (Department of Preparedness) at the Swedish Institute for Communicable Disease Control (SMI) for granting me permission to conduct this research at SMI.

References

1. Agah A, Aghajan M, Mashayekhi F, Amini S, Davis R. W, Plummer J. D, Ronaghi M, Griffin P. B. 2004. A multi-enzyme model for Pyrosequencing. Nucleic Acids Res. 32, e166.
2. Ahmadian A, Ehn M, Hober S. 2006. Pyrosequencing: history, biochemistry and future. Clin Chim Acta. 363:83–94.
3. Anthony RM, Schuitema ARJ, Bergval IL, Brown TJ, Oskam L, et al. 2005. Acquisition of rifabutin resistance by a rifampicin resistant mutant of *Mycobacterium tuberculosis* involves an unusual spectrum of mutations and elevated frequency. Ann Clin Microbiol Antimicrob. 4:9.
4. Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, et al. 2001. *Structural mechanism for rifampicin inhibition of bacterial RNA polymerase*. Cell 104: 901–912.
5. Comas I et al. 2005. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. Nature Genetics 44, 106-110.
6. Fakhrai-Rad H, Pourmand N, Ronaghi M. 2002. Pyrosequencing: an accurate

detection platform for single nucleotide polymorphisms. Hum. Mutat. 19:479-485.

7. Gandhi N.R. et al. 2010. Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. Lancet 375, 1830-1843.
8. Isola D. et al. 2005. A Pyrosequencing assay for rapid detection of SNPs in *Mycobacterium tuberculosis* emB306 region. J. Microbiol. Methods 62:113-120.
9. Jureen P, Engstrand L, Eriksson S, Alderborn A, Krabbe M, Hoffner SE. 2006. Rapid detection of rifampin resistance in *Mycobacterium tuberculosis* by Pyrosequencing technology. J Clin Microbiol. 44:1925–1929.
10. Ramaswamy S, Musser J. M. 1998. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. Tuber. Lung Dis. 79**:**3-29.
11. Sandgren A, Strong M, Muthukrishnan P, Weiner B. K, Church G. M, Murray M. B. 2009. Tuberculosis drug resistance mutation database. PLoS Med. 6**:**e2.
12. Tuohy M. J. et al. 2005. Pyrosequencing as a tool for the identification of common isolates of *Mycobacterium* sp. Diagn. Microbiol. Infect. Dis. 51:245-250.
13. Zhao, J. R, Bai Y. J, Zhang Q. H, Wang Y, Luo M, Yan X. J. 2005. Pyrosequencing-based approach for rapid detection of rifampin-resistant *Mycobacterium tuberculosis*. Diagn. Microbiol. Infect. Dis. 51:135-137.

Supporting Information

Figure S1 – Sanger sequencing genotypic data

Figure S2 – Sanger method, PCR

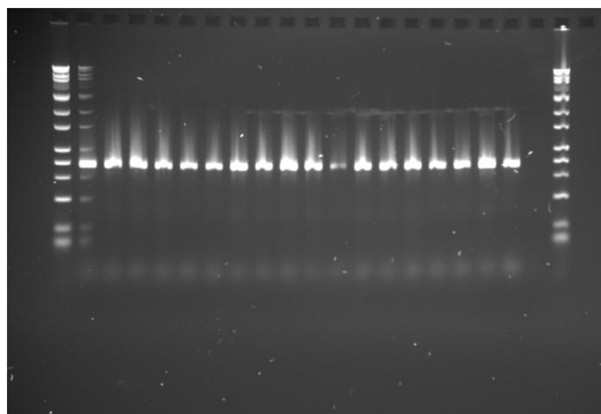

Figure S3 – Sanger method, second PCR

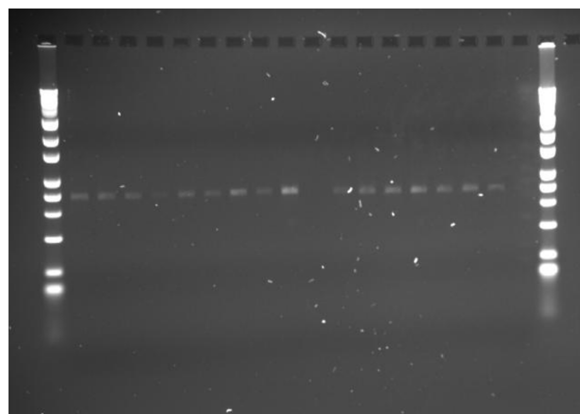

Figure S4 – Pyrosequencing PCR

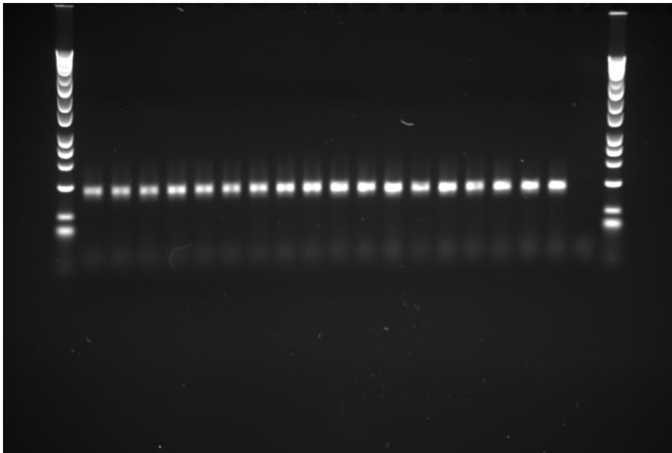

Table S1 – Strains included in study

Strains		Details
BTB 02-069		Estonia
BTB 03-157		KS
BTB 04-007		KUS
BTB 04-370		Spain
BTB 04-416		Göteborg
BTB 05-580		Armenia
BTB 05-585		Peru
BTB 05-593		Bangladesh
BTB 05-595		Bangladesh
BTB 05-596		Benin
BTB 05-597		Peru
BTB 05-598		Russia
BTB 05-650		KUS
BTB 06-041		Argentina
BTB 06-049		Argentina
BTB 06-159		Germany

Table S2 – NanoDrop data

Strains		260/280	Conc. (ng/µl)
BTB 02-069		1.39	37.7
BTB 03-157		1.55	74.6
BTB 04-007		1.79	59.0
BTB 04-370		1.76	257.7
BTB 04-416		2.60	323.1
BTB 05-580		1.50	273.0
BTB 05-585		1.83	396.0
BTB 05-593		1.62	69.6
BTB 05-595		1.90	79.7
BTB 05-596		1.81	123.0
BTB 05-597		2.32	303.28
BTB 05-598		1.70	48.5
BTB 05-650		1.80	65.6
BTB 06-041		2.26	154.8
BTB 06-049		2.77	584.1
BTB 06-159		1.99	411.0

Table S3 – Sanger method mutations

Strains		Forward	Reverse
BTB 02-069		TCG531TTG	TCG531TTG
BTB 03-157		Wild-type	Wild-type
BTB 04-007		Wild-type	Wild-type
BTB 04-370		Wild-type	Wild-type
BTB 04-416		Wild-type	Wild-type
BTB 05-580		TCG531TTG	TCG531TTG
BTB 05-585		GAC516GTC	GAC516GTC
BTB 05-593		TCG531TTG	TCG531TTG
BTB 05-595		TCG531TYG	TCG531TYG
BTB 05-596		N/A	N/A
BTB 05-597		TCG531TGG	TCG531TGG
BTB 05-598		GAC516TAC	GAC516TAC
BTB 05-650		TCG531TTG	TCG531TTG
BTB 06-041		TCG531TTG	TCG531TTG
BTB 06-049		Wild-type	Wild-type
BTB 06-159		TCG531TTG	TCG531TTG

Table S4 – Pyrosequencing mutations

Strains		F1	F9
BTB 02-069		Wild-type	TCG531TTG
BTB 03-157		Wild-type	Wild-type
BTB 04-007		Wild-type	Wild-type
BTB 04-370		Wild-type	Wild-type
BTB 04-416		Wild-type	N/A
BTB 05-580		Wild-type	N/A
BTB 05-585		GAC516GTC	Wild-type
BTB 05-593		Wild-type	TCG531TTG
BTB 05-595		Wild-type	TCG531TTG
BTB 05-596		Wild-type	Wild-type
BTB 05-597		Wild-type	TCG531TGG
BTB 05-598		GAC516TAC	Wild-type
BTB 05-650		Wild-type	TCG531TTG
BTB 06-041		Wild-type	TCG531TTG
BTB 06-049		Wild-type	Wild-type
BTB 06-159		Wild-type	TCG531TTG

Table S5 – Phenotypic data

Strains		Rifampin
BTB 02-069		Resistant
BTB 03-157		Susceptible
BTB 04-007		Resistant
BTB 04-370		Susceptible
BTB 04-416		Resistant
BTB 05-580		Resistant
BTB 05-585		Resistant
BTB 05-593		Resistant
BTB 05-595		Resistant
BTB 05-596		Resistant
BTB 05-597		Resistant
BTB 05-598		Susceptible
BTB 05-650		Resistant
BTB 06-041		Resistant
BTB 06-049		Resistant
BTB 06-159		Resistant