medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

A rapid pharmacogenomic assay to detect *NAT2* polymorphisms and guide isoniazid dosing for tuberculosis treatment

Renu Verma[1], Sunita Patil[1], Nan Zhang[2], Flora M.F. Moreira[3], Marize T. Vitorio[3], Andrea da S. Santos[3], Ellen Wallace[4], Devasena Gnanashanmugam[4], David Persing[4], Rada Savic[2], Julio Croda[5,6], Jason R. Andrews[1*]

1. Division of Infectious Diseases and Geographic Medicine, Stanford University School of Medicine, Stanford, CA, USA
2. Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA, USA
3. Federal University of Grande Dourados, Dourados, Brazil
4. Cepheid Inc., Sunnyvale, California, USA
5. Postgraduate Program in Infectious and Parasitic Diseases, Federal University of Mato Grosso do Sul, Mato Grosso do Sul, Brazil
6. Oswaldo Cruz Foundation Mato Grosso do Sul, Mato Grosso do Sul, Brazil

Correspondence*

Jason Andrews, MD

Division of Infectious Diseases and Geographic Medicine

Biomedical Innovations Building, Room 3458

Stanford University School of Medicine

Stanford, CA 94305

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Email: jandr@stanford.edu

Phone: +1 650 497 2679

Running title: A rapid pharmacogenomic assay for isoniazid dosing

Summary: This manuscript describes the development and validation of point-of-care multiplex pharmacogenomic assay to guide personalized dosing of isoniazid for treatment or prevention of tuberculosis.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

Rationale: Standardized weight-based dose of anti-tubercular drugs contributes to a substantial incidence of toxicities, inadequate treatment response, and relapse, in part due to variable drug levels achieved. Single nucleotide polymorphisms (SNPs) in the N-acetyltransferase-2 *(NAT2)* gene explain the majority of interindividual pharmacokinetic variability of isoniazid (INH). However, an obstacle to implementing pharmacogenomic-guided dosing is the lack of a point-of-care assay.

Objectives: To develop and test a *NAT2* classification algorithm, validate its performance in predicting isoniazid clearance, and develop a prototype pharmacogenomic assay.

Methods: We trained random forest models to predict *NAT2* acetylation genotype from unphased SNP data using a global collection of 8,561 phased genomes. We enrolled 48 pulmonary TB patients, performed sparse pharmacokinetic sampling, and tested the acetylator prediction algorithm accuracy against estimated INH clearance. We then developed a cartridge-based multiplex qPCR assay on the GeneXpert platform and assessed its analytical sensitivity on whole blood samples from healthy individuals.

Measurements and Main Results: With a 5-SNP model trained on two-thirds of the data (n=5,738), out-of-sample acetylation genotype prediction accuracy on the remaining third (n=2,823) was 100%. Among the 48 TB patients, predicted acetylator types were: 27 (56.2%) slow, 16 (33.3%) intermediate and 5 (10.4%) rapid. INH clearance rates were lowest in predicted slow acetylators (median 19.3 L/hr), moderate in intermediate acetylators (median 41.0 L/hr) and highest in fast acetylators (median 46.7 L/hr). The cartridge-based assay accurately detected all allele patterns directly from 25ul of whole blood.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Conclusions: An automated pharmacogenomic assay on a platform widely used globally for tuberculosis diagnosis could enable personalized dosing of isoniazid.

Keywords: tuberculosis; isoniazid; pharmacogenomic; molecular diagnostic; *NAT2*

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

INTRODUCTION

Despite the availability of effective chemotherapeutic regimens for treatment and prevention of tuberculosis, a substantial proportion of patients experience toxicities, fail treatment or develop recurrent disease (1-3). Standardized, weight-based dosing of anti-tuberculosis treatment has been the conventional approach to therapy, despite mounting evidence that inter-individual variability in metabolism leads to highly variable drug levels (4,5). High drug levels are strongly associated with risk of toxicity, while low drug levels are a determinant of treatment failure, slow response, and emergence of drug resistance. Hepatotoxicity is the most common adverse effect, affecting up to 33% of patients receiving standard four-drug therapy (6) and leading to regimen changes in up to 10% of patients (7). This toxicity is associated with increased costs, morbidity, and occasional mortality, particularly among HIV co-infected individuals (8). Additionally, as many as 3% of new tuberculosis cases experience treatment failure, and between 6-10% relapse within 2 years (9,10). Pharmacokinetic variability to a single drug is associated with treatment failure and acquired drug resistance (11,12). One study found that individuals with at least one drug below the recommended AUC threshold had a 14-fold increased risk of poor outcomes (13)

More than 50 years after its introduction, isoniazid (INH) remains one of the major first line drugs used to treat active and latent tuberculosis infections (14,15). There has been an increasing number of genetic markers identified that predict metabolism and toxicities from various antimicrobials. INH is among the most well characterized of these, with more than 80% of its pharmacokinetic variability explained by mutations in the gene encoding arylamine N-acetyltransferase 2 (NAT2), responsible for its metabolism in the liver (16-18). The primary step in the metabolism of INH is acetylation, catalyzed by the NAT2 enzyme, resulting in the

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

formation of acetyl-INH. The NAT2 enzyme displays genetic polymorphism, and its activity is expressed at highly variable levels. A high correlation between INH acetylator phenotype and seven most frequent SNPs in *NAT2* gene has been demonstrated by several studies (19-22). Individuals can be classified into three phenotypes—rapid, intermediate, and slow acetylators—according to whether they carry polymorphisms on neither, one, or both copies of this gene, respectively. Rapid acetylators typically have the lowest plasma INH concentrations, while slow acetylators have high concentrations (23). A worldwide population survey on *NAT2* acetylation phenotype reported that more than half of the global population are slow or rapid acetylators (18). A meta-analysis of 14 studies, comprising 474 cases and 1446 controls based on *NAT2* polymorphisms found that rapid acetylators are twice as likely to have microbiological failure and acquired drug resistance. Additionally, a significant association has been consistently observed between *NAT2* slow acetylators and the risk of anti-tuberculosis drug-induced liver injury (17). Additional meta-analyses have identified a three- to four-fold increased risk of hepatotoxicity among slow acetylators (24). A randomized trial of pharmacogenomic guided dosing for tuberculosis treatment found that, compared with standard dosing, it reduced hepatotoxicity among slow acetylators and increased treatment response at 8 weeks among rapid acetylators (25)

Despite this evidence, pharmacogenomic testing and guided treatment has not entered the mainstream of clinical practice for tuberculosis. Few clinical laboratories perform *NAT2* genotyping, which requires detection of multiple polymorphisms and testing for heterozygous allele patterns. Such testing is not widely available in resource-constrained environments where the majority of tuberculosis burden falls. To address this gap, we developed an algorithm from

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

unphased SNP patterns, derived from globally representative genomic data, to predict INH metabolism phenotype using fewer SNPs while retaining high accuracy. Based on the SNP combination derived from this model, we further developed a prototype *NAT2* pharmacogenomic (NAT2-PGx) assay on a commercial, automated PCR platform (GeneXpert) to detect *NAT2* polymorphisms. We demonstrate that this tool can accurately predict INH clearance rates directly from clinical samples and can be easily performed with minimal training and hands-on time.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

METHODS

Datasets

The datasets used to develop the NAT2 classifier was obtained from the IGSR (International Genome Sample Resource, 1000 genomes project) and a meta-analysis by Sabbagh et. al (18, 26). Population information on the combined dataset is provided in **Supplementary Table 1.**

NAT2 acetylator phenotype prediction classifier

Phased genomes from 8,561 individuals were used and haplotypes were labeled based on seven most frequent SNPs (both synonymous and non-synonymous) which are reported to affect the acetylation on INH - rs1801279 (191G>A), rs1041983 (282C>T), rs1801280 (341T>C), rs1799929 (481C>T), rs1799930 (590G>A), rs1208 (803A>G) and rs1799931 (857G>A) (19) in the *NAT2* gene, following an international consensus nomenclature (27). We then constructed an unphased dataset containing only information on whether each sample was wild type for both alleles, homozygous variant for both alleles, or heterozygous. We stratified the dataset by geographic region and then drew a random sample from two-thirds of each stratum for a training set and one third for an out-of-sample test set, to ensure geographic representativeness in the training and test sets. We trained random forest models on the training set using the *caret* package in R and assessed classification performance on the held-out test set. We began with a 7 SNP model and eliminated SNPs in sequential models according to the lowest variable importance factor.

Ethics statement

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The clinical study was approved by the institutional review boards of the Stanford University School of Medicine and Federal University of Grande Dourados. All participants were over the age of 18 and provided written informed consent. For assay optimization and validation on GeneXpert cartridge, anonymized blood samples from healthy individuals were obtained from Stanford blood center.

Sample collection

Sputum and plasma samples from 48 newly diagnosed patients with active pulmonary tuberculosis were collected at the Federal University of Grande Dourados, Brazil. All participants were treated with standardized, weight-based doses of isoniazid, rifampicin, pyrazinamide and ethambutol. Plasma samples were collected at 1 hour and 8 hours after the first dose and after 1 hour on day 14. Plasma drug concentrations for isoniazid and acetyl-isoniazid were quantified by high-performance liquid chromatography coupled to tandem mass spectrometry (HPLC-MS) as previously described (28).

Pharmacokinetic analysis of INH clearance in tuberculosis patients

The population PK analysis was performed using the non-linear mixed effects modeling approach using NONMEM (version 7.4.3; ICON plc, Gaithersburg, MD, USA), PsN and R-based Xpose (version 4.7 and higher) (29,30). One-compartment model with a first-order absorption with a lognormal distribution for inter-individual variability (IIV) on different PK parameter(s) as well as an additive and/or proportional model for the residual error were tested for the base model selection. Mixture models with two or three subpopulations representing different clearance rate were then evaluated. The first-order conditional estimation with

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

interaction method (FOCEI) was applied and the model-building procedure and model selection was based on the log-likelihood criterion (the difference in the minimum OFV between hierarchical models was assumed to be Chi-square distributed with degrees of freedom equal to the difference in the number of parameters between models), goodness-of-fit plots (e.g. relevant residuals against time randomly distributed around zero), and scientific plausibility of the model. Visual predictive check was conducted to evaluate whether the final model with estimated fixed-effect parameters and covariates adequately describe data.

Sputum processing and host DNA extraction

Spontaneously expectorated sputum samples from confirmed pulmonary tuberculosis patients were collected in 10mL of guanidine thiocyanate (GTC) solution. The samples were needle sheared and centrifuged at 3000 rpm for 30 min. The supernatant was collected in Trizol LS and host DNA was extracted from the supernatant using a manual extraction method described previously (31).

Primers and probes for melt curve analysis

We first developed single-plex melt curve assays based on molecular beacon probe chemistry for five NAT2 polymorphisms. Using same primer and probe sequences, we further developed multiplex NAT2-PGx assay compatible on GeneXpert platform. Three sets of primers and five molecular beacon probes spanning the *NAT2* gene were used. Primers and probes were designed using Beacon Designer (Premier Biosoft International, Palo, CA; version 8). A list of primers and probes sequences with their corresponding fluorophores and quenchers is provided in **Supplementary Table 2.**

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

NAT2-genotyping on pulmonary TB samples

Host genomic DNA extracted from 48 sputum samples from TB positive patients was used to perform single-plex qPCR assays developed in-house. The assays were validated using TaqMan commercial genotyping assays (NAT2 TaqMan® SNP Genotyping Assays, Applied Biosystems). The genotyping was performed on StepOne Plus Real Time PCR (Applied Biosystems). The NAT2-genotypes derived from the assays were used to predict INH acetylator phenotype by our 5-SNP model.

Automated NAT2-PGx Multiplex PCR and melt curve analysis

We combined the five single-plex NAT2 melt curve assays validated on TB samples into one and developed a multiplex assay (NAT2-PGx) on the FleXible Cartridge (Flex cart-01, Cepheid) platform. Flex cart allows automated DNA extraction from whole blood followed by PCR amplification and melt curve analysis to detect SNPs in *NAT2* gene in a single run. NAT2-PGx assay was performed on a GeneXpert IV instrument using GeneXpert Dx 4.8 software (Cepheid, Sunnyvale). Briefly, 100ul of whole blood was mixed with 900ul of GeneXpert lysis buffer for whole blood (Cepheid), incubated for 2 min at room temperature and loaded into sample preparation chamber of the flex cart for automated DNA extraction. A 70ul of PCR mastermix was simultaneously loaded in the PCR reaction chamber of the flex cart-01. PCR and melt conditions were optimized using mastermix prepared in house.

We validated the multiplex assay on blood samples from 20 healthy individuals. The accuracy of multiplex NAT2-PGx assay in SNP calling was subsequently validated by Sanger sequencing (see methods). We assessed analytical sensitivity of the assay and robustness to input blood volume by performing it on varying volumes of whole blood (200 ul, 100 ul, 50 ul, 25 ul) and

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

comparing the Tm results and standard deviation for each position across blood volumes. Additional details on the qPCR run method and melt curve analysis is provided in an online data supplement.

RESULTS

SNP selection and development of acetylation prediction model

Complete phased data for the seven polymorphisms that define acetylation haplotypes were available for 8,561 individuals from 59 populations. The dataset contains 3,573 (41.7%) individuals with a slow genotype, 3,428 (40.0%) individuals with an intermediate genotype, and 1,560 (18.2%) individuals with a rapid genotype (See **Table 1**). The highest proportion of rapid acetylators were in East Asia (40%), and three regions had prevalence of slow acetylator phenotypes over 50% (Central and South Asia, Europe and North Africa). We used these phased allele data to select SNPs for inclusion in an assay measuring unphased SNPs. Using a random forest model trained on two thirds of the data (n=5,738), out-of-sample phenotype prediction accuracy from unphased data on the remaining one third (n=2,823) was 100% for models using 7, 6 or 5 SNPs. With 4 SNPs, prediction accuracy was 98.0% (95% CI: 97.4-98.5%), and a 3 SNP model had similar performance (98.0%; 95% CI: 97.4-98.4%) (**Table 2**). However, both of these models performed poorly on data from Sub-Saharan Africa (4 SNP model accuracy: 82.5%, 95% CI: 78.1-86.4%); 3 SNP model accuracy: 81.3%, 95% CI: 76.8-85.3%). Based on these results, we selected the 5 SNP model (191G>A, 282C>T, 341T>C, 590G>A and 857G>A) to take forward for clinical validation and diagnostic development.

Genotype correlation with isoniazid clearance in patients with tuberculosis

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

We enrolled a cohort of 48 patients with newly diagnosed pulmonary tuberculosis and collected plasma at 1 hour and 8 hours after dose on day 1 and at 1 hour after dose on day 14. To detect five *NAT2* polymorphisms identified by our classifier, we performed single-plex melt curve qPCR assays developed in-house using host DNA extracted from sputum samples. Additionally, we used commercial 7-SNP single-plex genotyping assays and compared the results with 5-SNP single-plex PCR to validate the melt curve accuracy in SNP detection. There was 100% concordance in terms of SNP detection between single-plex melt curve and commercial 7-SNP assays. Of the 48 individuals for whom *NAT2* genotypes were profiled, DNA-probe hybrid melting temperature difference (ΔTm) (°C) between wild-type and mutant alleles for positions 191, 282, 341, 590 and 857 were found to be 4.38, 4.04, 2.40, 3.63 and 3.68 respectively. Both mutant and wild type probes had a minimum 2.40°C Tm difference which allowed SNP calling with high accuracy **(Table 3).**

We further predicted phenotypes from 5-SNP using the algorithm described above as well as a publicly available tool (NAT2Pred) (32), which uses a 6 SNP model that excludes rs1801279 (191G>A) and includes other two sites from 7-SNP model- rs1799929 (481C>T) and rs1208 (803A>G) when compared with our 5-SNP model. Among the 48 participants, predicted acetylator types from the 5 SNP assay were: 27 (56.2%) slow, 16 (33.3%) intermediate and 5 (10.4%) rapid. NAT2Pred classified 4 samples as intermediate that were classified as rapid (n=1) or slow (n=3) by the 5 SNP classifier. Among those classified as slow by the 5 SNP classifier and intermediate by NAT2Pred, acetyl-INH to INH ratios at 8 hours were 0.61, 0.38, 0.41, consistent with slow acetylation (median: 0.76, range 0.36-1.55) rather than intermediate acetylation (median 6.67, range 3.32-22.21) and suggesting misclassification by NAT2Pred. The

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

sample classified as intermediate by NAT2PRed and rapid by the 5 SNP classifier had an acetyl-INH to INH ratio of 9.8, which fell between the median values, and within both ranges, for intermediate and rapid acetylators (range 8.09 - ∞) (**Supplementary Table-3**). Phenotypes predicted by the 5 SNP classifier were strongly predictive of INH acetylation and clearance (**Figure 1a and 1b**). INH clearance rates were lowest in slow acetylators (median 19.3 L/hr), moderate in intermediate acetylators (median 41.0 L/hr) and highest in fast acetylators (median 46.7 L/hr).

Development of an automated pharmacogenomic assay

We validated the NAT2-PGx assay on 20 whole blood samples from healthy individuals. The hands-on time for the cartridge preparation per sample was 5 minutes followed by 140 minutes for overall run that included automated DNA extraction (**Figure 2**). Mutant, wild-type and heterozygous alleles were manually called based on peak patterns and Tm values detected in melt curves. Negative derivative transformed melt curves from five *NAT2* gene polymorphisms are shown in **Figure 3**. The assay detected all polymorphisms with 100% accuracy (average SD in Tm across all probes = 0.34°C) compared with Sanger sequencing. The *NAT2* genotypes corresponding to 20 blood samples covered all three categories - mutant, wild-type and heterozygous for five *NAT2* positions except for NAT2-191 for which all samples were all wild-type. The 191G>A mutation is highly prevalent in African and African-American populations and is less common in other populations (33). Among the 20 samples, predicted acetylator types using the 5-SNP classifier were: 8 (40%) slow, 10 (50%) intermediate and 2 (10%) rapid (**Supplementary Table 4a**).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

We further assessed the analytic performance of the NAT2-PGx assay at lower sample volumes. Our assay could accurately detect all melt peaks with as low as 25ul of sample volum**e**. The variability in Tm from five *NAT2* probes for sample volumes 200ul-25ul is shown in **Figure 4**. *NAT2* polymorphisms were accurately detected at all volumes (**Supplementary Table 4b)** and demonstrated robustness to variation in input volume**.**

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

DISCUSSION

Despite availability of effective treatment for drug-sensitive tuberculosis, a substantial proportion of population encounters drug associated toxicity or treatment failure, much of which could be averted through dosing guided by genetic markers of drug metabolism (34). We previously found that pharmacogenomic guided dosing of isoniazid could be highly cost-effective in low- and middle-income countries (35). A major barrier to its implementation has been the lack of a simple, scalable assay that could be used at points of care where tuberculosis is treated in resource-constrained settings. To address this gap, we used globally representative genomic data to identify patterns of 5 SNPs that enable accurate prediction of isoniazid acetylator phenotype, validating this with pharmacokinetic data of patients receiving tuberculosis treatment. We then developed a prototype automated pharmacogenomic assay on the GeneXpert platform, which is widely available globally but had never been applied to pharmacogenomics. We found that this assay could robustly distinguish wild type, mutant and heterozygous alleles from a range of blood volumes as low as 25 ul, making it suitable for use with venous blood samples or finger-stick blood samples. The assay requires minimal hands-on time for sample preparation, which would facilitate its use in resource-constrained settings.

An earlier model ("NAT2Pred") predicted NAT2 acetylation phenotype from unphased genomic data; however, it had moderate error rates in distinguishing intermediate from rapid acetylators (33). Moreover, error rates among individuals from Sub-Saharan Africa were 14%, in part due to the exclusion of the G191A (R64Q) SNP, common to the NAT2 *14 allele cluster, which is frequent in Africans and African-Americans, but virtually absent in Caucasian, Indian, and Korean populations (36). A recent study using the set of 6 NAT2 SNPs included in NAT2Pred,

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

but not including G191A, found no correlation between *NAT2* genotype and INH acetylation phenotype in HIV-infected, Zulu individuals with culture-confirmed tuberculosis in Durban, South Africa (37), underscoring the importance of including this SNP in genotype predictions in this region.

Indeed, ethnic differences in SNP frequencies are responsible for the differences in frequency of rapid, intermediate and slow acetylator *NAT2* haplotypes (18, 23). We trained our SNP classifier with globally representative data, which resulted in the selection and inclusion of the G191A SNP in our model and assay. This is particularly important as Sub-Saharan Africa bears a substantial burden of tuberculosis disease and mortality as well as HIV co-infection, which is independently associated with greater pharmacokinetic variability and tuberculosis treatment toxicity (38, 39).

The association between acetylation polymorphisms and INH metabolism was first demonstrated in 1959, and their importance was well characterized in subsequent decades through phenotypic descriptions (40-42). Subsequent genotypic descriptions confirmed that *NAT2* polymorphisms predicted INH early bactericidal activity, and clinical outcomes including hepatotoxicity, relapse and acquisition of drug resistance. Further dosing studies demonstrated that provision of lower doses to slow acetylators and higher doses to rapid acetylators could achieve target concentrations (43). One randomized trial of pharmacogenomic-guided dosing of INH during active tuberculosis treatment found that it significantly reduced toxicities (among slow acetylators) and treatment non-response (among rapid acetylators). Taken together, the evidence for pharmacogenomic guided dosing to achieve consistent drug levels and improve clinical

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

outcomes is strong. Automated, easy-to-use assays could enable pharmacogenomic guided isoniazid dosing in resource constrained settings, where a substantial burden of the world's tuberculosis occurs.

The findings of this study are subject to several limitations. We tested the assays on 48 individuals with active tuberculosis and 20 healthy individuals with a diverse representation of polymorphisms, but the number of participants with G191A mutations was limited (n=5). A larger validation study involving testing on whole blood, including from finger stick capillary blood, is needed to assess real-world performance of this assay under field conditions. Due to unavailability of whole blood samples from TB patients, we used sputum samples to extract host DNA for genotyping. Further studies should also investigate testing on non-invasive samples including sputum, saliva or oral swabs, from which DNA is abundant. Second, we focused on *NAT2* polymorphisms, as they explain majority of interindividual pharmacokinetic variability, though polymorphisms in several other genes have been associated with hepatotoxicity. However, these associations have been comparatively modest and somewhat inconsistent (44-45). We focused on INH and did not include other important tuberculosis drugs, such as rifampicin. The evidence base for pharmacogenomic markers predicting rifampin pharmacokinetics is less robust, and findings concerning clinical outcomes such as toxicities or treatment response are limited (46-49). However, given the importance of this drug class in treatment of active and latent tuberculosis, and emerging evidence supporting greater efficacy of higher doses of rifampin, further investigation of pharmacogenomic markers in rifampicin is needed. Future assays may include polymorphisms influencing rifampicin metabolism to further optimize treatment of tuberculosis.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Since the demonstration of the efficacy of six month, short-course chemotherapy in 1979, standardized treatment for drug susceptible tuberculosis using weight-based doses has remained essentially unchanged. Additionally, INH remains a major component of regimens for treatment of latent tuberculosis, which is recommended by the WHO for young children, HIV-infected individuals and household contacts of tuberculosis cases (50). More than half the world's population have slow or rapid acetylation phenotypes, which put them at risk for excessive drug levels resulting in drug toxicities or insufficient drug levels putting them at risk of acquired drug resistance or disease relapse. Dose adjustment based on NAT2 acetylation genotyping can achieve consistent, target drug levels and reduce the incidence of poor clinical outcomes. We developed a prototype automated, cartridge-based assay that can reliably predict acetylation phenotype directly from as low as 25 ul of whole blood. By developing this for the GeneXpert platform, which is widely used in low- and middle-income countries for tuberculosis diagnosis, this assay could make personalized tuberculosis treatment dosing available in resource-constrained settings. Further studies are needed to evaluate its accuracy and clinical impact in real-world clinical settings.

Funding

FleXcartridges and technical support for their use were provided by Cepheid. This study was supported by the Institute for Immunity, Transplantation, and Infection and the Department of Medicine at Stanford University.

Data availability

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Data supporting the findings of this manuscript are available in the Supplementary Information files or from the corresponding author upon request.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Acknowledgements

We thank Veronique Dartois for performing the plasma drug level assays.

Author contributions

JRA and RV conceived of the study. RV, JC, and JRA designed the experiments. RV, SP, FMM, MTV, ASS and JC collected data. RV, NZ, RS and JRA analyzed data. EW, DG and DP provided technical guidance on the assay development. RV and JRA wrote the first draft of the manuscript, and all authors contributed to the final version.

Conflicts of Interest

JRA and RV are named co-inventors on a provisional patent (Application number 62/991,477) for a *NAT2* pharmacogenomic assay.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

References

1. Gillespie SH, Crook AM, McHugh TD, Mendel CM, Meredith SK, Murray SR, Pappas F, Phillips PP, Nunn AJ; REMoxTB Consortium. Four-month moxifloxacin-based regimens for drug-sensitive tuberculosis. N Engl J Med. 2014 Oct 23;371(17):1577-87. doi: 10.1056/NEJMoa1407426. Epub 2014 Sep 7.
2. Merle CS, Fielding K, Sow OB, Gninafon M, Lo MB, Mthiyane T, Odhiambo J, Amukoye E, Bah B, Kassa F, N'Diaye A, Rustomjee R, de Jong BC, Horton J, Perronne C, Sismanidis C, Lapujade O, Olliaro PL, Lienhardt C; OFLOTUB/Gatifloxacin for Tuberculosis Project. A four-month gatifloxacin-containing regimen for treating tuberculosis. N Engl J Med. 2014 Oct 23;371(17):1588-98. doi: 10.1056/NEJMoa1315817. Erratum in: N Engl J Med. 2015 Apr 23;372(17):1677.
3. Jindani A, Harrison TS, Nunn AJ, Phillips PP, Churchyard GJ, Charalambous S, Hatherill M, Geldenhuys H, McIlleron HM, Zvada SP, Mungofa S, Shah NA, Zizhou S, Magweta L, Shepherd J, Nyirenda S, van Dijk JH, Clouting HE, Coleman D, Bateson AL, McHugh TD, Butcher PD, Mitchison DA; RIFAQUIN Trial Team. High-dose rifapentine with moxifloxacin for pulmonary tuberculosis. N Engl J Med. 2014 Oct 23;371(17):1599-608.
4. Pratt VM, Everts RE, Aggarwal P, Beyer BN, Broeckel U, Epstein-Baak R, Hujsak P, Kornreich R, Liao J, Lorier R, Scott SA, Smith CH, Toji LH, Turner A, Kalman LV. Characterization of 137 Genomic DNA Reference Materials for 28 Pharmacogenetic Genes: A GeT-RM Collaborative Project. J Mol Diagn. 2016 Jan;18(1):109-23.
5. García-Martín E. Interethnic and intraethnic variability of NAT2 single nucleotide polymorphisms. Curr Drug Metab. 2008 Jul;9(6):487-97.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

6. Schaberg T, Rebhan K, Lode H. Risk factors for side-effects of isoniazid, rifampin and pyrazinamide in patients hospitalized for pulmonary tuberculosis. Eur Respir J. 1996 Oct;9(10):2026-30.
7. Saukkonen JJ, Cohn DL, Jasmer RM, Schenker S, Jereb JA, Nolan CM, Peloquin CA, Gordin FM, Nunes D, Strader DB, Bernardo J, Venkataramanan R, Sterling TR; ATS (American Thoracic Society) Hepatotoxicity of Antituberculosis Therapy Subcommittee. An official ATS statement: hepatotoxicity of antituberculosis therapy. Am J Respir Crit Care Med. 2006 Oct 15;174(8):935-52.
8. Schutz C, Ismail Z, Proxenos CJ, Marais S, Burton R, Kenyon C, Maartens G, Wilkinson RJ, Meintjes G. Burden of antituberculosis and antiretroviral drug-induced liver injury at a secondary hospital in South Africa. S Afr Med J. 2012 Mar 2;102(6):506-11.
9. Luzze H, Johnson DF, Dickman K, Mayanja-Kizza H, Okwera A, Eisenach K, Cave MD, Whalen CC, Johnson JL, Boom WH, Joloba M; Tuberculosis Research Unit. Relapse more common than reinfection in recurrent tuberculosis 1-2 years post treatment in urban Uganda. Int J Tuberc Lung Dis. 2013 Mar;17(3):361-7.
10. Romanowski K, Balshaw RF, Benedetti A, Campbell JR, Menzies D, Ahmad Khan F, Johnston JC. Predicting tuberculosis relapse in patients treated with the standard 6-month regimen: an individual patient data meta-analysis. Thorax. 2019 Mar;74(3):291-297.
11. Pasipanodya JG, Srivastava S, Gumbo T. Meta-analysis of clinical studies supports the pharmacokinetic variability hypothesis for acquired drug resistance and failure of antituberculosis therapy. Clin Infect Dis. 2012 Jul;55(2):169-77.
12. Gumbo T, Louie A, Liu W, Brown D, Ambrose PG, Bhavnani SM, Drusano GL. Isoniazid bactericidal activity and resistance emergence: integrating pharmacodynamics and

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

pharmacogenomics to predict efficacy in different ethnic populations. Antimicrob Agents Chemother. 2007 Jul;51(7):2329-36.

13. Pasipanodya JG, McIlleron H, Burger A, Wash PA, Smith P, Gumbo T. Serum drug concentrations predictive of pulmonary tuberculosis outcomes. J Infect Dis. 2013 Nov 1;208(9):1464-73.
14. Hall RG, Leff RD, Gumbo T. Treatment of active pulmonary tuberculosis in adults: current standards and recent advances. Insights from the Society of Infectious Diseases Pharmacists. Pharmacotherapy. 2009 Dec;29(12):1468-81
15. Sterling TR, Njie G, Zenner D, et al. Guidelines for the Treatment of Latent Tuberculosis Infection: Recommendations from the National Tuberculosis Controllers Association and CDC, 2020. MMWR Recomm Rep 2020;69(No. RR-1):1–11.
16. Kinzig-Schippers M, Tomalik-Scharte D, Jetter A, Scheidel B, Jakob V, Rodamer M, Cascorbi I, Doroshyenko O, Sörgel F, Fuhr U. Should we use N-acetyltransferase type 2 genotyping to personalize isoniazid doses? Antimicrob Agents Chemother. 2005 May;49(5):1733-8.
17. Du H, Chen X, Fang Y, Yan O, Xu H, Li L, Li W, Huang W. Slow N-acetyltransferase 2 genotype contributes to anti-tuberculosis drug-induced hepatotoxicity: a meta-analysis. Mol Biol Rep. 2013 May;40
18. Sabbagh A, Langaney A, Darlu P, Gérard N, Krishnamoorthy R, Poloni ES. Worldwide distribution of NAT2 diversity: implications for NAT2 evolutionary history. BMC Genet. 2008 Feb 27;9:21.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

19. Wang PY, Xie SY, Hao Q, Zhang C, Jiang BF. NAT2 polymorphisms and susceptibility to anti-tuberculosis drug-induced liver injury: a meta-analysis. Int J Tuberc Lung Dis. 2012 May;16(5):589-95.
20. Parkin DP, Vandenplas S, Botha FJ, Vandenplas ML, Seifart HI, van Helden PD, van der Walt BJ, Donald PR, van Jaarsveld PP. Trimodality of isoniazid elimination: phenotype and genotype in patients with tuberculosis. Am J Respir Crit Care Med. 1997 May;155(5):1717-22.
21. Gross M, Kruisselbrink T, Anderson K, Lang N, McGovern P, Delongchamp R, Kadlubar F. Distribution and concordance of N-acetyltransferase genotype and phenotype in an American population. Cancer Epidemiol Biomarkers Prev. 1999 Aug;8(8):683-92.
22. Mrozikiewicz PM, Cascorbi I, Brockmöller J, Roots I. Determination and allelic allocation of seven nucleotide transitions within the arylamine N-acetyltransferase gene in the Polish population. Clin Pharmacol Ther. 1996 Apr;59(4):376-82.
23. Hein DW, Doll MA. Accuracy of various human NAT2 SNP genotyping panels to infer rapid, intermediate and slow acetylator phenotypes. Pharmacogenomics. 2012 Jan;13(1):31-41.
24. Snider DE Jr, Caras GJ. Isoniazid-associated hepatitis deaths: a review of available information. Am Rev Respir Dis. 1992 Feb;145(2 Pt 1):494-7.
25. Azuma J, Ohno M, Kubota R, Yokota S, Nagai T, Tsuyuguchi K, Okuda Y, Takashima T, Kamimura S, Fujio Y, Kawase I; Pharmacogenetics-based tuberculosis therapy research group. NAT2 genotype guided regimen reduces isoniazid-induced liver injury and early treatment failure in the 6-month four-drug standard treatment of tuberculosis: a randomized

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

controlled trial for pharmacogenetics-based therapy. Eur J Clin Pharmacol. 2013 May;69(5):1091-101.

26. McInnes G and Russ A. Drug Response Pharmacogenetics for 200,000 UK Biobank Participants.
27. *Database of arylamine N-acetyltransferases* (http://nat.mbg.duth.gr)
28. Prideaux B, Via LE, Zimmerman MD, Eum S, Sarathy J, O'Brien P, Chen C, Kaya F, Weiner DM, Chen PY, Song T, Lee M, Shim TS, Cho JS, Kim W, Cho SN, Olivier KN, Barry CE 3rd, Dartois V. The association between sterilizing activity and drug distribution into tuberculosis lesions. Nat Med. 2015 Oct;21(10):1223-7.
29. Beal SL. Et al. NONMEM 7.4 Users Guides. (1989– 2019). ICON plc, Gaithersburg, MD. https://nonmem.iconplc.com/nonmem744
30. Lindbom L, Ribbing J, Jonsson EN. Perl-speaks-NONMEM (PsN)--a Perl module for NONMEM related programming. Comput Methods Programs Biomed. 2004 Aug;75(2):85-94.
31. Chomczynski P. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. Biotechniques. 1993 Sep;15(3):532-4, 536-7. PMID: 7692896.
32. Kuznetsov IB, McDuffie M, Moslehi R. A web server for inferring the human N-acetyltransferase-2 (NAT2) enzymatic phenotype from NAT2 genotype. Bioinformatics. 2009 May 1;25(9):1185-6.
33. Sabbagh A, Darlu P, Vidaud M. Evaluating NAT2PRED for inferring the individual acetylation status from unphased genotype data. BMC Med Genet. 2009 Dec 31;10:148.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

34. Ramachandran G, Swaminathan S. Role of pharmacogenomics in the treatment of tuberculosis: a review. Pharmgenomics Pers Med. 2012;5:89-98

35. Rens NE, Uyl-de Groot CA, Goldhaber-Fiebert JD, Croda J, Andrews JR. Cost-effectiveness of a Pharmacogenomic Test for Stratified Isoniazid Dosing in Treatment of Active Tuberculosis. Clin Infect Dis. 2020 Dec 15;71(12):3136-3143.

36. Sabbagh A, Darlu P, Crouau-Roy B, Poloni ES. Arylamine N-acetyltransferase 2 (NAT2) genetic diversity and traditional subsistence: a worldwide population survey. PLoS One. 2011 Apr 6;6(4):e18507.

37. Mthiyane T, Millard J, Adamson J, Balakrishna Y, Connolly C, Owen A, Rustomjee R, Dheda K, McIlleron H, Pym AS. *N*-Acetyltransferase 2 Genotypes among Zulu-Speaking South Africans and Isoniazid and *N*-Acetyl-Isoniazid Pharmacokinetics during Antituberculosis Treatment. Antimicrob Agents Chemother. 2020 Mar 24;64(4):e02376-19.

38. van Oosterhout JJ, Dzinjalamala FK, Dimba A, Waterhouse D, Davies G, Zijlstra EE, Molyneux ME, Molyneux EM, Ward S. Pharmacokinetics of Antituberculosis Drugs in HIV-Positive and HIV-Negative Adults in Malawi. Antimicrob Agents Chemother. 2015 Oct;59(10):6175-80.

39. Wilkins JJ, Langdon G, McIlleron H, Pillai G, Smith PJ, Simonsson US. Variability in the population pharmacokinetics of isoniazid in South African tuberculosis patients. Br J Clin Pharmacol. 2011 Jul;72(1):51-62.

40. Iselius L, Evans DA. Formal genetics of isoniazid metabolism in man. Clin Pharmacokinet. 1983 Nov-Dec;8(6):541-4.

41. Evans DA. Genetic variations in the acetylation of isoniazid and other drugs. Ann N Y Acad Sci. 1968 Jul 31;151(2):723-33.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

42. Evans DA. An improved and simplified method of detecting the acetylator phenotype. J Med Genet. 1969 Dec;6(4):405-7.

43. Naidoo A, Chirehwa M, Ramsuran V, McIlleron H, Naidoo K, Yende-Zuma N, Singh R, Ncgapu S, Adamson J, Govender K, Denti P, Padayatchi N. Effects of genetic variability on rifampicin and isoniazid pharmacokinetics in South African patients with recurrent tuberculosis. Pharmacogenomics. 2019 Mar;20(4):225-240.

44. Wilkins JJ, Langdon G, McIlleron H, Pillai GC, Smith PJ, Simonsson US. Variability in the population pharmacokinetics of pyrazinamide in South African tuberculosis patients. Eur J Clin Pharmacol. 2006 Sep;62(9):727-35.

45. Lee SW, Chung LS, Huang HH, Chuang TY, Liou YH, Wu LS. NAT2 and CYP2E1 polymorphisms and susceptibility to first-line anti-tuberculosis drug-induced hepatitis. Int J Tuberc Lung Dis. 2010 May;14(5):622-6.

46. Roy B, Ghosh SK, Sutradhar D, Sikdar N, Mazumder S, Barman S. Predisposition of antituberculosis drug induced hepatotoxicity by cytochrome P450 2E1 genotype and haplotype in pediatric patients. J Gastroenterol Hepatol. 2006 Apr;21(4):784-6.

47. Simon T, Becquemont L, Mary-Krause M, de Waziers I, Beaune P, Funck-Brentano C, Jaillon P. Combined glutathione-S-transferase M1 and T1 genetic polymorphism and tacrine hepatotoxicity. Clin Pharmacol Ther. 2000 Apr;67(4):432-7.

48. Strange RC, Jones PW, Fryer AA. Glutathione S-transferase: genetics and role in toxicology. Toxicol Lett. 2000 Mar 15;112-113:357-63.

49. Mukonzo JK, Kengo A, Kutesa B, Nanzigu S, Pohanka A, McHugh TD, Zumla A, Aklillu E. Role of pharmacogenetics in rifampicin pharmacokinetics and the potential effect on TB-

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

rifampicin sensitivity among Ugandan patients. Trans R Soc Trop Med Hyg. 2020 Feb 7;114(2):107-114.

50. Haley CA. Treatment of Latent Tuberculosis Infection. Microbiol Spectr. 2017 Apr;5(2).

Figure 1. Predicted *NAT2* phenotype from sputum samples and associated acetylation ratio and isoniazid clearance rates from patients receiving tuberculosis treatment. The (a) 8 hour acetyl-INH to INH ratio and (b) isonazid clearance rates, according to acetylation phenotype predicted from 5 SNPs, measured in sputum samples from 48 patients receiving treatment for active tuberculosis.

Figure 2. Schemata for the automated NAT2 Pharmacogenomic assay. 1-2 drops of blood is collected in an Eppendorf tube and mixed with lysis buffer to a total of 1 ml, which is then loaded onto a GeneXpert Flex01 cartridge and placed into a GeneXpert instrument for automated DNA extraction, qPCR and meltcurve analysis. Allele patterns for each of the 5 SNPs are determined by Tm analysis, and the resulting data are used to predict acetylator phenotype.

Figure 3. Negative derivative transformed melt curves for the five NAT2 gene polymorphisms. The shift in melt curve temperature is observed during a nucleotide exchange. Molecular beacon probes are first hybridized and then melted off of their *NAT2* target amplicon. The melt curves indicate wild type alleles at positions 191(red), 341(green) and 857 (orange); and mutant alleles at positions 282(Blue) and 590(Purple).

Figure 4. Effect of whole blood sample volume on melting temperature for wild type and mutant alleles at 5 positions in *NAT2*: *NAT2* polymorphisms were accurately detected at all

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

volumes with sufficient different in melting temperature (Tm) to distinguish wild type from mutant alleles. None of the individuals in this dataset had mutations at position 191.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 1. Summary of populations included in genomic analysis and their acetylation genotypes.

		Acetylation Genotype, n (%)		
Region	**Number of individuals**	**Slow**	**Intermediate**	**Rapid**
Americas	1,112	432 (39%)	463 (42%)	217 (20%)
Central and South Asia	588	355 (60%)	198 (34%)	35 (6%)
East Asia	2,308	340 (15%)	1049 (45%)	919 (40%)
Europe	3,458	1966 (57%)	1249 (36%)	243 (7%)
North Africa	44	30 (68%)	10 (23%)	4 (9%)
sub-Saharan Africa	1,051	450 (43%)	459 (44%)	142 (14%)
Total	8,561	3573 (42%)	3428 (40%)	1560 (18%)

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 2. Out-of-sample prediction accuracy of unphased NAT2 SNP data for acetylation phenotype in random forest models. Models were trained with 5,738 individuals and tested on 2,823 individuals. Sens: sensitivity. Spec: specificity.

NAT2 SNP Positions	SNP number	Accuracy	95% CI	Sens. Rapid	Spec. Rapid	Sens. Slow	Spec Slow
191, 282, 341, 481, 590, 803, 857	7	1.000	(0.999-1.000)	1.000	1.000	1.000	1.000
191, 282, 341, 481, 590, 857	6	1.000	(0.999-1.000)	1.000	1.000	1.000	1.000
191, 282, 341, 590, 857	5	1.000	(0.999-1.000)	1.000	1.000	1.000	1.000
282, 341, 590, 857	4	0.978	(0.972-0.983)	0.996	0.988	0.969	0.999
341, 590, 847	3	0.976	(0.970-0.982)	1.000	0.986	0.967	1.000
341, 590	2	0.852	(0.838-0.865)	1.000	0.889	0.832	1.000

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 3. Melting temperature (Tm) values for five *NAT2* polymorphisms derived from DNA-probe hybrid melts using single-plex assays validated on 48 pulmonary TB patients

	Mutant	**Wild type**	**Het**	**Mutant**	**Wild type**	
NAT2 SNP position	Total samples analyzed	Total samples analyzed	Total samples analyzed	Tm ± SD	Tm ± SD	ΔTm (WT-MT)
191	0	44	5	66.3 ± 0.28	61.92 ± 0.68	4.38
282	6	24	19	62.81 ± 0.40	66.85 ± 0.34	4.04
341	12	23	14	66.30± 0.48	68.70 ± 0.38	2.40
590	0	35	14	68.0 ± 0.28	64.37 ± 0.28	3.63
857	1	41	7	66.4± 0.11	62.72 ± 0.29	3.68

SD: Standard deviation, ΔTm: DNA-probe hybrid melting temperature difference, WT: Wild type, MT: mutant

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Supplementary Table 1. Population information on the dataset used to develop the 5-SNP classifier.

Supplementary Table 2. List of primers and probes sequences for multiplex assay with their corresponding fluorophores and quenchers

Supplementary Table 3. *NAT2* polymorphism profiles of 48 pulmonary TB patients detected from 5-SNP and 7-SNP genotyping assays. Acetylator phenotypes predicted from 5-SNP, 6-SNP and 7-SNP classifiers and plasma INH and acetyl-INH levels recorded at various time-points.

Supplementary Table 4. (a) NAT2 genotype of whole blood samples (n=20) analyzed on 5-plex NAT2-PGx assay and validated on Sanger sequencing. **(b)** NAT2 genotypes of ten whole blood samples analyzed using 200ul, 100ul, 50ul and 25ul samples

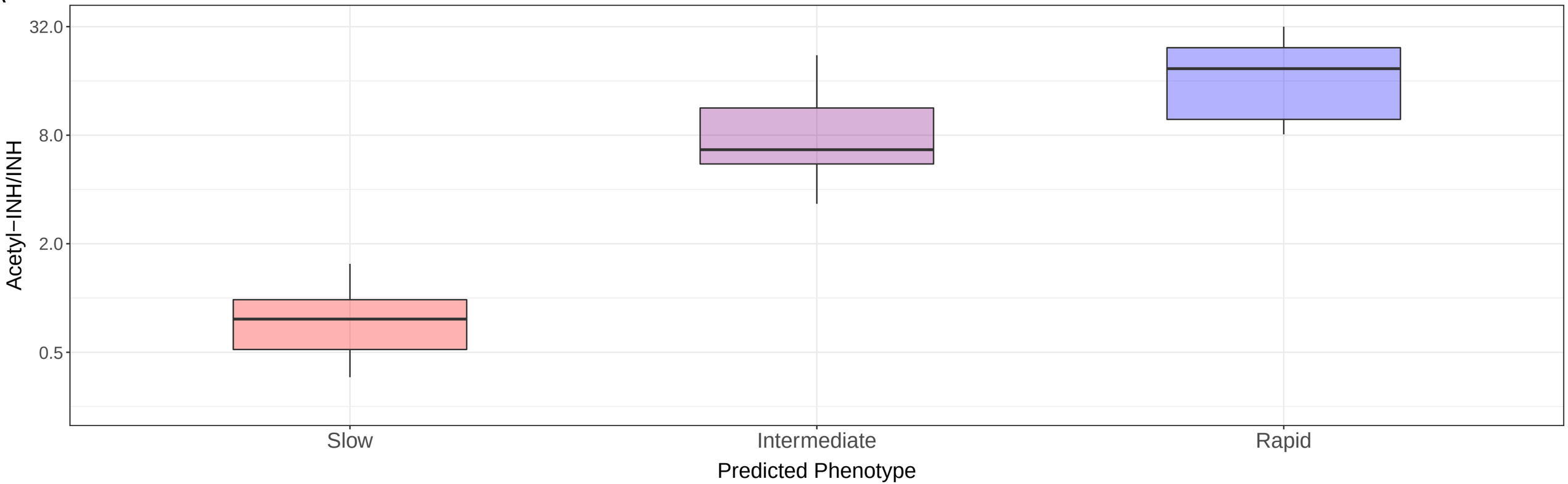
medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

At a Glance Commentary

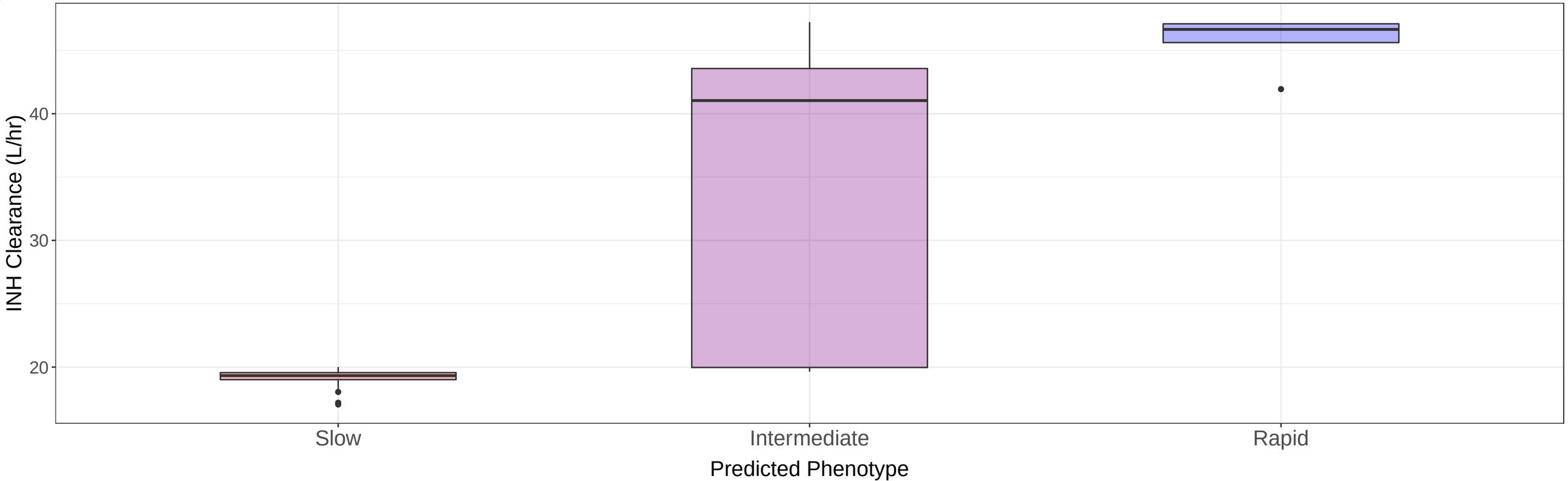
Scientific Knowledge on the Subject: Standardized weight-based dose of isoniazid (INH) may result in variable drug levels in individuals leading to drug-induced toxicities, treatment failure and relapse. Polymorphisms in the *NAT2* gene explain the majority of interindividual pharmacokinetic variability of INH. Pharmacogenomic guided dosing may improve treatment outcomes, but scalable assays are not available.

What This Study Adds to the Field: Earlier *NAT2* acetylation prediction algorithms had diminished accuracy in African populations by failing to include a common polymorphism. We developed an algorithm on globally representative data that predicted acetylation genotype with 100% accuracy using 5 unphased SNPs. We tested the accuracy of this acetylator prediction algorithm against INH clearance rates in 48 pulmonary TB patients and found very strong high correlation. We then developed a cartridge-based multiplex assay for these 5 SNPs on the GeneXpert platform that can accurately detect *NAT2* genotype from 25ul whole blood. Because this platform is widely used in low- and middle-income countries for tuberculosis diagnosis, this assay could make personalized isoniazid dosing available in resource-constrained settings.

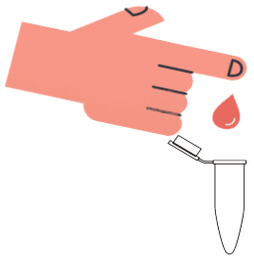
A


32.0
8.0
2.0
0.5
Acetyl−INH/INH
Slow
Intermediate
Rapid
Predicted Phenotype

B

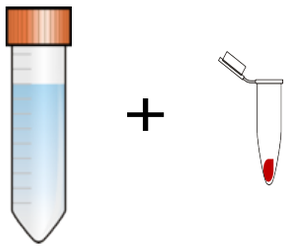

40
30
20
INH Clearance (L/hr)
Slow
Intermediate
Rapid
Predicted Phenotype

① Sample collection



Collect 1-2 drops of whole blood (approx. 50ul)

② Whole blood pretreatment



Mix 950ul lysis buffer, vortex for 2 sec and incubate at RT for 2min

③ Flex cart01 preparation

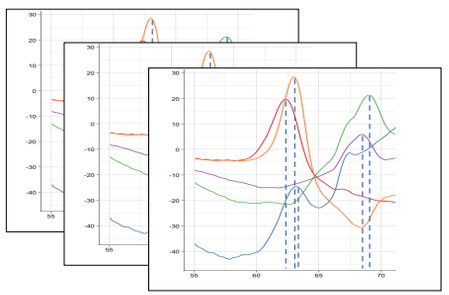
Pipette 1ml pretreated sample and 70ul PCR mastermix into the cartridge

④ Automated DNA extraction & qPCR

Total run time= 140min

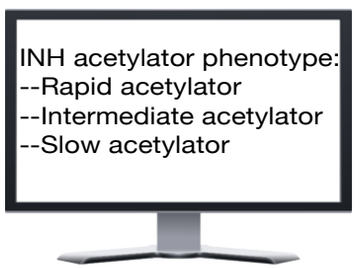
Load the cartridge in the GeneXpert machine and start the run

⑤ Data aquisition



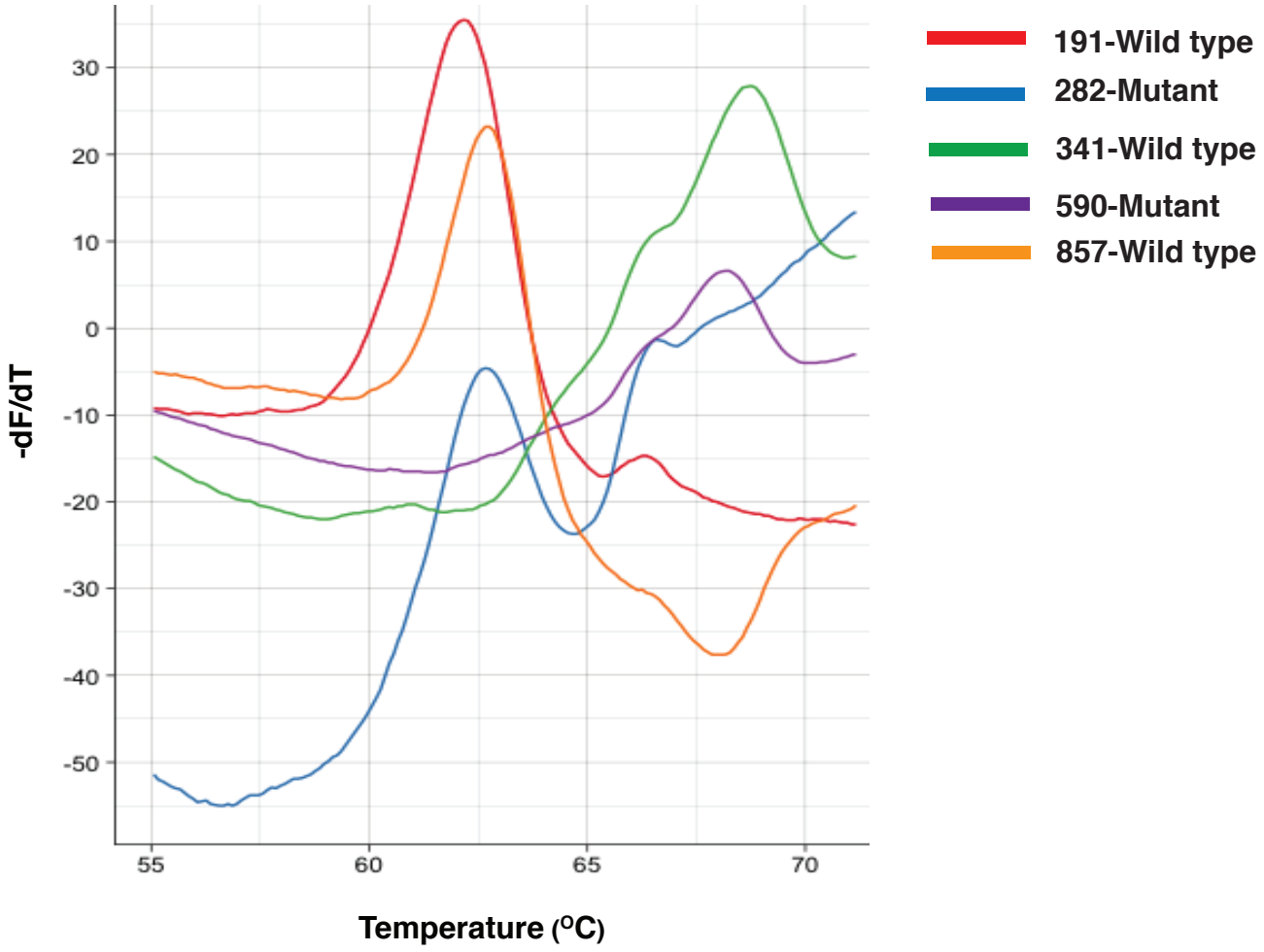
Acquire melt curve data for NAT2 gene polymorphism detection

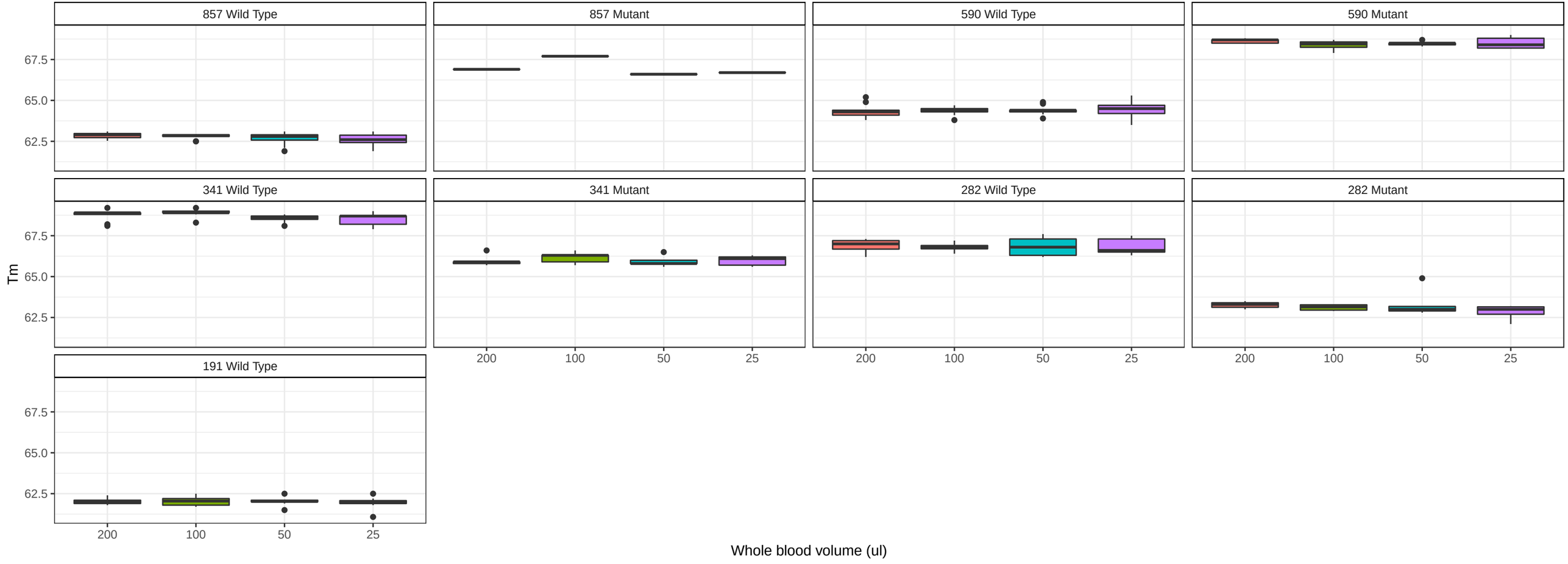
⑥ INH acetylator phenotype prediction



Upload NAT2 polymorphism data on phenotype prediction algorithm

medRxiv preprint doi: https://doi.org/10.1101/2021.01.17.21249995; this version posted March 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.




857 Wild Type
857 Mutant
590 Wild Type
590 Mutant
341 Wild Type
341 Mutant
282 Wild Type
282 Mutant
191 Wild Type
Tm
67.5
65.0
62.5
200
100
50
25
Whole blood volume (ul)