RESEARCH ARTICLE

Mycobacterium tuberculosis Pyrazinamide Resistance Determinants: a Multicenter Study

Paolo Miotto,[a] Andrea M. Cabibbe,[a] Silke Feuerriegel,[b,c] Nicola Casali,[d,e] Francis Drobniewski,[d,e] Yulia Rodionova,[f] Daiva Bakonyte,[g] Petras Stakenas,[g] Edita Pimkina,[h] Ewa Augustynowicz-Kopeć,[i] Massimo Degano,[j] Alessandro Ambrosi,[k] Sven Hoffner,[l] Mikael Mansjö,[l] Jim Werngren,[l] Sabine Rüsch-Gerdes,[m] Stefan Niemann,[b,c] Daniela M. Cirillo[a]

Emerging Bacterial Pathogens Unit, Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, Milan, Italy[a]; Molecular Mycobacteriology,[b] German Centre for Infection Research (DZIF),[c] Borstel, Germany; Infectious Diseases, Imperial College, London, United Kingdom[d]; Clinical TB and HIV Group, Public Health England (PHE) National Mycobacterium Reference Laboratory, Blizard Institute, Queen Mary University of London, London, United Kingdom[e]; Samara TB Service, Samara, Russian Federation[f]; Department of Immunology and Cell Biology, Institute of Biotechnology, Vilnius University, Vilnius, Lithuania[g]; National Tuberculosis Reference Laboratory, Infectious Diseases and Tuberculosis Hospital, Vilnius University Hospital Santariskiu Klinikos, Vilnius, Lithuania[h]; Department of Microbiology, National Tuberculosis and Lung Diseases Research Institute, Warsaw, Poland[i]; Biocrystallography Unit, Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, Milan Italy[j]; Vita-Salute San Raffaele University, Milan, Italy[k]; Public Health Agency of Sweden, Solna, and Department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institute, Stockholm, Sweden[l]; National Reference Center for Mycobacteria, Research Center Borstel, Borstel, Germany[m]

S.N. and D.M.C. contributed equally to this work.

ABSTRACT **Pyrazinamide (PZA) is a prodrug that is converted to pyrazinoic acid by the enzyme pyrazinamidase, encoded by the *pncA* gene in *Mycobacterium tuberculosis*. Molecular identification of mutations in *pncA* offers the potential for rapid detection of pyrazinamide resistance (PZA^r). However, the genetic variants are highly variable and scattered over the full length of *pncA*, complicating the development of a molecular test. We performed a large multicenter study assessing *pncA* sequence variations in 1,950 clinical isolates, including 1,142 multidrug-resistant (MDR) strains and 483 fully susceptible strains. The results of *pncA* sequencing were correlated with phenotype, enzymatic activity, and structural and phylogenetic data. We identified 280 genetic variants which were divided into four classes: (i) very high confidence resistance mutations that were found only in PZA^r strains (85%), (ii) high-confidence resistance mutations found in more than 70% of PZA^r strains, (iii) mutations with an unclear role found in less than 70% of PZA^r strains, and (iv) mutations not associated with phenotypic resistance (10%). Any future molecular diagnostic assay should be able to target and identify at least the very high and high-confidence genetic variant markers of PZA^r; the diagnostic accuracy of such an assay would be in the range of 89.5 to 98.8%.**

IMPORTANCE **Conventional phenotypic testing for pyrazinamide resistance in *Mycobacterium tuberculosis* is technically challenging and often unreliable. The development of a molecular assay for detecting pyrazinamide resistance would be a breakthrough, directly overcoming both the limitations of conventional testing and its related biosafety issues. Although the main mechanism of pyrazinamide resistance involves mutations inactivating the *pncA* enzyme, the highly diverse genetic variants scattered over the full length of the *pncA* gene and the lack of a reliable phenotypic gold standard hamper the development of molecular diagnostic assays. By analyzing a large number of strains collected worldwide, we have classified the different genetic variants based on their predictive value for resistance which should lead to more rapid diagnostic tests. This would assist clinicians in improving treatment regimens for patients.**

Received 20 August 2014 **Accepted** 26 September 2014 **Published** 21 October 2014

Citation Miotto P, Cabibbe AM, Feuerriegel S, Casali N, Drobniewski F, Rodionova Y, Bakonyte D, Stakenas P, Pimkina E, Augustynowicz-Kopeć E, Degano M, Ambrosi A, Hoffner S, Mansjö M, Werngren J, Rüsch-Gerdes S, Niemann S, Cirillo DM. 2014. *Mycobacterium tuberculosis* pyrazinamide resistance determinants: a multicenter study. mBio 5(5): e01819-14. doi:10.1128/mBio.01819-14.

Editor Carol A. Nacy, Sequella, Inc.

Address correspondence to Paolo Miotto, miotto.paolo@hsr.it.

Pyrazinamide (PZA) is a key drug in current and future tuberculosis (TB) treatment regimens. It has a high sterilizing capacity *in vivo*, but it is not active against *Mycobacterium tuberculosis* complex (MTBC) strains growing at neutral pH (1–4).

In addition to its crucial role in the standard short-course regimen for TB treatment, PZA is used in the treatment of patients infected with strains that are multidrug resistant (MDR) (resistant to at least isoniazid and rifampin). Here PZA has a strong impact on the success rates of MDR treatment and may allow a shortening of current MDR therapy (5). Finally, PZA is the only first-line drug that will be maintained in all regimens in the near future (6). These new regimens aim at reducing the treatment duration of susceptible, drug-resistant (especially MDR TB and extensively resistant) strain variants.

The essential role of PZA underlines the need for accurate and rapid detection of PZA resistance that is very difficult with current

phenotypic tests (7). The difficulties with culture-based PZA susceptibility testing result from several factors, including suboptimal test media with unreliable pH and larger inocula that reduce the activity of PZA (8, 9). Furthermore, the critical concentration itself may result in inconsistent results for isolates with a PZA MIC close to this concentration (10). While for isoniazid and rifampin, highly reliable culture-based drug susceptibility testing (DST) techniques and rapid molecular assays such as the line probe assay MTBDR*plus* (Hain Lifescience GmbH, Nehren, Germany) and the cartridge-based Xpert MTB/RIF assay (Cepheid, Sunnyvale, CA) are available (11), no commercial molecular assays are currently marketed for PZA.

Great efforts have been made in understanding molecular resistance mechanisms. PZA is a prodrug that needs to be converted to an active compound, pyrazinoic acid, by the bacterial pyrazinamidase (PZase) (encoded by *pncA*). Mutations/variations in *pncA* leading to the loss of PZase activity are the major mechanism leading to PZA resistance (PZA^r) (4, 12). However, while high numbers of PZA^r cases can be related to inactivation of the PZase, the genetic variants, including single nucleotide polymorphisms (SNPs) and small deletions, are highly diverse and scattered over the full length of the 561 bp of the *pncA* gene (4, 12). This complicates the development of molecular tests, as no "hot spot region" comprising the majority of mutations is present in the *pncA* gene, as is present in *rpoB* for rifampin resistance.

Accordingly, future molecular approaches to detect PZA^r in clinical isolates need to cover at least a significant number of possible variants, if not the complete gene, to reach a high sensitivity (e.g., using approaches based on classical Sanger sequencing or next-generation genome sequencing). These techniques must be combined with an appropriate interpretation algorithm/database that distinguishes SNPs clearly associated with drug resistance from those for which the impact for developing PZA^r is unclear, e.g., due to phylogenic variants found in members of the MTBC (13, 14). In-depth knowledge of the variants found in PZA^r strains combined with evidence-based correlation with resistance phenotypes are needed to develop large-scale databases ensuring valid data interpretation. The fact that such a valid data basis is currently lacking represents a substantial limitation for molecular PZA DST.

To tackle this question, we performed a large multicenter study assessing *pncA* sequence variations in 1,950 MTBC pansusceptible strains and PZA^r strains. The strains were classified in phylogenetic lineages to identify variants that are phylogenetically informative but not likely to be involved in PZA^r and those that are occurring in strains from different groups and are obviously under positive selection. Using this comprehensive approach, we could catalog 239 high-confidence PZA^r mutations and a number of *pncA* variants most likely not involved in PZA^r.

RESULTS

We studied 1,950 clinical isolates, including 1,142 MDR strains and 483 fully susceptible strains (see Table S1 in the supplemental material). By phenotypic DST, 1,107 clinical isolates were susceptible to PZA, whereas 843 were classified as PZA^r. Genotyping data were available for 1,853 isolates (95.0%). Predominant lineages among the strains investigated were Beijing (47.8%), LAM (9.0%), Ural (7.7%), and Haarlem (5.0%). Other lineages found (Ghana, EAI, Delhi/CAS, H37Rv_like, Uganda I and II, West African 1 and 2, S, Cameroon, Sierra Leone 1 and 2, *Mycobacterium bovis*, H37Rv, UZB_H37Rv_like, New-1, X, CAS, TUR,and *Mycobacterium microti* vole) each represented less than 5% of isolates (Fig. 1). Six percent were classified as EuroAmerican strains not belonging to a valid lineage described previously ("other undefined," as described in reference 15), and 5.2% of strains were classified as "unknown," because it was not possible to assign a defined lineage.

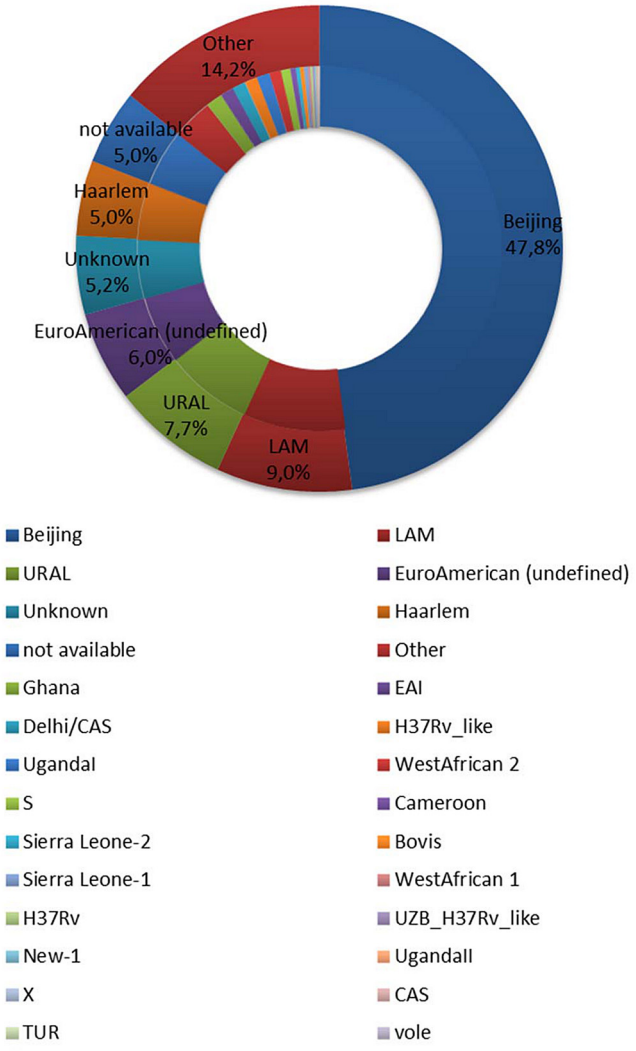

FIG 1 Pie chart reporting percentages of lineages for isolates included in the study.

Considering the sequencing results, 1,062 (54.5%) isolates were found to be wild type (WT) for the *pncA* gene, whereas 888 harbored variations in the *pncA* gene that amounted to a total of 280 genetic variants comprising 67 insertions or deletions (indels) and 213 SNPs (see Table S2 in the supplemental material). The PZase enzymatic activity was available for 251 clinical isolates accounting for 90 different genetic variants. Considering the distribution of the mutations across the entire gene, 73 (39.0%) codons were not affected by mutations, whereas the remaining 114 codons presented one or more mutations (Fig. 2). Only 50 codons showed a frequency of mutation over the mean value of 0.5%, but despite this, a clear hot spot region could not be found; the most

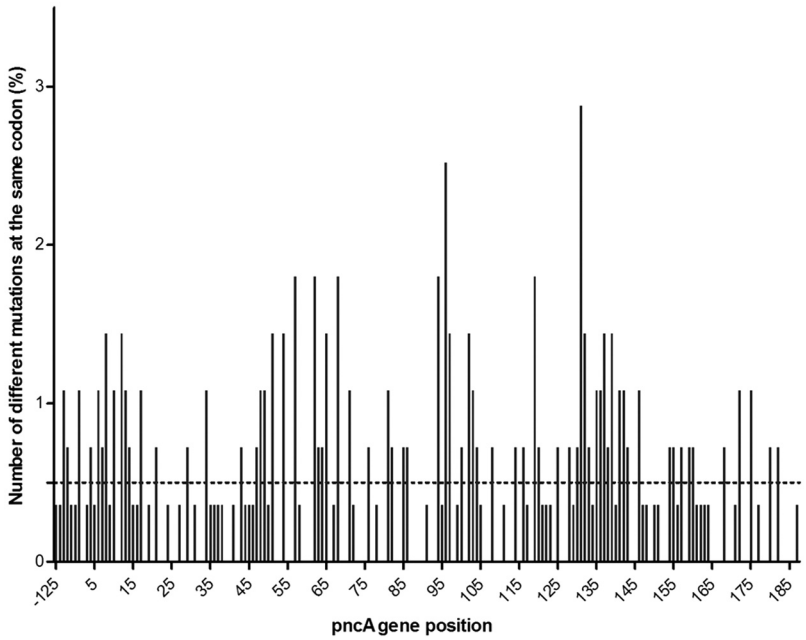

FIG 2 Number of different mutations found at each codon. Note that multiple mutations and IS*6110* are not included. The broken line indicates mean value.

frequently affected regions (representing more than 70% of mutated strains) were found at the promoter (positions −13 to −3) and at codons 6 to 15, 50 to 70, 90 to 100, 130 to 145, and 170 to 175 (Fig. 3).

For mutations found in both PZA-sensitive (PZA^s) and PZA^r isolates, enzymatic activity and structural analysis results were used to adjust for possible errors in phenotypic DST whenever possible and to obtain a "revised DST" (included as "DST rev" in Table S1 in the supplemental material). Accordingly, 56 clinical isolates originally reported as PZA^s were reclassified as PZA^r (Table S1). The final distribution of mutations among revised PZA^s and PZA^r isolates is summarized in Table 1. To further validate the classification data, we analyzed the homoplastic occurrence of particular mutations (e.g., the emergence in strains of two phylogenetic lineages [16]). As the homoplasy level is rather low in MTBC genomes, this confirms that these mutations are most likely under positive selection and involved in the development of PZA^r.

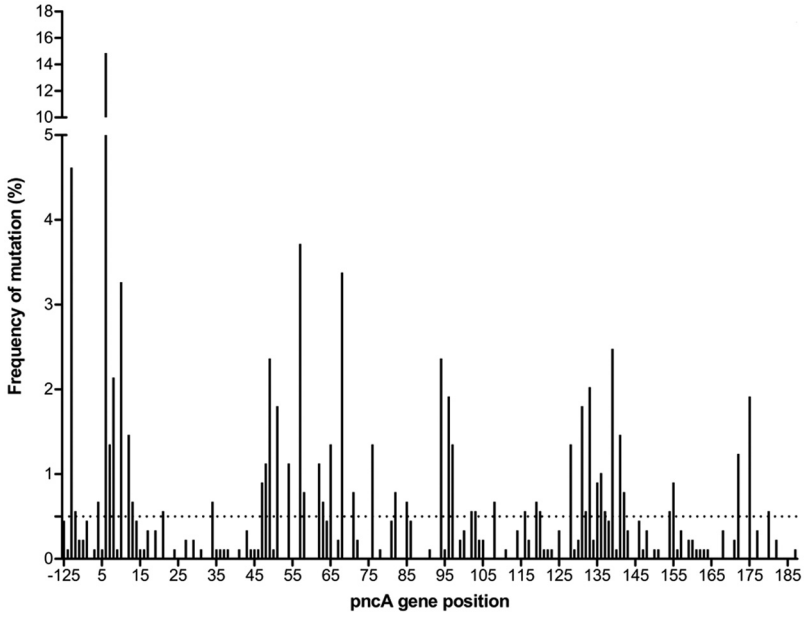

FIG 3 Frequency of mutations found at each codon (calculated with 888 mutated isolates). Note that multiple mutations and IS*6110* are not included. The dotted line indicates mean value.

TABLE 1 Distribution of mutations among PZA^s and PZA^r clinical isolates

pncA gene	No. of isolates (%)	
	PZA^s (n = 1,051)	PZA^r (n = 899)
WT	893[a] (85.0)	158 (15.0)
Mutant	200 (22.2)	699 (77.8)

[a] Includes 19 isolates harboring silent mutations or mutations at the distal region of the promoter (>100 nucleotides upstream of the start codon).

Using this procedure, four classes of genetic variants were identified: (i) very high confidence resistance mutations that were found only in PZA^r strains (category A), (ii) high-confidence resistance mutations found in more than 70% of PZA^r strains (category E), (iii) mutations with an unclear role found in less than 70% in PZA^r strains (category D), and (iv) genetic variants (including the wild type) not involved in phenotypic resistance (category B). Table S2 in the supplemental material summarizes these clinically relevant categories; a graphical overview is provided in Fig. 4.

Mutations conferring PZA^r at very high confidence. Out of the 280 sequence variants identified in *pncA*, 239 (85.4%) mutations found in 644 clinical isol ates (644/1,950 [33.0%]) were classified as very high confidence variants associated with phenotypic PZA^r (category A) (see Table S2 in the supplemental material). Several mutations affect the catalytic residues and amino acids recruited in the scaffold of the active site or directly/indirectly involved in the coordination of the Fe^{2+} ion (Asp8Gly/Ala/Glu/Asn, His51Gln/Tyr, His71Arg, Asp49Glu/Asn/Ala, His57Arg/Tyr/Gln/Pro, Trp68Arg/Gly/Cys/Stop/Leu, Gln10Pro/Arg, and His137Pro/Arg/Asp) or residues engaged in the hydrophobic core (Ile6Thr, Val44Gly, Val139Gly/Leu, Met175Thr/Val, and Phe94Cys/Ser/Leu). Out of the 90 variants tested, 87 variants, including nucleotide substitutions at position −11, were also associated with negative PZase activity, and 55 genetic variants (detected in 332 isolates) were found in strains of at least 2 different lineages, indicating homoplasy (data not shown). Table 2 reports the mutations mapping in the most frequently affected regions.

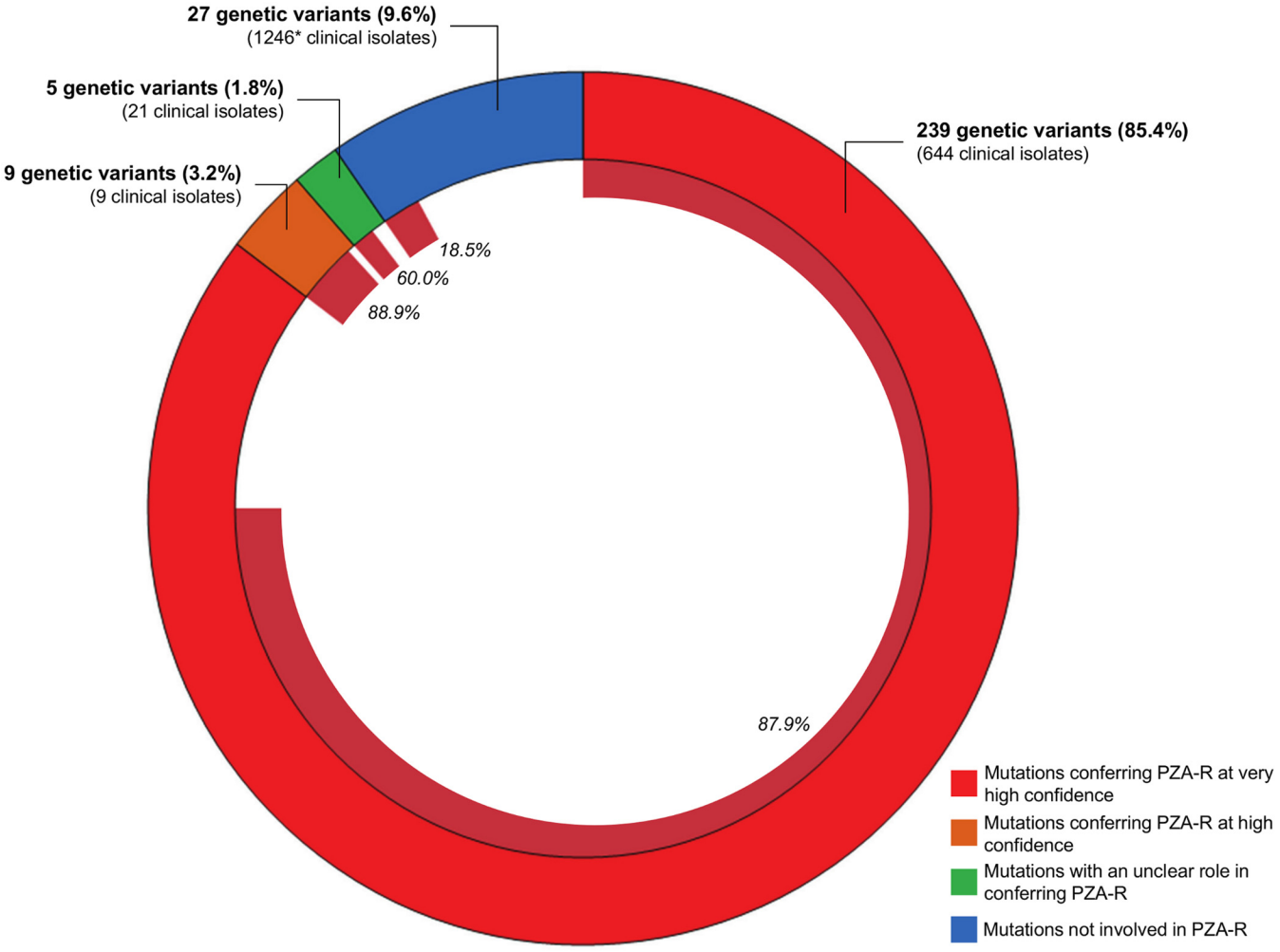

FIG 4 Distribution of genetic variants across the four categories identified: (i) very high confidence resistance mutations, (ii) high-confidence resistance mutations, (iii) mutations with an unclear role, and (iv) mutations not involved in phenotypic resistance. The number of isolates belonging to each category is also reported. The inner ring shows the percentages of mutations affecting the structure of the enzyme for each category of genetic variants. PZA-R, PZA resistance. *, including wild-type isolates for the *pncA* gene.

TABLE 2 Mutations for PZA^r affecting the most frequently affected regions of *pncA* gene and representing more than 70% of mutated cases

Nucleotide change[a]	Result of the mutation[b]	p.S[c]	p.R[d]	No. of cases
A-11C	Promoter −11	0.01497006	0.98502994	5
A-11G	Promoter −11	0.01497006	0.98502994	35
A-11T	Promoter −11	0.01497006	0.98502994	1
T-7C	Promoter −7	0.01497006	0.98502994	4
T-7G	Promoter −7	0.01497006	0.98502994	1
Del-5 → G	Promoter (del)	0.01497006	0.98502994	2
ATC6ACC	Ile6Thr	0.01497006	0.98502994	3
Del14 → TCATCG	FSC 6 (del)	0.01497006	0.98502994	1
GTC7GGC	Val7Gly	0.01497006	0.98502994	9
GTC7TTC	Val7Phe	0.01497006	0.98502994	2
WT + GTC7GGC	WT + Val7Gly	0.01497006	0.98502994	1
GAC8AAC	Asp8Asn	0.01497006	0.98502994	3
GAC8GAA	Asp8Glu	0.01497006	0.98502994	6
GAC8GCC	Asp8Ala	0.01497006	0.98502994	1
GAC8GGC	Asp8Gly	0.01497006	0.98502994	9
GTG9GGG	Val9Gly	0.01497006	0.98502994	1
CAG10AAG	Gln10Lys	0.01497006	0.98502994	2
CAG10CCG	Gln10Pro	0.01497006	0.98502994	21
CAG10CGG	Gln10Arg	0.01497006	0.98502994	6
GAC12AAC	Asp12Asn	0.01497006	0.98502994	
GAC12GAG	Asp12Glu	0.01497006	0.98502994	2
GAC12GCC	Asp12Ala	0.01497006	0.98502994	6
Ins37 → GACT	FSC 13 (ins)	0.01497006	0.98502994	1
TTC13TCC	Phe13Ser	0.01497006	0.98502994	2
TTC13TTG	Phe13Leu	0.01497006	0.98502994	3
TGC14CGC	Cys14Arg	0.01497006	0.98502994	1
TGC14TGA	Cys14Stop	0.01497006	0.98502994	2
WT + TGC14CGC	WT + Cys14Arg	0.01497006	0.98502994	1
Ins44 → C	FSC 15 (ins)	0.01497006	0.98502994	1
Del150 → T	FSC 50 (del)	0.01497006	0.98502994	1
CAC51CAA	His51Gln	0.01497006	0.98502994	7
CAC51CCC	His51Pro	0.01497006	0.98502994	2
CAC51CGC	His51Arg	0.01497006	0.98502994	4
CAC51TAC	His51Tyr	0.01497006	0.98502994	3
CCG54CAG	Pro54Gln	0.01497006	0.98502994	4
CCG54CGG	Pro54Arg	0.01497006	0.98502994	1
CCG54CTG	Pro54Leu	0.01497006	0.98502994	4
CCG54TCG	Pro54Ser	0.01497006	0.98502994	1
CAC57CAG	His57Gln	0.01497006	0.98502994	1
CAC57CCC	His57Pro	0.01497006	0.98502994	1
CAC57CGC	His57Arg	0.01497006	0.98502994	14
CAC57GAC	His57Asp	0.01497006	0.98502994	10
CAC57TAC	His57Tyr	0.01497006	0.98502994	5
WT + CAC57CGC	WT + His57Arg	0.01497006	0.98502994	2
TTC58CTC	Phe58Leu	0.01497006	0.98502994	7
CCG62CTG	Pro62Leu	0.01497006	0.98502994	3
Del186 → C	FSC 62 (del)	0.01497006	0.98502994	3
Ins185 → 4 nt	FSC 62 (ins)	0.01497006	0.98502994	1
Ins186 → A	FSC 62 (ins)	0.01497006	0.98502994	1
GAC63GGC	Asp63Gly	0.01497006	0.98502994	4
Ins192 → A	FSC 64 (ins)	0.01497006	0.98502994	1
TAT64TAG	Tyr64stop	0.01497006	0.98502994	3
Ins193 → A	FSC 65 (ins)	0.01497006	0.98502994	1
Ins193 → TATCAGG	FSC 65 (ins)	0.01497006	0.98502994	1
TCG67CCG	Ser67Pro	0.01497006	0.98502994	2
TGG68CGG	Trp68Arg	0.01497006	0.98502994	7
TGG68GGG	Trp68Gly	0.01497006	0.98502994	16
TGG68TAG	Trp68stop	0.01497006	0.98502994	1
TGG68TGC	Trp68Cys	0.01497006	0.98502994	5
TGG68TGT	Trp68Cys	0.01497006	0.98502994	1
GAG91TAG	Glu91Stop	0.01497006	0.98502994	1
TTC94CTC	Phe94Leu	0.01497006	0.98502994	8
TTC94TCC	Phe94Ser	0.01497006	0.98502994	3
TTC94TGC	Phe94Cys	0.01497006	0.98502994	6
TTC94TTA	Phe94Leu	0.01497006	0.98502994	2

(Continued on following page)

TABLE 2 (Continued)

Nucleotide change[a]	Result of the mutation[b]	p.S[c]	p.R[d]	No. of cases
TTC94TTG	Phe94Leu	0.01497006	0.98502994	1
WT + TTC94CTC	WT + Phe94Leu	0.01497006	0.98502994	1
TAC95TAG	Tyr95stop	0.01497006	0.98502994	1
AAG96AAC	Lys96Asn	0.01497006	0.98502994	1
AAG96ACG	Lys96Thr	0.01497006	0.98502994	2
AAG96AGG	Lys96Arg	0.01497006	0.98502994	1
AAG96CAG	Lys96Gln	0.01497006	0.98502994	1
AAG96GAC	Lys96Glu	0.01497006	0.98502994	6
Ins288 → T	FSC 96 (ins)	0.01497006	0.98502994	2
Ins288 → 33 nt	FSC 96 (ins)	0.01497006	0.98502994	4
Del291 → T	FSC 97 (del)	0.01497006	0.98502994	1
GGT97AGT	Gly97Ser	0.01497006	0.98502994	6
GGT97GAT	Gly97Asp	0.01497006	0.98502994	4
GGT97GCT	Gly97Ala	0.01497006	0.98502994	1
TAC99TAA	Tyr99stop	0.01497006	0.98502994	2
ACC100CCC	Thr100Pro	0.01497006	0.98502994	2
ACC100GCC	Thr100Ala	0.01497006	0.98502994	1
GTG130GCG	Val130Ala	0.01497006	0.98502994	1
GTG130GGG	Val130Gly	0.01497006	0.98502994	1
Ins391 → G	FSC 131 (ins)	0.01497006	0.98502994	3
Ins391 → GG	FSC 131 (ins)	0.01497006	0.98502994	2
Ins392 → G	FSC 131 (ins)	0.01497006	0.98502994	2
Ins392 → GG	FSC 131 (ins)	0.01497006	0.98502994	4
Ins393 → G	FSC 131 (ins)	0.01497006	0.98502994	2
Ins393 → GG	FSC 131 (ins)	0.01497006	0.98502994	1
Ins394 → ATGTGGTCG	FSC 131 (ins)	0.01497006	0.98502994	1
TGC131GGTGC	FSC 131 (ins)	0.01497006	0.98502994	1
GGT132AGT	Gly132Ser	0.01497006	0.98502994	1
GGT132GAT	Gly132Asp	0.01497006	0.98502994	1
GGT132GCT	Gly132Ala	0.01497006	0.98502994	1
GGT132TGT	Gly132Cys	0.01497006	0.98502994	2
ATT133ACT	Ile133Thr	0.01497006	0.98502994	17
Del398 → T	FSC 133 (del)	0.01497006	0,98502994	1
GCC134GTC	Ala134Val	0.01497006	0.98502994	2
ACC135AAC	Thr135Asn	0.01497006	0.98502994	3
ACC135CCC	Thr135Pro	0.01497006	0.98502994	4
GAT136TAT	Asp136Tyr	0.01497006	0.98502994	3
Ins408 → A	FSC 136 (ins)	0.01497006	0.98502994	4
CAT137CCT	His137Pro	0.01497006	0.98502994	1
CAT137CGT	His137Arg	0.01497006	0.98502994	1
CAT137GAT	His137Asp	0.01497006	0.98502994	1
TGT138CGT	Cys138Arg	0.01497006	0.98502994	3
TGT138TGG	Cys138Trp	0.01497006	0.98502994	1
Del417 → G	FSC 139 (del)	0.01497006	0.98502994	1
GTG139CTG	Val139Leu	0.01497006	0.98502994	3
GTG139GGG	Val139Gly	0.01497006	0.98502994	5
CGC140CCC	Arg140Pro	0.01497006	0.98502994	1
CAG141CCG	Gln141Pro	0.01497006	0.98502994	11
CAG141TAG	Gln141stop	0.01497006	0.98502994	1
Ins423 → CAGACGGCGCCAG	FSC 141 (ins)	0.01497006	0.98502994	1
ACG142AAG	Thr142Lys	0.01497006	0.98502994	1
ACG142ATG	Thr142Met	0.01497006	0.98502994	3
ACG142GCG	Thr142Ala	0.01497006	0.98502994	3
GCC143GGC	Ala143Gly	0.01497006	0.98502994	1
CTG172CCG	Leu172Pro	0.01497006	0.98502994	9
Del514 → C	FSC 172 (del)	0.01497006	0.98502994	1
Ins516 → CG	FSC 172 (ins)	0.01497006	0.98502994	1
ATG175ACG	Met175Thr	0.01497006	0.98502994	1
ATG175ATA	Met175Ile	0.01497006	0.98502994	10
ATG175GTG	Met175Val	0.01497006	0.98502994	6

[a] A-11C, nucleotide change A to C in position −11; Del-5 → G, deletion of nucleotide G in position −5; ATC6AAC, ATC at codon 7 changed to AAC; WT + GTC7GGC, double pattern wild-type + GTC at codon 7 changed to GGC; Ins37 → GACT, GACT inserted at codon 37.

[b] Promoter −11, nucleotidic mutation affecting the promoter region at position −11; del, deletion; FSC, frameshift codon; ins, insertion.

[c] p.S, probability associated with the susceptible phenotype.

[d] p.R, probability associated with the resistant phenotype.

Mutations conferring PZA^r at high confidence. Nine genetic variants (32 strains, category E) were found in both PZA^r and PZA^s isolates, but with a proportion higher than 70% in PZA^r strains. These mutations were mainly associated with an increase in free energy and/or structural constraints and were most frequently associated with PZA^r (93.5% of cases). We confirmed a reduced but still present PZase activity for some of these variants as a development of faint color during the enzymatic assay. Whereas Leu172Pro was found to be associated with homoplasy, for other substitutions, the number of cases was too low to consider convergent evolution in different lineages.

Mutations with an unclear role in conferring PZA^r. Five genetic variants (21 cases, category D) were found in both PZA^r and PZA^s isolates but at a proportion less than 70% in PZA^r strains. Two genetic variants (15 cases) showed borderline behavior in terms of structure/free energy variation and enzyme activity. Homoplasy was found for the Val139Ala mutation, thus suggesting a putative role in phenotypic resistance or at least in increasing the MIC. Pro62Arg, Asp63Ala, and Ser65Pro substitutions (6 cases) represent another group of mutations belonging to this ambiguous category. Further characterization of these mutations is needed to better understand their correlation with the PZA phenotype.

Mutations not involved in phenotypic resistance. Twenty-seven genetic variants were not associated with PZA^r according to our classification. Eighteen mutations (163 cases, category C) were most frequently associated with PZA^s (91.4% of cases). It should be noted that the Val21Ala mutation was also found associated with other mutations in category A responsible for PZA^r/PZase negativity. Interestingly, all these mutations were found to be associated with single lineages; thus, no homoplasy was observed. Further characterization of these mutations is needed to better understand their role (if any) in PZA susceptibility. The remaining genetic variants (27 cases, category B) did not affect the amino acid sequence of the PZase enzyme. We observed two silent mutations: TCC65TCT (Ser65Ser), GCG38GCC (Ala38Ala). The Ser65Ser silent mutation was found associated with the Delhi/CAS lineage. In some cases, sequencing of the upstream region of *pncA* allowed the identification of a deletion at position −125 or an insertion at nucleotide −3; however, isolates harboring these genetic variants were found associated with both phenotypic resistance and susceptibility. According to these data, and supported by the lack of homoplasy for these mutations, the indels detected do not represent a marker for PZA^r.

A total of 1,062 clinical isolates (1,062/1,950 [55.4%]) showed a WT sequence for the PZase enzyme (included in category B), and the sequence was associated with PZA^s in more than 80% of cases. Enzymatic assay results were not available for all: 17 isolates (out of 138 tested; 12.3%) gave a negative PZase enzymatic activity, indicating that a WT PZase does not exclude phenotypic resistance *a priori*.

DISCUSSION

PZA DST is crucial for successful management of patients with susceptible and drug-resistant TB, especially with MDR TB. Furthermore, future shorter regimens for both drug-resistant and drug-susceptible TB will include PZA as a key drug for achieving both sterilization and prevention of the development of drug resistance to new drugs (17, 18). Thus, reliable PZA^r data for clinical isolates are crucial for guiding the clinical management of patients. Phenotypic tests, however, have a long turnaround time, are expensive, and are considered poorly reliable. As a consequence, the design of a molecular test for predicting PZA^r is a priority. The development of a rapid molecular PZA DST is hampered by the diverse nature of resistance-associated mutations mainly scattered over the full length of the *pncA* gene, and by the fact that the impact of individual mutations has not been systematically investigated (4, 12). Therefore, we performed a large-scale study linking *pncA* sequence diversity with phenotypic, structural biology and population biology data to develop the first encyclopedia of *pncA* sequence variations linked to either a PZA^r or PZA^s phenotype. This is likely to pave the way for application of new genome-based sequencing technologies for predicting PZA^r, allowing for personalized treatment algorithms.

Strikingly, nearly 85% of the genetic variants identified in the *pncA* gene were associated with phenotypic resistance to PZA and were classified as "high-confidence" PZA resistance mutations. All in-frame and frameshift indel mutations within the coding region were included in this group. We found that nearly 90% of observed mutations are associated with protein structural destabilization that causes loss of enzymatic activity. Equally importantly, we described 27 mutations most likely not involved in PZA resistance that should be "filtered out" in future molecular tests and labeled as not "clinically relevant" (Fig. 4). Only five mutations cannot be classified by our approach and remain without clear association with a resistance or susceptible phenotype. These mutations need further validation for association with PZA^r and/or with a specific genetic background by an allelic exchange procedure as performed for other drugs (19).

Discrepancies between molecular and phenotypic DST are confusing for clinicians managing patients; 180 isolates investigated here showed discrepant results between phenotypic and genotypic tests (WT *pncA* gene sequence and resistance by Bactec MGIT 960 DST). It has been reported that the Bactec MGIT 960 mycobacterial detection system may overestimate resistance even in the best laboratory settings (due to changes in the medium pH and/or variability in the inoculum size). Alternatively, a different mechanism of resistance, such as mutations in *rpsA*, could also be hypothesized for a few cases, although these were not clearly confirmed in clinical isolates (data not shown) (20–23). In Fig. S1 in the supplemental material, we modeled the impact of these "discrepant cases" in different hypothetical diagnostic scenarios to provide worst and best performances of *pncA* sequencing-based assay as follows. If all 180 cases were truly susceptible, the diagnostic accuracy of a molecular test for PZA based on sequence would be 98.77% (95% confidence interval [95% CI], 98.18 to 99.17%) (Fig. S1C) in the range of the rifampin and isoniazid test results (11). If the 180 cases were truly PZA^r strains (wrongly predicted as PZA^s by *pncA* gene sequence), the diagnostic accuracy of *pncA* sequencing in detecting PZA^r would be 89.54% (95% CI, 89.21 to 90.82%), in the range of isoniazid resistance (Fig. S1B) (11).

Based on our findings, any future molecular test for PZA resistance should be able not only to detect the absence of the wild-type sequence but also to identify the specific SNPs. We found, indeed, a relevant number (10%) of mutations previously not reported as associated with drug resistance (DR) and the degree of variability in terms of indel mutations. In addition, we found mutations not associated with DR, including the previously reported lineage-specific genetic variants (e.g., TCC65TCT in Delhi/CAS) (14). Ac-

cordingly, only an assay with the capacity to provide in-depth sequence information could comply with the minimal requirements for a new molecular PZA^r test. Fully automated, low-cost medium-density arrays and user-friendly whole-gene/whole-genome sequencing-based approaches will become a reality in the very near future and will be the most suitable assays to fulfill this task. In particular, new next-generation sequencing (NGS)-based diagnostics could represent innovative tools to reduce false PZA^r cases and to improve safe and fast detection of drug resistances by molecular DST (24). Our work has generated the minimum sets of mutations that should be included in any molecular test for PZA and provide a start point for a *pncA* genetic variation encyclopedia needed for the valid interpretation of data generated by massive sequencing approaches.

An additional aspect that is highlighted by our study is the great advantage in sharing large data sets generated by several groups. The establishment of a common database combining data from clinical isolates collected in a large number of settings was crucial to improve our understanding the role of *pncA* gene mutations in determining the PZA susceptibility phenotype of *M. tuberculosis*. The high number of samples providing sufficient reiteration of less frequent mutations together with the inclusion of different parameters (phenotype, genotype, enzymatic activity, structure, and free energy analyses) in a decision tree allowed us to define specific operational categories of mutations relevant from a clinical point of view. This enabled us to build a user-friendly diagnostic algorithm through the classification of specific SNPs in a shared database collecting more-complex information. These large shared databases of mutations involved in drug resistance could contribute to a better understanding of molecular mechanisms of resistance, improved molecular diagnostics, new diagnostic algorithms, and better public health control of drug-sensitive and drug-resistant TB.

MATERIALS AND METHODS

Strain selection. Strains were made available by six TB National/Supranational Reference and partner laboratories within the TB-PANNET Consortium to provide wide coverage for most of the lineages observed for the *M. tuberculosis* complex. Strains were tested for PZA susceptibility and included in the study regardless of testing for other antitubercular drugs. PZA drug susceptibility testing (DST) was performed by using a Bactec MGIT 960 mycobacterial detection system and MGIT 960 PZA kits (BD, Franklin Lakes, NJ, USA) according to the manufacturer's instructions. A total of 1,950 clinical isolates were incorporated in the database. Whenever available, genotyping information (spoligotyping and/or mycobacterial interspersed repetitive-unit−variable-number tandem-repeat [MIRU-VNTR] typing [25]) were collected. The MIRU-VNTR*plus* web tool (26, 27) was used to define lineage information (similarity search settings for identification: 0.17; distance measure for MIRU-VNTR: categorical, weighting 1; distance measure for spoligotyping: categorical, weighting 1).

***pncA* gene sequencing.** DNA was extracted as described elsewhere (28). The *pncA* gene (*Rv2043c*, NCBI gene identifier [ID] 888260), including the proximal promoter region, was amplified. On a subset of samples, the distal promoter region (>100 bp upstream of the start codon) was also included in the amplified region according to the protocol described in reference 29. Amplicons were sequenced with an automated DNA sequencer. The *pncA* gene sequence of isolates from Samara, Russian Federation, was determined from whole-genome sequencing data as previously described (15). Mutations in the *pncA* gene were identified by comparison with the wild-type *M. tuberculosis* H37Rv *pncA* gene sequence.

PZase assay. PZase activity was evaluated as described by Singh et al. (30). Briefly, a Middlebrook 7H9 (BD, Franklin Lakes, NJ, USA) 1.5% agarose containing PZA (Sigma-Aldrich Corporation, Saint Louis, MO, USA) at a final concentration of 400 μg/ml was prepared. Melted PZA agar was distributed in glass tubes by using an agarose base to obtain a semitransparent medium allowing the detection of a faint pink band against a white background. A heavy loopful of actively growing culture was carefully inoculated on the surface of the PZA agar medium and incubated at 37°C for 4 days. One milliliter of ferrous ammonium sulfate (1%) was added to each tube after incubation and observed for 4 h for the appearance of a pink band (positive) in the subsurface agar. PZA-resistant isolates of *M. bovis* (negative by the PZase test) were used as negative controls, and the PZA-susceptible strain *M. tuberculosis* H37Rv was used as a positive control. All isolates showing discrepant results (namely, *pncA* mutant and PZase positive or WT *pncA* and PZase negative) were retested at 4 and 10 days (31).

PZase structure. For each amino acid substitution, we performed an *in silico* analysis of the free energy variation associated with the specific mutation taking into account an acidic environmental pH (6.0), very close to the one required for PZA activity. The crystal structure of the PZase enzyme determined to 2.2-Å resolution (PDB code 3PL1) (32) was used in conjunction with the program FoldX (33). Mean free energy variation was calculated for triplicates of predicted structures, and based on statistical analysis, a free energy variation greater than 2 kJ/mol was considered to destabilize the enzyme. Frameshifts and mutations affecting the promoter region were not considered. Free energy variation was then integrated with a visual structural analysis in order to identify substitutions tolerated by the free energy term but detrimental for the specific activity of the enzyme.

Statistical analysis. For understanding the significance of each mutation, we predicted the DST by fitting a conditional inference tree model considering results of sequencing, activity, and the combination of structure and energy analyses as predictors. In the model, we applied recursive partitioning based on conditional permutation tests. Furthermore, at each step, *P* values were adjusted for multiplicity by the procedure of Benjamini and Yekutiely (34). The majority of recursive partitioning algorithms introduced since 1963 (35), such as CHAID and CART, yield trees with too many branches and can also fail to pursue branches which can add significantly to the overall fit. This leads to potential drawbacks: overfitting and a selection bias toward covariates with many possible splits or missing values (36, 37).

This approach is able to address missing data, since it uses surrogate splits to determine the daughter node where the observations with missing values in the primary split variable are sent (for further details, see references 38 and 39). As output of the model, given an isolate's profile, a conditional probability of being PZA resistant is given. As a general rule, adjusted *P* values of less than 0.05 were considered significant. In the model, we applied recursive partitioning based on conditional permutation tests. In fact, when splitting, the use of the conditional distribution of the statistics ensures an unbiased selection of the covariates. This statistical approach prevented overfitting and overgrown trees, and no further pruning or cross-validation was needed.

Further details on the rationale used for the analysis is available in the supplemental material.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.01819-14/-/DCSupplemental.

Table S1, XLSX file, 0.2 MB.
Table S2, XLSX file, 0.04 MB.
Figure S1, DOCX file, 0.3 MB.
Text S1, DOCX file, 0.01 MB.

ACKNOWLEDGMENTS

We thank Federica Piana (S. Croce and Carle Hospital, Cuneo, Italy) for providing additional *M. tuberculosis* complex isolates. We thank the fol-

lowing for helpful discussions: Vladyslav Nikolayevskyy, Yanina Balabanova, Irina Kontsevaya, and Olga Ignatyeva.

This work was supported by the European Community's Seventh Framework Programme (EU FP7/2007–2013) under grant agreement FP7-223681 to D.M.C. (http://www.tbpannet.org) and the TB-PANNET Consortium and partly by EU FP7 grant 201483 to the TB-EUROGEN consortium.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. **Girling DJ.** 1984. The role of pyrazinamide in primary chemotherapy for pulmonary tuberculosis. Tubercle **65:**1–4. http://dx.doi.org/10.1016/0041-3879(84)90024-2.
2. **Mitchison DA.** 1985. The action of antituberculosis drugs in short-course chemotherapy. Tubercle **66:**219–225. http://dx.doi.org/10.1016/0041-3879(85)90040-6.
3. **Mitchison DA.** 2000. Role of individual drugs in the chemotherapy of tuberculosis. Int. J. Tuberc. Lung Dis. **4:**796–806.
4. **Zhang Y, Yew WW.** 2009. Mechanisms of drug resistance in *Mycobacterium tuberculosis*. Int. J. Tuberc. Lung Dis. **13:**1320–1330.
5. **Lange C, Abubakar I, Alffenaar JW, Bothamley G, Caminero JA, Carvalho AC, Chang KC, Codecasa L, Correia A, Crudu V, Davies P, Dedicoat M, Drobniewski F, Duarte R, Ehlers C, Erkens C, Goletti D, Günther G, Ibraim E, Kampmann B, Kuksa L, de Lange W, van Leth F, van Lunzen J, Matteelli A, Menzies D, Monedero I, Richter E, Rüsch-Gerdes S, Sandgren A, Scardigli A, Skrahina A, Tortoli E, Volchenkov G, Wagner D, van der Werf MJ, Williams B, Yew WW, Zellweger JP, Cirillo DM, TBNET.** 2014. Management of patients with multidrug-resistant/extensively drug-resistant tuberculosis in Europe: a TBNET consensus statement. Eur. Respir. J. **44:**23–63. http://dx.doi.org/10.1183/09031936.00188313.
6. **Diacon AH, Dawson R, von Groote-Bidlingmaier F, Symons G, Venter A, Donald PR, van Niekerk C, Everitt D, Winter H, Becker P, Mendel CM, Spigelman MK.** 2012. 14-day bactericidal activity of PA-824, bedaquiline, pyrazinamide, and moxifloxacin combinations: a randomised trial. Lancet **380:**986–993. http://dx.doi.org/10.1016/S0140-6736(12)61080-0.
7. **Horne DJ, Pinto LM, Arentz M, Lin SY, Desmond E, Flores LL, Steingart KR, Minion J.** 2013. Diagnostic accuracy and reproducibility of WHO-endorsed phenotypic drug susceptibility testing methods for first-line and second-line antituberculosis drugs. J. Clin. Microbiol. **51:**393–401. http://dx.doi.org/10.1128/JCM.02724-12.
8. **Hoffner S, Angeby K, Sturegård E, Jönsson B, Johansson A, Sellin M, Werngren J.** 2013. Proficiency of drug susceptibility testing of *Mycobacterium tuberculosis* against pyrazinamide: the Swedish experience. Int. J. Tuberc. Lung Dis. **17:**1486–1490. http://dx.doi.org/10.5588/ijtld.13.0195.
9. **Piersimoni C, Mustazzolu A, Giannoni F, Bornigia S, Gherardi G, Fattorini L.** 2013. Prevention of false resistance results obtained in testing the susceptibility of *Mycobacterium tuberculosis* to pyrazinamide with the Bactec MGIT 960 system using a reduced inoculum. J. Clin. Microbiol. **51:**291–294. http://dx.doi.org/10.1128/JCM.01838-12.
10. **Zhang Y, Permar S, Sun Z.** 2002. Conditions that may affect the results of susceptibility testing of *Mycobacterium tuberculosis* to pyrazinamide. J. Med. Microbiol. **51:**42–49.
11. **Ling DI, Zwerling AA, Pai M.** 2008. Rapid diagnosis of drug-resistant TB using line probe assays: from evidence to policy. Expert Rev. Respir. Med **2:**583–588. http://dx.doi.org/10.1586/17476348.2.5.583.
12. **Singh P, Mishra AK, Malonia SK, Chauhan DS, Sharma VD, Venkatesan K, Katoch VM.** 2006. The paradox of pyrazinamide: an update on the molecular mechanisms of pyrazinamide resistance in mycobacteria. J. Commun. Dis. **38:**288–298.
13. **Köser CU, Comas I, Feuerriegel S, Niemann S, Gagneux S, Peacock SJ.** 2014. Genetic diversity within *Mycobacterium tuberculosis* complex impacts on the accuracy of genotypic pyrazinamide drug-susceptibility assay. Tuberculosis (Edinb.) **94:**451–453. http://dx.doi.org/10.1016/j.tube.2014.04.002.
14. **Feuerriegel S, Köser CU, Niemann S.** 2014. Phylogenetic polymorphisms in antibiotic resistance genes of the *Mycobacterium tuberculosis* complex. J. Antimicrob. Chemother. **69:**1205–1210. http://dx.doi.org/10.1093/jac/dkt535.
15. **Casali N, Nikolayevskyy V, Balabanova Y, Harris SR, Ignatyeva O, Kontsevaya I, Corander J, Bryant J, Parkhill J, Nejentsev S, Horstmann RD, Brown T, Drobniewski F.** 2014. Evolution and transmission of drug-resistant tuberculosis in a Russian population. Nat. Genet. **46:**279–286. http://dx.doi.org/10.1038/ng.2878.
16. **Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, Parkhill J, Malla B, Berg S, Thwaites G, Yeboah-Manu D, Bothamley G, Mei J, Wei L, Bentley S, Harris SR, Niemann S, Diel R, Aseffa A, Gao Q, Young D, Gagneux S.** 2013. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. Nat. Genet. **45:**1176–1182. http://dx.doi.org/10.1038/ng.2744.
17. **Chang KC, Leung CC, Yew WW, Leung EC, Leung WM, Tam CM, Zhang Y.** 2012. Pyrazinamide may improve fluoroquinolone-based treatment of multidrug-resistant tuberculosis. Antimicrob. Agents Chemother. **56:**5465–5475. http://dx.doi.org/10.1128/AAC.01300-12.
18. **Zumla AI, Gillespie SH, Hoelscher M, Philips PP, Cole ST, Abubakar I, McHugh TD, Schito M, Maeurer M, Nunn AJ.** 2014. New antituberculosis drugs, regimens, and adjunct therapies: needs, advances, and future prospects. Lancet Infect. Dis. **14:**327–340. http://dx.doi.org/10.1016/S1473-3099(13)70328-1.
19. **Nebenzahl-Guimaraes H, Jacobson KR, Farhat MR, Murray MB.** 2014. Systematic review of allelic exchange experiments aimed at identifying mutations that confer drug resistance in *Mycobacterium tuberculosis*. J. Antimicrob. Chemother. **69:**331–342. http://dx.doi.org/10.1093/jac/dkt358.
20. **Shi W, Zhang X, Jiang X, Yuan H, Lee JS, Barry CE, III, Wang H, Zhang W, Zhang Y.** 2011. Pyrazinamide inhibits trans-translation in *Mycobacterium tuberculosis*. Science **333:**1630–1632. http://dx.doi.org/10.1126/science.1208813.
21. **Zhang S, Chen J, Shi W, Liu W, Zhang W, Zhang Y.** 2013. Mutations in *panD* encoding aspartate decarboxylase are associated with pyrazinamide resistance in *Mycobacterium tuberculosis*. Emerging Microbes Infect. **2:**e34. http://dx.doi.org/10.1038/emi.2013.38.
22. **Alexander DC, Ma JH, Guthrie JL, Blair J, Chedore P, Jamieson FB.** 2012. Gene sequencing for routine verification of pyrazinamide resistance in *Mycobacterium tuberculosis*: a role for *pncA* but not *rpsA*. J. Clin. Microbiol. **50:**3726–3728. http://dx.doi.org/10.1128/JCM.00620-12.
23. **Tan Y, Hu Z, Zhang T, Cai X, Kuang H, Liu Y, Chen J, Yang F, Zhang K, Tan S, Zhao Y.** 2014. Role of *pncA* and *rpsA* gene sequencing in detection of pyrazinamide resistance in *Mycobacterium tuberculosis* isolates from southern China. J. Clin. Microbiol. **52:**291–297. http://dx.doi.org/10.1128/JCM.01903-13.
24. **Köser CU, Ellington MJ, Cartwright EJ, Gillespie SH, Brown NM, Farrington M, Holden MT, Dougan G, Bentley SD, Parkhill J, Peacock SJ.** 2012. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. PLoS Pathog. **8:**e1002824. http://dx.doi.org/10.1371/journal.ppat.1002824.
25. **Barnes PF, Cave MD.** 2003. Molecular epidemiology of tuberculosis. N. Engl. J. Med. **349:**1149–1156. http://dx.doi.org/10.1056/NEJMra021964.
26. **Allix-Béguec C, Harmsen D, Weniger T, Supply P, Niemann S.** 2008. Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of *Mycobacterium tuberculosis* complex isolates. J. Clin. Microbiol. **46:**2692–2699. http://dx.doi.org/10.1128/JCM.00540-08.
27. **Weniger T, Krawczyk J, Supply P, Niemann S, Harmsen D.** 2010. MIRU-VNTRplus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. Nucleic Acids Res. **38:**W326–W331. http://dx.doi.org/10.1093/nar/gkq351.
28. **Miotto P, Piana F, Penati V, Canducci F, Migliori GB, Cirillo DM.** 2006. Use of genotype MTBDR assay for molecular detection of rifampin and isoniazid resistance in *Mycobacterium tuberculosis* clinical strains isolated in Italy. J. Clin. Microbiol. **44:**2485–2491. http://dx.doi.org/10.1128/JCM.00083-06.
29. **Juréen P, Werngren J, Toro JC, Hoffner S.** 2008. Pyrazinamide resistance and *pncA* gene mutations in *Mycobacterium tuberculosis*. Antimicrob. Agents Chemother. **52:**1852–1854. http://dx.doi.org/10.1128/AAC.00110-08.
30. **Singh P, Wesley C, Jadaun GP, Malonia SK, Das R, Upadhyay P, Faujdar J, Sharma P, Gupta P, Mishra AK, Singh K, Chauhan DS, Sharma VD, Gupta UD, Venkatesan K, Katoch VM.** 2007. Comparative evaluation of Löwenstein-Jensen proportion method, BacT/ALERT 3D system, and enzymatic pyrazinamidase assay for pyrazinamide susceptibility testing of *Mycobacterium tuberculosis*. J. Clin. Microbiol. **45:**76–80. http://dx.doi.org/10.1128/JCM.00951-06.

31. **Mirabal NC, Yzquierdo SL, Lemus D, Madruga M, Milián Y, Echemendía M, Takiff H, Martin A, Van der Stuyf P, Palomino JC, Montoro E.** 2010. Evaluation of colorimetric methods using nicotinamide for rapid detection of pyrazinamide resistance in *Mycobacterium tuberculosis*. J. Clin. Microbiol. **48:**2729–2733. http://dx.doi.org/10.1128/JCM.00311-10.
32. **Petrella S, Gelus-Ziental N, Maudry A, Laurans C, Boudjelloul R, Sougakoff W.** 2011. Crystal structure of the pyrazinamidase of *Mycobacterium tuberculosis*: insights into natural and acquired resistance to pyrazinamide. PLoS One **6:**e15785. http://dx.doi.org/10.1371/journal.pone.0015785.
33. **Guerois R, Nielsen JE, Serrano L.** 2002. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J. Mol. Biol. **320:**369–387. http://dx.doi.org/10.1016/S0022-2836(02)00442-4.
34. **Benjamini Y, Yekutieli D.** 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat. **24:**1165–1188. http://dx.doi.org/10.1214/aos/1013699998.
35. **Morgan JN, Sonquist JA.** 1963. Problems in the analysis of survey data, and a proposal. J. Am. Stat. Assoc. **58:**415–434. http://dx.doi.org/10.1080/01621459.1963.10500855.
36. **Breiman L, Friedman JH, Olshen RA, Stone CJ.** 1984. Classification and regression trees, 1st ed. Chapman and Hall/CRC, New York, NY.
37. **Mingers J.** 1987. Expert systems rule induction with statistical data. J. Oper. Res. Soc. **38:**39–47. http://dx.doi.org/10.2307/2582520.
38. **Hothorn T, Hornik K, Zeileis A.** 2006. Unbiased recursive partitioning: a conditional inference framework. J. Comput. Graph. Stat. **15:**651–674. http://dx.doi.org/10.1198/106186006X133933.
39. **Strasser H, Weber C.** 1999. On asymptotic theory permutation statistics. Math. Methods Stat. **8:**220–250.