Journal of Antimicrobial Chemotherapy

J Antimicrob Chemother 2010; **65**: 1359–1367
doi:10.1093/jac/dkq120 Advance Access publication 28 April 2010

embCAB sequence variation among ethambutol-resistant *Mycobacterium tuberculosis* isolates without *embB*306 mutation

Claudia Plinke[1], Helen S. Cox[2], Nana Zarkua[3], Hamraev A. Karimovich[4], Kai Braker[5], Roland Diel[6], Sabine Rüsch-Gerdes[7], Silke Feuerriegel[1]*† and Stefan Niemann[1]†

[1]*Molecular Mycobacteriology, Research Center Borstel, Borstel, Germany;* [2]*Macfarlane Burnet Institute for Medical Research and Public Health, Melbourne, Australia;* [3]*Médecins Sans Frontières, Tashkent, Uzbekistan;* [4]*Ministry of Health, Nukus, Karakalpakstan, Uzbekistan;* [5]*Médecins Sans Frontières, Berlin, Germany;* [6]*Department of Pneumology, Medical School Hannover (MHH), Hannover, Germany;* [7]*National Reference Center for Mycobacteria, Research Center Borstel, Borstel, Germany*

*Corresponding author. Tel: +49-4537-188274; Fax: +49-4537-188311; E-mail: sfeuerriegel@fz-borstel.de
†S. F. and S. N. contributed equally to this work.

Received 28 January 2010; returned 1 March 2010; revised 19 March 2010; accepted 21 March 2010

Objectives: Mechanisms of resistance to ethambutol in *Mycobacterium tuberculosis* remain inadequately described. Although there is mounting evidence that mutations of codon 306 in *embB* play a key role, a significant number of phenotypically ethambutol-resistant strains do not carry mutations in this codon. Here, other mutations in the *embCAB* operon are suggested to be involved in resistance development.

Methods: The entire *embCAB* operon (~10 kb) was analysed in 34 phenotypically ethambutol-resistant *M. tuberculosis* strains without mutations in *embB*306 and in 12 ethambutol-susceptible strains. Furthermore, 106 control strains were investigated for the presence of particular mutations only.

Results: Overall, 18 non-synonymous mutations in 15 distinct codons of the *embCAB* operon were identified in ethambutol-resistant strains but not in ethambutol-susceptible isolates. The majority occurred in the *embB* gene (10 distinct codons), in a 570 bp region also encompassing *embB*306. Mutations in *embC* and *embA* were found rarely and in most cases in combination with polymorphisms in *embB*. One synonymous mutation (*embA* 228 bp) and two non-synonymous mutations (*embC*Val981Leu and *embC*Arg738Gln) were found in ethambutol-susceptible strains as well as resistant strains and were confirmed to represent phylogenetic markers for strains of the Beijing, Haarlem and Delhi/CAS genotypes, respectively.

Conclusions: Besides mutations in *embB*306, mutations in *embB*406 and *embB*497 were confirmed as hot spots for genomic variation in ethambutol-resistant clinical isolates. Of all resistant strains 70.6% carry a mutation in a relatively short region in *embB*, which therefore represents a promising target for inclusion in molecular assays for rapid detection of ethambutol resistance.

Keywords: resistance testing, new mutations, genotyping, phylogenetic markers, Beijing genotype

Introduction

The worldwide emergence of drug-resistant *Mycobacterium tuberculosis* complex (MTBC) strains complicates the treatment of patients suffering from tuberculosis (TB). In particular multidrug-resistant (MDR) and extensively drug-resistant (XDR) TB represents a serious challenge for TB control.[1]

The rapid detection of drug resistance is essential to design appropriate treatment regimens, prevent treatment failure and thus reduce the further spread of drug-resistant isolates. While the development of liquid media-based systems has reduced turnaround times for drug susceptibility testing, molecular assays (e.g. line-probe assays like GenoType MTBDR; Hain Lifescience, Nehren, Germany) have the potential to dramatically reduce delay. However, these molecular assays require precise knowledge of the genetic variation involved in the development of resistance to particular anti-TB drugs.

Ethambutol is a key component in anti-TB therapy. Previous studies suggest that 69% of ethambutol-resistant *M. tuberculosis* strains carry a mutation in the *embB* gene, mutations in codon *embB*306 occurring most frequently.[2] Consequently mutations in this codon have been suggested as molecular markers for the rapid detection of ethambutol resistance.[3–6] Furthermore, two recent studies demonstrated a clear causative relationship between *embB*306 point mutations and *in vitro* ethambutol resistance by allelic-exchange experiments.[7,8]

However, approximately one-third of ethambutol-resistant clinical isolates do not carry a mutation in *embB*306 and are therefore not detectable by using molecular methods based on the determination of polymorphisms in *embB*306 only (e.g. DNA sequencing, real-time PCR or strip technologies). Although other mutations in the *embCAB* operon are suggested to confer resistance, only limited data have been available until now, with most studies analysing only a short fragment of the *embB* gene encompassing codon 306.[2,3,5,9,10]

To add further data to this important question we sequenced the entire *embCAB* operon from isolates phenotypically ethambutol resistant, but lacking a mutation in *embB*306, that were identified from previous investigations. These were compared with ethambutol-susceptible isolates included as controls.

Materials and methods

M. tuberculosis clinical isolates

A total of 152 clinical *M. tuberculosis* isolates were included in this study. The strain collection for analysis of the entire *embCAB* operon consisted of 30 (24 ethambutol resistant and 6 ethambutol susceptible) clinical *M. tuberculosis* isolates obtained from patients in Germany and 16 (10 ethambutol resistant and 6 ethambutol susceptible) from patients in Uzbekistan. The ethambutol-resistant strains were previously identified as having no mutations in codon *embB*306. To assess the significance of particular mutations for ethambutol resistance, 106 additional strains (104 ethambutol susceptible and 2 ethambutol resistant) collected during 1 year (2007) in Hamburg, Germany, were analysed for the presence of specific mutations.

Drug susceptibility testing

For all strains, resistance to the key antimycobacterial drugs isoniazid, rifampicin, ethambutol and streptomycin was determined. Resistance determination for isoniazid, rifampicin, ethambutol and streptomycin was performed by using the proportion method on Löwenstein–Jensen (LJ) medium (critical concentration used for ethambutol was 2.0 mg/L). If growth was insufficient, drug susceptibility testing was performed by using the modified proportion method in BACTEC 460TB (Becton-Dickinson; critical concentration used for ethambutol was 3.75 mg/L) or Bactec MGIT 960 system (5.0 mg/L for ethambutol).

IS6110 profiling and spoligotyping

Extraction of genomic DNA from mycobacterial strains and DNA fingerprinting using IS*6110* as a probe were performed according to a standardized protocol as described previously.[11] Additionally, all isolates were analysed by the spoligotyping technique as described previously by Kamerbeek *et al.*[12] The molecular typing data were analysed using the Bionumerics software (version 5.1; Applied Maths, Sint-Martens-Latem, Belgium) as instructed by the manufacturer. Spoligotyping data were used to additionally confirm strain relationships and for identification of Beijing genotype isolates (no hybridization to spacers 1–34, hybridization to spacers 35–43).

PCR amplification and sequencing strategy

The 34 ethambutol-resistant as well as the 12 ethambutol-susceptible clinical *M. tuberculosis* isolates were sequenced in the whole *embCAB* operon. The 106 strains collected in Hamburg in 2007 were analysed with regard to *embC*Arg738Gln as well as nucleotides −8, −12, −16, −25 and −32 in the *embCA* intergenic region. Furthermore the *embB* region encompassing codons 306–497 was analysed. DNA amplification was performed using the primers and conditions listed in Table 1. The PCR products obtained were sequenced using an ABI Prism 3130*xl* Genetic Analyser (Applied Biosystems, CA, USA) and the ABI Prism BigDye Terminator Kit v.3.1, according to the manufacturer's instructions. Analysis of sequence data was performed using SeqScape v2.6 software.

Results

The entire *embCAB* operon (~10 kb) of 34 ethambutol-resistant and 12 ethambutol-susceptible clinical *M. tuberculosis* isolates was sequenced. All ethambutol-resistant strains showed additional resistance at least to isoniazid, rifampicin and streptomycin ($n=25$) or isoniazid and streptomycin ($n=9$). Overall 33 different mutations, non-synonymous (resulting in an amino acid replacement) as well as synonymous (no amino acid replacement), were identified in *embCAB* of resistant strains only (Table 2). Most mutations occurred in *embB* outside of codon *embB*306 (13 non-synonymous and three synonymous mutations). The remaining regions of the operon were less frequently affected; seven mutations in *embA* (three non-synonymous and four synonymous mutations), six nucleotide variations in the *embC–embA* intergenic region and four mutations in *embC* (three non-synonymous and one synonymous mutation). A total of 26 ethambutol-resistant strains carried at least one mutation resulting in an amino acid replacement. Only three ethambutol-resistant isolates did not show any mutations across the entire operon. Three mutations occurred in resistant and susceptible isolates and are more likely to be phylogenetic rather than resistance markers.

In the following sections, the mutations detected in resistant strains only (Figures 1 and 2 and Table 2) as well as the mutations found in both resistant and susceptible isolates (phylogenetic markers; Figure 3) are presented in detail.

Polymorphisms in the embC gene

In the *embC* gene three non-synonymous mutations, possibly associated with ethambutol resistance, were identified in four ethambutol-resistant strains (see Table 2). Two strains had an amino acid replacement located in codon 738. Amino acid changes in *embC* in codons 270 and 406 were found in one isolate each. One of the strains with Arg738Gln replacement as well as the strain with Thr270Ile also had other mutations in *embCAB* resulting in amino acid changes. These additional mutations in *embA*, *embB* and the *embCA* intergenic region were also exclusively found in ethambutol-resistant isolates (see Table 2). One ethambutol-resistant isolate had a silent mutation at nucleotide position 1239 bp. Furthermore, one ethambutol-susceptible isolate showed a mutation at codon 102.

Polymorphisms in the embC–embA intergenic region

Nine ethambutol-resistant isolates had nucleotide changes located in the *embC–embA* intergenic region (see Table 2 and Figure 1). In four ethambutol-resistant strains a nucleotide substitution was identified at position −12 bp, but this mutation was also present in one ethambutol-susceptible isolate. The other nucleotide variations in the intergenic region were found in

Table 1. PCR primers and conditions used for amplification and sequencing

Primer	Primer sequence (5′→3′)	Size (bp)	PCR conditions[a]		
			D (s)	A (°C, s)	E (s)
Rv3795anewF	CTG GGG ATC GGT GGA GCA GTA	969	30	66.0, 45	60
Rv3795anewR	GCG TCG GTC AGG GTG AAG G				
Rv3795bF	TGG ACG GGC GGG GCT CAA T	1306	60	65.0, 60	90
Rv3795bcR	GCA AAC AGG GCG AAA AAG A				
Rv3795cF	TCC TGG CGG CGT TAT TCT T	1653	60	65.0, 60	90
Rv3795cS1[b]	TGG ACG GCG ATT CGG GTT CT				
Rv3795cS2[b]	GGA CTG GGC GGT CGG TTT G				
Rv3795cR	CAA CCG GGG TGA TGA TGG C				
embIR-F[c]	CTG GTG GTC GCG GTG ATC AT	907	30	64.0, 45	60
embIR-newR	ACG GTC GCT GGC AGG GGA AGT T				
embIR-newF	CGC CTA TGA CCC GAA CCT GAG	787	30	65.0, 45	60
embIR-R[c]	AAT TGG CGT CCT TGC CTT				
embA1-F[c]	GTG ACT CGC AGC GGG CTG TG	1223	30	68.0, 30	90
embA1-R[c]	CGG TGA ACA CAG CGA CCC GG				
embA2-F[c]	TGG ACC GGC TCA GCA GGG G	1500	30	67.0, 30	90
embA2-R[c]	TCA GGT TGG CCT TGG CGG TG				
embC1-F[c]	CCC AAC CAG CCC AAT GTT C	890	40	64.5, 30	60
embC1-R[c]	GGC GGT GTC CAG GAT GTG				
embC2-F[c]	GCT GCA CAT CCT GGA CAC	914	40	60.0, 30	60
embC2-R[c]	ACG ACA TTG CCA CCG ATA C				
embC3-F[c]	GTA TCG GTG GCA ATG TCG T	1176	40	60.0, 30	60
embC3-R[c]	CGG GAT GGC GGA CAG TGG T				
embC4-F[c]	ACC ACT GTC CGC CAT CCC G	635	30	67.0, 30	45
embC4-R[c]	GAC GAC GGC TGC TAG GCG TG				
embC1-F2	GAA GCC GCC CCA CCC CGT ATC G	996	30	68.0, 45	60
embC1-R2	AGC CCA CAG CGC CAG CAG GTC GTA				
embC2-F2	GCG GGC ATG TTT CTG GCT GTC TGG	1058	30	64.0, 45	60
embC2-R2	CAC GCC GGG TAC TGG GAA ATC ATC				
embC3-F2	CGC GCC CGG CTG CCC TAC AAC	464	30	68.0, 45	60
embC3-R2	GCC AGC CCC ACC AGC CAG TCC A				
embC_5′ (+2002)	CTG ACG GTG CTG GTG CTG CTG CTA	372	30	65.0, 30	30
embC_3′ (+2373)	GGG AAT GCC GTT GGG TGT GAA GG				
embCA_5′ (+3134)	TGG CCA GCT ACC TCA AAG ACG ACT	437	30	60.0, 30	30
embCA_3′ (+200)	GGC GCC CCG GAT ACC AGA				

[a]D, length of denaturation at 94°C; A, primer annealing conditions (temperature in °C and length in s); E, length of extension at 72°C. All PCRs were 35 cycles, were preceded by an initial denaturation step at 95°C for 15 min and included a final extension step at 72°C for 5 min.
[b]Sequencing primer.
[c]See Ramaswamy *et al.*[15]

resistant isolates only and occurred at positions −8, −16, −25 and −32 bp. In one strain no other mutation in *embCAB* was identified, so this single nucleotide polymorphism (SNP) at −8 bp might confer ethambutol resistance.

Polymorphisms in the embA gene

Three ethambutol resistance-associated amino acid replacements were identified in two isolates (see Table 2 and Figure 1). One of these isolates carried two non-synonymous mutations located in codons 468 and 639. Furthermore a silent mutation in *embA* at position 988 bp was detected in this strain as well. Another isolate carried a mutation at codon 576. Three isolates had a silent mutation in *embA* at positions 1851, 1995 and 2124 bp, respectively. All strains, but one, carrying a mutation in *embA* had additional mutations in *embB*, *embC* or the intergenic region (see Table 2). In the ethambutol-susceptible isolates no resistance-associated mutations were detected in *embA* (except the phylogenetic markers discussed below).

Polymorphisms in the embB gene

The *embB* gene was most frequently affected by mutations resulting in amino acid exchanges. Overall non-synonymous

Table 2. Mutations detected in the *embCAB* operon of the 34 investigated ethambutol-resistant clinical *M. tuberculosis* isolates[a]

	Mutations in *embCAB*			
Total number of strains	*embC*	intergenic region *embCA*	*embA*	*embB*
1	c→T, 809 bp (Thr270Ile)	—	c→T, 988 bp (silent) t→C, 1403 bp (Val468Ala) c→T, 1915 bp (Pro639Ser)	a→C, 1133 bp (Glu378Ala) g→A, 1216 bp (Gly406Ser) c→T, 2982 bp (silent)
1	a→G, 1216 bp (Ile406Val)	—	—	—
1	g→A, 1239 bp (silent)	deletion G, −32 bp	—	a→G, 1490 bp (Gln497Arg)
1	g→A, 2213 bp (Arg738Gln)[b]	—	—	—
1	g→A, 2213 bp (Arg738Gln)[b]	deletion A, −25 bp	—	g→A, 1217 bp (Gly406Asp)
1	—	c→A, −8 bp	—	—
1	—	c→T, −8 bp	g→A, 1726 bp (Ala576Thr)	a→G, 1490 bp (Gln497Arg)
3	—	c→T, −12 bp	—	—
1	—	c→T, −12 bp	—	g→A, 1060 bp (Asp354Asn) g→A, 3081 bp (silent)
1	—	c→T, −16 bp	c→T, 2124 bp (silent)	g→C, 1217 bp (Gly406Ala) g→A, 1995 bp (silent)
1	—	—	a→G, 1851 bp (silent)	g→A, 1217 bp (Gly406Asp)
1	—	—	c→T, 1995 bp (silent)	—
1	—	—	—	t→G, 221 bp (Leu74Arg) a→G, 983 bp (Asp328Gly)
1	—	—	—	t→G, 221 bp (Leu74Arg) g→A, 1360 bp (Ala454Thr)
1	—	—	—	c→G, 1204 bp (Leu402Val)
1	—	—	—	c→T, 1210 bp (Pro404Ser)
3	—	—	—	g→A, 1217 bp (Gly406Asp)
1	—	—	—	g→C, 1217 bp (Gly406Ala)
1	—	—	—	c→G, 1350 bp (Ile450Met)
1	—	—	—	c→A, 1489 bp (Gln497Lys)
7	—	—	—	a→G, 1490 bp (Gln497Arg)
3	—	—	—	—

[a]Other phylogenetic mutations have been excluded from this table.
[b]The R738Q mutation has been identified as a phylogenetic marker for the Delhi/CAS genotype.

mutations in 10 distinct codons were found in 24 ethambutol-resistant isolates (see Table 2 and Figure 2). Three isolates even had two amino acid replacements in *embB* (see Table 2). One of these three strains also had a silent mutation at position 2982 bp. Silent mutations at positions 1995 and 3081 bp, respectively, were also identified in two isolates. The majority of the mutations were detected in *embB*406 ($n=8$) and *embB*497 ($n=10$). No mutations in *embB* were detected among ethambutol-susceptible isolates.

Combination of mutations

Overall nine isolates had more than one mutation in the *embCAB* region (see Table 2). In one isolate seven different mutations were detected; one non-synonymous mutation in *embC* and two non-synonymous mutations and one synonymous mutation in both *embA* and *embB*. Four strains were found to have two mutations resulting in an amino acid exchange. In two of these strains both mutations occurred in *embB* and the other isolates carried one mutation in *embB* as well as one in *embA* and *embC*, respectively. The other four strains had synonymous mutations or nucleotide variations in the intergenic region in addition to a non-synonymous mutation in *embB*.

Genotype-specific SNPs

Overall, three different mutations were found in both susceptible and resistant isolates and, thus, are potentially not involved in ethambutol resistance. In 22 ethambutol-resistant clinical isolates a synonymous mutation at nucleotide position 228 bp in *embA* was detected (see Figure 3). This mutation also occurred in six ethambutol-susceptible strains, suggesting that it might be a phylogenetic marker. A dendrogram calculated on the basis of IS*6110* DNA fingerprint and spoligotyping analyses revealed that all strains carrying the mutation at position 228 bp in *embA* belong to the Beijing genotype (Figure 3). Interestingly, nine of the Beijing strains that formed a sub-group within the Beijing branch had an additional mutation at position 114 bp in *embA*. Furthermore an amino acid replacement in *embC* at codon 981 was preferentially found in Haarlem genotype strains and might be specific for this phylogenetic lineage.

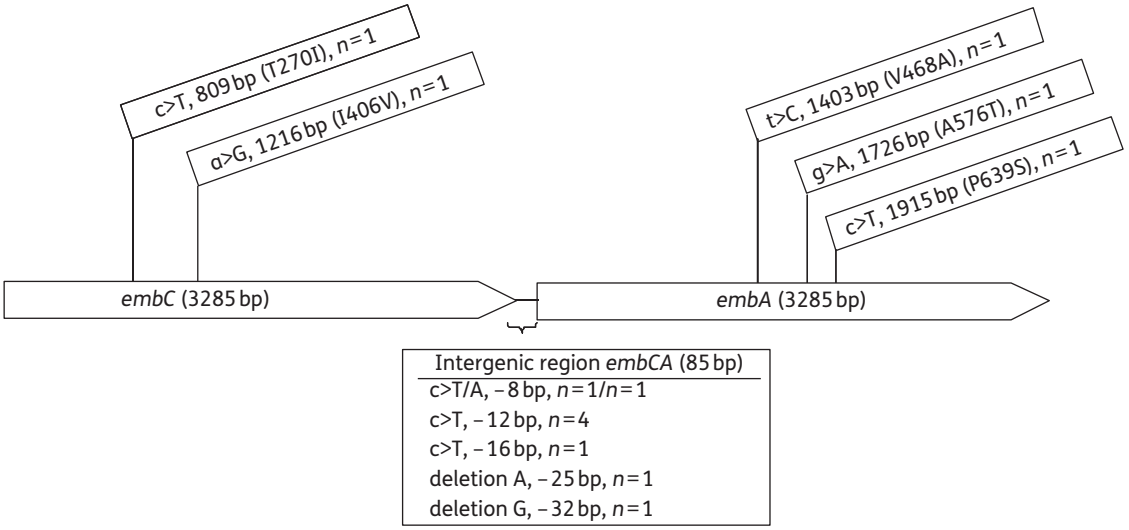

Figure 1. Non-synonymous mutations detected in *embC* and *embA* as well as in the intergenic region of the 34 ethambutol-resistant clinical *M. tuberculosis* isolates analysed. Two different mutations were found in *embC*. Out of the two *embC* mutants one isolate had no other mutation in the entire *embCAB* operon. In the *embA* gene three different mutations were found. Two of these mutations (V468A and P639S) occurred in one isolate. All of the *embA* mutants had additional mutations in *embC*, *embB* and/or the intergenic region *embCA*. Six different mutations in the intergenic region were identified in nine isolates of which two had no other mutations. Exchange of the base pairs as well as of the amino acids and the positions of nucleotides are indicated. The total number of strains in which the specific mutation was detected is also given.

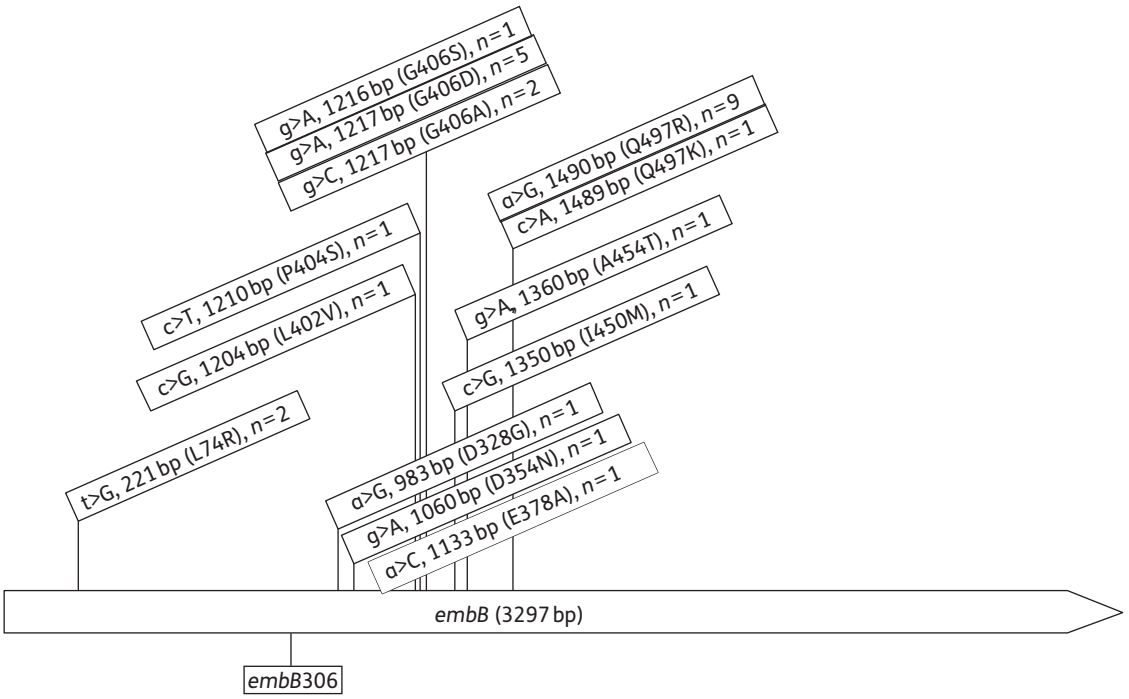

Figure 2. Non-synonymous mutations detected in the *embB* gene of the 34 ethambutol-resistant clinical *M. tuberculosis* isolates analysed. Overall, 24 of the ethambutol-resistant strains carried a mutation in *embB* resulting in an amino acid exchange. In three strains more than one non-synonymous mutation was found. The previously described hot spot codon *embB*306 is marked in a box. Exchange of the base pairs as well as of the amino acids and the positions of nucleotides are indicated. The total number of strains in which the specific mutation was detected is also given.

The mutation in *emb*C at codon 738 occurred in two ethambutol-resistant isolates. In one of those two strains this mutation was the only mutation detected in the whole *embCAB* operon, suggesting this SNP as a resistance-mediating polymorphism. However, IS*6110* DNA fingerprint and spoligotyping analyses showed that both strains belong to the Delhi/CAS genotype.

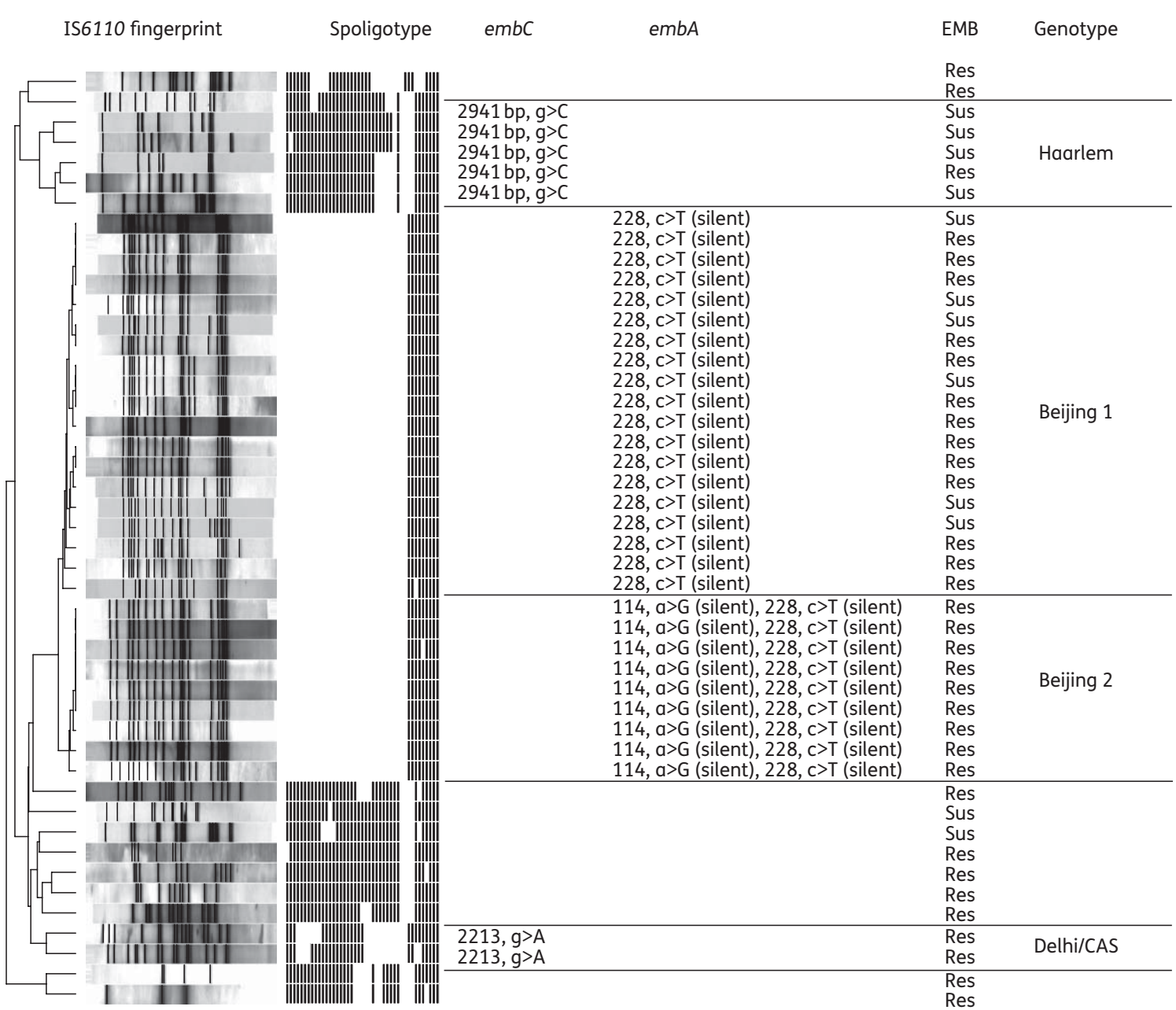

Figure 3. IS*6110* DNA fingerprint and spoligotype patterns as well as phylogenetic informative polymorphisms in *embA* and *embC* of the 46 strains investigated. The position of each IS*6110* band is normalized, so that banding patterns of all strains are mutually comparable. The strains' genotypes are ordered in a dendrogram based on the similarity of their IS*6110* DNA fingerprint patterns. EMB, ethambutol; Res, resistant; Sus, susceptible.

To further analyse the significance of Arg738Gln as a resistance-mediating SNP or phylogenetic marker onward sequencing work has been performed (see next paragraph).

Further verification of putative resistance-mediating mutations by analysis of 106 additional strains

The above section already demonstrated that SNPs are not necessarily involved in the development of resistance. Therefore we extended the investigation of selected mutations to 106 strains (104 ethambutol susceptible and 2 ethambutol resistant) from a population-based study collected in Hamburg during 2007. The population structure of these strains is highly diverse with an overall cluster rate of ~20% (21 strains in nine clusters). The majority of strains belong to the Haarlem genotype ($n=23$), followed by strains of the Beijing ($n=14$), East African Indian ($n=7$), Latin American Mediterranean ($n=5$), Cameroon ($n=4$) and Delhi/CAS ($n=4$) genotypes. The remaining 49 strains belong to various other genotypes or could not be allocated to previously described *M. tuberculosis* complex phylogenetic lineages.

First of all, the mutation at *embC*Arg738Gln was chosen to be screened in the panel of the above-mentioned strains. Four strains were found to have mutations in *embC* at codon 738, with only one of these an ethambutol-resistant strain, suggesting that this mutation may be a phylogenetic marker rather than associated directly with ethambutol resistance. To test this hypothesis, IS*6110* DNA fingerprint and spoligotyping patterns of the strains were compared. The analysis revealed that all strains belonging to the Delhi/CAS genotype carry the *embC*Arg738Gln mutation while strains of all other genotypes do not, irrespective of being susceptible or resistant to ethambutol. Hence it can be concluded that *embC*738 is a phylogenetic SNP for the Delhi/CAS genotype instead of an ethambutol resistance marker.

For verification of the other *embCAB* mutations occurring in ethambutol-resistant strains only, the *embCA* intergenic region (nucleotides −8, −12, −16, −25 and −32) as well as the *embB* hotspot region (codons 306–497, namely codons 328, 354, 378, 402, 404, 450 and 454) have been sequenced in the 106 additional strains as well. Sequence analyses showed that none of the susceptible strains analysed in the control panel carried a mutation in any of those positions, confirming a role for the respective SNPs in ethambutol resistance development. In addition, IS*6110* DNA fingerprint and spoligotyping analyses of the ethambutol-resistant strains revealed that those mutations analysed occurred in strains belonging to various genotypes (when occurring in more than one strain), further affirming their potential role in ethambutol resistance mediation (data not shown).

Discussion

Mutations in *embB*306 were found to be associated with ethambutol resistance in previous studies and polymorphisms in this codon are suggested as molecular markers for the rapid detection of ethambutol-resistant isolates.[3,4,6] However, between 31% and 50% of ethambutol-resistant clinical isolates have no mutation in the *embB*306 region suggesting a role for variations in other parts of the *embCAB* operon or other genomic regions for the development of ethambutol resistance. In this study we demonstrate that a high proportion (91.2%) of the 34 ethambutol-resistant strains without an *embB*306 polymorphism had other mutations in the remaining *embCAB* operon. Overall, we identified 18 non-synonymous mutations in 15 distinct codons of the *embCAB* operon that were exclusively found in the ethambutol-resistant strains but not in the ethambutol-susceptible control isolates.

The majority of mutations occurred in *embB*, supporting the idea that mutations in this gene play a key role in the development of ethambutol resistance. A total of 24 ethambutol-resistant strains investigated carried at least one non-synonymous mutation in *embB*. In accordance with previous studies, multiple mutations resulting in two or three different amino acid exchanges were identified in codons *embB*406 and *embB*497 representing additional hot spots for mutation besides *embB*306.[3,13–15]

The observation that multiple mutations affect the same codon resulting in different amino acid replacements is one example of positive Darwinian selection by antibiotic pressure. In a previous study, strains with non-synonymous mutations in *embB*406 and *embB*497 were correlated with high ethambutol MICs (≥30 mg/L) comparable to those for organisms carrying *embB*306 mutations.[2,7]

In the current study, further mutations were detected in *embB* codons 74, 328, 354, 378, 402, 404, 450 and 454. To our knowledge the polymorphisms at codons 74, 402, 404, 450 and 454 have not been described previously. Sequence variations in *embB*354 have already been reported before, but in contrast to our study a replacement of aspartic acid by alanine instead of asparagine was observed.[8,16] The mutations at codons 328 and 378 have also been described previously.[9,15,16] Since all mutations located in the *embB* hotspot region (codons 306–497) have not been detected in the panel of ethambutol-susceptible control strains, they possibly play a role in ethambutol resistance development.

Few publications provide data about the correlation between ethambutol MICs and these less frequent distinct mutations. The ethambutol MIC for a mutant carrying a mutation at codon 354 (Asp→Ala) was shown to be 20 mg/L,[8] while a mutation at codon *embB*328 has been shown to yield an ethambutol MIC ranging between 20 and 30 mg/L.[8,15] For strains harbouring a Glu378Ala amino acid exchange, an ethambutol MIC of 2 mg/L has previously been determined.[9] These MIC values were obtained from the analysis of small numbers of strains, carrying additional mutations, which might also be responsible for the MIC increase.

It is noticeable that in the current study the majority of the detected mutations were concentrated in a small region of ~570 bp in the *embB* gene. This region of the EmbB protein, including the hot spot codons 306, 406 and 497, is predicted to be the recognition site of the enzyme.[17] Only two strains had mutations in *embB* outside the binding site, but these strains carried additional mutations. These data suggest that only mutations located in the recognition site are associated with ethambutol resistance. Structural changes in the enzyme, caused by mutations in the recognition site, may prevent binding of ethambutol leading to resistance development. The finding that mutations in ethambutol-resistant clinical isolates without an *embB*306 mutation were concentrated in a

~570 bp encompassing region has practical implications for TB diagnostics. Analysing this part of the *embB* gene would probably enhance the sensitivity for the detection of ethambutol-resistant isolates.

In two previous studies,[5,18] we investigated a total of 150 ethambutol-resistant strains for variation in the *embB*306 region. Overall 72.7% of the ethambutol-resistant isolates could be detected by sequence variations in *embB*306. Additionally, seven strains were found to have other mutations (codons 319, 328, 332 and 334) in the analysed *embB* fragment. A short extension of the sequence analysis to a ~600 bp *embB* fragment (codons 297–497) would have enhanced the overall sensitivity of the molecular analysis for detection of ethambutol resistance to 93.3%, which is comparable to the data described for other first-line drugs such as isoniazid, rifampicin or pyrazinamide.[14]

However, considering the overall broad spectrum of observed mutations, it will be necessary to validate the potential of particular mutations to confer (alone or in combination) significant levels of ethambutol resistance by further MIC measurements on more strains. This is especially important in light of the increasing usage of molecular tests for the detection of drug resistance in high-incidence settings.

Interestingly, mutations in *embC* and *embA* occurred rarely and in most cases were combined with mutations in *embB*, suggesting that mutations in these genes are less important for ethambutol resistance. Apparently the arabinosyltransferases EmbC and EmbA are less susceptible to the inhibiting effect of ethambutol.[19–22] The differential mode of action of ethambutol on the Emb proteins can be explained by their involvement in different pathways. While EmbC is involved in the formation of arabinan in lipoarabinomannan, the enzymes EmbA and EmbB participate in the synthesis of arabinan in arabinogalactan.[23,24] The reason for the lower susceptibility of EmbA in comparison with EmbB, although participating in the same biosynthetic pathway, is not clear. It can be speculated that both Emb proteins function as a dimer and EmbB is the more important part of this dimer or that EmbB compensates an inhibiting effect of ethambutol on EmbA.[24]

The detected mutations in *embC* were located in codons 270 and 406 and in two strains in 738. To our knowledge the mutation in *embC*406 has not been described so far. No other mutations were found in this strain suggesting that sequence variations in *embC*406 may confer an ethambutol resistance phenotype. The mutations in *embC*270 and *embC*738 are already known from previous studies.[9,15,25,26] In one study, an ethambutol MIC of 40 mg/L for a strain carrying a mutation at codon 738 was reported.[15] Since this strain had additional mutations (e.g. *embB*306) it is unclear whether *embC*738 actually contributes to the resistance phenotype. The additional sequence analysis performed in this study resulted in three susceptible strains and one resistant strain carrying the *embC*738 mutation. The fact that all of these strains belong to the Delhi/CAS genotype and that strains belonging to other genotypes do not carry the *embC*738 mutation, irrespective of being susceptible or resistant to ethambutol, clearly demonstrates that this mutation is a newly identified phylogenetic marker instead of being associated with ethambutol resistance.

The strain carrying a mutation in *embC*270 was found to have additional mutations. A poor correlation between *embC*270 mutation and the degree of ethambutol resistance has been observed previously.[2,9,25,26] As varying results for ethambutol MIC determination among *embC*270 mutant strains carrying additional mutations in *embB* have been reported, it is not clear if there is an additive effect. The combination of *embB*306 and *embC*270 mutations may increase the ethambutol MIC,[9,25,26] but further investigations are necessary.

Mutations in *embA* were detected in codons 468, 576 and 639. None of them has been described previously. The mutations in codons 468 and 639 occurred together in one isolate that had additional mutations in *embC* and *embB*. The strain showing the mutation in codon 576 carried further mutations in the intergenic region and in *embB*. As the *embA* mutations appeared only in combination with other mutations their relevance for ethambutol resistance is unclear and needs to be further investigated.

Novel mutations in the *embC–embA* intergenic region were also found. Besides the known nucleotide variations, which were located 12 and 16 bp upstream of the *embA* gene, mutations were detected at positions −8, −25 and −32 bp.[15] The nucleotide substitutions at positions −8, −12 and −16 bp are located within/adjacent to a predicted TATA box.[15] Mutations in this promoter region may be involved in the development of ethambutol resistance by increasing the gene expression of *embA* and/or *embB*. With respect to the additional sequence analysis of the *embCA* intergenic region none of the susceptible strains analysed carried a mutation at those specific positions. Therefore, these results strengthen the hypothesis that those mutations, alone or in combination, contribute to the development of ethambutol resistance.

In addition to the *embC*738 mutation, we identified further genotype-specific mutations in the *embC* and *embA* genes. In *embA* two synonymous mutations were associated with the Beijing genotype and a sub-group of the Beijing genotype, respectively. The mutation in *embC* at codon 981 was already mentioned previously and suggested to be a phylogenetic rather than a resistance marker.[2,15] A total of eight ethambutol-resistant isolates were without non-synonymous mutations and three of these did not have any mutations in the entire *embCAB* operon. In these cases, mutations in other so far unknown genes might be involved in the development of ethambutol resistance.

In conclusion, this study demonstrates a strikingly high variety of mutations in *embCAB* that might have a potential role in the development of ethambutol resistance in clinical *M. tuberculosis* complex isolates and may be valuable markers for the identification of strains belonging to different genotypes. Our results clearly demonstrate that a short stretch of the *embB* gene encompasses the most important mutation hot spots (codons 306, 406 and 497). Although more data from other high TB drug resistance settings are required, this mutational hot spot region is of high interest for inclusion in molecular assays.

Our data also demonstrate that results from molecular association studies should be interpreted carefully. Large, well-chosen control groups are necessary to allow conclusions concerning a possible involvement in development of resistance. In addition, further validation should be sought by *in vitro* selection experiments or by generation of isogenic mutants via allelic exchange.

Acknowledgements

We would like to thank I. Radzio, T. Ubben, L. Dost and P. Vock for excellent technical assistance.

Funding

This work was supported by the European Union TM-REST (FP7-202145) and the TB-PAN-NET (FP7-223681) projects.

Transparency declarations

None to declare.

References

1 WHO. *Anti-Tuberculosis Drug Resistance in the World. Report Number 4*. Geneva: WHO, 2008.

2 Sreevatsan S, Stockbauer KE, Pan X *et al*. Ethambutol resistance in *Mycobacterium tuberculosis*: critical role of *embB* mutations. *Antimicrob Agents Chemother* 1997; **41**: 1677–81.

3 Lee AS, Othman SN, Ho YM *et al*. Novel mutations within the *embB* gene in ethambutol-susceptible clinical isolates of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 2004; **48**: 4447–9.

4 Mokrousov I, Otten T, Vyshnevskiy B *et al*. Detection of *embB306* mutations in ethambutol-susceptible clinical isolates of *Mycobacterium tuberculosis* from northwestern Russia: implications for genotypic resistance testing. *J Clin Microbiol* 2002; **40**: 3810–3.

5 Plinke C, Rusch-Gerdes S, Niemann S. Significance of mutations in *embB* codon 306 for prediction of ethambutol resistance in clinical *Mycobacterium tuberculosis* isolates. *Antimicrob Agents Chemother* 2006; **50**: 1900–2.

6 van Rie A, Warren R, Mshanga I *et al*. Analysis for a limited number of gene codons can predict drug resistance of *Mycobacterium tuberculosis* in a high-incidence community. *J Clin Microbiol* 2001; **39**: 636–41.

7 Safi H, Sayers B, Hazbon MH *et al*. Transfer of *embB* codon 306 mutations into clinical *Mycobacterium tuberculosis* strains alters susceptibility to ethambutol, isoniazid, and rifampin. *Antimicrob Agents Chemother* 2008; **52**: 2027–34.

8 Starks AM, Gumusboga A, Plikaytis BB *et al*. Mutations at *embB* codon 306 are an important molecular indicator of ethambutol resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 2009; **53**: 1061–6.

9 Srivastava S, Garg A, Ayyagari A *et al*. Nucleotide polymorphism associated with ethambutol resistance in clinical isolates of *Mycobacterium tuberculosis*. *Curr Microbiol* 2006; **53**: 401–5.

10 Wu XQ, Liang JQ, Zhang JX *et al*. Detection and evaluation of the mutations of *embB* gene in ethambutol-susceptible and resistant *Mycobacterium tuberculosis* isolates from China. *Chin Med J (Engl)* 2005; **118**: 1739–41.

11 Van Embden JD, Cave MD, Crawford JT *et al*. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol* 1993; **31**: 406–9.

12 Kamerbeek J, Schouls L, Kolk A *et al*. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* 1997; **35**: 907–14.

13 Ahmad S, Jaber AA, Mokaddas E. Frequency of *embB* codon 306 mutations in ethambutol-susceptible and -resistant clinical *Mycobacterium tuberculosis* isolates in Kuwait. *Tuberculosis (Edinb)* 2007; **87**: 123–9.

14 Parsons LM, Somoskovi A, Urbanczik R *et al*. Laboratory diagnostic aspects of drug resistant tuberculosis. *Front Biosci* 2004; **9**: 2086–105.

15 Ramaswamy SV, Amin AG, Goksel S *et al*. Molecular genetic analysis of nucleotide polymorphisms associated with ethambutol resistance in human isolates of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 2000; **44**: 326–36.

16 Sekiguchi J, Miyoshi-Akiyama T, Augustynowicz-Kopec E *et al*. Detection of multidrug resistance in *Mycobacterium tuberculosis*. *J Clin Microbiol* 2007; **45**: 179–92.

17 Zhang N, Torrelles JB, McNeil MR *et al*. The Emb proteins of mycobacteria direct arabinosylation of lipoarabinomannan and arabinogalactan via an N-terminal recognition region and a C-terminal synthetic region. *Mol Microbiol* 2003; **50**: 69–76.

18 Plinke C, Cox HS, Kalon S *et al*. Tuberculosis ethambutol resistance: concordance between phenotypic and genotypic test results. *Tuberculosis (Edinb)* 2009; **89**: 448–52.

19 Deng L, Mikusova K, Robuck KG *et al*. Recognition of multiple effects of ethambutol on metabolism of mycobacterial cell envelope. *Antimicrob Agents Chemother* 1995; **39**: 694–701.

20 Goude R, Amin AG, Chatterjee D *et al*. The critical role of *embC* in *Mycobacterium tuberculosis*. *J Bacteriol* 2008; **190**: 4335–41.

21 Khoo KH, Douglas E, Azadi P *et al*. Truncated structural variants of lipoarabinomannan in ethambutol drug-resistant strains of *Mycobacterium smegmatis*. Inhibition of arabinan biosynthesis by ethambutol. *J Biol Chem* 1996; **271**: 28682–90.

22 Mikusova K, Slayden RA, Besra GS *et al*. Biogenesis of the mycobacterial cell wall and the site of action of ethambutol. *Antimicrob Agents Chemother* 1995; **39**: 2484–9.

23 Escuyer VE, Lety MA, Torrelles JB *et al*. The role of the *embA* and *embB* gene products in the biosynthesis of the terminal hexaarabinofuranosyl motif of *Mycobacterium smegmatis* arabinogalactan. *J Biol Chem* 2001; **276**: 48854–62.

24 Shi L, Berg S, Lee A *et al*. The carboxy terminus of EmbC from *Mycobacterium smegmatis* mediates chain length extension of the arabinan in lipoarabinomannan. *J Biol Chem* 2006; **281**: 19512–26.

25 Jadaun GP, Das R, Upadhyay P *et al*. Role of *embCAB* gene mutations in ethambutol resistance in *Mycobacterium tuberculosis* isolates from India. *Int J Antimicrob Agents* 2009; **33**: 483–6.

26 Srivastava S, Ayyagari A, Dhole TN *et al*. *emb* nucleotide polymorphisms and the role of *embB*306 mutations in *Mycobacterium tuberculosis* resistance to ethambutol. *Int J Med Microbiol* 2009; **299**: 269–80.