**embCAB** sequence variation among ethambutol-resistant *Mycobacterium tuberculosis* isolates without embB306 mutation

Claudia Plinke1, Helen S. Cox2, Nana Zarkua3, Hamraev A. Karimovich4, Kai Braker5, Roland Diel6, Sabine Rüsch-Gerdes7, Silke Feuerriegel1*† and Stefan Niemann1†

1Molecular Mycobacteriology, Research Center Borstel, Borstel, Germany; 2Macfarlane Burnet Institute for Medical Research and Public Health, Melbourne, Australia; 3 Médecins Sans Frontières, Tashkent, Uzbekistan; 4Ministry of Health, Nukus, Karakalpakstan, Uzbekistan; 5 Médecins Sans Frontières, Berlin, Germany; 6Department of Pneumology, Medical School Hannover (MHH), Hannover, Germany; 7National 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 embB306 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 embB306. 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 (embCVal981Leu and embCArg738Gln) 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 embB306, mutations in embB406 and embB497 were confirmed as hotspots 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 embB306 occurring most frequently.2 Consequently mutations in this codon have been suggested as molecular markers for the rapid detection of ethambutol resistance.1–6 Furthermore, two recent studies demonstrated a clear causative relationship between embB306 point mutations and in vitro ethambutol resistance by allelic-exchange experiments.7,8

© The Author 2010. Published by Oxford University Press on behalf of the British Society for Antimicrobial Chemotherapy. All rights reserved. For Permissions, please e-mail: journals.permissions@oxfordjournals.org
However, approximately one-third of ethambutol-resistant clinical isolates do not carry a mutation in embB306 and are therefore not detectable by using molecular methods based on the determination of polymorphisms in embB306 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 embB306, 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) clinical isolates collected in Hamburg in 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 IS6110 as a probe were performed according to a standardized protocol as described previously.\(^1\)\(^1\)\(^1\)\(^2\) Additionally, all isolates were analysed by the spoligotyping technique as described previously by Kamerbeek et al.\(^1\)\(^2\)\(^2\) 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 embCArg738Gln 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 3130xI 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 embB306 (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...
resistant isolates only and occurred at positions 28, 216, 225 and 232 bp. In one strain no other mutation in embCAB was identified, so this single nucleotide polymorphism (SNP) at 28 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 1. PCR primers and conditions used for amplification and sequencing

<table>
<thead>
<tr>
<th>Primer</th>
<th>Primer sequence (5′→3′)</th>
<th>Size (bp)</th>
<th>D (s)</th>
<th>A (°C, s)</th>
<th>E (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rv3795anewF</td>
<td>CTG GGG ATC GGT GGA GCA GTA</td>
<td>969</td>
<td>30</td>
<td>66.0, 45</td>
<td>60</td>
</tr>
<tr>
<td>Rv3795anewR</td>
<td>GCG TCG GTC AGG GTG TGA A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv3795bF</td>
<td>TGG ACG GGC GGG GCT CAA T</td>
<td>1306</td>
<td>60</td>
<td>65.0, 60</td>
<td>90</td>
</tr>
<tr>
<td>Rv3795bCR</td>
<td>GCA AAC AGG GGC AAA AAG A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv3795cF</td>
<td>TTC TGG CCG GTG TAT TCT T</td>
<td>1653</td>
<td>60</td>
<td>65.0, 60</td>
<td>90</td>
</tr>
<tr>
<td>Rv3795cS1a</td>
<td>TGG ACG GGC ATT CGG GTT CT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv3795cS2a</td>
<td>GGA CTG GGC GGT CGG TTG G</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv3795cR</td>
<td>CAA CGG GGG TGA TGA TGG C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embIR-F</td>
<td>CTG GTG GTC GGG GTG ATG AT</td>
<td>907</td>
<td>30</td>
<td>64.0, 45</td>
<td>60</td>
</tr>
<tr>
<td>embIR-newR</td>
<td>ACG GTC GCT GGC AGG GGA AGT T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embIR-newF</td>
<td>CGC CTA TGA CCC GAA CCT GAG</td>
<td>787</td>
<td>30</td>
<td>65.0, 45</td>
<td>60</td>
</tr>
<tr>
<td>embIR-R</td>
<td>AAT GCG CCT TGC CTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embA1-F</td>
<td>GTG ACT CGC AGG GGC CTG G</td>
<td>1223</td>
<td>30</td>
<td>68.0, 30</td>
<td>90</td>
</tr>
<tr>
<td>embA1-R</td>
<td>CGG TGA ACA CAG CGA CCC G</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embA2-F</td>
<td>TGG ACC GGC TCA GCA GGG G</td>
<td>1500</td>
<td>30</td>
<td>67.0, 30</td>
<td>90</td>
</tr>
<tr>
<td>embA2-R</td>
<td>TCA GGT TGG CCT TGG CGG T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embC1-F</td>
<td>CCC AAC CAG CCC AAT GTT C</td>
<td>890</td>
<td>40</td>
<td>64.5, 30</td>
<td>60</td>
</tr>
<tr>
<td>embC1-R</td>
<td>GCC GGT GTC CAG GAT GTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embC2-F</td>
<td>GCT GCA CAT CCT GGA CAC</td>
<td>914</td>
<td>40</td>
<td>60.0, 30</td>
<td>60</td>
</tr>
<tr>
<td>embC2-R</td>
<td>ACG ACA TTG CCA CCG ATA C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embC3-F</td>
<td>GTA TGG GTG GCA ATG TTC T</td>
<td>1176</td>
<td>40</td>
<td>60.0, 30</td>
<td>60</td>
</tr>
<tr>
<td>embC3-R</td>
<td>CGG GAT GGC GGA CAG TGG T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embC4-F</td>
<td>ACC ACT GTT CCG CAT CCC G</td>
<td>635</td>
<td>30</td>
<td>67.0, 30</td>
<td>45</td>
</tr>
<tr>
<td>embC4-R</td>
<td>GAC GAC GTC GTC GAG TGG T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embC1-F2</td>
<td>GAA GCC GCC CCA CCC CGT ATC G</td>
<td>996</td>
<td>30</td>
<td>68.0, 45</td>
<td>60</td>
</tr>
<tr>
<td>embC1-R2</td>
<td>AGC CCA CAG CGC CAG GTC GTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embC2-F2</td>
<td>GCG GGC ATG TTT CTG GTC GTC TGG</td>
<td>1058</td>
<td>30</td>
<td>64.0, 45</td>
<td>60</td>
</tr>
<tr>
<td>embC2-R2</td>
<td>CAC GCC GGG TAC TGG GAA ATC A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embC3-F2</td>
<td>CGC GCC GCG CTG CCC TAC A</td>
<td>464</td>
<td>30</td>
<td>68.0, 45</td>
<td>60</td>
</tr>
<tr>
<td>embC3-R2</td>
<td>GCC AGC CCC ACC AGC CAG TCC A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embC_5 (+2002)</td>
<td>CTG ACG GTG CTG CTG CTG CTG CTG</td>
<td>372</td>
<td>30</td>
<td>65.0, 30</td>
<td>30</td>
</tr>
<tr>
<td>embC_3 (+2373)</td>
<td>GGG AAT GCC GTG GGG TGT GAA GG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>embCA_5 (+3134)</td>
<td>TGG CCA GTC ACC TCA AAG ACG ACT</td>
<td>437</td>
<td>30</td>
<td>60.0, 30</td>
<td>30</td>
</tr>
<tr>
<td>embCA_3 (+200)</td>
<td>GCC GCC CCG GAT ACC AGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*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.*

*Sequencing primer.

*See Ramaswamy et al.*15
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 embB406 (n=8) and embB497 (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 IS6110 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.
The mutation in embC 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, IS6110 DNA fingerprint and spoligotyping analyses showed that both strains belong to the Delhi/CAS genotype.
**Figure 3.** IS6110 DNA fingerprint and spoligotype patterns as well as phylogenetic informative polymorphisms in embA and embC of the 46 strains investigated. The position of each IS6110 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 IS6110 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 Mycobacterium tuberculosis other genotypes (when occurring in more than one strain), further affirming the idea that mutations in this gene play a key role in the development of ethambutol resistance. 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 hypothesis 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 embB406 and embB497 representing additional hot spots for mutation besides embB306.

The observation that multiple mutations affect the same codon result 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 embB406 and embB497 were correlated with high ethambutol MICs (≥30 mg/L) comparable to those for organisms carrying embB306 mutations.

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 embB304 have already been reported before, but in contrast to our study a replacement of aspartic acid by alanine instead of asparagine was observed. The mutations at codons 328 and 378 have also been described previously. 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, while a mutation at codon embB328 has been shown to yield an ethambutol MIC ranging between 20 and 30 mg/L. For strains harbouring a Glu378Ala amino acid exchange, an ethambutol MIC of 2 mg/L has previously been determined. 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. 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 embB306 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, we investigated a total of 150 ethambutol-resistant strains for variation in the embB306 region. Overall 72.7% of the ethambutol-resistant isolates could be detected by sequence variations in embB306. 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.

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. 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. 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.

The detected mutations in embC were located in codons 270 and 406 and in two strains in 738. To our knowledge the mutation in embC406 has not been described so far. No other mutations were found in this strain suggesting that sequence variations in embC406 may confer an ethambutol resistance phenotype. The mutations in embC270 and embC378 are already known from previous studies. In one study, an ethambutol MIC of 40 mg/L for a strain carrying a mutation at codon 738 was reported. Since this strain had additional mutations (e.g. embB306) it is unclear whether embC738 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 embC738 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 embC738 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 embC270 was found to have additional mutations. A poor correlation between embC270 mutation and the degree of ethambutol resistance has been observed previously. As varying results for ethambutol MIC determination among embC270 mutant strains carrying additional mutations in embB have been reported, it is not clear if there is an additive effect. The combination of embB306 and embC270 mutations may increase the ethambutol MIC 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. The nucleotide substitutions at positions –8, –12 and –16 bp are located within/adjacent to a predicted TATA box. 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 embC738 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 embA at codon 981 was already mentioned previously and suggested to be a phylogenetic marker rather than a resistance marker. 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