Antibiotic Resistance and SNP Cluster Grouping Type

In a Multi-national Sample of M. tuberculosis Resistant Isolates

M. Brimacombe1*, M. Hazbon2, A.S. Motiwala2, D. Alland2
Department of Preventive Medicine and Community Health1 and Department of Medicine2, New Jersey Medical School - UMDNJ

* corresponding author.

Running title: Antibiotic resistance and SNP cluster grouping.

Word Count (Abstract): 50
Word Count (Text): 981
ABSTRACT

A single nucleotide polymorphism-based cluster grouping (SCG) classification system for *Mycobacterium tuberculosis* is used to examine antibiotic resistance type and resistance mutations in relationship to specific evolutionary lineages. Drug resistance and resistance mutations were seen across all SCG. SCG2 had higher proportions of *katG*315 mutations and resistance to four drugs.
INTRODUCTION

Isoniazid (INH) is an effective agent for treatment of infections with *Mycobacterium tuberculosis*. Increases in INH-resistant and multi drug-resistant (MDR) tuberculosis jeopardize drug effectiveness (7, 24) and development of INH-resistance is often a first step in MDR (3, 8). Mutations in specific genes have been linked to INH related resistance (10, 11), including *katG* (26), *inhA* (14), codon 315 of *katG*, (13, 16, 19, 23), *ahpC* (20), the *inhA* open reading frame (ORF) and *inhA* promoter (14, 17, 25) and *ndh* (22).

Recent studies of drug-resistant *M. tuberculosis* have found associations among *M. tuberculosis* strains, drug-resistance and specific gene mutations. *M. tuberculosis* strains belonging to the Beijing family were associated with drug-resistance in Iran, Afghanistan and Russia (15, 18, 12, 6), though not in Venezuela (2). These associations have also been supported through genetic laboratory studies (1, 21). Thus, it is possible that specific types of drug resistance or drug resistance mutations might occur more commonly in certain evolutionary lineages of *M. tuberculosis*.

The single nucleotide polymorphism cluster group (SCG) classification system defined in (8) gives rise to seven phylogenetically distinct groups and three subgroups that can be used to infer evolutionary pattern in *M. tuberculosis*. Here, we analyze 428 *M. tuberculosis* isolates resistant to at least INH collected across 10 countries and report the prevalence of various INH resistance-associated mutations and prevalences of two, three and four-drug resistance according to the major SCG-defined phylogenetic lineages of *M. tuberculosis*. 
METHODS

*M. tuberculosis* isolates resistant to at least INH were obtained from laboratories in major medical centers in Australia, Colombia, India, Mexico, New York City, Spain and Texas (Table 1). The study population and sample selection has been described (10, 11) with each collection site performing susceptibility testing to INH, rifampin (RIF), streptomycin (STR) and ethambutol (EMB).

Isolates were tested for virtually all single nucleotide polymorphism (SNP) mutations in the *M. tuberculosis* *katG*, *kasA*, *mabA*, *inhA*, *oxyR*, *ahpC* and *ndh* genes found to be associated with INH-resistance in published studies (11). A total of 204 INH-resistance associated alleles were detected (9) with confirmatory testing carried out on alleles, identified mutations and drug-resistance.

Strain types are reported in terms of SCG grouping, based on observed genomic level clustering among identified SNPs (8). SCG assignment was performed by testing each isolate for nine SNPs, previously determined to replicate larger SNP-based phylogeny as described in (4, 5).

The genetic, resistance and SCG data from each set of country specific isolates was entered into a common database. Prevalence was reported by SCG type, country, selected genes, and resistant type. Chi-square tests of differences in proportions were employed as appropriate. The STATA statistical package was used for all calculations.

Results

The complete sample was tested for 240 alleles previously reported to be associated with INH-resistance. Country specific breakdowns by various types of resistance can be
found (10, 11). The overall prevalence of SCG groups is given in Table 1. The Beijing family (strongly associated with SCG-2) is present in 8.9% of the collected isolates. Most prevalent SCG types were SCG-5, SCG-3b and SCG-6a and rarest (excluding \textit{M. bovis}) SCG-6b, SCG-1, SCG-3c and SCG-4. SCG prevalence by country shows varying distributions. SCG-5 type is among the three most prevalent SCG types in Spain, Columbia, Mexico and Texas. SCG-3b (Mexico, Columbia, Texas), SCG-6a (Texas, Spain, Mexico). SCG-3a (India). SCG-2 (Australia, NYC) perhaps reflecting Asian immigrant populations. SCG-3c, SCG-4, SCG-6b had low prevalence in all countries.

Table 2 reports prevalence of mutations in a selected set of genes. Only KatG and KatG315 mutations occur across all SCG types (excluding \textit{M. bovis}). SCG types with highest number of mutations were SCG-5, SCG-3b and SCG-1. SCG-1 and SCG-5 had all mutations of interest. \textit{kasA} and \textit{ndh} mutations were found in the fewest number of SCG types (three).

As these measures could be influenced by the proportional representation of each SCG in the sample, we examined the relative proportion of each mutation within each SCG (Fig. 1). The highest proportion of KatG315 mutations occurred in SCG-3c and SCG-2. SCG-6b, SCG-2 and SCG-3c had the highest proportion of mutations in \textit{katG}. Mutations in \textit{inhA} and \textit{ahpC} promoters were found in all SCGs with n>17.

To examine antibiotic resistance on a cumulative scale, we examined resistance within each SCG classification type in relation to one, two or more resistant antibiotics. All SCG types (excepting \textit{M. bovis}) had at least 46% of isolates with resistance to 1 or 2 of these antibiotics. The range across all SCG types was 46% to 82%. SCG-2 had the highest prevalence (29%) of resistance to all four antibiotics.
DISCUSSION

This study provides insights into the association of drug-resistant *M. tuberculosis*, gene mutation and genomic SCG based phylogenetic lineage, utilizing a large number of resistant isolates and examining patterns of relationships across SCG classification type, resistance, gene mutation and country.

The sample contained 8.9% SCG-2 (Beijing) isolates, similar to Asian samples (18) and comparable to other SCG types excepting SCG-3b, SCG-5 and SCG-6a. SCG-2 isolates accounted for more than 10% of the KatG and KatG315 mutations in this study. Most mutations were prevalent in all SCGs.

The prevalence of resistance is high across all SCG types. Table 3 shows SCG-2 with the highest prevalence of resistance to all antibiotics, but all SCG types display high resistance levels to one and two antibiotics. SCG-6a displayed high prevalence of isolates resistant to all four antibiotics. As the SCG classification reflects regions, the presence of antibiotic resistance in so many SCG classifications may reflect several ongoing evolutionary processes and implies a need to maintain a broad perspective on *M. tuberculosis* antibiotic resistance.

Limitations

This study analyzed *M. tuberculosis* isolates for mutations associated with INH-resistance in previous studies. The results therefore may not be completely representative of mutations relevant to resistance. The overall patterns observed in the SCG classification may have been restricted by method of sample selection, but size and diversity of the sample makes this unlikely.
ACKNOWLEDGEMENTS

This work was supported by Public Health Service grants AI-46669 and AI-49352.
REFERENCES


Table 1: *M. tuberculosis* SCG Types Overall and by Country (Count (%))

<table>
<thead>
<tr>
<th></th>
<th>M. bovis</th>
<th>SCG 1</th>
<th>SCG 2</th>
<th>SCG 3a</th>
<th>SCG 3b</th>
<th>SCG 3c</th>
<th>SCG 4</th>
<th>SCG 5</th>
<th>SCG 6a</th>
<th>SCG 6b</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>0</td>
<td>9 (30.0)</td>
<td>11 (36.7)</td>
<td>3 (10.0)</td>
<td>3 (10.0)</td>
<td>0</td>
<td>0</td>
<td>3 (10.0)</td>
<td>0</td>
<td>1 (3.3)</td>
<td>30 (100)</td>
</tr>
<tr>
<td>Columbia</td>
<td>0</td>
<td>0</td>
<td>1 (0.7)</td>
<td>33 (23.1)</td>
<td>1 (0.7)</td>
<td>6 (4.2)</td>
<td>84 (58.7)</td>
<td>17 (11.9)</td>
<td>1 (0.7)</td>
<td>143 (100)</td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>0</td>
<td>2 (6.7)</td>
<td>1 (3.3)</td>
<td>23 (76.7)</td>
<td>1 (3.3)</td>
<td>0</td>
<td>2 (6.7)</td>
<td>1 (3.3)</td>
<td>0</td>
<td>30 (100)</td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td>2 (1.6)</td>
<td>0</td>
<td>1 (0.8)</td>
<td>44 (35.8)</td>
<td>11 (8.9)</td>
<td>14 (11.4)</td>
<td>31 (25.2)</td>
<td>15 (12.2)</td>
<td>4 (3.3)</td>
<td>123 (100)</td>
<td></td>
</tr>
<tr>
<td>NYC</td>
<td>0</td>
<td>1 (3.9)</td>
<td>17 (65.3)</td>
<td>0</td>
<td>2 (7.7)</td>
<td>1 (3.9)</td>
<td>0</td>
<td>2 (7.7)</td>
<td>3 (11.5)</td>
<td>0</td>
<td>26 (100)</td>
</tr>
<tr>
<td>Spain</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1 (9.1)</td>
<td>0</td>
<td>0</td>
<td>8 (72.7)</td>
<td>2 (18.2)</td>
<td>0</td>
<td>11 (100)</td>
<td></td>
</tr>
<tr>
<td>Texas</td>
<td>0</td>
<td>5 (7.7)</td>
<td>8 (12.3)</td>
<td>0</td>
<td>8 (12.3)</td>
<td>5 (7.7)</td>
<td>6 (9.2)</td>
<td>16 (24.6)</td>
<td>15 (23.1)</td>
<td>2 (3.1)</td>
<td>65 (100)</td>
</tr>
<tr>
<td>Total</td>
<td>2 (0.5)</td>
<td>17 (4.0)</td>
<td>38 (8.9)</td>
<td>28 (6.5)</td>
<td>92 (21.5)</td>
<td>18 (4.2)</td>
<td>26 (6.1)</td>
<td>146 (34.1)</td>
<td>53 (12.4)</td>
<td>8 (1.8)</td>
<td>428 (100)</td>
</tr>
</tbody>
</table>
Table 2: Prevalence of Selected Genes within SCG Type (Count (%))

<table>
<thead>
<tr>
<th></th>
<th>M. bovis</th>
<th>SCG 1</th>
<th>SCG 2</th>
<th>SCG 3a</th>
<th>SCG 3b</th>
<th>SCG 3c</th>
<th>SCG 4</th>
<th>SCG 5</th>
<th>SCG 6a</th>
<th>SCG 6b</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>KatG</td>
<td>0</td>
<td>12 (5.0)</td>
<td>29 (12.1)</td>
<td>12 (5.0)</td>
<td>51 (21.3)</td>
<td>12 (5.0)</td>
<td>16 (6.7)</td>
<td>81 (39.9)</td>
<td>22 (9.2)</td>
<td>4 (1.8)</td>
<td>239 (100)</td>
</tr>
<tr>
<td>KatG315</td>
<td>0</td>
<td>12 (5.4)</td>
<td>25 (11.5)</td>
<td>11 (5.1)</td>
<td>45 (20.1)</td>
<td>12 (5.4)</td>
<td>14 (6.4)</td>
<td>78 (35.9)</td>
<td>17 (7.8)</td>
<td>3 (1.4)</td>
<td>217 (100)</td>
</tr>
<tr>
<td>KasA</td>
<td>0</td>
<td>8 (16.7)</td>
<td>1 (2.1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>39 (81.2)</td>
<td>0</td>
<td>0</td>
<td>48 (100)</td>
</tr>
<tr>
<td>InhA</td>
<td>0</td>
<td>3 (7.1)</td>
<td>4 (9.5)</td>
<td>2 (4.8)</td>
<td>10 (23.8)</td>
<td>1 (2.4)</td>
<td>2 (4.8)</td>
<td>12 (28.6)</td>
<td>8 (19.0)</td>
<td>0</td>
<td>42 (100)</td>
</tr>
<tr>
<td>InhProm</td>
<td>0</td>
<td>3 (7.9)</td>
<td>4 (10.5)</td>
<td>2 (5.3)</td>
<td>9 (23.7)</td>
<td>1 (2.6)</td>
<td>2 (5.3)</td>
<td>10 (26.3)</td>
<td>7 (18.4)</td>
<td>0</td>
<td>38 (100)</td>
</tr>
<tr>
<td>InhORF</td>
<td>0</td>
<td>1 (11.1)</td>
<td>0</td>
<td>1 (11.1)</td>
<td>2 (22.2)</td>
<td>0</td>
<td>1 (11.1)</td>
<td>3 (33.3)</td>
<td>4 (11.1)</td>
<td>0</td>
<td>9 (100)</td>
</tr>
<tr>
<td>AhpC</td>
<td>0</td>
<td>1 (1.7)</td>
<td>1 (1.7)</td>
<td>17 (28.3)</td>
<td>10 (16.7)</td>
<td>0</td>
<td>11 (18.2)</td>
<td>13 (21.7)</td>
<td>6 (10.0)</td>
<td>1 (1.7)</td>
<td>60 (100)</td>
</tr>
<tr>
<td>Ndh</td>
<td>0</td>
<td>1 (7.1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1 (7.1)</td>
<td>0</td>
<td>12 (85.7)</td>
<td>0</td>
<td>0</td>
<td>14 (100)</td>
</tr>
</tbody>
</table>
## Table 3: Number of Individual Antibiotics Showing Resistance within SCG Type (Count (%))

| M. bovis | SCG 1 (Count) | SCG 2 (Count) | SCG 3a (Count) | SCG 3b (Count) | SCG 3c (Count) | SCG 4 (Count) | SCG 5 (Count) | SCG 6a (Count) | SCG 6b (Count) | Total (Count) |
|----------|----------------|---------------|----------------|----------------|----------------|---------------|---------------|---------------|---------------|---------------|--------------|
| 1        | 2 (70.59)      | 12 (31.58)    | 12 (42.86)     | 26 (28.26)     | 8 (44.44)      | 8 (30.77)     | 28 (19.18)    | 14 (26.42)    | 3 (37.5)      | 125 (29.21)   |
| 2        | 0 (11.76)      | 10 (26.32)    | 9 (32.14)      | 30 (32.61)     | 4 (22.22)      | 12 (46.15)    | 39 (26.71)    | 16 (30.19)    | 2 (25.0)      | 124 (28.97)   |
| 3        | 0 (11.76)      | 5 (13.16)     | 6 (21.43)      | 21 (22.83)     | 5 (27.78)      | 4 (15.38)     | 53 (36.30)    | 10 (18.87)    | 3 (37.50)     | 109 (25.47)   |
| 4        | 0 (5.88)       | 11 (28.95)    | 1 (3.57)       | 15 (16.3)      | 1 (5.56)       | 2 (7.69)      | 26 (17.81)    | 13 (24.53)    | 0 (0.00)      | 70 (16.36)    |
| Total    | 2 (100)        | 38 (100)      | 28 (100)       | 92 (100)       | 18 (100)       | 26 (100)      | 146 (100)     | 53 (100)      | 8 (100)       | 428 (100)     |
Figure 1. Distribution of mutations by SCG.