Genetic diversity of isoniazid-resistant *Mycobacterium tuberculosis* isolates in Poland, assessed by spoligotyping

EWA AUGUSTYNOWICZ-KOPEĆ, TOMASZ JAGIELSKI, ZOFIA ZWOLSKA*

*Department of Microbiology, National Tuberculosis and Lung Diseases Research Institute, Warsaw, Poland*
CORRESPONDENT FOOTNOTE

Mailing address: National Tuberculosis and Lung Diseases Research Institute, Department of Microbiology, Płocka 26, 01-138 Warsaw, Poland. Phone: 48 22 43 12 382. Fax: 48 22 43 12 182. E-mail: z.zwolska@igichp.edu.pl.

ACCEPTED
The genetic composition of 71 isoniazid-resistant *Mycobacterium tuberculosis* strains from Poland was determined by spoligotyping. Nearly 80% of the isolates belonged to either T or Haarlem family. The genotypic diversity was largely due to variation within those families. Scarcity of imported genotypes suggested an endemic nature of the *M. tuberculosis* population studied.
Isoniazid (INH) is one of the most important drugs for both therapy and prophylaxis of tuberculosis (TB). However, strains of *Mycobacterium tuberculosis*, the causative agent of TB, resistant to INH have been isolated increasingly throughout the world. Globally, the burden of INH resistance is estimated at 13.3% (25). In Poland, in 2000, the overall rate of resistance to INH reached 6%, or 1.5 times higher than that in 1997 (2). In 2004, 5% of all TB cases were reported as INH-resistant. INH monoresistance is the most prevalent type of TB resistance in Poland, with the median rate of primary and acquired resistance of 2.4% and 3.3%, respectively (25).

In recent years, molecular typing methods have been successfully employed to define the genetic relationships between *M. tuberculosis* strains and to delineate transmission patterns of TB in human populations (16). Currently, one of the methods most frequently used to differentiate clinical isolates of *M. tuberculosis* is spoligotyping, which detects polymorphisms in the *M. tuberculosis* complex direct repeat (DR) chromosomal locus containing a series of 36-bp DRs interspersed with 35- to 41-bp unique spacer sequences (14).

The aim of this study was to use spoligotyping to assess the genetic diversity of *M. tuberculosis* strains with INH-monoresistant phenotype in Poland in 2004. (Part of the results included in this study was presented in abstract form [D13] at the 8th International Meeting on Microbial Epidemiological Markers, Zakopane, Poland, 14-17 May, 2008).

The study included 71 *M. tuberculosis* isolates from 71 non-related, adult TB patients with pulmonary TB (52 men and 19 women; age range, 14-85 years; median age, 48 years) residing in 13 different regions (voivodeships) of Poland. The analysed population...
represented 87% of all culture-proven INH-resistant TB cases notified in Poland in 2004, and close to 30% of the total number of drug-resistant pulmonary TB cases recorded in Poland throughout 2004.

Primary isolation was performed using the Löwenstein-Jensen (LJ) medium and the BACTEC 460-TB system (Becton-Dickinson, Sparks, MD, USA), and the species identification was done by means of niacin test, gene probes (AccuProbe, GenProbe, San Diego, USA) and mycolic acids HPLC analysis. Drug susceptibility testing was performed using the proportion method on LJ medium. The criterion used for drug resistance was growth of 1% or more of the bacterial population on critical concentrations of the drugs tested (i.e. 0.2 µg/ml for isoniazid (INH), 40 µg/ml for rifampicin (RMP), 4 µg/ml for streptomycin, and 2 µg/ml for ethambutol (EMB)) (2). Bacterial DNA was extracted from LJ slants by the cetyl-trimethyl-ammonium bromide (CTAB) method (22).

Spoligotyping was performed with a commercially available kit (Isogen Bioscience BV, Maarssen, The Netherlands), according to the manufacturer’s instructions and as described earlier (14). Spoligotypes with 100% similarity were considered clusters, whereas non-clustered spoligotypes were referred to as unique. All spoligotypes obtained were compared to the world spoligotyping database (SpolDB4) at the Pasteur Institute of Guadeloupe (www.pasteur-guadeloupe.fr/tb/spoldb4). The isolates whose spoligotype patterns were already recorded in SpolDB4, were assigned ‘shared types’, whereas those of spoligotypes identified for the first time, were designated either as ‘new shared types’ (if 2 or more) or as ‘orphans’ (if occurred only once). Clade assignment of the spoligotypes not found in SpolDB4 (orphan types) was done with SpotClust, an algorithm
based on the SpolDB3 database, whose principle was described previously (23) and available online (http://cgi2.cs.rpi.edu/~bennek/SPOTCLUST.html).

A total of 30 different genetic profiles were identified among the 71 *M. tuberculosis* isolates, resulting in an overall diversity (the number of spoligotypes divided by the number of isolates) of 42%. Twenty-one (30%) isolates exhibited unique patterns and the remaining 50 (70%) isolates were grouped into 9 clusters, with 2-24 isolates per cluster. A comparison of the profiles obtained with the international spoligotyping database SpolDB4 allowed the attribution of the spoligotype ST (Shared Type) designations. Of the 21 unique patterns, 12 (57%) were already described in SpolDB4, while the remaining 9 (43%) were previously unreported orphan types. Among the 9 clusters, 5 containing three or more isolates each were considered as major spoligotypes and represented 42 (59%) of the isolates. Almost half of the clustered isolates (24; 34% of all isolates) belonged to ST53, which is the most prevalent genotype in Europe. All strains that were labeled with an ST number fell into four spoligotype-defined clades: T (STs 37, 40, 44, 53, 253, 280, 498, 612, 1278), Haarlem (STs 47, 50, 262, 382, 1640), U (STs 602, 775, 1410, 1498), and S (ST1253), comprising 36 (51%), 15 (21%), 4 (6%), and 2 (3%) isolates, respectively. Of the 30 spoligotypes found in this study, 12 (40%) had been reported in Poland previously (Table 1). Eleven spoligotypes (37%), including 14 (20%) isolates, were new and not found elsewhere in the world. In order to assign these spoligotypes to the existing clades, they were subjected to SpotClust analysis. Thus, spoligotypes designated D, E, F, G and K were linked to the T clade, whereas spoligotypes B, H and I were linked to the Latin American and Mediterranean (LAM) family. Three isolates harboured spoligotypes that were closely related to that of H37Rv.
Finally, two spoligotypes, represented by one isolate each, were shown to belong to the Haarlem and Central Asian (CAS) clades (Table 2).

Overall, the family assignment demonstrated that a major proportion of the analyzed strains belonged to either the T (58%) or Haarlem (22%) family.

The spoligotyping method has been widely accepted as a valuable tool for epidemiological studies of TB. It has proven useful for tracking outbreaks (10, 19), laboratory cross contaminations (6), and describing global spread of TB (7). Spoligotyping has considerable advantages in that it is simple, robust and highly reproducible. Since the technique is PCR-based, it can be applied directly to clinical samples, thus allowing fingerprinting of a large number of isolates to be performed in a very short time (14). An important aspect of the method is the binary result format, that makes the generated data easily interpretable, computerizable and comparable between laboratories (5). The establishment and development of an international spoligotyping database (SpolDB) has tremendously contributed to a better understanding of the genetic structure of the global *M. tuberculosis* population and its evolution. Recently, an updated SpolDB4 version, including 1,939 different spoligotypes (ST) identified worldwide and separated into phylogenetic clades (families), has been launched (4). Moreover, a web-based program SpotClust has been devised to assign spoligotypes (especially those not found in the SpolDB4 database) to ST families (23). Altogether, spoligotyping combined with bioinformatic analyses provide an efficient approach for determining the diversity of circulating *M. tuberculosis* strains, and for assessing the extent and dynamics of TB transmission.
In the present study we have employed spoligotyping to investigate the genetic diversity of INH-resistant *M. tuberculosis* isolates in Poland. Among the 71 isolates tested, 30 different spoligotype patterns were identified, indicating significant heterogeneity of the population studied. However, upon phylogenetic analysis, the vast majority of the genotypes were allocated to only two major clades, namely T and Haarlem, covering, respectively, 41 (58%) and 16 (22%) of the analyzed *M. tuberculosis* strains. Both these families are highly prevalent in Europe. According to the SpolDB4 database, the ill-defined T family encompasses spoligotypes which likely represent relatively old genotypes prevalent in Europe, whereas the Haarlem lineage has a European origin and is essentially localized in northern European countries (7, 8). Our results have clearly shown the T and Haarlem families to make up the backbone of the genetic structure of the *M. tuberculosis* population studied. The observed propensity of specific mycobacterial lineages to spread within particular geographical areas has recently been linked to the adaptation of the pathogen to different human host populations (9).

The three most prevalent spoligotypes (ST53, ST50, and ST47) accounted for more than a half of the TB cases, suggesting an important role of these genotypes in local TB transmissions in Poland. This could be further supported by the fact that the same three types were found dominant in two previous spoligotyping studies in Poland (1, 17). This observation indicates the continuing activity of these genotypes in the country. Conversely, 22 (31%) of INH-resistant *M. tuberculosis* strains yielded spoligotypes previously unreported in Poland. These spoligotypes had either already been defined in the global database or they were identified for the first time. Whereas finding of the
former may be indicative of the transmission of imported *M. tuberculosis* strains, the
presence of the latter suggests Polish phylogeographic specificity. Interestingly, one of
the Poland-specific spoligotypes has been demonstrated to belong to the Central Asian
(CAS) clade, which is very poorly represented in Europe. A possible explanation for this
may relate to phylogenetic convergence, which relies upon independent acquisition of
two similar structures without common ancestors (24).

Finally, clustering of the genotypes obtained with spoligotyping may suggest active
transmission of INH-resistant TB in Poland. However, whereas unique spoligotypes can
be confidently assumed to represent different strains, and thus reflect reactivation of
remotely acquired infection, spoligotype clustering as a proxy for recent or ongoing
transmission has to be interpreted with caution. This is because spoligotyping is
substantially less discriminatory than IS6110 restriction fragment length polymorphism
(RFLP) typing method which is referred to as a “gold standard” for *M. tuberculosis*
molecular epidemiology (15, 22). Several studies have shown that spoligotyping,
compared to the IS6110-RFLP, overestimates the number of isolates with identical DNA
fingerprints (12, 13, 15), and that this overestimation can be as large as 50% (11).

Nevertheless, a combination of spoligotyping with other PCR-based techniques, such as
double-repetitive-element PCR (DRE-PCR) (20) or mycobacterial interspersed repetitive
unit-variable-number tandem repeat (MIRU-VNTR) typing (21) has been demonstrated
to afford the resolution capacity close to that of IS6110-RFLP. It has to be emphasized
however that none of the currently used genetic markers provides an accurate and
sufficiently discriminatory genotyping system for *M. tuberculosis*, which could therefore
reliably estimate the extent of recent transmission. Moreover, clustering of *M.*
tuberculosis isolates, measured by DNA clonality, might not be attributed to recently transmitted infection but to coincident reactivations of strains endemic to the area or independent evolutionary convergence (3). Thus, genotypic clustering is not always consistent with the epidemiological relatedness of the analysed cases and it is of great importance to interpret genotyping results in conjunction with conventional contact tracing methodology.

Altogether, the results of spoligotyping, as the sole typing method, are not conclusive in terms of determining the amount of recent transmission in population-based studies (12, 18). Although useful as an initial screening test, spoligotyping has to be followed by other fingerprinting methods of a higher discrimination ability. Consequently, evaluation of the actual rate of ongoing INH-resistant TB transmission in Poland, requires the use of additional genetic markers, preferably coupled with conventional contact tracing data.

Nonetheless spoligotyping alone remains a highly informative genotyping method, providing insight into the genetic structure, evolution and spreading dynamics of the global M. tuberculosis population.

In conclusion, our spoligotyping has revealed the diversity of M. tuberculosis in Poland similar to typical European country with the predominance of T and Haarlem families, both accounting for nearly 80% of the TB cases studied. The diversity of the genotypes was largely due to variation within those families. At the same time, several new, previously unreported genotypes have been identified. Scarcity of cases harbouring imported genotypes may be indicative of endemic nature of the M. tuberculosis population studied.

ACKNOWLEDGEMENTS
The authors are deeply indebted to Dr. Vitali Sintchenko (University of Sydney, Australia) for critical revision of the manuscript.

REFERENCES


Narbonne, O. Narvskaya, A. Nastasi, S. Ngo Niobe-Eyangoh, J. W. Pape, V.
Rasolofo-Razanamparany, M. Ridell, M. Lucia Rossetti, F. Stauffer, P. N.
and N. Ragosti. 2003. Snapshot of moving and expanding clones of
*Mycobacterium tuberculosis* and their global distribution assessed by

Cataldi, R. C. Cooksey, D. V. Cousins, J. W. Dale, O. A. Dellagostin, F.
Drobniewski, G. Engelmann, S. Ferdinand, D. Gascoyne-Binzi, M. Gordon,
M. C. Gutierrez, W. H. Haas, H. Heersma, G. Källenius, E. Kassa-Kelembho,
T. Koivula, H. M. Ly, A. Makristathis, C. Mammina, G. Martin, P. Moström,
I. Mokrousov, V. Narbonne, O. Narvskaya, A. Nastasi, S. N. Niobe-Eyangoh,
J. W. Pape, V. Rasolofo-Razanamparany, M. Ridell, M. L. Rossetti, F.

Narayanan, M. Nicol, S. Niemann, K. Kremer, M. C. Gutierrez, M. Hilty, P.
C. Hopewell, and P. M. Small. 2006. Variable host-pathogen compatibility in

10. Goguet de la Salmonière, Y. O., H. M. Li, G. Torrea, A. Bunschoten, J. D. A.


methodology to IS6110-fingerprinting for epidemiological studies of tuberculosis.


2001. Automated high-throughput genotyping for study of global epidemiology of
*Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units.


TABLE I. Spoligotypes identified for *M. tuberculosis* isolates evaluated in this study and their clustering.

<table>
<thead>
<tr>
<th>Clade</th>
<th>ST&lt;sup&gt;a&lt;/sup&gt;</th>
<th>No. of isolates</th>
<th>Geographic distribution&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ubiquitous</td>
</tr>
<tr>
<td>T1</td>
<td>53</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>H3</td>
<td>50</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>47</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>1278</td>
<td>3</td>
<td>AUT, CZE, ESP, ITA, POL, USA</td>
</tr>
<tr>
<td>ND&lt;sup&gt;c&lt;/sup&gt;</td>
<td>A</td>
<td>3</td>
<td>POL</td>
</tr>
<tr>
<td>T1</td>
<td>253</td>
<td>2</td>
<td>ARG, DNK, FIN, FXX, IDN, NLD, POL, RUS, USA, VEN</td>
</tr>
<tr>
<td>T1_RUS2</td>
<td>280</td>
<td>2</td>
<td>AUS, AUT, DEU, EST, FIN, GEO, LVA, POL, RUS, SWE, TUR, USA</td>
</tr>
<tr>
<td>S</td>
<td>1253</td>
<td>2</td>
<td>ARG, DEU, FXX, RUS, TUR</td>
</tr>
<tr>
<td>ND</td>
<td>B</td>
<td>2</td>
<td>POL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>37</td>
<td>1</td>
<td>Ubiquitous</td>
</tr>
<tr>
<td>T4</td>
<td>40</td>
<td>1</td>
<td>AUT, BEL, CZE, DEU, DNK, ESP, ETH, FIN, FXX, GBR, GEO, GLP, IRN, ITA, LBY, MDG, NLD, POL, RUS, USA</td>
</tr>
<tr>
<td>T5</td>
<td>44</td>
<td>1</td>
<td>Ubiquitous</td>
</tr>
<tr>
<td>H4</td>
<td>262</td>
<td>1</td>
<td>POL, RUS, SWE, TUR, USA</td>
</tr>
<tr>
<td>H1</td>
<td>382</td>
<td>1</td>
<td>AUT, EST, FXX, POL, USA</td>
</tr>
<tr>
<td>T1</td>
<td>498</td>
<td>1</td>
<td>AUT, CZE, DEU, GBR, USA</td>
</tr>
<tr>
<td>U</td>
<td>602</td>
<td>1</td>
<td>AUT, BEL, BRA, DEU, FIN, GEO, IDN, IRN, ITA, NLD, NZL, RUS, SWE, USA, VNM, ZAF</td>
</tr>
<tr>
<td>T1</td>
<td>612</td>
<td>1</td>
<td>ARG, FXX, GBR, USA</td>
</tr>
<tr>
<td>U</td>
<td>775</td>
<td>1</td>
<td>AUT, CZE, HUN, NLD, SWE, USA</td>
</tr>
<tr>
<td>U</td>
<td>1410</td>
<td>1</td>
<td>AUT, BGD, BRA, POL, PRT, SAU</td>
</tr>
<tr>
<td>U</td>
<td>1498</td>
<td>1</td>
<td>MYS, USA</td>
</tr>
<tr>
<td>H3</td>
<td>1640</td>
<td>1</td>
<td>FXX, NLD</td>
</tr>
<tr>
<td>ND</td>
<td>C-K</td>
<td>1</td>
<td>POL</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>30</td>
<td>71</td>
</tr>
</tbody>
</table>

<sup>a</sup>ST – Shared Type, designation of the spoligotype in the world spoligotype database (SpolDB4);<br>

<sup>b</sup>Ubiquitous – spoligotype found in all eight geographic regions (AFR, Africa; CAM, Central America; EUR, Europe; FEA, Far-East Asia; MECA, Middle-East and Central Asia; NAM, North America; OCE, Oceania; SAM, South America); ARG, Argentina; AUS, Australia; AUT, Austria; BEL, Belgium; BGD, Bangladesh; BRA, Brazil; CZE, Czech Republic; DEU, Germany; DNK, Denmark; ESP, Spain; EST,
Estonia; ETH, Ethiopia; FIN, Finland; FXX, Metropolitan France; GBR, United Kingdom; GEO, Georgia;
GLP, Guadeloupe; HUN, Hungary; IDN, Indonesia; IRN, Iran; ITA, Italy; LBY, Libya; LVA, Latvia;
MDG, Madagascar; MYS, Malaysia; NLD, Netherlands; NZL, New Zealand; POL, Poland; PRT, Portugal;
RUS, Russia; SAU, Saudi Arabia; SWE, Sweden; TUR, Turkey; USA, United States; VEN, Venezuela;
VNM, Viet Nam; ZAF, South Africa. Country codes following the ISO 3166 specifications
(ftp://ftp.ripe.net/iso3166-countrycodes.txt);
Not Defined in the SpolDB4.
### TABLE II. Spoligotypes specific to Poland and their clade assignment.

<table>
<thead>
<tr>
<th>Spoligotype&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Clade&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Probability&lt;sup&gt;c&lt;/sup&gt;</th>
<th>No. of isolates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Binary</strong></td>
<td><strong>Octal</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>777741077760771</td>
<td>H&lt;sub&gt;37&lt;/sub&gt;Rv</td>
<td>0.98</td>
</tr>
<tr>
<td>B</td>
<td>777741003760771</td>
<td>LAM9</td>
<td>0.94</td>
</tr>
<tr>
<td>C</td>
<td>000017600003771</td>
<td>CAS</td>
<td>0.99</td>
</tr>
<tr>
<td>D</td>
<td>70777777760711</td>
<td>T1</td>
<td>0.99</td>
</tr>
<tr>
<td>E</td>
<td>74077777760700</td>
<td>T2</td>
<td>0.98</td>
</tr>
<tr>
<td>F</td>
<td>770000777360771</td>
<td>T3</td>
<td>0.99</td>
</tr>
<tr>
<td>G</td>
<td>77734777760771</td>
<td>T1</td>
<td>0.99</td>
</tr>
<tr>
<td>H</td>
<td>777737607420771</td>
<td>LAM9</td>
<td>0.99</td>
</tr>
<tr>
<td>I</td>
<td>777761007760731</td>
<td>LAM9</td>
<td>0.97</td>
</tr>
<tr>
<td>J</td>
<td>777771374020771</td>
<td>H1</td>
<td>0.99</td>
</tr>
<tr>
<td>K</td>
<td>77777777740031</td>
<td>T1</td>
<td>0.99</td>
</tr>
</tbody>
</table>

<sup>a</sup> Spoligotype, A-K, arbitrary spoligotype designation;

<sup>b</sup> SpotClust-assigned clade;

<sup>c</sup> Probability of the spoligotype pattern to belong to the clade.