Molecular epidemiologic studies of tuberculosis (TB) use genetic markers to identify genotypically similar clinical isolates in the face of a diverse background of strains. This approach has been successfully applied to confirm outbreaks, to define acquired drug resistance, and to estimate the proportion of TB cases due to recent transmission (1, 12, 15, 34, 35, 39). Through these studies, a number of widespread families of genotypically similar strains have been described throughout the world, including the Beijing family and its subset strain W (2, 4, 18, 40).

In high-prevalence areas, the predominance of a particular family or clone is likely due to both ongoing transmission of an epidemic strain and reactivation of an old endemic strain. In the absence of exquisite detail epidemiologic investigations, these events are indistinguishable. As a result, it is difficult to infer the success of this family and to determine whether its high prevalence stems from enhanced virulence.

In the province of Quebec (Canada), it has been observed that 6.2% of Mycobacterium tuberculosis isolates in the Canadian-born population are resistant to pyrazinamide (PZA) alone, a highly unusual phenotype (14, 21). Because PZA resistance has been attributed to mutations within the pncA gene (24, 28, 32, 33, 37), a previously published study examined the pncA gene in 21 isolates from Quebec. These isolates shared the same unique pncA mutation profile and showed similar IS6110 restriction fragment length polymorphism (RFLP) patterns. The data were thus interpreted to suggest ongoing transmission of a PZA-resistant (PZA-R) strain within that community (10).

Because of the low incidence of TB in the Canadian-born population of Quebec (1.9 per 100,000), we instead hypothesized that the PZA-R isolates in Quebec were clonal members of a PZA-R family and accounted for a significant number of cases due to remote transmission. To test our hypothesis, we assembled a case-control study of 77 PZA-R isolates with the “Quebec” mutation profile and 253 PZA-susceptible (PZA-S) control isolates. Multiple molecular markers were used to characterize the genetic similarity between isolates. To determine the possibility of an outbreak, public health chart review and geographic analyses were undertaken.

(Part of the information presented in this report was presented at the American Thoracic Society International Conference, May 2002, in Atlanta, Ga. [abstr. LB2-D21], and at the International Union against Tuberculosis and Lung Disease World Conference, September 2002, in Montreal, Quebec, Canada [abstr. 343-PS].)

MATERIALS AND METHODS

Setting. The province of Quebec measures 1.5 million km² (600,000 mi²). It has a population of 7.2 million, 80% of whom live in urban centers. Montreal is the largest city (1.8 million). The overall incidence of TB is 5 cases per 100,000, but there are only 1.9 cases per 100,000 in the non-Aboriginal Canadian-born population.

Isolates. All culture-positive TB cases (approximately 90% of reported cases) in the province of Quebec are reported and sent to the Laboratoire de santé publique du Québec (LSPQ) for culture confirmation and routine drug susceptibility testing (including PZA). Clinical isolates for this study were received between 1990 and 2000 at the LSPQ. Testing for susceptibility to PZA was performed using the radiometric method (BACTEC 460) (23). Resistance to PZA was defined as a MIC of >100 mg/liter and a negative pyrazinamide activity assay (45). A subset of isolates were further tested at a MIC of >300 mg/liter. Additional tests (niacin, nitrate reduction, and resistance to thiophencarboxylylhydrazide) were performed to exclude Mycobacterium bovis (20).

Study design. A case-control design was used. Cases were defined as PZA-R M. tuberculosis isolates with the specific “Quebec” mutation profile in the pncA gene.
gene (defined below). Controls were PZA-S M. tuberculosis isolates from Canadian-born subjects randomly selected from the study period (1990 to 2000). Controls were chosen among Canadian-born individuals because all patients in the “case” category at the time of the study design were Canadian-born. We included 4 controls per case given an expected 64 cases, representing approximately 20% of all M. tuberculosis isolates in Canadian-born subjects. PZA-R isolates without the “Quebec” mutation profile were not included in the remainder of the case-control study.

Genotyping. M. tuberculosis DNA was extracted from clinical isolates by using standardized methods. The “Quebec” mutation profile in the pncA gene—an 8-bp deletion (ATGGCTTT at position 446) and a point mutation (C to A) at position 418—was identified by PCR-RFLP. Briefly, the pncA gene was amplified by using primers L (5′-GGCATGCAAGCCTATATC-3′) and R (5′-CAAC AGTTATCGGGTTC-3′) and was digested with HinfI and BsrD1. The pncA gene in PZA-R isolates without the “Quebec” mutation profile was sequenced to identify other mutations. PCR confirmation of region RD4, a 1,031-bp region shown to be consistently deleted in M. bovis but present in other members of the M. tuberculosis complex, was performed for all PZA-R isolates (9, 19, 29). IS6110 RFLP by Southern blotting (38) and spoligotyping (Isogen Bioscience B.V.) (26) were performed using standardized methods. Major genetic groups (as described by Streeter et al. [36]) were determined based on katG943 and gyrA95 polymorphisms by using molecular beacon assays (11) with real-time PCR (ABI Prism 7700 sequence detection system; Perkin-Elmer Applied Biosystems) software.

Genotype analysis. IS6110 RFLP and PCR-RFLP results were scanned into a Syngene (Synoptics Ltd., Cambridge, United Kingdom) gel documentation system and digitized for computer-assisted visual reading by three independent readers. IS6110 RFLP cluster analysis was restricted to high-copy-number isolates (more than four bands) and was performed with MFA (Molecular Finger-print Analyzer J, version 2.0; Stanford Center for Tuberculosis Research) and Gel Compar II (Applied Maths, Kortrijk, Belgium). The Dice coefficient was used to estimate the similarity between isolates within 2% tolerance. Dendrograms were constructed with the Dice coefficient using the UPGMA (unweighted pair group method with arithmetic mean) algorithm. Spoligotype patterns were coded and analyzed manually as an Excel document by using the sort function. Molecular beacon results were analyzed with Sequence Analyzer (Perkin-Elmer Applied Biosystems) software.

Demographic and epidemiologic data. The provincial database of reportable diseases was used for demographic information, and public health charts were studied to identify reported epidemiologic links. French-Canadian heritage was used for demographic information, and public health charts were studied to identify reported epidemiologic links. French-Canadian heritage was used for demographic information, and public health charts were studied to identify reported epidemiologic links. French-Canadian heritage was used for demographic information, and public health charts were studied to identify reported epidemiologic links. French-Canadian heritage was used for demographic information, and public health charts were studied to identify reported epidemiologic links.

RESULTS

We identified 103 PZA-R M. tuberculosis clinical isolates and 256 PZA-S isolates from the LSPQ (Fig. 1). All PZA-R isolates were classified as M. tuberculosis and not M. bovis by conventional testing. Two PZA-S isolates were excluded from analysis due to laboratory contamination with the reference strain H37Ra. Another PZA-S isolate and two PZA-R isolates were excluded because the subjects lived outside of Quebec at the time of diagnosis, leaving 101 PZA-R isolates and 253 PZA-S isolates for study.

PCR-RFLP of the pncA gene in all isolates revealed that 77 of 101 isolates had the “Quebec” mutation profile and were thus considered cases. Seventy-six out of 77 cases were PZA monoresistant, and one was multidrug resistant. The RD4 region, described to be consistently absent in M. bovis, was present in all these PZA-R isolates, thus confirming that they are not M. bovis. Furthermore, it is worth noting that the pncA gene mutation identified here is distinctly different from the mutation seen in M. bovis (C to G at position 169). In the remaining 24 PZA-R isolates, sequencing of the pncA gene revealed 14 different mutations in the pncA gene or promoter region. PCR could not be performed for one isolate. These 24 isolates were excluded from further analysis. All 253 PZA-S controls had the pncA wild-type profile by PCR-RFLP.

Three hundred twenty-six out of 330 isolates were classified according to the major genetic groups by using molecular beacon assays. All cases with the Quebec mutation profile belonged to major genetic group 2, whereas the PZA-S controls were distributed among all three groups: group 1 (2%), group 2 (75%), and group 3 (23%). The molecular beacon assay was unsuccessful with four isolates (one PZA-R case and three PZA-S controls).

IS6110 RFLP of the 77 case isolates revealed a high degree of genetic similarity, though they were not all identical: by use of identical RFLP patterns to define a match, 34 out of 77 isolates (44%) fell into 12 clusters of 2 to 4. Allowing for a one-band difference to define a near-identical match resulted in a large cluster comprising 62 of the 77 cases. In contrast, 49 out of 230 control isolates (21% of all controls with more than 4 bands) fell into 18 clusters of 2 to 8 by use of identical RFLP patterns to define a match. When controls were matched by near-identical patterns, the largest cluster had only 18 members. By using the Dice coefficient as a measure of genetic similarity, the 77 case isolates had a median similarity index of 86% (standard deviation, 4%) compared to 52% for the controls (Fig. 2).

The spoligotype patterns of all 77 cases had a common “signature” deletion of spacers 9 and 10. Beyond the signature deletion, the variability in the spoligotype patterns among the cases provided evidence that the expansion of this clone was
not likely to have been recent. A total of 13 patterns were observed, with pattern A shared by 42 isolates (55%). Patterns B (shared by 14 isolates), C (2 isolates), and D (2 isolates) are each one deletion away from pattern A (Fig. 3). In contrast, spoligotype patterns seen in the PZA-S controls showed greater diversity, with a total of 101 patterns identified (Fig. 3).

The cases and controls were comparable with respect to the patients’ ages, places of residence (in or outside of Montreal), and prior history of TB. Demographic information obtained from the public health records is presented in Table 1. Epidemiologic investigations revealed no evidence of a large outbreak. Epidemiologic links to another case of TB were identified for 12% of the 77 cases versus 8% of the PZA-S controls (not statistically significant). The 77 cases occurred over a 10-year period, with a median of 7 cases per year and a peak incidence of 15 cases in 1995 (no statistically significant differences from year to year).

Postal codes were missing for 1 case and 15 controls, leaving 76 cases and 238 controls available for geographic mapping and analysis. The median nearest neighbor between controls was estimated by random repeated resampling of the control group. The median nearest neighbor between all 76 cases was 6 km (25th to 75th percentile, 2 to 15 km) compared to the mean nearest neighbor between controls of 5.9 km. This suggests that the cases were no more clustered in space than a random selection of Canadian-born controls. This conclusion held true even when the cases with identical IS6110 RFLP matches were compared to the controls without identical RFLP matches: these cases had a median nearest neighbor of 28 km (25th to 75th percentile, 3 to 21 km), whereas the controls had a mean nearest neighbor of 10 km (25th to 75th percentile, 7 to 12 km).

**DISCUSSION**

The use of molecular markers in epidemiologic studies of TB has been particularly informative in redefining transmis-
sion dynamics and documenting important epidemic strains and clones. The best described of these is the Beijing family, along with its subset, the W strain family. The clonal nature of members of this family was established by specific IS6110 insertion sites, high IS6110 RFLP similarity, and a unique spoligotype pattern (3, 5). Due to the high transmission rate in countries such as China, or in high-risk settings such as inner-city neighborhoods in the United States, the predominance of such an M. tuberculosis family is most often attributed to ongoing transmission. However, certain authors have raised the possibility that the preponderance of a family in certain settings may instead reflect an endemic strain (7, 8).

We describe a family of PZA-R strains which appears to be specific to the region of Quebec. This family is defined by a unique mutation profile in the pncA gene that confers the PZA-R phenotype. The clonality of these isolates is demonstrated by (i) the shared pncA deletion, (ii) the high similarity between the IS6110 RFLP patterns, (iii) the common major genetic group 2, and (iv) a shared spoligotype deletion of direct variable regions 9 and 10.

In a low-incidence setting such as the Canadian-born population in Quebec (1.9 cases per 100,000) (22), it is remarkable that a single strain accounts for 6.2% of annual TB cases and 76% of all PZA-R isolates. In this setting, it appeared unlikely that a large outbreak of a highly unusual drug-resistant strain would expand unrecognized over a decade. Instead, we hypothesized that the prevalence of this family was primarily due to reactivation of an old endemic strain in a stable population. Individuals of French-Canadian heritage represent more than 80% of the province’s overall population and 90% of the study cases and controls. This population has expanded over the past three centuries with little incoming migration, except for urban areas such as Montreal. If a population has remained secluded over time without the introduction of genetically diverse M. tuberculosis strains, one or a few clones may become endemic. Their prevalence today may then be due to reactivation of remotely acquired infection.

Support for the latter possibility comes from the observation that the DNA fingerprints by both IS6110 RFLP and spoligotyping are similar but not necessarily identical. Since the mo-

**TABLE 1.** Demographic characteristics of cases and controls

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cases ($n = 77$)</th>
<th>Controls ($n = 253$)</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (yr) (SD)</td>
<td>59.4 (16.8)</td>
<td>60.7 (19.9)</td>
<td>NS</td>
</tr>
<tr>
<td>No. (%) female</td>
<td>25 (32.5)</td>
<td>90 (35.6)</td>
<td>NS</td>
</tr>
<tr>
<td>No. (%) Canadian-born</td>
<td>69/70 (98.6)</td>
<td>253 (100)</td>
<td>NS</td>
</tr>
<tr>
<td>No. (%) of French Canadian heritage</td>
<td>69 (89)</td>
<td>234 (92)</td>
<td>NS</td>
</tr>
<tr>
<td>No. (%) living in Montreal at time of diagnosis</td>
<td>8 (10.4)</td>
<td>43 (17)</td>
<td>NS</td>
</tr>
<tr>
<td>No. (%) with prior diagnosis of TB</td>
<td>14/71 (19.7)</td>
<td>41/247 (16.6)</td>
<td>NS</td>
</tr>
</tbody>
</table>
molecular clock of IS6110 RFLP has been estimated to be approximately 2 to 3.2 years (13, 42) and that of spoligotypes is believed to be slower (30), the genotyping data are consistent with the hypothesis that these strains are linked to a common but remote ancestor strain and have since diverged genetically.

To further determine the evidence for and against ongoing spread, we examined the evidence from conventional contact investigation for transmission in the case group and in a control group of Canadian-born patients with PZA-S. We investigated for transmission in the case group and in a control group with a known low incidence of TB and ongoing transmission (27). Cases and controls have comparable rates of epidemiologic links and shared demographic characteristics (older age, nonurban dwellers) different from those described for groups at high risk for ongoing transmission (1, 17, 34). However, a recognized limitation of this analysis is that TB spread is often unrecognized by traditional contact investigation (17, 34). For additional evidence against an outbreak, we utilized spatial aggregation as a measure of possible (or improbable) transmission: the closer two cases are to each other, the more likely there is ongoing transmission, and vice versa (4, 6). The cases were as geographically dispersed as a random selection of controls, and cases with identical IS6110 patterns were as dispersed as controls with different IS6110 patterns. Together these data provide compelling evidence that these genetically related isolates do not represent an unrecognized outbreak.

This prevalent PZA-R strain family does, however, raise several interesting questions regarding the origins of the drug resistance phenotype and its success. PZA monoresistance in this setting appeared to be spontaneous, without antibiotic pressure, based on a number of observations. Previous episodes of active TB in 13 out of 14 subjects occurred between 1941 and 1972, prior to the introduction of PZA into common practice in the past 20 years. The one subject treated more recently (1990) did not receive PZA treatment at that time. Given this historic context and the epidemiologic profile of reactivation disease, we postulate that acquired drug resistance was unlikely and that the pncA mutation antedated drug selective pressure. The spontaneous loss of pyrazinamidase activity had apparently not conferred any major survival disadvantage, a finding supported by the innate resistance of virulent M. bovis to PZA.

Furthermore, two observations pertain to the apparent success of this strain family. First, the prevalence of this family may be greater than might be estimated from the 77 PZA-R cases described here; numerous PZA-S control isolates were noted to have IS6110 RFLP patterns similar (but not identical) to those of the case isolates as well as deletions of DVR 9 and 10 in their spoligotype patterns (Fig. 3). Although the clonal link between the PZA-R cases and those controls remains to be proven by further genetic analysis, one might postulate that this subset of the PZA-S controls from the Quebec population shares a common ancestor with the PZA-R case isolates. Furthermore, this strain family, as defined by a unique pncA mutation profile described only in Quebec, has a spoligotype signature pattern previously described in several other spoligotype databases (11, 16, 41). The clonal link between our strain family and these strains still remains unproven, however. Second, the success of this historic strain family stands in contrast to widespread strains currently associated with ongoing transmission. In the absence of an outbreak, two possible explanations present themselves: either this was the predominant strain in a past epidemic, or this strain has an enhanced capacity to cause reactivation disease.

ACKNOWLEDGMENTS

We thank Annie Gatwood and Michael Purdy for assistance with genotyping and genetic analysis. This work was supported by grants from the Canadian Institutes for Health Research (CHIR) (grant 43896 to P.B.) and from the Sequella Global Tuberculosis Foundation (to M.A.B.). D.N. is a recipient of an FRSQ Bourse de Formation, K.S. is a Chercheur-Clinicien Boursier of the FRSQ, and M.A.B. is a New Investigator of CHIR.

REFERENCES


