MAJOR ARTICLE

Stability of DNA Patterns and Evidence of *Mycobacterium tuberculosis* Reactivation Occurring Decades after the Initial Infection

Troels Lillebaek,[1] Asger Dirksen,[3] Emilia Vynnycky,[4] Inga Baess,[1] Vibeke Ø. Thomsen,[1] and Åse B. Andersen[2]

[1]International Reference Laboratory of Mycobacteriology, Statens Serum Institut, the National Institute for Prevention and Control of Infectious Diseases and Congenital Disorders, and [2]Clinic of Infectious Diseases, Rigshospitalet University Hospital, Copenhagen, and [3]Department of Pulmonary Medicine, Gentofte University Hospital, Gentofte, Denmark; [4]Infectious Disease Epidemiology Unit, London School of Hygiene and Tropical Medicine, London, United Kingdom

Two hundred three freeze-dried strains of *Mycobacterium tuberculosis* collected during the 1960s were compared with 4102 strains collected during the 1990s, and 14 DNA patterns identified among the "historical strains" were 100% identical to patterns identified among the "recent strains." They were isolated from 41 and 40 patients who had tuberculosis during the 1960s and 1990s, respectively. The patients' mean age differed by >30 years, a finding strongly suggesting that the patients from the 1990s experienced reactivation of *M. tuberculosis* infection acquired during the 1960s. The half-life of IS*6110* DNA patterns during latency was estimated to be 36 years (95% confidence interval, 25–54 years). Thus, this comparison of historical and recent strains yields molecular epidemiologic evidence of *M. tuberculosis* reactivation spanning decades and suggests that the rate of change of DNA patterns during latency is much longer than that during active disease. This has important implications for the interpretation of clustering, especially for the extent of recent transmission.

During the past decade, the understanding of the molecular genetics of *Mycobacterium tuberculosis* has improved dramatically [1]. Molecular subtyping (e.g., by use of the international standardized method of restriction fragment–length polymorphism [RFLP]) allows the characterization of specific strains of *M. tuberculosis* on the basis of their DNA pattern, and many studies have shown that epidemiologically unrelated strains have different patterns, whereas related strains generally have identical patterns [2]. This recognition has led to an increasing number of studies measuring the proportion of tuberculosis cases that are "clustered" (i.e., that share identical DNA patterns), and many studies have typically concluded that the extent of recent transmission is more common than previously thought [3, 4]—for example, 57% of all patients with tuberculosis in Denmark with onset of disease during 1992–1995 were infected with strains of *M. tuberculosis* that were part of a cluster, and disease in most of these patients has been attributed to recent transmission [5].

Such conclusions are based on the assumption that the half-life of DNA patterns is short. If the half-life of DNA patterns is very long, for example, then clustering cannot be assumed to reflect recent transmission, given that even third- and fourth- or fifth-generation cases in the same chain of transmission would share identical patterns with the primary case. Although it is recognized that the half-life of patterns of strains involved in active disease can be as short as 3 years [6, 7], the half-life of strains involved in latent infection is not known and may be longer [8]. A long half-life also has important implications for the interpretation of clustering among the indigenous elderly populations of many Western countries, where some of the cases thus may be clustered as a result of simultaneous reactivation of a strain that was prevalent many years ago, rather

Received 20 February 2003; accepted 11 April 2003; electronically published 11 September 2003.

Financial support: Danish Lung Association; European Community Program "Quality of Life and the Management of Living Resources" (grant 2000-00630); UK Medical Research Council (support to E.V.).

Reprints or correspondence: Dr. Troels Lillebaek, Statens Serum Institut, International Reference Laboratory of Mycobacteriology, Artillerivej 5, DK-2300 Copenhagen S, Denmark (tll@ssi.dk).

The Journal of Infectious Diseases 2003;188:1032–9

0022-1899/2003/18807-0014$15.00

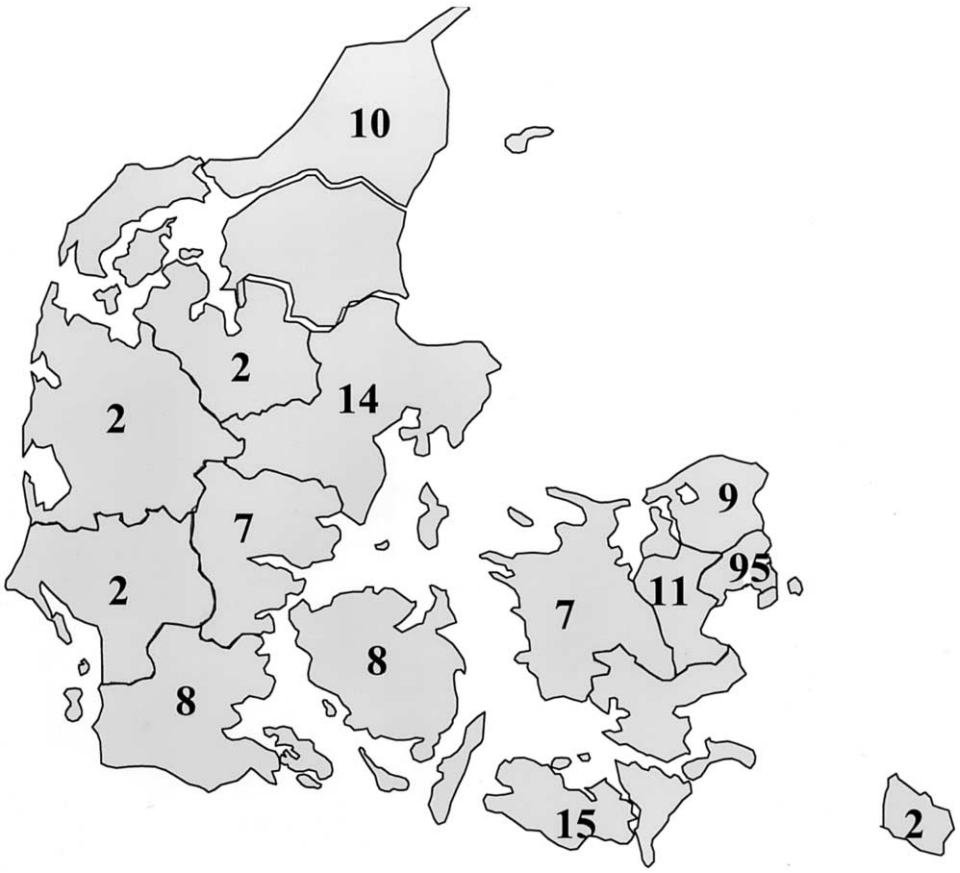

Figure 1. Map of Denmark showing the geographical distribution of 192 of the 205 historical strains of *Mycobacterium tuberculosis*. Nos. indicate the no. of strains provided from each of the 14 counties in Denmark.

than as a result of recent transmission [9]. Given the absence of molecular epidemiological data from many years ago, the extent to which this is true is still unknown.

Here, we present data on strains collected in Denmark 3 decades before the implementation of molecular-subtyping methods, which occurred during the early 1990s. Two hundred three freeze-dried strains collected during the 1960s were recultured, and their RFLP patterns were compared with those of 4102 strains collected during the 1990s, yielding new information about latent infection, reactivation, and the rate of change of DNA patterns.

METHODS

Data collection and analysis. Since 1922, microbiological analyses of mycobacteria have been performed at the International Reference Laboratory of Mycobacteriology at Statens Serum Institut, in Copenhagen. This is the only laboratory that performs culture-based tuberculosis diagnosis for the Danish Kingdom (Denmark, Greenland, and the Faroe Islands). It also serves as an international reference laboratory for Iceland and Lithuania. Because all specimens are processed in a single laboratory and because of the long-standing, mandatory, centralized tuberculosis notification system, we believe our that data are both nearly complete and highly representative of culture-positive tuberculosis from the areas covered [10]. This is a factor of major importance for interpretation of DNA pattern clustering [11, 12].

Between 1961 and 1967, 205 strains of *M. tuberculosis* were collected in Denmark. The majority were collected from patients living in the capital city (Copenhagen) and its surroundings, where most new cases of tuberculosis were, and still are, found (figure 1) [13]. The strains, all of which were collected from patients suspected to be part of various chains of local transmission, were divided into 2 groups by a bacteriophage BK1, and the results were compared with data on epidemiological linkage [13, 14]. They were then stored as freeze-dried samples for 33–39 years, until 203 of the 205 strains were successfully recultured (table 1) [15]. Their DNA patterns have been analyzed (figure 2) and compared with those from 4102 strains collected since 1992, when DNA analysis of *M. tuberculosis* was first implemented on a nationwide basis in Denmark, by use of the RFLP method [5, 16]. The 4102 strains

Table 1. Basic characteristics of patients and historical and recent strains of *Mycobacterium tuberculosis* included in the present study.

Variable	Historical strains (1961–1967) from Denmark[a]	Recent strains (1992–2001)		
		Denmark[a]	Foreign born	Total
Patients notified[b]	5111[c]	1756	2606	4362
Patients in study[d]	201	1672	2264	3936
Age, mean years (SD)	31.4 (21.3)	51.4 (19.5)	32.0 (15.5)	40.4 (19.8)
RFLP-typed strains[e,f]	203	1768	2334	4102
Resistant strains[g]	14 (6.9)	158 (8.9)	218 (9.3)	376 (9.2)
From male patients	115 (57.2)	1134 (64.1)	1252 (53.6)	2386 (58.2)
In cluster	128 (63.1)	1218 (68.9)	1141 (48.9)	2359 (57.5)
In low-copy cluster	13 (6.4)	102 (5.8)	341 (14.6)	443 (10.8)
In cluster 1	0 (0)	170 (9.6)	14 (0.6)	184 (4.5)
In cluster 2	0 (0)	258 (14.6)	14 (0.6)	272 (6.6)

NOTE. Data are no. (%) of subjects, unless otherwise noted.

[a] Including persons from Greenland and the Faroe islands living in Denmark.
[b] Patients diagnosed with tuberculosis in Denmark, including patients not verified by culture result.
[c] Only culture-positive cases, estimated from [21].
[d] Patients with culture-positive tuberculosis included in the present study.
[e] DNA analysis performed by use of the IS*6110* restriction fragment–length polymorphism (RFLP) method [22].
[f] Some patients had >1 *M. tuberculosis* strain examined (see Methods).
[g] RFLP-typed strains resistant for at least 1 of the drugs normally used for treatment: streptomycin, para-amino salicylic acid, and/or isoniazid (during the 1960s) and isoniazid, rifampicin, ethambutol, and/or pyrazinamide (during the 1990s).

were collected from the 3936 patients notified as having tuberculosis in Denmark between 1992 and 2001. More strains than patients were included because, if bacteria were found during the later stages of treatment or if bacteria were isolated from both pulmonary and extrapulmonary sites, RFLP was performed on >1 strain, which occurred for a small number of patients. All strains collected during the 1960s and 1990s (which will be referred to as "historical strains" and "recent strains," respectively) were processed as described elsewhere [15]. The study was approved by the local medical-ethics committees (project journal no. 11-087/99) and the Danish Data Protection Agency (project journal no. 2001-41-1018).

Definitions of terms. "Latent tuberculosis infection" was defined as subclinical infection with *M. tuberculosis* complex bacteria, without clinical, bacteriological, or radiological signs or symptoms of manifest disease [17, 18]. A "cluster" was defined as ⩾2 strains exhibiting 100%-identical DNA patterns, by use of the insertion element IS*6110* as a probe for strain differentiation. "Low-copy cluster strains" harbored only 1–5 IS*6110* copies (bands), whereas "unique strains" differed by at least 1 band from any other strains analyzed [11]. In addition, clusters identified among strains collected during the 1960s were designated "historical clusters." The half-life of IS*6110* DNA patterns was defined as the average time required for 50% of patterns to gain or lose 1 band.

Estimates of the rate of change of DNA patterns during latency. A recent case of tuberculosis was assumed to have been caused by reactivation of an infection with a historical strain acquired during the 1960s when all of the following 3 criteria were met:

1. The strain from the patient was clustered with a historical strain or its pattern was different by 1 band or 2 bands, compared with that of a historical strain (table 2);
2. The patient was alive and in Denmark during the 1960s;
3. The patient had no known direct contact, during the 1990s, with any other patient who was a possible source of infection (table 2).

However, these criteria were not intended to stand alone but were "tested" by the epidemiological data presented in the Results and Discussion sections.

The rate of change (assumed to be constant over time) and the half-life of DNA patterns during latency were estimated by maximum likelihood, by fitting expressions for the expected numbers of recent reactivation cases, as defined above, to those observed. Further details are provided in the Appendix.

RESULTS

Historical and recent cases. Table 1 summarizes the basic characteristics of the historical and recent strains. In brief, between 1961 and 1967, 5111 cases of culture-positive tuberculosis were reported in Denmark: 203 historical strains of *M. tuberculosis* collected from 201 (3.9%) of 5111 patients were included in the present study. The strains were from Danish-born patients with a mean age of 31.4 years (SD, 21.3 years).

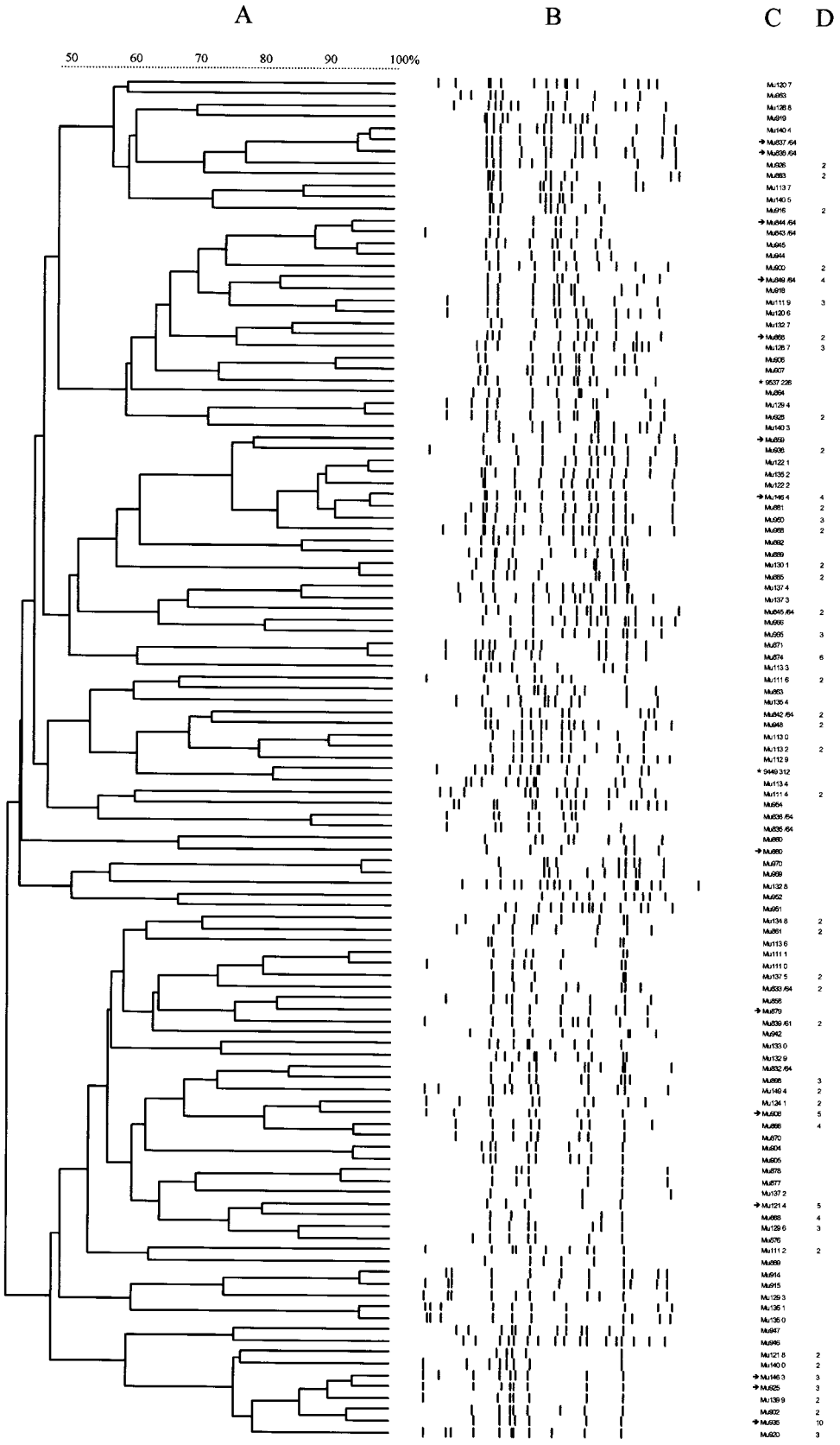

Figure 2. Dendrogram with similarity values in percent *(A)* and 1 branch for each computerized IS*6110* restriction fragment–length polymorphisms pattern *(B)* of the historical strains of *Mycobacterium tuberculosis*. Also included are the 2 most-frequent Danish clusters in the 1990s, from top cluster 2 and cluster 1 *(stars);* the strain numbers *(C);* and the number of strains with each pattern *(D)*. The arrows pinpoint the 14 100%-identical patterns described in table 2.

Table 2. Epidemiological features of the recent *Mycobacterium tuberculosis* strains from patients whose DNA patterns were either identical or different by 1 band or 2 bands, compared with those of the historical strains.

Variable	100% identical	1 band different	2 bands different
IS*6110* RFLP patterns[a]	14	17	7
Patients	40	24	7
Not born or not in Denmark during 1960s	7	4	2
Alive and in Denmark during 1960s	33	20	5
With unique strains during 1990s[b]	4	10	5
With clustered strains during 1990s[b]	29	10	0
Most likely infected during 1990s[c]	1	0	0
Most likely reactivated during 1990s	32	20	5
With epidemiological linkage information[d]	5	0	0
Age, mean years (SD)	62.9 (16.6)	65.3 (16.0)	78.8 (9.1)

NOTE. Data are no. of subjects, unless otherwise noted.

[a] Analysis performed by the IS*6110* restriction fragment–length polymorphism (RFLP) method [22].
[b] See definitions in Methods.
[c] Direct contact to acid-fast bacilli smear–positive patient in same cluster.
[d] Same family name and/or specific geographical origin as the patient(s) with historical strain.

Between 1992 and 2001, 4362 cases of tuberculosis were reported in Denmark: 3936 patients were *M. tuberculosis*–complex positive and were tested by RFLP; 4102 recent strains from these cases were included in the present study, with 1672 (42.5%) of 3936 collected from Danish-born patients. The mean age was 51.4 years (SD, 19.5 years), for the Danish-born patients, and 32.0 years (SD, 15.5 years), for the foreign-born patients. The 2 most frequent clusters (1 and 2) among recent strains were not found among historical strains (figure 2).

***Comparison between the historical and recent strains of* M. tuberculosis.** In total, 14 DNA patterns identified in the strains collected during the 1960s were 100% identical with DNA patterns identified in strains collected during the 1990s (table 2 and figure 2). The identical patterns spanned a period of up to 39 years and accounted for 41 and 40 patients during the 1960s and 1990s, respectively (table 2). During the 1960s, they were found in 8 historical clusters, with 2–10 patients/cluster, and in 6 patients harboring unique strains. During the 1990s, they were found in 10 clusters, with 2–7 patients/cluster, and in 4 patients harboring unique strains.

Seven of the 40 patients from the 1990s were not in Denmark and/or were not yet born during the 1960s, and 5 of these patients were clustered with at least 1 sputum smear–positive patient who was a possible source of infection, from the 1990s. Of the remaining 2 patients who were not accounted for, 1 was a 20-year-old Greenlander, and the other was a 50-year-old Turkish-born man; both were clustered with older patients with extrapulmonary tuberculosis only.

In addition to the 7 patients who could not have been infected during the 1960s, 1 62-year-old Danish-born woman, who had onset of tuberculosis during 1998, was most likely infected by her 66-year-old Danish-born partner, who had onset of sputum smear–positive tuberculosis in 1997 (table 2). This leaves 32 of the 40 patients who fulfilled the criteria for reactivation—that is, who were both alive and in Denmark during the 1960s and were probably not infected during the 1990s (table 2). Only 1 of these patients had onset of tuberculosis within the first 11 months after the start of nationwide RFLP typing in Denmark, in 1992, rendering it less likely that they were infected immediately before the start of RFLP typing. The mean ages of the 32 patients with clustered strains collected during the 1990s and the patients with clustered strains collected during the 1960s were 62.9 years (SD, 16.6 years) and 28.7 years (SD, 21.5 years), respectively (table 2).

In addition to the 14 DNA patterns from strains collected during the 1990s that were 100% identical with patterns identified in strains collected during the 1960s, we found 17 and 7 patterns that were different by only 1 band and 2 bands, respectively (table 2). Their characteristics are also described in table 2. In brief, the patterns different by 1 band and 2 bands accounted for 24 and 7 patients, respectively (table 2), among whom 20 and 5, respectively, were alive and in Denmark during the 1960s (table 2) and had probably not been infected during the 1990s (table 2); 10 of the 20 strains different by 1 band were found in 3 clusters, 1 of which was a low-copy cluster with 2 strains, whereas the remainder were unique (table 2). In total, 57 (32 + 20 + 5) patients with either identical strains or strains different by 1 band or 2 bands most likely developed tuberculosis as a result of reactivation of an infection acquired during the 1960s (table 2).

Estimates of the rate of change of DNA patterns. Assuming that the average time between infection and reactivation

was 30 years and using the numbers of identical patterns and patterns different by 1 band and 2 bands, compared with those of the historical strains observed among reactivated cases (see Appendix), the rate of change of DNA patterns was estimated to be 1.94%/year (95% CI, 1.29%–2.82%). This rate of change corresponds to a half-life of DNA patterns of 36 years (95% CI, 25–54 years).

DISCUSSION

Studies that use DNA-subtyping techniques typically assume that strains of *M. tuberculosis* with matching DNA patterns are epidemiologically related and represent recent (or ongoing) transmission. This assumption is consistent with a short half-life of DNA patterns, as implied by studies of serial isolates collected from patients during relatively short time periods [6, 7]. By use of unique data collected during the 1960s, our study has provided evidence of reactivation of infections acquired several decades ago and suggests that the half-life of DNA patterns of strains involved in latent infection may be much longer than those of strains involved in active disease (36 vs. 3 years). These findings have important implications for the interpretation of studies on clustering, which are increasingly being performed in many populations.

Our estimates of the half-life of DNA patterns relied on several criteria for attributing a case to being the result of reactivation. The criteria included that the patient (1) was alive and in Denmark during the 1960s, (2) harbored strains which were either identical or different by only 1 band or 2 bands, compared with those of the historical strains, and (3) had no known, direct contact with any other patient who was a possible source of infection, during the 1990s. We believe that these criteria were adequate for several reasons.

First, the patients with recent transmission whose patterns were either identical or different by 1 band or 2 bands, compared with those of the historical strains, were, on average, >30 years older than the patients in the historical cohort. Given the fact that contact between individuals is strongly age dependent, with individuals typically acquiring infection from others of a similar age [19], the observed age difference suggests that they could have been infected >30 years before the recent onset of disease and developed disease through reactivation. The assumption that disease among these individuals is attributable to reactivation was further supported by the fact that, in Denmark, the annual risk of infection was very low during the 1990s, compared with that during the 1960s, a finding suggesting that very few of these patients would have been newly infected or reinfected during the 1990s [20].

Second, no epidemiological links were identified between the 57 patients with onset of disease during the 1990s whose isolates either were identical or different by 1 band or 2 bands, compared with those of the historical strains (table 2). Although it is impossible to establish all links between the patients from the 1990s, the finding of no epidemiological links is consistent with the assumption that their initial infection occurred during the 1960s.

Third, the strains collected during the 1990s that either were identical or different by 1 band or 2 bands, compared with those of the historical strains, were found in relatively small clusters and occurred among elderly patients harboring unique strains (table 2). It is likely that, had any of these cases originated from recent active transmission, even during the 1980s, some of the unique strains would have been clustered, given the scale of the DNA subtyping carried out in Denmark, and some of the small clusters would have been larger and included younger patients.

Fourth, 5 of the patients shared family names and/or came from the same geographical region as their clustered "partner" from the 1960s. A direct contact has been confirmed for 1 of the 5 pairs and was reported as the first documented piece of evidence of *M. tuberculosis* reactivation after 33 years of latent infection, on the basis of molecular typing [4]. Attempts to prove that direct contact between the other pairs occurred were unsuccessful, since the patients did not respond to letters, and further investigations were not allowed by Danish law.

The present results indicate that, as the time window increases, the proportion of clustering that involves reactivation increases and that this phenomenon appears to be important [5]. In a rural area in Arkansas, epidemiological links could be found for only 42% of clustered cases during the early 1990s, a finding suggesting that clustering was attributable to simultaneous reactivation of identical strains that had been prevalent in the past [9]. It is also likely that, many years from now, clustering of strains with the 2 patterns that are currently the most prevalent in Denmark (cluster 1 and cluster 2) will be observed as a result of current transmission events (table 1 and figure 2). In the present study, 57 patients with tuberculosis during the 1990s had no epidemiological links to recent cases of infection, but they all had partners with identical or almost identical patterns and clinical disease during the 1960s—that is, tuberculosis up to 39 years before. Thus, our findings emphasize that, with respect to recent transmission, clustering should be interpreted cautiously.

It is essential to know the rate at which IS*6110* DNA patterns change, particularly for long-term population-based studies, such as the present study. If the pace of the "molecular clock" is too rapid for the time window of an epidemiological investigation, linked cases will be missed. Conversely, the long-term stability of patterns may lead to an overestimation of recent transmission [8]. In a previous study, we reported the first case of an IS*6110* DNA pattern that remained stable for nearly 4 decades [15], and here we present a comprehensive analysis of all our historical data. Patterns cannot be assumed to change

at a constant rate, and our observations support the recent suggestion that the rate of change is composed of a high rate during the early disease phase and a low rate during latency [8]. In the present study, the estimated half-life of patterns of strains involved in latent infection was based on 57 cases, and, in all of these cases, latency had lasted for >2 decades, which may have biassed the result toward a longer half-life. Nevertheless, it is noteworthy that, for all these cases, the strains of *M. tuberculosis* harbored in vivo for up to 39 years were almost as genetically stable as the freeze-dried samples. DNA patterns remain very stable during latency.

In conclusion, this comparison between historical and recent strains of *M. tuberculosis* yields convincing molecular, epidemiologic evidence that reactivation occurs up to 39 years after initial infection. It provides new information about the degree of polymorphism and stability of IS*6110* DNA patterns, and it emphasizes that considerable caution should be exercised when interpreting clustering as evidence of recent transmission. On the basis of data on DNA patterns that will accumulate during the next decades, it will be possible, in the future, to give a more precise picture of both latency and reactivation in tuberculosis.

Acknowledgments

We thank Peter Henrik Andersen, for providing the map of Denmark, and Grethe Danielsen, Pia Kristiansen, and Jette Nielsen, for skillful and patient laboratory work. We also thank Birthe Petersen for being capable of sustained effort retrieving data from old cellar archives.

APPENDIX

Estimating the rate of change of DNA patterns. It can be shown (see below) that the numbers of recent patients experiencing reactivation of an infection with a historical strain acquired during the 1960s, whose isolates either are identical (n_0) or different by 1 band (n_1) or 2 bands (n_2), compared with those of the historical strains, are given by the following expressions:

$$n_0 = e_{90}\, e^{-30r}\ , \tag{1}$$

$$n_1 = 30 e_{90} r e^{-30r}\ , \tag{2}$$

$$n_2 = e_{90}(30r)^2 e^{-30r}/2\ , \tag{3}$$

where r is the rate at which 1-band changes (additions/transpositions/deletions) occur in DNA patterns, and e_{90} is the total number of recent patients experiencing disease as a result of reactivation of an infection with a historical strain acquired during the 1960s, irrespective of whether the pattern had changed since infection. These expressions implicitly assume that, for these patients, (1) the time between infection and disease onset was 30 years, (2) once patterns have lost or gained a band in a given position, subsequent band changes do not occur in the same position, and (3) none of the 1- or 2-band differences in the patterns observed during the 1990s had existed during the 1960s.

Estimates of e_{90} and r were derived by fitting the above expressions for n_0, n_1, and n_2, by use of maximum likelihood, to the observed numbers of patients with recent reactivation whose isolates were either identical or different by 1 band and 2 bands, compared with those of the historical strains (denoted by $\hat{n}_0$, $\hat{n}_1$, and $\hat{n}_2$ respectively). The maximum-likelihood method used an algorithm based on the simplex method of Nelder and Mead [23] and involved identifying values for e_{90} and r, which resulted in the maximum value for the following log likelihood expression:

$$L = \sum_{i=0}^{2} -n_i + \hat{n}_i ln(n_i) - ln(\hat{n}_i!)\ .$$

The 95% CIs for e_{90} and r were calculated by profile likelihood—that is, the upper and lower confidence limit for a given parameter was obtained by fixing the other parameter at its best-fitting value and identifying the value for the parameter of interest for which the value for the log likelihood differed by 1.92 from the optimal log likelihood. The half-life of DNA patterns (defined as the time until 50% of patterns have changed by 1 band) is related to the rate of change in patterns, through the following expression:

$$h_{1/2} = -\log_e(0.5)/r\ .$$

Derivation of the expressions for $\mathbf{n}_0$, $\mathbf{n}_1$, and $\mathbf{n}_2$. The above expressions (1–3) can be derived by solving the following expressions, by use of standard integration methods, for the rate of change in the proportion of patients who were infected at the same time with historical strains and who reactivated at a given time (t) after infection with isolates that are either identical or different by 1 band or 2 bands, compared with those of the historical strain (denoted by $f_0(t)$, $f_1(t)$, and $f_2(t)$, respectively):

$$\frac{df_0}{dt} = -rf_0(t)\ ,$$

$$\frac{df_1}{dt} = rf_0(t) - rf_1(t)\ ,$$

$$\frac{df_1}{dt} = rf_0(t) - rf_1(t)\ .$$

The expressions for n_0, n_1, and n_2 follow after substituting $t = 30$ into the final expressions for $f_0(t)$, $f_1(t)$ and $f_2(t)$ and multiplying them by e_{90}.

Results. The rate of change of DNA patterns was estimated to be 1.94%/year (95% CI, 1.29%–2.82%), and the total number of recent patients experiencing disease through reactivation of an infection with a historical strain acquired during the 1960s, with or without the pattern of the strain having changed since infection (e_{90}), was 58 (95% CI, 44–75 patients). The fit of the model to the data was adequate (deviance of 0.116 on 1 *df*) [23].

References

1. Godfrey-Faussett P. DNA fingerprinting: a powerful new tool for the study of tuberculosis. In: Davies PDO, ed. Clinical tuberculosis. London: Chapman & Hall Medical, **1994**:391–400.
2. Small PM, van Embden JDA. Molecular epidemiology of tuberculosis. In: Bloom B, ed. Tuberculosis pathohenesis, protection, and control. Washington, DC: American Society for Microbiology Press, **1994**: 569–82.
3. van Soolingen D. Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements. J Intern Med **2001**; 249:1–26.
4. van Helden PD. Molecular epidemiology of TB: challenging dogmas and asking new questions. IUBMB Life **2002**; 53:219–23.
5. Bauer J, Yang Z, Poulsen S, Andersen AB. Results from 5 years of nationwide DNA fingerprinting of Mycobacterium tuberculosis complex isolates in a country with a low incidence of *M. tuberculosis* infection. J Clin Microbiol **1998**; 36:305–8.
6. de Boer AS, Borgdorff MW, de Haas PE, Nagelkerke NJ, van Embden JD, van Soolingen D. Analysis of rate of change of IS6110 RFLP patterns of *Mycobacterium tuberculosis* based on serial patient isolates. J Infect Dis **1999**; 180:1238–44.
7. Yeh RW, Ponce dL, Agasino CB, et al. Stability of *Mycobacterium tuberculosis* DNA genotypes. J Infect Dis **1998**; 177:1107–11.
8. Warren RM, van der Spuy GD, Richardson M, et al. Calculation of the stability of the IS6110 banding pattern in patients with persistent *Mycobacterium tuberculosis* disease. J Clin Microbiol **2002**; 40:1705–8.
9. Braden CR, Templeton GL, Cave MD, et al. Interpretation of restriction fragment length polymorphism analysis of *Mycobacterium tuberculosis* isolates from a state with a large rural population. J Infect Dis **1997**; 175: 1446–52.
10. Lillebaek T, Andersen AB, Bauer J, et al. Risk of *Mycobacterium tuberculosis* transmission in a low-incidence country due to immigration from high-incidence areas. J Clin Microbiol **2001**; 39:855–61.
11. Glynn JR, Bauer J, de Boer AS, et al. Interpreting DNA fingerprint clusters of *Mycobacterium tuberculosis*. European Concerted Action on Molecular Epidemiology and Control of Tuberculosis. Int J Tuberc Lung Dis **1999**; 3:1055–60.
12. Murray M, Alland D. Methodological problems in the molecular epidemiology of tuberculosis. Am J Epidemiol **2002**; 155:565–71.
13. Baess I. Subdivision of *M. tuberculosis* by means of bacteriophages: with special reference to epidemiological studies. Acta Pathol Microbiol Scand **1969**; 76:464–74.
14. Baess I. A bacteriophage for subdividing the species *M. tuberculosis*. Am Rev Respir Dis **1966**; 93:622–3.
15. Lillebaek T, Dirksen A, Baess I, Strunge B, Thomsen VO, Andersen ÅB. Molecular evidence of endogenous reactivation of *Mycobacterium tuberculosis* after 33 years of latent infection. J Infect Dis **2002**; 185:401–4.
16. Yang ZH, de Haas PE, van Soolingen D, van Embden JD, Andersen AB. Restriction fragment length polymorphism *Mycobacterium tuberculosis* strains isolated from Greenland during 1992: evidence of tuberculosis transmission between Greenland and Denmark. J Clin Microbiol **1994**; 32:3018–25.
17. Broekmans J, Migliori GB, Rieder HL, et al. European framework for tuberculosis control and elimination in countries with a low incidence. Eur Respir J **2002**; 19:765–75.
18. Orme M. The latent tuberculosis bacillus (I'll let you know if I ever meet one). Int J Tuberc Lung Dis **2001**; 5:589–93.
19. Borgdorff MW, Nagelkerke NJ, de Haas PE, van Soolingen D. Transmission of *Mycobacterium tuberculosis* depending on the age and sex of source cases. Am J Epidemiol **2001**; 154:934–43.
20. Styblo K. The relationship between the risk of tuberculosis infection and the risk of developing infectious tuberculosis. Bull Int Union Tuberc **1985**; 60:117–9.
21. Kok-Jensen A, Tørning K. Aspects of the tuberculosis situation in Denmark, especially chemoresistance. Scand J Respir Dis Suppl **1969**; 69:17–20.
22. 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.
23. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes in C: the art of scientific computing. Cambridge: Cambridge University Press, **1992**:408–12.