Molecular epidemiology of tuberculosis after declining incidence, New York City, 2001–2003

C. R. DRIVER[1]*, B. KREISWIRTH[2], M. MACARAIG[1], C. CLARK[1],
S. S. MUNSIFF[1,3], J. DRISCOLL[4] AND B. ZHAO[1]

[1] *New York City Department of Health and Mental Hygiene, Bureau of Tuberculosis Control, New York, NY, USA*
[2] *Public Health Research Institute, Newark, NJ, USA*
[3] *Centers for Disease Control and Prevention, Division of Tuberculosis Elimination, Atlanta, GA, USA*
[4] *Wadsworth Center, New York State Department of Health, Albany, NY, USA*

SUMMARY

Tuberculosis incidence in New York City (NYC) declined between 1992 and 2000 from 51·1 to 16·6 cases per 100 000 population. In January 2001, universal real-time genotyping of TB cases was implemented in NYC. Isolates from culture-confirmed tuberculosis cases from 2001 to 2003 were genotyped using IS*6110* and spoligotype to describe the extent and factors associated with genotype clustering after declining TB incidence. Of 2408 (91·8%) genotyped case isolates, 873 (36·2%) had a pattern indistinguishable from that of another study period case, forming 212 clusters; 248 (28·4%) of the clustered cases had strains believed to have been widely transmitted during the epidemic years in the early 1990s in NYC. An estimated 27·4% (873 minus 212) of the 2408 cases were due to recent infection that progressed to active disease during the study period. Younger age, birth in the United States, homelessness, substance abuse and presence of TB symptoms were independently associated with greater odds of clustering.

BACKGROUND

Tuberculosis (TB) genotyping became available in the late 1980s, coinciding with the resurgence of TB in the late 1980s and early 1990s in the United States. In New York City (NYC), epidemiological investigations of TB outbreaks in hospitals [1–4] and prisons [5] were aided by the availability of the new genotyping tools. As many as one-third of TB cases in NYC were attributed to recent transmission [6]. These studies of TB genotyping in NYC were either hospital-based or conducted among selected subpopulations.

Before 2001, the NYC tuberculosis control programme used selective genotyping of TB cases to investigate TB outbreaks and clusters, potential false-positive cultures, conduct surveillance of TB drug resistance, and in periodic cross-sectional surveys of genotype clustering [6–8]. In addition, several hospitals (the TB Network) performed genotyping of all TB cases diagnosed at these hospitals in 1992–1994 [9, 10] and 1996–1997 [11] and one hospital performed genotyping of all isolates from 1990 onwards [12].

In 2001, the NYC TB control programme began universal genotyping of *Mycobacterium tuberculosis* isolates to increase the identification of TB transmission, efficiency of epidemiological investigation

* Author for correspondence: C. R. Driver, DrPH, RN, New York City Department of Health and Mental Hygiene, Bureau of TB Control, Epidemiology Office, 225 Broadway, 22nd Floor, New York, NY 10007, USA.
(Email: cdriver@health.nyc.gov)

and identification of false-positive cultures. The purpose of the present analyses was to determine the extent of genotype clustering and factors associated with genotype clustering after nearly a decade of declining TB incidence.

METHODS

The study population included all incident, culture-positive TB cases in NYC from 1 January, 2001 to 31 December, 2003. The first *M. tuberculosis* isolate or a sub-culture was submitted to the Public Health Laboratory by the clinical laboratory in which *M. tuberculosis* was first identified. Specimens were sent to the genotyping laboratories for IS*6110* Southern blot hybridization at the Public Health Research Institute (PHRI), and for spacer oligonucleotide typing (spoligotyping) at the Wadsworth Center following standard procedures [13, 14]. Cases were classified as having a clustered genotype if both the IS*6110* restriction fragment length polymorphism (RFLP) and the spoligotype of the isolate were indistinguishable from that of ≥1 isolate in the study period. Other cases were classified as having non-clustered genotypes. Strains known to have been transmitted in the early 1990s were identified from the genotyping laboratory database. The methods for investigating false-positive cultures have been reported elsewhere [15].

We compared the characteristics of case-patients having clustered genotypes to those with unique genotypes. The characteristics included demographic, socio-behavioural, clinical, and bacteriological features of the case-patients, such as infectiousness, presence of acid-fast bacilli (AFB) from a respiratory source and highest number of bacilli on smear microscopy, results of chest radiographs, and prior history of TB. Multidrug-resistant TB (MDR TB) was defined as an isolate having resistance to at least isoniazid and rifampin. Homelessness was defined as the lack of fixed, regular housing or living in a public or private shelter or single-room-occupancy hotel at any time before or at diagnosis, or during TB treatment. Substance abuse included injection and non-injection of illicit drugs during the 12 months before diagnosis. Having prior TB disease was defined as diagnosis and treatment of TB disease in NYC ≥12 months before the current diagnosis. We assumed that one index case per cluster represented reactivated disease and that clustered cases minus one case per cluster represented recently acquired disease [16, 17].

Data collection

We used patient information from the TB registry and genotype information from the TB genotype database of the NYC Department of Health and Mental Hygiene. Demographic and clinical information for each patient was obtained from patient interview and medical-record review, by trained Bureau of Tuberculosis Control staff, on standard data collection forms.

Genotype clusters were investigated for the presence of epidemiological links between cases. Information from the initial patient and contact interviews such as shared contacts, potential sources of TB, potential locations of transmission such as prior hospitalization or shelter residence, prior history of TB infection and disease, characteristics of clustered case-patients, and classification of epidemiological links were entered into a molecular cluster investigation database. If a link was not found from the available information, a re-interview was attempted using a structured questionnaire. An epidemiological link between cases was defined as naming the other person as a contact, having contacts in common, or reporting having been in the same location prior to diagnosis.

Statistical analysis

Statistical analyses were performed with PC SAS software version 8.02 (SAS Institute, Cary, NC, USA). Frequencies and percentages of clustered isolates according to case characteristics were determined. A χ^2 analysis was used for comparison of categorical variables. The Wilcoxon rank-sum test was used for comparison of medians of continuous variables. Differences were considered significant at $P<0.05$. Odds ratios and 95% confidence intervals for an isolate being clustered were derived employing the generalized estimating equation (GEE) regression method, to account for individual characteristics correlated by having the same genotype [18, 19].

The genotyping activities and this analysis received ethical oversight and approval by the New York City Department of Health and Mental Hygiene Institutional Review Board and was reviewed by the Associate Director for Science of the National Center for HIV, STD, and TB Prevention of the CDC; it was determined not to be human-subjects research requiring review.
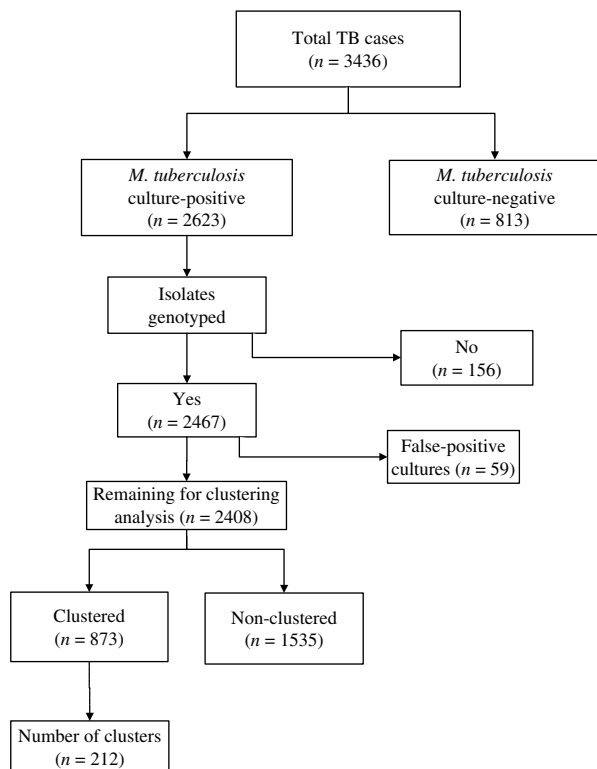
Fig. 1. Classification of TB cases, NYC 2001–2003.

RESULTS

During the study period, 3436 TB cases were reported in NYC. Of these, 2623 (76·3%) were culture-positive and 2467 (94%) were genotyped; 59 (2·4%) isolates were false-positive cultures and 2408 were included in subsequent analyses (Fig. 1). Compared to isolates that were not genotyped, those that were genotyped

were more likely to be from Hispanic TB patients (30% *vs.* 19%, $P=0.001$), from patients with only pulmonary disease (69% *vs.* 59%, $P=0.002$), and patients with AFB smear-positive disease from a respiratory source (43% *vs.* 35%, $P=0.030$). Isolates from Asian TB patients were less likely than other TB cases to be genotyped (28% *vs.* 42%, $P=0.001$).

Among the 2408 TB case isolates that were genotyped, 873 (36·2%) had a pattern that was indistinguishable to that of another TB case within the study period (i.e. the clustered cases). Thirty-one percent (272/873) of the clustered isolates had fewer than four copies of IS*6110*. The 873 clustered cases formed 212 genotype clusters; the median cluster size was 2 (range 2–85) (Fig. 2). There were 266 (11%) cases with strains believed to have been widely transmitted in the early 1990s. Twenty percent (176/873) of clustered cases had one or more epidemiological links to another case in the cluster; among clustered cases with historical strains, 17% had epidemiological links to another case in the cluster. The difference in the proportion of cases that were epidemiologically linked among historical strain cases compared to other strain cases was not statistically significant. An estimated 27·4% (873 minus 212) of the 2408 cases were due to recent infection that progressed to active disease during the study period.

The characteristics of cases and the proportion of cases having a clustered genotype according to these characteristics are shown in Table 1. The crude and adjusted odds ratios for factors associated with having a clustered genotype using GEE logistic regression are shown in Table 2. The following factors were independently associated with genotype
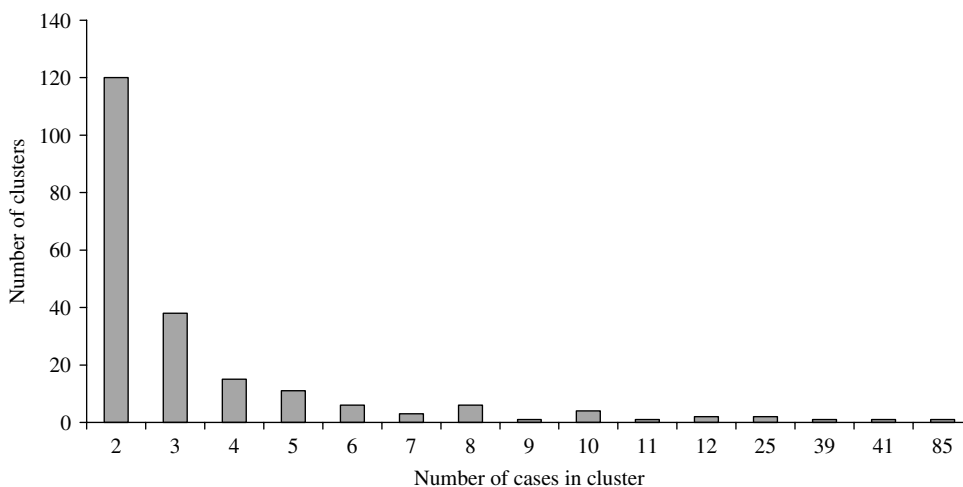

Fig. 2. Frequency of TB genotype clusters by cluster size, NYC 2001–2003 ($n=212$).

Table 1. M. tuberculosis *culture-positive TB cases by selected characteristics, New York City 2001–2003*

	Clustered (n=873)		Non-clustered (n=1535)		Total (n=2408)
	No.	%	No.	%	
Age (yr)*					
<18	43	5	38	2	81
18–34	266	30	504	33	770
35–44	225	26	335	22	560
45–54	165	19	212	14	377
55–64	96	11	161	10	257
≥65	78	9	285	19	363
Sex					
Male	562	64	953	62	1515
Female	311	36	582	38	893
Race*					
Asian	120	14	570	37	690
Hispanic	308	35	431	28	739
Non-Hispanic Black	386	44	413	27	799
Non-Hispanic White	59	7	121	8	180
Country of origin*					
US	419	48	331	22	750
Non-US-born	444	51	1207	79	1651
Unknown	10	1	6	0	16
Years in US, among non-US born					
<1	57	13	179	15	236
1–5	118	27	327	27	445
>5–10	66	15	183	15	249
>10	139	31	385	32	524
Unknown	64	14	131	11	195
Health-care worker in prior 2 years					
Yes	32	4	75	5	107
No	841	96	1460	95	2301
Homeless*					
Yes	113	13	64	4	177
No	750	86	1481	96	2231
Substance abuse*					
Yes	155	18	73	5	228
No	708	81	1472	96	2180
Site of disease*					
Pulmonary only	625	72	1042	68	1667
Extra-pulmonary only	151	17	349	23	500
Both sites	97	11	144	9	241
Cavitary lesions, among cases with pulmonary disease					
Yes	156	22	282	24	438
No	566	78	904	76	1470
Respiratory AFB smear positive*					
Yes	420	48	619	40	1039
No	453	52	916	60	1369
Highest smear grade in first 30 days, among respiratory AFB smear-positive					
1	102	24	173	28	275
2	67	16	114	18	181

Table 1 (*cont.*)

| | Clustered | | Non-clustered | | |
| | ($n=873$) | | ($n=1535$) | | Total |
	No.	%	No.	%	($n=2408$)
3	148	35	200	32	348
4	86	20	103	17	189
Unknown	17	4	29	5	46
HIV serostatus*					
Infected	225	26	177	12	402
Uninfected	449	51	832	54	1281
Unknown	199	23	526	34	725
Multidrug-resistant TB*					
Yes	36	4	36	2	72
No	827	95	1503	98	2330
TB symptoms present at diagnosis*					
Yes	572	66	888	58	1460
No	301	34	647	42	948
Median weeks of symptoms	1·1	0–25·6	1·3	0–26·7	1·4
Prior TB disease*					
Yes	26	3	19	1	45
No	847	97	1516	99	2363

AFB, Acid fast bacilli.

* $P<0.05$.

clustering after adjusting for the variables that were associated with clustering in the bivariate analyses: younger age, birth in the United States, homelessness, substance abuse and presence of TB symptoms. Among non-US-born patients, the number of years of residence in the United States was not associated with having a clustered genotype. Fourteen percent (236/1651) had been in the United States $\leqslant 1$ year before diagnosis; of these only 24 (10%) were examined at entry as part of the immigration screening, three were clustered cases. Of the 212 clusters, 170 (80·1%) had $\geqslant 1$ non-US-born patient in the cluster. Among the 207 clusters in which all patients in the cluster had known country of origin, 90 (42·5%) clusters had only non-US-born patients in the cluster, 37 (17·5%) had only US-born patients, and 80 (37·7%) had both US- and non-US-born patients in the cluster (Fig. 3). In 19 (21%) of the 90 clusters with only non-US-born patients, transmission was believed to have occurred in NYC based on epidemiological links between two or more of the cases in the cluster; 17 were clusters of 2 or 3 cases, the others had 8 and 10 cases respectively. No epidemiological links were identified among the cases in the remaining 71 clusters.
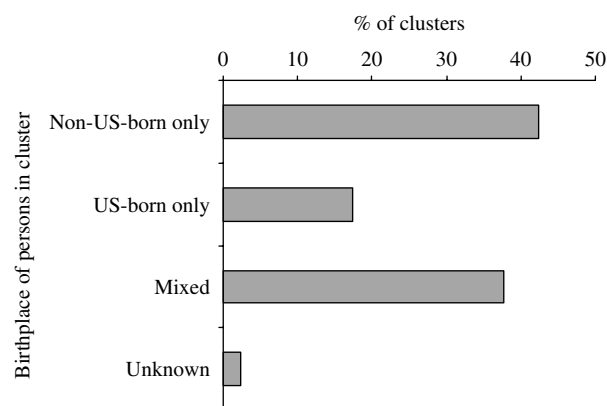

Fig. 3. Birthplace of TB patients in genotype clusters, NYC 2001–2003 ($n=212$).

Eight clusters had more than 10 cases in the cluster; five of these were clusters caused by strains known to have been transmitted in the early 1990s. The IS*6110* gel images and spoligotype patterns for these clusters are shown in Figure 4(*a*, *b*). Two of the largest genotype clusters ($n=85$ and $n=39$ respectively) were associated with recent outbreaks of TB in homeless persons in NYC. Only one of these was highly localized in one single-room-occupancy hotel. The other

Table 2. *Crude and Adjusted GEE odds ratio (OR) for genotype clustering by characteristics of TB cases, New York City 2001–2003*

	Crude		Adjusted	
	OR	(95% CI)	OR	(95% CI)
Age (yr)				
<18	4·13	(2·28–7·40)	3·65	(1·98–6·74)
18–34	1·92	(1·38–2·69)	2·04	(1·40–2·99)
35–44	2·45	(1·70–3·52)	1·90	(1·32–2·72)
45–54	2·84	(1·89–4·27)	2·04	(1·40–2·98)
55–64	2·17	(1·47–3·22)	1·78	(1·19–2·67)
>65	Referent			
Sex				
Male	1·10	(0·90–1·34)		
Female	Referent			
Race				
Asian	0·43	(0·26–0·69)	1·45	(0·90–2·34)
Hispanic	1·46	(0·94–2·27)	0·61	(0·37–1·00)
Black	1·91	(1·25–2·93)	1·42	(0·94–2·15)
White	Referent			
Country of origin				
US	3·53	(2·37–5·27)	2·25	(1·63–3·10)
Non-US-born	Referent			
Unknown	4·53	(1·57–13·03)	4·96	(1·40–17·22)
Years in US, among non-US-born				
<1	Referent			
1–5	1·13	(0·77–1·66)		
6–10	1·13	(0·76–1·68)		
>10	1·13	(0·77–1·66)		
Unknown	1·51	(0·98–2·31)		
Health-care worker				
Yes	0·74	(0·45–1·21)		
No	Referent			
Homeless				
Yes	3·50	(2·11–5·83)	1·55	(1·01–2·38)
No	Referent			
Substance abuse				
Yes	4·42	(2·87–6·80)	1·72	(1·22–2·42)
No	Referent			
Site of disease				
Pulmonary only	1·38	(1·11–1·73)	1·26	(0·97–1·63)
Extrapulmonary only	Referent			
Both sites	1·55	(1·11–2·16)	1·12	(0·78–1·61)
Cavitary lesions, among cases with pulmonary site				
Yes	0·96	(0·78–1·18)		
No	Referent			
Respiratory AFB smear-positive				
Yes	1·37	(1·13–1·65)	1·10	(0·89–1·35)
No	Referent			
Highest smear grade in first 30 days, among respiratory AFB smear-positive				
1	Referent			
2	0·59	(0·67–1·46)		
3	1·25	(0·87–1·80)		

Table 2 (*cont.*)

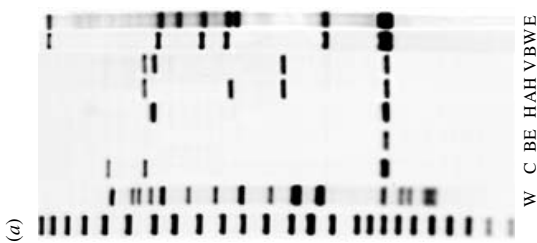|	Crude		Adjusted	
	OR	(95% CI)	OR	(95% CI)
4	1·41	(0·96–2·07)		
Unknown	0·99	(0·51–1·90)		
HIV serostatus				
Infected	1·35	(1·59–3·48)	1·19	(0·86–1·65)
Uninfected	Referent			
Unknown	0·70	(0·57–0·85)	0·90	(0·72–1·12)
Multidrug-resistant TB				
Yes	1·79	(0·81–3·91)	1·76	(0·85–3·63)
No	Referent			
TB symptoms present				
Yes	1·38	(1·16–1·64)	1·21	(1·00–1·47)
No	Referent			
Prior TB disease				
Yes	2·44	(1·35–4·43)	1·78	(0·98–3·22)
No	Referent			

GEE, Generalized estimating equation; AFB, acid fast bacilli.

was a strain known to have been widely transmitted in the early 1990s and associated with stays in shelters for the homeless [19]. Sixteen clusters had one or more MDR TB cases in the cluster; in four (25%) clusters, all with two cases each, both cases had MDR TB isolates; the remaining 12 (75%) clusters had both MDR and non-MDR TB cases in the cluster. There were 12 strain W [21, 22] cases, all of which had MDR isolates.

DISCUSSION

The proportion of clustered cases in the study period was comparable to that seen in NYC at the height of the TB epidemic, when extensive transmission was occurring [2–4, 6, 9, 10, 16, 23–27]. This was probably due to under-ascertainment of genotype clustering in prior studies and highlights the importance of the duration of the study period and sampling on the level of clustering [28, 29]. A longer time period during which cases can enter the sample, and a wider population base, increase the likelihood of identifying clustered cases [30, 31]. The effect of duration is thought to level off after 2–3 years. Our study sample, comprising all incident TB cases in NYC over a 3-year period, covered a longer time span and a larger population than did prior studies, which were either hospital-based or of much shorter duration. One of these studies reported 39 (37·5%) clustered cases of 104 cases diagnosed at a large urban hospital from

1989 to 1992 [16]. A city-wide study of prevalent culture-positive cases, during a 1-month period in 1991, found that 37% (126/344) of cases were clustered; strains of ≤3 IS*6110* bands were excluded from the analysis [6]. Reports of TB cases diagnosed at the TB Network hospitals found 68 (40·7%) of 167 cases diagnosed during 1992–1993 [10] were clustered; 94 (31%) of 302 diagnosed in 1992–1994 [9] and 97 (54%) of 180 in 1996–1997 [11] had clustered genotypes. A 10-year study of TB cases diagnosed at one hospital reported that 63% of cases diagnosed in 1993 had clustered genotypes; the proportion of clustered cases decreased significantly, to 31%, among cases diagnosed in 1999, supporting the notion that genotype clustering was likely to have been higher in earlier years [12]. Compared to other cities in the United States, the proportion of TB cases with clustered genotypes in NYC is lower than that seen in previously published studies from St Louis (MO), Baltimore (MD) and Los Angeles (CA), 39%, 46% and 59% respectively [32–34]. The NYC clustering rate was much higher than that seen in San Fransisco (CA), however, where 19% were clustered in a large population-based study over a 7-year period [35]. Some of the difference in the level of clustering across studies may be due to differences in the proportion of non-US-born TB patients and definitions of clustering. Cases in non-US-born persons are less likely to have isolates with clustered genotypes. San Fransisco had the highest proportion of non-US-born TB

patients, 63 %; on the other hand, the San Fransisco study used a shorter time frame for defining clustering (i.e. two or more isolates within 1 year with indistinguishable isolates). Differences in genotyping methods may also have contributed to the levels of clustering.

Historical strains transmitted in the early 1990s still contributed significantly to genotype clustering of TB cases in NYC in 2001–2003. At least one-third of clustered cases were due to such strains. The extent to which these cases are due to reactivation of infection which occurred in the early 1990s when transmission was widespread, as opposed to recent transmission is not known. This poses a challenge for real-time investigation of genotype clusters, the purpose of which is to identify recent transmission and opportunities for intervention to prevent further spread. Cluster investigations are complex and resource intensive; information on timing and results of previous tuberculin skin tests and absence of overlapping stays among non-US-born patients can be useful for confirming or excluding recent transmission. However, the number of epidemiological links found between cases in this densely populated and mobile city is small. The presence of epidemiological links among cases in clusters of historical strains suggests that these clusters also require investigation for recent transmission.

Our experience in NYC has shown us that universal, real-time TB genotyping is feasible in a large urban centre [15]. Laboratory participation and coverage has been high; the median time from specimen collection to spoligotype result is currently 39 days and 68 days for IS6110 RFLP. The added value from universal genotyping was 57 additional links, 17 additional sites of transmission, four additional investigations in congregate settings in which additional contacts and four secondary cases were identified. Length of unnecessary treatment decreased among patients with false-positive cultures. In addition, genotyping has allowed us to rule out transmission among TB cases that are clustered in place and time but having different genotypes. In such instances, smaller contact investigations can be conducted, rather than the larger case-finding efforts that would be required in settings with transmission confirmed by genotype [15].

There are groups in which genotype clustering continues to be high. US-born, homeless and substance-abusing patients had high rates of clustering, compared to patients without these characteristics.

(b)

IS6110	Spoligotype	Spoligotype pattern	Octal code	SpolDB4* Family	SpolDB4* Frequency
W	S00034		0000000000003771	Beijing	4051
C	S00030		7000367777760731	X3-variant	2244
BE	S00075		7777640776060601	X2-variant	1125
J	S00009		7777677776060601	X2	1069
AH	S00003		7777677776760771	X1	956
V	S00003		7777677776760771	X1	956
BW	S00005		7777777772020771	Haarlem3	1963
E	S00005		7777777772020771	Haarlem3	1963

(a)

W C BE HAH V BWE

Fig. 4. (a) IS6110 gel images for clusters with >10 cases, NYC 2001–2003 (n=8). Note: Cs00030 and BEs00075 were associated with outbreaks in the homeless community and single-room-occupancy hotels. (b) Spoligotype pattern for genotype clusters with >10 cases, NYC 2001–2003 (n=8). * Data from ref. [20].

The number of homeless TB cases increased from 89 in 2001 to 109 in 2003 in NYC (data not shown). Separate investigations have shown evidence of increased transmission of TB in three residences for homeless individuals over the past 3 years: one a large 1001-bed shelter [36], another a facility for homeless persons, and a third a single-room-occupancy hotel (New York City Department of Health and Mental Hygiene, unpublished data).

In summary, despite declining TB incidence TB transmission continues to occur in NYC, among persons who are US-born, homeless and substance abusers. Continued TB control efforts that focus on interrupting transmission in these groups are ongoing. Universal TB genotyping assisted the TB control programme to better understand the dynamics of TB transmission in the city.

ACKNOWLEDGEMENTS

DECLARATION OF INTEREST

None.

REFERENCES

1. **CDC.** Nosocomial transmission of multidrug-resistant tuberculosis among HIV-infected persons – Florida and New York, 1988–1991. *Morbidity and Mortality Weekly Report* 1991; **40**: 129–131.
2. **Edlin BR, et al.** An outbreak of multidrug-resistant tuberculosis among hospitalized patients with the acquired immunodeficiency symdrome. *New England Journal of Medicine* 1992; **326**: 1514–1521.
3. **Pearson ML, et al.** Nosocomial transmission of multi-drug-resistant Mycobacterium tuberculosis: a risk to patients and health care workers. *Annals of Internal Medicine* 1992; **117**: 191–196.
4. **Frieden TR, et al.** A multi-institutional outbreak of highly drug-resistant tuberculosis: epidemiology and clinical outcomes. *Journal of the American Medical Association* 1996; **276**: 1229–1235.
5. **Valway SE, et al.** Outbreak of multi-drug-resistant tuberculosis in a New York State Prison, 1991. *American Journal of Epidemiology* 1994; **140**: 113–122.
6. **Frieden TR, et al.** The molecular epidemiology of tuberculosis in New York City: the importance of nosocomial transmission and laboratory error. *Tubercle and Lung Disease* 1996; **77**: 407–413.
7. **Frieden TR, et al.** The emergence of drug-resistant tuberculosis in New York City. *New England Journal of Medicine* 1993; **328**: 521–526.
8. **Munsiff SS, et al.** Molecular epidemiology of multidrug-resistant tuberculosis, New York City, 1995–1997. *Emerging Infectious Diseases* 2002; **8**: 1230–1238.
9. **Tornieporth NG, et al.** Tuberculosis among foreign-born persons in New York City, 1992–1994: implications for tuberculosis control. *International Journal of Tuberculosis and Lung Diseases* 1997; **1**: 528–535.
10. **Friedman CR, et al.** Transmission of multidrug-resistant tuberculosis in a large urban setting. *American Journal of Respiratory Critical Care Medicine* 1995; **152**: 355–359.
11. **Magnani J, et al.** Molecular epidemiology of tuberculosis among eight hospitals in New York City, 1996–1997. *International Journal of Infectious Diseases* 2001; **5**: 126–132.
12. **Geng E, et al.** Changes in the transmission of tuberculosis in New York City from 1990 to 1999. *New England Journal of Medicine* 2002; **346**: 1453–1458.
13. **Kamerbeek J, et al.** Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. *Journal of Clinical Microbiology* 1997; **35**: 907–914.
14. **Kwara A, et al.** Evaluation of the epidemiologic utility of secondary typing methods for differentiation of Mycobacterium tuberculosis isolates. *Journal of Clinical Microbiology* 2003; **41**: 2683–2685.
15. **Clark CM, et al.** Universal genotyping in tuberculosis control program, New York City, 2001–2003. *Emerging Infectious Diseases* 2006; **12**: 719–724.
16. **Alland D, et al.** Transmission of tuberculosis in New York City. An analysis by DNA fingerprinting and conventional epidemiologic methods. *New England Journal of Medicine* 1994; **330**: 1710–1716.
17. **Ellis BA, et al.** National Tuberculosis Genotyping and Surveillance Network Work Group. Molecular epidemiology of tuberculosis in a sentinel surveillance population. *Emerging Infectious Diseases* 2002; **8**: 1197–1209.
18. **Liang K-Y, Zeger SL.** Longitudinal data analysis using generalized linear models. *Biometrika* 1986; **73**: 13–22.
19. **Zeger SL, Liang K-Y.** Longitudinal data analysis for discrete and continous outcomes. *Biometrics* 1986; **42**: 121–130.
20. **Brudey K, et al.** Mycobacterium tuberculosis complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiology* 2006; **6**: 23.
21. **Munsiff SS, et al.** Persistence of a hightly resistant strain of tuberculosis in New York City during 1990–1999. *Journal of Infectious Diseases* 2003; **188**: 356–363.
22. **Plikaytis BB, et al.** Multiplex PCR assay specific for the mulidrug-resistant strain W of *Mycobacterium tuberculosis*. *Journal of Clinical Microbiology* 1994; **32**: 1542–1546.

23. **Friedman CR,** *et al.* Widespread dissemination of a drug-susceptible strain of *Mycobacterium tuberculosis*. *Journal of Infectious Diseases* 1997; **176**: 478–484.

24. **Coronado VG,** *et al.* Transmission of multidrug-resistant *Mycobacterium tuberculosis* among persons with human immunodeficiency virus infection in an urban hospital: epidemiologic and restriction fragment length polymorphism analysis. *Journal of Infectious Diseases* 1993; **168**: 1052–55.

25. **Nivin B,** *et al.* A continuing outbreak of multidrug-resistant tuberculosis, with transmission in a hospital nursery. *Clinical Infectious Diseases* 1998; **26**: 303–307.

26. **Shafer RW,** *et al.* Temporal trends and transmission patterns during the emergence of multidrug-resistant tuberculosis in New York City: a molecular epidemiologic assessment. *Journal of Infectious Diseases* 1995; **171**: 170–176.

27. **Small PM,** *et al.* Exogenous reinfection with multidrug-resistant *Mycobacterium tuberculosis* in patients with advanced HIV infection. *New England Journal of Medicine* 1993; **328**: 1137–1144.

28. **Vynnycky E,** *et al.* The effect of age and study duration on the relationship between 'clustering' of DNA fingerprint patterns and the proportion of tuberculosis disease attributable to recent transmission. *Epidemiology and Infection* 2001; **126**: 43–62.

29. **van Soolingen D.** Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements. *Journal of Internal Medicine* 2001; **249**: 1–26.

30. **Glynn JR, Vynnycky E, Fine PE.** Influence of sampling on estimates of clustering and recent transmission of Mycobacterium tuberculosis derived from DNA fingerprinting techniques. *American Journal of Epidemiology* 1999; **149**: 366–371.

31. **van Soolingen D,** *et al.* Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. *Journal of Infectious Diseases* 1999; **180**: 726–736.

32. **McConkey SJ,** *et al.* Prospective use of molecular typing of *Mycobacterium tuberculosis* by use of restriction fragment-length polymorphism in a public tuberculosis-control program. *Clinical Infectious Diseases* 2002; **34**: 612–619.

33. **Bishai WR,** *et al.* Molecular and geographic patterns of tuberculosis transmission after 15 years of directly observed therapy. *Journal of the American Medical Association* 1998; **280**: 1679–1684.

34. **Barnes PF,** *et al.* Patterns of tuberculosis transmission in Central Los Angeles. *Journal of the American Medical Associationv* 1997; **278**: 1159–1163.

35. **Jasmer RM,** *et al.* A molecular epidemiologic analysis of tuberculosis trends in San Francisco, 1991–1997. *Annals of Internal Medicine* 1999; **130**: 971–978.

36. **CDC.** Tuberculosis transmission in a homeless shelter population – New York, 2000–2003. *Morbidity and Mortality Weekly Report* 2005; **54**: 149–152.