Epidemiologic Characterization of Culture Positive *Mycobacterium tuberculosis* Patients by *katG*-*gyrA* Principal Genetic Grouping

Carolyn Z. Grimes,* Larry D. Teeter,^{†‡}
Lu-Yu Hwang,* and Edward A. Graviss^{†‡}

From the University of Texas Health Science Center at Houston School of Public Health,* Division of Epidemiology and Disease Control, Houston; the Center for Molecular and Translational Human Infectious Diseases Research,[†] The Methodist Hospital Research Institute, Houston; and the Department of Pathology,[‡] Baylor College of Medicine, Houston, Texas

Molecular typing techniques make it possible to genetically characterize *Mycobacterium tuberculosis* isolates. Public health strategies to control the spread of tuberculosis are enhanced by the use of molecular data to study tuberculosis transmission dynamics within populations. This study compared epidemiological and clinical characteristics of three *M. tuberculosis* groups based on polymorphisms at *katG* codon 463 and *gyrA* codon 95 in 1893 culture-positive patients by a retrospective nested case-comparison design. Study participants, diagnosed from 1995 to 2001 in the Houston, Texas metropolitan area, were ≥ 18 years old, 70% male, 66% U.S.-born, 40% Black, 29% Hispanic, 19% White, and 12% Asian/Pacific Islander. The prevalence of each principal genetic group (GG) was 30% (GG1), 52% (GG2), and 18% (GG3). Multiple logistic regression analysis showed that GG1 participants were more likely to be Asian, male, and have a history of homelessness, as compared with participants with either GG2 or GG3 isolates. GG2 participants were more likely to be Hispanic, have streptomycin-resistant isolates, and be infected with HIV than either GG1 or GG3 participants. GG3 participants were more likely to be Black or Hispanic, report illicit drug use, and live in a congregative facility at the time of diagnosis, than GG1 or GG2 participants. Ethnicity and sociodemographic findings were significant, prompting additional research into social networks, genetic susceptibility, immunology, and virulence factors. (*J Mol Diagn* 2009, 11:472–481; DOI: 10.2353/jmoldx.2009.080171)

The World Health Organization estimated that there were over 9 million new cases of tuberculosis in 2006, including 1.5 million deaths (http://www.who.int/tb/publications/global_report/2008/key_points/en/index.html, accessed

May 12, 2008). One-third of the world's population, or about 2 billion people, have tuberculosis infection (<http://www.who.int/features/factfiles/tuberculosis/en/index.html>, accessed May 12, 2008), and 5% to 10% of these people will develop active disease, in their lifetime, reflecting a vast reservoir of potential tuberculosis cases able to infect others, if not treated. The global tuberculosis (TB) epidemic continues to be studied, clinically and epidemiologically, to find new strategies to slow and eliminate transmission from person to person. Molecular epidemiology is a valuable tool being used to investigate the complexity of the host, environmental, and organism interplay, by identifying and classifying genetic relationships among isolates of *Mycobacterium tuberculosis* (MTB).

Techniques for the molecular characterization of the mycobacterium species that cause TB, such as IS6110 restriction fragment length polymorphism analysis, spoligotyping (spacer oligonucleotide typing), mycobacterial interspersed repetitive units, and principal genetic grouping, have allowed for in-depth studies of populations who have TB disease. These techniques characterize MTB according to its genetic code; therefore isolates of MTB can be distinguished due to variations identified in their genetic make-up. In outbreak situations, this is of particular use because the isolate being transmitted can be tracked from person to person, creating an outbreak cluster, suggesting recent transmission, and necessitating public health action.^{1–6}

MTB can be classified into three principal genetic groups, GG1, GG2, and GG3, based on combinations of polymorphisms at *katG* codon 463 and *gyrA* codon 95.⁶ The study of these principal genetic groups has focused on the evolutionary history of MTB and its dissemination world-wide.^{7–11} While these studies have been important, there also has been an interest in exploring epidemiological links and characteristics of individuals affected by tuberculosis, so as to determine within a given population, which genetic groups are commonly found, and if

Supported in part by federal funds from the National Institute of Allergy and Infectious Diseases and the National Institute of Drug Abuse, National Institutes of Health, under contracts N01-A0-02738 and DA09238.

No author has declared a conflict of interest.

Accepted for publication April 15, 2009.

Address reprint requests to Dr. Edward A. Graviss, The Methodist Hospital Research Institute, 6565 Fannin, MGJ3-012, Houston, Texas 77030. E-mail: eagraviss@tmhs.org.

patterns of characteristics associated with one genetic group versus another give clues to subpopulations that need targeted interventions. The current study aimed to examine characteristics of patients who had been diagnosed with culture-confirmed tuberculosis, whose isolate had been typed into principal genetic groups, GG1, GG2 or GG3, in a well-defined population-based cohort in Harris County (Houston), Texas.

Materials and Methods

Study Population

The current study was a nested case-comparison study, with data derived from an established cohort of TB patients from the Houston Tuberculosis Initiative (HTI) project. The HTI project was a population based study, started in 1995, in which data were prospectively collected on persons with clinically suspected tuberculosis and laboratory diagnosed tuberculosis (positive culture for MTB) in Houston, Texas, and surrounding Harris County. This study was approved by the Institutional Review Board at Baylor College of Medicine. Trained study personnel obtained informed consent, and interviewed patients in the language of their choice, using a standardized questionnaire to collect information on demographics, socioeconomic factors, medical history, drug use, and sexual history. Available MTB isolates from these patients were molecularly characterized. The current study used data collected from October, 1995 through December, 2001, and was limited to adult patients (18 years or older) with an MTB isolate available for study in Harris County. Since 1995, 85% of all reported tuberculosis cases and nearly 90% of all culture positive tuberculosis cases were enrolled.^{12,13} Pediatric and clinically diagnosed cases were excluded. All extrapulmonary and pulmonary cases of tuberculosis were included in this study.

Laboratory Methods

The Houston Department of Health and Human Services laboratory and additional reference laboratories around Houston acquired MTB isolates from surrounding hospitals and laboratories for identification and susceptibility testing, and then transferred the isolates to the HTI project for additional profiling. Molecular characterization of MTB isolates by the standardized IS6110 restriction fragment length polymorphism analysis,¹⁴ and spoligotyping, a supplement to restriction fragment length polymorphism analysis^{10,11,15} was conducted as part of the overall HTI molecular characterization study. As principal genetic grouping is the focus of this paper the other genotyping methods will not be discussed here.

MTB isolates were assigned to their principal genetic group, (GG1, GG2, or GG3), based on nucleotide polymorphisms located on codon 463 of the *katG* gene encoding catalase-peroxidase and codon 95 of the *gyrA* gene encoding the A subunit of DNA gyrase. This methodology was conducted by using DNA sequencing technology.⁶ Drug susceptibility testing was performed by

using BACTEC 460 radiometric culture system at the hospital or reference laboratories supplying isolates to the HTI.¹⁶

Data Analysis

Questionnaire and laboratory data were entered into an EpiInfo database (version 6.02b, Centers for Disease Control and Prevention), and subsequently into a Microsoft Access database (Redmond, Washington). Analysis was completed with the use of STATA 10.0 (College Station, Texas). Bivariate analyses with covariates, known as risk factors for tuberculosis, were conducted for GG1 versus GG2, GG1 versus GG3, and GG2 versus GG3. Covariates included variables describing subject demographics, drug and alcohol history, history of smoking, history of incarceration, medical history, site of tuberculosis disease, and drug resistance. Type of residence was dichotomized into single dwelling (house, trailer, or condominium) versus multiple dwelling (hotel, shelter, treatment facility, or institution). Any drug resistance was defined as having an isolate resistant to any one of isoniazid, rifampin, ethambutol, ethionamide, or streptomycin. Mono drug resistance was analyzed for isoniazid, rifampin, or streptomycin only, as the other drugs did not have enough data to analyze. Cluster status was not included as an independent variable because the three principal genetic groups, along with the restriction fragment length polymorphism analysis and spoligotyping analyses helped define cluster status in the HTI project. Covariates with a *P* value of 0.2 or less, or that were biologically plausible, were carried forward into a multiple logistic regression analysis for each set of comparisons. Covariates with a *P* value of 0.05 or less were considered significant in the final model.

Results

Between 1995 and 2001, there were 2358 adult tuberculosis cases identified and enrolled in the greater Houston area by the HTI. From these, 268 cases (11%) were without an MTB isolate available for culture, and were excluded. Consequently, the total number of adult cases, with a positive culture available for analysis, consisted of 2090 observations (89%) for this study. Of the 2090 isolates available, 1893 (80%) had genetic typing procedures performed. The remaining isolates were not typed due to contamination or lack of secondary growth. The prevalence of each principal genetic group was 30% (563/1893) for GG1, 52% (991/1893) for GG2, and 18% (339/1893) for GG3. The majority of the 1893 cases genetically characterized were Black (40%), male (70%), US-born (65%), and between 18 and 45 years of age (57%).

Collinearity between country of birth status (U.S. born/foreign born) and homelessness, as well as country of birth status and race were found when conducting multiple logistic analyses. A decision was made to report bivariate results for country of birth status and withdraw this variable from the multiple logistic models so that race

Table 1. Bivariate and Multiple Logistic Analysis of Explanatory Variables for GG2 Isolates Compared with GG1 Isolates

Total variable	Total N(P)	GG1 N(P)	GG2 N(P)	OR N(P)	Adjusted OR (95% CI)
Overall	1554	563 (0.36)	991 (0.64)		
Age (Quartiles)					
18–35	478 (0.31)	163 (0.29)	315 (0.32)	Referent	
36–45	409 (0.26)	159 (0.28)	250 (0.25)	1.23 (0.93–1.62)	
46–55	289 (0.19)	102 (0.18)	187 (0.19)	1.05 (0.78–1.43)	
>55	378 (0.24)	139 (0.25)	239 (0.24)	1.12 (0.85–1.49)	
Race					
White	301 (0.19)	99 (0.18)	202 (0.20)	Referent	Referent
Black	585 (0.38)	204 (0.36)	381 (0.39)	1.09 (0.81–1.47)	1.27 (0.93–1.71)
Hispanic	446 (0.29)	81 (0.14)	365 (0.37)	0.45 (0.32–.64)*	.56 (0.39–.79)*
Asian/Pac Is	215 (0.14)	177 (0.32)	38 (0.04)	9.50 (6.21–14.54)*	15.50 (9.72–24.72)*
Sex					
Female	485 (0.31)	172 (0.31)	313 (0.32)	Referent	
Male	1069 (0.69)	391 (0.69)	678 (0.68)	1.05 (0.84–1.31)	1.38 (1.06–1.81)*
Birthplace					
Foreign-born	571 (0.37)	228 (0.41)	343 (0.35)	Referent	
US-born	983 (0.63)	335 (0.59)	648 (0.65)	.78 (0.63–.96)*	
Residence at diagnosis					
Single dwelling	1392 (0.94)	495 (0.93)	897 (0.94)	Referent	
Multi-dwelling	94 (0.06)	36 (0.07)	58 (0.06)	1.12 (0.73–1.73)	
History of homelessness					
No	1275 (0.82)	445 (0.79)	830 (0.84)	Referent	
Yes	277 (0.18)	118 (0.21)	159 (0.16)	1.38 (1.06–1.80)*	1.81 (1.36–2.42)*
History of incarceration					
No	780 (0.50)	298 (0.53)	482 (0.49)	Referent	
Yes	768 (0.50)	264 (0.47)	504 (0.51)	.85 (0.69–1.04)	
History of smoking					
No	546 (0.35)	208 (0.37)	338 (0.34)	Referent	
Yes	1004 (0.65)	355 (0.63)	649 (0.66)	.89 (0.72–1.10)	
History of alcohol use					
No	343 (0.22)	149 (0.26)	194 (0.20)	Referent	
Yes	1207 (0.78)	414 (0.74)	793 (0.80)	.68 (0.53–.87)*	
History of drug use					
No	986 (0.64)	361 (0.64)	625 (0.63)	Referent	
Yes	564 (0.36)	202 (0.36)	362 (0.37)	.97 (0.78–1.18)	
Diagnosed with HIV/AIDS [†]					
No	1242 (0.81)	465 (0.83)	777 (0.79)	Referent	
Yes	299 (0.19)	94 (0.17)	205 (0.21)	.77 (0.58–1.00) [†]	
TST positive [‡]					
No	329 (0.29)	103 (0.24)	226 (0.32)	Referent	
Yes	790 (0.71)	318 (0.76)	472 (0.68)	1.48 (1.13–1.94)*	
History of BCG vaccine					
No	987 (0.72)	341 (0.68)	646 (0.74)	Referent	
Yes	391 (0.28)	162 (0.32)	229 (0.26)	1.34 (1.05–1.70)*	
Site of disease					
Extrapulmonary	354 (0.23)	132 (0.23)	222 (0.22)	Referent	
Pulmonary	1198 (0.77)	430 (0.77)	768 (0.78)	.94 (0.74–1.20)	
Smear status					
Negative	656 (0.42)	247 (0.44)	409 (0.41)	Referent	
Positive	897 (0.58)	315 (0.56)	582 (0.59)	.90 (0.73–1.10)	
Past TB disease					
No	1434 (0.93)	508 (0.91)	926 (0.95)	Referent	
Yes	103 (0.07)	50 (0.09)	53 (0.05)	1.72 (1.15–2.57)*	
History of MOTT [§]					
No	1532 (0.99)	551 (0.99)	981 (0.99)	Referent	
Yes	10 (0.01)	6 (0.01)	4 (0.01)	2.67 (0.75–9.50)	
History of diabetes					
No	1299 (0.84)	463 (0.83)	836 (0.85)	Referent	
Yes	244 (0.16)	96 (0.17)	148 (0.15)	1.17 (0.88–1.55)	
Cough					
No	369 (0.24)	141 (0.25)	228 (0.23)	Referent	
Yes	1174 (0.76)	417 (0.75)	757 (0.77)	.89 (0.70–1.13)	
Hemoptysis					
No	1245 (0.81)	446 (0.80)	799 (0.81)	Referent	
Yes	298 (0.19)	112 (0.20)	186 (0.19)	1.08 (0.83–1.40)	
Chest x-ray					
Not cavitory	626 (0.52)	233 (0.53)	393 (0.51)	Referent	
Cavitory	586 (0.48)	207 (0.47)	379 (0.49)	.92 (0.73–1.16)	

(Continued)

Table 1. (Continued)

Total variable	Total N(P)	GG1 N(P)	GG2 N(P)	OR N(P)	Adjusted OR (95% CI)
Resistance to any one drug					
No	1408 (0.91)	512 (0.91)	896 (0.90)	Referent	
Yes	146 (0.09)	51 (0.09)	95 (0.10)	.94 (0.66–1.34)	
Isoniazid resistance					
No	1487 (0.96)	532 (0.94)	955 (0.96)	Referent	
Yes	67 (0.04)	31 (0.06)	36 (0.04)	1.55 (0.95–2.53) [†]	
Rifampin resistance					
No	1526 (0.98)	549 (0.98)	977 (0.99)	Referent	
Yes	28 (0.02)	14 (0.02)	14 (0.01)	1.78 (0.84–3.76)	
Streptomycin resistance					
No	1458 (0.94)	540 (0.96)	918 (0.93)	Referent	
Yes	96 (0.06)	23 (0.04)	73 (0.07)	.54 (0.33–.87)*	.28 (0.16–.51)*

*P value < 0.05. [†]0.10 > P value > 0.05.

P = Proportions; [†]Verified by medical chart review; [§]History of *Mycobacterium* other than tuberculosis (MOTT).

and homelessness could be accurately assessed in these models.

Principal Genetic Group 1 versus Principal Genetic Group 2

There were 1554 observations for this subanalysis, 563 (36%) patients in GG1 and 991 (64%) in GG2. The results for the bivariate analysis, presented in Table 1, showed the likelihood of having an infection with a GG1 MTB organism was significantly higher in the Asian/Pacific Islander race, and in those persons who had a history of homelessness or past TB disease. Conversely, the likelihood of cases having an isolate in GG1 versus GG2 was significantly lower in the Hispanic population, in those who were US-born, those with a streptomycin resistant organism and in those with a history of consuming alcohol. The study participants that had a positive tuberculin skin test (TST) test or a *bacille Calmette-Guerin* (BCG) vaccine were also more likely to have a GG1 isolate versus GG2 isolate.

Multiple logistic analysis results indicated that Asian/Pacific Islanders, males, and those who had a history of homelessness were more likely to have an isolate belonging to GG1 versus GG2 (Table 1). Hispanics and persons having a streptomycin resistant organism were less likely to have an isolate typed as GG1 versus GG2 (Table 1). These independent factors remained significant after adjusting for age, race, and history of incarceration, smoking, drug and alcohol use history, and type of living situation (single dwelling residence versus a multidwelling residence), homelessness, HIV, past TB history, and smear result. Due to the potential for complex interactions between homelessness, race, and sex in this model, interaction terms were created and explored, however, none of these relationships were significant.

As greater than 10% of the observations were missing for the variables, BCG vaccine history (11%; $n = 1378$) and positive TST history (27%; $n = 1119$), models were re-analyzed because of their statistical significance in the bivariate analysis and their biological contribution to tuberculosis infection and disease. A decrease in the like-

lihood of having a GG1 isolate versus GG2 isolate (Odds Ratio [OR] 0.64, 95% Confidence Interval [CI] = 0.45–.97) was apparent when the participant had a positive history for BCG vaccination with the addition of BCG vaccine history into the model, while gender no longer showed statistical significance. The final model was not altered with the addition of a positive history of TST (data not shown).

Principal Genetic Group 1 versus Principal Genetic Group 3

There were 902 observations for this subanalysis, 563 (62%) patients in GG1 and 339 (38%) in GG3. In the bivariate analysis, Asian/Pacific Islanders were more likely to have an isolate in GG1 versus GG2, as well as participants that had had a BCG vaccine (Table 1). In contrast, the likelihood of individuals having an isolate in GG1 versus GG3 was significantly lower in Hispanics, in U.S. born, in individuals living in a hotel, shelter, treatment facility, or institution (multidwelling residence), incarcerated, and who had a history of smoking, used alcohol, and/or used illicit drugs (Table 2).

A statistically significant increase in the likelihood of GG1 isolates versus GG3 isolates among Asian/Pacific Islanders was found in the multiple logistic analysis. Conversely, a decrease in the likelihood of having an isolate typed as GG1 versus GG3 was found among Hispanics and those who, at the time of diagnosis with tuberculosis, were living in a multidwelling residence (Table 2) after adjusting for age, race, sex, incarceration, smoking, drug and alcohol histories, single-dwelling versus multidwelling residences, homelessness, HIV and past TB history, cough and smear result. Interaction terms were created to analyze possible effect modification by where the person lived with race and HIV infection. A significant interaction term, Asian and living in a single-dwelling residence versus multidwelling residence was found when placed into the model, and a likelihood ratio test performed. A cross tab analysis demonstrated that an unstable cell (1 observation) was created with these two variables, making the interpretation and significance of

Table 2. Bivariate and Multiple Logistic Analyses of Explanatory Variables for GG3 Isolates Compared with GG1 Isolates

Variable	Total N(P)	GG1 N(P)	GG3 N(P)	OR (95% CI)	Adjusted OR (95% CI)
Overall	902	563 (0.62)	339 (0.38)		
Age (Quartiles)					
18–35	250 (0.28)	163 (0.29)	87 (0.25)	Referent	
36–45	260 (0.29)	159 (0.28)	101 (0.30)	.84 (0.59–1.20)	
46–55	173 (0.19)	102 (0.18)	71 (0.21)	.77 (0.51–1.14)	
>55	219 (0.24)	139 (0.25)	80 (0.24)	.93 (0.64–1.35)	
Race					
White	155 (0.17)	99 (0.18)	56 (0.16)	Referent	Referent
Black	367 (0.41)	204 (0.36)	163 (0.48)	.69 (0.47–1.02) [†]	.74 (0.49–1.12)
Hispanic	192 (0.21)	81 (0.14)	111 (0.33)	.40 (0.26–.62)*	.43 (0.27–.69)*
Asian/Pacific Islander	186 (0.21)	177 (0.32)	9 (0.03)	10.90 (5.18–22.97)*	12.18 (5.65–26.26)*
Sex					
Female	261 (0.29)	172 (0.31)	89 (0.26)	Referent	
Male	641 (0.71)	391 (0.69)	250 (0.74)	.81 (0.60–1.09)	
Birthplace					
Foreign-born	313 (0.35)	228 (0.41)	85 (0.25)	Referent	
US-born	589 (0.65)	335 (0.59)	254 (0.75)	.49 (0.37–.66)*	
Residence at diagnosis					
Single dwelling	772 (0.91)	495 (0.93)	277 (0.86)	Referent	
Multi-dwelling	81 (0.09)	36 (0.07)	45 (0.14)	.45 (0.28–.71)*	.52 (0.32–.85)*
History of homelessness					
No	697 (0.77)	445 (0.79)	252 (0.75)	Referent	
Yes	204 (0.23)	118 (0.21)	86 (0.25)	.78 (0.57–1.07)	
History of incarceration					
No	430 (0.48)	298 (0.53)	132 (0.39)	Referent	
Yes	470 (0.52)	264 (0.47)	206 (0.61)	.57 (0.43–.75)*	
History of smoking					
No	307 (0.34)	208 (0.37)	99 (0.29)	Referent	
Yes	595 (0.66)	355 (0.63)	240 (0.71)	.70 (0.53–.94)*	
History of alcohol use					
No	199 (0.22)	149 (0.26)	50 (0.15)	Referent	
Yes	702 (0.78)	414 (0.74)	288 (0.85)	.48 (0.34–.69)*	
History of drug use					
No	542 (0.60)	361 (0.64)	181 (0.54)	Referent	
Yes	358 (0.40)	202 (0.36)	156 (0.46)	.65 (0.49–.85)*	
Diagnosed with HIV/AIDS [‡]					
No	749 (0.84)	465 (0.83)	284 (0.84)	Referent	
Yes	148 (0.16)	94 (0.17)	54 (0.16)	1.06 (0.74–1.53)	1.41 (0.94–2.13) [†]
TST positive [‡]					
No	172 (0.27)	103 (0.24)	69 (0.30)	Referent	
Yes	476 (0.73)	318 (0.76)	158 (0.70)	1.35 (0.94–1.93) [†]	
History of BCG vaccine					
No	591 (0.74)	341 (0.68)	250 (0.84)	Referent	
Yes	208 (0.26)	162 (0.32)	46 (0.16)	2.58 (1.79–3.72)*	
Smear status					
Negative	379 (0.42)	247 (0.44)	132 (0.39)	Referent	
Positive	522 (0.58)	315 (0.56)	207 (0.61)	.81 (0.62–1.07)	
Site of disease					
Extrapulmonary	214 (0.24)	132 (0.23)	82 (0.24)	Referent	
Pulmonary	687 (0.76)	430 (0.77)	257 (0.76)	1.04 (0.76–1.43)	
Past TB disease					
No	816 (0.91)	508 (0.91)	308 (0.92)	Referent	
Yes	77 (0.09)	50 (0.09)	27 (0.08)	1.12 (0.69–1.83)	
History of MOTT [§]					
No	885 (0.99)	551 (0.99)	334 (0.99)	Referent	
Yes	7 (0.01)	6 (0.01)	1 (0.01)	3.64 (0.44–30.34)	
History of diabetes					
No	740 (0.83)	463 (0.83)	277 (0.82)	Referent	
Yes	156 (0.17)	96 (0.17)	60 (0.18)	.96 (0.67–1.37)	
Cough					
No	210 (0.23)	141 (0.25)	69 (0.20)	Referent	
Yes	686 (0.77)	417 (0.75)	269 (0.80)	.76 (0.55–1.05) [†]	
Hemoptysis					
No	719 (0.80)	446 (0.80)	273 (0.81)	Referent	
Yes	175 (0.20)	112 (0.20)	63 (0.19)	1.09 (0.77–1.53)	
Chest x-ray					
Not cavitory	340 (0.48)	207 (0.47)	133 (0.49)	Referent	
Cavitory	374 (0.52)	233 (0.53)	141 (0.51)	1.06 (0.79–1.44)	

(Continued)

Table 2. (Continued)

Variable	Total N(P)	GG1 N(P)	GG3 N(P)	OR (95% CI)	Adjusted OR (95% CI)
Resistance to any one drug					
No	829 (0.92)	512 (0.91)	317 (0.94)	Referent	
Yes	73 (0.08)	51 (0.09)	22 (0.06)	1.44 (0.85–2.41)	
Isoniazid resistance					
No	858 (0.95)	532 (0.94)	326 (0.96)	Referent	
Yes	44 (0.05)	31 (0.06)	13 (0.04)	1.46 (0.75–2.83)	
Rifampin resistance					
No	885 (0.98)	549 (0.98)	336 (0.99)	Referent	
Yes	17 (0.02)	14 (0.02)	3 (0.01)	2.86 (0.81–10.01)	
Streptomycin resistance					
No	869 (0.96)	540 (0.96)	329 (0.97)	Referent	
Yes	33 (0.04)	23 (0.04)	10 (0.03)	1.40 (0.66–2.98)	

**P* value < 0.05. †0.10 > *P* value > 0.05.P = Proportions; ‡Verified by medical chart review; §History of *Mycobacterium* other than tuberculosis (MOTT).

this interaction impossible to explore. This instability could be due to collinearity of these two variables versus the interaction. The OR of the two variables involved in the interaction did not change direction when the interaction term was added.

Since greater than 10% of observations were missing for the variables, BCG vaccine history (11%; *n* = 799) and positive TST history (28%; *n* = 648), models were re-analyzed, including these two variables individually, because of their statistical significance in the bivariate analysis and their biological contribution to tuberculosis infection and disease. The addition of BCG vaccine history to the model did not alter the final model. Living in a multidwelling residence was significant in the final model, but failed to show statistical significance with the addition of TST positivity. Also, with the addition of TST positivity, Blacks demonstrated a statistically decreased likelihood of having a GG1 versus GG3 isolate (OR 0.67, 95% CI 0.45–0.99).

Principal Genetic Group 2 versus Principal Genetic Group 3

There were 1330 observations for this subanalysis, 991 observations in GG 2 (75%) and 339 observations in GG 3 (25%). The bivariate analysis showed a higher likelihood of having an isolate typed as GG2 versus GG3 for participants with an HIV or AIDS diagnosis, a streptomycin resistant organism or a BCG vaccine (Table 3). In contrast, those aged 36 to 45, Black, born in the U.S., living in an institution or treatment facility, hotel or shelter, and those having a history of homelessness, incarceration, alcohol consumption, and illicit drug use, had a significantly lower likelihood of having an isolate in GG2 versus GG3 (Table 3).

Two variables maintained statistical significance with an increased likelihood of having a GG2 isolate versus a GG3 isolate in the multiple logistic analysis: having a medical diagnosis of HIV or AIDS and having a streptomycin resistant organism. On the other hand, Blacks, living in a multidwelling residence at the time of diagnosis, and having a positive illicit drug history, all presented

with a significantly lower likelihood of having a GG2 isolate versus a GG3 isolate (Table 3) after adjusting for age, race, sex, incarceration, smoking, drug and alcohol histories, single-dwelling residence versus a multidwelling residence, homelessness, HIV and past TB history, and smear result. No interactions were found between history of illicit drug use, history of living on the streets, and place where the patient resided with race, or HIV/AIDS diagnosis.

BCG vaccine history (12%; *n* = 1171) and TST positivity (30%; *n* = 925) had more than 10% missing data. BCG vaccine history was statistically significant when included in the model, as individuals had a greater likelihood of having an isolate in GG2 versus GG3 when having a BCG vaccine (OR 1.59, 95% CI = 1.06–2.40). Another variable that became significant was lifetime drug use, as those with a history of drug use were less likely to have an isolate in GG2 versus GG3 (OR 0.55, 95% CI 0.39–0.78) (data not shown). Lifetime TST positivity, when added in the model, did not alter the outcome of the final model.

Discussion

We analyzed epidemiological and clinical characteristics against principal genetic groups based on polymorphisms at *katG* codon 463 and *gyrA* codon 95 in 1893 study participants with culture-confirmed tuberculosis from the Houston metropolitan area. A summary of our results indicated that study participants who had an infection with MTB genetically typed as GG1 were more likely to be Asian, male, and have a history of homelessness. Study participants who had an infection with MTB typed as GG2 were more likely to be Hispanic (when compared with GG1 MTB infections, not GG3 MTB infections), have a diagnosis of HIV or AIDS, and be streptomycin resistant. And, participants with an infection with MTB typed as GG3 were more likely to be Black or Hispanic, have history of illicit drug use, and be living in a congregative (hotel, shelter, treatment facility, or other treatment institution) facility at the time of diagnosis.

Table 3. Bivariate and Multiple Logistic Analyses of Explanatory Variables for GG3 Isolates Compared with GG2 Isolates

Variable	Total N(P)	GG2 N(P)	GG3 N(P)	OR (95% CI)	Adjusted OR (95% CI)
Overall	1330	991 (0.75)	339 (0.25)		
Age (Quartiles)					
18–35	402 (0.30)	315 (0.32)	87 (0.26)	Referent	
36–45	351 (0.26)	250 (0.25)	101 (0.30)	.68 (0.49–.95)*	
46–55	258 (0.20)	187 (0.19)	71 (0.21)	.73 (0.51–1.04) [†]	
>55	319 (0.24)	239 (0.24)	80 (0.23)	.83 (0.58–1.17)	
Race					
White	258 (0.19)	202 (0.20)	56 (0.16)	Referent	Referent
Black	544 (0.41)	381 (0.39)	163 (0.48)	.63 (0.45–.90)*	.64 (0.44–.93)*
Hispanic	476 (0.36)	365 (0.37)	111 (0.33)	.89 (0.62–1.28)	.80 (0.53–1.18)
Asian/Pacific Islander	47 (0.04)	38 (0.04)	9 (0.03)	1.14 (0.52–2.50)	1.00 (0.45–2.28)
Sex					
Female	402 (0.30)	313 (0.32)	89 (0.26)	Referent	
Male	928 (0.70)	678 (0.68)	250 (0.74)	.77 (0.58–1.02) [†]	
Birthplace					
Foreign-born	428 (0.32)	343 (0.35)	85 (0.25)	Referent	
US-born	902 (0.68)	648 (0.65)	254 (0.75)	.63 (0.48–.84)*	
Residence at diagnosis					
Single dwelling	1174 (0.92)	897 (0.94)	277 (0.86)	Referent	
Multi-dwelling	103 (0.08)	58 (0.06)	45 (0.14)	.40 (0.26–.60)*	.48 (0.31–.74)*
History of homeless					
No	1082 (0.82)	830 (0.84)	252 (0.75)	Referent	
Yes	245 (0.18)	159 (0.16)	86 (0.25)	.56 (0.42–.76)*	
History of incarceration					
No	614 (0.46)	482 (0.49)	132 (0.39)	Referent	
Yes	710 (0.54)	504 (0.51)	206 (0.61)	.67 (0.52–.86)*	
History of smoking					
No	437 (0.33)	338 (0.34)	99 (0.29)	Referent	
Yes	889 (0.67)	649 (0.66)	240 (0.71)	.79 (0.61–1.04) [†]	
History of alcohol use					
No	244 (0.18)	194 (0.20)	50 (0.15)	Referent	
Yes	1081 (0.82)	793 (0.80)	288 (0.85)	.71 (0.51–1.00)*	
History of drug use					
No	806 (0.61)	625 (0.63)	181 (0.54)	Referent	
Yes	518 (0.39)	362 (0.37)	156 (0.46)	.67 (0.52–.86)*	.64 (0.48–.86)*
Diagnosed with HIV/AIDS [‡]					
No	1061 (0.80)	777 (0.79)	284 (0.84)	Referent	
Yes	259 (0.20)	205 (0.21)	54 (0.16)	1.39 (1.00–1.93)*	1.81 (1.25–2.61)*
TST positive [‡]					
No	295 (0.32)	226 (0.32)	69 (0.30)	Referent	
Yes	630 (0.68)	472 (0.68)	158 (0.70)	.91 (0.66–1.26)	
History of BCG vaccine					
No	896 (0.77)	646 (0.74)	250 (0.84)	Referent	
Yes	275 (0.23)	229 (0.26)	46 (0.16)	1.93 (1.36–2.73)*	
Smear status					
Negative	541 (0.41)	409 (0.41)	132 (0.39)	Referent	
Positive	789 (0.59)	582 (0.59)	207 (0.61)	.91 (0.71–1.17)	
Site of disease					
Extrapulmonary	304 (0.23)	222 (0.22)	82 (0.24)	Referent	
Pulmonary	1025 (0.77)	768 (0.78)	257 (0.76)	1.10 (0.83–1.48)	
Past TB disease					
No	1234 (0.94)	926 (0.95)	308 (0.92)	Referent	
Yes	80 (0.06)	53 (0.05)	27 (0.08)	.65 (0.40–1.06) [†]	.61 (0.36–1.01) [†]
History of MOTT [§]					
No	1315 (0.99)	981 (0.99)	334 (0.99)	Referent	
Yes	5 (0.01)	4 (0.01)	1 (0.01)	1.36 (0.15–12.23)	
History of diabetes					
No	1113 (0.84)	836 (0.85)	277 (0.82)	Referent	
Yes	208 (0.16)	148 (0.15)	60 (0.18)	.82 (0.59–1.14)	
Cough					
No	297 (0.22)	228 (0.23)	69 (0.20)	Referent	
Yes	1026 (0.78)	757 (0.77)	269 (0.80)	.85 (0.63–1.15)	
Hemoptysis					
No	1072 (0.81)	799 (0.81)	273 (0.81)	Referent	
Yes	249 (0.19)	186 (0.19)	63 (0.19)	1.01 (0.73–1.39)	
Chest x-ray					
Not cavitary	512 (0.49)	379 (0.49)	133 (0.49)	Referent	
Cavitary	534 (0.51)	393 (0.51)	141 (0.51)	.98 (0.74–1.29)	

(Continued)

Table 3. (Continued)

Variable	Total N(P)	GG2 N(P)	GG3 N(P)	OR (95% CI)	Adjusted OR (95% CI)
Resistance to any one drug					
No	1213 (0.91)	896 (0.90)	317 (0.94)	Referent	
Yes	117 (0.09)	95 (0.10)	22 (0.06)	1.53 (0.94–2.47)	
Isoniazid resistance					
No	1281 (0.96)	955 (0.96)	326 (0.96)	Referent	
Yes	49 (0.04)	36 (0.04)	13 (0.04)	.95 (0.50–1.80)	
Rifampin resistance					
No	1313 (0.99)	977 (0.99)	336 (0.99)	Referent	
Yes	17 (0.01)	14 (0.01)	3 (0.01)	1.60 (0.46–5.62)	
Streptomycin resistance					
No	1247 (0.94)	918 (0.93)	329 (0.97)	Referent	
Yes	83 (0.06)	73 (0.07)	10 (0.03)	2.62 (1.34–5.13)	2.57 (1.30–5.08)*

* P value < 0.05. †0.10 > P value > 0.05.

P = Proportions; ‡Verified by medical chart review; §History of *Mycobacterium* other than tuberculosis (MOTT).

Ethnicity was a significant finding in all analyses between principal genetic groups. Asian/Pacific Islander ethnicity has been found consistently in other studies to have a higher risk of being infected with an isolate in GG1 versus GG2 or GG3, and this finding stands when looking at ethnicity and MTB lineage classifications.^{5,12,17–19} Hispanic race was significantly associated with having an infection with an isolate in GG2 or GG3, and Black race was significantly associated with having an infection with an isolate in GG3. Other studies have not found a relationship between any principal genetic group and being Hispanic or Black. The group of MTB strains found in GG1, referred to as the ancestral group, have been traced back to the Asian provinces, and account for the established relationship between Asian race and an infection with a GG1 MTB organism.^{11,18} Even though other studies involving principal genetic grouping as the main outcome variable did not find a significant relationship between Hispanics and Blacks, other studies involving more advanced typing methods have. Re-categorization by spoligotyping of our MTB organisms occurred on a small number of participants when a participant's isolate had less than 6 IS6110 restriction fragment length polymorphism bands, but the numbers were so small, no meaningful data could be analyzed. However, our results are consistent with these other study results, as Hispanics and Blacks may have an affinity for GG2 and/or GG3 because of the geographical disbursement of GG2 and GG3 isolates, the modern strains of MTB, from the Euro-American lineages.^{7,18–20} Two studies have shown genetic particularity of MTB for genetically similar individuals in the same geographical places.^{3,18} As the world's population has expanded and travel and relocation has become easier, the spread of all lineages to different parts of the world will now be observed, and the study of the origins of all lineages will help determine what MTB organisms cause disease in different geographical regions of the world and which ethnicities are greatly affected.^{21,22}

Genetic susceptibility may be different between males and females, and in this study, males were more likely to have an isolate in GG1 versus females. This finding has not been confirmed in other studies.

Having a history of drug use or living in a congregative setting was significantly associated with having an infection with a MTB GG3 isolate, and a history of homelessness was significantly associated with an infection with a MTB GG1 isolate. These are new findings. Tuberculosis spreads rapidly in close environments. All three of these risk factors have the potential to increase spread because of where people use drugs, ie, crack houses, where homeless people are found, ie, crowded shelters, and in institutions or hospitals, because of crowding; however, the other thing that these three risk factors have in common is possible immune suppression. Drug users and homeless people may not have fully functioning immune systems because of their poor health status altered by the use of illicit drugs and/or living on the streets. Certain illicit drugs alter immune susceptibility.²³ Tuberculosis is a social disease, and certain principal genetic group isolates may cluster in different races, drug and alcohol users, homeless individuals, or other social or cultural groups, because these groups tend to congregate in the same environments and social settings. Further study into the link between immunological and genetic factors versus environmental factors will help define the relationship between GG1 and GG3 and these highly susceptible populations.

HIV infected individuals, at the time of TB diagnosis, were more likely to be infected with MTB GG2 isolates in this study, and MTB GG1 isolates, though this latter result was of borderline significance. Other studies that have looked specifically at principal genetic groups by polymorphisms at *katG* codon 463 and *gyrA* codon 95 lacked enough HIV-positive cases to analyze any type of relationship between principal genetic group and HIV.^{2,3,5} One other study found a significant association between HIV infection and the Indo-Oceanic lineage, or GG1.¹⁸ Our data may not have had enough HIV cases to detect significant differences, or perhaps if we had had the capability to split our three principal genetic groups into the six lineages, a particular lineage may have corroborated Gagneux's findings.¹⁸

This study did not find an association between smear status and principal genetic group, which is consistent with other studies.^{5,17}

Drug resistance analysis by principal genetic group showed that having an infection with a GG2 isolate was significantly associated with streptomycin resistance. On further investigation, the majority of the streptomycin-resistant GG2 isolates were foreign born. Limited data are inconsistent with regards to drug resistance, one study found no correlation between molecular characterization and drug resistance,²⁴ but streptomycin resistance was prevalent in GG2 and GG3 isolates in this study. Another study found that the LAM and T lineages, which correspond to GG2 and GG3 isolates, were associated with multidrug resistance,²⁵ including streptomycin. Bifani et al found that streptomycin resistance was associated with the W strain, principally, GG1, in New York City, majority of which was found in foreign born.²⁶ Clearly, drug resistance is related to immigration into an area, and which lineage or GG isolate is prevalent in the same area.

There are some limitations to this study. Recall bias was countered by medical chart review to confirm accuracy of the interviewer and collect missing data. This study excluded those individuals with clinically diagnosed tuberculosis (an MTB culture is necessary for molecular typing), so findings in this study should only apply to populations with culture confirmed tuberculosis. Principal genetic groups were used to categorize MTB isolates in this study versus lineage categorization, as was the current technology at the time of this study. Cross-references can be made however, as in this article, to determine where principal genetic groups fall into the different lineage groups,^{19,27} therefore, possible to compare, on a broad scale, lineage and principal genetic group results. Obviously, significant factors could have been masked, or changed, if lineages were analyzed in this cohort versus the three principal genetic groups. The collinearity found between country of birth (US-born/foreign-born) and a history of homelessness, and between country of birth and race, was an additional limitation. We could not conduct a *priori* hypothesis testing to include US-born and foreign-born status and race variables in a model stratified by homelessness.

Tuberculosis is a complex disease, with interactions between host, environment, and organism. Researchers in basic science circles continue to develop new molecular typing techniques,^{6,8,14,19,28} that are used to tie molecular characteristics of MTB to infected patients and better describe genetic susceptibility of individuals with certain MTB isolates and/or virulence of the isolate. Studies conducted to date, including this study, have conflicting results,^{2,3,5,6,12} and new epidemiological studies analyzing relationships between populations and lineage or principal genetic grouping of MTB will help to tease out real significant differences and will help target individuals at highest risk.

Acknowledgments

Special thanks go to Dr. Richard Grimes for manuscript review, Dr. Ralph Frankowski for data analysis review, and Dr. George Delclos for reviewing work before manuscript preparation.

References

1. Bishai WR, Graham NM, Harrington S, Pope DS, Hooper N, Astemborski J, Sheely L, Vlahov D, Glass GE, Chaisson RE: Molecular and geographic patterns of tuberculosis transmission after 15 years of directly observed therapy. *JAMA* 1998, 280:1679–1684
2. de Bruyn G, Adams GJ, Teeter LD, Soini H, Musser JM, Graviss EA: The contribution of ethnicity to *Mycobacterium tuberculosis* strain clustering. *Int J Tuberc Lung Dis* 2001, 5:633–641
3. Dolzani L, Rosato M, Sartori B, Banfi E, Lagatolla C, Predominato M, Fabris C, Tonin E, Gombac F, Monti-Bragadin C: *Mycobacterium tuberculosis* isolates belonging to katG gyrA group 2 are associated with clustered cases of tuberculosis in Italian patients. *J Med Microbiol* 2004;53:155–159
4. Lari N, Rindi L, Sola C, Bonanni D, Rastogi N, Tortoli E, Garzelli C: Genetic Diversity. Determined on the Basis of katG463 and gyrA95 Polymorphisms, spoligotyping, and IS6110 typing, of *mycobacterium tuberculosis* complex isolates from Italy. *J Clin Microbiol* 2005, 43:1617–1624
5. Rhee JT, Platek AS, Small PM, Harris LM, Chaparro SV, Kramer FR, Alland D: Molecular epidemiologic evaluation of transmissibility and virulence of *Mycobacterium tuberculosis*. *J Clin Microbiol* June 1999;37:1764–1770
6. Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whittam TS, Musser JM: Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci USA* 1997, 94:9869–9874
7. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson K, Kremer K, Parsons LM, Pym AS, Samper S, van Soolingen D, Cole ST: A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci USA* 2002, 99:3684–3689
8. Gutacker MM, Smoot JC, Migliaccio CA, Ricklefs SM, Hua S, Cousins DV, Graviss EA, Shashkina E, Kreiswirth BN, Musser JM: Genome-wide analysis of synonymous single nucleotide polymorphisms in *Mycobacterium tuberculosis* complex organisms: resolution of genetic relationships among closely related microbial strains. *Genetics* 2002, 162:1533–1543
9. Musser JM, Amin A, Ramaswamy S: Negligible genetic diversity of *Mycobacterium tuberculosis* host immune system protein targets: evidence of limited selective pressure. *Genetics* 2000, 155:7–16
10. Soini H, Pan X, Amin A, Graviss EA, Siddiqui A, Musser JM: Characterization of *Mycobacterium tuberculosis* isolates from patients in Houston Texas, by Spoligotyping. *J Clin Microbiol* 2000, 38:669–676
11. van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H, Portaels F, Qing HZ, Enkhsaikan D, Nymadawa P, van Embden J: Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of East Asia. *J Clin Microbiol* 1995, 33:3234–3238
12. El Sahly HM, Adams GJ, Soini H, Teeter L, Musser JM, Graviss EA: Epidemiologic Differences between United States- and Foreign-Born Tuberculosis Patients in Houston Texas. *J Infect Dis* 2001, 183:461–468
13. El Sahly HM, Teeter LD, Pawlak RR, Musser JM, Graviss EA: Drug-resistant tuberculosis: a disease of target populations in Houston Texas. *J Infect* 2006, 53:5–11
14. van Soolingen D, de Haas P, Hermans P, Groenen P, van Embden J: Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of *Mycobacterium tuberculosis*. *J Clin Microbiol* 1993, 31:1987–1995
15. Hayward A, Watson J: Typing of mycobacteria using spoligotyping. *Thorax* 1998, 53:329–330
16. Woods GL: Susceptibility testing for mycobacteria. *Clin Infect Dis* 2000, 31:1209–1215
17. Kong Y, Cave MD, Zhang L, Foxman B, Marrs CF, Bates JH, Yang ZH: Association between *Mycobacterium tuberculosis* Beijing/W lineage strain infection and extrathoracic tuberculosis: insights from epidemiologic and clinical characterization of the three principal genetic groups of *M. tuberculosis* clinical isolates. *J Clin Microbiol* 2007, 45:409–414
18. Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, Nicol M, Niemann S, Kremer K, Gutierrez MC, Hilty M, Hopewell PC, Small PM: Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* 2006, 103:2869–2873

19. Gagneux S, Small PM: Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infect Dis* 2007, 7:328–337
20. Wirth T, Hildebrand F, Allix-Beguec C, Wolbeling F, Kubica T, Kremer K, van Soolingen D, Rusch-Gerdes S, Locht C, Brisse S, Meyer A, Supply P, Niemann S: Origin, spread and demography of the *Mycobacterium tuberculosis* complex. *PLoS Pathog* 2008, 4:e1000160
21. Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajj SA, Allix C, Aristimuno L, Arora J, Baumanis V, Binder L, Cafrune P, Cataldi A, Cheong S, Diel R, Ellermeier C, Evans JT, Fauville-Dufaux M, Ferdinand S, de Viedma DG, Garzelli C, Gazzola L, Gomes HM, Gutierrez MC, Hawkey PM, van Helden PD, Kadival GV, Kreiswirth B, Kremer K, Kubin M, Kulkarni SP, Liens B, Lillebaek T, Ly HM, Martin C, Martin C, Mokrousov I, Narvskaia O, Ngeow YF, Naumann L, Niemann S, Parwati I, Rahim Z, Rasolofo-Razanamparany V, Rasolonalana T, Rossetti ML, Rusch-Gerdes S, Sajduda A, Samper S, Shemyakin IG, Singh UB, Somoskov A, Skuce RA, van Soolingen D, Streicher EM, Suffys PN, Tortoli E, Tracevska T, Vincent V, Victor TC, Warren RM, Yap SF, Zaman K, Portaels F, Rastogi N, Sola C: *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol* 2006, 6:23
22. Hirsh AE, Tsolaki AG, DeRiemer K, Feldman MW, Small PM: Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. *Proc Natl Acad Sci USA* 2004, 101:4871–4876
23. Friedman H, Pross S, Klein TW: Addictive drugs and their relationship with infectious diseases. *FEMS Immunol Med Microbiol* 2006, 47:330–342
24. Ramaswamy SV, Dou SJ, Rendon A, Yang Z, Cave MD, Graviss EA: Genotypic analysis of multidrug-resistant *Mycobacterium tuberculosis* isolates from Monterrey, Mexico. *J Med Microbiol* 2004, 53:107–113
25. Kibiki GS, Mulder B, Dolmans WM, de Beer JL, Boeree M, Sam N, van Soolingen D, Sola C, van der Zanden A: *M. tuberculosis* genotypic diversity and drug susceptibility pattern in HIV-infected and non-HIV-infected patients in northern Tanzania. *BMC Microbiol* 2007, 7:51
26. Bifani P, Mathema B, Campo M, Moghazeh S, Nivin B, Shashkina E, Driscoll J, Munsiff SS, Frothingham R, Kreiswirth BN: Molecular identification of streptomycin monoresistant *Mycobacterium tuberculosis* related to multidrug-resistant W strain. *Emerg Infect Dis* 2001, 7:842–848
27. Kulkarni S, Sola C, Filliol I, Rastogi N, Kadival G: Spoligotyping of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in Mumbai India. *Res Microbiol* 2005, 156:588–596
28. Cowan LS, Mosher L, Diem L, Massey JP, Crawford JT: Variable-number tandem repeat typing of *Mycobacterium tuberculosis* isolates with low copy numbers of IS6110 by using mycobacterial interspersed repetitive units. *J Clin Microbiol* 2002, 40:1592–1602