TRANSMISSION OF TB IN GEORGIA, 2009-2015: A MOLECULAR EPIDEMIOLOGICAL APPROACH

by

EDWARD ABDUL SHERIFF

(Under the Direction of Juliet Nabbuye Sekandi)

ABSTRACT

Statement of Problem: As opposed to the United States, the epidemiology of TB in Georgia is fueled by recent transmission. If current trends in TB incidence continue, elimination of TB in Georgia and the United States in accordance with the 2050 global targets for tuberculosis will not be met. **Goal**: Conduct a secondary analysis of TB data to expand knowledge on the molecular epidemiology of TB in Georgia by examining the lineage diversity, transmission dynamic and persistence of Mycobacterium tuberculosis strains. **Methods**: Secondary analysis of TB surveillance and molecular data from 2009 to 2015. Cases surveillance data was linked to molecular data using a state case number. In Aim 1, we examined the diversity of lineages and their association with patient characteristics. In Aim 2 we examined the risk factors associated with recent transmission of tuberculosis and in Aim 3 we examined the persistence of TB strains. **Results:** The most predominant lineage in our study was lineage 4 or Euro-American lineage. A patient's race and country of origin were significantly associated with Mycobacterium tuberculosis lineage. In aim 2 we found that 30% of the isolates in Georgia from 2009 to 2015 were clustered. The rate of transmission was estimated to be 25%. US-born cases that are Black, homeless or have resided in correctional facilities were more likely to be involved in a chain of recent transmission. Isoniazid resistance was associated with recent transmission and persistence of Mycobacterium tuberculosis strains. **Conclusion**: Our study shows a significant diversity of TB strains and a geographical relationship between these lineages and their host's country of origin. TB in Georgia is mainly transmitted among racial minorities born in the United States who are either homeless or have been incarcerated. TB control efforts in Georgia should focus more resources on addressing drug resistance and transmission in homeless shelters and other congregate settings.

INDEX WORDS: Tuberculosis, Molecular Epidemiology, Clustering of isolates, Strain persistence, Recent Transmission

TRANSMISSION OF TB IN GEORGIA, 2009-2015: A MOLECULAR EPIDEMIOLOGIC APPROACH

by

EDWARD ABDUL SHERIFF

BS, George Washington University, 2008

MPH, George Washington University, 2011

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2018

TRANSMISSION OF TB IN GEORGIA, 2009-2015: A MOLECULAR

EPIDEMIOLOGIC APPROACH

by

EDWARD ABDUL SHERIFF

Major Professor:	Juliet Nabbuye Sekandi
Committee:	Christopher Whalen
	Kevin Dobbin
	Andrea Swartzendruber

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2018

DEDICATION

Dedicated to Nana and Munda. Thanks for letting me see the light of day. You’re not here today, but you live on. “Hakuna Matata”

“To Realize One’s Destiny is a person’s Only Obligation”

-Paulo Coelho

ACKNOWLEDGEMENTS

This study was made possible through collaboration and support of many individuals and organizations. This study would have not been possible without the help of the Georgia Department of health. I would especially like to express gratitude to Antoine Perrymon from DPH for helping me navigate the data request process and answering my many data questions. I would like to express my heartfelt gratitude to my dissertation committee for their patience and guidance throughout this process. None of this would have been possible without your continuous support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

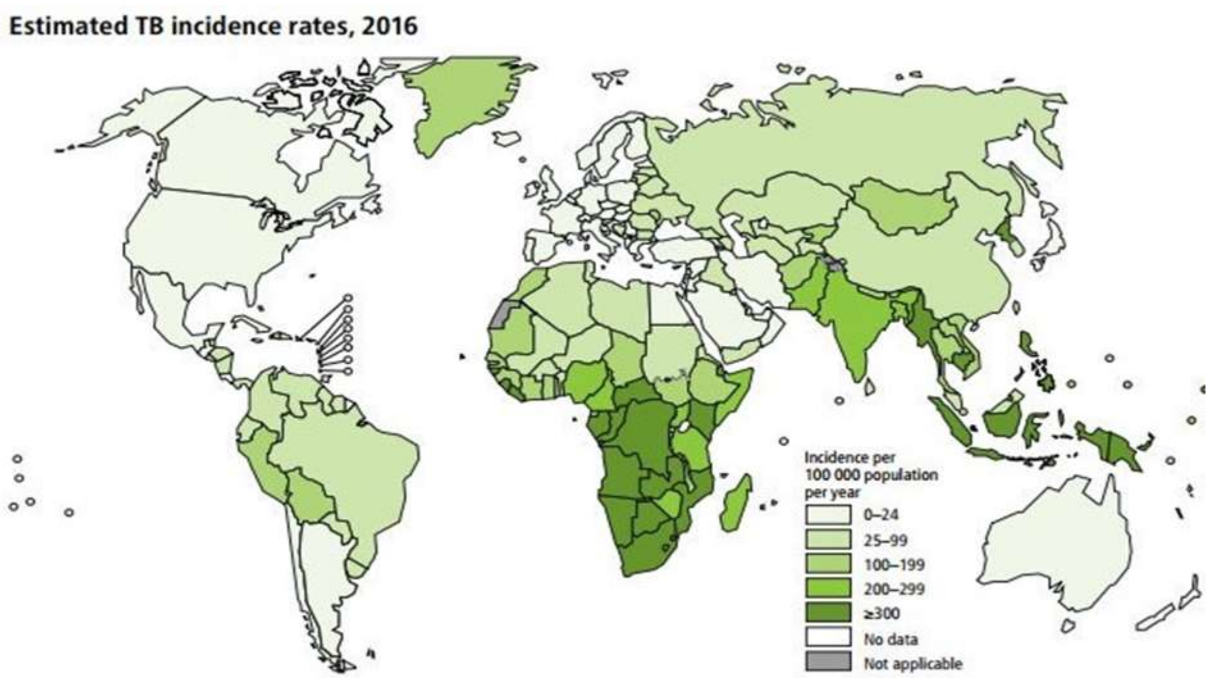
CHAPTER 1

INTRODUCTION

Epidemiology of Tuberculosis

Tuberculosis (TB) is an infectious disease caused by the bacillus Mycobacterium tuberculosis. Mycobacterium tuberculosis typically infects the lungs (pulmonary TB), but can also infect other sites (extra pulmonary TB). TB is an important public health problem across the globe in spite of the resources being invested in its control and eventual elimination. It is the ninth leading cause of death worldwide, and the leading cause from a single infectious agent (WHO, 2017). An estimated 10.4 million people became ill with TB in 2016 (WHO, 2017).

Preventive therapy for TB has expanded among vulnerable populations like people living with HIV/AIDS, and children under 5 years of age (WHO, 2017). However, most people eligible for TB preventive treatment do not have access to it. The burden of TB is highest in countries that are highly populated with high levels of poverty. Figure 1 shows the estimated global incidence rates of TB in 2016. The HIV/AIDS epidemic, the emergence of resistance Mycobacterium and active transmission has exacerbated the TB crisis around the world.



Source: WHO Report, 2017

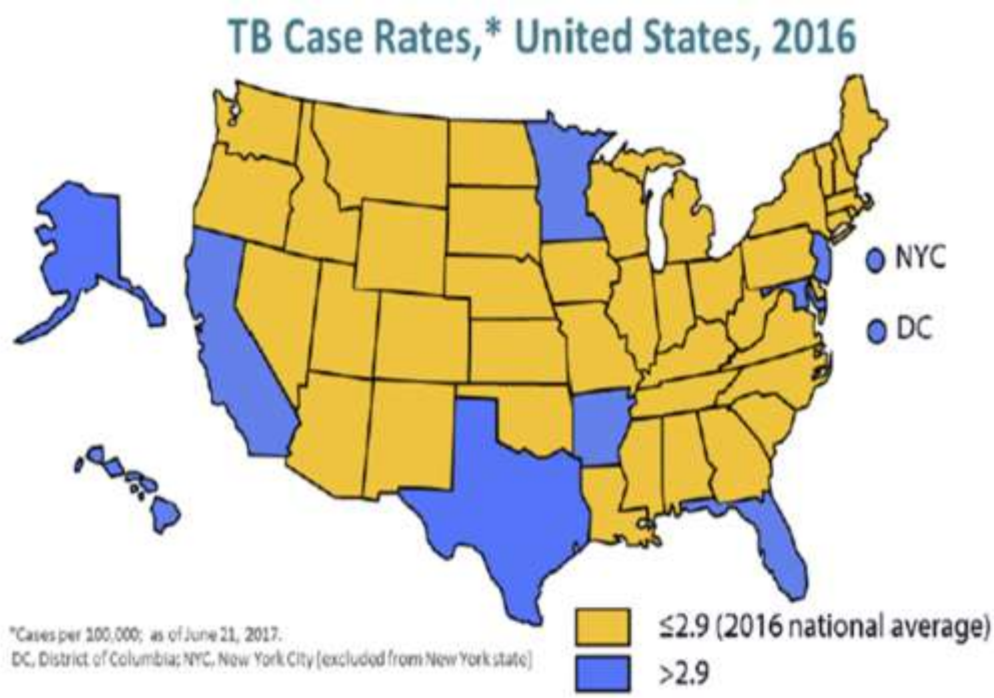
Figure 1: Estimated global incidence rates of tuberculosis (2016)

TB burden in the United States

In the United States, 9,272 cases of TB were reported to the Centers for Disease Control and Prevention (CDC) in 2016. This represents a 2.9% decline in cases from 2015 (CDC, 2017). The national incidence rate was 2.9 cases per 100,000 persons and cases were reported from all 50 states (CDC, 2017). TB case counts were highest in California, Texas, New York (including New York City) and Florida. These four states accounted for more than 50% of the total cases in the United States (CDC, 2017). Figure 2 shows TB case rates by states, in the United States in 2016.

In the United States country of birth continues to be a risk factor for TB and 68.5% of TB cases were reported in non-U.S.-born persons (CDC, 2017). U.S.-born persons had a TB incidence rate of 1.1 cases per 100,000 compared to 14.7 cases per 100,000 among non-U.S.-born persons (CDC, 2017). Among non-U.S.-born persons reported with TB in 2016, the top 5 countries of birth were Mexico, the Philippines, India, Vietnam, and China (CDC, 2017). Among U.S.-born persons reported with TB disease in 2016, non-Hispanic Blacks were most commonly represented (36%) followed by non-Hispanic whites (31.5%) and Hispanics (20.8%) (CDC, 2017). Among non-U.S.-born persons reported with TB disease in 2016, non-Hispanic Asians (47.9%) were the largest group, followed by Hispanics (31.3%) and non-Hispanic blacks (14.3%) (CDC, 2017).

Residence in congregate settings like homeless shelters and prisons remain a risk factor for TB infection, which can subsequently progress to TB disease (CDC, 2017). In 2016, 4.0% of all TB cases in the United States were residents of correctional facilities and 4.9% experienced homelessness in the year preceding diagnosis (CDC, 2017).



Source: Division of Tuberculosis Elimination, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention

Figure 2: TB Case Rates in United States, 2016

TB burden in Georgia

In 2016, Georgia reported 301 TB cases, which equates to a 6% decline from TB cases in 2015 (DPH, 2017). This represents a case rate of 2.9 cases per 100,000 persons (DPH, 2017). According to the Centers for Disease Control and Prevention, Georgia ranked 6th in the United States for number of new TB cases and 10th for TB case rate among the 50 reporting states in 2016 (CDC, 2017). The state of Georgia is made up of 159 counties, but only 4 counties accounted for 53% of the TB cases reported in 2016 (DPH, 2017). These counties are DeKalb, Fulton, Gwinnett, and Cobb counties.

In 2016, TB among persons born outside of the United States accounted for 48% of all TB cases in Georgia. Most non-US-born cases came from countries were TB is endemic. The number of non-US-born cases is disproportionately distributed in Georgia, with 72% of all non-US-born cases being reported by only 4 counties-DeKalb, Fulton, Gwinnett and Cobb counties (DPH, 2017). Reports show that 18 people died of TB in Georgia in 2016. The age-adjusted TB mortality rate of 0.2 deaths per 100,000 (DPH, 2017).

18 health districts have oversight responsibility for public health in the 159 counties. DeKalb Health District had the highest TB case rate in 2016 (7.8 per 100,000) followed by Columbus District (4.5 per 100,000) (DPH, 2017). Fulton and Clayton health district both had a TB case rate of 4.3 cases per 100,000 (DPH, 2017). Both Georgia and the United States have not seen any significant decline in TB case rate since 2013.

TB infection rates in the United States are low compared to other areas of the world. However, it continues to pose a real threat to the country’s public health infrastructure. TB is preventable and treatable, but remains a serious problem in the United States.

Statement of Problem

The epidemiology of TB in Georgia is different from the United States. TB cases in the United States are a result of reactivation of latent TB infection. On the contrary, TB in Georgia is mostly a result of recent transmission. Recent transmission of TB is more common among US-born persons. More than half of the TB cases in Georgia are reported among US-born persons. It is therefore important to understand the patterns of transmission of TB in Georgia.

TB and poverty are closely linked. Risk factors associated with poverty-homelessness, incarceration, increase both the probability of becoming infected and the probability of developing clinical disease. Together, poverty and Mycobacterium tuberculosis form a vicious cycle. Poor people live in close, unhygienic quarters where TB flourishes; TB decreases people’s capacity to work and adds treatment expenses, thereby exacerbating their poverty. Meanwhile, the poor receive inadequate health care, preventing the diagnosis of their tuberculosis. In the United States, treatment cost per case ranges from $18,000 up to $513,000 for extensively drug-resistant TB (CDC, 2017). TB is an economic burden not only on the individual but the community as a whole. Understanding the transmission pattern of TB in Georgia will lead to initiatives that will interrupt TB transmissions and eventually lead to the elimination of TB in accordance with the End TB strategy (WHO, 2017).

Molecular epidemiologic analysis has been used to respond to outbreaks of TB, assess the global dissemination of TB strains, identify groups at risk for recent transmission and examine TB clusters in major cities in the United States. This is done by examining the extent of genetic diversity of clinical strains of Mycobacterium tuberculosis. The molecular epidemiology of TB in the state of Georgia is yet to be fully elucidated. Exploring the molecular epidemiology of TB in Georgia is relevant, because it will provide knowledge on the diversity of Mycobacterium tuberculosis complex (MTBC) strains in Georgia and how strains have persisted over time. Due to the natural history of TB, where individuals may remain latently infected for decades before they develop active disease. The TB population in urban cities like Atlanta, consists of strains contracted not only in a diversity of locations but also over an extended period of time. Hence, such centers are concentrations of geographically and temporally disparate samples of TB as well as potential crossroads in global transmission.

The number of homeless persons in the city of Atlanta has increased. Studies have shown that in cities were homelessness is a common problem, the homeless shelters often serve as hot spots for TB transmission (Bamrah et al., 2013; Barnes et al., 1996; Powell et al., 2017). The objective of this study is to provide a comprehensive description of the circulating strains of Mycobacterium tuberculosis, their persistence in the population over time, and transmission patterns of TB in Georgia from 2009 to 2015. Molecular epidemiologic studies have added much-needed accuracy and precision in describing transmission dynamics, and they have facilitated investigation of previously unresolved issues, such as estimates of recent transmission versus reactivation of latent infection and the extent of exogenous infection (Jasmer et al., 1999; Kempf et al., 2005; Murray, 2002). The results from this study can influence control measures directed to stopping transmission of TB by providing more insight into the transmission patterns of Mycobacterium tuberculosis in the state of Georgia.

Purpose and Objectives

The goal of this dissertation is to contribute to the understanding of the molecular epidemiology of Mycobacterium tuberculosis in Georgia and the United States. To achieve this goal, we used surveillance and molecular data collected by the Georgia Department of Public Health Tuberculosis Prevention and Control Program from 2009 to 2015. The objectives of this research were categorized into three specific aims.

Specific Aim 1: To examine the association between Mycobacterium tuberculosis lineages and TB patient characteristics in Georgia from 2009 to 2015.

Rationale: Genetically distinct pathogens have different degrees of fitness, virulence, and clinical outcomes (Visser et al., 2012). TB patients, too, differ in terms of exposure to risk factors, place of origin, and the presence of other risk factors. A better understanding of the possible correlation between Mycobacterium tuberculosis lineages and host characteristics may identify factors predictive of ongoing transmission. It will also shed some light onto the biology of Mycobacterium tuberculosis.

Analytic Approach: We used Chi-square test to test the statistical significance of differences between groups in binary variables like gender and country of birth. We used Kruskal-Wallis rank test for the continuous variable age. P-values less than 0.05 were considered statistically significant. We used multinomial logistic regression to examine patient characteristics associated with Mycobacterium tuberculosis lineages. All statistical analyses were performed using SAS 9.4 (Cary, NC).

Specific Aims 2: To examine the risk factors of clustering of isolates from TB patients in Georgia from 2009 to 2015.

Rationale: Clustering of TB isolates has been used as a proxy for recent transmission of TB. Patients with identical isolates of *Mycobacterium tuberculosis* are involved in a chain of recent transmission, provided epidemiologic data supports this claim. Recent transmission of TB indicated by clustering is responsible for a significant number of TB cases in Georgia, and, therefore, increased emphasis should be placed on the identification of transmission sites and patient characteristics associated with clustering of TB.

Analytic Approach: Isolates were divided into 2 groups: clustered and non-clustered isolates as indicated by genetic testing and the plausible source case approach (France, Grant, Kammerer, & Navin, 2015). Descriptive statistics, including cross tabulations of demographic and clinical characteristics were computed. Chi square test or the Fisher's exact tests when applicable were used in a bivariate analysis to assess the relationship between clustering of TB isolates and patient characteristics. Binomial logistic regression was used to compare patient characteristics between clustered and unique isolates. P-values less than 0.05 were considered statistically significant and all statistical analyses were performed using SAS 9.4 (Cary, NC).

Specific Aim 3: To examine the persistence of MTBC strains in Georgia from 2009 to 2015.

Rationale: Given that TB occurs disproportionately in certain demographic groups like homeless and prison inmates, the persistence of any particular strain of MTBC will depend in which demographic group it first appeared. One will expect strains that first appear in urban areas to spread more rapidly than those in rural areas. However, in rural communities, certain strains of TB may persist because of limited mobility in the population. Understanding the persistence of strains in terms of their frequency and duration in the population will guide our understanding into the transmission dynamic of TB in Georgia. It is important to understand the patient demographic characteristics associated with the most persistent strains of MTBC in Georgia.

Analytic Approach: Clustered isolates were divided into three groups: persistently successful strains, transiently successful strains, and unsuccessful strains based on criteria that incorporates their duration in the community and the number of cases associated with that strain. Bivariate analyses employed Chi-square tests for independence. Multinomial logistic regression analysis was used to compare persistently successful strains with transiently successful and unsuccessful strains. All statistical tests were two-sided at alpha=0.005. All statistical analyses were done using SAS 9.4 (Cary, NC).

Significance of Research

This study provides insight into the extent to which lineages of Mycobacterium tuberculosis have adapted to specific human populations and how differences in genetic contents of each lineage affects the presentation of disease. This research provides a comprehensive picture of the transmission dynamic of TB in Georgia. For the first time we are examining the success of certain strains of Mycobacterium tuberculosis in Georgia. This analysis will provide a broader view of the TB problem in Georgia by using multiple years of surveillance and molecular data. Tuberculosis control requires research into new preventive measures and better tools are needed to control TB worldwide. This research highlights the markers for high-risk individuals, in terms of TB development in the state of Georgia. Additionally, there is inconclusive evidence about the role if any of the strains of Mycobacterium tuberculosis in disease manifestation in low incidence settings. To our knowledge, no research of this magnitude and scope has been undertaken in the state of Georgia.

Study Outline

Chapter 2 of this dissertation details the literature that is relevant to a full understanding of this molecular approach to studying TB transmission and gaps in literature. Chapter 3 details the methods that were used in conducting this research, including data sources and statistical analysis. Chapters 4, 5, and 6 will describe specific aims 1, 2, and 3 respectively in manuscript format. Chapter 7 provides a summary and conclusion. The sections in chapters 4, 5, and 6 will include background, methods, results, discussions and conclusions.

References: Chapter 1

Bamrah, S., Yelk Woodruff, R., Powell, K., Ghosh, S., Kammerer, J., & Haddad, M. (2013). Tuberculosis among the homeless, United States, 1994–2010. *The International Journal of Tuberculosis and Lung Disease, 17*(11), 1414-1419.

Barnes, P. F., El-Hajj, H., Preston-Martin, S., Cave, M. D., Jones, B. E., Otaya, M., . . . Eisenach, K. D. (1996). Transmission of tuberculosis among the urban homeless. *Jama, 275*(4), 305-307.

CDC. (2017). Reported Tuberculosis in the United States, 2016.

DPH. (2017). 2016 Georgia Tuberculosis Report.

Jasmer, R. M., Hahn, J. A., Small, P. M., Daley, C. L., Behr, M. A., Moss, A. R., . . . Hopewell, P. C. (1999). A molecular epidemiologic analysis of tuberculosis trends in San Francisco, 1991–1997. *Annals of internal medicine, 130*(12), 971-978.

Kempf, M.-C., Dunlap, N. E., Lok, K. H., Benjamin, W. H., Keenan, N. B., & Kimerling, M. E. (2005). Long-term molecular analysis of tuberculosis strains in Alabama, a state characterized by a largely indigenous, low-risk population. *Journal of clinical microbiology, 43*(2), 870-878.

Murray, M. B. (2002). Molecular epidemiology and the dynamics of tuberculosis transmission among foreign-born people. *Canadian Medical Association Journal, 167*(4), 355-356.

Powell, K. M., VanderEnde, D. S., Holland, D. P., Haddad, M. B., Yarn, B., Yamin, A. S., . . . Burns-Grant, G. (2017). Outbreak of Drug-Resistant Mycobacterium tuberculosis Among Homeless People in Atlanta, Georgia, 2008-2015. *Public Health Reports*, *132*(2), 231-240.

WHO. (2017). Global tuberculosis report 2017.

CHAPTER 2

LITERATURE REVIEW

Molecular Epidemiology of Tuberculosis

TB is caused by a group of closely related bacteria referred to as the Mycobacterium tuberculosis complex (MTBC) (Gagneux, 2012; Nelson & Williams, 2013; Smith, Hewinson, Kremer, Brosch, & Gordon, 2009). MTBC includes obligate human pathogens such as *Mycobacterium tuberculosis* and *Mycobacterium africanum,* as well as organisms adapted to various other species of mammals (Brites & Gagneux, 2015). One of the widely held beliefs was that clinical manifestation of TB was primarily driven by host variables and environmental factors as opposed to bacterial factors (Comas & Gagneux, 2009). Early DNA sequencing studies that reported very limited genetic diversity in MTBC compared with other bacterial pathogen reinforced this notion (Sreevatsan et al., 1997). It was also believed that TB was mainly a consequence of reactivation of latent infections rather than ongoing disease transmission. Mixed infections and exogenous reinfections were thought to be unlikely. The advent of molecular techniques to differentiate between strains of *Mycobacterium tuberculosis* has made it possible to re-address some of these points. Molecular markers for Mycobacterium tuberculosis have been developed since the early 1990s. These markers reveal different levels of genetic polymorphism and have different applications. The results from TB genotyping when combined with epidemiologic data can help identify persons with TB disease involved in the same chain of recent transmission (Barnes et al., 1996; Borgdorff et al., 1998; Small et al., 1994). Some of the applications of TB genotyping include:

1. Detect contamination events involving clinical specimens or laboratory cultures, or errors in specimen handling and labeling

2. Distinguish new TB disease from recurrence or relapse of previously diagnosed disease

3. Discover unsuspected transmission relationships between TB patients and confirm known transmission relationships

4. Determine completeness of contact investigations

5. Uncover inter-jurisdictional transmission

6. Identify outbreaks

In 1993, the CDC funded regional genotyping laboratories to support TB control programs for outbreak detection and epidemiologic studies. In 1996, CDC started the National Tuberculosis Genotyping Surveillance Network (NTGSN). In 2004, based on the knowledge gained from the network and related studies, CDC established the National TB Genotyping Service (NTGS) and funded two national genotyping laboratories, located in Michigan and California to genotype at least one M. tuberculosis isolate collected from each culture-positive TB case reported in the United States. Currently, most states genotype more than 90% of all TB cases annually. In 2010, the TB Genotyping information Management System (TB GIMS) was launched as a secure Web-based system to support ongoing use of TB genotyping data in TB control activities. TB GIMS facilitates systematic data collection of TB genotyping results and integrates genotyping results with epidemiologic data collected by the National TB Surveillance System (NTSS) to form a national and centralized database.

Genotyping Methods

Three genotyping methods are currently widely used in molecular epidemiologic studies of TB: IS6110 restriction fragment length polymorphism (RFLP) genotyping, spacer oligonucleotide typing (Spoligotyping), and mycobacterial interspersed repetitive-unit-variable-number-tandem-repeat (MIRU-VNTR) (Barnes & Cave, 2003). The most common method of DNA fingerprinting was RFLP. However, an important practical limitation of IS6110 RFLP is that results are usually obtained weeks to months after the initial diagnosis of TB (Barnes & Cave, 2003). This limitation stems from the need to grow large numbers of bacteria to extract DNA of sufficient quantity and quality for RFLP analysis (Mathema, Kurepina, Bifani, & Kreiswirth, 2006). This technique is therefore laborious, and time consuming (Allix-Béguec, Fauville-Dufaux, & Supply, 2008; Hanekom et al., 2008; Kristin Kremer et al., 2005). In contrast, a number of PCR-based typing modalities have been recently developed, including Spoligotyping and MIRU-VNTR, which offer the possibility of obtaining DNA fingerprints from small numbers of bacteria or even directly from clinical specimen (Barnes & Cave, 2003; Mathema et al., 2006; Roetzer et al., 2011; Varma-Basil et al., 2011). The NTGS laboratories primarily use two genotyping methods: Spoligotyping and MIRU-VNTR.

Spoligotyping

The genome of *Mycobacterium tuberculosis* contains 10 to 50 copies of a 36-bp direct repeat, which are separated from one another by spacers that have different sequences (Barnes & Cave, 2003; Brudey et al., 2006). However, the spacer sequences between any two specific direct repeats are conserved among strains. Because strains differ in terms of the presence or absence of specific spacers, the pattern of spacers in a strain can be used for genotyping using Spoligotyping (Barnes & Cave, 2003; Kamerbeek et al., 1997). To perform Spoligotyping, only small amounts of DNA are required and it can be performed on clinical samples or on strains of M. tuberculosis shortly after their inoculation into liquid culture (Kamerbeek et al., 1997). The results of Spoligotyping, which are expressed as positive or negative for the presence or absence of each spacer respectively, can be expressed in a digital format (Dale et al., 2001). However, Spoligotyping has less power to discriminate among M. tuberculosis strains than IS6110-based genotyping (K Kremer et al., 1999; Varma-Basil et al., 2011).

Mycobacterium Interspersed Repetitive Units Variable-Number-Tandem-Repeats (MIRU-VNTR)

The genome of *Mycobacterium tuberculosis* contains many mycobacterial interspersed repeat units (MIRU), some containing identical repeat units and others containing repeats that vary slightly in sequence and length (Frothingham & Meeker-O'Connell, 1998; Mazars et al., 2001). MIRU genotyping categorizes the number and size of the repeats in each of 12 or 24 independent MIRU’s with the use of polymerase chain reaction assay, followed by gel electrophoresis (Barnes & Cave, 2003). The discriminatory power of MIRU-VNTR is almost as great as that of IS6110 RFLP (Allix-Béguec et al., 2008; Mazars et al., 2001; Supply et al., 2001). Unlike IS6110-based genotyping, MIRU-VNTR analysis can be automated and can be used to evaluate large numbers of strains, yielding digital results that can be easily catalogued on a computer database. A website has been created so that a worldwide database of MIRU patterns can be made available to researchers (Supply et al., 2001). MIRU genotyping is technically simpler than IS6110 – based genotyping and can be applied directly to M. tuberculosis cultures without DNA purification (Barnes & Cave, 2003; Hanekom et al., 2008; Kristin Kremer et al., 2005).

Diversity of Human-associated MTBC strains

Mycobacterium tuberculosis complex has a global phylo-geographic population structure consisting of six main phylogenetic lineages (Comas, Homolka, Niemann, & Gagneux, 2009; Gagneux et al., 2006; Gutacker et al., 2002). These six lineages are associated with specific geographic regions and human populations (Baker, Brown, Maiden, & Drobniewski, 2004; Gagneux et al., 2006; Gagneux & Small, 2007; Hirsh, Tsolaki, DeRiemer, Feldman, & Small, 2004). Table 2.1 shows the lineages and their geographic areas.

Table 2.1: Mycobacterium tuberculosis lineages and their associated geographical regions in the world

Table 2.1: Mycobacterium tuberculosis lineages and their associated geographical regions in the world	
Mycobacterium tuberculosis Lineages	**Geographical region**
Indo-Oceanic Lineage (L1)	Indian Ocean
East Asian Lineage (L2)	East Asia
East African Indian Lineage (L3)	South East Asia
Euro-American Lineage (L4)	Europe, Americas, Africa
Africanum (L5 and L6)	Africa

Traditionally, the view is that Mycobacterium tuberculosis strains are equally virulent (Barnes & Cave, 2003). However, population-based genotyping has demonstrated that a small percentage of strains cause a disproportionate large number of cases, implying that some strains are spread more effectively than others (Caws et al., 2008; Click, Moonan, Winston, Cowan, & Oeltmann, 2012; Mireilla Coscolla & Gagneux, 2010; De Jong et al., 2008; Hanekom et al., 2011). For example, the Beijing strain, which belongs to lineage 2, has repeatedly been associated with drug resistance in a wide range of settings (Iwamoto, Yoshida, Suzuki, & Wada, 2008; Lasunskaia et al., 2010; Rajapaksa & Perera, 2011). Evidence suggests that the epidemiology of TB may be influenced by variations in the bacterial strain (Click et al., 2012; Mireilla Coscolla & Gagneux, 2010; De Jong et al., 2008). For example, the wide spatial distribution and dominance of the Beijing genotype has been associated with increased transmissibility relative to other Mycobacterium tuberculosis lineages in certain regions of the world (Mireia Coscolla & Gagneux, 2014; Hanekom et al., 2010; Kato-Maeda et al., 2012). The Beijing strains have caused outbreaks of TB throughout the world and constitute the dominant family of strains in multiple locations in Asia and North America.

The fact that phenotypic diversity exists between strains of TB is no longer in dispute; however, there is increased interest in establishing the relevance of this diversity to human disease. Recent studies have suggested that strains belonging to lineage 2, which include the W/Beijing phenotype, possess unique attributes that confer an increased ability to cause disease and to be transmitted within populations or geographic settings or certain patient ethnicities (Parwati, van Crevel, & van Soolingen, 2010). Studies have reported that TB transmission occurs more frequently in sympatric (sharing similar geographical area) host-pathogen combinations compared to allopatric (occurring in separate non-overlapping geographical areas) host-pathogen combinations (Gagneux, 2012; Gagneux et al., 2006). Similarly, experimental models have shown that Mycobacterium tuberculosis strains differ in virulence and immunogenicity (Mireilla

Coscolla & Gagneux, 2010), and epidemiologic studies have demonstrated that in addition to host and environmental factors, strain diversity contributes to the variable outcome of TB infection and disease (Nicol & Wilkinson, 2008). Studies using animal models have shown that many tuberculosis strains isolated from tuberculosis patients in India were less virulent and less infectious than strains from UK (Mireilla Coscolla & Gagneux, 2010; Gagneux & Small, 2007). Various studies have reported that lineage 2 (includes the Beijing genotype) is associated with drug resistance (Borrell & Gagneux, 2009; Fenner et al., 2012; Pang et al., 2012; Parwati et al., 2010). A few studies however could not find evidence of such association between lineage 2 and drug resistance (Iwamoto et al., 2008; Lasunskaia et al., 2010; Rajapaksa & Perera, 2011).

According to a study conducted in Nepal, gender was a significant predictor of MTBC lineage and drug resistance was more common with lineage 2 (Malla et al., 2012). This study found no significant relationship between age, previous history of TB or Bacille Calmette-Guerin (BCG) vaccine and MTBC lineages (Malla et al., 2012). This study was not a population based study and may have therefore missed some TB patients who were not diagnosed at the recruitment site. Studies in San Francisco, London and Montreal have shown that there exists a sympatric host-pathogen association that persists in urban settings (Hirsh et al., 2004; Reed et al., 2009). A study in Switzerland found that HIV infection was strongly associated with allopatric Mycobacterium tuberculosis lineages among European-born TB cases (Fenner et al., 2013). Among the allopatric lineages, the researchers found that lineages 1, 2 and 3 were more likely to be found in HIV-infected patients compared to HIV-negative ones (Fenner et al., 2013). It is still unclear to what extent different lineages of Mycobacterium tuberculosis have adapted to specific human populations, however, urban cities like Atlanta are said to have strains of Mycobacterium tuberculosis from diverse locations around the globe. Furthermore, due to the natural history of TB, where individuals may remain latently infected for decades before they develop active disease, the TB population in urban cities consists of strains contracted not only in a diversity of locations but also over an extended period of time. Hence, such centers are concentrations of geographically and temporally disparate samples of TB as well as potential crossroads in global transmission. It is important therefore to understand the relationship between the circulating lineages and the demographic characteristics of the host population.

Transmission Dynamics

The natural history of TB makes it difficult to understand the transmission dynamic of the disease. In most populations, TB involves a long latency period with symptoms of infection occurring anywhere from 3 months to decades after the establishment of infection (Comstock, 1975; Kaufmann, 2001). Since most infections with Mycobacterium tuberculosis are followed by a variable latency period, estimating the timing of transmission is difficult. This latency period is one of the main hallmarks of Mycobacterium tuberculosis infection and pathogenesis. For this reason, cases of TB disease may occur as a result of recent transmission from an infectious case or via reactivation of remotely acquired latent infection. The long and variable latency period and the high reproductive number of TB make it difficult to understand the transmission dynamics (Mathema et al., 2006).

TB control programs still rely on contact tracing and screening of contacts of TB patients. In contact tracing, individuals mentioned by the index case are screened for latent TB using purified protein derivative-based tuberculin skin test or using chest X rays and if TB is indicated, treatment is recommended (CDC, 2000). This method of understanding transmission dynamic of TB has been useful in low incidence countries (Frieden et al., 1996). However, this method is often imprecise and tends to underestimate the level of transmission in the population (Bifani et al., 1999; Bishai et al., 1998; Mathema et al., 2002; Small et al., 1994). Various molecular epidemiologic studies suggest that the imprecision of contact tracing may in part be due to complex transmission patterns in which casual contact may account for a considerable proportion of molecularly clustered yet epidemiologically unlinked cases (Mathema et al., 2002; Small et al., 1994; Valway et al., 1998; Yaganehdoost et al., 1999). This highlights one of the advantages of performing population-based molecular epidemiologic studies: identifying high-risk groups or areas where transmission is ongoing.

There is a broad variability in the genotypes of Mycobacterium tuberculosis isolates from patients with epidemiologically unrelated tuberculosis, whereas the genotypes of isolates from patients who were infected by a common source are identical (Barnes & Cave, 2003). Therefore, TB patients with identical strains of M. tuberculosis as disclosed by genotyping methods are considered clustered cases and cases with distinctive genotypes represent a reactivation of infection acquired in the distant past (Alland et al., 1994; Small et al., 1994; Yuen, Kammerer, Marks, Navin, & France, 2016). From a TB management point of view, characterizing Mycobacterium tuberculosis strains at the genomic level enables TB programs to track the transmission of specific M. tuberculosis strains, follow epidemics, and identify new outbreaks (Rodwell, Kapasi, Barnes, & Moser, 2012). Combining contact tracing from conventional surveillance with molecular techniques in population based studies have enabled public health researchers to determine which patient demographic and clinical factors might be associated with clusters of related strains and how interventions might be formulated to prevent those strains from being transmitted in the future (Yuen et al., 2016). Molecular epidemiologic studies have shown that transmission of tuberculosis varies by geographic region (Barnes & Cave, 2003). In cities were homelessness is common, shelters are often the foci of tuberculosis transmission (Barnes et al., 1997; Weis et al., 2002). In communities where TB is endemic, households are usually important sites of TB transmissions (Guwatudde et al., 2003; Whalen et al., 2011). In other locations, health care facilities and bars have been important sites of transmission (Alland et al., 1994; Small et al., 1994; Yaganehdoost et al., 1999). Therefore, local efforts to identify high-risk populations and transmission sites are crucial for the effective control of TB. The number of cases of TB diagnosed in U.S-born persons in Georgia, has declined by 67% since 2000, whereas the number of foreign-born cases has declined by only 3.3% (DPH, 2017).

Atlanta, the capital of the state of Georgia, is a considerably large city with specific urban attributes such as high homelessness, diverse cultures and considerable number of immigrants. Given an increase in the number of cases resulting from outbreaks and extensive transmission, an in-depth analysis into the transmission dynamic of TB using multiple years of data was warranted. A single infectious TB patient may have devastating effects on TB control and therefore ways to interrupt transmission should be directed towards specific populations at risk of acquiring and transmitting the disease. Recent transmission of TB as indicated by clustering is responsible for a significant number of TB cases in Georgia and therefore increased emphasis should be placed on the identification of sites of transmission and the application of environmental controls to include novel approach to contact tracing targeted to specific populations. This study will provide a more comprehensive analysis of the level of transmission by the use of genetic methods and the plausible-source case approach. This study will help identify sites of active transmission of TB in Georgia and influence control measures towards TB.

Persistence of Mycobacterium tuberculosis strains

TB presents with a variety of clinical features like duration of disease, severity and also manifests differently in certain populations (Hanekom et al., 2011). Despite ongoing research, the mechanisms that control the pathogenicity and factors influencing the degree of disease variability remain largely unknown. Even though several host and environmental factors have been associated with this variability, there is substantial evidence that bacterial factors also contribute to this variability in disease manifestation, frequency of transmission and treatment outcomes (Caminero et al., 2001; Caws et al., 2008; Valway et al., 1998). Certain strains of MTBC are more successful in establishing infection and disease in a large number of people within a short period of time. Available data demonstrates variation in evolutionary fitness and that some strains spread more or less quickly through populations (Cohen & Murray, 2004; Dye & Williams, 2009; Parwati et al., 2010). It is possible that these strains are constitutionally more transmissible than other strains due to the fact that they are more likely to produce sputum smear positive disease (Glynn et al., 2008). They could be associated with a more dangerous onset of clinical symptoms or they are more virulent and therefore produce more secondary cases within a short time period (Glynn et al., 2008; Valway et al., 1998).

For this study, the persistence or success of a strain takes into account the occurrence of the strain from year to year, the average number of cases per year and the presence of the strain in 2015. Given that TB occurs disproportionately in certain demographic groups like homeless and prison inmates, the persistence of a strain will also depend on the demographic population in which it first appeared. We will also expect strains that first appear in urban areas to spread to more people than those in rural areas. However, in rural communities, certain strains of TB may persist because of limited mobility in the population. Therefore, in rural communities, a single strain may be responsible for most of the TB diagnosed. Since transmission occurs more easily and frequently in crowded settings, we expect strains that emerge in these populations to spread more easily and therefore increase in size in a shorter period compared to strains that first emerge in the general population. A study conducted in a high burden HIV and TB community in South Africa found a wide diversity of *Mycobacterium tuberculosis* strains and that strain success was not associated with social factors (Middelkoop et al., 2014). HIV infection was said to be slightly associated with successful strains of *Mycobacterium tuberculosis* (Middelkoop et al., 2014). According to the authors, this association suggests a link between the success of individual strains and the immunological characteristics of the host. This study found that only 4% of the 311 distinct strains isolated were considered successful strains despite the high rate of transmission in this community (Middelkoop et al., 2014). This study was conducted in population with high burden of TB and HIV and most of the cases were too sick to provide sputum samples. Because some of the patients had more severe disease, they died before sputum samples could be collected.

Few population-based studies have examined the persistence of MTBC strains in communities. Some studies have examined the determinants of cluster sizes and have determined that the size of clusters could depend on factors favoring transmission or on differences in the strains themselves (Giordano et al., 2004; Glynn et al., 2008; Lillebaek, Dirksen, Kok-Jensen, & Andersen, 2004). The size of a cluster can be an indication of how transmissible that strain is but doesn't tell us how long the strain has been in circulation. Among the major challenges in achieving TB elimination in the United States are preventing, detecting and responding to TB outbreaks. Identifying high-risk settings and applying effective control measures to reduce TB transmission are basic principles of TB control. However, outbreaks in community settings have continued to occur, calling for increased vigilance in understanding and controlling TB transmission. There have been documented outbreaks of TB in Atlanta since 2008 and it has mostly been associated with residents of homeless shelters (Powell et al., 2017).

Understanding the factors associated with the success of strains will inform adjunct control strategies for TB. Drug-resistant TB also threatens global TB care and prevention, and remains a major public health concern in many countries. Drug–resistant TB has been reported since the early days of the introduction of chemotherapy (Espinal, 2003; Mathema et al., 2006). In vitro studies have shown that spontaneous mutations in Mycobacterium tuberculosis can be associated with drug resistance, while selective antibiotic pressure can lead to enhanced accumulation of these drug-resistant mutants (David, 1971; David & Newman, 1971). The efficient selection of drug resistance in the presence of a single antibiotic led investigators to recommend combination therapy using more than one antibiotic to reduce the emergence of drug resistance during treatment

(Councils, 1973). Studies have shown that TB control has been effective when adequate drug supplies are available and combination therapy is properly managed (Bass Jr et al., 1994; Kochi, Vareldzis, & Styblo, 1993).

Selection of drug-resistant mutants in patients mainly occurs when patients are treated inappropriately or are exposed to, even transiently, sub-therapeutic drug levels or conditions that may provide adequate positive selection pressure for the emergence and maintenance of drug-resistant organism (Mathema et al., 2006). One contributing factor is the exceptional length of chemotherapy required to treat and cure infection with Mycobacterium tuberculosis (Kaufmann, 2001). The need to maintain high drug levels over many months of treatment, combined with the inherent toxicity of the agents, results in reduced patient compliance and subsequently higher likelihood of acquisition of drug resistance (Davies, 1998). Therefore, in addition to identifying new anti-tuberculosis agents, the need for shortening the length of chemotherapy is paramount, as it would greatly impact clinical management and the emergence of drug resistance.

An alarming trend and growing source of public health concern has been the emergence of resistance to multiple drugs (MDR-TB). MDR-TB is defined as an isolate that is resistant to at least isoniazid (INH) and rifampin (RIF), the two most potent anti-tuberculosis drugs (Iseman & Madsen, 1989; Vareldzis et al., 1994). Drug-resistant TB is a persistent threat in the fight against the disease with 490,000 cases of MDR-TB emerging in 2016 and an additional 110,000 cases that were susceptible to isoniazid but resistant to rifampin (WHO, 2017). Whilst the treatment for MDR-TB has greatly improved in resource-rich settings, it is generally more difficult to treat and has been associated with very high morbidity and mortality, prolonged treatment to cure, and an increased risk of transmitting drug-resistant isolates in the community (Bifani et al., 1996; Crofton et al., 1997). The persistence of certain strains might be due to their ability to resist treatment with Isoniazid or Rifampin. If this is the case, these resistant strains will transmit despite the use of DOT and eventually become persistent in the community

References: Chapter 2

Alland, D., Kalkut, G. E., Moss, A. R., McAdam, R. A., Hahn, J. A., Bosworth, W., . . . Bloom, B. R. (1994). Transmission of tuberculosis in New York City--an analysis by DNA fingerprinting and conventional epidemiologic methods. *New England Journal of Medicine, 330*(24), 1710-1716.

Allix-Béguec, C., Fauville-Dufaux, M., & Supply, P. (2008). Three-year population-based evaluation of standardized mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing of Mycobacterium tuberculosis. *Journal of clinical microbiology, 46*(4), 1398-1406.

Baker, L., Brown, T., Maiden, M. C., & Drobniewski, F. (2004). Silent nucleotide polymorphisms and a phylogeny for Mycobacterium tuberculosis. *Emerging infectious diseases, 10*(9), 1568.

Barnes, P. F., & Cave, M. D. (2003). Molecular epidemiology of tuberculosis. *New England Journal of Medicine, 349*(12), 1149-1156.

Barnes, P. F., El-Hajj, H., Preston-Martin, S., Cave, M. D., Jones, B. E., Otaya, M., . . . Eisenach, K. D. (1996). Transmission of tuberculosis among the urban homeless. *Jama, 275*(4), 305-307.

Barnes, P. F., Yang, Z., Preston-Martin, S., Pogoda, J. M., Jones, B. E., Otaya, M., . . . Cave, M. D. (1997). Patterns of tuberculosis transmission in Central Los Angeles. *Jama, 278*(14), 1159-1163.

Bass Jr, J. B., Farer, L. S., Hopewell, P. C., O'Brien, R., Jacobs, R., Ruben, F., . . . Thornton, G. (1994). Treatment of tuberculosis and tuberculosis infection in adults and children. American Thoracic Society and The Centers for Disease Control and Prevention. *American journal of respiratory and critical care medicine, 149*(5), 1359-1374.

Bifani, P. J., Mathema, B., Liu, Z., Moghazeh, S. L., Shopsin, B., Tempalski, B., . . . Alcabes, P. (1999). Identification of a W variant outbreak of Mycobacterium tuberculosis via population-based molecular epidemiology. *Jama, 282*(24), 2321-2327.

Bifani, P. J., Plikaytis, B. B., Kapur, V., Stockbauer, K., Pan, X., Lutfey, M. L., . . . Kaplan, M. H. (1996). Origin and interstate spread of a New York City multidrug-resistant Mycobacterium tuberculosis clone family. *Jama, 275*(6), 452-457.

Bishai, W. R., Graham, N. M., Harrington, S., Pope, D. S., Hooper, N., Astemborski, J., . . . Chaisson, R. E. (1998). Molecular and geographic patterns of tuberculosis transmission after 15 years of directly observed therapy. *Jama, 280*(19), 1679-1684.

Borgdorff, M. W., Nagelkerke, N., van Soolingen, D., de Haas, P. E., Veen, J., & van Embden, J. D. (1998). Analysis of tuberculosis transmission between nationalities in the Netherlands in the period 1993–1995 using DNA fingerprinting. *American journal of epidemiology, 147*(2), 187-195.

Borrell, S., & Gagneux, S. (2009). Infectiousness, reproductive fitness and evolution of drug-resistant Mycobacterium tuberculosis [State of the art]. *The International Journal of Tuberculosis and Lung Disease, 13*(12), 1456-1466.

Brites, D., & Gagneux, S. (2015). Co-evolution of Mycobacterium tuberculosis and Homo sapiens. *Immunological reviews*, *264*(1), 6-24.

Brudey, K., Driscoll, J. R., Rigouts, L., Prodinger, W. M., Gori, A., Al-Hajoj, S. A., . . . Baumanis, V. (2006). Mycobacterium tuberculosis complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC microbiology, 6*(1), 23.

Caminero, J. A., Pena, M. J., Campos-Herrero, M. I., Rodriguez, J. C., Garcia, I., Cabrera, P., . . . Afonso, O. (2001). Epidemiological evidence of the spread of a Mycobacterium tuberculosis strain of the Beijing genotype on Gran Canaria Island. *American journal of respiratory and critical care medicine, 164*(7), 1165-1170.

Caws, M., Thwaites, G., Dunstan, S., Hawn, T. R., Lan, N. T. N., Thuong, N. T. T., . . . Loc, T. H. (2008). The influence of host and bacterial genotype on the development of disseminated disease with Mycobacterium tuberculosis. *PLoS pathogens, 4*(3), e1000034.

CDC. (2000). Targeted tuberculin testing and treatment of latent tuberculosis infection. *Am J Respir Crit Care Med, 161*, S221-S247.

Click, E. S., Moonan, P. K., Winston, C. A., Cowan, L. S., & Oeltmann, J. E. (2012). Relationship between Mycobacterium tuberculosis phylogenetic lineage and clinical site of tuberculosis. *Clinical infectious diseases, 54*(2), 211-219.

Cohen, T., & Murray, M. (2004). Modeling epidemics of multidrug-resistant M. tuberculosis of heterogeneous fitness. *Nature medicine, 10*(10), 1117.

Comas, I., & Gagneux, S. (2009). The past and future of tuberculosis research. *PLoS pathogens, 5*(10), e1000600.

Comas, I., Homolka, S., Niemann, S., & Gagneux, S. (2009). Genotyping of genetically monomorphic bacteria: DNA sequencing in Mycobacterium tuberculosis highlights the limitations of current methodologies. *PLoS One, 4*(11), e7815.

Comstock, G. W. (1975). FROST REVISITED: THE MODERN EPIDEMIOLOGY OF TUBERCULOSIS THE THIRD WADE HAMPTON FROST LECTURE. *American journal of epidemiology, 101*(5), 363-382.

Coscolla, M., & Gagneux, S. (2010). Does M. tuberculosis genomic diversity explain disease diversity? *Drug Discovery Today: Disease Mechanisms, 7*(1), e43-e59.

Coscolla, M., & Gagneux, S. (2014). *Consequences of genomic diversity in Mycobacterium tuberculosis.* Paper presented at the Seminars in immunology.

Councils, B. M. R. (1973). Controlled clinical trial of four short-course (6-month) regimens of chemotherapy for treatment of pulmonary tuberculosis: second report. *The Lancet, 301*(7816), 1331-1339.

Crofton, S. J., Chaulet, P., Maher, D., Grosset, J., Harris, W., Horne, N., . . . Watt, B. (1997). Guidelines for the management of drug-resistant tuberculosis.

Dale, J., Brittain, D., Cataldi, A. A., Cousins, D., Crawford, J., Driscoll, J., . . . Rastogi, N. (2001). Spacer oligonucleotide typing of bacteria of the Mycobacterium tuberculosis complex: recommendations for standardised nomenclature [The Language of Our Science]. *The International Journal of Tuberculosis and Lung Disease, 5*(3), 216-219.

David, H. L. (1971). Resistance to D-cycloserine in the tubercle bacilli: mutation rate and transport of alanine in parental cells and drug-resistant mutants. *Applied microbiology, 21*(5), 888-892.

David, H. L., & Newman, C. M. (1971). Some Observations on the Genetics of Isoniazid Resistance in the Tubercle Bacilli 1, 2. *American Review of Respiratory Disease, 104*(4), 508-515.

Davies, J. (1998). *Antibiotic resistance in mycobacteria.* Paper presented at the Genetics and Tuberculosis: Novartis Foundation Symposium 217.

De Jong, B. C., Hill, P. C., Aiken, A., Awine, T., Martin, A., Adetifa, I. M., . . . Gagneux, S. (2008). Progression to active tuberculosis, but not transmission, varies by Mycobacterium tuberculosis lineage in The Gambia. *Journal of Infectious Diseases, 198*(7), 1037-1043.

DPH. (2017). 2016 Georgia Tuberculosis Report.

Dye, C., & Williams, B. G. (2009). Slow elimination of multidrug-resistant tuberculosis. *Science translational medicine, 1*(3), 3ra8-3ra8.

Espinal, M. A. (2003). The global situation of MDR-TB. *Tuberculosis, 83*(1), 44-51.

Fenner, L., Egger, M., Bodmer, T., Altpeter, E., Zwahlen, M., Jaton, K., . . . Bruderer, T. (2012). Effect of mutation and genetic background on drug resistance in

Mycobacterium tuberculosis. *Antimicrobial agents and chemotherapy, 56*(6), 3047-3053.

Fenner, L., Egger, M., Bodmer, T., Furrer, H., Ballif, M., Battegay, M., . . . Rieder, H. L. (2013). HIV infection disrupts the sympatric host–pathogen relationship in human tuberculosis. *PLoS Genet, 9*(3), e1003318.

Frieden, T. R., Sherman, L. F., Maw, K. L., Fujiwara, P. I., Crawford, J. T., Nivin, B., . . . Alland, D. (1996). A multi-institutional outbreak of highly drug-resistant tuberculosis: epidemiology and clinical outcomes. *Jama, 276*(15), 1229-1235.

Frothingham, R., & Meeker-O'Connell, W. A. (1998). Genetic diversity in the Mycobacterium tuberculosis complex based on variable numbers of tandem DNA repeats. *Microbiology, 144*(5), 1189-1196.

Gagneux, S. (2012). Host–pathogen coevolution in human tuberculosis. *Phil. Trans. R. Soc. B, 367*(1590), 850-859.

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., De Jong, B. C., Narayanan, S., . . . Gutierrez, M. C. (2006). Variable host–pathogen compatibility in Mycobacterium tuberculosis. *Proceedings of the National academy of Sciences of the United States of America, 103*(8), 2869-2873.

Gagneux, S., & Small, P. M. (2007). Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development. *The Lancet infectious diseases, 7*(5), 328-337.

Giordano, T. P., Soini, H., Teeter, L. D., Adams, G. J., Musser, J. M., & Graviss, E. A. (2004). Relating the size of molecularly defined clusters of tuberculosis to the duration of symptoms. *Clinical infectious diseases, 38*(1), 10-16.

Glynn, J. R., Crampin, A. C., Traore, H., Chaguluka, S., Mwafulirwa, D. T., Alghamdi, S., . . . Fine, P. E. (2008). Determinants of cluster size in large, population-based molecular epidemiology study of tuberculosis, northern Malawi. *Emerging infectious diseases, 14*(7), 1060.

Gutacker, M. M., Smoot, J. C., Migliaccio, C. A. L., Ricklefs, S. M., Hua, S., Cousins, D. V., . . . Musser, J. M. (2002). Genome-wide analysis of synonymous single nucleotide polymorphisms in Mycobacterium tuberculosis complex organisms: resolution of genetic relationships among closely related microbial strains. *Genetics, 162*(4), 1533-1543.

Guwatudde, D., Nakakeeto, M., Jones-Lopez, E., Maganda, A., Chiunda, A., Mugerwa, R., . . . Whalen, C. (2003). Tuberculosis in household contacts of infectious cases in Kampala, Uganda. *American journal of epidemiology, 158*(9), 887-898.

Hanekom, M., Mata, D., van Pittius, N. G., van Helden, P., Warren, R., & Hernandez-Pando, R. (2010). Mycobacterium tuberculosis strains with the Beijing genotype demonstrate variability in virulence associated with transmission. *Tuberculosis, 90*(5), 319-325.

Hanekom, M., Van Der Spuy, G., van Pittius, N. G., McEvoy, C., Hoek, K., Ndabambi, S., . . . Warren, R. (2008). Discordance between mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing and IS6110 restriction fragment length polymorphism genotyping for analysis of Mycobacterium tuberculosis Beijing strains in a setting of high incidence of tuberculosis. *Journal of clinical microbiology, 46*(10), 3338-3345.

Hanekom, M., Van Pittius, N. G., McEvoy, C., Victor, T., Van Helden, P., & Warren, R. (2011). Mycobacterium tuberculosis Beijing genotype: a template for success. *Tuberculosis, 91*(6), 510-523.

Hirsh, A. E., Tsolaki, A. G., DeRiemer, K., Feldman, M. W., & Small, P. M. (2004). Stable association between strains of Mycobacterium tuberculosis and their human host populations. *Proceedings of the National academy of Sciences of the United States of America, 101*(14), 4871-4876.

Iseman, M., & Madsen, L. (1989). Drug-resistant tuberculosis. *Clinics in chest medicine, 10*(3), 341-353.

Iwamoto, T., Yoshida, S., Suzuki, K., & Wada, T. (2008). Population structure analysis of the Mycobacterium tuberculosis Beijing family indicates an association between certain sublineages and multidrug resistance. *Antimicrobial agents and chemotherapy, 52*(10), 3805-3809.

Kamerbeek, J., Schouls, L., Kolk, A., Van Agterveld, M., Van Soolingen, D., Kuijper, S., . . . Goyal, M. (1997). Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. *Journal of clinical microbiology, 35*(4), 907-914.

Kato-Maeda, M., Shanley, C. A., Ackart, D., Jarlsberg, L. G., Shang, S., Obregon-Henao, A., . . . Barrozo, J. C. (2012). Beijing sublineages of Mycobacterium tuberculosis differ in pathogenicity in the guinea pig. *Clinical and vaccine immunology, 19*(8), 1227-1237.

Kaufmann, S. H. (2001). How can immunology contribute to the control of tuberculosis? *Nature Reviews Immunology, 1*(1), 20-30.

Kochi, A., Vareldzis, B., & Styblo, K. (1993). Multidrug-resistant tuberculosis and its control. *Research in microbiology, 144*(2), 104-110.

Kremer, K., Au, B. K. Y., Yip, P. C. W., Skuce, R., Supply, P., Kam, K. M., & van Soolingen, D. (2005). Use of variable-number tandem-repeat typing to differentiate Mycobacterium tuberculosis Beijing family isolates from Hong Kong and comparison with IS6110 restriction fragment length polymorphism typing and spoligotyping. *Journal of clinical microbiology, 43*(1), 314-320.

Kremer, K., Van Soolingen, D., Frothingham, R., Haas, W., Hermans, P., Martin, C., . . . Yakrus, M. (1999). Comparison of methods based on different molecular epidemiological markers for typing of Mycobacterium tuberculosis complex strains: interlaboratory study of discriminatory power and reproducibility. *Journal of clinical microbiology, 37*(8), 2607-2618.

Lasunskaia, E., Ribeiro, S. C., Manicheva, O., Gomes, L. L., Suffys, P. N., Mokrousov, I., . . . Otten, T. (2010). Emerging multidrug resistant Mycobacterium tuberculosis strains of the Beijing genotype circulating in Russia express a pattern of biological properties associated with enhanced virulence. *Microbes and Infection, 12*(6), 467-475.

Lillebaek, T., Dirksen, A., Kok-Jensen, A., & Andersen, Å. (2004). A dominant Mycobacterium tuberculosis strain emerging in Denmark. *The International Journal of Tuberculosis and Lung Disease, 8*(8), 1001-1006.

Malla, B., Stucki, D., Borrell, S., Feldmann, J., Maharjan, B., Shrestha, B., . . . Gagneux, S. (2012). First insights into the phylogenetic diversity of Mycobacterium tuberculosis in Nepal. *PLoS One, 7*(12), e52297.

Mathema, B., Bifani, P. J., Driscoll, J., Steinlein, L., Kurepina, N., Moghazeh, S. L., . . . Mangura, B. (2002). Identification and evolution of an IS 6110 low-copy-number Mycobacterium tuberculosis cluster. *The Journal of infectious diseases, 185*(5), 641-649.

Mathema, B., Kurepina, N. E., Bifani, P. J., & Kreiswirth, B. N. (2006). Molecular epidemiology of tuberculosis: current insights. *Clinical microbiology reviews, 19*(4), 658-685.

Mazars, E., Lesjean, S., Banuls, A.-L., Gilbert, M., Vincent, V., Gicquel, B., . . . Supply, P. (2001). High-resolution minisatellite-based typing as a portable approach to global analysis of Mycobacterium tuberculosis molecular epidemiology. *Proceedings of the national academy of Sciences, 98*(4), 1901-1906.

Middelkoop, K., Bekker, L.-G., Mathema, B., Myer, L., Shashkina, E., Whitelaw, A., . . . Wood, R. (2014). Factors affecting tuberculosis strain success over 10 years in a high TB-and HIV-burdened community. *International journal of epidemiology, 43*(4), 1114-1122.

Nelson, K. E., & Williams, C. (2013). *Infectious disease epidemiology*: Jones & Bartlett Publishers.

Nicol, M. P., & Wilkinson, R. J. (2008). The clinical consequences of strain diversity in Mycobacterium tuberculosis. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 102*(10), 955-965.

Pang, Y., Zhou, Y., Zhao, B., Liu, G., Jiang, G., Xia, H., . . . Zhao, Y.-l. (2012). Spoligotyping and drug resistance analysis of Mycobacterium tuberculosis strains from national survey in China. *PLoS One, 7*(3), e32976.

Parwati, I., van Crevel, R., & van Soolingen, D. (2010). Possible underlying mechanisms for successful emergence of the Mycobacterium tuberculosis Beijing genotype strains. *The Lancet infectious diseases, 10*(2), 103-111.

Powell, K. M., VanderEnde, D. S., Holland, D. P., Haddad, M. B., Yarn, B., Yamin, A. S., . . . Burns-Grant, G. (2017). Outbreak of Drug-Resistant Mycobacterium tuberculosis Among Homeless People in Atlanta, Georgia, 2008-2015. *Public Health Reports, 132*(2), 231-240.

Rajapaksa, U., & Perera, A. (2011). Sublineages of Beijing strain of Mycobacterium tuberculosis in Sri Lanka. *Indian journal of microbiology, 51*(3), 410-412.

Reed, M. B., Pichler, V. K., McIntosh, F., Mattia, A., Fallow, A., Masala, S., . . . Menzies, D. (2009). Major Mycobacterium tuberculosis lineages associate with patient country of origin. *Journal of clinical microbiology, 47*(4), 1119-1128.

Rodwell, T. C., Kapasi, A. J., Barnes, R. F., & Moser, K. S. (2012). Factors associated with genotype clustering of Mycobacterium tuberculosis isolates in an ethnically diverse region of southern California, United States. *Infection, Genetics and Evolution, 12*(8), 1917-1925.

Roetzer, A., Schuback, S., Diel, R., Gasau, F., Ubben, T., di Nauta, A., . . . Niemann, S. (2011). Evaluation of Mycobacterium tuberculosis typing methods in a 4-year study in Schleswig-Holstein, Northern Germany. *Journal of clinical microbiology, 49*(12), 4173-4178.

Small, P. M., Hopewell, P. C., Singh, S. P., Paz, A., Parsonnet, J., Ruston, D. C., . . . Schoolnik, G. K. (1994). The Epidemiology of Tuberculosis in San Francisco--A Population-Based Study Using Conventional and Molecular Methods. *New England Journal of Medicine, 330*(24), 1703-1709.

Smith, N. H., Hewinson, R. G., Kremer, K., Brosch, R., & Gordon, S. V. (2009). Myths and misconceptions: the origin and evolution of Mycobacterium tuberculosis. *Nature Reviews Microbiology, 7*(7), 537.

Sreevatsan, S., Pan, X., Stockbauer, K. E., Connell, N. D., Kreiswirth, B. N., Whittam, T. S., & Musser, J. M. (1997). Restricted structural gene polymorphism in the Mycobacterium tuberculosis complex indicates evolutionarily recent global dissemination. *Proceedings of the national academy of Sciences, 94*(18), 9869-9874.

SUN, J. R., DOU, H. Y., LEE, S. Y., CHIUEH, T. S., & LU, J. J. (2011). Epidemiological studies of Beijing strains of Mycobacterium tuberculosis from Taipei and other Asian cities based on MIRU profiles. *Apmis, 119*(9), 581-587.

Supply, P., Lesjean, S., Savine, E., Kremer, K., Van Soolingen, D., & Locht, C. (2001). Automated high-throughput genotyping for study of global epidemiology of Mycobacterium tuberculosis based on mycobacterial interspersed repetitive units. *Journal of clinical microbiology, 39*(10), 3563-3571.

Valway, S. E., Sanchez, M. P. C., Shinnick, T. F., Orme, I., Agerton, T., Hoy, D., . . . Onorato, I. M. (1998). An outbreak involving extensive transmission of a virulent strain of Mycobacterium tuberculosis. *New England Journal of Medicine, 338*(10), 633-639.

Vareldzis, B. P., Grosset, J., De Kantor, I., Crofton, J., Laszlo, A., Felten, M., . . . Kochi, A. (1994). Drug-resistant tuberculosis: laboratory issues: World Health Organization recommendations. *Tubercle and Lung Disease, 75*(1), 1-7.

Varma-Basil, M., Kumar, S., Arora, J., Angrup, A., Zozio, T., Banavaliker, J. N., . . . Bose, M. (2011). Comparison of spoligotyping, mycobacterial interspersed repetitive units typing and IS6110-RFLP in a study of genotypic diversity of Mycobacterium tuberculosis in Delhi, North India. *Memorias do Instituto Oswaldo Cruz, 106*(5), 524-535.

Weis, S. E., Pogoda, J. M., Yang, Z., Cave, M. D., Wallace, C., Kelley, M., & Barnes, P. F. (2002). Transmission dynamics of tuberculosis in Tarrant County, Texas. *American journal of respiratory and critical care medicine, 166*(1), 36-42.

Whalen, C. C., Zalwango, S., Chiunda, A., Malone, L., Eisenach, K., Joloba, M., . . . Mugerwa, R. (2011). Secondary attack rate of tuberculosis in urban households in Kampala, Uganda. *PLoS One, 6*(2), e16137.

WHO. (2017). Global tuberculosis report 2017.

Yaganehdoost, A., Graviss, E. A., Ross, M. W., Adams, G. J., Ramaswamy, S., Wanger, A., . . . Musser, J. M. (1999). Complex transmission dynamics of clonally related virulent Mycobacterium tuberculosis associated with barhopping by predominantly human immunodeficiency virus-positive gay men. *The Journal of infectious diseases, 180*(4), 1245-1251.

Yuen, C. M., Kammerer, J. S., Marks, K., Navin, T. R., & France, A. M. (2016). Recent transmission of tuberculosis—United States, 2011–2014. *PLoS One, 11*(4), e0153728.

CHAPTER 3

METHODS

Study Settings and Study Population

The state of Georgia is the largest state east of the Mississippi River with 159 counties and 18 Health Districts with jurisdiction for TB control. The population of Georgia in 2016 was estimated at 10.3 million people and almost 10% of the population is non-US-born persons (Census, 2016). In 2016 Georgia ranked 6th in the United States for number of new TB cases and ranked 10th for TB case rate among the 50 reporting states (CDC, 2017). The study population consisted of all TB cases diagnosed in Georgia with at least one genotyped isolate from 2009 to 2015. Only cases with molecular data linked to surveillance data were included in our study. Patient demographic and risk factor information were used together with molecular data. The UGA IRB and Georgia DPH IRB approved this study.

Study Design

This was a secondary analysis of TB surveillance and molecular data in Georgia from 2009 to 2015. Enrolling the entire population of patients with TB disease from 2009 to 2015 overcame potential bias introduced by an analysis of only a sample of the population. Given that a long study period is more likely to capture transmission linked cases than a shorter one, our study provided a better opportunity for understanding the transmission of TB in Georgia than similar studies conducted in San Francisco, New York, Denmark and the Netherlands. All culture confirmed cases diagnosed between

January 1, 2009 and December 31, 2015 with genotyping data linked to surveillance data were selected for inclusion in the study. The methods used to examine the genetic contents of Mycobacterium tuberculosis isolates were Spoligotyping and a panel of 12 or 24 mycobacterial interspersed repetitive unit variable number of tandem repeats (MIRU-VNTR). These methods yield digital results that are readily analyzed and each isolate is assigned a genotype.

TB GIMS is a system that is managed by CDC and SendSS is the platform through which notifiable diseases and health conditions are reported to the State of Georgia's Department of Public Health. A state case number matched each case of TB in SendSS with its corresponding isolate in TB GIMS. Each isolate in TB GIMS is allocated a 'PCRtype' and 'GENType' designation. In TB GIMS, a genotype is defined as a unique combination of Spoligotyping and MIRU-VNTR analysis. Each unique combination of Spoligotyping and 12-loci MIRU-VNTR results is assigned a PCRtype. PCRtype is designated as ''PCR'' followed by five digits, which are assigned sequentially to every genotype identified in the United States (e.g., PCR01974). In April 2009, MIRU-VNTR analysis was expanded from 12 loci to 24 loci. The addition of these 12 loci increased the ability to distinguish between strains of M. tuberculosis and correctly identify chains of transmission. Each unique combination of spoligotyping and 24-locus MIRU-VNTR is assigned a GENType, formatted as "G" followed by five digits (e.g., G00056). In this study, clusters were defined as two or more cases of TB with identical PCRtypes or GENtypes.

This study relied on the underlying assumption that TB cases connected by recent transmission should have identical genotypes and those that are reactivation of LTBI should have unique genotypes (Barnes & Cave, 2003). Clustering of TB cases was used as a proxy for recent transmission and cases that were not clustered were assumed to be a result of reactivation of latent infection acquired in the past. To validate our estimation of recent transmission, we used a method called the plausible-source case approach (France, Grant, Kammerer, & Navin, 2015). This approach provides a standard method that will allow us to compare cases across time and geographic areas. Each case diagnosed in Georgia from 2009 to 2015 was compared to all other cases diagnosed in Georgia with available genotyping data to determine whether a plausible source case can be identified. If a plausible source case is identified for a certain case, the given case will be classified as belonging to a cluster or attributable to recent transmission provided both cases reside in the same county. Comparing cases in the longer time window enabled us to evaluate plausible-source cases over the same time unit for each case. A plausible source case was defined as a case that involves pulmonary TB in a patient over 4 years of age and was diagnosed within 2 years prior to the given case (France, Grant, Kammerer, & Navin, 2015). The plausible-source case must have resided within the same county as the given case and must have had the same genotyping pattern based on MIRU/VNTR and Spoligotyping.

The number of clusters was evaluated at the state level and among sub groups within the population. Risk factors associated with clustering of TB isolates were then examined. In TB GIMS, isolates are assigned a 'GENType' based on the MIRU/VNTR and Spoligotyping patterns and lineages are assigned to each isolate based on the GENType. The association between the lineages of MTBC and patient clinical and demographic risk factors were then examined. To examine the persistence of strains, all the isolates from 2009 to 2015 were grouped into persistently successful, transiently successful and unsuccessful strains. Persistently successful strains were strains that have been present for at least 3 consecutive years and were still present in 2015 and had greater than or equal to 3 cases per year (Middelkoop et al., 2014). Transiently successful strains were strains that did not persist to 2015 but occurred for at least 3 consecutive years, with a cluster size of at least 3 cases per year in the consecutive period (Middelkoop et al., 2014). Unsuccessful strains did not fall into the above categories and also includes clustered strains occurring for the first time. For each isolate, we recorded the year of first emergence and examined the epidemiology of that strain following emergence. Patients characteristics associated with the different groupings were then examined.

Analytic Strategy by Aim

Specific Aim # 1: To examine the association between Mycobacterium tuberculosis lineages and TB patient characteristics.

Research Question: Is there a statistically significant relationship between Mycobacterium tuberculosis lineages and any of the patient clinical or demographic variables?

HO: There is no association between MTBC lineages and patient demographic and clinical variables.

HA: There is an association between MTBC lineages and at least one patient demographic or clinical variable.

Figure 3.1 Multinomial Logistic Regression Model components for investigating the association between Mycobacterium tuberculosis lineages and TB patient characteristics		
Primary Outcome of Interest	Y_1= Mycobacterium tuberculosis lineages	1= Indo-Oceanic Lineage (L1) 2= East-Asian Lineage (L2) 3= East-African-Indian Lineage (L3) 4= Euro-American Lineage (L4)
Demographic Variables	Age	Age in Years
	Sex	0=Female 1=Male
	Race	1= American Indian/ Alaskan Native 2= Asian 3= Black 4= White
	Ethnicity	0=Hispanic 1=non-Hispanic
	Homelessness	0= No 1= Yes
	Prison	0=No 1=Yes
	Excess Alcohol Use	0=No 1=Yes
	Injection Drug Use	0=No 1=Yes
	Non-Injection Drug Use	0=No 1=Yes
	Country of Origin	0=No 1=Yes
	Long-term Care Facility	0=No 1=Yes
Clinical Variables	Sputum Smear	0=Negative 1=Positive
	Isoniazid Resistance	0= No 1= Yes
	Cavitary Disease	0=No 1=Yes
	Chest X-ray	0=Normal 1=Abnormal
	Case Status	0=Alive 1=Dead
	Previous TB	0=No 1=Yes

We used Chi-square test of independence to examine associations between MTBC lineages and TB patient characteristics. Kruskal Wallis rank test was used to determine if the median age was equal amongst all lineages. The null hypothesis was that the relative proportions of each of the demographic and clinical variables are independent of the MTBC lineages. Variables with significant relationships with Mycobacterium tuberculosis lineages were then included in a multinomial logistic model. A baseline-category model was used to describe the data with Euro-American Lineage as the baseline category. Logit equations were used describe the log-odds that an isolate is in a certain lineage instead of the Euro-American Lineage (L2).

With the Euro-American Lineage as the baseline or reference lineage, log-odds for all other lineages relative to the baseline lineage were calculated. The results from this multinomial logit will be presented as conditional odds ratios. The model fit will be evaluated using likelihood ratio chi-square to determine the presence of a relationship between the dependent variable and any of the independent variables. Associations were considered statistically significant at $\alpha \leq 0.05$ and all statistical analyses were performed using S.A.S. 9.4 (SAS Institute, Cary, NC, USA).

Specific Aim # 2: To examine the risk factors of clustering of isolates from TB patients in Georgia.

Research Question: Is there a statistically significant relationship between clustering of TB isolates and any of the predictor variables?

HO: There is no association between clustering of TB isolates and patient demographic and clinical variables.

HA: There is an association between clustering of TB isolates and at least one patient demographic and clinical variable.

Figure 3.2: Log Binomial Regression Model components for investigating the association between Clustering of TB patient Isolates and TB patient characteristics		
Primary Outcome of Interest	Y_1= Clustering of Isolates	0= Clustered 1= Unique
Demographic Variables	Age	Age in Years
	Sex	0=Female 1=Male
	Race	1= American Indian/ Alaskan Native 2= Asian 3= Black 4= White
	Ethnicity	0=Hispanic 1=non-Hispanic
	Homelessness	0= No 1= Yes
	Prison	0=No 1=Yes
	Excess Alcohol Use	0=No 1=Yes
	Injection Drug Use	0=No 1=Yes
	Non-Injection Drug Use	0=No 1=Yes
	Country of Origin	0=No 1=Yes
	Long-term Care Facility	0=No 1=Yes
Clinical Variables	Sputum Smear	0=Negative 1=Positive
	Isoniazid Resistance	0= No 1= Yes
	Cavitary Disease	0=No 1=Yes
	Chest X-ray	0=Normal 1=Abnormal
	Case Status	0=Alive 1=Dead
	Previous TB	0=No 1=Yes

The sample was divided into 2 groups: clustered isolates and non-clustered isolates as indicated by DNA fingerprinting and plausible source case approach. Descriptive statistics, including cross-tabulations of demographic and clinical characteristics were computed. The Chi-square test or the Fisher's exact test was used in a bivariate analysis to assess the relationship between recent transmission of TB and patient characteristics. Patient characteristics significant in the bivariate analysis were included in a multivariate logistic model. Statistical interaction between certain demographic and clinical variables was also investigated based on the data and prior knowledge. Manual backwards elimination was used for each model, eliminating the variables with the highest p-value and examining the Akaike Information Criterion (AIC) at each step. The model with the lowest AIC was retained. Associations were considered statistically significant at $\alpha \leq 0.05$.

Specific Aim # 3: To examine the persistence of MTBC strains in Georgia from 2009 to 2015

Research Question: Is there a statistically significant relationship between the most common strains of MTBC in Georgia and certain patient characteristics?

HO: The most persistent strains of TB in GA do not have any significant statistical relationship with certain patient demographic and clinical characteristics.

HA: The most persistent strains of TB in GA have a significant statistical relationship with certain patient characteristics that enable them to persist and spread in the community.

Figure 3.3 Multinomial Logistic Regression Model components for investigating the association between Persistence of TB Isolates and TB patient characteristics		
Primary Outcome of Interest	Y_1= Persistence of Strains	1= Unsuccessful Strains 2= Transiently Successful Strains 3= Persistently Successful Strains
Demographic Variables	Age	Age in Years
	Sex	0=Female 1=Male
	Race	1= American Indian/ Alaskan Native 2= Asian 3= Black 4= White
	Ethnicity	0=Hispanic 1=non-Hispanic
	Homelessness	0= No 1= Yes
	Prison	0=No 1=Yes
	Excess Alcohol Use	0=No 1=Yes
	Injection Drug Use	0=No 1=Yes
	Non-Injection Drug Use	0=No 1=Yes
	Country of Origin	0=No 1=Yes
	Long-term Care Facility	0=No 1=Yes
Clinical Variables	Sputum Smear	0=Negative 1=Positive
	Isoniazid Resistance	0= No 1= Yes
	Cavitary Disease	0=No 1=Yes
	Chest X-ray	0=Normal 1=Abnormal
	Case Status	0=Alive 1=Dead
	Previous TB	0=No 1=Yes

Strains clustered by genetic testing alone were further divided into persistently successful strains, transiently successful strains and unsuccessful strains as previously described. Chi-Square test for independence was used to examine if there exists a statistically significant relationship between these level of persistence of TB strains and certain patient characteristics. Variables that were significant in the bivariate analysis were included in a multinomial logistic regression model to examine factors associated with persistently successful, transiently successful and unsuccessful strains. With persistently successful strains as the baseline or reference category, log-odds for all alternative categories relative to the baseline were calculated. The results from this multinomial logit were presented as conditional odds ratios with 95% confidence intervals. Associations were considered statistically significant at $\alpha \leq 0.05$.

Data Integrity

The state of Georgia's Department of Public Health has the authority to collect data on reported TB cases to include demographic, clinical, risk factor, and contact information. TB is reported using case definitions provided by the CDC. The data collected by Georgia is included in the CDC annual TB report, which speaks to its accuracy and consistency. Due to confidentiality issues, the state does not release HIV data and city of residence for TB cases along with some of our predictor variables. The state maintains TB data in a secure database with restricted access only to authorized individuals.

Biomedical Ethics

This study was a secondary data analysis of TB data collected in the state of Georgia from 2009 to 2015. However, both the University of Georgia and the Georgia Department of Public Health required extensive IRB review. Both reviews concluded that this was not a study involving human subjects. Information collected as part of this project did not contain any identifiable patient information and therefore the potential risk of unintentional disclosure does not exist. All study investigators were trained in research ethics and the responsible conduct of research through certified training experiences (e.g., CITI). For any ethical study the benefits will outweigh the risks. The ratio of risks and benefits may be viewed from an individual or societal perspective. The benefits of this study will apply at the community level rather than individual. From a societal perspective, the main benefit may be an improved understanding of the dynamics of TB transmission within the state and how molecular epidemiology can improve our understanding of TB control. It is hoped that the information presented in this study will lead to novel and pragmatic solutions to TB control.

References: Chapter 3

Barnes, P. F., & Cave, M. D. (2003). Molecular epidemiology of tuberculosis. *New England Journal of Medicine, 349*(12), 1149-1156.

CDC. (2017). Reported Tuberculosis in the United States, 2016.

Census, B. (2016, 03/20/2017). State Quickfacts: Georgia. Retrieved from https://www.census.gov/quickfacts/table/PST045216/13

France, A. M., Grant, J., Kammerer, J. S., & Navin, T. R. (2015). A field-validated approach using surveillance and genotyping data to estimate tuberculosis attributable to recent transmission in the United States. *American journal of epidemiology*, *182*(9), 799-807.

Middelkoop, K., Bekker, L.-G., Mathema, B., Myer, L., Shashkina, E., Whitelaw, A., . . . Wood, R. (2014). Factors affecting tuberculosis strain success over 10 years in a high TB-and HIV-burdened community. *International journal of epidemiology, 43*(4), 1114-1122.

CHAPTER 4

EXAMINE THE ASSOCIATION BETWEEN MYCOBACTERIUM TUBERCULOSIS LINEAGES AND TB PATIENT CHARACTERISTICS IN GEORGIA, 2009-2015[1]

[1] Edward A. Sheriff, Kevin Dobbin, Andrea Swartzendruber, Christopher C. Whalen, Juliet N. Sekandi. To be submitted to *International Journal of Tuberculosis and Lung Disease*

Abstract

Setting: The State of Georgia and its counties, United States

Objectives: To examine the genetic diversity of Mycobacterium tuberculosis strain lineages and their associations with patient characteristics.

Methods: A secondary analysis of surveillance and molecular data collected by the Georgia department of Health was analyzed. Case surveillance data were linked to molecular data using a unique state case numbers. We used multinomial logistic regression model was used to compare patient characteristics associated with Mycobacterium tuberculosis lineages.

Results: of the 1,411 isolates, 1083 (77%) belonged to lineage 4. Lineage 4 was most common among US-born cases. All other lineages were common among non-US-born cases. We observed that race and country of origin are significant predictors of Mycobacterium tuberculosis lineages.

Conclusions: There is an apparent phylo-geographic relationship between strain lineages and host country of origin. A patients country of origin can be a good predictor of the lineage of TB they have.

Introduction

Tuberculosis (TB) is an infectious disease characterized by high morbidity and mortality in developing countries and in urban areas of developed countries. Infection with *Mycobacterium tuberculosis* is preventable and treatable. However, an estimated 1.7 million deaths occur globally each year (WHO, 2017). In the United States, 9,272 cases of TB were reported in 2016. This represents a 2.9% decline in the number of cases from 2015 (CDC, 2017). Georgia reported 301 TB cases with a case rate of 2.9 cases per 100,000 persons in 2016 (DPH, 2017). Georgia ranks 6th in the United States for number of new TB cases and 10th for TB case rate among the 50 reporting states in 2016 (DPH, 2017). Four counties in the metropolitan Atlanta area accounted for more than 50% of TB cases reported in Georgia in 2016. These counties are Fulton, DeKalb, Gwinnett and Cobb in descending order of cases reported (DPH, 2017). Among Georgia's 18 Health Districts, which has oversight responsibility for public health in the state's 159 counties, DeKalb Health District had the highest TB case rate in 2016 at 7.8 cases per 100,000, followed by Columbus Health District at 4.5 cases per 100,000 (DPH, 2017).

Research has identified six lineages of *Mycobacterium tuberculosis*, with most lineages having a strong association with specific human populations around the world (Gagneux et al., 2006). The six main phylogenetic lineages are: Lineage 1 (also known as Indo-Oceanic lineage), Lineage 2 (also known as the East Asian Lineage), Lineage 3 (also known as East African Asian Lineage), Lineage 4 (Euro-American Lineage), and Lineage 5 and 6 (West African Lineage 1 and 2). A growing body of evidence shows that, the epidemiology of TB may be influenced by the strain of *Mycobacterium tuberculosis* (Caws et al., 2008; Coscolla & Gagneux, 2010; De Jong et al., 2008).

Differences in virulence between *Mycobacterium tuberculosis* strains were reported as early as the 1960s, with the observation that *Mycobacterium tuberculosis* isolated in India were attenuated in the guinea pig model when compared with those isolated in the United Kingdom (Mitchison et al., 1960). The Beijing genotype for example has repeatedly been associated with drug resistance in a wide range of settings (Borrell & Gagneux, 2009; Fenner et al., 2012; Malla et al., 2012; Pang et al., 2012; Parwati, van Crevel, & van Soolingen, 2010). Strains of the Beijing/W genotype have caused large outbreaks of tuberculosis, sometimes involving multidrug resistance (Agerton et al., 1999; Bifani et al., 1996; CDC, 1993; Van Soolingen et al., 1995). A few studies however, did not find evidence of such an association (Iwamoto, Yoshida, Suzuki, & Wada, 2008; Lasunskaia et al., 2010; Rajapaksa & Perera, 2011).

TB research entered the genomic era in 1998 with the complete annotated genome of *Mycobacterium tuberculosis* strain H37Rv (Cole et al., 1998). Since then, a lot of other strains of *Mycobacteria* have been sequenced. Currently there are several methods for genotyping *Mycobacterium tuberculosis*. Spacer oligonucleotide typing (Spoligotyping) is a rapid polymerase chain reaction (PCR)-based method for genotyping strains of *Mycobacterium tuberculosis*. Spoligotyping data can be represented in absolute terms (digitally) and the results can be readily shared among laboratories thereby enabling the creation of large international databases (Driscoll, 2009). The discriminatory power of Spoligotyping is inferior compared to restriction fragment length polymorphism of insertion sequence 6110 (RFLP IS6110) (K Kremer et al., 1999). In 2001, a rapid and reproducible variable number of tandem repeat (VNTR) based typing using 12 mycobacterial interspersed repetitive units (MIRU) was developed. However, this method lacked the discriminatory power of IS6110-RFLP. More recently, a set of 24 MIRU-VNTR loci is reported to have greater discriminatory power than the original 12-locus system and may exceed that of IS6110-RFLP when combined with Spoligotyping (Christianson et al., 2010). MIRU-VNTR typing uses variations in repetitive sequences, which are not under selection pressure and evolve relatively rapidly making them suitable for prospective molecular epidemiologic investigations and surveillance purposes (Crawford, 2003; Jagielski et al., 2014). The discriminatory power of MIRU-VNTR is typically proportionally to the number of loci evaluated (Mathema et al., 2006). When more than 12 loci are used, or MIRU-VNTR analysis is combined with Spoligotyping, the discriminatory power approximates that of IS6110 RFLP analysis (Mathema et al., 2006). A comparative study of genotyping methods aimed at evaluating novel PCR-based typing techniques found MIRU-VNTR analysis to have the greatest discriminatory power among amplification-based approaches (Kristin Kremer et al., 2005). MIRU-VNTR genotyping has been used in a number of molecular epidemiologic studies, as well as to elucidate the phylogenetic relationships of clinical isolates (Kristin Kremer, Au, et al., 2005; Sola et al., 2003; Y. -J. Sun et al., 2004; Supply et al., 2003).

There are no studies on the phylogenetic diversity of Mycobacterium tuberculosis in Georgia. According to reports, majority of the TB cases in Georgia were born in the United States, this is different from the scenario at the national level where majority of the cases were born outside of the United States. We sought out to understand which lineages of TB are being transmitted in Georgia and among which groups of people. The aim of this study was to describe the *Mycobacterium tuberculosis* lineages in Georgia and associations with patient demographic and clinical characteristics using surveillance and molecular data collected between January 1st 2009 and December 31st 2015.

Methods

The population of Georgia in 2015 was estimated at 10,310,371 with 9.8% (1,010,416) being foreign-born persons (Census, 2016). We analyzed cases of culture confirmed pulmonary TB diagnosed in Georgia between January 1, 2009 and December 31, 2015. We included only cases with molecular data linked to surveillance report for the specified time period. We analyzed socio-demographic, clinical, and molecular data for each case of TB. Data for this project was collected by the Georgia Department of Public Health as part of an ongoing surveillance for TB. The data for this analysis were housed in two different databases. In one database we had case surveillance data with demographic and clinical variables and in the other we had molecular data/genotyping data. A unique state case number linked cases and isolates in both databases. Molecular analysis of all isolates were previously done using Spoligotyping and MIRU-VNTR technology. Isolates with matching Spoligotyping and MIRU-VNTR patterns were considered clustered and more likely to be part of a chain of relatively recent transmission. For each case of TB with molecular data, Mycobacterium tuberculosis lineage was also provided. We excluded cases with no state case numbers or cases that could not be linked to their corresponding molecular data.

Statistical Analysis

Descriptive statistics were used to examine the distribution of the outcome and predictor variables. We used Chi-square to test the differences between groups in categorical variables, and the Kruskal Wallis rank test for continuous variables. The null hypothesis was that the relative proportions of each of the demographic and clinical variables are independent of the *Mycobacterium tuberculosis* lineages. Multiple analysis on the same dependent variable increases the likelihood of coming about a significant result by pure chance. To correct for this, a Bonferroni correction was used. Multinomial logistic regression models were used to compare patient characteristics associated with the various lineages. Odds ratios and corresponding 95% confidence intervals were reported when comparing lineages. Associations were considered statistically significant at $\alpha \leq 0.05$. All statistical analyses were performed using S.A.S. 9.4 (SAS Institute, Cary, NC, USA).

Results

From 2009 to 2015 there were 2,521 cases of TB reported in Georgia. Molecular data was available for 76% (1,918) of these cases. Our final sample comprised of 1,411 cases of TB.

Bivariate Analysis

Genotyping results showed the presence of four *Mycobacterium tuberculosis* lineages in Georgia. The most frequent of the lineages were lineage 4 otherwise referred to as the "Euro-American" lineage with 1,083 isolates (77%) and lineage 2 or East Asian lineage with 139 isolates (13%). Table 4.1 shows baseline characteristics of TB patients and their distribution across the four lineages in Georgia from 2009 to 2015. Clustering of isolates was more common among isolates from lineage 4 followed by isolates from lineage 2. 84% of the clustered isolates were from lineage 4 compared to 11% from lineage 2, 3% from lineage 1, and 2% from lineage 3, χ^2 (3, N = 1411) = 78.5334, p = 0.0016.

We observed a significant relationship between gender and *Mycobacterium tuberculosis* lineages, χ^2 (3, N = 1411) = 9.5805, p = 0.0225. There were predominantly more male cases across all four lineages. The median age for all cases in our study was 46 years (interquartile range [IQR] 32-57). Females were significantly younger than males (median age 39 versus 47 years, $p < .0016$). Non-US-born cases were significantly younger than US-born cases (median age 38 versus 51 years, $p < .0016$). There was no significant difference in the median age among all 4 lineages ($p < .0.4818$). Among Black cases, 92% of the isolates were in lineage 4. Isolates from Asian cases were predominantly in Lineage 1 and lineage 2. The proportion of non-Hispanics was higher across all lineages, χ^2 (3, N = 1411) = 35.2513, $p < .0016$.

There was a significant relationship between Mycobacterium tuberculosis lineages and patient's country of origin, χ^2 (3, N = 1411) = 226.6758, $p < .0016$. 91% of the isolates from US-born cases were from lineage 4 and 59% of the isolates from non-US-born cases were from lineage 4. Compared to other lineages, majority of the cases resident in correctional facilities had isolates from lineage 4, χ^2 (3, N = 1411) = 19.9099, p = 0.0016. 13.4% of the isolates were from cases that had been homeless in the 12 months preceding diagnosis of TB. 95% of the homeless cases had isolates from lineage 4 χ^2 (3, N = 1411) = 41.4592, $p < .0016$. The prevalence of excess alcohol consumption in our sample was 17.9% and 88% of the cases that consumed excess alcohol had isolates that were assigned to lineage 4 χ^2 (6, N = 1411) = 25.3354, $p < .0016$.

Geographical Trends

Figure 4.1 shows the distribution of cases diagnosed in Georgia, from 2009 to 2015 across six WHO regions. Non-US-born cases represented 67 different countries from all WHO regions except Eastern Mediterranean region. 75% of the isolates in our study came from the WHO Americas region, 15% from South-East Asia region and 7% from the Africa region. When the sample was stratified by patient's country of birth there was a regional trend with respect to *Mycobacterium tuberculosis* lineages. This trend can be summarized as follows: Lineage 1 or the Indo-Oceanic lineage was predominantly isolated in cases from the South East Asia region. Lineage 2, which includes the Beijing strain was often isolated in the Americas, South East Asia and Western Pacific region. Lineage 3 was most commonly isolated from cases from South East Asia. 88% of the isolates from the Americas belonged to lineage 4 and 78% of the isolates from the African and European region belonged to lineage 4. In our study, majority of the isolates assigned to the lineage 1 were linked to cases from India, Philippines and Vietnam. The predominant lineage in our study- Lineage 4 was identified in cases from all regions and it was also identified in patients from 56 of the 68 countries represented. The regions that were predominantly Euro-American include Europe, the Americas (including the Caribbean), and the sub-regions in Africa. Among non-US-born cases, Mexico, Vietnam and India accounted for majority of the isolates in our study.

Association between M. tuberculosis lineages and patient characteristics

Table 4.3 and table 4.4 show unadjusted and adjusted multinomial logistics regression odds ratios respectively and corresponding 95% confidence intervals. Since lineage 4 was the most common lineage in our sample, we compared patient characteristics in lineage 4 to other lineages. Multinomial logistic regression showed that Asian cases were more likely to be associated with lineage 1 compared to Black cases, [aOR: 31.5; 95% CI: 13.8-71.9], after controlling for country of origin, ethnicity, non-injection drug use, excess alcohol consumption , homelessness, prison residence and clustering of isolates. After controlling for all other independent variables, Asians were more likely than Blacks to belong to lineage 2 than lineage 4, [aOR: 61.9; 95% CI: 25.6-149.3]. Asians were 5.4 times more likely than Blacks to belong to lineage 3 rather than lineage 4, [aOR: 5.4; 95% CI: 2.7-10.7] after controlling for all other independent variables.

Cases born outside of the United States were 4.0 times more likely to belong to lineage 1 than cases born in the United States, when comparing lineage 1 to lineage 4, [aOR: 4.0; 95% CI: 1.7-9.3]. Non-US-born cases were 13.8 times more likely than US-born cases to belong to lineage 3 rather than lineage 4, [aOR: 13.8; 95% CI: 4.9-38.8] after controlling for all other independent variables. After controlling for all other independent variables, clustered isolates were 2.2 times more likely to belong to lineage 2 than unique isolates when comparing lineage 2 to lineage 4, [aOR: 2.2; 95% CI: 1.4-3.6]. The final model includes race, country of birth and clustering of isolates.

Discussion

Transmission and incidence rates of TB in Georgia are low compared to other areas. It is difficult to examine the relationship between *Mycobacterium tuberculosis* lineages and TB patient characteristics when more than 70% of the cases belong to only one lineage. A remarkable degree of concordance exists between our study and the San Francisco study (Gagneux et al., 2006) with respect to the phylogeographic associations of the major mycobacterium tuberculosis lineages. This finding is even more striking when we consider that the patient distribution between the two populations are quite different.

Majority of the TB cases in Georgia are US-born as opposed to San Francisco. Gagneux et al. proposed that major Mycobacterium tuberculosis lineages have evolved so as to become adapted to specific host genetic backgrounds and are much more likely to transmit and cause disease among patients of the same ethnicity (Gagneux et al., 2006). Since most TB cases in Georgia are in persons born in the United States, we expected to see most of the isolates belonging to lineage 4. In the San Francisco study, majority of the isolates in lineages 1, 2, and 3 were from cases born outside of the United States. For both the Georgia and San Francisco patient cohorts, the majority of isolates were determined to be part of the Euro-American lineage that comprises 76% and 48% of Georgia and San Francisco isolates, respectively. Since TB in San Francisco is more common among non-US-born persons, we see that the combined proportion of Lineages 1, 2, and 3 is greater than that of lineage 4. However, in Georgia, were TB is more common among US-born persons, we see that the combined proportion of the imported lineages is less than that of lineage 4. In both studies, lineage 4 was isolated from TB patients originating from each of the WHO regions represented in this study. In this respect, the Euro-American lineage is unique, and its peculiar distribution pattern may be a function of multiple migration routes involving the individuals infected with these strains (Stucki et al., 2016). For both cohorts, lineage 4 is clearly predominant throughout the Americas, Europe, and Africa. US-born cases together with patients from Mexico, Guatemala, Ethiopia, Honduras, Peru and Kenya comprise more than 80% of all cases involving lineage 4 in our study.

We expected to see a higher proportion of cases associated with lineage 2 among Asians, due to the fact that the East Asian lineage is more widespread and possibly well suited to spread among Asian persons. Instead, among US-born cases, we saw a high proportion of isolates from lineage 2 among Black cases. Lineage 2 is the only allopatric lineage that is responsible for a significant number of TB cases among US-born persons. Lineage 1 and lineage 3 do not pose a significant disease burden among US-born cases.

Multivariate analysis showed that clustered isolates were more likely to belong to lineage 2 than lineage 4. This finding is even more important if we consider the fact that isolates from lineage 2 only make up 13% of isolates in our study. More than 50% of the isolates in lineage 2 were clustered. Among non-US-born cases, clustered isolates from lineage 2 were mostly Asians and among US-born cases, clustered isolates from lineage 2 were mostly Black. Since lineage 2 is most commonly isolated in cases from East Asia, it's surprising that a majority of the clustered cases among US-born cases in Lineage 2 were not Asians. This is a disruption of the sympatric host pathogen relationship that exists between TB lineages and their hosts. Studies have suggested that HIV/AIDS can possibly alter this sympatric relationship. Cases with HIV/AIDS have a compromised immune system and can therefore be infected by any lineage regardless of their country of origin. In the absence of HIV/AIDS data we couldn't test this hypothesis. Also looking at the clustered isolates among non-US-born cases in lineage 2, we found that most of the clustered isolates came from cases with different countries of origin. Making it unlikely that these cases belong to a chain of recent transmission. Interpreting clustering of isolates among non-US-born cases should be done with caution or supported by contact tracing methods.

Our study found a similar geographical relationship between Mycobacterium tuberculosis lineages and patients country of origin. However, we found no evidence of lineage effect on disease manifestation. The state of Georgia should continue to monitor trends in other lineages besides lineage 4, as a significant increase in these lineages may indicate an increase in the number of imported TB cases or an increase in transmission occurring between non-US-born and US-born cases.

One limitation of this study is that molecular data was not available for all cases due to the gradual uptake of molecular testing by state TB control programs. 76% of the cases diagnosed with TB in Georgia from 2009 to 2015 had molecular testing done. Another limitation is the fact that more than 70% of the isolates in our study belonged to a single lineage making it difficult to accurately examine associations. We were unable to explore the role of HIV/AIDS infection in the association of Mycobacterium tuberculosis lineages and TB patient characteristics. Some studies have shown that US-born cases infected with allopatric strains (Lineage 1, Lineage 2 and lineage 3) were more likely to be HIV-positive. Despite the above limitations, the study provides the first description of the diversity of Mycobacterium tuberculosis lineages using data spanning 7 years. This is the first study addressing the question of host-pathogen adaptation of Mycobacterium tuberculosis in the state of Georgia. This study provides the basis for future analysis of Mycobacterium tuberculosis lineages using more robust molecular techniques like whole genome sequencing. Future studies should examine the role of HIV/AIDS in lineage diversity and disease manifestation.

Table 4.1:Baseline characteristics of TB cases by lineages in Georgia, 2009-2015 (N=1,411)				
Characteristics	**L1[1] (n=126) n (%)**	**L2[2] (n=139) n (%)**	**L3[3] (n=63) n (%)**	**L4[4] (n=1083) n (%)**
Age (Median, IQR[1])	44 (32-61)	46 (32-61)	42 (27-56)	46 (32-57)
Cluster				
No	104 (14.7)	64 (9.1)	47 (6.7)	492 (69.6)
Yes	22 (3.1)	75 (10.7)	16 (2.3)	591 (83.9)
Sex				
Female	43 (10.1)	43(10.1)	29 (6.8)	311 (73)
Male	83 (8.4)	96 (9.8)	34 (3.5)	772 (78.4)
Race				
Asian	81 (32.8)	80 (32.4)	39 (15.8)	47 (19.1)
AI/AN[3]	2 (8.7)	0 (0.0)	2 (8.7)	19 (82.6)
Black	10 (1.4)	30 (4.2)	19 (2.7)	644 (91.6)
White	31 (7.2)	28 (6.5)	3 (0.7)	370 (85.7)
Unknown	2 (33.3)	1 (16.7)	0 (0.0)	3 (50.0)
US-Born				
No	112 (18.3)	84 (13.8)	57 (9.3)	358 (58.6)
Yes	14 (1.8)	55 (6.9)	6 (0.8)	725 (90.6)
Case Status				
Alive	125 (9.1)	135 (9.8)	59 (4.3)	1058 (76.8)
Dead	1 (2.9)	4 (11.8)	4 (11.8)	25 (73.5)
Sputum Smear				
Negative	53 (10.7)	49 (9.9)	26 (5.3)	367 (74.1)
Positive	71 (8.2)	86 (9.9)	32 (3.4)	679 (78.2)
Not Done	2 (4.2)	4 (8.3)	5 (10.4)	37 (77.1)
LTCF[6]				
No	126 (9.1)	138 (9.9)	62 (4.5)	1067 (76.6)
Yes	0 (0.0)	1 (5.9)	1(5.9)	15 (88.2)
Prison				
No	124 (9.4)	138 (10.4)	63 (4.8)	1001 (75.5)
Yes	2 (1.6)	1(0.7)	0 (0.0)	82 (7.5)
Previous TB				
No	119 (9.0)	128 (9.6)	60 (4.5)	1023 (76.9)
Yes	6 (7.5)	11 (13.8)	3 (3.8)	60 (75.0)
Drug Use (non-inj.)				
No	125 (10.0)	125 (10.0)	63 (5.1)	935 (74.9)
Yes	1 (0.6)	14 (8.9)	0 (0.0)	142 (90.5)
Unknown	0 (0.0)	0 (0.0)	0 (0.0)	6 (100.0)
Drug Use (inj.)				

No	126 (9.1)	139 (10.0)	63 (4.5)	1065 (76.5)
Yes	0 (0.0)	0 (0.0)	0 (0.0)	12 (1.10)
Unknown	0 (0.0)	0 (0.0)	0 (0.0)	6 (100.0)
Chest CT				
Abnormal	67 (8.8)	68 (8.8)	28 (3.7)	603 (78.8)
Normal	1 (5.0)	2 (10.0)	2 (10.0)	15 (75.0)
Not Done	24 (11.4)	27 (12.9)	14 (6.7)	145 (69.1)
INH[5] Resistance				
Resistant	14 (7.3)	15 (7.7)	5 (2.6)	159 (82.4)
Sensitive	112 (9.3)	123(10.2)	58 (4.8)	916 (75.8)
Not Done	0 (0.0)	1 (11.1)	0 (0.0)	8 (88.9)
Cavitary Disease				
No	80 (9.2)	86 (9.9)	44 (5.1)	660 (75.9)
Yes	37 (8.5)	45 (10.3)	16 (3.7)	340 (77.6)
Unknown	1 (11.1)	0 (0.0)	1 (11.1)	7 (77.8)
Homeless				
No	124 (10.2)	131 (10.7)	63 (5.2)	902 (73.9)
Yes	2 (1.1)	8 (4.2)	0 (0.0)	180 (94.7)
Unknown	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)
Excess Alcohol				
No	117 (10.2)	121 (10.5)	60 (5.2)	853 (74.1)
Yes	9 (3.6)	18 (7.1)	3 (1.2)	223 (88.1)
Unknown	0 (0.0)	0 (0.0)	0 (0.0)	6 (100.0)
Ethnicity				
Non-Hispanic	103 (8.9)	132 (11.4)	63 (5.5)	857 (74.2)
Hispanic	23 (9.0)	7 (2.8)	0 (0.0)	225 (88.2)
Unknown	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)

Figure 4.1: Distribution of Mycobacterium tuberculosis lineages in Georgia by regions, 2009-2015

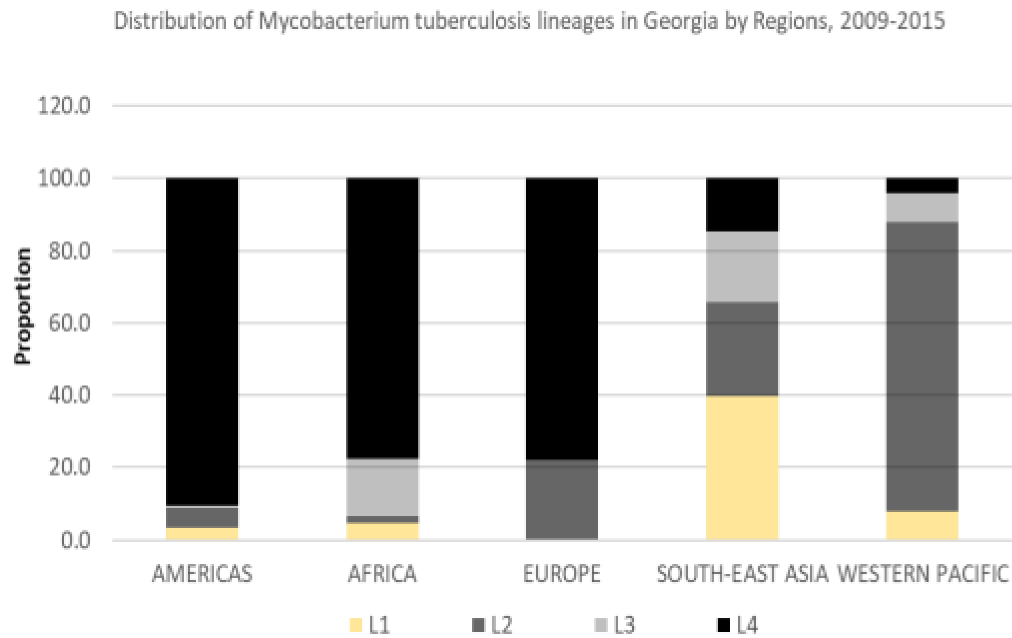


Table 4.2. Unadjusted Odds Ratio and 95% Confidence Intervals from Multinomial Logistic Regression comparing Euro-American Lineage to all other lineages

Patient Characteristics	L1[1] vs L4[2]	L2[3] vs L4	L3[4] vs L4
Ethnicity (non-Hispanic vs. Hispanic)	0.8 (0.5-1.3)	0.2 (0.1-0.4)*	
Race (AI/AN vs Black)	3.4 (0.4-27.9)		1.8 (0.2-14.1)
Race (Asian vs Black)	113.1 (55.1-231.8)*	36.3 (21.8-60.6)*	28.3 (15.2-52.7)*
Race (White vs Black)	5.5 (2.7-11.4) *	1.6 (0.9-2.7)	0.3 (0.1-0.9)
Country of Birth (Non-US vs. US)	16.4 (9.3-28.9)	3.1 (2.2-4.5) *	19.3 (8.2-45.1)*
Drug Use (non-injection) (No vs. Yes)	19.1 (2.6-137.4)*	1.3 (0.8-2.4)	
Excess Alcohol Use (No vs. Yes)	3.4 (1.7-6.8)*	1.7 (1.0-2.9)	5.2 (1.6-16.7)*
Homelessness (No vs. Yes)	12.4 (3.0-50.7)*	3.3 (1.6-6.8)*	
Prison (No vs. Yes)	5.1 (1.2-21.0)*	11.3 (1.6-81.5)*	
Clustering (Unique vs. Cluster)	0.2 (0.1-0.3)	1.0 (0.7-1.4)	0.2 (0.2-0.5)

[1] Indo-Oceanic Lineage, [2] Euro-American Lineage, [3] East Asian Lineage, [4] East African Indian Lineage, *Significant odds ratio

Table 4.3. Adjusted Odds Ratio and 95% Confidence Intervals from Multinomial Logistic Regression comparing Euro-American Lineage to all other lineages

Patient Characteristics	**L1[1] vs L4[2]**	**L2[3] vs L4**	**L3[4] vs L4**
Ethnicity (non-Hispanic vs. Hispanic)	0.6 (0.2-1.4)	0.4 (0.1-1.2)*	
Race (AI/AN vs Black)	1.6 (0.2-14.9)		23.2 (0.8-710.4)
Race (Asian vs Black)	31.5 (13.8-71.9)*	61.9 (25.6-149.3)*	5.4 (2.7-10.7)*
Race (White vs Black)	4.4 (1.8-10.8)*	2.9 (1.6-5.3)*	0.9 (0.2-3.2)
Country of Birth (Non-US vs. US)	4.0 (1.7-9.3)*	0.8 (0.3-1.8)	13.8 (4.9-38.8)*
Drug Use (non-injection) (N0 vs. Yes)	3.1 (0.4-24.9)	0.6 (0.3-1.2)	
Excess Alcohol Use (No vs. Yes)	0.8 (0.4-1.9)	0.8 (0.5-1.6)	1.0 (0.3-3.7)
Homelessness (No vs. Yes)	3.0 (0.7-13.6)	2.3 (1.0-5.4)	
Prison (No vs. Yes)	3.0 (0.7-13.0)	4.6 (0.6-34.3)	
Clustering (Clustered vs. Unique)	0.6 (0.4-1.1)	2.2 (1.4-3.6)*	1.0 (0.5-2.0)

*[1] Indo-Oceanic Lineage, [2] Euro-American Lineage, [3] East Asian Lineage, [4] East African Indian Lineage, *Significant odds ratio*

Appendix

Figure 4.2: Distribution of Mycobacterium tuberculosis lineages in Georgia by patient's country of origin, 2009-2015, (n=1103) (all isolates)

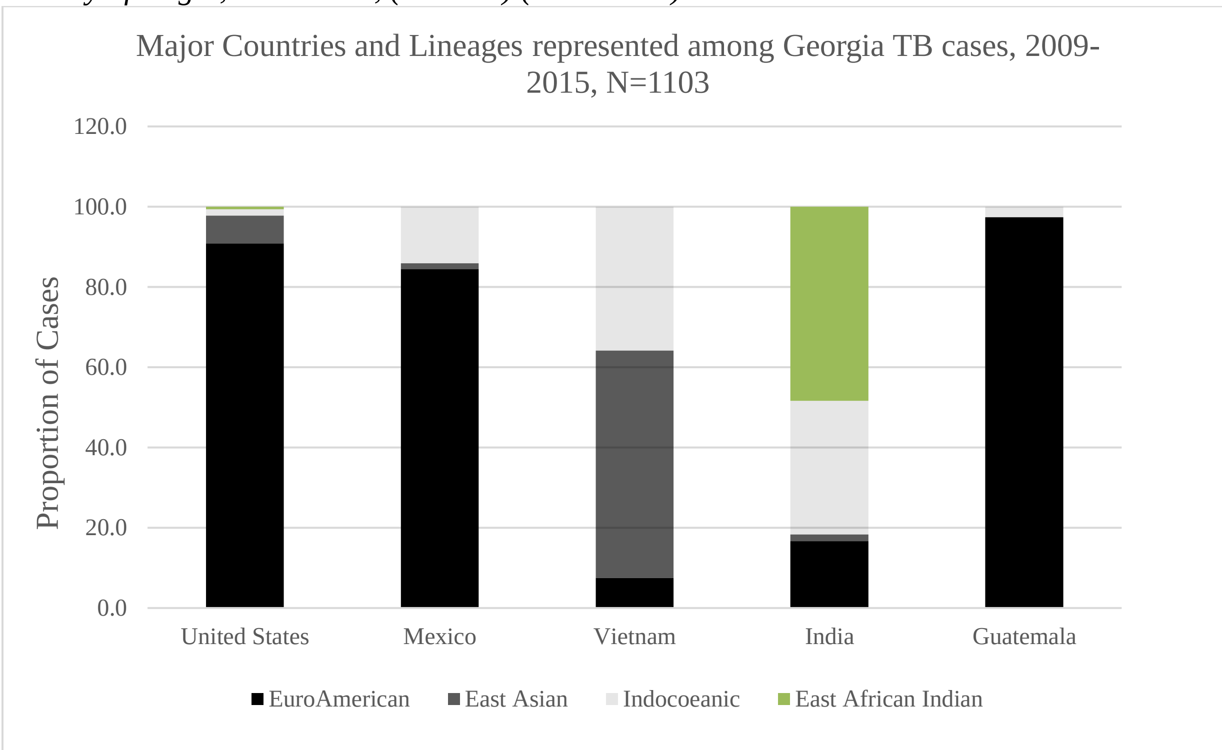


Figure 4.3: Distribution of clustered Mycobacterium tuberculosis lineages by patient's country of origin, 2009-2015, (n=1103)

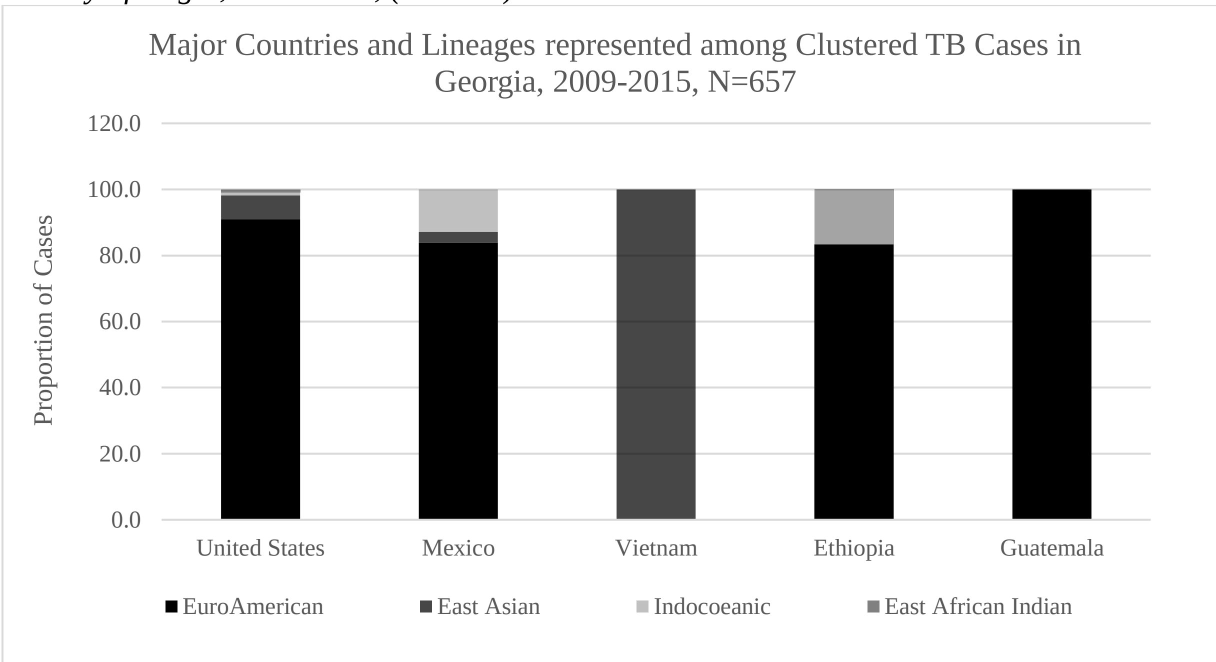
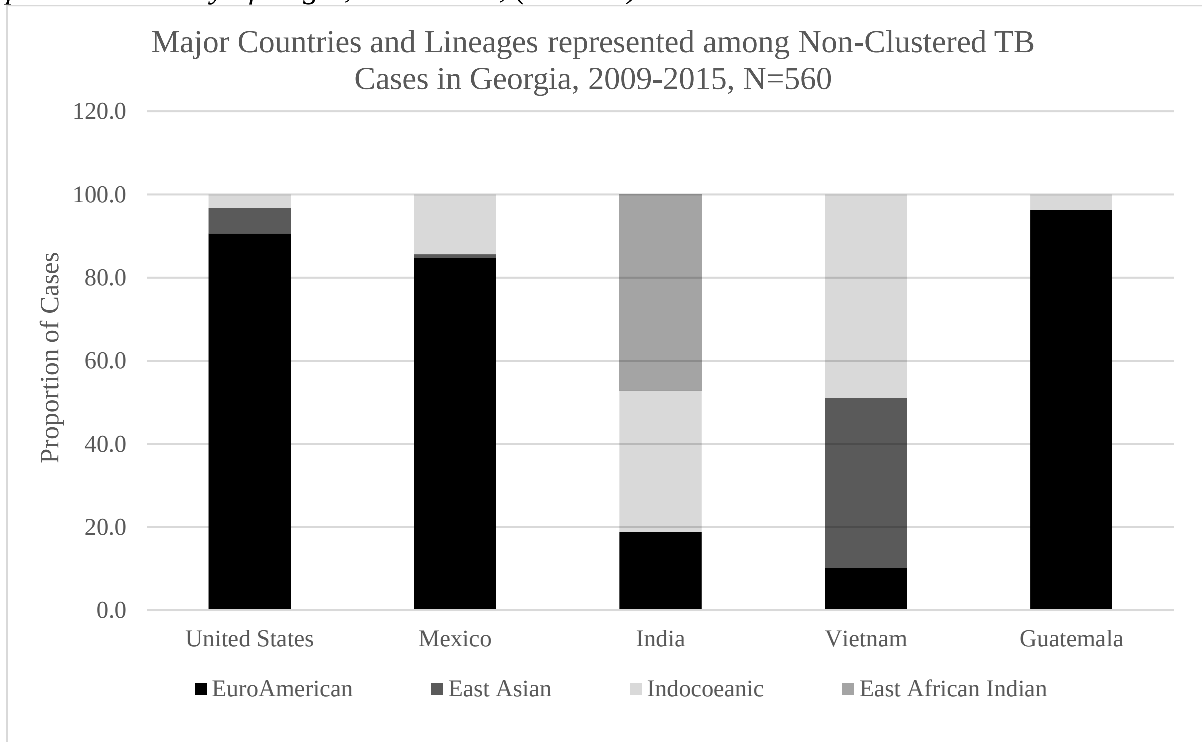


Figure 4.4: Distribution of unique Mycobacterium tuberculosis lineages in Georgia by patients' country of origin, 2009-2015, (n=1103)



References: Chapter 4

Agerton, T. B., Valway, S. E., Blinkhorn, R. J., Shilkret, K. L., Reves, R., Schluter, W. W., . . . Woodley, C. (1999). Spread of strain W, a highly drug-resistant strain of Mycobacterium tuberculosis, across the United States. *Clinical infectious diseases, 29*(1), 85-92.

Bifani, P. J., Plikaytis, B. B., Kapur, V., Stockbauer, K., Pan, X., Lutfey, M. L., . . . Kaplan, M. H. (1996). Origin and interstate spread of a New York City multidrug-resistant Mycobacterium tuberculosis clone family. *Jama, 275*(6), 452-457.

Borrell, S., & Gagneux, S. (2009). Infectiousness, reproductive fitness and evolution of drug-resistant Mycobacterium tuberculosis [State of the art]. *The International Journal of Tuberculosis and Lung Disease, 13*(12), 1456-1466.

Caws, M., Thwaites, G., Dunstan, S., Hawn, T. R., Lan, N. T. N., Thuong, N. T. T., . . . Loc, T. H. (2008). The influence of host and bacterial genotype on the development of disseminated disease with Mycobacterium tuberculosis. *PLoS pathogens, 4*(3), e1000034.

CDC. (1993). Outbreak of multidrug-resistant tuberculosis at a hospital--New York City, 1991. *MMWR. Morbidity and mortality weekly report, 42*(22), 427, 433.

CDC. (2017). Reported Tuberculosis in the United States, 2016.

Census, B. (2016, 03/20/2017). State Quickfacts: Georgia. Retrieved from https://www.census.gov/quickfacts/table/PST045216/13

Christianson, S., Wolfe, J., Orr, P., Karlowsky, J., Levett, P. N., Horsman, G. B., . . . Sharma, M. K. (2010). Evaluation of 24 locus MIRU-VNTR genotyping of Mycobacterium tuberculosis isolates in Canada. *Tuberculosis, 90*(1), 31-38.

Cole, S., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., . . . Barry Iii, C. (1998). Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature, 393*(6685), 537.

Coscolla, M., & Gagneux, S. (2010). Does M. tuberculosis genomic diversity explain disease diversity? *Drug Discovery Today: Disease Mechanisms, 7*(1), e43-e59.

Crawford, J. (2003). Genotyping in contact investigations: a CDC perspective. *The International Journal of Tuberculosis and Lung Disease, 7*(12), S453-S457.

De Jong, B. C., Hill, P. C., Aiken, A., Awine, T., Martin, A., Adetifa, I. M., . . . Gagneux, S. (2008). Progression to active tuberculosis, but not transmission, varies by Mycobacterium tuberculosis lineage in The Gambia. *Journal of Infectious Diseases, 198*(7), 1037-1043.

DPH. (2017). 2016 Georgia Tuberculosis Report.

Driscoll, J. R. (2009). Spoligotyping for molecular epidemiology of the Mycobacterium tuberculosis complex *Molecular Epidemiology of Microorganisms* (pp. 117-128): Springer.

Fenner, L., Egger, M., Bodmer, T., Altpeter, E., Zwahlen, M., Jaton, K., . . . Bruderer, T. (2012). Effect of mutation and genetic background on drug resistance in Mycobacterium tuberculosis. *Antimicrobial agents and chemotherapy, 56*(6), 3047-3053.

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., De Jong, B. C., Narayanan, S., . . . Gutierrez, M. C. (2006). Variable host–pathogen compatibility in Mycobacterium tuberculosis. *Proceedings of the National academy of Sciences of the United States of America, 103*(8), 2869-2873.

Iwamoto, T., Yoshida, S., Suzuki, K., & Wada, T. (2008). Population structure analysis of the Mycobacterium tuberculosis Beijing family indicates an association between certain sublineages and multidrug resistance. *Antimicrobial agents and chemotherapy, 52*(10), 3805-3809.

Jagielski, T., Van Ingen, J., Rastogi, N., Dziadek, J., Mazur, P. K., & Bielecki, J. (2014). Current methods in the molecular typing of Mycobacterium tuberculosis and other mycobacteria. *BioMed research international, 2014*.

Kremer, K., Arnold, C., Cataldi, A., Gutiérrez, M. C., Haas, W. H., Panaiotov, S., . . . van Soolingen, D. (2005). Discriminatory power and reproducibility of novel DNA typing methods for Mycobacterium tuberculosis complex strains. *Journal of clinical microbiology, 43*(11), 5628-5638.

Kremer, K., Van Soolingen, D., Frothingham, R., Haas, W., Hermans, P., Martin, C., . . . Yakrus, M. (1999). Comparison of methods based on different molecular epidemiological markers for typing of Mycobacterium tuberculosis complex strains: interlaboratory study of discriminatory power and reproducibility. *Journal of clinical microbiology, 37*(8), 2607-2618.

Lasunskaia, E., Ribeiro, S. C., Manicheva, O., Gomes, L. L., Suffys, P. N., Mokrousov, I., . . . Otten, T. (2010). Emerging multidrug resistant Mycobacterium tuberculosis strains of the Beijing genotype circulating in Russia express a pattern of biological properties associated with enhanced virulence. *Microbes and Infection, 12*(6), 467-475.

Malla, B., Stucki, D., Borrell, S., Feldmann, J., Maharjan, B., Shrestha, B., . . . Gagneux, S. (2012). First insights into the phylogenetic diversity of Mycobacterium tuberculosis in Nepal. *PLoS One, 7*(12), e52297.

Mitchison, D., Wallace, J., Bhatia, A., Selkon, J., Subbaiah, T., & Lancaster, M. (1960). A comparison of the virulence in guinea-pigs of South Indian and British tubercle bacilli. *Tubercle, 41*(1), 1-22.

Pang, Y., Zhou, Y., Zhao, B., Liu, G., Jiang, G., Xia, H., . . . Zhao, Y.-l. (2012). Spoligotyping and drug resistance analysis of Mycobacterium tuberculosis strains from national survey in China. *PLoS One, 7*(3), e32976.

Parwati, I., van Crevel, R., & van Soolingen, D. (2010). Possible underlying mechanisms for successful emergence of the Mycobacterium tuberculosis Beijing genotype strains. *The Lancet infectious diseases, 10*(2), 103-111.

Rajapaksa, U., & Perera, A. (2011). Sublineages of Beijing strain of Mycobacterium tuberculosis in Sri Lanka. *Indian journal of microbiology, 51*(3), 410-412.

Stucki, D., Brites, D., Jeljeli, L., Coscolla, M., Liu, Q., Trauner, A., . . . Luo, T. (2016). Mycobacterium tuberculosis lineage 4 comprises globally distributed and geographically restricted sublineages. *Nature genetics, 48*(12), 1535.

Van Soolingen, D., Qian, L., De Haas, P., Douglas, J. T., Traore, H., Portaels, F., . . . Van Embden, J. (1995). Predominance of a single genotype of Mycobacterium tuberculosis in countries of east Asia. *Journal of clinical microbiology, 33*(12), 3234-3238.

WHO. (2017). Global tuberculosis report 2017.

CHAPTER 5

EXAMINE THE ASSOCIATION BETWEEN CLUSTERING OF TB ISOLATES AND PATIENT CHARACTERISTICS IN GEORGIA, 2009-2015[2]

[2] Edward A. Sheriff, Juliet N. Sekandi, Kevin Dobbin, Andrea Swartzendruber, Christopher C. Whalen. To be submitted to *International Journal of Tuberculosis and Lung Disease*

Abstract

Statement of Problem: The epidemiology of TB in Georgia is different from that in the United States. TB in Georgia is more common among US-born cases and TB among US-born cases is usually as a result of recent transmission.

Objectives: To examine the relationship between clustering of isolates and TB patient characteristics.

Methods: A secondary analysis of surveillance and molecular data collected by the Georgia department of Health from 2009 to 2015 was done. Case surveillance data were linked to molecular data using a unique state case numbers. Isolates were classified as clustered and unique isolates based on molecular analysis and a plausible source case approach. We used Chi-square to test the differences between groups in categorical variables. Log binomial regression model was used to compare patient characteristics associated with clustering of isolates.

Results: of the 1,450 isolates, 433 (30%) were clustered or could be attributable to recent transmission. US-born cases, Black cases, homeless cases and cases with isolates resistant to isoniazid were significant predictors of clustering or recent transmission.

Conclusions: TB transmission in Georgia is more concentrated in congregate settings and among vulnerable minority populations. Interrupting TB transmission should be Georgia's urgent priority if they're to achieve elimination. Aggressive case findings, contact tracing and effective treatment programs should be implemented in counties with homeless shelters.

Introduction

In 2016, 6.3 million new cases of tuberculosis (TB) were reported to the World Health Organization (WHO, 2017). This is equivalent to 61% of the estimated 10.4 million cases every year (WHO, 2017). 1.7 million deaths worldwide were attributed to TB in 2016 (WHO, 2017). In the United States, 9,272 TB cases with a case rate of 2.9 cases per 100,000 persons were reported in 2016. The current TB case count in the United States represents a decline from 2015 and the lowest case count on record in the United States (CDC, 2017). The case rate of 2.9 cases per 100,000 persons is a 3.6% decrease from 2015 (CDC, 2017). In Georgia, 301 new TB cases were reported in 2016 with a case rate of 2.9 cases per 100,000 persons (DPH, 2017). The incidence rate of TB in Georgia is equal to that of the national incidence, and represents a 6% decline from the incidence rate of 3.1 cases per 100,000 persons in 2015 (DPH, 2017).

The time from TB infection to TB disease varies from one individual to another. Incident TB cases are comprised of reactivation of a historic infection or the result of a recent transmission event (Vynnycky & Fine, 1997; Vynnycky et al., 2001). The treatment of reactive TB and TB resulting from recent transmission are similar, but cases resulting from recent transmission are of more public health importance because they may be part of an ongoing outbreak or a series of transmission events that require public health control measures (Mathema, Kurepina, Bifani, & Kreiswirth, 2006). By virtue of this variability in latency of infection, the timing of transmission events often remains elusive, making it challenging to study transmission dynamics of *Mycobacterium tuberculosis*.

Molecular typing of M. tuberculosis has been used to study the epidemiology of tuberculosis to include estimating the fraction of cases attributable to recent transmission or reactivation (Barnes et al., 1996; Borgdorff et al., 1998; Small et al., 1994). Using spacer oligonucleotide typing (Spoligotyping) and mycobacterial interspersed repetitive units variable number tandem repeats (MIRU-VNTR) analysis; isolates from TB cases can be classified as being clustered or unique. Isolates with identical Spoligotyping and MIRU-VNTR patterns are considered clustered isolates. Therefore, patients harboring M. tuberculosis strains with identical Spoligotyping or MIRU-VNTR patterns suggest an epidemiologic link between such clustered cases.

Population based molecular epidemiologic studies from San Francisco, New York, The Netherlands, and Denmark noted that molecular clustering of TB isolates ranged on average between 35% and 45%, indicating a substantial proportion due to recent transmission (Alland et al., 1994; Bauer, Yang, Poulsen, & Andersen, 1998; van Soolingen et al., 1999). It is important to note that clustered TB cases do not always represent recent transmission. Some strains of tuberculosis may remain stable for years and the same strain may appear among patients who have never been in contact with other current cases. Despite this caveat, clustering of strains is still accepted as the best indicator for recent transmission of TB. Population studies have shown that contact cases of tuberculosis often share the same strain of tuberculosis as the source case (McConkey et al., 2002; Small et al., 1994; Whalen et al., 2011). TB in Georgia as opposed to the United States is more common among US-born persons. We know that TB among US-born persons is usually as a result of recent transmission. Therefore, understanding the transmission pattern of Mycobacterium tuberculosis will provided useful information that can be used to interrupt transmission and eventually eliminate TB in Georgia and the United states.

We conducted an in-depth analysis of the transmission dynamic of TB in Georgia. This long-term population-based study was designed to describe the socio-demographic and clinical risk factors associated with clustered versus unique strains of TB in 7 years (2009 to 2015). The null hypothesis was that the proportions of each of the demographic and clinical variables are independent of clustering among isolates. The alternative hypothesis was that patient demographic and clinical characteristics have an effect on clustering of isolates.

Methods

We analyzed cases of pulmonary TB reported in Georgia from January 1, 2009 to December 31, 2015. TB patients were diagnosed and classified according to the Centers for Disease Control and Prevention's (CDC) clinical case and laboratory criteria (Wharton, Chorba, Vogt, Morse, & Buehler, 1990). For each case of TB included in this study, we obtained demographic, clinical and molecular data from the Georgia, electronic notifiable disease surveillance system. Only culture-confirmed cases with molecular data were selected for inclusion in this study. Case surveillance data were linked to their respective molecular data via a state case number. Cases without state case numbers were excluded from this study. The state of Georgia is the largest state east of the Mississippi River with 159 counties and 18 Health Districts. The population in Georgia in 2015 was estimated at 10,310,371 with 9.8% being non-US born persons (Census, 2016). Of the 159 counties in Georgia, only 4 counties accounted for 55% of all TB cases reported in Georgia in 2015.

Molecular analysis and cluster definition

The methods used to characterize *mycobacterium tuberculosis* isolates were Spoligotyping and MIRU VNTR. We used a plausible source case approach to define clusters. This approach has been validated by the CDC (France, Grant, Kammerer, & Navin, 2015). Using this approach, each case of TB in our study was compared to other cases with the same Spoligotyping and MIRU-VNTR patterns to determine whether a source case can be identified (France et al., 2015). A plausible source case was defined as a case of pulmonary TB in a patient over 4 years of age, lived in the same county as the given case, and diagnosed within 2 years prior to the given case (France et al., 2015). The plausible source case approach allows us to compare isolates across time and geographic area using a standard method. Each case in our study was compared to other cases with the same genetic patterns based on Spoligotyping and MIRU-VNTR analysis to determine whether a source case can be identified. Clustered isolates had identical Spoligotyping and MIRU-VNTR patterns and met the plausible source case definition. Cases of TB were divided into clustered and non-clustered isolates and patient characteristics evaluated.

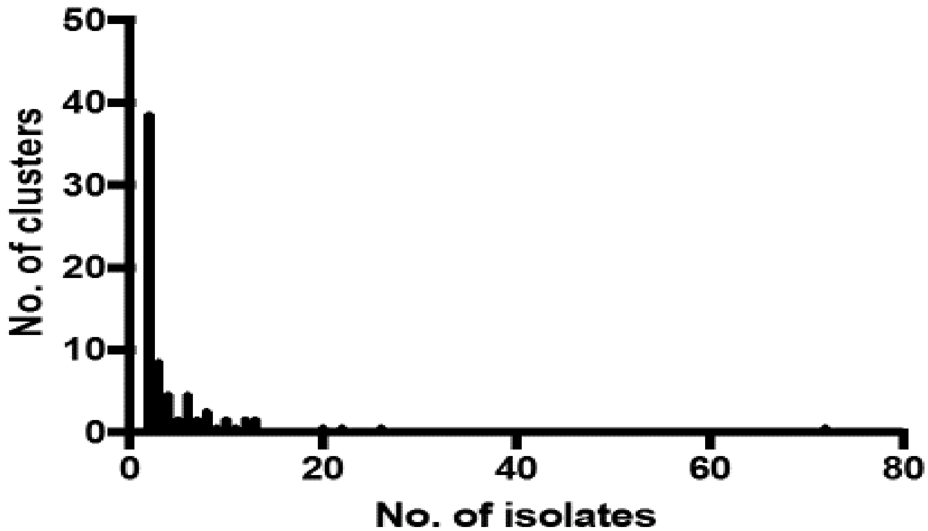
Statistical Analysis

Chi-square tests were performed to test the association of clustering with categorical predictor variables. Kruskal Wallis rank test was used to compare the median ages of clustered cases to non-clustered cases. To correct for multiple comparison, we used Bonferroni correction. Predictor variables that were significantly associated with clustering ($P<0.05$) were included in a log-binomial regression model. Manual backwards elimination was used for variable selection. Associations were considered statistically significant at $\alpha \leq 0.05$. All statistical analyses were performed using S.A.S. 9.4 (SAS Institute, Cary, NC, USA).

Results

From January 1, 2009 to December 31, 2015 we evaluated 1,450 cases of pulmonary TB with molecular data linked to case surveillance data. Based on unique combinations of Spoligotyping and 24-locus MIRU-VNTR we found that 710 of the 1450 (49%) of the isolates were clustered resulting in 154 clusters. When the plausible source case approach was implemented, we found that 433 (29.9%) isolates were clustered and 1017 (70.1%) were unique isolates. This resulted in 77 clusters and the number of cases within a cluster varied from 2 to 76. Using the plausible source case approach, a total of 277 (39.0%) isolates that were initially clustered based on genetic analysis alone were no longer clustered. Figure 5.1 shows the distribution of clusters by number of isolates in Georgia from 2009 to 2015. Majority of the clusters were made up 2 isolates representing 50.6% of all clusters. 70% (1012) of the isolates in our study came from male cases and 30% (438) from female cases. 50% (719) of all cases were black, 31% (445) were white, 18% (264) were Asian and 2% (22) was American Indians or Alaskan Natives. 56% (811) of the cases in our study were US-born cases and 82% (1186) of the cases were non-Hispanics. The median age was 46 (interquartile range [IQR] 32-57). Females were significantly younger than males (median age 39 versus 47 years, $p < .0001$). Non-US-born cases were significantly younger than US-born-cases (median age 38 versus 51, $p < .0001$). There was no significant difference in the median ages of clustered isolates and unique isolates (median age 48 versus 45 years, $p = 0.8331$).

Figure 5.1: Sizes and frequencies of 77 clusters consisting of 433 TB patient isolates in Georgia, US, 2009-2015



Bivariate Analysis

76% of the male cases were clustered compared to 24% of the female cases, (1, N = 1450) = 9.6028, p = .0304. 44% of all Black cases were clustered compared to 9% of Asian cases and 20% of White cases, (3, N = 1450) = 150.4262, p = .0016. Non-Hispanic cases were more likely to have clustered isolates than Hispanic cases, (1, N = 1450) = 9.6028, p = .0016. Amongst US-born cases, 45% of the isolates were clustered compared to 10% among non-US-born cases, (1, N = 1450) = 48.1277, p = .0016. Isolates that were resistant to isoniazid were more likely to be clustered compared to isolates sensitive to isoniazid, (1, N = 1450) = 62.1320, p = .0016. Homelessness, excess alcohol consumption, non-injection drug use and sputum smear positivity were significantly associated with clustering of isolates. Table 5.1 shows patient characteristics associated with clustering of tuberculosis isolates in Georgia from 2009 to 2015.

In a multivariate regression model, we used log binomial regression controlling for independent variables that were significantly associated with clustering based on bivariate analysis. Race, sex, ethnicity, alcohol consumption, drug use, country of origin, homelessness, sputum smear positivity, and isoniazid resistance were included in the multivariate analysis. Since the outcome of clustering was greater than 10%, it was more desirable to estimate a relative risk. Table 5.2 shows unadjusted relative risks and corresponding 95% confidence intervals. After controlling for all the other variables in the model, the risk of clustering of isolates is 1.6 times higher in Blacks compared to all other races combined, [RR: 1.5626; 95% CI: 1.2848-1.9004]. After controlling for all other variables in the model, the risk of clustering of isolates was 3 times higher in US-born cases compared to non-US-born cases, [RR: 3.0505; 95% CI: 2.3455-3.9673]. Amongst homeless cases, the risk of clustering of isolates was 1.4 times higher compared to non-homeless cases, after controlling for all other variables in the model, [RR: 1.4207; 95% CI: 1.2177-1.6575]. After controlling for all other variables in the model, the risk of clustering of isolates was 1.5 times higher in cases resistance to isoniazid compared to cases susceptible to isoniazid, [RR: 1.4915; 95% CI: 1.2836-1.7332]. Variables that were not significant in the multivariate analysis were removed from the model. The final model included race, country of origin, homelessness, and isoniazid resistance. Table 5.3 shows adjusted relative risks and corresponding 95% confidence intervals.

Discussion

In this secondary analysis of 1,450 patient isolates of M. tuberculosis collected over a 7-year period, we found that 49% of the isolates were considered clustered based on Spoligotyping and 24-locus MIRU-VNTR alone. Studies have shown that using

Spoligotyping and 12-locus MIRU-VNTR could overestimate the cluster proportion in a cosmopolitan set of isolates (Maes, Kremer, van Soolingen, Takiff, & de Waard, 2008; Supply et al., 2006). Applying the plausible source case method, we further reduced the number of clustered cases to 433 within the 7-year period. 30% of the cases in our study were clustered or could be attributed to recent transmission. The rate of transmission was calculated to be 25% within a 2 year period. Therefore, 25% of the cases of TB in Georgia could be attributed to recent transmission. This is in agreement with other studies conducted in San Francisco, New York, Alabama and some countries in Europe and Africa (Alland et al., 1994; Bauer et al., 1998; Genewein et al., 1993; Kempf et al., 2005; Pfyffer et al., 1998; Small et al., 1994; Whalen et al., 2011; Wilkinson et al., 1997).

Cases belonging to large clusters share demographic characteristics like homelessness, being US-born and or Black race. The occurrence of large clusters in Georgia, may indicate a gap in the tuberculosis control activities. The high proportion of clustered isolates among Blacks underscores the idea that TB disproportionately affects racial/ethnic minorities in Georgia. This is alarming since according to 2016 census estimates, only 32% of Georgian residents are Black (Census, 2016). According to the state’s TB surveillance report, TB case rate among Black non-Hispanics is more than six times higher than among white non-Hispanics. Although rates of TB in Blacks have declined substantially over the past decade, the disparity still remains. Among the Black population, socioeconomic factors impact health outcomes and are associated with poverty, including limited access to quality health care, unemployment, housing, and transportation. These factors can directly or indirectly increase the risk for TB disease and present barriers to treatment. Ongoing efforts are needed to address the persistent disparities that exist among racial and ethnic minorities in the United States. TB control efforts should focus on increasing the awareness about TB in Black communities and design new ways to address health system barriers specifically for Blacks with TB and those at risk of the disease.

Homelessness has long been implicated as an important risk factor for the spread of TB infection (Anderson, Story, Brown, Drobniewski, & Abubakar, 2010; Powell et al., 2017). The risk factors for homelessness often overlap with the risk factors for TB, and homeless people may be particularly vulnerable due to poor nutrition and weakened immunity. The fact that most homeless shelters are crowded and poorly ventilated may be a contributory factor in the spread of TB in this setting. Other studies assessing the burden of TB in specific vulnerable populations such as prisoners, homeless and certain ethnic minorities show that there is a strong association between social deprivation and TB risk (Barnes et al., 1996; Curtis et al., 2000; Zolopa et al., 1994). In our study, we found that extensive transmission of isoniazid-resistant TB occurs among homeless cases. The homeless population in Georgia represents an important risk group among U.S-born TB patients. To achieve TB elimination, ongoing efforts are needed to address the disproportionate number of TB cases among this high-risk population. TB control program in Georgia should actively collaborate with other public health organizations to improve screening, diagnosis, and treatment for people experiencing homelessness. Improving social conditions like housing affordability and acces to medical care for poor people may indirectly reduce the burden of TB in Georgia and the United States.

We found that isoniazid resistance strains were more likely to be clustered than isoniazid-sensitive strains. This is in agreement with other studies conducted in the

United States (Alland et al., 1994; Beck-Sagué et al., 1992; Frieden, Fujiwara, Washko, & Hamburg, 1995; Small et al., 1994; Snider Jr, Kelly, Cauthen, Thompson, & Kilburn, 1985). However, this contradicts results from a similar study conducted in Netherlands that found that drug-resistant strains were less likely to be clustered than drug-sensitive strains. In addition, most animal models have shown that isoniazid resistance may lead to reduced virulence of Mycobacterium tuberculosis (Middlebrook, 1954; Wolinsky, SMITH, & Steenken Jr, 1956). In our study we found that 62% of the clustered isolates that were resistant to isoniazid came from homeless cases. Even though drug resistance may lead to reduction in virulence, the close proximity of homeless persons in shelters may play a vital role in the transmission of a rather less virulent drug-resistant strain of TB.

The low proportion of clustering of isolates among non-US-born cases underscores the idea that an immigrant's risk of tuberculosis in the United States is associated with tuberculosis incidence in their country of origin (Ricks, Cain, Oeltmann, Kammerer, & Moonan, 2011). Majority of the clustered isolates came from US-born cases and risk factors such as homeless and Black race were more common among US-born cases. The highest priority should be given to detection, evaluation, and reporting of homeless persons who have current symptoms of active TB and completion of an appropriate course of treatment by those diagnosed with active TB. It is important to screen and provide preventive therapy for homeless persons who have, or are suspected of having, HIV infection. Since drug resistant isolates in Georgia are more likely to spread, priority should be given to the examination and appropriate treatment of persons with recent TB that has been inadequately treated. Contacts of homeless cases are difficult to ascertain, therefore all residents in homeless facility should be screened for TB infection when an infectious case is identified.

The present study did not link traditional contact investigation data with molecular data and clustering of isolates by molecular testing does not necessarily mean the cases were contacts or had a common contact. However, we used a combination of two robust molecular analysis technologies in combination with the plausible source case approach to classify isolates into clusters and we evaluated the isolates in a 7-year time period. Due to the application of such stringent rules, we may have underestimated the true rate of transmission of TB in Georgia. However, a transmission rate of 25% is within the range expected in the United States. To our knowledge, this is the first validated approach to estimating clustering of isolates in the state of Georgia. Whole genome sequencing will be the next phase of molecular analysis and future research using this molecular technique should provide a more accurate means of classifying isolates since it analyzes the whole genome of *Mycobacterium tuberculosis*.

Table 5.1: Patient demographic and clinical characteristics and their association with clustering of TB isolates in Georgia, 2009-2015 (N=1450)					
Characteristics	**Clustered n (%)**	**Unique n (%)**	**Total**	**Chi-Square**	**p-value[1]**
Age (Median years, IQR[2])			46 (32-57)	0.01893	1.00
Race				150.4262	0.0016
AI/AN[4]	2 (9.1)	20 (90.9)	22		
Asian	24 (9.1)	240 (90.9)	264		
Black/AA[5]	318 (44.2)	401 (55.8)	719		
White	89 (20.0)	356 (80.0)	445		
Sex				9.6028	0.0304
Female	106 (24.2)	332 (75.8)	438		
Male	327 (75.5)	685 (67.8)	1012		
Country of Origin				204.8073	0.0016
Non-US-born	67 (10.5)	572 (89.5)	639		
US-born	366 (45.1)	445 (54.9)	811		
Ethnicity				48.1277	0.0016
Hispanic	32 (12.2)	231 (87.8)	263		
Non-Hispanic	401 (33.8)	785 (66.2)	1186		
Missing			1		
Alcohol Abuse				35.1046	0.0016
No	316 (26.7)	868 (73.3)	1184		
Yes	117 (45.4)	141 (54.6)	258		
Missing			8		
Homeless				103.3505	0.0016
No	316 (25.1)	942 (74.9)	1258		
Yes	117 (61.3)	74 (38.7)	191		
Missing			1		
Cavity				1.8427	0.1811
No	255 (28.9)	628 (71.1)	883		
Yes	149 (32.5)	310 (67.5)	459		
Missing			108		
INH[6] Resistance				62.132	0.0016
Resistant	105 (54.1)	89 (45.9)	194		
Sensitive	327 (26.2)	919 (73.8)	1246		
Missing			10		
Drug Use (Injecting)				0.8814	1
No	428 (29.9)	1003 (70.1)	1431		
Yes	4 (33.3)	8 (66.7)	12		
Unknown	1 (14.3)	6 (85.7)	7		
Drug Use (Non-injecting)				54.6261	0.0016

No	345 (26.9)	939 (73.1)	1284		
Missing			7		
Yes	88 (55.4)	71 (44.6)	159		
Previous TB				2.0959	1
No	403 (29.5)	965 (70.5)	1368		
Missing			1		
Yes	30 (37.0)	51 (63.0)	81		
Prison				0.4087	1
No	405 (29.7)	960 (70.3)	1365		
Yes	28 (32.9)	57 (67.1)	85		
LTCF[7]				3.1221	1
No	431 (30.1)	1001 (69.9)	1432		
Yes	2 (11.8)	15 (88.2)	17		
Missing			1		
Sputum Smear				17.8616	0.0016
Negative	118 (23.3)	388 (76.7)	506		
Positive	305 (34.1)	589 (65.9)	894		
Case Status				4.1547	0.6656
Alive	428 (30.3)	987 (69.7)	1415		
Dead	5 (14.3)	30 (85.7)	35		

[1] Bonferroni Correction, [2] Interquartile Range, [3] Kruskal Wallis test, [4] American Indian/Alaskan Native, [5] African American, [6] Isoniazid,

Table 5.2. Unadjusted relative risks and 95% confidence intervals from binomial log regression examining the relationship between clustering of isolates and patient characteristics

Patient Characteristics	**Relative Risk**	**95% CI**
Race (Blacks vs All Other)	2.8114	2.8114-2.3324*
Sex (Male vs. Female)	1.3352	1.1061-1.6117*
US (US vs. NON-US)	4.3041	3.3894-5.4657*
Sputum Smear (Pos. vs. Neg.)	1.4630	1.2191-1.7556*
Homelessness	2.4386	2.1037-2.8269*
Ethnicity (non-Hispanic vs. Hispanic)	2.7789	1.9891-3.8821*
Excess Alcohol	1.6991	1.4423-2.0017*
Non-Injection Drug	2.0598	1.7444-2.4324*
INH[1] Resistant (Resistant vs. Susceptible)	2.0623	1.7582-2.4190*

[1] Isoniazid * Significant Relative Risk

Table 5.3. Adjusted relative risk and 95% confidence intervals from binomial log regression examining the relationship between clustering of isolates and patient characteristics

Patient Characteristics	**Odds Ratio**	**95% CI**
Race (Blacks vs All Other)	1.5595	1.2591-1.9314*
Sex (Male vs. Female)	1.1383	0.9469-1.3683
US (US vs. NON-US)	2.9689	2.2469-3.9230*
Sputum Smear (Pos. vs. Neg.)	1.1465	0.9798-1.3416
Homeless	1.3614	1.1625-1.5943*
Ethnicity (non-Hispanic vs. Hispanic)	1.0862	0.7398-1.5947
Excess Alcohol	0.9970	0.8671-1.1463
Non-Injection Drug	0.9814	0.8396-1.1472
INH[1] Resistant (Resistant vs. Susceptible)	1.4452	1.2414-1.6824*

[1] Isoniazid * Significant Relative Risk

References: Chapter 5

WHO. Global tuberculosis report 2017. 2017.

CDC. Reported Tuberculosis in the United States, 2016. 2017.

DPH. 2016 Georgia Tuberculosis Report. 2017.

Vynnycky E, Fine P. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. Epidemiology & Infection. 1997;119(2):183-201.

Vynnycky E, Nagelkerke N, Borgdorff M, Van Soolingen D, Van Embden J, Fine P. 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 & Infection. 2001;126(1):43-62.

Mathema B, Kurepina NE, Bifani PJ, Kreiswirth BN. Molecular epidemiology of tuberculosis: current insights. Clinical microbiology reviews. 2006;19(4):658-85.

Borgdorff MW, Nagelkerke N, van Soolingen D, de Haas PE, Veen J, van Embden JD. Analysis of tuberculosis transmission between nationalities in the Netherlands in the period 1993–1995 using DNA fingerprinting. American journal of epidemiology. 1998;147(2):187-95.

Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, Ruston DC, et al. The Epidemiology of Tuberculosis in San Francisco--A Population-Based Study Using Conventional and Molecular Methods. New England Journal of Medicine. 1994;330(24):1703-9.

Barnes PF, El-Hajj H, Preston-Martin S, Cave MD, Jones BE, Otaya M, et al. Transmission of tuberculosis among the urban homeless. Jama. 1996;275(4):305-7.

Alland D, Kalkut GE, Moss AR, McAdam RA, Hahn JA, Bosworth W, 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(24):1710-6.

Bauer J, Yang Z, Poulsen S, Andersen ÅB. Results from 5 years of nationwide DNA fingerprinting of Mycobacterium tuberculosis complex isolates in a country with a low incidence of M. tuberculosis infection. Journal of clinical microbiology. 1998;36(1):305-8.

van Soolingen D, Borgdorff MW, de Haas PE, Sebek MM, Veen J, Dessens M, et al. Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. Journal of Infectious Diseases. 1999;180(3):726-36.

McConkey SJ, Williams M, Weiss D, Adams H, Cave MD, Yang Z, 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(5):612-9.

Whalen CC, Zalwango S, Chiunda A, Malone L, Eisenach K, Joloba M, et al. Secondary attack rate of tuberculosis in urban households in Kampala, Uganda. PloS one. 2011;6(2):e16137.

Wharton M, Chorba TL, Vogt RL, Morse DL, Buehler JW. Case definitions for public health surveillance. 1990.

Census B. State Quickfacts: Georgia 2016 [updated 03/20/2017. Available from: https://www.census.gov/quickfacts/table/PST045216/13.

France AM, Grant J, Kammerer JS, Navin TR. A field-validated approach using surveillance and genotyping data to estimate tuberculosis attributable to recent transmission in the United States. American journal of epidemiology. 2015;182(9):799-807.

Maes M, Kremer K, van Soolingen D, Takiff H, de Waard JH. 24-locus MIRU-VNTR genotyping is a useful tool to study the molecular epidemiology of tuberculosis among Warao Amerindians in Venezuela. Tuberculosis. 2008;88(5):490-4.

Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, Willery E, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. Journal of clinical microbiology. 2006;44(12):4498-510.

Kempf M-C, Dunlap NE, Lok KH, Benjamin WH, Keenan NB, Kimerling ME. Long-term molecular analysis of tuberculosis strains in Alabama, a state characterized by a largely indigenous, low-risk population. Journal of clinical microbiology. 2005;43(2):870-8.

Pfyffer G, Strassle A, Rose N, Wirth R, Brandli O, Shang H. Transmission of tuberculosis in the metropolitan area of Zurich: a 3 year survey based on DNA fingerprinting. European Respiratory Journal. 1998;11(4):804-8.

Genewein A, Telenti A, Bernasconi C, Schopfer K, Bodmer T, Mordasini C, et al. Molecular approach to identifying route of transmission of tuberculosis in the community. The Lancet. 1993;342(8875):841-4.

Wilkinson D, Pillay M, Crump J, Lombard C, Davies GR, Sturm AW. Molecular epidemiology and transmission dynamics of Mycobacterium tuberculosis in rural Africa. Tropical Medicine & International Health. 1997;2(8):747-53.

Kiers A, Drost A, Van Soolingen D, Veen J. Use of DNA fingerprinting in international source case finding during a large outbreak of tuberculosis in The Netherlands. The International Journal of Tuberculosis and Lung Disease. 1997;1(3):239-45.

Frieden TR, Fujiwara PI, Washko RM, Hamburg MA. Tuberculosis in New York City—turning the tide. New England Journal of Medicine. 1995;333(4):229-33.

Snider Jr DE, Kelly GD, Cauthen GM, Thompson NJ, Kilburn JO. Infection and disease among contacts of tuberculosis cases with drug-resistant and drug-susceptible bacilli. American Review of Respiratory Disease. 1985;132(1):125-32.

Beck-Sagué C, Dooley SW, Hutton MD, Otten J, Breeden A, Crawford JT, et al. Hospital outbreak of multidrug-resistant Mycobacterium tuberculosis infections: factors in transmission to staff and HIV-infected patients. Jama. 1992;268(10):1280-6.

Middlebrook G. Isoniazid-resistance and catalase activity of tubercle bacilli; a preliminary report. American Review of Tuberculosis. 1954;69(3):471.

Wolinsky E, SMITH MM, Steenken Jr W. Isoniazid Susceptibility, Gatalase Activity, and Guinea Pig Virulence of recently isolated Cultures of Tubercle Bacilli. American Review of Tuberculosis and Pulmonary Diseases. 1956;73(5):768-72.

Anderson C, Story A, Brown T, Drobniewski F, Abubakar I. Tuberculosis in UK prisoners: a challenge for control. Journal of epidemiology and community health. 2010:jech. 2009.094375.

Powell KM, VanderEnde DS, Holland DP, Haddad MB, Yarn B, Yamin AS, et al. Outbreak of Drug-Resistant Mycobacterium tuberculosis Among Homeless People in Atlanta, Georgia, 2008-2015. Public Health Reports. 2017;132(2):231-40.

Curtis AB, Ridzon R, Novick L, Driscoll J, Blair D, Oxtoby M, et al. Analysis of Mycobacterium tuberculosis transmission patterns in a homeless shelter outbreak. The International Journal of Tuberculosis and Lung Disease. 2000;4(4):308-13.

Zolopa AR, Hahn JA, Gorter R, Miranda J, Wlodarczyk D, Peterson J, et al. HIV and tuberculosis infection in San Francisco's homeless adults: prevalence and risk factors in a representative sample. Jama. 1994;272(6):455-61.

Ricks PM, Cain KP, Oeltmann JE, Kammerer JS, Moonan PK. Estimating the burden of tuberculosis among foreign-born persons acquired prior to entering the US, 2005–2009. PLoS One. 2011;6(11):e27405.

CHAPTER 6

EXAMINE THE PERSISTENCE OF MYCOBACTERIUM TUBERCULOSIS STRAINS IN GEORGIA, 2009-2015[3]

[3] Edward A. Sheriff, Kevin Dobbin, Andrea Swartzendruber, Christopher C. Whalen, Juliet N. Sekandi. To be submitted to *International Journal of Tuberculosis and Lung Disease*

Abstract

Statement of Problem: TB in Georgia is mainly as a result of recent transmission and certain strains of TB continue to produce secondary cases year after year.

Objectives: To examine the relationship between strain persistence and TB patient characteristics.

Methods: Clustered isolates were classified as persistently successful, transiently successful and unsuccessful strains based on cluster size and duration in the state. We used Chi-square to test the differences between groups in categorical variables. Multinomial logistic regression model was used to compare patient characteristics associated with strain persistence.

Results: of the 706 isolates clustered based on molecular testing, 198 (28%) were persistently successful strains. 95% of all persistently successful strains were isolated from non-US-born cases. 81% of all resistant strains were persistently successful. US-born cases, Blacks cases, homeless or prison cases, and cases with isolates resistant to isoniazid were significant predictors of strain persistence.

Conclusions: TB transmission in Georgia is more concentrated in congregate settings and among vulnerable minority populations. Strains of TB that have persisted in Georgia from 2009 to 2017 have mostly been associated with drug resistance and congregate settings. The close proximity of persons in homeless shelters and prisons increases the likelihood of transmission given an infectious case. TB control in Georgia should provide a comprehensive policy to improve the management of TB cases from congregate settings.

Introduction

Tuberculosis (TB) continues to pose significant health problems around the world and control strategies are failing to contain the spread, even in countries with low incidence (WHO, 2017). With the advent of molecular epidemiology, we have expanded our ability to explore TB epidemics and study strains of *M. tuberculosis* (Barnes & Cave, 2003; Mathema, Kurepina, Bifani, & Kreiswirth, 2006). Community based studies have reported a high diversity of MTBC strains in endemic settings and identified clinical characteristics associated with *Mycobacterium tuberculosis* strain clustering and cluster size (Glynn et al., 2008; Lockman et al., 2001; Van der Spuy et al., 2009; Verver et al., 2004; Warren et al., 1996). Cases of TB resulting from recent transmission are particularly concerning from a public health standpoint because they represent the possibility of ongoing transmission from unrecognized infectious cases and the presence of recently infected contacts who would benefit from preventive therapy (Yuen, Kammerer, Marks, Navin, & France, 2016). The size of clusters could depend on factors favoring transmission or on physiological differences in the strains themselves (Glynn et al., 2008). *Mycobacterium tuberculosis* strains found in persons with smear-positive disease, many contacts, or persons experiencing delays in diagnosis and treatment are particularly likely to be transmitted. In Georgia, TB disproportionately occurs in certain demographic groups like Blacks and homeless persons and the proportion of clustered isolates are higher in these groups. The epidemiology of TB in Georgia is different from that in the United States. Given that TB burden is highest among US-born persons and occurs mainly as a result of recent transmission, if recent transmission of TB is left undetected or unchecked, outbreaks can occur. This is evident in Georgia, where multiple outbreaks of TB have been reported since 2008 (Bamrah et al., 2013; Powell et al., 2017). In Georgia, 4 counties account for more than 50% of the TB cases reported annually and most outbreaks of TB in Georgia are concentrated among certain vulnerable populations like the homeless (DPH, 2017). The success of a strain of M. tuberculosis may depend on which demographic group that strain first appeared in. Previous analyses have described patient characteristics associated with genotypic clustering, using shared genotypes within a geographically defined population as an indicator for recent transmission (Alland et al., 1994; Bauer, Yang, Poulsen, & Andersen, 1998; Kempf et al., 2005; Small et al., 1994; van Soolingen et al., 1999). Based on a review of TB outbreaks in the United States from 2002 to 2008, the CDC found that most of the outbreaks during this time period involved US-born persons who use illicit drugs or consumed excess alcohol (Mitruka, 2011). There is not enough data describing the host, environmental and bacterial factors associated with the persistence of M. tuberculosis strains in Georgia. We therefore analyzed data from 2009 to 2015 to examine persistence of Mycobacterium tuberculosis strains and the risk factors associated with the most persistent strains of M. tuberculosis in Georgia.

Methods

This was a secondary analysis of all TB cases diagnosed in Georgia between 2009 and 2015. All culture confirmed cases diagnosed between January 1, 2009 and December 31, 2015 with genotyping data were selected for inclusion in the study. Georgia ranks fifth in the number of new TB cases and tenth in the case rates among the 50 reporting states in the United States in 2016 (CDC, 2017). Even though there are 159 counties in Georgia, TB is disproportionality reported from four counties-Fulton, DeKalb, Gwinnett and Cobb counties. These counties accounted for 55% of all TB cases reported in Georgia in 2015 (DPH, 2017).

Spacer Oligonucleotide typing (spoligotyping) and Mycobacterial interspersed repetitive units-variable number tandem repeats (MIRU-VNTR) were used to characterize M. tuberculosis isolates. Case surveillance data was linked to isolate molecular data via a “State Case Number”. Cases without state case numbers were excluded. For this analysis, clustered cases were further classified into persistently successful, transiently successful and, unsuccessful strains. Persistently successful strains were strains that had been present for at least 3 consecutive years from 2009 and were still present in 2015 and had greater than or equal to 2 cases per year (Middelkoop et al., 2014). Transiently successful strains were strains that did not persist to 2015 but occurred for at least 3 consecutive years from 2009, with a cluster size of at least 2 cases per year in the consecutive period (Middelkoop et al., 2014). Unsuccessful strains did not fall into the above categories and also includes clustered strains occurring for the first time. Patient clinical and demographic characteristics associated with strain success were examined.

Statistical Analysis

Data were analyzed using SAS 9.4 (SAS Institute, Cary, NC, USA). Descriptive statistics were used to examine the distribution of the outcome and predictor variables. Kruskal-Wallis rank test was used to determine if the median age was equal amongst groups of the dependent variable. Categorical variables were summarized using frequencies. Bivariate analysis employed Chi-square test of independence to determine whether the proportions for each categorical independent variable are different among values of the dependent variable (strain persistence). Variables that were significant in the bivariate analysis were included in a multinomial logistic regression model to examine factors associated with persistently successful, transiently successful and unsuccessful strains. Odds ratios and corresponding 95% confidence intervals were reported when comparing persistently successful strains to transiently successful and unsuccessful strains. Associations were considered statistically significant at $\alpha \leq 0.05$.

Results

Based on molecular analysis using Spoligotyping and MIRU-VNTR, 706 isolates were clustered. We observed six different strains of Mycobacterium tuberculosis for which cases were reported every year from 2009 to 2015. These six strains accounted for 198 isolates. Amongst the isolates that occur every year, 96% were from US-born cases and 85% were from Black cases. The prevalence of isoniazid resistance amongst strains that occur every year was 54% and the prevalence of sputum smear positivity was 68%. Figure 6.1 shows the distribution of these six strains and the number of years in which they occur. The strain with the highest number of isolates occurring every year was designated 'G05625'. There were 76 isolates associated with this strain and 87% of these isolates were from Black cases. 93.4% were from US-born cases, 82% were from homeless cases and 100% of the isolates were resistant to isoniazid. Figure 6.2 shows the distribution of isolates of 'G05625' from 2009 to 2015.

Based on our criteria for persistence, 28% (198) were found to be persistently successful, 19% (131) were transiently successful and 53% (377) were unsuccessful strains. The median age was 48 (interquartile range [IQR] 34-56). Females were significantly younger than males (median age 43 versus 50 years, $p < .0001$). Non-US-born cases were significantly younger than US-born cases (median age 38 versus 49, $p < .0001$). There was no significant association between age and persistence of strains ($p=0.6370$).

Table 6.1 Shows patient characteristics associated with persistently successful strains compared with those of transiently successful and successful strains. One striking finding was that 81% (106) of the strains resistant to isoniazid were categorized as persistently successful strains and only 3.1% and 16 % were transiently successful and unsuccessful strains respectively, χ^2 (1, N = 706) = 224.3, $p < .0001$. Majority of the persistently successful strains were isolated from black cases, χ^2 (2, N = 706) = 68.1, $p < .0001$. There was a significant association between homelessness and persistence of strains. 57% of all homeless cases in our study, had persistently successful strains, χ^2 (2, N = 706) = 78.7, $p < .0001$. Majority of the cases resident in a correctional facility had persistently successful strains, χ^2 (2, N = 706) = 23.5, $p < .0001$. Country of origin was a significant predictor of persistence of strains. 34% of all US-born cases had persistently successful strains compared to only 5% of non-US-born cases, χ^2 (2, N = 706) = 74.3, $p < .0001$.

Multivariate Analysis

Multinomial logistic regression was used to examine the association between persistence of strains and patient demographic and clinical characteristics. From previous analysis, we know that TB disproportionately affects Black persons in Georgia, so we compared blacks to all other races in multivariate analysis.

Persistently successful vs. unsuccessful strains

Multivariate analysis showed that race, non-injection drug use, residence in a correctional facility, country of origin, and the interaction of isoniazid resistance and homelessness were significant predictors of strain persistence. Country of birth was a significant predictor of persistence. Persistence strains were more likely to come from US-born cases than non-US-born cases, [OR: 7.8; 95% CI: 2.9-20.9]. After controlling for other variables, Blacks cases were 3 times more likely to have persistently successful strains compared to other races, [OR: 2.7; 95% CI: 1.5 – 5.1]. Isolates from cases resident in a correctional facility were 5 times more likely to have persistently successful strains compared to isolates from non-incarcerated cases, [OR: 5.5; 95% CI: 1.9-16.0].

Isolates from homeless cases that are resistant to Isoniazid were 33 times more likely to be categorized as persistently successful strains. Even in the absence of homelessness, isolates resistant to isoniazid were most 19 times more likely to be categorized as persistently successful strains. Tables 6.2 and 6.3 show patient characteristics associated with strain persistence along with unadjusted and adjusted odds ratios and corresponding 95% confidence intervals.

Transiently successful strains vs. Unsuccessful Strains

Cases that use non-injection drugs were 2 time more likely to have transiently persistent strains than cases that do not use drugs, [OR: 2.3; 95% CI: 1.3-4.2]. Cases that were incarcerated were 5 times more likely to have transiently successful strains than non-incarcerated cases, [OR: 5.4; 95% CI: 2.0-14.9]. There was no significant relationship between strain persistence and race, homelessness, isoniazid resistance and country of birth when comparing transiently successful strains to unsuccessful strains.

Discussion

This study extended traditional cluster investigation by further analyzing 710 clustered isolates in Georgia from 2009 to 2015. We incorporated the concept of cluster density and time in assessing *Mycobacterium tuberculosis* strain persistence in Georgia during a 7-year period. Of the 710 isolates examined, 28% (198) were persistently successful. In our study, strain success or persistence was strongly associated with patient's country of origin. We found that persistently successful strains were more common among US-born cases. We have shown that clustering of isolates is more common among US-born cases based on our analysis in aim 2. However, this study shows that clusters of TB isolates that are persistently successful are almost entirely restricted to US-born cases. 95% of all persistently successful strains were isolated from US-born cases. This agrees with findings from a study assessing recent transmission of TB in the United States from 2011 to 2014 (Yuen et al., 2016). This study found a negative association between foreign birth and extensive recent transmission of TB (Yuen et al., 2016).

Homeless shelters are important aspects of TB transmission in the state of Georgia. Our study shows that resistant isolates are more likely to transmit in homeless shelters. Therefore, isolates from homeless cases that are resistant to isoniazid were more likely than other isolates to persist over time. An investigation of drug-resistant TB in Georgia found that from March 2008 to December 2009, there was an outbreak of TB in Georgia involving 1 homeless shelter (Powell et al., 2017). From January 2010 through December 2013 the number of cases from this outbreak waned until January 2014 through December 2015, when multiple homeless facilities were involved (Powell et al., 2017). Our analysis agrees with the findings from this study. In 2009 there were 15 cases involving persistently successful isolates resistant to isoniazid but this number decreased to 9 in 2013. In 2014, there, were 24 cases of TB with persistently successful strains and in 2015, there were 25 cases. The strain responsible for this outbreak is today the most predominant strain of *Mycobacterium tuberculosis* in Georgia. Even though acquiring resistance is said to decrease the virulence of strains, the fact that these resistant strains are common among homeless shelter residents increases the likelihood that such strains will circulate for years. No matter the effect of resistance on the virulence of the strain, constant and extended exposure may lead to transmission. Treatment adherence is also a huge problem among homeless cases which leads to even more resistant strains.

We found that persistently successful strains were positively associated with residence in a correctional facility. This study showed that TB in Georgia is concentrated among certain demographic populations, like people living in congregate settings like homeless shelters and correctional facilities. Among these groups of people, transmission of TB and resistance to isoniazid was common. Studies have shown that TB patients who abuse substances or are homeless often experience treatment failure and remain infectious longer because treatment failure presumable extends periods of infectiousness (Friedman, Williams, Singh, & Frieden, 1996; Reichman, Felton, & Edsall, 1979). Homeless persons, prisoners and substances abusers may have less access to routine medical care, potentially leading to delayed diagnosis. As the disease progresses, patients tend to be more contagious, thereby increasing probability of transmission of TB (Friedman et al., 1996).

In Georgia, the strains that cause the most disease are concentrated among vulnerable or under-served populations. Our study shows that a strain of TB first appearing among homeless persons or prison inmates can spread faster due to the close proximity of individuals and eventually become persistently successful. Efforts to reduce TB transmission in Georgia are based largely on contact investigations. Although contact investigations are an important component of the overall tuberculosis strategy, for many reasons, they are not conducted well even in successful TB control programs (Reichler et al., 2002; Weis, 2002). Treatment completion via direct observed treatment is also difficult for homeless persons and substance abusers. This eventually leads to drug resistant TB. The state of Georgia should provide standardized guidelines at each homeless shelter based on recommendations from the Centers of Disease Control and Prevention (CDC). These state guidelines should require TB screening for all residents and staff of homeless shelters to aid in early identification of suspected cases or active TB. This will ensure rapid evaluation of suspected cases by appropriate health care providers, and timely transportation to an appropriate health care facility if an evaluation cannot be done at the shelter. Shelter staff should be provided with the appropriate training and decision-making tools to identify new cases based on symptoms and previous records. In homeless shelters, the probability of transmission is affected by building ventilation, therefore recirculated air may contribute to transmission within a shelter. Homeless shelters should strive to have adequate ventilation supplemented by upper room germicidal ultraviolet air disinfection.

In this study, clustering by genotyping was not validated by contact tracing information and therefore we cannot be sure that all clustered isolates were in a chain of recent transmission of TB. However, since we further divided the clustered isolates based on cluster size and time, we can be sure that persistently successful strains were true clusters based on size and association with certain demographic and behavioral variables.

Table 6.1. Patient Characteristics and their association with strain persistence in Georgia, 2009-2015 (N=706)

Characteristics	Persistently Successful (n=198)	Transiently Successful (n=131)	Unsuccessful (n=377)	χ2	P-value
Age (Median, IQR[1])	48.0 (34.0-54.0)	47.5 (34.0-54.0)	48.0 (32.0-58.0)	0.90[2]	0.6370
Race				68.1	<. 0001
AI/AN[3]	0 (0)	0 (0)	4 (100.0)		
Asian	2 (3.2)	16 (25.8)	44 (71.0)		
Black	168 (36.0)	93 (19.9)	206 (44.1)		
White	28 (16.3)	21 (12.2)	123 (71.5)		
Ethnicity				34.1	<. 0001
Hispanic	4 (5.6)	6 (8.4)	61 (86.0)		
Non-Hispanic	194 (30.1)	125 (19.7)	316 (49.8)		
Gender				15.4	0.0005
Female	31 (16.9)	37 (20.2)	115 (62.8)		
Male	167 (31.9)	94 (18.0)	262 (50.1)		
US-Born				74.3	<. 0001
No	8 (5.2)	18 (11.7)	128 (83.1)		
Yes	190 (34.4)	113 (20.5)	249 (45.1)		
Alcohol Abuse				12.5	0.0143
No	137 (25.8)	95 (17.9)	300 (56.4)		
Unknown	0 (0.0)	1 (100.0)	0 (0.0)		
Yes	61(35.3)	35 (20.2)	77 (44.1)		
Homeless				78.7	<. 0001
No	113 (20.3)	115 (20.7)	329 (59.1)		
Yes	85 (57.1)	16 (10.8)	48 (32.2)		
Cavity				6.1	0.1890
No	119 (28.9)	71 (17.2)	222 (53.9)		
Unknown	0 (0)	0 (0)	4 (100.0)		
Yes	64 (26.5)	54 (22.3)	124 (51.2)		
INH[5] Resistance				224.3	<. 0001
Not Done	0 (0)	1 (50.0)	1 (50.0)		
Resistant	106 (80.9)	4 (3.1)	21 (16.0)		
Sensitive	92 (16.1)	126 (22.0)	355 (62.0)		
Drug Use (inj.)				11.3	0.0231

No	194 (27.9)	129 (18.6)	372 (53.5)		
Unknown	0 (0)	2 (100.0)	0 (0.0)		
Yes	4 (44.4)	0 (0)	5 (55.6)		
Drug Use (non-Inj.)				30.9	<.0001
No	149 (25.5)	97 (16.6)	338 (57.9)		
Unknown	0 (0)	1 (100.0)	0 (0.0)		
Yes	49 (40.5)	33 (27.3)	39 (32.2)		
Previous TB				3.7	0.1595
No	191 (28.8)	120 (18.1)	352 (53.1)		
Yes	7 (16.3)	11 (25.6)	25 (58.1)		
Prison				23.5	<.0001
No	177 (26.7)	116 (17.5)	369 (55.7)		
Yes	21 (47.7)	15 (34.1)	8 (18.2)		
LTCF[6]				2.8	0.2446
No	196 (30.2)	131 (18.8)	370 (53.1)		
Yes	2 (22.2)	0 (0)	7 (77.8)		
Sputum Smear				5.9	0.2065
Negative	61 (29.8)	30 (14.6)	114 (55.6)		
Not Done	3 (16.7)	2 (11.1)	13 (72.2)		
Positive	134 (27.7)	99 (20.5)	250 (51.8)		
Case Status				4.2	0.1222
Alive	197 (28.6)	128 (18.6)	378 (52.9)		
Dead	1 (6.3)	3 (18.8)	12 (75.0)		

[1] Interquartile range, [2] Kruskal Wallis test, [3] American Indian/Alaskan Native, [5] Isoniazid, [6] Long-term care facility

Table 6.2. Unadjusted odds Ratio and 95% Confidence Intervals from Multinomial Logistic Regression comparing unsuccessful and transiently successful strains to successful strains

Patient Characteristics	Successful Vs. Unsuccessful	Transient Vs. Unsuccessful
Race (All[1] vs. Black)	4.6 (2.9-7.1)*	2.1 (1.4-3.3)
Ethnicity (non-Hispanic vs. Hispanic)	0.1 (0.0-0.3)*	0.3 (0.1-0.6)*
Sex (Female vs. Male)	2.4 (1.5-3.7)*	1.1 (0.7-1.8)
Resist (Susceptible vs. Resistant)	19.5 (11.6-32.8)*	0.5 (0.2-1.6)
Drug (Non-injecting)	2.8 (1.8-4.5)*	2.8(1.6-4.7)*
Prison (No vs. Yes)	5.5 (2.4-12.6)*	6.1 (2.5-14.8)*
Country of Birth (Non-US vs. US)	12.3 (5.9-25.7)*	3.1 (1.8-5.4)*
Homelessness (No vs. Yes)	5.1 (3.4-7.8)*	1.0 (0.5-1.8)
Alcohol (No vs. Yes)	1.7 (1.2-2.6)*	1.4 (0.9-2.2)

[1] Includes Asian, American Indians and Whites, *Significant Confidence interval

Table 6.3. Adjusted odds Ratio and 95% Confidence Intervals from Multinomial Logistic Regression comparing unsuccessful and transiently successful strains to successful strains		
Patient Characteristics	**Successful vs. Unsuccessful**	**Transient vs. Unsuccessful**
Race (All[1] vs. Black)	2.7 (1.5-5.1)*	1.4 (0.8-2.2)
Ethnicity (non-Hispanic vs. Hispanic)	0.6 (0.1-2.5)	0.4 (0.1-1.3)
Sex (Female vs. Male)	1.5 (0.9-2.6)	1.2 (0.7-1.9)
Drug (Non-injecting)	1.1 (0.6-2.0)	2.3 (1.3-4.2)*
Prison (No vs. Yes).	5.5 (1.9-16.0)*	5.4 (2.0-14.9)*
Country of Birth (Non-US vs. US)	7.8 (2.9-20.9)*	1.9 (1.0-3.8)
Alcohol (No vs. Yes)	1.0 (0.6-1.7)	1.0 (0.6-1.7)
Resistance * Homelessness (resistance=1, homelessness=1)	33.2 (10.6-104.0)	0.6 (0.1-5.9)
Resistance * Homelessness (resistance=1, homelessness=0)	18.6 (8.4-41.4)	0.7 (0.2-2.7)

*[1] Includes Asian, American Indians and Whites, *Significant Confidence interval*

Appendix

Figure 6.1: Distribution of Mycobacterium tuberculosis strains isolated annually in Georgia, 2009-2015

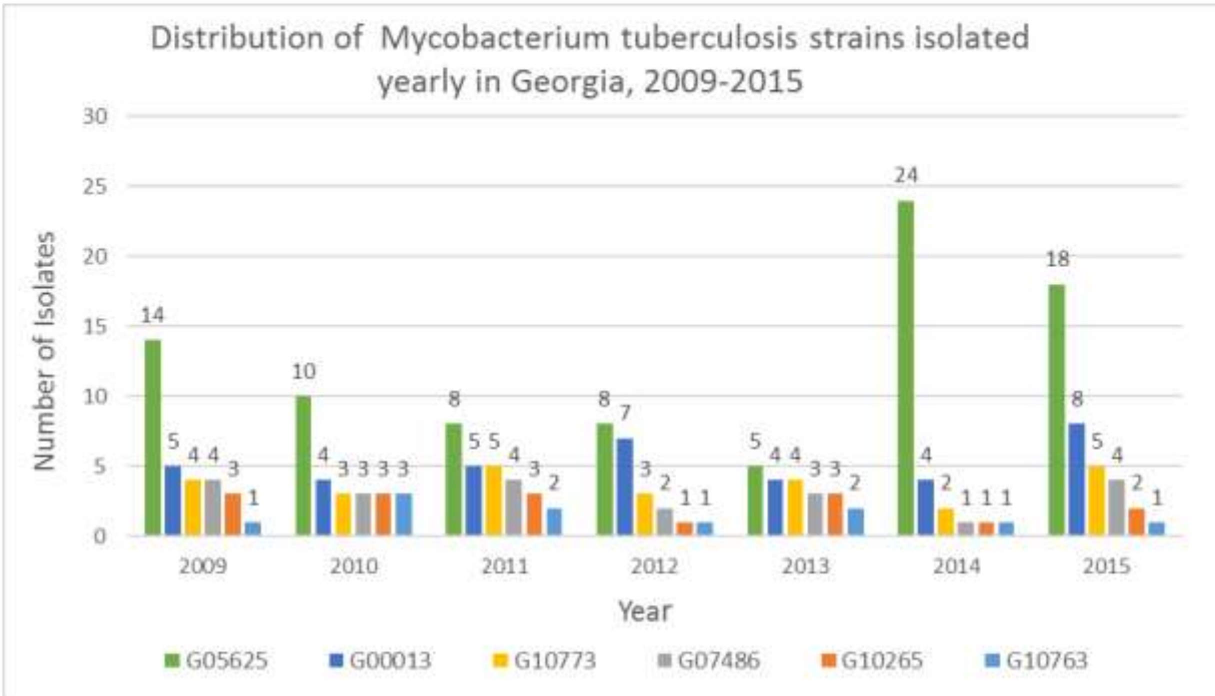
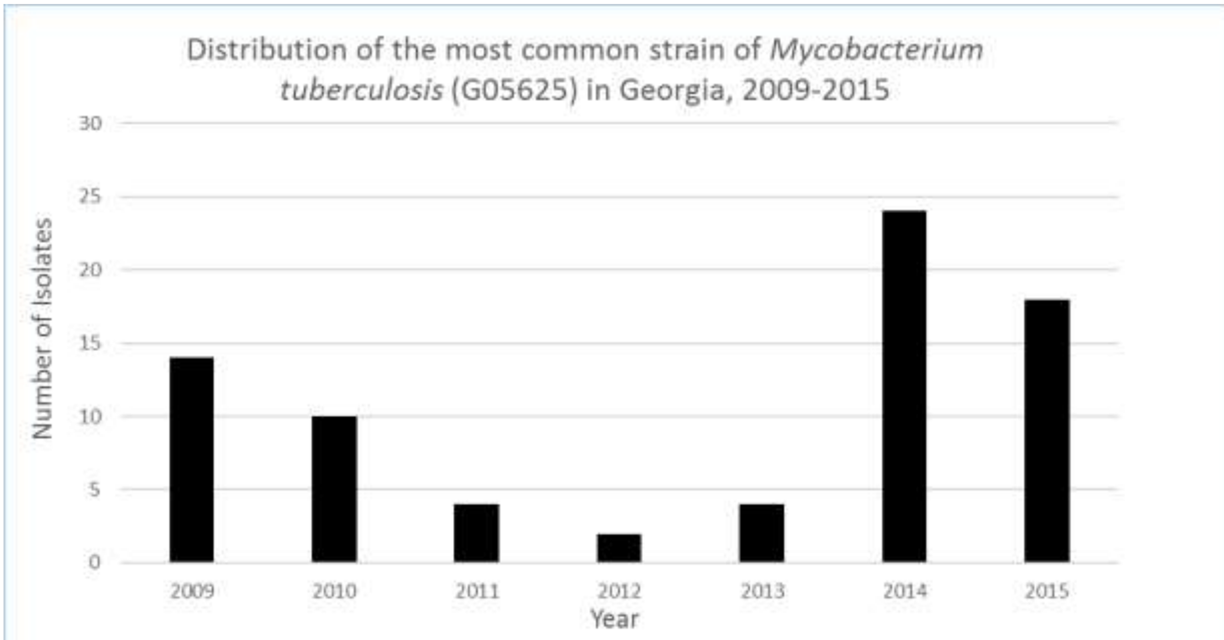


Figure 6.2: Distribution of the most common strain of Mycobacterium tuberculosis in Georgia, 2009-2015



References: Chapter 6

Alland, D., Kalkut, G. E., Moss, A. R., McAdam, R. A., Hahn, J. A., Bosworth, W., . . . Bloom, B. R. (1994). Transmission of tuberculosis in New York City--an analysis by DNA fingerprinting and conventional epidemiologic methods. *New England Journal of Medicine, 330*(24), 1710-1716.

Bamrah, S., Yelk Woodruff, R., Powell, K., Ghosh, S., Kammerer, J., & Haddad, M. (2013). Tuberculosis among the homeless, United States, 1994–2010. *The International Journal of Tuberculosis and Lung Disease, 17*(11), 1414-1419.

Barnes, P. F., & Cave, M. D. (2003). Molecular epidemiology of tuberculosis. *New England Journal of Medicine, 349*(12), 1149-1156.

Bauer, J., Yang, Z., Poulsen, S., & Andersen, Å. B. (1998). Results from 5 years of nationwide DNA fingerprinting of Mycobacterium tuberculosis complex isolates in a country with a low incidence of M. tuberculosis infection. *Journal of clinical microbiology, 36*(1), 305-308.

CDC. (2017). Reported Tuberculosis in the United States, 2016.

DPH. (2017). 2016 Georgia Tuberculosis Report.

Friedman, L. N., Williams, M. T., Singh, T. P., & Frieden, T. R. (1996). Tuberculosis, AIDS, and death among substance abusers on welfare in New York City. *New England Journal of Medicine, 334*(13), 828-833.

Glynn, J. R., Crampin, A. C., Traore, H., Chaguluka, S., Mwafulirwa, D. T., Alghamdi, S., . . . Fine, P. E. (2008). Determinants of cluster size in large, population-based molecular epidemiology study of tuberculosis, northern Malawi. *Emerging infectious diseases, 14*(7), 1060.

Kempf, M.-C., Dunlap, N. E., Lok, K. H., Benjamin, W. H., Keenan, N. B., & Kimerling, M. E. (2005). Long-term molecular analysis of tuberculosis strains in Alabama, a state characterized by a largely indigenous, low-risk population. *Journal of clinical microbiology, 43*(2), 870-878.

Lockman, S., Sheppard, J. D., Braden, C. R., Mwasekaga, M. J., Woodley, C. L., Kenyon, T. A., . . . Kesupile-Reed, M. (2001). Molecular and Conventional Epidemiology ofMycobacterium tuberculosis in Botswana: a Population-Based Prospective Study of 301 Pulmonary Tuberculosis Patients. *Journal of clinical microbiology, 39*(3), 1042-1047.

Mathema, B., Kurepina, N. E., Bifani, P. J., & Kreiswirth, B. N. (2006). Molecular epidemiology of tuberculosis: current insights. *Clinical microbiology reviews, 19*(4), 658-685.

Middelkoop, K., Bekker, L.-G., Mathema, B., Myer, L., Shashkina, E., Whitelaw, A., . . . Wood, R. (2014). Factors affecting tuberculosis strain success over 10 years in a high TB-and HIV-burdened community. *International journal of epidemiology, 43*(4), 1114-1122.

Mitruka, K. (2011). Tuberculosis Outbreak Investigations in the United States, 2002–2008-Volume 17, Number 3—March 2011-Emerging Infectious Disease journal-CDC.

Powell, K. M., VanderEnde, D. S., Holland, D. P., Haddad, M. B., Yarn, B., Yamin, A. S., . . . Burns-Grant, G. (2017). Outbreak of Drug-Resistant Mycobacterium tuberculosis Among Homeless People in Atlanta, Georgia, 2008-2015. *Public Health Reports, 132*(2), 231-240.

Reichler, M. R., Reves, R., Bur, S., Thompson, V., Mangura, B. T., Ford, J., . . . Onorato, I. M. (2002). Evaluation of investigations conducted to detect and prevent transmission of tuberculosis. *Jama, 287*(8), 991-995.

Reichman, L. B., Felton, C. P., & Edsall, J. R. (1979). Drug dependence, a possible new risk factor for tuberculosis disease. *Archives of internal medicine, 139*(3), 337-339.

Small, P. M., Hopewell, P. C., Singh, S. P., Paz, A., Parsonnet, J., Ruston, D. C., . . . Schoolnik, G. K. (1994). The Epidemiology of Tuberculosis in San Francisco--A Population-Based Study Using Conventional and Molecular Methods. *New England Journal of Medicine, 330*(24), 1703-1709.

Van der Spuy, G., Kremer, K., Ndabambi, S., Beyers, N., Dunbar, R., Marais, B., . . . Warren, R. (2009). Changing Mycobacterium tuberculosis population highlights clade-specific pathogenic characteristics. *Tuberculosis, 89*(2), 120-125.

van Soolingen, D., Borgdorff, M. W., de Haas, P. E., Sebek, M. M., Veen, J., Dessens, M., . . . van Embden, J. D. (1999). Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. *Journal of Infectious Diseases, 180*(3), 726-736.

Verver, S., Warren, R. M., Munch, Z., Vynnycky, E., van Helden, P. D., Richardson, M., . . . Behr, M. A. (2004). Transmission of tuberculosis in a high incidence urban community in South Africa. *International journal of epidemiology, 33*(2), 351-357.

Warren, R., Hauman, J., Beyers, N., Richardson, M., Schaaf, H., & Donald, P. (1996). Unexpectedly high strain diversity of Mycobacterium tuberculosis in a high-incidence community. *South African medical journal= Suid-Afrikaanse tydskrif vir geneeskunde, 86*(1), 45-49.

Weis, S. (2002). Contact investigations: how do they need to be designed for the 21st century? : Am Thoracic Soc.

WHO. (2017). Global tuberculosis report 2017.

Yuen, C. M., Kammerer, J. S., Marks, K., Navin, T. R., & France, A. M. (2016). Recent transmission of tuberculosis—United States, 2011–2014. *PLoS One, 11*(4), e0153728.

CHAPTER 7

SUMMARY AND CONCLUSION

Summary

The aim of this study was to examine the transmission dynamic of TB in Georgia by looking at the genetic diversity of *Mycobacterium tuberculosis* strains, extent of recent transmission and persistence of isolates. We found that majority of the TB isolates in Georgia belong to lineage 4. TB in Georgia is concentrated among US-born and non-Hispanic Blacks. Homeless shelters are the most likely hot spots for TB transmission in Georgia. Clustering of isolates was more common in homeless shelters and among US-born cases. Strains of *Mycobacterium tuberculosis* that have persisted in Georgia from 2009 to 2015 were mostly associated with homelessness, Black race, isoniazid resistance and residence in a correctional facility.

Aim 1: We examined the genetic diversity of *Mycobacterium tuberculosis* strains lineages and their associations with patient demographic and clinical characteristics in Georgia from 2009 to 2015. There were four lineages of *Mycobacterium tuberculosis* prevalent in Georgia. Lineage 4, otherwise known as the Euro-American lineage was the most common lineage, accounting for more than 70% of all isolates. Isolates from lineage 2 (East Asian) were more likely to cluster than isolates from lineage 4. We found a phylo-geographic relationship between *Mycobacterium tuberculosis* lineages and hosts. This relationship suggests that Mycobacterium tuberculosis strains have evolved to successfully infect certain populations. In Georgia, TB is more common among US-born cases and the fact lineage 4 is the most predominant lineage underscores this notion.

Aim 2: We examined the relationship between clustering of isolates and TB patient characteristics in Georgia from 2009 to 2015. We found that the 30% of the cases in our study were clustered or could be attributed to recent transmission. The rate of transmission was calculated to be 25% within a 2 year period. This is in agreement with other studies conducted in San Francisco, New York, Alabama and some countries in Europe and Africa (Alland et al., 1994; Bauer, Yang, Poulsen, & Andersen, 1998; Genewein et al., 1993; Kempf et al., 2005; Pfyffer et al., 1998; Small et al., 1994; Whalen et al., 2011; Wilkinson et al., 1997). Isolates that were resistant to Isoniazid were more likely to be clustered. Homelessness and Black race were also significant predictors of clustering. Clustering of isolates was more common among US-born than non-US-born cases. This underscores the idea that TB in non-US-born cases is usually associated with TB infection in their country of origin.

Aim 3: We analyzed data from 2009 to 2015 to examine persistence of Mycobacterium tuberculosis strains and the risk factors associated with the most persistent strains of M. tuberculosis in Georgia. The success of strains was strongly associated with patient's country of origin. We found that persistently successful strains were more common among US-born cases. We found that isolates from homeless cases that were resistant to isoniazid were more likely to be successful strains. Half of all persistently successful strains were resistant to Isoniazid compared to 4% of transiently successful strains and 5% of the unsuccessful strains. Congregate settings like prisons were also associated with persistently successful strains.

Interpretation of findings

There is not enough evidence that lineages of *Mycobacterium tuberculosis* have any effect on disease manifestation. However, the phylo-geographic relationship between strain lineages and host country of origin was apparent. A patient's race and country of origin can be good indicators of the strain of *Mycobacterium tuberculosis* they could be harboring. We saw a higher proportion of cases in lineage 4 because TB in Georgia is mainly as a result of recent transmission and recent transmission is more common among US-born cases. US-born cases make up the majority of the cases of TB and US-born cases usually have isolates from lineage 4. Since, TB in non-US-born cases is usually a result of reactivation of latent infection acquired in their country of origin, clustering of isolates across imported lineages like lineage 2 should be interpreted with caution.

TB transmission in Georgia is more concentrated in congregate settings and among vulnerable minority populations. Transmission of TB in the Georgia is more common among US-born cases than non-US born cases, and US-born cases were more likely to be homeless or incarcerated. Homeless shelters are considered hot spots in Georgia for the transmission of Isoniazid-resistant tuberculosis. The TB burden in Georgia is highest among Blacks and persons residing in homeless shelters. TB control efforts should focus on immediate diagnosis and treatments of cases in homeless shelters and congregate settings. Improving access to care among vulnerable populations will allow early detection and treatment of cases and consequently stopping transmission.

The most successful strains of TB in Georgia were resistant to isoniazid and were isolated from US-born homeless cases. Residence in a correctional facility and substance abuse were also indicators of successful strains. The most successful strains are found among vulnerable and underserved populations. Homeless persons, prisoners and substance abusers may have less access to routine medical care, potentially leading to delayed diagnosis. As disease progresses, patients tend to be more contagious, thereby increasing probability of transmission of TB. Adherence to treatment is also an issue among certain vulnerable populations like substance abusers and prison inmates. Therefore, strains that first appear among vulnerable populations like homeless and prison inmates are more likely to persist over time and become drug resistant. In Georgia, we see that congregate settings are important aspects in the epidemiology of TB and more efforts should be directed to improving the early diagnosis and effective treatment of cases.

Strengths and Weaknesses

This is the first study addressing the genetic diversity and the host-pathogen adaptation of Mycobacterium tuberculosis in Georgia. It is also the first study examining the transmission dynamic and persistence of strains in Georgia. We included 7 years' worth of TB data, thereby increasing our sample size to be able to capture any effect that may exist. We used a plausible source cases approach-a field validated method to divide isolates into clusters or non-clustered isolates. This study has provided a comprehensive description of the transmission dynamic of TB in Georgia from 2009 to 2015. This study has provided the basis for future research using more robust molecular methods. Despite the above strengths, there are a few limitations attached to this study. Not all cases in our study had molecular data available. We therefore had to exclude some cases from the analysis. This study was unable to examine the role of HIV/AIDS in the transmission dynamic of TB and in the association between lineages and patient characteristics.

Future studies should use whole genome sequencing to provide a clear delineation of strains. Future studies should also include HIV/AIDS data since HIV/AIDS according to studies, significantly affects the outcome of TB.

Conclusion

Although TB incidence is decreasing in Georgia, epidemiologic modeling by the CDC projects that the goal of TB elimination will not be attained in this century with the current rates of decline. Current control programs such as early identification of TB cases, completion of TB treatment by directly observed therapy, and contact investigation should be improved to accommodate high-risk individuals such as persons born in countries with a high prevalence of TB and persons who live or work in high-risk congregate settings. Since more than half of all TB cases are reported from only four counties, aggressive cases findings, contact investigations and treatment programs should be implemented in these counties.

The most urgent priority for controlling TB in Georgia is interrupting new transmission of Mycobacterium tuberculosis. The state of Georgia should focus on reducing the transmission of drug resistant TB in communities that have homeless shelters. Given that most resistant strains of TB in Georgia are isolated from cases residing in homeless shelters, early diagnosis and treatment of homeless cases will prove important in stopping transmission. Local public health staff should collaborate with homeless shelter and prison staff for staff training, information sharing and discharge planning. This will ensure that no cases of TB are left unmanaged or untreated.

References: Chapter 7

Alland, D., Kalkut, G. E., Moss, A. R., McAdam, R. A., Hahn, J. A., Bosworth, W., . . . Bloom, B. R. (1994). Transmission of tuberculosis in New York City--an analysis by DNA fingerprinting and conventional epidemiologic methods. *New England Journal of Medicine, 330*(24), 1710-1716.

Bauer, J., Yang, Z., Poulsen, S., & Andersen, Å. B. (1998). Results from 5 years of nationwide DNA fingerprinting of Mycobacterium tuberculosis complex isolates in a country with a low incidence of M. tuberculosis infection. *Journal of clinical microbiology, 36*(1), 305-308.

Genewein, A., Telenti, A., Bernasconi, C., Schopfer, K., Bodmer, T., Mordasini, C., . . . Rieder, H. (1993). Molecular approach to identifying route of transmission of tuberculosis in the community. *The Lancet, 342*(8875), 841-844.

Kempf, M.-C., Dunlap, N. E., Lok, K. H., Benjamin, W. H., Keenan, N. B., & Kimerling, M. E. (2005). Long-term molecular analysis of tuberculosis strains in Alabama, a state characterized by a largely indigenous, low-risk population. *Journal of clinical microbiology, 43*(2), 870-878.

Pfyffer, G., Strassle, A., Rose, N., Wirth, R., Brandli, O., & Shang, H. (1998). Transmission of tuberculosis in the metropolitan area of Zurich: a 3 year survey based on DNA fingerprinting. *European Respiratory Journal, 11*(4), 804-808.

Small, P. M., Hopewell, P. C., Singh, S. P., Paz, A., Parsonnet, J., Ruston, D. C., . . . Schoolnik, G. K. (1994). The Epidemiology of Tuberculosis in San Francisco--A Population-Based Study Using Conventional and Molecular Methods. *New England Journal of Medicine, 330*(24), 1703-1709.

Whalen, C. C., Zalwango, S., Chiunda, A., Malone, L., Eisenach, K., Joloba, M., . . . Mugerwa, R. (2011). Secondary attack rate of tuberculosis in urban households in Kampala, Uganda. *PLoS One, 6*(2), e16137.

Wilkinson, D., Pillay, M., Crump, J., Lombard, C., Davies, G. R., & Sturm, A. W. (1997). Molecular epidemiology and transmission dynamics of Mycobacterium tuberculosis in rural Africa. *Tropical Medicine & International Health, 2*(8), 747-753.

BIBLIOGRAPHY

1. WHO. Global tuberculosis report 2017. 2017.

2. CDC. Reported Tuberculosis in the United States, 2016. 2017.

3. DPH. 2016 Georgia Tuberculosis Report. 2017.

4. Bamrah S, Yelk Woodruff R, Powell K, Ghosh S, Kammerer J, Haddad M. Tuberculosis among the homeless, United States, 1994–2010. The International Journal of Tuberculosis and Lung Disease. 2013; 17 (11):1414-9.

5. Barnes PF, El-Hajj H, Preston-Martin S, Cave MD, Jones BE, Otaya M, et al. Transmission of tuberculosis among the urban homeless. Jama. 1996; 275(4):305-7.

6. Powell KM, VanderEnde DS, Holland DP, Haddad MB, Yarn B, Yamin AS, et al. Outbreak of Drug-Resistant Mycobacterium tuberculosis Among Homeless People in Atlanta, Georgia, 2008-2015. Public Health Reports. 2017; 132(2):231-40.

7. Jasmer RM, Hahn JA, Small PM, Daley CL, Behr MA, Moss AR, et al. A molecular epidemiologic analysis of tuberculosis trends in San Francisco, 1991–1997. Annals of Internal Medicine. 1999; 130(12):971-8.

8. Kempf M-C, Dunlap NE, Lok KH, Benjamin WH, Keenan NB, Kimerling ME. Long-term molecular analysis of tuberculosis strains in Alabama, a state characterized by a largely indigenous, low-risk population. Journal of clinical microbiology. 2005; 43(2):870-8.

9. Murray MB. Molecular epidemiology and the dynamics of tuberculosis transmission among foreign-born people. Canadian Medical Association Journal. 2002; 167(4):355-6.

10. Gagneux S. Host–pathogen coevolution in human tuberculosis. Phil Trans R Soc B. 2012; 367(1590):850-9.

11. Smith NH, Hewinson RG, Kremer K, Brosch R, Gordon SV. Myths and misconceptions: the origin and evolution of Mycobacterium tuberculosis. Nature Reviews Microbiology. 2009;7(7):537.

12. Nelson KE, Williams C. Infectious disease epidemiology: Jones & Bartlett Publishers; 2013.

13. Brites D, Gagneux S. Co-evolution of Mycobacterium tuberculosis and Homo sapiens. Immunological reviews. 2015; 264(1):6-24.

14. Comas I, Gagneux S. The past and future of tuberculosis research. PLoS pathogens. 2009;5(10):e1000600.

15. Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whittam TS, et al. Restricted structural gene polymorphism in the Mycobacterium tuberculosis complex indicates evolutionarily recent global dissemination. Proceedings of the National Academy of Sciences. 1997; 94(18):9869-74.

16. Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, Ruston DC, et al. The Epidemiology of Tuberculosis in San Francisco--A Population-Based Study Using Conventional and Molecular Methods. New England Journal of Medicine. 1994; 330(24):1703-9.

17. Borgdorff MW, Nagelkerke N, van Soolingen D, de Haas PE, Veen J, van Embden JD. Analysis of tuberculosis transmission between nationalities in the Netherlands in the period 1993–1995 using DNA fingerprinting. American journal of epidemiology. 1998; 147(2):187-95.

18. Barnes PF, Cave MD. Molecular epidemiology of tuberculosis. New England Journal of Medicine. 2003; 349(12):1149-56.

19. Mathema B, Kurepina NE, Bifani PJ, Kreiswirth BN. Molecular epidemiology of tuberculosis: current insights. Clinical microbiology reviews. 2006; 19(4):658-85.

20. Allix-Béguec C, Fauville-Dufaux M, Supply P. Three-year population-based evaluation of standardized mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing of Mycobacterium tuberculosis. Journal of clinical microbiology. 2008; 46(4):1398-406.

21. Kremer K, Au BKY, Yip PCW, Skuce R, Supply P, Kam KM, et al. Use of variable-number tandem-repeat typing to differentiate Mycobacterium tuberculosis Beijing family isolates from Hong Kong and comparison with IS6110 restriction fragment length polymorphism typing and spoligotyping. Journal of clinical microbiology. 2005; 43(1):314-20.

22. Hanekom M, Van Der Spuy G, van Pittius NG, McEvoy C, Hoek K, Ndabambi S, et al. Discordance between mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing and IS6110 restriction fragment length polymorphism genotyping for analysis of Mycobacterium tuberculosis Beijing strains in a setting of high incidence of tuberculosis. Journal of clinical microbiology. 2008; 46(10):3338-45.

23. Varma-Basil M, Kumar S, Arora J, Angrup A, Zozio T, Banavaliker JN, et al. Comparison of spoligotyping, mycobacterial interspersed repetitive units typing and IS6110-RFLP in a study of genotypic diversity of Mycobacterium tuberculosis in Delhi, North India. Memorias do Instituto Oswaldo Cruz. 2011; 106(5):524-35.

24. Roetzer A, Schuback S, Diel R, Gasau F, Ubben T, di Nauta A, et al. Evaluation of Mycobacterium tuberculosis typing methods in a 4-year study in Schleswig-Holstein, Northern Germany. Journal of clinical microbiology. 2011; 49(12):4173-8.

25. Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajoj SA, et al. Mycobacterium tuberculosis complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. BMC microbiology. 2006; 6(1):23.

26. Kamerbeek J, Schouls L, Kolk A, Van Agterveld M, Van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. Journal of clinical microbiology. 1997; 35(4):907-14.

27. Dale J, Brittain D, Cataldi AA, Cousins D, Crawford J, Driscoll J, et al. Spacer oligonucleotide typing of bacteria of the Mycobacterium tuberculosis complex: recommendations for standardised nomenclature [The Language of Our Science]. The International Journal of Tuberculosis and Lung Disease. 2001; 5(3):216-9.

28. Kremer K, Van Soolingen D, Frothingham R, Haas W, Hermans P, Martin C, et al. Comparison of methods based on different molecular epidemiological markers for typing of Mycobacterium tuberculosis complex strains: interlaboratory study of discriminatory power and reproducibility. Journal of clinical microbiology. 1999; 37(8):2607-18.

29. Frothingham R, Meeker-O'Connell WA. Genetic diversity in the Mycobacterium tuberculosis complex based on variable numbers of tandem DNA repeats. Microbiology. 1998; 144(5):1189-96.

30. Mazars E, Lesjean S, Banuls A-L, Gilbert M, Vincent V, Gicquel B, et al. High-resolution minisatellite-based typing as a portable approach to global analysis of Mycobacterium tuberculosis molecular epidemiology. Proceedings of the national academy of Sciences. 2001; 98(4):1901-6.

31. Supply P, Lesjean S, Savine E, Kremer K, Van Soolingen D, Locht C. Automated high-throughput genotyping for study of global epidemiology of Mycobacterium tuberculosis based on mycobacterial interspersed repetitive units. Journal of clinical microbiology. 2001; 39(10):3563-71.

32. Gagneux S, DeRiemer K, Van T, Kato-Maeda M, De Jong BC, Narayanan S, et al. Variable host–pathogen compatibility in Mycobacterium tuberculosis. Proceedings of the National academy of Sciences of the United States of America. 2006; 103(8):2869-73.

33. Comas I, Homolka S, Niemann S, Gagneux S. Genotyping of genetically monomorphic bacteria: DNA sequencing in Mycobacterium tuberculosis highlights the limitations of current methodologies. PloS one. 2009; 4(11):e7815.

34. Gutacker MM, Smoot JC, Migliaccio CAL, Ricklefs SM, Hua S, Cousins DV, et al. 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(4):1533-43.

35. Gagneux S, Small PM. Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development. The Lancet infectious diseases. 2007; 7(5):328-37.

36. Baker L, Brown T, Maiden MC, Drobniewski F. Silent nucleotide polymorphisms and a phylogeny for Mycobacterium tuberculosis. Emerging infectious diseases. 2004; 10(9):1568.

37. Hirsh AE, Tsolaki AG, DeRiemer K, Feldman MW, Small PM. Stable association between strains of Mycobacterium tuberculosis and their human host populations. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101(14):4871-6.

38. SUN JR, DOU HY, LEE SY, CHIUEH TS, LU JJ. Epidemiological studies of Beijing strains of Mycobacterium tuberculosis from Taipei and other Asian cities based on MIRU profiles. Apmis. 2011; 119(9):581-7.

39. Coscolla M, Gagneux S. Does M. tuberculosis genomic diversity explain disease diversity? Drug Discovery Today: Disease Mechanisms. 2010; 7(1):e43-e59.

40. Hanekom M, Van Pittius NG, McEvoy C, Victor T, Van Helden P, Warren R. Mycobacterium tuberculosis Beijing genotype: a template for success. Tuberculosis. 2011; 91(6):510-23.

41. Click ES, Moonan PK, Winston CA, Cowan LS, Oeltmann JE. Relationship between Mycobacterium tuberculosis phylogenetic lineage and clinical site of tuberculosis. Clinical Infectious Diseases. 2012; 54(2):211-9.

42. Caws M, Thwaites G, Dunstan S, Hawn TR, Lan NTN, Thuong NTT, et al. The influence of host and bacterial genotype on the development of disseminated disease with Mycobacterium tuberculosis. PLoS pathogens. 2008; 4(3):e1000034.

43. De Jong BC, Hill PC, Aiken A, Awine T, Martin A, Adetifa IM, et al. Progression to active tuberculosis, but not transmission, varies by Mycobacterium tuberculosis lineage in The Gambia. Journal of Infectious Diseases. 2008; 198(7):1037-43.

44. Rajapaksa U, Perera A. Sublineages of Beijing strain of Mycobacterium tuberculosis in Sri Lanka. Indian journal of microbiology. 2011; 51(3):410-2.

45. Iwamoto T, Yoshida S, Suzuki K, Wada T. Population structure analysis of the Mycobacterium tuberculosis Beijing family indicates an association between certain sublineages and multidrug resistance. Antimicrobial agents and chemotherapy. 2008; 52(10):3805-9.

46. Lasunskaia E, Ribeiro SC, Manicheva O, Gomes LL, Suffys PN, Mokrousov I, et al. Emerging multidrug resistant Mycobacterium tuberculosis strains of the Beijing genotype circulating in Russia express a pattern of biological properties associated with enhanced virulence. Microbes and Infection. 2010;12(6):467-75.

47. Hanekom M, Mata D, van Pittius NG, van Helden P, Warren R, Hernandez-Pando R. Mycobacterium tuberculosis strains with the Beijing genotype demonstrate variability in virulence associated with transmission. Tuberculosis. 2010; 90(5):319-25.

48. Kato-Maeda M, Shanley CA, Ackart D, Jarlsberg LG, Shang S, Obregon-Henao A, et al. Beijing sublineages of Mycobacterium tuberculosis differ in pathogenicity in the guinea pig. Clinical and vaccine immunology. 2012; 19(8):1227-37.

49. Coscolla M, Gagneux S, editors. Consequences of genomic diversity in Mycobacterium tuberculosis. Seminars in immunology; 2014: Elsevier.

50. Parwati I, van Crevel R, van Soolingen D. Possible underlying mechanisms for successful emergence of the Mycobacterium tuberculosis Beijing genotype strains. The Lancet infectious diseases. 2010; 10(2):103-11.

51. Nicol MP, Wilkinson RJ. The clinical consequences of strain diversity in Mycobacterium tuberculosis. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2008; 102(10):955-65.

52. Borrell S, Gagneux S. Infectiousness, reproductive fitness and evolution of drug-resistant Mycobacterium tuberculosis [State of the art]. The International Journal of Tuberculosis and Lung Disease. 2009; 13(12):1456-66.

53. Pang Y, Zhou Y, Zhao B, Liu G, Jiang G, Xia H, et al. Spoligotyping and drug resistance analysis of Mycobacterium tuberculosis strains from national survey in China. PLoS One. 2012; 7(3):e32976.

54. Fenner L, Egger M, Bodmer T, Altpeter E, Zwahlen M, Jaton K, et al. Effect of mutation and genetic background on drug resistance in Mycobacterium tuberculosis. Antimicrobial agents and chemotherapy. 2012; 56(6):3047-53.

55. Malla B, Stucki D, Borrell S, Feldmann J, Maharjan B, Shrestha B, et al. First insights into the phylogenetic diversity of Mycobacterium tuberculosis in Nepal. PLoS One. 2012; 7(12):e52297.

56. Reed MB, Pichler VK, McIntosh F, Mattia A, Fallow A, Masala S, et al. Major Mycobacterium tuberculosis lineages associate with patient country of origin. Journal of clinical microbiology. 2009; 47(4):1119-28.

57. Fenner L, Egger M, Bodmer T, Furrer H, Ballif M, Battegay M, et al. HIV infection disrupts the sympatric host–pathogen relationship in human tuberculosis. PLoS Genet. 2013; 9(3):e1003318.

58. Kaufmann SH. How can immunology contribute to the control of tuberculosis? Nature Reviews Immunology. 2001; 1(1):20-30.

59. Comstock GW. FROST REVISITED: THE MODERN EPIDEMIOLOGY OF TUBERCULOSIS THE THIRD WADE HAMPTON FROST LECTURE. American journal of epidemiology. 1975; 101(5):363-82.

60. CDC. Targeted tuberculin testing and treatment of latent tuberculosis infection. Am J Respir Crit Care Med. 2000; 161:S221-S47.

61. Frieden TR, Sherman LF, Maw KL, Fujiwara PI, Crawford JT, Nivin B, et al. A multi-institutional outbreak of highly drug-resistant tuberculosis: epidemiology and clinical outcomes. Jama. 1996; 276(15):1229-35.

62. Bifani PJ, Mathema B, Liu Z, Moghazeh SL, Shopsin B, Tempalski B, et al. Identification of a W variant outbreak of Mycobacterium tuberculosis via population-based molecular epidemiology. Jama. 1999; 282(24):2321-7.

63. Bishai WR, Graham NM, Harrington S, Pope DS, Hooper N, Astemborski J, et al. Molecular and geographic patterns of tuberculosis transmission after 15 years of directly observed therapy. Jama. 1998; 280(19):1679-84.

64. Mathema B, Bifani PJ, Driscoll J, Steinlein L, Kurepina N, Moghazeh SL, et al. Identification and evolution of an IS 6110 low-copy-number Mycobacterium tuberculosis cluster. The Journal of infectious diseases. 2002; 185(5):641-9.

65. Valway SE, Sanchez MPC, Shinnick TF, Orme I, Agerton T, Hoy D, et al. An outbreak involving extensive transmission of a virulent strain of Mycobacterium tuberculosis. New England Journal of Medicine. 1998; 338(10):633-9.

66. Yaganehdoost A, Graviss EA, Ross MW, Adams GJ, Ramaswamy S, Wanger A, et al. Complex transmission dynamics of clonally related virulent Mycobacterium tuberculosis associated with barhopping by predominantly human immunodeficiency virus-positive gay men. The Journal of infectious diseases. 1999; 180(4):1245-51.

67. Alland D, Kalkut GE, Moss AR, McAdam RA, Hahn JA, Bosworth W, 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(24):1710-6.

68. Yuen CM, Kammerer JS, Marks K, Navin TR, France AM. Recent transmission of tuberculosis—United States, 2011–2014. PloS one. 2016; 11(4):e0153728.

69. Rodwell TC, Kapasi AJ, Barnes RF, Moser KS. Factors associated with genotype clustering of Mycobacterium tuberculosis isolates in an ethnically diverse region of southern California, United States. Infection, Genetics and Evolution. 2012; 12(8):1917-25.

70. Barnes PF, Yang Z, Preston-Martin S, Pogoda JM, Jones BE, Otaya M, et al. Patterns of tuberculosis transmission in Central Los Angeles. Jama. 1997; 278(14):1159-63.

71. Weis SE, Pogoda JM, Yang Z, Cave MD, Wallace C, Kelley M, et al. Transmission dynamics of tuberculosis in Tarrant County, Texas. American journal of respiratory and critical care medicine. 2002; 166(1):36-42.

72. Guwatudde D, Nakakeeto M, Jones-Lopez E, Maganda A, Chiunda A, Mugerwa R, et al. Tuberculosis in household contacts of infectious cases in Kampala, Uganda. American journal of epidemiology. 2003; 158(9):887-98.

73. Whalen CC, Zalwango S, Chiunda A, Malone L, Eisenach K, Joloba M, et al. Secondary attack rate of tuberculosis in urban households in Kampala, Uganda. PloS one. 2011; 6(2):e16137.

74. Caminero JA, Pena MJ, Campos-Herrero MI, Rodriguez JC, Garcia I, Cabrera P, et al. Epidemiological evidence of the spread of a Mycobacterium tuberculosis strain of the Beijing genotype on Gran Canaria Island. American journal of respiratory and critical care medicine. 2001; 164(7):1165-70.

75. Dye C, Williams BG. Slow elimination of multidrug-resistant tuberculosis. Science translational medicine. 2009; 1(3):3ra8-3ra8.

76. Cohen T, Murray M. Modeling epidemics of multidrug-resistant M. tuberculosis of heterogeneous fitness. Nature medicine. 2004; 10(10):1117.

77. Glynn JR, Crampin AC, Traore H, Chaguluka S, Mwafulirwa DT, Alghamdi S, et al. Determinants of cluster size in large, population-based molecular epidemiology study of tuberculosis, northern Malawi. Emerging infectious diseases. 2008; 14(7):1060.

78. Middelkoop K, Bekker L-G, Mathema B, Myer L, Shashkina E, Whitelaw A, et al. Factors affecting tuberculosis strain success over 10 years in a high TB-and HIV-burdened community. International journal of epidemiology. 2014; 43(4):1114-22.

79. Giordano TP, Soini H, Teeter LD, Adams GJ, Musser JM, Graviss EA. Relating the size of molecularly defined clusters of tuberculosis to the duration of symptoms. Clinical infectious diseases. 2004; 38(1):10-6.

80. Lillebaek T, Dirksen A, Kok-Jensen A, Andersen Å. A dominant Mycobacterium tuberculosis strain emerging in Denmark. The International Journal of Tuberculosis and Lung Disease. 2004; 8(8):1001-6.

81. Espinal MA. The global situation of MDR-TB. Tuberculosis. 2003; 83(1):44-51.

82. David HL. Resistance to D-cycloserine in the tubercle bacilli: mutation rate and transport of alanine in parental cells and drug-resistant mutants. Applied microbiology. 1971; 21(5):888-92.

83. David HL, Newman CM. Some Observations on the Genetics of Isoniazid Resistance in the Tubercle Bacilli 1, 2. American Review of Respiratory Disease. 1971; 104(4):508-15.

84. Councils BMR. Controlled clinical trial of four short-course (6-month) regimens of chemotherapy for treatment of pulmonary tuberculosis: second report. The Lancet. 1973; 301(7816):1331-9.

85. Kochi A, Vareldzis B, Styblo K. Multidrug-resistant tuberculosis and its control. Research in microbiology. 1993; 144(2):104-10.

86. Bass Jr JB, Farer LS, Hopewell PC, O'Brien R, Jacobs R, Ruben F, et al. Treatment of tuberculosis and tuberculosis infection in adults and children. American Thoracic Society and the Centers for Disease Control and Prevention. American journal of respiratory and critical care medicine. 1994; 149(5):1359-74.

87. Davies J, editor Antibiotic resistance in mycobacteria. Genetics and Tuberculosis: Novartis Foundation Symposium 217; 1998: Wiley Online Library.

88. Vareldzis BP, Grosset J, De Kantor I, Crofton J, Laszlo A, Felten M, et al. Drug-resistant tuberculosis: laboratory issues: World Health Organization recommendations. Tubercle and lung disease. 1994; 75(1):1-7.

89. Iseman M, Madsen L. Drug-resistant tuberculosis. Clinics in chest medicine. 1989; 10(3):341-53.

90. Bifani PJ, Plikaytis BB, Kapur V, Stockbauer K, Pan X, Lutfey ML, et al. Origin and interstate spread of a New York City multidrug-resistant Mycobacterium tuberculosis clone family. Jama. 1996; 275(6):452-7.

91. Crofton SJ, Chaulet P, Maher D, Grosset J, Harris W, Horne N, et al. Guidelines for the management of drug-resistant tuberculosis. 1997.

92. Census B. State Quickfacts: Georgia 2016 [updated 03/20/2017. Available from: https://www.census.gov/quickfacts/table/PST045216/13.

93. France AM, Grant J, Kammerer JS, Navin TR. A field-validated approach using surveillance and genotyping data to estimate tuberculosis attributable to recent transmission in the United States. American journal of epidemiology. 2015; 182(9):799-807.

94. Cole S, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature. 1998; 393(6685):537.

95. Driscoll JR. Spoligotyping for molecular epidemiology of the Mycobacterium tuberculosis complex. Molecular Epidemiology of Microorganisms: Springer; 2009. p. 117-28.

96. Christianson S, Wolfe J, Orr P, Karlowsky J, Levett PN, Horsman GB, et al. Evaluation of 24 locus MIRU-VNTR genotyping of Mycobacterium tuberculosis isolates in Canada. Tuberculosis. 2010; 90(1):31-8.

97. Jagielski T, Van Ingen J, Rastogi N, Dziadek J, Mazur PK, Bielecki J. Current methods in the molecular typing of Mycobacterium tuberculosis and other mycobacteria. BioMed research international. 2014; 2014.

98. Crawford J. Genotyping in contact investigations: a CDC perspective. The International Journal of Tuberculosis and Lung Disease. 2003; 7(12):S453-S7.

99. Kremer K, Arnold C, Cataldi A, Gutiérrez MC, Haas WH, Panaiotov S, et al. Discriminatory power and reproducibility of novel DNA typing methods for Mycobacterium tuberculosis complex strains. Journal of clinical microbiology. 2005; 43(11):5628-38.

100. Vynnycky E, Fine P. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. Epidemiology & Infection. 1997; 119(2):183-201.

101. Vynnycky E, Nagelkerke N, Borgdorff M, Van Soolingen D, Van Embden J, Fine P. 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 & Infection. 2001; 126(1):43-62.

102. Barnes PF, El-Hajj H, Preston-Martin S, Cave MD, Jones BE, Otaya M, et al. Transmission of tuberculosis among the urban homeless. Journal of the American Medical Association. 1996; 275(4):305-7.

103. Bauer J, Yang Z, Poulsen S, Andersen ÅB. Results from 5 years of nationwide DNA fingerprinting of Mycobacterium tuberculosis complex isolates in a country with a low incidence of M. tuberculosis infection. Journal of clinical microbiology. 1998; 36(1):305-8.

104. van Soolingen D, Borgdorff MW, de Haas PE, Sebek MM, Veen J, Dessens M, et al. Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. Journal of Infectious Diseases. 1999; 180(3):726-36.

105. McConkey SJ, Williams M, Weiss D, Adams H, Cave MD, Yang Z, 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(5):612-9.

106. Wharton M, Chorba TL, Vogt RL, Morse DL, Buehler JW. Case definitions for public health surveillance. 1990.

107. Maes M, Kremer K, van Soolingen D, Takiff H, de Waard JH. 24-locus MIRU-VNTR genotyping is a useful tool to study the molecular epidemiology of tuberculosis among Warao Amerindians in Venezuela. Tuberculosis. 2008; 88(5):490-4.

108. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, Willery E, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. Journal of clinical microbiology. 2006; 44(12):4498-510.

109. Pfyffer G, Strassle A, Rose N, Wirth R, Brandli O, Shang H. Transmission of tuberculosis in the metropolitan area of Zurich: a 3 year survey based on DNA fingerprinting. European Respiratory Journal. 1998; 11(4):804-8.

110. Genewein A, Telenti A, Bernasconi C, Schopfer K, Bodmer T, Mordasini C, et al. Molecular approach to identifying route of transmission of tuberculosis in the community. The Lancet. 1993; 342(8875):841-4.

111. Wilkinson D, Pillay M, Crump J, Lombard C, Davies GR, Sturm AW. Molecular epidemiology and transmission dynamics of Mycobacterium tuberculosis in rural Africa. Tropical Medicine & International Health. 1997; 2(8):747-53.

112. Ricks PM, Cain KP, Oeltmann JE, Kammerer JS, Moonan PK. Estimating the burden of tuberculosis among foreign-born persons acquired prior to entering the US, 2005–2009. PLoS One. 2011; 6(11):e27405.

113. Frieden TR, Fujiwara PI, Washko RM, Hamburg MA. Tuberculosis in New York City—turning the tide. New England Journal of Medicine. 1995; 333(4):229-33.

114. Snider Jr DE, Kelly GD, Cauthen GM, Thompson NJ, Kilburn JO. Infection and disease among contacts of tuberculosis cases with drug-resistant and drug-susceptible bacilli. American Review of Respiratory Disease. 1985; 132(1):125-32.

115. Beck-Sagué C, Dooley SW, Hutton MD, Otten J, Breeden A, Crawford JT, et al. Hospital outbreak of multidrug-resistant Mycobacterium tuberculosis infections: factors in transmission to staff and HIV-infected patients. Jama. 1992; 268(10):1280-6.

116. Anderson C, Story A, Brown T, Drobniewski F, Abubakar I. Tuberculosis in UK prisoners: a challenge for control. Journal of epidemiology and community health. 2010: jech. 2009.094375.

117. Curtis AB, Ridzon R, Novick L, Driscoll J, Blair D, Oxtoby M, et al. Analysis of Mycobacterium tuberculosis transmission patterns in a homeless shelter outbreak. The International Journal of Tuberculosis and Lung Disease. 2000; 4(4):308-13.

118. Zolopa AR, Hahn JA, Gorter R, Miranda J, Wlodarczyk D, Peterson J, et al. HIV and tuberculosis infection in San Francisco's homeless adults: prevalence and risk factors in a representative sample. Jama. 1994; 272(6):455-61.

119. CDC. Tuberculosis outbreaks in prison housing units for HIV-infected inmates--California, 1995-1996. MMWR Morbidity and mortality weekly report. 1999; 48(4):79.

120. McElroy PD, Fortenberry ER, Levine EC, Diem LA, Woodley CL, Williams PM, et al. Outbreak of tuberculosis among homeless persons coinfected with human immunodeficiency virus. Clinical infectious diseases. 2003; 36(10):1305-12.

121. Mitruka K. Tuberculosis Outbreak Investigations in the United States, 2002–2008-Volume 17, Number 3—March 2011-Emerging Infectious Disease journal-CDC. 2011.

122. Lockman S, Sheppard JD, Braden CR, Mwasekaga MJ, Woodley CL, Kenyon TA, et al. Molecular and Conventional Epidemiology ofMycobacterium tuberculosis in Botswana: a Population-Based Prospective Study of 301 Pulmonary Tuberculosis Patients. Journal of Clinical Microbiology. 2001; 39(3):1042-7.

123. Van der Spuy G, Kremer K, Ndabambi S, Beyers N, Dunbar R, Marais B, et al. Changing Mycobacterium tuberculosis population highlights clade-specific pathogenic characteristics. Tuberculosis. 2009; 89(2):120-5.

124. Verver S, Warren RM, Munch Z, Vynnycky E, van Helden PD, Richardson M, et al. Transmission of tuberculosis in a high incidence urban community in South Africa. International journal of epidemiology. 2004; 33(2):351-7.

125. Warren R, Hauman J, Beyers N, Richardson M, Schaaf H, Donald P. Unexpectedly high strain diversity of Mycobacterium tuberculosis in a high-incidence

community. South African medical journal= Suid-Afrikaanse tydskrif vir geneeskunde. 1996; 86(1):45-9.

126. Reichman LB, Felton CP, Edsall JR. Drug dependence, a possible new risk factor for tuberculosis disease. Archives of Internal Medicine. 1979; 139(3):337-9.

127. Friedman LN, Williams MT, Singh TP, Frieden TR. Tuberculosis, AIDS, and death among substance abusers on welfare in New York City. New England Journal of Medicine. 1996; 334(13):828-33.

128. Weis S. Contact investigations: how do they need to be designed for the 21st century? : Am Thoracic Soc; 2002.

129. Reichler MR, Reves R, Bur S, Thompson V, Mangura BT, Ford J, et al. Evaluation of investigations conducted to detect and prevent transmission of tuberculosis. Jama. 2002; 287(8):991-5.