Infection, Genetics and Evolution 16 (2013) 362–368

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Current trends of *Mycobacterium tuberculosis* molecular epidemiology in Saudi Arabia and associated demographical factors

Sahal Al-Hajoj [a,*], Bright Varghese [a], Farah Al-Habobe [b], Mohammed M. Shoukri [c], Arnout Mulder [d], Dick van Soolingen [d,e]

[a] *Mycobacteriology Research Section, Department of Infection and Immunity, King Faisal Specialist Hospital and Research Centre, Riyadh 11211, Saudi Arabia*
[b] *Department of Medical Microbiology, Utrecht University, Utrecht, The Netherlands*
[c] *National Biotechnology Centre, King Faisal Specialist Hospital and Research Centre, Riyadh 11211, Saudi Arabia*
[d] *Tuberculosis Reference Laboratory, National Institute for Public Health and the Environment (RIVM), P.O. Box 1, 3720 BA Bilthoven, The Netherlands*
[e] *Department of Pulmonary Diseases/Department of Clinical Microbiology, Radboud University Nijmegen Medical Centre, P.O. Box 9101, Nijmegen, The Netherlands*

ARTICLE INFO

Article history:
Received 3 December 2012
Received in revised form 20 February 2013
Accepted 5 March 2013
Available online 22 March 2013

Keywords:
Tuberculosis
Saudi Arabia
Population dynamics
Phylogenetic diversity
Genotyping

ABSTRACT

Data are scarce on demographical factors related to the population structure of *Mycobacterium tuberculosis* in Saudi Arabia. A study was conducted on 902 clinical isolates to explore current trends in the phylogeography and associated demographical factors of tuberculosis by using spoligotyping and 24 loci based MIRU-VNTR typing. Young male patients (aged 16–29 and 30–44) were predominant in this cohort. The phylogenetic diversity among *M. tuberculosis* isolates was found high, as almost all known genetic lineages were identified. Delhi/CAS (26.4%), EAI (13.7%) and Haarlem (11.3%) were the most common lineages observed, particularly among the low age groups (16–29 and 30–44 years), whereas elderly patients (>60 years) showed a predominance in the lineages S, Ghana, TUR and Uganda-I. A statistically significant association was observed between gender of the patients and lineages of EAI (p value 0.026) and LAM (p value 0.005). Overall, molecular strain cluster rate was 34.4% with an elevated rate among patients aged below 15 years (43.1%), while cases among the elderly (>60 years) showed the lowest degree of clustering (12.5%). The largest level of clustering was noticed among cases caused by strains of the lineages Haarlem (59.8%), Beijing (55.8%) and LAM (42.8%). The current population structure of *M. tuberculosis* in Saudi Arabia is highly diverse with significant associations to demography, transmission dynamics and origin of the patients. The difference in genotype distributions among low and high aged patients reflects the ongoing change in the strain population structure in the country.

1. Introduction

Tuberculosis (TB) is one of the most ancient infectious diseases of mankind caused by *Mycobacterium tuberculosis*. Molecular studies indicate that the history of this disease, affecting animals and humans, dates back at least 17,000 years (Rothschild et al., 2001). Despite the existence of anti-tuberculosis drugs and mass application of the bacillus Calmette-Guerin (BCG) vaccination in the last six decades, TB continues to be a major threat worldwide with a high morbidity and mortality.

Saudi Arabia is the third biggest Arab country and has a moderate rate of new TB cases (24/100,000 population/year) according to a recent report (World Health Organization, 2012). Despite the improved living standards for the population as a whole, availability of free anti-tuberculosis medication, and the application of BCG vaccination at birth, tuberculosis remains as an endemic disease in this country and may even run out of control in the coming years. Many factors may contribute to the above mentioned situation as Saudi Arabia has a highly diverse human population with 8.4 million expatriates among the 26 million inhabitants. Moreover, most foreign-born originate from South East Asia, the Indian sub continent, and countries in the West and Horn of Africa where TB is considered to be pandemic according to WHO reports (World Health Organization, 2012). In addition to the massive number of expatriates' living in the country, Saudi Arabia hosts approximately 10 million visitors per year for the Islamic rituals. This huge gathering of peoples from around the globe has a significant effect on transmission of TB, as shown in previous studies (Alzeer, 2009; Wilder-Smith et al., 2005; Al-Jasser et al., 2012; Alzahrani et al., 2012). Recent molecular epidemiological studies revealed that Saudi Arabia harbors a highly diverse population structure of *M.*

---

* Corresponding author. Address: Mycobacteriology Research Section, Department of Infection and Immunity, King Faisal Specialist Hospital and Research Centre, P.O. Box 3354, Riyadh 11211, Saudi Arabia. Tel.: +966 1 4424992; fax: +966 1 4627858.

E-mail addresses: hajoj@kfshrc.edu.sa (S. Al-Hajoj), bvarghese99@kfshrc.edu.sa (B. Varghese), farah-88@hotmail.com (F. Al-Habobe), shoukri@kfshrc.edu.sa (M.M. Shoukri), Dick.van.Soolingen@rivm.nl (D. van Soolingen).

1567-1348/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2013.03.019

tuberculosis, including most major genotype families. Furthermore, the noticed high strain clustering rate may reflect a high degree of ongoing transmission in the country (Al-Hajoj et al., 2007; Varghese et al., 2012, 2013).

Molecular characterization of *M. tuberculosis* integrated with the clinical and epidemiological information creates an important study platform to investigate the possible associations of strains and (sub) lineages with clinical and epidemiological characteristics of the disease (Dale et al., 2005; Kong et al., 2007; Nava-Aguilera et al., 2011). Most molecular epidemiological studies so far showed significant associations between the demography of the patients, clustering of strains, clinical manifestations and drug resistance, even though some of the observations could not be readily explained (Fok et al., 2008; Ramazanzadeh et al., 2006; Dolzani et al., 2004; Rhee et al., 1999).

Nevertheless, the current TB control strategy and transmission models are developed based on the assumption that all *M. tuberculosis* clinical strains are equal in virulence, correlation with resistance and transmissibility (Sreevatsan et al., 1997). However, there are strong indications that the genetic lineages have important differences in biological features, and this may has a major influence on the epidemiology of TB, as previously described for the Beijing genotype of *M. tuberculosis* (Parwati et al., 2010). Thus, a better understanding of the clinical and epidemiological relevance of *M. tuberculosis* lineages and their dynamics may facilitate improved strategies for TB control. Such kind of studies are so far lacking in Saudi Arabia. In this study we sought to evaluate the molecular epidemiology of tuberculosis in Saudi Arabia, based on a representative sample collection and taking the demography of patients into consideration to determine the impact of different age groups, gender and infection type to strain lineages of *M. tuberculosis* among the newly diagnosed TB cases.

2. Materials and methods

2.1. Study population

The study utilized a representative isolate collection retrieved in the period July 2009–June 2010, from all the regional TB reference laboratories in Saudi Arabia. The national tuberculosis control program of the country depends mainly on these reference laboratories for diagnosis of suspected TB. Cases diagnosed with positive cultures were selected for the study, regardless of the type of infection and smear results. As the study aimed to evaluate the current molecular epidemiology of TB, only new cases, defined as cases with a treatment history of less than one month, were enrolled. The collection of isolates represents 35% of all the reported new cases in Saudi Arabia in the study period. However, the definition of new cases depended on available medical records and patients self statements. Patient demographical data were collected by using standard data collection form and the completeness of information was cross-checked with the national tuberculosis registry maintained by the Ministry of Health.

2.2. Laboratory methods

Genomic DNA from heat killed *M. tuberculosis* isolates was prepared by using standard spin-column technique [QIAamp DNA mini kit, QIAGEN, GmbH, Germany] according to manufacturer's protocol. The spoligotyping was performed by using commercially available membranes [Ocimum Biosolutions, Hyderabad, India] according to the protocol described previously (Kamerbeek et al., 1997). The 24 loci based MIRU-VNTR typing was carried out using a commercial kit [Genoscreen, Lille, France] taking advantage of the 48 capillaries ABI 3730 genetic analyzer (Applied Biosystems, CA, USA) and as described previously (Supply et al., 2006).

2.3. Data analysis

We used four international online database tools to identify genetic lineages among *M. tuberculosis* isolates on basis of both spoligo and MIRU VNTR typing results. Primarily we analyzed all the data using the MIRU database (www.miru-vntrplus.org) as previously described (Weniger et al., 2010). Fingerprints which showed no matching results in this database were re-analyzed for their phylogeny by using other established databases, SpolDB-4, SITVIT web and the TB lineage identification tool of centre for disease control, USA- TB Insight/TB lineage tool (Demay et al., 2012; Shabbeer et al., 2012; Aminian et al., 2010; Brudey et al., 2006). Identified lineages were divided into three principal genetic groups (PGG-1, PGG-2, and PGG-3) as described previously (Sreevatsan et al., 1997). The PGG-1 group consisted of the ancestral strains, Delhi/CAS, EAI and Beijing. PGG-2 and included the major lineages LAM and Haarlem, together with Cameroon, TUR, URAL, X, S and New-I. PGG-3 mainly consisted of the lineages belonging to the T clades, Ghana, Uganda-I and Uganda-II. *Mycobacterium bovis* is related to the PGG-1, but was analyzed separately. A dendogram based on VNTR and spoligotyping data was produced to maximize the support on predictions on the phylogeny. A cluster was defined as two or more patient isolates exhibiting identical fingerprint patterns. Dendograms were generated using the categorical coefficient and based on unweighted pair group method with arithmetic mean. The epidemiological and clinical data were analyzed against the lineages observed in the study. Pearson's chi square test was applied to calculate the statistical significance. A p value <0.05 was considered significant. The statistical analysis of data was performed by using the software SPSS version-19.0 (IBM Corporation, NY, USA).

3. Results

3.1. Study population

During the indicated study period, 902 isolates from the nine reference laboratories were enrolled. The proportion of cases in the age groups 16–29 (33.3%) and 30–44 (29.6%) were predominant. The proportion of male cases (68.3%) exceeded that of females by far (31.7%) with a male/female ratio of 2.1. The number of Saudi born patients (51.3%) almost equaled that of foreign born (48.7%). The non Saudi patients consisted of immigrants from 33 countries and they were further divided into four larger groupings according to geographical origin. These groupings were; Africa (36%), consisting of patients from 22 African countries, the Indian sub-continent (35.3%), South East Asia (18.4%) and West Asia-Yemen (10.2%). The proportion of pulmonary cases (77.7%) was dominated in the study (Table 1).

3.2. Diversity of the strain lineages

The *M. tuberculosis* population structure was found highly diverse and consisted of more than 15 lineages. Delhi/CAS (26.4%), EAI (13.7%), Haarlem (11.3%), Cameroon (6.9%) and Ghana (6.6%) were the most predominant lineages. The Beijing genotype was found in 52 (5.8%) cases. In addition, 8 other Euro American lineages (LAM, S, TUR, Uganda-I, Uganda-II, URAL, X, New-I) were also observed. *M. bovis* BCG was found as the causative agent of extrapulmonary TB in 11 (1.2%) cases. In total, lineages of 104 (11.4%) cases could not be defined using the MIRU-VNTR plus database (Fig. 1). The isolates of undefined lineages were further analyzed