Infection, Genetics and Evolution 14 (2013) 361–368

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

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

Molecular, ethno-spatial epidemiology of leprosy in China: Novel insights for tracing leprosy in endemic and non endemic provinces

Xiaoman Weng [a], Yan Xing [a], Jian Liu [a], Yonghong Wang [b], Yong Ning [c], Ming Li [d], Wenbin Wu [e], Lianhua Zhang [f], Wei Li [g], Vander Heiden Jason [g], Varalakshmi Vissa [g,*]

[a] Beijing Friendship Hospital – Affiliate of Capital University of Medical Sciences, Beijing Tropical Medicine Research Institute, 95 Yang An Road, Beijing 100050, China
[b] Skin Diseases Control Station of XIngyi City, #54 Fuxingxi Road, Jishan New District, Xingyi, Guizhou Province 562400, China
[c] Institute of Dermatology and Venereology, Sichuan Academy of Medical Sciences, #12 Sidao Street, Chengdu, Sichuan Province 610031, China
[d] Guangdong Provincial Center for Skin Disease and STI Control, Guangdong Provincial Dermatology Hospital, #2 Lujing Road, Guangzhou, Guangdong 510095, China
[e] Fujian Provincial Center for Disease Prevention and Control, #76 Jintai Road, Fuzhou, Fujian 350001, China
[f] Jiangsu Provincial Center for Disease Prevention and Control, #172 Jiangsu Road, Nanjing, Jiangsu 210009, China
[g] Department of Microbiology, Immunology and Pathology Colorado State University, Fort Collins, CO 80526, USA

ARTICLE INFO

Article history:
Received 3 October 2012
Received in revised form 4 December 2012
Accepted 10 December 2012
Available online 4 January 2013

Keywords:
VNTR
Strain typing
Mycobacterium leprae
Leprosy
Phylogeography
Maritime silk routes

ABSTRACT

Leprosy continues to be detected at near stable rates in China even with established control programs, necessitating new knowledge and alternative methods to interrupt transmission. A molecular epidemiology investigation of 190 patients was undertaken to define *Mycobacterium leprae* strain types and discern genetic relationships and clusters in endemic and non-endemic regions spanning seventeen provinces and two autonomous regions. The findings support multiple locus variable number of tandem repeat (VNTR) analysis as a useful tool in uncovering characteristic patterns across the multiethnic and divergent geographic landscape of China. Several scenarios of clustering of leprosy from township to provincial to regional levels were recognized, while recent occupational or remote migration showed geographical separation of certain strains. First, prior studies indicated that of the four major *M. leprae* subtypes defined by single nucleotide polymorphisms (SNPs), only type 3 was present in China, purportedly entering from Europe/West/Central Asia *via* the Silk Road. However, this study revealed VNTR linked strains that are of type 1 in Guangdong, Fujian and Guangxi in southern China. Second, a subset of VNTR distinguishable strains of type 3, co-exist in these provinces. Third, type 3 strains with *rpoT* VNTR allele of 4, detected in Japan and Korea were discovered in Jiangsu and Anhui in the east and in western Sichuan bordering Tibet. Fourth, considering the overall genetic diversity, strains of endemic counties of Qiubei, Yunnan; Xing Yi, Guizhou; and across Sichuan in southwest were related. However, closer inspection showed distinct local strains and clusters. Altogether, these insights, primarily derived from VNTR typing, reveal multiple and overlooked paths for spread of leprosy into, within and out of China and invoke attention to historic maritime routes in the South and East China Sea. More importantly, new concepts and approaches for prospective case finding and tracking of leprosy from county to national level have been introduced.

1. Introduction

China is a large country, historically involved in international trade and migration through many centuries. Though the leprosy epidemic has been recorded since the time of Confucius (551–479 BCE), how leprosy spread into, within and from China are not completely resolved. Furthermore, as China has claimed elimination status for leprosy at the national level, active case finding efforts by mass, contact or group clue surveys have been either reduced or eliminated and a decentralized control program largely based on voluntary reporting and dermatological clinics has been adopted (Chen et al., 2001). The World Health Organization reported that 1324 new cases of leprosy cases detected in 2010, 85% of them being of the multibacillary form (WER, 2011). While, leprosy patients are still diagnosed nationwide, pockets of endemicity in the three ethnically diverse, mountainous and underdeveloped southwest provinces of Yunnan (YN), Guizhou (GZ) and Sichuan (SC) account for the major proportion of leprosy in China (Yu et al., 2010; Shen et al., 2008). In addition, migration of patients from highly endemic areas to other non-endemic areas with faster social-economic development prospects is ongoing (Shen et al., 2008). The case detection rates and delay in diagnosis are dependent on multiple factors such as age, occupation, nationality,

* Corresponding author. Tel.: +1 970 223 4981; fax: +1 970 491 1825.
E-mail address: vvissa@gmail.com (V. Vissa).

1567-1348
http://dx.doi.org/10.1016/j.meegid.2012.12.009

endemicity, leprosy type and detection method (Chen et al., 2000; Shen et al., 2010, 2011). Therefore, newer tools and information can be incorporated to fully understand incidence and transmission of the various strains of leprosy in low and high endemic areas. In this regard, a molecular epidemiology approach utilizing strain typing of the pathogen *Mycobacterium leprae* linked to geographic, social, cultural and economic factors can provide useful complementary information. For instance, it will be helpful in case finding, in that once a cluster is identified by genotyping, it may be a clue to detect other undiagnosed leprosy patients by focusing on the geographical distribution or specific communities where the clusters are found, and taking into consideration their familial, social and occupational interactions. On the other hand, strain types that are novel to a region, may be explained by occupational migration, which provides opportunities for further spread of disease.

The present study attempts to expand on the knowledge of strain type diversity in China. Currently, a number of variable number of tandem repeat (VNTR) and single nucleotide polymorphisms (SNPs) have been discovered and applied to describe *M. leprae* strains for different geographical scales (Groathouse et al., 2004; Zhang et al., 2005; Monot et al., 2005, 2009; Truman et al., 2011). Strains within different regions and countries have been distinguished by applying VNTRs (Weng et al., 2007; Cardona-Castro et al., 2009; Fontes et al., 2009; Kimura et al., 2009; Phetsuksiri et al., 2012; Sakamuri et al., 2009; Srisungnam et al., 2009). Utilizing the SNP typing markers, in China, a single SNP type, i.e., 3K which is one of sixteen major types (Weng et al., 2007; Monot et al., 2009) has been detected. VNTR typing enabled further resolution of such type 3 strains at township, village, ethnic, and family scales to detect clusters of transmission in an endemic county in Yunnan Province (Weng et al., 2007, 2011). The present study extends our analyses and investigations in China and discusses strain types and genetic markers of similarity and differentiation amongst Chinese *M. leprae* strains covering a wider geographical range informative to provincial and national leprosy control programs.

2. Materials and methods

2.1. Patient inclusion, sample collection and M. leprae strain typing

During leprosy diagnosis, the collection of skin biopsies is routine and is performed by doctors of the Province or County level Skin Disease Control Stations (SDCS) of Centers of Disease Control (CDC) across China. A portion of the biopsies was placed in 70% ethanol for molecular studies. The use of these samples for research was approved by the ethical committee of the Beijing Tropical Medicine Research Institute. A subset of the patients belonged to earlier study cohorts (Weng et al., 2006, 2007, 2011; Xing et al., 2009).

2.2. M. leprae VNTR strain typing

M. leprae were strain typed using previously described methods (Kimura et al., 2009; Sakamuri et al., 2009; Weng et al., 2007, 2011). DNA extraction from biopsies followed by multiplex-PCR was performed. Fragment length analysis was applied to determine the VNTR patterns of each isolate based on the multiplex-PCR products (Main Research Center at Beijing Academy of Agriculture and Forestry Sciences or Proteomics and Metabolomics Facility, Colorado State University). Capillary electrophoresis was performed on an ABI 3730XL DNA Analyzer. Samples were typed at the following 17 VNTR loci: (AC)8b, (GTA)9, (GGT)5, (AT)17, 6–3a (*rpoT*), 21–3, (AC)9, (AT)15, (AC)8a, 27–5, 6–7, (TA)18, (TTC)21, 18–8, 23–3, 12–5 and (TA)10 (Groathouse et al., 2004; Kimura et al., 2009).

2.3. SNP typing of M. leprae

The SNP type (1–4) was identified for a subset of these samples ($n = 101$). A PCR-RFLP based procedure was used for differentiation of SNP types 1–4 as described by Sakamuri et al. (2009); sequencing of SNPs was performed for samples that could not be typed by the RFLP assay (Monot et al., 2005). SNP subtyping of *M. leprae* was performed by sequencing of the PCR products using published primer sequences (Monot et al., 2009) for a subset of samples. Subtype 3 K strains were identified by first sequencing SNPs at positions 2312059 (distinguishes SNP subtypes A–J from subtypes K–P) and 413902 (distinguishes subtypes A–K from subtypes L–P); if the strains were of 3 K, there was no requirement to map the H/I and I/J SNPs to resolve the subtype. 1A–D were identified by sequencing SNPs at positions 8453, 313361 and 61425. PCR products were sequenced using ABI BigDye® Terminator v3.1 and an ABI 3730XL DNA Analyzer (Sangon Biotech, Shanghai Co., Ltd.).

2.4. Analysis of strain types

The population structure of 144 samples having complete data (shown in bold in Supplementary Table 1) was explored using principal component analysis (PCA), using the complete panel of VNTR loci. The SNP type of a subset of typed samples was not used as a locus in the analysis methods, but was used as a categorical variable. This analysis was performed at Colorado State University in the U.S. PCA was performed on individuals using the correlation matrix within the software Minitab 16.

Maximum parsimony analysis was performed using the software PAUP* 4.0b (Swofford, 2003) and the methods described by Sakamuri et al. (2009). The heuristic search method was used, with tree bisection-reconnection and 1000 maximum trees. Loci were assigned weights corresponding to the reciprocal of Nei's unbiased gene diversity (Nei, 1987). A step matrix was implemented to apply a step-wise mutation model; a difference of one repeat equated to one step. The out group was defined as the set of SNP type 1 samples. From the 1000 trees generated by PAUP* a consensus network was built using SplitsTree4 v4.11.3 (Huson and Bryant, 2006).

3. Results

3.1. Genetic diversity of M. leprae in China

The *M. leprae* strain types in the skin biopsies of 190 leprosy patients diagnosed in 17 different provinces (including Shanghai) and two autonomous regions (Tibet and Xinjiang), shown in Fig. 1 (Fig. 1A and B) were determined. A panel of 17 VNTR loci was mapped by multiplex PCR-fragment length analyses methodology. Complete *M. leprae* VNTR profiles available for 144 patients were further analyzed. Nearly all DNA samples except those from Yunnan and Guizhou patients were also typed according to the SNP 1–4 system. The available demographic, clinical and *M. leprae* genotyping information for all the 190 patients has been compiled in Supplementary Table 1.

As a first approach, principal component analysis (PCA) was performed to explore the genetic variation within the VNTR dataset. Post analysis labeling of individuals according to the province, SNP type and selected VNTR loci alleles reveal several novel trends regarding the strain types and their geographical distributions in China (Fig. 2). As seen in Fig. 2A, the VNTR profiles of strains from different provinces separate into at least three subpopulations which are depicted within dotted lined circles (labeled as Ch1, Ch2 and Ch3) and a smaller group (labeled as Ch4) straddling Ch1 and Ch3.