Gene 582 (2016) 128–136

Contents lists available at ScienceDirect

Gene

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

Research paper

Whole genome analysis of an MDR Beijing/W strain of *Mycobacterium tuberculosis* with large genomic deletions associated with resistance to isoniazid

Qiufen Zhang [a], Baoshan Wan [a], Aiping Zhou [b], Jinjing Ni [a], Zhihong Xu [a], Shuxian Li [a], Jing Tao [a,*], YuFeng Yao [a,*]

[a] Laboratory of Bacterial Pathogenesis, Department of Microbiology and Immunology, Institute of Medical Sciences, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China
[b] Department of Laboratory Medicine, East Hospital Affiliated to Tongji University, Shanghai 200120, China

ARTICLE INFO

Article history:
Received 23 October 2015
Received in revised form 30 January 2016
Accepted 2 February 2016
Available online 4 February 2016

Keywords:
Mycobacterium tuberculosis
Isoniazid resistance
Large sequence polymorphisms
Beijing/W genotype
Genome sequencing

ABSTRACT

Mycobacterium tuberculosis (*M.tb*) is one of the most prevalent bacterial pathogens in the world. With geographical wide spread and hypervirulence, Beijing/W family is the most successful *M.tb* lineage. China is a country of high tuberculosis (TB) and high multiple drug-resistant TB (MDR-TB) burden, and the Beijing/W family strains take the largest share of MDR strains. To study the genetic basis of Beijing/W family strains' virulence and drug resistance, we performed the whole genome sequencing of *M.tb* strain W146, a clinical Beijing/W genotype MDR isolated from Wuxi, Jiangsu province, China. Compared with genome sequence of *M.tb* strain H37Rv, we found that strain W146 lacks three large fragments and the missing of *furA-katG* operon confers isoniazid resistance. Besides the missing of *furA-katG* operon, strain W146 harbored almost all known drug resistance-associated mutations. Comparison analysis of single nucleotide polymorphisms (SNPs) and indels between strain W146 and Beijing/W genotype strains and non-Beijing/W genotype strains revealed that strain W146 possessed some unique mutations, which may be related to drug resistance, transmission and pathogenicity. These findings will help to understand the large sequence polymorphisms (LSPs) and the transmission and drug resistance related genetic characteristics of the Beijing/W genotype of *M.tb*.

1. Introduction

Drug-resistant strains of *Mycobacterium tuberculosis* (*M.tb*) pose a major threat to the control of TB worldwide (Tomioka and Namba, 2006; Nodieva et al., 2010). Although the curative anti-tuberculosis therapy has been established a half century ago, *M.tb* resistance to the most important anti-tuberculosis drugs emerged quickly due to inappropriate treatment (van der Werf et al., 2012; Langendam et al., 2012). The drug-resistant tuberculosis spread quickly owing to the convergence of resistant strains of *M.tb* especially in high-risk patients such as those with HIV/AIDS and in high-risk environments such as hospitals (Moro et al., 1998; Pedersen et al., 1997). Meanwhile, compared with patients with drug-susceptible tuberculosis, patients with drug-resistant tuberculosis were more likely to generate new secondary cases (Prasad, 2010).

The molecular mechanisms of drug resistance have been elucidated in several studies. Spontaneous chromosomal mutations account for drug resistance in *M.tb* and mutations in a few genes of *M.tb* result in the occurrence of resistance to rifampin (RIF), isoniazid (INH), streptomycin (SM), and ethambutol (EMB), the main components of the first-line multidrug therapy for tuberculosis (Sintchenko et al., 1999; Kozhamkulov et al., 2011; Ramaswamy et al., 2003; Tracevska et al., 2004; Wong et al., 2011). More than 97% of RIF-resistant cases is due to mutations within the 81 bp "core region" of the *rpoB* gene, called the RIF resistance-determining region (RRDR) (Cole, 1996). As a prodrug, INH inhibits mycolic acid synthesis and its primary target is believed to be the enoyl-acyl carrier protein InhA. Mutations in the promoter and ORF region of *inhA* lead to a low level of INH resistance. In contrast, mutations in *katG*, which encodes the catalase–peroxidase enzyme KatG, result in a high level resistance (Oliveira et al., 2006; O'Brien et al., 1996). Nearly 70% of the streptomycin-resistant clinical isolates present mutations associated with *rpsL* and *rrs* genes, which encode the ribosomal protein S12 and 16S rRNA respectively (Tracevska et al., 2004). Mutations in codon 43 in the *rpsL* gene are the most common case. 47–65% mutations associated with resistance to EMB are found in the *embB* gene (Cuevas-Cordoba et al., 2015).

The World Health Organization (WHO) estimates that there were 480,000 new cases of MDR-TB worldwide in 2014, defined as strains that are resistant to at least the two most powerful first-line anti-TB drugs (isoniazid and rifampicin) (Zumla et al., 2015). Among patients with pulmonary TB who were notified in 2013, an estimated 300,000

Abbreviations: M.tb, Mycobacterium tuberculosis; TB, tuberculosis; MDR, multiple drug-resistant; SNPs, single nucleotide polymorphisms; LSPs, large sequence polymorphisms; RIF, rifampin; INH, isoniazid; SM, streptomycin; EMB, ethambutol; PZA, pyrazinamide; RRDR, RIF resistance-determining region; DTM-PCR, deletion-targeted multiplex PCR.

* Corresponding authors.
E-mail addresses: taojing00@sjtu.edu.cn (J. Tao), yfyao@sjtu.edu.cn (Y. Yao).

http://dx.doi.org/10.1016/j.gene.2016.02.003
0378-1119/

(range: 230,000–380,000) had MDR-TB. More than half of these patients were in China, India and the Russian Federation (Zumla et al., 2015).The Beijing/W genotype of *M.tb* is one of the most successful *M.tb* lineages in China even in the world (Zumla et al., 2015). It has been reported that several sublineages of the Beijing/W genotype are more frequently associated with MDR-TB (Iwamoto et al., 2008).

Gao and his colleagues reported that Beijing/W genotype strains have an increased ability to transmit infection, with Beijing/W genotype strains progressing more rapidly to active tuberculosis (Yang et al., 2012). However, the mechanisms underlying these epidemiological findings remain to be clearly elucidated. It is important to search for molecular basis for transmission, pathogenicity warrant and drug resistance in Beijing/W genotype. The whole genome sequencing is an efficient method for acquiring the genetic information of *M.tb*. Sequencing of the whole genomic DNA of *M.tb* strain H37Rv was completed in 1998 and it opened the way to understand the biology, metabolism and evolution of this pathogen in the genomic level (Brosch et al., 1998). In recent years, some Beijing/W genotype strains including drug-sensitive and multidrug-resistant strains' complete genome sequences have been determined (Wu et al., 2013; Zhang et al., 2011). However, the whole genome sequence of MDR-*M.tb* with large deletions is still lacking.

In this study, we sequenced the genome of one clinical MDR isolate-strain W146 isolated from Wuxi, Jiangsu province, China. We found W146 missing three large sequences and the missing of *furA-katG* operon conferred isoniazid resistance. We also analyzed the SNPs and indels of strain W146. We performed the comparative genome analysis of W146 with other Beijing/W genotype strains and non-Beijing/W genotype strains. We hope our report could promote the understanding of LSPs (large sequence polymorphisms) of *M.tb* and the genome polymorphism of the Beijing/W family.

2. Materials and methods

2.1. Ethics statement

The study was reviewed and approved by the local ethics committee (The ethics committee of Wuxi People's Hospital). Written informed consent was obtained from participants prior to their enrollment in the study.

2.2. Strain isolation and drug susceptibility testing

The clinical isolate of *M.tb* strain W146 was obtained from patient with pulmonary tuberculosis in Wuxi. The isolate was grown in Lowenstein–Jenson (L–J) at 37 °C for approximately 3–4 weeks. The isolate was tested for drug susceptibility according to the Clinical and Laboratory Standards Institute (CLSI). Strain W146 was analyzed by agar proportion method utilizing Lowenstein–Jenson (L–J) supplemented individually with the following drugs: RIF (1 μg/ml), INH (0.2, 1, and 5 μg/ml), EMB (5 μg/ml), SM (1.0 μg/ml), PZA (100 μg/ml).

2.3. Genomic DNA extraction and genotyping

DNA was extracted from Lowenstein–Jenson (L–J) subcultures using QIAamp DNA Mini Kit (Catalog number: 51,306) in accordance with the manufacturer's instruction. The genotype of strain W146 was identified by deletion-targeted multiplex PCR (DTM-PCR) on RD105 as described by Chen et al. (2007). Primers used in this experiment were listed in the Table 1.

2.4. Genomic sequencing and bioinformatics analysis

Complete genomic sequencing of strain W146 was performed by Illumina Hiseq2000. The genome was mapped by Bowtie2 according to the sequence of *M.tb* strain H37Rv. Gene pathways were annotated with KEGG. Comparative genome analysis was performed with the genome sequence of *M.tb* H37Rv (accession number NC_000962.2) as reference, using the Mauve version 2.3.1 (Wu et al., 2013). A Venn diagram for SNPs was drawn by R software using the VennDiagram package version 1.5.1. The numbers of mutated genes (SNPs or indels) in COGs and pathways were subjected to enrichment analysis using a classical hypergeometric distribution statistical analysis (Wu et al., 2013). A p value less than 0.05 was considered to have statistical significance. The phylogenetic tree was analyzed using the neighbor joining method.

Table 1
Primers used for PCR and sequencing.

Primer name	Primer sequence(5′ → 3′)
Beijing P1	GGAGTCGTTGAGGGTGTTCATCAGCTCAGTC
Beijing P2	CGCCAAGGCCGCATAGTCACGGTCG
Beijing P3	GGTTGCCCACTGGTCGATATGGTGGACTT
katG F	AACGGCTTCCTGTTGGACGAG
katG R	GGGTCTGACAAATCGCGCCG
Rv1592c F	ATGGTAGAGCCCGGCAATTT
Rv1592c R	TCAGAGCGGACGGCGGCTCA
ropB F	CTTCTCCGGGTCGATGTCGTTG
ropB R	CGCGCTTGTCGACGTCAAACTC
rpsL F	GTCAAGACCGCGGCTCTGAA
rpsL R	TTCTTGACACCCTGCGTATC
rrs F	TCACCATCGACGAAGCTCCG
rrs R	CTAGACGCGTCCTGTGCATG
gidB F	ATGTCTCCGATCGAGCCCGC
gidB R	TCACGCCGTCCCTCCACTCG
embB F	CGGCATGCGCCGGCTGATTC
embB R	TCCACAGACTGGCGTCGCTG
gyrA F1	ATGACAGACACGACGTTGCC
gyrA R1	ATCAGTGCAATGACCTCGTC
gyrA F2	TACGTTGACCACCAACTCGA
gyrA R2	TTAATTGCCCGTCTGGTCTG
LGD1 F	TTCAGGCTACGGTAGCGGTG
LGD1 R	CGGGTCGGTGGTCACCAATT
LGD3 F	CTCGGCGGCCGCGGTGAT
LGD3 R	TCGGCGGTCTTGGTCTCCAT
LGD3 walking F	CCAAGCGCGGGCCACCTTTG
cut1 F	ATGCCGGGGCGGTTCAGAGA
cut1 R	GCCGACCGGCGGCAAACACT
Rv1760 F	TTGGCGGGGTTGGGGATTTC
Rv1760 R	CGGCTGGTCATCACTGCACT
Rv1761c F	ATGTCTGATTTCGATACCGAGCG
Rv1761c R	TCAGAACCGGTAGTCGGTGCC
Rv1762c F	ATGCAATCAAGCTCGCTCGATCC
Rv1762c R	TCAGCTGTCCATCGGCAGGAC

2.5. PCR amplification for detecting mutations and the large deletions

Oligonucleotide primers used for detecting the drug resistance-related mutations and confirming the large deletions were designed from the *M.tb* strain H37Rv genome sequence (Table 1) (Brosch et al., 1998). For drug resistance-related mutations detection, genes of *katG*, *Rv1592c*, *ropB*, *rpsL*, *rrs*, *gidB*, *embB*, *gyrA* were amplified by PCR and were sequenced by an ABI 377 automatic DNA sequencer (Applied Biosystems) using the appropriate sequencing primers. To further confirming large deletions' genomic locations to the base pair, PCR and sequencing were performed. The amplification primers for confirming LGD1 (large genomic deletion) were LGD1 F and LGD1 R. LGD1 F and LGD1 R were also used for LGD1 PCR product sequencing. The amplification primers for confirming LGD3 were LGD3 F and LGD3 R. LGD3 walking F was used for LGD3 PCR product sequencing.

2.6. Accession number

This whole genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LJBH00000000. The version described in this paper is version LJBH01000000.