

两株超级广泛耐药结核分枝杆菌全基因组比较分析

林楠¹ 周杰² 周盈³ 汪世华^{1*}

(1. 福建农林大学 生命科学学院 福建 福州 350001)

(2. 佛山市第四人民医院 广东 佛山 528000)

(3. 中国科学院生物物理研究所 非编码核酸重点实验室 北京 100101)

摘要: 【目的】结合现有数据,通过对两株临床超级广泛耐药的结核分枝杆菌全基因组的测序和分析,发现其型别相关的突变位点,解释发生广泛耐药的基因组突变机制。【方法】利用 Solexa 第二代测序技术对两株广泛耐药结核分枝杆菌(FJ05194 和 GuangZ0019)进行全基因组测序分析。以 H37Rv 为参考序列得到两株广泛耐药菌株的单核苷酸多态性(SNPs),构建系统发育树鉴定菌株型别,判断突变位点中型别相关和非型别相关的 SNPs。定位 SNPs 所在的基因组区域,对型别相关的突变基因进行 KEGG 通路的富集分析,对非型别相关的突变基因和间隔区判断是否与耐药相关。【结果】两株广泛耐药菌株分别属于 Lineage2 和 Lineage4 型别,两菌株在碱基替换方面存在差异性,Lineage2 型别相关的基因功能富集于 ABC 转运蛋白和核苷酸切除修复的通路。耐药方面,发现了已知的耐药相关基因的突变(*rpoB*、*katG*、*rpsL*、*gyrA*、*gyrB*、*embB* 和 *ethA* 等),但卷曲霉素和卡那霉素相关的 *rrs*、*tlyA* 和 *eis* 启动子区域未发生突变,不足以解释其耐药性的产生。与最新报道的候选耐药基因比较,发现了卷曲霉素和卡那霉素相关的突变(Rv1393c、Rv0265c 和 *narX* 等)和外排泵相关的 *pstB*、Rv2333c 和 Rv2687c 突变。【结论】结核分枝杆菌 Lineage2 型别相关的 SNPs 中含有影响结核分枝杆菌突变率和耐药性的突变。对于两株超级广泛耐药的结核菌,已知的激活药物或药靶相关的单耐药基因突变集合不能完全解释其广泛耐药性,还涉及新候选结核耐药基因、外排泵和补偿等其他潜在机制的相关基因突变。

关键词: 单核苷酸多态性, 结核分枝杆菌, 耐药性, 比较基因组

Comparative genome analysis of two super-extensive drug-resistant isolates of *Mycobacterium tuberculosis*

LIN Nan¹ ZHOU Jie² ZHOU Ying³ WANG Shi-Hua^{1*}

(1. College of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou, Fujian 350001, China)

(2. The 4th Peoples' Hospital, Foshan, Guangdong 528000, China)

(3. Key Laboratory of Non-coding RNA, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China)

Abstract: [Objective] Based on the genome sequences of two super extensive drug resistant isolates

基金项目: 国家 973 计划项目(No. 2012CB518700); 中国科学院重点部署项目(No. KSZD-EW-Z-006)

*通讯作者: Tel/Fax: 86-591-87984471; ✉: wshyy1@sina.com

收稿日期: 2013-12-26; 接受日期: 2014-01-22; 优先数字出版日期(www.cnki.net): 2014-02-26

of *Mycobacterium tuberculosis*, we attempted to identify the lineage-specific SNPs and found the drug resistant-associated mutations in them. **[Methods]** By using the next-generation sequencing technology on two super-extensive drug resistant isolates of mycobacterium tuberculosis (FJ05194 and GuangZ0019), we compared them with the reference sequence to identify the single nucleotide polymorphisms (SNP) and divided the SNPs into lineage-specific and isolate-specific SNPs. Check the coordinate of SNPs on the genome, then we did the enrichment analysis of the lineage-specific SNPs, and were able to found out the drug resistant mutations from the isolate-specific SNPs by overlapping with the candidate drug resistance associated regions. **[Results]** The Lineage2-specific SNPs was significantly enriched on ABC transporters and nucleotide excision repair KEGG pathways, and the extensive drug resistance were mainly attributed to the mutation of the well-known genes (*rpoB*, *katG*, *rpsL*, *gyrA*, *gyrB*, *embB*, *ethA* and so on), but the capreomycin and kanamycin resistance was probably due to some candidate resistant-associated genes (Rv1393c, Rv0265c, *narX*) and the efflux pump genes (*pstB*, Rv2333c and Rv2687c). **[Conclusion]** The Lineage2-specific-SNPs possible result in the high mutation rate and drug resistance of the Lineage2. And the mechanism of super extensive drug resistance is more complex, it not only involve with the drug target and drug activation genes, but also some unclear candidate genes such as compensatory genes and efflux pump genes.

Keywords: Single nucleotide polymorphisms, *Mycobacterium tuberculosis*, Drug resistance, Comparative genome analysis

结核分枝杆菌于 20 世纪四五十年代初随着一些抗结核药物的陆续出现,得到了阶段性的控制。但是由于公共卫生监督力度的不足,在部分抗生素的广泛滥用与艾滋病并发感染等多种因素下,结核病疫情死灰复燃,尤其是在发展中国家,如印度、中国等更是成为耐多药结核病重灾区^[1]。在中国约有 1/10 的结核病患者是耐多药结核病(至少对一线药物异烟肼和利福平耐药,MDR),在所有新发生结核病患者中 MDR 的比例更是达到 5.7%,且 8%的 MDR 可被定义为广泛耐药结核病(即至少对一线药物异烟肼和利福平耐药外,还同时对卡那霉素、阿米卡星或卷曲霉素中任何一种注射类药物和任何一种喹诺酮类药物耐药,XDR)^[2],甚至部分地区已经出现全耐药的结核分枝杆菌。在流行病学方面,中国地区的结核病人主要受到 Lineage2 和 Lineage4 两种型别的结核分枝杆菌侵染,尤其是 Lineage2 在东亚的传播范围更广,由于其相关的耐药菌株患者多,且表现易突变的表型特征,因此对 Lineage2 的研究也备受关注。

本文选取了两株超级的广泛耐药分枝杆菌(同时耐 4 种一线药物和 4 种二线药物),生物信息学

方法初步探索这两株来自 Lineage2 和 Lineage4 的菌株之间的基因组水平差异,分析两株超级广泛耐药分枝杆菌的耐药性是否由已知报道的单耐药相关基因突变累加导致,进一步解释其耐药机制。

1 材料与方法

1.1 材料

菌株:本实验所涉及的 2 株临床广泛耐药结核菌株由中国疾病预防控制中心传染病所结核室提供,分别来自福建和广东地区。

1.2 实验方法

1.2.1 药敏试验和菌株 DNA 的提取: L-J 比例法检测 4 种一线/4 种二线抗结核药的药敏情况,参考 2008 WHO 标准^[3],终浓度分别为利福平(RMP, 40.0 mg/L),异烟肼(INH, 0.2 mg/L),链霉素(STR, 4.0 mg/L),乙胺丁醇(EMB, 2.0 mg/L),卷曲霉素(CPM, 40 mg/L),卡那霉素(KAN, 30 mg/L),氧氟沙星(OFX, 2.0 mg/L)和乙硫异烟胺(ETH, 40 mg/L),并采用改良的 CTAB 法提取基因组 DNA^[4]。

1.2.2 全基因组测序、组装和注释:用 2 个 DNA 样品分别构建文库,并采用 Illumina HiSeq 2000 平台的二代测序。SOAPdenovo v1.05 软件对得到

的 Reads 数据进行拼接和补缺^[5](原始 Reads 数据上传至 SRA 数据库, 登录号: SRA068167), 完成 2 株菌株的基因组草图, 然后 Glimmer v3.02 对基因进行预测^[6], BLAST 比对 NCBI 的 NR、COG 和 KEGG 等数据库以完成对预测基因的注释, RNAmmer 和 tRNAscan-SE 1.21 对 RNA 水平进行注释^[7-8]。

1.2.3 单核苷酸多态性的验证和系统发育树的构建: 以 H37Rv 为参考序列, 2 株菌株的 Clean reads 分别进行 SOAP2 比对^[9], 当某位点的异于参考序列碱基的 Reads 数大于该位点总 Reads 的 70%, 且总 Reads 大于 10, 则该位点确定为 SNP, 为了评估该标准的 SNP 可靠性, 本文设计 15 对引物(13 个区域)的扩增和 Sanger 测序验证。系统发育树是根据 SNPs 在所有菌株中情况, 以 *Mycobacterium canettii* 为外群, 通过 TreeBeST 软件的邻接法(Neighbor-Joining)进行构建(<http://treesoft.sourceforge.net/treebest.shtml>)。

1.2.4 生物信息学的注释和分析: 以参考序列 H37Rv (NC00962.2)的基因组注释信息, 分别对 2 株广泛耐药菌株 SNPs 的突变情况进行注释, 并统计其位点在不同型别(Lineage2 和 Lineage4)菌株中的突变情况。利用 DAVID 进行富集分析^[10], 其中 *P* value 即 EASE (校正后的 Fisher 精确检验)小于 0.05 的为具有统计学显著意义, 并对其中发生非同

义突变的 SNPs (nsSNPs)进行 SIFT 预测^[11], Score 分值小于 0.05 的为预测影响蛋白功能的有义突变。

2 结果与分析

2.1 比例法检测药敏结果和基因组情况

经药敏试验鉴定, 这 2 株菌株确定为对 8 个检测药物全耐, 为超级广泛耐药的菌株。Solexa 测序后拼接获得 2 株广泛耐药结核菌的基因组初步信息, 两者全基因组大小为 4 363 305 bp (FJ05194) 和 4 373 311 bp (GuangZ0019), 其 GC 含量平均约为 65.5%, 与参考序列 H37Rv 基因组情况基本一致。拼接和注释结果上传至 DDBJ/EMBL/GenBank 数据库, 登录号: ANBL00000000 (FJ05194) 和 ANFI00000000 (GuangZ0019)。

2.2 单核苷酸多态性的验证和偏好性

FJ05194 和 GuangZ0019 的 SNPs 位点数量分别为 1 287 和 795, 并对 2 个菌株的 13 个区域进行 Sanger 普通测序验证, 结果表明与二代测序得到的序列情况完全一致, 证明二代测序的 Call SNPs 方法可信^[12]。在这些 SNPs 中, 结核分枝杆菌的基因组发生转换的突变比例较高, 转换水平(FJ05194: 804, 62.67%; GuangZ0019: 516, 65.15%)明显高于颠换水平(FJ05194: 479, 37.33%; GuangZ0019: 276, 34.85%), 尤其 AT 之间的颠换率均低于 2% (FJ05194 1.76%, GuangZ0019 1.63%), 见图 1。

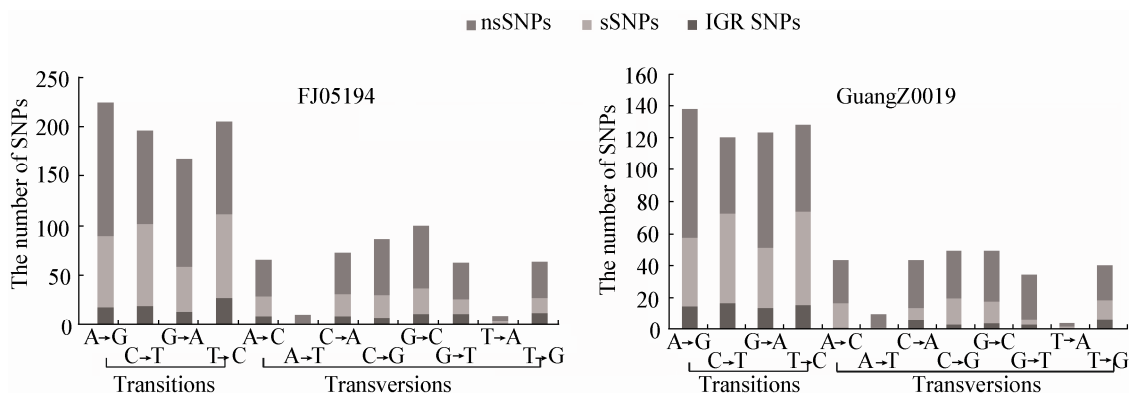


图 1 FJ05194 和 GuangZ0019 的碱基置换突变情况
Figure 1 Genetic variations between FJ05194 and GuangZ0019

计算 G:C>T:A 颠换率与 G:C>A:T 转换率的比值,因为当 DNA 遭遇活性氧(ROS)时,会将 DNA 碱基鸟嘌呤转变为 8-羟基鸟嘌呤,8-羟基鸟嘌呤在 DNA 复制时容易与腺嘌呤发生错误的配对,即导致了 G:C>T:A 颠换的发生,所以该比值越高被认为发生的突变与氧化损伤越有关^[13-14]。FJ05194 的该比值 0.37 要略高于 GuangZ0019 的比值 0.32,该结果初步表明 FJ05194 的基因组要比 GuangZ0019 因氧化损伤而导致的碱基突变比例略大。

2.3 Lineage2 和 Lineage4 型别相关位点的分析

整合 NCBI 中有分型背景的 22 株敏感结核分枝杆菌全基因组数据^[15],与本研究的 2 株广泛耐药菌株构建系统发育树,得出的进化树结构与前文一致^[15],2 株广泛耐药菌株分别属于 Lineage2 和 Lineage4 (图 2)。对于 FJ05194 和 GuangZ0019 菌株的 SNPs,首先过滤掉在所有菌株中都发生突变的 SNPs (268 个位点),这些滤掉的位点碱基情况实际上为参考序列 H37Rv 特有的;然后统计过滤后的每个 SNPs 分别在各自所属的型别中突变菌株

数(图 3)。如图 3 所示,一方面,在 6 株 Lineage2 菌株中,我们发现统计为一个菌株突变的 SNPs 是 310 个(IGR SNPs : 42, sSNPs : 89, nsSNPs : 179),即这些 SNPs 只发生在 FJ05194 菌株上,这些 SNPs 是非型别相关的位点。其他的 SNPs 则分别同时发生在 2-6 株 Lineage2 菌株中,表示这些 SNPs 为型别相关的位点,尤其是其中的 348 个 SNPs 是同时发生在所有 Lineage2 菌株上都发生了(IGR SNPs : 50, sSNPs : 108, nsSNPs : 190),这 348 个 SNPs 表明为 Lineage2 型别相关 SNPs;另一方面,因为参考序列 H37Rv 同属于 Lineage4,所以在 Lineage4 的另外 4 株菌株中,主要存在 GuangZ0019 菌株特有的 271 个非型相关 SNPs (IGR SNPs : 41, sSNPs : 91, nsSNPs : 139)。

此外,Lineage2 结核菌具有易耐药^[16-17]和高突变率的表型特征^[18-19],因此这里对 6 株 Lineage2 型别相关的 185 个发生非同义突变的基因(nsSNPs : 190)进行研究,着力于寻找 Lineage2 型别相关的突变中是否存在导致 Lineage2 表型特

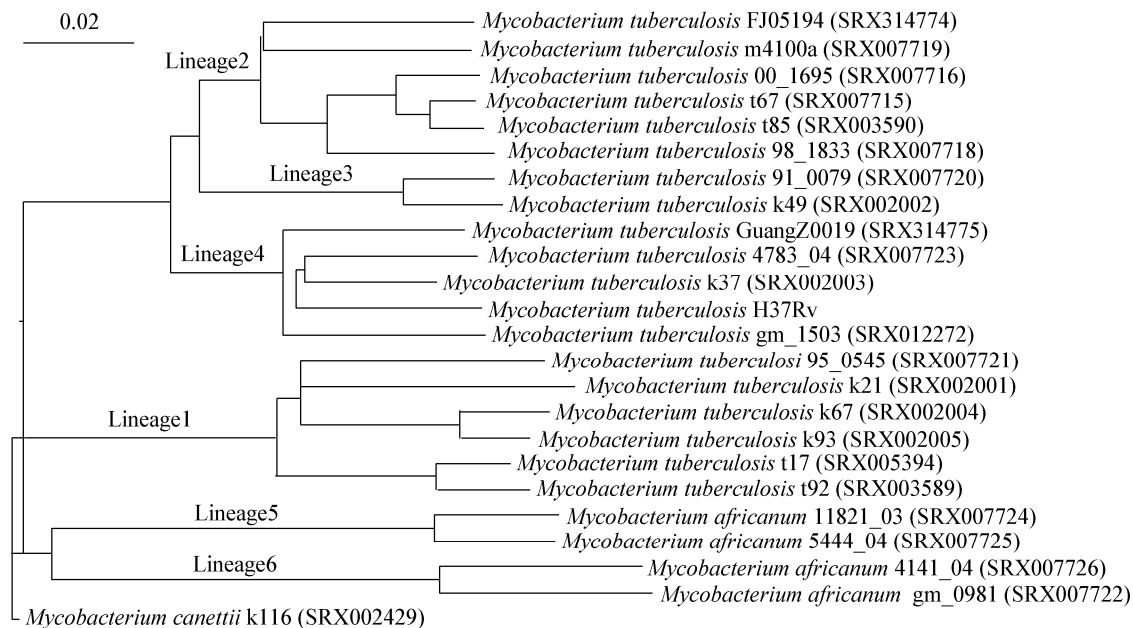


图 2 24 株结核分枝杆菌复合群的系统发育树

Figure 2 The Neighbor-Joining phylogeny of 24 *M. tuberculosis* complex isolates

注:系统发育树的标尺表示单位遗传距离,即单位长度上突变的 SNP 数。括号内为序列 SRA 接受号。

Note: The scale bars represent substitutions per SNP site. The numbers in the brackets are reads data accession numbers in SRA.

征的突变。在 DAVID 富集分析结果中得知，Lineage2 型别相关的基因中 ABC 转运蛋白 ($P=0.0096$) 和核苷酸切除修复 ($P=0.041$) 相比结核分枝杆菌基因组背景存在显著性水平差异 (表 1)。其中 ABC 转运蛋白的 *subI* 突变、核苷酸修复功能相关的 *uvrC* 和 *polA* 的突变根据 SIFT 方法判断它们为有害突变 (SIFT 值 <0.05 , 表 1), 即这些突变预测会影响对应编码蛋白的结构和功能。

2.4 FJ05194 和 GuangZ0019 非型别相关位点的耐药分析

报道称结核菌的耐药性主要是由编码区的非同义突变或间隔区的突变引起的^[20], 因此在耐药分析中, 针对 FJ05194 和 GuangZ0019 广泛耐药菌株中所特有、非型别相关的非同义突变和间隔区突变进行分析 (FJ05194 : 42 IGR SNPs, 179 nsSNPs ; GuangZ0019 : 41 IGR SNPs, 139 nsSNPs)^[12,20]。本

文从药敏试验得知 2 株广泛耐药菌株对 8 种抗结核药物全耐药, 因此将这两菌株 SNPs 突变所在区域与 8 个药物已知报道相关的所有耐药基因^[21] 进行比较, 结果发现大部分的药物能找到相应的已知突变 (表 2)。但是, 同样作为氨基糖甙类抗生素的卡那霉素和卷曲霉素, 在这两株广泛耐药菌株中却未变^[22-24]。在与最新报道的新耐药候选基因比较后 (图 4)^[20,25], 在 FJ05194 找到了 CPM/KAN 相关的候选基因 Rv1393c 的 H45R 突变, GuangZ0019 中发生了 Rv0265c 的 A235T, *narX* 的 D48Y 突变, 这一结果也刚好表明结核菌的耐药机制比原先认为的要复杂, 广泛耐药菌株的耐药性非仅仅由已知报道的基因突变可以解释^[26]。除此之外, 再进一步与外排泵相关基因进行比较^[27-28], 在 FJ05194 中发现 *pstB* 的 S252T 突变, Rv2333c 的 P514A 突变, 和 GuangZ0019 中 Rv2687c 的 P37S 突变。

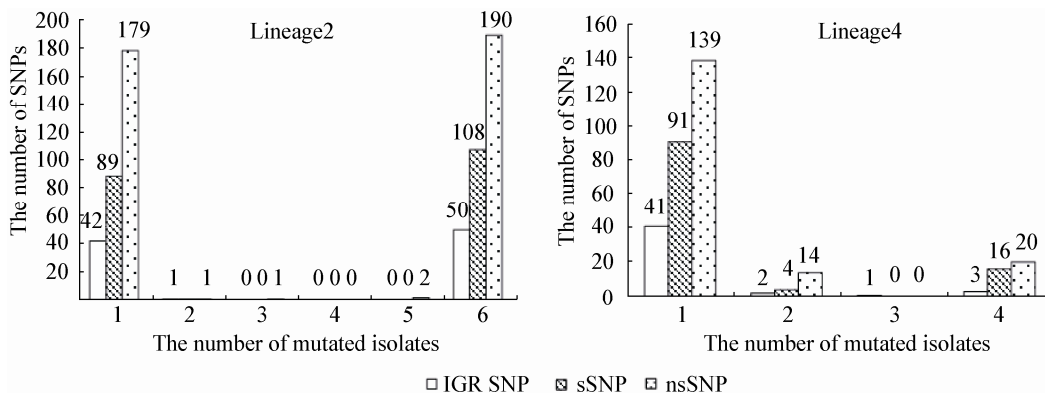


图 3 两菌株 SNPs 分别在各自 Lineage 中突变菌株数的统计情况

Figure 3 The statistics of mutated isolates containing common SNPs in two lineages respectively

表 1 Lineage2 型别相关 nsSNPs 的富集分析和 SIFT 预测

Table 1 The enrichment analysis and predicting the effect of the Lineage2-specific nsSNPs

| Rv 序号 Rv name | 基因名称 Gene name | 密码子突变 Mutations | SIFT | KEGG 通路/P 值 KEGG pathway/P value |
|------------------|-------------------|--------------------|------------|--|
| Rv0194 | - | P1098L | 0.00 | mtu02010:ABC transporters/0.0096 |
| Rv1218c | - | Q243R | 1.00 | mtu02010:ABC transporters/0.0096 |
| Rv1458c | - | T133A | 0.00 | mtu02010:ABC transporters/0.0096 |
| Rv2400c | <i>subI</i> | A76T | 0.03 | mtu02010:ABC transporters/0.0096 |
| Rv3666c | <i>dppA</i> | Q4R | 0.46 | mtu02010:ABC transporters/0.0096 |
| Rv3759c | <i>proX</i> | L85P | 1.00 | mtu02010:ABC transporters/0.0096 |
| Rv1420 | <i>uvrC</i> | V434A | 0.00 | mtu03420:Nucleotide excision repair/0.0410 |
| Rv1629 | <i>polA</i> | T186P, R188G | 0.03, 0.18 | mtu03420:Nucleotide excision repair/0.0410 |
| Rv3731 | <i>ligC</i> | R313H | 1.00 | mtu03420:Nucleotide excision repair/0.0410 |

表 2 FJ05194 和 GuangZ0019 中已知耐药相关基因的突变情况
Table 2 Previously-reported drug resistance-associated genes in our isolates

| 药物 Related drug | 相关耐药区域 Associated regions | FJ05194 Lineage2 | GuangZ0019 Lineage4 |
|--------------------|--|---------------------|------------------------|
| 异烟肼 INH | Rv1908c (<i>katG</i>) | G421S | S315N |
| | Rv1909c (<i>furA</i>) | — | T40A |
| | Rv3566c (<i>nat</i>) | — | G207R |
| | Rv1592c | E60D | — |
| | Rv2427c-Rv2428 (<i>proA-ahpC</i>) | C-52T | — |
| 利福平 RMP | Rv0667 (<i>rpoB</i>) | H526Y | S531L |
| 链霉素 STR | Rv0682 (<i>rpsL</i>) | K43R | — |
| | Rv3919c (<i>gidB</i>) | — | K163X |
| | Rv3794 (<i>embA</i>) | — | A366E |
| | Rv3795 (<i>embB</i>) | E504D | — |
| | Rv3795 (<i>embB</i>) | D1024N | — |
| 乙胺丁醇 EMB | Rv3124 | — | C175G |
| | Rv3793-Rv3794 (<i>embC-embA</i>) | — | C-11A |
| | Rv0005 (<i>gyrB</i>) | N538T | — |
| | Rv0006 (<i>gyrA</i>) | D94G | D94G |
| 氧氟沙星 OFX | Rv3854c (<i>ethA</i>) | A89E | H166P |

3 讨论

Lineage2 和 Lineage4 为中国主要流行的现代型结核分枝杆菌^[29], 尤其 Lineage2 又被认为与较强的耐药性有关^[16-17](虽然还存在争论^[30])。本文从 G:C>T:A 的颠换与 G:C>A:T 转换的比值中得知, 属于 Lineage2 的 FJ05194 因为氧化损伤导致的突变比例较 GuangZ0019 要大, 而氧化损伤又被报道同结核菌的耐药性相关^[14], 因此 Lineage2 基因组易发生氧化损伤的突变也为 Lineage2 易突变和易耐药的原因提供了一条可能的线索。从单核苷酸多态性水平看, Lineage2 和 Lineage4 存在一定的遗传距离^[31], 它们之间型别相关的 SNPs 可以作为分子标记, 为今后的快速诊断和控制各自型别的结核分枝杆菌提供参考。在对 Lineage2 型别相关的 nsSNPs 富集分析中有两条显著富集的 KEGG 通路, 分别为 ABC 转运蛋白和核苷酸切除修复通路, 分别为 ABC 转运蛋白作为一种外排泵蛋白, 正好可用作

解释 Lineage2 易发生耐多药的另一个潜在可能原因, 而核苷酸切除修复的基因突变则与 Lineage2 的高突变率特性相吻合。因此这类 Lineage2 型别相关的 SNPs 具有一定的重要性。

在对于两株广泛耐药菌株的耐药分析中, 它们对 6 种药物(INH, RMP, STR, EMB, OFX 和 ETH) 的耐药性可以由已知的耐药基因突变来解释, 如链霉素相关的 *rpsL*, 喹诺酮类相关的 *gyrA* 等(表 2)。但这两株菌株在 CPM 和 KAN 已知报道的耐药相关区域 *rrs*、*tlyA* 和 *eis* 启动子, 均未见突变。在与新发现的候选基因比较中^[20,25], 发现 CPM/ KAN 耐药性相关的新候选基因 Rv1393c、Rv0265c 和 *narX* 的突变(图 4) 而它们之间真正的功能和关联, 则亟需进一步验证, 尤其是作为超级广泛耐药的菌株, 能同时耐多种药物, 并在临床上存活和传播, 一定存在着一些突变基因是和补偿机制有关的, 用来抵御耐药有害突变而带来的适应性代价(Fitness

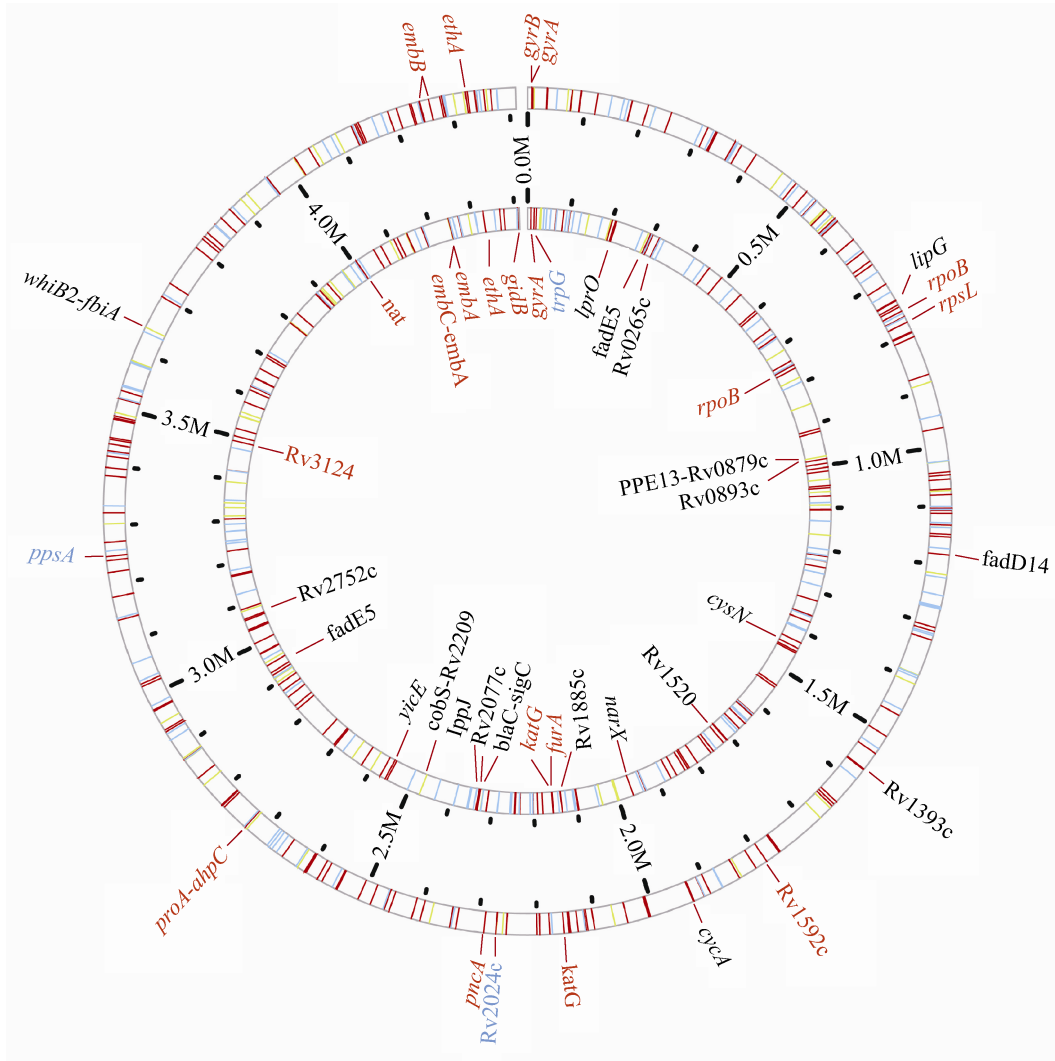


图 4 两菌株特有的 SNPs 和耐药相关突变基因的分布情况

Figure 4 Distribution of SNPs and drug resistance genes/IGRs in the two super-XDR genomes

注：圈图中外圈为 FJ05194 的非亚型相关 SNPs，内圈为 GuangZ0019 的非亚型相关 SNPs，其中圈内红色条纹表示非同义突变，蓝色条纹表示同义突变，黄色条纹为间隔区突变。用红色标记的基因名为已知报道耐药相关基因突变情况^[21]，黑色和蓝色分别表示两组最新报道的候选耐药相关基因在两菌株中的突变情况^[20-25]。

Note: The outer circle shows SNPs unique to FJ05194 and the inner circle shows those unique to GuangZ0019. Non-synonymous SNPs are represented by red bars, synonymous SNPs by blue bars, and intergenic SNPs by yellow bars. The genes and intergenic regions labelled with red words are well known regions to be associated with drug resistance^[21], labeled with black words are novel drug resistance associated regions reported by Bi L. et al^[20], with blue words are identified by Murray M. et al^[25].

cost)^[32]。如已报道的 *ahpC* 启动子区域和 *rpoC* 为对耐异烟肼、利福平菌株的补偿突变。在本文中，FJ05194 在 *katG* 的 G421S 和 *ahpC* 启动子区域同时突变，GuangZ0019 则是在 *katG* 的常见的密码子 315 位置发生非同义突变，在 *ahpC* 启动子区域

无相关的突变；同样，在利福平方面，FJ05194 和 GuangZ0019 中分别发生了 *rpoB* 的 H526Y 和 S531L 的常见位点突变，*rpoA*、*rpoC* 基因上均无发生相关的非同义突变，这点刚好吻合 GagneuxSebastien 课题组认为的 *rpoB* 在密码子 526

和 531, 和 *katG* 的密码子 315 这些常见密码子位置上突变时本身带给菌株的适应性代价小, 就无需补偿突变的发生; 反之, *katG* 其他位置的突变, 如 FJ05194 的密码子 421 突变, 就需要 *ahpC* 启动子的补偿突变^[32]。

另外, *pstB*、Rv2333c 和 Rv2687c 等结核分枝杆菌的外排泵相关基因在两株广泛耐药菌株中分别发生了突变, 由于外排泵常常认为和病原菌的耐药有关^[33], 如本文发现突变的 Rv2687c, 它编码的外排泵 ABC 转座蛋白被前人通过实验验证确实可以导致喹诺酮类的耐药发生^[28], 因此这类在耐药菌株中发生突变的外排泵相关基因同样值得关注。

随着二代测序通量的广泛应用, 关于结核分枝杆菌的全基因组水平研究愈来愈烈, 除了在流行病学的应用^[34-35], 与结核分枝杆菌的耐药性的基因组研究开始崭露头角^[20,25]。基因组学的不断深入, 结核分枝杆菌的耐药机制也被认为远比我们原本想象的复杂^[26], 尤其是二线药物的耐药相关基因突变了解更是局限, 除了与传统认为的激活药物或药物靶位的相关基因突变外, 还需要考虑不同型别的结核分枝杆菌差异性和一些间接耐药相关的基因突变, 以满足耐药或者补偿性的突变^[26]。如何从这些海量的数据中, 结合有效的菌株背景资料, 挖掘出有意义的结论和试验设计验证是一个新的挑战。本研究发现, 中国主要流行的两个现代型别的 Lineage2 和 Lineage4 结核分枝杆菌均出现了超级广泛耐药的情况, 基于两株耐药菌株的全基因组比较分析, 解释了两个不同型别超级广泛耐药菌株上的突变情况, 并为将来更全面地挖掘、验证不同型别的结核菌耐药机制提供重要参考依据。

参 考 文 献

- [1] WHO. Global tuberculosis control 2011[R]. WHO, 2011, WHO/HTM/TB/2011.16.
- [2] Zhao Y, Xu S, Wang L, et al. National survey of drug-resistant tuberculosis in China[J]. The New England Journal of Medicine, 2012, 366(23): 2161-2170.
- [3] WHO. Policy guidance on TB drug susceptibility testing (DST) of second-line drugs[R]. WHO, 2008, WHO/HTM/TB/2008392.
- [4] van Soolingen D, Hermans PW, de Haas PE, et al. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis[J]. Journal of Clinical Microbiology, 1991, 29(11): 2578-2586.
- [5] Rohde H, Qin J, Cui Y, et al. Open-source genomic analysis of Shiga-toxin-producing *Escherichia coli* O104:H4[J]. New England Journal of Medicine, 2011, 365(8): 718-724.
- [6] Delcher AL, Bratke KA, Powers EC, et al. Identifying bacterial genes and endosymbiont DNA with Glimmer[J]. Bioinformatics, 2007, 23(6): 673-679.
- [7] Lagesen K, Hallin P, Rodland EA, et al. RNAmmer: consistent and rapid annotation of ribosomal RNA genes[J]. Nucleic Acids Research, 2007, 35(9): 3100-3108.
- [8] Schattner P, Brooks AN, Lowe TM. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs[J]. Nucleic Acids Research, 2005, 33(Web Server issue): W686-689.
- [9] Li R, Yu C, Li Y, et al. SOAP2: an improved ultrafast tool for short read alignment[J]. Bioinformatics, 2009, 25(15): 1966-1967.
- [10] Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources[J]. Nature Protocols, 2009, 4(1): 44-57.
- [11] Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm[J]. Nature Protocols, 2009, 4(7): 1073-1081.
- [12] Lin N, Liu Z, Zhou J, et al. Draft genome sequences of two super-XDR isolates of *Mycobacterium tuberculosis* from China[J]. FEMS Microbiology Letters, 2013, 347(2), 93-96.
- [13] Aust AE, Eveleigh JF. Mechanisms of DNA oxidation[J]. Experimental Biology and Medicine, 1999, 222(3): 246-252.
- [14] Wang ZY, Xiong M, Fu LY, et al. Oxidative DNA damage is important to the evolution of antibiotic resistance: evidence of mutation bias and its medicinal implications[J]. Journal of Biomolecular Structure and Dynamics, 2013, 31(7): 729-733.
- [15] Comas I, Chakravarti J, Small PM, et al. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved[J]. Nature Genetics, 2010, 42(6): 498-503.
- [16] Drobniewski F, Balabanova Y, Nikolayevsky V, et al. Drug-resistant tuberculosis, clinical virulence, and the dominance of the Beijing strain family in Russia[J]. The Journal of the American Medical Association, 2005, 293(22): 2726-2731.
- [17] Parwati I, van Crevel R, van Soolingen D. Possible underlying mechanisms for successful emergence of the *Mycobacterium tuberculosis* Beijing genotype strains[J]. The Lancet Infectious Diseases, 2010, 10(2): 103-111.
- [18] Ford CB, Shah RR, Maeda MK, et al. *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of

- drug-resistant tuberculosis[J]. *Nature Genetics*, 2013, 45(7): 784-790.
- [19] Rad ME, Bifani P, Martin C, et al. Mutations in putative mutator genes of *Mycobacterium tuberculosis* strains of the W-Beijing family[J]. *Emerging Infectious Diseases*, 2003, 9(7): 838.
- [20] Zhang H, Li D, Zhao L, et al. Genome sequencing of 161 *Mycobacterium tuberculosis* isolates from China identifies genes and intergenic regions associated with drug resistance[J]. *Nature Genetics*, 2013, 45(10): 1255-1260.
- [21] Sandgren A, Strong M, Muthukrishnan P, et al. Tuberculosis drug resistance mutation database[J]. *PLoS Medicine*, 2009, 6(2): e2.
- [22] Maus CE, Plikaytis BB, Shinnick TM. Molecular analysis of cross-resistance to capreomycin, kanamycin, amikacin, and viomycin in *Mycobacterium tuberculosis*[J]. *Antimicrobial Agents and Chemotherapy*, 2005, 49(8): 3192-3197.
- [23] Maus CE, Plikaytis BB, Shinnick TM. Mutation of *thyA* confers capreomycin resistance in *Mycobacterium tuberculosis*[J]. *Antimicrobial Agents and Chemotherapy*, 2005, 49(2): 571-577.
- [24] Zaunbrecher MA, Sikes RD Jr, Metchock B, et al. Overexpression of the chromosomally encoded aminoglycoside acetyltransferaseeis confers kanamycin resistance in *Mycobacterium tuberculosis*[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(47): 20004-20009.
- [25] Farhat MR, Shapiro BJ, Kieser KJ, et al. Genomic analysis identifies targets of convergent positive selection in drug-resistant *Mycobacterium tuberculosis*[J]. *Nature Genetics*, 2013, 45(10): 1183-1189.
- [26] Warner DF, Mizrahi V. Complex genetics of drug resistance in *Mycobacterium tuberculosis*[J]. *Nature Genetics*, 2013, 45(10): 1107-1108.
- [27] Raman K, Chandra N. *Mycobacterium tuberculosis* interactome analysis unravels potential pathways to drug resistance[J]. *BMC Microbiology*, 2008, 8: 234.
- [28] Pasca MR, Gugliera P, Arcesi F, et al. Rv2686c-Rv2687c-Rv2688c, an ABC fluoroquinolone efflux pump in *Mycobacterium tuberculosis*[J]. *Antimicrobial Agents and Chemotherapy*, 2004, 48(8): 3175-3178.
- [29] Hershberg R, Lipatov M, Small PM, et al. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography[J]. *PLoS Biology*, 2008, 6(12): e311.
- [30] Yang C, Luo T, Sun G, et al. *Mycobacterium tuberculosis* Beijing strains favor transmission but not drug resistance in China[J]. *Clinical Infectious Diseases*, 2012, 55(9): 1179-1187.
- [31] Tsolaki AG, Gagneux S, Pym AS, et al. Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of *Mycobacterium tuberculosis*[J]. *Journal of Clinical Microbiology*, 2005, 43(7): 3185-3191.
- [32] Muller B, Borrell S, Rose G, et al. The heterogeneous evolution of multidrug-resistant *Mycobacterium tuberculosis*[J]. *Trends in Genetics: TIG*, 2013, 29(3): 160-169.
- [33] Ryan BM, Dougherty TJ, Beaulieu D, et al. Efflux in bacteria: what do we really know about it?[J]. *Expert Opinion on Investigational Drugs*, 2001, 10(8): 1409-1422.
- [34] Jennifer LG, James CJ, Shannan J Ho S, et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak[J]. *The New England Journal of Medicine*, 2011, 364(8): 730-739.
- [35] Timothy MW, Camilla LCIP, Ruth HH, et al. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study[J]. *The Lancet Infectious Diseases*, 2013, 13(2): 137-146.