

DOI: 10.3969/j.issn.1671-9638.2017.01.001

· 论 著 ·

全基因组测序分析耐甲氧西林金黄色葡萄球菌利奈唑胺耐药相关变异位点

姚伟明, 陈 重, 蒲彰雅, 王红燕, 程 航, 李多云, 郑金鑫, 邓向斌, 刘晓军, 邓启文, 余治健

(深圳大学南山区人民医院 内源性感染诊治研究重点实验室, 广东 深圳 518052)

[摘要] 目的 通过全基因组测序研究利奈唑胺(LZD)敏感株和诱导耐药耐甲氧西林金黄色葡萄球菌(MRSA)之间基因变异位点差异,了解 LZD 耐药基因变异位点。方法 采用不同浓度梯度的 LZD 对遗传背景清晰的 MRSA-MS4 母株进行诱导,获得 LZD 耐药菌株 MRSA-MS4-LZD100,测定最低抑菌浓度(MIC),采用普通聚合酶链反应(PCR)扩增 MRSA-MS4-LZD100 23S rRNA V 区及核糖体蛋白 L3/L4 基因,测序后与野生株比较,获得相应的突变位点;采用 Illumina HiSeq 2000 测序技术对样品 DNA 进行 paired-end(PE)测序,构建 Illumina PE 文库,利用生物信息学完成该菌株的全基因组测序。结果 经过 32 代诱导,获得 MRSA-MS4-LZD100 菌株,其 LZD MIC 为 96 $\mu\text{g}/\text{mL}$ 。PCR 测序分析提示其 V 区多拷贝基因均存在 G2447T 突变,L3 蛋白存在 Gly113Val 突变;全基因组包含 2 744 315 bp 碱基对,注释后共有 2 509 个基因,11 个 tRNA 编码基因,以及 2 个完整的 rRNA 基因编码操纵子,获得 PubMed 全基因组序列号(JXMJ00000000);共找出 101 个 SNP 和 6 个 Small indel 突变,发生于外显子的 SNP 占 16 个,SNP 后导致氨基酸序列改变的蛋白质包括 IstB ATP 结合区域包含蛋白、凝集因子 A 及转座子 IS1272 等,而发生于外显子的 Small indel 占 3 个,Small indel 突变后导致氨基酸序列改变的蛋白质包括假设蛋白、30S 核糖体蛋白 S1、凝集因子 A。结论 经 LZD 诱导获得 LZD 耐药 MRSA-MS4-LZD100 菌株,测序分析提示该菌株存在除 23S rRNA V 区及 L3 蛋白基因以外的突变位点,为进一步研究隐性 LZD 耐药机制指明了方向。

[关键词] 耐甲氧西林金黄色葡萄球菌;利奈唑胺;高通量测序;全基因组序列

[中图分类号] R378.1⁺1 **[文献标识码]** A **[文章编号]** 1671-9638(2017)01-0001-05

Whole genome sequencing for analyzing mutation sites in linezolid-resistant methicillin-resistant *Staphylococcus aureus*

YAO Wei-ming, CHEN Zhong, PU Zhang-ya, WANG Hong-yan, CHENG Hang, LI Duo-yun, ZHENG Jin-xin, DENG Xiang-bin, LIU Xiao-jun, DENG Qi-wen, YU Zhi-jian (Department of Infectious Diseases and Key Laboratory for Research on Endogenous Infections, Nanshan District People's Hospital of Shenzhen University, Shenzhen 518052, China)

[Abstract] **Objective** To understand genetic mutation sites in linezolid (LZD)-sensitive and inducible resistant strains of methicillin-resistant *Staphylococcus aureus* (MRSA) using whole-genome sequencing, and realize mutation sites of LZD-resistant gene. **Methods** MRSA-MS4 with explicit genotype and whole-genome sequences was induced by LZD of different concentration gradients, LZD-resistant strain MRSA-MS4-LZD100 was obtained, minimum inhibitory concentration(MIC) was detected, domain V of 23S rRNA and ribosomal proteins L3/L4 gene in MRSA-MS4-LZD100 were amplified by polymerase chain reaction (PCR), the sequenced products obtained the corresponding mutation site in contrast with the wild-type strain; Illumina PE library was constructed through paired-end sequencing by Illumina HiSeq 2000 technique, and whole genome sequencing was completed based on bioinformatics.

Results MRSA-MS4-LZD100 strain was induced after 32 passages, MIC of LZD was 96 $\mu\text{g}/\text{mL}$. Sequencing of

[收稿日期] 2016-08-26

[基金项目] 深圳市重点学科资助项目(201506093);深圳市科技创新课题(No. JCYJ20150402152130173, No. JCYJ20150402152130167);深圳市重点学科经费资助;深圳市南山区课题资助(No. 2015019, 2015022)

[作者简介] 姚伟明(1988-),男(汉族),湖南省岳阳市人,医师,主要从事细菌耐药机制研究。

[通信作者] 余治健 E-mail: yuzhijiansmu@163.com

PCR products indicated the genetic variations were G2447T mutation in multiple copies of domain V of 23S rRNA gene, and Gly113Val mutation in L3 protein respectively; the whole genome of MRSA-MS4-LZD100 contained 2 744 315 bp, annotation of the whole genome found a total of 2 509 genes, 11 tRNA-encoding genes and 2 entire rRNA-encoding operons. The data were submitted to the PubMed, and the GeneBank accession number JXMJ000000000 was assigned; a total of 101 SNPs and 6 Small indels were found, 16 of 101SNP mutations occurred in exon, of which the variant proteins with amino acid sequence alterations included IstB ATP binding domain-containing protein, clumping factor A, IS1272 transposase and so on; 3 of 6 Small indel mutations occurred in exon, of which the variant proteins with amino acid sequence alterations included hypothetical protein, 30S ribosomal protein S1, and clumping factor A. **Conclusion** LZD-resistant strain MRSA-MS4-LZD100 was successfully induced by LZD; beside 23S rRNA V domain and ribosomal L3 protein, the other mutant site exist in this resistant strain, which provide some direction for subsequent study of recessive LZD resistance mechanism.

[**Key words**] methicillin-resistant *Staphylococcus aureus*; linezolid; high-throughput sequencing; complete genome sequence

[Chin J Infect Control, 2017, 16(1): 1-5]

以利奈唑胺(linezolid, LZD)为代表的噁唑烷酮类(oxazolidinone)药物是目前多重耐药革兰阳性细菌治疗的“最后防线”,已成为治疗耐甲氧西林金黄色葡萄球菌(methicillin-resistant *Staphylococcus aureus*, MRSA)感染,控制 MRSA 医院感染暴发流行的主要药物之一。然而,随着该药大规模上市使用,耐药现象日益突出,LZD 耐药 MRSA 的感染已成为一种新的难治性感染,其耐药机制也已成为研究热点之一。目前,关于 LZD 的耐药机制主要有 3 个方面:一是细菌 23S rRNA V 区多拷贝基因突变;二是细菌核糖体 23S rRNA 修饰如甲基转移酶基因 *cfr* 的获得;三是编码细菌 50S 核糖体蛋白基因 L3 和 L4 的变异^[1]。近年来,随着全基因组测序技术的日趋成熟和推广,对 MRSA 抗菌药物耐药相关机制正逐渐被揭示出来,对细菌染色体编码的甲基转移酶、ABC 转运子等耐药基因和各种毒力基因的进化和演变均有更清晰的认识^[2-3]。本研究拟用 LZD 对我院临床样本来源的、基因分型和全基因组序列清晰的 MRSA-MS4 母株进行耐药诱导,产生的耐药菌株通过 paired-end(PE)高通量测序,并与母株基因组序列进行比较,以期获得完整的基因组精细信息,以便发现 LZD 耐药的基因变异特点。

1 对象与方法

1.1 研究对象 分离自我院患者创口脓液的耐甲氧西林金黄色葡萄球菌母株(MRSA-MS4),已完成其 SCCmec-spa-MLST 分型(III-t3592-ST59 型)和全基因组测序,并获得 GeneBank 号(CP009828)。体外多步法诱导 LZD 耐药:将 MRSA-MS4 接种于

血琼脂平板培养基中复苏,挑取多个单克隆菌落,依次接种到 LZD 浓度(起始浓度为 1 $\mu\text{g}/\text{mL}$)成倍递增的 M-H 培养基中,37 $^{\circ}\text{C}$ 过夜培养 16 h,每个药物浓度培养 3~5 代,当细菌生长不良时降低 LZD 浓度重复传代培养,每个药物浓度最多培养 10 代。待药物浓度为 4 $\mu\text{g}/\text{mL}$ 时反复传代数次后(不超过 10 代),逐渐成倍增加浓度至 256 $\mu\text{g}/\text{mL}$,筛选阳性克隆并对诱导出的耐药菌株使用 LZD E-test 条测定菌株最低抑菌浓度(MIC),最终鉴定得到耐甲氧西林金黄色葡萄球菌 LZD 耐药菌株 MRSA-MS4-LZD100。

1.2 DNA 提取 将 MRSA-MS4-LZD100 接种于血琼脂平板培养基中复苏,挑取单克隆菌落,接种到 LB 液体培养基中,37 $^{\circ}\text{C}$ 振荡培养至对数生长中期,通过商品化的 DNA 提取试剂盒(Wizard Genomic DNA Purification Kit, Promega)提取基因组 DNA。

1.3 23S rRNA V 区、核糖体蛋白 L3/L4 扩增和测序 采用聚合酶链反应(PCR)方法扩增该菌株的 23S rRNA 的 6 个同分异构体和核糖体蛋白 L3/L4 基因,引物及 PCR 扩增反应均见参考文献[4]。Marker 为 DL10000 DNA Marker(TaKaRa 公司),PCR 产物送华大基因公司纯化后测序,测序结果与野生株序列进行比对,分析软件为 DNAMAN, full version, v5. 2. 2,标记变异的核苷酸及氨基酸位点。

1.4 基因组测序、注释及序列分析 对提取的全基因组 DNA 采用 Illumina HiSeq 2000 进行测序,具体工作交由上海美吉生物医药科技有限公司完成。简要步骤如下:应用 Illumina HiSeq 2000 技术对样品的 DNA 进行 PE 测序,构建 300 bp 文库;对测序结果进行质量剪切后,利用 SOAPdenovo v2. 04(ht-

tp://soap.genomics.org.cn/) 拼接软件对优化序列进行多个 Kmer 参数的拼接, 得到最优的组装结果。其次, 应用 GapCloser v1.12 软件对组装结果进行局部内洞填充和碱基校正; 应用 Barrnap 0.4.2 和 tRNAscan-SE v1.3.1 软件预测获得的拼接结果基因组中包含的 rRNA 和 tRNA, 用 Glimmer 3.02 (<http://www.cbcb.umd.edu/software/glimmer/>) 软件进行细菌的基因预测, 将预测基因的蛋白序列分别与 Nr、genes、string 和 GO 数据库进行 blastp 比对 (BLAST 2.2.28+), 从而获得预测基因的注释信息; 以细菌基因组 MRSA-MS4 为参考序列, 利用 BWA 比对软件将测序片段比对回参考基因组序列, 利用 GATK 校正比对结果, 应用 VarScan 进行 SNP、small indel 的检测, 并过滤掉测序深度和比对质量值较低的位点, 得到高可信度的 SNP 和 small indel 数据集。

2 结果

2.1 MRSA-MS4 LZD 耐药性诱导 在体外诱导 LZD 耐药试验中逐渐成倍增加 LZD 浓度至 256 $\mu\text{g}/\text{mL}$, 当 MS4 传代到 32 代时, 筛选得到阳性克隆菌株, 使用 LZD E-test 条测定菌株 MIC, 最终鉴定得到 LZD 耐药菌株 MRSA-MS4-LZD100, 其 LZD MIC 为 96 $\mu\text{g}/\text{mL}$ 。

2.2 MRSA-MS4-LZD100 23S rRNA V 区及 L3/L4 蛋白基因测序结果 通过普通 PCR 扩增 MRSA-MS4-LZD100 23S rRNA V 区及 L3/L4 蛋白基因, 所得结果与母株序列进行比对, 结果提示所有 23S rRNA V 区均存在 G2447T 位点变异, 核糖体 L3 蛋白存在 Gly113Val 位点变异, 而核糖体 L4 蛋白未发现突变位点。

2.3 MRSA-MS4-LZD100 基因组概述 通过 Illumina HiSeq 2000 测序技术进行测序提示, 完整的 MRSA-MS4-LZD100 基因组大小为 2 744 315 bp, GC 含量 33.5%; 总共有 2 509 个基因, 11 个 tRNA 的编码基因, 2 个完整的 rRNA 基因编码操纵子。对预测基因进行 COG 功能分类预测, 共有 2 013 个 Gene 被分为 22 种 COG 分类; 共有 1 805 个基因注释上 42 条 GO 功能分类, 样本基因的功能在 biological process 分类中主要聚集于 cellular process 和 metabolic process; 在 cellular component 主要聚集于 cell 和 cell part; 在 molecular function 分类中主要聚集于 binding 和 catalytic activity。对基因进

行 KEGG 注释, 1 424 个基因共注释上 153 个 Pathway, 最多见的 5 个 pathway 注释分别是代谢通路 (metabolic pathways)、继发代谢产物生物合成通路 (biosynthesis of secondary metabolites)、不同环境中微生物代谢通路 (microbial metabolism in diverse environments)、氨基酸生物合成通路 (biosynthesis of amino acids)、ABC 转运蛋白通路 (ABC transporters)。目前该菌株全基因组序列已上传至 PubMed 数据库, GeneBank 号: JXMJ00000000。

2.4 SNP、Small indel 注释 应用 VarScan 进行 SNP、small indel 检测, 结果显示: 共有 101 个 SNP 和 6 个 Small indel 突变, 其中 SNP 发生于外显子的同义突变有 12 个, 非同义突变有 4 个, 而 Small indel 发生于外显子的框移插入有 2 个, 非框移插入有 1 个。见表 1。SNP 突变后导致氨基酸序列改变的蛋白质包括 IstB ATP 结合区域包含蛋白、凝集因子 A 及转座子 IS1272 等, 而 Small indel 突变后导致氨基酸序列改变的蛋白质包括有假设蛋白、30S 核糖体蛋白 S1、凝集因子 A 等。见表 2。

表 1 MRSA-MS4-LZD100 SNP、Small indel 注释结果
Table 1 Annotation results of SNP and Small indel in MRSA-MS4-LZD100

位置	SNP 突变	Small indel 突变
下游	16	
外显子	16	3
上游	1	
上游; 下游	2	3
基因间区	66	
合计	101	6

表 2 高通量测序分析 MRSA-MS4-LZD100 基因组变异结果
Table 2 Genome variation results of MRSA-MS4-LZD100 through high-throughput sequencing

基因名称	MRSA-MS4 位点	MRSA-MS4-L100 突变类型
IstB ATP 结合区域包含蛋白	orf00076	G379T A127S
凝集因子 A	orf00740	A1991G N664S
转座酶 IS1272	orf02107	T852C G284G
凝集因子 A	orf00740	2179-2196del 727-732del
假设蛋白	orf00074	938insA E313fs
30S 核糖体蛋白 S1	orf00332	937-938insAE313fs

3 讨论

MRSA 之所以常表现为多重耐药, 是因其能同时携带多种耐药相关基因, 如耐氨基糖苷类药物的

aac(6')/aph(2'')、aph(3')-III、ant(4',4'')基因、产 β -内酰胺酶基因 TEM、编码 PBP2a 的 *mecA* 基因、耐红霉素 *erm*、耐四环素 *tetM*、耐万古霉素 *vanA*、耐消毒剂 *qacA/B* 基因等,且检出率高^[5];而 MRSA 感染之所以常导致较高的致死致残率,则是因其可携带众多毒性强大的毒力基因,如杀白细胞素(PVL)、肠毒素 C(sec)、肠毒素 H(seh)、 α -溶血素(hla)、 β -溶血素(hlb)、黏附素凝聚因子(*clfA*、*clfB*)、纤连蛋白结合蛋白(*fnbA*、*fnbR*)、酚溶调制肽(PSM)等^[6]。毒力基因的高检出率与菌株致病性高度相关,而耐药基因的高检出率则与菌株多药耐药表型一致,因此,对 MRSA 毒力及耐药基因进行深入研究,对指导临床诊断和治疗具有重要的意义。本研究以临床样本来源、基因分型及全基因组序列清晰的 MRSA-MS4 为标准株,体外进行 LZD 耐药诱导。通过全基因组高通量测序,旨在发现更多的与 LZD 耐药相关因子,为解决临床细菌耐药突破提供更多的参考依据。

细菌 23S rRNA V 区位点突变以及核糖体蛋白 L3/L4 编码基因 *rplC/rplD* 变异为目前明确导致 LZD 耐药的分子机制^[1]。体外研究提示,细菌产生 LZD 耐药是一个缓慢的过程,获得高水平 LZD 耐药性需长时间暴露于治疗浓度,同时可能需要其他多种蛋白(包括核糖体蛋白等)的共同参与^[7],而部分细菌(如结核分枝杆菌和耻垢分枝杆菌等)体外诱导试验发现所获得的 LZD 低水平耐药菌株均与核糖体 23S rRNA V 区突变无关,因此基于实验结果,推测发生 LZD 耐药可能存在其他机制^[8-9]。本例诱导 LZD 耐药的 MRSA 通过普通 PCR 扩增 23S rRNA V 区及 L3/L4 蛋白基因后与母株序列相比对,发现所有 23S rRNA V 区均存在 G2447T 位点变异,此外核糖体 L3 蛋白也存在 G113V 突变,提示 V 区及 L3 突变仍为 LZD 耐药的主要机制,与先前报道^[4]一致。而 SNP、Small indel 检测及注释结果显示,共有 101 个 SNP 和 6 个 Small indel 突变,其中 SNP 发生于外显子的同义突变有 12 个,非同义突变有 4 个,而 Small indel 发生于外显子的框移插入有 2 个,非框移插入有 1 个;SNP 突变后导致氨基酸序列改变的蛋白质包括 IstB ATP 结合区域包含蛋白、凝集因子 A 及转座子 IS1272 等,而 Small indel 突变后导致氨基酸序列改变的蛋白质包括假设蛋白、30S 核糖体蛋白 S1、凝集因子 A 等。

凝集因子 A 为金黄色葡萄球菌重要的毒力因子之一,与组织中的纤维蛋白原结合后,介导细菌黏附于感染部位,且通过被纤维蛋白原包裹,而使细菌抵抗调理素接近和沉着,从而起到抗吞噬作用,现已成为抗感染疫苗的研究热点^[10]。IS1272 插入序列可介导高分子量低亲和力的青霉素结合蛋白 PBP2a 在肠球菌与葡萄球菌及不同葡萄球菌亚种之间的结合转移,在细菌种属之间耐药基因的进化过程中扮演重要作用^[11]。30S 核糖体蛋白 S1 属于核糖体蛋白 SIP 家族,其功能为结合 mRNA,有利于识别起始点,并与一个富含嘌呤的短 SD 序列结合翻译 mRNA。姜葵等^[12]应用蛋白组学技术研究幽门螺杆菌对不同水平克拉霉素耐药的机制表明,30S 核糖体蛋白 S1 在耐药菌株中表达较野生株下调。而谢丽玲等^[13]研究发现大黄醇提液、甘草醇提液和金银花醇提液可通过抑制嗜水气单胞菌 30S 核糖体蛋白 S1 等蛋白的表达,起到抑制嗜水气单胞菌生长的作用。然而,目前关于上述蛋白突变后与 LZD 耐药的研究尚未见报道,二者之间的具体关系仍不明确,本研究的结果为后续进一步深入了解 LZD 耐药机制提供了探索方向。

目前,快速发展的全基因组测序技术能全面、精确的分析病原微生物全基因组的碱基序列,破解其所包含的所有遗传信息,如微生物的耐药、毒力相关因子等,并通过与不同个体或群体的比对,发现二者之间的遗传特性差异及基因突变特点,从而加深对各种疾病发生、发展的认识,并制定相应的治疗策略,提高疾病信息及其特征的检出率,也给临床带来了很大的机遇和挑战^[14]。相对而言,全基因组测序仍因工作量大,耗费时间长,费用较高,测序后生物信息学分析更是错综复杂,故目前尚难以广泛应用于临床各种疾病的诊断,但在科研领域已取得了举足轻重的地位^[15]。本研究尚未对上述筛选得到的、有潜在意义的突变候选基因进行基因重组与转染等功能学研究,以最终确立其与 LZD 耐药发生的关系,但相关工作目前正在进行阶段。

总之,通过对标准菌株进行耐药诱导,产生耐药菌株后行全基因精细测序,显示其存在的基因变异位点,并对变异位点基因进行重组等后续功能学研究,是一条较为理想和可行的技术路线,可应用于多种药物耐药相关机制的研究,以不断提高医务人员对药物耐药机制的认识和丰富药物耐药的临床资料库。

[参 考 文 献]

- [1] Long KS, Vester B. Resistance to linezolid caused by modifications at its binding site on the ribosome[J]. Antimicrob Agents Chemother, 2012, 56(2): 603-612.
- [2] Billal DS, Feng J, Leprohon P, et al. Whole genome analysis of linezolid resistance in *Streptococcus pneumoniae* reveals resistance and compensatory mutations[J]. BMC Genomics, 2011, 12: 512.
- [3] Feng J, Lupien A, Gingras H, et al. Genome sequencing of linezolid-resistant *Streptococcus pneumoniae* mutants reveals novel mechanisms of resistance[J]. Genome Res, 2009, 19(7): 1214-1223.
- [4] Locke JB, Hilgers M, Shaw KJ. Novel ribosomal mutations in *Staphylococcus aureus* strains identified through selection with the oxazolidinones linezolid and torezolid (TR-700)[J]. Antimicrob Agents Chemother, 2009, 53(12): 5265-5274.
- [5] 文细毛, 任南, 吴安华, 等. 全国医院感染监控网细菌耐药情况及变化趋势[J]. 中国感染控制杂志, 2009, 8(6): 389-396.
- [6] 于淼, 赵自云, 刘成玉, 等. 医院获得性耐甲氧西林金黄色葡萄球菌中 PSM- α 基因检测及相关研究[J]. 中国感染控制杂志, 2015, 14(7): 443-446.
- [7] Bøsling J, Poulsen SM, Vester B, et al. Resistance to the peptidyl transferase inhibitor tiamulin caused by mutation of ribosomal protein L3[J]. Antimicrob Agents Chemother, 2003, 47(9): 2892-2896.
- [8] Hillemann D, Rüscher-Gerdes S, Richter E. In Vitro-selected linezolid-resistant *Mycobacterium tuberculosis* mutants[J]. Antimicrob Agents Chemother, 2008, 52(2): 800-801.
- [9] Richter E, Rüscher-Gerdes S, Hillemann D. First linezolid-resistant clinical isolates of *Mycobacterium tuberculosis*[J]. Antimicrobial Agents Chemother, 2007, 51(4): 1534-1536.
- [10] 向华, 才红军. 金黄色葡萄球菌凝集因子 A 研究进展[J]. 中国兽药杂志, 2013, 47(8): 62-65.
- [11] Archer GL, Thanassi JA, Niemeyer DM, et al. Characterization of IS1272, an insertion sequence-like element from *Staphylococcus haemolyticus* [J]. Antimicrob Agents Chemother, 1996, 40(4): 924-929.
- [12] 姜葵, 赵飞, 王邦茂, 等. 幽门螺杆菌 26695 及其不同克拉霉素耐药水平衍生株的比较蛋白质组学分析[J]. 胃肠病学, 2007, 12(9): 545-550.
- [13] 谢丽玲, 唐伟, 李培培, 等. 三种中药的醇提液对嗜水气单胞菌的抑制作用(英文)[J]. 中国生物化学与分子生物学报, 2014, 30(10): 1007-1016.
- [14] Klemm E, Dougan G. Advances in understanding bacterial pathogenesis gained from whole-genome sequencing and phylogenetics[J]. Cell Host Microbe, 2016, 19(5): 599-610.
- [15] 郭奕斌. 基因诊断中测序技术的应用及优缺点[J]. 遗传, 2014, 36(11): 1121-1130.

(本文编辑:左双燕)

· 信 息 ·



关注中国感染控制杂志微信公众号
及时了解世界感染控制新进展