



论文

类 SCC 样甲氧西林敏感性金黄色葡萄球菌的基因组学分析

李德志^{①②†}, 楚亚男^{③④†}, 任鲁风^{③④}, 李欣刚^{③④}, 袁丽娜^{③④}, 康禹^③, 张伟^⑤, 杨宇^⑥, 王绪敏^③, Baillie J. Kenneth^{⑦⑧}, 于军^{③④*}, 高占成^{①*}

① 北京大学人民医院呼吸与危重症医学科, 北京 100044;

② 山东大学附属省立医院呼吸科, 济南 250021;

③ 中国科学院北京基因组研究所基因组科学及信息重点实验室, 北京 100029;

④ 中国科学院北京基因组研究所 DNA 序列测定技术研究开发中心, 北京 100029;

⑤ 南昌大学第一附属医院呼吸科, 南昌 330006;

⑥ 中国检验检疫科学研究院, 北京 100025;

⑦ Division of Genetics and Genomics, The Roslin Institute, University of Edinburgh, Roslin, EH25 9RG, UK;

⑧ Department of Critical Care Medicine, University of Edinburgh, Edinburgh, EH16 4TJ, UK

† 同等贡献

* 联系人, E-mail: junyu@big.ac.cn; zcgao@bjmu.edu.cn

收稿日期: 2012-10-30; 接受日期: 2012-12-10

国家高技术研究发展计划(批准号: 2006AA02Z4A9)、国家科技部重大专项(批准号: 2009ZX10004, 2012ZX10004206)、国家自然科学基金(批准号: 30971610, 30900053)和中国科学院青年创新促进会基金资助项目

摘要 金黄色葡萄球菌(*Staphylococcus aureus*, *S. aureus*), 作为一种常见的致病菌常可导致严重感染性疾病. 金黄色葡萄球菌的主要特点是易产生耐药性. 临床上运用多重 PCR 方法对 *S. aureus* 耐药菌株进行分型. SCC 作为一种 *mecA* 基因的转载体, 介导金黄色葡萄球菌产生耐药性. 本研究报道了一株新甲氧西林敏感性金黄色葡萄球菌的全基因组序列信息, 即类 SCC 样 MSSA463 菌株, 经多重 PCR 方法, 该菌株被误鉴定为耐甲氧西林金黄色葡萄球菌, 结果与药敏结果相冲突. 为此, 运用焦磷酸测序方法完成了该菌株的全基因组测序, 并与已知金黄色葡萄球菌的基因组序列信息不同, 发现可读框(CZ049; AB037671)存在于 *attL*, *attR* 反向重复序列的临近下游. 这些结果提示, 在金黄色葡萄球菌与其他微生物之间可能存在一种水平基因转移, 并改变了金黄色葡萄球菌对药物的敏感性, 推测 *attL*, *attR* 反向重复序列可能作为外源性基因的插入部位而存在.

关键词

金黄色葡萄球菌
金黄色葡萄球菌
盒式染色体
mec(SCC *mec*)
基因组学
焦磷酸测序

引用格式: 李德志, 楚亚男, 任鲁风, 等. 类 SCC 样甲氧西林敏感性金黄色葡萄球菌的基因组学分析. 中国科学: 生命科学, 2013, 43: 247-253, doi: 10.1360/scis-2012-0161.R4

英文版见: Li D Z, Chu Y N, Ren L F, et al. Complete genome sequence of methicillin-sensitive *Staphylococcus aureus* containing a heterogeneic staphylococcal cassette chromosome element. Sci China Life Sci, 2013, 56: 268-274, doi: 10.1007/s11427-013-4453-9

金黄色葡萄球菌(*Staphylococcus aureus*, *S. aureus*)作为一种常见的病原菌经常会导致严重感染的发生, 这些感染包括坏死性肺炎、脓毒血症等。近年来, 耐甲氧西林金黄色葡萄球菌及耐万古霉素、耐利奈唑烷抗药性菌株的出现限制了人们对抗生素的选择^[1-3]。因此, 迫切需要对金黄色葡萄球菌耐药性分子机制进行更深入的研究。

金黄色葡萄球菌的一个主要特点在于其可以获得抗生素抗性。迄今为止, 金黄色葡萄球菌盒式染色体(SCC)结构是目前所知的唯一 *mecA* 基因的转载体, 该基因编码了葡萄球菌菌属对甲氧西林的药物抗药性。研究证实, SCCmec 结构的水平转移导致了细菌耐药性的获得^[4]。很长一段时间, 人们一直将 SCCmec 作为一个整体来看待, 然而, 越来越多的研究发现, 在甲氧西林敏感性菌株(MSSA)中依然可以检测到 SCC 组分的存在, 暗示 SCC 结构组成中耐药基因组在菌株中丢失的可能, 继而使得耐甲氧西林金黄色葡萄球菌菌株(MRSA)转变为甲氧西林敏感性菌株(MSSA)^[5]。高通量全基因组测序的方法, 尤其是新一代测序平台为研究细菌耐药性分子机制提供了一个全新的方法。已有研究通过对一株 MSSA 菌株 MSSA476 测序发现, 该菌株携带一个新的 SCCmec 样组分(命名为 SCC476), 该组分所整合的基因组位置与 MRSA 菌株整合的位置相同, 但是包含了编码夫西地酸钠抗性蛋白的基因^[6]。这些研究结果为探究 MSSA 细菌中 SCCmec 样组分的分子基础提供了有益的信息。

在前期多重 PCR 方法获得的分子流行病学信息中发现, MSSA463 菌株存在 85/2082 菌株的 ORF CZ049(AB37671)。该结构曾被认为是 SCCmecIII 型 MRSA 菌株所特有。这种结构的出现也暗示在 MSSA463 菌株中存在着片段的转移或其他结构的改变^[7]。本研究通过全基因组测序方法, 分析了 SCC 样 MSSA 菌株的主要特征, 进一步探讨了金黄色葡萄球菌耐药性获得与丢失的分子机制。

1 材料与方法

1.1 伦理声明

本实验获得北京大学人民医院伦理委员会审核批准(批准号: 00001384), 同时获得患者知情授权。

该项目按照《赫尔辛基宣言》执行。

1.2 菌株培养与鉴定

在前期分子流行病学信息研究中检测到 MSSA 463 菌株含有 SCCmecIII 型 MRSA 菌株特有的 ORF CZ049(AB037671)(MRSA85/2082)片段, 因此命名为 SCC 样 MSSA463 菌株。该菌株分离自南昌一位 74 岁患者, 临床诊断为慢性阻塞性肺病急性加重合并支气管扩张和社区获得性肺炎。该菌株对多数常用抗生素敏感, 如阿莫西林/克拉维酸、氨苄西林/舒巴坦、头孢氨噻肟、头孢曲松、环丙沙星、克林霉素、红霉素、夫西地酸、庆大霉素、亚胺培南、左氧氟沙星、利奈唑胺、莫西沙星、苯唑西林、替考拉宁、四环素、甲氧苄啶/磺胺和万古霉素。但是对氨苄青霉素和青霉素 G 有耐药性(网络版附表 1)。

1.3 基因组测序与拼接

DNA 提取后采用罗氏 454GSFLX 焦磷酸测序平台完成全基因组测序^[8], 全基因组覆盖度达到 70×, 使用配套拼接软件 Newbler 完成序列拼接。利用 3730 测序技术完成全基因组组装与补洞, 完成全基因组图谱绘制。使用 Glimmer3.02 软件进行可读框(open reading frames, ORFs)预测^[9], 运用 BLASTP 将所有预测蛋白序列与 NCBI 非冗余蛋白数据库进行比对, *E* 值设定为 1×10^{-5} 。通过 Swissprot 数据库(<http://www.psc.edu/general/software/packages/swiss/swiss.php>, 20091208 版, *E* 值设定 1×10^{-10})^[10]、京都基因与基因组百科全书数据库(Kyoto Encyclopedia of Genes and Genomes, KEGG)(Release 50, *E* 值设定 1×10^{-10})^[11]、直系同源基因簇数据库(Clusters of Orthologous Groups, COG)(*E* 值设定 1×10^{-10})、Interpro 数据库(Interproscan 4.3, Release16.0)^[12]和 Gene Ontology (GO)数据库^[13](网络版附表 2)等对全基因组序列进行注释。

1.4 比较基因组学分析

采用软件 Artemis Comparison Tool^[14], 完成基因组间的比较分析。所选取的参考序列包括 MSSA476 (社区获得性 MSSA), MRSA252(GenBank 登录号: BX571856.1, 医院获得性 MRSA), Mu50(GenBank 登录号: BA000017.4, 医院获得性 VISA), N315 (GenBank 登录号: BA000018.3, 医院获得性 MRSA)

和 USA300_FPR3757(GenBank 登录号: CP000255.1, 社区获得性 MRSA)等。

2 结果

2.1 基因组拼接与注释

使用 GS *De Novo* Assembler 软件^[8]完成全基因组序列拼接(原始序列平均读长 352 bp, 覆盖度达 70x(网络版附表 3), 拼接获得 33 个 contigs(其中 22 个大于 500 kb), 覆盖 98%的全基因组序列(网络版附表 3)。

经过补洞测序后, MSSA463 菌株全基因组全长测定为 2771498 bp, 预测到 2563 可读框; 基因分布密度为 0.92 kb, 重复序列占全基因组长度的 0.93%, G+C 含量为 32.45%, 符合典型的厚壁菌门特征。基因组中最长 ORF 为 28236 bp, 基因平均长度为 903 bp, 包含 5 个 rRNA 操纵子, 59 个 tRNA 及 29 个 small RNA 编码序列(网络版附表 4 和图 1)。

CDS 对应编码基因在 COG 数据库中进行同源基

因注释分类, 得到包括细胞生物学过程及信号传导、信息的储存与处理、代谢过程相关编码基因及未知功能的编码基因(网络版附图 1)。GO 完成所预测基因在细胞定位、分子功能及基因所参与的生物学过程进行聚类分析的结果见网络版附图 2。

2.2 全基因组比对

应用 Artemis Comparison Tool 软件^[14], 将 MSSA463 全基因组与其他 4 株菌株进行全基因组比对: MSSA476, 社区获得性 MSSA; Mu50(GenBank 登录号: BA000017.4, 医院获得性 VISA); N315 (GenBank 登录号: BA000018.3), 医院获得性 MRSA; USA300_FPR3757(GenBank 登录号: CP000255.1), 社区获得性 MRSA。所用菌株的基因组大小均在 2.8 Mb 左右, G+C 含量 32%。MSSA463 菌株中有 2235 个基因(87.2%)在其他 4 种菌株中均有, 81 个基因(3.4%)在其中 3 株菌株中存在, 40 个基因(1.6%)只在两株菌株中存在, 56 个基因(2.2%)只在单一菌株中存在。共发现 145 个基因是 MSSA463 菌株所独有(网络版附图 3)。

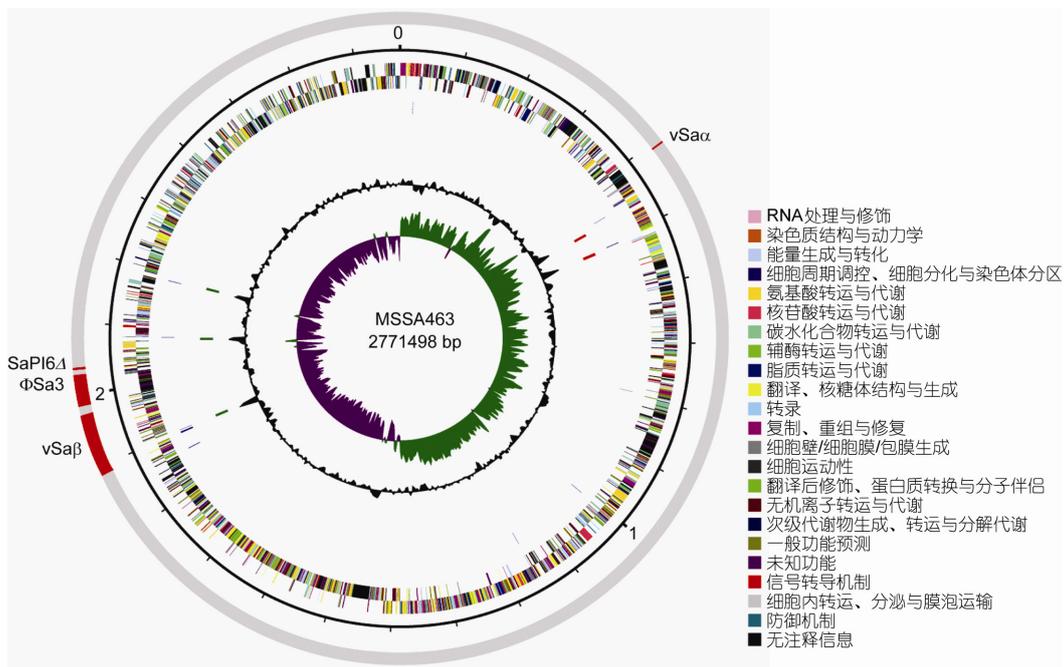


图1 SCC-like MSSA463 基因组图

自外向内, 环 1: MSSA463 基因组上已确认的 *S. aureus* 基因组岛(红色标记); 环 2: 基因组位置标尺; 环 3: 正链上的基因; 环 4: 负链上的基因, 不同颜色代表不同 COG 功能, 见右侧标注; 环 5: tRNA; 环 6: 正链上 rRNA 基因位置; 环 7: 负链上 rRNA 基因位置; 环 8: GC 含量变化曲线; 环 9: GC skew 变化峰图(>0%为绿色, <0%为紫色)

2.3 移动元件分析

通过与 MSSA476 和 MRSA252 菌株序列比较分析, 检测到 MSSA463 菌株基因组中的基因岛和可移动元件, 这些基因可能与毒力因子及抗药性因子的获得相关^[15]. 首先, 在 MSSA463 菌株基因组中并未发现典型的 SCC 结构(表 1), 该菌株含有与其他菌株中普遍存在的 Φ Sa3 噬菌体片段; 其次, MSSA463 菌株基因组中含有 Sa α -vSa β 基因岛, 包含 4 种葡萄球菌外毒素编码基因(*spl*, *lukDE*, *hysA* 和 *bsa*), 与其他 6 株金黄色葡萄球菌一致(MSSA476, MRSA252, N315, Mu50, MW2 和 NCTC8325); 再次, 发现致病岛 SaPI6 存在于 MSSA463 和 MSSA476 菌株, 但在 8325, COL, USA300 和 Mw2 菌株中并未发现, 而 2 个致病岛 SaPI4 和 SaPI1028 仅在 MRSA252 菌株中发现, 未在 MSSA463 和 MSSA476 中发现. 另外, 在 MSSA463 和 MSSA476 菌株中均未发现插入序列和转座子.

2.4 致病性相关编码序列与细胞膜相关抗性基因

通过 KEGG 完成 MSSA463 菌株疾病相关基因的预测. 在 MSSA463 菌株基因组中, 发现众多与感染

性疾病相关的基因, 包括与细胞侵入内皮细胞过程相关的编码基因 *fmbB* 和 *fmbA*. 另外, 还注释到与幽门螺旋杆菌感染侵袭上皮细胞过程相关的 *K08303*, *urea*, *ureB* 和 *ureC* 基因, 与致病性霍乱弧菌感染相关的 *luxS* 基因, 以及与阿米巴原虫感染过程相关的编码基因 *rocF*(K0416; EC3.5.3.1)和 *arg*(网络版附表 5).

通常细菌的致病因子和抗原蛋白锚定于细胞壁或细胞膜上^[16]. 在 MSSA463 菌株基因组中, 发现众多与细胞膜相关致病因子的 ORF, 如外源凝集素编码基因 *SdrD*(adhesin *SdrD*)、 β 孔道形成性溶血素(β -channel forming cytolysin)及凝集因子 ClfB (clumping factor B). 尽管 MSSA463 菌株对甲氧西林敏感, 但仍在其基因组中发现了一些其他抗性蛋白的编码基因, 如双环霉素抗性蛋白 *TcaB* 和氯霉素抗性蛋白等, 编码耐药基因与相应的抗生素耐药性相对应.

2.5 MSSA463 的系统发育关系与类 SCC 元件分析

采用 iTOL 软件构建系统发育树^[17](网络版附图 3). 在系统进化上, MSSA463 菌株更接近于 MRSA 252 菌株(gi 49240382, 图 2).

表 1 MSSA463、MSSA476 与 MRSA252 菌株中的移动元件

移动元件	长度(kbs)	毒力或耐药性基因	MSSA463	MSSA476	MRSA252
SCC					
II 型 SCCmec	~53.00	<i>mecA</i> , <i>mecR1</i> , <i>mecI</i>			+
SCC476	~23.00	<i>Far</i>		+	
噬菌体					
Φ Sa2	~45.80	<i>lukSF-PV</i> , Φ I2			+
Φ Sa3	~45.00	<i>sak</i> , <i>sep</i> , <i>sea</i> , <i>seg2</i> , <i>sek2</i>	+	+	+
Φ Sa4	~44.70	无		+	
基因岛					
vSa α	~3.80	<i>Set</i>	+	+	+
vSa β	~155.80	<i>spl</i> , <i>lukDE</i> , <i>hysA</i> , <i>bsa</i> , <i>extoxin</i>	+	+	+
致病岛					
SaPI4	~15.10	无			+
SaPI1028	~15.60	无			+
SaPI6 Δ	~3.14	无	+	+	
SaPI15305	~17.00	<i>aad</i> , <i>fosB</i>			
插入序列					
IS1272	~2.00	无			+
IS431	~0.80	无		+	+
ISX	~1.10	无			+
转座子					
Tn554	~6.70	<i>ermA</i> , <i>spc</i>			+
Tn552	~6.50	<i>blaZ</i>			+
Tn916-like	~7.60	无			+

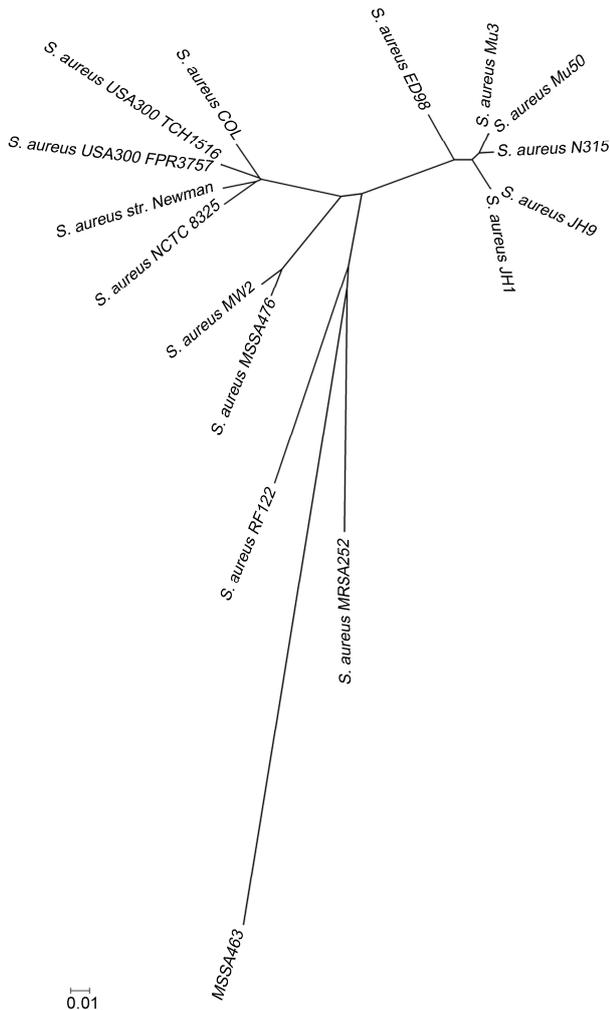


图2 部分金黄色葡萄球菌系统发育树

选取相应菌株的全基因组序列, 使用 SPLITSTREE Version 3.1 构建完成

选取了 MRSA85/2082, MSSA476 及 MRSA252 菌株的 SCC 结构元件作为参考序列^[6], 与 MSSA463 的类 SCC 结构进行比较分析(网络版附图 4). 所选取的 SCCmec 结构均代表性地整合于基因组的相同位置. 与其他 MRSA 菌株相似, MSSA463 菌株中也存在相同的反向重复序列, 但缺乏 SCCmec 结构中相对保守的位点特异性重组酶 *ccrA* 与 *ccrB*. 前期诊断分型检测到的 SCCmecIII 型特异性结构 ORF CZ049 (AB37671) 序列片段, 在 MSSA463 菌株中存在于反向重复序列 *attL* 和 *attR* 下游约 1680 bp 左右的位置, 而并非像常规 MRSA 菌株中处于反向重复序列 *attL* 和 *attR* 之间.

3 讨论

自第一株金黄色葡萄球菌 N315 测序完成以来, 已完成全基因组测序的菌株已达 22 株(<http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi>). 本研究所完成的金黄色葡萄球菌 MSSA463 并非典型 SCC 菌株, 由于前期分子流行病学方法筛选出来的特征, 由此命名为类 SCC 样 MSSA463 菌株.

MSSA463 菌株作为致病性菌株, 同多数致病菌一样, 可以侵入吞噬细胞和非吞噬细胞, 并在这些细胞内定植, 继而扩散至邻近组织^[18-21]. 在非吞噬性组织细胞中, 侵袭性致病菌通过锌指模型及触发模型机制诱导宿主非吞噬细胞摄取该类细菌, 如李斯特菌、链球菌、耶尔森菌等都是通过锌指模型侵入宿主细胞. MSSA463 菌株中 *fnbB* 和 *fnbA* 基因的存在, 提示金黄色葡萄球菌可通过相应的编码蛋白产物 *fnbB* 及 *fnbA* 与细胞表面纤维素结合蛋白相互作用, 通过锌指模型诱导宿主非吞噬细胞摄取该类细菌, 进而启动细菌侵入上皮细胞的级联反应过程. 同时, *ureA/ureB/ureC*, *luxSE3.5.3.1*, *rocF* 和 *arg* 蛋白编码基因的存在, 表明金黄色葡萄球菌的侵袭过程与幽门螺旋杆菌、霍乱弧菌及阿米巴原虫感染的级联反应信号系统存在一定的交叉^[16,22-24]. 这也暗示着病原菌之间乃至细菌与真核生物之间致病基因水平转移的可能^[25,26].

随着测序技术的发展以及基因组信息的不断积累, 对耐药机制的研究也有了不断的更新与完善. Katayama 等人^[27]发现, MRSA 菌株的耐甲氧西林抗药基因是由 SCCmec 结构携带, 该结构在细菌染色体上的整合和移除都是通过 *ccrA* 和 *ccrB* 基因编码的重组酶基因完成, 并首次定义了 SCCmec 元件的结构. 之后, 研究人员根据 SCCmec 元件中 *mec* 和 *ccr* 基因复合体组成和“junkyard”区中 DNA 片段的差异对 MRSA 菌株进行亚型分型^[28,29]. 在 MRSA 菌株中, SCCmec 结构的上下游边界分别是 *attL* 和 *attR* 反向重复序列. 通常这一对反向重复序列被认定与 SCCmec 是一个整体, 在基因横向转移过程中的共同出现, 并且导致耐药性改变. 然而在 MSSA476 基因组研究中发现的 SCC476 结构表明, 在 MSSA 菌株中也存在一些类似 SCC 元件的结构. 这些结构可能在 *S. aureus* 中普遍存在, 也可能是 MRSA 菌株结构变异丢失的结果, 仍需进一步研究来阐明其机制, 同时对金黄色

葡萄球菌抗药性的分子机制进行深入认识^[7].

先前针对下呼吸道感染 *S. aureus* 的分子流行病学检测主要应用多重 PCR 和 SCCmec 亚型鉴定^[7]. ORF CZ049 一直被认定为 SCCmecIII型 MRSA 菌株所特有的结构, 并作为 *S. aureus* 分型的依据^[28]. 在 MRSA85/2082 菌株中, ORF CZ049 也确实位于 *attL* 和 *attR* 反向重复序列之间. 本实验在基因比较研究中发现, 这种被认定为 SCCmecIII型菌株特有的结构不仅在其他亚型 MRSA 菌株中存在, 如 MRSA 株 JCSC3624(WIS)^[30], 同样在 MSSA463 菌株的结构中也发现了 ORF CZ049 结构.

这种结构上的差异产生的机制可能包括以下 3 个方面. 首先, MSSA463 是由一株 MRSA 型菌株经过基因转移后演变而来, 而在这株祖先 MRSA 菌株中, ORF CZ049 最初的位置可能就位于 SCC 结构附

近; 其次, SCCmec 元件也可能通过基因转移插入到 ORF CZ049 上游, 并且之后丢失了内部的基因, 只留下两侧反向重复序列; 再者, 也不排除 ORF CZ049 通过染色体重组事件进行移位的可能^[31,32]. ORF CZ049 位于反向重复序列上游, 究竟是短片段插入现象所致, 还是 SCCmec 结构变异后的继发性结构变异所致, 仍需进一步研究.

综上所述, 通过全基因组测序及比较基因组学分析, 进一步加深了对金黄色葡萄球菌致病性、耐药性及侵袭毒力相关分子机制和遗传基础的了解. 同时, SCCmec 结构中反向重复序列的作用是否也可以作为鉴定 *S. aureus* 的另一标准仍有待进一步确定. 因此, 随着更多金黄色葡萄球菌菌株全基因组序列测序工作的完成, 金黄色葡萄球菌耐药基因的结构、功能及其耐药性的产生和发生机制会逐渐清晰.

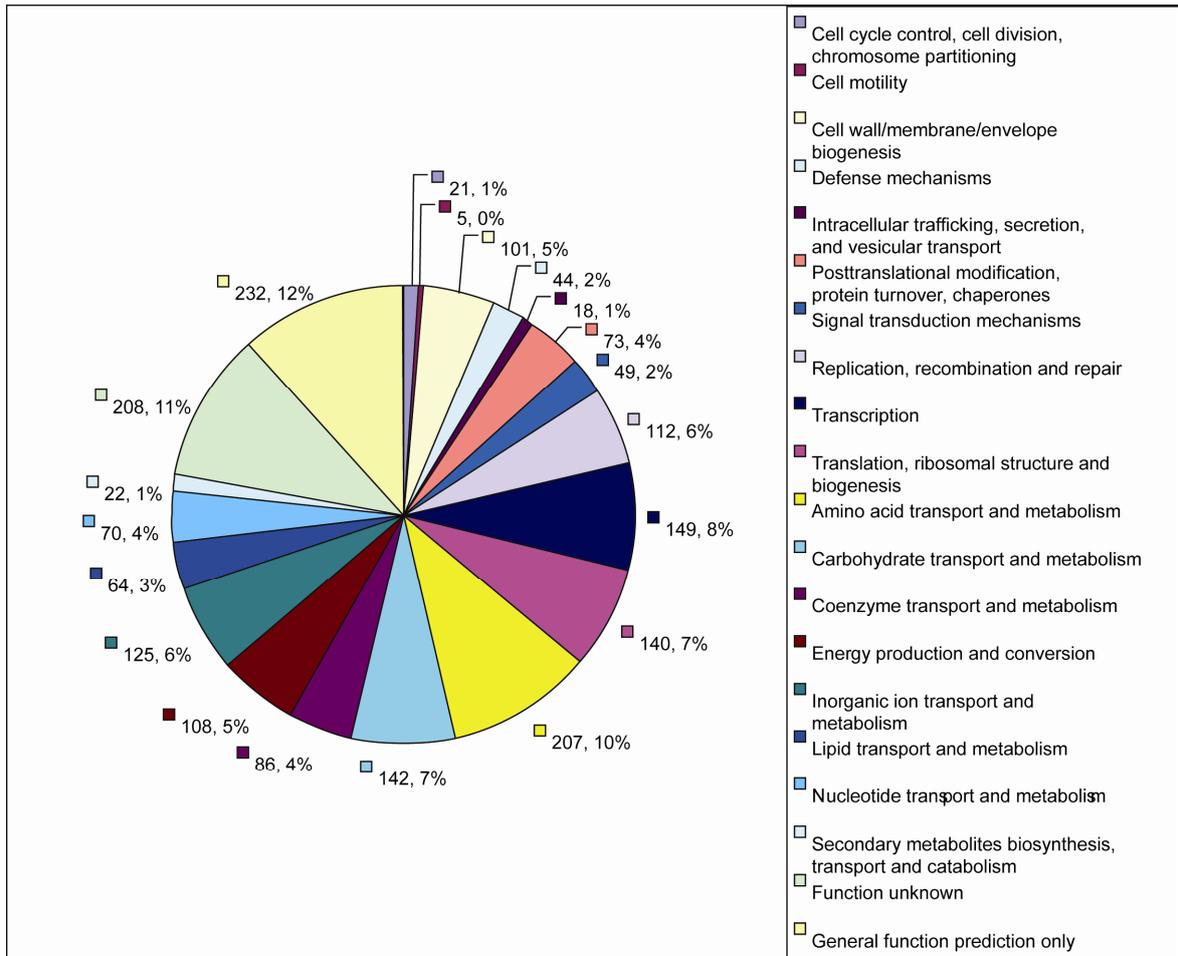
参考文献

- 1 Hiramatsu K, Aritaka N, Hanaki H, et al. Dissemination in Japanese hospitals of strains of *Staphylococcus aureus* heterogeneously resistant to vancomycin. *Lancet*, 1997, 350: 1670–1673
- 2 Vancomycin-resistant *Staphylococcus aureus*. New York, 2004, MMWR Morb Mortal Wkly Rep, 2004, 53: 322–323
- 3 Jones R N, Kohno S, Ono Y, et al. ZAAPS International Surveillance Program (2007) for linezolid resistance: results from 5591 Gram-positive clinical isolates in 23 countries. *Diagn Microb Infec Dis*, 2009, 64: 191–201
- 4 Pantosti A, Sanchini A, Monaco M. Mechanisms of antibiotic resistance in *Staphylococcus aureus*. *Future Microbiol*, 2007, 2: 323–334
- 5 Corkill J E, Anson J J, Griffiths P, et al. Detection of elements of the staphylococcal cassette chromosome (SCC) in a methicillin-susceptible (*mecA* gene negative) homologue of a fucidin-resistant MRSA. *J Antimicrob Chemother*, 2004, 54: 229–231
- 6 Holden M T, Feil E J, Lindsay J A, et al. Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. *Proc Natl Acad Sci USA*, 2004, 101: 9786–9791
- 7 Li D Z, Chen Y S, Yang J P, et al. Preliminary molecular epidemiology of the *Staphylococcus aureus* in lower respiratory tract infections: a multicenter study in China. *Chin Med J(Engl)*, 2011, 124: 687–692
- 8 Margulies M, Egholm M, Altman W E, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 2005, 437: 376–380
- 9 Delcher A L, Bratke K A, Powers E C, et al. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics*, 2007, 23: 673–679
- 10 Bairoch A, Apweiler R. The SWISS-PROT protein sequence data bank and its new supplement TrEMBL. *Nucleic Acids Res*, 1996, 24: 21–25
- 11 Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*, 2000, 28: 27–30
- 12 Mulder N J, Apweiler R, Attwood T K, et al. Inter Pro: an integrated documentation resource for protein families, domains and functional sites. *Brief Bioinform*, 2002, 3: 225–235
- 13 Harris M A, Clark J, Ireland A, et al. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res*, 2004, 32(Database issue): D258–D261
- 14 Carver T J, Rutherford K M, Berriman M, et al. ACT: the artemis comparison tool. *Bioinformatics*, 2005, 21: 3422–3423
- 15 Lindsay J A, Holden M T. *Staphylococcus aureus*: superbug, super genome? *Trends Microbiol*, 2004, 12: 378–385
- 16 Hsiao A, Liu Z, Joelsson A, et al. *Vibrio cholerae* virulence regulator-coordinated evasion of host immunity. *Proc Natl Acad Sci USA*, 2006, 103: 14542–14547
- 17 Letunic I, Bork P. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res*, 2011, 39: W475–W478
- 18 Mostowy S, Cossart P. Cytoskeleton rearrangements during *Listeria* infection: clathrin and septins as new players in the game. *Cell Motil*

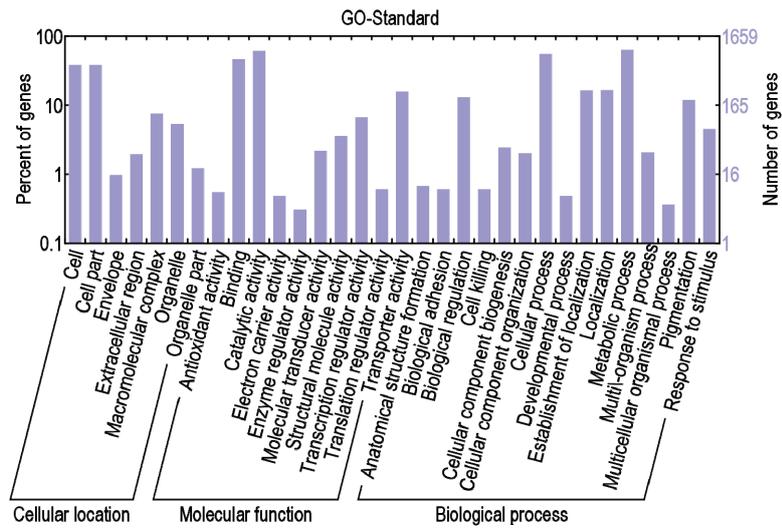
- Cytoskel, 2009, 66: 816–823
- 19 Wang B N, Li S Y, Dedhar S, et al. Paxillin phosphorylation: bifurcation point downstream of integrin-linked kinase (ILK) in streptococcal invasion. *Cell Microbiol*, 2007, 9: 1519–1528
 - 20 Matsumoto H, Young G M. Translocated effectors of *Yersinia*. *Curr Opin Microbiol*, 2009, 12: 94–100
 - 21 Iwai H, Kim M, Yoshikawa Y, et al. A bacterial effector targets Mad2L2, an APC inhibitor, to modulate host cell cycling. *Cell*, 2007, 130: 611–623
 - 22 Petri W A, Haque R, Mann B J. The bittersweet interface of parasite and host: lectin-carbohydrate interactions during human invasion by the parasite *Entamoeba histolytica*. *Annu Rev Microbiol*, 2002, 56: 39–64
 - 23 Gerhard M, Rad R, Prinz C, et al. Pathogenesis of *Helicobacter pylori* infection. *Helicobacter*, 2002, 7: 17–23
 - 24 Montecucco C, Rappuoli R. Living dangerously: how *Helicobacter pylori* survives in the human stomach. *Nat Rev Mol Cell Bio*, 2001, 2: 457–466
 - 25 Klotz M G, Loewen P C. The molecular evolution of catalatic hydroperoxidases: evidence for multiple lateral transfer of genes between prokaryota and from bacteria into Eukaryota. *Mol Biol Evol*, 2003, 20: 1098–1112
 - 26 Keeling P J, Palmer J D. Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet*, 2008, 9: 605–618
 - 27 Katayama Y, Ito T, Hiramatsu K. A new class of genetic element, staphylococcus cassette chromosome mec, encodes methicillin resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother*, 2000, 44: 1549–1555
 - 28 Zhang K Y, McClure J A, Elsayed S, et al. Novel multiplex PCR assay for characterization and concomitant subtyping of staphylococcal cassette chromosome mec types I to V in methicillin-resistant *Staphylococcus aureus*. *J Clin Microbiol*, 2005, 43: 5026–5033
 - 29 Higuchi W, Takano T, Teng L J, et al. Structure and specific detection of staphylococcal cassette chromosome mec type VII. *Biochem Bioph Res Co*, 2008, 377: 752–756
 - 30 Ito T, Ma X X, Takeuchi F, et al. Novel type v staphylococcal cassette chromosome mec driven by a novel cassette chromosome recombinase, ccrC. *Antimicrob Agents Chemother*, 2004, 48: 2637–2651
 - 31 Robinson D A, Enright M C. Evolution of *Staphylococcus aureus* by large chromosomal replacements. *J Bacteriol*, 2004, 186: 1060–1064
 - 32 Price L B, Stegger M, Hasman H, et al. *Staphylococcus aureus* CC398: host adaptation and emergence of methicillin resistance in livestock. *Mbio*, 2012, 3: e00305-11

论文

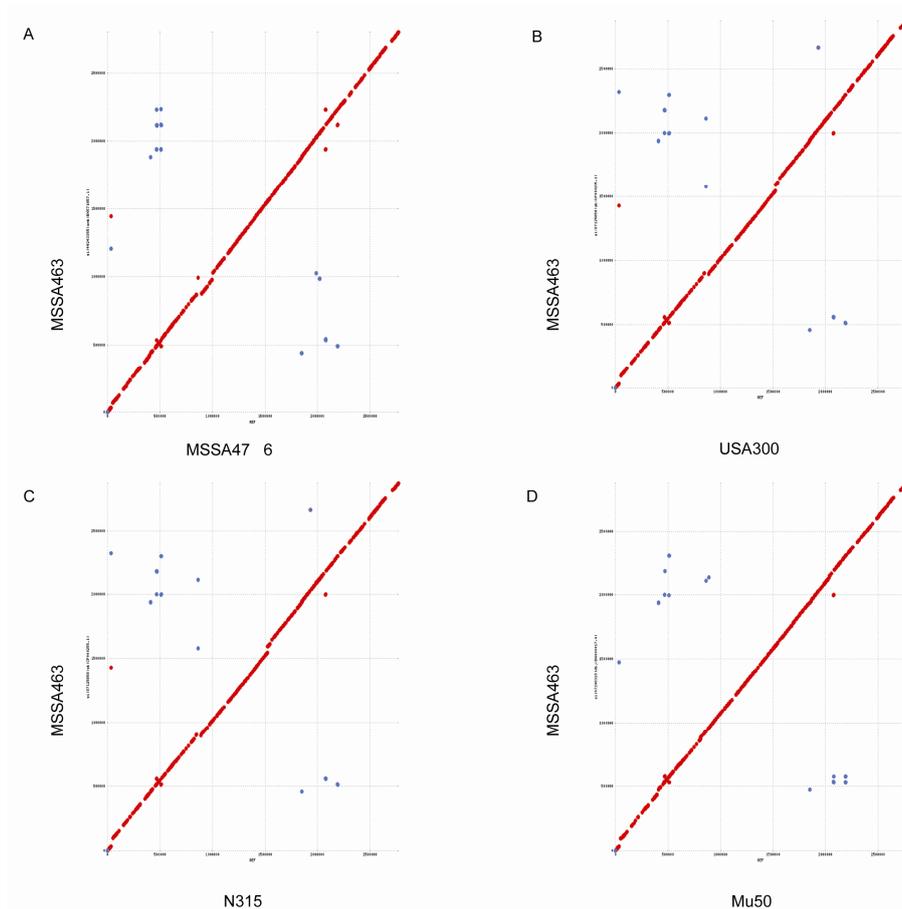
附录



附图1 Clusters of Orthologous Groups (COG) 注释

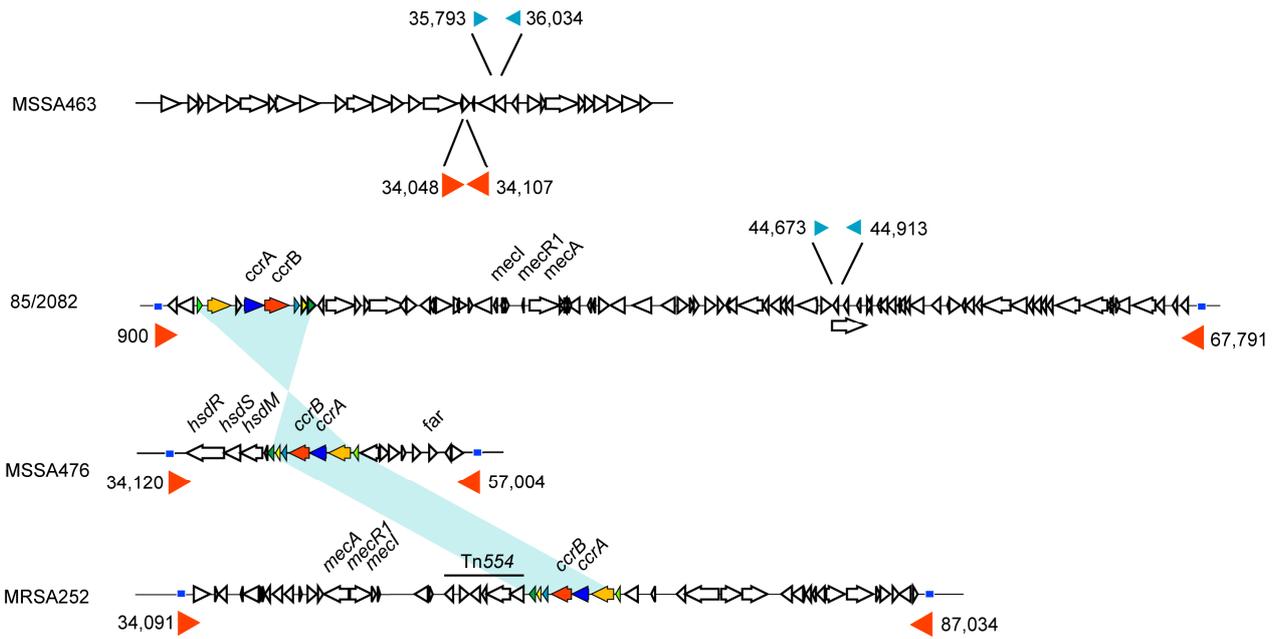


附图 2 Gene Ontology 分类注释



附图 3 全基因组同源性比对

采用 MUMmer-based 同源比对软件, 选取的 MSSA463 的全基因组序列与参考序列 MSSA476 (A), USA300 (B), N315 (C), and Mu50 (D) 进行比对分析. 红线表示在基因组之间位置相同的同源基因, 蓝线表示在基因组之间位置相反的同源基因



附图 4 SCC 结构比较

箭头指示基因的转录方向. SCC 组分保守区段以蓝色阴性做出标识. SCC 成份末端的深蓝色标记为左右两侧的结合部位(attL、attR)末端重复序列以相反方向的橘黄色箭头表示

附表 1 MSSA463 菌株耐药性^{a)}

Antibiotics	Drug sensitivity
Amoxicillin/CA	S
Ampicillin	R
Ampicillin/sulbactam	S
Cefaclor	S
Cefotaxime	S
Ceftriaxone	S
Ceftriaxone-Sodium	S
Ciprofloxacin	S
Clindamycin	S
Erythromycin	S
Fosfomicin	S
Fusidic Acid	S
Gentamicin	S
Imipenem	S
Levofloxacin	S
Linezolid	S
Moxifloxacin	S
Nitrofurantoin	S
Norfloxacin	S
Oxacillin	S
GPenicillin-G	R
Rifampin	S
Teicoplanin	S
Tetracycline	R
Tobramycin	S
Trimethoprim/Sulfa	S
Vancomycin	S

a) S: 敏感; R: 抗药

附表 2 分析使用数据库以及相应参数

Database	MSSA463		E cutoff	Database version
	of Annotated	%		
Total genes	2563	–	–	–
Nr	2526	98.6	1e-5	20091003
Swissprot	1894	73.9	1e-10	20091208
Kegg	2500	97.5	1e-10	Release50
COG	2370	92.5	1e-10	No Version
Interpro	2133	83.2	Interproscan 4.3	Release16.0
GO	1660	64.8		

附表 3 测序数据统计信息

Features	MSSA463
Average read length (bp)	352
Coverage of sequence	70×
Number of contigs (assembly)	33/25 (1000 24/24)
Contig N50 (bp)	233, 435
Largest contig (bp)	535, 747
GC%	32.3
Number of assembled bases	2, 773, 547/2, 771, 829
Predicted genes	2595

附表 4 MSSA463 基因组整体特征

Feature	Strain: MSSA463
Length of sequence (bp)	2771498
G+C content	32.45%
Repeat content length (%)	0.93%
Simple repeats length (coverage)	657 (0.02%)
Low complexity length (coverage)	15310 (0.55%)
Repeat element length (coverage)	553 (0.02%)
Small RNA # (length, coverage)	29 (9218, 0.33%)
ORF (Length in Genome)	2313501
Feature	
ORF #	2563
Genome coverage (%)	83.47%
Genome density (genes/Kb)	0.92
Average gene length (bp)	903
Max gene length (bp)	28236
Function	
ORF with assigned functions	2370
Conserved hypothetical protein	2526
ORF without database match	37
RNA	
rRNA	
5s rRNA # (length)	6 (113)
16s rRNA # (length)	5 (1539)
23s rRNA # (length)	5 (2920)
tRNA # (total length)	59 (4503)

附表5 KEGG 代谢通路中与致病性相关基因注释情况

Human Diseases	Numbers	CDS predicted
Cancers; Pathways in cancer	1	MSSA463 1727:E4.2.1.2B, fumC
Cancers; Renal cell carcinoma	1	MSSA463 1727:E4.2.1.2B, fumC
Immune System Diseases; Primary immunodeficiency	1	MSSA463 0515:UNG, UDG
Infectious Diseases; Amebiasis	1	MSSA463 2014:E3.5.3.1, rocF, arg
Infectious Diseases; Bacterial invasion of epithelial cells	2	MSSA463 2341:fnbB MSSA463 2342:fnbA
Infectious Diseases; Epithelial cell signaling in Helicobacter pylori infection	4	MSSA463 1505:K08303 MSSA463 2132:ureA MSSA463 2133:ureB MSSA463 2134:ureC
Infectious Diseases; Vibrio cholerae pathogenic cycle	1	MSSA463 1988:luxS
Metabolic Diseases; Type I diabetes mellitus	1	MSSA463 1883:groEL, HSPD1
Metabolic Diseases; Type II diabetes mellitus	1	MSSA463 1583:PK, pyk
Neurodegenerative Diseases; Alzheimer's disease	2	MSSA463 0715:GAPDH, gapA MSSA463 1573:GAPDH, gapA
Neurodegenerative Diseases; Amyotrophic lateral sclerosis(ALS)	1	MSSA463 1238:katE, CAT
Neurodegenerative Diseases; Huntington's disease	1	MSSA463 1448:E1.15.1.1A, sodA, sodB, SOD2