



创伤弧菌毒力相关因子的全基因组关联分析研究

张嘉鑫^{1,2}, 武雅蓉², 杨超², 康琳², 王景林^{1,2*}, 崔玉军^{1,2*}

¹安徽医科大学, 安徽 合肥 230032

²军事科学院军事医学研究院微生物流行病学研究所, 病原微生物生物安全国家重点实验室, 北京 100071

摘要:【目的】创伤弧菌是致死率最高的弧菌物种,但目前尚无在全基因组层面挖掘毒力相关因子的研究。本研究以创伤弧菌分离来源(临床和环境)作为不同表型,通过与260株基因组序列进行关联分析,挖掘毒力相关因子,从而进一步了解创伤弧菌致病因素。【方法】对139株创伤弧菌分离株进行高通量测序,获取其全基因组序列;与公共数据库已公开发表的121株基因组整合,使用pyseer软件进行全基因组关联分析,对与不同分离来源显著相关的基因进行注释和解读。【结果】共发现11个基因与临床分离株显著相关,其中9个是本研究新发现的创伤弧菌潜在毒力相关因子。【结论】本研究使用群体基因组学和统计遗传学方法,在全基因组范围扫描挖掘了创伤弧菌毒力相关因子,为深入揭示该物种致病机制、设计新的疫苗和治疗靶点提供了重要依据。

关键词: 创伤弧菌, 全基因组测序, 全基因组关联研究, 毒力因子

创伤弧菌(*Vibrio vulnificus*),与霍乱弧菌、副溶血弧菌并称为人类三大致病性弧菌^[1]。创伤弧菌可引发原发性败血症,是目前发现的致死率最高的致病弧菌^[2],美国95%以上因食用海鲜而致死的病例是该物种引起的^[3]。该病原对男性患者(约85%)和老年患者(>40岁)的影响尤为严重,特别是患有肝病、糖尿病和免疫紊乱等潜在疾病的人群^[2]。创伤弧菌的分布与温度有关,随着全球变暖和海水温度升高,其感染的病例数和地理区域不断扩大,明确创伤弧菌的毒力因子、探索其致病

机制具有重要意义^[4]。

全基因组关联分析(genome-wide association study, GWAS)是近年来迅速兴起的一种遗传机制分析方法。该方法将基于全基因组范围鉴定获得的遗传多态性位点与可观测性状(即表型)进行群体水平统计分析,从而筛选出最有可能影响该性状的遗传变异位点^[5]。与传统关联分析比较, GWAS具有如下优势:(1)不需要建立前提假设和筛选候选基因,可直接探查与某表型关联的遗传变异机制^[6];(2)检测范围更全面,可在整个基因

基金项目:国家重点研发计划(2018YFC1603902, 2017YFC1601500)

*通信作者。E-mail: 崔玉军, cuiyujun.new@gmail.com; 王景林, wjlwj10801@sina.com

收稿日期: 2019-12-31; 修回日期: 2020-02-05; 网络出版日期: 2020-03-12

组水平检测跟表型相关的所有变异位点^[7]; (3) 可同时对大量样本进行全基因组水平的关联分析, 实现精细定位, 节省时间和成本^[8]。GWAS 在人类复杂疾病遗传机制和动植物育种方面已取得丰硕成果, 打开了基因组学研究的新大门。但限于克隆繁殖特性所致的种群结构, 细菌全基因组关联研究(bacterial genome wide association study, BGWAS)开展较晚, 目前主要应用于细菌宿主适应性、毒力和抗生素耐药等表型与基因型的相关性研究中^[9]。

本研究首次将 BGWAS 引入创伤弧菌的研究领域, 对分离自临床和环境样本中共计 260 株创伤弧菌全基因组序列与其分离来源进行分析, 在全基因组的范围内筛选可能与创伤弧菌毒力相关的变异位点或功能基因, 为进一步探索创伤弧菌的致病机理, 以及该病原的毒力检测和防治工作提供科学依据。

1 材料和方法

1.1 菌株来源

本研究共收集 260 株创伤弧菌的全基因组序列, 包括 139 个新测序的全基因组序列和 121 个公共基因组序列。新测序的菌株来源于: 台湾成功大学(26 个)、温州医科大学附属第一医院(5 个)、中国科学院动物研究所(1 个)、北京出入境检验检疫局(32 个)、广东检验检疫技术中心卫检室(7 个)、中国海洋微生物菌种保藏管理中心(2 个)、浙江疾病预防控制中心(33 株)和上海交通大学(33 株), 所有菌株均保存于 -80°C ; 公共序列下载于 NCBI 数据库(<https://www.ncbi.nlm.nih.gov/genome/genomes/189>)。菌株背景信息见 <https://doi.org/10.6084/m9.figshare.11758704.v1>。

1.2 接种与培养

将收集的菌种在哥伦比亚血琼脂(Columbia blood agar, CBA)平板划线后 37°C 生化培养箱中培养, 挑取单菌落接种于 3 mL 心浸液肉汤(Heart infusion, HI)液体培养基中, 过夜培养至指数生长期。

1.3 全基因组测序

取 1.8 mL 过夜培养菌液, 使用细菌基因组 DNA 提取试剂盒(QIAGEN UltraClean[®] Microbial DNA Isolation kit), 参照试剂盒说明书提取全基因组 DNA。经检测质量合格后由测序公司(诺禾致源, 北京)进行细菌全基因组重测序, 测序平台为 Novaseq 6000, 测序文库大小为 500 bp, 测序 reads 读长为双端 150 bp, 平均测序深度 257X (80–290X)。

1.4 质控及组装

使用 FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/)软件中 FASTQ Quality Filter 模块过滤低质量数据, 并采用 shovill (<https://github.com/tseemann/shovill>)软件对过滤后的 reads 进行组装。对组装结果进行统计, 得到本次测序菌株的平均基因组大小为 5.03 Mb, 平均 GC 含量为 46.71% (<https://doi.org/10.6084/m9.figshare.11759601.v1>)。组装得到的基因组草图序列已上传至 NCBI, Bioproject 编号为 PRJNA588308, 详细信息见 <https://doi.org/10.6084/m9.figshare.11758704.v1>。

1.5 变异检测和注释

单核苷酸多态性(single nucleotide polymorphism, SNP)位点的鉴定方法参照已发表的相关工作^[10]: 以创伤弧菌 CECT4999 株基因组完成图为参考序列, 使用 MUMmer 软件包^[11]将 260 个创伤弧菌基因组序列分别比对到参考株基因组, 获得核心基因组(全部菌株中均能找到对应碱基状态)中的 SNP 状态信息。为保证 SNP 鉴定

结果的可靠性, 位于重复区和测序质量值低的 SNP 未纳入后续分析。

采用 Prokka 软件^[12]对各个菌株基因组序列进行基因预测, 将基因序列分别与 COG 库(clusters of orthologous groups) 和 KEGG 库 (kyoto encyclopedia of genes and genomes)进行蛋白质同源序列比对, 完成功能注释。使用 Roary 软件^[13]构建泛基因组序列和基因获得缺失矩阵, 进而判定基因获得缺失。

1.6 系统发育分析

使用 TreeBest 软件(<http://treesoft.sourceforge.net/treebest.shtml>)对 SNP 数据集构建邻接 (Neighbour-Joining, NJ)树, bootstrap 值设为 1000, 并使用在线工具 iTOL^[14]进行可视化展示。

1.7 已知毒力因子分析

统计 VFDB 数据库(virulence factors database)^[15]和已发表文献中毒力因子的存在缺失情况。基于开源 R 软件的分析工具包进行统计学检验, 采用列联表 χ^2 检验来判断创伤弧菌已知毒力因子在环境组和临床组中的频率分布差异, $P < 0.05$ 认为有显著性差异。

1.8 全基因组关联分析

采用 pyseer 软件^[16], 将创伤弧菌的分离来源信息(临床和环境)分别与全基因组范围 SNPs、基因存在缺失以及 Kmers (将全基因组序列切分成几十到上百 bp 的短片段)进行关联分析。首先, 使用 Mash 软件^[17]计算样本两两之间的遗传距离, 用于校正种群结构。然后, 分别从 SNPs、基因以及 Kmers 三个水平, 采用 pyseer 的固定效应模型, 过滤掉最小等位基因频率(MAF, Minor Allele Frequency) < 5% 的变异, 利用多维尺度 (multidimensional scaling, MDS)来评估每一个 SNP、

基因或 Kmer 对表型的影响, 并计算相应 P 值。

2 结果和讨论

2.1 创伤弧菌种群结构及地理分布

本研究共分析了 260 株创伤弧菌的全基因组序列, 这些菌株分离自 1977 年–2018 年间, 覆盖 4 个大洲共 10 个国家。变异检测发现 229462 个 SNP 变异, 其中 103117 个位于基因间区, 124999 个是同义 SNPs (碱基突变不引起氨基酸编码改变), 40165 个是非同义 SNPs (碱基突变引起氨基酸编码改变), 1255 个是无义 SNPs (碱基突变引起氨基酸编码改变成终止密码子)。基于所有 SNPs 构建的系统发育树(图 1)表明: 创伤弧菌可划分为 5 个系统发育谱系(L1、L2、L3、L4、L5), 与过去报道一致^[18]; 系统发育树内部分支较长且 bootstraps 值低, 表明该物种具有较高的同源重组率。菌株集中分布于 L1、L2 谱系(占 93.1%), 地理分布呈现一定的地域聚集性: L1 谱系主要包含中国分离株(占 86.8%)且临床株占比高(62%); L2 谱系主要包含美国分离株(占 42.9%)且环境株占比高(62%)。相比 L1 谱系, L2 谱系菌株的地理分布更为广泛, 分离国家包括西班牙、法国、丹麦、日本、中国、德国、韩国、以色列和巴西等。目前 L3、L4 和 L5 谱系分离株总数很少, 但临床株占比高(72.2%), 因此值得在未来监测中加以关注。

2.2 已知毒力因子与创伤弧菌分离来源的关联分析

目前应用最为广泛的毒力因子数据库是刘波等在 2005 年建立的 VFDB 数据库, 我们在该数据库的核心库(set A, 仅包含经实验验证过的毒力基因)中, 共找到 226 个存在等位基因的毒力因子。另外, 我们通过文献调研找到 56 个其他的

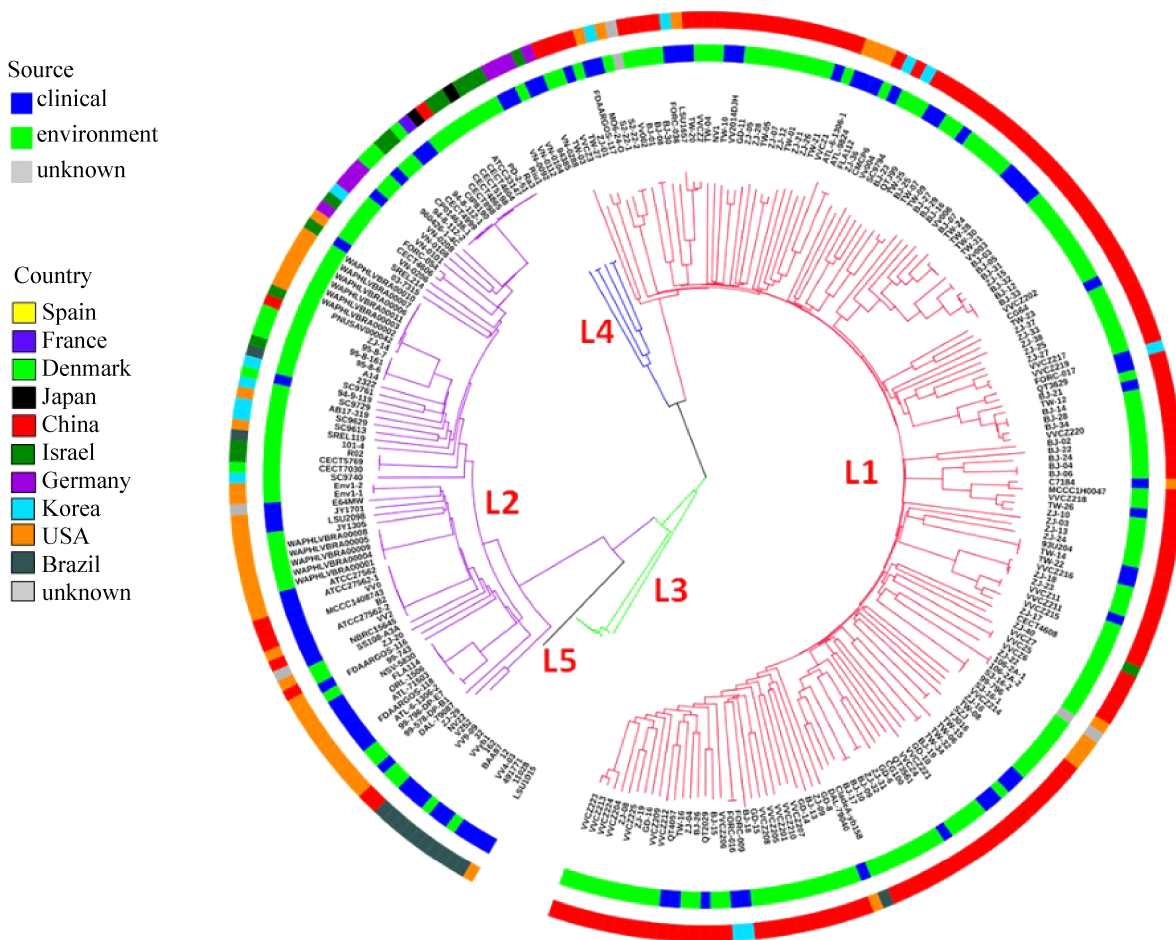


图 1. 种群结构图

Figure 1. Population structure of *V. vulnificus*. A NJ tree of 260 *V. vulnificus* strains constructed based on 229462 SNPs. The ring colors from inner to outer indicate source of sampling and isolated location, respectively. Branch colors indicate different populations, pink for L1, purple for L2, green for L3, blue for L4, black for L5.

毒力因子。使用 χ^2 检验来计算 282 个已知毒力因子在创伤弧菌环境分离株和临床分离株之间的频率分布差异(<https://doi.org/10.6084/m9.figshare.11760372.v1>)。结果表明, 仅有 2 个已知毒力因子 *pldA*、*vgrG* 在环境组和临床组中的分布有显著性差异($P < 0.05$), 因此, 目前已知的大多数毒力因子无法解释创伤弧菌环境分离株与临床分离株之间的差异。

2.3 使用 BGWAS 方法鉴定毒力相关因子

为全面挖掘创伤弧菌毒力相关元件, 我们应用 BGWAS 方法, 分别从 SNPs、基因获得缺失和

Kmers 三个不同水平, 在全基因组范围对 260 株创伤弧菌的分离来源表型与基因型进行关联分析, 检测临床菌株中的特异性遗传元件。为降低假阳性率, 我们结合先验知识, 采用严格阈值对分析结果进行判定, 即仅将统计检验 P 值低于已知毒力因子 *vgrG* 的遗传变异纳入最终结果 [$-\log(P\text{-value}) > 10.4$, 图 2]。在该阈值下, BGWAS 共鉴定出 11 个遗传变异位点, 分别位于 11 个编码基因上, 其中 9 个基因是本研究中首次发现与创伤弧菌毒力相关的遗传因子。全部 11 个基因中, 有 10 个基因在 SNPs 和 Kmers 水平被同时检测到,

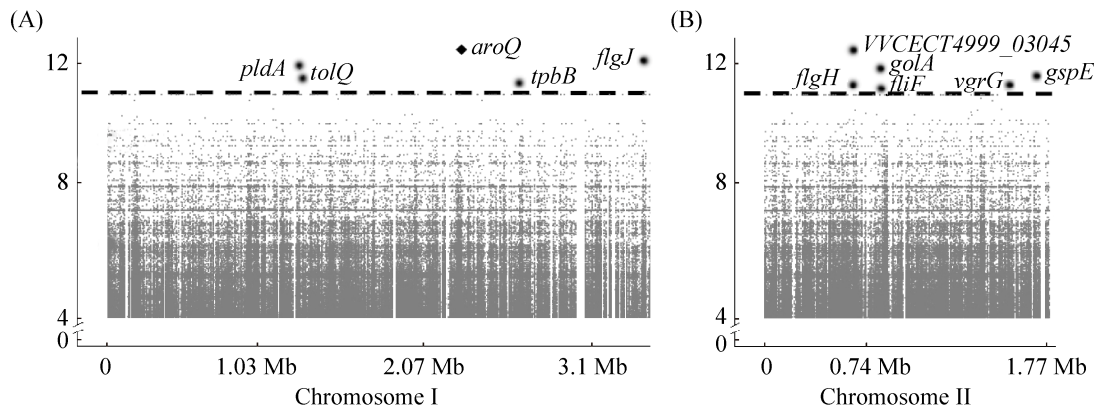


图 2. 菌株来源(环境/临床)与菌株 SNPs/Gene/Kmers 位点 BGWAS 分析的 Manhattan 图

Figure 2. Manhattan plots of BGWAS results for *V. vulnificus*. The X-axis of panel A and B separately indicate the position of chromosome I and the chromosome II of the reference genome (*V. vulnificus* CECT4999), and Y-axis is the $-\log_{10}(P\text{-value})$ score of BGWAS. Dash lines indicate the threshold used to identify the significant results. Grey dots, the $-\log_{10}(P\text{-value})$ of the variation sites at the corresponding position of the chromosomes; Black diamond, gene that carried significant Kmers; Black dots, genes that carried both significant Kmers and SNPs detected in current study.

提示统计检验结果是高度可靠的, 值得注意的是, 在基因获得缺失层面, 未发现与分离来源相关的变异位点。这提示创伤弧菌的毒力主要与核心基因组相关, 其附属基因组(包括基因组岛、质粒等)对其影响较弱。

2.4 显著相关基因的功能

我们对鉴定出的 11 个统计显著基因进行了注释和解读(表 1)。*pldA* 和 *vgrG* 基因已被报道与创伤弧菌致病性相关: *pldA* 基因编码磷脂酶 A1, 通过研究创伤弧菌细胞溶血素阴性突变株 CVD707, 发现磷脂酶 A1 是创伤弧菌的主要溶血和致死因子^[19]; *vgrG* 基因是六型分泌系统(Type VI secretion system, T6SS)蛋白编码基因, 在细菌感染宿主过程中发挥重要作用^[20-21], 该基因通过肌动蛋白交联域(actin cross-linking domains)发挥作用, 以影响创伤弧菌毒力, 在大肠杆菌和霍乱弧菌中也有相同的功能^[22]。本研究新发现的 9 个基因可能以直接或间接方式影响创伤弧菌的致病能力。其中 2 个基因编码代谢蛋白: *aroQ* 基因编码 3-脱氢奎

表 1. 统计显著相关基因

Table 1. Genes significantly associated with clinical strains

Genes	Function	Supported variations (number)
<i>pldA</i>	Phospholipase A1	K-mers (3), SNP (2)
<i>vgrG</i>	Type VI secretion system protein VgrG	K-mers (53), SNP (3)
<i>aroQ</i>	3-dehydroquinate dehydratase	K-mers (2)
<i>glpA</i>	Glycerol-3-phosphate dehydrogenase subunit A	K-mers (2), SNP (3)
<i>gspE</i>	Type II secretion system protein GspE	K-mers (2), SNP (3)
<i>flgJ</i>	Flagellar assembly peptidoglycan hydrolase FlgJ	K-mers (76), SNP (3)
<i>flgH</i>	Flagellar L-ring protein	K-mers (11), SNP (1)
<i>fliF</i>	Flagellar basal body M-ring protein FliF	K-mers (10), SNP (1)
<i>tpbB</i>	Diguanylate cyclase	K-mers (48), SNP (2)
<i>tolQ</i>	Protein TolQ	K-mers (8), SNP (2)
<i>VVCECT4999_03045</i>	Insulinase M16	K-mers (100), SNP (7)

尼酸脱水酶, 参与苯丙氨酸的合成, 苯丙氨酸的寡聚作用可影响创伤弧菌毒力因子溶血素的细胞裂解活性和凋亡活性^[23]; *glpA* 基因编码 3-磷酸甘油醛脱氢酶亚基 A, 3-磷酸甘油醛脱氢酶(GAPDH) 可修复 DNA、调节组蛋白、促进细胞凋亡^[24]、参与细菌黏附宿主组织、激发宿主保护性免疫反应^[25], 进而影响创伤弧菌毒力。5 个基因编码结构蛋白: *gspE* 基因是 II 型分泌系统(Type II Secretion System, T2SS)蛋白编码基因, T2SS 是广泛分布于革兰氏阴性细菌中的蛋白质分泌系统, 创伤弧菌 T2SS 可将多种蛋白转运到菌体细胞外, 直接与宿主细胞发生相互作用而发挥其毒性^[26]; *tolQ* 基因编码 TolQ 蛋白, 与 TolR、TolA 和 TolB 可构成 Tol 系统, 通过影响革兰氏阴性菌的生化 and 代谢机制, 增强细菌毒性, 在爱德华氏菌、小肠沙门氏菌、大肠杆菌、杜氏嗜血杆菌和霍乱弧菌中也有相同功能^[27-28]; *flgJ*、*flgH*、*fliF* 基因编码鞭毛蛋白, 是创伤弧菌鞭毛的组成成分, 与细菌的侵袭、黏附能力和致病性有关^[29]。2 个基因编码信息传递蛋白: *VVCECT4999_03045* 基因编码胰岛素酶 M16, 胰岛素酶可通过调控宿主内分泌信号为细菌提供有利生存条件, 促进创伤弧菌的体内存活能力^[30]; *tpbB* 基因编码二鸟苷酸环化酶, 二鸟苷酸环化酶含有 GGDEF 结构域的蛋白质, 可参与调控因子 c-di-GMP 的合成和降解, 干扰创伤弧菌生物膜形成, 影响粘附因子、运动性和毒力基因的表达^[31], 进而影响创伤弧菌毒力。

3 结论

本研究结合高通量测序技术和群体遗传学方法, 对 260 株创伤弧菌分离株进行了系统发育重建和毒力相关因子挖掘研究。研究结果表明创伤

弧菌由 5 个系统发育谱系组成, 谱系与菌株地域分布存在关联: 其中中国菌株主要聚集在 L1 分支, 美国菌株主要聚集在 L2 分支。BGWAS 分析发现 11 个遗传因子与临床或环境分离来源具有强相关性, 表明这些基因可能与创伤弧菌临床毒力或致病性相关。其中 9 个基因的毒力相关性在本研究首次得到证明, 为下一步深入探索创伤弧菌致病机制、设计新的疫苗或者药物提供了研究靶标。

参考文献

- [1] Wang MY, Hu CJ. Pathogenicity and virulence factors of *Vibrio vulnificus*: research advances. *Chinese Journal of Microecology*, 2017, 29(12): 1470-1473. (in Chinese)
王明义, 胡成进. 创伤弧菌致病性及其毒力因子研究进展. *中国微生态学杂志*, 2017, 29(12): 1470-1473.
- [2] Baker-Austin C, Oliver JD. *Vibrio vulnificus*: new insights into a deadly opportunistic pathogen. *Environmental Microbiology*, 2018, 20(2): 423-430.
- [3] Hoffmann S, Macculloch B, Batz M. Economic burden of major foodborne illnesses acquired in the United States. US: US Department of Agriculture, Economic Research Service, 2015.
- [4] Yun NR, Kim DM. *Vibrio vulnificus* infection: a persistent threat to public health. *The Korean Journal of Internal Medicine*, 2018, 33(6): 1070-1078.
- [5] He SB, Qu YP, Yu XQ, Cao J, Ye CR. Research status and prospect of genome-wide association study. *Shanghai Journal of Animal Husbandry and Veterinary Medicine*, 2017, (3): 12-15. (in Chinese)
何随彬, 瞿瑜萍, 俞向前, 曹杰, 叶承荣. 全基因组关联分析研究现状及展望. *上海畜牧兽医通讯*, 2017, (3): 12-15.
- [6] Power RA, Parkhill J, de Oliveira T. Microbial genome-wide association studies: lessons from human GWAS. *Nature Reviews Genetics*, 2017, 18(1): 41-50.
- [7] Wang SB, Feng JY, Ren WL, Huang B, Zhou L, Wen YJ, Zhang J, Dunwell JM, Xu SZ, Zhang YM. Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. *Scientific Reports*, 2016, 6(1): 19444.
- [8] Xiao YJ, Liu HJ, Wu LJ, Warburton M, Yan JB. Genome-wide

- association studies in maize: praise and stargaze. *Molecular Plant*, 2017, 10(3): 359–374.
- [9] Yang C, Yang RF, Cui YJ. Bacterial genome-wide association study: methodologies and applications. *Hereditas (Beijing)*, 2018, 40(1): 57–65. (in Chinese)
杨超, 杨瑞馥, 崔玉军. 细菌全基因组关联研究的方法与应用. *遗传*, 2018, 40(1): 57–65.
- [10] Cui YJ, Yu C, Yan YF, Li DF, Li YJ, Jombart T, Weinert LA, Wang ZY, Guo ZB, Xu LZ, Zhang YJ, Zheng HC, Qin N, Xiao X, Wu MS, Wang XY, Zhou DS, Qi ZZ, Du ZM, Wu HL, Yang XW, Cao HZ, Wang H, Wang J, Yao SS, Rakin A, Li YR, Falush D, Balloux F, Achtman M, Song YJ, Wang J, Yang RF. Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(2): 577–582.
- [11] Delcher AL, Salzberg SL, Phillippy AM. Using MUMmer to identify similar regions in large sequence sets. *Current Protocols in Bioinformatics*, 2003, (1): 10.3.1–10.3.18.
- [12] Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 2014, 30(14): 2068–2069.
- [13] Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, 2015, 31(22): 3691–3693.
- [14] Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, 2016, 44(W1): W242–W245.
- [15] Liu B, Zheng DD, Jin Q, Chen LH, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Research*, 2018, 47(D1): D687–D692.
- [16] Lees JA, Galardini M, Bentley SD, Weiser JN, Corander J. Pyseer: a comprehensive tool for microbial pangenome-wide association studies. *Bioinformatics*, 2018, 34(24): 4310–4312.
- [17] Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biology*, 2016, 17(1): 132.
- [18] Roig FJ, González-Candelas F, Sanjuán E, Fouz B, Feil EJ, Llorens C, Baker-Austin C, Oliver JD, Danin-Poleg Y, Gibas CJ, Kashi Y, Gulig PA, Morrison SS, Amaro C. Phylogeny of *Vibrio vulnificus* from the analysis of the core-genome: implications for intra-species taxonomy. *Frontiers in Microbiology*, 2018, 8: 2613.
- [19] Koo BS, Lee JH, Kim SC, Yoon HY, Kim KA, Kwon KB, Kim HR, Park JW, Park BH. Phospholipase A as a potent virulence factor of *Vibrio vulnificus*. *International Journal of Molecular Medicine*, 2007, 20(6): 913–918.
- [20] Liu X, Wang SH, Meng QM, Han XG, Xu X, Yang DH, Ding C, Peng DX, Yu SQ. Effects of type VI secretion system 2 core component VgrG on the pathogenicity of avian *Escherichia coli*. *Microbiology China*, 2016, 43(9): 2106–2113. (in Chinese)
刘新, 王少辉, 孟庆美, 韩先干, 许漩, 杨登辉, 丁铲, 彭大新, 于圣青. VI型分泌系统2核心组分VgrG对禽致病性大肠杆菌致病性的影响. *微生物学通报*, 2016, 43(9): 2106–2113.
- [21] Church SR, Lux T, Baker-Austin C, Buddington SP, Michell SL. *Vibrio vulnificus* type 6 secretion system 1 contains anti-bacterial properties. *PLoS One*, 2016, 11(10): e0165500.
- [22] Kudryashova E, Heisler DB, Kudryashov DS. Pathogenic mechanisms of actin cross-linking toxins: peeling away the layers//Mannherz HG. *The Actin Cytoskeleton and Bacterial Infection*. Cham: Springer, 2016: 87–112.
- [23] Kashimoto T, Ueno S, Ehara H, Fukudome S, Komai M, Susa N. Oligomerization is essential for apoptotic activity of *Vibrio vulnificus* hemolysin. *Journal of Veterinary Medical Science*, 2009, 71(10): 1403–1406.
- [24] Sukhanov S, Hou XW, Snarski P, Yoshida T, Delafontaine P. Smooth muscle specific glyceraldehyde-3'-phosphate dehydrogenase (GAPDH) reduces DNA damage, decreases cell apoptosis, suppresses atherosclerosis and promotes the stable plaque phenotype. *The FASEB Journal*, 2019, 33(S1): 522.1.
- [25] Rosinha GMS, Myioshi A, Azevedo V, Splitter GA, Oliveira SC. Molecular and immunological characterisation of recombinant *Brucella abortus* glyceraldehyde-3-phosphate-dehydrogenase, a T-and B-cell reactive protein that induces partial protection when co-administered with an interleukin-12-expressing plasmid in a DNA vaccine formulation. *Journal of Medical Microbiology*, 2002, 51(8): 661–671.
- [26] Ou P. Progress in Bacterium Type II Secretion System. *Journal of Hezhou University*, 2007, 23(3): 137–140. (in Chinese)
欧平. 细菌II型分泌系统的研究进展. *贺州学院学报*, 2007, 23(3): 137–140.
- [27] Abdelhamed H, Lu JJ, Lawrence ML, Karsi A. Involvement of *tolQ* and *tolR* genes in *Edwardsiella ictaluri* virulence. *Microbial Pathogenesis*, 2016, 100: 90–94.
- [28] Song J, Hou HM, Wu HY, Li KX, Wang Y, Zhou QQ, Zhang

- GL. Transcriptomic analysis of *Vibrio parahaemolyticus* reveals different virulence gene expression in response to benzyl isothiocyanate. *Molecules*, 2019, 24(4): 761.
- [29] Phippen BL, Oliver JD, Payne SM. Role of Anaerobiosis in capsule production and biofilm formation in *Vibrio vulnificus*. *Infection and Immunity*, 2015, 83(2): 551–559.
- [30] Kim IH, Kim IJ, Wen YC, Park NY, Park JY, Lee KW, Koh A, Lee JH, Koo SH, Kim KS. *Vibrio vulnificus* secretes an insulin-degrading enzyme that promotes bacterial proliferation *in vivo*. *Journal of Biological Chemistry*, 2015, 290(30): 18708–18720.
- [31] Nakhamchik A, Wilde C, Rowe-Magnus DA. Cyclic-di-GMP regulates extracellular polysaccharide production, biofilm formation, and rugose colony development by *Vibrio vulnificus*. *Applied and Environmental Microbiology*, 2008, 74(13): 4199–4209.

Genome-wide association study on virulence factors of *Vibrio vulnificus*

Jiaxin Zhang^{1,2}, Yarong Wu², Chao Yang², Lin Kang², Jinglin Wang^{1,2*}, Yujun Cui^{1,2*}

¹ Anhui Medical University, Hefei 230032, Anhui Province, China

² State Key Laboratory of Pathogen and Biosecurity, Institute of Microbiology and Epidemiology, Academy of Military Medical Sciences, Beijing 100071, China

Abstract: [Objective] *Vibrio vulnificus* is a pathogen with the highest mortality in *Vibrio* genus. However, currently there is lacking knowledge of virulence-associated factors across the whole genome wide for this species. By using bacterial genome-wide association study (BGWAS), we investigated the genes that are related with different sources (clinical/environment) of *V. vulnificus* isolates, aiming to provide further understanding on genetic factors that potentially influence its pathogenicity. [Methods] In total 260 *V. vulnificus* genomes were analyzed, including 139 newly sequenced in this study and 121 published genomes. We identified the genomes variations, including SNPs and accessory genes from the dataset, and then by using software pyseer, we determined the associations between sources of isolates and SNPs, genes, and Kmers, separately. We annotated the function of the statistically significant genes and inferred their possible influence on virulence of the bacteria. [Results] We found that 11 genes to be significantly correlated with clinical isolates of *V. vulnificus*, among which 9 were newly identified in this study. [Conclusion] By using methods of population genomics and statistical genetics, we screened the potential virulence-associated factors of *V. vulnificus* across the whole genome. Our findings provided new targets on vaccine designing and drug development, and would help to get further insight on pathogenic mechanism of this pathogen.

Keywords: *Vibrio vulnificus*, whole-genome sequencing, genome-wide association study, virulence factor

(本文责编: 张晓丽)

Supported by the National Key R&D Program of China (2018YFC1603902, 2017YFC1601500)

*Corresponding authors. E-mail: Yujun Cui, cuiyujun.new@gmail.com; Jinglin Wang, wjlwjl0801@sina.com

Received: 31 December 2019; Revised: 5 February 2020; Published online: 12 March 2020