

文章编号:1001-764X(2013)10-728-03

生物信息学技术在细菌耐药机制研究中的应用

糜祖煌^{1,2}(1. 无锡新区虎马生物信息工作室, 江苏无锡 214026; 2. 无锡市克隆遗传技术研究所, 江苏无锡 214026)

摘要:生物信息学(bioinformatics)是基于分子生物学理论,并与生物学、数学、计算机科学、信息学融合而成的交叉复合学科。收集、整理、解析生物学实验数据是生物信息学的主要任务。本文扼要介绍了该技术与细菌耐药相关基因分析中的应用。包括可将测得序列与数据库比对,测得的变异基因序列翻译成氨基酸序列委托 Swiss-Model Workspace 作三维结构同源建模,进行可视化分析与分子对接分析。将测序所得细菌全基因组序列上传至基因自动注释网站,可获得编码基因序列和相应的蛋白质序列;上传至日本京都基因和基因组百科全书(KEGG)在线注释工具 KAAS,可获得细菌代谢途径示意图。利用基因组数据库也可进行全基因组数据分析。

关键词:生物信息学;细菌;耐药机制

中图分类号:R446.5

文献标志码:A

生物信息学(bioinformatics)是建立在分子生物学理论基础上的由生物学、数学、计算机科学、信息学构成的交叉融合学科。生物信息学不仅包括收集、组织与整理、储存、分析和解释分子生物学实验数据,也包括分子生物学实验数据的可视化。这些实验数据可以是基因与基因组(含转录组)数据、蛋白质组数据或代谢组数据。在近二三十年中,生物信息学专家已推出了众多的工具软件和形形色色的数据库,为生命科学研究提供了强大工具。

细菌耐药机制研究大致可分为两类。一类为对所收集细菌菌株进行耐药相关基因的检测分析;另一类为对耐药菌株进行全基因组(含质粒)测序分析。这两类研究均需要生物信息学技术作支撑。本文就生物信息学技术在细菌耐药机制研究中的应用作简单介绍。

1 生物信息学技术与细菌耐药相关基因检测分析中的应用

细菌耐药形成因素主要有两类。(1)抗菌药物作用靶位编码基因突变,如肺炎链球菌青霉素耐药为青霉素作用靶位青霉素结合蛋白(PBPs)编码基因突变所致;结核分枝杆菌链霉素耐药为链霉素作用靶位核糖体小亚单位编码基因突变所致;细菌耐喹诺酮类药物为喹诺酮类药物作用靶位 DNA 旋转酶编码基因突变所致^[1]。(2)细菌借助于可移动遗传元件,包括细菌的质粒、噬菌体、转座子(含插入序列)和整合子等获得耐药基因,所表达的产物或灭活抗菌药物、或保护抗菌药物作用靶位、或外排抗

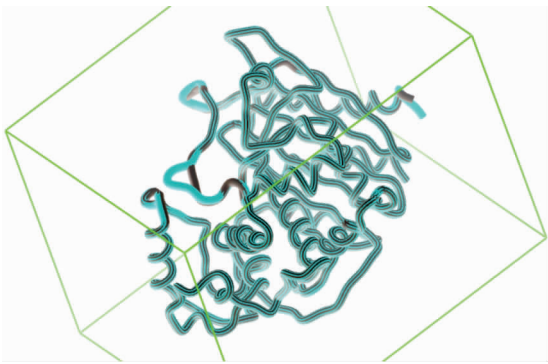
菌药物而导致耐药^[1-4]。这两种耐药相关基因的分析均需要采用 PCR 及 DNA 测序,测得 DNA 序列经 DNA 测序读序工具软件人工删除信号不佳的序列,并上网与核酸数据库(<http://www.ncbi.nlm.nih.gov/nucleotide>)比对方可确认。

1.1 DNA 测序读序工具软件与网上核酸数据库比对 目前 DNA 测序均用全自动荧光法。测得的 DNA 序列首先应经读序工具软件校正。前 10~20 个碱基荧光信号弱,以及 PCR 扩增时在 Taq DNA 聚合酶作用下 3' 末端额外延伸碱基 A 序列,因而需要经 DNA 测序读序工具软件人工删除相关片段。DNA 测序读序工具软件常用的有 Chromas 2.23(<http://www.technelysium.com.au/chromas.html>)和分子进化遗传分析(Molecular Evolutionary Genetics Analysis)工具包 MEGA 5.2(<http://www.megasoftware.net>)。经读序工具软件读序校正后的序列可直接点击 BLAST Search 与美国国立生物信息中心(NCBI)核酸数据库比对。并获得提交序列与相匹配序列的两两比对信息,其中高得分(score)与低期望值(expect value)的提示序列一致性(identities)高。点击 BLAST 页面 Distance tree of results 即可获得 NCBI BLAST 提供的提交序列与核酸数据库中不同程度匹配序列的分子进化分析图。

目前,全自动荧光 DNA 测序仪每次测序长度仅为 700~800 个碱基。鲍曼不动杆菌 β -内酰胺类药物作用靶位 PBPI 编码基因长达 2 556 个碱基。因此,PBPI 编码基因需重叠分段 PCR 扩增后作双向测序^[5],然后用 DNAMAN 工具软件(<http://www>。

lynonn.com) 拼接成全长序列。DNAMAN 工具软件也能读序校正和作分子进化分析。

1.2 分子三维结构同源建模与分子可视化 如果所研究的抗菌药物作用靶位或抗菌药物灭活酶已有三维结构报道, 那测得变异的基因序列翻译成氨基酸序列后, 可委托瑞士蛋白质数据库的瑞士模型工作区 (Swiss-Model Workspace, <http://www.swissmodel.expasy.org>) 作三维结构同源建模, 得到结果后可用分子可视化工具软件打开并与已有三维结构进行比较分析。分子可视化工具软件有 SWISS-MODEL (www.swissmodel.expasy.org)、ArgusLab 4.1 (<http://www.arguslab.com>)、UCSF Chimera 1.6.2 (<http://www.cgl.ucsf.edu/chimera>) 等。凌月明等^[6]在泛耐药鲍曼不动杆菌中发现了 β -内酰胺酶基因新亚型 ADC-65、ADC-66。典型的 ADC 型 β -内酰胺酶基因长度为 1 152 个碱基, 而 ADC-65 型基因长度为 1 158 个碱基, 可翻译为 385 个氨基酸 (相对于 ADC-66, ADC-65 插入了 6 个碱基)。委托 Swiss-Model Workspace 作三维结构同源建模后, ADC-65、ADC-66 分子可视化的相似性分析见图 1^[6]。



注: ADC-65 型 β -内酰胺酶蛋白质分子立体结构为天蓝色; ADC-66 型 β -内酰胺酶蛋白质分子立体结构为黑色; 天蓝色与黑色相嵌为两个分子重叠部分。本图引用得到原作者的许可。

图 1 ADC-65、ADC-66 分子可视化的相似性分析

1.3 分子对接 (molecular docking) 分析 变异序列作同源建模后不仅可用作可视化分析, 并可借助分子对接工具软件进一步作大分子 (蛋白质分子) 与小分子 (化合物) 相互作用分析。

分子对接是依据配体与受体作用的“锁-钥原理”, 模拟小分子配体与受体生物大分子的相互作用。主要作用力包括静电作用、氢键、疏水作用、范德华力等非共价键^[7]。在美国 NCBI 的 PubChem Compound 库 (<http://www.ncbi.nlm.nih.gov>) 可获取小分子化合物三维结构, 使用 ArgusLab 4.1 (<http://www.arguslab.com>) 软件中的 DOCK 模块即可

作大分子与小分子分子对接分析。国内已有多篇报道关于 β -内酰胺酶基因新亚型与 β -内酰胺类药物结合的自由能分析^[7-10], 结合自由能下降值大, 即催化活性最高^[7]。

分子对接分析还可用在喹诺酮类药物耐药机制研究中。喹诺酮类药物为一类人工合成的化合物, 作用于细菌 DNA 旋转酶或拓扑异构酶 IV, 使得细菌 DNA 无法复制而起抑菌作用。其中, DNA 旋转酶 A 亚单位编码基因 *gyrA* 第 83、87 位密码子最常见突变, 该区域被称为喹诺酮耐药决定区 (QRDR)。Vashist 等^[11]所作的喹诺酮类药物与细菌 DNA 旋转酶 A 亚单位分子对接报告显示, *gyrA* 编码基因第 83、87 位密码子突变后, 喹诺酮类药物与之结合位置发生改变, 所形成原子对的数量也减少。最近, 笔者等^[12]也报道了 2 种肺炎克雷伯菌 DNA 旋转酶 A 亚单位 QRDR 变异型与各种底物结合能力变化, 结论与 Vashist 等相近。

2 耐药菌株全基因组测序分析

随着新一代测序 (next-generation sequencing, NGS) 技术的进步与廉价化, 细菌全基因组测序技术已被应用到耐药机制研究中^[13-15]。经 NGS 测序获得某株细菌全基因组序列只需上传至基因自动注释网站 GeneMarkS (<http://www.exon.gatech.edu/genemarks.cgi>), 即可获得编码基因序列和相应的蛋白质序列。若将细菌全基因组序列上传至日本京都基因和基因组百科全书 (Kyoto Encyclopedia of Genes and Genomes, KEGG, <http://genome.jp/kegg>) 在线注释工具 KAAS, 不仅可获得基因自动注释, 并可获得代谢途径示意图 (pathway mapping)。

研究者获得细菌全基因组基因自动注释后, 可对抗菌药物作用靶位编码基因以及与耐药相关的膜孔蛋白编码基因、外排泵蛋白编码基因进行突变分析, 并搜索分析质粒、噬菌体、转座子 (含插入序列) 和整合子等可移动遗传元件, 获得耐药基因序列, 即可解析该菌耐药形成机制。

研究者可以对细菌全基因组数据库进行数据挖掘分析。截止到 2013 年 7 月 23 日, 基因组数据库 www.biocyc.org 已收录了 2 988 个全基因组数据。运用该基因组数据库作直系同源基因搜索分析十分方便。如: 该基因组数据库收录了鲍曼不动杆菌 SDF 株、ATCC 19606 株、ATCC 17978 株、AB307-0294 株、AB0057 株、6014059 株、6013150 株、6013113 株、AYE 株、ACICU 株等 10 株全基因组数据。经多基因组比

对,仅 SDF 株、6014059 株、6013150 株、6013113 株等 4 株无 AmpC 酶编码基因。研究者可将测得的细菌全基因组数据与数据库中数据进行比较分析。朱健铭等^[17]将测得的肺炎克雷伯菌全基因组数据与美国核酸数据库 (<http://www.ncbi.nlm.nih.gov/nucleotide>) 中已登录的 6 株肺炎克雷伯菌全基因组数据分别进行 9 种管家基因分子进化分析、多位点序列分型以及基因组系统学分析,结果提示在菌株亲缘性关系分析中,基于全基因组序列信息的基因组系统学分析法优于单个基因序列的分子进化分析法和多位点序列分型。

3 结语

生物信息学理论与技术仍在快速发展,工具软件不断改版,各种数据库也不断更新。随着 NGS 技术的进一步廉价化,细菌耐药机制研究会更多地进行全基因组分析与全转录组分析,菌株间的亲缘性分析也会基于更为可靠的全基因组聚类分析。

4 参考文献

- [1] Nikaido H. Multidrug resistance in bacteria[J]. *Annu Rev Biochem*, 2009, 78:119-146.
- [2] 翁幸璧,糜祖煌.可移动遗传元件:耐药基因的载体[J]. *中国人兽共患病学报*, 2013, 29(4):389-397.
- [3] Colomer-Lluch M, Jofre J, Muniesa M. Antibiotic resistance genes in the bacteriophage DNA fraction of environmental samples[J]. *PLoS One*, 2011, 6(3):e17549.
- [4] Colomer-Lluch M, Imamovic L, Jofre J, et al. Bacteriophages carrying antibiotic resistance genes in fecal waste from cattle, pigs, and poultry[J]. *Antimicrob Agents Chemother*, 2011, 55(10):4908-4911.
- [5] 翁幸璧,糜祖煌.泛耐药鲍曼不动杆菌的 PBPIA、CarO 及 β -内酰胺酶的编码基因分析[J]. *中华临床感染病杂志*, 2012, 5(4):215-219.
- [6] 凌月明,兰小鹏,王建福.泛耐药鲍曼不动杆菌中发现 β -内酰胺酶基因新的变异型 ADC-65、ADC-66[J]. *中国抗生素杂志*, 2013,

38(2):155-158.

- [7] 糜祖煌,翁幸璧,秦玲. KPC 型碳青霉烯酶分子进化及与底物结合自由能分析[J]. *中华临床感染病杂志*, 2010, 3(3):134-137.
- [8] 周军,王玉月,张秋娣. ADC-57 型头孢菌素酶分子进化及与底物结合自由能分析[J]. *中华临床感染病杂志*, 2012, 5(2):77-80.
- [9] 朱健铭,姜如金,肖丹宇,等. KPC-12 型碳青霉烯酶分子进化及与 β -内酰胺类药物分子对接分析[J]. *中华临床感染病杂志*, 2013, 6(1):31-34.
- [10] 原鸿雁,孙强,尹晶平,等. 耐药大肠埃希菌 β -内酰胺酶基因检测与 LAP-2 型 β -内酰胺酶分子对接研究[J]. *中华医院感染学杂志*, 2013, 23(10):2267-2270.
- [11] Vashist J, Vishvanath, Kapoor R, et al. Interaction of nalidixic acid and ciprofloxacin with type and mutated quinolone-resistance-determining region of DNA gyrase A[J]. *Indian J Biochem Biophys*, 2009, 46(2):147-153.
- [12] 糜祖煌,翁幸璧,高军晖.肺炎克雷伯菌 DNA 旋转酶 A 亚单位喹诺酮耐药决定区变异型与底物分子对接分析[J]. *中华传染病杂志*, 2013, 31(4):208-211.
- [13] Zhou H, Zhang T, Yu D, et al. Genomic analysis of the multidrug-resistant *Acinetobacter baumannii* strain MDR-ZJ06 widely spread in China[J]. *Antimicrob Agents Chemother*, 2011, 55(10):4506-4512.
- [14] Suzuki H, Lefebvre T, Bitar PP, et al. Comparative genomic analysis of the genus *Staphylococcus* including *Staphylococcus aureus* and its newly described sister species *Staphylococcus simiae*[J]. *BMC Genomics*, 2012, 13:38.
- [15] Köser CU, Holden MTG, Ellington MJ, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak[J]. *N Engl J Med*, 2012, 366(24):2267-2275.
- [16] Harris SR, Feil EJ, Holden MT, et al. Evolution of MRSA during hospital transmission and intercontinental spread[J]. *Science*, 2010, 327(5964):469-474.
- [17] 朱健铭,姜如金,吴晋兰,等.7 株肺炎克雷伯菌 9 种管家基因分子进化分析、多位点序列分型与基因组系统学研究初步报道[J]. *中华微生物和免疫学杂志*, 2013, 33(4):257.

(收稿日期:2013-07-24)

(本文编辑:刘群,陈维忠)