

• 微生物组测序与分析专题 •

王荃 女，博士，天津医科大学教授、博士生导师。长期围绕细菌与宿主的相互作用，研究病原细菌的致病机制和抗感染策略，通过建立病原菌株的精细分子分型方法进行其流行病学和进化分析。在 *Journal of Pathology*、*Journal of Clinical Microbiology*、*Molecular Microbiology* 等杂志发表通讯或第一作者文章 20 余篇，影响因子超过 90，引用 230 余次。参与 Springer 出版社出版的《Endotoxins: Structure, Function and Recognition》专著的编写。主持国家科技重大专项子课题、国家自然科学基金、天津市自然科学基金等多项课题。作为骨干成员获得天津市自然科学一等奖。获选天津市“131”第一层次人才。担任中华医学会微生物与免疫学分会青年委员，中华医学会细胞生物学分会青年委员。



耐药基因数据库概述

杨兵^{1*}，梁晶^{2*}，刘林梦¹，李雪佩¹，王荃³，任一¹

- 1 上海美吉生物医药科技有限公司，上海 201318
- 2 上海交通大学附属第一人民医院 急诊危重病科，上海 200080
- 3 天津医科大学 免疫学系，天津 300070

杨兵，梁晶，刘林梦，等. 耐药基因数据库概述. 生物工程学报, 2020, 36(12): 2582–2597.

Yang B, Liang J, Liu LM, et al. Overview of antibiotic resistance genes database. Chin J Biotech, 2020, 36(12): 2582–2597.

摘要: 抗生素是人类历史上的革命性发现，其临床应用挽救了无数患者的生命。但是随着抗生素的广泛使用和滥用，越来越多的病原菌产生了耐药性，甚至出现了具有多重耐药性的“超级细菌”。在人类与病原菌斗争的军备竞赛中，人类即将面临无药可用的境地。针对微生物的耐药性、耐药机制及耐药性传播的研究吸引了众多科研工作者的目光，各种耐药基因数据库以及耐药基因分析工具应运而生。文中对目前耐药领域的基因数据库进行收集整理，从数据库类型、数据特征、耐药基因预测模型以及可分析序列的类型等方面对这些数据库进行论述和介绍。此外，文中对抗金属离子和抗杀菌剂的基因数据库也有所涉及，将为如何选择及使用耐药基因数据库提供参考和帮助。

关键词: 抗生素，耐药基因，超级细菌，可移动元件

Received: June 24, 2020; **Accepted:** September 22, 2020

Supported by: Science & Technology Development Fund of Tianjin Education Commission for Higher Education (No. 2017ZD12).

Corresponding author: Quan Wang. Tel: +86-22-83336817; E-mail: wangquan@tmu.edu.cn

*These authors contributed equally to this study.

天津市教委科研计划重点项目 (No. 2017ZD12) 资助。

网络出版时间: 2020-11-09

网络出版地址: <https://kns.cnki.net/kcms/detail/11.1998.q.20201106.1008.001.html>

Overview of antibiotic resistance genes database

Bing Yang^{1*}, Jing Liang^{2*}, Linmeng Liu¹, Xuepei Li¹, Quan Wang³, and Yi Ren¹

¹ Shanghai Majorbio Biomedical Technology Co. Ltd., Shanghai 201318, China

² Department of Critical Care Medicine, Shanghai General Hospital, Shanghai Jiao Tong University, Shanghai 200080, China

³ Department of Immunology, Tianjin Medical University, Tianjin 300070, China

Abstract: The discovery of antibiotics is a big revolution in human history, and its clinical application has saved countless lives. However, with the widespread and abuse of antibiotics, many pathogens have developed resistance, and even “Super Bacteria” resistance to multiple drugs have evolved. In the arms race between humans and pathogens, humans are about to face a situation where no medicine is available. Research on microbial antibiotic resistance genes, resistance mechanisms, and the spread of resistance has attracted the attention of many scientific researchers, and various antibiotic resistance gene databases and analysis tools have emerged. In this review, we collect the current databases that focus on antibiotics resistance genes, and discuss these databases in terms of database types, data characteristics, antibiotics resistance gene prediction models and the types of analyzable sequences. In addition, a few gene databases of anti-metal ions and anti-biocides are also involved. It is believed that this summary will provide a reference for how to select and use antibiotic resistance gene databases.

Keywords: antibiotics, antibiotic resistance genes, super bacteria, mobile genetic elements

抗生素是人类历史上的革命性发现,自1929年第一种抗生素青霉素被发现以来,抗生素得到广泛使用,并挽救了无数患者的生命^[1]。但是随着抗生素的滥用,大量微生物产生了抗生素耐药性,细菌耐药的日益增长对全球公共卫生产生极大威胁。据推测,到2050年,全球死于抗生素耐药细菌的人数每年将超千万,花费约100万亿美元^[2-4]。研究者对2013年抗生素使用的调查中发现,中国抗生素的总产量和总使用量分别为24.8万t和16.2万t,为世界最高水平,其中48%的抗生素是人类使用的,其余为动物使用的^[5]。由于抗生素在医疗和畜牧业中的滥用和误用,大部分抗生素(约有30%–90%)被释放到各种环境基质中^[6-8],环境中残留的抗生素在长时间的作用下会加速环境中抗生素抗性细菌(Antibiotic resistance bacteria, ARB)和抗生素抗性基因(Antibiotic resistance genes, ARGs,也被称为耐药基因)的产生和传播。近年来,在我国的不同环境样品中都发现了ARGs的存在,如地表水^[9-10]、城市废水^[11-12]、地下水^[13]、医院废水^[14-15]、沉积物^[16]、淤泥^[17-18]、土壤^[19-20]、粪便^[21-22]以及空气中^[23]等。在自然环境中,抗生

素抗性基因可通过质粒、整合子、转座子等可移动元件在不同的细菌之间进行传播,甚至在非致病菌与致病菌之间传播^[24],形成“超级细菌”,如耐甲氧西林金黄色葡萄球菌(Methicillin-resistant *Staphylococcus aureus*, MRSA)、万古霉素肠球菌(Vancomycin-resistant *Enterococcus*, VRE)、携带有NDM-1基因的大肠杆菌和肺炎克雷伯菌等。超级细菌经常有医院获得感染或社区获得感染,由于大部分抗生素对其不起作用,超级细菌对人类生命健康已造成极大的危害^[25-26]。

常见的细菌耐药机制包括主动外排、膜通透性降低、靶蛋白改变、产生灭活酶等。主动外排是指细胞膜上的外排泵可以主动将进入细胞内的抗生素泵出体外:研究发现在大肠杆菌中存在9种可用于四环素外排的基因(*tetA*、*tetB*、*tetC*、*tetD*、*tetE*、*tetG*、*tetJ*、*tetL*和*tetY*)^[27],从而产生四环素耐药;膜通透性降低,是由于细胞膜或细胞壁结构发生改变,从而使药物难以进入细胞内:如某些大肠杆菌会发生变异导致孔蛋白OmpF缺少使得喹诺酮类药物无法进入细胞从而产生耐药性^[28];与抗生素相结合的靶蛋白发生突

变或者被修饰后,影响了抗生素与靶点的结合:如 DNA 解旋酶和拓扑异构酶基因发生突变后,可能会改变与喹诺酮类药物的结合位点,从而降低此类药物的敏感性^[27];细菌会产生一种或多种水解酶或钝化酶来水解或修饰进入细胞内的抗生素,使其到达靶位之前失去活性:如某些细菌可表达 β -内酰胺酶, β -内酰胺酶可裂解 β -内酰胺环使得此类抗菌药物失去杀菌活性,从而导致细菌对此类抗菌药物产生耐药性^[28]。

ARGs 作为一种新兴的污染物^[29],因其对生态和人类健康的巨大威胁而受到世界范围内广泛关注。当前,联合国环境署已经将耐药菌及耐药基因列为 6 种新型环境污染物中的首位^[29],世界卫生组织 (WHO) 于 2014 年发表了关于抗生素耐药对公共卫生造成严重的全球性威胁的报告^[30]。随后我国在国家卫生计生委、发展改革委等 14 个部门联合下,制定出台了《遏制细菌耐药国家行动计划 (2016–2020 年)》^[31],旨在从国家层面多个领域打出组合拳,有效遏制细菌耐药,维护人民群众身体健康,促进经济社会可持续发展。之后在 2017 年杭州举办的 G20 峰会上,各国代表达成了一项联合抗击抗生素耐药的国际合作协议。

随着测序技术的发展,基因组的测序成本逐步降低,越来越多的微生物基因组和宏基因组测序完成^[32]。通过基因组测序,科研工作者可以快速了解目标微生物携带的耐药基因和可移动元件等。而耐药基因的生物信息学分析有赖于各类耐药基因分析工具和耐药基因数据库。

本文对目前主要的基因数据库进行检索和梳理,从数据库类型、数据特征、耐药基因预测模型以及可分析序列的类型等方面对这些数据库进行分析和介绍。结合各数据库的特性,我们对数据库进行了评分,为用户更好更准确地选择数据库提供指引。此外本文也对抗金属离子和抗杀菌剂的基因数据库有所涉及。相信文中总结将为该领域的科研工作者在如何选择耐药基因数据库方

面提供参考和帮助。

1 数据库类型及数据特征

本文收集整理了 25 个微生物抗性基因数据库 (大部分为耐药基因数据库) (具体数据作为附件,可在网络版中下载,见附表 1),重点针对其中 16 个常用数据库进行阐述 (表 1),根据数据库中收录的抗性基因的类型,将其分为通用耐药型、特异型 (药物特异和菌种特异)、综合型及其他抗性类型等。

1.1 通用耐药型数据库

通用耐药型数据库,即相对药物特异或菌种特异的特异型数据库而言,收录多物种和多种类药物的耐药基因数据库。在此,我们主要介绍一些常见的数据库,并概述其数据特征。

最早的通用耐药型数据库可追溯到 2009 年,Mihai Pop 等将公开发表的微生物耐药基因及其相关信息进行了整合和人工校对,构建了 ARDB 数据库 (Antibiotic resistance genes database, ARDB)^[33],其数据主要来源于公开发表的文献和数据库。作者收集了来自 NCBI 和 Swiss-Prot 数据库的 13 254 个耐药基因的序列,在进行数据合并和去冗余之后,保留了 4 554 条完整的耐药基因序列,然后根据耐药机制对这些基因进行了归类,从而得到数据库的核心架构。除此之外, ARDB 数据库还收录了 12 个抗生素药物靶标基因和其相关信息 (*16S rRNA*^[34]、*23S rRNA*、*gyrA*^[35]、*gyrB*、*parC*、*parE*、*rpoB*、*katG*、*pncA*、*embB*、*forP*、*fdr*),这些靶标对于研究细菌耐药性具有重要作用。目前 ARDB 已经不再维护,其数据库信息已经被整合入 CARD 数据库。

2012 年,Zankari 等在 ARDB 数据库的基础上搭建了 ResFinder 数据库^[36],整理了通过水平转移获得的耐药基因。对序列进行去冗余后,根据抗生素类型对基因进行分类,共保留了 15 种抗生素的 2 156 个耐药基因。最初 ResFinder 只提供

表 1 16 个常见耐药基因数据库特征表

Table 1 Characteristic of 16 common resistance genes databases

Database	Year	Description/characteristic	Link	Database type	Anti-bacterial/Species type	Statistics	Tool	Prediction model	Input	Last update	Citation	Score
ARDB	2009	The first comprehensive database of antibiotic-resistant genes collected from variety of microorganisms.	http://ar.db.cbcb.umd.edu/	UAR	A	4 554 antibiotic resistance proteins, 12 drug targets	Blast	Blast	Genome/ Gene	2009. 07	671	7
ResFinder	2012	Resistance genes obtained by horizontal transfer were collected and classified according to antibiotics, and the chromosome mutation information of microorganisms was also collected.	https://cbs.utd.edu/sFinder/	UAR	A	2 156 ARGs, 15 antibiotics, mutation information of 11 strains	Resfinder/ PointFinder	Blast	Genome/ Gene/PE, SE Reads	2020. 02	1 736	11
CARD	2013	Ontology-based database that provides comprehensive information of ARGs and their resistance mechanisms, home SNP information, AMR detection models, and Mobile Genetic Elements information.	https://card.mcmaster.ca/	UAR	A	4 465 ontology terms, 2 984 reference sequences, 1 334 SNPs, 3 032 AMR models, 311 antibiotics, 92 mobile genetic elements	RG/Blast	Blast, HMMs	Genome/ Gene	2020. 06	1 393	14
Resfams	2014	A curated database of antibiotic resistance genes based on protein families and associated profile HMMs, and the function and category were organized by ontology.	http://www.dant.aslab.org/resfams/	UAR	A	166 HMMs	/	HMMs	/	2018. 02	1	1
ARG-ANNOT	2014	To detect existing and possible new antibiotic resistance genes in bacterial genomes, and to identify chromosomal mutations.	http://www.med.illinois.edu/infocion.com/ar	UAR	A	1 689 ARGs	ARG-ANNOT	Blast	Genome/ Gene	2019. 07	482	5

续表 1

Database	Year	Description/characteristic	Link	Database type	Anti-bacterial/Species type	Statistics	Tool	Prediction model	Input	Last update	Citation	Score
FARME	2017	Functional Antibiotic Resistance Metagenomic Element Database, provides a resource to better understand AR in uncultured bacteria and the relationship between environmental AR sequences and antibiotic resistant genes derived from clinical cultured isolates.	http://staff.washington.edu/jwallace/farme/index.html	UAR	A	24 530 non-overlapping HMMs, 2 250 different models, 8 478 AR HMMs, 1 369 transcriptional regulator HMMs and 360 mobile genetic element HMMs	Blast	Blast	Genome/ Gene	2017	11	3
DeepARG-DB	2018	The DeepARG database greatly expands the available ARGs individually available in the currently most widely deeparg used CARD, ARDB, and UNIPROT databases, including their existing sequence content and extensive metadata. DeepARG provides a publicly-available database structured into a simple category and group hierarchy for each ARG.	http://bench.cs.vt.edu/	UAR	A	30 antibiotic categories, 2 149 groups, 14 933 reference sequences	DeepARG	Deep Learning	Metagenome /PE Reads	/	92	6
SARG (v2)	2018	The SARG database also known as Structured Antibiotic Resistance Gene database is a collection of antimicrobial resistance genes. It can classify the resistance genes into different types or subtypes. The ARGs-OAP tool can	http://smile.hku.hk/SARGs	UAR	A	23 types, 1 227 subtypes and 4 246 reference sequences	ARGs-OAP (v2)	Blastx, SARGfam	Metagenome /PE Reads	2018	37	4

续表 1

Database	Year	Description/characteristic	Link	Database type	Anti-bacterial/Species type	Statistics	Tool	Prediction model	Input	Last update	Citation	Score
ARGO	2005	quickly identify and quantitatively analyze antibiotic-resistant genes from the metagenome. The earliest database of antibiotic resistance genes, which only contains resistance genes of beta-lactamases and vancomycin.	http://www.argodb.org/	AS	Beta, Van	555 Beta-bactamase and 115 Vancomycin resistance genes	/	/	/	/	39	2
Lahey list of β -lactamases	2010	It updated the functional classification of beta-lactamases into 3 groups according to the hydrolysis and inhibition of different enzymes, and takes into account substrate and inhibitor profiles in an attempt to group the enzymes in ways that can be correlated with their phenotype in clinical isolates.	https://externalwebapps.lahey.org/studies/	AS	Beta	Group 1 (Class C) cephalosporinases; Group 2 (Class A and D) broad-spectrum, inhibitor-resistant, and extended-spectrum β -lactamases and serine carbenemases; Group 3 (Class B) metallo- β -lactamases.	/	/	/	/	1 287	6
BLDB	2017	Beta-Lactamase Database provides up-to-date structural and functional information on the enzyme superfamily, listed by category (A, B, C, and D), family and subfamily, and also shows all three-dimensional structures, representative mutants, and hydrolytic profiles of these beta-lactamases.	http://bladb.eu/	AS	Beta	4 940 unique enzymes, 1 230 three-dimensional structures, 167 mutants and 47 hydrolytic profiles	Blast	Blast	Genome/ Gene	2020.06	64	4
TBDReaMDB	2009	Tuberculosis Drug Resistance Database is a comprehensive resource on drug resistance mutations in <i>Mycobacterium</i>	http://www.tbdr.com/	SS	MT	946 unique mutations associated with 7 different drug classes and spread	/	/	/	/	422	5

续表 1

Database	Year	Description/characteristic	Link	Database type	Anti-bacterial/Species type	Statistics	Tool	Prediction model	Input	Last update	Citation	Score
		<i>tuberculosis</i> . The database provides complete codon changes for each mutation at both the nucleotide and amino acid level.				Over 36 genes						
MUBII-TB-DB	2014	MUBII-TB-DB is a highly structured text-based database that contains a set of <i>Mycobacterium tuberculosis</i> mutations occurring at seven loci of major therapeutic value.	https://mr5558-proka.univ-lyon1.fr/mubii/mubii-i.n.cgi	SS	MT	Seven loci: <i>rpoB</i> , <i>pncA</i> , <i>katG</i> , <i>mabA</i> (<i>fabG1</i>)- <i>inhA</i> , <i>gyrA</i> , <i>gyrB</i> , and <i>rrs</i>	Blast	Blast	Genome/ Gene	/	50	4
u-CARE	2015	User-friendly Comprehensive Antibiotic resistance Repository of <i>Escherichia coli</i> is a manually curated catalogue of antibiotics with reported resistance, genes, transcription factors and SNPs involved in multiple drug resistance of this pathogen.	http://www.e-biointformatics.net/ucare/	SS	EC	52 antibiotics and 107 genes	Blast	Blast	Genome/ Gene	/	6	3
BacMet	2014	BacMet is an easy-to-use bioinformatics resource of antibacterial biocide- and metal-resistance genes. BacMet consists of two databases: experimentally confirmed and predicted resistance genes databases.	http://cmet.bio.e.gu.se/	Other	M, B	753 Experiment Confirmed resistance genes (biocide 268, metal 420, both biocide- and metal-65), 155 512 predicted resistance genes	Blast	Blast	Genome/ Gene	2018.03	179	5
MEGARes (v2.0)	2017	A database that incorporates standardized accessions for a comprehensive set of previously published resistance determinants for antimicrobial drugs, biocides and metals.	https://egares.mglab.org/	Comprehensive	A, M, B	57 classes, 220 mechanisms of resistance and 1 345 ARGs	AmrPlusPlus	Blast, HMMs	Metagenome / Genome	2019	115	6

UAR: Universal Antibiotic Resistance; AS: Antibiotic-Specific; SS: Species-Specific; A: Antibiotic; M: Metals; B: Biocide; EC: *Escherichia coli*; MT: *Mycobacterium tuberculosis*; Beta: β -lactamase; Va: Vancomycin.

已知的耐药基因的查询和比对服务,不能分析与耐药基因有关的染色体靶基因的点突变。但在2017年时,Resfinder更新了其基于web的服务,可以使用PointFinder^[37]识别染色体突变信息。目前ResFinder只收录了11个菌株/物种的突变信息,包括弯曲杆菌*Campylobacter*、大肠杆菌*Escherichia coli*、结核分枝杆菌*Mycobacterium tuberculosis*、淋病奈瑟菌*Neisseria gonorrhoeae*、恶性疟原虫*Plasmodium falciparum*、沙门氏菌*Salmonella*、粪肠球菌*Enterococcus faecalis*、屎肠球菌*Enterococcus faecium*、幽门螺杆菌*Helicobacter pylori*、克雷伯氏菌*Klebsiella*和金黄色葡萄球菌*Staphylococcus aureus*。这些突变信息的识别,对于理清耐药突变的发生机制和开发新型抗生素方面具有重要的意义。

随着新的耐药菌株和耐药基因的不断出现,各数据库的更新速度成为科研者关注的重点。2013年,一个基于志愿者贡献数据的共享平台CARD (Comprehensive antibiotic resistance database)^[38]的出现解决了此项问题。CARD不仅包含了ARDB数据库的所有耐药基因信息,且每月更新一次以确保数据的时效性。CARD数据库以ARO (Antibiotic resistance ontology)为分类单位即term (共4 498个),并关联抗生素(共311种)及其靶信息、耐药机制(共7种)、基因变异(共1 337个SNPs)等信息。在该数据库中,还可以获取到一些毒力基因的信息、与耐药/毒力基因相关的可移动元件信息(共92个),以及AMR (Anti-microbial resistance)检测模型(共3 030个)。用户基于CARD不仅可以检索目标菌株中已确定的耐药基因,还可以利用CARD提供的RGI工具预测潜在的耐药基因。RGI提供了3种预测标准,即Perfect、Strict和Loose;通过选择不同的标准,可以得到不同可信度和数量的潜在耐药基因,对新型耐药基因的鉴定具有重要的参考意义。CARD数据库已成为目前最受科研工作者青睐的耐药基因研究工具之一。

ARG-ANNOT数据库^[39]通过对公共数据库进行整合并去除冗余之后得到1 689个耐药基因。用户可以检索/预测细菌基因组中已知的和潜在的耐药基因,以及包括预测由靶基因突变产生的耐药性。但该数据库目前处于关闭状态。

2014年,Dantas实验室在CARD和LacED (Lactamase engineering database)等数据库的基础上,搭建了基于蛋白家族和隐马尔可夫模型(Hidden markov model, HMMs)的Resfams数据库^[40],共含有166个HMMs。同基于Blast搜索ARDB、CARD数据库预测耐药基因的方法相比,Resfams HMM能够预测更多的未知耐药基因。通过分析6 000余种微生物基因组,研究者发现耐药基因在不同系统发育分类和不同生态环境中存在富集现象。

为更好地了解不可培养细菌中的耐药基因情况,以及耐药基因在环境菌株与临床培养分离株之间的关联,Wallace等构建了FARME (Functional antibiotic resistance metagenomic element database)数据库^[41]。通过将GenBank中的20个功能宏基因组项目的蛋白质或核酸序列数据与NCBI-NR或NCBI-NT数据库进行Blast比对,从而根据项目信息获得每个序列的AR (Antibiotic resistance)分类信息;然后使用HMMER软件的HMMSEARCH模块在蛋白质序列中搜索Pfam、Resfams数据库中AR保守域(Domain)的隐马尔可夫模型HMMs,在score值大于模型最小得分阈值的HMMs中,选择bitscore最高的HMM作为该蛋白唯一的注释。最终获得24 530个不重叠的HMMs,其中包括:耐药基因(共8 478个)、调控元件(共1 369个)和可移动元件(共360个)。

DeepARG-DB^[42]是对CARD、ARDB和Uniprot数据库中的耐药基因进行信息整合而得到的数据库。首先根据相关关键词从Uniprot数据库中提取候选的耐药基因序列,然后与CARD、ARDB数据库合并,进行序列的聚类 and 去冗余。DeepARG-DB对ARDB和CARD中的抗生素分

类信息进行了手工校对,得到 30 个抗生素类别 (Antibiotic categories)。通过序列聚类信息 (Cluster) 和 text-mining 手段,补充了来自 Uniprot 数据库的序列注释信息,并依据其与 CARD/ARDB 中序列比对情况对注释信息的可信度进行分级。最终该数据库为每一个 ARG 提供了抗生素类别 (Antibiotic category) 和分组 (Group) 信息,共包含 14 933 个耐药基因、30 种 categories、2 149 个 groups。此外,DEEP-ARG 提供了基于机器学习算法的耐药基因分类鉴定工具,主要用于鉴定宏基因组中的 ARGs。

SARG (v2)^[43]是一个 Type-subtype-reference sequence 结构化 ARG 数据库。首先该数据库将 ARDB 和 CARD 数据库进行合并去冗余,得到 SARG v1.0;然后以 SARG v1.0 作为 ARG 种子序列,与 NCBI-NR 数据库进行 Blastp 比对,基于序列完整度、相似度及关键词匹配对 NR 库中的潜在的 ARG 序列进行过滤和分类,最后将得到的新序列与 SARG v1.0 合并后得到最终的数据库,共包含 23 个 types、1 227 个 subtypes 和 4 246 个 reference sequences。该数据库主要用来鉴定并定量分析宏基因组中的耐药基因。

1.2 特异型数据库

特异型数据库分为药物特异和菌种特异两种,即数据库是针对某一特定抗生素或特定种属的细菌构建的数据库。

ARGO 数据库 (Antibiotic resistance genes online)^[44]是最早出现的关于耐药基因的数据库,但只包含 β -内酰胺酶 (共 555 个) 和万古霉素 (共 115 个) 这两类抗生素的耐药基因。 β -内酰胺类抗生素的耐药机制研究是耐药领域的一大热点,因此涌现出很多与其相关的耐药基因数据库。其中最早且最著名的当属 Lahey 实验室整理的 Lahey list of β -lactamases^[45],该数据库打破了之前基于蛋白序列对酶进行分类的分子分型方法 (该方法将 β -内酰胺酶分为 class A、B、C、D

四种类型),综合考虑其底物和抑制剂的特征对 β -内酰胺酶进行功能分类,使得各功能类别可以与临床分离株的表型相关联。Lahey 数据库将 β -内酰胺酶划分为 3 个类型 group,其中 group1 (Class C) 为头孢霉素酶,group2 (Class A 和 D) 为广谱的、抑制剂抵抗的内酰胺酶和丝氨酸碳青霉烯酶,group3 (Class B) 为金属 β -内酰胺酶。目前收录信息较全的 β -内酰胺酶数据库是 BLDB (β -lactamase database)^[46],该数据库汇总整理了多个数据库中的 β -内酰胺酶信息,提供包括超家族分类、蛋白序列、三维结构、突变、菌株信息、抗生素水解动力学特征谱、基因位置、抑制剂敏感性等多方面的信息。BLDB 数据库共包括 4 940 个基因、1 230 个结构、167 个变体、47 种水解动力学特征谱。

除了针对特定抗生素的耐药基因数据库之外,还有针对特定菌株构建了耐药基因的数据库,如 TBDReaMDB (Tuberculosis drug resistance mutation database)^[47]、MUB II -TB-DB^[48] 和 u-CARE^[49]。TBDReaMDB 和 MUB II -TB-DB 是针对结核分枝杆菌构建的耐药基因数据库。其中,TBDReaMDB 收录了结核分枝杆菌中与耐药相关的基因突变信息,包含与 7 种不同药物相关的 946 个独特突变;而 MUB II -TB-DB 收录了对结核病具有重要治疗价值的 7 个基因的突变信息,包括 *rpoB*、*pncA*、*katG*、*fabG1*、*gyrA*、*gyrB* 和 *rrs* 等。u-CARE^[49] 是一个大肠杆菌耐药基因数据库,包括 107 个耐药基因和 52 种抗生素。

特异型数据库由于其针对性强,在研究特定菌种和特定抗生素的耐药性方面具有重大意义^[50],同时这些数据库方便快捷有效地管理新的抗生素耐药基因和与之相关的染色体突变信息^[51]。

1.3 其他抗性数据库

在医疗环境和自然环境中,一些金属离子和杀菌剂的使用也是微生物获得抗性的来源,2014 年,Pal 等构建了 BacMet (Antibacterial biocide and

metal resistance genes database)^[52]数据库, 包含了抗金属离子的基因和抗杀菌剂的基因, 共计 753 个基因 (其中抗金属离子的基因 420 个、抗杀菌剂的基因 268 个、两者都抗的基因有 65 个)。基于该数据库, 用户不仅可以识别目标菌株中的抗金属离子和抗杀菌剂的基因, 还可以鉴定到同时具有两种抗性的基因, 对于菌株抗性研究具有重要意义。该数据库除参考库 (收录经实验验证的基因) 之外, 还提供一个未经过实验验证的预测基因库。预测库包含 155 512 个基因, 仅供用户作为参考, 其准确性还需要通过实验做进一步验证。

1.4 综合型数据库

综合型数据库是指除耐药基因信息之外, 还收录抗金属离子或抗杀菌剂等抗性基因的数据库。研究显示, 在重金属污染的环境中细菌不仅具备重金属抗性, 并且具备多种抗生素抗性。在抗生素、重金属离子和杀菌剂等杀菌物质交叉污染的复杂环境中, 存在抗性基因的协同选择、交叉作用和共调控等多种机制导致微生物产生多重抗性^[53]。

2017 年 Lakin 等^[54]搭建了 MEGARes 数据库, 该数据库中包含抗生素、金属离子和杀菌剂的抗性基因共 1 345 个。根据可抵抗物质的类型不同将基因划分为耐药、抗金属、抗杀菌剂和抗多化合物 4 类, 其中有 461 个基因属于抗多化合物类型, 如 *cmeABC* 基因, 具有对多种抗生素、金属和杀菌剂的抗性。

综合型数据库除了可以鉴定微生物具有多抗性的特征之外, 对于多种抗性之间的协同作用、交叉作用以及共调控作用的研究具有重大意义。

2 预测模型

目前耐药基因的预测, 主要有 3 大类模型和方法 (表 1): (1) 基于序列相似度的预测模型, 如 Blast; (2) 基于保守序列或保守结构域的预测模型, 如 HMMs; (3) 基于机器学习的预测模型,

如 DeepARG。

Blast 是常见且使用率较高的基于序列相似度比对的基因功能预测工具, 用于多个耐药基因数据库中, 如 BacMet、Resfinder 以及 ARG-ANNOT 等。但通过 Blast 预测得到的结果多为功能相似且比对一致性较高的序列。根据序列相似度程度, 可以对结果的可信度进行如下分类^[38]: (a) perfect: 序列完全一致; (b) high: score > score (ref vs ref)*90%; (c) modular: identity > 80% AND align_coverage > 80%; (d) low modular: identity > 80% OR align_coverage > 80%; (e) other: e-value cutoff (default 1e-5)。

基于序列相似度的预测模型, 当查询基因同数据库中的参考基因之间的进化距离太大, 导致序列相似度不高时, 经常失效。为了避免上述情况, MEGARes、CARD、Resfams 和 ARGs-OAP 2.0 等数据库增加了耐药基因的隐性马尔可夫预测模型。HMM 预测模型将蛋白质序列多重比对结果转化成位点特异性的打分矩阵, 由于考虑了不同保守度的氨基酸在相应位置的权重, 它可以更为敏感地检测到进化距离较远的蛋白质的相关性, 得到比 Blast 序列预测模式方法更为灵敏的结果。HMMs 可以识别大量新的耐药基因和已知耐药基因的远源基因, 对确定未知环境样本中的耐药基因具有重要的意义。HMMs 方法也有其缺点, 在一些情况下, 它可能无法区分具有密切相关功能的序列之间的特征, 从而产生假阳性结果^[51]。

机器学习算法直接从数据中“学习”信息, 而不依赖于预定方程模型, 在图像处理、信号处理、语音识别、计算生物学等领域得到广泛应用。2018 年 Arango-Argoty 等提出了一种基于深度学习的耐药基因预测方法 DeepARG^[42], 它针对所有已知类别的耐药基因创建相异性矩阵 (Dissimilarity matrix), 然后分别建立针对短序列和全基因长度的深度学习模型 DeepARG-SS 和 DeepARG-LS。这种深度学习模型考虑到序列在

整个耐药基因数据库中的相异性分布,而不是只考虑“Best hit”,它能够比传统方法预测出更多的耐药基因,特别是在各种未知环境样本和宏基因组样本中。

以上3种预测模型各有优缺点,如何实现快速有效且科学的耐药基因预测,仍然是目前需要研究并加强的部分。

3 可分析序列的类型

随着高通量测序技术的广泛应用,产生了海量的细菌基因组和宏基因组数据。耐药基因的分析不再局限在单个基因的水平,针对细菌基因组草图、基因组完成图,宏基因组拼接结果、宏基因组 binning 结果等水平进行耐药分析的需求日益增多。特别是由于宏基因组数据过于复杂,经常导致基于拼接的宏基因组数据不能很好地代表样本本身,因而产生了很多基于读长 (Reads) 的功能分析需求。我们对各耐药数据库自带预测工具的可输入序列类型进行了总结 (表 1)。从中可以看出,目前各数据库可支持的输入数据类型也逐渐多样化,由最初的仅支持单个菌株的基因组进行耐药基因预测 (如 ARDB 和 CARD 等),到后来可以支持宏基因组中耐药基因的预测 (如 FARME 和 SARG v2 等)。除此之外, Resfinder、DeepARG 和 ARG-OAP (v2) 能够支持 Paired-end reads 作为输入数据,其中 Resfinder 还支持 Single end reads 作为输入数据,以便与细菌基因组中已知的获得性耐药基因进行比较。

输入数据的多元化满足了用户从基因到宏基因组水平预测耐药基因的不同需求,使得我们能够挖掘细菌耐药与细菌群落之间更深层次联系。

4 数据库评价

为了方便用户对各耐药数据库有个直观的认识,我们对所列数据库按照以下6条标准进行评分。

①输入数据:以“metagenome, genome/gene, reads”

为准,每一个+1。②支持根据功能域进行预测:本地化+2分,在线+1分,不支持或未知+0分。③数据库更新频率:以“月,年,不更新”为准,月+2分,年+1分,不更新或未知+0分。④引用情况:≥1 000+4分,[500, 1 000)+3分,[100, 500)+2分,[10, 100)+1分,<10次或未知+0分。⑤是否包含突变信息:是+1分,否+0分。⑥数据来源:整合之前数据库并进行扩展+3分,文献搜索整理+2分,整合之前数据库+1分,未知+0分。

根据以上标准得到每个数据库的分数为1-14分不等 (表 1, 表 2),其中 CARD 数据库的评分最高,为 14 分,该数据库更新较为频繁,为每个月更新一次,保证了数据的时效性。Resfinder 数据库紧随其后,评分为 11 分,而且 Resfinder 的引用次数 (截至 2020 年 4 月 20 日)高于 CARD 数据库 (分别为 1 736 次和 1 393 次),可能是由于 Resfinder 数据库不仅支持基因组/基因数据的分析,还支持未经拼接的高通量测序数据作为其输入数据,能够满足大量环境样本研究者的需求。有少数数据库得分为 1-2 分,主要是由于其停更、从未更新或更新信息不明确等。

5 展望

随着抗生素的大量使用,越来越多的微生物产生了耐药性,细菌耐药问题对全球公共卫生产生极大威胁。细菌耐药表型的预测、耐药机制的发现以及新抗生素的研发是世界范围内广大科研工作者的研究热点,相信未来几年在该领域会有更多的科研成果推出。基于前文对耐药基因数据库的概述以及对当前耐药领域的研究热点追踪,我们从以下几方面对未来的研究进行展望。

5.1 基于组学数据的细菌耐药预测

随着耐药数据的不断积累,目前已有一些研究基于全基因组数据,利用机器学习的算法对菌株的耐药基因和耐药表型 (如最小抑菌浓度 (Minimum inhibitory concentration, MIC) 进行建

模和预测。

Erol 等利用 1 595 株结核分枝杆菌 *Mycobacterium tuberculosis* 的基因组和耐药表型数据, 采用支持向量机 (Support vector machines, SVM) 的机器学习方法, 成功预测了结核分枝杆菌的 33 个已知的耐药基因并预测了 24 个新的耐药基因 (很多涉及代谢和细胞壁形成)。该算法还提供了一份针对每种药物耐药性的基因突变列表, 并按其重要性进行了排序。这些信息可以帮助医生为结核病患者选择针对性更强的正确药物^[55]。

Mitchell 等利用 78 个已报道的肠杆菌科菌株的基因组序列及其对 12 种抗生素的表型数据, 尝试通过基因组水平的基因型预测各菌株的抗生素敏感性^[56], 研究显示机器学习的算法同传统实验方法有 90.3% 的一致性。

Nguyen 等采用支持向量机 (SVM) 的机器学习方法分别预测了非伤寒沙门氏菌 *Nontyphoidal salmonella* 和肺炎克雷伯菌 *Klebsiella pneumoniae* 对 15 种和 20 种抗生素的 MIC 值。研究发现这种方法得到的结果具有较高的准确度和较低的错误率, 而且不需要提前验证菌株包含的耐药基因和表型特征^[57-58]。此方法可应用于其他常见临床病原菌, 对研究不同病原菌的耐药机制具有重要意义。

通过机器学习的算法基于全基因组序列可以快速预测临床菌株的抗生素耐药基因以及其耐药表型和敏感性。但是否使用这种方法用于初始治疗决策, 在学术界仍然存有异议, 有学者认为未知的参数会影响细菌基因型和表型之间的联系; 同时, 抗生素敏感性的某些变化也可能是由于代谢活性因素引起的^[59], 因此如何构建一个高效科学的表型模型在未来的研究领域会是一个重要焦点。

5.2 特异型耐药数据库逐渐增多

药物特异性数据库除当前已知的 β -内酰胺类较多之外, 也有一些关于四环素、磺胺类、氟喹

诺酮、大环内酯类和氨基糖苷类的研究, 尤其在畜牧业和农业生产领域^[60-62]; 同时随着人类对抗生素滥用以及由此产生的耐药基因、耐药细菌对人类生活安全威胁的重视, 针对特定细菌如肺炎链球菌 *Streptococcus pneumoniae*^[63]、金黄色葡萄球菌 *Staphylococcus aureus*^[64] 以及乳酸杆菌 *Lactobacillus*^[65] 等微生物的耐药基因及机制的研究成果正陆续发表出来, 相信未来将会有更多的药物特异型和菌株特异型数据库出现, 以满足科研工作者和临床医生的迫切需要。

5.3 可移动元件与耐药基因

可移动元件 (如质粒、基因组岛、噬菌体、转座子以及整合子等) 是耐药基因水平转移的重要载体。当前的耐药数据库更多关注耐药基因的预测问题, 仅少量数据库涉及可移动元件的注释。CARD 数据库整理了 6 大类型共 92 个可移动元件 (包括质粒、转座元件、噬菌体、基因组岛以及 Group I 和 Group II 内含子) (表 1); RAC (Antibiotic resistance cassettes) 数据库收集了 130 个含有耐药基因的基因盒^[66] (附表 1, 见网络版附录); FARME 数据库中包含了 360 个可移动元件的 HMM 模型 (表 1)。

但是以上数据库仅针对耐药基因和可移动元件分别提供分析, 缺乏对两者的整合分析。由于耐药基因并不总在可移动元件的内部, 而且同一类耐药基因可能借助多种不同的可移动元件进行传播, 这增加了整合分析的难度。

REFERENCES

- [1] Xing JY, Zhang YJ, Zhou LL. An overview on bacteria resistance of antibiotic and disinfectants in drinking water distribution system. *Chem Enterprise Manage*, 2020(4): 69-70, 79 (in Chinese). 邢佳夷, 张永吉, 周玲玲. 抗生素存在下给水管网中细菌抗生素抗性和消毒剂抗性研究进展. *化工管理*, 2020(4): 69-70, 79.
- [2] Tackling drug-resistant infections globally: final

- report and recommendations. [EB/OL]. [2020-06-24]. https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf.
- [3] Brogan DM, Mossialos E. A critical analysis of the review on antimicrobial resistance report and the infectious disease financing facility. *Global Health*, 2016, 12: 8.
- [4] Tackling a crisis for the health and wealth of nations. [EB/OL]. [2020-06-24]. https://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf.
- [5] Zhang QQ, Ying GG, Pan CG, et al. Comprehensive evaluation of antibiotics emission and fate in the river basins of China: source analysis, multimedia modeling, and linkage to bacterial resistance. *Environ Sci Technol*, 2015, 49(11): 6772–6782.
- [6] Zhang N, Li M, Liu X. Distribution and transformation of antibiotic resistance genes in soil. *China Environ Sci*, 2018, 38(7): 2609–2617 (in Chinese).
张宁, 李森, 刘翔. 土壤中抗生素抗性基因的分布及迁移转化. *中国环境科学*, 2018, 38(7): 2609–2617.
- [7] Bu QW, Wang B, Huang J, et al. Pharmaceuticals and personal care products in the aquatic environment in China: a review. *J Hazard Mater*, 2013, 262: 189–211.
- [8] Liu JL, Wong MH. Pharmaceuticals and personal care products (PPCPs): a review on environmental contamination in China. *Environ Int*, 2013, 59: 208–224.
- [9] Xu Y, Xu J, Mao DQ, et al. Effect of the selective pressure of sub-lethal level of heavy metals on the fate and distribution of ARGs in the catchment scale. *Environ Pollut*, 2017, 220: 900–908.
- [10] Zheng J, Gao RX, Wei YY, et al. High-throughput profiling and analysis of antibiotic resistance genes in East Tiaoxi River, China. *Environ Pollut*, 2017, 230: 648–654.
- [11] Xu YB, Hou MY, Li YF, et al. Distribution of tetracycline resistance genes and AmpC β -lactamase genes in representative non-urban sewage plants and correlations with treatment processes and heavy metals. *Chemosphere*, 2017, 170: 274–281.
- [12] Yang FX, Huang L, Li LY, et al. Discharge of KPC-2 genes from the WWTPs contributed to their enriched abundance in the receiving river. *Sci Total Environ*, 2017, 581–582: 136–143.
- [13] Chen QL, Li H, Zhou XY, et al. An underappreciated hotspot of antibiotic resistance: the groundwater near the municipal solid waste landfill. *Sci Total Environ*, 2017, 609: 966–973.
- [14] Li JN, Cheng WX, Xu LK, et al. Antibiotic-resistant genes and antibiotic-resistant bacteria in the effluent of urban residential areas, hospitals, and a municipal wastewater treatment plant system. *Environ Sci Pollut Res*, 2015, 22(6): 4587–4596.
- [15] Rodriguez-Mozaz S, Chamorro S, Marti E, et al. Occurrence of antibiotics and antibiotic resistance genes in hospital and urban wastewaters and their impact on the receiving river. *Water Res*, 2015, 69: 234–242.
- [16] Zhu YG, Zhao Y, Li B, et al. Continental-scale pollution of estuaries with antibiotic resistance genes. *Nat Microbiol*, 2017, 2: 16270.
- [17] Yang Y, Zhang T, Zhang XX, et al. Quantification and characterization of β -lactam resistance genes in 15 sewage treatment plants from East Asia and North America. *Appl Microbiol Biotechnol*, 2012, 95(5): 1351–1358.
- [18] Zhang T, Zhang M, Zhang XX, et al. Tetracycline resistance genes and tetracycline resistant lactose-fermenting *Enterobacteriaceae* in activated sludge of sewage treatment plants. *Environ Sci Technol*, 2009, 43(10): 3455–3460.
- [19] Wang FH, Qiao M, Lü ZE, et al. Impact of reclaimed water irrigation on antibiotic resistance in public parks, Beijing, China. *Environ Pollut*, 2014, 184: 247–253.
- [20] Zhou X, Qiao M, Wang FH, et al. Use of commercial organic fertilizer increases the abundance of antibiotic resistance genes and antibiotics in soil. *Environ Sci Pollut Res*, 2017,

- 24(1): 701–710.
- [21] Ben WW, Wang J, Pan X, et al. Dissemination of antibiotic resistance genes and their potential removal by on-farm treatment processes in nine swine feedlots in Shandong Province, China. *Chemosphere*, 2017, 167: 262–268.
- [22] Zhu YG, Johnson TA, Su JQ, et al. Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proc Natl Acad Sci USA*, 2013, 110(9): 3435–3440.
- [23] Pal C, Bengtsson-Palme J, Kristiansson E, et al. The structure and diversity of human, animal and environmental resistomes. *Microbiome*, 2016, 4(1): 54.
- [24] Zhao WX, Wang B, Yu G. Antibiotic resistance genes in China: occurrence, risk, and correlation among different parameters. *Environ Sci Pollut Res*, 2018, 25(22): 21467–21482.
- [25] Wang XL, Li FZ, He JL, et al. The research progress of environmental microbial antibiotic and heavy metal resistance. *Environ Sci Technol*, 2019, 32(1): 59–62 (in Chinese).
王小垒, 李凤姿, 何家乐, 等. 环境微生物抗生素与重金属抗性研究进展. *环境科技*, 2019, 32(1): 59–62.
- [26] Luo X, Zhang WL, Yuan LX, et al. Correlation between resistance genes and microbial community in polluted rivers. *China Environ Sci*, 2019, 39(6): 2606–2613 (in Chinese).
罗晓, 张文丽, 袁立霞, 等. 纳污河流抗性基因和微生物群落相关性. *中国环境科学*, 2019, 39(6): 2606–2613.
- [27] Han TF, Liu N, Zhang QQ, et al. Resistance mechanism of animal-derived bacteria to seven antibiotics and prevalence status of drug-resistant genes in China. *Chin Anim Health Inspect*, 2019, 36(11): 53–58 (in Chinese).
韩天飞, 刘娜, 张青青, 等. 七类常见抗菌药耐药机制及耐药基因国内流行现状. *中国动物检疫*, 2019, 36(11): 53–58.
- [28] Tan Y, Fang ZP. Research progress on the mechanism of antibacterial drugs and the mechanism of bacterial resistance. *World Notes Antibiot*, 2003, 24(2): 65–69 (in Chinese).
谭艳, 方治平. 抗菌药物的作用机制及细菌耐药性机制的研究进展. *国外医药(抗生素分册)*, 2003, 24(2): 65–69.
- [29] Pruden A, Pei RT, Storteboom H, et al. Antibiotic resistance genes as emerging contaminants: studies in northern Colorado. *Environ Sci Technol*, 2006, 40(23): 7445–7450.
- [30] Shankar PR. Antimicrobial resistance: global report on surveillance. *Australasian Med J*, 2014, 7(5): 238–239.
- [31] 耐药国家行动计划(2016-2020年) [EB/OL]. [2010-06-25]. http://www.gov.cn/xinwen/2016-08/25/content_5102348.htm.
- [32] Didelot X, Bowden R, Wilson DJ, et al. Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet*, 2012, 13(9): 601–612.
- [33] Liu B, Pop M. ARDB-antibiotic resistance genes database. *Nucleic Acids Res*, 2009, 37(Suppl.1): D443–D447.
- [34] Bilgin N, Richter AA, Ehrenberg M, et al. Ribosomal RNA and protein mutants resistant to spectinomycin. *EMBO J*, 1990, 9(3): 735–739.
- [35] Ruiz J, Moreno A, De Anta MTJ, et al. A double mutation in the *gyrA* gene is necessary to produce high levels of resistance to moxifloxacin in *Campylobacter* spp. clinical isolates. *Int J Antimicrob Agents*, 2005, 25(6): 542–545.
- [36] Zankari E, Hasman H, Cosentino S, et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother*, 2012, 67(11): 2640–2644.
- [37] Zankari E, Allesøe R, Allesøe KG, et al. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. *J Antimicrob Chemother*, 2017, 72(10): 2764–2768.
- [38] Alcock BP, Raphenya AR, Lau TTY, et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database.

- Nucleic Acids Res, 2020, 48(D1): D517–D525.
- [39] Gupta SK, Padmanabhan BR, Diene SM, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother*, 2014, 58(1): 212–220.
- [40] Gibson MK, Forsberg KJ, Dantas G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J*, 2015, 9(1): 207–216.
- [41] Wallace JC, Port JA, Smith MN, et al. FARME DB: a functional antibiotic resistance element database. *Database (Oxford)*, 2017, 2017: baw165.
- [42] Arango-Argoty G, Garner E, Pruden A, et al. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome*, 2018, 6(1): 23.
- [43] Yin XL, Jiang XT, Chai BL, et al. ARGs-OAP v2.0 with an expanded SARG database and hidden Markov models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. *Bioinformatics*, 2018, 34(13): 2263–2270.
- [44] Scaria J, Chandramouli U, Verma SK. Antibiotic Resistance Genes Online (ARGO): a database on vancomycin and β -lactam resistance genes. *Bioinformation*, 2005, 1(1): 5–7.
- [45] Bush K, Jacoby GA. Updated functional classification of β -lactamases. *Antimicrob Agents Chemother*, 2010, 54(3): 969–976.
- [46] Naas T, Oueslati S, Bonnin RA, et al. Beta-Lactamase DataBase (BLDB)-structure and function. *J Enzyme Inhib Med Chem*, 2017, 32(1): 917–919.
- [47] Sandgren A, Strong M, Muthukrishnan P, et al. Tuberculosis drug resistance mutation database. *PLoS Med*, 2009, 6(2): e1000002.
- [48] Flandrois JP, Lina G, Dumitrescu O. MUB II -TB-DB: a database of mutations associated with antibiotic resistance in *Mycobacterium tuberculosis*. *BMC Bioinformatics*, 2014, 15: 107.
- [49] Saha SB, Uttam V, Verma V. u-CARE: user-friendly comprehensive antibiotic resistance repository of *Escherichia coli*. *J Clin Pathol*, 2015, 68(8): 648–651.
- [50] Cox G, Wright GD. Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions. *Int J Med Microbiol*, 2013, 303(6/7): 287–292.
- [51] Boolchandani M, D'Souza AW, Dantas G. Sequencing-based methods and resources to study antimicrobial resistance. *Nat Rev Genet*, 2019, 20(6): 356–370.
- [52] Pal C, Bengtsson-Palme J, Rensing C, et al. BacMet: antibacterial biocide and metal resistance genes database. *Nucleic Acids Res*, 2014, 42(D1): D737–D743.
- [53] Chen S, Zou HY, Gao FZ, et al. Co-selection mechanism of antibiotic, metal and biocide resistance. *Asian J Ecotoxicol*, 2020, 15(2): 1–10 (in Chinese).
陈帅, 邹海燕, 高方舟, 等. 抗生素、重金属和杀生剂抗性共选择机制. *生态毒理学报*, 2020, 15(2): 1–10.
- [54] Lakin SM, Dean C, Noyes NR, et al. MEGARes: an antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Res*, 2017, 45(D1): D574–D580.
- [55] Kavvas ES, Catoi E, Mih N, et al. Machine learning and structural analysis of *Mycobacterium tuberculosis* pan-genome identifies genetic signatures of antibiotic resistance. *Nat Commun*, 2018, 9(1): 4306.
- [56] Pesesky MW, Hussain T, Wallace M, et al. Evaluation of machine learning and rules-based approaches for predicting antimicrobial resistance profiles in gram-negative bacilli from whole genome sequence data. *Front Microbiol*, 2016, 7: 1887.
- [57] Nguyen M, Long SW, McDermott PF, et al. Using machine learning to predict antimicrobial minimum inhibitory concentrations and associated genomic features for nontyphoidal *Salmonella*. *J Clin Microbiol*, 2019, 57(2): e01260–18.
- [58] Nguyen M, Brettin T, Long SW, et al. Developing an

- in silico* minimum inhibitory concentration panel test for *Klebsiella pneumoniae*. Sci Rep, 2018, 8: 421.
- [59] Kaldalu N, Hauryliuk V, Tenson T. Persisters-as elusive as ever. Appl Microbiol Biotechnol, 2016, 100(15): 6545–6553.
- [60] Xie WY, Yang XP, Li Q, et al. Changes in antibiotic concentrations and antibiotic resistome during commercial composting of animal manures. Environ Pollut, 2016, 219: 182–190.
- [61] Tian Z, Zhang Y, Yu B, et al. Changes of resistome, mobilome and potential hosts of antibiotic resistance genes during the transformation of anaerobic digestion from mesophilic to thermophilic. Water Res, 2016, 98: 261–269.
- [62] Wichmann F, Udikovic-Kolic N, Andrew S, et al. Diverse antibiotic resistance genes in dairy cow manure. mBio, 2014, 5(2): e01017-13.
- [63] Aliberti S, Cook GS, Babu BL, et al. International prevalence and risk factors evaluation for drug-resistant *Streptococcus pneumoniae* pneumonia. J Infect, 2019, 79(4): 300–311.
- [64] Vestergaard M, Leng BF, Haaber J, et al. Genome-wide identification of antimicrobial intrinsic resistance determinants in *Staphylococcus aureus*. Front Microbiol, 2016, 7: 2018.
- [65] Campedelli I, Mathur H, Salvetti E, et al. Genus-Wide Assessment of Antibiotic Resistance in *Lactobacillus* spp. Appl Environ Microbiol, 2019, 85(1): e01738–18.
- [66] Tsafnat G, Copty J, Partridge SR. RAC: repository of antibiotic resistance cassettes. Database (Oxford), 2011, 2011: bar054.

(本文责编 陈宏宇)