



基于非核糖体肽的文献计量分析

刘易 谭娅文 江辉*

浙江大学生命科学院 浙江 杭州 310012

摘要: 许多临床上的重要抗生素来源于微生物生产的非核糖体肽类天然产物或者聚酮-非核糖体肽杂合体类天然产物, 本文选取了近 5 年 Web of Science 上关于非核糖体肽的国际期刊文献, 采用文献计量、统计分析等方法展示非核糖体肽研究领域的热点方向, 探究了该领域的发展趋势, 以为为进一步研究提供参考。

关键词: 非核糖体肽, 文献计量学, Web of Science

Bibliometric analysis based on nonribosomal peptides

LIU Yi TAN Yawen JIANG Hui*

College of Life Sciences, Zhejiang University, Hangzhou, Zhejiang 310012, China

Abstract: Many clinical medicines are derived from the microbial natural products of nonribosomal peptides or polyketide/nonribosomal peptide hybrids. The international literatures about nonribosomal peptides on Web of Science in the past five years were analyzed by bibliometric and statistical methods. Our results here show the hot topics and the development trend on the research field about nonribosomal peptides.

Keywords: nonribosomal peptide, bibliometric, Web of Science

临床药物有 1/2 来源于天然产物或者天然产物衍生物, 而天然产物又有 2/3 是由微生物产生的^[1]。非核糖体肽和聚酮化合物是天然产物的两大家族, 许多微生物产生的非核糖体肽已经在临床上得到了广泛的应用, 如青霉素、头孢霉素、万古霉素、达托霉素、卡泊芬净、米卡芬净、阿尼芬净、环孢菌素 A、博来霉素、他克莫司等, 因此微生物产生的非核糖体肽具有重要的研究价值^[2]。

国内外每年都有大量关于微生物产生的非核糖体肽的论文, 但论文质量、研究内容和研究方向各有差异, 给该领域的学者带来了困扰。为了对非核糖体肽的发展趋势以及当前热点有更明确的认识和了解, 我们选择文献收录范围较广的 Web of Science 为研究对象, 以“Nonribosomal Peptide”为关键词进行检索, 并对相应检索文献进行分析和统计, 概括目前该领域较前沿的发展方向, 以为为进一步研究提供参考。

Foundation items: National Natural Science Foundation of China (31870005); State Key Laboratory of Microbial Metabolism, Shanghai Jiao Tong University for Open Projects (MMLKF17-06)

*Corresponding author: Tel: 86-571-88206995; E-mail: jianghuisioc@zju.edu.cn

Received: 08-02-2021; Accepted: 27-03-2021; Published online: 27-04-2021

基金项目: 国家自然科学基金(31870005); 上海交通大学微生物代谢国家重点实验室开放课题(MMLKF17-06)

*通信作者: Tel: 0571-88206995; E-mail: jianghuisioc@zju.edu.cn

收稿日期: 2021-02-08; 接受日期: 2021-03-27; 网络首发日期: 2021-04-27

1 数据来源与方法

1.1 数据来源

以 Web of Science 核心合集为检索数据库, 分别以“Nonribosomal Peptide or Non-Ribosomal Peptide”和“Nonribosomal Peptide”为主题词进行预检, 检索发现以“Nonribosomal Peptide”为主题词所得结果更为准确, 因此选取“Nonribosomal Peptide”为主题词, 检索方式为普通检索, 检索时间限定为 1990–2021 年。使用数据库中的“文献类型精炼”选项对检索结果进行提炼, 纳入“Article”“Review”类文献, 排除了会议摘要、评论、论文集论文、来信等不符合纳入标准的文章类型, 再以“全记录与引用的参考文献”及“纯文本”格式导出检索内容进行一般结果分析。另外, 使用“出版年精炼”选项在以上结果里选取近 5 年数据进行期刊分布、作者共现和关键词共现分析。

1.2 研究方法

VOSviewer 是一款通过可视化方法展现一定时期内某学科或领域的研究现状, 并能预测未来发展趋势的文献计量分析软件。本文运用 VOSviewer 软件对所检索文献进行作者和关键词共现分析, 将目标文献纯文本格式导入其中, 并对数据进行清洗、筛选绘制可视化图谱。不同颜色节点表示不同聚类, 节点大小表示出现的频次, 节点的距离则表示主题词之间的联系程度。此外, 借助 Web of Science 数据库提供的检索结果分

析功能, 对所检索文献的发表时间、期刊分布等数据信息进行统计分析。

2 结论

2.1 一般结果

Web of Science 核心合集中 1990–2021 年间发表含“Nonribosomal Peptide”主题词的文献为 2 817 篇。文献发表量统计结果显示, Web of Science 上发表文章数目从 1995 年开始总体趋势显示为逐年增多, 2000 年后增速较快, 其中 2018 年发表数目最多, 为 225 篇(图 1)。非核糖体多肽领域近期相关热点论文主要集中在微生物学领域, 研究方向主要集中在挖掘新结构非核糖体肽、解析非核糖体肽合成途径, 以及对非核糖体肽合成酶的重新设计等方面。

2.2 期刊分布

在检索的近 5 年关于非核糖体肽的文章中, 既包括 *Nature*、*Science*、*PNAS* 等自然科学领域顶级期刊, 也包括微生物学领域的专科期刊 *Frontiers in Microbiology*, 其中发表论文数目最多的是 *ACS Chemical Biology*, 有 58 篇, 占论文总量的 5.390%, 发表 10 篇以上的期刊分布见表 1。

2.3 关键词共现网络

将所选文献导入 VOSviewer 软件, 统计题目和摘要中的所有关键词, 采用二进制计数(Binary Counting)的方式, 筛选出现 10 次以上的词汇共 466 个, 另外去掉重复和无实际意义的

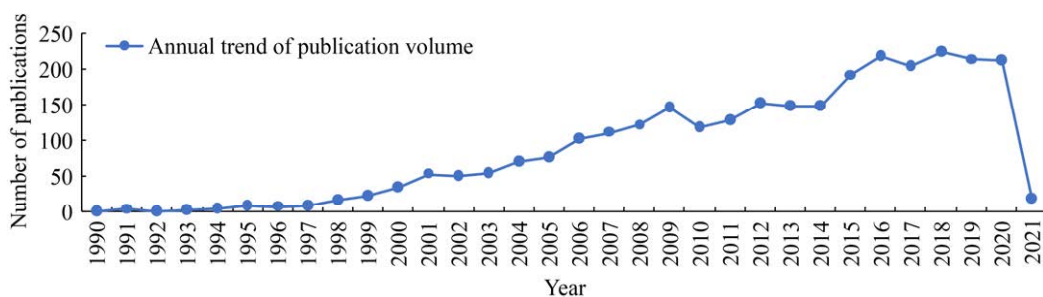


图 1 非核糖体肽研究领域文献发表量统计结果

Figure 1 Statistics of publications in the field of nonribosomal peptides research

表 1 非核糖体多肽领域论文期刊分布

Table 1 Distribution of journals in the field of nonribosomal peptides

排名 Rank	期刊 Journal	论文数量 Number of papers	占比 Percentage (%)
1	<i>ACS Chemical Biology</i>	58	5.390
2	<i>Chembiochem</i>	33	3.067
3	<i>Natural Product Reports</i>	33	3.067
4	<i>Angewandte Chemie-International Edition</i>	32	2.974
5	<i>Biochemistry</i>	28	2.602
6	<i>Frontiers in Microbiology</i>	28	2.602
7	<i>JACS</i>	26	2.416
8	<i>Scientific Reports</i>	25	2.323
9	<i>Journal of Natural Products</i>	24	2.230
10	<i>Organic Letters</i>	24	2.230
11	<i>Nature Chemical Biology</i>	23	2.138
12	<i>PNAS</i>	23	2.138
13	<i>Applied and Environmental Microbiology</i>	19	1.766
14	<i>Chemical Science</i>	19	1.766
15	<i>ACS Synthetic Biology</i>	16	1.487
16	<i>Journal of Biological Chemistry</i>	15	1.394
17	<i>BMC Genomics</i>	14	1.301
18	<i>Journal of Industrial Microbiology & Biotechnology</i>	14	1.301
19	<i>Marine Drugs</i>	13	1.208
20	<i>PLoS One</i>	13	1.208
21	<i>Environmental Microbiology</i>	12	1.115
22	<i>Nature Communications</i>	12	1.115
23	<i>Applied Microbiology and Biotechnology</i>	11	1.022
24	<i>Cell Chemical Biology</i>	11	1.022

词, 我们最终选择 78 个关键词进行共现网络分析(图 2)。图 2 中右下角有一颜色从深到浅排列的刻度线表示 2017 年 6 月以前出现的关键词, 在图中显示为蓝色, 而 2018 年 4 月以后出现的关键词显示为黄色。对图 2 中出现的关键词进行归纳分析发现, 最常出现(节点最大)的几个词, 如 Nonribosomal Peptide Synthetase、Domian、Gene Cluster 等均是围绕非核糖体肽合成酶展开的; 而 Natural Product 可具体到铁载体, 其与 Biosynthesis 又都出现在铁载体聚类网络中; 另外, Environment、Bacteria、Metabolite 等词则是微生物的相关内容, 同时, 代表微生物的节点是黄色, 表明其也是近年来的研究热点。综上, 我们选择非核糖体肽合成酶、铁载体、微生物(图 2 中以红色箭头标出)这 3 个关键词展开具体讨论。

2.3.1 非核糖体肽合成酶

非核糖体肽由非核糖体肽合成酶(NRPS)合成, NRPS 是在细菌和真菌中进化出的分子量巨大的多功能酶, 可合成高度复杂且具有生物活性的非核糖体肽次级代谢产物。一个 NRPS 由若干个模块(Module)组成, 每个模块由数个功能域(Domain)组成, 其中 3 个不可或缺的功能域是缩合功能域、腺苷化功能域和肽酰基载体蛋白功能域(也称硫醇化功能域), 有些模块中还含有其他功能域, 如差向异构功能域、甲基化功能域、硫酯酶功能域^[3]。

过去的几十年里, 大量的生化和结构生物学研究使人们对 NRPS 各个功能域, 以及功能域之间的相互作用方式有了较清晰的认识。随着 DNA 测序技术的迅速发展以及生物信息学预测能力的

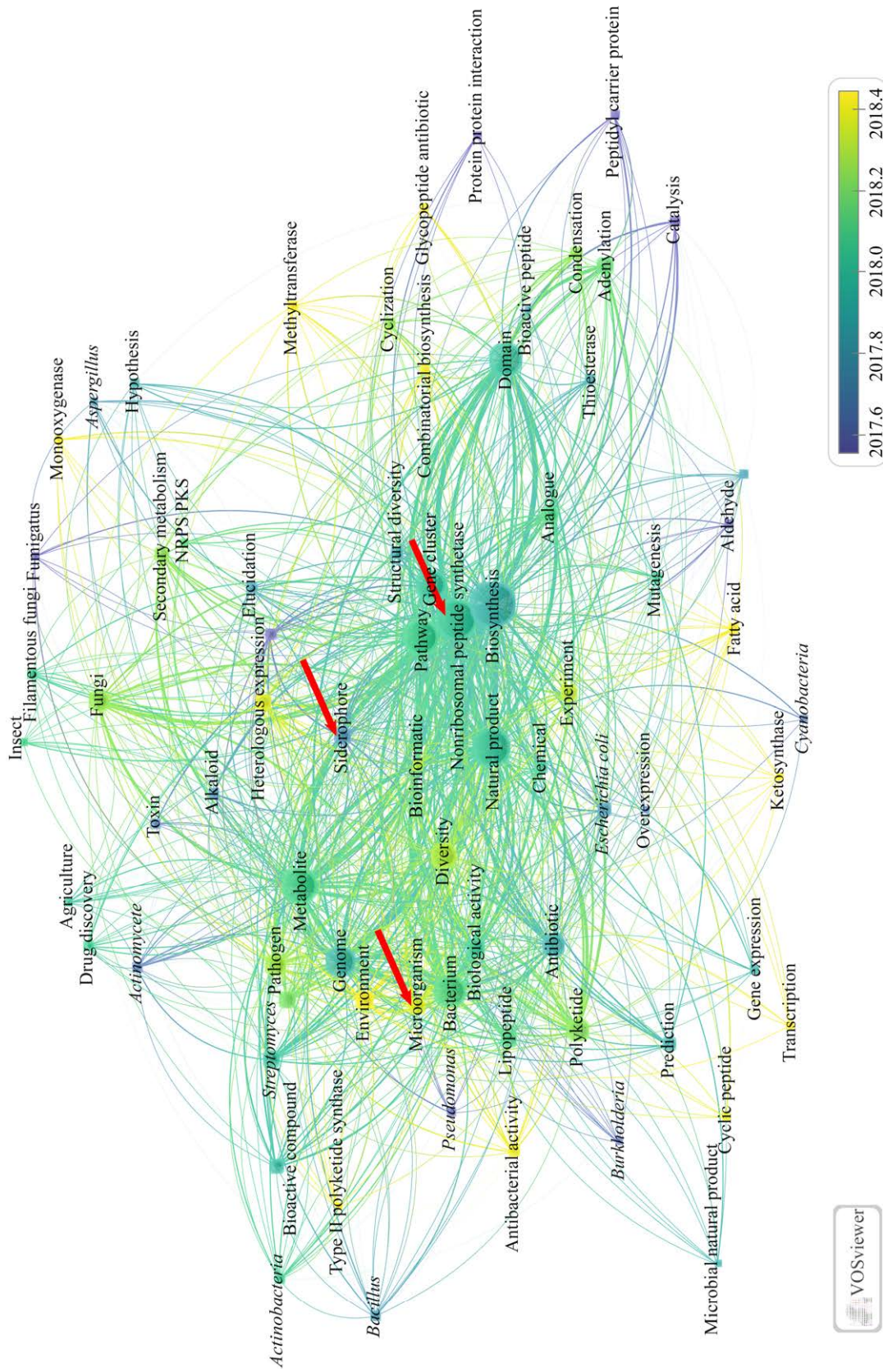


图 2 关键词共引关系网络图谱
Figure 2 Co-occurrence network of the key words of nonribosomal peptides research

注：箭头所指为本文介绍的关键词

Note: The arrow indicates the keywords introduced in this article

不断提高,许多新非核糖体肽抗生素被发现,如 Chu 等利用宏基因组测序—生物信息学预测—化学合成方法鉴定了抗生素 Humimycin,其对金黄色葡萄球菌、链球菌生长都有明显的抑制作用^[4]。

2.3.2 铁载体

铁载体是微生物或植物合成并分泌的一种用于摄取铁元素的小分子化合物。目前报道的铁载体生物合成途径有2种,最常见的是由 NRPS 负责合成,如最早在大肠杆菌中发现的含有儿茶酚结构的铁载体肠杆菌素和红色糖多孢菌中的氧脲酸盐类铁载体 Erythrochelin^[5];而另一种则不依赖于 NRPS 的铁载体合成酶合成,如天蓝色链霉菌 M145 中发现的去铁胺和金黄色葡萄球菌中发现的 Staphyloferrin B^[6]。

人们对于铁载体的结构、转运机制、生物合成途径等方面都进行了深入细致的研究,其在药物开发领域的巨大潜力也逐渐被关注。例如,抗生素-铁载体耦合物可以通过“特洛伊木马程序”进入细菌细胞内,从而达到抑制细菌增殖或直接杀死病原菌的效果^[7]。目前,具有铁载体功能的新型抗生素 Fetroja 已经用于临床,可有效治疗耐药革兰氏阴性菌感染^[8]。在抗生素滥用引起的细菌耐药性增强和新型抗生素发掘困难的背景下,研究铁载体对细菌耐药性的影响以及优化设计铁载体偶联抗生素无疑会受到更多人的关注。

2.3.3 微生物

微生物天然产物中有 2/3 是由放线菌纲产生的,70%的放线菌天然产物由链霉菌属产生^[9]。目前大部分研究集中在链霉菌上,菌株来源的单一性导致发掘出的天然产物分子结构具有类似性,近年来人们将研究对象扩大到非链霉菌属的微生物。

红球菌与链霉菌同属于放线菌,但链霉菌基因组 GC 含量往往在 70%以上,其次级代谢产物合成基因簇大多与萜烯化合物、聚酮化合物生物合成有关^[10],而红球菌基因组 GC 含量往往低于 70%,基因组中 NRPS 基因簇更多^[11,17],暗示红球菌的天然产物可能与链霉菌不尽相同。目前,红

球菌来源的天然产物数目非常之少,而且大部分分子结构都较简单,如丁二酰海藻糖脂^[12]、生物絮凝剂^[13]、生物表面活性剂^[14]、类胡萝卜素^[15]、铁载体^[16]等。McLeod 等发现红球菌 RHA1 的基因组大小(9.7 Mb)、蛋白数量(9 145 个)、次级代谢合成基因簇数量(>31 个)都明显大于天蓝色链霉菌(基因组 8.7 Mb,蛋白数量 7 769 个,次级代谢合成基因簇 22 个),其中包括 24 个 NRPS 基因簇和 7 个聚酮合酶(PKS)基因簇^[17],表明红球菌可能是天然产物尤其是非核糖体肽的潜在重要来源,值得关注与进一步研究。

2.4 作者共现网络

将所选文献导入 VOSviewer 软件统计文章中的关键词,采用 Full Counting 的方式进行计数,将发表文章数量达 5 篇以上的作者绘制成了作者共现标签视图(图 3)。发文量排名前 5 的分别为: Cryle Max J (24 篇),其工作内容主要集中在分析 NRPS 各模块之间的相互作用,并与 Süssmuth Roderich D.合作共同解析了糖肽类抗生素生物合成过程中缩合结构域的选择性^[18-19]; Komaki Hisayuki (19 篇),其研究内容大多是通过基因组测序挖掘链霉菌中 NRPS、PKS 基因簇; Hertweck Christian (16 篇)的研究方向则是解析某些非核糖体肽类天然产物生物合成过程;唐奕(15 篇)与美国斯科利普斯研究所沈奔、中国科学院上海有机所的唐功利、刘文课题组合作,主要解析真菌天然产物的生物合成途径与酶学机制研究,利用生物合成理念、微生物基因组数据,筛选、发掘新颖结构的天然产物; Bode Helge B. (13 篇)通过建立更高精度、高效率的异源表达系统激活沉默基因簇,发现和鉴定以前未表征的次级代谢产物。通过图谱可以看出在非核糖体肽领域,形成了以上述作者为核心的多个合作子网络,子网络内各学者之间有着较强的学术联系,分析图谱有助于发现某领域的研究团体,拓展与自身研究方向不同的其他研究人员交流与合作。

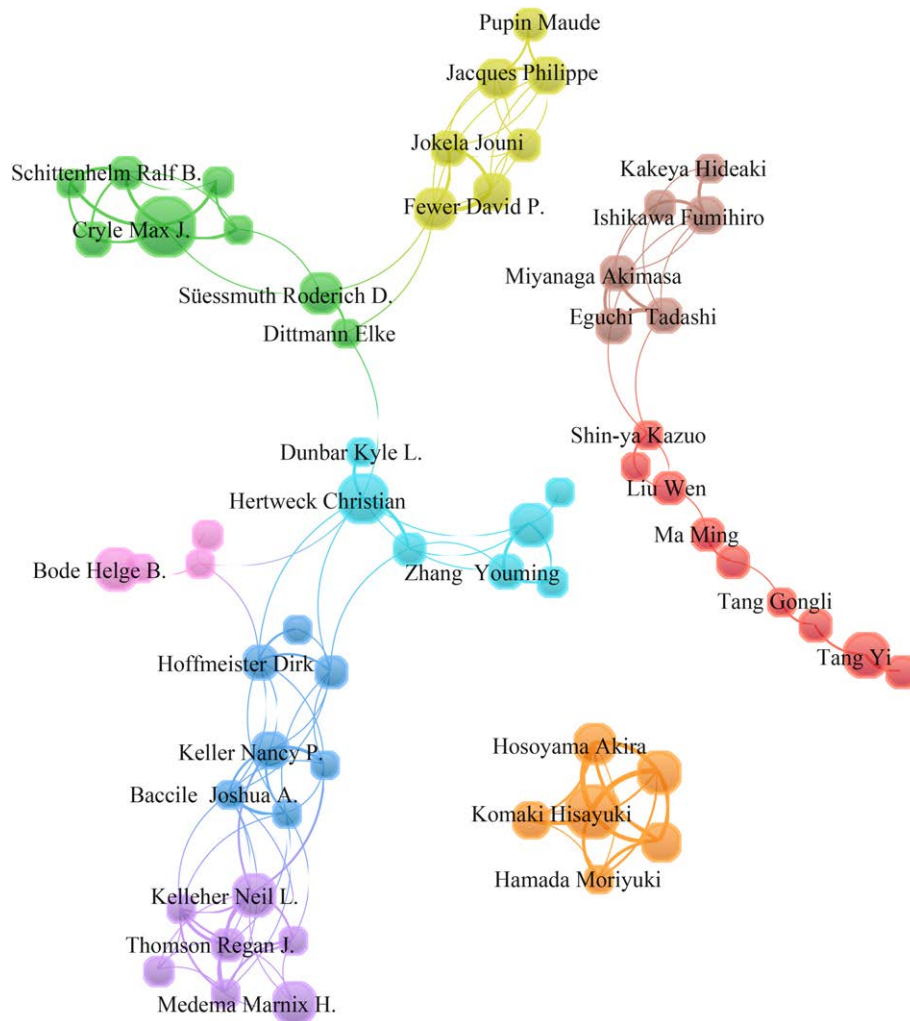


图3 非核糖体肽研究的作者共现网络

Figure 3 Co-occurrence network of authors of nonribosomal peptides research

3 总结

文献计量学是指用数学和统计学的方法,定量分析一切知识载体的交叉科学。其是集数学、统计学、文献学为一体、注重量化的综合性知识体系。本文通过文献计量学分析 Web of Science 核心合集中收录的非核糖体肽相关文献,展示了该领域发表文章的期刊分布、研究热点,以及在此领域发表论文相对较多的作者,基于此也可发现虽然前人对非核糖体多肽进行了大量工作,但还有很多问题和难点亟需解决,如何基于生物信息学基础更好地发掘出新型非核糖体肽类抗生素,优化构效关系后获得

多种活性较高的衍生物,使其具有新药开发潜力是目前仍要密切关注的热点。另外,国内有关于非核糖体肽的研究较少,科研投入仍有待加强,同时还需要加大与其他国家科研机构合作,以提高我国在该领域的国际影响力。

REFERENCES

- [1] Li JWH, Vederas JC. Drug discovery and natural products: end of an era or an endless frontier?[J]. *Science*, 2009, 325(5937): 161-165
- [2] Wei XY, Liu X, Yu HW, Guo RF. Advances in biosynthetic gene cluster mining and biosynthesis of microbial nonribosomal peptides[J]. *Chinese Journal of Antibiotics*, 2020, 45(6): 1-9 (in Chinese)

- 魏小雅, 刘欣, 于宏伟, 郭润芳. 微生物非核糖体多肽的合成基因簇挖掘及其合成研究进展[J]. 中国抗生素杂志, 2020, 45(6): 1-9
- [3] Marahiel MA. A structural model for multimodular NRPS assembly lines[J]. *Natural Product Reports*, 2016, 33(2): 136-140
- [4] Chu J, Vila-Farres X, Inoyama D, Ternei M, Cohen LJ, Gordon EA, Reddy BVB, Charlop-Powers Z, Zebroski HA, Gallardo-Macias R, et al. Discovery of MRSA active antibiotics using primary sequence from the human microbiome[J]. *Nature Chemical Biology*, 2016, 12(12): 1004-1006
- [5] Robbel L, Knappe TA, Linne U, Xie XL, Marahiel MA. Erythrochelin: a hydroxamate-type siderophore predicted from the genome of *Saccharopolyspora erythraea*[J]. *FEBS Journal*, 2010, 277(3): 663-676
- [6] Cheung J, Beasley FC, Liu S, Lajoie GA, Heinrichs DE. Molecular characterization of staphyloferrin B biosynthesis in *Staphylococcus aureus*[J]. *Molecular Microbiology*, 2009, 74(3): 594-608
- [7] Huang YZ, Jiang YF, Wang HY, Wang JX, Shin MC, Byun Y, He HN, Liang YQ, Yang VC. Curb challenges of the “Trojan Horse” approach: smart strategies in achieving effective yet safe cell-penetrating peptide-based drug delivery[J]. *Advanced Drug Delivery Reviews*, 2013, 65(10): 1299-1315
- [8] Wunderink RG, Matsunaga Y, Ariyasu M, Clevenbergh P, Echols R, Kaye KS, Kollef M, Menon A, Pogue JM, Shorr AF, et al. Cefiderocol versus high-dose, extended-infusion meropenem for the treatment of Gram-negative nosocomial pneumonia (APEKS-NP): a randomised, double-blind, phase 3, non-inferiority trial[J]. *Lancet Infectious Diseases*, 2021, 21(2): 213-225
- [9] Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Meier-Kolthoff JP, Klenk HP, Clément C, Ouhdouch Y, Van Wezel GP. Taxonomy, physiology, and natural products of actinobacteria[J]. *Microbiology and Molecular Biology Reviews*, 2016, 80(1): 1-43
- [10] Bentley SD, Chater KF, Cerdeño-Tárraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)[J]. *Nature*, 2002, 417(6885): 141-147
- [11] Bell KS, Philp JC, Aw DW, Christofi N. The genus *Rhodococcus*[J]. *Journal of Applied Microbiology*, 1998, 85(2): 195-210
- [12] Inaba T, Tokumoto Y, Miyazaki Y, Inoue N, Maseda H, Nakajima-Kambe T, Uchiyama H, Nomura N. Analysis of genes for succinoyl trehalose lipid production and increasing production in *Rhodococcus* sp. strain SD-74[J]. *Applied and Environmental Microbiology*, 2013, 79(22): 7082-7090
- [13] Peng LY, Yang CP, Zeng GM, Wang L, Dai CH, Long ZY, Liu HY, Zhong YY. Characterization and application of biofloculant prepared by *Rhodococcus erythropolis* using sludge and livestock wastewater as cheap culture media[J]. *Applied Microbiology and Biotechnology*, 2014, 98(15): 6847-6858
- [14] Ivshina I, Kostina L, Krivoruchko A, Kuyukina M, Peshkur T, Anderson P, Cunningham C. Removal of polycyclic aromatic hydrocarbons in soil spiked with model mixtures of petroleum hydrocarbons and heterocycles using biosurfactants from *Rhodococcus ruber* IEGM 231[J]. *Journal of Hazardous Materials*, 2016, 312: 8-17
- [15] Osawa A, Kasahara A, Mastuoka S, Gassel S, Sandmann G, Shindo K. Isolation of a novel carotenoid, OH-chlorobactene glucoside hexadecanoate, and related rare carotenoids from *Rhodococcus* sp. CIP and their antioxidative activities[J]. *Bioscience, Biotechnology, and Biochemistry*, 2011, 75(11): 2142-2147
- [16] Bosello M, Robbel L, Linne U, Xie XL, Marahiel MA. Biosynthesis of the siderophore rhodochelin requires the coordinated expression of three independent gene clusters in *Rhodococcus jostii* RHA1[J]. *Journal of the American Chemical Society*, 2011, 133(12): 4587-4595
- [17] Mcleod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, et al. The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, 103(42): 15582-15587
- [18] Kaniusaite M, Tailhades J, Marschall EA, Goode RJA, Schittenhelm RB, Cryle MJ. A proof-reading mechanism for non-proteinogenic amino acid incorporation into glycopeptide antibiotics[J]. *Chemical Science*, 2019, 10(41): 9466-9482
- [19] Schoppet M, Peschke M, Kirchberg A, Wiebach V, Süßmuth RD, Stegmann E, Cryle MJ. The biosynthetic implications of late-stage condensation domain selectivity during glycopeptide antibiotic biosynthesis[J]. *Chemical Science*, 2019, 10(1): 118-133