



极端酸性矿山环境微生物基于 CRISPR 系统的适应性免疫机制

黄珊珊¹, 马丽媛^{1,2*}, 王红梅^{1,3}, 刘学端²

¹中国地质大学(武汉)环境学院, 湖北 武汉 430074

²中南大学生物冶金教育部重点实验室, 湖南 长沙 410083

³中国地质大学(武汉)生物地质与环境地质国家重点实验室, 湖北 武汉 430074

摘要:【目的】通过对酸性矿山环境中嗜酸硫杆菌属(*Acidithiobacillus*)、脱硫弧菌属(*Desulfovibrio*)、钩端螺旋菌属(*Leptospirillum*)、硫化杆菌属(*Sulfobacillus*)、酸原体属(*Acidiplasma*)和铁质菌属(*Ferroplasma*)的 100 株冶金微生物基因组中 CRISPR-Cas 系统的结构特征和同源关系进行生物信息学分析, 在基因组水平上解析冶金微生物基于 CRISPR 系统对极端环境的适应性免疫机制。【方法】从 NCBI 网站下载基因组序列, 采用 CRISPR Finder 定位基因组中潜在的 CRISPR 簇。分析 CRISPR 系统的组成结构与功能: 利用 Clustal Omega 对重复序列(repeat)分类; 将间隔序列(spacer)分别与 nr 数据库、质粒数据库和病毒数据库比对, 获得注释信息; 根据 Cas 蛋白的种类和同源性对酸性矿山环境微生物的 CRISPR-Cas 系统分型。【结果】在 100 株冶金微生物基因组中共鉴定出 415 个 CRISPR 簇, 在 176 个 cCRISPR 簇中共有 80 种不同的重复序列和 4147 条间隔序列。对重复序列分类, 发现 12 类重复序列均能形成典型的 RNA 二级结构, Cluster 10 中的重复序列在冶金微生物中最具有代表性。间隔序列注释结果表明, 这些微生物曾遭受来自细菌质粒与病毒的攻击, 并通过不同的防御机制抵抗外源核酸序列的入侵。冶金微生物细菌的大部分 CRISPR-Cas 系统属于 I-C 和 I-E 亚类型, 而古菌的 CRISPR-Cas 系统多为 I-D 亚类型, 两者基于 CRISPR-Cas 系统的进化过程中存在显著差异。【结论】酸性矿山环境微生物的 CRISPR 结构可能采用不同免疫机制介导外源核酸序列与 Cas 蛋白的相互作用, 为进一步揭示极端环境微生物的适应性进化机理奠定了基础。

关键词: 酸性矿山环境, CRISPR-Cas 系统, 重复序列, 间隔序列, 适应机制

基金项目: 国家重点研发计划重点专项(2019YFC1803604); 生物冶金教育部重点实验室开放基金(MOEKLB02); 阳泉市水利局省级水利发展资金项目(2019046477); 中国地质大学(武汉)实验技术研究项目(SJ-201903)

*通信作者。E-mail: maly@cug.edu.cn

收稿日期: 2020-01-21; 修回日期: 2020-03-19; 网络出版日期: 2020-04-18

酸性矿山排水(acid mine drainage, AMD)是硫化矿物在表生条件下, 与水、大气和微生物等相互作用而形成的典型陆地特殊水生环境^[1]。酸性矿山环境的典型特征是 pH 值低至 3 以下, 铁、铜、铅、锌等重金属阳离子以及硫酸根阴离子含量超标^[2], 对矿山周边生态环境造成严重污染, 导致生物多样性急剧降低, 给人类健康带来危害^[3]。

在这种寡营养极端环境中, 仍有多种嗜酸性微生物存在, 并在铁、硫等元素的地球化学循环过程中扮演着重要角色。例如典型的嗜酸硫杆菌属(*Acidithiobacillus*), 该属中的嗜酸氧化亚铁硫杆菌 *A. ferrooxidans* 能够将体系中的亚铁离子、单质硫和无机硫化物(inorganic sulfur compounds, ISCs)快速氧化并获取能量, 有关其铁硫代谢系统和氧化金属硫化矿物的动力学模型已被广泛研究^[4-6]。钩端螺旋菌属(*Leptospirillum*)中的嗜铁钩端螺旋菌 *L. ferrooxidans* 是 AMD 环境中通过氧化亚铁获取能量的优势菌群^[7-8], 而革兰氏阴性的脱硫弧菌属(*Desulfovibrio*)和革兰氏阳性的硫化杆菌属(*Sulfobacillus*)在硫代谢过程中发挥着重要作用。此外, 酸性矿山环境中的嗜酸古菌也被广泛报道, 如酸原体属(*Acidiplasma*)和铁质菌属(*Ferroplasma*)^[9-10]常作为群落的优势物种参与含铁硫化矿的氧化溶解过程。

成簇规律的间隔短回文重复序列(clustered regularly interspaced short palindromic repeats, CRISPR)是微生物应对入侵外源核酸序列进化形成的特殊防御结构^[11]。据估计, 这种结构普遍存在于约 45%的细菌基因组以及 90%的古菌基因组中^[12]。微生物能够利用特异性的识别方式, 靶向定位并降解入侵的质粒、噬菌体或其他核酸序列, 其防御机理与真核生物的 RNA 干扰作用(RNA

interference, RNAi)有所类似^[11]。

CRISPR 簇主要由前导序列(leader sequence)、重复序列(repeat)和间隔序列(spacer)构成, CRISPR 相关(CRISPR-associated, Cas)蛋白经常存在于其侧翼区域中, 它们共同构成 CRISPR-Cas 系统^[13]。前导序列是一段位于 CRISPR 结构 5'端上游、长度约为几百个碱基的 AT 富含区域, 可能作为 CRISPR 转录的启动子(promoter), 或作为识别位点帮助获取新的 spacer 序列^[14]。重复序列具有高度保守性, 长度一般为 23-50 bp, 在不同基因组中数量差异较大, 其中存在着短回文序列, 与 RNA 二级结构的形成有关^[15]。对于其作用机制, 当重复序列形成以“环”为主的二级结构, 转录过程中间隔序列通过单个“repeat-spacer”结构单元与外源 DNA 或 RNA 发生交互作用, 在该过程中, 重复序列可能介导外源遗传物质与 Cas 编码蛋白的偶联切割作用; 当重复序列形成以“茎”为主的二级结构, 间隔序列可能通过相邻两个重复序列互补来完成与外源 DNA 或 RNA 的相互作用^[15]。间隔序列具有高度特异性, 长度为 17-84 bp 不等, 研究表明这些插入短片段多来自于外源组分, 只有极少数是自身 DNA 复制产生的^[16]。Cas 蛋白具有核酸酶、解螺旋酶和整合酶等多种酶的活性, 能够帮助识别并降解外源遗传物质^[17], 是系统发挥作用不可或缺的重要组成部分。根据 *cas* 基因的系统发育树信息与比较基因组分析, CRISPR-Cas 系统主要分为 Type I 型(标志性蛋白-Cas3)、Type II 型(标志性蛋白-Cas9)和 Type III 型(标志性蛋白-Cas10)三种类型, 并进一步分为 I-A、I-B、I-C、I-D、I-E、I-F、II-A、II-B、III-A 和 III-B 等亚类型^[17]。

有关原核生物 CRISPR 结构的研究主要集中于大肠杆菌 (*Escherichia coli*) 和嗜热链球菌

(*Streptococcus thermophilus*)等, 如对大肠杆菌 I-E 型 CRISPR-Cas 系统免疫机制的研究有助于对其耐药性提供应对策略^[18], 对嗜热链球菌 CRISPR 结构的检测则可能为生物发酵工业中的噬菌体污染提供新的解决方案^[19]。而酸性矿山作为极端微生物研究的模式系统, 为原核生物 CRISPR 簇的研究提供了一个天然场所。大量单菌基因组数据随着高通量测序技术的快速发展而涌现, 为在分子水平上深入解析微生物对极端环境的适应机制提供了可能。本文以 AMD 环境典型菌属的 100 株冶金微生物基因组为研究对象, 通过对其 CRISPR-Cas 系统的结构与功能进行分析, 解析嗜酸微生物基于 CRISPR 系统与外源核酸序列的相互作用以及对极端环境的适应性免疫机制。

1 材料和方法

1.1 基因组的下载

从 NCBI 网站(<http://www.ncbi.nlm.nih.gov>)下载酸性矿山环境典型菌属的 100 株微生物基因组序列。其中嗜酸硫杆菌属 35 株, 脱硫弧菌属 17 株, 钩端螺旋菌属 24 株, 硫化杆菌属 16 株, 酸原体属 5 株, 铁质菌属 3 株。

1.2 CRISPR 簇的鉴定

通过 CRISPR Finder (<https://crispr.i2bc.paris-saclay.fr/Server/>)^[20]对 Complete 基因组序列和 Draft 基因组序列(排除拼接质量较差、参考价值低的基因组序列)重叠群(contig)进行 CRISPR 预测, 包括 confirmed CRISPR (cCRISPR) 簇和 questionable CRISPR (qCRISPR)簇, 筛选并获得 cCRISPR 信息(repeat 序列、spacer 序列和侧翼序列)。

1.3 重复序列的分类

采用 Clustal Omega (default transition matrix: Gonnet, gap opening penalty: 6 bits, gap extension: 1 bit)对重复序列进行多重序列比对和系统发育树构建^[21], 根据比对结果与进化距离将重复序列分类。利用 WebLogo(<http://weblogo.threepiusone.com/>)在线软件对同一类别的同向重复序列进行差异性分析^[22], 基于 RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>)完成代表性重复序列 RNA 二级结构的预测^[23]。

1.4 间隔序列的注释

构建本地 NCBI-nr 数据库, 基于 Blast 对间隔序列进行注释。在此基础上, 从 NCBI 数据库中筛选质粒和病毒序列, 构建质粒数据库(Refseq 基因组 18598 个)和病毒数据库(Refseq 基因组 12144 个), 利用 BLASTn (-gapopen 10-gapextend 2-penalty-1-reward 1-word_size 7-evalue 0.01)^[24]将间隔序列分别与上述数据库进行比对。剔除与自身匹配的情况, 筛选满足下列条件的结果: 覆盖率 ≥ 0.5 , 相似率 ≥ 0.85 , e-value ≤ 0.01 。根据公式(1)对比对结果打分, 选取得分最高的注释信息^[25]。

$$\text{得分} = 3 \times \text{覆盖率} + \text{相似率} \quad \text{公式(1)}$$

1.5 Cas 蛋白的鉴定与分类

通过 BLASTx 分别将 cCRISPR 结构上游与下游约 10000 个碱基的区域与本地 Cas 蛋白数据库进行比对, 鉴定 Cas 蛋白的种类, 对 AMD 环境微生物的 CRISPR-Cas 系统分型。

2 结果和分析

2.1 典型嗜酸菌 CRISPR 簇分布概况

在 100 株嗜酸微生物基因组中共鉴定出 415

个 CRISPR 簇(附表 1, <http://journals.im.ac.cn/html/actamicrocn/2020/9/20200916.htm>), 其中 176 个 cCRISPR 簇分布于 73 株微生物基因组中, 239 个为仅有 2 或 3 个重复序列的 qCRISPR 簇。*Acidithiobacillus* 属的 35 株菌中, 17 株菌的基因组中鉴定出 cCRISPR 簇, 而 *Acidiplasma* 属的 5 株菌均具有 cCRISPR 簇, 在其他 4 个属中 65% 以上的菌株具有 cCRISPR 结构。这表明 AMD 环境不同属种间嗜酸菌的 CRISPR 簇分布是不均匀的, *Acidiplasma* 属的菌株具有较多 cCRISPR 簇, 平均 4.8 个, 其次是 *Sulfobacillus* 属(2.9 个), 而其他属内的菌株平均具有 1–2 个 cCRISPR 结构。

在 cCRISPR 簇中共存在着 4147 条间隔序列, 长度从 15 bp 到 54 bp 不等。不同属种菌株的间隔序列数量与其 cCRISPR 簇数量不成正相关。例如, *Acidithiobacillus* 属和 *Sulfobacillus* 属分别具有 40 多个 cCRISPR 簇, 平均每个 cCRISPR 簇中含有

10 条间隔序列, 而 *Ferroplasma* 属的 4 个 cCRISPR 簇中平均每个具有超过 40 条间隔序列。在含有 cCRISPR 结构的菌株中, 超过 50% (40/73) 的菌株间隔序列低于 40 条, 有 5 株菌拥有超过 150 条间隔序列, 分别是 *Desulfovibrio termitidis* HI1、*Leptospirillum rubarum* Leptoll、*Acidiplasma aeolicum* V、*Acidiplasma aeolicum* VT 和 *Ferroplasma acidiphilum* Y。由于间隔序列主要来自于宿主获取的外源遗传物质, 推测这几株菌可能与外源核酸序列发生过比较频繁的相互作用, 并与同种其他菌株在基因组水平上具有较大的潜在进化差异^[16]。

2.2 典型嗜酸菌重复序列的分类

在 176 个 cCRISPR 簇中共发现了 80 种不同的重复序列, 可分成 33 个类别(cluster), 12 个类别(命名为 Cluster 1–12)中涵盖了 51 条重复序列(表 1), 来自不同属的 66 株菌种。其中 Cluster 10 中的重复序列来自于 21 个菌株, 在嗜酸菌中分布最

表 1. 12 个最大的 CRISPR 重复序列类别

Table 1. The top 12 largest clusters in CRISPR repeat clusters

Cluster	Repeats numbers (Spacers numbers, Species numbers ^a)	Representative repeat sequence (5'→3')	Genus included	CRISPR-Cas system subtype
Cluster 1	4 (135, 7)	GTCGTCCCCGCGTGCGGGGGAAAACC	<i>Acidithiobacillus</i>	NA
Cluster 2	4 (51, 5)	GTCTGCCCTGCACGCGCAGGGATGAAAGG	<i>Acidithiobacillus</i>	NA
Cluster 3	3 (130, 4)	GTCTTAATCCCTTATTGATCAGGGCAATGATTCCTAC	<i>Acidithiobacillus</i>	III-A, III-B
Cluster 4	7 (588, 10)	GTCGCCCCCACGCGGGGGCGTGGATTGAAAC	<i>Desulfovibrio</i>	I-C
Cluster 5	6 (559, 13)	GTTTGTAGCCTACCTATGAGGAATTGAAAC	<i>Sulfobacillus</i>	III-B
Cluster 6	3 (21, 2)	GGTCCATCCCCGCGTGTGCGGGGGAAAAC	<i>Sulfobacillus</i>	I-E
Cluster 7	3 (99, 3)	CTTCTCAATCGCTACGCGGCGATAAAC	<i>Acidithiobacillus</i> , <i>Desulfovibrio</i>	I-F
Cluster 8	4 (110, 5)	GAAACACCCCCACATGCGTGGGGAAGAC	<i>Acidithiobacillus</i> , <i>Leptospirillum</i>	I-E
Cluster 9	3 (70, 5)	GTCTCCCGGCACTAACCTGTCGGGAGTGGATTGAAAC	<i>Acidithiobacillus</i> , <i>Sulfobacillus</i>	I-C
Cluster 10	7 (1118, 21)	GTATCCCCACGTTTCGTGGGGATGAACCG	<i>Acidithiobacillus</i> , <i>Desulfovibrio</i> , <i>Leptospirillum</i>	I-E
Cluster 11	4 (815, 6)	GTAAAATAGAACCTTAATAGGATTGAAAG	<i>Acidiplasma</i> , <i>Ferroplasma</i>	I-D
Cluster 12	3 (58, 2)	GTTTCAATTCCTCATAGGTAGGCTACAAAC	<i>Sulfobacillus</i> , <i>Ferroplasma</i>	NA

^a: Number of microorganisms with repeats in this Cluster; NA: not accessible.

为广泛。12 个 Cluster 中的代表性重复序列均能形成典型的茎环状二级结构(图 1), 在系统行使免疫功能时发挥作用, 且茎(结构化区域)和环(非结构化区域)的大小和功能存在差异。结构化区域对维持 RNA 二级结构的稳定性起着关键作用, 茎部的长度和 G-C 碱基对的分布不同, 导致最小自由能 MFE 不同, MFE 值的降低意味着二级结构趋向于稳定^[15]。环形区域对于识别目标 DNA 并与相应 Cas 蛋白结合非常重要, 体现 CRISPR 结构的免疫功能^[17]。此外, 重复序列一般都具有共同的保守 3'GA AA(C/G)末端^[15], 本研究进一步证实了这一观点。

大部分 Cluster 的重复序列可以形成形状相似的保守哑铃状 RNA 二级结构(图 1), 中间是一个“茎”, 在两极各有一个“环”, “茎”的长度通常为 6–7 bp。Cluster 9 中重复序列 RNA 二级结构的“茎”主干长度为 9 bp, 即以“茎”为主的茎环结构, 推测在转录过程中, 间隔序列对入侵核酸序列的防御作用可能通过相邻的 2 个重复序列互补来实现。而 Cluster 11 中来自不同古菌属种 CRISPR 重复序列所形成的二级结构, “茎”的长度仅有 3 bp, 即以“环”为主的茎环结构, 推测间隔序列通过单个“repeat-spacer”单元与外源核酸序列发生相互作用, 重复序列可能在转录过程中发挥介导外源遗传物质与 Cas 蛋白偶联切割的作用。另外, Cluster 3、5、12 中重复序列所形成的二级结构, 在茎的一端有 2 个环, 茎与环依次形成, 其 MFE 值都较高, 不如哑铃状二级结构稳定。特别的, 在 Cluster 1、3、9 中重复序列茎部区域存在补偿性的 G-U 碱基对, 这是 RNA 二级结构的典型特征。Kunin 等发现 CRISPR 结构中重复序列能够形成比随机序列更加稳定的二级结构^[15], 稳定的二级结构有利于 CRISPR-Cas 系统发挥免疫功能。本研究中

AMD 环境的细菌往往比古菌形成更为稳定的重复序列二级结构, 表明其 CRISPR-Cas 系统与外源入侵核酸序列相互作用的免疫机制存在差异。

重复序列分类结果表明, 同属内菌株的重复序列能够分布于不同类别中, 形成不同的 RNA 二级结构, 如 *Acidithiobacillus* 和 *Sulfobacillus*。*Acidithiobacillus* 广泛存在于矿坑水、沉积物和生物膜等环境中^[26], 而 *Sulfobacillus* 在 pH 0.8–1.4, 27–32 °C 环境中约占微生物群落的 6%–8%^[27], 其基于 CRISPR-Cas 系统发挥核酸免疫作用来适应环境变化时可能具有多样性。此外, 不同属菌株的重复序列可分布于同一类别中, 表现出较高的同源性, 如 Cluster 12 中的重复序列来自细菌 *Sulfobacillus* 与古菌 *Ferroplasma*, 研究表明这两个属的菌株可能在极端条件下协同参与矿物的氧化性溶解过程^[26], 推测菌株的亲缘关系和 CRISPR 同源性不是直接对应的, 可能与 CRISPR 结构序列的水平基因转移有所关联^[28]。

2.3 典型嗜酸菌间隔序列的注释

2.3.1 nr 数据库比对结果: 通过比对, 有 538 条间隔序列在 nr 数据库匹配的物种注释信息为同种内的其他菌株, 尤其是细菌的间隔序列, 可以推测, 细菌属内的菌株之间可能存在着相互作用关系。例如, 菌株 *Leptospirillum* sp. Group II CF-1 广泛“攻击”了同属内的 14 个不同菌株, 其中大部分菌株的 CRISPR-Cas 系统结构高度相似, 推测这些微生物在抵御外源核酸序列入侵过程中实现了共演化, 且细菌可能利用间隔序列来识别其他同种属菌株, 使自身能够更好地适应寡营养极端生境^[14]。

2.3.2 质粒数据库比对结果: 共有 111 条间隔序列与质粒数据库相匹配, 且这些质粒均来自于细菌。基于细菌目的分类单元进行统计, 主要组成为: *Desulfobibrionales* (*Desulfobivrio vulgaris*

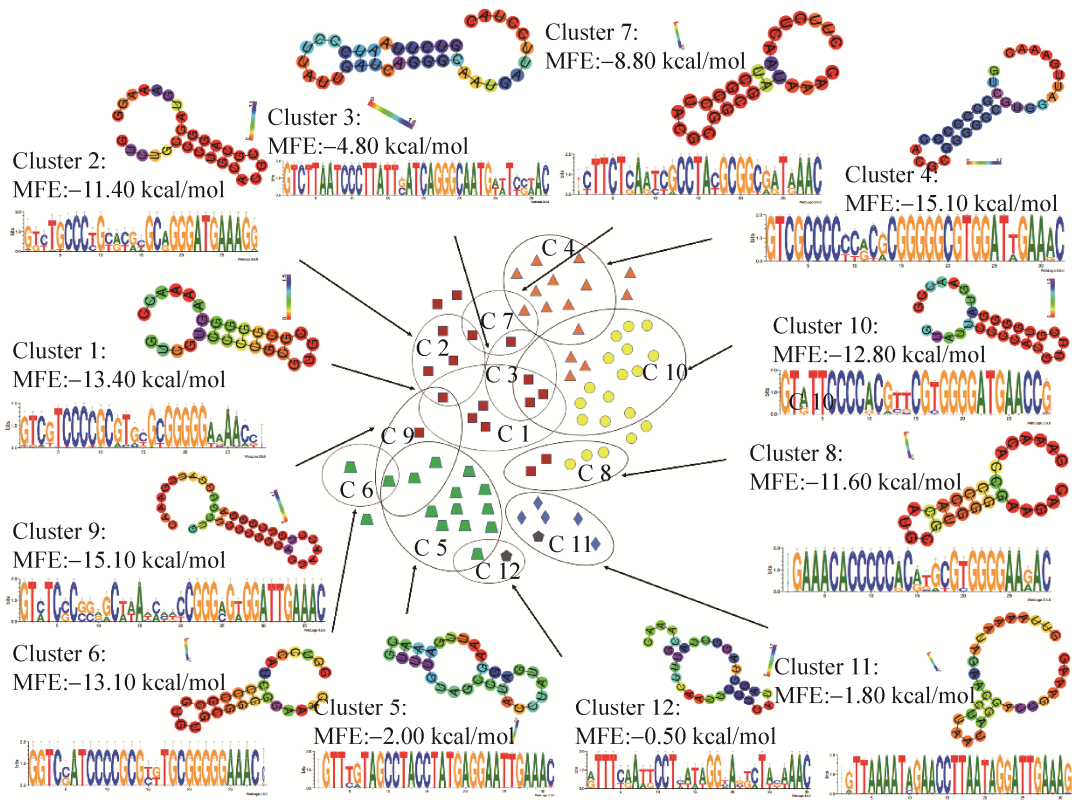


图 1. 12 个最大的 CRISPR 重复序列类别

Figure 1. The top 12 largest clusters in CRISPR repeat clusters. Dots denote individual strain, square represents *Acidithiobacillus*, triangle represents *Desulfovibrio*, round represents *Leptospirillum*, trapezium represents *Sulfobacillus*, rhombus represents *Acidiplasma*, pentagon represents *Ferroplasma*. The strains that belong to the same cluster are in the same circle, and clusters are indicated by circles together with their sequence logos and sample secondary structures where applicable.

Hildenborough), 27 条间隔序列(均来自于菌株 *Desulfovibrio vulgaris* RCH1 的一个 CRISPR 簇); *Rhizobiales*, 14 条间隔序列; *Mycoplasmatales*, 8 条间隔序列(图 2-A)。结合 nr 数据库的注释结果: 菌株 *Desulfovibrio vulgaris* Hildenborough 位于同一 CRISPR 簇的 27 条间隔序列注释信息为菌株 *Desulfovibrio vulgaris* RCH1 的质粒, 表明在进化过程中这两株菌之间可能存在相互“攻击”。以上注释结果表明, 相比古菌和真核生物质粒, AMD 环境嗜酸菌可能更容易遭受来自细菌质粒的攻击。一些细菌与这些嗜酸菌处于相似环境中, 如

分布于酸性矿山环境的 *Acidiphilium cryptum* JF-5 以及从酸性土壤分离出的 *Acidobacteriaceae bacterium* SBC82, 这为其质粒的攻击行为提供了潜在机会。此外, 一些细菌常见于农田或分离自大豆根瘤、水稻茎部, 如 *Burkholderia Gladioli* BSR3, 推测酸性矿山排水流经矿区周边农田时, 嗜酸菌会遭受某些寄生于植物细菌质粒的攻击。

2.3.3 病毒数据库比对结果: 通过计算, 91 条间隔序列的注释结果满足筛选标准。根据 ICTV (International Committee on Taxonomy of Viruses Classification), 有 85 条间隔序列的注释信息能分

类至病毒的科(图 2-B), 主要为: 38 条间隔序列注释信息为 Siphoviridae; 7 条间隔序列注释信息为 Myoviridae; 2 条间隔序列注释信息为 Podoviridae。其中 Siphoviridae 给乳制品发酵工业造成污染, 且这三类噬菌体颈部亚基的三维结构相似^[29], 对其侵染微生物可能具有重要作用。属于 Siphoviridae 的部分噬菌体分离自污水中, 如 Mycobacterium phage Adawi, 为其对酸性矿山排水微生物的侵染提供了可能。50.5%符合筛选标准的间隔序列(46/91)来自 *Acidiplasma* 和 *Ferroplasma*, 推测古菌更易与环境中的病毒发生相互作用。此外, 菌株 *Acidiplasma cupricumulans* BH2 和 *Acidiplasma cupricumulans* JCM 13668 多条位于同一 CRISPR 簇的间隔序列均具有相同的注释信息, 推测这 2 株菌应对病毒攻击的防御机制可能比较相似。

2.4 典型嗜酸菌 CRISPR 系统的基因结构与分型

Makarova 等^[17]基于 *cas* 基因的结构顺序和特点, 提出将 CRISPR-Cas 系统主要分为 Type I、Type II、Type III 三种类型以及多种亚类型。而嗜酸氧

化亚铁硫杆菌 *Acidithiobacillus ferrooxidans* ATCC23270 的 CRISPR-Cas 系统不含 Cas1、Cas2 蛋白, 没有其他分类特征被定义为 Type U^[17]。对 109 个存在 *cas* 基因的 cCRISPR 系统分型, 共有 7 种不同亚类型及 Type U (图 3)。Type I 型 CRISPR-Cas 系统广泛分布于不同属菌株基因组序列中, 少数菌株具有 Type III 型分类系统, 均不含 II 型 CRISPR-Cas 系统。

细菌属菌株 CRISPR-Cas 系统主要的亚类型属于 Subtype I-C 和 I-E, 其中 *Acidithiobacillus* 属菌株 CRISPR-Cas 系统类型最为多样, 包括 Type U 和 I-C、I-E、I-F、III-A、III-B 这些亚型。具有 I-E 型 CRISPR-Cas 系统的 5 株菌均从铜矿中分离出来, 此亚型对于其适应重金属离子的极端环境可能具有重要意义。值得注意的是, 菌株 *Acidithiobacillus ferrivorans* YL15 具有两类 CRISPR-Cas 系统, 推测其 III-B 型 CRISPR-Cas 系统靶向并降解外源 DNA 序列, 而 I-F 型 CRISPR-Cas 系统能够靶向定位并降解入侵的 DNA 和 RNA 序列^[30], 这表明菌株基于 CRISPR

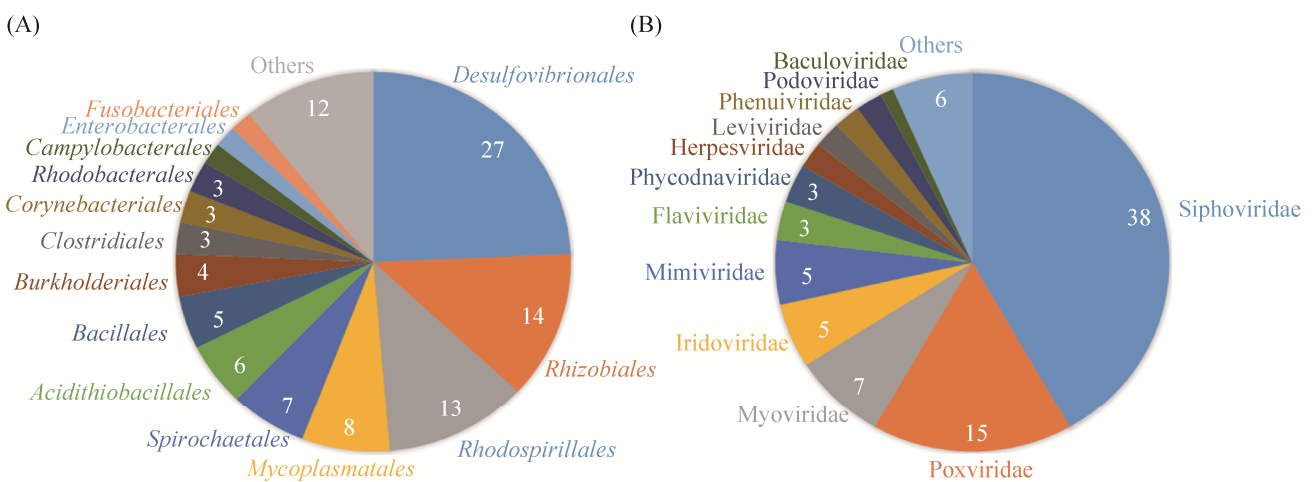


图 2. 来自于质粒(A)和病毒(B)的间隔序列物种分布图
Figure 2. The distribution of spacers from plasmid (A) and virus (B).

系统对极端环境的适应过程中进化出多种免疫机制。*Desulfovibrio* 属菌株具有一个 I-F 型 CRISPR-Cas 系统, 其余均为 I-C 和 I-E 亚类型。此属是中度嗜酸的硫酸盐还原菌(sulfate reducing bacteria, SRB), 在极低 pH 条件下, 能够代谢硫酸盐产生 H₂S 溢出体系以调节酸度保护自身^[27]。研究表明, 常用于发酵的 48 株醋酸菌大部分 CRISPR-Cas 系统为 Subtype I-C 和 I-E^[31], 推测这两种亚类型对于菌株适应酸性环境具有关键作用。同种不同菌株的 I-C 亚型 CRISPR 结构之间, *cas* 基因排列呈现多样性, 如 *Desulfovibrio vulgaris* Hildenborough 的 CRISPR-Cas 系统中 *cas* 基因排列顺序为 *cas3-cas5-cas8c-csd3-cas1* (图 3), 而 *Desulfovibrio vulgaris* RCH1 的为 *cas3-cas5-csd1-cas1*, 推测在应对不同外源核酸序列入侵过程中 Cas 编码蛋白的作用机制具有一定差异, 这

与文献报道一致^[32]。并且为 I-C 亚型的 CRISPR 簇大小均在 1800 bp 以上, 重复序列同源性较高, 属于 Cluster 4, 这表明 I-C 亚型 CRISPR-Cas 系统在本属细菌抵御外源核酸过程中发挥着重要作用。*Leptospirillum* 属菌株的 CRISPR-Cas 系统亚型只包含 I-E, 其中有 7 株菌 CRISPR-Cas 系统的重复序列与 *cas* 基因排列顺序均一致, 且都分离自美国 Iron Mountain 的里士满矿山(pH 0.3–0.7, 30–50 °C)^[27], 表明这几株嗜温细菌在适应极低 pH 环境防御入侵核酸序列时可能形成了相同的免疫体系。*Sulfobacillus* 属中存在多种类型 CRISPR-Cas 系统, 包括 I-C、I-E、III-A 和 III-B 亚型。彼此相似的 *cmr* 基因排列于菌株 *Sulfobacillus thermosulfidooxidans* CBAR-13 的 III-B 型 CRISPR-Cas 系统中(图 3), 推测可能是通过水平基因转移获取并且与上游间隔序列共同构成抵抗

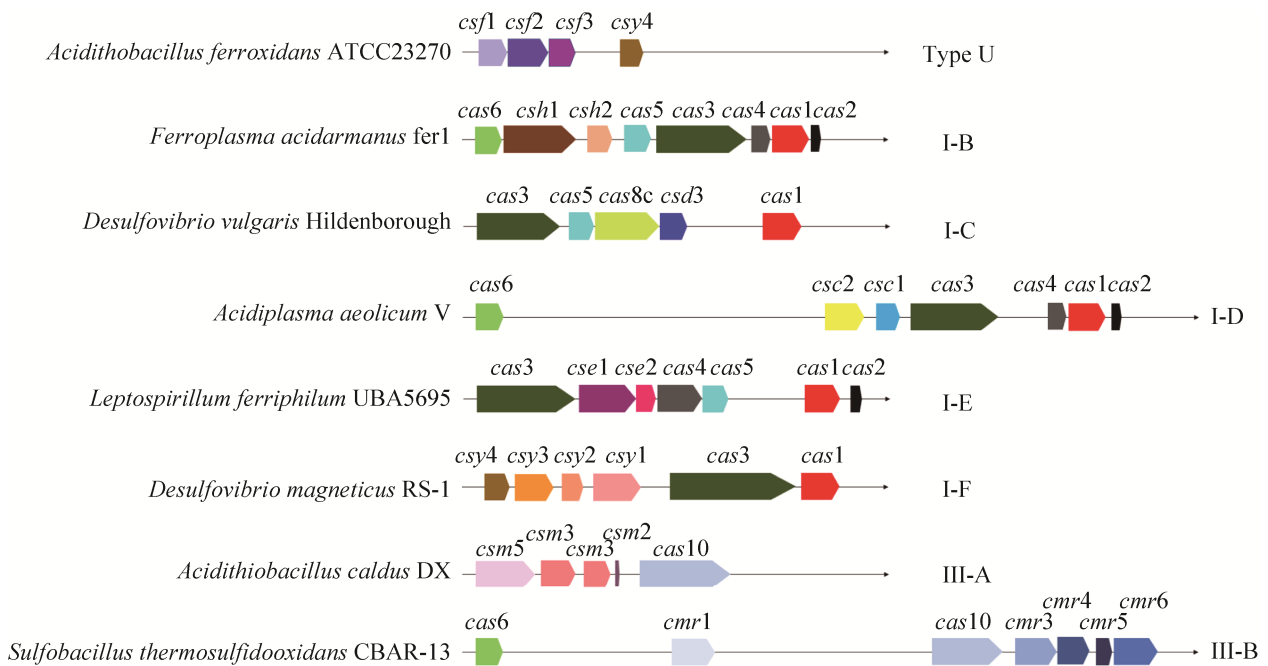


图 3. 100 株冶金微生物基因组中鉴定出的 8 种 CRISPR-Cas 系统分类和分布情况

Figure 3. Classification and distribution of 8 types CRISPR-Cas system identified in genomes of 100 bioleaching microorganisms. The distance between arrows represents the relative position of genes, and the length of arrows represents the relative size of genes.

外源病毒核酸序列的防御体系^[33]。

在古菌 *Acidiplasma* 和 *Ferroplasma* 中, 除菌株 *Ferroplasma acidarmanus* fer1 具有一个 I-B 型 CRISPR 结构, 其余 CRISPR-Cas 系统均属于 I-D 亚类型。据报道, 菌株 *Ferroplasma acidiphilum* Y 的 I-D 型 CRISPR-Cas 系统与 III 型 CRISPR-Cas 系统具有相似性, 推测其重序列形成的二级结构能够被 Cas6 蛋白识别并于 3' 端断裂^[34]。这两种亚类型在 4 个细菌属中未见分布, 在对极端酸性矿山环境适应过程中, 古菌可能进化出了独特的防御体系, 这为深入研究 AMD 环境古菌的适应性免疫机制提供了思路。

3 讨论

自 Haft^[35] 等首次命名了 CRISPR 并发现了与其作用相关的基因后, 该系统愈加受到人们的广泛关注, 其对外源序列的抵抗作用主要包括适应、表达和干扰 3 个阶段^[13]。同时越来越多的 CRISPR 信息随着细菌和古菌全基因组序列的陆续报道而涌现出来, 如对 5 株极端嗜盐古菌 CRISPR 结构的分析显示, 30 bp 长重复序列可形成两类不同的 RNA 二级结构, 可能对介导外源核酸序列与 Cas 蛋白的相互作用具有重要意义^[36]。CRISPR-Cas 系统演变迅速, 并逐渐发展成为强有力的基因编辑工具, 例如 CRISPR-Cas 9 系统被广泛应用于改造基因组的结构和功能^[13]。而对极端酸性矿山环境中典型微生物 CRISPR 结构的研究, 将有助于对细菌与古菌的适应性机制研究与菌种改造提供新思路。

本文在极端酸性矿山环境典型属的 100 株菌中共鉴定出 176 个 cCRISPR 簇, 在大部分细菌基因组中, 普遍含有 1-2 个小 cCRISPR 簇, 古菌

Acidiplasma 菌株平均具有超过 4 个 cCRISPR 簇, *Ferroplasma* 属菌株的 4 个 cCRISPR 簇则平均分布有较多(40 余条)间隔序列。对重复序列进行分类后, 发现同属菌株具有不同类别的重复序列, 形成多种 RNA 二级结构发挥免疫功能, 揭示了其基于 CRISPR 系统抵抗外源入侵核酸序列时可能具有多态性。而同源性较高的重复序列能够广泛分布于不同属菌株中, 其防御系统可能有所类似, 能够有效抵御外界干扰并且维持自身基因组的稳定性。CRISPR-Cas 系统分类结果表明, 在 51 株能够鉴定出 CRISPR-Cas 系统类型的微生物基因组中, 10 株菌同时具有 2 套不同类型的 *cas* 基因, 可能在应对不同核酸序列入侵时分别发挥免疫作用。对大肠杆菌 CRISPR 簇的研究表明重复序列与 *cas* 基因之间具有一致性^[18], 本研究中属于相同亚型 CRISPR 结构的重复序列往往分类至同一 Cluster 中, 推测重复序列与相应的 *cas* 基因可能有一定的结合偏嗜关系, 对于 CRISPR 系统免疫功能的发挥具有重要作用, 二者之间可能存在着密切的功能联系与协同进化关系。

在微生物对环境的漫长适应过程中, CRISPR 簇与质粒、病毒 DNA 等实现了“共进化”^[16]。据报道, 在 51 株产甲烷古菌的 CRISPR 簇中, 一些 GC 含量为 30%-35% 的嗜热菌普遍具有大量间隔序列, 可能遭受了频繁的外源组分攻击^[25]。本文对间隔序列的同源性分析结果表明, 古菌 *Acidiplasma* 和 *Ferroplasma* 有 46 条间隔序列与病毒数据库相匹配。AMD 环境细菌的基因组 GC 含量普遍高于 50%, 而古菌基因组的 GC 含量为 33%-38%, 且无细菌细胞壁的肽聚糖结构, 其热稳定性较差且不如细菌细胞壁坚韧,

故外源组分可能更易进攻成功^[10]。此外, 这些古菌基因组中分布有其他 4 个细菌属种不具有的 I-B 型和 I-D 型 CRISPR-Cas 系统, 其在抵御外源入侵核酸序列时, 可能进化形成了独特的防御机制。

总之, 本研究对酸性矿山环境典型细菌与古菌 CRISPR-Cas 系统的结构与功能进行了生物信息学分析, 增进了对嗜酸微生物基于 CRISPR-Cas 系统维持自身遗传稳定性的认识, 有助于进一步揭示微生物对极端环境的适应性进化机制。

参 考 文 献

- [1] Huang L, Feng XL, Du QS, Dong XZ, Liu SJ, Wen MZ, Dai X. Focusing on key scientific issues of microbiome research in hydrosphere: NSFC major research plan for microbes in hydrosphere. *Bulletin of Chinese Academy of Sciences*, 2017, 32(3): 266–272. (in Chinese)
黄力, 冯雪莲, 杜全生, 东秀珠, 刘双江, 温明章, 戴欣. 水圈微生物重大研究计划: 聚焦水圈微生物组研究的核心科学问题. *中国科学院院刊*, 2017, 32(3): 266–272.
- [2] Zhang X, Liu XD, Liang YL, Fan FL, Zhang XX, Yin HQ. Metabolic diversity and adaptive mechanisms of iron- and/or sulfur-oxidizing autotrophic acidophiles in extremely acidic environments. *Environmental Microbiology Reports*, 2016, 8(5): 738–751.
- [3] Zhang X, Liu XD, Liang YL, Xiao YH, Ma LY, Guo X, Miao B, Liu HW, Peng DL, Huang WK, Yin HQ. Comparative genomics unravels the functional roles of co-occurring acidophilic bacteria in bioleaching heaps. *Frontiers in Microbiology*, 2017, 8: 790.
- [4] Ma LY, Wang XJ, Liu XD, Wang SQ, Wang HM. Intensified bioleaching of chalcopyrite by communities with enriched ferrous or sulfur oxidizers. *Bioresource Technology*, 2018, 268: 415–423.
- [5] Ma LY, Wang HM, Wu JJ, Wang YG, Zhang D, Liu XD. Metatranscriptomics reveals microbial adaptation and resistance to extreme environment coupling with bioleaching performance. *Bioresource Technology*, 2019, 280: 9–17.
- [6] Fariq A, Blazier JC, Yasmin A, Gentry TJ, Deng YJ. Whole genome sequence analysis reveals high genetic variation of newly isolated *Acidithiobacillus ferrooxidans* IO-2C. *Scientific Reports*, 2019, 9(1): 13049.
- [7] Goltsman DSA, Deneff VJ, Singer SW, VerBerkmoes NC, Lefsrud M, Mueller RS, Dick GJ, Sun CL, Wheeler KE, Zemla A, Baker BJ, Hauser L, Land M, Shah MB, Thelen MP, Hettich RL, Banfield JF. Community genomic and proteomic analyses of chemoautotrophic iron-oxidizing “*Leptospirillum rubrum*” (Group II) and “*Leptospirillum ferrodiazotrophum*” (Group III) bacteria in acid mine drainage biofilms. *Applied and Environmental Microbiology*, 2009, 75(13): 4599–4615.
- [8] Ferrer A, Bunk B, Spröer C, Biedendieck R, Valdés N, Jahn M, Jahn D, Orellana O, Levicán G. Complete genome sequence of the bioleaching bacterium *Leptospirillum* sp. Group II strain CF-1. *Journal of Biotechnology*, 2016, 222: 21–22.
- [9] Bulaev AG, Kanygina AV, Manolov AI. Genome analysis of *Acidiplasma* sp. MBA-1, a polyextremophilic archaeon predominant in the microbial community of a bioleaching reactor. *Microbiology*, 2017, 86(1): 89–95.
- [10] Yelton AP, Comolli LR, Justice NB, Castelle C, Deneff VJ, Thomas BC, Banfield JF. Comparative genomics in acid mine drainage biofilm communities reveals metabolic and structural differentiation of co-occurring archaea. *BMC Genomics*, 2013, 14: 485.
- [11] Faure G, Makarova KS, Koonin EV. CRISPR-Cas: complex functional networks and multiple roles beyond adaptive immunity. *Journal of Molecular Biology*, 2019, 431(1): 3–20.
- [12] Grissa I, Vergnaud G, Pourcel C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics*, 2007, 8: 172.
- [13] Li WJ, Tian Y. CRISPR/Cas—a new tool for molecular genetic research. *Acta Microbiologica Sinica*, 2017, 57(11): 1653–1664. (in Chinese)
李文均, 田野. CRISPR/Cas 工具——分子遗传研究的新刃. *微生物学报*, 2017, 57(11): 1653–1664.

- [14] Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 2007, 315(5819): 1709–1712.
- [15] Kunin V, Sorek R, Hugenoltz P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biology*, 2007, 8(4): R61.
- [16] Sorek R, Kunin V, Hugenoltz P. CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature Reviews Microbiology*, 2008, 6(3): 181–186.
- [17] Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, Moineau S, Mojica FJM, Wolf YI, Yakunin AF, van der Oost J, Koonin EV. Evolution and classification of the CRISPR-Cas systems. *Nature Reviews Microbiology*, 2011, 9(6): 467–477.
- [18] Díez-Villaseñor C, Almendros C, García-Martínez J, Mojica FJM. Diversity of CRISPR loci in *Escherichia coli*. *Microbiology*, 2010, 156(5): 1351–1361.
- [19] Li W, Liang HZ, Zhang DQ, Wang NN, Tang YR, Li BL, Huo GC. Detection of CRISPR-Cas system in *Streptococcus thermophilus*. *Acta Microbiologica Sinica*, 2016, 56(4): 680–688. (in Chinese)
李婉, 梁宏彰, 张丹青, 王娜娜, 唐雅茹, 李柏良, 霍贵成. 嗜热链球菌 CRISPR-Cas 系统的检测. *微生物学报*, 2016, 56(4): 680–688.
- [20] Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research*, 2007, 35(S2): W52–W57.
- [21] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li WZ, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 2011, 7: 539.
- [22] Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Research*, 2004, 14(6): 1188–1190.
- [23] Mathews DH, Sabina J, Zuker M, Turner DH. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of Molecular Biology*, 1999, 288(5): 911–940.
- [24] Biswas A, Gagnon JN, Brouns SJ, Fineran PC, Brown CM. CRISPRTarget: bioinformatic prediction and analysis of crRNA targets. *RNA Biology*, 2013, 10(5): 817–827.
- [25] Fang J, Hou JL, Zhang Y, Wang FP, He Y. Distribution of CRISPR arrays in methanogenic archaea. *Microbiology China*, 2016, 43(11): 2353–2365. (in Chinese)
方静, 侯佳林, 张宇, 王风平, 何莹. 产甲烷古菌中 CRISPR 簇的研究. *微生物学通报*, 2016, 43(11): 2353–2365.
- [26] Méndez-García C, Peláez AI, Mesa V, Sánchez J, Golyshina OV, Ferrer M. Microbial diversity and metabolic networks in acid mine drainage habitats. *Frontiers in Microbiology*, 2015, 6: 475.
- [27] Baker BJ, Banfield JF. Microbial communities in acid mine drainage. *FEMS Microbiology Ecology*, 2003, 44(2): 139–152.
- [28] Koonin EV, Wolf YI. Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. *Nucleic Acids Research*, 2008, 36(21): 6688–6719.
- [29] Iwasaki T, Yamashita E, Nakagawa A, Enomoto A, Tomihara M, Takeda S. Three-dimensional structures of bacteriophage neck subunits are shared in *Podoviridae*, *Siphoviridae* and *Myoviridae*. *Genes to Cells*, 2018, 23(7): 528–536.
- [30] Peng TJ, Ma LY, Feng X, Tao JM, Nan MH, Liu YD, Li JK, Shen L, Wu XL, Yu RL, Liu XD, Qiu GZ, Zeng WM. Genomic and transcriptomic analyses reveal adaptation mechanisms of an *Acidithiobacillus ferrivorans* strain YL15 to alpine acid mine drainage. *PLoS One*, 2017, 12(5): e0178008.
- [31] Xia K, Liang XL, Li YD. Comparative genomics and evolutionary analysis of CRISPR loci in acetic acid bacteria. *Hereditas*, 2015, 37(12): 1242–1250. (in Chinese)
夏凯, 梁新乐, 李余动. 醋酸菌中 CRISPR 位点的比较基因组学与进化分析. *遗传*, 2015, 37(12): 1242–1250.
- [32] Crispim JS, Dias RS, Vidigal PMP, de Sousa MP, da Silva CC, Santana MF, de Paula SO. Screening and characterization of prophages in *Desulfovibrio* genomes. *Scientific Reports*, 2018, 8(1): 9273.
- [33] Guo X, Yin HQ, Liang YL, Hu Q, Zhou XS, Xiao YH, Ma

- LY, Zhang X, Qiu GZ, Liu XD. Comparative genome analysis reveals metabolic versatility and environmental adaptations of *Sulfobacillus thermosulfidooxidans* strain ST. *PLoS One*, 2014, 9(6): e99417.
- [34] Golyshina OV, Tran H, Reva ON, Lemak S, Yakunin AF, Goesmann A, Nechitaylo TY, LaCono V, Smedile F, Slesarev A, Rojo D, Barbas C, Ferrer M, Yakimov MM, Golyshin PN. Metabolic and evolutionary patterns in the extremely acidophilic archaeon *Ferroplasma acidiphilum* Y^T. *Scientific Reports*, 2017, 7(1): 3682.
- [35] Haft DH, Selengut J, Mongodin EF, Nelson KE. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Computational Biology*, 2005, 1(6): e60.
- [36] Zhang F, Zhang B, Xiang H, Hu SN. Comparative analysis of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) loci in the genomes of halophilic archaea. *Acta Microbiologica Sinica*, 2009, 49(11): 1445–1453. (in Chinese)
张帆, 张兵, 向华, 胡松年. 极端嗜盐古菌中 CRISPR 结构的生物信息学分析. *微生物学报*, 2009, 49(11): 1445–1453.
- [37] Castro M, Moya-Beltrán A, Covarrubias PC, Gonzalez M, Cardenas JP, Issotta F, Nuñez H, Acuña LG, Encina G, Holmes DS, Johnson DB, Quatrini R. Draft genome sequence of the type strain of the sulfur-oxidizing acidophile, *Acidithiobacillus albertensis* (DSM 14366). *Standards in Genomic Sciences*, 2017, 12: 77.
- [38] Zhang X, Liu XD, He Q, Dong WL, Zhang XX, Fan FL, Peng DL, Huang WK, Yin HQ. Gene turnover contributes to the evolutionary adaptation of *Acidithiobacillus caldus*: insights from comparative genomics. *Frontiers in Microbiology*, 2016, 7: 1960.
- [39] Hedrich S, Johnson DB. *Acidithiobacillus ferridurans* sp. nov., an acidophilic iron-, sulfur- and hydrogen-metabolizing chemolithotrophic gammaproteobacterium. *International Journal of Systematic and Evolutionary Microbiology*, 2013, 63(11): 4018–4025.
- [40] Ccorahua-Santo R, Eca A, Abanto M, Guerra G, Ramírez P. Physiological and comparative genomic analysis of *Acidithiobacillus ferrivorans* PQ33 provides psychrotolerant fitness evidence for oxidation at low temperature. *Research in Microbiology*, 2017, 168(5): 482–492.
- [41] Chen P, Yan L, Wu ZR, Xu RX, Li SY, Wang NB, Liang N, Li HY. Draft genome sequence of extremely acidophilic bacterium *Acidithiobacillus ferrooxidans* DLC-5 isolated from acid mine drainage in northeast China. *Genomics Data*, 2015, 6: 267–268.
- [42] Zhang X, Feng X, Tao JM, Ma LY, Xiao YH, Liang YL, Liu XD, Yin HQ. Comparative genomics of the extreme acidophile *Acidithiobacillus thiooxidans* reveals intraspecific divergence and niche adaptation. *International Journal of Molecular Sciences*, 2016, 17(8): 1355.
- [43] Quatrini R, Escudero LV, Moya-Beltrán A, Galleguillos PA, Issotta F, Acosta M, Cárdenas JP, Nuñez H, Salinas K, Holmes DS, Demergasso C. Draft genome sequence of *Acidithiobacillus thiooxidans* CLST isolated from the acidic hypersaline gorcea salt flat in northern Chile. *Standards in Genomic Sciences*, 2017, 12(1): 84.
- [44] Abildgaard L, Nielsen MB, Kjeldsen KU, Ingvorsen K. *Desulfovibrio alkalitolerans* sp. nov., a novel alkalitolerant, sulphate-reducing bacterium isolated from district heating water. *International Journal of Systematic and Evolutionary Microbiology*, 2006, 56(5): 1019–1024.
- [45] Lobo SAL, Melo AMP, Carita JN, Teixeira M, Saraiva LM. The anaerobe *Desulfovibrio desulfuricans* ATCC 27774 grows at nearly atmospheric oxygen levels. *FEBS Letters*, 2007, 581(3): 433–436.
- [46] Spring S, Sorokin DY, Verborg S, Rohde M, Woyke T, Kyrpides NC. Sulfate-reducing bacteria that produce exopolymers thrive in the calcifying zone of a hypersaline cyanobacterial mat. *Frontiers in Microbiology*, 2019, 10: 862.
- [47] Day LA, De León KB, Kempfer ML, Zhou JZ, Wall JD. Complete genome sequence of *Desulfovibrio desulfuricans* IC1, a sulfonate-respiring anaerobe. *Microbiology Resource Announcements*, 2019, 8(31): e00456–19.
- [48] Brown SD, Gilmour CC, Kucken AM, Wall JD, Elias DA, Brandt CC, Podar M, Chertkov O, Held B, Bruce DC, Detter JC, Tapia R, Han CS, Goodwin LA, Cheng JF, Pitluck S,

- Woyke T, Mikhailova N, Ivanova NN, Han J, Lucas S, Lapidus AL, Land ML, Hauser LJ, Palumbo AV. Genome sequence of the mercury-methylating strain *Desulfovibrio desulfuricans* ND132. *Journal of Bacteriology*, 2011, 193(8): 2078–2079.
- [49] Nakazawa H, Arakaki A, Narita-Yamada S, Yashiro I, Jinno K, Aoki N, Tsuruyama A, Okamura Y, Tanikawa S, Fujita N, Takeyama H, Matsunaga T. Whole genome sequence of *Desulfovibrio magneticus* strain RS-1 revealed common gene clusters in magnetotactic bacteria. *Genome Research*, 2009, 19(10): 1801–1808.
- [50] Heidelberg JF, Seshadri R, Haveman SA, Hemme CL, Paulsen IT, Kolonay JF, Eisen JA, Ward N, Methe B, Brinkac LM, Daugherty SC, Deboy RT, Dodson RJ, Durkin AS, Madupu R, Nelson WC, Sullivan SA, Fouts D, Haft DH, Selengut J, Peterson JD, Davidsen TM, Zafar N, Zhou LW, Radune D, Dimitrov G, Hance M, Tran K, Khouri H, Gill J, Utterback TR, Feldblyum TV, Wall JD, Voordouw G, Fraser CM. The genome sequence of the anaerobic, sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. *Nature Biotechnology*, 2004, 22(5): 554–559.
- [51] Justice NB, Norman A, Brown CT, Singh A, Thomas BC, Banfield JF. Comparison of environmental and isolate *Sulfobacillus* genomes reveals diverse carbon, sulfur, nitrogen, and hydrogen metabolisms. *BMC Genomics*, 2014, 15: 1107.
- [52] Zhang X, Liu XD, Liang YL, Guo X, Xiao YH, Ma LY, Miao B, Liu HW, Peng DL, Huang WK, Zhang YG, Yin HQ. Adaptive evolution of extreme acidophile *Sulfobacillus thermosulfidooxidans* potentially driven by horizontal gene transfer and gene loss. *Applied and Environmental Microbiology*, 2017, 83(7): e03098–16.
- [53] Golyshina OV, Yakimov MM, Lünsdorf H, Ferrer M, Nimtz M, Timmis KN, Wray V, Tindall BJ, Golyshin PN. *Acidiplasma aeolicum* gen. nov., sp. nov., a euryarchaeon of the family *Ferroplasmaceae* isolated from a hydrothermal pool, and transfer of *Ferroplasma cupricumulans* to *Acidiplasma cupricumulans* comb. nov. *International Journal of Systematic and Evolutionary Microbiology*, 2009, 59(11): 2815–2823.
- [54] Dopson M, Baker-Austin C, Hind A, Bowman JP, Bond PL. Characterization of *Ferroplasma* isolates and *Ferroplasma acidarmanus* sp. nov., extreme acidophiles from acid mine drainage and industrial bioleaching environments. *Applied and Environmental Microbiology*, 2004, 70(4): 2079–2088.

Adaptive immunity mechanisms of microorganisms in extreme acid mine environment based on CRISPR system

Shanshan Huang¹, Liyuan Ma^{1,2*}, Hongmei Wang^{1,3}, Xueduan Liu²

¹ School of Environmental Studies, China University of Geosciences, Wuhan 430074, Hubei Province, China

² Key Laboratory of Biometallurgy of Ministry of Education, Central South University, Changsha 410083, Hunan Province, China

³ State Key Laboratory of Geobiology and Environmental Geology, China University of Geosciences, Wuhan 430074, Hubei Province, China

Abstract: [Objective] To investigate the adaptive mechanism of bioleaching microorganisms to extreme environment based on CRISPR system at the genomic level, bioinformatic analyses were performed on the structural characteristics and homology of CRISPR-Cas system in genomes of 100 species from genera (*Acidithiobacillus*, *Desulfovibrio*, *Leptospirillum*, *Sulfobacillus*, *Acidiplasma* and *Ferroplasma*) in acid mine environment. [Methods] We downloaded genome sequences from NCBI website and identified potential CRISPR arrays by using CRISPR Finder. The composition, structure and function of each CRISPR system were analyzed, where repeats were classified by Clustal Omega and spacers were aligned and annotated according to nr database, plasmid database and viral database, respectively. The CRISPR-Cas system of microorganisms in acid mine environment was classified based on the types and homology of Cas proteins. [Results] Among the genomes of 100 bioleaching microorganisms, we found 415 CRISPR arrays. There were 80 different repeats and 4147 spacers in 176 confirmed CRISPR arrays. All the 12 types of repeat sequences in each cluster could form typical RNA secondary structure and the sequence of cluster 10 was the most representative one among all the bioleaching microorganisms. The annotation results showed that these microorganisms have been attacked by bacterial plasmids and virus, and have resisted the invasion of foreign nucleic acid sequences through different defense mechanisms. Most of the CRISPR-Cas systems of bioleaching bacteria belong to I-C and I-E subtypes, while most of the CRISPR-Cas systems of archaea belong to I-D subtype. There are significant differences between them in the evolution process based on the CRISPR-Cas system. [Conclusion] The CRISPR structure of acid mine environment microorganisms may mediate the interaction between foreign nucleic acid sequence and Cas protein based on different immune mechanisms, which provides a foundation for further revealing the adaptive evolution mechanism of extreme environment microorganisms.

Keywords: acid mine environment, CRISPR-Cas system, repeat, spacer, adaptive mechanism

(本文责编: 张晓丽)

Supported by the National Key Research and Development Program of China (2019YFC1803604), by the Research Fund Program of Key Laboratory of Biometallurgy, Ministry of Education, Central South University (MOEKL02), by the Provincial Water Resources Development Fund Project of Yangquan Municipal Water Resources Bureau (2019046477) and by the Research Fund Program of Experimental Technology, China University of Geosciences (Wuhan) (SJ-201903)

*Corresponding author. E-mail: maly@cug.edu.cn

Received: 21 January 2020; Revised: 19 March 2020; Published online: 8 April 2020



马丽媛, 博士, 中国地质大学(武汉)环境学院副教授, “地大学者”青年优秀人才, 硕士生导师。2017年毕业于中南大学, 获理学博士学位。主要从事微生物介导的元素地球化学循环、微生物功能组学及微生物生态学研究, 研究载体涉及酸性矿山、镉污染矿区、煤矿老窑水以及喀斯特洞穴生态系统等。主持国家重点研发计划重点专项子课题、湖北省自然科学基金、中央高校基本科研业务费专项资金杰出人才培育基金等科研项目7项, 作为项目骨干(排名第三)参与水圈重点项目“岩溶地下水系统洞穴微生物驱动的甲烷循环及其生态效应”, 参与编写专著和教材2部。近五年发表SCI论文40余篇, 以第一作者或通讯作者在 *Bioresourcetechnology*、*Hydrometallurgy* 和 *Applied Microbiology and Biotechnology* 等刊物上发表学术论文11篇。