

嗜酸硫杆菌 *Acidithiobacillus* 抗砷基因多样性分析

赵永红, 张静, 周丹, 彭艳平, 余水静

江西理工大学资源与环境工程学院, 江西省矿冶环境污染控制重点实验室, 江西 赣州 341000

摘要: 为探讨嗜酸硫杆菌(*Acidithiobacillus*)抗砷基因多样性特征, 利用生物信息学方法挖掘出嗜酸氧化亚铁硫杆菌(*A. ferrooxidans*)、喜温嗜酸硫杆菌(*A. caldus*)和耐冷嗜酸硫杆菌(*A. ferrivorans*)共49个抗砷基因, 主要包括构成 ars 操纵子模式的 $arsA$ 、 $arsB$ 、 $arsC$ 、 $arsD$ 、 $arsH$ 和 $arsR$ 基因, 为嗜酸硫杆菌重要的抗砷遗传因子。研究了49个抗砷基因的聚类与进化关系、理化性质、二级结构、跨膜区、亚细胞定位、砷结合位点等特征。结果表明: 这些抗砷基因聚类为11个组, 第2组抗砷基因($arsB$)为嗜酸硫杆菌所共有, 编码泵砷内膜蛋白(ArsB), 该基因可能为嗜酸硫杆菌抗砷机制所必需; 聚类和进化分析表明其中5个抗砷基因可能源于基因水平迁移; 脂溶指数分析显示25个抗砷蛋白脂溶指数大于100, 为亲水性蛋白, 其余24个抗砷蛋白脂溶指数均小于100, 为疏水性蛋白; 从不稳定指数分析结果可知, 大部分抗砷转录调控因子ArsR是不稳定的, 49个抗砷基因蛋白二级结构均由 α -螺旋、无规则卷曲和扩展链结构3种形式组成。亚细胞定位分析发现抗砷蛋白在细菌内膜、细菌外膜和细胞周质空间均有出现, ArsB具有10-11个跨膜区; 所有抗砷蛋白均无信号肽。对嗜酸硫杆菌的ArsC、ArsD和ArsH抗砷功能蛋白与砷结合位点进行预测和建模, 发现活性位点均存在Cys残基; 与其他抗砷蛋白结合位点(含Cys残基)不同的是, Acife1530(ArsA)2个甲硫氨酸(Met)与As(III)直接结合, 暗示着Met残基中存在与Cys残基功能相同的巯基, 由此可见, Cys残基或Met残基的巯基对于构成嗜酸硫杆菌抗砷功能蛋白与砷结合的活性位点是必不可少的。

关键词: 嗜酸硫杆菌; 抗砷基因; 聚类分析; 多样性分析; 生物信息学

中图分类号: Q938.1; X753

文献标志码: A

文章编号: 1674-5906(2013)07-1141-07

引用格式: 赵永红, 张静, 周丹, 彭艳平, 余水静. 嗜酸硫杆菌 *Acidithiobacillus* 抗砷基因多样性分析[J]. 生态环境学报, 2013, 22(7): 1141-1147.

ZHAO Yonghong, ZHANG Jing, ZHOU Dan, PENG Yanping, YU Shuijing. Diversity analysis of arsenic resistance genes in *Acidithiobacillus* genus [J]. Ecology and Environmental Sciences, 2013, 22(7): 1141-1147.

嗜酸硫杆菌(*Acidithiobacillus*)是一类极其重要的极端微生物, 可通过氧化硫和还原态无机硫化物获得能量, 能在极酸(pH1-3)、高浓度重金属环境中生长繁殖^[1]。嗜酸硫杆菌属主要代表菌有中温嗜酸氧化亚铁硫杆菌(*A. ferrooxidans*)、喜温嗜酸硫杆菌(*A. caldus*)和耐冷嗜酸硫杆菌(*A. ferrivorans*)。砷是广泛存在于自然界的一种有毒元素, 伴随着矿产资源的开采, 大量的砷被带到环境中, 产生了严重的环境问题。研究表明, 嗜酸硫杆菌能含砷环境中生长, 生物浸出高砷原生硫化铜矿时, 嗜酸氧化亚铁硫杆菌对砷的耐受性比喜温嗜酸硫杆菌高^[2]。在进化过程中嗜酸硫杆菌产生了对抗砷毒的生理机制, 而这些机制的产生往往由其遗传学基础所决定。

ars 操纵子(Arsenic-resistance operon)可很好地阐释微生物抗砷机制, 该操纵子编码转录调控蛋白(Regulatory protein, ArsR)、砷酸盐透性酶(Arsenate

permease, ArsB)和砷酸盐还原酶(Arsenate reductase, ArsC)等。ArsR 是一个抑制子, 可结合到启动区域并调节 ars 操纵子的表达水平, ArsB 是一种跨膜砷外排泵蛋白, ArsC 负责将砷酸盐还原为亚砷酸盐^[3-5]。研究发现, 嗜酸硫杆菌属内各菌种对砷的耐受性有明显的差异^[2], 抗砷细菌 $arsH$ 基因在不同的物种之间也存在基因水平转移现象^[6], 因此, 可以推测嗜酸硫杆菌抗砷基因必然存在着多样性以适应复杂的含砷极端环境, 而嗜酸硫杆菌抗砷基因多样性特征分析仍未见报道。

许多科学工作者试图从基因组学角度探究微生物遗传型和表型之间的相关性。从生物信息学角度探究嗜酸硫杆菌抗砷基因多样性特征, 有助于理解这些菌株生长代谢及其对极端环境适应性分子机制。目前, 嗜酸硫杆菌基因组测序已完成的有 *A. ferrooxidans*、*A. caldus* 和 *A. ferrivorans* 菌株。因此, 本文拟对已测序的嗜酸硫杆菌抗砷基因进行挖

基金项目: 国家自然科学基金项目(51064007); 江西省科技厅对外科技合作计划项目(20111bdh80032); 江西省科技厅科技支撑计划项目(20123bbf60170; 20121BBG70004); 江西省教育厅科学技术研究项目(GJJ13397)

作者简介: 赵永红(1977年生), 男, 副教授, 博士, 主要从事矿冶环境污染评测与控制研究。E-mail: szgl@mail.jxust.cn

收稿日期: 2013-05-05

表1 嗜酸硫杆菌基因组信息及抗砷基因数目

Table 1 The information of genomes and arsenic resistance genes from *Acidithiobacillus* genus

嗜酸硫杆菌	基因组登录号	基因组大小(M bp)	总基因数	抗砷相关基因数	抗砷相关转录调控基因数
<i>A. caldus</i> SM1	NC_015850	3.24	3240	2	8
<i>A. caldus</i> ATCC 51756	—	2.95	2878	2	6
<i>A. ferrivorans</i> SS3	NC_015942	3.21	3335	4	9
<i>A. ferrooxidans</i> ATCC 23270	NC_011761	2.98	3303	3	5
<i>A. ferrooxidans</i> ATCC 53993	NC_011206	2.89	2959	2	8

掘并系统地分析,从抗砷基因聚类和进化关系、理化性质、二级结构、砷结合位点等方面阐释嗜酸硫杆菌抗砷基因多样性特征。

1 材料和方法

1.1 嗜酸硫杆菌抗砷基因挖掘

从美国国立生物技术信息中心网站(National Center for Biotechnology Information, NCBI) (<http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi>) 收集嗜酸硫杆菌 *A. caldus* ATCC 51756、*A. caldus* SM-1、*A. ferrivorans* SS3、*A. ferrooxidans* ATCC 23270、*A. ferrooxidans* ATCC 53993 共 5 个菌株基因组信息(包括后缀名为 .ffn、.faa、.ptt 的 3 种类型文件),利用关键词“ars”对 .ptt 文件进行全面扫描搜索抗砷相关基因,再对 .faa 文件进行检索相应抗砷基因的氨基酸序列。

1.2 抗砷基因聚类和进化分析

利用 OrthoMCL 数据库网站 (<http://www.orthomcl.org>) 对嗜酸硫杆菌抗砷相关基因进行聚类分析^[7];利用 MEGA 4.1 软件对抗砷基因进行多重序列比对,并利用 Neighbor-joining 法构建进化树。

1.3 抗砷基因生物信息学分析

对抗砷基因进行生物信息学分析,包括利用 ProtParam 在线工具计算抗砷基因的理化性质,如氨基酸数目、相对分子量、理论等电点、分子式、脂溶指数和不稳度指数等^[8];蛋白二级结构使用 Hnn Secondary Prediction Method 程序预测 (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hnn.html)^[9];利用 TMpredict 进行蛋白质序列跨膜区分析 (<http://www.cbs.dtu.dk/services/TMHMM/>)^[10];利用 SignalP v4.0 进行信号肽预测 (<http://www.cbs.dtu.dk/services/SignalP/>)^[11];利用 PSORT (<http://psort.hgc.jp/form.html>) 进行亚细胞定位分析^[12]。

1.4 抗砷功能蛋白与砷结合位点预测

将抗砷氨基酸序列提交至 NCBI 网站 PDB 数据库进行 PSI-BLAST 比对,以获得同源性 PDB-ID,用 SWISS-MODEL 在线服务器预测抗砷功能蛋白的分子模型^[13],将建立的分子模型存为后缀名为 .pdb 文件,上传至 Q-SiteFinder 在线工具

(<http://www.modelling.leeds.ac.uk/qsitefinder/>) 预测砷结合位点^[14]。

2 结果与分析

2.1 嗜酸硫杆菌抗砷基因挖掘

从 NCBI 网站获得 *A. caldus* SM1、*A. caldus* ATCC 51756、*A. ferrivorans* SS3、*A. ferrooxidans* ATCC 23270 和 *A. ferrooxidans* ATCC 53993 共 5 株嗜酸硫杆菌基因组信息,利用关键词“ars”对 .ptt 文件进行全面扫描搜索,挖掘出 49 个抗砷基因,包括 13 个抗砷功能基因和 36 个抗砷转录调控基因(arsR)(表 1 和表 2)。从抗砷基因数量分布来看,*A. ferrivorans* SS3 菌株抗砷功能基因和抗砷转录调控基因数量最多,暗示着该菌在含砷环境中具有较好的适应能力。

表 2 嗜酸硫杆菌 13 个抗砷基因功能注释

Table 2 Functional annotation of 13 arsenic resistance genes in *Acidithiobacillus* genus

基因号	基因功能注释
Lferr 2478	砷抗性蛋白(arsenical resistance protein ArsH)
Lferr 2479	泵砷膜蛋白(arsenical pump membrane protein ArsB)
AFE 2857	砷抗性蛋白(arsenical resistance protein ArsH)
AFE 2858	泵砷膜蛋白(arsenical pump membrane protein ArsB)
AFE 2860	砷酸盐还原酶(arsenate reductase ArsC)
ACA 2815	砷酸盐还原酶(arsenate reductase ArsC)
ACA 2817	砷泵出蛋白(arsenic efflux pump protein ArsB)
Atc 0975	砷酸盐还原酶(arsenate reductase ArsC)
Atc 0977	砷泵出蛋白(arsenic efflux pump protein ArsB)
Acife 0028	砷酸盐还原酶(arsenate reductase ArsC)
Acife 1530	亚砷酸盐-激活 ATP 酶(arsenite-activated ATPase ArsA)
Acife 1531	砷抗性操纵子抑制因子(Arsenical resistance operon trans-acting repressor ArsD)
Acife 2213	泵砷膜蛋白(arsenical pump membrane protein ArsB)

从表 2 抗砷基因功能注释可以看出,嗜酸硫杆菌 13 个抗砷基因功能蛋白主要为亚砷酸盐-激活 ATP 酶(ArsA)、泵砷膜蛋白(ArsB)、砷酸盐还原酶(ArsC)、砷抗性操纵子抑制因子(ArsD)和砷抗性蛋白(ArsH)。这些抗砷功能基因与抗砷转录调控基因(arsR)共同构成一个 ars 操纵子,成为嗜酸硫杆菌重

表3 嗜酸硫杆菌49个抗砷基因聚类分析
Table 3 Ortholog analysis of 49 arsenic resistance genes in *Acidithiobacillus* genus

Groups	<i>A. caldus</i> ATCC 51756	<i>A. caldus</i> SM1	<i>A. ferrivorans</i> SS3	<i>A. ferrooxidans</i> ATCC 23270	<i>A. ferrooxidans</i> ATCC 53993
1	—	—	Acife 1530	—	—
2	ACA 2817	Atc 0977	Acife 2213	AFE 2858	Lferr 2479
3	ACA 2815	Atc 0975	—	AFE 2860	—
4	—	—	Acife 0028	—	—
5	—	—	—	AFE 2857	Lferr 2478
6	—	—	Acife 1531	—	—
7	—	Atc 1809	Acife 0390	AFE 2796	Lferr 2420
8	ACA 2816	Atc 0976	Acife 2214	AFE 2859	Lferr 2480
	ACA 2532	Atc 1751	Acife 2421	AFE 2369	Lferr 2002
	ACA 2791	Atc 0951	Acife 1344	AFE 2026	Lferr 1691
	ACA 0638	Atc 0550	Acife 1334	AFE 2641	Lferr 2267
	—	Atc 1903	—	—	—
	—	Atc 0558	Acife 0826	—	Lferr 0187
9*	ACA 0632	—	—	—	Lferr 0208
	—	Atc 1895	—	—	Lferr 0191
10	—	—	Acife 1478	—	—
11	ACA 0465	—	—	—	—

*: 根据图 1 抗砷基因进化关系对第 9 组抗砷转录调控基因进行了细分

要的抗砷模式^[15-16]。

2.2 抗砷基因聚类分析

对 49 个抗砷基因进行 OrthoMCL 聚类分析, 分别对抗砷功能基因和抗砷转录调控基因进行聚类, 结果如表 3 和图 1 所示, 共聚类为 11 组, 其中 1-6 组为抗砷功能基因, 7-11 组为抗砷转录调控基因。

从抗砷功能基因聚类分析结果 (1-6 组) 可以看出, 第 2 组抗砷功能基因为 5 株嗜酸硫杆菌所共有, 该组基因编码泵砷内膜蛋白 (ArsB), 推测该基因为嗜酸硫杆菌抗砷所必需基因; 而 Acife 1530、Acife 0028 和 Acife 1531 基因为 *A. ferrivorans* SS3 菌株所独有; *A. caldus* 和 *A. ferrooxidans* 均含有砷酸盐还原酶 (ArsC)。从抗砷功能基因的数量来看, 显示出 *A. ferrivorans* SS3 菌株具有较好的抗砷性能。

从抗砷转录调控基因聚类分析结果 (7-11 组) 可以看出, 第 8 组抗砷转录调控基因为 5 株嗜酸硫杆菌所共有; 第 9 组的转录调控基因最多, 共有 24 个基因, 其中有 3 个亚组为 5 株嗜酸硫杆菌所共有; 第 10 组 Acife 1478 基因为 *A. ferrivorans* SS3 菌株所独有; 第 11 组 ACA 0465 为 *A. caldus* ATCC 51756 菌株所独有。

2.3 抗砷基因进化分析

用 MEGA 4.1 软件对 49 个抗砷基因构建了进化树 (见图 1), 从图 1 中可以看出抗砷基因进化关系的远近。其中, Acife 1530、Acife 1531、Acife 0028、

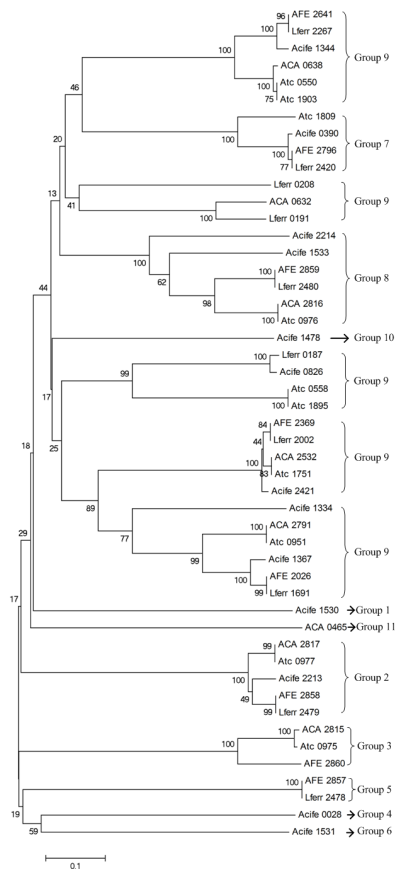


图 1 嗜酸硫杆菌 49 个抗砷基因的系统发育树, 其中右侧大括号标注抗砷基因所属的聚类组别, 图例为遗传距离
Fig.1 Neighbour-joining phylogenetic tree based on 49 arsenic resistance genes in *Acidithiobacillus* genus. The brackets clustering groups of arsenic resistance genes were labeled in the right side of the brackets. Bar, 0.1 substitutions per amino acid position

Acife 1478 和 ACA 0465 基因在进化树上形成一个独立的小分枝, 在聚类分析中也为一个独立的组, 推测这些基因可能来源于物种间的基因水平迁移; Lferr 1691 与 AFE 2026、Atc 0951 与 ACA 2791、Lferr 2002 与 AFE 2369、Atc 1751 与 ACA 2532、Lferr 2480 与 AFE 2859、Lferr 2479 与 AFE 2858、Atc 0977 与 ACA 2817、Lferr 2478 与 AFE 2857、Atc 0976 与 ACA 2816、Atc 0558 与 Atc 1895、Atc 0550 与 Atc 1903、Lferr 2420 与 AFE 2796 共 12 对基因序列两两高度相似。对比抗砷基因的进化关系和聚类分析, 发现进化分支和聚类分析有较好的相关性, 见表 3 和图 1 所示。

2.4 抗砷蛋白理化性质分析

对嗜酸硫杆菌抗砷基因编码的蛋白理化性质进行了分析 (数据结果没有显示), 等电点分析显示: 其中 18 个抗砷蛋白等电点小于 7.0, 呈酸性, Acife 1531 等电点最低(4.71); 31 个抗砷蛋白等电点大于 7.0, 呈碱性, Acife 1367 最高(11.14)。脂溶指数分析显示: 其中 25 个抗砷蛋白脂溶指数大于 100, 为亲水性蛋白, 第 2 组抗砷基因 Lferr 2479、Atc 0977、Acife 2213、ACA 2817 和 AFE 2858 编码的泵砷内膜蛋白脂溶指数均大于 100, 为亲水性蛋白; 其余 24 个抗砷蛋白脂溶指数均小于 100, 为疏水性蛋白。不稳定指数分析表明: 31 个抗砷蛋白不稳定指数大于 40, 为不稳定蛋白; 18 个抗砷蛋白不稳定指数小于 40, 为稳定蛋白, 其中, 第 2 组抗砷功能基因 Lferr 2479、Atc 0977、Acife 2213、ACA 2817 和 AFE 2858 编码的泵砷内膜蛋白不稳定指数均小于 40, 为稳定蛋白。从不稳定指数分析结果可知, 大部分抗砷蛋白是不稳定的, 其中多数为抗砷相关转录调控因子 ArsR, 由此可以推测, 嗜酸硫杆菌抗砷基因表达调控可能受外界环境条件刺激诱导产生。

2.5 抗砷基因蛋白二级结构分析

对 49 个嗜酸硫杆菌抗砷基因蛋白质二级结构分析表明 (数据结果没有显示), 这些抗砷基因蛋白二级结构均由 α -螺旋、无规则卷曲和扩展链结构 3 种形式组成。其中 10 个抗砷基因蛋白 Lferr 2478、

Atc 0975、AFE 2641、AFE 2860、ACA 2815、ACA 0632、Acife 1334、Acife 1478、Acife 1531、Acife 1530 中各组成的百分比为无规则卷曲> α -螺旋>扩展链, 其余 39 个抗砷基因蛋白的百分比皆为 α -螺旋>无规则卷曲>扩展链。

2.6 抗砷基因蛋白跨膜区预测和亚细胞定位分析

对 49 个抗砷基因编码的蛋白序列进行亚细胞定位分析 (数据结果没有显示), 跨膜区预测结果表明: 嗜酸硫杆菌共有的第 2 组抗砷功能基因所编码的泵砷内膜蛋白均存在跨膜区, 其中 Lferr 2479、Atc 0977 和 Acife 2213 存在 10 个跨膜区, ACA 2817 和 AFE 2858 存在 11 个跨膜区; 此外, Acife 0826 和 Acife 0826 各有 1 个跨膜区, 其余 42 个抗砷基因编码的蛋白均不存在跨膜区。

亚细胞定位分析结果表明: Lferr 2479、Atc 0977、Acife 2213、ACA 2817 和 AFE 2858 五个基因编码的泵砷内膜蛋白位于细菌内膜, Lferr 0187 和 Atc 1809 也可能在细菌内膜出现; 令人惊讶的预测结果是 2 个抗砷转录调控因子 Lferr 2420 和 AFE 2796 出现在周质空间或细菌外膜上, 这 2 个转录调控因子如何去实施其对 *ars* 操纵子的调控功能, 是一个值得令人思考的科学问题; 其余 40 个抗砷基因蛋白均分布于细胞质中; 此外, 信号肽分析结果表明, 所有抗砷基因蛋白均无信号肽。

2.7 抗砷功能基因蛋白与砷结合位点预测

研究表明, 抗砷功能基因蛋白存在 1~3 个 Cys 残基的巯基与 As(III)直接结合^[17-18]。本研究对嗜酸硫杆菌 *ars* 操纵子上表达的 ArsA、ArsB、ArsC、ArsD 和 ArsH 5 个抗砷功能蛋白进行了砷结合位点预测, 通过 PSI-BLAST 比对分析获得同源性 PDB-ID, 进一步分析获得 5 个抗砷功能基因蛋白与砷结合位点, 见表 4, 与砷结合位点的预测模型, 见图 2。分析预测结果发现, 与其他抗砷蛋白结合位点 (Cys 残基) 不同的是, Acife 1530 (ArsA) 2 个甲硫氨酸 (MET) 的巯基与 As(III)直接结合, 而 AFE 2858 (ArsB) 未能预测出模型结果, 可能存在其它模式的结合位点或与选用的预测软件有关, 有

表 4 嗜酸硫杆菌抗砷功能基因蛋白与砷结合位点预测

Table 4 Binding sites of ArsA-As, ArsC-As, ArsD-As and ArsH-As in *Acidithiobacillus* genus

基因号	PDB-ID	与砷结合位点相关氨基酸 (数字为该氨基酸在蛋白序列中的位置)
Acife 1530(ArsA)	1F48_A	ASP459、ARG466、HIS478、TYR479、THR480、THR481、PRO482、 <u>MET483*</u> 、 <u>MET484</u> 、GLN485
AFE 2860(ArsC)	1JL3_A	<u>CYS10</u> 、THR11、GLY12、ASN13、SER14、 <u>CYS15</u> 、ARG16、SER17、ILE18、ASP106、PRO107、ALA108
Acife 1531(ArsD)	3MWH_A	<u>CYS12</u> 、CYS13、SER14、THR15、GLY16、VAL17、 <u>CYS18</u> 、GLY19、PRO20、THR21、ILE22、ASP23
AFE 2857(ArsH)	2Q62_A	GLY38、SER39、ASN40、ARG41、 <u>CYS43</u> 、SER44、TYR45、SER46、PRO103、GLU104、ARG105、THR110、GLY111、VAL112、VAL140、 <u>CYS141</u> 、GLY142、ALA174、PRO75、ASP76、ASP118

*: 下划线标注的氨基酸为与砷直接结合的 MET 或 CYS

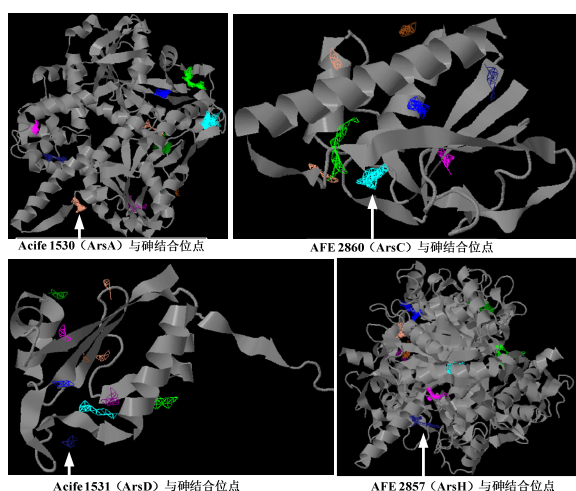


图2 嗜酸硫杆菌抗砷功能蛋白与砷结合位点预测模型

Fig.2 The structure of binding sites of ArsA-As, ArsC-As, ArsD-As and ArsH-As in *Acidithiobacillus* genus

待于进一步深入研究。

3 讨论

原核生物大多数基因表达调控是通过操纵子机制实现的。常见的 *ars* 操纵子有 4 类: *arsRDABC*、*arsCRBH*、*arsRBC*、抗砷基因簇 (*arsADRC* 和 *arsR2M*)，这些操纵子除了含有必须的 *arsR* 调控基因之外，大部分操纵子都含有 *arsB* 基因，该基因编码泵砷内膜蛋白(ArsB)。研究表明，空肠弯曲菌 (*Campylobacter jejuni*) 和喜温嗜酸硫杆菌 (*A. caldus*) 突变株 *arsB* 基因缺失导致其抗砷性能大大降低^[19-20]，说明该基因在生物体抗砷机制中扮演着非常重要的角色。ArsB 是一个能够把 As(III) 泵出细菌体外的阴离子泵，ArsC 是将 As(IV) 转换为 As(III) 的砷还原酶，砷酸盐只有通过 ArsC 还原成亚砷酸盐后才可以被 ArsB 泵出细胞外^[21]。本研究对嗜酸硫杆菌抗砷基因聚类分析发现，第 2 组基因 *arsB* 为嗜酸硫杆菌所共有，该基因编码泵砷内膜蛋白(ArsB)，推测该基因为嗜酸硫杆菌抗砷性能所必需基因。浸矿细菌嗜铁钩端螺旋菌 (*Leptospirillum ferriphilum*) 中也存在这些功能基因 *arsA*、*arsB*、*arsC* 和调节基因 *arsR* 和 *arsD*，以及一个磷酸转移系统的大的抗砷基因簇^[22]。此外，在喜温嗜酸硫杆菌 (*A. caldus*) 中还发现源于转位子 (Tn*AtcArs*) 的携带一组抗砷基因的 *ars* 操纵子^[23]，推测生物体一些抗砷基因可能发生基因水平迁移，大肠杆菌 (*Escherichia coli*) 菌株间发生质粒传递抗砷性的现象已被证实^[24]。本研究通过进化分析发现 5 个嗜酸硫杆菌抗砷基因在进化树上形成一个小分支，在聚类分析中也为一个独立的组，推测这些基因可能来源于物种间的基因水平迁移。

ArsR 是一个抑制子，通过结合到启动区域来调

节 *ars* 操纵子的表达水平。ArsR 蛋白家族呈现出一个共同的二级结构特征($\alpha 1-\alpha 2-\alpha 3-\alpha 4-\beta 1-\beta 2-\alpha 5$)，并包含与 DNA 结合的翼状螺旋家族($\alpha 3$ -turn- $\alpha 4$)^[25]。本文预测的 ArsR 二级结构均由 α -螺旋、无规则卷曲和扩展链结构 3 种形式组成，呈现出多样性特征；嗜酸氧化亚铁硫杆菌 (*A. ferrooxidans*) 响应 As(III) 的转录调控与半胱氨酸残基 (Cys) 直接相关，3 个 Cys 残基与 As(III) 结合，可释放启动子解除抑制启动抗砷基因表达^[26-27]。从本研究结果可知，多数抗砷相关转录调控因子 ArsR 是不稳定的，暗示着嗜酸硫杆菌抗砷基因表达调控受外界环境条件中砷刺激诱导。研究证实，当培养条件中添加亚砷酸盐 As(III) 或砷酸盐 As(V) 时，嗜酸氧化亚铁硫杆菌 (*A. ferrooxidans*) 抗砷机制被诱导启动表达^[16]。ArsD 是一个伴侣蛋白，作为较弱的第 2 抑制因子调控 *ars* 操纵子，通过将砷转运至 ArsAB 外排泵来增加对砷的抗性^[28]。

嗜酸硫杆菌抗砷功能蛋白与砷结合是抗砷机制中至关重要的一步。研究发现，枯草芽孢杆菌 (*Bacillus subtilis*) 砷酸盐还原酶 ArsC 与砷结合的活性位点存在 Cys10、Cys82、Cys89 以及 Arg16 和 Asp105 必需氨基酸残基^[29]，大肠杆菌 (*Escherichia coli*) 砷酸盐还原酶 ArsC 与砷结合的活性位点存在 Cys12 以及 Arg60、Arg94 和 Arg107 四个残基，在与砷相互作用时，Cys 残基的巯基与砷直接结合^[17]，嗜酸氧化亚铁硫杆菌 (*A. ferrooxidans*) 转录调控因子 ArsR 与 As(III) 结合过程中也有 3 个 Cys 残基参与^[27]。本研究对嗜酸硫杆菌 *ars* 操纵子上表达的 ArsC、ArsD 和 ArsH 抗砷功能蛋白与砷结合位点预测结果中均发现 Cys 残基的存在。此外，我们还发现 Acife 1530 (ArsA) 的 2 个甲硫氨酸 (MET) 残基可能与 As(III) 直接结合，暗示着甲硫氨酸 (MET) 残基中存在与 Cys 残基功能相同的巯基，由此可推测 Cys 残基或 MET 残基的巯基对于构成生物体抗砷功能蛋白与砷结合的活性位点是必不可少的。

4 结论

1) 对挖掘出的嗜酸硫杆菌抗砷基因进行聚类进化分析，发现一些抗砷基因为嗜酸硫杆菌所共有的基因，说明这类基因为嗜酸硫杆菌抗砷性能所必需基因；发现一些抗砷基因为一些菌株所独有，显示出抗砷基因的多样性，有助于这些菌株极端环境适应能力。

2) 对抗砷基因蛋白的理化性质、二级结构、亚细胞定位、信号肽和跨膜区分析，显示出嗜酸硫杆菌抗砷基因具有多样性特征，暗示了嗜酸硫杆菌抗砷机制的复杂性。

3) 对抗砷基因蛋白与砷结合位点分析发现, Cys 残基或 Met 残基的巯基对于构成嗜酸硫杆菌抗砷功能蛋白与砷结合的活性位点是必不可少的。

参考文献:

- [1] RAWLINGS D E. Characteristics and adaptability of iron- and sulfur-oxidizing microorganisms used for the recovery of metals from minerals and their concentrates[J]. *Microbial Cell Factories*, 2005, 4 (1): 13.
- [2] 赖绍师, 覃文庆, 杨聪仁, 等. 高砷原生硫化铜矿细菌浸出试验研究[J]. *矿冶工程*, 2011, 31 (3): 55-58.
- [3] CERVANTES C, JI G, RAMIREZ J L, et al. Resistance to arsenic compounds in microorganisms[J]. *FEMS Microbiology Reviews*, 1994, 15 (4): 355-367.
- [4] JI G, SILVER S. Reduction of arsenate to arsenite by the ArsC protein of the arsenic resistance operon of *Staphylococcus aureus* plasmid p1258[J]. *Proceedings of the National Academy of Sciences U S A*, 1992, 89 (20): 9474-9478.
- [5] SALTNIKOV C W, NEWMAN D K. Genetic identification of a respiratory arsenate reductase[J]. *Proceedings of the National Academy of Sciences U S A*, 2003, 100 (19): 10983-10988.
- [6] 历丽, 夏金兰, 徐爱玲, 等. 抗砷细菌 arsH 基因多样性分析[J]. *现代生物医学进展*, 2010, 10 (9): 1650-1653.
- [7] FISCHER S, BRUNK B P, CHEN F, et al. Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups[J]. *Current Protocols in Bioinformatics*, 2011, Chapter 6, Unit 6.12: 1-19.
- [8] WILKINS M R, GASTEIGER E, BAIROCH A, et al. Protein identification and analysis tools in the ExpASY server[J]. *Methods in Molecular Biology*, 1999, 112: 531-552.
- [9] COMBET C, BLANCHET C, GEOURJON C, et al. NPS@: network protein sequence analysis[J]. *Trends in Biochemical Sciences*, 2000, 25 (3): 147-150.
- [10] KROGH A, LARSSON B, VON HEIJNE G, et al. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes[J]. *Journal of Molecular Biology*, 2001, 305 (3): 567-580.
- [11] PETERSEN T N, BRUNAK S, VON HEIJNE G, et al. SignalP 4.0: discriminating signal peptides from transmembrane regions[J]. *Nature methods*, 2011, 8 (10): 785-786.
- [12] NAKAI K, HORTON P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization[J]. *Trends in Biochemical Sciences*, 1999, 24 (1): 34-36.
- [13] BENKERT P, BIASINI M, SCHWEDE T. Toward the estimation of the absolute quality of individual protein structure models[J]. *Bioinformatics*, 2011, 27 (3): 343-350.
- [14] LAURIE A T, JACKSON R M. Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites[J]. *Bioinformatics*, 2005, 21 (9): 1908-1916.
- [15] KOTZE A A, TUFFIN I M, DEANE S M, et al. Cloning and characterization of the chromosomal arsenic resistance genes from *Acidithiobacillus caldus* and enhanced arsenic resistance on conjugal transfer of ars genes located on transposon TnAtcArs[J]. *Microbiology*, 2006, 152 (Pt 12): 3551-3560.
- [16] BUTCHER B G, RAWLINGS D E. The divergent chromosomal ars operon of *Acidithiobacillus ferrooxidans* is regulated by an atypical ArsR protein[J]. *Microbiology*, 2002, 148 (Pt 12): 3983-3992.
- [17] MARTIN P, DEMEL S, SHI J, et al. Insights into the structure, solvation, and mechanism of ArsC arsenate reductase, a novel arsenic detoxification enzyme[J]. *Structure*, 2001, 9 (11): 1071-1081.
- [18] BHATTACHARJEE H, ROSEN B P. Spatial proximity of Cys113, Cys172, and Cys422 in the metalloactivation domain of the ArsA ATPase[J]. *The Journal of Biological Chemistry*, 1996, 271 (40): 24465-24470.
- [19] SHEN Z, HAN J, WANG Y, et al. The Contribution of ArsB to Arsenic Resistance in *Campylobacter jejuni*[J]. *PLoS One*, 2013, 8(3): e58894.
- [20] VAN ZYL L J, VAN MUNSTER J M, RAWLINGS D E. Construction of arsB and tetH mutants of the sulfur-oxidizing bacterium *Acidithiobacillus caldus* by marker exchange[J]. *Applied and Environmental Microbiology*, 2008, 74 (18): 5686-5694.
- [21] MATEOS L M, ORDONEZ E, LETEK M, et al. *Corynebacterium glutamicum* as a model bacterium for the bioremediation of arsenic[J]. *International Microbiology*, 2006, 9 (3): 207-215.
- [22] 车媛媛. 嗜铁钩端螺旋菌抗砷基因簇的生物信息学分析及改造[D]. 济南: 山东大学, 2010: 21-69.
- [23] TUFFIN I M, DE GROOT P, DEANE S M, et al. An unusual Tn21-like transposon containing an ars operon is present in highly arsenic-resistant strains of the biomining bacterium *Acidithiobacillus caldus*[J]. *Microbiology*, 2005, 151 (Pt 9): 3027-3039.
- [24] HEDGES R W, BAUMBERG S. Resistance to arsenic compounds conferred by a plasmid transmissible between strains of *Escherichia coli*[J]. *Journal of Bacteriology*, 1973, 115 (1): 459-460.
- [25] BUSENLEHNER L S, PENNELLA M A, GIEDROC D P. The SmtB/ArsR family of metalloregulatory transcriptional repressors: Structural insights into prokaryotic metal resistance[J]. *FEMS Microbiology Reviews*, 2003, 27 (2/3): 131-143.
- [26] HODAR C, MORENO P, DI GENOVA A, et al. Genome wide identification of *Acidithiobacillus ferrooxidans* (ATCC 23270) transcription factors and comparative analysis of ArsR and MerR metal regulators[J]. *Biometals*, 2012, 25 (1): 75-93.
- [27] QIN J, FU H L, YE J, et al. Convergent evolution of a new arsenic binding site in the ArsR/SmtB family of metalloregulators[J]. *The Journal of Biological Chemistry*, 2007, 282 (47): 34346-34355.
- [28] LIN Y F, YANG J, ROSEN B P. ArsD: an As(III) metallochaperone for the ArsAB As(III)-translocating ATPase[J]. *Journal of Bioenergetics and Biomembranes*, 2007, 39 (5/6): 453-458.
- [29] BENNETT M S, GUAN Z, LAURBERG M, et al. *Bacillus subtilis* arsenate reductase is structurally and functionally similar to low molecular weight protein tyrosine phosphatases[J]. *Proceedings of the National Academy of Sciences U S A*, 2001, 98 (24): 13577-13582.

Diversity analysis of arsenic resistance genes in *Acidithiobacillus* genus

ZHAO Yonghong, ZHANG Jing, ZHOU Dan, PENG Yanping, YU Shuijing

School of Resources and Environment, Jiangxi University of Science and Technology, Jiangxi Key Laboratory of Environmental Pollution Control of Mining and Metallurgy, Ganzhou, Jiangxi 341000, China

Abstract: To explore the diversity of arsenic resistance genes in *Acidithiobacillus* genus, 49 arsenic resistance genes were mined by bioinformatics method from species of *A. ferrooxidans*, *A. caldus*, and *A. ferrivorans*. These 49 genes are *ars*-related including *arsA*, *arsB*, *arsC*, *arsD*, *arsH* and *arsR*, which form the arsenic-resistance operon. The characteristic of arsenic resistance genes were analyzed on basis of their clustering and evolution, physicochemical properties, secondary structure, transmembrane regions, subcellular localization, and arsenic binding sites. Results showed that all 49 arsenic resistance genes were clustered into 11 groups, of which the *arsB* gene belonging to the second group was common with *Acidithiobacillus* genus. The *arsB* gene, coding for the arsenical pump membrane protein (ArsB) may be essential to the arsenic resistance of *Acidithiobacillus* strains. The clustering and evolution analyses showed that five arsenic resistance genes may originate from the horizontal gene transformation. Aliphatic index analysis showed that 25 arsenic resistance proteins were hydrophilic and the other 24 were hydrophobic. The instability indices revealed that most of ArsRs were unstable. The secondary structures of corresponding 49 arsenic resistance proteins were comprised of α -helix, random coil, and extended chain. Subcellular localization analysis implied that arsenic resistance proteins might exist in inner and outer membrane, periplasmic space, and cytoplasm. ArsB proteins contained 10-11 transmembrane regions, and no signal peptide was found in all arsenic resistance proteins. Based on the arsenic binding site prediction and modeling results, we found that the Cys residues occurred in all active sites of ArsC, ArsD and ArsH proteins. Compared with other arsenic resistance proteins, the active sites of protein Acife1530 (i.e. ArsA) possessed two Met residues binding As (III), implying that the thiol probably existed in Met and played a same role in binding As (III) directly as that in Cys residue. Therefore, the thiol in Cys or Met residue is vital for the arsenic binding site of arsenic resistant proteins in *Acidithiobacillus* microorganisms.

Key words: *Acidithiobacillus*; arsenic resistance gene; ortholog analysis; diversity analysis; bioinformatics