

高效产氢菌 B49 菌株 *adh* 和 *L-ldh* 基因 克隆及序列分析

林海龙 任南琪* 郑国香 张 坤

(哈尔滨工业大学市政环境工程学院 哈尔滨 150090)

摘 要: 设计细菌 *adh* 基因和 *L-ldh* 基因简并引物, 采用降落 PCR 并结合 Cassette PCR 方法从高效产氢菌 B49 中扩增全长基因。共获得两段基因组 DNA 片段。其中一段序列长 1902 bp, 包括 *adh* 基因开放阅读框 1101 bp, 共编码 366 个氨基酸, 编码产物分子量为 39.71 kD, 理论等电点为 pH 5.93。该基因与 *Clostridium thermocellum* ATCC 27405 的 *adh* 位点序列一致性为 73%。另一段序列长 2490 bp, 包括 *L-ldh* 基因开放阅读框 951 bp, 共编码 316 个氨基酸, 编码产物分子量为 34.23 kD, 理论等电点为 pH 6.09。该基因与 *Bacillus megaterium* 的 *L-ldh* 位点一致性为 74%。试验结果表明, 采用 CodeHop 和 Genefisher 程序化设计的简并引物可信性强, 阳性率高, CodeHop 设计简并引物效果要好于 Genefisher。*adh* 和 *L-ldh* 的成功克隆为通过代谢工程手段敲除 *adh* 和 *L-ldh* 基因提供了科学依据, 从而为进一步提高 B49 产氢能力的代谢工程研究奠定基础。

关键词: 引物设计, 乙醇脱氢酶, L-乳酸脱氢酶, *Ethanoligenens harbinense* B49

Clone the Alcohol Dehydrogenase and L-lactate Dehydrogenase of Bio-hydrogen Producing Bacteria *Ethanoligenens harbinense* B49 and Its Sequence Analysis

LIN Hai-Long REN Nan-Qi* ZHENG Guo-Xiang ZHANG Kun

(School of Municipal and Environmental Engineering, Harbin Institute of Technology, Harbin 150090)

Abstract: Designed the degenerate primers of alcohol dehydrogenase and L-lactate dehydrogenase to augment *Ethanoligenens harbinense* B49 genomic DNA, and obtained about 780 bp and 610 bp PCR product respectively. Augmented flank sequences of the two PCR fragments with the Cassette PCR method. Similarity alignment showed that the products of the cloned DNA were very high similar to those of alcohol dehydrogenase genes and L-lactate dehydrogenase genes respectively. One of the two sequences was 1902 bp long, and the ORF of *adh* was 1101 bp long and encoded 366 amino acids. Its putative molecular weight was about 39.71 kD, its calculational isoelectric point was pH 5.93. The maximal identity and positive was 51% and 73% with *Clostridium thermocellum* ATCC 27405 *adh*. The other one was 2490 bp long, and the ORF of *adh* was 951 bp long and encoded 316 amino acids. Its putative molecular weight was 34.23 kD, its calcula-

tional isoionic point was pH 6.09. The maximal identity and positive was 55% and 74% with *Bacillus megaterium* *L-ldh*. Successfully cloning these two genes would not only enrich the gene resources of L-lactate dehydrogenase and alcohol dehydrogenase genes, but also give the scientific warrant for the metabolic engineering research and the construction of the gene-engineering bacteria.

Keywords: Design primer, Alcohol dehydrogenase, L-lactate dehydrogenase, *Ethanoligenens harbinense* B49

近年来由于生物制氢技术可以克服常规制氢方法的诸多缺陷而倍受关注, 其在新能源开发中的地位也越来越重要, 世界上已有许多国家投入了大量的人力物力对生物制氢技术进行开发研究, 以期早日实现其产业化^[1]。生物制氢技术按微生物产氢机制分为两种类型, 光合法生物制氢与发酵法生物制氢。与光合法生物制氢相比, 发酵法生物制氢有着无需光照、产氢稳定等优点, 具有更为广阔的前景^[2]。

为了实现发酵法生物制氢的产业化, 国内外研究者分离了大量的高效产氢发酵菌种^[3-10]。其中产氢菌 *Ethanoligenens harbinense* B49 是本实验室从生物制氢反应器的乙醇型发酵活性污泥中分离出的一株高产氢乙醇型发酵菌株, 其酵解葡萄糖的主要发酵产物为乙醇、乙酸、 H_2 、 CO_2 和少量乳酸^[2], 它是新发现的乙醇型发酵产氢菌属—*Ethanoligenens* 属^[6]的典型代表菌种之一。

乙醇和乳酸都是 *E.harbinens* B49 发酵的末端产物。根据现有的 NADH 产氢代谢途径理论^[11,12] 和已知厌氧菌的中心代谢途径, Chen 等^[13] 从代谢工程的角度, 运用化学平衡计量的方法对厌氧菌的最大理论产氢量进行了分析。研究结果表明: *Clostridium butyricum* (典型的厌氧产氢菌) 在厌氧状态和葡萄糖最终都代谢为乙酸的情况下, 其最大理论产氢量为 3.26 mol H_2 /mol 葡萄糖。而根据 NADH 产氢理论, 乙醇和乳酸的生成会消耗 NADH, 将减少用于生成氢气的 NADH 量, 从而导致氢气量的减少。众所周知, 葡萄糖在生成乙醇的过程中, 乙醇脱氢酶是控制乙醇生成的关键酶。而乳酸脱氢酶 (*ldh*) 则是丙酮酸生成乳酸的关键酶。因此, 若能克隆到 *E.harbinense* B49 生成乙醇和乳酸的相关酶基因, 则可以认为通过基因敲除的手段阻断乙醇和乳酸的生成提供科学依据, 从而为提高 B49 的产氢能力奠定基础。

本实验采用 ClustalX^[14,15] 比对分析软件和 CodeHop^[16-20] 在线引物设计软件设计简并引物同源克隆乙醇脱氢酶基因。采用 Giegerich 等^[21] 开发的 Genefisher 软件设计简并引物同源克隆 L-乳酸脱氢酶基因。

1 材料与方法

1.1 材料

1.1.1 产氢菌 B49: 2002 年林明从产氢反应器的厌氧活性污泥中分离到高效产氢菌 B49, 其 16S rDNA 序列在 EMBL 核酸数据库中注册号为 AF481148, 在中国微生物菌种保藏中心的保藏号为 CGMCC1153。采用林明提出的产氢菌富集培养方法^[22] 获得 B49 的纯培养物。

1.1.2 实验材料: 大肠杆菌 DH5 α 本实验室保存, pMD18-T 载体购自宝生物工程(大连)有限公司。细菌基因组 DNA 提取纯化试剂盒、小量胶回收试剂盒购自上海华舜生物工程有限公司, TaKaRa LA PCRTM *in vitro* cloning Kit 购自于宝生物工程(大连)有限公司, 引物由上海生工生物工程技术有限公司合成。Taq 酶和 dNTP 均购自宝生物工程(大连)有限公司。

1.2 方法

1.2.1 乙醇脱氢酶和 L-乳酸脱氢酶部分片段的克隆: 在 NCBI 蛋白质数据库中, 挑选了 2 个梭菌属 *Clostridium thermocellum* ATCC 27405 和 *Clostridium* sp. *OhILAs* 菌种的乙醇脱氢酶蛋白质序列。将这 2 个序列提交 CodeHop^[17] 在线引物设计网站设计简并引物。经过筛选后选用了一对引物, 命名为 SADHJ (见表 1)。PCR 反应条件为: 预变性 95 $^{\circ}C$, 5 min; 94 $^{\circ}C$, 30 s, 64 $^{\circ}C$ 降到 55 $^{\circ}C$, 每次降 0.3 $^{\circ}C$, 40 s, 72 $^{\circ}C$, 30 s, 共 30 个循环; 然后 94 $^{\circ}C$, 30 s, 55 $^{\circ}C$, 40 s, 72 $^{\circ}C$, 2 min, 共 10 个循环; 最后在 72 $^{\circ}C$ 延伸 10 min。

以 B49 菌株的 16S rDNA 序列为递交序列, 用 Blast2.0 进行 16S rDNA 同源性比对, 获得与 B49 相似性较高且在 NCBI 蛋白质数据库中存在其乳酸脱氢酶蛋白质序列的菌种。它们是 *Clostridium acetobutylicum* CAC0267, *Clostridium acetobutylicum* CAC3552, *Clostridium perfringens* CPE0103, *Clostridium tetani* E88 CTC01998, *Clostridium thermocellum*, *Thermoanaerobacter saccharolyticum*, *Moorella thermoacetica* ATCC 39073 共 7 个菌种的乳酸脱氢酶蛋白质序列。把这 7 个蛋白质序列输入到 Genefisher 在线软件中, 选用 ClustalX 方法进行比对分析, 并对引物长度(18 bp~30 bp)、引物 GC 含量(35%~85%)、溶解温度(35 ~75)、PCR 产物长度(90 bp~900 bp)、3 夹处 GC 含量(35%~85%)、引物最大简并度(5000)、3 夹最大碱基数(4 bp)、3 夹最大简并度(5000)、末端残基(A)、序列的唯一性(1)进行

设定。运行后, 共得到 43 对得分最高的引物。经过筛选后选用了一对引物, 命名为 LDHJ(见表 1)。简并 PCR 反应条件为: 预变性 95 , 5 min; 94 , 30 s, 62 降到 50 , 每次降 1 , 40 s, 72 , 2 min, 共 12 个循环; 然后 94 , 30 s, 50 , 40 s, 72 , 2 min, 共 23 个循环; 最后在 72 延伸 10 min。

1.2.2 乙醇脱氢酶和 L-乳酸脱氢酶部分片段两侧序列的克隆: 采用 Oligo6.0 软件设计克隆 *adh* 片段两侧序列的引物(Sadhs11、Sadhs21、Sadhs12、Sadhs22)见表 1, 克隆 L-*ldh* 脱氢酶片段两侧序列的引物(LDHS11、LDHS12 和 LDHS21、LDHS22)见表 2。基因组 DNA 用已克隆片段中没有的限制性内切酶(如 *EcoR* , *Hind* 、 *BamH* 和 *Xba* 等)酶切后加相应的接头进行 PCR 扩增。PCR 条件参见说明书, 但根据引物 Tm 值不同, 退火温度也稍有变动。

表 1 用于克隆乙醇脱氢酶片段的引物

Table 1 The primers used to clone alcohol dehydrogenase

Primers	Sequences	Degeneracy
SADHJF	5-GATCTACTTCGGCAGAACTGCRTNMARRARAA-3	128
SADHJR	5-GATGTTGTGGAAGTAGGTGTAGCARTANCCCATNCC-3	32
Sadhs11	5-GCGAGATAGCCTTCCAGCGAGTGCCTGAAAGCGTC-3	
Sadhs12	5-AAACGGTTTTGGAATAACTCTTCGGGCTGCATGTCGTTG-3	
Sadhs21	5-GAGGTCACACCCTACAGTGTATCCTGCGCCGCGACCTGC-3	
Sadhs22	5-GTGGACGCTTTCACGCACTCGCTGGAAGGCTATCT-3	
SadhwF	5-GCGCCGTAACGGCCGGCTGGAATAGTTGAG-3	
SadhwR	5-CACTCGCCAGGGCCGGACGCTGTGAAGCGG-3	

表 2 用于克隆 L-乳酸脱氢酶片段的引物

Table 2 The primers used to clone L-lactate dehydrogenase

Primers	Sequences	Degeneracy
LDHJF	5-GCNRANGGNGARGCNHTGGA-3	3072
LDHJR	5-RBANACRTCNTTDTATNCCRTA-3	4068
LDHS11	5-CACATACAGGTTGTCCGGCTTTGC-3	
LDHS12	5-GCGACTATTTCAAGGTGGACCCGCGCAACGTG-3	
LDHS21	5-TTGCAAAGCCGGACAACCTGTATGTGACGTAG-3	
LDHS22	5-CCCGCGTAGATTTTCATGTGGTGC-3	
LDHwF	5-CGTCCTCCAGGTTCCATTGGCGGCTCTGGATGATCTCGG-3	
LDHwR	5-CTCCACGGCATCAAACCCCGCTCTGCGGCGCGCTTGGTTCG-3	

1.2.3 乙醇脱氢酶和 L-乳酸脱氢酶的基因全长克隆: 根据拼接的乙醇脱氢酶全长序列, 在 485 bp 和 2387 bp 各截取一段序列作为引物 SadhwF、SadhwR, 其具体序列见表 1。PCR 反应条件: 94 变性 5 min; 94 变性 30 s, 66 复性 30 s, 72 延伸 2 min, 35 个循环; 72 延伸 10 min。

根据拼接的 L-乳酸脱氢酶全长序列, 设计了用

于全长验证的引物 LDHwF 和 LDHwR, 见表 2。其 PCR 反应条件如下: 94 变性 5 min; 94 变性 30 s, 69 复性 30 s, 72 延伸 2 min, 35 个循环; 72 延伸 10 min。

1.2.4 PCR 产物的纯化、克隆和检测: 目标条带切下, 胶回收操作方法见华舜小量胶回收试剂盒说明书。克隆过程参照宝生物工程(大连)有限公司

pMD18-T载体说明书进行。PCR产物与pMD18-T载体连接, 转化大肠杆菌DH5 α 。转化产物 37 过夜, 经蓝白斑筛选, 挑选白色菌落, 进行菌落PCR验证得到阳性克隆, 取两个阳性克隆送交上海生工生物工程技术有限公司测序。测序的克隆片段经NCBI上去除载体序列后, 输入GenBank中进行同源性比对分析。

2 结果

2.1 简并 PCR 电泳结果

乙醇脱氢酶简并 PCR 电泳结果见图 1-A。PCR片段大小在 750 bp 到 1000 bp 之间, 更靠近 750 bp, 大约 780 bp 左右。这与我们预期片段大小相一致。经测序、比对后与乙醇脱氢酶的相似性较高。而且 PCR 结果仅为一条带, 没有其它杂带。其基因全长 PCR 电泳结果见图 1.B。L-乳酸脱氢酶简并 PCR 电泳结果见图 2。共得到 4 个 DNA 条带, 其中处于 500 bp~750 bp 之间的 DNA 条带最亮, 并与目标条带大小相符(大约 600 bp), 其余条带可能是非特异性扩增的结果。在克隆 *L-ldh* 时, 采用 Genefisher 设计的上、下游简并引物的简并度分别为 3072 和 4068, 引物长度分别为 20 bp、21 bp(见表 1), 也成功地扩增到 *L-ldh* 基因片段, 但是非特异性扩增较多。而 CodeHop 设计的简并引物简并度分别为 128 和 32, 引物长度分别为 33 bp、36 bp(见表 1), 成功地扩增到 *adh* 基因片段, 且特异性非常好, 在使用降落 PCR 的情况下只扩增出一条带。

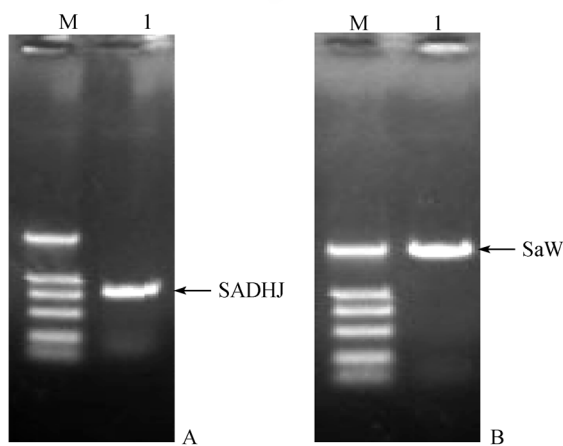


图 1 乙醇脱氢酶部分片段及全长基因扩增的电泳结果
Fig. 1 Agarose gel analysis of *adh* fragment and whole *adh* gene

A: 部分片段扩增结果; B: 全长基因扩增结果

A: Agarose gel analysis of *adh*; B: Agarose gel analysis of whole *adh* gene

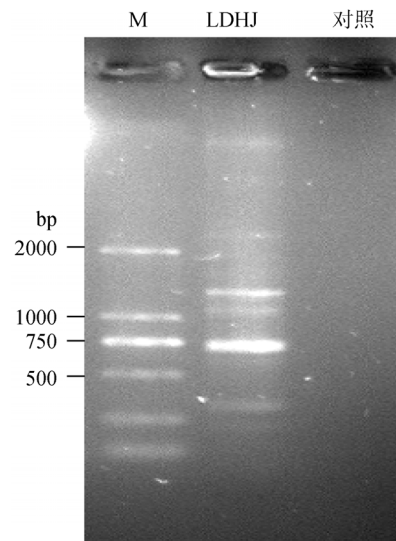


图 2 简并引物扩增 *L-ldh* 基因保守区的电泳图

Fig. 2 Agarose gel analysis of *L-ldh* fragment from *Ethanologens harbinense* B49 fragment

2.2 乙醇脱氢酶及 L-乳酸脱氢酶基因全长克隆

在克隆 L-乳酸脱氢酶基因全长中, 共获得长 2490 bp 的 DNA 序列, 该序列在 GenBank 注册号为 DQ179104。其中 L-乳酸脱氢酶基因的开放读码框 951 bp, 在 991 bp 处有起始密码子 ATG, 在 1941 bp 处有终止密码子 TAA, 该序列无内含子, 共编码 316 个氨基酸。编码产物分子量为 34.23 kD, 理论等电点为 6.09。其开放读码框及推导的氨基酸序列如图 3。

乙醇脱氢酶基因全长 PCR 扩增出符合预期全长序列大小(1900 bp)的特异 DNA 条带。其测序结果如图 4。其在 GenBank 注册号为 EU196512。该乙醇脱氢酶基因的开放阅读框位于第 352 bp 到第 1452 bp 之间, 共 1101 bp 长, 其中 A 碱基 242 个, T 碱基 204 个, C 碱基 348 个, G 碱基 307 个, A+T 含量为 40.51%, C+G 含量为 59.49%。起始密码子 ATG 位于第 352 bp 到第 354 bp, 终止密码子 TGA 位于第 1450 bp 到第 1452 bp。在起始密码子-7 bp 处存在一个预测的核糖体结合位点 AGCAGG(第 339 至第 344), 它与 16S rRNA 3 端互补。在起始密码子之前, 即第 297 bp 到第 346 bp 处存在一个预测的启动子(CTG TTGATTGGCTGCGCACACCTCGTCATTTTATTTT CAAAAAGCAGGTT)。在终止密码子 TGA 之后, 即第 1478 bp 到第 1507 bp 处存在一个预测的终止子(GGCCGGTACCCGTGAAGACGGGCG CCGGCC)。

```

1 ATGTGCACCGACAACCGCAAAGTCGTGCTGGTAGCCACGGACTGGTGGGCATGAGTTTGCCTACGCCCTGCTCAACCAGCAGCATGTGACGAACGGTCTC
1 M C T D N R K V V L V G T G L V G M S F A Y A L L N Q H A C D E L V L
106 ATTGATATCAACAAACAGCGTGTGAGGGTAGGGCGATGGACCTCAACCATGGTCTGGCGTTTTCCGGCACCACATGAAATCTACGGGGCGATTACAAGGAC
36 I D I N K Q R A E G E A M D L N H G L A F S G T N M K I Y A G D Y K D
211 TGGCCGACGCGGCATCGTGGCCATCTGCGCGGGCGTGGCGCAGAAACCGGGCGAGCCGATGGATCTGCTGACGGCAACACCGCGTGTCAAAATCCATC
71 C A D A D I V A I C A G V A Q K P G E S R M D L L Q R N T A V F K S I
316 GTGGACCGGTGGTGGCTTCGGGATTTTCCGGGTGTTCTTGTGGCCACCAATCCGGTGGACATCATGCTCTACGTACATACAGGTTGTCGGCTTTGCAAAA
106 V E P V V A S G F S G V F L V A T N P V D I M S Y V T Y R L S G F A K
421 GGCCGATCGTGGCCACCGCACCGCTGGACACCGCCCGCTCGCTACCTGCTGGCGACTATTTCAGGTGGACCCCGCAGCAACGTGCATGCTTATGTGATG
141 G R I V G T G T T L D T A R L R Y L L G D Y F K V D P R N V H A Y V M
526 GCGGACGCGGCACGTGAGTTTGGCCGTTGCGAGGCGCTATAGCCACCCCGCGTATGGGGCTTTCGCTGAAAAACCAGCCCGGACTATAAAGCC
176 G E H G D S E F V P W S Q A L I A T R P V M G L C V E N H G P D Y K A
631 GGCATGCTCCACATCGGTGAGGATCCGACCGCTGCCACCGCATCGAAGCCAAAAAGCCACTATTACGGGATCGGCATGGCCATGTTGCGGTGGCC
211 G M L H I G E E V R T A A Y R I E A K K A T Y Y G I G M A M V R V M
736 CGCGCCATTCGCGCGGCAAAACAGCGTGTACCGTTTCTCGTGTGCGACGACTACGGCACCCCAAGTCTATGCCGCGTCCGCTCCATCGTCAGC
246 R A I L G G E N S V L T V S S L L D D D Y G T P K V Y A G V P S I V S
841 CGGGCGGGCGTACGCGTATCATTGCGTTTCGCTCACACCGGAAGAAAATCAGTGTGCAAGATTCTGTGCCAACTGGAGCAGCCGCTCCCGGTTGACA
281 R R G V S R I I R L S L T P E E N Q L M Q D S C A K L E Q A V S G L T
946 TTTTAA
316 F *

```

图3 L-乳酸脱氢酶基因 DNA 序列和对应的氨基酸序列

Fig. 3 The DNA and amino acid sequence of the L-lactate dehydrogenase

```

1 TGCCCGTAAACGGCGCTGGGAATAGTTGAGTCTCCCGCGTTTTGTAAATTCAGGCAGCGGGATACATACAAAAATGTTTTAGATGTTTCAGTTCACATA
106 AACCTTTTCCCGACGCGCCCTTCCCGTGCCCGGAACCATATGCACCTATCATAACATGTTTTCCGGCAGCAGCGGCATCGCGCAGGATCGTCAATCAC
211 AAAGTCTGATGAAACCATCATCTCTCGCGATTTTCTTTTACGAAGGATATGGTATCCTATAGAAAGGATAGTCCGGGTTTTGCCTGTTGATGGCTGCGCAC
316 ACCTCGTCAATTTATTTCAAAAAGCAGGTTTTTATATGAAGTATCAGTATTTATGCCACAAGGCTCGCGTTCGGACGAAACTGCATACAGAAAAACCGCGAC
1 M K Y Q Y F M P T R L A F G R N C I Q E N A D
421 CGGTGGCCGATTCGGTAGCCGCGCTTCAATTTGACAGGCGGCATTCGCCAAGGCTGCGGTGCGCTGGATGATGTCACCGCGCTCTCGAGAGCGGGAT
24 R L A A F G S R A F I L T G R H S A K A C G A L D D V T A A L E S R D
526 ATTGCTTACCAGGTC TTCAATGAAATCGAAAGCAATCCGCTCGTTGAAACCGTCTGAAAGCCGGCACAATCGCAAAAGCATTGCGCGCGCAATTGCTCATCGGG
59 I A Y Q V F N E I E S N P S L E T V V K A G T I A K A F G A E F V I G
631 ATCGCGGGGGTCTCCTCTGGACGCGCGGAAGGCGCTCGTGTCTCAGGCAACGACATGCAGCCCAAGAGTTATTCAAAAACCGTTTTCCCAAGCGCGCG
94 I G G G S P L D A A K A V A V L T A N D M Q P E E L F Q N R F P K R P
736 CTCCGATCGTGGCGTTCCACCACCGCGGACAGGCACTGAGGTCACACCTACAGTGCATCCTGCGCCGCGACTGACAGCAAGGTCAGTCTCGGCACC
129 L P I V G V P T T A G T G S E V T P Y S V I L R R D L Q T K V S F G T
841 CCCGAAACGTTTCCGCTCTCGCGTGTGGACAGCGGTACACCGCACAACCTGCCCGCACCGTAACGATGGACACTGCGGTGGACGCTTTCAGCACTCGGTG
164 P E T F P S L A L L D S R Y T A Q L P R T V T M D T A V D A F T H S L
946 GAAGGCTATCTCCCGCGGAGCACCAGGCTCAGCAGTACTGGACATGGAAGCATCCGGCTCTCGGTGAGTGTGCCAACTGCTCTCGTTGAAATT
199 E G Y L A R R S T A L S S V L A L E G I R L F G E V L P N L L S F E I
1051 GACGACCGTGGCGACAAACTGCTTACGCTCCACCGTCCGCGGCATCGTATCGCGCAGCCGCGTCAACATCGACATGGCATGGGTATTGCTACACC
234 D D T V R D K L L H A S T V G G I V I A Q A G V T I A H G M G Y C Y T
1156 TTTTCAAACACGCGCACGGCGTCCCAACGGTTCTTATCGGGGAATACCTCAAGTCAACCAGCCCGCGCGCAGAGAAGATGCACGCGCCCTGAAG
269 F F K N T P H G R A N G F L M R E Y L K F N Q P A A A E K I D T A L K
1261 CTGTTGGCATGGCGACATAGACGCGTTCGGGATGTCTGGAACGTCTGATCGCAAAGCCCGTGTGACATCGGACGAAGTGGATTCTTCACGGAAGT
304 L L G M A D I D A F G D V L E R L I G K A P V L T S D E V D S F T E L
1366 ACAGTCTCAGAAAGGAGCAGTCCCAACAACGACGCGGAGCTTCCAAAGAAGATATCCATCGGCTGTGAAAAACAACGCGTGCACCGCTTCATCAAGCATAT
339 T L L Q K G S I A N N A R E L S K E D I H R L W K N N A *
1471 TCCTATCGGCGGTACCGTGAAGACGGCGCGCGCTTTCGCGTTCGCGCAAAAGAGCGTTCGCGCGCGCGCGTTCCTTTGGTTTAAACCGGTGATACCGTTCA
1576 GATAGTCTTCGGTGAATCCACCGTGGCAATCACACCGCTTTTCAAGCAGGATTCGTCGATGTCGAACTTCGGTTATGCTGCGCGCGCGGTGCCAACCGTT
1681 GGTTCGCGACCCCGAAGAACATAAACACCGGGATACCGTCCAGATACTGTGCCATACTGTCGACCCCGCAATCGCGGCATGGAAGCCAGGACGCGTTTT
1786 CGAGCCGACGGGTGATCGATGCGGACGCAACCGCTGACGGCTGCATCGTTGACAGTGGGAATGCCCTGCCTGATAGGTGATTTCCGCTTCACAGCGTCCGG
1891 CCCTGGCGAGTG

```

图4 乙醇脱氢酶基因 DNA 序列和对应的氨基酸序列

Fig. 4 The DNA and amino acid sequence of alcohol dehydrogenase

该乙醇脱氢酶共编码 366 个氨基酸, 其预测蛋白质分子量为 39.71 kD, 理论等电点为 pH 5.93。

2.3 乙醇脱氢酶和 L-乳酸脱氢酶基因的保守性分析

通过对包括 *E.harbinense* B49 乙醇脱氢酶在内的 8 个细菌乙醇脱氢酶蛋白质序列进行的比对分析(结果见图 5), 发现它们共同存在 G⁴⁹, L⁵⁸, P⁷⁵, P¹⁷⁶, D¹⁸², R²⁴⁶, G²⁵⁷, H²⁶⁹, H²⁸⁴, G²⁸⁵ 共 10 个比较保守氨基酸位点, 以及(G⁹⁹GGXDXK¹⁰⁸), (T¹⁴⁵XGTXE¹⁵¹), (T¹⁹⁷XXDAXHXE²⁰⁷) 共 3 个保守区域。这些位点与区域可能参与该乙醇脱氢酶蛋白的空间折叠、底物结合和酶催化等酶活性功能。其中包含 1 个核酸结合位点保守区 GXGXG(GIGGG, 93-97) 和 1 个

GXGXXVXXXA 保守区(GTGSEVPYSV, 134-144)^[23], GXGXXVXXXA 保守区可能涉及到 NAD(H)-NADP(H) 辅酶的结合, 虽然 *E.harbinense* B49 的 *adh* 蛋白中是不保守的 V, 而不是保守的 A, 但是这种保守位点的取代作用并没有改变结构的变化。

通过比对包括 *E.harbinense* B49 在内的共 9 个细菌 *L-ldh* 蛋白质序列分析了 *E.harbinense* B49 乳酸脱氢酶的保守区域和氨基酸活性位点(见图 6)。在 *E.harbinense* B49 的 L-乳酸脱氢酶中共找到 5 个保守区域: 在氨基端存在 1 个 NAD 结合位点(即 GXGXG 保守区), 包含了 3 个严格保守的甘氨酸残基。位置在第 16 到第 21 个氨基酸之间, 序列为

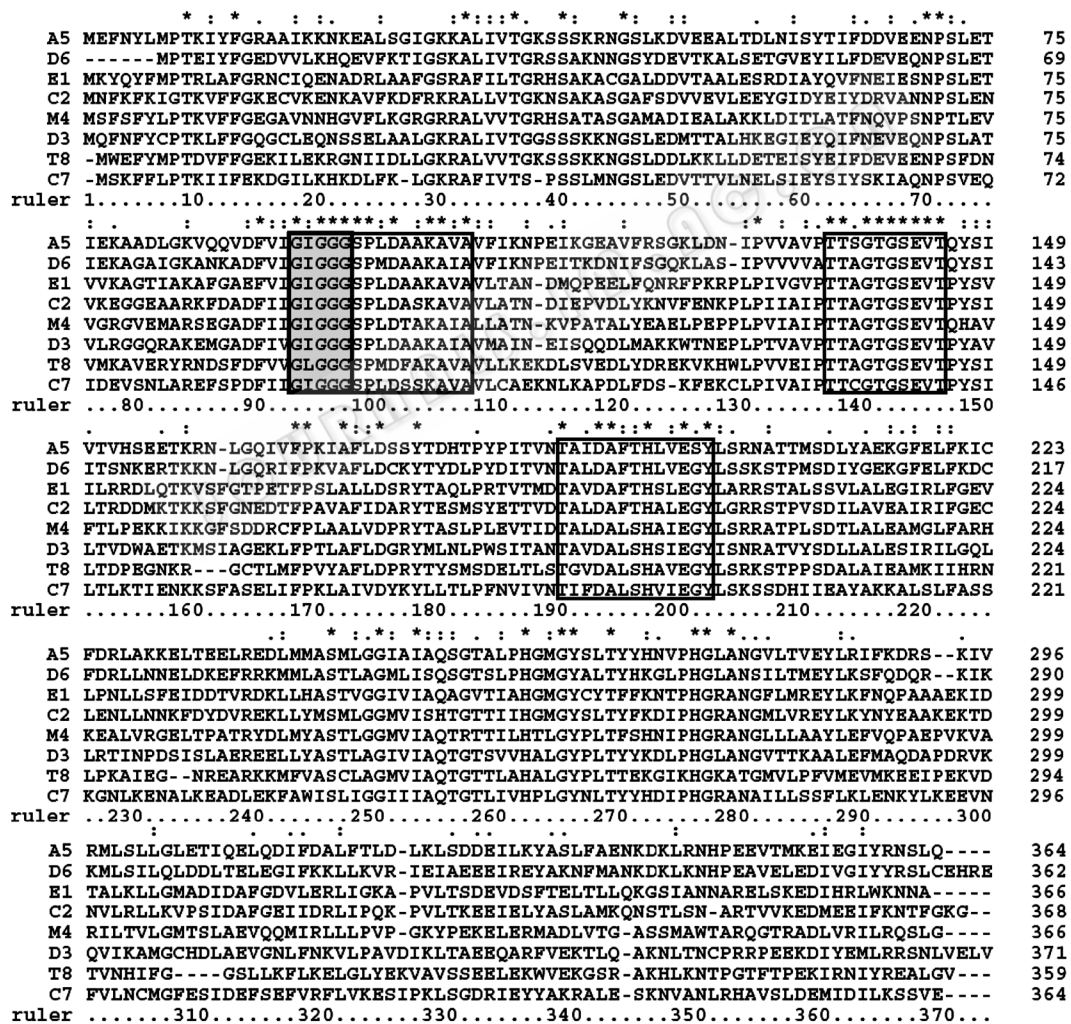


图 5 *E.harbinense* B49 的乙醇脱氢酶保守区分析

Fig. 5 The conservation analysis of *adh* from *E.harbinense* B49 with other bacteria *adhs*

E1: *Ethanoligenens harbinense* B49 (PID EU196512); C2: *Clostridium thermocellum* ATCC 27405 (PID YP_001038972); D3: *Desulfotomaculum reducens* MI-1 (PID YP_001114005); M4: *Moorella thermoacetica* ATCC 39073 (PID YP_429881); A5: *Alkaliphilus oremlandii* OhILAs (PID YP_001512177); D6: *Desulfitobacterium hafniense* DCB-2 (PID ZP_01373007); C7: *Caldicellulosiruptor saccharolyticus* DSM 8903 (PID YP_001180284); T8: *Thermotoga maritima* MSB8 (PID NP_228728)

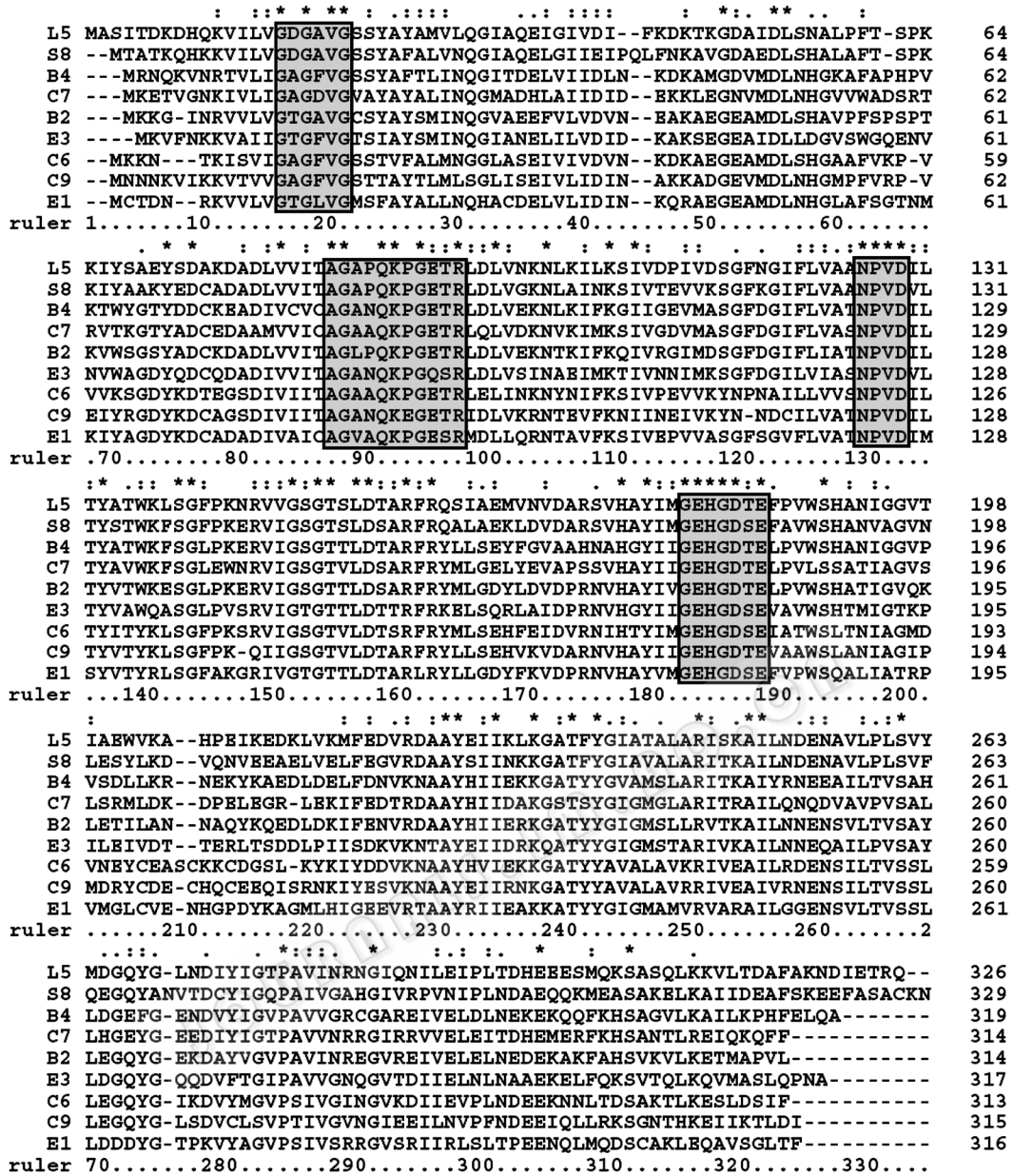


图 6 *E. harbinense* B49 的 L-乳酸脱氢酶保守区分析

Fig. 6 The conservation analysis of L-ldh from *E. harbinense* B49 with other bacteria L-ldhs

E1: *Ethanoligenens harbinense* B49 (PID DQ179104); B2: *Bacillus cereus* AH187 (PID ZP_02255649); E3: *Enterococcus faecalis* V583 (NP_814395); B4: *Bacillus licheniformis* ATCC 14580 (YP_077608); L5: *Lactobacillus casei* ATCC 334 (YP_807714); C6: *Clostridium acetobutylicum* ATCC 824 (NP_346908); C7: *Corynebacterium glutamicum* R (YP_001139733); S8: *Streptococcus bovis* (Q59828); C9: *Clostridium thermocellum* (AAM29186)

GTGVLG^[24,25]。 1 个催化 loop 区域 (AGX₄PG [X₂/X₇]_R, 见图 7), 包含 1 个保守的精氨酸残基。位置在第 87 到第 97 个氨基酸之间, 序列为 AGVAQKPGESR。可能参与质子供体和受体转换时的结构变化^[26]。 严格保守的多肽 NPVG, 包含 1 个保守的天冬酰胺残基 (可能的 NAD 或底物结合位点)。位置在第 129 到第 132 个氨基酸之间, 序列为

NPVG。 GXGTXLDXXR 保守区域, 包含 1 个严格保守的精氨酸残基 (可能的底物结合位点)。位置在第 151 到第 160 个氨基酸之间, 序列为 GTGTTLDTAR。 1 个高度保守的 GEHGDSE 多肽, 包含 1 个保守的组氨酸残基 (可能的质子供体或受体)。位置在第 182 到第 188 个氨基酸之间, 序列为 GEHGDSE^[25]。

```

87                               97
A G F T K A P G K S D K E W N R
A G A N Q K P G E T . . . . . R
A G V A Q K P G E S . . . . . R
A G . . . . P G . . . . . R

```

图7 L-乳酸脱氢酶的催化loop区域

Fig. 7 Protein sequence alignment in the catalytic loop region

2.4 乙醇脱氢酶和 L-乳酸脱氢酶基因的同源性分析

获得的乙醇脱氢酶序列以检索工具 Blastx 检索 GenBank 数据库进行相似性比较。由表 3 可知, 克隆的菌株 B49 所编码的氨基酸序列与 *Desulfotomaculum*、*Moorella*、*Thermotoga*、*Clostridia* 和 *Eubacterium* 属一些菌种的乙醇脱氢酶有高度相似性, 其中与 *Clostridium thermocellum* ATCC 27405 的

乙醇脱氢酶序列一致性达 51%, 位点序列一致性为 73%; 与 *Clostridium beijerinckii* NCIMB 8052 乙醇脱氢酶序列一致性达 30%, 位点序列一致性为 48%。通过序列比对分析表明, 这段 DNA 为菌株 B49 的另一个新乙醇脱氢酶的基因。该核酸序列在 NCBI 上的注册号为 EU196512。

B49 的 L-乳酸脱氢酶以 Blastx 检索 GenBank 进行相似性比对结果见表 4。由表 4 可知, 克隆的菌株 B49 所编码的氨基酸序列与 *Bacillus*、*Lactobacillus*、*Exiguobacterium*、*Geobacillus* 和 *Clostridia* 属某些菌种的乳酸脱氢酶有高度相似性, 其中与 *Bacillus megaterium* 乳酸脱氢酶的相似度最高, 其序列一致性达 54%, 位点相似性达 74%; 与 *Clostridium cellulolyticum* H10 乳酸脱氢酶的序列一致性也达到 48%, 位点相似性达到 66%。

表 3 菌 B49 的 *adh* 基因用 Blastx 检索 NCBI 的同源性结果
Table 3 Result of *adh* gene similarity alignment research on the NCBI web using Blastx

Strain	Identities(%)	Positives(%)	Score	Accession
<i>Clostridium thermocellum</i> ATCC 27405	51	73	397	YP_001038972
<i>Desulfotomaculum reducens</i> MI-1	43	66	322	YP_001114005
<i>Moorella thermoacetica</i> ATCC 39073	41	61	290	YP_429881
<i>Clostridium</i> sp. <i>OhILAs</i>	41	62	283	ZP_01362759
<i>Desulfobacterium hafniense</i> Y51	42	60	279	YP_516856
<i>Thermotoga maritima</i> MSB8	36	58	227	NP_228728
<i>Eubacterium ventriosum</i> ATCC 27560	34	56	214	ZP_02027305

表 4 菌 B49 的 *L-ldh* 基因用 Blastx 检索 NCBI 的同源性结果
Table 4 Result of *L-ldh* gene similarity alignment research on the NCBI web using Blastx

Strain	Identities(%)	Positives(%)	Score	Accession
<i>Bacillus megaterium</i>	54	74	289	P00345
<i>Bacillus cereus</i> ATCC 10987	55	74	285	NP_981325.1
<i>Geobacillus kaustophilus</i> HTA426	57	73	282	YP_146328.1
<i>Exiguobacterium sibiricum</i> 255-15	50	69	265	ZP_00538596.1
<i>Lactobacillus plantarum</i>	49	68	262	CAA50277.1
<i>Clostridium cellulolyticum</i> H10	49	66	261	ZP_01574113.1

3 讨论

本文选用两个梭菌属菌种的乙醇脱氢酶蛋白质序列和 CodeHop 软件来设计 B49 乙醇脱氢酶的简并引物, 引物长度分别为 33 bp 和 36 bp, 简并度分别为 128 和 32 倍, 简并 PCR 获得了预期的结果, 该简并引物成功扩增出一条目标带, 无杂带。采用

Genefisher 软件设计 B49 的 L-乳酸脱氢酶简并引物, 引物长度分别为 20 bp 和 21 bp, 简并度分别为 3072 和 4068 倍, 也成功地克隆到 *L-ldh* 基因片段。但是, Genefisher 软件设计的简并引物 PCR 结果中有较多杂带。这说明用 CodeHop 软件设计的简并引物质量好, 特异性较强, 成功率高。采用 CodeHop 软件设计简并引物并结合 Cassette PCR 法克隆全长基因是

同源克隆家族蛋白基因优化策略之一。

由于简并引物存在简并性问题, 所以设计的引物通常简并性过高, 导致有效引物利用率过低, PCR产物非特异性增高。史兆兴等^[27]认为简并引物的简并度应在 1000~10000 倍之间较好, 而王洪振等^[28]认为简并度最好不要超过 516 倍。史兆兴和王洪振等人在简并引物设计时, 只说明了简并度对简并引物的影响。而作者却认为简并度和引物长度都会影响简并引物的质量。引物长度至少应该在 27 bp 以上, 才能设计出特异性好的引物。

克隆到的乙醇脱氢酶基因具有起始密码子、终止密码子、启动子、RBS 和终止子等开放阅读框和转录单元特征, 它可能是独立转录的。克隆到的 L-乳酸脱氢酶基因没有明显的启动子、RBS 和终止子转录单元特征, 它可能存在于一个操纵子中, 而不是独立转录的。

这两个基因的成功克隆, 为下一步进行高效产氢菌 B49 的代谢工程研究和构建高效产氢基因工程菌研究提供了科学依据。该基因的表达和酶学特征研究有待下一步工作来完成。

参 考 文 献

[1] 樊耀亭, 李晨林, 侯红卫, 等. 天然厌氧微生物氢发酵生产生物氢气的研究. 中国环境科学, 2002, 22(4): 370-374.

[2] 任南琪, 林明, 马汐平, 等. 厌氧高效产氢细菌的筛选及其耐酸性研究. 太阳能学报, 2003, 24(1): 80-84.

[3] Oh YK, Seol EH, Kim JR, et al. Fermentative biohydrogen production by a new chemoheterotrophic bacterium *Citrobacter* sp. Y19. *Int J Hydrogen Energy*, 2003, 28(12): 1353-1359.

[4] Oh YK, Seol EH, Lee EY, et al. Fermentative hydrogen production by a new chemoheterotrophic bacterium *Rhodospseudomonas palustris* P4. *Int J Hydrogen Energy*, 2002, 27(11-12): 1373-1379.

[5] Kumar N, Das D. Enhancement of hydrogen production by *Enterobacter cloacae* IIT-BT08. *Proc Biochem*, 2000, 35(6): 589-593.

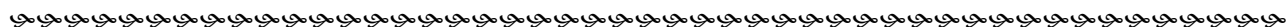
[6] Xing DF, Ren NQ, Li QB, et al. *Ethanoligenens harbinense* gen. nov., sp. nov., isolated from molasses wastewater. *Int J System Evolut Microbiol*, 2006, 56(4): 755-760.

[7] Yokoi H, Tokushige T, Hirose J, et al. Hydrogen produc-

tion by immobilized cells of *Enterobacter aerogenes* strain HO-39. *J Ferment Bioeng*, 1997, 83(5): 481-484.

- [8] Taguchi F, Mizukami N, Saito-Taki T, et al. Isolation of a hydrogen-producing bacterium *Clostridium beijerinckii* strain AM21B, from termites. *Can J Microbiol*, 1993, 39(7): 726-730.
- [9] Taguchi F, Mizukami N, Hasegawa K, et al. Microbial conversion of arabinose and xylose to hydrogen by a new isolated *Clostridium* sp. No.2. *Can J Microbiol*, 1994, 40(3): 228-233.
- [10] Tanisho S, Suzuki Y, Wakao N. Fermentative hydrogen evolution by *Enterobacter aerogenes* strain E82005. *Int J Hydrogen Energy*, 1987, 12(9): 623-627.
- [11] Tanisho S, Kuromoto M, Kadokura N. Effect of CO₂ removal on hydrogen production by fermentation. *Int J Hydrogen Energy*, 1998, 23(7): 559-563.
- [12] Tanisho S. Biological hydrogen production by fermentation. *Proceedings of the Second International Conference on New Energy Systems & Conv*, Istanbul, 1995, pp. 311-316.
- [13] Chen X, Sun YQ, Xiu ZL, et al. Stoichiometric analysis of biological hydrogen production by fermentative bacteria. *Int J Hydrogen Energy*, 2006, 31(4): 539-549.
- [14] Higgins DG, Sharp PM. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene*, 1988, 73(1): 237-244.
- [15] Thompson JD, Gibson TJ, Plewniak F, et al. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*, 1997, 25(24): 4876-4882.
- [16] Henikoff S, Henikoff GJ, Alford JW, et al. Block+: a non-redundant database of protein alignment blocks derived from multiple compilations. *Bioinformatics*, 1999, 15(6): 471-479.
- [17] Rose MT, Schultz RE, Henikoff GJ, et al. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic acids research*, 1998, 26(7): 1628-1635.
- [18] Wei XT, Kuhn DN, Narasimhan G. Degenerate primer design via clustering. bioinformatics conference, CSB. *Proceedings of the 2003 IEEE*, 2003, 11(14): 75-83.
- [19] Wilks FF, Kurban RR, Hovens CM, et al. The application of polymerase chain reaction to cloning members of the protein tyrosine kinase family. *Gene*, 1989, 58(1): 67-74.
- [20] 黄菁, 王少丽, 乔传令. 程序化设计简并引物与克隆小菜蛾酯酶基因. 昆虫知识, 2002, 39(6): 458-461.
- [21] Giegerich R, Meyer F, Schleiermacher C. *GeneFisher* - software support for the detection of postulated genes. *Proceedings of the Fourth International Conference on*

- Intelligent Systems for Molecular Biology, AAAI Press, 1996, 4: 68-77.
- [22] 林明, 任南琪, 马汐平, 等. 产氢发酵细菌培养基的选择和改进. 哈尔滨工业大学学报, 2003, 35(4): 398-402.
- [23] Arnau J, Jorgensen F, Madsen SM, *et al.* Cloning of the *Lactococcus lactis adhE* gene, encoding a multifunctional alcohol dehydrogenase, by complementation of a fermentative mutant of *Escherichia coli*. *J Bacteriol*, 1998, 180(12): 3049-3055.
- [24] Wierenga RK, Terpstra P, Hol WGJ. Prediction of the occurrence of the ADP-binding bab-fold in proteins using an aminoacid sequence fingerprint. *Mol Biol*, 1986, 187: 101-107.
- [25] Jado I, Fenoll A, Casal J, *et al.* Nucleotide sequence and chromosomal location of L-lactate dehydrogenase gene from *Streptococcus pneumoniae*. *Curr Microbiol*, 1998, 37(2): 77-79.
- [26] Turgut-Balik D, Shoemark KD, Sessions BR, *et al.* Mutagenic exploration of the active site of lactate dehydrogenase from *Plasmodium falciparum*. *Biotechnol Lett*, 2001, 23(11): 923-927.
- [27] 史兆兴, 王恒樑, 苏国富, 等. 简并 PCR 及其应用. 生物技术通讯, 2004, 15 (2): 172-175.
- [28] 王洪振, 周晓馥, 宋朝霞, 等. 简并 PCR 技术及其在基因克隆中的应用. 遗传, 2003, 25(2): 201-204.



科技信息

某些微生物术语名称中以姓氏命名的“氏”不能省去

最近, 中国大百科全书(二版)正在按 A、B、C…… 顺序排版, 涉及微生物术语名称有关姓氏问题。提到沙门氏菌属(*Salmonella*)中的“氏”是否可以省除? 长期以来, 在微生物学中发现的某菌均有以发现者姓氏命名的还不少, 在术语名称中这个“氏”不可去掉。对某种新发现的微生物, 以姓氏命名者如已提到的沙门氏杆菌, 加上“氏”(相当中国人的姓)以表示对发现者的尊重与纪念。这个属细菌是美国细菌学家 DE Salmon 发现的而得以命名的, 译成中文名称即沙门氏菌属, 曾译为沙门氏杆菌属, 其中“杆”字倒可省去, 而“氏”应保留。还有如巴斯德氏菌属(*Pasteurella*)、志贺氏菌属(*Shigella*)等等。这里着重强调的是, 菌的发现者并以其命名加个“氏”的内涵不仅仅表示对发现者的尊重与纪念, 而且避免与地名汉译术语相混淆; 再者, 带“氏”的菌名术语名称在教科书中久已通用成为习惯, 突然将“氏”省去, 似感到不适应; 第三, 菌的发现者所做出的贡献可以被后人铭记并激励后人, 在未来的生命科学, 特别是微生物学的发展进程中, 地球、宇宙间客观存在的那些未知微生物或超微生命体, 能培养或不能培养的等等, 都有待生命科学工作者积极探究并做出更多更大的贡献。

(柯为 供稿)