

曲霉 N1-14' 高产 L-苹果酸相关基因的克隆与分析

乐贤松^{1,2,3}, 吴军林^{2,4}, 吴清平², 张菊梅², 郭伟鹏²

(1. 中国科学院南海海洋研究所, 广东广州 510301) (2. 广东省微生物研究所, 省部共建华南应用微生物国家重点实验室, 广东省菌种保藏与应用重点实验室, 广东省微生物新技术公共实验室, 广东广州 510070) (3. 中国科学院大学, 北京 100049) (4. 广东环凯微生物科技有限公司, 广东广州 510643)

摘要: 为了解曲霉(*Aspergillus* sp.)N1-14'的高产 L-苹果酸(LMA)机制,提取其总 DNA 作模板,设计同源引物扩增包含丙酮酸羧化酶基因(*pyc*)和苹果酸脱氢酶基因(*mdh*)全长的片段并测序,再参考测序结果找到丙酮酸羧化酶(PYC)和苹果酸脱氢酶(MDH)的编码区,设计引物从 N1-14'总 cDNA 扩增出其编码序列并通过 TA 克隆测序。测序结果显示 *pyc* 基因编码区长 3582 bp, 编码 1193aa, 分析结果显示 PYC 氨基酸序列在曲霉属内相当保守, 同源性高达 90%以上, 其中 N1-14'在两个保守位点出现突变, 833 位的 A 位于一个环区和 1022 位的 F 位于 α -螺旋中部, 可能与其高产酸活性相关; *mdh* 编码区全长 1023 bp, 编码 340aa。MDH 氨基酸序列高度保守, 突变株同样有两个保守区出现氨基酸点突变, 且两点均出现在 α -螺旋区域。本试验主要克隆两个曲霉 N1-14'产 L-苹果酸通路关键酶基因, 分析其种属的特异性及预测特异氨基酸位点的功能, 为继续探究 N1-14'的高产 LMA 机制及相应的基因工程改造提升产酸水平奠定基础。

关键词: 曲霉 N1-14'; PYC; MDH; 序列分析

文章编号: 1673-9078(2017)2-76-82

DOI: 10.13982/j.mfst.1673-9078.2017.2.012

Cloning and Characterization of the Genes Related to High Production of L-malate in *Aspergillus* sp. N1-14'

LE Xian-song^{1,2,3}, WU Jun-lin^{2,4}, WU Qing-ping², ZHANG Ju-mei², GUO Wei-peng²

(1. South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China) (2. Guangdong Institute of Microbiology, State Key Laboratory of Applied Microbiology Southern China, Guangdong Provincial Key Laboratory of Microbial Culture Collection and Application, Guangdong Open laboratory of Applied Microbiology, Guangzhou 510070, China) (3. University of the Chinese Academy of Sciences, Beijing 100049, China) (4. Guangdong Huankai Microbial Sci. & Tech. Co. Ltd., Guangzhou 510643, China)

Abstract: In order to understand the mechanism behind the high production of L-malate (LMA) by *Aspergillus* sp. N1-14', its total DNA was extracted to be used as the template, homologous primers were designed to amplify the segment containing the full length of the pyruvate carboxylase gene (*pyc*) and the malate dehydrogenase gene (*mdh*), and sequencing was then performed. Based on the results of this sequencing, the coding regions of *pyc* and *mdh* were identified, their coding sequences were amplified from the total cDNA of N1-14' using the designed primers, and sequencing was carried out through TA cloning. The results of the sequencing showed that the open reading frame (ORF) of *pyc* had 3582 bp and it encoded 1193 amino acids. The analysis of the results showed that the amino acid sequence of pyruvate carboxylase (PYC) was quite conserved in *Aspergillus* species as the sequence similarity with other *Aspergillus* species was higher than 90%. Mutations occurred at two conserved sites of N1-14', 833(A) and 1022(F), which were located in a loop and in the middle of α -helix, respectively, suggesting that those mutations may be relevant to the high production of LMA. The ORF of *mdh* had 1023 bp and it encoded 340 amino acids. The amino acid sequence of malate dehydrogenase (MDH) was also conserved across the *Aspergillus* species, and there were two amino acid mutation sites in the conserved domain. Both of them were situated in the α -helix. In this experiment, two genes related to the key process of producing LMA in

收稿日期: 2016-01-20

基金项目: 国家自然科学基金项目 (31271940)

作者简介: 乐贤松 (1989-), 男, 硕士研究生, 研究方向: 食品微生物基因工程; 吴军林 (1978-), 男, 博士, 研究员, 研究方向: 微生物活性物质结构与功能 (共同第一作者)

通讯作者: 吴清平 (1962-), 男, 博士, 研究员, 研究方向: 食品微生物安全监测与控制研究

N1-14' were cloned, the specificity of these two genes was analyzed, and the function of the specific amino acid sites were predicted. This study provides a basis for further study on the mechanisms behind the high production of LMA and genetic manipulations required for increasing its production.

Key words: *Aspergillus* N1-14'; pyruvate carboxylase; malate dehydrogenase; sequence analysis

苹果酸, 又名 2-羟基丁二酸, 自然界存在三种形式的苹果酸, D-苹果酸、L-苹果酸(L-malic acid, LMA) 和 DL-苹果酸混合物。LMA 是继柠檬酸之后的一种极有发展前途的有机酸产品, 在食品工业中, L-苹果酸是继柠檬酸和乳酸之后用量排第 3 位的食品添加剂, 目前主要用作饮料及食品工业中的酸味剂和增味剂^[1]。LMA 还具有多种的生理功能, 能够有效的提高运动能力, 具有抗疲劳、保护心脏、抗氧化和降低抗癌药物毒副作用等生理功能, 可应用于功能性食品及医药产品中。微生物发酵生产的 L-苹果酸, 不含 D-苹果酸, 原料来源丰富、产品成本低廉(约为酶法生产的 1/4)、杂酸含量低(约为酶法生产的 1/5)和食用安全性高等特点^[2]。

Osmani 和 Scrutton^[3,4]发现一些丝状真菌, 如棒曲霉和米曲霉, 拥有细胞质型的 PYC 和 MDH, 正是这些胞质酶的作用大量产若干种有机酸(苹果酸和富马酸)。[1-¹³C]标记葡萄糖作为底物, 直接证明了在黄曲霉和米曲霉中, 还原性 TCA 循环(包含胞质型 PYC 和 MDH)是产富马酸和 L-苹果酸的主要途径^[5]。酶活比较试验同样证明 rTCA 通路相关酶在产酸过程中表现活跃, 酶活受到诸如 pH 等环境因素的影响与有机酸产量正相关^[6]。Zelle 等人^[7]通过基因工程手段在酿酒酵母中过表达自身的 *pyc* 和 *mdh*, LMA 产量提升 3 倍。Brown 等人^[8]过表达米曲霉 NRRL3448 的自身基因(胞质 *pyc* 和 *mdh*)结合苹果酸转运蛋白基因(malate transporter), L-苹果酸的产量提升 27%, 发酵 164 h 苹果酸浓度达到 154 g/L, 表明 PYC 和 MDH 对发酵产 LMA 中的重要作用。本研究团队从 1988 年起, 一直从事糖质原料直接发酵生产 L-苹果酸的研究, 选育出高产 L-苹果酸的突变株曲霉 N1-14' (*Aspergillus* sp. N1-14'), 并且经过多年发酵条件优化, 摇瓶发酵 140 h 产 LMA 达 94 g/L, 平均产酸速率达 0.67 g/(L·h), 发酵液属实际无毒级^[9,10]。周小燕^[11]等人连续监测 N1-14' 的产酸速率与酶活的关系, 证明胞质 PYC 相关性达到 0.9563, MDH 活性同样与 L-苹果酸生产速率正相关, 说明 N1-14' 的高产 LMA 特性与其独特的胞质 PYC、MDH 有关, 这两种酶可能有别于其他种的同工酶, 能够提高葡萄糖到 L-苹果酸的转化效率。

随着曲霉属多种菌株完成全基因组测序, 欲打破 L-苹果酸生产菌株的产酸瓶颈, 应用生物信息学结合

代谢工程技术对菌株进行代谢途径改造, 成为提高 L-苹果酸产量的新途径^[12]。为进一步了解 N1-14' 胞质 PYC 和 MDH 的特异性, 克隆曲霉 N1-14' 的 *pyc* 和 *mdh* 基因及其编码序列并测序显得尤为重要, 本文首次对高效曲霉 N1-14' 的 *pyc* 和 *mdh* 进行克隆、测序和分析, 为后续的进一步从酶活角度解析 N1-14' 高产酸机制, 分析胞质 PYC 和 MDH 的特异性及为菌种代谢工程改造试验奠定基础。

1 材料和方法

1.1 菌种

曲霉 N1-14' (*Aspergillus* sp. N1-14'), 本研究团队选育^[9]。

1.2 培养基

葡萄糖, 80 g/L; (NH₄)₂SO₄, 3 g/L; KH₂PO₄, 0.3 g/L; ZnSO₄·7H₂O, 0.044 g/L; FeCl₂·6H₂O, 0.3 g/L; 酵母粉, 1 g/L; 轻质 CaCO₃, 10 g/L。如果是配制固体培养基, 另加琼脂 20 g/L, 自然 pH, 其中(NH₄)₂SO₄ 分开灭菌。

1.3 试剂

Ezup 柱式基因组 DNA 抽提试剂盒(真菌)为上海生工生物公司产品, TransFast Taq DNA Polymerase 用于扩增 18S 片段(北京全式金), DNA 凝胶回收试剂盒(东盛生物), Total RNA 提取试剂盒(HiPure Fungal RNA Kit, 广州美基生物科技有限公司), pMD®19-T Simple Vector、PrimeScript® Reverse Transcriptase(宝生物工程(大连)有限公司), TransStart FastPfu Fly DNA Polymerase(北京全式金)

1.4 引物

上游引物: FR1: 5'CTCTCAATCTGTCAA TCCTTATT3'、Pyc-F: 5'TTCTGGGACCAACGTC TAGCC3'、Mdh-F: 5'AGCGTCGGCTGTATAG TGTCGC3'、c-PYC-F: 5'ATGGCGGCTCCGTTTC GTCAG3'、c-MDH-F: 5'ATGTTTCGCTGCTCGCCA3'。下游引物: FF2: 5'GGTTCTATTTTGTGGTTTCT3'、Pyc-R: 5'TTCTTCCTCAGATCCCGTTCC3'、Mdh-R:

5'ATTAAAACTCCGTCGCCACC3'、c-PYC-R :
5'TTACGCTTTGACGATCTTGACAG3'、c-MDH-R :
5'TTAAGGGTTGGCCTTGACG3'。

1.5 方法

1.5.1 克隆比较曲霉 N1-14' 的 18S rDNA 序列

总 DNA 提取参照基因组 DNA 抽提试剂盒, 取液体发酵 72 h 的菌丝体, 采用液氮研磨法提取总 DNA, 稀释 100 倍后用作模板, 引物为 FR1/FF2, 扩增 18S rDNA 序列, PCR 反应体系为: 模板, 2 μ L; 引物, 1 μ L; 10 \times TransFast Taq Buffer, 5 μ L; 2.5 mmol/L dNTPs, 4 μ L; TransFast Taq DNA Polymerase, 1 μ L; ddH₂O, 36 μ L。扩增程序: 95 $^{\circ}$ C, 5 min; 95 $^{\circ}$ C, 30 s; 55 $^{\circ}$ C, 30 s; 72 $^{\circ}$ C, 30 s; 35 个循环; 72 $^{\circ}$ C, 7 min; 12 $^{\circ}$ C 保存。反应完毕后取 5 μ L 于 1.5% 琼脂糖凝胶电泳检测分析。切胶回收目的条带, 按照 DNA 凝胶回收试剂盒说明书进行, 回收所得产物送交华大基因进行测序。利用 NCBI 网站的 BLAST 在线分析功能及 DNAMAN 软件对所得 DNA 序列进行系统发育分析。

1.5.2 克隆曲霉 N1-14' 的 pyc 基因和 mdh 基因全长并测序

鉴于 N1-14' 与烟曲霉的高度同源性, 参考烟曲霉 (*Aspergillus fumigatus* Af293) 的 pyc 和 mdh 序列, 设 2 对同源性引物用于扩增 N1-14' 的丙酮酸羧化酶基因 (pyc) 和苹果酸脱氢酶基因 (mdh) 全长。总 DNA 为模板, 反应体系 (TransFast Taq DNA Polymerase) 如下, 扩增程序: 95 $^{\circ}$ C、5 min; 95 $^{\circ}$ C、30 s; 退火温度分别为 60 $^{\circ}$ C、30 s (pyc) 和 58 $^{\circ}$ C、30 s (mdh); 延伸程序为 72 $^{\circ}$ C、3 min (pyc) 和 72 $^{\circ}$ C、2 min (mdh); 30 个循环; 72 $^{\circ}$ C、7 min; 12 $^{\circ}$ C 保存。PCR 产物经琼脂糖凝胶电泳后切胶回收送交华大基因测序, Blast 比对以确定 pyc 和 mdh 的全长。

1.5.3 总 RNA 提取及 pyc 和 mdh 相应编码序列 (CDS) 克隆测序

用总 RNA 提取试剂盒, 参照试剂说明书, 将少量 N1-14' 菌丝体研磨后提取总 RNA, 参照 PrimeScript[®] Reverse Transcriptase 使用说明反转录得到总 cDNA。以上总 cDNA 稀释 50 倍作为模板, 参照 pyc 和 mdh 开放阅读框起始与终止位置序列各设计一对引物 c-PYC(F/R) 和 c-MDH(F/R) 用于扩增其编码序列, 扩增程序参照说明书, 退火温度分别为 60 $^{\circ}$ C 和 57 $^{\circ}$ C, 延伸时间为 2 min 和 45 s。PCR 产物经琼脂糖凝胶电泳之后, 产物回收并通过 TA 克隆插入 pMD-19t 载体 (参照试剂说明书), 挑取白斑进行 PCR 菌落验证, 将阳性菌株送交华大测序并上传至 NCBI。

2 结果与讨论

2.1 18S rDNA 序列比对分析

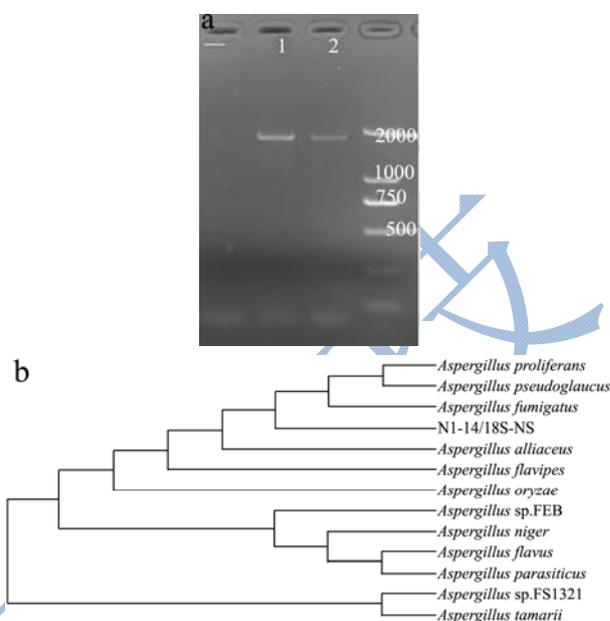


图1 曲霉 N1-14' 18S rDNA 电泳图 (a) 和进化树分析 (b)

Fig.1 Electrophoresis (a) and analysis of phylogenetic analysis (b) of 18S rDNA sequence of N1-14'

注: 图 1a 中 “-” 为阴性对照; 1 和 2 为 Total DNA 扩增泳道, 均扩出一条 1800 bp 左右的条带。图 1b 为曲霉 N1-14' 与其他曲霉属菌种的 18S rDNA 同源性分析。 *Aspergillus proliferans*, GeneID: 12862460; *Aspergillus pseudoglaucus*, GeneID: 12862461; *Aspergillus fumigatus* strain WL002, GeneID: 634869525; N1-14', GeneID: KT935264; *Aspergillus alliaceus*, GeneID: 12862448; *Aspergillus flavipes* strain CBS 260.73, GeneID: 323371154; *Aspergillus oryzae* RIB40, GeneID: 83775367; *Aspergillus sp.* FE8, GeneID: 47678946; *Aspergillus niger*, GeneID: 134057889; *Aspergillus flavus*, GeneID: 945411; *Aspergillus parasiticus*, GeneID: 945414; *Aspergillus sp.* FS132, GeneID: 117066818; *Aspergillus tamarii*, GeneID: 29126062。

曲霉 N1-14' 经 FR1/FF2 两条引物特异性扩增后, 琼脂糖凝胶电泳检测结果显示如图: 琼脂糖凝胶电泳后出现一条 1800 bp 左右的单一明亮条带, 切胶回收后并测序, 登录号: KT935264, 经 18S rDNA 序列同源性检索出同源性最高的 12 条 DNA 序列用于 Mega 进化树制作, 如图 1b 所示, N1-14' 在进化关系上与洋葱曲霉 (*Aspergillus alliaceus*) 最相近, 相似度为 99.60%, 其后依次烟曲霉 (*Aspergillus fumigatus* strain WL002), 相似度为 99.32%; 假灰绿曲霉 (*Aspergillus pseudoglaucus*), 相似度为 99.27%; 黄柄曲霉 (*Aspergillus flavipes* strain CBS 260.73), 相似度为

99.21%；多育曲霉 (*Aspergillus proliferans*)，相似度为 99.15%，考虑到烟曲霉有上传基因组序列，将其基因组作为参考。

2.2 pyc 和 mdh 全长分析

克隆测序并同源比对后，测序结果包含 pyc 基因的序列总长为 5419 bp 和 mdh 基因的序列总长为 2892 bp。将包含 pyc 和 mdh 的序列上传至 GeneBank，在线 (<http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi>) 结合同源基因的表达框分析 pyc 和 mdh 的开放阅读框起始与终止的位置。pyc 基因从反向 275 bp 处 ATG 开始编码，到 3919 bp 处 TAA 终止，包含内含子部分全长 3645 bp；而 mdh 基因从正向 53 bp 处 ATG 开始编码到 1360 bp 处 TAA 终止，包含内含子部分全长 1308 bp。

2.3 pyc 和 mdh 编码区序列分析

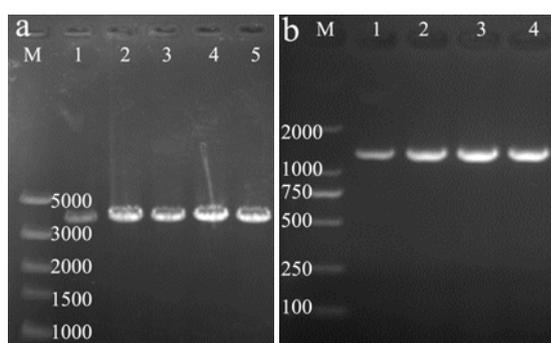


图 2 pyc 编码区电泳图 (a) 和 mdh 编码区电泳图 (b)

Fig.2 Electrophoresis of the coding region of pyc (a) and Electrophoresis of the coding region of mdh (b)

注：图 2a 为 pfu 酶扩增后得到的 pyc 编码区电泳图，大小约为 3500 bp；图 2b 为 pfu 酶扩增后得到的 mdh 编码区电泳图，大小约为 1200 bp。

扩增 pyc 基因和 mdh 基因编码区所得产物经电泳后，结果图 2 所示。

阳性 TA 克隆测序后拼接得到 pyc 编码区(cPYC)，登录号为：KT935265，cPYC 全长 3582 bp，编码一个含 1193 个氨基酸的多肽，其分子量约为 35.6 ku，等电点约为 8.5，总平均亲水性为-0.019。推导的丙酮酸羧化酶氨基酸序列作进化树分析如下：

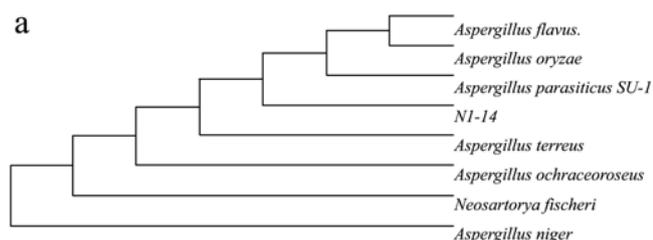


图 3 PYC 进化树分析 (a) 和 PYC 氨基酸序列比对分析 (b)

Fig.3 Phylogenetic analysis of PYC (a) and BLAST analysis of the amino acid sequence of PYC (b)

注：图 3a 为 N1-14'PYC 及同源蛋白氨基酸序列构建的进化树，PYC 登录号从上至下依次为 *Aspergillus flavus* (UniProtKB-B8N9R6)、*Aspergillus oryzae* (UniProtKB-Q2UGL1)、*Aspergillus parasiticus* SU-1 (UniProtKB-A0A0F0ILW2)、N1-14' (gi:KT935265)、*Aspergillus terreus* (UniProtKB-O93918)、*Aspergillus terreus* (UniProtKB-A0A0F8U4L3)、*Neosartorya fischeri* (UniProtKB-A1CXK2)、*Aspergillus niger* (UniProtKB-G3Y8N0)。图 3b 为 PYC 同源蛋白序列比对分析。比对序列从上至下序列来自 N1-14' (gi:KT935265)、*Aspergillus parasiticus* SU-1 (gi:770309847)、*Aspergillus oryzae* RIB 40 (gi:169773675)、*Aspergillus terreus* (gi:51701707)、*Aspergillus* sp. (gi:143360658)、*Neosartorya fischeri* (gi:119500988)、*Neosartorya udagawae* (gi:849273887)、*Aspergillus terreus* (gi:816334271)。

分析得知，N1-14' 与黄曲霉 (UniProtKB-B8N9R6)、米曲霉 (UniProtKB-Q2UGL1)、寄生曲霉 (UniProtKB-A0A0F0ILW2) 和土曲霉 (UniProtKB-O93918) 的丙酮酸羧化酶氨基酸序列具有很高的同源性，分别为 98.6%、98.6%、98.7%和 94.3%，其中与寄生曲霉的 PYC 同源性最高。经氨基酸序列 BLAST 分析 (图 3b)，突变株 N1-14' 第 833 位氨基酸为 A (疏水性氨基酸)，而其他同源的丙酮酸羧化酶为 E (亲水性氨基酸)，SMART (<http://smart.embl-heidelberg.de/>) 结构域预测其位于 TIM 桶状结构域 (图 5a) 的一个环中。N1-14' 的 1022 位为 F，而其他种均为 Y，SMART 结构域预测该氨基酸位于草酰乙酸脱羧酶功能域 (图 5e) 的一个 α 螺旋的开端。在线 (SWISS-MODEL) 进行同源建模，模板为单增李斯特菌的丙酮酸羧化酶的晶体结构 (SMTL id 4qsk.1)，相似度为 51.63%，GMEQ 值为 0.74，PredictProtein 分析显示，PYC 编码多肽包含 39 个 α -helices 和 50 个 β -strands，其中 1022 位的 F 位于 α 螺旋中部，SWISS-MODEL 建模，3-D 结构如图 5f。

GEARDALVNR IQFGGDEVVKAKDGAGSA TLMAMAGARMAES LLRAAQGEEGVVE PT FVDS PLYKDQGVDF FASKVEL GPNG	300
GKTRDELVNR IQFGGDEVVKAKDGAGSA TLMAMAGARMAES LLKAAQGEEGVVE PT FVDS PLYKDQGVDF FASKVEL GPNG	300
GKTRDELVNR IQFGGDEVVKAKDGAGSA TLMAMAGARMAES LLKAAQGEEGVVE PT FVDS PLYKDQGVDF FASKVEL GPNG	300
GKTRDELVNR IQFGGDEVVKAKDGAGSA TLMAMAGARMAES LLKAAQGEEGVVE PT FVDS PLYKDQGVDF FASKVEL GPNG	300
GKTRDELVNR IQFGGDEVVKAKDGAGSA TLMAMAGARFAES LLKAAQGEEGVIE PT FVDS PLYKDQGVDF FASRVEL GPNG	300
GETRDALVNR IQFGGDEVVKAKDGAGSA TLMAMAGARFAES LLKAAQGEEGVIE PT FV DS PLYKDQGVDF FASRVEL GPNG	300
GETRDALVNR IQFGGDEVVKAKDGAGSA TLMAMAGARFAES LLKAAQGEEGVIE PT FV ES PLYKDQGVDF FASRVEL GPNG	300
GETRDLVNR IQFGGDEVVKAKDGAGSA TLMAMAGARFAES LLKAAQGEEGVIE PT FV ES PLYKDQGVDF FASRVEL GPNG	300
GETRDLVNR IQFGGDEVVKAKDGAGSA TLMAMAGARFAES LLKAAQGEEGVIE PT FV ES PLYKDQGVDF FASRVEL GPNG	300

图 4 MDH 氨基酸序列比对分析

Fig.4 BLAST analysis of the amino acid sequence of MDH

注：图 4，N1-14' MDH 及同源蛋白氨基酸序列比对分析，从上至下序列来自 N1-14'(KT949002)、Aspergillus oryzae RIB 40 (gi 169766066)、Aspergillus flavus AF70(gi 768707059)、Aspergillus flavus NRRL3357 (gi 238482699)、Aspergillus terreus NIH2624 (gi 115399894)、Neosartorya udagawae (gi 849270975)、Aspergillus fumigatus var.RP-2014 (gi 666427958)、Aspergillus fumigatus Af293 (gi 143360658)、Neosartorya fischeri NRRL181 (gi 119482986)。

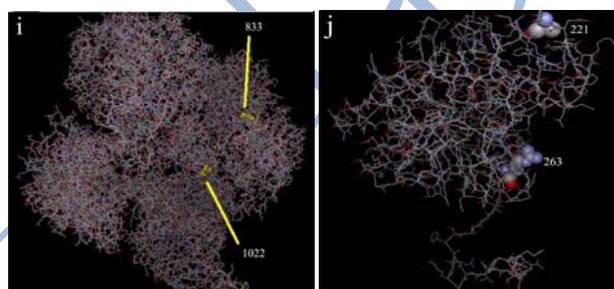
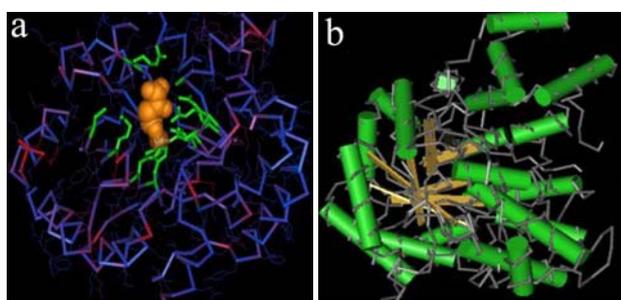
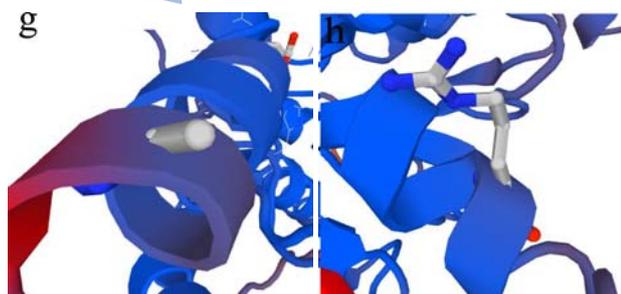
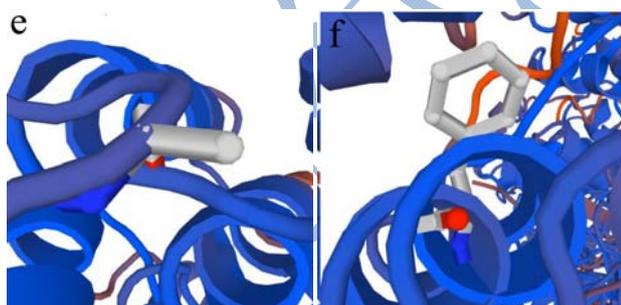
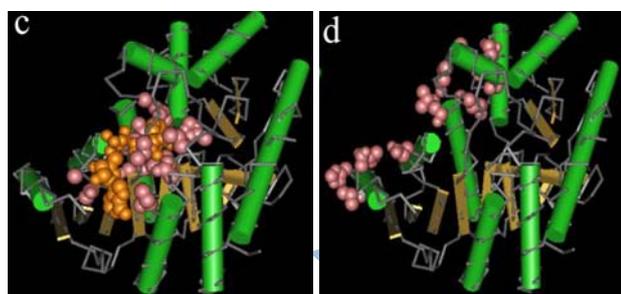


图 5 特异氨基酸 3D 结构分析

Fig.5 Analysis of the 3-dimensional structure of specific amino acids

注：图 5a 表示 TIM 桶状结构域(TIM- barrel domain)，黄色部分代表 CO²⁺-生物素辅基；图 5b 表示草酰乙酸脱羧酶活性中心(OADA domain)；图 5c 表示 NAD 结合功能域(NAD binding domain)；图 5d 表示 MDH 二聚化界面(dimerization interface)；图 5e 和 f 表示 PYC 中，A₈₃₃ 和 F₁₀₂₂ 的二级结构；图 5g 和 h 表示 MDH 中，A₂₂₁ 和 R₂₆₃ 的二级结构；图 5i 和 j 分别是特异氨基酸在两种酶蛋白三维结构上的位置。



阳性 TA 克隆测序后拼接得到 *mdh* 编码区 (cMDH)，登录号为：KT949002，cMDH 全长 1023 bp，编码一个含 340 个氨基酸的蛋白，其分子量约为 113 ku，等电点约为 6.0，总平均亲水性为-0.202。经同源蛋白氨基酸序列系统发育树分析（未列出）发现，苹果酸脱氢酶在各种曲霉中高度保守，N1-14'的 MDH 与米曲霉 MDH(UniProtKB-Q2USG3)寄生曲霉 MDH (UniProtKB-A0A0F0I277) 的同源性最高，达到 97.4%，黄曲霉 (UniProtKB-A0A0D9N180) 为 97.1%，土曲霉 (UniProtKB-Q0CIX6) 和费希新萨托菌 (UniProtKB-A1DCR4) 均为 92.9%。经氨基酸序列 BLAST 分析 (图 4)，突变株 N1-14'的第 263 位氨基酸为 R 而其它种均为 K，另外 N1-14'的 221 位为 A，而其他种均为 T，这两个突变氨基酸均位于 α-螺旋上 (图 5g 和 h)。SMART 结构域预测这两个氨基酸残基

分别位于 MDH 的 NADP-binding 功能域的 Rossmann 折叠和 MDH 二聚化界面。在线 (SWISS-MODEL) 进行同源建模, 模板为人 2 型苹果酸脱氢酶晶体结构 (SMTL id 2dfd.1.A), 序列相似性为 58.77%, GMEQ 值为 0.78, 3D 模型包含 12 个 α -helices 和 13 个 β -strands, PredictProtein 分析显示, K₂₂₀ 和 R₂₂₂ 为 NAD-binding 活性中心。

根据 N1-14' 的 18S rDNA 序列同源性构建进化树, 发现 N1-14' 在进化关系上与洋葱曲霉 (*Aspergillus alliaceus*) 最相近, 但与其他种的相似度也同样超过 99%, 因此仅通过 18S rDNA 序列比对难以鉴定, 进一步的鉴定工作需要借助于菌丝体微观形态比较和生理生化试验。PYC 对应氨基酸序列同源比对显示, N1-14' 的 PYC 与曲霉属其它种的同源性很高, 均达到 90%, 其中最相近的是寄生曲霉 (UniProtKB-A0A0F0ILW2)。一级结构比对中发现, N1-14' 有两个氨基酸残基在属内其他种中相当保守, 而在 N1-14' 中出现变异, 分别是 A₈₃₃ 和 F₁₀₂₂, 丙氨酸位于一个保守的丙糖磷酸异构酶桶状结构域 (TIM-barrel domain, 如图 5a), 该结构域含有一个包含 Co²⁺ 结合位点的催化中心, 催化中心的三个核心氨基酸为 D₈₁₄、E₈₃₃ 和 R₈₄₇, N1-14' 的亲水性氨基酸 E₈₃₃ 电离后带负电结合 Co²⁺-生物素, 对转羧基活性起重要作用; 变成疏水性氨基酸 A₈₃₃ 后, 理论上这样的变化可能改变其转羧基酶活性, 从而影响丙酮酸结合 CO₂ 的过程^[13]。F₁₀₂₂ 位于 PYC 草酰乙酸脱羧酶功能域 α 链 (OADA 区), SWISS-MODEL 模拟 (图 5b) 显示 F₁₀₂₂ 位于 PYC 外侧的一个凹槽外围, 高保守的 T 处于 OADA 活性中心外侧, 其羟基亲水端对维持 OADA 活性中心的高级结构非常重要; 突变成 F 后, 这个残基无法处于该活性中心外侧的水环境中而通过疏水作用向内侧延伸, 改变 OADA 活性中心高级结构, 介于曲霉 N1-14' PYC 高活性表现, 可能突变后形成另一种更高活性的高级结构, 至于这种突变与酶活之间关系仍在研究中。PYC 是还原性 TCA 循环产 L-苹果酸过程中重要酶, 本课题组通过实验证明突变菌株表现出更高的 PYC 活性, 可能与氨基酸突变改变蛋白的三维空间结构有关, 这两个氨基酸位点的变异可能是曲霉 N1-14' 胞质 PYC 酶活提升的原因, 进一步研究其与 PYC 酶活的关系需要借助于氨基酸定点突变和结构生物学方法。

MDH 氨基酸序列同源性比对结果显示, N1-14' 的 MDH 与米曲霉 MDH (UniProtKB-Q2USG3) 和寄生曲霉 MDH (UniProtKB-A0A0F0I277) 的同源性最高, 属内进化关系上高度保守, 同样在两个相当保守的位点, N1-14' 出现氨基酸位点变异。3D Molecule

Viewer 显示 A₂₂₁ 位于 MDH 的 NAD-binding 功能域一个 α -螺旋上, 米曲霉 (*Aspergillus oryzae* RIB40) MDH 的 K₂₂₀ 和 R₂₂₂ 均为 NAD-binding 活性中心的起辅助作用氨基酸, 毗邻活性中心的 A₂₂₁ 突变理论上会一定程度改变 MDH 的 NAD-binding 活性中心。R₂₆₃ 位于二聚化界面 (dimerization interface), 已知的所有 MDH 都是多聚体形式^[14], 尽管 R₂₆₃ 不是活性中心, 但氨基酸大小的变化理论上会不同程度的因空间位阻变化影响亚基相互之间的结合, 其变异与酶活及其产 L-苹果酸之间的关系, 有待进一步验证。

3 结论

本研究中, 作者通过克隆测序分析曲霉 N1-14' 还原性 TCA 循环中两种产 L-苹果酸关键酶, 从进化生物学角度分析了 PYC 和 MDH 的特异性, 结合其在曲霉 N1-14' 中的高活性表型, 分析其可能的高活性原因, 三维模拟特异氨基酸位点与酶活性中心的位置关系及其可能产生的影响, 为后续研究 PYC、MDH 两种特异酶和对 N1-14' 进行基因工程改造以提升 LMA 产量提供理论依据和参考。

参考文献

- [1] Zelle R M, De Hulster E, Van Winden W A, et al. Malic acid production by *Saccharomyces cerevisiae*: engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export [J]. Applied and Environmental Microbiology, 2008, 74(9): 2766-2777
- [2] 刘建龙, 刘建军, 杨连生. 酶转化法及微生物发酵法生产 L-苹果酸的研究进展 [J]. 中国酿造, 2005, 24(7): 5-7
LIU Jian-long, LIU Jian-jun, YANG Lian-sheng. Research advancement of L-malic acid production in enzyme conversion and fermentation [J]. China Brewing, 2005, 24(7): 5-7
- [3] Osmani Stephen A, Michael C Scrutton. The sub-cellular localisation of pyruvate carboxylase and of some other enzymes in *Aspergillus nidulans* [J]. European Journal of Biochemistry, 1983, 133(3): 551-560
- [4] Osmani, Stephen A, Michael C Scrutton. The sub-cellular localisation and regulatory properties of pyruvate carboxylase from *Rhizopus arrhizus* [J]. European Journal of Biochemistry, 1985, 147(1): 119-128
- [5] Peleg Yoav, Barry Stieglitz, Israel Goldberg. Malic acid accumulation by *Aspergillus flavus* [J]. Applied Microbiology and Biotechnology 1988, 28(1): 69-75
- [6] Yang L, Lübeck M, Ahring B K, et al. Enhanced succinic

- acid production in *Aspergillus saccharolyticus* by heterologous expression of fumarate reductase from *Trypanosoma brucei* [J]. Applied Microbiology and Biotechnology, 2015: 1-11
- [7] Zelle R M, de Hulster E, van Winden W A, et al. Malic acid production by *Saccharomyces cerevisiae*: engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export [J]. Applied and Environmental Microbiology, 2008, 74(9): 2766-2777
- [8] Brown, Stephen H, L Bashkirova, R Berka, et al. Metabolic engineering of *Aspergillus oryzae* NRRL 3488 for increased production of L-malic acid [J]. Applied Microbiology and Biotechnology, 2013, 97(20): 8903-8912
- [9] 吴清平,周小燕,钟瑜,等.L-苹果酸产生菌筛选及高产突变株诱变选育[J].真菌学报,1993,4:304-312
- WU Qing-ping, ZHOU Xiao-yan, ZHONG Yu, et al. Screening of L- malic acid producing strain and mutation breeding of high yield mutant [J]. Acta Mycologica Sinica, 1993, 4: 304-312
- [10] 吴清平,周小燕,陈素云,等.中国,专利号 CN 1112160A [P] 1995.11.22,
- WU Qing-ping, ZHOU Xiao-yan, CHEN Su-yun, et al. China, patent numb CN 1112160A [P] 1995.11.22
- [11] 周小燕,吴清平,蔡芷荷,曲霉 N1-14'胞质酶活性与产 L-苹果酸能力的关系[J].微生物学报,2000,40(5):500-506
- ZHOU Xiao-yan, WU Qing-ping, CAI Zhi-he, et al. Relationship between aspergillus N1-14' cytoplasmic enzyme activity and production capacity of L-malic acid [J]. Acta Microbiologica Sinica, 2000, 40(5): 500-506
- [12] Arnaud M B, Chibucos M C, Costanzo M C, et al. The *Aspergillus* genome database, a curated comparative genomics resource for gene, protein and sequence information for the *Aspergillus* research community [J]. Nucleic Acids Research, 2012, 38(suppl 1): D420-D427
- [13] Hall P R, Zheng R, Pusztai-Carey M, et al. Expression and crystallization of several forms of the *Propionibacterium shermanii* transcarboxylase 5S subunit [J]. Acta Crystallographica Section D: Biological Crystallography, 2004, 60(3): 521-523
- [14] Madern D. Molecular evolution within the L-malate and L-lactate dehydrogenase super-family [J]. Journal of Molecular Evolution, 2002, 54(6): 825-840