



(12) 发明专利申请

(10) 申请公布号 CN 104031902 A

(43) 申请公布日 2014. 09. 10

(21) 申请号 201310067821. 2

C12R 1/645(2006. 01)

(22) 申请日 2013. 03. 04

C12R 1/89(2006. 01)

(71) 申请人 中国科学院大连化学物理研究所

地址 116023 辽宁省大连市中山路 457 号

(72) 发明人 赵宗保 林心萍 张素芳 王雅南

(74) 专利代理机构 中科专利商标代理有限责任

公司 11021

代理人 陈晓娜

(51) Int. Cl.

C12N 9/88(2006. 01)

C12N 15/60(2006. 01)

C12N 15/63(2006. 01)

C12N 1/19(2006. 01)

C12N 1/13(2006. 01)

C12N 1/21(2006. 01)

C12R 1/865(2006. 01)

权利要求书1页 说明书10页

序列表11页 附图5页

(54) 发明名称

一种产油酵母 ATP 柠檬酸裂解酶及其应用

(57) 摘要

本发明涉及自圆红冬孢酵母分离得到的 ATP 柠檬酸裂解酶 (ACL) 及其编码核苷酸序列和重组载体。利用圆红冬孢酵母转录组测序结果,以圆红冬孢酵母菌 CGMCC2. 1389 总 RNA 为模板,利用 RT-PCR,分离得到 Rt-ACL cDNA 序列,如 SEQ ID No :1 所示,其编码蛋白的氨基酸序列如 SEQ ID No :3 所示。本发明通过在红酵母中表达 Rt-ACL 基因组 DNA 序列,能够提高产油酵母细胞内的油脂含量;另外还通过构建到产油细菌或者产油微藻中的表达载体中,在其中表达来源于 Rt-ACL 的 cDNA 序列,也获得油脂含量的积累。本发明的 ATP 柠檬酸裂解酶基因可以用于构建微生物油脂生产细胞。

1. 一种 ATP 柠檬酸裂解酶,所述 ATP 柠檬酸裂解酶是由如下 (a) 或 (b) 的蛋白质:
 - (a) 由 SEQ ID NO :3 所示的氨基酸序列组成的蛋白质;
 - (b) 将 SEQ ID NO :3 的氨基酸经过一个或几个氨基酸的取代和 / 或缺失和 / 或添加且与具有油脂贮运和代谢调控相关活性的由 SEQ ID NO :3 衍生的蛋白质。
2. 编码权利要求 1 所述的 ATP 柠檬酸裂解酶的核苷酸序列。
3. 根据权利要求 2 所述的核苷酸序列,其为 SEQ ID NO :1。
4. 包含权利要求 2 或 3 的核苷酸序列的重组载体。
5. 一种重组细胞,所述细胞导入权利要求 4 所述的重组载体。
6. 一种构建油脂生产重组细胞的方法,所述方法包括:将权利要求 2 或 3 所述的核苷酸序列导入宿主细胞中,得到油脂生产重组细胞。
7. 根据权利要求 6 所述的方法,其中通过用权利要求 4 所述的重组载体转化宿主细胞而导入权利要求 2 或 3 所述的核苷酸序列。
8. 根据权利要求 6 所述的方法,其中所述宿主细胞选自酵母、微藻、红球菌或大肠杆菌。
9. 根据权利要求 8 所述的方法,其中所述酵母选自酿酒酵母 (*Saccharomyces cerevisiae*) 或圆红冬孢酵母 (*R. toruloides*),所述微藻选自衣藻或聚球藻,所述红球菌是浑浊红球菌 (*R. opacus*)。
10. 根据权利要求 9 所述的方法,其中所述衣藻是莱茵衣藻 (*C. reinhardtii*)。

一种产油酵母 ATP 柠檬酸裂解酶及其应用

技术领域

[0001] 本发明涉及ATP 柠檬酸裂解酶ACL,具体地说是自圆红冬孢酵母分离的ATP 柠檬酸裂解酶基因的克隆、重组载体构建,以及在基因工程研究中的应用。

背景技术

[0002] 国际石油价格屡创新高,低油价的时代一去不复返。为了保障国家石油安全,缓解石油过度依赖进口的危机,同时有效降低机动车污染物排放,我国正大力推动生物柴油的发展。生物柴油主要指以动植物油脂为原料,用甲醇或乙醇在催化剂作用下经脂交换制成的脂肪酸酯。然而利用动植物油脂为原料存在着与人争粮与农争地的问题,而微生物因其生长快、基本不依赖耕地、可连续生产、便于改造等特点,近年来显示出其突出的特点 [赵宗保. 中国生物工程杂志 2005, 25 (2), 8-11]。因此利用微生物生产油脂为生物柴油提供原料是现今可选替代方案之一。

[0003] 自然界中部分微生物在极端情况(如氮源缺乏)能在胞内积累超过其细胞干重 20% 的油脂,其中以甘油三酯为主,这些微生物被称为产油微生物。主要包括细菌、酵母、霉菌、藻类等。其中产油酵母包括 *Rhodotorula*, *Candida*, *Cryptococcus*, *Rhizopus*, *Trichosporon* 和 *Yarrowia* 属中的某些菌株 [Ratledge C, Wynn J P. *Adv Appl Microbiol* 2002, 51, 1-51]。产油细菌包括 *Colwellia*, *Shewanella*, *Alteromonas*, *Pseudoalteromonas* 和 *Ferrimonas* 属中的某些菌株 [相光明. 粮食与油脂 2008 (6), 7-11]。产油微藻包括 *Chlorella*, *Nannochloris* 和 *Phaeodactylum* 属中的某些菌株 [郑洪立, 中国生物工程学报 2009, 29, 110-116。王立柱, 微生物学通报 2010, 37, 336-347。]

[0004] 关于产油酵母,尤其是红酵母基因工程改造的研究相对匮乏,其在特定条件下积累油脂的机制仍未从分子水平上进行阐明仅仅从生化阐明产油酵母在调控油脂合成代谢过程中可能是ATP 柠檬酸裂解酶(ACL)和苹果酸脱氢酶(ME)起到关键作用 [Evans, C. T. and C. Ratledge. *Canadian Journal of Microbiology* 1985, 31 :1000-1005.]。ACL 存在于细胞质中,催化柠檬酸裂解为乙酰辅酶 A 和草酰乙酸。它广泛存在于油脂酵母和霉菌中。在非产油酵母(即积累油脂不超过细胞干重 20%)中不存在这个酶。近年来,有研究表明,红发夫酵母中 ACL 基因在限氮条件下,随着氮源的减少活性增大,其活性的提高有助于虾青素的积累 [Cipriano Chávez-Cabrera, Zoila R, et al. *Appl Microbiol Biotechnol* 2010, 85 :1953-1960.]。另有研究表明,在高山被孢霉 (*Mortierella alpina*) 中表达来源于自身的 ACL 基因,能够脂肪酸含量得到较大的提高 [专利号 :200880113428. X]。

[0005] 但是,至今为止未有人进行过产油酵母的 ACL 基因的分离及其基因功能的鉴定,也没有人使用这个基因进行过油脂的生产。

发明内容

[0006] 本发明的目的是提供一种产油酵母 ATP 柠檬酸裂解酶蛋白、其编码核苷酸序列以

及应用。具体来说,是一种来源于圆红冬孢酵母 CGMCC2.1389 的 ATP 柠檬酸裂解酶新基因 Rt-ACL 的克隆和生物功能分析,以及该基因在构建重组载体和构建基因工程菌株领域的应用。

[0007] 为实验上述目的,本发明采用的技术方案为:

[0008] 首先,利用圆红冬孢酵母转录组测序结果,设计 cDNA 扩增的引物,以圆红冬孢酵母 CGMCC2.1389 总 RNA 为模板,经 RT-PCR 得到 ATP 柠檬酸裂解酶基因 Rt-ACL,其 cDNA 具有如序列表 SEQ ID NO:1 所示的序列,其基因组 DNA 具有如序列表 SEQ ID NO:2 所示的序列,其氨基酸具有如序列表 SEQ ID NO:3 所示的序列。而后通过将其导入酿酒酵母 (*Saccharomyces cerevisiae*) 中,由于酿酒酵母中不存在 ACL 基因,因此能够对 ACL 的活性进行检测 [Beth L.Fatland,Jinshan Ke,et al.PlantPhysiology2002,130:740-756.]。最后通过将这个功能确定的基因在红酵母自身、其他产油酵母及其他产油细菌、微藻中进行过表达,获得了油脂含量大量提高的菌株。

[0009] 本发明提供的产油酵母 ATP 柠檬酸裂解酶,名称为 Rt-ACL(*R.toruloides* ATP-citrate lyase,圆红冬孢酵母 ATP 柠檬酸裂解酶),来源于圆红冬孢酵母 *R.toruloides* CGMCC2.1389。所述产油酵母 ATP 柠檬酸裂解酶是由如下 (a) 或 (b) 的蛋白质:

[0010] (a) 由 SEQ ID NO:3 所示的氨基酸序列组成的蛋白质;

[0011] (b) 将 SEQ ID NO:3 的氨基酸经过一个或几个氨基酸的取代和 / 或缺失和 / 或添加且与具有油脂贮运和代谢调控相关活性的由 SEQ ID NO:3 衍生的蛋白质。

[0012] 为了使 (a) 中的蛋白便于纯化,可以由 SEQ ID NO:3 所示的氨基酸序列组成的蛋白质的氨基末端或羧基末端连接上如表 1 所示的标签。

[0013] 表 1 标签及其序列

[0014]

标签	氨基酸残基数	序列
Poly-Arg	5-6(通常为 5 个)	RRRRR
FLAG	8	DYKDDDDK
Poly-His	2-10(通常为 6 个)	HHHHHH
C-myc	10	EQKLISEEDL
Strep-tag II	8	WSHPQFEK
Poly-Phe	11	FFFFFFFFFFF

[0015] 上述 (b) 中的蛋白可人工合成,也可先合成其编码基因,再进行常规蛋白质表达得到。上述 (b) 中的蛋白的编码基因可能通过将 SEQ ID NO:1 或 SEQ ID NO:2 所示的 DNA 序列中缺失一个或几个编码氨基酸的密码子,和 / 或进行一个或几个碱基对的错义突变,和 / 或在其 5' 端和 3' 端连上表 1 所示的标签的编码序列而获得。

[0016] 所述产油酵母 ATP 柠檬酸裂解酶 Rt-ACL 的氨基酸序列和编码核苷酸序列 (rtacl,

RHTO_03915. t1) 均属于本发明的保护范围。优选地,所述蛋白的编码核苷酸序列为序列表中 SEQ ID NO :1 所示的 DNA 分子。

[0017] 含有所述核苷酸序列的重组表达载体、表达盒、转基因细胞系或重组菌株均属于本发明的保护范围。

[0018] 因此,本发明还提供包含编码本发明的产油酵母 ATP 柠檬酸裂解酶的核苷酸序列的重组载体,优选是重组表达载体;提供导入了(例如,通过转化或转染技术导入)所述重组载体的重组细胞。

[0019] 本领域技术人员应该理解,将编码产油酵母 ATP 柠檬酸裂解酶的核苷酸序列或包含编码产油酵母 ATP 柠檬酸裂解酶的核苷酸序列的重组载体导入宿主细胞可以按照本领域的常规技术进行,例如,可以通过转化、转染(例如,农杆菌介导的转染)或电穿孔等常规技术进行。

[0020] 本发明还保护扩增所述基因的引物。

[0021] 可用现有的产油酵母、产油细菌和产油微藻表达载体构建含有所述编码基因(rtacl)的重组表达载体。所述的还包括根瘤农杆菌介导基因重组的载体和可用于基因敲除线性载体等。所述的表达载体还包含外源基因的3'端非翻译区域,即包含聚腺苷酸信号和任何其它参与 mRNA 加工或基因表达的 DNA 片段。所述聚腺苷酸信号可引导聚腺苷酸加入到 mRNA 前体的3'端,如农杆菌冠瘿瘤诱导(Ti)质粒基因(如胭脂合成酶 Nos 基因)、圆红冬孢酵母基因(如 R. toruloides 的 G3PDH 基因)3'端转录的非翻译区均具有类似功能。为了便于对转基因细胞系或重组菌株进行鉴定及筛选,可对所用载体进行修饰,如引入可在细胞中表达的编码产生颜色变化的酶(如绿色荧光蛋白)或发光化合物的基因(如 GUS 基因、荧光素酶基因等)、具有抗性的抗生素标记基因(如卡那霉素标记基因、博来霉素标记基因、潮霉素标记基因)或抗化学试剂标记基因(如抗除草剂基因)、以及营养筛选标记基因(如 LEU2、URA3、URA3、LYS2、LYS5、MET15)等。

[0022] 本发明的另一个目的是提供一种构建油脂生产重组细胞的方法,所述方法包括:将编码本发明的产油酵母 ATP 柠檬酸裂解酶的核苷酸序列导入宿主细胞,得到油脂生产重组细胞。利用任何一种可以启动外源基因在产油酵母中表达的载体,将本发明所提供 rtacl 导入产油酵母细胞中,得到油脂生产重组酵母。或者利用任何一种可以启动外源基因在产油细菌中表达的载体,将本发明所提供 rtacl 导入产油细菌细胞中,得到油脂生产重组细菌。。或者利用任何一种可以启动外源基因在产油微藻中表达的载体,将本发明所提供 rtacl 导入产油微藻中,得到油脂生产重组微藻菌株。

[0023] 实验证明,本发明提供的产油酵母 ATP 柠檬酸裂解酶可显著提高重组细胞的油脂含量。

[0024] 综上所述,本发明提供下述:

[0025] 1. 一种 ATP 柠檬酸裂解酶,所述 ATP 柠檬酸裂解酶是由如下(a)或(b)的蛋白质:

[0026] (a) 由 SEQ ID NO :3 所示的氨基酸序列组成的蛋白质;

[0027] (b) 将 SEQ ID NO :3 的氨基酸经过一个或几个氨基酸的取代和/或缺失和/或添加且与具有油脂贮运和代谢调控相关活性的由 SEQ ID NO :3 衍生的蛋白质。

[0028] 2. 编码第 1 项所述的 ATP 柠檬酸裂解酶的核苷酸序列。

[0029] 3. 根据第 2 项所述的核苷酸序列,其为 SEQ ID NO :1。

- [0030] 4. 包含第 2 项或第 3 项所述的核苷酸序列的重组载体。
- [0031] 5. 一种重组细胞,所述细胞导入第 4 项所述的重组载体。
- [0032] 6. 一种构建油脂生产重组细胞的方法,所述方法包括:将第 2 项或第 3 项所述的核苷酸序列导入宿主细胞中,得到油脂生产重组细胞。
- [0033] 7. 根据第 6 项所述的方法,其中通过用第 4 项所述的重组载体转化宿主细胞而导入第 2 项或第 3 项所述的核苷酸序列。
- [0034] 8. 根据第 6 项所述的方法,其中所述宿主细胞选自酵母、微藻、红球菌或大肠杆菌。
- [0035] 9. 根据第 8 项所述的方法,其中所述酵母选自酿酒酵母 (*Saccharomyces cerevisiae*) 或圆红冬孢酵母 (*R. toruloides*),所述微藻选自衣藻或聚球藻,所述红球菌是浑浊红球菌 (*R. opacus*)。
- [0036] 10. 根据第 9 项所述的方法,其中所述衣藻是莱茵衣藻 (*C. reinhardtii*)。
- [0037] 序列表说明
- [0038] SEQ ID NO :1ATP 柠檬酸裂解酶基因 Rt-ACL (cDNA 序列)
- [0039] SEQ ID NO :2ATP 柠檬酸裂解酶基因的基因组 DNA 序列
- [0040] SEQ ID NO :3ATP 柠檬酸裂解酶的氨基酸序列

附图说明

- [0041] 图 1 为扩增 Rt-ACL cDNA 的 RT-PCR 结果,扩增条带约为 3.5kb。M 为分子量标准,泳道 1 为 *R. toruloides* Rt-ACL cDNA 扩增产物。
- [0042] 图 2 为 Rt-ACL cDNA 与其它物种 ATP 柠檬酸裂解酶同源序列多重比对的分析结果。
- [0043] 图 3 为扩增 Rt-ACL gDNA 的 RT-PCR 结果,扩增条带约为 4.3kb。M 为分子量标准,泳道 1 为 *R. toruloides* Rt-ACL gDNA 扩增产物。
- [0044] 图 4 为 Rt-ACL cDNA 在酿酒酵母中表达载体 pYX212-Rt-ACL 的构建。
- [0045] 图 5 为在酿酒酵母中测定 *rtac1p* 活性结果。每组数据均为三株重组菌株测定平均值。pYX212 为仅表达 pYX212 空载体的酿酒酵母菌株,pYX212-ACL 为表达 pYX212-ACL 的酿酒酵母菌株。

具体实施方式

- [0046] 以下实施例将有助于本领域的普通技术人员进一步理解本发明,但不以任何形式限制本发明。下述实施例中的实验方法,如无特殊说明,均为常规方法。下述实施例中所用的实验材料,如无特殊说明,均为自常规生化试剂公司购买得到的。
- [0047] 圆红冬孢酵母 (*R. toruloides*) CGMCC2.1389:购自中国普通微生物菌种保藏中心 (CGMCC, 保藏号为 CGMCC2.1389),源自东京大学应用微生物研究所 (IF08766),由 IF00559 和 IF00880 经配合而成的二倍体,等同于 CBS6016 或 NBRC8766 或 NRRL Y-6987。
- [0048] 圆红冬孢酵母 (*R. toruloides*) ATCC10788:购自美国典型培养物保藏中心 (ATCC),源自 CBS-KNAW fungal biodiversity centre,等同于 IF00559 或 JCM3792 或 CCRC20306 或 DBVPG6740 或 IAM13469 或 IGC4416 或 MUCL30249 或 NCYC921 或 NRRL Y-1091 或 VKM Y-334 或 PYCC4416。

[0049] 以下实施例中所涉及的培养基配方及用途如下：

[0050] (1) YEPD 培养基：酵母粉 10g/L, 蛋白胨 10g/L, 葡萄糖 20g/L, pH6.0, 固体培养基则再加入琼脂粉 15g/L；用于菌种活化培养、种子液制备和菌种短期保藏。

[0051] (2) 限氮培养基：葡萄糖 70g/L, 酵母粉 0.75g/L, (NH₄)₂SO₄ 0.1g/L, KH₂PO₄ 0.1g/L, MgSO₄·7H₂O 1.5g/L, pH5.6, 并补加 1% (V/V) 的微量元素液体 (4.0g/L CaCl₂·2H₂O, 0.55g/L FeSO₄·7H₂O, 0.52g/L citric acid·H₂O, 0.10g/L ZnSO₄·7H₂O, 0.076g/L MnSO₄·H₂O 和 100 μl 18MH₂SO₄), 用于菌株油脂积累和富含脂滴菌体的培养。

[0052] 实施例 1：圆红冬孢酵母的多组学研究

[0053] 圆红冬孢酵母 (*R. toruloides*) 全基因组从头测序：在 YEPD 培养基中培养 *R. toruloides* ATCC10788, 温度为 30℃, 转速 200rpm, 培养时间 48h。收集菌体, 按标准基因组 DNA 提取方法 [精编分子生物学实验指南 (第四版), 奥斯伯等著, 颜子颖等译, 科学出版社出版] 制备基因组 DNA, 委托华大基因利用 solexa 测序仪进行分析, 最终获得 94 个 scaffold (基因组构架), 基因组全长 20.2Mb。

[0054] 圆红冬孢酵母 (*R. toruloides*) 转录组测序：在 2L 发酵罐中 (初始培养基量 1.7L), 温度为 30℃, pH 为 5.6 的条件下恒化培养 *R. toruloides* ATCC10788, 通气速率为 100L/h, 溶氧保持为 73% 饱和 (溶氧与搅拌联动), YEPD 培养基流加速率为 0.828L/h (稀释率为 0.53) (反应器工作体积为 1.7L)。约 10 个工作体积后取样, 离心收集菌体, 立即于液氮中速冻, -70℃ 保存, 干冰低温保存运输至华大基因 (全名: 北京六合华大基因科技股份有限公司深圳分公司)。采用液氮研磨 -RNAiso 法提取总 RNA, RNA-seq 样品库的制备使用 mRNA-Seq8 Sample Preparation Kit (购自 Illumina, San Diego, CA, USA), 具体步骤参考产品使用说明。之后经 Illumina HiSeq2000 测序得到 90bp 的双末端序列。

[0055] 基于 *R. toruloides* 全基因组从头测序和转录组测序数据, 组合基因从头预测, 共注释了 8171 个蛋白编码基因。利用 Bioedit 软件建立 *R. toruloides* 本地基因组数据库和蛋白质组数据库。基因和蛋白命名规则如下: 根据注释基因所在 scaffold (基因组构架) 由大到小进行编码, 编号前统一加上 *R. toruloides* 的两个单词前两个字母缩写, 转录本和蛋白质后分别加 .t1 和 .t1p 的后缀。比如, scaffold1 上 5' 端起第一个编码基因命名为 RHT0_00001, 对应的转录本为 RHT0_00001.t1, 相应编码的蛋白质序列为 RHT0_00001.t1p。数据库包含 94 个 scaffold (基因组构架), 8171 个编码基因及相应 8171 个蛋白。

[0056] 实施例 2: 圆红冬孢酵母 ACL 基因的发现

[0057] *R. toruloides* ATCC10788 于 YEPD 固体培养基上复苏, 30℃ 倒置培养 48h, 挑单菌落接种于 50ml YEPD 液体培养基 (装液于 250ml 容量的三角瓶), 30℃, 200rpm 培养 28h。培养液分别以 1:50 (V/V) 比例转接于自动补料的 2L 连续发酵罐 (工作体积 1.7L), 培养基分别为 MM 和 MM-N。恒化培养温度为 30℃, pH 通过滴加 10.0M NaOH 或 2M HCl 自动控制在 5.6, 转速保持为 600rpm, 通气速率为 100L/h (约 0.98VVM), 溶氧保持为 85% 饱和, 稀释率为 0.085。约 15 个工作体积后取样, 样品经 4℃, 8000rpm 离心 5min 收集, 每 45ml 培养物能收集到约 0.7g 湿菌体, MM 和 MM-N 菌体样品均立即于液氮中速冻, -70℃ 保存。菌体样品用于数字基因表达谱分析, 以及胞内油脂含量分析, 培养上清用于残氮分析 [沈宏伟, 靳国杰, 胡翠敏, 龚志伟, 白凤武, 赵宗保. 生物工程学报 2012, 28(1), 57-65]。

[0058] 在实施例 1 基础上, 菌体样品委托华大基因提取 RNA 后进行 *R. toruloides* 在限氮

和非限氮培养条件下的数字基因表达谱分析。样品重复分析三次,数据分析基于三次重复分析中鉴定到的差异表达基因累加。

[0059] 结果显示,R. toruloides np11 使用 MM 和 MM-N 这两种培养基进行恒化培养时,油脂含量有显著的差别,限氮培养 (MM-N 培养) 条件下,细胞积累的油脂量为 33.3%,比非限氮培养 (MM 培养) 条件下提高了 22.8%。通过检测培养液上清中氮的浓度表明非限氮条件下氮浓度为 46.97mM,而限氮 (MM-N) 条件下,其氮的浓度已经低于检测限,表明菌体培养时采用了一种稳定的限氮与非限氮环境。

[0060] 剔除含 N 碱基和低质量数据后,数字基因表达谱分析得到 119192 (MM-N) 和 92644 条 (MM) 条差异的 clean Tag。在这些 clean Tag 的基础下进行了下面的分析,将 Tag 定位匹配到参考基因上,以获得基因的表达量,无法定位匹配到基因上的 Tag,则将其定位匹配到基因组上,最后均不能匹配的 Tag 则认为是未知的 Tag。过滤后得到 56846 个 MM Tag (总数 3454492 个) 和 60010 个 MM-N Tag (总数为 3600362 个)。经过归一化处理,和统计检验,选取表达丰度比 ≥ 2 ,控制假阳性 $FDR \leq 0.001$ 最终得到表达上调基因 1176 个,其中有三个表达丰度比 ≥ 10 的 Tag :CATGCCGCACTCGCCGGCCCC, CATGCAGGACATCGAGGTCGT, 和 ATCTTCATCCAGCCCTTCA, 分别来源于三个核苷酸序列长度为 3.8kb,8.8kb 和 3.4kb 的转录本 (R. toruloides 本地转录组数据库中编号为 RHTO_02032. t1, RHTO_02139. t1, 和 RHTO_03915. t1), 分别编码脂肪酸合酶 beta 亚基, alpha 亚基及 ATP 柠檬酸裂解酶蛋白。其中, RHTO_03915. t1 表达量随脂滴内油脂含量增加而增加,其编码的氨基酸序列如 SEQ ID NO :3 所示,命名为 rtac1p。其编码基因命名为 RT-ACL cDNA。

[0061] 实施例 3 :圆红冬孢酵母 AS2. 1389RT-ACL 基因组 DNA 的获取

[0062] 根据 Rt-ACL 的核苷酸序列设计基因特异性引物如下 :

[0063] ACL-orf-P1 :5' -ATGTCGGCAAAGCCTATCCGCGA-3'

[0064] ACL-orf-P2 :5' -CTACTGGCGCTGCTGGACGAGGAT-3'

[0065] 按照“精编分子生物学实验指南 (第四版) (奥斯伯等著,颜子颖等译,科学出版社出版)”描述方法提取圆红冬孢酵母 AS2. 1389 的基因组 DNA,紫外 / 可见光光谱分析测得样品 $OD_{260}/OD_{280} = 1.88$,表明基因组 DNA 质量很好。冻存于 -20°C ,备用。

[0066] 以圆红冬孢酵母 AS2. 1389 的基因组 DNA 为模板,利用 ACL-orf-P1 和 ACL-orf-P2 两条引物,按照常规方法进行 PCR 扩增,得到约 4.3kb 的 PCR (图 3)。PCR 扩增产物按照实施例 2 的操作步骤回收、克隆到 pMD18-T 载体中,并进行测序,测序结果表明,所扩增到的核苷酸序列为 SEQ ID NO :2,长度为 4428bp。将 SEQ ID NO :2 所示的核苷酸命名为 RT-ACLgDNA,获得的阳性重组质粒命名为 pMD18T-RTACLg。

[0067] 实施例 4 :圆红冬孢酵母 RT-ACL 基因的功能验证

[0068] 根据 RT-ACL 的核苷酸序列设计以下引物,用于构建 RT-ACL 在酿酒酵母中表达的载体 (引物 ACL-PYX-BamH I 的下划线部分为 BamH I 酶切位点,引物 ACL-PYX-HindIII 的下划线部分为 HindIII 酶切位点) :

[0069] ACL-PYX-BamH I :5' -TCAGGGGATCCAAAAATGTCGGCAAAGCCTATCCGCGA-3'

[0070] ACL-PYX-HindIII :5' -ACTGAAGCTTCTAATGGTGATGGTGATGATGCTGGCGCTGCTGGACGAGG-3'

[0071] 步骤一 RT-ACL 酵母重组载体 pYX212-Rt-ACL 的构建

[0072] 采用 ACL-PYX-BamH I 和 ACL-PYX-HindIII 两个引物扩增片段,经 BamHI/HindIII 双酶切后连入同样酶切的 pYX212(购自 Biovector)。将经 BamHI/HindIII 酶切鉴定可释放出约 3.5kb 插入片段的重组质粒送 TaKaRa 测序,测序正确的重组质粒命名为 pYX212-ACL(图 4)。

[0073] 步骤二 RT-ACL 重组酵母表达菌株的构建

[0074] 将 pYX212-ACL 转化至 *Saccharomyces cerevisiae* INVSc1(购自 Invitrogen, 基因型: MATa/ α , his3 Δ 1/his3 Δ 1, leu2/leu2, trp1-289/trp1-289, ura3-52/ura3-52) 中。转化方法如下:单个酵母菌落接种于 5ml YPD 培养基(20g/L 蛋白胨,10g/L 酵母提取物和 20g/L 葡萄糖, pH6.0),过夜培养;以 1:50 接种于 100ml 新鲜的 YPD 培养基,培养 8h 左右,此时 OD 值约为 1.0-1.2,冰浴 15min;4 $^{\circ}$ C,2000 \times g 离心 10min 收集菌体,悬浮于 50ml 冰冷的 ddH₂O,2000 \times g 离心 10min 离心收集菌体,然后悬浮于 20ml 冰冷的 1M 山梨醇,2000 \times g 离心 10min 离心收集菌体,倒尽溶液,菌体悬浮于 0.5-1.0ml 冰冷的山梨醇,此时 OD 值约为 100-200。50 μ l 电感受态细胞中加入不多于 5 μ l 质粒 DNA(0.5-1 μ g),冰上放置 10min,转移至冰冷的电转杯中,1500V 电压转化,此条件下电击时间约为 5ms,电击完加入 1ml 冰冷的山梨醇,于 30 $^{\circ}$ C 摇床温浴 2h,涂于 SC-Uracil 平板(20g/L glucose,6.7g/L YNB W/O amino acids but with ammonium sulfate(无氨基酸酵母基础氮源,含硫酸铵,购自 BD Difco,货号:291920),1.9g/L Yeast Synthetic Drop-out Medium Supplements(购自 sigma,货号 Y1501)),30 $^{\circ}$ C 培养 3-4 天至长出转化子。转化子接种于 SC-Ura 液体培养基(0.67% YNB W/O amino acids but with ammonium sulfate(同上),2%葡萄糖,0.01%亮氨酸,0.01%色氨酸,0.005%组氨酸)过夜,3-4ml 菌液离心收集菌体,玻璃珠破碎细胞,提取质粒 DNA 转化 *E. coli* DH5a,转化子提取质粒, BamHI/HindIII 酶切验证正确的质粒来源重组菌则命名为 INVSc1-pYX212-ACL。

[0075] 步骤三酶液的制备

[0076] 将正确的重组菌株 INVSc1-pYX212-ACL 接种于 5ml SC-Ura 液体培养基作为种子夜,培养过夜。而后以 1:100(v/v) 的接种量接入 100ml SC-Ura 液体培养基中,至 OD 值达到 4.2。

[0077] 将菌液离心弃上清,并用同体积蒸馏水洗涤一遍。用提取缓冲液(50mM Tris-HCl(pH8.0)、1mM EDTA、10mM DTT、1mM PMSF)5ml 悬浮,使用费氏压碎器用 12kPa 破碎菌体。20,000g,4 $^{\circ}$ C 离心分离 15min,收集上清。于填充了 Shehadex G-25 的 PD-10 色谱柱(GEHEALTHCAREBIO-SCIENCES 公司),用洗脱缓冲液(10mM 磷酸钠(pH7.4)、1mM MgCl₂、0.1mM EDTA、1mM DTT)洗脱,洗脱液作为酶液。

[0078] 步骤四 ACL 活性测定

[0079] ACL 的活性是通过其催化柠檬酸生成的草酰乙酸的量,进一步由苹果酸脱氢酶催化草酰乙酸与 NADH 反应过程中 NADH 减少的量来求出的。NADH 的变化量可以通过在 A₃₄₀ 处的变化来求出。

[0080] 反应式如下:

[0081] (1) 柠檬酸 + CoA + ATP \rightarrow 草酰乙酸 + ADP + Pi + 乙酰 CoA

[0082] (2) 草酰乙酸 + NADH \rightarrow 苹果酸 + NAD⁺

[0083] 反应液的组成为 10mM Tris-HCl(pH8.4)、10mM MgCl₂、1mM DTT、10mM ATP、10mM 柠

檬酸、0.2mM CoA、6 单位苹果酸脱氢酶、0.1mMNADH、50 μ l 酶液,总量为 1ml。添加 CoA 使反应开始,反应在 28°C 下进行。

[0084] 结果如图 5 所示,可以看出仅在酿酒酵母中表达空载体 pYX212 的菌株不能测定到 ACL 的活性,相反,表达了 pYX212-ACL 的菌株显示出 ACL 的活性,并且具有显著性。因此我们可以判定表达的 Rt-ACL 蛋白能够在酿酒酵母中起到分解柠檬酸裂解为草酰乙酸和乙酰 CoA 的作用,即具有 ATP 柠檬酸裂解酶的活性。

[0085] 实施例 5 :Rt-ACL 表达重组圆红冬孢酵母菌株的构建

[0086] 根据 Rt-ACL 的核苷酸序列设计以下引物(引物 FBA-F-SpeI 的下划线部分为 SpeI 酶切位点,引物 ACL-R-EcoRV 的下划线部分为 EcoRV 酶切位点):

[0087] FBA-F-SpeI :5' -ACTGACTAGTCTCTGCTCTCGCTCGCTGTGGCTTG-3'

[0088] FBA-R :5' -GATTCCACGCACCTTTGCCGACATTGTAGCTAGTTAGTGTTAGAAGTG-3'

[0089] ACL-F :5' -CACTTCTAACACTAACTAGCTACAATGTCGGCAAAGGTGCGTGGGAATC-3'

[0090] ACL-R-EcoRV :5' -CCGTGATATCCTACTGGCGCTGCTGGACGAGG-3'

[0091] 引物方向均为 5' 至 3' 方向。

[0092] 以实施例 3 中构建的克隆载体 pMD18T-RTACLg 为模板,利用引物 ACL-F 和 ACL-R-EcoRV 扩增 RT-ACL 序列,另以圆红冬孢酵母 AS2.1389 基因组为模板,采用引物 FBA-F-SpeI 和 FBA-R 扩增 FBA 启动子区域。两个 PCR 片段利用文献 [Zhou YJ, Gao W, et al. J. Am. Chem. Soc 2012, 134 :3234-3241.] 中的方法进行融合。获得的融合片段经 EcoRV/SpeI 双酶切,连接入同样双酶切的 pRH203 [Liu Y, Koh CM, Sun L, Hlaing MM, Du M, Peng N, Ji L. Appl Microbiol Biotechnol. 2013 Jan ;97 (2) :719-729] 载体中。连接产物转化 E. coli DH5a 化学感受态细胞,挑选链霉素抗性转化子进行增菌培养、质粒提取(碧云天质粒小量提取试剂盒,货号:D0003)。重组质粒样品送至 TaKaRa 公司测序,经测序验证正确构建的重组质粒命名为 pRH203-RTACLg。

[0093] 将 pRH203-RTACLg 转化农杆菌 AGL1 [Lazo GR, Stein PA, Ludwig RA. Biotechnology (N Y). 1991, 9(10), 963-967.]。挑取壮观霉素抗性转化子,利用引物 ACL-F 和 ACL-R-EcoRV 进行 PCR 鉴定,阳性重组子命名为 AGL1/pRH203-RTACLg。

[0094] 将重组农杆菌 AGL1/pRH203-RTACLg 按 Yanbin Liu 等人方法转化导入圆红冬孢酵母 R. toruloides ATCC10788,挑取潮霉素抗性转化子进行 PCR 鉴定,阳性重组子命名为 R. toruloides ATCC10788RTACLg [Liu Y, Koh CM, Sun L, Hlaing MM, Du M, Peng N, Ji L. Appl Microbiol Biotechnol. 2013 Jan ;97 (2) :719-729]。

[0095] 将出发菌株圆红冬孢酵母 R. toruloides ATCC10788 和重组圆红冬孢酵母 R. toruloides ATCC10788RTACLg 分别接种 YEPD 培养基,30°C,200rpm,培养 3d,每隔 24h 取样,留做油脂含量分析和荧光观测 [Wu SG, Zhao X, Shen HW, Wang Q, Zhao ZK. Bioresour. Technol. 2011, 102(2), 1803-1807.]。结果发现,在不利于油脂积累的 YEPD 培养条件下,与对照菌株 ATCC10788 相比,重组圆红冬孢酵母 ATCC10788RTACLg 胞内油脂含量由 14% 增加到了 35%,增加了 2.5 倍。这证明 RL-ACL 基因可促进圆红冬孢酵母油脂含量显著增加。

[0096] 实施例 6 :Rt-ACL 重组浑浊红球菌的构建

[0097] 根据 Rt-ACL 的编码核苷酸序列设计以下引物(引物 ACL-F-BamHI 的下划线部分为 BamHI 酶切位点,引物 ACL-R-XbaI 的下划线部分为 XbaI 酶切位点):

[0098] ACL-F-BamHI :5' -ATGCGGATCCATGTTCGGCAAAGGTGCGTGGGAATC-3'

[0099] ACL-R-XbaI :5' -TGACTCTAGACTACTGGCGCTGCTGGACGAGGAT-3'

[0100] 引物方向均为 5' 至 3' 方向。

[0101] 以实施例 2 中构建的克隆载体 pMD19T-RTACLc 为模板,利用引物 ACL-F-BamHI 和 ACL-R-XbaI 扩增 RT-ACL 编码区序列,PCR 片段经 BamHI/XbaI 双酶切,利用 DNA 回收试剂盒(购自北京舟鼎国,货号:NEP013-2),按照供应商建议步骤纯化酶切大片段;pJAM2 质粒 [Triccas, J. A., T. Parish, W. J. Britton, and B. Giquel. FEMS Microbiol. Lett. 1998, 167, 151-156] 利用 BamHI/XbaI 双酶切后回收大片段。将 RT-ACL cDNA 酶切大片段 pJAM2 酶切大片段按 5 : 1 的摩尔比混合,利用 T4 DNA 连接酶连接环化,转化 E. coli DH5a 化学感受态细胞,挑选 Amp 抗性转化子进行增菌培养、质粒提取(碧云天质粒小量提取试剂盒,货号:D0003)。重组质粒样品送至 TaKaRa 公司测序,经测序验证正确构建的重组质粒命名为 pJAM-RL-ACLc。

[0102] 利用电击转化方法,将 pJAM-RL-ACLc 质粒转化至浑浊红球菌 (*R. opacus*) ATCC51881(购自 ATCC:American Tissue Culture Colection) 感受态细胞,挑取抗性转化子进行 PCR 鉴定,PCR 鉴定阳性转化子命名为 PD630/pJAM-RL-ACLc [Kalscheuer R, M. Arenskötter, and A Steinbüchel. Appl Microbiol Biotechnol 1999, 52, 508-515.]。

[0103] 将重组浑浊红球菌 PD630/pJAM-RL-ACLc 接种于含 0.01g/1 NH₄Cl 和 10g/1 葡萄糖酸钠的 MSM 培养基 [Schlegel H G, H. Kaltwasser, and G. Gottschalk. Arch. Mikrobiol. 1961, 38, 209-222.], 28 °C, 200rpm, 培养 24h, 每隔 6h 取样,留做油脂含量分析 [Jan Hänisch, Marc Wältermann, Horst Robenek, Alexander Steinbüchel. Appl. Environ. Microbiol. 2006, 72(10), 6743.]。结果发现,与对照菌株 *R. opacus* PD630 相比,重组浑浊红球菌 *R. opacus* PD630/pJAM-RL-ACLc 胞内油脂含量由 30% 增加到了 74%。证明 RL-ACL 基因可促进浑浊红球菌油脂含量显著增加。

[0104] 实施例 7 :Rt-ACL 重组莱茵衣藻的构建

[0105] 根据 rldp1 的核苷酸序列设计以下引物(引物 ACL-F-XbaI 的下划线部分为 XbaI 酶切位点,引物 ACL-R-NotI 的下划线部分为 NotI 酶切位点):

[0106] ACL-F-XbaI :5' -ATGCTCTAGAATGTTCGGCAAAGGTGCGTGGGAATC-3'

[0107] ACL-R-NotI :5' -TGACGCGGCCGCCTACTGGCGCTGCTGGACGAGGAT-3'

[0108] 引物方向均为 5' 至 3' 方向。

[0109] 以实施例 2 中构建的克隆载体 pMD19T-RTACLc 为模板,利用引物 ACL-F-XbaI 和 ACL-R-NotI 扩增 RT-ACL 编码区序列,PCR 片段经 XbaI/NotI 双酶切,连接入同样双酶切的 pChlamy_1(购自 Invitrogen,货号:A14258) 载体(回收酶切大片断用于连接),转化 E. coli DH5a 化学感受态细胞,挑选 Amp 抗性转化子进行增菌培养、质粒提取(碧云天质粒小量提取试剂盒,货号:D0003)。重组质粒样品送至 TaKaRa 公司测序,经测序验证正确构建的重组质粒命名为 pChlamy_1-RLACLc。利用 ScaI 将重组质粒 pChlamy_1-RLDP1 线性化,再利用 DNA 回收试剂盒(购自北京舟鼎国,货号:NEP013-2),按照供应商建议步骤(NEP013-2 说明书)纯化 PCR 产物,-20 °C 保存备用。

[0110] 莱茵衣藻 (*C. reinhardtii*) ATCC-50049(购自 ATCC:American Tissue Culture Colection) 利用 GibcoR TAP 培养基(购自 Invitrogen,货号:A1379801) 培养至 OD750 值

为 0.5,按照试剂盒说明书(货号:A14258)制备莱茵衣藻 *C. reinhardtii* CC-849 电击感受态细胞,取 250 μ l 感受态细胞,加入 2 μ g 前述线性化的 pChlamy_1-RLACLc 重组质粒,室温下温育 5min,按照如下参数(电压 600V,电容 50 μ F,电阻无穷大),利用 Bio-RadGene Pulser II 进行电击转化。转化后立即加入 5ml TAP-40mM 蔗糖溶液(TAP 购自 Invitrogen,货号:A1379801,补加 40mM 蔗糖),于 6 孔培养板 28 $^{\circ}$ C,50 μ E m-2s-1 条件下复苏 24h,再 2000rpm 离心 5min 收集藻体,涂布于 TAP-agar-Hygromycin 平板(TAP 购自 Invitrogen,货号:A1379801,补加 2%琼脂粉,100 μ g/ml 潮霉素),28 $^{\circ}$ C,50 μ E m-2s-1 条件下培养 5d,挑取潮霉素抗性转化子,利用 ACL-F-XbaI 和 ACL-R-Not 进行 PCR 鉴定,阳性克隆命名为 CC-849/pChlamy_1-RLACLc。

[0111] 重组莱茵衣藻 CC-849/pChlamy_1-RLACLc 接种 TAP 限氮培养基(缩写为 TAP-N:Tris 碱 20mM, MgSO₄ · 7H₂O 0.83mM, CaCl₂ · 2H₂O 0.45mM, K₂HPO₄ 0.165mM, KH₂PO₄ 0.05mM, 冰醋酸 0.1% (V/V), 痕量元素溶液 1ml/1 培养基,痕量元素溶液配方:Na₂EDTA · 2H₂O 5g/100ml, ZnSO₄ · 7H₂O 2.2g/100ml, H₃BO₃ 0.14g/100ml, MnCl₂ · 4H₂O 0.5g/100ml, FeSO₄ · 7H₂O 0.5g/100ml, CoCl₂ · 6H₂O 0.16g/100ml, CuSO₄ · 5H₂O 0.16g/100ml, (NH₄)₆Mo₇O₂₄ · 4H₂O 0.11g/100ml) 28 $^{\circ}$ C,持续光照 70 μ E m-2s-1 条件下培养 4d,每隔 24h 取样,留做油脂含量分析和荧光观测 [Moellering EricR., Benningl Christoph, Eukaryotic Cell 2010,9,97-106]。结果发现,与对照菌株莱茵衣藻 CC-849 相比,重组莱茵衣藻 CC-849/pChlamy_1-RLACLc 胞内油脂含量由 18%增加到了 39%。证明 RL-ACL 基因可促进莱茵衣藻油脂含量显著增加。

[0112] 应该理解,尽管参考其示例性的实施方案,已经对本发明进行具体地显示和描述,但是本领域的普通技术人员应该理解,在不背离由权利要求书所定义的本发明的精神和范围的条件下,可以在其中进行各种形式和细节的变化,可以进行各种实施方案的任意组合。

[0001]

序列表

<110> 中国科学院大连化学物理研究所
 <120> 一种产油酵母 ATP 柠檬酸裂解酶及其编码基因与应用
 <130> IB130263
 <160>
 <170> PatentIn version 3.5
 <210> 1
 <211> 3474
 <212> DNA
 <213> *Rhodospiridium torulooides* CGMCC 2.1389

 <400> 1
 atgtcggcaa agcctatccg cgagtagcag gccaaagttgc tcctcgccta ccacctcgcc 60
 cgcgcaaccga ccgccggctc caaggcagtc gcacgggacg gcttccagtc ccccgaagtc 120
 aaggttgctc aggtctcgtg ggaccccgag accaaccagg tcacgcctga tgcggcgtg 180
 cctcaactggg tcttcaccga gaagctcgtc gtcaagcccg accagctcat caagcgtcgt 240
 ggaaaggccg gtctcctcgc actgaacaag acgtgggccc agggcaagca gtggatcgt 300
 gagcgcgcgg gcaagcaagt tcaggtcag aagacgacgg gcacgctcaa caacttcac 360
 gtcgagccgt tctgcctca cccttcggac gccgagtact acatctgcat caactcggtc 420
 cgcgagggcg acgtgatcct gttcaccac gagggcgggtg tcgacgtcgg cgatgtcgac 480
 gccaaaggcg tcacgctcct cgtcccgcgc ggccggcagc tcccctcgcg cgacgagatc 540
 cgctcgcagc tcctcaagca cgtcacggc gccgaacgcc aggaggccct catcgactac 600
 atcatecgcc tctactcggg ctacgtcagc ctccactttg cctacctega gatcaaccgg 660
 ctcgtcgcgg tcgagaaacc ctcgaccggc aagaccgaca tctttctacct cgacatggcc 720
 gccaaagctc accagacggc cgagtagctc gtcggtecca agtgggcat cgccgcgcac 780
 cctccatca tcaacccgc cgcgcgccc atgtcgaacg gcaagatctc ggccgacaag 840
 ggcccggcca tgtttctggc cctcccttt ggtcgtgacc tcacgaagga ggaggcctac 900
 atcgccaagc tcgacggctc gaccggcgc tcgctcaagc tcaccgtctt gaacgtcag 960
 ggacgcatct ggaccatggt tgccggctgt ggtgcttcgg tcgtctactc ggacgccatc 1020
 gcccgccacg gattcgtca cgagctcgc aactacggcg aatactcggg cgcccccacc 1080
 cagaccaga cctacgagta cgccaagacc atcctcgaact tgatgaccgg cggcacgcc 1140
 aaccgcgagg gcaagctcct ctctcctcgc ggccgtatcg ccaacttcac caacgtcgc 1200
 gcgaccttca agggcatcat cacggcgcctc aaggagtacc agcacctctc tcaggagcac 1260
 aaggctcgc tctttgttcg ccgcccgggc cccaactacc agggaggcct caaggccatg 1320
 cgcctcctcg gcgagacct cggcgtcag atccaggtct ttggccccga gaccacatc 1380

[0002]

acctcgatcg	tcccgcctcgg	actcggcttg	atcaagtcgg	tgcacgacgc	tctcaaggtc	1440
cccggcgccc	gcgcccgcgc	cgatgcgacc	ggcaccctca	cccccgccc	gggctcggcc	1500
aagtcgcgcg	ccgcccagct	cccgaccggc	gcgtcgacgc	cctcggccca	gcagcccag	1560
gacaacatcg	tcagcttctc	ggacaaggtc	cacgcgcccg	actcgggtcg	cccgtggtac	1620
cgcccctttg	acgagacgac	gcgctcgatc	gtgtaeggcc	tocagcctcg	cgcgatccag	1680
ggcatgctcg	actttgactt	tgctgcgga	cgcgagacgc	cctcggtcgc	ggccatggtc	1740
taccctctcg	gcggccacca	cgccagaag	ttctactggg	gaaccaagga	gacctcctc	1800
cccgtcttca	cctcgatgaa	ggaggccgtc	gccaaagtgc	ccgacgcga	cgtcgctgtc	1860
aactttgcgt	cgtcgcgctc	ggtctaccag	tgcaccctcg	aggcgtcga	gttccccag	1920
atcaaggcca	tgcacctcat	cgccgagggt	gtcccgcgac	gccacgctcg	cgagatcctc	1980
cacctcgcca	agaagaagga	ggtcatcacc	atcggctccc	cgaccgtcgg	cgccatcaag	2040
cccggctgct	tcggtatcgg	caacacgggc	ggcatgaacg	agaacatcct	ctcgtccaag	2100
ctgtaccgcg	ccggttccgt	cggtaccgtc	tccaagtccg	gaggcatgtc	gaacgagctc	2160
aacaacatcc	tctcgtctac	gaccgacggc	gcgtacgagg	gcacgcccac	cgccgggtgac	2220
cgctaccggg	gcaccacctt	catcgaccac	ctcctccgct	acgaggccga	cccgaactgc	2280
aagatgctcg	tcctcctcgg	agaggtcggt	ggcgtcgagg	agtaccgctg	catcgaggcc	2340
gtcaagtcgg	gccagatcaa	gaagcccacc	gtcgcgtggg	ccatcggcac	ctgcgccaaag	2400
atggtcgcca	ccgacgtcca	gttcggccac	gccggctcga	tggccaactc	ggacctcgag	2460
acggccgagg	ccaagaacaa	cgccatgccc	gccgccggct	tcacgtgccc	cccgaacttc	2520
gaggagctcc	cgcaggtcct	cgccgagacc	taccagaagc	tcgtcggcga	cggcacgata	2580
cagcccaagg	ccgaggttcc	ccctcctcag	atcccgatgg	actacaactg	ggcccagacg	2640
ctcggcatgg	tccgcaagcc	cgccgccttc	atctogacca	tctcggacga	gcgcggccag	2700
gagctcctct	acgcccggcat	gcccattctc	aaggtctttg	aggaggacat	cgccatcggc	2760
ggtgtcgtct	cgctcctctg	gttcaagcgt	cgctcctccg	cctacgcgac	caagttcctc	2820
gagatggctc	tcatgctcac	ggccgaccac	ggaccgcggc	tctcgggcgc	catgaccacc	2880
gtcatcacca	cccgtgccgg	caaggacctc	gtctcgtcgc	tcgtcgcggg	tctcctcacc	2940
atcggcgacc	gcttcgggtg	cgcgctcgac	ggcgccggc	aggagtttac	gcgcgcgttc	3000
gaggctggcc	tcacgccccg	cgagtttgtc	gactcgatgc	gcaaggccaa	caagctcacc	3060
cccgttatcg	gacacaaggt	caagtccaag	gccaaaccgg	acaagcgtgt	cgagctcgtc	3120

[0003]

aagaactacg tcttcaagca cttcccgctc gccaaagctcc tcgagtagcc tctcggcgtc 3180
gaggacgtca cgagcgccaa gaaggacacg ctcatcctca acgtcgacgg agcaatcgcg 3240
gtttctttcg tagatttact taagaactcc ggcgcgttca ccgccgagga ggtgcccag 3300
tacatgaaga tcggcacctt gaacggcttg ttcttctcgc gacgctccat cggttcatt 3360
gccaccacc tcgaccagaa ggcctcaag cagccgctct accgtcacc gcccgacgac 3420
atcttcatcc agcccttcaa caccgaccgc atcctcgtcc agcagcgcca gtag 3474

<210> 2
<211> 4428
<212> DNA
<213> *Rhodospiridium toruloides* CGMCC 2.1389

<400> 2
atgtcggcaa aggtgcgtgg aatcgcttc cgccaacctc gagaaacgac tttcctgact 60
ttgtccccc cctggcctc tctgcctga tccctgatgc cctgctcac actctcgacg 120
accacgaac ccgcccgag cctatcccg agtacgacgc caagttgctc ctgcctacc 180
acctcgccc cgaccgacc gccggtcca aggcagtcgc acgggacggc ttcagtccc 240
ccgaagtcaa ggttgcgag gtacgttcat ctacgagcc tacaccgct ctgaggagga 300
cgagacgagg ttgagtcgtg ccggatgggg accgggtttc gatcaggctg gatcttgctc 360
acacaacgct ggatggactg gatactgctg gacagtcgct acgtagactc tggggaactt 420
tctacacgct cgagcgggca ccagagcgcg tcgtcggcgt tcgcgactcg ctgcccacc 480
ccaccaccat cgacgggtc ccccaaact cgtcccac acccgtgctg acctcgtatg 540
cgtgcatgga caggtctcgt gggaccccga gaccaaccag gtcacgcctg atcgggcgtc 600
gcctcaetgg gtctcaccg agaagctcgt cgtcaagccc gaccagctca tcaagcgtc 660
tggaaaggcc ggtctcctgt gcgccgacg ctgttctcgt cgtcctgtag catcttgctg 720
actccctcga cgtcgcagc gcaactgaaca agacgtgggc cgagggcaag cagtggatcg 780
ctgagcgcgc gggcaagcaa gttcaggtgc gtcacatct tccacctcgc ctctttctg 840
cagaaaggac acgcgactg aactgcgaa cgaacgcag tgcgcaggtc gagaagacga 900
cgggcacgct caacaacttc atcgtcagc cgttctgccc tcaccttcg gacgccagtc 960
actacatctg catcaactcg gtccgcgagg gcgacgtgat cctgttcacc cacgagggcg 1020
gtgtcgacgt cggcgatgtc gacgccaaag cgtcacgct cctcgtcccc gtcggcggcg 1080
agctcccctc gcgcgacgag atccgctcgc agctcctcaa gcacgtcacc ggcgcgaa 1140
gccaggaggc cctcatcgac tacatcatcc gcctctactc ggtctacgtc gacctcact 1200

[0004]

ttgcctacct	cgagatcaac	ccgctcgtcg	ccgtcgagaa	cccctcgacc	ggcaagaccg	1260
acatcttcta	cctcgacatg	gccgccaaagc	tcgaccagac	ggccgagtac	gtcgtcggtc	1320
ccaagtgggc	catcgcgcgc	gacccctcca	tcataaacc	cgccgcgcgc	cccatgtcga	1380
acggcaagat	ctcggccgac	aagggcccgc	ccatgtttcg	gccccctccc	tttggtcgtg	1440
acctcaccga	ggaggaggcc	tacatcgcca	agctcgacgg	ctcgaccggc	gcctcgetca	1500
agctcaccgt	cttgaacgct	gagggacgca	tctggaccat	ggttgccggt	ggtggtgctt	1560
cggtcgtcta	ctcggacgcc	atcgccgcgc	acggattcgc	tcacgagctc	gccaaactacg	1620
gcgaatactc	ggggcgcgcc	accagacccc	agacctacga	gtacgccaaag	accatccgtg	1680
cgtttcggcc	ttctcttttc	cgctctcaga	ccgcctcgct	aacgctcgac	ggtctctctg	1740
agtcgacttg	atgaccgcgc	gcacgcccac	cccgcagggc	aagctcctct	tcatcggcgg	1800
cggtatcgcc	aacttcacca	acgtcgcgcg	gaccttcaag	ggcatcatca	cggegetcaa	1860
ggagtaccag	caccgtcttc	aggagcaca	ggtccgcata	tttgttcgcc	gcggcggccc	1920
caactaccag	gagggcctca	aggccatgcg	ccccttcggc	gagaccctcg	gcgtcgagat	1980
ccaggtcttt	ggccccgaga	cccacatcac	ctcgatcgtc	ccgctcggac	tcggcttgat	2040
caagtgggtc	gacgacgctc	tcaaggtccc	cggcgcgccg	gccgcgcgcc	atgcgaccgg	2100
caacctcacc	cccgtcccgg	gctcgcacca	gtcgcgcgcc	gcccagctcc	cgaccggcgc	2160
gtcgcgcgcc	tcgcccagc	agccccagga	caacatcgtc	agcttctcgg	acaaggteca	2220
cgcgcccggc	tcgggtcgcc	cggtgtaccg	cccctttgac	gagacgacgc	gctcgatcgt	2280
gtacggcctc	cagcctcggc	cgatccaggc	catgctcgac	tttgactttg	cctgcggaag	2340
cgagacgccc	tcggtcgcgg	ccatggtcta	ccccttcggc	ggccaccacg	tcagaagtt	2400
ctactgggga	accaaggaga	ccctcctccc	cgcttccacc	tcgatgaagg	aggcctcgcg	2460
caagtgcctc	gacgcgcgac	tcgtcgtcaa	ctttgcgtcg	tcgctcggc	tcctaccagtc	2520
gaccctcgag	gcgctcgagt	tccccagat	caaggccatc	gccctcctcg	ccgagggtgt	2580
ccccgagcgc	cacgctcgcg	agatcctcca	ccctcgccaag	aagaaggagg	tcctcatcat	2640
cggtcccgcg	accgtcggcg	gcataaagcc	cggtgcttc	cgtatcggca	acaccggcgg	2700
catgaacgag	aacatcctct	cgccaagct	gtaccgcgcc	ggttcgcgtc	gctacgtctc	2760
caagtccgga	ggcatgtcga	acgagctcaa	caacatcctc	tcgctcacga	ccgacggcgc	2820
gtacgagggc	atcggcatcg	gcggtgaccg	ctaccggggc	accaccttca	tcgaccacct	2880
ccctcgcctac	gagggccgacc	cgaactgcaa	gatgctcgtc	ctcctcggag	aggctcggtg	2940
cgctcagggg	taccgcgtca	tcgaggccgt	caagtcgggc	cagatcaaga	agcccacgtt	3000

[0005]

cgcggtgggccc atcgggcacct ggcgccaagat gttcgccacc gacgtccagt tcggcccacgc 3060
 cggtctgatg gccaaactcgg acctcgagac ggccgaggcc aagaacaacg ccatgcgcgc 3120
 cgccggettc atcgtgcccc cgacgttcga ggagctcccg caggctctcg ccgagacctt 3180
 ccagaagctc gtcggcgacg gcacgatcca gcccagccc gaggttcccc ctctcagat 3240
 cccgatggac tacaactggg ccagacgct cggcatggtc cgcaagccc cgccttcat 3300
 ctcgaccatc tcggacgagc gcggccagga gctcctctac gccggcatgc ccatctccaa 3360
 ggtctttgag gaggacatcg gcatcggcgg tgtcgtctcg ctctctgggt tcaagcgtcg 3420
 cctccccgcc tacgggacca agttcctcga gatggctctc atgctcacgg ccgaccacgg 3480
 acccgcgcgc tcggggcgca tgaccacgct catcaccacc cgtgcgggca aggacctcgt 3540
 ctcgtcgcgc gtcgcccgtc tctcaccat cggcgaccgc ttcgggtggc cgcctcagcg 3600
 cgccgcgcag gaggttacgc gcgcttcga ggctggcctc accccccgcg agtttgctga 3660
 ctcgatgcgc aaggccaaca agctcatccc cggtatcggc cacaaggcca agtccaaggc 3720
 caaccgggac aagcgtgtcg agctcgtcaa gaactacgtc ttcaagcact tcccgtctgc 3780
 caagctcctc gagtacgctc tcgcccgcga ggacgtcacg agcgccaaga aggacacgct 3840
 catcctcaac gtcgacggag gtgagcctcc acttctcct cctcgacttt tccttgact 3900
 gacgcttctt cctcgtccag caatcgcggt ttctttcgta gatttactta agaagaacgc 3960
 tagtctcat gcgtaaaaa tcaggacgaa ggagaactaa cccacgcgcg ttccctcccg 4020
 cagctccggc gcgttcaccg ccgaggaggc tgccgagtac atgaaggaga gttcacgtaa 4080
 actgcaactc ctgtttctaca gcagaagcgc tgttctacgg cgttcacccc tcgdcgcgtc 4140
 gcctccccct cttegcgcgc acctcccgtt ctcgacttcc gctgacccta cccttcgctt 4200
 gcagatcggc accttgaacg gcttgttcgt tctcggacgc tccatcggct tcattgcca 4260
 ccacctcgac cagaagcgc tcaagcagcc gctctaccgt caccgcgcgc agtacgctcg 4320
 acctctctt cgcaccgcgc actttgcact gaccogctcc tcttcgcag cgacatctt 4380
 atccagccct tcaacaccga ccgcatcctc gtccagcagc gccagtag 4428

<210> 3

<211> 1157

<212> PRT

<213> *Rhodospiridium toruloides* CGMCC 2.1389

<400> 3

Met. Ser Ala Lys Pro Ile Arg Glu Tyr Asp Ala Lys Leu Leu Leu Ala

[0006]

1	5	10	15
Tyr His Leu Ala Arg Ala Pro Thr Ala Gly Ser Lys Ala Val Ala Arg	20	25	30
Asp Gly Phe Gln Ser Pro Glu Val Lys Val Ala Gln Val Ser Trp Asp	35	40	45
Pro Glu Thr Asn Gln Val Thr Pro Asp Ala Ala Leu Pro His Trp Val	50	55	60
Phe Thr Glu Lys Leu Val Val Lys Pro Asp Gln Leu Ile Lys Arg Arg	65	70	75
Gly Lys Ala Gly Leu Leu Ala Leu Asn Lys Thr Trp Ala Glu Gly Lys	85	90	95
Gln Trp Ile Ala Glu Arg Ala Gly Lys Gln Val Gln Val Glu Lys Thr	100	105	110
Thr Gly Thr Leu Asn Asn Phe Ile Val Glu Pro Phe Cys Pro His Pro	115	120	125
Ser Asp Ala Glu Tyr Tyr Ile Cys Ile Asn Ser Val Arg Glu Gly Asp	130	135	140
Val Ile Leu Phe Thr His Glu Gly Gly Val Asp Val Gly Asp Val Asp	145	150	155
Ala Lys Ala Leu Thr Leu Leu Val Pro Val Gly Gly Glu Leu Pro Ser	165	170	175
Arg Asp Glu Ile Arg Ser Gln Leu Leu Lys His Val Thr Gly Ala Glu	180	185	190
Arg Gln Glu Ala Leu Ile Asp Tyr Ile Ile Arg Leu Tyr Ser Val Tyr	195	200	205
Val Asp Leu His Phe Ala Tyr Leu Glu Ile Asn Pro Leu Val Ala Val			

[0007]

210	215	220
Glu Asn Pro Ser Thr Gly Lys Thr Asp Ile Phe Tyr Leu Asp Met Ala 225	230	235 240
Ala Lys Leu Asp Gln Thr Ala Glu Tyr Val Val Gly Pro Lys Trp Ala 245	250	255
Ile Ala Arg Asp Pro Ser Ile Ile Asn Pro Ala Ala Ala Pro Met Ser 260	265	270
Asn Gly Lys Ile Ser Ala Asp Lys Gly Pro Pro Met Phe Trp Pro Pro 275	280	285
Pro Phe Gly Arg Asp Leu Thr Lys Glu Glu Ala Tyr Ile Ala Lys Leu 290	295	300
Asp Gly Ser Thr Gly Ala Ser Leu Lys Leu Thr Val Leu Asn Ala Glu 305	310	315 320
Gly Arg Ile Trp Thr Met Val Ala Gly Gly Gly Ala Ser Val Val Tyr 325	330	335
Ser Asp Ala Ile Ala Ala His Gly Phe Ala His Glu Leu Ala Asn Tyr 340	345	350
Gly Glu Tyr Ser Gly Ala Pro Thr Gln Thr Gln Thr Tyr Glu Tyr Ala 355	360	365
Lys Thr Ile Leu Asp Leu Met Thr Arg Gly Thr Pro Asn Pro Gln Gly 370	375	380
Lys Leu Leu Phe Ile Gly Gly Gly Ile Ala Asn Phe Thr Asn Val Ala 385	390	395 400
Ala Thr Phe Lys Gly Ile Ile Thr Ala Leu Lys Glu Tyr Gln His Arg 405	410	415
Leu Gln Glu His Lys Val Arg Ile Phe Val Arg Arg Gly Gly Pro Asn		

[0008]

	420		425		430											
Tyr	Gln	Glu	Gly	Leu	Lys	Ala	Met	Arg	Leu	Leu	Gly	Glu	Thr	Leu	Gly	
	435						440					445				
Val	Glu	Ile	Gln	Val	Phe	Gly	Pro	Glu	Thr	His	Ile	Thr	Ser	Ile	Val	
	450					455						460				
Pro	Leu	Gly	Leu	Gly	Leu	Ile	Lys	Ser	Val	Asp	Asp	Ala	Leu	Lys	Val	
465					470					475					480	
Pro	Gly	Ala	Arg	Ala	Ala	Ala	Ala	Asp	Ala	Thr	Gly	Thr	Leu	Thr	Pro	Val
				485						490					495	
Pro	Gly	Ser	Pro	Lys	Ser	Arg	Ala	Ala	Gln	Leu	Pro	Thr	Gly	Ala	Ser	
			500						505						510	
Thr	Pro	Ser	Arg	Gln	Gln	Pro	Gln	Asp	Asn	Ile	Val	Ser	Phe	Ser	Asp	
	515						520							525		
Lys	Val	His	Ala	Pro	Asp	Ser	Gly	Arg	Pro	Trp	Tyr	Arg	Pro	Phe	Asp	
	530					535						540				
Glu	Thr	Thr	Arg	Ser	Ile	Val	Tyr	Gly	Leu	Gln	Pro	Arg	Ala	Ile	Gln	
545					550					555					560	
Gly	Met	Leu	Asp	Phe	Asp	Phe	Ala	Cys	Gly	Arg	Glu	Thr	Pro	Ser	Val	
				565					570						575	
Ala	Ala	Met	Val	Tyr	Pro	Phe	Gly	Gly	His	His	Val	Gln	Lys	Phe	Tyr	
			580					585						590		
Trp	Gly	Thr	Lys	Glu	Thr	Leu	Leu	Pro	Val	Phe	Thr	Ser	Met	Lys	Glu	
	595						600						605			
Ala	Val	Ala	Lys	Cys	Pro	Asp	Ala	Asp	Val	Val	Val	Asn	Phe	Ala	Ser	
	610					615						620				
Ser	Arg	Ser	Val	Tyr	Gln	Ser	Thr	Leu	Glu	Ala	Leu	Glu	Phe	Pro	Gln	

[0009]

625	630	635	640
Ile Lys Ala Ile Ala Leu Ile Ala Glu Gly Val Pro Glu Arg His Ala	645	650	655
Arg Glu Ile Leu His Leu Ala Lys Lys Lys Glu Val Ile Ile Ile Gly	660	665	670
Pro Ala Thr Val Gly Gly Ile Lys Pro Gly Cys Phe Arg Ile Gly Asn	675	680	685
Thr Gly Gly Met Asn Glu Asn Ile Leu Ser Ser Lys Leu Tyr Arg Ala	690	695	700
Gly Ser Val Gly Tyr Val Ser Lys Ser Gly Gly Met Ser Asn Glu Leu	705	710	715
Asn Asn Ile Leu Ser Leu Thr Thr Asp Gly Ala Tyr Glu Gly Ile Ala	725	730	735
Ile Gly Gly Asp Arg Tyr Pro Gly Thr Thr Phe Ile Asp His Leu Leu	740	745	750
Arg Tyr Glu Ala Asp Pro Asn Cys Lys Met Leu Val Leu Leu Gly Glu	755	760	765
Val Gly Gly Val Glu Glu Tyr Arg Val Ile Glu Ala Val Lys Ser Gly	770	775	780
Gln Ile Lys Lys Pro Ile Val Ala Trp Ala Ile Gly Thr Cys Ala Lys	785	790	795
Met Phe Ala Thr Asp Val Gln Phe Gly His Ala Gly Ser Met Ala Asn	805	810	815
Ser Asp Leu Glu Thr Ala Glu Ala Lys Asn Asn Ala Met Arg Ala Ala	820	825	830
Gly Phe Ile Val Pro Pro Thr Phe Glu Glu Leu Pro Gln Val Leu Ala			

[0010]

835	840	845
Glu Thr Tyr Gln Lys Leu Val Gly Asp Gly Thr Ile Gln Pro Lys Pro 850	855	860
Glu Val Pro Pro Pro Glu Ile Pro Met Asp Tyr Asn Trp Ala Gln Thr 865	870	875
Leu Gly Met Val Arg Lys Pro Ala Ala Phe Ile Ser Thr Ile Ser Asp 885	890	895
Glu Arg Gly Gln Glu Leu Leu Tyr Ala Gly Met Pro Ile Ser Lys Val 900	905	910
Phe Glu Glu Asp Ile Gly Ile Gly Gly Val Val Ser Leu Leu Trp Phe 915	920	925
Lys Arg Arg Leu Pro Ala Tyr Ala Thr Lys Phe Leu Glu Met Val Leu 930	935	940
Met Leu Thr Ala Asp His Gly Pro Ala Val Ser Gly Ala Met Thr Thr 945	950	955
Val Ile Thr Thr Arg Ala Gly Lys Asp Leu Val Ser Ser Leu Val Ala 965	970	975
Gly Leu Leu Thr Ile Gly Asp Arg Phe Gly Gly Ala Leu Asp Gly Ala 980	985	990
Ala Gln Glu Phe Thr Arg Ala Phe Glu Ala Gly Leu Thr Pro Arg Glu 995	1000	1005
Phe Val Asp Ser Met Arg Lys Ala Asn Lys Leu Ile Pro Gly Ile 1010	1015	1020
Gly His Lys Val Lys Ser Lys Ala Asn Pro Asp Lys Arg Val Glu 1025	1030	1035
Leu Val Lys Asn Tyr Val Phe Lys His Phe Pro Ser Ala Lys Leu		

[0011]

1040	1045	1050
Leu Glu Tyr Ala Leu Ala Val Glu Asp Val Thr Ser Ala Lys Lys 1055	1060	1065
Asp Thr Leu Ile Leu Asn Val Asp Gly Ala Ile Ala Val Ser Phe 1070	1075	1080
Val Asp Leu Leu Lys Asn Ser Gly Ala Phe Thr Ala Glu Glu Ala 1085	1090	1095
Ala Glu Tyr Met Lys Ile Gly Thr Leu Asn Gly Leu Phe Val Leu 1100	1105	1110
Gly Arg Ser Ile Gly Phe Ile Ala His His Leu Asp Gln Lys Arg 1115	1120	1125
Leu Lys Gln Pro Leu Tyr Arg His Pro Ala Asp Asp Ile Phe Ile 1130	1135	1140
Gln Pro Phe Asn Thr Asp Arg Ile Leu Val Gln Gln Arg Gln 1145	1150	1155

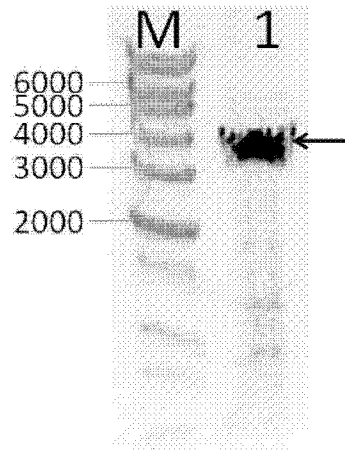


图 1

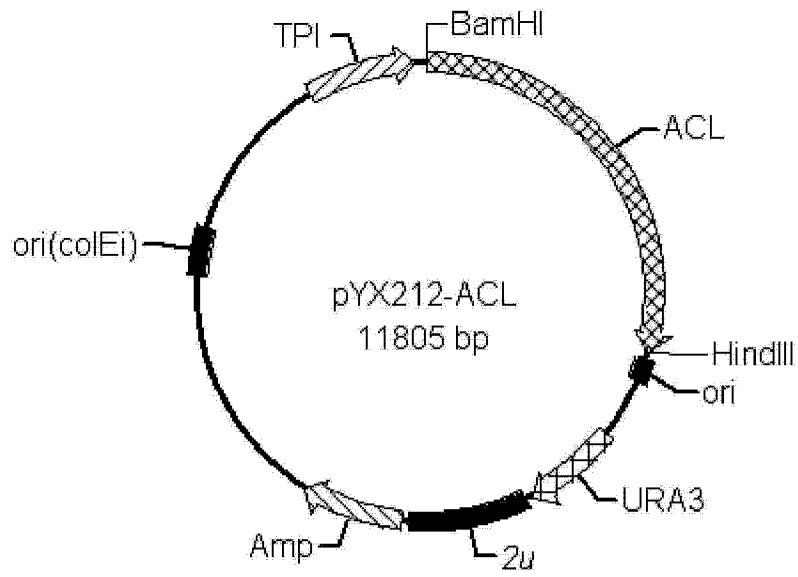


图 4

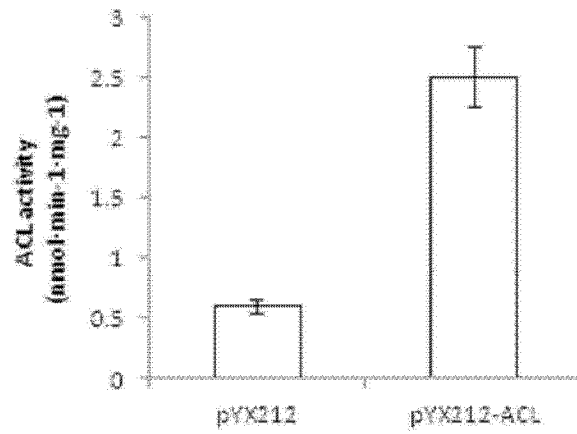


图 5