



(12)发明专利

(10)授权公告号 CN 102822343 B

(45)授权公告日 2017. 10. 20

(21)申请号 201080064882.8

PTA-9616 2008.11.18

(22)申请日 2010.12.28

(73)专利权人 赛诺菲疫苗技术公司

地址 法国里昂

(65)同一申请的已公布的文献号

申请公布号 CN 102822343 A

(72)发明人 A-C.V.贝恩 J.C.利普梅尔

K.E.埃普特 R.E.泽克尔

(43)申请公布日 2012.12.12

(30)优先权数据

61/290,441 2009.12.28 US

61/290,469 2009.12.28 US

61/413,353 2010.11.12 US

(74)专利代理机构 北京坤瑞律师事务所 11494

代理人 封新琴

(85)PCT国际申请进入国家阶段日

2012.08.28

(56)对比文件

CN 101484466 A,2009.07.15,

CN 1354792 A,2002.06.19,

Bilsel,P.A et al.GenBank:AAA48384.1.

《GenBank》.2002,1.

(86)PCT国际申请的申请数据

PCT/US2010/062276 2010.12.28

Springgs,M.K et al.UniProtKB/Swiss-

Prot:P06828.2.《GenBank》.2009,1.

(87)PCT国际申请的公布数据

W02011/090731 EN 2011.07.28

Fuhrmann M et al.Production of

antigens in Chlamydomonas reinhardtii:

green microalgae as a novel source of

recombinant proteins.《Methods Mol Med》

.2004,第194卷第191-195页.

(83)生物保藏信息

PTA-9695 2009.01.07

PTA-9696 2009.01.07

PTA-9697 2009.01.07

PTA-9698 2009.01.07

PTA-10212 2009.07.14

PTA-10213 2009.07.14

PTA-10214 2009.07.14

PTA-10215 2009.07.14

PTA-10209 2009.09.25

PTA-10210 2009.09.25

PTA-10211 2009.09.25

PTA-10208 2009.07.14

PTA-9613 2008.11.18

PTA-9614 2008.11.18

PTA-9615 2008.11.18

Schickli,J.H et al.GenBank:AF389118_

1.《GenBank》.2002,1.

Schickli,J.H et al.GenBank:AF389118_

1.《GenBank》.2002,1.

Springgs,M.K et al.UniProtKB/Swiss-

Prot:P06828.2.《GenBank》.2009,1.

Bilsel,P.A et al.GenBank:AAA48384.

《GenBank》.2002,1.

审查员 杜玉娟

权利要求书1页 说明书37页

序列表64页 附图30页

(54)发明名称

微藻中异源多肽的生产,微藻胞外体、组合物及其制备方法和用途

本发明涉及重组微藻细胞及其在异源蛋白生产中的用途,在微藻胞外体中生产异源多肽的方法,包含异源多肽的微藻胞外体,和包含它们的组合物。

(57)摘要

1. 一种生产病毒血凝素 (HA) 蛋白的方法, 所述方法包括:
 - (A) 提供包含核酸分子的重组微藻细胞, 所述核酸分子包含 (1) 编码包含HA膜结构域的病毒HA蛋白的多核苷酸序列, 和 (2) 有活性的调节性调控元件;
 - (B) 在培养基中培养重组微藻细胞以使得所述病毒HA蛋白分泌到培养基中;
 - (C) 从培养基回收所述分泌的重组HA蛋白。
2. 依照权利要求1所述的方法, 其中所述重组病毒HA蛋白是全长病毒HA蛋白。
3. 依照权利要求1或2所述的方法, 其中所述重组病毒HA蛋白是流感HA蛋白。
4. 依照权利要求3所述的方法, 其中所述重组病毒HA蛋白选自下组: H1、H2、H3、H4、H5、H6、H7、H8、H9、H10、H11、H12、H13、H14、H15、和H16型流感HA蛋白。
5. 依照权利要求3所述的方法, 其中所述重组病毒HA蛋白选自下组: H1、H2、H3、H5、H6、H7、和H9型流感HA蛋白。
6. 依照权利要求1所述的方法, 其中所述重组病毒HA蛋白是全长流感HA蛋白。
7. 依照权利要求1所述的方法, 其中所述重组病毒HA蛋白与SEQ ID NO:77是至少90%相同的。
8. 依照权利要求1所述的方法, 其中所述重组病毒HA蛋白具有SEQ ID NO:77的序列。
9. 依照权利要求1所述的方法, 其中所述重组病毒HA蛋白在培养中在没有外源蛋白酶的情况下被切割以产生HA1和HA2片段。
10. 依照权利要求1所述的方法, 其中所述微藻细胞是破囊壶菌目 (Thraustochytriales) 成员。
11. 依照权利要求1所述的方法, 其中所述微藻细胞是裂殖壶菌属 (Schizochytrium) 或破囊壶菌属 (Thraustochytrium) 的。
12. 一种制备疫苗组合物的方法, 所述方法包括:
 - 提供依照前述权利要求任一项所述的方法制备的重组病毒HA蛋白; 和
 - 将所述重组病毒HA蛋白配制为疫苗组合物, 所述疫苗组合物用于通过至少一种选自下组的途径施用: 肌肉内、静脉内、皮下、肺内、气管内、透皮、眼内、鼻内、吸入、腔内、导管内、和实质内施用。

微藻中异源多肽的生产,微藻胞外体、组合物及其制备方法和用途

[0001] 发明背景

技术领域

[0002] 本发明涉及重组微藻细胞及其在异源多肽生产中的用途,在微藻胞外体中生产异源多肽的方法,包含异源多肽的微藻胞外体,以及包含它们的组合物。

背景技术

[0003] 生物技术和分子生物学的进步使在微生物、植物和动物细胞中生产蛋白成为可能,其中许多蛋白以前只能通过从人类和其它动物的组织、血液或尿液中提取。如今市售的生物制剂通常在哺乳动物细胞,例如中国仓鼠卵巢 (CHO) 细胞,或在微生物细胞,例如酵母或大肠杆菌细胞系中制备。

[0004] 通过微生物发酵生产蛋白相对于现有的系统,例如植物和动物细胞培养呈现几项优势。例如,基于微生物发酵的方法可以提供:(i) 高浓度蛋白的快速生产;(ii) 使用无菌、充分受控的生产条件(例如良好生产规范(GMP)的条件)的能力;(iii) 使用简单、化学上确定成分的生长培养基以允许更简单的发酵和更少杂质的能力;(iv) 人或动物病原体污染的规避;(v) 易于回收蛋白(例如通过从发酵培养基中分离)。此外,发酵设施通常比细胞培养设施建设的成本更低。

[0005] 微藻,例如网粘菌门(Labyrinthulomycota)的破囊壶菌(thraustochytrids),能够以标准的发酵设备,非常短的培养周期(例如,1-5天),低廉的确定成分培养基和最低限度的纯化(如果有的话)进行培养。此外,某些微藻,例如裂殖壶菌属(Schizochytrium),其生物质及自其衍生的脂质具有已经证明的用于食品应用的安全历史。例如,来自此类微生物的富含DHA的甘油三酯油已获得美国食品和药物管理局的GRAS(一般认为安全)状态。

[0006] 微藻已被证明能够表达重组蛋白。例如,美国专利第7,001,772号公开了第一个适于转化破囊壶菌,包括裂殖壶菌属成员的重组构建体。此出版物公开了乙酰乳酸合酶、乙酰乳酸合酶启动子和终止子区、 α -微管蛋白启动子、来自聚酮合酶(PKS)系统的启动子、和脂肪酸去饱和酶启动子的裂殖壶菌属核酸和氨基酸序列等。美国申请公开号2006/0275904和2006/0286650随后公开了肌动蛋白、延伸因子1 α (ef1 α)和甘油醛-3-磷酸脱氢酶(GAPDH)启动子和终止子的裂殖壶菌序列。

[0007] 对于鉴定在微藻中表达异源多肽的方法以及治疗应用的替代组合物的需求一直存在。

[0008] 发明概述

[0009] 本发明涉及一种生产病毒蛋白的方法,所述病毒蛋白选自下组:血凝素(HA)蛋白、神经氨酸酶(NA)蛋白、融合(F)蛋白、糖蛋白(G)蛋白、包膜(E)蛋白、120kDa的糖蛋白(gp120)、41kDa的糖蛋白(gp41)、基质蛋白及其组合,该方法包括在培养基中培养重组微藻细胞以产生病毒蛋白,其中所述重组微藻细胞包含含有编码病毒蛋白的多核苷酸序列的核

酸分子。在一些实施方案中,病毒蛋白被分泌。在一些实施方案中,从培养集中回收病毒蛋白。在一些实施方案中,病毒蛋白在微藻细胞中积累。在一些实施方案中,病毒蛋白在微藻细胞的膜中积累。在一些实施方案中,病毒蛋白是HA蛋白。在一些实施方案中,HA蛋白与SEQ ID NO:77是至少90%相同的。在一些实施方案中,微藻细胞能够在不存在外源蛋白酶的情况下,翻译后加工HA蛋白以产生HA1和HA2多肽。在一些实施方案中,病毒蛋白是NA蛋白。在一些实施方案中,NA蛋白与SEQ ID NO:100是至少90%相同的。在一些实施方案中,病毒蛋白是F蛋白。在一些实施方案中,F蛋白与SEQ ID NO:102是至少90%相同的。在一些实施方案中,病毒蛋白是G蛋白。在一些实施方案中,G蛋白与SEQ ID NO:103是至少90%相同的。在一些实施方案中,微藻细胞是破囊壶菌目 (Thraustochytriales) 成员。在一些实施方案中,微藻细胞是裂殖壶菌属 (Schizochytrium) 或破囊壶菌属 (Thraustochytrium) 的。在一些实施方案中,编码病毒蛋白的多核苷酸序列进一步包含HA膜结构域。在一些实施方案中,核酸分子进一步包含选自下组的多核苷酸序列:SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:38、SEQ ID NO:42、SEQ ID NO:43、SEQ ID NO:44、SEQ ID NO:45、SEQ ID NO:46及其组合。

[0010] 本发明涉及由上述方法中的任一种产生的分离的病毒蛋白。

[0011] 本发明涉及一种重组微藻细胞,其包含含有编码病毒蛋白的多核苷酸序列的核酸分子,所述病毒蛋白选自下组:血凝素 (HA) 蛋白、神经氨酸酶 (NA) 蛋白、融合 (F) 蛋白、糖蛋白 (G) 蛋白、包膜 (E) 蛋白、120kDa的糖蛋白 (gp120)、41kDa的糖蛋白 (gp41)、基质蛋白及其组合。在一些实施方案中,病毒蛋白是HA蛋白。在一些实施方案中,HA蛋白与SEQ ID NO:77是至少90%相同的。在一些实施方案中,微藻细胞能够在不存在外源蛋白酶的情况下,翻译后加工HA蛋白以产生HA1和HA2多肽。在一些实施方案中,病毒蛋白是NA蛋白。在一些实施方案中,NA蛋白与SEQ ID NO:100是至少90%相同的。在一些实施方案中,病毒蛋白是F蛋白。在一些实施方案中,F蛋白与SEQ ID NO:102是至少90%相同的。在一些实施方案中,病毒蛋白是G蛋白。在一些实施方案中,G蛋白与SEQ ID NO:103是至少90%相同的。在一些实施方案中,微藻细胞是破囊壶菌目成员。在一些实施方案中,微藻细胞是裂殖壶菌属 (Schizochytrium) 或破囊壶菌属 (Thraustochytrium) 的。在一些实施方案中,编码病毒蛋白的多核苷酸序列进一步包含HA膜结构域。在一些实施方案中,核酸分子进一步包含选自下组的多核苷酸序列:SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:38、SEQ ID NO:42、SEQ ID NO:43、SEQ ID NO:44、SEQ ID NO:45、SEQ ID NO:46及其组合。

[0012] 本发明涉及一种生产包含异源多肽的微藻胞外体的方法,该方法包括:(a) 在微藻宿主细胞中表达异源多肽,其中所述异源多肽包含膜结构域,和 (b) 将微藻宿主细胞在足以产生包含异源多肽的胞外体的条件下培养,其中所述胞外体与宿主细胞质膜是不连续的。

[0013] 本发明涉及一种生产包含微藻胞外体和异源多肽的组合物的方法,该方法包括:(a) 在微藻宿主细胞中表达异源多肽,其中所述异源多肽包含膜结构域,和 (b) 将微藻宿主细胞在足以产生包含异源多肽的胞外体的条件下培养,其中所述胞外体与宿主细胞质膜是不连续的,其中所述组合物作为包含所述胞外体的培养上清产生。在一些实施方案中,该方法进一步包括从组合物去除培养上清和在含水液体载体中重悬胞外体。本发明涉及由该方法产生的组合物。

[0014] 在一些实施方案中,生产微藻胞外体和异源多肽的方法,或者生产包含微藻胞外体和异源多肽的组合物的方法,包括宿主细胞,其是网粘菌门 (Labyrinthulomycota) 宿主

细胞。在一些实施方案中,宿主细胞是裂殖壶菌属或破囊壶菌属宿主细胞。

[0015] 本发明涉及一种包含异源多肽的微藻胞外体,其中所述胞外体与微藻细胞的质膜是不连续的。在一些实施方案中,胞外体是囊泡、胶束、膜片段、膜聚集体、或其混合物。在一些实施方案中,胞外体是囊泡和膜片段的混合物。在一些实施方案中,胞外体是囊泡。在一些实施方案中,异源多肽包含膜结构域。在一些实施方案中,异源多肽是糖蛋白。在一些实施方案中,糖蛋白包含高甘露寡糖。在一些实施方案中,糖蛋白基本上不含唾液酸。

[0016] 本发明涉及一种包含上述权利要求中任一项所述的胞外体和含水液体载体的组合物。在一些实施方案中,含水液体载体是培养上清。

[0017] 附图简述

[0018] 图1显示了编码甲型流感病毒(A/Puerto Rico/8/34/Mount Sinai (H1N1))血凝素(HA)蛋白的多核苷酸序列(SEQ ID NO:76),其经密码子优化以供在裂殖壶菌属菌种ATCC 20888中表达。

[0019] 图2显示了pCL0143的质粒图谱。

[0020] 图3显示了用于分析CL0143-9克隆的流程。

[0021] 图4显示了转基因裂殖壶菌属CL0143-9(“E”)进行的HA蛋白的分泌。图4A显示了在抗H1N1免疫印迹中从不同温度(25°C、27°C、29°C)和pH(5.5、6.0、6.5、7.0)的培养物的低速上清(即无细胞上清(“CFS”))回收的重组HA蛋白(如箭头所示)。图4B显示了在非还原或还原条件下从60%蔗糖级分的抗-H1N1免疫印迹(“IB:抗-H1N1”)和考马斯染色凝胶(“考马斯”)回收的重组HA蛋白。

[0022] 图5显示了来自转基因裂殖壶菌属CL0143-9(“E”)的重组HA蛋白的血凝素活性。图5A显示了无细胞上清(“CFS”)中的血凝素活性。图5B显示了在可溶(“US”)和不溶性(“UP”)级分中的血凝素活性。“[蛋白质]”指蛋白质的浓度,随样品稀释增加而从左至右降低。“-”指缺乏HA的阴性对照。“+”指流感血凝素阳性对照。“C”指阴性对照裂殖壶菌属菌种ATCC 20888野生型菌株。“HAU”指基于从左至右的样品倍数稀释的血凝素活性单位。“2”指第一孔中的2倍稀释的样品;从左至右的随后的孔代表在上一孔之上的加倍稀释,从而使从左至右从第一孔到最后的孔的倍数稀释是2、4、8、16、32、64、128、256、512、1024、2048和4096。

[0023] 图6显示了存在于转基因裂殖壶菌属CL0143-9(“E”)60%蔗糖级分中的HA蛋白的表达和血凝素活性。图6A显示了考马斯染色凝胶(“考马斯”)和抗-H1N1免疫印迹(“IB:抗-H1N1”)中所示的来自60%蔗糖级分的回收的重组HA蛋白(如箭头所示)。图6B显示了相应的血凝素活性。“-”指缺乏HA的阴性对照。“+”指流感HA蛋白阳性对照。“C”指阴性对照裂殖壶菌属菌种ATCC20888野生型菌株。“HAU”指基于样品倍数稀释的血凝素活性单位。

[0024] 图7显示了回收的重组HA蛋白的肽序列分析,这是通过覆盖完整HA蛋白序列(SEQ ID NO:77)超过42%的共27条肽(与肽相关的氨基酸以粗体突出显示)确定的。HA1多肽由总共17条肽确定,而HA2多肽由总共9条肽确定。

[0025] 图8显示了例示裂殖壶菌属中HA蛋白糖基化的考马斯染色凝胶(“考马斯”)和抗-H1N1免疫印迹(“IB:抗-H1N1”)。“EndoH”和“PNGase F”指转基因裂殖壶菌属CL0143-9的60%蔗糖级分使用各种酶的酶处理。“NT”指不用酶温育但在与EndoH和PNGase F处理相同条件下的转基因裂殖壶菌属CL0143-9。

[0026] 图9显示了对应于时间(小时)的全部裂殖壶菌属菌种ATCC 20888培养上清蛋白

(g/L)。

[0027] 图10在11-15泳道显示了全部裂殖壶菌属菌种ATCC 20888培养上清蛋白的SDS-PAGE,其中在如图9所示第37-68小时中的6个时间点中的五个,除了第52小时,收集了上清。鉴定为肌动蛋白或凝溶胶蛋白的条带(通过质谱肽测序)以箭头标记。泳道11以2.4μg总蛋白上样,其余孔以5μg总蛋白上样。

[0028] 图11显示了来自裂殖壶菌属菌种ATCC 20888("C:20888")和转基因裂殖壶菌属CL0143-9("E:CL0143-9")负染色的囊泡。

[0029] 图12显示了来自裂殖壶菌属菌种ATCC 20888("C:20888")和转基因裂殖壶菌属CL0143-9("E:CL0143-9")的抗-H1N1免疫金标记的囊泡。

[0030] 图13显示了基于使用SignalP算法预测的裂殖壶菌属天然的信号锚序列。参见,例如,Bendsten等,J.Mol.Biol.340:783-795(2004);Nielsen,H.和Krogh,A.Proc.Int.Conf.Intell.Syst.Mol.Biol.6:122-130(1998);Nielsen,H.等,Protein Engineering 12:3-9(1999);Emanuelsson,O.等,Nature Protocols2:953-971(2007)。

[0031] 图14显示了预测的裂殖壶菌属中的I型膜蛋白,其是基于针对与来自其它生物的已知I型膜蛋白具有同源性并在蛋白的极C-端区具有跨膜区的基因,对裂殖壶菌属基因组和EST DNA数据库进行BLAST检索。

[0032] 图15显示了pCL0120的质粒图谱。

[0033] 图16显示了裂殖壶菌属的密码子选择表。

[0034] 图17显示了pCL0130的质粒图谱。

[0035] 图18显示了pCL0131的质粒图谱。

[0036] 图19显示了pCL0121的质粒图谱。

[0037] 图20显示了pCL0122的质粒图谱。

[0038] 图21显示了为在裂殖壶菌属菌种ATCC 20888中表达优化的编码瘤胃真菌属菌种(*Piromyces* sp.)E2木糖异构酶蛋白“XylA”的多核苷酸序列(SEQ ID NO:92),对应于GenBank登录号CAB76571。

[0039] 图22显示了为在裂殖壶菌属菌种ATCC 20888中表达优化的编码瘤胃真菌属菌种E2木酮糖激酶蛋白“XylB”的多核苷酸序列(SEQ ID NO:93),对应于GenBank登录号AJ249910。

[0040] 图23显示了pCL0132的质粒图谱。

[0041] 图24显示了pCL0136的质粒图谱。

[0042] 图25A显示了pCL0140的质粒图谱而图25B显示了pCL0149的质粒图谱。

[0043] 图26A显示了为在裂殖壶菌属菌种ATCC 20888中表达优化的编码甲型流感病毒(A/Puerto Rico/8/34/Mount Sinai(H1N1))的神经氨酸酶(NA)蛋白的多核苷酸序列(SEQ ID NO:100)。图26B显示了为在裂殖壶菌属菌种ATCC20888中表达优化的编码甲型流感病毒(A/Puerto Rico/8/34/Mount Sinai(H1N1))的NA蛋白,接有V5标签和多聚组氨酸标签的多核苷酸序列(SEQ ID NO:101)。

[0044] 图27显示了用于分析CL0140和CL0149克隆的流程表。

[0045] 图28显示了来自转基因裂殖壶菌属菌株CL0140-16,-17,-20,-21,-22,-23,-24,-26,-28的重组NA的神经氨酸酶活性。通过测量4-甲基伞形酮的荧光来测定活性,所述荧光

在唾液酸酶水解(4-甲基伞形酮)- α -D-N-乙酰神经氨酸(4-methylumbelliferyl)- α -D-N-Acetylneuraminate;4-MUNANA)之后产生(激发(Exc):365nm,发射(Em):450nm)。以浓缩的无细胞上清(cCFS,每个克隆最左边的柱)和无细胞提取物(CFE,每个克隆最右边的柱)中每 μ g蛋白的相对荧光单位(RFU)来表示活性。与转基因菌株以相同方式培养和制备的裂殖壶菌属菌种ATCC 20888野生型菌株(“-”)和用pCL0140转化的裂殖壶菌属的PCR阴性菌株(“27”)用作阴性对照。

[0046] 图29显示了来自转基因裂殖壶菌属菌株CL0140-26的重组NA蛋白的部分纯化。各级分的神经氨酸酶活性如图29A所示。“cCFS”指浓缩的无细胞上清。“D”指以洗涤缓冲液稀释的cCFS,“FT”指流过物(flow-through),“W”指洗液,“E”指洗脱液和“cE”指浓缩的洗脱级分。每级分12.5 μ L的考马斯染色凝胶(“考马斯”)如图29B所示。箭头指向鉴定为NA蛋白的条带。以MOPSSDS运行缓冲液在NuPAGE® Novex®12%的bis-tris凝胶上使用SDS-PAGE分离蛋白。

[0047] 图30显示了回收的重组NA蛋白的肽序列分析,NA蛋白是通过覆盖蛋白序列(SEQ ID NO:10)25%的共9条肽(以粗红突出显示)确定的。

[0048] 图31显示了转基因裂殖壶菌属菌株CL0149-10,-11,-12的神经氨酸酶活性和对应的考马斯染色凝胶(“考马斯”)和抗-V5免疫印迹(“免疫印迹:抗-V5”)。图31A显示了神经氨酸酶活性,如通过测量4-甲基伞形酮的荧光来测定的,荧光在唾液酸酶水解4-MUNANA之后产生(激发(Exc):365nm,发射(Em):450nm)。以无细胞上清(CFS)中每 μ g蛋白的相对荧光单位(RFU)来表示活性。将与转基因菌株以相同方式培养和制备的裂殖壶菌属菌种ATCC20888野生型菌株(“-”)用作阴性对照。图31B显示了3株转基因裂殖壶菌属CL0149菌株(“10”、“11”、和“12”)的12.5 μ L CFS的考马斯染色凝胶和对应的抗-V5免疫印迹。将Positope™抗体对照蛋白用作阳性对照(“+”)。将与转基因菌株以相同方式培养和制备的裂殖壶菌属菌种ATCC 20888野生型菌株(“-”)用作阴性对照。

[0049] 图32显示了用CL0140和CL0143共转化的转基因裂殖壶菌属的无细胞上清中的流感HA和NA的酶活性。显示了克隆CL0140-143-1,-3,-7,-13,-14,-15,-16,-17,-18,-19,-20的数据。图32A显示了神经氨酸酶活性,如通过测量4-甲基伞形酮的荧光来测定的,所述荧光由唾液酸酶水解4-MUNANA之后产生(激发(Exc):365nm,发射(Em):450nm)。以25 μ L CFS中的相对荧光单位(RFU)来表示活性。将与转基因菌株以相同方式培养和制备的裂殖壶菌属菌种ATCC 20888野生型菌株(“-”)用作阴性对照。图32B显示了血凝素活性。“-”指缺少HA的阴性对照。“+”指流感HA阳性对照。“HAU”指基于样品倍数稀释的血凝素活性单位。

[0050] 发明详述

[0051] 本发明涉及在微藻宿主细胞中生产异源多肽的方法。本发明也涉及所述方法产生的异源多肽、包含所述异源多肽的微藻细胞、和包含所述异源多肽的组合物。本发明也涉及在微藻宿主细胞中异源多肽的产生,其中所述异源多肽与微藻胞外体相关,所述微藻胞外体与宿主细胞质膜是不连续的。本发明也涉及包含异源多肽的微藻胞外体的产生,以及包含它们的组合物的产生。本发明进一步涉及包含异源多肽的微藻胞外体、组合物及其用途。

[0052] 微藻宿主细胞

[0053] 微藻,又称微小藻类(microscopic algae),经常发现于淡水和海洋系统。微藻是单细胞的但也可以生长成链和组。单个细胞的大小从几微米到几百微米。因为细胞能够生

长在含水悬液中,它们能有效地获得营养和进入液体环境。

[0054] 在一些实施方案中,微藻宿主细胞是不等鞭毛体(heterokont)或原生藻菌(stramenopile)。

[0055] 在一些实施方案中,微藻宿主细胞是网粘菌门(Labyrinthulomycota)成员。在一些实施方案中,网粘菌门宿主细胞是破囊壶菌目(Thraustochytriales)或网粘菌目(Labyrinthulales)成员。根据本发明,术语“破囊壶菌(thraustochytrid)”是指破囊壶菌目(Thraustochytriales)的任何成员,包括破囊壶菌科(Thraustochytriaceae),而术语“网粘菌(labyrinthulid)”是指网粘菌目(Labyrinthulales)的任何成员,包括网粘菌科(Labyrinthulaceae)。网粘菌科的成员从前曾被认为是破囊壶菌目的成员,但是在最近此类生物体的分类学归类的修订版中,网粘菌科现在被认为是网粘菌目的成员。网粘菌目和破囊壶菌目均被认为是网粘菌门的成员。分类学理论现在通常将这两个群的微生物与原生藻菌世系中的藻或藻样原生生物放在一起。破囊壶菌和网粘菌当前的分类学地位可以归纳如下:

[0056] 界:藻菌界(Stramenopila)(管毛生物界(Chromista))

[0057] 门:网粘菌门(Labyrinthulomycota)(不等毛门(Heterokonta))

[0058] 纲:网粘菌纲(Labyrinthulomycetes)(Labyrinthulae)

[0059] 目:网粘菌目(Labyrinthulales)

[0060] 科:网粘菌科(Labyrinthulaceae)

[0061] 目:破囊壶菌目(Thraustochytriales)

[0062] 科:破囊壶菌科(Thraustochytriaceae)

[0063] 对于本发明的目的而言,作为破囊壶菌描述的菌株包括如下生物体:目:破囊壶菌目;科:破囊壶菌科;属:破囊壶菌属(种:菌种,arudimentale,aureum,benthicola,globosum,kinnei,motivum,multirudimentale,pachydermum,proliferum,roseum,striatum)、吾肯氏壶菌属(Ulkenia)(种:菌种,amoeboidea,kerguelensis,minuta,profunda,radiata,sailens,sarkariana,schizochytrids,visurgensis,yorkensis)、裂殖壶菌属(种:菌种,aggregatum,limnaceum,mangrovei,minutum,octosporum), Japonochytrium属(种:菌种,marinum)、Aplanochytrium属(种:菌种,haliotidis,kerguelensis,profunda,stocchini)、Althornia属(种:菌种,crouchii)、或Elina属(种:菌种,marisalba,sinorifica)。出于本发明的目的,吾肯氏壶菌属中描述的物种将被看作是破囊壶菌属的成员。Aurantiochytrium、Oblongichytrium、Botryochytrium、Parietichytrium、和Sicyoidochytrium是本发明网粘菌门包括的额外的属。

[0064] 本发明作为网粘菌描述的菌株包括如下生物体:目:网粘菌目,科:网粘菌科,属:网粘菌属(种:菌种,algeriensis,coenocystis,chattonii,macrocystis,macrocystis atlantica,macrocystis macrocystis,marina,minuta,roscoffensis,valkanovii,vitellina,vitellina pacifica,vitellina vitellina,zopfii),Labyrinthuloides属(种:菌种,haliotidis,yorkensis),Labyrinthomyxa属(种:菌种,marina),Diplophrys属(种:菌种,archeri),Pyrrhosorus属(种:菌种,marinus),Sorodiplophrys属(种:菌种,stercorea)或Chlamydomyxa属(种:菌种,labyrinthuloides,montana)(尽管关于Pyrrhosorus、Sorodiplophrys或Chlamydomyxa的精确分类学地位目前尚没有达成一致)。

[0065] 网粘菌门的微藻细胞包括但不限于保藏菌株PTA-10212、PTA-10213、PTA-10214、PTA-10215、PTA-9695、PTA-9696、PTA-9697、PTA-9698、PTA-10208、PTA-10209、PTA-10210、PTA-10211、以SAM2179保藏的微生物(保藏者命名为“吾肯氏壶菌SAM2179”)、任何破囊壶菌属菌种(包括前述的吾肯氏壶菌属菌种,例如*U.visurgensis*, *U.amoeboidea*, *U.sarkariana*, *U.profunda*, *U.radiata*, *U.minuta*和吾肯氏壶菌菌种BP-5601),并且包括*Thraustochytrium striatum*, *Thraustochytrium aureum*, *Thraustochytrium roseum*;和任何*Japonochytrium*属菌种。破囊壶菌目菌株包括但不限于破囊壶菌属菌种(23B) (ATCC 20891); *Thraustochytrium striatum* (Schneider) (ATCC 24473); *Thraustochytrium aureum* (Goldstein) (ATCC 34304); *Thraustochytrium roseum* (Goldstein) (ATCC 28210);和*Japonochytrium*属菌种(L1) (ATCC 28207)。裂殖壶菌属包括但不限于*Schizochytrium aggregatum*, *Schizochytrium limacinum*, 裂殖壶菌属菌种(S31) (ATCC 20888), 裂殖壶菌属菌种(S8) (ATCC 20889), 裂殖壶菌属菌种(LC-RM) (ATCC 18915), 裂殖壶菌属菌种(SR 21), 保藏菌株ATCC 28209,和保藏的*Schizochytrium limacinum*菌株IFO 32693。在一些实施方案中,细胞是裂殖壶菌属(*Schizochytrium*)或破囊壶菌属(*Thraustochytrium*)的。裂殖壶菌属能够通过连续二等分裂和通过形成孢子囊两种方式复制,最终释放游动孢子。但是,破囊壶菌属仅通过形成孢子囊复制,然后释放游动孢子。

[0066] 在一些实施方案中,微藻宿主细胞是网粘菌纲(也称作*Labyrinthulomycetes*)。网粘菌纲产生特有的结构称为“质外网”(ectoplasmic nets)。这些结构是质膜分叉的、管状的延伸物,极大地有助于增加质膜的表面积。参见,例如,Perkins, *Arch. Mikrobiol.* 84:95-118 (1972); Perkins, *Can. J. Bot.* 51:485-491 (1973)。质外网是从称为sagenosome或生网体(bothrosome)的独特的细胞结构形成的。质外网将网粘菌纲细胞附着于表面并能够穿过表面。分别参见,例如, Coleman和Vestal, *Can. J. Microbiol.* 33:841-843 (1987),和Porter, *Mycologia* 84:298-299 (1992)。例如,已经观察到裂殖壶菌属菌种ATCC 20888当在固体培养基上培养时,产生延伸到琼脂中的质外网(数据未显示)。在此情况下质外网看起来充当类伪根(pseudorhizoid)。此外,在某些质外网膜延伸中发现丰富的肌动蛋白纤丝。参见,例如, Preston, *J. Eukaryot. Microbiol.* 52:461-475 (2005)。基于肌动蛋白纤丝在其他生物的细胞骨架结构中的重要性,预计例如肌动蛋白的细胞骨架元件在质外网膜延伸的形成和/或完整性中发挥作用。

[0067] 产生类伪根延伸的其他生物包括称为壶菌(chytrids)的生物,分类学上将这些生物分类在包括壶菌门(*Chytridiomycota*)或须霉属(*Phycomyces*)的各组中。属的例子包括*Chytridium*, *Chytriumyces*, *Cladochytrium*, *Lacustromyces*、根生壶菌属(*Rhizophyidium*)、*Rhizophyctidaceae*, *Rozella*, *Olpidium*和*Lobulomyces*。

[0068] 在一些实施方案中,微藻宿主细胞包含膜延伸物。在一些实施方案中,微藻宿主细胞包含类伪根。在一些实施方案中,微藻宿主细胞包含质外网。在一些实施方案中,微藻宿主细胞包含sagenosome或生网体。

[0069] 在一些实施方案中,微藻宿主细胞是破囊壶菌。在一些实施方案中,微藻宿主细胞是裂殖壶菌属或破囊壶菌属细胞。

[0070] 在一些实施方案中,微藻宿主细胞是网粘菌。

[0071] 在一些实施方案中,微藻宿主细胞是能够通过常规分泌途径加工多肽的真核细

胞,例如网粘菌门成员,包括裂殖壶菌属、破囊壶菌属和其他破囊壶菌。例如,已经认识到网粘菌门成员比CHO细胞产生更少的丰富分泌的蛋白,导致使用裂殖壶菌比例如CHO细胞具有优势。此外,不像大肠杆菌,网粘菌门成员,例如裂殖壶菌属,进行某些蛋白的生物活性所必需的蛋白糖基化,例如N-联糖基化。已经确定破囊壶菌,例如裂殖壶菌属,展示的N-联糖基化比酵母糖基化更紧密类似于哺乳动物的糖基化模式。

[0072] 用于本发明宿主细胞的有效培养条件包括但不限于允许蛋白产生和/或重组的有效培养基、生物反应器、温度、pH和氧条件。有效的培养基指任何通常培养微藻细胞,如破囊壶菌目细胞,例如裂殖壶菌属宿主细胞的培养基。这种培养基通常包含具有可同化的碳、氮、和磷酸盐源、以及适当的盐、无机物、金属和其他营养物质,如维生素的含水培养基。合适的培养基和培养条件的非限制性实例公开在实施例部分。适合破囊壶菌目微生物的非限制性培养条件也描述于,例如,美国专利第5,340,742号,将其通过提述整体并入本文。本发明的细胞可以培养于常规的发酵生物反应器、摇瓶、试管、微滴定板和培养皿中。可以在适于重组细胞的温度、pH和氧含量下进行培养。

[0073] 在一些实施方案中,本发明的微藻宿主细胞含有包含编码选择标记的核酸序列的重组载体。在一些实施方案中,选择标记允许选择经转化的微生物。显性选择标记的例子包括降解具有抗生素或杀真菌活性的化合物的酶,如,例如来自*Streptoalloteichus hindustanus*的编码SEQ ID NO:5所示的“博莱霉素结合蛋白”的Sh ble基因。显性选择标记的另一个例子包括破囊壶菌乙酰乳酸合酶序列,例如SEQ ID NO:6的多核苷酸序列的突变形式。可以修饰、突变或以其他方式选择乙酰乳酸合酶以使其对磺脲类化合物、咪唑啉酮类抑制剂和/或嘧啶基氧代苯甲酸酯(pyrimidinyl oxybenzoates)的抑制有抗性。破囊壶菌乙酰乳酸合酶序列的代表性例子包括但不限于如SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10的氨基酸序列,或通过在一个或多个以下位置的氨基酸缺失、插入或取代而不同于SEQ ID NO:7的氨基酸序列:116G、117A、192P、200A、251K、358M、383D、592V、595W、或599F,和例如SEQ ID NO:11、SEQ ID NO:12、或SEQ ID NO:13的多核苷酸序列,以及与任一代表性序列具有至少90%、至少95%、至少96%、至少97%、至少98%、或至少99%同一性的序列。可以包含于用于转化微藻细胞的重组载体的选择标记进一步的例子包括ZEOCINTM、巴龙霉素、潮霉素、杀稻瘟素或任何其他合适的抗性标记。

[0074] 术语“转化”用于指任何一种方法,通过这种方法外源核酸分子(即,重组核酸分子)可被插入微生物细胞。在微生物系统中,术语“转化”用于描述由于微生物获得外源核酸所致的一种遗传性改变而且可基本上与术语“转染”同义。用于将外源核酸分子引入微藻宿主细胞的合适的转化技术包括但不限于,粒子轰击、电穿孔、显微注射、脂转染、吸附、感染和原生质体融合。例如,可以在指数生长期或当微藻细胞在600nm的光密度达到1.5-2时,将外源核酸分子,包括重组载体,引入处于的静息期的微藻细胞。也可以在通过电穿孔将核酸分子导入宿主细胞前,用具有蛋白酶活性的酶预处理微藻宿主细胞。

[0075] 在一些实施方案中,遗传修饰宿主细胞以引入或缺失涉及与转运和/或合成碳水化合物相关的生物合成途径的基因,包括涉及糖基化的那些。例如,可以通过缺失内源糖基化基因和/或插入人或动物的糖基化基因以允许更紧密类似于人的糖基化模式来修饰宿主细胞。酵母中的糖基化修饰可参见,例如,美国专利第7,029,872号和美国申请公开号2004/0171826,2004/0230042,2006/0257399,2006/0029604和2006/0040353。本发明的宿主细胞

包括其中使用RNA病毒元件来增加或调节基因表达的细胞。

[0076] 表达系统

[0077] 用于在微藻宿主细胞中表达异源多肽的表达系统包含在微藻细胞中有活性的调节性调控元件。在一些实施方案中,表达系统包含在网粘菌门细胞中有活性的调节性调控元件。在一些实施方案中,表达系统包含在破囊壶菌中有活性的调节性调控元件。在一些实施方案中,表达系统包含在裂殖壶菌属或破囊壶菌属中有活性的调节性调控元件。许多调节性调控元件,包括各种启动子,在许多不同物种中具有活性。因此,调节序列可以在与分离出所述调节序列的细胞相同的细胞类型中应用,或者也可以在与分离出所述调节序列的细胞不同的细胞类型中应用。这种表达盒的设计和构建使用本领域熟练技术人员已知的标准分子生物学技术。参见,例如,Sambrook等,2001,Molecular Cloning:A Laboratory Manual,第三版。

[0078] 在一些实施方案中,用于在微藻细胞中的异源多肽产生的表达系统包含衍生自网粘菌门序列的调节元件。在一些实施方案中,在微藻细胞中用于产生异源多肽的表达系统包含衍生自非网粘菌门序列的调节元件,包括衍生自非网粘菌门藻类序列的序列。在一些实施方案中,表达系统包含编码异源多肽的多核苷酸序列,其中所述多核苷酸序列与在微藻宿主细胞中有功能的任何启动子序列、任何终止子序列、和/或任何其他调节序列相关。可以使用诱导性或组成型活性的序列。合适的调节性调控元件也包括与本文描述的核酸分子相关的任何调节性调控元件。

[0079] 本发明也涉及在微藻宿主细胞中用于表达异源多肽的表达盒。本发明也涉及包含在宿主细胞中表达异源多肽的表达盒的任何一种上述宿主细胞。在一些实施方案中,表达系统包含含有遗传元件,例如至少是启动子、编码序列和终止子区的表达盒,它们以在宿主细胞中有功能的方式可操作地连接。在一些实施方案中,表达盒包含至少一种分离的如本发明所述的核酸分子。在一些实施方案中,表达盒的所有遗传元件是与分离的核酸分子关联的序列。在一些实施方案中,调控序列是可诱导的序列。在一些实施方案中,编码异源多肽的核酸序列整合到宿主细胞的基因组中。在一些实施方案中,编码异源多肽的核酸序列稳定整合到宿主细胞的基因组中。

[0080] 在一些实施方案中,编码要表达的异源多肽的分离的核酸序列可操作地连接启动子序列和/或终止子序列,在宿主细胞中两者均是有功能的。编码要表达的异源多肽的分离的核酸序列可操作地连接的启动子和/或终止子序列可以包括任何启动子和/或终止子序列,包括但不限于本文公开的核酸序列、美国专利第7,001,772号公开的调节序列、美国申请公开号2006/0275904和2006/0286650公开的调节序列、美国申请公开号2010/0233760号和W02010/107709公开的调节序列、或其他与编码异源多肽的分离的多核苷酸序列可操作地连接、在其转化的宿主细胞中有功能的调节序列。在一些实施方案中,编码异源多肽的核酸序列是为特定的微藻宿主细胞密码子优化的,以最大化翻译效率。

[0081] 本发明也涉及包含本发明的表达盒的重组载体。重组载体包括但不限于质粒、噬菌体和病毒。在一些实施方案中,重组载体是线性化载体。在一些实施方案中,重组载体是表达载体。本文所用的短语“表达载体”指适于生产编码产物(如感兴趣的蛋白)的载体。在一些实施方案中,将编码要生产的产物的核酸序列插入重组载体以产生重组核酸分子。将编码要产生的异源多肽的核酸序列以将该核酸序列与载体中的调节序列(如破囊壶菌目启

动子)可操作地连接的方式插入载体,其使得所述核酸序列在重组微生物中能够转录和翻译。在一些实施方案中,选择标记,包括任何本文所述的选择标记,使得能够选择已将本发明的重组核酸分子成功引入的重组微生物。

[0082] 在一些实施方案中,本发明的宿主细胞产生的异源多肽以商业规模生产。商业规模包括从在规格 $\geq 100\text{L}$ 、 $\geq 1,000\text{L}$ 、 $\geq 10,000\text{L}$ 或 $\geq 100,000\text{L}$ 的通气发酵罐中生长的微生物来产生异源多肽。在一些实施方案中,在带有搅拌的通气发酵罐中进行商业规模生产。

[0083] 在一些实施方案中,本发明的宿主细胞产生的异源多肽可以积累在细胞内或可以从细胞分泌,例如作为可溶性异源多肽分泌到培养基中。

[0084] 在一些实施方案中,本发明生产的异源多肽从细胞、从培养基、或细胞生长的发酵培养基中回收。在一些实施方案中,异源多肽是作为可溶性异源多肽回收自培养基的分泌的异源多肽。在一些实施方案中,异源多肽是包含信号肽的分泌的蛋白。

[0085] 在一些实施方案中,本发明生产的异源多肽包含指导其在内质网中滞留、指导其胞外分泌、或指导其到其他细胞器或细胞区室的定位信号。在一些实施方案中,异源多肽包含信号肽。在一些实施方案中,异源多肽包含Na/Pi-IIb2转运蛋白信号肽或Sec1转运蛋白。在一些实施方案中,信号肽包含SEQ ID NO:1或SEQ ID NO:37的氨基酸序列。在一些实施方案中,包含具有SEQ ID NO:1或SEQ ID NO:37的氨基酸序列的信号肽的异源多肽分泌到培养基中。在一些实施方案中,在分泌过程中从蛋白切除信号肽,产生蛋白的成熟形式。

[0086] 在一些实施方案中,本发明的宿主细胞产生的异源多肽是糖基化的。在一些实施方案中,本发明产生的异源多肽的糖基化模式比酵母或大肠杆菌中产生的蛋白更紧密类似于哺乳动物的糖基化模式。在一些实施方案中,本发明的微藻宿主细胞产生的异源多肽包含N-联糖基化模式。当糖基化模式与主题生物中发现的糖基化模式相似时,用于治疗目的的糖基化蛋白不太可能促进抗糖型的免疫应答。相反地,具有不符合主题生物特征的连接或糖的糖基化蛋白更可能是有抗原性的。也可以通过特定的糖型调节效应功能。例如,IgG可以调节促-或抗-炎反应,这与其在Fc区糖型上分别不存在或存在末端唾液酸相关(Kaneko等,Science 313:670-3 (2006))。

[0087] 本发明进一步涉及一种生产重组异源多肽的方法,该方法包括在足够表达编码异源多肽的多核苷酸序列的条件下培养本发明的重组微藻宿主细胞。在一些实施方案中,重组异源多肽是从宿主细胞分泌的并从培养基中回收。在一些实施方案中,从细胞分泌的异源多肽包含分泌信号肽。根据用于生产的载体和宿主系统,本发明的重组异源多肽可以保留在重组细胞中,可以分泌到发酵培养基中,可以分泌到两层细胞膜之间的空间,或者可以保留在细胞膜的外表面。本文使用的短语“回收蛋白”指收集包含蛋白的发酵培养基而不需暗指额外的分离或纯化步骤。本发明的方法产生的异源多肽可以使用各种标准的蛋白纯化技术,例如但不限于,亲和层析、离子交换层析、过滤、电泳、疏水相互作用层析、凝胶过滤层析、反相层析、刀豆蛋白A层析、层析聚焦和差相溶解。在一些实施方案中,本发明的方法产生的异源多肽以“基本上纯”的形式分离。如本文使用的“基本上纯”指允许有效使用异源多肽作为商业产品的纯度。在一些实施方案中,重组异源多肽在细胞内积累并从细胞回收。在一些实施方案中,本方法的宿主细胞是破囊壶菌。在一些实施方案中,本方法的宿主细胞是裂殖壶菌属或破囊壶菌属的。在一些实施方案中,重组异源多肽是治疗性蛋白、食品酶或工业酶。在一些实施方案中,重组微藻宿主细胞是裂殖壶菌属的且重组异源多肽是包含分泌

信号序列的治疗性蛋白。

[0088] 在一些实施方案中,本发明的重组载体是靶向载体。本文使用的短语“靶向载体”指用于将特定核酸分子递送到重组细胞内的载体,其中该核酸分子被用来缺失或失活宿主细胞内的内源基因(即用于靶向基因破坏或敲除技术)。这种载体还被称作“敲除”载体。在一些实施方案中,靶向载体的一部分具有与宿主细胞中靶基因(即,靶向以缺失或失活的基因)的核酸序列同源的核酸序列。在一些实施方案中,插入载体的核酸分子(即插入序列)与靶基因是同源的。在一些实施方案中,载体插入序列的核酸序列被设计成与靶基因结合,使得靶基因和插入序列进行同源重组,藉此使内源靶基因缺失、失活或削弱(即,通过突变或缺失内源靶基因的至少一部分)。

[0089] 分离的核酸分子

[0090] 根据本发明,分离的核酸分子是已自其天然环境移出的核酸分子(即,已经过人为操作),其天然环境是该核酸分子在自然界中存在于其中的基因组或染色体。如此,“分离的”并不必然反映核酸分子已被纯化的程度,而是表示该分子不包括该核酸分子在自然界中存在于其中的整个基因组或整个染色体。分离的核酸分子可以包括DNA、RNA(例如mRNA),或DNA或RNA的衍生物(例如cDNA)。虽然短语“核酸分子”主要指物理的核酸分子而短语“核酸序列”或“多核苷酸序列”主要指核酸分子中的核苷酸序列,但所述短语可以互换使用,尤其是涉及能够编码异源多肽的核酸分子、多核苷酸序列、或者核酸序列。在一些实施方案中,利用重组DNA技术(例如,聚合酶链式反应(PCR)扩增,克隆)或化学合成产生本发明的分离的核酸分子。分离的核酸分子包括天然的核酸分子及其同源物,包括但不限于,天然的等位基因变体和修饰的核酸分子,其中已经将核苷酸以如下方式插入、缺失、取代和/或倒置,该方式使得这样的修饰在所编码的异源多肽的序列、功能和/或生物活性上提供期望的效果。

[0091] 启动子序列、终止子序列、信号肽序列、或任何其他序列的核酸序列互补物指与具有启动子序列、终止子序列、信号肽序列、或任何其他序列的链互补的核酸链的核酸序列。双链DNA最好包含单链DNA和其具有作为所述单链DNA互补物的序列的互补链。由此,核酸分子可以是双链或单链的,并包括那些与本发明的序列和/或本发明的序列的互补物,在“严紧”杂交条件下形成稳定杂交体的核酸分子。推断互补序列的方法是本领域熟练技术人员已知的。

[0092] 术语“多肽”包括单链多肽分子以及其中单独的组分多肽通过共价或非共价手段连接形成的多重多肽复合物。根据本发明,分离的多肽是已自其自然的环境移出的多肽(即,已经过人为操作),并可以包括例如纯化的蛋白、纯化的肽、部分纯化的蛋白、部分纯化的肽、重组生产的蛋白或肽,和合成产生的蛋白或肽。

[0093] 如本文使用的,重组微生物具有使用重组技术由其正常(即野生型或天然发生的)形式修饰(即突变或改变)的基因组。根据本发明的重组微生物可以包括这样的微生物,其中以使一种或多种修饰在该微生物中提供所需效果的方式,在其中插入、缺失、或修饰了(即突变,例如,通过核苷酸的插入、缺失、取代和/或倒置)核酸分子。如本文使用的,导致基因表达、基因功能、或基因产物(即由基因编码的蛋白)的功能降低的遗传修饰,可称作基因的失活(全部或部分)、缺失、间断、阻断或下调。例如导致基因编码蛋白的功能降低的基因遗传修饰可以是下列造成的:基因的完全缺失(即在重组微生物中不存在该基因,并因此重

组微生物中不存在该蛋白),基因突变导致蛋白翻译不完全或者不翻译(例如蛋白不表达),基因突变降低或完全消除蛋白的天然功能(例如,表达的蛋白具有降低的活性或无活性(例如酶活或作用)。导致基因表达或功能增加的遗传修饰可称作基因的扩增、过生产、过表达,活化、增强、添加或上调。

[0094] 启动子

[0095] 启动子是指导关联的编码区的转录的DNA区。

[0096] 在一些实施方案中,启动子来自网粘菌门的微生物。在一些实施方案中,启动子来自破囊壶菌,包括但不限于:以SAM2179保藏的微生物(保藏者命名为“吾肯氏壶菌属SAM2179”)、吾肯氏壶菌属(*Ulkenia*)或破囊壶菌属或裂殖壶菌属的微生物。裂殖壶菌属包括但不限于*Schizochytrium aggregatum*、*Schizochytrium limacinum*、裂殖壶菌属菌种(S31)(ATCC 20888)、裂殖壶菌属菌种(S8)(ATCC 20889)、裂殖壶菌属菌种(LC-RM)(ATCC 18915)、裂殖壶菌属菌种(SR 21)、保藏的裂殖壶菌菌株ATCC 28209、和保藏的裂殖壶菌菌株IFO 32693。

[0097] 启动子可以至少在破囊壶菌中具有启动子活性,并包括全长启动子序列及其功能片段、融合序列,和自然发生的启动子的同源物。启动子的同源物与自然发生的启动子区别在于至少一个、两个、三个或几个核苷酸已被缺失、插入、倒置、取代和/或衍生化。虽然活性可以增加、减少、或使其依赖于某些刺激,启动子的同源物至少在破囊壶菌中能够保留作为启动子的活性。启动子可以包含一个或多个赋予发育和组织特异性调节性调控或表达的序列元件。

[0098] 在一些实施方案中,本文所述的分离的核酸分子包含PUFA PKS OrfC启动子(“PKS OrfC启动子”,也称作PFA3启动子),例如,SEQ ID NO:3代表的多核苷酸序列。PKS OrfC启动子包括与自然发生的PKS OrfC启动子序列足够相似的PKS OrfC启动子同源物,同源物的核酸序列能够在适中、高或非常高的严紧条件下与自然发生的PKS OrfC启动子,例如SEQ ID NO:3,或以ATCC登录号PTA-9615保藏的pCL0001的OrfC启动子的核酸序列的互补序列杂交。

[0099] 在一些实施方案中,本发明分离的核酸分子包含EF1短启动子(“EF1短”或“EF1-S”启动子)或EF1长(“EF1长”或“EF1-L”启动子),例如,SEQ ID NO:42代表的EF1短启动子,或SEQ ID NO:43代表的EF1长启动子。EF1短启动子或EF1长启动子包括分别与自然发生的EF1短和/或长启动子序列足够相似的EF1短或长启动子的同源物,同源物的核酸序列能够在适中、高或非常高的严紧条件下与自然发生的EF1短和/或长启动子,例如分别是SEQ ID NO:42和/或SEQ ID NO:43,或以ATCC登录号PTA-9616保藏的pAB0018的EF1长启动子的核酸序列的互补序列杂交。

[0100] 在一些实施方案中,本发明分离的核酸分子包含60S短启动子(“60S短”或“60S-S”启动子)或60S长启动(“60S长”或“60S-L”启动子)例如,SEQ ID NO:44代表的60S短启动子,或具有SEQ ID NO:45代表的多核苷酸序列的60S长启动子。在一些实施方案中,60S短或60S长启动子包括分别与自然发生的60S短或60S长启动子序列足够相似的60S短或60S长启动子的同源物,同源物的核酸序列能够在适中、高或非常高的严紧条件下与自然发生的60S短和/或60S长启动子,例如分别是SEQ ID NO:44和/或SEQ ID NO:45,或以ATCC登录号PTA-9614保藏的pAB0011的60S长启动子的核酸序列的互补序列杂交。

[0101] 在一些实施方案中,分离的核酸分子包含Sec1启动子(“Sec1启动子”),例如,SEQ

ID NO:46代表的多核苷酸序列。在一些实施方案中,Sec1启动子包括与自然发生的Sec1启动子序列足够相似的Sec1启动子同源物,同源物的核酸序列能够在适中、高或非常高的严紧条件下与自然发生的Sec1启动子,例如SEQ ID NO:46,或以ATCC登录号PTA-9613保藏的pAB0022的Sec1启动子的核酸序列的互补序列杂交。

[0102] 终止子

[0103] 终止子区是遗传序列的一部分,其标志基因组DNA中的基因序列转录的结束。

[0104] 在一些实施方案中,终止子区来自网粘菌门的微生物。在一些实施方案中,终止子区来自破囊壶菌。在一些实施方案中,终止子区来自裂殖壶菌属或破囊壶菌属。裂殖壶菌属包括但不限于*Schizochytrium aggregatum*,*Schizochytrium limacinum*,裂殖壶菌属菌种(S31) (ATCC 20888),裂殖壶菌属菌种(S8) (ATCC 20889),裂殖壶菌属菌种(LC-RM) (ATCC 18915),裂殖壶菌属菌种(SR 21),保藏菌株ATCC 28209,和保藏菌株IFO 32693。在一些实施方案中,终止子区是异源终止子区,例如,异源SV40的终止子区。

[0105] 终止子区至少在破囊壶菌中具有终止活性并包括全长终止子序列及其功能片段、融合序列,和自然发生的终止子区的同源物。终止子的同源物与自然发生的终止子区别在于至少一个或几个,但不限于一个或几个核苷酸已被缺失、插入、倒置、取代和/或衍生化。在一些实施方案中,虽然活性可以增加、减少、或使其依赖于某些刺激,终止子的同源物至少在破囊壶菌中保留作为终止子区的活性。

[0106] 在一些实施方案中,分离的核酸分子可以包含PUFA PKS OrfC基因的终止子区(“PKS OrfC终止子区”,也称作PFA3终止子),例如,SEQ ID NO:4代表的多核苷酸序列。SEQ ID NO:4公开的终止子区是来自破囊壶菌微生物的自然发生的(野生型)终止子序列,并特别是裂殖壶菌属PKS OrfC终止子区并被称为“OrfC终止子元件1”。在一些实施方案中,PKS OrfC终止子区包括与自然发生的PUFA PKS OrfC终止子区足够相似的PKS OrfC终止子区同源物,所述同源物的核酸序列能够在适中、高或非常高的严紧条件下与自然发生的PKS OrfC终止子区,例如SEQ ID NO:4,或以ATCC登录号PTA-9614保藏的pAB0011的OrfC终止子区的核酸序列的互补序列杂交。

[0107] 信号肽

[0108] 在一些实施方案中,分离的核酸分子可以包含编码来自网粘菌门微生物的分泌蛋白的信号肽的多核苷酸序列。在一些实施方案中,微生物是破囊壶菌。在一些实施方案中,微生物是裂殖壶菌属或破囊壶菌属的。

[0109] 信号肽可以在破囊壶菌中具有分泌信号活性,并包括全长肽及其功能片段、融合肽和自然发生的信号肽的同源物。信号肽的同源物与自然发生的信号肽区别在于至少一个或少数几个,但不限于一个或少数几个氨基酸已被缺失(例如蛋白的截短版本,如肽或片段)、插入、倒置、取代和/或衍生化(例如,通过糖基化、磷酸化、乙酰化、豆蔻酰化、异戊烯化、棕榈酰化、酰胺化、和/或加入糖基磷脂酰肌醇)。在一些实施方案中,虽然活性可以是提高的、降低的、或使其依赖于某些刺激,信号肽的同源物至少在破囊壶菌中保留作为信号肽的活性。

[0110] 在一些实施方案中,分离的核酸分子包含编码Na/Pi-II2转运蛋白信号肽的多核苷酸序列。Na/Pi-II2转运蛋白信号肽至少在破囊壶菌中至少对Na/Pi-II2转运蛋白具有信号靶向活性,并包括全长肽及其功能片段、融合肽和自然发生的Na/Pi-II2转运蛋白信号肽

的同源物。在一些实施方案中,Na/Pi-II2转运蛋白信号肽具有SEQ ID NO:1代表的氨基酸序列。在一些实施方案中,Na/Pi-II2转运蛋白信号肽具有SEQ ID NO:15代表的氨基酸序列。在一些实施方案中,分离的核酸分子包含编码分离的包含SEQ ID NO:1或SEQ ID NO:15功能性片段的氨基酸序列的多核苷酸序列,所述功能性片段至少在破囊壶菌中,至少对于Na/Pi-II2转运蛋白是作为信号肽发挥功能。在一些实施方案中,分离的核酸分子包含SEQ ID NO:2。

[0111] 本发明也涉及包含Na/Pi-II2转运蛋白信号肽氨基酸序列的分离的多肽。

[0112] 在一些实施方案中,分离的核酸分子包含编码 α -1,6-甘露糖基转移酶 (ALG12) 信号肽的多核苷酸序列。ALG12信号肽至少在破囊壶菌中,至少对ALG12蛋白能够具有信号靶向活性,并包括全长肽及其功能片段、融合肽和自然发生的ALG12信号肽的同源物。在一些实施方案中,ALG12信号肽具有SEQ ID NO:59代表的氨基酸序列。在一些实施方案中,分离的核酸分子包含编码分离的含有SEQ ID NO:59功能性片段的氨基酸序列的多核苷酸序列,所述功能性片段至少在破囊壶菌中,至少对于ALG12蛋白是作为信号肽发挥功能。在一些实施方案中,分离的核酸分子包含SEQ ID NO:60。

[0113] 本发明也涉及包含ALG12信号肽氨基酸序列的分离的多肽。

[0114] 在一些实施方案中,分离的核酸分子包含编码免疫球蛋白结合蛋白 (BiP) 信号肽的多核苷酸序列。BiP信号肽至少在破囊壶菌中,至少对BiP蛋白能够具有信号靶向活性,并包括全长肽及其功能片段、融合肽和自然发生的BiP信号肽的同源物。在一些实施方案中,BiP信号肽具有SEQ ID NO:61代表的氨基酸序列。在一些实施方案中,分离的核酸分子包含编码含有SEQ ID NO:61功能性片段的分离的氨基酸序列的多核苷酸序列,所述功能性片段至少在破囊壶菌中,至少对于BiP蛋白是作为信号肽发挥功能。在一些实施方案中,分离的核酸分子包含SEQ ID NO:62。

[0115] 本发明也涉及包含BiP信号肽氨基酸序列的分离的多肽。

[0116] 在一些实施方案中,分离的核酸分子包含编码 α -1,3-葡糖苷酶 (GLS2) 信号肽的多核苷酸序列。GLS2信号肽至少在破囊壶菌中,至少对GLS2蛋白能够具有信号靶向活性,并包括全长肽及其功能片段、融合肽和自然发生的GLS2信号肽的同源物。在一些实施方案中,GLS2信号肽具有SEQ ID NO:63代表的氨基酸序列。在一些实施方案中,分离的核酸分子包含编码含有SEQ ID NO:63功能性片段的分离的氨基酸序列的多核苷酸序列,所述功能性片段至少在破囊壶菌中,至少对于GLS2蛋白是作为信号肽发挥功能。在一些实施方案中,分离的核酸分子包含SEQ ID NO:64。

[0117] 本发明也涉及包含GLS2信号肽氨基酸序列的分离的多肽。

[0118] 在一些实施方案中,分离的核酸分子包含编码 α -1,3-1,6-甘露糖苷酶样信号肽的多核苷酸序列。 α -1,3-1,6-甘露糖苷酶样信号肽至少在破囊壶菌中,至少对 α -1,3-1,6-甘露糖苷酶样蛋白能够具有信号靶向活性,并包括全长肽及其功能片段、融合肽和自然发生的 α -1,3-1,6-甘露糖苷酶样信号肽的同源物。在一些实施方案中, α -1,3-1,6-甘露糖苷酶样信号肽具有SEQ ID NO:65代表的氨基酸序列。在一些实施方案中,分离的核酸分子包含编码含有SEQ ID NO:65功能性片段的分离的氨基酸序列的多核苷酸序列,所述功能性片段至少在破囊壶菌中,至少对于 α -1,3-1,6-甘露糖苷酶样蛋白是作为信号肽发挥功能。在一些实施方案中,分离的核酸分子包含SEQ ID NO:66。

[0119] 本发明也涉及包含 α -1,3-1,6-甘露糖苷酶样信号肽氨基酸序列的分离的多肽。

[0120] 在一些实施方案中,分离的核酸分子包含编码 α -1,3-1,6-甘露糖苷酶样#1信号肽的多核苷酸序列。 α -1,3-1,6-甘露糖苷酶样#1信号肽至少在破囊壶菌中,至少对 α -1,3-1,6-甘露糖苷酶样#1蛋白能够具有信号靶向活性,并包括全长肽及其功能片段、融合肽和自然发生的 α -1,3-1,6-甘露糖苷酶样#1信号肽的同源物。在一些实施方案中, α -1,3-1,6-甘露糖苷酶样#1信号肽具有SEQ ID NO:67代表的氨基酸序列。在一些实施方案中,分离的核酸分子包含编码含有SEQ ID NO:67功能性片段的分离的氨基酸序列的多核苷酸序列,所述功能性片段至少在破囊壶菌中,至少对于 α -1,3-1,6-甘露糖苷酶样#1蛋白是作为信号肽发挥功能。在一些实施方案中,分离的核酸分子包含SEQ ID NO:68。

[0121] 本发明也涉及包含 α -1,3-1,6-甘露糖苷酶样#1信号肽氨基酸序列的分离的多肽。

[0122] 在一些实施方案中,分离的核酸分子包含编码 α -1,2-甘露糖苷酶样信号肽的多核苷酸序列。 α -1,2-甘露糖苷酶样信号肽至少在破囊壶菌中,至少对 α -1,2-甘露糖苷酶样蛋白能够具有信号靶向活性,并包括全长肽及其功能片段、融合肽和自然发生的 α -1,2-甘露糖苷酶样信号肽的同源物。在一些实施方案中, α -1,2-甘露糖苷酶样信号肽具有SEQ ID NO:69代表的氨基酸序列。在一些实施方案中,分离的核酸分子包含编码含有SEQ ID NO:69功能性片段的分离的氨基酸序列的多核苷酸序列,所述功能性片段至少在破囊壶菌中,至少对于 α -1,2-甘露糖苷酶样蛋白是作为信号肽发挥功能。在一些实施方案中,分离的核酸分子包含SEQ ID NO:70。

[0123] 本发明也涉及包含 α -1,2-甘露糖苷酶样信号肽氨基酸序列的分离的多肽。

[0124] 在一些实施方案中,分离的核酸分子包含编码 β -木糖苷酶样信号肽的多核苷酸序列。 β -木糖苷酶样信号肽至少在破囊壶菌中,至少对 β -木糖苷酶样蛋白能够具有信号靶向活性,并包括全长肽及其功能片段、融合肽和自然发生的 β -木糖苷酶样信号肽的同源物。在一些实施方案中, β -木糖苷酶样信号肽具有SEQ ID NO:71代表的氨基酸序列。在一些实施方案中,分离的核酸分子包含编码含有SEQ ID NO:71功能性片段的分离的氨基酸序列的多核苷酸序列,所述功能性片段至少在破囊壶菌中,至少对于 β -木糖苷酶样蛋白是作为信号肽发挥功能。在一些实施方案中,分离的核酸分子包含SEQ ID NO:72。

[0125] 本发明也涉及包含 β -木糖苷酶样信号肽氨基酸序列的分离的多肽。

[0126] 在一些实施方案中,分离的核酸分子包含编码胡萝卜素合酶信号肽的多核苷酸序列。胡萝卜素合酶信号肽至少在破囊壶菌中,至少对胡萝卜素合酶蛋白能够具有信号靶向活性,并包括全长肽及其功能片段、融合肽和自然发生的胡萝卜素合酶信号肽的同源物。在一些实施方案中,胡萝卜素合酶信号肽具有SEQ ID NO:73代表的氨基酸序列。在一些实施方案中,分离的核酸分子包含编码含有SEQ ID NO:73功能性片段的分离的氨基酸序列的多核苷酸序列,所述功能性片段至少在破囊壶菌中,至少对于胡萝卜素合酶蛋白是作为信号肽发挥功能。在一些实施方案中,分离的核酸分子包含SEQ ID NO:74。

[0127] 本发明也涉及包含胡萝卜素合酶信号肽氨基酸序列的分离的多肽。

[0128] 在一些实施方案中,分离的核酸分子包含编码Sec1蛋白(“Sec1”)信号肽的多核苷酸序列。Sec1信号肽至少在破囊壶菌中,至少对Sec1蛋白能够具有信号靶向活性,并包括全长肽及其功能片段、融合肽和自然发生的Sec1信号肽的同源物。在一些实施方案中,Sec1信号肽由SEQ ID NO:37代表。在一些实施方案中,分离的核酸分子包含编码含有SEQ ID NO:

37功能性片段的分离的氨基酸序列的多核苷酸序列,所述功能性片段至少在破囊壶菌中,至少对于Sec1蛋白是作为信号肽发挥功能。在一些实施方案中,分离的核酸分子包含SEQ ID NO:38。

[0129] 本发明也涉及包含Sec1信号肽氨基酸序列的分离的多肽。

[0130] 在一些实施方案中,分离的核酸分子可以包含与本文描述的任何启动子、终止子和/或信号肽序列是至少90%、95%、96%、97%、98%、或99%相同的启动子序列、终止子序列、和/或信号肽序列。

[0131] 在一些实施方案中,分离的核酸分子包含OrfC启动子、EF1短启动子、EF1长启动子、60S短启动子、60S长启动子、Sec1启动子、PKS OrfC终止子区、编码Na/Pi-IIb2转运蛋白信号肽的序列、或编码Sec1转运蛋白信号肽的序列,上述序列与编码异源多肽的核酸序列的5'端可操作地连接。重组载体(包括但不限于表达载体)、表达盒、和宿主细胞也可以包含OrfC启动子、EF1短启动子、EF1长启动子、60S短启动子、60S长启动子、Sec1启动子、PKS OrfC终止子区、编码Na/Pi-IIb2转运蛋白信号肽的序列、或编码Sec1转运蛋白信号肽的序列,上述序列与编码异源多肽的核酸序列的5'端可操作地连接。

[0132] 如本文所使用的,除非另有说明,百分比(%)同一性(和%同一性)涉及一种同源性评估,其是如下进行的:(1)BLAST 2.0Basic BLAST同源性搜索,其中使用blastp进行氨基酸搜索和使用blastn进行核酸搜索,均使用标准缺省参数,其中缺省对查询序列进行低复杂性区(low complexity regions)过滤(参见,例如Altschul,S.等,Nucleic Acids Res.25:3389-3402(1997),本文引入其全部内容作为参考);(2)使用下文所述的参数进行BLAST 2比对;(3)和/或PSI-BLAST(Position-Specific Iterated BLAST),使用标准缺省参数。需要注意,由于BLAST 2.0BasicBLAST和BLAST 2之间标准参数有一些差异,所以两个具体的序列可能在使用BLAST 2程序时被识别为具有显著同源性,而用其中一个序列作为查询序列在BLAST 2.0Basic BLAST中进行搜索时,第二个序列可能不会被确定为最匹配的结果之一。此外,PSI-BLAST提供了一种自动、易用的“谱”(“profile”)搜索形式,这是一种灵敏的寻找序列同源物的方法。程序首先执行有缺口的(gapped)BLAST数据库搜索。PSI-BLAST程序使用来源于任何返回的显著性匹配(significant alignments)的信息构建位置特异性评分矩阵(position-specific score matrix),其取代查询序列用于下一轮数据库搜索。因此,可以理解,百分比同一性可使用这些程序中的任何一种来确定。

[0133] 两个具体的序列可以使用BLAST 2序列相互比对,如例如Tatusova和Madden,FEMS Microbiol Lett.174:247-250(1999)所述的那样,本文引入其全部内容作为参考。在blastp或blastn中进行BLAST 2序列比对,使用BLAST2.0算法在两个序列之间进行有缺口的BLAST搜索(Gapped BLAST)(BLAST 2.0),允许在所得比对中引入缺口(缺失和插入)。在一些实施方案中,BLAST 2序列比对使用如下的标准缺省参数进行。

[0134] 对于blastn,使用0BLOSUM62矩阵:

[0135] 匹配加分(reward for match)=1

[0136] 错配罚分(penalty for mismatch)=-2

[0137] 开放缺口(open gap)(5)和延伸缺口(extension gap)(2)罚分

[0138] 缺口(gap)x_脱落(dropoff)(50)期望(expect)(10)

[0139] 字长(word size)(11)过滤(filter)(on)

[0140] 对于blastp使用OBLOSUM62矩阵：

[0141] 开放缺口 (11) 和延伸缺口 (1) 罚分

[0142] 缺口x脱落 (50) 期望 (10) 字长 (3) 过滤 (on)。

[0143] 如本文使用的, 杂交条件是指使用核酸分子鉴定相似的核酸分子的标准杂交条件。参见, 例如Sambrook J和Russell D, Molecular cloning: A laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 本文其全部内容作为参考。此外, 在例如Meinkoth等, Anal. Biochem. 138, 267-284 (1984) 中公开了计算的合适杂交和洗涤条件, 以实现允许不同程度的核苷酸错配的杂交的公式, 本文其全部内容作为参考。本领域技术人员可以使用Meinkoth等中的公式, 例如, 计算合适的杂交和洗涤条件, 以实现特定的核苷酸错配水平。这些条件会根据所形成的是DNA:RNA还是DNA:DNA杂交体而变化。DNA:DNA杂交体算得的解链温度比DNA:RNA杂交体低10℃。在具体的实施方案中, 对于DNA:DNA杂交体的严紧杂交条件包括在20℃到35℃之间 (低严紧度), 在28℃到40℃之间 (更严紧的), 在35℃到45℃之间 (更严紧的) 的温度下, 在6X SSC (0.9M Na⁺) 的离子强度下进行杂交, 以及合适的洗涤条件。在具体的实施方案中, 对于DNA:RNA杂交体的严紧杂交条件包括在30℃到45℃之间, 在38℃到50℃之间, 在45℃到55℃之间的温度下, 在6X SSC (0.9MNa⁺) 的离子强度下进行杂交, 以及相似的严紧洗涤条件。这些值是基于大于约100个核苷酸的分子、0%甲酰胺、以及大约40%的G+C含量的解链温度的计算结果。或者, T_m可以根据Sambrook等提出的那样凭经验计算。一般地, 洗涤条件应尽可能严紧, 并应适合于所选的杂交条件。例如, 杂交条件可包括盐和温度条件的组合, 其中温度条件比特定杂交体的计算T_m低大约20-25℃, 而洗涤条件通常包括盐和温度条件的组合, 其中温度条件比特定杂交体的计算T_m低大约12-20℃。适合于DNA:DNA杂交体使用的杂交条件的一个例子包括在42℃在6X SSC (50%甲酰胺) 中杂交2-24小时, 随后进行洗涤步骤, 其包括在室温下在2X SSC中洗涤1次或多次, 随后在更高温度和更低离子强度下进行进一步的洗涤 (例如, 在37℃和0.1X-0.5X SSC中进行至少一次洗涤, 随后在68℃和0.1X-0.5X SSC中进行至少一次洗涤)。

[0144] 异源多肽

[0145] 本文使用的术语“异源”指非天然存在于微藻宿主细胞中的序列。在一些实施方案中, 本发明的重组宿主细胞产生的异源多肽包括但不限于治疗性蛋白。本文使用的术语“治疗性蛋白”包括在动物或人中用于治疗或预防疾病、状况或失调的蛋白。

[0146] 在某些实施方案中, 治疗性蛋白包括但不限于生物活性蛋白, 例如酶、抗体、或抗原蛋白。

[0147] 在一些实施方案中, 本发明的重组宿主细胞产生的异源多肽包括但不限于工业酶。工业酶包括但不限于用于产品, 包括食品、医疗、化学、机械和其他工业产品的生产、配制、保存、营养动员、或加工的酶。

[0148] 在一些实施方案中, 本发明的重组宿主细胞产生的异源多肽包括营养缺陷标记、显性选择标记 (例如, 降解抗生素活性的酶), 或另一种参与转化选择的蛋白、作为报道子发挥功能的蛋白、参与蛋白糖基化的酶、和参与细胞代谢的酶。

[0149] 在一些实施方案中, 本发明的重组宿主细胞产生的异源多肽包括选自下组的病毒蛋白: H或HA (血凝素) 蛋白、N或NA (神经氨酸酶) 蛋白、F (融合) 蛋白、G (糖蛋白) 蛋白、E或env (包膜) 蛋白、gp120 (120kDa糖蛋白)、和gp41蛋白 (41kDa糖蛋白)。在一些实施方案中, 本发

明的重组宿主细胞产生的异源多肽是病毒基质蛋白。在一些实施方案中,本发明的重组宿主细胞产生的异源多肽是选下组的病毒基质蛋白:M1、M2(膜通道蛋白)、Gag及其组合。在一些实施方案中,HA、NA、F、G、E、gp120、gp41或基质蛋白来自病毒来源,如流感病毒或麻疹病毒。

[0150] 流感是由于呼吸道病毒导致人类死亡的首要原因。常见的症状包括发烧、喉咙痛、呼吸急促、和肌肉酸痛等。流感病毒是从被感染的哺乳动物和鸟类细胞的质膜出芽的包膜病毒。根据核蛋白和基质蛋白抗原的存在,将它们分为甲、乙或丙型。甲型流感病毒根据存在的HA和NA表面糖蛋白的组合可以进一步划分为亚型。HA是抗原性糖蛋白,并在病毒与其感染的细胞的结合中发挥功能。NA从宿主细胞上的聚糖链和病毒表面蛋白去除末端的唾液酸残基,从而防止病毒聚集并利于病毒的运动。

[0151] 流感病毒HA蛋白是同源三聚体,在每个单体的球状头部上具有受体结合口袋,而流感病毒NA蛋白是四聚体,在每个单体的头部具有酶活位点。目前,认可了16个HA(H1-H16)和9个NA(N1-N9)亚型。每种甲型流感病毒呈现一种类型的HA和一种类型的NA糖蛋白。一般来说,每个亚型具有物种特异性,例如,已知所有的HA和NA亚型感染鸟类,而只有H1、H2、H3、H5、H7、H9、H10、N1、N2、N3和N7亚型已被证明感染人类。流感病毒是由其携带的HA和NA的类型表征的,如H1N1、H5N1、H1N2、H1N3、H2N2、H3N2、H4N6、H5N2、H5N3、H5N8、H6N1、H7N7、H8N4、H9N2、H10N3、H11N2、H11N9、H12N5、H13N8、H15N8、H16N3等。亚型进一步分为株;每个遗传上不同的病毒分离株通常被认为是一个单独的株系,例如,甲型流感/Puerto Rico/8/34/Mount Sinai(H1N1)和甲型流感/Vietnam/1203/2004(H5N1)。在本发明的某些实施方案中,HA来自流感病毒,例如HA来自甲型流感、乙型流感、或者甲型流感的亚型,其选自下组:H1、H2、H3、H4、H5、H6、H7、H8、H9、H10、H11、H12、H13、H14、H15、和H16。在另一个实施方案中,HA来自甲型流感病毒,其选自下组:H1、H2、H3、H5、H6、H7和H9。在一个实施方案中,HA来自流感H1N1亚型。

[0152] 流感病毒HA蛋白在细胞中作为单一的蛋白翻译,其切除信号肽后是约62kDa的蛋白(通过概念翻译),称为HA0(即血凝素前体蛋白)。为了病毒活化,血凝素前体蛋白(HA0)必须由胰蛋白酶样丝氨酸内切蛋白酶在特定位点切割,通常是由蛋白的HA1和HA2多肽间的一个单一的碱性氨基酸(通常是精氨酸)编码。在A/Puerto Rico/8/34株的特定实例中,这种切割发生于氨基酸343位的精氨酸和氨基酸344位的甘氨酸之间。切割后,两个二硫键-键合的蛋白多肽产生蛋白亚基的成熟形式,作为融合及进而病毒感染所必需的构象变化的先决条件。

[0153] 在一些实施方案中,本发明的HA蛋白是经切割的,例如,将本发明的HA0蛋白切割成HA1和HA2。在一些实施方案中,在微藻宿主细胞,如裂殖壶菌属中表达HA蛋白,结果HA0蛋白适当的切割为功能性HA1和HA2多肽,而无需添加外源蛋白酶。非脊椎动物表达系统中的这种不添加外源蛋白酶的血凝素切割之前一直没有被证实过。

[0154] 病毒F蛋白可以在C-端附近包括单次跨膜结构域。可以将F蛋白在Furin切割位点(氨基酸109)处分为两个肽。称为F2的蛋白第一部分包含完整F蛋白的N-端部分。含有F蛋白C-端部分的病毒F蛋白的其余部分称为F1。F1和/或F2区可以单独融合至异源序列,例如编码异源信号肽的序列。包含病毒F蛋白的F1和F2部分的载体可以单独表达或联合表达。表达完整F蛋白的载体可以与会在Furin切割位点切割蛋白的Furin酶共表达。或者,可以用编码

不同蛋白酶识别并切割的其他蛋白酶切割位点的序列替换编码F蛋白Furin切割位点的序列。包含替代的蛋白酶切割位点的F蛋白可以与相应的识别和切割替代蛋白酶切割位点的蛋白酶共表达。

[0155] 在一些实施方案中,HA、NA、F、G、E、gp120、gp41或基质蛋白是其全长蛋白、片段、变体、衍生物、或类似物。在一些实施方案中,HA、NA、F、G、E、gp120、gp41或基质蛋白是包含与代表性病毒蛋白的已知序列是至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%相同或100%相同的氨基酸序列的多肽,或其多核苷酸编码上述多肽,其中所述多肽是通过特异性结合已知序列的抗体可识别的。例如HA序列可以是全长HA蛋白,其基本上由胞外(ECD)结构域、跨膜(TM)结构域和胞质(CYT)结构域组成;或是完整HA蛋白的片段,其基本上由HA1多肽和HA2多肽(例如由切割全长HA产生)组成;或是完整HA蛋白的片段,其基本上由HA1多肽、HA2多肽和TM结构域组成;或是完整HA蛋白的片段,其基本上由CYT结构域组成;或是完整HA蛋白的片段,其基本上由TM结构域组成;或是完整HA蛋白的片段,其基本上由HA1多肽组成;或是完整HA蛋白的片段,其基本上由HA2多肽组成。HA序列也可以包括HA1/HA2切割位点。HA1/HA2切割位点可以位于HA1和HA2多肽之间,也可以相对于其他多核苷酸或多肽构建体序列按任何顺序安排。病毒蛋白可以来自病原性病毒株。

[0156] 在一些实施方案中,本发明的异源多肽是包含全长HA、NA、F、G、E、gp120、gp41或基质蛋白,或其片段、变体、衍生物、或类似物的融合多肽。

[0157] 在一些实施方案中,异源多肽是包含HA0多肽、HA1多肽、HA2多肽、TM结构域、其片段和其组合的融合多肽。在一些实施方案中,异源多肽包含来自不同亚型或不同病毒株,例如流感病毒的不同亚型或病毒株的两个或两个以上的HA1多肽、HA2多肽、TM结构域、或其片段的组合。在一些实施方案中,异源多肽包含来自不同病毒,例如流感病毒和麻疹病毒的两个或两个以上的HA1多肽、HA2多肽、TM结构域、或其片段的组合。

[0158] 可以通过测量红细胞的凝集确定血凝活性。血凝和红细胞随后的沉淀是血凝素被吸附到红细胞表面的结果。红细胞的集群,肉眼分辨为堆、肿块、和/或团块,在血凝过程中形成。血凝是由红细胞中存在的凝集原与含有凝集素的血浆相互作用导致的。每种凝集原具有相应的凝集素。血凝反应用于,例如,确定病毒的抗血清活性或类型。由红血细胞上的因子的直接作用导致的主动血凝和由针对红血细胞先前吸附的抗原的特异性的抗血清引起的被动血凝之间具有区别。可以测量样品中血凝活性的量,例如,以血凝活性单位(HAU)。血凝可能由例如,结核、鼠疫和兔热病的致病细菌的多糖,大肠杆菌的多糖,和流感病毒、流行性腮腺炎病毒、白鼠肺炎病毒、猪和马流感病毒、天花病毒疫苗、黄热病病毒、和其他血凝诱导性疾病病毒引起。

[0159] 微藻胞外体

[0160] 本发明也涉及微藻胞外体,其中所述胞外体与质膜是不连续的。“与质膜是不连续的”是指微藻胞外体没有与宿主细胞的质膜相连接。在一些实施方案中,胞外体是膜。在一些实施方案中,胞外体是囊泡、胶束、膜片段、膜聚集体、或其混合物。本文使用的术语“囊泡”指包含脂双层(单位膜)的封闭结构,例如细胞膜形成的泡状结构。本文使用的术语“膜聚集体”指作为单一物质形成关联的任何膜结构的集合。膜聚集体可以是单一类型膜结构的集合,例如但不限于膜囊泡的集合,或者也可以是多于一种类型膜结构的集合,例如但不

限于,至少囊泡、胶束或膜片段中两种的集合。本文使用的术语“膜片段”指能够包含本文所述的异源多肽的膜的任何部分。在一些实施方案中,膜片段是膜片层。在一些实施方案中,胞外体是囊泡和膜片段的混合物。在一些实施方案中,胞外体是囊泡。在一些实施方案中,囊泡是崩解的囊泡。在一些实施方案中,囊泡是病毒样颗粒。在一些实施方案中,胞外体是包含宿主细胞所产生的天然和异源多肽的生物材料的聚集体。在一些实施方案中,胞外体是天然和异源多肽的聚集体。在一些实施方案中,胞外体是异源多肽的聚集体。

[0161] 在一些实施方案中,微藻宿主细胞的质外网在微藻宿主细胞培养过程中变为片段化,导致微藻胞外体的形成。在一些实施方案中,通过微藻宿主细胞质外网的片段化形成微藻胞外体,片段化是在搅拌的培养基中流体动力物理剪切质外网膜延伸物的结果。

[0162] 在一些实施方案中,微藻胞外体是通过挤压微藻膜形成的,例如但不限于挤压质膜、质外网、类伪根、或其组合,其中挤出膜使其变得与质膜分离。

[0163] 在一些实施方案中,微藻胞外体是具有不同直径的囊泡或胶束、具有不同长度的膜片段,或其组合。

[0164] 在一些实施方案中,胞外体是具有10nm-2500nm,10nm-2000nm,10nm-1500nm,10nm-1000nm,10nm-500nm,10nm-300nm,10nm-200nm,10nm-100nm,10nm-50nm,20nm-2500nm,20nm-2000nm,20nm-1500nm,20nm-1000nm,20nm-500nm,20nm-300nm,20nm-200nm,20nm-100nm,50nm-2500nm,50nm-2000nm,50nm-1500nm,50nm-1000nm,50nm-500nm,50nm-300nm,50nm-200nm,50nm-100nm,100nm-2500nm,100nm-2000nm,100nm-1500nm,100nm-1000nm,100nm-500nm,100nm-300nm,100nm-200nm,500nm-2500nm,500nm-2000nm,500nm-1500nm,500nm-1000nm,2000nm或更小,1500nm或更小,1000nm或更小,500nm或更小,400nm或更小,300nm或更小,200nm或更小,100nm或更小,或50nm或更小的直径的囊泡。

[0165] 破囊壶菌宿主细胞产生微藻胞外体的非限制性发酵条件如下表1所示。

[0166] 表1:容器培养基

	<u>成分</u>	<u>浓度</u>	<u>范围</u>
	Na ₂ SO ₄	g/L 13.62	0-50, 15-45, 或 25-35
	K ₂ SO ₄	g/L 0.72	0-25, 0.1-10, 或 0.5-5
	KCl	g/L 0.56	0-5, 0.25-3, 或 0.5-2
[0167]	MgSO ₄ ·7H ₂ O	g/L 2.27	0-10, 1-8, 或 2-6
	(NH ₄) ₂ SO ₄	g/L 17.5	0-50, 0.25-30, 或 5-20
	CaCl ₂ ·2H ₂ O	g/L 0.19	0.1-5, 0.1-3, 或 0.15-1
	KH ₂ PO ₄	g/L 6.0	0-20, 0.1-10, 或 1-7

[0168] 高压灭菌后(金属)

	柠檬酸	mg/L	3.50	0.1-5000, 1-3000, 或 3-2500
	FeSO ₄ ·7H ₂ O	mg/L	51.5	0.1-1000, 1-500, 或 5-100
	MnCl ₂ ·4H ₂ O	mg/L	3.10	0.1-100, 1-50, 或 2-25
	ZnSO ₄ ·7H ₂ O	mg/L	6.20	0.1-100, 1-50, 或 2-25
[0169]	CoCl ₂ ·6H ₂ O	mg/L	0.04	0-1, 0.001-0.1, 或 0.01-0.1
	Na ₂ MoO ₄ ·2H ₂ O	mg/L	0.04	0.001-1, 0.005-0.5, 或 0.01-0.1
	CuSO ₄ ·5H ₂ O	mg/L	2.07	0.1-100, 0.5-50, 或 1-25
	NiSO ₄ ·6H ₂ O	mg/L	2.07	0.1-100, 0.5-50, 或 1-25

[0170] 高压灭菌后(维生素)

[0171] 硫胺素** mg/L 9.75 0.1-100, 1-50, 或 5-25

[0172] 维生素B12** mg/L 0.16 0.01-100, 0.05-5, 或 0.1-1.0

[0173] Ca1/2-泛酸盐** mg/L 3.33 0.1-100, 0.1-50, 或 1-10

[0174] **过滤灭菌并在高压灭菌后加入

[0175] 高压灭菌后(碳)

[0176] 葡萄糖 g/L 20.0 5-150, 10-100, 或 20-50

[0177] 氮给料:

[0178] 成分 浓度

[0179] NH₄OH mL/L 23.6 5-150, 10-100, 15-50

[0180] 生产微藻胞外体的常用培养条件包括下列:

[0181] pH: 5.5-9.5, 6.5-8.0, 或 6.3-7.3

[0182] 温度: 15°C-45°C, 18°C-35°C 或 20°C-30°C

[0183] 溶解氧: 0.1%-100%饱和, 5%-50%饱和, 或 10%-30%饱和

[0184] 受控的葡萄糖: 5g/L-100g/L, 10g/L-40g/L, 或 15g/L-35g/L。

[0185] 在一些实施方案中, 微藻胞外体产自网粘菌门宿主细胞。在一些实施方案中, 微藻胞外体产自网粘菌目宿主细胞。在一些实施方案中, 微藻胞外体产自破囊壶菌宿主细胞。在一些实施方案中, 微藻胞外体产自裂殖壶菌属或破囊壶菌属。

[0186] 本发明也涉及包含异源多肽的微藻胞外体, 其中所述胞外体与微藻宿主细胞的质膜是不连续的。

[0187] 在一些实施方案中, 本发明的微藻胞外体包含的多肽也与微藻宿主细胞的质膜相关联。在一些实施方案中, 关联微藻宿主细胞质膜的多肽包括天然膜多肽、异源多肽及其组合。

[0188] 在一些实施方案中, 异源多肽包含于微藻胞外体中。

[0189] 在一些实施方案中, 异源多肽包含膜结构域。本文所述的术语“膜结构域”是指使多肽靶向膜和/或允许多肽保持与膜相关联的多肽内的任何结构域, 而且包括但不限于, 跨膜结构域(例如单次或多次跨膜区)、完整单境域(integral monotopic domain)、信号锚序列、ER信号序列、N-端或内部或C-端终止转移信号、糖基磷脂酰肌醇锚, 及其组合。膜结构域可以位于多肽的任何位置, 包括多肽的N-端、C-端或中间。膜结构域可以与永久或暂时将多

肽附着于膜相关。在一些实施方案中,膜结构域可以从膜蛋白上裂解。在一些实施方案中,膜结构域是信号锚序列。在一些实施方案中,膜结构域是如图13所示的任何信号锚序列或其衍生的锚序列。在一些实施方案中,膜结构域是病毒信号锚序列。

[0190] 在一些实施方案中,异源多肽是天然包含膜结构域的蛋白。在一些实施方案中,异源多肽不是天然包含膜结构域而是已通过重组与膜结构域融合。在一些实施方案中,异源多肽是融合膜结构域的其他 (otherwise) 可溶的蛋白。

[0191] 在一些实施方案中,膜结构域是微藻膜结构域。在一些实施方案中,膜结构域是网粘菌门膜结构域。在一些实施方案中,膜结构域是破囊壶菌膜结构域。在一些实施方案中,膜结构域是裂殖壶菌属或破囊壶菌属膜结构域。在一些实施方案中,膜结构域包括来自裂殖壶菌属 α -1,3-甘露糖基- β -1,2-GlcNac-转移酶-I-样蛋白#1 (SEQ ID NO:78),裂殖壶菌属 β -1,2-木糖基转移酶样蛋白#1 (SEQ ID NO:80),裂殖壶菌属 β -1,4-木糖苷酶样蛋白 (SEQ ID NO:82),或裂殖壶菌属半乳糖基转移酶样蛋白#5 (SEQ ID NO:84) 的信号锚序列。

[0192] 在一些实施方案中,异源多肽是膜蛋白。本文使用的术语“膜蛋白”指与细胞膜相关或相结合的任何蛋白。例如,如Chou和Elrod, *Proteins: Structure, Function and Genetics* 34:137-153 (1999) 描述的,膜蛋白可分为几种基本类型:

[0193] 1) I型膜蛋白:这些蛋白在成熟蛋白中具有单跨膜结构域。N-末端在胞外,而C-末端在胞质内。蛋白N-末端特征性地具有典型的信号肽序列以引导蛋白进入内质网。这些蛋白细分为Ia型(含可切割信号序列)和Ib型(无可切割信号序列)。I型膜蛋白的例子包括但不限于:流感HA、胰岛素受体、血型糖蛋白、LDL受体、和病毒G蛋白。

[0194] 2) II型膜蛋白:对于这些单膜结构域蛋白,C-末端在胞外,而N-末端在胞质内。N-末端可以具有信号锚序列。这种蛋白类型的例子包括但不限于:流感神经氨酸酶、高尔基体半乳糖基转移酶、高尔基体唾液酸转移酶、蔗糖酶-异麦芽糖酶前体、去唾液酸糖蛋白受体和转铁蛋白受体。

[0195] 3) 多次跨膜蛋白 (multipass transmembrane protein):在I型和II型膜蛋白中,多肽跨越脂双层1次,而在多次跨膜蛋白中多肽穿过膜多次。多次跨膜蛋白也细分为IIIa型和IIIb型。IIIa型蛋白具有可切割的信号序列。IIIb型蛋白将其氨基末端暴露在膜的外表面,但不具有可切割的信号序列。IIIa型蛋白包括但不限于光反应中心的M和L肽。IIIb型蛋白包括但不限于大肠杆菌细胞色素P450和前导肽酶。多次跨膜蛋白的其他例子是膜转运蛋白,如糖转运蛋白(葡萄糖、木糖),和离子转运蛋白。

[0196] 4) 脂质链锚定的膜蛋白:这些蛋白借由一个或更多共价连接的脂肪酸链或其他类型的称为异戊二烯基的脂质链与膜双层相连。

[0197] 5) GPI锚定的膜蛋白:这些蛋白通过糖基磷脂酰肌醇 (GPI) 锚与膜结合。

[0198] 6) 外周膜蛋白:这些蛋白通过与其他膜蛋白的非共价相互作用间接与膜结合。

[0199] 在一些实施方案中,膜结构域是HA蛋白的膜结构域。

[0200] 在一些实施方案中,异源多肽包含来自对应于异源多肽的野生型多肽的天然信号锚序列或天然膜结构域。在一些实施方案中,异源多肽与同天然信号锚序列或天然膜结构域不同的异源信号锚序列或异源膜结构域融合。在一些实施方案中,异源多肽包含异源信号锚序列或异源膜结构域,然而对应于异源多肽的野生型多肽不包含任何信号锚序列或膜结构域。在一些实施方案中,异源多肽包含裂殖壶菌属信号锚序列。在一些实施方案中,异

源多肽包含HA膜结构域。在一些实施方案中,异源多肽是治疗性多肽。

[0201] 在一些实施方案中,膜结构域是来自图14所示的任何I型膜蛋白的膜结构域,或由其衍生的膜结构域。在一些实施方案中,本发明的异源多肽是图14所示的任何膜蛋白C-末端内跨膜区的融合多肽。在一些实施方案中,通过以来自病毒蛋白的相似区取代进一步修饰跨膜区的C-末端侧。

[0202] 在一些实施方案中,异源多肽是糖蛋白。在一些实施方案中,异源多肽具有特征为在网粘菌细胞中表达的糖基化模式。在一些实施方案中,异源多肽具有特征为在破囊壶菌细胞中表达的糖基化模式。在一些实施方案中,在微藻宿主细胞中表达的异源多肽是具有比酵母或大肠杆菌中产生的蛋白更紧密类似于哺乳动物糖基化模式的糖基化模式的糖蛋白。在一些实施方案中,糖基化模式包括N-联糖基化模式。在一些实施方案中,糖蛋白包含高甘露寡糖。在一些实施方案中,糖蛋白基本上不含唾液酸。本文使用的术语“基本上不含唾液酸”是指低于10%、低于9%、低于8%、低于7%、低于6%、低于5%、低于4%、低于3%、低于2%、或低于1%的唾液酸。在一些实施方案中,糖蛋白中不存在唾液酸。

[0203] 在一些实施方案中,包含异源多肽的本发明的微藻胞外体以商业或工业规模生产。

[0204] 本发明还涉及包含如本文所述的本发明的任何微藻胞外体和含水液体载体的组合物。

[0205] 在一些实施方案中,从生长微藻宿主细胞的培养基或发酵培养基中回收包含异源多肽的本发明的微藻胞外体。在一些实施方案中,本发明的微藻胞外体可以“基本上纯”的形式分离。如本文所使用,“基本上纯”指允许有效使用微藻胞外体作为商业或工业产品的纯度。

[0206] 本发明也涉及一种生产包含异多肽的微藻胞外体的方法,该方法包括:(a)在微藻宿主细胞中表达异源多肽,其中所述异源多肽包含膜结构域,和(b)在足以产生包含所述异源多肽的微藻胞外体的培养条件下培养宿主细胞,其中所述胞外体与宿主细胞质膜是不连续的。

[0207] 本发明也涉及一种生产包含微藻胞外体和异源多肽的组合物方法,该方法包括:(a)在微藻宿主细胞中表达异源多肽,其中所述异源多肽包含膜结构域,和(b)在足以产生包含所述异源多肽的微藻胞外体的培养条件下培养宿主细胞,其中所述胞外体与宿主细胞质膜是不连续的,其中所述组合物是作为包含所述胞外体的培养上清生成的。在一些实施方案中,该方法进一步包括去除培养上清并在含水液体载体中重悬胞外体。在一些实施方案中,组合物用作疫苗。

[0208] 包含病毒多肽的微藻胞外体

[0209] 病毒包膜蛋白是形成病毒颗粒外层的膜蛋白。这些蛋白的合成利用膜结构域,例如细胞靶向信号,以指导蛋白到质膜上。包膜外壳蛋白分为几个主要的组,其包括但不限于:H或HA(血凝素)蛋白、N或NA(神经氨酸酶)蛋白、F(融合)蛋白、G(糖蛋白)蛋白、E或env(包膜)蛋白、gp120(120kDa糖蛋白)和gp41(41kDa糖蛋白)。结构蛋白通常称作“基质”蛋白,发挥功能以帮助稳定病毒。基质蛋白包括但不限于,M1、M2(膜通道蛋白)和Gag。包膜和基质蛋白均可以参与病毒的装配和功能。例如,单独或与病毒基质蛋白结合表达病毒包膜外壳蛋白可以导致病毒样颗粒(VLPs)的形成。

[0210] 病毒疫苗通常是从与为了预防的疾病对应的病毒培养物的灭活或减毒制备物制备的,而且通常留有病毒物质,例如病毒遗传物质。一般说来,病毒是从与野生病毒可能感染的细胞类型相同或相似的细胞培养而来。这种细胞培养价格昂贵而且往往难以形成规模。为了解决此问题,某些特定的病毒蛋白抗原取而代之以通过转基因宿主表达,这可以使培养成本更低并且更适于形成规模。然而,病毒蛋白通常是存在于病毒包膜的完整的膜蛋白。由于膜蛋白的大量产生是非常困难的,通常修饰这些病毒蛋白以产生蛋白的可溶形式。这些病毒包膜蛋白是建立宿主免疫的关键,但是在异源系统中全部或部分地表达病毒包膜蛋白的大量尝试仅能获得有限的成功,可能是为了有足够的免疫原性,在病毒包膜范围内的蛋白必须呈递给免疫系统。因此,需要新的异源表达系统,例如本发明的表达系统,其能够规模化并能够呈递病毒抗原,而除了所需的病毒抗原外,不含或基本上不含有关联的病毒物质,例如病毒遗传物质。本文使用的术语“基本上不含有关联的病毒物质”指关联的病毒物质低于10%、低于9%、低于8%、低于7%、低于5%、低于4%、低于3%、低于2%、或低于1%。

[0211] 在一些实施方案中,微藻胞外体包含异源多肽,异源多肽是选自下组的病毒糖蛋白:H或HA(血凝素)蛋白、N或NA(神经氨酸酶)蛋白、F(融合)蛋白、G(糖蛋白)蛋白、E或env(包膜)蛋白、gp120(120kDa糖蛋白)和gp41(41kDa糖蛋白),及其组合。在一些实施方案中,微藻胞外体包含是病毒基质蛋白的异源多肽。在一些实施方案中,微藻胞外体包含选自下组的病毒基质蛋白:M1、M2(膜通道蛋白)、Gag、及其组合。在一些实施方案中,微藻胞外体包含选自下组的两种或多种病毒蛋白的组合:H或HA(血凝素)蛋白、N或NA(神经氨酸酶)蛋白、F(融合)蛋白、G(糖蛋白)蛋白、E或env(包膜)蛋白、gp120(120kDa糖蛋白)、gp41(41kDa糖蛋白),以及病毒基质蛋白。

[0212] 在一些实施方案中,本发明的微藻胞外体包含缺乏唾液酸的病毒糖蛋白,否则唾液酸可能会干扰蛋白的积累或功能。

[0213] 在一些实施方案中,微藻胞外体是VLP。

[0214] 本文使用的术语“VLP”是指形态与感染性病毒相似,当病毒蛋白过表达时,能够通过病毒蛋白自发的自组装形成的颗粒。VLP已经在酵母、昆虫和哺乳动物细胞中产生,并表现为亚单位疫苗的有效和更安全的类型,因为它们模拟病毒颗粒的整体结构却不含有传染性的遗传物质。这种类型的疫苗递送系统已经在刺激细胞和体液应答中取得成功。

[0215] 对副粘病毒的研究表明,当多种病毒蛋白共表达时,产生的VLP在大小和密度上与真正的病毒体是非常相似的。对于VLP的形成,仅仅表达基质蛋白(M)是必要且足够的。在副粘病毒中,仅仅表达HN导致VLP的形成效率非常低。其他蛋白单独不足以使NDV出芽。HN是II型膜糖蛋白,作为穗状四聚体存在于病毒体和受感染的细胞表面。参见,例如,Collins PL和Mottet G, *J. Virol.* 65:2362-2371 (1991); Mirza AM等, *J. Biol. Chem.* 268:21425-21431 (1993); 和Ng D等, *J. Cell. Biol.* 109:3273-3289 (1989)。HN和NP蛋白并入VLP是由与M蛋白的相互作用引起的。参见,例如,Pantua等, *J. Virology* 80:11062-11073 (2006)。

[0216] 乙型肝炎病毒(HBV)或人类乳头状瘤病毒(HPV)的VLP是简单的VLP,无包膜,并且是通过表达一种或两种衣壳蛋白产生的。更复杂的非包膜VLP包括例如为蓝舌病开发的VLP等颗粒。在此情况下,在昆虫细胞中同时表达4种来自蓝舌病毒(BTV,呼肠孤病毒科(Reoviridae))的主要结构蛋白。来自具有脂质包膜的病毒的VLP也已产生(例如丙型肝炎和甲型流感)。也有VLP样结构,例如可以重复展示抗原表位的自组装多肽纳米颗粒(SAPN)。

这些已被用来设计潜在的疟疾疫苗。参见,例如,Kaba SA等,J. Immunol. 183 (11):7268-7277 (2009)。

[0217] 因为具有与活减毒或灭活病毒相媲美的产生免疫的潜能,VLP具有显著的优势,并被认为是具有高度的免疫原性,因为它们的颗粒性质,也因为它们以密集的重叠阵列展示表面表位。例如,已经假设B细胞特异性地将具有间距50Å至100Å的表位的抗原识别为外源的。参见Bachman等,Science 262:1448 (1993)。VLP也具有被认为大大促进树突状细胞和巨噬细胞摄取的颗粒大小。此外,20nm至200nm的颗粒自由扩散到淋巴结,而500至2000nm的颗粒则不能。至少有两个已批准的用于人类的VLP疫苗,乙型肝炎疫苗 (HBV) 和人类乳头状瘤病毒 (HPV)。然而,基于病毒的VLP,例如基于杆状病毒的VLP,往往包含大量的病毒物质,其需要从VLP进一步的纯化。

[0218] 在一些实施方案中,微藻胞外体是包含病毒糖蛋白的VLP,糖蛋白选自下组:H或HA (血凝素) 蛋白、N或NA (神经氨酸酶) 蛋白、F (融合) 蛋白、G (糖蛋白) 蛋白、E或env (包膜) 蛋白、gp120 (120kDa糖蛋白) 和gp41 (41kDa糖蛋白),及其组合。在一些实施方案中,微藻胞外体是包含病毒基质蛋白的VLP。在一些实施方案中,微藻胞外体是包含选自下组的病毒基质蛋白的VLP:M1、M2 (膜通道蛋白)、Gag、及其组合。在一些实施方案中,微藻胞外体是包含选自下组的两种或多种病毒蛋白的组合物VLP:H或HA (血凝素) 蛋白、N或NA (神经氨酸酶) 蛋白、F (融合) 蛋白、G (糖蛋白) 蛋白、E或env (包膜) 蛋白、gp120 (120kDa糖蛋白)、gp41 (41kDa糖蛋白),和病毒基质蛋白。

[0219] 利用微藻胞外体的方法

[0220] 在一些实施方案中,本发明的微藻胞外体可用作蛋白活性或功能的运载体。在一些实施方案中,蛋白活性或功能与存在于胞外体内部或胞外体上的异源多肽相关。在一些实施方案中,异源多肽是膜蛋白。在一些实施方案中,蛋白活性或功能与连接于存在于胞外体中的膜蛋白的多肽相关。在一些实施方案中,蛋白质是可溶的时没有功能,但是本发明的胞外体的一部分时有功能。在一些实施方案中,通过将糖捕获在囊泡内,含糖转运蛋白 (例如木糖、蔗糖、或葡萄糖转运蛋白) 的微藻胞外体可用于消耗含糖或其他低分子量溶质、痕量的糖的混合物的培养基,然后所述囊泡可以通过包括过滤或离心的各种方法分离。

[0221] 本发明也包括本发明包含异源多肽的任何微藻胞外体及其组合物用于在动物或人类中从预防性治疗到对疾病的治疗应用。

[0222] 术语“处理”和“治疗”指治疗性治疗和预防或防止性措施,其中目标是防止或缓解 (减轻) 不希望的生理状况、疾病或失调,或是获得有益或需要的临床结果。对于本发明目的,有益或需要的临床结果包括但不限于缓解或消除与某一病症、疾病或失调相关的症状或征兆;减轻某一病症、疾病或失调的程度;稳定某一病症、疾病或失调 (即其中所述病症、疾病或失调不恶化);延迟所述病症、疾病或失调的发生或进展;改善所述病症、疾病或失调;缓和 (无论部分或全部且无论可检测或不可检测) 所述病症、疾病或失调;或者改良或者改进某一病症、疾病或失调。治疗包括引出临床上显著的反应而没有过多的副作用。治疗也包括与若不接受治疗的预期存活期相比,延长存活。

[0223] 在一些实施方案中,从培养上清中回收本发明包含异源多肽的任何微藻胞外体以直接用作动物或人类疫苗。

[0224] 在一些实施方案中,根据感兴趣的用途的要求,例如作为疫苗施用,纯化包含异源

多肽的微藻胞外体。对于典型的人类疫苗应用,低速上清通过浓缩(例如切向流过滤接超滤)、层析分离(例如阴离子交换层析)、大小排阻层析、和灭菌(例如0.2 μ m过滤)进行初始纯化。在一些实施方案中,本发明的疫苗没有潜在致敏的遗留蛋白,例如,鸡蛋中的蛋白。在一些实施方案中,与胞外体相关的病毒多肽不同,包含本发明的胞外体的疫苗没有任何病毒物质。

[0225] 根据公开的方法,施用包含异源多肽的微藻胞外体或其组合物可以,例如,通过肌肉内(i.m.)、静脉内(i.v.)、皮下(s.c.)、或肺内途径。施用的其他合适途径包括但不限于气管内、透皮、眼内、鼻内、吸入、腔内、导管内(例如到胰腺内)、和实质内(例如到任何组织)施用。透皮递送包括但不限于皮内(例如进入真皮或表皮)、透皮(例如,经皮)、和经粘膜施用(例如进入或通过皮肤或粘膜组织)。腔内施用包括但不限于施用到口腔、阴道、直肠、鼻、腹膜、和肠腔、以及鞘内(例如进入椎管)、室内(例如到脑室或心室)、房内(例如进入心房)和蛛网膜下(例如脑蛛网膜下腔)施用。

[0226] 在一些实施方案中,本发明包括包含微藻胞外体的组合物,所述胞外体包含异源多肽。在一些实施方案中,组合物包括含水液体载体。在进一步的实施方案中,含水液体载体是培养上清。在一些实施方案中,本发明的组合物包括本领域已知的常规的药学上可接受的赋形剂,例如但不限于,人血清白蛋白、离子交换剂、氧化铝、卵磷脂、缓冲物质如磷酸盐、甘氨酸、山梨酸、山梨酸钾、和盐或电解质,例如硫酸鱼精蛋白,以及例如,Remington: The Science and Practice of Pharmacy,第21版.(2005)所列的赋形剂。

[0227] 或者,本文所述的涉及微藻胞外体的任一实施方案可涉及壶菌胞外体。

[0228] 本发明的组合物最有效的施用模式和剂量方案依赖于疾病的严重程度和进程、个体的健康和对治疗的反应和主治医生的判断。因此,组合物的剂量应对个体受试者边增加边观察(titrate)。然而,本发明的组合物的有效剂量可以从1mg/kg-2000mg/kg,1mg/kg-1500mg/kg,1mg/kg-1000mg/kg,1mg/kg-500mg/kg,1mg/kg-250mg/kg,1mg/kg-100mg/kg,1mg/kg-50mg/kg,1mg/kg-25mg/kg,1mg/kg-10mg/kg,500mg/kg-2000mg/kg,500mg/kg-1500mg/kg,500mg/kg-1000mg/kg,100mg/kg-2000mg/kg,100mg/kg-1500mg/kg,100mg/kg-1000mg/kg,或100mg/kg-500mg/kg。

[0229] 已经一般描述了本发明,参考本文提供的例子可获得进一步的理解。这些例子仅出于说明目的而不是为了限制本发明。

[0230] 实施例1

[0231] pCL0143表达载体的构建

[0232] 合成了pCL0143表达载体并由DNA 2.0 (Menlo Park,CA)通过Sanger测序验证了序列。pCL0143载体包括驱动HA转基因的表达的来自裂殖壶菌属延伸因子-1基因(EF1)的启动子、接着HA转基因的OrfC终止子(又称为PFA3终止子),和赋予对抗生素巴龙霉素抗性的选择标记盒。

[0233] SEQ ID NO:76(图1)编码甲型流感病毒(A/Puerto Rico/8/34/Mount Sinai (H1N1))的HA蛋白。蛋白序列与GenBank登录号AAM75158匹配。在图16所示的裂殖壶菌属密码子选择表的指导下,对特定核酸序列SEQ ID NO:76进行密码子优化并由DNA 2.0合成以在裂殖壶菌属中表达。使用替代的信号肽也产生了一个构建体,其中去除了SEQ ID NO:76的信号肽(前51个核苷酸)并以编码裂殖壶菌属Sec1信号肽的多核苷酸序列取代(SEQ ID

NO:38)。

[0234] 实施例2

[0235] 裂殖壶菌属中产生的HA蛋白的表达和表征

[0236] 使用裂殖壶菌属菌种ATCC 20888作为以Biolistic™粒子轰击器 (particle bombardier) (BioRad,Hercules,CA) 转化pCL043载体的宿主细胞。简言之,裂殖壶菌属菌种ATCC号20888的培养物生长于M2B培养基,M2B培养基由10g/L葡萄糖,0.8g/L (NH₄)₂SO₄,5g/L Na₂SO₄,2g/L MgSO₄ · 7H₂O,0.5g/L KH₂PO₄,0.5g/L KCl,0.1g/L CaCl₂ · 2H₂O,0.1M MES (pH 6.0),0.1%PB26金属,和0.1%PB26维生素 (v/v) 组成。PB26维生素由50mg/mL维生素B12,100μg/mL硫胺素和100μg/mL泛酸钙组成。PB26金属调整到pH 4.5并由3g/L FeSO₄ · 7H₂O,1g/L MnCl₂ · 4H₂O,800mg/mL ZnSO₄ · 7H₂O,20mg/mL CoCl₂ · 6H₂O,10mg/mL Na₂MoO₄ · 2H₂O,600mg/mL CuSO₄ · 5H₂O,和800mg/mL NiSO₄ · 6H₂O组成。PB26贮液分别过滤灭菌后加入到高压灭菌后的培养液。在混合前,将葡萄糖、KH₂PO₄和CaCl₂ · 2H₂O与剩余的培养液成分分开各自高压灭菌,以防止盐析出和糖焦化。所有培养基成分均购自SigmaChemical (St.Louis,MO)。裂殖壶菌属培养物生长到对数期并以Biolistic™粒子轰击器 (BioRad,Hercules,CA) 转化。Biolistic™转化步骤与先前描述的基本相同 (参见Apt等,J.Cell.Sci.115 (Pt 21):4061-9 (1996) 和美国专利第7,001,772号)。原代转化体在27°C温育2-6天后在含20g/L琼脂 (VWR, West Chester,PA),10μg/mL甲噻磺隆 (SMM) (Chem Service,Westchester,PA) 的固体M2B培养基上选择。

[0237] 提取pCL0143原代转化体的gDNA并纯化并用作PCR模板以检测转基因的存在。

[0238] 用于裂殖壶菌属的基因组DNA提取方案-裂殖壶菌属转化体在50mL培养基中生长。将25mL培养物无菌吸到50mL锥形瓶中并在3000x g离心4分钟形成沉淀。去除上清并将沉淀储存在-80°C直到使用。在50mL锥形瓶中将沉淀重悬于约4-5倍体积的含20mM Tris pH8,10mM EDTA,50mMNaCl,0.5%SDS和100μg/ml蛋白酶K的溶液中。沉淀在50°C温和振荡温育1小时。溶解后,加入100μg/ml的核糖核酸酶A并将溶液在37°C振荡温育10分钟。接着,加入2体积的苯酚:氯仿:异戊醇并将溶液在室温下振荡1小时,然后在8000x g离心15分钟。将上清转移到一个干净的管中。再次,加入2体积的苯酚:氯仿:异戊醇并将溶液在室温下振荡1小时,然后在8000x g离心15分钟,并将上清转移到一个干净的管中。生成的上清中加入等体积的氯仿并将溶液在室温下振荡30分钟。溶液以8000x g离心15分钟并将上清转移到一个干净的管中。生成的上清中加入等体积的氯仿并将溶液在室温下振荡30分钟。溶液以8000x g离心15分钟并将上清转移到一个干净的管中。上清中加入0.3体积的3M NaOAc和2倍体积100%EtOH,轻轻摇晃几分钟。用无菌玻璃棒卷出DNA并浸入70%EtOH中1-2分钟。DNA转入1.7ml离心管并允许其风干10分钟。DNA中加入多至0.5mL预热的EB并将其在4°C放置过夜。

[0239] 转基因裂殖壶菌属 (用pCL043转化的) 的冷冻贮存物在M50-20中生长到汇合,然后在50mL带挡板的摇瓶中在培养基中在27°C,200rpm增殖48小时 (h),除非另有说明,所述培养基包含下列 (每L):

	Na ₂ SO ₄	13.62 g
[0240]	K ₂ SO ₄	0.72 g
	KCl	0.56 g

	MgSO ₄ ·7H ₂ O	2.27 g
	(NH ₄) ₂ SO ₄	3 g
	CaCl ₂ ·2H ₂ O	0.19 g
[0241]	一水 MSG	3 g
	MES	21.4 g
	KH ₂ PO ₄	0.4 g

[0242] 用去离子水将体积调整到900mL并将pH调整到6.5 (除非另有说明),然后在高温灭菌35分钟。然后加入过滤灭菌的葡萄糖 (50g/L)、维生素 (2mL/L) 和痕量金属 (2mL/L) 到培养基中并将体积调整到1L。维生素溶液含有0.16g/L维生素B12,9.75g/mL硫胺素和3.33g/L泛酸钙。痕量金属溶液 (pH2.5) 含有1.00g/L柠檬酸,5.15g/L FeSO₄·7H₂O,1.55g/L MnCl₂·4H₂O,1.55g/L ZnSO₄·7H₂O,0.02g/L CoCl₂·6H₂O,0.02g/L Na₂MoO₄·2H₂O,1.035g/L CuSO₄·5H₂O,和1.035g/L NiSO₄·6H₂O。

[0243] 将裂殖壶菌属培养物转移到50mL锥形管中并以3000×g或4500×g离心15分钟。参见图3。离心产生的上清,称为“无细胞上清”(CFS),用于免疫印迹分析和血凝素活性分析。

[0244] 无细胞上清进一步以100,000×g超速离心1小时。参见图3。生成的包含HA蛋白的沉淀(不溶级分或“UP”)重悬于PBS,pH7.4。将这种悬液在含15-60%蔗糖溶液的不连续蔗糖密度梯度上离心(120,000×g,18h,4°C)。见图3。包含HA蛋白的60%蔗糖级分用于肽序列分析、糖基化分析以及电子显微镜分析。

[0245] 免疫印迹分析

[0246] 通过免疫分析接标准免疫印迹程序验证了转基因裂殖壶菌属CL0143-9 (“E”)的重组HA蛋白的表达。除非另有说明,以MOPS SDS电泳缓冲液在还原条件下在NuPAGE® Novex®12%的bis-tris凝胶(Invitrogen,Carlsbad,CA)上通过十二烷基硫酸钠聚丙烯酰胺凝胶电泳(SDS-PAGE)分离来自无细胞上清(CFS)的蛋白。然后用考马斯亮蓝(SimplyBlue Safe Stain,Invitrogen,Carlsbad,CA)进行蛋白染色或转移到聚偏氟乙烯膜并用来自兔的抗-甲型流感/Puerto Rico/8/34 (H1N1) 病毒的抗血清(1:1000稀释,惠赠自Dr.Albert D.M.E.Osterhaus;Fouchier R.A.M.等,J.Virol.79:2814-2822(2005)),接着用偶联碱性磷酸酶的抗-兔IgG(Fc)二抗(1:2000稀释,#S3731,Promega Corporation,Madison,WI),探针检测HA蛋白的存在。然后使用制造商的方案(KPL,Gaithersburg,MD)用5-溴-4-氯-3-吡啶-磷酸盐/四氮唑蓝溶液(BCIP/NBT)处理膜。生长于不同pH(5.5,6.0,6.5和7.0)和不同温度(25°C,27°C,29°C)的转基因裂殖壶菌属CL0143-9 (“E”)的抗-H1N1免疫印迹如图4A所示。阴性对照(“C”)是裂殖壶菌属菌种ATCC 20888野生型菌株。在pH 6.5在无细胞上清中检测重组HA蛋白(图4A)并最高在pH6.5,27°C(图4A)检测血凝素活性。在非还原和还原条件下生长于pH6.5,27°C的pCL0143-9 (“E”)的考马斯亮蓝-染色凝胶(“考马斯”)和相应的抗-H1N1免疫印迹(“IB:抗-H1N1”)如图4B所示。阴性对照(“C”)是裂殖壶菌属菌种ATCC 20888野生型菌株。

[0247] HA活性

[0248] 通过血凝活性测定评价裂殖壶菌属中产生的HA蛋白的活性。功能性HA蛋白显示出

血凝活性,其易于通过标准血凝活性分析测定。简言之,在96孔微滴定板中制备了50 μ L低速上清在PBS中的2倍稀释物。然后每孔中加入等体积的在PBS,pH7.4中的约1%的鸡红细胞溶液(Fitzgerald Industries,Acton,MA),接着在室温下温育30分钟。然后视觉分析凝集程度。血凝活性单位(HAU)定义为导致孔中可见血凝的最高稀释度。

[0249] 在转基因裂殖壶菌属CL0143-9(“E”)无细胞上清的512HAU量级发现了典型的活性(图5A)。PBS(“-”)或与转基因株系以相同方式生长和制备的野生型裂殖壶菌属菌种ATCC 20888菌株(“C”)用来作为阴性对照并没有表现出任何血凝活性。来自甲型流感/Vietnam/1203/2004(H5N1)(Protein Sciences Corporation,Meriden,CT,在PBS中1:1000稀释)的重组HA蛋白用来作为阳性对照(“+”)。

[0250] 转基因裂殖壶菌属CL0143-9菌株无细胞上清的可溶和不溶级分通过血凝测定的分析表明HA蛋白主要在不溶级分中存在(图5B)。在可溶级分(“US”)的16HAU量级中和不可溶级分(“UP”)的256HAU量级中发现了典型的活性。

[0251] 2L培养物中HA蛋白的活性水平展示了与培养于恒定pH6.5的相同培养基中的摇瓶培养物相似的活性。

[0252] 在另一项独立实验中,去除HA的天然信号肽并由裂殖壶菌属Sec1信号肽(SEQ ID NO:37,由SEQ ID NO:38编码)替换。用此替代构建体获得的转基因裂殖壶菌属显示出与含有pCL0143构建体的转基因裂殖壶菌属相似的血凝素活性和重组蛋白分布(数据未显示)。

[0253] 肽序列分析

[0254] 来自无细胞上清的100,000x g离心生成的不溶级分(“UP”)进一步在蔗糖密度梯度上分级,并且通过SDS-PAGE分离通过血凝活性分析(图6B)表明的包含HA蛋白的级分并用考马斯亮蓝染色或转移到PVDF并用兔的抗-H1N1抗血清进行免疫印迹(图6A),如上所述。从考马斯亮蓝染色的凝胶切下对应免疫印迹中交叉反应的条带(HA1和HA2)并进行肽序列分析。简言之,在50%乙醇,5%醋酸中清洗/脱色感兴趣的条带。然后在乙腈中脱水凝胶块,在SpeedVac®(Thermo Fisher Scientific,Inc.,Waltham,MA)中干燥,并通过在50mM的碳酸氢铵中加入5 μ L 10ng/ μ L的胰蛋白酶用胰蛋白酶消化并于室温温育过夜。用两等份30 μ L含5%甲酸的50%乙腈从聚丙烯酰胺提取形成的肽。合并这些提取物并在SpeedVac®中蒸发到<10 μ L,然后在1%醋酸中重悬以调整到约30 μ L的终体积以供LC-MS分析。LC-MS系统是Finnigan™ LTQ™线性离子阱质谱仪(Thermo Electron Corporation,Waltham,MA)。HPLC柱是自包装的9cm $\times$ 75 μ m Phenomenex Jupiter™ C18反相毛细管层析柱(Phenomenex,Torrance,CA)。然后,注射 μ L体积的提取物并通过乙腈/0.1%甲酸梯度以0.25 μ L/分钟的流速从柱洗脱肽并引入到线上质谱仪的源中。在2.5kV操作微电喷雾离子源。使用选择性反应(SRM)实验分析消化物,其中在LC实验整个过程中质谱片段化一系列m/z比。然后将感兴趣的肽的片段化模式用于产生色谱图。确定每个肽的峰面积并以内标均一化。在此分析中使用的内标是在研究的样本之间具有不变的丰度的蛋白。通过比较每个蛋白均一化的峰比值确定两个系统之间的最终比较。然后对NCBI数据库搜索碰撞-诱导的解离谱。通过覆盖蛋白质序列超过42%的共27个肽确定了HA蛋白。测序的特定的肽在图7中以粗体突出显示。更具体地,通过总共17个肽确定HA1和通过总共9个肽确定HA2。这与在位置397之前截短的HA N-端多肽符合。在整个HA蛋白氨基酸序列中显示了对于HA1和HA2而言确定的肽的布置。HA中推定的切割位点位于氨基酸343和344之间(以R[^]G表示)。在氨基酸402开始的斜体肽序列与

HA2多肽相关但出现于在HA1中鉴定的肽中,可能是由于在HA1切下的带中痕量携带HA2肽。参见,例如,Wright等,BMC Genomics10:61(2009)的图3。

[0255] 糖基化分析

[0256] 通过酶处理评估HA蛋白上聚糖的存在。转基因裂殖壶菌属“CL0143-9”的60%蔗糖级分使用制造商的方案(New England Biolabs,Ipswich,MA)用EndoH或PNGase F消化。然后通过用MOPS SDS电泳缓冲液的NuPAGE®Novex®12%bis-tris凝胶(Invitrogen, Carlsbad,CA)上的SDS-PAGE分离蛋白,接着用考马斯亮蓝(“考马斯”)染色或通过用抗-H1N1抗血清(“IB:抗-H1N1”) (图8)的免疫印迹时,预期的流动性偏移来确定聚糖的去除。酶处理的阴性对照为未用酶温育的转基因裂殖壶菌属“CL0143-9”(“NT”=未处理)。在HA1水平的免疫印迹中能够鉴定至少5种不同的种类而在HA2水平的免疫印迹中能够鉴定2种不同的种类。这与如文献报道的HA1具有多个糖基化位点,HA2只有单一的糖基化位点相一致。

[0257] 实施例3

[0258] 裂殖壶菌属培养上清中的蛋白的表征

[0259] 如上所述在典型发酵条件下培养裂殖壶菌属菌种ATCC 20888。从培养的20h到52h以4h间隔收集培养上清样品,最后在68h收集。

[0260] 通过标准Bradford分析确定基于每份样品的培养上清中的总蛋白。见图9。

[0261] 使用图3的方法从37h、40h、44h、48h和68h的培养上清样品中分离蛋白。蛋白的SDS-PAGE胶如图10所示。泳道11以2.4μg总蛋白上样,其余泳道以5μg总蛋白上样。鉴定为肌动蛋白或凝溶胶蛋白的大量条带(通过质谱肽测序)在图10中以箭头标记。

[0262] 实施例4

[0263] 培养上清物质的负染色和电子显微镜

[0264] 在如上所述典型的摇瓶条件下培养裂殖壶菌属菌种ATCC 20888(对照)和转基因裂殖壶菌属CL0143-9(实验)。将培养物转移到50mL锥形管中并以3000×g或4500×g离心15分钟。这种无细胞上清进一步以100,000×g超速离心1小时并将获得的沉淀重悬于PBS, pH7.4。将这种悬液在不连续的15%到60%蔗糖梯度上离心(120,000×g,18h,4°C)并将60%蔗糖级分用于负染色和电子显微镜检测。

[0265] 对照物质负染色物质的电子显微镜观察包含膜片段、膜聚集体和直径从几百纳米到<50nm的囊泡的混合物(统称为“胞外体”)。见图11。囊泡的形状从圆形到拉长(管状)不等,并且囊泡的边缘是光滑或不规则的。囊泡内部出现轻微染色,表明存在有机物质。较大的囊泡具有增厚的膜,表明在制备过程中囊泡的边缘重叠。膜聚集体和碎片在形状和大小上是非常不规则的。如在通过超速离心纯化的膜中与肌动蛋白具有很强的相关性表明的,膜物质可能源自质外网。

[0266] 类似地,来自表达异源蛋白的转基因裂殖壶菌属CL0143-9培养物的无细胞上清的负染物质的电子显微镜观察表明该物质是膜片段、膜聚集体、和直径从几百纳米到<50nm的囊泡的混合物。见图11。

[0267] 使用实施例22的免疫印迹分析描述的H1N1抗血清和12nm金颗粒,如在Perkins等,J.Virol.82:7201-7211(2008)描述的也对这种物质进行了免疫定位。分离自转基因裂殖壶菌属CL0143-9的细胞外膜体是通过附接于抗血清的金颗粒高度修饰的(图12),表明在胞外体中存在抗体识别的HA蛋白。在没有膜物质的区域观察到最小的背景。很少或根本没有金

颗粒结合分离自对照物质的胞外体(图12)。

[0268] 实施例5

[0269] 木糖转运蛋白、木糖异构酶和木酮糖激酶表达载体的构建

[0270] 以HindIII消化载体pAB0018(ATCC登录号PTA-9616)、用绿豆核酸酶处理、纯化,然后再进一步以KpnI消化,产生四种不同大小的片段。通过标准电泳技术在琼脂凝胶中分离了2552bp的片段并使用商用DNA纯化试剂盒纯化。然后使用PmeI和Kpn对pAB0018进行二次消化。从该消化物中分离并纯化了6732bp的片段并与2552bp的片段连接。然后使用制造商的方案将所述连接产物用于转化商业提供的感受态DH5- α 大肠杆菌细胞菌株(Invitrogen)。将来自氨苄青霉素抗性克隆的质粒增殖、纯化,然后通过限制性消化或PCR筛选确认连接产生了预期的质粒结构。将一个经验证的质粒命名为pCL0120。参见图15。

[0271] 按照图16所示的裂殖壶菌属密码子选择表的指导,对编码中间假丝酵母(*Candida intermedia*)木糖转运蛋白GXS1(GenBank登录号AJ875406)和拟南芥木糖转运蛋白At5g17010(GenBank登录号BT015128)的序列进行密码子优化和合成(Blue Heron Biotechnology, Bothell, WA)。SEQ ID NO:94是GSX1的密码子优化的核酸序列,而SEQ ID NO:95是At5g17010的密码子优化的核酸序列。

[0272] 按照标准技术使用5'和3'限制位点BamHI和NdeI进行插入和连接将SEQ ID NO:94和SEQ ID NO:95分别克隆到pCL0120。产生的载体pCL0130和pCL0131的图谱分别如图17和图18所示。

[0273] 通过连接通过用HindIII和KpnI消化从pCL0120释放的5095bp片段与设计赋予博来霉素(zeocin)或巴龙霉素抗性的合成选择标记盒构建载体pCL0121和pCL0122。这些盒包含驱动sh ble基因(博来霉素)或npt基因(巴龙霉素)表达的 α -微管蛋白启动子。两种选择标记基因的转录物均由SV40终止子终止。载体pCL0121和pCL0122的全序列分别如SEQ ID NO:90和SEQ ID NO:91所示。载体pCL0121和pCL0122的图谱分别如图19和图20所示。

[0274] 按照图16所示的裂殖壶菌属密码子选择表的指导,对编码瘤胃真菌属菌种E2木糖异构酶(CAB76571)和瘤胃真菌属菌种E2木酮糖激酶(AJ249910)的序列进行密码子优化和合成(Blue Heron Biotechnology, Bothell, WA)。“XylA”(SEQ ID NO:92)是CAB76571密码子优化的核酸序列,而“XylB”(SEQ ID NO:93)是AJ249910密码子优化的核酸序列(图22)。

[0275] 通过插入BamHI和NdeI位点,将SEQ ID NO:92克隆到载体pCL0121产生命名为pCL0132的载体(图23)并将SEQ ID NO:21克隆到载体pCL0122产生命名为pCL0136的载体(图24)。

[0276] 实施例6

[0277] 裂殖壶菌属中产生的木糖转运蛋白、木糖异构酶和木酮糖激酶蛋白的表达和表征

[0278] 使用裂殖壶菌属菌种ATCC 20888作为分别用pCL030、pCL0131、pCL0132和pCL0136载体转化的宿主细胞。

[0279] 伴有酶预处理的电穿孔-将细胞在50mL M50-20培养基中(见美国申请公开号2008/0022422)在30℃以200rpm在摇床上培养2天。将细胞按1:100稀释到M2B培养基(见下文段落)并培养过夜(16-24h),以期达到中-对数期生长(OD_{600} 1.5-2.5)。将细胞在50mL锥形管中以3000 $\times$ g离心5分钟。移去上清并将细胞重悬于合适体积的1M甘露醇、pH5.5中,以达到20 OD_{600} 单位的终浓度。将5mL细胞分装到25mL摇瓶并以10mM $CaCl_2$ (1.0M贮液,过滤灭菌)

和0.25mg/mL蛋白酶XIV(10mg/mL贮液,过滤灭菌;Sigma-Aldrich,St.Louis,MO)修正。将摇瓶在30℃和100rpm的摇床上温育4小时。在显微镜下监测细胞以确定原生质体程度(degree of protoplasting),理想的是单个的细胞。将细胞在圆底管(即14mL Falcon™管,BD Biosciences, San Jose, CA)中以2500x g离心5分钟。除去上清液,轻轻地用5mL冰冷的10%甘油悬浮细胞。将细胞在圆底管中以约2500x g重新离心5分钟。除去上清并使用大口径枪头以500μL冰冷的10%甘油轻轻地悬浮细胞。将90μL的细胞分装入预冷的电穿孔小杯(electro-cuvette) (GenePulser®小杯-0.2cm间隙,Bio-Rad,Hercules,CA)。向小杯中添加1μg-5μg的DNA(以小于或等于10μL体积),用枪头轻轻混合,并置于冰上5分钟。将细胞在200ohms(电阻),25μF(电容)和500V下电穿孔。将0.5mL M50-20培养基立即加入小杯。然后将细胞转移到25mL摇瓶中的4.5mL M50-20培养基中并在30℃和100rpm的摇床上温育2-3小时。将细胞在圆底管中以2500x g离心5分钟。移除上清并将细胞沉淀重悬于0.5mL M50-20培养基中。将细胞铺板到适当数量(2-5个)的含有适当选择(如需要)的M2B平板上并于30℃温育。

[0280] M2B培养基由10g/L葡萄糖,0.8g/L (NH₄)₂SO₄,5g/L Na₂SO₄,2g/L MgSO₄ · 7H₂O,0.5g/L KH₂PO₄,0.5g/L KCl,0.1g/L CaCl₂ · 2H₂O,0.1M MES (pH 6.0),0.1%PB26金属,和0.1%PB26维生素(v/v)组成。PB26维生素由50mg/mL维生素B12、100μg/mL硫胺素和100μg/mL泛酸钙组成。PB26金属调整到pH 4.5并由3g/L FeSO₄ · 7H₂O,1g/L MnCl₂ · 4H₂O,800mg/mL ZnSO₄ · 7H₂O,20mg/mL CoCl₂ · 6H₂O,10mg/mL Na₂MoO₄ · 2H₂O,600mg/mL CuSO₄ · 5H₂O,和800mg/mL NiSO₄ · 6H₂O组成。PB26贮液分别过滤灭菌后加入到高压灭菌后的培养液。在混合前,将葡萄糖、KH₂PO₄和CaCl₂ · 2H₂O与剩余的培养液成分分开各自高压灭菌,以防止盐析出和糖焦化。所有培养基成分均购自Sigma Chemical (St.Louis,MO)。

[0281] 在含有合适抗生素的固体培养基上选择转化体的生长。将每种载体的20-100个原代转化体重铺到“木糖-SSFM”固体培养基上,其与SSFM(下面描述的)相同除了它包含木糖代替葡萄糖作为唯一碳源并且不加入抗生素。在这些条件下未观察到任何克隆生长。

[0282] SSFM培养基:50g/L葡萄糖,13.6g/L Na₂SO₄,0.7g/L K₂SO₄,0.36g/L KCl,2.3g/L MgSO₄ · 7H₂O,0.1M MES (pH 6.0),1.2g/L (NH₄)₂SO₄,0.13g/L谷氨酸钠,0.056g/L KH₂PO₄,和0.2g/L CaCl₂ · 2H₂O。维生素从由0.16g/L维生素B12,9.7g/L硫胺素和3.3g/L泛酸钙组成的贮液,以1mL/L加入。痕量金属从由1g/L柠檬酸,5.2g/L FeSO₄ · 7H₂O,1.5g/L MnCl₂ · 4H₂O,1.5g/L ZnSO₄ · 7H₂O,0.02g/L CoCl₂ · 6H₂O,0.02g/L Na₂MoO₄ · 2H₂O,1.0g/L CuSO₄ · 5H₂O,和1.0g/L NiSO₄ · 6H₂O组成的贮液,以2mL/L加入,调整到pH 2.5。

[0283] 提取pCL0130和pCL0131原代转化体的gDNA、纯化并用作PCR模板以检测转基因的存在。

[0284] 按照实施例2描述的方法进行基因组DNA提取。

[0285] 另外,在核糖核酸酶A温育后,按照制造商的方案使用Qiagen Genomic tip 500/G柱(Qiagen, Inc USA, Valencia, CA)进一步纯化DNA。

[0286] PCR-用于检测GXS1转基因的引物是5' CL0130 (CCTCGGGCGGCGTCCTCTT) (SEQ ID NO:96)和3' CL0130 (GGCGGCCTTCTCTGTTGC) (SEQ ID NO:97)。用于检测At5g17010转基因的引物是5' CL0131 (CTACTCCGTTGTTGCCGCCATCCT) (SEQ ID NO:98)和3' CL0131 (CCGCCGACCATACCGAGAACGA) (SEQ ID NO:99)。

[0287] 使用pCL0130、pCL0132和pCL0136一起的组合(“pCL01310”系列)或pCL0131、pCL0132和pCL0136一起的组合(“pCL0131”系列),共转化裂殖壶菌属野生型菌株(ATCC 20888)。转化子直接铺于固体木糖培养基上,3-5周后,挑取菌落并在液体木糖-SSFM中进一步增殖。在含木糖的液体培养基中的几轮连续转移提高了转化体的生长速率。pCL0130系列或pCL0131系列的共转化体也铺入含有SMM、博来霉素或巴龙霉素的固体SSFM培养基。铺入这些培养基的所有转化体对每一种测试的抗生素有抗性,表明转化体包含所有三种它们各自的载体。以木糖转运蛋白、木糖异构酶和木酮糖激酶转化的裂殖壶菌属能够在包含木糖作为唯一碳源的培养基中生长。

[0288] 在进一步的实验中,对选择的SMM-抗性转化克隆(单独的pCL0130或pCL0131转化体,或pCL0130系列共转化体,或pCL0131系列共转化体)的摇瓶培养物的无细胞提取物和无细胞上清进行Western印迹并显示两种转运蛋白均得到表达并且在两个级分中均存在,表明这些膜结合蛋白以与本文描述的其它膜蛋白观察到的相似的方式与胞外囊泡相连。此外,进行Western印迹,其显示在希望它们存在的所有克隆的无细胞提取物中表达了木糖异构酶和木酮糖激酶。通过捕捉糖到囊泡内,胞外体例如含有木糖转运蛋白的囊泡,可用于消耗包含糖或其他低分子量溶质、痕量木糖的混合物的培养基,然后所述囊泡可以通过包括过滤或离心的各种方法分离。

[0289] 实施例7

[0290] pCL0140和pCL0149表达载体的构建

[0291] 以BamHI和NdeI消化载体pCL0120产生837碱基对(bp)和8454bp的两条片段。通过标准电泳技术在琼脂凝胶中分级8454bp的片段,使用商用DNA纯化试剂盒纯化,并连接到也已经用BamHI和NdeI消化的合成的序列(SEQ ID NO:100或SEQ ID NO:101;见图26)。SEQ ID NO:100(图26)编码甲型流感病毒(A/Puerto Rico/8/34/Mount Sinai(H1N1))的NA蛋白。蛋白序列与GenBank登录号NP_040981匹配。按照图16所示的裂殖壶菌属密码子选择表的指导,对特定核酸序列SEQ ID NO:100进行密码子优化和通过DNA 2.0合成以在裂殖壶菌属中表达。SEQ ID NO:101(图26)编码与SEQ ID NO:100相同的NA蛋白,但是包括在编码区C-端的V5标签序列和多聚组氨酸序列。

[0292] 然后使用制造商的方案将连接产物用于转化商业提供的感受态DH5- α 大肠杆菌细胞菌株(Invitrogen, Carlsbad, CA)。然后通过限制消化或PCR筛选这些质粒以确认连接产生了预期的质粒结构。此方法产生的质粒载体由DNA 2.0(Menlo Park, CA)使用Sanger测序进行了验证并命名为pCL0140(图25A),其包含SEQ ID NO:100,和pCL0149(图25B),其包含SEQ ID NO:101。pCL0140和pCL0149载体包括驱动NA转基因表达的来自裂殖壶菌属延伸因子-1基因(EF1)的启动子、接着HA转基因的OrfC终止子(又称为PFA3终止子),和赋予对甲磺隆的抗性的选择标记盒。

[0293] 实施例8

[0294] 裂殖壶菌属中产生的NA蛋白的表达和表征

[0295] 如实施例2所述,使用裂殖壶菌属菌种ATCC 20888作为以BiolisticTM粒子轰击器(particle bombardier)(BioRad, Hercules, CA)转化pCL0140和pCL0149载体的宿主细胞。在含有合适抗生素的固体培养基上选择转化体生长。如之前所述(实施例2),提取原代转化体的gDNA并纯化并用作PCR模板以检测转基因的存在。

[0296] 如实施例2所述,转基因裂殖壶菌属(用pCL0140和pCL0149转化的)的冷冻贮存物在M50-20中生长到汇合,然后在50mL带挡板的摇瓶中扩增。

[0297] 将裂殖壶菌属培养物转移到50mL锥形管中并以3000×g离心15分钟。参见图27。离心产生的上清,称为“无细胞上清”(CFS)。使用Centriprep™重力浓缩器(Millipore, Billerica, MA)将CFS级分浓缩50-100倍并称为“浓缩的无细胞上清”(cCFS)。用水洗涤离心产生自离心的细胞沉淀并在重悬于2倍沉淀重量的裂解缓冲液(由50mM磷酸钠(pH7.4), 1mM EDTA, 5%甘油和1mM新鲜苯甲基磺酰氟组成)和2倍沉淀重量的0.5mm玻璃珠(Sigma, St. Louis, MO)之前冻存于液氮中。然后通过多管涡旋器(VWR, Westchester, PA)上以最高速4℃涡旋3小时裂解细胞沉淀混合物。然后将产生的细胞裂解物在4℃以5500x g离心10分钟。产生的上清保留并在4℃以5500x g重新离心10分钟。产生的上清本文定义为“无细胞提取物”(CFE)。cCFS和CFE中的蛋白浓度通过标准的Bradford测定法(Bio-Rad, Hercules, CA)确定。这些级分用于神经氨酸酶活性测定以及免疫印迹分析。

[0298] 功能性流感NA蛋白显示神经氨酸酶的活性,可以通过标准的荧光NA活性分析进行检测,该分析基于唾液酸酶水解(4-甲基伞形酮)- α -D-N-乙酰神经氨酸钠(4-MUNANA)底物(Sigma-Aldrich, St. Louis, MO)释放游离4-甲基伞形酮,其在365nm激发之后在450nm处有荧光发射。简言之,按照Potier等,Anal. Biochem. 94:287-296 (1979)描述的方法,使用25 μ L CFS和75 μ L 40 μ M 4-MUNANA或75 μ L ddH₂O作对照,分析转基因裂殖壶菌属菌株的CFS、cCFS或CFE。反应在37℃温育30分钟并以FLUOstar Omega多模式酶标仪(BMG LABTECH, Offenbourg, Germany)测量荧光。

[0299] 在来自9株用CL0140转化的转基因裂殖壶菌属菌株的浓缩的无细胞上清(cCFS)和无细胞提取物(CFE)中观察到的典型活性如图28所示。与转基因菌株以相同方式生长和制备的裂殖壶菌属菌种ATCC 20888野生型菌株(“-”)和用pCL0140转化的裂殖壶菌属的PCR阴性菌株(“27”)用作阴性对照。在浓缩的无细胞上清中发现了大部分活性,表明功能性流感神经氨酸酶成功的由裂殖壶菌属表达和分泌到外部环境。

[0300] 肽序列分析

[0301] 转基因裂殖壶菌属菌株CL0140-26用于流感NA蛋白的部分纯化以通过肽序列分析确认其成功的表达和分泌。纯化过程改编自Tarigan等, JITV14 (1):75-82 (2008),并接着如上所述测量NA活性(图29A)。简言之,转基因菌株CL0140-26的无细胞上清进一步以100,000x g 4℃离心1小时。使用Centriprep™重力浓缩器(Millipore, Billerica, MA)将产生的上清液浓缩100倍(图29A中“CCFS”级分)并用含0.1%Triton X-100的0.1M碳酸氢钠缓冲液(pH 9.1)稀释到原来的体积(图29A中“D”级分)。稀释后的样品用于通过亲和层析纯化。N-(对-氨基苯基)草氨酸琼脂糖(Sigma-Aldrich, St. Louis, MO)装填入PD-10柱(BioRad, Hercules, CA),通过用6个柱体积(CV)的含0.1%Triton X-100的0.1M碳酸氢钠缓冲液(pH9.1)洗涤后,接着用5CV含0.1%Triton X-100的0.05M醋酸钠缓冲液pH 5.5洗后活化。稀释后的样品(级分“D”)上样到柱上,通过用10CV含0.1%Triton X-100的0.15M醋酸钠缓冲液洗柱去除未结合物质(图29A中的级分“W”)。用5CV含0.1%Triton X-100和2mM CaCl₂的0.1M碳酸氢钠缓冲液从柱洗脱结合的NA(图29A中的级分“E”)。使用10-kDa-分子截留旋转浓缩器将级分E的富含NA的溶液浓缩至原体积的约10%以产生级分cE。

[0302] 以MOPS SDS电泳缓冲液在还原条件下在NuPAGE® Novex®12%的bis-tris凝

胶 (Invitrogen, Carlsbad, CA) 上的十二烷基硫酸钠聚丙烯酰胺凝胶电泳 (SDS-PAGE) 从每个级分分离蛋白。然后用考马斯亮蓝 (SimplyBlue Safe Stain, Invitrogen, Carlsbad, CA) 进行蛋白染色。从考马斯亮蓝染色凝胶中切出泳道“cE” (图29B) 中可见的蛋白条带并如实施例2所述进行肽序列分析。通过覆盖蛋白序列25%的共9条肽 (113个氨基酸) 确定了含NA蛋白的蛋白条带 (“cE”泳道中箭头所示)。测序的特定的肽在图30中以粗体突出显示。

[0303] 免疫印迹分析

[0304] 通过按照标准的免疫印迹程序的免疫印迹分析测试了来自转基因裂殖壶菌属 CL0149 (克隆10、11、12) 的重组NA蛋白的表达 (图31B)。以MOPSSDS电泳缓冲液在还原条件下在NuPAGE® Novex® 12%的bis-tris凝胶 (Invitrogen, Carlsbad, CA) 上的十二烷基硫酸钠聚丙烯酰胺凝胶电泳 (SDS-PAGE) 分离得自无细胞上清 (CFS) 的蛋白。然后用考马斯亮蓝 (SimplyBlue Safe Stain, Invitrogen, Carlsbad, CA) 进行蛋白染色或转移到聚偏氟乙烯膜并用抗-V5-AP缀合的小鼠单克隆抗体 (1:1000稀释, #962-25, Invitrogen, Carlsbad, CA) 探查NA蛋白的存在。然后根据制造商的说明 (KPL, Gaithersburg, MD) 用5-溴-4-氯-3-吡啶-磷酸盐/四氮唑蓝溶液 (BCIP/NBT) 处理膜。在克隆11的物细胞上清中检测到重组NA蛋白 (图31B)。阴性对照 (“-”) 是裂殖壶菌属菌种ATCC 20888野生型菌株。阳性对照 (“+”) 是Positope™抗体对照蛋白 (#R900-50, Invitrogen, Carlsbad, CA)。相应的神经氨酸酶活性如图31A所示。

[0305] 实施例9

[0306] 裂殖壶菌属中流感HA和NA的同时表达

[0307] 如实施例2所述, 使用裂殖壶菌属菌种ATCC 20888作为宿主细胞以Biolistic™粒子轰击器 (particle bombardier) (BioRad, Hercules, CA) 同时用pCL0140 (图25A) 和pCL0143 (图2) 载体转化。

[0308] 如实施例2所述培养和处理转基因裂殖壶菌属 (用pCL0140和pCL0143转化的) 的冷冻贮存物。分别如实施例2和7所述测量血凝和神经氨酸酶活性, 并示于图32。用pCL0140和pCL0143转化的转基因裂殖壶菌属证明了与HA和NA相关的活性。

[0309] 实施例10

[0310] 裂殖壶菌属中产生的包含副流感F蛋白的胞外体的表达和表征

[0311] 使用裂殖壶菌属菌种ATCC 20888作为宿主细胞用包含编码人副流感3病毒NIH 47885株 (GenBank登录号P06828) 的F蛋白的序列的载体转化。副F蛋白的代表序列如SEQ ID NO:102所提供。以包含编码F蛋白相关的天然信号肽序列的序列的载体转化某些细胞。以包含编码不同的信号肽序列 (例如裂殖壶菌属信号锚序列) 的载体转化其他细胞, 信号肽序列融合于编码F蛋白的序列, 以使得F蛋白与异源信号肽序列一起表达。以包含编码不同膜结构域的序列 (例如HA膜结构域) 的载体转化其他细胞, 膜结构域序列融合于编码F蛋白的序列, 以使得F蛋白与异源膜结构域一起表达。F蛋白在C-末端附近包括单次跨膜结构域。F蛋白可以在Furin切割位点 (氨基酸109) 处分为两个肽。称为F2的蛋白第一部分包含完整F蛋白的N-端部分。F2区可以单独融合于编码异源信号肽的序列。病毒F蛋白含有F蛋白C-端部分的其余部分被称为F1。F1区可以单独融合于编码异源信号肽的序列。包含病毒F蛋白的F1和F2部分的载体可以单独表达或联合表达。表达完整F蛋白的载体可以与在furin切割位点切割蛋白的furin酶共表达。另外, 可以用编码由不同蛋白酶识别或切割的其他蛋白酶切割

位点的序列替换编码F蛋白furin切割位点的序列。包含替代的蛋白酶切割位点的F蛋白可以与相应的识别和切割替代蛋白酶切割位点的蛋白酶共表达。

[0312] 进行转化,并按照本文所述的任何方法培养和增殖冷冻贮存物。将裂殖壶菌属培养物转移到50mL锥形管中并以 $3000\times g$ 或 $4500\times g$ 离心15分钟以产生低速上清。进一步将低速上清以 $100,000\times g$ 超速离心1h。参见图3。生成的包含F蛋白的不溶级分的沉淀重悬于磷酸盐缓冲盐水(PBS)并如实施例2所述用于肽序列分析和糖基化分析。

[0313] 如实施例2所述根据标准免疫印迹程序通过免疫分析,使用抗-F抗血清和适当稀释的二抗验证了F蛋白自转基因裂殖壶菌属的表达。在低速上清和不溶级分检测到了重组F蛋白。此外,来自表达F蛋白的转基因裂殖壶菌属的无细胞提取物中检测到了重组F蛋白。

[0314] 通过F活性分析评估裂殖壶菌属中产生的F蛋白的活性。通过标准F活性分析,功能性F蛋白显示出易于检测的F活性。

[0315] 使用根据实施例4产生的负染色物质进行电子显微镜观察以确认胞外体的存在。进行免疫金标记以确认蛋白与胞外体的关联。

[0316] 实施例11

[0317] 裂殖壶菌属中产生的包含G水泡性口炎病毒G蛋白的胞外体的表达和表征

[0318] 使用裂殖壶菌属菌种ATCC 20888作为宿主细胞用包含水泡性口炎病毒G(VSV-G)蛋白编码序列的载体转化。VSV-G蛋白的代表序列如SEQ IDNO:103(来自GenBank登录号M35214)所提供。以包含编码与VSV-G蛋白相关的天然信号肽序列的载体转化某些细胞。以包含不同的信号肽序列(例如裂殖壶菌属信号锚序列)的载体转化其他细胞,信号肽序列融合于VSV-G蛋白的编码序列,以使得VSV-G蛋白与异源信号肽序列一起表达。以包含编码不同膜结构域(例如HA膜结构域)的序列的载体转化其他细胞,膜结构域序列融合于VSV-G蛋白的编码序列,以使得VSV-G蛋白与异源膜结构域一起表达。进行转化,并且按照本文所述的任何方法培养和增殖冷冻贮存物。将裂殖壶菌属培养物转移到50mL锥形管中并以 $3000\times g$ 或 $4500\times g$ 离心15分钟以产生低速上清。进一步将低速上清以 $100,000\times g$ 超速离心1h。参见图3。生成的包含VSV-G蛋白的不溶级分的沉淀重悬于磷酸盐缓冲盐水(PBS)并如实施例2所述用于肽序列分析和糖基化分析。

[0319] 如实施例2所述根据标准免疫印迹程序通过免疫分析,使用抗-VSV-G抗血清和适当稀释的二抗验证了VSV-G蛋白自转基因裂殖壶菌属的表达。在低速上清和不溶级分检测到重组VSV-G蛋白。此外,来自表达VSV-G蛋白的转基因裂殖壶菌属的无细胞提取物中检测到重组VSV-G蛋白。

[0320] 通过VSV-G活性分析评估裂殖壶菌属中产生的VSV-G蛋白的活性。通过标准VSV-G活性分析,功能性VSV-G蛋白显示出易于检测的VSV-G活性。

[0321] 使用根据实施例4产生的负染色物质进行电子显微镜观察以确认胞外体的存在。进行免疫金标记以确认蛋白与胞外体的关联。

[0322] 实施例12

[0323] 裂殖壶菌属中产生的包含eGFP融合蛋白的胞外体的表达和表征

[0324] 已经描述了使用包含编码eGFP的多核苷酸序列的载体转化裂殖壶菌属菌种ATCC 20888和在转化的裂殖壶菌属中eGFP的表达。参见美国公开第2010/0233760号和WO 2010/107709,以整体引用的方式并入本文作为参考。

[0325] 在进一步的实验中,使用裂殖壶菌属菌种ATCC 20888作为宿主细胞用包含eGFP和膜结构域,例如裂殖壶菌属膜结构域或如HA膜结构域的病毒膜结构域之间的融合蛋白编码序列的载体转化。代表性裂殖壶菌属膜结构域如图13和图14所提供。进行转化,并且按照本文所述的任何方法培养并增殖冷冻贮存物。将裂殖壶菌属培养物转移到50mL锥形管中并以 $3000\times g$ 或 $4500\times g$ 离心15分钟以产生低速上清。进一步将低速上清以 $100,000\times g$ 超速离心1h。参见图3。生成的包含来自转基因裂殖壶菌属的eGFP融合蛋白的不溶级分的沉淀重悬于磷酸盐缓冲盐水(PBS)并如实施例2所述用于肽序列分析和糖基化分析。

[0326] 如实施例2所述根据标准免疫印迹程序通过免疫分析,使用抗-eGFP融合蛋白抗血清和适当稀释的二抗验证了转基因裂殖壶菌属的eGFP融合蛋白的表达。在低速上清和不溶级分中检测到重组eGFP融合蛋白。此外,在来自表达eGFP融合蛋白的转基因裂殖壶菌属的无细胞提取物中检测到重组eGFP融合蛋白。

[0327] 通过eGFP融合蛋白活性分析评估裂殖壶菌属中产生的eGFP融合蛋白的活性。通过标准eGFP融合蛋白活性分析,功能性eGFP融合蛋白显示出易于检测的eGFP融合蛋白活性。

[0328] 使用根据实施例4产生的负染色物质进行电子显微镜观察以确认胞外体的存在。进行免疫金标记以确认蛋白与胞外体的关联。

[0329] 实施例13

[0330] 破囊壶菌属培养物中产生的异源多肽的检测

[0331] 在适当发酵条件下,在发酵罐中制备包含至少一种异源多肽的破囊壶菌宿主细胞的培养物。发酵罐以含有例如,碳(葡萄糖)、氮、磷、盐、痕量金属和维生素的培养基成批生产(batch with a media)。发酵罐接种典型种子培养物,然后培养72-120小时,并补料碳(例如葡萄糖)补料。碳补料在整个发酵过程中被补入和消耗。72-120小时后,收获发酵罐并离心培养液以从上清中分离生物质。

[0332] 通过标准分析,例如Bradford或BCA测定生物质和无细胞上清的蛋白含量。通过标准SDS-PAGE和Western印迹进一步分析蛋白以确定在代表性生物质和无细胞上清级分中异源多肽的表达。通过常规染色程序(例如负染色和免疫金标记)和随后的电子显微镜观察来显示包含膜结构域的异源多肽与微藻胞外体的关联。

[0333] 实施例14

[0334] 从微藻培养物制备病毒样颗粒

[0335] 在上述条件下在微藻宿主细胞中异源表达一种或多种病毒包膜多肽,使得病毒多肽定位于培养条件下产生的微藻胞外体。当使用适当的培养条件和调节性调控元件过表达时,在微藻胞外体中的病毒包膜多肽自发自组装成与传染性病毒形态相似的颗粒。

[0336] 同样,在上述条件下在微藻宿主细胞中异源表达一种或多种病毒包膜多肽和一种或多种病毒基质多肽,使得病毒多肽定位于培养条件下产生的微藻胞外体。当使用适当的培养条件和调节性调控元件过表达时,在微藻胞外体中的病毒多肽自发自组装成与传染性病毒形态相似的颗粒。

[0337] 本文所述的各个方面、实施例和选项均可以与任何和所有的变化相结合。

[0338] 只要个人出版物、专利或专利申请是专门单独表示将纳入本文作为参考的,在本文中提到的所有出版物、专利和专利申请,以相同的程度纳入本文作为参考。

序列表

<110> DSM IP 资产公司

<120> 微藻中异源多肽的生产，微藻胞外体、组合物及其制备方法和用途

<130> 2715.238PC01/JUK/SAS/BNC

<140> 待指定

<141> 同此

<150> 61/413,353

<151> 2010-11-12

<150> 61/290,469

<151> 2009-12-28

<150> 61/290,441

<151> 2009-12-28

<160> 103

<170> PatentIn version 3.5

<210> 1

<211> 35

<212> PRT

<213> 裂殖壶菌属

<400> 1

Met Ala Asn Ile Met Ala Asn Val Thr Pro Gln Gly Val Ala Lys Gly
1 5 10 15

Phe Gly Leu Phe Val Gly Val Leu Phe Phe Leu Tyr Trp Phe Leu Val
20 25 30

Gly Leu Ala
35

<210> 2

<211> 105

<212> DNA

<213> 裂殖壶菌属

<400> 2

atggccaaca tcatggccaa cgtcacgcc cagggcgctcg ccaaggcctt tggcctcttt 60

gtcgcgctgc tcttctttct ctactgggtc cttgtcgccc tggccc 105

<210> 3

<211> 1999

<212> DNA

<213> 裂殖壶菌属

<400> 3

ccgcgaatca agaagtagg cgcgcctgca ggcgcggcgg cggagcggag cgaggagag 60

ggagaggag agagaggag agagacgtcg ccgcggcggg gcctggcctg gcctggtttg 120

gcttggtcag ccgcggccttg tccgagcgtg cagctggagt tgggtggatt catttggtt 180

tctttttgtt ttgtttttt tctctttccc ggaaagtgtt ggcgggtcgg tgtttttgt 240

tttgatttct tcaaaagttt tggtagtttg ttctctctct tggctctctg ttaggcggtc 300

cgtccacgc ccgggcctct cctctccctc cctctccctc cctctccgtg cgtatacgtg 360

cgtacgtttg tatacgtaca tacatccgc ccgcgcgtgc ggcgagggtt tgcacgcct 420

[0001]

	ggagcaatgc gatgcgatgc gatgcgatgc gacgcgacgc gacgcgagtc actggttgcg	480
	gctgtggcctg tggatttgctt gcttaacttgc ttctgagctc tcccgcttctc ttctttactt	540
	ctcacgccac caccacagaa agaagatcgg ccccggcacg ccgctgagaa gggctggcgg	600
	cgatgaaggg acgcgcgcgc getgcacgt tggcgctgcg tgcgtgctg getgctgctg	660
	ctgctgctgc tgcgtgctg getgctgctt ctgcgcgcag gctttgccac gaggcggcg	720
	tgttgccgcg tgcgccttc agtccgcgtg gagagatcga atgagagata aactggatgg	780
	attcctgagc ggatgaatga acgatggctg gatgcctttt tcttttttca ggtecaacgc	840
	gggaagcagg agcgcgtgaa tctgcgcga tccgcatacg tctgcctgc atgcctgcg	900
	atgcacgcat cgtcgcgcg gagccacaga cggcgacag ggcggccagc cagccaggca	960
	gccagccagg caggcaccag agggccagag agcgcgcctc acgcacgcgc cgcagtgcg	1020
	gcctgcctgc cagtgcagac ctgattccc cgcgcggatc tccgcgagcc cgaacagaa	1080
	agcgcctgac gggccctacc tagcgtgcgc tgcacgcga tgcctgcga tgcgttccc	1140
	tagagagtag tactgcagca aggcacctt tccgcctgc tcttcggcgc gatcgaggcc	1200
	cccgccgcgc ccacgctgc ggcgcgcgc gccttgccgc cggccctggc gctcgcctg	1260
	gcgcgcgcgc cggcgctgc gccttgccgc gcgcgcgcgc cgcaggagga gcggcagcgc	1320
	ctgctgcgcg ccagagaagg agcgcgcgc gcgcgcgcgc ggcaggagga gagaaggaga	1380
	agcgcgcga ggcgcgcgc aaagagaaga ccttgacat gaacgcgaag aagaagaaga	1440
	aggagaagaa gtgaagaag aagaagaaga aggagaggaa gtgaagaag acgaggagca	1500
	ggcgccttc aaggcgcgtt ctcttcgcga ggcgccttc agctgcgcgc gcggggcggg	1560
[0002]	ctgcgggcgc ggcgcgcgc cggctgcgcg cagaggggac gcgcgcgcgc aggcggaggg	1620
	ggcgcgcgcg ggcgcctgc tgcctgcgcg cgcgcgcgc gcagctgtgc cgcgcgcgc	1680
	gacggagggc acgacgcag cgcgcgcgc gacaaaggcc gcgcgcgcgc cgcgcggaag	1740
	gccccgcgcg gacaggggc ggcgcgcgc caagcgcgc gacaggggc aggcgcgcgc	1800
	gcgggagcgc gacgggcgc cgcgcgcgc ggcagaacc aatcgagatc cagagcgcgc	1860
	agagggcgcgc cgcgcgcgc agccgcgcgc gcagatcact agtaccctg cggatcaca	1920
	gcagcgcgcg cgcgcgcgc agcgcgcgc gcagcgcgc cagccacgc agggagataa	1980
	agaaaaagc gcagagacg	1999
	<210> 4	
	<211> 325	
	<212> DNA	
	<213> 裂殖壶菌属	
	<400> 4	
	gacccgaaag tgaaccttgc cctaaccgc cagcgaatgg cgggaggggg cgggctaaaa	60
	gacgttatta catagtattt ttccctact ctttgtgttt gtcttttttt tttttgaa	120
	gcattcaagc caattgtctt ggtttacttg ttgtttgct tgcctgctg cttgcttgc	180
	tgtttcttgg tcagacgcgc caaaaaggg aaaaattca ttcattggc acataagaaa	240
	aagaaaaagt ttgtgacca cgtcctcag aaagcaagag aagagaaaca ctgcgcctc	300
	cattctcgt cgcgttaagaa tctta	325
	<210> 5	
	<211> 372	

<212> DNA

<213> Streptoalloteichus hindustanus

<400> 5

atggccaagt tgaccagtgc cgttccgggtg ctcaaccgcg ggcacgtcgc eggagcggtc	60
gagttctcga ccgaccgggt egggttctcc cgggacttgc tggaggacga ctccgcgggt	120
gtggtccggg acgacgtgac cctgttcac agcccggtcc aggaccaggt ggtgccggac	180
aacaccctgg cctgggtgtg ggtgcgcgcg ctggacgagc tgtacgccga gtggtcggag	240
gtcgtgtcca cgaacttccg ggaagcctcc gggcgggcca tgaccgagat cggcgagcag	300
ccgtgggggc gggagtgcgc cctgcgcgac ccggcgggca actgcgtgca ctctgtggcc	360
gaggagcagg ac	372

<210> 6

<211> 2055

<212> DNA

<213> 裂殖壶菌属

<400> 6

[0003]

atgagcgca ccgcgcggc gacgaggaca gggcgggcgc tgtctcggc gcgacgacg	60
cctgtaaacg agcagcagca gcagcagctg cgcgtaggcg eggcgtcggc acggtcggcg	120
gccgcggcgt tctcgtccgg cagggcgga gacgcggcca agaaggcggc cgcgcgagg	180
gcgttctcca cgggacggcg ccccaacgcg acaacgaga agagctcgct ggccacggtc	240
caggcgcgca cggacgatgc gcgttcgtc ggcccgacc gcgccaaat ctctcatgag	300
ctcatgcgcg agcaccaggt ggacaccatc ttggctaac ctggcggcgc cttctgcgc	360
gtttttgatg ccatttttga gattgacgcc ttcaagtcca ttctcgctcg ccacgagcag	420
ggcgccggcc acatggccga ggcctacgcg cgcgccacgg gcaagcccg cgttgctctc	480
gtcaccctgg gccctggagc caccacacac atcacccgca tcatggatgc ttacatggac	540
ggtacgcgcg tctctggtgt caacggccag gtgcccaact ctgtgtctgg caaggacgt	600
ttccaggagt gtgacattgt tggcaccagc cgcgcgtgca ccaagtggaa cgtcatggtc	660
aaggacgtga aggagctccc gccccgcatc aatgagccct ttgagattgc catgagcggc	720
cgcgcgggtc ccgtgctcgt cgacltccct aaggatgtga ccgcgttga gctcaaggaa	780
atgcccgaca gctcccccca gttgtctgtg cgcagaaagc aaaaggctga gcttttcac	840
aaggagcgca ttggcgctcc tggcacggcc gacttcaage tcattgcga gatgacaa	900
cgtgcggagc gaccgctcat ctatgtctgc caggggtgca tgcagagccc gttgaatgga	960
ccgctgtgca tcaaggagtt cgcggagaag gccaacatc cgtgaccac caccatgcag	1020
ggctctggcg gctttgacga gcgtagtcgc ctctccctca agatgctcgg catgacgggc	1080
ctgcctacg ccaactactc gatgcagaac gccgatctta tcttggegt cgggtcccgc	1140
tttgatgac gtgtgacggg ccgcgttgc gcccttgctc cggaggtctg ccgtgccgag	1200
cgcgagggac ggggtggcat cgttacctt gagatttccc ccaagaaact ccacaaggtc	1260
gtcagccca ccgtcgcggt cctcgcgac gtggtcgaga acctcgccaa cgtcacgccc	1320
cacgtgcagc gccaggagcg ccagccgtgg ttgtgcaga tgcgcattg gaaggagaag	1380
cacccttttc tgcctgagtc tgttgattcg gacgacaagg ttctcaagcc gcagcaggtc	1440
ctcacggagc ttaacaagca gattctcgag attcaggaga aggacgcga ccaggaggtc	1500
tacatccca cggcgctcgg aagccaccag atgcaggcag cgcagttct taccctggac	1560

```

aagcgcgcgc agtggatctc ctccgggtgc gcgcgcacta tgggctaagg ccttcactcg 1620
gccattggcg ccaagattgc caagccgat gctattgtta ttgacatga tggatgatct 1680
tcttattcga tgaccggat ggaattgate acagcagccg aattcaaggt tggcgtgaag 1740
attcttcttt tgcagaacaa ctttcagggc atggtcaaga actggcagga tctcttttac 1800
gacaagcgct acfcgggcac cgcctatgtr aaccgcgcct tcgacaaggt cgcgatgcg 1860
atgcgtgcc aagggtctta ctgcgcgaaa cagtcggagc tcaaggacaa gatcaaggag 1920
ttctctgagt acgatgaggg tcccgctctc ctgcagggtt tctgtgacaa ggacacgctc 1980
gtcttgccca tggtcctcgc tggctttcgc ctccacgaga tggctctcga gcctcttaag 2040
cccaaggacg cctaa 2055

```

<210> 7
 <211> 684
 <212> PPT
 <213> 裂殖壶菌属

<400> 7

```

Met Ser Ala Thr Arg Ala Ala Thr Arg Thr Ala Ala Ala Leu Ser Ser
1          5          10          15

```

```

Ala Leu Thr Thr Pro Val Lys Gln Gln Gln Gln Gln Leu Arg Val
20          25          30

```

```

Gly Ala Ala Ser Ala Arg Leu Ala Ala Ala Ala Phe Ser Ser Gly Thr
35          40          45

```

[0004]

```

Gly Gly Asp Ala Ala Lys Lys Ala Ala Ala Ala Arg Ala Phe Ser Thr
50          55          60

```

```

Gly Arg Gly Pro Asn Ala Thr Arg Glu Lys Ser Ser Leu Ala Thr Val
65          70          75          80

```

```

Gln Ala Ala Thr Asp Asp Ala Arg Phe Val Gly Leu Thr Gly Ala Gln
85          90          95

```

```

Ile Phe His Glu Leu Met Arg Glu His Gln Val Asp Thr Ile Phe Gly
100         105         110

```

```

Tyr Pro Gly Gly Ala Ile Leu Pro Val Phe Asp Ala Ile Phe Glu Ser
115         120         125

```

```

Asp Ala Phe Lys Phe Ile Leu Ala Arg His Glu Gln Gly Ala Gly His
130         135         140

```

```

Met Ala Glu Gly Tyr Ala Arg Ala Thr Gly Lys Pro Gly Val Val Leu
145         150         155         160

```

```

Val Thr Ser Gly Pro Gly Ala Thr Asn Thr Ile Thr Pro Ile Met Asp
165         170         175

```

```

Ala Tyr Met Asp Gly Thr Pro Leu Leu Val Phe Thr Gly Gln Val Pro
180         185         190

```

```

Thr Ser Ala Val Gly Thr Asp Ala Phe Gln Gln Cys Asp Ile Val Gly
195         200         205

```

Ile	Ser	Arg	Ala	Cys	Thr	Lys	Trp	Asn	Val	Met	Val	Lys	Asp	Val	Lys	210	215	220
Glu	Leu	Pro	Arg	Arg	Ile	Asn	Glu	Ala	Phe	Glu	Ile	Ala	Met	Ser	Gly	225	230	235
Arg	Pro	Gly	Pro	Val	Leu	Val	Asp	Leu	Pro	Lys	Asp	Val	Thr	Ala	Val			
				245					250					255				
Glu	Leu	Lys	Glu	Met	Pro	Asp	Ser	Ser	Pro	Gln	Val	Ala	Val	Arg	Gln			
			260					265					270					
Lys	Gln	Lys	Val	Glu	Leu	Phe	His	Lys	Glu	Arg	Ile	Gly	Ala	Pro	Gly			
		275					280					285						
Thr	Ala	Asp	Phe	Lys	Leu	Ile	Ala	Glu	Met	Ile	Asn	Arg	Ala	Glu	Arg			
	290					295					300							
Pro	Val	Ile	Tyr	Ala	Gly	Gln	Gly	Val	Met	Gln	Ser	Pro	Leu	Asn	Gly			
305					310					315					320			
Pro	Ala	Val	Leu	Lys	Glu	Phe	Ala	Glu	Lys	Ala	Asn	Ile	Pro	Val	Thr			
				325					330					335				
Thr	Thr	Met	Gln	Gly	Leu	Gly	Gly	Phe	Asp	Glu	Arg	Ser	Pro	Leu	Ser			
			340					345						350				
Leu	Lys	Met	Leu	Gly	Met	His	Gly	Ser	Ala	Tyr	Ala	Asn	Tyr	Ser	Met			
		355					360					365						
Gln	Asn	Ala	Asp	Leu	Ile	Leu	Ala	Leu	Gly	Ala	Arg	Phe	Asp	Asp	Arg			
		370				375					380							
Val	Thr	Gly	Arg	Val	Asp	Ala	Phe	Ala	Pro	Glu	Ala	Arg	Arg	Ala	Glu			
385					390					395					400			
Arg	Glu	Gly	Arg	Gly	Gly	Ile	Val	His	Phe	Glu	Ile	Ser	Pro	Lys	Asn			
				405					410					415				
Leu	His	Lys	Val	Val	Gln	Pro	Thr	Val	Ala	Val	Leu	Gly	Asp	Val	Val			
			420					425					430					
Glu	Asn	Leu	Ala	Asn	Val	Thr	Pro	His	Val	Gln	Arg	Gln	Glu	Arg	Glu			
		435					440					445						
Pro	Trp	Phe	Ala	Gln	Ile	Ala	Asp	Trp	Lys	Glu	Lys	His	Pro	Phe	Leu			
		450				455					460							
Leu	Glu	Ser	Val	Asp	Ser	Asp	Asp	Lys	Val	Leu	Lys	Pro	Gln	Gln	Val			
465					470					475					480			
Leu	Thr	Glu	Leu	Asn	Lys	Gln	Ile	Leu	Glu	Ile	Gln	Glu	Lys	Asp	Ala			
				485					490					495				
Asp	Gln	Glu	Val	Tyr	Ile	Thr	Thr	Gly	Val	Gly	Ser	His	Gln	Met	Gln			
			500					505					510					

Ala Ala Gln Phe Leu Thr Trp Thr Lys Pro Arg Gln Trp Ile Ser Ser
 515 520 525
 Gly Gly Ala Gly Thr Met Gly Tyr Gly Leu Pro Ser Ala Ile Gly Ala
 530 535 540
 Lys Ile Ala Lys Pro Asp Ala Ile Val Ile Asp Ile Asp Gly Asp Ala
 545 550 555 560
 Ser Tyr Ser Met Thr Gly Met Glu Leu Ile Thr Ala Ala Glu Phe Lys
 565 570 575
 Val Gly Val Lys Ile Leu Leu Leu Gln Asn Asn Phe Gln Gly Met Val
 580 585 590
 Lys Asn Trp Gln Asp Leu Phe Tyr Asp Lys Arg Tyr Ser Gly Thr Ala
 595 600 605
 Met Phe Asn Pro Arg Phe Asp Lys Val Ala Asp Ala Met Arg Ala Lys
 610 615 620
 Gly Leu Tyr Cys Ala Lys Gln Ser Glu Leu Lys Asp Lys Ile Lys Glu
 625 630 635 640
 Phe Leu Glu Tyr Asp Glu Gly Pro Val Leu Leu Glu Val Phe Val Asp
 645 650 655
 Lys Asp Thr Leu Val Leu Pro Met Val Pro Ala Gly Phe Pro Leu His
 660 665 670
 Glu Met Val Leu Glu Pro Pro Lys Pro Lys Asp Ala
 675 680
 <210> 8
 <211> 684
 <212> PRT
 <213> 人工序列
 <220>
 <223> 突变的 ALS 1
 <400> 8
 Met Ser Ala Thr Arg Ala Ala Thr Arg Thr Ala Ala Ala Leu Ser Ser
 1 5 10 15
 Ala Leu Thr Thr Pro Val Lys Gln Gln Gln Gln Gln Leu Arg Val
 20 25 30
 Gly Ala Ala Ser Ala Arg Leu Ala Ala Ala Ala Phe Ser Ser Gly Thr
 35 40 45
 Gly Gly Asp Ala Ala Lys Lys Ala Ala Ala Ala Arg Ala Phe Ser Thr
 50 55 60
 Gly Arg Gly Pro Asn Ala Thr Arg Glu Lys Ser Ser Leu Ala Thr Val
 65 70 75 80
 Gln Ala Ala Thr Asp Asp Ala Arg Phe Val Gly Leu Thr Gly Ala Gln
 85 90 95

[0006]

	Ile Phe His Glu Leu Met Arg Glu His Gln Val Asp Thr Ile Phe Gly	100	105	110
	Tyr Pro Gly Gly Ala Ile Leu Pro Val Phe Asp Ala Ile Phe Glu Ser	115	120	125
	Asp Ala Phe Lys Phe Ile Leu Ala Arg His Glu Gln Gly Ala Gly His	130	135	140
	Met Ala Glu Gly Tyr Ala Arg Ala Thr Gly Lys Pro Gly Val Val Leu	145	150	155
	Val Thr Ser Gly Pro Gly Ala Thr Asn Thr Ile Thr Pro Ile Met Asp	165	170	175
	Ala Tyr Met Asp Gly Thr Pro Leu Leu Val Phe Thr Gly Gln Val Pro	180	185	190
	Thr Ser Ala Val Gly Thr Asp Ala Phe Gln Glu Cys Asp Ile Val Gly	195	200	205
	Ile Ser Arg Ala Cys Thr Lys Trp Asn Val Met Val Lys Asp Val Lys	210	215	220
	Glu Leu Pro Arg Arg Ile Asn Glu Ala Phe Glu Ile Ala Met Ser Gly	225	230	235
[0007]	Arg Pro Gly Pro Val Leu Val Asp Leu Pro Lys Asp Val Thr Ala Val	245	250	255
	Glu Leu Lys Glu Met Pro Asp Ser Ser Pro Gln Val Ala Val Arg Gln	260	265	270
	Lys Gln Lys Val Glu Leu Phe His Lys Glu Arg Ile Gly Ala Pro Gly	275	280	285
	Thr Ala Asp Phe Lys Leu Ile Ala Glu Met Ile Asn Arg Ala Glu Arg	290	295	300
	Pro Val Ile Tyr Ala Gly Gln Gly Val Met Gln Ser Pro Leu Asn Gly	305	310	315
	Pro Ala Val Leu Lys Glu Phe Ala Glu Lys Ala Asn Ile Pro Val Thr	325	330	335
	Thr Thr Met Gln Gly Leu Gly Gly Phe Asp Glu Arg Ser Pro Leu Ser	340	345	350
	Leu Lys Met Leu Gly Met His Gly Ser Ala Tyr Ala Asn Tyr Ser Met	355	360	365
	Gln Asn Ala Asp Leu Ile Leu Ala Leu Gly Ala Arg Phe Asp Asp Arg	370	375	380
	Val Thr Gly Arg Val Asp Ala Phe Ala Pro Glu Ala Arg Arg Ala Glu	385	390	395
				400

[0008]

```

Arg Glu Gly Arg Gly Ile Val His Phe Glu Ile Ser Pro Lys Asn
    405                      410                      415

Leu His Lys Val Val Gln Pro Thr Val Ala Val Leu Gly Asp Val Val
    420                      425                      430

Glu Asn Leu Ala Asn Val Thr Pro His Val Gln Arg Gln Glu Arg Glu
    435                      440                      445

Pro Trp Phe Ala Gln Ile Ala Asp Trp Lys Glu Lys His Pro Phe Leu
    450                      455                      460

Leu Glu Ser Val Asp Ser Asp Asp Lys Val Leu Lys Pro Gln Gln Val
    465                      470                      475                      480

Leu Thr Glu Leu Asn Lys Gln Ile Leu Glu Ile Gln Glu Lys Asp Ala
    485                      490                      495

Asp Gln Glu Val Tyr Ile Thr Thr Gly Val Gly Ser His Gln Met Gln
    500                      505                      510

Ala Ala Gln Phe Leu Thr Trp Thr Lys Pro Arg Gln Trp Ile Ser Ser
    515                      520                      525

Gly Gly Ala Gly Thr Met Gly Tyr Gly Leu Pro Ser Ala Ile Gly Ala
    530                      535                      540

Lys Ile Ala Lys Pro Asp Ala Ile Val Ile Asp Ile Asp Gly Asp Ala
    545                      550                      555                      560

Ser Tyr Ser Met Thr Gly Met Glu Leu Ile Thr Ala Ala Glu Phe Lys
    565                      570                      575

Val Gly Val Lys Ile Leu Leu Leu Gln Asn Asn Phe Gln Gly Met Val
    580                      585                      590

Lys Asn Val Gln Asp Leu Phe Tyr Asp Lys Arg Tyr Ser Gly Thr Ala
    595                      600                      605

Met Phe Asn Pro Arg Phe Asp Lys Val Ala Asp Ala Met Arg Ala Lys
    610                      615                      620

Gly Leu Tyr Cys Ala Lys Gln Ser Glu Leu Lys Asp Lys Ile Lys Glu
    625                      630                      635                      640

Phe Leu Glu Tyr Asp Glu Gly Pro Val Leu Leu Glu Val Phe Val Asp
    645                      650                      655

Lys Asp Thr Leu Val Leu Pro Met Val Pro Ala Gly Phe Pro Leu His
    660                      665                      670

Glu Met Val Leu Glu Pro Pro Lys Pro Lys Asp Ala
    675                      680

<210> 9
<211> 684
<212> PRT
<213> 人工序列

```

<220>

<223> 突变的 ALS 2

<400> 9

Met Ser Ala Thr Arg Ala Ala Thr Arg Thr Ala Ala Ala Leu Ser Ser
1 5 10 15

Ala Leu Thr Thr Pro Val Lys Gln Gln Gln Gln Gln Leu Arg Val
20 25 30

Gly Ala Ala Ser Ala Arg Leu Ala Ala Ala Ala Phe Ser Ser Gly Thr
35 40 45

Gly Gly Asp Ala Ala Lys Lys Ala Ala Ala Ala Arg Ala Phe Ser Thr
50 55 60

Gly Arg Gly Pro Asn Ala Thr Arg Glu Lys Ser Ser Leu Ala Thr Val
65 70 75 80

Gln Ala Ala Thr Asp Asp Ala Arg Phe Val Gly Leu Thr Gly Ala Gln
85 90 95

Ile Phe His Glu Leu Met Arg Glu His Gln Val Asp Thr Ile Phe Gly
100 105 110

Tyr Pro Gly Gly Ala Ile Leu Pro Val Phe Asp Ala Ile Phe Glu Ser
115 120 125

[0009]

Asp Ala Phe Lys Phe Ile Leu Ala Arg His Glu Gln Gly Ala Gly His
130 135 140

Met Ala Glu Gly Tyr Ala Arg Ala Thr Gly Lys Pro Gly Val Val Leu
145 150 155 160

Val Thr Ser Gly Pro Gly Ala Thr Asn Thr Ile Thr Pro Ile Met Asp
165 170 175

Ala Tyr Met Asp Gly Thr Pro Leu Leu Val Phe Thr Gly Gln Val Gln
180 185 190

Thr Ser Ala Val Gly Thr Asp Ala Phe Gln Glu Cys Asp Ile Val Gly
195 200 205

Ile Ser Arg Ala Cys Thr Lys Trp Asn Val Met Val Lys Asp Val Lys
210 215 220

Glu Leu Pro Arg Arg Ile Asn Glu Ala Phe Glu Ile Ala Met Ser Gly
225 230 235 240

Arg Pro Gly Pro Val Leu Val Asp Leu Pro Lys Asp Val Thr Ala Val
245 250 255

Glu Leu Lys Glu Met Pro Asp Ser Ser Pro Gln Val Ala Val Arg Gln
260 265 270

Lys Gln Lys Val Glu Leu Phe His Lys Glu Arg Ile Gly Ala Pro Gly
275 280 285

Thr Ala Asp Phe Lys Leu Ile Ala Glu Met Ile Asn Arg Ala Glu Arg
 290 295 300
 Pro Val Ile Tyr Ala Gly Gln Gly Val Met Gln Ser Pro Leu Asn Gly
 305 310 315 320
 Pro Ala Val Leu Lys Glu Phe Ala Glu Lys Ala Asn Ile Pro Val Thr
 325 330 335
 Thr Thr Met Gln Gly Leu Gly Gly Phe Asp Glu Arg Ser Pro Leu Ser
 340 345 350
 Leu Lys Met Leu Gly Met His Gly Ser Ala Tyr Ala Asn Tyr Ser Met
 355 360 365
 Gln Asn Ala Asp Leu Ile Leu Ala Leu Gly Ala Arg Phe Asp Asp Arg
 370 375 380
 Val Thr Gly Arg Val Asp Ala Phe Ala Pro Glu Ala Arg Arg Ala Glu
 385 390 395 400
 Arg Glu Gly Arg Gly Gly Ile Val His Phe Glu Ile Ser Pro Lys Asn
 405 410 415
 Leu His Lys Val Val Gln Pro Thr Val Ala Val Leu Gly Asp Val Val
 420 425 430
 Glu Asn Leu Ala Asn Val Thr Pro His Val Gln Arg Gln Glu Arg Glu
 435 440 445
 Pro Trp Phe Ala Gln Ile Ala Asp Trp Lys Glu Lys His Pro Phe Leu
 450 455 460
 Leu Glu Ser Val Asp Ser Asp Asp Lys Val Leu Lys Pro Gln Gln Val
 465 470 475 480
 Leu Thr Glu Leu Asn Lys Gln Ile Leu Glu Ile Gln Glu Lys Asp Ala
 485 490 495
 Asp Gln Glu Val Tyr Ile Thr Thr Gly Val Gly Ser His Glu Met Gln
 500 505 510
 Ala Ala Gln Phe Leu Thr Trp Thr Lys Pro Arg Gln Trp Ile Ser Ser
 515 520 525
 Gly Gly Ala Gly Thr Met Gly Tyr Gly Leu Pro Ser Ala Ile Gly Ala
 530 535 540
 Lys Ile Ala Lys Pro Asp Ala Ile Val Ile Asp Ile Asp Gly Asp Ala
 545 550 555 560
 Ser Tyr Ser Met Thr Gly Met Glu Leu Ile Thr Ala Ala Glu Phe Lys
 565 570 575
 Val Gly Val Lys Ile Leu Leu Leu Gln Asn Asn Phe Gln Gly Met Val
 580 585 590
 Lys Asn Trp Gln Asp Leu Phe Tyr Asp Lys Arg Tyr Ser Gly Thr Ala

[0010]

595	600	605
Met Phe Asn Pro Arg Phe Asp Lys Val Ala Asp Ala Met Arg Ala Lys		
610	615	620
Gly Leu Tyr Cys Ala Lys Gln Ser Glu Leu Lys Asp Lys Ile Lys Glu		
625	630	635
Phe Leu Glu Tyr Asp Glu Gly Pro Val Leu Leu Glu Val Phe Val Asp		
645	650	655
Lys Asp Thr Leu Val Leu Pro Met Val Pro Ala Gly Phe Pro Leu His		
660	665	670
Glu Met Val Leu Glu Pro Pro Lys Pro Lys Asp Ala		
675	680	
<210> 10		
<211> 684		
<212> PRT		
<213> 人工序列		
<220>		
<223> 突变的 ALS 3		
<400> 10		
Met Ser Ala Thr Arg Ala Ala Thr Arg Thr Ala Ala Ala Leu Ser Ser		
1	5	10
15		
[0011]		
Ala Leu Thr Thr Pro Val Lys Gln Gln Gln Gln Gln Gln Leu Arg Val		
20	25	30
Gly Ala Ala Ser Ala Arg Leu Ala Ala Ala Ala Phe Ser Ser Gly Thr		
35	40	45
Gly Gly Asp Ala Ala Lys Lys Ala Ala Ala Ala Arg Ala Phe Ser Thr		
50	55	60
Gly Arg Gly Pro Asn Ala Thr Arg Glu Lys Ser Ser Leu Ala Thr Val		
65	70	75
80		
Gln Ala Ala Thr Asp Asp Ala Arg Phe Val Gly Leu Thr Gly Ala Gln		
85	90	95
Ile Phe His Glu Leu Met Arg Glu His Gln Val Asp Thr Ile Phe Gly		
100	105	110
Tyr Pro Gly Gly Ala Ile Leu Pro Val Phe Asp Ala Ile Phe Glu Ser		
115	120	125
Asp Ala Phe Lys Phe Ile Leu Ala Arg His Glu Gln Gly Ala Gly His		
130	135	140
Met Ala Glu Gly Tyr Ala Arg Ala Thr Gly Lys Pro Gly Val Val Leu		
145	150	155
160		
Val Thr Ser Gly Pro Gly Ala Thr Asn Thr Ile Thr Pro Ile Met Asp		
165	170	175

Ala Tyr Met Asp Gly Thr Pro Leu Leu Val Phe Thr Gly Gln Val Gln
 180 185 190
 Thr Ser Ala Val Gly Thr Asp Ala Phe Gln Glu Cys Asp Ile Val Gly
 195 200 205
 Ile Ser Arg Ala Cys Thr Lys Trp Asn Val Met Val Lys Asp Val Lys
 210 215 220
 Glu Leu Pro Arg Arg Ile Asn Glu Ala Phe Glu Ile Ala Met Ser Gly
 225 230 235 240
 Arg Pro Gly Pro Val Leu Val Asp Leu Pro Lys Asp Val Thr Ala Val
 245 250 255
 Glu Leu Lys Glu Met Pro Asp Ser Ser Pro Gln Val Ala Val Arg Gln
 260 265 270
 Lys Gln Lys Val Glu Leu Phe His Lys Glu Arg Ile Gly Ala Pro Gly
 275 280 285
 Thr Ala Asp Phe Lys Leu Ile Ala Glu Met Ile Asn Arg Ala Glu Arg
 290 295 300
 Pro Val Ile Tyr Ala Gly Gln Gly Val Met Gln Ser Pro Leu Asn Gly
 305 310 315 320
 Pro Ala Val Leu Lys Glu Phe Ala Glu Lys Ala Asn Ile Pro Val Thr
 325 330 335
 Thr Thr Met Gln Gly Leu Gly Gly Phe Asp Glu Arg Ser Pro Leu Ser
 340 345 350
 Leu Lys Met Leu Gly Met His Gly Ser Ala Tyr Ala Asn Tyr Ser Met
 355 360 365
 Gln Asn Ala Asp Leu Ile Leu Ala Leu Gly Ala Arg Phe Asp Asp Arg
 370 375 380
 Val Thr Gly Arg Val Asp Ala Phe Ala Pro Glu Ala Arg Arg Ala Glu
 385 390 395 400
 Arg Glu Gly Arg Gly Gly Ile Val His Phe Glu Ile Ser Pro Lys Asn
 405 410 415
 Leu His Lys Val Val Gln Pro Thr Val Ala Val Leu Gly Asp Val Val
 420 425 430
 Glu Asn Leu Ala Asn Val Thr Pro His Val Gln Arg Gln Glu Arg Glu
 435 440 445
 Pro Trp Phe Ala Gln Ile Ala Asp Trp Lys Glu Lys His Pro Phe Leu
 450 455 460
 Leu Glu Ser Val Asp Ser Asp Asp Lys Val Leu Lys Pro Gln Gln Val
 465 470 475 480
 Leu Thr Glu Leu Asn Lys Gln Ile Leu Glu Ile Gln Glu Lys Asp Ala

[0012]

485 490 495
 Asp Gln Glu Val Tyr Ile Thr Thr Gly Val Gly Ser His Gln Met Gln
 500 505 510
 Ala Ala Gln Phe Leu Thr Trp Thr Lys Pro Arg Gln Trp Ile Ser Ser
 515 520 525
 Gly Gly Ala Gly Thr Met Gly Tyr Gly Leu Pro Ser Ala Ile Gly Ala
 530 535 540
 Lys Ile Ala Lys Pro Asp Ala Ile Val Ile Asp Ile Asp Gly Asp Ala
 545 550 555 560
 Ser Tyr Ser Met Thr Gly Met Glu Leu Ile Thr Ala Ala Glu Phe Lys
 565 570 575
 Val Gly Val Lys Ile Leu Leu Leu Gln Asn Asn Phe Gln Gly Met Val
 580 585 590
 Lys Asn Val Gln Asp Leu Phe Tyr Asp Lys Arg Tyr Ser Gly Thr Ala
 595 600 605
 Met Phe Asn Pro Arg Phe Asp Lys Val Ala Asp Ala Met Arg Ala Lys
 610 615 620
 Gly Leu Tyr Cys Ala Lys Gln Ser Glu Leu Lys Asp Lys Ile Lys Glu
 625 630 635 640
 Phe Leu Glu Tyr Asp Glu Gly Pro Val Leu Leu Glu Val Phe Val Asp
 645 650 655
 Lys Asp Thr Leu Val Leu Pro Met Val Pro Ala Gly Phe Pro Leu His
 660 665 670
 Glu Met Val Leu Glu Pro Pro Lys Pro Lys Asp Ala
 675 680

[0013]

<210> 11
 <211> 2052
 <212> DNA
 <213> 人工序列
 <220>
 <223> 突变的 ALS 1
 <400> 11
 atgagcgcga ccgcgcgcgc gaagaggaca ggcgcgcgc tgcctcgcgc gctgaagcgc 60
 ectgtaaagc agcagcagca gcagcagctg cgcgtaggcg cgcgctgcgc acggetggcg 120
 gccgcgcgcgt tctcgtccgc caccggcgga gacgcggcca agaagcggcg ccgcgcgagg 180
 ggccttccca cgggacgcgc ccccaacgcg acacgcgaga agagctcgcg ggccaagctc 240
 caggcgcgcga cggacgatgc gcgcttcgct ggcctgacgc gcgcgcgaat ctttcagtag 300
 ctcatgcgcg agcaaccaggt ggacaccatc ttggetacc ctcgcgcgcg cttctgcgcc 360
 gtttttgatg ccatttttga gactgaagcc ttcaagttca ttctcgcgcg ccacgagcag 420
 gggcgcgcgc acatggccga ggcctacgcg cgcgcgaagg gcaagcccgg cgttgctctc 480
 gtcacctcgc gccctggagc caccaacacc atcacccga tcatggatgc ttacatggac 540

[0014]

```

ggtaacgcgc tgcctgtgtt caccggccag gtgccacct ctctgtcgg caccgacgt 600
ttccaggagt gtgacattgt tggcatcagc cgcgcgtgca ccaagtggaa egtcattgtc 660
aaggacgtga aggagctccc gcgcgcctc aatgaggcct ttgagattgc catgagcggc 720
cgcccggtc ccgtgctcgt cgtcttctct aaggatgtga ccgcggttga gctcaaggaa 780
atgcccgaca gctcccccca ggttctgtg cgcacagaagc aaaaggctga gcttttccac 840
aaggagcgca ttgctctcc tggcaaggcc gacttcaagc tcatgtccga gatgaacaa 900
cgtcgggagc gaccctcat ctatcttggc cagggtgtca tgcagagccc gttgaatggc 960
ccgctgttgc tcaaggagtt cgcggagaag gccaacatc ccgtgaccac caccatgcag 1020
ggtctcggcg gctttgacga gcttctccc ctctctctca agatgtctcg catgcacggc 1080
cttgcctacg ccaactactc gatgcagaa gccgatctta tcttggcgt ccgtgcgcc 1140
tttgatgate gtgtgacggg ccgcgttgac gctttgtct cggaggtctg ccgtgcggag 1200
cgcgaggccc gcggtggcat cgttctctt gagatttccc ccaagaacct ccacaaggct 1260
gtccagccca ccgtcgcgtt cctcggcgac gtgtctgaga acctcgccaa cgtcacgccc 1320
cacgtgcagc gccaggagcg cgagccgtgg ttgtcgcaga tgcctgattg gaaggagaag 1380
cacccttttc tctctgagtc tgttattctg gacgacaagg ttctcaagcc gcagcaggtc 1440
ctcaggagc ttaacaagca gattctcgag attcaggaga aggacgcgca ccaggaggct 1500
tacatccca cggcgctcgg aagccaccag atgcaggcag cgcagttctt taactggacc 1560
aagccgcgcc agtggatctc ctccgggtgg gccggcacta tgggtctcgg ccttctctcg 1620
gccattggcg ccaagattgc caagcccgat gctattgtta ttgacatga tgggtatgct 1680
cttatttctg tgaccgttat ggaattgact acagcagccg aattcaaggt tggcgtgaag 1740
attctttttt tgcagaacaa ctctcaggcg atgttcaaga acgttcaaga tctcttttac 1800
gacaagegct actcgggcac cgcctgttc aaccgcgctt tgcacaaggi ccgcgatgcg 1860
atgcgtgcc aagggtctct ctgcgcgaaa cgtctggagc tcaaggacaa gatcaaggag 1920
ttctctgagt acgatgaggg tccctctctc ctctgaggtt tcttgacaaa ggacaagctc 1980
gttttgccta tggctccccc ttgctttctg ctccacgaga tggctctga gctctctaag 2040
cccaaggagc cc 2052

```

<210> 12

<211> 2052

<212> DNA

<213> 人工序列

<220>

<223> 突变的 ALS 2

<400> 12

```

atgagcgaga cccgcgcgga gacgaggaca gggcgcggc tgtctctggc gctgacgacg 60
ctgttaagc agcagcagca gacgagctg cgcgtaggcg cggcgtcggc accgtggcg 120
gccgcggcgt tctgtctcgg caccggcgga gacgcggcca agaaggcggc ccgcgcgagg 180
gctttctcca cgggacgcgg ccccaacgcg acacgcgaga agagctcgtt ggccaaggct 240
caggcgcgga cggacgatgc gcgtttctgc ggcttgaccg gcgcccaaat ctttcatgag 300
ctcatgcgag agcaccaggt ggacaccatc ttggcttacc ctggcgcggc cattctcccc 360
gtttttgatg ccatttttga gattgaagcc ttcaagtcca ttctctctcg ccacgagcag 420

```

	ggcgccggcc acatggccga ggctacgcg cgcgccacgg gcaagcccgg cgttgctctc	480
	gtcacctcgg gccctggage caccacacac atcacccaga tcatggatgc ttacatggac	540
	ggfacgcgcg tgcctgtgtt caccggccag gtgcagacct ctgtgtcgg caeggacgt	600
	ttccaggagt gtgacattgt tggcatcagc cgcgcgtgca ccaagtggaa cgtcatggtc	660
	aaggacgtga aggagctccc gcgcgcac acatgagcct ttgagattgc catgagcggc	720
	cgcgcgggtc ccgtgtcgtt cgtcttctt aaggatgtga ccgcgttga gctcaaggaa	780
	atgcgccaca gctccccca ggttgctgtg cgcagaagc aaaaggctga gcttttccac	840
	aaggagcgca ttggcgctcc tggcaggcc gacttcagc tcattgcca gatgatcaac	900
	cgtgaggaga gaccctcat ctatgtcggc cagggtgtca tgcagagccc gtgaatggc	960
	ccggtgtgic tcaaggatt ccgcgagaag gccacattc ccgtgaccac caccatgcag	1020
	ggtctcggcg gctttgaca gcgtagtccc ctctccctca agatgtcgg catgacggc	1080
	tctgcctacg ccaactac acatgcagac gccgatctt tctggcgtt cgtgcccgc	1140
	tttgatgac gtgtgacggg ccgcgttgac gcctttgctc cggaggtctg ccgtgccag	1200
	cgcgaggccc gcggtggcat cgttcacttt gagatttccc ccaagaacct ccacaaggtc	1260
	gtccagccca ccgtcgggt cctcggcgac gtggtcgaga acctgccaa cgtcacgccc	1320
	cacgtgcagc gccaggagcg cgcgcgtgg ttgcccaga tgcgcgattg gaaggagaag	1380
	caaccttttc tctcagatc tgttgattcg gacgacaagg ttctcaagcc gcagcaggtc	1440
	ctcaggaga ttaacagca gattctcgag attcaggaga aggagccga ccaggaggtc	1500
	tacatcacea cggcgctcgg aagccaccag atgcaggcag cgcagttcct tacttgacc	1560
[0015]	aagcccgccc agtggatctc ctcgggtgac gccgcacta tgggctaagg ccttccctcg	1620
	gccattggcg ccaagattgc caagccgac gctattgcta ttgacataga ttgtagtct	1680
	tcttattcga tgaccggtat ggaattgac acagcagcgc aattcaaggt tggcgtgaag	1740
	attcttcttt tgcagaacaa ctctcagggc atggtcaaga actggcagga tctcttttac	1800
	gacaagcgct acctggggac cgcctgttc aacccggctc tgcacaaggt ccgcgattcg	1860
	atgcgtacca aggtctctc ctcgcgaaa cagtccgagc tcaaggacaa gatcaaggag	1920
	ttctctgagt acgatgaggg tcccgtctc ctcgaggttt tegtggacaa ggacacgctc	1980
	gttttgccca ttgtccccc ttgctttccg ctccacgaga ttgtctctga gctcttaag	2040
	cccaaggacg cc	2052

<210> 13

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 突变的 ALS 3

<400> 13

atgagcgaga ccgcgcggc gacgaggaca gcgcgcgcgc tgtctctggc gctgacgacg 60 |cctgtaaage agcagcagca gcagcagctg cgcgtaggcg ccgcgtcggc accgctggcg 120 |gccgcggcgt tctcgtccg caccggcgga gacgcggcca agaaggcggc ccgcgcgagg 180 |gggtttctca cgggacggcg ccccaacgcg acacgcgaga agagctcgtt ggccacggtc 240 |caggcgcgga ccgacgatgc gcgttctgc ggctgacgc gcgcacaaat ctctcatgag 300 |ctcatgcgcg agcaccaggt ggacaccatc ttgtgctacc ctggcggcgc cttctgccc 360 |

	gtttttgatg ccatttttga gagtgacgee ttcaagtta ttctcgetcg ccaegagcag	420
	ggcgccggcc acatggccga gggtacgcg cgcgccacgg gcaagcccg cggtgtctc	480
	gtcaccctcg gccctggagc caccaacacc atcaccocga tcattgatgc ttacatggac	540
	ggtacgcgcg tgcctggtt caecggccag gtcagacct ctgctgtcgg caccgacgt	600
	ttccaggagt gtgacattgt tggcatcagc cgcgcgtgca ccaagtggaa cgtcatggtc	660
	aaggacgtga aggagctccc gcgcgcate aatgagacct ttgagattgc catgagcgcc	720
	cgcgccggtc cgtgtctgt cgtcttctct aaggatgtga ccgcggtga gctcaaggaa	780
	atgcgccaca gctccccca ggttgcgtg cgcagagaag aaaaggtega gcttttcac	840
	aaggagcgca ttggcgtccc tggcaccgcc gacttcaagc tcattgcga gatgatcaac	900
	cgtcgggagc gaccctcat ctatgciggc cagggtgca tgcagagccc gtlgaalggc	960
	ccgctgtgct tcaaggagtt cgcggagaag gccaacatc ccgtgaccac caccatgcag	1020
	ggtctcgccg gctttgacga gctagtccc ctctccctca agatgtctcg catgcacgc	1080
	tctgcctacg ccaactactc gatgcagaa gcgatctta tcttggcgt cgggtcccgc	1140
	tttgatgac gtgtgacgg ccgcgttac gctttgctc cggagctcg ccgtgccgag	1200
	cgcgagggcc gcggtggcat cgttcatctt gagatttcc ccaagaacct ccacaaggtc	1260
	gtccagccca ccgtcgggt cctcgccgac ggggtcgaga accctgccaa cgtcacgcc	1320
	cactgtcagc gccaggagcg cgcgcgtgg ttgcgcaga tgcgcgatt gaaggagaag	1380
	cacccttttc tgcctagtc tgttgattcg gacgacaagg ttctcaagcc gcagcaggtc	1440
	ctcaccgagc ttaacaagca gattctcgag attcaggaga aggacgcga ccaggaggtc	1500
[0016]	tacatccca cggcgctcgg aagccaccag atgcaggcag cgcagttct taactggacc	1560
	aagccgcgcc agtggatctc ctccgggtgc gccggcacta tgggtacgg ccttccctcg	1620
	gccattggcg ccaagattgc caagccgat gctattgta ttgacatga tggtagatgt	1680
	tcttattcga tgaccggtat ggaattgat acagcagccg aattcaaggt tggcgtgaag	1740
	attcttttt tgcagaacaa ctctcagggc atgttcaga acgttcagga tctctttac	1800
	gacaagcgt actcgggac cgcctatgtt aaccgcgct tcgacaaggt cgcctatgcg	1860
	atgcgtacca aggtctctc ctgcgcgaaa cgtcggagc tcaaggacaa gatcaggag	1920
	ttctctgagt acgatgagg tcctgctc ctccagggtt tctgtgacaa ggacacgtc	1980
	gtcttgccca tggctcccgc tggctttccg ctccacgaga tggctctga gctcttaag	2040
	cccaaggacg cctaa	2055

<210> 14

<211> 12

<212> DNA

<213> 裂殖壶菌属

<400> 14

cacgacgagt tg

12

<210> 15

<211> 53

<212> PRT

<213> 裂殖壶菌属

<400> 15

Met Ala Asn Ile Met Ala Asn Val Thr Pro Gln Gly Val Ala Lys Gly

1	5	10	15
Phe Gly Leu Phe Val Gly Val Leu Phe Phe Leu Tyr Trp Phe Leu Val			
20	25	30	
Gly Leu Ala Leu Leu Gly Asp Gly Phe Lys Val Ile Ala Gly Asp Ser			
35	40	45	
Ala Gly Thr Leu Phe			
50			

<210> 16
 <211> 21
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物 S4termF

<400> 16
 gatcccatgg caegtgtctac g 21

<210> 17
 <211> 22
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物 S4termR

<400> 17
 ggcaacatgt atgataagat ac 22

[0017]

<210> 18
 <211> 21
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物 C2mcsSmaF

<400> 18
 gatccceggg ttaagcttgg t 21

<210> 19
 <211> 20
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物 C2mcsSmaR

<400> 19
 actggggccc gtttaaactc 20

<210> 20
 <211> 34
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物 5'tubMCS_BglI

<400> 20
 gactagatct caatttttagg ccccccactg accg 34

<210> 21
 <211> 34

[0018]

<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成引物 3' SV40MCS_SaI	
<400>	21	
	gactgtcgac catgtaatgai aagatacatt gatg	34
<210>	22	
<211>	28	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成引物 5' A1SproNde3	
<400>	22	
	gactcatatg gccaggect actttcac	28
<210>	23	
<211>	34	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成引物 3' A1SternBglII	
<400>	23	
	gactagatct gggtaaggc agaagaattc egcc	34
<210>	24	
<211>	97	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成引物 sec. Gfp5' Ib	
<400>	24	
	tactgggtcc ttgtggcct egcccttctc ggcatggct tcaaggatcat egccgggtgac	60
	tcgccggta cgtcttctcat ggtgagcaag ggccgagg	97
<210>	25	
<211>	32	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成引物 sec. Gfp3' Spe	
<400>	25	
	gacgggtacc ggtgtctttt gttttgattt ct	32
<210>	26	
<211>	105	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成引物 sec. Gfp5' Bam	
<400>	26	
	taatggatcc atggccaaca tcatggccaa cgtcacgcc caggcgctcg ccaagggtt	60
	tggectcttt gfcggcgtgc tcttttttct ctactgggtc cttgt	105
<210>	27	
<211>	40	
<212>	DNA	

[0019]	<213>	人工序列	
	<220>		
	<223>	合成引物 ss.eGfpHELD3' RV	
	<400>	27	
		ccatgatatct tacaactcgt cgtggttgta cagctcgtcc	40
	<210>	28	
	<211>	105	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成引物 sec.Gfp5' Bam2	
	<400>	28	
		taatggatcc atggccaaca tcatggccaa cgtaacgcc caggcgctcg ccaagggtt	60
		tggectcttt gtggcggtgc tcttctttct ctactggttc cttgt	105
	<210>	29	
	<211>	27	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成引物 prREZ15	
	<400>	29	
		cggatccgc gaatcaagaa ggtagcc	27
	<210>	30	
	<211>	27	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成引物 prREZ16	
	<400>	30	
		cggatccgt ctctgccgt tttctt	27
	<210>	31	
	<211>	30	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成引物 prREZ17	
	<400>	31	
		cggatccgaa agtgaacctt gtcctaacc	30
	<210>	32	
	<211>	27	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成引物 prREZ18	
	<400>	32	
		ctctagacag atccgcaacca tcggccg	27
	<210>	33	
	<211>	32	
	<212>	DNA	
	<213>	人工序列	
	<220>		

[0020]

<223> 合成引物 5' eGFP kpn	
<400> 33	
gaatgggtacc atgggtgaagc aaggcgagag ag	32
<210> 34	
<211> 34	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成引物 3' eGFP_xba	
<400> 34	
gaattctaga ttacttgtag agctctgtcca tggc	34
<210> 35	
<211> 32	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成引物 5' ORFCproKpn-2	
<400> 35	
gacgggtacc ggtgttcttt gttttgattt ct	32
<210> 36	
<211> 31	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成引物 3' ORFCproKpn-2	
<400> 36	
gacgggtacc gtctctgccg ctttttcttt a	31
<210> 37	
<211> 20	
<212> PRT	
<213> 裂殖壶菌属	
<400> 37	
Met Lys Phe Ala Thr Ser Val Ala Ile Leu Leu Val Ala Asn Ile Ala	
1 5 10 15	
Thr Ala Leu Ala	
20	
<210> 38	
<211> 60	
<212> DNA	
<213> 裂殖壶菌属	
<400> 38	
atgaagttag cgacctggg cgaattttg attgtggcca acatagccac cgcctctggg	60
<210> 39	
<211> 28	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成引物 5' ss-X Bgl long	
<400> 39	
gactagatct atgaagttag cgacctcg	28

<210> 40
 <211> 33
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物 3' ritz_kap_bli_Bgl

<400> 40
 gactagatct tcagcaatca ccgggttaa agg 33

<210> 41
 <211> 702
 <212> DNA
 <213> 人工序列

<220>
 <223> 优化的 Sec1

<400> 41
 atgaagtccg cgacctcggt cgaattttg cttgtgcca acatagccac cgcctcgcg 60
 cagatctccc tcagccagtc ccccgccatc ctctccgtt ccccggtga gaaggtgacc 120
 atgacctgcc gcctagctc ctccgtctcg tacatccact ggttccagca gaagcccgcc 180
 tctcccccac agccttgat ctacgccacc tccaacctcg cctccggtgt tccggttct 240
 ttttccggtt ccggttccg cactctctac tctctacca tctcccggt cgaggccgag 300
 gatccgccca cctactctg ccagcagtcg accagcaacc cccccacct cggcggtggt 360
 accaagctcg agattaagcg caccgtgcc gcccccctcg tctcatttt tccccctcc 420
 gatgagcage tcaagtcgg tactccctcc gtgtttgccc tctcaacaa cttctacccc 480
 [0021] cgtgagccca aggtccagtg gaaggicgac aacgcgttc agtccggtaa ctcacaggag 540
 tccgtaccg agcaggatc gaaggacage acctactccc tctctccac cctcaccctc 600
 tcaagcccg actaagaa gcaaaaggtc tacccctcg aggtacgca ccagggtctt 660
 tctcccccg tcacgaagtc cttaaccgc ggtgagtgct ga 702

<210> 42
 <211> 853
 <212> DNA
 <213> 裂殖壶菌属

<400> 42
 ctccatcgat cgtcgggta aaaagaagg aagaagaag gaaaaagaa ggcgtgcga 60
 ccgagtcgc cgtgagcgc ccctcgcgg ccccgaggag cctcccggtt agtcccccgc 120
 ccgcgcgcg cagtcctccg ggaggaatc cgcacctc gcgcacctt cgcgcctcgc 180
 cgattcccg cctcccttt tccgttttt cgcgcctcc gctcgggccc ggtcgcgcg 240
 cgcctccgct cctatctgt ccccagggg gaactccga ccttttgcgc cgcctccgc 300
 cgcgcggccc gccccgcgc cctggtttcc ccccgagag cgcccgctc gccgcgaaa 360
 gactcgcgc gtgcgcgcc gagaacggg tggcgcgcc gccgcgcgc gccggcgcg 420
 gccgcgcgta gccggggcta gccgcgggt aggcgaaac cgcctcccg gccgcgcgc 480
 cgcctccgct agagcagtc ccgcgccga ccgccaacgc agagaccgag accgaggtac 540
 gtcgcgcgc agcacgcgc gacgcgcgc agggacgag agcacgacgc cgcctccgc 600
 cgcgcggggg gggggaggga gaggcaggac gccggagcga ggtgcatgt ttcgcgcga 660
 gacgacgcgc cgcgcgtcg agaggagata aggcgttgg atcgcgagag gccagaccag 720

gctggaggcg aaaaatgggtg gagaggatag tatcttgcgt gcttggacga ggagactgac	780
gaggaggacg gatacgtcga tgatgatgtg cacagagaag aagcagttcg aaagcgacta	840
ctagcaagca agg	853
<210> 43	
<211> 1064	
<212> DNA	
<213> 裂殖壶菌属	
<400> 43	
ctcttatctg cctcgccgcg ttgaccgcgc cttgactctt ggcgttgcg gctcgatcc	60
tgctctgctc ggcagggcgc gcgggcgagt gggtgggtcc gcagccttcc gcctcgccc	120
gctagctcgc tcgcgcgcgt ctgcagccag cagggcagca ccgcacggca ggcaggtccc	180
ggcgccgcat gatcgatcca tcgatccatc gatccatcga tcgtcggttc aaaaagaaag	240
gaagaagaaa ggaanaagaa aggcgtgcgc acccgagtgc gcctgagcgc cccgctcgcg	300
gtcccgcgga gcctcgcgt tagtccccgc ccgcgcgcgc gaagtcctcc gggaggcctc	360
ggcgcctctt cgcgcctccc tcgcgcctcg ccgattccc gcctcccttt ttccgtttct	420
tcgcgcctc cgcgcgcgc cgcgtgcgc gcgcgcgcgt ccttatctgc tcccagggg	480
ggcactcgc accttttgcg ccgcgtgcgc ccgcgcgcgc cgcgcgcgcg ccttggttc	540
ccccgcgagc ggcgcgcgt cgcgcgcga agactgcgc cgtgcgcgc cgcgaacgcg	600
gtggcgccgc cgcgcgcgc ggcgcgcgc ggcgcgcgt aggcggggt aggcgcgcgc	660
taggcgaac gcgcgcgcg ggcgcgcgc ccgcgcctc cagagcagtc gcgcgcgcg	720
accgcgaac cagagacga gaccgaggta cgtcgccgc ggcgcgcgc cgcgcgcgcg	780
cagggcgcg ggcgcgcgc ccgcgcgcgc ccgcgcgcgc ggcgcgcgcg agagcgagga	840
cgcgcgcgc agcgtgcgt ttccgcgcgc agacgcgcgc gcgcgcgcgc gagagcgat	900
aaggcgttg gctcgcgaga ggcgcgcgc ggcgcgcgc gaaatgggt ggagagcga	960
gtatcttgcg tcttgagcgc aggcgcgcgc cgcgcgcgc ggcgcgcgc atgatgatgt	1020
gcacagagaa gaagcagtc gaaagcgcgc actagcaagc aagg	1064
<210> 44	
<211> 837	
<212> DNA	
<213> 裂殖壶菌属	
<400> 44	
cttcgttttc tcaacctatc tggacagcaa tccgcctatt gcttgatcc ccttcgcgc	60
ctcaatcact cgtccacgt cctcttccc cctctcctc tccgtcttt ccttcgcgc	120
ccccccccgc ccgcgcgcgc cgcgcgcgc gcgcgcgcgc cgcgcgcgc tccatactat	180
cctcgctccc aaaaatgggt gccttatagg gcccggttag gcgaaagtct agcaggcact	240
tgtttggcgc agagccgcgc ccgcgcctc ttgcgcgcga tggagaggga gagagagccc	300
gcctcgataa gcagagacag acagtgcgc tgacagacag acagagagac tggcagaccg	360
gaatacctcg aggtgagtc ggcgcgcgc agcgggcgcgc agcgggcgcgc caagaggagc	420
ggcgcgcgc ggcgcgcctc cgcgcgcgc ccgcgtattc tcgtgcgcgc cgcgcgcgc	480
cgggcgcgc ggcgtggtga tggttgaagc ggggcgcgc gaaatgttag atgagatgat	540
catacagcgc ggtcgcgcgc tcttgctcgc cttgcctgcg ttggtgcgc ggcctgcgc	600
gtttgcgaga aagagatga ggagagcgc gaggaagcgc gagaagactg acgtgtaggg	660

	cgcgcatgg atgacgatt gatgatiga ttgattggtt gattggctgt gtggtcgatg	720
	aacgtttaga ctcaggagac gtggttaaat tgttcttgcg ccagacgcga ggactccacc	780
	cccttcttcc gccctttacac agcctttttg tgaagcaaca agaaagaaaa agccaag	837
	<210> 45	
	<211> 1020	
	<212> DNA	
	<213> 裂殖壶菌属	
	<400> 45	
	ctttttccgc tctgcataat cctaaaagaa agactatacc ctagtcactg tacaaatggg	60
	acattttctct cccgagcgat agctaaggat ttttgcctcg tgigcactgt gtgcctcgcc	120
	cgcgcatcga aagtcacagga tcttaetgtt tctctttcct ttcttttatt tctgtttctc	180
	ttcttgcctt tctcaaccta tctggacagc aatccgcacac ttgccttgat ccccttccgc	240
	gctcaaatca ctgcctccac gtccctcttc cccctctcca tctccgtgct ttctctcgcc	300
	ccccccccc ccgcgcggcg gtgcgcgcgc gtggcgccgc ggccgcgaca ctttccatac	360
	tatctctgct cccaaaatgg gtgcgcctat agggcccgcg taggcgaaag tctagcaggc	420
	acttgcttgg ccgagagccg ccgcgcgcgc tcttgcgcgc ggatggagag ggagagagag	480
	cccgctctga taagcagaga cagacagtgc gactgacaga cagacagaga gactggcaga	540
	ccggaatacc tgcaggtgag tgcgcgcgcg gcgagcgzgc ggcagcgsga gcgcaagagg	600
	gacggcgccg ccgcgcgcgc ctgcgcgcgc ccgcggcgta ttctctgtcg cagcgccgag	660
	cagcgggacg ggcggtcggc tgatggttga agcgggcgcg ggtgaaatgt tagatgagat	720
[0023]	gacatcagac gacggctcgt gctcttggcg tgcttggctt gcttggctg gcggccctgc	780
	cgtgtttgcg agaaagagga tgcggagagc gacgaggaag gacgagaaga ctgacgtgta	840
	ggcgcgcgga tggatgacg attgattgat tgattgattg gttgatttgc tgtgtggtcg	900
	atgaacgtgt agactcaggg agcgttggtt aattgttctt gcgcagagcg cgaggactcc	960
	accccttctt ttgccttcta cacagccttt ttgtgaagca acaagaaaga aaaagccaag	1020
	<210> 46	
	<211> 1416	
	<212> DNA	
	<213> 裂殖壶菌属	
	<400> 46	
	ccggtccttg acgcttccgc ttccgcgcgc gccatcgatt caattcaccc atccgatacg	60
	ttccgcgcgc tcaagtcctg ctgcgcgcga cccctgcacg accacgcaca ggccaacgcg	120
	ccgctcagct cagcttctcg acgagtcgca cgtcacatat ctcaagtcga ttgcctgcct	180
	gcctgcctgc ctgcctgcct gcctgcctgc ctgcctgcct cagctctctt ttgcctctct	240
	tgcggcgccc gctgcgacgc gctgtacagg agaatactc caggaagtgc ggttgggata	300
	cgcgtggcgc tggcgctga tgcgctgac ggcgcgcgcg cagcgcgcgc aggggttgag	360
	cagaggacga agcgaggcga gacgagacag gccagcgcg ggagcgctc gctgcctgta	420
	gcagcagacc agggcgcgag aatgtacttt tcttgcggga gcgagacga ggttgcgcgc	480
	tgtgtgctgc cgttgcctct gcacgcgcgc cccgaattgg cgtagcgttg accgcgcgcg	540
	ggcgcgcgcg tctgtctcgcg gtcgcttgg ccgtgtatcg accgtgcggg cttagacagg	600
	gatggcgga gttcagcact gctgcgaccc ctgcgcgcgc agaacgagga gacgcgaggc	660

```

cggtttcaag ttgaaagga gaggaaggca ggcaaggagc tggaagcttg ccgcggaagg 720
cgcaggcatg cgtcacgtga aaaaaaggga tttaagagt agtaagtagg taiggtctac 780
aagtcacctt ttcttacttc gcggaacgtg ggtgtctegt gcgggcttc atctgtttt 840
tgttttttt tccgttaggc gcgtgcattg ctgatgagt ctacagcttc gtcgcagcg 900
agggcaggaa aataaggcgc ccgtgccgtc gagcgacag gacgtgcaag cgccttgga 960
gcgcagcctc ctgcacggc gagcatagag accgcggcgc atggactcca gcgaggatt 1020
ttcgacctc tctatcaage tgccttgac agccgggaat ggcagcctga ggagagagg 1080
ggaaggaag gacttgagc aasagagta aggcacctc aatcacggcg cgtgaaagcc 1140
agtcacctt cgcagaaga agacaaaagc ggttttttg ttctgatgg aaagaattc 1200
ttagaggaag aagcgcaca cagactcgcg ccctgcagat ttctgcgcag ctgcgcata 1260
aaccaggaa gtgtgtctg cgcgcacta tcagggttag cgcacgaata ccaaacgat 1320
tactagctac gcgcctgtga cccgaggatc gggccacaga cgttgtctct tgcctacca 1380
cgacctgcca gcgagaagat cgtccattac icatcg 1416

```

<210> 47

<211> 24

<212> DNA

<213> 人工序列

<220>

<223> 合成引物 5' 60S-807

<400> 47

tegatttgcg gatacttgc caca

24

[0024]

<210> 48

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 合成引物 3' 60S-2821

<400> 48

gacgacctcg cccttgaca c

21

<210> 49

<211> 34

<212> DNA

<213> 人工序列

<220>

<223> 合成引物 5' 60Sp-1302-Kpn

<400> 49

gaetggtacc ttttccgct ctgcataatc etaa

34

<210> 50

<211> 32

<212> DNA

<213> 人工序列

<220>

<223> 合成引物 3' 60Sp-Bam

<400> 50

gactggatcc ttgcttttt ctttctgtt gc

32

<210> 51

<211> 24

<212> DNA

[0025]	<213> 人工序列	
	<220>	
	<223> 合成引物 5' EF1-68	
	<400> 51	
	cgccgttgac cgccgttga ctct	24
	<210> 52	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成引物 3' EF1-2312	
	<400> 52	
	cgggggtagc ctcggggatg gact	24
	<210> 53	
	<211> 34	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成引物 5' EF1-54-Kpn	
	<400> 53	
	gactggtaac tcttatctgc ctgcgcctg tgac	24
	<210> 54	
	<211> 37	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成引物 5' EF1-1114-Bam	
	<400> 54	
	gactggatcc ctgcttgcct agtagtgcct ttcgaa	37
	<210> 55	
	<211> 29	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成引物 5' SecIP-kpn	
	<400> 55	
	gactggtaac ccgtccttga cgccttgc	29
	<210> 56	
	<211> 31	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成引物 3' SecIP-ba	
	<400> 56	
	gactggatcc gatgagtaat ggacgatctt c	31
	<210> 57	
	<211> 1614	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> Secretion signal	

[0026]	<400> 57	
	ggatccatga agtctcgac ctcggctgca attttgcitg tggccaaat agccaccgce	60
	ctcggctgca tgaccacaaga gacctcggac cggccctctcg tgcactttac ccccaacaag	120
	ggttggatga acgatcccaa eggccctctgg taagacgaga aggatgetaa gtggcacctt	180
	tactttcagt acaaccccaa cgacaccgta tggggcaacc cgtctttctg ggccacgce	240
	acctccgacg acctcaccaa ctgggaggac cagcccatig ctatcgccc caagcgcaac	300
	gactcgggag ctttttccgg ttccatgggt gtgactaca acaacacctc cgtttttttt	360
	aacgacacca ttgacccccc ccagcgcgc gtcgcactct ggacctacaa caccgccgag	420
	agcgaggagc agtacatcag ctacagctt gatggaggct acacctttac cgagtaccag	480
	aagaaccttg tctcgcgcgc caactccacc cagtccgcgc acctaaagt ttttggtae	540
	gagccttccc agaagtggat taagaccgc gctaaagtcg aggaatacaa gatcgagatc	600
	tacagcagcg acgacctcaa gtcttgaag cttagatcgc cctttgccc caagggtttt	660
	ctcgataacc agtacagatg ccccggtctc atcgaggctc ccaccgagca ggaccgctc	720
	aagctcact gggtcagtt tatttccatc aacctggcg cccctgcgcg cggcagcttc	780
	aaccagtaet tctcggctc cttaacggc acgcatttg aggccttga caaccagtc	840
	cgcgtcgtcg acttcgcaa ggactactac ggcctccaga cctctttaa caccgacccc	900
	acctacgcaa ggcctcgcg tattgtttg gcctccact gggagtaet cgtttctgc	960
	cccactaac cctgggcgag ctgatgtc ctcgtcgcga agtttctgt taacaccgag	1020
	taaccaggca accccgagac cgagcttat aacctgaag ccgagctat tctaacatc	1080
	tccaacgctg gcccctgct ccgcttctgt actaacata cctcaccaa ggccaactc	1140
	tacaacgtcg atctctcaa ctccaccgtt acctttgagt ttgagctgt ctacccgtc	1200
	aacaccaccc agacctctc caagtcgtc ttcccgacc tctcctctg gttaagggc	1260
	cttgaggacc ccgaggagta cctgcgatg ggttttgagg tctcgcctc ctccttctc	1320
	ctcgatcgcg gtaactcaa ggttaagtt gtaaggaga acctactt tactaacgt	1380
	atgagcgtca acaaccgac cttaagtc gagaacgac ttagctacta caaggttac	1440
	ggctctctcg accagaacat tctcgatc tactttaac agggagatg cgtcagcacc	1500
	aacacctact ttatgaccac tggaaacgc ctcggcagcg tgaacatgac caccggagtc	1560
	gacaacctct ttacattga caagttcag gtccgcgag ttaagtaaca taig	1614

<210> 58

<211> 11495

<212> DNA

<213> 人工序列

<220>

<223> pCL0076

<400> 58

cttttatctg cctcgcgcgc ttgacgcgc ctgactctt ggccttgcg gctcgcctc

60

tgctcgcgc ggcagggcg gggcgaggt ggttgggtc gcagccttc ggcctgcgc

120

gctagctcgc tccgcgcgtg ctgcagccag cagggcagca ccgcacggca ggcaggtcc

180

ggcgcggatc gatcgtacca tgcacatc gatcctcga tcttgcggc aaaaaaaag

240

gaagaagaaa ggaagaaaga aggcgtgcg acccgagtgc ggcgtgagc cccgcctgcg

300

gtcccgcgga gctcgcggt tagtccccc cccgcgcgc gtagtcccc ggcaggaatc

360

	gagcaccctt cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc 420
	tcgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc 480
	ggcgcgcgc acccttttgc cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc 540
	ccccgcgcgc ggcgcgcgc cgcgcgcgc agcgcgcgc cgcgcgcgc cgcgcgcgc 600
	gtgggcgcgc cgcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc 660
	taggcgcgcgc ggcgcgcgc ggcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc 720
	accgcgcgc cgcgcgcgc ggcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc 780
	cagggcgcgc ggcgcgcgc cgcgcgcgc cgcgcgcgc ggcgcgcgc ggcgcgcgc 840
	cgcgcgcgc agcgcgcgc ttcgcgcgc agcgcgcgc ggcgcgcgc ggcgcgcgc 900
	aagggcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc 960
	gtatcttgcg tgccttgagc aggcgcgc cgcgcgcgc ggcgcgcgc ggcgcgcgc 1020
	gcacgcgcgc ggcgcgcgc ggcgcgcgc actgcgcgc aagggcgcgc tgcgcgcgc 1080
	gaactgcgcgc ggcgcgcgc tgcgcgcgc catgcgcgc ggcgcgcgc ggcgcgcgc 1140
	cgcgcgcgc ggcgcgcgc tgcgcgcgc tgcgcgcgc aagggcgcgc tgcgcgcgc 1200
	caaggcgcgc tgcgcgcgc aggcgcgc taagtcgcgc cttgcgcgc agtcgcgcgc 1260
	taagcgcgc ggcgcgcgc cgcgcgcgc cgcgcgcgc ggcgcgcgc ggcgcgcgc 1320
	caactgcgcgc ggcgcgcgc tgcgcgcgc cgcgcgcgc aagggcgcgc ggcgcgcgc 1380
	cgcgcgcgc ggcgcgcgc ggcgcgcgc cgcgcgcgc tgcgcgcgc ggcgcgcgc 1440
	cgcgcgcgc tgcgcgcgc tgcgcgcgc caagcgcgc ggcgcgcgc ggcgcgcgc 1500
[0027]	cgcgcgcgc cttgcgcgc ggcgcgcgc tgcgcgcgc cgcgcgcgc cgcgcgcgc 1560
	cgcgcgcgc accgcgcgc ggcgcgcgc ggcgcgcgc tgcgcgcgc cgcgcgcgc 1620
	gattatgcgc ggcgcgcgc cgcgcgcgc caagcgcgc cgcgcgcgc ggcgcgcgc 1680
	caagtcgcgc aagtcgcgc cgcgcgcgc caagcgcgc ttcgcgcgc accgcgcgc 1740
	ggcgcgcgc cgcgcgcgc cgcgcgcgc ggcgcgcgc cgcgcgcgc cgcgcgcgc 1800
	gttatcttgc atgcgcgc ggcgcgcgc cgcgcgcgc tgcgcgcgc cgcgcgcgc 1860
	ctcttttaac ggcgcgcgc tgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc 1920
	caagcgcgc tgcgcgcgc agcgcgcgc taagcgcgc cgcgcgcgc ggcgcgcgc 1980
	cgcgcgcgc tgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc 2040
	cgcgcgcgc tgcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc 2100
	gagcgcgcgc attgcgcgc aggcgcgcgc tgcgcgcgc atgcgcgcgc cgcgcgcgc 2160
	gtgcgcgcgc ggcgcgcgc cgcgcgcgc caagcgcgc tgcgcgcgc tgcgcgcgc 2220
	caagcgcgc ggcgcgcgc agcgcgcgc cgcgcgcgc ggcgcgcgc ggcgcgcgc 2280
	ctgcgcgcgc ggcgcgcgc cgcgcgcgc cgcgcgcgc ggcgcgcgc ggcgcgcgc 2340
	gtgcgcgcgc atgcgcgc aggcgcgcgc cgcgcgcgc tgcgcgcgc ggcgcgcgc 2400
	caagcgcgc tgcgcgcgc aggcgcgcgc cgcgcgcgc cgcgcgcgc ggcgcgcgc 2460
	gagcgcgcgc tgcgcgcgc atgcgcgcgc cgcgcgcgc ggcgcgcgc ggcgcgcgc 2520
	cgcgcgcgc cgcgcgcgc aggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc 2580
	cgcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc 2640
	tgacgcgcgc cgcgcgcgc aggcgcgcgc ggcgcgcgc ggcgcgcgc ggcgcgcgc 2700

[0028]

```

cittctctaa cccgacagcg aatggcggga gggggcgggc taaaagatcg tattacatag 2760
tatttttccc ctactctttg tgtttgtctt ttttttttt ttgaacgat tcaagccact 2820
tgtctgggtt tacttgittg ttgtcttctg tcttgccttg ctigccttct tcttggtcag 2880
acggcccaaa aaagggaaaa aattcattca tggcacagat aagaaaaaga aaaagtttgt 2940
cgaccaccgt catcagaag caagagaaga gaaacaetcg cgtcacatt ctgcctcgcg 3000
taagaatctt agccacgat aagaagtaat ttgtccatct ggagaatctt tacatgagcg 3060
tittcaaget ggagcgtgag atcatacctt tcttgatctg aatgttccaa ccttgcctag 3120
gectctgtgc gatccgctag caatgcgtcg tactcccggt gcaactgcgc cctgcctca 3180
ttgtgacgtg agttcagatt ctctctgaga ccttcgagcg ctgctaattt cgcctgacgc 3240
tcctttcttt gtgttccat gacacgcgc ttcaccgtgc gtccacttc ttctcagac 3300
atgcccttgg ctgcctcgac ctgcctggta agcttctctg taatctctc gatctcgaa 3360
ttcttcttgc cctccatca ctgcgcacca tacttggcag cctgttcaac acgtcattg 3420
aaaaactttt cattctcttc cagctccgca acccgcgctc gaagctcatt cacttcgcgc 3480
accacggett cggcctcgag cgcgaatca gtcccgaaac ttccgaaag atacaccaag 3540
gcccctcgcg tctgtctgag cagcgtcctc atcagtcgcg tgttatcttc ggcagattc 3600
tccacttctt cctgaagcag ctccacggtg gctcttctg tctgaggctt cactgtctgg 3660
attagegett gcagctcttg cagctcgtc agcttggag agctcgtaat catgctttg 3720
cacttgccca gacgtcgag agcgttcgag agcgccttcg cgttatctgc catggacgt 3780
ctgcgcctcg cggcctccct gacgacgtc tcttgagtt tcactagatc atgtccaatc 3840
agcttgcggt gacgtctcc aatcagttc tgcattctgt ttgtgtgtc gggcgcgcgc 3900
tcttcttgag atttgcgaat ttctctctcg agctcgcgtt cagctctcag ggcgccttta 3960
agtagctcga agtcagcgc cgttagcccc agctcgtcgc cgcgttcag acagtcggtt 4020
agcttgattc gattccgett ttccatggca agtttaagat cctggcccag ctgcacctcc 4080
tgcgccttgc gcatcctgc cgttccgcc tggcgcaaaa gcttcgagtc gtatcttgc 4140
tgcctgcga ggcgaatgga acgaagagc gacttgagtt gcaactatt cctgcgcgag 4200
atgagcagca ttctgatctg catgaacacc tctcagagtt cgtatctctc tgcctctcc 4260
agctctgcgg gcgagcgag ctctccttgc agatgaagcg agggccgcag gcctccgaag 4320
agcactcttt gcgcgagatc ctctcctgc gtgcctctcc gcaggattgc ggctgtgtcc 4380
gccatcttgc cgcacagca gcttttgcgc gctctgcacc ttcaatttct ggctgcgctg 4440
gtgcgcgtgg tgcgccttgt gtggtgtctg gtgtgtgtgc tggctgtgtt gccctgtct 4500
ggtgtctcca cagacacgc cgtctctgt gtgtctcttc cggccccctc gcgcgcgcgc 4560
cgagcccccg ccgcgcgcgc tgcctgggtt ctgcgcctc tccgggggtt cctgcgcctc 4620
ggctctgcgc tccgcgcga cgtctgcgc gcgatgggtg cggatctgt ctagaggggc 4680
cttgaaggtt aagctatcc ctacccctct cctcgtctc gattctacgc gtaccgtca 4740
tcacacccat caccattgag tttaaacggg cccacgacg tctacgaga ttctgattc 4800
accgcgcctt tctatgaaag gtggcgctc ggaatcttt tccgggaagc cggctggatg 4860
atctctcagc gcgggatct catgctggag ttcttcgcgc accccaactt gtttattgca 4920
gcttataatg gttacaaata aagcaatagc atcacaatt tcacaaataa agcatttttt 4980

```

	tcactgcatt ctagtgttg ttgtccaaa ctcacaaatg tatcttaica tacatggteg	5040
	accfgcagga accfgcafa atgaatcgcc caacgcgcgg ggagaggcgg ttfgcgtati	5100
	gggcgcctct ccgcttcctc gctcactgac tgcctgcgct cggctgttgc gctgcggcga	5160
	gggtatcag ctcactcaaa ggccgtaata cggttatcca cagaatcagg ggataacgca	5220
	ggaaagaaca tgtgagcaaa agccacagca aagccagga accgtaaaaa ggcgcgttg	5280
	ctggcgcttt tccataggtt cgcgcgcct gacgagcgc acaaaaatcg acgtcaagt	5340
	cagagggtgc gaaccccgac aggaactataa agataccagg cgtttccccc tggaaagctc	5400
	ctcgtgcgtt ctcctgttcc gacctgcgc cttaccggat accgtgcgc cttctccct	5460
	tccggaagcg tggcgcttcc tcatagctca cgtgttaggt atctcagtic gggttaggtc	5520
	gttcgtcca agctgggtg tgtgcagaa ccccccgtc agcccgaccg ctgcgcctta	5580
	tccgtaact atcgtcttga gtccaaaccg gtaagacag acttatcgcc actgcagca	5640
	gccactgcta acaggattag cagagcgagg tatgtaggcg gtgtacaga gtcttgaag	5700
	tgttgcccta actacgcta cactagaaga acagtattg gtatctgcgc tctgtgaag	5760
	ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaaaaaac caccgtggt	5820
	agcggtggtt ttttgttg caagcagcag attacgcgc gaaaaaagg atctcaagaa	5880
	gaccccttga tctttctac ggggtctgac gctcagtgga acgaaactc acgttaaggg	5940
	attttgtca tgagattatc aaaaaggatc ttcaactaga tcttttaaa ttaaaaatga	6000
	aglllilaal caactaaag laataatgag laaacctggc ctgacgtta ccaatgcila	6060
	atcagtgagg cacttatctc agcagctgt ctatttcgtt cctccatagt tgcctgactc	6120
[0029]	cccgctgtgt agataactac gatacgggag ggettaecat ctgccccag tctgcaatg	6180
	atcccgag acacagctc accgctcca gallatcag caataaaca gccagcgga	6240
	aggcccgag gcagaagtgg tctgcaact ttatccgct ccatccagtc tattaatgt	6300
	tgcgggaag ctagagtaag tagttcgca gtaaatagtt tgcgcaagt tgttgcatt	6360
	gtacagga tcggtgtc accgtctg tlggtatgg ctccatcag ctccgttcc	6420
	caacgataa ggcagttac atgaccccc atgttgtca aaaaagcgt tagctcttc	6480
	ggtctcga tctgttcag aagtaagtgt gcgcaggt tatcaactat ggttatgca	6540
	gcactcata attctctac tgtcatgca tccgtaagat gctttctgt gactggtgag	6600
	tactcaacca agtcattctg agaataggt atgcggcgac cgagttgctc ttgcccgcg	6660
	tcaatacggg ataataccgc gccacatagc agaacittaa aagtgcctat cattgaaaa	6720
	cgctcttcgg ggcgaaaact ctcaaggatc ttaccgtgt tgagatccag ttcatgtaa	6780
	cccactgtg caccccctg atcttcagca tctttactt tcaacagct tctgggtga	6840
	gcaaaaacag gaaggcaaaa tgcgcacaaa aagggaataa ggccgacacg gaaatgtga	6900
	atactcatac tcttctttt tcaatattat tgaagcatt atcagggtta ttgtctcatg	6960
	agcgataca tatttgaatg tatttagnaa aataaaaaaa taggggttcc ggcacattt	7020
	ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt aacctataa	7080
	aataggcgta tcaaggcc ctctctctc gcgcgttgc gtgatgacgg tgaacacctc	7140
	tgacacatgc agctccgga gacggtcaca gcttctctg aagcgatgc cgggagcaga	7200
	caagccgctc agggcgctc agcggtgtt ggccggtgc ggggtggct taactatgcg	7260
	gcacagagc agattgtact gagagtgcac caagcttgc ctcaacgcaa ctaggcccag	7320

[0030]

gccacttcc actgtgtctt gtcttgctt tcacaccgac cgagtgtgca caaccgtatt 7380
 ttgcacaaag cgcaagatgc tcactcgact gtgaagcaaa gggtgcgcgc aagcgactgc 7440
 gactgcgagg atgaggatga ctggcagcct gticaaaaa tgaaaaaccg cgaigggtca 7500
 gctgccattc ggcgatgacg cctgcgagag acaagttaac tctgttcaact ggcatgtect 7560
 agcatcttta cgcgagcaaa attcaatcgc tttatttttt cagtctcgta acctctcgc 7620
 aaccgcgaat cgcggtttca gccgactaa tctgcagctg cgtggcactg tcagtcagtc 7680
 agtcagtcgt ggcgcgtgtt ccagcaccga ggctgcgcgt cgcgcgcctt ggaccgtgc 7740
 tgctactget agtggcacgg caggtaggag ctgtttgccg gaacaccagc agccgccagt 7800
 cgacgccagc caggggaaag tcggcgctgc aaggagagg aaggcgctgt gtcaaaacta 7860
 acgttgacca ctggcgccgc ccgacacgag caggaagcag gcagctgcag agcgacgcgc 7920
 gcaagtgcag atgcgcgaa agatccactt gcgcgcggcg ggcgcgactt tgcggcgcg 7980
 gcgcggaaac gtgcggaaag gacgggtgca gacggcgcg agtgacagtg ggcgcaaac 8040
 cgcgcagtaa gcgcggcg ggaaacgtat acgcagtgcc gcgggcgcgc gcacacagaa 8100
 gtatagcggc gccgaagtgg ggcgtgcgc gcgggaagtgc cgaatggcg ggcaaggaaa 8160
 ggaggagacg gaaagaggcg gggaaagaga gagagagaga gtgaaaaag aaagaagaa 8220
 agaagaagag aagaagaact cggagccacg ccgcggggag agagagaaat gaaagcagg 8280
 caccgcaaac caaagcaaac cagaccagc cagaccagc cgaggaggga gcgcgcgcag 8340
 gaccgcgcgc gcgagcgagc gacacggcg ccgcgagcag cagcgagcgc agcgcgcgag 8400
 cgagcaagcg ttgtcgcgag cgtcgcgag agcgagcggc aaggatgagc gcgaccgcgc 8460
 cggcgagcag gacagcgcg gcgcgtctct cggcgctgac gacgcctgta aagcagcgc 8520
 agcagcagca gctgcgcgta ggcgcgcgct cggcacgctt gcgcgcgcgc gcttctctgt 8580
 ccggcacggg cggagacgcg gccaaagaag cggccgcgcg gaggcgcttc tcacggggac 8640
 gggcccccac cgcgacacgc gagaagagct cgttgccac ggtccaggcg gcgacggacg 8700
 atgcgcgctt cgtcgccctg accgcgcgc naatcttca tgaatcatg cgcgagcacc 8760
 aggtggacac catcttttgc tactctggcg gcgccattct gccgttttt gatgccattt 8820
 ttgagagtga cgcgttcaa gtteattctc gctgcacac agcaggcgcg ccgcccacatg 8880
 gccgaggget acgcgcgcgc caccggcaag ccgcggcttg tctcgicac ctccggccct 8940
 ggagccacca acaccatcac ccgacatcat gatcttaca tggacggtae gccgtctctc 9000
 gtgttcaccg gccaggtgca gacctctgt gtgcgcacgg accgtttcca ggagtctgac 9060
 attgttgcca tcagcgcgcg gtgcaccaag tggaactgca tggtaaggga cgtgaaggag 9120
 ctcccgccc gcataatga ggcctttgag attgccatga gcggcgccc gggtcccgtg 9180
 ctctcgatc ttcctaagga tgtgaacgac gttgagctca aggaatgcc cagacagctc 9240
 cccaggttg ctgtgcgcca gaagcaaaag gtcgagcttt tcacaagga ggcatttggc 9300
 gctcttgcca cggccgactt caagctcatt gccgagatga tcaaccgtgc ggagcgaccc 9360
 gtcattatg ctggccaggg tgcattgcag agcccgctga atggcccgcg tgtctcaag 9420
 gagtctcgcg agaaggccaa cattcccgct accaccacca tgcagggctt cggcgcttt 9480
 gacgagcgta gtcctctct cctcaagatg ctggcatgc accgctctgc ctacgcacac 9540
 tactcgatgc agaagcgca tcttatctgt gcctcggtg ccgctttga tgatctgtg 9600

acgggcgcgg ttgacgcctt tgcctcggag gctcgccttg ccgagcgga ggccgcgggt 9660
ggcatcgttc actttgagat tcccccaag aacctccaca aggtcgtcca gccacgcgc 9720
gggttcctcg gcagctgggt cgagaacctc gccaacgtca cgcgccacgt gcagcgccag 9780
gagcgcgagc cgtgggttgc gcagatcgcc gattggaagg agaagcacc cttctctctc 9840
gagttctgtt attcggaaga caaggcttc aagccgcagc aggtcctcac ggagctaac 9900
aagcagattc tcgagattca ggagaaggac gccgaccagg aggtctacat caccacgggc 9960
gtcggaagcc accagatgca gccagcgagc ttcttacct ggaccaagcc gcgccagtgg 10020
atctctctcg gtggcgccgg cactatgggc tacggccttc cctcgcccat tggcgccaag 10080
attgccaagc ccgatgctat tgattatgac atcgatgtg atgtctctta ttgatgacc 10140
ggtatggaat tgatcacagc agcgaattc aaggttggcg tgaagattct tcttttcag 10200
aacaacttcc agggcatgtt caagaacgtt caggatctct tttacgaca gctctactcg 10260
ggccacgcgc atgttcaac cgcgttcca caaggtcgcc gatcgcatgc gtccaaggcg 10320
tctctactgc gcgaaacagt cggagctcaa ggacaagatc aaggagtctc tcgagtacga 10380
tgagggtccc gtcctcctcg aggttttctt ggacaaggac acgtctgctt tgcctatgtt 10440
ccccgcggc ttcccgctcc accagatgtt cctcgagcct cctaaagcca aggcgccta 10500
agttcttttt tccatggcgg gcgagcgagc gacgcgcga gcgcgcaagt gcgcaagcgc 10560
cttgccttgc ttgcttgcgc ttgccttgc ttgcttgcac acaacctaa gtaagattca 10620
agttttcttg ctgtcgcgcg atgcctgctt gccaccagc cagcctcctg gccgcgcgc 10680
cttgacgctt tcgttccgc cgcgcctc gattcaattc accatccga tacgttccgc 10740
[0031] cccctcagct cgtctgcgc acgacctcg cagaccagc ccaagccaa cgcgcgcgc 10800
agctcagctt gtcagcagc cgcacgtcac atatctcaga tgcatttggc ctgtgaggtt 10860
tattatgcca ctgacgcga acgatcttcg ggtctctcgc tcatgtcgc cgttcgggcc 10920
ctgcaggcgt ggacgcagc cgcgcgcgag acgtctcagc aggcgcctcc gacgcgagcg 10980
ctcgagctcg ccgcgcgcgc gcgatgtctg cctggcgccg actgatctct ggagcgcaag 11040
gaagacacgc cgcgcgcgag aggcgcgaag agagacgcgc ggggatgcag gatataccgc 11100
ggcgcggaac ttgttccgc atacctccc ccttcgagc ttgctctcc ttggcagagc 11160
cgagcggaac cgttccgaa cgcgcgaagg attttggctc tgggtgggtg actcagatcg 11220
agcgcgaggt tctcgcagc ttctgcagc ccgcagtag tcttagaaa tagggagtgc 11280
cggagtcttg acgcgcctta gctcctctc cgcgaagcg cgcctcgcgc ccatgcgcgc 11340
gtcccgctcg tcgtcgctt ggccgcgacc ggtcgcgcca gactacgaca gtgggacaga 11400
gtcgagggc acgcgaatcg ctcgggtgtt aagggttcca aggttcgggc gtcgtcgctt 11460
gccaaagtga aaatagtagg gggggggggg ggtac 11495

<210> 59

<211> 32

<212> PRT

<213> 裂殖壶菌属

<400> 59

Met Arg Thr Val Arg Gly Pro Gln Thr Ala Ala Leu Ala Ala Leu Leu
1 5 10 15

Ala Leu Ala Ala Thr His Val Ala Val Ser Pro Phe Thr Lys Val Glu

	20	25	30
<210>	60		
<211>	96		
<212>	DNA		
<213>	裂殖壶菌属		
<400>	60		
	atggcgacgg tgagggggcc gcaaacggcg gcactcgccg cccttcctggc acitggccgcg	60	
	acgcacgtgg ctgtgagccc gttcaccaag gtggag		96
<210>	61		
<211>	30		
<212>	PRT		
<213>	裂殖壶菌属		
<400>	61		
	Met Gly Arg Leu Ala Lys Ser Leu Val Leu Leu Thr Ala Val Leu Ala		
	1 5 10 15		
	Val Ile Gly Gly Val Arg Ala Glu Glu Asp Lys Ser Glu Ala		
	20 25 30		
<210>	62		
<211>	90		
<212>	DNA		
<213>	裂殖壶菌属		
<400>	62		
	atggcccgcc tcgcgaagtc gcttgigctg ctgacggccg tgetggccgt gatcgagagc	60	
[0032]	gtccggcccg aagaggacaa gtccgagccc		90
<210>	63		
<211>	38		
<212>	PRT		
<213>	裂殖壶菌属		
<400>	63		
	Met Thr Ser Thr Ala Arg Ala Leu Ala Leu Val Arg Ala Leu Val Leu		
	1 5 10 15		
	Ala Leu Ala Val Leu Ala Leu Leu Ala Ser Gln Ser Val Ala Val Asp		
	20 25 30		
	Arg Lys Lys Phe Arg Thr		
	35		
<210>	64		
<211>	114		
<212>	DNA		
<213>	裂殖壶菌属		
<400>	64		
	atgacgtcaa cggcgccgcg gctcgcgctc gtgggtgctt iggtgctcgc tetggctgtc	60	
	ttggcgctgc tagcgagcca aagcgtggcc gtggaccgca aaaagttcag gacc		114
<210>	65		
<211>	34		
<212>	PRT		
<213>	裂殖壶菌属		
<400>	65		

Met Leu Arg Leu Lys Pro Leu Leu Leu Leu Phe Leu Cys Ser Leu Ile
1 5 10 15

Ala Ser Pro Val Val Ala Trp Ala Arg Gly Gly Glu Gly Pro Ser Thr
20 25 30

Ser Glu

<210> 66
<211> 102
<212> DNA
<213> 裂殖壶菌属

<400> 66
atgttgcggc tcaagccact tttaactctc ttctctgtgt cgttgattgc ttgcctgtg 60
gttgcctggg caagaggagg agaaggccg tccacgagcg aa 102

<210> 67
<211> 29
<212> PRT
<213> 裂殖壶菌属

<400> 67

Met Ala Lys Ile Leu Arg Ser Leu Leu Leu Ala Ala Val Leu Val Val
1 5 10 15

Thr Pro Gln Ser Leu Arg Ala His Ser Thr Arg Asp Ala
20 25

[0033]

<210> 68
<211> 87
<212> DNA
<213> 裂殖壶菌属

<400> 68
atggccaaga tcttgcgcag ttgtctctg gcggccgtgc tcttggtgac tctcaatca 60
ctgcgtgtctc attcgacgcg ggaagca 87

<210> 69
<211> 36
<212> PRT
<213> 裂殖壶菌属

<400> 69

Met Val Phe Arg Arg Val Pro Trp His Gly Ala Ala Thr Leu Ala Ala
1 5 10 15

Leu Val Val Ala Cys Ala Thr Cys Leu Gly Leu Gly Leu Asp Ser Glu
20 25 30

Glu Ala Thr Tyr
35

<210> 70
<211> 108
<212> DNA
<213> 裂殖壶菌属

<400> 70
atgctgttgc ggcgcgtgcc atggcacgcg gcggcgacgc tggcgccctt ggtcgtggcc 60
tgcgcgaagt gtttaggcct gggactggac tcggaggagg ccacgtac 108

<210> 71
 <211> 30
 <212> PRT
 <213> 裂殖壶菌属

<400> 71

Met Thr Ala Asn Ser Val Lys Ile Ser Ile Val Ala Val Leu Val Ala
 1 5 10 15

Ala Leu Ala Trp Glu Thr Cys Ala Lys Ala Asn Tyr Gln Trp
 20 25 30

<210> 72
 <211> 90
 <212> DNA
 <213> 裂殖壶菌属

<400> 72

atgacagcta actcggtgaa aataagcacc gtggctgtgc tggtegcgc actggttgg 60

gaaacatgg caaaagctaa ctatcagtgg 90

<210> 73
 <211> 35
 <212> PRT
 <213> 裂殖壶菌属

<400> 73

Met Ala Arg Arg Ala Ser Arg Leu Gly Ala Ala Val Val Val Val Leu
 1 5 10 15

[0034]

Val Val Val Ala Ser Ala Cys Cys Trp Gln Ala Ala Ala Asp Val Val
 20 25 30

Asp Ala Gln
 35

<210> 74
 <211> 105
 <212> DNA
 <213> 裂殖壶菌属

<400> 74

atggcgcca gggtgcagc cctcgggccc ggcgtcgtc tgcctcctgt cgtcgtccc 60

tccgctgtc gtggcaagc cgtcgggac gtgtggagc cgcag 105

<210> 75
 <211> 1785
 <212> DNA
 <213> 人工序列

<220>

<223> 密码子优化的核酸序列

<400> 75

atgaagtgc cgacctcgt cgcaattttg cttgtggcca acatagccac cgcctcgcg 60

gcctccccct cgaatcagac ccgtgcctcc gtgtcattg attacaaagt cgtctctct 120

aacctctcca cctccccgaa cggcagcctc tttagaacct ggctcctcg cgcctcagtt 180

cctccccctc acgttcagat tggcgatccc tgcctccact acaccgatcc ctgcactggc 240

ctcttttcag tcgcttttct ccacgatggc tccggcattt cctcgcctac tactgaagac 300

	ctcgtacct acaaggatct caaccagggc aaccaggta cgtccccgg cggataaac	360
	gaacctgtcg ctgttttcca cggctccgic attccttcg gcattaacgg cctccctaac	420
	ctctctaca cctcgtcag ctacctcccc attcactggt cctccctata cccccgggt	480
	tcggagacgc agagcctggc tgtctccagc gatggtggt ccaactttac taagctcgac	540
	cagggtcccg ttattctcgg ccccccttt gctacaaag tcaccgcctt cggcgacccc	600
	taagtctttc agaacccac cctcgactcc cctctccact ccaagaacaa cactggtaac	660
	accgtcattt cgggtggcct ccacggcaag ggcctccccc agtttcttta cgtcagtaac	720
	gaacctgaat ttcagtaact ggagttcttc ggccagtggg ggcacgagcc taccaactcc	780
	accctgggca accgcacctg ggccggccgc tgggccttca acttcgagac cggcaacgtc	840
	ttttcgttg acgagtaagg ctacaacccc caaggccaga tcttctccac cattggcaac	900
	gagggtcccg accagccctg tgtccccag ctaactcca tccagatat gctttgggtc	960
	tcaggtaacg ttccgcgcaa cggatcggtt tcttccac ccaacatggc cggtctcttc	1020
	gaatggggtt tctgtctca cgcgcgcgag ggttaaggta tctcttccac gtcgtccccc	1080
	tcaccaagt cgggtcccc cgatecttc atttctgacg tttggtcttc cggcgacctc	1140
	tttgagcagg ctgagggctt tctacacac cagcagaact ggacgggac cctctcttc	1200
	ccccgtgagc tccggtctct ttacatcccc aacgtggttg ataagccct tgcgcgcgag	1260
	tcgggcctt cctggcagg cgtctctcc gatagctcg cggtaactgt ggagctccag	1320
	accctcgcaa ttccatcgc cgcgagacc aggcgcgc cctcttccgg caactcgttc	1380
	actagctccg accgcactct taactctcc ggctctgttc cctttaaggg tccccctcc	1440
[0035]	gagaagtttt tctctcttc cgcacagctc tcttccccc cctcgcctcg cggctcgggc	1500
	cctagctccg gcttccagat tcttctccc gactcagat ccaacagggt ctactaccg	1560
	tttagcaacg agtccatcat cgtcgaccgc agcaacacca ggcgcgcgc cgtactaac	1620
	gaaggtaacg actctccgc cagggccgga aagctccgc tctttgagct cctcaacggc	1680
	ggcgagcagg ctattgagac cctcgacctt accctcgtcg ttgataatc cgtgtcgag	1740
	atttacgcca accgtctgtt cgcctttcc accctgggtc gctaa	1785
	<210> 76	
	<211> 1682	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 密码子优化的 HA	
	<400> 76	
	atgaaggcta acctctctgt tcttttttcc gctctcgtg ctgggatgc egacaccatc	60
	tgcattggtt accaagctaa caacagcagc gacaccgtg atactgtctt ggagaagaac	120
	gttacccgac ccattcggtc aacctcttgg aggacagcca caacggcaag ctctgcgttc	180
	ttaaggcat cgcctccctc cagctcgcaa agtgcacat cgcgggtgg cctctcgcaa	240
	accgggagt cgateccctc tccccgttc cctctggctg tacattgttg agactccgaa	300
	cagcgagaac ggtatctgtt acccgggca ttttatcgac tacgaggagc tcccgagca	360
	gtctctctcc gtgtccagct tgagcgttcc gagatttttc cgaaggagtc ctgttgccc	420
	aaccacaaca ccaacggcgt caccgcgcgc tgcctccacg agggcaagtc gactttttac	480
	cgcacccgct ttggctcacc cgagaagggg gttcttacc taagctcaag aactcgtacg	540

teaacaagaa ggcaaggag gtctctgtcc tctggggcat ccaccatccc ccgaacagca	600
aggagcagca gaacatctac cagaacgaga acgccacgtt tcggtaggta cgtcgaacta	660
caacgcgcgc ttactctctg agatcgccga gcgcgcccaag gtgcgcgacc aggcctggccg	720
catgaactac tactggaccc tcttaagcc cggtgacacg atatctttga ggccaacggc	780
aaccttatcg cgcccatgta cgcgttcgcc ctctcccgcg gctttggtag cggcatcatt	840
accagcaacg ccagcatgca cgagtgaac acgaagtgcc agaccocgcc ggtgccatca	900
acagcagcct gccttaccag aacatccacc ccgtaccat cggtgagtgc ccgaagtacg	960
tgcgtcggc caagctcgc atggtaacgg gcctccgcaa cactctctcg atcagcccg	1020
cggcctcttc ggcccatgtt ccggtttcat cgaggcgccg tggacggcca tgatecagcg	1080
ctggtaacgc taccaccacc agaacgagca gggtcccggt tacgcccgcg accagaagtc	1140
caccagaacg ccacaaagg cattactaac aaggtaaca cggtcctega gaagatgaac	1200
attcagitta ccgtctcgg caaggagttc aacaagctgg agaagcgc at ggagaacctc	1260
aacaagaagg ggacgatggt ttcccgaca ttgggacctt caacgccgag ctctctgtgc	1320
tcttgagaa cgagcgtacc ctgacttccc acgactccaa cgtaagaac ctctacgaga	1380
aggtaagtc gcagctcaga acaacgccaa ggagattggc aacggttgc tgcagittta	1440
ccacaagtc gacaacgagt gcattggatc cgtccgcaac ggcaacctac actaccgaa	1500
gtactccgag gagicgaagc tgaacgcgag aaggtaggac gctgaagct ggagtcctg	1560
ggcatctacc agatctctgc cattactcgc acggttgcc tgcgtctcgt cctctctgtc	1620
tcctctgggtg cgatttcgtt ctggatgtgc tgaacggcag ccttcagtc cgcattctga	1680
tc	1682

[0036]

<210> 77
 <211> 565
 <212> PRT
 <213> 甲型流感病毒

<400> 77

Met Lys Ala Asn Leu Leu Val Leu Leu Ser Ala Leu Ala Ala Asp
 1 5 10 15

Ala Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
 20 25 30

Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn
 35 40 45

Leu Leu Glu Asp Ser His Asn Gly Lys Leu Cys Arg Leu Lys Gly Ile
 50 55 60

Ala Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu Gly
 65 70 75 80

Asn Pro Glu Cys Asp Pro Leu Leu Pro Val Arg Ser Trp Ser Tyr Ile
 85 90 95

Val Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp Phe
 100 105 110

	Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe	115	120	125
	Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn	130	135	140
	Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe	145	150	155
	Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys	165	170	175
	Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu	180	185	190
	Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Ile Tyr	195	200	205
	Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg	210	215	220
	Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala	225	230	235
	Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile	245	250	255
[0037]	Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala	260	265	270
	Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Ser Met	275	280	285
	His Glu Cys Asn Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn Ser	290	295	300
	Ser Leu Pro Tyr Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys Pro	305	310	315
	Lys Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn	325	330	335
	Thr Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe	340	345	350
	Ile Glu Gly Gly Trp Thr Gly Met Ile Asp Gly Trp Tyr Gly Tyr His	355	360	365
	His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr	370	375	380
	Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Thr Val Ile Glu	385	390	395
	Lys Met Asn Ile Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu	405	410	415

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

[0038]

<210> 78
 <211> 51
 <212> PRT
 <213> 裂殖壶菌属
 <220>
 <223> GlcNac-转移酶-I 样蛋白
 <400> 78
 Met Arg Gly Pro Gly Met Val Gly Leu Ser Arg Val Asp Arg Glu His
 1 5 10 15
 Leu Arg Arg Arg Gln Gln Gln Ala Ala Ser Glu Trp Arg Arg Trp Gly
 20 25 30
 Phe Phe Val Ala Thr Ala Val Val Leu Leu Val Phe Leu Thr Val Tyr
 35 40 45
 Pro Asn Val
 50
 <210> 79
 <211> 153
 <212> DNA
 <213> 裂殖壶菌属
 <220>
 <223> 信号锚序列
 <400> 79
 atgcgcggcc cggcgaatggt cggcctcagc cggctggacc gcgagcaact ggcggggggg 60

cagcagcagg cggcgagcga atggcgggcgc tgggggttct tcgtcgcgac ggccgctgta 120
ctgctcgtct ttctcacctg ataccggaac gta 153

<210> 80

<211> 66

<212> PRT

<213> 裂殖壶菌属

<220>

<223> β -1,2- 木糖基转移酶样蛋白

<400> 80

Met Arg Thr Arg Gly Ala Ala Tyr Val Arg Pro Gly Gln His Glu Ala
1 5 10 15

Lys Ala Leu Ser Ser Arg Ser Ser Asp Glu Gly Tyr Thr Thr Val Asn
20 25 30

Val Val Arg Thr Lys Arg Lys Arg Thr Thr Val Ala Ala Leu Val Ala
35 40 45

Ala Ala Leu Leu Val Thr Gly Phe Ile Val Val Val Val Phe Val Val
50 55 60

Val Val
65

[0039]

<210> 81

<211> 198

<212> DNA

<213> 裂殖壶菌属

<220>

<223> 信号锚序列

<400> 81

atgcgcacgc ggggcgcggc gtacgtcgg ccgggacagc acgaggcgaa ggcgtctctg 60

tcaaggagca gcgacgaggg atatacgacg gtcaacgttg tcaggaccaa gcgaaagagg 120

aacctgttag ccgcgcttgt agccgcggcg ctgcctgtga cggcttttat cgtcgtcgtc 180

gtcttcgtcg tcgttgtt 198

<210> 82

<211> 64

<212> PRT

<213> 裂殖壶菌属

<220>

<223> β -1,4-木糖苷样蛋白

<400> 82

Met Glu Ala Leu Arg Glu Pro Leu Ala Ala Pro Pro Thr Ser Ala Arg
1 5 10 15

Ser Ser Val Pro Ala Pro Leu Ala Lys Glu Glu Gly Glu Glu Glu Asp
20 25 30

Gly Glu Lys Gly Thr Phe Gly Ala Gly Val Leu Gly Val Val Ala Val
35 40 45

Leu Val Ile Val Val Phe Ala Ile Val Ala Gly Gly Gly Gly Asp Ile
50 55 60

<210> 83
<211> 192
<212> DNA
<213> 裂殖壶菌属

<220>
<223> 信号锚序列

<400> 83

atggaggccc tgcgcgagcc ctiggetgcg ccgccaacgt cgcccgagatc gtcggtgcca 60
ggcgcgcctcg cgaaggagga gggcgaggag gaggacggcg aaaaagggac glltggggcg 120
ggggtcctcg gtgtcgtggc ggtgcctcgc atcgtggtgt ttgcgacgt ggcgggagcc 180
ggagggcgata tt 192

<210> 84
<211> 73
<212> PRT
<213> 裂殖壶菌属

<220>
<223> 半乳糖基转移酶样蛋白

<400> 84

Met Leu Ser Val Ala Gln Val Ala Gly Ser Ala His Ser Arg Pro Arg
1 5 10 15

[0040]

Arg Gly Gly Glu Arg Met Gln Asp Val Leu Ala Leu Glu Glu Ser Ser
20 25 30

Arg Asp Arg Lys Arg Ala Thr Ala Arg Pro Gly Leu Tyr Arg Ala Leu
35 40 45

Ala Ile Leu Gly Leu Pro Leu Ile Val Phe Ile Val Trp Gln Met Thr
50 55 60

Ser Ser Leu Thr Thr Ala Pro Ser Ala
65 70

<210> 85
<211> 219
<212> DNA
<213> 裂殖壶菌属

<220>
<223> 信号锚序列

<400> 85

atgttgagcg tagcacaagt cgcggggtcg gcccactcgc ggcgcgagcg aggtggtgag 60
cggtatgaag acgtgcigge cctggaggaa agcagcagag atcgaaaacg agcaacagca 120
aggeccgggc tatatgcgc atttgcgatt ctggggctgc cgtcctcgt attcctegta 180
tgccaatga ctactccct cactactgcc ccgagcgcc 219

<210> 86

<211> 997
 <212> PRT
 <213> 裂殖盘菌属

<220>
 <223> EMC1

<400> 86

Met Gly Thr Thr Thr Ala Arg Met Ala Val Ala Val Leu Ala Ala Ala
 1 5 10 15

Val Ser Val Ala His Gly Leu His Glu Asp Gln Ala Gly Val Asn Asp
 20 25 30

Trp Thr Val Arg Asn Leu Gly Ala Tyr Ala His Gly Val Phe Leu Asp
 35 40 45

Asp Asp Leu Ala Leu Val Ala Thr Thr Gln Ala Thr Val Gly Ala Val
 50 55 60

Arg Met Thr Asp Gly Glu Val Val Trp Arg Glu Thr Leu Pro Thr Ala
 65 70 75 80

Arg Ser Ala Pro Leu Ala Ser Gln Val Lys His Glu Leu Phe Ala Thr
 85 90 95

Ala Ser Ala Asp Ala Cys Val Ile Glu Leu Trp Ala Thr Pro Ser Gly
 100 105 110

[0041]

Asp Val Met Thr Ser Asp Ser Arg Gln Ala Gly Leu Glu Trp Asp Ala
 115 120 125

Lys Ile Cys Asp Asn Thr Asp Ala Asp Ala Thr Gly Val Leu Glu Leu
 130 135 140

Leu Asp Asn Asp Phe Asn Asn Asp Gly Thr Pro Asp Val Ala Ala Leu
 145 150 155 160

Thr Pro Phe Gln Phe Val Ile Leu Asp Gly Val Ser Gly Arg Val Leu
 165 170 175

His Glu Val Asp Leu Asp Lys Thr Ile Ala Trp Gln Gly Leu Val Glu
 180 185 190

Ala Ala Gly Ser Ala Thr Gly Gly Lys Arg Lys Arg Pro Ser Ile Met
 195 200 205

Ala Tyr Gly Val Asp Ile Lys Thr Gly Lys Leu Glu Val Arg Lys Leu
 210 215 220

Ala Asn Ser Gly Ala Thr Leu Asp Pro Val Ser Gly Leu Glu Gly Val
 225 230 235 240

Ser Ala Asp Glu Ile Thr Val Leu Lys Ser Gly Val Ala Lys Val Gly
 245 250 255

Ser Ala Leu Leu Phe Val Arg Lys Glu Ser Gly Ala Leu Val Ala Phe
 260 265 270

Asp Cys Val Ala Asn Gln Leu Gln Glu Leu Thr Asn Ala Pro Ser Ile
 275 280 285
 Lys Gly Ser Val Gln Ser Leu Gly Ser Ala Arg Phe Phe Ala Thr Asp
 290 295 300
 Ala Gly Val Ile Tyr Ala Val Asp Gly Glu Leu Lys Ile Ala Glu Thr
 305 310 315 320
 Leu Lys Gly Val Glu Ala Ala Ala Ile Gly Val Ser Gly Ala Ser Val
 325 330 335
 Ile Ala Ala Val Gln Ser Ser Thr Ala Ser Gly Thr Gly Asp Glu Ala
 340 345 350
 Gln Cys Gly Pro Ile Ser Arg Val Leu Val Gln Ser Ala Ser Gly Val
 355 360 365
 Thr Glu Ile Ala Phe Pro Gln Gln Gln Gly Gln Ser Gly Ala Arg Gly
 370 375 380
 Leu Val Glu Lys Ile Ile Val Gly Asp Ser Ser Thr Gly Thr Arg Ala
 385 390 395 400
 Ile Phe Val Phe Glu Asp Ala Ser Ala Val Gly Ile Glu Ile Glu Ser
 405 410 415
 [0042] Gly Ala Ser Glu Ala Ser Thr Leu Phe Val Arg Glu Glu Ala Leu Ala
 420 425 430
 Asn Val Val Glu Ala Val Ala Val Asp Leu Pro Pro Thr Asp Glu Val
 435 440 445
 Gly Ser Leu Gly Asp Glu Ala Ala His Val Phe Ala His Gly Ser His
 450 455 460
 Ala Ser Ile Phe Met Phe Arg Leu Lys Asp Gln Val Arg Thr Val Gln
 465 470 475 480
 Arg Phe Val Gln Ser Leu Phe Gly Ala Ala Thr Gln His Leu Ser Glu
 485 490 495
 Phe Val Ala Ser Gln Gly Lys Thr Leu Val Gln Ala Ile Arg Gly Glu
 500 505 510
 Leu Pro Arg Ala Glu Ser Leu Ser Gln Ser Glu Met Phe Ser Phe Gly
 515 520 525
 Phe Arg Arg Val Leu Val Leu Arg Ser Ala Ser Gly Lys Val Phe Gly
 530 535 540
 Leu Asn Ser Ala Asp Gly Ser Leu Leu Trp Ala Ala Gln Ser Pro Gly
 545 550 555 560
 Ser Arg Leu Phe Val Thr Arg Ala Arg Glu Ala Gly Leu Asp His Pro
 565 570 575

Ala Glu Val Ala Ile Val Asp Glu Ala His Gly Arg Val Thr Trp Arg
 580 585 590
 Asn Ala Ile Thr Gly Ala Val Thr Arg Val Glu Asp Ile Asp Thr Pro
 595 600 605
 Leu Ala Gln Ile Ala Val Leu Pro Gly Asp Ile Phe Pro Ser Thr Ala
 610 615 620
 Ser Ser Glu Glu Asp Val Ser Pro Ala Ala Val Leu Ile Ala Leu Asp
 625 630 635 640
 His Ala Gln Arg Val His Ile Leu Pro Ser Ser Arg Thr Glu Ser Val
 645 650 655
 Leu Gln Leu Glu Asp Leu Leu Arg Ala Leu His Phe Val Val Tyr Ser
 660 665 670
 Asn Glu Thr Gly Ala Leu Thr Gly Tyr Ala Val Asp Pro Ser Gln Arg
 675 680 685
 Ala Gly Val Glu Leu Trp Ser Met Ile Val Pro Ala Ser Gln Thr Leu
 690 695 700
 Leu Ala Val Glu Gly Gln Ser Gly Gly Ala Leu Asn Asn Pro Gly Ile
 705 710 715 720
 Lys Arg Gly Asp Gly Ala Val Leu Val Lys Phe Val Asp Pro His Leu
 725 730 735
 Leu Met Val Ala Thr Gln Ser Gly Pro His Leu Gln Val Ser Ile Leu
 740 745 750
 Asn Gly Ile Ser Gly Arg Val Ile Ser Arg Phe Thr His Lys Lys Ser
 755 760 765
 Thr Gly Pro Val His Ala Val Leu Ala Asp Asn Thr Val Thr Tyr Ser
 770 775 780
 Phe Trp Asn Gln Val Lys Ser Arg Gln Glu Val Ser Val Val Gly Leu
 785 790 795 800
 Phe Glu Gly Glu Ile Gly Pro Arg Glu Leu Asn Met Trp Ser Ser Arg
 805 810 815
 Pro Asn Met Gly Ser Gly Lys Ala Met Ser Ala Phe Asp Asp Ser Met
 820 825 830
 Met Pro Asn Val Gln Gln Lys Thr Phe Tyr Thr Glu Arg Ala Ile Ala
 835 840 845
 Ala Leu Gly Val Thr Lys Thr Arg Phe Gly Ile Ala Asp Arg Arg Val
 850 855 860
 Leu Ile Gly Thr Ala Asn Gly Ala Val Asn Met Gln Val Pro Gln Ile
 865 870 875 880
 Leu Ser Pro Arg Arg Pro Val Gly Lys Leu Ser Asp Met Glu Lys Glu

[0043]

885	890	895
Glu Gly Leu Met Leu Tyr Ala Pro	Glu Leu Pro Leu Ile Pro Thr Gln	
900	905	910
Thr Ile Thr Tyr Tyr Glu Ser Ile Pro Gln Leu Arg Leu Ile Arg Ser		
915	920	925
Phe Ala Thr Arg Leu Glu Ser Thr Ser Leu Val Leu Ala Ala Gly Leu		
930	935	940
Asp Ile Phe Tyr Thr Arg Val Met Pro Ser Arg Gly Phe Asp Val Leu		
945	950	955
Asp Glu Asp Phe Ala Ser Gly Leu Leu Leu Ala Leu Ile Ala Ala Leu		
965	970	975
Leu Ala Leu Thr Ile Tyr Leu Ser Lys Ala Val Gly Lys Ser Thr Leu		
980	985	990
Asp Glu Thr Trp Lys		
995		

<210> 87
 <211> 731
 <212> PRT
 <213> 裂殖壶菌属

[0044]

<220>
 <223> Nicastrin 样

<400> 87

Met Gly Ala Ala Arg Arg Ser Met Gly Ala Ala Arg Lys Ala Leu Ala		
1	5	10
Ala Ser Ala Thr Leu Ala Ala Leu Ala Leu Ala Gly Leu Gln Pro Ala		
20	25	30
Arg Ala Glu Val Asn Gly Val Asn Ala Met Thr Glu Ala Met Leu Thr		
35	40	45
Glu Tyr Ala Ser Leu Pro Cys Val Arg Ser Ile Ala Arg Asp Gly Ala		
50	55	60
Val Gly Cys Gly Ser Pro Ser Asp Arg Ser Val Ala Glu Gly Gly Ala		
65	70	75
Leu Phe Leu Val Glu Ser Val Glu Asp Val Thr Gly Leu Ile Glu Asn		
85	90	95
Ala Gln Gly Leu Asp Ala Val Ala Leu Val Val Asp Asp Ala Leu Leu		
100	105	110
His Gly Asp Ser Leu Arg Ala Met Gln Asp Leu Ala Lys Lys Ile Arg		
115	120	125
Val Thr Ala Val Ile Val Thr Val Glu Glu Asp Gly Ser Pro Gln Glu		
130	135	140

	Pro	Pro	Arg	Ser	Ser	Ala	Ala	Pro	Thr	Thr	Trp	Ile	Pro	Ser	Gly	Asp
	145					150					155					160
	Gly	Leu	Leu	Asn	Glu	Thr	Val	Ser	Phe	Val	Val	Thr	Arg	Leu	Arg	Asn
				165						170					175	
	Ala	Thr	Gln	Ser	Glu	Glu	Ile	Arg	Ala	Leu	Ala	Ala	Ser	Asn	Arg	Asp
				180					185						190	
	Arg	Gly	Tyr	Val	Asp	Ala	Val	Phe	Gln	His	Ser	Ala	Arg	Tyr	Gln	Phe
		195						200					205			
	Tyr	Leu	Gly	Lys	Glu	Thr	Ala	Thr	Ser	Leu	Ser	Cys	Leu	Ala	Ser	Gly
	210					215						220				
	Arg	Cys	Asp	Pro	Leu	Gly	Gly	Leu	Ser	Val	Trp	Ala	Ser	Ala	Gly	Pro
	225				230						235					240
	Val	Pro	Val	Asn	Ser	Ala	Lys	Glu	Thr	Val	Leu	Leu	Thr	Ala	Asn	Leu
				245						250					255	
	Asp	Ala	Ala	Ser	Phe	Phe	His	Asp	Val	Val	Pro	Ala	Arg	Asp	Thr	Thr
				260					265						270	
	Ala	Ser	Gly	Val	Ala	Ala	Val	Leu	Leu	Ala	Ala	Lys	Ala	Leu	Ala	Ser
		275						280					285			
[0045]	Val	Asp	Glu	Ser	Val	Leu	Glu	Ala	Leu	Ser	Lys	Gln	Ile	Ala	Val	Ala
	290					295						300				
	Leu	Phe	Asn	Gly	Glu	Val	Trp	Ser	Arg	Ala	Gly	Ser	Arg	Arg	Phe	Val
	305				310						315					320
	His	Asp	Val	Ala	Leu	Gly	Glu	Cys	Leu	Ser	Pro	Gln	Thr	Ala	Ser	Pro
				325						330					335	
	Tyr	Asn	Glu	Ser	Thr	Cys	Ala	Asn	Pro	Pro	Val	Tyr	Ala	Leu	Ala	Trp
		340							345					350		
	Thr	Ser	Leu	Gly	Leu	Asp	Asn	Ile	Thr	Asp	Val	Val	Ser	Val	Asn	Asn
		355					360						365			
	Val	Ala	Gly	Ser	Glu	Ser	Gly	Ala	Phe	Tyr	Val	His	Thr	Ala	Ala	Gly
		370					375					380				
	Thr	Ala	Ser	Ala	Asn	Ala	Ala	Ala	Ala	Leu	Gln	Ser	Val	Ala	Ser	Ser
	385				390						395					400
	Ser	Thr	Asp	Val	Asp	Val	Ser	Ile	Thr	Gly	Ala	Thr	Thr	Ser	Gly	Val
				405						410					415	
	Val	Pro	Pro	Ser	Pro	Leu	Asp	Ser	Phe	Leu	Ala	Ala	Glu	Met	Glu	Thr
				420					425					430		
	Asp	Val	Ser	Phe	Ser	Gly	Ala	Gly	Leu	Val	Val	Ser	Gly	Phe	Asp	Ala
		435					440						445			

Ala Ile Thr Asp Ala Asn Pro Arg Tyr Ser Ser Arg Tyr Asp Arg Arg
 450 455 460
 Asp Lys Gly Pro Glu Ala Asp Asp Ala Glu Ala Leu Thr Ala Ala Arg
 465 470 475 480
 Ile Ala Asp Val Ala Thr Leu Leu Ala Arg His Ala Phe Val Gln Ala
 485 490 495
 Gly Gly Ser Ile Ser Asp Ala Val Asn Phe Val Leu Val Asp Gly Thr
 500 505 510
 His Ala Ala Glu Leu Trp Asp Cys Leu Thr Lys Asp Phe Ala Cys Thr
 515 520 525
 Leu Val Ala Asp Val Ile Gly Ala Glu Asp Thr Thr Ala Val Ala Asp
 530 535 540
 Phe Met Gly Ser Thr Leu Leu Ala Ala Ser Glu Gly Val Ala Gly Gly
 545 550 555 560
 Ala Pro Asn Phe Phe Ser Gly Ile Tyr Ser Pro Phe Pro Val Glu Asn
 565 570 575
 Asn Val Met Arg Pro Val Pro Leu Phe Val Arg Asp Tyr Leu Ala Gln
 580 585 590
 Tyr Gly Arg Asn Ala Ser Leu Ile Glu Lys Val Thr Glu Ser Ala Lys
 595 600 605
 Tyr Ala Cys Ala Gln Asp Leu Asp Cys Met Val Met Thr Glu Pro Pro
 610 615 620
 Ala Cys Glu Leu Gly Arg Ser Ala Leu Ala Cys Leu Arg Gly Gly Cys
 625 630 635 640
 Val Cys Ser Asn Ala Tyr Phe His Asp Ala Val Ser Pro Ala Leu Val
 645 650 655
 Tyr Glu Asp Gly Ala Phe Ser Val Asp Ala Gln Lys Leu Thr Asp Asp
 660 665 670
 Asp Gly Leu Trp Thr Glu Pro Arg Trp Ser Asp Gly Thr Leu Thr Leu
 675 680 685
 Tyr Thr Ser Ala Asn Ser Ala Ser Thr Thr Ile Ala Leu Leu Val Cys
 690 695 700
 Gly Ile Leu Leu Thr Ile Gly Cys Val Phe Ala Leu Arg Lys Ala Gln
 705 710 715 720
 Gly Met Leu Asp Asn Thr Lys Tyr Lys Leu Asn
 725 730

[0046]

<210> 88
 <211> 232
 <212> PRT

<213> 裂殖壶菌属

<220>

<223> Emp24

<400> 88

Met Ala Thr Thr Glu Asn Glu Ala Arg Leu Pro Pro Gly Lys Gln Arg
1 5 10 15

Leu Gly Arg Arg Arg Arg Gly Arg Val Ser Lys Ala Ser Gly Trp Gly
20 25 30

Thr Thr Leu Ala Leu Ala Ala Ala Val Leu Val Phe Ser Val Asp Arg
35 40 45

Ala Ser Gly Val Arg Phe Glu Val Ala Ser Thr Glu Glu Arg Cys Ile
50 55 60

Phe Asp Val Leu Arg Lys Asp Gln Leu Val Thr Gly Glu Phe Glu Val
65 70 75 80

His Ala Asp Gly Asp Asp Val Asn Met Asp Ile His Val Thr Gly Pro
85 90 95

Leu Gly Glu Glu Val Phe Ser Lys Gln Asn Ser Lys Met Ala Lys Phe
100 105 110

[0047]

Gly Phe Thr Ala Glu Ala Ala Gly Glu His Val Leu Cys Leu Arg Asn
115 120 125

Asn Asp Met Ile Met Arg Glu Val Gln Val Lys Leu Arg Ser Gly Val
130 135 140

Glu Ala Lys Asp Leu Thr Glu Val Val Gln Arg His His Leu Lys Pro
145 150 155 160

Leu Ser Ala Glu Val Ile Arg Ile Gln Glu Thr Ile Arg Asp Val Arg
165 170 175

His Glu Leu Thr Ala Leu Lys Gln Arg Glu Ala Glu Met Arg Asp Met
180 185 190

Asn Glu Ser Ile Asn Thr Arg Val Ser Leu Phe Ser Phe Phe Ser Ile
195 200 205

Ala Val Val Gly Ser Leu Gly Ala Trp Gln Ile Met Tyr Leu Lys Ser
210 215 220

Tyr Phe Gln Arg Lys Lys Leu Ile
225 230

<210> 89

<211> 550

<212> PRT

<213> 裂殖壶菌属

<220>

<223> Calnexin 样

<400> 89

Met Arg Thr Thr Phe Val Ala Ala Tyr Ala Ala Val Ala Ala Leu Ala
1 5 10 15

Leu Gly Gln Cys Glu Ala Ile Asn Phe Arg Glu Ser Phe Glu Gly Ala
20 25 30

Asn Val Glu Lys Glu Trp Val Lys Ser Ala Ser Asp Arg Tyr Ala Gly
35 40 45

Ser Glu Trp Ala Phe Asp Thr Ser Lys Asp Thr Gly Asp Val Gly Leu
50 55 60

Gln Thr Val Lys Pro His Lys Phe Tyr Gly Ile Ser Arg Lys Phe Glu
65 70 75 80

Asn Pro Ile Pro Val Gly Asp Gly Glu Lys Pro Phe Val Ala Gln Tyr
85 90 95

Glu Val Lys Phe Thr Glu Gly Val Ser Cys Ser Gly Ala Tyr Leu Lys
100 105 110

Leu Leu Glu Gln Asp Asp Ala Phe Thr Pro Lys Asp Leu Val Glu Ser
115 120 125

[0048]

Ser Pro Tyr Ser Ile Met Phe Gly Pro Asp Asn Cys Gly Ala Asn Asn
130 135 140

Lys Val His Leu Ile Phe Arg Gln Glu Asn Pro Val Thr Lys Glu Tyr
145 150 155 160

Glu Glu Lys His Met Thr Lys Lys Val Thr Ser Val Arg Asp Arg Thr
165 170 175

Ser His Val Tyr Thr Leu Glu Val His Pro Asp Asn Thr Phe Lys Val
180 185 190

Lys Val Asp Gly Lys Val Glu Ala Glu Gly Ser Leu Thr Asp Asp Glu
195 200 205

Ala Phe Ser Pro Pro Phe Gln Gln Pro Lys Glu Ile Asp Asp Pro Asn
210 215 220

Asp Glu Lys Pro Asp Asp Trp Val Asp Gln Ala Lys Ile Pro Asp Pro
225 230 235 240

Glu Ala Ser Lys Pro Asp Asp Trp Asp Glu Asp Ala Pro Lys Arg Ile
245 250 255

Ala Asp Pro Asp Ala Val Lys Pro Glu Gly Trp Leu Asp Asp Glu Pro
260 265 270

Asp Gln Val Pro Asp Pro Ala Ala Ser Glu Pro Glu Asp Trp Asp Glu
275 280 285

Glu Asp Asp Gly Ile Trp Glu Ala Pro Leu Val Ala Asn Pro Lys Cys
 290 295 300
 Thr Ala Gly Pro Gly Cys Gly Glu Trp Asn Ala Pro Met Ile Glu Asn
 305 310 315 320
 Pro Asn Tyr Lys Gly Lys Trp Ser Ala Pro Met Ile Asp Asn Pro Glu
 325 330 335
 Tyr Lys Gly Val Trp Lys Pro Arg Arg Ile Glu Asn Pro Ala Tyr Phe
 340 345 350
 Glu Glu Ser Ser Pro Val Thr Thr Ile Lys Pro Ile Gly Ala Val Ala
 355 360 365
 Ile Glu Ile Leu Ala Asn Asp Lys Gly Ile Arg Phe Asp Asn Ile Ile
 370 375 380
 Ile Gly Asn Asp Val Lys Glu Ala Ala Glu Phe Ile Asp Lys Glu Phe
 385 390 395 400
 Leu Ala Lys Gln Ala Asp Glu Lys Ala Lys Val Lys Glu Glu Ala Ala
 405 410 415
 Gln Ala Ala Gln Asn Ser Arg Trp Glu Glu Tyr Lys Lys Gly Ser Ile
 420 425 430
 Gln Gly Tyr Val Met Trp Tyr Ala Gly Asp Tyr Ile Asp Tyr Val Met
 435 440 445
 Glu Leu Tyr Glu Ala Ser Pro Ile Ala Val Gly Val Gly Ala Ala Ala
 450 455 460
 Ala Gly Leu Ala Val Leu Val Ala Leu Met Val Met Cys Met Ser Gly
 465 470 475 480
 Ala Pro Glu Glu Tyr Asp Asp Asp Val Ala Leu His Lys Lys Asp Asp
 485 490 495
 Asp Ala Ala Ala Gly Asp Asp Asp Glu Ala Glu Ala Glu Ala Glu Asn
 500 505 510
 Asp Ala Ala Asp Glu Asp Glu Asp Glu Glu Asp Asp Asp Asp Glu Glu
 515 520 525
 Asp Glu Asp Glu Glu Glu Asp Glu Asp Glu Ala Thr Gly Pro Arg Arg
 530 535 540
 Arg Val Asn Arg Ala Asn
 545 550

[0049]

<210> 90
 <211> 6175
 <212> DNA
 <213> 人工的
 <220>

<223> pCL0121	
<400> 90	
ctcttatctg cctcgcgcgc ttgaecgcgc cttagactctt ggccgtttgce gctcgcaccc	60
tgcctcgcgc ggcagcgccg gcggcgagct gggtgggtcc gcagccttcc gcctcgcgcc	120
gctagctcgc tcgcgcgcgt ctgcagccag cagggcagca ccgcacggca ggcaggtccc	180
ggcgcggatc gatcgatcca tcgatccatc gatccatcca tcgtgcggtc aaaaagaaag	240
gaagaagaaa ggaagaaaga aggcgtgcgc acccgagtcg gcgtcagagc cccgctcgcg	300
gtcccgcgga gcctcgcgt tagtccccc cccgcgcgc gcagtccccc gggagggatc	360
gcgcacctct cgcgcgcgc tcgcgcctcg ccgattcccc gctccctctt ttccgtctct	420
tcgcgcctc cgtcgcgcgc cgcgtcgcgc gcgcgcgcct ccttatctgc tccccagggg	480
ggcactcgc acattttgcg ccgcgcgcgc ccgcgcgcgc ggcgcgcgc ccctggttcc	540
ccccgcgagc ggcgcgcgt cgcgcgcga agactcgcgc cgtgcgcgc gcagcaacgc	600
glggcgccgc cgcgcgcgc ggcgcgcgc ggcgcgcgc aggcgcgcgc aggcgcgcgc	660
taggggaaa gcgcgcgcgc ggcgcgcgc cgcgcgcgc cagagcagtc gccgcgcgc	720
accgcacaac cagagcagc gacgcagga cgtcgcgcgc ggcgcgcgc gcagcgcgc	780
cagggcagc ggcgcgcgc cgcgcgcgc cgcgcgcgc ggcgcgcgc aggcgcagga	840
cgcgcgcgc agcgtgcagc ttccgcgcgc agcgcgcgc gcgcgcgcgc gagagagat	900
aaggcgttgc gatcgcgaga ggcgcgcgc ggcgtggagc gaaaatgggt ggagaggata	960
gtatcttgcg tcgttgagc aggcagctga cgcagagga ggatagctgc atgatgatgt	1020
gcacagagaa gaagcagtc gaagcagct actagcaagc aaggatcca tgaagtctgc	1080
gaactcgcgc gcaattttgc ttgtgccaac catagccacc gccctcgcgc agagcagtcg	1140
ctgcacccc accgaccaga cgtgtgtgag caaggcgagc ggcgtgttca cgggggtggt	1200
gccatcctg gtcgagctgc accgcagct aacgcgcac aagttcagc tgcgcgcgc	1260
gggcgagggc gatgcacct accgcagct gacctgaag ttcatctgca ccaccgcaa	1320
gcgcgcgcgc ccctgcca ccctgcgcgc caccgcgcgc taccgcgcgc agtccttcag	1380
ccctacccc gaccacatga agcagcagc ctctctcaag tcgcacatgc ccgaaggata	1440
cgtcagagc cgcacacct ttctcaagc gcagcgcac tacaagacc gcgcgcaggt	1500
gaagtgcagc ggcgcacccc tggtaacgc catcgagctg aaggcagtc actcaagga	1560
ggcgcgcac atcttgggc aacagctgga gtacaactac aacagcaca acgtctatat	1620
catggccgac aagcagaaga aggcacaa ggtagaactc aagatccgc acaacatga	1680
ggcgcgcgc gcgcgcgcgc ccgcgcgcgc ccgcgcgcgc cgcgcgcgc gcgcgcgcgc	1740
cgtctgctgc ccgcgcac actacctgag caccagtcgc gccctgagca aagacccaa	1800
cgagaagcgc gacacatgc tcctgcgcgc gtcgcgcgc gcgcgcgcgc tcatctcgc	1860
catggacgag ctgtacaagc accacatca ccaccactaa catatgagtt atgagatccg	1920
aaagtgaac ttgtctaac ccgcgcgcgc atggcgagc ggcgcgcgc aagatcgt	1980
attacatagt atttttccc tactctttgt gtttctctt tttttttt tgaacgcatt	2040
caagccactt gtcgtgggtt acttctttgt ttgttctt gtctcttgc ttgcctgctt	2100
cttggtcaga cgcgcacaaa aaggcaaaa attcttctat ggcacagata agaaaaagaa	2160
aaagtctgc gaccacgctc atcagaagc aagagaagag aaacactcgc gctcacattc	2220

[0050]

	tcgctcgcgt aagaatctta gccacgcata cgaagtaatt tgtccatctg gcgaatcttt	2280
	acatgagcgt ttccaagctg gacgctgaga tcataaccitt ctgatcgtta atgttccaac	2340
	cttgcatagg cctcgttgcg atccgctage aatgcgtcgt actcccttig caactgcgcc	2400
	atcgcccat tgtgacgtga gtccagattc ttctcgagac ctccgagcgc tgctaatttc	2460
	gcctgacgt cctttctttg tcttccatg acacgcgcgt tcaccgtgcg ttccacttct	2520
	tcctcagaca tgcccttgge tgcctcgacc tgcctcgtaa aaeggccccc agcacgtgct	2580
	acgagatttc gattccaccg ccgccttcta tgaaggttg ggcttcggaa tegtttccg	2640
	ggacgcgcgc tggatgatcc tcacgcgcgc ggatctcatg ctggagtctt tgcgccccc	2700
	caacttggtt atgcagctt ataatggtta caaataaagc aatagcataa caaatttcaa	2760
	aaataaagca tttttttcac tgcattctag ttgtggtttg tccaaactca tcaatgtatc	2820
	ttatcataca tggtegacct gcaggaaacct gcattaatga atcgcccaac gcgcggggag	2880
	aggcggtttg cgtattgggc gctcttcgc ttccctgcct actgactcgc tgcgctcgtt	2940
	cgttcgctg cgcgcgcgcg tatcagctca ctcaaaaggc gtaatacgtt tateccacaga	3000
	atcaggggat aacgcaggaa agaacatgtg agcaaaagcc cagcaaaagg ccaggaaacc	3060
	taaaaaggcc gcgttctgtg cgtttttcca taggtctcgc ccccttgacg agcatacaaa	3120
	aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accagcggtt	3180
	tcccccctga agctccctcg tgcgctctcc tgttccgacc ctgcgcgtta ccggataacct	3240
	gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctacgctt gtaggtatct	3300
	cagttcggtg taggtcgttc gctccaagct gggcigtgtg caagaacccc ccgttcagcc	3360
[0051]	cgaccgctgc gccttatccg gtaactatcg tcttgagtc aaeccggtta gacacgaatt	3420
	atgcgcacig gcagcagcca ctgglaacag gattagcaga gcgagglaig tagggcgigc	3480
	tacagagttc ttgaagtgtt ggcttaacta cggctacact agaagaacag tatttggtat	3540
	ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa	3600
	acaaaccacc gcigttagcg gtggtttttt tgtttgcaag cagcagattt ccgcagaaaa	3660
	aaaaggaatc caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaaacg	3720
	aaactcaegt taagggattt tggctatgag attatcaaaa aggatcttca cctagatcct	3780
	ttlaaatlaa aaatgaagtt tlaaatcaat claaagtata tatgaglaaa ctgggtctga	3840
	cagttaccaa tgcctaatca gtgaggcacc tatctcagcg atctgtctat ttcttcaatc	3900
	catagttgcc tgactccccc tctgttagat aactacgata cgggagggct taccatctgg	3960
	ccccagtgtt gcaatgatac ccgcagaccc acgtcacccg gtccagattt tatcagcaat	4020
	aaaccagcca gcgggaaggc ccgagcgcag aagtggctct gcaactttat ccgcctcat	4080
	ccagttctat aattgttgcg gggaagctag agtaagtatg tgcacagtta atagtttgcg	4140
	caacgttgtt gccattgcta caggcatcgt ggtgtcacgc tctctgtttg gtaigtcttc	4200
	attcagctcc gggtcccaac gatcaaggcg agttacatga tccccatgt tctgcacaaa	4260
	agcgggttag tctttcggtc ctccgactgt tctcagaagt aagtggccg cagtgttate	4320
	actcatggtt atggcagcac tgcataattc tcttacgtc atgccatccg taagatgctt	4380
	ttctgtgaat ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag	4440
	ttgtctctgc ccggcgtcaa tacgggataa taaccgcgca catagcagaa ctttaaaagt	4500
	gtcctacatt ggaaaacgtt ctctggggcg aaaactctca aggatcttac ccctgttgag	4560

[0052]

```

atcagttcg afgtaaccca ctctgcacc caactgacct tcagcatctt ttactttcac 4620
cagcgtttct gggtagacaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc 4680
gacacggaaa tgitgaatac tcatactctt cctttttcaa tattattgaa gcattttatca 4740
gggtttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 4800
ggttccgcgc acatttcccc gaaaagtgc acctgacgtc taagaaacca ttattatcat 4860
gacattaaac tataaaaaata ggcgtatcac gaggccttt cgtctcgcgc gtttcgggta 4920
tgacggtgaa aaacctctgac acatgcagct cccggagacg gtcacagctt gtcgtgaagc 4980
ggatgccggg agcagacaag cccgtcaggc cgcgtcagcg ggtgttagcg ggtgtcgggg 5040
ctggcttaac tatggcgcat cagagcagat tgtactgaga gtgcaccaag ctccaattt 5100
taggcccccc acigaccgag gtcgtcgtat aaaccacitt tccatlgatt ticcaggltt 5160
cgtttaactca tgccactgag caaaacttcg gtcttttcta acaaaagctc tccctacaaa 5220
gcattgcgcg gcaacggacg tgcctcata ctccactgcc acacaaggtc gataaactaa 5280
gtctctcaca aatagaggag aattccactg acaactgaaa acaatgtatg agagacgac 5340
accactggag cgggcgcggc gttggcgcg gaggtcgcca gcaaaaaaca ggcactcgcc 5400
gagcaaaccc gaatcagcct tcagacggtc gtgcctaaca acacgcggtt ctaccccgcc 5460
ttcttcgcgc ccttctcgct ccaagcatcc ttcaagttaa tctctctagt tcaacttcaa 5520
gaagaacaa accaccaaca ccatggccaa gttgaccagt gccgttcggg tgcacacgc 5580
gcgcgacgtc gccggagcgg tcagattctg gaccgacgg ctggggttct cccgggaact 5640
cgiggaggac gaatttcgcg gtgtggtcgc ggacgacgtg acctgttca tcagcgcggt 5700
ccaggaccag gtgtgcgcg acaacacct ggctcgggtg tgggtgcgcg gccgtggaca 5760
gtctacgcc gagtggtcgg aggtcgtgtc caggaaactt cgggacgctt cgggcgcgc 5820
catgaccgag atcggcgagc agccgtgggg ggggagttc gccctgcgcg acccggcgcg 5880
caactgcgtg caattctgtg ccgaggagca ggactgacac gtgctacgag atttcgatc 5940
caccgcgcgc ttctatgaaa ggttgggctt cggatcgtt ttccgggacg ccggtgggat 6000
gatcctccag cgggggacat tcatgtgga gttcttgcgc caccacaact tgtttatgc 6060
agcttataat ggttacaaat aaagcaatag caccacaaat ttacaaata aagcattttt 6120
ttactgcat tctagttgtg gtttgtcaa acitacaaat gtatcttate ggtac 6175

```

<210> 91
 <211> 6611
 <212> DNA
 <213> 人工的

<220>
 <223> pCL0122

```

<400> 91
ctcttatctg cctcgcgcgc ttgaccgcgc cttgaactct ggcgcttgcg gctcgcaccc 60
tgctctgcgc gcgcaggcgc gggggcaggt ggttgggtcc gcagccttcc gcctcgcgc 120
gctagctcgc tcgcgcgcgt ctgcagccag cagggcagca ccgcacggca ggcaggctcc 180
ggcgcggatc gatcgatcca tcgatccatc gatccatcga tcgtgcggtc aaaaagaaag 240
gaagagaaaa ggaagaaaga aggcgtgcgc acccgagtgc gcgctgagcg cccgctcgcg 300
gtccgcggga gcctcgcgct tagtcccgcc cccgcgcgc gcagtcaccc gggaggcac 360

```

	ggcaccatct cgcgcgcgc tcgcgcctcg ccgattcccc gcctccccc ttccgcttct	420
	tcgcgcgcctc cgcgcgcgcgc cgcgcgcgc gcgcgcgcct ccttatctgc tcccagggg	480
	ggcactccgc accttttgcg ccgcgtgcg ccgcgcgcgc cgcgcgcgc ccttggtttc	540
	ccccgcgagc ggcgcgcggt cgcgcgcgca agactcgcgc cgtgcgcgc cgagcaacgg	600
	gtggcgcgcg cgcgcgcgcg ggcgcgcgc ggcgcgcgc aggcgcgcgc aggcgcgcgc	660
	taggcgaaac gcgcgcgcgc ggcgcgcgc ccgcgcgc cagagcagtc gcgcgcgcag	720
	accgcgaacg cagagacgca gaccgaggtc cgtgcgcgc gacgcgcgc cgacgcgcgc	780
	caggcgagcg gacgcgcgcg ccgcgcgcgc ccgcgcgcgc ggcgcgcgcgc agagcgagga	840
	cgcgcgcgcgc agcgtgcag ttccgcgcgc agacgcgcgc gcgcgcgcgc gagagcgat	900
	aaggcgcttg gatcgcgcgc ggcgcgcgc ggcgcgcgc gaaatcggt ggagcgagga	960
	gtatcttgcg tgcctggacg aggcagctga cgcgcgcgc gcgcgcgcgc atgatgatgt	1020
	gcacagagaa gaagcagtc gaaagcagc actagcagc aaggcagca tgaagtgcgc	1080
	gaactcgcgc gcaattttgc ttgtgcgcgc catagcgcgc gcgcgcgcgc agagcgatgg	1140
	ctgcaccccc accgaccaga cgcgtgtgag caagcgcgag ggcgcgcgc cgcgcgcgc	1200
	gcgcgcgcgc gtcgcgcgcg aggcgcgcgc aaacgcgcgc aagtcagcg tgcgcgcgc	1260
	ggcgaggggc gatgcacac accgcagct gacgcgcgc ttcacgcgc ccacgcgcgc	1320
	gctgcgcgc cctgcgcgc cctgcgcgc caccgcgcgc tgcgcgcgc agtcgcgcgc	1380
	ccgcgcgcgc gaccacatga agcgcgcgc ctttcgcgc tgcgcgcgc cgcgcgcgc	1440
	cgtccagcgc cgcgcgcgc ttttcgcgc cgcgcgcgc tacaagcgc gcgcgcgcgc	1500
[0053]	gaagtgcgc ggcgcgcgc ttgtgcgcgc cgcgcgcgc aagcgagcg acttcgcgc	1560
	ggcgcgcgc aaccgcgcgc aaagcgcgc gacgcgcgc aaagcgcgc aagtcgcgc	1620
	cagcgcgac aagcgcgcgc accgcgcgc ggcgcgcgc aagtcgcgc aacacgcgc	1680
	ggcgcgcgc gtcgcgcgc cgcgcgcgc ccgcgcgcgc accgcgcgc ggcgcgcgc	1740
	cgtgcgcgc ccgcgcgcgc actgcgcgc cgcgcgcgc gcgcgcgcgc aagcgcgcgc	1800
	cgcgcgcgc gatgcacatg tgcgcgcgc gtcgcgcgc gcgcgcgcgc tgcgcgcgc	1860
	cagcgagcg cgtgcgcgc accgcgcgc ccgcgcgcgc catatgcgc atgcgcgcgc	1920
	aaagtcgcgc ttgcgcgcgc cgcgcgcgc aggcgcgcgc ggcgcgcgcgc aaagtcgcgc	1980
	attacatagt attttgcgc tgcgcgcgc gttgcgcgc tttttgcgc tgaacgcgc	2040
	caagcgcgc gtcgcgcgc attgcgcgc ttgcgcgcgc gtcgcgcgc ttgcgcgcgc	2100
	cgtgcgcgc cgcgcgcgcgc aagcgcgcgc attgcgcgc ggcgcgcgc agaaaaagaa	2160
	aaagttgcgc gaccgcgcgc atgcgcgcgc aagcgcgcgc aaacgcgcgc gtcgcgcgc	2220
	tcgcgcgcgc aagaatctta gccgcgcgc cgaagtcgc tgcgcgcgc gcgcgcgcgc	2280
	acatgcgcgc ttgcgcgcgc ggcgcgcgc tgcgcgcgc cttgcgcgc atgcgcgcgc	2340
	cgtgcgcgc cctgcgcgc atgcgcgcgc aatgcgcgc actgcgcgc caactgcgcgc	2400
	atgcgcgcgc tgcgcgcgc gtcgcgcgc ttgcgcgcgc ctcgcgcgc tgcgcgcgc	2460
	gcgcgcgcgc ccttcgcgc tgcgcgcgc acgcgcgcgc tgcgcgcgc ttgcgcgcgc	2520
	tcgcgcgcgc tgcgcgcgc tgcgcgcgc tgcgcgcgc aagcgcgcgc agcgcgcgc	2580
	acgcgcgcgc gtcgcgcgc cgcgcgcgc tgaagcgcgc ggcgcgcgcgc tgcgcgcgc	2640
	ggcgcgcgcgc tgcgcgcgc tgcgcgcgc ggcgcgcgc ctcgcgcgcgc tgcgcgcgc	2700

[0054]

```

caacttgttt attgcagctt ataattggtt caataaage aatagcatca caaatftcac 2760
aaataaagca tttttttcac tgcattctag ttgttggttg tccaaactca tcaatgtatc 2820
ttatcataca tggtegacct gcaggaacct gcattaatga atcggecaac ggcggggag 2880
agggcggttg cgtattgggc gcttttcgc tctctcgctc actgactcgc tgcctcggt 2940
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga 3000
atcaggggat aacgcaggaa agaacatgtg agcaaaagcg cagcaaaagg ccaggaaaccg 3060
taaaaagccc ggttgcctgg cgtttttcca taggtctcgc cccctgacg agcateacaa 3120
aaatcgacgc tcaagtccga ggtggcgaaa cccgacagga ctataaagat accaggcggt 3180
tccccctgga agctccctcg tgcctctccc tgttcgacc ctgcctgta ccggatacct 3240
gtccgccttt ctccttcgg gaugcglggc gctttctcat agctcacgt giaggtaict 3300
cagttcggtg taggtcgttc gctccaagct gggtcgtg cagcaacccc ccgttcagcc 3360
cgacctgcgc gccttatccg gtaactatcg tcttgagtc aacctcgtaa gacacgactt 3420
atcgcaactg gcagcagcca ctggaacag gattagcaga gcgaggtatg taggcggtgc 3480
tacagagttc ttgaagtgtt ggcctaacct cggctacact agaagaacag tatttggtat 3540
ctgcctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 3600
acaaaccacc gctggtagcg gtgtttttt tgttgcaag cagcagatta cgcgcagaaa 3660
aaaaggatct caagaagatc cttgatctt ttctacggcg tctgacgctc agtggaaaga 3720
aaactcagct taagggtatt tggctatgag attatacaaa aggatcttca cctagatcct 3780
ttfaaattaa aaatgaagtt ttaaatacat ctaaagtata tatgagtaaa ctfggtctga 3840
cagttaccaa tgcctaatca gtgagccacc tatctcagcg atctgtctat ttcttctate 3900
catagtctgc tgaactcccg tctgttagat aactacgata cgggagggtt taccatctgg 3960
ccccagttgt gcaatgatac cgcgagaccc acgtcacccg gctccagatt tatcagcaat 4020
aaaccagcca gcccgaaggc ccgagcgcag aagtggctct gcaactttat ccgcctccat 4080
ccagtctatt aattgttgc gggaaugctag agtaagtagt tccagatta atagtctgcg 4140
caacgttgtt gccattgcta caggcatcgt ggtgtcacgc tctgtgttg gtatgcttc 4200
attcagctcc ggttcccaac gataaaggcg agttacatga tccccatgt tgtgcaaaaa 4260
agcggtttagc tcttctggtc ctccgactgt tctcagaagt aagttggcgc cagtgttate 4320
actcatggtt atggcagcac tgcataatte tcttactgta atgcatccg taagatgctt 4380
ttctgtgact ggtgagtaact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag 4440
ttgccttgc cggcgctcaa tacgggataa tacgcgcaca catagcagaa cttaaaaagt 4500
gtctatcatt gaaaaagctt ctteggggcg aaaaactctca aggatcttca cgtgttgag 4560
atccagttcg atgtaaccca ctctgcacc caactgatct tcagcatctt ttactttcac 4620
cagcgtttct ggttgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc 4680
gacacggaaa tgttgaatac tcatactctt cctttttcaa taitattgaa gcattttatca 4740
gggttatgtt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 4800
ggttccgcgc acatttcccc gaaaagtccc acctgacgta taagaaacca ttattatcat 4860
gacattaaac tataaaaaata ggcgtatcac gaggcctttt cgtctcgcgc gtttcgggtga 4920
tgaagggtgaa aacctctgac acatgcagct cccggagacg gtcacagctt gtcgtgaagc 4980

```

ggatgcggg agcagacaag cccgtcagg cgcgtcagcg ggtgttggcg ggtgtcgggg 5040
 ctggcttaac taigcggcat cagagcagat tgaetgaga gtgaaccaag ctccaattt 5100
 taggcceccc actgaccgag gtctgtcgat aatccacitt tccattgatt ttccaggttt 5160
 cgtaactca tgcactigag caaaacttcc gtctttctca acaaaagctc tcttcacaaa 5220
 gcattgcgcg gcaacggagc gtctctcata ctccactgcc acacaaggic gataaactaa 5280
 gctctcaca aatagaggag aattccactg acaactgaaa acaatgtaig agagacgac 5340
 accactggag cggcggcgcg gttggcgcg gaggtcggca gcaaaaacaa gcgactcgcc 5400
 gagcaaaccc gaatcagcct tcagacggic gtgcctaaac acacgcgttt ctacccegcc 5460
 ttcttcgcgc cccitcgctt ccaagcctcc ttcaagttaa tctctctagt tcaacttcaa 5520
 gaagaacaac accaccuaca ccatgattga acaagatgga ttgcacgcag gtctccggc 5580
 cgcttgggtg gagagctat tggctatga ctgggcacaa cagacaatcg gctgtctga 5640
 tgcgcgcgtg ttcggctgt cagcgcagg gcgcgcggtt ctttttgta agaccgacct 5700
 gtccgttccc ctgaatgaac tgcaggacga ggcagcgcgg ctatcgttgc tggccacgac 5760
 gggcgttctt tgcgcagctg tgcctgactg tgcactgaa ggcggaaggg actggtctgt 5820
 attggcgcaa gtgcggggc aggatctctt gtcctctcac ctgtctctg ccgagaaagt 5880
 atccatcatg gctgatgcaa tgcggcggtt gcatacgtt gatccggcta cctgccatt 5940
 cgaccacaa gcgaacac gcctcgagcg agcactaant cggatggaag ccgtcttctt 6000
 cgatcaggat gatctggacg aagagcaca gggcctcgcg ccagccgaac tgttcgccag 6060
 gtcaaggcg cgcctcccg accgcgatga tctcgtctg acccatggcg atgctctgtt 6120
 gccgaatate atggtggaaa atggcgctt ttctggatc ategactgtg gccgctggg 6180
 tglggcgga cgtatcagg acatagcgtt ggctaccgtt gatattgtg aagagcttg 6240
 cggcgaaatg gctgacgctt tctcgtctt ttacggtatc gccgtctccg attcgcagcg 6300
 categcctt tatcgctt ttgacgatt ctctgacac gtctacgag atttcgatt 6360
 caccgcgcgc ttctatgaaa ggttgggtt ccgaatggtt ttcgggacg ccggttgat 6420
 gatctccag cgcggggatc tcatgctgga gtctctgcc caccaccaact tgtttatgc 6480
 agcttataat ggttacaaat aaagcaatag catcacaat ttacaaaata aagcattttt 6540
 ttacactcat tctagttgtg gtttgcacaa actcactaat gtaacttact atgtctgaat 6600
 tccgggata c 6611

[0055]

<210> 92

<211> 1314

<212> DNA

<213> 人工的

<220>

<223> 密码子优化的 Isomerase

<400> 92

atggttaagg agtaattccc ccagatccag aagattaat tgcagggtaa ggacagcaag 60
 aaccgcctcg cctttcattt ctacagccc gagaaggagg tgatgggcaa gaagatgaag 120
 gactggcttc gctttgctat gctttgttg cacactctct gcctgaggg cgcggaccag 180
 ttggcgcgcg gtacgaagag ctttctgtg aacgagggca ctgacgtat tgagattgct 240
 aagcagaagg ttgacgttgg ttctgagatt atgcagaagc tgggtattcc gtaactactgc 300
 ttacacgatg tgaactctgt ttccgagggc aactcgatcg aggagtacga gtccaacctc 360

```

aaggctgtgg ttgcttaect caaggagaag cagaaggaga ccggaatcaa gctctcttgg 420
agcaccgccca acgtttttcgg ccacaagcgc tacatgaacg ggcctctcac caaccttgac 480
ttcgatgttg ttgcccgcgc tattgtccag attaagaacg ccctcgacgc tggatctgag 540
ctcggagccg agaactacgt tttttggggc ggacgcgagg gttacatgtc cctctcaaac 600
accgaccaga agcgtgagaa ggagcactg gccactatgc ttaccaatgc ccgcgactac 660
gcctgcagca agggtttttaa gggtactttt ctctttgagc cgaagcccat ggagccgacc 720
aagcaccagt aagcgtcga caccgagacc gccattgget tcttaaggc ccacaacctt 780
gacaaggatt ttaaggtaga cctcgagggt aaccacgcta cgtttgcgg ccacaacctt 840
gagcatgagc tcgctgcgc tgttgacgc ggaatgctt gttccattga cgcacaacgc 900
ggcgactacc agaaccgctg ggacaccgac cagtttcga ttgaccagta cgagctcgtc 960
caggccctga tggagatcat ccgttggtga ggctttgtta ccgttggtac gaacttcgac 1020
gccaaagacc gccgttaacg caccgacctc gaggacatca tctttgctc tgtgtcgggc 1080
atggacgcca tggctgcgc ctctgagaa gcctctaagc tctctcagga gacccctac 1140
acgaagtga agaaggagc ctacgcctg ttgacagcg gaatcgtaa ggatttcgag 1200
gatggcaagc tcacctgga gcaggtgtac gactacgga agaagaacgc cgagccgaag 1260
cagaccagcg gcaagcagga gctctacgag gccattgtc ccatgtacca gtag 1314

```

[0056]

<210> 93
 <211> 1485
 <212> DNA
 <213> 人工的
 <220>
 <223> 密码子优化的 Kinase

```

<400> 93
atgaagaccg tcgcggcat cgtcttga acccagcca tgaagggtgt cttttacgac 60
taagagaaga aggagatcat cgagtcgc tcgtgccta tggagctat tagcgatcg 120
gacggaaacc gcagcagac gactgagtg ttigacaagg gtctcgaggt gtgctttgga 180
aagctctcgc ctgataaaa gaagaccatt gagcgattg gcctctcgg ccagctccac 240
ggcttcgtcc ctctcgatgc gaacggaaag ggcctctaca acatcaagct ctggtcgac 300
accgccactg tggaggatg caagatcatt acgacgcgc ccggcgcgca caaggctgtc 360
atcgacgcgc tcggcaacct catgctcac ggattcacgc cccgaagat tctctggctc 420
aagcgcaaca agcccgagge ctttctaac ctcaagtaca ttatgtgcc ccacgattac 480
ctcaactgga agctgactgg agactacgtc atggagtac ggcagcctc cggaaccgcc 540
ctttttgatt cgaagaaccg ctgctgtgag aagaagattt gcacattat tgatcttaag 600
ctgctcgacc ttctccctaa gctcattgag ccctcgccc ccgcggtaa ggtaaacgac 660
gaggccgcca aggcgtacgc cttcccgcc ggaatcccc tttccgtgg ccgcgggat 720
aacatgatgg gtgcggtcgc tactggcacc gtcgtgacg gattctcac gatgagcatg 780
ggcaacctcg gaaetcttta eggtactcg gacaagccta ttccgaccc ggctaacggc 840

```

ctcagcgget tetgcagctc caccggcgcc tggetteccc tcttttgac catgaactgc 900
 accgtcgcca ccgagttcgt ccgcaacctt ttccagaagg atatacagga gctgaacgtc 960
 gaggttgeta agtccccctg cggcagcgag ggcgttcctg tcattctttt ctccaacggc 1020
 gagcgcccc cgaacctccc caacggcgcc gctcgattta ccggcctcac ctccgcgaac 1080
 acgtcccgcg ccaacatcgc tcgcgcctcc ttgagtcgg ccgtctttgc catgcgcggt 1140
 ggctcgatg cgtttcgtaa gctcgattc cagcccaagg agattcgct catcgcggt 1200
 ggttcgaagt ccgacctcgt gcgccagatc gctgctgaca ttatgaacct tcccatccgt 1260
 gtcccccttc tcgaggagcg cgcgcctc ggccgagctg tccaggcct ttggtgcctt 1320
 aagaaccagt ccgtaagtg ccacatcgtc gagctttgca aggagcctat caagattgac 1380
 ggtccaaga accccaacc gattgcgag aacgtcgccg tgtacgataa ggctacgat 1440
 gactactga aggtcgtaa cactcagc cctctgtag cctaa 1485

<210> 94

<211> 1569

<212> DNA

<213> 人工的

<220>

<223> 密码子优化的转运蛋白

<400> 94

atgggcctcg aggataaccg catgtttaag cgttttgca acgtggcgga gaagaaggcc 60
 ggtagcaacg ccattgccat catgtttgca ctcttcggcg cctcgggcgg cgtctcttc 120
 ggtacgaca ccggcactat ctggcgctc atgactatgg actacgttct cgcctcgac 180
 [0057] cctccaaca agcactctt caccgtgac gactgctgc tcatgtttc cttcttttg 240
 gtcggcactt tcttcggcgc ccttcggcc ccgttctca acgataccct cggccgcgc 300
 tgggtctca tcttcagcgc cctcatgtc tttaacatcg gcgccatctt ccaggtcatt 360
 tccaccgcca tccccctgct ctgcgcggcg ccggttatcg ccgatttcgg tctcgccctc 420
 atttcgcca ccatcccgct ctaccagtc gagactgct cgaagtggat tcgcgcgcgc 480
 atcgtttct gctaccagt gcccatcact atcgacttt tctcgtctc ctgctcaac 540
 aagggcaccg agcactgac caactcggg tcttaccgta ttctctgga cctccagtc 600
 ctctgggcgc tcatcttgg tatggcag atttctcc ctgagaccc ccgtctcgg 660
 atttcgaagg gcaaccagga gaaggccgc gactccctcg ccgtctcgg caagctcccc 720
 atcgaccatc ctgatagcct tgaggagctt ccgafatta ctgcgccta ccagttcgag 780
 accgtctacg gtaagtcag ctggtccag gtctttccc acaagaacca tcagctcaag 840
 cgcctctta ccggcgttgc ctttcaggcc ttccagcgc tcccgaggt taactttatc 900
 ttttactacg gcaccaacct ttttaagcg gccggagtc accgattcac ctcagcctt 960
 gccaccaaca tctttaacgt ccgcagcact attcccgca ttcttctcat ggaggtctc 1020
 ggccgcgcca acatgctcat ggccggtgcc accggcatgt cgtgtcgca gcttatcgct 1080
 gccattgtcg gatttgccac gtccgagaa aacnagtcg gccagtcggt cctctcgct 1140
 ttctctgca tctttatcgc tttttttgc gccacctgg gtccctgcgc ctgggtcgct 1200
 gtccggcagc tctttccct tcgcactcg gctaaagtcg ttccctctg caaccgctc 1260
 aactggctct ggaactgggg catgtctac gccacccctt acatggtcga ccaggataag 1320
 ggtaacctcg gcagcaacgt ttttttatt tggggaggt tcaacctcgc ttgcgtctt 1380

```

ttcgcgtggg atttcattta cgcagccaag ggcctttccc tcgagcaggt tgaatgagctc 1440
tacgagcatg ttctgaagge gtggaagicc aagggttttg tcccgtecaa gcactccitt 1500
cgcgagcagg tcgaccagca gatggacicc aagaccgagg ccattatgag cgaggaggcg 1560
tcggtttaa 1569

```

<210> 95
 <211> 1512
 <212> DNA
 <213> 人工的

<220>
 <223> 密码子优化的转运蛋白

```

<400> 95
atggccctcg accctgacea gcagcagccc atttccctcg tgcgcgcga gtttggtaag 60
tcgtccgggtg agatctcccc cgcgcgtag cctctcatta aggagaacca cgtccccag 120
aactaccccg ttgttgcgc cactctcccc ttctcttccc cggccctggg tggcctcctt 180
tacggitaag agattggcgc tacgtctgc gctacgattt ccttcagtc cccctccctc 240
tcggcatct cctggtacaa cctctctccc gtctgatttg gcctctcac ttccggttcc 300
cctacggtg ctctgtttgg ctccattgtt gccttcacca ttgccagct tattggcgt 360
cgcaaggagc ttatctctgc tctctctccc tacctctctg gtccctctg tacctctctc 420
gccctacgt actccgttct cactctctgc cgtctctt accgtgttcc cgtctctt 480
gccatgatg ctccctctat gtacatcgcg gagaccgccc cgtccctct accgcccag 540
ctcgtttccc tcaaggagtt ttctatctgt ctctgtatgg tcggcggata cggcatttgt 600
tccttcacgg tcaacgtcca ctccggttgg cgtctctat accgtctct cgttccctc 660
gtctgtatca tggcatttgg catgtgttgg ctctctctt cccctcttgc gctctctc 720
cgcgtctatc aggttaaggg taacgttgcg aaccagcgcg aggtctcat taagtcctc 780
tctctctctc gtgtctctgc ctctctctgc tcggcgcgc agcaggtcaa cgagattctc 840
gccgagctta cctctcttgg cgcagataag gaggtcact tcggcagct ctctcagga 900
aagtcctcca agccctctat tctctctgc ggccttctt tcttctcca gatcaccgt 960
cagcttctgg tctctctatc cgcctctctg atctctcaga ctctctctt ctctctctc 1020
ggcagctcta cccctcttct cttctctctc ggcctctcga agctctctat gacccgtct 1080
gccgtctctg ttatctctgc tctctctctg cgcctctctc tctctctctg agtctctgt 1140
atgttcttct cgtctctctt ctttctctgc tactctctt tctctctctc ttctctctc 1200
gtctctcttgc tcctctctct tctctctctg ggtctctctc agctctctt tgcctctct 1260
ggctctctta tgaattctga gatcttctcc ctcaagctcc gtgtctctgc actctctct 1320
gccgtctctg tcaacttctg tgcctctctc ctctctctct ttctctctc cctctctc 1380
gagctctctg gcctctctat cctgttctgc ggttctctgc ttatctctgt tctctctct 1440
gttttctctt ttctctctgt cccgtctctc aaggtctctc cgtctctctg gatctctct 1500
aagtcctct aa 1512

```

[0058]

<210> 96
 <211> 19
 <212> DNA
 <213> 人工的

[0059]

<220>		
<223>	引物 5' CL0130	
<400>	96	
cctcggggcg egtcctctt		19
<210>	97	
<211>	20	
<212>	DNA	
<213>	人工的	
<220>		
<223>	引物 3' CL0130	
<400>	97	
ggcggccttc tcttggttgc		20
<210>	98	
<211>	24	
<212>	DNA	
<213>	人工的	
<220>		
<223>	引物 5' CL0131	
<400>	98	
ctactcctgt gtgcggcca tcct		24
<210>	99	
<211>	22	
<212>	DNA	
<213>	人工的	
<220>		
<223>	引物 3' CL0131	
<400>	99	
ccgccgacca taccgagaa ga		22
<210>	100	
<211>	1362	
<212>	DNA	
<213>	人工的	
<220>		
<223>	密码子优化的 NA	
<400>	100	
atgaacccca accagaagat tactactatc ggtagcattt gectcgtctg ttgacttata		60
tcccttattc ttacagattg taacattata tccatttggg tctcgcatac ctttcagacc		120
ggctcccaga accacaccgg catttgcac cagaacatta ttaattacaa gaactccact		180
tgggtcaagg acactactag cgttatcttt accggtaact cgtcgttttg cctatttcgc		240
ggctgggcta ttacagcaa ggacaactcg atccgcacg gtagcaaggg cgaagtcttt		300
gicacccgtg agccttttat ttctgcagc caccctcagt gccgtacttt tttctgact		360
cagggcctc tctcaacga taagcattcc aacggcactg tcaaggatcg cagccctac		420
cgcgccctta tgcctgcgc tctcggcgag gctcccagcc cctacaactc cgttttgag		480
tccgttgcct ggctcccgag cgcctgcac gacggaatgg gatggctac taatggatct		540
tccggccctg ataacggcgc tctcggcgtc cttaagtaca accgcattat caccgagacc		600
atcagtcctt ggagtaagaa gatctctcgc acccaggagt ccgagtgcgc ctgcgtcaac		660

	ggragetget teacgattat gacagacgce cccctccgacg gcctcgettc ctacaagatt	720
	tttaagattg agaagggtaa ggtaacgaag tccatcgagc ttaacgcccc gaactccac	780
	tacgaggagt gcctctgcta cctgacact ggcaaggta tgtgcgtctg ccgcgataac	840
	tggcatgget ccaaccgccc ctgggttagc ttcgatcaga accttgacta ccagattgga	900
	tacatttget cgggtgtttt tggcgacaa ccgcgccccg aggatggaac tggttcgtgc	960
	ggctcgtttt acgttgacgg cgcacaacgc gttaagggtt ttctctaccg ttacggtaac	1020
	ggagcttgga tggcgccac caagtcgac agctcgccgc acggatttga gatgatctgg	1080
	gaccccaacg gatggactga gacgatctcc aagtttagcg ttccgcagga tgtcgttget	1140
	atgaccgatt ggtcgggata ctccggttcc ttgtgcagc acctgagct caccggcctt	1200
	gactgcaltg gcccttgctt ttgggtcgag ctcatctcg gtcgccctaa ggagaagact	1260
	atttggacct ccgccagcag catttccitt tgcggcgata actccgacac cgtcgactgg	1320
	tcttgccccc atggcgccga gcttcccttt tccattgata ag	1362
	<210> 101	
	<211> 1431	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 密码子优化的 NA, 具有 V5 标签和多组氨酸标签	
	<400> 101	
[0060]	atgaacccca accagaagat tactactatc ggtagcattt gcctcgtcgt tggacttate	60
	tcccttatte ttcagattgg taacattate tccatttga tctcgcatag cttcagacc	120
	ggctcccaaga accaaccgg catttgcac cagaacatta ttacttaca gaactccact	180
	tgggtcaagg acactaetag cgttatctt accgtaact cgtcgtttg cctattcgc	240
	ggctgggcta ttacagcaa ggacaactcg atccgcacg gtagcaaggc cgaagttttt	300
	gtcctcgtg agccttttat ttctgcagc cactctgagt gcgtacttt tttctgact	360
	caggcgctc tctcaacga taagcattcc aacggcactg tcaaggatcg cagccctac	420
	cgcgcctta tgtctgcgc tgtcgccgag gctcccagc cctacaactc ccgtttgag	480
	tcctttgcct ggctcgccag cgcctgcac gacggaatgg gatggctac tatggattt	540
	tcggccctg ataacggcgc tgtcgccgtc cttaagtaca acggcattat caccgagacc	600
	atcaagtctt ggctgaaga gactctcgc acccaggagt ccgagtgcgc ctgcgtcaac	660
	ggcagetget tcaagattat gacagacgce cccctccgacg gcctcgettc ctacaagatt	720
	tttaagattg agaagggtaa ggtaacgaag tccatcgagc ttaacgcccc gaactccac	780
	tacgaggagt gctcctgcta cctgacact ggcaaggta tgtgcgtctg ccgcgataac	840
	tggcatgget ccaaccgccc ctgggttagc ttcgatcaga accttgacta ccagattgga	900
	tacatttget cgggtgtttt tggcgacaa ccgcgccccg aggatggaac tggttcgtgc	960
	ggctcgtttt acgttgacgg cgcacaacgc gttaagggtt ttctctaccg ttacggtaac	1020
	ggagcttgga tggcgccac caagtcgac agctcgccgc acggatttga gatgatctgg	1080
	gaccccaacg gatggactga gacgatctcc aagtttagcg ttccgcagga tgtcgttget	1140
	atgaccgatt ggtcgggata ctccggttcc ttgtgcagc acctgagct caccggcctt	1200
	gactgcaltg gcccttgctt ttgggtcgag ctcatctcg gtcgccctaa ggagaagact	1260

```

atttggacct cggccagcag catttcttt tgcggcgta actccgacac cgtcgactgg 1320
tcgtggcccg atggcgccga gcttcccttt tccatigata agggtaagcc tatccctaac 1380
cctctctctg gctctgattc tacgggtacc ggtaacatc accatcaacca t 1431

```

<210> 102
 <211> 539
 <212> PRT
 <213> 人副流感 3 病毒
 <400> 102

```

Met Pro Thr Ser Ile Leu Leu Ile Ile Thr Thr Met Ile Met Ala Ser
1          5          10          15

```

```

Phe Cys Gln Ile Asp Ile Thr Lys Leu Gln His Val Gly Val Leu Val
20          25          30

```

```

Asn Ser Pro Lys Gly Met Lys Ile Ser Gln Asn Phe Glu Thr Arg Tyr
35          40          45

```

```

Leu Ile Leu Ser Leu Ile Pro Lys Ile Glu Asp Ser Asn Ser Cys Gly
50          55          60

```

```

Asp Gln Gln Ile Lys Gln Tyr Lys Arg Leu Leu Asp Arg Leu Ile Ile
65          70          75          80

```

```

Pro Leu Tyr Asp Gly Leu Arg Leu Glu Lys Asp Val Ile Val Ser Asn
85          90          95

```

[0061]

```

Gln Glu Ser Asn Glu Asn Thr Asp Pro Arg Thr Lys Arg Phe Phe Gly
100          105          110

```

```

Gly Val Ile Gly Thr Ile Ala Leu Gly Val Ala Thr Ser Ala Gln Ile
115          120          125

```

```

Thr Ala Ala Val Ala Leu Val Glu Ala Lys Gln Ala Arg Ser Asp Ile
130          135          140

```

```

Gln Lys Leu Lys Glu Ala Ile Arg Asp Thr Asn Lys Ala Val Gln Ser
145          150          155          160

```

```

Val Gln Ser Ser Ile Gly Asn Leu Ile Val Ala Ile Lys Ser Val Gln
165          170          175

```

```

Asp Tyr Val Asn Lys Glu Ile Val Pro Ser Ile Ala Arg Leu Gly Cys
180          185          190

```

```

Glu Ala Ala Gly Leu Gln Leu Gly Ile Ala Leu Thr Gln His Tyr Ser
195          200          205

```

```

Glu Leu Thr Asn Ile Phe Gly Asp Asn Ile Gly Ser Leu Gln Glu Lys
210          215          220

```

```

Gly Ile Lys Leu Gln Gly Ile Ala Ser Leu Tyr Arg Thr Asn Ile Thr
225          230          235          240

```

```

Gln Ile Phe Thr Thr Ser Thr Val Asp Lys Tyr Asp Ile Tyr Asp Leu
245          250          255

```

```

Leu Phe Thr Glu Ser Ile Lys Val Arg Val Ile Asp Val Asp Leu Asn
    260                265                270

Asp Tyr Ser Ile Thr Leu Glu Val Arg Leu Pro Leu Leu Thr Arg Leu
    275                280                285

Leu Asn Thr Gln Ile Tyr Arg Val Asp Ser Ile Ser Tyr Asn Ile Gln
    290                295                300

Asn Arg Glu Trp Tyr Ile Pro Leu Pro Ser His Ile Met Thr Lys Gly
    305                310                315                320

Ala Phe Leu Gly Gly Ala Asp Val Lys Glu Cys Ile Glu Ala Phe Ser
    325                330                335

Ser Tyr Ile Cys Pro Ser Asp Pro Gly Phe Val Leu Asn His Glu Met
    340                345                350

Glu Ser Cys Leu Ser Gly Asn Ile Ser Gln Cys Pro Arg Thr Val Val
    355                360                365

Lys Ser Asp Ile Val Pro Arg Tyr Ala Phe Val Asn Gly Gly Val Val
    370                375                380

Ala Asn Cys Ile Thr Thr Thr Cys Thr Cys Asn Gly Ile Gly Asn Arg
    385                390                395                400

Ile Asn Gln Pro Pro Asp Gln Gly Val Lys Ile Ile Thr His Lys Glu
    405                410                415

Cys Asn Thr Ile Gly Ile Asn Gly Met Leu Phe Asn Thr Asn Lys Glu
    420                425                430

Gly Thr Leu Ala Phe Tyr Thr Pro Asn Asp Ile Thr Leu Asn Asn Ser
    435                440                445

Val Ala Leu Asp Pro Ile Asp Ile Ser Ile Glu Leu Asn Lys Ala Lys
    450                455                460

Ser Asp Leu Glu Glu Ser Lys Glu Trp Ile Arg Arg Ser Asn Gln Lys
    465                470                475                480

Leu Asp Ser Ile Gly Asn Trp His Gln Ser Ser Thr Thr Ile Ile Ile
    485                490                495

Val Leu Ile Met Ile Ile Ile Leu Phe Ile Ile Asn Val Thr Ile Ile
    500                505                510

Ile Ile Ala Val Lys Tyr Tyr Arg Ile Gln Lys Arg Asn Arg Val Asp
    515                520                525

Gln Asn Asp Lys Pro Tyr Val Leu Thr Asn Lys
    530                535

```

[0062]

<210> 103

<211> 511
 <212> PRT
 <213> 水泡性口炎病毒 Indiana
 <400> 103
 Met Lys Cys Leu Leu Tyr Leu Ala Phe Leu Phe Ile Gly Val Asn Cys
 1 5 10 15
 Lys Phe Thr Ile Val Phe Pro His Asn Gln Lys Gly Asn Trp Lys Asn
 20 25 30
 Val Pro Ser Asn Tyr His Tyr Cys Pro Ser Ser Ser Asp Leu Asn Trp
 35 40 45
 His Asn Asp Leu Val Gly Thr Ala Leu Gln Val Lys Met Pro Lys Ser
 50 55 60
 His Lys Ala Ile Gln Ala Asp Gly Trp Met Cys His Ala Ser Lys Trp
 65 70 75 80
 Val Thr Thr Cys Asp Phe Arg Trp Tyr Gly Pro Lys Tyr Ile Thr His
 85 90 95
 Ser Ile Arg Ser Phe Thr Pro Ser Val Glu Gln Cys Lys Glu Ser Ile
 100 105 110
 Glu Gln Thr Lys Gln Gly Thr Trp Leu Asn Pro Gly Phe Pro Pro Gln
 115 120 125
 [0063] Ser Cys Gly Tyr Ala Thr Val Thr Asp Ala Glu Ala Ala Ile Val Gln
 130 135 140
 Val Thr Pro His His Val Leu Val Asp Glu Tyr Thr Gly Glu Trp Val
 145 150 155 160
 Asp Ser Gln Phe Ile Asn Gly Lys Cys Ser Asn Asp Ile Cys Pro Thr
 165 170 175
 Val His Asn Ser Thr Thr Trp His Ser Asp Tyr Lys Val Lys Gly Leu
 180 185 190
 Cys Asp Ser Asn Leu Ile Ser Met Asp Ile Thr Phe Phe Ser Glu Asp
 195 200 205
 Gly Glu Leu Ser Ser Leu Gly Lys Lys Gly Thr Gly Phe Arg Ser Asn
 210 215 220
 Tyr Phe Ala Tyr Glu Thr Gly Asp Lys Ala Cys Lys Met Gln Tyr Cys
 225 230 235 240
 Lys His Trp Gly Val Arg Leu Pro Ser Gly Val Trp Phe Glu Met Ala
 245 250 255
 Asp Lys Asp Leu Phe Ala Ala Ala Arg Phe Pro Glu Cys Pro Glu Gly
 260 265 270
 Ser Ser Ile Ser Ala Pro Ser Gln Thr Ser Val Asp Val Ser Leu Ile
 275 280 285

Gln Asp Val Glu Arg Ile Leu Asp Tyr Ser Leu Cys Gln Glu Thr Trp
 290 295 300

Ser Lys Ile Arg Ala Gly Leu Pro Ile Ser Pro Val Asp Leu Ser Tyr
 305 310 315 320

Leu Ala Pro Lys Asn Pro Gly Thr Gly Pro Val Phe Thr Ile Ile Asn
 325 330 335

Gly Thr Leu Lys Tyr Phe Glu Thr Arg Tyr Ile Arg Val Asp Ile Ala
 340 345 350

Ala Pro Ile Leu Ser Arg Met Val Gly Met Ile Ser Gly Thr Thr Thr
 355 360 365

Glu Arg Val Leu Trp Asp Asp Trp Ala Pro Tyr Glu Asp Val Gly Ile
 370 375 380

Gly Pro Asn Gly Val Leu Arg Thr Ser Ser Gly Tyr Lys Phe Pro Leu
 385 390 395 400

Tyr Met Ile Gly His Gly Met Leu Asp Ser Asp Leu His Leu Ser Ser
 405 410 415

Lys Ala Gln Val Phe Glu His Pro His Ile Gln Asp Ala Ala Ser Gln
 420 425 430

Leu Pro Asp Gly Glu Thr Leu Phe Phe Gly Asp Thr Gly Leu Ser Lys
 435 440 445

Asn Pro Ile Glu Phe Val Glu Gly Trp Phe Ser Ser Trp Lys Ser Ser
 450 455 460

Ile Ala Ser Phe Phe Phe Thr Ile Gly Leu Ile Ile Gly Leu Phe Leu
 465 470 475 480

Val Leu Arg Val Gly Ile Tyr Leu Cys Ile Lys Leu Lys His Thr Lys
 485 490 495

Lys Arg Gln Ile Tyr Thr Asp Ile Glu Met Asn Arg Leu Gly Thr
 500 505 510

[0064]

ATGAAGGCTAACCTCCTCGTTCCTTCTTTCCGCTCTCGCTGCTGCGGATGCCGACACCA
TCTGCATTGGCTACCACGCTAACAAACAGCACGGACACCGTCGATACTGTCCTGGAGA
AGAACGTTACCGCACCCATTCCGGTCAACCTCCTGGAGGACAGCCACAACGGCAAGC
TCTGCCGTCTTAAGGGCATCGCCCCCTCCAGCTCGGCAAGTGCAACATCGCCGGCT
GGCTCCTCGGCAACCCGGAGTGCGATCCCTCCTCCCGTTTCGCTCCTGGTCGTACAT
TGTGGAGACTCCGAACAGCGAGAACGGTATCTGCTACCCCGGCGATTTTATCGACTA
CGAGGAGCTCCGCGAGCAGCTCTCCTCCGTGTCCAGCTTGAGCGTTTCGAGATTTTT
CCGAAGGAGTCCTCGTGGCCCAACCACAACACCAACGGCGTCACCGCCGCTGCTC
CCACGAGGGCAAGTCGAGCTTTTACCGCAACCTGCTTTGGCTCACCGAGAAGGGGG
TTCGTACCCTAAGCTCAAGAACTCGTACGTCAACAAGAAGGGCAAGGAGGTCTCTCG
TCCTCTGGGGCATCCACCATCCCCCGAACAGCAAGGAGCAGCAGAACATCTACCAG
AACGAGAACGCCACGTTTCGGTGGTTCACGTGCAACTACAACCGCCGCTTCACTCCTG
AGATCGCCGAGCGCCCCAAGGTGCGCGACCAGGCTGGCCGCATGAACTACTACTGG
ACCCTCCTTAAGCCCCGTGACACGATATCTTTGAGGCCAACGGCAACCTTATCGCGC
CCATGTACGCGTTTCGCCCTCTCCCGCGGCTTTGGTAGCGGCATCATTACCAGCAACG
CCAGCATGCACGAGTGCAACACGAAGTGCCAGACCCCGCCGGTGCCATCAACAGCA
GCCTGCCTTACCAGAACATCCACCCCGTCACCATCGGTGAGTGCCCGAAGTACGTGC
GCTCGGCCAAGCTCCGCATGGTCACGGGCCTCCGCAACACTCCTTCGATCCAGCCCC
CGGCCTCTTCGGCGCCATTGCCGGTTTCATCGAGGGCGGCTGGACGGGCATGATCGA
CGGCTGGTACGGCTACCACCACCAGAACGAGCAGGGCTCCGGTTACGCCGCGGACC
AGAAGTCCACCAGAACGCCATCAACGGCATTACTAACAAGGTCAACACGGTCATCG
AGAAGATGAACATTCAGTTTACCGCTGTGCGCAAGGAGTTCAACAAGCTGGAGAAG
CGCATGGAGAACCTCAACAAGAAGGGGACGATGGTTTCCTGGACATTGACCTAC
AACGCCGAGCTCCTCGTGCTCCTTGAGAACGAGCGTACCCTCGACTTCCACGACTCC
AACGTCAAGAACCTCTACGAGAAGGTCAAGTCGCAGCTCAGAACAACGCCAAGGAG
ATTGGCAACGGTTGCTTCGAGTTTACCACAAGTGCGACAACGAGTGCATGGAGTCC
GTCCGCAACGGCACCTACGACTACCCGAAGTACTCCGAGGAGTCGAAGCTGAACGC
GAGAAGGTGGACGGCGTGAAGCTGGAGTCCATGGGCATCTACCAGATCCTCGCCAT
TTACTCGACGGTTGCCTCGTCGCTCGTCCTCCTTGCTCTCCCTCGGTGCGATTTCGTCT
GGATGTGCTGAACGGCAGCCTTCAGTGCCGCATCTGCATC (SEQ ID NO: 76)

图1

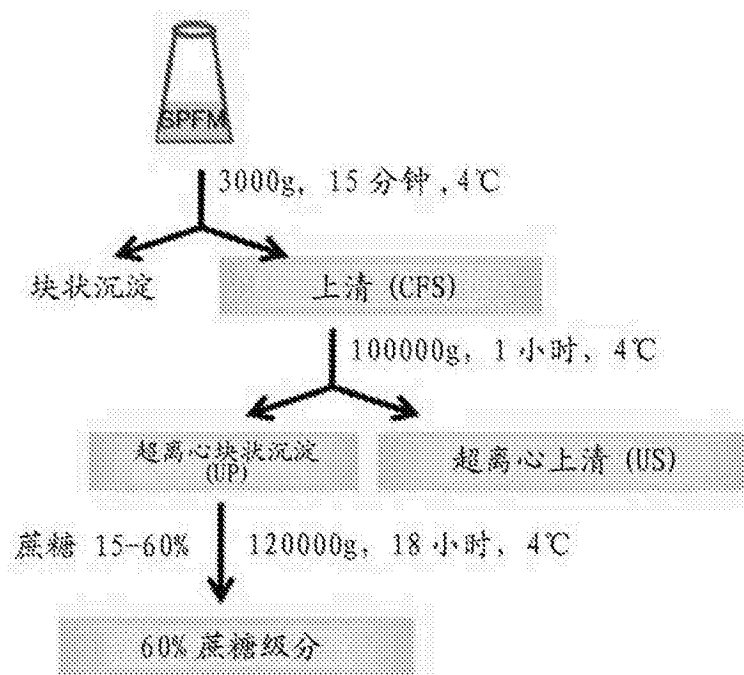


图3

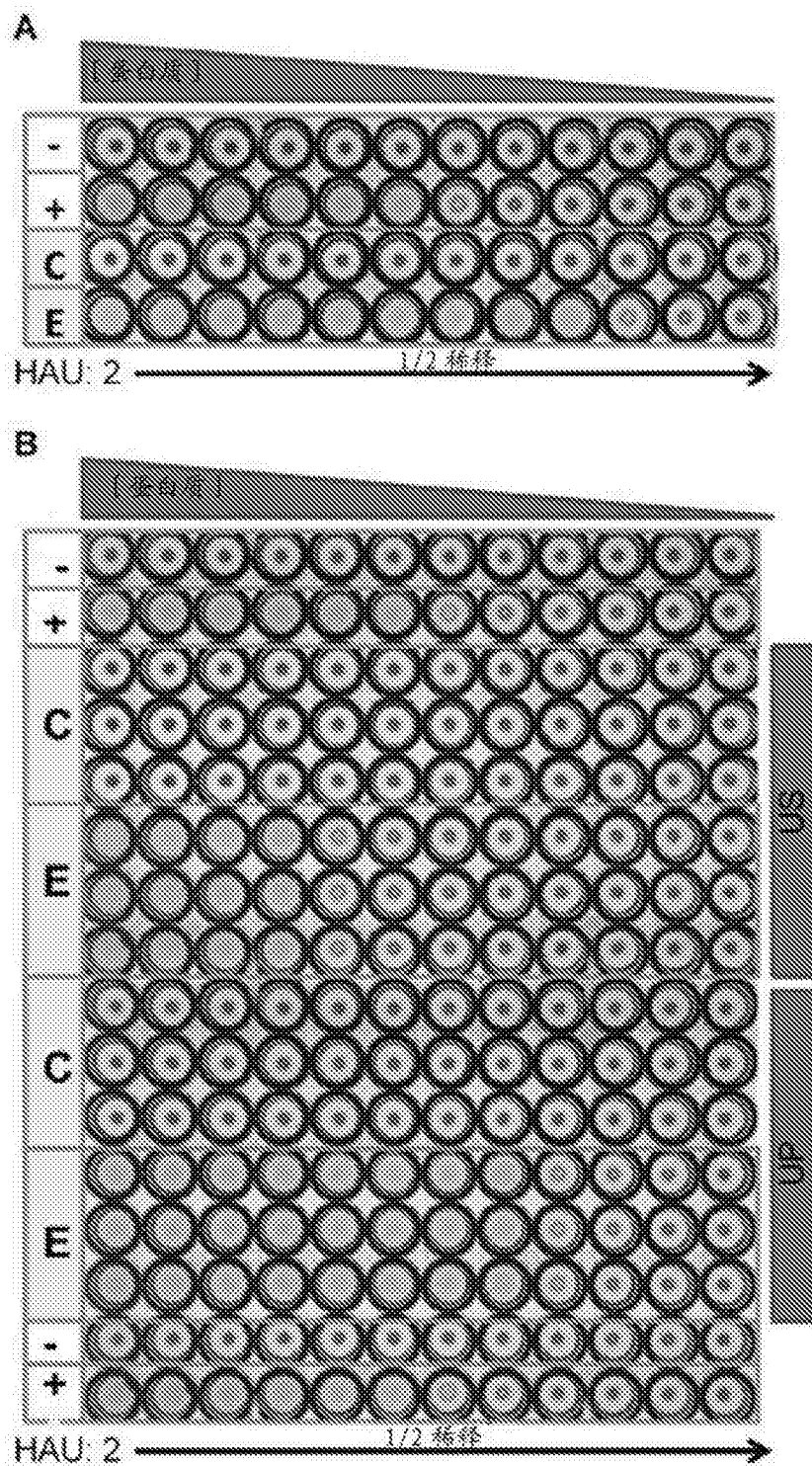


图5

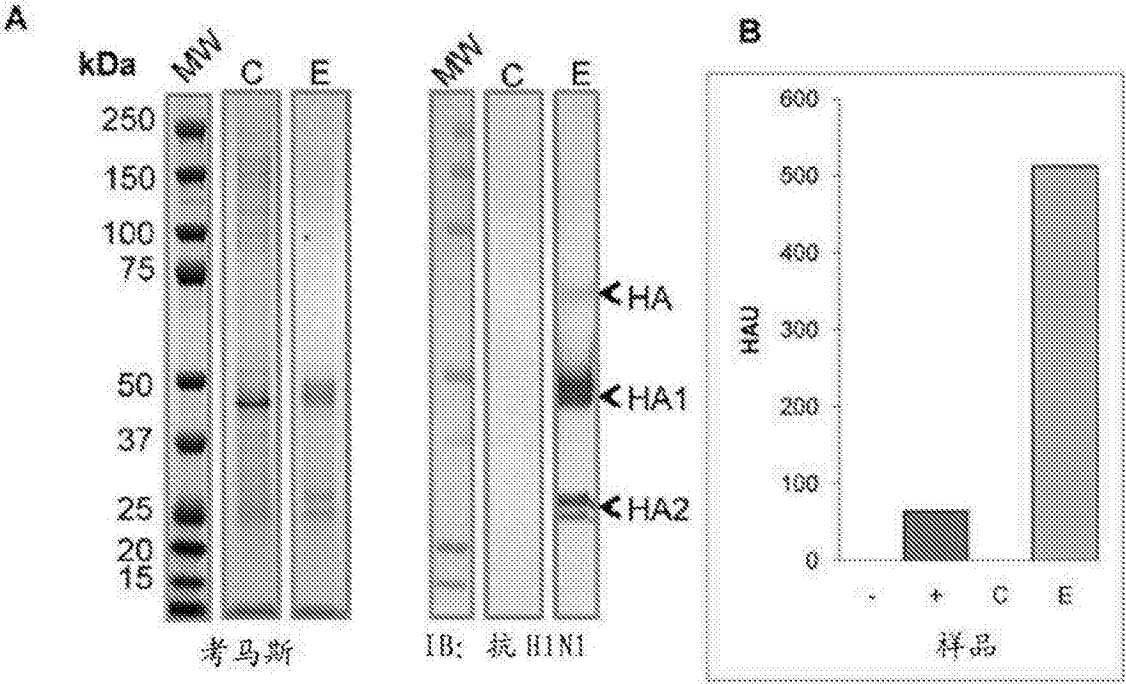


图6

物种: 甲型流感病毒 (A/Puerto Rico/8/34/Mount Sinai (H1N1))

名称: 血凝素 HA

在 HA1 中鉴定的肽

1 MKANLLVLLS ALAAADADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCR
61 LKGIAPLQLG KCNIAGWLLG NPECDPLLPV RSWSYIVETP NSENGICYPG DFIDYEELRE
121 QLSSVSSFER FEIFPKESSW PNHNTNGVTA ACSHEGKSSF YRNLLWLTEK EGSYPKLKNS
181 YVNKKGKEVL VLWGIIHPPN SKEQQNIYQN ENAYVSVVTS NYNRRFTPEI AERP KVRDQA
241 GRMNYWYWTLL KPGDTIIFEA NGNLIAPMYA FALS RGFSG ITSNASMHE CNTKCQTPLG
301 AINSSLPYQN IHPVTIGCEP KYVRS AKLRM VTGLRNTPSI QSR^{*}GLFGAIA GFIEGGWTGM
361 IDGWYGYHHQ NEQSGGYAAD QKSTQNAING ITNKVNTVIE **KMNIQFTAVG** KEFNKLEKRM
421 ENLNKKVDDG FLDIWTYNAE LLVLL ENERT LDFHDSNVKN LYEKVKSQ LK NNAKEIGNGC
481 FEFYHKCDNE CMESVRNGTY DYPKYSEESK LNREKVDGVK LESMGYQIL AIYSTVASSL
541 VLLVSLGAIS FWMCSNGSLQ CRICI (SEQ ID NO: 77)

中部: 在 HA2 中鉴定的肽

1 MKANLLVLLS ALAAADADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCR
61 LKGIAPLQLG KCNIAGWLLG NPECDPLLPV RSWSYIVETP NSENGICYPG DFIDYEELRE
121 QLSSVSSFER FEIFPKESSW PNHNTNGVTA ACSHEGKSSF YRNLLWLTEK EGSYPKLKNS
181 YVNKKGKEVL VLWGIIHPPN SKEQQNIYQN ENAYVSVVTS NYNRRFTPEI AERP KVRDQA
241 GRMNYWYWTLL KPGDTIIFEA NGNLIAPMYA FALS RGFSG ITSNASMHE CNTKCQTPLG
301 AINSSLPYQN IHPVTIGCEP KYVRS AKLRM VTGLRNTPSI QSR^{*}GLFGAIA GFIEGGWTGM
361 IDGWYGYHHQ NEQSGGYAAD **QKSTQNAING** ITNKVNTVIE **KMNIQFTAVG** KEFNKLEKRM
421 ENLNKKVDDG FLDIWTYNAE LLVLL ENERT **LDFHDSNVKN** LYEKVKSQ LK NNAKEIGNGC
481 **FEFYHKCDNE** CMESVRNGTY DYPKYSEESK LNREKVDGVK LESMGYQIL AIYSTVASSL
541 VLLVSLGAIS FWMCSNGSLQ CRICI (SEQ ID NO: 77)

图7

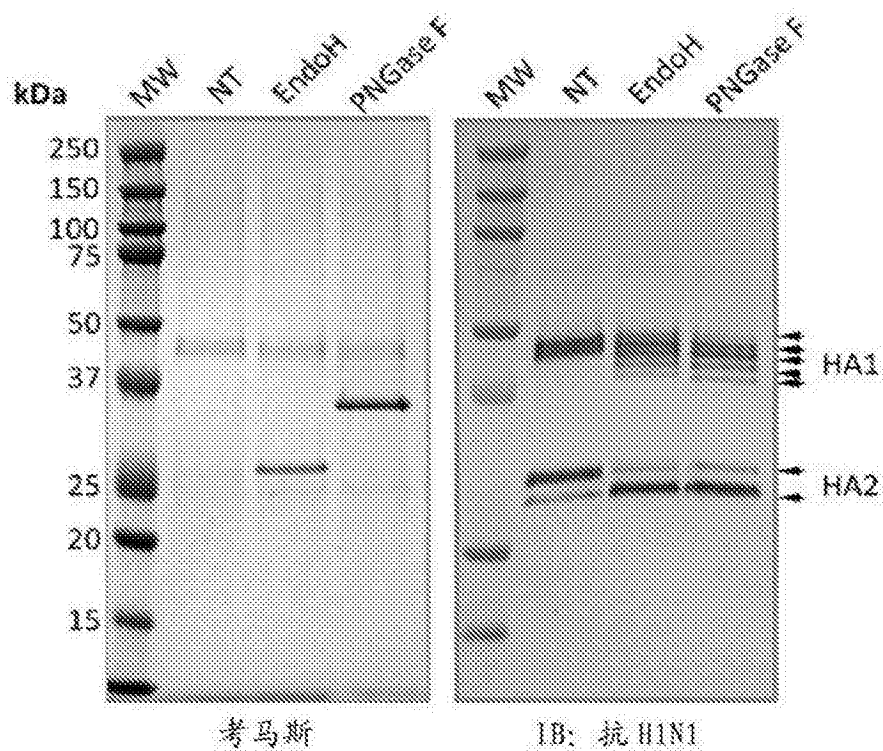


图8

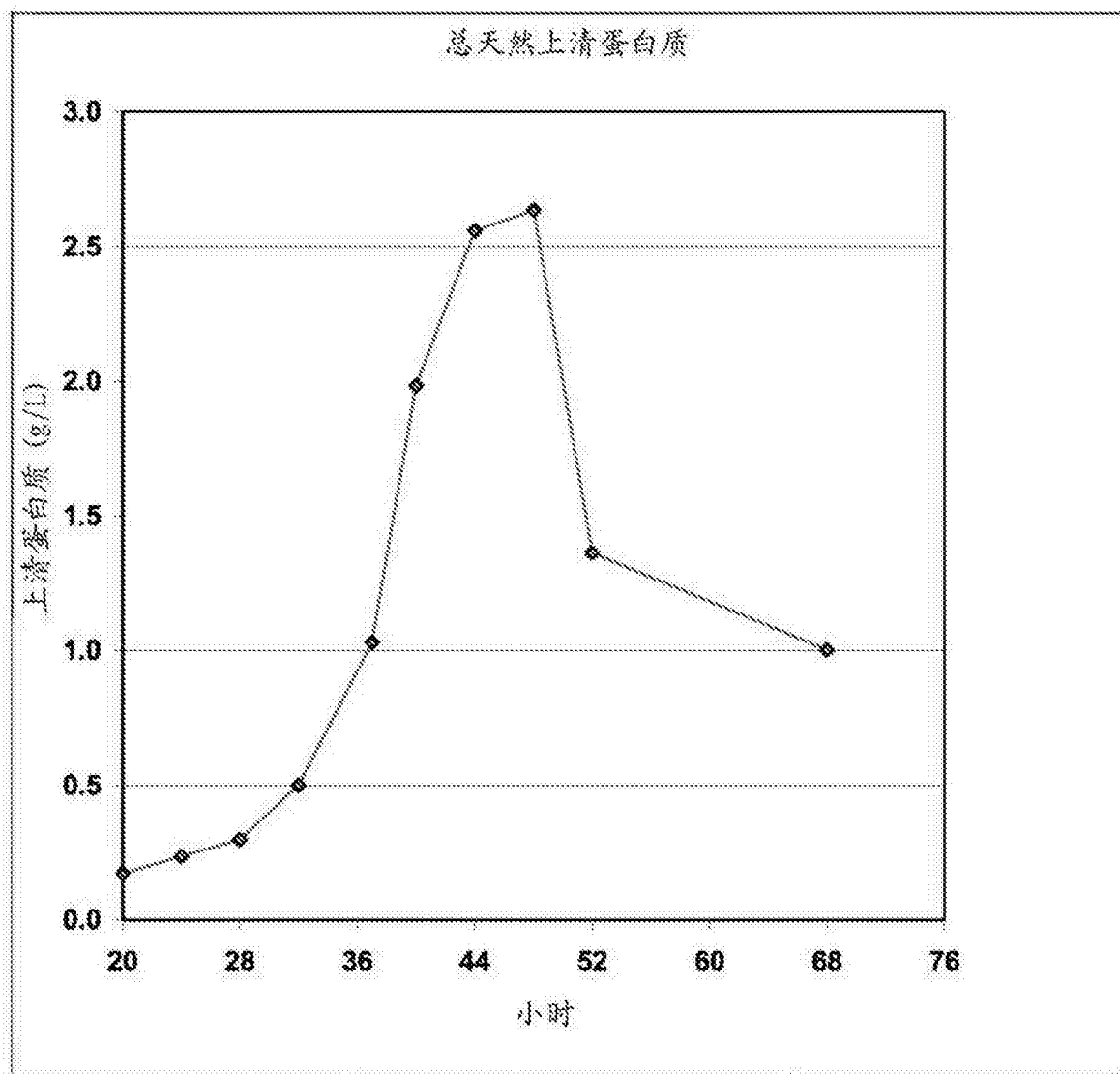


图9

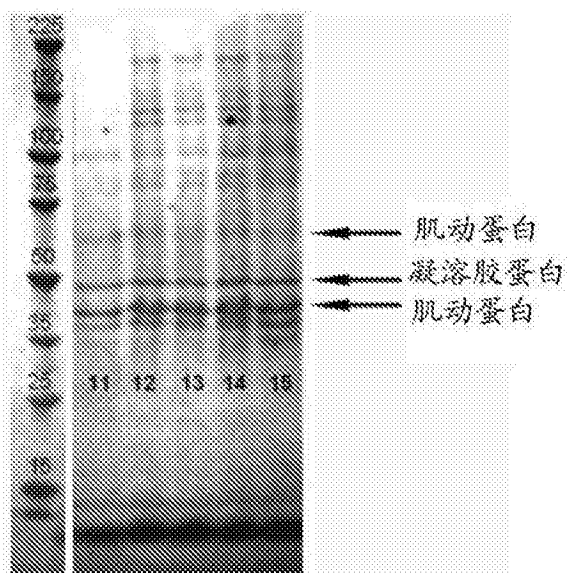


图10

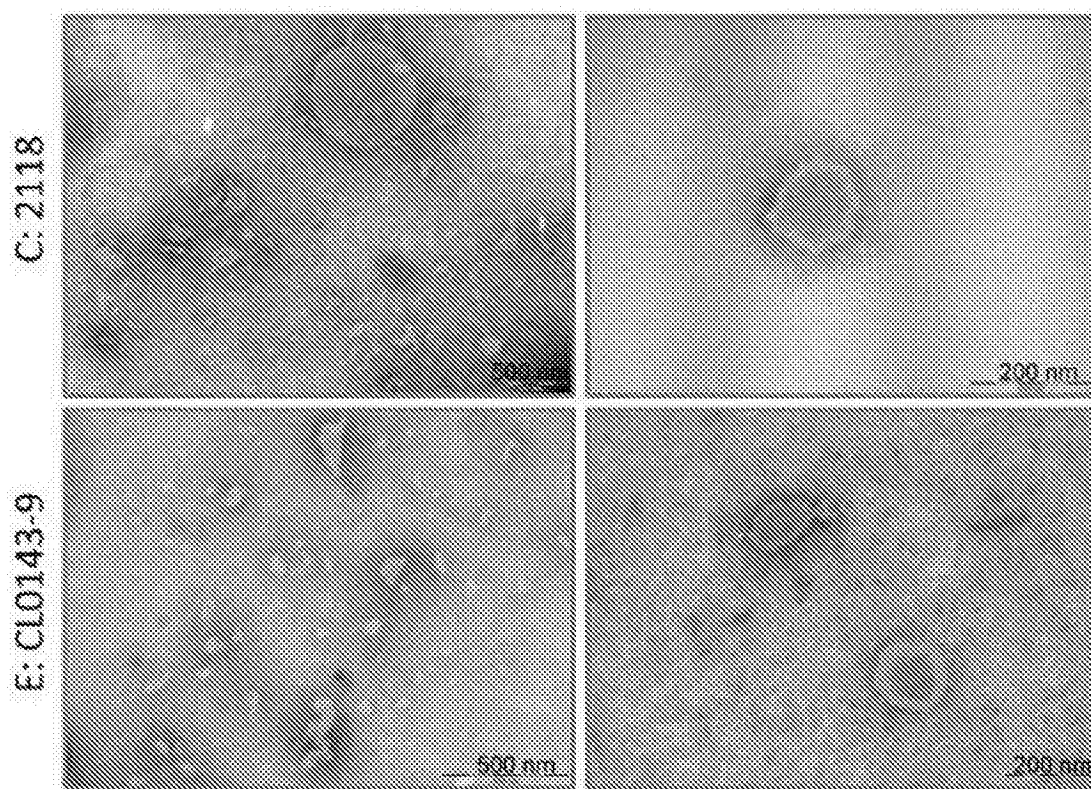


图11

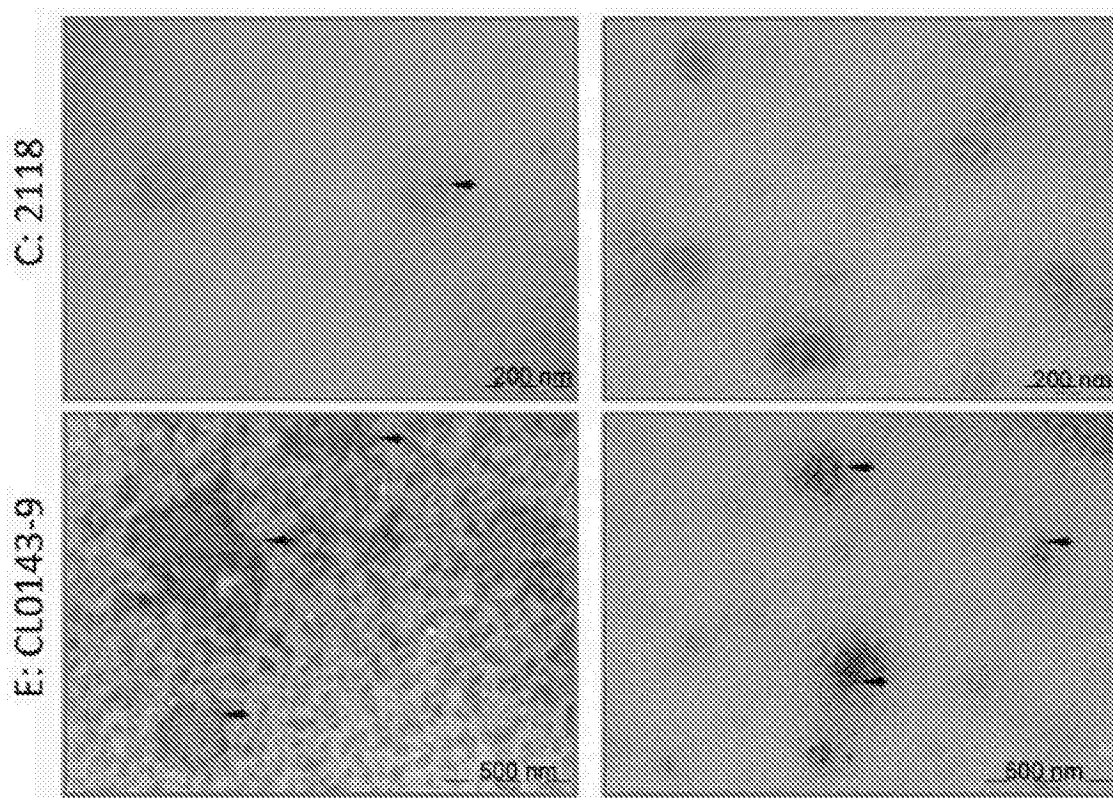


图12

裂殖壶菌属预测的信号锚序列

α -1,3 甘露糖基- β -1,2-GlcNAc- 转移酶 I 样蛋白 #1

MRGPGMVGLSRVDREHLRRRQQQAASEWRRWGFFVATAVLLVFLTVYPNV
(SEQ ID NO: 78)

ATGCGCGGCCCGGGCATGGTCGGCCTCAGCCGCGTGGACCGCGAGCACCTGCGGCGGGCGGC
AGCAGCAGGCGGCGAGCGAATGGCGGCGCTGGGGTTCTTCGTCGCGACGGCCGTCGTCCT
GCTCGTCTTCTCACCGTATACCCGAACGTA (SEQ ID NO: 79)

β -1,2- 木糖基转移酶样蛋白 #1

MRTRGAAYVRPGQHEAKALSSRSSDEGYTTVNVRTRKRRTVAALVAAALLVTGFIVVVVFV
VVV (SEQ ID NO: 80)

ATGCGCACGCGGGGCGCGGCGTACGTGCGGCCGGGACAGCACGAGGCGAAGGCGCTCTCGT
CAAGGAGCAGCGACGAGGGATATACGACGGTCAACGTTGTCAGGACCAAGCGAAAGAGGA
CCACTGTAGCCGCGCTTGTAAGCCGCGGCGCTGCTGGTGACGGGCTTTATCGTCGTCGTCGTCT
TCGTCGTCGTTGTT (SEQ ID NO: 81)

β -1,4- 木糖苷酶样蛋白

MEALREPLAAPPTSARSSVPAPLAKEEGEEEDGEKGTFGAGVLGVVAVLVIVVFAIVAGGGGDI
(SEQ ID NO: 82)

ATGGAGGCCCTGCGCGAGCCCTTGCTGCGCCGCCAACGTCGGCGCGATCGTCGGTGCCAGC
GCCGCTCGCGAAGGAGGAGGGGgAGGAGGAGGACGGGGAAAAaGGGACGTTGGGGCGGG
GGTCCTCGGTGTCGTGGCGGTGCTCGTCATCGTGTTGCGATCGTGCGGGAGGCGGAG
GCGATATT (SEQ ID NO: 83)

半乳糖基转移酶样蛋白 #5

MLSVAQVAGSAHSRPRRGGERMQDVLALAESSRDRKRATARPGLYRALAILGLPLIVFIVWQMT
SSLTTAPSA (SEQ ID NO: 84)

ATGTTGAGCGTaGCACAAGTCGCGGGGTCGGCCCACTCGCGGCCGAGACGAGGTGGTGAGC
GGATGCAAGACGTGCTGGCCCTGGAGGAAAGCAGCAGAGATCGAAAACGAGCAACAGCAA
GGCCCGGGCTATATCGCGCACTTGCGATTCTGGGGCTGCCGCTCATCGTATTCATCGTATGG
CAAATGACTAgCTCCCTCACGACTGCCCCGAGCGCC (SEQ ID NO: 85)

图13

裂殖壶菌属预测的 I 型膜蛋白

EMC1

MGTTTARMAYAVLAAAVSVAHGLHEDQAGVNDWTVRNLGAYAHGVFLDDDLALVATTQATVGAVRM
 TDGEVYVWRETLPTARSAPLASQVKHELFATASADACVIELWATPSGDVMTSDSRQAGLEWDKICDNTDA
 DATGVLELLDNDNFNDGTPDVAALTPFQFVILDGVSGRVLHEVDLDKTIWQGLVEAAGSATGKKRPS
 IMAYGVDIKTGKLEVRKLANSGATLDPVSGLEGVSADEITVLKSGVAKVGSALLFVRKESGALVAFDCVA
 NQLQELTNAPSIKOSVQSLGSRFFATDAGVIYAVDDELKIAETLKGVEAAAIGVSGASVIAAVQSSTASGT
 GDEAQCQGISRVLVQSASGVTEIAFPEQQQSGARGLVKIHVGDSSTGTTRAFVFEDASAVGIEJESGASEAS
 TLFVREEALANVVEAVAVDLPPTDEVGSLGDEAAHVFAHGSHASIFMFRLLKDQVRTVQRFVQSLFGAATQ
 HLSEFVASQOKTLVQAIRGELPRAESLSQSEMFSFGFRRLVLRASGKVFGLNSADGSLWAAQSPGSRLLF
 VTRAREAGLDHPAEVAIVDEAHGRVTWRNAITGAVTRVEDIDTPLAQIAVLPGDIFPSTASSEEDVSPAVAL
 IALDHAQRVHILPSSRTESVLQLEDLLRALHFVVYSNETGALTGYAVDPSQRAGVELWSMIVPASQTLLAV
 EGQSGGALNNPGIKRODGAVALVKFVDPHLLMVATQSGPHLQVSILNGISGRVISRFTHKKSTGPVHAVLAD
 NTVTYSFVNQVKSQREVSVVGLFEGEIGPRELNMWSSRPNMGSGKAMSAFDDSMMPNVQKFTYTERAI
 AALGVTKTRFGIADRRVLIGTANGAVNMQVPQILSPRRPVGKLSMEKEEGLMLYAPELPLIPTQTITYYES
 IPQLRLIRSFATRLESTSLVLAAGLDIFYTRVMPSRGFDVLDEDFASGLLLALIAALLALTIYLSKAVGKSTL
 DETWK (SEQ ID NO: 86)

果蛋白样

MGAARRSMGAARKALAASATLAALALAGLQPARAEVNGVNMTEAMLTEYASLPCVRSIARDGAVGCC
 SPSDRSVAEGGALFLVESVEDVTGLIENAQGLDAVALVDDALLHGDSLRLAMQDLAKKIRVTAVIVTVEE
 DGSPQEPFRSSAAPTTPWPSGDGLLNETVSFVTVRLRNATQSEEIRALAASNRDRGYVDAVFQHSARYQFY
 LGKETATSLSLASGRCDPLGGLSVWASAGPVVNSAKETVLLTANLDAASFFHDVVPARDTTASGVAAV
 LLAAKALASVDESLEALSKQIAVALFNGEVWSRAGSRFRVHDVALGECLSPQTASPYNESTCANPPVYAL
 AWTSLGLDNITDVSVNNVAGSEGAFFVHTAAGTASANAALQSVASSSTDVDVSIITGATTSGVPPSP
 LDSFLAAEMETDVSFSGAGLVVSGFDAATDANPRYSSRYDRRDKGPEADDAEALTAARIADVATLLARH
 AFVQAGGSISDAVNFVLVDGTHAAELWDCLTKDFACTLVADVIGAEDTTAVADFMGSTLLAASEGVAGG
 APNFFSGIYSPFPVENNVMPVPLFVRDYLAQYGRNASLIEKVTESAKYACAQDLDCMVMTEPPACELGRS
 ALACLRGGCVCSNAYFHDVSPALVYEDGAFSVDAQKLTDDDGWLWTEPRWSDGTLTLYTSANSASTIAL
 LVCGILLTIGCVFALRKAQGMLDNTKYKLN (SEQ ID NO: 87)

Emp24

MATTENEARLPQKQRLGRRRRGRVSKASGWGTTLALAAAVLVFSVDRASGVRFEEVASTEERCIFDVLRK
 DQLVTGEFEVHADGDDVNMDIHVTGPLGEEVFSKQNSKMAKFGFTAEEAAGEHVLCLRNNDMIMREVQV
 KLRSGVEAKDLTEVVQRHHLKPLSAEVIRIQETIRDVRHELTALKQREAE MRDMNESINTRVSLFSFPIAV
 VGSGLGAWQIMYLSYFQKRLI (SEQ ID NO: 88)

钙联接蛋白样

MRTTFVAAYAAVAALALGQCEAINFRESFEGANVEKEWKSASDRYAGSEWAFDTSKDTGDVGLQTVKP
 HKFYGISRKFNPIPVGDGEKPFVAQYEVKFTEGVSCSGAYLKLLEQDDAFTPKDLVESSPYSIMFGPDNCG
 ANNKVLHIFRQENPVTKYEYEEKHMTKKVTSVRDRTSHVYTLVHPDNTFKVKVDGKVEAEGSLTDDEAPS
 PPFQPKKEIDDPNDEKPDWVDQAKIPDPEASKPDWDEDAPKRIADPDVAVKPEGWLDDEPDQVPDPAAS
 EPEDWDEEDDGIWEAPLVANPKCTAGPGCGEWNAPMIENPNYKKGKWSAPMIDNPEYKGVWKPRRIENPA
 YFEESPVTTIKFIGAVAIELANDKGIRFDNIIIGNDVKEAAEFIDKEFLAKQADEKAKVKEEAAQAAQNSR
 WEEYKKGSIQGYVMWYAGDYIDYVMELYEASPIAVGVGAAAAGLAVLVALMVMCMMSGAPPEYDDDDV
 ALHKKDDDAAGDDDEAEAEAEAEADADEDEDEDEDEDEDEDEDEDEDEDEDEDEATGPRRRRVNAN (SEQ ID
 NO: 89)

图14

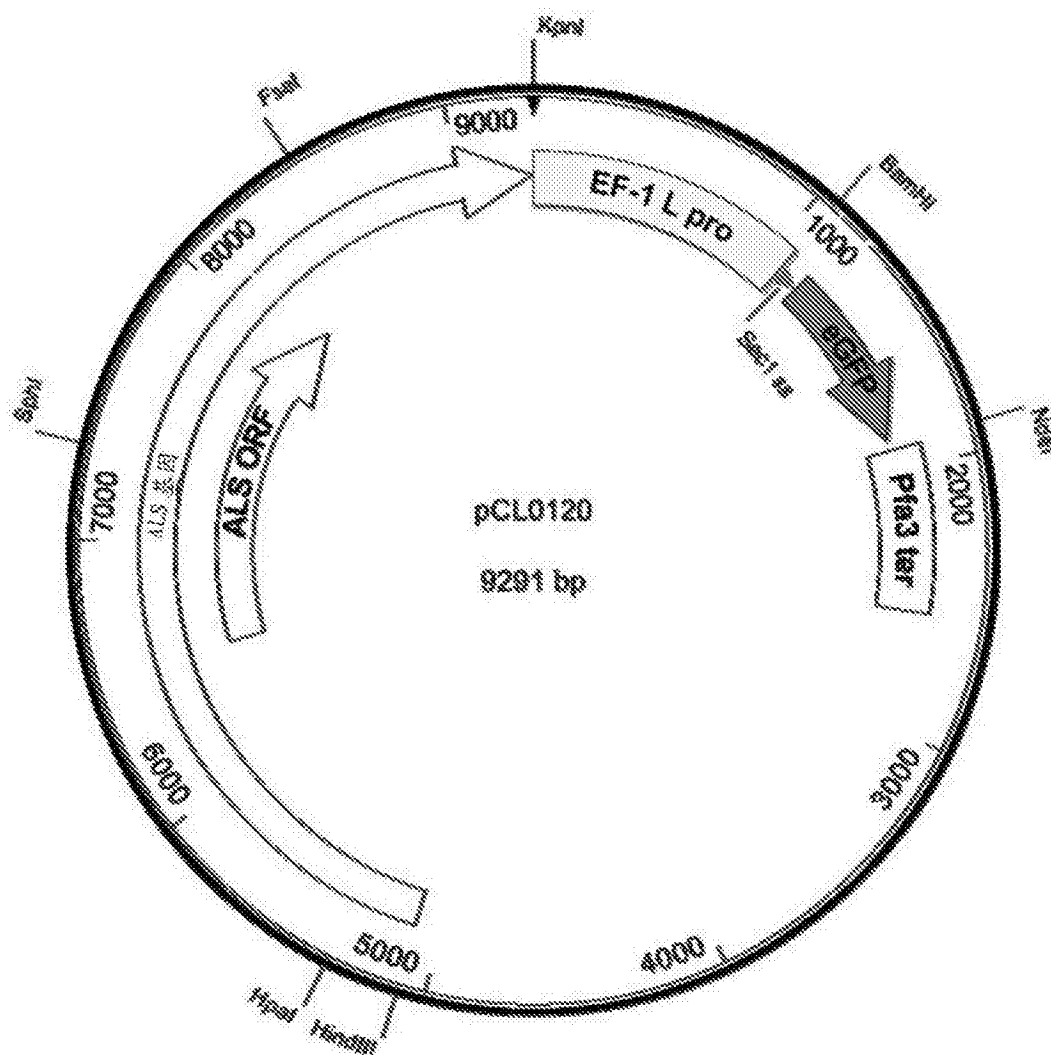


图15

AA*	密码子	分数	AA*	密码子	分数	AA*	密码子	分数	AA*	密码子	分数
Ala	GCC	0.64	终止	TAA	0.34	Leu	CTT	0.16	Ser	TCG	0.33
Ala	GCA	0.03	终止	TGA	0.33	Leu	TTG	0.02	Ser	TCC	0.31
Ala	GCT	0.18	终止	TAG	0.33	Leu	CTG	0.12	Ser	AGT	0.03
Ala	GCG	0.16	Gln	CAA	0.08	Leu	CTC	0.69	Ser	TCA	0
Arg	CGG	0.01	Gln	CAG	0.92	Leu	TTA	0	Ser	TCT	0.09
Arg	AGA	0	Glu	GAA	0.09	Leu	CTA	0	Thr	ACG	0.3
Arg	CGC	0.8	Glu	GAG	0.91	Lys	AAA	0.04	Thr	ACC	0.54
Arg	CGA	0.01	Gly	GGA	0.1	Lys	AAG	0.96	Thr	ACA	0.02
Arg	AGG	0	Gly	GGT	0.2	Met	ATG	1	Thr	ACT	0.14
Arg	CGT	0.17	Gly	GGG	0	Phe	TTT	0.45	Trp	TGG	1
Asn	AAC	0.94	Gly	GGC	0.7	Phe	TTC	0.55	Tyr	TAC	0.94
Asn	AAT	0.06	His	CAC	0.83	Pro	CCT	0.21	Tyr	TAT	0.06
Asp	GAT	0.24	His	CAT	0.17	Pro	CCG	0.34	Val	GTC	0.62
Asp	GAC	0.76	Ile	ATC	0.7	Pro	CCC	0.43	Val	GTA	0
Cys	TGC	0.95	Ile	ATA	0	Pro	CCA	0.02	Val	GTT	0.14
Cys	TGT	0.05	Ile	ATT	0.3	Ser	AGC	0.24	Val	GTG	0.24

*AA=氨基酸

图16

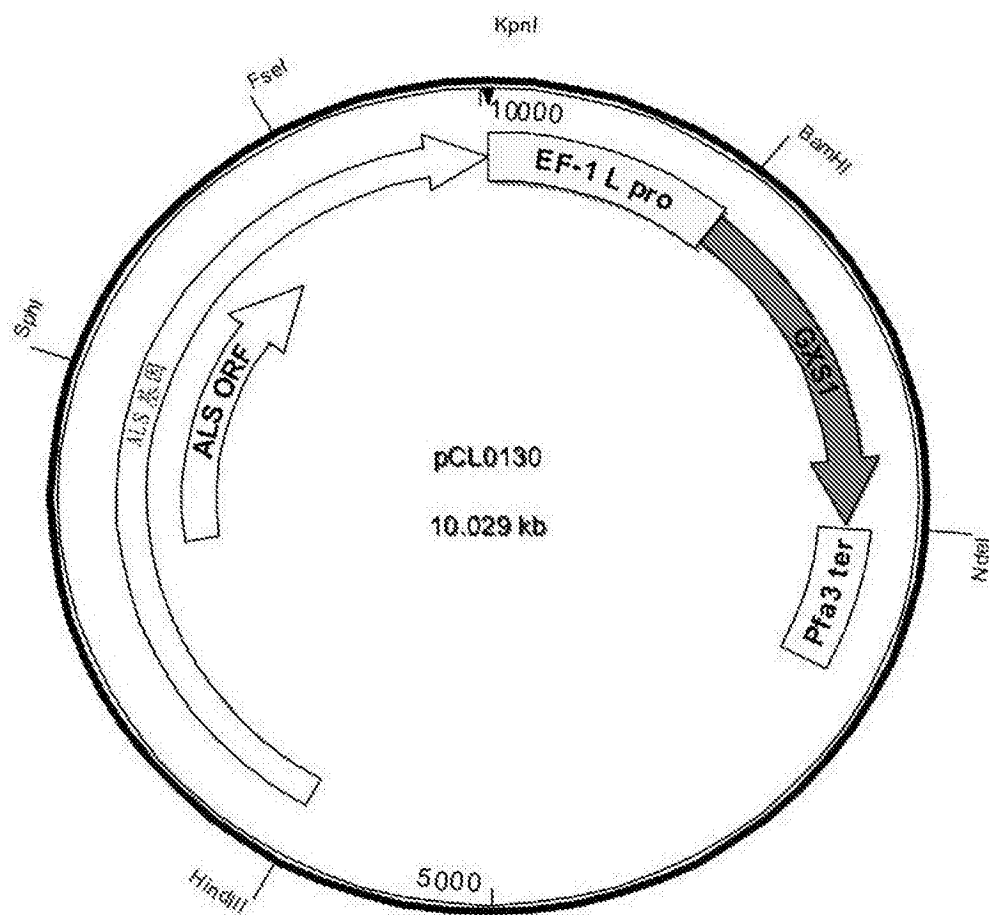


图17

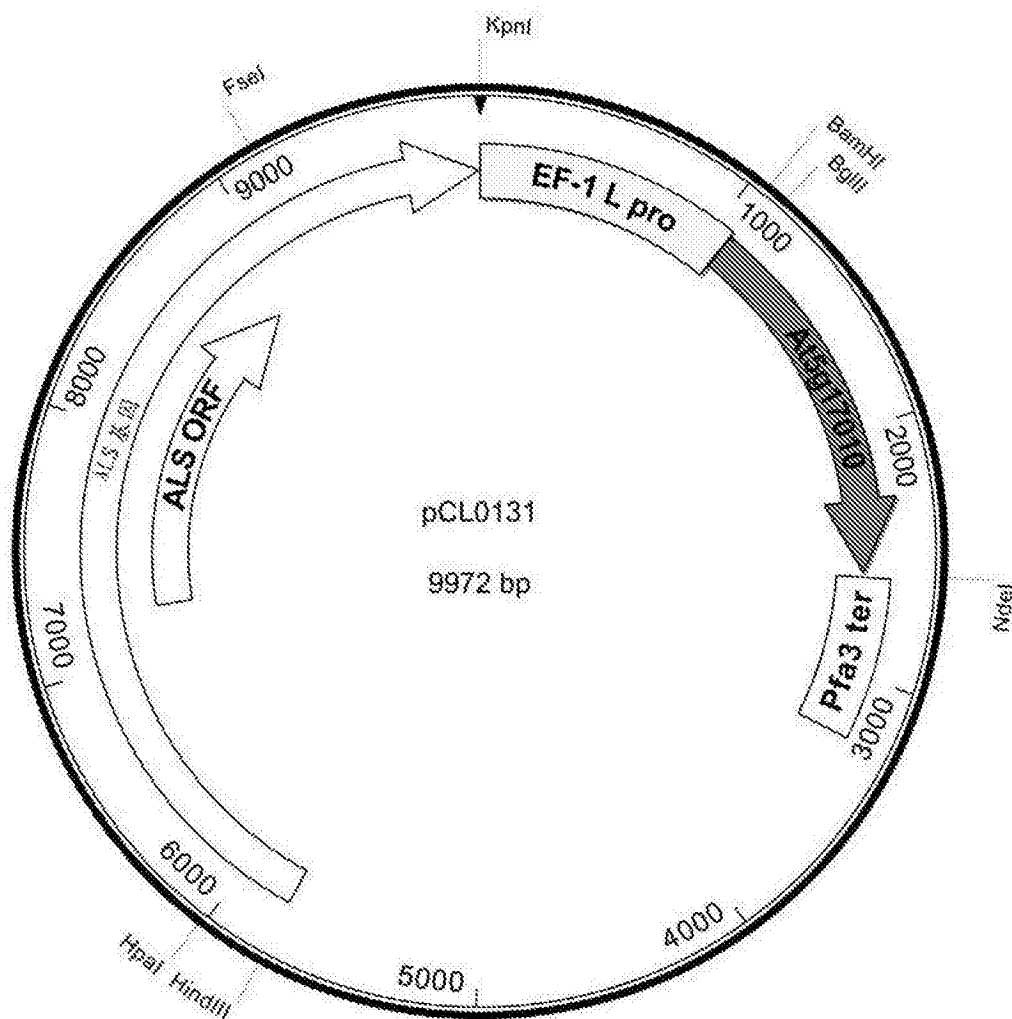


图18

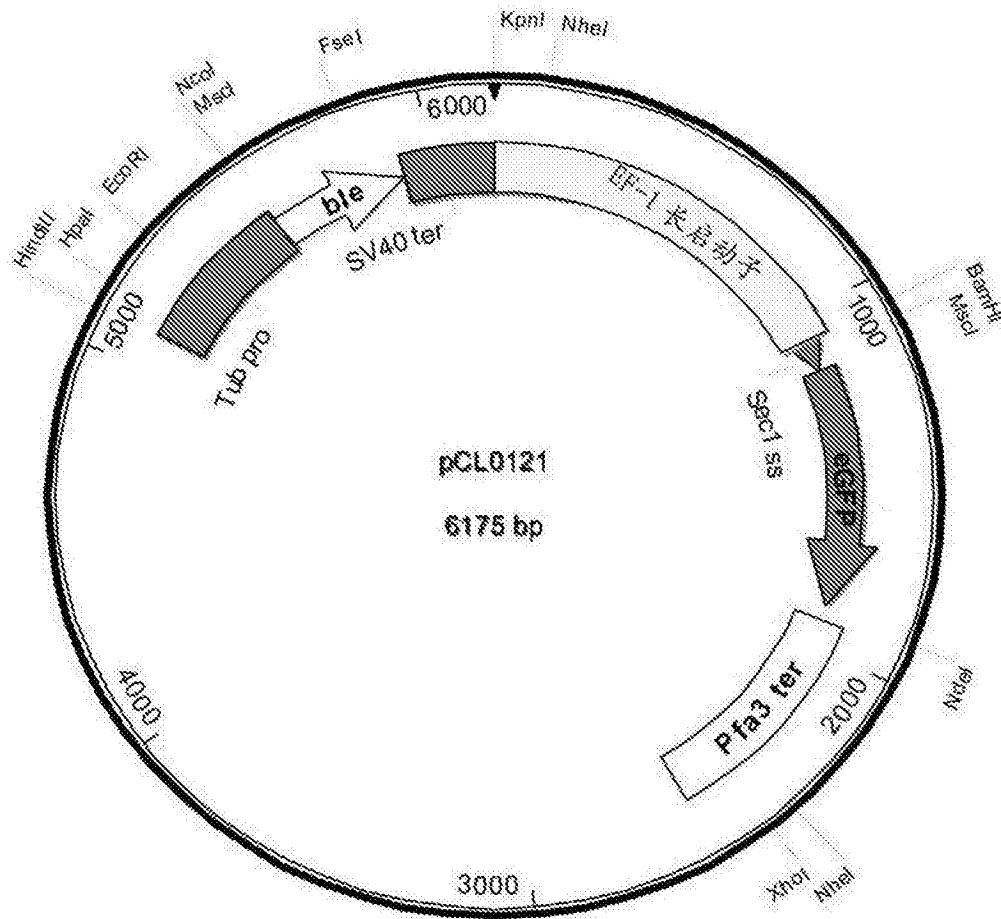


图19

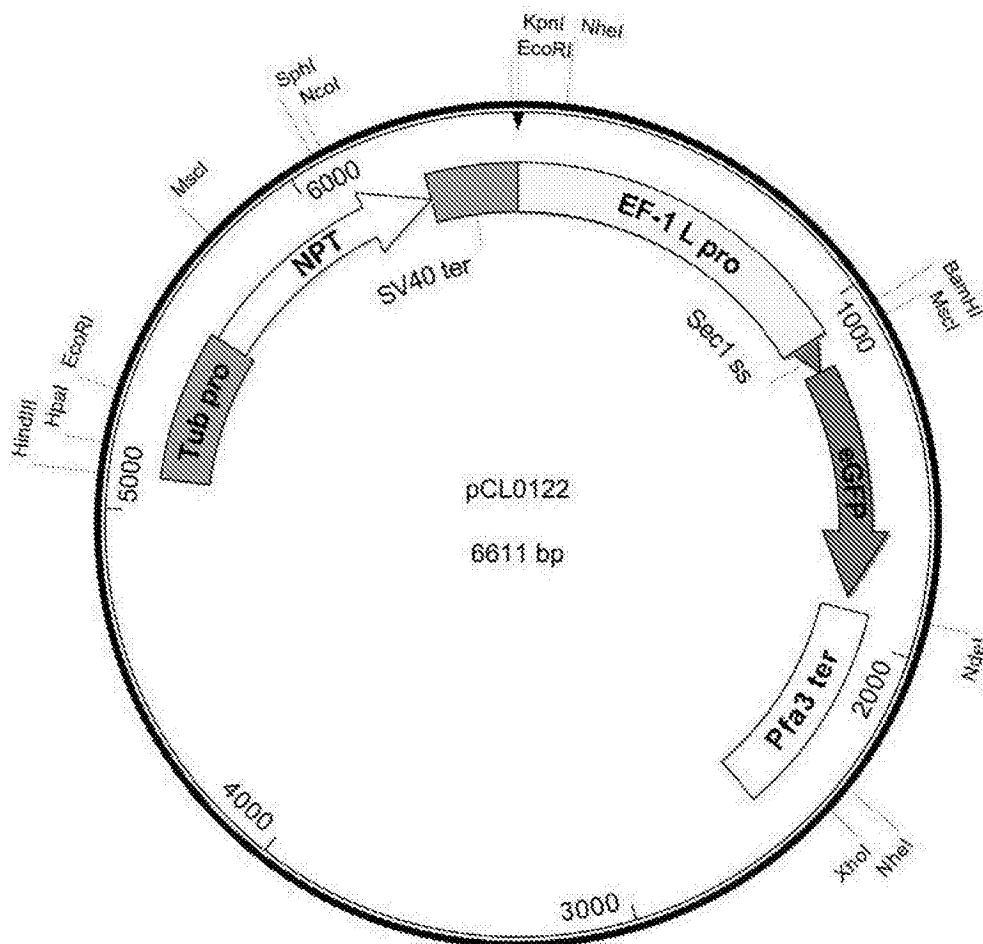


图20

ATGGCTAAGGAGTACTTCCCCCAGATCCAGAAGATTAAGTTCGAGGGTAAGGACAGCA
AGAACCCGCTCGCCTTTCATTACTACGACGCCGAGAAGGAGGTGATGGGCAAGAAG
ATGAAGGACTGGCTTCGCTTTGCTATGGCTTGGTGGCACACTCTCTGCGCTGAGGGCGC
GGACCAGTTTGGCGGCGGTACGAAGAGCTTTCGTTGGAACGAGGGCACTGACGCTAT
TGAGATTGCTAAGCAGAAGGTTGACGCTGGTTTCGAGATTATGCAGAAGCTCGGTATTCC
GTACTACTGCTTTCACGATGTCGACCTCGTTTCCGAGGGCAACTCGATCGAGGAGTACG
AGTCGAACCTCAAGGCTGTGGTTGCCTACCTCAAGGAGAAGCAGAAGGAGACCGGA
ATCAAGCTCCTCTGGAGCACCGCCAACGTTTTCGGCCACAAGCGCTACATGAACGGC
GCCTCCACCAACCCTGACTTCGATGTTGTTGCCCGCGCTATTGTCCAGATTAAAGAACGC
CATCGACGCTGGTATCGAGCTCGGAGCCGAGAAGTACGTTTTTGGGGCGGACGCGA
GGGTACATGTCCCTCCTCAACACCGACCAGAAGCGTGAGAAGGAGCACATGGCCA
CTATGCTTACCATGGCCCCGCGACTACGCCCCGAGCAAGGGTTTAAGGGTACTTTTCTCA
TTGAGCCGAAGCCCATGGAGCCGACCAAGCACCAGTACGACGTCGACACCGAGAC
CGCCATTGGCTTCCTTAAGGCCCAACCTTGACAAGGATTTTAAGGTGAACATCGAGG
TTAACCACGCTACGCTTGCCGGCCACACCTTTGAGCATGAGCTCGCCTGCGCTGTTGA
CGCCGGAATGCTTGGTTCATTGACGCCAACC GCGGCGACTACCAGAACGGCTGGG
ACACCGACCAGTTTCCGATTGACCAAGTACGAGCTCGTCCAGGCCTGGATGGAGATCAT
CCGTGGTGGAGGCTTTGTTACCGGTGGTACGAACTTCGACGCCAAGACGCGCCGTA
CAGCACGGACCTCGAGGACATCATCATTGCTCATGTGTCGGGCATGGACGCCATGGC
TCGCGCCCTTGAGAACGCTGCTAAGCTCCTCCAGGAGAGCCCCTACACGAAGATGA
AGAAGGAGCGCTACGCGTCGTTTGACAGCGGAATCGGTAAAGACTTCGAGGATGGC
AAGCTCACCTTGAGCAGGTGTACGAGTACGGTAAGAAGAACGGCGAGCCGAAGC
AGACCAGCGGCAAGCAGGAGCTCTACGAGGCCATTGTCGCCATGTACCAGTAG

(SEQ ID NO: 92)

图21

ATGAAGACCGTCGCCGGCATCGATCTTGGAACCCAGTCCATGAAGGTTGTCATTACGA
CTACGAGAAGAAGGAGATCATCGAGTCCGCCTCGTGCCCTATGGAGCTCATTAGCGA
GTCGGACGGAACCCGCGAGCAGACGACTGAGTGGTTTGACAAGGGTCTCGAGGTGT
GCTTTGGAAAGCTCTCCGCTGATAACAAGAAGACCATTGAGGCGATTGGCATCTCCGG
CCAGCTCCACGGCTTCGTCCCTCTCGATGCGAACGGAAAGGCGCTCTACAACATCAA
GCTCTGGTGCGACACCGCCACTGTGGAGGAGTGCAAGATCATTACTGACGCCGCCG
GCGGCGACAAGGCTGTCATCGACGCGCTCGGCAACCTCATGCTCACCGGATTCACC
GCCCCGAAGATTCTCTGGCTCAAGCGCAACAAGCCCGAGGGCCTTTGCTAACCTCAAG
TACATTATGCTGCCCCACGATTACCTCAACTGGAAGCTGACTGGAGACTACGTCATGGA
GTACGGCGACGCCCTCCGGCACCGCCCTTTTGGATTGCAAGAACCGCTGCTGGTCGAA
GAAGATTTGCGACATTATTGATCCTAAGCTGCTCGACCTTCTCCCTAAGCTCATTGAGCCC
TCGGCCCCCGCCGGTAAGGTCAACGACGAGGCCGCCAAGGCGTACGGCATTCCC
GCCGGAATCCCCGTTTCCGCTGGCGGCGGTGATAACATGATGGGTGCGGTGCGTACT
GGCACCGTCGCTGACGGATTCTCACGATGAGCATGGGACACCTCCGGAACCTCTTAC
GGCTACTCGGACAAGCCTATTTCCGACCCGGCTAACGGCCTCAGCGGCTTCTGCAGC
TCCACGGGGCGGCTGGCTTCCCCTCCTTTCACCATGAACCTGCACCGTCGCCACCGAG
TTCGTCCGCAACCTTTTTCAGATGGATATCAAGGAGCTGAACGTGAGGGCTGCTAAGTCC
CCCTGCGGCAGCGAGGGCGTTCTTGTCATTCTTTCTTCAACGGCGAGCGCACCCCGA
ACCTCCCCAACGGCCGCGCCTCGATTACCGGCCTCACCTCCGCGAACACGTCCTCG
CGCCAACATCGCTCGCGCCTCCTTTGAGTCGGCCGCTTTTGCCATGCGCGGTGGCCTC
GATGCGTTTCGTAAGCTCGGATTCCAGCCCAAGGAGATTGCGCTCATCGGCGGTGGTTC
GAAGTCCGACCTCTGGCGCCAGATCGCTGCTGACATTATGAACCTTCCCATCCGTGTCC
CCCTTCTCGAGGAGGCCGCCGCCCTCGGCGGAGCTGTCCAGGCCCTTTGGTGCCCTTA
AGAACCAGTCCGGTAAGTGCGACATCGTCGAGCTTTGCAAGGAGCATATCAAGATTGA
CGAGTCCAAGAACGCCAACCCGATTGCCGAGAACGTCGCCGTGTACGATAAGGCCT
ACGATGAGTACTGCAAGGTGCTTAACACGCTCAGCCCTCTGTACGCCTAA

(SEQ ID NO: 93)

图22

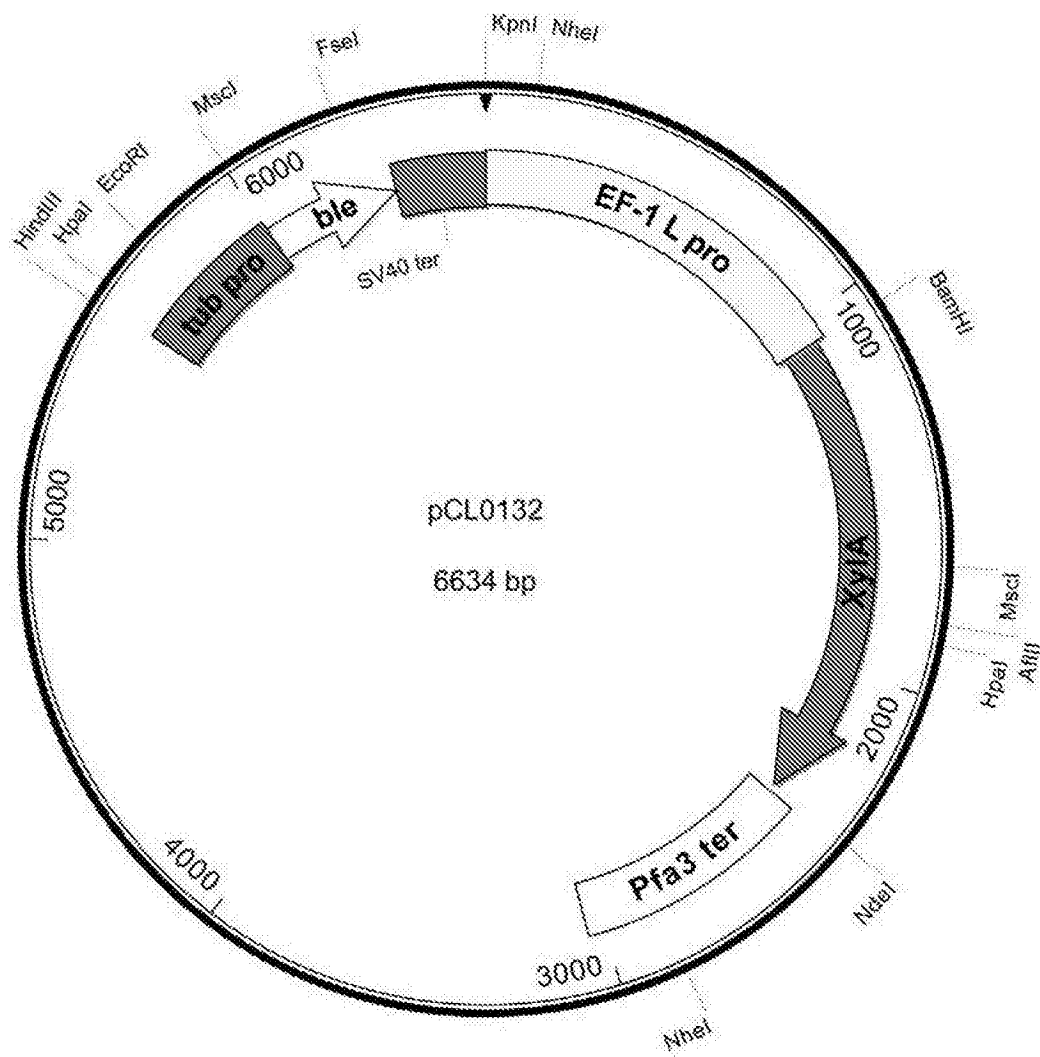


图23

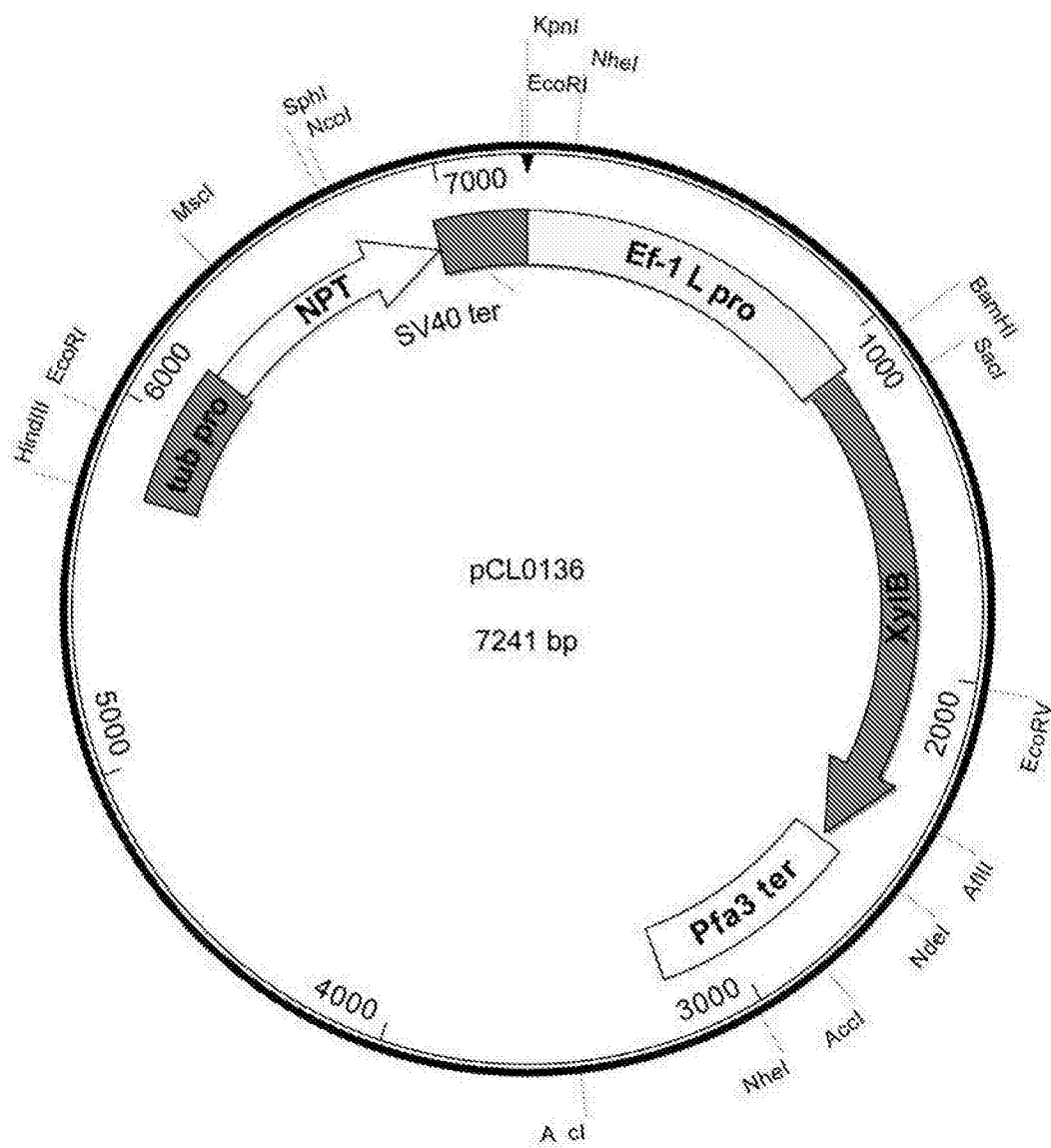


图24

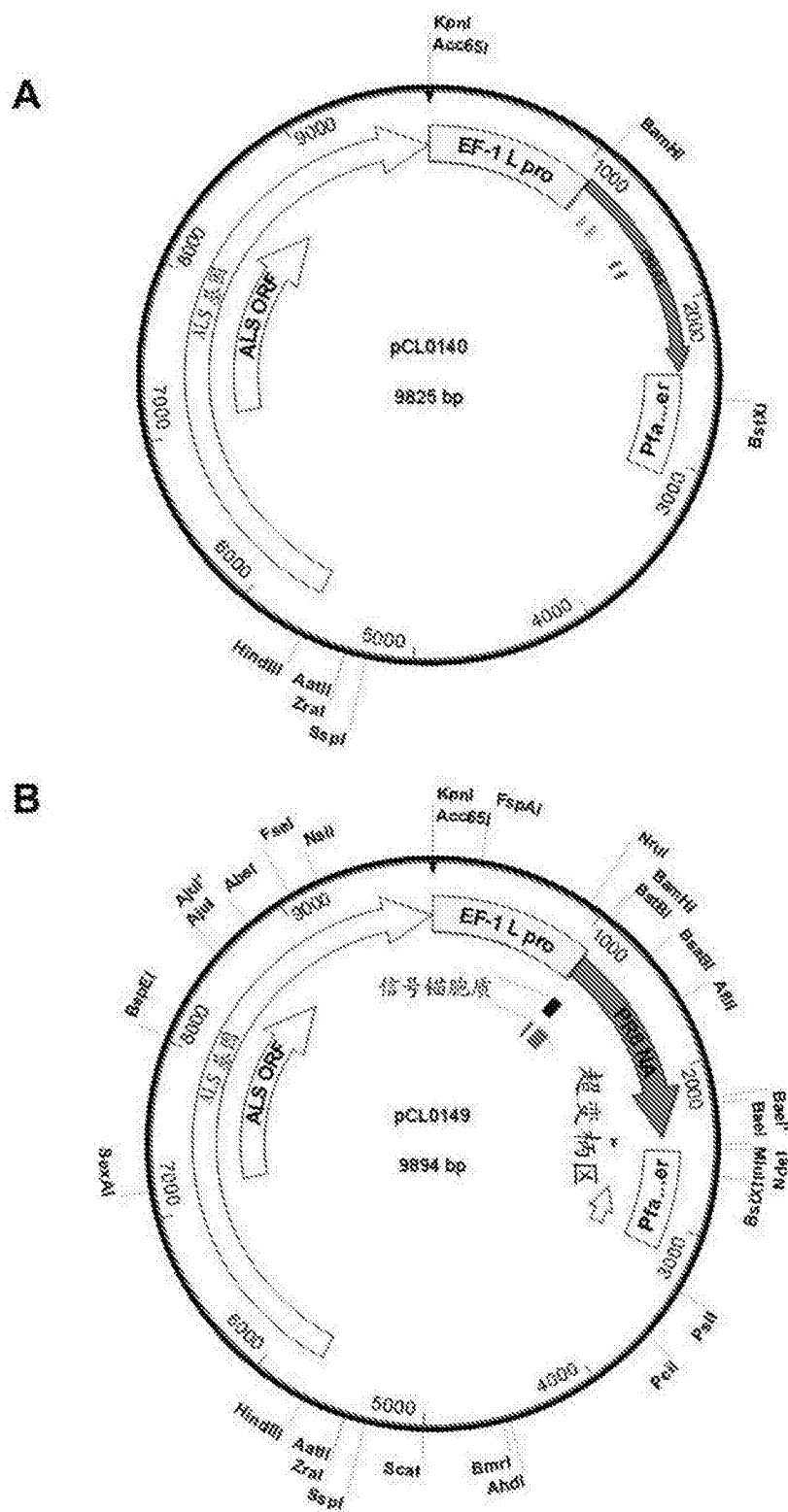


图25

ATGAACCCCAACCAGAAGATTACTACTATCGGTAGCATTTCGCTCGTCGTTGGACTTATCTC
CCTTATTCTTCAGATTGGTAACATTATCTCCATTGATCTCGCATAGCATTTCAGACCGGCT
CCCAGAACCACACCGGCATTGCAACCAGAACATTATTACTTACAAGAAGTCCACTTGGGT
CAAGGACACTACTAGCGTTATTCTTACCGGTAACCTCGTCGCTTTGCCCTATTTCGCGGCTGG
GCTATTTACAGCAAGGACAACCTCGATCCGCATCGGTAGCAAGGGCGACGTTTTTGTATCC
GTGAGCCTTTTATTTCTGAGCCACCTCGAGTGCCGTAATTTTTTCTGACTCAGGGCGC
TCTCCTCAACGATAAGCATTCCAACGGCACTGTCAAGGATCGCAGCCCTACCGCGCCCT
TATGTCCTGCCCTGTGCGCGAGGCTCCCAGCCCTACAACCTCCGTTTTGAGTCCGTTGC
CTGGTCCGCGCAGCGCTGCCACGACGGAATGGGATGGCTCACTATTGGTATTTCCGGCCC
TGATAACGGCGCTGTGCGCGTCCTTAAGTACAACGGCATTATCACCGAGACCATCAAGTCC
TGGCGTAAGAAGATCCTCCGCACCCAGGAGTCCGAGTGCGCCTGCGTCAACGGCAGCTG
CTTCACGATTATGACCGACGGCCCTCCGACGGCCTCGCTTCCTACAAGATTTTAAGATT
GAGAAGGGTAAGGTCACGAAGTCCATCGAGCTTAACGCCCCGAAGTCCACTACGAGGAG
TGCTCCTGCTACCTGACACTGGCAAGGTGATGTGCGTCTGCCGCGATAACTGGCATGGC
TCCAACCGCCCTGGGTTAGCTTCGATCAGAACCTTGACTACCAGATTGGATACATTTGCT
CCGGTGTTTTTGGCGACAACCCGCGCCCGAGGATGGAACTGGTTTCGTGCGGTCCTGTTT
ACGTTGACGGCGCCAACGGCGTTAAGGGTTTTTCTACCGTTACGGTAACGGAGTCTGGA
TCGGCCGCGACCAAGTCGCACAGCTCGCGCCACGGATTTGAGATGATCTGGGACCCCAAC
GGATGGACTGAGACCGATTCCAAGTTTAGCGTTCCGCCAGGATGTCGTTGCTATGACCGATT
GGTCGGGATACTCCGGTTCCTTTGTGCAGCACCTGAGCTCACCGGCCTTGACTGCATGC
GCCCTTGCTTTTGGGTGAGCTCATTGCGGTCGCCCTAAGGAGAAGACTATTTGGACCT
CCGCCAGCAGCATTTCCTTTTGGCGCGTTAACTCCGACACCGTCGACTGGTCGTGGCCCG
ATGGCGCCGAGCTTCCTTTTCCATTGATAAG (SEQ ID NO: 100)

ATGAACCCCAACCAGAAGATTACTACTATCGGTAGCATTTCGCTCGTCGTTGGACTTATCTC
CCTTATTCTTCAGATTGGTAACATTATCTCCATTGATCTCGCATAGCATTTCAGACCGGCT
CCCAGAACCACACCGGCATTGCAACCAGAACATTATTACTTACAAGAAGTCCACTTGGGT
CAAGGACACTACTAGCGTTATTCTTACCGGTAACCTCGTCGCTTTGCCCTATTTCGCGGCTGG
GCTATTTACAGCAAGGACAACCTCGATCCGCATCGGTAGCAAGGGCGACGTTTTTGTATCC
GTGAGCCTTTTATTTCTGAGCCACCTCGAGTGCCGTAATTTTTTCTGACTCAGGGCGC
TCTCCTCAACGATAAGCATTCCAACGGCACTGTCAAGGATCGCAGCCCTACCGCGCCCT
TATGTCCTGCCCTGTGCGCGAGGCTCCCAGCCCTACAACCTCCGTTTTGAGTCCGTTGC
CTGGTCCGCGCAGCGCTGCCACGACGGAATGGGATGGCTCACTATTGGTATTTCCGGCCC
TGATAACGGCGCTGTGCGCGTCCTTAAGTACAACGGCATTATCACCGAGACCATCAAGTCC
TGGCGTAAGAAGATCCTCCGCACCCAGGAGTCCGAGTGCGCCTGCGTCAACGGCAGCTG
CTTCACGATTATGACCGACGGCCCTCCGACGGCCTCGCTTCCTACAAGATTTTAAGATT
GAGAAGGGTAAGGTCACGAAGTCCATCGAGCTTAACGCCCCGAAGTCCACTACGAGGAG
TGCTCCTGCTACCTGACACTGGCAAGGTGATGTGCGTCTGCCGCGATAACTGGCATGGC
TCCAACCGCCCTGGGTTAGCTTCGATCAGAACCTTGACTACCAGATTGGATACATTTGCT
CCGGTGTTTTTGGCGACAACCCGCGCCCGAGGATGGAACTGGTTTCGTGCGGTCCTGTTT
ACGTTGACGGCGCCAACGGCGTTAAGGGTTTTTCTACCGTTACGGTAACGGAGTCTGGA
TCGGCCGCGACCAAGTCGCACAGCTCGCGCCACGGATTTGAGATGATCTGGGACCCCAAC
GGATGGACTGAGACCGATTCCAAGTTTAGCGTTCCGCCAGGATGTCGTTGCTATGACCGATT
GGTCGGGATACTCCGGTTCCTTTGTGCAGCACCTGAGCTCACCGGCCTTGACTGCATGC
GCCCTTGCTTTTGGGTGAGCTCATTGCGGTCGCCCTAAGGAGAAGACTATTTGGACCT
CCGCCAGCAGCATTTCCTTTTGGCGCGTTAACTCCGACACCGTCGACTGGTCGTGGCCCG
ATGGCGCCGAGCTTCCTTTTCCATTGATAAGGGTAAGCCTATCCCTAACCTCTCCTCG
TCTCGATTCTACGCGTACCGGTCATCATCACCATCACCAT (SEQ ID NO: 101)

图26

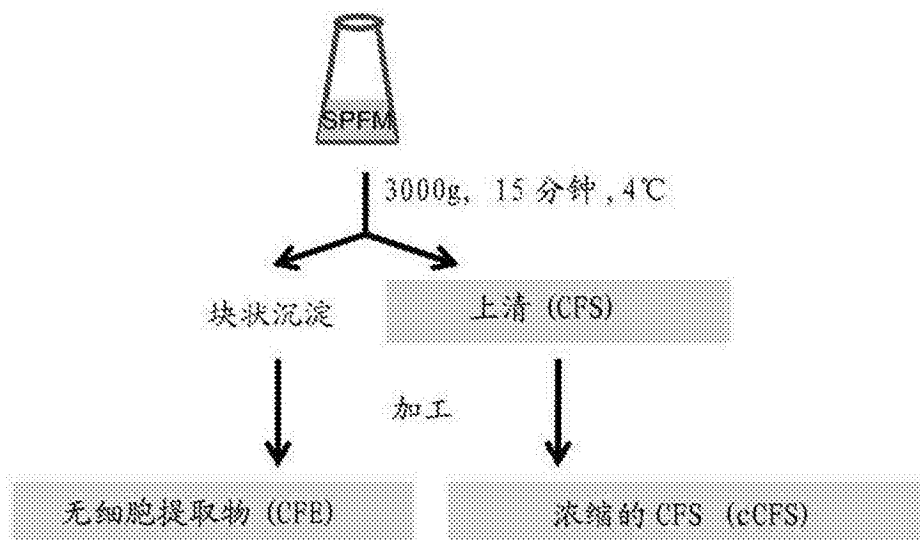


图27

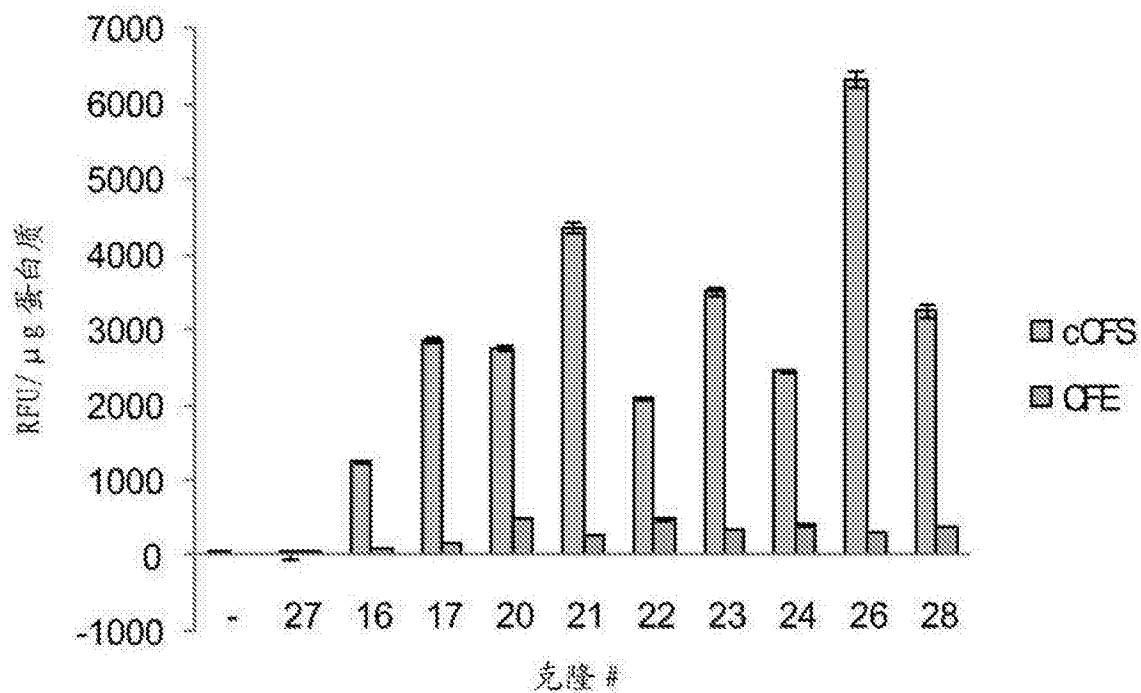


图28

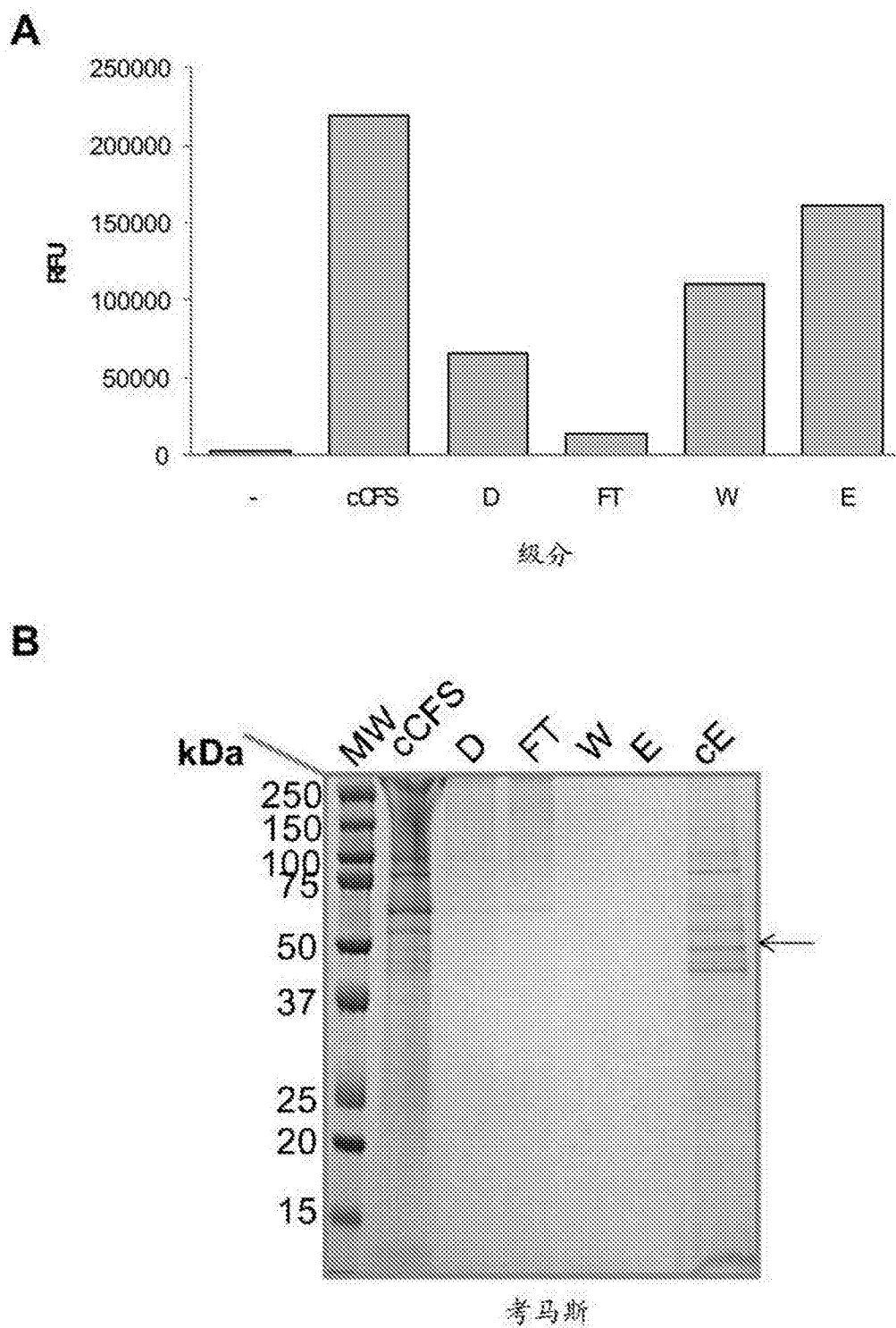


图29

Species: 甲型流感病毒 (A/Puerto Rico/8/34/Mount Sinai (H1N1))

名称: 神经氨酸酶

```
1  MNPNQKIITI  GSICLVVGLI  SLILQIGNII  SIWISHSIQT  GSNHTGICN
QNIITYKNST
61  WVKDTTSVIL  TGNSSLCPIR  GWAIYSKDNS  IRIGSKGDVF  VIREPFISCS
HLECRTFFLT
121 QGALLNDRHS  NGTVKDRSPY  RALMSCPVGE  APSPYNSRFE  SVAWSASACH
DGMGWLTIGI
181 SGPDNGAVAV  LKYNGIITET  IKSWRKKILR  TQESECACVN  GSCFTIMTDG
PSDGLASYKI
241 FKIEKGKVTK  SIELNAPNSH  YEECSCYPDT  GKVMCVCRDN  WHGSNRPWVS
FDQNLDYQIG
301 YICSGVFGDN  PRPKDGTGSC  GPVYVDGANG  VKGFSYRYGN  GVWIGRTKSH
SSRHGFEMIW
361 DPNGWTETDS  KFSVRQDVVA  MTDWSGYSGS  FVQHPELTGL  DCIRPCFWVE
LIRGRPKEKT
421 IWTSASSISF  CGVNSDTVDW  SWPDGAELPF  TIDK
```

(SEQ ID NO: 100)

图30

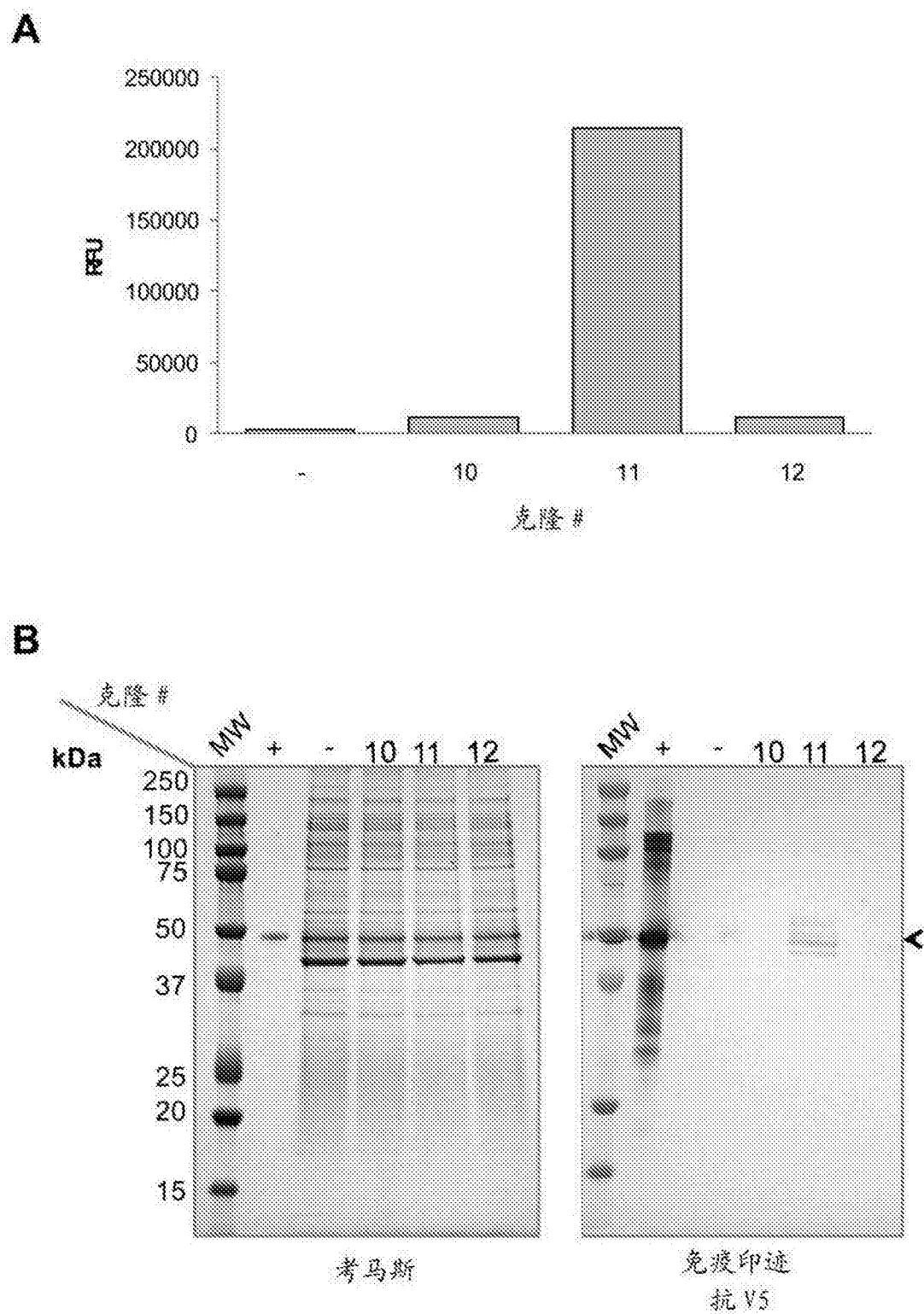


图31

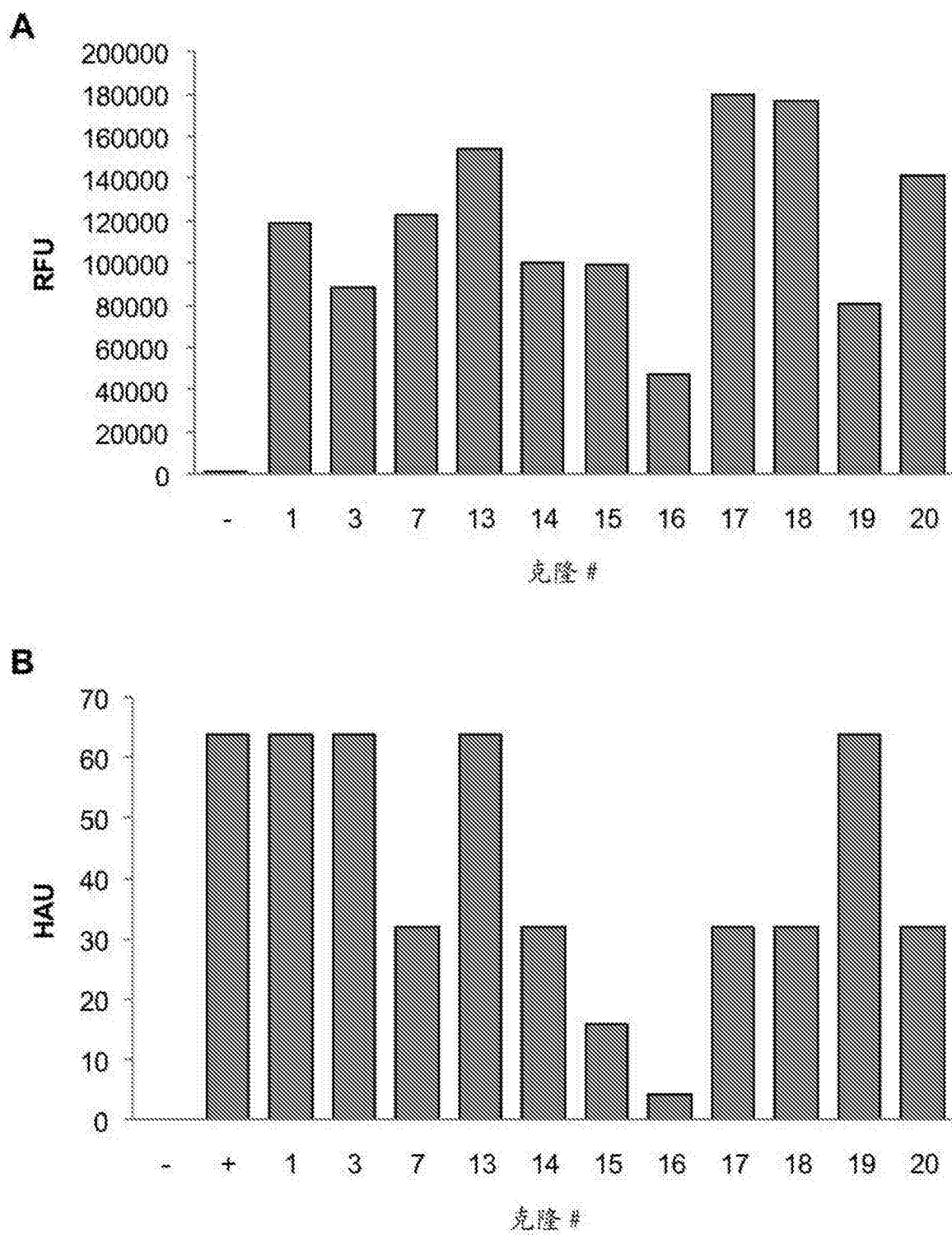


图32