



(19)中華民國智慧財產局

(12)發明說明書公告本

(11)證書號數：TW I504749 B

(45)公告日：中華民國 104 (2015) 年 10 月 21 日

(21)申請案號：099107452

(22)申請日：中華民國 99 (2010) 年 03 月 15 日

(51)Int. Cl. : C12N15/87 (2006.01)

C12P21/00 (2006.01)

(30)優先權：2009/03/16 美國

61/160,618

2009/12/28 美國

61/290,441

(71)申請人：D S M智慧財產有限公司 (荷蘭) DSM IP ASSETS B. V. (NL)

荷蘭

(72)發明人：阿普特 卡克 E APT, KIRK E. (US)；利普麥爾 詹姆斯 C LIPPMEIER, JAMES CASEY (US)；王鈞 WANG, JUN (US)；威恩 詹姆斯 P WYNN, JAMES P.

(US)；札可 羅斯 ZIRKLE, ROSS (US)；辛普森 大衛 SIMPSON, DAVID (US)

(74)代理人：憚軼群；陳文郎

(83)生物材料寄存：

財團法人食品工業發展研究所	BCRC 940588	2010 年 05 月 27 日
財團法人食品工業發展研究所	BCRC 940586	2010 年 05 月 27 日
財團法人食品工業發展研究所	BCRC 940589	2010 年 05 月 27 日
財團法人食品工業發展研究所	BCRC 940587	2010 年 05 月 27 日
財團法人食品工業發展研究所	BCRC 940590	2010 年 05 月 27 日
財團法人食品工業發展研究所	BCRC 930128	2010 年 05 月 06 日
財團法人食品工業發展研究所	BCRC 930129	2010 年 05 月 06 日
財團法人食品工業發展研究所	BCRC 930130	2010 年 05 月 06 日
財團法人食品工業發展研究所	BCRC 930131	2010 年 05 月 06 日
財團法人食品工業發展研究所	BCRC 930133	2010 年 06 月 10 日

(56)參考文獻：

AU 2008254837B2

審查人員：吳姵諄

申請專利範圍項數：18 項 圖式數：45 共 190 頁

(54)名稱

於網黏菌門微生物中生產蛋白質之技術

PROTEIN PRODUCTION IN MICROORGANISMS OF THE PHYLUM LABYRINTHULOMYCOTA

(57)摘要

本發明係關於網黏菌門(Labyrinthulomycota)之重組細胞及微生物及其用於異源蛋白質製造之用途。本發明也涵蓋自重組宿主細胞及微生物有效製造及選擇性地分泌多肽用的新穎啟動子、終結子、及信號序列。

The present invention relates to recombinant cells and microorganisms of the phylum Labyrinthulomycota and their use in heterologous protein production. Novel promoter, terminator, and signal sequences for efficient production and, optionally, secretion of polypeptides from recombinant host cells and microorganisms are also encompassed by the present invention.

六、發明說明：

【發明所屬之技術領域】

發明領域

本發明係關於網黏菌門(Labyrinthulomycota)之重組細胞及微生物及其用於異源蛋白質製造之用途。本發明也涵蓋自重組宿主細胞及微生物有效製造及選擇性地分泌多肽用的新穎啟動子、終結子、及信號序列。

【先前技術】

發明背景

生物技術及分子生物學的進展已經允許於微生物、植物及動物細胞中製造蛋白質，其中多種蛋白質先前只可自人及其它動物的組織、血液或尿液萃取得。今日市面上可得之生物製劑典型係於哺乳動物細胞諸如中國倉鼠卵巢(CHO)細胞製造，或於微生物細胞諸如酵母或大腸桿菌(E. coli)細胞系製造。

透過微生物發酵來製造蛋白質具有優於既有系統諸如，植物及動物細胞培養的若干優勢。舉例言之，基於微生物發酵法可提供：(i)快速製造高濃度蛋白質；(ii)可使用無菌且經良好控制之製造條件(諸如優良製造規範(GMP)條件)；(iii)可使用簡單的化學上經界定的生長培養基，允許更單純發酵及較少雜質；(iv)不存在有污染之人或動物病原；及(v)容易回收蛋白質(例如透過自發酵培養基分離而回收)。此外，發酵設施的建造典型比細胞培養設施的建造成本低。

美國公告案第2003/0166207號(現為美國專利案第7,001,772號)首次揭示適合用於轉形破囊壺菌(thraustochytrids)包括裂殖壺菌屬(裂殖壺菌屬)之成員之重組構成體。本公告案也揭示裂殖壺菌屬的乙醯乳酸合成酶、乙醯乳酸合成酶啟動子及終結子區、 α -微管素(tubulin)啟動子、得自聚縮酮合成酶(PKS)系統之啟動子、及脂肪酸去飽和酶啟動子之核酸序列及胺基酸序列。美國專利公告案第2006/0270275904及2006/0286650號，全文皆係以引用方式併入此處，隨後揭示裂殖壺菌屬的肌動蛋白(actin)、延長因子1 α (efl α)、及甘油醛3-磷酸去氫酶(gadph)啟動子及終結子之序列。

持續需要識別於破囊壺菌微生物表現蛋白質之額外調節控制元體，包括於不同時間點或於某些生長條件下或回應於化學刺激或環境刺激有不同表現之調節控制元體。也需要識別可有效指導自網黏菌門及破囊壺菌目(Thraustochytriales)，諸如裂殖壺菌屬及其它破囊壺菌之微生物分泌蛋白質之分泌信號序列。

【發明內容】

發明概要

本發明係針對包含SEQ ID NO:3多核苷酸序列之一已單離之核酸分子。

本發明也係針對包含SEQ ID NO:4多核苷酸序列之一已單離之核酸分子。

本發明也係針對包含編碼多肽之多核苷酸序列之一已

單離之核酸分子，其中該多肽包含SEQ ID NO:1之胺基酸序列。於若干實施例中，該多核苷酸序列包含SEQ ID NO:2。

本發明也係針對包含編碼多肽之多核苷酸序列之一已單離之核酸分子，其中該多肽包含SEQ ID NO:37之胺基酸序列。於若干實施例中，該多核苷酸序列包含SEQ ID NO:38。

本發明亦係針對包含SEQ ID NO:42之多核苷酸序列之一已單離之核酸分子。於若干實施例中，該多核苷酸序列包含SEQ ID NO:43。

本發明亦係針對包含SEQ ID NO:44之多核苷酸序列之一已單離之核酸分子。於若干實施例中，該多核苷酸序列包含SEQ ID NO:45之多核苷酸序列。

本發明亦係針對一種包含SEQ ID NO:46之多核苷酸序列之已單離之核酸分子。

本發明亦係針對一種包含與前述任一種多核苷酸序列完全互補之多核苷酸序列之已單離之核酸分子。

本發明亦係針對包含前述任一種已單離之核酸分子之重組核酸分子。於若干實施例中，該重組核酸分子為載體。

於若干實施例中，該已單離之核酸分子係工作式地鏈接至編碼蛋白質之多核苷酸序列。於若干實施例中，該蛋白質係工作式地鏈接至分泌信號。

本發明亦係針對包含前述任一種已單離之核酸分子或重組核酸分子或其組合之宿主細胞。於若干實施例中，該宿主細胞為破囊壺菌目之一成員。於若干實施例中，該宿

主細胞為裂殖壺菌屬或破囊壺菌屬(*Thraustochytrium*)。

本發明亦係針對一種用於製造蛋白質之方法，包含於培養基中培養破囊壺菌目重組微生物，其中該重組微生物包含前述已單離之核酸分子中之任一種工作式地鏈接編碼該蛋白質之多核苷酸序列來製造該蛋白質。於若干實施例中，該蛋白質係自己單離之破囊壺菌目生質回收。於若干實施例中，該蛋白質係蓄積於微生物體內部。於若干實施例中，該蛋白質係蓄積於微生物之膜內。於若干實施例中，該蛋白質係自培養基中回收。於若干實施例中，該蛋白質係分泌出。

本發明亦係針對包含SEQ ID NO:1之胺基酸序列之一已單離之多肽。

本發明亦係針對包含SEQ ID NO:15之胺基酸序列之一已單離之多肽。

圖式簡單說明

第1圖顯示裂殖壺菌屬Na/Pi-IIIb2轉運子蛋白質信號肽胺基酸序列(SEQ ID NO:1)。

第2圖顯示編碼SEQ ID NO:1之信號肽之多核苷酸序列(SEQ ID NO:2)。

第3圖顯示裂殖壺菌屬PUFA PKS OrfC啟動子區之多核苷酸序列(SEQ ID NO:3)。

第4圖顯示裂殖壺菌屬PUFA PKS OrfC終結子元件-1之多核苷酸序列(SEQ ID NO:4)。

第5圖顯示pSchizE之質體拼圖。

第6圖顯示pSchiz-sG之質體拼圖，也稱作為pCO0001。

第7圖顯示pSchiz-sGr之質體拼圖。

第8圖顯示pSchiz-cG之質體拼圖。

第9圖顯示於裂殖壺菌屬細胞之胞質及內質網(ER)的eGFP表現。第9A、9C及9E圖為螢光顯微相片。第9B、9D及9F圖為光學顯微相片。第9A及9B圖顯示以pSchiz-sGr轉形之細胞的相同視野。第9C及9D圖顯示以pSchiz-cG轉形之細胞的相同視野。第9E及9F圖顯示以pSchiz-E轉形之細胞的相同視野。

第10A圖顯示於經以pSchiz-sGr轉形之裂殖壺菌屬細胞中經ER靶定之eGFP及核酸專一性染色4',6-二脒-2-苯基吡啶(DAPI)之複合螢光定位。第10B及10C圖分別顯示用於製造複合顯微相片之eGFP-ER染色及DAPI-核染色。第10D圖顯示相同視野之光學顯微相片。如第10A圖指示，ER膜包住細胞之各個核，各細胞含有多個核。指示一個細胞中一個核之相關特徵。

第11圖顯示得自四個裂殖壺菌屬轉形體純株之不含細胞之上清液及不含細胞之萃取樣本之西方墨點及相對應之經庫馬西(Coomassie)-染色之SDS-PAGE凝膠。「sup」及「cyto」分別係指不含細胞之上清液及不含細胞之萃取物。「sG」係指得自以pSchiz-sG轉形細胞之樣本。「sGr」係指得自以pSchiz-sGr轉形細胞之樣本。「cG」係指得自以pSchiz-cG轉形細胞之樣本。「vec(-)」係指得自以pSchiz-E10轉形細胞之樣本。「GFP(+)」係指已純化之重組GFP標準品

(克隆泰公司(Clontech)，加州山景市)。凝膠載荷以3微克不含細胞之上清液蛋白質及1微克不含細胞之萃取物蛋白質樣本。於各對樣本亦即重組GFP標準品與分子量標記間發現有一空白線道。

第12圖顯示裂殖壺菌屬Sec1蛋白質轉運子蛋白質之首30個胺基酸。胺基酸1至20組成信號肽序列(SEQ ID NO:37)。

第13圖顯示編碼SEQ ID NO:37之信號肽之多核苷酸序列(SEQ ID NO:38)。

第14圖顯示pSchiz-Cpt-sleGFP之質體拼圖，也稱作為pCL0001。

第15A圖顯示已分泌之eGFP蛋白質之西方墨點，及第15B圖顯示得自於不同發酵條件(例如「B26」、「B27」或「B28」發酵條件，如第15圖定義)下生長的三個裂殖壺菌屬培養之相對應經庫馬西染色之SDS-PAGE凝膠。線道1-19載荷指示量之蛋白質。LH=發酵時間，單位為小時。第15A圖之線道20載荷10奈克，及第15B圖之線道20載荷0.5奈克已純化之重組GFP標準品；得自裂殖壺菌屬之eGFP帶由於含有聯結子序列故該帶係略大於對照帶。

第16圖顯示pSchiz-Cpt-slkappabh之質體拼圖。

第17圖顯示藉裂殖壺菌屬分泌κ抗體亞單位之西方墨點。L=不含細胞之萃取物；S=不含細胞之上清液。

第18A圖顯示表現κ抗體亞單位之西方墨點。獲得培養上清液樣本之培養時間係指示於第18圖頂。「wt」係指「野

生型」(亦即未經轉形之)裂殖壺菌屬。「+cptS1kappa」係指以含有編碼人 κ 抗體片段之密碼子最佳化基因、ORFC啟動子及終結子、及編碼Sec1信號肽之序列之載體轉形之裂殖壺菌屬。第18B圖顯示於以cptS1kappa轉形之裂殖壺菌屬培養上清液中總蛋白質(根據布萊佛(Bradford)檢定分析)及抗體 κ 鏈(透過ELISA檢定分析)之蓄積。

第19圖顯示裂殖壺菌屬EF1短啟動子多核苷酸序列 (SEQ ID NO:42)。

第20圖顯示裂殖壺菌屬EF1長啟動子多核苷酸序列 (SEQ ID NO:43)。

第21圖顯示裂殖壺菌屬60S短啟動子多核苷酸序列 (SEQ ID NO:44)。

第22圖顯示裂殖壺菌屬60S長啟動子多核苷酸序列 (SEQ ID NO:45)。

第23圖顯示裂殖壺菌屬Sec1啟動子多核苷酸序列 (SEQ ID NO:46)。

第24圖顯示pAB0011之質體拼圖。

第25圖顯示pAB0018之質體拼圖。

第26圖顯示pAB0022之質體拼圖。

第27圖顯示取自於以含有藉EF1啟動子(短版本)、EF1啟動子(長版本)、60S啟動子(短版本)、60S啟動子(長版本)、SEC1啟動子、及OrfC啟動子分別驅動之eGFP基因的表現載體轉形之裂殖壺菌屬培養之不含細胞上清液樣本中之eGFP之西方墨點。

第28圖顯示與藉OrfC啟動子(pCL0001-4)或EF1-L啟動子(AB0018-9及-10)驅動eGFP表現相關聯之轉形株細胞系之螢光顯微術。

第29圖顯示藉全甲基化N-聚糖之NSI-全MS分析測得之於天然裂殖壺菌屬所分泌之蛋白質上檢測得之N-聚糖結構。

第30圖顯示藉全甲基化N-聚糖之NSI-全離子映射測得之於天然裂殖壺菌屬所分泌之蛋白質上檢測得之N-聚糖結構。

第31圖顯示pSchiz-EPCT(+)-s1Suc2_CL0076之質體拼圖也定名為pCL0076。

第32圖顯示得自於蔗糖-SSFM上生長之以pCL0076轉形之裂殖壺菌種屬ATCC 20888培養之細胞丸粒之乾重(克/升)。轉形株稱作為1-1、1-3、1-24、3-1、3-2、3-5、3-21、4-1、4-24、及4-31。「對照組」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬ATCC 20888細胞。

第33圖顯示得自於蔗糖-SSFM上生長之以pCL0076轉形之裂殖壺菌種屬ATCC 20888培養之細胞丸粒之脂肪含量(以乾細菌量之重量百分比表示)。轉形株稱作為1-1、1-3、1-24、3-1、3-2、3-5、3-21、4-1、4-24、及4-31。「對照組」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬ATCC 20888細胞。

第34圖顯示得自於蔗糖-SSFM上生長之以pCL0076轉形之裂殖壺菌種屬ATCC 20888培養之細胞丸粒隨時間之經過測量得之乾重(克/升)。轉形株稱作為1-3及3-5。「對照組」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬

ATCC 20888細胞。

第35圖顯示得自於蔗糖-SSFM上生長之二轉形株培養之細胞丸粒之脂肪含量(以乾細菌量之重量百分比表示)。轉形株稱作為1-3及3-5。「對照組」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬 ATCC 20888細胞。

第36圖顯示得自以pCL0076轉形且於蔗糖-SSFM中生長2日或7日後收穫之裂殖壺菌屬株系B76-32培養之細胞丸粒之乾重(克/升)。「2118*」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬 ATCC 20888細胞之次單離株。

「B76-32**」係指於葡萄糖-SSFM上生長之B76-32親代株系。

第37圖顯示得自以pCL0076轉形且於蔗糖-SSFM中生長2日或7日後收穫之裂殖壺菌屬株系B76-32培養之細胞丸粒之脂肪含量。各樣本之最右欄顯示脂肪含量，以乾生質之重量百分比表示。各樣本之最左欄顯示由含18個或更少個碳(淺灰)或20個或更多個碳(中灰)之醯基所組成之總脂肪百分比。「2118*」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬 ATCC 20888細胞之次單離株。「B76-32**」係指於葡萄糖-SSFM上生長之B76-32親代株系。

第38A圖顯示反錄酶蛋白質之西方墨點及第38B圖顯示相對應之經庫馬西染色之SDS-PAGE凝膠。如西方墨點頂端指示，釀酒酵母(*S. cerevisiae*)反錄對照組及pCL0076轉形株1-3之3日培養之不含細胞的上清液分別係以5微克、2.5微克、1.25微克及0.625微克之數量載荷。

第39A圖顯示藉反應速率呈蔗糖濃度之函數示例顯示

之反錄酶活性檢定分析。第39B圖顯示用於計算 K_m 及 V_{max} 之標準蘭威佛-伯克(Lineweaver-Burk)作圖。

第40A圖顯示藉全甲基化N-聚糖之NSI總離子映射測得於經裂殖壺菌屬分泌之蛋白質上檢測得之N-聚糖結構式。

第40B圖顯示經由全甲基化N-聚糖之NSI-總離子映射所得之聚糖種類列表。

第41A圖及第41B圖顯示基於使用信號P演繹法則預測得裂殖壺菌屬之天然信號序列。例如參考Bendsten等人, J. Mol. Biol. 340: 783-795 (2004); Nielsen及Krogh, Proc. Int. Conf. Intell. Syst. Mol. Biol. 6:122-130 (1998); Nielsen等人, 蛋白質基因改造12:3-9 (1999); Emanuelsson等人, 自然方案2:953-971 (2007)。

第42圖顯示裂殖壺菌屬之密碼子使用表。

第43圖顯示pCL0120之質體拼圖。

第44圖顯示編碼得自裂殖壺菌屬Sec1信號肽融合至得自黑麴菌(*Aspergillus niger*)之成熟Suc1反錄酶(基因存庫存取號碼S33920)之密碼子最佳化核酸序列(SEQ ID NO:75)。

第45圖顯示pCL0137_EPCT(+)-s1Suc1之質體拼圖, 也定名為pCL0137。

【實施方式】

較佳實施例之詳細說明

網黏菌門之成員諸如裂殖壺菌屬、破囊壺菌屬、及其它破囊壺菌類為可透過習知分泌徑路處理多肽的真核生物。瞭解此等微生物比較CHO細胞也產生較少數大量分泌

的蛋白質，結果獲得使用裂殖壺菌屬優於CHO細胞之優點。此外，不似大腸桿菌，網黏菌門之成員諸如裂殖壺菌屬進行蛋白質糖化諸如N-鏈接的糖化作用，此乃某些蛋白質之生物活性所需。已經判定破囊壺菌諸如裂殖壺菌屬所具有之N-鏈接的糖化作用比較酵母糖化作用更加類似哺乳動物的糖化樣式。

重組蛋白質的有效製造也包括：(i)轉形經選定之宿主細胞之方法，(ii)用於選擇轉形株之選擇標記，及(iii)包含調節元體其可於特定宿主細胞發揮功能之表現卡匣。此種調節元體包括對於控制多核苷酸序列的表現上相當重要的啟動子及終結子序列。根據本發明，調節元體、調節控制元體及調節序列等術語可互換使用，及包括但非限於屬於啟動子、增強子、轉錄終結子、信號序列、核糖體結合位置、阻遏物結合位置、莖-環、及內含子剪接位置之序列及/或分子。若期望蛋白質分泌入培養基，則也可利用指導蛋白質分泌之信號肽(也稱作為信號序列、分泌信號肽、或先導子序列)。

宿主細胞

本發明係針對屬於網黏菌門微生物之宿主細胞內蛋白質的製造。於若干實施例中，該網黏菌門宿主細胞係用作為蛋白質製造的生物工廠。

於若干實施例中，該網黏菌門重組宿主細胞為破囊壺菌類諸如裂殖壺菌屬或破囊壺菌屬。根據本發明，「破囊壺菌類」係指破囊壺菌目(Thraustochytriales)中之任何成員，

其包括破囊壺菌科(Thraustochytriaceae)；而「網黏菌類」係指網黏菌目(Labyrinthulales)中之任何成員其包括網黏菌科(Labyrinthulaceae)。網黏菌科之成員先前被視為破囊壺菌目之成員，但更為晚近修正此等有機體之分類學類別，網黏菌科今日被視為網黏菌目之成員。網黏菌目及破囊壺菌目屬於網黏菌門之成員。分類學理論學家今日將此兩類微生物置於金黃藻菌鞭毛菌(Stramenopile)世系的藻菌或藻菌類原生生物。破囊壺菌類及網黏菌類之目前分類學位置摘述如下：

界：金黃藻菌鞭毛菌界(Stramenopila)(藻菌界(Chromista))

門：網黏菌門(Labyrinthulomycota)(鞭毛菌門(Heterokonta))

綱：網黏菌綱(Labyrinthulomycetes(Labyrinthulae))

目：網黏菌目(Labyrinthulales)

科：網黏菌科(Labyrinthulaceae)

目：破囊壺菌目(Thraustochytriales)

科：破囊壺菌科(Thraustochytriaceae)

用於本發明之目的，描述為破囊壺菌類之株系包括下列有機體：目：破囊壺菌目；科：破囊壺菌科；屬：破囊壺菌屬(Thraustochytrium)(種：種屬(sp.)、無基本(arudimentale)、金黃(aureum)、班西氏(benthicola)、球狀(globosum)、肯尼(kinnei)、運動(motivum)、多基本(multirudimentale)、巴奇氏(pachydermum)、增生(proliferum)、玫瑰(roseum)、紋狀(striatum))、吾肯氏壺菌

屬 (*Ulkenia*)(種：種屬、阿米巴 (*amoeboides*)、克氏 (*kerguelensis*)、微小 (*minuta*)、深海 (*profunda*)、星狀 (*radiata*)、塞氏 (*sailens*)、沙氏 (*sarkariana*)、裂殖 (*schizochytrids*)、威瑟氏 (*visurgensis*)、約肯氏 (*yorkensis*)、裂殖壺菌屬 (種：種屬、凝聚 (*aggregatum*)、利馬 (*limnaceum*)、蒙氏 (*mangrovei*)、微小 (*minuta*)、八孢節 (*octosporum*)、日本壺菌屬 (*Japonochytrium*)(種：種屬、海洋 (*marinum*)、亞普蘭壺菌屬 (*Aplanochytrium*)(種：種屬、哈里提迪 (*haliotidis*)、克古蘭喜 (*Kerguelensis*)、深海 (*profunda*)、史托奇諾 (*stocchinoi*)、亞松壺菌屬 (*Althornia*)(種：種屬、克魯奇 (*crouchii*)、或伊利那壺菌屬 (*Elina*)(種：種屬、馬里沙巴 (*marisalba*)、希諾里菲卡 (*sinorifica*)。用於本發明之目的，吾肯氏壺菌屬所述之種可視為屬於破囊壺菌屬之成員。奧蘭氏壺菌屬 (*Aurantiachytrium*)及橢孢節壺菌屬 (*Oblogospora*)屬於本發明中網黏菌門所涵蓋之額外二菌屬。

本發明中描述為網黏菌之株係包括下列有機體：目：網黏菌目，科：網黏菌科，屬：網黏菌屬 (種：種屬、藻 (*algeriensis*)、多核 (*coenocystis*)、查妥奈 (*chattonii*)、巨藻 (*macrocystis*)、大西洋巨藻 (*macrocystis atlantica*)、巨藻巨藻 (*macrocystis macrocystis*)、海洋 (*marina*)、微小 (*minuta*)、羅可芬喜 (*roscoffensis*)、沃卡諾威 (*valkanovii*)、蛋黃 (*vitellina*)、太平洋蛋黃 (*vitellina pacifica*)、蛋黃蛋黃 (*vitellina vitellina*)、左普菲 (*zopfii*)、擬網黏菌屬

(Labyrinthuloides)(種：種屬、哈里提迪(haliotidis)、約肯氏(yorkensis))、黏液網黏菌屬(Labyrinthomyxa)(種：種屬、海洋(marina))、二孢網黏菌屬(Diplophrys)(種：種屬、亞奇瑞(archeri))、紅網黏菌屬(pyrrhosorus)(種：種屬、海洋(marinus))、獨二孢網黏菌屬(Sorodiplophrys)(種：種屬、斯德科銳(sterocorea))或單胞網黏菌屬(Chlamydomyxa)(種：種屬、擬網黏菌(labyrinthuloides)、蒙大拿(montana))(但目前對於紅網黏菌屬、獨二孢網黏菌屬、或單胞網黏菌屬的確切分類學定位仍有爭議)。

網黏菌門之宿主細胞包括但非限於寄存株系PTA-10212、PTA10213、PTA-10214、PTA-10215、PTA-9695、PTA-9696、PTA-9697、PTA-9698、PTA-10208、PTA-10209、PTA-10210、PTA-10211、寄存為SAM2179之微生物(由寄存者命名為「吾肯氏壺菌屬SAM2179」)、任何破囊壺菌種屬(包括前者吾肯氏壺菌種屬諸如威瑟氏吾肯氏壺菌、阿米巴吾肯氏壺菌、沙氏吾肯氏壺菌、深海吾肯氏壺菌、星狀吾肯氏壺菌、微小吾肯氏壺菌及吾肯氏壺菌種屬BP-5601)，及包括紋狀破囊壺菌、金黃破囊壺菌、玫瑰破囊壺菌；及任何日本壺菌種屬。破囊壺菌目之株系包括但非限於破囊壺菌種屬(23B)(ATCC20891)；紋狀破囊壺菌(Schneider)(ATCC 24473)；金黃破囊壺菌(高斯坦(Goldstein))(ATCC 34304)；玫瑰破囊壺菌(高斯坦)(ATCC 28210)及日本破囊壺菌種屬(L1) (ATCC 28207)。裂殖壺菌屬包括但非限於凝聚裂殖壺菌、利馬裂殖壺菌、裂殖壺菌

種屬 (S31)(ATCC 20888)、裂殖壺菌種屬 (S8)(ATCC 20889)、裂殖壺菌種屬 (LC-RM)(ATCC 18915)、裂殖壺菌種屬 (S21)、已寄存株系 ATCC 28209 及已寄存利馬裂殖壺菌株系 IFO 32693。於若干實施例中，宿主細胞為裂殖壺菌屬或破囊壺菌。裂殖壺菌屬可藉連續二分法及經由形成孢子囊二者而複製，最終釋放出遊動孢子。但破囊壺菌屬只經由形成孢子囊複製，然後釋放出游動孢子。於若干實施例中，本發明之宿主細胞為重組宿主細胞。

本發明之宿主細胞之有效培養條件包括但非限於允許蛋白質製造及/或重組之有效培養基、生物反應器、溫度、pH及氧氣條件。有效培養基係指破囊壺菌目細胞例如裂殖壺菌屬宿主細胞典型培養於其中之任一種培養基。此等培養基典型包含具有可同化碳源、氮源、及磷酸來源，以及適當鹽類、礦物質、金屬、及其它營養素諸如維生素之水性培養基。適當培養基及培養條件之非限制性實例係揭示於實例章節。適合用於破囊壺菌目微生物之非限制性培養條件也說明於美國專利案第5,340,742號，全文以引用方式併入此處。本發明之細胞可於習知發酵生物反應器、振搖瓶、試管、微力價皿、及培養平板中培養。培養可於適合用於重組細胞之溫度、pH、及氧含量進行。於若干實施例中，本發明之網黏菌門宿主細胞含有包含編碼一選擇標記之核酸序列的重組載體。於若干實施例中，該選擇標記允許選擇已轉形之微生物。顯性選擇標記之實例包括可分解具抗生素活性或殺菌活性化合物之酶諸如得自印度鏈異

壁菌(*Streptomyces hindustanus*)之Sh ble基因，其編碼由SEQ ID NO:5表示之「薄萊黴素(bleomycin)-結合蛋白質」。於若干實施例中，編碼顯性選擇標記之核酸序列包含破囊壺菌乙醯乳酸酯合成序列諸如SEQ ID NO:6之多核苷酸序列之突變版本。於若干實施例中，該乙醯乳酸酯合成酶已經經過改性、突變、或以其它方式選擇來對於由磺胺嘧啶化合物、咪唑啉酮類抑制劑、及/或嘧啶基酮苯甲酸酯之抑制作用具有抗性。於若干實施例中，該乙醯乳酸酯合成酶為天然乙醯乳酸酯合成酶之同系物。於若干實施例中，已經以包含該乙醯乳酸酯合成酶之重組載體轉染之破囊壺菌微生物對磺胺嘧啶化合物、咪唑啉酮類抑制劑、及/或嘧啶基酮苯甲酸酯之敏感度減低。於若干實施例中，該重組載體包含編碼乙醯乳酸酯合成酶蛋白質之核酸序列，其包含之胺基酸序列與SEQ ID NO:7之差異為於下列位置中之一者或更多者：116G、117A、192P、200A、251K、358M、383D、592V、595W、或599F的一個胺基酸刪失、插入、或取代。於若干實施例中，已突變之乙醯乳酸酯合成酶蛋白質具有選自於下列所組成之組群之胺基酸序列：SEQ ID NO:8、SEQ ID NO:9、及SEQ ID NO:10。於若干實施例中，該重組載體包含與SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、或SEQ ID NO:10至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列，其中該多核苷酸序列編碼至少於破囊壺菌作為顯性選擇標記

之功能之胺基酸序列。於若干實施例中，該重組載體包含編碼SEQ ID NO:7之一功能片段之多核苷酸序列，其至少於破囊壺菌係作為顯性選擇標記。於若干實施例中，該重組載體包含選自於由下列所組成之組群：SEQ ID NO:11、SEQ ID NO:12、及SEQ ID NO:13中之多核苷酸序列。

根據本發明，「轉形」一詞用來指外生核酸分子(亦即重組核酸分子)可藉此插入微生物細胞之任一種方法。於微生物系統中，「轉形」一詞係用來描述該微生物因獲得外生性核酸所繼承的變化，且大致上與「轉染」一詞為同義詞。用於將外生核酸分子導入網黏菌門宿主細胞之適當轉形技術包括但非限於粒子撞擊、電穿孔、顯微注射、脂質體轉染、吸附、感染、及原生質體融合。於若干實施例中，包括重組載體之外生核酸分子導入於靜止期之微生物細胞。於若干實施例中，包括重組載體之外生核酸分子導入於對數生長期之微生物細胞。於若干實施例中，包括重組載體之外生核酸分子導入於600奈米達1.5至2之光密度的細胞。

本發明亦係針對一種轉形宿主細胞之方法包含：(a)以具有蛋白酶活性之酶前處理該宿主細胞，及(b)藉電穿孔將核酸分子導入該宿主細胞。於若干實施例中，該宿主細胞係於酶前處理後而於電穿孔前以比較未經酶前處理更高效率轉形。該酶包括但非限於與蝸牛丙酮(snail acetone)粉末、蛋白酶IX、蛋白酶XIV、硫酸酶、 β -葡萄糖醛酸酶、及其組合相關聯之酶活性。於若干實施例中，該宿主細胞係以約0.05毫克/毫升，約0.1毫克/毫升，約0.15毫克/毫升，約

0.2毫克/毫升，約0.25毫克/毫升，約0.3毫克/毫升，約0.4毫克/毫升，約0.5毫克/毫升，約0.6毫克/毫升，0.7毫克/毫升，0.8毫克/毫升，0.9毫克/毫升，或約1毫克/毫升蝸牛丙酮粉末、蛋白酶IX、蛋白酶XIV、或其組合前處理。於若干實施例中，該宿主細胞係以約0.05毫克/毫升至約1毫克/毫升、約0.1毫克/毫升至約1毫克/毫升、約0.1毫克/毫升至約0.5毫克/毫升、或約0.05毫克/毫升至約0.5毫克/毫升蝸牛丙酮粉末、蛋白酶IX、蛋白酶XIV、或其組合處理。於若干實施例中，該宿主細胞係以0.05X、0.1X、0.2X、0.3X、0.4X、0.5X、0.6X、0.7X、0.8X、0.9X、或1X硫酸酶、 β -葡萄糖醛酸酶或其組合處理。於若干實施例中，該宿主細胞係以約0.05X至約1X、約0.1X至約1X、約0.1X至約0.5X、或約0.05X至約0.5X硫酸酶、 β -葡萄糖醛酸酶或其組合處理。於若干實施例中，蛋白酶前處理包含使用於前述濃度中之任一者之蛋白酶IX、蛋白酶XIV、蝸牛丙酮粉末、硫酸酶、 β -葡萄糖醛酸酶或其組合前處理。於若干實施例中，電穿孔係於約100伏特至約500伏特電壓經歷0.1厘米或0.2厘米透光管間隙距離。於若干實施例中，電穿孔係於約100伏特、150伏特、200伏特、250伏特、300伏特、350伏特、400伏特、450伏特、或500伏特電壓經歷0.1厘米或0.2厘米透光管間隙距離。

於本發明之若干實施例中，宿主細胞經基因修改來導入或刪失與碳水化合物之轉運及/或合成相關聯之生物徑路的基因，包括涉及糖化作用的基因。舉例言之，宿主細

胞可藉刪失內生性糖化基因及/或插入人或動物糖化基因而修飾來允許糖化樣式更佳類似於人類的糖化樣式。於酵母之糖化作用修飾例如可參考美國專利案第7,029,872號及美國公告案第2004/0171826、2004/0230042、2006/0257399、2006/0029604、及2006/0040353號。本發明之宿主細胞也包括其中RNA病毒元體用來增加或調節基因表現之細胞。

表現系統

於若干實施例中，用於宿主細胞表現蛋白質之本發明之表現系統包含於藻菌細胞具有活性之調節控制元體。於若干實施例中，用於宿主細胞表現蛋白質之本發明之表現系統包含於網黏菌門細胞具有活性之調節控制元體。於若干實施例中，用於宿主細胞表現蛋白質之本發明之表現系統包含於破囊壺菌類具有活性之調節控制元體。於若干實施例中，用於宿主細胞表現蛋白質之本發明之表現系統包含於裂殖壺菌屬或破囊壺菌屬具有活性之調節控制元體。多種藻菌調節控制元體包括多個啟動子於多種種屬具有活性。因此，揭示作為本發明之面相之新穎調節序列可用於與其自其中單離之該細胞相同之細胞類型，或可用於與其單離之該細胞相異之細胞類型。此種表現卡匣之設計及組成係使用熟諳技藝人士已知之標準分子生物技術。例如參考Sambrook等人，2001年，分子選殖：實驗室手冊，第3版。

於若干實施例中，用於網黏菌門細胞製造蛋白質之表現系統包含衍生自網黏菌門序列之調節元體。於若干實施

例中，用於網黏菌門細胞製造蛋白質之表現系統包含衍生自非網黏菌門序列之調節元體，包括衍生自非網黏菌門藻菌序列之序列。於若干實施例中，本發明之表現系統包含編碼蛋白質之多核苷酸序列，其中該多核苷酸序列係與於網黏菌門宿主細胞有功能之任何啟動子序列、任何終結子序列、及/或任何其它調節序列相關聯。可使用可誘導的或組成性活性序列。適當調節控制元體也包括與此處所述之核苷酸分子相關聯之任一種調節控制元體。

本發明亦係針對於宿主細胞用於蛋白質表現之表現卡匣。本發明亦係針對包含於宿主細胞表現蛋白質之表現卡匣之前述任一種宿主細胞。於若干實施例中，該表現系統包含含下列基因元體諸如至少一啟動子、一編碼序列及一終結子區工作式地聯結，因此可於宿主細胞發揮功能之一表現卡匣。於若干實施例中，該表現卡匣包含如此處所述之本發明已單離之核酸分子中之至少一者。於若干實施例中，該表現卡匣之全部基因元體皆為與該已單離之核酸分子相關聯之序列。於若干實施例中，控制序列為可誘導序列。於若干實施例中，編碼該蛋白質之核酸序列係整合入該宿主細胞之基因體。於若干實施例中，編碼該蛋白質之核酸序列係穩定地整合入該宿主細胞之基因體。

於若干實施例中，編碼欲表現之蛋白質之已單離之核酸序列係工作式地聯結至一啟動子序列及/或一終結子序列，二者皆於該宿主細胞有功能。編碼欲表現之蛋白質之已單離的核酸序列所工作式聯結之啟動子序列及/或終結

子序列可包括任一種啟動子及/或終結子序列，包括但非限於本發明之新穎核酸序列、已核發之美國專利案第7,001,722號之調節序列、美國公告案第2006/0275904及2006/0286650號所揭示之調節序列、或其中被轉形而工作式聯結至編碼蛋白質之已單離之多核苷酸序列之於宿主內有功能之其它調節序列。於若干實施例中，編碼該蛋白質之核酸序列對該特定網黏菌門宿主細胞為密碼子最佳化來提高轉譯效率。

本發明亦係針對包括本發明之表現卡匣之重組載體。重組載體包括但非限於質體、噬菌體及病毒。於若干實施例中，該重組載體為線性化的載體。於若干實施例中，該重組載體為表現載體。如此處使用，「表現載體」一詞係指適合用於製造所編碼的產物(例如感興趣之蛋白質)之載體。於若干實施例中，編碼欲製造之產物之核酸序列係插入該重組載體來製造重組核酸分子。編碼該欲製造的蛋白質之核酸序列係以工作式聯結該核酸序列之載體內之調節序列(例如破囊壺菌目啟動子)之方式而插入該載體，允許該核酸序列於該重組微生物之轉錄及轉譯。於若干實施例中，可選擇標記包括此處所述任一種可選擇標記允許選擇其中已經成功地導入本發明之重組核酸分子之重組微生物。

於若干實施例中，藉本發明之重組宿主細胞製造之蛋白質包括但非限於治療性蛋白質。如此處使用，「治療性蛋白質」包括可用於動物體及人體之疾病、病況或病症之治療或預防之蛋白質。「處理」及「處理」等詞係指治療性處

理及預防性處理或預防措施，其中該物件可防止或減慢(減輕)非期望之生理病況、疾病或病症，或獲得有利的或期望的臨床結果。用於本發明之目的，有利的或期望的臨床結果包括但非限於與病況、疾病或病症相關聯之症狀或病徵的減輕；病況、疾病或病症程度的減少；病況、疾病或病症的穩定化(亦即該病況、疾病或病症並未惡化)；病況、疾病或病症之開始或進行延遲；病況、疾病或病症的改善；病況、疾病或病症的緩解(部分或全部緩解及包括可檢測或非可檢測)；或病況、疾病或病症之增進或改善。處理包括提引出臨床上顯著反應而無過多副作用。處理也包括比較若未經處理之預期存活率可延長存活。

於若干實施例中，治療性蛋白質包括但非限於生物活性蛋白質例如酶、抗體、或抗原性蛋白質。於若干實施例中，治療性蛋白質包括但非限於：蛋白質A、人生長激素、干擾素、抑肽酶、人 α 抗胰蛋白酶、親脂性蛋白質、人血清白蛋白、麩胺酸脫羧酶、胃脂解酶、乳鐵蛋白/溶菌酶、反錄酶、抗體(包括但非限於VEGF單株抗體(阿瓦斯汀(AVASTIN))及HER2單株抗體(賀癌平(HERCEPTIN))、人疫苗、動物疫苗、及動物治療劑。

於若干實施例中，藉本發明之重組宿主細胞製造之蛋白質包括但非限於工業酶。工業酶包括但非限於用於製造、製備、保藏、營養素移動、或產品包括食品、藥品、化學品、機械產品、及其它工業產品加工用之酶。工業酶包括但非限於： α -澱粉酶、 α -半乳糖苷酶、 β -澱粉酶、纖

維素、 β -葡聚糖酶、右旋葡聚糖酶、糊精酶、葡萄糖澱粉酶、半纖維素酶/戊聚糖酶、木聚糖酶、反錄酶、乳酸酶、柚苷酶、果膠酶、支鏈澱粉酶、酸性蛋白酶、鹼性蛋白酶、鳳梨蛋白酶、木瓜蛋白酶、胃蛋白酶、胺肽酶、內肽酶(胰蛋白酶、胰凝乳蛋白酶、胃蛋白酶、彈力蛋白酶)、凝乳酶(rennet/rennin/chymosin)、枯草桿菌蛋白酶(subtilisin)、嗜熱蛋白酶(thermolysin)、胺醯化酶、麩胺酶、溶菌酶、青黴素醯化酶(penicillin acylase)、三酸甘油酯酶類、磷脂酶類、前胃酯酶類、植酸酶、醯胺酶類、異構酶類、醇去氫酶、胺基酸氧化酶、催化酶、氯過氧化酶、過氧化酶、乙醯乳酸酯脫羧酶、天冬酸 β -脫羧酶、組胺酸酶、環糊精糖基轉移酶、根毛黴菌蛋白酶(fromase)、植酸酶、及凝乳酶。

於若干實施例中，藉本發明之重組宿主細胞製造之蛋白質包括營養缺陷型標記、涉及轉形選擇之顯性選擇標記(諸如分解抗生素活性之酶)或其它蛋白質、作為轉運子之蛋白質、涉及蛋白質糖化之酶、及涉及細胞代謝之酶。

於本發明實施例之任一個實施例中，由本發明之宿主細胞所製造之蛋白質可為「輸出蛋白質」或「異源性輸出蛋白質」。如此處使用「輸出蛋白質」或「異源性輸出蛋白質」係指未涉及修飾可製造該蛋白質之宿主細胞且係由該宿主細胞所製造用於隨後單離之異源性重組蛋白質。如此處定義「輸出蛋白質」不包括由通報子基因所編碼之蛋白質。

由本發明之重組宿主細胞所製造之異源性輸出蛋白質不包括可選擇標記諸如吉歐黴素(Zeocin)抗性基因(諸如得

自印度鏈異壁菌之ble基因及大腸桿菌新黴素(Neomycin)磷酸轉移酶(npt)、及轉座子Tn5、得自土麴菌(*Aspergillus terreus*)之殺稻瘟菌素(blasticidin)去胺酶(bsdR)、得自破囊壺菌屬T23B之PUFA合成酶ORFA、得自破囊壺菌屬T23B之PUFA合成酶ORFB、得自破囊壺菌屬T23B之PUFA合成酶ORFC、衍生自維多利亞水母(*Aequorea victoria*)之合成eGFP、編碼與選自於由廿二碳六烯酸(DHA)、廿二碳五烯酸(DPA)、廿碳五烯酸(EPA)及花生四烯酸(ARA)所組成之組群之脂肪酸合成相關聯之蛋白質的天然基因、脂肪酸合成酶、脂肪酸去飽和酶、脂肪酸延長酶、與聚縮酮合成酶複體相關聯之蛋白質及與脂肪酸摻混入磷脂質或摻混入三醯基甘油分子相關聯之蛋白質、 ω -3脂肪酸去飽和酶、多烯脂肪酸異構酶、HMG-CoA合成酶、HMG-CoA還原酶、鮫鯊烯合成酶、植烯合成酶、植烯去飽和酶、類胡蘿蔔素環化酶、類胡蘿蔔素羥化酶、類胡蘿蔔素酮醇酶、維生素E及硫辛酸、與類異戊間二烯生合成徑路相關聯之蛋白質，及涉及宿主細胞製造多元不飽和脂肪酸或類胡蘿蔔素之酶。

於若干實施例中，由本發明之宿主細胞所製造之蛋白質係以商業規模生產。商業規模包括自尺寸 ≥ 100 升、 $\geq 1,000$ 升、 $\geq 10,000$ 升、或 $\geq 100,000$ 升之通氣發酵槽生長之微生物製造蛋白質。於若干實施例中，商業規模製造係於伴以攪動之通氣發酵槽進行。

於若干實施例中，由本發明之宿主細胞所製造之蛋白質可蓄積於細胞內部或可呈可溶性蛋白質由細胞分泌入例

如培養基。

於若干實施例中，本發明所製造之蛋白質係自細胞、培養基或該細胞生長於其中之發酵介質回收。於若干實施例中，該蛋白質為自該培養基呈可溶性蛋白質回收之分泌蛋白質。於若干實施例中，該蛋白質為包含信號肽之分泌蛋白質。

於若干實施例中，本發明所製造之蛋白質包含指導其保留於內質網、指導其胞外分泌或指導其至其它小器官或細胞隔間之靶定信號。於若干實施例中，該蛋白質包含信號肽。於若干實施例中，該蛋白質包含Na/Pi-IIb2轉運子信號肽或Sec 1轉運蛋白質。於若干實施例中，該信號肽包含SEQ ID NO:1或SEQ ID NO:37之胺基酸序列。於若干實施例中，包含具有SEQ ID NO:1或SEQ ID NO:37之胺基酸序列之信號肽的蛋白質係分泌入培養基。於若干實施例中，該信號肽係於分泌過程自該蛋白質裂解，結果獲得蛋白質之成熟形式。

於若干實施例中，由本發明之宿主細胞所製造之蛋白質係經糖基化。於若干實施例中，由本發明所製造之蛋白質之糖基化樣式比較於酵母或大腸桿菌中所製造之蛋白質更密切類似哺乳動物糖基化樣式。於若干實施例中，由本發明之網黏菌門宿主細胞所製造之蛋白質包含N-鏈接糖化樣式。用於治療目的之糖化蛋白質當其糖化樣式類似於主角有機體內所見糖化樣式時較不可能促成抗糖型免疫反應。相反地，具有非屬主角有機體之特性的鍵聯或糖之糖

化蛋白質較可能具有抗原性。效應物功能也藉特定糖型式調整。舉例言之，IgG可媒介與Fc區糖型上分別為不存在有或存在有末端唾液酸相關的促炎反應或抗炎反應(Kaneko 等人，科學313(5787):670-3 (2006))。

本發明進一步係針對一種製造重組蛋白質之方法，該方法包含於足夠表現編碼該蛋白質之多核苷酸序列之條件下培養本發明之重組網黏菌門宿主細胞。於若干實施例中，該重組蛋白質係自該宿主細胞分泌及自該培養基中回收。於若干實施例中，自該細胞分泌之蛋白質包含分泌信號肽。依據載體及用於製造之宿主系統，本發明之重組蛋白質可維持於重組細胞內，可分泌入發酵介質內，可分泌入二細胞膜間之空間，或可保留於細胞膜外表面上。如此處使用，「回收蛋白質」一詞係指收集含該蛋白質之發酵介質而無需暗示額外分離或純化步驟。藉本發明方法製造之蛋白質可使用多種標準蛋白質純化技術純化，諸如但非限於親和層析術、離子交換層析術、過濾、電穿孔、斥水性交互作用層析術、凝膠過濾層析術、反相層析術、刀豆蛋白A層析術、色層聚焦、及差異溶解。於若干實施例中，藉本發明方法製造之蛋白質係以「實質上純質」形式單離。如此處使用，「實質上純質」係指允許蛋白質有效用作為商品之純度。於若干實施例中，該重組蛋白質係蓄積於細胞內部且係自細胞回收。於若干實施例中，該方法之宿主細胞為破囊壺菌。於若干實施例中，該方法之宿主細胞為裂殖壺菌屬或破囊壺菌屬。於若干實施例中，該重組蛋白質

為治療性蛋白質、食物酶、或工業酶。於若干實施例中，該重組網黏菌門宿主細胞為裂殖壺菌屬，及該重組蛋白質為包含分泌信號序列之治療性蛋白質。

於若干實施例中，本發明之重組載體為靶定載體。如此處使用「靶定載體」一詞係指用來將特定核酸分子遞送入重組細胞的載體，其中該核酸分子係用來刪失或鈍化該宿主細胞內部之內生基因(換言之，用於被靶定的基因破壞或擊倒技術)。此種載體也稱作為「擊倒」載體。於若干實施例中，部分靶定載體具有與宿主細胞標靶基因(亦即被靶定欲刪失或鈍化之基因)之核酸序列同源的核酸序列。於若干實施例中，插入該載體之核酸分子(亦即插子)係與該標靶基因同源。於若干實施例中，該載體插子之核酸序列係設計成結合至該標靶基因，使得該標靶基因與該插子進行同源重組，藉此該內生標靶基因被刪失、鈍化或衰減(亦即藉至少部分內生標靶基因被突變或刪失而衰減)。

已單離之核酸分子

本發明亦係針對可用於重組宿主細胞調節基因表現及/或指導蛋白質分泌之已單離之核酸分子或多核苷酸序列。此處所述核酸序列包括啟動子、終結序列、及編碼信號肽之核酸序列，及可用於調節同源蛋白質或異源蛋白質之轉錄及/或分泌。

根據本發明，已單離之核酸分子為已經自其天然周圍環境移出的(亦即，已經接受人為操縱的)核酸分子，其天然周圍環境為該核酸分子天然出現於其中之基因體或染色

體。如此，「已單離之」並非必要反映該核酸分子之純化程度，反而指示該分子不包括該核酸分子天然出現於其中的完整基因體或完整染色體。已單離之核酸分子可包括DNA、RNA（例如mRNA）、或DNA或RNA之衍生物（例如cDNA）。雖然「核酸分子」一詞主要係指物理核酸分子，及「核酸序列」或「多核苷酸序列」等詞主要係指核酸分子上的核苷酸序列，但特別就核酸分子、多核苷酸序列或可編碼蛋白質之核酸序列而言，該等術語係互換使用。於若干實施例中，本發明之經單離之核酸分子係使用重組DNA技術（例如聚合酶連鎖反應(PCR)擴增、選殖）或化學合成製造。已單離之核酸分子包括天然核酸分子及其同系物，包括但非限於天然同位變異株及改性核酸分子其中已經以此等改性可提供對所編碼之肽或蛋白質之序列、功能及/或生物活性具有期望的效果之方式而插入、刪失、取代、及/或反轉核苷酸。

本發明之啟動子序列、終結子序列、信號肽序列、或任何其它序列之核酸序列補體係指與本發明之啟動子序列、終結子序列、信號肽序列或任何其它序列之該股核酸序列互補的核酸股之核酸序列。須瞭解含有本發明之序列之雙股DNA包含單股DNA，及其互補股具有與該單股DNA互補的序列。如此，本發明之核酸分子可為雙股或單股，及包括於「苛刻」雜交條件下與本發明序列，及/或與本發明序列之補體形成穩定雜交體之該等核酸序列。推定互補序列之方法為熟諳技藝人士所已知。

「蛋白質」一詞包括單鏈多肽分子及多多肽複體，此處個別組成多肽係藉共價手段或非共價手段鏈接。「多肽」一詞包括典型具有超過5、10、或20個胺基酸長度之兩個或多個胺基酸之胜肽。

本發明之新穎核酸分子可用於其可發揮功能之任一種微生物。於若干實施例中，該等核酸分子係用於網黏菌門之重組微生物。於若干實施例中，該等重組核酸分子係用於破囊壺菌目之重組微生物。於若干實施例中，該等重組核酸分子係用於裂殖壺菌屬或破囊壺菌屬微生物。如此處使用，重組微生物具有使用重組技術而自其正常(亦即野生型或天然出現)形式改性(亦即突變或改變)的基因體。根據本發明之重組微生物可包括其中核酸分子已經經過插入、刪失、或改性(亦即例如經由核苷酸之插入、刪失、取代、及/或反轉而突變)使得此種改性或此等改性於微生物體內提供期望效果之一種微生物。如此處使用，可導致基因表現、基因功能或基因產物功能(亦即基因所編碼之蛋白質)減少之基因改性可稱作為基因的鈍化(完全或部分)、刪失、中斷、阻斷或向下調節。舉例言之，其中導致由該基因所編碼之蛋白質功能減低之基因中的遺傳改性可能為基因完全刪失(亦即該基因不存在於重組微生物，因此該蛋白質不存在於重組微生物)、基因突變而結果導致蛋白質未完全轉譯或未轉譯(例如蛋白質未表現)，或基因突變而減少或消除該蛋白質之天然功能(例如蛋白質之表現具有減少活性或無活性，例如酶活性或酶作用)的結果。導致基因表現或功

能增加之遺傳改性可稱作為基因之擴增、過度製造、過度表現、活化、增強、加成、或向上調節。

啟動子

本發明亦係針對屬於啟動子之新穎調節控制元體。本發明之啟動子為指導相關聯編碼區轉錄之DNA之一區。

於若干實施例中，該啟動子係得自網黏菌門微生物。於若干實施例中，該啟動子係得自破囊壺菌包括但非限於：寄存為SAM2179之微生物(由寄存者定名為「吾肯氏壺菌SAM2179」)、吾肯氏壺菌屬或破囊壺菌屬或裂殖壺菌屬微生物。裂殖壺菌屬包括但非限於凝聚裂殖壺菌、利馬裂殖壺菌、裂殖壺菌種屬(S31)(ATCC 20888)、裂殖壺菌種屬(S8)(ATCC 20889)、裂殖壺菌種屬(LC-RM)(ATCC 18915)、裂殖壺菌種屬(S21)、已寄存株系ATCC 28209及已寄存利馬裂殖壺菌株系IFO 32693。

本發明之啟動子可至少於破囊壺菌具有啟動子活性，及包括全長啟動子序列及其功能片段、融合序列、及天然啟動子同系物。限剪酶可用於消化本發明之核酸分子，接著適當檢定分析來測定啟動子活性所需之最小序列。此等片段本身個別表示本發明之實施例。啟動子同系物與天然啟動子之差異在於至少一、二、三或數個核苷酸已經經過刪失、插入、反轉、取代及/或衍生。啟動子同系物至少於破囊壺菌可保有作為啟動子之活性，但依據某些刺激而定，活性可增高、減低或促成。本發明之啟動子可包含提供發育方面及/或組織專一性調節控制或表現的一個或多

個序列元體。

於若干實施例中，本發明之已單離之核酸分子包含 PUFA PKS OrfC 啟動子(「PKS OrfC 啟動子」)。本發明之 PKS OrfC 啟動子為天然位在 OrfC 編碼區上游(朝向 5' 端)及指導 OrfC 轉錄之 DNA 一區。於若干實施例中，PKS OrfC 啟動子具有 SEQ ID NO:3 表示之多核苷酸序列。於若干實施例中，本發明之已單離之核酸分子包含與 SEQ ID NO:3 至少約 65%、至少約 70%、至少約 75%、至少約 80%、至少約 85%、至少約 90%、至少約 95%、至少約 96%、至少約 97%、至少約 98%、或至少約 99% 相同之多核苷酸序列，其中該多核苷酸序列具有啟動子活性(亦即至少對 PUFA PKS OrfC 序列具有基本啟動子轉錄活性)，至少於破囊壺菌係如此。於至少 10、至少 20、至少 30、至少 40、至少 50、至少 100、至少 200、至少 300、至少 400、至少 500、至少 1000、至少 1500 個核苷酸之序列，或完整序列發現同系度(或%相同度)。

本發明亦係針對包含雜交至 SEQ ID NO:3 或雜交至與 SEQ ID NO:3 至少 95% 相同之多核苷酸序列之多核苷酸序列之一單離之核酸分子。於若干實施例中，該已單離之核酸分子包含與 SEQ ID NO:3 完全互補之多核苷酸序列或與 SEQ ID NO:3 至少 95% 相同之多核苷酸序列。於若干實施例中，本發明之 PKS OrfC 啟動子包括一 PKS OrfC 啟動子同系物，其係充分類似於天然 PKS OrfC 啟動子序列，因此該同系物之核酸分子可於中、高或極高苛刻條件(容後詳述)下雜交至天然 PKS OrfC 啟動子之核酸序列之補體。於若干實施

例中，本發明之PUFA PKS OrfC啟動子序列於中、高或極高苛刻條件下雜交至SEQ ID NO:3之補體。

於若干實施例中，本發明之啟動子包含寄存於ATCC存取號碼PTA-9615之pCL0001之OrfC啟動子。

於若干實施例中，本發明之已單離之核酸分子包含EF1短啟動子(「EF1短」或「EF1-S」啟動子)或EF1長啟動子(「EF1長」或「EF1-L」啟動子)。本發明之EF1短或長啟動子為天然位在EF1編碼區上游(朝向5'端)及指導EF1轉錄之DNA的一區。於若干實施例中，該EF1短啟動子具有SEQ ID NO:42表示之多核苷酸序列。於若干實施例中，該EF1長啟動子具有SEQ ID NO:43表示之多核苷酸序列。於若干實施例中，本發明之已單離之核酸分子包含與SEQ ID NO:42或SEQ ID NO:43至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列，其中該多核苷酸序列至少於破囊壺菌具有啟動子活性(亦即至少對EF1短或長啟動子序列分別具有基本啟動子轉錄活性)。於至少10，至少20，至少30，至少40，至少50，至少100，至少200，至少300，至少400，或至少500個核苷酸之序列或整個序列具有同系度(或%相同度)。

本發明亦係針對一種已單離之核酸分子其包含雜交至SEQ ID NO:42及/或SEQ ID NO:43或雜交至與SEQ ID NO:42及/或SEQ ID NO:43至少95%相同之多核苷酸序列之一多核苷酸序列。於若干實施例中，該已單離之核酸分子

包含與SEQ ID NO:42或SEQ ID NO:43完全互補之多核苷酸序列或與SEQ ID NO:42或SEQ ID NO:43至少95%相同之多核苷酸序列。於若干實施例中，本發明之EF1短啟動子或EF1長啟動子包括分別充分類似天然EF1短及/或長啟動子序列之一EF1短或長啟動子同系物，使得於中高或極高苛刻條件(容後詳述)下，該同系物之核酸序列可雜交至該天然EF1短及/或長啟動子之核酸序列之補體。於若干實施例中，本發明之EF1短及/或長啟動子序列於中、高或極高苛刻條件下雜交至SEQ ID NO:42及/或SEQ ID NO:43分別之補體。

於若干實施例中，本發明之啟動子包含寄存於ATCC存取號碼PTA-9616之pAB0018之EF1長啟動子。

於若干實施例中，本發明之已單離之核酸分子包含60S短啟動子(「60S短」或「60S-S」啟動子)或60S長啟動子(「60S長」或「60S-L」啟動子)。本發明之60S短或長啟動子為天然位在60S編碼區下游(朝向5'端)及指導60S轉錄之DNA的一區。於若干實施例中，該60S短啟動子具有SEQ ID NO:44所表示之多核苷酸序列。於若干實施例中，該60S長啟動子具有SEQ ID NO:45所表示之多核苷酸序列。於若干實施例中，本發明之已單離之核酸分子包含多核苷酸序列其係與SEQ ID NO:44或SEQ ID NO:45至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同，其中該多核苷酸序列至少於破囊壺菌具有啟動

子活性(亦即至少對60S短或長啟動子序列分別具有基本啟動子轉錄活性)。同系度(或%相同度)出現於至少10、至少20、至少30、至少40、至少50、至少100、至少200、至少300、至少400、或至少500個核苷酸之序列或整個序列。

本發明亦係針對一種已單離之核酸分子包含雜交至SEQ ID NO:44及/或SEQ ID NO:45或雜交至與SEQ ID NO:44及/或SEQ ID NO:45至少95%相同之多核苷酸序列之一多核苷酸序列。於若干實施例中，該已單離之核酸分子包含與SEQ ID NO:44及/或SEQ ID NO:45完全互補或與SEQ ID NO:44及/或SEQ ID NO:45至少95%相同之多核苷酸序列互補之多核苷酸序列。於若干實施例中，本發明之60S短或60S長啟動子包括60S短或60S長啟動子同系物，其分別係充分類似天然60S短或60S長啟動子序列，使得該同系物之核酸分子可於中、高或極高苛刻條件(容後詳述)下，分別雜交至天然60S短及/或60S長啟動子之核酸序列的補體。於若干實施例中，本發明之60S短及/或60S長啟動子序列於中、高或極高苛刻條件下分別雜交至SEQ ID NO:44及/或SEQ ID NO:45之補體。

於若干實施例中，本發明之啟動子包含寄存於ATCC存取號碼PTA-9614之pAB0011之60S長啟動子。

於若干實施例中，本發明之已單離之核酸分子包含Sec1啟動子(「Sec1啟動子」)。本發明之Sec1啟動子為天然存在於Sec1編碼區上游(朝向5'端)及指導Sec1轉錄子DNA的一區。於若干實施例中，該Sec1啟動子具有SEQ ID NO:46

表示之多核苷酸序列。於若干實施例中，本發明之已單離之核酸分子包含與SEQ ID NO:46至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列，其中該多核苷酸序列具有至少於破囊壺菌之啟動子活性(亦即至少對SecI序列具有基本啟動子轉錄活性)。同系度(亦即%相同度)可出現於至少10、至少20、至少30、至少40、至少50、至少100、至少200、至少300、至少400、或至少500個核苷酸之序列或整個序列。

本發明亦係針對包含一多核苷酸序列之已單離之核酸分子，該序列係雜交至SEQ ID NO:46或雜交至與SEQ ID NO:46至少95%相同之多核苷酸序列。於若干實施例中，該已單離之核酸分子包含與SEQ ID NO:46完全互補之多核苷酸序列或與SEQ ID NO:46至少95%相同之多核苷酸序列完全互補之多核苷酸序列。於若干實施例中，本發明之SecI啟動子包括SecI啟動子同系物其係充分類似於天然SecI啟動子序列，使得該同系物之核酸序列可於中、高或極高苛刻條件(容後詳述)下雜交至天然SecI啟動子之核酸序列的補體。於若干實施例中，本發明之SecI啟動子於中、高或極高苛刻條件下雜交至SEQ ID NO:46之補體。

於若干實施例中，本發明之啟動子包含寄存於ATCC存取號碼PTA-9613之pAB0022之SecI啟動子。

終結子

本發明亦係針對屬於轉錄終結子之新穎調節控制元

體。本發明之終結子區為標記基因體DNA中之基因序列轉錄終點之遺傳序列的一區段。

於若干實施例中，該終結子區係得自網黏菌門微生物。於若干實施例中，該終結子區係得自破囊壺菌。於若干實施例中，該終結子區係得自裂殖壺菌屬或破囊壺菌屬。裂殖壺菌屬包括但非限於凝聚裂殖壺菌、利馬裂殖壺菌、裂殖壺菌種屬(S31)(ATCC 20888)、裂殖壺菌種屬(S8)(ATCC 20889)、裂殖壺菌種屬(LC-RM)(ATCC 18915)、裂殖壺菌種屬(S21)、已寄存株系ATCC 28209及已寄存利馬裂殖壺菌株系IFO 32693。

本發明之終結子區可至少於破囊壺菌具有終結子活性，包括全長終結子序列及其功能片段、融合序列、及天然終結子區之同系物。終結子同系物與天然終結子之差異在於已經刪失、插入、反轉、取代及/或衍生至少一個或數個但非限於一個或數個核苷酸。於若干實施例中，本發明之終結子同系物保有至少於破囊壺菌作為終結子區的活性，但依據某些刺激該活性可增高、減低或促成。

於若干實施例中，本發明包含一已單離之核酸分子其包含PUFA PKS OrfC基因之終結子區(「PKS OrfC終結子區」)。本發明之PKS OrfC終結子區為標記於基因體內DNA之OrfC基因序列轉錄終點之遺傳序列的一區段。於若干實施例中，該終結子區具有SEQ ID NO:4表示之多核苷酸序列。SEQ ID NO:4揭示之終結子區為得自破囊壺菌微生物之天然(野生型)終結子序列，特定言之為裂殖壺菌屬PKS

OrfC終結子區且定名為「OrfC終結子元體1」。於若干實施例中，本發明之已單離之核酸分子包含與SEQ ID NO:4至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同且至少於破囊壺菌發揮至少作為PUFA PKS OrfC終結子區之功能的多核苷酸序列。同系度(亦即%相同度)出現於至少10、至少20、至少30、至少40、至少50、至少100、至少150、或至少200個核苷酸序列或整個核苷酸序列。

本發明亦係針對包含一多核苷酸序列之已單離之核酸分子，該序列係雜交至SEQ ID NO:4或雜交至與SEQ ID NO:4至少95%相同之多核苷酸序列。於若干實施例中，該已單離之核酸分子包含與SEQ ID NO:4完全互補之多核苷酸序列或與SEQ ID NO:4至少95%相同之多核苷酸序列完全互補之多核苷酸序列。於若干實施例中，本發明之PKS OrfC終結子區包括PKS OrfC終結子區同系物其係充分類似於天然PUFA PKS OrfC終結子區序列，使得該同系物之核酸序列可於中、高或極高苛刻條件(容後詳述)下雜交至天然PKS OrfC終結子區之核酸序列的補體。於若干實施例中，本發明之PKS OrfC終結子區於中、高或極高苛刻條件下雜交至SEQ ID NO:4之補體。

於若干實施例中，本發明之終結子包含寄存於ATCC存取號碼PTA-9614之pAB0011之OrfC終結子區。

信號肽

本發明亦係針對編碼信號肽之新穎核酸分子。

於若干實施例中，本發明係針對包含編碼得自網黏菌門微生物之分泌蛋白質的信號肽之多核苷酸序列之一種已單離之核酸分子。於若干實施例中，該微生物為破囊壺菌。於若干實施例中，該微生物為裂殖壺菌屬或破囊壺菌屬。

本發明之信號肽可於破囊壺菌具有分泌信號活性，及包括全長肽及其功能片段、融合肽、及天然信號肽之同系物。信號肽同系物與天然信號肽之差異在於，至少一個或數個，但非限於一個或數個胺基酸已經刪失(例如蛋白質之截頭版本，諸如肽或片段)、插入、反轉、取代及/或衍生(例如經由糖化、磷酸化、乙醯化、肉豆蔻醯化、戊二烯化、棕櫚醯化、醯胺化、及/或添加糖基磷脂基肌糖醇而衍生)。於若干實施例中，信號肽同系物至少與破囊壺菌保有作為信號之活性，但該活性可依據某些刺激增減或促成。

於若干實施例中，該已單離之核酸分子包含編碼Na/Pi-IIb2轉運子蛋白質信號肽之多核苷酸序列。Na/Pi-IIb2轉運子蛋白質信號肽可具有至少於破囊壺菌至少對Na/Pi-IIb2轉運子蛋白質之信號靶定活性，及包括全長肽及其功能片段、融合肽、及天然Na/Pi-IIb2轉運子蛋白質信號肽之同系物。於若干實施例中，該Na/Pi-IIb2轉運子蛋白質信號肽具有SEQ ID NO:1表示之胺基酸序列。於若干實施例中，該Na/Pi-IIb2轉運子蛋白質信號肽具有SEQ ID NO:15表示之胺基酸序列。於若干實施例中，本發明之已單離之核酸分子包含編碼胺基酸序列之多核苷酸序列其係與SEQ ID

NO:1 或 SEQ ID NO:15 為至少約 60%、至少約 65%、至少約 70%、至少約 75%、至少約 80%、至少約 85%、至少約 90%、至少約 95%、至少約 96%、至少約 97%、至少約 98%、或至少約 99% 相同，其中該多核苷酸序列編碼至少於破囊壺菌至少對 Na/Pi-IIb2 轉運子蛋白質作為信號肽功能之胺基酸序列。於若干實施例中，該已單離之核酸分子包含一多核苷酸序列，該序列編碼已單離之胺基酸序列其包含至少於破囊壺菌至少對 Na/Pi-IIb2 轉運子蛋白質作為信號肽功能之 SEQ ID NO:1 或 SEQ ID NO:15 之功能片段。於若干實施例中，該已單離之核酸分子包含 SEQ ID NO:2。於若干實施例中，該已單離之核酸分子包含雜交至下列任一者之多核苷酸序列：(i) SEQ ID NO:2；(ii) 與 SEQ ID NO:2 至少約 60%、至少約 65%、至少約 70%、至少約 75%、至少約 80%、至少約 85%、至少約 90%、至少約 95%、至少約 96%、至少約 97%、至少約 98%、或至少約 99% 相同之多核苷酸序列；(iii) 與 SEQ ID NO:1 至少約 60%、至少約 65%、至少約 70%、至少約 75%、至少約 80%、至少約 85%、至少約 90%、至少約 95%、至少約 96%、至少約 97%、至少約 98%、或至少約 99% 相同之多核苷酸序列；及 (vi) 與 SEQ ID NO:15 至少約 60%、至少約 65%、至少約 70%、至少約 75%、至少約 80%、至少約 85%、至少約 90%、至少約 95%、至少約 96%、至少約 97%、至少約 98%、或至少約 99% 相同之多核苷酸序列。於若干實施例中，該已單離之核酸分子包含多核苷酸序列其係與下列任一者完全互補：(i) SEQ ID NO:2；(ii) 與 SEQ

ID NO:2至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；(iii)與SEQ ID NO:1至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(vi)與SEQ ID NO:15至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。

本發明亦係針對一種包含Na/Pi-IIb2轉運子信號胺基酸序列之已單離之多肽。於若干實施例中，該已單離之多肽包含選自於由下列所組成之組群之一胺基酸序列：(i) SEQ ID NO:1，(ii) SEQ ID NO:15，及(iii)與SEQ ID NO:1或SEQ ID NO:15至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之胺基酸序列，其係至少於破囊壺菌至少對Na/Pi-IIb2轉運子作為信號序列功能。於若干實施例中，該已單離之多肽包含SEQ ID NO:15之頭20、21、22、23、24、25、26、27、28、29、30、31、32、33、34或35個胺基酸殘基(亦即其中於此系列之C端最末1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、或18個胺基酸已經被刪失之SEQ ID NO:15)。位在SEQ ID NO:15之C端之18個胺基酸

預測為成熟Na/Pi轉運子蛋白質之一部分，及信號序列之裂解位置預測出現於SEQ ID NO:15之胺基酸殘基35與36間。但包括至多成熟Na/Pi轉運子蛋白質之全部18個胺基酸(亦即SEQ ID NO:15之最末18個胺基酸殘基)之信號肽為本發明可採用且所涵蓋。根據本發明，已單離之多肽為已經自其天然周圍環境移開(亦即已經接受人為操縱)之多肽及可包括例如已純化之蛋白質、已純化之胜肽、部分純化之蛋白質、部分純化之胜肽重組製造之蛋白質或胜肽，及天然製造之蛋白質或胜肽。如此，「已單離之」一詞並未反映該多肽之純化程度。於若干實施例中，本發明之已單離之Na/Pi-IIb2轉運子信號肽係重組製造。

於若干實施例中，該已單離之核酸分子包含編碼 α -1,6-甘露糖基轉移酶(ALG12)信號肽之多核苷酸序列。ALG12信號肽至少於破囊壺菌中至少對ALG12蛋白質可具有信號靶定活性，及包括天然ALG12信號肽之全長肽及其片段、融合肽及同系物。於若干實施例中，該ALG12信號肽具有SEQ ID NO:59表示之胺基酸序列。於若干實施例中，本發明之已單離之核酸分子包含編碼胺基酸序列之多核苷酸序列，該胺基酸序列係與SEQ ID NO:59至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同，其中該多核苷酸序列編碼胺基酸序列，其至少於破囊壺菌至少對ALG12蛋白質作為信號肽功能。於若干實施例中，該已單離之核酸分子包含多核

苷酸序列其係編碼包含SEQ ID NO:59之一功能片段的已單離之胺基酸序列，該功能片段係至少於破囊壺菌至少對ALG12蛋白質發揮作為信號肽之功能。於若干實施例中，該已單離之核酸分子包含SEQ ID NO:60。於若干實施例中，該已單離之核酸分子包含雜交至下列任一者之多核苷酸序列：(i) SEQ ID NO:60；(ii)與SEQ ID NO:60至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:59至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。於若干實施例中，該已單離之核酸分子包含與下列任一者完全互補之多核苷酸序列：(i) SEQ ID NO:60；(ii)與SEQ ID NO:60至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:59至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。

本發明亦係針對一種包含ALG12信號肽胺基酸序列之已單離之多肽。於若干實施例中，該已單離之多肽包含選

自於由下列所組成之組群中之胺基酸序列：(i) SEQ ID NO:59；及(ii)與SEQ ID NO:59至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之胺基酸序列其係至少於破囊壺菌至少對ALG12蛋白質發揮作為信號序列之功能。於若干實施例中，該已單離之多肽包含SEQ ID NO:59之首24、25、26、27、28、29、30、31、或32個胺基酸殘基。於若干實施例中，本發明之已單離之ALG12信號肽係重組製造。

於若干實施例中，該已單離之核酸分子包含結合免疫球蛋白蛋白質(BiP)信號肽之多核苷酸序列。BiP信號肽至少於破囊壺菌至少對BiP蛋白質可具有信號靶定活性，及包括天然BiP信號肽之全長肽及其片段、融合肽及同系物。於若干實施例中，該BiP信號肽具有SEQ ID NO:61表示之胺基酸序列。於若干實施例中，本發明之已單離之核酸分子包含編碼胺基酸序列之多核苷酸序列，該胺基酸序列係與SEQ ID NO:61至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同，其中該多核苷酸序列編碼胺基酸序列，其至少於破囊壺菌至少對BiP蛋白質作為信號肽功能。於若干實施例中，該已單離之核酸分子包含多核苷酸序列其係編碼包含SEQ ID NO:61之一功能片段的已單離之胺基酸序列，該功能片段係至少於破囊壺菌至少對BiP蛋白質發揮作為信號肽之功

能。於若干實施例中，該已單離之核酸分子包含SEQ ID NO:62。於若干實施例中，該已單離之核酸分子包含雜交至下列任一者之多核苷酸序列：(i) SEQ ID NO:62；(ii)與SEQ ID NO:62至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:61至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。於若干實施例中，該已單離之核酸分子包含與下列任一者完全互補之多核苷酸序列：(i) SEQ ID NO:62；(ii)與SEQ ID NO:62至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:61至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。

本發明亦係針對一種包含BiP信號肽胺基酸序列之已單離之多肽。於若干實施例中，該已單離之多肽包含選自於由下列所組成之組群中之胺基酸序列：(i) SEQ ID NO:61；及(ii)與SEQ ID NO:61至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約

90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之胺基酸序列其係至少於破囊壺菌至少對BiP蛋白質發揮作為信號序列之功能。於若干實施例中，該已單離之多肽包含SEQ ID NO:61之首23、24、25、26、27、28、29、30、或31個胺基酸殘基。於若干實施例中，本發明之已單離之BiP信號肽係重組製造。

於若干實施例中，該已單離之核酸分子包含編碼 α -1,3-葡萄糖苷酶(GLS2)信號肽之多核苷酸序列。GLS2信號肽至少於破囊壺菌至少對GLS2蛋白質可具有信號靶定活性，及包括天然GLS2信號肽之全長肽及其片段、融合肽及同系物。於若干實施例中，該GLS2信號肽具有SEQ ID NO:63表示之胺基酸序列。於若干實施例中，本發明之已單離之核酸分子包含編碼胺基酸序列之多核苷酸序列，該胺基酸序列係與SEQ ID NO:63至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同，其中該多核苷酸序列編碼胺基酸序列，其至少於破囊壺菌至少對GLS2蛋白質作為信號肽功能。於若干實施例中，該已單離之核酸分子包含多核苷酸序列其係編碼包含SEQ ID NO:63之一功能片段的已單離之胺基酸序列，該功能片段係至少於破囊壺菌至少對GLS2蛋白質發揮作為信號肽之功能。於若干實施例中，該已單離之核酸分子包含SEQ ID NO:64。於若干實施例中，該已單離之核酸分子包含雜交至下列任一者之多核苷酸序列：(i) SEQ ID

NO:64；(ii)與SEQ ID NO:64至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:63至少約64%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。

於若干實施例中，該已單離之核酸分子包含與下列任一者完全互補之多核苷酸序列：(i) SEQ ID NO:64；(ii)與SEQ ID NO:64至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:63至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。

本發明亦係針對一種包含GLS2信號肽胺基酸序列之已單離之多肽。於若干實施例中，該已單離之多肽包含選自於由下列所組成之組群中之胺基酸序列：(i) SEQ ID NO:63；及(ii)與SEQ ID NO:63至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之胺基酸序列其係至少於破囊壺菌至少對GLS2蛋白質發揮作為信號序列之功能。於若干實施例

中，該已單離之多肽包含SEQ ID NO:63之首30、31、32、33、34、35、36、37或38個胺基酸殘基。於若干實施例中，本發明之已單離之GLS2信號肽係重組製造。

於若干實施例中，該已單離之核酸分子包含編碼仿 α -1,3-1,6-甘露糖苷酶信號肽之多核苷酸序列。仿 α -1,3-1,6-甘露糖苷酶信號肽至少於破囊壺菌至少對仿 α -1,3-1,6-甘露糖苷酶蛋白質可具有信號靶定活性，及包括天然仿 α -1,3-1,6-甘露糖苷酶信號肽之全長肽及其片段、融合肽及同系物。於若干實施例中，該仿 α -1,3-1,6-甘露糖苷酶信號肽具有SEQ ID NO:65表示之胺基酸序列。於若干實施例中，本發明之已單離之核酸分子包含編碼胺基酸序列之多核苷酸序列，該胺基酸序列係與SEQ ID NO:65至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同，其中該多核苷酸序列編碼胺基酸序列，其至少於破囊壺菌至少對仿 α -1,3-1,6-甘露糖苷酶蛋白質作為信號肽功能。於若干實施例中，該已單離之核酸分子包含多核苷酸序列其係編碼包含SEQ ID NO:65之一功能片段的已單離之胺基酸序列，該功能片段係至少於破囊壺菌至少對仿 α -1,3-1,6-甘露糖苷酶蛋白質發揮作為信號肽之功能。於若干實施例中，該已單離之核酸分子包含SEQ ID NO:66。於若干實施例中，該已單離之核酸分子包含雜交至下列任一者之多核苷酸序列：

(i) SEQ ID NO:66；(ii)與SEQ ID NO:66至少約60%、至少約

65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:65至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。於若干實施例中，該已單離之核酸分子包含與下列任一者完全互補之多核苷酸序列：(i) SEQ ID NO:66；(ii)與SEQ ID NO:66至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:65至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。

本發明亦係針對一種包含仿 α -1,3-1,6-甘露糖苷酶信號肽胺基酸序列之已單離之多肽。於若干實施例中，該已單離之多肽包含選自於由下列所組成之組群中之胺基酸序列：(i) SEQ ID NO:65；及(ii)與SEQ ID NO:65至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之胺基酸序列其係至少於破囊壺菌至少對仿 α -1,3-1,6-甘露糖苷酶蛋白質發揮作為信號序列之功能。於若干實施例中，該已單離之多肽包含SEQ ID

NO:65之首24、25、26、27、28、29、30、31、32、33或34個胺基酸殘基。於若干實施例中，本發明之已單離之仿 α -1,3-1,6-甘露糖苷酶信號肽係重組製造。

於若干實施例中，該已單離之核酸分子包含編碼仿 α -1,3-1,6-甘露糖苷酶#1信號肽之多核苷酸序列。仿 α -1,3-1,6-甘露糖苷酶#1信號肽至少於破囊壺菌至少對仿 α -1,3-1,6-甘露糖苷酶#1蛋白質可具有信號靶定活性，及包括天然仿 α -1,3-1,6-甘露糖苷酶#1信號肽之全長肽及其片段、融合肽及同系物。於若干實施例中，該仿 α -1,3-1,6-甘露糖苷酶#1信號肽具有SEQ ID NO:67表示之胺基酸序列。於若干實施例中，本發明之已單離之核酸分子包含編碼胺基酸序列之多核苷酸序列，該胺基酸序列係與SEQ ID NO:67至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同，其中該多核苷酸序列編碼胺基酸序列，其至少於破囊壺菌至少對仿 α -1,3-1,6-甘露糖苷酶#1蛋白質作為信號肽功能。於若干實施例中，該已單離之核酸分子包含多核苷酸序列其係編碼包含SEQ ID NO:67之一功能片段的已單離之胺基酸序列，該功能片段係至少於破囊壺菌至少對仿 α -1,3-1,6-甘露糖苷酶#1蛋白質發揮作為信號肽之功能。於若干實施例中，該已單離之核酸分子包含SEQ ID NO:68。於若干實施例中，該已單離之核酸分子包含雜交至下列任一者之多核苷酸序列：(i) SEQ ID NO:68；(ii)與SEQ ID NO:68至少約

60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:67至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。於若干實施例中，該已單離之核酸分子包含與下列任一者完全互補之多核苷酸序列：(i) SEQ ID NO:68；(ii)與SEQ ID NO:68至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:67至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。

本發明亦係針對一種包含仿 α -1,3-1,6-甘露糖苷酶#1信號肽胺基酸序列之已單離之多肽。於若干實施例中，該已單離之多肽包含選自於由下列所組成之組群中之胺基酸序列：(i) SEQ ID NO:67；及(ii)與SEQ ID NO:67至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之胺基酸序列其係至少於破囊壺菌至少對仿 α -1,3-1,6-甘露糖苷酶#1蛋白質發揮作為信

號序列之功能。於若干實施例中，該已單離之多肽包含SEQ ID NO:67之首23、24、25、26、27、28、或29個胺基酸殘基。於若干實施例中，本發明之已單離之仿 α -1,3-1,6-甘露糖苷酶#1信號肽係重組製造。

於若干實施例中，該已單離之核酸分子包含編碼仿 α -1,2-甘露糖苷酶信號肽之多核苷酸序列。仿 α -1,2-甘露糖苷酶信號肽至少於破囊壺菌至少對仿 α -1,2-甘露糖苷酶蛋白質可具有信號靶定活性，及包括天然仿 α -1,2-甘露糖苷酶信號肽之全長肽及其片段、融合肽及同系物。於若干實施例中，該仿 α -1,2-甘露糖苷酶信號肽具有SEQ ID NO:69表示之胺基酸序列。於若干實施例中，本發明之已單離之核酸分子包含編碼胺基酸序列之多核苷酸序列，該胺基酸序列係與SEQ ID NO:69至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同，其中該多核苷酸序列編碼胺基酸序列，其至少於破囊壺菌至少對仿 α -1,2-甘露糖苷酶蛋白質作為信號肽功能。於若干實施例中，該已單離之核酸分子包含多核苷酸序列其係編碼包含SEQ ID NO:69之一功能片段的已單離之胺基酸序列，該功能片段係至少於破囊壺菌至少對仿 α -1,2-甘露糖苷酶蛋白質發揮作為信號肽之功能。於若干實施例中，該已單離之核酸分子包含SEQ ID NO:66。於若干實施例中，該已單離之核酸分子包含雜交至下列任一者之多核苷酸序列：(i) SEQ ID NO:66；(ii)與SEQ ID NO:66至

少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:69至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。於若干實施例中，該已單離之核酸分子包含與下列任一者完全互補之多核苷酸序列：(i) SEQ ID NO:66；(ii)與SEQ ID NO:66至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:69至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。

本發明亦係針對一種包含仿 α -1,2-甘露糖苷酶信號肽胺基酸序列之已單離之多肽。於若干實施例中，該已單離之多肽包含選自於由下列所組成之組群中之胺基酸序列：(i) SEQ ID NO:69；及(ii)與SEQ ID NO:69至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之胺基酸序列其係至少於破囊壺菌至少對仿 α -1,2-甘露糖苷酶蛋白質發揮作為信號序列之

功能。於若干實施例中，該已單離之多肽包含SEQ ID NO:69之首26、27、28、29、30、31、32、33、34、35或36個胺基酸殘基。於若干實施例中，本發明之已單離之仿 α -1,2-甘露糖苷酶信號肽係重組製造。

於若干實施例中，該已單離之核酸分子包含編碼仿 β -木糖苷酶信號肽之多核苷酸序列。仿 β -木糖苷酶信號肽至少於破囊壺菌至少對仿 β -木糖苷酶蛋白質可具有信號靶定活性，及包括天然仿 β -木糖苷酶信號肽之全長肽及其片段、融合肽及同系物。於若干實施例中，該仿 β -木糖苷酶信號肽具有SEQ ID NO:71表示之胺基酸序列。於若干實施例中，本發明之已單離之核酸分子包含編碼胺基酸序列之多核苷酸序列，該胺基酸序列係與SEQ ID NO:71至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同，其中該多核苷酸序列編碼胺基酸序列，其至少於破囊壺菌至少對仿 β -木糖苷酶蛋白質作為信號肽功能。於若干實施例中，該已單離之核酸分子包含多核苷酸序列其係編碼包含SEQ ID NO:71之一功能片段的已單離之胺基酸序列，該功能片段係至少於破囊壺菌至少對仿 β -木糖苷酶蛋白質發揮作為信號肽之功能。於若干實施例中，該已單離之核酸分子包含SEQ ID NO:72。於若干實施例中，該已單離之核酸分子包含雜交至下列任一者之多核苷酸序列：(i) SEQ ID NO:72；(ii)與SEQ ID NO:72至少約60%、至少約65%、至少約70%、至少約

75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:71至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。於若干實施例中，該已單離之核酸分子包含與下列任一者完全互補之多核苷酸序列：(i) SEQ ID NO:72；(ii)與SEQ ID NO:72至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:71至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。

本發明亦係針對一種包含仿 β -木糖苷酶信號肽胺基酸序列之已單離之多肽。於若干實施例中，該已單離之多肽包含選自於由下列所組成之組群中之胺基酸序列：(i) SEQ ID NO:71；及(ii)與SEQ ID NO:71至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之胺基酸序列其係至少於破囊壺菌至少對仿 β -木糖苷酶蛋白質發揮作為信號序列之功能。於若干實施例中，該已單離之多肽包含SEQ ID NO:71之首24、25、

26、27、28、29、或30個胺基酸殘基。於若干實施例中，本發明之已單離之仿 β -木糖苷酶信號肽係重組製造。

於若干實施例中，該已單離之核酸分子包含編碼胡蘿蔔素合成酶信號肽之多核苷酸序列。胡蘿蔔素合成酶信號肽至少於破囊壺菌至少對胡蘿蔔素合成酶蛋白質可具有信號靶定活性，及包括天然胡蘿蔔素合成酶信號肽之全長肽及其片段、融合肽及同系物。於若干實施例中，該胡蘿蔔素合成酶信號肽具有SEQ ID NO:73表示之胺基酸序列。於若干實施例中，本發明之已單離之核酸分子包含編碼胺基酸序列之多核苷酸序列，該胺基酸序列係與SEQ ID NO:73至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同，其中該多核苷酸序列編碼胺基酸序列，其至少於破囊壺菌至少對胡蘿蔔素合成酶蛋白質作為信號肽功能。於若干實施例中，該已單離之核酸分子包含多核苷酸序列其係編碼包含SEQ ID NO:73之一功能片段的已單離之胺基酸序列，該功能片段係至少於破囊壺菌至少對胡蘿蔔素合成酶蛋白質發揮作為信號肽之功能。於若干實施例中，該已單離之核酸分子包含SEQ ID NO:74。於若干實施例中，該已單離之核酸分子包含雜交至下列任一者之多核苷酸序列：(i) SEQ ID NO:74；(ii)與SEQ ID NO:74至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約

99%相同之多核苷酸序列；及(iii)與SEQ ID NO:73至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。於若干實施例中，該已單離之核酸分子包含與下列任一者完全互補之多核苷酸序列：(i) SEQ ID NO:74；(ii)與SEQ ID NO:74至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:73至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。

本發明亦係針對一種包含胡蘿蔔素合成酶信號肽胺基酸序列之已單離之多肽。於若干實施例中，該已單離之多肽包含選自於由下列所組成之組群中之胺基酸序列：(i) SEQ ID NO:73；及(ii)與SEQ ID NO:73至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之胺基酸序列其係至少於破囊壺菌至少對胡蘿蔔素合成酶蛋白質發揮作為信號序列之功能。於若干實施例中，該已單離之多肽包含SEQ ID NO:73之首15、16、17、18、19、20、21、29、30、31、32、33或34個胺基酸殘基。於若干實施例中，本發明之已單離之

胡蘿蔔素合成酶信號肽係重組製造。

於若干實施例中，該已單離之核酸分子包含編碼Sec1蛋白質(「Sec1」)信號肽之多核苷酸序列。Sec1信號肽至少於破囊壺菌至少對Sec1蛋白質可具有信號靶定活性，及包括天然Sec1信號肽之全長肽及其片段、融合肽及同系物。於若干實施例中，該Sec1信號肽具有SEQ ID NO:37表示之胺基酸序列。於若干實施例中，本發明之已單離之核酸分子包含編碼胺基酸序列之多核苷酸序列，該胺基酸序列係與SEQ ID NO:37至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同，其中該多核苷酸序列編碼胺基酸序列，其至少於破囊壺菌至少對Sec1蛋白質作為信號肽功能。於若干實施例中，該已單離之核酸分子包含多核苷酸序列其係編碼包含SEQ ID NO:37之一功能片段的已單離之胺基酸序列，該功能片段係至少於破囊壺菌至少對Sec1蛋白質發揮作為信號肽之功能。於若干實施例中，該已單離之核酸分子包含SEQ ID NO:38。於若干實施例中，該已單離之核酸分子包含雜交至下列任一者之多核苷酸序列：(i) SEQ ID NO:38；(ii) 與SEQ ID NO:38至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:37至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少

約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。於若干實施例中，該已單離之核酸分子包含與下列任一者完全互補之多核苷酸序列：(i) SEQ ID NO:38；(ii)與SEQ ID NO:38至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列；及(iii)與SEQ ID NO:37至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之多核苷酸序列。

本發明亦係針對一種包含Sec1信號肽胺基酸序列之已單離之多肽。於若干實施例中，該已單離之多肽包含選自於由下列所組成之組群中之胺基酸序列：(i) SEQ ID NO:37；及(ii)與SEQ ID NO:37至少約60%、至少約65%、至少約70%、至少約75%、至少約80%、至少約85%、至少約90%、至少約95%、至少約96%、至少約97%、至少約98%、或至少約99%相同之胺基酸序列其係至少於破囊壺菌至少對Sec1蛋白質發揮作為信號序列之功能。於若干實施例中，該已單離之多肽包含SEQ ID NO:37之首18或19個胺基酸殘基(亦即SEQ ID NO:37，其中於本系列C端之末1或2個胺基酸經刪失)。於若干實施例中，本發明之已單離之Sec1信號肽係重組製造。

於若干實施例中，本發明之已單離之核酸分子包含本

發明之OrfC啟動子、EF1短啟動子、EF1長啟動子、60S短啟動子、60S長啟動子、Sec1啟動子、Na/Pi-IIb轉運子蛋白質信號肽之編碼序列、或Sec1轉運蛋白質信號肽之編碼序列其係工作式鏈接至編碼蛋白質之核酸序列的5'端。本發明也涵蓋重組載體(包括但非限於表現載體)、表現卡匣及宿主細胞包含本發明之OrfC啟動子、EF1短啟動子、EF1長啟動子、60S短啟動子、60S長啟動子、Sec1啟動子、Na/Pi-IIb轉運子蛋白質信號肽之編碼序列、或Sec1轉運蛋白質信號肽之編碼序列其係工作式鏈接至編碼蛋白質之核酸序列的5'端。

包含前述本發明之已單離之核酸分子中之任一者(例如包含OrfC啟動子、EF1短啟動子、EF1長啟動子、60S短啟動子、60S長啟動子、Sec1啟動子、Na/Pi-IIb轉運子蛋白質信號肽之編碼序列、或Sec1轉運蛋白質信號肽之編碼序列之核酸分子)的重組載體(包括但非限於表現載體)、表現卡匣、宿主細胞及微生物係涵蓋於本發明之範圍；將該等載體及/或表現卡匣導入宿主細胞及重組微生物之方法也係涵蓋於本發明之範圍。適當載體及表現卡匣可經選擇或組成，因而含有適宜的調節序列、終結子片段、聚腺苷酸化序列、增強子序列、標記基因、及其它適當序列。有關載體、表現卡匣、及宿主細胞之額外細節係說明於此處。

除非另行載明，否則如此處使用，述及百分比(%)相同度(及%相同)係指使用下列所進行之同系度評估：(1)以標準內設值參數，使用blastp用於胺基酸搜尋及blastn用於核酸

搜尋進行BLAST 2.0基本BLAST同系度搜尋，其中查詢序列藉內設值過濾低複雜度區(例如參考Altschul S.等人，核酸研究25:3389-3402 (1997)，全文以引用方式併入此處)；(2)使用下述參數做BLAST 2校準；(3)及/或以標準內設值參數進行PSI-BLAST (位置專一性迭代BLAST)。須注意BLAST 2.0基本BLAST與BLAST 2之標準參數間有若干差異，兩個特定序列須使用BLAST 2程式辨識為具有顯著同系度，而使用該等序列中之一者作為查詢序列於BLAST 2.0基本BLAST進行搜尋可能無法識別於頂端匹配的第二序列。此外，PSI-BLAST提供「側寫資料」搜尋之自動化容易使用版本，此乃尋找序列同系物的敏感方式。該程式首先進行有間隙之BLAST資料庫搜尋。PSI-BLAST程式使用得自任何顯著校準資訊返回構成一位置專一性分數矩陣，其替代查詢序列用於資料庫的下一回合搜尋。因此，須瞭解百分比相同度可使用此等程式中之任一者測定。

例如如於Tatusova及Madden，FEMS Microbiol. Lett. 174:247-250 (1999)，全文以引用方式併入此處，所述使用BLAST 2序列的兩個特定序列可彼此校準。BLAST 2序列校準係使用BLAST 2.0演繹法則於blastp或blastn進行，來進行兩個序列間之有間隙BLAST搜尋(BLAST 2.0)，允許於所得校準導入間隙(刪失及插入)。於若干實施例中，BLAST 2序列校準係使用標準內設值參數進行如下。

對blastn使用0 BLOSUM62矩陣：

匹配加分 = 1

不匹配扣分 = -2

開放間隙(5)及延伸間隙(2)扣分 gap x_dropoff (50)預期(10)字大小(11)過濾(on)。

對blastp使用 0 BLOSUM62矩陣：

開放間隙(11)及延伸間隙(1)扣分 gap x_dropoff (50)預期(10)字大小(3)過濾(on)。

如此處使用，雜交條件係指核酸分子用來識別相似的核酸分子之標準雜交條件。例如參考Sambrook J.及Russell D. (2001)分子選殖：實驗室手冊第3版，冷泉港實驗室出版社，冷泉港，紐約，全文以引用方式併入此處。此外，揭示計算達成雜交允許核苷酸之各種不匹配程度的適當雜交及洗滌條件之公式，例如參考Meinkoth等人，Anal. Biochem. 138, 267-284 (1984)，全文以引用方式併入此處。

更特定言之，如此處所述中等苛刻雜交及洗滌條件係指允許與雜交反應中用來探測之核酸分子具有至少約70%核酸序列相同度之核酸分子單離的條件(亦即允許約30%或以下核苷酸不匹配之條件)。如此處所述高度苛刻雜交及洗滌條件係指允許與雜交反應中用來探測之核酸分子具有至少約80%核酸序列相同度之核酸分子單離的條件(亦即允許約20%或以下核苷酸不匹配之條件)。如此處所述極高度苛刻雜交及洗滌條件係指允許與雜交反應中用來探測之核酸分子具有至少約90%核酸序列相同度之核酸分子單離的條件(亦即允許約10%或以下核苷酸不匹配之條件)。如前文討論，熟諳技藝人士可使用於Meinkoth等人(舉例)之公式來計

算適當雜交及洗滌條件來達成此種特定程度的核苷酸不匹配。此等條件將依所形成者為DNA:RNA或DNA:DNA雜交體而改變。DNA:DNA雜交體計算所得熔點比DNA:RNA雜交體低 10°C 。於特定實施例中，DNA:DNA雜交體之苛刻雜交條件包括使用適當洗滌條件，於約 20°C 至約 35°C （較低苛刻）、約 28°C 至約 40°C （較高苛刻）、及約 35°C 至約 45°C （又更高苛刻）間之溫度，於6X SSC (0.9 M Na^{+})之離子強度雜交。於特定實施例中，DNA:RNA雜交體之苛刻雜交條件包括使用類似苛刻的洗滌條件，於約 30°C 至約 45°C 、約 38°C 至約 50°C 、及約 45°C 至約 55°C 間之溫度，於6X SSC (0.9 M Na^{+})之離子強度雜交。此等數值係基於對大於約100核苷酸，0%甲醯胺，及約40%G+C含量之分子的熔點計算。另外，可如Sambrook等人所述憑經驗算出 T_m 。一般而言，洗滌條件應儘可能地苛刻，且應適合所選用之雜交條件。舉例言之，雜交條件可包括比特定雜交體計算所得 T_m 低約 $20\text{-}25^{\circ}\text{C}$ 之鹽與溫度條件的組合，及洗滌條件典型地包括比該特定雜交體計算所得之 T_m 低約 $12\text{-}20^{\circ}\text{C}$ 之鹽與溫度條件的組合。適合用於DNA:DNA雜交體之雜交條件之一個實例包括於6 SSC (50%甲醯胺)於約 42°C 雜交2至24小時，接著為洗滌步驟包括於約2X SSC於室溫一次或多次洗滌，接著為於較高溫及較低離子強度額外洗滌(例如於約0.1X-0.5X SSC於約 37°C 至少一次洗滌，接著於約0.1X-0.5X SSC於約 68°C 至少一次洗滌)。

已經大致上說明本發明，經由參考此處提供之實例可

對本發明獲得進一步瞭解。此等實例僅供舉例說明之目的而非意圖限制性。

實例1

裂殖壺菌屬蛋白質表現載體pSchizE之構成

a. p07074#6之構成：

pSP73載體(波米迦(Promega), acc#X65333)以XbaI及綠豆核酸酶消化，及然後經純化及接合來形成載體p070604#3。然後p070604#3進一步以SphI、HpaI、及綠豆核酸酶消化，及然後經純化及再度接合而形成載體p070704#6。

b. pSchiz1之構成：

如WO02/083869之揭示，pTUBzeo11-2載體係以BamHI消化來釋放含裂殖壺菌屬 α -微管素啟動子、ble基因、及SV40終結子區之1122鹼基對(bp)片段。此片段經凝膠純化及接合入載體pYES2/CT(英維仇貞(Invitrogen))，該載體事先已經使用BamHI消化。然後所得構成體以SmaI、SphI、及綠豆核酸酶消化來釋放含 α -微管素啟動子之540 bp片段。該片段接合入事先已經以BamHI、SmaI、及綠豆接合酶消化之pUC19(基因存庫存取號碼L09137)來形成pSchiz1。

c. pSchiz2之構成

於分開反應中，PCR係用來使用下列引子自pTUBzeo11-2產生編碼SV40終結子之擴增子，該等引子結合NcoI及PciI限剪位置(以斜體字顯示)用於接合：

引子 5'-GATCCCATGGCACGTGCTACG (SEQ ID NO:16)

引子 5'-GGCAACATGTATGATAAGATAC (SEQ ID NO:17)

所得擴增子以 NcoI 及 PciI 消化而暴露出黏端。然後此 265 bp 片段接合入先前已經以 NcoI 消化之 pSchiz1。所得質體定名為 pSchiz2，含有該 α -微管素啟動子接著為 SV40 終結子。
d. pSchiz3 之構成：

於分開反應中，使用下列設計來將 SmaI 位置(以斜體字顯示)添加至任一端的下列引子，PCR 用來自 pYES2/CT 擴增多個選殖位置(MCS)：

引子 5'-GATCCCCGGGTTAAGCTTGGT (SEQ ID NO:18)

引子 5'-ACTGGGGCCCGTTTAACTC (SEQ ID NO:19)

然後所得 MCS 擴增子以 SmaI 消化及接合入先前已經以 NcoI 及綠豆核酸酶消化之 pSchiz2。所得載體定名為 pSchiz3，含有 α 微管素啟動子、SV40 終結子、及 pYES2/CT MCS。

e. pSchiz0.5#4 之構成

以下列引子及 pSchiz3 作為樣板，PCR 用來產生編碼 MCS 卡匣之一擴增子，該 MCS 卡匣係由 α -微管素啟動子 pYES2/CT MCS，及 SV40 終結子區所組成。

引子 5'-GACTAGATCTCAATTTTAGGCCCCCACTGACCG (SEQ ID NO:20)

引子 5'-GACTGTCGACCATGTATGATAAGATACATTGATG (SEQ ID NO:21)

此等引子係設計來將 BglII 及 SalI 限剪位置(以斜體字顯示)添加至 MCS 卡匣擴增子末端。所得 PCR 片段以 BglII 及 SalI 消化及接合入先前也已經以 BglII 及 SalI 消化之

p070704#6來產生pSchiz0.5#4。

f. pSchizE之構成

使用如美國專利案第7001,772號所述之載體pMON50203作為樣板，PCR用來產生編碼ALS基因(包括其啟動子區及終結子區)之4776 bp擴增子。含有NdeI及BglII限剪位置(以斜體字顯示)之下列引子用於本PCR反應：

引子5'-GACTCATATGGCCCAGGCCTACTTTCAC (SEQ ID NO:22)

引子5'-GACTAGATCTGGGTCAAGGCAGAAGAATTCCGCC (SEQ ID NO:23)

然後所得ALS擴增子以BglII及NdeI消化。pSchiz0.5#4同樣以BglII及NdeI消化，較大的3171 bp帶經凝膠純化及接合至已純化之ALS PCR擴增子。所得載體定名為pSchizE係藉定序驗證為含有ALS基因(包括啟動子區及終結子區)接著為一表現卡匣其含有裂殖壺菌屬 α -微管素啟動子、pYES2/CT MCS、及SV40終結子區(參考第5圖)。

實例2

裂殖壺菌屬蛋白質表現載體pSchiz-sG之構成

eGFP之表現及分泌係使用pSchiz-sG達成(參考第6圖)。本質體也定名為pCO0001，及2008年11月18日，寄存於美國種型培養物保藏所，專利寄存館，維吉尼亞州20110-2209曼納沙司，大學大道10801號，獲得ATCC存取號碼PTA-9617。本載體含有(i)裂殖壺菌 α -微管素啟動子序列接著為(ii)附接有成熟轉運子蛋白質N-端部之一片段(SEQ ID NO:15)之裂殖壺菌屬Na/Pi轉運子信號序列之一序列融合至eGFP編碼序列接著為(iii) MCS其餘部分及(iv)

SV40終結子區。本載體也含有裂殖壺菌屬ALS基因作為可選擇標記。該載體係如後文說明構成。

選用來編碼該信號肽之序列(SEQ ID NO:2)係得自由裂殖壺菌屬EST存庫中單離的Na/Pi轉運子。編碼信號肽及eGFP之核苷酸序列係藉PCR融合，使用含eGFP之質體pPha-T1-eGFP (Apt 等人，細胞科學期刊115:4061-4069(2002))，及三個引子設計來加上信號序列。第一PCR反應採用含eGfp之質體作為樣板，於eGfp 3'端之小引子(引子sec.Gfp3'Spe，含有SpeI位置，以斜體字顯示如下)，及旁出於eGfp 5'端及信號序列3'端之100 bp引子(引子sec.Gfp5'1b, sec)。該等引子序列如下：

引子 5'-TACTGGTTCCTTGTCGGCCTCGCCCTTCTCGGCGAT
GGCTTCAAGGTCATCGCCGGTGACTCCGCCGGTACGCTCTTCATGGT
GAGCAAGGGCGAGG (SEQ ID NO:24)
引子5'-CGTCACTAGTTTACTTGTACAGCTCGTCCATGCC (SEQ ID NO:25)

於第二PCR反應中，使用第一PCR之擴增子產物作為樣板使用相同3'引子連同結合信號序列其餘部分之第二100 bp 5'引子(sec.Gfp5'Bam)。第二5'引子序列含有BamHI位置(以斜體字顯示如下)且為下列：引子5'-TAATGGATCCATGGCCAACATCATGGCC
AACGTCACGCCCCAGGGCGTCGCCAAGGGCTTTGGCCTCTTTGT
CGGCGTGCTCTTCTTTCTCTACTGGTTCCTTGT (SEQ ID NO:26)

由本第二PCR反應所得PCR產物含有用於選殖之BamHI位置及SpeI位置。

第二PCR反應之擴增子及pSchizE皆係以BamHI及SpeI

消化及彼此接合來形成載體pSchiz-sG。

實例3

裂殖壺菌屬蛋白質表現載體pSchiz-sGr之構成

pSchiz-sGr載體含有 α -微管素啟動子、帶有編碼ER保留信號之一序列之eGFP核苷酸序列、SV40終結子區、及已突變之ALS可選擇標記。

常見ER保留信號胺基酸序列HDEL經反轉譯，編碼本保留信號之序列 (SEQ ID NO:14) 使用得自實例2之pSchiz-sG作為樣板，藉PCR而融合至eGFP。寡核苷酸引子設計為包括與中止密碼子(顯示為框起)同框架之HDEL編碼序列(於如下ss.eGfpHELD3'RV引子序列中下方劃線的反訊息補體)加上於一框架之BamHI位置(以斜體字顯示)及於另一個框架的EcoRV位置(以斜體字顯示)。

引子 ss.eGfpHELD3'RV: 5'-CCTGATATCTTACAACTCGTCGTGGTTGTACAGCTCGTCC (SEQ ID NO:27)

引子 sec.Gfp5'Bam2: 5'-TAATGGATCCATGGCCAACATCATGGCCAACGTCACGCCCCAGGGCGTCGCCAAGGGCTTTGGCCTCTTTGTCGGCGTGCTCTTCTTTCT CTACTGGTTCCTTGT (SEQ ID NO:28)

所得PCR產物以BamHI及EcoRV消化及接合至以BamHI及EcoRV消化pSchizE(如實例1所述)獲得之較大型片段。所得載體定名為pSchiz-sGr(參考第7圖)。

實例4

裂殖壺菌蛋白質表現載體pSchiz-cG之構成

作為比較性對照，pSchiz-cG載體構成可表現eGFP使

得螢光產物將蓄積於細胞的胞質。pSchiz-cG質體包含裂殖壺菌OrfC啟動子、編碼eGFP之多核苷酸序列、SV40終結子區、及已突變之裂殖壺菌屬ALS可選擇標記。

首先，裂殖壺菌屬ORFC上游的2000 bp以得自裂殖壺菌種屬ATCC 20888之基因體DNA的引子進行PCR擴增：

引子prREZ15: 5'-CGGTACCCGCGAATCAAGAAGGTAGGC (SEQ ID NO:29)

引子prREZ16: 5'-CGGATCCCGTCTCTGCCGCTTTTTCTT (SEQ ID NO:30)

prREZ15及prREZ16引子分別係含於KpnI及BamHI序列(斜體化)。所得擴增子係以BamHI及KpnI消化及藉凝膠純化。

其次，裂殖壺菌屬ORFC上游的1985 bp以得自裂殖壺菌種屬ATCC 20888之基因體DNA的引子進行PCR擴增：

引子prREZ17: 5'-CGGATCCGAAAGTGAACCTTGTCCTAACCC (SEQ ID NO:31)

引子prREZ18: 5'-CTCTAGACAGATCCGCACCATCGGCCG (SEQ ID NO:32)

prREZ17及prREZ18引子分別含有BamHI序列及XbaI序列(斜體化)。所得擴增子係以BamHI及XbaI消化及藉凝膠純化。

其次，載體pBluescript SK(+) (史徹特基因(Stratagene)，acc# X52328)以KpnI及XbaI消化及藉凝膠純化。本載體及如上產生的兩個擴增子全部同時接合而製造載體pREZ22。

然後載體pSchizE (實例1)以BamHI消化及以綠豆核酸酶處理，經管柱純化，以XbaI消化及然後經凝膠純化。然後以下列引子進行PCR來產生含eGFP編碼區之擴增子(使

用樣板pPha-T1-eGFP)：

引子 5'eGFP_kpn: 5'-GACTGGTACCATGGTGAAGCAAGGGCGAGGAG
(SEQ ID NO:33)

引子 3'eGFP_xba: 5'-GACTTCTAGATTACTTGTACAGCTCGTCCATGCC
(SEQ ID NO:34)

然後本擴增子以XbaI消化及接合至前述pSchizE片段來形成載體pSchizE-eGFP。

然後以pREZ22作為樣板使用PCR來產生編碼ORFC之啟動子之擴增子。各自含有KpnI限剪序列(以斜體字顯示)之下列引子用於本PCR：

引子5'ORFCproKpn-2: 5'- GATCGGTACCGGTGTTCTTTGTTTTGATTCT
(SEQ ID NO:35)

引子3'ORFCproKpn-2: 5'- GATCGGTACCGTCTCTGCCGCTTTTCTTTA
(SEQ ID NO:36)

然後本擴增子以KpnI消化。pSchizE-eGFP載體也以KpnI消化而產生二片段。較大的片段(7554 bp)經凝膠純化及接合至如上經KpnI消化之擴增子來產生pSchiz-cG載體，其含有該裂殖壺菌屬ORFC啟動子序列接著為eGFP序列及SV40終結子區(參考第8圖)。

實例5

裂殖壺菌屬之轉形及隨後之蛋白質表現

除非另行指示，否則全部載體及構成體皆係使用適合給定的培養規模之奎亞貞(Qiagen)套件組(加州汎倫西亞)，於用於質體純化之大腸桿菌UltraMax DH5-α FT化學

勝任細胞(英維仇貞，加州卡爾司貝)增殖。

裂殖壺菌種屬 ATCC 第 20888 號之培養係於 M2B 培養基生長，該培養基之組成為 10 克/升葡萄糖、0.8 克/升 $(\text{NH}_4)_2\text{SO}_4$ 、5 克/升 Na_2SO_4 、2 克/升 $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.5 克/升 KH_2PO_4 、0.5 克/升 KCl 、0.1 克/升 $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 、0.1 M MES (pH 6.0)、0.1% PB26 金屬、及 0.1% PB26 維生素(v/v)。PB26 維生素之組成為 50 毫克/毫升維生素 B12、100 微克/毫升噻胺、及 100 微克/毫升泛酸鈣。PB26 金屬經調整至 pH 4.5 及其組成為 3 克/升 $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、1 克/升 $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ 、800 毫克/毫升 $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 、20 毫克/毫升 $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ 、10 毫克/毫升 $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ 、600 毫克/毫升 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 、及 800 毫克/毫升 $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$ 。PB26 儲備溶液經分開過濾滅菌及於高壓蒸氣滅菌後添加至培養醪。於混合前，葡萄糖、 KH_2PO_4 、及 $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 各自與其餘培養醪成分分開高壓蒸氣滅菌以防止鹽沈澱及碳水化合物的焦糖化。全部培養基成分皆係購自希格瑪化學公司(Sigma Chemical)(密蘇里州聖路易)。裂殖壺菌屬培養生長至對數期及使用載體 pSchiz-E1 (實例 1)、pSchiz-sG (實例 2)、pSchiz-sGr (實例 3)、或 pSchiz-cG (實例 4)，使用拜歐利提(Biolistic)粒子撞擊器(拜雷公司(BioRad)，加州赫克利斯)轉形。拜歐利提轉形程序大致上與前文說明相同(參考 Apt 等人，J. Cell. Sci. 115(Pt 21):4061-9 (1996)及美國專利案第 7,001,772 號)。於 27°C 培養 2 日至 6 日後，一次轉形株於含 20 克/升瓊脂(VWR，賓州西契斯特)、10 微克/毫升蘇佛美突隆(Sulfometuron)甲基

(SMM)(化學服務公司(Chem Service)，賓州西契斯特)之固體M2B培養基上選擇。全部一次轉形株皆以人工轉移至含SMM之新鮮M2B平板上。

藉螢光顯微術及光學顯微術分析一次轉形株群落。一次轉形株群落也用來接種50微升M2B-SMM液體培養基。於27°C培養2日至5日後，經由於5500 x g離心15分鐘收穫培養物。不含細胞之上清液使用桑翠普(Centriprep)重力濃縮機(密里波(Millipore)，麻省比雷瑞卡)濃縮100倍，及細胞丸粒於水中洗滌及於液態氮中冷凍隨後再懸浮於丸粒重量兩倍之溶解緩衝液(組成為50 mM磷酸鈉(pH 7.4)、1 mM EDTA、5%甘油、及1 mM新鮮苯基甲基磺醯氟)及丸粒重量兩倍的0.5毫米玻璃珠(希格瑪，密蘇里州聖路易)。然後細胞丸粒混合物經由於多管渦旋器(VWR，賓州西契斯特)以最大速度於4°C渦旋3小時(h)而溶解。然後細胞溶解產物於於4°C於5500 x g離心10分鐘。保留上清液及於4°C於5500 x g再離心10分鐘。所得上清液於此處定義為「不含細胞之萃取物」。不含細胞之上清液及不含細胞之萃取物二者之蛋白質使用布萊佛檢定分析套件組(希格瑪，密蘇里州聖路易)定量，及於載荷至4-12%聚丙烯醯胺Bis-Tris凝膠(拜雷，加州赫克利斯)前，根據製造商之指示(拜雷，加州赫克利斯)與XT樣本緩衝液呈混合物沸騰。

SDS-PAGE凝膠係使用庫馬希染料染色或藉西方墨點移轉至PVDF。於打點之後，PVDF膜以經Tris緩衝之食鹽水(TBS)(希格瑪，密蘇里州聖路易)清洗及於室溫以5%脫脂乾

乳(NFDM)於TBS處理2小時。對感興趣之蛋白質具有專一性之一次抗體根據製造商指示於5% NFDM-TBS中稀釋。若有所需，此溶液經移出，以新鮮5% NFDM-TBS替代，該NFDM-TBS中添加綴合至鹼性磷酸酶且對第一抗體具有專一性之二次抗體。若未使用二次抗體，則一次抗體已經綴合至鹼性磷酸酶。然後經抗體處理之PVDF以TBS清洗及以5-溴-4-氯-3-吡啶基磷酸/硝基藍四唑鎢溶液(BCIP/NBT)(KPL，馬里蘭州蓋茲堡)處理。

如第9圖所示，以pSchizGr轉形之裂殖壺菌屬具有eGFP位在ER(也參考第10圖)，而以pSchizcG轉形之裂殖壺菌屬顯示eGFP遍布於胞質。以pSchizE轉形之裂殖壺菌屬(空白載體對照組)未顯示eGFP之表現。如第11圖所示，eGFP係於不含細胞之上清液樣本(亦即胞外)檢測，以及對以pSchiz-sG轉形之裂殖壺菌屬於不含細胞之萃取物中檢測。以pSchiz-sGr轉形之裂殖壺菌屬於不含細胞之萃取物中含有eGFP，及於不含細胞之上清液中含有eGFP至較少程度。最後，以pSchizcG轉形之裂殖壺菌屬排它地幾乎皆係含有eGFP於不含細胞之萃取物。

實例6

Sec1信號肽之識別

裂殖壺菌屬之基因體序列事先產生，組裝及格式化用於使用工業標準技術進行BLAST搜尋。得自於氮氣充滿條件下生長的裂殖壺菌屬培養上清液經濃縮及於SDS-PAGE上跑。由凝膠上切下主分泌蛋白質帶及用於胺基酸定序。

經由所得胺基酸序列與裂殖壺菌屬基因體之BLAST比較(演繹法則-tBLASTn，低複雜度過濾-關，預期-1000，矩陣-PAM30，無間隙校準-開)，識別相對應之ORF。ORF之5'部分使用SignalP演繹法則分析。例如參考Bendsten等人，J, Mol. Biol. 340:783-795 (2004); Nielsen及Krogh, Proc. Int. Conf. Intell. Syst. Mol. Biol. 6:122-130 (1998); Nielsen等人，蛋白質基因改造12:3-9 (1999); Emanuelsson等人，自然方案2:953-971 (2007)。Sec1蛋白質之5'區根據本分析識別為分泌信號。參考第12圖(包含SEQ ID NO:37)及第13圖(SEQ ID NO:38)。

實例7

裂殖壺菌屬載體pSchiz-Cpt之構成

pSchiz-Cpt載體含有OrfC啟動子及終結子及ALS可選擇標記。簡言之，經由首先以KpnI/XbaI消化pSchizE質體及凝膠純化所得6.89 kb片段而構成本載體。此消化也獲得自pSchizE質體中移除微管素啟動子及大部分多聯結子；ALS編碼序列維持完好。所得6.8 kb SchizE主鏈內部接合4 kb KpnI/XbaI片段，其含有裂殖壺菌屬OrfC上游之2 kb序列加裂殖壺菌屬OrfC下游之2 kb序列且有BamHI位置分開上游節段與下游節段。自質體pREZ22切下4 kb KpnI/XbaI片段(參考實例4)。6.8 kb pSchizE主鏈與4 kb KpnI/XbaI片段的接合獲得pSchiz-Cpt。

實例8

裂殖壺菌屬蛋白質表現載體pSchizCpt-s1eGFP之構成

pSchizCpt-sleGFP質體包含裂殖壺菌屬OrfC啟動子、eGFP編碼序列前方之Sec1信號序列、及OrfC終結子及已突變之裂殖壺菌屬ALS可選擇標記序列。參考第14圖。本質體也定名為pCL0001，及2008年11月18日，寄存於美國種型培養物保藏所，專利寄存館，維吉尼亞州20110-2209曼納沙司，大學大道10801號，獲得ATCC存取號碼PTA-9615。

實例9

於發酵槽藉裂殖壺菌屬表現及分泌eGFP

如實例5所述，裂殖壺菌屬細胞以pSchizCpt-sleGFP轉形(參考實例8)。SMM抗性細胞系於M2B瓊脂平板上單離及轉移入振搖瓶的M2B液體培養(含10微克/毫升SMM)內及於27.5°C使用於150 rpm之振搖培養。培養72小時至168小時後，藉離心(5000 x g 10分鐘)收穫培養物，不含細胞之上清液使用桑翠普及麥可康(Microcon)濃縮器(MWCO 10000)濃縮(約250倍)樣本(1微升至7微升)於SDS-PAGE上跑，分離的蛋白質移轉至PVDF膜上。經阻斷且經過洗滌之膜以免抗-eGFP IgG(拜維森(Biovision))探測，經由使用鹼性磷酸酶綴合山羊抗兔IgG(fc)(波米迦)探測及使用BCIP/NBT試劑處理讓反應性蛋白質帶變成目測可見。選出表現最高量eGFP之細胞系。

高製造細胞系中之一者於2.0升(工作體積)發酵槽培養。有擋板的接種物瓶含有150毫升HD1培養基，於29.5°C培養24小時至48小時伴以於200 rpm振搖。接種物培養用來接種發酵槽含有：50克/升葡萄糖、13.62克/升Na₂SO₄、0.72

克/升 K_2SO_4 、0.56克/升 KCl 、2.27克/升 $MgSO_4 \cdot 7H_2O$ 、1.8克/升 KH_2PO_4 、17.5克/升 $(NH_4)_2SO_4$ 、0.19克/升 $CaCl_2 \cdot 2H_2O$ 、51.5毫克 $FeSO_4 \cdot 7H_2O$ 、3.1克/升 $MnCl_2 \cdot 4H_2O$ 、6.2克/升 $ZnSO_4 \cdot 7H_2O$ 、0.04毫克 $CoCl_2 \cdot 6H_2O$ 、0.04毫克 Na_2MoO_4 、2.07克/升 $CuSO_4 \cdot 5H_2O$ 、2.07克/升 $NiSO_4 \cdot 6H_2O$ 、9.75毫克噻胺、0.16毫克維生素B12、及3.33毫克泛酸鈣。培養期間之溫度為 $29.5^\circ C$ ， $dO_2\%$ 控制於20%，葡萄糖濃度維持於15-20克/升(一旦初濃度落入於此範圍及pH維持於6.5)。於一定間隔以無菌方式移出樣本進行分析。

不含細胞上清液(含有0.5微克至5.0微克總蛋白質)之未經濃縮樣本藉SDS-PAGE分離，蛋白質移轉至PVDF膜。經阻斷且經洗滌之膜如前文說明探測分泌的eGFP。參考第15圖。

實例10

裂殖壺菌屬蛋白質表現載體pSchizCpt-s1kappabh之構成

pSchizCpt-s1kappabh質體包含裂殖壺菌屬OrfC啟動子、編碼IgG κ 亞單位之多核苷酸序列前方之Sec1信號序列，及OrfC終結子區及已突變之裂殖壺菌屬ALS可選擇標記序列(參考第16圖)。表現載體係如下製造：

編碼IgG κ 鏈之再合成的(密碼子-經最佳化的)基因係使用第42圖之密碼子使用表而自藍喜隆生技公司(Blue Heron Biotechnologies)出品的相對應胺基酸序列製造，5' Sec1分泌信號序列係選殖入orfC啟動子與orfC終結子間之質體pSchiz-Cpt。簡言之，載體pSchiz-Cpt係根據酶製造商指示(新英格蘭生物實驗室(New England Biolabs))以BamHI

及鹼性磷酸酶消化。此載體接合至以BglIII(NEB)消化之擴增子，及該擴增子係使用下列引子：

引子5'ss-X Bgl long: GACTagatctATGAAGTTCGCGACCTCG (SEQ ID NO:39)

引子 3'ritx_kap_bh_Bgl: gactagatctTCAGCACTCACCGCGGTAAAGG (SEQ ID NO: 40)

及藍喜隆生技公司提供的樣板形成，該樣板載有合成最佳化之DNA分子其含有Sec1信號肽之多核苷酸序列接著為κ多核苷酸序列(SEQ ID NO:41)。所得細菌性轉形株群落篩檢載體插子，適當校準用於表現。撿拾一個純株(定名為pSchizCPT-slkappabh)用於進一步分析，驗證序列及用於裂殖壺菌屬之轉形。

實例11

藉裂殖壺菌屬表現及分泌抗體亞單位κ

轉形株細胞系之生成

裂殖壺菌種屬ATCC 20888於含50毫升M2B培養基之250毫升振搖瓶內於27°C培養24小時。於600奈米測量吸光率，具有吸光率值為1體積相當於1毫升之培養於8,000 x g離心5分鐘。細胞丸粒懸浮於100微升M2B，以2厘米直徑圓展開於M2B瓊脂平板上。

平板含有展開的裂殖壺菌屬細胞之區以1100 Psi壓力使用經pSchizCptS 1kappa(以XbaI線性化)的M10珠粒撞擊。

撞擊後，M2B平板於27°C培養16小時，撿取長在裂殖壺菌屬展布平板中心的群落及展開於含10微克/毫升SMM之M2B平板。經24小時至72小時後，撿取20個群落及接種

於25毫升培養管內含10毫克/毫升SMM之5毫升M2B。培養管係於27°C於135 rpm培養72小時。

κ表現之檢測

選出如上所得十個最混濁的振搖瓶培養物接受表現分析。250毫升燒瓶內含10微克/毫升SMM之M2B 50毫升接種0.5毫升得自培養管之培養。培養物於27°C以於135 rpm之振搖培養24小時，隨後經由於5500 x g離心10分鐘分離細胞丸粒及不含細胞之上清液。

細胞丸粒懸浮於40毫升dH₂O，於5500 x g離心10分鐘然後懸浮於其濕重兩倍的萃取緩衝液。添加丸粒濕重兩倍的玻璃珠及培養管於4°C振搖3小時。所得細胞均化物於5000 x g離心10分鐘及保留上清液作為不含細胞之萃取物。

不含細胞之上清液使用桑翠普及麥可康濃縮機(亞米康(Amicon))具有10000分子量截留自約50毫升濃縮至大於200微升。

所選定之各轉形株之蛋白質(不含細胞之萃取物及已濃縮之不含細胞之上清液) 3微克於4-12% Bis Tris SDS PAGE(艾克里提昂(xt criterion)，拜雷公司)使用MOPS流動緩衝液於200伏特跑約45分鐘(直到染料鋒剛剛跑出凝膠底部之外)。使用紐沛吉(Nupage)轉移緩衝液於70伏特將蛋白質帶移轉至PVDF膜歷90分鐘。膜以5%(w/v)乳粉於含0.1%吐溫(Tween) 20之TBS中封阻，使用鹼性磷酸酶綴合抗人κ IgG(希格瑪公司)透過西方墨點法定位抗體亞單位κ。以BCIP/NBT磷酸鹽酶基質(KPL)讓陽性帶變成目測可見。

抗體亞單位 κ 可清晰檢測主要係得自如前述培養的振搖瓶培養物之不含細胞的上清液(參考第17圖)。 κ 蛋白質出現於不含細胞之上清液且具有與真正 κ 標準品無法區別之分子量係符合分泌 κ 亞單位及進行適當後轉譯改性(Sec1分泌信號的裂解)。

實例12

藉裂殖壺菌屬於發酵槽表現抗體亞單位 κ
培養

顯然可表現最高量胞外 κ 亞單位之兩株純株(1及3)於2.0升(工作體積)發酵槽內培養。有擋板之接種瓶含有150微升HD1培養基及於29.5°C培養24小時至48小時伴以於200 rpm之振搖。用於接種發酵槽之接種培養物含有：50克/升葡萄糖、13.62克/升 Na_2SO_4 、0.72克/升 K_2SO_4 、0.56克/升 KCl 、2.27克/升 $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、1.8克/升 KH_2PO_4 、17.5克/升 $(\text{NH}_4)_2\text{SO}_4$ 、0.19克/升 $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 、51.5毫克 $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、3.1克/升 $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ 、6.2克/升 $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.04毫克 $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ 、0.04毫克 Na_2MoO_4 、2.07克/升 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 、2.07克/升 $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$ 、9.75毫克噻胺、0.16毫克維生素B12、及3.33毫克泛酸鈣。培養期間之溫度為29.5°C， $\text{dO}_2\%$ 控制於20%，葡萄糖濃度維持於15-20克/升，一旦初濃度落入於此範圍，及pH維持於6.5。樣本係以間隔時間以無菌方式移出接受分析。

κ 表現之檢測

不含細胞之上清液未經濃縮進行分析，使用布萊佛方

法(希格瑪布萊佛試劑)以BSA作為標準品測定總蛋白質。驗證κ亞單位表現之西方分析係如前文對振搖瓶培養所述進行。

κ表現之定量係使用人κ-B+F Elisa定量套件組(比席爾實驗室公司(Bethyl Laboratories Inc.))及福洛歐美加(Fluoro Omega)平板讀取器(BMG萊柏泰克公司(BMG Labtech))進行，皆係根據製造商的指示進行。參考第18圖。

實例13

pAB0011表現載體之構成

使用萃取自裂殖壺菌種屬ATCC 20888之gDNA作為樣板，使用下列引子進行PCR來形成長2015 bp之擴增子：

5' 60S-807: TCGATTTGCGGATACTTGCTCACA (SEQ ID NO:47)

3' 60S-2821: GACGACCTCGCCCTTGGACAC (SEQ ID NO:48)

此擴增子經凝膠純化及用作為樣板用於隨後使用下列引子之PCR：

5' 60Sp-1302-Kpn: GACTggtaccTTTTCCGCTCTGCATAATCCTAA
(SEQ ID NO:49)

3' 60Sp-Bam: GACTggatccTTGGCTTTTTCTTTCTTGTTGC (SEQ ID NO:50)

所得1017 bp擴增子經純化，以KpnI及BamHI消化及接合至pSchiz-CPT(+)-slGFP(6h)，該載體先前經純化且以KpnI及BamHI消化。接合產物用於轉形大腸桿菌，質體經純化及藉得自所得群落之限剪消化物產物篩檢。具有預期之限剪消化物樣式及包括60S長啟動子之一個質體純株(#4.1)係藉珊哲(Sanger)定序驗證及命名為pAB0011用於裂殖壺菌屬之轉形。參考第24圖。該載體pAB0011係於2008

年11月18日，寄存於美國種型培養物保藏所，專利寄存館，維吉尼亞州20110-2209曼納沙司，大學大道10801號，獲得ATCC存取號碼PTA-9614。

實例14

pAB0018表現載體之構成

使用萃取自裂殖壺菌種屬 ATCC 20888之gDNA作為樣板，使用下列引子進行PCR來形成長2268 bp之擴增子：

5' EF1-68: CGCCGTTGACCGCCGCTTGACTCT (SEQ ID NO:51)

3' EF1-2312: CGGGGGTAGCCTCGGGGATGGACT (SEQ ID NO:52)

此擴增子經凝膠純化及用作為樣板用於隨後使用下列引子之PCR：

5' EF1-54-Kpn: GACTggtaccTCTTATCTGCCTCGCGCCGTTGAC (SEQ ID NO: 53)

3' EF1-1114-Bam:GACTggatccCTTGCTTGCTAGTAGTCGCTTTCGAAC (SEQ ID NO:54)

所得1060 bp擴增子經純化，以KpnI及BamHI消化及接合至pSchiz-CPT(+)-s1GFP(6h)，該載體先前經純化且以KpnI及BamHI消化。接合產物用於轉形大腸桿菌，質體經純化及藉得自所得群落之限剪消化物產物篩檢。具有預期之限剪消化物樣式及包括EF-1長啟動子之一個質體純株(#6.1)係藉珊哲定序驗證及命名為pAB0018用於裂殖壺菌屬之轉形。參考第25圖。該載體pAB0018係於2008年11月18日，寄存於美國種型培養物保藏所，專利寄存館，維吉尼亞州20110-2209曼納沙司，大學大道10801號，獲得ATCC

存取號碼PTA-9616。

實例15

pAB0022表現載體之構成

裂殖壺菌屬培養物之不含細胞的上清液藉SDS-PAGE分析，及由此選出單一蛋白帶標示為Sec1p(表示Sec1蛋白質)選用來純化至均質。此蛋白質之胜肽片段序列使用質譜技術識別，及使用BLAST演繹法則(tBLASTn，無間隙校準，低複雜度過濾OFF，預期=10000，矩陣=PAM30)(<ftp://ncbi.nlm.nih.gov/blast>)與裂殖壺菌屬全基因體序列之構想轉譯交互關聯。識別編碼全部胜肽序列之一個開放讀碼框及標示為Sec1g(表示Sec1基因)。於起始ATG上游的推定之啟動子序列也經識別與合成(藍喜隆生技公司)。含有合成Sec1g啟動子之載體使用下列引子用作為PCR之樣板：

5' Sec1P-kpn: GACTggtaccCCGTCCTTGACGCCTTCGC (SEQ ID NO:55)

3' Sec1P-bam: GACTggatccGATGAGTAATGGACGATCTTC (SEQ ID NO:56)

所得1438 bp擴增子經純化，以KpnI及BamHI消化及接合至pSchiz-CPT(+)-s1GFP(6h)，該載體先前經純化且以KpnI及BamHI消化。接合產物用於轉形大腸桿菌，質體經純化及藉得自所得群落之限剪消化物產物篩檢。具有預期之限剪消化物樣式及包括Sec1啟動子之一個質體純株(#8.1)係藉珊哲定序驗證及命名為pAB0022用於裂殖壺菌屬之轉形。參考第26圖。該載體pAB0022係於2008年11月18日，寄存於美國種型培養物保藏所，專利寄存館，維吉尼亞州20110-2209曼納沙司，大學大道10801號，獲得ATCC存取

號碼PTA-9613。

實例16

異源基因之轉錄

編碼延長因子1(EF1)及60S核糖體單元之基因之啟動子選用作為於裂殖壺菌屬異源基因轉錄的啟動子。搜尋裂殖壺菌屬之基因體序列，識別對二基因已公告之序列顯示同源性之基因。各啟動子之二版本(一短一長)透過PCR選殖。

SEC1基因之啟動子驅動表現也選用作為於裂殖壺菌屬中異源基因轉錄之啟動子。由於SEC1基因編碼至目前為止唯一識別於裂殖壺菌屬培養中之天然分泌的經糖化蛋白質，故此種啟動子可時間表現異源蛋白質至最適合用於製造分泌的已糖化蛋白質之生長期。

含有S1eGFP構成體(亦即於N端有SEC1分泌信號之eGFP基因，藉OrfC啟動子-載體CL0001驅動表現)之載體cpt(+)係經修改來切除OrfC啟動子，且以下列序列中之一者置換此元體：

EF1啟動子(短版本)=EF1-S得自載體AB0015

EF1啟動子(長版本)=EF1-L得自載體AB0018

60S啟動子(短版本)=60S-S得自載體AB0010

60S啟動子(長版本)=60S-L得自載體AB0011

SEC1啟動子=Sec1得自載體AB0022

裂殖壺菌種屬20888透過如前文說明之粒子撞擊各以5個載體轉形。隨機選定各個轉形之十個可存活細胞系用於分析。對5個載體各自進行以CL0001轉形來作為對照。隨機

選出以CL0001轉形之5個可存活細胞系用於分析。

轉形株細胞系於含50毫升M2B之250毫升振搖瓶內於29.5℃培養72小時，伴以於200 rpm連續振搖。經由於5000 x g離心10分鐘移出生質，不含細胞之上清液使用桑翠普離心機(MWCO 10000)濃縮至約1毫升。使用布萊佛方法測量不含細胞之上清液樣本之蛋白質濃度。

整份不含細胞之上清液於還原條件下於SDS丙烯醯胺凝膠(艾克里提昂)上分離，分離所得蛋白質帶移轉至PVDF膜上。eGFP使用與AP綴合之抗eGFP抗體檢測及使用NCIP/NBT試劑變成目測可見。

於各線道分離最大量已濃縮之上清液(7微升)來測定哪一個細胞系表現及分泌eGFP(表1)。雖然eGFP之表現程度於多個啟動子間各異，及對個別啟動子於多個細胞系間各異，但33細胞系(得自對所比較之6個啟動子分析55細胞系)顯示可表現eGFP(資料未顯示)。

表1.得自各次轉形可表現可檢測量之分泌的eGFP之細胞系數目

Promoter	OrfC (CL0001)	EF1-S AB0015	EF1-L AB0018	60S-S AB0010	60S-L AB0011	SEC1 AB0022
#表現 eGFP/# 總分析	2/5	6/10	5/10	6/10	6/10	8/10

雖然並非全部細胞系皆表現分泌的eGFP，但至少約半量細胞系確實產生可檢測濃度之eGFP，各個啟動子可指導eGFP的表現。

發現可表現與分泌eGFP之細胞系經進一步分析來比較相對啟動子長度。得自各培養物之不含細胞的上清液且測定可表現eGFP之蛋白質載荷至SDS丙烯醯胺凝膠上，且經

規度化成為每個線道1微克。蛋白質藉電穿孔分離，已分離之蛋白質移轉至PVDF膜及使用如前文說明之經AP綴合之抗eGFP抗體探測eGFP。參考第27圖。載荷於初步篩檢eGFP之上清液蛋白質數量比較用於產生第27圖之實驗中之用量約大十倍。如此，異源性eGFP蛋白質於第27圖之全部線道皆不顯著，原因在於於載荷進行實驗的蛋白質濃度，其對若干樣本係低於檢測濃度。於1微克蛋白質/線道之負載量，得自OrfC啟動子之分泌的eGFP之表現係低於可檢測極限。但於至少若干所產生之細胞系中由全部其它啟動子驅動的eGFP之表現為可檢測。如此驗證比較OrfC啟動子，所選出之啟動子亦即EF1、60S及SEC1皆為「強」啟動子。特定言之，於得自EF1-L轉形株之某些細胞系中，可見分泌的eGFP之表現係大於任何其它啟動子構成體之表現，指示此種啟動子為所測試的啟動子中最強而有力者。

經由螢光顯微鏡檢下觀察於EF1-L轉形株細胞系AB0018-9及AB0018-10中之表現且與典型OrfC轉形株之表現做比較，獲得EF1-L啟動子強度之證實。參考第28圖。而CL0001-4細胞系具有溫和螢光(Fluo:ISO 200 1.1秒平板)，AB0018-9及AB0018-10細胞系具有顯著螢光，指示胞內eGFP之實質上蓄積。

實例17

於裂殖壺菌屬之糖化側寫資料

測定天然裂殖壺菌屬蛋白質之N-糖化。發現裂殖壺菌屬與哺乳動物及昆蟲之糖化徑路有共用的步驟，且未觀察

得利用酵母之高甘露糖化徑路特性。第29圖及第30圖顯示藉裂殖壺菌屬分泌蛋白質之質譜術分析檢測得之聚糖結構。特定言之，觀察得特性峰，GlcNAc₂Man₅於m/z 1580、GlcNAc₂Man₆於m/z 1785、及GlcNAc₂Man₇於m/z 1991。

實例18

藉電穿孔轉形裂殖壺菌屬

裂殖壺菌種屬 ATCC 20888細胞於振搖器上於M50-20培養基(參考美國公告案第2008/0022422號)於200 rpm於29℃生長48小時。細胞於新鮮培養基稀釋1:100及生長隔夜。細胞經離心及再懸浮於1 M甘露糖醇及10 mM CaCl₂(pH 5.5)至2 OD₆₀₀單位之終濃度。5毫升細胞混合0.25毫克/毫升蛋白酶XIV(希格瑪化學公司)及於振搖器上培養4小時。細胞以10%冰冷甘油洗兩次及再懸浮於500微升冷10%甘油。90微升整份置於預先經過冷激之0.2厘米間隙電透光管(拜雷165-2086)。添加10微升DNA(1-5微克)至透光管，溫和混合及維持於冰上。細胞於200歐姆(電阻)、25微法拉第(μF)、及自0伏特至500伏特(用於0.1厘米透光管間隙距離)或500伏特(用於0.2厘米透光管間隙距離)範圍之電壓，細胞於重組載體電穿孔。即刻添加0.5毫升培養基至透光管。然後細胞移至4.5毫升M50-20培養基及於振搖器上於100 rpm培養2至3小時。細胞經離心及再懸浮於0.5毫升培養基，及以適當選擇(若有所需)接種至2片至5片M2B平板上及於29℃培養。

表2顯示於以不同酶組合物前處理(300伏特及0.1厘米透光管間隙距離參數)後所產生之裂殖壺菌種屬 ATCC

20888轉形株之數目。

表3顯示於以不同酶組合物及電壓前處理(0.1厘米透光管間隙距離)後所產生之裂殖壺菌種屬ATCC 20888轉形株之數目。

表4顯示使用不同電穿孔透光管間隙距離所產生之裂殖壺菌種屬ATCC 20888轉形株之數目。細胞係以0.25毫克/毫升蛋白酶XIV前處理。

表5顯示使用不同電穿孔電壓所產生之裂殖壺菌種屬ATCC 20888轉形株之數目。細胞係使用0.1毫克/毫升，蝸牛丙酮粉末(Snail Acetone Powder)+0.25毫克/毫升蛋白酶XIV(0.1厘米透光管間隙距離)前處理。

表2.以不同酶組合物前處理後所產生之裂殖壺菌屬轉形株

處理	#轉形株
無	0
0.1毫克/毫升，蝸牛丙酮粉末+0.25毫克/毫升蛋白酶XIV	450
蝸牛丙酮粉末，蛋白酶XIV添加前	225
0.1X硫酸酯酶+蛋白酶XIV	240
0.1X β -葡萄糖糖苷酶+蛋白酶XIV	430
0.5X硫酸酯酶+蛋白酶XIV	200
0.5X β -葡萄糖糖苷酶+蛋白酶XIV	375

表3.以不同酶組合物及電壓前處理後所產生之裂殖壺菌屬轉形株

處理	#轉形株於200伏特	#轉形株於250伏特
0.1毫克/毫升，蝸牛丙酮粉末+0.25毫克/毫升蛋白酶XIV	315	380
蛋白酶XIV	820	1360
0.5X β -葡萄糖糖苷酶+蛋白酶XIV	650	110
0.1X β -葡萄糖糖苷酶+蛋白酶XIV	400	560

表4.使用不同電穿孔透光管間隙距離(以0.25毫克/毫升蛋白
酶XIV前處理)所產生之裂殖壺菌屬轉形株

透光管間隙	#轉形株
0.1厘米(250伏特)	345
0.2厘米(500伏特)	530

表5.使用不同電穿孔電壓(以0.1毫克/毫升蝸牛丙酮粉末+0.25
毫克/毫升蛋白質XIV)前處理所產生之裂殖壺菌屬轉形株

電壓(伏特)	0	100	150	200	300	400	500
#轉形株	0	4	190	1320	794	156	100

實例19

反錄酶於裂殖壺菌屬之表現

載體pAB0018以BamHI及NdeI消化，獲得長838 bp及9879 bp之片段。9879 bp片段於瓊脂凝膠藉標準電穿孔技術分選，使用商用DNA純化套件組純化，及接合至一序列(SEQ ID NO:57)，該序列包含編碼裂殖壺菌屬Sec1蛋白質之天然分泌信號之多核苷酸序列，接著為編碼釀酒酵母之成熟反錄酶蛋白質(SUC2)之合成序列，該SUC2係經密碼子最佳化用於裂殖壺菌屬表現(參考第42圖)。含有Sec1信號肽編碼序列及釀酒酵母反錄酶蛋白質編碼序列之一融合序列插入裂殖壺菌屬載體pSchiz，接著以BamHI及NdeI消化來獲得SEQ ID NO:57。

接合產物用來使用製造商之方案轉形市面上供應的勝任DH5α大腸桿菌細胞株系(英維仇貞公司)。若干所得純株經過增殖，其質體經萃取及純化。然後此等質體藉限剪消化或PCR篩檢來驗證接合產生預期的質體載體。由與SEQ ID NO:57接合所得之一個此種質體載體係藉珊哲定序證實

及命名為pCL0076 (SEQ ID NO:58)。參考第31圖。

裂殖壺菌種屬 ATCC 20888 及遺傳改性裂殖壺菌屬衍生物標示為 B76-32 之培養物於 M2B 培養基生長，該培養基之組成為 10 克/升葡萄糖、0.8 克/升 $(\text{NH}_4)_2\text{SO}_4$ 、5 克/升 Na_2SO_4 、2 克/升 $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.5 克/升 KH_2PO_4 、0.5 克/升 KCl 、0.1 克/升 $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 、0.1 M MES (pH 6.0)、0.1% PB26 金屬、及 0.1% PB26 維生素 (v/v)。PB26 維生素之組成為 50 毫克/毫升維生素 B12、100 微克/毫升噻胺、及 100 微克/毫升泛酸鈣。PB26 金屬調整至 pH 4.5 及其組成為 3 克/升 $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、1 克/升 $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ 、800 毫克/毫升 $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 、20 毫克/毫升 $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ 、10 毫克/毫升 $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ 、600 毫克/毫升 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 、及 800 毫克/毫升 $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$ 。PB26 儲備溶液經分開過濾滅菌及於高壓蒸氣滅菌後添加至培養醪。葡萄糖、 KH_2PO_4 、及 $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 各自係與培養醪之其餘成分分開高壓蒸氣滅菌，隨後混合以防鹽沈澱及碳水化合物焦糖化。全部培養基成分皆係購自希格瑪化學公司(密蘇里州聖路易)。株系 B76-32 為根據美國專利案第 7,211,418 號及美國專利公告案第 2008/0022422 及 2008/0026434 號基因改造之裂殖壺菌種屬 ATCC 20888 之衍生物。

裂殖壺菌種屬 ATCC 20888 及 B76-32 之培養物生長至對數期，及使用電穿孔以後文說明之帶有酶前處理以載體 pCL0076 轉形。

帶有酶前處理之電穿孔-細胞於振搖器上於 50 毫升

M50-20培養基(參考美國公告案第2008/0022422號)於200 rpm於30°C生長2日。細胞以1:100稀釋入M2B培養基(參考下段)及生長隔夜(16小時至24小時)，意圖達到對數生長中期(OD_{600} 為1.5-2.5)。細胞於約3000 x g於50毫升錐形管離心5分鐘。移出上清液及細胞再懸浮於適當體積之1 M甘露糖醇、pH 5.5來達2 OD_{600} 單位之終濃度。5毫升細胞整份添加至25毫升振搖瓶內，使用10 mM $CaCl_2$ (1.0 M儲備溶液，過濾滅菌)及0.25毫克/毫升蛋白酶XIV(10毫克/毫升儲備溶液，過濾滅菌；希格瑪-亞利須公司(Sigma-Aldrich)，密蘇里州聖路易)改性。振搖瓶於振搖器上於30°C及約100 rpm培養4小時。細胞於顯微鏡下監視測定原生質體化程度，期望為單細胞。細胞於圓底試管(亦即14毫升法爾康(Falcon)試管，BD生科公司，加州聖荷西市)於2500 x g離心5分鐘。移出上清液，細胞以5毫升冰冷10%甘油溫和再懸浮。細胞於圓底試管於約2500 x g再離心5分鐘。移出上清液及細胞使用廣口徑滴量管梢端以500微升冰冷10%甘油溫和再懸浮。90微升細胞計量入經過預冷激之電透光管(基因帕塞(Gene Pulser)透光管-0.1厘米間隙或0.2厘米間隙，拜雷公司，加州赫克利斯)。1微克至5微克DNA(小於或等於10微升體積)添加至透光管，以滴量管梢端溫和混合及置於冰上5分鐘。細胞於200歐姆(電阻)、25微法拉第(電容)、及250伏特(用於0.1厘米間隙)或500伏特(用於0.2厘米間隙)電穿孔。0.5毫升M50-20培養基即刻添加至透光管。然後細胞移轉入25毫升振搖瓶之4.5毫升M50-20培養基內，及於振搖器

上於30°C及約100 rpm培養2小時至3小時。細胞於圓底試管內於2500 x g離心5分鐘。移出上清液及細胞丸粒再懸浮於0.5毫升M50-20培養基。細胞以適當選擇(若有所需)接種於適當數目的(2至5個) M2B平板上及於30°C培養。

選出轉形株用於M2B+SMM培養基生長，或經由接種於MSFM+蔗糖上直接選擇於蔗糖上生長。用於MSFM+蔗糖選擇，經1週至2週後，群落再度接種，伴以數次於含新鮮蔗糖的培養基上繼代培養。測定反錄酶之表現可用作為破囊壺菌群落於蔗糖作為唯一碳源上生長的可選擇標記。

用於下列實驗，於27°C培養2日至6日後，選出一次轉形株用來於含20克/升瓊脂(VWR，賓州西契斯特)、10微克/毫升SMM(化學服務公司，賓州西契斯特)之固體M2B培養基上生長。全部一次轉形株以手動移轉至含SMM之新鮮M2B平板。一週後，群落移至不含SMM之MSFM及5克/升蔗糖。1週後，最大的群落移轉至新鮮MSFM/蔗糖培養基平板。選出10個於蔗糖上生長之裂殖壺菌種屬ATCC 20888轉形株用於進一步決定特徵且分別標示為1-1、1-3、1-24、3-1、3-2、3-5、3-21、4-1、4-24、及4-31。選出於蔗糖上生長之B76-32轉形株中之九株接受進一步特徵化及標示為B76-32#2、#12、#19、326、#30、#39、#42、#56、及#61。

於蔗糖上生長之群落(1-1、1-3、1-24、3-1、3-2、3-5、3-21、4-1、4-24、4-31)使用接種環自平板上移出，移轉入含5毫升蔗糖培養基之培養試管內，及於振搖器上於29°C生長4日。2毫升此培養物用於接種250毫升燒瓶內之50毫升培

養基(MSFM或SSFM)及於29℃於振搖器上於200 rpm生長。

親代株系裂殖壺菌種屬 ATCC 20888之對照瓶係以相同方式但使用含蔗糖之培養基生長。7日後收穫細胞。細胞經離心及以50%異丙醇：蒸餾水混合物洗滌。已丸粒化之細胞經凍乾，稱重，及進行脂肪酸甲酯(FAME)分析。裂殖壺菌種屬 ATCC 20888或B76-32之CL0076轉形株之生長及脂肪含量係如先前於美國公告案第2008/0022422號(全文以引用方式併入此處)所述藉重量分析及已衍生油類之氣相層析術檢定分析。結果示於表6至表9。得自轉形株及親代株系之振搖瓶培養物之丸粒的乾重及脂肪含量顯示於第32圖至第37圖。

SSFM培養基：50克/升蔗糖或葡萄糖，13.6克/升 Na_2SO_4 、0.7克/升 K_2SO_4 、0.36克/升 KCl 、2.3克/升 $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.1M MES(pH 6.0)、1.2克/升 $(\text{NH}_4)_2\text{SO}_4$ 、0.13克/升麩胺酸一鈉、0.056克/升 KH_2PO_4 、及0.2克/升 $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 。維生素係以1毫升/升自儲備溶液添加，儲備溶液之組成為0.16克/升維生素B12、9.7克/升噻胺、及3.3克/升泛酸鈣。微量金屬係以2毫升/升自儲備溶液添加，儲備溶液之組成為1克/升檸檬酸、5.2克/升 $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、1.5克/升 $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ 、1.5克/毫升 $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.02克/毫升 $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ 、0.02克/毫升 $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ 、1.0克/毫升 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 、及1.0克/毫升 $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$ 調整至pH 2.5。

改性SFM(MSFM)培養基：10克/升葡萄糖或蔗糖，25.0克/升 NaCl 、1.0克/升 KCl 、0.2克/升 $(\text{NH}_4)_2\text{SO}_4$ 、5克/升、5.0

克 / 升 $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、 0.1 克 / 升 KH_2PO_4 、 0.3 克 / 升 $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 、 0.1M HEPES(pH 7.0)、0.1% PB26金屬、及0.1% PB26維生素(v/v)。維生素係以2毫升/升自儲備溶液添加，儲備溶液之組成為0.16克/升維生素B12、9.7克/升噻胺、及3.3克/升泛酸鈣。微量金屬係以2毫升/升自儲備溶液添加，儲備溶液之組成為1克/升檸檬酸、5.2克/升 $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、1.5克/升 $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ 、1.5克/毫升 $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.02克/毫升 $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ 、0.02克/毫升 $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ 、1.0克/毫升 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 、及1.0克/毫升 $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$ 調整至pH 2.5。

表6顯示於含葡萄糖、果糖、蔗糖或未添加碳源之MSFM中生長之裂殖壺菌種屬ATCC 20888之生長及脂肪含量。

表7顯示於含葡萄糖之MSFM培養基(對照)生長之裂殖壺菌種屬ATCC 20888及於含蔗糖之MSFM培養基生長之裂殖壺菌種屬ATCC 20888轉形細胞系之乾重及%脂肪酸。

表8顯示於含葡萄糖之SSFM培養基(對照)生長之裂殖壺菌種屬ATCC 20888及於含蔗糖之SSFM培養基生長之裂殖壺菌種屬ATCC 20888轉形細胞系之乾重及%脂肪酸。

表9顯示於含葡萄糖之SSFM培養基(對照)生長之裂殖壺菌屬B76-32及於含蔗糖之SSFM培養基生長之裂殖壺菌屬B76-32轉形細胞系之乾重及%脂肪酸。

表6.於含葡萄糖、果糖、蔗糖或未添加碳源之MSFM中生長之裂殖壺菌種屬ATCC 20888之生長及脂肪含量。]

	葡萄糖	果糖	蔗糖	未添加碳
DW(克/升)	2.84	2.65	0.16	0.11
% FA	66.5	65.3	ND	ND

DW=乾重
FA=脂肪酸

表 7. 於含蔗糖之 MSFM 培養基中生長之裂殖壺菌種屬
ATCC 20888 轉形細胞系

	20888 對照	1-1	1-3	3-2	3-5	3-21	4-1	4-24	4-31
DW (克/升)	2.94	2.49	2.79	2.21	2.60	2.64	2.44	3.05	2.24
% FA	70.87	70.79	72.36	67.97	69.78	71.05	68.84	73.85	73.66

DW=乾重
FA=脂肪酸

表 8. 於含蔗糖之 SSFM 培養基中生長之裂殖壺菌種屬 ATCC
20888 轉形細胞系

	20888 對照	1-1	1-3	1-24	3-1	3-2	3-5	3-21	4-1	4-24	4-31
DW (克/升)	11.24	10.04	10.51	9.99	8.40	10.29	9.03	8.34	8.16	10.63	10.92
% FA	78.22	78.20	76.29	77.10	77.37	77.71	74.97	73.44	73.65	80.05	79.82

DW=乾重
FA=脂肪酸

表 9. 於含蔗糖之 SSFM 培養基之 B76-32 轉形細胞系

	B76-32 對照	#2	#12	#19	#26	#30	#39	#42	#56	#61
7-day DW (克/升)	10.56	13.37	10.21	13.26	7.88	10.26	11.81	10.47	12.84	8.97
% FA	62.8	74.3	75.2	65.4	66.9	65.1	64.8	71.4	77.9	73.7

DW=乾重
FA=脂肪酸

免疫墨點-於 SSFM 生長 3 日之 50 毫升振搖瓶培養物之
不含細胞之上清液(參考美國公告案第 2008/0022422 號)係
於培養物於 5000 x g 離心後收集。培養上清液直接用於
SDS-PAGE，或使用裝配有可滲透膜允許比 10 kDa 更重的全
部組分濃縮之市售濃縮機濃縮 50 倍至 100 倍。總蛋白質濃度
藉布萊佛檢定分析(拜雷公司)測量。然後遵照標準免疫墨點
法程序(Sambrook 等人)藉免疫墨點分析證實反錄酶之表
現。簡言之，於 bis-tris 凝膠(英維仇貞，美國加州卡爾司貝)

上藉SDS-PAGE分離蛋白質(0.625微克至5微克)。然後以庫馬西藍(單純藍安全染色, 英維仇貞公司)染色蛋白質, 或蛋白質移轉至聚亞乙烯氟膜上, 及使用衍生自己經注射釀酒酵母反錄酶之純質製劑(希格瑪公司)之兔衍生所得之反錄酶抗血清(開放生物系統公司(Open Biosystems))來探測是否存在有反錄酶蛋白質。隨後膜與偶合至鹼性磷酸酶之小鼠抗兔二次抗體(波米迦公司)培養。然後根據製造商之指示(KPL, 馬里蘭州蓋塞堡), 膜以5-溴-4-氯-3-吡啶基-磷酸酯/硝基藍四唑鎢溶液(BCIP/NBT)處理。一個實例示於第38圖。抗反錄酶免疫墨點法及相對應之經庫馬西藍染色之凝膠分別示於A圖及B圖。於純株1-3之培養上清液中所見四個主帶中, 只有一者顯示與抗反錄酶抗血清反應。藉胜肽序列分析證實蛋白質身分。

功能檢定分析-酶EC 3.2.1.26為可催化蔗糖水解成為果糖及葡萄糖之蔗糖酶之反錄酶類型。蔗糖酶活性係藉果糖及葡萄糖自蔗糖釋放之速率測量。經由添加蔗糖至發酵醪上清液粗略進行檢定分析, 葡萄糖/果糖濃度係藉HPLC測量。

裂殖壺菌株系B76-32 #3係於MSFM(含蔗糖)於50毫升振搖瓶內於29°C生長至OD達到約4。細胞於4500 x g離心15分鐘, 及測量於上清液之反錄酶活性。經由添加0.1 M蔗糖至不等體積之發酵醪及將終體積調整至1毫升, 進行反錄酶之檢定分析。反應係於55°C培養3分鐘。於100°C反應10分鐘結束反應, 然後冷凍至分析為止, 分析包含藉HPLC測定葡萄糖、果糖及蔗糖。使用已修改之方法版本如Liu等人食

品科學28:293-296(2007)進行HPLC。簡言之，使用有魯納(Luna)NH₂管柱之HPLC分離單醣及雙醣及使用RID(折射率檢測器)檢測。經由比較其滯留時間與標準品之滯留時間進行識別。藉外部標準校準進行定量。反應速率呈蔗糖濃度之函數顯示於第39A圖。Km(33.4 mM)及Vmax(6.8 mM葡萄糖/分鐘)係自標準Lineweaver-Burk作圖求出。參考第39B圖。

糖化分析-上清液蛋白質係於4%至12% bis-tris凝膠(英維仇貞)藉SDS-PAGE分離。然後蛋白質以庫馬西藍(單純藍安全染色，英維仇貞)染色。自凝膠切下感興趣之已染色的蛋白質及將切片切成小塊(約1立方毫米)及以40 mM碳酸氫銨(AmBic)及100%乙腈交錯脫染色直到顏色轉透明無色。已脫染色之凝膠於55°C於10 mM DTT於40 mM碳酸氫銨再度溶脹1小時。DTT溶液交換55 mM碘乙醯胺(IAM)及於暗處培養45分鐘。培養接著以40 mM碳酸氫銨及100%乙腈交錯洗滌兩次。已脫水之凝膠最初於冰上使用胰蛋白酶溶液(胰蛋白酶於40 mM碳酸氫銨)再度溶脹45分鐘，及於37°C進行蛋白質消化隔夜。上清液移至另一根試管。自凝膠一系列使用下列萃取劑來萃取胜肽及糖肽：20%乙腈於5%甲酸，50%乙腈於5%甲酸及然後80%乙腈於5%甲酸。樣本溶液經乾燥及組合入一根試管。已萃取之胰蛋白酶消化產物通過C18賽普-派克(sep-pak)卡匣且以5%乙酸洗滌去除污染物(諸如鹽類及SDS)。胜肽及糖肽以20%異丙醇於5%乙酸、40%異丙醇於5%乙酸、及100%異丙醇一系列洗提及於加速真空濃縮機內乾燥。已乾燥之樣本經組合然後以50 mM磷

酸鈉緩衝液(pH 7.5)重新調製及於100°C加熱5分鐘來鈍化胰蛋白酶。胰蛋白酶消化產物與PNGase F於37°C培養隔夜來釋放出N-聚糖。消化後，樣本通過C18賽普-派克卡匣，碳水化合物選分以5%乙酸洗提及藉凍乾乾燥。釋放出之與N鏈接的寡糖基於Anumula及Taylor，Anal Biochem. 203:101-108 (1992)之方法經全甲基化及藉質譜術決定側寫資料。遵照複雜碳水化合物研究中心(Aoki K等人，J. Biol. Chem. 282:9127-42 (2007))發展之方法進行質譜分析。質量分析係使用NSI-LTQ/MS_n測定。簡言之，全甲基化聚糖溶解於1 mM氫氧化鈉於50%甲醇，以0.4微升/分鐘恆定流速直接擴散入儀器(LTQ，瑟摩費尼根(Thermo Finnigan))。MS分析係以陽離子模式進行。用於總離子映射，自500至2000之m/z範圍之自動化MS/MS分析(於35碰撞能)以連續2.8個質量單位窗其係重疊前一窗達2質量單位掃描。

進行總離子映射來檢查是否存在有指示聚糖之斷片離子。自m/z 500至m/z 2000之全部MS/MS資料取得及以人工分析原始資料。藉NSI-總離子映射所得之各種層析圖及表係顯示於第40A圖及第40B圖。此層析圖藉掃描濾波器處理；m/z 139之中性損耗乃高甘露糖型聚糖之特徵。總離子映射顯示本樣本含有具有長甘露糖鏈之一串列高甘露糖型聚糖。此等結果係類似於天然裂殖壺菌屬分泌蛋白質上檢測得之N-聚糖結構，如藉實例17之相同方法測定(參考第30圖)。

實例20

黑麴菌反錄酶於裂殖壺菌屬之表現

載體 pAB0018(ATCC 存取號碼 PTA-9616)以 HindIII 消化，以綠豆核酸酶處理、純化及以可產生四個不等大小片段之 KpnI 進一步消化。2552 bp 片段係於瓊脂凝膠藉標準電穿孔技術分離及使用商用 DNA 純化套件組純化。然後進行以 PmeI 及 Kpn 進行 pAB0018 之第二次消化。由此消化產物分離及純化 6732 bp 片段及接合至 2552 bp 片段。然後遵照製造商之方案，接合產物用來轉形商業供應的勝任 DH5- α 大腸桿菌細胞株系(英維仇貞公司)。增殖得自安比西林(ampicillin)抗藥性純株之質體、純化及然後藉限剪消化或 PCR 篩檢來證實接合產生期望的質體結構。一個經證實的質體定名為 pCL0120。參考第 43 圖。

得自真菌黑麴菌(基因存庫存取號碼 S33920)之 Suc1 反錄酶蛋白質之成熟形式使用第 42 圖之裂殖壺菌屬密碼子使用表經密碼子最佳化供於裂殖壺菌表現(密碼子最佳化係藉藍喜隆生技公司，華盛頓州波特海爾進行)。已經密碼子最佳化之序列經合成，所得多核苷酸序列融合至編碼裂殖壺菌屬 Sec1 信號肽(「Sec1 ss」)之多核苷酸序列作為 N 端先導子來替代內生性信號肽。所得「s1Suc1」核酸序列(SEQ ID NO:75)之編碼區顯示於第 44 圖。此已經密碼子最佳化之 s1Suc1 多核苷酸根據標準技術使用 5' 及 3' 限剪位置 BamHI 及 NdeI 用於插入及接合而選殖入載體 pCL0120。所得載體 pCL0137 之質體拼圖顯示於第 45 圖。野生型株系裂殖壺菌種屬 ATCC 20888 係以此載體轉形，所得純株係於含 SMM 之固

體SSFM培養基上選擇。然後SMM抗性純株再接種至含蔗糖作為唯一碳源之SSFM固體培養基來檢定分析生長情況。依據轉形實驗而定，50%至90%之SMM-抗性一次轉形株係可於蔗糖培養基上生長。

此處所述全部各個面相、實施例及選項皆可以任一種變化及全部變化組合。

本說明書所述全部公告案、專利案、及專利申請案皆係以引用方式併入此處，彷彿個別公告案、專利案、或專利申請案係特別地個別指示以引用方式併入此處。

【圖式簡單說明】

第1圖顯示裂殖壺菌屬Na/Pi-IIIb2轉運子蛋白質信號肽胺基酸序列(SEQ ID NO:1)。

第2圖顯示編碼SEQ ID NO:1之信號肽之多核苷酸序列(SEQ ID NO:2)。

第3圖顯示裂殖壺菌屬PUFA PKS OrfC啟動子區之多核苷酸序列(SEQ ID NO:3)。

第4圖顯示裂殖壺菌屬PUFA PKS OrfC終結子元件-1之多核苷酸序列(SEQ ID NO:4)。

第5圖顯示pSchizE之質體拼圖。

第6圖顯示pSchiz-sG之質體拼圖，也稱作為pCO0001。

第7圖顯示pSchiz-sGr之質體拼圖。

第8圖顯示pSchiz-cG之質體拼圖。

第9圖顯示於裂殖壺菌屬細胞之胞質及內質網(ER)的eGFP表現。第9A、9C及9E圖為螢光顯微相片。第9B、9D

及9F圖為光學顯微相片。第9A及9B圖顯示以pSchiz-sGr轉形之細胞的相同視野。第9C及9D圖顯示以pSchiz-cG轉形之細胞的相同視野。第9E及9F圖顯示以pSchiz-E轉形之細胞的相同視野。

第10A圖顯示於經以pSchiz-sGr轉形之裂殖壺菌屬細胞中經ER靶定之eGFP及核酸專一性染色4',6-二脒-2-苯基吡啶(DAPI)之複合螢光定位。第10B及10C圖分別顯示用於製造複合顯微相片之eGFP-ER染色及DAPI-核染色。第10D圖顯示相同視野之光學顯微相片。如第10A圖指示，ER膜包住細胞之各個核，各細胞含有多個核。指示一個細胞中一個核之相關特徵。

第11圖顯示得自四個裂殖壺菌屬轉形體純株之不含細胞之上清液及不含細胞之萃取樣本之西方墨點及相對應之經庫馬西(Coomassie)-染色之SDS-PAGE凝膠。「sup」及「cyto」分別係指不含細胞之上清液及不含細胞之萃取物。「sG」係指得自以pSchiz-sG轉形細胞之樣本。「sGr」係指得自以pSchiz-sGr轉形細胞之樣本。「cG」係指得自以pSchiz-cG轉形細胞之樣本。「vec(-)」係指得自以pSchiz-E10轉形細胞之樣本。「GFP(+)」係指已純化之重組GFP標準品(克隆泰公司(Clontech)，加州山景市)。凝膠載荷以3微克不含細胞之上清液蛋白質及1微克不含細胞之萃取物蛋白質樣本。於各對樣本亦即重組GFP標準品與分子量標記間發現有一空白線道。

第12圖顯示裂殖壺菌屬Sec1蛋白質轉運子蛋白質之首

30個胺基酸。胺基酸1至20組成信號肽序列(SEQ ID NO:37)。

第13圖顯示編碼SEQ ID NO:37之信號肽之多核苷酸序列(SEQ ID NO:38)。

第14圖顯示pSchiz-Cpt-sleGFP之質體拼圖，也稱作為pCL0001。

第15A圖顯示已分泌之eGFP蛋白質之西方墨點，及第15B圖顯示得自於不同發酵條件(例如「B26」、「B27」或「B28」發酵條件，如第15圖定義)下生長的三個裂殖壺菌屬培養之相對應經庫馬西染色之SDS-PAGE凝膠。線道1-19載荷指示量之蛋白質。LH=發酵時間，單位為小時。第15A圖之線道20載荷10奈克，及第15B圖之線道20載荷0.5奈克已純化之重組GFP標準品；得自裂殖壺菌屬之eGFP帶由於含有聯結子序列故該帶係略大於對照帶。

第16圖顯示pSchiz-Cpt-slkappabh之質體拼圖。

第17圖顯示藉裂殖壺菌屬分泌κ抗體亞單位之西方墨點。L=不含細胞之萃取物；S=不含細胞之上清液。

第18A圖顯示表現κ抗體亞單位之西方墨點。獲得培養上清液樣本之培養時間係指示於第18圖頂。「wt」係指「野生型」(亦即未經轉形之)裂殖壺菌屬。「+cptSlkappa」係指以含有編碼人κ抗體片段之密碼子最佳化基因、ORFC啟動子及終結子、及編碼Sec1信號肽之序列之載體轉形之裂殖壺菌屬。第18B圖顯示於以cptSlkappa轉形之裂殖壺菌屬培養上清液中總蛋白質(根據布萊佛(Bradford)檢定分析)及抗

體κ鏈(透過ELISA檢定分析)之蓄積。

第19圖顯示裂殖壺菌屬EF1短啟動子多核苷酸序列 (SEQ ID NO:42)。

第20圖顯示裂殖壺菌屬EF1長啟動子多核苷酸序列 (SEQ ID NO:43)。

第21圖顯示裂殖壺菌屬60S短啟動子多核苷酸序列 (SEQ ID NO:44)。

第22圖顯示裂殖壺菌屬60S長啟動子多核苷酸序列 (SEQ ID NO:45)。

第23圖顯示裂殖壺菌屬Sec1啟動子多核苷酸序列 (SEQ ID NO:46)。

第24圖顯示pAB0011之質體拼圖。

第25圖顯示pAB0018之質體拼圖。

第26圖顯示pAB0022之質體拼圖。

第27圖顯示取自於以含有藉EF1啟動子(短版本)、EF1啟動子(長版本)、60S啟動子(短版本)、60S啟動子(長版本)、SEC1啟動子、及OrfC啟動子分別驅動之eGFP基因的表現載體轉形之裂殖壺菌屬培養之不含細胞上清液樣本中之eGFP之西方墨點。

第28圖顯示與藉OrfC啟動子(pCL0001-4)或EF1-L啟動子(AB0018-9及-10)驅動eGFP表現相關聯之轉形株細胞系之螢光顯微術。

第29圖顯示藉全甲基化N-聚糖之NSI-全MS分析測得之於天然裂殖壺菌屬所分泌之蛋白質上檢測得之N-聚糖結構。

第30圖顯示藉全甲基化N-聚糖之NSI-全離子映射測得之於天然裂殖壺菌屬所分泌之蛋白質上檢測得之N-聚糖結構。

第31圖顯示 pSchiz-EPCT(+)-s1Suc2_CL0076 之質體拼圖也定名為 pCL0076。

第32圖顯示得自於蔗糖-SSFM上生長之以 pCL0076 轉形之裂殖壺菌種屬 ATCC 20888 培養之細胞丸粒之乾重(克/升)。轉形株稱作為 1-1、1-3、1-24、3-1、3-2、3-5、3-21、4-1、4-24、及 4-31。「對照組」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬 ATCC 20888 細胞。

第33圖顯示得自於蔗糖-SSFM上生長之以 pCL0076 轉形之裂殖壺菌種屬 ATCC 20888 培養之細胞丸粒之脂肪含量(以乾細菌量之重量百分比表示)。轉形株稱作為 1-1、1-3、1-24、3-1、3-2、3-5、3-21、4-1、4-24、及 4-31。「對照組」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬 ATCC 20888 細胞。

第34圖顯示得自於蔗糖-SSFM上生長之以 pCL0076 轉形之裂殖壺菌種屬 ATCC 20888 培養之細胞丸粒隨時間之經過測量得之乾重(克/升)。轉形株稱作為 1-3 及 3-5。「對照組」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬 ATCC 20888 細胞。

第35圖顯示得自於蔗糖-SSFM上生長之二轉形株培養之細胞丸粒之脂肪含量(以乾細菌量之重量百分比表示)。轉形株稱作為 1-3 及 3-5。「對照組」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬 ATCC 20888 細胞。

第36圖顯示得自以pCL0076轉形且於蔗糖-SSFM中生長2日或7日後收穫之裂殖壺菌屬株系B76-32培養之細胞丸粒之乾重(克/升)。「2118*」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬ATCC 20888細胞之次單離株。「B76-32**」係指於葡萄糖-SSFM上生長之B76-32親代株系。

第37圖顯示得自以pCL0076轉形且於蔗糖-SSFM中生長2日或7日後收穫之裂殖壺菌屬株系B76-32培養之細胞丸粒之脂肪含量。各樣本之最右欄顯示脂肪含量，以乾生質之重量百分比表示。各樣本之最左欄顯示由含18個或更少個碳(淺灰)或20個或更多個碳(中灰)之鹼基所組成之總脂肪百分比。「2118*」係指於葡萄糖-SSFM上生長之野生型裂殖壺菌種屬ATCC 20888細胞之次單離株。「B76-32**」係指於葡萄糖-SSFM上生長之B76-32親代株系。

第38A圖顯示反錄酶蛋白質之西方墨點及第38B圖顯示相對應之經庫馬西染色之SDS-PAGE凝膠。如西方墨點頂端指示，釀酒酵母(*S. cerevisiae*)反錄對照組及pCL0076轉形株1-3之3日培養之不含細胞的上清液分別係以5微克、2.5微克、1.25微克及0.625微克之數量載荷。

第39A圖顯示藉反應速率呈蔗糖濃度之函數示例顯示之反錄酶活性檢定分析。第39B圖顯示用於計算 K_m 及 V_{max} 之標準蘭威佛-伯克(Lineweaver-Burk)作圖。

第40A圖顯示藉全甲基化N-聚糖之NSI總離子映射測得於經裂殖壺菌屬分泌之蛋白質上檢測得之N-聚糖結構式。

第40B圖顯示經由全甲基化N-聚糖之NSI-總離子映射

所得之聚糖種類列表。

第41A圖及第41B圖顯示基於使用信號P演繹法則預測得裂殖壺菌屬之天然信號序列。例如參考Bendsten等人，J. Mol. Biol. 340: 783-795 (2004)；Nielsen及Krogh，Proc. Int. Conf. Intell. Syst. Mol. Biol. 6:122-130 (1998)；Nielsen等人，蛋白質基因改造12:3-9 (1999)；Emanuelsson等人，自然方案2:953-971 (2007)。

第42圖顯示裂殖壺菌屬之密碼子使用表。

第43圖顯示pCL0120之質體拼圖。

第44圖顯示編碼得自裂殖壺菌屬Sec1信號肽融合至得自黑麴菌(*Aspergillus niger*)之成熟Suc1反錄酶(基因存庫存取號碼S33920)之密碼子最佳化核酸序列(SEQ ID NO:75)。

第45圖顯示pCL0137_EPCT(+)-s1Suc1之質體拼圖，也定名為pCL0137。

【主要元件符號說明】

(無)

序列表

<110> Apt, Kirk E.
Lippmeier, James Casey
Wang, Jun
Wynn, James P.

<120> 於網黏菌門微生物中生產蛋白質之技術

<130> 2715.013TWOB/JUK/SAS/BNC

<140> To Be Assigned
<141> Herewith

<150> 61/290,441
<151> 2009-12-28

<150> 61/160,618
<151> 2009-03-16

<160> 75

<170> PatentIn version 3.3

<210> 1
<211> 35
<212> PRT
<213> 裂殖壺菌屬

<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> 裂殖壺菌屬

<400> 2
atggccaaca tcatggccaa cgtcacgccc cagggcgctg ccaagggctt tggcctcttt 60
gtcggcggtgc tcttctttct ctactggttc cttgtcggcc tcgcc 105

<210> 3
<211> 1999
<212> DNA
<213> 裂殖壺菌屬

<400> 3
ccgcgaatca agaaggtagg cgcgctgcga ggcgcggcgg cggagcggag cgagggagag 60
ggagagggag agagagggag ggagacgtcg ccgcggcggg gcctggcctg gcctggtttg 120
gcttggtcag cgcggccttg tccgagcgtg cagctggagt tgggtggatt catttggtt 180
ttcttttggt tttgtttttc tctctttccc ggaaagtgtt ggccggtcgg tgttctttgt 240
tttgatttct tcaaaagttt tgggtggttg ttctctctct tggctctctg tcaggcggtc 300
cgggtccacgc cccggcctct cctctcctct cctctcctct cctctccgtg cgtatacgta 360
cgtacgtttg tatacgtaca tacatccgc ccgccgtgcc ggcgagggtt tgctcagcct 420

```

ggagcaatgc gatgcgatgc gatgcgatgc gacgcgacgc gacgcgagtc actggttcgc 480
gctgtggctg tggcttgctt gcttacttgc tttcgagctc tcccgctttc ttctttcctt 540
ctcacgccac caccaacgaa agaagatcgg ccccggcacg ccgctgagaa gggctggcgg 600
cgatgacggc acgcgcgccc gctgccacgt tggcgctcgc tgctgctgct gctgctgctg 660
ctgctgctgc tgctgctgct gctgctgctt ctgcgcgcag gctttgccac gaggccggcg 720
tgctggccgc tgccgcttcc agtccgcgtg gagagatcga atgagagata aactggatgg 780
attcatcgag ggatgaatga acgatggttg gatgcctttt tcctttttca ggtccacagc 840
gggaagcagg agcgcgtgaa tctgccgcca tccgcatacg tctgcatcgc atcgcatcgc 900
atgcacgcat cgctcgccgg gagccacaga cgggcgacag ggcggccagc cagccaggca 960
gccagccagg caggcaccag agggccagag agcgcgcctc acgcacgcgc cgcagtgcgc 1020
gcatcgctcg cagtgcagac ctigattccc cgcgcggatc tccgcgagcc cgaaacgaag 1080
agcgcggtac gggcccatcc tagcgtcgcc tcgcaccgca tcgcatcgca tcgcgttccc 1140
tagagagtag tactcgacga aggcaccatt tccgcgctcc tcttcggcgc gatcgaggcc 1200
cccggcgccg cgacgatcgc ggcggccgcg gcgctggcgg cggccctggc gctcgcgctg 1260
gcggccgccg cgggcgtctg gccctggcgc gcgcgggcgc gcgaggagga gcggcagcgg 1320
ctgctcgccg ccagagaagg agcgcgcggg gcccggggag ggacggggag gagaaggaga 1380
aggcgcgcaa ggcggccccg aaagagaaga ccctggactt gaacgcgaag aagaagaaga 1440
aggagaagaa gttgaagaag aagaagaaga aggagaggaa gttgaagaag acgaggagca 1500
ggcgcgttcc aaggcgcgtt ctcttccgga ggcgcgttcc agctgcggcg gcggggcggg 1560
ctgcggggcg ggcgcggggc cgggtgcggg cagaggggac gcgcgcgcgg aggcggaggg 1620
ggccgagcgg gagcccctgc tgctgcgggg cgcgcgggccc gcagggtgtg gcgcgcgcac 1680
gacggaggcg acgacgccag cggccgcgac gacaaggccg gcggcgctcg cgggcggaag 1740
gccccgcgcg gagcaggggc gggagcagga caaggcgcag gagcaggagc agggccggga 1800
gcgggagcgg gagcgggcgg cggagcccga ggcagaacct aatcgagatc cagagcgagc 1860
agaggccggc cgcgagcccg agcccgcgcc gcagatcact agtaccgctg cggaatcaca 1920
gcagcagcag cagcagcagc agcagcagca gcagcagcag cagccacgag agggagataa 1980
agaaaaagcg gcagagacg 1999

```

<210> 4
 <211> 325
 <212> DNA
 <213> 裂殖壺菌屬

```

<400> 4
gatccgaaag tgaaccttgt octaaccgca cagcgaatgg cgggaggggg cgggctaaaa 60
gatcgtatta catagtattt ttcccctact ctttgtgttt gtcttttttt ttttttgaac 120
gcattcaagc cacttgtctt gggttacttg tttgtttgct tgcttgcttg cttgcttgcc 180
tgcttcttgg tcagacggcc caaaaaaggg aaaaaattca ttcattggcac agataagaaa 240
aagaaaaagt ttgtcgacca ccgtcatcag aaagcaagag aagagaaaca ctgcgctca 300
cattctcgct cgcgtaagaa totta 325

```

<210> 5
<211> 372
<212> DNA
<213> 印度鏈異壁菌

<400> 5
atggccaagt tgaccagtgc cgttccggtg ctcaccgcgc gcgacgtcgc cggagcggtc 60
gagttctgga ccgaccggct cgggttctcc cgggacttcg tggaggacga cttcgccgggt 120
gtgggtccggg acgacgtgac cctgttcatc agcgcgggtcc aggaccaggt ggtgccggac 180
aacaccctgg cctgggtgtg ggtgcgcggc ctggacgagc tgtacgccga gtggtcggag 240
gtcgtgtcca cgaacttccg ggacgcctcc gggccggcca tgaccgagat cggcgagcag 300
ccgtgggggc gggagttcgc cctgcgcgac ccggccggca actgcgtgca cttcgtggcc 360
gaggagcagg ac 372

<210> 6
<211> 2055
<212> DNA
<213> 裂殖壺菌屬

<400> 6
atgagcgcga cccgcgcggc gacgaggaca gcggcggcgc tgtcctcggc gctgacgacg 60
cctgtaaagc agcagcagca gcagcagctg cgcgtaggcg cggcgtcggc acggctggcg 120
gccgcggcgt tctcgtccgg cacgggcgga gacgcggcca agaaggcggc cgcggcgagg 180
gcgttctcca cgggacgcgg cccaacgcg acacgcgaga agagctcgct ggccacggtc 240
caggcggcga cggacgatgc gcgcttcgtc ggctgaccg gcgccccaat ctttcatgag 300
ctcatgcgcg agcaccaggt ggacaccatc tttggctacc ctggcggcgc cattctgccc 360
gtttttgatg ccatttttga gagtgcgcgc ttcaagttca ttctcgtcgc ccacgagcag 420
ggcgcgggcc acatggccga gggctacgcg cgcgccacgg gcaagcccgg cgttgtcctc 480
gtcacctcgg gccctggagc caccaacacc atcaccccga tcatggatgc ttacatggac 540
ggtacgccgc tgctcgtgtt caccggccag gtgccacact ctgctgtcgg cacggacgct 600
ttccaggagt gtgacattgt tggcatcagc cgcgcgtgca ccaagtggaa cgtcatggtc 660
aaggacgtga aggagctccc gcgcccatc aatgaggcct ttgagattgc catgagcggc 720
cgccccggtc ccgtgctcgt cgatcttctt aaggatgtga ccgccgttga gctcaaggaa 780
atgcccgaaca gctccccca ggttgctgtg cgccagaagc aaaaggtcga gcttttccac 840
aaggagcgca ttggcgctcc tggcacggcc gacttcaagc tcattgccga gatgatcaac 900
cgtgcgggagc gaccgctcat ctatgctggc cagggtgtca tgcagagccc gttgaatggc 960
ccggctgtgc tcaaggagtt cgoggagaag gccaacattc ccgtgaccac caccatgcag 1020
ggtctcggcg gctttgacga gcgtagtccc ctctccctca agatgctcgg catgcacggc 1080
tctgcctacg ccaactactc gatgcagaac gccgatctta tcctggcgct cgggtccccg 1140
tttgatgatc gtgtgacggg ccgcgttgac gcctttgctc cggaggctcg ccgtgccgag 1200
cgcgagggcc gcggtggcat cgttcacttt gagatttccc ccaagaacct ccacaaggtc 1260
gtccagccca ccgtcgcggc cctcggcgac gtggctgaga acctcgccaa cgtcacgccc 1320
cacgtgcagc gccaggagcg cgagccgtgg tttgcgcaga tcgccgattg gaaggagaag 1380
cacccttttc tgctcgagtc tgttgattcg gacgacaagg ttctcaagcc gcagcaggtc 1440

ctcacggagc ttaacaagca gattctcgag attcaggaga aggacgccga ccaggaggtc 1500
 tacatcacca cgggcgtcgg aagccaccag atgcaggcag cgagttcct tacctggacc 1560
 aagccgcgcc agtggatctc ctcggtggc gccggcacta tgggctacgg ccttcctcg 1620
 gccattggcg ccaagattgc caagcccgat gctattgta ttgacatoga tggatgatgct 1680
 tcttattoga tgaccggtat ggaattgatc acagcagccg aattcaaggt tggcgtgaag 1740
 attcttcttt tgcagaacaa ctttcagggc atgggtcaaga actggcagga tctcttttac 1800
 gacaagcgct actcgggcac cgccatgttc aaccgcgcgt tcgacaaggt cgccgatgcg 1860
 atgcgtgcc aaggtctcta ctgcgcgaaa cagtcggagc tcaaggacaa gatcaaggag 1920
 tttctcgagt acgatgaggg tcccgctctc ctcgaggttt tctgtggaaa ggacacgctc 1980
 gtcttgccca tgggtccccgc tggctttccg ctccacgaga tggctcctoga gcctcctaag 2040
 cccaaggacg cctaa 2055

<210> 7
 <211> 684
 <212> PRT
 <213> 裂殖壺菌屬

<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

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

6

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> 人造序列

<220>
 <223> 已突變的 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 65	Arg	Gly	Pro	Asn	Ala 70	Thr	Arg	Glu	Lys	Ser 75	Ser	Leu	Ala	Thr	Val 80
Gln	Ala	Ala	Thr	Asp 85	Asp	Ala	Arg	Phe	Val 90	Gly	Leu	Thr	Gly	Ala 95	Gln
Ile	Phe	His	Glu 100	Leu	Met	Arg	Glu	His 105	Gln	Val	Asp	Thr	Ile 110	Phe	Gly
Tyr	Pro	Gly 115	Gly	Ala	Ile	Leu	Pro 120	Val	Phe	Asp	Ala	Ile 125	Phe	Glu	Ser
Asp	Ala 130	Phe	Lys	Phe	Ile	Leu 135	Ala	Arg	His	Glu	Gln 140	Gly	Ala	Gly	His
Met 145	Ala	Glu	Gly	Tyr	Ala 150	Arg	Ala	Thr	Gly	Lys 155	Pro	Gly	Val	Val	Leu 160
Val	Thr	Ser	Gly	Pro 165	Gly	Ala	Thr	Asn	Thr 170	Ile	Thr	Pro	Ile	Met 175	Asp
Ala	Tyr	Met	Asp 180	Gly	Thr	Pro	Leu	Leu 185	Val	Phe	Thr	Gly	Gln 190	Val	Pro
Thr	Ser	Ala 195	Val	Gly	Thr	Asp	Ala 200	Phe	Gln	Glu	Cys	Asp 205	Ile	Val	Gly
Ile	Ser 210	Arg	Ala	Cys	Thr	Lys 215	Trp	Asn	Val	Met	Val 220	Lys	Asp	Val	Lys
Glu 225	Leu	Pro	Arg	Arg	Ile 230	Asn	Glu	Ala	Phe	Glu 235	Ile	Ala	Met	Ser	Gly 240
Arg	Pro	Gly	Pro	Val 245	Leu	Val	Asp	Leu	Pro 250	Lys	Asp	Val	Thr	Ala 255	Val
Glu	Leu	Lys	Glu 260	Met	Pro	Asp	Ser	Ser 265	Pro	Gln	Val	Ala	Val 270	Arg	Gln
Lys	Gln	Lys 275	Val	Glu	Leu	Phe	His 280	Lys	Glu	Arg	Ile	Gly 285	Ala	Pro	Gly
Thr	Ala 290	Asp	Phe	Lys	Leu	Ile 295	Ala	Glu	Met	Ile	Asn 300	Arg	Ala	Glu	Arg
Pro 305	Val	Ile	Tyr	Ala	Gly 310	Gln	Gly	Val	Met	Gln 315	Ser	Pro	Leu	Asn	Gly 320
Pro	Ala	Val	Leu	Lys 325	Glu	Phe	Ala	Glu	Lys 330	Ala	Asn	Ile	Pro	Val 335	Thr
Thr	Thr	Met	Gln 340	Gly	Leu	Gly	Gly	Phe 345	Asp	Glu	Arg	Ser	Pro 350	Leu	Ser

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 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> 人造序列

<220>
<223> 已突變的 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

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 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> 10
<211> 684
<212> PRT
<213> 人造序列

<220>
<223> 已突變的 ALS 3

<400> 10

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

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

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 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> 11
<211> 2052

<212> DNA
<213> 人造序列

<220>
<223> 已突變的 ALS 1

<400> 11
atgagcgcg cccgcgcggc gacgaggaca gcggcggcgc tgcctcggc gctgacgacg 60
cctgtaaagc agcagcagca gcagcagctg cgcgtaggcg cggcgtcggc acggctggcg 120
gccgcggcgt tctcgtccgg cacgggcgga gacgcggcca agaaggcggc cgcggcgagg 180
gcgttctcca cgggacgcgg cccaacgcg acacgcgaga agagctcgct ggccacggtc 240
caggcggcga cggacgatgc gcgcttcgtc ggctgaccg gcgccccaat cttcatgag 300
ctcatgcgcg agcaccaggc ggacaccatc ttgggtacc ctggcggcgc cattctgccc 360
gtttttgatg ccatttttga gaggtagcgc ttcaagttca ttctcgctcg ccacgagcag 420
ggcgccggcc acatggccga gggctacgcg cgcgccacgg gcaagcccgg cgttgtcctc 480
gtcacctcgg gccctggagc caccaacacc atcacccga tcatggatgc ttacatggac 540
ggtacgcgcg tgctcgtgtt caccggccag gtgcccact ctgctgtcgg cacggacgct 600
ttccaggagt gtgacattgt tggcatcagc cgcgcgtgca ccaagtggaa cgtcatggtc 660
aaggacgtga aggagctccc gcgcgcgcatc aatgaggcct ttgagattgc catgagcggc 720
cgcccgggtc ccgtgctcgt cgatcttcct aaggatgtga ccgccgttga gctcaaggaa 780
atgcccagca gctccccca ggttgctgtg cgccagaagc aaaaggctga gcttttccac 840
aaggagcgca ttggcgctcc tggcacggcc gacttcaagc tcattgccga gatgatcaac 900
cgtgcggagc gaccgcgtat ctatgctggc cagggtgtca tgcagagccc gttgaatggc 960
ccggctgtgc tcaaggagt cgcggagaag gccaacattc ccgtgaccac caccatgcag 1020
ggtctcggcg gctttgacga gcgtagtccc ctctccctca agatgctcgg catgcacggc 1080
tctgcctacg ccaactactc gatgcagaac gccgatctta tcctggcgct cggtgcccgc 1140
tttgatgata gtgtgacggg ccgcgttgac gcctttgctc cggaggctcg ccgtgccgag 1200
cgcgagggcc gcggtggcat cgttcacttt gagatttccc ccaagaacct ccacaaggtc 1260
gtccagccca ccgtcgcggc cctcggcgac gtggctcaga acctcgccaa cgtcacgccc 1320
cacgtgcagc gccaggagcg cgagccgtgg tttgcgcaga tcgccgattg gaaggagaag 1380
cacccttttc tgctcgagtc tgttgattcg gacgacaagg ttctcaagcc gcagcaggtc 1440
ctcacggagc ttaacaagca gattctcgag attcaggaga aggacgccga ccaggaggtc 1500
tacatcacca cgggcgtcgg aagccaccag atgcaggcag cgcagttcct tacctggacc 1560
aagccgcgcc agtggatctc ctcggtggc gccggcacta tgggctacgg ccttccctcg 1620
gccattggcg ccaagattgc caagcccgat gctattgtta ttgacatga tggatgatgct 1680
tcttattcga tgaccggtat ggaattgatc acagcagccg aattcaaggt tggcgtgaag 1740
attcttcttt tgcagaacaa ctttcagggc atgggtcaaga acgttcagga tctcttttac 1800
gacaagcgct actcgggcac cgccatgttc aaccgcgct tcgacaaggt cgccgatgcg 1860
atgcgtgcca agggctctta ctgcgcgaaa cagtcggagc tcaaggacaa gatcaaggag 1920
tttctcgagt acgatgaggg tcccgctctc ctogaggttt tcgtggacaa ggacacgctc 1980
gtcttgccca tgggtccccg tggctttccg ctccacgaga tggctcctga gcctcctaag 2040

cccaaggacg cc

2052

<210> 12
<211> 2052
<212> DNA
<213> 人造序列

<220>
<223> 已突變的 ALS 2

<400> 12
atgagcgcga cccgcgcggc gacgaggaca gcgcgcgcgc tgtcctcggc gctgacgacg 60
cctgtaaagc agcagcagca gcagcagctg cgcgtaggag cggcgtcggc acggctggcg 120
gccgcggcgt tctcgtccgg cacgggcgga gacgcggcca agaaggcggc cgcggcgagg 180
gcgttctcca cgggacgcgg cccaacgcg acacgcgaga agagctcgct ggccacggtc 240
caggcggcga cggacgatgc gcgcttcgtc ggcttgaccg gcgccccaaat ctttcatgag 300
ctcatgcgag agcaccaggc ggacaccatc tttggctacc ctggcgcgcg cattctgccc 360
gtttttgatg ccatttttga gaggtagcgc ttcaagttca ttctcgctcg ccacgagcag 420
ggcgccggcc acatggccga gggctacgcg cgcgccacgg gcaagcccgg cgttgtcctc 480
gtcacctcgg gccctggagc caccaacacc atcacccga tcatggatgc ttacatggac 540
ggtacgccgc tgctcgtgtt caccggccag gtgcagacct ctgctgtcgg cacggacgct 600
ttccaggagt gtgacattgt tggcatcagc cgcgcgtgca ccaagtggaa cgtcatggtc 660
aaggacgtga aggagctccc gcgcgcgcatc aatgaggcct ttgagattgc catgagcggc 720
cgcccgggtc ccgtgctcgt cgatcttcct aaggatgtga ccgccgttga gctcaaggaa 780
atccccgaca gctccccca ggttgctgtg cgccagaagc aaaaggtcga gcttttccac 840
aaggagcgca ttggcgctcc tggcacggcc gacttcaagc tcattgccga gatgatcaac 900
cgtgcggagc gacccgtcat ctatgctggc cagggtgtca tgcagagccc gttgaatggc 960
ccggctgtgc tcaaggagt cgcggagaag gccaacattc ccgtgaccac caccatgcag 1020
ggtctcggcg gctttgacga gcgtagtccc ctctccctca agatgctcgg catgcacggc 1080
tctgcctacg ccaactactc gatgcagaac gccgatctta tcctggcgct cgggtgcccgc 1140
tttgatgatc gtgtgacggg ccgcgttgac gcctttgctc cggaggctcg ccgtgccgag 1200
cgcgagggcc gcggtggcat cgttcacttt gagatttccc ccaagaacct ccacaaggtc 1260
gtccagccca ccgtcgcggt cctcggcgac gtggtogaga acctcgccaa cgtcacgccc 1320
cacgtgcagc gccaggagcg cgagccgtgg tttgcgcaga tcgccgattg gaaggagaag 1380
cacccttttc tgctcgagtc tgttgattcg gacgacaagg ttctcaagcc gcagcaggtc 1440
ctcacggagc ttaacaagca gattctcgag attcaggaga aggacgccga ccaggaggtc 1500
tacatcacca cgggcgtcgg aagccaccag atgcaggcag cgagttcct tacctggacc 1560
aagccgcgcc agtggatctc ctcggtgggc gccggcacta tgggctacgg ccttccctcg 1620
gccattggcg ccaagattgc caagcccgat gctattgtta ttgacatcga tggatgatgct 1680
tcttattcga tgaccggtat ggaattgatc acagcagccg aattcaaggc tggcgtgaag 1740
attcttcttt tgcagaacaa ctttcagggc atggtcaaga actggcagga tctcttttac 1800
gacaagcgct actcgggcac cgccatgttc aaccgcgcgt tcgacaaggc cgccgatgcg 1860
atgcgtgccg agggctctcta ctgcgcgaaa cagtcggagc tcaaggacaa gatcaaggag 1920

tttctogagt acgatgaggg tcccgtcctc ctogaggttt tcgtggacaa ggacacgctc 1980
gtcttgccca tgggtccccgc tggctttccg ctccacgaga tggctcctga gcctcctaag 2040
cccaaggacg cc 2052

<210> 13
<211> 2055
<212> DNA
<213> 人造序列

<220>
<223> 已突變的 ALS 3

<400> 13
atgagcgcgga cccgcgcggc gacgaggaca gcggcggcgc tgtcctcggc gctgacgacg 60
cctgtaaagc agcagcagca gcagcagctg cgcgtaggcg cggcgtcggc acggctggcg 120
gccgcggcgt tctcgtccgg cacgggcgga gacgcggcca agaaggcggc cgcggcgagg 180
gcgttctcca cgggacgcgg cccaacgcg acacgcgaga agagctcgtt ggccacggtc 240
caggcggcga cggacgatgc gcgcttcgtc ggcttgaccg gcgccccaat cttcatgag 300
ctcatgcgcg agcaccaggt ggacaccatc tttggctacc ctggcggcgc cattctgccc 360
gtttttgatg ccatttttga gagtgcgccc ttcaagttca ttctcgtctg ccacgagcag 420
ggcgccggcc acatggccga gggctacgcg cgcgccacgg gcaagcccg cgttgtcctc 480
gtcacctcgg gccctggagc caccaacacc atcacccga tcattggatgc ttacatggac 540
ggtacgcgcg tgctcgtgtt caccggccag gtgcagacct ctgctgtcgg cacggacgct 600
ttccaggagt gtgacattgt tggcatcagc cgcgcgtgca ccaagtggaa cgtcatggtc 660
aaggacgtga aggagctccc gcgccgcacg aatgaggcct ttgagattgc catgagcggc 720
cgcccggttc ccgtgctcgt cgatcttcct aaggatgtga ccgccgttga gctcaaggaa 780
atgcccga ca gctccccca ggttgctgtg cgcgcagaagc aaaaggctga gcttttccac 840
aaggagcgca ttggcgcctc tggcacggcc gacttcaagc tcattgccga gatgatcaac 900
cgtgcggagc gacccgtcat ctatgctggc cagggtgtca tgcagagccc gttgaatggc 960
ccggctgtgc tcaaggagtt cgcggagaag gccaacattc ccgtgaccac caccatgcag 1020
ggtctcggcg gctttgacga gcgtagtccc ctctccctca agatgctcgg catgcacggc 1080
tctgcctacg ccaactactc gatgcagaac gccgatctta tcctggcgtt cgggtgccgc 1140
tttgatgacg gtgtgacggg ccgcgttgac gcctttgctc cggaggctcg ccgtgccgag 1200
cgcgagggcc gcggtggcat cgttcacttt gagatttccc ccaagaacct ccacaaggtc 1260
gtccagccca ccgtcgcggt cctcggcgac gtggtcgaga acctcgcaa cgtcacgccc 1320
cacgtgcagc gccaggagcg cgagccgtgg tttgcgcaga tcgccgattg gaaggagaag 1380
cacccttttc tgctcgagtc tgttgattcg gacgacaagg ttctcaagcc gcagcaggtc 1440
ctcacggagc ttaacaagca gattctcgag attcaggaga aggacgccga ccaggaggtc 1500
tacatcacca cgggcgtcgg aagccaccag atgcaggcag cgcagttcct tacctggacc 1560
aagccgcgcc agtggatctc ctcggtggc gccggcacta tgggctacgg ccttccctcg 1620
gccattggcg ccaagattgc caagcccgat gctattgtta ttgacatga tggatgatgct 1680
tcttattcga tgaccggtat ggaattgatc acagcagccg aattcaaggt tggcgtgaag 1740

attcttcttt tgcagaacaa ctttcagggc atggtcaaga acgttcagga tctcttttac 1800
gacaagcgct actcgggcac cgccatgttc aaccgcgcgt tcgacaaggt cgccgatgcg 1860
atgcgtgcca aggggtctcta ctgcgcgaaa cagtcggagc tcaaggacaa gatcaaggag 1920
tttctcgagt acgatgaggg tcccgtcctc ctcgaggttt tcgtggacaa ggacacgctc 1980
gtcttgccca tgggtccccgc tggctttccg ctccacgaga tggtcctcga gcctcctaag 2040
ccaaggacg cctaa 2055

<210> 14
<211> 12
<212> DNA
<213> 裂殖壺菌屬

<400> 14
cacgacgagt tg 12

<210> 15
<211> 53
<212> PRT
<213> 裂殖壺菌屬

<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> 人造序列

<220>
<223> 合成引子 S4termF

<400> 16
gatcccatgg cacgtgctac g 21

<210> 17
<211> 22
<212> DNA
<213> 人造序列

<220>
<223> 合成引子 S4termR

<400> 17
ggcaacatgt atgataagat ac 22

<210> 18
<211> 21
<212> DNA
<213> 人造序列

<220>

<223>	合成引子 C2mcsSmaF	
<400>	18	
	gatccccggg ttaagcttgg t	21
<210>	19	
<211>	20	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 C2mcsSmaR	
<400>	19	
	actggggccc gtttaaactc	20
<210>	20	
<211>	34	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 5' tubMCS_BglI	
<400>	20	
	gactagatct caattttagg cccccactg accg	34
<210>	21	
<211>	34	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 3' SV40MCS_Sal	
<400>	21	
	gactgtcgac catgtatgat aagatacatt gatg	34
<210>	22	
<211>	28	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 5' ALSproNde3	
<400>	22	
	gactcatatg gcccaggcct actttcac	28
<210>	23	
<211>	34	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 3' ALStermBglII	
<400>	23	
	gactagatct ggggtcaaggc agaagaattc cgcc	34
<210>	24	
<211>	97	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 sec. Gfp5' 1b	
<400>	24	



tactggttcc ttgtcggcct cgcccttctc ggcgatggct tcaaggtcat cgccggtgac	60
tccgccggtgta cgctcttcat ggtgagcaag ggcgagg	97
<210> 25	
<211> 32	
<212> DNA	
<213> 人造序列	
<220>	
<223> 合成引子 sec. Gfp3' Spe	
<400> 25	
gatcggtacc ggtgttcttt gttttgattt ct	32
<210> 26	
<211> 105	
<212> DNA	
<213> 人造序列	
<220>	
<223> 合成引子 sec. Gfp5' Bam	
<400> 26	
taatggatcc atggccaaca tcatggccaa cgtcacgccc cagggcgtcg ccaagggtt	60
tggcctcttt gtcggcgtgc tcttctttct ctactggttc cttgt	105
<210> 27	
<211> 40	
<212> DNA	
<213> 人造序列	
<220>	
<223> 合成引子 ss. eGfpHELD3' RV	
<400> 27	
cctgatatct tacaactcgt cgtggttgta cagctcgtcc	40
<210> 28	
<211> 105	
<212> DNA	
<213> 人造序列	
<220>	
<223> 合成引子 sec. Gfp5' Bam2	
<400> 28	
taatggatcc atggccaaca tcatggccaa cgtcacgccc cagggcgtcg ccaagggtt	60
tggcctcttt gtcggcgtgc tcttctttct ctactggttc cttgt	105
<210> 29	
<211> 27	
<212> DNA	
<213> 人造序列	
<220>	
<223> 合成引子 prREZ15	
<400> 29	
cgtacccgc gaatcaagaa ggtaggc	27
<210> 30	
<211> 27	
<212> DNA	
<213> 人造序列	
<220>	

<223>	合成引子 prREZ16	
<400>	30	
	cggatccggt ctctgccgct ttttctt	27
<210>	31	
<211>	30	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 prREZ17	
<400>	31	
	cggatccgaa agtgaacctt gtcctaacct	30
<210>	32	
<211>	27	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 prREZ18	
<400>	32	
	ctctagacag atccgcacca tcggccg	27
<210>	33	
<211>	32	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 5' eGFP_kpn	
<400>	33	
	gactgttacc atggtgaagc aagggcgagg ag	32
<210>	34	
<211>	34	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 3' eGFP_xba	
<400>	34	
	gacttctaga ttacttgtac agctcgtcca tgcc	34
<210>	35	
<211>	32	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 5' ORFCproKpn-2	
<400>	35	
	gatcgttacc ggtgttcttt gttttgattt ct	32
<210>	36	
<211>	31	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 3' ORFCproKpn-2	
<400>	36	



gatcgggtacc gtctctgccg ctttttcttt a		31
<210> 37		
<211> 20		
<212> PRT		
<213> 裂殖壺菌屬		
<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> 裂殖壺菌屬		
<400> 38		
atgaagttcg cgacctcggg cgcaattttg cttgtggcca acatagccac cgccctcgcg		60
<210> 39		
<211> 28		
<212> DNA		
<213> 人造序列		
<220>		
<223> 合成引子 5' ss-X Bgl long		
<400> 39		
gactagatct atgaagttcg cgacctcg		28
<210> 40		
<211> 33		
<212> DNA		
<213> 人造序列		
<220>		
<223> 合成引子 3' ritx_kap_bh_Bgl		
<400> 40		
gactagatct tcagcactca ccgcgggtaa agg		33
<210> 41		
<211> 702		
<212> DNA		
<213> 人造序列		
<220>		
<223> 最佳化序列		
<400> 41		
atgaagttcg cgacctcggg cgcaattttg cttgtggcca acatagccac cgccctcgcg		60
cagatcgtcc tcagccagtc ccccgccatc ctttcggtt cccccgggga gaaggtgacc		120
atgacctgcc gcgctagctc ctccgtctcg tacatccact ggttccagca gaagcccggc		180
tcgtcccccag gccctggat ctacgccacc tccaacctcg cctccggtgt tcccgttcgt		240
ttttccggtt ccggttccgg cacctcctac tccctcacca tctcccgcgt cgaggccgag		300
gatgccgcca cctactactg ccagcagtg accagcaacc ccccccactt cggcggtggt		360
acgaagctcg agattaagcg caccgtcgcc gcccctccg tcttcatttt tccccctcc		420
gatgagcagc tcaagtccgg taccgcctcc gtcgtttgcc tcctcaacaa cttctacccc		480

cgtgaggcca aggtccagtg gaaggctgac aacgcgcttc agtccggtaa ctcccaggag 540
tccgtcaccg agcaggattc gaaggacagc acctactccc tctcctccac cctcaccctc 600
tccaaggccg actacgagaa gcacaaggtc tacgcctgcg aggtcacgca ccagggtctt 660
tcctcccccg tcacgaagtc ctttaaccgc ggtgagtgct ga 702

<210> 42
<211> 853
<212> DNA
<213> 裂殖壺菌屬

<400> 42
ctccatcgat cgtgcggtca aaaagaaagg aagaagaaag gaaaaagaaa ggcgtgcgca 60
cccgagtgcg cgctgagcgc ccgctcgcgg ccccgcgagg cctccgcgtt agtccccgcc 120
ccgcgcgcgc cagtcccccg ggaggcatcg cgacactctc gccgccccct cgcgctcgc 180
cgattccccg cctccccctt tccgcttctt cgccgcctcc gctcgcggcc gcgtcgcgccg 240
cgccccgctc cctatctgct ccccaggggg gcaactccga ccttttgccg ccgctgccgc 300
cgccgcggcc gcccgccgc cctggtttcc cccgcgagcg cggccgcgtc gccgcgcaaa 360
gactcgcgcc gtgccgcccc gagcaacggg tggcgggcgg gcggcgggcg gcggggcgcg 420
gcggcgcgta ggcggggcta ggcgccggct aggcgaaacg ccgcccccg gcgccgccgc 480
cgcccgctcc agagcagtcg ccgcgccaga ccgccaacgc agagaccgag accgaggtag 540
gtcgcgcccc agcacgccgc gacgcgcggc agggacgagg agcacgacgc cgcgccgcgc 600
cgcgcggggg gggggaggga gaggcaggac gcgggagcga gcgtgcatgt ttccgcgcga 660
gacgacgccg cgcgcgctgg agaggagata aggcgcttgg atcgcgagag ggccagccag 720
gctggaggcg aaaatgggtg gagaggatag tatcttgctg gcttggacga ggagactgac 780
gaggaggacg gatacgtcga tgatgatgtg cacagagaag aagcagttcg aaagcgacta 840
ctagcaagca agg 853

<210> 43
<211> 1064
<212> DNA
<213> 裂殖壺菌屬

<400> 43
ctcttatctg cctcgcgcgc ttgaccgccg cttgactctt ggcgcttgcc gctcgcaccc 60
tgctcgcctc gcgcaggcgg gcgggcgagt ggggtgggtcc gcagccttcc gcgctcgcgc 120
gctagctcgc tcgcgccgtg ctgcagccag cagggcagca ccgcacggca ggcagggtccc 180
ggcgcggtat gatcgatcca tcgatccatc gatccatcga tcgtgcggtc aaaaagaaag 240
gaagaagaaa ggaaaaagaa aggcgtgcgc acccgagtgc gcgctgagcg cccgctcgcg 300
gtcccgcgga gcctcccggt tagtccccgc cccgcgcgc gcagtcccc gggaggcatc 360
gcgcacctct cgccgcccc tcgcgcctcg ccgattcccc gcctccccct ttccgcttct 420
tcgccgcctc cgctcgcggc cgcgtcgcgc gcgccccgct ccctatctgc tccccagggg 480
ggcactccgc accttttgcg cccgctgccg ccgcgcggc cgccccgccc ccctggtttc 540
ccccgcgagc gcggcccggt cgccgcgcaa agactcgccg cgtgccgccc cgagcaacgg 600
gtggcgggcg gcggcgggcg ggcgggggcg ggcgggcggt aggcgggggt aggcgccggc 660

taggcgaaac gccgccccg ggcgccgccc cgcgccgctc cagagcagtc gccgcgccag 720
accgccaacg cagagaccga gaccgaggta cgtcgcgccc gagcacgccg cgacgcgcgg 780
cagggacgag gagcacgacg ccgcgcccgc cgcgcggggg ggggggaggg agaggcagga 840
cgcgggagcg agcgtgcatg tttccgcgcg agacgacgcc gcgcgcgctg gagaggagat 900
aaggcgcttg gatcgcgaga gggccagcca ggctggaggc gaaaatgggt ggagaggata 960
gtatcttgcg tgcttgacg aggagactga cgaggaggac ggatacgtcg atgatgatgt 1020
gcacagagaa gaagcagttc gaaagcgact actagcaagc aagg 1064

<210> 44
<211> 837
<212> DNA
<213> 裂殖壺菌屬

<400> 44
cttcgctttc tcaacctatc tggacagcaa tccgccactt gccttgatcc ccttccgcgc 60
ctcaatcaact cgctccacgt ccctcttccc cctcctcacc tccgtgcttt ctctgcccc 120
cccccccccg ccgcggcggtg gcgcgcgctg gcgccgcggc cgcgacacct tccatactat 180
cctcgctccc aaaatgggtt gcgctatagg gcccgctag gcgaaagtct agcaggcact 240
tgcttggcgc agagccgccg cggccgctcg ttgcccgga tggagaggga gagagagccc 300
gcctcgataa gcagagacag acagtgcgac tgacagacag acagagagac tggcagaccg 360
gaatacctcg aggtgagtgc ggcgcgggcg agcgggcggg agcgggagcg caagagggac 420
ggcgcgggcg ggcgggccctg gcgcagcccg cggcgctatt tcgtgcgcag cgccgagcag 480
cgggacgggc ggctggctga tggttgaagc ggggcggggg gaaatgtag atgagatgat 540
catcgacgac ggtccgtgcg tcttggctgg cttggctggc ttggctggcg ggctgccgt 600
gtttgcgaga aagaggatga ggagagcgac gaggaaggac gagaagactg acgtgtaggg 660
cgcgcgatgg atgatcgatt gattgattga ttgattgggt gattggctgt gtggtcgatg 720
aacgtgtaga ctcaggagc gtggttaaat tgttcttgcg ccagacgcga ggactccacc 780
cccttctttc gcctttacac agcctttttg tgaagcaaca agaaagaaaa agccaag 837

<210> 45
<211> 1020
<212> DNA
<213> 裂殖壺菌屬

<400> 45
ctttttccgc tctgcataat cctaaaagaa agactataacc ctagtactg tacaaatggg 60
acatttctct cccgagcgat agctaaggat ttttgcttcg tgtgactgt gtgctctggc 120
cgcgcatcga aagtccagga tcttactgtt tctctttcct ttcctttatt tctgtttctc 180
ttcttcgctt tctcaacctt tctggacagc aatccgccac ttgccttgat ccccttccgc 240
gcctcaatca ctcgctccac gtccctcttc cccctcctca tctccgtgct ttctctcgcc 300
cccccccccg ccgccgcggc gtgcgcgcgc gtggcgccgc ggccgcgaca ccttcatac 360
tatcctcgct cccaaaatgg gttgcgctat agggcccggc taggcgaaag tctagcaggc 420
acttgcttgg cgagagccg ccgcggccgc tcgttgccgc ggatggagag ggagagagag 480
cccgctcga taagcagaga cagacagtgc gactgacaga cagacagaga gactggcaga 540
ccggaatacc tcgaggtagg tgcggcgcgg gcgagcgggc gggagcggga gcgcaagagg 600

gacggcgcg ggcggcgcc ctgcgcgacg ccgcgcgta ttctcgtgcg cagcgccgag 660
cagcgggacg ggcggctggc tgatggttga agcgggcg ggtgaaatgt tagatgagat 720
gatcatcgac gacggtccgt gcgtcttggc tggcttggct ggcttggctg gcgggcctgc 780
cgtgtttgcg agaaagagga tggagagagc gacgaggaag gacgagaaga ctgacgtgta 840
gggcgcgcga tggatgatcg attgattgat tgattgattg gttgattggc tgtgtggctg 900
atgaacgtgt agactcaggg agcgtggta aattgttctt gcgccagacg cgaggactcc 960
accccttctt ttgccttta cacagccttt ttgtgaagca acaagaaaga aaaagccaag 1020

<210> 46
<211> 1416
<212> DNA
<213> 裂殖壺菌屬

<400> 46
cccgtccttg acgccttgcg ttccggcgcg gccatcgatt caattcaccg atccgatacg 60
ttccgcccc tcacgtccgt ctgcgcacga cccctgcacg accacgcaa ggccaacgcg 120
ccgtcagct cagcttgcg acgagtcgca cgtcacatat ctgagatgca ttgctgcct 180
gcctgcctgc ctgcctgcct gcctgcctgc ctgcctgcct cagcctctct ttgctctctc 240
tgcgggcgcc gctgcgacgc gctgtacagg agaatgactc caggaagtgc ggctgggata 300
cgcgctggcg tcggccgtga tgcgcgtgac ggcggcgggg cagggccggc acgggttgag 360
cagaggacga agcaggcgga gacgagacag gccaggcgcg gggagcgctc gctgccgtga 420
gcagcagacc agggcgcagg aatgtacttt tcttgcggga gcggagacga ggctgccggc 480
tgctggctgc cggttgctct gcacgcgccg ccgacttgg cgtagcgtgg acgcgcggcg 540
gcggccgcgcg tctcgtcgcg gtcggctttg ccgtgtatcg acgctgcggg cttgacacgg 600
gatggcgga gttcagcatc gctgcgatcc ctgcgcgcgc agaacgagga gagcgaggc 660
cggcttcaag tttgaaagga gaggaaggca ggcaaggagc tggaagcttg ccgcggaagg 720
cgaggcatg cgtcacgtga aaaaaaggga tttcaagagt agtaagtagg tatggtctac 780
aagtccccta ttcttacttc gcggaacgtg ggctgctcgt gcgggcgtcc atcttgtttt 840
tgtttttttt tccgctaggc gcgtgcattg cttgatgagt ctgagcgttc gtctgcagcg 900
agggcaggaa aataagcggc ccgtgccgtc gagcgcacag gacgtgcaag gcgcttgcga 960
gcgcagcatc cttgcacggc gagcatagag accgcggccg atggactcca gcgaggaatt 1020
ttcgacctc tctatcaagc tgcgcttgac agccgggaat ggcagcctga ggagagaggg 1080
gcgaaggaag ggacttggag aaaagaggta aggcaccctc aatcacggcg cgtgaaagcc 1140
agtcacccct cgcaaagaaa agacaaaagc gggttttttg tttcgatggg aaagaatttc 1200
ttagaggaag aagcggcaca cagactcgcg ccatgcagat ttctgcgcag ctgcgatca 1260
aaccaggaac gtggtcgtg cgcgccacta tcaggggtag cgacgaata ccaaacgcat 1320
tactagctac gcgcctgtga cccgaggatc gggccacaga cgttgtctct tgccatccca 1380
cgacctggca gcgagaagat cgtccattac tcatcg 1416

<210> 47
<211> 24
<212> DNA
<213> 人造序列

<220>		
<223>	合成引子 5' 60S-807	
<400>	47	
	tcgatttgcg gatacttgct caca	24
<210>	48	
<211>	21	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 3' 60S-2821	
<400>	48	
	gacgacctcg cccttggaca c	21
<210>	49	
<211>	34	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 5' 60Sp-1302-Kpn	
<400>	49	
	gactggtacc tttttccgct ctgcataatc ctaa	34
<210>	50	
<211>	32	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 3' 60Sp-Bam	
<400>	50	
	gactggatcc ttggcttttt ctttcttggt gc	32
<210>	51	
<211>	24	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 5' EF1-68	
<400>	51	
	cgccgttgac cgccgcttga ctct	24
<210>	52	
<211>	24	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 3' EF1-2312	
<400>	52	
	cgggggtagc ctcggggatg gact	24
<210>	53	
<211>	34	
<212>	DNA	
<213>	人造序列	
<220>		
<223>	合成引子 5' EF1-54-Kpn	

<400> 53	
gactggtacc tcttatctgc ctcgcgccgt tgac	34
<210> 54	
<211> 37	
<212> DNA	
<213> 人造序列	
<220>	
<223> 合成引子 5' EF1-1114-Bam	
<400> 54	
gactggatcc cttgcttgct agtagtcgct ttogaac	37
<210> 55	
<211> 29	
<212> DNA	
<213> 人造序列	
<220>	
<223> 合成引子 5' Sec1P-kpn	
<400> 55	
gactggtacc ccgtccttga cgccttcgc	29
<210> 56	
<211> 31	
<212> DNA	
<213> 人造序列	
<220>	
<223> 合成引子 3' Sec1P-ba	
<400> 56	
gactggatcc gatgagtaat ggacgatctt c	31
<210> 57	
<211> 1614	
<212> DNA	
<213> 人造序列	
<220>	
<223> 分泌信號	
<400> 57	
ggatccatga agttcgcgac ctcggtcgca attttgcttg tggccaacat agccaccgcc	60
ctcgcgtcga tgaccaacga gacctcggac cgccctctcg tgcaactttac cccaacaag	120
ggttgatga acgatcccaa cggcctctgg tacgacgaga aggatgctaa gtggcacctt	180
tactttcagt acaaccctaa cgacaccgtc tggggcacc cgtcttctg gggccacgcc	240
acctccgacg acctcaccaa ctgggaggac cagccattg ctatcgccc caagcgcaac	300
gactcgggag ctttttccgg ttccatggtt gtggactaca acaacacctc cggttttttt	360
aacgacacca ttgacccccg ccagcgctgc gtcgccatct ggacctaca cacgcccgag	420
agcgaggagc agtacatcag ctacagcctt gatggaggct acacctttac cgagtaccag	480
aagaaccctg tcctcgccgc caactccacc cagttccgcg accctaaggt tttttgtac	540
gagccttccc agaagtggat tatgaccgcc gctaagtcgc aggattaca gatcgagatc	600
tacagcagcg acgacctcaa gtcctggaag cttgagtccg cttttgcaa cgagggtttt	660
ctcgatacc agtacgagt ccccggtctc atcgagggtcc ccaccgagca ggaccgtcc	720
aagtcctact gggatcatgt tattttccatc aacctggcg cccctgccgg cggcagcttc	780



aaccagtact tcgtcggctc ctttaacggc acgcattttg aggccttcga caaccagtcc 840
cgcgctcgtcg acttcggcaa ggactactac gccctccaga ctttctttaa caccgacccc 900
acctacggca gcgccctcgg tattgcttgg gcctccaact gggagtactc cgctttcgtc 960
cccactaacc cctggcgcag ctcgatgtcc ctcgccgca agttttcgtc taacaccgag 1020
taccaggcca accccgagac cgagcttatt aacctgaagg cggagcctat tctcaacatc 1080
tccaacgctg gcccttggtc ccgctttgct actaacacta ccctcaccaa ggccaactcc 1140
tacaacgtcg atctctccaa ctccaccggt actcttgagt ttgagctcgt ctacgccgtc 1200
aacaccaccc agaccatctc caagtccgtc ttcgccgacc tctccctctg gttcaagggc 1260
cttgaggacc ccgaggagta cctgocgatg ggttttgagg tctccgcctc ctctttcttc 1320
ctcgatcgcg gtaactccaa ggtaagttt gtcaaggaga acccctactt tactaaccgt 1380
atgagcgta acaaccagcc ctttaagtcc gagaacgatc ttagctacta caaggtttac 1440
ggcctcctcg accagaacat tctcgagctc tactttaacg acggagatgt cgtcagcacc 1500
aacacctact ttatgaccac tggaaacgcc ctcggcagcg tgaacatgac caccggagtc 1560
gacaacctct ttacattga caagtttcag gttcgcgagg ttaagtaaca tatg 1614

<210> 58
<211> 11495
<212> DNA
<213> 人造序列

<220>
<223> pCL0076

<400> 58
ctcttatctg cctcgcgccg ttgaccgccg cttgactott ggcgcttgcc gctcgcattcc 60
tgccctcgctc gcgcaggcgg gcggggcgagt ggggtgggtcc gcagccttcc gcgctcgccc 120
gctagctcgc tcgcgccgtg ctgcagccag cagggcagca ccgcacggca ggcagggtccc 180
ggcgcggtat gatcgatcca tcgatccatc gatccatcga tcgtgcggtc aaaaagaaag 240
gaagaagaaa ggaaaaagaa aggcgtgcgc acccgagtgc gcgctgagcg cccgctcgcg 300
gtcccgcgga gcctccgct tagtccccgc cccgcgcgc gcagtcccc gggaggcatc 360
gcgcacctct cgccgcccc tcgocctcg ccgattcccc gcctcccctt ttccgcttct 420
tcgccgcctc cgctcgcggc cgogtcgccc gcgccccgct ccctatctgc tcccagggg 480
ggcactccgc accttttgcg cccgctgccg ccgccgggc cgccccgcg ccctggtttc 540
ccccgcgagc gcggccgctg cgccgcgcaa agactcgccg cgtgccgccc cgagcaacgg 600
gtggcgggcg cgcgccggcg ggccggggcg ggccggcggt aggcgggggt aggcgccggc 660
taggcgaaac gccgcccccg ggcccgccg ccgcccgctc cagagcagtc gccgcgccag 720
accgccaacg cagagaccga gaccgaggta cgtcgcgcc gagcacgcc cgacgcgcgg 780
cagggacgag gagcacgac ccgcgcgcgc ccgcgcgggg ggggggaggg agaggcagga 840
cgcgggagcg agcgtgcatg tttccgcgcg agacgacgc gcgcgcgctg gagaggagat 900
aaggcgcttg gatcgcgaga gggccagcca ggctggaggc gaaaatgggt ggagaggata 960
gtatcttgcg tgcttgagc aggagactga cgaggaggac ggatacgtcg atgatgatgt 1020
gcacagagaa gaagcagttc gaaagcgact actagcaagc aagggatcca tgaagttcgc 1080
gacctcggtc gcaattttgc ttgtggccaa catagccacc gccctcgct cgatgaccaa 1140

cgagacctog gaccgccctc togtgcactt tacccccaac aagggttga tgaacgatcc	1200
caacggcctc tggtacgacg agaaggatgc taagtggcac ctttactttc agtacaaccc	1260
taacgacacc gtctggggca ccccgctctt ctggggccac gccacctccg acgacctcac	1320
caactgggag gaccagccca ttgctatcgc cccaagcgc aacgactcgg gagctttttc	1380
cggttccatg gttgtggact acaacaacac ctccggtttt tttaacgaca ccattgaccc	1440
ccgccagcgc tgcgtcgcca tctggaccta caacacgccc gagagcgagg agcagtacat	1500
cagctacagc cttgatggag gctacacctt taccgagtac cagaagaacc ctgtcctcgc	1560
cgccaactcc acccagttcc gcgaccctaa ggttttttgg tacgagcctt cccagaagtg	1620
gattatgacc gccgctaagt cgcaggatta caagatcgag atctacagca gcgacgacct	1680
caagtcctgg aagcttgagt ccgcctttgc caacgagggt tttctcggat accagtacga	1740
gtgccccggg ctcatcgagg tccccaccga gcaggaccgc tccaagtcct actgggtcat	1800
gtttatttcc atcaaccctg gcgcccctgc cggcggcagc ttcaaccagt acttcgtcgg	1860
ctcctttaac ggcacgcatt ttgaggcctt cgacaaccag tcccgcgtcg tcgacttcgg	1920
caaggactac tacgccctcc agaccttctt taacaccgac cccacctacg gcagcgcctt	1980
cggatttgct tgggcctcca actgggagta ctccgctttc gtccccacta acccctggcg	2040
cagctcgatg tccctcgtcc gcaagttttc gcttaacacc gagtaccagg ccaaccccg	2100
gaccgagctt attaacctga aggccgagcc tattctcaac atctccaacg ctggcccctg	2160
gtcccgcctt gctactaaca ctaccctcac caaggccaac tcctacaacg tcgatctctc	2220
caactccacc ggtactcttg agtttgagct cgtctacgcc gtcaacacca cccagaccat	2280
ctccaagtcc gtcttcgccg acctctccct ctggttcaag ggccttgagg accccgagga	2340
gtacctgcgc atgggttttg aggtctccgc ctccctcctt ttccctcgatc gcggtaactc	2400
caaggttaag tttgtcaagg agaaccctta ctttactaac cgtatgagcg tcaacaacca	2460
gccctttaag tccgagaacg atcttagcta ctacaaggtt tacggcctcc tcgaccagaa	2520
cattctcgag ctctacttta acgacggaga tgtcgtcagc accaacacct actttatgac	2580
cactggaaac gccctcggca gcgtgaacat gaccaccgga gtcgacaacc tcttttacat	2640
tgacaagttt caggttcgcg aggttaagta acatatgtta tgagagatcc gaaagtgaac	2700
cttgcctaa cccgacagcg aatggcggga gggggcgggc taaaagatcg tattacatag	2760
tatttttccc ctactctttg tgtttgtctt tttttttttt ttgaacgcat tcaagccact	2820
tgtctgggtt tacttgtttg tttgcttgct tgcttgcttg cttgcctgct tcttggtcag	2880
acggcccaaa aaaggga aaaattcattca tggcacagat aagaaaaaga aaaagtttgt	2940
cgaccaccgt catcagaaag caagagaaga gaaacactcg cgctcacatt ctcgctcgcg	3000
taagaatctt agccacgcat acgaagtaat ttgtccatct ggccaatctt tacatgagcg	3060
tttcaagct ggagcgtgag atcatacctt tcttgatcgt aatgttccaa ccttgcatag	3120
gcctcgttgc gatccgctag caatgcgtcg tactcccgtt gcaactgcgc catcgctca	3180
ttgtgacgtg agttcagatt cttctcgaga ccttcgagcg ctgctaattt cgcctgacgc	3240
tccttctttt gtgcttccat gacacgccg ttcaccgtgc gttccacttc ttcctcagac	3300
atgcccttgg ctgcctcgac ctgctcggta agcttcgtcg taatctcctc gatctcggaa	3360



ttcttcttgc cctccatcca ctccggcacca tacttggcag cctgttcaac acgctcattg	3420
aaaaactttt catttctctt cagctccgca acccgcgctc gaagctcatt cacttccgcc	3480
accacggctt cggcatcgag cgccgaatca gtcgccgaac tttccgaaag atacaccacg	3540
gccctccgc tgctgctgcg cagcgtcatc atcagtcgog tgttatcttc gcgcagattc	3600
tccacctgct ccgtaagcag cttcacggtg gcctcttgat tctgagggct cacgtcgttg	3660
attagcgctt gcagctcttg cagctccgtc agcttggag agctcgtaat catggctttg	3720
cacttgcca gacgtcgag agcgttcgag agccgcttcg cgttatctgc catggacgct	3780
tctgcgctcg cggcctccct gacgacagtc tcttgcagtt tcactagatc atgtccaatc	3840
agcttgcggt gcagctctcc aatcacgttc tgcatcttgt ttgtgtgtcc gggccgcgcc	3900
tcgtcttgcg atttgcaat ttccctctcg agctcgcgtt cgagctccag ggcgcttta	3960
agtagctcga agtcagccgc cgtagcccc agctccgtcg ccgcgttcag acagtcggtt	4020
agcttgattc gattccgctt ttccatggca agttaagat cctggcccag ctgcacctcc	4080
tgcgcttgc gcatcatgcg cggttccgcc tggcgcaaaa gcttcgagtc gtatcctgcc	4140
tgccatgcca gcgcaatggc acgcacgagc gacttgagtt gccaaactatt catcgccgag	4200
atgagcagca ttttgatctg catgaacacc tcgtcagagt cgtcatcctc tgccctctcc	4260
agctctgcgg gcgagcgacg ctctccttgc agatgaagcg agggccgcag gcctccgaag	4320
agcacctctt gcgcgagatc ctctccgtc gtcgcccctc gcaggattgc ggtcgtgtcc	4380
gccatcttgc cgcacagca gcttttgctc gctctgcacc ttcaatttct ggtgccgctg	4440
gtgccgctgg tgccgcttgt gctggtgctg gtgctggtgc tgggtgctggt gccttgtgct	4500
ggtgctgcca cagacaccgc cgctcctgct gctgctcttc cggccccctc gccgccgccg	4560
cgagcccccg ccgcgcgccg tgccctgggt ctccgcgctc tccgcgggct cctcggcctc	4620
ggcctcgccg tccgcgacga cgtctgcgcg gccgatgggt cggtatctgct ctagagggcc	4680
cttcgaaggt aagcctatcc ctaaccctct cctcgggtct gattctacgc gtaccggtca	4740
tcatcaccat caccattgag tttaaacggg cccagcacg tgctacgaga tttcgattcc	4800
accgccgcct tctatgaaag gttgggcttc ggaatcgttt tccgggacgc cggctggatg	4860
atcctccagc gcggggatct catgctggag ttcttcgccc accccaactt gtttattgca	4920
gcttataatg gttacaaata aagcaatagc atcacaaatt tcacaaataa agcatttttt	4980
tactgcatt ctagttgtgg tttgtccaaa ctcatcaatg tatcttatca tacatggtcg	5040
acctgcagga acctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt	5100
gggcgtctt ccgcttcctc gtcactgac tcgctgcgct cggctcgttc gctgcggcga	5160
gcggtatcag ctactcaaa ggccgtaata cggttatcca cagaatcagg ggataacgca	5220
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg	5280
ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt	5340
cagagggtggc gaaacccgac aggactataa agataccagg cgtttcccc tggaagctcc	5400
ctcgtgcgct ctctgtttcc gacctgccg cttaccggat acctgtccgc ctttctccct	5460
tcgggaagcg tggcgctttc tcatagctca cgtgttaggt atctcagttc ggtgtaggtc	5520
gttcgctcca agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta	5580
tccggttaact atcgtcttga gtccaaccgc gtaagacacg acttatcgcc actggcagca	5640

gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 5700
 tgggtggccta actacggcta cactagaaga acagtatttg gtatctgcgc tctgctgaag 5760
 ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaaaaaac caccgctggt 5820
 agcgggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 5880
 gatcctttga tcttttctac ggggtctgac gctcagtggg acgaaaactc acgttaaggg 5940
 attttgggtca tgagattatc aaaaaggatc ttcacctaga tcctttttaa ttaaaaatga 6000
 agttttaaat caatctaaag tatatatgag taaacttggg ctgacagtta ccaatgctta 6060
 atcagtggag cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 6120
 cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 6180
 ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 6240
 agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 6300
 tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 6360
 gctacaggca tcgtgggtgc acgctcgtcg tttgggtatgg cttcattcag ctccggttcc 6420
 caacgatcaa ggcgagttac atgatcccc atgttgtgca aaaaagcggg tagctccttc 6480
 ggtcctccga tcgttgcag aagtaagttg gccgcagtgt tatcactcat ggttatggca 6540
 gcactgcata attctcttac tgtcatgccg tccgtaagat gcttttctgt gactgggtgag 6600
 tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg 6660
 tcaatacggg ataataccgc gccacatagc agaactttaa aagtgtcat cattggaaaa 6720
 cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa 6780
 cccactcgtg caccgaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga 6840
 gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga 6900
 atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg 6960
 agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt 7020
 ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa 7080
 aataggcgta tcacgaggcc ctttcgtctc gcgcgtttcg gtgatgacgg tgaaaacctc 7140
 tgacacatgc agctcccgga gacggtcaca gcttgtctgt aagcggatgc cgggagcaga 7200
 caagcccgtc agggcgcgctc agcgggtggt ggcggggtgtc ggggctggct taactatgcg 7260
 gcatcagagc agattgtact gagagtgcac caagctttgc ctcaacgcaa ctaggccag 7320
 gcctactttc actgtgtctt gtcttgctt tcacaccgac cgagtgtgca caaccgtgtt 7380
 ttgcacaaag cgcaagatgc tactcgact gtgaagcaaa ggttgcgcgc aagcgactgc 7440
 gactgcgagg atgaggatga ctggcagcct gttcaaaaac tgaaaatccg cgatgggtca 7500
 gctgccattc gcgcatgaoc cctgcgagag acaagttaac tcgtgtcact ggcatgtcct 7560
 agcatcttta cgcgagcaaa attcaatcgc tttatTTTTT cagtttctga accttctgc 7620
 aaccgcgaat cgccgtttca gcctgactaa tctgcagctg cgtggcactg tcagtcagtc 7680
 agtcagtctg gcgcgctgtt ccagcaccga ggtcgcgcgt cgccgcgcct ggaccgctgc 7740
 tgctactgct agtggcacgg caggtaggag cttgttgccg gaacaccagc agccgccagt 7800
 cgacgccagc caggggaaaag tccggcgtcg aaggagagg aaggcggcgt gtgcaaaacta 7860

acgttgacca	ctggcgcccc	ccgacacgag	caggaagcag	gcagctgcag	agcgacgcgc	7920
gcaagtgcag	aatgcgcgaa	agatccactt	gcgcgcggcg	ggcgcgcaact	tgcgggcgcg	7980
gcgcggaaca	gtgcggaaag	gagcgggtgca	gacggcgcg	agtgacagt	ggcgcaaagc	8040
cgcgcagtaa	gcagcggcgg	ggaacggtat	acgcagtgcc	gcgggccgcc	gcacacagaa	8100
gtatacgcgg	gccgaagtgg	ggcgctcgcg	gcgggaagt	cggaatggcg	ggcaaggaaa	8160
ggaggagacg	gaaagagggc	gggaaagaga	gagagagaga	gtgaaaaaag	aaagaaagaa	8220
agaaagaaa	aaagaaagct	cggagccacg	ccgcggggag	agagagaaat	gaaagcacgg	8280
cacggcaaa	caaagcaaag	cagaccacgc	cagaccacgc	cgaggagga	gcgcgcgcag	8340
gaccgcgcg	gcgagcgagc	gagcacggcg	cgcgagcgag	cgagcgagcg	agcgcgcgag	8400
cgagcaaggc	ttgctgcgag	cgatcgagcg	agcgagcggg	aaggatgagc	gcgaccgcg	8460
cggcgacgag	gacagcggcg	gcgctgtcct	cggcgctgac	gacgcctgta	aagcagcagc	8520
agcagcagca	gctgcgcgta	ggcgcgcggt	cggcacggct	ggcggccgcg	gcgttctcgt	8580
ccggcacggg	cggagacgcg	gccaaagaag	cggccgcggc	gagggcgttc	tccacgggac	8640
gcggcccaa	cgcgacacgc	gagaagagct	cgctggccac	ggtccaggcg	gcgacggacg	8700
atgcgcgctt	cgtcggcctg	accggcgccc	aaatctttca	tgagctcatg	cgcgagcacc	8760
aggtggacac	catctttggc	taccctggcg	gcgccattct	gcccgttttt	gatgccattt	8820
ttgagagtga	cgcgcttcaa	gttcattctc	gctcgccacg	agcagggcgc	cggccacatg	8880
gccgagggct	acgcgcgcgc	cacgggcaag	cccggcgttg	tcctcgtcac	ctcggggcct	8940
ggagccacca	acaccatcac	cccgatcatg	gatgcttaca	tggacggtac	gccgctgctc	9000
gtgttcaccg	gccagggtgca	gacctctgct	gtcggcacgg	acgctttcca	ggagtgtgac	9060
attgttggca	tcagccgcgc	gtgcaccaag	tggaaagtca	tggtaaggga	cgtgaaggag	9120
ctcccgcgcc	gcatcaatga	ggcctttgag	attgccatga	gcggccgccc	gggtcccgtg	9180
ctcgtcgatc	ttcctaagga	tgtgaccgcc	gttgagctca	aggaaatgcc	cgacagctcc	9240
ccccagggtg	ctgtgcgcc	gaagcaaaag	gtcgagcttt	tccacaagga	gcgcattggc	9300
gctcctggca	cggccgactt	caagctcatt	gccgagatga	tcaaccgtgc	ggagcgaccc	9360
gtcatctatg	ctggccaggg	tgtcatgcag	agcccggtga	atggcccggc	tgtgctcaag	9420
gagttcgcgg	agaaggccaa	cattcccgtg	accaccacca	tgagggtct	cggcggcctt	9480
gacgagcgta	gtcccctctc	cctcaagatg	ctcggcatgc	acggctctgc	ctacccaac	9540
tactcgatgc	agaacgccga	tcttatcctg	gcgctcgggtg	cccgccttga	tgatcgtgtg	9600
acgggccgcg	ttgacgcctt	tgctccggag	gctcgccgtg	ccgagcgcga	gggccgcggt	9660
ggcatcggtc	actttgagat	ttcccccaag	aacctccaca	aggtcgtcca	gcccaccgtc	9720
gcggtcctcg	gcgacgtggt	cgagaacctc	gccaacgtca	cgcacacgt	gcagcgccag	9780
gagcgcgagc	cgtggtttgc	gcagatcgcc	gattggaagg	agaagcacc	ttttctgctc	9840
gagtcgtgtg	attcgagacga	caaggttctc	aagccgcagc	aggtcctcac	ggagcttaac	9900
aagcagattc	tcgagattca	ggagaaggac	gccgaccagg	aggtctacat	caccacgggc	9960
gtcggaaagg	accagatgca	ggcagcgag	ttccttacct	ggaccaaggc	gcgccagtgg	10020
atctcctcgg	gtggcgccgg	cactatgggc	tacggccttc	cctcgcccat	tggcgccaag	10080
attgccaagc	ccgatgctat	tgttattgac	atcgatggtg	atgcttctta	ttcgatgacc	10140

ggtatggaat tgatcacagc agccgaattc aaggttggcg tgaagattct tcttttgcag 10200
aacaactttc agggcatggt caagaacgtt caggatctct tttacgacaa gcgctactcg 10260
ggccaccgcc atgttcaacc cgcgcttcga caaggtcgcc gatgcgatgc gtgccaaggg 10320
tctctactgc gcgaaacagt cggagctcaa ggacaagatc aaggagtttc tcgagtacga 10380
tgagggtccc gtctctctcg aggttttctg ggacaaggac acgctcgtct tgcccatggt 10440
ccccgctggc tttccgctcc acgagatggt cctcgagcct cctaagccca aggacgccta 10500
agttcttttt tccatggcgg gcgagcgagc gagcgcgcga gcgcgcaagt gcgcaagcgc 10560
cttgccctgc tttgcttcgc ttgcttttgc tttgcttcac acaacctaag tatgaattca 10620
agttttcttg ctgtctggcg atgcctgcct gccaaaccagc cagccatccg gccggccgtc 10680
cttgacgcct tcgcttccgg cgcggccatc gattcaattc acccatccga tacgttccgc 10740
cccctcacgt ccgtctgcgc acgaccctg cagcaccacg ccaaggccaa cgcgccgctc 10800
agctcagctt gtcgacgagt cgcacgtcac atatctcaga tgcatttgga ctgtgagtgt 10860
tattatgcca ctagcacgca acgatcttcg gggtcctcgc tcattgcatc cgttcggggc 10920
ctgcaggcgt ggacgcgagt cgcgcgcgag acgctgcagc aggccgctcc gacgcgaggg 10980
ctcgagctcg ccgcgcccgc gcgatgtctg cctggcgccg actgatctct ggagcgcaag 11040
gaagacacgg cgacgcgagg aggaccgaag agagacgctg gggtatgcag gatatacccg 11100
gggcggggaca ttggttccgc atacactccc ccattcgagc ttgctcgtcc ttggcagagc 11160
cgagcgcgaa cggttccgaa cgcggcaagg attttgctc tgggtgggtg actccgatcg 11220
aggcgcaggt tctccgcagg ttctcgagc cgggcagtgg tcgttagaaa tagggagtgc 11280
cggagtcttg acgcgcctta gctcactctc cgccacgcg cgcacgcgcg ccattgccgc 11340
gtcccgtctg tcgctgcgct ggccgcgacc ggctgcgcca gactacgaca gtgggacaga 11400
gctcgaggcg acgcgaatcg ctggggttgt aagggtttca agggtcgggc gtcgtcgcgt 11460
gccaaagtga aaatagtagg gggggggggg ggtac 11495

<210> 59
<211> 32
<212> PRT
<213> 裂殖壺菌屬

<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> 裂殖壺菌屬

<400> 60

atgcgcacgg tgagggggcc gcaaacggcg gcactcgccg cccttctggc acttgccgcg 60
acgcacgtgg ctgtgagccc gttcaccaag gtggag 96

<210> 61

<211> 30
<212> PRT
<213> 裂殖壺菌屬

<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> 裂殖壺菌屬

<400> 62
atgggccgcc tcgcgaagtc gcttgtgctg ctgacggccg tgctggccgt gatcggaggc 60
gtccgcgccg aagaggacaa gtccgaggcc 90

<210> 63
<211> 38
<212> PRT
<213> 裂殖壺菌屬

<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> 裂殖壺菌屬

<400> 64
atgacgtcaa cggcgcgcgc gctcgcgctc gtgcgtgctt tgggtgctcgc tctggctgtc 60
ttggcgctgc tagcgagcca aagcgtggcc gtggaccgca aaaagttcag gacc 114

<210> 65
<211> 34
<212> PRT
<213> 裂殖壺菌屬

<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> 裂殖壺菌屬

<400> 66
atgttgcggc tcaagccact ttctactcctc ttctctctgct cgttgattgc ttgcgcctgtg 60
gttgccctggg caagaggagg agaagggccg tccacgagcg aa 102

<210> 67
<211> 29
<212> PRT
<213> 裂殖壺菌屬

<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

<210> 68
<211> 87
<212> DNA
<213> 裂殖壺菌屬

<400> 68
atggccaaga tcttgcgagc ttgtctcctg gcggccgtgc tcgtggtgac tcctcaatca 60
ctgcgtgctc attcgacgcg ggacgca 87

<210> 69
<211> 36
<212> PRT
<213> 裂殖壺菌屬

<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> 裂殖壺菌屬

<400> 70
atgggtgttc ggcgcggtgcc atggcacggc gcggcgacgc tggcggcctt ggtcgtggcc 60
tgcgcgacgt gtttaggcct gggactggac tcggaggagg ccacgtac 108

<210> 71
<211> 30
<212> PRT
<213> 裂殖壺菌屬

<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> 裂殖壺菌屬

<400> 72
atgacagcta actcgggtgaa aataagcatc gtggctgtgc tggtcgcggc actggcttgg 60
gaaacatgcg caaaagctaa ctatcagtgg 90

<210> 73
<211> 35
<212> PRT
<213> 裂殖壺菌屬

<400> 73
Met Ala Arg Arg Ala Ser Arg Leu Gly Ala Ala Val Val Val Val Leu
1 5 10 15

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> 裂殖壺菌屬

<400> 74
atggcgcgca gggcgtcgcg cctcggcgcc gccgtcgtcg tcgtcctcgt cgtcgtcggc 60
tccgcctgct gctggcaagc cgctgcggac gtcgtggacg cgacg 105

<210> 75
<211> 1785
<212> DNA
<213> 人造

<220>
<223> 密碼子已最佳化之核酸序列

<400> 75
atgaagtctg cgacctcggc cgcaattttg cttgtggcca acatagccac cgccctcgcg 60
gcctccccct cgatgcagac ccgtgcctcc gtcgtcattg attacaacgt cgctcctcct 120
aacctctcca ccctcccgaa cggcagcctc tttgagacct ggcgtcctcg cgcccacgtt 180
cttcccccta acggtcagat tggcgatccc tgcctccact acaccgatcc ctcgactggc 240
ctctttcaag tcggctttct ccaagatggc tccggcattt cctccgccac tactgacgac 300
ctcgttacct acaaggatct caaccagggc aaccagggtca tcgtccccgg cggtatcaac 360
gaccctgtcg ctgttttcga cggctccgtc attccttcg gcattaacgg cctccctacc 420
ctcctctaca cctccgtcag ctacctcccc attcaactggc ccatccccta caccgcgggt 480
tccgagacgc agagcctggc tgtctccagc gatgggtggc ccaactttac taagctcgac 540
cagggccccg ttattcctgg ccccccttt gcctacaacg tcaccgcctt ccgcgacccc 600

tacgtctttc agaacccccac cctcgactcc ctctccact ccaagaacaa cacctggtac	660
accgtcattt cgggtggcct ccacggcaag ggccccgcc agtttcttta ccgtcagtac	720
gaccccgact ttcagtactg ggagttcctc ggccagtggg ggcaagagcc taccaactcc	780
acctggggca acggcacctg ggccggccgc tgggccttca acttcgagac cggcaacgtc	840
tttctgcttg acgagtacgg ctacaacccc cacggccaga tcttctccac cattggcacc	900
gagggctccg accagcccgt tgtcccccag ctacactcca tccacgatat gctttgggtc	960
tccggtaacg tttcgcgcaa cggatcggtt tccttcactc ccaacatggc cggcttcctc	1020
gactgggggt tctcgtccta cgccgccgc ggtaagggtt ttccttccac gtcgctcccc	1080
tccaccaagt ccggtgcccc cgatcgcttc atttcgtacg tttggctctc cggcgacctc	1140
tttgagcagg ctgagggtt tcctaccaac cagcagaact ggaccggcac cctcctcctc	1200
ccccgtgagc tccgctcct ttacatcccc aacgtggttg ataacgcct tgcgcgcgag	1260
tccggcgctt cctggcaggt cgtctcctcc gatagctcg ccggtactgt ggagctccag	1320
accctcggca tttccatcgc ccgcgagacc aaggccgcc tcctgtccgg cacctcgttc	1380
actgagtccg accgcactct taactcctcc ggcgctggtt cctttaagcg tccccctcc	1440
gagaagtttt tcgtcctctc cggccagctc tccttccccg cctccgcccg cggctcgggc	1500
ctcaagtccg gcttcagat tctttcctcc gagctcgagt ccaccacggt ctactaccag	1560
tttagcaacg agtccatcat cgtcgaccgc agcaacacca gcgccgccgc ccgtactacc	1620
gacggtatcg actcctccgc cgaggccggc aagctccgcc tctttgacgt cctcaacggc	1680
ggcgagcagg ctattgagac cctcgacctt accctcgtcg ttgataactc cgtgctcgag	1740
atttacgcca acggtcggtt cgcgctttcc acctgggttc gctaa	1785

發明專利說明書

(本說明書格式、順序，請勿任意更動，※記號部分請勿填寫)

※申請案號：99107452

※申請日：99.3.15

※IPC 分類：

C12N 15/87 (2006.01)

C12P 21/00 (2006.01)

一、發明名稱：(中文/英文)

於網黏菌門微生物中生產蛋白質之技術

PROTEIN PRODUCTION IN MICROORGANISMS OF THE PHYLUM
LABYRINTHULOMYCOTA

二、中文發明摘要：

本發明係關於網黏菌門(Labyrinthulomycota)之重組細胞及微生物及其用於異源蛋白質製造之用途。本發明也涵蓋自重組宿主細胞及微生物有效製造及選擇性地分泌多肽用的新穎啟動子、終結子、及信號序列。

三、英文發明摘要：

The present invention relates to recombinant cells and microorganisms of the phylum Labyrinthulomycota and their use in heterologous protein production. Novel promoter, terminator, and signal sequences for efficient production and, optionally, secretion of polypeptides from recombinant host cells and microorganisms are also encompassed by the present invention.

MANIMANVTPQGVAKGFLFVGVLFFLY
WFLVGLA (SEQ ID NO:1)

第1圖

atggccaaca tcatggccaa cgtcacgccc cagggcgctg ccaagggctt tggcctcttt 60
gtcggcgtgc tcttctttct ctactgggtc cttgtcggcc tcgcc 105

(SEQ ID NO:2)

第 2 圖

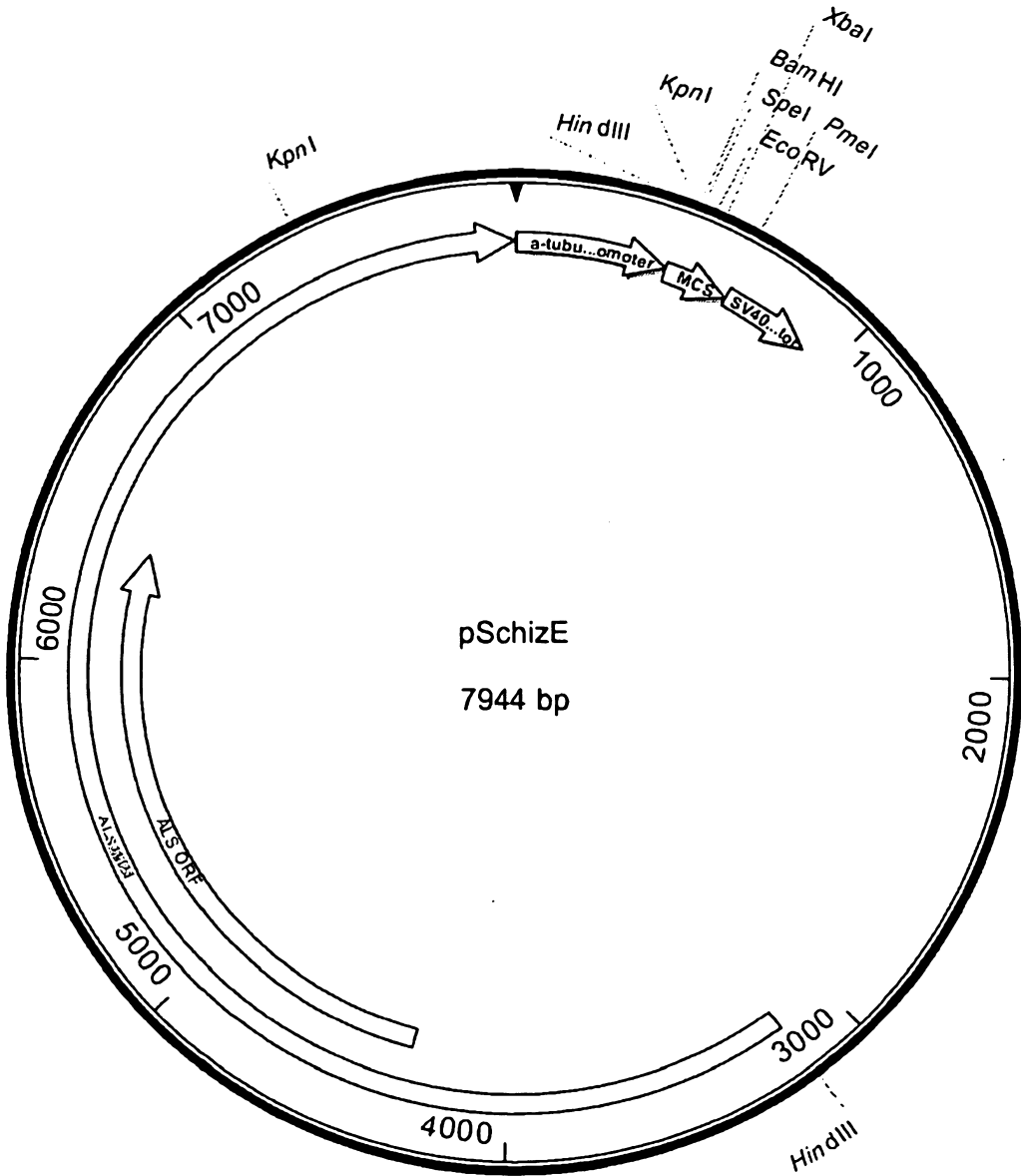
(SEQ ID NO:3)

gatecgaaag tgaaccttgt cctaaccga cagcgaatgg cgggaggggg cgggctaaaa	60
gategtatta catagtattt tteccctact ctttgtgttt gtcttttttt tttttgaa	120
cgcatccaag ccacttgtct tggtttactt gtttgtttgc ttgcttgctt gcttgcttgc	180
ctgcttcttg gtcagacggc ccaaaaaagg gaaaaaatc attcatggca cagataagaa	240
aaagaaaaag tttgtcgacc accgtcatca gaaagcaaga gaagagaaac actcgcgctc	300
acattctcgc tcgcgtaaga atctta	326

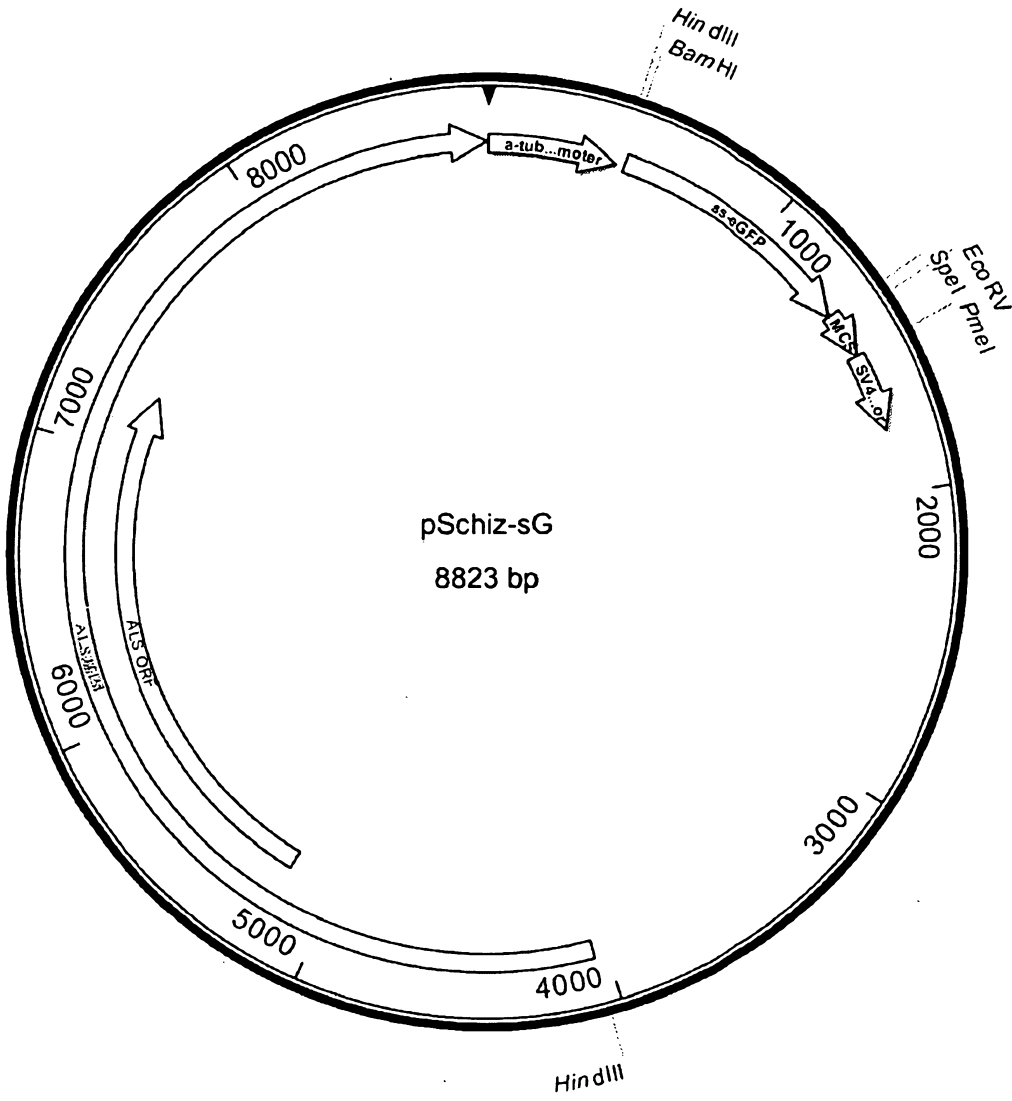
(SEQ ID NO:4)

第 4 圖

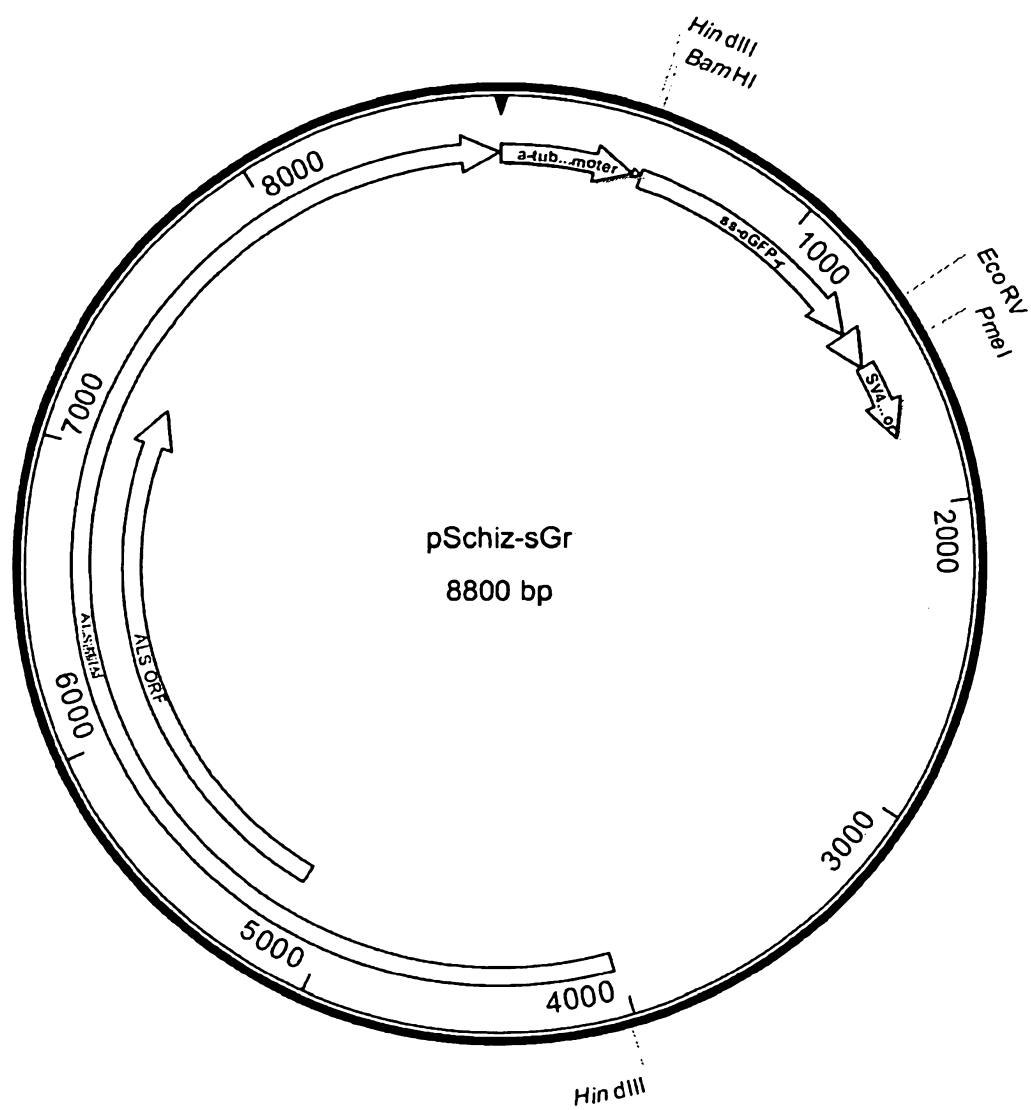




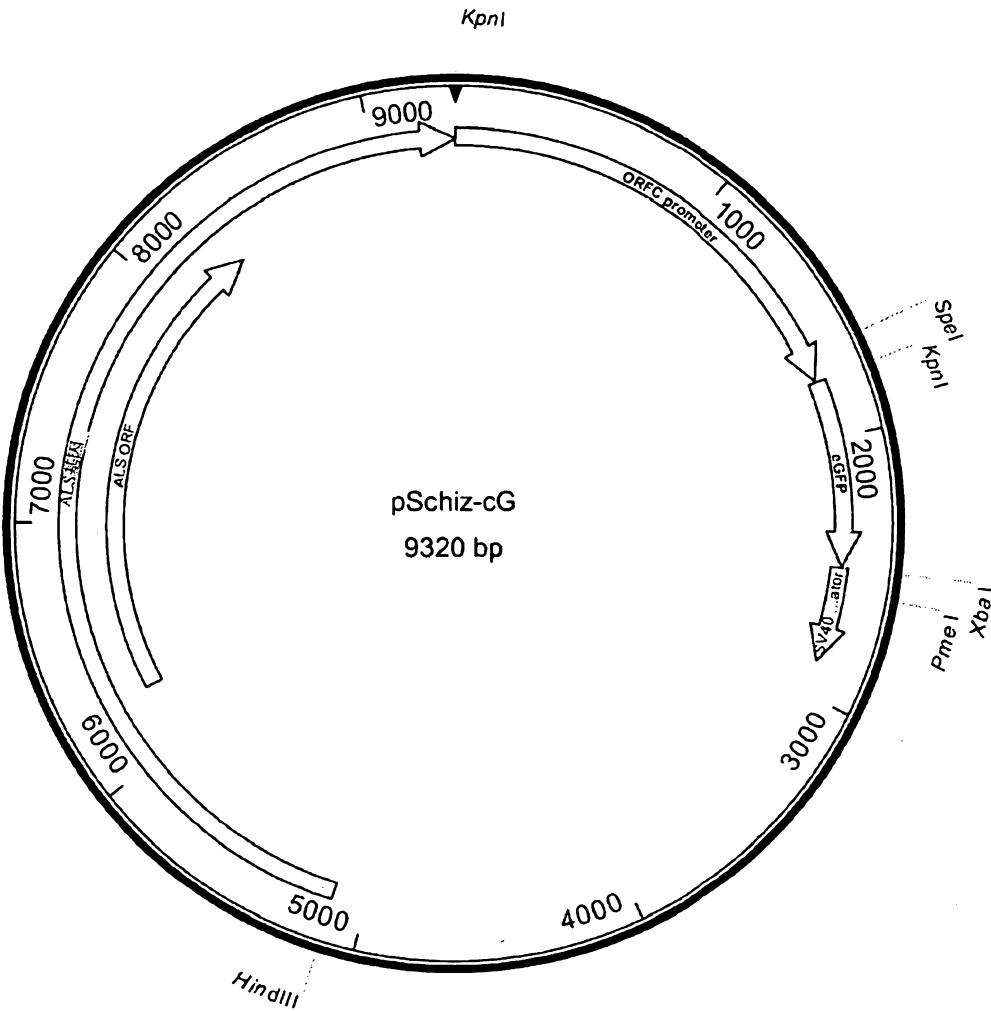
第 5 圖



第 6 圖



第 7 圖



第 8 圖

ctccatcgat cgtgcggtca aaaagaaagg aagaagaaag gaaaaagaaa ggcgtgcgca	60
cccgagtgcg cgctgagcgc ccgctcgcgg ccccgcgag cctccgcgtt agtccccgcc	120
ccgcgccgcg cagtcccccg ggaggcatcg cgcacctctc gccgccccct cgcgcctcgc	180
cgattccccg cctccccctt tccgcttctt cgccgcctcc gctcgcggcc gcgtcgcgcg	240
cgccccgctc cctatctgct cccaggggg gcactccgca ccttttgcg cgcgtgccgc	300
cgccgcggcc gccccgccgc cctggtttcc cccgcgagcg cggccgcgtc gccgcgcaaa	360
gactcgcgcg gtgccgcccc gagcaacggg tggcggcggc gcggcggcgg gcggggcgcg	420
gcggcgcgta ggcggggcta ggcgccggct aggcgaaacg ccgcccccg gcgccgccgc	480
cgccccgctc agagcagtcg ccgcgccaga ccgccaacgc agagaccgag accgaggtac	540
gtcgcgcccc agcacgccgc gacgcgcggc agggacgagg agcacgacgc cgcgccgcgc	600
cgcgcggggg gggggaggga gaggcaggac gcgggagcga gcgtgcatgt ttccgcgcga	660
gacgacgccg cgcgcgctgg agaggagata aggcgcttgg atcgcgagag ggccagccag	720
gctggaggcg aaaatgggtg gagaggatag tatcttgctt gcttggacga ggagactgac	780
gaggaggacg gatacgtcga tgatgatgtg cacagagaag aagcagttcg aaagcgacta	840
ctagcaagca agg	853

(SEQ ID NO:42)

第 19 圖

ctcttatctg cctcgcgcgcg ttgaccgccg cttgactctt ggcgcttgcc gctcgcaccc	60
tgccctcgctc gcgcaggcgg gcggggcgagt ggggtgggtcc gcagccttcc gcgctcgccc	120
gctagctcgc tcgcgccgtg ctgcagccag cagggcagca ccgcacggca ggcaggtccc	180
ggcgcggatc gatcgatcca tcgatccatc gatccatcga tcgtgcggtc aaaaagaaag	240
gaagaagaaa ggaaaaagaa aggcgtgcgc acccgagtgc gcgctgagcg cccgctcgcg	300
gtcccgcgga gcctccgcgt tagtccccgc cccgcgccgc gcagtcccc gggaggcatc	360
gcgcacctct cgccgcccc tcgcgcctcg ccgattcccc gcctcccctt ttccgcttct	420
tcgcgcctc cgctcgcggc cgcgtcgccc gcgccccgct ccctatctgc tccccagggg	480
ggcactccgc accttttgcg cccgctgccg ccgcgcgggc cgccccgcgc ccctggtttc	540
ccccgcgagc gcggccgcgt cgccgcgcaa agactcgccg cgtgcgcgcc cgagcaacgg	600
gtggcggcgg cgcgggcgcg ggcggggcg gcggcgcggt aggcgggggt aggcgcgggc	660
taggcgaaac gccgccccg ggcgcgcgcg ccgcccgtc cagagcagtc gccgcgccag	720
accgccaacg cagagaccga gaccgaggta cgtcgcgccc gagcacgccg cgacgcgcgg	780
cagggacgag gagcacgacg ccgcgcgcgc ccgcgcgggg ggggggaggg agaggcagga	840
cgcgggagcg agcgtgcatg ttccgcgcgc agacgacgcc gcgcgcgctg gagaggagat	900
aaggcgcttg gatcgcgaga gggccagcca ggctggaggc gaaaatgggt ggagaggata	960
gtatcttgcg tgcttgacg aggagactga cgaggaggac ggatacgtcg atgatgatgt	1020
gcacagagaa gaagcagttc gaaagcgact actagcaagc aagg	1064

(SEQ ID NO:43)

第 20 圖

cttcgctttc tcaacctatc tggacagcaa tccgccactt gccttgatcc ccttccgcgc 60
ctcaatcact cgctccacgt ccctcttccc cctcctcacc tccgtgcttt ctctgcccc 120
cccccccccg ccgcggcggtg cgcgcgcggtg gcgccgcggc cgcgacacct tccatactat 180
cctcgctccc aaaatggggtt gcgctatagg gcccggttag gcgaaagtct agcaggcact 240
tgcttggcgc agagccgccg cggccgctcg ttgccgcgga tggagaggga gagagagccc 300
gcctcgataa gcagagacag acagtgcgac tgacagacag acagagagac tggcagaccg 360
gaatacctcg aggtgagtgc ggcgcgggcg agcgggcggg agcgggagcg caagagggac 420
ggcgcgggcg ggcggccctg cgcgacgccg cggcgtattc tcgtgcgcag cgccgagcag 480
cgggacgggc ggctggctga tgggtgaagc ggggcggggt gaaatgttag atgagatgat 540
catcgacgac ggtccgtgcg tcttggctgg cttggctggc ttggctggcg ggctgccgt 600
gtttgcgaga aagaggatga ggagagcgac gaggaaggac gagaagactg acgtgtaggg 660
cgcgcgatgg atgatcgatt gattgattga ttgattggtt gattggctgt gtggtcgatg 720
aacgtgtaga ctcagggagc gtggttaaata tgttcttgcg ccagacgcga ggactccacc 780
cccttctttc gcctttacac agcctttttg tgaagcaaca agaaagaaaa agccaag 837

(SEQ ID NO:44)

第 21 圖

ctttttccgc	tctgcataat	cctaaaagaa	agactatacc	ctagtcactg	tacaaatggg	60
acattttctct	cccgagcgat	agctaaggat	ttttgcttcg	tgtgcactgt	gtgctctggc	120
cgcgcacgca	aagtccagga	tcttactgtt	tctctttcct	ttcctttatt	tcctgttctc	180
ttcttcgctt	tctcaaccta	tctggacagc	aatccgccac	ttgccttgat	ccccttcgcg	240
gcctcaatca	ctcgctccac	gtccctcttc	cccctcctca	tctcogtgct	ttctctcgcc	300
ccccccccc	ccgccgggc	gtgcgcgcgc	gtggcgccgc	ggccgcgaca	ccttcatac	360
tatcctcgct	cccaaatgg	gttgcgctat	agggcccggc	taggcgaaag	tctagcaggc	420
acttgcttgg	cgcagagccg	ccgcggccgc	tcgttgccgc	ggatggagag	ggagagagag	480
cccgctcga	taagcagaga	cagacagtgc	gactgacaga	cagacagaga	gactggcaga	540
ccggaatacc	tcgaggtgag	tgcggcgcg	gcgagcgggc	gggagcggga	gcgcaagagg	600
gacggcgcg	cgcggcgcc	ctgcgcgacg	ccgcggcgta	ttctcgtgcg	cagcgccgag	660
cagcgggacg	ggcggctggc	tgatgggtga	agcggggcgg	ggtgaaatgt	tagatgagat	720
gatcatcgac	gacggtccgt	gcgtcttggc	tggcttggct	ggcttggctg	gcgggcctgc	780
cgtgtttgcg	agaaagagga	tgaggagagc	gacgaggaag	gacgagaaga	ctgacgtgta	840
gggcgcgcga	tggatgatcg	attgattgat	tgattgattg	gttgattggc	tgtgtggtcg	900
atgaacgtgt	agactcaggg	agcgtggtta	aattgttctt	gcgccagacg	cgaggactcc	960
accccttct	ttcgcttta	cacagccttt	ttgtgaagca	acaagaaaga	aaaagccaag	1020

(SEQ ID NO:45)

第 22 圖

CCCGTCCTTGACGCCTTCGCTTCCGGCGCGGCCATCGATTCAATTCACCCATCCGATACGTTCC
GCCCCCTCACGTCCGTCTGCGCACGACCCCTGCACGACCACGCCAAGGCCAACGCGCCGCTCAG
CTCAGCTTGTCGACGAGTCGCACGTACATATCTCAGATGCATTGCCTGCCTGCCTGCCTGCCT
GCCTGCCTGCCTGCCTGCCTGCCTGCCTCAGCCTCTCTTTGCTCTCTCTGCGGCGGCCGCTGCG
ACGCGCTGTACAGGAGAATGACTCCAGGAAGTGCGGCTGGGATACGCGCTGGCGTCGGCCGTGA
TGCGCGTGACGGGCGGCGGGCACGGCCGGCACGGGTTGAGCAGAGGACGAAGCGAGGCGAGACG
AGACAGGCCAGGCGCGGGGAGCGCTCGCTGCCGTGAGCAGCAGACCAGGGCGCAGGAATGTACT
TTTCTTGCGGGAGCGGAGACGAGGCTGCCGGCTGCTGGCTGCCGGTTGCTCTGCACGCGCCGCC
CGACTTGCGGTAGCGTGACGCGCGGGCGGGCGGCCGCGTCTCGTCGCGGTTCGGCTTTGCCGTGT
ATCGACGCTGCGGGCTTGACACGGGATGGCGGAAGTTCAGCATCGCTGCGATCCCTCGCGCCGC
AGAACGAGGAGAGCGCAGGCCGGCTTCAAGTTTGAAAGGAGAGGAAGGCAGGCAAGGAGCTGGA
AGCTTGCCGCGGAAGGCGCAGGCATGCGTCACGTGAAAAAAGGGATTTCAAGAGTAGTAAGTA
GGTATGGTCTACAAGTCCCCTATTCTTACTTCGCGGAACGTGGGCTGCTCGTGCGGGCGTCCAT
CTTGTTTTTGTTTTTTTTTTCCGCTAGGCGCGTGCAATTGCTTGATGAGTCTCAGCGTTCGTCTGC
AGCGAGGGCAGGAAAATAAGCGGCCCCGTGCCGTGAGCGCACAGGACGTGCAAGCGCCTTGCGA
GCGCAGCATCCTTGACGCGGCGAGCATAGAGACCGCGCCGATGGACTCCAGCGAGGAATTTTCG
ACCCTCTCTATCAAGCTGCGCTTGACAGCCGGGAATGGCAGCCTGAGGAGAGAGGGGCGAAGGA
AGGGACTTGGAGAAAAGAGGTAAGGCACCCTCAATCACGGCGCGTGAAAGCCAGTCATCCCTCG
CAAAGAAAAGACAAAAGCGGGTTTTTTTGTTCGATGGGAAAGAATTTCTTAGAGGAAGAAGCGG
CACACAGACTCGCGCCATGCAGATTTCTGCGCAGCTCGCGATCAAACCAGGAACGTGGTTCGCTG
CGCGCCACTATCAGGGGTAGCGCACGAATACCAAACGCATTACTAGCTACGCGCCTGTGACCCG
AGGATCGGGCCACAGACGTTGTCTCTTGCCATCCCACGACCTGGCAGCGAGAAGATCGTCCATT
ACTCATCG

(SEQ ID NO:46)

第 23 圖

提示之 結構式	質量 [M+Na]	Chage 態 [z]	76-3-11
Hex ₅ HexNAc ₂	1579	1	√
	802	2	
Hex ₆ HexNAc ₂	1783	1	√
	903	2	
Hex ₇ HexNAc ₂	1987	1	√
	1005	2	
Hex ₈ HexNAc ₂	1108	2	√
Hex ₉ HexNAc ₂	1210	2	√
Hex ₁₀ HexNAc ₂	1312	2	√
Hex ₁₁ HexNAc ₂	1407	2	√
Hex ₁₂ HexNAc ₂	1516	2	√
Hex ₁₃ HexNAc ₂	1618	2	√

√..... 已檢測
n.d..... 未檢測

第 40A 圖

裂殖壺菌屬預測之信號序列(基於信號P程式)

α -1,6-甘露糖基轉移酶(ALG12) :

MRTVRGPQTAALAALLALAATHVAVSPFTKVE (SEQ ID NO:59)

ATGCGCACGGTGAGGGGGCCGCAAACGGCGGCACTCGCCGCCCTTCTGGCACTTGCCGCGACGC
ACGTGGCTGTGAGCCCGTTCACCAAGGTGGAG (SEQ ID NO:60)

BiP:

MGRLAKSLVLLTAVLAVIGGVRAEEDKSEA (SEQ ID NO:61)

ATGGgCCGCCTCGCGAAGTCGCTTGTGCTGCTGACGGCCGTGCTGGCCGTGATCGGAGGCGTCCG
CGCCGAAGAGGACAAGTCCGAGGCC (SEQ ID NO:62)

α -1,3-葡萄糖苷(GLS2) :

MTSTARALALVRALVLALAVLALLASQSVAVDRKKFRT (SEQ ID NO:63)

ATGACGTCAACGGCGCGCGCGCTCGCGCTCGTGCGTGCTTTGGTGCTCGCTCTGGCTGTCTTGGC
GCTGCTAGCGAGCCAAAGCGTGGCCGTGGACCGAAAAAGTTCAGGACC (SEQ ID NO:64)

仿 α -1,3-1,6-甘露糖苷酶 :

MLRLKPLLLFLCSLIASPVVAWARGGEGPSTSE (SEQ ID NO:65)

ATGTTGCGGCTCAAGCCACTTTTACTCCTCTTCCTCTGCTCGTTGATTGCTTCGCCTGTGGTTGCC
TGGGCAAGAGGAGGAGAAGGGCCGTCCACGAGCGAA (SEQ ID NO:66)

第 41A 圖

仿 α -1,3-1,6-甘露糖苷酶#1：

MAKILRSLLLA AVL VVTPQSLRAHSTRDA (SEQ ID NO:67)

ATGGCCAAGATCTTGC GCAGTTTGCTCCTGGCGGCCGTGCTCGTGGTGACTCCTCAATCACTGCG
TGCTCATTCGACGCGGGACGCA (SEQ ID NO:68)

仿 α -1,2-甘露糖苷酶：

MVFRRVPWHGAATLAALVVACATCLGLGLDSEEATY (SEQ ID NO:69)

ATGGTGTTTCGGCGCGTGCCATGGCACGGCGCGGCGACGCTGGCGGCCTTGGTCGTGGCCTGCG
CGACGTGTTTAGGCCTGGGACTGGACTCGGAGGAGGCCACGTAC (SEQ ID NO:70)

仿 β 木糖苷酶蛋白質：

MTANSVKISIVAVLVAAALAWETCAKANYQW (SEQ ID NO:71)

ATGACAGCTAACTCGGTGAAAATAAGCATCGTGGCTGTGCTGGTCGCGGCACTGGCTTGGGAAA
CATGCGCAAAAGCTAACTATCAGTGG (SEQ ID NO:72)

胡蘿蔔素合成酶：

MARRASRLGAAVVVVLVVVASACCWQAAADVVD AQ (SEQ ID NO:73)

ATGGCGCGCAGGGCGTCGCGCCTCGGCGCCGCGTCGTGCTCGTCCTCGTCGTCGTCGCTCCGC
CTGCTGCTGGCAAGCCGCTGCGGACGTCGTGGACGCGCAG (SEQ ID NO:74)

第 41B 圖

胺基酸	密碼子	分量	胺基酸	密碼子	分量	胺基酸	密碼子	分量	胺基酸	密碼子	分量
Ala	GCC	0.64	End	TAA	0.34	Leu	CTT	0.16	Ser	TCG	0.33
Ala	GCA	0.03	End	TGA	0.33	Leu	TTG	0.02	Ser	TCC	0.31
Ala	GCT	0.18	End	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

第 42 圖

atgaagttcgcgacctcgggtcgcaattttgcttgtggccaacatagccaccgccctcgGCTCCCCCTCGATGCAGACCCGTGCCTCCGTC
GTCATTGATTACAACGTCGCTCCTCCTAACCTCTCCACCCTCCCGAACGGCAGCCTCTTTGAGACCTGGCGTCCTCGCGC
CCACGTTCTTCCCCCTAACGGTCAGATTGGCGATCCCTGCCTCCACTACACCGATCCCTCGACTGGCCTCTTTCACGTCG
GCTTTCTCCACGATGGCTCCGGCATTTCCTCCGCCACTACTGACGACCTCGCTACCTACAAGGATCTCAACCAGGGCAAC
CAGGTCATCGTCCCCGGCGGTATCAACGACCCTGTCGCTGTTTTCGACGGCTCCGTCATTCTTCCGGCATTAAACGGCCT
CCCTACCCTCCTCTACACCTCCGTACGCTACCTCCCATCACTGGTCCATCCCTACACCCGCGGTTCCGAGACGCAGA
GCCTGGCTGTCTCCAGCGATGGTGGCTCCAACTTTACTAAGCTCGACCAGGGCCCCGTTATTCTGGCCCCCCTTTGCC
TACAACGTCAACGCCTTCCGCGACCCCTACGTCTTTCAGAACCCACCCCTCGACTCCCTCCTCCACTCCAAGAACAAACAC
CTGGTACACCGTCATTTCCGGGTGGCCTCCACGGCAAGGGCCCCGCCAGTTTCTTTACCGTCAGTACGACCCGACTTT
CAGTACTGGGAGTTCCTCGGCCAGTGGTGGCAGAGCCTACCAACTCCACCTGGGGCAACGGCACCTGGGCCGGCCG
CTGGGCCTTCAATTTCGAGACCGGCAACGTCTTTTCGCTTGACGAGTACGGCTACAACCCCCACGGCCAGATCTTCTCC
ACCATTTGGCACCGAGGGCTCCGACCAGCCCGTTGTCCCCCAGCTCACCTCCATCCACGATATGCTTTGGGTCTCCGGTA
ACGTTTCGCGCAACGGATCGGTTTCCTTCACTCCCAACATGGCCGGCTTCCTCGACTGGGGTTTCTCGTCTACGCCGCC
GCGGGTAAGGTTCTTCTTCCACGTCGCTCCCCTCCACCAAGTCCGGTGCCCCCGATCGCTTCATTTCTGTACGTTTGGCT
CTCCGGCGACCTCTTTGAGCAGGCTGAGGGCTTTCCTACCAACCAGCAGAACTGGACCGGCACCCTCCTCCTCCCCCGT
GAGCTCCGCGTCCTTTACATCCCCAACGTGGTTGATAACGCCCTTGCGCGCGAGTCCGGCGCTTCTGGCAGGTCGTCT
CCTCCGATAGCTCGGCCGGTACTGTGGAGCTCCAGACCCTCGGCATTTCCATCGCCCCGCGAGACCAAGGCCGCCCTCCT
GTCCGGCACCTCGTTCACTGAGTCCGACCGCACTCTTAACTCCTCCGGCGTCGTTCCCTTTAAGCGTTCCCCCTCCGAGA
AGTTTTTCGTCTCTCCGCCAGCTCTCCTTCCCCGCCTCCGCCGCGGCTCGGGCCTCAAGTCCGGCTTCCAGATTCTTT
CCTCCGAGCTCGAGTCCACCACGGTCTACTACCAGTTTAGCAACGAGTCCATCATCGTCGACCGCAGCAACACCAGCGC
CGCCGCCCGTACTACCGACGGTATCGACTCCTCCGCCGAGGCCGGCAAGCTCCGCCTCTTTGACGTCTCAACGGCGGC
GAGCAGGCTATTGAGACCCTCGACCTTACCCTCGTCGTTGATAACTCCGTGCTCGAGATTTACGCCAACGGTCGTTTCG
CGCTTTCACCTGGGTTTCGCTAA

(SEQ ID NO.: 75)

第 44 圖



四、指定代表圖：

(一)本案指定代表圖為：第 (3) 圖。

(二)本代表圖之元件符號簡單說明：

(無)

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

(無)

公告本

七、申請專利範圍：

1. 一種已單離之重組核酸分子，包含：
 - (a) 與SEQ ID NO:38之多核苷酸序列具有至少90%序列相同度之多核苷酸序列，其中該多核苷酸序列編碼有作為信號肽功能之多肽；
 - (b) 編碼多肽之多核苷酸序列，其中該多肽包含與SEQ ID NO:37之胺基酸序列具有至少90%序列相同度之胺基酸序列或其片段，其中該胺基酸序列或其片段有作為信號肽之功能；或
 - (c) 與(a)或(b)之多核苷酸序列完全互補之多核苷酸序列。
2. 如申請專利範圍第1項之重組核酸分子，其中該多核苷酸序列與SEQ ID NO:38之多核苷酸序列具有至少95%序列相同度。
3. 如申請專利範圍第1項之重組核酸分子，其中該多核苷酸序列係為SEQ ID NO:38。
4. 如申請專利範圍第1項之重組核酸分子，其中該多核苷酸序列編碼包含與SEQ ID NO:37之胺基酸序列具有至少95%序列相同度之胺基酸序列的多肽。
5. 如申請專利範圍第1項之重組核酸分子，其中該多核苷酸序列編碼包含SEQ ID NO:37之胺基酸序列的多肽。
6. 如申請專利範圍第1項之重組核酸分子，其中該多核苷酸序列編碼包含SEQ ID NO:37之首18個胺基酸殘基的多肽。

7. 如申請專利範圍第1項之重組核酸分子，其中該多核苷酸序列編碼包含SEQ ID NO:37之首19個胺基酸殘基的多肽。
8. 如申請專利範圍第1項之重組核酸分子，其中該重組核酸分子為載體。
9. 如申請專利範圍第1項之重組核酸分子，其中該多核苷酸序列係可操作地聯結至編碼蛋白質之多核苷酸序列。
10. 一種宿主細胞，包含如申請專利範圍第1項之重組核酸分子。
11. 如申請專利範圍第10項之宿主細胞，其中該宿主細胞為破囊壺菌目(Thraustochytriales)之一成員。
12. 如申請專利範圍第11項之宿主細胞，其中該宿主細胞為裂殖壺菌屬(Schizochytrium)或破囊壺菌屬(Thraustochytrium)。
13. 一種用於製造如申請專利範圍第9項之重組核酸分子所編碼之蛋白質之方法，包含：
 - a) 於培養基中培養破囊壺菌目之一重組微生物，其中該重組微生物包含如申請專利範圍第9項之已單離之核酸分子；及
 - b) 單離於步驟a)中產生的該蛋白質。
14. 如申請專利範圍第13項之方法，其中該蛋白質係自己單離之破囊壺菌目生質回收。
15. 如申請專利範圍第13項之方法，其中該蛋白質係蓄積於微生物體內。

16. 如申請專利範圍第13項之方法，其中該蛋白質係蓄積於微生物之膜。
17. 如申請專利範圍第13項之方法，其中該蛋白質係自該培養基中回收。
18. 如申請專利範圍第13項之方法，其中該蛋白質為經分泌之蛋白質。