



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẢNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0049265

(51)⁷ A01H 5/00

(13) B

(21) 1-2015-00134

(22) 14/06/2013

(86) PCT/AU2013/000639 14/06/2013

(87) WO 2013/185184 19/12/2013

(30) 61/660,392 15/06/2012 US; 61/663,344 22/06/2012 US; 61/697,676 06/09/2012 US;
61/782,680 14/03/2013 US

(45) 25/07/2025 448

(43) 27/07/2015 328A

(73) 1. COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH
ORGANISATION (AU)

Clunies Ross St, Acton, Australian Capital Territory 2601, Australia

2. GRAINS RESEARCH AND DEVELOPMENT CORPORATION (AU)

Level 4, 4 National Circuit, Barton, Australian Capital Territory 2600, Australia

3. Nuseed Nutritional Australia Pty Ltd (AU)

103-109 Pipe Road, Laverton North, Victoria 3026, Australia

(72) PETRIE, James Robertson (AU); SINGH, Surinder Pal (AU); DE FEYTER, Robert
Charles (AU).

(74) Công ty TNHH T&T INVENMARK Sở hữu trí tuệ Quốc tế (T&T INVENMARK
CO., LTD.)

(54) LIPIT CHIẾT TÁCH TỪ CÂY CHỨA CÁC AXIT BÉO Ở DẠNG ĐƯỢC ESTE
HOÁ VÀ QUY TRÌNH SẢN XUẤT LIPIT NÀY

(21) 1-2015-00134

(57) Sáng chế đề xuất phương pháp sản xuất axit béo nhiều liên kết chưa bão hòa mạch dài, cụ thể là axit docosahexaenoic, trong tế bào thực vật tái tổ hợp.

Lĩnh vực sáng chế đề cập

Sáng chế đề cập đến phương pháp sản xuất axit béo nhiều liên kết chưa bão hòa mạch dài, cụ thể là axit docosahexaenoic, trong tế bào thực vật tái tổ hợp.

Tình trạng kỹ thuật của sáng chế

Axit béo omega-3 nhiều liên kết chưa bão hòa mạch dài (LC-PUFA) hiện nay đã được biết một cách rộng rãi là các hợp chất quan trọng cho sức khỏe con người và động vật. Các axit béo này có thể có được từ các nguồn thức ăn hoặc bằng cách chuyển hóa axit béo linoleic (LA, 18:2 ω 6) hoặc α -linolenic (ALA, 18:3 ω 3), cả hai axit này đều được coi là các axit béo thiết yếu trong chế độ ăn của người. Trong khi người và các loại động vật có xương sống khác có thể chuyển hóa LA hoặc ALA, thu được từ các nguồn gốc thực vật thành C22, chúng tiến hành sự chuyển hóa này với tốc độ rất chậm. Hơn nữa, phần lớn trong xã hội hiện đại, chế độ ăn uống là mất cân bằng trong đó ít nhất 90% axit béo nhiều liên kết chưa bão hòa (PUFA) là thuộc nhóm axit béo ω 6, thay vì tỷ lệ axit béo ω 6: ω 3 được coi là lý tưởng là 4:1 hoặc thấp hơn (Trautwein, 2001). Nguồn từ thức ăn trực tiếp của LC-PUFA như axit eicosapentaenoic (EPA, 20:5 ω 3) và axit docosahexaenoic (DHA, 22:6 ω 3) cho người phần lớn là từ cá hoặc dầu cá. Do đó, các chuyên gia về sức khỏe đã khuyến cáo việc thường xuyên đưa cá có mức LC-PUFA cao vào chế độ thức ăn cho người. Ví dụ ngày càng nhiều các loại dầu LC-PUFA có nguồn gốc từ cá được đưa vào các sản phẩm thức ăn và trong các chế phẩm công thức cho trẻ. Tuy nhiên, do sự sụt giảm của ngư trường trên toàn cầu và cả ở mức quốc gia, cần có các nguồn khác cho các loại dầu tăng cường sức khỏe này.

Trái ngược với động vật, các cây ra hoa không có khả năng tổng hợp axit béo nhiều liên kết chưa bão hòa có chiều dài mạch dài hơn 18 carbon. Cụ thể là, hoa màu và các cây vườn với các cây hạt kín khác không có các enzym cần thiết để tổng hợp các axit béo ω 3 mạch dài như EPA, axit docosapentaenoic (DPA, 22:5 ω 3) và DHA có nguồn gốc từ ALA. Do đó, mục đích quan trọng trong ngành công nghệ sinh học thực vật là tạo ra các cây hoa màu sinh ra các lượng đủ LC-PUFA, do đó là nguồn thay thế các hợp chất này.

Con đường sinh tổng hợp LC-PUFA

Con đường sinh tổng hợp LC-PUFA ở các sinh vật như vi tảo, rêu và nấm thường xảy ra là một loạt các phản ứng kéo dài và khử bão hòa phụ thuộc vào oxy (Fig.1). Con đường thông thường nhất sinh ra EPA ở các sinh vật này bao gồm phản ứng khử bão hòa- $\Delta 6$, phản ứng kéo dài- $\Delta 6$ và phản ứng khử bão hòa- $\Delta 5$ (gọi là con đường khử bão hòa- $\Delta 6$) trong khi con đường ít thông dụng hơn sử dụng phản ứng khử bão hòa- $\Delta 9$, phản ứng khử bão hòa- $\Delta 8$ và phản ứng khử bão hòa- $\Delta 5$ (có tên gọi là con đường khử bão hòa vị trí $\Delta 9$). Các phản ứng khử bão hòa và kéo dài liên tiếp có thể bắt đầu hoặc với cơ chất axit béo $\omega 6$ LA, được thể hiện dạng biểu đồ là phần trên bên trái trong Fig.1 ($\omega 6$) hoặc với cơ chất ALA $\omega 3$ cho đến EPA, thể hiện là phần dưới bên phải trong Fig.1 ($\omega 3$). Nếu phản ứng khử bão hòa – vị trí $\Delta 6$ đầu tiên được thực hiện trên cơ chất $\omega 6$ LA, sản phẩm LC-PUFA của một loạt ba enzym này sẽ là axit béo $\omega 6$ ARA. Các sinh vật tổng hợp LC-PUFA có thể chuyển hóa axit $\omega 6$ thành axit béo $\omega 3$ bằng $\omega 3$ -desaturaza, thể hiện là bước $\Delta 17$ -desaturaza trong Fig.1 để chuyển hóa axit arachidonic (ARA, 20:4 $\omega 6$) thành EPA. Một số enzym trong họ $\omega 3$ -desaturaza có thể tác dụng trên nhiều cơ chất từ LA đến ARA. Các $\omega 3$ -desaturaza thực vật thường xúc tác đặc hiệu phản ứng khử bão hòa- $\Delta 15$ khử LA thành ALA, trong khi $\omega 3$ -desaturaza của nấm và nấm men có thể xúc tác đặc hiệu cho phản ứng khử bão hòa- $\Delta 17$ khử ARA thành EPA (Pereira et al., 2004a; Zank et al., 2005). Một số báo cáo gợi ý rằng các $\omega 3$ -desaturaza không đặc hiệu có thể tồn tại, chúng có thể chuyển hóa một loại rất nhiều cơ chất $\omega 6$ thành các sản phẩm $\omega 3$ tương ứng của chúng (Zhang et al., 2008).

Sự chuyển hóa EPA thành DHA ở các sinh vật này xảy ra bởi phản ứng kéo dài- $\Delta 5$ EPA để tạo ra DPA, sau đó phản ứng khử bão hòa- $\Delta 4$ để tạo ra DHA (Fig.1). Ngược lại, động vật có vú sử dụng con đường "Sprecher" chuyển hóa DPA thành DHA bởi ba phản ứng riêng biệt độc lập với $\Delta 4$ -desaturaza (Sprecher et al., 1995).

Các tiền-desaturaza thường được tìm thấy ở thực vật, rêu, vi tảo, và các động vật bậc thấp như *Caenorhabditis elegans* phần lớn chấp nhận cơ chất axit béo được este hóa ở vị trí *sn*-2 của cơ chất phosphatidylcholin (PC). Do đó, các desaturaza được biết là các tiền-desaturaza axyl-PC, được liên kết lipid, (Domergue et al., 2003). Ngược lại, tiền-desaturaza của động vật cao cấp hơn thường chấp nhận cơ chất axyl-CoA, trong đó cơ chất axit béo được liên kết với CoA mà không phải là PC (Domergue et al., 2005). Một số desaturaza của vi tảo và một desaturaza thực vật đã được biết là sử dụng cơ chất

axit béo được este hóa với CoA (Bảng 2).

Mỗi phản ứng kéo dài PUFA gồm có bốn bước được xúc tác bởi phức protein đa thành phần: đầu tiên, phản ứng ngưng tụ gây kết hợp đơn vị 2C từ malonyl-CoA vào axit béo, gây hình thành chất trung gian β -ketoaxyl. Sau đó nó được khử bởi NADPH, sau đó dehydrat hóa để thu được chất trung gian enoyl. Chất trung gian này cuối cùng được khử lần thứ hai để tạo ra axit béo kéo dài. Người ta thường cho rằng bước ngưng tụ trong bốn phản ứng này là đặc hiệu về cơ chất trong khi các bước khác thì không. Trên thực hành, điều này có nghĩa là cơ chế kéo dài ở cây tự nhiên có khả năng kéo dài PUFA với điều kiện là enzym ngưng tụ (đặc trưng được gọi là 'elongaza') đặc hiệu với PUFA được đưa vào, mặc dù hiệu quả của cơ chế kéo dài trong cây tự nhiên để kéo dài các cơ chất PUFA không có trong tự nhiên có thể là thấp. Năm 2007, đã công bố việc xác định và đặc điểm của dehydrataza kéo dài chu kỳ của nấm men (Denic và Weissman, 2007).

Phản ứng khử bão hòa PUFA ở thực vật, rêu và vi tảo xảy ra trong tự nhiên với cơ chất axit béo chủ yếu là trong nhóm axyl-PC trong khi phản ứng kéo dài xảy ra với cơ chất trong nhóm axyl-CoA. Việc vận chuyển axit béo từ phân tử axyl-PC đến chất mang CoA được tiến hành bởi phospholipaza (PLA) trong khi việc vận chuyển axit béo axyl-CoA đến chất mang PC được tiến hành bởi lysophosphatidyl-cholin axyltransferaza (LPCAT) (Fig.21) (Singh et al., 2005).

Sản xuất LC-PUFA theo thiết kế

Hầu như việc thiết kế LC-PUFA chuyển hóa dc tiến hành bởi con đường khử bão hòa/kéo dài $\Delta 6$ ưa khí. Sự sinh tổng hợp axit γ -linolenic (GLA, 18:3n-6) ở cây thuốc lá lần đầu tiên được báo cáo vào năm 1996 sử dụng $\Delta 6$ -desaturaza từ vi khuẩn lam *Synechocystis* (Reddy và Thomas, 1996). Gần đây, GLA đã được tạo ra trong các cây hoa màu như cây rum (73% GLA trong dầu ép hạt; Knauf et al., 2006) và đậu nành (28% GLA; Sato et al., 2004). Việc sản xuất LC-PUFA như EPA và DHA liên quan đến việc thiết kế phức tạp hơn do gồm tăng số bước khử bão hòa và kéo dài hơn. Việc sản xuất EPA trong cây mặt đất lần đầu tiên được báo cáo bởi Qi et al. (2004) người đưa gen mã hóa $\Delta 9$ -elongaza từ *Isochrysis galbana*, $\Delta 8$ -desaturaza từ *Euglena gracilis* và $\Delta 5$ -desaturaza từ *Mortierella alpina* vào *Arabidopsis* thu được đến 3% EPA. Sau đó là Abbadi et al. (2004) báo cáo về việc tạo ra đến 0,8% EPA trong hạt cây lanh sử dụng gen mã hóa $\Delta 6$ -desaturaza và $\Delta 6$ -elongaza từ *Physcomitrella patens* và $\Delta 5$ -desaturaza

từ *Phaeodactylum tricornutum*.

Báo cáo đầu tiên về việc sản xuất DHA, và cho đến nay mức tạo ra VLC-PUFA cao nhất là, là trong WO 04/017467 trong đó mô tả việc tạo ra 3% DHA ở mầm đậu nành, mà không phải là hạt, bằng cách đưa gen mã hóa $\Delta 6$ -desaturaza *Saprolegnia diclina*, $\Delta 6$ -desaturaza *Mortierella alpina*, $\Delta 5$ -desaturaza *Mortierella alpina*, $\Delta 4$ -desaturaza *Saprolegnia diclina*, $\Delta 17$ -desaturaza *Saprolegnia diclina*, $\Delta 6$ -elongaza *Mortierella alpina* và $\Delta 5$ -elongaza *Pavlova lutheri*. Mức EPA tối đa trong mầm cũng sinh ra mức DHA là 19,6%, thể hiện rằng hiệu quả chuyển hóa EPA thành DHA là kém (WO 2004/071467). Phát hiện này tương tự với những điều đã được công bố bởi Robert et al. (2005), trong đó lượng từ EPA thành DHA là thấp, với hiệu suất 3% EPA và 0,5% DHA ở *Arabidopsis* sử dụng $\Delta 5/6$ -desaturaza *Danio rerio*, $\Delta 6$ -elongaza *Caenorhabditis elegans*, và $\Delta 5$ -elongaza và $\Delta 4$ -desaturaza *Pavlova salina*. Cũng trong năm 2005, Wu et al. công bố việc tạo ra 25% ARA, 15% EPA, và 1,5% DHA ở *Brassica juncea* sử dụng $\Delta 6$ -desaturaza *Pythium irregulare*, $\Delta 5$ -desaturaza *Thraustochytrid*, $\Delta 6$ -elongaza *Physcomitrella patens*, của $\Delta 12$ - desaturaza *Calendula officianalis*, $\Delta 5$ -elongaza *Thraustochytrid*, $\Delta 17$ -desaturaza *Phytophthora infestans*, LC-PUPA elongaza *Oncorhynchus mykiss*, $\Delta 4$ -desaturaza của *Thraustochytrid* và LPCAT của *Thraustochytrid* (Wu et al., 2005). Sơ lược về nỗ lực để tạo ra các hoa màu có hạt dầu tổng hợp $\omega 3$ LCPUFA được nêu trong Venegas- Caleron et al. (2010) và Ruiz-Lopez et al. (2012). Như được chỉ ra bởi Ruiz-Lopez et al. (2012), các kết quả thu được cho đến nay để tạo ra DHA ở thực vật biến đổi gen chưa đạt đến được mức như trong dầu cá.

Do đó vẫn tồn tại nhu cầu về quy trình sản xuất LC-PUFA hiệu quả hơn nữa ở các tế bào tái tổ hợp, đặc biệt là DHA trong hạt của các cây hạt dầu.

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất phương pháp và cây để tạo ra lipit có mức DHA cao.

Theo phương án đầu tiên, sáng chế đề xuất lipit chiết tách từ cây, chứa axit béo ở dạng este hóa, các axit béo bao gồm axit oleic, axit palmitic, axit béo $\omega 6$ bao gồm axit linoleic (LA), axit béo $\omega 3$ bao gồm axit alinolenic (ALA), và axit docosahexaenoic (DHA), và tùy chọn một hoặc nhiều axit stearidonic (SDA), axit eicosapentaenoic (EPA), axit docosapentaenoic (DPA) và axit eicosatetraenoic (ETA), trong đó mức DHA trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 7% đến

20%.

Theo một phương án, lipit chiết tách có một hoặc nhiều hoặc tất cả các đặc điểm sau

i) mức axit palmitic trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 2% đến 18%, nằm trong khoảng từ 2% đến 16%, hoặc nằm trong khoảng từ 2% đến 15%,

ii) mức axit myristic (C14:0) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 6%, thấp hơn khoảng 3%, thấp hơn khoảng 2%, hoặc thấp hơn khoảng 1%,

iii) mức axit oleic trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 1% đến 30%, nằm trong khoảng từ 3% đến 30%, nằm trong khoảng từ 6% đến 30%, nằm trong khoảng từ 1% đến 20%, nằm trong khoảng từ 30% đến 60%, nằm trong khoảng từ 45% đến 60%, hoặc là khoảng 30%,

iv) mức axit linoleic (LA) trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 4% đến 35%, nằm trong khoảng từ 4% đến 20%, hoặc nằm trong khoảng từ 4% đến 17%,

v) mức axit α -linolenic (ALA) trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 4% đến 40%, nằm trong khoảng từ 7% đến 40%, nằm trong khoảng từ 10% đến 35%, nằm trong khoảng từ 20% đến 35%, nằm trong khoảng từ 4% đến 16%, hoặc nằm trong khoảng từ 2% đến 16%,

vi) mức axit γ -linolenic (GLA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 4%, thấp hơn khoảng 3%, thấp hơn khoảng 2%, thấp hơn khoảng 1%, thấp hơn khoảng 0,5%, nằm trong khoảng từ 0,05% đến 7%, nằm trong khoảng từ 0,05% đến 4%, nằm trong khoảng từ 0,05% đến 3%, hoặc nằm trong khoảng từ 0,05% đến 2%,

vii) mức axit stearidonic (SDA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 7%, thấp hơn khoảng 6%, thấp hơn khoảng 4%, thấp hơn khoảng 3%, nằm trong khoảng từ 0,05% đến 7%, nằm trong khoảng từ 0,05% đến 6%, nằm trong khoảng từ 0,05% đến 4%, nằm trong khoảng từ 0,05% đến 3%, hoặc nằm trong khoảng từ 0,05% đến 2%,

viii) mức axit eicosatetraenoic (ETA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 6%, thấp hơn khoảng 5%, thấp hơn khoảng 4%, thấp hơn

khoảng 1%, thấp hơn khoảng 0,5%, nằm trong khoảng từ 0,05% đến 6%, nằm trong khoảng từ 0,05% đến 5%, nằm trong khoảng từ 0,05% đến 4%, nằm trong khoảng từ 0,05% đến 3%, hoặc nằm trong khoảng từ 0,05% đến 2%,

ix) mức axit eicosatrienoic (ETrA) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn khoảng 4%, thấp hơn khoảng 2%, thấp hơn khoảng 1%, nằm trong khoảng từ 0,05% đến 4%, nằm trong khoảng từ 0,05% đến 3%, nằm trong khoảng từ 0,05% đến 2%, hoặc nằm trong khoảng từ 0,05% đến 1%,

x) mức axit eicosapentaenoic (EPA) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn khoảng 4%, thấp hơn khoảng 3%, thấp hơn khoảng 2%, nằm trong khoảng từ 0,05% đến 10%, nằm trong khoảng từ 0,05% đến 5%, nằm trong khoảng từ 0,05% đến 3%, hoặc nằm trong khoảng từ 0,05% đến 2%,

xi) mức axit docosapentaenoic (DPA) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn khoảng 4%, thấp hơn khoảng 3%, thấp hơn khoảng 2%, nằm trong khoảng từ 0,05% đến 8%, nằm trong khoảng từ 0,05% đến 5%, nằm trong khoảng từ 0,05% đến 3%, hoặc nằm trong khoảng từ 0,05% đến 2%,

xii) mức DHA trong tổng hàm lượng axit béo của lipid chiết tách là khoảng 8%, khoảng 9%, khoảng 10%, khoảng 11%, khoảng 12%, khoảng 13%, khoảng 14%, khoảng 15%, khoảng 16%, khoảng 17%, khoảng 18%, nằm trong khoảng từ 8% đến 20%, nằm trong khoảng từ 10% và 20%, nằm trong khoảng từ 11% đến 20%, nằm trong khoảng từ 10% đến 16%, hoặc nằm trong khoảng từ 14% đến 20%,

xiii) lipid chứa axit ω 6-docosapentaenoic (22:5^{Δ4,7,10,13,16}) trong hàm lượng axit béo của nó,

xiv) lipid này về cơ bản không có axit ω 6-docosapentaenoic (22:5^{Δ4,7,10,13,16}) trong thành phần axit béo của nó,

xv) lipid này về cơ bản không có SDA, EPA và ETA trong thành phần axit béo của nó,

xvi) mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 25%, nằm trong khoảng từ 4% đến 20%, nằm trong khoảng từ 6% đến 20%, nằm trong khoảng từ 4% đến 60%, nằm trong khoảng từ 30% đến 60%, hoặc nằm trong khoảng từ 45% đến 60%,

xvii) mức toàn bộ axit béo một liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 35%, nằm trong khoảng từ 8% đến

25%, hoặc nằm trong khoảng từ 8% đến 22%,

xviii) mức toàn bộ axit béo nhiều liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 20% đến 75%, nằm trong khoảng từ 50% đến 75%, hoặc nằm trong khoảng từ 60% đến 75%,

xix) mức toàn bộ axit béo $\omega 6$ trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 35% đến 50%, nằm trong khoảng từ 20% đến 35%, nằm trong khoảng từ 6% đến 20%, thấp hơn khoảng 20%, thấp hơn khoảng 16%, thấp hơn khoảng 10%, nằm trong khoảng từ 1% đến 16%, nằm trong khoảng từ 2% đến 10%, hoặc nằm trong khoảng từ 4% đến 10%,

xx) mức axit béo $\omega 6$ mới trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 10%, thấp hơn khoảng 8%, thấp hơn khoảng 6%, thấp hơn 4%, nằm trong khoảng từ 1% đến 20%, nằm trong khoảng từ 1% đến 10%, nằm trong khoảng từ 0,5% đến 8%, hoặc nằm trong khoảng từ 0,5% đến 4%,

xxi) mức toàn bộ axit béo $\omega 3$ trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 36% đến 65%, nằm trong khoảng từ 40% đến 60%, nằm trong khoảng từ 20% đến 35%, nằm trong khoảng từ 10% đến 20%, đến 25%, đến 30%, đến 35% hoặc đến 40%,

xxii) mức axit béo $\omega 3$ mới trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 9% đến 33%, nằm trong khoảng từ 10% đến 20%, nằm trong khoảng từ 20% đến 30%, nằm trong khoảng từ 12% đến 25%, đến 13%, đến 15%, đến 17% hoặc đến 20%,

xxiii) tỷ lệ toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$ trong hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 1,0 đến 3,0, nằm trong khoảng từ 0,1 đến 1, nằm trong khoảng từ 0,1 đến 0,5, thấp hơn khoảng 0,50, thấp hơn khoảng 0,40, thấp hơn khoảng 0,30, thấp hơn khoảng 0,20, thấp hơn khoảng 0,15, khoảng 1,0, khoảng 0,1 hoặc khoảng 0,2,

xxiv) tỷ lệ của axit béo $\omega 6$ mới: axit béo $\omega 3$ mới trong hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 1,0 đến 3,0, nằm trong khoảng từ 0,1 đến 1, nằm trong khoảng từ 0,1 đến 0,5, thấp hơn khoảng 0,50, thấp hơn khoảng 0,40, thấp hơn khoảng 0,30, thấp hơn khoảng 0,20, thấp hơn khoảng 0,15, khoảng 0,1, khoảng 0,2 hoặc khoảng 1,0,

xxv) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa axit oleic thành

LA bởi $\Delta 12$ -desaturaza ít nhất là khoảng 60%, ít nhất khoảng 70%, ít nhất khoảng 80%, nằm trong khoảng từ 60% đến 98%, nằm trong khoảng từ 70% đến 95%, hoặc nằm trong khoảng từ 75% đến 90%,

xxvi) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa ALA thành SDA bởi $\Delta 6$ -desaturaza ít nhất là khoảng 30%, ít nhất khoảng 40%, ít nhất khoảng 50%, ít nhất khoảng 60%, ít nhất khoảng 70%, nằm trong khoảng từ 30% đến 70%, nằm trong khoảng từ 35% đến 60%, hoặc nằm trong khoảng từ 50% đến 70%,

xxvii) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa SDA thành axit ETA bởi $\Delta 6$ -elongaza ít nhất là khoảng 60%, ít nhất khoảng 70%, ít nhất khoảng 75%, nằm trong khoảng từ 60% đến 95%, nằm trong khoảng từ 70% đến 88%, hoặc nằm trong khoảng từ 75% đến 85%,

xxviii) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa ETA thành EPA bởi $\Delta 5$ -desaturaza ít nhất là khoảng 60%, ít nhất khoảng 70%, ít nhất khoảng 75%, nằm trong khoảng từ 60% đến 99%, nằm trong khoảng từ 70% đến 99%, hoặc nằm trong khoảng từ 75% đến 98%,

xxix) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa EPA thành DPA bởi $\Delta 5$ -elongaza ít nhất là khoảng 80%, ít nhất khoảng 85%, ít nhất khoảng 90%, nằm trong khoảng từ 50% đến 95%, hoặc nằm trong khoảng từ 85% đến 95%,

xxx) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa DPA thành DHA bởi $\Delta 4$ -desaturaza ít nhất là khoảng 80%, ít nhất khoảng 90%, ít nhất khoảng 93%, nằm trong khoảng từ 50% đến 95%, nằm trong khoảng từ 80% đến 95%, hoặc nằm trong khoảng từ 85% đến 95%,

xxxi) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa axit oleic thành DHA ít nhất là khoảng 10%, ít nhất khoảng 15%, ít nhất khoảng 20%, nằm trong khoảng từ 10% đến 50%, nằm trong khoảng từ 10% đến 30%, hoặc nằm trong khoảng từ 10% đến 25%,

xxxii) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa LA thành DHA ít nhất là khoảng 15%, ít nhất khoảng 20%, ít nhất khoảng 22%, ít nhất khoảng 25%, nằm trong khoảng từ 15% đến 50%, nằm trong khoảng từ 20% đến 40%, hoặc nằm trong khoảng từ 20% đến 30%,

xxxiii) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa ALA thành DHA ít nhất là khoảng 17%, ít nhất khoảng 22%, ít nhất khoảng 24%, nằm trong

khoảng từ 17% đến 55%, nằm trong khoảng từ 22% đến 35%, hoặc nằm trong khoảng từ 24% đến 35%,

xxxiv) toàn bộ axit béo trong lipid chiết tách có lượng C20:1 ít hơn 1%,

xxxv) hàm lượng triaxylglycerol (TAG) trong lipid ít nhất là khoảng 70%, ít nhất khoảng 80%, ít nhất khoảng 90%, ít nhất là 95%, nằm trong khoảng từ 70% đến 99%, hoặc nằm trong khoảng từ 90% đến 99%,

xxxvi) lipid chứa diaxylglycerol (DAG),

xxxvii) lipid chứa ít hơn khoảng 10%, ít hơn khoảng 5%, ít hơn khoảng 1%, hoặc nằm trong khoảng từ 0,001% đến 5%, là axit béo tự do (không este hóa) và/hoặc phospholipid, hoặc về cơ bản không có chúng,

xxxviii) ít nhất 70%, hoặc ít nhất 80%, DHA được este hóa ở dạng TAG là ở vị trí *sn*-1 hoặc *sn*-3 của TAG,

xxxix) kiểu TAG chứa DHA nhiều nhất trong lipid là DHA/18:3/18:3 (TAG 58:12), và

xl) lipid chứa tri-DHA TAG (TAG 66:18).

Theo phương án khác, lipid chiết tách là ở dạng dầu, trong đó ít nhất khoảng 90%, hoặc ít nhất khoảng 95%, ít nhất khoảng 98%, hoặc nằm trong khoảng từ 95% và khoảng 98%, theo trọng lượng của dầu là lipid.

Theo một phương án được ưu tiên, lipid hoặc dầu, tốt hơn là dầu hạt, có các đặc điểm sau: trong tổng hàm lượng axit béo của lipid hoặc dầu, mức DHA nằm trong khoảng từ 7% đến 20%, mức axit palmitic nằm trong khoảng từ 2% đến 16%, mức axit myristic là thấp hơn khoảng 6%, mức axit oleic nằm trong khoảng từ 1% đến 30%, mức LA nằm trong khoảng từ 4% đến 35%, có mặt ALA, có mặt GLA, mức SDA nằm trong khoảng từ 0,05% đến 7%, mức ETA là thấp hơn khoảng 4%, mức EPA nằm trong khoảng từ 0,05% đến 10%, mức DPA nằm trong khoảng từ 0,05% đến 8%, mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 25%, mức toàn bộ axit béo một liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 35%, mức toàn bộ axit béo nhiều liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 20% đến 75%, tỷ lệ của toàn bộ axit béo ω 6: toàn bộ axit béo ω 3 trong hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 0,05 đến 3,0, tỷ lệ của axit béo ω 6 mới: axit béo ω 3 mới trong hàm lượng axit béo của lipid chiết tách nằm trong

khoảng từ 0,03 đến 3,0, tốt hơn là thấp hơn khoảng 0,50, thành phần axit béo của lipid dựa trên hiệu suất chuyển hóa axit oleic thành LA bởi $\Delta 12$ -desaturaza ít nhất là khoảng 60%, hiệu suất chuyển hóa SDA thành axit ETA bởi $\Delta 6$ -elongaza ít nhất là khoảng 60%, hiệu suất chuyển hóa EPA thành DPA bởi $\Delta 5$ -elongaza nằm trong khoảng từ 50% đến 95%, hiệu suất chuyển hóa DPA thành DHA bởi $\Delta 4$ -desaturaza nằm trong khoảng từ 50% đến 95%, hiệu suất chuyển hóa axit oleic thành DHA ít nhất là khoảng 10%, và hàm lượng triaxylglyxerol (TAG) của lipid ít nhất là khoảng 70%, và tùy chọn lipid về cơ bản không có cholesterol và/hoặc lipid chứa tri-DHA TAG (TAG 66:18).

Theo một phương án được ưu tiên hơn, lipid hoặc dầu, tốt hơn là dầu hạt, có các đặc điểm sau: trong tổng hàm lượng axit béo của lipid, mức DHA nằm trong khoảng từ 7% đến 20%, mức axit palmitic nằm trong khoảng từ 2% đến 16%, mức axit myristic là thấp hơn khoảng 2%, mức axit oleic nằm trong khoảng từ 1% đến 30%, mức LA nằm trong khoảng từ 4% đến 35%, mức ALA nằm trong khoảng từ 7% đến 40%, mức GLA là thấp hơn khoảng 4%, mức SDA nằm trong khoảng từ 0,05% đến 7%, mức ETA là thấp hơn khoảng 4%, mức ETrA nằm trong khoảng từ 0,05% đến 4%, mức EPA nằm trong khoảng từ 0,05% đến 10%, mức DPA nằm trong khoảng từ 0,05% đến 8%, mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 25%, mức toàn bộ axit béo một liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 35%, mức toàn bộ axit béo nhiều liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 20% đến 75%, mức axit béo $\omega 6$ mới trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 0,5% đến 10%, mức toàn bộ axit béo $\omega 3$ trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 36% đến 75%, mức axit béo $\omega 3$ mới trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 9% đến 33%, tỷ lệ của toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$ trong hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 0,05 đến 3,0, tỷ lệ của axit béo $\omega 6$ mới: axit béo $\omega 3$ mới trong hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 0,03 đến 3,0, thành phần axit béo của lipid dựa trên hiệu suất chuyển hóa axit oleic thành LA bởi $\Delta 12$ -desaturaza ít nhất là khoảng 60%, hiệu suất chuyển hóa SDA thành axit ETA bởi $\Delta 6$ -elongaza ít nhất là khoảng 60%, hiệu suất chuyển hóa ETA thành EPA bởi $\Delta 5$ -desaturaza ít nhất là khoảng 60%, hiệu suất chuyển hóa EPA thành DPA bởi $\Delta 5$ -elongaza nằm trong khoảng từ 50% đến 95%, hiệu suất chuyển hóa DPA thành DHA

bởi $\Delta 4$ -desaturaza nằm trong khoảng từ 50% đến 95%, hiệu suất chuyển hóa axit oleic thành DHA ít nhất là khoảng 10%, hiệu suất chuyển hóa LA thành DHA ít nhất là khoảng 15%, hiệu suất chuyển hóa ALA thành DHA ít nhất là khoảng 17%, và tổng hàm lượng axit béo trong lipid chiết tách có C20:1 với lượng ít hơn 1%, hàm lượng triaxylglyxerol (TAG) của lipid ít nhất khoảng 70%, lipid về cơ bản không có cholesterol, và lipid chứa tri-DHA TAG (TAG 66:18). Tốt hơn là, lipid hoặc dầu là dầu cải và/hoặc không bị xử lý bằng quy trình chuyển hóa este sau khi nó được chiết từ cây hoặc bộ phận của cây. Theo một phương án cụ thể, lipid hoặc dầu cải về cơ bản có thể được xử lý để chuyển hóa các axit béo trong dầu thành các alkyl este như metyl hoặc etyl este. Các quy trình xử lý tiếp theo có thể được áp dụng để làm giàu DHA trong lipid hoặc dầu.

Theo một phương án, lipid hoặc dầu, tốt hơn là dầu hạt, có các đặc điểm sau: trong tổng hàm lượng axit béo của lipid, mức DHA nằm trong khoảng từ 7% đến 20%, mức axit palmitic nằm trong khoảng từ 2% đến 16%, mức axit myristic là thấp hơn khoảng 2%, mức axit oleic nằm trong khoảng từ 30% đến 60%, tốt hơn nằm trong khoảng từ 45% đến 60%, mức LA nằm trong khoảng từ 4% đến 20%, mức ALA nằm trong khoảng từ 2% đến 16%, mức GLA là thấp hơn khoảng 3%, mức SDA là thấp hơn khoảng 3%, mức ETA là thấp hơn khoảng 4%, mức ETrA thấp hơn khoảng 2%, mức EPA là thấp hơn khoảng 4%, mức DPA là thấp hơn khoảng 4%, mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 25%, mức toàn bộ axit béo một liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 30% đến 60%, hoặc nằm trong khoảng từ 40% đến 60%, mức toàn bộ axit béo nhiều liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 20% đến 75%, mức axit béo $\omega 6$ mới trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 0,5% đến 10%, mức toàn bộ axit béo $\omega 3$ trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 10% đến 20%, mức axit béo $\omega 3$ mới trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 9% đến 20%, tỷ lệ của toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$ trong hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 0,05 đến 3,0, tốt hơn là thấp hơn khoảng 0,50, tỷ lệ của axit béo $\omega 6$ mới: axit béo $\omega 3$ mới trong hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 0,03 đến 3,0, hàm lượng triaxylglyxerol (TAG) của lipid ít nhất là khoảng 70%, lipid về cơ bản không có

cholesterol, và lipit chứa tri-DHA TAG (TAG 66:18). Tốt hơn là, lipit hoặc dầu về cơ bản không có SDA, EPA và ETA và/hoặc là dầu cải và/hoặc không bị xử lý bằng quy trình chuyển hóa este sau khi nó được chiết từ cây hoặc bộ phận cây. Theo một phương án cụ thể, lipit hoặc dầu cải sau đó có thể được xử lý để chuyển hóa các axit béo trong dầu thành alkyl este như methyl hoặc ethyl este. Việc xử lý tiếp theo nữa có thể được áp dụng để làm giàu hàm lượng DHA trong lipit hoặc dầu.

Theo một phương án được ưu tiên hơn, lipit hoặc dầu, tốt hơn là dầu hạt, có các đặc điểm sau: trong tổng hàm lượng axit béo của lipit hoặc dầu, mức DHA nằm trong khoảng từ 7% đến 20%, mức axit palmitic nằm trong khoảng từ 2% đến 16%, mức axit myristic là thấp hơn khoảng 6%, mức axit oleic nằm trong khoảng từ 1% đến 30%, mức LA nằm trong khoảng từ 4% đến 35%, có mặt ALA, có mặt GLA, mức SDA nằm trong khoảng từ 0,05% đến 7%, mức ETA là thấp hơn khoảng 6%, mức EPA nằm trong khoảng từ 0,05% đến 10%, mức DPA nằm trong khoảng từ 0,05% đến 8%.

Theo phương án mở rộng, lipit chiết tách còn bao gồm một hoặc nhiều sterol, tốt hơn là sterol thực vật.

Theo phương án khác, lipit chiết tách là ở dạng dầu, và chứa sterol với lượng ít hơn khoảng 10mg sterol/g dầu, ít hơn khoảng 7mg sterol/g dầu, nằm trong khoảng từ 1,5mg đến 10mg sterol/g dầu, hoặc nằm trong khoảng từ 1,5mg đến 7mg sterol/g dầu.

Ví dụ về sterol có thể có trong lipit chiết tách bao gồm, nhưng không giới hạn ở, một hoặc nhiều hoặc tất cả campesterol/24-methylcholesterol, Δ^5 -stigmasterol, eburicol, β -sitosterol/24-ethylcholesterol, Δ^5 -avenasterol/isofucosterol, Δ^7 -stigmasterol/stigmast-7-en-3 β -ol, và Δ^7 -avenasterol.

Theo một phương án, loài thực vật được nêu trong Bảng 26, như cải dầu, và mức sterol là như được nêu trong Bảng 26 đối với các loài thực vật cụ thể.

Theo một phương án, lipit chiết tách chứa cholesterol với lượng ít hơn khoảng 0,5mg cholesterol/g dầu, ít hơn khoảng 0,25mg cholesterol/g dầu, nằm trong khoảng từ 0mg đến 0,5mg cholesterol/g dầu, hoặc nằm trong khoảng từ 0mg đến 0,25mg cholesterol/g dầu, hoặc về cơ bản không có cholesterol.

Theo phương án mở rộng, lipit là dầu, tốt hơn là dầu từ hạt dầu. Ví dụ về các dầu như vậy bao gồm, nhưng không hạn chế ở, dầu *Brassica sp.* như dầu cải, dầu *Gossypium hirsutum*, dầu *Linum usitatissimum*, dầu *Helianthus sp.*, dầu *Carthamus tinctorius*, dầu *Glycine max*, dầu *Zea mays*, dầu *Arabidopsis thaliana*, dầu *Sorghum*

bicolor, dầu *Sorghum vulgare*, dầu *Avena sativa*, dầu *Trifolium sp.*, dầu *Elaeagnus guineensis*, dầu *Nicotiana benthamiana*, dầu *Hordeum vulgare*, dầu *Lupinus angustifolius*, dầu *Oryza sativa*, dầu *Oryza glaberrima*, dầu *Camelina sativa*, dầu *Crambe abyssinica*, dầu *Miscanthus x giganteus*, hoặc dầu *Miscanthus sinensis*.

Sáng chế cũng đề xuất lipit chiết tách từ cây, tốt hơn là dầu chiết từ hạt cải dầu, chứa các axit béo ở dạng este hóa, các axit béo chứa axit oleic, axit palmitic, axit béo $\omega 6$ chứa axit linoleic (LA), axit béo $\omega 3$ bao gồm axit α -linolenic (ALA), và axit docosahexaenoic (DHA), và tùy chọn một hoặc nhiều axit stearidonic (SDA), axit eicosapentaenoic (EPA), axit docosapentaenoic (DPA) và axit eicosatetraenoic (ETA), trong đó lipit có các đặc điểm sau trong toàn bộ thành phần axit béo của lipit;

- (i) mức DHA là khoảng 3%, khoảng 4%, khoảng 5%, khoảng 6% hoặc khoảng 7%
- (ii) mức axit palmitic nằm trong khoảng từ 2% đến 16%,
- (iii) mức axit myristic là thấp hơn khoảng 2%,
- (iv) mức axit oleic nằm trong khoảng từ 30% đến 60%, tốt hơn nằm trong khoảng từ 45% đến 60%,
- (v) mức LA nằm trong khoảng từ 4% đến 20%,
- (vi) mức ALA nằm trong khoảng từ 2% đến 16%,
- (vii) mức GLA là thấp hơn khoảng 4%,
- (viii) mức SDA là thấp hơn khoảng 6%, hoặc thấp hơn khoảng 4%,
- (ix) mức ETA là thấp hơn khoảng 6%, hoặc thấp hơn khoảng 4%,
- (x) mức ETrA thấp hơn khoảng 1%,
- (xi) mức EPA là thấp hơn khoảng 10% và/hoặc mức EPA là từ 0,5-2,0 lần so với mức DHA,
- (xii) mức DPA là thấp hơn khoảng 4%,
- (xiii) mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 4% đến 25%,
- (xiv) mức toàn bộ axit béo một liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 30% đến 70%,
- (xv) mức toàn bộ axit béo nhiều liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 15% đến 75%, tốt hơn nằm trong khoảng từ 15% đến 30%,

(xvi) mức axit béo $\omega 6$ mới trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 0,5% đến 10%,

(xvii) mức toàn bộ axit béo $\omega 3$ trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 10% đến 20%,

(xviii) mức axit béo $\omega 3$ mới trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 3% đến 20%,

(xix) tỷ lệ của toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$ trong hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 0,05 đến 3,0, tốt hơn là thấp hơn khoảng 0,50,

(xx) tỷ lệ của axit béo $\omega 6$ mới: axit béo $\omega 3$ mới trong hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 0,03 đến 3,0,

(xxi) hàm lượng triaxylglycerol (TAG) trong lipit ít nhất là khoảng 70%, và (xxii) lipit về cơ bản không có cholesterol. Theo một phương án, lipit chứa tri- DHA TAG (TAG 66:18). Tốt hơn nữa là, lipit về cơ bản không có SDA và ETA, và/hoặc không bị xử lý bằng quy trình chuyển hóa este sau khi nó được chiết từ cây hoặc bộ phận cây.

Theo khía cạnh khác, sáng chế đề xuất lipit chiết tách từ cây, chứa các axit béo ở dạng este hóa, axit béo bao gồm axit oleic, axit palmitic, axit béo $\omega 6$ bao gồm axit linoleic (LA), axit béo $\omega 3$ bao gồm axit α -linolenic (ALA) và axit docosahexaenoic (DHA), và một hoặc nhiều axit stearidonic (SDA), axit eicosapentaenoic (EPA), axit docosapentaenoic (DPA) và axit eicosatetraenoic (ETA), trong đó (i) mức DHA trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 7% đến 20%, (ii) mức axit palmitic trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 2% đến 16%, (iii) mức axit myristic (C14:0) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 6%, (iv) mức axit oleic trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 1% đến 30% hoặc nằm trong khoảng từ 30% đến 60%, (v) mức axit linoleic (LA) trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 4% đến 35%, (vi) mức axit α -linolenic (ALA) trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 4% đến 40%, (vii) mức axit eicosatrienoic (ETrA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 4%, (viii) mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 4% và 25%, (ix) tỷ lệ của toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$ trong hàm

lượng axit béo của lipid chiết tách nằm trong khoảng từ 1,0 đến 3,0 hoặc nằm trong khoảng từ 0,1 đến 1, (x) hàm lượng triaxylglyxerol (TAG) trong lipid ít nhất là 70%, và (xi) ít nhất 70% DHA được este hóa ở dạng TAG là ở vị trí *sn*-1 hoặc *sn*-3 của TAG.

Theo một phương án, lipid theo sáng chế có một hoặc nhiều hoặc tất cả các đặc điểm sau:

- i) mức axit palmitic trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 2% đến 15%,
- ii) mức axit myristic (C14:0) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn 1%,
- iii) mức axit oleic trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 3% đến 30%, nằm trong khoảng từ 6% đến 30%, từ 1% đến 20%, nằm trong khoảng từ 45% đến 60%, hoặc là khoảng 30%,
- iv) mức axit linoleic (LA) trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 20%, hoặc nằm trong khoảng từ 4% đến 17%,
- v) mức axit α -linolenic (ALA) trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 7% đến 40%, nằm trong khoảng từ 10% đến 35%, nằm trong khoảng từ 20% đến 35%, hoặc nằm trong khoảng từ 4% và 16%,
- vi) mức axit γ -linolenic (GLA) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn 4%, thấp hơn khoảng 3%, thấp hơn khoảng 2%, thấp hơn khoảng 1%, thấp hơn khoảng 0,5%, nằm trong khoảng từ 0,05% đến 7%, từ 0,05% đến 4%, hoặc nằm trong khoảng từ 0,05% đến 3%, hoặc nằm trong khoảng từ 0,05% đến 2%,
- vii) mức axit stearidonic (SDA) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn khoảng 4%, thấp hơn khoảng 3%, nằm trong khoảng từ 0,05% đến 7%, nằm trong khoảng từ 0,05% đến 4%, nằm trong khoảng từ 0,05% đến 3%, hoặc nằm trong khoảng từ 0,05% đến 2%,
- viii) mức axit eicosatetraenoic (ETA) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn khoảng 4%, thấp hơn khoảng 1%, thấp hơn khoảng 0,5%, nằm trong khoảng từ 0,05% đến 5%, từ 0,05% đến 4%, từ 0,05% đến 3%, hoặc nằm trong khoảng từ 0,05% đến 2%,
- ix) mức axit eicosatrienoic (ETrA) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn khoảng 2%, thấp hơn khoảng 1%, nằm trong khoảng từ 0,05% đến 4%, nằm trong khoảng từ 0,05% đến 3%, hoặc nằm trong khoảng từ 0,05% đến

2%, hoặc nằm trong khoảng từ 0,05% đến 1%,

x) mức axit eicosapentaenoic (EPA) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn 4%, thấp hơn khoảng 3%, thấp hơn khoảng 2%, nằm trong khoảng từ 0,05% đến 10%, nằm trong khoảng từ 0,05% và 5%, hoặc nằm trong khoảng từ 0,05% đến 3%, hoặc nằm trong khoảng từ 0,05% đến 2%,

xi) mức axit docosapentaenoic (DPA) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn 4%, thấp hơn khoảng 3%, thấp hơn khoảng 2%, nằm trong khoảng từ 0,05% đến 8%, từ 0,05% đến 5%, hoặc nằm trong khoảng từ 0,05% từ 3%, hoặc nằm trong khoảng từ 0,05% đến 2%,

xii) mức DHA trong tổng hàm lượng axit béo của lipid chiết tách là khoảng 8%, khoảng 9%, khoảng 10%, khoảng 11%, khoảng 12%, khoảng 13%, khoảng 14%, khoảng 15%, khoảng 16%, khoảng 17%, khoảng 18%, nằm trong khoảng từ 8% đến 20%, nằm trong khoảng từ 10% đến 20%, nằm trong khoảng từ 11% đến 20%, nằm trong khoảng từ 10% đến 16%, hoặc nằm trong khoảng từ 14% đến 20%,

xiii) lipid chứa axit ω 6-docosapentaenoic ($22:5^{\Delta 4,7,10,13,16}$) trong hàm lượng axit béo của nó,

xiv) lipid về cơ bản không có axit ω 6-docosapentaenoic ($22:5^{\Delta 4,7,10,13,16}$) trong thành phần axit béo của nó,

xv) lipid về cơ bản không có SDA, EPA và ETA trong thành phần axit béo của nó,

xvi) mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 20%, hoặc nằm trong khoảng từ 6% đến 20%,

xvii) mức toàn bộ axit béo một liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 35%, từ 8% đến 25%, hoặc nằm trong khoảng từ 8% đến 22%,

xviii) mức toàn bộ axit béo nhiều liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 20% đến 75%, nằm trong khoảng từ 50% đến 75%, hoặc nằm trong khoảng từ 60% đến 75%,

xix) mức toàn bộ axit béo ω 6 trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 35% đến 50%, nằm trong khoảng từ 20% đến 35%, nằm trong khoảng từ 6% và 20%, thấp hơn 20%, thấp hơn khoảng 16%, thấp hơn khoảng 10%, nằm trong khoảng từ 1% đến 16%, nằm trong khoảng từ 2% đến 10%, hoặc nằm

trong khoảng từ 4% đến 10%,

xx) mức axit béo $\omega 6$ mới trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 10%, thấp hơn khoảng 8%, thấp hơn khoảng 6%, thấp hơn 4%, nằm trong khoảng từ 1% đến 20%, nằm trong khoảng từ 1% đến 10%, nằm trong khoảng từ 0,5% đến 8%, hoặc nằm trong khoảng từ 0,5% đến 4%,

xxi) mức toàn bộ axit béo $\omega 3$ trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 36% đến 65%, between 40% đến 60%, nằm trong khoảng từ 20% đến 35%, nằm trong khoảng từ 10% đến 20%, khoảng 25%, khoảng 30%, khoảng 35% hoặc khoảng 40%,

xxii) mức axit béo $\omega 3$ mới trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 9% đến 33%, nằm trong khoảng từ 10% đến 20%, nằm trong khoảng từ 20% đến 30%, nằm trong khoảng từ 12% đến 25%, khoảng 13%, khoảng 15%, khoảng 17% hoặc khoảng 20%,

xxiii) tỷ lệ của toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$ trong hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 0,1 đến 0,5, thấp hơn khoảng 0,50, thấp hơn khoảng 0,40, thấp hơn khoảng 0,30, thấp hơn khoảng 0,20, thấp hơn khoảng 0,15, khoảng 1,0, khoảng 0,1 hoặc khoảng 0,2,

xxiv) tỷ lệ của axit béo $\omega 6$ mới: axit béo $\omega 3$ mới trong hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 1,0 khoảng 3,0, nằm trong khoảng từ 0,1 khoảng 1, nằm trong khoảng từ 0,1 khoảng 0,5, thấp hơn khoảng 0,50, thấp hơn khoảng 0,40, thấp hơn khoảng 0,30, thấp hơn khoảng 0,20, thấp hơn khoảng 0,15, khoảng 0,1, khoảng 0,2 hoặc khoảng 1,0,

xxv) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa axit oleic thành DHA ít nhất là khoảng 10%, ít nhất khoảng 15%, ít nhất khoảng 20%, nằm trong khoảng từ 10% đến 50%, nằm trong khoảng từ 10% đến 30%, hoặc nằm trong khoảng từ 10% đến 25%,

xxvi) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa LA thành DHA ít nhất là khoảng 15%, ít nhất khoảng 20%, ít nhất khoảng 22%, ít nhất khoảng 25%, nằm trong khoảng từ 15% đến 50%, nằm trong khoảng từ 20% đến 40%, hoặc nằm trong khoảng từ 20% đến 30%,

xxvii) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa ALA thành DHA ít nhất là khoảng 17%, ít nhất khoảng 22%, ít nhất khoảng 24%, nằm trong

khoảng từ 17% đến 55%, nằm trong khoảng từ 22% đến 35%, hoặc nằm trong khoảng từ 24% đến 35%,

xxviii) toàn bộ axit béo trong lipit chiết tách có C20:1 với lượng ít hơn 1%,

xxix) hàm lượng triaxylglyxerol (TAG) trong lipit ít nhất là khoảng 80%, ít nhất khoảng 90%, ít nhất 95%, nằm trong khoảng từ 70% đến 99%, hoặc nằm trong khoảng từ 90% đến 99%,

xxx) lipit chứa diaxylglyxerol (DAG),

xxxi) lipit chứa ít hơn khoảng 10%, thấp hơn khoảng 5%, thấp hơn khoảng 1%, hoặc nằm trong khoảng từ 0,001% đến 5%, axit béo tự do (không este hóa) và/hoặc phospholipit, hoặc về cơ bản không có chúng,

xxxii) ít nhất 80%, DHA được este hóa ở dạng TAG là ở vị trí *sn*-1 hoặc vị trí *sn*-3 trong TAG,

xxxiii) kiểu TAG chứa DHA nhiều nhất trong lipit là DHA/18:3/18:3 (TAG 58:12), và

xxxiv) lipit chứa tri-DHA TAG (TAG 66:18).

Đối với khía cạnh trên, theo một phương án

i) lipit là ở dạng dầu, trong đó dầu này chứa một hoặc nhiều sterol như một hoặc nhiều hoặc tất cả các campesterol, Δ 5-stigmasterol, eburicol, β -sitosterol, Δ 5-avenasterol, Δ 7-stigmasterol và Δ 7-avenasterol, và tùy chọn dầu này chứa ít hơn 10mg sterol/g dầu và/hoặc dầu về cơ bản không có cholesterol, và/hoặc

ii) lipit là ở dạng dầu từ hạt dầu như là hạt dầu *Brassica sp* hoặc hạt cải dầu.

Theo khía cạnh khác, sáng chế đề xuất quy trình sản xuất lipit chiết tách từ cây, quy trình này bao gồm các bước

i) thu hoạch bộ phận cây chứa lipit, lipit này chứa các axit béo ở dạng este hóa, các axit béo bao gồm axit oleic, axit palmitic, axit béo ω 6 bao gồm axit linoleic (LA), axit béo ω 3 bao gồm axit α -linolenic (ALA), và axit docosahexaenoic (DHA), và tùy chọn một hoặc nhiều axit eicosapentaenoic (EPA), axit stearidonic (SDA), axit docosapentaenoic (DPA) và axit eicosatetraenoic (ETA), trong đó mức DHA trong tổng hàm lượng axit béo của lipit chiết được trong bộ phận cây là khoảng từ 7% đến 20%, và

ii) chiết lipit từ bộ phận cây này, trong đó mức DHA trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 7% đến 20%.

Theo một phương án được ưu tiên, lipit chiết tách có một hoặc nhiều đặc điểm nêu trên.

Theo một phương án, trong đó bộ phận cây là hạt, tốt hơn là hạt dầu. Ví dụ về hạt này bao gồm, nhưng không hạn chế ở, hạt *Brassica sp.*, *Gossypium hirsutum*, *Linum usitatissimum*, *Helianthus sp.*, *Carthamus tinctorius*, *Glycine max*, *Zea mays*, *Arabidopsis thaliana*, *Sorghum bicolor*, *Sorghum vulgare*, *Avena sativa*, *Trifolium sp.*, *Elaeis guineensis*, *Nicotiana benthamiana*, *Hordeum vulgare*, *Lupinus angustifolius*, *Oryza sativa*, *Oryza glaberrima*, *Camelina sativa*, hoặc *Crambe abyssinica*, tốt hơn là hạt *Brassica napus*, *B. juncea* hoặc *C. sativa*.

Theo phương án khác, hạt này chứa DHA với lượng ít nhất khoảng 18mg, ít nhất khoảng 22mg, ít nhất khoảng 26mg, nằm trong khoảng từ 18mg đến 100mg, nằm trong khoảng từ 22mg đến 70mg, hoặc nằm trong khoảng từ 24mg đến 50mg, trong mỗi gam hạt.

Theo phương án mở rộng, bộ phận cây chứa polynucleotit ngoại sinh mã hóa một trong các bộ enzym sau;

- i) ω 3-desaturaza, Δ 6-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 6-elongaza và Δ 5-elongaza,
 - ii) Δ 15-desaturaza, Δ 6-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 6-elongaza và Δ 5-elongaza,
 - iii) Δ 12-desaturaza, Δ 6-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 6-elongaza và Δ 5-elongaza,
 - iv) Δ 12-desaturaza, ω 3-desaturaza hoặc Δ 15-desaturaza, Δ 6-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 6-elongaza và Δ 5-elongaza,
 - v) ω 3-desaturaza, Δ 8-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 9-elongaza và Δ 5-elongaza,
 - vi) Δ 15-desaturaza, Δ 8-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 9-elongaza và Δ 5-elongaza,
 - vii) Δ 12-desaturaza, Δ 8-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 9-elongaza và Δ 5-elongaza, hoặc
 - viii) Δ 12-desaturaza, ω 3-desaturaza hoặc Δ 15-desaturaza, Δ 8-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 9-elongaza và Δ 5-elongaza,
- và trong đó mỗi polynucleotit được liên kết điều khiển với một hoặc nhiều

gen khởi đầu có khả năng điều khiển quá trình biểu hiện của polynucleotit trong tế bào của bộ phận cây.

Theo một phương án rộng hơn, bộ phận cây có một hoặc nhiều hoặc tất cả các đặc điểm sau:

i) $\Delta 12$ -desaturaza chuyển hóa axit oleic thành axit linoleic trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 60%, ít nhất khoảng 70%, ít nhất khoảng 80%, nằm trong khoảng từ 60% đến 98%, nằm trong khoảng từ 70% đến 95%, hoặc nằm trong khoảng từ 75% đến 90%,

ii) $\omega 3$ -desaturaza chuyển hóa axit béo $\omega 6$ thành axit béo $\omega 3$ trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 65%, ít nhất khoảng 75%, ít nhất khoảng 85%, nằm trong khoảng từ 65% đến 95%, nằm trong khoảng từ 75% đến 95%, hoặc nằm trong khoảng từ 80% đến 95%,

iii) $\Delta 6$ -desaturaza chuyển hóa ALA thành SDA trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 30%, ít nhất khoảng 40%, ít nhất khoảng 50%, ít nhất khoảng 60%, ít nhất khoảng 70%, nằm trong khoảng từ 30% đến 70%, nằm trong khoảng từ 35% đến 60%, hoặc nằm trong khoảng từ 50% đến 70%,

iv) $\Delta 6$ -desaturaza chuyển hóa axit linoleic thành axit γ -linolenic trong một hoặc nhiều tế bào của cây với hiệu suất thấp hơn khoảng 5%, thấp hơn khoảng 2,5%, thấp hơn khoảng 1%, nằm trong khoảng từ 0,1% đến 5%, nằm trong khoảng từ 0,5% đến 2,5%, hoặc nằm trong khoảng từ 0,5% đến 1%,

v) $\Delta 6$ -elongaza chuyển hóa SDA thành ETA trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 60%, ít nhất khoảng 70%, ít nhất khoảng 75%, nằm trong khoảng từ 60% đến 95%, nằm trong khoảng từ 70% đến 88%, hoặc nằm trong khoảng từ 75% đến 85%,

vi) $\Delta 5$ -desaturaza chuyển hóa ETA thành EPA trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 60%, ít nhất khoảng 70%, ít nhất khoảng 75%, ít nhất khoảng 80%, ít nhất khoảng 90%, nằm trong khoảng từ 60% đến 99%, nằm trong khoảng từ 70% đến 99%, hoặc nằm trong khoảng từ 75% đến 98%,

vii) $\Delta 5$ -elongaza chuyển hóa EPA thành DPA trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 80%, ít nhất khoảng 85%, ít nhất khoảng 90%, nằm trong khoảng từ 50% đến 95%, hoặc nằm trong khoảng từ 85% đến 95%,

viii) $\Delta 4$ -desaturaza chuyển hóa DPA thành DHA trong một hoặc nhiều tế bào

của cây với hiệu suất ít nhất là khoảng 80%, ít nhất khoảng 90%, ít nhất khoảng 93%, nằm trong khoảng từ 50% đến 95%, nằm trong khoảng từ 80% đến 95%, hoặc nằm trong khoảng từ 85% đến 95%,

ix) hiệu suất chuyển hóa axit oleic thành DHA trong một hoặc nhiều tế bào của bộ phận cây ít nhất là khoảng 10%, ít nhất khoảng 15%, ít nhất khoảng 20%, nằm trong khoảng từ 10% đến 50%, nằm trong khoảng từ 10% đến 30%, hoặc nằm trong khoảng từ 10% đến 25%,

x) hiệu suất chuyển hóa LA thành DHA trong một hoặc nhiều tế bào bộ phận cây ít nhất là khoảng 15%, ít nhất khoảng 20%, ít nhất khoảng 22%, ít nhất khoảng 25%, nằm trong khoảng từ 15% đến 50%, nằm trong khoảng từ 20% đến 40%, hoặc nằm trong khoảng từ 20% đến 30%,

xi) hiệu suất chuyển hóa ALA thành DHA trong một hoặc nhiều tế bào bộ phận cây ít nhất là khoảng 17%, ít nhất khoảng 22%, ít nhất khoảng 24%, nằm trong khoảng từ 17% đến 55%, nằm trong khoảng từ 22% đến 35%, hoặc nằm trong khoảng từ 24% đến 35%,

xii) một hoặc nhiều tế bào của bộ phận cây chứa ít nhất khoảng 15%, ít nhất khoảng 20%, nằm trong khoảng từ 15% đến 30%, hoặc nằm trong khoảng từ 22,5% đến 27,5%, nhiều axit béo $\omega 3$ hơn các tế bào tương ứng không có polynucleotit ngoại sinh này,

xiii) $\Delta 6$ -desaturaza ưu tiên khử bão hòa axit α -linolenic (ALA) tương ứng thành axit linoleic (LA),

xiv) $\Delta 6$ -elongaza cũng có hoạt tính $\Delta 9$ -elongaza,

xv) $\Delta 12$ -desaturaza cũng có hoạt tính $\Delta 15$ -desaturaza,

xvi) $\Delta 6$ -desaturaza cũng có hoạt tính $\Delta 8$ -desaturaza,

xvii) $\Delta 8$ -desaturaza cũng có hoạt tính $\Delta 6$ -desaturaza hoặc không có hoạt tính $\Delta 6$ -desaturaza,

xviii) $\Delta 15$ -desaturaza cũng có hoạt tính $\omega 3$ -desaturaza trên GLA,

xix) $\omega 3$ -desaturaza cũng có hoạt tính $\Delta 15$ -desaturaza trên LA,

xx) $\omega 3$ -desaturaza khử bão hòa cả LA và/hoặc GLA,

xxi) $\omega 3$ -desaturaza ưu tiên khử bão hòa GLA tương ứng thành LA,

xxii) mức DHA trong bộ phận cây là dựa trên hiệu suất chuyển hóa axit oleic thành DHA trong bộ phận cây này ít nhất là khoảng 10%, ít nhất khoảng 15%, ít nhất

khoảng 20%, nằm trong khoảng từ 10% đến 50%, nằm trong khoảng từ 15% đến 30%, hoặc nằm trong khoảng từ 20% đến 25%,

xxiii) mức DHA trong bộ phận cây là dựa trên hiệu suất chuyển hóa LA thành DHA trong bộ phận cây ít nhất là khoảng 15%, ít nhất khoảng 20%, ít nhất khoảng 22%, nằm trong khoảng từ 15% đến 60%, nằm trong khoảng từ 20% đến 40%, hoặc nằm trong khoảng từ 22% đến 30%,

xxiv) mức DHA trong bộ phận cây là dựa trên hiệu suất chuyển hóa ALA thành DHA trong bộ phận cây ít nhất là khoảng 17%, ít nhất khoảng 22%, ít nhất khoảng 24%, nằm trong khoảng từ 17% đến 65%, nằm trong khoảng từ 22% đến 35%, hoặc nằm trong khoảng từ 24% đến 35%

xxx) một hoặc nhiều hoặc tất cả các desaturaza hoạt tính trên cơ chất axyl-CoA cao hơn so với cơ chất axyl-PC tương ứng,

xxxi) $\Delta 6$ -desaturaza có hoạt tính $\Delta 6$ -desaturaza trên ALA cao hơn so với LA là cơ chất axit béo,

xxxii) $\Delta 6$ -desaturaza có hoạt tính $\Delta 6$ -desaturaza trên ALA-CoA là cơ chất axit béo cao hơn so với trên ALA được gắn với vị trí *sn*-2 của PC là cơ chất cơ chất axit béo,

xxxiii) $\Delta 6$ -desaturaza có hoạt tính $\Delta 6$ -desaturaza cao hơn ít nhất khoảng 2 lần, hoạt tính cao hơn ít nhất 3 lần, hoạt tính cao hơn ít nhất 4 lần, hoặc hoạt tính cao hơn ít nhất 5 lần, trên ALA là cơ chất so với LA,

xxxiv) $\Delta 6$ -desaturaza có hoạt tính trên ALA-CoA làm cơ chất axit béo cao hơn so với trên ALA được gắn với vị trí *sn*-2 của PC làm cơ chất axit béo,

xxxv) $\Delta 6$ -desaturaza có hoạt tính cao hơn ít nhất khoảng 5 lần so với hoạt tính $\Delta 6$ -desaturaza hoặc ít nhất 10 lần, trên ALA-CoA làm cơ chất axit béo so với trên ALA được gắn với vị trí *sn*-2 của PC làm cơ chất axit béo,

xxxvi) desaturaza là tiền-desaturaza,

xxxvii) $\Delta 6$ -desaturaza có hoạt tính $\Delta 5$ -desaturaza trên ETA không phát hiện được.

Theo một phương án rộng hơn, bộ phận cây có một hoặc nhiều hoặc tất cả các đặc điểm sau:

i) $\Delta 12$ -desaturaza bao gồm các axit amin có trình tự như được nêu trong SEQ ID NO: 10, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất

là 50% với SEQ ID NO: 10,

ii) ω 3-desaturaza bao gồm các axit amin có trình tự như được nêu trong SEQ ID NO: 12, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 50% với SEQ ID NO: 12,

iii) Δ 6-desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 16, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 50% với SEQ ID NO: 16,

iv) Δ 6-elongaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 25, mảnh hoạt tính sinh học của nó như SEQ ID NO: 26, hoặc trình tự axit amin tương đồng ít nhất là 50% với SEQ ID NO: 25 và/hoặc SEQ ID NO: 26,

v) Δ 5-desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 30, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 50% với SEQ ID NO: 30,

vi) Δ 5-elongaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 37, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 50% với SEQ ID NO: 37,

vii) Δ 4-desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 41, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 50% với SEQ ID NO: 41.

Theo một phương án, bộ phận cây còn bao gồm polynucleotit ngoại sinh mã hóa diacylglycerol axyltransferaza (DGAT), monoacylglycerol axyltransferaza (MGAT), glycerol-3-phosphat axyltransferaza (GPAT), 1-axyl-glycerol-3-phosphat axyltransferaza (LPAAT) tốt hơn là LPAAT có thể sử dụng cơ chất axyl-CoA béo nhiều liên kết chưa bão hòa C22, axyl-CoA:lysophosphatidylcholin axyltransferaza (LPCAT), phospholipaza A₂ (PLA₂), phospholipaza C (PLC), phospholipaza D (PLD), CDP-cholin diacylglycerol cholin phosphotransferaza (CPT), phosphatidylcholin diacylglycerol axyltransferaza (PDAT), phosphatidylcholin:diacylglycerol cholin phosphotransferaza (PDCT), axyl-CoA synthaza (ACS), hoặc kết hợp hai hoặc nhiều enzym này.

Theo phương án khác, bộ phận cây còn bao gồm polynucleotit ngoại sinh hoặc đột biến được đưa vào điều hòa xuôi dòng sự sản sinh và/hoặc hoạt tính của enzym nội sinh trong tế bào của bộ phận cây được chọn từ FAE1, DGAT, MGAT, GPAT,

LPAAT, LPCAT, PLA₂, PLC, PLD, CPT, PDAT, thioesteraza như FATB, hoặc $\Delta 12$ -desaturaza, hoặc kết hợp hai hoặc nhiều enzym này.

Theo phương án mở rộng, ít nhất một, hoặc tất cả các gen khởi đầu là gen khởi đầu đặc hiệu với hạt. Theo một phương án, ít nhất một, hoặc tất cả các gen khởi đầu được thu lấy từ quá trình sinh tổng hợp dầu hoặc gen tích lũy như oleosin, hoặc từ các gen protein tích trữ trong hạt như conlinin.

Theo phương án khác, (các) gen khởi đầu điều khiển quá trình biểu hiện polynucleotit ngoại sinh mã hóa $\Delta 4$ -desaturaza và $\Delta 5$ -elongaza khởi động quá trình biểu hiện polynucleotit trong hạt đang phát triển của bộ phận cây trước khi, hoặc đạt đến ngưỡng biểu hiện trước khi, (các) gen khởi đầu điều khiển quá trình biểu hiện polynucleotit ngoại sinh mã hóa $\Delta 12$ -desaturaza và $\omega 3$ -desaturaza.

Theo phương án mở rộng, polynucleotit ngoại sinh được liên kết đồng hóa trị trong phân tử ADN, tốt hơn là phân tử T-ADN, được sát nhập vào hệ gen của tế bào của bộ phận cây và tốt hơn là trong đó số lượng phân tử ADN được sát nhập vào hệ gen của tế bào của bộ phận cây là không nhiều hơn một, hai hoặc ba, hoặc là hai hoặc ba phân tử.

Theo một phương án khác nữa, cây bao gồm ít nhất hai polynucleotit ngoại sinh, độc lập mỗi polynucleotit mã hóa $\Delta 6$ -desaturaza có trình tự axit amin giống hoặc khác nhau.

Theo phương án mở rộng, tổng hàm lượng dầu của bộ phận cây bao gồm polynucleotit ngoại sinh ít nhất là khoảng 40%, hoặc ít nhất khoảng 50%, hoặc ít nhất khoảng 60%, hoặc ít nhất khoảng 70%, hoặc nằm trong khoảng từ 50% đến 80% tổng hàm lượng dầu của bộ phận cây tương ứng không có polynucleotit ngoại sinh này. Theo các phương án này, hàm lượng dầu tối đa có thể là khoảng 100% hàm lượng dầu của bộ phận cây kiểu dại tương ứng.

Theo phương án khác, lipit là ở dạng dầu, tốt hơn là dầu hạt từ hạt dầu, và trong đó triaxylglyxerol có mặt với lượng ít nhất khoảng 90%, hoặc ít nhất khoảng 95%, ít nhất khoảng 98%, hoặc nằm trong khoảng từ 95% đến 98%, theo trọng lượng lipit.

Theo phương án mở rộng, quy trình này còn bao gồm việc xử lý lipit để làm tăng mức DHA ở dạng phần trăm tổng hàm lượng axit béo. Ví dụ, việc xử lý là bước chuyển hóa este. Ví dụ, lipit như dầu cải có thể được xử lý để chuyển hóa các axit béo trong dầu thành alkyl este như metyl hoặc etyl este, sau đó có thể được cất phân đoạn để làm

giàu DHA trong lipit hoặc dầu.

Hơn nữa, sáng chế đề xuất quy trình tạo ra lipit chiết tách từ cây, quy trình này bao gồm các bước

i) thu hoạch bộ phận cây, tốt hơn là hạt cải dầu, bao gồm lipit, lipit này bao gồm các axit béo ở dạng este hóa, các axit béo này bao gồm axit oleic, axit palmitic, axit béo $\omega 6$ bao gồm axit linoleic (LA), axit béo $\omega 3$ bao gồm axit α -linolenic (ALA), và axit docosahexaenoic (DHA), và tùy chọn một hoặc nhiều axit eicosapentaenoic (EPA), axit stearidonic (SDA), axit docosapentaenoic (DPA) và axit eicosatetraenoic (ETA), trong đó mức DHA trong tổng hàm lượng axit béo của lipit chiết được trong bộ phận cây là khoảng 3%, khoảng 4%, khoảng 5%, khoảng 6% hoặc khoảng 7%, và

ii) chiết lipit từ bộ phận cây này,

trong đó lipit chiết tách có các đặc điểm sau trong tổng hàm lượng axit béo của lipit;

i) mức DHA là khoảng 3%, khoảng 4%, khoảng 5%, khoảng 6% hoặc khoảng 7%,

ii) mức axit palmitic nằm trong khoảng từ 2% đến 16%,

iii) mức axit myristic là thấp hơn khoảng 2%,

iv) mức axit oleic nằm trong khoảng từ 30% đến 60%, tốt hơn nằm trong khoảng từ 45% đến 60%,

v) mức LA nằm trong khoảng từ 4% đến 20%,

vi) mức ALA nằm trong khoảng từ 2% đến 16%,

vii) mức GLA là thấp hơn khoảng 4%,

viii) mức SDA là thấp hơn khoảng 6%, hoặc thấp hơn khoảng 4%,

ix) mức ETA là thấp hơn khoảng 6%, hoặc thấp hơn khoảng 4%,

x) mức ETrA thấp hơn khoảng 1%,

xi) mức EPA là thấp hơn khoảng 10% và/hoặc mức EPA là 0,5-2,0 lần mức DHA,

xii) mức DPA là thấp hơn khoảng 4%,

xiii) mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 4% đến 25%,

xiv) mức toàn bộ axit béo một liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 30% đến 70%,

xv) mức toàn bộ axit béo nhiều liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 15% đến 75%, tốt hơn nằm trong khoảng từ 15% đến 30%,

xvi) mức axit béo $\omega 6$ mới trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 0,5% đến 10%,

xvii) mức toàn bộ axit béo $\omega 3$ trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 10% đến 20%,

xviii) mức axit béo $\omega 3$ mới trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 3% đến 20%,

xix) tỷ lệ của toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$ trong hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 0,05 đến 3,0, tốt hơn là thấp hơn khoảng 0,50,

xx) tỷ lệ của axit béo $\omega 6$ mới: axit béo $\omega 3$ mới trong hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 0,03 đến 3,0,

xxi) hàm lượng triaxylglyxerol (TAG) trong lipid ít nhất là khoảng 70%, và

xxii) lipid về cơ bản không có cholesterol. Theo một phương án, lipid chứa tri-DHA TAG (TAG 66:18). Tốt hơn nữa là, lipid về cơ bản không có SDA và ETA, và/hoặc không bị xử lý bằng quy trình chuyển hóa este sau khi nó được chiết từ cây hoặc bộ phận cây.

Sáng chế cũng đề xuất quy trình tạo ra lipid chiết tách từ cây, quy trình này bao gồm các bước sau:

(i) thu hoạch bộ phận cây chứa lipid, lipid bao gồm các axit béo ở dạng este hóa, các axit béo này bao gồm axit oleic, axit palmitic, axit béo $\omega 6$ bao gồm axit linoleic (LA), axit béo $\omega 3$ bao gồm axit α -linolenic (ALA) và axit docosahexaenoic (DHA), và một hoặc nhiều axit stearidonic (SDA), axit eicosapentaenoic (EPA), axit docosapentaenoic (DPA) và axit eicosatetraenoic (ETA), trong đó (i) mức DHA trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 7% đến 20%, (ii) mức axit palmitic trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 2% đến 16%, (iii) mức axit myristic (C14:0) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn 6%, (iv) mức axit oleic trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 1% đến 30% hoặc từ 30% đến 60%, (v) mức axit linoleic (LA) trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 35%, (vi) mức axit α -linolenic (ALA) trong tổng hàm lượng axit béo của lipid chiết tách

nằm trong khoảng từ 4% đến 40%, (vii) mức axit eicosatrienoic (ETrA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 4%, (viii) mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 4% và 25%, (ix) tỷ lệ của toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$ trong hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 1,0 đến 3,0 hoặc nằm trong khoảng từ 0,1 đến 1, (x) hàm lượng triaxylglyxerol (TAG) trong lipit ít nhất là 70%, và (xi) ít nhất 70% DHA được este hóa ở dạng TAG là ở vị trí *sn*-1 hoặc *sn*-3 trong TAG, và

(ii) chiết lipit từ bộ phận cây này, trong đó mức DHA trong tổng hàm lượng axit béo của lipit chiết tách là khoảng 7% đến 20%.

Sáng chế cũng đề xuất lipit, hoặc dầu chứa lipit, được tạo ra bởi quy trình theo sáng chế.

Theo khía cạnh khác, sáng chế đề xuất quy trình để tạo ra etyl este của axit béo nhiều liên kết chưa bão hòa, quy trình này bao gồm bước chuyển hóa este triaxylglyxerol trong lipit chiết tách từ cây, trong đó lipit chiết tách từ cây này bao gồm các axit béo ở dạng este hóa, các axit béo này bao gồm axit oleic, axit palmitic, axit béo $\omega 6$ bao gồm axit linoleic (LA), axit béo $\omega 3$ bao gồm axit α -linolenic (ALA), và axit docosahexaenoic (DHA), và tùy chọn một hoặc nhiều axit stearidonic (SDA), axit eicosapentaenoic (EPA), axit docosapentaenoic (DPA) và axit eicosatetraenoic (ETA), trong đó mức DHA trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 7% đến 20%, bằng cách ấy tạo ra etyl este.

Theo một phương án được ưu tiên, lipit chiết tách có một hoặc nhiều đặc điểm nêu trên.

Theo khía cạnh khác, sáng chế đề xuất quy trình để tạo ra etyl este của axit béo nhiều liên kết chưa bão hòa, quy trình này bao gồm chuyển hóa este triaxylglyxerol trong lipit chiết tách từ cây, trong đó lipit chiết tách từ cây này bao gồm các axit béo được este hóa ở dạng triaxylglyxerol, các axit béo bao gồm axit oleic, axit palmitic, axit béo $\omega 6$ bao gồm axit linoleic (LA), axit béo $\omega 3$ bao gồm axit α -linolenic (ALA) và axit docosahexaenoic (DHA), và một hoặc nhiều axit stearidonic (SDA), axit eicosapentaenoic (EPA), axit docosapentaenoic (DPA) và axit eicosatetraenoic (ETA), trong đó (i) mức DHA trong tổng hàm lượng axit béo của lipit chiết tách là khoảng 3%, khoảng 4%, khoảng 5%, khoảng 6% hoặc nằm trong khoảng từ 7% đến 20%, (ii) mức

axit palmitic trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 2% đến 16%, (iii) mức axit myristic (C14:0) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn 6%, (iv) mức axit oleic trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 1% đến 30% hoặc nằm trong khoảng từ 30% đến 60%, (v) mức axit linoleic (LA) trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 35%, (vi) mức axit α -linolenic (ALA) trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 40%, (vii) mức axit eicosatrienoic (ETra) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn 4%, (viii) mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 25%, (ix) tỷ lệ của toàn bộ axit béo ω 6: toàn bộ axit béo ω 3 trong hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 1,0 và 3,0 hoặc nằm trong khoảng từ 0,1 đến 1, (x) hàm lượng triaxylglyxerol (TAG) trong lipid ít nhất là 70%, và (xi) ít nhất 70% DHA được este hóa ở dạng TAG là ở vị trí *sn*-1 hoặc *sn*-3 trong TAG, bằng cách đó tạo ra etyl este.

Theo một phương án, lipid chiết tách từ cây này có một hoặc nhiều hoặc tất cả các đặc điểm sau:

i) mức axit palmitic trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 2% đến 15%,

ii) mức axit myristic (C14:0) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn 1%,

xxxv) mức axit oleic trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 3% đến 30%, nằm trong khoảng từ 6% đến 30%, từ 1% đến 20%, nằm trong khoảng từ 45% đến 60%, hoặc là khoảng 30%,

xxxvi) mức axit linoleic (LA) trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 4% đến 20%, hoặc nằm trong khoảng từ 4% đến 17%,

xxxvii) mức axit α -linolenic (ALA) trong tổng hàm lượng axit béo của lipid chiết tách nằm trong khoảng từ 7% đến 40%, nằm trong khoảng từ 10% đến 35%, nằm trong khoảng từ 20% đến 35%, hoặc nằm trong khoảng từ 4% đến 16%,

xxxviii) mức axit γ -linolenic (GLA) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn 4%, thấp hơn khoảng 3%, thấp hơn khoảng 2%, thấp hơn khoảng 1%, thấp hơn khoảng 0,5%, nằm trong khoảng từ 0,05% đến 7%, từ 0,05% và 4%, hoặc từ 0,05% đến 3%, hoặc từ 0,05% đến 2%,

xxxix) mức axit stearidonic (SDA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 4%, thấp hơn khoảng 3%, nằm trong khoảng từ 0,05% đến 7%, nằm trong khoảng từ 0,05% đến 4%, nằm trong khoảng từ 0,05% đến 3%, hoặc từ 0,05% đến 2%,

xl) mức axit eicosatetraenoic (ETA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 4%, thấp hơn khoảng 1%, thấp hơn khoảng 0,5%, nằm trong khoảng từ 0,05% đến 5%, nằm trong khoảng từ 0,05% đến 4%, nằm trong khoảng từ 0,05% đến 3%, hoặc nằm trong khoảng từ 0,05% đến 2%,

xli) mức axit eicosatrienoic (ETrA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 2%, thấp hơn khoảng 1%, nằm trong khoảng từ 0,05% đến 4%, từ 0,05% đến 3%, hoặc từ 0,05% đến 2%, hoặc từ 0,05% đến 1%,

xlii) mức axit eicosapentaenoic (EPA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 4%, thấp hơn khoảng 3%, thấp hơn khoảng 2%, từ 0,05% đến 10%, từ 0,05% đến 5%, hoặc từ 0,05% đến 3%, hoặc từ 0,05% đến 2%,

xliii) mức axit docosapentaenoic (DPA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 4%, thấp hơn khoảng 3%, thấp hơn khoảng 2%, từ 0,05% đến 8%, từ 0,05% đến 5%, hoặc từ 0,05% đến 3%, hoặc từ 0,05% đến 2%,

xliv) mức DHA trong tổng hàm lượng axit béo của lipit chiết tách là khoảng 8%, từ 9%, từ 10%, từ 11%, từ 12%, từ 13%, từ 14%, từ 15%, từ 16%, từ 17%, từ 18%, nằm trong khoảng từ 8% đến 20%, nằm trong khoảng từ 10% đến 20%, nằm trong khoảng từ 11% đến 20%, nằm trong khoảng từ 10% đến 16%, hoặc nằm trong khoảng từ 14% đến 20%,

xlvi) lipit chứa axit ω 6-docosapentaenoic ($22:5^{\Delta 4,7,10,13,16}$) trong hàm lượng axit béo của nó,

xlvi) lipit về cơ bản không có axit ω 6-docosapentaenoic ($22:5^{\Delta 4,7,10,13,16}$) trong hàm lượng axit béo của nó,

xlvi) lipit về cơ bản không có SDA, EPA và ETA trong thành phần axit béo của nó,

xlvi) mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 4% đến 20%, hoặc nằm trong khoảng từ 6% đến 20%,

xlvi) mức toàn bộ axit béo một liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 4% đến 35%, nằm trong khoảng từ 8% đến

25%, hoặc từ 8% đến 22%,

l) mức toàn bộ axit béo nhiều liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 20% đến 75%, nằm trong khoảng từ 50% đến 75%, hoặc nằm trong khoảng từ 60% đến 75%,

li) mức toàn bộ axit béo $\omega 6$ trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 35% đến 50%, nằm trong khoảng từ 20% đến 35%, nằm trong khoảng từ 6% đến 20%, thấp hơn 20%, thấp hơn khoảng 16%, thấp hơn khoảng 10%, nằm trong khoảng từ 1% đến 16%, nằm trong khoảng từ 2% đến 10%, hoặc nằm trong khoảng từ 4% đến 10%,

lii) mức axit béo $\omega 6$ mới trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 10%, thấp hơn khoảng 8%, thấp hơn khoảng 6%, thấp hơn 4%, nằm trong khoảng từ 1% đến 20%, nằm trong khoảng từ 1% đến 10%, nằm trong khoảng từ 0,5% đến 8%, hoặc nằm trong khoảng từ 0,5% đến 4%,

liii) mức toàn bộ axit béo $\omega 3$ trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 36% đến 65%, từ 40% đến 60%, nằm trong khoảng từ 20% đến 35%, nằm trong khoảng từ 10% đến 20%, khoảng 25%, khoảng 30%, khoảng 35% hoặc khoảng 40%,

liv) mức axit béo $\omega 3$ mới trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 9% đến 33%, nằm trong khoảng từ 10% đến 20%, nằm trong khoảng từ 20% đến 30%, nằm trong khoảng từ 12% đến 25%, khoảng 13%, khoảng 15%, khoảng 17% hoặc khoảng 20%,

lv) tỷ lệ của toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$ trong hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 0,1 đến 0,5, thấp hơn khoảng 0,50, thấp hơn khoảng 0,40, thấp hơn khoảng 0,30, thấp hơn khoảng 0,20, thấp hơn khoảng 0,15, khoảng 1,0, khoảng 0,1 hoặc khoảng 0,2,

lvi) tỷ lệ của axit béo $\omega 6$ mới: axit béo $\omega 3$ mới trong hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ 1,0 đến 3,0, nằm trong khoảng từ 0,1 đến 1, nằm trong khoảng từ 0,1 đến 0,5, thấp hơn khoảng 0,50, thấp hơn khoảng 0,40, thấp hơn khoảng 0,30, thấp hơn khoảng 0,20, thấp hơn khoảng 0,15, khoảng 0,1, khoảng 0,2 hoặc khoảng 1,0,

lvii) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa axit oleic thành DHA ít nhất là khoảng 10%, ít nhất khoảng 15%, ít nhất khoảng 20%, nằm trong

khoảng từ 10% đến 50%, nằm trong khoảng từ 10% đến 30%, hoặc nằm trong khoảng từ 10% đến 25%,

lviii) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa LA thành DHA ít nhất là khoảng 15%, ít nhất khoảng 20%, ít nhất khoảng 22%, ít nhất khoảng 25%, nằm trong khoảng từ 15% đến 50%, nằm trong khoảng từ 20% đến 40%, hoặc nằm trong khoảng từ 20% đến 30%,

lix) thành phần axit béo của lipit là dựa trên hiệu suất chuyển hóa ALA thành DHA ít nhất là khoảng 17%, ít nhất khoảng 22%, ít nhất khoảng 24%, nằm trong khoảng từ 17% đến 55%, nằm trong khoảng từ 22% đến 35%, hoặc nằm trong khoảng từ 24% đến 35%,

lx) toàn bộ axit béo trong lipit chiết tách có C20:1 với lượng ít hơn 1%,

lxi) hàm lượng triaxylglyxerol (TAG) trong lipit ít nhất là khoảng 80%, ít nhất khoảng 90%, ít nhất 95%, nằm trong khoảng từ 70% đến 99%, hoặc nằm trong khoảng từ 90% đến 99%,

lxii) lipit chứa diaxylglyxerol (DAG),

lxiii) lipit chứa ít hơn khoảng 10%, thấp hơn khoảng 5%, thấp hơn khoảng 1%, hoặc nằm trong khoảng từ 0,001% đến 5%, axit béo tự do (không este hóa) và/hoặc phospholipit, hoặc về cơ bản không có chúng,

lxiv) ít nhất 80%, DHA được este hóa ở dạng TAG là ở vị trí *sn*-1 hoặc *sn*-3 trong TAG,

lxv) kiểu TAG chứa DHA nhiều nhất trong lipit là DHA/18:3/18:3 (TAG 58:12), và

lxvi) lipit chứa tri-DHA TAG (TAG 66:18).

Với khía cạnh nêu trên, theo một phương án, một hoặc nhiều hoặc tất cả các đặc điểm sau đây:

i) lipit là ở dạng dầu, trong đó dầu bao gồm một hoặc nhiều sterol như một hoặc nhiều hoặc tất cả campesterol, Δ 5-stigmasterol, eburicol, β -sitosterol, Δ 5-avenasterol, Δ 7-stigmasterol và Δ 7-avenasterol, và tùy chọn dầu này bao gồm thấp hơn 10mg sterol/g dầu và/hoặc dầu này về cơ bản không có cholesterol,

ii) lipit là ở dạng dầu từ hạt dầu như hạt dầu là hạt dầu *Brassica sp* hoặc hạt cải dầu,

iii) mức DHA trong tổng hàm lượng axit béo của lipit chiết tách từ cây là

khoảng 3%, khoảng 4%, khoảng 5%, khoảng 6%, hoặc nằm trong khoảng từ 7% đến 20%.

Theo khía cạnh khác, sáng chế đề xuất cấu trúc gen khảm chứa theo thứ tự gen thứ nhất, gen thứ hai, gen thứ ba, gen thứ tư, gen thứ năm và gen thứ sáu tất cả được liên kết đồng hóa trị trên phân tử ADN đơn,

trong đó gen thứ nhất, thứ hai và thứ ba được gắn với nhau thành cụm gen thứ nhất và gen thứ tư, thứ năm, và thứ sáu được gắn với nhau thành cụm gen thứ hai,

trong đó mỗi gen chứa gen khởi đầu, vùng mã hóa và gen kết thúc phiên mã và/hoặc vùng polyadenyl hóa sao cho mỗi gen khởi đầu được liên kết điều khiển với vùng mã hóa và gen kết thúc phiên mã và/hoặc vùng polyadenyl hóa,

trong đó mỗi gen khởi đầu độc lập giống hoặc khác với gen khởi đầu khác sao cho phân tử ADN bao gồm ba, bốn, năm hoặc sáu gen khởi đầu khác nhau,

trong đó một hoặc nhiều hoặc tất cả gen khởi đầu là khác loại với vùng mã hóa mà nó được liên kết điều khiển vào đó,

trong đó hướng phiên mã của gen thứ nhất là xa với gen thứ ba và ngược với hướng phiên mã của gen thứ ba,

trong đó hướng phiên mã của gen thứ tư là xa với gen thứ sáu và ngược với hướng phiên mã của gen thứ sáu,

trong đó hướng phiên mã của gen thứ hai là giống với gen thứ nhất hoặc gen thứ ba,

trong đó hướng phiên mã của gen thứ năm là giống với gen thứ tư hoặc gen thứ sáu,

trong đó gen kết thúc phiên mã và/hoặc vùng polyadenyl hóa của gen thứ hai được ngăn cách với gen khởi đầu của gen thứ nhất hoặc gen thứ ba, bất cứ gen nào gần hơn, bởi gen đệm đầu tiên có cỡ từ 0,2 đến 3,0 kilobazơ,

trong đó cụm gen thứ nhất được ngăn cách với cụm gen thứ hai bởi gen đệm thứ hai có cỡ từ 1,0 đến 10,0 kilobazơ, và

trong đó gen kết thúc phiên mã và/hoặc vùng polyadenyl hóa của gen thứ năm được ngăn cách với gen khởi đầu của gen thứ tư hoặc thứ sáu, bất kỳ gen nào gần hơn, bởi gen đệm thứ ba có cỡ từ 0,2 đến 3,0 kilobazơ.

Theo một phương án, phân tử ADN bao gồm gen thứ bảy được ngăn cách với cụm gen thứ nhất hoặc cụm gen thứ hai, bất kỳ gen nào gần hơn, bởi gen đệm có cỡ từ

1,0 đến 10,0 kilobazơ.

Theo phương án khác, phân tử ADN bao gồm hai hoặc nhiều gen kết thúc phiên mã và/hoặc vùng polyadenyl hóa khác nhau.

Theo một phương án rộng hơn, ít nhất một trong các gen đệm chứa vùng gắn cơ chất (MAR).

Theo phương án mở rộng, phân tử ADN bao gồm vùng biên phải và biên trái nằm hai bên sườn các gen và là phân tử T-ADN.

Theo phương án khác, cấu trúc gen là trong tế bào *Agrobacterium* hoặc được sát nhập vào hệ gen của tế bào thực vật.

Theo một phương án được ưu tiên, ít nhất một trong các gen mã hóa gen desaturaza axit béo hoặc gen elongaza axit béo.

Theo phương án khác, cấu trúc gen bao gồm các gen mã hóa bộ các enzym được nêu trong đây, và/hoặc trong đó một hoặc nhiều gen mã hóa enzym được nêu trong đây.

Theo khía cạnh khác, sáng chế đề xuất polynucleotit phân tách và/hoặc ngoại sinh bao gồm:

i) trình tự nucleotit được chọn từ SEQ ID NO bất kỳ: 1 đến 9, 11, 14, 18, 22, 23, 28, 34, 35, 39 hoặc 45, và/hoặc

ii) trình tự nucleotit tương đồng ít nhất 95% hoặc tương đồng 99% với một hoặc nhiều trình tự được nêu trong SEQ ID NO: 1 đến 9, 11, 14, 18, 22, 23, 28, 34, 35, 39 hoặc 45.

Theo phương án được ưu tiên cụ thể, polynucleotit phân tách và/hoặc ngoại sinh bao gồm:

i) trình tự nucleotit của SEQ ID NO: 2, và/hoặc

ii) trình tự nucleotit tương đồng ít nhất 95% hoặc tương đồng 99% với trình tự được nêu trong SEQ ID NO: 2.

Theo khía cạnh khác, sáng chế đề xuất vector hoặc cấu trúc gen bao gồm polynucleotit theo sáng chế và/hoặc cấu trúc gen theo sáng chế.

Theo một phương án, trình tự nucleotit được chọn từ SEQ ID NO bất kỳ: 11, 14, 18, 22, 23, 28, 34, 35, 39 hoặc 45, hoặc trình tự nucleotit tương đồng ít nhất là 95% hoặc tương đồng 99% với một hoặc nhiều trình tự được nêu trong SEQ ID NO: 11, 14, 18, 22, 23, 28, 34, 35, 39 hoặc 45, được liên kết điều khiển với gen khởi đầu.

Theo khía cạnh khác, sáng chế đề xuất tế bào chủ bao gồm polynucleotit ngoại

sinh mã hóa một trong các bộ enzym sau;

- i) ω 3-desaturaza, Δ 6-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 6-elongaza và Δ 5-elongaza,
- ii) Δ 15-desaturaza, Δ 6-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 6-elongaza và Δ 5-elongaza,
- iii) Δ 12-desaturaza, Δ 6-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 6-elongaza và Δ 5-elongaza,
- iv) Δ 12-desaturaza, ω 3-desaturaza hoặc Δ 15-desaturaza, Δ 6-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 6-elongaza và Δ 5-elongaza,
- v) ω 3-desaturaza, Δ 8-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 9-elongaza và Δ 5-elongaza,
- vi) Δ 15-desaturaza, Δ 8-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 9-elongaza và Δ 5-elongaza,
- vii) Δ 12-desaturaza, Δ 8-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 9-elongaza và Δ 5-elongaza, hoặc
- viii) Δ 12-desaturaza, ω 3-desaturaza hoặc Δ 15-desaturaza, Δ 8-desaturaza, Δ 5-desaturaza, Δ 4-desaturaza, Δ 9-elongaza và Δ 5-elongaza,

và trong đó mỗi polynucleotit được liên kết điều khiển với một hoặc nhiều gen khởi đầu có khả năng điều khiển quá trình biểu hiện của polynucleotit nêu trên trong tế bào.

Theo một phương án, tế bào này bao gồm lipid như được xác định trên, hoặc trong đó một hoặc nhiều hoặc tất cả desaturaza hoặc elongaza có một hoặc nhiều đặc điểm như được xác định trên.

Theo khía cạnh khác, sáng chế đề xuất tế bào chủ bao gồm:

- i) polynucleotit ngoại sinh đầu tiên mã hóa Δ 12-desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 10, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 50% với SEQ ID NO: 10, và
- ii) polynucleotit ngoại sinh thứ hai mã hóa ω 3-desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 12, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 50% với SEQ ID NO: 12,

trong đó mỗi polynucleotit được liên kết điều khiển với một hoặc nhiều gen khởi đầu có khả năng điều khiển quá trình biểu hiện của polynucleotit nêu trên trong tế bào.

Theo khía cạnh khác, sáng chế đề xuất tế bào chủ bao gồm một hoặc nhiều

polynucleotit theo sáng chế, cấu trúc gen theo sáng chế, hoặc vector hoặc cấu trúc gen theo sáng chế.

Theo một phương án, tế bào này là tế bào cây, trong một bộ phận cây và/hoặc là tế bào hạt cây trưởng thành.

Theo một phương án, cây hoặc hạt cây lần lượt là cây hạt dầu hoặc hạt dầu.

Sáng chế cũng đề xuất sinh vật biến đổi gen không phải là người bao gồm tế bào theo sáng chế. Tốt hơn là, sinh vật biến đổi gen không phải là người là cây biến đổi gen, tốt hơn là cây hạt dầu hoặc *Arabidopsis thaliana*. Theo một phương án, cây là cây *Brassica*, tốt hơn là *B. napus* hoặc *B. juncea*, hoặc cây không phải là *Arabidopsis thaliana*.

Theo khía cạnh khác, sáng chế đề xuất cây hạt dầu bao gồm:

- a) lipit trong hạt của nó, lipit chứa các axit béo ở dạng este hóa, và
- b) polynucleotit ngoại sinh mã hóa một trong các bộ enzym sau:
 - i) $\Delta 12$ -desaturaza, $\omega 3$ -desaturaza nấm và/hoặc $\Delta 15$ -desaturaza nấm, $\Delta 6$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 6$ -elongaza và $\Delta 5$ -elongaza, hoặc
 - ii) $\Delta 12$ -desaturaza, $\omega 3$ -desaturaza nấm và/hoặc $\Delta 15$ -desaturaza nấm, $\Delta 8$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 9$ -elongaza và $\Delta 5$ -elongaza,

trong đó mỗi polynucleotit được liên kết điều khiển với một hoặc nhiều gen khởi đầu đặc hiệu với hạt có khả năng điều khiển quá trình biểu hiện của polynucleotit này trong hạt đang phát triển của hạt, trong đó các axit béo bao gồm axit oleic, axit palmitic, axit béo $\omega 6$ bao gồm axit linoleic (LA) và axit γ -linolenic (GLA), axit béo $\omega 3$ bao gồm axit α -linolenic (ALA), axit stearidonic (SDA), axit docosapentaenoic (DPA) và axit docosahexaenoic (DHA), và tùy chọn axit eicosapentaenoic (EPA) và/hoặc axit eicosatetraenoic (ETA), và trong đó mức DHA trong tổng hàm lượng axit béo của lipit nằm trong khoảng từ 7% đến 20%.

Ví dụ về cây hạt dầu bao gồm, nhưng không hạn chế ở, *Brassica sp.*, *Gossypium hirsutum*, *Linum usitatissimum*, *Helianthus sp.*, *Carthamus tinctorius*, *Glycine max*, *Zea mays*, *Arabidopsis thaliana*, *Sorghum bicolor*, *Sorghum vulgare*, *Avena sativa*, *Trifolium sp.*, *Elaeagnus guineensis*, *Nicotiana benthamiana*, *Hordeum vulgare*, *Lupinus angustifolius*, *Oryza sativa*, *Oryza glaberrima*, *Camelina sativa*, hoặc *Crambe abyssinica*. Theo một phương án, cây hạt dầu là cây cải dầu, *Glycine max*, *Camelina sativa* hoặc *Arabidopsis thaliana*. Theo một phương án thay thế, cây hạt dầu không phải

là *A. thaliana*.

Theo một phương án, một hoặc nhiều desaturaza có khả năng sử dụng cơ chất axyl-CoA. Theo một phương án được ưu tiên, một hoặc nhiều $\Delta 6$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza và $\Delta 8$ -desaturaza, nếu có, có khả năng sử dụng cơ chất axyl-CoA, tốt hơn là mỗi i) $\Delta 6$ -desaturaza, $\Delta 5$ -desaturaza và $\Delta 4$ -desaturaza hoặc ii) $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza và $\Delta 8$ -desaturaza có khả năng sử dụng cơ chất axyl-CoA. Theo một phương án, $\Delta 12$ -desaturaza và/hoặc $\omega 3$ -desaturaza có khả năng sử dụng cơ chất axyl-CoA. Tốt hơn là, cơ chất axyl-CoA là ALA-CoA, ETA-CoA, DPA-CoA, ETrA-CoA, LA-CoA, GLA-CoA, hoặc ARA-CoA.

Theo một phương án, hạt già, được thu hoạch của cây có hàm lượng DHA ít nhất là khoảng 28mg trong mỗi gam hạt, tốt hơn là ít nhất khoảng 32mg trong mỗi gam hạt, ít nhất khoảng 36mg trong mỗi gam hạt, ít nhất khoảng 40mg trong mỗi gam hạt, tốt hơn nữa là ít nhất khoảng 44mg trong mỗi gam hạt hoặc ít nhất khoảng 48mg trong mỗi gam hạt. Hàm lượng DHA tối đa có thể nằm trong khoảng từ 80 đến 100mg trong mỗi gam hạt, hoặc khoảng 80mg hoặc khoảng 100mg trong mỗi gam hạt.

Theo khía cạnh khác, sáng chế đề xuất cây *Brassica napus*, *B. juncea* hoặc *Camelina sativa* có khả năng tạo ra hạt chứa DHA, trong đó hạt già, được thu hoạch của cây này có hàm lượng DHA ít nhất là khoảng 28mg trong mỗi gam hạt, tốt hơn là ít nhất khoảng 32mg trong mỗi gam hạt, ít nhất khoảng 36mg trong mỗi gam hạt, ít nhất khoảng 40mg trong mỗi gam hạt, tốt hơn nữa là ít nhất khoảng 44mg trong mỗi gam hạt hoặc ít nhất khoảng 48mg trong mỗi gam hạt. Hàm lượng DHA có thể nằm trong khoảng từ 80 đến 100mg trong mỗi gam hạt, hoặc khoảng 80mg hoặc khoảng 100mg trong mỗi gam hạt.

Theo khía cạnh khác, sáng chế đề xuất tế bào cây của cây theo sáng chế bao gồm polynucleotit ngoại sinh.

Sáng chế cũng đề xuất bộ phận cây, tốt hơn là hạt, có một hoặc nhiều đặc điểm sau:

- i) từ cây theo sáng chế,
- ii) bao gồm lipid được nêu trong đây,
- iii) được sử dụng trong quy trình theo sáng chế,
- iv) bao gồm cấu trúc gen theo sáng chế, hoặc
- v) bao gồm bộ polynucleotit ngoại sinh được nêu trong đây.

Theo khía cạnh khác nữa, sáng chế đề xuất hạt *Brassica napus*, *B. juncea* hoặc

Camelina sativa trưởng thành, được thu hoạch bao gồm DHA và hàm lượng ẩm nằm trong khoảng từ 4% đến 15% theo trọng lượng, trong đó hàm lượng DHA của hạt ít nhất khoảng 28mg trong mỗi gam hạt, tốt hơn là ít nhất khoảng 32mg trong mỗi gam hạt, ít nhất khoảng 36mg trong mỗi gam hạt, ít nhất khoảng 40mg trong mỗi gam hạt, tốt hơn nữa là ít nhất khoảng 44mg trong mỗi gam hạt hoặc ít nhất khoảng 48mg trong mỗi gam hạt. Hàm lượng DHA có thể nằm trong khoảng 80mg đến 100mg trong mỗi gam hạt, hoặc khoảng 80mg hoặc khoảng 100mg trong mỗi gam hạt.

Theo một phương án, tế bào theo sáng chế, sinh vật biến đổi gen theo sáng chế, cây hạt dầu theo sáng chế, cây *Brassica napus*, *B. juncea* hoặc *Camelina sativa* theo sáng chế, bộ phận cây theo sáng chế, hoặc hạt theo sáng chế, có thể được sử dụng để tạo ra lipit chiết tách có một hoặc nhiều hoặc tất cả các đặc điểm được nêu trong đây.

Theo khía cạnh khác nữa, sáng chế đề xuất phương pháp tạo ra tế bào theo sáng chế, phương pháp này bao gồm

a) đưa vào tế bào này, tốt hơn là tế bào không có khả năng tổng hợp LC-PUFA, cấu trúc gen theo sáng chế, polynucleotit phân tách và/hoặc ngoại sinh theo sáng chế, vector hoặc cấu trúc gen theo sáng chế, một hoặc nhiều tổ hợp của polynucleotit ngoại sinh nêu trong đây,

b) tùy chọn, biểu hiện gen hoặc (các) polynucleotit trong tế bào này;

c) tùy chọn, phân tích thành phần axit béo của tế bào, và

d) tùy chọn, lựa chọn tế bào biểu hiện gen hoặc (các) polynucleotit.

Theo một phương án, lipit trong tế bào này có một hoặc nhiều đặc điểm nêu trong đây.

Theo phương án khác, cấu trúc gen, polynucleotit phân tách và/hoặc ngoại sinh, vector, cấu trúc gen hoặc sự kết hợp của các polynucleotit ngoại sinh, được sát nhập ổn định vào hệ gen của tế bào.

Theo phương án mở rộng, tế bào là tế bào cây, và phương pháp này còn bao gồm bước phục hồi cây biến nạp từ tế bào ở bước a).

Theo phương án khác, gen và/hoặc polynucleotit ngoại sinh được biểu hiện tạm thời trong tế bào này.

Sáng chế cũng đề xuất tế bào được tạo ra bởi phương pháp theo sáng chế.

Theo khía cạnh khác, sáng chế đề xuất phương pháp tạo ra hạt, phương pháp này

bao gồm các bước

a) trồng cây theo sáng chế, hoặc cây cung cấp bộ phận được nêu trong đây, tốt hơn là trên cánh đồng nằm trong quần thể ít nhất 1000 cây này hoặc trong diện tích ít nhất 1 héc ta được trồng với mật độ trồng cây chuẩn,

b) thu hoạch hạt từ cây hoặc các cây, và

c) tùy chọn, chiết lipid từ hạt này, tốt hơn là để tạo ra dầu với hiệu suất DHA tổng cộng ít nhất là 60kg DHA/héc ta.

Theo một phương án, cây, tế bào cây, bộ phận cây hoặc hạt theo sáng chế có một hoặc nhiều đặc điểm sau

i) dầu như được nêu trong đây,

ii) bộ phận cây hoặc hạt có thể được sử dụng trong quy trình theo sáng chế,

iii) polynucleotit ngoại sinh nằm trong cấu trúc gen theo sáng chế,

iv) polynucleotit ngoại sinh bao gồm polynucleotit ngoại sinh theo sáng chế,

v) tế bào cây là tế bào theo sáng chế, và

vi) hạt được tạo ra theo phương pháp theo sáng chế.

Theo khía cạnh khác, sáng chế đề xuất phương pháp tạo ra một hoặc nhiều desaturaza axit béo và/hoặc elongaza axit béo, hoặc một hoặc nhiều desaturaza axit béo và một hoặc nhiều elongaza axit béo, phương pháp này bao gồm biểu hiện trong tế bào hoặc hệ biểu hiện không phải tế bào cấu trúc gen theo sáng chế, polynucleotit phân tách và/hoặc ngoại sinh theo sáng chế, vector hoặc cấu trúc gen theo sáng chế, một hoặc nhiều tổ hợp của polynucleotit ngoại sinh nêu trong đây, tốt hơn là trong hạt dầu đang phát triển trong cây hạt dầu trên cánh đồng.

Theo khía cạnh khác, sáng chế đề xuất lipid, hoặc dầu, được tạo ra bằng, hoặc thu được từ, sử dụng quy trình theo sáng chế, tế bào theo sáng chế, sinh vật biến đổi gen theo sáng chế, cây hạt dầu theo sáng chế, cây *Brassica napus*, *B. juncea* hoặc *Camelina sativa* theo sáng chế, bộ phận cây theo sáng chế, hạt theo sáng chế, hoặc cây, tế bào cây, bộ phận cây hoặc hạt theo sáng chế.

Theo một phương án, lipid hoặc dầu thu được bằng cách chiết dầu từ hạt dầu. Ví dụ về dầu từ hạt dầu bao gồm, nhưng không hạn chế ở, dầu cải (*Brassica napus*, *Brassica rapa* ssp.), dầu mù tạc (*Brassica juncea*), dầu họ Brassica khác, dầu hướng dương (*Helianthus annuus*), dầu hạt lanh (*Linum usitatissimum*), dầu đậu nành (*Glycine max*), dầu rum (*Carthamus tinctorius*), dầu ngô (*Zea mays*), dầu cây thuốc lá (*Nicotiana*

tabacum), dầu lạc (*Arachis hypogaea*), dầu cọ, dầu hạt bông (*Gossypium hirsutum*), dầu dừa (*Cocos nucifera*), dầu lê tàu (*Persea americana*), dầu ôliu (*Olea europaea*), dầu cây điều (*Anacardium occidentale*), dầu mácca (*Macadamia integrifolia*), dầu quả hạnh (*Prunus amygdalus*) hoặc dầu hạt *Arabidopsis* (*Arabidopsis thaliana*).

Theo khía cạnh khác, sáng chế đề xuất axit béo được tạo ra bởi, hoặc thu được từ, bằng cách sử dụng quy trình theo sáng chế, tế bào theo sáng chế, sinh vật biến đổi gen theo sáng chế, cây hạt dầu theo sáng chế, cây *Brassica napus*, *B. juncea* hoặc *Camelina sativa* theo sáng chế, bộ phận cây theo sáng chế, hạt theo sáng chế, hoặc cây, tế bào cây, bộ phận cây hoặc hạt theo sáng chế. Tốt hơn là, axit béo là DHA. Axit béo có thể là ở dạng hỗn hợp các axit béo có thành phần axit béo như mô tả trong đây. Theo một phương án, axit béo không được este hóa.

Sáng chế cũng đề xuất bột hạt thu được từ hạt theo sáng chế. Bột hạt được ưu tiên bao gồm, nhưng không hạn chế ở, bột hạt *Brassica napus*, *B. juncea*, *Camelina sativa* hoặc *Glycine max*. Theo một phương án, bột hạt bao gồm polynucleotit ngoại sinh và/hoặc cấu trúc gen được nêu trong đây.

Theo khía cạnh khác, sáng chế đề xuất chế phẩm chứa một hoặc nhiều lipid hoặc dầu theo sáng chế, axit béo theo sáng chế, cấu trúc gen theo sáng chế, polynucleotit phân tách và/hoặc ngoại sinh theo sáng chế, vectơ hoặc cấu trúc gen theo sáng chế, tế bào theo sáng chế, sinh vật biến đổi gen theo sáng chế, cây hạt dầu theo sáng chế, cây *Brassica napus*, *B. juncea* hoặc *Camelina sativa* theo sáng chế, bộ phận cây theo sáng chế, hạt theo sáng chế, cây, tế bào cây, bộ phận cây hoặc hạt theo sáng chế, hoặc bột hạt theo sáng chế. Theo các phương án, sáng chế này bao gồm chất mang thích hợp cho ứng dụng dược phẩm, thực phẩm hoặc nông nghiệp, hợp chất xử lý hạt, phân bón, thực phẩm hoặc thành phần thức ăn khác, hoặc protein hoặc vitamin bổ sung.

Sáng chế cũng đề xuất sản phẩm dinh dưỡng, mỹ phẩm hoặc hóa chất bao gồm một hoặc nhiều lipid hoặc dầu theo sáng chế, axit béo theo sáng chế, cấu trúc gen theo sáng chế, polynucleotit phân tách và/hoặc ngoại sinh theo sáng chế, vectơ hoặc cấu trúc gen theo sáng chế, tế bào theo sáng chế, sinh vật biến đổi gen theo sáng chế, cây hạt dầu theo sáng chế, cây *Brassica napus*, *B. juncea* hoặc *Camelina sativa* theo sáng chế, bộ phận cây theo sáng chế, hạt theo sáng chế, cây, tế bào cây, bộ phận cây hoặc hạt theo sáng chế, bột hạt theo sáng chế, hoặc chế phẩm theo sáng chế.

Theo khía cạnh khác, sáng chế đề xuất phương pháp tạo ra sản phẩm dinh dưỡng,

phương pháp này bao gồm trộn một hoặc nhiều lipit hoặc dầu theo sáng chế, axit béo theo sáng chế, cấu trúc gen theo sáng chế, polynucleotit phân tách và/hoặc ngoại sinh theo sáng chế, vector hoặc cấu trúc gen theo sáng chế, tế bào theo sáng chế, sinh vật biến đổi gen theo sáng chế, cây hạt dầu theo sáng chế, cây *Brassica napus*, *B. juncea* hoặc *Camelina sativa* theo sáng chế, bộ phận cây theo sáng chế, hạt theo sáng chế, cây, tế bào cây, bộ phận cây hoặc hạt theo sáng chế, bột hạt theo sáng chế, hoặc chế phẩm theo sáng chế, với ít nhất một thành phần thức ăn khác.

Theo khía cạnh khác, sáng chế đề xuất phương pháp điều trị hoặc phòng ngừa tình trạng bệnh mà PUFA sẽ có lợi, phương pháp này bao gồm sử dụng cho đối tượng một hoặc nhiều lipit hoặc dầu theo sáng chế, axit béo theo sáng chế, cấu trúc gen theo sáng chế, polynucleotit phân tách và/hoặc ngoại sinh theo sáng chế, vector hoặc cấu trúc gen theo sáng chế, tế bào theo sáng chế, sinh vật biến đổi gen theo sáng chế, cây hạt dầu theo sáng chế, cây *Brassica napus*, *B. juncea* hoặc *Camelina sativa* theo sáng chế, bộ phận cây theo sáng chế, hạt theo sáng chế, cây, tế bào cây, bộ phận cây hoặc hạt theo sáng chế, bột hạt theo sáng chế, chế phẩm theo sáng chế, hoặc sản phẩm dinh dưỡng theo sáng chế.

Ví dụ về tình trạng bệnh mà PUFA sẽ có lợi bao gồm, nhưng không hạn chế ở, chứng loạn nhịp tim, bệnh nguyên bào mạch, bệnh viêm, bệnh hen, bệnh vẩy nến, chứng loãng xương, bệnh sỏi thận, AIDS, bệnh xơ cứng rải rác, viêm khớp dạng thấp, bệnh Crohn, bệnh tâm thần phân liệt, bệnh ung thư, hội chứng ngộ độc rượu ở thai nhi, rối loạn tăng động giảm chú ý, xơ nang, axit phenylpyruvic-niệu, trầm cảm đơn cực, hội chứng xung-hấn, bệnh loạn dưỡng chất trắng thượng thận, bệnh mạch vành tim, bệnh cao huyết áp, bệnh đái tháo đường, bệnh béo phì, bệnh Alzheimer, bệnh tắc phổi mãn tính, bệnh viêm ruột kết mạn loét, bệnh tái phát hẹp van tim sau bệnh nguyên bào mạch, bệnh chàm, bệnh huyết áp cao, bệnh kết tụ tiểu cầu, bệnh chảy máu dạ dày, bệnh lạc nội mạc tử cung, hội chứng tiền kinh nguyệt, hội chứng mệt mỏi kinh niên, suy nhược mãn tính sau khi nhiễm virus hoặc bệnh mất.

Sáng chế cũng đề xuất việc sử dụng một hoặc nhiều lipit hoặc dầu theo sáng chế, axit béo theo sáng chế, cấu trúc gen theo sáng chế, polynucleotit phân tách và/hoặc ngoại sinh theo sáng chế, vector hoặc cấu trúc gen theo sáng chế, tế bào theo sáng chế, sinh vật biến đổi gen theo sáng chế, cây hạt dầu theo sáng chế, cây *Brassica napus*, *B. juncea* hoặc *Camelina sativa* theo sáng chế, bộ phận cây theo sáng chế, hạt theo sáng

ché, cây, tế bào cây, bộ phận cây hoặc hạt theo sáng chế, bột hạt theo sáng chế, chế phẩm theo sáng chế, hoặc sản phẩm dinh dưỡng theo sáng chế để sản xuất thuốc để điều trị hoặc phòng tình trạng bệnh lý mà PUFA sẽ có lợi. Việc sản xuất thuốc có thể bao gồm trộn dầu theo sáng chế với chất mang được dụng, để điều trị tình trạng bệnh như mô tả trong đây. Phương pháp này có thể bao gồm đầu tiên là tinh chế dầu và/hoặc chuyển hóa este, và/hoặc cắt phân đoạn dầu để làm tăng mức DHA. Theo một phương án cụ thể, phương pháp này bao gồm việc xử lý lipit hoặc dầu như dầu cải để chuyển hóa axit béo trong dầu thành alkyl este như metyl hoặc etyl este. Việc xử lý tiếp như cắt phân đoạn hoặc có thể chưng cất để làm giàu DHA trong lipit hoặc dầu. Theo một phương án được ưu tiên, thuốc này bao gồm etyl este của DHA. Theo một phương án ưu tiên hơn nữa, mức etyl este của DHA trong thuốc nằm trong khoảng từ 30% đến 50%. Thuốc này còn có thể bao gồm etyl este của EPA, như từ 30% đến 50% tổng hàm lượng axit béo trong thuốc này. Thuốc này thích hợp để sử dụng cho đối tượng người hoặc động vật để điều trị các tình trạng bệnh lý như mô tả trong đây.

Theo khía cạnh khác, sáng chế đề xuất phương pháp kinh doanh hạt, phương pháp này bao gồm việc thu lấy theo sáng chế, và kinh doanh hạt thu được vì mục đích mục đích doanh thu.

Theo một phương án, việc thu lấy hạt bao gồm việc trồng cây theo sáng chế và/hoặc thu hoạch hạt từ cây.

Theo phương án khác, việc thu lấy hạt còn bao gồm việc đưa hạt vào vật chứa và/hoặc bảo quản hạt.

Theo phương án mở rộng, việc thu lấy hạt còn bao gồm vận chuyển hạt đến một địa điểm khác.

Theo một phương án khác nữa, phương pháp này còn bao gồm vận chuyển hạt đến một địa điểm khác sau khi hạt được bán.

Theo phương án mở rộng, việc mua bán được thực hiện bằng phương tiện điện tử như máy tính.

Theo khía cạnh khác nữa, sáng chế đề xuất quy trình đóng hộp hạt bao gồm:

a) xén, vun thành luống và/hoặc gạt phần trên mặt đất của cây chứa hạt theo sáng chế,

b) đập và/hoặc sàng các bộ phận của cây để tách riêng hạt khỏi bộ phận cây còn lại, và

c) sàng và/hoặc phân loại được tách riêng trong bước b), và đóng hạt đã được sàng và/hoặc phân loại vào các hộp, theo đó tạo ra các hộp chứa hạt.

Theo một phương án, khi thích hợp, lipit hoặc dầu, tốt hơn là dầu hạt, theo sáng chế hoặc hữu dụng theo sáng chế, có độ béo được nêu trong Bảng trong phần Ví dụ, như hạt 14 của Bảng 16.

Phương án bất kỳ ở đây sẽ được đưa vào áp dụng có sửa đổi thích hợp thành phương án khác bất kỳ trừ khi có quy định khác.

Sáng chế không bị giới hạn phạm vi ở các phương án cụ thể được mô tả trong đây, các phương án được chỉ nhằm mục đích minh họa. Các sản phẩm, chế phẩm, và phương pháp tương đương về mặt chức năng rõ ràng sẽ nằm trong phạm vi sáng chế, như mô tả trong đây.

Trong toàn bộ bản mô tả, trừ khi được ghi chú khác, khi nhắc đến một bước riêng lẻ, chế phẩm chứa vật chất, nhóm các bước hoặc nhóm các chế phẩm chứa vật chất sẽ bao gồm cả nghĩa số ít và số nhiều (tức là một hoặc nhiều) bước, chế phẩm chứa vật chất, nhóm các bước hoặc nhóm các chế phẩm chứa vật chất.

Sau đây sáng chế sẽ được mô tả bởi các ví dụ không giới hạn phạm vi sau đây và tham khảo các hình vẽ kèm theo.

Mô tả vắn tắt các hình vẽ

Fig.1. Con đường sinh tổng hợp DHA ưa khí.

Fig.2. Bản đồ vùng lòng T-ADN ở giữa biên phải và biên trái của pJP3416-GA7. RB ký hiệu là biên phải; LB, biên trái; TER, gen kết thúc phiên mã/ vùng polyadenyl hóa; PRO, gen khởi đầu; Vùng mã hóa được thể hiện bên trên mũi tên, gen khởi đầu và gen kết thúc bên dưới. Micpu- $\Delta 6D$, *Micromonas pusilla* $\Delta 6$ -desaturaza; Pyrco- $\Delta 6E$, *Pyramimonas cordata* $\Delta 6$ -elongaza; Pavsa- $\Delta 5D$, $\Delta 5$ -desaturaza *Pavlova salina*; Picpa- $\omega 3D$, $\omega 3$ -desaturaza *Pichia pastoris*; Pavsa- $\Delta 4D$, $\Delta 4$ -desaturaza *P. salina*; Lack1- $\Delta 12D$, $\Delta 12$ -desaturaza *Lachancea kluyveri*; Pyrco- $\Delta 5E$, $\Delta 5$ -elongaza *Pyramimonas cordata*. NOS để chỉ gen kết thúc phiên mã/vùng polyadenyl hóa *Agrobacterium tumefaciens* nopal synthaza; FP1, gen khởi đầu cắt napin *Brassica napus*; FAE1, gen khởi đầu *Arabidopsis thaliana* FAE1; Lectin, gen kết thúc phiên mã/vùng polyadenyl hóa *Glycine max* lectin; Cnl1 và Cnl2 để chỉ gen khởi đầu hoặc gen kết thúc *Linum usitatissimum* conlinin1 hoặc conlinin2. MAR để chỉ vùng gắn cơ chất Rb7 từ

Nicotiana tabacum.

Fig.3. Bản đồ vùng lồng T-ADN giữa biên phải và biên trái của pJP3404. Các ký hiệu là như trong Fig.2.

Fig.4. Bản đồ vùng lồng giữa biên phải và biên trái của pJP3367. Các ký hiệu là như trong Fig.2.

Fig.5. Mức DHA là phần trăm của toàn bộ axit béo trong lipid hạt từ nhiều hạt *Arabidopsis thaliana* biến đổi gen độc lập ở cả thể hệ T₂ và T₃. Các cây T₂ trong ngoặc được chuyển sang T₃. Các cây đối chứng *A. thaliana* Columbia và *fad2* đột biến cũng được thể hiện.

Fig.6. Hàm lượng dầu (trọng lượng/trọng lượng) so với hàm lượng DHA, là phần trăm phần trăm toàn bộ axit béo của lipid từ hạt *Arabidopsis thaliana* biến đổi gen.

Fig.7. Gel RT-PCR đặc trưng thể hiện sự biểu hiện kém của gen $\Delta 6$ -desaturaza tương ứng với các gen biến đổi khác trong T-ADN của mầm *B. napus* được gây biến nạp bằng pJP3416-GA7. Các vệt từ bên trái thể hiện sản phẩm RT-PCR: 1, gen đánh dấu kích thước ADN; vệt 2, $\Delta 12$ desaturaza; vệt 3, $\omega 3$ -desaturaza; vệt 4, $\Delta 6$ -desaturaza (biểu hiện kém); vệt 5, $\Delta 6$ -elongaza; vệt 6, $\Delta 5$ -desaturaza; vệt 7, $\Delta 5$ -elongaza; vệt 8, $\Delta 4$ -desaturaza.

Fig.8. Phần trăm của ALA được vẽ đồ thị với phần trăm của axit oleic, mỗi đường là phần trăm của toàn bộ axit béo trong lipid thu được từ mầm sinh dưỡng *Brassica napus* biến đổi gen 35S:LEC2.

Fig.9. Phân bố vị trí phân tích bởi NMR trên A) dầu cá ngừ và, B) dầu hạt *Arabidopsis* biến đổi gen DHA. Các đỉnh ký hiệu 'DHA-alpha' thể hiện lượng DHA có ở vị trí *sn*-1 và vị trí *sn*-3 của TAG (không có đối chứng vị trí, nó sẽ bằng 66% tổng DHA) trong khi các đỉnh đánh dấu 'DHA-beta' thể hiện lượng DHA ở vị trí *sn*-2 của TAG (không có đối chứng nó sẽ bằng 33% DHA).

Fig.10. Phân tích LC-MS của dạng triaxylglyxerol chứa DHA trong hạt *A. thaliana* biến đổi gen đang phát triển (màu xám) và già (màu đen). Con số đứng sau DHA để chỉ tổng số nguyên tử carbon và tổng số liên kết đôi trong hai axit béo khác. Do đó, DHA/34:1 cũng có thể được ký hiệu là TAG 56:7, v.v..

Fig.11. Bản đồ vùng lồng T-ADN ở giữa biên phải và biên trái của pORE04+11ABGBEC_Đậu đũa_EPA_đoạn lồng. Ký hiệu là như trong Fig.2; SSU, gen khởi đầu của tiểu đơn vị nhỏ *Arabidopsis thaliana rubisco*.

Fig.12. Bản đồ vector kép pJP3364 thể hiện vị trí giới hạn *NotI* trong đó $\Delta 12$ -desaturaza nghiên cứu được tách dòng vào đó.

Fig.13. Biểu đồ dạng hộp thu được bằng SigmaPlot thể hiện phần trăm của axit béo 20:4 ω 6 (ARA) trong lipid hạt của nhóm hạt *Arabidopsis* T2 được gây biến nạp bằng pFN045-pFN050. Đường biên của mỗi hộp gần nhất với mức 0 chỉ mức phân vị phần trăm 25, đường bên trong mỗi hộp đánh dấu mức trung bình, và đường biên của mỗi hộp xa với mức 0 nhất chỉ mức phân vị phần trăm 75. Thanh báo lỗi thể hiện trên và dưới mỗi hộp chỉ ra mức phân vị phần trăm 90 và 10.

Fig.14. Mức trung bình của ARA ở dạng phần trăm tổng hàm lượng axit béo trong lipid hạt của hạt *Arabidopsis* T2 được gây biến nạp bằng pFN045-pFN050.

Fig.15. Biểu đồ dạng hộp thể hiện phần trăm axit béo 20:2 ω 6 (EDA) trong lipid hạt của nhóm hạt *Arabidopsis* T2 được gây biến nạp bằng pFN045-pFN050. Biểu đồ dạng hộp thể hiện giá trị như mô tả trong Fig.13.

Fig.16. Biểu đồ dạng hộp thể hiện phần trăm của ARA trong lipid hạt của nhóm hạt *Arabidopsis* T4 được gây biến nạp bằng pFN045-pFN050. Biểu đồ dạng hộp thể hiện giá trị như mô tả trong Fig.13.

Fig.17. Mức trung bình của ARA ở dạng phần trăm tổng hàm lượng axit béo trong lipid hạt của nhóm hạt *Arabidopsis* T4 được gây biến nạp bằng pFN045-pFN050.

Fig.18. Biểu đồ dạng hộp thể hiện phần trăm EDA trong lipid hạt của nhóm hạt *Arabidopsis* T4 được gây biến nạp bằng pFN045-pFN050. Biểu đồ dạng hộp thể hiện giá trị như mô tả trong Fig.13.

Fig.19. (A) Cấu trúc sterol thực vật cơ bản có vòng và chuỗi bên đánh số. (B) Cấu trúc hóa học của một số sterol thực vật.

Fig.20. Cây hệ thống sinh học của các LPAAT đã biết.

Fig.21. Các enzym trao đổi axyl khác nhau sẽ vận chuyển các axit béo giữa PC, nhóm CoA, và nhóm TAG. Được sửa đổi từ Singh et al. (2005).

Ký hiệu danh sách trình tự

SEQ ID NO: 1 – trình tự nucleotit pJP3416-GA7.

SEQ ID NO: 2 – trình tự nucleotit pGA7- mod_B.

SEQ ID NO: 3 – trình tự nucleotit pGA7- mod_C.

SEQ ID NO: 4 – trình tự nucleotit pGA7- mod_D.

SEQ ID NO: 5 – trình tự nucleotit pGA7- mod_E.

SEQ ID NO: 6 – trình tự nucleotit pGA7- mod_F.

SEQ ID NO: 7 – trình tự nucleotit pGA7- mod_G.

SEQ ID NO: 8 - pORE04+11ABGBEC_Đậu đũa_EPA_trình tự nucleotit lồng.

SEQ ID NO: 9 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 12$ desaturaza *Lachancea kluyveri* ở thực vật.

SEQ ID NO: 10 - *Lachancea kluyveri* $\Delta 12$ -desaturaza.

SEQ ID NO: 11 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\omega 3$ desaturaza *Pichia pastoris* ở thực vật.

SEQ ID NO: 12 - $\omega 3$ desaturaza *Pichia pastoris*.

SEQ ID NO: 13 - Khung đọc mở mã hóa $\Delta 6$ -desaturaza *Micromonas pusilla*.

SEQ ID NO: 14 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 6$ -desaturaza *Micromonas pusilla* ở thực vật (phiên bản 1).

SEQ ID NO: 15 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 6$ -desaturaza *Micromonas pusilla* ở thực vật (phiên bản 2).

SEQ ID NO: 16 - $\Delta 6$ -desaturaza *Micromonas pusilla*.

SEQ ID NO: 17 - Khung đọc mở mã hóa $\Delta 6$ -desaturaza *Ostreococcus lucimarinus*.

SEQ ID NO: 18 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 6$ -desaturaza *Ostreococcus lucimarinus* ở thực vật.

SEQ ID NO: 19 - $\Delta 6$ -desaturaza *Ostreococcus lucimarinus*.

SEQ ID NO: 20 - $\Delta 6$ -desaturaza *Ostreococcus tauri*.

SEQ ID NO: 21 - Khung đọc mở mã hóa *Pyramimonas cordata* $\Delta 6$ -elongaza.

SEQ ID NO: 22 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 6$ -elongaza *Pyramimonas cordata* ở thực vật (bị cắt cụt ở đầu 3' và mã hóa elongaza chức năng) (phiên bản 1).

SEQ ID NO: 23 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 6$ -elongaza *Pyramimonas cordata* ở thực vật (bị cắt cụt ở đầu 3' và mã hóa elongaza chức năng) (phiên bản 2).

SEQ ID NO: 24 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 6$ -elongaza *Pyramimonas cordata* ở thực vật (bị cắt cụt ở đầu 3' và mã hóa elongaza chức năng) (phiên bản 3).

SEQ ID NO: 25 - $\Delta 6$ -elongaza *Pyramimonas cordata*.

SEQ ID NO: 26 – $\Delta 6$ -elongaza *Pyramimonas cordata* bị cắt cụt.

SEQ ID NO: 27 - Khung đọc mở mã hóa *Pavlova salina* $\Delta 5$ -desaturaza.

SEQ ID NO: 28 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 5$ -desaturaza *Pavlova salina* ở thực vật (phiên bản 1).

SEQ ID NO: 29 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 5$ -desaturaza *Pavlova salina* ở thực vật (phiên bản 2).

SEQ ID NO: 30 - $\Delta 5$ -desaturaza *Pavlova salina*.

SEQ ID NO: 31 - Khung đọc mở mã hóa $\Delta 5$ -desaturaza *Pyramimonas cordata*.

SEQ ID NO: 32 - $\Delta 5$ -desaturaza *Pyramimonas cordata*.

SEQ ID NO: 33 - Khung đọc mở mã hóa $\Delta 5$ -elongaza *Pyramimonas cordata*.

SEQ ID NO: 34 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 5$ -elongaza *Pyramimonas cordata* ở thực vật (phiên bản 1).

SEQ ID NO: 35 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 5$ -elongaza *Pyramimonas cordata* ở thực vật (phiên bản 2).

SEQ ID NO: 36 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 5$ -elongaza *Pyramimonas cordata* ở thực vật (phiên bản 3).

SEQ ID NO: 37 - $\Delta 5$ -elongaza *Pyramimonas cordata*.

SEQ ID NO: 38 - Khung đọc mở mã hóa $\Delta 4$ -desaturaza *Pavlova salina*.

SEQ ID NO: 39 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 4$ -desaturaza *Pavlova salina* ở thực vật (phiên bản 1).

SEQ ID NO: 40 - Khung đọc mở tối ưu hóa theo codon để biểu hiện $\Delta 4$ -desaturaza *Pavlova salina* ở thực vật (phiên bản 2).

SEQ ID NO: 41 - $\Delta 4$ -desaturaza *Pavlova salina*.

SEQ ID NO: 42 - Khung đọc mở mã hóa $\Delta 9$ -elongaza *Isochrysis galbana*.

SEQ ID NO: 43 - $\Delta 9$ -elongaza *Isochrysis galbana*.

SEQ ID NO: 44 - Khung đọc mở mã hóa $\Delta 9$ -elongaza *Emiliana huxleyi* CCMP1516.

SEQ ID NO: 45 - Khung đọc mở tối ưu hóa theo codon để biểu hiện *Emiliana huxleyi* $\Delta 9$ -elongaza ở thực vật.

SEQ ID NO: 46 - $\Delta 9$ -elongaza *Emiliana huxleyi* CCMP1516.

SEQ ID NO: 47 - Khung đọc mở mã hóa $\Delta 9$ -elongaza *Pavlova pinguis*.

SEQ ID NO: 48 - $\Delta 9$ -elongaza *Pavlova pinguis*.

SEQ ID NO: 49 - Khung đọc mở mã hóa $\Delta 9$ -elongaza *Pavlova salina*.

SEQ ID NO: 50 - $\Delta 9$ -elongaza *Pavlova salina*.

SEQ ID NO: 51 - Khung đọc mở mã hóa $\Delta 8$ -desaturaza *Pavlova salina*.

SEQ ID NO: 52 - $\Delta 8$ -desaturaza *Pavlova salina*.

SEQ ID NO: 53 – gen ức chế virus P19.

SEQ ID NO: 54 – Gen ức chế virus V2.

SEQ ID NO: 55 – Gen ức chế virus P38.

SEQ ID NO: 56 – Gen ức chế virus Pe-P0.

SEQ ID NO: 57 – Gen ức chế virus RPV-P0.

SEQ ID NO: 58 – Khung đọc mở mã hóa gen ức chế virus P19.

SEQ ID NO: 59 – Khung đọc mở mã hóa gen ức chế virus V2.

SEQ ID NO: 60 – Khung đọc mở mã hóa gen ức chế virus P38.

SEQ ID NO: 61 – Khung đọc mở mã hóa gen ức chế virus Pe-P0.

SEQ ID NO: 62 – Khung đọc mở mã hóa gen ức chế virus RPV-P0.

SEQ ID NO: 63 - LPAAT2 *Arabidopsis thaliana*.

SEQ ID NO: 64 – LPAAT *Limnanthes alba*.

SEQ ID NO: 65 – LPAAT *Saccharomyces cerevisiae*.

SEQ ID NO: 66 – LPAAT *Micromonas pusilla*.

SEQ ID NO: 67 – LPAAT *Mortierella alpina*. SEQ ID NO: 68 – LPAAT

Brassica napus.

SEQ ID NO: 69 – LPAAT *Brassica napus*.

SEQ ID NO: 70 - $\omega 3$ desaturaza *Phytophthora infestans*. SEQ ID NO: 71 - $\omega 3$ desaturaza *Thalassiosira pseudonana*.

SEQ ID NO: 72 - $\omega 3$ desaturaza *Pythium irregulare*.

Mô tả chi tiết sáng chế

Các kỹ thuật tổng quát và định nghĩa

Trừ khi có quy định khác, tất cả các thuật ngữ kỹ thuật và khoa học sử dụng ở đây sẽ được hiểu theo nghĩa như được hiểu thông thường với các chuyên gia trong lĩnh vực này (ví dụ, trong nuôi cấy tế bào, di truyền phân tử, tổng hợp axit béo, cây biến đổi

gen, hóa học protein, và hóa sinh).

Trừ khi có quy định khác, kỹ thuật protein tái tổ hợp, nuôi cấy tế bào, và miễn dịch học sử dụng trong sáng chế là các quy trình chuẩn, đã biết rõ với các chuyên gia trong lĩnh vực này. Các kỹ thuật này được mô tả và giải thích trong các nguồn tài liệu như, J. Perbal, *A Practical Guide to Molecular Cloning*, John Wiley and Sons (1984), J. Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory Press (1989), T.A. Brown (người biên tập), *Essential Molecular Biology: A Practical Approach*, Volumes 1 and 2, IRL Press (1991), D.M. Glover and B.D. Hames (người biên tập), *DNA Cloning: A Practical Approach*, Volumes 1-4, IRL Press (1995 và 1996), F.M. Ausubel et al. (người biên tập), *Current Protocols in Molecular Biology*, Greene Pub. Associates and Wiley-Interscience (1988, bao gồm tất cả các tài liệu cập nhật cho đến nay), Ed Harlow và David Lane (người biên tập), *Antibodies: A Laboratory Manual*, Cold Spring Harbour Laboratory, (1988), và J.E. Coligan et al. (người biên tập), *Current Protocols in Immunology*, John Wiley & Sons (bao gồm tất cả các tài liệu cập nhật cho đến nay).

Thuật ngữ “và/hoặc”, ví dụ, “X và/hoặc Y” sẽ được hiểu có nghĩa là hoặc “X và Y” hoặc “X hoặc Y” và sẽ được hiểu theo cả nghĩa “và” và “hoặc”.

Như sử dụng trong đây, thuật ngữ “khoảng”, trừ khi có quy định khác, là để chỉ +/- 10%, tốt hơn nữa là +/- 5%, tốt hơn nữa là +/- 1% giá trị được nêu.

Trong suốt bản mô tả, từ "bao gồm", hoặc các từ tương tự như "chứa", sẽ được hiểu là ngụ ý sự gồm có một yếu tố, số nguyên hoặc bước đã nêu, hoặc nhóm các yếu tố, số nguyên hoặc bước, nhưng không có nghĩa loại trừ yếu tố, số nguyên hoặc bước bất kỳ khác, hoặc nhóm các yếu tố, số nguyên hoặc bước.

Định nghĩa

Như sử dụng trong đây, thuật ngữ "lipit chiết tách từ cây" và “lipit thực vật được phân tách” là để chỉ chế phẩm lipit đã được chiết tách từ, ví dụ bằng cách ép, cây hoặc bộ phận của nó như hạt. Lipit chiết tách có thể là chế phẩm tương đối thô thu được bằng cách, ví dụ, ép hạt cây, hoặc chế phẩm tinh khiết hơn trong đó phần lớn, nếu không phải là tất cả, một hoặc nhiều hoặc mỗi trong số nước, axit nucleic, protein và carbohydrat có nguồn gốc từ nguyên liệu cây đã được loại bỏ. Ví dụ về các phương pháp tinh chế được mô tả sau đây. Theo một phương án, lipit thực vật được phân tách hoặc được chiết tách bao gồm lipit với lượng ít nhất khoảng 60%, ít nhất khoảng 70%, ít nhất khoảng

80%, ít nhất khoảng 90%, hoặc ít nhất khoảng 95% (trọng lượng/trọng lượng) theo trọng lượng của chế phẩm. Lipit có thể là rắn hoặc lỏng ở nhiệt độ phòng, khi ở dạng lỏng nó được coi là dầu. Theo một phương án, lipit chiết tách theo sáng chế không được phối hợp với lipit khác như DHA được sinh ra bởi nguồn khác (ví dụ, DHA từ dầu cá). Theo một phương án, sau khi chiết, một hoặc nhiều hoặc tất cả các tỷ lệ sau, axit oleic với DHA, axit palmitic với DHA, axit linoleic với DHA, và toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$, không biến đổi nhiều (ví dụ, biến đổi không quá 10% hoặc 5% so với tỷ lệ trong hạt hoặc tế bào nguyên vẹn. Theo phương án khác, lipit chiết tách từ cây được tiếp xúc với quy trình, như quy trình hydro hóa hoặc phân đoạn, những quy trình này có thể làm thay đổi tỷ lệ của một hoặc nhiều hoặc tất cả, axit oleic với DHA, axit palmitic với DHA, axit linoleic với DHA, và toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$, so với tỷ lệ trong tế bào hoặc hạt nguyên vẹn. Khi lipit chiết tách từ cây theo sáng chế ở trong dầu, dầu này còn có thể chứa các phân tử không phải axit béo như sterol.

Như sử dụng trong đây, thuật ngữ "dầu thực vật chiết tách" và "dầu thực vật phân tách" là để chỉ một chất hoặc chế phẩm chứa lipit chiết tách từ cây hoặc lipit thực vật được phân tách và là chất lỏng ở nhiệt độ phòng. Dầu này thu được từ cây hoặc bộ phận của nó như hạt. Dầu chiết tách hoặc phân tách có thể là chế phẩm tương đối thô thu được bằng cách, ví dụ, ép hạt cây, hoặc chế phẩm tinh khiết hơn trong đó hầu hết, nếu không phải là tất cả, một hoặc nhiều hoặc mỗi thành phần nước, axit nucleic, protein và carbohydrat có nguồn gốc từ nguyên liệu cây đã được loại bỏ. Chế phẩm này có thể bao gồm các thành phần khác có thể là lipit hoặc không phải lipit. Theo một phương án, chế phẩm dầu bao gồm lipit với lượng nằm trong khoảng từ ít nhất khoảng 60%, ít nhất khoảng 70%, ít nhất khoảng 80%, ít nhất khoảng 90%, hoặc ít nhất khoảng 95% (trọng lượng/trọng lượng) chiết từ cây. Theo một phương án, dầu chiết tách theo sáng chế không được phối hợp với dầu khác như DHA được tạo ra bởi nguồn khác (ví dụ, DHA từ dầu cá). Theo một phương án, sau khi chiết tách, tỷ lệ của một hoặc nhiều hoặc tất cả, axit oleic với DHA, axit palmitic với DHA, axit linoleic với DHA, và toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$, không được biến đổi đáng kể (ví dụ, sự biến đổi không cao hơn 10% hoặc 5%) so với tỷ lệ trong tế bào hoặc hạt nguyên vẹn. Theo phương án khác, dầu thực vật chiết tách không được tiếp xúc với quy trình, như hydro hóa hoặc phân đoạn, quy trình này có thể biến đổi tỷ lệ của một hoặc nhiều hoặc tất cả, axit oleic với DHA, axit palmitic với DHA, axit linoleic với DHA, và toàn bộ axit béo $\omega 6$: toàn

bộ axit béo $\omega 3$, so với tỷ lệ trong tế bào hoặc hạt nguyên vẹn. Dầu thực vật chiết tách theo sáng chế có thể chứa phân tử không phải axit béo như sterol.

Như sử dụng trong đây, "dầu" là chế phẩm chứa chủ yếu là lipid và là chất lỏng ở nhiệt độ phòng. Ví dụ, dầu theo sáng chế tốt hơn là chứa lipid với lượng ít nhất là 75%, ít nhất 80%, ít nhất 85% hoặc ít nhất 90% theo trọng lượng. Đặc trưng, dầu tinh khiết chứa ít nhất 90% triaxylglycerol (TAG) theo trọng lượng của lipid trong dầu. Các thành phần phụ của dầu như diaxylglycerol (DAG), các axit béo tự do (FFA), phospholipit và sterol có thể có mặt như mô tả trong đây.

Như sử dụng trong đây, thuật ngữ "axit béo" là để chỉ axit (hoặc axit hữu cơ), thường là có đuôi dài béo, hoặc bão hòa hoặc chưa bão hòa. Đặc trưng các axit béo có chuỗi liên kết cacbon-cacbon có chiều dài ít nhất 8 nguyên tử, tốt hơn nữa là có chiều dài ít nhất 12 carbon. Các axit béo có nhiều trong tự nhiên nhất có số chẵn nguyên tử carbon chẵn do sự sinh tổng hợp chúng bao gồm axetat có hai nguyên tử carbon. Các axit béo có thể có thể ở trạng thái tự do (không este hóa) hoặc ở dạng este hóa như một phần của dạng gắn kết triglycerit, diaxylglycerit, monoaxylglycerit, axyl- CoA (thio-ester) hoặc dạng gắn kết khác. Axit béo có thể được este hóa thành phospholipit như dạng phosphatidylcholin, phosphatidyletanolamin, phosphatidylserin, phosphatidylglycerol, phosphatidylinositol hoặc diphosphatidylglycerol.

"Axit béo bão hòa" không chứa liên kết đôi bất kỳ hoặc các nhóm chức khác trong mạch. Thuật ngữ "bão hòa" là để chỉ hydro, trong đó tất cả các carbon (ngoài nhóm axit carboxylic $[-COOH]$) chứa nhiều hydro nhất có thể. Nói cách khác, dầu omega (ω) chứa 3 hydro (CH_3-) và mỗi carbon trong mạch chứa 2 hydro ($-CH_2-$).

"Axit béo chưa bão hòa" là có dạng tương tự như axit béo bão hòa, ngoại trừ là có một hoặc nhiều nhóm chức alken trong mạch, với mỗi alken thế phần liên kết đơn $-CH_2-CH_2-$ của mạch bằng phần liên kết đôi $-CH=CH-$ (đó là, liên kết đôi carbon với carbon khác). Hai nguyên tử carbon bên cạnh trong mạch được liên kết với bên khác của liên kết đôi có thể có cấu hình *cis* hoặc *trans*.

Như sử dụng trong đây, thuật ngữ "axit béo một liên kết chưa bão hòa" là để chỉ axit béo bao gồm ít nhất 12 nguyên tử carbon trong mạch carbon của nó và chỉ có một nhóm alken (liên kết đôi cacbon-cacbon) trong mạch. Như sử dụng trong đây, thuật ngữ "axit béo nhiều liên kết chưa bão hòa" hoặc "PUFA" là để chỉ axit béo bao gồm ít nhất 12 nguyên tử carbon trong mạch carbon của nó và ít nhất hai nhóm alken (liên kết đôi

cacbon-cacbon).

Như sử dụng trong đây, thuật ngữ "axit béo nhiều liên kết chưa bão hòa mạch dài" và "LC-PUFA" là để chỉ axit béo bao gồm ít nhất 20 nguyên tử carbon trong mạch carbon của nó và ít nhất hai liên kết đôi cacbon-cacbon, và do đó bao gồm VLC-PUFA. Như sử dụng trong đây, thuật ngữ "axit béo nhiều liên kết chưa bão hòa mạch rất dài" và "VLC-PUFA" là để chỉ axit béo bao gồm ít nhất 22 nguyên tử carbon trong mạch carbon của nó và ít nhất ba liên kết đôi cacbon-cacbon. Thông thường, số lượng nguyên tử carbon trong mạch carbon của các axit béo là để chỉ mạch carbon không phân nhánh. Nếu mạch carbon phân nhánh, số lượng nguyên tử carbon không bao gồm nguyên tử carbon trong các nhóm bên. Theo một phương án, axit béo nhiều liên kết chưa bão hòa mạch dài là axit béo ω_3 , đó là có vị trí chưa bão hòa (liên kết đôi cacbon-cacbon) trong liên kết cacbon-cacbon thứ ba từ đầu metyl của axit béo này. Theo phương án khác, axit béo nhiều liên kết chưa bão hòa mạch dài là axit béo ω_6 , nó có vị trí chưa bão hòa (liên kết đôi cacbon-cacbon) trong liên kết cacbon-cacbon thứ sáu từ đầu metyl của axit béo. Theo phương án mở rộng, axit béo nhiều liên kết chưa bão hòa mạch dài được chọn từ nhóm bao gồm; axit arachidonic (ARA, 20:4 Δ 5,8,11,14; ω_6), axit eicosatetraenoic (ETA, 20:4 Δ 8,11,14,17, ω_3), axit eicosapentaenoic (EPA, 20:5 Δ 5,8,11,14,17; ω_3), axit docosapentaenoic (DPA, 22:5 Δ 7,10,13,16,19, ω_3), hoặc axit docosahexaenoic (DHA, 22:6 Δ 4,7,10,13,16,19, ω_3). LC-PUFA cũng có thể là axit dihomog- γ -linoleic (DGLA) hoặc axit eicosatrienoic (ETrA, 20:3 Δ 11,14,17, ω_3). Đã biết rõ rằng LC-PUFA được tạo ra theo sáng chế có thể là hỗn hợp của bất kỳ chất nêu trên hoặc tất cả chúng và có thể bao gồm LC-PUFA khác hoặc dẫn xuất của bất kỳ LC-PUFA. Theo một phương án được ưu tiên, axit béo ω_3 ít nhất là DHA, tốt hơn là, DPA và DHA, hoặc EPA, DPA và DHA.

Hơn nữa, như sử dụng ở đây thuật ngữ "axit béo nhiều liên kết chưa bão hòa mạch dài" và "axit béo nhiều liên kết chưa bão hòa mạch rất dài" là để chỉ axit béo ở trạng thái tự do (không este hóa) hoặc ở dạng este hóa như phần liên kết với triglycerit, diacylglycerit, monoacylglycerit, axyl-CoA hoặc dạng liên kết khác. Axit béo có thể được este hóa là phospholipit như dạng phosphatidylcholin (PC), phosphatidyletanolamin, phosphatidylserin, phosphatidylglycerol, phosphatidylinositol hoặc diphosphatidylglycerol. Do đó, LC-PUFA có thể là hỗn hợp của các dạng trong lipit của tế bào hoặc dầu tinh khiết hoặc lipit được chiết từ tế bào, mô hoặc một cơ

quan. Theo các phương án được ưu tiên, sáng chế đề xuất dầu chứa ít nhất 75% hoặc ít nhất 85% triaxylglyxerol, có phần còn lại là ở dạng khác của lipit như đã nêu, có ít nhất triaxylglyxerol nêu trên chứa LC-PUFA. Dầu này sau đó có thể được tinh chế tiếp hoặc xử lý tiếp, ví dụ bằng cách thủy phân bằng bazơ mạnh để giải phóng các axit béo tự do, hoặc bằng cách chưng cất hoặc tương tự.

Như sử dụng trong đây, “toàn bộ axit béo $\omega 6$ ” hoặc “tổng hàm lượng axit béo $\omega 6$ ” hoặc tương tự là để chỉ tổng của tất cả axit béo $\omega 6$, este hóa và không este hóa, trong lipit chiết tách, dầu, tế bào tái tổ hợp, bộ phận cây hoặc hạt, là hàm lượng xác định, thể hiện là phần trăm của tổng hàm lượng axit béo. Các axit béo $\omega 6$ bao gồm (nếu có) LA, GLA, DGLA, ARA, EDA và $\omega 6$ -DPA, và ngoại trừ axit béo $\omega 3$ bất kỳ và axit béo một liên kết chưa bão hòa.

Như sử dụng trong đây, “axit béo $\omega 6$ mới” hoặc “hàm lượng axit béo $\omega 6$ mới” hoặc tương tự là để chỉ tổng của tất cả các axit béo $\omega 6$ trừ LA, este hóa và không este hóa, trong lipit chiết tách, dầu, tế bào tái tổ hợp, bộ phận cây hoặc hạt, là hàm lượng xác định, thể hiện là phần trăm của tổng hàm lượng axit béo. Các axit béo $\omega 6$ mới là các axit béo được tạo ra trong tế bào, cây, bộ phận cây và hạt theo sáng chế bằng cách biểu hiện cấu trúc gen (polynucleotit ngoại sinh) được đưa vào trong tế bào, và bao gồm (nếu có) GLA, DGLA, ARA, EDA và $\omega 6$ -DPA, nhưng ngoại trừ LA và axit béo $\omega 3$ bất kỳ và các axit béo một liên kết chưa bão hòa. Tổng hàm lượng axit béo $\omega 6$ và hàm lượng axit béo $\omega 6$ mới ví dụ được xác định bằng cách chuyển hóa axit béo trong mẫu thành FAME và phân tích bởi GC, như mô tả trong Ví dụ 1.

Như sử dụng trong đây, “toàn bộ axit béo $\omega 3$ ” hoặc “tổng hàm lượng axit béo $\omega 3$ ” hoặc tương tự là để chỉ tổng của tất cả các axit béo $\omega 3$, este hóa và không este hóa, trong lipit chiết tách, dầu, tế bào tái tổ hợp, bộ phận cây hoặc hạt, là hàm lượng xác định, thể hiện là phần trăm của tổng hàm lượng axit béo. Các axit béo $\omega 3$ bao gồm (nếu có) ALA, SDA, ETrA, ETA, EPA, DPA và DHA, và ngoại trừ axit béo $\omega 6$ bất kỳ và các axit béo một liên kết chưa bão hòa.

Như sử dụng trong đây, “axit béo $\omega 3$ mới” hoặc “tổng hàm lượng axit béo $\omega 3$ mới” hoặc tương tự là để chỉ tổng của tất cả các axit béo $\omega 3$ trừ ALA, este hóa và không este hóa, trong lipit chiết tách, dầu, tế bào tái tổ hợp, bộ phận cây hoặc hạt, là hàm lượng xác định, thể hiện là phần trăm của tổng hàm lượng axit béo. Các axit béo

$\omega 3$ mới là các axit béo được tạo ra trong tế bào, cây, bộ phận cây và hạt theo sáng chế bằng cách biểu hiện cấu trúc gen (polynucleotit ngoại sinh) được đưa vào trong tế bào, và bao gồm (nếu có) SDA, ETrA, ETA, EPA, DPA và DHA, nhưng ngoại trừ ALA và axit béo $\omega 6$ bất kỳ và các axit béo một liên kết chưa bão hòa. Tổng hàm lượng axit béo $\omega 3$ và tổng hàm lượng axit béo $\omega 3$ mới ví dụ được xác định bằng cách chuyển hóa axit béo trong mẫu thành FAME và phân tích bằng GC, như mô tả trong Ví dụ 1.

Các protein desaturaza, elongaza và axyl transferaza và gen mã hóa chúng có thể được sử dụng theo sáng chế là enzym bất kỳ đã biết trong lĩnh vực hoặc đồng phân hoặc dẫn xuất của chúng. Ví dụ về gen này và kích thước protein mã hóa được nêu trong Bảng 1. Các enzym desaturaza đã được chỉ ra là tham gia vào quá trình sinh tổng hợp LC-PUFA tất cả các chất trong nhóm có tên là "tiền" desaturaza.

Như sử dụng trong đây, thuật ngữ "tiền-desaturaza" là để chỉ một trong nhóm enzym đưa liên kết đôi vào giữa nhóm carboxyl và một phần chưa bão hòa đã tồn tại trước đó của mạch axyl của lipid, có đặc trưng cấu trúc là sự có mặt của khu vực sắc tố bảo vệ tận cùng là N, cùng với khu vực desaturaza axit béo đặc trưng bao gồm ba hộp histidin có tính bảo tồn cao (Napier et al., 1997).

Hoạt tính của elongaza hoặc desaturaza bất kỳ để sử dụng theo sáng chế có thể được thử nghiệm bằng cách biểu hiện gen mã hóa enzym trong tế bào, ví dụ, tế bào nấm men, tế bào thực vật hoặc tốt hơn là trong mầm sinh dưỡng hoặc cây biến đổi gen, và xác định rằng tế bào, mầm hoặc cây có tăng khả năng sản sinh LC-PUFA so với tế bào, mầm hoặc cây có thể so sánh trong đó enzym không được biểu hiện hay không.

Theo một phương án, một hoặc nhiều desaturaza và/hoặc elongaza để sử dụng theo sáng chế có thể được tinh chế từ vi tảo, tức là tương đồng về trình tự axit amin với polypeptit có thể được tinh chế từ vi tảo.

Trong khi một số enzym nhất định được mô tả cụ thể ở đây là "có chức năng kép", sự không có mặt của thuật ngữ này tất nhiên không có ý là enzym cụ thể không có hoạt tính không phải như được xác định cụ thể.

Desaturaza

Như sử dụng trong đây, thuật ngữ "desaturaza" là để chỉ enzym có khả năng đưa liên kết đôi cacbon-cacbon vào trong nhóm axyl của cơ chất axit béo đặc trưng ở dạng este hóa như, ví dụ, axyl-CoA este. Nhóm axyl có thể được este hóa thành phospholipit như phosphatidylcholin (PC), hoặc thành protein mang axyl (ACP), hoặc theo một

phương án được ưu tiên thành CoA. Desaturaza thường có thể được phân nhóm thành ba nhóm sau đây. Theo một phương án, desaturaza là tiền-desaturaza.

Như sử dụng trong đây, " $\Delta 4$ -desaturaza" là để chỉ protein tiến hành phản ứng desaturaza mà đưa liên kết đôi cacbon-cacbon ở liên kết cacbon-cacbon thứ 4 từ đầu carboxyl của cơ chất axit béo. " $\Delta 4$ -desaturaza" ít nhất là có khả năng chuyển hóa DPA thành DHA. Bước khử bão hòa để tạo ra DHA từ DPA được xúc tác bởi $\Delta 4$ -desaturaza ở sinh vật không phải là động vật có vú, và gen mã hóa enzyme này đã được phân tách từ loài sinh vật đơn bào nước ngọt *Euglena gracilis* và loài nước mặn *Thraustochytrium* sp. (Qiu et al., 2001; Meyer et al., 2003). Theo một phương án, $\Delta 4$ -desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 41, hoặc $\Delta 4$ -desaturaza *Thraustochytrium* sp., mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 80% với SEQ ID NO: 41.

Bảng 1. Gen tách dòng liên quan đến quy trình sinh tổng hợp LC-PUFA

Enzym	Kiểu sinh vật	Loài	Số hiệu lưu trữ:	Protein (aa)	Tài liệu tham khảo
A4-desaturaza	Sinh vật đơn bào	<i>Euglena gracilis</i>	AY278558	541	Meyer et al., 2003
	Tảo	<i>Pavlova lutherii</i>	AY332747	445	Tonon et al., 2003
		<i>Isochrysis galbana</i>	AAV33631	433	Pereira et al., 2004b
		<i>Pavlova salina</i>	AAAY15136	447	Zhou et al., 2007
	Thraustochytrid	<i>Thraustochytrium aureum</i>	AAN75707 AAN75708 AAN75709 AAN75710	515	N/A
		<i>Thraustochytrium sp.</i>	AAM09688	519	Qiu et al. 2001
	Động vật có vú	Người	AF199596	444	Cho et al., 1999b Leonard et al., 2000b
	Giun	<i>Caenorhabditis elegans</i>	AF11440, NM_069350	447	Michaelson et al., 1998b; Watts và Browse, 1999b
	Nấm	<i>Mortierella alpina</i>	AF067654	446	Michaelson et al., 1998a; Knutzon et al., 1998
		<i>Pythium irregulare</i>	AF419297	456	Hong et al., 2002a
		<i>Dictyostelium discoideum</i>	AB022097	467	Saito et al., 2000
		<i>Saprolegnia diclina</i>		470	WO02081668
	Tảo cát	<i>Phaeodactylum tricornutum</i>	AY082392	469	Domergue et al., 2002
	Tảo	<i>Thraustochytrium sp</i>	AF489588	439	Qiu et al., 2001
		<i>Thraustochytrium aureum</i>		439	WO02081668
		<i>Isochrysis galbana</i>		442	WO02081668
	Rêu	<i>Marchantia polymorpha</i>	AY583465	484	Kajikawa et al., 2004

Enzym	Kiểu sinh vật	Loài	Số hiệu lưu trữ..	protein (aa)	Tài liệu tham khảo
Δ6-desaturaza	Động vật có vú	Người	NM_013402	444	Cho et al., 1999a; Leonard et al., 2000
		<i>Mus musculus</i>	NM_019699	444	Cho et al., 1999a
	Giun tròn	<i>Caenorhabditis elegans</i>	Z70271	443	Napier et al., 1998
	Thực vật	<i>Borago officinales</i>	U79010	448	Sayanova et al., 1997
		<i>Echium</i>	AY055117 AY055118		Garcia-Maroto et al., 2002
		<i>Primula vialii</i>	AY234127	453	Sayanova et al., 2003
		<i>Anemone leveillei</i>	AF536525	446	Whitney et al., 2003
	Rêu	<i>Ceratodon purpureus</i>	AJ250735	520	Sperling et al., 2000
		<i>Marchantia polymorpha</i>	AY583463	481	Kajikawa et al., 2004
		<i>Physcomitrella patens</i>	CAA11033	525	Girke et al., 1998
	Nấm	<i>Mortierella alpina</i>	AF110510 AB020032	457	Huang et al., 1999; Sakuradani et al., 1999
		<i>Pythium irregulare</i>	AF419296	459	Hong et al., 2002a
		<i>Mucor circinelloides</i>	AB052086	467	NCBI*
		<i>Rhizopus sp.</i>	AY320288	458	Zhang et al., 2004
		<i>Saprolegnia diclina</i>		453	WO02081668
	Tảo cát	<i>Phaeodactylum tricornutum</i>	AY082393	477	Domergue et al., 2002
	Vì khuẩn	<i>Synechocystis</i>	L11421	359	Reddy et al., 1993
	Tảo	<i>Thraustochytrium aureum</i>		456	WO02081668
Δ5/Δ6-desaturaza có chức năng kép	cá	<i>Danio rerio</i>	AF309556	444	Hastings et al., 2001
C20 Δ8-desaturaza	Tảo	<i>Euglena gracilis</i>	AF139720	419	Wallis và Browse, 1999
	Thực vật	<i>Borago officinales</i>	AAG43277	446	Sperling et al., 2001

Enzym	Kiểu sinh vật	Loài	Số hiệu lưu trữ:	protein (aa)	Tài liệu tham khảo
A6-elongaza	Giun tròn	<i>Caenorhabditis elegans</i>	NM_069288	288	Beaudoin et al., 2000
		<i>Physcomitrella patens</i>	AF428243	290	Zank et al., 2002
	Rêu	<i>Marchantia polymorpha</i>	AY583464	290	Kajikawa et al., 2004
		<i>Mortierella alpina</i>	AF206662	318	Parker-Barnes et al., 2000
	Nấm	<i>Pavlova lutheri</i> **		501	WO 03078639
	Tảo	<i>Thraustochytrium</i>	AX951565	271	WO 03093482
		<i>Thraustochytrium sp</i> **	AX214454	271	WO 0159128
	Động vật có vú	Người	AF231981	299	Leonard et al., 2000b; Leonard et al., 2002
		<i>Rattus norvegicus</i>	AB071985	299	Imagaki et al., 2002
		<i>Rattus norvegicus</i> **	AB071986	267	Imagaki et al., 2002
PUFA-elongaza		<i>Mus musculus</i>	AF170907	279	Tvrđik et al., 2000
		<i>Mus musculus</i>	AF170908	292	Tvrđik et al., 2000
	Cá	<i>Danio rerio</i>	AF532782	291 (282)	Agaba et al., 2004
		<i>Danio rerio</i> **	NM_199532	266	Lo et al., 2003
	Giun	<i>Caenorhabditis elegans</i>	Z68749	309	Abbott et al., 1998 Beaudoin et al., 2000
		<i>Thraustochytrium aureum</i> **	AX464802	272	WO 0208401-A2
	Tảo	<i>Pavlova lutheri</i> **		320	WO 03078639
	Tảo	<i>Isocorysis galbana</i>	AF390174	263	Qi et al., 2002
		<i>Engelena gracilis</i>		258	WO 08/128241
	Tảo	<i>Ostreococcus tauri</i>	AAV67798	300	Meyer et al., 2004
A5-elongaza		<i>Pyramimonas cordata</i>		268	WO 2010/057246
		<i>Pavlova sp. CCM P459</i>	AAV33630	277	Pereira et al., 2004b
		<i>Pavlova salina</i>	AAV15135	302	Robert et al., 2009
		<i>Thalassiosira pseudonana</i>	AAV67800	358	Meyer et al., 2004
	Tảo cát	<i>Oncorhynchus mykiss</i>	CAM55862	295	WO 06/008099
	Cá				
	Rêu	<i>Marchantia polymorpha</i>	BAE71129	348	Kajikawa et al., 2006

* <http://www.ncbi.nlm.nih.gov/> ** Chức năng chưa được chứng minh

Như sử dụng trong đây, " $\Delta 5$ -desaturaza" là để chỉ protein thực hiện phản ứng desaturaza mà đưa liên kết đôi cacbon-cacbon vào liên kết cacbon-cacbon thứ năm tính từ đầu carboxyl của cơ chất axit béo. Ví dụ về $\Delta 5$ -desaturaza được nêu trong Ruiz-Lopez et al. (2012) và Petrie et al. (2010a) và trong Bảng 1 trong đây. Theo một phương án, $\Delta 5$ -desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 30, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 80% với SEQ ID NO: 30. Theo phương án khác, $\Delta 5$ -desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 32, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 53% với SEQ ID NO: 32. Theo phương án khác, $\Delta 5$ -desaturaza là từ loài *Thraustochytrium* hoặc *Emiliana huxleyi*.

Như sử dụng trong đây, " $\Delta 6$ -desaturaza" là để chỉ protein thực hiện phản ứng desaturaza đưa liên kết đôi cacbon-cacbon vào liên kết cacbon-cacbon thứ sáu tính từ đầu carboxyl của cơ chất axit béo. Ví dụ về $\Delta 6$ -desaturaza được nêu trong Ruiz-Lopez et al. (2012) và Petrie et al. (2010a) và trong Bảng 1 trong đây. $\Delta 6$ -desaturaza được ưu tiên là từ *Micromonas pusilla*, *Pythium irregulare* hoặc *Ostreococcus taurii*.

Theo một phương án, $\Delta 6$ -desaturaza còn có đặc trưng là có ít nhất hai, tốt hơn là tất cả ba đặc điểm sau và tốt hơn là trong tế bào cây: i) hoạt tính $\Delta 6$ -desaturaza trên axit α -linolenic (ALA, 18:3 $\Delta 9,12,15$, $\omega 3$) cao hơn so với axit linoleic (LA, 18:2 $\Delta 9,12$, $\omega 6$) làm cơ chất axit béo; ii) hoạt tính $\Delta 6$ -desaturaza trên ALA-CoA làm cơ chất axit béo cao hơn so với trên ALA được gắn với vị trí *sn*-2 của PC làm cơ chất axit béo; và iii) hoạt tính $\Delta 8$ -desaturaza trên ETrA. Ví dụ về $\Delta 6$ -desaturaza như vậy được nêu trong Bảng 2.

Theo một phương án, $\Delta 6$ -desaturaza có hoạt tính cao hơn trên cơ chất $\omega 3$ so với cơ chất $\omega 6$ tương ứng và có hoạt tính trên ALA để tạo ra axit octadecatetraenoic (axit stearidonic, SDA, 18:4 $\Delta 6,9,12,15$, $\omega 3$) với hiệu suất ít nhất là 30%, tốt hơn nữa ít nhất là 40%, hoặc tốt nhất hơn là ít nhất 50% khi được biểu hiện từ polynucleotit ngoại sinh trong tế bào tái tổ hợp như tế bào cây, hoặc ít nhất 35% khi được biểu hiện trong tế bào nấm men. Theo một phương án, $\Delta 6$ -desaturaza có hoạt tính cao hơn, ví dụ, ít nhất hoạt tính $\Delta 6$ -desaturaza cao hơn khoảng 2 lần, trên cơ chất axit béo ALA hơn là LA. Theo phương án khác, $\Delta 6$ -desaturaza có hoạt tính cao hơn, ví dụ, hoạt tính $\Delta 6$ -desaturaza cao hơn ít nhất khoảng 5 lần hoặc hoạt tính cao hơn ít nhất 10 lần trên ALA-CoA làm cơ

chất axit béo hơn so với trên ALA được gắn với vị trí *sn*-2 của PC làm cơ chất axit béo. Theo phương án mở rộng, $\Delta 6$ -desaturaza có hoạt tính trên cả cơ chất axit béo ALA-CoA và trên ALA được gắn với vị trí *sn*-2 của PC.

Bảng 2. Desaturaza chứng minh là có hoạt tính trên cơ chất axyl-CoA

Enzym	Kiểu sinh vật	Loài	Số hiệu lưu trữ:	Cỡ protein (aa)	Tài liệu tham khảo
$\Delta 6$ - desaturaza	Tảo	<i>Mantoniella squamata</i>	CAQ30479	449	Hoffmann et al., 2008
		<i>Ostreococcus tauri</i>	AAW7015 9	456	Domergue et al., 2005
		<i>Micromonas pusilla</i>	EEH58637		Petrie et al., 2010a (SEQ ID NO: 13)
$\Delta 5$ - desaturaza	Tảo	<i>Mantoniella squamata</i>	CAQ30478	482	Hoffmann et al., 2008
	Thực vật	<i>Anemone leveillei</i>	N/A		Sayanova et al., 2007
$\omega 3$ - desaturaza	Nấm	<i>Pythium aphanidermatum</i>	FW362186. 1	359	Xue et al., 2012; WO2008/054565
	Nấm (oomycete)	<i>Phytophthora sojae</i>	FW362214. 1	363	Xue et al., 2012; WO2008/054565
	Nấm (oomycete)	<i>Phytophthora ramorum</i>	FW362213. 1	361	Xue et al., 2012; WO2008/054565

Theo một phương án, $\Delta 6$ -desaturaza có hoạt tính $\Delta 5$ -desaturaza không phát hiện được trên ETA. Theo phương án khác, $\Delta 6$ -desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 16, SEQ ID NO: 19 hoặc SEQ ID NO: 20, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 77% với SEQ ID NO: 16, SEQ ID NO: 19 hoặc SEQ ID NO: 20. Theo phương án khác, $\Delta 6$ -desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 19 hoặc SEQ ID NO: 20, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 67% với một hoặc cả hai SEQ ID NO: 19 hoặc SEQ ID NO: 20. $\Delta 6$ -desaturaza cũng có thể có hoạt tính $\Delta 8$ -desaturaza.

Như sử dụng trong đây, " $\Delta 8$ -desaturaza" là để chỉ protein thực hiện phản ứng desaturaza đưa liên kết đôi cacbon-cacbon vào liên kết cacbon-cacbon thứ 8 tính từ đầu carboxyl của cơ chất axit béo. $\Delta 8$ -desaturaza ít nhất là có khả năng chuyển hóa ETrA thành ETA. Ví dụ về $\Delta 8$ -desaturaza được nêu trong Bảng 1. Theo một phương án, $\Delta 8$ -desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 52, mảnh

hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 80% với SEQ ID NO: 52.

Như sử dụng trong đây, "ω3-desaturaza" là để chỉ protein thực hiện phản ứng desaturaza đưa liên kết đôi cacbon-cacbon vào liên kết cacbon-cacbon thứ ba tính từ đầu metyl của cơ chất axit béo. Do đó, ω3-desaturaza có thể chuyển hóa LA thành ALA và GLA thành SDA (tất cả các axit béo C18), hoặc DGLA thành ETA và/hoặc ARA thành EPA (các axit béo C20). Một số ω3-desaturaza (nhóm I) có hoạt tính chỉ trên cơ chất C18, như ω3-desaturaza thực vật và khuẩn lam. ω3-desaturaza như vậy cũng là Δ15-desaturaza. ω3-desaturaza khác có hoạt tính trên cơ chất C20 không có hoạt tính (nhóm II) hoặc một số hoạt tính (nhóm III) trên cơ chất C18. ω3-desaturaza như vậy cũng là Δ17-desaturaza. ω3-desaturaza được ưu tiên là kiểu nhóm III chuyển hóa LA thành ALA, GLA thành SDA, DGLA thành ETA và ARA thành EPA, như ω3-desaturaza *Pichia pastoris* (SEQ ID NO: 12). Ví dụ về ω3-desaturaza bao gồm ω3-desaturaza được mô tả bởi Pereira et al. (2004a) (ω3-desaturaza *Saprolegnia diclina*, nhóm II), Horiguchi et al. (1998), Berberich et al. (1998) và Spychalla et al. (1997) (ω3-desaturaza *C. elegans*, nhóm III). Theo một phương án được ưu tiên, ω3-desaturaza là ω3-desaturaza nấm. Như sử dụng trong đây, "ω3-desaturaza nấm" là để chỉ ω3-desaturaza là từ nguồn gốc nấm, bao gồm nguồn gốc mốc nước, hoặc biến thể của nó mà trình tự axit amin của biến thể tương đồng ít nhất là 95% với nó. Gen mã hóa một số ω3-desaturaza đã được phân tách từ nguồn gốc nấm như, ví dụ, từ *Phytophthora infestans* (Số hiệu lưu trữ: CAJ30870, WO2005083053; SEQ ID NO: 70), *Saprolegnia diclina* (Số hiệu lưu trữ: AAR20444, Pereira et al., 2004a & Patent Mỹ số 7211656), *Pythium irregulare* (WO2008022963, Group II; SEQ ID NO: 72), *Mortierella alpina* (Sakuradani et al., 2005; Số hiệu lưu trữ: BAD91495; WO2006019192), *Thalassiosira pseudonana* (Armbrust et al., 2004; Số hiệu lưu trữ: XP_002291057; WO2005012316, SEQ ID NO: 71), *Lachancea kluyveri* (cũng có tên là *Saccharomyces kluyveri*; Oura et al., 2004; Số hiệu lưu trữ: AB118663). Xue et al. (2012) mô tả ω3-desaturaza từ mốc nước *Pythium aphanidermatum*, *Phytophthora sojae*, và *Phytophthora ramorum* là những enzym có khả năng chuyển hóa cơ chất axit béo ω6 thành axit béo ω3 tương ứng, với ưu tiên cho cơ chất C20, tức là chúng có hoạt tính Δ17-desaturaza mạnh hơn hoạt tính Δ15-desaturaza hoạt tính. Các enzym này không có hoạt tính Δ12-desaturaza,

nhưng có thể sử dụng các axit béo trong cả axyl- CoA và phần phospholipit làm cơ chất.

Theo một phương án được ưu tiên hơn, ω 3-desaturaza nấm là ω 3-desaturaza/ Δ 15-desaturaza *Pichia pastoris* (cũng có tên là *Komagataella pastoris*) (Zhang et al., 2008; Số hiệu lưu trữ: EF116884; SEQ ID NO: 12), hoặc polypeptit tương đồng ít nhất là 95% với nó.

Theo một phương án, ω 3-desaturaza ít nhất là có khả năng chuyển hóa ARA thành EPA, DGLA thành ETA, GLA thành SDA, cả ARA thành EPA và DGLA thành ETA, cả ARA thành EPA và GLA thành SDA, hoặc tất cả ba khả năng này.

Theo một phương án, ω 3-desaturaza có hoạt tính Δ 17-desaturaza trên axit béo C20 có ít nhất ba liên kết đôi cacbon-cacbon, tốt hơn là ARA. Theo phương án khác, ω 3-desaturaza có hoạt tính Δ 15-desaturaza trên axit béo C18 có ba liên kết đôi carbon-carbon, tốt hơn là GLA. Tốt hơn là, có cả hai hoạt tính này.

Như sử dụng trong đây, " Δ 12-desaturaza" là để chỉ protein thực hiện phản ứng desaturaza đưa liên kết đôi cacbon-cacbon vào liên kết cacbon-cacbon thứ 12 tính từ đầu carboxyl của cơ chất axit béo. Đặc trưng, Δ 12-desaturaza lần lượt chuyển hóa hoặc oleoyl-phosphatidylcholin hoặc oleoyl-CoA thành linoleoyl- phosphatidylcholin (18:1-PC) hoặc linoleoyl-CoA (18:1-CoA). Phân nhóm sử dụng cơ chất liên kết PC được gọi là Δ 12-desaturaza phụ thuộc phospholipit, phân nhóm sau là Δ 12-desaturaza phụ thuộc axyl-CoA. Δ 12-desaturaza thực vật và nấm thường thuộc phân nhóm đầu tiên, trong khi Δ 12-desaturaza động vật thuộc nhóm sau, ví dụ Δ 12-desaturaza được mã hóa bởi các gen c tách dòng từ côn trùng bởi Zhou et al. (2008). Nhiều trình tự Δ 12-desaturaza khác có thể được xác định dễ dàng bằng cách tra cứu dữ liệu trình tự.

Như sử dụng trong đây, " Δ 15-desaturaza" là để chỉ protein thực hiện phản ứng desaturaza đưa liên kết đôi cacbon-cacbon vào liên kết cacbon-cacbon thứ 15 tính từ đầu carboxyl của cơ chất axit béo. Một số gen mã hóa Δ 15-desaturaza đã được tách dòng từ loài thực vật và nấm. Ví dụ, US5952544 mô tả axit nucleic mã hóa Δ 15-desaturaza thực vật (FAD3). Các enzym bao gồm môtip axit amit có đặc trưng là Δ 15-desaturaza thực vật. WO200114538 mô tả gen mã hóa FAD3 đậu nành. Nhiều trình tự Δ 15-desaturaza khác có thể được xác định dễ dàng bằng cách tra cứu dữ liệu trình tự.

Như sử dụng trong đây, " Δ 17-desaturaza" là để chỉ protein thực hiện phản ứng

desaturaza đưa liên kết đôi cacbon-cacbon vào liên kết cacbon-cacbon thứ 17 tính từ đầu carboxyl của cơ chất axit béo. $\Delta 17$ -desaturaza cũng được coi là $\omega 3$ -desaturaza nếu nó hoạt động trên cơ chất C20 để đưa liên kết $\omega 3$ thành chưa bão hòa.

Theo một phương án được ưu tiên, $\Delta 12$ -desaturaza và/hoặc $\Delta 15$ -desaturaza là $\Delta 12$ -desaturaza nấm hoặc $\Delta 15$ -desaturaza nấm. Như sử dụng trong đây, “ $\Delta 12$ -desaturaza nấm” hoặc “ $\Delta 15$ -desaturaza nấm” là để chỉ $\Delta 12$ -desaturaza hoặc $\Delta 15$ -desaturaza từ nguồn gốc nấm, bao gồm nguồn gốc mốc nước, hoặc biến thể của nó có trình tự axit amin của biến thể tương đồng ít nhất là 95% với nó. Các gen mã hóa một số desaturaza đã được phân tách từ nguồn gốc nấm. US 7211656 mô tả $\Delta 12$ desaturaza từ *Saprolegnia diclina*. WO2009016202 mô tả desaturaza nấm từ *Helobdella robusta*, *Laccaria bicolor*, *Lottia gigantea*, *Microcoleus chthonoplastes*, *Monosiga brevicollis*, *Mycosphaerella fijiensis*, *Mycosphaerella graminicola*, *Naegleria gruberi*, *Nectria haematococca*, *Nematostella vectensis*, *Phycomyces blakesleeanus*, *Trichoderma reesei*, *Physcomitrella patens*, *Postia placenta*, *Selaginella moellendorffii* và *Microdochium nivale*. WO2005/012316 mô tả $\Delta 12$ -desaturaza từ *Thalassiosira pseudonana* và nấm khác. WO2003/099216 mô tả gen mã hóa $\Delta 12$ -desaturaza nấm và $\Delta 15$ -desaturaza được phân tách từ *Neurospora crassa*, *Aspergillus nidulans*, *Botrytis cinerea* và *Mortierella alpina*. WO2007133425 mô tả $\Delta 15$ desaturaza nấm được phân tách từ: *Saccharomyces kluyveri*, *Mortierella alpina*, *Aspergillus nidulans*, *Neurospora crassa*, *Fusarium graminearum*, *Fusarium moniliforme* và *Magnaporthe grisea*. $\Delta 12$ desaturaza được ưu tiên là từ *Phytophthora sojae* (Ruiz-Lopez et al., 2012).

Một phân nhóm khác của $\Delta 12$ -desaturaza nấm, và $\Delta 15$ -desaturaza nấm, là $\Delta 12/\Delta 15$ -desaturaza nấm có chức năng kép. Gen mã hóa enzym đã được tách dòng từ *Fusarium moniliforme* (Số hiệu lưu trữ: DQ272516, Damude et al., 2006), *Acanthamoeba castellanii* (Số hiệu lưu trữ: EF017656, Sayanova et al., 2006), *Perkinsus marinus* (WO2007042510), *Claviceps purpure* (Số hiệu lưu trữ: EF536898, Meesapyodsuk et al., 2007) và *Coprinus cinereus* (Số hiệu lưu trữ: AF269266, Zhang et al., 2007).

Theo phương án khác, $\omega 3$ -desaturaza có ít nhất một số hoạt tính trên, tốt hơn là hoạt tính cao hơn, trên cơ chất axyl-CoA so với cơ chất axyl-PC tương ứng. Như sử dụng trong đây, “cơ chất axyl-PC tương ứng” là để chỉ axit béo được este hóa ở vị trí

sn-2 của phosphatidylcholin (PC) trong đó axit béo là axit béo giống như trong cơ chất axyl-CoA. Ví dụ, cơ chất axyl-CoA có thể là ARA-CoA và cơ chất axyl-PC tương ứng là trong *sn*-2 ARA-PC. Theo một phương án, hoạt tính cao hơn ít nhất là 2 lần. Tốt hơn là, ω 3-desaturaza có ít nhất một số hoạt tính trên cả cơ chất axyl-CoA và cơ chất axyl-PC tương ứng của nó và có hoạt tính trên cả cơ chất C18 và C20. Ví dụ về ω 3-desaturaza đã được biết trong số các desaturaza nấm đã được tách dòng nêu trên.

Theo phương án mở rộng, ω 3-desaturaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 12, tương đồng ít nhất là 60% với SEQ ID NO: 12, tốt hơn là tương đồng ít nhất 90% hoặc ít nhất 95% với SEQ ID NO: 12.

Theo một phương án rộng hơn, desaturaza để sử dụng theo sáng chế có hoạt tính cao hơn trên cơ chất axyl-CoA so với cơ chất axyl-PC tương ứng. Theo phương án khác, desaturaza để sử dụng theo sáng chế có hoạt tính trên cơ chất axyl-PC cao hơn so với cơ chất axyl-CoA tương ứng, nhưng có một số hoạt tính trên cả hai cơ chất. Như đã nêu sơ lược trên, “cơ chất axyl-PC tương ứng” là để chỉ axit béo được este hóa ở vị trí *sn*-2 của phosphatidylcholin (PC) trong đó axit béo là giống như axit béo trong cơ chất axyl-CoA. Theo một phương án, hoạt tính cao hơn ít nhất là hai lần. Theo một phương án, desaturaza là Δ 5 hoặc Δ 6-desaturaza, hoặc ω 3-desaturaza, ví dụ về nó được đưa ra, nhưng không giới hạn ở, các enzym nêu trong Bảng 2. Để kiểm tra cơ chất mà desaturaza hoạt động trên đó, cụ thể là cơ chất axyl-CoA hoặc axyl-PC, các thử nghiệm được tiến hành trong tế bào nấm men như mô tả trong Domergue et al. (2003) và (2005). Cũng có thể suy ra cơ chất axyl-CoA tiềm năng của desaturaza khi elongaza được biểu hiện cùng với desaturaza, có hiệu suất chuyển hóa enzym trong tế bào thực vật ít nhất là khoảng 90% trong đó elongaza xúc tác sự kéo dài sản phẩm của desaturaza. Trên cơ sở này, Δ 5-desaturaza và Δ 4-desaturaza được biểu hiện từ cấu trúc GA7 (Ví dụ 2 và 3) và biến thể của nó (Ví dụ 5) có khả năng khử bão hòa cơ chất axyl-CoA, ETA-CoA và DPA-CoA tương ứng của chúng.

Elongaza

Bằng chứng hóa sinh gợi ý rằng quá trình kéo dài axit béo gồm có 4 bước: hóa lỏng, khử, dehydro hóa và lần khử thứ hai. Trong phạm vi của sáng chế, “elongaza” là để chỉ polypeptit xúc tác bước hóa lỏng trong sự có mặt của các thành phần khác của phức kéo dài, trong điều kiện sinh lý thích hợp. Cũng đã chỉ ra rằng đòi hỏi sự biểu hiện khác loại hoặc cùng loại trong tế bào chỉ của thành phần hóa lỏng (“elongaza”) của

phức protein kéo dài để kéo dài mạch axyl đặc trưng. Do đó, elongaza đưa ra có khả năng thúc đẩy thành công hoạt tính khử và dehydro hóa từ thể chủ biến đổi gen để thực hiện phản ứng kéo dài axyl thành công. Tính đặc hiệu của phản ứng kéo dài theo độ dài mạch và độ bão hòa của cơ chất axit béo được cho là nằm ở thành phần hóa lỏng. Thành phần này cũng được cho là bị hạn chế tốc độ trong phản ứng kéo dài.

Như sử dụng trong đây, " $\Delta 5$ -elongaza" ít nhất là có khả năng chuyển hóa EPA thành DPA. Ví dụ về $\Delta 5$ -elongaza bao gồm enzym được bộc lộ trong WO2005/103253. Theo một phương án, $\Delta 5$ -elongaza có hoạt tính trên EPA để tạo ra DPA với hiệu suất ít nhất là 60%, tốt hơn nữa là ít nhất 65%, tốt hơn nữa là ít nhất 70% hoặc tốt nhất là ít nhất 80% hoặc 90%. Theo phương án mở rộng, $\Delta 5$ -elongaza bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 37, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 47% với SEQ ID NO: 37. Theo phương án mở rộng, $\Delta 6$ -elongaza là từ *Ostreococcus taurii* hoặc *Ostreococcus lucimarinus* (US2010/088776).

Như sử dụng trong đây, " $\Delta 6$ -elongaza" ít nhất là có khả năng chuyển hóa SDA thành ETA. Ví dụ về $\Delta 6$ -elongaza bao gồm enzym được nêu trong Bảng 1. Theo một phương án, elongaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 25, mảnh hoạt tính sinh học của nó (như mảnh như nêu trong SEQ ID NO: 26), hoặc trình tự axit amin tương đồng ít nhất là 55% với một hoặc cả hai SEQ ID NO: 25 hoặc SEQ ID NO: 26. Theo một phương án, $\Delta 6$ -elongaza là từ *Physcomitrella patens* (Zank et al., 2002; Số hiệu lưu trữ: AF428243) hoặc *Thalassiosira pseudonana* (Ruiz- Lopez et al., 2012).

Như sử dụng trong đây, " $\Delta 9$ -elongaza" ít nhất là có khả năng chuyển hóa ALA thành ETrA. Ví dụ về $\Delta 9$ -elongaza bao gồm các enzym được nêu trong Bảng 1. Theo một phương án, $\Delta 9$ -elongaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 43, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 80% với SEQ ID NO: 43. Theo phương án khác, $\Delta 9$ -elongaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 46, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 81% với SEQ ID NO: 46. Theo phương án khác, $\Delta 9$ -elongaza bao gồm axit amin có trình tự như được nêu trong SEQ ID NO: 48, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 50% với SEQ ID NO: 48. Theo phương án khác, $\Delta 9$ -elongaza bao gồm axit amin có trình tự như được

nêu trong SEQ ID NO: 50, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 50% với SEQ ID NO: 50. Theo phương án mở rộng, $\Delta 9$ -elongaza có hoạt tính trên cơ chất $\omega 6$ cao hơn cơ chất $\omega 3$ tương ứng, hoặc ngược lại.

Như sử dụng trong đây, thuật ngữ “có hoạt tính trên cơ chất $\omega 6$ cao hơn cơ chất $\omega 3$ tương ứng” là để chỉ hoạt tính tương đối của enzym này trên cơ chất là khác với tác dụng của $\omega 3$ desaturaza. Tốt hơn là, cơ chất $\omega 6$ là LA và cơ chất $\omega 3$ là ALA.

Elongaza có hoạt tính $\Delta 6$ -elongaza và $\Delta 9$ -elongaza ít nhất là có khả năng (i) chuyển hóa SDA thành ETA và (ii) chuyển hóa ALA thành ETrA và có hoạt tính $\Delta 6$ -elongaza hoạt tính cao hơn hoạt tính $\Delta 9$ -elongaza. Theo một phương án, elongaza có hiệu suất chuyển hóa trên SDA để tạo ra ETA ít nhất là 50%, tốt hơn nữa là ít nhất 60%, và/hoặc hiệu suất chuyển hóa trên ALA để tạo ra ETrA ít nhất là 6% hoặc tốt hơn nữa là ít nhất 9%. Theo phương án khác, elongaza có hoạt tính $\Delta 6$ -elongaza cao hơn ít nhất khoảng 6,5 lần so với hoạt tính $\Delta 9$ -elongaza. Theo phương án mở rộng, elongaza có hoạt tính $\Delta 5$ -elongaza không phát hiện được.

Các enzym khác

Như sử dụng trong đây, thuật ngữ "1-axyl-glycerol-3-phosphat axyltransferaza" (LPAAT), cũng có tên là axit lysophosphatidic-axyltransferaza hoặc axylCoA-lysophosphatidat-axyltransferaza, là để chỉ protein axyl hóa *sn*-1-axyl-glycerol-3-phosphat (*sn*-1 G-3-P) ở vị trí *sn*-2 để tạo ra axit phosphatidic (PA). Do đó, thuật ngữ "hoạt tính 1-axyl-glycerol-3-phosphat axyltransferaza" là để chỉ sự axyl hóa (*sn*-1 G-3- ở vị trí *sn*-2 để tạo ra PA (EC 2.3.1.51). LPAAT được ưu tiên là các enzym có thể sử dụng C22 axyl-CoA nhiều liên kết chưa bão hòa làm cơ chất để chuyển nhóm axyl C22 nhiều liên kết chưa bão hòa vào vị trí *sn*-2 của LPA, tạo ra PA. Các LPAAT như vậy được minh họa trong Ví dụ 13 và có thể được thử nghiệm như mô tả trong đây. Theo một phương án, LPAAT hữu dụng theo sáng chế bao gồm axit amin có trình tự như được nêu trong trình tự bất kỳ SEQ ID NO: 63 đến 69, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 40% với một hoặc nhiều trình tự bất kỳ trong số SEQ ID NO: 63 đến 69. Theo một phương án được ưu tiên, LPAAT hữu dụng theo sáng chế bao gồm axit amin có trình tự như được nêu trong trình tự bất kỳ trong số SEQ ID NO: 64, 65 và 67, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 40% với một hoặc nhiều trình tự bất kỳ trong số SEQ ID NO: 64, 65 và 67.

Như sử dụng trong đây, thuật ngữ “diacylglycerol axyltransferaza” (EC 2.3.1.20; DGAT) là để chỉ chuyển nhóm axyl béo từ axyl-CoA vào cơ chất diacylglycerol để tạo ra triacylglycerol. Do đó, thuật ngữ “hoạt tính diacylglycerol axyltransferaza” là để chỉ sự chuyển axyl-CoA vào diacylglycerol để tạo ra triacylglycerol. Có ba kiểu DGAT đã biết lần lượt có tên gọi là DGAT1, DGAT2 và DGAT3. DGAT1 polypeptit đặc trưng có 10 khu vực xuyên màng, DGAT2 đặc trưng có 2 khu vực xuyên màng, trong khi đó DGAT3 đặc trưng là tan được. Ví dụ về DGAT1 polypeptit bao gồm polypeptit được mã hóa bởi gen DGAT1 từ *Aspergillus fumigatus* (Số hiệu lưu trữ: XP_755172), *Arabidopsis thaliana* (CAB44774), *Ricinus communis* (AAR11479), *Vernicia fordii* (ABC94472), *Vernonia galamensis* (ABV21945, ABV21946), *Euonymus alatus* (AAV31083), *Caenorhabditis elegans* (AAF82410), *Rattus norvegicus* (NP_445889), Người (NP_036211), cũng như biến thể và/hoặc thể đột biến của nó. Ví dụ về DGAT2 polypeptit bao gồm polypeptit được mã hóa bởi gen DGAT2 từ *Arabidopsis thaliana* (Số hiệu lưu trữ: NP_566952), *Ricinus communis* (AAY16324), *Vernicia fordii* (ABC94474), *Mortierella ramanniana* (AAK84179), Người (Q96PD7, Q58HT5), *Bos taurus* (Q70VD8), *Mus musculus* (AAK84175), *Micromonas* CCMP1545, cũng như biến thể và/hoặc thể đột biến của nó. Ví dụ về DGAT3 polypeptit bao gồm polypeptit được mã hóa bởi gen DGAT3 từ cây lạc (*Arachis hypogaea*, Saha, et al., 2006), cũng như biến thể và/hoặc thể đột biến của nó.

Polypeptit/Peptit

Thuật ngữ "tái tổ hợp" đề cập đối với polypeptit là để chỉ polypeptit được tạo ra bởi tế bào, hoặc trong hệ biểu hiện không phải tế bào, với lượng điều chỉnh hoặc ở tốc độ được điều chỉnh, so với trạng thái tự nhiên của nó nếu nó được tạo ra theo cách tự nhiên. Theo một phương án, tế bào này là tế bào trong tự nhiên không tạo ra polypeptit. Tuy nhiên, tế bào này có thể là tế bào chứa gen không phải nội sinh, gen này khiến cho lượng được điều chỉnh của polypeptit được sản xuất. Polypeptit tái tổ hợp theo sáng chế bao gồm polypeptit trong tế bào, mô, cơ quan hoặc sinh vật, hoặc hệ biểu hiện không phải tế bào, trong đó nó được tạo ra tức là polypeptit không được tinh chế hoặc được tách từ khối thành phần khác của tế bào biến đổi gen (tái tổ hợp) mà nó được tạo ra trong tế bào này, và polypeptit được tạo ra trong các tế bào này hoặc hệ biểu hiện không phải tế bào, sau đó hệ này được tinh chế ít nhất một số thành phần khác.

Thuật ngữ "polypeptit" và "protein" thường được sử dụng thay thế nhau.

Polypeptit hoặc nhóm polypeptit có thể được xác định bởi mức tương đồng (% độ tương đồng) của trình tự axit amin của nó với trình tự axit amin đối chứng, hoặc có % độ tương đồng cao với một trình tự axit amin đối chứng cao hơn với trình tự khác. % độ tương đồng của polypeptit với trình tự axit amin đối chứng đặc trưng được xác định bằng phân tích GAP (Needleman và Wunsch, 1970; chương trình GCG) với các thông số khả năng tạo ra $GAP=5$, và khả năng mở rộng của $GAP=0,3$. Trình tự truy vấn có chiều dài ít nhất là 15 axit amin, và phân tích trật tự GAP hai trình tự trên một vùng ít nhất 15 axit amin. Tốt hơn nữa là, trình tự truy vấn có chiều dài ít nhất là 50 axit amin, và phân tích GAP sắp xếp hai trình tự trên một vùng có chiều dài ít nhất là 50 axit amin. Tốt hơn nữa là, trình tự truy vấn có chiều dài ít nhất là 100 axit amin và GAP phân tích sắp xếp hai trình tự trên vùng có chiều dài ít nhất là 100 axit amin. Tốt hơn nữa là, trình tự truy vấn có chiều dài ít nhất là 250 axit amin và phân tích GAP sắp xếp hai trình tự trên vùng có chiều dài ít nhất là 250 axit amin. Tốt hơn nữa là, GAP phân tích sắp xếp hai trình tự trên toàn bộ chiều dài của chúng. Polypeptit hoặc nhóm polypeptit có thể có chung hoạt tính enzym, hoặc hoạt tính khác so với, hoặc không có hoạt tính, của polypeptit đối chứng. Tốt hơn là, polypeptit có hoạt tính enzym ít nhất là 10%, ít nhất 50%, ít nhất 75% hoặc ít nhất 90% hoạt tính của polypeptit đối chứng.

Như sử dụng ở đây mảnh có "hoạt tính sinh học" là một phần của polypeptit nêu ở đây giữ được hoạt tính xác định của polypeptit đối chứng có chiều dài đầy đủ, ví dụ có hoạt tính desaturaza và/hoặc elongaza hoặc hoạt tính enzym khác. Mảnh hoạt tính sinh học như sử dụng ở đây không phải là polypeptit có chiều dài đầy đủ. Mảnh hoạt tính sinh học có thể là mảnh có cỡ bất kỳ miễn sao chúng giữ được hoạt tính xác định. Tốt hơn là, mảnh hoạt tính sinh học giữ được ít nhất 10%, ít nhất 50%, ít nhất 75% hoặc ít nhất 90%, hoạt tính của protein chiều dài đầy đủ.

Liên quan đến polypeptit hoặc enzym đã xác định, sẽ biết rõ là con số % tương đồng cao hơn con số được nêu ở đây sẽ là phương án được ưu tiên. Do đó, khi có thể áp dụng, khi nói đến con số phần trăm % tối thiểu, tốt hơn là polypeptit/enzym bao gồm trình tự axit amin tương đồng ít nhất là 60%, tốt hơn nữa là ít nhất 65%, tốt hơn nữa là ít nhất 70%, tốt hơn nữa là ít nhất 75%, tốt hơn nữa là ít nhất 76%, tốt hơn nữa là ít nhất 80%, tốt hơn nữa là ít nhất 85%, tốt hơn nữa là ít nhất 90%, tốt hơn nữa là ít nhất 91%, tốt hơn nữa là ít nhất 92%, tốt hơn nữa là ít nhất 93%, tốt hơn nữa là ít nhất 94%, tốt hơn nữa là ít nhất 95%, tốt hơn nữa là ít nhất 96%, tốt hơn nữa là ít nhất 97%, tốt hơn

nữa là ít nhất 98%, tốt hơn nữa là ít nhất 99%, tốt hơn nữa là ít nhất 99,1%, tốt hơn nữa là ít nhất 99,2%, tốt hơn nữa là ít nhất 99,3%, tốt hơn nữa là ít nhất 99,4%, tốt hơn nữa là ít nhất 99,5%, tốt hơn nữa là ít nhất 99,6%, tốt hơn nữa là ít nhất 99,7%, tốt hơn nữa là ít nhất 99,8%, và tốt hơn nữa là tương đồng ít nhất 99,9% với SEQ ID NO có liên quan được chọn.

Biến thể/thể đột biến trình tự axit amin của polypeptit được nêu ở đây có thể được điều chế bằng cách đưa các thay đổi nucleotit thích hợp vào axit nucleic nêu trong đây, hoặc bằng cách tổng hợp *in vitro* polypeptit mong muốn. Biến thể/thể đột biến bao gồm, ví dụ, đoạn xóa, đoạn lồng hoặc đoạn thể chứa các gốc nằm trong trình tự axit amin. Có thể kết hợp đoạn xóa, đoạn lồng và đoạn thể để tạo ra cấu trúc cuối cùng, với điều kiện là sản phẩm peptit cuối cùng có hoạt tính enzym mong muốn.

Peptit đột biến (biến đổi) có thể được điều chế bằng cách sử dụng kỹ thuật bất kỳ đã biết trong lĩnh vực này. Ví dụ, polynucleotit nêu ở đây có thể được đưa vào kỹ thuật đột biến gen *in vitro* hoặc kỹ thuật chuyển dịch ADN như đã được mô tả rộng rãi bởi Harayama (1998). Các sản phẩm có nguồn gốc từ ADN đột biến/biến đổi có thể được sàng lọc dễ dàng bởi các kỹ thuật đã được mô tả ở đây để xác định liệu chúng có, ví dụ, hoạt tính desaturaza hoặc elongaza hay không.

Khi thiết kế thể đột biến trình tự axit amin, việc định vị vị trí đột biến và bản chất của đột biến sẽ phụ thuộc vào các đặc điểm cần được điều chỉnh. Các vị trí đột biến có thể được điều chỉnh riêng biệt hoặc hàng loạt, ví dụ, bằng cách (1) đầu tiên thể bởi các lựa chọn axit amin bảo tồn và sau đó bởi nhiều lựa chọn cơ bản hơn phụ thuộc vào kết quả thu được, (2) xóa các gốc đích, hoặc (3) lồng các gốc khác vào gần kề với vị trí đã đặt.

Đoạn xóa trình tự axit amin thường là từ 1 đến 15 gốc, tốt hơn nữa là từ 1 đến 10 gốc và đặc trưng từ khoảng 1 đến 5 gốc gần nhau.

Các đột biến đoạn thể có ít nhất một gốc axit amin trong phân tử polypeptit bị loại và gốc khác được lồng vào vị trí này. Các vị trí được quan tâm nhất trong kỹ thuật đột biến đoạn thể bao gồm vị trí không được bảo tồn nằm giữa desaturaza hoặc elongaza có trong tự nhiên. Tốt hơn là, các vị trí này được thể theo cách tương đối bảo tồn để giữ được hoạt tính enzym. Các phép thể bảo tồn như vậy được trình bày trong Bảng 3 dưới tiêu đề "các gốc thể ví dụ".

Theo một phương án được ưu tiên, polypeptit đột biến/biến thể chỉ có, hoặc

không có quá, một hoặc hai hoặc ba hoặc bốn thay đổi axit amin bảo tồn so với polypeptit có trong tự nhiên. Chi tiết về các thay đổi axit amin bảo tồn được cung cấp trong Bảng 3. Các chuyên gia trong lĩnh vực này sẽ biết rõ, các thay đổi nhỏ như vậy có thể dự đoán là không làm thay đổi hoạt tính của polypeptit khi được biểu hiện trong tế bào tái tổ hợp.

Polypeptit có thể được tạo ra theo nhiều cách, bao gồm việc tạo ra và thu hồi polypeptit tự nhiên hoặc polypeptit tái tổ hợp theo phương pháp đã biết trong lĩnh vực này. Theo một phương án, polypeptit tái tổ hợp được tạo ra bằng cách nuôi cấy tế bào có khả năng biểu hiện polypeptit trong các điều kiện có hiệu quả để sinh ra polypeptit này, như tế bào chủ nêu trong đây. Tế bào được ưu tiên hơn để sinh ra polypeptit là tế bào cây, đặc biệt là hạt của cây.

Polynucleotit

Sáng chế cũng đề xuất và/hoặc sử dụng polynucleotit, ví dụ, có thể là gen, polynucleotit tách, cấu trúc gen khám như phân tử T-ADN, hoặc ADN khám. Nó có thể là ADN hoặc ARN của dòng gốc hệ gen hoặc tổng hợp, sợi kép hoặc sợi đơn, và kết hợp với carbohydrat, lipid, protein hoặc nguyên liệu khác để thực hiện tác dụng cụ thể nêu trong đây. Thuật ngữ "polynucleotit" được sử dụng thay thế ở đây với thuật ngữ "phân tử axit nucleic". Thuật ngữ "polynucleotit tách" có nghĩa là polynucleotit mà, nếu thu được từ nguồn gốc tự nhiên thì trước đó đã được tách từ trình tự polynucleotit mà nó kết hợp hoặc liên kết cùng khi ở trạng thái tự nhiên của nó, hoặc polynucleotit không có trong tự nhiên. Tốt hơn là, polynucleotit tách ít nhất là 60% không có, tốt hơn nữa là ít nhất 75% không có, và tốt hơn nữa là ít nhất 90% không có các thành phần khác mà nó liên kết cùng trong tự nhiên.

Bảng 3. Các gốc thế ví dụ.

Gốc ban đầu	Gốc thế ví dụ
Ala (A)	val; leu; ile; gly
Arg (R)	lys
Asn (N)	gln; his
Asp (D)	glu
Cys (C)	ser
Gln (Q)	asn; his

Glu (E)	asp
Gly (G)	pro, ala
His (H)	asn; gln
Ile (I)	leu; val; ala
Leu (L)	ile; val; met; ala; phe
Lys (K)	arg
Met (M)	leu; phe
Phe (F)	leu; val; ala
Pro (P)	gly
Ser (S)	thr
Thr (T)	ser
Trp (W)	tyr
Tyr (Y)	trp; phe
Val (V)	ile; leu; met; phe, ala

Theo một phương án, polynucleotit theo sáng chế là không có trong tự nhiên. Ví dụ về polynucleotit không có trong tự nhiên bao gồm, nhưng không hạn chế ở, các polynucleotit được được gây đột biến (như bằng cách sử dụng các phương pháp mô tả trong đây), và polynucleotit trong đó khung đọc mở mã hóa protein được liên kết điều khiển với gen khởi đầu mà trong tự nhiên nó không gắn vào (như trong cấu trúc mô tả trong đây).

Như sử dụng trong đây, thuật ngữ "gen" được đưa vào với ý nghĩa rộng nhất của nó và bao gồm trình tự deoxyribonucleotit bao gồm vùng phiên mã và, nếu được phiên mã, vùng mã hóa protein, của gen cấu trúc và bao gồm trình tự được đặt sát với vùng mã hóa trên cả đầu 5' và 3' với khoảng cách ít nhất là khoảng 2 kb với hai đầu này và có vai trò trong sự biểu hiện gen. Về vấn đề này, gen này bao gồm các trình tự tín hiệu kiểm soát như trình tự tín hiệu khởi đầu, trình tự tín hiệu tăng cường, tín hiệu kết thúc và/hoặc polyadenyl hóa thường đi cùng với gen nêu ra, hoặc trình tự tín hiệu kiểm soát khác loại trong trường hợp này gen sẽ được gọi là "gen khảm". Các trình tự được đặt ở đầu 5' của vùng mã hóa protein và có mặt trên mRNA được gọi là các trình tự 5' không dịch mã. Trình tự được đặt ở đầu 3' hoặc xuôi dòng của vùng mã hóa protein và có mặt trên mRNA được gọi là trình tự 3' không dịch mã. Thuật ngữ "gen" bao gồm cả dạng cDNA và hệ gen của gen. Dạng hệ gen hoặc dòng vô tính của gen chứa vùng mã hóa có

thể bị ngắt quãng bởi các trình tự không mã hóa có tên là "intron" hoặc "vùng xen vào" hoặc "trình tự xen vào." Các intron là các đoạn gen được phiên mã vào trong ARN nhân (hnRNA). Intron có thể chứa các yếu tố điều hòa như gen tăng cường. Intron được tách riêng ra hoặc "cắt ra" từ bản sao trong nhân hoặc bản sao sơ bộ; do đó các intron là không có trong bản sao ARN thông tin (mARN). mARN hoạt động trong quá trình dịch mã để xác định rõ trình tự hoặc thứ tự của axit amin trong polypeptit mới sinh. Thuật ngữ "gen" bao gồm phân tử tổng hợp hoặc dung hợp mã hóa tất cả hoặc một phần protein theo sáng chế được mô tả ở đây và trình tự nucleotit bổ sung với trình tự bất kỳ nêu trên.

Như sử dụng trong đây, "ADN khảm" hoặc "cấu trúc gen khảm" là để chỉ phân tử ADN bất kỳ không phải là phân tử ADN tự nhiên ở vị trí tự nhiên của nó, cũng được gọi ở đây là "cấu trúc ADN". Đặc trưng, ADN khảm hoặc gen khảm bao gồm trình tự điều hòa và phiên mã hoặc trình tự mã hóa protein không được thấy là liên kết điều khiển cùng nhau trong tự nhiên tức là chúng là khác loại với nhau. Theo đó, ADN khảm hoặc gen khảm có thể bao gồm trình tự điều hòa và trình tự mã hóa có nguồn gốc từ các nguồn khác nhau, hoặc trình tự điều hòa và trình tự mã hóa có nguồn gốc từ cùng nguồn gốc, nhưng sắp xếp theo cách khác như được tìm thấy trong tự nhiên.

Thuật ngữ "nội sinh" được sử dụng ở đây là để chỉ một chất bình thường có mặt hoặc được tạo ra trong, ví dụ, cây không biến đổi ở cùng giai đoạn phát triển như cây đang nghiên cứu. "Gen nội sinh" là để chỉ một gen tự nhiên ở vị trí tự nhiên của chúng trong hệ gen của một sinh vật. Như sử dụng trong đây, "phân tử axit nucleic tái tổ hợp", "polynucleotit tái tổ hợp" hoặc cách diễn đạt khác của chúng là để chỉ phân tử axit nucleic đã được xây dựng cấu trúc hoặc biến đổi bằng kỹ thuật ADN tái tổ hợp. Thuật ngữ "polynucleotit ngoại lai" hoặc "polynucleotit ngoại sinh" hoặc "polynucleotit khác loại" và tương tự là để chỉ axit nucleic bất kỳ được đưa vào trong hệ gen của tế bào bằng phương pháp thí nghiệm. Gen ngoại lai hoặc ngoại sinh có thể là các gen được lồng vào trong sinh vật không tự nhiên, gen tự nhiên đưa vào một vị trí mới trong thể chủ tự nhiên, hoặc gen khảm. "Gen biến đổi" là gen đã được đưa vào trong hệ gen bằng quy trình biến nạp. Thuật ngữ "biến đổi di truyền", "biến đổi gen" và cách diễn đạt khác của chúng bao gồm việc đưa các gen vào trong tế bào bằng cách biến nạp hoặc tải nạp, gây đột biến gen trong các tế bào và biến đổi hoặc điều chỉnh quá trình điều hòa gen trong tế bào hoặc sinh vật mà các tác động này được thực hiện hoặc trong thế hệ con

của chúng. “Vùng hệ gen” như sử dụng ở đây là để chỉ vị trí nằm trong hệ gen trong đó gen biến đổi, hoặc nhóm các gen biến đổi (cũng được gọi là cụm gen), đã được lồng vào trong tế bào, hoặc dạng tiền thân của nó. Vùng như vậy chỉ bao gồm các nucleotit đã được đưa vào bằng cách can thiệp của con người như các phương pháp được mô tả trong đây.

Thuật ngữ "ngoại sinh" trong nội dung polynucleotit là để chỉ polynucleotit khi có mặt trong tế bào với lượng biến đổi so với trạng thái tự nhiên của nó. Theo một phương án, tế bào này là tế bào không phải tự nhiên bao gồm polynucleotit này. Tuy nhiên, tế bào này có thể là tế bào chứa polynucleotit không phải là nội sinh gây biến đổi lượng sản sinh polypeptit đã mã hóa. Polynucleotit ngoại sinh theo sáng chế bao gồm polynucleotit chưa được tách khỏi các thành phần khác của tế bào biến đổi gen (tái tổ hợp), hoặc hệ biểu hiện không phải tế bào, trong đó nó có mặt, và polynucleotit được tạo ra trong các tế bào như vậy hoặc hệ không phải tế bào như vậy sau đó được tinh chế khỏi ít nhất một số thành phần khác. Polynucleotit ngoại sinh (axit nucleic) có thể là sợi duỗi thẳng sát nhau của các nucleotit có trong tự nhiên, hoặc chứa hai hoặc nhiều sợi duỗi thẳng sát nhau của các nucleotit từ các nguồn khác nhau (có trong tự nhiên và/hoặc tổng hợp) được gắn để tạo ra polynucleotit đơn lẻ. Polynucleotit khảm như vậy đặc trưng bao gồm ít nhất một khung đọc mở mã hóa polypeptit theo sáng chế được liên kết điều khiển với gen khởi đầu thích hợp để điều khiển sự phiên mã của khung đọc mở này trong tế bào nghiên cứu.

Như sử dụng trong đây, thuật ngữ “polynucleotit ngoại sinh khác” hoặc cách diễn đạt khác có nghĩa là trình tự nucleotit của mỗi polynucleotit là khác biệt ít nhất một, tốt hơn là nhiều, nucleotit. Polynucleotit mã hóa các ARN có thể hoặc có thể không được dịch mã thành protein trong tế bào. Theo một ví dụ, ưu tiên rằng mỗi polynucleotit mã hóa một protein với một hoạt tính khác nhau. Theo ví dụ khác, mỗi polynucleotit ngoại sinh là tương đồng ít hơn 95%, ít hơn 90%, hoặc ít hơn 80% với polynucleotit ngoại sinh khác. Tốt hơn là, polynucleotit ngoại sinh mã hóa các protein/enzym chức năng. Hơn nữa, ưu tiên là polynucleotit ngoại sinh khác nhau là không trùng lên nhau trong đó mỗi polynucleotit là một vùng riêng biệt, ví dụ, của axit nucleic vận chuyển ngoài nhiễm sắc thể, vùng này không có đoạn trùng với polynucleotit ngoại sinh khác. Tối thiểu, mỗi ngoại sinh polnucleotit có vị trí khởi động và ngừng phiên mã, cũng như gen khởi đầu được chỉ định. Một polynucleotit ngoại sinh độc lập có thể bao gồm hoặc có thể không

bao gồm intron.

Đối với polynucleotit đã xác định, sẽ rõ ràng con số % tương đồng cao hơn những con số đã nêu trên sẽ là phương án được ưu tiên. Do đó, khi có thể áp dụng, khi nói đến con số % tương đồng tối thiểu, ưu tiên rằng polynucleotit bao gồm trình tự polynucleotit tương đồng ít nhất là 60%, tốt hơn nữa là ít nhất 65%, tốt hơn nữa là ít nhất 70%, tốt hơn nữa là ít nhất 75%, tốt hơn nữa là ít nhất 80%, tốt hơn nữa là ít nhất 85%, tốt hơn nữa là ít nhất 90%, tốt hơn nữa là ít nhất 91%, tốt hơn nữa là ít nhất 92%, tốt hơn nữa là ít nhất 93%, tốt hơn nữa là ít nhất 94%, tốt hơn nữa là ít nhất 95%, tốt hơn nữa là ít nhất 96%, tốt hơn nữa là ít nhất 97%, tốt hơn nữa là ít nhất 98%, tốt hơn nữa là ít nhất 99%, tốt hơn nữa là ít nhất 99,1%, tốt hơn nữa là ít nhất 99,2%, tốt hơn nữa là ít nhất 99,3%, tốt hơn nữa là ít nhất 99,4%, tốt hơn nữa là ít nhất 99,5%, tốt hơn nữa là ít nhất 99,6%, tốt hơn nữa là ít nhất 99,7%, tốt hơn nữa là ít nhất 99,8%, và tốt hơn nữa là ít nhất 99,9% với SEQ ID NO có liên quan được chọn.

Polynucleotit theo sáng chế có thể lai chọn lọc, trong các điều kiện nghiêm ngặt, thành polynucleotit mã hóa polypeptit theo sáng chế. Như sử dụng trong đây, các điều kiện nghiêm ngặt là (1) sử dụng trong quá trình lai tác nhân gây biến tính như formamit, ví dụ, formamit 50% (thể tích/thể tích) với albumin huyết thanh bò 0,1% (trọng lượng/thể tích), Ficoll 0,1%, polyvinylpyrrolidon 0,1%, đệm natri phosphat 50mM ở độ pH =6,5 với NaCl 750mM, natri xitrat 75mM ở 42°C; hoặc (2) sử dụng formamit 50%, 5 x SSC (NaCl 0,75M, natri xitrat 0,075M), natri phosphat 50mM (độ pH =6,8), natri pyrophosphat 0,1%, dung dịch 5 x Denhardt's, ADN tinh trùng cá hồi được chiếu âm (50 g/ml), SDS 0,1% và dextran sulfat 10% ở 42°C trong 0,2 x SSC và SDS 0,1% và/hoặc (3) sử dụng cường độ ion thấp và nhiệt độ cao để rửa sạch, ví dụ, NaCl 0,015 M/natri xitrat 0,0015M/SDS 0,1% ở 50°C.

Polynucleotit theo sáng chế có thể có, so với các phân tử có trong tự nhiên, một hoặc nhiều đột biến là đoạn xóa, đoạn lỏng, hoặc đoạn thể gốc nucleotit. Polynucleotit có các đột biến so với trình tự đối chứng hoặc là có trong tự nhiên (được nói rằng, được tách từ nguồn tự nhiên) hoặc tổng hợp (ví dụ, bằng cách thực hiện phép đột biến gen trực tiếp tại vị trí hoặc kỹ thuật chuyển dịch ADN trên axit nucleic như mô tả ở trên). Do đó, rõ ràng là polynucleotit theo sáng chế hoặc là có thể từ nguồn gốc có trong tự nhiên hoặc có thể là tái tổ hợp. Polynucleotit được ưu tiên là các polynucleotit có các vùng mã hóa được tối ưu hóa theo để dịch mã trong tế bào thực vật, như đã biết trong

lĩnh vực này.

Vector tái tổ hợp

Một phương án theo sáng chế bao gồm vector tái tổ hợp, bao gồm ít nhất một phân tử polynucleotit nêu trong đây, lồng vào trong vector bất kỳ có khả năng phân phối phân tử polynucleotit vào trong tế bào chủ. Vector tái tổ hợp bao gồm vector biểu hiện. Vector tái tổ hợp chứa trình tự polynucleotit khác loại, đó là, trình tự polynucleotit không có trong tự nhiên bên cạnh phân tử polynucleotit nêu trong đây, mà tốt hơn là chúng có nguồn gốc từ loài không phải loài mà (các) polynucleotit này bắt nguồn từ đó. Vector có thể hoặc là ARN hoặc là ADN và đặc trưng là plasmit.

Vector plasmit đặc trưng bao gồm các trình tự axit nucleic bổ sung để lựa chọn, khuếch đại, và biến nạp dễ dàng cắt xét biểu hiện trong tế bào nhân không điển hình, ví dụ, vector có nguồn gốc từ pUC, vector có nguồn gốc từ pSK, vector có nguồn gốc từ pGEM, vector có nguồn gốc từ pSP, vector có nguồn gốc từ pBS, hoặc tốt hơn là vector kép chứa một hoặc nhiều vùng T-ADN. Trình tự axit nucleic bổ sung bao gồm các dòng gốc sao chép để tự sao chép vector, gen đánh dấu chọn lọc được, tốt hơn là mã hóa tính kháng kháng sinh hoặc thuốc diệt cỏ, vị trí đa tách dòng đơn nhất mang lại nhiều vị trí để lồng trình tự axit nucleic hoặc gen mã hóa vào trong cấu trúc axit nucleic, và trình tự làm tăng cường mức biến nạp của tế bào nhân không điển hình và nhân điển hình (đặc biệt là thực vật). Vector tái tổ hợp có thể bao gồm nhiều hơn một polynucleotit nêu trong đây, ví dụ ba, bốn, năm hoặc sáu polynucleotit nêu ở đây kết hợp, tốt hơn là cấu trúc gen khả năng theo sáng chế, mỗi polynucleotit được liên kết điều khiển với trình tự điều khiển biểu hiện có thể điều khiển trong tế bào quan tâm. Tốt hơn là, nhiều hơn một polynucleotit nêu trong đây, ví dụ 3, 4, 5 hoặc 6 polynucleotit, được liên kết đồng hóa trị với nhau trong vector tái tổ hợp riêng, tốt hơn là trong một phân tử T-ADN đơn, phân tử này sau đó có thể được đưa vào dạng phân tử đơn vào trong tế bào để tạo ra tế bào tái tổ hợp theo sáng chế, và tốt hơn là được sát nhập vào hệ gen của tế bào tái tổ hợp, ví dụ ở cây biến đổi gen. Theo đó, polynucleotit được nối như vậy sẽ cùng nhau được thừa hưởng locus gen đơn lẻ trong thế hệ con của tế bào hoặc cây tái tổ hợp. Vector hoặc cây tái tổ hợp có thể bao gồm hai hoặc nhiều vector tái tổ hợp, mỗi vector chứa nhiều polynucleotit, ví dụ trong đó mỗi vector tái tổ hợp chứa 3, 4, 5 hoặc 6 polynucleotit.

"Liên kết điều khiển" như sử dụng ở đây là để chỉ một liên hệ chức năng giữa hai hoặc nhiều đoạn axit nucleic (ví dụ, ADN). Đặc trưng, nó là để chỉ liên hệ chức năng

của yếu tố điều hòa phiên mã (gen khởi đầu) với trình tự được phiên mã. Ví dụ, gen khởi đầu được liên kết điều khiển với trình tự mã hóa, như polynucleotit nêu trong đây, nếu nó kích thích hoặc điều hòa sự phiên mã của trình tự mã hóa trong tế bào thích hợp. Nói chung, các yếu tố điều hòa phiên mã gen khởi đầu được liên kết điều khiển với trình tự được phiên mã, chúng ở bên cạnh về mặt vật lý với trình tự được phiên mã, tức là, chúng hoạt động *cis*. Tuy nhiên, một số yếu tố điều hòa phiên mã, như gen tăng cường, cần phải không ở cạnh nhau về mặt vật lý hoặc không đặt gần với trình tự mã hóa mà chúng làm tăng sự phiên mã.

Khi có nhiều gen khởi đầu, mỗi gen khởi đầu độc lập có thể giống hoặc khác nhau. Các phân tử tái tổ hợp như ADN khảm hoặc cấu trúc gen cũng có thể chứa (a) một hoặc nhiều trình tự tín hiệu tiết mã hóa trình tự peptit tín hiệu, khiến cho polypeptit biểu hiện ở đây có thể được tiết từ tế bào sinh ra polypeptit này hoặc chúng cung cấp cho việc định hướng của polypeptit biểu hiện, ví dụ để giữ polypeptit trong lưới nội bào (ER) trong tế bào hoặc chuyển vào lập thể, và/hoặc (b) chứa trình tự dung hợp dẫn đến sự biểu hiện phân tử axit nucleic là protein dung hợp. Ví dụ về các đoạn trình tự tín hiệu thích hợp bao gồm đoạn trình tự tín hiệu bất kỳ có khả năng điều khiển sự tiết hoặc định hướng của polypeptit nêu trong đây. Các phân tử tái tổ hợp cũng có thể bao gồm các trình tự xen vào và/hoặc trình tự không được dịch mã xung quanh và/hoặc nằm trong trình tự axit nucleic của phân tử axit nucleic nêu trong đây.

Để xác định đặc điểm thuận lợi thể biến nạp, cấu trúc axit nucleic mong muốn bao gồm gen đánh dấu chọn lọc được hoặc sàng lọc được, hoặc ngoài chúng ra, như polynucleotit ngoại lai hoặc ngoại sinh. "Gen đánh dấu" có nghĩa là gen truyền phenotyp riêng biệt cho tế bào biểu hiện gen đánh dấu này và do đó cho phép các tế bào đã biến nạp được phân biệt với các tế bào không có gen đánh dấu này. Gen đánh dấu chọn lọc được mang đặc điểm khiến cho có thể "lựa chọn" dựa trên tính kháng với tác nhân chọn lọc bởi đặc điểm này (ví dụ, thuốc diệt cỏ, kháng sinh, chiếu xạ, nhiệt, hoặc các phương thức xử lý khác có hại cho tế bào không được gây biến nạp). Gen đánh dấu có thể sàng lọc được (hoặc là gen báo cáo) có đặc điểm là có thể xác định thông qua việc quan sát hoặc kiểm tra, tức là, bằng cách "sàng lọc" (ví dụ, β -glucuronidaza, luciferaza, GFP hoặc enzym hoạt tính khác không có trong tế bào không được gây biến nạp). Gen đánh dấu và trình tự nucleotit được quan tâm không cần phải được liên kết. Việc lựa chọn thực tế gen đánh dấu là không bắt buộc miễn sao nó có chức năng (tức là, có tính chọn

lọc) khi kết hợp với các tế bào được chọn như tế bào cây.

Ví dụ về gen đánh dấu chọn lọc được của vi khuẩn là các gen đánh dấu mang tính kháng kháng sinh như kháng ampicilin, erythromycin, cloramphenicol hoặc tetracyclin, tốt hơn là kháng kanamycin. Các gen đánh dấu chọn lọc được ví dụ để lựa chọn thể biến nạp thực vật bao gồm, nhưng không hạn chế ở, gen *hyg* mã hóa tính kháng B hygromycin; gen neomycin phosphotransferaza (*nptII*) mang tính kháng kanamycin, paromomycin, G418; gen glutathion-S-transferaza từ gan chuột mang tính kháng glutathion có nguồn gốc từ thuốc diệt cỏ, ví dụ, được mô tả trong EP 256223; gen glutamin synthetaza khi được biểu hiện, mang tính kháng chất ức chế glutamin synthetaza như phosphinothrixin, ví dụ, được mô tả trong WO 87/05327, gen axetyltransferaza từ *Streptomyces viridochromogenes* mang tính kháng tác nhân chọn lọc phosphinothrixin, ví dụ, như được mô tả trong EP 275957, gen mã hóa 5-enolshikimat-3-phosphat synthaza (EPSPS) mang tính dung nạp N-phosphonometylglyxin, ví dụ, như được mô tả trong tài liệu Hinchee et al. (1988), gen *bar* mang tính kháng bialaphos, ví dụ, như được mô tả trong WO91/02071; gen nitrilaza như *bxn* từ *Klebsiella ozaenae* mang tính kháng bromoxynil (Stalker et al., 1988); gen dihydrofolat reductaza (DHFR) mang tính kháng metotrexat (Thillet et al., 1988); gen axetolactat synthaza (ALS) đột biến, mang tính kháng imidazolinon, sulfonylure hoặc các hóa chất ức chế ALS (EP 154,204); gen anthranilat synthaza đột biến mang tính kháng 5-metyl tryptophan; hoặc gen dalapon dehalogenaza mang tính kháng thuốc diệt cỏ.

Các gen đánh dấu có thể sàng lọc bao gồm, nhưng không hạn chế ở, gen *uidA* mã hóa enzym β -glucuronidaza (GUS) của cơ chất tạo màu đã được biết, gen protein huỳnh quang xanh (Niedz et al., 1995) hoặc dẫn xuất của chúng; gen luciferaza (*luc*) (Ow et al., 1986), cho phép phát hiện tín hiệu phát quang sinh học, và các gen khác đã biết trong lĩnh vực này. "Phân tử báo cáo" được sử dụng trong sáng chế có nghĩa là phân tử mà, với bản chất hóa học của nó, tạo ra tín hiệu phân tích có thể xác định cho phép xác định gen khởi đầu hoạt tính bởi sản phẩm protein.

Tốt hơn là, cấu trúc axit nucleic được sát nhập ổn định vào trong hệ gen của tế bào, như tế bào cây. Theo đó, axit nucleic có thể bao gồm các yếu tố thích hợp cho phép phân tử này được sát nhập vào trong hệ gen, tốt hơn là trình tự biên phải và biên trái của phân tử T-ADN, hoặc cấu trúc này được đặt vào trong vectơ thích hợp có thể

được sát nhập vào nhiễm sắc thể của tế bào.

Biểu hiện

Như sử dụng trong đây, vector biểu hiện là vector ADN có khả năng gây biến nạp tế bào chủ và biểu hiện hiệu quả một hoặc nhiều phân tử polynucleotit xác định. Vector biểu hiện được ưu tiên của sáng chế có thể điều khiển quá trình biểu hiện gen ở nấm men và/hoặc tế bào thực vật. Vector biểu hiện hữu dụng theo sáng chế chứa các trình tự điều hòa như trình tự kiểm soát phiên mã, trình tự kiểm soát dịch mã, dòng gốc sao chép, và các trình tự điều hòa khác tương thích với tế bào tái tổ hợp và kiểm soát sự biểu hiện của các phân tử polynucleotit của sáng chế. Cụ thể, polynucleotit hoặc vector hữu dụng theo sáng chế bao gồm trình tự kiểm soát phiên mã. Trình tự kiểm soát phiên mã là các trình tự kiểm soát quá trình khởi động, kéo dài, và kết thúc phiên mã. Trình tự kiểm soát phiên mã đặc biệt quan trọng là các trình tự kiểm soát quá trình khởi động phiên mã, như gen khởi đầu và trình tự tăng cường. Trình tự kiểm soát phiên mã thích hợp bao gồm trình tự kiểm soát phiên mã có thể hoạt động trong ít nhất một tế bào tái tổ hợp theo sáng chế. Việc lựa chọn trình tự điều hòa để sử dụng phụ thuộc vào sinh vật đích như cây và/hoặc cơ quan đích hoặc mô đích được quan tâm. Trình tự điều hòa như vậy có thể thu được từ sinh vật nhân điển hình như cây hoặc virus cây, hoặc có thể là được tổng hợp hóa học. Một loạt trình tự kiểm soát phiên mã như vậy đã được các chuyên gia trong lĩnh vực này biết rõ. Trình tự kiểm soát phiên mã được ưu tiên đặc biệt là gen khởi đầu hoạt động trong việc điều khiển sự phiên mã ở thực vật, đặc hiệu theo cấu trúc hoặc theo giai đoạn và/hoặc theo mô, phụ thuộc vào việc sử dụng cây hoặc bộ phận của cây.

Một số vector thích hợp để chuyển nạp ổn định tế bào thực vật hoặc để tạo ra cây biến đổi gen đã được mô tả trong, ví dụ, Pouwels et al., *Cloning Vector: A Laboratory Manual*, 1985, supp. 1987; Weissbach và Weissbach, *Methods for Plant Molecular Biology*, Academic Press, 1989; và Gelvin et al., *Plant Molecular Biology Manual*, Kluwer Academic Publishers, 1990. Đặc trưng, vector biểu hiện của cây bao gồm, ví dụ, một hoặc nhiều gen thực vật được tách dòng dưới sự kiểm soát phiên mã của trình tự điều hòa 5' và 3' và gen đánh dấu chọn lọc tính trội. Vector biểu hiện của cây cũng chứa vùng điều hòa gen khởi đầu (ví dụ, vùng điều hòa kiểm soát quá trình biểu hiện do tác động hoặc quá trình biểu hiện cơ bản, điều hòa kiểu môi trường - hoặc kiểu phát triển, hoặc đặc hiệu tế bào hoặc đặc hiệu mô), vị trí bắt đầu khởi động phiên mã, vị trí gắn kết

ribosom, tính hiệu xử lý ARN, vị trí kết thúc phiên mã, và/hoặc tín hiệu polyadenyl hóa.

Một số gen khởi đầu cấu trúc hoạt động trong tế bào thực vật đã được mô tả. Gen khởi đầu thích hợp để biểu hiện cấu trúc ở thực vật bao gồm, nhưng không hạn chế ở, gen khởi đầu 35S virus gây bệnh khảm trên súp lơ (CaMV), 35S của virus gây bệnh khảm Figwort (FMV), gen khởi đầu của virus que trên mía, gen khởi đầu của virus đốm vàng họ Thài lài, gen khởi đầu cảm ứng với ánh sáng từ tiểu đơn vị nhỏ của ribulose-1,5-bis-phosphat carboxylaza, gen khởi đầu triosephosphat isomerase trong dịch bào tương của lúa, gen khởi đầu adenin phosphoribosyltransferaza của *Arabidopsis*, gen khởi đầu actin 1 của lúa, gen khởi đầu mannopin synthaza và octopin synthaza, gen khởi đầu Adh, gen khởi đầu sucroza synthaza, gen khởi đầu phức gen R, và gen khởi đầu protein gắn kết diệp lục α/β .

Với mục đích biểu hiện trong các nguồn mô của cây, như là lá, hạt, rễ hoặc thân, ưu tiên rằng gen khởi đầu sử dụng trong sáng chế có mức biểu hiện tương đối cao trong các mô cụ thể này. Với mục đích này, người ta có thể chọn từ một số gen khởi đầu cho các gen có sự biểu hiện tăng cường hoặc đặc hiệu với mô hoặc tế bào. Ví dụ về các gen khởi đầu này được báo cáo trên lý thuyết bao gồm gen khởi đầu glutamin synthetaza GS2 diệp lục từ đậu, gen khởi đầu fructoza-1,6-biphosphataza diệp lục từ lúa mì, gen khởi đầu ST-LS1 tổng hợp ánh sáng trong nhân từ khoai tây, gen khởi đầu serin/threonin kinaza và gen khởi đầu glucoamylaza (CHS) từ *Arabidopsis thaliana*. Cũng được báo cáo là có hoạt tính tổng hợp ánh sáng là gen khởi đầu ribuloza -1,5-bisphosphat carboxylaza, và gen khởi đầu Cab.

Một loạt các gen khởi đầu thực vật được điều hòa đáp ứng với các tín hiệu môi trường, hormon, hóa học, và/hoặc phát triển, cũng được sử dụng để biểu hiện gen ở tế bào thực vật, bao gồm gen khởi đầu được điều hòa bởi (1) nhiệt, (2) ánh sáng, ví dụ, gen khởi đầu RbcS-3A ở đậu, gen khởi đầu RbcS ở ngô; (3) hormon, như axit abscisic, (4) vết thương (ví dụ, WunI); hoặc (5) hóa chất, như methyl jasmonat, axit salixylic, hormone steroid, rượu, Safeners (WO97/06269), hoặc cũng có thể thuận lợi khi sử dụng (6) gen khởi đầu đặc hiệu theo cơ quan.

Như sử dụng trong đây, thuật ngữ “gen khởi đầu đặc hiệu với hạt cây” hoặc cách diễn đạt khác là để chỉ gen khởi đầu ưu tiên, so với các mô cây khác, điều khiển sự phiên mã gen ở các hạt đang phát triển của cây. Theo một phương án, gen khởi đầu đặc

hiệu với hạt được biểu hiện mạnh hơn ít nhất là 5 lần trong hạt đang phát triển của cây so với lá và/hoặc thân cây, và tốt hơn là được biểu hiện mạnh hơn trong mầm của hạt đang phát triển so với các mô cây khác. Tốt hơn là, gen khởi đầu chỉ điều khiển quá trình biểu hiện của gen quan tâm trong hạt đang phát triển, và/hoặc biểu hiện gen quan tâm ở các bộ phận khác của cây như lá là không phát hiện được bởi phân tích thẩm tách ARN và/hoặc RT-PCR. Đặc trưng, gen khởi đầu điều khiển quá trình biểu hiện gen trong quá trình phát triển và tăng trưởng của hạt, cụ thể là trong pha tổng hợp và tích lũy của các hợp chất được trữ trong hạt. Gen khởi đầu này có thể điều khiển quá trình biểu hiện gen trong toàn bộ cơ quan trữ của thực vật hoặc chỉ bộ phận của nó như vỏ hạt, hoặc lá mầm, tốt hơn là trong mầm, trong hạt của cây hai lá mầm hoặc nội nhũ hoặc lớp hạt aluron của hạt cây một lá mầm.

Gen khởi đầu được ưu tiên để biểu hiện đặc hiệu hạt bao gồm i) gen khởi đầu từ gen mã hóa enzym có vai trò trong sinh tổng hợp axit béo và tích lũy trong hạt, như desaturaza và elongaza, ii) gen khởi đầu từ gen mã hóa protein lưu trữ trong hạt, và iii) gen khởi đầu từ gen mã hóa enzym có vai trò trong sinh tổng hợp carbohydrat và tích lũy trong hạt. Gen khởi đầu đặc hiệu với hạt thích hợp là gen khởi đầu napin trong dầu hạt nho (US5,608,152), gen khởi đầu *Vicia faba* USP (Baumlein et al., 1991), gen khởi đầu *Arabidopsis* oleosin (WO98/45461), gen khởi đầu *Phaseolus vulgaris* phaseolin (US5,504,200), gen khởi đầu *Brassica* Bce4 (WO91/13980) hoặc gen khởi đầu Legumin LeB4 từ *Vicia faba* (Baumlein et al., 1992), và gen khởi đầu dẫn đến sự biểu hiện đặc hiệu hạt trong các cây một lá mầm như ngô, lúa mạch, lúa mỳ, lúa mạch đen, lúa và tương tự. Gen khởi đầu đáng chú ý thích hợp là gen khởi đầu (lpt2 hoặc lpt1 lúa mạch WO95/15389 và WO95/23230) hoặc gen khởi đầu được mô tả trong WO99/16890 (gen khởi đầu từ gen hordein lúa mạch, gen glutelin lúa, gen oryzin lúa, gen prolamin lúa, gen gliadin lúa mỳ, gen glutelin lúa mỳ, gen zein ngô, gen glutelin yến mạch đen, gen kasirin lúa miến, gen secalin lúa mạch đen). Gen khởi đầu khác bao gồm các gen khởi đầu được mô tả bởi Broun et al. (1998), Potenza et al. (2004), US20070192902 và US20030159173. Theo một phương án, gen khởi đầu đặc hiệu với hạt được biểu hiện ưu tiên trong các phần hạt xác định như mầm, lá mầm hoặc nội nhũ. Ví dụ về các gen khởi đầu đặc hiệu bao gồm, nhưng không hạn chế ở, gen khởi đầu FP1 (Ellerstrom et al., 1996), gen khởi đầu legumin đậu hà lan (Perrin et al., 2000), gen khởi đầu ngưng kết tố hồng cầu thực vật đậu (Perrin et al., 2000), gen khởi đầu conlinin 1 và

conlinin 2 của các gen mã hóa protein trữ 2S cây lanh (Cheng et al., 2010), gen khởi đầu của gen FAE1 từ *Arabidopsis thaliana*, gen khởi đầu BnGLP của gen protein tương tự globulin của *Brassica napus*, gen khởi đầu LPXR của gen peroxiredoxin từ *Linum usitatissimum*.

Trình tự dẫn đầu không dịch mã 5' có thể có nguồn gốc từ gen khởi đầu được lựa chọn để biểu hiện trình tự gen khác loại của polynucleotit theo sáng chế, hoặc tốt hơn là khác loại với vùng mã hóa của enzym được tạo ra, và có thể được biến đổi theo cách riêng biệt nếu muốn miễn sao làm tăng sự dịch mã mRNA. Để xem xét về việc tối ưu hóa sự biểu hiện của gen biến đổi, xem Koziel et al. (1996). Cũng có thể thu được các vùng không dịch mã 5' từ ARN virus thực vật (virus gây bệnh khảm thuốc lá, virus gây bệnh ký axit thuốc lá, virus gây bệnh khảm lùn ở ngô, virus khảm cỏ linh lăng, trong số các virus khác) từ gen của nhân điển hình thích hợp, gen thực vật (gen dẫn đầu protein gắn kết diệt lục a/b của lúa mì và ngô), hoặc từ trình tự gen tổng hợp. Sáng chế không giới hạn ở các cấu trúc trong đó vùng không được dịch mã có nguồn gốc từ trình tự không được dịch mã 5' đi cùng với trình tự gen khởi đầu. Trình tự dẫn đầu cũng có thể có nguồn gốc từ trình tự gen khởi đầu không được dịch mã hoặc trình tự mã hóa. Trình tự dẫn đầu hữu dụng trong phạm vi sáng chế bao gồm gen dẫn đầu Hsp70 (US5,362,865 và US5,859,347), và yếu tố omega TMV.

Việc kết thúc phiên mã được thực hiện bởi trình tự ADN không được dịch mã 3' được liên kết điều khiển trong vector khảm của polynucleotit được quan tâm. Vùng không được dịch mã 3' của phân tử ADN tái tổ hợp chứa trình tự tín hiệu polyadenyl hóa có chức năng ở thực vật gây bổ sung các adenylat nucleotit vào đầu 3' của ARN. Có thể thu được vùng không được dịch mã 3' từ các gen khác nhau được biểu hiện trong tế bào thực vật. Vùng không được dịch mã 3' của nopal synthase, vùng không được dịch mã 3' từ gen Rubisco tiểu đơn vị nhỏ của đậu Hà Lan, vùng không được dịch mã 3' từ gen protein lưu trữ trong hạt 7S của đậu nành hoặc gen conlinin của cây lanh thường được sử dụng trong khả năng này. Vùng được phiên mã, không được dịch mã 3', chứa trình tự tín hiệu polyadenylat của gen plasmit *Agrobacterium* (Ti) gây khối u cũng là thích hợp.

Các kỹ thuật ADN tái tổ hợp có thể được sử dụng để cải thiện mức biểu hiện của polynucleotit đã được gây biến nạp bằng cách thực hiện xử lý, ví dụ, một số bản sao của phân tử polynucleotit trong tế bào chủ, hiệu suất mà các phân tử polynucleotit này

được phiên mã, hiệu suất mà các bản sao thu được được dịch mã, và hiệu suất biến đổi sau dịch mã. Các kỹ thuật tái tổ hợp hữu dụng để làm tăng sự biểu hiện của các phân tử polynucleotit ở đây bao gồm, nhưng không hạn chế ở, sát nhập phân tử polynucleotit vào trong một hoặc nhiều nhiễm sắc thể tế bào chủ, thêm trình tự ổn định vào mRNA, thể hoặc biến đổi trình tự tín hiệu kiểm soát phiên mã (ví dụ, gen khởi đầu, gen chỉ huy, gen tăng cường), thể hoặc biến đổi trình tự tín hiệu kiểm soát dịch mã (ví dụ, vị trí gắn kết ribosom, trình tự Shine-Dalgarno), biến đổi phân tử polynucleotit để tương ứng với khả năng codon của tế bào chủ, và xóa trình tự biến tính bản sao.

Tế bào tái tổ hợp

Sáng chế cũng đề xuất tế bào tái tổ hợp, tốt hơn là tế bào cây tái tổ hợp, là tế bào chủ được gây biến nạp bằng một hoặc nhiều phân tử tái tổ hợp, như polynucleotit, cấu trúc gen khả hoặc vector tái tổ hợp nêu trong đây. Tế bào tái tổ hợp có thể bao gồm tổ hợp bất kỳ của chúng, như hai hoặc ba vector tái tổ hợp, hoặc vector tái tổ hợp và một hoặc nhiều polynucleotit hoặc ADN khả bổ sung. Tế bào thích hợp theo sáng chế bao gồm tế bào bất kỳ có thể được gây biến nạp bằng polynucleotit, ADN khả hoặc vector tái tổ hợp theo sáng chế, như ví dụ, phân tử mã hóa polypeptit hoặc enzym được mô tả trong đây. Tốt hơn là, tế bào này là tế bào có khả năng được sử dụng để tạo ra LC-PUFA. Tế bào tái tổ hợp có thể là tế bào nuôi cấy, tế bào *in vitro*, hoặc trong một sinh vật như ví dụ cây, hoặc trong một bộ phận như ví dụ hạt hoặc lá. Tốt hơn là, tế bào này là ở cây hoặc bộ phận cây, tốt hơn nữa là ở hạt của cây.

Tế bào chủ mà (các) polynucleotit được đưa vào trong đó hoặc có thể là tế bào không được gây biến nạp hoặc có thể là tế bào đã được gây biến nạp bằng ít nhất một phân tử axit nucleic. Các phân tử axit nucleic như vậy có thể liên quan đến quá trình tổng hợp LC-PUFA, hoặc không liên quan. Tế bào chủ theo sáng chế hoặc có thể là có khả năng sinh ra protein nội sinh (tức là, tự nhiên) nêu trong đây, trong trường hợp này tế bào tái tổ hợp có nguồn gốc từ đó có khả năng tạo ra mức polypeptit tăng lên, hoặc có thể có khả năng tạo ra protein chỉ sau khi được gây biến nạp bằng ít nhất một polynucleotit theo sáng chế. Theo một phương án, tế bào tái tổ hợp theo sáng chế có khả năng tổng hợp axit béo nhiều liên kết chưa bão hòa mạch dài tăng lên. Như sử dụng trong đây, thuật ngữ "tế bào có khả năng tổng hợp axit béo nhiều liên kết chưa bão hòa mạch dài tăng lên" là một thuật ngữ tương đối trong đó tế bào tái tổ hợp theo sáng chế được so sánh với tế bào chủ không có (các) polynucleotit theo sáng chế, với tế bào tái

tổ hợp tạo ra axit béo nhiều liên kết chưa bão hòa mạch dài hơn, hoặc hàm lượng LC-PUFA cao hơn như DHA (so với các axit béo khác), cao hơn tế bào tự nhiên. Tế bào có khả năng tổng hợp sản phẩm khác tăng lên, như ví dụ axit béo khác, lipit, carbohydrat như tinh bột, phân tử ARN, polypeptit, sản phẩm hoặc dược phẩm khác có ý nghĩa tương tự.

Tế bào chủ theo sáng chế có thể là tế bào bất kỳ có khả năng tạo ra ít nhất một protein được mô tả trong đây, và bao gồm tế bào vi khuẩn, nấm, (bao gồm nấm men), ký sinh trùng, động vật chân đốt, động vật và thực vật. Các tế bào này có thể là nhân không điển hình hoặc nhân điển hình. Tế bào chủ được ưu tiên là tế bào nấm men và tế bào thực vật. Theo một phương án được ưu tiên, tế bào cây là tế bào hạt, cụ thể là tế bào trong lá mầm hoặc nội nhũ của hạt. Theo một phương án, tế bào này là tế bào động vật hoặc tế bào tảo. Tế bào động vật có thể là kiểu động vật bất kỳ như, ví dụ, tế bào động vật không phải là người, tế bào động vật có xương sống không phải là người không phải là người, tế bào động vật có vú không phải là người, hoặc tế bào động vật dưới nước như, cá hoặc giáp xác, động vật không có xương sống, côn trùng, v.v. Các tế bào có thể là của sinh vật thích hợp cho quy trình lên men. Như sử dụng trong đây, thuật ngữ "quy trình lên men" là để chỉ quy trình lên men bất kỳ hoặc quy trình bao gồm bước lên men. Ví dụ về vi sinh vật lên men bao gồm sinh vật nấm, như nấm men. Như sử dụng trong đây, "nấm men" bao gồm *Saccharomyces* spp., *Saccharomyces cerevisiae*, *Saccharomyces carlbergensis*, *Candida* spp., *Kluveromyces* spp., *Pichia* spp., *Hansenula* spp., *Trichoderma* spp., *Lipomyces starkey*, và *Yarrowia lipolytica*. Nấm men được ưu tiên bao gồm chủng *Saccharomyces* spp., và đặc biệt, *Saccharomyces cerevisiae*.

Cây biến đổi gen

Sáng chế cũng đề xuất cây bao gồm tế bào theo sáng chế, như cây biến đổi gen bao gồm một hoặc nhiều polynucleotit theo sáng chế. Thuật ngữ "cây" như sử dụng ở đây là một danh từ chỉ toàn bộ cây, nhưng khi sử dụng như một tính từ là để chỉ một dạng vật chất có trong, thu được từ, có nguồn gốc từ, hoặc có liên quan đến cây, như ví dụ, một bộ phận của cây (ví dụ lá, thân, rễ, hoa), tế bào riêng lẻ (ví dụ phấn hoa), hạt, tế bào thực vật và tương tự. Thuật ngữ "bộ phận cây" là để chỉ tất cả các bộ phận cây bao gồm ADN cây, bao gồm các cấu trúc sinh dưỡng như, ví dụ, lá hoặc thân, rễ, bộ phận hoặc cấu trúc hoa, phấn hoa, hạt, phần của hạt như mầm, nội nhũ, vảy hoặc vỏ hạt, mô

cây như, ví dụ, mô mạch, tế bào và thể hệ con của chúng, miễn sao bộ phận cây này tổng hợp lipit theo sáng chế.

"Cây biến đổi gen", "cây biến đổi di truyền" hoặc cách diễn đạt khác của chúng là để chỉ cây chứa cấu trúc gen("gen biến đổi") không tìm thấy được ở cây kiểu dại của cùng loài, giống hoặc cây trồng. Cây biến đổi gen như được xác định trong phạm vi sáng chế bao gồm cây và thể hệ con của chúng đã được biến đổi gen bằng kỹ thuật tái tổ hợp để tạo ra lipit hoặc ít nhất một polypeptit nêu ở đây trong cây mong muốn hoặc bộ phận cây. Tế bào thực vật biến đổi gen và bộ phận cây biến đổi gen có ý nghĩa tương đương. "Gen biến đổi" như được nhắc đến ở đây có ý nghĩa thông thường trong lĩnh vực công nghệ và bao gồm trình tự gen đã được tạo ra hoặc biến đổi bởi kỹ thuật tái tổ hợp ADN hoặc ARN và được đưa vào trong tế bào theo sáng chế, tốt hơn là tế bào cây. Gen biến đổi có thể bao gồm trình tự gen có nguồn gốc từ tế bào cây có thể là cùng loài, giống hoặc cây trồng với tế bào cây mà gen biến đổi được đưa vào đó hoặc khác loài, giống hoặc cây, hoặc từ tế bào không phải tế bào cây. Đặc trưng, gen biến đổi đã đưa vào tế bào, như tế bào cây, bằng các thao tác do con người làm, ví dụ, bằng cách biến nạp nhưng có thể sử dụng phương pháp bất kỳ mà các chuyên gia trong lĩnh vực này đã biết.

Thuật ngữ "hạt" và "hột" được sử dụng thay thế cho nhau. "Hạt" là để chỉ hạt già như hạt đã được thu hoạch hoặc hạt vẫn còn trên cây nhưng đã sẵn sàng để thu hoạch, nhưng cũng có thể là để chỉ hạt sau khi hút ẩm hoặc nảy mầm, theo bản mô tả. Hạt hoặc hột già thường có hàm lượng ẩm thấp hơn khoảng 18-20%. "Hạt đang phát triển" như sử dụng ở đây là để chỉ hạt trước khi già, đặc trưng thấy trong các cấu trúc sinh sản của cây sau khi thụ phấn hoặc nở hoa, nhưng cũng có thể là để chỉ các hạt trước khi trưởng thành được tách ra từ cây.

Như sử dụng trong đây, thuật ngữ "thu hoạch bộ phận cây" hoặc "thu lấy hạt" là để chỉ cách thức bất kỳ để thu hoạch bộ phận cây hoặc hạt, lần lượt, bao gồm thu hoạch bộ phận cây hoặc hạt từ cây trên cánh đồng hoặc trong nơi chứa như nhà kính hoặc buồng phát triển, hoặc bằng cách mua hoặc nhận từ nhà cung cấp các bộ phận cây hoặc hạt. Hạt có thể thích hợp để trồng tức là có thể nảy mầm và sinh ra các cây con, hoặc cách khác đã được xử lý theo cách mà không thể nảy mầm nữa, ví dụ làm vỡ, xát hoặc nghiền hạt hữu dụng để làm thức ăn hoặc thực phẩm, hoặc để chiết lipit theo sáng chế.

Như sử dụng trong đây, thuật ngữ "cơ quan tích trữ của cây" là để chỉ một bộ

phận của cây chuyên dụng để trữ năng lượng ở dạng, ví dụ, protein, carbohydrat, axit béo và/hoặc dầu. Ví dụ về cơ quan tích trữ của cây là hạt, quả, rễ củ, và củ. Cơ quan tích trữ của cây được ưu tiên theo sáng chế là hạt.

Như sử dụng trong đây, thuật ngữ “bình thường về phenotip” là để chỉ cây hoặc bộ phận của cây đã biến đổi gen, cụ thể là bộ phận lưu trữ như hạt, củ hoặc quả theo sáng chế không có khả năng phát triển và sinh sôi bị giảm đáng kể so với cây hoặc bộ phận cây không biến đổi. Theo một phương án, cây hoặc bộ phận cây biến đổi gen bình thường về phenotip bao gồm polynucleotit ngoại sinh mã hóa gen ức chế gây cân được liên kết điều khiển với gen khởi đầu đặc hiệu với cơ quan tích trữ của cây và có khả năng phát triển hoặc sinh sôi về cơ bản giống với cây hoặc bộ phận cây cùng dòng không chứa polynucleotit này. Tốt hơn là, sinh khối, tốc độ phát triển, tốc độ nảy mầm, kích thước cơ quan tích trữ, cỡ hạt và/hoặc số lượng hạt sống tạo ra là không thấp hơn 90% so với các chỉ số này ở cây không có polynucleotit ngoại sinh này khi phát triển trong các điều kiện giống nhau. Thuật ngữ này không bao gồm các đặc điểm của cây có thể khác với cây kiểu dại mà chúng không ảnh hưởng đến sự có ích của cây với mục đích thương mại như, ví dụ, phenotip kiểu *ballerina* của lá cây con.

Cây được đề xuất hoặc được dự tính để sử dụng trong thực hành sáng chế bao gồm cây một lá mầm và cây hai lá mầm. Theo các phương án được ưu tiên, cây theo sáng chế là cây hoa màu (ví dụ, ngũ cốc và đậu Hà lan, ngô, lúa mì, khoai tây, bột sắn, lúa, lúa miến, kê, sắn, lúa mạch, hoặc đậu Hà lan), hoặc các cây quả đậu khác. Cây cũng có thể phát triển để tạo ra rễ ăn được, củ, lá, thân, hoặc hoa hoặc quả. Cây có thể là cây rau hoặc cây cảnh. Cây theo sáng chế có thể là: ngô (*Zea mays*), cải dầu (*Brassica napus*, *Brassica rapa* ssp.), mù tạc (*Brassica juncea*), lanh (*Linum usitatissimum*), cỏ linh lăng (*Medicago sativa*), lúa (*Oryza sativa*), lúa mạch đen (*Secale cereale*), lúa miến (*Sorghum bicolor*, *Sorghum vulgare*), hướng dương (*Helianthus annuus*), lúa mì (*Triticum aestivum*), đậu nành (*Glycine max*), thuốc lá (*Nicotiana tabacum*), khoai tây (*Solanum tuberosum*), lạc (*Arachis hypogaea*), bông (*Gossypium hirsutum*), khoai lang (*Lopmoea batatus*), sắn (*Manihot esculenta*), cà phê (*Cofea* spp.), dừa (*Cocos nucifera*), dứa (*Anana comosus*), cây họ cam (*Citrus* spp.), dừa (*Theobroma cacao*), chè (*Camellia senensis*), chuối (*Musa* spp.), lê tàu (*Persea americana*), sung (*Ficus casica*), ổi (*Psidium guajava*), xoài (*Mangifer indica*), ôliu (*Olea europaea*), đu đủ (*Carica papaya*), đào lộn hột (*Anacardium occidentale*), mắc ca

(*Macadamia intergrifolia*), quả hạnh (*Prunus amygdalus*), củ cải đường (*Beta vulgaris*), yền mạch, hoặc lúa mạch.

Theo một phương án được ưu tiên, cây là cây hạt kín.

Theo một phương án, cây là cây hạt dầu, tốt hơn là cây hoa màu có hạt dầu. Như sử dụng trong đây, "cây hạt dầu" là loài cây sử dụng để sản xuất thương mại dầu từ hạt cây. Cây hạt dầu có thể là nho hạt dầu (như cải dầu), ngô, hướng dương, đậu nành, lúa miến, lanh (hạt lanh) hoặc củ cải đường. Hơn nữa, cây hạt dầu có thể là *Brassicas khác*, bông, lạc, anh túc, mù tạc, hạt thầu dầu, vừng, rum, hoặc cây tạo ra quả hạch khác. Cây có thể tạo ra các mức cao dầu trong quả của nó, như ôliu, dầu cọ hoặc dầu dừa. Các cây vườn mà sáng chế có thể áp dụng vào là rau diếp, rau diếp quăn, hoặc các cây họ cải khác bao gồm bắp cải, bông cải xanh, hoặc súp lơ. Sáng chế cũng có thể áp dụng cho cây thuốc lá, cây bầu bí, cà rốt, dâu tây, khoai tây, hoặc hạt tiêu.

Theo một phương án được ưu tiên hơn, cây không biến đổi gen sử dụng để tạo ra cây biến đổi gen theo sáng chế để tạo adầu, đặc biệt là trong hạt, nó có i) các axit béo 18:2 có lượng thấp hơn 20%, thấp hơn 10% hoặc thấp hơn 5% và/hoặc ii) các axit béo 18:3 với lượng thấp hơn 10% hoặc thấp hơn 5%.

Theo một phương án được ưu tiên, cây biến đổi gen là đồng hợp tử với nhau và mỗi gen đã được đưa vào (gen biến đổi) sao cho cây con của nó không được phân ly với phenotip mong muốn. Cây biến đổi có thể là dị hợp tử với (các) gen biến đổi được đưa vào), tốt hơn là dị hợp tử đồng dạng với gen biến đổi, như ví dụ trong thế hệ F1 được trồng từ hạt lai. Cây như vậy có thể mang lại các thuận lợi như độ khỏe ưu thế giống lai, đã biết trong lĩnh vực này.

Khi thích hợp, cây biến đổi gen cũng có thể bao gồm thêm các gen biến đổi mã hóa enzyme có vai trò trong sản xuất LC-PUFA như, nhưng không giới hạn ở, $\Delta 6$ -desaturaza, $\Delta 9$ -elongaza, $\Delta 8$ -desaturaza, $\Delta 6$ -elongaza, $\Delta 5$ -desaturaza, $\omega 3$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 5$ -elongaza, diacylglycerol acyltransferaza, LPAAT, $\Delta 17$ -desaturaza, $\Delta 15$ -desaturaza và/hoặc $\Delta 12$ desaturaza. Ví dụ về các enzyme này có một trong số nhiều hoạt tính đã biết rõ trong lĩnh vực này và bao gồm các enzyme được mô tả trong đây. Trong các ví dụ cụ thể, cây biến đổi gen bao gồm ít nhất polynucleotit ngoại sinh mã hóa;

a) $\Delta 4$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 6$ -desaturaza, $\Delta 5$ -elongaza và $\Delta 6$ -elongaza,

b) $\Delta 4$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 8$ -desaturaza, $\Delta 5$ -elongaza và $\Delta 9$ -elongaza,

c) $\Delta 4$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 6$ -desaturaza, $\Delta 5$ -elongaza, $\Delta 6$ -elongaza, và $\Delta 15$ -desaturaza,

d) $\Delta 4$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 8$ -desaturaza, $\Delta 5$ -elongaza, $\Delta 9$ -elongaza, và $\Delta 15$ -desaturaza,

e) $\Delta 4$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 6$ -desaturaza, $\Delta 5$ -elongaza, $\Delta 6$ -elongaza, và $\Delta 17$ -desaturaza, hoặc

f) $\Delta 4$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 8$ -desaturaza, $\Delta 5$ -elongaza, $\Delta 9$ -elongaza, và $\Delta 17$ -desaturaza.

Theo một phương án, polynucleotit ngoại sinh mã hóa bộ polypeptit là $\Delta 6$ -desaturaza *Pythium irregulare*, $\Delta 5$ -desaturaza *Thraustochytrid* hoặc $\Delta 5$ -desaturaza *Emiliana huxleyi*, $\Delta 6$ -elongaza *Physcomitrella patens*, $\Delta 5$ -elongaza *Thraustochytrid* hoặc $\Delta 5$ -elongaza *Ostreococcus taurii*, $\omega 3$ -desaturaza *Phytophthora infestans* hoặc $\omega 3$ -desaturaza *Pythium irregulare*, và $\Delta 4$ -desaturaza *Thraustochytrid*.

Theo một phương án, cây theo sáng chế được trồng trên cánh đồng, tốt hơn là một quần thể ít nhất 1.000 hoặc 1.000.000 cây về cơ bản là giống nhau, hoặc trên một diện tích ít nhất 1 héc ta. Mật độ trồng khác nhau theo loài thực vật, loài cây, khí hậu, điều kiện đất, tốc độ phân bón và các yếu tố khác đã biết trong lĩnh vực này. Ví dụ, cải dầu đặc trưng được trồng với mật độ trồng là 1,2-1,5 triệu cây trên một héc ta. Cây được thu hoạch như đã biết trong lĩnh vực này, có thể bao gồm xén, vun thành luống và/hoặc gặt cây, sau đó đập và/hoặc sấy nguyên liệu cây để tách hạt từ bộ phận cây còn lại thường là ở dạng trấu. Cách khác, hạt có thể được thu hoạch từ cây trên cánh đồng trong một quy trình duy nhất, cụ thể là gom lại.

Biến nạp cây

Cây biến đổi gen có thể được tạo ra bằng các kỹ thuật đã biết trong lĩnh vực này, như các kỹ thuật được mô tả chung trong A. Slater et al., *Plant Biotechnology - The Genetic Manipulation of Plants*, Oxford University Press (2003), và P. Christou và H. Klee, *Handbook of Plant Biotechnology*, John Wiley và Sons (2004).

Như sử dụng trong đây, thuật ngữ “gây biến nạp ổn định”, “được gây biến nạp ổn định” và cách diễn đạt khác của chúng là để chỉ sự hợp nhất của phân tử axit nucleic

ngoại sinh vào trong hệ gen của tế bào sao cho chúng được chuyển vào tế bào thể hệ con trong quá trình phân chia tế bào mà không cần chủ động lựa chọn sự có mặt của chúng. Thể biến nạp ổn định, hoặc thể hệ con của chúng, có thể được lựa chọn bởi phương pháp bất kỳ đã biết trong lĩnh vực như phương pháp thẩm tách ADN trên ADN nhiễm sắc thể hoặc phép lai tại chỗ hệ gen ADN.

Sự vận chuyển gián tiếp qua *Agrobacterium* là một hệ có thể ứng dụng rộng rãi để đưa gen vào trong tế bào thực vật do ADN có thể được đưa vào tế bào trong mô của cây nguyên vẹn hoặc bộ phận của cây hoặc mảnh cây trong môi trường nuôi cấy mô, để hoặc là biểu hiện tạm thời hoặc để sát nhập ổn định ADN vào trong hệ gen của tế bào cây. Việc sử dụng vector cây trung gian qua *Agrobacterium* để đưa ADN vào trong tế bào thực vật đã được biết rõ trong lĩnh vực này (xem, ví dụ, US 5177010, US 5104310, US 5004863 hoặc US 5159135) bao gồm phương pháp nhúng chìm hoa bằng *Agrobacterium* hoặc vi khuẩn khác có thể vận chuyển ADN vào trong tế bào thực vật. Vùng ADN vận chuyển được xác định bằng trình tự biên, và ADN xen vào (T-DNA) thường được lồng vào trong hệ gen của cây. Hơn nữa, sự sát nhập vào của T-ADN là một quy trình tương đối chính xác trong một số quy trình sắp xếp. Trong các loài cây mà sự biến nạp qua *Agrobacterium* có hiệu quả, đó là phương pháp được lựa chọn do sự thuận tiện và bản chất đã xác định của sự chuyển gen. Vector biến nạp *Agrobacterium* được ưu tiên có khả năng sao chép trong *E. coli* cũng như *Agrobacterium*, cho phép thao tác thuận tiện như mô tả (Klee et al., In: Plant ADN Infectious Agents, Hohn và Schell, eds., Springer-Verlag, New York, pp. 179-203 (1985).

Các phương pháp gia tốc có thể được sử dụng bao gồm, ví dụ, phương pháp chuyển gen bằng súng bắn gen và tương tự. Một ví dụ về phương pháp để phân phối phân tử axit nucleic gây biến nạp vào tế bào thực vật là phương pháp chuyển gen bằng súng bắn gen. Phương pháp này đã được nêu bởi Yang et al., Particle Bombardment Technology for Gene Transfer, Oxford Press, Oxford, England (1994). Các hạt không phải sinh học (vi đạn) có thể được bao bằng axit nucleic và phân phối vào trong tế bào bằng lực đẩy. Các hạt ví dụ bao gồm các hạt là wonfam, vàng, platin, và tương tự. Một lợi điểm riêng của phương pháp chuyển gen bằng súng bắn gen, ngoài việc nó là phương pháp hiệu quả để gây biến nạp cây một lá mầm có thể lặp lại, đó là đòi hỏi nó không phân tách thể nguyên sinh, cũng như không dễ gây nhiễm *Agrobacterium*.

Theo phương án thay thế khác, lập thể có thể được biến nạp ổn định. Các phương pháp bộc lộ về biến nạp lập thể ở các cây cao hơn bao gồm phương pháp bắn hạt ADN chứa gen đánh dấu chọn lọc được và chọn đích ADN đến hệ gen lập thể thông qua kỹ thuật tái tổ hợp cùng loài (US5, 451,513, US5,545,818, US5,877,402, US5,932,479, và WO99/05265).

Phương pháp khác để biến nạp tế bào cũng có thể được sử dụng và bao gồm nhưng không hạn chế ở phương pháp đưa ADN vào cây bằng cách chuyển ADN trực tiếp vào phần hoa, bằng cách bơm trực tiếp ADN vào trong cơ quan sinh sản của cây, hoặc bằng cách bơm trực tiếp ADN vào tế bào của mầm chưa trưởng thành sau khi hydrat hóa lại các mầm đã khô.

Việc phục hồi, phát triển, và nuôi trồng cây từ thể biến nạp nguyên sinh thực vật đơn hoặc từ nhiều mảnh cây được gây biến nạp khác nhau đã được biết rõ trong lĩnh vực này (Weissbach et al., In: *Methods for Plant Molecular Biology*, Academic Press, San Diego, Calif., (1988)). Quy trình phục hồi và phát triển đặc trưng bao gồm các bước lựa chọn tế bào gây biến nạp, nuôi cấy các tế bào đã biệt hóa qua các giai đoạn phát triển mầm thông thường trong suốt giai đoạn cây non mọc rễ. Mầm biến đổi gen và hạt được phục hồi tương tự. Các chồi biến đổi gen mọc rễ thu được sau đó được trồng trong môi trường phát triển cây thích hợp như đất.

Việc phát triển hoặc phục hồi cây chứa gen ngoại lai, gen ngoại sinh cũng đã được biết rõ trong lĩnh vực này. Tốt hơn là, các cây được phục hồi là tự thụ phấn để sinh ra cây biến đổi gen đồng hợp tử. Cách khác, phần hoa thu được từ cây được phục hồi được lai chéo với cây trồng từ hạt của dòng quan trọng trong nông nghiệp. Ngược lại, phần hoa từ cây của các dòng quan trọng này được sử dụng để thụ phấn cho cây được phục hồi. Cây biến đổi gen theo sáng chế chứa axit nucleic ngoại sinh mong muốn được nuôi trồng bằng các phương pháp đã biết với các chuyên gia trong lĩnh vực này.

Để khẳng định sự có mặt của gen biến đổi trong tế bào và cây biến đổi gen, phương pháp khuếch đại phản ứng chuỗi polymeraza (PCR) hoặc phương pháp phân tích thâm tách ADN có thể được thực hiện bằng phương pháp đã biết với các chuyên gia trong lĩnh vực này. Các sản phẩm biểu hiện của gen biến đổi có thể được phát hiện bởi bất kỳ trong số hàng loạt cách, phụ thuộc vào bản chất của sản phẩm, và bao gồm phân tích thâm tách protein và thử nghiệm enzym. Khi đã thu được cây biến đổi gen,

chúng có thể được trồng để tạo ra mô cây hoặc các bộ phận của cây có phenotip mong muốn. Mô cây hoặc bộ phận cây, có thể được thu hoạch, và/hoặc thu lấy hạt. Hạt có thể được sử dụng làm nguồn để trồng thêm cây có mô hoặc bộ phận có đặc điểm mong muốn.

Cây biến đổi gen tạo ra bằng phương pháp phương pháp biến nạp *Agrobacterium* hoặc phương pháp khác đặc trưng chứa một locus gen đơn trên một nhiễm sắc thể. Cây biến đổi gen này có thể được coi là bán hợp tử với (các) gen được bổ sung vào. Được ưu tiên hơn là cây biến đổi gen đồng hợp tử với (các) gen được bổ sung vào; tức là, cây biến đổi gen chứa hai gen bổ sung, một gen trên cùng locus trên mỗi nhiễm sắc thể của cặp nhiễm sắc thể. Cây biến đổi gen có thể là đồng hợp tử thu được bằng cách tự thụ phấn cây biến đổi gen bán hợp tử, nảy mầm một số hạt tạo ra và phân tích cây thu được về gen được quan tâm.

Cũng hiểu rằng hai cây biến đổi gen khác nhau chứa hai gen ngoại sinh được phân ly độc lập hoặc locus cũng có thể được lai chéo (giao phối) để tạo ra cây con chứa cả bộ gen hoặc locus. Việc tự thụ phấn của cây con F1 thích hợp có thể tạo ra các cây đồng hợp tử cả với gen hoặc locus ngoại sinh. Lai ngược lại với cây mẹ và lai tiếp với cây không biến đổi gen cũng được dự tính, là sự nhân giống sinh dưỡng. Mô tả về các phương pháp sinh sản khác thường được sử dụng cho các đặc tính và hoa màu khác có thể thấy trong Fehr, In: *Breeding Methods for Cultivar Development*, Wilcox J. ed., American Society of Agronomy, Madison Wis. (1987).

Tăng mức ARN ngoại sinh và biểu hiện ổn định ARN ngoại sinh

Gen ức chế gây cảm

Theo một phương án, tế bào, cây hoặc bộ phận cây theo sáng chế chứa polynucleotit ngoại sinh mã hóa protein ức chế làm cảm.

Sự làm gập gen sau phiên mã (PTGS) là một cơ chế bảo vệ đặc hiệu với trình tự nucleotit có thể có đích ở cả mARN tế bào và virus để làm giảm sự PTGS xảy ra ở thực vật hoặc nắm một cách ổn định hoặc tạm thời được gây biến nạp bằng ADN ngoại sinh (khác loại) hoặc ADN nội sinh và khiến cho sự tích lũy giảm phân tử ARN có sự tương tự trình tự với axit nucleic được đưa vào.

Đã được xem xét một cách rộng rãi là sự biểu hiện đồng thời gen ức chế gây cảm với gen biến đổi được quan tâm sẽ làm tăng mức ARN có trong tế bào được phiên mã từ gen biến đổi. Trong khi điều này đã được chứng minh là đúng với các tế bào *in vitro*,

thì quan sát thấy các tác dụng phụ đáng kể ở nhiều nghiên cứu biểu hiện đồng thời ở cây nguyên vẹn. Cụ thể hơn, như mô tả trong Mallory et al. (2002), Chapman et al. (2004), Chen et al. (2004), Dunoyer et al. (2004), Zhang et al. (2006), Lewsey et al. (2007) và Meng et al. (2008) cây biểu hiện gen ức chế gây câm, thường là dưới gen khởi đầu cấu trúc, thường có phenotip bất thường đến mức không thể sử dụng chúng để sản xuất thương mại.

Gần đây, đã phát hiện ra rằng mức phân tử ARN có thể tăng lên, và/hoặc mức phân tử ARN ổn định qua nhiều thế hệ, bằng cách hạn chế sự biểu hiện của gen ức chế gây câm ở hạt của cây hoặc bộ phận của nó (WO2010/057246). Như sử dụng trong đây, “protein ức chế làm câm” hoặc SSP là polypeptit bất kỳ có thể được biểu hiện trong tế bào cây làm tăng mức biểu hiện từ gen biến đổi khác trong tế bào cây, đặc biệt là các thế hệ lặp lại quá nhiều lần từ cây biến nạp đầu tiên. Theo một phương án, SSP là gen ức chế làm câm của virus hoặc thể đột biến của nó. Một số lượng lớn gen ức chế làm câm của virus đã biết trong lĩnh vực này và bao gồm, nhưng không hạn chế ở P19, V2, P38, Pe-Po và RPV-P0. Theo một phương án, gen ức chế làm câm của virus bao gồm axit amin có trình tự như được nêu trong trình tự bất kỳ từ SEQ ID NO 53 đến 57, mảnh hoạt tính sinh học của nó, hoặc trình tự axit amin tương đồng ít nhất là 50% với một hoặc nhiều trình tự bất kỳ SEQ ID NO 53 từ 57 và có hoạt tính là gen ức chế gây câm.

Như sử dụng trong đây, thuật ngữ “biểu hiện ổn định”, “được biểu hiện ổn định”, và cách diễn đạt khác của chúng là để chỉ mức phân tử ARN về cơ bản giống hoặc cao hơn trong các cây thế hệ con so với các thế hệ lặp lại, ví dụ ít nhất ba, ít nhất năm hoặc ít nhất 10 thế hệ, so với cây cùng dòng không có polynucleotit ngoại sinh mã hóa gen ức chế gây câm. Tuy nhiên, thuật ngữ này không loại trừ khả năng các thế hệ lặp lại quá nhiều sẽ có sự giảm mức phân tử ARN so với thế hệ trước đó, ví dụ sự giảm không ít hơn 10% trong mỗi thế hệ.

Gen ức chế có thể được chọn từ nguồn bất kỳ ví dụ cây, virus, động vật có vú v.v. Xem WO2010/057246 về danh sách các virus mà có thể thu được gen ức chế từ đó và protein (ví dụ B2, P14 v.v.) hoặc thiết kế vùng mã hóa gen ức chế từ mỗi virus cụ thể. Nhiều bản sao của gen ức chế có thể được sử dụng. Các gen ức chế khác nhau có thể được sử dụng cùng nhau (ví dụ, theo kiểu trước sau).

Phân tử ARN

Về cơ bản, phân tử ARN bất kỳ được mong muốn là biểu hiện ở hạt cây có thể được biểu hiện đồng thời với gen ức chế gây câm. Polypeptit mã hóa có thể có vai trò trong chuyển hóa dầu, tinh bột, carbohydrat, chất dinh dưỡng, v.v., hoặc có thể có vai trò trong quá trình tổng hợp protein, peptit, axit béo, lipid, sáp, dầu, tinh bột, đường, carbohydrat, chất điều vị, chất tạo mùi, độc tố, carotenoid, hormon, polyme, flavonoid, protein tích trữ, axit phenolic, alkaloid, lignin, tanin, xenluloza, glycoprotein, glycolipit, v.v., tốt hơn là quá trình sinh tổng hợp hoặc lắp ráp TAG.

Trong ví dụ cụ thể, cây tạo ra các mức enzym tăng lên để sinh ra dầu ở thực vật như Brassicas, ví dụ cải dầu hoặc hướng dương, rum, lanh, bông, đậu nành, *Camelina* hoặc ngô.

Mức LC-PUFA tạo ra

Mức LC-PUFA hoặc hỗn hợp các LC-PUFA được tạo ra trong tế bào tái tổ hợp hoặc bộ phận cây như hạt là rất quan trọng. Mức này có thể được thể hiện là thành phần (phần trăm) toàn bộ axit béo là LC-PUFA cụ thể hoặc nhóm của LC-PUFA cùng loại, ví dụ $\omega 3$ LC-PUFA hoặc $\omega 6$ LC-PUFA, hoặc VLC-PUFA, hoặc axit béo khác có thể được xác định bởi phương pháp đã biết trong lĩnh vực này. Mức này cũng có thể được thể hiện là hàm lượng LC-PUFA, như ví dụ phần trăm của LC-PUFA trong trọng lượng khô của nguyên liệu chứa tế bào tái tổ hợp, ví dụ phần trăm theo trọng lượng hạt là LC-PUFA. Sẽ biết rõ rằng LC-PUFA được tạo ra trong hạt dầu có thể cao hơn nhiều về hàm lượng LC-PUFA so với trong cây hoặc hạt không được trồng để sản xuất dầu, cả hai có thể có thành phần LC-PUFA giống nhau, và cả hai có thể được sử dụng làm nguồn LC-PUFA cho nhu cầu tiêu thụ của người hoặc động vật.

Các mức LC-PUFA có thể được xác định bằng phương pháp bất kỳ đã biết trong lĩnh vực này. Theo một phương án được ưu tiên, toàn bộ lipid được chiết từ tế bào, mô hoặc sinh vật và axit béo chuyển hóa thành este metyl trước khi phân tích bằng sắc ký khí (GC). Các kỹ thuật này được mô tả trong Ví dụ 1. Vị trí đỉnh trong sắc ký đồ có thể được sử dụng để xác định mỗi axit béo cụ thể, và kết hợp diện tích dưới mỗi đỉnh để định lượng. Như sử dụng trong đây, trừ khi có quy định khác, phần trăm của axit béo cụ thể trong mẫu được xác định là diện tích dưới đỉnh của axit béo là phần trăm của tổng diện tích của các axit béo trong sắc ký đồ. Nó về cơ bản tương ứng với phần trăm trọng lượng (trọng lượng/trọng lượng). Độ tương đồng của các axit béo có thể được xác nhận

bằng GC-MS. Toàn bộ lipit có thể được phân tách bởi các kỹ thuật đã biết trong lĩnh vực này để tinh chế các phần như phần TAG. Ví dụ, sắc ký lớp mỏng (TLC) có thể thực hiện ở quy mô phân tích để phân tách TAG từ các phân đoạn lipit như DAG, axyl-CoA hoặc phospholipit để xác định thành phần axit béo cụ thể của TAG.

Theo một phương án, toàn bộ tổng của ARA, EPA, DPA và DHA trong axit béo trong lipit chiết tách nằm trong khoảng từ 7% đến 25% của toàn bộ axit béo trong tế bào. Theo phương án mở rộng, toàn bộ axit béo trong tế bào có C20:1 với lượng ít hơn 1%. Theo các phương án được ưu tiên, TAG có thể chiết được trong tế bào chứa các axit béo ở mức được đề cập trong đây. Cũng bao gồm mỗi cách kết hợp có thể các đặc điểm xác định lipit như mô tả trong đây.

Mức sản sinh LC-PUFA trong tế bào tái tổ hợp, cây hoặc bộ phận cây như hạt cũng có thể được thể hiện là phần trăm chuyển hóa của cơ chất axit béo thành một hoặc nhiều sản phẩm axit béo, cũng được gọi là “hiệu suất chuyển hóa” hoặc “hiệu suất enzym”. Thông số này là dựa trên thành phần axit béo trong lipit chiết tách từ tế bào, cây, bộ phận cây hoặc hạt, tức là, lượng LC-PUFA tạo thành (bao gồm LC-PUFA khác có nguồn gốc từ chúng) là phần trăm của một hoặc nhiều cơ chất axit béo (bao gồm tất cả các axit béo có nguồn gốc từ chúng). Công thức chung cho phần trăm chuyển hóa là: $100 \times (\text{tổng các phần trăm của sản phẩm LC-PUFA và tất cả sản phẩm có nguồn gốc từ chúng}) / (\text{tổng các phần trăm của cơ chất axit béo và tất cả các sản phẩm có nguồn gốc từ chúng})$. Về DHA, ví dụ, nó có thể được thể hiện là tỷ lệ của mức DHA (là phần trăm trong tổng hàm lượng axit béo trong lipit) với mức cơ chất axit béo (ví dụ OA, LA, ALA, SDA, ETA hoặc EPA) và tất cả các sản phẩm không phải là DHA có nguồn gốc từ cơ chất này. Phần trăm chuyển hóa hoặc hiệu suất chuyển hóa có thể được thể hiện cho bước enzym riêng lẻ trong con đường chuyển hóa, hoặc một phần hoặc toàn bộ con đường chuyển hóa.

Hiệu suất chuyển hóa cụ thể được tính toán ở đây theo công thức:

1. $OA \text{ thành DHA} = 100 \times (\%DHA) / (\text{tổng \% của OA, LA, GLA, DGLA, ARA, EDA, ALA, SDA, ETrA, ETA, EPA, DPA và DHA})$.
2. $LA \text{ thành DHA} = 100 \times (\%DHA) / (\text{tổng \% của LA, GLA, DGLA, ARA, EDA, ALA, SDA, ETrA, ETA, EPA, DPA và DHA})$.
3. $ALA \text{ thành DHA} = 100 \times (\%DHA) / (\text{tổng \% của ALA, SDA, ETrA, ETA, EPA, DPA và DHA})$.

4. EPA thành DHA = $100 \times (\% \text{DHA}) / (\text{tổng } \% \text{ của EPA, DPA và DHA})$.
5. DPA thành DHA (hiệu suất $\Delta 4$ -desaturaza) = $100 \times (\% \text{DHA}) / (\text{tổng } \% \text{ của DPA và DHA})$.
6. hiệu suất $\Delta 12$ -desaturaza = $100 \times (\text{tổng } \% \text{ của LA, GLA, DGLA, ARA, EDA, ALA, SDA, ETrA, ETA, EPA, DPA và DHA}) / (\text{tổng } \% \text{ của OA, LA, GLA, DGLA, ARA, EDA, ALA, SDA, ETrA, ETA, EPA, DPA và DHA})$.
7. hiệu suất $\omega 3$ -desaturaza = $100 \times (\text{tổng } \% \text{ của ALA, SDA, ETrA, ETA, EPA, DPA và DHA}) / (\text{tổng } \% \text{ của LA, GLA, DGLA, ARA, EDA, ALA, SDA, ETrA, ETA, EPA, DPA và DHA})$.
8. OA thành ALA = $100 \times (\text{tổng } \% \text{ của ALA, SDA, ETrA, ETA, EPA, DPA và DHA}) / (\text{tổng } \% \text{ của OA, LA, GLA, DGLA, ARA, EDA, ALA, SDA, ETrA, ETA, EPA, DPA và DHA})$.
9. Hiệu suất $\Delta 6$ -desaturaza (trên cơ chất $\omega 3$ ALA) = $100 \times (\text{tổng } \% \text{ của SDA, ETA, EPA, DPA và DHA}) / (\% \text{ALA, SDA, ETrA, ETA, EPA, DPA và DHA})$.
10. Hiệu suất $\Delta 6$ -elongaza (trên cơ chất $\omega 3$ SDA) = $100 \times (\text{tổng } \% \text{ của ETA, EPA, DPA và DHA}) / (\text{tổng } \% \text{ của SDA, ETA, EPA, DPA và DHA})$.
11. Hiệu suất $\Delta 5$ -desaturaza (trên cơ chất $\omega 3$ ETA) = $100 \times (\text{tổng } \% \text{ của EPA, DPA và DHA}) / (\text{tổng } \% \text{ của ETA, EPA, DPA và DHA})$.
12. Hiệu suất $\Delta 5$ -elongaza (trên cơ chất $\omega 3$ EPA) = $100 \times (\text{tổng } \% \text{ của DPA và DHA}) / (\text{tổng } \% \text{ của EPA, DPA và DHA})$.

Thành phần axit béo của lipit, tốt hơn là dầu hạt, theo sáng chế, cũng có đặc trưng bởi tỷ lệ của axit béo $\omega 6$:axit béo $\omega 3$ trong tổng hàm lượng axit béo, hoặc là tỷ lệ của toàn bộ axit béo $\omega 6$:toàn bộ axit béo $\omega 3$ hoặc tỷ lệ của axit béo $\omega 6$ mới:axit béo $\omega 3$ mới. Thuật ngữ toàn bộ axit béo $\omega 6$, toàn bộ axit béo $\omega 3$, axit béo $\omega 6$ mới và axit béo $\omega 3$ mới có ý nghĩa như được nêu trong đây. Các tỷ lệ này được tính toán từ thành phần axit béo trong lipit được chiết từ tế bào, cây, bộ phận cây hoặc hạt, theo cách được minh họa trong đây. Mong muốn có mức axit béo $\omega 3$ cao hơn axit béo $\omega 6$ trong lipit, và do đó ưu tiên tỷ lệ $\omega 6$: $\omega 3$ là thấp hơn 1,0. Tỷ lệ 0,0 chỉ ra rằng hoàn toàn không có axit béo $\omega 6$ được nêu; đạt được tỷ lệ 0,03 như mô tả trong Ví dụ 6. Các tỷ lệ thấp này có thể được đạt được bằng cách sử dụng kết hợp $\Delta 6$ -desaturaza ưu tiên cơ chất $\omega 3$ cùng với $\omega 3$ -desaturaza, đặc biệt là $\omega 3$ -desaturaza nấm như $\omega 3$ -desaturaza *Pichia pastoris* như

được minh họa trong đây.

Hiệu suất của LC-PUFA trên trọng lượng hạt cũng có thể tính trên tổng hàm lượng dầu trong hạt và %DHA trong dầu. Ví dụ, nếu hàm lượng dầu của hạt cải dầu là khoảng 40% (trọng lượng/trọng lượng) và khoảng 12% tổng hàm lượng axit béo của dầu là DHA, hàm lượng DHA của hạt là khoảng 4,8% hoặc khoảng 48mg trong mỗi gam hạt. như mô tả trong Ví dụ 2, hàm lượng DHA của hạt *Arabidopsis* có khoảng 9% DHA, nó có hàm lượng dầu thấp hơn cải dầu, là khoảng 25mg/g hạt. Với hàm lượng DHA khoảng 7%, hạt cải dầu hoặc hạt *Camelina sativa* có hàm lượng DHA là khoảng 28mg trong mỗi gam hạt. Do đó, sáng chế đề xuất cây *Brassica napus*, *B. juncea* và *Camelina sativa*, và hạt thu được từ cây này, chứa ít nhất khoảng 28mg DHA trong mỗi gam hạt. Hạt có hàm lượng ẩm theo chuẩn với hạt già thu hoạch được sau khi sấy khô (độ ẩm 4-15%). Sáng chế cũng đề xuất quy trình để thu được dầu, quy trình này bao gồm thu lấy hạt và chiết dầu từ hạt, và sử dụng dầu và phương pháp thu được hạt bao gồm thu hoạch hạt từ cây theo sáng chế.

Lượng DHA tạo ra trên mỗi héc ta cũng có thể được tính nếu hạt thu được trên mỗi héc ta được biết hoặc có thể ước tính. Ví dụ, cải dầu ở Úc đặc trưng có hiệu suất khoảng 2,5 tấn hạt trên mỗi héc ta, với hàm lượng dầu 40% sẽ cho khoảng 1000kg dầu. Với hàm lượng 12% DHA trong toàn bộ dầu, nó sẽ cho khoảng 120kg DHA trên mỗi héc ta. Nếu hàm lượng dầu giảm 50%, nó vẫn cung cấp khoảng 60kg DHA/ha.

Bằng chứng cho đến nay gợi ý rằng một số desaturaza được biểu hiện khác loại trong nấm men hoặc cây có hoạt tính tương đối thấp khi kết hợp với một số elongaza. Điều này có thể được giảm đi bằng cách mang lại cho desaturaza khả năng sử dụng dạng axyl-CoA của axit béo làm cơ chất trong quá trình tổng hợp LC-PUFA, và điều này được cho là thuận lợi trong tế bào tái tổ hợp đặc biệt là trong tế bào thực vật. Một cách kết hợp bất kỳ thuận lợi để tổng hợp DHA hiệu quả là ω 3-desaturaza nấm, ví dụ như ω 3-desaturaza *Pichia pastoris* (SEQ ID NO: 12), với Δ 6-desaturaza có sự ưu tiên cho cơ chất ω 3 axyl như, ví dụ, 6-desaturaza *Micromonas pusilla* Δ (SEQ ID NO: 13), hoặc biến thể của nó có độ tương đồng trình tự axit amin ít nhất là 95%.

Như sử dụng trong đây, thuật ngữ “về cơ bản không có” có nghĩa là chế phẩm (ví dụ lipid hoặc dầu) chứa ít (ví dụ, thấp hơn khoảng 0,5%, thấp hơn khoảng 0,25%, thấp hơn khoảng 0,1%, hoặc thấp hơn khoảng 0,01%) hoặc không có thành phần được nêu. Theo một phương án, “về cơ bản không có” có nghĩa là thành phần này không thể phát

hiện được bằng kỹ thuật phân tích thông thường, ví dụ axit béo cụ thể (như axit ω -docosapentaenoic) không thể được phát hiện bằng phương pháp sắc ký khí như sơ lược trong Ví dụ 1.

Sản xuất dầu

Các kỹ thuật được thực hành thông thường trong lĩnh vực này có thể được sử dụng để chiết, xử lý, và phân tích dầu được tạo ra bởi tế bào, cây, hạt, v.v. của theo sáng chế. Đặc trưng, hạt cây được đun, ép, và chiết để tạo ra dầu thô, sau đó dầu thô được loại gôm, tinh lọc, tẩy màu, và khử mùi. Thông thường, các kỹ thuật ép hạt đã được biết trong lĩnh vực này. Ví dụ, hạt dầu có thể được gia nhiệt bằng cách phun nước vào chúng để tăng hàm lượng ẩm đến, ví dụ, 8,5%, và cán thành tấm bằng máy cán chạy êm với khoảng cách được đặt là từ 0,23 đến 0,27mm. Phụ thuộc vào kiểu hạt, có thể không bổ sung nước trước khi ép. Áp dụng nhiệt để bất hoạt enzym, thúc đẩy làm vỡ tế bào, kết tụ các giọt dầu, và kết tụ các hạt protein, tất cả chúng thuận lợi cho quy trình chiết.

Theo một phương án, phần lớn dầu hạt được giải phóng bằng cách đi qua một máy ép kiểu vít. Các bánh được đùn từ máy ép kiểu vít sau đó được chiết dung môi, ví dụ, bằng hexan, sử dụng cột ghi nhiệt. Cách khác, dầu thô tạo ra bằng quy trình ép có thể được cho đi qua bề lằng phía trên có dây hút nước có rãnh để loại chất rắn được ép ra cùng với dầu trong quá trình ép. Dầu trong có thể được đi qua đĩa và khung lọc để loại bỏ các hạt rắn mịn còn lại bất kỳ. Nếu muốn, dầu thu hồi được từ quy trình chiết có thể được kết hợp với dầu trong để tạo thành hỗn hợp dầu thô.

Khi dung môi được tháo ra khỏi dầu thô, các phần đã ép và chiết được gộp lại và đưa vào quy trình xử lý dầu bình thường. Như sử dụng trong đây, thuật ngữ "tinh khiết/đã tinh chế" khi được sử dụng với lipit hoặc dầu theo sáng chế đặc trưng có nghĩa là lipit chiết tách hoặc dầu đã được đưa vào một hoặc nhiều bước xử lý để làm tăng độ tinh khiết của thành phần lipit/dầu. Ví dụ, bước tinh chế có thể bao gồm một hoặc nhiều hoặc tất cả các bước sau: loại gôm, khử mùi, tẩy màu, sấy và/hoặc cất phân đoạn dầu chiết được. Tuy nhiên, như sử dụng trong đây, thuật ngữ "tinh khiết/đã tinh chế" không bao gồm quy trình chuyển hóa este hoặc quy trình khác làm thay đổi thành phần axit béo của lipit hoặc dầu theo sáng chế sao cho làm tăng hàm lượng DHA theo phần trăm tổng hàm lượng axit béo. Nói cách khác, thành phần axit béo của lipit hoặc dầu tinh chế về cơ bản giống với lipit hoặc dầu không được tinh chế.

Loại gôm

Quy trình loại gôm là quy trình trước trong quá trình tinh lọc dầu và mục đích chính của nó là loại phần lớn phospholipit khỏi dầu, chúng có thể có mặt với lượng xấp xỉ 1-2% của toàn bộ lipit chiết tách. Bổ sung ~2% là nước, đặc trưng chứa axit phosphoric, ở 70–80°C vào dầu thô sẽ gây tách phần lớn phospholipit đi kèm theo kim loại vi lượng và thuốc nhuộm. Nguyên liệu không tan được loại đi phần lớn là hỗn hợp của phospholipit và triaxylglyxerol và cũng được biết là lexitin. Việc loại gôm có thể được thực hiện bằng cách bổ sung axit phosphoric đặc vào dầu hạt thô để chuyển hóa các phosphatit không hydrat hóa được thành dạng hydrat hóa được, và để tạo chelat các kim loại vi lượng có trong đó. Gôm được tách khỏi dầu hạt bằng cách ly tâm.

Tinh chế bằng kiềm (tinh cất dầu)

Quy trình tinh chế kiềm là một quy trình tinh lọc để xử lý dầu thô, đôi khi cũng được gọi là quy trình tinh cất dầu. Nó thường là sau quy trình loại gôm và trước quy trình tẩy màu. Sau khi loại gôm, dầu hạt có thể được xử lý bằng cách bổ sung một lượng vừa đủ dung dịch kiềm để chuẩn độ tất cả các axit béo và axit phosphoric, và loại bỏ xà phòng tạo ra theo cách này. Nguyên liệu kiềm thích hợp bao gồm natri hydroxit, kali hydroxit, natri carbonat, liti hydroxit, canxi hydroxit, canxi carbonat và amoni hydroxit. Quy trình này đặc trưng được tiến hành ở nhiệt độ phòng và loại phần axit béo tự do. Xà phòng được loại bỏ bằng cách ly tâm hoặc bằng cách chiết vào trong dung môi chứa xà phòng, và dầu trung hòa được rửa bằng nước. Nếu đòi hỏi, kiềm dư bất kỳ trong dầu có thể được trung hòa bằng axit thích hợp như axit clohydric hoặc axit sulphuric.

Tẩy màu

Quy trình tẩy màu là một quy trình lọc trong đó dầu được làm ấm ở nhiệt độ 90–120°C trong 10–30 phút trong sự có mặt của đất tẩy màu (0,2–2,0%) và trong sự không có mặt của oxy bằng cách xử lý với nitơ hoặc hơi nước hoặc trong chân không. Bước này trong quá trình xử lý dầu được lập nên để loại các chất màu không mong muốn (carotenoit, diệp lục, gossypol v.v.), và quy trình này cũng loại những sản phẩm oxy hóa, kim loại vi lượng, hợp chất lưu huỳnh và vết xà phòng.

Khử mùi

Quá trình khử mùi là việc xử lý dầu và mỡ ở nhiệt độ cao (200–260°C) và áp suất thấp (0,1–1mm Hg). Đặc trưng có thể khử mùi bằng cách đưa hơi nước vào trong

dầu hạt với tốc độ khoảng 0,1 mL/phút/100ml dầu hạt. Sau khoảng 30 phút sục, dầu hạt được để cho lạnh trong điều kiện chân không. Đặc trưng, dầu hạt được chuyển vào vật chứa thủy tinh và xịt bằng argon trước khi được bảo quản trong điều kiện làm lạnh. Việc xử lý này làm tăng màu của dầu hạt và loại đi phần lớn chất dễ bay hơi hoặc các hợp chất có mùi bao gồm các axit béo tự do còn dư bất kỳ, monoaxylglyxerol và sản phẩm oxy hóa.

Đông cứng

Việc đông cứng (tạo môi trường có nhiệt độ thấp hơn môi trường) là quy trình đôi khi được sử dụng trong sản xuất thương mại dầu để tách dầu và mỡ thành các đoạn rắn (stearin) và lỏng (olein) bằng cách kết tinh ở nhiệt độ thấp hơn nhiệt độ môi trường. Nó được áp dụng đầu tiên với dầu hạt bông để tạo ra sản phẩm không có chất rắn. Đặc trưng, nó được sử dụng để làm giảm hàm lượng axit béo bão hòa của dầu.

Chuyển hóa este

Quy trình chuyển hóa este là một quy trình trao đổi các axit béo trong và giữa các TAG hoặc chuyển axit béo vào rượu khác để tạo ra một este, bắt đầu bằng việc giải phóng axit béo từ TAG hoặc ở dạng axit béo tự do hoặc este của axit béo, thường là este metyl hoặc este etyl của axit béo. Khi kết hợp với quy trình cắt phân đoạn, quy trình chuyển hóa este có thể được sử dụng để biến đổi chế phẩm axit béo của lipid (Marangoni et al., 1995). Quy trình chuyển hóa este có thể sử dụng phương pháp hóa học (ví dụ xúc tác bằng axit hoặc bazơ mạnh) hoặc phương pháp enzym, phương pháp enzym sử dụng lipaza có thể đặc hiệu ở vị trí (đặc hiệu ở *sn*-1/3 hoặc *sn*-2) với axit béo trên TAG, hoặc có sự ưu tiên với một số axit béo hơn so với các axit béo khác (Speranza et al, 2012). Việc cắt phân đoạn axit béo để làm tăng nồng độ của LC-PUFA trong dầu có thể đạt được bằng phương pháp bất kỳ đã biết trong lĩnh vực này, như, ví dụ, phương pháp kết tinh đông lạnh, tạo phức bằng ure, chưng cất phân tử, chiết lỏng siêu tới hạn và tạo phức ion bạc. Phương pháp tạo phức bằng urê là phương pháp được ưu tiên vì sự đơn giản của nó và hiệu quả làm giảm mức axit béo bão hòa và một liên kết chưa bão hòa trong dầu (Gamez et al., 2003). Ban đầu, TAG của dầu được chia thành các axit béo cấu trúc của chúng, thường là ở dạng este của axit béo, bằng cách thủy phân trong điều kiện phản ứng được xúc tác bởi axit hoặc bazơ, bằng cách đó 1 mol TAG phản ứng với ít nhất 3 mol rượu (ví dụ etanol với este etyl hoặc metanol với este metyl) với rượu dư được sử dụng để có thể tách alkyl este đã tạo ra và glyxerol

cũng thường được tạo ra, hoặc bởi lipaza. Các axit béo tự do hoặc este của axit béo, là những thành phần thường không bị biến đổi trong thành phần axit béo bởi sự xử lý, sau đó có thể được phối hợp với dung dịch etanol chứa urê để tạo phức. Các axit béo bão hòa và một liên kết chưa bão hòa tạo phức dễ dàng với urê và kết tinh xuống khi làm lạnh và sau đó có thể được loại đi bằng cách lọc. Phần phức không phải với urê được làm giàu bằng cách này với LC-PUFA.

Sản phẩm dinh dưỡng

Sáng chế bao gồm chế phẩm có thể sử dụng làm sản phẩm dinh dưỡng. Theo mục đích của sáng chế, "sản phẩm dinh dưỡng" bao gồm thức ăn bất kỳ hoặc chế phẩm cho người hoặc động vật khi được đưa vào cơ thể (a) sử dụng để nuôi dưỡng hoặc xây dựng mô hoặc cung cấp năng lượng; và/hoặc (b) duy trì, giữ hoặc củng cố trạng thái đầy đủ dinh dưỡng hoặc chức năng chuyển hóa. Sản phẩm dinh dưỡng theo sáng chế bao gồm các chế phẩm dinh dưỡng cho trẻ em và/hoặc trẻ nhỏ như, ví dụ, chế phẩm công thức cho trẻ em, và bột hạt theo sáng chế.

Sản phẩm dinh dưỡng theo sáng chế bao gồm, ví dụ, tế bào theo sáng chế, cây theo sáng chế, bộ phận cây theo sáng chế, hạt theo sáng chế, sản phẩm chiết theo sáng chế, sản phẩm của phương pháp theo sáng chế, sản phẩm của quy trình lên men theo sáng chế, hoặc chế phẩm cùng với (các) chất mang thích hợp. Thuật ngữ "chất mang" được sử dụng với nghĩa rộng nhất của nó để bao gồm thành phần bất kỳ có thể có hoặc không có giá trị dinh dưỡng. Các chuyên gia trong lĩnh vực này sẽ biết rõ, chất mang này phải thích hợp để sử dụng (hoặc được sử dụng với hàm lượng thấp vừa đủ) trong sản phẩm dinh dưỡng sao cho nó không có ảnh hưởng bất lợi lên sinh vật tiêu thụ sản phẩm dinh dưỡng này.

Sản phẩm dinh dưỡng theo sáng chế bao gồm dầu, este của axit béo, hoặc axit béo được tạo ra trực tiếp hoặc gián tiếp bằng cách sử dụng phương pháp, tế bào hoặc cây bộc lộ trong đây. Chế phẩm này có thể ở dạng lỏng hoặc rắn. Ngoài ra, sản phẩm này có thể bao gồm các nguyên tố dinh dưỡng đa lượng tiêu hóa được, protein, carbohydrat, vitamin, và/hoặc khoáng chất với các lượng mong muốn cho mục đích sử dụng cụ thể. Lượng của các thành phần này sẽ thay đổi phụ thuộc vào việc liệu sản phẩm này được dự tính để sử dụng cho cá thể bình thường hay sử dụng cho cá thể có nhu cầu đặc biệt, như cá thể mắc bệnh chuyển hóa và bệnh tương tự.

Ví dụ về các chất mang thích hợp có giá trị dinh dưỡng bao gồm, nhưng không

hạn chế ở, nguyên tố dinh dưỡng đa lượng như các chất béo ăn được, carbohydrat và protein. Ví dụ về chất béo ăn được bao gồm, nhưng không hạn chế ở, dầu dừa, dầu borage, dầu nấm, dầu nho đen, dầu đậu nành, và mono- và diglycerit. Ví dụ về carbohydrat như vậy bao gồm (nhưng không hạn chế ở): glucoza, lactoza tiêu hóa được, và tinh bột thủy phân. Ngoài ra, ví dụ về protein có thể được sử dụng trong chế phẩm dinh dưỡng theo sáng chế bao gồm (nhưng không hạn chế ở) protein đậu nành, nước sữa được điện thẩm tích, váng sữa được điện thẩm tích, nước sữa, hoặc các sản phẩm thủy phân chứa các protein này.

Về vitamin và khoáng chất, những chất sau có thể được bổ sung vào sản phẩm dinh dưỡng theo sáng chế: canxi, phospho, kali, natri, clorua, magie, mangan, iron, đồng, kẽm, selen, iot, và vitamin A, E, D, C, và phức vitamin B. Các vitamin và khoáng chất khác cũng có thể được bổ sung vào.

Các thành phần sử dụng trong sản phẩm dinh dưỡng theo sáng chế có thể là có nguồn gốc bán tinh chế hoặc tinh chế. Bán tinh chế hoặc tinh chế có nghĩa là nguyên liệu đã được điều chế bằng cách tinh chế nguyên liệu tự nhiên hoặc bằng cách tổng hợp *de novo*.

Sản phẩm dinh dưỡng theo sáng chế cũng có thể được bổ sung vào thức ăn thậm chí khi không đòi hỏi bổ sung thêm vào chế độ ăn. Ví dụ, sản phẩm này có thể được vào thức ăn dạng bất kỳ, bao gồm (nhưng không giới hạn ở): bơ thực vật, bơ chế biến, pho mát, sữa, sữa chua, sôcôla, kẹo, đồ ăn nhẹ, dầu salat, dầu nấu ăn, mỡ nấu ăn, thịt lợn, cá và đồ uống.

Giống *Saccharomyces* spp được sử dụng trong cả việc ủ bia và rượu và cũng là tác nhân làm bánh, cụ thể là bánh mỳ. Các nấm men khác như nấm men dầu bao gồm, ví dụ, *Yarrowia* spp, cũng hữu dụng trong việc sản xuất LC-PUFA. Nấm men có thể được sử dụng làm chất phụ gia trong thức ăn cho động vật, như trong ngành nuôi trồng thủy sản. Rõ ràng là các chủng nấm men biến đổi gen có thể được tạo ra được biến đổi phù hợp để tổng hợp LC-PUFA như mô tả trong đây. Các chủng nấm men này, hoặc tạo ra LC-PUFA trong đây, sau đó có thể sử dụng trong các sản phẩm thức ăn và trong sản xuất rượu và bia để thu được sản phẩm có hàm lượng axit béo tăng lên.

Ngoài ra, các axit béo được tạo ra theo sáng chế hoặc tế bào chủ được gây biến nạp để chứa và biểu hiện gen đích cũng có thể được sử dụng làm nguồn bổ sung thức ăn cho động vật để thay đổi thành phần axit béo của mô, trứng hoặc sữa của động vật đến

mức mong muốn hơn để tiêu thụ ở người hoặc động vật. Ví dụ về động vật này bao gồm cừu, gia súc, ngựa, gia cầm như gà và tương tự.

Hơn nữa, sản phẩm dinh dưỡng theo sáng chế có thể được sử dụng trong nuôi trồng thủy sản để làm tăng mức axit béo trong cá hoặc động vật giáp xác như, ví dụ, tôm pandan để tiêu thụ ở người hoặc động vật. Cá được ưu tiên là cá hồi.

Sản phẩm dinh dưỡng được ưu tiên theo sáng chế là cây, hạt và bộ phận cây khác như lá và thân có thể được sử dụng trực tiếp làm thức ăn cho người hoặc động vật khác. Ví dụ, động vật có thể ăn trực tiếp trên cây đang trồng trên cánh đồng hoặc được cho ăn thêm các lượng xác định khi chăn nuôi kiểm soát. Sáng chế bao gồm việc sử dụng cây và bộ phận cây như vậy làm thức ăn để làm tăng mức LC-PUFA ở người và động vật khác.

Chế phẩm

Sáng chế cũng đề xuất các chế phẩm, đặc biệt là dược phẩm, chứa một hoặc nhiều axit béo và/hoặc các dầu thu được bằng phương pháp theo sáng chế.

Dược phẩm có thể chứa một hoặc nhiều axit béo và/hoặc dầu, kết hợp với chất mang, tá dược hoặc tá dược lỏng được dụng chuẩn, đã biết rõ, không độc như nước muối đệm phosphat, nước, etanol, rượu polyhydric, dầu thực vật, tác nhân tạo ẩm hoặc nhũ tương như nhũ tương dầu/nước. Dược phẩm này có thể là ở dạng rắn hoặc dạng lỏng. Ví dụ, dược phẩm có thể là dạng viên nén, viên nang, bột hoặc chất lỏng tiêu hóa được, tiêm được, thuốc hoặc kem dùng tại chỗ. Có thể duy trì độ lỏng, ví dụ, bằng cách duy trì cỡ hạt yêu cầu trong trường hợp là hệ phân tán và bằng cách sử dụng chất hoạt động bề mặt. Cũng có thể mong muốn đưa các tác nhân đẳng trương vào, ví dụ, đường, natri clorua, và tương tự. Bên cạnh các chất pha loãng trở như vậy, dược phẩm cũng có thể bao gồm các tá dược, như tác nhân tạo ẩm, tác nhân nhũ hóa và tạo huyền phù, tác nhân làm ngọt, tác nhân điều vị và tác nhân tạo mùi.

Huyền phù, ngoài các chất hoạt động, có thể bao gồm tác nhân tạo huyền phù như etoxyl hóa rượu isostearic, polyoxyetylen sorbitol và sorbitan este, xenluloza vi tinh thể, nhôm metahydroxit, bentonit, aga-aga, và nhựa tragacan hoặc hỗn hợp của các chất này.

Dạng liều rắn như viên nén và viên nang có thể được bào chế bằng các kỹ thuật đã biết trong lĩnh vực này. Ví dụ, các axit béo được sản xuất theo sáng chế có thể được tạo viên nén với các chất nền viên nén thông dụng như lactoza, sucroza, và tinh bột ngô

kết hợp với chất gắn kết như acacia, tinh bột ngô hoặc gelatin, tá dược rã như tinh bột khoai tây hoặc axit alginic, và tá dược trơn chảy như axit stearic hoặc magie stearat. Viên nang có thể được bào chế bằng cách đưa các tá dược này vào viên nang gelatin cùng với chất chống oxy hóa và (các) axit béo liên quan.

Với đường dùng trong tĩnh mạch, các axit béo được tạo ra theo sáng chế hoặc dẫn xuất của chúng có thể được đưa vào các chế phẩm thương mại.

Liều đặc trưng của axit béo cụ thể nằm trong khoảng từ 0,1mg đến 20 g, dùng từ 1 đến 5 lần mỗi ngày (lên đến 100g hằng ngày) và tốt hơn nằm trong khoảng từ 10mg đến 1, 2, 5, hoặc 10g hằng ngày (dùng một hoặc nhiều liều). Đã biết trong lĩnh vực này, mong muốn tối thiểu khoảng 300 mg axit béo/ngày, đặc biệt là LC-PUFA,. Tuy nhiên, cũng biết rõ là lượng axit béo bất kỳ đều có lợi cho đối tượng sử dụng.

Đường dùng hữu dụng của dược phẩm theo sáng chế bao gồm, ví dụ, đường dùng trong ruột (ví dụ, đường miệng và trực tràng) và ngoài đường tiêu hóa. Ví dụ, chế phẩm lỏng có thể được sử dụng qua đường miệng hoặc tại trực tràng. Ngoài ra, hỗn hợp đồng nhất có thể được phân bố hoàn toàn trong nước, trộn trong điều kiện vô khuẩn với chất pha loãng sinh lý, chất bảo quản, chất đệm hoặc chất đẩy để tạo ra thuốc phun hoặc thuốc hít.

Liều dược phẩm sử dụng cho bệnh nhân có thể được xác định bởi các chuyên gia trong lĩnh vực này và phụ thuộc các yếu tố khác nhau trọng lượng của bệnh nhân, tuổi của bệnh nhân, sức khỏe tổng thể của bệnh nhân, tiền sử của bệnh nhân, trạng thái miễn dịch của bệnh nhân, v.v.

Ngoài ra, dược phẩm theo sáng chế có thể được sử dụng với mục đích làm đẹp. Nó có thể được bổ sung vào các chế phẩm mỹ phẩm đã có sao cho tạo thành hỗn hợp hoặc axit béo tạo ra theo sáng chế có thể được sử dụng làm thành phần "hoạt động" duy nhất trong mỹ phẩm.

Ví dụ thực hiện sáng chế

Ví dụ 1. Nguyên liệu và phương pháp

Biểu hiện gen ở tế bào thực vật trong hệ biểu hiện tạm thời

Các cấu trúc gen ngoại sinh được biểu hiện trong các tế bào thực vật trong hệ biểu hiện tạm thời như được mô tả về cơ bản bởi Voinnet et al. (2003) và Wood et al.

(2009). Plasmid chứa vùng mã hóa được biểu hiện từ gen khởi đầu có tính bảo tồn mạnh như gen khởi đầu CaMV 35S được đưa vào trong chủng *Agrobacterium tumefaciens* AGL1. Gen khảm 35S:p19 để biểu hiện gen ức chế làm câm của virus p19 được đưa riêng biệt vào trong AGL1, như mô tả trong WO 2010/057246. Tế bào *Agrobacterium* tái tổ hợp được phát triển ở 28°C trong nước thịt LB được bổ sung 50mg/L kanamycin và 50mg/l rifampixin cho đến pha tĩnh. Sau đó vi khuẩn được tạo viên bằng cách ly tâm ở tốc độ 5000 g trong 15 phút ở nhiệt độ phòng trước khi được tạo huyền phù lại đến giá trị OD600 = 1,0 trong đệm thẩm lọc chứa 10mM MES độ pH =5,7, MgCl₂ 10mM và acetosyringon 100µM. Sau đó, các tế bào được ủ ở 28°C đồng thời lắc trong 3 giờ trước khi các thể tích tương đương của *Agrobacterium* nuôi cấy chứa 35S:p19 và (các) cấu trúc khảm thử nghiệm đang quan tâm được phối hợp trước khi thẩm vào trong mô lá. Đặc trưng, cây được phát triển trong 5 ngày nữa sau khi thẩm lọc trước khi thu lấy phiến lá và làm đông lạnh để phân tích GC các axit béo.

Este metyl của axit béo (FAME) của toàn bộ lipid lá trong các mẫu đông lạnh được tạo ra bằng cách ủ mẫu trong dung dịch metanol/HCl/diclorometan (10/1/1 thể tích/thể tích) trong 2 giờ ở 80°C cùng với một lượng đã biết axit hexadecanoic là chất chuẩn nội. FAME được chiết trong hexan/DCM, cô đến thể tích nhỏ trong hexan và bơm vào GC. Lượng các axit béo riêng biệt và toàn bộ axit béo có trong các phân đoạn lipid được định lượng dựa trên lượng chất chuẩn nội đã biết.

Phân tích sắc ký khí (GC) axit béo

FAME được phân tích bằng sắc ký khí sử dụng Agilent Technologies 7890A GC (Palo Alto, California, USA) được trang bị cột 30m SGE-BPX70 (70% xyanopropyl polysilphenylen-siloxan, đường kính trong 0,25mm, chiều dày màng 0,25mm), FID, bơm tách/không tách và máy lấy mẫu và bơm tự động Agilent Technologies 7693 Series. Heli được sử dụng làm khí mang. Các mẫu được bơm theo kiểu tách (tỷ lệ 50:1) ở nhiệt độ lò là 150°C. Sau khi bơm, nhiệt độ lò được giữ ở 150°C trong 1 phút sau đó nâng lên 210°C với tốc độ 3°C.phút⁻¹, nâng tiếp đến 240°C với tốc độ 50°C.phút⁻¹ và cuối cùng giữ trong 1,4 phút ở 240°C. Các đỉnh được định lượng bằng phần mềm Agilent Technologies ChemStation (Rev B.04.03 (16), Palo Alto, California, USA) dựa trên đáp ứng của lượng đã biết của chất chuẩn ngoại GLC- 411 (Nuchek) và chất chuẩn nội C17:0-ME.

Phân tích sắc ký lỏng- Quang phổ khối lượng (LC-MS) của lipid

Toàn bộ lipit được chiết từ hạt đang phát triển đã được làm đông lạnh, mười hai ngày sau khi ra hoa (daf), và các hạt già sau khi bổ sung một lượng đã biết tri-C17:0-TAG làm chất chuẩn nội. Lipit chiết tách được hòa tan vào trong 1mL hydroxytoluen butyl hóa 10mM trong butanol:metanol (tỷ lệ 1:1 thể tích/thể tích) trên mỗi 5mg nguyên liệu khô và phân tích bằng LC-MS phổ ba tứ cực ion hóa tia điện Agilent 1200 series LC và 6410b. Các lipit được phân tách bằng sắc ký cột Ascentis Express RP- Amide (50 mm x 2,1 mm, 2,7 μ m, Supelco) hoạt động ở gradient kép với tốc độ dòng là 0,2mL/phút. Pha động là: A. Amoni format 10mM trong H₂O:metanol: tetrahydrofuran (50:20:30 thể tích/thể tích/thể tích); B. Amoni format 10mM trong H₂O:metanol: tetrahydrofuran (5:20:75, thể tích/thể tích/thể tích). Danh sách điều kiện giám sát nhiều phản ứng (MRM) dựa trên các axit béo chính sau đây: 16:0, 18:0, 18:1, 18:2, 18:3, 18:4, 20:1, 20:2, 20:3, 20:4, 20:5, 22:4, 22:5, 22:6 sử dụng năng lượng va chạm là 30V và năng lượng gây vỡ mảnh là 60V. MRM TAG riêng được xác định dựa trên tiền ion amoni hóa và sản phẩm ion từ sự mất trung hòa là 22:6. TAG được định lượng bằng chất chuẩn ngoại tristearin 10 μ M.

Xác định profin axit béo của hạt và hàm lượng dầu

Khi hàm lượng dầu trong hạt được xác định, các hạt được sấy khô trong bình hút ẩm trong 24 giờ và khoảng 4mg hạt được chuyển vào ống thủy tinh 2 ml Teflon được nối với nút xoáy. 0,05mg triheptadecanoin hòa tan trong 0,1ml toluen được bổ sung vào ống này làm chất chuẩn nội.

FAME trong hạt được điều chế bằng cách bổ sung 0,7ml metanolic HCl 1N (Supelco) vào ống chứa nguyên liệu hạt, quay nhanh và ủ ở 80°C trong 2 giờ. Sau khi làm mát đến nhiệt độ phòng, 0,3ml NaCl 0,9% (trọng lượng/thể tích) và 0,1ml hexan được bổ sung vào ống này và trộn kỹ trong 10 phút trong Heidolph Vibramax 110. FAME được gộp lại vào trong ống lồng thủy tinh 0,3 ml và phân tích bằng GC bằng bộ cảm biến ion hóa ngọn lửa (FID) như đã nêu trên.

Diện tích đỉnh của FAME riêng lẻ đầu tiên được hiệu chỉnh trên cơ sở kết quả diện tích đỉnh của lượng đã biết của các FAME tương tự có trong chất chuẩn GLC-411 thương mại (NU-CHEK PREP, INC., USA). GLC-411 chứa các lượng bằng nhau của 31 axit béo (% theo trọng lượng), nằm trong khoảng từ C8:0 đến C22:6. Trong trường hợp axit béo không có trong chất chuẩn, các tác giả lấy kết quả diện tích đỉnh của FAME giống nhất. Ví dụ, kết quả diện tích đỉnh của các FAME là 16:1d9 được sử dụng

cho 16:1d7 và kết quả FAME của C22:6 được sử dụng cho C22:5. Các diện tích đã hiện chỉnh được sử dụng để tính khối lượng của mỗi FAME trong mẫu bằng cách so sánh khối lượng chất chuẩn nội. Dầu được lưu trữ chủ yếu ở dạng TAG và khối lượng của nó được tính toán dựa trên khối lượng FAME. Số mol tổng cộng của glycerol được xác định bằng cách tính số mol của mỗi FAME và chia tổng số mol của FAME cho ba. TAG được tính là tổng số của glycerol và các gốc axyl béo bằng hệ thức: % dầu theo trọng lượng = $100 \times ((41 \times \text{tổng số mol FAME}/3) + (\text{tổng số g FAME} - (15 \times \text{tổng số mol FAME}))) / \text{g hạt}$, trong đó 41 và 15 lần lượt là trọng lượng phân tử của gốc glycerol và gốc methyl.

Phân tích hàm lượng sterol trong mẫu dầu

Các mẫu chứa xấp xỉ 10mg dầu cùng với lượng nhỏ C24:0 monol bổ sung làm chất chuẩn nội được saponin hóa bằng 4mL KOH 5% trong MeOH 80% và gia nhiệt trong 2 giờ ở 80°C trong ống thủy tinh Teflon có nút xoáy. Sau khi hỗn hợp phản ứng được làm mát, 2mL nước Milli-Q được bổ sung vào và sterol được chiết vào trong 2mL hexan: diclometan (4:1 thể tích/thể tích) bằng cách lắc và lắc xoáy. Hỗn hợp này được ly tâm và phần chiết sterol được loại riêng và rửa bằng 2mL nước Milli-Q. Phần chiết sterol sau đó được loại riêng bằng cách lắc và ly tâm. Phần chiết này được bốc hơi bằng dòng khí nitơ và sterol silyl hóa bằng 200mL BSTFA và gia nhiệt trong 2 giờ ở 80°C.

Để phân tích GC/GC-MS các sterol, dẫn xuất sterol-OTMSi được sấy khô trong điều kiện dòng khí nitơ trên khối nhiệt ở 40°C và sau đó được hòa tan lại trong cloform hoặc hexan ngay trước khi phân tích GC/GC-MS. Các dẫn xuất sterol-OTMS được phân tích bằng sắc ký khí (GC) sử dụng Agilent Technologies 6890A GC (Palo Alto, California, USA) được khớp với Supelco Equity™-1 cột mao dẫn được nhồi silic oxit (đường kính bên trong 15m x 0,1mm, chiều dày màng 0,1µm), FID, bơm tách/không tách và máy lấy mẫu và bơm tự động Agilent Technologies 7683B Series. Heli là khí mang. Các mẫu được bơm theo kiểu không tách ở nhiệt độ lò là 120°C. Sau khi bơm, nhiệt độ lò được nâng lên 270°C ở tốc độ 10°C phút⁻¹ và cuối cùng là đến 300°C ở tốc độ 5°C phút⁻¹. Các đỉnh được định lượng bằng phần mềm Agilent Technologies ChemStation (Palo Alto, California, USA). Kết quả GC có sai số ±5% diện tích từng thành phần riêng biệt.

Phân tích GC-quang phổ khối lượng (GC-MS) được tiến hành trên GC-MS Finnigan Thermoquest GCQ và GC-MS Finnigan Thermo Electron Corporation; cả hai

hệ thống đều được khớp với bơm trên cột và phần mềm Thermoquest Xcalibur (Austin, Texas, USA). Mỗi GC được khớp với cột mao dẫn có độ phân cực tương đương với được mô tả trên. Các thành phần riêng biệt được xác định bằng dữ liệu phổ khối lượng và bằng cách so sánh số liệu thời gian lưu với số liệu thu được của các chất chuẩn đáng tin cậy và chất chuẩn phòng thí nghiệm. Tiến hành phân tích mẫu trắng trong quy trình đầy đủ đồng thời với mẻ mẫu.

Điều kiện RT-PCR

Đặc trưng, sự khuếch đại PCR-sao chép ngược (RT-PCR-Reverse transcription-PCR) được tiến hành bằng hệ RT-PCR Superscript III One-Step (Invitrogen) trong thể tích 25 μ L sử dụng 10 pmol đoạn môi xuôi và 30 pmol đoạn môi ngược, MgSO₄ đến nồng độ cuối là 2,5mM, 400ng toàn bộ ARN với đệm và các thành phần nucleotit theo hướng dẫn của nhà sản xuất. Quy trình nhiệt đặc trưng là: 1 chu kỳ 45°C trong 30 phút để xảy ra sao chép ngược; sau đó 1 chu kỳ 94°C trong 2 phút sau đó là 40 chu kỳ 94°C trong 30 giây, 52°C trong 30 giây, 70°C trong 1 phút; sau đó 1 chu kỳ 72°C trong 2 phút trước khi làm lạnh hỗn hợp phản ứng đến 5°C.

Tạo ra mầm sinh dưỡng *B. napus* bằng cách gây nhiễm với 35S-LEC2

Các hạt *B. napus* (cv. Oscar) được tiệt khuẩn bằng khí clo như mô tả trong tài liệu: Attila Kereszt et al., 2007. Hạt đã tiệt khuẩn được cho nảy mầm trên môi trường MS mạnh 1/2 (Murashige và Skoog, 1962) với 0,8% thạch điều chỉnh đến độ pH=5,8, và phát triển ở 24°C trong điều kiện ánh sáng huỳnh quang (50 μ E/m²s) với chu kỳ ánh sáng 18/6 giờ (sáng/tối) trong 6-7 ngày. Cuống lá của lá mầm có chiều dài thân 2-4mm được tách vô khuẩn từ các cây giống con này và sử dụng làm mảnh cấy. Môi trường nuôi cấy các chủng *A. tumefaciens* AGL1 đã biến nạp, là một chủng mang vector kép đặc hiệu với hạt và chủng thứ hai với cấu trúc 35S-LEC2 được ủ từ các khuẩn lạc đơn từ các đĩa sạch và phát triển trong 10 mL môi trường LB với các kháng sinh thích hợp và để phát triển qua đêm ở 28°C và trộn với tốc độ 150 vòng/phút. Tế bào vi khuẩn được tập trung bằng cách ly tâm ở tốc độ 4000 vòng/phút trong 5 phút, rửa bằng môi trường MS chứa 2% sucroza và tạo huyền phù lại trong 10mL môi trường này và phát triển cùng với kháng sinh để lựa chọn thích hợp trong 4 giờ sau khi bổ sung axetosyringon đến 100 μ M. Hai giờ trước khi bổ sung vào các mô cây, spermidin được bổ sung đến nồng độ cuối là 1,5mM và mật độ vi khuẩn cuối cùng được điều chỉnh đến giá trị OD 600nm = 0,4 bằng môi trường sạch. Hai chủng vi khuẩn nuôi cấy, một

chúng mang cấu trúc đặc hiệu với hạt và chúng khác mang 35S-AtLEC2, được trộn với tỷ lệ từ 1:1 đến 1:1,5.

Cuống lá của lá mầm *B. napus* vừa được tách được gây nhiễm bằng 20mL *A. tumefaciens* nuôi cấy trong 6 phút. Cuống lá của lá mầm được thấm trên giấy lọc vô khuẩn để loại *A. tumefaciens* dư và sau đó được chuyển vào môi trường nuôi cấy đồng thời (môi trường MS với TDZ 1mg/L, NAA 0,1mg/L, axetosyringon 100 μ M được bổ sung L-xystein (50mg/L), axit ascorbic (15mg/l) và MES (250mg/l)). Các đĩa được che kín bằng băng dính có lỗ siêu nhỏ và ủ trong bóng tối ở 24°C trong 48 giờ. Mảnh cây được nuôi cấy đồng thời được chuyển vào môi trường lựa chọn sơ bộ (MS chứa TDZ 1mg/L, NAA 0,1mg/L, AgNO₃ 3mg/L, cefotaxim 250mg/l và timentin 50mg/L) và nuôi cấy trong 4-5 ngày ở 24°C với chu kỳ ánh sáng 16 giờ/8 giờ. Mảnh cây sau đó được chuyển vào môi trường lựa chọn (MS chứa TDZ 1mg/L, NAA 0,1mg/L, AgNO₃ 3mg/L, cefotaxim 250mg/l và timentin 50mg/L) theo gen đánh dấu lựa chọn được trên vector đặc hiệu với hạt và nuôi cấy trong 2-3 tuần ở 24°C với chu kỳ ánh sáng 16 giờ/8 giờ. Mảnh cây có mô thể chai xanh có khả năng sinh mầm được chuyển vào môi trường MS không có hormon (MS với AgNO₃ 3mg/L, cefotaxim 250mg/l, timentin 50mg/L và tác nhân lựa chọn) và nuôi cấy thêm trong 2-3 tuần. Mầm giai đoạn lá mầm được phân tách từ mảnh cây còn sống trên môi trường lựa chọn được phân tích thành phần axit béo trong toàn bộ lipit của chúng bằng GC.

Ví dụ 2. Biểu hiện ổn định của con đường tổng hợp DHA biến đổi gen ở hạt *Arabidopsis thaliana*

Xây dựng cấu trúc vector kép

Vector kép pJP3416-GA7 và pJP3404, mỗi vector chứa bảy gen sinh tổng hợp axit béo khác loại, mã hóa 5 desaturaza và 2 elongaza, và gen đánh dấu chọn lọc được của cây nằm giữa các gen lặp biên phải và biên trái của T-ADN có trong mỗi vector (Fig.2 và 3). SEQ ID NO: 1 đưa ra trình tự nucleotit của vùng T-ADN của pJP3416- GA7 từ trình tự biên phải sang trái. Cả hai cấu trúc gen này chứa các gen tối ưu hóa theo codon thực vật mã hóa Δ 12-desaturaza *Lachancea kluyveri* (bao gồm các nucleotit 14143-16648 của SEQ ID NO: 1), ω 3-desaturaza *Pichia pastoris* (bao gồm các nucleotit 7654-10156 của SEQ ID NO: 1), Δ 6-desaturaza *Micromonas pusilla* (bao gồm các nucleotit 226-2309 của SEQ ID NO: 1), Δ 5- và Δ 4-desaturaza *Pavlova salina* (lần lượt bao gồm

các nucleotit 4524-6485 và 10157-14142 của SEQ ID NO: 1) và $\Delta 6$ - và $\Delta 5$ -elongaza *Pyramimonas cordata* (lần lượt bao gồm các nucleotit 2310-4523 và 17825-19967 của SEQ ID NO: 1). Các vùng đặc hiệu của T-ADN (Hướng: trình tự biên từ phải sang trái) của vectơ kép pJP3416-GA7 đối với SEQ ID NO: 1 là như sau:

Nucleotit 1-163: Biên phải; 480-226, gen kết thúc nopalin synthaza *Agrobacterium tumefaciens* (TER_NOS); 1883-489, $\Delta 6$ -desaturaza *Micromonas pusilla*; 2309-1952, gen khởi đầu cắt napin *Brassica napus* (PRO_FP1); 2310-3243, gen khởi đầu FAE1 *Arabidopsis thaliana* (PRO_FAE1); 3312-4181, $\Delta 6$ -elongaza *Pyramimonas cordata*; 4190-4523, gen kết thúc lectin *Glycine max* (TER_Lectin); 4524-4881, PRO_FP1; 4950-6230: $\Delta 5$ -desaturaza *Pavlova salina*; 6231-6485: TER_NOS; 7653-6486, *Nicotiana tabacum* Rb7 vùng gắn cơ chất (MAR); 8387-7654, gen kết thúc conlinin1 *Linum usitatissimum* (TER_Cnl1); 9638-8388, $\omega 3$ -desaturaza *Pichia pastoris*; 10156- 9707, gen khởi đầu conlinin1 *Linum usitatissimum* (PRO_Cnl1); 10157-12189, gen khởi đầu conlinin1 *Linum usitatissimum*; 12258-13604, $\Delta 4$ -desaturaza *Pavlova salina*; 13605-14142, gen kết thúc conlinin2 *Linum usitatissimum*; 14143-14592, PRO_Cnl1; 14661-15914, $\Delta 12$ -desaturaza *Lachancea kluyveri*; 15915-16648, TER_Cnl1; 17816-16649, MAR; 17825-18758, PRO_FAE1; 18827-19633, $\Delta 5$ -elongaza *Pyramimonas cordata*; 19634-19967, TER_Lectin; 19990-20527, gen khởi đầu 35S của virus gây bệnh khảm trên súp lơ với vùng tăng cường nhân đôi; 20537-21088, phosphinothrixin- N-axetyltransferaza *Streptomyces viridochromogenes*; 21097-21349, TER_NOS; 21367-21527, biên trái.

Bảy vùng mã hóa trong các cấu trúc này mỗi vùng đều ở dưới sự kiểm soát của gen khởi đầu đặc hiệu với hạt- ba gen khởi đầu khác nhau được sử dụng, cụ thể là gen khởi đầu napin *Brassica napus* bị cắt cụt (pBnFP1), gen khởi đầu FAE1 *Arabidopsis thaliana* (pAtFAE1) và gen khởi đầu conlinin 1 *Linum usitatissimum* (pLuCnl1). Bảy gen sinh tổng hợp axit béo cùng được mã hóa cho con đường tổng hợp DHA nguyên vẹn được thiết kế để chuyển hóa 18:1 ^{$\Delta 9$} (axit oleic) thành 22:6 ^{$\Delta 4,7,10,13,16,19$} (DHA). Cả hai vectơ kép chứa vùng mã hóa gen đánh dấu chọn lọc được của cây BAR được liên kết điều khiển với gen khởi đầu 35S CaMV (Cauliflower Mosaic Virus) với vùng tăng cường nhân đôi và vùng polyadenyl hóa *A. tumefaciens nos3'* - gen kết thúc phiên mã. Gen đánh dấu chọn lọc được của cây được đặt sát cạnh biên trái của vùng T-ADN, do đó được xác định là cách xa trên T-ADN với hướng di chuyển của T-ADN vào trong tế

bào thực vật. Nó làm tăng khả năng vận chuyển từng phần T-ADN, phần này có khả năng không chứa gen đánh dấu lựa chọn được, sẽ không được chọn. pJP3416-GA7 và pJP3404, mỗi gen chứa gen gốc RiA4 để sao chép từ *Agrobacterium rhizogenes* (Hamilton, 1997).

pJP3416-GA7 được tạo ra bằng cách tổng hợp vùng ADN tương ứng với nucleotit 226-19975 của SEQ ID NO: 1 (vùng GA7) và lồng vùng này vào trong vector kép thể nhận pJP3416 ở vị trí *PspOMI*. Mỗi gen sinh tổng hợp axit béo trên GA7 bao gồm trình tự vùng không được dịch mã của virus khảm thuốc lá (Tobacco Mosaic Virus 5') (5'UTR), trình tự này được liên kết điều khiển với mỗi vùng mã hóa, nằm giữa gen khởi đầu và ATG khởi động dịch mã, để tối đa hóa hiệu suất dịch mã của mARN được tạo ra từ các gen này. Cấu trúc GA7 cũng bao gồm hai trình tự của vùng gắn cơ chất Rb7 *Nicotiana tabacum* (MAR), như mô tả bởi Hall et al. (1991). Các trình tự MAR, đôi khi kết thúc bởi các vùng gắn vào nhân, được biết là gắn kết đặc hiệu với chất nền của nhân *in vitro* và có thể là trung gian gắn kết của chất nhuộm sắc với chất nền của nhân *in vivo*. MAR cũng được cho là có chức năng làm giảm sự can thiệp do biến đổi gen. Trong pJP3416-GA7, các MAR cũng được lồng vào và có vị trí trong vùng T-ADN để hoạt động như là gen đệm ADN để ngăn cách các cát xét biểu hiện biến đổi gen. Vector pJP3416 trước khi lồng vùng GA7 chỉ chứa cát xét gen đánh dấu chọn lọc được của cây giữa các biên.

Cấu trúc gen pJP3404 được tạo ra bằng cách lồng liên tiếp các enzym giới hạn trong đó các cát xét gen được bổ sung vào vector kép, pJP3367, gen này bao gồm các gen để tạo ra SDA trong hạt. Cấu trúc này chứa các gen mã hóa $\Delta 12$ -desaturaza *L. kluyveri* và $\omega 3$ -desaturaza *P. pastoris*, cả hai gen này được biểu hiện bởi gen khởi đầu cắt napin *B. napus* (FP1), và $\Delta 6$ -desaturaza *M. pusilla* được biểu hiện bởi gen khởi đầu FAE1 *A. thaliana* (Fig.4). Đầu tiên, intron FAD2 *A. thaliana* được đặt hai bên là vị trí *EcoRI* và tách dòng vào vị trí *MfeI* của pJP3367 để sinh ra pJP3395. Một mảnh chứa cát xét $\Delta 6$ - và $\Delta 5$ -elongaza *P. cordata* lần lượt được điều khiển bởi gen khởi đầu FAE1 và FP1, được tách dòng vị trí *KasI* của pJP3395 để sinh ra pJP3398. Sau đó, pJP3399 được tạo ra bằng cách thay thế gen gốc RK2 để sao chép trong pJP3398 bằng gen gốc RiA4 để sao chép. Vector kép cuối cùng, pJP3404, được tạo ra bằng cách tách dòng mảnh hai bên sườn là *SbfI* chứa cát xét $\Delta 5$ - và $\Delta 4$ -desaturaza *P. salina* lần lượt được điều khiển bởi gen khởi đầu FP1 và FAE1, vào vị trí *SbfI* của pJP3399.

Biến nạp *A. thaliana* và phân tích thành phần axit béo

Vectơ khảm được đưa vào trong chủng *A. tumefaciens* AGL1 và các tế bào từ môi trường nuôi cấy của *Agrobacterium* biến nạp được sử dụng để xử lý cây *A. thaliana* (kiểu sinh thái Columbia và đột biến *fad2*) sử dụng phương pháp nhúng chìm hoa để biến nạp (Clough và Bent, 1998). Sau khi trưởng thành, các hạt T₁ từ cây đã được xử lý được thu hoạch và cho lên các đĩa MS chứa PPT để lựa chọn cây chứa gen đánh dấu chọn lọc được BAR. Các cây giống con T₁ sống, khỏe mạnh được chuyển vào đất. Sau khi phát triển cây này đến trưởng thành và để chúng tự thụ phấn, các hạt T₂ từ cây này được thu hoạch và thành phần axit béo của lipid hạt này được phân tích bởi phân tích GC như mô tả trong Ví dụ 1.

Số liệu về mức DHA trong lipid hạt được trình bày trong Fig.5 (các dòng được đánh dấu T₂) đối với 13 thể biến nạp sử dụng pJP3416-GA7 vào trong chất chuẩn di truyền Columbia, và đối với sáu thể biến nạp sử dụng biến thể *fad2*. Cấu trúc pJP3416-GA7 khiến tạo ra các mức DHA hơi cao hơn, là phần trăm của tổng hàm lượng axit béo, trên mức trung bình so với cấu trúc pJP3404. Bảng 4 thể hiện thành phần axit béo của toàn bộ lipid hạt từ dòng T₂ có mức DHA cao nhất. Hiệu suất chuyển hóa tính toán được với mỗi bước enzym trong quá trình tạo ra DHA từ axit oleic trong cùng các hạt được trình bày trong Bảng 5. Hiệu suất chuyển hóa được tính toán là (% sản phẩm x 100)/(% cơ chất còn lại + % sản phẩm), được thể hiện bằng tỷ lệ phần trăm.

Mức DHA cao nhất quan sát thấy được tạo ra trong các dòng T₂ biến nạp pJP3416-GA7 là 6,2%, ngoài ra với 0,5% EPA và 0,2% DPA (dòng #14). Các hạt T₂ vẫn ở trạng thái phân ly với gen biến đổi tức là vẫn chưa đồng hợp tử đồng đều. Số liệu được thu thập từ profin toàn bộ lipid hạt từ hạt biến đổi gen độc lập (Bảng 4) được trình bày trong Bảng 6. Mức axit béo ω 3 tạo ra là kết quả của gen biến đổi trong các hạt này (toàn bộ axit béo ω 3 mới, trừ mức ALA được tạo ra nội sinh ở dòng chuẩn Columbia) là 10,7% trong khi mức axit béo ω 6 (toàn bộ axit béo ω 6 mới trừ 18:2^{Δ9,12}) là 1,5%. Điều này thể hiện tỷ lệ đặc biệt ưu tiên của axit béo ω 3 mới:axit béo ω 6 mới, cụ thể là 7,3:1.

Các hạt T₂ của các dòng được chọn biến nạp với pJP3416-GA7, cụ thể là với các dòng được ký hiệu là 7, 10, 14, 22 và 34 ở dòng chuẩn Columbia và đối với các dòng được ký hiệu 18, 21 và 25 trong thể đột biến chuẩn *fad2*, được cấy đĩa lên môi trường MS chứa PPT để lựa chọn cây giống con biến đổi gen *in vitro*. Hai mươi cây

giống con kháng PPT của mỗi dòng được chuyển vào đất chuyển vào đất và phát triển đến khi trưởng thành sau khi tự thụ phấn. Các cây này có khả năng cao là đồng hợp tử với gen đánh dấu lựa chọn được, và do đó có ít nhất một đoạn lồng T-ADN trong hệ gen của cây. Hạt T₃ từ các cây này được thu hoạch và phân tích thành phần axit béo trong dầu hạt của chúng bằng GC. Số liệu được trình bày trong Bảng 7. Phân tích này bộc lộ rằng cấu trúc pJP3416-GA7 tạo ra các mức DHA ω 3 LC-PUFA trong hạt T₃ của cây đồng hợp tử cao hơn so với trong các hạt T₂ phân ly. Quan sát thấy lên đến khoảng 13,9% DHA trong dòng biến nạp pJP3416-GA7 T₃ ký hiệu là 22.2 trong dòng chuẩn Columbia tăng từ khoảng 5,5% trong hạt bán hợp tử T₂, với mức tổng cộng là khoảng 24,3% là axit béo ω 3 mới là phần trăm của toàn bộ axit béo trong hàm lượng lipid hạt. Axit béo ω 6 mới có mức là 1,1% của toàn bộ axit béo, thể hiện một tỷ lệ rất ưu tiên của axit béo ω 3 mới:axit béo ω 6 mới, cụ thể là khoảng 22:1. Tương tự, thể biến nạp trong thể đột biến chuẩn *fad2* thu được tổng cộng axit béo ω 3 mới với mức 20,6%, bao gồm 11,5% DHA, là phần trăm của toàn bộ axit béo trong hàm lượng lipid hạt.

Bảng 4. Thành phần axit béo của toàn bộ lipid hạt từ hạt *Arabidopsis* T₂ biến đổi gen độc lập có mức DHA ở giới hạn trên của khoảng quan sát được. ‘Col’ để chỉ kiểu sinh thái Columbia và ‘FAD2’ để chỉ kiểu sinh thái đột biến *fad2*. ‘GA7’ để chỉ sự biến nạp bởi T-ADN của vector pJP3416-GA7, pJP3404 với T-ADN của vectơ pJP3404. Axit béo 20:1n-9 và 20:1n-11 không được phân tích trong phân tích GC. Các axit béo “thứ yếu khác” bao gồm 14:0, 16:1n7, 16:1n9, 16:1n13t, 16:2n6, 16:3n3, i18:0, 18:1n5, 20:1n5, 22:0, 22:1n7, 22:1n11/n13, 24:0, 24:1n9.

	pJP3404_Col_#1	pJP3404_FAD2_#31	GA7_Col_#7	GA7_Col_#34	GA7_Col_#2	GA7_Col_#10	GA7_Col_#22	GA7_Col_#14	GA7_FAD2_#25	GA7_FAD2_#21	GA7_FAD2_#18
16:0	9,6	7,8	8,7	8,2	8,7	8,6	8,3	9,7	7,2	8,5	7,5
18:0	2,9	3,9	3,7	3,9	3,6	3,3	3,4	3,6	3,2	3,9	3,0
18:1d11	2,2	1,8	2,0	1,9	2,0	2,3	2,3	2,7	1,9	2,0	1,8
20:0	1,6	2,3	2,0	2,0	2,1	1,6	1,6	1,8	1,6	2,2	1,5
20:1d13	2,2	1,8	1,6	1,5	1,7	1,6	1,5	1,7	1,5	1,7	1,4
20:1d9/d11	13,0	15,9	16,1	16,1	16,3	15,0	13,9	13,5	18,3	15,9	17,0
22:1d13	1,1	1,2	1,1	1,1	1,3	1,0	1,0	1,0	1,0	1,3	1,2

Axit thứ khác	béo yếu											
		1,9	1,5	1,5	1,4	1,5	1,3	1,6	1,7	1,6	1,4	1,6
18:1d9		10,8	14,0	10,6	10,6	10,1	11,1	10,0	7,7	26,0	8,2	20,9
18:2ω6		28,9	28,3	16,4	16,1	18,2	13,7	13,7	11,4	6,6	16,6	4,3
18:3ω3		16,6	14,9	29,6	29,6	27,5	32,4	30,4	32,8	21,9	27,7	30,1
18:3ω6		0,7	0,5	0,1	0,1	0,1	0,1	0,2	0,1	0,1	0,2	0,1
20:2ω6		1,6	1,5	1,1	1,2	1,3	1,0	1,0	1,0	0,4	1,4	0,4
20:3ω6		0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
20:4ω6		0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
22:4ω6		1,6	0,6	0,3	0,3	0,3	0,4	0,5	0,4	0,5	0,4	0,4
22:5ω6		0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
18:4ω3		1,0	0,5	1,2	1,1	1,1	1,5	2,7	2,7	1,9	1,8	1,7
20:3ω3		0,0	0,0	0,0	0,6	0,0	0,0	0,6	0,7	0,0	0,8	0,6
20:4ω3		0,4	0,6	0,6	0,7	0,5	0,8	0,8	0,4	1,0	0,8	0,8
20:5ω3		0,2	0,2	0,3	0,3	0,3	0,3	0,7	0,5	0,6	0,4	0,5
22:5ω3		0,0	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,3	0,3	0,3
22:6ω3		3,6	2,4	3,0	3,1	3,3	3,9	5,5	6,2	4,3	4,4	4,8

112

[illegible]

Bảng 6. Số liệu thu thập từ profin toàn bộ lipid hạt từ hạt biến đổi gen độc lập được nêu trong Bảng 2. Kết quả tính toán không bao gồm ‘các axit béo thứ yếu’ trong Bảng 4.

Thông số	pJP3404_Col_#1	pJP3404_FAD2_#31	GA7_Col_#7	GA7_Col_#34	GA7_Col_#2	GA7_Col_#10	GA7_Col_#22	GA7_Col_#14	GA7_FAD2_#25	GA7_FAD2_#21	GA7_FAD2_#18
Tổng ω3 (% trong FA tổng cộng)	21,8	18,8	34,9	35,6	32,9	39,1	40,9	43,5	30,0	36,2	38,8
Tổng ω6 (% trong FA tổng cộng)	32,9	31,0	17,9	17,7	19,9	15,2	15,4	12,9	7,6	18,6	5,2
Tỷ lệ ω3/ω6	0,66	0,61	1,95	2,01	1,65	2,57	2,66	3,37	3,95	1,95	7,46
Tỷ lệ ω6/ω3	1,51	1,65	0,51	0,50	0,60	0,39	0,38	0,30	0,25	0,51	0,13
Tổng ω3 mới (% trong FA tổng cộng)	5,2	3,9	5,3	6,0	5,4	6,7	10,5	10,7	8,1	8,5	8,7
Tổng ω6 mới (% trong FA tổng cộng)	4,0	2,7	1,5	1,6	1,7	1,5	1,7	1,5	1,0	2,0	0,9
Tỷ lệ ω3/ω6 mới	1,30	1,44	3,53	3,75	3,18	4,47	6,18	7,13	8,10	4,25	9,67
Tỷ lệ ω6/ω3 mới	0,77	0,69	0,28	0,27	0,31	0,22	0,16	0,14	0,12	0,24	0,10
Hiệu suất OA thành EPA	4,8%	3,5%	4,3%	4,4%	4,7%	5,4%	7,9%	8,8%	6,3%	6,4%	6,7%
Hiệu suất OA thành DHA	4,5%	3,0%	3,7%	3,8%	4,1%	4,8%	6,8%	7,9%	5,2%	5,5%	5,8%
Hiệu suất LA thành EPA	6,9%	5,6%	6,6%	6,8%	7,2%	8,1%	11,4%	12,2%	13,8%	9,3%	12,7%
Hiệu suất LA thành DHA	6,6%	4,8%	5,7%	5,8%	6,3%	7,2%	9,8%	11,0%	11,4%	8,0%	10,9%
Hiệu suất ALA thành EPA	17,4%	14,9%	10,0%	10,1%	11,6%	11,3%	15,6%	15,9%	17,3%	14,1%	14,4%
Hiệu suất ALA thành DHA	16,5%	12,8%	8,6%	8,7%	10,0%	10,0%	13,4%	14,3%	14,3%	12,2%	12,4%
Tổng lượng axit bão hòa	14,1	14,0	14,4	14,1	14,4	13,5	13,3	15,1	12,0	14,6	12,0
Tổng lượng một liên kết chưa bão hòa	29,3	34,7	31,4	31,2	31,4	31,0	28,7	26,6	48,7	29,1	42,3
Tổng lượng axit béo nhiều liên kết chưa bão hòa	54,7	49,8	52,8	53,3	52,8	54,3	56,3	56,4	37,6	54,8	44,0
Tổng cộng C20	17,4	20	19,7	20,4	20,1	18,7	18,5	17,8	21,8	21	20,7
Tổng cộng C22	6,4	4,5	4,6	4,7	5,1	5,5	7,2	7,8	6,1	6,4	6,7
Tỷ lệ C20/C22	2,72	4,44	4,28	4,34	3,94	3,40	2,57	2,28	3,57	3,28	3,09

Bảng 7. Thành phần axit béo của toàn bộ lipid hạt từ hạt cây con *Arabidopsis* T₃ và T₄ biến đổi gen độc lập thu được từ dòng cây trong Bảng 3. Sai số nêu trong thể hệ T₄ biểu thị giá trị SD của n=10.

	GA7_Col_7,2	GA7_Col_34,2	GA7_Col_10,13	GA7_Col_22,2	GA7_Col_14,19	GA7_FAD2-	GA7_FAD2-	GA7_FAD2-	T ₄ Col_22,2 (trung bình ±SD)	T ₄ Col_22,2 dòng tốt nhất
16:0	9,8	9,0	9,5	11,2	10,4	8,1	10,7	7,7	10,6±0,9	12,2
18:0	4,0	3,8	4,2	3,4	3,5	3,5	3,8	3,3	3,5±0,4	3,6
18:1n7	2,0	1,9	2,2	2,9	2,5	1,7	2,2	1,6	2,3±0,2	2,6
20:0	2,2	1,9	1,7	1,4	2,3	1,8	2,0	1,9	1,9±0,3	2,0
20:1d13	1,4	1,3	1,2	1,6	2,5	1,2	1,4	1,3	1,6±0,2	1,9
20:1d9/11	13,6	14,7	12,4	9,5	13,0	15,7	12,4	18,4	11,7±1,7	9,5
22:1d13	1,2	1,2	0,8	0,6	1,6	1,0	1,1	1,5	0,9±0,1	0,8
Axit béo thứ yếu khác	1,8	1,5	1,5	2,1	2,6	1,7	1,9	1,6	1,9±0,1	2,3
18:1d9	5,5	6,7	6,8	4,6	6,9	11,3	4,2	11,5	4,6±1,0	3,3
18:2ω6	7,5	7,9	7,4	5,6	14,8	5,8	8,9	5,6	5,3±0,9	4,3
18:3ω3	33,7	33,7	36,1	31,5	26,1	28,3	28,9	30,8	31,0±1,1	29,5
18:3ω6	0,2	0,2	0,2	0,4	0,1	0,3	0,6	0,1	0,4±0,1	0,4
20:2ω6	1,0	1,0	0,7	0,7	1,4	0,6	1,2	0,6	0,9±0,1	0,9
18:3ω6	0,2	0,2	0,2	0,4	0,1	0,3	0,6	0,1	0,4±0,1	0,4
20:2ω6	1,0	1,0	0,7	0,7	1,4	0,6	1,2	0,6	0,9±0,1	0,9
20:3ω6	0	0	0	0	0	0	0	0		
20:4ω6	0	0	0	0	0	0	0	0		
22:4ω6	0	0	0	0	0	0	0	0	0,1±0,0	0,1
22:5ω6	0	0	0	0	0	0	0	0		
18:4ω3	3,1	2,6	3,0	5,3	3,3	3,7	5,2	2,6	4,8±0,9	5,5
20:3ω3	1,4	1,3	1,2	1,3	1,2	1,1	1,3	1,3	1,5±0,2	1,7
20:4ω3	0,7	0,6	0,6	0,9	0,2	1,7	0,9	0,9	0,8±0,2	0,8

20:5 ω 3	0,9	0,9	0,7	1,9	0,8	1,2	1,0	0,8	1,5 $\pm$ 0,3	1,8
22:5 ω 3	0,7	0,6	0,6	1,0	0,4	0,8	0,6	0,5	1,1 $\pm$ 0,2	1,5
22:6 ω 3	9,5	9,2	9,4	13,9	6,6	10,3	11,5	7,9	13,3 $\pm$ 1,6	15,1

Hiệu suất chuyển hóa enzym của mỗi bước enzym trong con đường tạo ra DHA từ axit oleic được trình bày trong Bảng 8 đối với hạt T₃ có các mức DHA cao hơn. Hiệu suất chuyển hóa Δ 12-desaturaza trong hạt của dòng 22.2 là 81,6% và hiệu suất ω 3-desaturaza là 89,1%, cả hai đều cao đáng kể và chỉ ra rằng các enzym nấm (nấm men) là có thể có hoạt động tốt trong các hạt đang phát triển. Hoạt tính của các enzym ngoại sinh khác trong con đường tổng hợp DHA là cao tương đương với cơ chất ω 3 với hoạt tính Δ 6-desaturaza có hiệu suất là 42,2%, Δ 6-elongaza là 76,8%, Δ 5-desaturaza là 95,0%, Δ 5-elongaza là 88,7% và Δ 4-desaturaza ở hiệu suất là 93,3%. Hoạt tính Δ 6-desaturaza trên cơ chất ω 6 LA là thấp hơn nhiều, so với Δ 6-desaturaza hoạt động chỉ ở mức hiệu suất 0,7% trên LA. GLA được thể hiện ở mức độ chỉ là 0,4% và là sản phẩm duy nhất ω 6 mới ngoài 20:2 ω 6 là sản phẩm được phát hiện trong hạt T₃ có hàm lượng DHA cao nhất. Số liệu thu thập từ profin toàn bộ lipid hạt từ hạt biến đổi gen độc lập (Bảng 7) được trình bày trong Bảng 9. Số liệu của dòng có mức DHA cao nhất bao gồm tỷ lệ ω 6 FA tổng cộng (bao gồm LA) với ω 3 FA tổng cộng (bao gồm ALA) là 0,10. Tỷ lệ ω 6 FA mới (trừ LA) với ω 3 FA mới (trừ ALA) trong lipid của dòng này là 0,05. Mức toàn bộ axit béo nhiều liên kết chưa bão hòa là cao hơn 50% trong các dòng này, và cao hơn 60% trong ít nhất 4 trong số các dòng này. Hiệu suất chuyển hóa toàn bộ được tính là: OA thành EPA = 21,8%, OA thành DHA = 18,0%, LA thành EPA = 26,9%, LA thành DHA = 22,2%, ALA thành EPA = 30,1%, ALA thành DHA = 24,9%.

Bảng 8. Hiệu suất chuyển hóa của các bước enzym riêng biệt để tạo ra DHA từ axit oleic, quan sát trong toàn bộ lipid hạt từ hạt *Arabidopsis T₃* biến đổi gen trong Bảng 7.

	GA7_Col_7.2	GA7_Col_34.2	GA7_Col_10.13	GA7_Col_22.2	GA7_Col_14.19	GA7_FAD2-25.10	GA7_FAD2-21.2	GA7_FAD2-18.14	T ₄ Col_22.2 (trung bình)	T ₄ Col_22.2 dòng tốt nhất
d12-des	75,4%	73,1%	75,7%	81,6%	73,4%	66,6%	78,5%	63,1%	67,6%	82,7%
d15-des	85,3%	84,4%	86,2%	89,1%	70,2%	87,5%	82,2%	87,6%	81,0%	90,9%
d6-des	0,3%	0,3%	0,3%	0,7%	0,3%	0,6%	1,0%	0,2%	1,3%	0,7%
(d9-elo)	1,7%	1,7%	1,2%	1,2%	2,6%	1,1%	2,0%	1,3%	1,6%	1,5%
d6-elo										
d5-des										
d5-elo										
d4-des										
Omega-6										
d6-des	30,7%	29,3%	28,2%	42,2%	30,2%	38,5%	40,0%	29,2%	41,0%	45,7%
(d9-elo)	2,7%	2,7%	2,3%	2,4%	3,0%	2,3%	2,7%	2,9%	2,8%	3,1%
d6-elo	79,0%	81,1%	79,0%	76,8%	70,9%	79,2%	73,2%	79,1%	77,5%	77,7%
d5-des	94,0%	94,6%	94,5%	95,0%	97,9%	87,8%	93,3%	91,1%	95,0%	95,8%
d5-elo	91,9%	91,7%	93,6%	88,7%	89,5%	89,9%	92,2%	91,6%	90,8%	90,2%
d4-des	93,2%	93,7%	94,4%	93,3%	93,7%	92,5%	95,0%	93,9%	92,2%	90,9%
Omega-3										

Bảng 9. Số liệu thu thập từ profin toàn bộ lipid hạt từ hạt biến đổi gen độc lập nêu trong Bảng 5. Kết quả tính toán không bao gồm 'các axit béo thứ yếu' trong Bảng 7.

Thông số	GA7-Col_7.2	GA7-Col_34.2	GA7-Col_10.13	GA7-Col_22.2	GA7-Col_14.19	GA7-FAD2-25.10	GA7-FAD2-21.2	GA7-FAD2-18.14	T ₄ Col_22.2 (trung bình±SD)	T ₄ Col_22.2 dòng tốt nhất
Tổng ω3 (% trong FA tổng cộng)	50,0	48,9	51,6	55,8	38,6	47,1	49,4	44,8	54,0	55,9
Tổng ω6 (% trong FA tổng cộng)	8,7	9,1	8,3	6,7	16,3	6,7	10,7	6,3	6,7	5,7
Tỷ lệ ω3/ω6	5,75	5,37	6,22	8,33	2,37	7,03	4,62	7,11	8,06	9,81
Tỷ lệ ω6/ω3	0,17	0,19	0,16	0,12	0,42	0,14	0,22	0,14	0,12	0,10
Tổng ω3 mới (% trong FA tổng cộng)	16,3	15,2	15,5	24,3	12,5	18,8	20,5	14,0	23,0	26,4
Tổng ω6 mới (% trong FA tổng cộng)	1,2	1,2	0,9	1,1	1,5	0,9	1,8	0,7	1,4	1,4
Tỷ lệ ω3/ω6 mới	13,58	12,67	17,22	22,09	8,33	20,89	11,39	20,00	16,43	18,86
Tỷ lệ ω6/ω3 mới	0,07	0,08	0,06	0,05	0,12	0,05	0,09	0,05	0,06	0,05
Hiệu suất OA thành EPA	14,1%	13,3%	13,4%	21,8%	10,2%	15,0%	16,8%	11,2%	20,4%	24,5%
Hiệu suất OA thành DHA	12,0%	11,4%	11,8%	18,0%	8,6%	12,6%	14,8%	9,6%	17,1%	20,1%
Hiệu suất LA thành EPA	18,9%	18,4%	17,9%	26,9%	14,2%	22,9%	21,8%	18,0%	26,2%	29,9%
Hiệu suất LA thành DHA	16,2%	15,9%	15,7%	22,2%	12,0%	19,1%	19,1%	15,5%	21,9%	24,5%
Hiệu suất ALA thành EPA	22,2%	21,9%	20,7%	30,1%	20,2%	26,1%	26,5%	20,5%	29,4%	32,9%
Hiệu suất ALA thành DHA	19,0%	18,8%	18,2%	24,9%	17,1%	21,9%	23,3%	17,6%	24,6%	27,0%
Tổng lượng axit béo hòa	16,0	14,7	15,4	16,0	16,2	13,4	16,5	12,9	16,0	17,8
Tổng lượng một liên kết chưa bão hòa	23,7	25,8	23,4	19,2	26,5	30,9	21,3	34,3	21,1	18,1
Tổng lượng axit béo nhiều liên kết chưa bão hòa	58,7	58,0	59,9	62,5	54,9	53,8	60,1	51,1	60,7	61,6
Toàn bộ C20	19	19,8	16,8	15,9	19,1	21,5	18,2	23,3	18	16,6
Toàn bộ C22	11,4	11	10,8	15,5	8,6	12,1	13,2	9,9	15,4	17,5
Tỷ lệ C20/C22	1,67	1,80	1,56	1,03	2,22	1,78	1,38	2,35	1,17	0,95

Hạt T₃ từ dòng pJP3416-GA7 22.2 trong dòng chuẩn Columbia, là thể hệ con từ dòng T₂ 22, được gieo trực tiếp vào đất và thành phần axit béo của hạt già từ cây T₃ thu được được phân tích bằng GC. Mức DHA trung bình của hạt này là $13,3\% \pm 1,6$ (n=10) là phần trăm của các toàn bộ axit béo trong lipid hạt. Như trình bày trong Bảng 6 (cột bên tay trái), dòng có mức DHA cao nhất chứa 15,1% DHA trong toàn bộ axit béo của lipid hạt. Hiệu suất chuyển hóa enzym được trình bày trong Bảng 8 cho mỗi bước trong quá trình tạo ra DHA từ axit oleic.

Tỷ lệ $\omega 6$ FA tổng cộng (bao gồm LA) với $\omega 3$ FA (bao gồm ALA) trong dòng có mức DHA cao nhất là 0,102. Tỷ lệ $\omega 6$ FA mới (ngoại trừ LA) với $\omega 3$ FA mới (ngoại trừ ALA) trong dòng có mức DHA cao nhất là 0,053. Mức toàn bộ axit béo bão hòa là khoảng 17,8% và mức axit béo một liên kết chưa bão hòa là khoảng 18,1%. Mức toàn bộ axit béo $\omega 6$ là khoảng 5,7% và mức axit béo $\omega 3$ là khoảng 55,9%. Hiệu suất chuyển hóa toàn phần được tính là: OA thành EPA = 24,5%, OA thành DHA = 20,1%, LA thành EPA = 29,9%, LA thành DHA = 24,5%, ALA thành EPA = 32,9%, ALA thành DHA = 27,0%. Toàn bộ axit béo omega-3 được tìm thấy tích lũy đến 55,9% toàn bộ axit béo trong đó axit béo omega-6 là 5,7% toàn bộ profin.

Phân tích lai thẩm tách ADN được tiến hành. Kết quả thể hiện rằng các dòng có độ tích lũy DHA cao là bản sao đơn hoặc bản sao kép của T-ADN từ cấu trúc pJP3416-GA7 ngoại trừ dòng biến đổi gen Columbia#22, dòng này có ba đoạn gen lồng T-ADN trong hệ gen của cây *Arabidopsis*. Hạt thể hệ T5 cũng được phân tích và thấy có đến 13,6% DHA trong toàn bộ lipid hạt. Cấu trúc GA7 được thấy là ổn định qua nhiều thế hệ về khả năng sản xuất DHA.

Xác định hàm lượng dầu ở các dòng *A. thaliana* DHA biến đổi gen DHA

Hàm lượng dầu của hạt *A. thaliana* biến đổi gen có mức DHA khác nhau được xác định bởi GC như mô tả trong Ví dụ 1. Số liệu được trình bày trong Fig.6, vẽ đồ thị hàm lượng dầu (% dầu theo trọng lượng của hạt) so với hàm lượng DHA (là phần trăm của toàn bộ axit béo). Quan sát thấy lên đến 26,5mg DHA trong mỗi gam hạt (Bảng 10). Hàm lượng dầu của hạt *Arabidopsis* biến đổi gen được thấy là tỷ lệ nghịch với hàm lượng DHA. Lượng DHA tính trên trọng lượng của hạt là cao hơn ở hạt biến nạp có mức DHA khoảng 9% so với hạt có khoảng 14% DHA. Chưa xác định được điều này có đúng với hạt không phải là *Arabidopsis* hay không.

Bảng 10. Tỷ lệ và lượng DHA trong hạt *Arabidopsis* GA7-biến nạp.

	Hàm lượng DHA (% TFA)	Hàm lượng dầu (% dầu trên 1g hạt)	Hàm lượng DHA trọng lượng (mg/g hạt)
GA7/col 22.2-1	14,2	14,89	20,2
GA7/col 22.2-2	14,3	15,02	20,5
GA7/col 22.2-3	14,0	15,92	21,2
GA7/col 10.15-1	8,7	30,23	25,06
GA7/col 10.15-2	8,6	31,25	25,77
GA7/col 10.15-3	8,8	31,70	26,49

Ví dụ 3. Biểu hiện ổn định con đường tổng hợp DHA biến đổi gen ở hạt *Camelina sativa*

Vectơ kép pJP3416-GA7 như mô tả trên được đưa vào chủng *A. tumefaciens* AGL1 và tế bào từ môi trường nuôi cấy chứa *Agrobacterium* biến nạp được sử dụng để xử lý cây đang ra hoa *C. sativa* bằng phương pháp nhúng chìm hoa để biến nạp (Lu và Kang, 2008). Sau khi cây phát triển và trưởng thành, hạt T₁ từ cây đã được xử lý cây đã được xử lý được thu hoạch, gieo lên trên đất và cây thu được được xử lý bằng cách phun thuốc diệt cỏ BASTA để lựa chọn cây biến đổi gen và biểu hiện gen đánh dấu chọn lọc được *bar* có trên T-ADN của pJP3416-GA7. Cây T₁ sống sót sẽ có khả năng chịu được thuốc diệt cỏ được nuôi đến khi trưởng thành sau khi để chúng tự thụ phấn, và hạt T₂ thu được được thu hoạch. Năm cây biến đổi gen được thu lấy, chỉ ba trong số chúng chứa T-ADN nguyên vẹn.

Lipit được chiết từ nhóm khoảng hai mươi hạt từ mỗi trong số ba cây chứa T-ADN nguyên vẹn. Hai trong số ba mẫu được lấy nhóm chứa mức DHA rất thấp, khó phát hiện được, nhưng nhóm thứ ba chứa khoảng 4,7% DHA (Bảng 12). Do đó, lipit được chiết từ 10 hạt T₂ riêng biệt từ cây này và chế phẩm axit béo được phân tích bởi GC. Dữ liệu thành phần axit béo của các hạt riêng biệt của dòng biến nạp này cũng được trình bày trong Bảng 11. Số liệu thu thập từ profin toàn bộ lipit hạt (Bảng 11) được trình bày trong Bảng 12.

Bảng 11. Thành phần axit béo của toàn bộ lipid hạt từ hạt T₂ *Camelina sativa* biến đổi gen được biến nạp bởi T-ADN từ pIP3416-GA7. Thành phần axit béo được trình bày theo mẻ hạt được phân nhóm (FD5.46) và với 10 hạt riêng lẻ được xếp loại (trái sang phải) từ mức DHA cao nhất đến thấp nhất.

Axit béo	FD5.46 được gộp lại	# 2	# 4	# 8	# 7	# 9	# 1	# 3	# 5	# 6	# 10
14:0	0	0,2	0,2	0,1	0,2	0,2	0,2	0,2	0,1	0,2	0,2
16:0	11,6	12,1	12,3	12,1	13,2	12,3	12,8	11,9	11,4	11,5	11,7
16:1	0,2	0,0	0,1	0,1	0,0	0,2	0,0	0,2	0,2	0,2	0,2
16:3	0,3	0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
18:0	3,7	3,3	3,2	3,2	3,0	3,1	3,2	3,3	3,1	3,2	3,2
18:1	10,8	8,0	8,0	8,6	8,5	9,4	11,0	10,2	8,3	9,4	8,6
18:1d11	1,7	1,3	1,4	1,4	1,7	1,4	1,5	1,3	1,3	1,3	1,3
18:2	24,7	18,2	19,5	19,2	18,5	20,1	23,8	32,2	30,3	29,8	31,6
18:3ω3	27,4	26,7	26,6	27,3	28,9	28,2	27,4	28,3	29,2	29,5	28,2
18:3ω6	0,2	1,4	0,3	0,3	0,4	0,2	0,5	0,0	0,5	0,4	0,6
20:0	1,6	1,4	1,3	1,4	1,2	1,4	1,4	1,8	2,1	1,9	2,0
18:4ω3	2,2	6,8	6,4	5,7	7,2	5,7	4,1	0,0	0,0	0,0	0,0
20:1d11	5,3	4,4	4,6	4,8	3,3	4,1	3,5	4,4	6,1	5,8	5,5
20:1iso	0,4	0,3	0,3	0,3	0,3	0,3	0,0	0,5	0,6	0,5	0,5
20:2ω6	0,8	0,8	0,9	0,8	0,6	0,8	0,7	1,3	1,5	1,4	1,4
20:3ω3	0,6	0,8	0,8	0,8	0,7	0,8	0,7	0,6	0,7	0,7	0,6
22:0	0,4	0,5	0,5	0,5	0,4	0,5	0,5	0,6	0,6	0,6	0,6
20:4ω3	0,2	0,3	0,3	0,3	0,4	0,4	0,5	0,0	0,0	0,0	0,0
22:1	1,1	1,1	1,2	1,1	0,5	0,9	0,8	1,6	2,2	1,9	2,0

20:5ω3	0,7	1,3	1,6	1,5	1,6	1,1	1,7	0,0	0,0	0,0	0,1
22:2ω6	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,3	0,2	0,2
22:4ω6+22:3ω	0,3	0,2	0,3	0,3	0,0	0,3	0,0	0,4	0,6	0,5	0,5
3	0,3	0,3	0,3	0,3	0,0	0,3	0,0	0,4	0,4	0,4	0,4
24:0	0,3	0,4	0,4	0,3	0,0	0,3	0,0	0,5	0,6	0,5	0,5
22:5ω3	0,3	1,1	1,2	1,1	1,1	0,9	0,8	0,0	0,0	0,0	0,0
22:6ω3	4,7	9,0	8,5	8,3	8,3	7,1	4,9	0,0	0,0	0,0	0,0

Bảng 12. Số liệu thu thập từ profin toàn bộ lipid hạt từ hạt biến đổi gen trình bày trong Bảng 11. Kết quả tính toán không bao gồm ‘các axit béo thứ yếu’ trong Bảng 11.

Thông số	FD5.46 được gộp lại	# 2	# 4	# 8	# 7	# 9	# 1	# 3	# 5	# 6	# 10
Tổng ω3 (% trong FA tổng cộng)	36,1	46	45,4	45	48,2	44,2	40,1	28,9	29,9	30,2	28,9
Tổng ω6 (% trong FA tổng cộng)	25,8	20,4	20,7	20,3	19,5	21,1	25	33,7	32,6	31,8	33,8
Tỷ lệ ω3/ω6	1,40	2,25	2,19	2,22	2,47	2,09	1,60	0,86	0,92	0,95	0,86
Tỷ lệ ω6/ω3	0,71	0,44	0,46	0,45	0,40	0,48	0,62	1,17	1,09	1,05	1,17
Tổng ω3 mới (% trong FA tổng cộng)	8,1	18,5	18	16,9	18,6	15,2	12	0	0	0	0,1
Tổng ω6 mới (% trong FA tổng cộng)	1,1	2,2	1,2	1,1	1	1	1,2	1,5	2,3	2	2,2
Tỷ lệ ω3/ω6 mới	7,36	8,41	15,00	15,36	18,60	15,20	10,00				0,05
Tỷ lệ ω6/ω3 mới	0,14	0,12	0,07	0,07	0,05	0,07	0,10				22,00
Hiệu suất OA thành EPA	8,2%	15,6%	15,5%	15,1%	15,1%	12,8%	10,5%	0,0%	0,0%	0,0%	0,1%
Hiệu suất OA thành DHA	6,7%	12,3%	11,6%	11,5%	11,4%	10,0%	7,0%	0,0%	0,0%	0,0%	0,0%
Hiệu suất LA thành EPA	9,2%	17,2%	17,1%	16,7%	16,2%	13,9%	11,4%	0,0%	0,0%	0,0%	0,2%
Hiệu suất LA thành DHA	7,6%	13,6%	12,9%	12,7%	12,3%	10,9%	7,5%	0,0%	0,0%	0,0%	0,0%
Hiệu suất ALA thành EPA	15,8%	24,8%	24,9%	24,2%	22,8%	20,6%	18,5%	0,0%	0,0%	0,0%	0,3%
Hiệu suất ALA thành DHA	13,0%	19,6%	18,7%	18,4%	17,2%	16,1%	12,2%	0,0%	0,0%	0,0%	0,0%

Tổng lượng axit bão hòa	17,6	17,8	17,8	17,6	18	17,8	18,1	18,2	17,7	17,8	18,1
Tổng lượng một liên kết chưa bão hòa	19,8	15,5	16	16,6	14,3	16,6	16,8	18,7	19,3	19,6	18,6
Tổng lượng axit béo nhiều liên kết chưa bão hòa	62,5	66,6	66,4	65,6	67,7	65,6	65,1	63	63,1	62,5	63,2
Toàn bộ C20	9,6	9,3	9,8	9,9	8,1	8,9	8,5	8,6	11	10,3	10,1
Toàn bộ C22	5,4	10,3	10	9,7	9,4	8,3	5,7	0,6	0,9	0,7	0,7
Tỷ lệ C20/C22	1,78	0,90	0,98	1,02	0,86	1,07	1,49	14,33	12,22	14,71	14,43

DHA có ở trong sáu trong số 10 hạt riêng lẻ. Bốn hạt khác không có DHA và được dự đoán là thể phân ly vô hiệu không có T-ADN, dựa trên tình trạng bán hợp tử của đoạn gen lồng T-ADN ở cây mẹ. Lipit chiết tách từ hạt riêng lẻ có mức DHA cao nhất là 9,0% DHA trong khi phần trăm tổng cộng của EPA, DPA và DHA là 11,4%. Phần trăm tổng cộng của axit béo $\omega 3$ mới được tạo ra trong hạt này là kết quả của quá trình biến nạp (SDA, ETrA, ETA, EPA, DPA, DHA) là 19,3% trong khi tổng số tương ứng của axit béo $\omega 6$ mới (GLA, EDA, DGLA, ARA và sản phẩm kéo dài của $\omega 6$ bất kỳ) là 2,2% - chỉ GLA và EDA được phát hiện là axit béo $\omega 6$ mới. Tỷ lệ của toàn bộ $\omega 6$ FA (bao gồm LA) với $\omega 3$ FA (bao gồm ALA) được thấy là 0,44. Tỷ lệ $\omega 6$ FA mới (trừ LA) với $\omega 3$ FA mới (trừ ALA) trong hạt này có mức DHA cao nhất là 0,12. Mức toàn bộ axit béo bão hòa là khoảng 17,8% và mức axit béo một liên kết chưa bão hòa là khoảng 15,5%. Mức toàn bộ axit béo $\omega 6$ là khoảng 20,4% và mức axit béo $\omega 3$ là khoảng 46%. Hiệu suất chuyển hóa toàn phần được tính là: OA thành EPA = 15,6%, OA thành DHA = 12,3%, LA thành EPA = 17,2%, LA thành DHA = 13,6%, ALA thành EPA = 24,8%, ALA thành DHA = 19,6%.

Thu được hạt đồng hợp tử dòng này ở thể hệ T4. Lên đến 10,3% DHA được tạo ra ở trường hợp FD5-46-18-110 với mức trung bình 7,3% DHA quan sát được trong toàn bộ thể hệ T4.

Hạt đồng hợp tử được trồng trong nhiều nhà kính để thu được tổng cộng là hơn 600 cây cá thể. Dầu được chiết từ hạt bằng một loạt các phương pháp bao gồm bao gồm chiết bằng hệ thống chiết Soxhlet, chiết axeton và hexan.

Do số lượng dòng biến nạp độc lập của *C. sativa* thu được như mô tả trên là thấp, người ta tiến hành thêm các thử nghiệm để biến nạp *C. sativa* bằng pJP3416- GA7. Các tác giả sáng chế dự đoán rằng sẽ đạt được mức DHA cao 10% so với phần trăm của toàn bộ axit béo trong dầu hạt ở các dòng biến nạp thêm này, và cây là đồng hợp tử của T-ADN với mức 20% DHA. Hai mươi trường hợp *C. sativa* GA7_modH được tạo ra và hạt được phân tích hàm lượng DHA. Ba trường hợp GA7_modB được tạo ra và phân tích hạt T1 từ trường hợp CMD17.1 bộc lộ rằng hàm lượng DHA từ hạt lấy nhóm là 9,8%. Giá trị DHA trong hạt đơn lẻ cao nhất được thấy là 13,5%.

Ví dụ 4. Biểu hiện ổn định con đường tổng hợp DHA biến đổi gen trong hạt *Brassica napus*

Biến nạp *B. napus* và phân tích thành phần axit béo bằng vector đơn

Vector kép pJP3416-GA7 được sử dụng để tạo ra cây *Brassica napus* biến nạp và hạt từ cây này. Vector pJP3416-GA7 như mô tả trên được đưa vào chủng *Agrobacterium tumefaciens* AGL1 qua quy trình điện thẩm chuẩn. Môi trường nuôi cấy tế bào *Agrobacterium* biến đổi gen được cho phát triển qua đêm ở 28°C trong môi trường LB và khuấy trộn ở tốc độ 150 vòng/phút. Tế bào vi khuẩn được tập trung bằng cách ly tâm ở tốc độ 4000 vòng/phút trong 5 phút, rửa bằng môi trường Winans AB (Winans, 1988) và tạo huyền phù lại trong 10 mL môi trường Winans AB (độ pH =5,2) và tiếp tục phát triển qua đêm trong sự có mặt của kanamycin (50mg/l), rifampixin (25mg/l) và axetosyringon 100µM. Hai giờ trước khi bơm tế bào *Brassica*, spermidin (120mg/l) được bổ sung vào và mật độ cuối cùng của vi khuẩn được điều chỉnh đến giá trị OD ở 600nm là 0,3-0,4 bằng môi trường AB sạch. Cuống lá vừa được phân tách của lá mầm từ cây giống con *Brassica napus* 8 ngày tuổi trên 1/2 MS (Murashige và Skoog, 1962) hoặc các đoạn trụ dưới lá mầm đã được làm quen trước khoảng 3-4 ngày trên môi trường MS có thidiazuron 1mg/l (TDZ) và axit α -naphthalenaxetic 0,1mg/l (NAA) được gây nhiễm bằng 10mL chủng nuôi cấy *Agrobacterium* trong 5 phút. Sau đó, mảnh cây đã gây nhiễm bởi *Agrobacterium* được chấm trên giấy lọc vô khuẩn để loại *Agrobacterium* dư và chuyển vào môi trường nuôi cấy đồng thời (môi trường MS có TDZ 1mg/L, NAA 0,1mg/L và axetosyringon 100µM) được bổ sung hoặc không có các chất chống oxy hóa khác (L-xystein 50mg/l và ascorbic 15mg/l). Tất cả các đĩa được đậy kín bằng màng phim parafilm và ủ trong bóng tối ở 23-24°C trong 48 giờ.

Sau đó, các mảnh cây đã được xử lý được rửa bằng nước cất vô khuẩn chứa cefotaxim 500mg/l và timentin 50mg/L trong 10 phút, tráng trong nước cất vô khuẩn trong 10 phút, chấm khô trên giấy lọc vô khuẩn, chuyển vào môi trường lựa chọn sơ bộ (MS chứa TDZ 1mg/L, NAA 0,1mg/L, adenin sulphat 20mg/l (ADS), AgNO₃ 1,5mg/l, cefotaxim 250mg/l và timentin 50mg/L) và nuôi cấy trong 5 ngày ở 24°C với chu kỳ ánh sáng 16 giờ/8 giờ. Sau đó chúng được chuyển vào môi trường lựa chọn (MS chứa TDZ 1mg/L, NAA 0,1mg/L, ADS 20mg/l, AgNO₃ 1,5mg/l, cefotaxim 250mg/l và timentin 50mg/L) với glufosinat amoni 1,5mg/l làm tác nhân lựa chọn tế bào biến nạp, và nuôi cấy trong 4 tuần ở 24°C với chu kỳ ánh sáng 16 giờ/8 giờ và nuôi cấy cấp hai hai tuần một lần trên cùng môi trường. Mảnh cây có vết chai màu xanh được chuyển

vào môi trường nảy chồi (MS chứa kinetin 1mg/l, ADS 20mg/l, AgNO₃ 1,5mg/l, cefotaxim 250mg/l, timentin 50mg/L và glufosinat amoni 1,5mg/l) và nuôi cấy trong 2-3 tuần nữa. Các chồi nhú lên từ mảnh cây còn lại được chuyển vào môi trường kéo dài chồi (môi trường MS có axit gibberelic 0,1mg/l, ADS 20mg/l, AgNO₃ 1,5mg/l, cefotaxim 250mg/l và glufosinat amoni 1,5mg/l) và nuôi cấy trong thêm hai tuần. Các chồi khỏe mạnh có chiều dài 2-3cm được lựa chọn và chuyển vào môi trường kích thích nảy rễ (1/2 MS chứa NAA 1mg/l, ADS 20mg/l, AgNO₃ 1,5mg/l và cefotaxim 250mg/l) và nuôi cấy trong 2-3 tuần. Các chồi đã ổn định với rễ được chuyển vào các bình chứa hỗn hợp dành cho cây trồng từ hạt và phát triển trong buồng phát triển trong hai tuần và sau đó chuyển vào nhà kính. Thu được khoảng 40 cây (T₀) được gây biến nạp bằng cấu trúc GA7 bằng phương pháp này.

Cây được phát triển đến thành thực sau khi tự thụ phấn. Hạt thu được từ cây biến nạp được phân tích về thành phần axit béo trong dầu hạt của chúng như mô tả trong Ví dụ 1. Số liệu về dòng gây biến nạp có mức DHA cao nhất được trình bày trong Bảng 13. Các mức DHA trung bình là thấp hơn đáng kể trong dầu hạt của hạt *B. napus* được gây biến nạp bằng T-ADN từ pJP3416-GA7 so với hạt *A. thaliana* (Ví dụ 2) hoặc hạt *Camelina* (Ví dụ 3) được gây biến nạp bằng cùng một cấu trúc. Thấy rằng mức DHA cao nhất trong khoảng 40 dòng là 1,52% với phần lớn các dòng biến đổi gen có mức DHA có thể phát hiện được. Đáng chú ý là sự tích lũy lượng lớn ALA, khoảng 35% trong số toàn bộ axit béo, trong các hạt không được chuyển hóa hiệu quả thành SDA hoặc các sản phẩm sau đó trong con đường này.

Phân tích profin axit béo của hạt *B. napus* riêng lẻ từ trường hợp T₁, CT125-2, được tiến hành để xác định thêm về lượng DHA tạo ra trong hạt biến đổi gen. Các hạt được thấy là chứa DHA với lượng nằm trong khoảng từ 0% (hạt vô hiệu) và 8,5% (Bảng 13).

Một số hạt từ dòng cây CT116 cũng như các dòng biến đổi gen khác thể hiện sự sản sinh DHA được gieo để tạo ra cây con. RT-PCR được thực hiện trên toàn bộ ARN được tách từ mầm đang phát triển từ cây để xác định lý do cấu trúc GA7 thể hiện sự sản sinh DHA tương ứng với *A. thaliana* và *C. sativa* biến đổi gen có cấu trúc tương tự, và tương đối thấp tổ hợp các gen trên pJP3115 và pJP3116 (sau đây). RT-PCR được thực hiện trên toàn bộ ARN bằng kit RT-PCR một bước (Invitrogen) và các đoạn môi đặc hiệu với gen chọn đích trên mỗi gen biến đổi. Điều này khẳng định rằng mỗi gen trong

cấu trúc GA7 được biểu hiện tốt trong thể biến nạp *B. napus* ngoại trừ $\Delta 6$ - desaturaza được biểu hiện kém trong phần lớn hạt gây biến nạp. Các gen khác từ cấu trúc này hoạt động tốt ở cả hạt *B. napus* và *A. thaliana*, ví dụ $\Delta 12$ - và $\Delta 15$ -desaturaza thực hiện chức năng tạo ra các mức LA và ALA cao hơn ở hạt trong khi làm giảm mức axit oleic. Gel RT-PCR đặc trưng được trình bày trong Fig.7 thể hiện rõ ràng sự biểu hiện thấp $\Delta 6$ -desaturaza so với các gen biến đổi khác từ pJP3416-GA7.

Cây và hạt biến đổi gen là đồng hợp tử của các gen biến đổi được tạo ra bằng cách trồng cây con từ các dòng có mức DHA cao nhất.

Bảng 13. Thành phần axit béo theo phần trăm của toàn bộ axit béo trong dầu hạt từ hạt *Brassica napus* T₁ độc lập được gây biến nạp bởi pJP3416-GA7, dòng CT116-11 và CT-125-2 so với đối chứng kiểu dại (không được gây biến nạp). 22:6 ω 3 là DHA. Số liệu từ hạt *B. napus* CT125-2 riêng lẻ được đánh dấu là ‘SS’.

	Đối chứng	CT116-11	CT125-2	CT125-2 #2 SS	CT125-2 #3 SS	CT125-2 #10 SS
14:0	0,1	0,2	0,1	0,1	0,1	0,1
16:0	4,3	7,2	5,2	6,5	4,7	7,7
16:1	0,2	0,5	0,4	0,3	0,3	0,8
16:3	0,2	0,2	0,2	0,1	0,2	0,2
18:0	2,1	2,2	2,4	2,3	2,3	2,8
18:1d9	59,1	27,0	38,1	34,0	19,3	14,8
18:1d11	3,7	6,6	4,2	4,4	4,3	9,6
18:2	19,7	14,1	16,6	13,9	10,2	10,2
18:3 ω 3	8,3	35,2	27,7	34,1	49,5	37,9
20:0	0,6	0,5	0,6	0,4	0,3	0,7
18:4 ω 3	0,0	0,9	0,3	0,5	0,6	2,6
20:1d11	1,2	1,1	1,0	1,0	0,8	0,6
20:1is0		0,2		0,1		0,2
20:2 ω 6	0,1	0,1	0,1	0,1	0,1	0,1
20:3 ω 3		1,3	0,7	0,8	1,6	0,9
22:0	0,3	0,4	0,3	0,1	0,1	0,4
20:4 ω 3		0,1	0,3	0,4	0,6	0,5
22:1						
20:5 ω 3					0,1	0,3
22:3 ω 3					0,1	
24:0	0,2	0,4	0,3	0,1	0,1	0,3
24:1	0,1	0,3	0,1	0,1	0,2	0,1
22:5 ω 3		0,1	0,1	0,1	0,1	0,5
22:6 ω 3		1,52	1,2	1,3	2,7	8,5

Biến nạp *B. napus* và phân tích thành phần axit béo bằng hai vectơ

Trong thí nghiệm khác ở *B. napus* và là định dạng khác để đưa các gen biến đổi vào, các vector kép pJP3115 và pJP3116 như mô tả trong WO 2010/057246 được sử dụng để tạo ra các cây *B. napus* biến nạp riêng biệt và hạt biến nạp thu được từ cây này. T-ADN trên pJP3115 chứa gen khảm mã hóa $\Delta 12$ -desaturaza *Crepis palestina*, $\Delta 6$ -desaturaza *Micromonas pusilla*, $\Delta 6$ -elongaza *Pyramimonas cordata* và $\Delta 5$ -desaturaza *Pavlova salina* và T-ADN trên pJP3116 chứa gen khảm mã hóa $\Delta 15$ -desaturaza *Perilla frutescens*, $\Delta 5$ -elongaza *Pyramimonas cordata* và $\Delta 4$ -desaturaza *Pavlova salina*. Hai T-ADN, khi cùng có mặt và được biểu hiện ở các hạt đang phát triển, tạo ra con đường 7-gen để sinh ra DHA từ axit oleic nội sinh. Các vector được đưa vào chủng *Agrobacterium tumefaciens* AGL1 thông qua quy trình điện thẩm chuẩn và tế bào biến nạp được sử dụng độc lập để gây biến nạp *B. napus* bằng phương pháp như mô tả ở trên để tạo ra cây biến nạp ổn định T₀. Thu được 29 thể biến nạp pJP3115 và 19 thể biến nạp pJP3116 và các cây được phát triển đến trưởng thành và hạt thu được sau khi tự thụ phấn được phân tích thành phần axit béo trong dầu hạt của chúng. Việc biến nạp bằng T-ADN từ pJP3115 được hy vọng là sinh ra EPA từ ALA nội sinh trong khi việc biến nạp bởi T-ADN từ pJP3116 được hy vọng là sinh ra ALA nhiều hơn từ LA. Nhiều cây được xác định là biểu hiện các phenotyp của chúng. Phần lớn các cây biểu hiện phenotyp OA giảm/LA tăng do sự khử bão hòa $\Delta 12$ khiến mức tạo ra EPA thấp. Quan sát thấy lên đến 2,6% EPA trong ở hạt nhóm pJP3115 biến đổi gen. Tương tự, phần lớn pJP3116 được thấy là có phenotyp ALA tăng cao do hoạt tính $\Delta 15$ -desaturaza. Thấy là mức ALA lên đến 18,5% trong nhóm hạt được gây biến nạp bằng T-ADN từ pJP3116.

Cây T₁ từ các dòng có mức EPA và ALA cao nhất được lai chéo và hạt cây con (F1) từ 24 cây phục hồi được phân tích hàm lượng DHA. DHA được tìm thấy trong 17 trong số cây này có mức lên đến 1,9% DHA trong nhóm hạt từ các cây này. Tiến hành phân tích từng hạt riêng lẻ để xác định khoảng sản sinh DHA – số liệu được trình bày trong Bảng 14. Một khoảng rộng của mức DHA được quan sát thấy ở cây con được lai chéo, có thể là do bản chất bán hợp tử của T-ADN của các cây thế hệ gốc, khiến cho các hạt này không nhận được cả hai T-ADN. Có đến 6,7% DHA trong toàn bộ lipit hạt.

Bảng 14. Thành phần axit béo ở dạng phần trăm của toàn bộ axit béo trong dầu hạt từ hạt riêng lẻ *B. napus* F1 được tạo ra khi lai chéo cây biến đổi gen T-ADN từ pJP3115 với cây biến đổi gen T-ADN từ pJP3116. B1, B2 và B4 ký hiệu các biến cố. 0,0 =

không phát hiện được bằng phương pháp GC.

	B1.1	B1.2	B1.3	B1.4-g	B1.5-g	B2.1	B2.2	B2.3g	B2.4g	B2.5g	B3.1	B3.2
14:0	0,1	0,1	0,1	0,2	0,2	0,1	0,1	0,2	0,2	0,1	0,1	0,1
16:0	6,6	6,4	4,5	12,3	7,9	5,1	5,0	10,1	8,5	6,8	5,3	7,2
16:1	0,4	0,5	0,2	1,0	0,6	0,4	0,4	0,6	1,1	0,5	0,5	0,6
16:3	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,2	0,1	0,1	0,2
18:0	2,3	2,6	2,2	1,6	2,9	2,9	3,4	2,2	1,8	2,9	3,4	2,4
18:1	34,1	39,3	46,9	14,9	20,7	41,6	46,3	14,4	23,4	38,3	43,6	32,0
18:1d11	4,6	5,8	2,7	6,8	6,2	3,8	4,9	5,9	8,7	4,5	5,5	5,1
18:2	33,6	30,7	30,4	29,2	34,4	31,7	27,7	33,2	23,9	33,3	27,9	33,4
18:3ω6	0,2	0,3	0,1	0,4	0,4	0,2	0,2	0,7	0,1	0,2	0,2	0,3
18:3ω3	10,3	7,1	7,7	18,7	14,9	8,2	5,9	14,8	28,1	6,3	7,3	10,0
20:0	0,6	0,7	0,6	0,5	0,7	0,8	0,9	0,6	0,4	0,7	0,9	0,7
18:4ω3	0,2	0,1	0,1	0,8	0,5	0,2	0,2	0,8	0,0	0,2	0,2	0,2
20:1d11	1,0	1,1	1,1	0,7	0,8	1,1	1,1	0,5	0,9	1,1	1,1	0,9
20:1iso	0,1	0,1	0,0	0,1	0,1	0,1	0,1	0,1	0,3	0,1	0,1	0,1
20:2ω6	0,4	0,3	0,2	0,5	0,5	0,4	0,3	0,4	0,5	0,5	0,3	0,5
20:3ω6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
20:4ω6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0
20:3ω3	1,8	1,6	1,1	2,8	2,1	1,1	1,0	2,7	0,7	1,4	0,9	1,6
22:0	0,3	0,4	0,3	0,3	0,4	0,4	0,5	0,3	0,3	0,4	0,5	0,4
20:4ω3	0,3	0,2	0,2	0,4	0,4	0,1	0,1	0,5	0,0	0,2	0,1	0,2
22:1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
20:5ω3	0,0	0,0	0,0	0,1	0,1	0,0	0,0	0,2	0,0	0,0	0,0	0,0
22:2ω6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
22:4ω6	0,1	0,2	0,1	0,2	0,2	0,1	0,1	0,4	0,2	0,2	0,1	0,2
24:0	0,3	0,4	0,2	0,2	0,3	0,3	0,3	0,3	0,4	0,4	0,4	0,3
22:5ω6	0,1	0,2	0,1	0,2	0,3	0,1	0,1	0,5	0,0	0,2	0,1	0,2
24:1	0,2	0,2	0,2	0,3	0,3	0,2	0,2	0,3	0,3	0,2	0,2	0,2
22:5ω3	0,7	0,7	0,3	2,1	1,6	0,3	0,4	3,2	0,0	0,5	0,4	1,2
22:6ω3	1,4	1,0	0,5	5,5	3,9	0,8	0,7	6,7	0,0	1,1	0,8	2,0

Bảng 15. Số liệu thu thập từ profin toàn bộ lipid hạt từ hạt biến đổi gen trình bày trong Bảng 14. Kết quả tính toán không bao gồm các 'axit béo thứ yếu' trong Bảng 14.

Thông số	B1.1	B1.2	B1.3	B1.4-g	B1.5-g	B2.1	B2.2	B2.3g	B2.4g	B2.5g	B3.1
Tổng $\omega 3$ (% trong FA tổng cộng)	4,6	3,9	2,3	12,1	9	2,7	2,6	14,8	0,8	3,6	2,6
Tổng $\omega 6$ (% trong FA tổng cộng)	44,5	38,5	38,5	48,8	50,3	40,5	34,1	49,4	52,7	40,5	35,7
Tỷ lệ $\omega 3/\omega 6$	0,10	0,10	0,06	0,25	0,18	0,07	0,08	0,30	0,02	0,09	0,07
Tỷ lệ $\omega 6/\omega 3$	9,67	9,87	16,74	4,03	5,59	15,00	13,12	3,34	65,88	11,25	13,73
Tổng $\omega 3$ mới (% trong FA tổng cộng)	2,6	2	1,1	8,9	6,5	1,4	1,4	11,4	0	2	1,5
Tổng $\omega 6$ mới (% trong FA tổng cộng)	10,5	7,5	7,9	19,1	15,4	8,4	6,1	15,8	28,3	6,7	7,5
Tỷ lệ $\omega 3/\omega 6$ mới	0,25	0,27	0,14	0,47	0,42	0,17	0,23	0,72	0,00	0,30	0,20
Tỷ lệ $\omega 6/\omega 3$ mới	4,04	3,75	7,18	2,15	2,37	6,00	4,36	1,39		3,35	5,00
Hiệu suất OA thành EPA	2,5%	2,1%	0,9%	10,1%	6,9%	1,3%	1,3%	12,8%		1,9%	1,4%
Hiệu suất OA thành DHA	1,7%	1,2%	0,6%	7,2%	4,8%	0,9%	0,8%	8,5%		1,3%	1,0%
Hiệu suất LA thành EPA	4,3%	4,0%	2,0%	12,6%	9,4%	2,5%	3,0%	15,7%		3,6%	3,1%
Hiệu suất LA thành DHA	2,9%	2,4%	1,2%	9,0%	6,6%	1,9%	1,9%	10,4%		2,5%	2,1%
Hiệu suất ALA thành EPA	47,7%	44,7%	36,4%	68,1%	65,9%	44,0%	45,8%	72,1%		47,1%	50,0%
Hiệu suất ALA thành DHA	31,8%	26,3%	22,7%	48,7%	45,9%	32,0%	29,2%	47,9%		32,4%	33,3%
Tổng lượng axit béo hòa	10,2	10,6	7,9	15,1	12,4	9,6	10,2	13,7	11,6	11,3	10,6
Tổng lượng một liên kết không bão hòa	40,4	47	51,1	23,8	28,7	47,2	53	21,8	34,7	44,7	51
Tổng lượng axit béo nhiều liên kết chưa bão hòa	49,2	42,5	40,9	61	59,4	43,3	36,8	64,3	53,7	44,2	38,4
Toàn bộ C20	4,2	4	3,2	5,1	4,7	3,6	3,5	5,1	2,8	4	3,4
Toàn bộ C22	2,6	2,5	1,3	8,3	6,4	1,7	1,8	11,1	0,5	2,4	1,9
Tỷ lệ C20/C22	1,62	1,60	2,46	0,61	0,73	2,12	1,94	0,46	5,60	1,67	1,79

Số liệu thu thập từ profin toàn bộ lipit (Bảng 14) được trình bày trong Bảng 15. Từ số liệu trong Bảng 15, tỷ lệ toàn bộ $\omega 6$ FA (bao gồm LA) với $\omega 3$ FA (bao gồm ALA) trong hạt có mức DHA cao nhất là 3,34. Tỷ lệ $\omega 6$ FA mới (trừ LA) với $\omega 3$ FA mới (trừ ALA) là 1,39. Mức toàn bộ axit béo bão hòa là khoảng 13,7% và mức axit béo một liên kết chưa bão hòa là khoảng 21,8%. Mức toàn bộ axit béo $\omega 6$ là khoảng 46,4% và mức axit béo $\omega 3$ là khoảng 14,8%. Hiệu suất chuyển hóa toàn phần được tính là: OA thành EPA = 12,8%, OA thành DHA = 8,5%, LA thành EPA = 15,7%, LA thành DHA = 10,4%, ALA thành EPA = 72,1%, ALA thành DHA = 47,9%. Quan sát thấy hiệu suất chuyển hóa axit béo $\omega 6$ thành axit béo $\omega 3$ giảm trong thử nghiệm này với tổ hợp pJP3115 và pJP3116 được cho là do hiệu suất thấp hơn của $\Delta 15$ -desaturaza thực vật so với $\Delta 15/\omega 3$ desaturaza nấm (Ví dụ 2 và 3) khi kết hợp với các gen chuyển hóa ALA thành DHA.

Cây con từ các dòng chứa DHA là đồng hợp tử của tất cả các gen biến đổi được đưa vào tạo ra cho phân tích này.

Ví dụ 5. Biến đổi T-ADN mã hóa con đường tổng hợp DHA ở hạt cây

Để cải thiện mức tạo ra DHA trong *B. napus* tăng hơn so với mức được mô tả trong Ví dụ 4, các vector kép pJP3416-GA7-modA, pJP3416-GA7-modB, pJP3416-GA7-modC, pJP3416-GA7-modD, pJP3416-GA7-modE và pJP3416-GA7-modF được xây dựng cấu trúc như sau. Các vector kép là biến thể của cấu trúc pJP3416-GA7 được mô tả trong Ví dụ 2 và được thiết kế để làm tăng hơn nữa khả năng tổng hợp DHA ở hạt cây, cụ thể là bằng cách cải thiện chức năng $\Delta 6$ -desaturaza và $\Delta 6$ -elongaza. SDA đã được quan sát thấy là tích lũy trong một số hạt được gây biến nạp bằng cấu trúc GA7 do hiệu suất kéo dài tương đối thấp so với $\Delta 5$ -elongaza, vì vậy trong số các biến đổi khác, hai vị trí gen elongaza được chuyển vào trong T-ADN.

Hai trình tự mã hóa elongaza trong pJP3416-GA7 được chuyển vào trong các vị trí của chúng trên T-ADN để thu được pJP3416-GA7-modA bằng cách đầu tiên là tách dòng cắt xét $\Delta 6$ -elongaza *P. cordata* mới nằm giữa các vị trí *SbfI* của pJP3416-GA7 để thay thế cắt xét $\Delta 5$ -elongaza *P. cordata*. Cấu trúc này được biến đổi tiếp bằng cách trao đổi gen khởi đầu FP1 điều khiển $\Delta 6$ -desaturaza *M. pusilla* với gen khởi đầu conlinin Cnl2 (pLuCnl2) để thu được pJP3416-GA7-modB. Sự biến đổi này được thực hiện nhằm cố gắng làm tăng mức biểu hiện $\Delta 6$ -desaturaza và do đó làm tăng hiệu suất

enzym. Cho rằng gen khởi đầu *Cnl2* có thể gây ra mức biểu hiện cao hơn ở *B. napus* so với gen khởi đầu cắt napin. pJP3416-GA7-modC được tạo ra bằng cách bổ sung cát xét $\Delta 6$ -desaturaza *M. pusilla* thứ hai có khả năng codon hơi khác một chút (SEQ ID NO: 15) và được điều khiển bởi gen khởi đầu FP1, nó được lồng ở vị trí *PmeI* ngay bên trong biên phải của pJP3416-GA7-modB. Cát xét $\Delta 6$ -desaturaza thứ hai được bổ sung vào cả pJP3416-GA7-modB và pJP3416-GA7-modF để làm tăng mức biểu hiện $\Delta 6$ -desaturaza và kéo dài thời gian của quá trình phát triển hạt để biểu hiện $\Delta 6$ -desaturaza bằng cách sử dụng nhiều gen khởi đầu. Các khả năng codon khác được sử dụng trong hai trình tự nucleotit để gây ra dịch mã cùng trình tự protein mà không gây ra nguy cơ ức chế đồng thời từ các vùng mã hóa trong cùng một T-ADN này. pJP3416-GA7-modD và pJP3416-GA7-modE là các biến thể tương tự nhau trong đó trình tự MAR thứ ba, tương ứng với các nucleotit 16649-17816 của SEQ ID NO: 1, lần lượt được bổ sung vào pJP3416-GA7 và pJP3416-GA7-modB, ở vị trí *PmeI*. pJP3416-GA7-modF được tạo ra bằng cách bổ sung cát xét $\Delta 6$ -desaturaza *M. pusilla* thứ hai chứa trình tự nucleotit $\Delta 6$ -desaturaza tự nhiên và được điều khiển bởi gen khởi đầu FP1 ở vị trí *PmeI* ở biên phải của pJP3416-GA7-modB. pJP3416-GA7-modG được tạo ra bằng cách đầu tiên thay thế cát xét $\Delta 6$ -desaturaza *M. pusilla* bằng cát xét $\Delta 5$ -elongaza *Cnl2:P. cordata* bằng cách tách dòng giới hạn ở các vị trí *AscI-PacI*. Sau đó, pJP3416-GA7-modG được tạo ra bằng cách thay thế cát xét gốc $\Delta 5$ -elongaza FAE1:*P. cordata* bằng cát xét $\Delta 6$ -desaturaza FAE1:*M. pusilla* bằng cách tách dòng giới hạn ở các vị trí *SbfI*. Trình tự nucleotit của T-ADN từ mỗi cấu trúc gen trong số này được trình bày là: pJP3416-GA7-modB (SEQ ID NO: 2), pJP3416-GA7-modC (SEQ ID NO: 3), pJP3416-GA7-modD (SEQ ID NO: 4), pJP3416-GA7-modE (SEQ ID NO: 5), pJP3416-GA7-modF (SEQ ID NO: 6) và pJP3416-GA7-modG (SEQ ID NO: 7).

Vectơ kép pJP3416-GA7-modB, pJP3416-GA7-modC, pJP3416-GA7-modD, pJP3416-GA7-modE, pJP3416-GA7-modF và pJP3416-GA7-modG được sử dụng để tạo ra mầm sinh dưỡng *Brassica* biến nạp và cây *Brassica napus*, *Camelina sativa* và *Arabidopsis thaliana* và hạt cây con. Số liệu của pJP3416-GA7-modB được trình bày trong Ví dụ tiếp theo.

Tám cây *A. thaliana* pJP3416-GA7-modB biến đổi gen và 15 cây *A. thaliana* pJP3416-GA7-modG biến đổi gen được tạo ra. Quan sát thấy mức DHA trong nhóm hạt pJP3416-GA7-modB nằm trong khoảng từ 3,4% đến 7,2% và mức DHA trong

nhóm hạt T2 pJP3416-GA7-modG nằm trong khoảng từ 0,6 đến 4,1%. Nhiều cây pJP3416-GA7-modB cao nhất được gieo trên môi trường lựa chọn và cây giống con sống sót được chuyển sang thể hệ tiếp theo. Hạt được phân tích hàm lượng DHA. Do hạt T1 phân nhóm là quần thể đang phân ly với gen biến đổi và bao gồm các cây phân ly vô hiệu bất kỳ, hy vọng là các hạt đồng hợp tử từ cây thể hệ con sẽ có mức DHA tăng lên, lên đến 20% tổng hàm lượng axit béo trong dầu hạt. Các cấu trúc biến đổi khác được sử dụng để gây biến nạp *A. thaliana*. Mặc dù chỉ thu được một số lượng nhỏ dòng biến nạp, không thu được các mức DHA nào cao hơn so với cấu trúc modB.

Cấu trúc pJP3416-GA7-modB cũng được sử dụng để tạo ra cây *B. napus* biến nạp của cây trồng Oscar và trong dòng gây giống được ký hiệu là NX005. Thu được mười cá thể cây biến nạp (T0) cho đến nay đối với cây biến nạp Oscar, và 20 dòng độc lập đối với NX005. Hạt (hạt T1) được thu hoạch từ các dòng biến đổi gen. Các nhóm hạt được thử nghiệm về mức DHA trong dầu hạt, và hai dòng thể hiện mức cao nhất được chọn, đó là các dòng được thiết kế CT132.5 (ở cây trồng Oscar) và CT133.15 (ở NX005). Hai mươi hạt từ CT132.5 và 11 hạt từ CT133.15 được để hút ẩm, sau hai ngày, dầu được chiết từ một nửa lá mầm từ mỗi hạt riêng lẻ. Một nửa lá mầm còn lại với trục mầm được giữ lại và nuôi cấy trên môi trường để duy trì dòng cây con đặc hiệu. Thành phần axit béo trong dầu được xác định; số liệu được trình bày trong Bảng 16 đối với CT132.5. Mức DHA của mười trong số 20 hạt được phân tích nằm trong khoảng từ 7-20% tổng hàm lượng axit béo như được xác định bởi phân tích GC. Các hạt khác có mức DHA thấp hơn 7% và có thể chứa bản sao một phần (không hoàn toàn) của T-ADN từ pJP3416-GA7-modB. Dòng biến đổi gen dường như chứa nhiều đoạn lồng biến đổi gen không được liên kết về di truyền. Hạt của dòng biến đổi gen CT133.15 biểu hiện mức DHA nằm trong khoảng từ 0-5%. Hạt không có DHA có khả năng là của cây phân ly vô hiệu. Số liệu này khẳng định rằng cấu trúc modB thể hiện tốt sự sản sinh DHA trong hạt cải dầu.

Cấu trúc pJP3416-GA7-modB và pJP3416-GA7-modF cũng được sử dụng để tạo ra cây *Camelina sativa* biến nạp. Thu được ít nhất 24 cây biến nạp độc lập được (T0) và thử nghiệm chi tiết hơn bằng cách phân tích cây con. Hạt (hạt T1) được thu hoạch từ các dòng biến đổi gen. Các nhóm hạt được thử nghiệm về mức DHA trong dầu hạt, và 6 dòng thể hiện mức DHA cao nhất (từ 6% đến 9%) được lựa chọn. Mức DHA trong 20 hạt T1 từ mỗi dòng là các hạt được phân tích nhiều nhất thể hiện mức

DHA nằm trong khoảng từ 6 đến 14% tổng hàm lượng axit béo như được xác định bởi phân tích GC. Thành phần axit béo trong dầu được xác định; số liệu được trình bày trong Bảng 17 của nhiều hạt biến đổi gen. Các số liệu này khẳng định rằng cả hai cấu trúc modB và modF biểu hiện tốt khả năng tạo ra DHA ở hạt *Camelina*.

Bảng 16. Profil axit béo của một nửa lá mầm của hạt *B. napus* biến đổi gen T1 nảy mầm chứa cấu trúc modB. Quan sát thấy lên đến 18,1% DHA ở một số mẫu chứa cao hơn 10% DHA.

Hạt	14:0	16:0	16:1d3?	16:1	16:3	18:0	18:1	18:1d11	18:2	18:3n6	18:3n3	20:0	18:4n3	C20:1d11	20:1d13	C20:2n6	C20:3n3	C22:0	20:4n3	20:5n3	22:3n3	C24:0	C24:1	22:5n3	C22:6n3
1	0,1	4,2	0,1	0,1	0,2	1,8	29,9	2,5	9,9	0,1	38,4	0,5	0,8	1,0	0,0	0,1	2,1	0,3	2,8	0,3	0,1	0,2	0,2	0,5	3,9
2	0,1	4,7	0,1	0,1	0,2	4,0	23,0	2,3	7,4	0,3	29,3	1,0	4,3	1,1	0,0	0,1	1,9	0,4	6,9	1,0	0,0	0,3	0,1	1,7	9,5
3	0,1	3,7	0,2	0,1	0,2	1,8	55,1	1,9	4,7	0,2	15,2	0,8	1,8	1,4	0,0	0,1	0,3	0,5	11,3	0,0	0,0	0,3	0,2	0,0	0,0
4	0,1	4,6	0,2	0,2	0,2	2,9	22,1	1,8	6,6	0,4	26,5	1,0	7,2	1,0	0,0	0,1	0,8	0,5	11,2	1,9	0,0	0,2	0,2	1,7	8,7
5	0,1	4,0	0,1	0,1	0,2	1,7	27,4	2,1	8,1	0,3	26,4	0,6	2,8	1,0	0,0	0,1	1,5	0,3	7,6	1,5	0,0	0,1	0,1	1,8	12,2
6	0,1	3,5	0,1	0,1	0,2	1,6	59,8	2,0	4,3	0,1	18,5	0,6	0,5	1,3	0,0	0,0	0,7	0,3	6,0	0,0	0,0	0,2	0,1	0,0	0,0
7	0,1	6,0	0,3	0,3	0,3	1,7	16,6	2,6	23,9	1,0	23,2	0,6	5,4	0,8	0,0	0,2	0,6	0,4	2,6	1,1	0,0	0,3	0,3	1,7	9,9
8	0,1	4,9	0,1	0,1	0,2	2,7	12,9	1,4	11,7	0,3	34,3	0,9	5,0	0,9	0,0	0,2	2,4	0,5	4,1	1,3	0,0	0,2	0,2	1,8	13,8
9	0,1	3,9	0,1	0,1	0,1	2,4	41,6	1,7	21,5	0,0	23,4	0,7	0,0	1,2	0,0	0,1	2,2	0,4	0,0	0,0	0,1	0,3	0,2	0,0	0,0
10	0,1	3,7	0,2	0,1	0,1	2,1	30,9	1,7	19,2	0,4	23,6	0,7	2,1	1,1	0,0	0,1	1,5	0,4	3,6	0,6	0,0	0,2	0,1	0,7	6,9
11	0,1	5,7	0,4	0,3	0,2	3,8	41,2	2,4	26,7	2,1	7,2	1,3	0,3	1,2	0,0	0,2	0,3	0,8	4,8	0,0	0,0	0,6	0,3	0,0	0,0
12	0,1	4,6	0,0	0,1	0,2	2,4	25,5	1,7	16,1	0,3	28,9	0,8	3,9	1,1	0,0	0,1	1,9	0,4	3,9	0,6	0,0	0,2	0,0	1,1	6,2
13	0,1	4,3	0,1	0,1	0,1	4,2	19,4	1,6	9,2	0,1	45,5	1,0	0,2	1,1	0,0	0,1	5,2	0,4	2,6	0,3	0,2	0,2	0,1	0,4	3,4
14	0,1	6,3	0,2	0,2	0,2	4,0	10,5	2,3	8,4	0,3	31,1	1,3	3,9	0,8	0,0	0,1	2,3	0,6	4,6	1,8	0,1	0,3	0,2	2,5	18,1
15	0,1	5,1	0,1	0,2	0,2	3,3	16,8	2,4	11,2	0,3	28,8	1,0	4,5	0,9	0,0	0,1	2,1	0,6	3,2	1,5	0,1	0,3	0,1	1,8	15,1
16	0,1	4,4	0,1	0,1	0,2	4,0	16,2	1,5	11,6	0,2	33,5	0,9	2,8	1,1	0,0	0,2	3,7	0,4	4,6	0,7	0,1	0,3	0,1	1,3	12,1
17	0,2	7,2	0,2	0,2	0,2	4,9	15,0	2,1	8,9	0,3	25,9	1,4	5,1	0,9	0,0	0,0	1,6	0,8	4,9	2,1	0,0	0,6	0,3	2,2	15,0
18	0,1	4,0	0,1	0,1	0,2	2,3	64,8	1,2	7,2	0,1	12,5	1,0	3,5	1,5	0,0	0,1	0,0	0,7	0,0	0,0	0,0	0,5	0,2	0,0	0,0
19	0,1	3,9	0,1	0,1	0,2	4,6	36,9	1,7	7,1	0,2	28,6	1,2	1,8	1,2	0,0	0,1	1,4	0,5	4,3	0,4	0,0	0,4	0,1	0,8	4,3
20	0,1	4,8	0,1	0,1	0,2	6,0	18,5	1,2	12,8	0,2	34,8	1,4	2,4	1,1	0,0	0,1	3,4	0,6	3,2	0,4	0,1	0,3	0,1	0,7	7,6

Bảng 17. Profin axit béo của hạt *C. sativa* biến đổi gen T1 chứa cấu trúc modB hoặc modF

	C14:0	C16:0	C16:1	C18:0	C18:1	C18:1d11	C18:2	C18:3n6	C18:3n3	C20:0	18:4n3	C20:1d11	20:1d13	C20:2n6	C20:3n6	C20:4n6	C20:3n3	C22:0	20:4n3	C22:1	20:5n3	C22:2n6	22:3n3	C24:0	C24:1	22:5n3	C22:6n3
123-8	0,1	7,3	0,0	5,2	7,9	1,0	7,7	0,7	29,9	2,3	6,0	7,1	0,4	0,7	0,0	0,0	0,9	0,4	1,3	1,0	4,6	0,0	0,1	0,2	0,3	1,5	13,3
123-12	0,1	8,3	0,0	5,3	7,2	1,2	8,7	0,9	27,2	2,5	5,7	6,9	0,5	0,7	0,0	0,1	0,9	0,5	1,5	1,2	5,0	0,0	0,1	0,2	0,4	1,5	13,2
5-8	0,1	8,3	0,1	3,5	9,4	1,3	8,1	1,1	29,0	1,0	9,3	7,9	0,4	0,6	0,0	0,0	0,8	0,2	0,4	0,8	3,4	0,0	0,1	0,2	0,4	0,9	12,6
5-9	0,1	8,1	0,0	3,5	9,4	1,2	8,4	1,2	29,2	1,0	9,0	8,1	0,3	0,6	0,0	0,0	0,8	0,2	0,5	0,8	3,5	0,0	0,1	0,1	0,3	0,9	12,6
17-10	0,1	8,7	0,1	4,1	8,4	1,3	5,5	1,2	26,1	1,6	11,8	7,2	0,3	0,0	0,4	0,03	0,8	0,3	0,4	0,7	5,5	0,0	0,0	0,2	0,3	1,3	13,5
17-26	0,1	8,8	0,1	5,5	5,0	1,3	7,6	0,9	27,8	2,7	10,1	6,2	0,3	0,0	0,7	0,03	1,1	0,6	0,5	1,0	4,7	0,1	0,1	0,3	0,4	1,0	13,1

Các tác giả sáng chế cần nhắc rằng, thông thường, hiệu quả của hoạt tính enzym hạn gây hạn chế tốc độ trong con đường tổng hợp DHA trong thể biến nạp nhiều bản sao T-ADN có thể cao hơn so với thể biến nạp một bản sao T-ADN, hoặc có thể tăng lên bằng cách lồng vào trong gen phức T-ADN mã hóa enzym có thể bị hạn chế trong con đường này. Bằng chứng về tầm quan trọng có thể của thể biến nạp đa phiên bản có thể thấy ở hạt *Arabidopsis* được gây biến nạp bằng cấu trúc GA7 (Ví dụ 2), trong đó cây có DHA thu được cao nhất có ba T-ADN lồng vào trong hệ gen của thể chủ. Gen phức có thể là giống nhau, hoặc tốt hơn nếu chúng là các biến thể khác nhau mã hóa cùng một polypeptit, hoặc dưới sự kiểm soát của gen khởi đầu khác nhau có mẫu hình biểu hiện trùng nhau. Ví dụ, có thể đạt được sự biểu hiện tăng lên bằng cách biểu hiện vùng đa mã hóa $\Delta 6$ -desaturaza, thậm chí tại đó tạo ra cùng một protein. Trong pJP3416-GA7-modF và pJP3416-GA7-modC, ví dụ, hai phiên bản của *M. pusilla* $\Delta 6$ -desaturaza có mặt và được biểu hiện bởi các gen khởi đầu khác nhau. Trình tự mã hóa có khả năng codon khác nhau và do đó trình tự nucleotit khác nhau, làm giảm khả năng làm câm gen hoặc hiệu quả đồng ức chế nhưng cùng tạo ra một protein.

Ví dụ 6. Hoạt tính của cấu trúc đặc hiệu hạt trong mầm sinh dưỡng

Nhằm tạo ra hệ thử nghiệm nhanh để dự đoán sự biểu hiện của cấu trúc gen ở hạt dưới sự điều khiển của gen khởi đầu đặc hiệu với hạt, hệ mầm sinh dưỡng được thiết lập đối với *Brassica napus*. Nó sử dụng vector để biểu hiện yếu tố phiên mã LEC2 có vai trò khởi động quá trình phát triển mầm sinh dưỡng. Để chứng minh, vector kép 35S:LEC2 và pJP107 (Petrie et al., 2010a và b) được đưa vào trong chủng *Agrobacterium tumefaciens* AGL1 thông qua quy trình điện thẩm chuẩn và thể biến nạp *Agrobacterium* sử dụng để đồng biến nạp *Brassica napus* bằng cách nuôi cấy đồng thời. Vùng T-ADN của pJP107 chứa gen mã hóa $\Delta 9$ -elongaza *Isochrysis galbana*, $\Delta 8$ -desaturaza *P. salina* và $\Delta 5$ -desaturaza *P. salina* với mỗi gen được biểu hiện bởi gen khởi đầu đặc hiệu với hạt. Thể biến nạp đối chứng sử dụng vector 35S:LEC2 đơn độc. Sự biểu hiện 35S:LEC2 khiến cho tạo ra mầm sinh dưỡng trong môi trường nuôi cấy mô trực tiếp từ mô thể chai *B. napus* biến nạp như mô tả trong Ví dụ 1.

Phân tích axit béo chỉ ra rằng các gen đặc hiệu hạt trên T-ADN của cấu trúc pJP107 được biểu hiện trong mầm sinh dưỡng biến đổi gen trong sự có mặt của gen LEC2 đồng biến nạp và có tác động để tạo ra ARA ($20:4^{\Delta 5,8,11,14}$) từ LA và EPA ($20:5^{\Delta 5,8,11,14,17}$) từ ALA. Số liệu của ba mầm sinh dưỡng được biến nạp đồng thời được

trình bày trong Bảng 18 và thành phần axit béo của mỗi mầm sinh dưỡng so với thành phần axit béo của dầu hạt từ hạt *Brassica napus* được biến đổi gen, và biểu hiện, ở T-ADN của pJP107 (Petrie et al., 2010a và b). Quan sát thấy tổng số phần trăm tương tự nhau của ARA và các axit béo trung gian EDA (20:2 ω 6) và DGLA (20:3 ω 6), cũng như hiệu suất chuyển hóa, trong mô mầm sinh dưỡng so với profin hạt được biến nạp ổn định. Quan sát thấy kết quả tương tự trong thành phần axit béo của hạt T₂ và mầm sinh dưỡng biến đổi gen ổn định: axit béo ω 6 lần lượt ở mức 26,6% và 25,6% (trung bình), trong khi mức ARA lần lượt là 9,7% và 10,6% (trung bình).

Khi 35S:LEC2 đơn độc được đưa vào và mầm sinh dưỡng được phân tích trong một khoảng thời gian, thấy rằng profin axit béo thay đổi sang profin giống với mầm hơn với mức 18:3 ^{Δ 9,12,15} giảm và 18:1 ^{Δ 9} tăng theo kiểu tỷ lệ nghịch (Fig.8). Các kết quả chỉ ra rằng mầm sinh dưỡng thực sự đã trở nên giống với hạt về đặc điểm và các gen trên T-ADN từ pJP107 đã được biểu hiện. Điều này chứng minh rằng hệ mầm sinh dưỡng cho phép xác định đặc điểm nhanh chóng cấu trúc biến đổi gen đặc hiệu với hạt ở *B. napus* mà không đòi hỏi một quy trình hoàn chỉnh để tạo ra cây biến đổi gen và, từ đó, hạt già.

Bảng 18. Thành phần axit béo của lipid thu được từ mầm sinh dưỡng *Brassica napus* tạo ra bằng cách đồng biến nạp pJP107 bằng 35S:LEC2, so với đối chứng không được gây biến nạp (WT) và hạt T₂ được gây biến nạp bằng pJP107. Hiệu suất chuyển hóa enzym riêng biệt được trình bày trong ngoặc đơn sau thứ tự enzym tương ứng. D9-elo là Δ9-elongaza, D8-Des là Δ8- desaturaza và D5-Des là Δ5-desaturaza.

	WT	Hạt biến đổi gen pJP107 T2	LEC2:#45	LEC2:#57	LEC2:#58
18:1 ^{Δ9}	57,2	45,7	3,8	2,5	1,9
18:2 ^{Δ9,12}	19,1	8,7	10	10,6	10
18:3 ^{Δ9,12,15}	10,2	4,1	22,5	27,5	24,2
20:2 ^{Δ11,14}		7,1 ± 1,9 (67% D9-elo)	5,2 (61,8% D9-elo)	3,7 (56,7% D9-elo)	4,6 (61,8% D9-elo)
20:3 ^{Δ8,11,14}		1,1 ± 0,2 (60% D8-des)	0,4 (67% D8-des)	0,2 (73% D8-des)	0,4 (73% D8-des)
20:4 ^{Δ5,8,11,14}		9,7 ± 0,9 (90% D5-des)	10,6 (98% D5-des)	10 (96% D5-des)	11,2 (97% D5-des)
20:3 ^{Δ11,14,17}		4,0 ± 0,8	9,9	5,5	7,3
20:4 ^{Δ8,11,14,17}		0,3 ± 0,1	0,4	0,3	0,4
20:5 ^{Δ5,8,11,14,17}		2,4 ± 0,2	7,6	6,4	7,9
Tổng sản phẩm mới		24,6	34,1	26,1	31,8

Sử dụng hệ tương tự để tạo ra mầm sinh dưỡng, tế bào *Brassica napus* được gây biến nạp riêng biệt bằng pJP3416-GA7-modB và pJP3416-GA7-modD. Thu được 42 mầm, 18 mầm đối với modB và 24 mầm đối với modD. Toàn bộ lipit được chiết từ mầm và phân tích thành phần axit béo. Các mầm này chứa DHA với lượng nằm trong khoảng từ 0% và lên đến 16,9% (Bảng 19). Các kết quả với 0% DHA được cho là do sự sát nhập chỉ T-ADN một phần hoặc sự lồng vào vùng phiên mã của hệ gen. Thấy rằng tỷ lệ toàn bộ $\omega 3$ FA (bao gồm ALA) với toàn bộ $\omega 6$ FA (bao gồm LA) là 2,3 đối với mầm #270 và 11,96 đối với mầm #284. Tỷ lệ toàn bộ $\omega 6$ FA (bao gồm LA) với toàn bộ $\omega 3$ FA (bao gồm ALA) là 0,08 đối với #284. Tỷ lệ $\omega 6$ FA mới (trừ LA) với $\omega 3$ FA mới (trừ ALA) là 0,03 đối với #284. Hiệu suất chuyển hóa toàn phần được tính là: (đối với mầm #270, #284) OA thành EPA = 14,0%, 29,8%; OA thành DHA = 9,7%, 24,2%; LA thành EPA = 15,4%, 30,7%; LA thành DHA = 10,7%, 25,0%; ALA thành EPA = 22,1%, 33,3%; ALA thành DHA = 15,3%, 27,0%. Các hiệu suất này là tương đương, hoặc cao hơn ở trường hợp #284, so với các hiệu suất thấy được với dòng pJP3416-GA7 *Arabidopsis* T₃, điều này chỉ ra rằng vectơ pJP3416-GA7-modB có khả năng hoạt động tốt trong tế bào *B. napus*. Mức SDA thấp hơn 3,0%, chỉ ra rằng $\Delta 6$ -elongaza hoạt động thậm chí tốt hơn cấu trúc GA7. Hiệu suất enzym riêng biệt trong #284 là: $\Delta 12$ -desaturaza, 97,4%; $\omega 3$ -desaturaza, 92,3%; $\Delta 6$ -desaturaza, 38,2%; $\Delta 6$ -elongaza, 88,2%; $\Delta 5$ -desaturaza, 98,8%; $\Delta 5$ -elongaza, 94,1%; và $\Delta 4$ -desaturaza, 86,3%. Tổng lượng axit bão hòa là 21,2%, tổng lượng axit béo một liên kết chưa bão hòa là 10,2%, tổng lượng axit béo nhiều liên kết chưa bão hòa là 68,6%.

Các tác giả sáng chế tin rằng đây là mức DHA cao nhất đạt được ở tế bào *B. napus* cho đến nay, trừ các số liệu được mô tả dưới đây. Điều này cũng chứng minh là sự biến đổi pJP3416-GA7-modB tương ứng với pJP3416-GA7 là có hiệu quả trong việc làm tăng mức biểu hiện của gen $\Delta 6$ -desaturaza. Các vectơ kép pJP3416-GA7, pJP3416-GA7-modA, pJP3416-GA7-modC, pJP3416-GA7-modD, pJP3416-GA7-modE và pJP3416-GA7-modF như mô tả trên được đồng biến nạp bằng 35S:LEC2 để tạo ra mầm sinh dưỡng *B. napus* biến nạp. Thấy rằng lên đến 7,0% DHA trong mầm modD, 9,9% ở mầm modE, 8,3% ở mầm modF và 3,6% trong một số lượng ít mầm modG.

Bảng 19. Thành phần axit béo của dầu từ mầm sinh dưỡng *Brassica napus* #270 và #284 được tạo ra bằng cách đồng biến nạp cấu trúc axit DHA đặc hiệu hạt pJP3416-GA7-modB bằng 35S:LEC2, và #286 và #289 (pJP3416-GA7-modD).

	#270	#284	#286	#289
14:0	0,3	0,2	0,2	0,2
16:0	14,0	15,7	17,2	16,6
16:1d9	0,7	0,4	0,8	0,8
16:3	0,5	0,6	1,1	1,3
18:0	2,6	2,4	2,5	2,5
18:1d9	6,6	1,8	1,5	1,1
18:1d11	6,3	6,8	6,5	6,7
18:2	18,9	4,5	10,0	9,8
18:3 ω 6	0,7	0,8	0,3	0,3
18:3 ω 3	33,0	37,2	42,0	41,5
20:0	0,9	0,9	0,8	0,8
18:4 ω 3	1,9	2,8	3,6	4,5
20:1d11	0,2	0,1	0,1	0,1
20:2 ω 6	0,1	0,1	0,1	0,2
20:3 ω 3	0,5	0,0	0,5	0,6
22:0	0,8	1,5	0,6	0,7
20:4 ω 3	0,2	0,9	0,7	0,7
20:5 ω 3	0,7	0,2	0,3	0,3
22:2 ω 6	0,0	1,2	0,0	0,0
22:3 ω 3	0,0	0,1	0,0	0,1
24:0	0,8	1,0	1,0	1,0
24:1	0,8	1,0	0,7	0,9
22:5 ω 3	2,4	2,7	3,2	3,0
22:6 ω 3	7,0	16,9	6,1	6,4

Ví dụ 7. Phân tích TAG từ hạt *A. thaliana* biến đổi gen tạo ra DHA

Phân bố vị trí của DHA trên TAG từ hạt *A. thaliana* biến nạp được xác định bởi NMR. Toàn bộ lipid được chiết từ khoảng 200mg hạt bằng cách đầu tiên ép chúng dưới hexan trước khi chuyển hạt đã nghiền vào ống thủy tinh chứa 10mL hexan. Ống này được làm ấm ở khoảng 55°C trong bồn nước và sau đó được quay và ly tâm. Dung dịch hexan được loại và quy trình được lặp lại với thêm 4 x 10 mL. Dịch chiết được kết hợp lại, cô bằng cách bốc hơi quay và TAG trong lipid chiết tách được tinh chế đi từ các lipid phân cực bằng cách đi qua cột silica ngăn sử dụng 20mL diethyl ete 7% trong hexan. Phân bố vị trí nhóm axyl trên TAG tinh khiết được xác định định lượng như đã được mô tả (Petrie et al., 2010a và b).

Phân tích chỉ ra rằng phần lớn DHA trong toàn bộ dầu hạt nằm ở vị trí *sn*-1/3 của TAG và một phần ít ở vị trí *sn*-2 (Fig.9). Điều này là trái ngược với TAG từ các hạt sinh ARA, điều này chứng minh rằng 50% ARA (20:4 ^{Δ 5,8,11,14}) nằm ở vị trí *sn*-2 của dầu cải biến đổi gen trong khi chỉ 33% được hy vọng là phân bố ngẫu nhiên (Petrie et al.,

2012).

Phân bố vị trí của DHA trong TAG từ hạt *B. napus* được gây biến nạp bằng pJP3416-GA7 hoặc bằng hỗn hợp của pJP3115 và pJP3116 được xác định về cơ bản là bằng phương pháp tương tự.

Toàn bộ lipit từ hạt *A. thaliana* biến đổi gen cũng được phân tích bởi LC-MS ba phổ bốn cực để xác định kiểu triaxylglycerol (TAG) chứa DHA chính (Fig.10). Kiểu TAG chứa DHA nhiều nhất được thấy là DHA-18:3-18:3 (TAG 58:12; danh pháp này không có tính chất mô tả phân bố vị trí) có mức cao thứ hai là DHA-18:3-18:2 (TAG 58:11). Quan sát thấy Tri-DHA TAG (TAG 66:18) trong toàn bộ dầu hạt, mặc dù thấp nhưng vẫn ở mức có thể phát hiện được. Các loại TAG chứa DHA chính khác bao gồm DHA-34:3 (TAG 56:9), DHA-36:3 (TAG 58:9), DHA-36:4 (TAG 58:10), DHA-36:7 (TAG 58:13) và DHA-38:4 (TAG 60:10). Sự giống nhau của hai TAG chứa DHA chính được khẳng định thêm bởi Q-TOF MS/MS.

Ví dụ 8. Dự đoán mức sinh DHA trong hạt *B. napus*

Đã chứng minh mức tạo ra DHA hiệu quả ở hạt *Arabidopsis* ở mức 15% bởi cấu trúc gen GA7 trong Ví dụ 2. Cấu trúc tương tự ở hạt *Brassica napus* chỉ sinh ra khoảng 1,5% DHA trong nhiều thể biến nạp (nhưng không phải tất cả), chủ yếu là do sự biểu hiện kém gen $\Delta 6$ -desaturaza của GA7 trong loài này (Ví dụ 4). Dựa trên nhận biết rằng các biến đổi cấu trúc GA7 sẽ khắc phục vấn đề biểu hiện gen $\Delta 6$ -desaturaza kém (xem Ví dụ 5, như được chứng minh trong Ví dụ 6), việc tính toán được thực hiện để xác định profin axit béo có thể của hạt biến đổi gen *B. napus* biểu hiện gen từ biến thể của pJP3416-GA7, trong đó mỗi enzym mã hóa gen biến đổi hoạt động có hiệu quả như quan sát thấy ở *A. thaliana* có cấu trúc GA7. Thành phần axit béo dự đoán của ba công thức tính (#1, #2, #3) được trình bày trong Bảng 20. Điều này là dựa trên thành phần axit béo kiểu đại (không được gây biến nạp) của *B. napus* bao gồm 59% axit oleic, 20% LA và 8% ALA. Ba profin axit béo được dự đoán một phần được trình bày trong nửa dưới của Bảng là dựa trên hiệu suất chuyển hóa của mỗi bước enzym nêu trong nửa trên của Bảng. Trong dự đoán #2, kết hợp mức $\Delta 12$ -khử bão hòa ở hiệu suất 75%, mức $\Delta 15$ -khử bão hòa ở 75%, mức $\Delta 6$ -khử bão hòa ở 35%, mức $\Delta 6$ -kéo dài ở 80%, mức $\Delta 5$ -kéo dài ở 90%, mức $\Delta 5$ -kéo dài ở 90% và $\Delta 4$ -khử bão hòa ở 90% sẽ có mức sản sinh khoảng 10% DHA trong hạt cải dầu biến đổi gen đặc trưng. Tất cả các mức hiệu suất

này thấp hơn hoặc bằng với hiệu suất riêng biệt quan sát thấy ở *Arabidopsis*, do đó kết quả dự đoán #2 thể hiện giá trị ước tính có tính bảo tồn. Hiệu suất chuyển hóa được liệt kê trong #3 là các công thức tính gần đúng dựa trên mức chuyển hóa hiệu quả quan sát thấy ở *A. thaliana* được gây biến nạp bằng pJP3416-GA7. DHA được dự đoán là được tạo ra với hàm lượng khoảng 15% toàn bộ axit béo trong dầu hạt tạo ra ở hạt *B. napus*, kết quả này phản ứng mức tạo ra hiệu quả nhất quan sát thấy ở *A. thaliana*. Việc lồng nhiều T-ADN ở trạng thái đồng hợp tử được hy vọng là làm tăng mức DHA đến 20% ở *B. napus*.

Bảng 20. Thành phần axit béo dự đoán với các axit béo đã chọn thể hiện là phần trăm tổng hàm lượng axit béo trong dầu hạt từ *Brassica napus* được gây biến nạp bằng cấu trúc trong con đường tổng hợp DHA, tính trên hiệu suất enzym quan sát được ở *Arabidopsis* biến đổi gen. Các enzym được nêu theo thứ tự trong con đường tổng hợp DHA từ axit oleic. des = desaturaza, elo = elongaza. Thành phần axit béo dự đoán #1, #2 và #3 dựa trên hiệu suất trong nửa trên của Bảng.

Enzym		#1	#2	#3
d12-des		70%	75%	80%
d15-des		70%	75%	80%
d6-des (ω 3)		30%	35%	40%
d6-elo		80%	80%	90%
d5-des		80%	90%	90%
d5-elo		80%	90%	90%
d4-des		80%	90%	90%
Axit béo	WT	#1	#2	#3
18:1d9	59%	26%	22%	18%
18:2 ω 6	20%	19%	17%	14%
18:3 ω 6		1%	2%	3%
18:3 ω 3	8%	30%	32%	34%
18:4 ω 3		3%	3%	2%
20:4 ω 3		2%	1%	2%
20:5 ω 3		2%	1%	2%
22:5 ω 3		1%	1%	2%
22:6 ω 3		5%	10%	15%

Ví dụ 9. Biểu hiện ổn định con đường EPA biến đổi gen ở lá cây

Xây dựng cấu trúc vector kép

Vector kép, pORE04+11ABGBEC_Đậu đũa_EPA_đoạn lồng (SEQ ID NO: 8), được thiết kế để đưa T-ADN vào cây để tổng hợp EPA trong mô lá. Nó chứa gen khảm mã hóa các enzym: $\Delta 6$ -desaturaza *M. pusilla* (SEQ ID NO: 16), $\Delta 6$ -elongaza *P. cordata* (SEQ ID NO: 25) và $\Delta 5$ -desaturaza *P. salina* (SEQ ID NO: 30), mỗi enzym dưới sự kiểm soát của CaMV 35S và gen khởi đầu tiểu đơn vị nhỏ rubisco *A. thaliana* (SSU) (Fig.9). Vector kép được xây dựng cấu trúc bằng cách tổng hợp vùng 199- 10878 của SEQ ID 2 và tách dòng vào trong vector kép thể nhận pORE04 (Coutu et al., 1997) ở vị trí *Bsi*WI và *Kas*I. Ba gen sinh tổng hợp axit béo mã hóa cho các enzym cần có để chuyển hóa ALA, 18:3 ^{$\Delta 9,12,15$} thành EPA, 20:5 ^{$\Delta 5,8,11,14,17$} .

Biểu hiện tạm thời cấu trúc EPA trong tế bào lá *N. benthamiana*

Để kiểm tra cấu trúc này có đúng và có thể biểu hiện các gen hiệu quả trong mô lá hay không, vector khảm pORE04+11ABGBEC_Đậu đũa_EPA_đoạn lồng được đưa vào chủng *A. tumefaciens* AGL1. Vector khảm 35S:p19 cũng được đưa vào trong chủng *A. tumefaciens* AGL1 như mô tả trong Ví dụ 1. Tế bào từ môi trường nuôi cấy của được thẩm lọc vào các mô lá của cây *Nicotiana benthamiana* trong phòng tăng trưởng 24°C. Nhiều lá so sánh trực tiếp được thẩm bởi các mẫu để so sánh tại chỗ trên cả hai mặt của cùng một lá. Các thử nghiệm được tiến hành thành ba bản. Sau khi thẩm, cây được được phát triển thêm trong năm ngày trước khi thu lấy phiến lá để phân tích profin axit béo bằng GC như mô tả trong Ví dụ 1. Phân tích GC bộc lộ rằng vector EPA có chức năng để tạo ra EPA trong lá *Nicotiana benthamiana* (Bảng 21) với mức EPA cao nhất thấy là 10,7% trong lipid lá tổng cộng.

Biến nạp ổn định *Nicotiana tabacum*

Vector khảm pORE04+11ABGBEC_Đậu đũa_EPA_đoạn lồng được sử dụng để gây biến nạp ổn định *Nicotiana tabacum*. Vector này được đưa vào chủng *A. tumefaciens* AGL1 thông qua quy trình điện thẩm chuẩn. Tế bào đã gây biến nạp được phát triển trên môi trường LB rắn được bổ sung kanamycin (50mg/l) và rifampixin (25mg/l) và được ủ ở 28°C trong hai ngày. Một khuẩn lạc đơn được sử dụng để bắt đầu nuôi cấy sạch. Sau 48 giờ phát triển nuôi cấy mạnh, tế bào được tập trung bằng cách ly tâm ở tốc độ 2,000x g và phần nổi bề mặt được loại. Các tế bào được tạo huyền phù lại trong dung dịch sạch chứa 50% LB và 50% MS trung bình ở mật độ OD₆₀₀=0,5.

Bảng 21. Thành phần axit béo của lipid lá tổng cộng từ các trường hợp *Nicotiana benthamiana* biến đổi gen (tạm thời) và *Nicotiana tabacum* (thể gây biến nạp sơ bộ ổn

định) có mức EPA cao nhất từ mỗi thí nghiệm.

		<i>N. benthamiana</i>	<i>N. tabacum</i>
	14:0	0,1	0,1
	16:0	18,5	17,8
	16:1w13t	2,2	3,8
	16:1d9	0,1	0
	16:3	6,2	5,7
	18:0	3,4	3,2
	18:1d11	0,3	0,3
	20:0	0,5	0,5
	22:0	0,2	0,3
	24:0	0,1	0,4
	18:1	2,9	1,6
	18:2ω6	12,6	14,5
Omega-6	18:3ω6	2,3	2,9
	20:2ω6	0,0	0,0
	20:3ω6	0,1	0,0
	20:4ω6	0,3	0,7
Omega-3	18:3ω3	37,1	32,4
	18:4ω3	1,6	1,9
	20:3ω3	0,1	0,3
	20:4ω3	0,3	1,1
	20:5ω3	10,7	12,1
	22:5ω3	0,3	0,4

Các mẫu lá của cây W38 *N. tabacum* phát triển *in vitro* được cắt lấy và cắt thành các phần hình vuông có kích thước khoảng 0,5-1 cm² bằng dao mổ sắc trong khi vẫn nhúng trong dung dịch *A. tumefaciens*. Các mảnh lá *N. tabacum* bị cắt ngập trong dung dịch *A. tumefaciens* được để yên ở nhiệt độ phòng trong 10 phút trước khi được chấm trên giấy lọc vô khuẩn và chuyển lên các đĩa MS mà không bổ sung thêm gì. Sau giai đoạn nuôi cấy đồng thời hai ngày ở 24°C, các mảnh cấy được rửa ba lần bằng môi trường MS lỏng, vô khuẩn, sau đó chấm khô bằng giấy lọc vô khuẩn và đặt lên aga MS chọn lọc được bổ sung benzylaminopurin (BAP) 1,0mg/l, axit indolaxetic (IAA) 0,25mg/L, kanamycin 50mg/L và cefotaxim 250mg/l. Các đĩa được ủ ở 24°C trong hai tuần để phát triển chồi từ các mảnh lá *N. tabacum* đã được gây biến nạp.

Để các cây biến đổi gen bén rễ *in vitro*, các chồi xanh khỏe mạnh được cắt ra và chuyển vào các bình nuôi cấy mô 200 mL chứa môi trường aga MS được bổ sung

25µg/L IAA, 50mg/L kanamycin và cefotaxim 250mg/L. Các chồi biến đổi gen được chuyển vào đất sau khi nảy rễ và phát triển đến khi trưởng thành trong nhà kính. Sau đó, thu lấy phiến lá lớn từ 21 cây biến đổi gen trưởng thành và phân tích profin axit béo như mô tả trong Ví dụ 1. Tất cả các mẫu biến đổi gen được thấy là chứa EPA (Bảng 21) có mức cao nhất EPA trong thể biến nạp sơ bộ bán hợp tử là 12,1% toàn bộ lipid lá. Các mẫu lá cũng chứa một lượng nhỏ DPA (<0,5%) trong lipid của chúng, điều này là do sự kéo dài EPA nhờ hoạt tính $\Delta 5$ -kéo dài yếu của $\Delta 6$ -elongaza. Tỷ lệ toàn bộ $\omega 3$ FA (bao gồm ALA) với $\omega 6$ FA (bao gồm LA) là 2,7. Hiệu suất chuyển hóa toàn phần được tính là: OA thành EPA = 18,4%, LA thành EPA = 18,9%, ALA thành EPA = 25,9%. Việc sản sinh 12,1% EPA là đặc biệt đáng chú ý do các cây này là thể biến nạp bán hợp tử sơ bộ. Hiệu suất ALA thành EPA đặc biệt là rất gần với hiệu suất quan sát được ở hạt biến nạp ổn định. Điều đáng chú ý là cấu trúc này không chứa $\Delta 12$ hoặc $\Delta 15$ -desaturaza để làm tăng sự chuyển hóa OA và LA thành ALA. Hiệu suất tăng lên được hy vọng là do có thêm các hoạt tính này.

Hạt từ thể biến nạp bán hợp tử được thu hoạch và được gieo để tạo ra cây đồng hợp tử.

Tập hợp các hạt ở các dòng EPA trên cùng thể hiện bình thường và hạt từ các dòng từ #10 đến #17 nảy mầm để tạo ra thể hệ T_2 . Tỷ lệ của EPA với dòng vô hiệu (không có EPA) chỉ ra rằng trường hợp #28 là đơn-locus và do đó thể hệ T_3 của dòng này cũng được tạo ra. Phân tích profin axit béo của nhóm T_3 chỉ ra rằng các gen biến đổi là đồng hợp tử không với dòng vô hiệu tìm thấy và lượng ổn định EPA. Lượng EPA trung bình trong lipid lá tổng cộng trong toàn bộ nhóm T_3 là $9,4\% \pm 0,3$ (Bảng 22).

Bảng 22. Profin axit béo đặc trưng của lipid lá tổng cộng từ dòng kiểu dại (WT) và dòng biến đổi gen hoặc được gây biến nạp tạm thời độc lập (EPA). Các loài là *Nicotiana benthamiana* (biến nạp tạm thời), *N. tabacum* (nhóm T_3 được gây biến nạp ổn định), *Vigna unguiculata* (T_1 gây biến nạp ổn định). Sai số là độ lệch chuẩn của nhiều mẫu. Hiệu suất chuyển hóa biểu kiến thể hiện ở dưới dùng mô tả con đường $\omega 3$ và được tính toán là tổng của các FA sản phẩm/ tổng FA của cơ chất + sản phẩm.

		<i>N. benthamiana</i>		<i>N. tabacum</i>		<i>V. unguiculata</i>	
		WT	EPA	WT	EPA	WT	EPA
	16:0	17,7 $\pm$ 0,1	18,7 $\pm$ 0,2	15,0 $\pm$ 0,6	16,5 $\pm$ 0,5	18,0	18,2 $\pm$ 0,2
	16:1 ω 13t	3,2 $\pm$ 0,1	2,2 $\pm$ 0	3,5 $\pm$ 0,1	3,0 $\pm$ 0,3	3,8	2,0 $\pm$ 0,9
	16:3	6,8 $\pm$ 0,1	6,2 $\pm$ 0,1	5,2 $\pm$ 0,5	5,4 $\pm$ 0,3	—	—
	18:0	3,1 $\pm$ 0	3,5 $\pm$ 0,3	2,2 $\pm$ 0,2	2,6 $\pm$ 0,1	1,8	4,5 $\pm$ 0,4

	Thứ yếu	1,4±0	1,4±0,1	3,1±0,4	2,5±0,3	2,3	2,5±0,4
	OA	1,7±0,1	2,7±0,2	1,6±0,3	2,1±0,3	2,0	4,3±1,3
	LA	12,5±0,4	12,7±0,2	17,0±1,1	18,0±0,9	13,4	18,2±3,0
	ALA	53,3±0,2	37,2±0,2	52,2±1,9	34,0±0,6	58,6	38,2±0
Omega-6	GLA	—	2,3±0,1	—	2,3±0,3	—	0,6±0,2
	20:2ω6	0,1±0	—	0,1±0	0,1±0	—	0,1±0
	DGLA	0,1±0	0,1±0	—	—	—	—
	ARA	—	0,3±0	—	0,7±0,1	—	0,2±0
Omega-3	SDA	—	1,5±0,1	—	1,6±0,1	—	1,5±0
	20:3ω3	0,1±0	0,1±0	0,1±0	0,3±0	0,1±0	1,5±0,1
	ETA	—	0,4±0	—	1,1±0,1	—	0,3±0,2
	EPA	—	10,2±0,5	—	9,4±0,3	—	7,1±0,2
	DPA	—	0,3±0,1	—	0,4±0	—	0,8±0,1
Chuyển hóa Omega-3	Δ6-des		25%		27%		20%
	Δ6-elo		88%		87%		85%
	Δ5-des		97%		90%		96%
	Δ5-elo		3%		4%		10%

Các mẫu lá của cây *N. tabacum* đồng hợp tử T₃ được tiếp tục đưa vào phân tích hóa sinh. Toàn bộ lipid được chiết từ nguyên liệu lá đông lạnh và được chiết phân đoạn bởi sắc ký lớp mỏng (TLC -thin-layer chromatography). EPA được thấy là có mặt trong TAG *N. tabacum* ở mức lên đến 30,1% cũng như ở các lipid phân cực là ở mức 6.3% (Bảng 23). Đáng chú ý là EPA được tạo ra bởi con đường biến đổi gen có mặt trong tất cả các phần lipid được định lượng bao gồm TAG, MGDG, DGDG, SQDG, PG, PC, PE, PI và PS. Tất cả các nhóm lipid chứa các mức thấp của chất trung gian mới hoặc axit béo ω6 LC-PUFA có tỷ lệ TAG của axit béo ω3 với ω6 mới là 10:1.

Biến nạp ổn định đậu đũa

Vectơ khảm pORE04+11ABGBEC-đậu đũa-EPA-đoạn lồng được biến nạp vào đậu đũa (*Vigna unguiculata*) như sau. Hạt già khô là nguyên liệu bắt đầu được ưu tiên mặc dù cũng có thể sử dụng hạt thu hoạch từ quả đậu non có trọng lượng tươi tối đa. Hạt khô được đập vỡ bằng tay để tránh việc làm vỡ vỏ hạt và do đó làm giảm sự tạp nhiễm vi sinh vật.

Hạt khô hoặc quả non được ngâm trong etanol 70% trong hai phút và sau đó được xử lý trong 30 phút trong chất tẩy trên thị trường 20% (natri hypoclorit nồng độ cuối là 8,4g/L). Sau đó hạt được rửa nhiều lần bằng nước vô khuẩn. Hạt non được loại

bỏ vô khuẩn khỏi quả trong khi hạt già được để hút ẩm qua đêm. Có thể sử dụng hai mảnh cây khác nhau để tạo ra nhiều chồi, tức là trục mầm và chính lá mầm của nó, tốt hơn là lá mầm có trục mầm chia đôi được gắn vào. Chồi và chop rễ được tách riêng khỏi trục trước khi cắt ở mấu lá mầm, tức là điểm gắn của trục với lá mầm. Từ kết quả so sánh ban đầu của 19 cây trồng và dòng, rõ ràng là các dòng đậu đũa có thể gây biến nạp dễ dàng nhất, điều dự báo duy nhất là cần tối ưu hóa các điều kiện nuôi cấy mô khác nhau đối với mỗi dòng.

Bảng 23. Phân tích các phân lipid của lá non và trưởng thành (non/trưởng thành) triaxylglyxerol (TAG), toàn bộ lipid phân cực (PL), monogalactosyl[di]axylglyxerol (MGDG), digalactosyl[di]axylglyxerol (DGDG), sulfoquinovosyl[di]axylglyxerol (SQDG), phosphatidylglyxerol (PG), phosphatidylcholin (PC), phosphatidyletanolamin (PE), phosphatidylinositol (PI) và phosphatidylserin (PS) từ các mẫu lá *Nicotiana tabacum* biến đổi gen, Sai số là độ lệch chuẩn của nhiều mẫu, Quan sát thấy có đến 30% EPA ở lá TAG với EPA cũng được phân bố trong toàn bộ các lipid phân cực, Cũng quan sát thấy sự chênh lệch giữa prolin lá non và lá trưởng thành về nhiều axit béo,

		Cloplastidic					Extra-cloplastidic				
		TAG	PL	MGDG	DGDG	SQDG	PG	PC	PE	PI	PS
	16:0	9,8 18,3	17,8 23,8	3,1 3,2	18,0 16,8	48,3 50,0	21,0 26,4	22,9 30,0	24,0 30,5	38,7 43,3	31,9 36,2
	16:1ω13t	0 0	3,4 3,1	0 0	0 0	0 0	34,0 32,0	0 0	0 0	0 0	1,0 1,4
	16:3	0,2 0,9	5,6 6,4	14,8 19,4	1,2 1,8	0,4 1,2	0 0	0 0	0 0	0 0	0 0
	18:0	7,3 3,7	2,9 3,9	1,1 1,2	3,5 3,5	5,4 7,1	4,7 6,9	6,6 9,1	11,0 11,4	9,4 9,3	20,2 19,4
	Thư yếu	2,5 2,9	1,4 2,4	1,0 0,4	0,8 1,0	1,9 2,1	1,0 1,5	1,4 1,6	4,9 4,1	6,5 7,7	2,5 3,7
	OA	5,5 0,8	2,8 1,1	0,8 0,3	1,8 1,0	2,7 1,3	5,3 4,9	8,1 2,9	2,5 1,1	2,5 0,8	4,9 2,3
	LA	27,7 13,7	17,3 12,3	8,0 6,8	9,2 10,5	11,7 8,9	17,1 13,2	39,2 25,2	37,9 28,5	22,0 13,4	24,4 17,1
	ALA	9,6 17,2	39,0 34,4	60,3 51,9	61,2 58,6	23,7 21,5	15,7 14,1	7,3 18,2	5,5 10,5	7,6 10,0	4,8 10,5
	GLA	2,5 3,0	1,5 2,1	2,1 3,0	1,1 1,8	1,4 1,9	0,2 0	1,8 2,5	1,7 2,7	0,8 0,9	1,1 1,3
	20:2ω6	0 0	0,1 1,1	0 0	0 0	0 0	0 0	0 0	0,5 0	0 0	0 0
	DGLA	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
	ARA	0,6 0,9	0,1 0,2	0,2 0,4	0 0	0 0	0 0	0,3 0,3	0,4 0,4	0,4 0,6	0 0,2
	SDA	4,0 7,6	1,6 2,0	1,7 2,0	0,6 0,7	1,2 1,2	0 0	2,1 3,6	1,3 2,0	0,8 0,8	0,9 1,6
	20:3ω3	0,2 0,3	0,1 0,2	0 0	0,2 0,3	0 0	0 0,1	0,2 0	0,3 0,4	0 0	0 0
	ETA	0,9 0,2	0,2 0,3	0 0,2	0 0,3	0 0	0 0	0,2 0	0,4 0,2	0,1 0,2	0 0
	EPA	28,8 30,1	6,1 6,3	6,9 11,2	2,3 3,6	3,4 4,6	1,0 0,8	9,7 6,4	9,1 7,8	11,2 12,8	8,4 6,2
	DPA	0,4 0,5	0 0	0 0,1	0 0	0 0	0 0	0,3 0,4	0,5 0,4	0 0,2	0 0,1

Gen đánh dấu lựa chọn được, *bar* hoặc *NptII* có thể được sử dụng để biến nạp. Chủng *Agrobacterium tumefaciens* AGL1 là chủng được ưu tiên để biến nạp đậu đũa. *Agrobacterium* chứa vector pORE04+11ABGBEC-đậu đũa-EPA-đoạn lồng được nuôi cấy qua đêm ở 28°C trên máy rung lắc ở tốc độ 180 vòng/phút và huyền phù được ly tâm ở mức 8000g trong 10 phút và tạo huyền phù lại trong Môi trường 1 (môi trường MS-cơ bản được pha loãng mười lần và chứa 1 sucroza 30g/l, 2- MES 20mM, được điều chỉnh đến độ pH= 5,6 trước khi hấp tiệt khuẩn, bổ sung MS-vitamin đã lọc vô khuẩn, myo-inositol 100mg/l, BAP 1,7mg/l, GA₃ 0,25mg/l, axetosyringon 0,2mM, Na-thiosulphat 250mg/l, dithiothreitol 150mg/l và L-xystein 0,4g/l). Mảnh cây được nhúng mà không lắc trong huyền phù vi khuẩn trong một giờ sau khi cắt ở phần mô phân sinh bằng dao mổ. Các mảnh cây đã xử lý sau đó được chấm trên giấy lọc vô khuẩn và chuyển vào Môi trường 2 (Môi trường 1 chứa 0,8% aga) được phủ giấy lọc lên. Sau bốn ngày nuôi cấy đồng thời, mảnh cây được chuyển vào Môi trường 3 (môi trường MS đầy đủ, được bổ sung myo-inositol 100mg/l, timentin 150mg/L, sucroza 30g/L, MES 3mM, BAP 1,7mg/l, PPT 5mg/l hoặc 25-50mg/l genetixin hoặc kanamycin 150mg/l, aga 0,8 g/L và điều chỉnh đến độ pH =5,6) để thúc chồi và lựa chọn các chồi đã được biến nạp. Sau hai tuần, nhìn thấy chồi đầu tiên. Lá mầm được lấy từ vùng máu của lá mầm và phần nuôi cấy được chuyển vào môi trường sạch 3. Phần nuôi cấy được chuyển vào môi trường sạch 3 mỗi hai tuần sau khi loại bỏ mô chết đi. Bốn môi trường nuôi cấy cấp hai đầu tiên là trên môi trường lựa chọn kanamycin sau đó thay bởi genetixin và kanamycin. Sau sáu lần nuôi cấy cấp hai, các chồi xanh còn sống được chuyển vào Môi trường 4 (Môi trường 3 không có BAP nhưng được bổ sung GA₃ 0,5mg/l, asparagin 50mg/l, 0,1mg/l axit 3-indolaxetic (IAA), timentin 150mg/L, và hoặc PPT (10mg/l), genetixin (50mg/L) hoặc kanamycin (150mg/l), để kéo dài chồi. Các chồi được nuôi cấy cấp hai hai tuần một lần cho đến khi các chồi đơn dài hơn 1cm. Các chồi lớn hơn được chuyển từ đĩa petri vào các bình nuôi cấy (cao 80mm) để phát triển tiếp và lựa chọn.

Phần lớn chồi phục hồi có thể nảy rễ *in vitro*, và các cây đã mọc rễ được chuyển vào đất và để củng cố trong buồng có độ ẩm cao trong 14-21 ngày trước khi chuyển qua môi trường điều kiện nhà kính.

Để tăng cường di chuyển gen vào đậu đũa, bổ sung hợp chất thiol vào môi trường nuôi cấy đồng thời. Việc bổ sung L-xystein, dithiothreitol, và natri thiosulfat

giảm sự nâu hóa các mô bị tổn thương.

Một số lượng lớn mảnh cây đậu đũa có thể được xử lý trong quy trình đơn giản hóa. Ngắn gọn, quy trình này bao gồm các bước sau: cho hạt già sạch hút nước qua đêm, có được mảnh cây là kết quả việc cắt đôi theo chiều dọc các hạt của nó, trục mầm tách (với chồi và ngọn rễ được loại bỏ) vẫn gắn với lá mầm, gây nhiễm bằng chủng *Agrobacterium AGL1* bằng cách cắt tại chỗ ở vùng mô phân sinh, nuôi cấy đồng thời trên môi trường chứa hợp chất thiol trong 4 ngày ở 25°C trong ánh sáng, nảy chồi và kéo dài trên môi trường chứa tác nhân lựa chọn, chồi được bén rễ *in vitro* và chuyển vào điều kiện nhà kính để ra hoa và tạo hạt, phân tích PCR hoặc enzym các cây biến đổi gen giả định, và sàng lọc cho cây con thế hệ tiếp theo bằng PCR hoặc hoạt tính enzym.

Cây con của cây biến đổi gen T₀ là bình thường về phenotyp. Các gen biến đổi được truyền lại cho cây con và cây đồng hợp tử T₂ được xác định bằng cách sàng lọc cây con T₃ của chúng về hoạt tính enzym hoặc bằng PCR.

Bằng hệ biến nạp này, tạo ra khoảng 10 cây biến đổi gen trên mỗi 1000 mảnh cây, tần số này tương tự với tần số biến nạp của các cây họ đậu khác. Phụ thuộc vào cây trồng hoặc dòng cây được gây biến nạp, quy trình này đòi hỏi 5-8 tháng từ khi chuẩn bị mô cây đến khi thu hoạch hạt T₁.

Hệ biến nạp được sử dụng để đưa vector kép pORE04+11ABGBEC-đậu đũa-EPA-đoạn lồng vào cây đậu đũa phục hồi, được gây biến nạp.

Thực hiện sự biến đổi vector kép pORE04+11ABGBEC-đậu đũa-EPA-đoạn lồng trong đó gen mã hóa $\Delta 5$ -elongaza và $\Delta 4$ -desaturaza được bổ sung vào, để tạo ra cấu trúc gen mang lại khả năng chuyển hóa EPA sinh ra thành DHA cao hơn. Cấu trúc này được biến nạp vào cây để sinh DHA trong các mô thực vật.

EPA được thấy là có mặt trong một số cây sống sót khi lựa chọn hóa học. Dòng cao nhất chứa EPA $7,1\% \pm 0,2$ trong lipid lá tổng cộng. Tỷ lệ biến nạp thấp hơn mức thường thấy của đậu đũa chỉ với sáu dòng chắc chắn đã biến đổi gen. Vẫn chưa biết điều gì gây ra tác động này mặc dù đáng chú ý là tỷ lệ cao hơn thông thường các cây biến đổi gen sẽ chứa các vùng T-ADN không hoàn thiện. Có khả năng là kích thước cấu trúc lớn đã góp phần làm giảm hiệu suất này. Hiệu suất chuyển hóa biểu kiến của mỗi enzym trong số ba enzym biến đổi gen cũng được tính toán (Bảng 22). Các kết quả là khá giống nhau trong cả ba loài có mức chuyển hóa thành EPA tốt sau khi phản ứng $\Delta 6$ -khử bão hòa đầu tiên của ALA tự nhiên. Chú ý thấy một số phản ứng $\Delta 5$ -kéo dài

của EPA thành DPA cho dù không có mặt $\Delta 5$ -elongaza đặc hiệu. $\Delta 6$ -elongaza *P. cordata* đã được chỉ ra là có mức hoạt tính $\Delta 9$ -elongaza thấp (tức là chuyển hóa $18:3^{\Delta 9,12,15}$ thành $20:3^{\Delta 11,14,17}$) mặc dù hoạt tính $\Delta 5$ -elongaza không được phát hiện thấy trong thử nghiệm nấm men.

Ví dụ 10. Thử nghiệm các biến thể của gen $\Delta 12$ -desaturaza

Xây dựng cấu trúc vector kép

Một nỗ lực trong việc thử nghiệm và so sánh một loạt các gen $\Delta 12$ -desaturaza khảm, nhiều vector kép tạo ra sẽ được sử dụng để gây biến nạp *A. thaliana* và *B. napus*. Các vector kép pJP3365, pJP3366, pJP3367, pJP3368 và pJP3369 mỗi vector chứa các gen mã hóa enzym $\omega 3$ -desaturaza *P. pastoris* (SEQ ID NO: 12) và $\Delta 6$ -desaturaza *M. pusilla* (SEQ ID NO: 16), và một trong dãy $\Delta 12$ -desaturaza. $\Delta 12$ -desaturaza là từ *Cryptococcus neoformans* (Số hiệu lưu trữ: XP_570226 trong pJP3365), một phiên bản của $\Delta 12$ -desaturaza *Cryptococcus neoformans* chứa đột biến L151M để thử làm tăng hoạt tính gen (trong pJP3366), *Lachancea kluyveri* (SEQ ID NO: 10 trong pJP3367), *Synechocystis* PCC6803 (Số hiệu lưu trữ: BAA18169 trong pJP3368) và *Crepis palaestina* (Số hiệu lưu trữ: CAA76157, Lee et al., 1998, trong pJP3369). Desaturaza *Crepis* là desaturaza thực vật duy nhất trong các chuỗi này; các enzym còn lại là enzym nấm. Các vector được tạo ra bằng cách lồng vùng mã hóa protein được tối ưu hóa theo codon, ngoại trừ $\Delta 12$ -desaturaza *Crepis palestina* là kiểu dại, của mỗi $\Delta 12$ -desaturaza vào vị trí *NotI* của vector pJP3364 (xem Fig.12), theo hướng được liên kết điều khiển với gen khởi đầu FP1 để gây ra sự biểu hiện đặc hiệu với hạt của mỗi desaturaza. Vector pJP3364 cũng chứa gen khảm mã hóa $\omega 3$ -desaturaza *P. pastoris* và $\Delta 6$ -desaturaza *M. pusilla*, mỗi gen dưới sự kiểm soát của gen khởi đầu đặc hiệu với hạt (Fig.12). Sự kết hợp ba enzym sinh tổng hợp axit béo, cụ thể là $\Delta 12$ -desaturaza, $\omega 3$ -desaturaza và $\Delta 6$ -desaturaza, được thiết kế để tập trung vào con đường chuyển hóa axit oleic ($18:1^{\Delta 9}$) thành SDA ($18:4^{\Delta 6,9,12,15}$). Do đó, các thử nghiệm được tiến hành để định lượng mức sinh SDA ở hạt được gây biến nạp.

Biến nạp *A. thaliana* và *B. napus* và phân tích

Vector kép khảm được đưa vào chủng *A. tumefaciens* AGL1 và các tế bào từ môi trường nuôi cấy của *Agrobacterium* biến nạp được sử dụng để gây biến nạp cây *A. thaliana* đột biến *fad2* bằng phương pháp nhúng chìm hoa để biến nạp (Clough và Bent,

1998). Sau khi trưởng thành, hạt T_1 từ cây đã được xử lý được thu hoạch và cho lên đĩa MS chứa kanamycin để lựa chọn cây non có gen đánh dấu chọn lọc được *NptII* có trên T-ADN của mỗi vectơ khảm. Cây giống con T_1 sống sót được chuyển vào đất. Sau khi để cho cây tự thụ phấn và phát triển đến trưởng thành, hạt T_2 từ các cây này được thu hoạch và thành phần axit béo của lipit hạt được phân tích bởi GC.

Vectơ khảm pJP3367 cũng được sử dụng để gây biến nạp *B. napus* bằng phương pháp mô tả trong Ví dụ 4 để tạo ra 12 trường hợp biến đổi gen. SDA được tìm thấy nằm trong khoảng từ 0,6% đến 2,2% trong hạt được phân nhóm của cây, và chín hạt riêng lẻ từ cây biến đổi gen, cây biến đổi gen SDA cao nhất được phân tích về thành phần axit béo. Số liệu của thành phần axit béo từ phân tích này được trình bày trong Bảng 24.

Số liệu thể hiện là hoạt tính $\Delta 12$ -desaturaza biểu hiện từ mỗi T-ADN ở cả *A. thaliana* và *B. napus* là thấp ngoài mức dự đoán, tạo ra hiệu suất chuyển hóa enzym là khoảng 20% mà không phải là 70-80% quan sát thấy với cát xét biểu hiện tương tự trong cấu trúc GA7 (Ví dụ 2 và 3). Lý do cho sự biểu hiện tương đối yếu của gen $\Delta 12$ -desaturaza từ các vectơ này là chưa được rõ nhưng có thể có liên quan đến vị trí của gen trong toàn bộ cấu trúc.

Ngược lại, việc phân tích sự biểu hiện RT-PCR chứng minh rằng các gen $\omega 3$ -desaturaza *P. pastoris* và $\Delta 6$ -desaturaza *M. pusilla* trên T-ADN được biểu hiện tương đối tốt ở hạt biến nạp. Bảng 24 bao gồm hiệu suất chuyển hóa $\Delta 6$ -desaturaza trong hạt biến nạp, nằm trong khoảng từ 11% đến 25% ở một dòng *B. napus* biến nạp. Nó cao hơn đáng kể hiệu suất chuyển hóa $\Delta 6$ -desaturaza là khoảng 7% quan sát được ở hạt *B. napus* được gây biến nạp bằng cấu trúc GA7 (Ví dụ 4).

Bảng 24. Thành phần axit béo theo phần trăm của toàn bộ axit béo trong dầu hạt từ các hạt riêng lẻ từ cây *Brassica napus* T_1 được gây biến nạp bằng T-ADN từ pJP3367. SDA (18:4 ω 3) được in đậm.

Mẫu	CT110-3#1	CT110-3#2	CT110-3#3	CT110-3#4	CT110-3#5	CT110-3#6	CT110-3#7	CT110-3#8	CT110-3#9
C14:0	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1
C16:0	4,3	4,2	4,1	4,5	3,8	4,3	4,0	5,0	4,7
16:1d7	0,1	0,1	0,1	0,1	0,0	0,1	0,1	0,1	0,1
C16:1d9	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,3	0,3
16:3	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2

C18:0	1,9	1,9	1,3	1,8	2,1	1,8	2,4	3,1	2,2
C18:1	58,1	59,4	55,5	59,1	62,1	56,0	57,2	52,0	53,2
C18:1d11	3,5	3,6	3,0	3,2	2,9	3,6	3,2	4,4	3,5
C18:2	18,4	17,1	19,2	17,3	17,4	18,7	19,0	20,3	20,2
C18:3 ω 6	0,3	0,2	0,3	0,2	0,2	0,2	0,2	0,2	0,3
C18:3 ω 3	8,2	9,0	11,1	8,6	7,5	10,2	9,8	9,3	9,8
C20:0	0,5	0,5	0,4	0,5	0,6	0,5	0,6	0,7	0,6
18:4 ω 3	2,4	2,0	2,8	2,5	1,4	2,6	1,3	2,4	3,2
C20:1d11	1,1	1,1	1,2	1,2	1,2	1,1	1,2	1,1	1,1
20: liso	0,03	0,03	0,03	0,03	0,01	0,03	0,02	0,03	0,02
C20:2 ω 6	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1
C22:0	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,3	0,2
C24:0	0,1	0,1	0,1	0,2	0,1	0,1	0,2	0,3	0,2
C24:1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1
Δ 6-des %	22,9	17,9	20,3	22,8	15,8	20,2	11,7	20,9	24,9

Do đó, để tận dụng ưu điểm là thu được hiệu suất chuyển hóa Δ 6-desaturaza cao hơn có bởi T-ADN từ pJP3367, cây *B. napus* được gây biến nạp bằng T-ADN này được lai chéo với cây được gây biến nạp bằng T-ADN từ pJP3416-GA7 (Ví dụ 4) để tạo ra cây thế hệ con và hạt mang cả hai T-ADN. Thành phần axit béo của dầu được chiết từ hạt F1 được phân tích bởi GC về hàm lượng DHA và hàm lượng axit béo khác. Quan sát thấy mức DHA tăng lên là kết quả của sự biểu hiện tăng của Δ 6-desaturaza. Cây đồng hợp tử với cả hai T-ADN được tạo ra và sẽ cho mức DHA cao hơn.

Ví dụ 11. Sự tích lũy tăng lên của các axit béo bằng cách sử dụng protein ức chế làm câm

Xây dựng cấu trúc vector kép

WO 2010/057246 mô tả việc sử dụng protein ức chế làm câm (SSP) để làm tăng sự biểu hiện biến đổi gen ở hạt cây. Để chứng minh rằng việc sử dụng protein như vậy có thể làm tăng cường và ổn định sự sản sinh LC-PUFA trong hạt dầu qua nhiều thế hệ, nhiều SSP được lựa chọn để thử nghiệm, cụ thể là V2 (Số hiệu lưu trữ: GU178820.1), p19 (Số hiệu lưu trữ: AJ288943.1), p38 (Số hiệu lưu trữ: DQ286869.1) và P0^{PE} (Số hiệu lưu trữ: L04573.1). p19 là một protein ức chế từ TBSV (Tomato Bushy Stunt Virus – virus gây căn bụi cà chua) gắn kết với các siRNA chiều dài 21 nucleotit trước khi chúng chỉ đạo sự tách điều hướng Argonaute của ARN cùng loài (Voinnet et al., 2003). V2, protein ức chế từ virus làm xoắn vàng lá cà chua (TYLCV- Tomato Yellow

Leaf Curl Virus), gắn kết với protein thực vật SGS3 (Glick et al., 2008), protein được cho là cần có để tạo ra các ARN trung gian sợi kép từ cơ chất ssRNA (Beclin et al., 2002), hoặc gắn kết với cấu trúc dsRNA có phần nhô ra ở đầu 5' (Fukunaga et al., 2009). p38 là protein ức chế từ TCV (Turnip Crinkle Virus) là protein gây cản trở cơ chế gây câm ở cây bằng cách gắn kết với protein Dicer và Argonaute (Azevedo et al., 2010). Protein P0 như P0^{PE} và RPV-P0, từ Polerovirus, chọn đích protein Argonaut để tăng cường phân hủy (Baumberger et al., 2007; Bortolamiol et al., 2007; Fusaro et al., 2012). Do đó, các cấu trúc gen được tạo ra để biểu hiện các SSP ở hạt cây kết hợp với bộ gen sinh tổng hợp axit béo để tạo ra ARA (20:4^{Δ5,8,11,14}) từ LA (18:1^{Δ9,12}), như sau.

Gen sinh tổng hợp axit béo mã hóa Δ9-elongaza *Isochrysis galbana* và Δ8- và Δ5-desaturaza *Pavlova salina* và gen đánh dấu chọn lọc của vi khuẩn được thu lấy trên mảnh ADN đơn từ pJP3010 bằng cách cắt bằng *PmeI* và *AvrII* sinh ra mảnh 9560 bp. Vùng mã hóa Δ9-elongaza trên mảnh này được gắn với gen khởi đầu FAE1 *A. thaliana* (pAtFAE1) và vùng kết thúc phiên mã/polyadenyl hóa conlinin (LuCnl2-3'). Vùng mã hóa desaturaza mỗi vùng được gắn với gen khởi đầu napin FP1 bị cắt cụt (pBnFP1) và vùng kết thúc phiên mã/polyadenyl hóa *nos3'*. Ba gen sinh tổng hợp axit béo trên mảnh này được định hướng và được ngăn cách theo cách tương tự như trong pJP107 (Petrie et al., 2012) và mã hóa cùng protein là pJP107. Mảnh ADN cũng chứa gen pFP1:GFIP:*nos3'* từ pCW141 (xem WO2010/057246) mã hóa protein huỳnh quang xanh (GFP). Gen đánh dấu sàng lọc được sử dụng làm gen đánh dấu đặc hiệu với hạt nhận biết được bằng mắt, cho phép xác định một cách đơn giản và không phải về mặt cấu trúc và bằng cách đó lựa chọn hạt biến đổi gen chứa và biểu hiện gen này.

Mảnh *PmeI-AvrII* được lồng vào trong vị trí *PmeI-AvrII* của mỗi chuỗi năm vectơ, mỗi chuỗi chứa một SSP khác nhau (WO2010/057246), tạo ra các cấu trúc gen ký hiệu là pFN045, pFN046, pFN047, pFN048 và pFN049. Chúng lần lượt chứa các gen mã hóa SSPs P0^{PE}, p38, p19, 35S:V2 và V2. Mỗi gen SSP là dưới sự kiểm soát của gen khởi đầu FP1 và vùng kết thúc phiên mã/polyadenyl hóa *ocs3'* ngoại trừ trong cấu trúc pFN048 là V2 vùng mã hóa dưới sự kiểm soát của gen khởi đầu 35S cấu trúc CaMV. Gen SSP trong mỗi trường hợp nằm trong vùng T-ADN của các cấu trúc này, liền kề với biên phải (RB) của T-ADN. Cấu trúc thứ sáu, pFN050 không có trình tự mã hóa SSP bất kỳ, được tạo ra bằng cách cắt pFN045 bằng *AhdI* và *NheI* sau đó gửi thông tin lại với ADN ligaza để xóa gen FP1:P0^{PE}. Mỗi cấu trúc trong sáu cấu trúc này bao

gồm gen đánh dấu chọn lọc được *NptII* nằm trong T-ADN và liền kề với biên trái của T-ADN. Tất cả các cấu trúc này có dòng gốc RK2 sao chép để duy trì plasmit trong *Agrobacterium*.

Biến nạp *A. thaliana* bằng vector biểu hiện ARA kết hợp với SSPs

Để gây biến nạp genotip MC49 của *Arabidopsis*, là thể đột biến kép *fad2/fae1* có mức axit linoleic cao trong lipid hạt của nó, cây được xử lý bằng phương pháp nhúng chìm hoa (Clough và Bent, 1998) với chủng *A. tumefaciens* GV3101 được gây biến nạp riêng biệt bằng mỗi cấu trúc trong số sáu cấu trúc pFN045-pFN050. Cây đã xử lý được phát triển đến trưởng thành và hạt T₁ được thu hoạch từ chúng được cho lên đĩa trên môi trường MS chứa kanamycin để lựa chọn cây T₁ đã được gây biến nạp. Việc sàng lọc để biểu hiện GFP trong hạt cũng được sử dụng như phương pháp đánh dấu phát hiện bằng thị giác của hạt T₁ biến nạp. Cây giống con sống sót trên các đĩa MS/Kan hoặc cây thu được từ hạt GFP-dương tính được chuyển vào đất và phát triển đến khi trưởng thành để có hạt T₂. Số lượng cây biến nạp thu được lần lượt là 5, 14, 32, 8, 23 và 24 với việc biến nạp bằng pFN045, pFN046, pFN047, pFN048, pFN049 và pFN050. Phát hiện ra rằng ở giai đoạn này gen mã hóa p38 trong pFN046 là không có chức năng và do đó cây được gây biến nạp bằng vector pFN046 được coi là đối chứng phụ tức là về cơ bản giống đối với pFN050.

Thu lấy khoảng 100 hạt T₂ được chia nhóm từ mỗi cây đã được biến nạp để xác định thành phần axit béo của lipid hạt bằng chế phẩm FAME và phân tích GC. Sáu cây giống con T₂ từ mỗi dòng biến đổi gen cũng được trồng để tạo ra hạt T₃.

Thành phần axit béo trong tổng lượng lipid được chiết từ hạt T₂ được xác định bằng GC. Phân tích đã chỉ ra một khoảng ARA và chất trung gian EDA (20:2 ω 6) và DGLA (20:3 ω 6) trong nhóm T₂. Dữ liệu của ARA được trình bày trong Fig.13 và 14.

Fig.13 thể hiện phân tích đồ thị dạng hộp của mức ARA trong lipid của nhóm hạt T₂. Rõ ràng là mức trung bình (mức phân vị phần trăm 50) ARA trong nhóm hạt chứa gen FP1:p19 và 35S:V2 ngoài gen sinh tổng hợp ARA là cao hơn đáng kể so với ở hạt chứa gen khiếm khuyết FP1:p38 hoặc T-ADN đối chứng từ pFP050 không chứa gen SSP. Mức ARA trung bình của hạt được gây biến nạp bằng gen mã hóa p19 và V2 là cao hơn của hạt được gây biến nạp bằng gen p38 hoặc hạt không có SSP (Fig.14). Một dòng FP1:p19 và hai dòng FP1:V2 lần lượt biểu hiện khoảng 19%, 20% và 23% ARA. Có các giá trị lệch ra ngoài và do đó không được đưa vào tính toán theo phân tích đồ thị

dạng hộp. Một số ít hơn cây được gây biến nạp bằng T-ADN chứa các gen FP1:P0^{PE} và 35S:V2 sống sót so với các cấu trúc khác; người ta cho rằng các gen này có thể là gây bất lợi cho sức khỏe của cây trong MC49 chuẩn.

Không chỉ mức ARA khác biệt có ý nghĩa trong số các cấu trúc này, các mức lipid hạt của chất trung gian đầu tiên trong con đường từ LA thành ARA, cụ thể là EDA (20:2ω6), được quan sát thấy là thấp ở các dòng hoặc biểu hiện V2 hoặc biểu hiện p19 so với ở hạt không có SSP hoặc hạt chứa cấu trúc p38 (Fig.15). Ở hạt T₃, một nhóm chứa cấu trúc biểu hiện p19 thể hiện tỷ lệ 38% ARA là phần trăm của toàn bộ axit béo trong lipid hạt.

Một loạt các dòng biến đổi gen T₃ được phát triển đến thế hệ T₄. Mức ARA trong hạt T₄ biểu hiện V2 hoặc là bằng so với thế hệ trước, hoặc mức biểu hiện thực tế tăng lên so với thế hệ gốc T₃ của chúng (Fig.16). Các dòng biểu hiện p19 thể hiện mức ARA dao động hơn. Mức ARA giảm ở một số dòng trong khi ở các dòng khác mức này lại bằng hoặc tăng so với thế hệ gốc T₃. Ngược lại, các dòng chứa gen p38 khiếm khuyết hoặc không có SSP thường thể hiện sự giảm mức ARA và tăng mức các chất trung gian (Fig.18). Ở một số dòng, ARA giảm xuống khoảng 1% và mức EDA tăng lên khoảng 20%. Mức ARA trung bình ở hạt T₄ là cao hơn với các dòng biểu hiện p19 và V2 so với các dòng p38 hoặc không có SSP (Fig.17).

Thử nghiệm này đã chỉ ra rằng sự biểu hiện SSP ở hạt của cây biến đổi gen cùng với các gen khác của con đường sinh tổng hợp LC-PUFA không chỉ là làm tăng mức sản sinh axit béo mong muốn trong thế hệ cây con đầu tiên, mà còn ổn định mức sản sinh axit béo trong các thế hệ tiếp theo như thế hệ cây con thứ ba hoặc thứ tư. Việc sản sinh axit béo tăng lên kéo theo việc giảm các mức axit béo trung gian trong con đường sinh tổng hợp. p19 và V2 của SSP được biểu hiện từ gen khởi đầu đặc hiệu với hạt được ưu tiên. Cấu trúc thiết kế để biểu hiện p38 SSP là khiếm khuyết và không thu được số liệu hữu dụng thu với cấu trúc này. V2 SSP và gen cùng loài từ các virus khác được cho là được ưu tiên đặc biệt do chúng cho phép biểu hiện tối đa các gen trong con đường sinh tổng hợp và đồng thời làm câm các gen khác trong cùng tế bào ở hạt đang phát triển.

Ví dụ 12. Thử nghiệm hàm lượng sterol và thành phần trong dầu

Sterol thực vật từ 12 mẫu dầu thực vật mua được từ các nguồn trên thị trường tại

Úc có đặc trưng bởi các dẫn xuất ete O-trimethylsilyl (OTMSi-ether) trong phân tích GC và GC-MS như mô tả trong Ví dụ 1. Sterol được xác định bởi số liệu lưu, số liệu phổ khối lượng và so sánh với số liệu phổ khối lượng tiêu chuẩn lý thuyết và thực hành. Sterol được định lượng bằng cách sử dụng chất chuẩn nội 5 β (H)-cholan-24-ol. Cấu trúc sterol thực vật cơ bản và cấu trúc hóa học của một số sterol đã xác định được trình bày trong được trình bày trong Fig.19 và Bảng 25.

Dầu thực vật phân tích là từ: vừng (*Sesamum indicum*), ôliu (*Olea europaea*), hướng dương (*Helianthus annuus*), thầu dầu (*Ricinus communis*), cải dầu (*Brassica napus*), rum (*Carthamus tinctorius*), lạc (*Arachis hypogaea*), lanh (*Linum usitatissimum*) và đậu nành (*Glycine max*). Với lượng tích trữ giảm dần, trong tất cả các mẫu dầu, sterol thực vật chính là: β -sitosterol (nằm trong khoảng từ 28-55% của toàn bộ hàm lượng sterol), Δ 5-avenasterol (isofucosterol) (3-24%), campesterol (2-33%), Δ 5-stigmasterol (0.7-18%), Δ 7-stigmasterol (1-18%) và Δ 7-avenasterol (0.1-5%). Nhiều sterol thứ yếu khác được xác định, chúng là: cholesterol, brassicasterol, chalinasterol, campestanol và eburicol. Bốn sterol C29:2 và hai C30:2 sterol cũng được phát hiện, nhưng cần phải tìm hiểu thêm để hoàn thành việc xác định đặc điểm của các thành phần thứ yếu này. Ngoài ra, có nhiều sterol chưa được xác định khác có mặt trong một số dầu nhưng do hàm lượng của chúng rất thấp, phổ khối lượng không đủ mạnh để xác định cấu trúc của chúng.

Hàm lượng thể hiện là mg/g dầu với lượng giảm gần là: dầu cải (6,8 mg/g), dầu vừng (5,8 mg/g), dầu cây lanh (4,8-5,2 mg/g), dầu hướng dương (3,7-4,1 mg/g), dầu lạc (3,2 mg/g), dầu rum (3,0 mg/g), dầu đậu nành (3,0 mg/g), dầu ôliu (2,4 mg/g), dầu thầu dầu (1,9 mg/g). % thành phần sterol và tổng hàm lượng sterol được trình bày trong Bảng 26.

Bảng 25. Tên IUPAC/hệ thống của sterol đã xác định.

Sterol số	Tên thông thường	Tên IUPAC / hệ thống
1	cholesterol	cholest-5-en-3 β -ol
2	brassicasterol	24-metylcholesta-5,22E-dien-3 β -ol
3	chalinasterol/24-metylen cholesterol	24-metylcholesta-5,24(28)E-dien-3 β -ol
4	campesterol / 24-metylcholesterol	24-metylcholest-5-en-3 β -ol
5	campestanol / 24-metylcholestanol	24-metylcholestan-3 β -ol

7	Δ^5 -stigmasterol	24-etylcholesta-5,22E-dien-3 β -ol
9	ergost-7-en-3 β -ol	24-metylcholest-7-en-3 β -ol
11	eburicol	4,4,14-trimthylergosta-8,24(28)-dien-3 β -ol
12	β -sitosterol / 24-etylcholesterol	24-etylcholest-5-en-3 β -ol
13	D5-avenasterol / isofucosterol	24-etylcholesta-5,24(28)Z-dien-3 β -ol
19	D7-stigmasterol / stigmast-7-en-3 β -ol	24-etylcholest-7-en-3 β -ol
20	D7-avenasterol	24-etylcholesta 7,24(28)-dien-3 β -ol

Trong số tất cả các mẫu dầu hạt, sterol thực vật chính thường là β -sitosterol (nằm trong khoảng từ 30-57% của tổng hàm lượng sterol). Có khá nhiều dầu có tỷ lệ của sterol chính khác: campesterol (2-17%), Δ^5 -stigmasterol (0,7-18%), Δ^5 -avenasterol (4-23%), Δ^7 -stigmasterol (1-18%). Dầu từ các loài khác nhau có profin sterol khác với một số dầu có profin hoàn toàn khác biệt. Trong trường hợp dầu cải, nó có tỷ lệ campesterol cao nhất (33.6%), trong khi các loài khác thường có mức thấp hơn, ví dụ lên đến 17% trong dầu lạc. Dầu rum có tỷ lệ tương đối cao của Δ^7 -stigmasterol (18%), trong khi sterol thường thấp ở dầu của các loài khác, lên đến 9% trong dầu hướng dương. Do chúng khác biệt ở mỗi loài, profin sterol do đó có thể được sử dụng để hỗ trợ việc xác định đặc điểm của dầu cây hoặc dầu thực vật cụ thể và để kiểm tra tính xác thực hoặc sự pha trộn với các dầu khác.

Bảng 26. Hàm lượng sterol và thành phần dầu thực vật được tính mạch.

Sterol số*	Tên thông thường của sterol	Vùng	Ồiu	Hương duơng	Hương duơng Ép lạnh	Thần dầu	Cải dầu	Rum	Rum Ép lạnh	Lạc	Lạnh (hạt cây lạnh)	Lạnh (hạt cây lạnh)	Dầu nành
1	cholesterol	0,2	0,8	0,2	0,0	0,1	0,3	0,2	0,1	0,2	0,4	0,4	0,2
2	brassicasterol	0,1	0,0	0,0	0,0	0,3	0,1	0,0	0,0	0,0	0,2	0,2	0,0
3	chalinasterol/24-metylen cholesterol	1,5	0,1	0,3	0,1	1,1	2,4	0,2	0,1	0,9	1,5	1,4	0,8
4	campesterol/24-metylcholesterol	16,2	2,4	7,4	7,9	8,4	33,6	12,1	8,5	17,4	15,7	14,4	16,9
5	campestanol/24-metylcholestanol	0,7	0,3	0,3	0,1	0,9	0,2	0,8	0,8	0,3	0,2	0,2	0,7
6	C29:2*	0,0	0,0	0,1	0,2	0,0	0,1	0,5	0,5	0,0	1,2	1,3	0,1
7	$\Delta 5$ -stigmasterol	6,4	1,2	7,4	7,2	18,6	0,7	7,0	4,6	6,9	5,1	5,8	17,6
8	chưa biết	0,5	1,3	0,7	0,6	0,8	0,7	0,7	1,3	0,4	0,7	0,6	1,3
9	ergost-7-en- 3β -ol	0,1	0,1	1,9	1,8	0,2	0,4	2,7	4,0	1,4	1,4	1,4	1,0
10	Chưa biết	0,0	1,3	0,9	0,8	1,2	0,9	1,8	0,7	1,2	0,7	0,5	0,7
11	eburicol	1,6	1,8	4,1	4,4	1,5	1,0	1,9	2,9	1,2	3,5	3,3	0,9
12	β -sitosterol/24-ethylcholesterol	55,3	45,6	43,9	43,6	37,7	50,8	40,2	35,1	57,2	29,9	28,4	40,2
13	$\Delta 5$ -avenasterol/isofucosterol	8,6	16,9	7,2	4,1	19,3	4,4	7,3	6,3	5,3	23,0	24,2	3,3
14	rượu triterpenoid	0,0	2,4	0,9	1,1	0,0	0,0	1,6	1,9	0,0	0,0	0,0	0,9
15	rượu triterpenoid	0,0	0,0	0,7	0,6	0,0	0,0	2,8	1,8	0,0	0,0	0,3	0,0
16	C29:2*	0,0	0,5	0,7	0,7	1,5	1,2	2,8	1,9	2,0	1,0	0,7	0,5
17	C29:2*	1,0	0,9	2,3	2,4	0,6	0,4	1,3	1,9	0,9	1,0	1,0	1,0
18	C30:2*	0,0	0,0	0,0	0,0	1,9	0,0	0,0	0,0	0,0	0,0	0,0	0,0
19	$\Delta 7$ -stigmasterol/stigmast-7-en- 3β -ol	2,2	7,1	9,3	10,9	2,3	0,9	10,5	18,3	1,1	7,9	8,7	5,6
20	$\Delta 7$ -avenasterol	1,3	0,1	4,0	3,6	0,6	0,2	2,0	4,7	0,7	0,4	0,4	0,6
21	chưa biết	0,7	7,1	0,9	0,8	0,0	0,4	0,3	0,4	0,0	3,0	3,6	0,0
22	chưa biết	0,3	0,0	0,3	0,9	0,0	0,0	1,2	1,3	0,2	0,1	0,0	0,3
23	chưa biết	0,2	0,2	0,3	0,3	0,2	0,1	0,3	0,2	0,2	0,1	0,2	0,5
24	chưa biết	0,0	3,1	0,9	1,3	0,6	0,4	0,2	0,4	0,7	1,7	1,9	0,8
25	chưa biết	0,9	0,4	0,3	0,5	0,3	0,1	0,5	0,7	0,3	0,1	0,1	0,6
26	C30:2	2,2	6,0	4,6	5,7	1,4	0,6	1,0	1,2	1,2	1,2	1,1	5,2

27	chưa biết	0,0	0,4	0,4	0,3	0,3	0,2	0,1	0,2	0,3	0,1	0,0	0,3
	Tổng	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
	Tổng bộ sterol (mg/g dầu)	5,8	2,4	4,1	3,7	1,9	6,8	3,2	3,0	3,2	4,8	5,2	3,0

C29:2* và C30:2* ký hiệu lần lượt sterol C29 với hai liên kết đôi và C30 sterol có hai liên kết đôi

Hai mẫu hướng dương và rum, mỗi mẫu được so sánh, trong mỗi trường hợp nó được tạo ra bằng cách ép lạnh hạt và không lọc, trong khi các dầu khác không được ép lạnh và được lọc. Mặc dù quan sát thấy có một số sự khác biệt, hai nguồn dầu này có thành phần sterol và tổng hàm lượng sterol tương tự nhau, gợi ý rằng việc xử lý và lọc có rất ít tác dụng lên hai thông số này. Hàm lượng sterol trong các mẫu này thay đổi ba lần và nằm trong khoảng từ 1,9 mg/g đến 6,8 mg/g. Dầu cải có hàm lượng sterol cao nhất và dầu thầu dầu có hàm lượng sterol thấp nhất.

Ví dụ 13. Tăng cường tích lũy DHA ở vị trí *sn*-2 TAG

Các tác giả sáng chế đã xem xét rằng việc tích lũy DHA ở vị trí *sn*-2 trong TAG có thể tăng lên bởi sự biểu hiện đồng thời 1-axyl-glycerol-3-phosphat axyltransferaza (LPAAT) cùng với con đường sinh tổng hợp DHA như có được bởi cấu trúc GA7 hoặc biến thể của nó. Các LPAAT được ưu tiên là các LPAAT có thể tác dụng trên axyl-CoA béo nhiều liên kết chưa bão hòa C22 là cơ chất, khiến tăng mức lồng chuỗi nhiều liên kết chưa bão hòa C22 ở vị trí *sn*-2 trong LPA để tạo ra PA, tương ứng với LPAAT nội sinh. Các enzym LPAAT bào tương thường biểu hiện sự ưu tiên cơ chất khác nhau, cụ thể là khi loài này tổng hợp và tích lũy các axit béo bất thường trong TAG. LPAAT2 từ *Limnanthes douglasii* được thể hiện là sử dụng erucoyl-CoA (C22:1-CoA) là cơ chất để tổng hợp PA, ngược lại với LPAAT1 từ cùng loại không thể sử dụng cơ chất C22 (Brown et al., 2002).

Các LPAAT đã biết được xem xét và một số được chọn để thử nghiệm, bao gồm một số chất không được dự tính là làm tăng sự kết hợp DHA ở vị trí *sn*-2, làm đối chứng. Các LPAAT đã biết bao gồm: LPAAT2 *Arabidopsis thaliana*: (SEQ ID NO: 63, Số hiệu lưu trữ: ABG48392, Kim et al., 2005), LPAAT *Limnanthes alba* (SEQ ID NO: 64, Số hiệu lưu trữ: AAC49185, Lassner et al., 1995), Slc1p *Saccharomyces cerevisiae* (SEQ ID NO: 65, Số hiệu lưu trữ: NP_010231, Zou et al., 1997), LPAAT1 *Mortierella alpina* (SEQ ID NO: 67, Số hiệu lưu trữ: AED33305; Patent Mỹ số 7879591) và *Brassica napus* LPAAT (SEQ ID NO: 68 và SEQ ID NO: 69, Số hiệu lưu trữ lần lượt là: ADC97479 và ADC97478). Chúng được chọn để bao gồm ba nhóm enzym LPAAT: 1) LPAAT hạt cây đối chứng đặc trưng có tác dụng yếu trên axit béo nhiều liên kết chưa bão hòa mạch dài bất thường (bao gồm LPAAT *Arabidopsis* và *Brassica*), 2. LPAAT đã được chứng minh là có tác dụng trên các axit béo C22 bằng cách sử dụng C22 axyl-CoA làm cơ chất, trong trường hợp này axit erucic C22:1 (bao gồm LPAAT

Limnanthes và *Saccharomyces*), 3. LPAAT mà các tác giả sáng chế cho là có khả năng sử dụng axit béo nhiều liên kết chưa bão hòa mạch dài như EPA và DHA làm cơ chất (bao gồm LPAAT *Mortierella*).

LPAAT2 *Arabidopsis* (cũng ký hiệu là LPAT2) là enzym vị trí ở lưới nội bào thể hiện là có hoạt tính trên cơ chất C16 và C18, tuy nhiên hoạt tính trên cơ chất C20 hoặc C22 chưa được kiểm tra (Kim et al., 2005). LPAAT2 *Limnanthes alba* được chứng minh là lồng chuỗi axyl C22:1 vào vị trí *sn*-2 của PA, mặc dù khả năng sử dụng DHA làm cơ chất không được kiểm tra (Lassner et al., 1995). LPAAT Slc1p *S. cerevisiae* được chọn được thể hiện là có hoạt tính sử dụng 22:1-CoA ngoài 18:1-CoA làm cơ chất, chỉ ra rằng tính đặc hiệu cơ chất là lớn đối với chiều dài mạch (Zou et al., 1997). Một lần nữa, DHA-CoA và các LC-PUFA khác không được thử nghiệm làm cơ chất. LPAAT *Mortierella* đã được thể hiện là có hoạt tính trên cơ chất axit béo EPA và DHA trong *Yarrowia lipolytica* biến đổi gen (US 7879591).

Các LPAAT khác được xác định bởi các tác giả sáng chế. *Micromonas pusilla* là một vi tảo sản sinh và tích tụ DHA trong dầu của nó, mặc dù phân bố theo vị trí của DHA trên TAG trong loài này chưa được khẳng định. LPAAT *Micromonas pusilla* (SEQ ID NO: 66, Số hiệu lưu trữ: XP_002501997) được xác định bằng cách tìm kiếm trình tự hệ gen *Micromonas pusilla* bằng *Arabidopsis* LPAAT2 làm BLAST trình tự truy vấn. Nhiều trình tự xem xét xuất hiện và trình tự XP_002501997 được tổng hợp để thử nghiệm làm enzym LPAAT tiềm năng có hoạt tính trên C22 LC-PUFA. LPAAT *Ricinus communis* được chú thích là LPAAT giả định trong trình tự hệ gen thầu dầu (Chan et al., 2010). Bốn LPAAT xem xét từ hệ gen thầu dầu được tổng hợp và thử nghiệm trong dịch ly giải lá thô chứa mô lá *N. benthamiana* đã thẩm lọc. Trình tự đang xem xét được mô tả ở đây thể hiện hoạt tính LPAAT.

Một số LPAAT đang xem xét được sắp hàng với các LPAAT đã biết trên cây hệ thống sinh học (Fig.20). Chú ý rằng LPAAT *Micromonas* giả định không cụm lại với C22 LPAAT giả định mà là trình tự khác đi.

Đối với thử nghiệm ban đầu của các LPAAT khác nhau về khả năng sử dụng DHA-CoA làm cơ chất, các cấu trúc gen khảm được tạo ra để biểu hiện cơ bản LPAAT ngoại sinh trong lá *N. benthamiana*, mỗi cấu trúc dưới sự kiểm soát của gen khởi đầu 35S, như sau: 35S:Arath-LPAAT2 (*Arabidopsis* ER LPAAT); 35S:Ricco-LPAAT2; 35S:Limal-LPAAT (*Limnanthes alba* LPAAT); 35S:Sacce-Slc1p (*S. cerevisiae*

LPAAT); 35S:Micpu-LPAAT (*Micromonas pusilla* LPAAT); 35S:Moral- LPAAT1 (*Mortierella alpina* LPAAT). Cấu trúc 35S:p19 không có LPAAT ngoại sinh được sử dụng để làm đối chứng trong thử nghiệm này. Mỗi cấu trúc này được đưa vào qua *Agrobacterium* vào trong lá *N. benthamiana* như mô tả trong Ví dụ 1, và 5 ngày sau khi thâm lọc, vùng lá được xử lý này được cắt và nghiền để làm dịch ly giải lá. Mỗi dịch ly giải chứa LPAAT ngoại sinh cũng như enzym nội sinh để tổng hợp LPA. Các phản ứng *in vitro* được thiết lập bằng cách bổ sung riêng biệt ^{14}C -được đánh dấu- OA, -LA hoặc -ALA (cơ chất C18), -ARA (cơ chất C20) và -DHA (C22) vào dịch ly giải, thành ba bản lặp. Các phản ứng được ủ ở 25°C và mức đưa các axit béo được đánh dấu ^{14}C vào trong PA được xác định bởi TLC. Khả năng của mỗi LPAAT sử dụng DHA tương ứng với ARA và các axit béo C18 được tính toán. LPAAT của Meadowfoam, *Mortierella* và *Saccharomyces* được thấy là có hoạt tính trên cơ chất DHA, với PA được đánh dấu phóng xạ có ở các LPAAT này nhưng không có ở các LPAAT khác. Tất cả LPAAT được khẳng định là hoạt động bởi nguồn axit oleic tương tự.

Để thử nghiệm hoạt tính LPAAT trong hạt, nhiều trình tự mã hóa protein hoặc LPAAT được lồng vào trong vectơ kép dưới sự kiểm soát của gen khởi đầu conlinin (pLuCn11). Các cấu trúc gen thu được lần lượt chứa gen khảm, Cn11:Arath-LPAAT (đối chứng âm), Cn11:Limal-LPAAT, Cn1:Sacce-Slc1p, và Cn11:Moral-LPAAT, sau đó được sử dụng để biến nạp cây *B. napus* và *A. thaliana* để tạo ra thể biến nạp ổn định biểu hiện LPAAT theo cách đặc hiệu với hạt. Cây biến nạp có cấu trúc Cn11:LPAAT được lai chéo với cây biểu hiện cấu trúc GA7 hoặc biến thể của nó (Ví dụ 5), việc này tạo ra DHA trong hạt để làm tăng hợp nhất DHA ở vị trí *sn*-2 của TAG. Các cấu trúc này cũng được sử dụng để gây biến nạp cho cây *B. napus*, *C. sativa* và *A. thaliana* đã chứa cấu trúc GA7 và biến thể của nó (Ví dụ 2 đến 5) để tạo ra cây con mang cả cấu trúc gen gốc và LPAAT. Sự hợp nhất DHA ở vị trí *sn*-2 của TAG tăng lên được hy vọng là có liên quan đến sự hợp nhất DHA ở thực vật không có các gen biến đổi mã hóa LPAAT. Hàm lượng dầu cũng được cải thiện trong hạt, cụ thể là hạt tạo ra các mức DHA cao hơn, đi ngược lại xu hướng thấy ở hạt *Arabidopsis* như mô tả trong Ví dụ 2. Các chuyên gia trong lĩnh vực này biết rõ rằng có thể thực hiện các biến đổi hoặc điều chỉnh với sáng chế được nêu trong các phương án cụ thể mà không vượt quá phạm vi bản mô tả. Các phương án theo sáng chế chỉ nhằm mục đích minh họa mà không giới hạn phạm vi của sáng chế.

Tất cả các tài liệu bàn luận và/hoặc tham khảo ở đây được đưa hoàn toàn vào đây bằng cách viện dẫn.

Tài liệu tham khảo

- Abbadi et al. (2004) Plant Cell 16: 2734-2748.
- Abbott et al. (1998) Science 282:2012-2018.
- Abdullah et al. (1986) Biotech. 4:1087.
- Agaba et al. (2004) Marine Biotechnol. (NY) 6:251-261. Alvarez et al. (2000) Theor Appl Genet 100:319-327.
- Armbrust et al. (2004) Science 306:79-86.
- Attila Kereszt et al. (2007) Nature Protocols 2:948 – 952. Baumberger et al. (2007) Curr. Biol. 17:1609-1614.
- Baumlein et al. (1991) Mol. Gen. Genet. 225:459-467. Baumlein et al. (1992) Plant J. 2:233-239.
- Beaudoin et al. (2000) Proc. Natl. Acad. Sci. U.S.A. 97:6421-6426. Beclin et al. (2002) Curr. Biol. 12:684-688.
- Berberich. et al. (1998) Plant Mol. Biol. 36:297-306. Bortolamiol et al. (2007) Curr. Biol. 17:1615-1621. Broun et al. (1998) Plant J. 13:201-210.
- Brown et al. (2002) Biochem J. 364:795-805. Chapman et al. (2004) Gen. Dev. 18:1179-1186. Chen et al. (2004) The Plant Cell 16:1302-1313. Cheng et al. (1996) Plant Cell Rep. 15:653-657. Cheng et al. (2010) Transgenic Res 19: 221-229.
- Chikwamba et al. (2003) Proc. Natl. Acad. Sci. U.S.A. 100:11127-11132. Cho et al. (1999a) J. Biol. Chem. 274:471-477.
- Cho et al. (1999b) J. Biol. Chem. 274:37335-37339. Clough and Bent (1998) Plant J. 16:735-43.
- Coutu et al. (2007) Transgenic Res. 16: 771-781.
- Damude et al. (2006). Proc Natl Acad Sci USA 103: 9446-9451. Denic and Weissman (2007) Cell 130:663-677.
- Domergue et al (2002) Eur. J. Biochem. 269:4105-4113. Domergue et al. (2002) Eur. J. Biochem. 269:4105-4113. Domergue et al. (2003) J. Biol. Chem. 278: 35115-35126.
- Domergue et al. (2005) Biochem. J. 1 389: 483-490. Dunoyer et al. (2004) The Plant Cell 16:1235-1250. Ellerstrom et al. (1996) Plant Mol. Biol. 32:1019-1027.

- Fujimura et al. (1985) *Plant Tissue Culture Lett.* 2:74. Fukunaga (2009) *EMBO J.* 28:545-55.
- Gamez et al. (2003) *Food Res International* 36: 721-727.
- Garcia-Maroto et al. (2002) *Lipids* 37:417-426. Girke et al. (1998) *Plant J.* 15:39-48.
- Glick et al. (2008) *Proc. Natl. Acad. Sci U.S.A.* 105:157-161. Grant et al. (1995) *Plant Cell Rep.* 15:254-258.
- Hall et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:9320-9324 Hamilton and Baulcombe (1999) *Science* 286:950-952.
- Hamilton et al. (1997) *Gene* 200:107-16.
- Harayama (1998). *Trends Biotechnol.* 16: 76-82.
- Hastings et al. (2001) *Proc. Natl. Acad. Sci. U.S.A.* 98:14304-14309. Hinchee et al. (1988) *Biotechnology* 6:915-922.
- Hoffmann et al. (2008) *J Biol. Chem.* 283:22352-22362. Hong et al. (2002a) *Lipids* 37:863-868.
- Horiguchi et al. (1998) *Plant Cell Physiol.* 39:540-544.
- Horvath et al. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97:1914-1919. Huang et al. (1999) *Lipids* 34:649-659.
- Inagaki et al. (2002) *Biosci. Biotechnol. Biochem.* 66:613-621. Johansen and Carrington (2001) *Plant Physiol.* 126:930-938.
- Kajikawa et al. (2004) *Plant Mol. Biol.* 54:335-52. Kajikawa et al. (2006) *FEBS Lett* 580:149-154.
- Kim et al. (2005) *Plant Cell.* 2005 1073-89. Knutzon et al. (1998) *J. Biol Chem.* 273:29360-6. Koziel et al. (1996) *Plant Mol. Biol.* 32:393-405. Lassner (1995) *Plant Physiol.* 109:1389-94.
- Leonard et al. (2000) *Biochem. J.* 347:719-724. Leonard et al. (2000b) *Biochem. J.* 350:765-770. Leonard et al. (2002) *Lipids* 37:733-740.
- Lewsey et al. (2007) *Plant J.* 50:240-252. Lo et al. (2003) *Hê gene Res.* 13:455-466.
- Lu and Kang (2008) *Plant Cell Rep.* 27:273-8. Mallory et al. (2002) *Nat. Biotech.* 20:622-625.
- Marangoni et al. (1995) *Trends in Food Sci. Technol.* 6: 329-335. Meesapyodsuk

et al. (2007) *J Biol Chem* 282: 20191-20199.

Meng et al. (2008) *J. Gen. Virol.* 89:2349-2358. Meyer et al. (2003) *Biochem.* 42:9779-9788.

Meyer et al. (2004) *Lipid Res* 45:1899-1909.

Michaelson et al. (1998a) *J. Biol. Chem.* 273:19055-19059. Michaelson et al. (1998b) *FEBS Lett.* 439:215-218.

Murashige and Skoog (1962) *Physiologia Plantarum* 15:473-497. Napier et al. (1998) *Biochem. J.* 330:611-614.

Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443-453. Niedz et al. (1995) *Plant Cell Reports* 14:403.

Ow et al. (1986) *Science* 234:856-859.

Parker-Barnes et al. (2000) *Proc. Natl. Acad. Sci. USA* 97:8284-8289. Pereira et al. (2004a) *Biochem. J.* 378:665-671.

Pereira et al. (2004b) *Biochem. J.* 384:357-366. Perrin et al. (2000) *Mol Breed* 6:345-352.

Petrie et al. (2010a) *Metab. Eng.* 12:233-240. Petrie et al. (2010b) *Plant Methods* 11:6:8.

Petrie et al. (2012) *Transgenic Res.* 21:139-147.

Potenza et al. (2004) *In Vitro Cell Dev Biol – Plant* 40:1-22. Prasher et al (1985) *Biochem. Biophys. Res. Commun.* 127:31-36. Qi et al. (2002) *FEBS Lett.* 510:159-165.

Qi et al. (2004) *Nat. Biotech.* 22: 739-745.

Qiu et al. (2001) *J. Biol. Chem.* 276:31561-31566.

Reddy and Thomas (1996) *Nat. Biotech.* 14:639-642. Reddy et al. (1993) *Plant Mol. Biol.* 22:293-300.

Robert et al. (2005) *Func. Plant Biol.* 32:473-479. Robert et al. (2009) *Marine Biotech* 11:410-418.

Ruiz-Lopez et al. (2012) *Transgenic Res.* 21:139-147. Saha et al. (2006) *Plant Physiol.* 141:1533-1543.

Saito et al. (2000) *Eur. J. Biochem.* 267:1813-1818. Sakuradani et al. (1999) *Gene* 238:445-453.

Sato et al. (2004) *Crop Sci.* 44: 646-652.

Sakuradani et al. (2005) *Appl. Microbiol. Biotechnol.* 66:648-654. Sayanova et

al. (2006) *J Biol Chem* 281: 36533-36541.

Sayanova et al. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94:4211-4216. Sayanova et al. (2003) *FEBS Lett.* 542:100-104.

Sayanova et al. (2006) *Planta* 224:1269-1277. Sayanova et al. (2007) *Plant Physiol* 144:455-467. Singh et al. (2005) *Curr. Opin. in Plant Biol.* 8:197-203. Speranza et al. (2012) *Process Biochemistry* (In Press). Sperling et al. (2000) *Eur. J. Biochem.* 267:3801-3811.

Sperling et al. (2001) *Arch. Biochem. Biophys.* 388:293-8. Sprecher et al. (1995) *J. Lipid Res.* 36:2471-2477.

Spychalla et al. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94:1142-1147. Stalker et al. (1998) *J. Biol. Chem.* 263:6310-6314.

Thillet et al. (1988) *J. Biol. Chem* 263:12500-12508. Tonon et al. (2003) *FEBS Lett.* 553:440-444.

Toriyama et al. (1986) *Theor. Appl. Genet.* 205:34.

Trautwein (2001) *European J. Lipid Sci. and Tech.* 103:45-55. Tvrdik (2000) *J. Cell Biol.* 149:707-718.

Venegas-Caleron et al. (2010) *Prog. Lipid Res.* 49:108-119. Voinnet et al. (2003) *Plant J.* 33:949-956.

Wallis and Browse (1999) *Arch. Biochem. Biophys.* 365:307-316. Watts and Browse (1999b) *Arch. Biochem. Biophys.* 362:175-182. Weiss et al. (2003) *Int. J. Med. Microbiol.* 293:95-106.

Whitney et al. (2003) *Planta* 217:983-992.

Winans (1988) *J. Bacteriol.* 170:4047-54.

Wood (2009) *Plant Biotechnol J.* 7:914-24. Wu et al. (2005) *Nat. Biotech.* 23:1013-1017. Yang et al. (2003) *Planta* 216:597-603.

Zank et al. (2002) *Plant J.* 31:255-268. Zank et al. (2005) *WO* 2005/012316

Zhang et al. (2004) *FEBS Lett.* 556:81-85. Zhang et al. (2006) 20:3255-3268.

Zhang et al. (2007a) *FEBS Letters* 581: 315-319. Zhang et al. (2008) *Yeast* 25: 21-27.

Zhou et al. (2007) *Phytochem.* 68:785-796.

Zhou et al. (2008) *Insect Mol Biol* 17: 667-676. Zou et al. (1997) *Plant Cell.* 9:909-23.

YÊU CẦU BẢO HỘ

1. Lipit chiết tách từ cây, chứa các axit béo ở dạng được este hóa, các axit béo này bao gồm axit oleic, axit palmitic, axit béo $\omega 6$ bao gồm axit linoleic (LA), axit béo $\omega 3$ bao gồm axit α -linolenic (ALA) và axit docosahexaenoic (DHA), và tùy ý một hoặc nhiều trong số axit stearidonic (SDA), axit eicosapentaenoic (EPA), axit docosapentaenoic (DPA) và axit eicosatetraenoic (ETA), trong đó mức DHA trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 7% đến 20%, và trong đó mức axit palmitic trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 2 đến 16% và trong đó mức axit myristic (C14:0) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 1%.

2. Lipit theo điểm 1, trong đó lipit này có một hoặc nhiều hoặc tất cả các đặc điểm sau:

- i) mức axit palmitic trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 2% đến 15%,
- ii) mức axit oleic trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 3% đến khoảng 30%,
- iii) mức axit linoleic (LA) trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 4% đến khoảng 35%,
- iv) mức axit α -linolenic (ALA) trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 7% đến khoảng 40%,
- v) mức axit γ -linolenic (GLA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 4%,
- vi) mức axit stearidonic (SDA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 4%,
- vii) mức axit eicosatetraenoic (ETA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 4%,
- viii) mức axit eicosatrienoic (ETrA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 4%,
- ix) mức axit eicosapentaenoic (EPA) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 4%,
- x) mức axit docosapentaenoic (DPA) trong tổng hàm lượng axit béo của lipit

chiết tách là thấp hơn 4%,

xi) mức DHA trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 11% đến 20%,

xii) mức toàn bộ axit béo bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 4% đến khoảng 25%,

xiii) mức toàn bộ axit béo một liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 4% đến khoảng 35%,

xiv) mức toàn bộ axit béo nhiều liên kết chưa bão hòa trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 20% đến khoảng 75%,

xv) mức toàn bộ axit béo $\omega 6$ trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 20%,

xvi) mức axit béo $\omega 6$ mới trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn khoảng 10%,

xvii) mức toàn bộ axit béo $\omega 3$ trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 40% đến khoảng 60%,

xviii) mức axit béo $\omega 3$ mới trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 9% đến khoảng 33%,

xix) tỷ lệ của toàn bộ axit béo $\omega 6$: toàn bộ axit béo $\omega 3$ trong hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 1,0 đến khoảng 3,0,

xx) tỷ lệ của axit béo $\omega 6$ mới: axit béo $\omega 3$ mới trong hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 1,0 đến khoảng 3,0,

xxi) thành phần axit béo của lipit được dựa trên hiệu suất chuyển hóa axit oleic thành DHA ít nhất khoảng 10%,

xxii) thành phần axit béo của lipit được dựa trên hiệu suất chuyển hóa LA thành DHA ít nhất là khoảng 15%,

xxiii) thành phần axit béo của lipit được dựa trên hiệu suất chuyển hóa ALA thành DHA ít nhất là khoảng 17%,

xxiv) toàn bộ axit béo trong lipit chiết tách có C20:1 với lượng ít hơn 1%,

xxv) hàm lượng triaxylglyxerol (TAG) trong lipit ít nhất khoảng 70%,

xxvi) lipit chứa diaxylglyxerol (DAG),

xxvii) lipit chứa axit béo tự do (chưa được este hóa) và/hoặc phospholipit với lượng ít hơn khoảng 10% hoặc gần như không chứa chúng,

xxviii) ít nhất 70% DHA được este hóa ở dạng TAG là ở vị trí *sn*-1 hoặc *sn*-3 trong TAG,

xxix) lipid chứa tri-DHA TAG (TAG 66:18).

3. Lipit theo điểm 1 hoặc 2, trong đó lipit này còn chứa một hoặc nhiều sterol và/hoặc ở dạng dầu, và chứa ít hơn khoảng 10mg sterol/g dầu, ít hơn khoảng 7mg sterol/g dầu, từ khoảng 1,5mg đến khoảng 10mg sterol/g dầu, hoặc từ khoảng 1,5mg đến khoảng 7mg sterol/g dầu.

4. Lipit theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó lipit là dầu *Brassica sp.*, hoặc dầu *Camelina sativa*.

5. Quy trình sản xuất lipit chiết tách từ cây, bao gồm các bước:

i) thu hoạch bộ phận cây chứa lipit, trong đó lipit này chứa các axit béo ở dạng được este hóa, các axit béo này bao gồm axit oleic, axit palmitic, axit béo $\omega 6$ bao gồm axit linoleic (LA) và axit γ -linolenic (GLA), axit béo $\omega 3$ chứa axit α -linolenic (ALA), axit stearidonic (SDA), axit docosapentaenoic (DPA) và axit docosahexaenoic (DHA), và tùy ý một hoặc nhiều axit eicosapentaenoic (EPA) và axit eicosatetraenoic (ETA), trong đó mức DHA trong tổng hàm lượng axit béo của lipit chiết tách trong bộ phận cây nằm trong khoảng từ khoảng 7% đến 20%, và trong đó mức axit palmitic trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 2% đến 16% và trong đó mức axit myristic (C14:0) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 1%, và

(ii) chiết tách lipit từ bộ phận cây này,

trong đó mức DHA trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 7% đến 20% và trong đó mức axit palmitic trong tổng hàm lượng axit béo của lipit chiết tách nằm trong khoảng từ khoảng 2% đến 16% và trong đó mức axit myristic (C14:0) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 1%.

6. Quy trình theo điểm 5, trong đó lipit chiết tách có một hoặc nhiều đặc điểm được xác định trong các điểm từ điểm 2 đến 4.

7. Quy trình theo điểm 5 hoặc điểm 6, trong đó bộ phận cây bao gồm polynucleotit ngoại sinh mã hóa $\Delta 12$ -desaturaza, $\omega 3$ -desaturaza hoặc $\Delta 15$ -desaturaza, $\Delta 6$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 6$ -elongaza và $\Delta 5$ -elongaza, và trong đó mỗi polynucleotit được liên kết điều khiển với một hoặc nhiều gen khởi đầu có khả năng điều khiển quá trình biểu hiện polynucleotit này trong tế bào của bộ phận cây.

8. Quy trình theo điểm 7, trong đó bộ phận cây có một hoặc nhiều hoặc tất cả các đặc điểm sau:

- i) $\Delta 12$ -desaturaza chuyển hóa axit oleic thành axit linoleic trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 60%,
- ii) $\omega 3$ -desaturaza chuyển hóa axit béo $\omega 6$ thành axit béo $\omega 3$ trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 65%,
- iii) $\Delta 6$ -desaturaza chuyển hóa ALA thành SDA trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 30%,
- iv) $\Delta 6$ -desaturaza chuyển hóa axit linoleic thành axit γ -linolenic trong một hoặc nhiều tế bào của cây với hiệu suất thấp hơn khoảng 5%,
- v) $\Delta 6$ -elongaza chuyển hóa SDA thành ETA trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 60%,
- vi) $\Delta 5$ -desaturaza chuyển hóa ETA thành EPA trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 60%,
- vii) $\Delta 5$ -elongaza chuyển hóa EPA thành DPA trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 80%,
- viii) $\Delta 4$ -desaturaza chuyển hóa DPA thành DHA trong một hoặc nhiều tế bào của cây với hiệu suất ít nhất là khoảng 80%,
- ix) hiệu suất chuyển hóa axit oleic thành DHA trong một hoặc nhiều tế bào của bộ phận cây ít nhất là khoảng 10%,
- x) hiệu suất chuyển hóa LA thành DHA trong một hoặc nhiều tế bào của bộ phận cây ít nhất là khoảng 15%,
- xi) hiệu suất chuyển hóa ALA thành DHA trong một hoặc nhiều tế bào của bộ phận cây ít nhất là khoảng 17%,
- xii) một hoặc nhiều tế bào của bộ phận cây chứa nhiều axit béo $\omega 3$ hơn các tế

bào tương ứng thiếu polynucleotit ngoại sinh với lượng ít nhất khoảng 15%,

- xiii) $\Delta 6$ -desaturaza ưu tiên khử bão hòa axit α -linolenic (ALA) hơn so với axit linoleic (LA),
- xiv) $\Delta 6$ -elongaza cũng có hoạt tính $\Delta 9$ -elongaza,
- xv) $\Delta 12$ -desaturaza cũng có hoạt tính $\Delta 15$ -desaturaza,
- xvi) $\Delta 6$ -desaturaza cũng có hoạt tính $\Delta 8$ -desaturaza,
- xvii) $\omega 3$ -desaturaza cũng có hoạt tính $\Delta 15$ -desaturaza trên LA,
- xviii) $\omega 3$ -desaturaza khử bão hòa cả LA và/hoặc GLA,
- xix) $\omega 3$ -desaturaza ưu tiên khử bão hòa GLA hơn so với LA,
- xx) mức DHA trong bộ phận cây được dựa trên hiệu suất chuyển hóa axit oleic thành DHA trong bộ phận cây ít nhất là khoảng 10%,
- xxi) mức DHA trong bộ phận cây được dựa trên hiệu suất chuyển hóa LA thành DHA trong bộ phận cây ít nhất là khoảng 15%,
- xxii) mức DHA trong bộ phận cây được dựa trên hiệu suất chuyển hóa ALA thành DHA trong bộ phận cây ít nhất là khoảng 17%,
- xxiii) một hoặc nhiều hoặc tất cả desaturaza có hoạt tính trên cơ chất axyl-CoA cao hơn trên cơ chất axyl-PC tương ứng,
- xxiv) $\Delta 6$ -desaturaza có hoạt tính $\Delta 6$ -desaturaza trên cơ chất axit béo ALA cao hơn cơ chất axit béo LA,
- xxv) $\Delta 6$ -desaturaza có hoạt tính $\Delta 6$ -desaturaza trên cơ chất axit béo ALA-CoA cao hơn trên cơ chất axit béo ALA được gắn với vị trí *sn*-2 của PC,
- xxvi) $\Delta 6$ -desaturaza có hoạt tính $\Delta 6$ -desaturaza trên cơ chất ALA cao hơn ít nhất khoảng 2 lần so với LA,
- xxvii) $\Delta 6$ -desaturaza có hoạt tính trên cơ chất axit béo ALA-CoA cao hơn trên cơ chất axit béo ALA được gắn với vị trí *sn*-2 của PC,
- xxviii) $\Delta 6$ -desaturaza có hoạt tính $\Delta 6$ -desaturaza trên cơ chất axit béo ALA-CoA cao hơn ít nhất khoảng 5 lần so với trên cơ chất axit béo ALA được gắn với vị trí *sn*-2 của PC,
- xxix) desaturaza là desaturaza đầu trước,
- xxx) $\Delta 6$ -desaturaza không có hoạt tính $\Delta 5$ -desaturaza phát hiện được trên ETA.

9. Quy trình theo điểm bất kỳ trong số các điểm từ 5 đến 8, trong đó quy trình này có một hoặc nhiều dấu hiệu sau:

i) polynucleotit ngoại sinh được liên kết đồng hóa trị trong phân tử ADN, tốt hơn là phân tử T-ADN, được tích hợp vào hệ gen của các tế bào của bộ phận cây và tốt hơn là trong đó số lượng phân tử ADN như vậy được tích hợp vào hệ gen của các tế bào của bộ phận cây không nhiều hơn một, hai hoặc ba, hoặc là hai hoặc ba,

(ii) tổng hàm lượng dầu của bộ phận cây chứa polynucleotit ngoại sinh ít nhất là khoảng 40% tổng hàm lượng dầu của bộ phận cây tương ứng thiếu polynucleotit ngoại sinh,

(iii) lipit ở dạng dầu, tốt hơn là dầu hạt từ hạt dầu, và trong đó ít nhất khoảng 90% trọng lượng lipit là triaxylglyxerol,

(iv) quy trình này còn bao gồm bước xử lý lipit để làm tăng mức DHA dưới dạng phần trăm của tổng hàm lượng axit béo tốt hơn là bước xử lý này là chuyển hóa este.

10. Tế bào chủ chứa polynucleotit ngoại sinh mã hóa một trong các bộ enzym sau:

i) $\Delta 12$ -desaturaza, $\omega 3$ -desaturaza hoặc $\Delta 15$ -desaturaza, $\Delta 6$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 6$ -elongaza và $\Delta 5$ -elongaza,

ii) $\omega 3$ -desaturaza, $\Delta 6$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 6$ -elongaza và $\Delta 5$ -elongaza,

iii) $\Delta 15$ -desaturaza, $\Delta 6$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 6$ -elongaza và $\Delta 5$ -elongaza,

iv) $\Delta 12$ -desaturaza, $\Delta 6$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 6$ -elongaza và $\Delta 5$ -elongaza,

v) $\omega 3$ -desaturaza, $\Delta 8$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 9$ -elongaza và $\Delta 5$ -elongaza,

vi) $\Delta 15$ -desaturaza, $\Delta 8$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 9$ -elongaza và $\Delta 5$ -elongaza,

vii) $\Delta 12$ -desaturaza, $\Delta 8$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 9$ -elongaza và $\Delta 5$ -elongaza, hoặc

viii) $\Delta 12$ -desaturaza, $\omega 3$ -desaturaza hoặc $\Delta 15$ -desaturaza, $\Delta 8$ -desaturaza, $\Delta 5$ -

desaturaza, $\Delta 4$ -desaturaza, $\Delta 9$ -elongaza và $\Delta 5$ -elongaza,

và trong đó mỗi polynucleotit được liên kết điều khiển với một hoặc nhiều gen khởi đầu có khả năng điều khiển quá trình biểu hiện của polynucleotit nêu trên trong tế bào, và trong đó tế bào chứa lipid như được xác định trong điểm bất kỳ trong số các điểm từ 1 đến 4.

11. Cây biến đổi gen chứa tế bào theo điểm 10.

12. Cây hạt dầu chứa:

- a) lipid trong hạt của nó, lipid này chứa các axit béo ở dạng được este hóa, và
- b) polynucleotit ngoại sinh mã hóa một trong các bộ enzym sau;

(i) $\Delta 12$ -desaturaza, $\omega 3$ -desaturaza nằm và/hoặc $\Delta 15$ -desaturaza nằm, $\Delta 6$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 6$ -elongaza và $\Delta 5$ -elongaza, hoặc

(ii) $\Delta 12$ -desaturaza, $\omega 3$ -desaturaza nằm và/hoặc $\Delta 15$ -desaturaza nằm, $\Delta 8$ -desaturaza, $\Delta 5$ -desaturaza, $\Delta 4$ -desaturaza, $\Delta 9$ -elongaza và $\Delta 5$ -elongaza,

trong đó mỗi polynucleotit được liên kết điều khiển với một hoặc nhiều gen khởi đầu đặc hiệu với hạt mà có khả năng điều khiển quá trình biểu hiện polynucleotit này trong quá trình phát triển hạt của cây, trong đó các axit béo bao gồm axit oleic, axit palmitic, axit béo $\omega 6$ bao gồm axit linoleic (LA) và axit γ -linolenic (GLA), axit béo $\omega 3$ bao gồm axit α -linolenic (ALA), axit stearidonic (SDA), axit docosapentaenoic (DPA) và axit docosahexaenoic (DHA), và tùy ý axit eicosapentaenoic (EPA) và/hoặc axit eicosatetraenoic (ETA), và trong đó mức DHA trong tổng hàm lượng axit béo của lipid nằm trong khoảng từ khoảng 7% đến 20%, và trong đó mức axit palmitic trong tổng hàm lượng axit béo của của lipid chiết tách nằm trong khoảng từ khoảng 2% đến 16% và trong đó mức axit myristic (C14:0) trong tổng hàm lượng axit béo của lipid chiết tách là thấp hơn 1%.

13. Cây theo điểm 12, trong đó cây này có một hoặc nhiều đặc điểm sau:

- (i) cây là cây cải dầu, *Glycine max*, *Camelina sativa* hoặc *Arabidopsis thaliana*,
- (ii) một hoặc nhiều desaturaza có khả năng sử dụng cơ chất axyl-CoA,
- (ii) hạt chín, thu hoạch được của cây có hàm lượng DHA ít nhất là khoảng 28mg trong mỗi gam hạt.

14. Bộ phận cây có một hoặc nhiều đặc điểm sau:

- (i) chứa lipit như được xác định trong điểm bất kỳ trong số các điểm từ 1 đến 4,
- (ii) có nguồn gốc từ cây theo điểm bất kỳ trong số các điểm từ 11 đến 13, hoặc
- (ii) có thể được sử dụng trong quy trình theo điểm bất kỳ trong số các điểm từ 5 đến 9.

15. Bộ phận của cây theo điểm 14, trong đó bộ phận này là hạt.

16. Phương pháp tạo ra hạt, trong đó phương pháp này bao gồm các bước:

- a) trồng cây, trong đó cây này sinh ra bộ phận như được xác định trong điểm 14 hoặc điểm 15, hoặc cây này là cây theo điểm bất kỳ trong số các điểm từ 11 đến 13,
- b) thu hoạch hạt từ cây hoặc các cây, và
- c) tùy ý, chiết tách lipit từ hạt, tốt hơn là để tạo ra dầu với sản lượng DHA tổng cộng ít nhất là 60kg DHA/héc ta.

17. Bột hạt thu được từ hạt theo điểm 15.

18. Phương pháp sản xuất thức ăn gia súc, trong đó phương pháp này bao gồm các bước trộn một hoặc nhiều lipit hoặc dầu theo điểm bất kỳ trong số các điểm từ 1 đến 4, tế bào theo điểm 10, cây biến đổi gen theo điểm 11, cây hạt dầu theo điểm 12 hoặc điểm 13, bộ phận cây theo điểm 14 hoặc điểm 15, bột hạt theo điểm 17 với ít nhất một thành phần thức ăn khác.

19. Quy trình sản xuất etyl este của axit béo nhiều liên kết chưa bão hòa, trong đó quy trình này bao gồm các bước chuyển hóa este triaxylglyxerol trong lipit chiết tách từ cây, trong đó lipit chiết tách từ cây chứa các axit béo ở dạng được este hóa, các axit béo này bao gồm axit oleic, axit palmitic, axit béo $\omega 6$ bao gồm axit linoleic (LA), axit béo $\omega 3$ bao gồm axit α -linolenic (ALA), và axit docosahexaenoic (DHA), và tùy ý một hoặc nhiều trong số axit stearidonic (SDA), axit eicosapentaenoic (EPA), axit docosapentaenoic (DPA) và axit eicosatetraenoic (ETA), trong đó mức DHA trong tổng hàm lượng axit béo của lipit chiết tách là khoảng từ khoảng 7% đến 20%, và trong đó

mức axit palmitic trong tổng hàm lượng axit béo của của lipit chiết tách nằm trong khoảng từ khoảng 2% đến 16% và trong đó mức axit myristic (C14:0) trong tổng hàm lượng axit béo của lipit chiết tách là thấp hơn 1%, bằng cách đó tạo ra etyl este.

1/21

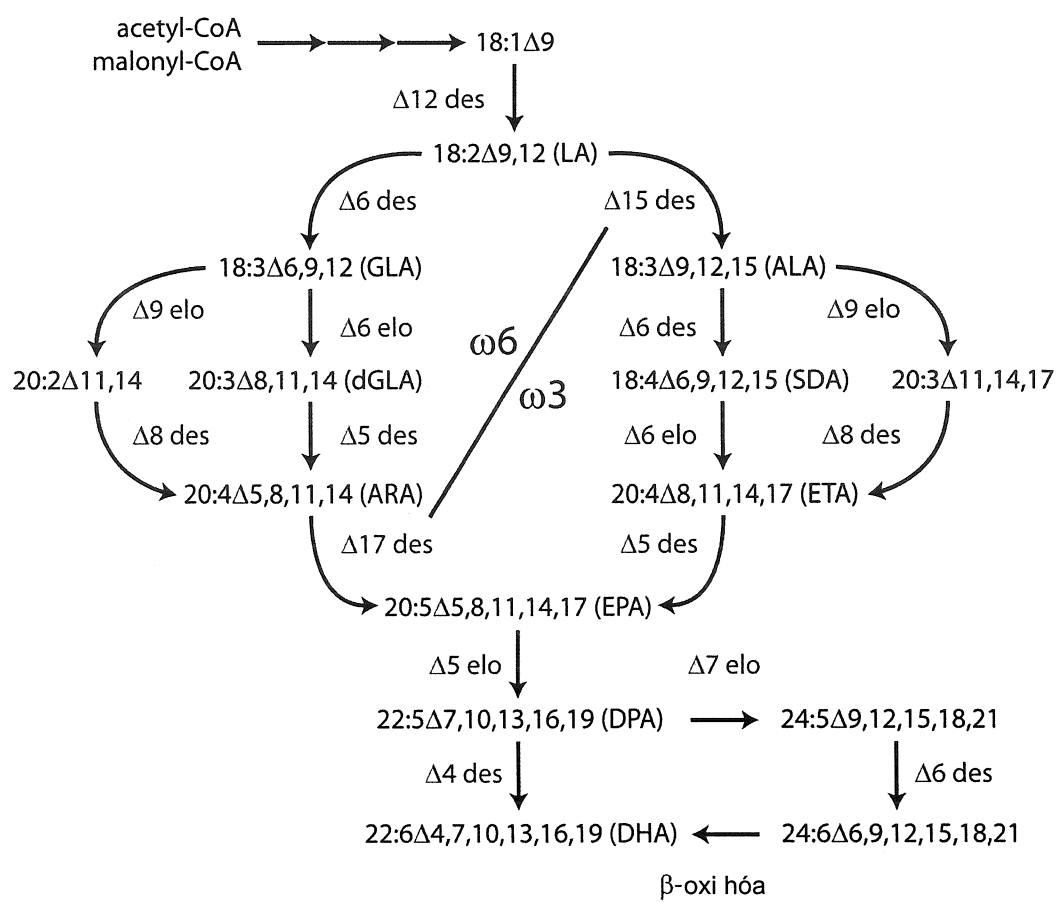


Fig. 1

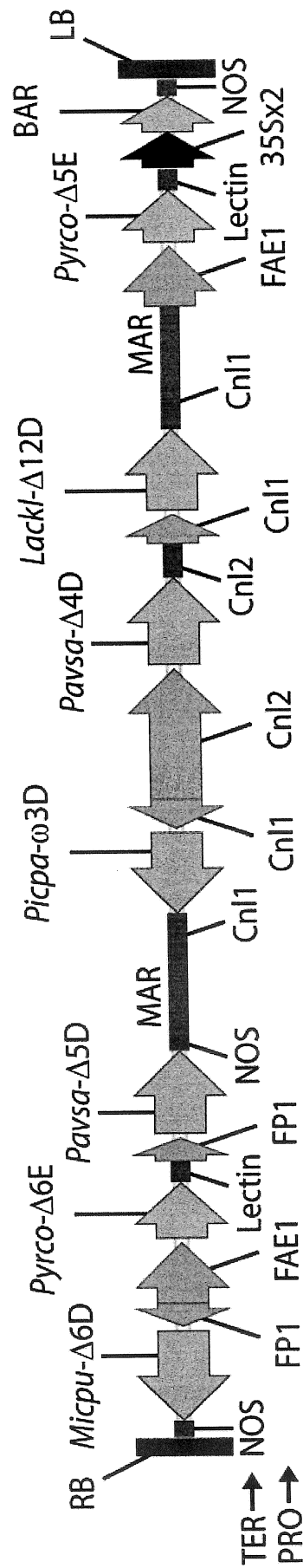
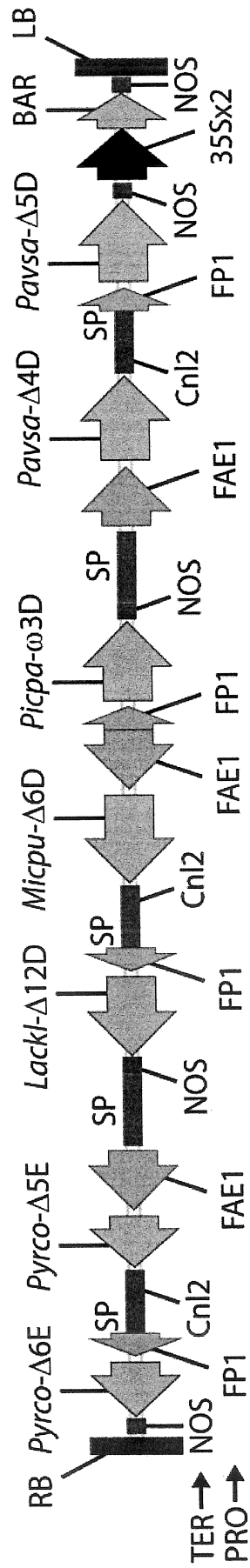


Fig. 2



4/21

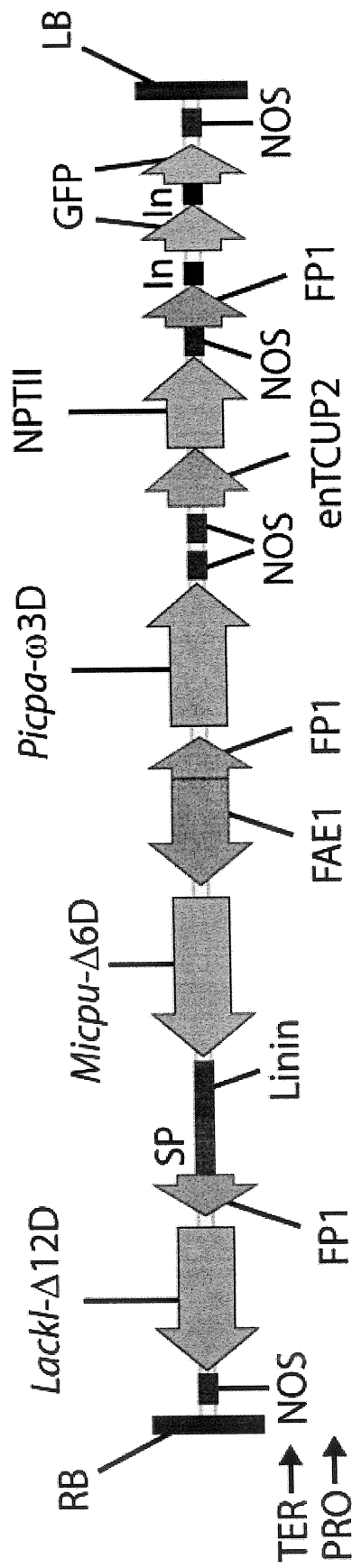


Fig. 4

5/21

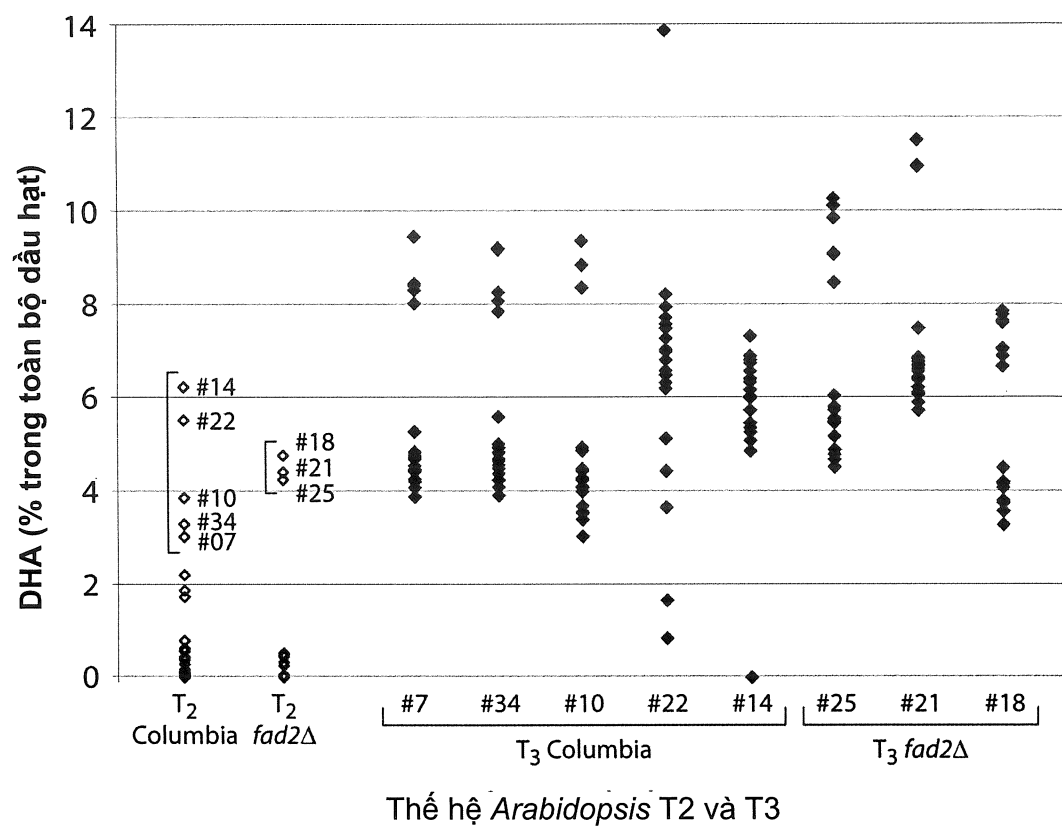
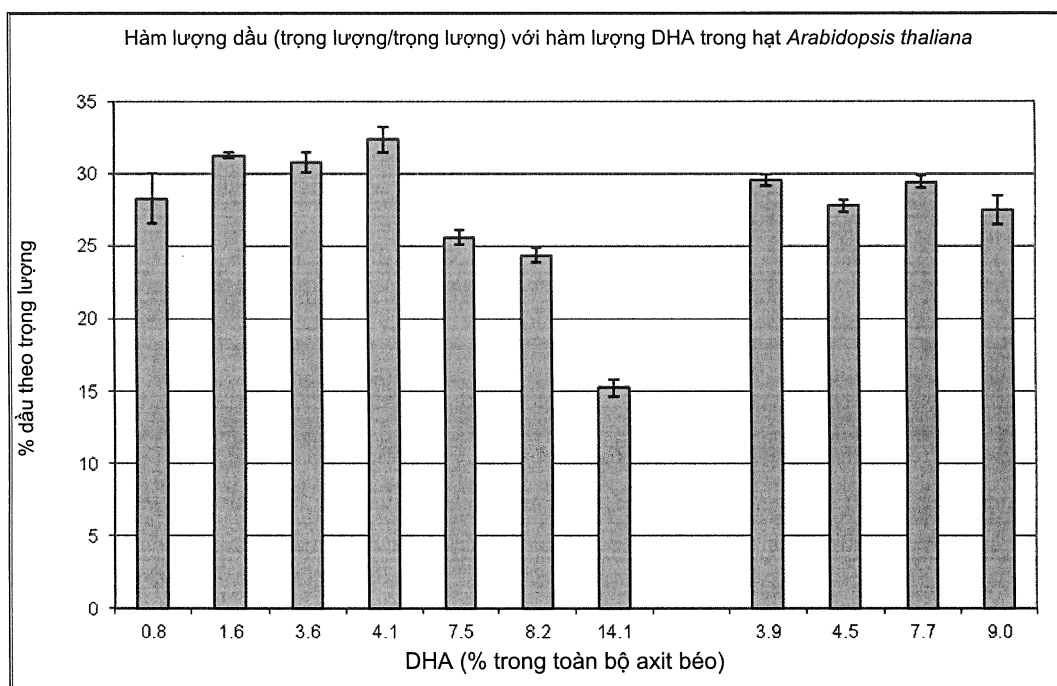
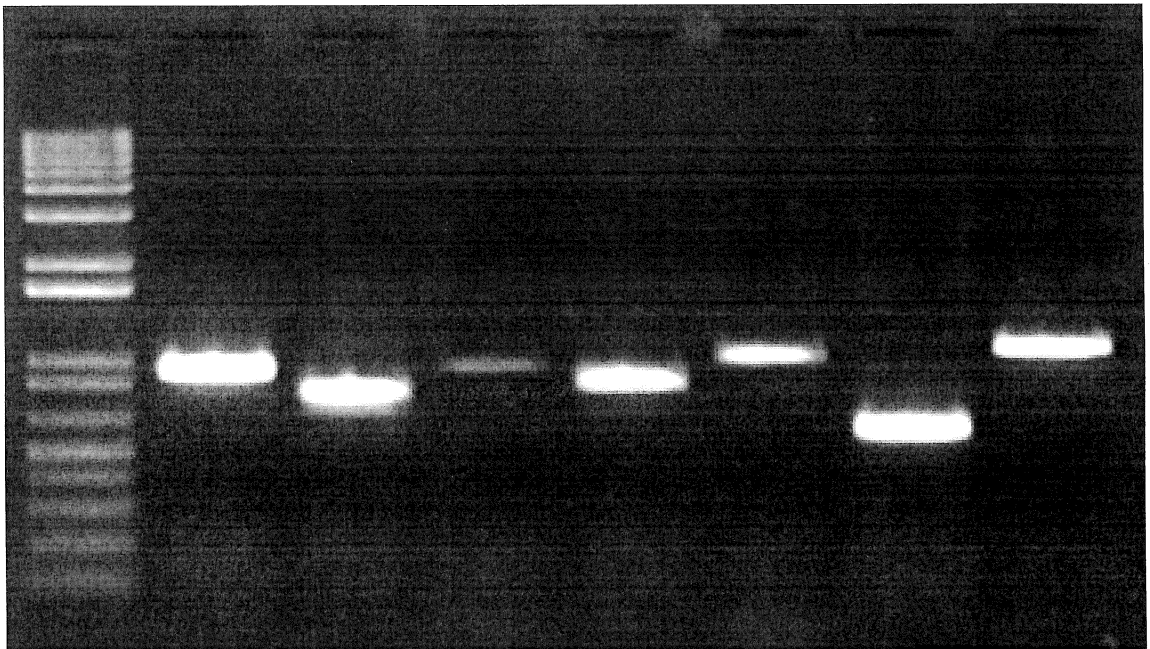


Fig. 5

6/21

**Fig. 6**

7/21

**Fig. 7**

8/21

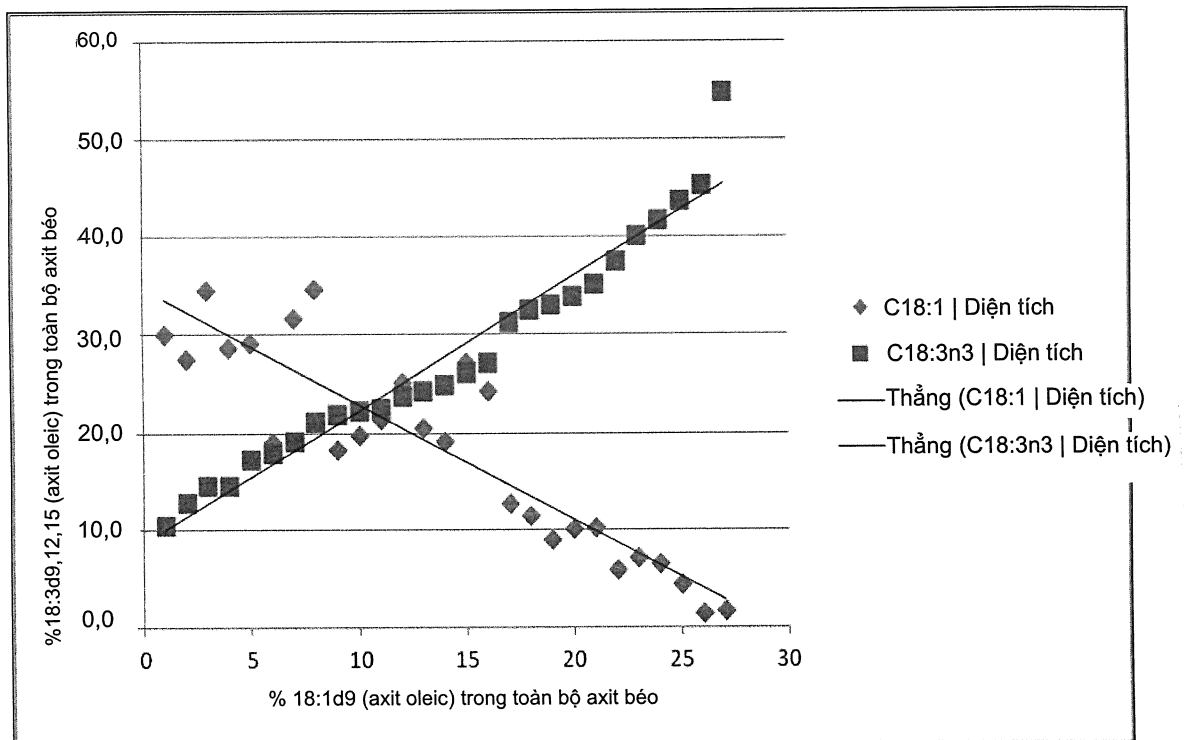
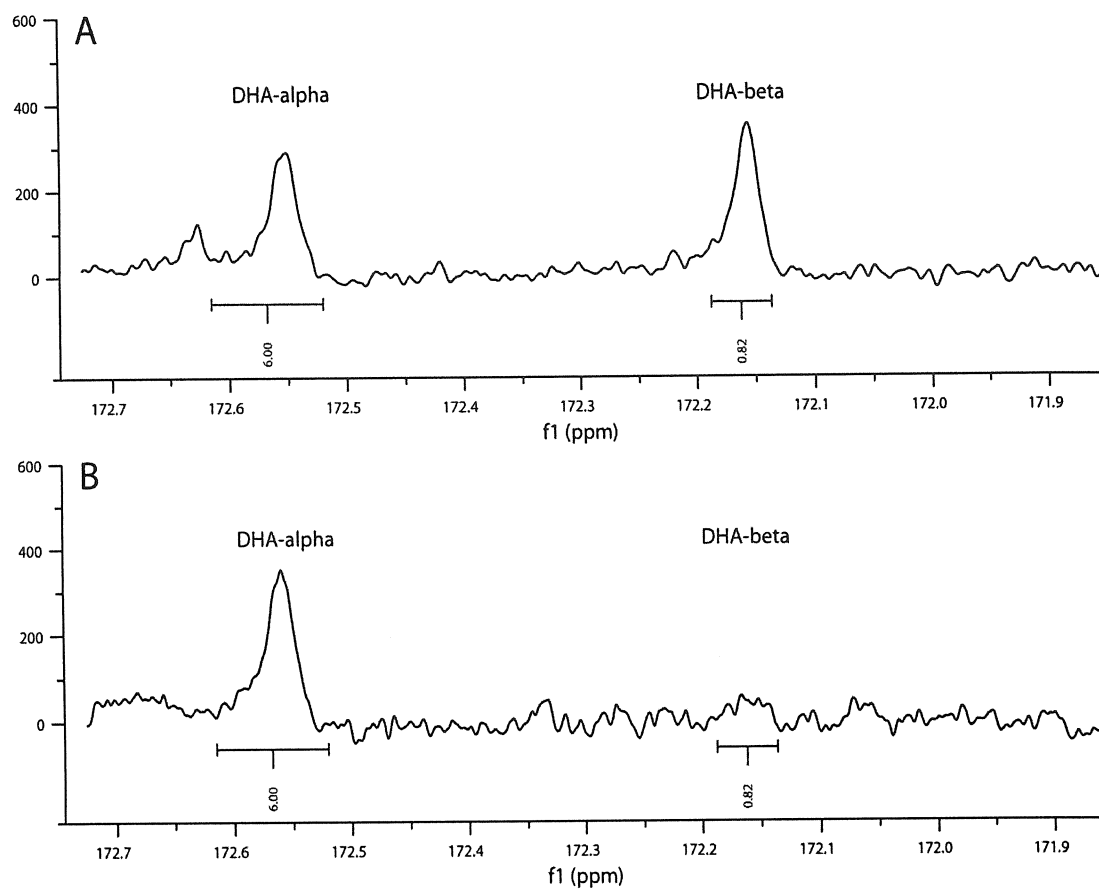
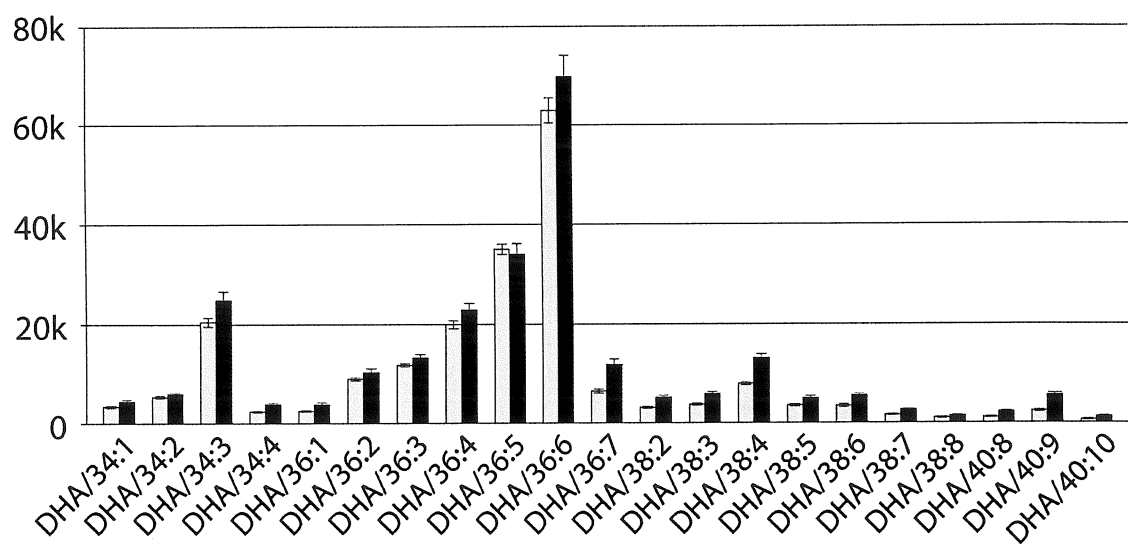


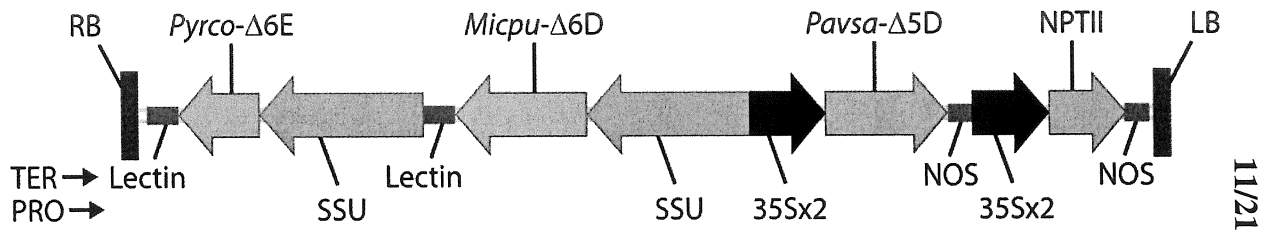
Fig. 8

9/21

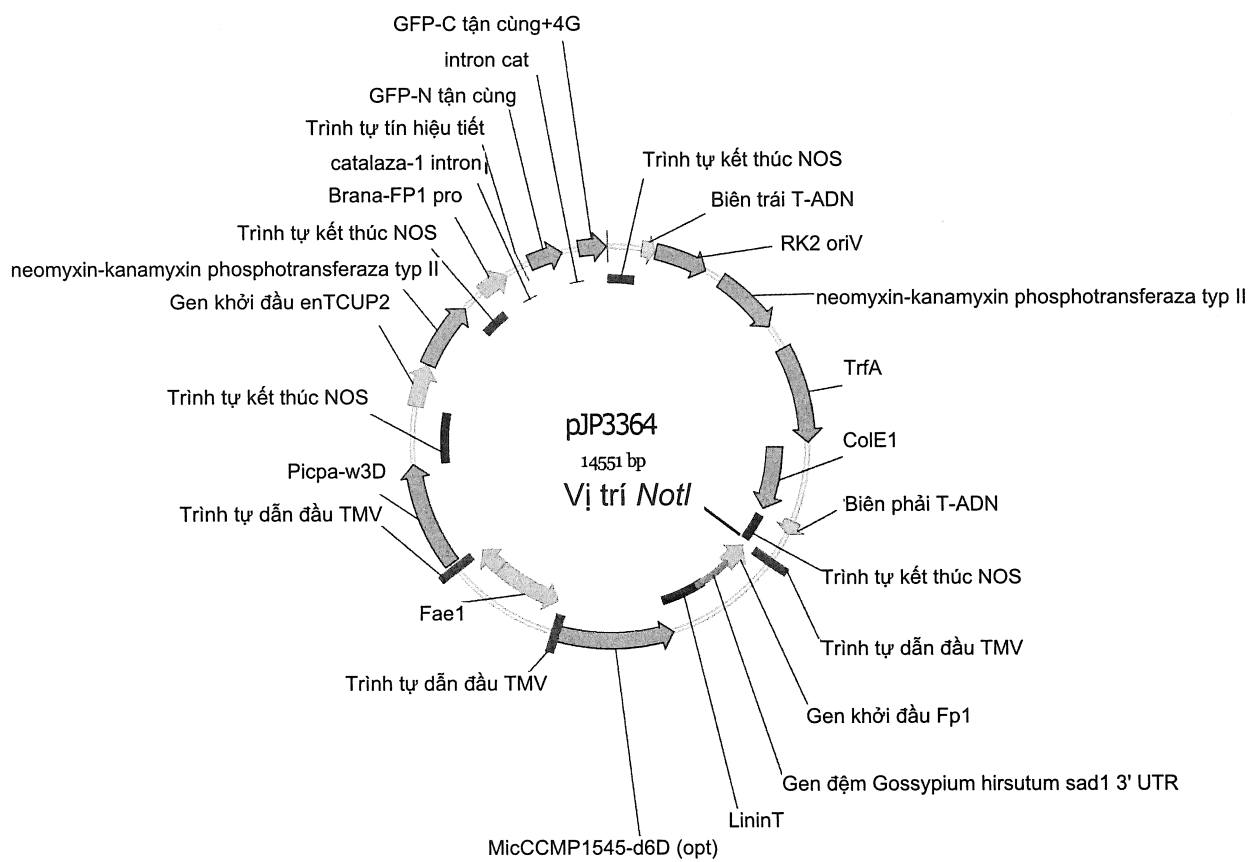
**Fig. 9**

10/21

**Fig. 10**

**Fig. 11**

12/21

**Fig. 12**

13/21

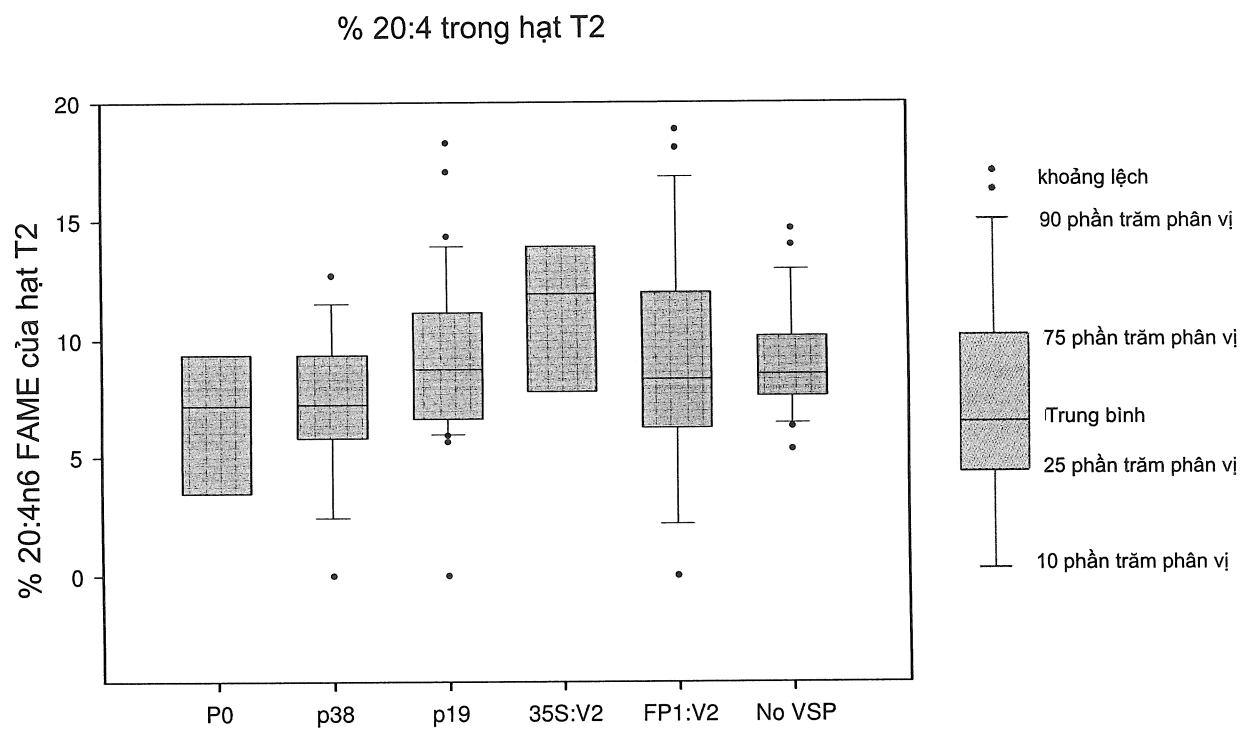
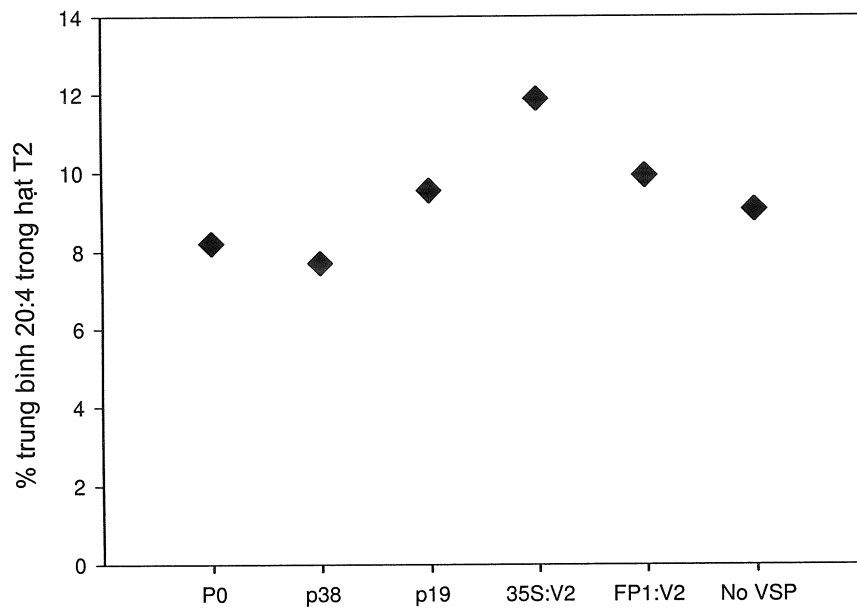


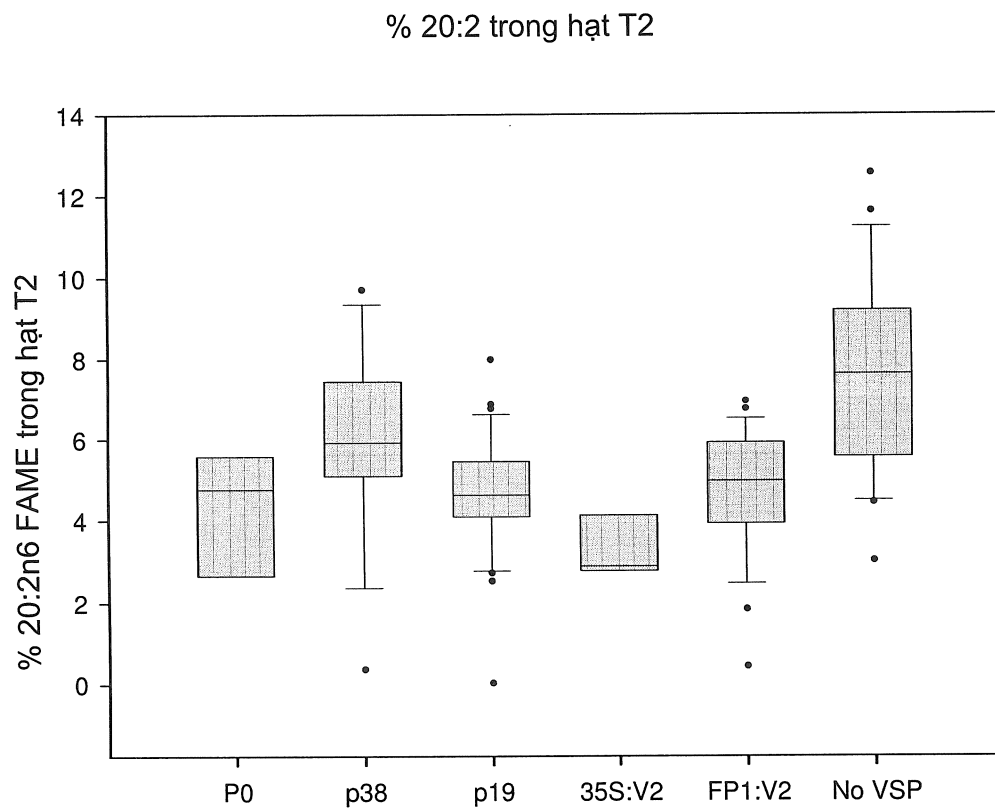
Fig. 13

14/21

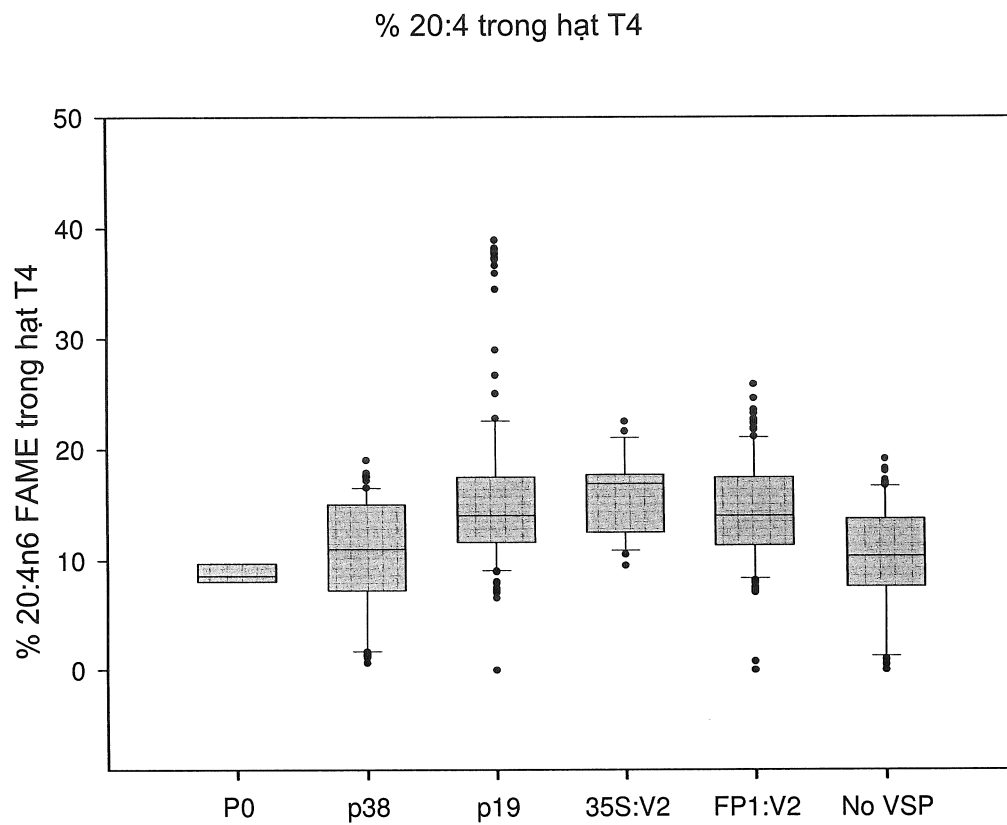
trung bình 20:4 trong hạt T2

**Fig. 14**

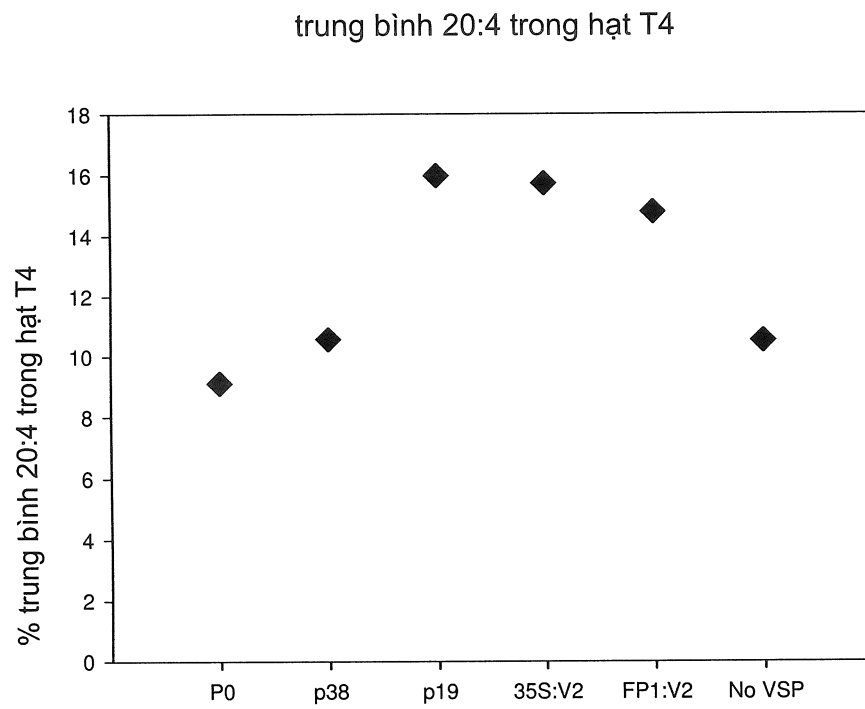
15/21

**Fig. 15**

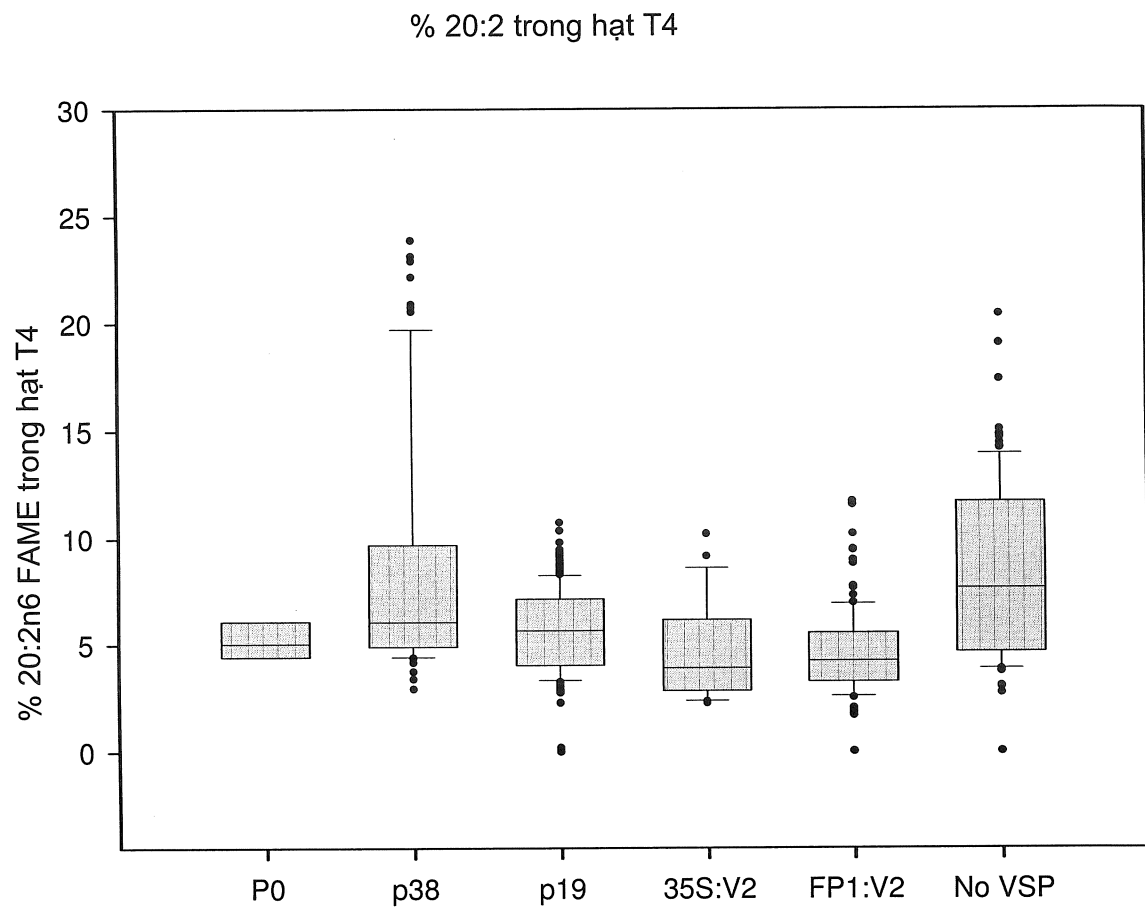
16/21

**Fig. 16**

17/21

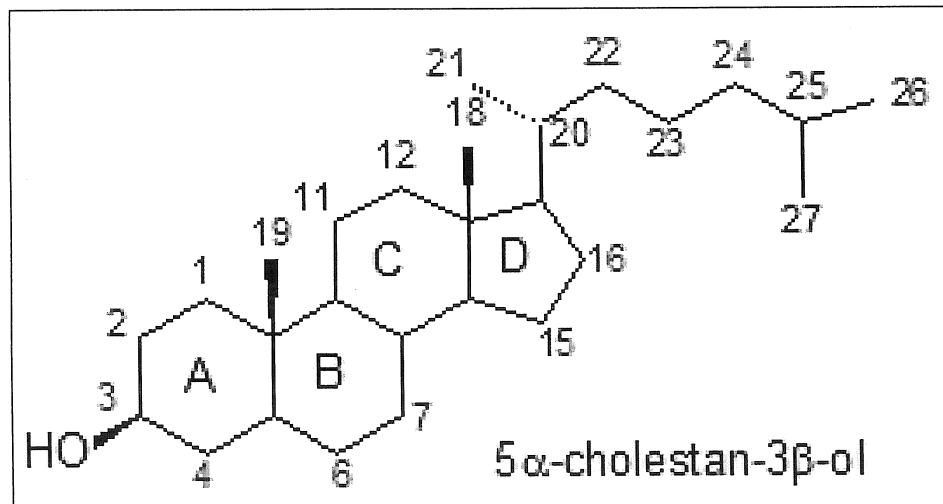
**Fig. 17**

18/21

**Fig. 18**

19/21

A)



B)

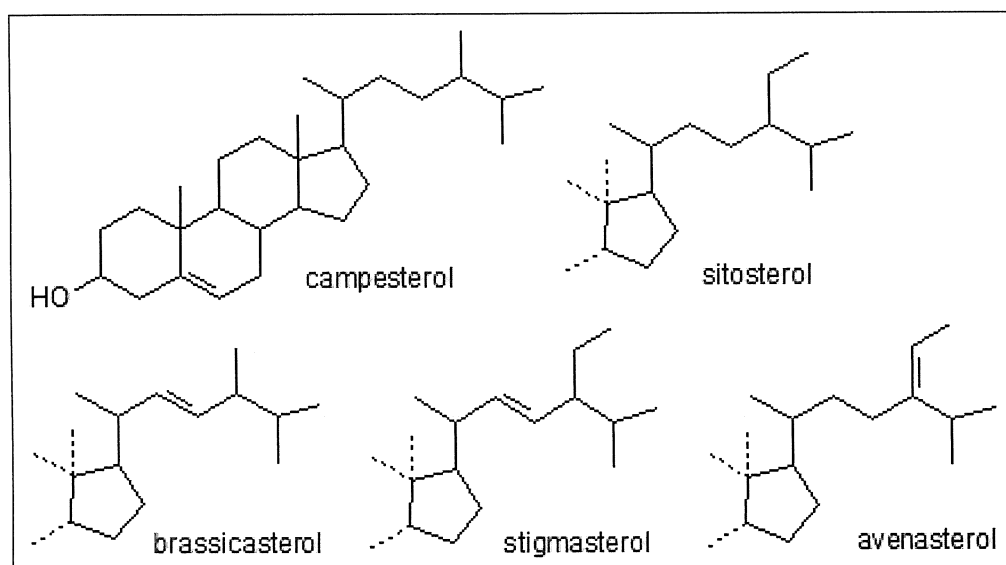
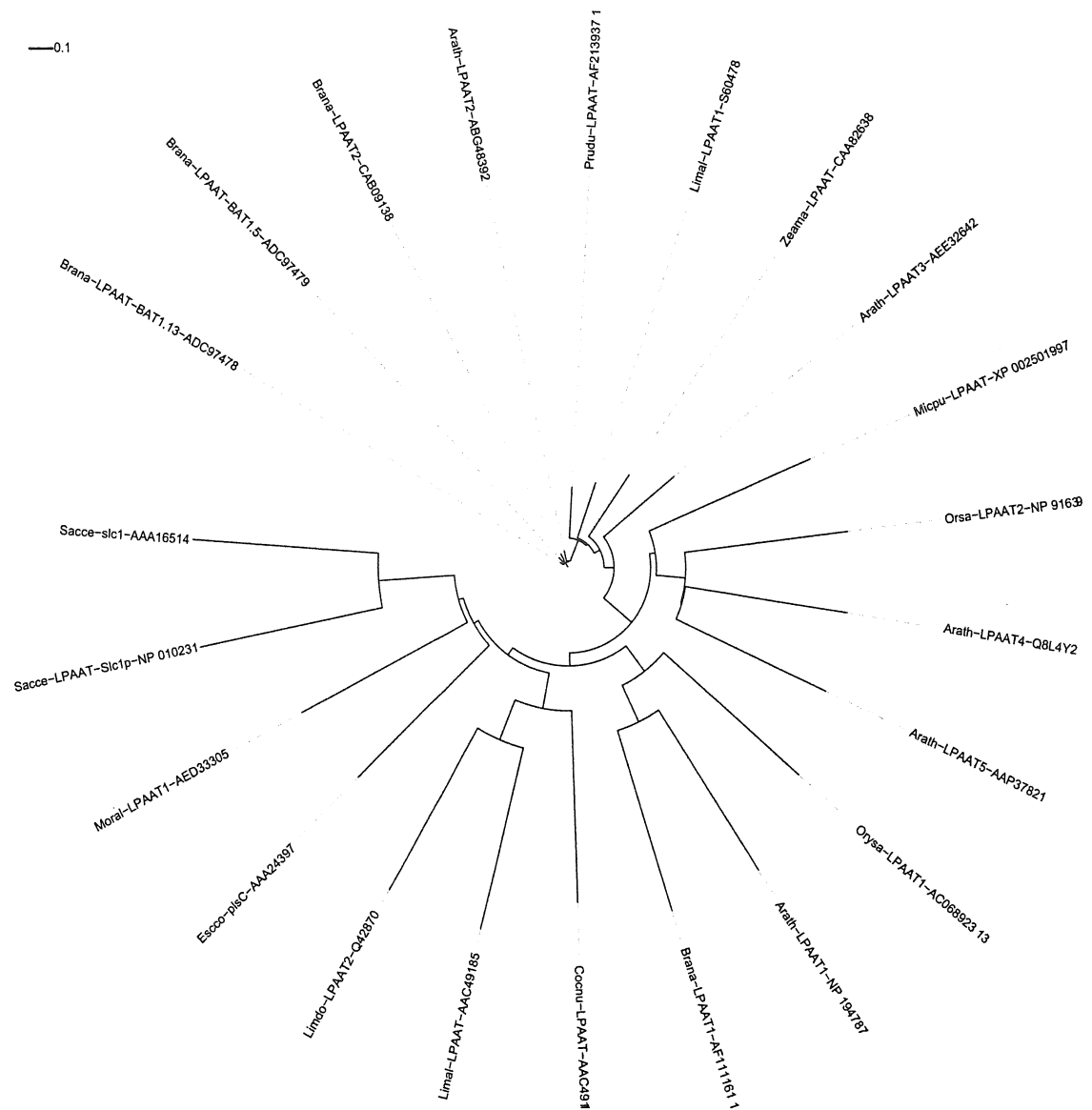
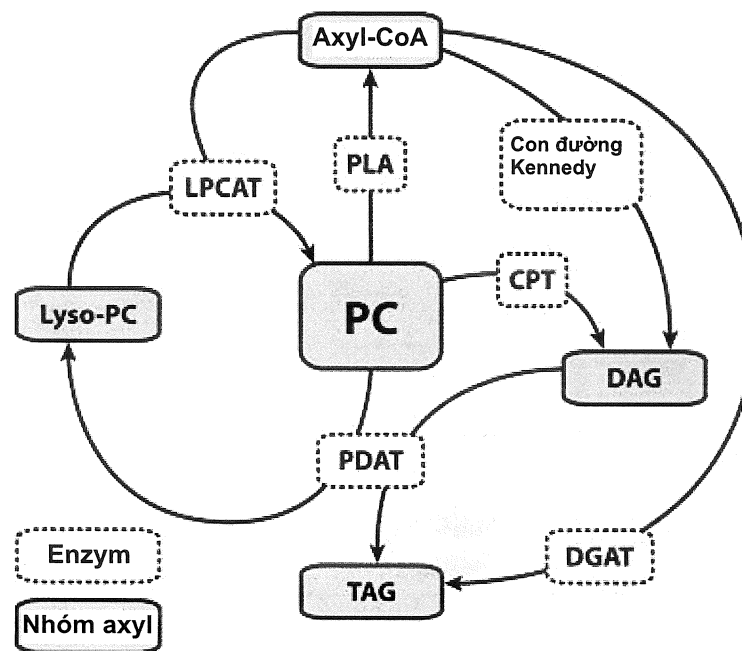


Fig. 19

20/21

**Fig. 20**

**Fig. 21**

1/143

DANH SÁCH TRÌNH TỰ

- <110> Nuseed Nutritional Australia Pty Ltd
Grain Research and Development Corporation
Commonwealth Scientific and Industrial Research
Organisation
- <120> Lipit chiết tách từ cây chứa các axit béo ở dạng được este hóa và quy trình sản xuất
lipit này
- <130> 513594
- <150> US 61/660,392
<151> 2012-06-15
- <150> US 61/663,344
<151> 2012-06-22
- <150> US 61/697,676
<151> 2012-09-06
- <150> US 61/782,680
<151> 2013-03-14
- <160> 72
- <170> PatentIn phiên bản 3.5
<210> 1
<211> 21527
<212> ADN
<213> Trình tự nhân tạo
- <220>
<223> Trình tự nucleotit pJP3416-GA7.
- <400> 1
tcctgtggtt ggcatgcaca tacaaatgga cgaacggata aacctttttca cgccctttta 60
aatatccgat tattctaata aacgctcttt tctcttaggt ttacccgcca atatatcctg 120
tcaaacactg atagttttaa ctgaaggcgg gaaacgacaa tctgctagtg gatctcccag 180
tcacgacgtt gtaaaacggg cgccccgcgg aaagcttgcg gccgcccgat ctagtaacat 240
agatgacacc gcgcgcgata atttatccta gtttgcgcg tatattttgt tttctatcgc 300
gtattaaatg tataattgcg ggactctaata cataaaaaacc catctcataa ataacgtcat 360
gcattacatg ttaattatta cgtgcttaac gtaattcaac agaaattata tgataatcat 420
cgcaagaccg gcaacaggat tcaatcttaa gaaactttat tgccaaatgt ttgaacgac 480
ggcgcgcctc attagtgagc cttctcagcc tttccgttaa cgtagtagtg ctgtcccacc 540
ttatcaaggt tagagaaagt agccttccaa gcaccgtagt aagagagcac cttgtagttg 600
agtccccact tcttagcgaa aggaacgaat cttctgctaa cctcaggctg tctgaattga 660
ggcatatcag ggaagaggtg gtggataacc tgacagttaa ggtatcccat aagccagttc 720
acgtatcctc tagaaggatc gatatcaacg gtgtgatcaa cagcgtagtt aaccaagaa 780
aggtgcttat cagatggaac aacagggagg tgagtatgag aagtagagaa gtgagcgaaa 840
aggtacatgt aagcgatcca gtttccgaaa gtgaaccacc agtaagcaac aggccaagag 900
tatccagtag caagcttgat aacagcgggt ctaacaacat gagaaacgag catccaagaa 960

2/143

gcctcttcgt agttcttctt acggagaact tgtctagggg ggagaacgta gatccagaaa	1020
gcttgaacaa gaagtccaga ggtaacagga acgaaagtcc aagcttgaag tctagcccaa	1080
gctctagaga atcctctagg tctgttatcc tcaacagcag tgttgaagaa agccacagca	1140
ggagtgggtat caagatccat atcgtgtcta accttttgag gggtagcatg gtgcttgta	1200
tgcattctggg tccacatctc accagaagta gaaagtccga atccacaagt catagcctga	1260
agtctcttgt ccacgtaaac agatccggta agagagttat gtccaccctc atgttgaacc	1320
catccacatc tagctccgaa gaaagcaccg taaacaacag aagcaatgat agggatatcca	1380
gcgtacataa gagcagttcc aagagcgaat gtagcaagaa gctcgagaag tctgtaagcc	1440
acatgggtga tagaaggctt gaagaatcca tctctctcaa gctcagcacg ccatctagcg	1500
aaatcctcaa gcataggagc atcctcagac tcagatctct tgatctcagc aggtctagaa	1560
ggcaaagctc taagcatctt ccaagccttg agagaacgca tgtggaattc tttgaaagcc	1620
tcagtagcat cagcaccagt gttagcaagc atgtagaaga tcacagatcc accagggtgc	1680
ttgaagttag tcacatcgta ctcaacgtcc tcaactctaa cccatctagt ctcgaaagta	1740
gcagcaagct catgaggctc aagagtctta agatcaacag gagcagtaga agcatcctta	1800
gcatcaagag cctcagcaga agatttagac ctggtaagtg gagatctagg agaagatctt	1860
ccatcagtct taggagggca catgggtatg taattgtaaa tgtaattgta atgttgtttg	1920
ttgtttgttg ttgttggtaa ttgttgtaaa agatcctcgt gtatgttttt aatcttgttt	1980
gtatcgatga gttttggttt gagtaaagag tgaagcggat gagttaattt ataggctata	2040
aaggagattt gcatggcgat cacgtgtaat aatgcatgca cgcattgat tgtatgtgtg	2100
tgctgtgaga gagaagctct taggtgtttg aaggagtgta caagtggcga agaaaaacaa	2160
ttctccgcgg ctgcatgcta tgtgtaacgt gtagctaattg ttctggcatg gcatcttatg	2220
aacgattctt tttaaaaaca aggtaaaaac ttaacttcat aaaattaaaa aaaaaaacgt	2280
ttactaagtt ggtttaaaag gggatgagac tagtagattg gttggttggt ttccatgtac	2340
cagaaggctt accctattag ttgaaagttg aaactttgtt ccctactcaa ttcctagttg	2400
tgtaaatgta tgtatatgta atgtgtataa aacgtagtac ttaaatgact aggagtggtt	2460
cttgagaccg atgagagatg ggagcagaac taaagatgat gacataatta agaacgaatt	2520
tgaaaggctc ttaggtttga atcctattcg agaattgttt tgtcaaagat agtggcgatt	2580
ttgaaccaa gaaaacattt aaaaaatcag tatccgggta cgttcatgca aatagaaagt	2640
ggcttaggat ctgattgtaa ttttagactt aaagagtctc ttaagattca atcctggctg	2700
tgtacaaaac tacaaataat atattttaga ctatttggcc ttaactaaac ttccactcat	2760
tatttactga ggtagagaa tagacttgcg aataaacaca ttcccagaa atactcatga	2820
tcccataatt agtcagaggg tatgccaatc agatctaaga acacacattc cctcaaattt	2880
taatgcacat gtaatcatag ttagcacaa ttcaaaaata atgtagtatt aaagacagaa	2940
atttgtagac ttttttttg cgttaaaaga agactaagtt tatacgtaca ttttatttta	3000
agtggaac cgaaattttc catcgaaata tatgaattta gtatatatat ttctgcaatg	3060

tactatTTTTg ctatTTTTggc aactttcagt ggactactac tttattacaa tgtgtatgga	3120
tgcatgagtt tgagtataca catgtctaaa tgcatgcttt gtaaaacgta acggaccaca	3180
aaagaggatc catacaaata catctcatag cttcctccat tattttccga cacaaacaga	3240
gcattttaca acaattacca acaacaacaa acaacaacaa acattacaat tacatttaca	3300
attaccatac catggaattc gcccagcctc ttgttgctat ggctcaagag caatacgctg	3360
ctatcgatgc tgttggtgct cctgctatct tctctgctac tgattctatc ggatggggac	3420
ttaagcctat ctcttctgct actaaggact tgcctcttgt tgagtctcct acacctctca	3480
tcctttcttt gcttgcttac ttcgctatcg ttggatctgg actcgtttac agaaaggttt	3540
tccttagaac cgtgaaggga caagatccat tccttttgaa ggctcttatg cttgctcaca	3600
acgtgttcct tatcggactt tctctttaca tgtgcctcaa gcttgtgtac gaggcttacg	3660
ttaacaagta ctctttctgg ggaaacgctt acaaccctgc tcaaactgag atggctaagg	3720
ttatctggat cttctacgtg agcaagatct acgagttcat ggataccttc atcatgctcc	3780
tcaagggaaa tgtaaccag gttagcttcc ttcacgttta ccatcacgga tctatctctg	3840
gaatctggtg gatgattact tacgctgctc ctggtggtga tgcttacttc tctgctgctc	3900
ttaactcttg ggttcacgtg tgtatgtaca cctactatTTt tatggctgcc gtgcttccta	3960
aggacgagaa aactaagaga aagtacctct ggtggggaag ataccttact caaatgcaga	4020
tgttccagtt cttcatgaac cttctccagg ctgtttacct tctctactct tcatctcctt	4080
accctaagtt tatcgctcag ctctcgtgg tgtacatggt tactcttctc atgcttttcg	4140
gaaacttcta ctacatgaag caccacgcta gcaagtgatg aggcgcgccg ggccgccgcc	4200
atgtgacaga tcgaaggaag aaagtgtaat aagacgactc tcactactcg atcgctagtg	4260
attgtcattg ttatatataa taatgttatc tttcacaact tatcgtaatg catgtgaaac	4320
tataacacat taatcctact tgtcatatga taacactctc cccattttaa actcttgtca	4380
atttaaagat ataagattct ttaaattgatt aaaaaaata tattataaat tcaatcactc	4440
ctactaataa attattaatt attatttatt gattaaaaaa atacttatac taatttagtc	4500
tgaatagaat aattagattc tagtctcatc ccctttttaa ccaacttagt aaacgttttt	4560
ttttttaatt ttatgaagtt aagtttttac cttgttttta aaaagaatcg ttcataagat	4620
gccatgccag aacattagct acacgttaca catagcatgc agccgcggag aattgttttt	4680
cttcgccact tgtcactccc ttcaaacacc taagagcttc tctctcacag cacacacata	4740
caatcacatg cgtgcatgca ttattacacg tgatcgccat gcaaactctc tttatagcct	4800
ataaattaac tcatccgctt cactctttac tcaaaccaa actcatcgat acaaacaaga	4860
ttaaaaacat acacgaggat cttttacaac aattaccaac aacaacaaac aacaacaac	4920
attacaatta catttacaat taccatacca tgcctccaag ggactcttac tcttatgctg	4980
ctctccttc tgctcaactt cacgaagttg atactcctca agagcacgac aagaaagagc	5040
ttgttatcgg agatagggct tacgatgtta ccaacttcgt taagagacac cctggtggaa	5100

4/143

agatcattgc ttaccaagtt ggaactgatg ctaccgatgc ttacaagcag ttccatgtta	5160
gatctgctaa ggctgacaag atgcttaagt ctcttccttc tcgtcctgtt cacaagggat	5220
actctccaag aagggtgat cttatcgctg atttccaaga gttcaccaag caacttgagg	5280
ctgagggaaat gttcgagcct tctcttcctc atgttgctta cagacttgct gaggttatcg	5340
ctatgcatgt tgctggtgct gctcttatct ggcatggata cactttcgct ggaatcgcta	5400
tgcttgaggt tgttcagggg agatgtggat ggcttatgca tgaggggtgga cattactctc	5460
tcactggaaa cattgctttc gacagagcta tccaagttgc ttgttacgga cttggatgtg	5520
gaatgtctgg tgcttggtgg cgtaaccagc ataacaagca ccatgctact cctcaaaagc	5580
ttcagcacga tgttgatctt gatacccttc ctctcgttgc tttccatgag agaatcgctg	5640
ctaaggttaa gtctcctgct atgaaggctt ggctttctat gcaagctaag cttttcgctc	5700
ctgttaccac tcttcttggt gctcttggtt ggcatgctta ccttcacctc agacacatgc	5760
tcaggactaa gcactacgat gagcttgcta tgctcggaat cagatacggg cttgttggtt	5820
accttgctgc taactacggt gctggatacg ttctcgcttg ttaccttctt tacgttcagc	5880
ttggagctat gtacatcttc tgcaacttcg ctgtttctca tactcacctc cctgttggtg	5940
agcctaacga gcatgctact tgggttgagt acgctgctaa ccacactact aactgttctc	6000
catcttggtg gtgtgattgg tggatgtctt accttaacta ccagatcgag caccacctt	6060
accttctat gcctcaattc agacacccta agatcgctcc tagagttaag cagcttttcg	6120
agaagcacgg acttcactac gatgttagag gatacttcga ggctatggct gatactttcg	6180
ctaaccttga taacgttgcc catgctcctg agaagaaaat gcagtaatga gatcggtcaa	6240
acatttgga ataaagtttc ttaagattga atcctgttg cggctcttgcg atgattatca	6300
tataatttct gttgaattac gtttaagcac taataattaa catgtaatgc atgacgttat	6360
ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa	6420
acaaaatata gcgcgcaaac taggataaat tatcgcgcg cgtgtcatct atgttactag	6480
atcggtcgat taaaaatccc aattatatat ggtctaattt agtttggtat tgagtaaaac	6540
aaattcgaac caaaccaaaa tataaatata tagtttttat atatatgcct ttaagacttt	6600
ttatagaatt ttcttttaaa aatatctaga aatatttgcg actcttctgg catgtaatat	6660
ttcgttaaat atgaagtgt ccatttttat taactttaaa taattgggtg tacgatcact	6720
ttcttatcaa gtgttactaa aatgcgtcaa tctctttgtt cttccatatt catatgtcaa	6780
aatctatcaa aattcttata tatctttttc gaatttgaag tgaaatttcg ataatttaaa	6840
attaaataga acatatcatt atttaggtat catattgatt ttatactta attactaaat	6900
ttggttaact ttgaaagtgt acatcaacga aaaattagtc aaacgactaa aataaataaa	6960
tatcatgtgt tattaagaaa attctcctat aagaatattt taatagatca tatgtttgta	7020
aaaaaaatta atttttacta acacatatat ttacttatca aaaatttgac aaagtaagat	7080
taaaataata ttcatctaac aaaaaaaaaa ccagaaaatg ctgaaaaccc ggcaaaaccg	7140
aaccaatcca aaccgatata gttgggttgg tttgattttg atataaacg aaccaactcg	7200

gtccatttgc acccctaatac ataatagctt taatatttca agatattatt aagttaacgt	7260
tgtcaatatac ctggaaattt tgcaaaatga atcaagccta tatggctgta atatgaattt	7320
aaaagcagct cgatgtggtg gtaatatgta atttacttga ttctaaaaaa atatcccaag	7380
tattaataat ttctgctagg aagaaggtta gctacgattt acagcaaagc cagaatacaa	7440
agaaccataa agtgattgaa gctcgaaata tacgaaggaa caaatatttt taaaaaata	7500
cgcaatgact tggaacaaaa gaaagtgata ttttttttgt tcttaaacia gcatccccctc	7560
taaagaatgg cagttttcct ttgcatgtaa ctattatgct cccttcgtta caaaaatttt	7620
ggactactat tgggaacttc ttctgaaaat agtgatagaa cccacacgag catgtgcttt	7680
ccattttaatt ttaaaaacca agaaacatac atacataaca ttccatcagc ctctctctct	7740
ttttattacg gttaatgact taaaacacat cttattatcc catccttaac acctagcagt	7800
gtctttatac gatctcatcg atcaccactt caaaaccatg cagactgctg ctgcccctgg	7860
agctggcatc ggctaggctg ggtgccgcac tgtcccgaa ggtccctagc gacttgttta	7920
gattgatggg accacctctc aacttcctgc tgctgtccct gctgctggat gtcctgcctc	7980
atctggccga ttgcacgctc cagtcccctg catgtgcact cgctcctcaa ttgcttaaga	8040
tcacgcagc agctatcgaa gtgctggctc tgttgccctc ctccacggcc ttggttgtag	8100
tagtagctgc cgccgccctt ctggactttt tcccacagga accgccgaat aattcgatag	8160
aaccacacga gcatgtgctt tcatttattt taaaaaccaa gaaacataca taacatttca	8220
tcagcctctc tctctctctc tctctctctc tctctctctc tctctctctc tctctcttta	8280
ttacagctgt tacactaact taaaacacat tcattctatt attattatta ttatccatcc	8340
ttaacaccta gcagtgtctt tgtacgatct cataatcgat cacccttca tcaggatatcc	8400
ttaggcttca ctccaacgtt gttgcagtta cggaacatgt acacaccatc atggttctca	8460
acgaactggc aagatctcca agttttccaa aggctaacc acatgttctc atcgggtgtgt	8520
ctgtagtgct ctcccataac tttcttgatg cactcggtag cttctctagc atggtagaat	8580
gggatccttg aaacgtagtg atggagcaca tgagtctcga tgatgtcatg gaagatgatt	8640
ccgaggattc cgaactctct atcgatagta gcagcagcac ccttagcgaa agtccactct	8700
tgagcatcgt aatgaggcat agaagaatcg gtgtgctgaa ggaaggtaac gaaaacaagc	8760
cagtggttaa caaggatcca aggacagaac catgtgatga aagtaggcca gaatccgaaa	8820
accttgtaag cggtgtaaac agaagtgagg gtagcaagga ttccaagatc agaaagaacg	8880
atgtaccagt agtccttctt atcgaaaaca gggctagaag gccagtagtg agacttgaag	8940
aacttagaaa caccagggtg aggttgctca gtagcgtag tagcaaggta aagagaaagt	9000
cctccaagct gttggaacia gagagcgaaa acagagtaga taggagtttc ctacgcgata	9060
tcgtgaaggc tggttaacttg gtgcttctct ttgaattcct cggcggtgta aggaacgaaa	9120
accatatctc tggcatgtg tccagtagcc ttatggtgct tagcatgaga gaacttccag	9180
ctgaagtaag gaaccataac aagagagtgg agaaccatc caacggtatc gttaacccat	9240

6/143

ccgtagttag	agaaagcaga	atgtccacac	tcatgtccaa	ggatccagat	tccgaatccg	9300
aaacaagaga	tagagaacac	gtaagcagac	caagcagcga	atctaaggaa	ttcgttaggg	9360
agaagagggg	tgtaggtaag	tccaacgtaa	gcatagcag	agatagccac	gatatctctc	9420
accacgtaag	acatagactt	cacgagagat	ctctcgtaac	agtgccttagg	gatagcgtca	9480
aggatatcct	tgatggtgta	atctggcacc	ttgaaaacgt	ttccgaagggt	atcgatagcg	9540
gtctttttgct	gcttgaaaga	tgcaacgttt	ccagaacgcc	taacgggtctt	agtagatccc	9600
tcaaggatct	cagatccaga	cacggtaacc	ttagacatgg	tatggtaatt	gtaaatgtaa	9660
ttgtaatggt	gtttgttggt	tgttgttggt	ggtaattggt	gtaaaatttt	tggtggtgat	9720
tggttcttta	agggtgtgaga	gtgagttgtg	agttgtgtgg	tgggtttggt	gagattgggg	9780
atgggtgggtt	tatatagtgg	agactgagga	atggggtcgt	gagtgttaac	tttgcatggg	9840
ctacacgtgg	gttcttttgg	gcttacacgt	agtattattc	atgcaaatgc	agccaataca	9900
tatacggtat	tttaataatg	tgtgggaata	caatatgccg	agtattttac	taattttggc	9960
aatgacaagt	gtacatttgg	attatcttac	ttggcctctc	ttgctttaat	ttggattatt	10020
tttattctct	taccttggcc	gttcatattc	acatccctaa	aggcaagaca	gaattgaatg	10080
gtggccaaaa	attaaaacga	tggatatgac	ctacatagtg	taggatcaat	taacgtcgaa	10140
ggaaaatact	gattctctca	agcatacgga	caagggtaaa	taacatagtc	accagaacat	10200
aataaacaaa	aagtgcagaa	gcaagactaa	aaaaattagc	tatggacatt	cagggtcata	10260
ttggaaacat	cattatccta	gtcttgtgac	catccttcct	cctgctctag	ttgagaggcc	10320
ttgggactaa	cgagaggta	gttgggatag	cagatcctta	tcctggacta	gcctttctgg	10380
tgtttcagag	tcttcgtgcc	gccgtctaca	tctatctcca	ttaggtctga	agatgactct	10440
tcacaccaac	gacgtttaag	gtctctatcc	tactcctagc	ttgcaatacc	tggcttgcaa	10500
tacctggagc	atcgtgcacg	atgattggat	actgtggagg	aggagtgttt	gctgatttag	10560
agctccccgt	tgggtgattt	gacttcgatt	tcagtttagg	cttggtgaaa	tttttcaggt	10620
tccattgtga	agcctttaga	gcttgagctt	ccttccatgt	taatgccttg	atcgaatact	10680
cctagagaaa	agggaagtcg	atctctgagt	attgaaatcg	aagtgcacat	ttttttcaa	10740
cgtgtccaat	caatccacaa	acaaagcaga	agacaggtaa	tctttcatac	ttatactgac	10800
aagtaatagt	cttaccgtca	tgcataataa	cgtctcgttc	cttcaagagg	ggttttccga	10860
catccataac	gacccgaagc	ctcatgaaag	cattagggaa	gaacttttgg	ttcttcttgt	10920
catggccttt	ataggtgtca	gccgagctcg	ccaattcccc	tccgactggc	tccgcaaaat	10980
attcgaacgg	caagttatgg	acttgcaacc	ataactccac	ggtattgagc	aggacctatt	11040
gtgaagactc	atctcatgga	gcttcagaat	gtggttgtca	gcaaaccaat	gaccgaaatc	11100
catcacatga	cggacgtcca	gtgggtgagc	gaaacgaaac	aggaagcgcc	tatctttcag	11160
agtcgtgagc	tccacaccgg	attccggcaa	ctacgtgttg	ggcaggcttc	gccgtattag	11220
agatatgttg	aggcagaccc	atctgtgcca	ctcgtacaat	tacgagagtt	gttttttttg	11280
tgattttcct	agtttctcgt	tgatggtgag	ctcatattct	acatcgtatg	gtctctcaac	11340

gtcgtttcct gtcattctgat atcccgtcat ttgcatccac gtgcgccgcc tcccgtgcc	11400
agtccttagg tgtcatgcac gccaaattgg tgggtggtgcg ggctgccctg tgcttcttac	11460
cgatgggtgg aggttgagtt tgggggtctc cgcggcgatg gtagtgggtt gacggtttgg	11520
tgtgggttga cggcattgat caatttactt cttgcttcaa attctttggc agaaaacaat	11580
tcattagatt agaactggaa accagagtga tgagacggat taagtcagat tccaacagag	11640
ttacatctct taagaaataa tgtaaccctt ttagacttta tatatttgca attaaaaaaa	11700
taatttaact ttttagacttt atatatagtt ttaataacta agtttaacca ctctattatt	11760
tatatcgaaa ctatttgtat gtctcccctc taaataaact tgggtatttg tttacagaac	11820
ctataatcaa ataataata ctcaactgaa gtttgtgcag ttaattgaag ggattaacgg	11880
ccaaaatgca ctagtattat caaccgaata gattcacact agatggccat ttccatcaat	11940
atcatcgccg ttcttcttct gtccacatat cccctctgaa acttgagaga cacctgcact	12000
tcattgtcct tattacgtgt taaaaaatga aacctatgca tccatgcaaa ctgaagaatg	12060
gcgcaagaac cttccccctc catttcttat gtggcgacca tccatttcac catctccgc	12120
tataaaacac ccccatcact tcacctagaa catcatcact acttgcttat ccatcaaaa	12180
gatacccact tttaacaaca ttaccaaca caacaacaa caaacaacat tacaattaca	12240
tttacaatta ccataccatg ccacctagcg ctgctaagca aatgggagct tctactggtg	12300
ttcatgctgg tgttactgac tcttctgctt tcaccagaaa ggatggttgc gatagacctg	12360
atctcaccat cgttggagat tctgtttacg atgctaaggc tttcagatct gagcatcctg	12420
gtggtgctca tttcgtttct ttgttcggag gaagagatgc tactgaggct ttcatggaat	12480
accatagaag ggcttggcct aagtctagaa tgtctagatt ccacgttgga tctcttgctt	12540
ctactgagga acctgttgct gctgatgagg gataccttca actttgtgct aggatcgcta	12600
agatggtgcc ttctgtttct tctggattcg ctctgcttc ttactgggtt aaggctggac	12660
ttatccttgg atctgctatc gctcttgagg cttacatgct ttacgctgga aagagacttc	12720
tcccttctat cgttcttggga tggcttttcg ctcttatcgg tcttaacatc cagcatgatg	12780
ctaaccatgg tgctttgtct aagtctgctt ctgttaacct tgctcttggga ctttgtcagg	12840
attggatcgg aggatctatg atcctttggc ttcaagagca tgttgttatg caccacctcc	12900
acactaacga tgttgataag gatcctgatc aaaaggctca cggtgctctt agactcaagc	12960
ctactgatgc ttggtcacct atgcattggc ttcagcatct ttaccttttg cctggtgaga	13020
ctatgtacgc tttcaagctt ttgttcctcg acatctctga gcttgttatg tggcggtggg	13080
agggtagacc tatctctaag cttgctggat acctctttat gccttctttg cttctcaagc	13140
ttaccttctg ggctagattc gttgctttgc ctctttacct tgctccttct gttcatactg	13200
ctgtgtgtat cgctgctact gttatgactg gatctttcta cctcgcttct ttcttcttca	13260
tctcccacaa cttcgagggg gttgcttctg ttggacctga tggatctatc acttctatga	13320
ctagaggtgc tagcttcctt aagagacaag ctgagacttc ttctaacgtt ggaggacctc	13380

ttcttgctac tcttaacggt ggactcaact accaaattga gcatcacttg ttccctagag	13440
ttcaccatgg attctaccct agacttgctc ctcttgtaa ggctgagctt gaggctagag	13500
gaatcgagta caagcactac cctactatct ggtctaacct tgcttctacc ctgagacata	13560
tgtacgctct tggaagaagg cctagatcta aggctgagta atgacaagct tatgtgacgt	13620
gaaataataa cggtaaaata tatgtaataa taataataat aaagccacaa agtgagaatg	13680
aggggaaggg gaaatgtgta atgagccagt agccggtggt gctaattttg tatcgtattg	13740
tcaataaatc atgaattttg tggtttttat gtgttttttt aaatcatgaa ttttaaattt	13800
tataaaataa tctccaatcg gaagaacaac attccatctc catgcatgga tgtttcttta	13860
cccaaataa gttcttgaga ggatgaagca tcaccgaaca gttctgcaac tatccctcaa	13920
aagcttttaa atgaacaaca aggaacagag caacgttcca aagatcccaa acgaaacata	13980
ttatctatac taatactata ttattaatta ctactgccg gaatcacaat ccctgaatga	14040
ttcctattaa ctacaagcct tgttggcggc ggagaagtga tcggcgcggc gagaagcagc	14100
ggactcggag acgaggcctt ggaagatctg agtcgaacgg gcagaatcag tattttcctt	14160
cgacgttaat tgatcctaca ctatgtaggt catatccatc gttttaattt ttggccacca	14220
ttcaattctg tcttgccctt agggatgtga atatgaacgg ccaaggtaag agaataaaaa	14280
taatccaaat taaagcaaga gaggccaagt aagataatcc aaatgtacac ttgtcattgc	14340
caaaattagt aaaatactcg gcatattgta ttccacaca ttattaaaat accgtatatg	14400
tattggctgc atttgcatga ataatactac gtgtaagccc aaaagaaccc acgtgtagcc	14460
catgcaaagt taacactcac gacccattc ctgagtctcc actatataaa cccaccatcc	14520
ccaatctcac caaaccacc acacaactca caactcactc tcacacctta aagaaccaat	14580
caccaccaa aattttacaa caattacca caacaacaaa caacaaacaa cattacaatt	14640
acatttacia ttaccatacc atgagcgctg ttaccgttac tggatctgat cctaagaaca	14700
gaggatcttc tagcaacacc gagcaagagg ttccaaaagt tgctatcgat accaacggaa	14760
acgtgttttc tgttcttgat ttcaccatca aggacatcct tggagctatc cctcatgagt	14820
gttacgagag aagattggct acctctctct actacgtgtt cagagatatc ttctgcatgc	14880
ttaccaccgg ataccttacc cataagatcc ttaccctct cctcatctct tacacctta	14940
acagcatcat caagttcact ttctgggccc ttacactta cgttcaagga cttttcgaa	15000
ccggaatctg ggttctcgct catgagtgtg gacatcaagc tttctctgat tacggaatcg	15060
tgaacgattt cgttggatgg acccttact cttaccttat ggttccttac ttcagctgga	15120
agtactctca tggaaagcac cataaggcta ctggacacat gaccagagat atggttttcg	15180
ttcctgccac caaagaggaa ttcaagaagt ctaggaactt cttcggtaac ctgctgagt	15240
actctgagga ttctccactt agaaccctt acgagcttct tgttcaacaa cttggaggat	15300
ggatcgctta cctcttcgtt aacgttacag gacaacctta ccctgatgtt ctttcttgga	15360
aatggaacca cttctggctt acctctccac ttttcgagca aagagatgct ctctacatct	15420
tcctttctga tcttggaatc ctcaccagg gaatcgttct tactctttgg tacaagaaat	15480

9/143

tcggaggatg gtcccttttc atcaactggg tcgttcctta catctgggtt aaccactggc	15540
tcgttttcat cacattcctt cagcacactg atcctactat gcctcattac aacgctgagg	15600
aatggacttt cgctaagggt gctgctgcta ctatcgatag aaagttcgga ttcacggac	15660
ctcacatctt ccatgatatc atcgagactc atgtgcttca ccactactgt tctaggatcc	15720
cattctacaa cgctagacct gcttctgagg ctatcaagaa agttatggga aagcactaca	15780
gggtctagcga cgagaacatg tggaagtcac tttggaagtc tttcaggtct tgccaatacg	15840
ttgacggtga taacggtggt ctcattgttcc gtaacatcaa caactgcgga gttggagctg	15900
ctgagaagta atgaaggggt gatcgattat gagatcgtac aaagacactg ctaggtgtta	15960
aggatggata ataataataa taatgagatg aatgtgtttt aagttagtgt aacagctgta	16020
ataaagagag agagagagag agagagagag agagagagag agagagagag agagaggctg	16080
atgaaatggt atgtatgttt cttggttttt aaaataaatg aaagcacatg ctcgtgtggt	16140
tctatcgaat tattcggcgg ttcctgtggg aaaaagtcca gaagggccgc cgcagctact	16200
actacaacca aggccgtgga ggagggcaac agagccagca cttcgatagc tgctgcgatg	16260
atcttaagca attgaggagc gagtgcacat gcaggggact ggagcgtgca atcggccaga	16320
tgaggcagga catccagcag caggggacagc agcaggaagt tgagagggtg tcccatcaat	16380
ctaaacaagt cgctagggac cttccgggac agtgcggcac ccagcctagc cgatgccagc	16440
tccaggggca gcagcagtct gcatggtttt gaagtgggtga tcgatgagat cgtataaaga	16500
cactgctagg tgtaaggat gggataataa gatgtgtttt aagtcattaa ccgtaataaa	16560
aagagagaga ggctgatgga atgttatgta tgtatgtttc ttggttttta aaattaaatg	16620
gaaagcacat gctcgtgtgg gttctatctc gattaaaaat cccaattata tttggtctaa	16680
tttagtttgg tattgagtaa aacaaattcg aaccaaacca aaatataaat atatagtttt	16740
tatatatatg cttttaagac tttttataga attttcttta aaaaatatct agaaatattt	16800
gcgactcttc tggcatgtaa ttttctgta aatatgaagt gctccatttt tattaacttt	16860
aaataattgg ttgtacgac actttcttat caagtgttac taaaatgcgt caatctcttt	16920
gttcttccat attcatatgt caaaatctat caaaattctt atatatcttt ttcgaatttg	16980
aagtgaaatt tcgataattt aaaattaaat agaacatatc attatttagg tatcatattg	17040
atttttatac ttaattacta aatttggtta actttgaaag tgtacatcaa cgaaaaatta	17100
gtcaaacgac taaaataaat aaatatcatg tgttattaag aaaattctcc tataagaata	17160
ttttaataga tcatatgttt gtaaaaaaaaa ttaattttta ctaacacata tttttactta	17220
tcaaaaaattt gacaaagtaa gattaaaata atattcatct aacaaaaaaaa aaaccagaaa	17280
atgctgaaaa cccggcaaaa ccgaaccaat ccaaaccgat atagttgggt tggtttgatt	17340
ttgatataaa ccgaaccaac tcggtccatt tgcacccta atcataatag ctttaatat	17400
tcaagatatt attaagttaa cgttgtcaat atcctggaaa ttttgcaaaa tgaatcaagc	17460
ctatatggct gtaatatgaa tttaaaagca gctcgatgtg gtggaatat gtaatttact	17520

10/143

tgattctaaa aaaatatccc aagtattaat aatttctgct aggaagaagg ttagctacga	17580
tttacagcaa agccagaata caaagaacca taaagtgatt gaagctcgaa atatacgaag	17640
gaacaaatat ttttaaaaaa atacgcaatg acttggaaca aaagaaagtg atatatTTTT	17700
tgttcttaaa caagcatccc ctctaaagaa tggcagtttt cctttgcatg taactattat	17760
gctcccttcg ttacaaaaat ttgggactac tattgggaac ttcttctgaa aatagtcctg	17820
caggctagta gattggttgg ttggtttcca tgtaccagaa ggcttaccct attagttgaa	17880
agttgaaact ttgttccta ctcaattcct agttgtgtaa atgtatgtat atgtaatgtg	17940
tataaaacgt agtacttaaa tgactaggag tggttcttga gaccgatgag agatgggagc	18000
agaactaaag atgatgacat aattaagaac gaatttgaaa ggctcttagg tttgaatcct	18060
attcgagaat gtttttgtca aagatagtgg cgattttgaa ccaaagaaaa catttaaaaa	18120
atcagtatcc ggttacgttc atgcaaataa aaagtggctt aggatctgat tgtaatttta	18180
gacttaaaga gtctcttaag attcaatcct ggctgtgtac aaaactacaa ataatatatt	18240
ttagactatt tggccttaac taaacttcca ctcatatttt actgaggtta gagaatagac	18300
ttgcgaataa acacattccc gagaaatact catgatccca taattagtca gagggtatgc	18360
caatcagatc taagaacaca cattccctca aattttaatg cacatgtaat catagtttag	18420
cacaattcaa aaataatgta gtattaaaga cagaaatttg tagacttttt tttggcgta	18480
aaagaagact aagtttatac gtacatttta ttttaagtgg aaaaccgaaa ttttccatcg	18540
aaatatatga atttagtata tatatttctg caatgtacta ttttgctatt ttggcaactt	18600
tcagtggact actactttat tacaatgtgt atggatgcat gagtttgagt atacacatgt	18660
ctaaatgcat gctttgtaaa acgtaacgga ccacaaaaga ggatccatac aaatacatct	18720
catagcttcc tccattatttt tccgacacaa acagagcatt ttacaacaat taccaacaac	18780
aacaacaac aaacaacatt acaattacat ttacaattac cataccatgg cctctatcgc	18840
tatccctgct gctcttgctg gaactcttg atacgttacc tacaatgtgg ctaaccctga	18900
tatcccagct tctgagaaag ttcctgctta cttcatgcag gttgagtact ggggacctac	18960
tatcggaaact attggatacc tcctcttcat ctacttcgga aagcgtatca tgcagaacag	19020
atctcaacct ttcggactca agaacgctat gctcgtttac aacttctacc agaccttctt	19080
caacagctac tgcattctacc ttttcgttac ttctcatagg gctcaggac ttaaggtttg	19140
gggaaacatc cctgatatga ctgctaactc ttggggaatc tctcaggta tctggcttca	19200
ctacaacaac aagtacgttg agcttctcga caccttcttc atggtgatga ggaagaagtt	19260
cgaccagctt tctttccttc acatctacca ccacactctt ctcatctggt catggttcgt	19320
tgttatgaag cttgagcctg ttggagattg ctacttcgga tcttctgtta acaccttctg	19380
gcacgtgatc atgtactctt actacggact tgctgctctt ggagttaact gtttctggaa	19440
gaagtacatc acccagatcc agatgcttca gttctgtatc tgtgcttctc actctatcta	19500
caccgcttac gttcagaata ccgctttctg gcttccttac cttcaactct gggttatggt	19560
gaacatgttc gttctcttcg ccaacttcta ccgtaagagg tacaagtcta agggtgctaa	19620

11/143

gaagcagtga taagggccgc cgccatgtga cagatcgaag gaagaaagtg taataagacg 19680
 actctcacta ctcgatcgct agtgattgtc attgttatat ataataatgt tatctttcac 19740
 aacttatcgt aatgcatgtg aaactataac acattaatcc tacttgatcat atgataacac 19800
 tctccccatt taaaactctt gtcaatttaa agatataaga ttctttaaat gattaaaaaa 19860
 aatatattat aaattcaatc actcctacta ataaattatt aattattatt tattgattaa 19920
 aaaaatactt atactaattt agtctgaata gaataattag attctagcct gcagggcggc 19980
 cgcggatccc atggagtcaa agattcaaat agaggaccta acagaactcg ccgtaaagac 20040
 tggcgaacag ttcatacaga gtctcttacg actcaatgac aagaagaaaa tcttcgtcaa 20100
 catggtggag cacgacacac ttgtctactc caaaaatatc aaagatacag ttcagaaga 20160
 ccaaagggca attgagactt ttcaacaaaag ggtaatatcc ggaaacctcc tcggattcca 20220
 ttgcccagct atctgtcact ttattgtgaa gatagtggaa aaggaaggtg gctcctacaa 20280
 atgccatcat tgcgataaag gaaaggccat cgttgaagat gcctctgccg acagtgggcc 20340
 caaagatgga cccccacca cgaggagcat cgtggaaaaa gaagacgttc caaccacgtc 20400
 ttcaaagcaa gtggattgat gtgatatctc cactgacgta agggatgacg cacaatccca 20460
 ctatccttcg caagaccctt cctctatata aggaagttca tttcatttgg agagaacacg 20520
 ggggactgaa ttaaatatga gccctgagag gcgtcctggt gaaatcagac ctgctactgc 20580
 tgctgatatg gctgctgttt gtgatatcgt gaaccactac atcgagactt ctaccgttaa 20640
 cttcagaact gagcctcaaa ctcctcaaga gtggatcgat gatcttgaga gactccaaga 20700
 tagataccct tggcttggtg ctgaggttga ggggtgtgtt gctggaatcg cttacgctgg 20760
 accttggaag gctagaaacg cttacgattg gactgttgag tctaccgttt acgtttcaca 20820
 cagacatcag agacttggac ttggatctac cttttactac caccttctca agtctatgga 20880
 agctcagga ttcaagtctg ttgttgctgt tatcggaact cctaacgatc cttctgttag 20940
 acttcatgag gctcttgat acactgctag aggaactctt agagctgctg gatacaagca 21000
 cgggtggatg catgatgttg gattctggca aagagatttc gagcttcctg ctcctcctag 21060
 acctgttaga ccagttactc agatctgaat ttgcgtgatc gttcaaacat ttggcaataa 21120
 agtttcttaa gattgaatcc tgttgccggg cttgcatga ttatcatata atttctgttg 21180
 aattacgtta agcatgtaat aattaacatg taatgcatga cgttatttat gagatgggtt 21240
 tttatgatta gagtcccgc attatacatt taatacgca tagaaaacaa aatatagcgc 21300
 gcaaactagg ataaattatc gcgcgcgggt tcatctatgt tactagatca ctagtgatgt 21360
 acggttaaaa ccaccccgat acattaaaaa cgtccgcaat gtgttattaa gttgtctaag 21420
 cgtcaatttg ttacaccac aatatatcct gccaccagcc agccaacagc tccccgaccg 21480
 gcagctcggc acaaaatcac cactcgatac aggcagccca tcagtcc 21527

<210> 2
 <211> 23512
 <212> ADN

12/143

<213> Trình tự nhân tạo

<220>

<223> Trình tự nucleotit pGA7- mod_B

<400> 2

tcctgtgggtt ggcatgcaca tacaaatgga cgaacggata aacctttttca cgccctttta	60
aatatccgat tattctaata aacgctcttt tctcttaggt ttacccgcca atatatcctg	120
tcaaacactg atagtttaaa ctgaaggcgg gaaacgacaa tctgctagt gatctcccag	180
tcacgacgtt gtaaaacggg cgccccgcgg aaagcttgcg gccgcggtac cgcccggtcg	240
actcagatct tccaaggcct cgtctccgag tccgctgctt ctcgccgcgc cgatcacttc	300
tccgccgcca acaaggcttg tagttaatag gaatcattca gggattgtga ttccgggacg	360
tagtaattaa taatatagta ttagtataga taatatgttt cgtttgggat ctttggaacg	420
ttgctctgtt ccttggtgtt cattttaaaag cttttgaggg atagttgcag aactgttcgg	480
tgatgcttca tcctctcaag aactagattt gggtaaagaa acatccatgc atggatatgg	540
aatgttggtt ttccgattgg agattatatt ataaaattta aaattcatga tttaaaaaaa	600
cacataaaaa ccacaaaatt catgatttat tgacaatacg atacaaaatt agcaccaccg	660
gctactggct cattacacat ttccccttcc cctcattctc actttgtggc tttattatta	720
ttattattac atatatttta ccgttattat ttcacgtcac ataagcttgt taattaatca	780
ttagtgagcc ttctcagcct ttccgttaac gtagtagtgc tgtcccacct tatcaagggt	840
agagaaagta gccttccaag caccgtagta agagagcacc ttgtagttga gtccccactt	900
cttagcgaaa ggaacgaatc ttctgctaac ctgaggctgt ctgaattgag gcatatcagg	960
gaagagggtg tggataacct gacagttaag gtatcccata agccagttca cgtatcctct	1020
agaaggatcg atatcaacgg tgtgatcaac agcgtagtta acccaagaaa ggtgcttatt	1080
agatggaaca acagggagggt gagtatgaga agtagagaag tgagcgaaaa ggtacatgta	1140
agcgatccag tttccgaaag tgaaccacca gtaagcaaca ggccaagagt atccagtagc	1200
aagcttgata acagcggttc taacaacatg agaaacgagc atccaagaag cctcttcgta	1260
gttcttctta cggagaactt gtctaggggt gagaacgtag atccagaaag cttgaacaag	1320
aagtccagag gtaacaggaa cgaaagtcca agcttgaagt ctagccaag ctctagagaa	1380
tcctctaggt ctgttatcct caacagcagt gttgaagaaa gccacagcag gagtggatc	1440
aagatccata tcgtgtctaa ccttttgagg ggtagcatgg tgcttggtat gcatctggtt	1500
ccacatctca ccagaagtag aaagtccgaa tccacaagtc atagcctgaa gtctcttgtc	1560
cacgtaaaca gatccggtta gagagttagt tccaccctca tgttgaacct atccacatct	1620
agctccgaag aaagcaccgt aaacaacaga agcaatgata gggatatccag cgtacataag	1680
agcagttcca agagcgaatg tagcaagaag ctcgagaagt ctgtaagcca catgggtgat	1740
agaaggcttg aagaatccat ctctctcaag ctcagcacgc catctagcga aatcctcaag	1800
cataggagca tcctcagact cagatctctt gatctcagca ggtctagaag gcaaagctct	1860
aagcatcttc caagccttga gagaacgcat gtggaattct ttgaaagcct cagtagcatc	1920

13/143

agcaccagtg ttagcaagca tgtagaagat cacagatcca ccagggtgct tgaagttagt	1980
cacatcgtac tcaacgtcct caactctaac ccatctagtc tcgaaagtag cagcaagctc	2040
atgaggctca agagtcttaa gatcaacagg agcagtagaa gcatccttag catcaagagc	2100
ctcagcagaa gatttagacc tggtaagtgg agatctagga gaagatcttc catcagtctt	2160
aggagggcac atggtatggt aattgtaaat gtaattgtaa tgttgtttgt tgtttgttgt	2220
tgttggtaat tgttgtaaaa ttaattaagt gggatatctt tggtatggata agcaagtagt	2280
gatgatgttc taggtgaagt gatgggggtg ttttatagcg ggagatggtg aaatggatgg	2340
tcgccacata agaaatggag ggggaagggtt cttgcgcat tcttcagttt gcatggatgc	2400
atgggtttca ttttgtaaca cgtaataagg acaatgaagt gcagggtgtct ctcaagtttc	2460
agaggggata tgtggacaga agaagaacgg cgatgatatt gatggaaatg gccatctagt	2520
gtgaatctat tcggttgata atactagtg c attttggccg ttaatccctt caattaactg	2580
cacaaacttc agttgagtat tgattatttg attataggtt ctgtaaacac aataccaagt	2640
ttatttagag gggagacata caaatagttt cgatataaat aatagagtgg ttaaacttag	2700
ttattaaaac tatatataaa gtctaaaagt taaattattt ttttaattgc aaatatataa	2760
agtctaaagg ggttacatta tttcttaaga gatgtaactc tgttggaatc tgacttaatc	2820
cgtctcatca ctctggtttc cagttcta at ctaatgaatt gttttctgcc aaagaatttg	2880
aagcaagaag taaattgatc aatgccgtca acccacacca aaccgtcaac cactaccat	2940
cgccgaggag acccccaaac tcaacctcca cccatcggtg agaagcacag ggcagccgc	3000
accaccacca atttggcgtg catgacacct agggacttgg cacgggaggc ggcgcacgtg	3060
gatgcaa atg acgggatatc agatgacagg aaacgacgtt gagagaccat acgatgtaga	3120
atatgagctc accatcaacg agaaactagg aaaatcaca aaaaaacaac tctcgtaatt	3180
gtacgagtgg cacagatggg tctgcctcaa catatctcta atacggcgaa gcctgccc aa	3240
cacgtagtgt ccggaatccg gtgtggagct cacgactctg aaagataggc gcttcctgtt	3300
tcgtttcgct caccactgg acgtccgtca tgtgatggat ttcggtcatt ggttgctga	3360
caaccacatt ctgaagctcc atgagatgag tcttcacaat aggtcctgct caataccgtg	3420
gagttatggt tgcaagtcca taacttgccg ttcgaatatt ttgcggagcc agtcggacgg	3480
gaattggcga gctcggctga cacctataaa ggcatgaca agaagaacca aaagttcttc	3540
ccta atgctt tcatgaggct tcgggtcgtt atggatgtcg gaaaaccct cttgaaggaa	3600
cgagacgtta ttatgcatga cggtaaact attacttgtc agtataagta tgaaagatta	3660
cctgtcttct gctttgtttg tggattgatt ggacacgttg aaaaaaatg tgcacttcga	3720
tttcaatact cagagatcga cttccctttt ctctaggagt attcgatcaa ggcattaaca	3780
tgaaggaag ctcaagctct aaaggcttca caatggaacc tgaaaaattt caacaagcct	3840
aaactgaaat cgaagtcaaa tcaccaacc gggagctcta aatcagcaaa cactcctcct	3900
ccacagtatc caatcatcgt gcacgatgct ccagggtattg caagccagggt attgcaagct	3960

14/143

aggagtagga tagagacctt aaacgtcggt ggtgtgaaga gtcattctca gacctaatgg	4020
agatagatgt agacggcggc acgaagactc tgaaacacca gaaaggctag tccaggataa	4080
ggatctgcta tcccaactga cctctcggtt gtcccaaggc ctctcaacta gagcaggagg	4140
aaggatgggt acaagactag gataatgatg tttccaatat gaacctgaat gtccatagct	4200
aatTTTTTTT gtcttgcttc tgcactTTTT gtttattatg ttctggtgac tatgttattt	4260
acccttgctc gtatgcttga gggtagccta gtagattggt tgggttggtt ccatgtacca	4320
gaaggcttac cctattagtt gaaagttgaa actttgttcc ctactcaatt cctagttgtg	4380
taaatgtatg tatatgtaat gtgtataaaa cgtagtactt aaatgactag gagtggttct	4440
tgagaccgat gagagatggg agcagaacta aagatgatga cataattaag aacgaatttg	4500
aaaggctctt aggtttgaat cctattcgag aatgtttttg tcaaagatag tggcgatttt	4560
gaaccaaaga aaacatttaa aaaatcagta tccggttacg ttcatgcaaa tagaaagtgg	4620
tctaggatct gattgtaatt ttagacttaa agagtctctt aagattcaat cctggctgtg	4680
tacaaaacta caaataatat attttagact atttggcctt aactaaactt ccactcatta	4740
tttactgagg ttagagaata gacttgcgaa taaacacatt cccgagaaat actcatgatc	4800
ccataattag tcagagggta tgccaatcag atctaagaac acacattccc tcaaatttta	4860
atgcacatgt aatcatagtt tagcacaatt caaaaataat gtagtattaa agacagaaat	4920
ttgtagactt ttttttggcg ttaaaagaag actaagttaa tacgtacatt ttattttaag	4980
tggaaaaccg aaattttcca tcgaaatata tgaatttagt atatatattt ctgcaatgta	5040
ctattttgct attttggcaa ctttcagtgg actactactt tattacaatg tgtatggatg	5100
catgagtttg agtatacaca tgtctaaatg catgctttgt aaaacgtaac ggaccacaaa	5160
agaggatcca tacaaatata tctcatagct tctccatta ttttccgaca caaacagagc	5220
attttacaac aattaccaac aacaacaaac aacaacaaac attacaatta catttacaat	5280
taccatacca tggcctctat cgctatccct gctgctcttg ctggaactct tggatacgtt	5340
acctacaatg tggctaacc ctagatccca gcttctgaga aagttcctgc ttacttcatg	5400
caggttgagt actggggacc tactatcgga actattggat acctcctctt catctacttc	5460
ggaaagcgta tcatgcagaa cagatctcaa ctttctggac tcaagaacgc tatgctcggt	5520
tacaacttct accagacctt cttcaacagc tactgcatct accttttcgt tacttctcat	5580
agggtcagg gacttaagggt ttggggaaac atccctgata tgactgctaa ctcttgggga	5640
atctctcagg ttatctggct tcactacaac aacaagtacg ttgagcttct cgacaccttc	5700
ttcatgggtg tgaggaagaa gttcgaccag ctttctttcc ttcacatcta ccaccacact	5760
cttctcatct ggtcatgggt cgttgttatg aagcttgagc ctgttgagga ttgctacttc	5820
ggatcttctg ttaacacctt cgtgcacgtg atcatgtact ctactacgg acttgctgct	5880
cttggagtta actgtttctg gaagaagtac atcaccaga tccagatgct tcagttctgt	5940
atctgtgctt ctactctat ctacaccgct tacgttcaga ataccgcttt ctggcttcct	6000
taccttcaac tctgggttat ggtgaacatg ttcgttctct tcgccaactt ctaccgtaag	6060

15/143

aggtacaagt ctaagggtgc taagaagcag tgataaggcg cgcggcgcg cgggccgccg	6120
ccatgtgaca gatcgaagga agaaagtgtg ataagacgac tctcactact cgatcgctag	6180
tgattgtcat tgttatatat aataatgtta tctttcacaa cttatcgtaa tgcattgtgaa	6240
actataacac attaatccta cttgtcatat gataacactc tccccattta aaactcttgt	6300
caatttaaag atataagatt ctttaaatga ttaaaaaaaaa tatattataa attcaatcac	6360
tcctactaat aaattattaa ttattattta ttgattaaaa aaatacttat actaatttag	6420
tctgaataga ataattagat tctagtctca tcccctttta aaccaactta gtaaactgtt	6480
ttttttttta ttttatgaag ttaagttttt accttgtttt taaaaagaat cgttcataag	6540
atgccatgcc agaacattag ctacacgtta cacatagcat gcagccgcg agaattgttt	6600
ttcttcgcca cttgtcactc cttcaaaca cctaagagct tctctctcac agcacacaca	6660
tacaatcaca tgcgtgcatg cattattaca cgtgatcgcc atgcaaactt cttttatagc	6720
ctataaatta actcatccgc ttcactcttt actcaaaca aaactcatcg atacaaacaa	6780
gattaaaaac atacacgagg atcttttaca acaattacca acaacaacaa acaacaacaa	6840
acattacaat tacatttaca attaccatac catgcctcca agggactctt actcttatgc	6900
tgctcctcct tctgctcaac ttcacgaagt tgatactcct caagagcacg acaagaaaga	6960
gcttgttatc ggagataggg cttacgatgt taccaacttc gttaagagac accctggtgg	7020
aaagatcatt gcttaccaag ttggaactga tgctaccgat gcttacaagc agttccatgt	7080
tagatctgct aaggctgaca agatgcttaa gtctcttcct tctcgtcctg ttcacaaggg	7140
atactctcca agaagggctg atcttatcgc tgatttcaa gagttcacca agcaacttga	7200
ggctgagggg atgttcgagc cttctcttcc tcatgttgct tacagacttg ctgagggtat	7260
cgctatgcat gttgctggtg ctgctcttat ctggcatgga tacactttcg ctggaatcgc	7320
tatgcttgga gttgttcagg gaagatgtgg atggcttatg catgaggggtg gacattactc	7380
tctcactgga aacattgctt tcgacagagc tatccaagtt gcttgttacg gacttggtatg	7440
tggaatgtct ggtgcttggt ggcgtaacca gcataacaag caccatgcta ctctcaaaa	7500
gcttcagcac gatgttgatc ttgataccct tcctctcggt gctttccatg agagaatcgc	7560
tgctaagggt aagtctcctg ctatgaaggc ttggctttct atgcaagcta agcttttcgc	7620
tcctgttacc actcttcttg ttgctcttg atggcagctt taccttcac ctagacacat	7680
gctcaggact aagcactacg atgagcttgc tatgctcgga atcagatacg gacttggttg	7740
ataccttgct gctaactacg gtgctggata cgctctcgct tgttaccttc ttacgttca	7800
gcttgagct atgtacatct tctgcaactt cgctgtttct catactcacc tccctgttgt	7860
tgagcctaac gagcatgcta cttgggttga gtacgtgct aaccacacta ctaactgttc	7920
tccatcttgg tgggtgtgatt ggtggatgtc ttaccttaac taccagatcg agcaccacct	7980
ttaccttct atgcctcaat tcagacaccc taagatcgct cctagagtta agcagctttt	8040
cgagaagcac ggacttcact acgatgttag aggatacttc gaggctatgg ctgatacttt	8100

16/143

cgctaacctt gataacgttg cccatgctcc tgagaagaaa atgcagtaat gagatcggtc	8160
aaacatttgg caataaagtt tcttaagatt gaatcctggt gccggtcttg cgatgattat	8220
catataatth ctgttgaatt acgttaagca cgtaataatt aacatgtaat gcatgacgtt	8280
atthtatgaga tgggttttta tgattagagt cccgcaatta tacatttaat acgcgataga	8340
aaacaaaata tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact	8400
agatcggtcg attaaaaatc ccaattatat ttggtctaatt ttagtttggt attgagtaaa	8460
acaaattcga accaaaccaa aatataaata tatagttttt atatatatgc ctthaagact	8520
ttttatagaa ttttctttaa aaaatatcta gaaatatttg cgactcttct ggcatgtaat	8580
atthcgttaa atatgaagtg ctccattttt attaaactta aataattggt tgtacgatca	8640
ctttcttatac aagtgttact aaaatgcgtc aatctctttg ttcttcata ttcatatgtc	8700
aaaatctatac aaaattctta tatatctttt tcgaatttga agtgaaattt cgataattta	8760
aaattaaata gaacatatca ttatttaggt atcatattga tttttatact taattactaa	8820
atthggttaa ctttgaaagt gtacatcaac gaaaaattag tcaaacgact aaaataaata	8880
aatatcatgt gttattaaga aaattctcct ataagaatat tthaatagat catatgtttg	8940
taaaaaaaaat taatttttac taacacatat atthacttat caaaaatttg acaaagtaag	9000
attaaaaata tattcatcta aaaaaaaaaa aaccagaaaa tgctgaaaac ccggcaaaac	9060
cgaaccaatc caaacgata tagttggttt ggtttgattt tgatataaac cgaaccaact	9120
cggtccattt gcaccctaa tcataatagc tthaatatth caagatatta ttaagttaac	9180
gttgtcaata tcctggaaat ttgcaaaat gaatcaagcc tatatggctg taatatgaat	9240
ttaaaagcag ctcatgttg tggtaatatg taatttactt gattctaaaa aaatatccca	9300
agtattaata atthctgcta ggaagaaggt tagctacgat ttacagcaa gccagaatac	9360
aaagaacat aaagtgattg aagctcgaaa tatacgaagg acaaatatt tthaaaaaaaa	9420
tacgcaatga cttggaacaa aagaaagtga tatatttttt gtctttaaac aagcatcccc	9480
tctaaagaat ggcagttttc ctttgcatgt aactattatg ctcccttcgt tacaaaaatt	9540
ttggactact attgggaact tcttctgaaa atagtgatag aaccacacg agcatgtgct	9600
ttccatttaa tthtaaaaac caagaaacat acatacataa cattccatca gcctctctct	9660
ctttttatta cggttaatga cttaaaacac atcttattat cccatcctta acacctagca	9720
gtgtctttat acgatctcat cgatcaccac ttcaaaacca tgcagactgc tgctgccct	9780
ggagctggca tcggctaggc tgggtgccgc actgtcccgg aaggcccta gcgacttggt	9840
tagattgatg ggaccacctc tcaacttcct gctgctgtcc ctgctgctgg atgtcctgcc	9900
tcctctggcc gattgcacgc tccagtcccc tgcattgtga ctgctcctc aattgcttaa	9960
gatcatcgca gcagctatcg aagtgtggc tctgttgccc tctccacgg ccttggttgt	10020
agtagtagct gccgccgcc ttctggactt tttccacag gaaccgccga ataattcgat	10080
agaaccacac gagcatgtgc tttcatttat tthaaaaacc aagaaacata cataacattt	10140
catcagctc tctctctctc tctctctctc tctctctctc tctctctctc tctctctctt	10200

17/143

tattacagct gttacactaa cttaaaacac attcatctca ttattattat tattatccat	10260
ccttaacacc tagcagtgtc tttgtacgat ctcataatcg atcacccctt catcagggtat	10320
ccttaggctt cactccaacg ttgttgcagt tacggaacat gtacacacca tcatggttct	10380
caacgaactg gcaagatctc caagttttcc aaaggctaac ccacatgttc tcatcggtgt	10440
gtctgtagtg ctctcccata actttcttga tgcactcggg agcttctcta gcatggtaga	10500
atgggacctt tgaaacgtag tgatggagca catgagcttc gatgatgtca tggaagatga	10560
ttccgaggat tccgaactct ctatcgatag tagcagcagc acccttagcg aaagtccact	10620
cttgagcatc gtaatgaggc atagaagaat cgggtgtgctg aaggaaggta acgaaaacaa	10680
gccagtgggtt aacaaggatc caaggacaga accatgtgat gaaagtaggc cagaatccga	10740
aaaccttgta agcgggtgtaa acagaagtga gggtagcaag gattccaaga tcagaaagaa	10800
cgatgtacca gtagtccttc ttatcgaaaa cagggctaga aggccagtag tgagacttga	10860
agaacttaga aacaccaggg taaggttgtc cagtagcggt agtagcaagg taaagagaaa	10920
gtcctccaag ctgttggaaac aagagagcga aaacagagta gataggagtt tcctcagcga	10980
tatcgtgaag gctggtaact tgggtgcttct ctttgaattc ctcggcgggtg taaggaacga	11040
aaaccatatc tctggtcatg tgtccagtag ctttatgggtg cttagcatga gagaacttcc	11100
agctgaagta aggaaccata acaagagagt ggagaaccca tccaacggta tcgttaaccc	11160
atccgtagtt agagaaagca gaatgtccac actcatgtcc aaggatccag attccgaatc	11220
cgaacaaga gatagagaac acgtaagcag accaagcagc gaatctaagg aattcgtag	11280
ggagaagagg gatgtaggta agtccaacgt aagcgatagc agagatagcc acgatatctc	11340
tcaccacgta agacatagac ttcacgagag atctctcgta acagtgccta gggatagcgt	11400
caaggatatc cttgatgggtg taatctggca cttgaaaac gtttccgaag gtatcgatag	11460
cggctctttt ctgcttgaaa gatgcaacgt ttccagaacg cctaacggtc ttagtagatc	11520
cctcaaggat ctcatatcca gacacggtaa ccttagacat ggtatggtaa ttgtaaattgt	11580
aattgtaattg ttgtttgttg tttgttgttg ttggttaattg ttgtaaaatt tttggtgggtg	11640
attggttctt taagggtgtga gagtgagttg tgagttgtgt ggtgggttg gtgagattgg	11700
ggatgggtggg tttatatagt ggagactgag gaatgggggtc gtgagtgtta actttgcatg	11760
ggctacacgt gggttctttt gggcttacac gtagtattat tcatgcaaata gcagccaata	11820
catatacggg attttaataa tgtgtgggaa tacaatatgc cgagtatttt actaattttg	11880
gcaatgacaa gtgtacattt ggattatctt acttggcctc tcttgcttta atttgatta	11940
tttttattct cttaccttgg ccgttcatat tcacatccct aaaggcaaga cagaattgaa	12000
tggtggccaa aaattaaaac gatggatatg acctacatag tgtaggatca attaacgtcg	12060
aaggaaaata ctgattctct caagcatacg gacaagggtg aataacatag tcaccagaac	12120
ataataaaca aaaagtgcag aagcaagact aaaaaatta gctatggaca ttcaggttca	12180
tattggaaac atcattatcc tagtcttgtg accatccttc ctctgctct agttgagagg	12240

18/143

ccttgggact aacgagaggt cagttgggat agcagatcct taccctggac tagcctttct	12300
ggtgtttcag agtcttcgtg ccgccgtcta catctatctc cattaggtct gaagatgact	12360
cttcacacca acgacgttta aggtctctat cctactccta gcttgcaata cctggcctgc	12420
aatacctgga gcatcgtgca cgatgattgg atactgtgga ggaggagtgt ttgctgattt	12480
agagctcccg gttgggtgat ttgacttcga tttcagttta ggcttggtga aatttttcag	12540
gttccattgt gaagccttta gagcttgagc ttccttccat gttaatgcct tgatcgaata	12600
ctcctagaga aaaggaagt cgatctctga gtattgaaat cgaagtgcac attttttttc	12660
aacgtgtcca atcaatccac aaacaaagca gaagacaggt aatctttcat acttatactg	12720
acaagtaata gtcttaccgt catgcataat aacgtctcgt tccttcaaga ggggttttcc	12780
gacatccata acgacccgaa gcctcatgaa agcattaggg aagaactttt ggttcttctt	12840
gtcatggcct ttataggtgt cagccgagct cgccaattcc cgtccgactg gctccgaaa	12900
atattcgaac ggcaagtatt ggacttgcaa ccataactcc acggtattga gcaggaccta	12960
ttgtgaagac tcatctcatg gagcttcaga atgtggttgt cagcaaacca atgaccgaaa	13020
tccatcacat gacggacgtc cagtgggtga gcgaaacgaa acaggaagcg cctatctttc	13080
agagtcgtga gctccacacc ggattccggc aactacgtgt tgggcaggct tcgccgtatt	13140
agagatatgt tgaggcagac ccactctgtc cactcgtaca attacgagag ttgttttttt	13200
tgtgattttc ctagtttctc gttgatggtg agctcatatt ctacatcgta tggctctctca	13260
acgtcgtttc ctgtcatctg atatcccgtc atttgcatcc acgtgcgccg cctcccgtc	13320
caagtcccta ggtgtcatgc acgccaaatt ggtggtggtg cgggctgccc tgtgcttctt	13380
accgatgggt ggaggttgag tttgggggtc tccgcggcga tggtagtggg ttgacggttt	13440
ggtgtgggtt gacggcattg atcaatttac ttcttgcttc aaattctttg gcagaaaaca	13500
attcattaga ttagaactgg aaaccagagt gatgagacgg attaagtcag attccaacag	13560
agttacatct cttaagaaat aatgtaacct ctttagactt tatatatattg caattaaaaa	13620
aataatttaa ctttttagact ttatatatag ttttaataac taagtttaac cactctatta	13680
tttatatcga aactatttgt atgtctcccc tctaaataaa cttggtattg tgtttacaga	13740
acctataatc aaataatcaa tactcaactg aagtttgtgc agttaattga agggattaac	13800
ggccaaaatg cactagtatt atcaaccgaa tagattcaca ctagatggcc atttccatca	13860
atatcatcgc cgttcttctt ctgtccacat atcccctctg aaacttgaga gacacctgca	13920
cttcattgtc cttattacgt gttacaaaat gaaacccatg catccatgca aactgaagaa	13980
tggcgcaaga acccttcccc tccatttctt atgtggcgac catccatttc accatctccc	14040
gctataaaac acccccatca cttcacctag aacatcatca ctacttgctt atccatccaa	14100
aagataccca cttttacaac aattaccaac aacaacaaac aacaaacaac attacaatta	14160
catttacaat taccatacca tgccacctag cgctgctaag caaatgggag cttctactgg	14220
tgttcatgct ggtgttactg actcttctgc tttcaccaga aaggatgttg ctgatagacc	14280
tgatctcacc atcgttggag attctgttta cgatgctaag gctttcagat ctgagcatcc	14340

19/143

tggtggtgct catttcgttt ctttgttcgg aggaagagat gctactgagg ctttcatgga	14400
ataccataga agggcttggc ctaagtctag aatgtctaga ttccacgttg gatctcttgc	14460
ttctactgag gaacctgttg ctgctgatga gggatacctt caactttgtg ctaggatcgc	14520
taagatggtg ctttctgttt cttctggatt cgctcctgct tcttactggg ttaaggctgg	14580
acttatcctt ggatctgcta tcgctcttga ggcttacatg ctttacgctg gaaagagact	14640
tctcccttct atcgttcttg gatggctttt cgctcttatt ggtcttaaca tccagcatga	14700
tgctaaccat ggtgctttgt ctaagtctgc ttctgttaac cttgctcttg gactttgtca	14760
ggattggatc ggaggatcta tgatcctttg gcttcaagag catgttggtta tgcaccacct	14820
ccacactaac gatgttgata aggatcctga tcaaaaggct cacggtgctc ttagactcaa	14880
gcctactgat gcttggtcac ctatgcattg gcttcagcat ctttaccttt tgcctggtga	14940
gactatgtac gctttcaagc ttttgttctt cgacatctct gagcttggtta tgtggcgttg	15000
ggagggtgag cctatctcta agcttgctgg atacctcttt atgccttctt tgcttctcaa	15060
gcttaccttc tgggctagat tcgttgcttt gcctctttac cttgctcctt ctgttcatac	15120
tgctgtgtgt atcgctgcta ctgttatgac tggatctttc tacctcgctt tcttcttctt	15180
catctccac aacttcgagg gtgttgcttc tgttggaact gatggatcta tcacttctat	15240
gactagaggt gctagcttcc ttaagagaca agctgagact tcttctaacg ttggaggacc	15300
tcttcttgct actcttaacg gtggactcaa ctaccaaatt gagcatcact tgttccctag	15360
agttcaccat ggattctacc ctagacttgc tcctcttggt aaggctgagc ttgaggctag	15420
aggaatcgag tacaagcact accctactat ctggtctaac cttgcttcta ccctcagaca	15480
tatgtacgct cttggaagaa ggcctagatc taaggctgag taatgacaag cttatgtgac	15540
gtgaaataat aacggtaaaa tatatgtaat aataataata ataaagccac aaagtgaagaa	15600
tgaggggaag gggaaatgtg taatgagcca gtagccggtg gtgctaattt tgtatcgtat	15660
tgtcaataaa tcatgaattt tgtgggtttt atgtgttttt ttaaatacatg aattttaaat	15720
tttataaaat aatctccaat cggaagaaca acattccata tccatgcatg gatgtttctt	15780
tacccaaatc tagttcttga gaggatgaag catcaccgaa cagttctgca actatccctc	15840
aaaagcttta aaatgaacaa caaggaacag agcaacgttc caaagatccc aaacgaaaca	15900
tattatctat actaatacta tattattaat tactactgcc cggaatcaca atccctgaat	15960
gattcctatt aactacaagc cttgttggtg gcggagaagt gatcggcgcg gcgagaagca	16020
gcggactcgg agacgaggcc ttggaagatc tgagtcgaac gggcagaatc agtattttcc	16080
ttcgacgtta attgatccta cactatgtag gtcatatcca tcgttttaat ttttggccac	16140
cattcaattc tgtcttgctt ttagggatgt gaatatgaac ggccaaggta agagaataaa	16200
aataatccaa attaaagcaa gagaggccaa gtaagataat ccaaatagtac acttgtcatt	16260
gccaaaatta gtaaaatact cggcatattg tattcccaca cattattaaa ataccgtata	16320
tgtattggct gcatttgcatt gaataatact acgtgtaagc ccaaagaac ccacgtgtag	16380

20/143

cccatgcaaa gttaacactc acgaccccat tcctcagtct ccactatata aaccaccat	16440
ccccaatctc accaaacca ccacacaact cacaactcac tctcacacct taaagaacca	16500
atcaccacca aaaattttac aacaattacc aacaacaaca aacaacaaac aacattacaa	16560
ttacatttac aattaccata ccatgagcgc tgttaccggt actggatctg atcctaagaa	16620
cagaggatct tctagcaaca ccgagcaaga ggttccaaaa gttgctatcg ataccaacgg	16680
aaacgtgttc tctgttcctg atttcaccat caaggacatc cttggagcta tccctcatga	16740
gtgttacgag agaagattgg ctacctctct ctactacgtg ttcagagata tcttctgcat	16800
gcttaccacc ggatacctta ccataagat cttttaccct ctctcatct cttacacctc	16860
taacagcatc atcaagttca ctttctgggc cttttacact tacgttcaag gacttttcgg	16920
aaccggaatc tgggttctcg ctcatgagtg tggacatcaa gctttctctg attacggaat	16980
cgtgaacgat ttcgttggat ggacccttca ctcttacctt atgggttcctt acttcagctg	17040
gaagtactct catggaaaagc accataaggc tactggacac atgaccagag atatggtttt	17100
cgttcctgcc accaaagagg aattcaagaa gtctaggaac ttcttcggta acctcgctga	17160
gtactctgag gattctccac ttagaacctt ttacgagctt cttgttcaac aacttgagg	17220
atggatcgct tacctcttcg ttaacgttac aggacaacct taccctgatg ttccttcttg	17280
gaaatggaac cacttctggc ttacctctcc acttttcgag caaagagatg ctctctacat	17340
cttcctttct gatcttgga tcctcaccca gggaatcgtt ctactcttt ggtacaagaa	17400
attcgaggga tgggtccctt tcatcaactg gttcgttcct tacatctggg ttaaccactg	17460
gctcgttttc atcacattcc ttcagcacac tgatcctact atgcctcatt acaacgctga	17520
ggaatggact ttcgctaagg gtgctgctgc tactatcgat agaaagtctg gattcatcgg	17580
acctcacatc ttccatgata tcatcgagac tcatgtgctt caccactact gttctaggat	17640
cccattctac aacgctagac ctgcttctga ggctatcaag aaagttatgg gaaagcacta	17700
caggtctagc gacgagaaca tgtggaagtc actttggaag tctttcaggt cttgccaata	17760
cgttgacggt gataacggtg ttctcatgtt ccgtaacatc aacaactgcg gagttggagc	17820
tgctgagaag taatgaaggg gtgatcgatt atgagatcgt acaaagacac tgctaggtgt	17880
taaggatgga taataataat aataatgaga tgaatgtgtt ttaagttagt gtaacagctg	17940
taataaagag agagagagag agagagagag agagagagag agagagagag agagagaggc	18000
tgatgaaatg ttatgtatgt ttcttggttt ttaaaataaa tgaaagcaca tgctcgtgtg	18060
gttctatcga attattcggc gggtcctgtg ggaaaaagtc cagaagggcc gccgcagcta	18120
ctactacaac caaggccgtg gaggagggca acagagccag cacttcgata gctgctgcga	18180
tgatcttaag caattgagga gcgagtgcac atgcagggga ctggagcgtg caatcggcca	18240
gatgaggcag gacatccagc agcagggaca gcagcaggaa gttgagaggt ggtcccatca	18300
atctaaacaa gtcgctaggg accttccggg acagtgcggc acccagccta gccgatgcca	18360
gctccagggg cagcagcagt ctgcatggtt ttgaagtggg gatcgatgag atcgtataaa	18420
gacactgcta ggtgttaagg atgggataat aagatgtgtt ttaagtcatt aaccgtaata	18480

aaaagagaga gaggctgatg gaatggtatg tatgtatggt tcttggtttt taaaattaaa	18540
tggaagcac atgctcgtgt gggttctatc tcgattaaaa atcccaatta tatttgggtct	18600
aatttagttt ggtattgagt aaaacaaatt cgaaccaaac caaaatataa atatatagtt	18660
tttatatata tgcctttaag actttttata gaattttctt taaaaaatat ctagaaatat	18720
ttgcgactct tctggcatgt aatatttcgt taaatatgaa gtgctccatt tttattaact	18780
ttaaataatt ggttgtaga tcactttctt atcaagtgtt actaaaatgc gtcaatctct	18840
ttgttcttcc atattcatat gtcaaaatct atcaaaattc ttatatatct ttttcgaatt	18900
tgaagtgaat tttcgataat ttaaaattaa atagaacata tcattattta ggtatcatat	18960
tgatttttat acttaattac taaatttggt taactttgaa agtgtagatc aacgaaaaat	19020
tagtcaaacg actaaaataa ataaatatca tgtgttatta agaaaattct cctataagaa	19080
tattttaata gatcatatgt ttgtaaaaaa aattaatttt tactaacaca tatatttact	19140
tatcaaaaat ttgacaaagt aagattaaaa taatattcat ctaacaaaaa aaaaaccaga	19200
aatgctgaa aacccggcaa aaccgaacca atccaaaccg atatagttgg tttggtttga	19260
ttttgatata aaccgaacca actcgggtcca tttgcacccc taatcataat agctttaata	19320
tttcaagata ttattaagtt aacgttgtca atatcctgga aattttgcaa aatgaatcaa	19380
gcctatatgg ctgtaatatg aatttaaaag cagctcgatg tgggtggaat atgtaattta	19440
cttgattcta aaaaaatc ccaagtatta ataatttctg ctaggaagaa ggtagctac	19500
gatttacagc aaagccagaa tacaagaac cataaagtga ttgaagctcg aaatatacga	19560
aggaacaaat atttttaaaa aaatacgcaa tgacttgga caaaagaaag tgatatattt	19620
tttggttcta aacaagcatc ccctctaaag aatggcagtt ttcctttgca tgtaactatt	19680
atgctccctt cgttacaaaa attttggtact actattggga acttcttctg aaaatagtcc	19740
tgaggctag tagattgggt ggttggtttc catgtaccag aaggcttacc ctattagttg	19800
aaagtgaat ctttggtccc tactcaattc ctagttgtgt aaatgtatgt atatgtaatg	19860
tgtataaaac gtagtactta aatgactagg agtggttctt gagaccgatg agagatggga	19920
gcagaactaa agatgatgac ataattaaga acgaatttga aaggctctta ggttgaatc	19980
ctattcgaga atgtttttgt caaagatagt ggcgattttg aaccaagaa aacatttaaa	20040
aatcagtat ccggttacgt tcatgcaa atgaagtgg ctaggatctg attgtaattt	20100
tagacttaaa gagtctctta agattcaatc ctggctgtgt acaaaactac aaataatata	20160
ttttagacta tttggcctta actaaacttc cactcattat ttactgaggt tagagaatag	20220
acttgcaat aaacacattc ccgagaaata ctcgatgcc cataattagt cagagggtat	20280
gccaatcaga tctaagaaca cacattccct caaattttta tgcacatgta atcatagttt	20340
agcacaattc aaaaataatg tagtattaaa gacagaaatt tgtagacttt tttttggcgt	20400
taaaagaaga ctaagtttat acgtacattt tattttaagt ggaaaaccga aattttccat	20460
cgaaatatat gaatttagta tatatatttc tgcaatgtac tattttgcta ttttggcaac	20520

22/143

tttcagtgga ctactacttt attacaatgt gtatggatgc atgagtttga gtatacacat	20580
gtctaaatgc atgcttttga aaacgtaacg gaccacaaaa gaggatccat acaaatacat	20640
ctcatagctt cctccattat tttccgacac aaacagagca ttttacaaca attaccaaca	20700
acaacaaaca acaaacaaca ttacaattac atttacaatt accataccat ggaatttgct	20760
caacctctcg ttgctatggc tcaagagcag tacgctgcta tcgatgctgt tgttgctcct	20820
gctatcttct ctgctaccga ctctattgga tggggactca agcctatctc ttctgctact	20880
aaggatctcc ctctcgttga atctcctacc cctcttatcc tttctctcct cgcttacttc	20940
gctatcgttg gttctggact cgtttaccgt aaagtgttcc ctagaaccgt taagggacag	21000
gatcctttcc ttctcaaggc tcttatgctc gtcacaacg ttttccttat cggactcagc	21060
ctttacatgt gcctcaagct cgtttacgag gcttacgtga acaagtactc cttctgggga	21120
aacgcttaca accctgctca aaccgagatg gctaagggtga tctggatctt ctacgtgtcc	21180
aagatctacg agttcatgga caccttcatc atgcttctca agggaaacgt taaccaggtt	21240
tccttcctcc atgtttacca ccacggatct atctctggaa tctggtggat gatcacttat	21300
gctgctccag gtggagatgc ttacttctct gctgctctca actcttgggt tcatgtgtgc	21360
atgtacacct actacttcat ggctgctgtt cttcctaagg acgaaaagac caagagaaaag	21420
tacctttggt ggggaagata ccttaccag atgcaaatgt tccagttctt catgaacctt	21480
ctccaggctg ttacctcct ctactcttct tctccttacc ctaagttcat tgctcaactc	21540
ctcgttgttt acatggttac cctcctcatg cttttcggaa acttctacta catgaagcac	21600
cacgcttcta agtgataagg gccgccgcca tgtgacagat cgaaggaaga aagtgtata	21660
agacgactct cactactcga tcgctagtga ttgtcattgt tatatataat aatgttatct	21720
ttcacaactt atcgtaatgc atgtgaaact ataacacatt aatcctactt gtcatatgat	21780
aacactctcc ccatttaaaa ctcttgtaaa tttaaagata taagattctt taaatgatta	21840
aaaaaaatat attataaatt caatcactcc tactaataaa ttattaatta ttattttattg	21900
attaaaaaaa tactttatact aatttagtct gaatagaata attagattct agcctgcagg	21960
gcggccgcgg atcccatgga gtcaaagatt caaatagagg acctaacaga actcgccgta	22020
aagactggcg aacagttcat acagagtctc ttacgactca atgacaagaa gaaaatcttc	22080
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca	22140
gaagacaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga	22200
ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc	22260
tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt	22320
gggtcccaaag atggaccccc acccacgagg agcatcgttg aaaaagaaga cgttccaacc	22380
acgtcttcaa agcaagtgga ttgatgtgat atctccactg acgtaaggga tgacgcacaa	22440
tcccactatc cttcgcaaga cccttcctct atataaggaa gttcatttca tttggagaga	22500
acacggggga ctgaattaaa tatgagccct gagaggcgtc ctgttgaaat cagacctgct	22560
actgctgctg atatggctgc tgtttgtgat atcgtgaacc actacatcga gacttctacc	22620

23/143

gttaacttca gaactgagcc tcaaactcct caagagtgga tcgatgatct tgagagactc 22680
 caagatagat acccttggct tgttgctgag gttgaggggtg ttgttgctgg aatcgcttac 22740
 gctggacctt ggaaggctag aaacgcttac gattggactg ttgagtctac cgtttacgtt 22800
 tcacacagac atcagagact tggacttgga tctacccttt acactcacct tctcaagtct 22860
 atggaagctc agggattcaa gtctgttggt gctgttatcg gactccctaa cgatccttct 22920
 gtttagacttc atgaggctct tggatacact gctagaggaa ctcttagagc tgctggatac 22980
 aagcacggtg gatggcatga tggttgattc tggcaaagag atttcgagct tcctgctcct 23040
 cctagacctg ttagaccagt tactcagatc tgaatttgcg tgatcggtca aacatttggc 23100
 aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc atataatttc 23160
 tgttgaatta cgttaagcat gtaataatta acatgtaatg catgacgtta tttatgagat 23220
 gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa aacaaaatat 23280
 agcgcgcaaa ctaggataaa ttatcgcgcg cgggtgtcatc tatgttacta gatcactagt 23340
 gatgtacggt taaaaccacc ccagtacatt aaaaacgtcc gcaatgtgtt attaagttgt 23400
 ctaagcgtca atttgtttac accacaatat atcctgccac cagccagcca acagctcccc 23460
 gaccggcagc tcggcacaaa atcaccactc gatacaggca gcccatcagt cc 23512

<210> 3

<211> 25787

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự nucleotit pGA7- mod_C

<400> 3

tcctgtggtt ggcattgcaca tacaaatgga cgaacggata aaccttttca cgccctttta 60
 aatatccgat tattctaata aacgctcttt tctcttaggt ttacccgcca atatatcctg 120
 tcaaacactg atagttttaa ctgaaggcgg gaaacgacaa tctgctagtg gatctcccag 180
 tcacgacgtt gtaaaacggg cgccccgac tagtaacata gatgacaccg cgcgcgataa 240
 tttatcctag ttgcgcgct atattttgtt ttctatcgcg tattaaatgt ataattgcgg 300
 gactctaadc ataaaaacc atctcataaa taacgtcatg cattacatgt taattattac 360
 gtgcttaacg taattcaaca gaaattatat gataatcatc gcaagaccgg caacaggatt 420
 caatcttaag aaactttatt gccaaatgtt tgaacgatct gcccggaagc ggccaactcg 480
 aaaatttaat taatcatcag tgagccttct cagcctttcc gttaacgtag tagtgctgtc 540
 caactttgtc gaggttgctg aaagtagcct tccaagcacc gtagtaagag agcaccttgt 600
 agttgagtcc ccacttctta gcgaaaggga cgaatcttct tgacacctca ggctgtctga 660
 attgaggcat atcaggggaag agatggtgga taacctggca gttaaggtat ccataagcc 720
 agttaacgta tccacgagaa ggatcgatgt caacggtgtg atcaacagcg tagttaaccc 780
 agctaagggt cttgtcagat ggaacaacag ggaggtgagt gtgagaagta gagaagttag 840

24/143

cgaagaggta catgtaagcg atccagtttc cgaaagtga ccaccagtaa gcaacaggcc	900
aagagtatcc ggtagcaagc ttgataacag cggttctaac aacgtgagaa acgagcatcc	960
aagaagcttc ctctagtttc ttcttcctga gcacctgtct aggatggaga acgtagatcc	1020
agaaagcctg aacgagaagt ccagaagtaa caggaacgaa ggtccaagct tgaagtctag	1080
cccaagctct agagaatccc ctaggtctat tatcctccac agcgggtgtg aagaaagcca	1140
cagcaggagt ggtatcaaga tccatgtcgt gtctaacttt ctgaggggta gcatgggtgt	1200
tgttatgcat ctggttccac atctctccgc tggtagaaag tccgaatccg caagtcatag	1260
cctgaagtct cttatccacg tacacagatc cggtaaagaga gttgtgtcca ccctcatgtt	1320
gaacccatcc acatctagct ccgaagaaag caccgtacac aacgctagca atgataggg	1380
atccagcgta cataagagcg gttccaagag cgaaagtagc aagaagctcc aaaagacgg	1440
aagcaacatg ggtgatagaa ggcttgaaga atccgtccct ctcaagttca gctctccacc	1500
tagcgaaatc ctcaagcata ggagcatcct cagactcaga tctcttgatc tcagcaggtc	1560
tagaaggcaa agctctaagc atcttccaag ccttgaggct acgcatgtga aattctttga	1620
aagcctcagt agcatcagca ccagtgttag caagcatgta gaagatcacg cttccaccag	1680
gatgtttgaa gttggtcacg tcgtactcaa catcctcaac cctaaccat ctagtctcga	1740
aggtagcagc aagttcatga ggctcaaggg tcttaagatc aacaggagcg gtagaagcat	1800
ccttagcatc aagagcctca gcagatgact tagacctggt gagaggagat ctaggagaag	1860
atcttccatc ggtcttagga ggacacatgg cgcgccgatt ttcgagatgg taattgtaaa	1920
tgtaattgta atgttgtttg ttgtttgttg ttgttggtaa ttgttgtaaa attcgagttg	1980
gccgcttccg gggatcctcg tgtatgtttt taatcttggt tgtatcgatg agttttggtt	2040
tgagtaaaga gtgaagcgga tgagttaatt tataggctat aaaggagatt tgcattggcg	2100
tcacgtgtaa taatgcatgc acgcatgtga ttgtatgtgt gtgctgtgag agagaagctc	2160
ttaggtgttt gaagggagtg acaagtggcg aagaaaaaca attctccgcg gctgcatgct	2220
atgtgtaacg tgtagctaat gttctggcat ggcattctat gaacgattct ttttaaaaac	2280
aaggtaaaaa cttaacttca taaaattaaa aaaaaaacg ttactaagt tggtttaaaa	2340
ggggatgaga ggcgccccgc ggaaagcttg ctagccaatt ggggcccaac gttctcgagt	2400
ttttctagaa ggaaactgaa ggcgggaaac gacaatctgc tagtgatctt cccagtcacg	2460
acgttgtaaa acgggcgccc cgcggaaagc ttgcggccgc ggtaccgccc gttcgactca	2520
gatcttccaa ggcctcgtct ccgagtccgc tgcttctcgc cgcgccgatc acttctccgc	2580
cgccaacaag gcttgtagtt aataggaatc attcagggat tgtgattccg ggcagtagta	2640
attaataata tagtattagt atagataata tgtttcgttt gggatctttg gaacgttgct	2700
ctgttccttg ttgttcattt taaagctttt gagggatagt tgcagaactg ttcgggtgatg	2760
cttcatcctc tcaagaacta gatttggtga aagaaacatc catgcatgga tatggaatgt	2820
tgttcttccg attggagatt attttataaa atttaaaatt catgatttaa aaaaacacat	2880
aaaaaccaca aaattcatga tttattgaca atacgataca aaattagcac caccggctac	2940

25/143

tggtcatta cacatttccc cttcccctca ttctcacttt gtggctttat tattattatt	3000
attacatata ttttaccgtt attatttcac gtcacataag cttgttaatt aatcattagt	3060
gagccttctc agcctttccg ttaacgtagt agtgctgtcc caccttatca aggttagaga	3120
aagtagcctt ccaagcaccg tagtaagaga gcaccttgta gttgagtccc cacttcttag	3180
cgaaaggaac gaatcttctg ctaacctcag gctgtctgaa ttgaggcata tcagggaaga	3240
gggtggtgat aacctgacag ttaaggtatc ccataagcca gttcacgtat cctctagaag	3300
gatcgatata aacggtgtga tcaacagcgt agttaacca agaaagggtgc ttatcagatg	3360
gaacaacagg gaggtgagta tgagaagtag agaagtgagc gaaaagggtac atgtaagcga	3420
tccagtttcc gaaagtgaac caccagtaag caacaggcca agagtatcca gtagcaagct	3480
tgataacagc ggttctaaca acatgagaaa cgagcatcca agaagcctct tcgtagttct	3540
tcttacggag aacttgtcta ggggtggagaa cgtagatcca gaaagcttga acaagaagtc	3600
cagaggtaac aggaacgaaa gtccaagctt gaagtctagc ccaagctcta gagaatcctc	3660
taggtctgtt atcctcaaca gcagtgttga agaaagccac agcaggagtg gtatcaagat	3720
ccatatcgtg tctaaccttt tgaggggtag catgggtgctt gttatgcatc tggttccaca	3780
tctcaccaga agtagaaaagt ccgaatccac aagtcatagc ctgaagtctc ttgtccacgt	3840
aaacagatcc ggtaagagag ttatgtccac cctcatgttg aacctatcca catctagctc	3900
cgaagaaagc accgtaaaaca acagaagcaa tgatagggta tccagcgtac ataagagcag	3960
ttccaagagc gaatgtagca agaagctcga gaagtctgta agccacatgg gtgatagaag	4020
gcttgaagaa tccatctctc tcaagctcag cacgccatct agcgaaatcc tcaagcatag	4080
gagcatcctc agactcagat ctcttgatct cagcaggctt agaaggcaaa gctctaagca	4140
tcttccaagc cttgagagaa cgcatgtgga attctttgaa agcctcagta gcatcagcac	4200
cagtgttagc aagcatgtag aagatcacag atccaccagg gtgcttgaag ttagtcacat	4260
cgtactcaac gtcctcaact ctaacctatc tagtctcgaa agtagcagca agctcatgag	4320
gctcaagagt cttaatatca acaggagcag tagaagcatc cttagcatca agagcctcag	4380
cagaagattt agacctggta agtggagatc taggagaaga tcttccatca gtcttaggag	4440
ggcacatggt atggtaattg taaatgtaat tgtaatgttg tttgttgttt gttgttgttg	4500
gtaattgttg taaaattaat taagtgggta tcttttggat ggataagcaa gtagtgatga	4560
tgttctaggt gaagtgatgg ggggtgttta tagcgggaga tggtgaaatg gatggtcgcc	4620
acataagaaa tggaggggaa gggttcttgc gccattcttc agtttgcatg gatgcatggg	4680
tttcattttg taacacgtaa taaggacaat gaagtgcagg tgtctctcaa gtttcagagg	4740
ggatatgtgg acagaagaag aacggcgtat atattgatgg aaatggccat ctagtgtgaa	4800
tctattcggg tgataatact agtgcatttt ggccgttaat cccttcaatt aactgcacaa	4860
acttcagttg agtattgatt atttgattat aggttctgta aacacaatac caagtttatt	4920
tagaggggag acatacaaat agtttcgata taaataatag agtggttaaa cttagttatt	4980

26/143

aaaactatat ataaagtcta aaagttaa	5040
aaaggggtta cattatttct taagagatgt	5100
catcactctg gtttccagtt ctaatcta	5160
agaagtaaat tgatcaatgc cgtcaacca	5220
cggagacccc caaactcaac ctccacccat	5280
caccaatttg gcgtgcatga cacctagggg	5340
aaatgacggg atatcagatg acaggaaacg	5400
agctcaccat caacgagaaa ctaggaaaa	5460
agtggcacag atgggtctgc ctcaacatat	5520
agttgccgga atccggtgtg gagctcacg	5580
tcgctcacc actggacgtc cgtcatgtga	5640
acattctgaa gctccatgag atgagtcctt	5700
atggttgcaa gtccataact tgccgttcga	5760
ggcgagctcg gctgacacct ataaaggcca	5820
tgctttcatg aggcctcggg tcgttatgga	5880
cgttattatg catgacggta agactattac	5940
cttctgcttt gtttgtggat tgattggaca	6000
atactcagag atcgacttcc cttttctcta	6060
ggaagctcaa gctctaaagg cttcacaatg	6120
gaaatcgaag tcaaatcacc caaccgggag	6180
gtatccaatc atcgtgcacg atgctccagg	6240
taggatagag accttaaacg tcggttggtg	6300
gatgtagacg gcggcacgaa gactctgaaa	6360
tgctatccca actgacctct cgttagtccc	6420
tggtcacaag actaggataa tgatgtttcc	6480
ttttagtctt gcttctgcac tttttgttta	6540
tgtccgtatg cttgagggta ccctagtaga	6600
cttaccctat tagttgaaag ttgaaacttt	6660
gtatgtatat gtaatgtgta taaaacgtag	6720
ccgatgagag atgggagcag aactaaagat	6780
ctcttagggt tgaatcctat tcgagaatgt	6840
aaagaaaaca tttaaaaaat cagtatccgg	6900
gatctgattg taattttaga cttaaagagt	6960
aactacaaat aatatatttt agactatttg	7020
tgagggttaga gaatagactt gcgaataaac	7080

27/143

attagtcaga gggatgcca atcagatcta agaacacaca ttccctcaaa ttttaatgca	7140
catgtaatca tagtttagca caattcaaaa ataatgtagt attaaagaca gaaatttgta	7200
gacttttttt tggcgttaaa agaagactaa gtttatacgt acatttttatt ttaagtggaa	7260
aaccgaaatt ttccatcgaa atatatgaat ttagtatata tatttctgca atgtactatt	7320
ttgctatttt ggcaactttc agtggactac tactttatta caatgtgtat ggatgcatga	7380
gtttgagtat acacatgtct aaatgcatgc tttgtaaaac gtaacggacc acaaaagagg	7440
atccatacaa atacatctca tagcttcctc cattattttc cgacacaaac agagcatttt	7500
acaacaatta ccaacaacaa caaacaacaa acaacattac aattacattt acaattacca	7560
taccatggcc tctatcgcta tccctgctgc tcttgctgga actcttgga acgttaccta	7620
caatgtggct aaccctgata tcccagcttc tgagaaagtt cctgcttact tcatgcaggt	7680
tgagtactgg ggacctacta tcggaactat tggatacctc ctcttcatct acttcggaaa	7740
gcgtatcatg cagaacagat ctcaaccttt cggactcaag aacgctatgc tcgtttacaa	7800
cttctaccag accttcttca acagctactg catctacctt ttcgttactt ctcatagggc	7860
tcagggactt aaggtttggg gaaacatccc tgatatgact gctaactctt ggggaatctc	7920
tcaggttatc tggcttact acaacaacaa gtacgttgag cttctcgaca cttcttcat	7980
ggtgatgagg aagaagttcg accagctttc tttccttcac atctaccacc acactcttct	8040
catctggta tggttcgttg ttatgaagct tgagcctggt ggagattgct acttcggatc	8100
ttctgttaac accttcgtgc acgtgatcat gtactcttac tacggacttg ctgctcttgg	8160
agttaactgt ttctggaaga agtacatcac ccagatccag atgcttcagt tctgtatctg	8220
tgcttctcac tctatctaca ccgcttacgt tcagaatacc gctttctggc ttccttacct	8280
tcaactctgg gttatggtga acatgttcgt tctcttcgcc aacttctacc gtaagaggta	8340
caagtctaag ggtgctaaga agcagtgata aggcgcgcgg cgcgccgggc cgccgcatg	8400
tgacagatcg aaggaagaaa gtgtaataag acgactctca ctactcgatc gctagtgtatt	8460
gtcattgtta tatataataa tgttatcttt cacaacttat cgtaatgcat gtgaaactat	8520
aacacattaa tctacttgt catatgataa cactctcccc atttaaaact cttgtcaatt	8580
taaagatata agattcttta aatgattaaa aaaaatatat tataaattca atcactccta	8640
ctaataaatt attaattatt atttattgat taaaaaata cttatactaa tttagtctga	8700
atagaataat tagattctag tctcatcccc ttttaaacca acttagtaaa cgtttttttt	8760
tttaatttta tgaagttaag tttttacctt gtttttaaaa agaatcgttc ataagatgcc	8820
atgccagaac attagctaca cgttacacat agcatgcagc cgcgagaaat tgtttttctt	8880
cgccacttgt cactcccttc aaacacctaa gagcttctct ctcacagcac acacatacaa	8940
tcacatgcgt gcatgcatta ttacacgtga tcgccatgca aatctccttt atagcctata	9000
aattaactca tccgcttcac tctttactca aacaaaaact catcgatata aacaagatta	9060
aaaacataca cgaggatctt ttacaacaat taccaacaac aacaacaac aaacaacatt	9120

28/143

acaattacat ttacaattac cataccatgc ctccaaggga ctcttactct tatgctgctc	9180
ctccttctgc tcaacttcac gaagttgata ctctcaaga gcacgacaag aaagagcttg	9240
ttatcggaga tagggcttac gatgttacca acttcgttaa gagacaccct ggtggaaaga	9300
tcattgctta ccaagttgga actgatgcta ccgatgctta caagcagttc catgttagat	9360
ctgctaaggc tgacaagatg cttaaagtctc ttccttctcg tcctgttcac aagggatact	9420
ctccaagaag ggctgatctt atcgctgatt tccaagagtt caccaagcaa cttgaggctg	9480
agggaaatgtt cgagccttct cttcctcatg ttgcttacag acttgctgag gttatcgcta	9540
tgcatgttgc tgggtctgct cttatctggc atggatacac tttcgctgga atcgctatgc	9600
ttggagttgt tcaggaaga tgtggatggc ttatgcatga ggggtggacat tactctctca	9660
ctggaaacat tgctttcgac agagctatcc aagttgcttg ttacggactt ggatgtggaa	9720
tgtctgggtgc ttgggtggcgt aaccagcata acaagcacca tgctactcct caaaagcttc	9780
agcacgatgt tgatcttgat acccttcctc tcgttgcttt ccatgagaga atcgctgcta	9840
aggttaagtc tcctgctatg aaggcttggc tttctatgca agctaagctt ttcgctcctg	9900
ttaccactct tcttgttgct cttggatggc agctttacct tcatcctaga cacatgctca	9960
ggactaagca ctacgatgag cttgctatgc tcggaatcag atacggactt gttggatacc	10020
ttgctgctaa ctacgggtct ggatacgttc tcgcttgcta ctttctttac gttcagcttg	10080
gagctatgta catcttctgc aacttcgctg tttctcatac tcacctcct gttgttgagc	10140
ctaacgagca tgctacttgg gttgagtacg ctgctaacca cactactaac tgttctccat	10200
cttgggtgtg tgattggtgg atgtcttacc ttaactacca gatcgagcac cacctttacc	10260
cttctatgcc tcaattcaga caccctaaga tcgctcctag agttaagcag cttttcgaga	10320
agcacggact tctactagat gttagaggat acttcgaggc tatggctgat actttcgcta	10380
accttgataa cgttgcccac gctcctgaga agaaaatgca gtaatgagat cgttcaaaca	10440
tttggaata aagtttctta agattgaatc ctggtgccgg tcttgcgatg attatcatat	10500
aatttctgtt gaattacgtt aagcacgtaa taattaacat gtaatgcatg acgttattta	10560
tgagatgggt ttttatgatt agagtcccg c aattatacat ttaatacgcg atagaaaaca	10620
aaatatagcg cgcaaactag gataaattat cgcgcgcggt gtcattctatg ttactagatc	10680
ggtcgattaa aaatcccaat tatatttggt ctaatttagt ttggtattga gtaaaacaaa	10740
ttcgaaccaa accaaaatat aaatatatag tttttatata tatgccttta agacttttta	10800
tagaattttc tttaaaaaat atctagaaat atttgcgact cttctggcat gtaatatttc	10860
gttaaatatg aagtgtcca tttttattaa ctttaaataa ttggttgtag gatcactttc	10920
ttatcaagtg ttactaaaat gcgtcaatct ctttgttctt ccatattcat atgtcaaaat	10980
ctatcaaaat tcttatatat ctttttcgaa tttgaagtga aatttcgata atttaaaatt	11040
aaatagaaca tatcattatt taggtatcat attgattttt atacttaatt actaaatttg	11100
gttaactttg aaagtgtaca tcaacgaaaa attagtcaaa cgactaaaat aaataaatat	11160
catgtgttat taagaaaatt ctctataag aatattttta tagatcatat gtttgtaaaa	11220

29/143

aaaattaatt tttactaaca catatatatta cttatcaaaa atttgacaaa gtaagattaa	11280
aataatattc atctaacaaa aaaaaaacca gaaaatgctg aaaacccggc aaaaccgaac	11340
caatccaaac cgatatagtt ggtttggttt gattttgata taaaccgaac caactcggtc	11400
catttgcacc cctaatacata atagctttaa tttttcaaga tattattaag ttaacgttgt	11460
caatatcctg gaaattttgc aaaatgaatc aagcctatat ggctgtaata tgaatttaaa	11520
agcagctcga tgtggtggtg atatgtaatt tacttgattc taaaaaata tccaagtat	11580
taataatttc tgctaggaag aaggtttagct acgatttaca gcaaagccag aatacaaaga	11640
accataaagt gattgaagct cgaaatatac gaaggaacaa atatttttaa aaaaatacgc	11700
aatgacttgg aacaaaagaa agtgatatac tttttgttct taaacaagca tcccctctaa	11760
agaatggcag ttttcctttg catgtaacta ttatgctccc ttcgttaca aaattttgga	11820
ctactattgg gaacttcttc tgaaaatagt gatagaacct acacgagcat gtgctttcca	11880
tttaatttta aaaaccaaga aacatacata cataacattc catcagcctc tctctctttt	11940
tattacggtt aatgacttaa aacacatctt attatcccat ccttaacacc tagcagtgtc	12000
tttatacgat ctcatcgatc accacttcaa aaccatgcag actgctgctg cccctggagc	12060
tggcatcggc taggctgggt gccgcactgt cccggaaggt ccctagcgac ttgtttagat	12120
tgatgggacc acctctcaac ttcctgctgc tgtccctgct gctggatgtc ctgcctcatc	12180
tggccgattg cacgctccag tcccctgcat gtgcactcgc tcctcaattg cttaagatca	12240
tcgcagcagc tatcgaagtg ctggctctgt tgccctcctc cacggccttg gttgtagtag	12300
tagctgccgc cgcccttctg gactttttcc cacaggaacc gccgaataat tcgatagaac	12360
cacacgagca tgtgctttca tttattttta aaaccaagaa acatacataa catttcatca	12420
gcctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctttatta	12480
cagctgttac actaacttaa aacacattca tctcattatt attattatta tccatcctta	12540
acacctagca gtgtctttgt acgatctcat aatcgatcac cccttcatca ggtatcctta	12600
ggcttcactc caacgttggt gcagttacgg aacatgtaca caccatcatg gttctcaacg	12660
aactggcaag atctccaagt tttccaaagg ctaaccaca tgttctcatc ggtgtgtctg	12720
tagtgctctc ccataacttt cttgatgcac tcggtagctt ctctagcatg gtagaatggg	12780
atccttgaaa cgtagtgatg gagcacatga gtctcgatga tgtcatggaa gatgattccg	12840
aggattccga actctctatc gatagtagca gcagcaccct tagcgaaagt ccactcttga	12900
gcatcgtaat gaggcataga agaatcgggt tgctgaagga aggtaacgaa aacaagccag	12960
tggttaacaa ggatccaagg acagaacat gtgatgaaag taggccagaa tccgaaaacc	13020
ttgtaagcgg tgtaaacaga agtgagggtg gcaaggattc caagatcaga aagaacgatg	13080
taccagtagt ctttcttatc gaaaacaggg ctagaaggcc agtagtgaga cttgaagaac	13140
ttagaaacac cagggttaagg ttgtccagta gcgttagtag caaggtaaag agaaagtcct	13200
ccaagctggt ggaacaagag agcgaaaaca gagtagatag gagtttcctc agcgatatcg	13260

30/143

tgaaggctgg taacttggtg cttctctttg aattcctcgg cggtgtaagg aacgaaaacc	13320
atatctctgg tcatgtgtcc agtagcctta tgggtgcttag catgagagaa cttccagctg	13380
aagtaaggaa ccataacaag agagtggaga acccatccaa cggtatcgtt aacccatccg	13440
tagttagaga aagcagaatg tccacactca tgtccaagga tccagattcc gaatccgaaa	13500
caagagatag agaacacgta agcagaccaa gcagcgaatc taaggaattc gttagggaga	13560
agagggatgt aggtaagtcc aacgtaagcg atagcagaga tagccacgat atctctcacc	13620
acgtaagaca tagacttcac gagagatctc tcgtaacagt gcttagggat agcgtcaagg	13680
atatccttga tgggtgtaatc tggcaccttg aaaacgtttc cgaaggatc gatagcggtc	13740
ttttgctgct tgaaagatgc aacgtttcca gaacgcctaa cggtccttagt agatccctca	13800
aggatctcag atccagacac ggtaacctta gacatggat ggtaattgta aatgtaattg	13860
taatgttggt tgttggttgt tgttggttgg aattgttgta aaatttttgg tggtgattgg	13920
ttctttaagg tgtgagagt agttgtgagt tgtgtggtgg gtttggtgag attggggatg	13980
gtgggtttat atagtggaga ctgaggaatg gggtcgtgag tgttaacttt gcatgggcta	14040
cacgtgggtt cttttgggct tacacgtagt attattcatg caaatgcagc caatacatat	14100
acggtatttt aataatgtgt gggaatacaa tatgccgagt attttactaa ttttggaat	14160
gacaagtgtta catttggtt atcttacttg gcctctcttg ctttaatttg gattattttt	14220
attctcttac ctggccgtt catattcaca tccctaaagg caagacagaa ttgaatgggtg	14280
gccaaaaatt aaaacgatgg atatgacct catagtgtag gatcaattaa cgtcgaagga	14340
aaatactgat tctctcaagc atacggacaa gggtaaataa catagtcacc agaacataat	14400
aaacaaaaag tgcagaagca agactaaaaa aattagctat ggacattcag gttcatattg	14460
gaaacatcat tctctagtc ttgtgacct ctttctctt gctctagtgt agaggccttg	14520
ggactaacga gaggtcagtt gggatagcag atccttatcc tggactagcc tttctggtgt	14580
ttcagagtct tcgtgccgcc gtctacatct atctccatta ggtctgaaga tgactcttca	14640
caccaacgac gtttaaggct tctatcctac tctagcttg caatacctgg cttgcaatac	14700
ctggagcatc gtgcacgatg attggatact gtggaggagg agtgtttgct gatttagagc	14760
tcccggttgg gtgatttgac ttcgatttca gtttaggctt gttgaaattt ttcaggttcc	14820
attgtgaagc ctttagagct tgagcttcct tccatgttaa tgccttgatc gaatactcct	14880
agagaaaagg gaagtcgatc tctgagtatt gaaatcgaag tgcacatttt ttttcaacgt	14940
gtccaatcaa tccacaaaca aagcagaaga caggtaatct ttcatactta tactgacaag	15000
taatagtctt accgtcatgc ataataacgt ctcgttcctt caagaggggt tttccgacat	15060
ccataacgac ccgaagcctc atgaaagcat tagggaagaa cttttggttc ttcttgatc	15120
ggcctttata ggtgtcagcc gagctcgcca attcccgtcc gactggctcc gcaaaatatt	15180
cgaacggcaa gttatggact tgcaaccata actccacggt attgagcagg acctattgtg	15240
aagactcatc tcatggagct tcagaatgtg gttgtcagca aaccaatgac cgaaatccat	15300
cacatgacgg acgtccagt ggtgagcgaa acgaaacagg aagcgcttat ctttcagagt	15360

31/143

cgtgagctcc acaccggatt ccggcaacta cgtgttgggc aggcttcgcc gtattagaga	15420
tatgttgagg cagaccatc tgtgccactc gtacaattac gagagtgtt ttttttgtga	15480
ttttcctagt ttctcgttga tggtagctc atattctaca tcgtatggc tctcaacgctc	15540
gtttcctgtc atctgatatc ccgtcatttg catccacgtg cgccgcctcc cgtgccaagt	15600
ccctaggtgt catgcacgcc aaattggtgg tggtagcggc tgccctgtgc ttcttaccga	15660
tgggtggagg ttgagtttg ggtctccgc ggcgatggtg gtgggttgac ggtttgggtg	15720
gggttgacgg cattgatcaa ttacttctt gcttcaaatt ctttggcaga aaacaattca	15780
ttagattaga actggaacc agagtgatga gacggattaa gtcagattcc aacagagtta	15840
catctcttaa gaaataatgt aacccttta gactttatat atttgcaatt aaaaaataa	15900
tttaactttt agactttata tatagtttta ataactaagt ttaaccactc tattatttat	15960
atcgaaacta tttgtatgtc tcccctctaa ataaacttgg tattgtgttt acagaaccta	16020
taatcaaata atcaatactc aactgaagt tgtgcagtta attgaaggga ttaacggcca	16080
aaatgcacta gtattatcaa ccgaatagat tcacactaga tggccatttc catcaatatc	16140
atcgccgttc ttcttctgtc cacatatccc ctctgaaact tgagagacac ctgcacttca	16200
ttgtccttat tacgtgttac aaaatgaaac ccatgcatcc atgcaaactg aagaatggcg	16260
caagaaccct tcccctccat ttcttatgtg gcgaccatcc atttcaccat ctcccgtat	16320
aaaacacccc catcacttca cctagaacat catcactact tgcttatcca tccaaaagat	16380
accactttt acaacaatta ccaacaaca caaacaaca acaacattac aattacattt	16440
acaattacca taccatgcca cctagcgctg ctaagcaa at gggagcttct actggtgttc	16500
atgctggtgt tactgactct tctgctttca ccagaaagga tgttgctgat agacctgatc	16560
tcaccatcgt tggagattct gtttacgatg ctaaggcttt cagatctgag catcctggtg	16620
gtgctcattt cgtttctttg ttcggaggaa gagatgctac tgaggctttc atggaatacc	16680
atagaagggc ttggcctaag tctagaatgt ctagattcca cgttggatct cttgcttcta	16740
ctgaggaacc tgttgctgct gatgagggat accttcaact ttgtgctagg atcgctaaga	16800
tggtagccttc tgtttcttct ggattcgctc ctgcttctta ctgggttaag gctggactta	16860
tccttgatc tgctatcgct cttgaggctt acatgcttta cgctggaaag agacttctcc	16920
cttctatcgt tcttgatgg cttttcgctc ttatcggctt taacatccag catgatgcta	16980
accatggtgc tttgtctaag tctgcttctg ttaaccttgc tcttgactt tgtcaggatt	17040
ggatcggagg atctatgatc ctttggcttc aagagcatgt tgttatgcac cacctccaca	17100
ctaacgatgt tgataaggat cctgatcaaa aggctcacgg tgctcttaga ctcaagccta	17160
ctgatgcttg gtcacctatg cattggcttc agcatcttta cttttgcct ggtgagacta	17220
tgtacgcttt caagcttttg ttcctcgaca tctctgagct tgttatgtgg cgttgggagg	17280
gtgagcctat ctctaagctt gctggatacc tctttatgcc ttctttgctt ctcaagctta	17340
ccttctgggc tagattcgtt gctttgcctc ttaccttgc tccttctggt catactgctg	17400

32/143

tgtgtatcgc tgctactggt atgactggat ctttctacct cgctttcttc ttcttcatct	17460
cccacaactt cgaggggtgtt gcttctgttg gacctgatgg atctatcact tctatgacta	17520
gaggtgctag cttccttaag agacaagctg agacttcttc taacgttgga ggacctcttc	17580
ttgctactct taacggtgga ctcaactacc aaattgagca tcacttggtc cctagagttc	17640
accatggatt ctaccctaga cttgctcctc ttgttaaggc tgagcttgag gctagaggaa	17700
tcgagtacaa gcactaccct actatctggt ctaaccttgc ttctaccctc agacatatgt	17760
acgctcttgg aagaaggcct agatctaagg ctgagtaatg acaagcttat gtgacgtgaa	17820
ataataacgg taaaatatat gtaataataa taataataaa gccacaaagt gagaatgagg	17880
ggaaggggaa atgtgtaatg agccagtagc cggtggtgct aattttgtat cgtattgtca	17940
ataaatcatg aattttgtgg tttttatgtg tttttttaaa tcatgaattt taaattttat	18000
aaaataatct ccaatcgga gaacaacatt ccataccat gcatggatgt ttctttaccc	18060
aaatctagtt cttgagagga tgaagcatca ccgaacagtt ctgcaactat ccctcaaaag	18120
ctttaaaatg aacaacaagg aacagagcaa cgttccaaag atcccaaacg aaacatatta	18180
tctatactaa tactatatta ttaattacta ctgcccggaa tcacaatccc tgaatgattc	18240
ctattaacta caagccttgt tggcggcgga gaagtgatcg gcgcggcgag aagcagcgga	18300
ctcggagacg aggccttgg agatctgagt cgaacgggca gaatcagtat tttccttcga	18360
cgtaattga tcctacacta ttaggtcat atccatcggt ttaatttttg gccaccattc	18420
aattctgtct tgcctttagg gatgtgaata tgaacggcca aggtaagaga ataaaaataa	18480
tccaaattaa agcaagagag gccaagtaag ataatccaaa tgtacacttg tcattgccaa	18540
aattagtaaa atactcggca tattgtattc ccacacatta ttaaaataacc gtatatgtat	18600
tggctgcatt tgcatgaata atactacgtg taagcccaaa agaaccacg tgtagcccat	18660
gcaaagttaa cactcacgac cccattcctc agtctccact atataaacc accatcccca	18720
atctcacaa acccaccaca caactcaca ctcactctca caccttaaag aaccaatcac	18780
cacaaaaat tttaaca aa ttaccaaaa caacaacaa caacaacat tacaattaca	18840
tttacaatta ccataccatg agcgtgtgta ccgttactgg atctgaccc aagaacagag	18900
gatcttctag caacaccgag caagagggtc caaaagttgc tatcgatacc aacggaaacg	18960
tgttctctgt tcctgatttc accatcaagg acatccttgg agctatccct catgagtgtt	19020
acgagagaag attggctacc tctctctact acgtgttcag agatatcttc tgcattgctta	19080
ccaccggata ccttaccat aagatccttt accctctcct catctcttac acctctaaca	19140
gcatcatcaa gttcactttc tgggcccttt acacttacgt tcaaggactt ttcggaaccg	19200
gaatctgggt tctcgctcat gagtgtggac atcaagcttt ctctgattac ggaatcgtga	19260
acgatttcgt tggatggacc cttcactctt accttatggt tccttacttc agctggaagt	19320
actctcatgg aaagcaccat aaggctactg gacacatgac cagagatatg gttttcggtc	19380
ctgccaccaa agaggaattc aagaagtcta ggaacttctt cggtaacctc gctgagtact	19440
ctgaggattc tccacttaga accctttacg agcttcttgt tcaacaactt ggaggatgga	19500

tcgcttacct cttcggttaac gttacaggac aaccttaccc tgatgttcct tcttggaat	19560
ggaaccactt ctggcttacc tctccacttt tcgagcaaag agatgctctc tacatcttcc	19620
tttctgatct tggaatcctc acccagggaa tcgttcttac tctttggtac aagaaattcg	19680
gaggatggtc ctttttcac aactgggttcg ttccttacat ctgggttaac cactggctcg	19740
ttttcatcac attccttcag cacactgatc ctactatgcc tcattacaac gctgaggaat	19800
ggactttcgc taagggtgct gctgctacta tcgatagaaa gttcggattc atcggacctc	19860
acatcttcca tgatatcatc gagactcatg tgcttcacca ctactgttct aggatcccat	19920
tctacaacgc tagacctgct tctgaggcta tcaagaaagt tatgggaaag cactacaggt	19980
ctagcgacga gaacatgtgg aagtcacttt ggaagtcttt caggctcttg caatacgttg	20040
acggtgataa cgggtgttctc atgttccgta acatcaacaa ctgcggagtt ggagctgctg	20100
agaagtaatg aaggggtgat cgattatgag atcgtaacaa gacactgcta ggtgttaagg	20160
atggataata ataataataa tgagatgaat gtgttttaag ttagtgtaac agctgtaata	20220
aagagagaga gagagagaga gagagagaga gagagagaga gagagagaga gaggctgatg	20280
aaatgttatg tatgtttctt gggttttaaa ataaatgaaa gcacatgctc gtgtggttct	20340
atcgaattat tcggcggttc ctgtgggaaa aagtccagaa gggccgccgc agctactact	20400
acaaccaagg ccgtggagga gggcaacaga gccagcactt cgatagctgc tgcgatgatc	20460
ttaagcaatt gaggagcgag tgcacatgca ggggactgga gcgtgcaatc ggccagatga	20520
ggcaggacat ccagcagcag ggacagcagc aggaagttga gaggtggtcc catcaatcta	20580
aacaagtcgc tagggacctt ccgggacagt gcggcaccca gcctagccga tgccagctcc	20640
aggggcagca gcagtctgca tggttttgaa gtggtgatcg atgagatcgt ataaagacac	20700
tgctaggtgt taaggatggg ataataagat gtgttttaag tcattaaccg taataaaaag	20760
agagagaggc tgatggaatg ttatgtatgt atgtttcttg gtttttaaaa ttaaatggaa	20820
agcacatgct cgtgtgggtt ctatctcgat taaaaatccc aattatattt ggtctaattt	20880
agtttggtat tgagtaaaac aaattcgaac caaaccaaaa tataaatata tagtttttat	20940
atatatgcct ttaagacttt ttatagaatt ttcttttaaaa aatatctaga aatatttgcg	21000
actcttctgg catgtaatat ttcgttaaat atgaagtgc ccatTTTTTat taactttaaa	21060
taattgggtg tacgatcact ttcttatcaa gtgttactaa aatgcgtcaa tctctttgtt	21120
cttccatatt catatgtcaa aatctatcaa aattcttata tatctttttc gaatttgaag	21180
tgaaatttcg ataatttaaa attaaataga acatatcatt atttaggtat catattgatt	21240
tttatactta attactaaat ttgggttaact ttgaaagtgt acatcaacga aaaattagtc	21300
aaacgactaa aataaataaa tatcatgtgt tattaagaaa attctcctat aagaatattt	21360
taatagatca tatgtttgta aaaaaatta atttttacta acacatatat ttacttatca	21420
aaaatttgac aaagtaagat taaaataata ttcatctaac aaaaaaaaaa ccagaaaatg	21480
ctgaaaaccc ggcaaaaccg aaccaatcca aaccgatata gttggtttgg tttgattttg	21540

34/143

atataaaccg aaccaactcg gtccatttgc acccctaadc ataatagctt taatatttca	21600
agatattatt aagttaacgt tgtcaatadc ctggaaattt tgcaaaatga atcaagccta	21660
tatggctgta atatgaattt aaaagcagct cgatgtggtg gtaatatgta atttacttga	21720
ttctaaaaaa atatcccaag tattaataat ttctgctagg aagaagggtta gctacgattt	21780
acagcaaagc cagaatacaa agaaccataa agtgattgaa gctcgaaata tacgaaggaa	21840
caaataattt taaaaaata cgcaatgact tggaacaaaa gaaagtgata tattttttgt	21900
tcttaaacaa gcatccctc taaagaatgg cagttttcct ttgcatgtaa ctattatgct	21960
cccttcgtta caaaaattt ggactactat tgggaacttc ttctgaaaat agtcctgcag	22020
gctagtagat tggttggtt gtttccatgt accagaaggc ttaccctatt agttgaaagt	22080
tgaaactttg ttccctactc aattcctagt tgtgtaaatg tatgtatatg taatgtgtat	22140
aaaacgtagt acttaaata ctaggagtg ttcttgagac cgatgagaga tgggagcaga	22200
actaaagatg atgacataat taagaacgaa ttgaaaggc tcttaggtt gaatcctatt	22260
cgagaatgtt ttgtcaaag atagtggcga tttgaacca aagaaaacat ttaaaaaatc	22320
agtatccggt tacgttcatt caaatagaaa gtggtctagg atctgattgt aatttttagac	22380
ttaaagagtc tcttaagatt caatcctggc tgtgtacaaa actacaaata atatatttta	22440
gactatttgg ccttaactaa acttcactc attatttact gaggttagag aatagacttg	22500
cgaataaaca cattcccgag aaatactcat gatcccataa ttagtcagag ggtatgccaa	22560
tcagatctaa gaacacacat tccctcaaat tttaatgcac atgtaatcat agtttagcac	22620
aattcaaaaa taatgtagta ttaaagacag aaatttgtag actttttttt ggcgttaaaa	22680
gaagactaag ttatatacgt cattttattt taagtggaaa accgaaattt tccatcgaaa	22740
tatatgaatt tagtatatat atttctgcaa tgtactattt tgctattttg gcaactttca	22800
gtggactact actttattac aatgtgtatg gatgcatgag ttgagtata cacatgtcta	22860
aatgcatgct ttgtaaaacg taacggacca caaaagagga tccatacaaa tacatctcat	22920
agcttcctcc attattttcc gacacaaaca gagcatttta caacaattac caacaacaac	22980
aaacaacaaa caacattaca attacattta caattaccat accatggaat ttgctcaacc	23040
tctcgttgct atggctcaag agcagtagc tgctatcgat gctgttggtg ctctgctat	23100
cttctctgct accgactcta ttggatggg actcaagcct atctcttctg ctactaagga	23160
tctccctctc gttgaatctc ctaccctct tacccttct ctctcgctt acttcgctat	23220
cgttggttct ggactcgtt accgtaaagt gttccctaga accgttaagg gacaggatcc	23280
tttccttctc aaggctctta tgctcgctca caacgtttc cttatcggac tcagccttta	23340
catgtgcctc aagctcgtt acgaggctta cgtgaacaag tactccttct ggggaaacgc	23400
ttacaaccct gctcaaaccg agatggctaa ggtgatctgg atcttctacg tgtccaagat	23460
ctacgagttc atggacacct tcatcatgct tctcaaggga aacgttaacc aggtttcctt	23520
cctccatgtt taccaccacg gatctatctc tggaatctgg tggatgatca cttatgctgc	23580
tccaggtgga gatgcttact tctctgctgc tctcaactct tgggttcatt tgtgcatgta	23640

35/143

cacctactac ttcattggctg ctgttcttcc taaggacgaa aagaccaaga gaaagtacct	23700
ttggtgggga agatacctta cccagatgca aatgttccag ttcttcatga accttctcca	23760
ggctgtttac ctctctact cttcttctcc ttaccctaag ttcattgctc aactcctcgt	23820
tgttttacatg gttaccctcc tcatgctttt cggaacttc tactacatga agcaccacgc	23880
ttctaagtga taagggccgc cgccatgtga cagatcgaag gaagaaaagtg taataagacg	23940
actctcacta ctcgatcgct agtgattgtc attgttatat ataataatgt tatctttcac	24000
aacttatcgt aatgcatgtg aaactataac acattaatcc tacttgtcat atgataacac	24060
tctccccatt taaaactctt gtcaatttaa agatataaga ttctttaaat gattaaaaaa	24120
aatatattat aaattcaatc actcctacta ataaattatt aattattatt tattgattaa	24180
aaaaatactt atactaattt agtctgaata gaataattag attctagcct gcagggcggc	24240
cgcgatccc atggagtcaa agattcaaat agaggaccta acagaactcg ccgtaaagac	24300
tggcgaacag ttcatacaga gtctcttacg actcaatgac aagaagaaaa tcttcgtcaa	24360
catggtggag cacgacacac ttgtctactc caaaaatatc aaagatacag ttcagaaga	24420
ccaaagggca attgagactt ttcaacaaag ggtaatatcc ggaaacctcc tcggattcca	24480
ttgcccagct atctgtcact ttattgtgaa gatagtggaa aaggaagggtg gctcctacaa	24540
atgccatcat tgcgataaag gaaaggccat cgttgaagat gcctctgccg acagtgggcc	24600
caaagatgga cccccacca cgaggagcat cgtggaaaaa gaagacgttc caaccacgtc	24660
ttcaaagcaa gtggattgat gtgatatctc cactgacgta agggatgacg cacaatccca	24720
ctatccttcg caagaccctt cctctatata aggaagttca tttcatttg agagaacacg	24780
ggggactgaa ttaaatatga gccctgagag gcgtcctggt gaaatcagac ctgctactgc	24840
tgctgatatg gctgctgttt gtgatatcgt gaaccactac atcgagactt ctaccgttaa	24900
cttcagaact gagcctcaaa ctctcaaga gtggatcgat gatcttgaga gactccaaga	24960
tagataccct tggcttggtg ctgagggtga ggggtgtgtt gctggaatcg cttacgctgg	25020
accttggaag gctagaaacg cttacgattg gactgttgag tctaccgttt acgtttcaca	25080
cagacatcag agacttggaac ttggatctac cttttacact caccttctca agtctatgga	25140
agctcaggga ttcaagtctg ttgttgctgt tatcggactc cctaacgac cttctgttag	25200
acttcatgag gctcttgga acactgctag aggaactctt agagctgctg gatacaagca	25260
cggtggatgg catgatgttg gattctggca aagagatttc gagcttcctg ctctcctag	25320
acctgttaga ccagttactc agatctgaat ttgctgtagc gttcaaacat ttggcaataa	25380
agtttcttaa gattgaatcc tgttgccggt cttgcatga ttatcatata atttctgttg	25440
aattacgtta agcatgtaat aattaacatg taatgcatga cgttatttat gagatgggtt	25500
tttatgatta gagtcccgca attatacatt taatacgca tagaaaacaa aatatagcgc	25560
gcaaactagg ataaattatc gcgcgcggtg tcatctatgt tactagatca ctagtgatgt	25620
acggttaaaa ccacccagc acattaaaaa cgtccgcaat gtgttattaa gttgtctaag	25680

36/143

cgatcaatttg ttacaccac aatatatcct gccaccagcc agccaacagc tccccgaccg 25740
gcagctcggc aaaaatcac cactcgatac aggcagccca tcagtcc 25787

<210> 4
<211> 22824
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit pGA7- mod_D

<400> 4
tcctgtggtt ggcattgcaca tacaaatgga cgaacggata aaccttttca cgccctttta 60
aatatccgat tattctaata aacgctcttt tctcttaggt ttacccgcca atatatcctg 120
tcaaacactg atagttttaa ctgaaggcgg gaaacgacaa tctgctagt gatctcccag 180
tcacgacgtt gtaaaacggg cgctcgtatt aaaaatccca attatatttg gtctaattta 240
gtttggtatt gagtaaaaca aattcgaacc aaacaaaaat ataaatatat agtttttata 300
tatatgcctt taagactttt tatagaattt tcttttaaaa atatctagaa atatttgcca 360
ctcttctggc atgtaatat tctgttaata tgaagtgtc catttttatt aacttttaaat 420
aattggttgt acgatcactt tcttatcaag tgttactaaa atgcgtcaat ctctttgttc 480
ttccatattc atatgtcaaa atctatcaaa attcttatat atctttttcg aatttgaagt 540
gaaatttcga taatttataa ttaaatagaa catatcatta tttagggtatc atattgattt 600
ttatacttaa ttactaaatt tggttactt tgaagtgtg catcaacgaa aaattagtc 660
aacgactaaa ataaataaat atcatgtgtt attaagaaaa ttctcctata agaatatatt 720
aatagatcat atgtttgtaa aaaaaattaa tttttactaa cacatatatt tacttatcaa 780
aaatttgaca aagtaagatt aaaataatat tcatctaaca aaaaaaaaac cagaaaaatgc 840
tgaaaacccg gcaaaaccga accaatccaa accgatatag ttggtttggt ttgattttga 900
tataaacgga accaactcgg tccatttgca cccctaata taatagcttt aatatttcaa 960
gatattatta agttaacgtt gtcaatatcc tggaaatttt gcaaaatgaa tcaagcctat 1020
atggctgtaa tatgaattta aaagcagctc gatgtggtgg taatatgtaa tttacttgat 1080
tctaaaaaaa tatcccaagt attaataatt tctgctagga agaaggtag ctacgattta 1140
cagcaaagcc agaatacaaa gaaccataaa gtgattgaag ctcgaaatat acgaaggaac 1200
aaatattttt aaaaaatac gcaatgactt ggaacaaaag aaagtgatat attttttggt 1260
cttaacaag catccctct aaagaatggc agttttcctt tgcattgtaac tattatgctc 1320
ccttcgttac aaaaattttg gactactatt gggaacttct tctgaaaata gtggcgcccc 1380
gcggaaagct tgctagccaa ttggggccca acgttctcga gtttttctag aaggaaactg 1440
aaggcgggaa acgacaatct gctagtggat ctcccagtca cgacgttgta aaacgggcgc 1500
cccgcggaaa gcttgcggcc gccgatcta gtaacataga tgacaccgcg cgcgataatt 1560
tatcctagtt tgcgcgctat attttgtttt ctatcgcgta ttaaatgtat aattgcggga 1620
ctctaatacat aaaaacccat ctcataaata acgtcatgca ttacatgtta attattacgt 1680

37/143

gcttaacgta attcaacaga aattatatga taatcatcgc aagaccggca acaggattca	1740
atcttaagaa actttattgc caaatgtttg aacgatcggc gcgcctcatt agtgagcctt	1800
ctcagccttt ccgttaacgt agtagtgctg tcccacctta tcaaggttag agaaagtagc	1860
cttccaagca ccgtagtaag agagcacctt gtagttgagt cccacttct tagcgaaagg	1920
aacgaatctt ctgctaacct caggctgtct gaattgaggc atatcagggg agagggtggtg	1980
gataacctga cagttaaggt atcccataag ccagttcacg tatcctctag aaggatcgat	2040
atcaacggtg tgatcaacag cgtagttaac ccaagaaagg tgcttatcag atggaacaac	2100
agggaggtga gtatgagaag tagagaagtg agcgaaaagg tacatgtaag cgatccagtt	2160
tccgaaagtg aaccaccagt aagcaacagg ccaagagtat ccagtagcaa gcttgataac	2220
agcggttcta acaacatgag aaacgagcat ccaagaagcc tcttcgtagt tcttcttacg	2280
gagaacttgt ctaggggtgga gaacgtagat ccagaaagct tgaacaagaa gtccagaggt	2340
aacaggaacg aaagtccaag cttgaagtct agcccaagct ctagagaatc ctctaggtct	2400
gttatcctca acagcagtgt tgaagaaagc cacagcagga gtggtatcaa gatccatatt	2460
gtgtctaacc ttttgagggg tagcatggtg cttgttatgc atctggttcc acatctcacc	2520
agaagtagaa agtccgaatc cacaagtcag agcctgaagt ctcttggtcca cgtaaacaga	2580
tccggtgaaga gagttatgtc caccctcatg ttgaacccat ccacatctag ctccgaagaa	2640
agcaccgtaa acaacagaag caatgatagg gtatccagcg tacataagag cagttccaag	2700
agcgaatgta gcaagaagct cgagaagtct gtaagccaca tgggtgatag aaggcttgaa	2760
gaatccatct ctctcaagct cagcacgcca tctagcgaaa tcctcaagca taggagcatc	2820
ctcagactca gatctcttga tctcagcagg tctagaaggc aaagctctaa gcatcttcca	2880
agccttgaga gaacgcatgt ggaattcttt gaaagcctca gtagcatcag caccagtgtt	2940
agcaagcatg tagaagatca cagatccacc aggggtgcttg aagttagtca catcgactc	3000
aacgtcctca actctaacc atctagtctc gaaagtagca gcaagctcat gaggctcaag	3060
agtcttaaga tcaacaggag cagtagaagc atccttagca tcaagagcct cagcagaaga	3120
tttagacctg gtaagtggag atctaggaga agatcttcca tcagtcttag gagggcacat	3180
ggtatggtaa ttgtaaatgt aattgtaatg ttgtttgttg tttgttggtg ttggttaattg	3240
ttgtaaaaga tcctcgtgta tgtttttaat cttgtttgta tcgatgagtt ttggtttgag	3300
taaagagtga agcggatgag ttaatttata ggctataaag gagatttgca tggcgatcac	3360
gtgtaataat gcatgcacgc atgtgattgt atgtgtgtgc tgtgagagag aagctcttag	3420
gtgtttgaag ggagtgacaa gtggcgaaga aaaacaattc tccgaggctg catgctatgt	3480
gtaacgtgta gctaattgtc tggcatggca tcttatgaac gattcttttt aaaaacaagg	3540
taaaaactta acttcataaa attaaaaaaa aaaacgttta ctaagttggt ttaaaagggg	3600
atgagactag tagattgggt gggttggttc catgtaccag aaggcttacc ctattagttg	3660
aaagttgaaa ctttgttccc tactcaattc ctagttgtgt aaatgtatgt atatgtaatg	3720

38/143

tgtataaaac gtagtactta aatgactagg agtggttctt gagaccgatg agagatggga	3780
gcagaactaa agatgatgac ataattaaga acgaatttga aaggctctta ggtttgaatc	3840
ctattcgaga atgtttttgt caaagatagt ggcgattttg aaccaaagaa aacatttaaa	3900
aaatcagtat ccggttacgt tcatgcaa atgaaagtgg ttaggatctg attgtaattt	3960
tagacttaaa gagtctctta agattcaatc ctggctgtgt aaaaaactac aaataatata	4020
ttttagacta tttagcctta actaaacttc cactcattat ttactgaggt tagagaatag	4080
acttgcaaat aaacacattc ccgagaaata ctcatgatcc cataattagt cagaggggat	4140
gccaatcaga tctaagaaca cacattccct caaattttaa tgcacatgta atcatagttt	4200
agcacaattc aaaaataatg tagtattaaa gacagaaatt ttagactttt ttttggcgt	4260
taaaagaaga ctaagtttat acgtacattt tattttaagt ggaaaaccga aattttccat	4320
cgaaatatat gaatttagta tatatatttc tgcaatgtac tattttgcta ttttggcaac	4380
tttcagtga ctactacttt attacaatgt gtatggatgc atgagtttga gtatacacat	4440
gtctaaatgc atgctttgta aaacgtaacg gaccacaaaa gaggatccat acaaatacat	4500
ctcatagctt cctccattat tttccgacac aaacagagca ttttacaaca attaccaaca	4560
acaacaaaca acaacaaca ttacaattac atttacaatt accataccat ggaattcgcc	4620
cagcctcttg ttgctatggc tcaagagcaa tacgctgcta tcgatgctgt tgttgctcct	4680
gctatcttct ctgctactga ttctatcgga tggggactta agcctatctc ttctgctact	4740
aaggacttgc ctcttggtga gtctcctaca cctctcatcc tttctttgct tgcttacttc	4800
gctatcgttg gatctggact cgtttacaga aagggtttcc ctagaaccgt gaagggacaa	4860
gatccattcc ttttgaaggc tcttatgctt gctcacaacg tgttccttat cggactttct	4920
ctttacatgt gcctcaagct tgtgtacgag gcttacgtta acaagtactc tttctgggga	4980
aacgcttaca accctgctca aactgagatg gctaagggtta tctggatctt ctacgtgagc	5040
aagatctacg agttcatgga taccttcata atgctcctca agggaaatgt taaccagggt	5100
agcttccttc acgtttacca tcacggatct atctctggaa tctggtggat gattacttac	5160
gctgctcctg gtggtgatgc ttacttctct gctgctctta actcttgggt tcacgtgtgt	5220
atgtacacct actattttat ggctgccgtg cttcctaagg acgagaaaac taagagaaag	5280
tacctctggt ggggaagata cttactcaa atgcagatgt tccagttctt catgaacctt	5340
ctccaggctg tttaccttct ctactcttca tctccttacc ctaagtttat cgctcagctc	5400
ctcgtggtgt acatgggttac tcttctcatg cttttcgga acttctacta catgaagcac	5460
cacgctagca agtgatgagg cgcgccgggc cgccgcatg tgacagatcg aaggaagaaa	5520
gtgtaataag acgactctca ctactcgatc gctagtgtt gtcattgtta tatataataa	5580
tgttatcttt cacaacttat cgtaatgcat gtgaaactat aacacattaa tcctacttgt	5640
catatgataa cactctcccc atttaaaact cttgtcaatt taaagatata agattcttta	5700
aatgattaaa aaaaatatat tataaattca atcactccta ctaataaatt attaattatt	5760
atttattgat taaaaaataa cttatactaa tttagtctga atagaataat tagattctag	5820

39/143

tctcatcccc	ttttaaacca	acttagtaaa	cgtttttttt	tttaatttta	tgaagttaag	5880
tttttacctt	gtttttaaaa	agaatcgttc	ataagatgcc	atgccagaac	attagctaca	5940
cgttacacat	agcatgcagc	cgcgagagaat	tgtttttctt	cgccacttgt	cactcccttc	6000
aaacacctaa	gagcttctct	ctcacagcac	acacatacaa	tcacatgcgt	gcatgcatta	6060
ttacacgtga	tcgccatgca	aatctccttt	atagcctata	aattaactca	tccgcttcac	6120
tctttactca	aacaaaaact	catcgataca	aacaagatta	aaaacataca	cgaggatctt	6180
ttacaacaat	taccaacaac	aacaaacaac	aaacaacatt	acaattacat	ttacaattac	6240
cataccatgc	ctccaaggga	ctcttactct	tatgctgctc	ctccttctgc	tcaacttcac	6300
gaagttgata	ctcctcaaga	gcacgacaag	aaagagcttg	ttatcggaga	tagggcttac	6360
gatgttacca	acttcgttaa	gagacaccct	ggtggaaaga	tcattgctta	ccaagttgga	6420
actgatgcta	ccgatgctta	caagcagttc	catgttagat	ctgctaaggc	tgacaagatg	6480
cttaagtctc	ttccttctcg	tcctgttcac	aagggatact	ctccaagaag	ggctgatctt	6540
atcgctgatt	tccaagagtt	caccaagcaa	cttgaggctg	agggaatggt	cgagccttct	6600
cttcctcatg	ttgcttacag	acttgctgag	gttatcgcta	tgcatgttgc	tggtgctgct	6660
cttatctggc	atggatacac	tttcgctgga	atcgctatgc	ttggagttgt	tcaggggaaga	6720
tgtggatggc	ttatgcatga	gggtggacat	tactctctca	ctggaaacat	tgctttcgac	6780
agagctatcc	aagttgcttg	ttacggactt	ggatgtggaa	tgtctggtgc	ttggtggcgt	6840
aaccagcata	acaagcacca	tgctactcct	caaaagcttc	agcacgatgt	tgatcttgat	6900
acccttcctc	tcgttgcttt	ccatgagaga	atcgctgcta	aggttaagtc	tcctgctatg	6960
aaggcttggc	tttctatgca	agctaagctt	ttcgctcctg	ttaccactct	tcttgttgct	7020
cttggatggc	agctttacct	tcacctctaga	cacatgctca	ggactaagca	ctacgatgag	7080
cttgctatgc	tcggaatcag	atacggactt	gttgataacc	ttgctgctaa	ctacggtgct	7140
ggatacgttc	tcgcttgtta	ccttctttac	gttcagcttg	gagctatgta	catcttctgc	7200
aacttcgctg	tttctcatac	tcacctccct	gttggtgagc	ctaacgagca	tgctacttgg	7260
gttgagtacg	ctgctaacca	cactactaac	tgttctccat	cttggtggtg	tgattggtgg	7320
atgtcttacc	ttactacca	gatcgagcac	cacctttacc	cttctatgcc	tcaattcaga	7380
caccctaaga	tcgctcctag	agttaagcag	cttttcgaga	agcacggact	tcactacgat	7440
gttagaggat	acttcgaggc	tatggctgat	actttcgcta	accttgataa	cgttgcccac	7500
gctcctgaga	agaaaatgca	gtaatgagat	cgttcaaaca	tttggaata	aagtttctta	7560
agattgaatc	ctggtgccgg	tcttgcgatg	attatcatat	aatttctggt	gaattacggt	7620
aagcacgtaa	taattaacat	gtaatgcatg	acgttattta	tgagatgggt	ttttatgatt	7680
agagtccgcg	aattatacat	ttaatacgcg	atagaaaaca	aaatatagcg	cgaaactag	7740
gataaattat	cgcgcgcggt	gtcatctatg	ttactagatc	ggtcgattaa	aatcccaat	7800
tatatttggt	ctaatttagt	ttggtattga	gtaaaacaaa	ttcgaaccaa	acaaaaatat	7860

aaatatatag tttttatata tatgccttta agacttttta tagaattttc tttaaaaaat	7920
atctagaaat atttgcgact cttctggcat gtaatatttc gttaaatatg aagtgtcca	7980
tttttattaa ctttaaataa ttggttgtag gatcactttc ttatcaagtg ttactaaaat	8040
gcgtcaatct ctttgttctt ccatattcat atgtcaaaat ctatcaaaat tcttatatat	8100
ctttttcgaa tttgaagtga aatttcgata attttaaatt aaatagaaca tatcattatt	8160
taggtatcat attgattttt atacttaatt actaaatttg gttaactttg aaagtgtaca	8220
tcaacgaaaa attagtcaaa cgactaaaat aaataaatat catgtgttat taagaaaatt	8280
ctcctataag aatattttta tagatcatat gtttgtaaaa aaaattaatt ttactaaca	8340
catatattta cttatcaaaa atttgacaaa gtaagattaa aataatattc atctaacaaa	8400
aaaaaaacca gaaaatgctg aaaaccggc aaaaccgaac caatccaaac cgatatagtt	8460
ggtttggttt gattttgata taaaccgaac caactcggtc ctttgcacc cctaataata	8520
atagctttta tatttcaaga tattattaag ttaacgttgt caatatcctg gaaattttgc	8580
aaaatgaatc aagcctatat ggctgtaata tgaatttaaa agcagctcga tgtggtggtg	8640
atatgtaatt tacttgattc taaaaaata tccaagtat taataatttc tgctaggaag	8700
aaggttagct acgatttaca gcaaagccag aatacaaaga accataaagt gattgaagct	8760
cgaaatatac gaaggaacaa atatttttaaa aaaaatacgc aatgacttgg aacaaaagaa	8820
agtgatatat tttttgttct taaacaagca tcccctctaa agaatggcag ttttcctttg	8880
catgtaacta ttatgtctcc ttcgttacaa aaattttgga ctactattgg gaacttcttc	8940
tgaaaatagt gatagaacct acacgagcat gtgctttcca ttttaatttta aaaaccaaga	9000
aacatacata cataacattc catcagctc tctctctttt tattacggtt aatgacttaa	9060
aacacatctt attatcccat ccttaacacc tagcagtgtc tttatacgat ctcatcgatc	9120
accacttcaa aaccatgcag actgctgctg cccctggagc tggcatcggc taggctgggt	9180
gccgcaactgt cccggaaggc ccctagcgac ttgttttagat tgatgggacc acctctcaac	9240
ttcctgctgc tgtccctgct gctggatgtc ctgcctcatc tggccgattg cacgctccag	9300
tcccctgcat gtgcactcgc tcctcaattg ctttaagatca tcgcagcagc tatcgaagtg	9360
ctggctctgt tgccctctc cacggccttg gttgtagtag tagctgccgc cgcccttctg	9420
gactttttcc cacaggaacc gccgaataat tcgatagaac cacacgagca tgtgctttca	9480
tttattttta aaaccaagaa acatacataa catttcatca gcctctctct ctctctctct	9540
ctctctctct ctctctctct ctctctctct ctctttatta cagctgttac actaaactaa	9600
aacacattca tctcattatt attattatta tccatcctta acacctagca gtgtctttgt	9660
acgatctcat aatcgatcac ccttcatca ggtatcctta ggcttcactc caacgttggt	9720
gcagttacgg aacatgtaca caccatcatg gttctcaacg aactggcaag atctccaagt	9780
tttccaaagg ctaaccaca tgttctcatc ggtgtgtctg tagtgctctc ccataacttt	9840
cttgatgcac tcggtagctt ctctagcatg gtagaatggg atccttgaaa cgtagtgatg	9900
gagcacatga gtctcgatga tgtcatggaa gatgattccg aggattccga actctctatc	9960

gatagtagca gcagcacct tagcgaaagt ccactcttga gcatcgtaat gaggcataga	10020
agaatcggtg tgctgaagga aggtaacgaa aacaagccag tggtaacaa ggatccaagg	10080
acagaacat gtgatgaaag taggccagaa tccgaaaacc ttgtaagcgg tgtaaacaga	10140
agtgagggtg gcaaggattc caagatcaga aagaacgatg taccagtagt ctttcttatc	10200
gaaaacaggg ctagaaggcc agtagtgaga cttgaagaac ttagaaacac cagggttaagg	10260
ttgtccagta gcgttagtag caaggtaaag agaaagtcct ccaagctggt ggaacaagag	10320
agcgaaaaca gagtagatag gagtttcctc agcgatatcg tgaaggctgg taacttggtg	10380
cttctctttg aattcctcgg cgggtgaagg aacgaaaacc atatctctgg tcatgtgtcc	10440
agtagcctta tgggtgcttag catgagagaa cttccagctg aagtaaggaa ccataacaag	10500
agagtggaga acccatccaa cggtatcggt aacctatccg tagttagaga aagcagaatg	10560
tccacactca tgtccaagga tccagattcc gaatccgaaa caagagatag agaacacgta	10620
agcagaccaa gcagcgaatc taaggaattc gttagggaga agagggatgt aggtaagtcc	10680
aacgtaagcg atagcagaga tagccacgat atctctcacc acgtaagaca tagacttcac	10740
gagagatctc tcgtaacagt gcttagggat agcgtcaagg atatccttga tgggtgtaatc	10800
tggcaccttg aaaacgtttc cgaaggatc gatagcggtc ttttgctgct tgaaagatgc	10860
aacgtttcca gaacgcctaa cggtcttagt agatccctca aggatctcag atccagacac	10920
ggtaacctta gacatggtat ggtaattgta aatgtaattg taatgttggt tgttggttgt	10980
tgttggttgg aattgttgta aaatttttgg tggtgattgg ttctttaagg tgtgagagtg	11040
agttgtgagt tgtgtggtgg gtttgggtgag attggggatg gtggggttat atagtggaga	11100
ctgaggaatg gggctcgtgag tgtaacttt gcatgggcta cacgtgggtt cttttgggct	11160
tacacgtagt attattcatg caaatgcagc caatacatat acggtatattt aataatgtgt	11220
gggaatacaa tatgccgagt attttactaa ttttggaat gacaagtgtta catttggtt	11280
atcttacttg gcctctcttg ctttaatttg gattatTTTT attctcttac cttggccgtt	11340
catattcaca tccctaaagg caagacagaa ttgaatggtg gccaaaaatt aaaacgatgg	11400
atatgaccta catagtgtag gatcaattaa cgtcgaagga aaatactgat tctctcaagc	11460
atacggacaa gggtaaataa catagtcacc agaacataat aaacaaaaag tgcagaagca	11520
agactaaaaa aattagctat ggacattcag gttcatattg gaaacatcat tatcctagtc	11580
ttgtgaccat cttcctcct gctctagttg agaggccttg ggactaacga gaggtcagtt	11640
gggtagacag atccttatcc tggactagcc tttctggtgt ttcagagtct tcgtgccgcc	11700
gtctacatct atctccatta ggtctgaaga tgactcttca caccaacgac gtttaaggtc	11760
tctatcctac tcctagcttg caatacctgg cttgcaatac ctggagcatc gtgcacgatg	11820
attggatact gtggaggagg agtgtttgct gatttagagc tcccggttgg gtgatttgac	11880
ttcgatttca gtttaggctt gttgaaattt ttcaggttcc attgtgaagc ctttagagct	11940
tgagcttcct tccatgttaa tgccttgatc gaatactcct agagaaaagg gaagtcgatc	12000

42/143

tctgagtatt gaaatcgaag tgcacat	12060
ttttcaacgt gtccaatcaa tccacaaaca	
aagcagaaga caggtaatct ttcatactta	12120
tactgacaag taatagtctt accgtcatgc	
ataataacgt ctcgttcctt caagaggggt	12180
tttccgacat ccataacgac ccgaagcctc	
atgaaagcat tagggaagaa cttttggttc	12240
ttcttgatcat ggcctttata ggtgtcagcc	
gagctcgcca attcccgtcc gactggctcc	12300
gcaaaatatt cgaacggcaa gttatggact	
tgcaaccata actccacggt attgagcagg	12360
acctattgtg aagactcatc tcatggagct	
tcagaatgtg gttgtcagca aaccaatgac	12420
cgaaatccat cacatgacgg acgtccagt	
ggtgagcgaa acgaaacagg aagcgcctat	12480
ctttcagagt cgtgagctcc acaccggatt	
ccggcaacta cgtgttgggc aggccttcgcc	12540
gtattagaga tatgttgagg cagacccatc	
tgtgccactc gtacaattac gagagtgtgt	12600
ttttttgtga ttttcctagt ttctcgttga	
tggtgagctc atattctaca tcgtatggtc	12660
tctcaacgtc gtttcctgtc atctgatatc	
ccgtcatttg catccacgtg cgccgcctcc	12720
cgtgccaaagt ccctaggtgt catgcacgcc	
aaattggtgg tggtgcgggc tgccctgtgc	12780
ttcttaccga tgggtggagg ttgagtttgg	
gggtctccgc ggcgatggta gtgggttgac	12840
ggtttggtgt gggttgacgg cattgatcaa	
tttacttctt gcttcaaatt ctttggcaga	12900
aaacaattca ttagattaga actggaaacc	
agagtgatga gacggattaa gtcagattcc	12960
aacagagtta catctcttaa gaaataatgt	
aaccccttta gactttatat atttgcaatt	13020
aaaaaataa ttttaactttt agactttata	
tatagtttta ataactaagt ttaaccactc	13080
tattatttat atcgaaacta tttgtatgtc	
tcccctctaa ataaacttgg tattgtgttt	13140
acagaaccta taatcaaata atcaatactc	
aactgaagtt tgtgcagtta attgaaggga	13200
ttaacggcca aaatgcacta gtattatcaa	
ccgaatagat tcacactaga tggccatttc	13260
catcaatatc atcgccgttc ttcttctgtc	
cacatatccc ctctgaaact tgagagacac	13320
ctgcacttca ttgtccttat tacgtgttac	
aaaatgaaac ccatgcatcc atgcaaactg	13380
aagaatggcg caagaaccct tcccctccat	
ttcttatgtg gcgaccatcc atttcacat	13440
ctcccgtat aaaacacccc catcacttca	
cctagaacat catcactact tgcttatcca	13500
tccaaaagat acccactttt acaacaatta	
ccaacaacaa caaacaacaa acaacattac	13560
aattacattt acaattacca taccatgcca	
cctagcgctg ctaagcaaat gggagcttct	13620
actggtgttc atgctggtgt tactgactct	
tctgctttca ccagaaagga tggtgctgat	13680
agacctgatc tcaccatcgt tggagattct	
gtttacgatg ctaaggcttt cagatctgag	13740
catcctgggtg gtgctcattt cgtttctttg	
ttcggaggaa gagatgctac tgaggctttc	13800
atggaatacc atagaagggc ttggcctaag	
tctagaatgt ctagattcca cgttggatct	13860
cttgcttcta ctgaggaacc tggtgctgct	
gatgagggat accttcaact ttgtgctagg	13920
atcgctaaga tgggtgccttc tgtttcttct	
ggattcgctc ctgcttctta ctgggttaag	13980
gctggactta tccttggatc tgctatcgct	
cttgaggctt acatgcttta cgctggaaag	14040
agacttctcc cttctatcgt tcttgatgg	
cttttcgctc ttatcgggtct taacatccag	14100
catgatgcta accatggtgc tttgtctaag	

tctgcttctg ttaaccttgc tcttggactt tgtcaggatt ggatcggagg atctatgatk	14160
ctttggcttc aagagcatgt tgttatgcac cacctccaca ctaacgatgt tgataaggat	14220
cctgatcaaa aggctcacgg tgctcttaga ctcaagccta ctgatgcttg gtcacctatg	14280
cattggcttc agcatcttta ccttttgcct ggtgagacta tgtacgcttt caagcttttg	14340
ttcctcgaca tctctgagct tgttatgtgg cgttgggagg gtgagcctat ctctaagctt	14400
gctggatacc tctttatgcc ttctttgctt ctcaagctta ccttctgggc tagattcggt	14460
gctttgcctc tttaccttgc tccttctgtt catactgctg tgtgtatcgc tgctactggt	14520
atgactggat ctttctacct cgctttcttc ttcttcatct cccacaactt cgagggtggt	14580
gcttctggtg gacctgatgg atctatcact tctatgacta gaggtgctag cttccttaag	14640
agacaagctg agacttcttc taacgttggg ggacctcttc ttgctactct taacggtgga	14700
ctcaactacc aaattgagca tcactgttc cctagagttc accatggatt ctaccctaga	14760
cttgctcctc ttgttaaggc tgagcttgag gctagaggaa tcgagtacaa gcactaccct	14820
actatctggt ctaaccttgc ttctaccctc agacatatgt acgctcttgg aagaaggcct	14880
agatctaagg ctgagtaatg acaagcttat gtgacgtgaa ataataacgg taaaatatat	14940
gtaataataa taataataaa gccacaaagt gagaatgagg ggaaggggaa atgtgtaatg	15000
agccagtagc cgggtggtgct aattttgtat cgtattgtca ataaatcatg aattttgtgg	15060
tttttatgtg tttttttaaa tcatgaattt taaattttat aaaataatct ccaatcggaa	15120
gaacaacatt ccatatccat gcatggatgt ttctttaccc aaatctagtt cttgagagga	15180
tgaagcatca ccgaacagtt ctgcaactat ccctcaaaag ctttaaaatg aacaacaagg	15240
aacagagcaa cgttccaaag atcccaaagc aaacatatta tctatactaa tactatatta	15300
ttaattacta ctgcccggaa tcacaatccc tgaatgattc ctattaacta caagccttgt	15360
tggcggcgga gaagtgatcg gcgcggcgag aagcagcgga ctcggagacg aggcccttga	15420
agatctgagt cgaacgggca gaatcagtat ttctcttcga cgttaattga tcctacacta	15480
tgtaggatcat atccatcggt ttaatttttg gccaccattc aattctgtct tgcctttagg	15540
gatgtgaata tgaacggcca aggtaagaga ataaaaataa tccaaattaa agcaagagag	15600
gccaagtaag ataatccaaa tgtacacttg tcattgccaa aattagtaaa atactcggca	15660
tattgtattc ccacacatta ttaaaataacc gtatatgtat tggctgcatt tgcattgaata	15720
atactacgtg taagcccaaa agaaccacg tgtagcccat gcaaagttaa cactcacgac	15780
cccattcctc agtctccact atataaacc accatcccca atctcaccaa acccaccaca	15840
caactcaca ctcactctca caccttaaag aaccaatcac caccaaaaat ttacaacaa	15900
ttaccaacaa caacaacaa caaacaacat tacaattaca ttacaatta ccataccatg	15960
agcgctgtta ccgttactgg atctgatcct aagaacagag gatcttctag caacaccgag	16020
caagagggtc caaaagttgc tatcgatacc aacggaaacg tgttctctgt tcctgatttc	16080
accatcaagg acatccttgg agctatccct catgagtgtt acgagagaag attggctacc	16140

tctctctact acgtgttcag agatatcttc tgcattgctta ccaccggata ccttaccat	16200
aagatccttt accctctcct catctcttac acctctaaca gcatcatcaa gttcactttc	16260
tgggcccttt acacttacgt tcaaggactt ttcggaaccg gaatctgggt tctcgctcat	16320
gagtgtggac atcaagcttt ctctgattac ggaatcgtga acgatttcgt tggatggacc	16380
cttcactctt accttatggt tccttacttc agctggaagt actctcatgg aaagcaccat	16440
aaggctactg gacacatgac cagagatatg gttttcgttc ctgccaccaa agaggaattc	16500
aagaagtcta ggaacttctt cggtaacctc gctgagtact ctgaggattc tccacttaga	16560
accttttacg agcttcttgt tcaacaactt ggaggatgga tcgcttacct cttcgttaac	16620
gttacaggac aaccttacc tgatgttcct tcttggaat ggaaccactt ctggcttacc	16680
tctccacttt tcgagcaaag agatgctctc tacatcttc tttctgatct tggaatcctc	16740
acctagggaa tcgttcttac tctttggtac aagaaattcg gaggatggtc cttttcatc	16800
aactggttcg ttccttacat ctgggttaac cactggctcg ttttcatcac attccttcag	16860
cacactgac ctactatgcc tcattacaac gctgaggaat ggactttcgc taagggtgct	16920
gctgctacta tcgatagaaa gttcggattc atcggacctc acatcttcca tgatatcatc	16980
gagactcatg tgcttcacca ctactgttct aggatcccat tctacaacgc tagacctgct	17040
tctgaggcta tcaagaaagt tatgggaaag cactacaggt ctagcgacga gaacatgtgg	17100
aagtcacttt ggaagtcttt caggtcttgc caatacgttg acggtgataa cgggtgttctc	17160
atgttccgta acatcaaaa ctgagggtt ggagctgctg agaagtaatg aaggggtgat	17220
cgattatgag atcgtacaaa gacactgcta ggtgttaagg atggataata ataataataa	17280
tgagatgaat gtgttttaag ttagtgtaac agctgtaata aagagagaga gagagagaga	17340
gagagagaga gagagagaga gagagagaga gaggctgatg aaatgttatg tatgtttctt	17400
ggtttttaaa ataatgaaa gcacatgctc gtgtggttct atcgaattat tcggcggttc	17460
ctgtgggaaa aagtccagaa gggccgccgc agctactact acaaccaagg ccgtggagga	17520
gggcaacaga gccagcactt cgatagctgc tgcgatgac ttaagcaatt gaggagcag	17580
tgcacatgca ggggactgga gcgtgcaatc ggccagatga ggcaggacat ccagcagcag	17640
ggacagcagc aggaagttga gaggtggtcc catcaatcta aacaagtcgc tagggacctt	17700
ccgggacagt gcggcaccca gcctagccga tgccagctcc aggggcagca gcagtctgca	17760
tggttttgaa gtggtgatcg atgagatcgt ataaagacac tgctaggtgt taaggatggg	17820
ataataagat gtgttttaag tcattaaccg taataaaaag agagagaggc tgatggaatg	17880
ttatgtatgt atgtttcttg gtttttaaaa ttaaatggaa agcacatgct cgtgtgggtt	17940
ctatctcgat taaaaatccc aattatattt ggtctaattt agtttggtat tgagtaaac	18000
aaattcgaac caaacaaaa tataaatata tagtttttat atatatgcct ttaagacttt	18060
ttatagaatt ttcttttaaaa aatatctaga aatatttgcg actcttctgg catgtaatat	18120
ttcgttaaat atgaagtgct ccatttttat taactttaaa taattggttg tacgatcact	18180
ttcttatcaa gtgttactaa aatgcgtcaa tctctttgtt cttccatatt catatgtcaa	18240

45/143

aatctatcaa aattcttata tatctttttc gaatttgaag tgaaatttcg ataatttaaa	18300
attaaataga acatatcatt atttaggat catattgatt tttatactta attactaaat	18360
ttggtttaact ttgaaagtgt acatcaacga aaaattagtc aaacgactaa aataaataaa	18420
tatcatgtgt tattaagaaa attctcctat aagaatattt taatagatca tatgtttgta	18480
aaaaaatta atttttacta acacatatat ttacttatca aaaatttgac aaagtaagat	18540
taaaataata ttcattctaac aaaaaaaaaa ccagaaaatg ctgaaaaccc ggcaaaaccg	18600
aaccaatcca aaccgatata gttggtttgg tttgattttg atataaacg aaccaactcg	18660
gtccatttgc acccctaatac ataatagctt taatatttca agatattatt aagttaacgt	18720
tgtcaatatc ctggaaattt tgcaaaatga atcaagccta tatggctgta atatgaattt	18780
aaaagcagct cgatgtggtg gtaatatgta atttacttga ttctaaaaaa atatcccaag	18840
tattaataat ttctgctagg aagaaggtta gctacgattt acagcaaagc cagaatacaa	18900
agaaccataa agtgattgaa gctcgaaata tacgaaggaa caaatatttt taaaaaata	18960
cgcaatgact tggaacaaaa gaaagtgata tttttttgt tcttaaaca gcatcccctc	19020
taaagaatgg cagttttcct ttgcatgtaa ctattatgct cccttcgtta caaaaatttt	19080
ggactactat tgggaacttc ttctgaaaat agtcctgcag gctagtagat tggttggttg	19140
gtttccatgt accagaaggc ttaccctatt agttgaaagt tgaaactttg ttccctactc	19200
aattcctagt tgtgtaaatg tatgtatatg taatgtgtat aaaacgtagt acttaaatga	19260
ctaggagtgg ttcttgagac cgatgagaga tgggagcaga actaaagatg atgacataat	19320
taagaacgaa tttgaaaggc tcttaggttt gaatcctatt cgagaatggt tttgtcaaag	19380
atagtggcga ttttgaacca aagaaaacat ttaaaaaatc agtatccggt tacgttcatg	19440
caaatagaaa gtggcttagg atctgattgt aatttttagac ttaaagagtc tcttaagatt	19500
caatcctggc tgtgtacaaa actacaaata atatatttta gactatttgg ccttaactaa	19560
acttcactc attatttact gaggttagag aatagacttg cgaataaaca cattcccag	19620
aaatactcat gatcccataa ttagtcagag ggtatgccaa tcagatctaa gaacacacat	19680
tccctcaaat tttaatgcac atgtaatcat agtttagcac aattcaaaaa taatgtagta	19740
ttaaagacag aaattttag actttttttt ggcgttaaaa gaagactaag tttatacgta	19800
cattttattt taagtggaaa accgaaattt tccatcgaaa tatatgaatt tagtatatat	19860
atttctgcaa tgtactattt tgctattttg gcaactttca gtggactact actttattac	19920
aatgtgtatg gatgcatgag tttgagtata cacatgtcta aatgcatgct ttgtaaaacg	19980
taacggacca caaagagga tccatacaaa tacatctcat agcttcctcc attattttcc	20040
gacacaaaca gagcatttta caacaattac caacaacaac aaacaacaaa caacattaca	20100
attacattta caattaccat accatggcct ctatcgctat ccctgctgct cttgctggaa	20160
ctcttgata cgttacctac aatgtggcta accctgatat cccagcttct gagaaagttc	20220
ctgcttactt catgcagggt gagtactggg gacctactat cggaactatt ggatacctcc	20280

46/143

tcttcatcta cttcggaag cgtatcatgc agaacagatc tcaacctttc ggactcaaga	20340
acgctatgct cgtttacaac ttctaccaga ctttcttcaa cagctactgc atctaccttt	20400
tcgttacttc tcatagggct cagggactta aggtttggg aaacatccct gatatgactg	20460
ctaactcttg gggaatctct caggttatct ggcttcacta caacaacaag tacgttgagc	20520
ttctcgacac cttcttcatg gtgatgagga agaagttcga ccagctttct ttccttcaca	20580
tctaccacca cactcttctc atctggtcat ggttcgttgt tatgaagctt gagcctgttg	20640
gagattgcta cttcggatct tctgttaaca cttcgtgca cgtgatcatg tactcttact	20700
acggacttgc tgctcttgga gttaactggt tctggaagaa gtacatcacc cagatccaga	20760
tgcttcagtt ctgtatctgt gcttctcact ctatctacac cgcttacgtt cagaataccg	20820
ctttctggct tccttacctt caactctggg ttatggtgaa catgttcgtt ctcttcgcca	20880
acttctaccg taagaggtag aagtctaagg gtgctaagaa gcagtataa gggccgccgc	20940
catgtgacag atcgaaggaa gaaagtgtaa taagacgact ctactactc gatcgctagt	21000
gattgtcatt gttatatata ataatgttat ctttcacaac ttatcgtaat gcatgtgaaa	21060
ctataacaca ttaatcctac ttgtcatatg ataacactct cccatttaa aactcttgct	21120
aatttaaaga tataagattc tttaaatgat taaaaaaaa atattataaa ttcaatcact	21180
cctactaata aattattaat tattatttat tgattaaaaa aatacttata ctaatttagt	21240
ctgaatagaa taattagatt ctagcctgca gggcggccgc ggatcccatg gagtcaaaga	21300
ttcaaataga ggacctaa gaactcgccg taaagactgg cgaacagttc atacagagtc	21360
tcttacgact caatgacaag aagaaaatct tcgtcaacat ggtggagcac gacacacttg	21420
tctactcaa aaatatcaa gatacagtct cagaagacca aagggaatt gagacttttc	21480
aacaaaggg aatatccga aacctcctg gattccattg ccagctatc tgtcacttta	21540
ttgtgaagat agtggaaaag gaaggtggct cctacaaatg ccatcattgc gataaaggaa	21600
aggccatcgt tgaagatgcc tctgccgaca gtggtcccaa agatggaccc ccaccacga	21660
ggagcatcgt ggaaaaagaa gacgttccaa ccacgtcttc aaagcaagtg gattgatgtg	21720
atatctccac tgacgtaagg gatgacgcac aatccacta tccttcgcaa gacccttcct	21780
ctatataagg aagttcattt catttgagga gaacacgggg gactgaatta aatatgagcc	21840
ctgagaggcg tcctgttgaa atcagacctg ctactgctgc tgatatggct gctgtttgtg	21900
atatcgtgaa ccactacatc gagacttcta ccgttaactt cagaactgag cctcaaactc	21960
ctcaagagtg gatcgatgat cttgagagac tccaagatag atacccttg cttgttgctg	22020
aggttgagg tggtgttgct ggaatcgctt acgctggacc ttggaaggct agaaacgctt	22080
acgattggac tggtgagtct accgtttacg tttcacacag acatcagaga cttggacttg	22140
gatctaccct ttactcac cttctcaagt ctatggaagc tcagggattc aagtctgttg	22200
ttgtgttat cggactccct aacgatcctt ctgtagact tcatgaggct cttggataca	22260
ctgctagagg aactcttaga gctgctggat acaagcacgg tggatggcat gatgttgat	22320
tctggcaaag agatttcgag cttcctgctc ctctagacc tgtagacca gttactcaga	22380

47/143

tctgaatttg cgtgatcggt caaacatttg gcaataaagt ttcttaagat tgaatcctgt	22440
tgccggtctt gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat	22500
taacatgtaa tgcgatgacgt tatttatgag atgggttttt atgattagag tcccgcgaatt	22560
atacatttaa tacgcatag aaaacaaaat atagcgcgca aactaggata aattatcgcg	22620
cgcggtgtca tctatgttac tagatcacta gtgatgtacg gttaaaacca cccagtaca	22680
ttaaaaacgt ccgcaatgtg ttattaagtt gtctaagcgt caatttgttt acaccacaat	22740
atatcctgcc accagccagc caacagctcc ccgaccggca gctcggcaca aaatcaccac	22800
tcgatacagg cagcccatca gtcc	22824

<210> 5
 <211> 24809
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Trình tự nucleotit pga7- mod_e

<400> 5	
tcctgtgggt ggcgatgcaca tacaaatgga cgaacggata aaccttttca cgccctttta	60
aatatccgat tattctaata aacgctcttt tctcttaggt ttaccgcga atatatcctg	120
tcaaacactg atagttaaaa ctgaaggcgg gaaacgacaa tctgctagtg gatctcccag	180
tcacgacgtt gtaaacggg cgctcgatt aaaaatccca attatatttg gtctaattta	240
gtttggtatt gagtaaaaca aattcgaacc aaacaaaat ataaatatat agtttttata	300
tatatgcctt taagactttt tatagaattt tctttaaaaa atatctagaa atatttgca	360
ctcttctggc atgtaatatt tcgttaata tgaagtgtc catttttatt aactttaaat	420
aattggttgc acgatcactt tcttatcaag tgttactaaa atgcgtcaat ctctttgttc	480
ttccatattc atatgtcaaa atctatcaaa attcttatat atctttttcg aatttgaagt	540
gaaatttcga taatttaaaa ttaaatagaa catatcatta tttagggtatc atattgattt	600
ttatacttaa ttactaaatt tggttaactt tgaaagtgt catcaacgaa aaattagtca	660
aacgactaaa ataaataaat atcatgtgtt attaagaaaa ttctcctata agaatatatt	720
aatagatcat atgtttgtaa aaaaaattaa tttttactaa cacatatatt tacttatcaa	780
aaatttgaca aagtaagatt aaaataatat tcatctaaca aaaaaaaac cagaaaatgc	840
tgaaaacccg gcaaaaccga accaatccaa accgatatag ttggtttgggt ttgattttga	900
tataaacgga accaactcgg tccatttgca cccctaatca taatagcttt aatatttcaa	960
gatattatta agttaacgtt gtcaatatcc tggaaatttt gcaaaatgaa tcaagcctat	1020
atggctgtaa tatgaattta aaagcagctc gatgtggtgg taatatgtaa ttacttgat	1080
tctaaaaaaa tatcccaagt attaataatt tctgctagga agaagggttag ctacgattta	1140
cagcaaagcc agaatacaaa gaaccataaa gtgattgaag ctcgaaatat acgaaggaac	1200
aaatatatttt aaaaaatac gcaatgactt ggaacaaaag aaagtgatat attttttggt	1260

48/143

cttaaacaag catcccctct aaagaatggc agttttcctt tgcatgtaac tattatgctc	1320
ccttcgttac aaaaattttg gactactatt ggggaacttct tctgaaaata gtggcgcccc	1380
gcggaaagct tgctagccaa ttggggccca acgtttctga gtttttctag aaggaaactg	1440
aaggcgggaa acgacaatct gctagtggat ctcccagtca cgacgttgta aaacgggcgc	1500
cccgcggaaa gcttgcggcc gcggtaccgc ccgttcgact cagatcttcc aaggcctcgt	1560
ctccgagtcc gctgcttctc gccgcgccga tcacttctcc gccgccaaca aggcctttag	1620
ttaataggaa tcattcaggg attgtgattc cgggcagtag taattaataa tatagtatta	1680
gtatagataa tatgtttcgt ttgggatcct tggaacgttg ctctgttcct tgttgttcat	1740
tttaaagctt ttgagggata gttgcagaac tgttcgggtga tgcttcatcc tctcaagaac	1800
tagatttggg taaagaaaca tccatgcatg gatattggaat gttgttcttc cgattggaga	1860
ttattttata aaatttaaaa ttcattgattt aaaaaaacac ataaaaacca caaaattcat	1920
gatttattga caatacgata caaaattagc accaccggct actggctcat tacacatttc	1980
cccttccctt cattctcact ttgtggcttt attattatta ttattacata tattttaccg	2040
ttattatttc acgtcacata agcttgtaa ttaatcatta gtgagccttc tcagcctttc	2100
cgттаacgta gtagtgctgt cccaccttat caaggtaga gaaagtagcc ttccaagcac	2160
cgtagtaaga gagcaccttg tagttgagtc cccacttctt agcgaaagga acgaatcttc	2220
tgctaacctc aggcgtgtctg aattgaggca tatcagggaa gaggtggtgg ataacctgac	2280
agttaaggta tcccataagc cagttcacgt atcctctaga aggatcgata tcaacggtgt	2340
gatcaacagc gtagttaacc caagaaagggt gcttatcaga tggaacaaca gggagggtgag	2400
tatgagaagt agagaagtga gcgaaaagggt acatgtaagc gatccagttt ccgaaagtga	2460
accaccagta agcaacaggc caagagtatc cagtagcaag cttgataaca gcggttctaa	2520
caacatgaga aacgagcatc caagaagcct cttcgtagtt cttcttacgg agaacttgtc	2580
taggggtggag aacgtagatc cagaaagctt gaacaagaag tccagaggta acaggaacga	2640
aagtccaagc ttgaagtcta gcccaagctc tagagaatcc tctaggtctg ttatcctcaa	2700
cagcagtgtt gaagaaagcc acagcaggag tggatatcaag atccatatcg tgtctaacct	2760
tttgaggggt agcatggtgc ttgttatgca tctggttcca catctcacca gaagtagaaa	2820
gtccgaatcc acaagtcata gcctgaagtc tcttgccac gtaaacagat ccggttaagag	2880
agttatgtcc accctcatgt tgaaccatc cacatctagc tccgaagaaa gcaccgtaaa	2940
caacagaagc aatgataggg tatccagcgt acataagagc agttccaaga gcgaatgtag	3000
caagaagctc gagaagtctg taagccacat gggtagatga aggcctgaag aatccatctc	3060
tctcaagctc agcacgccat ctagcgaaat cctcaagcat aggagcatcc tcagactcag	3120
atctcttgat ctacgcaggt ctagaaggca aagctctaag catcttccaa gccttgagag	3180
aacgcagtgt gaattctttg aaagcctcag tagcatcagc accagtgtta gcaagcatgt	3240
agaagatcac agatccacca ggggtgcttg agttagtcac atcgactca acgtcctcaa	3300
ctctaacca tctagtctcg aaagtagcag caagctcatg aggcctcaaga gtcttaagat	3360

caacaggagc agtagaagca tccttagcat caagagcctc agcagaagat ttagacctgg	3420
taagtggaga tctaggagaa gatcttccat cagtcttagg agggcacatg gtatggtaat	3480
tgtaaagtga attgtaattgt tgtttggtgt ttggtgtgtg tggtaattgt tgtaaaatta	3540
attaagtggg tatcttttgg atggataagc aagtagtgat gatgttctag gtgaagtgat	3600
gggggtgttt tatagcggga gatggtgaaa tggatggctg ccacataaga aatggagggg	3660
aagggttctt gcgccattct tcagtttgca tggatgcatg ggtttcattt tgtaacacgt	3720
aataaggaca atgaagtgca ggtgtctctc aagtttcaga ggggatattgt ggacagaaga	3780
agaacggcga tgatattgat ggaaatggcc atctagtgtg aatctattcg gttgataata	3840
ctagtgcatt ttggccgtta atcccttcaa ttaactgcac aaacttcagt tgagtattga	3900
ttatttgatt ataggttctg taaacacaat accaagttta tttagagggg agacatacaa	3960
atagtttcga tataaataat agagtgggta aacttagtta ttaaaactat atataaagtc	4020
taaaagttaa attatttttt taattgcaaa tatataaagt cttaaaggggt tacattattt	4080
cttaagagat gtaactctgt tggaatctga cttaatccgt ctcatcactc tggtttccag	4140
ttctaactta atgaattggt ttctgccaaa gaatttgaag caagaagtaa attgatcaat	4200
gccgtcaacc cacaccaaac cgtcaacca ctaccatcgc cgcggagacc cccaaactca	4260
acctccacc atcggtgaaga agcacagggc agcccgacc accaccaatt tggcgtgcat	4320
gacacctagg gacttggcac gggaggcggc gcacgtggat gcaaagacg ggatatcaga	4380
tgacaggaaa cgacgttgag agaccatacg atgtagaata tgagctcacc atcaacgaga	4440
aactaggaaa atcacaaaa aaacaactct cgtaattgta cgagtggcac agatgggtct	4500
gcctcaacat atctctaata cggcgaagcc tgcccaacac gtagttgccg gaatccggtg	4560
tggagctcac gactctgaaa gataggcgct tcctgtttcg tttcgctcac cactggacg	4620
tccgtcatgt gatggatttc ggtcattggt ttgctgacaa ccacattctg aagctccatg	4680
agatgagtct tcacaatagg tcctgtcaa taccgtggag ttatggttgc aagtccataa	4740
cttgccgttc gaatatatttg cggagccagt cggacgggaa ttggcgagct cggctgacac	4800
ctataaaggc catgacaaga agaaccaaaa gttcttcctt aatgctttca tgaggcttcg	4860
ggtcgttatg gatgtcggaa aacctctctt gaaggaacga gacgttatta tgcacgacg	4920
taagactatt acttgtcagt ataagtatga aagattacct gtcttctgct ttgtttgtgg	4980
attgattgga cacgttgaaa aaaaatgtgc acttcgattt caatactcag agatcgactt	5040
cccttttctc taggagtatt cgatcaaggc attaacatgg aaggaagctc aagctctaaa	5100
ggcttcacaa tggaaacctga aaaatttcaa caagcctaaa ctgaaatcga agtcaaata	5160
cccaaccggg agctctaaat cagcaaacac tcctcctcca cagtatcaa tcacgtgca	5220
cgatgctcca ggtattgcaa gccagggtatt gcaagctagg agtaggatag agaccttaaa	5280
cgtcgttggt gtgaagagtc atcttcagac ctaatggaga tagatgtaga cggcggcacg	5340
aagactctga aacaccagaa aggctagtcc aggataagga tctgctatcc caactgacct	5400

50/143

ctcgttagtc ccaaggcctc tcaactagag caggaggaag gatggtcaca agactaggat	5460
aatgatgttt ccaatatgaa cctgaatgtc catagctaatt ttttttagtc ttgcttctgc	5520
actttttgtt tattatgttc tgggtgactat gttatttacc cttgtccgta tgcttgaggg	5580
taccctagta gattggttgg ttggttttcca tgtaccagaa ggcttaccct attagttgaa	5640
agttgaaact ttgtttcccta ctcaattcct agttgtgtaa atgtatgtat atgtaatgtg	5700
tataaaacgt agtacttaaa tgactaggag tggttcctga gaccgatgag agatgggagc	5760
agaactaaag atgatgacat aattaagaac gaatttgaaa ggctccttagg tttgaatcct	5820
attcgagaat gtttttgtca aagatagtgg cgattttgaa ccaaagaaaa catttaaaaa	5880
atcagtatcc ggttacgttc atgcaaatag aaagtgggtct aggatctgat tgtaatttta	5940
gacttaaaga gtctcttaag attcaatcct ggctgtgtac aaaactacaa ataatatatt	6000
ttagactatt tggccttaac taaacttcca ctcatatttt actgagggtta gagaatagac	6060
ttgcgaataa acacattccc gagaaatact catgatccca taattagtca gagggtatgc	6120
caatcagatc taagaacaca cattccctca aattttaatg cacatgtaat catagtttag	6180
cacaattcaa aaataatgta gtattaaaga cagaaatttg tagacttttt tttggcggtta	6240
aaagaagact aagtttatac gtacatttta ttttaagtgg aaaaccgaaa ttttccatcg	6300
aaatatatga atttagtata tatatttctg caatgtacta ttttgctatt ttggcaactt	6360
tcagtggact actactttat tacaatgtgt atggatgcat gagtttgagt atacacatgt	6420
ctaaatgcat gctttgtaaa acgtaacgga ccacaaaaga ggatccatac aaatacatct	6480
catagcttcc tccattattt tccgacacaa acagagcatt ttacaacaat taccaacaac	6540
aacaaacaac aaacaacatt acaattacat ttacaattac cataccatgg cctctatcgc	6600
tatccctgct gctcttgctg gaactcttg atacgttacc tacaatgtgg ctaaccctga	6660
tatcccagct tctgagaaag ttcttgctta cttcatgcag gttgagtact ggggacctac	6720
tatcggaact attggatacc tcctcttcat ctacttcgga aagcgatatca tgcagaacag	6780
atctcaacct ttcggactca agaacgctat gctcgtttac aacttctacc agaccttctt	6840
caacagctac tgcattctacc ttttcgttac ttctcatagg gctcagggaac ttaaggtttg	6900
gggaaacatc cctgatatga ctgctaactc ttgggggaatc tctcagggtta tctggcttca	6960
ctacaacaac aagtacgttg agcttctcga cacttcttc atggtgatga ggaagaagtt	7020
cgaccagctt tctttccttc acatctacca ccacactctt ctcatctggt catggttcgt	7080
tgttatgaag cttgagcctg ttggagattg ctacttcgga tcttctgtta acaccttcgt	7140
gcacgtgatc atgtactctt actacggact tgctgctctt ggagttaact gtttctggaa	7200
gaagtacatc acccagatcc agatgcttca gttctgtatc tgtgcttctc actctatcta	7260
caccgcttac gttcagaata ccgctttctg gcttccttac cttcaactct gggttatggt	7320
gaacatgttc gttctcttcg ccaacttcta ccgtaagagg tacaagtcta agggtgctaa	7380
gaagcagtga taaggcgcg gcgcgcccgg gccgccgcca tgtgacagat cgaaggaaga	7440
aagtgtata agacgactct cactactcga tcgctagtga ttgtcattgt tatatataat	7500

51/143

aatgttatct ttcacaactt atcgtaatgc atgtgaaact ataacacatt aatcctactt	7560
gtcatatgat aacactctcc ccattttaaaa ctcttgtcaa tttaaagata taagattctt	7620
taaatgatta aaaaaaatat attataaatt caatcactcc tactaataaa ttattaatta	7680
ttattttattg attaaaaaaa tactttatact aatttagtct gaatagaata attagattct	7740
agtctcatcc cctttttaaac caacttagta aacgtttttt tttttaattt tatgaagta	7800
agtttttacc ttgtttttta aaagaatcgt tcataagatg ccatgccaga acattagcta	7860
cacgttacac atagcatgca gccgcggaga attgtttttc ttcgccactt gtcactccct	7920
tcaaacacct aagagcttct ctctcacagc acacacatac aatcacatgc gtgcatgcat	7980
tattacacgt gatcgccatg caaatctcct ttatagccta taaattaact catccgcttc	8040
actctttact caaaccaaaa ctcatcgata caaacaagat taaaaacata cacgaggatc	8100
ttttacaaca attaccaaca acaacaaaca acaaacaaca ttacaattac atttacaatt	8160
accataccat gcctccaagg gactcttact cttatgctgc tcctccttct gctcaacttc	8220
acgaagttga tactcctcaa gagcacgaca agaaagagct tgttatcgga gatagggctt	8280
acgatgttac caacttcggt aagagacacc ctggtggaaa gatcattgct taccaagttg	8340
gaactgatgc taccgatgct tacaagcagt tccatgttag atctgctaag gctgacaaga	8400
tgcttaagtc tcttccttct cgtcctgttc acaagggata ctctccaaga agggctgatc	8460
ttatcgctga tttccaagag ttcaccaagc aacttgaggc tgagggaatg ttcgagcctt	8520
ctcttcctca tgttgcttac agacttgctg aggttatcgc tatgcatggt gctggtgctg	8580
ctcttatctg gcatggatac actttcgctg gaatcgctat gcttgaggtt gttcagggaa	8640
gatgtggatg gcttatgcat gaggggtggac attactctct cactggaaac attgctttcg	8700
acagagctat ccaagttgct tgttacggac ttggatgtgg aatgtctggt gcttggtggc	8760
gtaaccagca taacaagcac catgctactc ctcaaaagct tcagcacgat gttgatcttg	8820
atacccttcc tctcgttgct ttccatgaga gaatcgctgc taagggttaag tctcctgcta	8880
tgaaggcttg gctttctatg caagctaagc ttttcgctcc tgttaccact cttcttggtg	8940
ctcttgatg gcagctttac cttcatccta gacacatgct caggactaag cactacgatg	9000
agcttgctat gctcggaatc agatacggac ttgttgata ccttgctgct aactacggtg	9060
ctggatacgt tctcgcttgt taccttcttt acgttcagct tggagctatg tacatcttct	9120
gcaacttcgc tgtttctcat actcacctcc ctgttggtga gcctaacgag catgctactt	9180
gggttgagta cgctgctaac cacactacta actgttctcc atcttggtgg tgtgattggt	9240
ggatgtctta ccttaactac cagatcgagc accaccttta cccttctatg cctcaattca	9300
gacaccctaa gatcgctcct agagttaagc agcttttcga gaagcacgga cttcactacg	9360
atggttagagg atacttcgag gctatggctg atactttcgc taaccttgat aacgttgccc	9420
atgctcctga gaagaaaatg cagtaatgag atcgttcaaa catttggtgcaa taaagtttct	9480
taagattgaa tcctgttgcc ggtcttgcca tgattatcat ataatttctg ttgaattacg	9540

52/143

ttaagcacgt aataattaac atgtaatgca tgacgttatt tatgagatgg gtttttatga	9600
ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag cgcgcaaact	9660
aggataaatt atcgcgcgcg gtgtcatcta tgttactaga tcggtcgatt aaaaatccca	9720
attatatattg gtctaattta gtttggtatt gagtaaaaca aattcgaacc aaacaaaaat	9780
ataaatatat agtttttata tatatgcctt taagactttt tatagaattt tctttaaaaa	9840
atatctagaa atatttgcca ctcttctggc atgtaatat tctgttaaata tgaagtgctc	9900
cattttttatt aacttttaa atttggttgt acgatcactt tcttatcaag tgttactaaa	9960
atgcgtcaat ctctttgttc ttccatattc atatgtcaaa atctatcaaa attcttatat	10020
atctttttcg aatttgaagt gaaatttcga taatttaaaa ttaaatagaa catatcatta	10080
tttaggtatc atattgattt ttatacttaa ttactaaatt tggttaactt tgaaagtgtg	10140
catcaacgaa aaattagtca aacgactaaa ataaataaat atcatgtgtt attaagaaaa	10200
ttctcctata agaatatattt aatagatcat atgtttgtaa aaaaaattaa tttttactaa	10260
cacatatatt tactttatcaa aaatttgaca aagtaagatt aaaataatat tcatctaaca	10320
aaaaaaaaac cagaaaatgc tgaaaacccg gcaaaaccga accaatccaa accgatatag	10380
ttggtttggg ttgattttga tataaaccga accaactcgg tccatttgca cccctaatac	10440
taatagcttt aatatttcaa gatattatta agttaacgtt gtcaatatcc tggaaatttt	10500
gcaaaatgaa tcaagcctat atggctgtaa tatgaattta aaagcagctc gatgtggtgg	10560
taatatgtaa ttacttgat tctaaaaaaa tatcccaagt attaataatt tctgctagga	10620
agaaggttag ctacgattta cagcaaagcc agaatacaaa gaaccataaa gtgattgaag	10680
ctcgaaatat acgaaggaac aaatattttt aaaaaatac gcaatgactt ggaacaaaag	10740
aaagtgatat attttttgtt cttaacaag catcccctct aaagaatggc agttttcctt	10800
tgcatgtaac tattatgctc ctttcgttac aaaaattttg gactactatt gggaaacttct	10860
tctgaaaata gtgatagaac ccacacgagc atgtgctttc catttaattt taaaaaccaa	10920
gaaacataca tacataacat tccatcagcc tctctctctt ttattacgg ttaatgactt	10980
aaaacacatc ttattatccc atccttaaca cctagcagtg tctttatacg atctcatcga	11040
tcaccacttc aaaaccatgc agactgctgc tgcccctgga gctggcatcg gctaggctgg	11100
gtgccgcact gtcccgaag gtccctagcg acttgtttag attgatggga ccacctctca	11160
acttctgct gctgtccctg ctgctggatg tcctgcctca tctggccgat tgcacgtcc	11220
agtcccctgc atgtgcactc gctcctcaat tgcttaagat catcgcagca gctatcgaag	11280
tgctggctct gttgccctcc tccacggcct tggtttagt agtagctgcc gccgcccttc	11340
tggacttttt cccacaggaa ccgccgaata attcgataga accacacgag catgtgcttt	11400
cattttattt aaaaaccaag aaacatacat aacatttcat cagcctctct ctctctctct	11460
ctctctctct ctctctctct ctctctctct ctctctttat tacagctgtt aactaactt	11520
aaaacacatt catctcatta ttattattat tatccatcct taacacctag cagtgtcttt	11580
gtacgatctc ataatcgatc accccttcat caggatcct taggcttcac tccaacgttg	11640

53/143

ttgcagttac ggaacatgta cacaccatca tggttctcaa cgaactggca agatctccaa	11700
gttttccaaa ggctaacca catgttctca tcggtgtgtc tgtagtgctc tcccataact	11760
ttcttgatgc actcggtagc ttctctagca tggtagaatg ggatccttga aacgtagtga	11820
tggagcacat gagtctcgat gatgtcatgg aagatgattc cgaggattcc gaactctcta	11880
tcgatagtag cagcagcacc cttagcgaaa gtccactctt gagcatcgta atgaggcata	11940
gaagaatcgg tgtgctgaag gaaggtaacg aaaacaagcc agtggttaac aaggatccaa	12000
ggacagaacc atgtgatgaa agtaggccag aatccgaaaa ccttgtaagc ggtgtaaaca	12060
gaagtgaggg tagcaaggat tccaagatca gaaagaacga tgtaccagta gtccttctta	12120
tcgaaaacag ggctagaagg ccagtagtga gacttgaaga acttagaaac accagggtaa	12180
ggttgtccag tagcgttagt agcaaggtaa agagaaagtc ctccaagctg ttggaacaag	12240
agagcgaaaa cagagtagat aggagtttcc tcagcgatat cgtgaaggct ggtaacttgg	12300
tgcttctctt tgaattcctc ggcggtgtaa ggaacgaaaa ccatatctct ggtcatgtgt	12360
ccagtagcct tatggtgctt agcatgagag aacttccagc tgaagtaagg aaccataaca	12420
agagagtgga gaacccatcc aacggtatcg ttaacccatc cgtagttaga gaaagcagaa	12480
tgtccacact catgtccaag gatccagatt ccgaatccga aacaagagat agagaacacg	12540
taagcagacc aagcagcgaa tctaaggaat tcgttaggga gaagagggat gtaggtaagt	12600
ccaacgtaag cgatagcaga gatagccacg atatctctca ccacgtaaga catagacttc	12660
acgagagatc tctcgtaaca gtgcttaggg atagcgtcaa ggatatacctt gatggtgtaa	12720
tctggcacct tgaaaacgtt tccgaaggta tcgatagcgg tcttttgctg cttgaaagat	12780
gcaacgtttc cagaacgcct aacggtctta gtagatccct caaggatctc agatccagac	12840
acggtaacct tagacatggt atggttaattg taaatgtaat tgtaatgttg tttgttgttt	12900
gttgttgttg gtaattgttg taaaattttt ggtggtgatt ggttctttaa ggtgtgagag	12960
tgagtgtgta gttgtgtggt gggtttggtg agattgggga tgggtgggtt atatagtgga	13020
gactgaggaa tggggtcgtg agtgtaact ttgcatgggc tacacgtggg ttcttttggg	13080
cttacacgta gtattattca tgcaaatgca gccaatatcat atacggtatt ttaataatgt	13140
gtgggaatac aatatgccga gtattttact aattttggca atgacaagtg tacatttgga	13200
ttatcttact tggcctctct tgctttaatt tggattatct ttattctctt accttggccg	13260
ttcatattca catccctaaa ggcaagacag aattgaatgg tggccaaaaa ttaaaacgat	13320
ggatatgacc tacatagtgt aggatcaatt aacgtcgaag gaaaatactg attctctcaa	13380
gcatacggac aagggtaaat aacatagtca ccagaacata ataaacaaaa agtgcagaag	13440
caagactaaa aaaattagct atggacattc aggttcatat tggaacatc attatcctag	13500
tcttgtgacc atccttcctc ctgctctagt tgagaggcct tgggactaac gagaggtcag	13560
ttgggatagc agatccttat cctggactag ctttctggt gtttcagagt cttcgtgccg	13620
ccgtctacat ctatctccat taggtctgaa gatgactctt cacaccaacg acgtttaagg	13680

54/143

tctctatcct actcctagct tgcaatacct ggcttgcaat acctggagca tcgtgcacga	13740
tgattggata ctgtggagga ggagtgtttg ctgatttaga gctcccgggtt ggggtgatttg	13800
acttcgattt cagtttaggc ttgttgaaat ttttcagggtt ccattgtgaa gccttttagag	13860
cttgagcttc cttccatggt aatgccttga tcgaatactc ctagagaaaa gggaagtcga	13920
tctctgagta ttgaaatcga agtgcacatt ttttttcaac gtgtccaatc aatccacaaa	13980
caaagcagaa gacaggtaat ctttcatact tatactgaca agtaatagtc ttaccgtcat	14040
gcataataac gtctcgtttcc ttcaagaggg gttttccgac atccataacg acccgaagcc	14100
tcatgaaagc attaggggaag aacttttgggt tcttcttgtc atggccttta taggtgtcag	14160
ccgagctcgc caattcccgt ccgactggct ccgcaaaaata ttcgaacggc aagttatgga	14220
cttgcaacca taactccacg gtattgagca ggacctattg tgaagactca tctcatggag	14280
cttcagaatg tggttgtcag caaaccaatg accgaaatcc atcacatgac ggacgtccag	14340
tgggtgagcg aaacgaaaca ggaagcgcct atctttcaga gtcgtgagct ccacaccgga	14400
ttccggcaac tacgtgttgg gcaggcttcg ccgtattaga gatatgttga ggcagaccca	14460
tctgtgccac tcgtacaatt acgagagttg ttttttttgt gattttccta gtttctcgtt	14520
gatggtgagc tcatattcta catcgtatgg tctctcaacg tcgtttcctg tcatctgata	14580
tcccgtcatt tgcattccacg tgcgccgcct cccgtgccaa gtccctaggt gtcattgcacg	14640
ccaaattgggt ggtggtgcgg gctgccctgt gcttcttacc gatgggtgga ggttgagttt	14700
gggggtctcc gcggcgatgg tagtgggttg acggtttgggt gtgggttgac ggcattgatc	14760
aatttacttc ttgcttcaaa ttctttggca gaaaacaatt cattagatta gaactggaaa	14820
ccagagtgat gagacggatt aagtcagatt ccaacagagt tacatctctt aagaaataat	14880
gtaacccctt tagactttat atatttgcaa ttaaaaaaat aatttaactt ttagacttta	14940
tatatagttt taataactaa gtttaaccac tctattattt atatcgaaac tatttgtatg	15000
tctcccctct aaataaactt ggtattgtgt ttacagaacc tataatcaaa taatcaatac	15060
tcaactgaag tttgtgcagt taattgaagg gattaacggc caaatgcac tagtattatc	15120
aaccgaatag attcacacta gatggccatt tccatcaata tcatcgccgt tcttcttctg	15180
tccacatatc ccctctgaaa cttgagagac acctgcactt cattgtcctt attacgtgtt	15240
acaaaatgaa acccatgcat ccatgcaaac tgaagaatgg cgcaagaacc cttcccctcc	15300
atttcttatg tggcgaccat ccatttcacc atctcccgt ataaaacacc cccatcactt	15360
cacctagaac atcatcacta cttgcttatc catccaaaag ataccactt ttacaacaat	15420
taccaacaac aacaacaac aaacaacatt acaattacat ttacaattac cataccatgc	15480
cacctagcgc tgctaagcaa atgggagctt ctactggtgt tcatgctggt gttactgact	15540
cttctgcttt caccagaaag gatgttgctg atagacctga tctcaccatc gttggagatt	15600
ctgtttacga tgctaaggct ttcagatctg agcatcctgg tgggtgctcat ttcgtttctt	15660
tgttcggagg aagagatgct actgaggctt tcatggaata ccatagaagg gcttggccta	15720
agtctagaat gtctagattc cacgttggat ctcttgcttc tactgaggaa cctgttgctg	15780

ctgatgaggg ataccttcaa ctttgtgcta ggatcgctaa gatggcgcct tctgtttcct	15840
ctggattcgc tcctgcttct tactgggtta aggctggact tatccttgga tctgctatcg	15900
ctcttgaggc ttacatgctt tacgctggaa agagacttct ccttctatc gttcttgga	15960
ggcttttcgc tcttatcggt cttaacatcc agcatgatgc taaccatggg gctttgtcta	16020
agtctgcttc tgtaacctt gctcttgga tttgtcagga ttggatcgga ggatctatga	16080
tcctttggct tcaagagcat gttgttatgc accacctcca cactaacgat gttgataagg	16140
atcctgatca aaaggctcac ggtgctctta gactcaagcc tactgatgct tggtcaccta	16200
tgcattggct tcagcatctt taccttttgc ctggtgagac tatgtacgct ttcaagcttt	16260
tgttcctcga catctctgag cttgttatgt ggcgttgga gggtagcct atctctaagc	16320
ttgctggata cctctttatg ccttcttgc ttctcaagct taccttctgg gctagattcg	16380
ttgctttgcc tctttacctt gctccttctg ttcatactgc tgtgtgtatc gctgctactg	16440
ttatgactgg atctttctac ctgctttct tcttcttcat ctcccacaac ttcgaggggtg	16500
ttgcttctgt tggacctgat ggatctatca cttctatgac tagagggtgct agcttcctta	16560
agagacaagc tgagacttct tctaacgttg gaggacctt tcttgctact cttaacgggtg	16620
gactcaacta ccaaattgag catcacttgt tccctagagt tcaccatgga ttctacccta	16680
gacttgctcc tcttgtaag gctgagcttg aggctagagg aatcgagtac aagcactacc	16740
ctactatctg gtctaacctt gcttctacct tcagacatat gtacgctctt ggaagaaggc	16800
ctagatctaa ggctgagtaa tgacaagctt atgtgacgtg aaataataac ggtaaaatat	16860
atgtaataat aataataata aagccacaaa gtgagaatga ggggaagggg aaatgtgtaa	16920
tgagccagta gccggtgggtg ctaattttgt atcgtattgt caataaatca tgaattttgt	16980
ggtttttatg tgttttttta aatcatgaat tttaaatttt ataaaataat ctccaatcgg	17040
aagaacaaca ttccatatcc atgcatggat gtttctttac ccaaacttag ttcttgagag	17100
gatgaagcat caccgaacag ttctgcaact atccctcaaa agcttttaaa tgaacaacaa	17160
ggaacagagc aacgttccaa agatcccaa cgaaacatat tatctatact aatactatat	17220
tattaattac tactgcccgg aatcacatc cctgaatgat tcctattaac tacaagcctt	17280
gttgccggcg gagaagtgat cggcgccggc agaagcagcg gactcggaga cgaggccttg	17340
gaagatctga gtcgaacggg cagaatcagt attttccttc gacgttaatt gatcctacac	17400
tatgtaggtc atatccatcg ttttaatttt tggccaccat tcaattctgt cttgccttta	17460
gggatgtgaa tatgaacggc caaggtaaga gaataaaaaat aatccaaatt aaagcaagag	17520
aggccaagta agataatcca aatgtacact tgtcattgcc aaaattagta aaatactcgg	17580
catattgtat tcccacacat tattaataa ccgtatatgt attggctgca tttgcatgaa	17640
taatactacg tgtaagccca aaagaaccca cgtgtagccc atgcaaagtt aacactcacg	17700
acccattcc tcagtctcca ctatataaac ccaccatccc caatctcacc aaaccacca	17760
cacaactcac aactcactct cacaccttaa agaaccaatc accacaaaa attttacaac	17820

aattaccaac aacaacaaac aacaaacaac attacaatta catttacaat taccatacca	17880
tgagcgctgt taccgttact ggatctgata ctaagaacag aggatcttct agcaacaccg	17940
agcaagaggt tccaaaagtt gctatcgata ccaacggaaa cgtgttctct gttcctgatt	18000
tcaccatcaa ggacatcctt ggagctatcc ctcatgagtg ttacgagaga agattggcta	18060
cctctctcta ctacgtgttc agagatatct tctgcatgct taccaccgga taccttacct	18120
ataagatcct ttaccctctc ctcatctctt acacctctaa cagcatcatc aagttcactt	18180
tctgggccct ttacacttac gttcaaggac ttttcggaac cggaatctgg gttctcgctc	18240
atgagtgagg acatcaagct ttctctgatt acggaatcgt gaacgatttc gttggatgga	18300
cccttcactc ttaccttatg gttccttact tcagctggaa gtactctcat ggaaagcacc	18360
ataaggctac tggacacatg accagagata tggttttcgt tcctgccacc aaagaggaat	18420
tcaagaagtc taggaacttc ttcggtaacc tcgctgagta ctctgaggat tctccactta	18480
gaacccttta cgagcttctt gttcaacaac ttggaggatg gatcgcttac ctcttcgtta	18540
acgttacagg acaaccttac cctgatgttc cttcttgga atggaaccac ttctggctta	18600
cctctccact tttcgagcaa agagatgctc tctacatctt ctttctgat cttggaatcc	18660
tcaccaggg aatcgttctt actctttggt acaagaaatt cggaggatgg tcccttttca	18720
tcaactgggt cgttccttac atctgggtta accactggct cgttttcatc acattccttc	18780
agcacactga tcctactatg cctcattaca acgctgagga atggactttc gctaagggtg	18840
ctgctgtac tatcgataga aagttcggat tcacatcttc catgatata	18900
tcgagactca tgtgcttcac cactactgtt ctaggatccc attctacaac gctagacctg	18960
cttctgaggc tatcaagaaa gttatgggaa agcactacag gtctagcgac gagaacatgt	19020
ggaagtcact ttggaagtct ttcaggtctt gccaatcgt tgacggtgat aacggtgttc	19080
tcattgtccg taacatcaac aactgcggag ttggagctgc tgagaagtaa tgaaggggtg	19140
atcgattatg agatcgtaca aagacactgc taggtgttaa ggatggataa taataataat	19200
aatgagatga atgtgtttta agttagtgt acagctgtaa taaagagaga gagagagaga	19260
gagagagaga gagagagaga gagagagaga gagaggctga tgaaatgtta tgtatgtttc	19320
ttggttttta aaataaatga aagcacatgc tcgtgtgggt ctatcgaatt attcggcggg	19380
tcctgtggga aaaagtccag aagggccgcc gcagctacta ctacaaccaa ggccgtggag	19440
gagggcaaca gagccagcac ttcgatagct gctgcgatga tcttaagcaa ttgaggagcg	19500
agtgcacatg caggggactg gagcgtgcaa tcggccagat gaggcaggac atccagcagc	19560
agggacagca gcaggaagtt gagagggtgg cccatcaatc taaacaagtc gctagggacc	19620
ttccgggaca gtgcggcacc cagcctagcc gatgccagct ccaggggcag cagcagtctg	19680
catggttttg aagtgggtgat cgatgagatc gtataaagac actgctaggt gttaaggatg	19740
ggataataag atgtgtttta agtcattaac cgtaataaaa agagagagag gctgatggaa	19800
tgttatgtat gtatgtttct tggtttttaa aattaaatgg aaagcacatg ctctgtgtgg	19860
ttctatctcg attaaaaatc ccaattatat ttggtctaatt ttagtttggt attgagtaaa	19920

acaaattcga accaaaccaa aatataaata tatagttttt atatatatgc ctttaagact	19980
ttttatagaa ttttctttaa aaaatatcta gaaatatattg cgactcttct ggcatgtaat	20040
atctcgtaa atatgaagtg ctccattttt attaacctta aataattggg tgtacgatca	20100
ctttcttatac aagtgttact aaaatgcgtc aatctctttg ttcttccata ttcatatgtc	20160
aaaatctatc aaaattctta tatatctttt tcgaatttga agtgaaattt cgataattta	20220
aaattaaata gaacatatca ttatttaggt atcatattga tttttatact taattactaa	20280
atctggtaa ctttgaaagt gtacatcaac gaaaaattag tcaaacgact aaaataaata	20340
aatatcatgt gttattaaga aaattctcct ataagaatat tttaatagat catatgtttg	20400
taaaaaaaaat taattttttac taacacatat atttacttat caaaaatttg acaaagtaag	20460
attaaaataa tattcatcta acaaaaaaaaa aaccagaaaa tgctgaaaac ccggcaaaac	20520
cgaaccaatc caaaccgata tagttgggtt ggtttgattt tgatataaac cgaaccaact	20580
cggtccattt gcacccttaa tcataatagc tttaatattt caagatatta ttaagttaac	20640
gttgtcaata tcctggaaat ttgcaaaat gaatcaagcc tatatggctg taatatgaat	20700
ttaaaagcag ctcgatgtgg tggtaatatg taatttactt gattctaaaa aaatatccca	20760
agtattaata atttctgcta ggaagaaggt tagctacgat ttacagcaaa gccagaatac	20820
aaagaaccat aaagtgattg aagctcgaat tatacgaagg acaaatatt tttaaaaaa	20880
tacgcaatga cttggaacaa aagaaagtga tatatttttt gttcttaaac aagcatcccc	20940
tctaaagaat ggagttttt ctttgcatgt aactattatg ctcccttcgt taaaaaatt	21000
ttggactact attgggaact tcttctgaaa atagtcctgc aggctagtag attggttggt	21060
tggtttccat gtaccagaag gcttacccta ttagttgaaa gttgaaactt tgttccctac	21120
tcaattccta gttgtgtaaa tgtatgtata tgtaatgtgt ataaaacgta gtacttaaat	21180
gactaggagt gggtcttgag accgatgaga gatgggagca gaactaaaga tgatgacata	21240
attaagaacg aatttgaaag gctcttaggt ttgaatccta ttcgagaatg tttttgtcaa	21300
agatagtggc gatcttgaac caaagaaaac atttaaaaaa tcagtatccg gttacgttca	21360
tgcaaataga aagtggctta ggatctgatt gtaatttttag acttaaagag tctcttaaga	21420
ttcaatcctg gctgtgtaca aaactacaaa taatatattt tagactattt ggccttaact	21480
aaacttccac tcattattta ctgagggttag agaataagact tgcgaataaa cacattcccc	21540
agaaatactc atgatcccat aattagtcag agggtagtgc aatcagatct aagaacacac	21600
attccctcaa attttaatgc acatgtaatc atagtttagc acaattcaaa aataatgtag	21660
tattaaagac agaaatttgt agactttttt ttggcgtaa aagaagacta agtttatacg	21720
tacattttat ttttaagtga aaaccgaaat tttccatcga aatatatgaa tttagtatat	21780
atatttctgc aatgtactat tttgctattt tggcaacttt cagtggacta ctactttatt	21840
acaatgtgta tggatgcatg agtttgagta tacacatgtc taaatgcatg ctttgtaaaa	21900
cgtaacggac caaaaagag gatccatata aatacatctc atagcttcct ccattatttt	21960

58/143

ccgacacaaa cagagcattt tacaacaatt accaacaaca acaaacaaca aacaacatta	22020
caattacatt tacaattacc ataccatgga atttgctcaa cctctcgttg ctatggctca	22080
agagcagtag gctgctatcg atgctgttgt tgctcctgct atcttctctg ctaccgactc	22140
tattggatgg ggactcaagc ctatctcttc tgctactaag gatctccctc tcgttgaatc	22200
tcctaccctt cttatccttt ctctcctcgc ttacttcgct atcgttggtt ctggactcgt	22260
ttaccgtaaa gtgttcctta gaaccgttaa gggacaggat cctttccttc tcaaggctct	22320
tatgctcgct cacaacgttt tccttatcgg actcagcctt tacatgtgcc tcaagctcgt	22380
ttacgaggct tacgtgaaca agtactcctt ctggggaaac gcttacaacc ctgctcaaac	22440
cgagatggct aaggatgatc ggatcttcta cgtgtccaag atctacgagt tcatggacac	22500
cttcatcatg cttctcaagg gaaacgttaa ccagggttcc ttcctccatg tttaccacca	22560
cggatctatc tctggaatct ggtggatgat cacttatgct gctccagggtg gagatgctta	22620
cttctctgct gctctcaact cttgggttca tgtgtgcatg tacacctact acttcatggc	22680
tgctgttctt cctaaggacg aaaagaccaa gagaaagtac ctttggtggg gaagatacct	22740
taccagatg caaatgttcc agttcttcat gaaccttctc caggctgttt acctcctcta	22800
ctcttcttct ccttacccta agttcattgc tcaactcctc gttgtttaca tggttaccct	22860
cctcatgctt ttcggaaact tctactacat gaagcaccac gcttctaagt gataagggcc	22920
gccgccatgt gacagatcga aggaagaaag tgtaataaga cgactctcac tactcgatcg	22980
ctagtgattg tcattgttat atataataat gttatctttc acaacttatc gtaatgcatg	23040
tgaaactata acacattaat cctacttgct atatgataac actctcccca tttaaaactc	23100
ttgtcaattt aaagatataa gattctttta atgattaaaa aaaatatatt ataaattcaa	23160
tcactcctac taataaatta ttaattatta ttatttgatt aaaaaatac ttatactaata	23220
ttagtctgaa tagaataatt agattctagc ctgcagggcg gccgcggatc ccatggagtc	23280
aaagattcaa atagaggacc taacagaact cgccgttaaag actggcgaac agttcataca	23340
gagtctctta cgactcaatg acaagaagaa aatcttcgtc aacatggtgg agcacgacac	23400
acttgtctac tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac	23460
ttttcaacaa agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca	23520
ctttattgtg aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa	23580
aggaaaggcc atcgttgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc	23640
cacgaggagc atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg	23700
atgtgatatc tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc	23760
ttcctctata taaggaagtt catttcattt ggagagaaca cgggggactg aattaaatat	23820
gagccctgag aggcgtcctg ttgaaatcag acctgctact gctgctgata tggctgctgt	23880
ttgtgatatc gtgaaccact acatcgagac ttctaccgtt aacttcagaa ctgagcctca	23940
aactcctcaa gagtggatcg atgatcttga gagactcaa gatagatacc cttggcttgt	24000
tgctgagggt gaggggtgtg ttgctggaat cgcttacgct ggaccttgga aggctagaaa	24060

59/143

cgcttacgat tggactgttg agtctaccgt ttacgtttca cacagacatc agagacttgg 24120
 acttgatct accctttaca ctcaccttct caagtctatg gaagctcagg gattcaagtc 24180
 tgttgttgct gttatcggac tccctaacga tccttctgtt agacttcatg aggctcttgg 24240
 atacactgct agaggaactc ttagagctgc tggatacaag cacggtggat ggcatgatgt 24300
 tggattctgg caaagagatt tcgagcttcc tgctcctcct agacctgtta gaccagttac 24360
 tcagatctga atttgcgtga tcgttcaaac atttggcaat aaagtttctt aagattgaat 24420
 cctgttgccg gtcttgcat gattatcata taatttctgt tgaattacgt taagcatgta 24480
 ataattaaca tgtaatgcat gacgttattt atgagatggg tttttatgat tagagtcccc 24540
 caattataca ttttaatacgc gatagaaaac aaaatatagc gcgcaaacta ggataaatta 24600
 tcgcgcgcgg tgtcatctat gttactagat cactagtgat gtacggttaa aaccacccca 24660
 gtacattaaa aacgtccgca atgtgttatt aagttgtcta agcgtcaatt tgtttacacc 24720
 acaatatatc ctgccaccag ccagccaaca gtcctccgac cggcagctcg gcacaaaatc 24780
 accactcgat acaggcagcc catcagtcc 24809

<210> 6
 <211> 26543
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Trình tự nucleotit pGA7- mod_F

<400> 6
 tcctgtggtt ggcatgcaca tacaaatgga cgaacggata aaccttttca cgccctttta 60
 aatatccgat tattctaata aacgctcttt tctcttaggt ttaccgcga atatatcctg 120
 tcaaactg atagttaaaa ctgaaggcgg gaaacgacaa tctgctagtg gatctcccg 180
 tcacgacgtt gtaaaacggg cgggcggccg cctagaatct aattattcta ttcagactaa 240
 attagtataa gtattttttt aatcaataaa tattaattaa taatttatta gtaggagtga 300
 ttgaatttat aatatatttt ttttaatcat ttaaagaatc ttatatcttt aaattgacaa 360
 gagttttaaa tggggagagt gttatcatat gacaagtagg attaagtgtg tatagtttca 420
 catgcattac gataagttgt gaaagataac attattatat ataacaatga caatcactag 480
 cgatcgagta gtgagagtcg tcttattaca ctttcttcct tcgatctgtc acatggcggc 540
 ggcccgcat cgcgataatt ctcatgctgc cttctccgcc ttgccgttga cgtagtagtg 600
 ctgcccagacc ttatccaagt tcgagaacgt cgccttccag gcgccgtaat aggacagcac 660
 cttgtagttc agccccact tcttcgcgaa cgggacgaac cgccggctca cctccggctg 720
 gcgaaactgc ggcatgtccg ggaacagggtg atgaatgacc tggcagttca gatatcccat 780
 caaccagttc acgtacccgc gcgacgggtc gatgtccacg gtgtgatcga ccgcgtagtt 840
 caccagctc aggtgcttat ccgagggcac gaccgggagg tgcgtgtggc tcgtggagaa 900
 gtgcgcgaag aggtacatgt acgcgatcca gttgccgaag gtgaaccacc agtacgcgac 960

60/143

gggccacgag taccccgtcg cgagtttaaat caccgcggtc ctgacgacgt gagagacgag	1020
catccacgac gcctcctcgt agttcttctt tcgcaacacc tgccgcgggt gcaggacgta	1080
gatccagaac gcctggacga gcagcccgga ggtcaccggg acgaacgtcc acgcctgaag	1140
ccgagcccac gcgcgggaga accccctcgg ccggttggtc tccacggcgg tgttaaaaaa	1200
cgccaccgcg ggggtcgtgt ccagggtccat gtcgtgcctc actttctgcg gcgtcgcgtg	1260
gtgcttattg tgcattctggt tccacatctc ccgctcgtg gacagcccga acccgcacgt	1320
catcgcttgg aggcgcttgt cgacgtagac ggaccccgtg agcgagttgt gcccgcctc	1380
gtgctggacc caaccgcacc gagcgccgaa gaacgcgccg tacacgacgg acgcgatgat	1440
cggttaccg gcgtacatga gggcggtgcc gagggcgaa gtcgcgagga gtcgagtaa	1500
ccgatacgcg acgtgcgtta tcgagggctt aaagaacccg tcgcgttcga gtcgcgcg	1560
ccaccgcgcg aaatcctcca acatcggcgc gtctcgtc tcgctgcgtt tgatctccgc	1620
ggggcgcgac ggcagcgctc tgagcatctt ccacgcctta agcgatcgca tgtggaactc	1680
cttgaacgcc tccgtggcgt ccgcgcccgt gttcgcgagc atgtagaata tcacgtgcc	1740
tcccgggtgt ttgaagtttg tgacgtcgta ctgacgtcc tccacgcgca cccatcgct	1800
ctcgaacgtc gccgcgagct cgtgcggctc gagcgttttg agatcgacgg gcgcggctga	1860
cgcgctccttg gcgtcgagcg cctccgcgga ggatttgctg cgcgtcagcg gcgatcgcg	1920
ggacgatcgg ccgtccgtct tcggcgggca catcgtcgcg cgcgcgactt aaaccgacga	1980
cggacggacg aacctgcaac ggcgaattat caattgacgc gttgctctgt ttgtgtcgga	2040
aaataatgga ggaagctatg agatgtattt gcatggatcc tcttttgtgg tccgttacgt	2100
tttgcaaagc atgcatttag acatgtgtat actcaaactc atgcatccat acacattgta	2160
ataaagtagt agtccactga aagttgcaa aatagcaaaa tagtacattg cagaaatata	2220
tatactaaat tcatatatat cgatggaaaa tttcggtttt ccacttaaaa taaaatgtac	2280
gtataaactt agtcttcctt taacgcaaaa aaaaagtcta caaatttctg tctttaatac	2340
tacattatit ttgaattgtg ctaaactatg attacatgtg cattaaaatt tgagggaatg	2400
tgtgttctta gatctgattg gcataccctc tgactaatta tgggatcatg agtatttctc	2460
gggaatgtgt ttattcgcaa gtctattctc taacctcagt aaataatgag tggaagtta	2520
gttaaggcca aatagtctaa aatatattat ttgtagtttt gtacacagcc aggattgaat	2580
cttaagagac tctttaagtc taaaattaca atcagatcct agaccacttt ctatttgc	2640
gaacgtaacc ggatactgat tttttaaatg ttttctttgg ttcaaatcg ccactatctt	2700
tgacaaaaac attctcgaat aggattcaaa cctaagagcc tttcaaatc gttcttaatt	2760
atgtcatcat ctttagttct gctcccatct ctcatcggtc tcaagaacca ctctagtca	2820
tttaagtact acgttttata cgcattacat atacatacat ttacacaact aggaattgag	2880
tagggaacaa agtttcaact ttcaactaat agggtaagcc ttctggtaca tggaaccaa	2940
ccaaccaatc tactaggcgg ccgcccgtcg ggatcttctg caagcatctc tatttctga	3000
aggtctaacc tcgaagattt aagatttaac tacgtttata attacaaaat tgattctagt	3060

atctttaatt taatgcttat acattattaa ttaatttagt actttcaatt tgttttcaga	3120
aattatttta ctatttttta taaaataaaa gggagaaaat ggctatttaa actgaaggcg	3180
ggaaacgaca atctgctagt ggatctccca gtcacgacgt tgtaaaacgg gcgccccgcg	3240
gaaagcttgc ggccgcggta ccgcccgttc gactcagatc ttccaaggcc tcgtctccga	3300
gtccgctgct tctcgccgcg ccgatcactt ctccgcccgc aacaaggctt gtagttaata	3360
ggaatcattc agggattgtg attccgggca gtagtaatta ataatatagt attagtatag	3420
ataatatgtt tcgtttggga tctttggaac gttgctctgt tccttgttgt tcattttaaa	3480
gcttttgagg gatagttgca gaactgttcg gtgatgcttc atcctctcaa gaactagatt	3540
tgggtaaaga aacatccatg catggatatg gaatgttggt cttccgattg gagattattt	3600
tataaaattt aaaattcatg atttaaaaaa acacataaaa accacaaaat tcatgattta	3660
ttgacaatac gatacaaaat tagcaccacc ggctactggc tcattacaca tttccccttc	3720
ccctcattct cactttgtgg ctttattatt attattatta catatatttt accgttatta	3780
tttcacgtca cataagcttg ttaattaatc attagtgagc cttctcagcc tttccgtaa	3840
cgtagtagtg ctgtcccacc ttatcaagggt tagagaaagt agccttccaa gcaccgtagt	3900
aagagagcac cttgtagttg agtccccact tcttagcgaa aggaacgaat cttctgctaa	3960
cctcaggctg tctgaattga ggcatatcag ggaagagggt gtggataacc tgacagttaa	4020
ggtatcccat aagccagttc acgtatcctc tagaaggatc gatatcaacg gtgtgatcaa	4080
cagcgtagtt aaccaagaa aggtgcttat cagatggaac aacagggagg tgagtatgag	4140
aagtagagaa gtgagcgaaa aggtacatgt aagcgatcca gtttccgaaa gtgaaccacc	4200
agtaagcaac aggccaagag tatccagtag caagcttgat aacagcgggt ctaacaacat	4260
gagaaacgag catccaagaa gcctcttcgt agttcttctt acggagaact tgtctaggg	4320
ggagaacgta gatccagaaa gcttgaacaa gaagtccaga ggtaacagga acgaaagtcc	4380
aagcttgaag tctagcccaa gctctagaga atcctctagg tctgttatcc tcaacagcag	4440
tgttgaagaa agccacagca ggagtgggtat caagatccat atcgtgtcta accttttgag	4500
gggtagcatg gtgcttggtt tgcatctggt tccacatctc accagaagta gaaagtccga	4560
atccacaagt catagcctga agtctcttgt ccacgtaaac agatccggta agagagtatt	4620
gtccaccctc atgttgaacc catccacatc tagctccgaa gaaagcaccg taaacaacag	4680
aagcaatgat agggatatcca gcgtacataa gagcagttcc aagagcgaat gtagcaagaa	4740
gctcgagaag tctgtaagcc acatgggtga tagaaggctt gaagaatcca tctctctcaa	4800
gctcagcacg ccatctagcg aaatcctcaa gcataggagc atcctcagac tcagatctct	4860
tgatctcagc aggtctagaa ggcaaagctc taagcatctt ccaagccttg agagaacgca	4920
tgtggaattc tttgaaagcc tcagtagcat cagcaccagt gttagcaagc atgtagaaga	4980
tcacagatcc accaggggtgc ttgaagttag tcacatcgta ctcaacgtcc tcaactctaa	5040
cccatctagt ctcgaaagta gcagcaagct catgaggctc aagagtctta agatcaacag	5100

gagcagtaga agcatcctta gcatcaagag cctcagcaga agatttagac ctggtaagtg	5160
gagatctagg agaagatctt ccatcagtct taggagggca catggtatgg taattgtaaa	5220
tgtaattgta atgttgtttg ttgtttgttg ttgttggtaa ttgttgtaaa attaatgaag	5280
tgggtatctt ttggatggat aagcaagtag tgatgatgtt ctaggtgaag tgatgggggt	5340
gttttatagc gggagatggg gaaatggatg gtcgccacat aagaaatgga ggggaagggt	5400
tcttgcgcca ttcttcagtt tgcattggatg catgggtttc attttgtaac acgtaataag	5460
gacaatgaag tgcaggtgtc tctcaagttt cagaggggat atgtggacag aagaagaacg	5520
gcgatgatat tgatggaaat ggccatctag tgtgaatcta ttcggttgat aatactagt	5580
cattttggcc gttaatccct tcaattaact gcacaaactt cagttgagta ttgattat	5640
gattataggt tctgtaaaca caataccaag tttatttaga ggggagacat acaaatagtt	5700
tcgatataaa taatagagt gttaaactta gttattaaaa ctatatataa agtctaaaag	5760
ttaaattatt tttttaattg caaatatata aagtctaaag gggttacatt atttcttaag	5820
agatgtaact ctggttgaat ctgacttaat ccgtctcatc actctggttt ccagttctaa	5880
tctaataaat tgttttctgc caaagaattt gaagcaagaa gtaaattgat caatgccgtc	5940
aaccacacc aaaccgtcaa cccactacca tcgccgcgga gacccccaaa ctcaacctcc	6000
acccatcggg aagaagcaca gggcagccc caccaccacc aatttgccgt gcatgacacc	6060
tagggacttg gcacgggagg cggcgcacgt ggatgcaaat gacgggatat cagatgacag	6120
gaaacgacgt tgagagacca tacgatgtag aatatgagct caccatcaac gagaaactag	6180
gaaaatcaca aaaaaaaca ctctcgtaat tgtacgagt gacacagatg gtctgcctca	6240
acatatctct aatacggcga agcctgccca acacgtagt gccggaatcc ggtgtggagc	6300
tcacgactct gaaagatagg cgcttcctgt ttcgtttcgc tcacccactg gacgtccgtc	6360
atgtgatgga tttcgggtcat tggtttgctg acaaccacat tctgaagctc catgagatga	6420
gtcttcacaa taggtcctgc tcaataccgt ggagttatgg ttgcaagtcc ataacttgcc	6480
gttcgaatat tttgcggagc cagtcggacg ggaattggcg agctcggctg acacctataa	6540
aggccatgac aagaagaacc aaaagttctt ccctaagtct ttcattgaggc ttcgggtcgt	6600
tatggatgtc ggaaaacccc tcttgaagga acgagacgtt attatgcatg acggtgaagc	6660
tattacttgt cagtataagt atgaaagatt acctgtcttc tgctttgttt gtggattgat	6720
tggacacgtt gaaaaaaaaat gtgcacttcg atttcaatac tcagagatcg acttcccttt	6780
tctctaggag tattcgtatc aggcattaac atggaaggaa gctcaagctc taaaggcttc	6840
acaatggaac ctgaaaaatt tcaacaagcc taaactgaaa tcgaagtcaa atcaccaac	6900
cgggagctct aaatcagcaa acactcctcc tccacagtat ccaatcatcg tgcacgatgc	6960
tccaggtatt gcaagccagg tattgcaagc taggagtagg atagagacct taaacgtcgt	7020
tggtgtgaag agtcatcttc agacctaatg gagatagatg tagacggcgg cacgaagact	7080
ctgaaacacc agaaaggcta gtccaggata aggatctgct atcccaactg acctctcggt	7140
agtcccaagg cctctcaact agagcaggag gaaggatggg cacaagacta ggataatgat	7200

63/143

gtttccaata tgaacctgaa tgtccatagc taatTTTTTT agtcttgctt ctgcactttt	7260
tgtttattat gttctggtga ctatgttatt tacccttgtc cgtatgcttg agggtagcct	7320
agtagattgg ttggttggtt tccatgtacc agaaggctta ccctattagt tgaaagttga	7380
aactttgttc cctactcaat tcctagttgt gtaaagtgtat gtatatgtaa tgtgtataaa	7440
acgtagtact taaatgacta ggagtgggtc ttgagaccga tgagagatgg gagcagaact	7500
aaagatgatg acataattaa gaacgaatTTT gaaaggctct taggtttgaa tcctattcga	7560
gaatgtTTTT gtcaaagata gtggcgattt tgaaccaaag aaaacattta aaaaatcagt	7620
atccggttac gttcatgcaa atagaaaagt gtctaggatc tgattgtaat tttagactta	7680
aagagtctct taagattcaa tcctggctgt gtacaaaact acaaataata tatttttagac	7740
tattttgcct taactaaact tccactcatt atttactgag gttagagaat agacttgcg	7800
ataaacacat tcccagagaaa tactcatgat ccataatta gtcagagggt atgccaatca	7860
gatctaagaa cacacattcc ctcaaatttt aatgcacatg taatcatagt ttagcacaat	7920
tcaaaaataa tgtagtatta aagacagaaa tttgtagact tttttttggc gttaaagaa	7980
gactaagttt atacgtacat tttattttta gtggaaaacc gaaattttcc atcgaaatat	8040
atgaatttag tatatatatt tctgcaatgt actattttgc tattttggca actttcagtg	8100
gactactact ttattacaat gtgtatggat gcatgagttt gagtatacac atgtctaaat	8160
gcatgctttg taaaacgtaa cggaccacaa aagaggatcc atacaaatac atctcatagc	8220
ttcctccatt attttccgac acaaacagag cattttacaa caattaccaa caacaacaaa	8280
caacaacaaa cattacaatt acattttacaa ttaccatacc atggcctcta tcgctatccc	8340
tgctgctctt gctggaactc ttggatacgt tacctacaat gtggctaacc ctgatatccc	8400
agcttctgag aaagttcctg cttacttcat gcaggttgag tactggggac ctactatcgg	8460
aactattgga tacctcctct tcatctactt cggaaagcgt atcatgcaga acagatctca	8520
acctttcgga ctcaagaacg ctatgctcgt ttacaacttc taccagacct tcttcaacag	8580
ctactgcac taccctttcg ttactttctca tagggctcag ggacttaagg tttggggaaa	8640
catccctgat atgactgcta actcttgggg aatctctcag gttatctggc ttcactacaa	8700
caacaagtac gttgagcttc tcgacacctt cttcatggtg atgaggaaga agttcgacca	8760
gctttctttc cttcacatct accaccacac tcttctcatc tggatcatggt tcgttggtat	8820
gaagcttgag cctggttgag attgctactt cggatcttct gttaacacct tcgtgcacgt	8880
gatcatgtac tcttactacg gacttgctgc tcttgaggtt aactgtttct ggaagaagta	8940
catcaccag atccagatgc ttcagttctg tatctgtgct tctcactcta tctacaccgc	9000
ttacgttcag aataccgctt tctggcttcc ttaccttcaa ctctgggtta tgggtgaacat	9060
gttcgttctc ttcgccaact tctaccgtaa gaggtacaag tctaagggtg ctaagaagca	9120
gtgataaggc gcgcggcgcg ccgggccgcc gccatgtgac agatcgaagg aagaaagtgt	9180
aataagacga ctctcactac tcgatcgcta gtgattgtca ttgttatata taataatgtt	9240

atctttcaca	acttatcgta	atgcatgtga	aactataaca	cattaatcct	acttgtcata	9300
tgataacact	ctccccat	aaaactcttg	tcaattttaa	gatataagat	tctttaaatg	9360
attaaaaaaaa	atatattata	aattcaatca	ctcctactaa	taaattatta	attattat	9420
attgattaaa	aaaatactta	tactaattta	gtctgaatag	aataattaga	ttctagtctc	9480
atcccccttt	aaaccaactt	agtaaacggt	ttttttttta	attttatgaa	gttaagtttt	9540
taccttgttt	ttaaaaagaa	tcgttcataa	gatgccatgc	cagaacatta	gctacacgtt	9600
acacatagca	tgagccgcg	gagaattggt	tttcttcgcc	acttgtcact	cccttcaaac	9660
acctaagagc	ttctctctca	cagcacacac	atacaatcac	atgcgtgcat	gcattattac	9720
acgtgatcgc	catgcaaatc	tcctttatag	cctataaatt	aactcatccg	cttcactctt	9780
tactcaaacc	aaaactcatc	gatacaaaca	agattaaaaa	catacacgag	gatcttttac	9840
aacaattacc	aacaacaaca	aacaacaaac	aacattacaa	ttacatttac	aattaccata	9900
ccatgcctcc	aagggactct	tactcttatg	ctgctcctcc	ttctgctcaa	cttcacgaag	9960
ttgatactcc	tcaagagcac	gacaagaaag	agcttggtat	cggagatagg	gcttacgatg	10020
ttaccaactt	cgtaagaga	caccctgggtg	gaaagatcat	tgcttaccaa	gttggaactg	10080
atgctaccga	tgcttacaag	cagttccatg	ttagatctgc	taaggctgac	aagatgctta	10140
agtctcttcc	ttctcgtcct	gttcacaagg	gatactctcc	aagaagggct	gatcttatcg	10200
ctgatttcca	agagttcacc	aagcaacttg	aggctgaggg	aatgttcgag	ccttctcttc	10260
ctcatgttgc	ttacagactt	gctgagggtta	tcgctatgca	tggtgctggt	gctgctctta	10320
tctggcatgg	atacactttc	gctggaatcg	ctatgcttgg	agttgttcag	ggaagatgtg	10380
gatggcttat	gcatgagggt	ggacattact	ctctcactgg	aaacattgct	ttcgacagag	10440
ctatccaagt	tgcttggtac	ggacttggat	gtggaatgtc	tggtgcttgg	tggcgtaacc	10500
agcataacaa	gcaccatgct	actcctcaaa	agcttcagca	cgatgttgat	cttgataccc	10560
ttcctctcgt	tgctttccat	gagagaatcg	ctgctaagggt	taagtctcct	gctatgaagg	10620
cttggctttc	tatgcaagct	aagcttttcg	ctcctgttac	cactcttctt	gttgctcttg	10680
gatggcagct	ttaccttcat	cctagacaca	tgctcaggac	taagcactac	gatgagcttg	10740
ctatgctcgg	aatcagatac	ggacttggtg	gataccttgc	tgctaactac	ggtgctggat	10800
acgttctcgc	ttgttacctt	ctttacgttc	agcttggagc	tatgtacatc	ttctgcaact	10860
tcgctgtttc	tcatactcac	ctccctgttg	ttgagcctaa	cgagcatgct	acttgggttg	10920
agtacgctgc	taaccacact	actaactggt	ctccatcttg	gtggtgtgat	tggtggatgt	10980
cttaccttaa	ctaccagatc	gagcaccacc	tttacccttc	tatgcctcaa	ttcagacacc	11040
ctaagatcgc	tcctagagtt	aagcagcttt	tcgagaagca	cggacttcac	tacgatgtta	11100
gaggatactt	cgaggctatg	gctgatactt	tcgctaacct	tgataacggt	gcccattgctc	11160
ctgagaagaa	aatgcagtaa	tgagatcggt	caaacatttg	gcaataaagt	ttcttaagat	11220
tgaatcctgt	tgccggtctt	gcgatgatta	tcatataatt	tctgttgaat	tacgttaagc	11280
acgtaataat	taacatgtaa	tgcatgacgt	tatttatgag	atgggttttt	atgattagag	11340

65/143

tcccgcaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca aactaggata	11400
aattatcgcg cgcggtgtca tctatgttac tagatcggtc gattaanaa cccaattata	11460
tttggctctaa tttagtttgg tattgagtaa aacaaattcg aaccaaacca aaatataaat	11520
atatagtttt tatatatatg cctttaagac tttttataga attttcttta aaaaatatct	11580
agaaatattt gcgactcttc tggcatgtaa tttttcgta aatatgaagt gctccatttt	11640
tattaacttt aaataattgg ttgtacgac actttcttat caagtgttac taaaatgcgt	11700
caatctcttt gtctctccat attcatatgt caaaatctat caaaattctt atatatcttt	11760
ttcgaatttg aagtgaatt tcgataattt aaaattaaat agaacatatc attatttagg	11820
tatcatattg atttttatac ttaattacta aatttggtta actttgaaag tgtacatcaa	11880
cgaaaaatta gtcaaacgac taaaataaat aaatatcatg tgttattaag aaaattctcc	11940
tataagaata ttttaataga tcatatgttt gtaaaaaaa ttaattttta ctaacacata	12000
tatttactta tcaaaaattt gacaaagtaa gattaanaa atattcatct aacaaaaaa	12060
aaaccagaaa atgctgaaaa cccggcaaaa ccgaaccaat ccaaaccgat atagttgggt	12120
tggtttgatt ttgatataaa ccgaaccaac tcggtccatt tgcacccta atcataatag	12180
ctttaatatt tcaagatatt attaagttaa cgttgtcaat atcctggaaa ttttgcaaaa	12240
tgaatcaagc ctatatggct gtaatatgaa tttaaaagca gctcgaatg gtggaatat	12300
gtaatttact tgattctaaa aaaatatccc aagtattaat aatttctgct aggaagaagg	12360
ttagctacga ttacagcaa agccagaata caaagaacca taaagtgatt gaagctcgaa	12420
atatacgaag gaacaaatat ttttaaaaaa atacgcaatg acttggaaca aaagaaagt	12480
atatattttt tgttcttaaa caagcatccc ctctaaagaa tggcagtttt cttttgcatg	12540
taactattat gctcccttcg ttacaaaaat tttggactac tattgggaac ttcttctgaa	12600
aatagtata gaaccacac gagcatgtgc tttccattta attttaaaaa ccaagaaaca	12660
tacatacata acattccatc agcctctctc tctttttatt acggttaatg acttaaaaca	12720
catcttatta tcccatcctt aacacctagc agtgtcttta tacgatctca tcgatcacca	12780
cttcaaaacc atgcagactg ctgctgcccc tggagctggc atcggctagg ctgggtgccc	12840
cactgtcccg gaaggctcct agcgacttgt ttagattgat gggaccacct ctcaacttcc	12900
tgctgtctgc cctgtgtctg gatgtcctgc ctcatctggc cgattgcacg ctccagtccc	12960
ctgcatgtgc actcgtctct caattgctta agatcatcgc agcagctatc gaagtgtctg	13020
ctctgttgcc ctctccacg gccttggtg tagtagtagc tgccgccgcc ctcttgact	13080
ttttccca ggaaccgccg aataattcga tagaaccaca cgagcatgtg ctttcattta	13140
ttttaaaaac caagaaacat acataacatt tcatcagcct ctctctctct ctctctctct	13200
ctctctctct ctctctctct ctctctctct ttattacagc tgttacacta acttaaaaca	13260
cattcatctc attattatta ttattatcca tccttaacac ctagcagtggt ctttgtacga	13320
tctcataatc gatcaccctt tcatcaggta tccttaggct tcaactcaac gttgttgacg	13380

66/143

ttacggaaca tgtacacacc atcatggttc tcaacgaact ggcaagatct ccaagttttc	13440
caaaggctaa cccacatggt ctcatcggtg tgtctgtagt gctctcccat aactttcttg	13500
atgcactcgg tagctttctct agcatggtag aatgggatcc ttgaaacgta gtgatggagc	13560
acatgagtct cgatgatgtc atggaagatg attccgagga ttccgaactc tctatcgata	13620
gtagcagcag cacccttagc gaaagtccac tcttgagcat cgtaatgagg catagaagaa	13680
tcggtgtgct gaaggaaggt aacgaaaaca agccagtggg taacaaggat ccaaggacag	13740
aaccatgtga tgaaagtagg ccagaatccg aaaaccttgt aagcgggtga aacagaagtg	13800
agggtagcaa ggattccaag atcagaaaga acgatgtacc agtagtcctt cttatcgaaa	13860
acagggctag aaggccagta gtgagacttg aagaacttag aaacaccagg gtaaggttgt	13920
ccagtagcgt tagtagcaag gtaaagagaa agtcctccaa gctgttgga caagagagcg	13980
aaaacagagt agataggagt ttcctcagcg atatcgtgaa ggctggtaac ttggtgcttc	14040
tctttgaatt cctcggcggg gtaaggaacg aaaaccatat ctctgggtcat gtgtccagta	14100
gccttatggg gcttagcatg agagaacttc cagctgaagt aaggaaccat aacaagagag	14160
tggagaaccc atccaacggg atcgttaacc catccgtagt tagagaaagc agaatgtcca	14220
cactcatgtc caaggatcca gattccgaat ccgaaacaag agatagagaa cacgtaagca	14280
gaccaagcag cgaatctaag gaattcggtt gggagaagag ggatgtaggt aagtccaacg	14340
taagcgatag cagagatagc cacgatatct ctaccacgt aagacataga cttcacgaga	14400
gatctctcgt aacagtgcct agggatagcg tcaaggatat ccttgatggg gtaatctggc	14460
accttgaaaa cgtttccgaa ggtatcgata gcggctcttt gctgcttgaa agatgcaacg	14520
tttccagaac gcctaacggg cttagtagat ccctcaagga tctcagatcc agacacggta	14580
accttagaca tggtaggta attgtaaag taattgtaat gttgtttggt gtttggtgtt	14640
gttggttaatt gttgtaaaat ttttggtggg gattggttct ttaagggtgtg agagtgagtt	14700
gtgagttgtg tgggtgggttt ggtgagattg gggatgggtg gtttatatag tggagactga	14760
ggaatggggg cgtgagtgtt aactttgcat gggctacacg tgggttcttt tgggcttaca	14820
cgtagtatta ttcatgcaaa tgcagccaat acatatacgg tattttaata atgtgtggga	14880
atacaatatg ccgagtattt tactaatttt ggcaatgaca agtgtacatt tggattatct	14940
tacttggcct ctcttgcttt aatttggatt atttttattc tcttaccttg gccgttcata	15000
ttcacatccc taaaggcaag acagaattga atggtggcca aaaattaaaa cgatggatat	15060
gacctacata gtgtaggata aattaacgtc gaaggaaaat actgattctc tcaagcatac	15120
ggacaagggt aaataacata gtcaccagaa cataataaac aaaaagtgca gaagcaagac	15180
taaaaaaatt agctatggac attcaggttc atattggaaa catcattatc ctagtcttgt	15240
gaccatcctt cctcctgctc tagttgagag gccttgggac taacgagagg tcagttggga	15300
tagcagatcc ttatcctgga ctacgctttc tgggtgttca gagtcttcgt gccgccgtct	15360
acatctatct ccattaggtc tgaagatgac tcttcacacc aacgacgttt aagggtctcta	15420
tcctactcct agcttgcaat acctggcttg caatacctgg agcatcgtgc acgatgattg	15480

gatactgtgg aggaggagtg tttgctgatt tagagctccc ggttgggtga tttgacttcg 15540
 atttcagttt aggcttggtg aaatttttca ggttccattg tgaagccttt agagcttgag 15600
 cttcctttcca tgtaaatgcc ttgatcgaat actcctagag aaaaggggaag tcgatctctg 15660
 agtattgaaa tcgaagtgc cttttttttt caacgtgtcc aatcaatcca caaacaagc 15720
 agaagacagg taatctttca tacttatact gacaagtaat agtcttaccg tcatgcataa 15780
 taacgtctcg ttccttcaag aggggttttc cgacatccat aacgacccga agcctcatga 15840
 aagcattagg gaagaacttt tggttcttct tgtcatggcc tttatagggtg tcagccgagc 15900
 tcgccaatcc cgtccgact ggctccgcaa aatattcgaa cggcaagtta tggacttgca 15960
 accataactc cacggtattg agcaggacct attgtgaaga ctcatctcat ggagcttcag 16020
 aatgtggttg tcagcaaacc aatgaccgaa atccatcaca tgacggacgt ccagtgggtg 16080
 agcgaaacga aacaggaagc gcctatcttt cagagtcgtg agctccacac cggattccgg 16140
 caactacgtg ttgggcaggc ttcgccgtat tagagatatg ttgaggcaga cccatctgtg 16200
 ccactcgtac aattacgaga gttgtttttt ttgtgatttt cctagtttct cgttgatgggt 16260
 gagctcatat tctacatcgt atgggtctctc aacgtcgttt cctgtcatct gatatcccgt 16320
 catttgcatc cacgtgcgcc gcctcccgtg ccaagtccct aggtgtcatg cacgccaat 16380
 tggtggtggg gcgggctgcc ctgtgcttct taccgatggg tggagggtga gtttgggggt 16440
 ctccgcggcg atggtagtg gttgacggtt tgggtgtgggt tgacggcatt gatcaattta 16500
 cttcttgctt caaattcttt ggcagaaaac aattcattag attagaactg gaaaccagag 16560
 tgatgagacg gattaagtca gattccaaca gagttacatc tcttaagaaa taatgtaacc 16620
 ccttttagact ttatatattt gcaattaaaa aaataattta acttttagac tttatatata 16680
 gttttaataa ctaagtttaa ccactctatt atttatatcg aaactatttg tatgtctccc 16740
 ctctaaataa acttggtatt gtgtttacag aacctataat caaataatca atactcaact 16800
 gaagtttgat cagttaattg aagggtattaa cggccaaaat gcactagtat tatcaaccga 16860
 atagattcac actagatggc catttccatc aatatcatcg ccgttcttct tctgtccaca 16920
 tatcccctct gaaacttgag agacacctgc acttcattgt ccttattacg tgttacaaaa 16980
 tgaaacccat gcatccatgc aaactgaaga atggcgcaag aacccttccc ctccatttct 17040
 tatgtggcga ccatccattt caccatctcc cgctataaaa ccccccatc acttcaccta 17100
 gaacatcatc actacttgct tatccatcca aaagataccc acttttacia caattaccaa 17160
 caacaacaaa caacaacaaa cattacaatt acatttacia ttaccatacc atgccaccta 17220
 gcgctgctaa gcaaatggga gcttctactg gtgttcatgc tgggtgttact gactcttctg 17280
 ctttcaccag aaaggatggt gctgatagac ctgatctcac catcgttgga gattctgttt 17340
 acgatgctaa ggctttcaga tctgagcatc ctgggtgtgc tcatttcgtt tctttgttcg 17400
 gaggaagaga tgctactgag gctttcatgg aataccatag aagggttgg cctaagtcta 17460
 gaatgtctag attccacgtt ggatctcttg cttctactga ggaacctgtt gctgctgatg 17520

68/143

agggataacct tcaactttgt gctaggatcg ctaagatggg gccttctgtt tcttctggat	17580
tcgctcctgc ttcttactgg gttaaggctg gacttatcct tggatctgct atcgctcttg	17640
aggcttacat gctttacgct ggaaagagac ttctcccttc tatcgttcctt ggatggcttt	17700
tcgctcttat cggctcttaac atccagcatg atgctaacca tgggtgctttg tctaagtctg	17760
cttctgttaa ccttgctctt ggactttgtc aggattggat cggaggatct atgacccctt	17820
ggcttcaaga gcatgttggt atgcaccacc tccacactaa cgatgttgat aaggatcctg	17880
atcaaaaggc tcacgggtgct cttagactca agcctactga tgcttgggtca cctatgcatt	17940
ggcttcagca tctttacctt ttgcctgggtg agactatgta cgctttcaag cttttgttcc	18000
tcgacatctc tgagcttggt atgtggcggt gggagggtga gcctatctct aagcttgctg	18060
gatacctctt tatgccttct ttgcttctca agcttacctt ctgggctaga ttcgttgctt	18120
tgctctttta ccttgctcct tctgttcata ctgctgtgtg tatcgctgct actgttatga	18180
ctggatcttt ctacctgct ttcttcttct tcatctccca caacttcgag ggtgttgctt	18240
ctgttggacc tgatggatct atcacttcta tgactagagg tgctagcttc cttaaagagac	18300
aagctgagac ttcttctaac gttggaggac ctcttcttgc tactcttaac ggtggactca	18360
actaccaaat tgagcatcac ttgttccta gagttcacca tggattctac cctagacttg	18420
ctcctcttgt taaggctgag cttgaggcta gaggaatcga gtacaagcac taccctacta	18480
tctggtctaa ccttgcttct accctcagac atatgtacgc tcttggaaga aggcctagat	18540
ctaaggctga gtaatgacaa gcttatgtga cgtgaaataa taacggtaaa atatatgtaa	18600
taataataat aataaagcca caaagtgaga atgaggggaa ggggaaatgt gtaatgagcc	18660
agtagccggt ggtgctaatt ttgtatcgta ttgtcaataa atcatgaatt ttgtggtttt	18720
tatgtgtttt tttaaactat gaatttttaa ttttataaaa taatctccaa tcggaagaac	18780
aacattccat atccatgcat ggatgtttct ttacccaaat ctagttcttg agaggatgaa	18840
gcatcaccga acagttctgc aactatccct caaaagcttt aaaatgaaca acaaggaaca	18900
gagcaacgtt ccaaagatcc caaacgaaac atattatcta tactaatact atattattaa	18960
ttactactgc ccggaatcac aatccctgaa tgattcctat taactacaag ccttggtggc	19020
ggcggagaag tgatcggcgc ggcgagaagc agcggactcg gagacgaggc cttggaagat	19080
ctgagtcgaa cgggcagaat cagtattttc cttcgacgtt aattgatcct acactatgta	19140
ggtcatatcc atcgttttta tttttggcca ccattcaatt ctgtcttgcc tttagggatg	19200
tgaatatgaa cggccaaagg aagagaataa aaataatcca aattaaagca agagaggcca	19260
agtaagataa tccaaatgta cacttgtcat tgccaaaatt agtaaaatac tcggcatatt	19320
gtattcccac acattattaa aataccgtat atgtattggc tgcatttgca tgaataatac	19380
tacgtgtaag cccaaaagaa cccacgtgta gcccatgcaa agttaacact cagacccca	19440
ttctcagtc tccactatat aaaccacca tcccaaatct caccaaacc accacacaac	19500
tcacaactca ctctcacacc ttaaagaacc aatcaccacc aaaaatttta caacaattac	19560
caacaacaac aaacaacaaa caacattaca attacattta caattacat accatgagcg	19620

69/143

ctgttacgt tactggatct gatcctaaga acagaggatc ttctagcaac accgagcaag	19680
aggttccaaa agttgctatc gataccaacg gaaacgtggt ctctgttcct gatttcacca	19740
tcaaggacat ccttgagct atccctcatg agtgttacga gagaagattg gctacctctc	19800
tctactacgt gttcagagat atcttctgca tgcttaccac cggatacctt acccataaga	19860
tcctttaccc tctcctcatc tcttacacct ctaacagcat catcaagttc actttctggg	19920
ccctttacac ttacgttcaa ggacttttcg gaaccggaat ctgggttctc gctcatgagt	19980
gtggacatca agctttctct gattacggaa tcgtgaacga tttcgttggg tggacccttc	20040
actcttacct tatgggttct tactttcagct ggaagtactc tcatggaaag caccataagg	20100
ctactggaca catgaccaga gatatggttt tcgttcctgc caccaaagag gaattcaaga	20160
agtctaggaa cttcttcggt aacctcgtg agtactctga ggattctcca cttagaacc	20220
tttacgagct tcttggtcaa caacttgag gatggatcgc ttacctctc gttaacgtta	20280
caggacaacc ttaccctgat gttccttctt ggaaatggaa ccacttctgg cttacctctc	20340
cacttttcga gcaaagagat gctctctaca tcttccttctc tgatcttggg atcctcacc	20400
agggaatcgt tcttactctt tgggtacaaga aattcggagg atgggtccctt ttcatcaact	20460
ggttcgttcc ttacatctgg gtttaaccact ggctcgtttt catcacattc cttcagcaca	20520
ctgatcctac tatgcctcat tacaacgctg aggaatggac tttcgtctag ggtgctgctg	20580
ctactatcga tagaaagtcc ggattcatcg gacctcacat cttccatgat atcatcgaga	20640
ctcatgtgct tcaccactac tgttctagga tcccattcta caacgctaga cctgcttctg	20700
aggctatcaa gaaagtattg ggaaagcact acaggtctag cgacgagAAC atgtggaagt	20760
cactttggaa gtctttcagg tcttgccaat acgttgacgg tgataacggg gttctcatgt	20820
tccgtaacat caacaactgc ggagttggag ctgctgagaa gtaatgaagg ggtgatcgat	20880
tatgagatcg tacaagaca ctgctaggtg ttaaggatgg ataataataa taataatgag	20940
atgaatgtgt ttttaagttag tgtaacagct gtaataaaga gagagagaga gagagagaga	21000
gagagagaga gagagagaga gagagagagg ctgatgaaat gttatgtatg tttcttggtt	21060
tttaaaataa atgaaagcac atgctcgtgt ggttctatcg aattattcgg cggttcctgt	21120
gggaaaaagt ccagaagggc cgccgcagct actactacaa ccaaggccgt ggaggagggc	21180
aacagagcca gcacttcgat agctgctgct atgatcttaa gcaattgagg agcgagtgc	21240
catgcagggg actggagcgt gcaatcggcc agatgaggca ggacatccag cagcagggac	21300
agcagcagga agttgagagg tgggtcccatc aatctaaaca agtcgctagg gaccttcgg	21360
gacagtgcgg caccagcct agccgatgcc agctccaggg gcagcagcag tctgcatggt	21420
tttgaagtgg tgatcgatga gatcgtataa agacactgct aggtgttaag gatgggataa	21480
taagatgtgt ttttaagtc taaaccgtaat aaaaagagag agaggctgat ggaatgttat	21540
gtatgtatgt ttcttggttt ttaaaattaa atggaaagca catgctcgtg tgggttctat	21600
ctcgattaa aatcccaatt atatttggtc taatttagtt tggattgag taaaacaaat	21660

70/143

tcgaaccaaa ccaaaatata aatatatagt ttttatatat atgcctttaa gactttttat	21720
agaattttct ttaaaaaata tctagaaata tttgcgactc ttctggcatg taatatttcg	21780
ttaaatatga agtgctccat ttttattaac tttaaataat tggttgtacg atcactttct	21840
tatcaagtgt tactaaaatg cgtcaatctc tttgttcttc catattcata tgtcaaaatc	21900
tatcaaaatt cttatatatc tttttcgaat ttgaagtga atttcgataa tttaaaatta	21960
aatagaacat atcattatct aggtatcata ttgattttta tacttaatta ctaaatttgg	22020
ttaactttga aagtgtacat caacgaaaaa ttagtcaaac gactaaaata aataaatatc	22080
atgtgttatt aagaaaattc tcctataaga atattttaat agatcatatg tttgtaaaaa	22140
aaattaatct ttactaacac atatatttac ttatcaaaaa tttgacaaag taagattaaa	22200
ataatattca tctaacaaaa aaaaaaccag aaaatgctga aaacccggca aaaccgaacc	22260
aatccaaacc gatatagttg gtttggtttg attttgatat aaaccgaacc aactcgggtcc	22320
atttgcacc ctaatcataa tagctttaat atttcaagat attattaagt taacgttgct	22380
aatatcctgg aaattttgca aaatgaatca agcctatatg gctgtaatat gaatttaaaa	22440
gcagctcgat gtggtggtaa tatgtaattt acttgattct aaaaaaatat cccaagtatt	22500
aataatttct gctaggaaga aggttagcta cgatttacag caaagccaga atacaaagaa	22560
ccataaagtg attgaagctc gaaatatacg aaggaacaaa tttttttaa aaaatacgca	22620
atgacttgga acaaaagaaa gtgatataat tttgttctt aaacaagcat cccctctaaa	22680
gaatggcagt tttcctttgc atgtaactat tatgctcct tcgttacaaa aattttggac	22740
tactattggg aacttcttct gaaaatagtc ctgcaggcta gtagattggt tggttggtt	22800
ccatgtacca gaaggcttac cctattagtt gaaagttgaa actttgttcc ctactcaatt	22860
cctagtgtg taaatgtatg tatatgtaat gtgtataaaa cgtagtactt aaatgactag	22920
gagtggttct tgagaccgat gagagatggg agcagaacta aagatgatga cataattaag	22980
aacgaatttg aaaggctctt aggtttgaat cctattcgag aatgtttttg tcaaagatag	23040
tggcgatttt gaaccaaaga aaacatttaa aaaatcgata tccggttacg ttcattgcaa	23100
tagaaagtgg tctaggatct gattgtaatt ttagacttaa agagtctctt aagattcaat	23160
cctggctgtg taaaaacta caaataatat attttagact atttggcctt aactaaactt	23220
ccactcatta ttactgagg ttagagaata gacttgcgaa taaacacatt cccgagaaat	23280
actcatgatc ccataattag tcagagggta tgccaatcag atctaagaac acacattccc	23340
tcaaatttta atgcacatgt aatcatagtt tagcacaatt caaaaataat gtagtattaa	23400
agacagaaat ttgtagactt ttttttggcg ttaaaagaag actaagtta tacgtacatt	23460
ttattttaag tggaaaaccg aaattttcca tcgaaatata tgaatttagt atatatattt	23520
ctgcaatgta ctattttgct attttggcaa ctttcagtgg actactactt tattacaatg	23580
tgtatggatg catgagtttg agtatacaca tgtctaaatg catgctttgt aaaacgtaac	23640
ggaccacaaa agaggatcca tacaataca tctcatagct tcctccatta ttttccgaca	23700
caaacagagc attttacaac aattaccaac aacaacaaac aacaacaaac attacaatta	23760

catttacaat taccatacca tggaatttgc tcaacctctc gttgctatgg ctcaagagca	23820
gtacgctgct atcgatgctg ttgttgctcc tgctatcttc tctgctaccg actctattgg	23880
atggggactc aagcctatct cttctgctac taaggatctc cctctcgttg aatctcctac	23940
ccctcttata ctttctctcc tcgcttactt cgctatcggt gggttctggac tcgtttaccg	24000
taaagtgttc cctagaaccg ttaagggaca ggatcctttc cttctcaagg ctcttatgct	24060
cgctcacaac gttttcctta tcggactcag cctttacatg tgcctcaagc tcgtttacga	24120
ggcttacgtg aacaagtact ctttctgggg aaacgcttac aaccctgctc aaaccgagat	24180
ggctaagggt atctggatct tctacgtgtc caagatctac gagttcatgg acaccttcat	24240
catgcttctc aagggaaacg ttaaccagggt ttccttcctc catgtttacc accacggatc	24300
tatctctgga atctggtgga tgatcactta tgctgctcca ggtggagatg cttacttctc	24360
tgctgctctc aactcttggg ttcattgtgt catgtacacc tactacttca tggctgctgt	24420
tcttcctaag gacgaaaaga ccaagagaaa gtacctttgg tggggaagat acctaccca	24480
gatgcaaag ttccagttct tcatgaacct tctccagggt gtttacctcc tctactcttc	24540
ttctccttac cctaagttca ttgctcaact cctcgttggt tacatggtta ccctcctcat	24600
gcttttcgga aacttctact acatgaagca ccacgcttct aagtataag ggccgccgcc	24660
atgtgacaga tcgaaggaag aaagtgtaat aagacgactc tcactactcg atcgctagt	24720
attgtcattg ttatatataa taatgttatc ttccacaact tatcgtaatg catgtgaaac	24780
tataacacat taatcctact tgtcatatga taacactctc cccatttaaa actcttgtca	24840
atttaaagat ataagattct ttaaagtatt aaaaaaata tattataaat tcaatcactc	24900
ctactaataa attattaatt attatttatt gattaaaaaa atacttatac taatttagtc	24960
tgaatagaat aattagattc tagcctgcag ggcggccgcg gatcccatgg agtcaaagat	25020
tcaaatagag gacctaacag aactcgccgt aaagactggc gaacagttca tacagagtct	25080
cttacgactc aatgacaaga agaaaatctt cgtcaacatg gtggagcacg acacacttgt	25140
ctactccaaa aatatcaaag atacagtctc agaagaccaa agggcaattg agacttttca	25200
acaaagggta atatccggaa acctcctcgg attccattgc ccagctatct gtcactttat	25260
tgtgaagata gtggaaaagg aagggtggctc ctacaaatgc catcattgcg ataaaggaaa	25320
ggccatcggt gaagatgcct ctgccgacag tgggtcccaa gatggacccc caccacgag	25380
gagcatcggt gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga	25440
tatctccact gacgtaaggg atgacgcaca atcccactat ctttcgcaag acccttcctc	25500
tatataagga agttcatttc atttgagag aacacggggg actgaattaa atatgagccc	25560
tgagaggcgt cctgttgaaa tcagacctgc tactgctgct gatatggctg ctgtttgtga	25620
tatcgtgaac cactacatcg agacttctac cgttaacttc agaactgagc ctcaaactcc	25680
tcaagagtgg atcgatgatc ttgagagact ccaagataga tacccttggc ttgttgctga	25740
ggttgaggggt gttgttgctg gaatcgctta cgctggacct tggaaggcta gaaacgctta	25800

72/143

cgattggact gttgagtcta ccgtttacgt ttcacacaga catcagagac ttggacttgg 25860
 atctaccctt tacactcacc ttctcaagtc tatggaagct cagggattca agtctgttgt 25920
 tgctgttata ggactcccta acgacccctt tgtttagactt catgaggctc ttggatacac 25980
 tgctagagga actcttagag ctgctggata caagcacggt ggatggcatg atgttggatt 26040
 ctggcaaaga gatttcgagc ttcctgctcc tcctagacct gttagaccag ttactcagat 26100
 ctgaatttgc gtgacggtc aaacatttgg caataaagtt tcttaagatt gaatcctgtt 26160
 gccggtcttg cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt 26220
 aacatgtaat gcatgacggt atttatgaga tgggttttta tgattagagt cccgcaatta 26280
 tacatttaac acgcataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc 26340
 gcggtgtcat ctatgttact agatcactag tgatgtacgg ttaaaaccac cccagtacat 26400
 taaaaacgtc cgcaatgtgt tattaagttg tctaagcgtc aatttgttta caccacaata 26460
 tatcctgccca ccagccagcc aacagctccc cgaccggcag ctcggcaca aatcaccact 26520
 cgatacaggc agcccatcag tcc 26543

<210> 7
 <211> 23760
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Trình tự nucleotit pGA7- mod_G

<400> 7
 tcctgtgggt ggcatgcaca tacaaatgga cgaacggata aaccttttca cgccctttta 60
 aatatccgat tattctaata aacgctcttt tctcttaggt ttaccgcga atatatcctg 120
 tcaaacactg atagttaaaa ctgaaggcgg gaaacgacaa tctgctagtg gatctccag 180
 tcacgacgtt gtaaaacggg cgccccgcgg aaagcttgcg gccgcggtac cgcccggtcg 240
 actcagatct tccaaggcct cgtctccgag tccgctgctt ctcgccgcgc cgatcacttc 300
 tccgccgcca acaaggcttg tagttaatag gaatcattca gggattgtga ttccgggcag 360
 tagtaattaa taatatagta ttagtataga taatatgttt cgtttgggat ctttggaacg 420
 ttgctctgtt ccttggtgtt cattttaaaag cttttgaggg atagttgcag aactgttcgg 480
 tgatgcttca tcctctcaag aactagattt gggtaaagaa acatccatgc atggatatgg 540
 aatgttggtt ttccgattgg agattatttt ataaaattta aaattcatga tttaaaaaaa 600
 cacataaaaa ccacaaaatt catgatttat tgacaatacg atacaaaatt agcaccaccg 660
 gctactggct cattaacat tcccccttc cctcattctc actttgtggc tttattatta 720
 ttattattac atatatttta ccgttattat ttcacgtcac ataagcttgt taattaatta 780
 tcaactgttc ttagcaccct tagacttgta cctcttacgg tagaagttgg cgaagagaac 840
 gaacatgttc accataacc agagttgaag gtaaggaagc cagaaagcgg tattctgaac 900
 gtaagcgggt tagatagagt gagaagcaca gatacagaac tgaagcatct ggatctgggt 960
 gatgtacttc ttccagaaac agttaactcc aagagcagca agtccgtagt aagagtacat 1020

gatcacgtgc acgaaggtgt taacagaaga tccgaagtag caatctccaa caggctcaag	1080
cttcataaca acgaaccatg accagatgag aagagtgtgg tggtagatgt gaaggaaaga	1140
aagctggtcg aacttcttcc tcatcaccat gaagaaggtg tcgagaagct caacgtactt	1200
gttgttgtag tgaagccaga taacctgaga gattccccaa gagttagcag tcatatcagg	1260
gatgtttccc caaaccttaa gtccctgagc cctatgagaa gtaacgaaaa ggtagatgca	1320
gtagctgttg aagaaggtct ggtagaagtt gtaaacgagc atagcgttct tgagtccgaa	1380
aggttgagat ctgttctgca tgatacgctt tccgaagtag atgaagagga ggtatccaat	1440
agttccgata gtaggtcccc agtactcaac ctgcatgaag taagcaggaa ctttctcaga	1500
agctgggata tcagggttag ccacattgta ggtaacgtat ccaagagttc cagcaagagc	1560
agcagggata gcgatagagg ccatggatg gtaattgtaa atgtaattgt aatgttgttt	1620
gttgttgtgt gtgtgttgta attgttgtaa aattaattaa gtgggtatct tttggatgga	1680
taagcaagta gtgatgatgt tctaggtgaa gtgatggggg tgttttatag cgggagatgg	1740
tgaaatggat ggtcgccaca taagaaatgg aggggaaggg ttcttgcgcc attcttcagt	1800
ttgcatggat gcatgggttt ctttttgtaa cacgtaataa ggacaatgaa gtgcaggtgt	1860
ctctcaagtt tcagagggga tatgtggaca gaagaagaac ggcgatgata ttgatggaaa	1920
tggccatcta gtgtgaatct attcggttga taatactagt gcattttggc cgттаатccc	1980
ttcaattaac tgcacaaact tcagttgagt attgattatt tgattatagg ttctgtaaac	2040
acaataccaa gtttatattag aggggagaca tacaataagt ttcgatataa ataatagagt	2100
ggttaaactt agttattaaa actatatata aagtctaaaa gttaaattat ttttttaatt	2160
gcaaataatat aaagtctaaa ggggttacat tatttcttaa gagatgtaac tctgttgtaa	2220
tctgacttaa tccgtctcat cactctggtt tccagttcta atctaataa ttgttttctg	2280
ccaaagaatt tgaagcaaga agtaaatgta tcaatgccgt caaccacac caaacctca	2340
accactacc atcgccgcgg agaccccaa actcaacctc caccatcgg taagaagcac	2400
agggcagccc gcaccaccac caatttggcg tgcacgacac ctagggactt ggcacgggag	2460
gcggcgcacg tggatgcaaa tgacgggata tcagatgaca ggaaacgacg ttgagagacc	2520
atacgatgta gaatatgagc tcaccatcaa cgagaaacta ggaaaatcac aaaaaaaca	2580
actctcgtaa ttgtacgagt ggcacagatg ggtctgcctc aacatatctc taatacggcg	2640
aagcctgccc aacacgtagt tgccggaatc cgggtgtggag ctacgactc tgaaagatag	2700
gcgcttctctg tttcgtttctg ctacccact ggacgtccgt catgtgatgg atttcgggtca	2760
ttggtttgct gacaaccaca ttctgaagct ccatgagatg agtcttcaca ataggtcctg	2820
ctcaataaccg tggagttatg gttgcaagtc cataacttgc cgttcgaata ttttgcgag	2880
ccagtcggac gggaaattggc gagctcggct gacacctata aaggccatga caagaagaac	2940
caaaagttct tccctaatagc tttcatgagg cttcgggtcg ttatggatgt cggaaaaccc	3000
ctcttgaagg aacgagacgt tattatgcat gacggtaaga ctattacttg tcagtataag	3060

74/143

tatgaaagat tacctgtctt ctgctttgtt tgtggattga ttggacacgt tgaaaaaaaa	3120
tgtgcacttc gatttcaata ctgagagatc gacttccctt ttctctagga gtattcgatc	3180
aaggcattaa catggaagga agctcaagct ctaaaggctt cacaatggaa cctgaaaaat	3240
ttcaacaagc ctaaactgaa atcgaagtca aatcacccaa ccgggagctc taaatcagca	3300
aacactcctc ctccacagta tccaatcatc gtgcacgatg ctccagggtat tgcaagccag	3360
gtattgcaag ctaggagtag gatagagacc ttaaacgtcg ttggtgtgaa gagtcatctt	3420
cagacctaata ggagatagat gtagacggcg gcacgaagac tctgaaacac cagaaaggct	3480
agtccaggat aaggatctgc tatcccaact gacctctcgt tagtcccaag gcctctcaac	3540
tagagcagga ggaaggatgg tcacaagact aggataatga tgtttccaat atgaacctga	3600
atgtccatag ctaatTTTTT tagtcttgct tctgcacttt ttgtttatta tgttctggtg	3660
actatgttat ttacccttgt ccgtatgctt gagggtagcc tagtagattg gttggttggt	3720
ttccatgtac cagaaggctt accctattag ttgaaagttg aaactttggt ccctactcaa	3780
ttcctagttg tgtaaagtga tgtatatgta atgtgtataa aacgtagtac ttaaagtact	3840
aggagtgggt cttgagaccg atgagagatg ggagcagaac taaagatgat gacataatta	3900
agaacgaatt tgaaaggctc ttaggtttga atcctattcg agaattgttt tgtcaaagat	3960
agtggcgatt ttgaaccaa gaaaacattt aaaaaatcag tatccggtta cgttcatgca	4020
aatagaaagt ggtctaggat ctgattgtaa ttttagactt aaagagtctc ttaagattca	4080
atcctggctg tgtacaaaac tacaaataat atattttaga ctatttggcc ttaactaaac	4140
ttccactcat tatttactga ggtagagaa tagacttgcg aataaacaca ttcccagaa	4200
atactcatga tcccataatt agtcagaggg tatgccaatc agatctaaga acacacattc	4260
cctcaaattt taatgcacat gtaatcatag tttagcaca ttcaaaaata atgtagtatt	4320
aaagacagaa atttgtagac ttttttttgg cgttaaaaga agactaagtt tatacgtaca	4380
ttttatttta agtggaacac cgaaattttc catcgaaata tatgaattta gtatatatat	4440
ttctgcaatg tactattttg ctattttggc aactttcagt ggactactac ttattacaa	4500
tgtgtatgga tgcattgatt tgagtataca catgtctaaa tgcattgctt gtaaacgta	4560
acggaccaca aaagaggatc catacaaata catctcatag cttcctccat tttttccga	4620
cacaaacaga gcattttaca acaattacca acaacaaca acaacaaca acattacaat	4680
tacatttaca attaccatac catggaattc gccagcctc ttgttgctat ggctcaagag	4740
caatacgctg ctatcgatgc tgttggtgct cctgctatct tctctgctac tgattctatc	4800
ggatggggac ttaagcctat ctcttctgct actaaggact tgcctcttgt tgagtctcct	4860
acacctctca tcctttcttt gcttgcttac ttcgctatcg ttggatctgg actcgtttac	4920
agaaagggtt tccctagaac cgtgaaggga caagatccat tccttttgaa ggctcttatg	4980
cttgctcaca acgtgttcct tatcggactt tctctttaca tgtgcctcaa gcttggtgac	5040
gaggcttacg ttaacaagta ctctttctgg ggaaacgctt acaaccctgc tcaaactgag	5100
atggctaagg ttatctggat ctctacgtg agcaagatct acgagttcat ggataccttc	5160

atcatgctcc tcaagggaaa tgtaaccag gttagcttcc ttcacgttta ccatcacgga	5220
tctatctctg gaatctggtg gatgattact tacgctgctc ctggtggtga tgcttacttc	5280
tctgctgctc ttaactcttg ggttcacgtg tgtatgtaca cctactatth tatggctgcc	5340
gtgcttccta aggacgagaa aactaagaga aagtacctct ggtggggaag ataccttact	5400
caaatgcaga tgttccagtt cttcatgaac cttctccagg ctgtttacct tctctactct	5460
tcatctcctt accctaagtt tatcgctcag ctctctgtgg tgtacatggg tactcttctc	5520
atgcttttctg gaaacttcta ctacatgaag caccacgcta gcaagtgatg aggcgcgccg	5580
ggccgccgcc atgtgacaga tcgaaggaag aaagtgtaat aagacgactc tcactactcg	5640
atcgctagtg attgtcattg ttatatataa taatgttatc tttcacaact tatcgtaatg	5700
catgtgaaac tataacacat taatcctact tgtcatatga taacactctc cccatttaaa	5760
actcttgtca atttaagat ataagattct ttaaattgatt aaaaaaata tattataaat	5820
tcaatcactc ctactaataa attattaatt attatttatt gattaaaaaa atacttatac	5880
taatttagtc tgaatagaat aattagattc tagtctcatc cccttttaaa ccaacttagt	5940
aaacgttttt ttttttaatt ttatgaagtt aagtttttac cttgttttta aaaagaatcg	6000
ttcataagat gccatgccag aacattagct acacgttaca catagcatgc agccgcggag	6060
aattgttttt cttcgccact tgtcactccc ttcaaacc taagagcttc tctctcacag	6120
cacacacata caatcacatg cgtgcatgca ttattacacg tgatcgccat gcaaatctcc	6180
tttatagcct ataaattaac tcatccgctt cactctttac tcaaaccaaa actcatcgat	6240
acaaacaaga ttaaaaacat acacgaggat cttttacaac aattaccaac aacaacaac	6300
aacaacaac attacaatta catttacaat taccatacca tgccccaag ggactcttac	6360
tcttatgctg ctctctcttc tgctcaactt cacgaagttg atactcctca agagcacgac	6420
aagaaagagc ttgttatcgg agatagggct tacgatgtta ccaacttcgt taagagacac	6480
cctggtggaa agatcattgc ttaccaagtt ggaactgatg ctaccgatgc ttacaagcag	6540
ttccatgtta gatctgctaa ggctgacaag atgcttaagt ctcttccttc tcgtcctgtt	6600
cacaagggat actctccaag aagggtgat cttatcgctg atttccaaga gttaccaag	6660
caacttgagg ctgagggaaat gttcgagcct tctcttcctc atgttgctta cagacttgct	6720
gaggttatcg ctatgcatgt tgctgggtgct gctcttatct ggcatggata cactttcgct	6780
ggaatcgcta tgcttgaggt tgttcagga agatgtggat ggcttatgca tgagggtgga	6840
cattactctc tcaactgaaa cattgctttc gacagagcta tccaagttgc ttgttacgga	6900
cttggtatgtg gaatgtctgg tgcttggtgg cgtaaccagc ataacaagca ccatgctact	6960
cctcaaaagc ttcagcacga tggtgatctt gatacccttc ctctcgttgc tttccatgag	7020
agaatcgctg ctaagggttaa gtctcctgct atgaaggctt ggctttctat gcaagctaag	7080
cttttcgctc ctgttaccac tcttcttggt gctcttggtt ggcagcttta ccttcacct	7140
agacacatgc tcaggactaa gcactacgat gagcttgcta tgctcggaat cagatacgga	7200

cttgttggat accttgctgc taactacggt gctggatacg ttctcgcttg ttaccttctt	7260
tacgttcagc ttggagctat gtacatcttc tgcaacttcg ctgtttctca tactcacctc	7320
cctgttggtg agcctaacga gcatgctact tgggttgagt acgctgctaa ccacactact	7380
aactgttctc catcttggtg gtgtgattgg tggatgtctt accttaacta ccagatcgag	7440
caccaccttt acccttctat gcctcaattc agacacccta agatcgctcc tagagttaag	7500
cagcttttcg agaagcacgg acttcactac gatgttagag gatacttcga ggctatggct	7560
gatactttcg ctaaccttga taacgttgcc catgctcctg agaagaaaat gcagtaatga	7620
gatcgttcaa acatttgga ataaagtttc ttaagattga atcctgttgc cggctttg	7680
atgattatca tataatttct gttgaattac gttaagcacg taataattaa catgtaatgc	7740
atgacgttat ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac	7800
gcgatagaaa acaaaatata gcgcgcaaac taggataaat tatcgcgcg	7860
atgttactag atcggctgat taaaaatccc aattatattt ggtctaattt agtttggtat	7920
tgagtaaaac aaattcgaac caaaccaaaa tataaatata tagtttttat atatatgcct	7980
ttaagacttt ttatagaatt ttctttaaaa aatatctaga aatatttgcg actcttctg	8040
catgtaatat ttcgttaaat atgaagtgc cattttttat taactttaaa taattgggtg	8100
tacgatcact ttcttatcaa gtgttactaa aatgcgtcaa tctctttggt cttccatatt	8160
catatgtcaa aatctatcaa aattcttata tatctttttc gaatttgaag tgaaatttcg	8220
ataatttaaa attaaataga acatatcatt atttaggtat catattgatt ttataactta	8280
attactaaat ttggttaact ttgaaagtgt acatcaacga aaaattagtc aaacgactaa	8340
aataaataaa tatcatgtgt tattaagaaa attctcctat aagaatattt taatagatca	8400
tatgtttgta aaaaaatta atttttacta acacatatat ttacttatca aaaatttgac	8460
aaagtaagat taaaataata ttcattctaac aaaaaaaaaa ccagaaaatg ctgaaaaccc	8520
ggcaaaaccg aaccaatcca aaccgatata gttgggtttg tttgattttg atataaaccg	8580
aaccaactcg gtccatttgc acccctaatac ataatagctt taatatttca agatattatt	8640
aagttaacgt tgtcaatatc ctggaaattt tgcaaaatga atcaagccta tatggctgta	8700
atatgaattt aaaagcagct cgatgtggtg gtaatatgta atttacttga ttctaaaaaa	8760
atatcccaag tattaataat ttctgctagg aagaaggtta gctacgattt acagcaaagc	8820
cagaatacaa agaaccataa agtgattgaa gctcgaaata tacgaaggaa caaatatttt	8880
taaaaaata cgcaatgact tggaacaaaa gaaagtgata tattttttgt tcttaaacia	8940
gcatccctc taaagaatgg cagttttcct ttgcatgtaa ctattatgct cccttcgtta	9000
caaaaatttt ggactactat tgggaacttc ttctgaaaat agtgatagaa cccacacgag	9060
catgtgcttt ccatttaatt ttaaaaacca agaaacatac atacataaca ttccatcagc	9120
ctctctctct ttttattacg gttaatgact taaaacacat cttattatcc catccttaac	9180
acctagcagt gtctttatac gatctcatcg atcaccactt caaaaccatg cagactgctg	9240
ctgcccctgg agctggcatc ggctaggctg ggtgccgcac tgtcccggaa ggtccctagc	9300

gacttgttta gattgatggg accacctctc aacttcctgc tgctgtccct gctgctggat	9360
gtcctgcctc atctggccga ttgcacgctc cagtcacctg catgtgcact cgctcctcaa	9420
ttgcttaaga tcatcgagc agctatcgaa gtgctggctc tgttgccctc ctccacggcc	9480
ttggtttag tagtagctgc cgccgccctt ctggactttt tcccacagga accgccgaat	9540
aattcgatag aaccacacga gcatgtgctt tcattttattt taaaaaccaa gaaacataca	9600
taacatttca tcagcctctc tctctctctc tctctctctc tctctctctc tctctctctc	9660
tctctcttta ttacagctgt tacactaact taaaacacat tcatctcatt attattatta	9720
ttatccatcc ttaacaccta gcagtgtctt tgtacgatct cataatcgat cacccttca	9780
tcaggatatcc ttaggcttca ctccaacggt gttgcagtta cggaacatgt acacaccatc	9840
atggttctca acgaactggc aagatctcca agttttcaa aggctaacc acatgttctc	9900
atcgggtgtgt ctgtagtgct ctcccataac tttcttgatg cactcggtag cttctctagc	9960
atggtagaat gggatccttg aaacgtagtg atggagcaca tgagtctcga tgatgtcatg	10020
gaagatgatt ccgaggattc cgaactctct atcgatagta gcagcagcac ccttagcgaa	10080
agtccactct tgagcatcgt aatgaggcat agaagaatcg gtgtgctgaa ggaaggtaac	10140
gaaaacaagc cagtggttaa caaggatcca aggacagaac catgtgatga aagtaggcca	10200
gaatccgaaa accttgtaag cgggtgtaa acgaagtgagg gtagcaagga ttccaagatc	10260
agaaagaacg atgtaccagt agtccttctt atcgaaaaca gggctagaag gccagtagtg	10320
agacttgaag aacttagaaa caccagggta aggttgtcca gtagcgttag tagcaaggta	10380
aagagaaagt cctccaagct gttggaacaa gagagcgaaa acagagtaga taggagtttc	10440
ctcagcgata tcgtgaaggc tggttaacttg gtgcttctct ttgaattcct cggcgggtga	10500
aggaacgaaa accatatctc tggatcatgtg tccagtagcc ttatggtgct tagcatgaga	10560
gaacttccag ctgaagtaag gaaccataac aagagagtgg agaaccatc caacgggtatc	10620
gttaacccat ccgtagttag agaaagcaga atgtccacac tcatgtccaa ggatccagat	10680
tccgaatccg aaacaagaga tagagaacac gtaagcagac caagcagcga atctaaggaa	10740
ttcgttaggg agaagagga tgtaggtaag tccaacgtaa gcgatagcag agatagccac	10800
gatattctct accacgtaag acatagactt cacgagagat ctctcgtaac agtgcttagg	10860
gatagcgtca aggatatcct tgatggtgta atctggcacc ttgaaaacgt ttccgaagg	10920
atcgatagcg gtcttttgct gcttgaaaga tgcaacgttt ccagaacgcc taacgggtctt	10980
agtagatccc tcaaggatct cagatccaga cacggtaacc ttagacatgg tatggtaatt	11040
gtaaatgtaa ttgtaatgtt gtttggtgtt tgttggtgtt ggtaattgtt gtaaaatttt	11100
tggtggtgat tggttcttta aggtgtgaga gtgagttgtg agttgtgtgg tgggtttggt	11160
gagattgggg atggtgggtt tatatagtgg agactgagga atggggctcg gagtgtaaac	11220
tttgcatggg ctacacgtgg gttcttttgg gcttacacgt agtattattc atgcaaatgc	11280
agccaataca tatacgggtat ttaataatg tgtgggaata caatatgccg agtattttac	11340

taattttggc aatgacaagt gtacatttgg attatcttac ttggcctctc ttgctttaat	11400
ttggattatt ttatttctct taccttggcc gttcatattc acatccctaa aggcaagaca	11460
gaattgaatg gtggccaaaa attaaaacga tggatatgac ctacatagtg taggatcaat	11520
taacgtcgaa ggaaaatact gatttctctca agcatacgga caagggtaaa taacatagtc	11580
accagaacat aataaacaaa aagtgcagaa gcaagactaa aaaaattagc tatggacatt	11640
caggttcata ttggaaacat cattatccta gtcttgtgac catccttcct cctgctctag	11700
ttgagaggcc ttgggactaa cgagagggtca gttgggtag cagatcctta tcctggacta	11760
gcctttctgg tgtttcagag tcttcgtgcc gccgtctaca tctatctcca ttaggtctga	11820
agatgactct tcacaccaac gacgtttaag gtctctatcc tactcctagc ttgcaatacc	11880
tggcttgcaa tacctggagc atcgtgcacg atgattggat actgtggagg aggagtgttt	11940
gctgatttag agctcccggg tgggtgattt gacttcgatt tcagtttagg cttgttgaaa	12000
tttttcagggt tccattgtga agcctttaga gcttgagctt ccttccatgt taatgccttg	12060
atcgaatact cctagagaaa agggaagtcg atctctgagt attgaaatcg aagtgcacat	12120
tttttttcaa cgtgtccaat caatccacaa acaaagcaga agacaggtaa tctttcatac	12180
ttatactgac aagtaatagt cttaccgtca tgcataataa cgtctcgttc cttcaagagg	12240
ggttttccga catccataac gaccgaagc ctcatgaaag cattagggaa gaacttttgg	12300
ttcttcttgt catggccttt atagggtgtca gccgagctcg ccaattcccg tccgactggc	12360
tccgcaaaat attcgaacgg caagttatgg acttgcaacc ataactccac ggtattgagc	12420
aggacctatt gtgaagactc atctcatgga gcttcagaat gtggttggtca gcaaaccaat	12480
gaccgaaatc catcacatga cggacgtcca gtgggtgagc gaaacgaaac aggaagcgcc	12540
tatctttcag agtcgtgagc tccacaccgg attccggcaa ctacgtgttg ggcaggcttc	12600
gccgtattag agatatgttg aggcagaccc atctgtgcca ctcgtacaat tacgagagtt	12660
gttttttttg tgattttcct agtttctcgt tgatgggtgag ctcatattct acatcgtatg	12720
gtctctcaac gtcgtttcct gtcattctgat atcccgtcat ttgcatccac gtgcgccgcc	12780
tcccgtgcca agtccctagg tgtcatgcac gccaaattgg tgggtggtgcg ggctgacctg	12840
tgcttcttac cgatgggtgg aggttgagtt tgggggtctc cgcggcgatg gtagtgggtt	12900
gacggtttgg tgtgggttga cggcattgat caatttactt cttgcttcaa attctttggc	12960
agaaaacaat tcattagatt agaactggaa accagagtga tgagacggat taagtcagat	13020
tccaacagag ttacatctct taagaaataa tgtaaccctt ttagacttta tatatttgca	13080
attaaaaaaa taatttaact tttagacttt atatatagtt ttaataacta agtttaacca	13140
ctctattatt tatatcgaaa ctatttgtat gtctcccctc taaataaact tggattgtg	13200
tttacagaac ctataatcaa ataatacaata ctcaactgaa gtttgtgcag ttaattgaag	13260
ggattaacgg ccaaaatgca ctagtattat caaccgaata gattcacact agatggccat	13320
ttccatcaat atcatcgccg ttcttcttct gtccacatat cccctctgaa acttgagaga	13380
cacctgcact tcattgtcct tattacgtgt tacaaaatga aacctatgca tccatgcaaa	13440

ctgaagaatg gcgcaagaac ctttcccctc catttcttat gtggcgacca tccatttcac	13500
catctcccg c tataaaacac ccccatcact tcacctagaa catcatcact acttgcttat	13560
ccatccaaaa gatacccact tttaacaaca ttaccaacaa caacaaacaa caaacaacat	13620
tacaattaca tttaacaatta ccataccatg ccacctagcg ctgctaagca aatgggagct	13680
tctactggtg ttcattgctgg tgttactgac tcttctgctt tcaccagaaa ggatgttgct	13740
gatagacctg atctcaccat cgttggagat tctgtttacg atgctaaggc tttcagatct	13800
gagcatcctg gtggtgctca tttcgtttct ttgttcggag gaagagatgc tactgaggct	13860
ttcatggaat accatagaag ggcttggcct aagtctagaa tgtctagatt ccacgttgga	13920
tctcttgctt ctactgagga acctgttgct gctgatgagg gataccttca actttgtgct	13980
aggatcgcta agatggtgcc ttctgtttct tctggattcg ctcttgcttc ttactgggtt	14040
aaggctggac ttatccttgg atctgctatc gctcttgagg cttacatgct ttacgctgga	14100
aagagacttc tcccttctat cgttcttgga tggcttttcg ctcttatcgg tcttaacatc	14160
cagcatgatg ctaaccatgg tgctttgtct aagtctgctt ctgttaacct tgctcttgga	14220
ctttgtcagg attggatcgg aggatctatg atcctttggc ttcaagagca tgttgttatg	14280
caccacctcc aactaacga tgttgataag gatcctgatc aaaaggctca cggtgctctt	14340
agactcaagc ctactgatgc ttggtcacct atgcattggc ttcagcatct ttaccttttg	14400
cctggtgaga ctatgtacgc tttcaagctt ttgttcctcg acatctctga gcttgttatg	14460
tggcgttggg agggtagacc tatctctaag cttgctggat acctctttat gccttctttg	14520
cttctcaagc ttaccttctg ggctagattc gttgctttgc ctctttacct tgctccttct	14580
gttcatactg ctgtgtgtat cgctgctact gttatgactg gatctttcta cctcgctttc	14640
ttcttcttca tctcccacaa cttcgagggt gttgcttctg ttggacctga tggatctatc	14700
acttctatga ctagaggatgc tagcttcctt aagagacaag ctgagacttc ttctaacggt	14760
ggaggacctc ttcttgctac tcttaacggt ggactcaact accaaattga gcatcacttg	14820
ttccctagag ttcacatgg attctaccct agacttgctc ctcttgtaa ggctgagctt	14880
gaggctagag gaatcgagta caagcactac cctactatct ggtctaacct tgcttctacc	14940
ctcagacata tgtacgtctt tggaagaagg cctagatcta aggctgagta atgacaagct	15000
tatgtgacgt gaaataataa cggtaaaata tatgtaataa taataataat aaagccacaa	15060
agtgagaatg aggggaagg gaaatgtgta atgagccagt agccggtggg gctaattttg	15120
tatcgtattg tcaataaatc atgaattttg tggtttttat gtgttttttt aaatcatgaa	15180
ttttaaattt tataaaataa tctccaatcg gaagaacaac attccatatc catgcatgga	15240
tgtttcttta cccaaatcta gttcttgaga ggatgaagca tcaccgaaca gttctgcaac	15300
tatccctcaa aagcttttaa atgaacaaca aggaacagag caacgttcca aagatcccaa	15360
acgaaacata ttatctatac taatactata ttattaatta ctactgcccg gaatcacaat	15420
ccctgaatga ttcctattaa ctacaagcct tgttggcggc ggagaagtga tcggcgcggc	15480

gagaagcagc ggactcggag acgaggcctt ggaagatctg agtcgaacgg gcagaatcag	15540
tattttcctt cgacgttaat tgatcctaca ctatgtaggt catatccatc gttttaat	15600
ttggccacca ttcaattctg tcttgccctt agggatgtga atatgaacgg ccaaggtaag	15660
agaataaaaa taatccaaat taaagcaaga gaggccaagt aagataatcc aaatgtacac	15720
ttgtcattgc caaaattagt aaaatactcg gcatattgta ttcccacaca ttattaaaat	15780
accgtatatg tattggctgc atttgcatga ataatactac gtgtaagccc aaaagaaccc	15840
acgtgtagcc catgcaaagt taacactcac gaccccatc ctcagtctcc actatataaa	15900
cccaccatcc ccaatctcac caaaccacc acacaactca caactcactc tcacacctta	15960
aagaaccaat caccaccaa aattttacaa caattacaa caacaacaaa caacaacaa	16020
cattacaatt acatttacia ttaccatacc atgagcgctg ttaccgttac tggatctgat	16080
cctaagaaca gaggatcttc tagcaacacc gagcaagagg ttccaaaagt tgctatcgat	16140
accaacggaa acgtgttctc tgttcctgat ttcaccatca aggacatcct tggagctatc	16200
cctcatgagt gttacgagag aagattggct acctctctct actacgtgtt cagagatatt	16260
ttctgcatgc ttaccaccgg ataccttacc cataagatcc ttaccctct cctcatctct	16320
tacacctcta acagcatcat caagttcact ttctgggccc ttacactta cgttcaagga	16380
cttttcggaa ccggaatctg ggttctcgct catgagtgtg gacatcaagc tttctctgat	16440
tacggaatcg tgaacgattt cgttggatgg acccttcact cttaccttat ggttccttac	16500
ttcagctgga agtactctca tggaaagcac cataaggcta ctggacacat gaccagagat	16560
atggttttcg ttctgcccac caaagaggaa ttcaagaagt ctaggaactt cttcggtaac	16620
ctcgctgagt actctgagga ttctccactt agaacccttt acgagcttct tgttcaacaa	16680
cttggaggat ggatcgctta cctcttcggt aacgttacag gacaacctta ccctgatgtt	16740
ccttcttgga aatggaacca cttctggctt acctctccac ttttcgagca aagagatgct	16800
ctctacatct tcctttctga tcttggaaatc ctcaccacag gaatcgctt tactctttgg	16860
tacaagaaat tcggaggatg gtcccttttc atcaactggt tcgttcctta catctgggtt	16920
aaccactggc tcgttttcat cacattcctt cagcacactg atcctactat gcctcattac	16980
aacgctgagg aatggacttt cgctaagggt gctgctgcta ctatcgatag aaagtccgga	17040
ttcatcggac ctcacatctt ccatgatatt atcgagactc atgtgcttca ccactactgt	17100
tctaggatcc cattctacia cgctagacct gcttctgagg ctatcaagaa agttatggga	17160
aagcactaca ggtctagcga cgagaacatg tggaagtcac tttggaagtc tttcaggtct	17220
tgccaatacg ttgacggtga taacggtggt ctcattgtcc gtaacatcaa caactgcgga	17280
gttggagctg ctgagaagta atgaaggggt gatcgattat gagatcgtag aaagacactg	17340
ctaggtgtta aggatggata ataataataa taatgagatg aatgtgtttt aagttagtgt	17400
aacagctgta ataaagagag agagagagag agagagagag agagagagag agagagagag	17460
agagaggctg atgaaatgtt atgtatgttt cttgggtttt aaaataaatg aaagcacatg	17520
ctcgtgtggt tctatcgaat tattcggcgg ttctgtggg aaaaagtcca gaagggccgc	17580

cgagctact actacaacca aggccgtgga ggagggcaac agagccagca cttcgatagc	17640
tgctgcatg atcttaagca attgaggagc gagtgcacat gcaggggact ggagcgtgca	17700
atcggccaga tgaggcagga catccagcag cagggacagc agcaggaagt tgagaggtgg	17760
tcccatcaat ctaaacaagt cgctagggac cttccgggac agtgcggcac ccagcctagc	17820
cgatgccagc tccaggggca gcagcagtct gcatggtttt gaagtgggtga tcgatgagat	17880
cgtataaaga cactgctagg tgттаaggat gggataataa gatgtgtttt aagtcattaa	17940
ccgtaataaa aagagagaga ggctgatgga atgttatgta tgtatgtttc ttggttttta	18000
aaattaaatg gaaagcacat gctcgtgtgg gttctatctc gattaaaaat cccaattata	18060
tttggcttaa tttagtttggt tattgagtaa aacaaattcg aaccaaacca aaatataaat	18120
atatagtttt tatatatatg cttttaagac tttttataga attttcttta aaaaatatct	18180
agaaatatatt gcgactcttc tggcatgtaa tatttcgtta aatatgaagt gctccatttt	18240
tattaacttt aaataattgg ttgtacgatc actttcttat caagtgttac taaaatgcgt	18300
caatctcttt gttcttccat attcatatgt caaaatctat caaaattctt atatatcttt	18360
ttcgaatttg aagtgaaatt tcgataattt aaaattaaat agaacatatc attatttagg	18420
tatcatattg atttttatac ttaattacta aatttggtta actttgaaag tgtacatcaa	18480
cgaaaaatta gtcaaacgac taaaataaat aaatatcatg tgttattaag aaaatttctc	18540
tataagaata ttttaataga tcatatgttt gtaaaaaaaaa ttaattttta ctaacacata	18600
tatttactta tcaaaaattt gacaaagtaa gattaaaata atattcatct aacaaaaaaaa	18660
aaaccagaaa atgctgaaaa cccggcaaaa ccgaaccaat ccaaaccgat atagttgggt	18720
tggtttgatt ttgatataaa ccgaaccaac tcggtccatt tgcaccccta atcataatag	18780
ctttaatatt tcaagatatt attaagttaa cgttgtcaat atcctggaaa ttttgcaaaa	18840
tgaatcaagc ctatatggct gtaatatgaa tttaaaagca gctcgaatgt gtggtaatat	18900
gtaatttact tgattctaaa aaaatatccc aagtattaat aatttctgct aggaagaagg	18960
ttagctacga tttacagcaa agccagaata caaagaacca taaagtgatt gaagctcgaa	19020
atatacgaag gaacaaatat ttttaaaaaa atacgcaatg acttggaaca aaagaaagtg	19080
atatattttt tgttcttaaa caagcatccc ctctaaagaa tggcagtttt ctttgcag	19140
taactattat gtcctcttcg ttacaaaaat tttggactac tattgggaac ttcttctgaa	19200
aatagtcctg caggctagta gattgggttg ttggtttcca tgtaccagaa ggcttaccct	19260
attagttgaa agttgaaact ttgttcccta ctcaattcct agttgtgtaa atgtatgtat	19320
atgtaatgcg tataaaacgt agtacttaaa tgactaggag tggttcttga gaccgatgag	19380
agatgggagc agaactaaag atgatgacat aattaagaac gaatttgaaa ggctcttagg	19440
tttgaatcct attcgagaat gtttttgtca aagatagtggt cgattttgaa ccaaagaaaa	19500
catttaaaaa atcagtatcc ggttacgttc atgcaaataa aaagtgggtct aggatctgat	19560
tgtaatttta gacttaaaga gtctcttaag attcaatcct ggctgtgtac aaaactacaa	19620

82/143

ataatatatt ttagactatt tggccttaac taaacttcca ctcattatit actgagggtta	19680
gagaatagac ttgcgaataa acacattccc gagaaatact catgatccca taattagtca	19740
gagggtatgc caatcagatc taagaacaca cattccctca aattttaatg cacatgtaat	19800
catagtttag cacaattcaa aaataatgta gtattaaaga cagaaatttg tagacttttt	19860
tttggcgtta aaggaagact aagttttatac gtacattttta ttttaagtgg aaaaccgaaa	19920
ttttccatcg aaatatatga atttagtata tatattttctg caatgtacta ttttgctatt	19980
ttggcaactt tcagtggact actactttat tacaatgtgt atggatgcat gagtttgagt	20040
atacacatgt ctaaatgcat gctttgcaaa acgtaacgga ccacaaaaga ggatccatgc	20100
aaatacatct catagcttcc tccattatit tccgacacaa acagagcaga ctctagagga	20160
tccccccgtt ttacaacaat taccaacaac aacaaacaac aaacaacatt acaattacat	20220
ttacaattac catcccaaat cggcgcgcca tgtgtcctcc taagaccgat ggaagatctt	20280
ctcctagatc tcctctcacc aggtctaagt catctgctga ggctcttgat gctaaggatg	20340
cttctaccgc tcctgttgat cttagaccc ttgagcctca tgaacttgct gctaccttcg	20400
agactagatg ggtaggggtt gaggatgttg agtacgacgt gaccaacttc aaacatcctg	20460
gtggaagcgt gatcttctac atgcttgcta acactgggtgc tgatgctact gaggctttca	20520
aagaatttca catgcgtagc ctcaaggctt ggaagatgct tagagctttg ctttctagac	20580
ctgctgagat caagagatct gagtctgagg atgtctctat gcttgaggat ttcgctaggt	20640
ggagagctga acttgagagg gacggattct tcaagccttc tatcacccat gttgcttacc	20700
gtcttttga gcttcttgct actttcgctc ttggaaccgc tcttatgtac gctggatacc	20760
ctatcattgc tagcgttggtg tacgggtgctt tcttcggagc tagatgtgga tgggttcaac	20820
atgagggttg acacaactct cttaccggat ctgtgtacgt ggataagaga cttcaggcta	20880
tgacttgagg attcggactt tctaccagcg gagagatgtg gaaccagatg cataacaagc	20940
accatgctac ccctcagaaa gttagacacg acatggatct tgataccact cctgctgtgg	21000
ctttcttcaa caccgctgtg gaggataata gacctagggg attctctaga gcttgggcta	21060
gacttcaagc ttggaccttc gttcctgtta cttctggact tctcgttcag gctttctgga	21120
tctacgttct ccctcctaga cagggtgctca ggaagaagaa ctacgaggaa gcttcttgga	21180
tgctcgtttc tcacgttggtt agaaccgctg ttatcaagct tgctaccgga tactcttggc	21240
ctgttgctta ctggtggttc actttcgga actggatcgc ttacatgtac ctcttcgctc	21300
acttctctac ttctcacact cacctccctg ttgttccatc tgacaagcac cttagctggg	21360
ttactacgc tgttgatcac accgttgaca tcgatccttc tcgtggatac gttactggc	21420
ttatgggata ccttaactgc caggttatcc accatctctt ccctgatatg cctcaattca	21480
gacagcctga ggtgtcaaga agattcgtcc ctttcgctaa gaagtgggga ctcaactaca	21540
aggtgctctc ttactacggg gcttggaagg ctactttcag caacctcgac aaagtggac	21600
agcactacta cgttaacgga aaggctgaga aggtcactg atgattaatt aaatttgggc	21660
tcgaaccggg tcgagcaagc ttatgtgacg tgaaataata acggtaaaat atatgtaata	21720

ataataataa taaagccaca aagtgagaat gaggggaagg ggaaatgtgt aatgagccag	21780
tagccggtgg tgctaatttt gtatcgtatt gtcaataaat catgaatttt gtggttttta	21840
tgtgtttttt taaatcatga attttaaatt ttataaaaata atctccaatc ggaagaacaa	21900
cattccatat ccatgcatgg atgtttcttt acccaaatct agttcttgag aggatgaagc	21960
atcaccgaac agttctgcaa ctatccctca aaagctttaa aatgaacaac aaggaacaga	22020
gcaacgttcc aaagatccca aacgaaacat attatctata ctaatactat attattaatt	22080
actactgccc ggaatcaciaa tccctgaatg attcctatta actacaagcc ttgttggcgg	22140
cggagaagtg atcggcgcgg cgagaagcag cggactcgga gacgaggcct tggaagatct	22200
cctgcagggc ggccgcggat cccatggagt caaagattca aatagaggac ctaacagaac	22260
tcgccgtaaa gactggcgaa cagttcatac agagtctctt acgactcaat gacaagaaga	22320
aaatcttcgt caacatggtg gagcacgaca cacttgtcta ctccaaaaat atcaaagata	22380
cagtctcaga agaccaaagg gcaattgaga cttttcaaca aagggttaata tccggaaacc	22440
tcctcggatt ccattgcca gctatctgtc actttattgt gaagatagtg gaaaaggaag	22500
gtggctccta caaatgccat cattgcgata aaggaaaggc catcgttgaa gatgcctctg	22560
ccgacagtgg tcccaaagat ggacccccac ccacgaggag catcgtggaa aaagaagacg	22620
ttccaaccac gtcttcaaag caagtggatt gatgtgatat ctccactgac gtaagggatg	22680
acgcacaatc ccactatcct tcgcaagacc cttcctctat ataaggaagt tcatttcatt	22740
tggagagaac acgggggact gaattaaata tgagccctga gaggcgtcct gttgaaatca	22800
gacctgctac tgctgctgat atggctgctg tttgtgatat cgtgaaccac tacatcgaga	22860
cttctaccgt taacttcaga actgagcctc aaactcctca agagtggatc gatgatcttg	22920
agagactcca agatagatac ccttggtctg ttgctgaggt tgaggggtgtt gttgctggaa	22980
tcgcttacgc tggaccttg aaggctagaa acgcttacga ttggactgtt gagtctaccg	23040
tttacgtttc acacagacat cagagacttg gacttgatc taccctttac actcaccttc	23100
tcaagtctat ggaagctcag ggattcaagt ctgttggtgc tgttatcgga ctccctaacg	23160
atccttctgt tagacttcat gaggtctctg gatacactgc tagaggaact cttagagctg	23220
ctggatacaa gcacggtgga tggcatgatg ttggattctg gcaaagagat ttcgagcttc	23280
ctgctcctcc tagacctgtt agaccagtta ctcagatctg aatttgctg atcgttcaaa	23340
catttgcaa taaagtttct taagattgaa tcctgttgcc ggtcttgca tgattatcat	23400
ataatttctg ttgaattacg ttaagcatgt aataattaac atgtaatgca tgacgttatt	23460
tatgagatgg gtttttatga ttagagtccc gcaattatac atttaatacg cgatagaaaa	23520
caaaatatag cgcgcaaact aggataaatt atcgcgcgcg gtgtcatcta tgttactaga	23580
tcactagtga tgtacggtta aaaccacccc agtacattaa aaacgtccgc aatgtgttat	23640
taagttgtct aagcgtcaat ttgtttacac cacaatatat cctgccacca gccagccaac	23700
agctccccga ccggcagctc ggcacaaaat caccactcga tacaggcagc ccatcagtcc	23760

84/143

<210> 8
 <211> 11042
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> trình tự nucleotit pORE04+11ABGBEC_Cowpea_EPA_đoạn lỏng

<400> 8
 tcctgtggtt ggcattgcaca tacaaatgga cgaacggata aaccttttca cgccctttta 60
 aatatccgat tattctaata aacgctcttt tctcttaggt ttacccgcca atatatacctg 120
 tcaaacactg atagttttaa ctgaaggcgg gaaacgacaa tctgctagtg gatctcccag 180
 tcacgacgtt gtaaaacggg cgccctagaa tctaattatt ctattcagac taaattagta 240
 taagtatttt tttaataaat aaataataat taataattta ttagtaggag tgattgaatt 300
 tataatatat tttttttaat catttaaaga atcttatatc tttaaattga caagagtttt 360
 aaatggggag agtgttatca tatcacaagt aggattaatg tgttatagtt tcacatgcat 420
 tacgataagt tgtgaaagat aacattatta tatataacaa tgacaatcac tagcgatcga 480
 gtagtgagag tcgtcttatt acactttctt ccttcgatct gtcacatggc ggcggcccga 540
 attctcatca cttagaagcg tgggtgcttca tgtagtagaa gtttccgaaa agcatgagaa 600
 gagtcacat gtacaccacg aggagctgag cgataaactt agggtaaggt gaagaagagt 660
 agaggaggta cacagcctgg agaaggttca tgaagaactg gaacatttgc atctgggtga 720
 ggtatcttcc ccaccagagg tactttctct tagttttctc atccttaggg agcacagcag 780
 ccataaagta gtaggtgtac atgcacacgt gcacccaaga gttgagagca gcagagaaat 840
 aagcatcacc acctggagca gcgtaggtga tcatccacca gattccagag atagatccgt 900
 gatggtacac gtggaggaat gacacctggt tcacatttcc cttgaggagc atgatgaagg 960
 tatccatgaa ctgtagatc tttgacacgt agaaaatcca gatcaccta gccatctcgg 1020
 tctgagcagg gttgtaagcg tttccccaga aagagtactt gttcacgtaa gcctcgtaca 1080
 cgagcttgag gcacatgtag agtgagagtc cgatgaggaa cacgttatga gcgagcatga 1140
 gagccttgag caagaatgga tcctgtccct tcacagttct agggaaacacc tttctgtaca 1200
 cgagtccaga tcccacgata gcgaagtaag cgaggagaga caagataaga ggggtaggag 1260
 attccacgag agggagatcc ttagtagcag aagagatagg cttgagtccc catccgatag 1320
 aatcggtagc agagaagata gcaggagcca caacagcatc gatagcagcg tattgttctt 1380
 gagccatagc cacgagaggc tgagcaaatt ccatgaattc tggtcttctt tactctttgt 1440
 gtgactgagg tttggtctag tgctttgggtc atctatatat aatgataaca acaatgagaa 1500
 caagctttgg agtgatcgga gggcttagga tacatgagat tcaagtggac taggatctac 1560
 accgttggtat tttgagtgtg gatatgtgtg aggttaattt tacttggtaa cgccacaaa 1620
 ggcctaagga gaggtgttga gaccttatc ggcttgaacc gctggaataa tgccacgtgg 1680
 aagataattc catgaatctt atcgttatct atgagtgaat ttgtgtgatg gtggagtggg 1740
 gcttgctcat tttacttgcc tgggtggactt ggccctttcc ttatggggaa tttatatattt 1800

86/143

gcgaagaggt acatgtaagc gatccagttt ccgaaggtga accaccaata agcaacaggc	3900
caagagtatc cggtagcgag cttgataaca gcggttctca caacgtgaga cagcagcatc	3960
caagaagcct cttcgtagtt cttctttctg agcacctgtc taggatggag aacgtagatc	4020
cagaaagcct gcacgagaag tccagaagtc acaggaacga aagtccaagc ctgaagtcta	4080
gccaagctc tagaaaatcc cctaggcctg ttatcctcaa cagcgggtgtt gaagaaagcc	4140
acagcaggag tggatcgcag atccatatca tgcctcacct tttgtggggt tgcgtggtgc	4200
ttgttgtgca tctggttcca catctcacca gaggtagaaa gtccgaatcc gcaagtcata	4260
gcctggagcc tcttatccac atacacagat ccggtgagag agttatgacc accctcgtgt	4320
tgaaccatc cacatctagc tccgaagaaa gcaccgtaca ccacagaagc gataataggg	4380
tatccagcat acatgagagc agttccgaga gcgaaagtag caagaagctc gagaagtctg	4440
tatgccacgt gggtagataga aggcttgaag aatccatccc tctcaagctc agctctccac	4500
ctagcgaat cttcgagcat aggagcatcc tcagactcag acctcttgat ctgagctggt	4560
ctagaaggca aagccctaag catcttccaa gccttgagag atctcatgtg aaattctttg	4620
aaagcctcag tagcatcagc accggtgtta gcgagcatgt agaagatcac agaaccacca	4680
gggtgcttga agttagtaac atcgtactca acatcctcaa ctctcaccca tctagtctcg	4740
aaggtagcag ccaactcatg aggctcaaga gtcttgagat ccacaggagc agtagaagca	4800
tccttagcat cgagagcctc agcagatgac ttagacctgg taagaggtga cctaggagaa	4860
gatcttccat cagtcttttg agggcacatg cggccgctgt tcttctttac tctttgtgtg	4920
actgaggttt ggtctagtgc tttggtcatc tatatataat gataacaaca atgagaacaa	4980
gctttggagt gatcggaggg tctaggatac atgagattca agtggactag gatctacacc	5040
gttggtttt gagtgtggat atgtgtgagg ttaattttac ttggtaacgg ccacaaaggc	5100
ctaaggagag gtgttgagac cttatcggc ttgaaccgct ggaataatgc cacgtggaag	5160
ataattccat gaatcttatc gttatctatg agtgaaattg tgtgatggtg gagtgggtgct	5220
tgctcatttt acttgccctgg tggacttggc ctttcccta tggggaattt atattttact	5280
tactatagag ctttcatacc ttttttttac cttggattta gttaatatat aatggtatga	5340
ttcatgaata aaaatgggaa atttttgaat ttgtactgct aaatgcataa gattaggtga	5400
aactgtggaa tatatatattt tttcatttaa aagcaaaatt tgccttttac tagaattata	5460
aatatagaaa aatatataac attcaaataa aaatgaaaat aagaactttc aaaaaacaga	5520
actatgttta atgtgtaaag attagtcgca catcaagtca tctgttacia tatgttacia	5580
caagtcataa gccaacaaa gttagcacgt ctaaaataaac taaagagtcc acgaaaatat	5640
tacaaatcat aagcccaaca aggttattga tcaaaaaaaaa aaaacgcca acaaagctaa	5700
acaaagtcca aaaaaaactt ctcaagtctc catcttcctt tatgaacatt gaaaactata	5760
cacaaaacaa gtcagataaa tctctttctg ggcctgtctt cccaacctcc tacatcactt	5820
ccctatcgga ttgaatgttt tacttgtacc ttttccgttg caatgatatt gatagtatgt	5880
ttgtgaaaac taataggggt aacaatcgaa gtcatggaat atggatttgg tccaagattt	5940

tccgagagct	ttctagtaga	aagcccatca	ccagaaat	ttt actagtaaaa	taaatcacca	6000
attaggtttc	ttattatgtg	ccaaattcaa	tataattata	gaggatattt	caaataaaaa	6060
cgtatgaatg	ttattagtaa	atggtcaggt	aagacattaa	aaaaatccta	cgtcagatat	6120
tcaactttta	aaattcgatc	agtgtggaat	tgtacaaaaa	tttgggatct	actatatata	6180
tataatgctt	tacaacactt	ggatTTTTTT	ttggaggctg	gaatTTTTta	tctacatatt	6240
tgTTTTggcc	atgcaccaac	tcattgttta	gtgtaatact	ttgattttgt	caaatatatg	6300
tgttcgtgta	tatttgtata	agaatttctt	tgaccatata	cacacacaca	tatatatata	6360
tatatatata	ttatatatca	tgcactttta	attgaaaaaa	taatatatat	atatatagtg	6420
cattttttct	aacaaccata	tatgttgcca	ttgatctgca	aaaatactgc	tagagtaatg	6480
aaaaatataa	tctattgctg	aaattatctc	agatgttaag	attttcttaa	agtaaattct	6540
ttcaaatttt	agctaaaagt	cttgtaataa	ctaaagaata	atacacaatc	tcgaccacgg	6600
aaaaaaaaaca	cataataaat	ttgggcgcgc	cgcgatttgg	ctagagcagc	ttgccaacat	6660
gggtggagcac	gacactctcg	tctactccaa	gaatatcaaa	gatacagtct	cagaagacca	6720
aagggtctatt	gagacttttc	aacaaagggt	aatatcgggg	aacctcctcg	gattccattg	6780
cccagctatc	tgtcacttca	tcaaaaggac	agtagaaaag	gaagggtggc	cctacaaatg	6840
ccatcattgc	gataaaggaa	aggctatcgt	tcaagatgcc	tctgccgaca	gtggtcccaa	6900
agatggaccc	ccaccacga	ggagcatcgt	ggaaaaagaa	gacgttccaa	ccacgtcttc	6960
aaagcaagtg	gattgatgtg	ataacatggt	ggagcacgac	actctcgtct	actccaagaa	7020
tatcaaagat	acagtctcag	aagaccaaag	ggctattgag	acttttcaac	aaagggtaat	7080
atcgggaaac	ctcctcggat	tccattgccc	agctatctgt	cacttcatca	aaaggacagt	7140
agaaaaggaa	gggtggcacct	acaaatgcca	tcattgcgat	aaaggaaagg	ctatcgttca	7200
agatgcctct	gccgacagtg	gtcccaaaga	tggaccccca	cccacgagga	gcacgttgga	7260
aaaagaagac	gttccaacca	cgtcttcaaa	gcaagtggat	tgatgtgata	tctccactga	7320
cgtaagggat	gacgcacaat	cccactatcc	ttcgcaagac	cttcctctat	ataaggaagt	7380
tcattttcatt	tggagaggac	acgctgaaat	caccagtctc	tctctacaaa	tctatctctg	7440
cgatcgcatg	cctcctaggg	attcttactc	ttacgctgct	cctccatctg	ctcagctcca	7500
tgaagttgat	actcctcaag	agcacgataa	gaaagaactc	gtgatcggag	atagggttta	7560
cgatgtgacc	aacttcgtga	agagacaccc	tgggtggaaag	attatcgctt	accagggttg	7620
aactgatgct	accgatgctt	acaagcagtt	ccacgtgaga	tctgctaagg	ctgataagat	7680
gctcaagtct	ctcccatcta	ggcctgtgca	caagggatat	tctccaagaa	gggctgatct	7740
tatcgtgat	ttccaagagt	tcaccaagca	gcttgaggct	gagggaatgt	tcgaaccttc	7800
tctccctcat	gtggcttaca	gactcgctga	ggttatcgct	atgcatgttg	ctggtgctgc	7860
tctcatctgg	cacggatata	ctttcgctgg	aatcgctatg	ctcggagtgg	ttcagggaag	7920
atgtggatgg	cttatgcatg	agggtggaca	ctactctctc	accggaaaca	ttgctttcga	7980

tagggctatc	caggtggcat	gctatggact	tggatgtgga	atgtctggtg	cttggtggag	8040
aaaccagcat	aacaagcacc	atgctacccc	tcaaaagctc	cagcatgatg	tggatctcga	8100
tactctccct	ctcgtggctt	tccatgagag	aatcgctgct	aagggtgaagt	ctcctgctat	8160
gaaggcttgg	ctctctatgc	aggctaagct	tttcgctcct	gtgactactc	ttctcgttgc	8220
tcttggaatg	cagctctacc	tccatcctag	acacatgctc	aggaccaagc	actacgatga	8280
gcttgctatg	ctcggtatca	gatacggact	cgttggatac	ctcgctgcta	attacggtgc	8340
tggatacgtt	ctcgcttgct	accttcttta	cgttcagctc	ggagctatgt	acatcttctg	8400
caacttcgct	gtgtctcaca	ctcatctccc	tgtggttgaa	cctaacgagc	atgctacttg	8460
ggttgagtac	gctgctaacc	acactaccaa	ctgctctcca	tcttggtggt	gtgattggtg	8520
gatgtcttac	ctcaactacc	agatcgagca	ccacctctac	ccttctatgc	ctcagttcag	8580
acaccctaag	atcgctccta	gagtgaagca	gcttttcgag	aagcacggac	tccactacga	8640
tgtgagagga	tactttgagg	ctatggctga	taccttcgct	aacctcgata	atgtggctca	8700
cgctcctgag	aagaaaatgc	agtgatgagc	gatcgcgatc	gttcaaacat	ttggcaataa	8760
agtttcttaa	gattgaatcc	tgttgccggt	cttgcgatga	ttatcatata	atttctgttg	8820
aattacgtta	agcatgtaat	aattaacatg	taatgcatga	cgttatttat	gagatggggt	8880
tttatgatta	gagtcccga	attatacatt	taatacgca	tagaaaacaa	aatatagcgc	8940
gcaaaactagg	ataaattatc	gcgcgcggtg	tcatctatgt	tactagatcc	ctgcagggcg	9000
tattggctag	agcagcttgc	caacatggtg	gagcacgaca	ctctcgtcta	ctccaagaat	9060
atcaaagata	cagtctcaga	agaccaaagg	gctattgaga	cttttcaaca	aagggttaata	9120
tcgggaaacc	tcctcggatt	ccattgccca	gctatctgtc	acttcatcaa	aaggacagta	9180
gaaaaggaag	gtggcaccta	caaatgccat	cattgcgata	aaggaaaggc	tatcgttcaa	9240
gatgcctctg	ccgacagtgg	tcccaaagat	ggacccccac	ccacgaggag	catcgtggaa	9300
aaagaagacg	ttccaaccac	gtcttcaaag	caagtggatt	gatgtgataa	catggtggag	9360
cacgacactc	tcgtctactc	caagaatatc	aaagatacag	tctcagaaga	ccaaagggct	9420
attgagactt	ttcaacaaag	ggtaatatcg	ggaaacctcc	tcggattcca	ttgcccagct	9480
atctgtcact	tcatcaaaag	gacagtagaa	aaggaagggtg	gcacctacaa	atgccatcat	9540
tgcgataaag	gaaaggctat	cgttcaagat	gcctctgccg	acagtgggtcc	caaagatgga	9600
ccccaccca	cgaggagcat	cgtggaaaaa	gaagacgttc	caaccacgtc	ttcaaagcaa	9660
gtggattgat	gtgatatctc	cactgacgta	agggatgacg	cacaatccca	ctatccttcg	9720
caagaccttc	ctctatataa	ggaagtcat	ttcatttgga	gaggacacgc	tgaaatcacc	9780
agtctctctc	tacaaatcta	tctctctcga	gatgattgaa	caagatggat	tgcacgcagg	9840
ttctccggcc	gcttgggtgg	agaggctatt	cggctatgac	tgggcacaac	agacaatcgg	9900
ctgctctgat	gccgccgtgt	tccggctgtc	agcgcagggg	aggccgggttc	tttttgtcaa	9960
gaccgacctg	tccggtgccc	tgaatgaact	tcaagacgag	gcagcgcggc	tatcgtggct	10020
ggccacgacg	ggcgttcctt	gcgcagctgt	gctcgacgtt	gtcactgaag	cgggaagggg	10080

ctggctgcta ttgggcgaag tgccggggca ggatctcctg tcatctcacc ttgctcctgc 10140
 cgagaaagta tccatcatgg ctgatgcaat gcggcggctg catacgcttg atccggctac 10200
 ctgcccattc gaccaccaag cgaaacatcg catcgagcga gcacgtactc ggatggaagc 10260
 cggtcttgtc gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact 10320
 gttcgccagg ctcaaggcgc gcatgcccga cggcgaggat ctctgctgta ctcatggcga 10380
 tgcctgcttg ccgaatatca tgggtggaaa tggccgcttt tctggattca tcgactgtgg 10440
 ccggctgggt gtggcggacc gctatcagga catagcgctg gctaccctg atattgctga 10500
 agagcttggc ggcgaatggg ctgaccgctt cctcgtgctt tacgggtatcg ccgctcccga 10560
 ttcgcagcgc atcgcttct atcgcttct tgacgagttc ttctgaaacg cgtgatcgtt 10620
 caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt gcgatgatta 10680
 tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa tgcgatgacgt 10740
 tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa tacgcgatag 10800
 aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca tctatgttac 10860
 tagatcgacg tccgtacggt taaaaccacc ccagtacatt aaaaacgtcc gcaatgtggt 10920
 attaagttgt ctaagcgtca atttgtttac accacaatat atcctgccac cagccagcca 10980
 acagctcccc gaccggcagc tcggcacaaa atcaccactc gatacaggca gcccatcagt 11040
 cc 11042

<210> 9

<211> 1254

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện *Lachancea kluyveri* 12 desaturaza ở cây

<400> 9

atgagcgtg ttaccgttac tggatctgat cctaagaaca gaggatcttc tagcaacacc 60
 gagcaagagg ttccaaaagt tgctatcgat accaacggaa acgtgttctc tgttcctgat 120
 ttcaccatca aggacatcct tggagctatc cctcatgagt gttacgagag aagattggct 180
 acctctctct actacgtgtt cagagatatc ttctgcatgc ttaccaccgg ataccttacc 240
 cataagatcc tttaccctct cctcatctct tacacctcta acagcatcat caagttcact 300
 ttctgggccc tttacactta cgttcaagga cttttcgga ccggaatctg ggttctcgct 360
 catgagtgtg gacatcaagc tttctctgat tacggaatcg tgaacgattt cgttggatgg 420
 acccttcact cttaccttat gggttccttac ttcagctgga agtactctca tggaaagcac 480
 cataaggcta ctggacacat gaccagagat atgggttttcg ttcctgccac caaagaggaa 540
 ttcaagaagt ctaggaactt cttcggtaac ctctgtagt actctgagga ttctccactt 600
 agaacccttt acgagcttct tgttcaacaa cttggaggat ggatcgctta cctcttcgtt 660
 aacgttacag gacaacctta ccctgatgtt cttcttgga aatggaacca cttctggctt 720

90/143

```

acctctccac ttttcgagca aagagatgct ctctacatct tcctttctga tcttggaatc      780
ctcaccagg gaatcgttct tactctttgg tacaagaaat tcggaggatg gtcccttttc      840
atcaactggt tcgttcctta catctggggt aaccactggc tcgttttcat cacattcctt      900
cagcactg atcctactat gcctcattac aacgctgagg aatggacttt cgctaagggt      960
gctgctgcta ctatcgatag aaagtccgga ttcacggac ctcacatctt ccatgatatc     1020
atcgagactc atgtgcttca ccactactgt tctaggatcc cattctacaa cgctagacct     1080
gcttctgagg ctatcaagaa agttatggga aagcactaca ggtctagcga cgagaacatg     1140
tggaagtcac tttggaagtc tttcaggtct tgccaatacg ttgacggtga taacggtggt     1200
ctcatgttcc gtaacatcaa caactgcgga gttggagctg ctgagaagta atga           1254

```

<210> 10
 <211> 416
 <212> PRT
 <213> Lachancea kluyveri

<400> 10

Met Ser Ala Val Thr Val Thr Gly Ser Asp Pro Lys Asn Arg Gly Ser
1 5 10 15

Ser Ser Asn Thr Glu Gln Glu Val Pro Lys Val Ala Ile Asp Thr Asn
20 25 30

Gly Asn Val Phe Ser Val Pro Asp Phe Thr Ile Lys Asp Ile Leu Gly
35 40 45

Ala Ile Pro His Glu Cys Tyr Glu Arg Arg Leu Ala Thr Ser Leu Tyr
50 55 60

Tyr Val Phe Arg Asp Ile Phe Cys Met Leu Thr Thr Gly Tyr Leu Thr
65 70 75 80

His Lys Ile Leu Tyr Pro Leu Leu Ile Ser Tyr Thr Ser Asn Ser Ile
85 90 95

Ile Lys Phe Thr Phe Trp Ala Leu Tyr Thr Tyr Val Gln Gly Leu Phe
100 105 110

Gly Thr Gly Ile Trp Val Leu Ala His Glu Cys Gly His Gln Ala Phe
115 120 125

Ser Asp Tyr Gly Ile Val Asn Asp Phe Val Gly Trp Thr Leu His Ser
130 135 140

Tyr Leu Met Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Gly Lys His
145 150 155 160

His Lys Ala Thr Gly His Met Thr Arg Asp Met Val Phe Val Pro Ala
165 170 175

Thr Lys Glu Glu Phe Lys Lys Ser Arg Asn Phe Phe Gly Asn Leu Ala
180 185 190

Glu Tyr Ser Glu Asp Ser Pro Leu Arg Thr Leu Tyr Glu Leu Leu Val
195 200 205

Gln Gln Leu Gly Gly Trp Ile Ala Tyr Leu Phe Val Asn Val Thr Gly

91/143

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

<210> 11
 <211> 1251
 <212> ADN
 <213> Pichia pastoris

<400> 11
 atgtctaagg ttaccgtgtc tggatctgag atccttgagg gatctactaa gaccgtagg 60
 cgttctggaa acgttgcatac tttcaagcag caaaagaccg ctatcgatac cttcggaac 120
 gttttcaagg tgccagatta caccatcaag gatataccttg acgctatccc taagcactgt 180
 tacgagagat ctctcgtgaa gtctatgtct tacgtggtga gagatatcgt ggctatctct 240
 gctatcgctt acgttggact tacctacatc cctcttctcc ctaacgaatt ccttagattc 300
 gctgcttggt ctgcttacgt gttctctatc tcttgtttcg gattcggaat ctggatcctt 360
 ggacatgagt gtggacattc tgctttctct aactacggat gggttaacga taccgttgga 420
 tgggttctcc actctcttgt tatgggtcct tacttcagct ggaagtctc tcattgctaag 480
 caccataagg ctactggaca catgaccaga gatatggttt tcgttcctta caccgccgag 540
 gaattcaaag agaagcacca agttaccagc cttcacgata tcgctgagga aactcctatc 600
 tactctgttt tcgctctctt gttccaacag cttggaggac tttctcttta ccttgctact 660
 aacgctactg gacaacctta ccctggtggt tctaagttct tcaagtctca ctactggcct 720
 tctagccctg ttttcgataa gaaggactac tggatcatcg ttctttctga tcttggaatc 780

92/143

cttgctaccc tcacttctgt ttacaccgct tacaaggttt tcggattctg gcctactttc 840
 atcacatggg tctgtccttg gacacctgtt aaccactggc ttgttttcgt taccttcctt 900
 cagcacaccg attcttctat gcctcattac gatgctcaag agtggacttt cgctaagggg 960
 gctgctgcta ctatcgatag agagttcgga atcctcgga tcatcttcca tgacatcatc 1020
 gagactcatg tgctccatca ctacgtttca aggatcccat tctaccatgc tagagaagct 1080
 accgagtgca tcaagaaagt tatgggagag cactacagac acaccgatga gaacatgtgg 1140
 gttagccttt ggaaaacttg gagatcttgc cagttcgttg agaaccatga tgggtgtgtac 1200
 atgttccgta actgcaacaa cgttggagtg aagcctaagg atacctgatg a 1251

<210> 12

<211> 415

<212> PRT

<213> *Pichia pastoris*

<400> 12

Met Ser Lys Val Thr Val Ser Gly Ser Glu Ile Leu Glu Gly Ser Thr
1 5 10 15

Lys Thr Val Arg Arg Ser Gly Asn Val Ala Ser Phe Lys Gln Gln Lys
20 25 30

Thr Ala Ile Asp Thr Phe Gly Asn Val Phe Lys Val Pro Asp Tyr Thr
35 40 45

Ile Lys Asp Ile Leu Asp Ala Ile Pro Lys His Cys Tyr Glu Arg Ser
50 55 60

Leu Val Lys Ser Met Ser Tyr Val Val Arg Asp Ile Val Ala Ile Ser
65 70 75 80

Ala Ile Ala Tyr Val Gly Leu Thr Tyr Ile Pro Leu Leu Pro Asn Glu
85 90 95

Phe Leu Arg Phe Ala Ala Trp Ser Ala Tyr Val Phe Ser Ile Ser Cys
100 105 110

Phe Gly Phe Gly Ile Trp Ile Leu Gly His Glu Cys Gly His Ser Ala
115 120 125

Phe Ser Asn Tyr Gly Trp Val Asn Asp Thr Val Gly Trp Val Leu His
130 135 140

Ser Leu Val Met Val Pro Tyr Phe Ser Trp Lys Phe Ser His Ala Lys
145 150 155 160

His His Lys Ala Thr Gly His Met Thr Arg Asp Met Val Phe Val Pro
165 170 175

Tyr Thr Ala Glu Glu Phe Lys Glu Lys His Gln Val Thr Ser Leu His
180 185 190

Asp Ile Ala Glu Glu Thr Pro Ile Tyr Ser Val Phe Ala Leu Leu Phe
195 200 205

Gln Gln Leu Gly Gly Leu Ser Leu Tyr Leu Ala Thr Asn Ala Thr Gly
210 215 220

Gln Pro Tyr Pro Gly Val Ser Lys Phe Phe Lys Ser His Tyr Trp Pro
225 230 235 240

93/143

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

<210> 13

<211> 1392

<212> ADN

<213> Micromonas pusilla

<400> 13

```

atgtgcccgc cgaagacgga cggccgatcg tccccgcgat cgccgctgac ggcgagcaaa      60
tcctccgcgg aggcgctcga cgccaaggac gcgtcgaccg cgcccgtcga tctcaaaacg      120
ctcgagccgc acgagctcgc ggcgacgttc gagacgcgat ggggtgcgct ggaggacgtc      180
gagtacgacg tcacaaactt caaacacccg ggaggcagcg tgatattcta catgctcgcg      240
aacacgggcg cggacgccac ggaggcggttc aaggagttcc acatgcgatc gcttaaggcg      300
tggaagatgc tcagagcgct gccgtcgcgc cccgcggaga tcaaacgcag cgagagcgag      360
gacgcgccga tggttgagga ttctgcgcgg tggcgcgcgg agctcgaacg cgacgggttc      420
ttaagccct cgataacgca cgtcgcgtat cggttactcg agctcctcgc gaccttcgcc      480
ctcggcaccg ccctcatgta cgccgggtac ccgatcatcg cgtccgtcgt gtacggcgcg      540
ttcttcggcg ctcggtgcgg ttgggtccag cagcagggcg ggcacaactc gctcacgggg      600
tccgtctacg tcgacaagcg cctccaagcg atgacgtgcg gggttcgggct gtccacgagc      660
ggggagatgt ggaaccagat gcacaataag caccacgcga cgccgcagaa agtgaggcac      720
gacatggacc tggacacgac ccccgcggtg gcgtttttta acaccgccgt ggaggacaac      780
cggccgaggg ggttctcccg cgcgtgggct cggcttcagg cgtggacgtt cgtcccgggtg      840
acctccgggc tgctcgtcca ggcgttctgg atctacgtcc tgcacccgcg gcaggtgttg      900
  
```

94/143

cgaaagaaga actacgagga ggcgtcgtgg atgctcgtct ctcacgtcgt caggaccgcg 960
 gtgattaaac tcgcgacggg gtactcgtgg cccgtcgcgt actggtgggtt caccttcggc 1020
 aactggatcg cgtacatgta cctcttcgcg cacttctcca cgagccacac gcacctcccg 1080
 gtcgtgccct cggataagca cctgagctgg gtgaactacg cggtcgatca caccgtggac 1140
 atcgacccgt cgcgcgggta cgtgaactgg ttgatgggat atctgaactg ccagggtcatt 1200
 catcacctgt tcccggacat gccgcagttt cgccagccgg aggtgagccg gcggttcgtc 1260
 ccgttcgcga agaagtgggg gctgaactac aaggtgctgt cctattacgg cgcctggaag 1320
 gcgacgttct cgaacttggg taaggtcggg cagcactact acgtcaacgg caaggcggag 1380
 aaggcgcact ga 1392

<210> 14

<211> 1395

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện *Micromonas pusilla* 6-desaturaza ở cây (phiên bản 1)

<400> 14

atgtgccctc ctaagactga tggaagatct tctcctagat ctccacttac cagggtctaaa 60
 tcttctgctg aggctcttga tgctaaggat gcttctactg ctctgttga tcttaagact 120
 cttgagcctc atgagcttgc tgctactttc gagactagat gggtttagagt tgaggacgtt 180
 gagtacgatg tgactaactt caagcaccct ggtggatctg tgatcttcta catgcttgct 240
 aacactgggtg ctgatgctac tgaggctttc aaagaattcc acatgcgttc tctcaaggct 300
 tggaagatgc ttagagcttt gccttctaga cctgctgaga tcaagagatc tgagtctgag 360
 gatgctccta tgcttgagga tttcgctaga tggcgtgctg agcttgagag agatggattc 420
 ttcaagcctt ctatcaccca tgtggcttac agacttctcg agcttcttgc tacattcgct 480
 cttggaactg ctcttatgta cgctggatac cctatcattg cttctgttgt ttacggtgct 540
 ttcttcggag ctagatgtgg atgggttcaa catgaggggtg gacataactc tcttaccgga 600
 tctgtttacg tggacaagag acttcaggct atgacttggtg gattcggact ttctacttct 660
 ggtgagatgt ggaaccagat gcataacaag caccatgcta cccctcaaaa ggtagacac 720
 gatatggatc ttgataccac tcctgctgtg gctttcttca aactgctgt tgaggataac 780
 agacctagag gattctctag agcttgggct agacttcaag cttggacttt cgttcctgtt 840
 acctctggac ttcttgttca agctttcttg atctacgttc tccaccctag acaagttctc 900
 cgtaagaaga actacgaaga ggcttcttgg atgctcgttt ctcattgtgt tagaaccgct 960
 gttatcaagc ttgctactgg atactcttgg cctgttgctt actggtgggtt cactttcgga 1020
 aactggatcg cttacatgta ccttttcgct cacttctcta cttctcatac tcacctccct 1080
 gttgttccat ctgataagca cctttcttgg gttactacg ctgttgatca caccgttgat 1140
 atcgatcctt ctagaggata cgtgaactgg cttatgggat accttaactg tcagggttatc 1200
 caccacctct tcctgatat gcctcaattc agacagcctg aggttagcag aagattcggtt 1260

95/143

cctttcgcta agaagtgggg actcaactac aaggtgctct cttactacgg tgcttggaag 1320
 gctactttct ctaaccttga taaggtggga cagcactact acgttaacgg aaaggctgag 1380
 aaggctcact aatga 1395

<210> 15
 <211> 1395
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện *Micromonas pusilla* 6-desaturaza ở cây (phiên bản 2)

<400> 15
 atgtgtcctc ctaagaccga tggaagatct tctcctagat ctctctcac caggtctaag 60
 tcattctgctg aggctcttga tgctaaggat gcttctaccg ctctgttga tcttaagacc 120
 cttgagcctc atgaacttgc tgctaccttc gagactagat gggttagggg tgaggatggt 180
 gagtacgacg tgaccaactt caaacatcct ggtggaagcg tgatcttcta catgcttgct 240
 aacactgggtg ctgatgctac tgaggctttc aaagaatttc acatgcgtag cctcaaggct 300
 tggaagatgc ttagagcttt gccttctaga cctgctgaga tcaagagatc tgagtctgag 360
 gatgctccta tgcttgagga tttcgctagg tggagagctg aacttgagag ggacggattc 420
 ttcaagcctt ctatcaccca tgttgcttac cgtcttttgg agcttcttgc tactttcgct 480
 cttggaaccg ctcttatgta cgctggatac cctatcattg ctagcgttgt gtacgggtgct 540
 ttcttcggag ctagatgtgg atgggttcaa catgaggggtg gacacaactc tcttaccgga 600
 tctgtgtacg tggataagag acttcaggct atgacttgcg gattcggact ttctaccagc 660
 ggagagatgt ggaaccagat gcataacaag caccatgcta cccctcagaa agttagacac 720
 gacatggatc ttgataccac tcctgctgtg gctttcttca acaccgctgt ggaggataat 780
 agacctaggg gatttcttag agcttgggct agacttcaag cttggacctt cgttcctggt 840
 acttctggac ttctcgttca ggctttcttg atctacgttc tccatcctag acagggtgctc 900
 aggaagaaga actacgagga agcttcttgg atgctcgttt ctacagttgt tagaaccgct 960
 gttatcaagc ttgctaccgg atactcttgg cctggtgctt actggtggtt cactttcggga 1020
 aactggatcg cttacatgta cctcttcgct cacttctcta cttctcacac tcacctccct 1080
 gttgttccat ctgacaagca ccttagctgg gtttaactac ctggtgatca caccgttgac 1140
 atcgatcctt ctgctggata cgtttaactg cttatgggat accttaactg ccagggttatc 1200
 caccatctct tccctgatat gcctcaattc agacagcctg aggtgtcaag aagattcgctc 1260
 cctttcgcta agaagtgggg actcaactac aaggtgctct cttactacgg tgcttggaag 1320
 gctactttca gcaacctcga caaagttgga cagcactact acgttaacgg aaaggctgag 1380
 aaggctcact gatga 1395

<210> 16
 <211> 463
 <212> PRT
 <213> *Micromonas pusilla*

96/143

<400> 16

Met Cys Pro Pro Lys Thr Asp Gly Arg Ser Ser Pro Arg Ser Pro Leu
 1 5 10 15
 Thr Arg Ser Lys Ser Ser Ala Glu Ala Leu Asp Ala Lys Asp Ala Ser
 20 25 30
 Thr Ala Pro Val Asp Leu Lys Thr Leu Glu Pro His Glu Leu Ala Ala
 35 40 45
 Thr Phe Glu Thr Arg Trp Val Arg Val Glu Asp Val Glu Tyr Asp Val
 50 55 60
 Thr Asn Phe Lys His Pro Gly Gly Ser Val Ile Phe Tyr Met Leu Ala
 65 70 75 80
 Asn Thr Gly Ala Asp Ala Thr Glu Ala Phe Lys Glu Phe His Met Arg
 85 90 95
 Ser Leu Lys Ala Trp Lys Met Leu Arg Ala Leu Pro Ser Arg Pro Ala
 100 105 110
 Glu Ile Lys Arg Ser Glu Ser Glu Asp Ala Pro Met Leu Glu Asp Phe
 115 120 125
 Ala Arg Trp Arg Ala Glu Leu Glu Arg Asp Gly Phe Phe Lys Pro Ser
 130 135 140
 Ile Thr His Val Ala Tyr Arg Leu Leu Glu Leu Leu Ala Thr Phe Ala
 145 150 155 160
 Leu Gly Thr Ala Leu Met Tyr Ala Gly Tyr Pro Ile Ile Ala Ser Val
 165 170 175
 Val Tyr Gly Ala Phe Phe Gly Ala Arg Cys Gly Trp Val Gln His Glu
 180 185 190
 Gly Gly His Asn Ser Leu Thr Gly Ser Val Tyr Val Asp Lys Arg Leu
 195 200 205
 Gln Ala Met Thr Cys Gly Phe Gly Leu Ser Thr Ser Gly Glu Met Trp
 210 215 220
 Asn Gln Met His Asn Lys His His Ala Thr Pro Gln Lys Val Arg His
 225 230 235 240
 Asp Met Asp Leu Asp Thr Thr Pro Ala Val Ala Phe Phe Asn Thr Ala
 245 250 255
 Val Glu Asp Asn Arg Pro Arg Gly Phe Ser Arg Ala Trp Ala Arg Leu
 260 265 270
 Gln Ala Trp Thr Phe Val Pro Val Thr Ser Gly Leu Leu Val Gln Ala
 275 280 285
 Phe Trp Ile Tyr Val Leu His Pro Arg Gln Val Leu Arg Lys Lys Asn
 290 295 300
 Tyr Glu Glu Ala Ser Trp Met Leu Val Ser His Val Val Arg Thr Ala
 305 310 315 320
 Val Ile Lys Leu Ala Thr Gly Tyr Ser Trp Pro Val Ala Tyr Trp Trp
 325 330 335
 Phe Thr Phe Gly Asn Trp Ile Ala Tyr Met Tyr Leu Phe Ala His Phe
 340 345 350

97/143

Ser Thr Ser His Thr His Leu Pro Val Val Pro Ser Asp Lys His Leu
 355 360 365
 Ser Trp Val Asn Tyr Ala Val Asp His Thr Val Asp Ile Asp Pro Ser
 370 375 380
 Arg Gly Tyr Val Asn Trp Leu Met Gly Tyr Leu Asn Cys Gln Val Ile
 385 390 395 400
 His His Leu Phe Pro Asp Met Pro Gln Phe Arg Gln Pro Glu Val Ser
 405 410 415
 Arg Arg Phe Val Pro Phe Ala Lys Lys Trp Gly Leu Asn Tyr Lys Val
 420 425 430
 Leu Ser Tyr Tyr Gly Ala Trp Lys Ala Thr Phe Ser Asn Leu Asp Lys
 435 440 445
 Val Gly Gln His Tyr Tyr Val Asn Gly Lys Ala Glu Lys Ala His
 450 455 460

<210> 17
 <211> 1449
 <212> ADN
 <213> *Ostreococcus lucimarinus*

<400> 17
 atgtgctgcg aaacgaccga aggcacatcg cgaacgatgg cgaacgaacg cacgagctcg 60
 tcgtcgtcgc tgagcgaagg cggaacgccg acggtgacgg tcgggatggg aagcgaagac 120
 gcggggaaga agactcgaaa cgcgagcgtc acggcgtgga cgaaagagtt ggagccgcac 180
 gcgatcgcga agacgttcga acggcggtac gtgacgatcg aaggcgtgga atacgatgtg 240
 acggatttta agcatcccgg aggatcggtt atttattaca tgctgtcgaa cacgggagcg 300
 gacgcgacgg aggcttttaa agagtttcat tatcggtcga aaaaggcgcg caaggcgttg 360
 gcggcggttg cgcataagcc agtggaacgc gcgacgcggg aaccgatcga agatgaggcg 420
 atgctgaagg atttcgcgca gtggcgcaag gaattggagc gtgagggatt ttttaagccc 480
 tcgccggcgc acgtggcgta tcgattcgcc gagctcgcg cgatgttcgc gctcggcacg 540
 gcgttgatgc acgcgcgttg gcacgtcgct tccgtgatcg tgtactcgtg tttcttcggc 600
 gcgcgatgcg gttgggtgca gcacgagggg gggcacaatt cgttgactgg aaacatttgg 660
 tgggacaagc gaatccaagc cttcgccgcg gggttcggct tggcgtcgag tggcgacatg 720
 tggaacaaca tgcacaacaa gcatcacgcg acgccccaaa aggtgcgaca cgatatggat 780
 ctcgacacca ctcccacggt ggcgttcttc aactccgcgg ttgaagaaaa tcgcccgcgg 840
 ggattcagta agttgtgggt gcgccttcaa gcgtggacct tcgtgcccgt gacgtccggt 900
 atggttttgt tcttctggat gttcgtcttg caccgcgta acgcgctgcg acgcaaaagc 960
 ttcgaagaag cggcttggat gttttccgcg cacgtcattc gcacggcggt tatcaaagcc 1020
 gtcaccggct actcctggat cgctcgtac ggcttggtcg cggcgacgat gtgggcgagc 1080
 ggatgttact tgttcgcgca cttttccacg tctcacacgc acttggtatgt cgtgccgagc 1140
 gataaacacc tctcgtgggt gcgatacgcc gtcgatcaca cgatcgacat caatccgaac 1200
 aacagcgtcg tcaactgggt gatgggctac ttgaactgcc aagtcatcca tcacctgttc 1260

98/143

ccggatatgc ctcaattccg ccaacccgaa gtctcccgcc gattcgtccc gtttgcaag 1320
 aagtggaaact taaactacaa ggtcttgacg tattatgggg cctggaaggc gacgttcggc 1380
 aacttgaacg acgtcgggaa gcactattac gtgcacggat ctcaagcgcgt caaatcaaag 1440
 tcggcgtga 1449

<210> 18
 <211> 1449
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện *Ostreococcus lucimarinus* 6-desaturaza ở cây

<400> 18
 atgtgtgttg agactactga gggaacctct agaactatgg ctaacgagag gacctcttct 60
 tcttcttcac tctctgaggg tggaactcct actgttactg tgggaatggg atctgaggat 120
 gctggaaaga aaaccagaaa cgcttctgtt actgcttgga ccaaagagct tgagcctcac 180
 gctatcgcta agaccttcga gagaagatac gttaccatcg aggggtgttg gtacgatgtg 240
 accgatttca aacaccctgg tggatctgtg atctactaca tgctctctaa cactggtgct 300
 gatgctactg aggctttcaa agagtccac taccgttcta agaaggctag aaaggctctt 360
 gctgctcttc ctcaacaggc tgttgatgct gctactagag agcctattga ggacgaggct 420
 atgcttaagg atttcgctca gtggagaaaa gagttggaga gagagggatt cttcaagcct 480
 tctcctgctc atgttgctta ccgtttcgct gaactcgctg ctatgttcgc tcttggaacc 540
 gctcttatgc atgctagatg gcacgttgct agcgttatcg tgtactcctg tttcttcgga 600
 gctagatgtg gatgggttca acatgagggt ggacacaact ctcttaccgg aaacatctgg 660
 tgggataaga gaatccaagc tttcgctgct ggattcggac ttgcttcttc tggtgacatg 720
 tggaacaaca tgcacaacaa gcaccatgct actcctcaga aagtgaagaca cgatatggat 780
 cttgatacca cccctaccgt tgctttcttc aactctgctg tggaggaaaa cagacctagg 840
 ggattctcta agctttggct cagacttcaa gcttgacatc tcgttcctgt tacctctgga 900
 atggtgctct tcttctggat gttcgttctc catcctagaa acgctctccg tcgtaagtct 960
 ttcgaagagg ctgcttggat gttctctgct cacgttatca gaaccgctgt tatcaaggct 1020
 gttaccggat actcttggat cgctagctac ggacttttcg ctgctactat gtgggcttct 1080
 ggatgctacc ttttcgctca cttctctact tctcacaccc acctcgatgt tgttccatct 1140
 gataagcacc ttagctgggt taggtacgct gttgatcaca ccatcgacat caaccctaac 1200
 aactctgttg tgaactggct tatgggatac cttaactgcc aggttatcca ccatctcttc 1260
 cctgatatgc ctcaattcag acagcctgag gtgtcaagaa gattcgtccc tttcgctaag 1320
 aagtggaacc tcaactacaa ggtgctcact tactacggtg cttggaaggc tactttcgga 1380
 aacctcaacg atgttggaag gcactactac gttcacggat ctcaagagag gaagagcaag 1440
 agcgcttga 1449

<210> 19

99/143

<211> 482
 <212> PRT
 <213> *Ostreococcus lucimarinus*

<400> 19

Met Cys Val Glu Thr Thr Glu Gly Thr Ser Arg Thr Met Ala Asn Glu
 1 5 10 15
 Arg Thr Ser Ser Ser Ser Ser Leu Ser Glu Gly Gly Thr Pro Thr Val
 20 25 30
 Thr Val Gly Met Gly Ser Glu Asp Ala Gly Lys Lys Thr Arg Asn Ala
 35 40 45
 Ser Val Thr Ala Trp Thr Lys Glu Leu Glu Pro His Ala Ile Ala Lys
 50 55 60
 Thr Phe Glu Arg Arg Tyr Val Thr Ile Glu Gly Val Glu Tyr Asp Val
 65 70 75 80
 Thr Asp Phe Lys His Pro Gly Gly Ser Val Ile Tyr Tyr Met Leu Ser
 85 90 95
 Asn Thr Gly Ala Asp Ala Thr Glu Ala Phe Lys Glu Phe His Tyr Arg
 100 105 110
 Ser Lys Lys Ala Arg Lys Ala Leu Ala Ala Leu Pro His Lys Pro Val
 115 120 125
 Asp Ala Ala Thr Arg Glu Pro Ile Glu Asp Glu Ala Met Leu Lys Asp
 130 135 140
 Phe Ala Gln Trp Arg Lys Glu Leu Glu Arg Glu Gly Phe Phe Lys Pro
 145 150 155 160
 Ser Pro Ala His Val Ala Tyr Arg Phe Ala Glu Leu Ala Ala Met Phe
 165 170 175
 Ala Leu Gly Thr Ala Leu Met His Ala Arg Trp His Val Ala Ser Val
 180 185 190
 Ile Val Tyr Ser Cys Phe Phe Gly Ala Arg Cys Gly Trp Val Gln His
 195 200 205
 Glu Gly Gly His Asn Ser Leu Thr Gly Asn Ile Trp Trp Asp Lys Arg
 210 215 220
 Ile Gln Ala Phe Ala Ala Gly Phe Gly Leu Ala Ser Ser Gly Asp Met
 225 230 235 240
 Trp Asn Asn Met His Asn Lys His His Ala Thr Pro Gln Lys Val Arg
 245 250 255
 His Asp Met Asp Leu Asp Thr Thr Pro Thr Val Ala Phe Phe Asn Ser
 260 265 270
 Ala Val Glu Glu Asn Arg Pro Arg Gly Phe Ser Lys Leu Trp Leu Arg
 275 280 285
 Leu Gln Ala Trp Thr Phe Val Pro Val Thr Ser Gly Met Val Leu Phe
 290 295 300
 Phe Trp Met Phe Val Leu His Pro Arg Asn Ala Leu Arg Arg Lys Ser
 305 310 315 320
 Phe Glu Glu Ala Ala Trp Met Phe Ser Ala His Val Ile Arg Thr Ala
 325 330 335

100/143

Val Ile Lys Ala Val Thr Gly Tyr Ser Trp Ile Ala Ser Tyr Gly Leu
 340 345 350
 Phe Ala Ala Thr Met Trp Ala Ser Gly Cys Tyr Leu Phe Ala His Phe
 355 360 365
 Ser Thr Ser His Thr His Leu Asp Val Val Pro Ser Asp Lys His Leu
 370 375 380
 Ser Trp Val Arg Tyr Ala Val Asp His Thr Ile Asp Ile Asn Pro Asn
 385 390 395 400
 Asn Ser Val Val Asn Trp Leu Met Gly Tyr Leu Asn Cys Gln Val Ile
 405 410 415
 His His Leu Phe Pro Asp Met Pro Gln Phe Arg Gln Pro Glu Val Ser
 420 425 430
 Arg Arg Phe Val Pro Phe Ala Lys Lys Trp Asn Leu Asn Tyr Lys Val
 435 440 445
 Leu Thr Tyr Tyr Gly Ala Trp Lys Ala Thr Phe Gly Asn Leu Asn Asp
 450 455 460
 Val Gly Lys His Tyr Tyr Val His Gly Ser Gln Arg Val Lys Ser Lys
 465 470 475 480
 Ser Ala

<210> 20
 <211> 456
 <212> PRT
 <213> *Ostreococcus lucimarinus*
 <400> 20

Met Cys Val Glu Thr Glu Asn Asn Asp Gly Ile Pro Thr Val Glu Ile
 1 5 10 15
 Ala Phe Asp Gly Glu Arg Glu Arg Ala Glu Ala Asn Val Lys Leu Ser
 20 25 30
 Ala Glu Lys Met Glu Pro Ala Ala Leu Ala Lys Thr Phe Ala Arg Arg
 35 40 45
 Tyr Val Val Ile Glu Gly Val Glu Tyr Asp Val Thr Asp Phe Lys His
 50 55 60
 Pro Gly Gly Thr Val Ile Phe Tyr Ala Leu Ser Asn Thr Gly Ala Asp
 65 70 75 80
 Ala Thr Glu Ala Phe Lys Glu Phe His His Arg Ser Arg Lys Ala Arg
 85 90 95
 Lys Ala Leu Ala Ala Leu Pro Ser Arg Pro Ala Lys Thr Ala Lys Val
 100 105 110
 Asp Asp Ala Glu Met Leu Gln Asp Phe Ala Lys Trp Arg Lys Glu Leu
 115 120 125
 Glu Arg Asp Gly Phe Phe Lys Pro Ser Pro Ala His Val Ala Tyr Arg
 130 135 140
 Phe Ala Glu Leu Ala Ala Met Tyr Ala Leu Gly Thr Tyr Leu Met Tyr
 145 150 155 160
 Ala Arg Tyr Val Val Ser Ser Val Leu Val Tyr Ala Cys Phe Phe Gly
 165 170 175

101/143

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

<210> 21

<211> 894

<212> ADN

<213> Pyramimonas cordata

<400> 21

atggagttcg ctcagcctct tgtggctatg gcacaggagc agtatgccgc aattgacgcg 60

gtggtagccc ctgcaatttt ctcagctacc gacagcatcg gttgggggtct taagccatt 120

agcagcgcga caaaggatct tcctctcgtt gagagtccga cgccgctcat actgagcctg 180

ttggcctatt ttgcgatcgt cggctctggg ctggtgtacc gcaaagtatt ccctgcaca 240

102/143

gtaaaggggc aagacccctt cctgctgaag gcgctcatgc ttgcgacaaa cgtgttcctc 300
 attggcctca gtctatacat gtgcttgaag cttgtctacg aggcttacgt caacaagtac 360
 tccttctggg gaaacgccta caaccccgca cagaccgaga tggcgaaggt catctggatt 420
 ttctacgtct ccaagatcta tgagttcatg gacacgttca tcatgctctt gaagggcaac 480
 gtcaaccagg tctctttcct gcatgtgtac catcatggct ccatctctgg tatctggtgg 540
 atgatcacct acgctgcccc tggcgggtgac gcgtacttct cggcggcgct caactcgtgg 600
 gtgcacgtgt gcatgtacac gtactacttc atggcggcggt tgctgccccaa ggacgagaag 660
 accaagcgca agtacctctg gtggggccgc tacctgacct agatgcagat gttccagttc 720
 ttcatgaacc tgctccaggc ggtctacctc ctctactcct ctagccccta cccaagtctc 780
 atcgcccagc tgctggtggt gtacatgggt acgctgctga tgctcttcgg caacttctac 840
 tacatgaagc accacgcgag caagaagcag aagctggcca gcaagaagca gtag 894

<210> 22

<211> 870

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện Pyramimonas
 cordata 6-elongaza ở cây (được cắt ở đầu 3' và mã hóa
 elongaza chức năng) (phiên bản 1)

<400> 22

atggaattcg cccagcctct tgttgctatg gctcaagagc aatacgctgc tatcgatgct 60
 gttgttgctc ctgctatctt ctctgctact gattctatcg gatggggact taagcctatc 120
 tcttctgcta ctaaggactt gcctcttggt gagtctccta cacctctcat ctttctttg 180
 cttgcttact tcgctatcgt tggatctgga ctcgtttaca gaaaggtttt ccctagaacc 240
 gtgaagggac aagatccatt ccttttgaag gctcttatgc ttgctcacia cgtgttcctt 300
 atcggacttt ctctttacat gtgcctcaag cttgtgtacg aggcttacgt taacaagtac 360
 tctttctggg gaaacgctta caaccctgct caaactgaga tggctaaggt tatctggatc 420
 ttctacgtga gcaagatcta cgagttcatg gataccttca tcatgctcct caagggaaat 480
 gtttaaccagg ttagcttcct tcacgtttac catcacggat ctatctctgg aatctggtgg 540
 atgattactt acgctgctcc tgggtggtgat gcttacttct ctgctgctct taactcttgg 600
 gttcacgtgt gtatgtacac ctactatttt atggctgccc tgcttcctaa ggacgagaaa 660
 actaagagaa agtacctctg gtggggaaga taccttactc aaatgcagat gttccagttc 720
 ttcatgaacc ttctccaggc tgtttacctt ctctactcct catctcctta ccctaagttt 780
 atcgctcagc tcctcgtggt gtacatgggt actcttctca tgcttttcgg aaacttctac 840
 tacatgaagc accacgctag caagtatga 870

<210> 23

<211> 870

<212> ADN

<213> Trình tự nhân tạo

<220>

103/143

<223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện *Pyramimonas cordata* 6-elongaza ở cây (được cắt ở đầu 3' và mã hóa elongaza chức năng) (phiên bản 2)

<400> 23
 atggaattcg cccagcctct tggtgctatg gctcaagagc aatacgctgc tatcgatgct 60
 gttgttgctc ctgctatctt ctctgctact gattctatcg gatggggact taagcctatc 120
 tcttctgcta ctaaggactt gcctcttggt gagtctccta cacctctcat cctttctttg 180
 cttgcttact tcgctatcgt tggatctgga ctcgtttaca gaaagggttt ccctagaacc 240
 gtgaagggac aagatccatt ccttttgaag gctcttatgc ttgctcaciaa cgtgttcctt 300
 atcggacttt ctctttacat gtgcctcaag cttgtgtacg aggcttacgt taacaagtac 360
 tctttctggg gaaacgctta caaccctgct caaactgaga tggctaaggt tatctggatc 420
 ttctacgtga gcaagatcta cgagttcatg gataccttca tcatgctcct caagggaaat 480
 gttaaccagg ttagcttcct tcacgtttac catcacggat ctatctctgg aatctggtgg 540
 atgattactt acgctgctcc tgggtggtgat gcttacttct ctgctgctct taactcttgg 600
 gttcacgtgt gtatgtacac ctactatctt atggctgccg tgcttcctaa ggacgagaaa 660
 actaagagaa agtacctctg gtggggaaga taccttactc aaatgcagat gttccagttc 720
 ttcatgaacc ttctccaggc tgtttacctt ctctactctt catctcctta ccctaagttt 780
 atcgctcagc tcctcgtggt gtacatggtt actcttctca tgcttttcgg aaacttctac 840
 tacatgaagc accacgctag caagtgatga 870

<210> 24
 <211> 870
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện *Pyramimonas cordata* 6-elongaza ở cây (được cắt ở đầu 3' và mã hóa elongaza chức năng) (phiên bản 3)

<400> 24
 atggaatttg ctcaacctct cgttgctatg gctcaagagc agtacgctgc tatcgatgct 60
 gttgttgctc ctgctatctt ctctgctacc gactctattg gatggggact caagcctatc 120
 tcttctgcta ctaaggatct cctctcgtt gaatctccta cccctcttat cctttctctc 180
 ctcgcttact tcgctatcgt tggttctgga ctcgtttacc gtaaagtgtt ccctagaacc 240
 gttaagggac aggatccttt ccttctcaag gctcttatgc tcgctcaciaa cgttttcctt 300
 atcggactca gcctttacat gtgcctcaag ctcgtttacg aggcttacgt gaacaagtac 360
 tccttctggg gaaacgctta caaccctgct caaaccgaga tggctaaggt gatctggatc 420
 ttctacgtgt ccaagatcta cgagttcatg gacaccttca tcatgcttct caagggaaac 480
 gttaaccagg tttccttcct ccatgtttac caccacggat ctatctctgg aatctggtgg 540
 atgatcactt atgctgctcc aggtggagat gcttacttct ctgctgctct caactcttgg 600
 gttcatgtgt gcatgtacac ctactacttc atggctgctg ttcttcctaa ggacgaaaag 660

104/143

accaagagaa agtacctttg gtggggaaga taccttacct agatgcaa atgttccagttc 720
 ttcatgaacc ttctccaggc tgtttacctc ctctactctt cttctcctta ccctaagttc 780
 attgctcaac tcctcggtgt ttacatgggt accctcctca tgcttttcgg aaacttctac 840
 tacatgaagc accacgcttc taagtataa 870

<210> 25
 <211> 297
 <212> PRT
 <213> Pyramimonas cordata

<400> 25

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

105/143

Lys Gln Lys Leu Ala Ser Lys Lys Gln
 290 295

<210> 26
 <211> 288
 <212> PRT
 <213> Pyramimonas cordata

<400> 26

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

<210> 27
 <211> 1278
 <212> ADN
 <213> Pavlova salina

<400> 27
 atgccgccgc gcgatagcta ctcgtacgcc gccccgccgt cggcccagct gcacgaggtc 60
 gataccccgc aggagcatga taagaaggag ctcgtcatcg gtgaccgcgc gtacgacgtg 120
 accaactttg tgaagcgcca cccgggtggc aagatcatcg cataccaggt tggcacagat 180
 gcgacggacg cgtacaagca gttccatgtg cggctctgcca aggcggacaa gatgctcaag 240
 tcgctgcctt cgcgcccggg gcacaagggc tactcgcccc gccgcgctga cctcattgcc 300
 gacttccagg agttcaccaa gcagctggag gcggagggca tgtttgagcc gtcgctgccg 360
 cacgtggcat accgcctggc ggaggtgatc gcgatgcacg tggccggcgc cgcgctcatc 420
 tggcacgggt acaccttcgc gggcattgcc atgctcggcg ttgtgcaggg ccgctgcggc 480
 tggctcatgc acgagggcgg ccactactcg ctcacgggca acattgcttt tgaccgtgcc 540
 atccaagtcg cgtgctacgg ccttggtcgc ggcattgtcg gcgctgggtg gcgcaaccag 600
 cacaacaagc accacgcgac gccgcagaag ttgcagcacg acgtcgacct cgacaccctc 660
 ccgctcgtcg ccttcacga gcggatagcc gccaaaggta agagccccgc gatgaaggcg 720
 tggcttagta tgcaggcgaa gctcttcgcg ccagtgacca cgctgctggt cgcgctgggc 780
 tggcagctgt acctgcacc cgcgccatat ctgcgcacca agcactacga cgagctcgcg 840
 atgctcggca ttcgctacgg ccttgctcggc tacctcgcgg cgaactacgg cgcggggtag 900
 gtgctcgcgt gctacctgct gtacgtgcag ctcggcgcca tgtacatctt ctgcaacttt 960
 gccgtgtcgc acacacacct gccggttgct gagcctaacg agcacgcaac gtgggtggag 1020
 tacgccgcga accacacgac caactgctcg ccctcgtggt ggtgcgactg gtggatgtcg 1080
 tacctcaact accagatcga gcaccacctc taccctcca tgccgcagtt ccgccacccg 1140
 aagattgcgc cgcgggtgaa gcagctcttc gagaagcacg gcctgcacta cgacgtgcgt 1200
 ggctacttcg aggccatggc ggacacgttt gccaaccttg acaacgctcg gcacgcgccg 1260
 gagaagaaga tgcagtga 1278

<210> 28
 <211> 1281
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện Pavlova
 salina 5-desaturaza ở cây (phiên bản 1)

<400> 28
 atgcctccaa gggactctta ctcttatgct gctcctcctt ctgctcaact tcacgaagtt 60
 gatactcctc aagagcacga caagaaagag cttgttatcg gagatagggc ttacgatgtt 120
 accaacttcg ttaagagaca ccctggtgga aagatcattg cttaccaagt tggaactgat 180
 gctaccgatg cttacaagca gttccatgtt agatctgcta aggctgacaa gatgcttaag 240
 tctcttcctt ctcgtcctgt tcacaaggga tactctcaa gaagggtga tcttatcgct 300
 gatttccaag agttcaccaa gcaacttgag gctgaggga tgttcgagcc ttctcttcct 360
 catgttgctt acagacttgc tgaggttatc gctatgcatg ttgctggtgc tgctcttata 420

107/143

tgccatggat acactttcgc tggaatcgct atgcttggag ttgttcaggg aagatgtgga	480
tggttatgc atgaggggtg acattactct ctactggaa acattgcttt cgacagagct	540
atccaagttg cttgttacgg acttggatgt ggaatgtctg gtgcttggtg gcgtaaccag	600
cataacaagc accatgctac tcctcaaaag cttcagcacg atgttgatct tgataccctt	660
cctctcgttg ctttccatga gagaatcgct gctaaggtta agtctcctgc tatgaaggct	720
tggttttcta tgcaagctaa gcttttcgct cctgttacca ctcttcttgt tgctcttgga	780
tggcagcttt accttcatcc tagacacatg ctacaggacta agcactacga tgagcttgct	840
atgctcggaa tcagatacgg acttgttgga taccttgctg ctaactacgg tgctggatac	900
gttctcgctt gttaccttct ttacgttcag cttggagcta tgtacatctt ctgcaacttc	960
gctgtttctc atactcacct ccctgttggt gagcctaacg agcatgctac ttgggttgag	1020
tacgctgcta accacactac taactgttct ccatcttggt ggtgtgattg gtggatgtct	1080
taccttaact accagatcga gcaccacctt tacccttcta tgcctcaatt cagacaccct	1140
aagatcgctc ctagagttaa gcagcttttc gagaagcacg gacttcacta cgatgttaga	1200
ggatacttcg aggctatggc tgatactttc gctaaccttg ataacgttgc ccatgctcct	1260
gagaagaaaa tgcagtaatg a	1281

<210> 29

<211> 1281

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện Pavlova
salina 5-desaturaza ở cây (phiên bản 2)

<400> 29

atgcctccta gggactctta ctcttacgct gctcctcctt ctgctcaact tcacgaggtt	60
gacactcctc aagagcacga caagaaagag cttgttatcg gagatagggc ttacgatgtg	120
accaacttcg ttaagagaca ccctgggtgga aagatcattg cttaccaagt gggaactgat	180
gctaccgatg cttacaagca gttccatgtg agatctgcta aggctgacaa gatgctcaag	240
tctctccctt ctagacctgt tcacaaggga tactctccta gaagagctga tcttatcgct	300
gacttccaag agttcactaa gcaacttgag gctgagggaa tggtcgaacc ttctctcctt	360
catgttgctt accgtcttgc tgagggtatc gctatgcatg ttgctgggtg tgctcttatc	420
tggcacggat acactttcgc tggaatcgct atgcttggag ttgttcaggg aagatgcgga	480
tggttatgc atgaggggtg acactactct cttaccggaa acattgcttt cgatagggct	540
atccaagttg cttgttacgg acttggatgc ggaatgtctg gtgcttggtg gagaaaccag	600
cataacaagc accatgctac tcctcaaaag ctccagcacg atgttgatct tgataccctc	660
cctctcgttg ctttccatga gagaatcgct gctaaggtta agtctcctgc tatgaaggct	720
tggtcttcca tgcaagctaa actcttcgct cctgttacca ctcttcttgt tgctcttgga	780
tggcagcttt accttcccc tagacacatg ctccagaacta agcactacga cgagcttgct	840
atgcttggta tcagatacgg acttgtggga taccttgctg ctaactacgg tgctggatac	900

108/143

```

gttcttgctt gctaccttct ctacgttcag cttggagcta tgtacatctt ctgcaacttc      960
gctgtttctc acactcatct ccctgttggt gagcctaacg agcatgctac ttgggttgag      1020
tacgctgcta accacactac taactgctct ccatcttggt ggtgtgattg gtggatgagc      1080
tacctcaact accagatcga gcatcacctt tacccttcta tgcctcagtt caggcatcct      1140
aagatcgctc ctagagtga gcaactcttc gagaagcacg gacttcacta cgatgtgcgt      1200
ggatacttcg aggctatggc tgatactttc gctaacctcg ataacgttgc tcatgctcct      1260
gagaagaaaa tgcaatgatg a                                              1281

```

<210> 30

<211> 425

<212> PRT

<213> Pavlova salina

<400> 30

```

Met Pro Pro Arg Asp Ser Tyr Ser Tyr Ala Ala Pro Pro Ser Ala Gln
1          5          10          15
Leu His Glu Val Asp Thr Pro Gln Glu His Asp Lys Lys Glu Leu Val
          20          25          30
Ile Gly Asp Arg Ala Tyr Asp Val Thr Asn Phe Val Lys Arg His Pro
          35          40          45
Gly Gly Lys Ile Ile Ala Tyr Gln Val Gly Thr Asp Ala Thr Asp Ala
          50          55          60
Tyr Lys Gln Phe His Val Arg Ser Ala Lys Ala Asp Lys Met Leu Lys
65          70          75          80
Ser Leu Pro Ser Arg Pro Val His Lys Gly Tyr Ser Pro Arg Arg Ala
          85          90          95
Asp Leu Ile Ala Asp Phe Gln Glu Phe Thr Lys Gln Leu Glu Ala Glu
          100          105          110
Gly Met Phe Glu Pro Ser Leu Pro His Val Ala Tyr Arg Leu Ala Glu
          115          120          125
Val Ile Ala Met His Val Ala Gly Ala Ala Leu Ile Trp His Gly Tyr
          130          135          140
Thr Phe Ala Gly Ile Ala Met Leu Gly Val Val Gln Gly Arg Cys Gly
145          150          155          160
Trp Leu Met His Glu Gly Gly His Tyr Ser Leu Thr Gly Asn Ile Ala
          165          170          175
Phe Asp Arg Ala Ile Gln Val Ala Cys Tyr Gly Leu Gly Cys Gly Met
          180          185          190
Ser Gly Ala Trp Trp Arg Asn Gln His Asn Lys His His Ala Thr Pro
          195          200          205
Gln Lys Leu Gln His Asp Val Asp Leu Asp Thr Leu Pro Leu Val Ala
          210          215          220
Phe His Glu Arg Ile Ala Ala Lys Val Lys Ser Pro Ala Met Lys Ala
225          230          235          240
Trp Leu Ser Met Gln Ala Lys Leu Phe Ala Pro Val Thr Thr Leu Leu
          245          250          255

```

109/143

Val Ala Leu Gly Trp Gln Leu Tyr Leu His Pro Arg His Met Leu Arg
 260 265 270
 Thr Lys His Tyr Asp Glu Leu Ala Met Leu Gly Ile Arg Tyr Gly Leu
 275 280 285
 Val Gly Tyr Leu Ala Ala Asn Tyr Gly Ala Gly Tyr Val Leu Ala Cys
 290 295 300
 Tyr Leu Leu Tyr Val Gln Leu Gly Ala Met Tyr Ile Phe Cys Asn Phe
 305 310 315 320
 Ala Val Ser His Thr His Leu Pro Val Val Glu Pro Asn Glu His Ala
 325 330 335
 Thr Trp Val Glu Tyr Ala Ala Asn His Thr Thr Asn Cys Ser Pro Ser
 340 345 350
 Trp Trp Cys Asp Trp Trp Met Ser Tyr Leu Asn Tyr Gln Ile Glu His
 355 360 365
 His Leu Tyr Pro Ser Met Pro Gln Phe Arg His Pro Lys Ile Ala Pro
 370 375 380
 Arg Val Lys Gln Leu Phe Glu Lys His Gly Leu His Tyr Asp Val Arg
 385 390 395 400
 Gly Tyr Phe Glu Ala Met Ala Asp Thr Phe Ala Asn Leu Asp Asn Val
 405 410 415
 Ala His Ala Pro Glu Lys Lys Met Gln
 420 425

<210> 31
 <211> 1329
 <212> ADN
 <213> Pyramimonas cordata

<400> 31
 atgggaaagg gaggcaatgc tagcgctcct actgcgaaga aggaggtggt gatcgagggg 60
 aagttttacg atgtcaccga cttcaggcac cccggtgggt cgatcatcaa gtttctctcg 120
 ggttctggtg ctgacgccac cgcttcctac cgcgaggttcc acgttaggtc agcgaaggca 180
 gacaagttct tgaagacgct gccctcccgc gaagccactc cccaggagct gaagcaggcg 240
 gttgagttct ccaagctcaa cccgccctcc gcggagagtg cctctgctcc cctgaccgac 300
 cttgccaaagg tggaagcgct gaacaaggac ttcgaggctt tccgtgagca gctcattcag 360
 gagggtttct ttaagcccaa tatcccgcat gtggtcaagc gcatcacgga agtcgtggcg 420
 atgatggccg tagcctcctg gatgatggtg cagaccaacg ctcttggtgt gaccctcgga 480
 gttctgatcc gcggcattgc acagggccgg tgcggttggc ttatgcacga gggcggccac 540
 tatagtctta ctgggaagat ctccattgat aggcgtctgc aggagtcaat ttacggattc 600
 ggctgtggaa tgtccggcgc ctggtggcgc aaccagcaca acaagcacca cgcaacccca 660
 cagaagctgc agcatgacgt cgacctggag acccttcctc tgatggcttt caacaacgct 720
 gttaccgata gacgcaaggt gaagcctggt agtctccagg ctctgtggct caagtaccag 780
 gccttcctct tcttccccgt gacctccctt ctggtcggcc tcggttgac caccgtcctc 840
 caccaccaggc acagcttgcg caccaagcac tatttcgagc tgctctgcat ggctgctcgt 900

110/143

tacgcgagtt tcgctgctct tttcgctccc aagtacggac ttgcaggagc tgccgggctc 960
 tacctcgcca ccttcgctgt cgggtgcaac tatattttca tcaacttctc ggtctctcac 1020
 actcacctgc ccgtgagcgg tgcgagcgag tacctgcatt gggtcgtgta ttcggccatc 1080
 cacaccacta acatcaaatac cagcatgctg tgcgattggt ggatgtcatt cctcaacttc 1140
 cagatcgagc atcacctggt cccttcaatg cccagttcc gccacaagat tatctccccg 1200
 cgtgtaaagg ccttgtttga gaagcacggt cttgtgtatg atgtgcgccc ctattggggg 1260
 gccatggctg acaccttcaa gaacttgaat gacgttgga ctcacgcata tcaactcaag 1320
 gcgcactag 1329

<210> 32

<211> 442

<212> PRT

<213> Pyramimonas cordata

<400> 32

Met Gly Lys Gly Gly Asn Ala Ser Ala Pro Thr Ala Lys Lys Glu Val
 1 5 10 15
 Leu Ile Glu Gly Lys Phe Tyr Asp Val Thr Asp Phe Arg His Pro Gly
 20 25 30
 Gly Ser Ile Ile Lys Phe Leu Ser Gly Ser Gly Ala Asp Ala Thr Ala
 35 40 45
 Ser Tyr Arg Glu Phe His Val Arg Ser Ala Lys Ala Asp Lys Phe Leu
 50 55 60
 Lys Thr Leu Pro Ser Arg Glu Ala Thr Pro Gln Glu Leu Lys Gln Ala
 65 70 75 80
 Val Glu Phe Ser Lys Leu Asn Pro Pro Ser Ala Glu Ser Ala Ser Ala
 85 90 95
 Pro Leu Thr Asp Leu Ala Lys Val Glu Ala Leu Asn Lys Asp Phe Glu
 100 105 110
 Ala Phe Arg Glu Gln Leu Ile Gln Glu Gly Phe Phe Lys Pro Asn Ile
 115 120 125
 Pro His Val Val Lys Arg Ile Thr Glu Val Val Ala Met Met Ala Val
 130 135 140
 Ala Ser Trp Met Met Val Gln Thr Asn Ala Leu Val Val Thr Leu Gly
 145 150 155 160
 Val Leu Ile Arg Gly Ile Ala Gln Gly Arg Cys Gly Trp Leu Met His
 165 170 175
 Glu Gly Gly His Tyr Ser Leu Thr Gly Lys Ile Ser Ile Asp Arg Arg
 180 185 190
 Leu Gln Glu Ser Ile Tyr Gly Phe Gly Cys Gly Met Ser Gly Ala Trp
 195 200 205
 Trp Arg Asn Gln His Asn Lys His His Ala Thr Pro Gln Lys Leu Gln
 210 215 220
 His Asp Val Asp Leu Glu Thr Leu Pro Leu Met Ala Phe Asn Asn Ala
 225 230 235 240

111/143

Val Thr Asp Arg Arg₂₄₅ Lys Val Lys Pro Gly₂₅₀ Ser Leu Gln Ala Leu₂₅₅ Trp
 Leu Lys Tyr Gln₂₆₀ Ala Phe Leu Phe Phe₂₆₅ Pro Val Thr Ser Leu₂₇₀ Leu Val
 Gly Leu Gly₂₇₅ Trp Thr Thr Val Leu₂₈₀ His Pro Arg His₂₈₅ Ser Leu Arg Thr
 Lys His₂₉₀ Tyr Phe Glu Leu Leu₂₉₅ Cys Met Ala Ala Arg₃₀₀ Tyr Ala Ser Phe
 Ala₃₀₅ Ala Leu Phe Ala Pro₃₁₀ Lys Tyr Gly Leu Ala₃₁₅ Gly Ala Ala Gly Leu₃₂₀
 Tyr Leu Ala Thr Phe₃₂₅ Ala Val Gly Cys Asn₃₃₀ Tyr Ile Phe Ile Asn₃₃₅ Phe
 Ser Val Ser His₃₄₀ Thr His Leu Pro Val₃₄₅ Ser Gly Ala Ser Glu₃₅₀ Tyr Leu
 His Trp Val₃₅₅ Val Tyr Ser Ala Ile₃₆₀ His Thr Thr Asn Ile₃₆₅ Lys Ser Ser
 Met Leu₃₇₀ Cys Asp Trp Trp Met₃₇₅ Ser Phe Leu Asn Phe₃₈₀ Gln Ile Glu His
 His₃₈₅ Leu Phe Pro Ser Met₃₉₀ Pro Gln Phe Arg His₃₉₅ Lys Ile Ile Ser Pro₄₀₀
 Arg Val Lys Ala₄₀₅ Leu Phe Glu Lys His Gly₄₁₀ Leu Val Tyr Asp Val₄₁₅ Arg
 Pro Tyr Trp Gly₄₂₀ Ala Met Ala Asp Thr₄₂₅ Phe Lys Asn Leu Asn₄₃₀ Asp Val
 Gly Thr His₄₃₅ Ala Ser His Ser Lys₄₄₀ Ala His

<210> 33

<211> 804

<212> ADN

<213> Pyramimonas cordata

<400> 33

atggcgtcta ttgcgattcc ggctgcgctg gcagggactc ttggttatgt gacgtacaat	60
gtcgcaaacc cagatattcc tgcattccgag aagggtgcctg cttactttat gcaggtcgag	120
tattgggggc caacgattgg gaccatcggg tatcttctgt tcatctactt tggtaaacgg	180
attatgcaaa acaggagcca gccgtttggc ctgaagaacg ctatgctggt gtacaacttc	240
tatcagactt tcttcaactc gtactgcata tacctttttg tcacgtcgca ccgcgctcag	300
gggctgaaag tttggggaaa catccccgat atgactgcca acagctgggg gatctcacag	360
gtgatctggc tgcactacaa caacaagtac gttgagctgc tggacacgtt cttcatggtc	420
atgcgcaaga agtttgacca gctttcggtc ctgcacattt accatcatac cctgttgatc	480
tggctcttggg tcgtgggtgat gaaattggag cccgttgggg actgctactt tggctctagc	540
gtcaacacgt ttgtgcacgt cattatgtac tcgtactatg gccttgccgc gctcgggggtg	600
aattgcttct ggaagaagta cattacgcag attcagatgc tgcagttctg tatctgcgct	660
tcgcactcga ttataaccgc ctatgtgcag aacaccgcgt tctggttgcc ttacttgacg	720

112/143

ctgtgggtga tgggtgaacat gttcgtgttg ttcgccaact tctatcgcaa gcgctacaag 780
agcaaggggtg ccaagaagca gtaa 804

<210> 34
<211> 807
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện Pyramimonas
cordata 5-elongaza ở cây (phiên bản 1)

<400> 34
atggcctcta tcgctatccc tgctgctctt gctggaactc ttggatacgt tacctacaat 60
gtggctaacc ctgatatccc agcttctgag aaagttcctg cttacttcat gcaggttgag 120
tactgggggac ctactatcgg aactattgga tacctcctct tcatctactt cggaaagcgt 180
atcatgcaga acagatctca acctttcgga ctcaagaacg ctatgctcgt ttacaacttc 240
taccagacct tcttcaacag ctactgcacg taccttttcg ttactttctca tagggctcag 300
ggacttaagg tttggggaaa catccctgat atgactgcta actcttgggg aatctctcag 360
gttatctggc ttactataca caacaagtac gttgagcttc tcgacacctt cttcatgggtg 420
atgaggaaga agttcgacca gctttctttc cttcacatct accaccacac tcttctcatc 480
tggtcatggg tcgttggtat gaagcttgag cctggttgag attgctactt cggatcttct 540
gttaacacct tcgtgcacgt gatcatgtac tcttactacg gacttgctgc tcttgaggtt 600
aactgtttct ggaagaagta catcaccag atccagatgc ttcagttctg tatctgtgct 660
tctcactcta tctacaccgc ttacgttcag aataccgctt tctggcttcc ttacattcaa 720
ctctgggtta tgggtgaacat gttcgttctc ttcgccaact tctaccgtaa gaggtacaag 780
tctaaggggtg ctaagaagca gtgataa 807

<210> 35
<211> 867
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện Pyramimonas
cordata 5-elongaza ở cây (phiên bản 2)

<400> 35
atggaatttg ctcaacctct cggtgctatg gctcaagagc agtacgctgc tatcgatgct 60
gttggtgctc ctgctatctt ctctgctacc gactctattg gatggggact caagcctatc 120
tcttctgcta ctaaggatct ccctctcgtt gaatctccta cccctcttat cctttctctc 180
ctcgcttact tcgctatcgt tggttctgga ctcgtttacc gtaaagtgtt ccctagaacc 240
gttaagggac aggatccttt ccttctcaag gctcttatgc tcgctcaca cgttttctt 300
atcggactca gcctttacat gtgcctcaag ctcgtttacg aggcttacgt gaacaagtac 360
tccttctggg gaaacgctta caaccctgct caaacgaga tggctaaggt gatctggatc 420
ttctacgtgt ccaagatcta cgagttcatg gacaccttca tcatgcttct caagggaaac 480

113/143

gttaaccagg tttccttcct ccatgtttac caccacggat ctatctctgg aatctggtgg 540
 atgatacatt atgctgctcc aggtggagat gcttacttct ctgctgctct caactcttgg 600
 gttcatgtgt gcatgtacac ctactacttc atggctgctg ttcttcctaa ggacgaaaag 660
 accaagagaa agtacctttg gtggggaaga taccttacc agatgcaa atgttccagttc 720
 ttcatgaacc ttctccaggc tgtttacctc ctctactctt cttctcctta ccctaagttc 780
 attgctcaac tcctcgttgt ttacatgggt accctcctca tgcttttcgg aaacttctac 840
 tacatgaagc accacgcttc taagtga 867

<210> 36
 <211> 807
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện *Pyramimonas cordata* 5-elongaza ở cây (phiên bản 3)

<400> 36
 atggcttcta tcgctatccc tgctgctctt gctggaactc ttggatacgt gacctacaac 60
 gtggctaacc ctgatattcc tgcttctgag aaggttccag cttacttcat gcaagtggag 120
 tactggggac ctactatcgg aactatcggg tacctcctct tcatctactt cggaaagcgt 180
 atcatgcaaa acagaagcca gcctttcggg ctttaagaacg ctatgctcgt gtacaacttc 240
 taccagacct tcttcaacag ctactgcac tacctcttcg ttacctctca tagggctcag 300
 ggacttaaaag tttggggaaa catccctgat atgaccgcta actcttgggg aatctctcag 360
 gttatctggc tccactacaa caacaagtac gtggagcttc tcgatacctt cttcatgggtg 420
 atgaggaaga agttcgacca gctttctttc cttcacatct accaccacac tcttctcatc 480
 tggatcatggg tcgtgggttat gaagctcgag cctggttgag attgctactt cggatctagc 540
 gttaacacct tcgtgcacgt gatcatgtac tcttactacg gacttgctgc tcttgagtt 600
 aactgcttct ggaagaagta catcaccag atccagatgc ttcagttctg tatctgcgct 660
 tctcactcta tctacaccgc ttacgttcag aacactgctt tctggcttcc ttaccttcag 720
 ctctgggtga tggttaacat gttcgtgctc ttcgctaact tctaccgtaa aaggtacaag 780
 agcaagggtg ctaagaagca gtgataa 807

<210> 37
 <211> 267
 <212> PRT
 <213> *Pyramimonas cordata*

<400> 37

Met Ala Ser Ile Ala Ile Pro Ala Ala Leu Ala Gly Thr Leu Gly Tyr
 1 5 10 15
 Val Thr Tyr Asn Val Ala Asn Pro Asp Ile Pro Ala Ser Glu Lys Val
 20 25 30
 Pro Ala Tyr Phe Met Gln Val Glu Tyr Trp Gly Pro Thr Ile Gly Thr
 35 40 45
 Ile Gly Tyr Leu Leu Phe Ile Tyr Phe Gly Lys Arg Ile Met Gln Asn

114/143

50 55 60
 Arg Ser Gln Pro Phe Gly Leu Lys Asn Ala Met Leu Val Tyr Asn Phe
 65 70 75 80
 Tyr Gln Thr Phe Phe Asn Ser Tyr Cys Ile Tyr Leu Phe Val Thr Ser
 85 90 95
 His Arg Ala Gln Gly Leu Lys Val Trp Gly Asn Ile Pro Asp Met Thr
 100 105 110
 Ala Asn Ser Trp Gly Ile Ser Gln Val Ile Trp Leu His Tyr Asn Asn
 115 120 125
 Lys Tyr Val Glu Leu Leu Asp Thr Phe Phe Met Val Met Arg Lys Lys
 130 135 140
 Phe Asp Gln Leu Ser Phe Leu His Ile Tyr His His Thr Leu Leu Ile
 145 150 155 160
 Trp Ser Trp Phe Val Val Met Lys Leu Glu Pro Val Gly Asp Cys Tyr
 165 170 175
 Phe Gly Ser Ser Val Asn Thr Phe Val His Val Ile Met Tyr Ser Tyr
 180 185 190
 Tyr Gly Leu Ala Ala Leu Gly Val Asn Cys Phe Trp Lys Lys Tyr Ile
 195 200 205
 Thr Gln Ile Gln Met Leu Gln Phe Cys Ile Cys Ala Ser His Ser Ile
 210 215 220
 Tyr Thr Ala Tyr Val Gln Asn Thr Ala Phe Trp Leu Pro Tyr Leu Gln
 225 230 235 240
 Leu Trp Val Met Val Asn Met Phe Val Leu Phe Ala Asn Phe Tyr Arg
 245 250 255
 Lys Arg Tyr Lys Ser Lys Gly Ala Lys Lys Gln
 260 265

<210> 38
 <211> 1344
 <212> ADN
 <213> Pavlova salina

<400> 38
 atgcctccga ggcgggcgaa gcagatgggc gcgagcacgg gcgtgcatgc gggcgtcaca 60
 gattcgtcgg ccttcacgcg caaggatgtc gccgacaggc cggacctcac gatcgtgggt 120
 gacagcgtgt acgatgcgaa ggcgttccgc tccgagcatc cgggtggcgc gcactttgtg 180
 tcgctgttcg gcgggcgaga tgccacggag gcgttcatgg agtaccaccg gcgcgcctgg 240
 cccaagtcgc gcatgtcgcg cttccacgtc ggctctcttg catcgaccga ggagcccgtc 300
 gccgccgatg agggctacct ccagctgtgc gctcgcacgc ccaagatggt gccgtcggtc 360
 agcagcgggt tcgcgccggc gtcgtactgg gtgaaggccg ggctgacccg cggctccgcg 420
 atcgcgctcg aggcgtacat gctgtacgcg ggcaagcgcc tgctcccgtc gatcgtgctc 480
 ggggtggctgt ttgcgctgat tggcctgaac atccagcacg atgccaacca cggcgcgctc 540
 tccaagtcgg cctcgggtcaa cctggcgctc gggttgtgcc aggactggat cggcggggagc 600
 atgatcctct ggctgcagga gcacgttgtc atgcaccact tgcacaccaa cgacgttgac 660

115/143

aaggacccgg accagaaggc gcacggcgcc ctgcggctca agccgaccga cgcgtggagc 720
ccgatgcact ggctgcagca cctctacctg ctgcctgggg agacgatgta cgccttcaag 780
ctgctgtttc tcgacatcag cgagctgggt atgtggcggt gggagggcga gcccatcagc 840
aagctggccg ggtacctctt catgccctcg ctgctcctca agctcacctt ctgggcgcgc 900
tttgtcgcgc tgccgctgta cctcgcgcc agcgtgcaca cggcgggtgtg catcgcggcg 960
acggtaatga cggggagctt ctacctcgcc ttcttcttct tcatctcgca caacttcgag 1020
ggcgtggcga gcgtcggacc ggacggcagc atcaccagca tgacgcgcgg cgcattcttc 1080
ctcaagcggc aggccgagac ctcgtccaac gtgggcggcc cgctgctcgc cacgctcaac 1140
ggcggcctca actacaaat cgagcaccac ctcttcccca gggtgcacca cggcttctac 1200
cctcgcctcg cgccgttgggt caaggcggag ctcgaggcgc gcggcattga gtacaagcac 1260
taccacacca tatggagcaa cctggcatcc acgctgaggc acatgtacgc gctcggccgc 1320
aggccgcga gcaaggcgga gtga 1344

<210> 39

<211> 1347

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện Pavlova
salina 4-desaturaza ở cây (phiên bản 1)

<400> 39

atgccaccta gcgctgctaa gcaaattgga gcttctactg gtgttcatgc tgggtgttact 60
gactcttctg ctttcaccag aaaggatggt gctgatagac ctgatctcac catcgttgga 120
gattctgttt acgatgctaa ggctttcaga tctgagcatc ctgggtgggtgc tcatttcggt 180
tctttgttcg gaggaagaga tgctactgag gctttcatgg aataccatag aagggttgg 240
cctaagtcta gaatgtctag attccacggt ggatctcttg cttctactga ggaacctgtt 300
gctgctgatg agggatacct tcaactttgt gctaggatcg ctaagatggt gccttctgtt 360
tcttctggat tcgctcctgc ttcttactgg gttaaggctg gacttatcct tggatctgct 420
atcgctcttg aggcttacat gctttacgct ggaaagagac ttctcccttc tatcgttctt 480
ggatggcttt tcgctcttat cggctttaac atccagcatg atgctaacca tgggtgctttg 540
tctaagtctg cttctgttaa ccttgctctt ggactttgtc aggattggat cggaggatct 600
atgatccttt ggcttcaaga gcatgttggt atgcaccacc tccacactaa cgatgttgat 660
aaggatcctg atcaaaaggc tcacggtgct ctagactca agcctactga tgcttgggtca 720
cctatgcatt ggcttcagca tctttacctt ttgcctgggtg agactatgta cgctttcaag 780
cttttgttcc tcgacatctc tgagcttggt atgtggcggt gggaggggtga gcctatctct 840
aagcttgctg gatacctctt tatgccttct ttgcttctca agcttacctt ctgggctaga 900
ttcgttgctt tgcctcttta ccttgctcct tctgttcata ctgctgtgtg tatcgtgct 960
actgttatga ctggatcttt ctacctgct ttcttcttct tcatctcca caacttcgag 1020
gggtgttgct ctgttggtgacc tgatggatct atcatttcta tgactagagg tgctagcttc 1080

116/143

cttaagagac aagctgagac ttcttctaac gttggaggac ctcttcttgc tactcttaac 1140
 ggtggactca actaccaaact tgagcatcac ttgttcccta gagttcacca tggattctac 1200
 cctagacttg ctcttcttgt taaggctgag cttgaggcta gaggaatcga gtacaagcac 1260
 taccctacta tctggtctaa ccttgcttct accctcagac atatgtacgc tcttggaaga 1320
 aggcctagat ctaaggctga gtaatga 1347

<210> 40
 <211> 1347
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện Pavlova
 salina 4-desaturaza ở cây (phiên bản 2)

<400> 40
 atgcctccat ctgctgctaa acagatggga gcttctactg gtgttcacgc tgggtgttacc 60
 gattcttctg ctttcaccag aaaggatgtg gctgatagac ctgatcttac catcgttggt 120
 gactctgtgt acgatgctaa ggctttcaga tctgagcatc ctggtggtgc tcatttcggt 180
 tcactcttcg gaggaagaga tgctactgag gctttcatgg aataccacag aagagcttgg 240
 cctaagtcta ggatgtctag gttccatggt ggatctcttg cttctaccga ggaacctgtt 300
 gctgctgatg aggataacct tcagctttgt gctaggatcg ctaagatggt gccttctgtg 360
 tcactctggat tcgctccagc ttcttactgg gtttaaggctg gacttatcct cggatctgct 420
 atcgctcttg aggcttacat gctctacgct ggaaagagac ttctcccttc tatcgttctt 480
 ggatggctct tcgctcttat cggacttaac atccagcatg acgctaacca tgggtgctttg 540
 tctaagtctg ctagcggtta ccttgctctt ggactttgtc aggattggat cggaggatct 600
 atgatccttt ggctccaaga gcatgttggt atgcaccacc tccacaccaa cgatgttgat 660
 aaggaccctg atcaaaaggc tcatggtgct cttagactca agcctaccga tgcttggtca 720
 cctatgcatt ggcttcagca cctttacctt ctccctggtg aaactatgta cgctttcaag 780
 ctcttcttcc tcgatatctc tgagcttgtg atgtggagat gggaggggtga acctatctct 840
 aagctcgctg gatacctctt catgccttct cttctcctca agcttacctt ctgggctaga 900
 ttcgttgctc ttctctttta cctcgtcctt tctgttcata ctgctgtgtg tatcgctgct 960
 actgttatga ccggaagctt ctaccttgct ttcttcttct tcatcagcca caacttcgag 1020
 ggtgttgctt ctgttgacc tgatggatct atcacctcta tgaccagggg agcttctttc 1080
 cttaagaggc aggctgagac ttcttctaact gtgggaggac ctcttcttgc tactcttaac 1140
 ggtggactca actaccaaact cgagcaccac cttttcccta gagttcacca cggattctac 1200
 cctagacttg ctcttcttgt gaaggctgaa cttgaggcta gaggaatcga gtacaagcac 1260
 taccctacca tctggtctaa cctcgttctt accctcagac atatgtacgc tcttggaaga 1320
 aggcctagat ctaaggctga gtgatga 1347

<210> 41
 <211> 447

117/143

<212> PRT
 <213> Pavlova salina

<400> 41

```

Met Pro Pro Ser Ala Ala Lys Gln Met Gly Ala Ser Thr Gly Val His
1      5      10      15
Ala Gly Val Thr Asp Ser Ser Ala Phe Thr Arg Lys Asp Val Ala Asp
20      25      30
Arg Pro Asp Leu Thr Ile Val Gly Asp Ser Val Tyr Asp Ala Lys Ala
35      40      45
Phe Arg Ser Glu His Pro Gly Gly Ala His Phe Val Ser Leu Phe Gly
50      55      60
Gly Arg Asp Ala Thr Glu Ala Phe Met Glu Tyr His Arg Arg Ala Trp
65      70      75      80
Pro Lys Ser Arg Met Ser Arg Phe His Val Gly Ser Leu Ala Ser Thr
85      90      95
Glu Glu Pro Val Ala Ala Asp Glu Gly Tyr Leu Gln Leu Cys Ala Arg
100     105     110
Ile Ala Lys Met Val Pro Ser Val Ser Ser Gly Phe Ala Pro Ala Ser
115     120     125
Tyr Trp Val Lys Ala Gly Leu Ile Leu Gly Ser Ala Ile Ala Leu Glu
130     135     140
Ala Tyr Met Leu Tyr Ala Gly Lys Arg Leu Leu Pro Ser Ile Val Leu
145     150     155     160
Gly Trp Leu Phe Ala Leu Ile Gly Leu Asn Ile Gln His Asp Ala Asn
165     170     175
His Gly Ala Leu Ser Lys Ser Ala Ser Val Asn Leu Ala Leu Gly Leu
180     185     190
Cys Gln Asp Trp Ile Gly Gly Ser Met Ile Leu Trp Leu Gln Glu His
195     200     205
Val Val Met His His Leu His Thr Asn Asp Val Asp Lys Asp Pro Asp
210     215     220
Gln Lys Ala His Gly Ala Leu Arg Leu Lys Pro Thr Asp Ala Trp Ser
225     230     235     240
Pro Met His Trp Leu Gln His Leu Tyr Leu Leu Pro Gly Glu Thr Met
245     250     255
Tyr Ala Phe Lys Leu Leu Phe Leu Asp Ile Ser Glu Leu Val Met Trp
260     265     270
Arg Trp Glu Gly Glu Pro Ile Ser Lys Leu Ala Gly Tyr Leu Phe Met
275     280     285
Pro Ser Leu Leu Leu Lys Leu Thr Phe Trp Ala Arg Phe Val Ala Leu
290     295     300
Pro Leu Tyr Leu Ala Pro Ser Val His Thr Ala Val Cys Ile Ala Ala
305     310     315     320
Thr Val Met Thr Gly Ser Phe Tyr Leu Ala Phe Phe Phe Phe Ile Ser
325     330     335
His Asn Phe Glu Gly Val Ala Ser Val Gly Pro Asp Gly Ser Ile Thr

```

118/143

340	345	350
Ser Met Thr Arg Gly Ala Ser Phe Leu Lys Arg Gln Ala Glu Thr Ser		
355	360	365
Ser Asn Val Gly Gly Pro Leu Leu Ala Thr Leu Asn Gly Gly Leu Asn		
370	375	380
Tyr Gln Ile Glu His His Leu Phe Pro Arg Val His His Gly Phe Tyr		
385	390	395
Pro Arg Leu Ala Pro Leu Val Lys Ala Glu Leu Glu Ala Arg Gly Ile		
405	410	415
Glu Tyr Lys His Tyr Pro Thr Ile Trp Ser Asn Leu Ala Ser Thr Leu		
420	425	430
Arg His Met Tyr Ala Leu Gly Arg Arg Pro Arg Ser Lys Ala Glu		
435	440	445

<210> 42
 <211> 792
 <212> ADN
 <213> Isochrysis galbana

<400> 42
 atggccctcg caaacgacgc gggagagcgc atctgggcgg ctgtgaccga cccggaaatc 60
 ctcatgggca ctttctcgta cttgctactc aaaccgctgc tccgcaattc cgggctgggtg 120
 gatgagaaga agggcgcata caggacgtcc atgatctggg acaacgttct gctggcgctc 180
 ttctctgctc tgagcttcta cgtgacggcg accgccctcg gctgggacta tggtaggggc 240
 gcgtgggtgc gcaggcaaac cggcgacaca ccgcagccgc tcttccagtg cccgtccccg 300
 gtttgggact cgaagctctt cacatggacc gccaaaggcat tctattactc caagtacgtg 360
 gagtacctcg acacggcctg gctgggtgctc aagggaaga ggggtctcctt tctccaggcc 420
 ttccaccact ttggcgcgcc gtgggatgtg tacctcggca ttcggctgca caacgagggc 480
 gtatggatct tcatgttttt caactcggtc attcacacca tcatgtacac ctactacggc 540
 ctcaccgccg ccgggtataa gttcaaggcc aagccgctca tcaccgcgat gcagatctgc 600
 cagttcgtgg gcggcttcct gttgggtctg gactacatca acgtcccctg cttcaactcg 660
 gacaaagggg agttgttcag ctgggctttc aactatgcat acgtcggctc ggtcttcttg 720
 ctcttctgcc acttttttcta ccaggacaac ttggcaacga agaaatcggc caaggcgggc 780
 aagcagctct ag 792

<210> 43
 <211> 263
 <212> PRT
 <213> Isochrysis galbana

<400> 43
 Met Ala Leu Ala Asn Asp Ala Gly Glu Arg Ile Trp Ala Ala Val Thr
 1 5 10 15
 Asp Pro Glu Ile Leu Ile Gly Thr Phe Ser Tyr Leu Leu Leu Lys Pro
 20 25 30
 Leu Leu Arg Asn Ser Gly Leu Val Asp Glu Lys Lys Gly Ala Tyr Arg
 35 40 45

119/143

Thr Ser Met Ile Trp Tyr Asn Val Leu Leu Ala Leu Phe Ser Ala Leu
 50 55 60
 Ser Phe Tyr Val Thr Ala Thr Ala Leu Gly Trp Asp Tyr Gly Thr Gly
 65 70 75 80
 Ala Trp Leu Arg Arg Gln Thr Gly Asp Thr Pro Gln Pro Leu Phe Gln
 85 90 95
 Cys Pro Ser Pro Val Trp Asp Ser Lys Leu Phe Thr Trp Thr Ala Lys
 100 105 110
 Ala Phe Tyr Tyr Ser Lys Tyr Val Glu Tyr Leu Asp Thr Ala Trp Leu
 115 120 125
 Val Leu Lys Gly Lys Arg Val Ser Phe Leu Gln Ala Phe His His Phe
 130 135 140
 Gly Ala Pro Trp Asp Val Tyr Leu Gly Ile Arg Leu His Asn Glu Gly
 145 150 155 160
 Val Trp Ile Phe Met Phe Phe Asn Ser Phe Ile His Thr Ile Met Tyr
 165 170 175
 Thr Tyr Tyr Gly Leu Thr Ala Ala Gly Tyr Lys Phe Lys Ala Lys Pro
 180 185 190
 Leu Ile Thr Ala Met Gln Ile Cys Gln Phe Val Gly Gly Phe Leu Leu
 195 200 205
 Val Trp Asp Tyr Ile Asn Val Pro Cys Phe Asn Ser Asp Lys Gly Lys
 210 215 220
 Leu Phe Ser Trp Ala Phe Asn Tyr Ala Tyr Val Gly Ser Val Phe Leu
 225 230 235 240
 Leu Phe Cys His Phe Phe Tyr Gln Asp Asn Leu Ala Thr Lys Lys Ser
 245 250 255
 Ala Lys Ala Gly Lys Gln Leu
 260

<210> 44
 <211> 801
 <212> ADN
 <213> Emiliania huxleyi

<400> 44
 atgctcgatc ggcctcgtc cgacgcggcc atctggctctg cgggtgtccga tccggaaatc 60
 ctgatcggca ctttctccta cctgctgctc aagccgctgc tacgcaactc agggctcgtg 120
 gacgagcgga aaggcgccta ccggacctcg atgatctggt acaacgtggt gctcgcgctc 180
 ttctccgcga cgagcttcta cgtgactgcg accgcgctcg ggtgggacaa gggcaccggc 240
 gagtggctcc gcagtctcac gggcgacagc ccgcagcagc tgtggcaatg cccgtcgagg 300
 gtatgggact ccaagctggt cctgtggacg gccaaaggcct tctactactc aaagtacgtg 360
 gagtacctcg acacggcgtg gctcgtcctc aaggggaaga aggtctcctt cctgcagggc 420
 ttccaccact ttggcgcgcc gtgggacgtg tacctgggca ttcggctgaa gaacgagggc 480
 gtgtggatct tcatgttctt caactcggtc atccacacgg tcatgtacac gtactacggc 540
 ctcaccgccg cgggctacaa gatccgcggc aagccgatca tcaccgcgat gcaaataagc 600

120/143

cagttcgtcg gcggctttgt cctagtgtgg gactacatca acgtgccgtg cttccacgcc 660
 gacgccgggc aggtcttcag ctgggtcttt aactatgctt acgtcggctc cgtctttctg 720
 ctgttctgcc acttcttcta catggacaac atcgcggaagg ccaaggccaa gaaggccgtc 780
 gctacccgca aggcgctgtg a 801

<210> 45

<211> 801

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Khung đọc mở được tối ưu hóa theo codon để biểu hiện Emiliana
 huxleyi 9-elongaza ở cây

<400> 45

atgcttgata gagcttcac tcatgtgtgct atttggagcg ctgtttctga tcctgagatc 60
 cttatcggaa ctttctctta ctttttgctt aagcctctcc tcagaaactc tggacttggt 120
 gatgagagaa agggagctta ccgtacttct atgatctggt acaacgttgt tcttgctctt 180
 ttctctgcta cctctttcta cgttactgct actgctcttg gatgggataa gggaactggt 240
 gagtggctta gatctcttac tgggtgattct cctcaacaac tttggcagtg cccttctaga 300
 gtttgggaca gcaaactctt cttgtggact gctaaagcct tctactactc caagtacgtt 360
 gagtaccttg atactgcttg gcttgttctc aagggaaaga aggtttcatt cctccaggga 420
 ttccatcatt tcggtgctcc atgggatggt taccttgga tcaggcttaa gaacgaggga 480
 gtttggatct tcatgttctt caacagcttc atccacactg ttatgtacac ttactacgga 540
 cttactgctg ctggatacaa gatcagagga aagcctatca tcaccgctat gcaaactctt 600
 caattcgttg gtggattcgt tcttggtgtg gactacatca acgttccttg tttccatgct 660
 gatgctggac aagttttctc ttgggtgttc aactacgctt atgtgggac tgttttcctt 720
 cttttctgcc acttcttcta catggacaac attgctaagg ctaaggctaa aaaggctgtt 780
 gctaccagaa aggcctcttg a 801

<210> 46

<211> 266

<212> PRT

<213> Emiliana huxleyi

<400> 46

Met Leu Asp Arg Ala Ser Ser Asp Ala Ala Ile Trp Ser Ala Val Ser
 1 5 10 15
 Asp Pro Glu Ile Leu Ile Gly Thr Phe Ser Tyr Leu Leu Leu Lys Pro
 20 25 30
 Leu Leu Arg Asn Ser Gly Leu Val Asp Glu Arg Lys Gly Ala Tyr Arg
 35 40 45
 Thr Ser Met Ile Trp Tyr Asn Val Val Leu Ala Leu Phe Ser Ala Thr
 50 55 60
 Ser Phe Tyr Val Thr Ala Thr Ala Leu Gly Trp Asp Lys Gly Thr Gly
 65 70 75 80
 Glu Trp Leu Arg Ser Leu Thr Gly Asp Ser Pro Gln Gln Leu Trp Gln

121/143

	85		90		95										
Cys	Pro	Ser	Arg	Val	Trp	Asp	Ser	Lys	Leu	Phe	Leu	Trp	Thr	Ala	Lys
			100					105					110		
Ala	Phe	Tyr	Tyr	Ser	Lys	Tyr	Val	Glu	Tyr	Leu	Asp	Thr	Ala	Trp	Leu
		115					120					125			
Val	Leu	Lys	Gly	Lys	Lys	Val	Ser	Phe	Leu	Gln	Gly	Phe	His	His	Phe
	130					135					140				
Gly	Ala	Pro	Trp	Asp	Val	Tyr	Leu	Gly	Ile	Arg	Leu	Lys	Asn	Glu	Gly
145					150					155					160
Val	Trp	Ile	Phe	Met	Phe	Phe	Asn	Ser	Phe	Ile	His	Thr	Val	Met	Tyr
				165					170					175	
Thr	Tyr	Tyr	Gly	Leu	Thr	Ala	Ala	Gly	Tyr	Lys	Ile	Arg	Gly	Lys	Pro
			180					185					190		
Ile	Ile	Thr	Ala	Met	Gln	Ile	Ser	Gln	Phe	Val	Gly	Gly	Phe	Val	Leu
		195					200					205			
Val	Trp	Asp	Tyr	Ile	Asn	Val	Pro	Cys	Phe	His	Ala	Asp	Ala	Gly	Gln
	210					215					220				
Val	Phe	Ser	Trp	Val	Phe	Asn	Tyr	Ala	Tyr	Val	Gly	Ser	Val	Phe	Leu
225					230					235					240
Leu	Phe	Cys	His	Phe	Phe	Tyr	Met	Asp	Asn	Ile	Ala	Lys	Ala	Lys	Ala
				245					250					255	
Lys	Lys	Ala	Val	Ala	Thr	Arg	Lys	Ala	Leu						
			260					265							

<210> 47

<211> 819

<212> ADN

<213> Pavlova pinguis

<400> 47

atggttgcgc caccatcac gctcgagtgg ctgctttcgc cgaagctcaa ggatgcagtg	60
ttcgggtgggg aggtgctcta cttctccatt gcctacctgt ttcttgccgc cattttgaag	120
cgcaccccggt tgggtggacac gcggaagggc gcgtataaga gtggtatgat cgcgtacaac	180
gtgatcatgt gcgtgttctc gctgggtgtgc ttcatctgcc agctcgcagc cctgggctat	240
gacatgggct acttgcagtg ggtgcgtgac ctacagggg acgagattgt cccctctac	300
caggacgtgt ccccgcccc cgcttctcc aacaagctct tcaagtattc gtctattgcc	360
ttccactact ccaagtatgt tgagtacatg gacaccgat ggctgggtgat gaagggcaag	420
cccggtgcct tgctccaggc cttccaccac tttggcgccg cctgggacac ctactttggc	480
atcaccttcc agaacgaggg catctacgtg ttctgtgtgc tcaacgcctt catccacacg	540
atcatgtacg catactacgc ggccactgcg gcgggtctca agttctcact gaagttcgtc	600
atcacgctca tgcagatcac ccaattcaac gtgggcttcg taatgggtgta tcactacatc	660
accctggagt acttccgcaa ctaccggag ctctcttctt cctacctttt caactatgcg	720
tacgtctgca cggttctcct cctcttcatg cagttcttct acatggacaa ctttggcaag	780
aagaaggccg ctgccgccgc gggcaagaag aagaagtag	819

122/143

<210> 48
 <211> 272
 <212> PRT
 <213> Pavlova pinguis

<400> 48

Met Val Ala Pro Pro Ile Thr Leu Glu Trp Leu Leu Ser Pro Lys Leu
 1 5 10 15
 Lys Asp Ala Val Phe Gly Gly Glu Val Leu Tyr Phe Ser Ile Ala Tyr
 20 25 30
 Leu Phe Leu Ala Pro Ile Leu Lys Arg Thr Pro Leu Val Asp Thr Arg
 35 40 45
 Lys Gly Ala Tyr Lys Ser Gly Met Ile Ala Tyr Asn Val Ile Met Cys
 50 55 60
 Val Phe Ser Leu Val Cys Phe Ile Cys Gln Leu Ala Ala Leu Gly Tyr
 65 70 75 80
 Asp Met Gly Tyr Leu Gln Trp Val Arg Asp Leu Thr Gly Asp Glu Ile
 85 90 95
 Val Pro Leu Tyr Gln Asp Val Ser Pro Ser Pro Ala Phe Ser Asn Lys
 100 105 110
 Leu Phe Lys Tyr Ser Ser Ile Ala Phe His Tyr Ser Lys Tyr Val Glu
 115 120 125
 Tyr Met Asp Thr Ala Trp Leu Val Met Lys Gly Lys Pro Val Ser Leu
 130 135 140
 Leu Gln Gly Phe His His Phe Gly Ala Ala Trp Asp Thr Tyr Phe Gly
 145 150 155 160
 Ile Thr Phe Gln Asn Glu Gly Ile Tyr Val Phe Val Val Leu Asn Ala
 165 170 175
 Phe Ile His Thr Ile Met Tyr Ala Tyr Tyr Ala Ala Thr Ala Ala Gly
 180 185 190
 Leu Lys Phe Ser Leu Lys Phe Val Ile Thr Leu Met Gln Ile Thr Gln
 195 200 205
 Phe Asn Val Gly Phe Val Met Val Tyr His Tyr Ile Thr Leu Glu Tyr
 210 215 220
 Phe Arg Asn Ser Pro Glu Leu Val Phe Ser Tyr Leu Phe Asn Tyr Ala
 225 230 235 240
 Tyr Val Cys Thr Val Leu Leu Leu Phe Met Gln Phe Phe Tyr Met Asp
 245 250 255
 Asn Phe Gly Lys Lys Lys Ala Ala Ala Ala Ala Gly Lys Lys Lys Lys
 260 265 270

<210> 49
 <211> 840
 <212> ADN
 <213> Pavlova salina

<400> 49

atggcgactg aagggatgcc ggcgataacg ctggactggc tgctctcgcc cgggctgaag 60
 gatgccgtaa ttggcgggga ggtgctctac ttttcgcttg ggtatctgct gctcgagccc 120
 atcctcaagc gctcaccgtt tgtggacaag cgcaagggcg cataccgcaa cggcatgatc 180

123/143

gcgtacaaca tcctcatgtg cggtttctcg ctggtatgct tcgtgtgcca gatggcggcg 240
 ctcggccttg atcgcggccca cctgcagttt gtccgcgacc tcacggggcga cagcgtggtg 300
 cagctctacc aggacgtgag cccatcccct gcattcgca acaagctctt ccggtactca 360
 gcggtggcgt tccactactc aaagtacgtg gagtacatgg acacagcgtg gcttggtgctg 420
 aagggcaagc ccgtctcggt cctgcagggc ttccaccact tcggcgccgc gtgggacacc 480
 tactttggca tcacgtttca gaacgagggc acctacgtct ttgtgctgct caacgcattc 540
 atccacacaa tcatgtacac ctactacggc gcgacggcag cgggcatcaa aatctcgatg 600
 aagccgctga tcaccctcat gcagatcacg cagttcctgc tgggcttcgc gctcgtctac 660
 ccgtacattg acctcggcta cttccgtgcg tcgcccagagc tcgtgtggag ctacctgttc 720
 aactatgctg acgtactcat ggtgctcttc ctcttcatgc gcttcttcta ccacgacaac 780
 tttagcaagc acaagccaat ctcgcgcatac gactccagca accgcatgaa aaccgagtag 840

<210> 50
 <211> 279
 <212> PRT
 <213> Pavlova salina

<400> 50

Met Ala Thr Glu Gly Met Pro Ala Ile Thr Leu Asp Trp Leu Leu Ser
 1 5 10 15
 Pro Gly Leu Lys Asp Ala Val Ile Gly Gly Glu Val Leu Tyr Phe Ser
 20 25 30
 Leu Gly Tyr Leu Leu Leu Glu Pro Ile Leu Lys Arg Ser Pro Phe Val
 35 40 45
 Asp Lys Arg Lys Gly Ala Tyr Arg Asn Gly Met Ile Ala Tyr Asn Ile
 50 55 60
 Leu Met Cys Gly Phe Ser Leu Val Cys Phe Val Cys Gln Met Ala Ala
 65 70 75 80
 Leu Gly Leu Asp Arg Gly His Leu Gln Phe Val Arg Asp Leu Thr Gly
 85 90 95
 Asp Ser Val Val Gln Leu Tyr Gln Asp Val Ser Pro Ser Pro Ala Phe
 100 105 110
 Ala Asn Lys Leu Phe Arg Tyr Ser Ala Val Ala Phe His Tyr Ser Lys
 115 120 125
 Tyr Val Glu Tyr Met Asp Thr Ala Trp Leu Val Leu Lys Gly Lys Pro
 130 135 140
 Val Ser Phe Leu Gln Gly Phe His His Phe Gly Ala Ala Trp Asp Thr
 145 150 155 160
 Tyr Phe Gly Ile Thr Phe Gln Asn Glu Gly Thr Tyr Val Phe Val Leu
 165 170 175
 Leu Asn Ala Phe Ile His Thr Ile Met Tyr Thr Tyr Tyr Gly Ala Thr
 180 185 190
 Ala Ala Gly Ile Lys Ile Ser Met Lys Pro Leu Ile Thr Leu Met Gln
 195 200 205

124/143

Ile Thr Gln Phe Leu Leu Gly Phe Ala Leu Val Tyr Pro Tyr Ile Asp
 210 215 220
 Leu Gly Tyr Phe Arg Ala Ser Pro Glu Leu Val Trp Ser Tyr Leu Phe
 225 230 235 240
 Asn Tyr Ala Tyr Val Leu Met Val Leu Phe Leu Phe Met Arg Phe Phe
 245 250 255
 Tyr His Asp Asn Phe Ser Lys His Lys Pro Ile Ser Arg Ile Asp Ser
 260 265 270
 Ser Asn Arg Met Lys Thr Glu
 275

<210> 51
 <211> 1284
 <212> ADN
 <213> Pavlova salina

<400> 51
 atgggacgcg gcggagacag cagtgggcag gcgcatccgg cggcggagct ggcggtcccg 60
 agcgaccgcg cggaggtgag caacgctgac agcaaagcgc tgcacatcgt gctgtatggc 120
 aagcgcgtgg atgtgaccaa gttccaacgc acgcaccgcg gtggtagcaa ggtcttccgg 180
 atcttccagg accgcatgac gacggagcag ttcgagtcct accactcgaa gcgcgcgacg 240
 aagatgatgg agggcatgct caagaagtct gaggatgctc ccgccgacac gcccttgccc 300
 tcccagtcac cgatggggaa ggacttcaag gcgatgatcg agcggcacgt tgcagcgggt 360
 tactacgatc catgcccgtc cgatgagctg ttcaagctca gcctcgtgct cctcccgacc 420
 tttgcgggca tgtacatgct caaggcgggc gtcggctccc cgctctgcgg cgccctcatg 480
 gtgagctttg gctggtacct cgatggctgg ctgcgcacg actatctgca ccaactccgtc 540
 ttcaaggggt ccgtcgcacg caccgtcggg tggacaacg cggcgggcta cttcctcggc 600
 ttcgtgcagg ggtatgcggt cgagtgggtg cgcgcgcggc ataacacgca ccacgtgtgc 660
 accaatgagg acggctcggg ccccgacatc aaaacggcgc cgctgctcat atacgtgcgc 720
 aacaagccga gcatcgccaa gcgcctgaac gccttccagc gctaccagca gtactactat 780
 gtgccgggtga tggcaatcct cgacctgtac tggcggctcg agtcgatcgc ctacgtcgcg 840
 atgcgcctgc cgaagatgct gccgcaggcc ctgcgactcg tcgcgcacta cgccatcgtc 900
 gcgtgggtct ttgcgggcaa ctaccacctg ctcccgtcgc tgacggttct gcgcgggttt 960
 ggcactggga tcaccgtttt cgcgacgcac tacggtgagg acattctcga cgcgaccag 1020
 gtgcgtcaca tgacgctcgt cgagcagacg gcactcacct cgcgcaacat ctcgggcggc 1080
 tggctcgtga acgtgctcac cggttctatc tctactgaga cggagcacca cctgttcccg 1140
 atgatgcaa ccggcaacct catgactatc cagcccgagg tgcgcgcctt cttcaagaag 1200
 cacggacttg agtaccgcga gggcaacctc attgagtgcg tgcggcagaa catccgtgcg 1260
 cttgcattcg agcacctgct ttga 1284

<210> 52
 <211> 427
 <212> PRT
 <213> Pavlova salina

125/143

<400> 52

Met Gly Arg Gly Gly Asp Ser Ser Gly Gln Ala His Pro Ala Ala Glu
 1 5 10 15
 Leu Ala Val Pro Ser Asp Arg Ala Glu Val Ser Asn Ala Asp Ser Lys
 20 25 30
 Ala Leu His Ile Val Leu Tyr Gly Lys Arg Val Asp Val Thr Lys Phe
 35 40 45
 Gln Arg Thr His Pro Gly Gly Ser Lys Val Phe Arg Ile Phe Gln Asp
 50 55 60
 Arg Asp Ala Thr Glu Gln Phe Glu Ser Tyr His Ser Lys Arg Ala Ile
 65 70 75 80
 Lys Met Met Glu Gly Met Leu Lys Lys Ser Glu Asp Ala Pro Ala Asp
 85 90 95
 Thr Pro Leu Pro Ser Gln Ser Pro Met Gly Lys Asp Phe Lys Ala Met
 100 105 110
 Ile Glu Arg His Val Ala Ala Gly Tyr Tyr Asp Pro Cys Pro Leu Asp
 115 120 125
 Glu Leu Phe Lys Leu Ser Leu Val Leu Leu Pro Thr Phe Ala Gly Met
 130 135 140
 Tyr Met Leu Lys Ala Gly Val Gly Ser Pro Leu Cys Gly Ala Leu Met
 145 150 155 160
 Val Ser Phe Gly Trp Tyr Leu Asp Gly Trp Leu Ala His Asp Tyr Leu
 165 170 175
 His His Ser Val Phe Lys Gly Ser Val Ala Arg Thr Val Gly Trp Asn
 180 185 190
 Asn Ala Ala Gly Tyr Phe Leu Gly Phe Val Gln Gly Tyr Ala Val Glu
 195 200 205
 Trp Trp Arg Ala Arg His Asn Thr His His Val Cys Thr Asn Glu Asp
 210 215 220
 Gly Ser Asp Pro Asp Ile Lys Thr Ala Pro Leu Leu Ile Tyr Val Arg
 225 230 235 240
 Asn Lys Pro Ser Ile Ala Lys Arg Leu Asn Ala Phe Gln Arg Tyr Gln
 245 250 255
 Gln Tyr Tyr Tyr Val Pro Val Met Ala Ile Leu Asp Leu Tyr Trp Arg
 260 265 270
 Leu Glu Ser Ile Ala Tyr Val Ala Met Arg Leu Pro Lys Met Leu Pro
 275 280 285
 Gln Ala Leu Ala Leu Val Ala His Tyr Ala Ile Val Ala Trp Val Phe
 290 295 300
 Ala Gly Asn Tyr His Leu Leu Pro Leu Val Thr Val Leu Arg Gly Phe
 305 310 315 320
 Gly Thr Gly Ile Thr Val Phe Ala Thr His Tyr Gly Glu Asp Ile Leu
 325 330 335
 Asp Ala Asp Gln Val Arg His Met Thr Leu Val Glu Gln Thr Ala Leu
 340 345 350

126/143

Thr Ser Arg Asn Ile Ser Gly Gly Trp Leu Val Asn Val Leu Thr Gly
 355 360 365
 Phe Ile Ser Leu Gln Thr Glu His His Leu Phe Pro Met Met Pro Thr
 370 375 380
 Gly Asn Leu Met Thr Ile Gln Pro Glu Val Arg Ala Phe Phe Lys Lys
 385 390 395 400
 His Gly Leu Glu Tyr Arg Glu Gly Asn Leu Ile Glu Cys Val Arg Gln
 405 410 415
 Asn Ile Arg Ala Leu Ala Phe Glu His Leu Leu
 420 425

<210> 53
 <211> 172
 <212> PRT
 <213> Virut gây lùn bụi cà chua

<400> 53

Met Glu Arg Ala Ile Gln Gly Asn Asp Ala Arg Glu Gln Ala Asn Ser
 1 5 10 15
 Glu Arg Trp Asp Gly Gly Ser Gly Gly Thr Thr Ser Pro Phe Lys Leu
 20 25 30
 Pro Asp Glu Ser Pro Ser Trp Thr Glu Trp Arg Leu His Asn Asp Glu
 35 40 45
 Thr Asn Ser Asn Gln Asp Asn Pro Leu Gly Phe Lys Glu Ser Trp Gly
 50 55 60
 Phe Gly Lys Val Val Phe Lys Arg Tyr Leu Arg Tyr Asp Arg Thr Glu
 65 70 75 80
 Ala Ser Leu His Arg Val Leu Gly Ser Trp Thr Gly Asp Ser Val Asn
 85 90 95
 Tyr Ala Ala Ser Arg Phe Phe Gly Phe Asp Gln Ile Gly Cys Thr Tyr
 100 105 110
 Ser Ile Arg Phe Arg Gly Val Ser Ile Thr Val Ser Gly Gly Ser Arg
 115 120 125
 Thr Leu Gln His Leu Cys Glu Met Ala Ile Arg Ser Lys Gln Glu Leu
 130 135 140
 Leu Gln Leu Ala Pro Ile Glu Val Glu Ser Asn Val Ser Arg Gly Cys
 145 150 155 160
 Pro Glu Gly Thr Glu Thr Phe Glu Lys Glu Ser Glu
 165 170

<210> 54
 <211> 116
 <212> PRT
 <213> Virut gây bệnh vàng xoắn lá cà chua

<400> 54

Met Trp Asp Pro Leu Leu Asn Glu Phe Pro Glu Ser Val His Gly Phe
 1 5 10 15
 Arg Cys Met Leu Ala Ile Lys Tyr Leu Gln Ser Val Glu Glu Thr Tyr
 20 25 30
 Glu Pro Asn Thr Leu Gly His Asp Leu Ile Arg Asp Leu Ile Ser Val

127/143

35 40 45
 Val Arg Ala Arg Asp Tyr Val Glu Ala Thr Arg Arg Tyr Asn His Phe
 50 55 60
 His Ala Arg Leu Glu Gly Ser Pro Lys Ala Glu Leu Arg Gln Pro Ile
 65 70 75 80
 Gln Gln Pro Cys Cys Cys Pro His Cys Pro Arg His Lys Gln Ala Thr
 85 90 95
 Ile Met Asp Val Gln Ala His Val Pro Glu Ala Gln Asn Ile Gln Asn
 100 105 110
 Val Ser Lys Pro
 115
 <210> 55
 <211> 351
 <212> PRT
 <213> Virut làm xoắn cây tulip
 <400> 55
 Met Glu Asn Asp Pro Arg Val Arg Lys Phe Ala Ser Glu Gly Ala Gln
 1 5 10 15
 Trp Ala Ile Lys Trp Gln Lys Lys Gly Trp Ser Ser Leu Thr Ser Arg
 20 25 30
 Gln Lys Gln Thr Ala Arg Ala Ala Met Gly Ile Lys Leu Ser Pro Val
 35 40 45
 Ala Gln Pro Val Gln Lys Val Thr Arg Leu Ser Ala Pro Val Ala Leu
 50 55 60
 Ala Tyr Arg Glu Val Ser Thr Gln Pro Arg Val Ser Thr Ala Arg Asp
 65 70 75 80
 Gly Ile Thr Arg Ser Gly Ser Glu Leu Ile Thr Thr Leu Lys Lys Asn
 85 90 95
 Thr Asp Thr Glu Pro Lys Tyr Thr Thr Ala Val Leu Asn Pro Ser Glu
 100 105 110
 Pro Gly Thr Phe Asn Gln Leu Ile Lys Glu Ala Ala Gln Tyr Glu Lys
 115 120 125
 Tyr Arg Phe Thr Ser Leu Arg Phe Arg Tyr Ser Pro Met Ser Pro Ser
 130 135 140
 Thr Thr Gly Gly Lys Val Ala Leu Ala Phe Asp Arg Asp Ala Ala Lys
 145 150 155 160
 Pro Pro Pro Asn Asp Leu Ala Ser Leu Tyr Asn Ile Glu Gly Cys Val
 165 170 175
 Ser Ser Val Pro Trp Thr Gly Phe Ile Leu Thr Val Pro Thr Asp Ser
 180 185 190
 Thr Asp Arg Phe Val Ala Asp Gly Ile Ser Asp Pro Lys Leu Val Asn
 195 200 205
 Phe Gly Lys Leu Ile Met Ala Thr Tyr Gly Gln Gly Ala Asn Asp Ala
 210 215 220
 Ala Gln Leu Gly Glu Val Arg Val Glu Tyr Thr Val Gln Leu Lys Asn
 225 230 235 240

128/143

Arg Thr Gly Ser Thr Ser Asp Ala Gln Ile Gly Asp Phe Ala Gly Val
 245 250 255
 Lys Asp Gly Pro Arg Leu Val Ser Trp Ser Lys Thr Lys Gly Thr Ala
 260 265 270
 Gly Trp Glu His Asp Cys His Phe Leu Gly Thr Gly Asn Phe Ser Leu
 275 280 285
 Thr Leu Phe Tyr Glu Lys Ala Pro Val Ser Gly Leu Glu Asn Ala Asp
 290 295 300
 Ala Ser Asp Phe Ser Val Leu Gly Glu Ala Ala Ala Gly Ser Val Gln
 305 310 315 320
 Trp Ala Gly Val Lys Val Ala Glu Arg Gly Gln Ser Val Lys Met Val
 325 330 335
 Thr Thr Glu Glu Gln Pro Arg Gly Lys Trp Gln Ala Leu Arg Ile
 340 345 350

<210> 56

<211> 303

<212> PRT

<213> Virus gây bệnh khảm trên đậu

<400> 56

Met His Gly Ile Glu Gln Pro Gln Leu Pro Leu Asp Tyr Val His Arg
 1 5 10 15
 Cys Ala Ser Thr Ser Phe Leu Leu Ala Ser Leu Asp Gly Leu Leu Ser
 20 25 30
 Glu Ala Arg Glu Leu Ser Gly Pro Leu Ala Leu Ile Thr Ser Ser Tyr
 35 40 45
 Tyr Leu Leu Val Ser Ile Ala Leu Cys Trp Ala Ile Pro Gly Ser Phe
 50 55 60
 Trp Tyr Arg Pro Gly Cys Trp Leu Gln Pro Val Ser Gly Arg Asn Leu
 65 70 75 80
 Ile Phe Cys Gly Pro Thr Glu Ala Leu Gln Arg Phe Arg Leu Tyr Ala
 85 90 95
 Ala Arg Leu Gly Leu Val Leu Ser Glu Asn Cys Pro Arg His Gly Gln
 100 105 110
 Ser Ala Ala Ile Thr Leu Gln Ser Tyr Trp Ala Leu Pro Asn Asn Ile
 115 120 125
 Trp Met Asp Met Ala Gln Leu Asp Leu Leu Thr Phe Ser Met Pro Ile
 130 135 140
 Ala Asn Thr Phe Ala Tyr Leu Ala Asp Cys Glu Ala Arg Phe Pro Pro
 145 150 155 160
 Ile Val Glu Gly Val Gly Ser Ala Tyr Tyr Val Pro Thr Leu Leu Gly
 165 170 175
 Leu Thr His Gln Asp Pro Arg Leu Tyr Leu Ala Leu Arg Arg Arg Asn
 180 185 190
 Leu Asp Leu Ser Gly Glu Pro His Arg Val Arg Pro Gly Val Leu Glu
 195 200 205
 Ser Met Ala Leu Leu Cys Ser Ser Val Arg Ser Thr Ser Arg Ser Arg
 210 215 220

129/143

Gln Ile Pro Pro Leu Tyr Gly Ser Val Leu His His Val Leu Gly Leu
 225 230 235 240
 Ala Glu Arg Asp Cys Ile Leu Phe Asp Thr Asp Ser Asn Tyr Ser Ser
 245 250 255
 Tyr Thr His Arg Val Leu Glu Gln Asp Arg Asn Arg Ala Asp Gln Ser
 260 265 270
 Leu Phe Ser Ile Asp Leu Glu Tyr Val His Asp Leu Glu Leu Ile Ala
 275 280 285
 Leu Gly Tyr Ser Asp Glu Asp Asp Glu Asp Leu Asp Asn Phe Phe
 290 295 300

<210> 57

<211> 256

<212> PRT

<213> virut gây bệnh vàng lùn ở ngũ cốc

<400> 57

Met Phe Ile Ala Gln Pro Cys Gly Arg Val Leu Val Phe Asp Val Ala
 1 5 10 15
 Ser Arg Thr Pro Ser Phe Phe Thr Arg Tyr Ser Val Glu Leu Ser Leu
 20 25 30
 Arg Val Leu Asp Pro Phe Phe Thr Arg Ala Val Thr Asp Phe Arg Tyr
 35 40 45
 Thr Gln Asn Glu Ile Asp Leu Phe Cys Val Ser Leu Gly Phe Leu Leu
 50 55 60
 Pro Ile Leu Leu Thr Gly Glu Ser Tyr Ser Trp Arg Gly His Leu Asn
 65 70 75 80
 Leu Pro Leu Ser Tyr Thr Glu Leu Leu Val Arg Trp Gly Leu Ala Val
 85 90 95
 Gly Tyr Phe Pro Thr Phe Ser Thr Asp Gly Asp Ile Arg Gln Asn Pro
 100 105 110
 Glu Leu Arg Ile Asp Leu Ser Thr Met Ser Thr Arg Ser Phe Tyr Glu
 115 120 125
 Gln Phe Leu Leu Arg Tyr Asn Thr Ser Gly Leu Ala Lys Ala Ile Val
 130 135 140
 Gly Gln Gln Glu Cys Phe Gln Ser Gly Met Glu Ser Phe Lys Arg Phe
 145 150 155 160
 Leu His Tyr Arg Leu Thr Cys Phe Glu Ser Cys Leu Pro Arg Pro Arg
 165 170 175
 Trp Glu Ser Pro Leu Ala Pro Gly Pro Tyr Leu Asp Arg Ala Phe Glu
 180 185 190
 Ala Thr Leu Leu Gly Arg Met Val Gly His Asn Gln Leu Leu Phe Thr
 195 200 205
 Gly Leu Ser Ser Asp Ile Thr Arg Tyr Tyr Asn Glu Leu Val Val Glu
 210 215 220
 Gly Val Pro Val Ala Phe Trp Asp Ala Ala Gly Ile Thr Leu His His
 225 230 235 240
 Ala Gly Glu Glu Tyr Phe Ser Asn Ser Tyr Ile Gln Lys Ile Leu Gln

130/143

245

250

255

<210> 58
 <211> 519
 <212> ADN
 <213> Virut gây lùn bụi cà chua

<400> 58
 atggaacgag ctatacaagg aaacgacgct agggaacaag ctaacagtga acgttgggat 60
 ggaggatcag gaggtaccac ttctcccttc aaacttcctg acgaaagtcc gagttggact 120
 gagtggcggc tacataacga tgagacgaat tcgaatcaag ataatcccct tggtttcaag 180
 gaaagctggg gtttcgggaa agttgtatct aagagatatt tcagatacga caggacggaa 240
 gcttcactgc acagagtcct tggatcttgg acgggagatt cggttaacta tgcagcatct 300
 cgatttttcg gtttcgacca gatcggatgt acctatagta ttcggtttcg aggagttagt 360
 atcaccgttt ctggagggtc gcgaactctt cagcatctct gtgagatggc aattcgggtct 420
 aagcaagaac tgctacagct tgccccaatc gaagtggaaa gtaatgtatc aagaggatgc 480
 cctgaaggta ctgagacctt cgaaaaagaa agcgagtaa 519

<210> 59
 <211> 351
 <212> ADN
 <213> Virut gây bệnh vàng xoắn lá cà chua

<400> 59
 atgtgggagc cacttctaaa tgaatttcct gaatctgttc acggatttcg ttgtatgtta 60
 gctattaaat atttgcagtc cgttgaggaa acttacgagc ccaatacatt gggccacgat 120
 ttaattaggg atcttatatc tgttgtaagg gcccgtagct atgtcgaagc gaccaggcga 180
 tataatcatt tccacgccc cctcgaagg tgcgcaagg ctgaacttcg acagcccata 240
 cagcagccgt gctgctgtcc ccattgtcca aggcacaaac aagcgacgat catggacgta 300
 caggcccatg taccggaagc ccagaatata cagaatgtat cgaagccctg a 351

<210> 60
 <211> 1056
 <212> ADN
 <213> Virut làm xoắn cây tulip

<400> 60
 atggaaaatg atcctagagt ccgaaagtgc gcatccgagg gcgccaatg ggcgataaag 60
 tggcagaaga agggctgggc atccctaacc agcagacaga aacagaccgc ccgcgacgag 120
 atggggatca agctctcccc tgtggcgcaa cctgtgcaga aagtgactcg actgagtgtc 180
 ccggtggctc tcgcctaccg cgaggtttcc acccagcctc gggtttctac tgccagggac 240
 ggcataacca gaagcgggtc tgaactgac acaaccctga agaagaacac tgacactgaa 300
 cctaagtaca ccacagctgt gcttaacca agcgaaccgc gaacattcaa ccaactcatc 360
 aaggaggcgg ccagtagatg aaaataccga ttcacgtcac tcagatttag gtactctccc 420
 atgagccctt caaccaccgg gggcaagggt gctctggcat tcgaccgaga cgctgcaaaa 480
 cctccgcca acgacctgc ttccctctac aacatagagg gttgtgtatc tagcgtgccc 540
 tggacagggt ttattttgac cgtccaaca gattctactg accgctttgt ggcggatggt 600

131/143

atcagcgcgtc caaagcttgt caatttcggc aagctcatca tggccaccta tggccaagga 660
 gccaatgatg ccgcccact cggtgaagtg cgagtcgagt acaccgtgca gctcaagaac 720
 agaactggct caaccagcga cgcccagatt ggggacttcg cgggtgttaa ggacggaccc 780
 aggttggctt cgtggtccaa gaccaaggga acagctgggt gggagcacga ttgtcatttt 840
 ctcggaaccg gaaacttctc gttgacactg ttctacgaga aggcgcccgt ttcggggcta 900
 gaaaacgcag acgcctctga cttctcggtc ctgggagaag ccgcagcagg tagtgtccaa 960
 tgggctggag tgaaggtagc agaaagggga caaagcgtga aaatggtcac aactgaggag 1020
 cagccaaggg gaaaatggca agcactcagg atttag 1056

<210> 61

<211> 912

<212> ADN

<213> Virut gây bệnh khảm trên đậu

<400> 61

atgcacggaa ttgagcagcc tcaactaccg ctagattacg ttcaccgttg cgcatacaacc 60
 tccttcttgc tcgcatcact agatggcctc ctttctgaag cccgtgaact ctcagggcct 120
 ctggctctca ttacttctag ctattactta cttgtttcta ttgccctctg ctgggcaatc 180
 cctggatcct tctggtatag gcctggctgc tggttgcagc cagtctcagg gcggaatctc 240
 atcttttgcg gccctaccga ggccttgcaa cgattccgtc tgtacgctgc cagacttggg 300
 ttggtcctgt cagagaactg cccaagacac ggccaatcag cagcaatcac ctttcaatca 360
 tactgggcac ttcctaacia catctggatg gacatggccc aattggactt gctcaccttc 420
 tcaatgccaa ttgctaatac atttgcctac ttggcagatt gtgaagcaag atttcctcct 480
 attgttgaag gagtgggacg tgcttactat gtgccaacgc tgctcggact tactcaccaa 540
 gaccccaggc tttatcttgc gcttcgcagg agaaaccttg atcttagtgg cgaacctcat 600
 agagttcgtc ctggtgtcct ggagtctatg gctttgctct gttctagtgt acgtagcaca 660
 agccgttcca ggcaaattcc tcctttatat ggcagcgttt tgcaccacgt tttgggcctg 720
 gccgagagag actgcatcct ctttgatacg gatagtaact actcctctta cactcatcgg 780
 gttcttgaac aagaccggaa tcgggctgat cagtcattgt ttagcattga cttggaatat 840
 gttcatgacc tggagcttat tgccctgggt tactctgatg aagatgatga agatcttgat 900
 aacttcttct ag 912

<210> 62

<211> 771

<212> ADN

<213> Virut gây bệnh vàng lùn ở ngũ cốc

<400> 62

atgttcatcg cccaaccttg cgggagagtt cttgtgttcg acgtcgccct ccgcacgcca 60
 tcgtttctca ctcgttatag tgttgaactc tcgctccgtg ttctagacct attcttcacg 120
 cgagcagtaa cagatttccg atacacccaa aatgaaatcg atttattttg tgtgtctctt 180
 ggctttctgt tgccaattct cctcacagga gaatcttact cttggcgcgg tcacctcaac 240

132/143

ctcccccttt cttacaccga attacttggt cgatgggggc tcgcagtggg gtacttcctt 300
 acctttctcca ctgatggtga cattcgacag aaccagaac tccgcatcga cctgtccacc 360
 atgtcaaccc gctctttcta cgagcagttc ctactcagat ataacacgag tgggttggca 420
 aaagctatcg tcggacagca agagtgcctt caaagcggca tggagtcttt taaaagattc 480
 ctacactacc gcctcacgtg ctttgaaagc tgccttcac gacctcgtt ggaaagtcct 540
 ttggctcctg gtccttatct ggacagggct tttgaggcaa ctcttctcgg ccgtatggtc 600
 ggtcataacc aactactctt taccggtttg tcttctgata tcactaggta ttataacgag 660
 ttggttggtg aaggcgtgcc ggtggctttt tgggacgctg ccggcattac tttgcatcac 720
 gctggtgaag aatatttttc gaattcttac attcaaaaga ttcttcaatg a 771

<210> 63

<211> 389

<212> PRT

<213> *Arabidopsis thaliana*

<400> 63

Met Val Ile Ala Ala Val Ile Val Pro Leu Gly Leu Leu Phe Phe
 1 5 10 15
 Ile Ser Gly Leu Ala Val Asn Leu Phe Gln Ala Val Cys Tyr Val Leu
 20 25 30
 Ile Arg Pro Leu Ser Lys Asn Thr Tyr Arg Lys Ile Asn Arg Val Val
 35 40 45
 Ala Glu Thr Leu Trp Leu Glu Leu Val Trp Ile Val Asp Trp Trp Ala
 50 55 60
 Gly Val Lys Ile Gln Val Phe Ala Asp Asn Glu Thr Phe Asn Arg Met
 65 70 75 80
 Gly Lys Glu His Ala Leu Val Val Cys Asn His Arg Ser Asp Ile Asp
 85 90 95
 Trp Leu Val Gly Trp Ile Leu Ala Gln Arg Ser Gly Cys Leu Gly Ser
 100 105 110
 Ala Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile Gly
 115 120 125
 Trp Ser Met Trp Phe Ser Glu Tyr Leu Phe Leu Glu Arg Asn Trp Ala
 130 135 140
 Lys Asp Glu Ser Thr Leu Lys Ser Gly Leu Gln Arg Leu Ser Asp Phe
 145 150 155 160
 Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe Thr
 165 170 175
 Glu Ala Lys Leu Lys Ala Ala Gln Glu Tyr Ala Ala Ser Ser Glu Leu
 180 185 190
 Pro Ile Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val Ser
 195 200 205
 Ala Val Ser Asn Met Arg Ser Phe Val Pro Ala Ile Tyr Asp Met Thr
 210 215 220
 Val Thr Ile Pro Lys Thr Ser Pro Pro Pro Thr Met Leu Arg Leu Phe
 225 230 235 240

Lys Gly Gln Pro Ser Val Val His Val Ile Lys Cys His Ser Met
245 250 255

Lys Asp Leu Pro Glu Ser Asp Asp Ala Ile Ala Gln Trp Cys Arg Asp
260 270

Gln Phe Val Ala Lys Asp Ala Leu Leu Asp Lys His Ile Ala Ala Asp
275 280 285

Thr Phe Pro Gly Gln Gln Glu Gln Asn Ile Gly Arg Pro Ile Lys Ser
290 295 300

Leu Ala Val Val Leu Ser Trp Ala Cys Val Leu Thr Leu Gly Ala Ile
305 310 315 320

Lys Phe Leu His Trp Ala Gln Leu Phe Ser Ser Trp Lys Gly Ile Thr
325 330 335

Ile Ser Ala Leu Gly Leu Gly Ile Ile Thr Leu Cys Met Gln Ile Leu
340 345

Ile Arg Ser Ser Gln Ser Glu Arg Ser Thr Pro Ala Lys Val Val Pro
355 360 365

Ala Lys Pro Lys Asp Asn His His Pro Glu Ser Ser Ser Gln Thr Glu
370 375 380

Thr Glu Lys Glu Lys
385

<210> 64
<211> 281
<212> PRT
<213> Limnanthes alba

<400> 64

Met Ala Lys Thr Arg Thr Ser Ser Leu Arg Asn Arg Arg Gln Leu Lys
1 5 10 15

Thr Ala Val Ala Ala Thr Ala Asp Asp Asp Lys Asp Gly Ile Phe Met
20 25 30

Val Leu Leu Ser Cys Phe Lys Ile Phe Val Cys Phe Ala Ile Val Leu
35 40 45

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

Pro Tyr Met Arg Ile Arg Leu Gly Asn Leu Tyr Gly His Ile Ile Gly
65 70 75 80

Gly Leu Val Ile Trp Leu Tyr Gly Ile Pro Ile Glu Ile Gln Gly Ser
85 90 95

Glu His Thr Lys Lys Arg Ala Ile Tyr Ile Ser Asn His Ala Ser Pro
100 105 110

Ile Asp Ala Phe Phe Val Met Trp Leu Ala Pro Ile Gly Thr Val Gly
115 120 125

Val Ala Lys Lys Glu Val Ile Trp Tyr Pro Leu Leu Gly Gln Leu Tyr
130 135 140

Thr Leu Ala His His Ile Arg Ile Asp Arg Ser Asn Pro Ala Ala Ala
145 150 155 160

Ile Gln Ser Met Lys Glu Ala Val Arg Val Ile Thr Glu Lys Asn Leu

134/143

165										170					175						
Ser	Leu	Ile	Met	Phe	Pro	Glu	Gly	Thr	Arg	Ser	Gly	Asp	Gly	Arg	Leu						
			180					185					190								
Leu	Pro	Phe	Lys	Lys	Gly	Phe	Val	His	Leu	Ala	Leu	Gln	Ser	His	Leu						
		195					200					205									
Pro	Ile	Val	Pro	Met	Ile	Leu	Thr	Gly	Thr	His	Leu	Ala	Trp	Arg	Lys						
		210				215					220										
Gly	Thr	Phe	Arg	Val	Arg	Pro	Val	Pro	Ile	Thr	Val	Lys	Tyr	Leu	Pro						
225					230				235						240						
Pro	Ile	Asn	Thr	Asp	Asp	Trp	Thr	Val	Asp	Lys	Ile	Asp	Asp	Tyr	Val						
				245					250					255							
Lys	Met	Ile	His	Asp	Ile	Tyr	Val	Arg	Asn	Leu	Pro	Ala	Ser	Gln	Lys						
			260					265					270								
Pro	Leu	Gly	Ser	Thr	Asn	Arg	Ser	Lys													
		275					280														

<210> 65
 <211> 303
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 65

Met	Ser	Val	Ile	Gly	Arg	Phe	Leu	Tyr	Tyr	Leu	Arg	Ser	Val	Leu	Val						
1				5					10					15							
Val	Leu	Ala	Leu	Ala	Gly	Cys	Gly	Phe	Tyr	Gly	Val	Ile	Ala	Ser	Ile						
			20					25					30								
Leu	Cys	Thr	Leu	Ile	Gly	Lys	Gln	His	Leu	Ala	Gln	Trp	Ile	Thr	Ala						
		35				40					45										
Arg	Cys	Phe	Tyr	His	Val	Met	Lys	Leu	Met	Leu	Gly	Leu	Asp	Val	Lys						
	50					55					60										
Val	Val	Gly	Glu	Glu	Asn	Leu	Ala	Lys	Lys	Pro	Tyr	Ile	Met	Ile	Ala						
65					70					75					80						
Asn	His	Gln	Ser	Thr	Leu	Asp	Ile	Phe	Met	Leu	Gly	Arg	Ile	Phe	Pro						
				85					90					95							
Pro	Gly	Cys	Thr	Val	Thr	Ala	Lys	Lys	Ser	Leu	Lys	Tyr	Val	Pro	Phe						
			100					105					110								
Leu	Gly	Trp	Phe	Met	Ala	Leu	Ser	Gly	Thr	Tyr	Phe	Leu	Asp	Arg	Ser						
		115					120					125									
Lys	Arg	Gln	Glu	Ala	Ile	Asp	Thr	Leu	Asn	Lys	Gly	Leu	Glu	Asn	Val						
		130				135					140										
Lys	Lys	Asn	Lys	Arg	Ala	Leu	Trp	Val	Phe	Pro	Glu	Gly	Thr	Arg	Ser						
145					150					155					160						
Tyr	Thr	Ser	Glu	Leu	Thr	Met	Leu	Pro	Phe	Lys	Lys	Gly	Ala	Phe	His						
				165					170					175							
Leu	Ala	Gln	Gln	Gly	Lys	Ile	Pro	Ile	Val	Pro	Val	Val	Val	Ser	Asn						
			180					185					190								
Thr	Ser	Thr	Leu	Val	Ser	Pro	Lys	Tyr	Gly	Val	Phe	Asn	Arg	Gly	Cys						
			195				200					205									

135/143

Met Ile Val Arg Ile Leu Lys Pro Ile Ser Thr Glu Asn Leu Thr Lys
 210 215 220
 Asp Lys Ile Gly Glu Phe Ala Glu Lys Val Arg Asp Gln Met Val Asp
 225 230 235 240
 Thr Leu Lys Glu Ile Gly Tyr Ser Pro Ala Ile Asn Asp Thr Thr Leu
 245 250 255
 Pro Pro Gln Ala Ile Glu Tyr Ala Ala Leu Gln His Asp Lys Lys Val
 260 265 270
 Asn Lys Lys Ile Lys Asn Glu Pro Val Pro Ser Val Ser Ile Ser Asn
 275 280 285
 Asp Val Asn Thr His Asn Glu Gly Ser Ser Val Lys Lys Met His
 290 295 300

<210> 66
 <211> 373
 <212> PRT
 <213> Micromonas pusilla

<400> 66

Met Thr Pro Tyr Gln Trp Phe Asn Val Val Ser Ser Leu Gly Tyr Val
 1 5 10
 Leu Phe Thr Ala Thr Thr Ser Thr Val Thr Met Leu Val Pro Ala Ile
 20 25 30
 Ile Leu Leu Arg Pro Val Ser Ala Asn Leu Tyr Ala Arg Cys Thr Ser
 35 40 45
 Trp Ile Phe Ala Cys Trp Trp Thr Ser Cys Leu Phe Ile Thr Glu Arg
 50 55 60
 Leu Asn Gly Val Lys Val Arg Val Thr Gly Asp Ala Leu Pro Leu Asn
 65 70 75 80
 Ala Pro Leu Leu Ile Met Ser Asn His Lys Cys Asn Leu Asp Trp Met
 85 90 95
 Phe Leu Trp Ser Ser Ala Ile Arg Thr Gly Ser Met Phe His Val Gly
 100 105 110
 Val Phe Lys Ala Val Ala Lys Ser Glu Ile Arg Val Ile Pro Ile Phe
 115 120 125
 Gly Trp Gly Cys Lys Leu Asn Gly Phe Ala Tyr Val Arg Arg Arg Trp
 130 135 140
 Ser Ser Asp Ala Ser His Leu Thr Ser Trp Ile Gln Ser Gln Ile Arg
 145 150 155 160
 Arg Arg Leu Asn Ala Asn Trp Thr Leu Ile Phe Pro Glu Gly Thr Arg
 165 170 175
 Tyr Thr Asp Arg Asn Lys Glu Arg Ser Asp Leu Ser Cys Ala Lys Asp
 180 185 190
 Gly Leu Glu Pro Met Ala Gly Glu Ile Leu Arg Pro Arg Thr Lys Gly
 195 200 205
 Leu Ala Leu Leu Leu Arg Glu Ser Ala Lys Gly Gly Gly Tyr Tyr Arg
 210 215 220
 Lys Ile Val Asp Met Thr Ile Gln Tyr Thr Asp Ala Asp Gly Lys Pro
 225 230 235 240

136/143

Leu Lys Gly Ala Ala Leu Gly Thr Arg Cys Phe Gly Gln Leu Ala Lys
 245 250 255
 Gly Gln Leu Pro Val Ala Thr Cys His Val His Phe Asp Val Phe Ser
 260 265 270
 His Lys Asp Val Pro Ala Gly Glu Asp Glu Asp Glu Val Glu Ala Trp
 275 280 285
 Val Trp Lys Arg Trp Arg Lys Lys Ala Asn Met Leu Glu Ala Cys Ala
 290 295 300
 Ser Ala Gly Gln Phe Glu Gly Val Arg Glu Trp Ser Thr Ser Gly Thr
 305 310 315 320
 Ala Val Pro Leu Lys Thr Gln Thr Ala Leu Arg Cys Phe Phe Val Leu
 325 330 335
 Gln Gly Leu Val Cys Val Gly Val Ala Cys Ser Ser Thr Ala Phe Leu
 340 345 350
 Ala Tyr Val Ala Cys Ala Ala Val Gly Ala Ala Val Ile Ala Gln Thr
 355 360 365
 Asp Pro Ala Trp Trp
 370

<210> 67
 <211> 314
 <212> PRT
 <213> Mortierella alpina

<400> 67

Met Ser Ile Gly Ser Ser Asn Pro Val Leu Leu Ala Ala Ile Pro Phe
 1 5 10 15
 Val Tyr Leu Phe Val Leu Pro Arg Val Leu Ala Phe Leu Pro Gln Lys
 20 25 30
 Ala Gln Phe Leu Ala Lys Cys Ile Val Val Leu Ile Ala Thr Leu Ile
 35 40 45
 Met Ser Val Ala Gly Cys Phe Ile Ser Ile Val Cys Ala Leu Leu Asp
 50 55 60
 Lys Arg Tyr Val Ile Asn Tyr Val Val Ser Arg Leu Phe Ser Phe Leu
 65 70 75 80
 Ala Ala Arg Pro Cys Gly Val Thr Tyr Lys Ile Val Gly Glu Glu His
 85 90 95
 Leu Asp Lys Tyr Pro Ala Ile Val Val Cys Asn His Gln Ser Ser Met
 100 105 110
 Asp Met Met Val Leu Gly Arg Val Phe Pro Lys His Cys Val Val Met
 115 120 125
 Ala Lys Lys Glu Leu Leu Tyr Phe Pro Phe Leu Gly Met Phe Met Lys
 130 135 140
 Leu Ser Asn Ala Ile Phe Ile Asp Arg Lys Asn His Lys Lys Ala Ile
 145 150 155 160
 Glu Ser Thr Thr Gln Ala Val Ala Asp Met Lys Lys His Asn Ser Gly
 165 170 175
 Ile Trp Ile Phe Pro Glu Gly Thr Arg Ser Arg Leu Asp Lys Ala Asp

137/143

180 185 190
 Leu Leu Pro Phe Lys Lys Gly Ala Phe His Leu Ala Ile Gln Ala Gln
 195 200 205
 Leu Pro Ile Leu Pro Ile Ile Ser Gln Gly Tyr Ser His Ile Tyr Asp
 210 215 220
 Ser Ser Lys Arg Tyr Phe Pro Gly Gly Glu Leu Glu Ile Arg Val Leu
 225 230 235 240
 Glu Pro Ile Pro Thr Thr Gly Leu Thr Thr Asp Asp Val Asn Asp Leu
 245 250 255
 Met Asp Lys Thr Arg Asn Leu Met Leu Lys His Leu Lys Glu Met Asp
 260 265 270
 Ser Gln Tyr Ser Ser Ser Thr Ala Glu Asn Gly Ser Thr His Ile Asp
 275 280 285
 Ala Asp Ile Ala Lys Ser Thr Ala Thr Ser Ile Gly Asn Thr Asp Asp
 290 295 300
 Ala Ile Thr Lys Arg Arg Thr Pro Lys Glu
 305 310
 <210> 68
 <211> 391
 <212> PRT
 <213> Braccisa napus
 <400> 68
 Met Ala Met Ala Ala Ala Ala Val Ile Val Pro Leu Gly Ile Leu Phe
 1 5 10 15
 Phe Ile Ser Gly Leu Val Val Asn Leu Leu Gln Ala Val Cys Tyr Val
 20 25 30
 Leu Ile Arg Pro Leu Ser Lys Asn Thr Tyr Arg Lys Ile Asn Arg Val
 35 40 45
 Val Ala Glu Thr Leu Trp Leu Glu Leu Val Trp Ile Val Asp Trp Trp
 50 55 60
 Ala Gly Val Lys Ile Gln Val Phe Ala Asp Asp Glu Thr Phe Asn Arg
 65 70 75 80
 Met Gly Lys Glu His Ala Leu Val Val Cys Asn His Arg Ser Asp Ile
 85 90 95
 Asp Trp Leu Val Gly Trp Ile Leu Ala Gln Arg Ser Gly Cys Leu Gly
 100 105 110
 Ser Ala Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile
 115 120 125
 Gly Trp Ser Met Trp Phe Ser Glu Tyr Leu Phe Leu Glu Arg Asn Trp
 130 135 140
 Ala Lys Asp Glu Ser Thr Leu Lys Ser Gly Leu Gln Arg Leu Asn Asp
 145 150 155 160
 Phe Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe
 165 170 175
 Thr Glu Ala Lys Leu Lys Ala Ala Gln Glu Tyr Ala Ala Ser Ser Gln
 180 185 190

138/143

Leu Pro Val Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val
 195 200 205
 Ser Ala Val Ser Asn Met Arg Ser Phe Val Pro Ala Ile Tyr Asp Met
 210 215 220
 Thr Val Ala Ile Pro Lys Thr Ser Pro Pro Thr Met Leu Arg Leu
 225 230 235 240
 Phe Lys Gly Gln Pro Ser Val Val His Val His Ile Lys Cys His Ser
 245 250 255
 Met Lys Asp Leu Pro Glu Ser Asp Asp Ala Ile Ala Gln Trp Cys Arg
 260 265 270
 Asp Gln Phe Val Ala Lys Asp Ala Leu Leu Asp Lys His Ile Ala Ala
 275 280 285
 Asp Thr Phe Pro Gly Gln Lys Glu His Asn Ile Gly Arg Pro Ile Lys
 290 295 300
 Ser Leu Ala Val Val Val Ser Trp Ala Cys Leu Leu Thr Leu Gly Ala
 305 310 315 320
 Met Lys Phe Leu His Trp Ser Asn Leu Phe Ser Ser Leu Lys Gly Ile
 325 330 335
 Ala Leu Ser Ala Leu Gly Leu Gly Ile Ile Thr Leu Cys Met Gln Ile
 340 345 350
 Leu Ile Arg Ser Ser Gln Ser Glu Arg Ser Thr Pro Ala Lys Val Ala
 355 360 365
 Pro Ala Lys Pro Lys Asp Lys His Gln Ser Gly Ser Ser Ser Gln Thr
 370 375 380
 Glu Val Glu Glu Lys Gln Lys
 385 390

<210> 69

<211> 390

<212> PRT

<213> Braccisa napus

<400> 69

Met Ala Met Ala Ala Ala Val Ile Val Pro Leu Gly Ile Leu Phe Phe
 1 5 10 15
 Ile Ser Gly Leu Val Val Asn Leu Leu Gln Ala Ile Cys Tyr Val Leu
 20 25 30
 Ile Arg Pro Leu Ser Lys Asn Thr Tyr Arg Lys Ile Asn Arg Val Val
 35 40 45
 Ala Glu Thr Leu Trp Leu Glu Leu Val Trp Ile Val Asp Trp Trp Ala
 50 55 60
 Gly Val Lys Ile Gln Val Phe Ala Asp Asn Glu Thr Phe Asn Arg Met
 65 70 75 80
 Gly Lys Glu His Ala Leu Val Val Cys Asn His Arg Ser Asp Ile Asp
 85 90 95
 Trp Leu Val Gly Trp Ile Leu Ala Gln Arg Ser Gly Cys Leu Gly Ser
 100 105 110
 Ala Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile Gly
 115 120 125

139/143

Trp Ser Met Trp Phe Ser Glu Tyr Leu Phe Leu Glu Arg Asn Trp Ala
 130 135 140
 Lys Asp Glu Ser Thr Leu Lys Ser Gly Leu Gln Arg Leu Asn Asp Phe
 145 150 155 160
 Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe Thr
 165 170 175
 Glu Ala Lys Leu Lys Ala Ala Gln Glu Tyr Ala Ala Ser Ser Glu Leu
 180 185 190
 Pro Val Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val Ser
 195 200 205
 Ala Val Ser Asn Met Arg Ser Phe Val Pro Ala Ile Tyr Asp Met Thr
 210 215 220
 Val Ala Ile Pro Lys Thr Ser Pro Pro Pro Thr Met Leu Arg Leu Phe
 225 230 235 240
 Lys Gly Gln Pro Ser Val Val His Val His Ile Lys Cys His Ser Met
 245 250 255
 Lys Asp Leu Pro Glu Ser Asp Asp Ala Ile Ala Gln Trp Cys Arg Asp
 260 265 270
 Gln Phe Val Ala Lys Asp Ala Leu Leu Asp Lys His Ile Ala Ala Asp
 275 280 285
 Thr Phe Pro Gly Gln Gln Glu Gln Asn Ile Gly Arg Pro Ile Lys Ser
 290 295 300
 Leu Ala Val Val Leu Ser Trp Ser Cys Leu Leu Ile Leu Gly Ala Met
 305 310 315 320
 Lys Phe Leu His Trp Ser Asn Leu Phe Ser Ser Trp Lys Gly Ile Ala
 325 330 335
 Phe Ser Ala Leu Gly Leu Gly Ile Ile Thr Leu Cys Met Gln Ile Leu
 340 345 350
 Ile Arg Ser Ser Gln Ser Glu Arg Ser Thr Pro Ala Lys Val Val Pro
 355 360 365
 Ala Lys Pro Lys Asp Asn His Asn Asp Ser Gly Ser Ser Ser Gln Thr
 370 375 380
 Glu Val Glu Lys Gln Lys
 385 390

<210> 70
 <211> 361
 <212> PRT
 <213> Phytophthora infestans

<400> 70

Met Ala Thr Lys Glu Ala Tyr Val Phe Pro Thr Leu Thr Glu Ile Lys
 1 5 10 15
 Arg Ser Leu Pro Lys Asp Cys Phe Glu Ala Ser Val Pro Leu Ser Leu
 20 25 30
 Tyr Tyr Thr Val Arg Cys Leu Val Ile Ala Val Ala Leu Thr Phe Gly
 35 40 45

140/143

Leu Asn Tyr Ala Arg Ala Leu Pro Glu Val Glu Ser Phe Trp Ala Leu
 50 55 60
 Asp Ala Ala Leu Cys Thr Gly Tyr Ile Leu Leu Gln Gly Ile Val Phe
 65 70 75 80
 Trp Gly Phe Phe Thr Val Gly His Asp Ala Gly His Gly Ala Phe Ser
 85 90 95
 Arg Tyr His Leu Leu Asn Phe Val Val Gly Thr Phe Met His Ser Leu
 100 105 110
 Ile Leu Thr Pro Phe Glu Ser Trp Lys Leu Thr His Arg His His His
 115 120 125
 Lys Asn Thr Gly Asn Ile Asp Arg Asp Glu Val Phe Tyr Pro Gln Arg
 130 135 140
 Lys Ala Asp Asp His Pro Leu Ser Arg Asn Leu Ile Leu Ala Leu Gly
 145 150 155 160
 Ala Ala Trp Leu Ala Tyr Leu Val Glu Gly Phe Pro Pro Arg Lys Val
 165 170 175
 Asn His Phe Asn Pro Phe Glu Pro Leu Phe Val Arg Gln Val Ser Ala
 180 185 190
 Val Val Ile Ser Leu Leu Ala His Phe Phe Val Ala Gly Leu Ser Ile
 195 200 205
 Tyr Leu Ser Leu Gln Leu Gly Leu Lys Thr Met Ala Ile Tyr Tyr Tyr
 210 215 220
 Gly Pro Val Phe Val Phe Gly Ser Met Leu Val Ile Thr Thr Phe Leu
 225 230 235 240
 His His Asn Asp Glu Glu Thr Pro Trp Tyr Ala Asp Ser Glu Trp Thr
 245 250 255
 Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr Gly Ala Leu
 260 265 270
 Ile Asp Asn Leu Ser His Asn Ile Gly Thr His Gln Ile His His Leu
 275 280 285
 Phe Pro Ile Ile Pro His Tyr Lys Leu Lys Lys Ala Thr Ala Ala Phe
 290 295 300
 His Gln Ala Phe Pro Glu Leu Val Arg Lys Ser Asp Glu Pro Ile Ile
 305 310 315 320
 Lys Ala Phe Phe Arg Val Gly Arg Leu Tyr Ala Asn Tyr Gly Val Val
 325 330 335
 Asp Gln Glu Ala Lys Leu Phe Thr Leu Lys Glu Ala Lys Ala Ala Thr
 340 345 350
 Glu Ala Ala Ala Lys Thr Lys Ser Thr
 355 360

<210> 71
 <211> 418
 <212> PRT
 <213> Thalassiosira pseudonana

<400> 71

Met Tyr Arg Leu Thr Ser Thr Phe Leu Ile Ala Leu Ala Phe Ser Ser

141/143

1	5	10	15
Ser Ile Asn Ala Phe Ser Pro Gln Arg Pro Pro Arg Thr Ile Thr Lys	20	25	30
Ser Lys Val Gln Ser Thr Val Leu Pro Ile Pro Thr Lys Asp Asp Leu	35	40	45
Asn Phe Leu Gln Pro Gln Leu Asp Glu Asn Asp Leu Tyr Leu Asp Asp	50	55	60
Val Asn Thr Pro Pro Arg Ala Gly Thr Ile Met Lys Met Leu Pro Lys	65	70	75
Glu Thr Phe Asn Ile Asp Thr Ala Thr Ser Leu Gly Tyr Phe Gly Met	85	90	95
Asp Met Ala Ala Val Val Ser Ser Met Thr Leu Leu Asn Ala Ile Val	100	105	110
Thr Ser Asp Gln Tyr His Ala Leu Pro Leu Pro Leu Gln Ala Ala Thr	115	120	125
Val Ile Pro Phe Gln Leu Leu Ala Gly Phe Ala Met Trp Cys Met Trp	130	135	140
Cys Ile Gly His Asp Ala Gly His Ser Thr Val Ser Lys Thr Lys Trp	145	150	155
Ile Asn Arg Val Val Gly Glu Val Ala His Ser Val Val Cys Leu Thr	165	170	175
Pro Phe Val Pro Trp Gln Met Ser His Arg Lys His His Leu Asn His	180	185	190
Asn His Ile Glu Lys Asp Tyr Ser His Lys Trp Tyr Ser Arg Asp Glu	195	200	205
Phe Asp Asp Ile Pro Gln Leu Tyr Lys Thr Phe Gly Tyr Asn Pro Arg	210	215	220
Met Met Gln Leu Pro Phe Leu Tyr Phe Met Tyr Leu Ala Leu Gly Ile	225	230	235
Pro Asp Gly Gly His Val Val Phe Tyr Gly Arg Met Trp Glu Gly Val	245	250	255
Ser Leu Gln Lys Lys Phe Asp Ala Ala Ile Ser Val Ala Val Ser Cys	260	265	270
Ala Thr Ala Gly Ser Leu Trp Met Asn Met Gly Thr Ala Asp Phe Thr	275	280	285
Val Val Cys Met Val Pro Trp Leu Val Leu Ser Trp Trp Leu Phe Met	290	295	300
Val Thr Tyr Leu Gln His His Ser Glu Asp Gly Lys Leu Tyr Thr Asp	305	310	315
Glu Thr Phe Thr Phe Glu Lys Gly Ala Phe Glu Thr Val Asp Arg Ser	325	330	335
Tyr Gly Lys Leu Ile Asn Arg Met Ser His His Met Met Asp Gly His	340	345	350
Val Val His His Leu Phe Phe Glu Arg Val Pro His Tyr Arg Leu Glu	355	360	365
Ala Ala Thr Glu Ala Leu Val Lys Gly Met Asp Glu Thr Gly Gln Lys			

142/143

370

375

380

His Leu Tyr Lys Tyr Ile Asp Thr Pro Asp Phe Asn Ala Glu Ile Val
 385 390 395 400

Asn Gly Phe Arg Asp Asn Trp Phe Leu Val Glu Glu Glu Asn Ile Lys
 405 410 415

Arg Glu

<210> 72
 <211> 363
 <212> PRT
 <213> Pythium irregulare

<400> 72

Met Ala Ser Thr Ser Ala Ala Gln Asp Ala Ala Pro Tyr Glu Phe Pro
 1 5 10 15

Ser Leu Thr Glu Ile Lys Arg Ala Leu Pro Ser Glu Cys Phe Glu Ala
 20 25 30

Ser Val Pro Leu Ser Leu Tyr Tyr Thr Ala Arg Ser Leu Ala Leu Ala
 35 40 45

Gly Ser Leu Ala Val Ala Leu Ser Tyr Ala Arg Ala Leu Pro Leu Val
 50 55 60

Gln Ala Asn Ala Leu Leu Asp Ala Thr Leu Cys Thr Gly Tyr Val Leu
 65 70 75 80

Leu Gln Gly Ile Val Phe Trp Gly Phe Phe Thr Val Gly His Asp Cys
 85 90 95

Gly His Gly Ala Phe Ser Arg Ser His Val Leu Asn Phe Ser Val Gly
 100 105 110

Thr Leu Met His Ser Ile Ile Leu Thr Pro Phe Glu Ser Trp Lys Leu
 115 120 125

Ser His Arg His His His Lys Asn Thr Gly Asn Ile Asp Lys Asp Glu
 130 135 140

Ile Phe Tyr Pro Gln Arg Glu Ala Asp Ser His Pro Val Ser Arg His
 145 150 155 160

Leu Val Met Ser Leu Gly Ser Ala Trp Phe Ala Tyr Leu Phe Ala Gly
 165 170 175

Phe Pro Pro Arg Thr Met Asn His Phe Asn Pro Trp Glu Ala Met Tyr
 180 185 190

Val Arg Arg Val Ala Ala Val Ile Ile Ser Leu Gly Val Leu Phe Ala
 195 200 205

Phe Ala Gly Leu Tyr Ser Tyr Leu Thr Phe Val Leu Gly Phe Thr Thr
 210 215 220

Met Ala Ile Tyr Tyr Phe Gly Pro Leu Phe Ile Phe Ala Thr Met Leu
 225 230 235 240

Val Val Thr Thr Phe Leu His His Asn Asp Glu Glu Thr Pro Trp Tyr
 245 250 255

Ala Asp Ser Glu Trp Thr Tyr Val Lys Gly Asn Leu Ser Ser Val Asp
 260 265 270

143/143

Arg Ser Tyr Gly Ala Leu Ile Asp Asn Leu Ser His Asn Ile Gly Thr
275 280 285

His Gln Ile His His Leu Phe Pro Ile Ile Pro His Tyr Lys Leu Asn
290 295 300

Asp Ala Thr Ala Ala Phe Ala Lys Ala Phe Pro Glu Leu Val Arg Lys
305 310 315 320

Asn Ala Ala Pro Ile Ile Pro Thr Phe Phe Arg Met Ala Ala Met Tyr
325 330 335

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

Glu Ala Lys Ala Ala Ala Lys Thr Lys Ser Ser
355 360