



- (12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11) CỤC SỞ HỮU TRÍ TUỆ
(51)⁸ A01H 5/00; C12Q 1/32; C12N 9/02; (13) B
C12N 15/82; C12N 5/04



1-0045425

-
- (21) 1-2019-00765 (22) 26/07/2017
(86) PCT/US2017/043990 26/07/2017 (87) WO2018/022777 01/02/2018
(30) 62/368,840 29/07/2016 US
(45) 26/05/2025 446 (43) 25/07/2019 376A
(71) MONSANTO TECHNOLOGY LLC (US)
800 North Lindbergh Boulevard, St. Louis, MO 63167, United States of America
(72) LARUE, Clayton, T. (US); REAM, Joel, E. (US); SHARIFF, Aabid (US); ZHANG,
Yuanji (US); ZHOU, Xuefeng (CN).
(74) Công ty Luật TNHH T&G (TGVN)
-

- (54) PHÂN TỬ ADN TÁI TỔ HỢP MÃ HÓA PROTEIN DUNG NẠP CHẤT DIỆT CỎ,
CÂY TRỒNG, HẠT, HOẶC TẾ BÀO CHUYỂN GEN CHỨA PHÂN TỬ NÀY,
PHƯƠNG PHÁP SẢN XUẤT CÂY TRỒNG DUNG NẠP CHẤT DIỆT CỎ VÀ
PHƯƠNG PHÁP KIỂM SOÁT SỰ SINH TRƯỞNG CỦA CỎ ĐẠI DUNG NẠP
CHẤT DIỆT CỎ

(21) 1-2019-00765

(57) Sáng chế đề cập đến phân tử ADN tái tổ hợp hữu dụng để mang lại sự biểu hiện hiệu quả của protein ở cây trồng chuyển gen, cũng như là hợp phần và phương pháp sử dụng phân tử ADN tái tổ hợp này. Theo các phương án cụ thể, sáng chế đề cập đến phân tử ADN tái tổ hợp và các cấu trúc có chứa các trình tự mã hóa peptit chuyển tiếp và được liên kết có điều khiển các trình tự mang lại tính dung nạp chất diệt cỏ.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập chung đến các lĩnh vực nông nghiệp, công nghệ sinh học thực vật, và sinh học phân tử. Cụ thể hơn, sáng chế đề cập đến hợp phần để biểu hiện protein tái tổ hợp ở cây trồng chuyển gen và phương pháp sử dụng chúng.

Tình trạng kỹ thuật của sáng chế

Sản xuất cây trồng nông nghiệp thường sử dụng cây trồng có hệ gen được cải biến, bao gồm các tính trạng chuyển gen được tạo ra bằng cách sử dụng các phương pháp sinh học phân tử. Ví dụ, gen khác loại, còn được biết đến dưới dạng gen chuyển, có thể được đưa vào trong hệ gen cây trồng. Sự biểu hiện của gen chuyển ở thực vật mang lại tính trạng, chẳng hạn như sự dung nạp chất diệt cỏ hoặc sự kiểm soát côn trùng, ở thực vật. Sự biểu hiện thành công của gen chuyển ở thực vật có thể đạt được bằng cách sử dụng các thành phần biểu hiện gen khác loại. Một ví dụ về nó là việc sử dụng peptit chuyển tiếp được liên kết có điều khiển với protein tái tổ hợp để đạt được sự định vị dưới tế bào của protein tái tổ hợp và do đó tăng cường sự biểu hiện hoặc chức năng protein. Do đó có nhu cầu về peptit chuyển tiếp mới có khả năng định vị một cách hiệu quả protein tái tổ hợp bên trong tế bào cây trồng.

Bản chất kỹ thuật của sáng chế

Theo một khía cạnh, sáng chế đề xuất phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên kết có điều khiển với trình tự ADN mã hóa protein dung nạp chất diệt cỏ khác loại, trong đó peptit chuyển tiếp có chứa trình tự axit amin có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266. Theo một phương án, protein dung nạp chất diệt cỏ khác loại có hoạt tính protoporphyrinogen oxidaza không nhạy cảm với chất diệt cỏ. Theo phương án khác, protein dung nạp chất diệt cỏ khác loại có chứa trình tự

axit amin có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:100-119, các SEQ ID NO:163-182, và các SEQ ID NO:224-228. Theo phương án khác, trình tự ADN mã hóa peptit chuyển tiếp có chứa trình tự axit nucleic có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:54-99 và các SEQ ID NO:267-297. Theo phương án khác nữa, trình tự ADN mã hóa protein dung nạp chất diệt cỏ khác loại có chứa trình tự axit nucleic có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:121-162 và các SEQ ID NO:183-223, các SEQ ID NO:229-235. Theo phương án khác nữa, phân tử ADN tái tổ hợp còn chứa vùng khởi đầu khác loại được liên kết có điều khiển với trình tự ADN mã hóa peptit chuyển tiếp.

Theo khía cạnh khác, sáng chế đề xuất cấu trúc ADN có chứa phân tử ADN được đề xuất trong bản mô tả này, chẳng hạn như phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên kết có điều khiển với trình tự ADN mã hóa protein dung nạp chất diệt cỏ khác loại, trong đó peptit chuyển tiếp có chứa trình tự axit amin có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266, được liên kết có điều khiển với vùng khởi đầu khác loại. Theo một phương án, protein dung nạp chất diệt cỏ khác loại có hoạt tính protoporphyrinogen oxidaza không nhạy cảm với chất diệt cỏ. Theo phương án khác, protein dung nạp chất diệt cỏ khác loại có chứa trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO:100-119, các SEQ ID NO:163-182, và các SEQ ID NO:224-228. Theo phương án khác nữa, cấu trúc ADN có mặt trong hệ gen của cây trồng, hạt, hoặc tế bào chuyển gen.

Theo khía cạnh khác, sáng chế đề xuất cây trồng, hạt, hoặc tế bào chuyển gen có chứa phân tử ADN tái tổ hợp được đề xuất trong bản mô tả này, chẳng hạn như phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên kết có điều khiển với trình tự ADN mã hóa protein dung nạp chất diệt cỏ khác loại, trong đó peptit chuyển tiếp có chứa trình tự axit amin có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266. Theo một phương án, cây trồng, hạt, hoặc tế bào này dung nạp đối với ít nhất một chất diệt cỏ PPO. Theo phương án khác, chất diệt cỏ PPO được chọn từ nhóm gồm có: acifluorfen, fomesafen, lactofen, floglycofen-ethyl, oxyfluorfen, flumioxazin, azafenidin,

carfentrazone-ethyl, sulfentrazone, fluthiacet-methyl, oxadiargyl, oxadiazon, pyraflufen-ethyl, saflufenacil, và S-3100. Theo phương án khác, cây trồng, hạt, hoặc tế bào chuyển gen dung nạp đối với ít nhất là chất diệt cỏ thứ hai.

Theo khía cạnh khác, sáng chế đề xuất protein tái tổ hợp có chứa trong sự liên kết có điều khiển: a) peptit chuyển tiếp có chứa trình tự axit amin có độ tương đồng bằng ít nhất là 95 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266; và b) protein dung nạp chất diệt cỏ khác loại. Theo một phương án, protein dung nạp chất diệt cỏ khác loại có hoạt tính protoporphyrinogen oxidaza không nhạy cảm với chất diệt cỏ. Theo khía cạnh khác, sáng chế đề xuất cây trồng, hạt, hoặc tế bào chuyển gen có chứa protein tái tổ hợp được đề xuất trong bản mô tả này.

Theo khía cạnh khác nữa, sáng chế đề xuất phương pháp sản xuất cây trồng dung nạp chất diệt cỏ có chứa các bước: a) biến nạp tế bào cây trồng với phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên kết có điều khiển với trình tự ADN mã hóa protein dung nạp chất diệt cỏ khác loại, trong đó peptit chuyển tiếp có chứa trình tự axit amin có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266; và b) tái tạo từ đó cây trồng dung nạp chất diệt cỏ mà có chứa phân tử ADN. Theo một phương án, protein dung nạp chất diệt cỏ khác loại có chứa trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO:100-119, các SEQ ID NO:163-182, và các SEQ ID NO:224-228. Theo phương án khác, phương pháp này còn chứa bước lai cây trồng đã được tái tạo với bản thân nó hoặc với cây trồng thứ hai để sản xuất một hoặc nhiều cây trồng thế hệ con. Theo phương án khác nữa, phương pháp này có thể còn chứa bước chọn lọc cây trồng thế hệ con mà dung nạp đối với ít nhất một chất diệt cỏ PPO. Theo các phương án nhất định, chất diệt cỏ PPO được chọn từ nhóm gồm có: acifluorfen, fomesafen, lactofen, floglycofen-ethyl, oxyfluorfen, flumioxazin, azafenidin, carfentrazone-ethyl, sulfentrazone, fluthiacet-methyl, oxadiargyl, oxadiazon, pyraflufen-ethyl, saflufenacil, và S-3100.

Theo khía cạnh khác, sáng chế đề xuất phương pháp sản xuất cây trồng hoặc hạt chuyển gen dung nạp với chất diệt cỏ có chứa bước lai cây trồng có chứa phân tử ADN tái tổ hợp được đề xuất trong bản mô tả này với bản thân nó hoặc cây trồng thứ hai để sản xuất cây trồng hoặc hạt chuyển gen dung nạp với chất diệt cỏ. Theo các phương án nhất định, phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên

kết có điều khiển với trình tự ADN mã hóa protein dung nạp chất diệt cỏ khác loại, trong đó peptit chuyển tiếp có chứa trình tự axit amin có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266.

Theo khía cạnh khác nữa, sáng chế đề xuất phương pháp biểu hiện protein dung nạp chất diệt cỏ khác loại ở cây trồng hoặc tế bào, phương pháp này có chứa bước trồng cây trồng hoặc tế bào mà có chứa phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên kết có điều khiển với trình tự ADN mã hóa protein dung nạp chất diệt cỏ khác loại, trong đó peptit chuyển tiếp có chứa trình tự axit amin có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266, trong đó việc trồng này dẫn đến sự biểu hiện của protein dung nạp chất diệt cỏ khác loại. Theo một phương án, protein dung nạp chất diệt cỏ khác loại có hoạt tính protoporphyrinogen oxidaza không nhạy cảm với chất diệt cỏ.

Theo khía cạnh khác, sáng chế đề xuất phương pháp kiểm soát hoặc ngăn chặn sự phát triển cỏ dại ở khu vực trồng cây có chứa bước dùng lượng hữu hiệu của ít nhất một chất diệt cỏ PPO cho khu vực trồng cây mà có chứa cây trồng hoặc hạt chuyển gen như được đề xuất trong bản mô tả này, chẳng hạn như cây trồng hoặc hạt chuyển gen có chứa phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên kết có điều khiển với trình tự ADN mã hóa protein dung nạp chất diệt cỏ khác loại, trong đó peptit chuyển tiếp có chứa trình tự axit amin có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266, trong đó cây trồng hoặc hạt chuyển gen dung nạp đối với chất diệt cỏ PPO. Theo các phương án nhất định, chất diệt cỏ PPO được chọn từ nhóm gồm có: acifluorfen, fomesafen, lactofen, floglycofen-ethyl, oxyfluorfen, flumioxazin, azafenidin, carfentrazone-ethyl, sulfentrazone, fluthiacet-methyl, oxadiargyl, oxadiazon, pyraflufen-ethyl, saflufenacil, và S-3100.

Theo khía cạnh khác, sáng chế đề xuất phương pháp kiểm soát sự sinh trưởng của cỏ dại dung nạp chất diệt cỏ có chứa bước: a) trồng ở khu vực trồng cây trồng hoặc hạt được đề xuất trong bản mô tả này, ví dụ cây trồng hoặc hạt có chứa phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên kết có điều khiển với trình tự ADN mã hóa protein dung nạp chất diệt cỏ khác loại, trong đó peptit chuyển tiếp có

chứa trình tự axit amin có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266; và b) dùng chất diệt cỏ PPO và ít nhất một chất diệt cỏ khác cho khu vực trồng cây, tại đó cây trồng hoặc hạt dung nạp đối với chất diệt cỏ PPO và ít nhất một chất diệt cỏ khác. Theo các phương án nhất định, chất diệt cỏ PPO được chọn từ nhóm gồm có acifluorfen, fomesafen, lactofen, floglycofen-etyl, oxyfluorfen, flumioxazin, azafenidin, carfentrazone-etyl, sulfentrazone, fluthiacet-metyl, oxadiargyl, oxadiazon, pyraflufen-etyl, saflufenacil, và S-3100. Theo phương án khác, chất diệt cỏ khác mà cây trồng hoặc hạt dung nạp với nó được chọn từ nhóm gồm có: chất ức chế ACCaza, chất ức chế ALS, chất ức chế EPSPS, auxin tổng hợp, chất ức chế quang hợp, chất ức chế glutamin synthetaza, chất ức chế HPPD, chất ức chế PPO, và chất ức chế axit béo mạch dài. Theo phương án khác, chất ức chế ACCaza là aryloxyphenoxy propionat hoặc xyclohexandion; chất ức chế ALS là sulfonylure, imidazolinon, triazolopyrimidin, hoặc triazolinon; chất ức chế EPSPS là glyphosat; auxin tổng hợp là chất diệt cỏ phenoxy, axit benzoic, axit carboxylic, hoặc semicarbazon; chất ức chế quang hợp là triazin, triazinon, nitril, benzothiadiazol, hoặc ure; chất ức chế glutamin synthetaza là glufosinat; chất ức chế HPPD là isoxazol, pyrazolon, hoặc triketon; chất ức chế PPO là diphenylete, N-phenylphthalimit, aryl triazinon, hoặc pyrimidinđion; hoặc chất ức chế axit béo mạch rất dài là cloaxetamit, oxyaxetamit, hoặc pyrazol.

Theo khía cạnh khác nữa, sáng chế đề xuất phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên kết có điều khiển với trình tự ADN mã hóa protein dung nạp chất diệt cỏ khác loại, trong đó peptit chuyển tiếp có chứa trình tự axit amin có độ tương đồng bằng ít nhất là 95 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:236-266. Theo một phương án, protein dung nạp chất diệt cỏ khác loại có hoạt tính protoporphyrinogen oxidaza không nhạy cảm với chất diệt cỏ. Theo phương án khác, protein dung nạp chất diệt cỏ khác loại có chứa trình tự axit amin có độ tương đồng bằng ít nhất là 95 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:100-119, các SEQ ID NO:163-182, và các SEQ ID NO:224-228. Theo phương án khác, trình tự ADN mã hóa peptit chuyển tiếp có chứa trình tự axit nucleic có độ tương đồng bằng ít nhất là 95 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:267-297. Theo phương án khác nữa, trình tự ADN mã hóa protein dung nạp chất diệt

có khác loại có chứa trình tự axit nucleic có độ tương đồng bằng ít nhất là 95 phần trăm với trình tự được chọn từ nhóm gồm có các SEQ ID NO:121-162 và các SEQ ID NO:183-223, các SEQ ID NO:229-235. Theo phương án khác nữa, phân tử ADN tái tổ hợp còn chứa vùng khởi đầu khác loại được liên kết có điều khiển với trình tự ADN mã hóa peptit chuyển tiếp.

Mô tả vắn tắt các trình tự

Các SEQ ID NO:1-2 và SEQ ID NO:236 là trình tự axit amin của peptit chuyển tiếp bạch tạng và có mắt màu xanh nhạt (APG6) *Arabidopsis thaliana*.

SEQ ID NO:3 là trình tự axit amin của peptit chuyển tiếp 12G088600TP ở cây bông.

Các SEQ ID NO:4-49 và các SEQ ID NO:237-266 là trình tự axit amin của peptit chuyển tiếp.

Các SEQ ID NO:50-52 và SEQ ID NO:267 là trình tự axit nucleic mã hóa peptit chuyển tiếp APG6.

SEQ ID NO:53 là trình tự axit nucleic mã hóa peptit chuyển tiếp 12G088600TP ở cây bông.

Các SEQ ID NO:54-99 và các SEQ ID NO:268-297 là các trình tự axit nucleic ví dụ lần lượt mã hóa các SEQ ID NO:4-49 và các SEQ ID NO:237-266.

Các SEQ ID NO:100-119 là trình tự axit amin của HemG protoporphyrinogen oxidaza.

SEQ ID NO:120 là trình tự axit amin của protoporphyrinogen oxidaza kiểu dại từ *Amaranthus tuberculatus* (cỏ waterhemp) (WH).

Các SEQ ID NO:121-162 và SEQ ID NO:229 là các trình tự axit nucleic ví dụ mã hóa các SEQ ID NO:100-119.

Các SEQ ID NO:163-182 và các SEQ ID NO:224-228 là trình tự axit amin của HemY protoporphyrinogen oxidaza.

Các SEQ ID NO:183-223 và các SEQ ID NO:230-235 là các trình tự axit nucleic ví dụ mã hóa các SEQ ID NO:163-182 và các SEQ ID NO:224-228.

Mô tả chi tiết sáng chế

Phần mô tả và các định nghĩa sau đây được nêu để mô tả sáng chế rõ hơn và hướng dẫn người có hiểu biết trung bình trong lĩnh vực thực hành sáng chế. Trừ khi được lưu ý khác, các thuật ngữ được hiểu theo cách sử dụng thông thường bởi người có hiểu biết trung bình trong lĩnh vực liên quan.

Việc liên kết có điều khiển peptit chuyển tiếp với protein khác loại sử dụng hệ thống định vị protein của tế bào cây trồng chuyển gen để đạt được sự định vị dưới tế bào của protein khác loại. Peptit chuyển tiếp được loại bỏ khỏi protein khác loại trong bước xử lý trong quá trình định vị của protein khác loại vào bào quan. Tính chất của dạng kết hợp của peptit chuyển tiếp cụ thể với protein khác loại khi được biểu hiện ở thực vật có thể không dự đoán được và bất ngờ. Ví dụ, hiệu quả của sự định vị dưới tế bào và hiệu quả của việc xử lý (việc loại bỏ peptit chuyển tiếp từ protein khác loại) thay đổi và có thể bị ảnh hưởng bởi trình tự axit amin của peptit chuyển tiếp, protein khác loại, hoặc cả hai. Các nhân tố biến đổi này ảnh hưởng đến chức năng và hàm lượng của protein khác loại và do đó ảnh hưởng đến kiểu hình của tế bào, cây trồng, hoặc hạt chuyển gen có chứa protein khác loại. Các peptit chuyển tiếp khác nhau được biết đến trong lĩnh vực để sử dụng ở cây trồng chuyển gen, nhưng do khả năng biến đổi về hiệu quả của sự định vị dưới tế bào và việc xử lý và sự phát triển liên tục của các tính trạng chuyển gen mới, nên cần có peptit chuyển tiếp mới.

Sáng chế đề xuất phân tử ADN tái tổ hợp, mới để hướng đích một cách hiệu quả protein khác loại bên trong tế bào cây trồng. Việc hướng đích hiệu quả của protein khác loại bao gồm sự định vị dưới tế bào hiệu quả của dạng kết hợp peptit chuyển tiếp và protein khác loại và việc xử lý của peptit chuyển tiếp từ protein khác loại. Mặc dù peptit chuyển tiếp để định vị protein khác loại bên trong tế bào được biết đến, mức độ của sự định vị và sự xử lý đối với dạng kết hợp peptit chuyển tiếp và protein khác loại bất kỳ thay đổi. Sự định vị và sự xử lý ảnh hưởng đến mức độ biểu hiện và chức năng của protein khác loại và do đó ảnh hưởng đến kiểu hình của tế bào, cây trồng, hoặc hạt có chứa protein khác loại. Ví dụ, sự định vị và sự xử lý không hiệu quả của dạng kết hợp peptit chuyển tiếp và protein dung nạp chất diệt cỏ có thể dẫn đến sự dung nạp chất diệt cỏ kém cho cây trồng chuyển gen.

Sáng chế khắc phục các trở ngại này bằng cách đề xuất phân tử ADN tái tổ hợp mới có khả năng tạo ra sự hướng đích hiệu quả của protein thông qua sự định vị và sự xử lý được cải thiện. Phân tử ADN tái tổ hợp theo sáng chế có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên kết có điều khiển với trình tự ADN mã hóa protein khác loại. Trong một ví dụ, phân tử ADN tái tổ hợp theo sáng chế bao gồm, nhưng không giới hạn ở, phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên kết có điều khiển với trình tự ADN mã hóa protoporphyrinogen oxidaza dung nạp với chất diệt cỏ. Hợp phần và phương pháp để sử dụng các phân tử ADN tái tổ hợp này cũng được đề xuất.

Phân tử tái tổ hợp

Như dùng trong bản mô tả này, thuật ngữ “tái tổ hợp” dùng để chỉ ADN, protein, tế bào, hạt, hoặc sinh vật không tự nhiên mà là kết quả của sự thiết kế di truyền và được tạo ra bởi sự can thiệp của con người. “Phân tử ADN tái tổ hợp” là phân tử ADN không có trong tự nhiên và như vậy là kết quả của sự can thiệp của con người, chẳng hạn như phân tử ADN tạo thành bởi dạng kết hợp của ít nhất là hai trình tự ADN khác loại đối với nhau. Ví dụ về phân tử ADN tái tổ hợp là phân tử ADN được đề xuất trong bản mô tả này mã hóa peptit chuyển tiếp theo sáng chế, chẳng hạn như peptit chuyển tiếp có chứa trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266, được liên kết có điều khiển với phân tử ADN mã hóa protein dung nạp chất diệt cỏ theo sáng chế, chẳng hạn như protoporphyrinogen oxidaza có chứa trình tự được chọn từ nhóm gồm có các SEQ ID NO:100-119, 163-182, và 224-228. “Protein tái tổ hợp” là protein được sản xuất là kết quả của sự can thiệp của con người mà không có trong tự nhiên. Ví dụ về protein tái tổ hợp là protein được đề xuất trong bản mô tả này có chứa peptit chuyển tiếp theo sáng chế, chẳng hạn như peptit chuyển tiếp có chứa trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266, được liên kết có điều khiển với protein khác loại, chẳng hạn như protein dung nạp chất diệt cỏ theo sáng chế, ví dụ, protoporphyrinogen oxidaza có chứa trình tự được chọn từ nhóm gồm có các SEQ ID NO:100-119, 163-182, và 224-228. Tế bào, hạt, hoặc sinh vật tái tổ hợp là tế bào, hạt, hoặc sinh vật có chứa ADN hoặc protein chuyển gen hoặc khác loại, ví dụ tế bào cây trồng, hạt, cây trồng, hoặc phần cây trồng chuyển gen có chứa phân tử ADN khác loại hoặc protein khác loại theo sáng chế.

Như dùng trong bản mô tả này, thuật ngữ “phân tử ADN được phân lập” có nghĩa là phân tử ADN có mặt một mình hoặc kết hợp với hợp phần khác nhưng không ở trong môi trường tự nhiên của nó. Phân tử ADN theo sáng chế là phân tử ADN được phân lập miễn là phân tử ADN không ở trong ADN của sinh vật ở vị trí trong hệ gen mà nó có mặt trong tự nhiên. Ví dụ, phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa protein và trình tự ADN peptit chuyển tiếp khác loại được coi là được phân lập khi nó được tìm thấy trong trường hợp mà không phải là hệ gen mà cả trình tự ADN mã hóa protein và trình tự ADN peptit chuyển tiếp khác loại được tìm thấy trong tự nhiên (chẳng hạn như hệ gen của cây trồng, hạt, phần cây trồng, hoặc tế bào chuyển gen).

Như dùng trong bản mô tả này, thuật ngữ “sự thiết kế di truyền” dùng để chỉ sự tạo ra, sự cải biến, hoặc sự sản xuất của phân tử ADN, protein, tế bào, hoặc sinh vật bằng cách sử dụng các kỹ thuật công nghệ sinh học (chẳng hạn như sinh học phân tử, hóa sinh protein, biến nạp vi khuẩn, và biến nạp thực vật). Do đó, sự thiết kế di truyền là kết quả của sự can thiệp của con người. Ví dụ, sự thiết kế di truyền có thể được sử dụng để tạo ra phân tử ADN tái tổ hợp mã hóa peptit chuyển tiếp có chứa trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49 và các SEQ ID NO:236-266 được liên kết có điều khiển với phân tử ADN mã hóa protein dung nạp chất diệt cỏ, chẳng hạn như protoporphyrinogen oxidaza có chứa trình tự được chọn từ nhóm gồm có các SEQ ID NO:100-119, 163-182, và 224-228 bằng cách sử dụng một hoặc nhiều kỹ thuật sinh học phân tử, chẳng hạn như tách dòng gen, nối ADN, và tổng hợp ADN. Phân tử ADN tái tổ hợp này tùy ý có thể còn chứa vùng khởi đầu khác loại có chức năng trong tế bào cây trồng.

Như dùng trong bản mô tả này, “sự dung nạp chất diệt cỏ” hoặc “dung nạp với chất diệt cỏ” đối với protein có nghĩa là khả năng để duy trì ít nhất là một số hoạt tính hoặc chức năng của nó trong sự có mặt của chất diệt cỏ. Ví dụ, protoporphyrinogen oxidaza (PPO) dung nạp với chất diệt cỏ nếu nó duy trì ít nhất là một số hoạt tính enzym của nó trong sự có mặt của một hoặc nhiều chất diệt cỏ PPO. Sự dung nạp chất diệt cỏ có thể đo được bằng phương thức bất kỳ đã biết trong lĩnh vực. Ví dụ, hoạt tính enzym của protoporphyrinogen oxidaza có thể đo được bằng thử nghiệm enzym trong đó sự sản xuất của sản phẩm của protoporphyrinogen oxidaza hoặc sự tiêu thụ của cơ chất của protoporphyrinogen oxidaza trong sự có mặt của một hoặc nhiều chất diệt cỏ PPO được

đo thông qua sự phát huỳnh quang, sắc ký lỏng hiệu năng cao (HPLC), hoặc đo phổ khối (MS). Ví dụ khác về thử nghiệm để đo hoạt tính enzym của protoporphyrinogen oxidaza là thử nghiệm vi khuẩn, chẳng hạn như các thử nghiệm sinh trưởng được mô tả trong bản mô tả này, bằng cách đó protoporphyrinogen oxidaza tái tổ hợp được biểu hiện ở tế bào vi khuẩn mất khác thiếu hoạt tính PPO và khả năng của protoporphyrinogen oxidaza tái tổ hợp để bù cho kiểu hình bất hoạt này được đo. Sự dung nạp chất diệt cỏ có thể là sự không nhạy cảm hoàn toàn hoặc một phần đối với chất diệt cỏ, và có thể được biểu diễn dưới dạng phần trăm (%) dung nạp hoặc không nhạy cảm đối với chất diệt cỏ PPO. Như dùng trong bản mô tả này, “dung nạp với chất diệt cỏ protoporphyrinogen oxidaza” thể hiện sự dung nạp chất diệt cỏ trong sự có mặt của một hoặc nhiều chất diệt cỏ PPO.

Như dùng trong bản mô tả này, “sự dung nạp chất diệt cỏ” hoặc “dung nạp với chất diệt cỏ” đối với sinh vật, cây trồng, hạt, mô, phần, hoặc tế bào có nghĩa là khả năng của sinh vật, cây trồng, hạt, mô, phần, hoặc tế bào để chống lại tác dụng của chất diệt cỏ khi được dùng. Ví dụ, cây trồng dung nạp chất diệt cỏ có thể sống hoặc tiếp tục phát triển trong sự có mặt của chất diệt cỏ. Sự dung nạp chất diệt cỏ của cây trồng, hạt, mô cây trồng, phần cây trồng, hoặc tế bào có thể được đo bằng cách so sánh cây trồng, hạt, mô cây trồng, phần cây trồng, hoặc tế bào với đối chứng thích hợp. Ví dụ, sự dung nạp chất diệt cỏ có thể được đo hoặc được đánh giá bằng cách dùng chất diệt cỏ cho cây trồng có chứa phân tử ADN tái tổ hợp mã hóa protein có khả năng mang lại sự dung nạp chất diệt cỏ (cây trồng thử nghiệm) và cây trồng không chứa phân tử ADN tái tổ hợp mã hóa protein có khả năng mang lại sự dung nạp chất diệt cỏ (cây trồng đối chứng) và sau đó so sánh tổn thương cây trồng của hai cây trồng này, trong đó sự dung nạp chất diệt cỏ của cây trồng thử nghiệm được chỉ ra bằng tỉ lệ tổn thương giảm khi so với tỉ lệ tổn thương của cây trồng đối chứng. Cây trồng, hạt, mô cây trồng, phần cây trồng, hoặc tế bào dung nạp chất diệt cỏ thể hiện đáp ứng giảm đối với tác dụng gây độc của chất diệt cỏ khi so với cây trồng đối chứng, hạt, mô cây trồng, phần cây trồng, hoặc tế bào. Như dùng trong bản mô tả này, “tính trạng dung nạp chất diệt cỏ” là tính trạng chuyển gen truyền sự dung nạp chất diệt cỏ được cải thiện cho cây trồng khi so với cây trồng kiểu dại. Cây trồng dự kiến mà có thể được sản xuất có tính trạng dung nạp chất diệt cỏ theo sáng chế có thể bao gồm, ví dụ, cây trồng bất kỳ bao gồm cây trồng mùa vụ chẳng hạn như cây đậu tương

(*Glycine max*), cây ngô (*Zea mays*), cây bông (*Gossypium sp.*), cây lúa mì (*Triticum spp.*), và cây cải *Brassica*, trong số những cây khác.

Như dùng trong bản mô tả này, “chúng bị bất hoạt *hemG*” có nghĩa là sinh vật hoặc tế bào của sinh vật, chẳng hạn như *E. coli*, mà thiếu hoạt tính HemG đến mức mà nó không thể sinh trưởng trên môi trường sinh trưởng không chứa hem, hoặc sao cho sự sinh trưởng của nó bị suy yếu đến mức có thể phát hiện được khi không có mặt hem so với chủng đẳng gen khác có chứa HemG chức năng. Chúng bị bất hoạt *hemG* của, ví dụ, *E. coli* có thể được điều chế theo hiểu biết trong lĩnh vực, ví dụ theo trình tự *hemG E. coli* (Mã Số Truy Cập Ecogene EG11485; Sasarman và cộng sự, “Nucleotide sequence of the *hemG* gene involved in the protoporphyrinogen oxidase activity of *Escherichia coli* K12” *Can J Microbiol* 39:1155-1161, 1993).

Thuật ngữ “gen chuyển” dùng để chỉ phân tử ADN được kết hợp theo cách nhân tạo vào hệ gen của sinh vật là kết quả của sự can thiệp của con người, chẳng hạn như bằng phương pháp biến nạp thực vật. Như dùng trong bản mô tả này, thuật ngữ “chuyển gen” có nghĩa là có chứa gen chuyển, ví dụ “cây trồng chuyển gen” dùng để chỉ cây trồng có chứa gen chuyển trong hệ gen của nó và “tính trạng chuyển gen” dùng để chỉ đặc điểm hoặc kiểu hình được truyền tải hoặc mang lại bởi sự có mặt của gen chuyển kết hợp trong hệ gen cây trồng. Kết quả của sự thay đổi hệ gen này là, cây trồng chuyển gen khác biệt ở mức độ nào đó với cây trồng kiểu đại liên quan và tính trạng chuyển gen là tính trạng không tìm thấy trong tự nhiên trong cây trồng kiểu đại. Cây trồng chuyển gen theo sáng chế có chứa phân tử ADN tái tổ hợp đề xuất bởi sáng chế.

Như dùng trong bản mô tả này, thuật ngữ “khác loại” dùng để chỉ mối quan hệ giữa hai hoặc nhiều thứ không thường kết hợp trong tự nhiên, ví dụ có nguồn gốc từ các nguồn khác nhau hoặc không thường được tìm thấy trong tự nhiên cùng nhau theo bất kỳ phương thức nào khác. Ví dụ, phân tử ADN hoặc protein có thể là khác loại đối với phân tử ADN, protein, tế bào, cây trồng, hạt, hoặc sinh vật khác nếu không thường được tìm thấy trong tự nhiên cùng nhau hoặc trong cùng hoàn cảnh. Theo các phương án nhất định, phân tử ADN thứ nhất là khác loại với phân tử ADN thứ hai nếu hai phân tử ADN này không thường được tìm thấy trong tự nhiên cùng nhau trong cùng hoàn cảnh, và protein là khác loại đối với protein được liên kết có điều khiển thứ hai, chẳng hạn như peptit chuyển tiếp, nếu dạng kết hợp này không thường được tìm thấy trong tự nhiên. Theo

phương án khác, phân tử ADN tái tổ hợp mã hóa peptit chuyển tiếp được liên kết có điều khiển với protoporphyrinogen oxidaza là khác loại đối với vùng khởi đầu được liên kết có điều khiển mà có chức năng trong tế bào thực vật nếu dạng kết hợp này không thường được tìm thấy trong tự nhiên. Phân tử ADN tái tổ hợp cũng có thể là khác loại đối với tế bào, hạt, hoặc sinh vật mà nó được cài xen vào đó khi trong tự nhiên nó không có ở trong tế bào, hạt, hoặc sinh vật đó. “Protein khác loại” là protein có mặt ở cây trồng, hạt, tế bào, mô, hoặc sinh vật mà trong tự nhiên nó không có mặt trong đó hoặc được liên kết có điều khiển với protein trong tự nhiên nó không được liên kết với. Ví dụ về protein khác loại là protein có chứa trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49, 236-266, 100-119, 163-182, và 224-228 mà được biểu hiện ở cây trồng, hạt, tế bào, mô, hoặc sinh vật mà trong tự nhiên nó không có mặt trong đó, hoặc được liên kết có điều khiển với protein thứ hai, chẳng hạn như peptit chuyển tiếp hoặc protein dung nạp với chất diệt cỏ, mà trong tự nhiên nó không được liên kết với. Trong ví dụ khác, protein khác loại, chẳng hạn như protein dung nạp chất diệt cỏ khác loại, ví dụ protoporphyrinogen oxidaza có thể được đưa vào trong tế bào cây trồng mà trong tự nhiên nó không có mặt trong đó bằng cách sử dụng kỹ thuật sinh học phân tử và biến nạp thực vật.

Như dùng trong bản mô tả này, thuật ngữ “phân tử ADN mã hóa protein” dùng để chỉ phân tử ADN có chứa trình tự ADN mà mã hóa protein. Như dùng trong bản mô tả này, “trình tự ADN mã hóa protein” có nghĩa là trình tự ADN mà mã hóa protein. Trình tự ADN mã hóa protein có thể là trình tự ADN bất kỳ mà mã hóa protein, ví dụ protein có chứa trình tự được chọn từ nhóm gồm có các SEQ ID NO:4-49, 236-266, 100-119, 163-182, và 224-228. Như dùng trong bản mô tả này, thuật ngữ “protein” dùng để chỉ chuỗi axit amin được liên kết bằng liên kết peptit (amit) và bao gồm cả hai chuỗi polypeptit mà được cuộn gấp hoặc được sắp xếp theo cách có chức năng sinh học và chuỗi polypeptit mà không được làm như vậy. “Trình tự” có nghĩa là sự sắp xếp tuần tự của nucleotit hoặc axit amin. Ranh giới của trình tự mã hóa protein thường được xác định bằng bộ ba mở đầu dịch mã ở đầu tận cùng 5' và bộ ba kết thúc dịch mã ở đầu tận cùng 3'.

Như dùng trong bản mô tả này, thuật ngữ “protein dung nạp chất diệt cỏ” có nghĩa là protein có khả năng mang lại sự dung nạp chất diệt cỏ cho tế bào, mô, phần cây trồng, hạt, hoặc sinh vật. Ví dụ về protein dung nạp chất diệt cỏ được biết rõ trong lĩnh vực và

bao gồm, nhưng không giới hạn ở, 5-enolpyruvyl shikimat 3-phosphat synthaza dung nạp glyphosat (ví dụ, CP4-EPSPS, 2mEPSPS), glyphosat oxidoreductaza (GOX), glyphosat N-axetyltransferaza (GAT), axetolactat synthaza (ALS) / axetohydroxyaxit synthaza (AHAS) dung nạp với chất diệt cỏ, 4-hydroxyphenylpyruvat dioxygenaza (HPPD) dung nạp với chất diệt cỏ, dicamba monooxygenaza (DMO), phosphinothrixin axetyl transferaza (PAT), glutamin synthetaza (GS) dung nạp với chất diệt cỏ, 2,4-diclophenoxypropionat dioxygenaza (TfdA), R-2,4-diclophenoxypropionat dioxygenaza (RdpA), S-2,4-diclophenoxypropionat dioxygenaza (SdpA), protoporphyrinogen oxidaza (PPO) dung nạp với chất diệt cỏ, và xytocrom P450 monooxygenaza. Ví dụ, protoporphyrinogen oxidaza có chứa trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO:100-119, các SEQ ID NO:163-182, và các SEQ ID NO:224-228 là protein dung nạp với chất diệt cỏ.

Như dùng trong bản mô tả này, “sự biểu hiện gen chuyển”, “biểu hiện gen chuyển”, “sự biểu hiện protein”, và “biểu hiện protein” có nghĩa là sự sản xuất của protein thông qua quy trình phiên mã phân tử ADN thành ARN thông tin (mARN) và dịch mã mARN thành chuỗi polypeptit, mà cuối cùng có thể được cuộn gấp hoặc không cuộn gấp thành protein. Phân tử ADN mã hóa protein có thể được liên kết có điều khiển với vùng khởi đầu khác loại trong cấu trúc ADN để sử dụng trong sự biểu hiện protein trong tế bào được biến nạp với, và do đó có chứa, phân tử ADN tái tổ hợp hoặc phần của chúng. Như dùng trong bản mô tả này, “được liên kết có điều khiển” có nghĩa là hai phân tử ADN hoặc protein được liên kết theo cách sao cho phân tử này có thể ảnh hưởng đến chức năng của phân tử kia. Các phân tử ADN được liên kết có điều khiển có thể là phần của phân tử tiếp giáp đơn lẻ và có thể liên kết hoặc không liên kết. Ví dụ, vùng khởi đầu được liên kết có điều khiển với phân tử ADN mã hóa protein trong cấu trúc ADN trong đó hai phân tử ADN này được sắp xếp sao cho vùng khởi đầu có thể ảnh hưởng đến sự biểu hiện của gen chuyển. Theo phương án khác, hai hoặc nhiều phân tử protein có thể được liên kết có điều khiển. Ví dụ, peptit chuyển tiếp có thể được liên kết có điều khiển với protein khác loại, chẳng hạn như protein dung nạp với chất diệt cỏ.

Theo một phương án, phân tử ADN tái tổ hợp theo sáng chế bao gồm trình tự ADN mã hóa protoporphyrinogen oxidaza (PPO) được liên kết có điều khiển với trình tự peptit chuyển tiếp. Như dùng trong bản mô tả này, “protoporphyrinogen oxidaza” hoặc

“PPO” có nghĩa là oxidaza có khả năng chuyển hóa protoporphyrinogen IX thành protoporphyrin IX. Protoporphyrinogen oxidaza này được biết đến trong lĩnh vực và bao gồm, ví dụ, trình tự protein được nêu dưới dạng các SEQ ID NO:100-119, các SEQ ID NO:163-182, và các SEQ ID NO:224-228.

Theo phương án khác, phân tử ADN tái tổ hợp theo sáng chế bao gồm trình tự ADN mã hóa trình tự peptit chuyển tiếp được liên kết có điều khiển với trình tự axit nucleic khác loại mã hóa protein mà có hoạt tính protoporphyrinogen oxidaza dung nạp với chất diệt cỏ, bằng cách đó trình tự peptit chuyển tiếp làm thuận lợi cho việc định vị phân tử protein ở trong tế bào. Peptit chuyển tiếp cũng được biết trong lĩnh vực dưới dạng trình tự tín hiệu, trình tự hướng đích, peptit hướng đích, và trình tự định vị. Ví dụ về peptit chuyển tiếp là peptit chuyển tiếp lục lạp (chloroplast transit peptide - CTP), trình tự hướng đích ty thể (mitochondrial targeting sequence - MTS), hoặc peptit hướng đích ty thể và lục lạp kép. Bằng cách làm thuận lợi cho sự định vị protein ở trong tế bào, chẳng hạn như vào ty thể hoặc lục lạp, peptit chuyển tiếp đảm bảo sự định vị của protein vào bào quan cho hoạt tính enzym tối ưu và có thể làm tăng sự tích lũy của protein và bảo vệ protein khỏi sự thoái hóa do phân giải protein, và/hoặc tăng cường mức độ của sự dung nạp chất diệt cỏ, và bằng cách đó làm giảm mức độ tổn thương trong tế bào, hạt, hoặc sinh vật chuyển gen sau khi dùng chất diệt cỏ. Khi chuyển vị vào bào quan, peptit chuyển tiếp thường được phân cắt khỏi protein, cũng được đề cập đến dưới dạng sự xử lý. Sự xử lý peptit chuyển tiếp có thể là hoàn toàn (có nghĩa là peptit chuyển tiếp hoàn chỉnh được phân cắt khỏi đầu tận cùng amino của protein), không hoàn toàn (có nghĩa là một hoặc nhiều axit amin của peptit chuyển tiếp vẫn còn ở trên đầu tận cùng amino của protein), hoặc dẫn đến việc loại bỏ một hoặc nhiều axit amin từ đầu tận cùng amino của protein. Sự xử lý hoàn toàn của peptit chuyển tiếp từ protoporphyrinogen oxidaza làm tăng mức độ của sự tích lũy protein, bằng cách đó làm tăng sự dung nạp chất diệt cỏ PPO và làm giảm mức độ tổn thương trong tế bào, hạt, hoặc sinh vật chuyển gen sau khi dùng chất diệt cỏ. Ví dụ, peptit chuyển tiếp có thể có chứa trình tự axit amin theo sáng chế, chẳng hạn như trình tự axit amin nêu bởi các SEQ ID NO:1-49 và các SEQ ID NO:236-266. Peptit chuyển tiếp này có thể được mã hóa bởi trình tự axit nucleic theo sáng chế, ví dụ như nêu bởi các SEQ ID NO:50-99 và các SEQ ID NO:267-297.

Phân tử ADN tái tổ hợp theo sáng chế có thể được tổng hợp và được cải biến bằng phương pháp đã biết trong lĩnh vực, hoàn toàn hoặc một phần, cụ thể là trong đó điều mong muốn là tạo ra các trình tự hữu dụng để thao tác ADN (chẳng hạn như vị trí nhận diện enzym giới hạn hoặc vị trí tách dòng dựa trên sự tái tổ hợp), các trình tự ưu tiên cây trồng (chẳng hạn như sử dụng bộ ba mã hóa cây trồng hoặc trình tự liên ứng Kozak), hoặc các trình tự hữu dụng để thiết kế cấu trúc ADN (chẳng hạn như các trình tự vùng đệm hoặc cầu nối). Sáng chế bao gồm phân tử ADN và protein có độ tương đồng trình tự bằng ít nhất là 90%, độ tương đồng trình tự bằng ít nhất là 91%, độ tương đồng trình tự bằng ít nhất là 92%, độ tương đồng trình tự bằng ít nhất là 93%, độ tương đồng trình tự bằng ít nhất là 94%, độ tương đồng trình tự bằng ít nhất là 95%, độ tương đồng trình tự bằng ít nhất là 96%, độ tương đồng trình tự bằng ít nhất là 97%, độ tương đồng trình tự bằng ít nhất là 98%, và độ tương đồng trình tự bằng ít nhất là 99% với trình tự phân tử ADN hoặc protein bất kỳ được đề xuất trong bản mô tả này dưới dạng các SEQ ID NO:1-297. Như dùng trong bản mô tả này, thuật ngữ “tỉ lệ phần trăm độ tương đồng trình tự” hoặc “% độ tương đồng trình tự” dùng để chỉ tỉ lệ phần trăm của các nucleotit hoặc axit amin giống nhau trong trình tự polynucleotit hoặc protein mạch thẳng của trình tự tham chiếu (“truy vấn”) (hoặc sợi bổ sung của nó) khi so với trình tự thử nghiệm (“chủ thể”) (hoặc sợi bổ sung của nó) khi hai trình tự này được sắp xếp thẳng hàng tối ưu (với tổng sự cài xen, sự làm khuyết, hoặc chỗ hở nucleotit hoặc axit amin thích hợp nhỏ hơn 20 phần trăm của trình tự tham chiếu trong cửa sổ so sánh). Sự sắp xếp thẳng hàng tối ưu của các trình tự để sắp xếp thẳng hàng của sổ so sánh được biết rõ đối với người có hiểu biết trung bình trong lĩnh vực và có thể được thực hiện bằng các công cụ chẳng hạn như thuật toán tương đồng cục bộ của Smith và Waterman, thuật toán sắp xếp thẳng hàng tương đồng của Needleman và Wunsch, phương pháp tra cứu độ tương tự của Pearson và Lipman, và bằng cách thực hiện vi tính hóa của các thuật toán này chẳng hạn như GAP, BESTFIT, FASTA, và TFASTA có sẵn dưới dạng một phần của gói phần mềm phân tích trình tự của GCG® Wisconsin Package® (Accelrys Inc., San Diego, CA), MEGAlign (DNASTar Inc., 1228 S. Park St., Madison, WI 53715), và MUSCLE (phiên bản 3.6) (Edgar, *Nucleic Acids Research* 32(5):1792-7, 2004) với các tham số mặc định. “Phân số độ tương đồng” đối với các đoạn được sắp xếp thẳng hàng của trình tự thử nghiệm và trình tự tham chiếu là số lượng của các thành phần giống nhau mà có chung giữa hai trình

tự được sắp xếp thẳng hàng chia cho tổng số thành phần trong đoạn trình tự tham chiếu, tức là, toàn bộ trình tự tham chiếu hoặc phần được xác định nhỏ hơn của trình tự tham chiếu. Tỷ lệ phần trăm độ tương đồng trình tự được thể hiện dưới dạng phân số độ tương đồng nhân với 100. Sự so sánh của một hoặc nhiều trình tự có thể là với trình tự chiều dài đầy đủ, hoặc phần của chúng, hoặc với trình tự dài hơn.

Như dùng trong bản mô tả này, “cấu trúc ADN” là phân tử ADN tái tổ hợp có chứa hai hoặc nhiều trình tự ADN khác loại. Các cấu trúc ADN hữu dụng để biểu hiện gen chuyển và có thể được chứa trong vật truyền và plasmit. Các cấu trúc ADN có thể được sử dụng trong vật truyền để biến nạp, là sự đưa ADN khác loại vào tế bào chủ, để sản xuất cây trồng và tế bào chuyển gen, và như vậy có thể cũng được chứa trong ADN plasmid hoặc ADN hệ gen của cây trồng, hạt, tế bào, hoặc phần cây trồng chuyển gen. Như dùng trong bản mô tả này, “vật truyền” có nghĩa là phân tử ADN tái tổ hợp bất kỳ mà có thể được sử dụng để biến nạp thực vật. Phân tử ADN như nêu trong danh mục trình tự, có thể, ví dụ, được cài xen vào vật truyền dưới dạng một phần của cấu trúc có phân tử ADN được liên kết có điều khiển với yếu tố biểu hiện gen mà có chức năng ở thực vật để ảnh hưởng đến sự biểu hiện của protein được mã hóa bởi phân tử ADN. Phương pháp tạo cấu trúc của cấu trúc ADN và vật truyền được biết rõ trong lĩnh vực. Các thành phần của cấu trúc ADN, hoặc vật truyền có chứa cấu trúc ADN, thường bao gồm một hoặc nhiều yếu tố biểu hiện gen được liên kết có điều khiển với trình tự ADN có thể phiên mã, chẳng hạn như các thành phần sau đây: vùng khởi đầu để biểu hiện ADN được liên kết có điều khiển, phân tử ADN mã hóa protein được liên kết có điều khiển, và vùng không dịch mã 3'. Yếu tố biểu hiện gen hữu dụng trong thực hành sáng chế bao gồm, nhưng không giới hạn ở, một hoặc nhiều loại yếu tố sau đây: vùng khởi đầu, vùng không dịch mã 5', vùng tăng cường, trình tự dẫn đầu, yếu tố hoạt hóa cis, intron, vùng không dịch mã 3', và một hoặc nhiều gen chuyển đánh dấu chọn lọc.

Cấu trúc ADN theo sáng chế có thể bao gồm vùng khởi đầu được liên kết có điều khiển với phân tử ADN mã hóa protein đề xuất bởi sáng chế, bằng cách đó vùng khởi đầu gây ra sự biểu hiện của phân tử protein khác loại. Các vùng khởi đầu hữu dụng trong thực hành sáng chế bao gồm các vùng khởi đầu mà có chức năng trong tế bào để biểu hiện polynucleotit được liên kết có điều khiển, chẳng hạn như vùng khởi đầu vi khuẩn hoặc thực vật. Các vùng khởi đầu thực vật thay đổi và đã biết rõ trong lĩnh vực và bao gồm các

vùng khởi đầu mà là cảm ứng, virus, tổng hợp, cấu tạo, được điều hòa về thời gian, được điều hòa về không gian, và/hoặc được điều hòa về không gian - thời gian.

Theo một phương án theo sáng chế, cấu trúc ADN được đề xuất trong bản mô tả này bao gồm trình tự ADN mã hóa peptit chuyển tiếp mà được liên kết có điều khiển với trình tự ADN khác loại mã hóa protein mà có hoạt tính protoporphyrinogen oxidaza dung nạp với chất diệt cỏ, bằng cách đó trình tự peptit chuyển tiếp làm thuận lợi cho việc định vị protein ở trong tế bào.

Như dùng trong bản mô tả này, “đối chứng” có nghĩa là đối chứng thí nghiệm được thiết kế nhằm mục đích so sánh. Ví dụ, cây trồng đối chứng trong phân tích cây trồng chuyển gen là cây trồng thuộc cùng loại như cây trồng thí nghiệm (tức là, cây trồng cần được thử nghiệm) nhưng không chứa phân tử ADN tái tổ hợp, cài xen chuyển gen, hoặc cấu trúc ADN của cây trồng thí nghiệm. Ví dụ về cây trồng đối chứng hữu dụng để so sánh với cây trồng chuyển gen bao gồm: đối với cây ngô, cây ngô LH244 không chuyển gen (số nộp lưu ATCC PTA-1173); để so sánh với cây đậu tương chuyển gen: cây đậu tương A3555 không chuyển gen (số nộp lưu ATCC PTA-10207); để so sánh với cây bông chuyển gen: Coker 130 không chuyển gen (Bảo Hộ Giống Cây Trồng (PVP) Số 8900252); để so sánh với cây cải canola hoặc cây *Brassica napus* chuyển gen: dòng Khôi phục giống *Brassica napus* không chuyển gen 65037 (Đơn đăng ký bảo hộ giống cây trồng Canada 06-5517); để so sánh với mầm nguyên sinh Samson cây lúa mì chuyển gen: giống lúa mì không chuyển gen (PVP 1994).

Như dùng trong bản mô tả này, “kiểu đại” có nghĩa là dạng tương tự, nhưng không giống hệt, có trong tự nhiên. “Phân tử ADN kiểu đại” hoặc “protein kiểu đại” là dạng có trong tự nhiên của phân tử ADN hoặc protein, tức là, dạng của phân tử ADN hoặc protein tồn tại từ trước trong tự nhiên. Ví dụ về protein kiểu đại hữu dụng để so sánh với protein được thiết kế đề xuất bởi sáng chế là protoporphyrinogen oxidaza từ *Arabidopsis thaliana*. “Cây trồng kiểu đại” là cây trồng không chuyển gen thuộc cùng loại như cây trồng chuyển gen, và như vậy khác biệt về mặt di truyền với cây trồng chuyển gen có chứa tính trạng dung nạp chất diệt cỏ. Ví dụ về cây trồng kiểu đại hữu dụng để so sánh bao gồm: đối với cây ngô chuyển gen, cây ngô LH244 không chuyển gen (số nộp lưu ATCC PTA-1173); để so sánh với cây đậu tương chuyển gen, cây đậu tương A3555 không chuyển gen (số nộp lưu ATCC PTA-10207); để so sánh với cây bông chuyển gen,

Coker 130 không chuyển gen (Bảo Hộ Giống Cây Trồng Số 8900252); để so sánh với cây cải canola hoặc cây *Brassica napus* chuyển gen, dòng Khôi phục giống *Brassica napus* không chuyển gen 65037 (Đơn đăng ký bảo hộ giống cây trồng Canada 06-5517); để so sánh với cây lúa mì chuyển gen, mầm nguyên sinh Samson giống lúa mì không chuyển gen (PVP 1994).

Cây trồng chuyển gen & chất diệt cỏ

Một khía cạnh theo sáng chế bao gồm tế bào cây trồng chuyển gen, mô cây trồng chuyển gen, cây trồng chuyển gen, và hạt chuyển gen mà có chứa phân tử ADN tái tổ hợp đề xuất bởi sáng chế. Các tế bào, mô, cây trồng, và hạt này có chứa phân tử ADN tái tổ hợp thể hiện sự dung nạp đối với một hoặc nhiều chất diệt cỏ PPO, và, tùy ý, dung nạp đối với một hoặc nhiều chất diệt cỏ khác.

Phương pháp thích hợp để biến nạp tế bào chủ thực vật để sử dụng bởi sáng chế bao gồm phương pháp gần như là bất kỳ mà bằng phương pháp này ADN có thể được đưa vào trong tế bào (ví dụ, trong đó cấu trúc ADN tái tổ hợp được tích hợp ổn định vào trong nhiễm sắc thể thực vật) và được biết rõ trong lĩnh vực. Phương pháp minh họa để đưa cấu trúc ADN tái tổ hợp vào cây trồng bao gồm hệ biến nạp *Agrobacterium* và bắn phá hạt ADN, cả hai đều được biết rõ đối với người có hiểu biết trung bình trong lĩnh vực. Phương pháp minh họa khác để đưa cấu trúc ADN tái tổ hợp vào cây trồng là cài xen cấu trúc ADN tái tổ hợp vào hệ gen cây trồng ở vị trí định trước bằng phương pháp tích hợp định hướng vị trí. Tích hợp định hướng vị trí có thể được thực hiện bằng phương pháp bất kỳ đã biết trong lĩnh vực, ví dụ, bằng cách sử dụng nucleaza ngón tay kềm, meganucleaza được thiết kế hoặc tự nhiên, TALE-endonucleaza, hoặc endonucleaza được dẫn hướng bằng ARN (ví dụ hệ thống CRISPR/Cas9). Cây trồng chuyển gen có thể được tái tạo từ tế bào cây trồng biến nạp bằng các phương pháp nuôi cấy tế bào cây trồng hoặc bằng cách lấy phần cắt từ cây trồng chuyển gen và cho ra rễ phần cắt để tạo thành dòng sinh dưỡng của cây trồng chuyển gen. Cây trồng chuyển gen đồng hợp tử đối với gen chuyển (tức là, hai bản sao alen của gen chuyển) có thể thu được bằng cách tự thụ phấn (tự thân) cây trồng chuyển gen mà chứa alen gen chuyển đơn lẻ với bản thân nó, ví dụ cây trồng R0, để sản xuất hạt R1. Một phần tư của số hạt R1 được sản xuất ra là đồng hợp tử đối với gen chuyển. Cây trồng mọc từ hạt R1 này mầm có thể được thử nghiệm về tình trạng tiếp hợp giao tử, thường là bằng cách sử dụng thử nghiệm SNP, xác định trình

tự ADN, hoặc thử nghiệm khuếch đại nhiệt mà cho phép phân biệt giữa thể dị hợp tử và thể đồng hợp tử, được đề cập đến dưới dạng thử nghiệm tình trạng tiếp hợp giao tử.

Như dùng trong bản mô tả này, “chất diệt cỏ” là phân tử bất kỳ mà được sử dụng để kiểm soát, ngăn chặn, hoặc gây trở ngại cho sự sinh trưởng của một hoặc nhiều cây trồng. Chất diệt cỏ ví dụ bao gồm chất ức chế axetyl-CoA carboxylaza (ACCase) (ví dụ aryloxyphenoxy propionat và xyclohexandion); chất ức chế axetolactat synthaza (ALS) (ví dụ sulfonylure, imidazolinon, triazolopyrimidin, và triazolinon); chất ức chế 5-enolpyruvylshikimat-3-phosphat synthaza (EPSPS) (ví dụ glyphosat), auxin tổng hợp (ví dụ phenoxy, axit benzoic, axit carboxylic, semicarbazon), chất ức chế quang hợp (hệ thống quang hóa II) (ví dụ triazin, triazinon, nitril, benzothiadiazol, và ure), chất ức chế glutamin synthetaza (GS) (ví dụ glufosinat và bialaphos), chất ức chế 4-hydroxyphenylpyruvat dioxygenaza (HPPD) (ví dụ isoxazol, pyrazolon, và triketon), chất ức chế protoporphyrinogen oxidaza (PPO) (ví dụ diphenylete, N-phenylphthalimit, aryl triazinon, và pyrimidindion), chất ức chế axit béo mạch rất dài (ví dụ cloaxetamit, oxyaxetamit, và pyrazol), chất ức chế sinh tổng hợp xenluloza (ví dụ indaziflam), chất ức chế hệ thống quang hóa I (ví dụ paraquat), chất ức chế lắp ráp vi ống (ví dụ pendimethalin), và chất ức chế phytoen desaturaza (PDS) (ví dụ norflurazon), trong số những chất khác.

Như dùng trong bản mô tả này, “chất diệt cỏ PPO” là chất hóa học mà hướng đích và ức chế hoạt tính enzym của protoporphyrinogen oxidaza (PPO), mà xúc tác cho sự khử hydro của protoporphyrinogen IX để tạo thành protoporphyrin IX, mà là tiền chất của hem và chất diệp lục. Sự ức chế của protoporphyrinogen oxidaza gây ra sự tạo thành của các loại oxy hoạt tính, dẫn đến sự phá vỡ màng tế bào và cuối cùng là cái chết của tế bào bị ảnh hưởng. Chất diệt cỏ PPO được biết rõ trong lĩnh vực và có sẵn trên thị trường. Ví dụ về chất diệt cỏ PPO bao gồm, nhưng không giới hạn ở, diphenylete (chẳng hạn như acifluorfen, muối và este của nó, aclonifen, bifenox, muối và este của nó, ethoxyfen, muối và este của nó, flonitrofen, furyloxyfen, halosafen, chlomethoxyfen, floglycofen, muối và este của nó, lactofen, muối và este của nó, oxyfluorfen, và fomesafen, muối và este của nó); thiadiazol (chẳng hạn như fluthiacet-metyl và thidiazimin); pyrimidindion hoặc phenyluraxin (chẳng hạn như benzfendizon, butafenacil, etyl [3-2-clo-4-flo-5-(1-metyl-6-triflometyl-2,4-di-oxo-1,2,3,4-tetrahydropyrimidin-3-yl)phenoxy]-2-

pyridyloxy]axetat (Số Đăng Ký CAS 353292-31-6 và được đề cập đến trong bản mô tả này dưới dạng S-3100), flupropacil, saflufenacil, và tiafenacil); phenylpyrazol (chẳng hạn như fluazolate, pyraflufen và pyraflufen-ethyl); oxadiazol (chẳng hạn như oxadiargyl và oxadiazon); triazolinon (chẳng hạn như azafenidin, bencarbazon, carfentrazon, muối và este của nó, và sulfentrazon); oxazolidinedion (chẳng hạn như pentoxazon); N-phenylphthalimit (chẳng hạn như cinidon-ethyl, flumiclorac, flumiclorac-pentyl, và flumioxazin); dẫn xuất benzoxazinon (chẳng hạn như 1,5-dimetyl-6-thioxo-3-(2,2,7-triflo-3,4-đihydro-3-oxo-4-prop-2-ynyl-2H-1,4-benzoxazin-6-yl)-1,3,5-triazinan-2,4-dion); flufenpyr và flufenpyr-ethyl; pyraclonil; và profluazol. Protoporphyrinogen oxidaza và tế bào, hạt, cây trồng, và phần cây trồng đề xuất bởi sáng chế thể hiện sự dung nạp chất diệt cỏ đối với một hoặc nhiều chất diệt cỏ PPO.

Cây trồng, hạt, phần cây trồng, mô cây trồng, và tế bào đề xuất bởi sáng chế thể hiện sự dung nạp chất diệt cỏ đối với một hoặc nhiều chất diệt cỏ PPO. (Các) chất diệt cỏ PPO có thể là được dùng cho khu vực trồng cây có chứa cây trồng và hạt đề xuất bởi sáng chế làm phương pháp kiểm soát cỏ dại. Cây trồng và hạt đề xuất bởi sáng chế có chứa tính trạng dung nạp chất diệt cỏ và như vậy dung nạp đối với việc dùng một hoặc nhiều chất diệt cỏ PPO. Việc dùng chất diệt cỏ có thể là ở tỉ lệ thương mại được khuyến nghị (1X) hoặc phân số hoặc bội số bất kỳ của chúng, chẳng hạn như hai lần tỉ lệ thương mại được khuyến nghị (2X). Tỉ lệ chất diệt cỏ có thể được biểu diễn dưới dạng gam cho mỗi hecta (g/h) hoặc pao cho mỗi mẫu Anh (pao/mẫu), đương lượng axit cho mỗi pao cho mỗi mẫu Anh (pao đương lượng axit/mẫu), đương lượng axit cho mỗi gam cho mỗi hecta (g đương lượng axit/ha), pao thành phần hoạt tính cho mỗi mẫu Anh (pao thành phần hoạt tính/mẫu), hoặc gam thành phần hoạt tính cho mỗi hecta (g thành phần hoạt tính/ha) tùy thuộc vào chất diệt cỏ và chế phẩm. Việc dùng chất diệt cỏ có chứa ít nhất một chất diệt cỏ PPO. Khu vực trồng cây có thể chứa hoặc không chứa cỏ dại vào lúc dùng chất diệt cỏ. Lượng hữu hiệu để diệt cỏ của chất diệt cỏ PPO để sử dụng trong khu vực để kiểm soát cỏ dại cần bao gồm khoảng giá trị từ khoảng 0,1X đến khoảng 30X tỉ lệ ghi trên nhãn trong mùa phát triển. 1X tỉ lệ ghi trên nhãn đối với một số chất diệt cỏ PPO ví dụ được nêu trong Bảng 1. Một (1) mẫu Anh tương đương với 2,47105 hecta và một (1) pao tương đương với 453,592 gam. Tỉ lệ chất diệt cỏ có thể được chuyển đổi giữa hệ Anh và hệ mét như sau: (pao thành phần hoạt tính/mẫu) nhân với 1,12 = (kg thành phần

hoạt tính/ha) và (kg thành phần hoạt tính/ha) nhân với 0,89 = (pao thành phần hoạt tính/mẫu).

Bảng 1: Chất diệt cỏ PPO minh họa

Chất diệt cỏ PPO	Họ hóa chất	Tỉ lệ 1X
acifluorfen	Điphenylete	420 g thành phần hoạt tính/ha
fomesafen	điphenylete	420 g thành phần hoạt tính/ha
lactofen	Điphenylete	70-220 g thành phần hoạt tính/ha
floglycofen-etyl	Điphenylete	15- 40 g thành phần hoạt tính/ha
oxyfluorfen	Điphenylete	0,28-2,24 kg thành phần hoạt tính/ha
flumioxazin	N-phenylphthalimit	70-105 g thành phần hoạt tính/ha
azafenidin	Triazolinon	240 g thành phần hoạt tính/ha
carfentrazon-etyl	Triazolinon	4-36 g thành phần hoạt tính/ha
sulfentrazon	Triazolinon	0,1-0,42 kg thành phần hoạt tính/ha
fluthiacet-metyl	Thiadiazol	3-15 g thành phần hoạt tính/ha
oxadiargyl	Oxadiazol	50-150 g thành phần hoạt tính/ha
oxadiazon	Oxadiazol	2,24-4,48 kg thành phần hoạt tính/ha

pyraflufen-ethyl	Phenylpyrazol	6-12 g thành phần hoạt tính/ha
saflufenacil	Pyrimidin đion	25-100 g/ha
S-3100	Pyrimidin đion	5-80 g/ha

Việc dùng chất diệt cỏ có thể là tuần tự hoặc trộn trong bể với một, hai, hoặc dạng kết hợp của một vài chất diệt cỏ hoặc chất diệt cỏ tương thích khác bất kỳ. Việc sử dụng nhiều lần của một chất diệt cỏ hoặc hai hoặc nhiều chất diệt cỏ, ở dạng kết hợp hoặc một mình, có thể được sử dụng trong mùa phát triển cho các khu vực có chứa cây trồng chuyển gen theo sáng chế để kiểm soát phổ rộng của cỏ dại hai lá mầm, cỏ dại một lá mầm, hoặc cả hai, ví dụ, việc sử dụng hai lần (chẳng hạn như sử dụng trước khi trồng và sử dụng sau khi nảy mầm hoặc sử dụng trước khi nảy mầm và sử dụng sau khi nảy mầm) hoặc sử dụng ba lần (chẳng hạn như sử dụng trước khi trồng, sử dụng trước khi nảy mầm, và sử dụng sau khi nảy mầm hoặc sử dụng trước khi nảy mầm và hai lần sử dụng sau khi nảy mầm).

Như dùng trong bản mô tả này, “cỏ dại” là cây không mong muốn bất kỳ. Cây trồng có thể được coi là nhìn chung là không mong muốn cho mục đích nông nghiệp hoặc nghề làm vườn (ví dụ, các loài *Amaranthus*) hoặc có thể được coi là không mong muốn trong tình huống cụ thể (ví dụ, cây trồng mùa vụ của một loài trên cánh đồng của các loài khác, còn được biết đến dưới dạng cây mọc tự nhiên).

Cây trồng chuyển gen, thế hệ con, hạt, tế bào cây trồng, và phần cây trồng theo sáng chế có thể cũng chứa một hoặc nhiều tính trạng khác. Tính trạng khác này có thể được đưa vào bằng cách lai cây trồng chứa gen chuyển có chứa phân tử ADN tái tổ hợp đề xuất bởi sáng chế với cây trồng khác chứa một hoặc nhiều tính trạng khác. Như dùng trong bản mô tả này, “lai” có nghĩa là gây giống hai cây trồng riêng lẻ để sản xuất cây trồng thế hệ con. Do đó, hai cây trồng có thể được lai để sản xuất thế hệ con mà chứa các tính trạng mong muốn từ mỗi bố mẹ. Như dùng trong bản mô tả này “thế hệ con” có nghĩa là con thuộc thế hệ bất kỳ của cây trồng bố mẹ, và thế hệ con chuyển gen có chứa cấu trúc ADN đề xuất bởi sáng chế và được di truyền từ ít nhất là một cây trồng bố mẹ. Tính trạng khác cũng có thể được đưa vào bằng cách đồng biến nạp cấu trúc ADN đối với (các) tính trạng chuyển gen khác đó với cấu trúc ADN có chứa phân tử ADN tái tổ

hợp đề xuất bởi sáng chế (ví dụ, với tất cả các cấu trúc ADN có mặt dưới dạng một phần của cùng vật truyền dùng để biến nạp cây trồng) hoặc bằng cách cài xen tính trạng khác vào cây trồng chuyển gen có chứa cấu trúc ADN đề xuất bởi sáng chế hoặc ngược lại (ví dụ, bằng cách sử dụng phương pháp biến nạp thực vật hoặc biên tập hệ gen bất kỳ trên cây trồng hoặc tế bào cây trồng chuyển gen). Tính trạng khác này bao gồm, nhưng không giới hạn ở, tính kháng côn trùng tăng, hiệu quả sử dụng nước tăng, năng suất sản lượng tăng, tính chịu hạn tăng, chất lượng hạt tăng, chất lượng dinh dưỡng tăng, sản xuất hạt lai, và dung nạp chất diệt cỏ, trong đó tính trạng này được đo đối với cây trồng kiểu dại. Tính trạng dung nạp chất diệt cỏ khác ví dụ có thể bao gồm sự dung nạp chuyển gen hoặc không chuyển gen đối với một hoặc nhiều chất diệt cỏ chẳng hạn như chất ức chế ACCaza (ví dụ aryloxyphenoxy propionat và xyclohexandion), chất ức chế ALS (ví dụ sulfonylure, imidazolinon, triazolopyrimidin, và triazolinon) chất ức chế EPSPS (ví dụ glyphosat), auxin tổng hợp (ví dụ phenoxy, axit benzoic, axit carboxylic, semicarbazon), chất ức chế quang hợp (ví dụ triazin, triazinon, nitril, benzothiadiazol, và ure), chất ức chế tổng hợp glutamin (ví dụ glufosinat), chất ức chế HPPD (ví dụ isoxazol, pyrazolon, và triketon), chất ức chế PPO (ví dụ diphenylete, N-phenylphthalimit, aryl triazinon, và pyrimidinđion), và chất ức chế axit béo mạch dài (ví dụ cloaxetamint, oxyaxetamit, và pyrazol), trong số những chất khác. Các tính trạng kháng côn trùng ví dụ có thể bao gồm tính kháng đối với một hoặc nhiều thành viên côn trùng trong một hoặc nhiều bộ Lepidoptera, Coleoptera, Hemiptera, Thysanoptera, Diptera, Hymenoptera, và Orthoptera, trong số những bộ khác. Tính trạng khác này được biết rõ đối với người có hiểu biết trung bình trong lĩnh vực; ví dụ, và danh mục của các tính trạng chuyển gen này được nêu bởi Cục Kiểm Dịch Động Vật và Thực Vật (Animal and Plant Health Inspection Service - APHIS) của Phòng Nông Nghiệp Hoa Kỳ (United States Department of Agriculture - USDA).

Tế bào được biến nạp với polynucleotit theo sáng chế, chẳng hạn như cấu trúc biểu hiện, có thể được chọn lọc về sự có mặt của polynucleotit hoặc hoạt tính enzym được mã hóa của nó trước hoặc sau khi tái tạo tế bào này thành cây trồng chuyển gen. Do đó cây trồng chuyển gen có chứa polynucleotit này có thể được chọn lọc ví dụ bằng cách xác định cây trồng chuyển gen mà có chứa polynucleotit hoặc hoạt tính enzym được mã

hóa, và/hoặc thể hiện tính trạng được thay đổi tương đối với cây trồng đối chứng đẳng gen khác. Tính trạng này có thể là, ví dụ, dung nạp đối với chất diệt cỏ PPO.

Cây trồng chuyển gen và thể hệ con mà chứa tính trạng chuyển gen đề xuất bởi sáng chế có thể được sử dụng bằng phương pháp gây giống bất kỳ mà thường là đã biết trong lĩnh vực. Trong các dòng cây trồng có chứa hai hoặc nhiều tính trạng chuyển gen, các tính trạng chuyển gen có thể là tách biệt một cách độc lập, được liên kết, hoặc dạng kết hợp của cả hai trong các dòng cây trồng có chứa ba hoặc hơn ba tính trạng chuyển gen. Việc lai ngược với cây trồng bố mẹ và lai xa với cây trồng không chuyển gen cũng được dự tính, như là nhân giống sinh dưỡng. Mô tả về phương pháp gây giống mà thường được sử dụng đối với các tính trạng và cây trồng khác nhau được biết rõ đối với người có hiểu biết trung bình trong lĩnh vực. Để xác nhận sự có mặt của (các) gen chuyển ở cây trồng hoặc hạt, nhiều thử nghiệm có thể được tiến hành. Các thử nghiệm này bao gồm, ví dụ, thử nghiệm sinh học phân tử, chẳng hạn như thẩm tách Southern và northern, PCR, và xác định trình tự ADN; thử nghiệm hóa sinh, chẳng hạn như phát hiện sự có mặt của sản phẩm protein, ví dụ, bằng phương pháp miễn dịch học (ELISA và thẩm tách western) hoặc bằng chức năng enzym; thử nghiệm phần cây trồng, chẳng hạn như thử nghiệm lá hoặc rễ; và, bằng cách phân tích kiểu hình của toàn bộ cây trồng. Để phân tích sự xử lý peptit chuyển tiếp ở cây trồng hoặc hạt chuyển gen, các thử nghiệm chẳng hạn như xác định trình tự theo phương pháp thoái hóa Edman hoặc phân tích đo phổ khối có thể được tiến hành trên protein protoporphyrinogen oxidaza khác loại thu được từ tế bào, cây trồng, hoặc hạt chuyển gen và dữ liệu trình tự thu được so với dữ liệu trình tự của protein protoporphyrinogen oxidaza.

Sự du nhập gen của tính trạng chuyển gen vào kiểu gen cây trồng đạt được là kết quả của quy trình biến đổi lai ngược. Kiểu gen cây trồng mà tính trạng chuyển gen đã được du nhập vào có thể được đề cập đến dưới dạng kiểu gen, dòng, con lai cùng dòng, hoặc thể lai được biến đổi bằng cách lai ngược. Tương tự, kiểu gen cây trồng thiếu tính trạng chuyển gen mong muốn có thể được đề cập đến dưới dạng kiểu gen, dòng, con lai cùng dòng, hoặc thể lai không được biến đổi.

Như dùng trong bản mô tả này, thuật ngữ “có chứa” có nghĩa là “bao gồm nhưng không giới hạn ở”.

Khi mô tả chi tiết sáng chế, rõ ràng là các biến thể, cải biến, và các phương án tương đương là có thể được mà không nằm ngoài phạm vi của sáng chế được xác định trong các yêu cầu bảo hộ kèm theo. Hơn nữa, cần hiểu rõ ràng các ví dụ trong sáng chế này được đề xuất dưới dạng các ví dụ không làm giới hạn sáng chế.

Ví dụ thực hiện sáng chế

Các ví dụ sau đây được bao gồm để chứng minh các phương án được ưu tiên theo sáng chế. Người có hiểu biết trung bình trong lĩnh vực nên hiểu rõ rằng các kỹ thuật được bộc lộ trong các ví dụ sau đây là theo các kỹ thuật tiêu biểu được phát hiện bởi các tác giả sáng chế để thực hiện tốt chức năng trong thực hành sáng chế, và do đó có thể được xem là thiết lập các phương thức ưu tiên để thực hành nó. Tuy nhiên, người có hiểu biết trung bình trong lĩnh vực nên, dựa trên bản mô tả sáng chế này, hiểu rõ rằng nhiều sự thay đổi có thể được thực hiện trong các phương án cụ thể mà được bộc lộ và vẫn đạt được cùng kết quả hoặc kết quả tương tự mà không nằm ngoài khái niệm, tinh thần và phạm vi của sáng chế. Cụ thể hơn là, cần hiểu rõ rằng các chất nhất định mà có liên quan cả về mặt hóa học và sinh lý học có thể được thay thế cho các chất được mô tả trong bản mô tả này để đạt được cùng kết quả hoặc kết quả tương tự. Tất cả các dạng thể và cải biến tương tự này rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực được cho là nằm trong tinh thần, phạm vi, và ý tưởng của sáng chế như được xác định bởi các yêu cầu bảo hộ kèm theo.

Ví dụ 1: Phát hiện peptit chuyển tiếp

Peptit chuyển tiếp mới được khai thác từ bộ sưu tập cơ sở dữ liệu trình tự cây trồng. Phương pháp và công cụ thông tin sinh học, chẳng hạn như mô hình Markov ẩn (hidden Markov model - HMM), cơ sở dữ liệu Pfam, và công cụ tra cứu sắp xếp thẳng hàng cục bộ cơ bản (basic local alignment search tool - BLAST), được sử dụng để xác định hàng nghìn EST và trình tự hệ gen được dự đoán để mã hóa protein đã biết để được định vị vào lục lạp và ty thể trong tế bào cây trồng, chẳng hạn như protoporphyrinogen oxidaza và protein sốc nhiệt. Sau đó, các trình tự này được phân tích, và trình tự mã hóa peptit chuyển tiếp được xác định. Hàng nghìn trình tự peptit chuyển tiếp giả định được xác định và được đánh giá về hiệu quả được dự đoán và sự đa dạng trình tự so sánh. Từ

đó, 60 peptit chuyển tiếp độc nhất được chọn để tách dòng và thử nghiệm trong tế bào cây trồng, với các biến thể được sản xuất ra cho một số trong số chúng (được chỉ ra dưới dạng “_var” trong bản mô tả này). Bảng 2 cung cấp SEQ ID NO tương ứng với các trình tự protein và nucleotit của mỗi peptit chuyển tiếp và biến thể của chúng.

Phân tử ADN tái tổ hợp mã hóa peptit chuyển tiếp được tổng hợp bằng cách sử dụng trình tự đối với mỗi peptit chuyển tiếp được dự đoán. Cấu trúc ADN được sản xuất ra liên kết có điều khiển mỗi peptit chuyển tiếp với vùng khởi đầu và trình tự mã hóa protein. Sau đó, các cấu trúc ADN này được sử dụng để biến nạp tế bào trần của cây trồng. Thử nghiệm tế bào trần được sử dụng với tế bào trần của cây trồng đã biến nạp để thử nghiệm peptit chuyển tiếp về hoạt tính chức năng của protein dung nạp chất diệt cỏ được liên kết có điều khiển trong sự có mặt của chất diệt cỏ. Sau đó các ứng viên thành công được cải tiến để biến nạp cây trồng để có thể thử nghiệm cây trồng chuyển gen.

Bảng 2. Peptit chuyển tiếp

Peptit chuyển tiếp	SEQ ID NO CỦA PRT	SEQ ID NO CỦA ADN
ADADI_1600	8	58
ALLCE_3035	16, 37, 46, 47, 237	66, 87, 96, 97, 268
AMACR_2643	33	83
AMAGR_5230	29	79
AMAPA_1826	12	62
AMAPA_4787	18	68
AMBTR_1537	30	80
ANDGE_6461	26	76
BRANA_6036	31	81
BRANA_9788	7	57
CAMSA_6215	21, 41	71, 91
CANRO_3271	24	74
CANRO_3976	4, 35	54, 85
CONCA_4103	11	61
CUCME_4756	22, 39, 48	72, 89, 98

DIGSA_5107	17	67
DIGSA_5109	27	77
ERATE_2090	25	75
ERATE_4149	23, 36, 45	73, 86, 95
ERATE_4824	28	78
KOCSC_1672	14	64
NICBE_5162	6	56
ROSHY_6783	32	82
ROSHY_8873	9	59
SEDAL_8241	20	70
SENOB_8832	5, 44	55, 94
SETIT_2080	15	65
SPIOL_0401	19	69
SPIOL_0410	13	63
TAROF_2111	34, 42, 38, 43	84, 92, 88, 93
XANST_27	10, 40, 49	60, 90, 99
ERATE_3481	28	29
SETIT_9796	239	270
ACAOS_3432	240	271
ADADI_0544	241	272
TAROF_9570	242	273
AMACR_2380	243	274
AMACR_2381	244	275
AMAHY_5254	245	276
AMAPA_22810	246	277
AMAPA_2811	247	278
AMAPA_6265_1	248	279
AMAPA_6265_2	249	280
AMAPA_2906	250	281
AMARU_1762	251	282

AMARU_1763	252	283
AMARU_1764	253	284
AMAVI_1826	254	285
AMAVI_1827	255	286
AMBTR_6334	256	287
CONCA_3910	257	288
CUCME_3420	258	289
KOCSC_5431	259	290
KOCSC_9516	260	291
KOCSC_0438	261	292
ROSHY_3269	262	293
SEDAL_6599	263	294
SEDAL_6601	264	295
SPIOL_1551	265	296
ALLCE_6618	266	297

Ví dụ 2: Phát hiện enzym PPO

Các protoporphyrinogen oxidaza HemG và HemY của vi sinh vật mới mà dung nạp đối với chất diệt cỏ PPO được xác định từ cơ sở dữ liệu trình tự vi sinh vật bằng cách sử dụng phương pháp thông tin sinh học và hệ thống sàng lọc vi khuẩn chất diệt cỏ mới. Hệ thống sàng lọc này sử dụng thử nghiệm sinh trưởng của chủng *E. coli* bị bất hoạt *hemG* trong môi trường lỏng LB với chất diệt cỏ PPO để xác nhận hoạt tính protoporphyrinogen oxidaza đối với enzym và để xác định protoporphyrinogen oxidaza mà không nhạy đối với chất diệt cỏ PPO. Một cách vắn tắt, chủng *E. coli* bị bất hoạt *hemG* được biến nạp với vật truyền biểu hiện vi khuẩn chứa protoporphyrinogen oxidaza giả định và được nuôi cấy trong môi trường lỏng LB. Dạng kết tinh đã tinh chế của một trong năm chất diệt cỏ PPO khác nhau (acifluorfen (1 mM), flumioxazin (0,5 mM), lactofen (0,5 mM), fomesafen (1 mM), và S-3100 (100 microM), thể hiện ba phân lớp hóa học PPO khác nhau, được bổ sung vào môi trường. Protein tái tổ hợp được biểu hiện và tốc độ sinh trưởng *E. coli* được đo. Đường cong sinh trưởng (OD600) được đo đối với các biến thể khác nhau trong sự có mặt và vắng mặt của chất diệt cỏ PPO ở các thời điểm

được chọn từ thời điểm không đến hai mươi tư giờ. Sự sinh trưởng của chủng *E. coli* bị bất hoạt *hemG* được biến nạp trong môi trường LB trong sự có mặt của chất diệt cỏ PPO chỉ ra rằng gen dùng để biến nạp *E. coli* mã hóa tính dung nạp với chất diệt cỏ protoporphyrinogen oxidaza. Chủng *E. coli* bị bất hoạt *hemG* biểu hiện protoporphyrinogen oxidaza của cỏ waterhemp (WH) (SEQ ID NO:120), mà nhạy cảm đối với cả năm chất diệt cỏ PPO, được sử dụng làm đối chứng để xác nhận rằng thử nghiệm này có thể phân biệt giữa protoporphyrinogen oxidaza nhạy cảm và dung nạp đối với mỗi chất diệt cỏ.

Protoporphyrinogen oxidaza mà là protein dung nạp với chất diệt cỏ được nêu dưới dạng các SEQ ID NO:100-119, các SEQ ID NO:163-182, và các SEQ ID NO:224-228 và được thể hiện trong Bảng 3. Trình tự ADN mã hóa protoporphyrinogen oxidaza có thể bao gồm ở đầu 5' bộ ba mã hóa methionin, thường gọi là bộ ba mở đầu, hoặc bộ ba mã hóa này (và tùy ý một vài axit amin ở đầu tận cùng amino, ví dụ từ 2 đến 7), có thể được loại trừ để làm thuận lợi cho sự liên kết có điều khiển của trình tự peptit chuyển tiếp với đầu 5' của trình tự mã hóa. Các trình tự ADN mã hóa protoporphyrinogen oxidaza tùy ý có thể được tổng hợp mà được tối ưu hóa để biểu hiện ở cây một lá mầm hoặc cây hai lá mầm. Bảng 3 cung cấp đối với mỗi trình tự ADN protoporphyrinogen oxidaza mà được tối ưu hóa để biểu hiện ở cây một lá mầm và cây hai lá mầm.

Bảng 3. Protoporphyrinogen oxidaza

Tên	Protein SEQ ID NO	SEQ ID NO của ADN vi khuẩn	SEQ ID NO của ADN được tối ưu hóa ở cây hai lá mầm	SEQ ID NO của ADN được tối ưu hóa ở cây một lá mầm
H_N10	103, 112	124	134, 143	156
H_N20	101, 111	122	132, 142, 151	154
H_N30	104, 113	125	135, 144	157
H_N40	105, 114	126	136, 145	158
H_N50	106, 115	127	137, 146	159
H_N60	102	123	133	155
H_N70	107	128	138	160

H_N90	100, 110, 117, 118	121	131, 141, 148, 149, 150, 229	153
H_N100	108, 116, 119	129	139, 147, 152	161
H_N110	109	130	140	162
WH PPO	120	không có	không có	không có
R2N30	163, 164	183	189, 190	195
R2N40	165, 224	184	191, 230	196
R2N40opt	166, 225	185	231, 232	không có
R2N70	167, 226	186	192, 233	197
R2N90	168, 227	187	193, 234	198
R2N100	169, 228	188	194, 235	199
R1N473	170, 175, 179	200	205, 216, 220	211
R1N533	171, 176, 180	201	206, 217, 221	212
R1N171	172, 177, 181	202	207, 218, 222	213
R1N311	173	203	208	214
R1N333	174, 178, 182	204	209, 210, 219, 223	215

Ví dụ 3: Peptit chuyển tiếp và protoporphyrinogen oxidaza thử nghiệm ở tế bào trần

Peptit chuyển tiếp được liên kết có điều khiển với protoporphyrinogen oxidaza được thử nghiệm ở tế bào trần của cây trồng về sự dung nạp chất diệt cỏ PPO. Vật truyền biến nạp thực vật được tạo cấu trúc có chứa phân tử ADN tái tổ hợp mã hóa protoporphyrinogen oxidaza H_N90 được liên kết có điều khiển với peptit chuyển tiếp. Sau đó, vật truyền này được sử dụng để biến nạp tế bào trần của cây trồng, mà được đánh giá về tính nhạy cảm đối với chất diệt cỏ PPO.

Vật truyền biến nạp thực vật được sản xuất ra có chứa (i) các thành phần biểu hiện cố định (vùng khởi đầu và UTR 3') được liên kết có điều khiển với peptit chuyển tiếp được liên kết có điều khiển với protoporphyrinogen oxidaza H_N90. Bằng cách sử dụng nó, 68 peptit chuyển tiếp được thử nghiệm và việc so sánh trực tiếp được thực hiện bằng cách sử dụng cùng protoporphyrinogen oxidaza và các thành phần biểu hiện khác trong

mỗi vật truyền. Vật truyền đối chứng với cùng thành phần biểu hiện cố định được sản xuất ra có chứa (i) protoporphyrinogen oxidaza H_N90 mà không có bất kỳ peptit chuyển tiếp nào (Đối chứng H_N90) hoặc (ii) Protein Huỳnh Quang Xanh (Green Fluorescent Protein - GFP) mà không có peptit chuyển tiếp (Đối chứng GFP).

Biến nạp tế bào trần đậu tương bằng cách sử dụng phương pháp tiêu chuẩn và trồng trong sự có mặt của chất diệt cỏ PPO S-3100 ở nồng độ 1,0 microM. Sau đó, phân tích tế bào trần về tính dung nạp chất diệt cỏ PPO, được biểu hiện tương đối với Đối chứng GFP (cho phép độ lệch của điểm dung nạp tương đối để có thể so sánh giữa các thí nghiệm). Thực hiện các thử nghiệm trong hai mẻ, được chỉ ra dưới dạng Thí nghiệm số 1 hoặc Thí nghiệm số 2. Thực hiện các thử nghiệm này trong bốn bản sao, tính trung bình các điểm dung nạp tương đối đối với mỗi peptit chuyển tiếp, và tính sai số chuẩn (SE). Peptit hướng đích bất kỳ có điểm dung nạp tương đối bằng 50 hoặc cao hơn được coi là hữu hiệu cao để mang lại sự định vị dưới tế bào và sự xử lý hiệu quả khi được liên kết có điều khiển với protein dung nạp chất diệt cỏ và điểm bằng 40-50 chỉ ra là rất tốt để tạo ra sự định vị dưới tế bào và sự xử lý hiệu quả khi được liên kết có điều khiển với protein dung nạp chất diệt cỏ. Thử nghiệm Đối chứng GFP có điểm dung nạp bằng 0, xác nhận rằng tế bào trần đậu tương không dung nạp đối với chất diệt cỏ PPO khi không có mặt protein dung nạp chất diệt cỏ. Thử nghiệm Đối chứng H_N90 có điểm dung nạp bằng 24 (Thí nghiệm 1, SE 4) và 11 (Thí nghiệm 2, SE 4), trong khi một vài peptit chuyển tiếp mang lại điểm dung nạp cao hơn, chỉ ra rằng peptit chuyển tiếp hiệu quả có thể làm tăng tính dung nạp chất diệt cỏ của tế bào trần của cây trồng. Ví dụ, ADADI_0544 và KOCSC_9516 được tính điểm là peptit hướng đích hữu hiệu cao và AMAPA_62652 được tính điểm là peptit hướng đích rất tốt. Dữ liệu được nêu trong Bảng 4.

Bảng 4. Kết quả thử nghiệm tế bào trần

Peptit chuyển tiếp	Điểm dung nạp	SE	Thí nghiệm
ADADI_0544	62	2	1
KOCSC_9516	60	1	1
ALLCE_3035_var	56	4	1
CAMSA_6215	56	3	1
AMAPA_2810	56	3	1

ALLCE_6618	56	2	1
AMARU_1764	56	3	1
AMBTR_6334	56	1	1
SETIT_9796	55	5	1
AMACR_2381	55	2	1
AMAVI_1827	54	4	1
CONCA_3910	54	1	1
ERATE_3481	53	2	1
ROSHY_3269	53	5	1
AMAPA_6265_1	53	2	1
AMAHY_5254	52	4	1
SEDAL_6599	52	2	1
AMACR_2380	51	3	1
CUCME_3420	51	3	1
AMARU_1762	51	5	1
SEDAL_6601	50	5	1
KOCSC_5431	48	4	1
AMAPA_6265_2	47	2	1
KOCSC_0438	47	3	1
AMAPA_2811	46	3	1
AMAVI_1826	45	4	1
ACAOS_3432	44	2	1
SPIOL_1551	43	4	1
AMAPA_2906	43	2	1
TAROF_9570	41	3	1
AMARU_1763	40	8	1
Đối chứng không H_N90	24	4	1
Không - GFP	0	4	1
ADADI_0544	60	1	2

SPIOL_1551	53	3	2
KOCSC_9516	51	4	2
ROSHY_3269	49	4	2
AMACR_2381	48	3	2
CAMSA_6215	46	2	2
CUCME_4756_var	46	1	2
CUCME_3420	46	3	2
CONCA_3910	45	4	2
AMAGR_5230	43	2	2
SENOB_8832	43	1	2
KOCSC_1672	42	3	2
CONCA_4103	36	5	2
ADADI_1600	36	4	2
BRANA_9788	33	1	2
CUCME_4756	33	4	2
ANDGE_6461	33	2	2
ALLCE_3035	33	3	2
AMAPA_4787	30	2	2
TAROF_2111	28	3	2
ROSHY_6783	26	4	2
CANRO_3976	25	4	2
TAROF_2111_var	25	5	2
XANST_27_var	24	2	2
NICBE_5162	24	3	2
XANST_27	22	3	2
SPIOL_0401	22	2	2
ERATE_2090	22	1	2
SPIOL_0410	21	2	2
CANRO_3271	20	2	2
AMAPA_1826	20	2	2

DIGSA_5109	20	2	2
DIGSA_5107	17	2	2
ERATE_4149	15	4	2
SETIT_2080	14	2	2
ROSHY_8873	12	4	2
AMBTR_1537	12	6	2
SEDAL_8241	11	6	2
Đối chứng không H_N90	11	4	2
ERATE_4824	9	5	2
ALLCE_3035_var	8	1	2
Không - GFP	0	4	2
AMACR_2643	0	4	2

Ví dụ 4: Peptit chuyển tiếp và protoporphyrinogen oxidaza thử nghiệm ở cây đậu tương

Peptit chuyển tiếp được liên kết có điều khiển với protoporphyrinogen oxidaza được thử nghiệm ở cây đậu tương chuyển gen về sự dung nạp chất diệt cỏ PPO. Vật truyền biến nạp thực vật được tạo cấu trúc có chứa cấu trúc ADN có chứa phân tử ADN tái tổ hợp được tối ưu hóa để biểu hiện ở cây hai lá mầm và mã hóa protoporphyrinogen oxidaza được liên kết có điều khiển với peptit chuyển tiếp. Sau đó, sử dụng vật truyền biến nạp thực vật này để biến nạp cây đậu tương, và tái tạo cây trồng và đánh giá về tính nhạy cảm của nó đối với chất diệt cỏ PPO.

Các gen mã hóa bảy HemG protoporphyrinogen oxidaza H_N10, H_N20, H_N30, H_N40, H_N50, H_N90, và H_N100 được liên kết có điều khiển với ba mươi bảy peptit chuyển tiếp khác nhau và được tách dòng vào vật truyền biến nạp thực vật cơ bản như được mô tả trong Ví dụ 3. Điều này cho phép so sánh cạnh nhau bảy HemG protoporphyrinogen oxidaza khác nhau với ba mươi bảy peptit chuyển tiếp khác nhau bằng cách sử dụng cùng các thành phần vùng khởi đầu và UTR 3' trong mọi cấu trúc ADN. Sử dụng các vật truyền biến nạp thực vật này để biến nạp phôi cắt của đậu tương (mầm nguyên sinh A3555) bằng cách sử dụng *A. tumefaciens* và phương pháp tiêu chuẩn

đã biết trong lĩnh vực. Cây ghép bốn trăm mô cây bên ngoài đối với mỗi cấu trúc. Sử dụng dung dịch chất diệt cỏ PPO tiết trùng để thử nghiệm sự dung nạp chất diệt cỏ. Dung dịch chất diệt cỏ gồm có 0,3 g S-3100 trong chất cô dầu của cây trồng (5,0 ml) và 495 ml nước khử ion.

Vào năm tuần sau biến nạp, phun cây trồng bằng hai lần phun dung dịch chất diệt cỏ PPO tiết trùng ở tỉ lệ 20g/ha. Đối với mỗi cấu trúc ADN được thử nghiệm, bốn vật chứa trong đó mỗi vật chứa có 30-40 cây trồng được biến nạp riêng lẻ được thử nghiệm. Sau đó các cây con đã được xử lý nhận được ít nhất là 15 giờ phơi sáng sau khi phun mỗi ngày trong bốn ngày. Khi kết thúc ngày bốn sau khi dùng S-3100, các cây con đã được xử lý được chụp ảnh và tính điểm trên thang trực quan của màu xanh (màu xanh đại diện cho mô cây trồng quang hợp khỏe mạnh khi so với mô bị tẩy trắng quang) so với sự hư hại. Giá trị điểm bằng 0 đối với sự dung nạp kém, sự hư hại cao, màu xanh ít; 1 đối với sự dung nạp một chút, sự hư hại trung bình, màu xanh vừa phải; và 2 đối với sự dung nạp tốt, sự hư hại ít, màu xanh nhiều. Sự tính điểm đối với mỗi cấu trúc được thể hiện trong Bảng 5, trong đó n.d. chỉ phân tích không được thực hiện. Kết quả chỉ ra rằng một vài cấu trúc mang lại sự dung nạp đối với chất diệt cỏ PPO.

Bảng 5. Điểm dung nạp vào 5 tuần ở cây đậu tương

Peptit chuyển tiếp	H_N 10	H_N 20	H_N 30	H_N 40	H_N5 0	H_N 90	H_N 100
APG6	không thực hiện (n.d.)	0	2	2	1	2	2
12G088600TP	không thực hiện	0	0	1	1	2	1
CANRO_3976	1	1	không thực hiện	1	1	2	1

SENOB_8832	không thực hiện	1	không thực hiện	2	1	1	không thực hiện
NICBE_5162	không thực hiện	không thực hiện	không thực hiện	không thực hiện	0	1	không thực hiện
BRANA_9788	0	1	0	không thực hiện	2	2	2
ADADI_1600	không thực hiện	2	1	2	1	2	2
ROSHY_8873	0	1	1	2	1	0	0
XANST_27	1	1	1	0	1	1	0
CONCA_4103	không thực hiện	không thực hiện	0	không thực hiện	1	2	1
AMAPA_1826	không thực hiện	1	1	1	0	2	0
SPIOL_0410	1	2	1	1	1	2	2
KOCSC_1672	1	2	1	1	1	2	0
SETIT_2080	0	0	không thực hiện	2	1	2	1
ALLCE_3035	không thực hiện	1	1	2	2	2	2
DIGSA_5107	1	1	không thực hiện	không thực hiện	0	1	1

AMAPA_4787	không thực hiện	2	1	1	1	2	1
SPIOL_0401	1	1	0	1	1	2	1
SEDAL_8241	0	1	0	1	0	1	1
CAMSA_6215	0	2	không thực hiện	không thực hiện	không thực hiện	2	2
CUCME_4756	0	0	không thực hiện	2	1	1	0
ERATE_4149	1	1	không thực hiện	không thực hiện	2	2	2
CANRO_3271	1	1	không thực hiện	1	1	1	2
ERATE_2090	0	1	1	0	1	1	0
ANDGE_6461	không thực hiện	1	không thực hiện	2	2	1	1
DIGSA_5109	0	1	0	1	1	0	không thực hiện
ERATE_4824	1	1	0	1	1	1	1
AMAGR_5230	không thực hiện	1	0	1	1	2	1
AMBTR_1537	không thực hiện	1	1	1	1	1	1

BRANA_6036	không thực hiện	1	không thực hiện	1	1	1	1
ROSHY_6783	1	1	không thực hiện	1	0	0	1
AMACR_2643	không thực hiện	0	không thực hiện	1	1	0	2
TAROF_2111	1	1	1	0	1	2	1
CANRO_3976_var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
ERATE_4149_var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	0	2	0
ALLCE_3035_var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	1	1
TAROF_2111_var	0	không thực hiện	không thực hiện	không thực hiện	0	2	1
CUCME_4756_var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
XANST_27_var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	0	không thực hiện	không thực hiện

Sau đó, các cây con trong vật chứa không được phun tương ứng với các cấu trúc có điểm bằng 2 được ghép ở xấp xỉ bảy tuần sau biến nạp và trồng dưới dạng cây trồng R0 bằng cách sử dụng phương pháp tiêu chuẩn đã biết trong lĩnh vực. Việc chọn lọc các

cây con tương ứng với điểm không dung nạp bằng 0 và 1 cũng được trồng để dùng làm đối chứng âm. Cây trồng R0 được trồng trong nhà kính trong điều kiện ươm cả ngày (18 giờ có ánh sáng ở nhiệt độ 80°F sau đó là 6 giờ trong tối ở nhiệt độ 74°F) trong xấp xỉ bốn tuần nữa. Vào mười một tuần sau biến nạp, cây trồng R0 được phun bằng hai lần phun của cùng dung dịch chất diệt cỏ được mô tả ở trên để có tỉ lệ sử dụng cuối cùng bằng 20g/ha. Đối với mỗi cấu trúc ADN được thử nghiệm, 15-30 cây trồng được biến nạp riêng lẻ được thử nghiệm. Xếp loại tổn thương cho chất diệt cỏ được tính điểm trực quan dựa trên lượng của tổn thương mô bên trên đất với 0% là không có tổn thương nhìn thấy và 100% là cây trồng chết hoàn toàn. Cây trồng đối chứng không chuyển gen được tính điểm xếp loại tổn thương là lớn hơn 30%. Sự dung nạp biên là tổn thương 30% hoặc nhỏ hơn, sự dung nạp tốt là tổn thương 20% hoặc nhỏ hơn, và sự dung nạp cực tốt được coi là tổn thương 10% hoặc nhỏ hơn. Điểm được thu thập bảy ngày sau khi xử lý và tính trung bình cho tất cả các cây trồng đối với mỗi cấu trúc ADN.

Kết quả của việc dùng dung nạp chất diệt cỏ ở mười một tuần cho cây trồng R0 xác nhận điểm xếp loại tổn thương có tỉ lệ phần trăm thấp quan sát được vào năm tuần. Đối với đánh giá vào mười một tuần, xếp loại tổn thương bất kỳ bằng 30% hoặc cao hơn là tương đương với xếp loại tổn thương cây đậu tương không chuyển gen. Một vài cấu trúc nổi bật vì mang lại sự dung nạp rất tốt đối với việc dùng chất diệt cỏ. Ví dụ, APG6 (SEQ ID NO:1) với PPO H_N90 (SEQ ID NO:110) chỉ có 3% tổn thương, mỗi APG6 (SEQ ID NO:1) với PPO H_N30 (SEQ ID NO:113) hoặc APG6 (SEQ ID NO:1) với PPO H_N40 (SEQ ID NO:114) chỉ có 5% tổn thương; peptit chuyển tiếp CAMSA_6215 (SEQ ID NO:21) với PPO H_N90 (SEQ ID NO:110) chỉ có 5% tổn thương. Ngược lại, peptit chuyển tiếp AMACR_2643 (SEQ ID NO:33) với PPO H_N90 (SEQ ID NO:110) có điểm tổn thương bằng 50%. Dữ liệu được nêu trong Bảng 6, trong đó n.d. chỉ ra phân tích không được thực hiện.

Bảng 6. Điểm dung nạp vào 11 tuần ở cây đậu tương

Peptit chuyển tiếp	H_N 20	H_N 30	H_N 40	H_N 50	H_N 90	H_N 100
APG6	không thực hiện	5	5	không thực hiện	3	15

12G088600TP	không thực hiện	không thực hiện	không thực hiện	không thực hiện	35	không thực hiện
CANRO_3976	không thực hiện	không thực hiện	không thực hiện	không thực hiện	30	không thực hiện
SENOB_8832	không thực hiện	không thực hiện	15	không thực hiện	không thực hiện	không thực hiện
NICBE_5162	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
BRANA_9788	25	không thực hiện	không thực hiện	40	25	30
ADADI_1600	20	không thực hiện	40	không thực hiện	15	30
ROSHY_8873	không thực hiện	không thực hiện	30	không thực hiện	40	không thực hiện
XANST_27	không thực hiện	35	không thực hiện	40	30	không thực hiện
CONCA_4103	không thực hiện	không thực hiện	không thực hiện	không thực hiện	30	35
AMAPA_1826	không thực hiện	35	không thực hiện	không thực hiện	30	không thực hiện

SPIOL_0410	20	không thực hiện	không thực hiện	không thực hiện	30	50
KOCSC_1672	20	không thực hiện	15	40	15	không thực hiện
SETIT_2080	không thực hiện	không thực hiện	35	40	25	không thực hiện
ALLCE_3035	30	35	30	40	35	30
DIGSA_5107	không thực hiện	không thực hiện	không thực hiện	không thực hiện	35	không thực hiện
AMAPA_4787	25	không thực hiện	không thực hiện	40	15	không thực hiện
SPIOL_0401	không thực hiện	không thực hiện	không thực hiện	không thực hiện	30	không thực hiện
SEDAL_8241	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
CAMSA_6215	20	không thực hiện	không thực hiện	không thực hiện	5	35
CUCME_4756	không thực hiện	không thực hiện	35	không thực hiện	25	không thực hiện
ERATE_4149	không thực hiện	không thực hiện	không thực hiện	40	30	30

CANRO_3271	không thực hiện	không thực hiện	không thực hiện	không thực hiện	30	35
ERATE_2090	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
ANDGE_6461	không thực hiện	không thực hiện	15	35	không thực hiện	không thực hiện
DIGSA_5109	không thực hiện	35	không thực hiện	không thực hiện	40	không thực hiện
ERATE_4824	không thực hiện	không thực hiện	không thực hiện	không thực hiện	35	không thực hiện
AMAGR_5230	không thực hiện	không thực hiện	không thực hiện	không thực hiện	30	35
AMBTR_1537	30	không thực hiện	không thực hiện	không thực hiện	không thực hiện	40
BRANA_6036	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
ROSHY_6783	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
AMACR_2643	không thực hiện	không thực hiện	không thực hiện	không thực hiện	50	40

TAROF_2111	không thực hiện	không thực hiện	không thực hiện	không thực hiện	25	không thực hiện
CANRO_3976_var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
ERATE_4149_var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	35	không thực hiện
ALLCE_3035_var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	15	35
TAROF_2111_var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	15	không thực hiện
CUCME_4756_var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
XANST_27_var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện

Các gen mã hóa mười protoporphyrinogen oxidaza HemY R2N30, R2N40, R2N40opt, R2N70, R2N90, R2N100, R1N473, R1N533, R1N171, R1N311, và R1N33 được liên kết có điều khiển với ba mươi chín peptit chuyển tiếp khác nhau và được tách dòng vào vật truyền biến nạp thực vật cơ bản như được mô tả trong Ví dụ 3. Điều này cho phép so sánh cạnh nhau mười HemY protoporphyrinogen oxidaza khác nhau với ba mươi chín peptit chuyển tiếp khác nhau bằng cách sử dụng cùng thành phần vùng khởi đầu và UTR 3' trong mọi cấu trúc ADN. Sử dụng các vật truyền biến nạp thực vật này để biến nạp phôi cắt của đậu tương (mầm nguyên sinh A3555) bằng cách sử dụng *A. tumefaciens* và phương pháp tiêu chuẩn đã biết trong lĩnh vực. Cây ghép bốn trăm mô cây bên ngoài đối với mỗi cấu trúc. Sử dụng dung dịch chất diệt cỏ PPO diệt trùng để thử

nghiệm sự dung nạp chất diệt cỏ. Dung dịch chất diệt cỏ gồm có 0,3 g S-3100 trong chất cô dầu cây trồng (5,0 ml) và 495 ml nước khử ion.

Vào năm tuần sau biến nạp, đối với mỗi vật chứa trong bốn vật chứa cấu trúc ADN (mỗi vật chứa có 30-40 cây trồng được biến nạp riêng lẻ) được phun bằng hai lần phun của dung dịch chất diệt cỏ PPO tiết trùng để có tỉ lệ sử dụng cuối cùng bằng 20 g/ha. Sau đó các cây con đã được xử lý nhận được ít nhất là 15 giờ phơi sáng sau khi phun mỗi ngày trong bốn ngày. Khi kết thúc ngày bốn sau khi dùng S-3100, các cây con đã được xử lý được chụp ảnh và tính điểm trên thang trực quan của màu xanh (màu xanh đại diện cho mô cây trồng quang hợp khỏe mạnh khi so với mô bị tẩy trắng quang) so với sự hư hại. Giá trị điểm bằng 0 đối với sự dung nạp kém, sự hư hại cao, màu xanh ít; 1 đối với sự dung nạp một chút, sự hư hại trung bình, màu xanh vừa phải; và 2 đối với sự dung nạp tốt, sự hư hại ít, màu xanh nhiều. Sự tính điểm đối với mỗi cấu trúc được thể hiện trong Bảng 7, trong đó n.d. chỉ phân tích không được thực hiện. Kết quả chỉ ra rằng một vài cấu trúc mang lại sự dung nạp đối với chất diệt cỏ PPO.

Bảng 7. Điểm dung nạp vào 5 tuần ở cây đậu tương

Peptit chuyển tiếp	R1N171	R1N473	R1N533	R2N30	R2N40	R2N40opt	R2N70	R2N90	R2N100	R1N333
APG6	0	2	0	2	không thực hiện	1	không thực hiện	không thực hiện	0	không thực hiện
12G088600TP	0	0	2	không thực hiện	không thực hiện	không thực hiện	2	0	0	0
CANRO_3976	0	1	0	1	không thực hiện	không thực hiện	1	không thực hiện	0	0
SENOB_8832	không thực hiện	1	0	2	không thực hiện	0	0	không thực hiện	0	0
NICBE_5162	1	không thực hiện	không thực hiện	không thực hiện	1	1	không thực hiện	0	0	không thực hiện
BRANA_9788	không thực hiện	1	1	không thực hiện	không thực hiện	1	0	không thực hiện	0	0

ADADI_1 600	0	1	0	1	không thực hiện	2	không thực hiện	không thực hiện	không thực hiện	0
ROSHY_ 8873	1	1	không thực hiện	2	0	1	0	1	1	0
XANST_2 7	1	1	không thực hiện	2	0	0	không thực hiện	1	không thực hiện	1
CONCA_ 4103	1	1	1	2	không thực hiện	không thực hiện	không thực hiện	0	1	không thực hiện
AMAPA_ 1826	0	0	0	2	không thực hiện	1	không thực hiện	không thực hiện	không thực hiện	0
SPIOL_04 10	0	1	0	1	không thực hiện	2	không thực hiện	1	0	1
KOCSC_1 672	0	0	0	không thực hiện	không thực hiện	0	không thực hiện	0	không thực hiện	0
SETIT_20 80	không thực hiện	1	1	1	không thực hiện	không thực hiện	không thực hiện	0	1	0
ALLCE_3 035	1	1	1	2	không thực hiện	1	không thực hiện	không thực hiện	0	0
DIGSA_5 107	1	1	2	2	không thực hiện	1	0	0	không thực hiện	0
AMAPA_ 4787	0	1	không thực hiện	1	không thực hiện	1	không thực hiện	không thực hiện	0	0
SPIOL_04 01	0	0	0	1	không thực hiện	0	không thực hiện	1	1	0
SEDAL_8 241	1	0	1	không thực hiện	2	1	không thực hiện	1	1	0

CAMSA_ 6215	0	1	1	2	không thực hiện	1	không thực hiện	0	không thực hiện	không thực hiện
CUCME_ 4756	0	0	không thực hiện	1	không thực hiện	không thực hiện	0	1	0	0
ERATE_4 149	không thực hiện	1	2	1	không thực hiện	không thực hiện	không thực hiện	0	0	0
CANRO_ 3271	1	1	1	1	không thực hiện	không thực hiện	không thực hiện	1	0	1
ERATE_2 090	không thực hiện	0	2	2	không thực hiện	không thực hiện	không thực hiện	0	0	0
ANDGE_ 6461	0	1	0	2	không thực hiện	1	không thực hiện	không thực hiện	0	0
DIGSA_5 109	1	0	1	1	không thực hiện	1	không thực hiện	không thực hiện	1	0
ERATE_4 824	0	1	0	không thực hiện	không thực hiện	2	không thực hiện	0	0	1
AMAGR_ 5230	0	2	0	2	không thực hiện	không thực hiện	không thực hiện	0	1	0
AMBTR_ 1537	0	0	1	1	không thực hiện	0	không thực hiện	0	0	1
BRANA_ 6036	1	1	không thực hiện	1	không thực hiện	0	không thực hiện	0	0	0
ROSHY_ 6783	1	1	1	1	0	0	1	0	1	1
AMACR_ 2643	0	1	1	1	không thực hiện	0	không thực hiện	0	0	0

TAROF_2 111	0	2	0	không thực hiện	2	1	0	0	0	0
CANRO_ 3976_var	không thực hiện	không thực hiện	không thực hiện	0	1	không thực hiện	không thực hiện	không thực hiện	không thực hiện	1
ERATE_4 149_var	0	0	1	1	1	1	không thực hiện	không thực hiện	không thực hiện	không thực hiện
ALLCE_3 035_var	không thực hiện	không thực hiện	0	1	1	không thực hiện	không thực hiện	không thực hiện	0	1
TAROF_2 111_var	0	0	0	1	1	2	không thực hiện	không thực hiện	0	không thực hiện
CUCME_ 4756_var	không thực hiện	không thực hiện	2	không thực hiện	2	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
XANST_2 7_var	1	1	2	1	2	1	không thực hiện	không thực hiện	không thực hiện	không thực hiện

Sau đó các cây con trong vật chứa không được phun tương ứng với các cấu trúc có điểm bằng 2 được ra ngôi vào xấp xỉ bảy tuần sau biến nạp và trồng dưới dạng cây trồng R0 bằng cách sử dụng phương pháp tiêu chuẩn đã biết trong lĩnh vực. Việc chọn lọc các cây con tương ứng với điểm không dung nạp bằng 0 và 1 cũng được trồng để dùng làm đối chứng âm. Cây trồng R0 được trồng trong nhà kính trong điều kiện ươm cả ngày (18 giờ có ánh sáng ở nhiệt độ 80°F sau đó là 6 giờ trong tối ở nhiệt độ 74°F) trong xấp xỉ bốn tuần nữa. Vào mười một tuần sau biến nạp, cây trồng R0 được phun bằng hai lần phun của cùng dung dịch chất diệt cỏ được mô tả ở trên để có tỉ lệ sử dụng cuối cùng bằng 20 g/ha. Đối với mỗi cấu trúc ADN được thử nghiệm, 15-30 cây trồng được biến nạp riêng lẻ được thử nghiệm. Xếp loại tổn thương cho chất diệt cỏ được tính điểm trực quan dựa trên lượng của tổn thương mô bên trên đất với 0% là không có tổn thương nhìn thấy và 100% là cây trồng chết hoàn toàn. Cây trồng đối chứng không chuyển gen được tính điểm xếp loại tổn thương là lớn hơn 30%. Sự dung nạp biên là tổn thương 30% hoặc nhỏ hơn, sự dung nạp tốt là tổn thương 20% hoặc nhỏ hơn, và sự dung nạp cực tốt được

coi là tổn thương 10% hoặc nhỏ hơn. Điểm được thu thập bảy ngày sau khi xử lý và tính trung bình cho tất cả các cây trồng đối với mỗi cấu trúc ADN.

Kết quả của việc dùng dung nạp chất diệt cỏ ở mười một tuần cho cây trồng R0 xác nhận điểm xếp loại tổn thương có tỉ lệ phần trăm thấp quan sát được vào năm tuần. Đối với đánh giá vào mười một tuần, xếp loại tổn thương bất kỳ bằng 30% hoặc cao hơn là tương đương với xếp loại tổn thương cây đậu tương không chuyển gen. Một ít cấu trúc nổi bật vì mang lại sự dung nạp rất tốt đối với việc dùng chất diệt cỏ. Ví dụ, peptit chuyển tiếp ANDGE_6461 (SEQ ID NO:26) với R2N30 (SEQ ID NO:163) chỉ có 7% tổn thương. Dữ liệu được nêu trong Bảng 8, trong đó n.d. chỉ ra phân tích không được thực hiện.

Bảng 8. Điểm dung nạp vào 11 tuần ở cây đậu tương

Peptit chuyển tiếp	R1N 171	R1N 473	R1N 533	R2 N30	R2 N40	R2 N40opt	R2 N70	R1N 333
APG6	không thực hiện	30	không thực hiện	17	không thực hiện	20	không thực hiện	không thực hiện
12G088600TP	không thực hiện	không thực hiện	40	không thực hiện	không thực hiện	không thực hiện	30	không thực hiện
CANRO_3976	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
SENOB_8832	không thực hiện	không thực hiện	không thực hiện	25	không thực hiện	không thực hiện	không thực hiện	không thực hiện
NICBE_5162	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
BRANA_9788	không thực hiện	35	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện

ADADI_1600	không thực hiện	không thực hiện	không thực hiện	25	không thực hiện	30	không thực hiện	không thực hiện
ROSHY_8873	không thực hiện	không thực hiện	không thực hiện	35	không thực hiện	30	không thực hiện	35
XANST_27	không thực hiện	không thực hiện	không thực hiện	20	không thực hiện	25	không thực hiện	35
CONCA_4103	không thực hiện	không thực hiện	không thực hiện	25	không thực hiện	không thực hiện	không thực hiện	không thực hiện
AMAPA_1826	không thực hiện	không thực hiện	không thực hiện	25	không thực hiện	không thực hiện	không thực hiện	không thực hiện
SPIOL_0410	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	35	không thực hiện	không thực hiện
KOCSC_1672	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
SETIT_2080	không thực hiện	không thực hiện	không thực hiện	20	không thực hiện	không thực hiện	không thực hiện	35
ALLCE_3035	không thực hiện	không thực hiện	không thực hiện	25	không thực hiện	không thực hiện	không thực hiện	không thực hiện
DIGSA_5107	30	40	35	35	không thực hiện	không thực hiện	không thực hiện	không thực hiện

AMAPA_4787	không thực hiện	không thực hiện	không thực hiện	25	không thực hiện	không thực hiện	không thực hiện	không thực hiện
SPIOL_0401	không thực hiện	không thực hiện	không thực hiện	15	không thực hiện	không thực hiện	không thực hiện	không thực hiện
SEDAL_8241	không thực hiện	không thực hiện	không thực hiện	không thực hiện	20	không thực hiện	không thực hiện	không thực hiện
CAMSA_6215	không thực hiện	không thực hiện	không thực hiện	15	không thực hiện	20	không thực hiện	không thực hiện
CUCME_4756	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
ERATE_4149	không thực hiện	không thực hiện	35	25	không thực hiện	không thực hiện	không thực hiện	không thực hiện
CANRO_3271	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
ERATE_2090	không thực hiện	không thực hiện	35	15	không thực hiện	không thực hiện	không thực hiện	không thực hiện
ANDGE_6461	không thực hiện	không thực hiện	không thực hiện	7	không thực hiện	không thực hiện	không thực hiện	không thực hiện
DIGSA_5109	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện

ERATE_4824	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	25	không thực hiện	không thực hiện
AMAGR_5230	không thực hiện	35	không thực hiện	35	không thực hiện	không thực hiện	không thực hiện	không thực hiện
AMBTR_1537	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
BRANA_6036	không thực hiện	không thực hiện	không thực hiện	25	không thực hiện	không thực hiện	không thực hiện	không thực hiện
ROSHY_6783	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
AMACR_2643	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
TAROF_2111	không thực hiện	40	không thực hiện	không thực hiện	20	không thực hiện	không thực hiện	không thực hiện
CANRO_3976 _var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
ERATE_4149_ var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện
ALLCE_3035_ var	không thực hiện	không thực hiện	không thực hiện	25	không thực hiện	không thực hiện	không thực hiện	không thực hiện

TAROF_2111_ var	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện	25	không thực hiện	không thực hiện
CUCME_4756 _var	không thực hiện	không thực hiện	35	không thực hiện	25	không thực hiện	không thực hiện	không thực hiện
XANST_27_va r	không thực hiện	30	35	không thực hiện	không thực hiện	không thực hiện	không thực hiện	không thực hiện

Các gen mã hóa HemG protoporphyrinogen oxidaza H_N90 được liên kết có điều khiển với 44 peptit chuyển tiếp khác nhau và được tách dòng vào vật truyền biến nạp thực vật cơ bản như được mô tả trong Ví dụ 3. Điều này cho phép so sánh cạnh nhau các peptit chuyển tiếp khác nhau bằng cách sử dụng cùng thành phần vùng khởi đầu, protein dung nạp chất diệt cỏ, và UTR 3' trong mọi cấu trúc ADN. Sử dụng các vật truyền biến nạp thực vật này để biến nạp phôi cắt của đậu tương (mầm nguyên sinh AG3555) bằng cách sử dụng *A. tumefaciens* và phương pháp tiêu chuẩn đã biết trong lĩnh vực. Từ bốn trăm đến 4.5000 cây trồng riêng lẻ chuyển gen được thử nghiệm đối với mỗi cấu trúc. Sử dụng dung dịch chất diệt cỏ PPO tiết trùng để thử nghiệm sự dung nạp chất diệt cỏ. Dung dịch chất diệt cỏ gồm có 0,3 g S-3100 trong chất cô dầu cây trồng (5,0 ml) và 495 ml nước khử ion.

Vào năm tuần sau biến nạp, phun cây trồng bằng hai lần phun của dung dịch chất diệt cỏ PPO tiết trùng để có tỉ lệ sử dụng cuối cùng bằng 20 g/ha. Đối với mỗi cấu trúc ADN được thử nghiệm, từ 400 đến 4.5000 bản sao được thực hiện. Sau đó, các cây con đã được xử lý nhận được ít nhất là 15 giờ phơi sáng sau khi phun mỗi ngày trong bốn ngày. Khi kết thúc ngày bốn sau khi dùng S-3100, các cây con đã được xử lý được tính điểm về tỉ lệ phần trăm của tần suất vượt qua tương đối (được xác định dưới dạng tỉ lệ phần trăm của tất cả các cây trồng riêng lẻ đối với cấu trúc ADN mà thể hiện trực quan sự dung nạp đối với việc dùng chất diệt cỏ tương đối với cây trồng đối chứng chuyển gen được phun bằng dung dịch chỉ có chất hoạt động bề mặt.). Các cây con trong vật chứa không được phun được ra ngôi vào xấp xỉ bảy tuần sau biến nạp và trồng dưới dạng cây trồng R0. Cây trồng R0 được trồng trong nhà kính trong điều kiện ươm cả ngày (18 giờ

có ánh sáng ở nhiệt độ 80°F sau đó là 6 giờ trong tối ở nhiệt độ 74°F) trong xấp xỉ bốn tuần nữa. Vào mười một đến mười hai tuần sau biến nạp, cây trồng R0 được phun bằng hai lần phun của cùng dung dịch chất diệt cỏ được mô tả ở trên ở tỉ lệ 20 g/ha. Đối với mỗi cấu trúc ADN được thử nghiệm, từ 15 đến 45 bản sao được thực hiện. Xếp loại tổn thương cho chất diệt cỏ được thu thập vào từ ba đến bảy ngày sau khi xử lý. Đối với đánh giá vào mười một tuần, tỉ lệ phần trăm của cây trồng bằng hoặc dưới 10% tổn thương và bằng hoặc dưới 20% tổn thương được ghi nhận. Ở các tỉ lệ dùng chất diệt cỏ được thử nghiệm, cây trồng chuyển gen biểu hiện protoporphyrinogen oxidaza H_N90 mà không có bất kỳ peptit chuyển tiếp được liên kết có điều khiển nào (Đối chứng PPO), sản xuất ra không cây trồng có tổn thương 20% hoặc nhỏ hơn. Một vài peptit chuyển tiếp được liên kết có điều khiển với protein dung nạp chất diệt cỏ H_N90 nổi bật vì mang lại sự dung nạp cực tốt hoặc rất tốt đối với việc dùng chất diệt cỏ. Ví dụ, ở lần phun vào mười một tuần trên 50% cây trồng có điểm tổn thương bằng hoặc dưới 20% khi biểu hiện H-N90 được liên kết có điều khiển với ALLCE_3035 (57%), KOCSC_9516 (59%), CAMSA_6215 (69%), ROSHY_3269 (70%), ADADI_0544 (75%), CUCME_3420 (80%), SPIOL_1551 (85%), CUCME_4756 (89%), hoặc CONCA_3910 (90%). Dữ liệu được nêu trong Bảng 9.

Bảng 9. Điểm dung nạp vào 5 và 11 tuần ở cây đậu tương

Peptit chuyển tiếp	tần suất vượt qua tương đối khi phun 5 tuần	% cây trồng khi phun 11 tuần $\leq 10\%$	% cây trồng khi phun 11 tuần $\leq 20\%$
CUCME_4756	27%	0%	0%
CANRO_3271	23%	0%	0%
DIGSA_5109	24%	0%	0%
CAMSA_6215	68%	62%	69%
AMACR_2381	30%	0%	0%
ROSHY_3269	54%	25%	70%
CUCME_3420	51%	20%	80%
ADADI_0544	30%	20%	75%
SPIOL_1551	40%	70%	85%
NICBE_5162	9%	0%	0%

CUCME_4756	28%	26%	89%
BRANA_9788	11%	0%	0%
SPIOL_0410	18%	0%	0%
XANST_0027	22%	0%	0%
SETIT_2080	3%	0%	0%
ERATE_4149	3%	0%	0%
TAROF_2111	3%	0%	0%
CONCA_4103	26%	0%	0%
CANRO_3976	6%	0%	0%
AMACR_2643	3%	0%	0%
SPIOL_0401	6%	0%	0%
ADADI_1600	30%	0%	0%
ANDGE_6461	47%	0%	0%
ERATE_2090	11%	0%	0%
12G088600TP	13%	0%	0%
ALLCE_3035	5%	0%	0%
SENOB_8832	52%	0%	0%
TAROF_2111	66%	0%	0%
ROSHY_8873	10%	0%	0%
KOCSC_1672	25%	12%	24%
AMBTR_1537	2%	0%	0%
AMAPA_1826	7%	0%	0%
BRANA_6036	5%	0%	0%
CONCA_3910	40%	60%	90%
AMAPA_4787	6%	0%	0%
ROSHY_6783	0%	0%	0%
ALLCE_3035	26%	35%	57%
ERATE_4824	12%	0%	0%
AMAGR_5230	2%	0%	0%
SEDAL_8241	5%	0%	0%

DIGSA_5107	11%	0%	0%
KOCSC_9516	27%	16%	59%
XANST_0027_var	3%	0%	0%
APG6	60%	30%	63%
Đối chứng không PPO	1%	0%	0%

Ví dụ 5: Peptit chuyển tiếp và protoporphyrinogen oxidaza thử nghiệm ở cây ngô

Peptit chuyển tiếp được liên kết có điều khiển với protoporphyrinogen oxidaza được thử nghiệm ở cây ngô chuyển gen về sự dung nạp chất diệt cỏ PPO. Vật truyền biến nạp thực vật được tạo cấu trúc có chứa cấu trúc ADN có chứa phân tử ADN tái tổ hợp được tối ưu hóa để biểu hiện ở cây một lá mầm và mã hóa protoporphyrinogen oxidaza được liên kết có điều khiển với peptit chuyển tiếp. Sau đó sử dụng vật truyền biến nạp thực vật này để biến nạp cây ngô, và cây trồng đã được tái tạo được đánh giá về tính nhạy cảm của nó đối với chất diệt cỏ PPO.

Các gen mã hóa protoporphyrinogen oxidaza H_N90 được liên kết có điều khiển với mười bốn peptit chuyển tiếp khác nhau và được tách dòng vào vật truyền biến nạp thực vật cơ bản với nhiều thành phần vùng khởi đầu và UTR 3'. Việc sử dụng cùng protoporphyrinogen oxidaza trong mỗi cấu trúc ADN cho phép so sánh cạnh nhau các peptit chuyển tiếp khác nhau. Vật truyền biến nạp thực vật cũng được sản xuất ra với protoporphyrinogen oxidaza H_N90 mà không có peptit chuyển tiếp được liên kết có điều khiển bất kỳ (Đối chứng PPO). Sử dụng các vật truyền biến nạp thực vật này để biến nạp cây ngô bằng cách sử dụng *A. tumefaciens* và phương pháp tiêu chuẩn đã biết trong lĩnh vực. Cây trồng R0 được tái tạo được trồng và sau đó được sàng lọc để đánh giá mức độ dung nạp được thể hiện đối với việc sử dụng S-3100 (tỉ lệ từ 40 đến 80 g/ha) vào xấp xỉ 10-14 tuần sau biến nạp. Sự dung nạp được đánh giá trực quan từ 3 đến 10 ngày sau khi sử dụng chất diệt cỏ. Cây trồng đã được phun được tính điểm trên tỉ lệ phần trăm của tổn thương đối với toàn bộ phần bên trên đất của cây trồng sau khi xử lý chất diệt cỏ, tương đối với đối chứng. Đối với mỗi cấu trúc ADN được thử nghiệm, từ 10 đến 120 cây trồng được thử nghiệm và tỉ lệ tổn thương được tính trung bình. Tỉ lệ phần trăm của cây trồng R0 vượt qua ở điểm tổn thương 20% hoặc nhỏ hơn được ghi nhận. Cây trồng

chuyển gen bất kỳ sản xuất ra cấu trúc ADN với 50% hoặc nhiều hơn có tổn thương 20% hoặc nhỏ hơn được coi là cấu trúc ADN dung nạp cao. Cây trồng chuyển gen bất kỳ sản xuất ra cấu trúc ADN với 20% hoặc nhiều hơn có tổn thương 20% hoặc nhỏ hơn được coi là cấu trúc ADN dung nạp. Ở các tỉ lệ dùng chất diệt cỏ được thử nghiệm (S-3100 ở tỉ lệ từ 40 đến 80 g/ha), cây trồng chuyển gen biểu hiện protoporphyrinogen oxidaza H_N90 mà không có peptit chuyển tiếp được liên kết có điều khiển bất kỳ (Đối chứng PPO), với XANST_27 hoặc với ALLCE_3035 tạo ra không cây trồng với tổn thương 20% hoặc nhỏ hơn. Tuy nhiên, một vài cây trồng chuyển gen sản xuất ra peptit chuyển tiếp biểu hiện protoporphyrinogen oxidaza H_N90 mà dung nạp cao hoặc dung nạp: ADADI_0544 (41%), ANDGE_6461 (60%), CAMSA_6215 (vượt qua 60% và 41%), CONCA_3910 (36% và 45%), ROSHY_3269 (64% và 74%), SPIOL_1551 (50% và 55%), SETIT_9796 (55%). Dữ liệu được nêu trong Bảng 10.

Bảng 10. Điểm dung nạp ở ngô

Vùng khởi động	Peptit chuyển tiếp	UTR 3'	Tỉ lệ phần trăm có tổn thương 20% hoặc nhỏ hơn
A	SETIT_9796	E	55%
A	ACAOS_3432	E	37%
A	ADADI_0544	E	41%
A	TAROF_9570	E	29%
A	ALLCE_6618	E	31%
D	ROSHY_3269	H	74%
B	ROSHY_3269	F	64%
D	CONCA_3910	H	36%
B	CONCA_3910	F	45%
D	SPIOL_1551	H	55%
B	SPIOL_1551	F	50%
D	CAMSA_6215	H	41%
B	CAMSA_6215	F	60%
B	ANDGE_6461	F	60%

B	ADADI_1600	F	11%
D	XANST_27_var	H	0%
C	XANST_27_var	G	0%
A	ALLCE_3035	E	0%
B	ALLCE_3035	F	0%
C	Đối chứng không PPO	G	0%

Ví dụ 6: Peptit chuyển tiếp và protoporphyrinogen oxidaza thử nghiệm ở cây bông

Peptit chuyển tiếp được liên kết có điều khiển với protoporphyrinogen oxidaza được thử nghiệm ở cây bông chuyển gen về sự dung nạp chất diệt cỏ PPO. Vật truyền biến nạp thực vật được tạo cấu trúc có chứa cấu trúc ADN có chứa phân tử ADN tái tổ hợp được tối ưu hóa để biểu hiện ở cây hai lá mầm và mã hóa protoporphyrinogen oxidaza được liên kết có điều khiển với peptit chuyển tiếp. Sau đó sử dụng vật truyền biến nạp thực vật này để biến nạp cây bông, và cây trồng đã được tái tạo được đánh giá về tính nhạy cảm của nó đối với chất diệt cỏ PPO.

Các gen mã hóa protoporphyrinogen oxidaza H_N20 và H_N90 được liên kết có điều khiển với bốn peptit chuyển tiếp khác nhau và được tách dòng vào vật truyền biến nạp thực vật cơ bản như được mô tả trong Ví dụ 3. Điều này cho phép so sánh cạnh nhau của các peptit chuyển tiếp khác nhau bằng cách sử dụng cùng thành phần vùng khởi đầu và UTR 3' trong mọi cấu trúc ADN. Sử dụng các vật truyền biến nạp thực vật này để biến nạp cây bông bằng cách sử dụng *A. tumefaciens* và phương pháp tiêu chuẩn đã biết trong lĩnh vực. Cây trồng được tái tạo được trồng và sau đó được sàng lọc để đánh giá mức độ dung nạp được thể hiện đối với việc sử dụng S-3100 (tỉ lệ 20 g/ha) vào xấp xỉ từ 11 đến 12 tuần sau biến nạp. Sự dung nạp được đánh giá trực quan từ 3 đến 10 ngày sau khi sử dụng chất diệt cỏ. Cây trồng đã được phun được tính điểm trên tỉ lệ phần trăm của tổn thương đối với toàn bộ phần bên trên đất của cây trồng sau khi xử lý chất diệt cỏ, tương đối với đối chứng. Đối với mỗi cấu trúc ADN được thử nghiệm, từ 10 đến 15 bản sao được thử nghiệm và tỉ lệ tổn thương trung bình được tính trung bình. Điểm tổn thương trung bình bằng 50% hoặc nhỏ hơn được coi là cấu trúc ADN dung nạp cao với chất diệt cỏ, và điểm tổn thương trung bình lớn hơn 50% nhưng nhỏ hơn 80% được coi là

cấu trúc ADN dung nạp biên với chất diệt cỏ. Điểm tổn thương trung bình bằng hoặc cao hơn 80% được coi là không phân biệt so với cây trồng đối chứng. Cây bông chuyển gen biểu hiện protoporphyrinogen oxidaza H_N90 được liên kết có điều khiển với cây trồng sản xuất ra CAMSA_6215 mà dung nạp cao với chất diệt cỏ với điểm tổn thương trung bình 38%. Cây bông chuyển gen biểu hiện protoporphyrinogen oxidaza H_N90 được liên kết có điều khiển với cây trồng sản xuất ra AMAPA_4787 mà dung nạp biên với chất diệt cỏ với điểm tổn thương trung bình bằng 63%.

YÊU CẦU BẢO HỘ

1. Phân tử ADN tái tổ hợp có chứa trình tự ADN mã hóa peptit chuyển tiếp được liên kết có điều khiển với trình tự ADN mã hóa protoporphyrinogen oxidaza khác loại, trong đó peptit chuyển tiếp có chứa trình tự axit amin có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự SEQ ID NO:21, và trong đó protoporphyrinogen oxidaza khác loại có chứa trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO:100-119, SEQ ID NO:163-182, và các SEQ ID NO:224-228.
2. Phân tử ADN tái tổ hợp theo điểm 1, trong đó trình tự ADN mã hóa peptit chuyển tiếp có chứa trình tự axit nucleic có độ tương đồng bằng ít nhất là 97 phần trăm với trình tự SEQ ID NO:71 hoặc SEQ ID NO:91.
3. Phân tử ADN tái tổ hợp theo điểm 1, trong đó trình tự ADN mã hóa protoporphyrinogen oxidaza khác loại có chứa trình tự axit nucleic được chọn từ nhóm gồm có các SEQ ID NO:121-162 và SEQ ID NO:183-223, SEQ ID NO:229-235.
4. Phân tử ADN tái tổ hợp theo điểm 1, còn chứa vùng khởi đầu khác loại được liên kết có điều khiển với trình tự ADN mã hóa peptit chuyển tiếp.
5. Cấu trúc ADN có chứa phân tử ADN theo điểm 1, trong đó phân tử ADN này được liên kết có điều khiển với vùng khởi đầu khác loại.
6. Cấu trúc ADN theo điểm 5, trong đó cấu trúc ADN có mặt trong hệ gen của cây trồng, hạt, hoặc tế bào chuyển gen.
7. Cây trồng, hạt, hoặc tế bào chuyển gen có chứa phân tử ADN tái tổ hợp theo điểm 1.
8. Cây trồng, hạt, hoặc tế bào chuyển gen theo điểm 7, trong đó cây trồng, hạt, hoặc tế bào này dung nạp đối với ít nhất một chất diệt cỏ PPO.

9. Cây trồng, hạt, hoặc tế bào chuyển gen theo điểm 8, trong đó chất diệt cỏ PPO được chọn từ nhóm gồm có: acifluorfen, fomesafen, lactofen, floglycofen-ethyl, oxyfluorfen, flumioxazin, azafenidin, carfentrazone-ethyl, sulfentrazone, fluthiacet-methyl, oxadiargyl, oxadiazon, pyraflufen-ethyl, saflufenacil, tiafenacil, 1,5-dimethyl-6-thioxo-3-(2,2,7-triflo-3,4-dihydro-3-oxo-4-prop-2-ynyl-2H-1,4-benzoxazin-6-yl)-1,3,5-triazinan-2,4-dion, và S-3100.

10. Cây trồng, hạt, hoặc tế bào chuyển gen theo điểm 7, trong đó cây trồng, hạt, hoặc tế bào này dung nạp đối với ít nhất một chất diệt cỏ bổ sung.

11. Phương pháp sản xuất cây trồng dung nạp chất diệt cỏ có chứa các bước:

- a) biến nạp tế bào cây trồng với phân tử ADN tái tổ hợp theo điểm 1; và
- b) từ đó tái tạo cây trồng dung nạp chất diệt cỏ mà có chứa phân tử ADN.

12. Phương pháp theo điểm 11, còn chứa bước lai cây trồng đã được tái tạo với bản thân nó hoặc với cây trồng thứ hai để sản xuất một hoặc nhiều cây trồng thế hệ con.

13. Phương pháp theo điểm 12, còn chứa bước chọn lọc cây trồng thế hệ con mà dung nạp với ít nhất một chất diệt cỏ PPO.

14. Phương pháp theo điểm 13, trong đó chất diệt cỏ PPO được chọn từ nhóm gồm có: acifluorfen, fomesafen, lactofen, floglycofen-ethyl, oxyfluorfen, flumioxazin, azafenidin, carfentrazone-ethyl, sulfentrazone, fluthiacet-methyl, oxadiargyl, oxadiazon, pyraflufen-ethyl, saflufenacil, tiafenacil, 1,5-dimethyl-6-thioxo-3-(2,2,7-triflo-3,4-dihydro-3-oxo-4-prop-2-ynyl-2H-1,4-benzoxazin-6-yl)-1,3,5-triazinan-2,4-dion, và S-3100.

15. Phương pháp kiểm soát hoặc ngăn chặn sự phát triển cỏ dại ở khu vực trồng cây có chứa bước dùng lượng hữu hiệu của ít nhất một chất diệt cỏ PPO cho khu vực trồng cây mà có chứa cây trồng hoặc hạt chuyển gen theo điểm 8, trong đó cây trồng hoặc hạt chuyển gen dung nạp với chất diệt cỏ PPO.

16. Phương pháp theo điểm 15, trong đó chất diệt cỏ PPO được chọn từ nhóm gồm có: acifluorfen, fomesafen, lactofen, floglycofen-ethyl, oxyfluorfen, flumioxazin, azafenidin, carfentrazone-ethyl, sulfentrazone, fluthiacet-methyl, oxadiargyl, oxadiazon, pyraflufen-ethyl, saflufenacil, tiafenacil, 1,5-dimethyl-6-thioxo-3-(2,2,7-trifluoro-3,4-dihydro-3-oxo-4-prop-2-ynyl-2H-1,4-benzoxazin-6-yl)-1,3,5-triazinan-2,4-dione, và S-3100.

17. Phương pháp kiểm soát sự sinh trưởng của cỏ dại dung nạp chất diệt cỏ có chứa bước:

- a) trồng cây trồng hoặc hạt theo điểm 10 ở khu vực trồng cây; và
- b) dùng chất diệt cỏ PPO và ít nhất một chất diệt cỏ khác cho khu vực trồng cây, trong đó cây trồng hoặc hạt dung nạp đối với chất diệt cỏ PPO và ít nhất một chất diệt cỏ khác.

18. Phương pháp theo điểm 17, trong đó chất diệt cỏ PPO được chọn từ nhóm gồm có: acifluorfen, fomesafen, lactofen, floglycofen-ethyl, oxyfluorfen, flumioxazin, azafenidin, carfentrazone-ethyl, sulfentrazone, fluthiacet-methyl, oxadiargyl, oxadiazon, pyraflufen-ethyl, saflufenacil, và S-3100.

19. Phương pháp theo điểm 17, trong đó ít nhất một chất diệt cỏ bổ sung được chọn từ nhóm gồm có: chất ức chế ACCase, chất ức chế ALS, chất ức chế EPSPS, auxin tổng hợp, chất ức chế quang hợp, chất ức chế glutamin synthetase, chất ức chế HPPD, chất ức chế PPO, và chất ức chế axit béo mạch dài.

20. Phương pháp theo điểm 19, trong đó chất ức chế ACCase là aryloxyphenoxy propionat hoặc cyclohexandione; chất ức chế ALS là sulfonylurea, imidazolinon, triazolopyrimidin, hoặc triazolinon; chất ức chế EPSPS là glyphosat; auxin tổng hợp là chất diệt cỏ phenoxy, axit benzoic, axit carboxylic, hoặc semicarbazone; chất ức chế quang hợp là triazin, triazinon, nitril, benzothiadiazol, hoặc ure; chất ức chế glutamin synthetase là glufosinat; chất ức chế HPPD là isoxazol, pyrazolon, hoặc triketon; hoặc chất ức chế axit béo mạch rất dài là cloxetamit, oxyxetamit, hoặc pyrazol.

DANH MỤC TRÌNH TỰ

<110> Monsanto Technology LLC

<120> PHÂN TỬ ADN TÁI TỔ HỢP MÃ HÓA PROTEIN DUNG NẠP CHẤT DIỆT CỎ, CÂY TRỒNG, HẠT, HOẶC TẾ BÀO CHUYỂN GEN CHỨA PHÂN TỬ NÀY, PHƯƠNG PHÁP SẢN XUẤT CÂY TRỒNG DUNG NẠP CHẤT DIỆT CỎ VÀ PHƯƠNG PHÁP KIỂM SOÁT SỰ SINH TRƯỞNG CỦA CỎ DẠY DUNG NẠP CHẤT DIỆT CỎ

<130> MONS:397WO

<150> US 62/368,840

<151> 29-07-2016

<160> 297

<170> PatentIn phiên bản 3.5

<210> 1

<211> 68

<212> PRT

<213> Arabidopsis thaliana

<400> 1

Met	Ala	Thr	Ala	Thr	Thr	Thr	Ala	Thr	Ala	Ala	Phe	Ser	Gly	Val	Val
1				5					10					15	

Ser	Val	Gly	Thr	Glu	Thr	Arg	Arg	Ile	Tyr	Ser	Phe	Ser	His	Leu	Gln
			20					25					30		

Pro	Ser	Ala	Ala	Phe	Pro	Ala	Lys	Pro	Ser	Ser	Phe	Lys	Ser	Leu	Lys
		35					40					45			

Leu	Lys	Gln	Ser	Ala	Arg	Leu	Thr	Arg	Arg	Leu	Asp	His	Arg	Pro	Phe
	50					55					60				

Val	Val	Arg	Cys
65			

<210> 2

<211> 68

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 2

Met	Ala	Ser	Ser	Thr	Thr	Thr	Ala	Thr	Ala	Ala	Phe	Ser	Gly	Val	Val
1				5					10					15	

Ser Val Gly Thr Glu Thr Arg Arg Ile Tyr Ser Phe Ser His Leu Gln
20 25 30

Pro Ser Ala Ala Phe Pro Ala Lys Pro Ser Ser Phe Lys Ser Leu Lys
35 40 45

Leu Lys Gln Ser Ala Arg Leu Thr Arg Arg Leu Asp His Arg Pro Phe
50 55 60

Val Val Arg Cys
65

<210> 3
<211> 35
<212> PRT
<213> *Gossypium raimondii*

<400> 3

Met Leu Asn Ile Ala Pro Ser Cys Val Leu Ala Ser Gly Ile Ser Lys
1 5 10 15

Pro Val Thr Lys Met Ala Ser Thr Glu Asn Lys Asp Asp His Ser Ser
20 25 30

Ala Lys Arg
35

<210> 4
<211> 45
<212> PRT
<213> *Canavalia rosea*

<400> 4

Met Val Ala Val Phe Asn Asp Val Val Phe Pro Pro Ser Gln Thr Leu
1 5 10 15

Leu Arg Pro Ser Phe His Ser Pro Thr Phe Phe Phe Ser Ser Pro Thr
20 25 30

Pro Lys Phe Thr Arg Thr Arg Pro Asn Arg Ile Leu Arg
35 40 45

<210> 5
<211> 60
<212> PRT
<213> *Senna obtusifolia*

<400> 5

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

Pro Thr Gln Thr Leu Val His Ser Asn Thr Ser Phe Ile Ser Ser Arg
20 25 30

Thr Cys Phe Val Cys Pro Ile Ile Pro Phe Pro Ser Arg Ser Gln Leu
35 40 45

Asn Arg Arg Ile Ala Cys Ile Arg Ser Asn Val Arg
50 55 60

<210> 6

<211> 55

<212> PRT

<213> *Nicotiana benthamiana*

<400> 6

Met Thr Thr Thr Pro Val Ala Asn His Pro Asn Ile Phe Thr His Arg
1 5 10 15

Ser Pro Pro Ser Ser Ser Ser Ser Ser Pro Ser Ala Phe Leu Thr Arg
20 25 30

Thr Ser Phe Leu Pro Phe Ser Ser Ile Cys Lys Arg Asn Ser Val Asn
35 40 45

Cys Asn Gly Trp Arg Thr Arg
50 55

<210> 7

<211> 34

<212> PRT

<213> *Brassica napus*

<400> 7

Met Asp Phe Ser Leu Leu Arg Pro Ala Ser Thr Gln Pro Phe Leu Ser
1 5 10 15

Pro Phe Ser Asn Pro Phe Pro Arg Ser Arg Pro Tyr Lys Pro Leu Asn
20 25 30

Leu Arg

<210> 8
 <211> 37
 <212> PRT
 <213> *Adansonia digitata*

<400> 8

Met Ala Ile Leu Ile Asp Leu Ser Leu Leu Arg Ser Ser Pro Ser Val
 1 5 10 15

Phe Ser Phe Ser Lys Pro Asn His Arg Ile Pro Pro Arg Ile Tyr Lys
 20 25 30

Pro Phe Lys Leu Arg
 35

<210> 9
 <211> 47
 <212> PRT
 <213> *Rosa hybrida osiana*

<400> 9

Met Thr Thr Leu Ser Arg Leu Ala Asp Leu Pro Ser Phe Ala Ala Pro
 1 5 10 15

Pro Pro Leu Leu Thr His Arg Pro Pro Pro Ser Val Phe Leu Thr Pro
 20 25 30

Lys Pro Thr Lys Pro Ser Pro Pro His His Phe Phe Lys Leu Arg
 35 40 45

<210> 10
 <211> 46
 <212> PRT
 <213> *Xanthium strumarium*

<400> 10

Met Ser Ser Leu Thr Asp Leu Pro Ser Leu Asn His Tyr Arg Thr Cys
 1 5 10 15

Ser Pro Arg Pro Phe Pro Ile Ser Arg Gln Thr Ser Ser Ser Ile Asn
 20 25 30

Pro Asn Asn Leu Thr Thr Ser Asn Arg Trp Arg Arg Phe Arg
 35 40 45

<210> 11

<211> 35
 <212> PRT
 <213> *Conyza canadensis*

<400> 11

Met Thr Ser Leu Thr Asn Phe Thr Pro Leu Lys Leu Thr Asn Pro Asn
 1 5 10 15

Tyr Leu Asn Thr Thr Thr Thr Tyr Asn His Arg Lys Leu Ser Asn Phe
 20 25 30

Arg Phe Arg
 35

<210> 12
 <211> 45
 <212> PRT
 <213> *Amaranthus palmeri*

<400> 12

Met Ser Ala Met Ala Leu Ser Ser Ser Ile Leu Gln Cys Pro Pro His
 1 5 10 15

Ser Asp Ile Ser Phe Arg Phe Ser Ala Tyr Thr Ala Thr Arg Ser Pro
 20 25 30

Phe Phe Phe Gly Arg Pro Arg Lys Leu Ser Tyr Ile His
 35 40 45

<210> 13
 <211> 47
 <212> PRT
 <213> *Spinacia oleracea*

<400> 13

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

Ser Ser Met Ser Leu Ser His Cys Arg His Asn Arg Ile Thr Ile Leu
 20 25 30

Ile Pro Ser Ser Ser Leu Arg Arg Arg Gly Gly Ser Ser Ile Arg
 35 40 45

<210> 14
 <211> 61
 <212> PRT

<213> Kochia Scoparia

<400> 14

Met Ser Ala Met Ala Ser Pro Ser Ile Ile Pro Gln Ser Phe Leu Gln
1 5 10 15

Arg Ser Pro Thr Ser Leu Gln Ser Arg Ser Asn Tyr Ser Lys Asn His
20 25 30

Ile Ile Ile Ser Ile Ser Thr Pro Cys Ser His Gly Lys Asn Gln Arg
35 40 45

Arg Phe Leu Arg Lys Thr Thr His Phe Arg Ser Ile His
50 55 60

<210> 15

<211> 34

<212> PRT

<213> Setaria italica

<400> 15

Met Val Ala Ala Ala Met Ala Thr Ala Pro Ser Ala Gly Val Pro Pro
1 5 10 15

Leu Arg Gly Thr Arg Gly Pro Ala Arg Phe Arg Ile Arg Gly Val Ser
20 25 30

Val Arg

<210> 16

<211> 27

<212> PRT

<213> Allium cepa

<400> 16

Met Ala Thr Thr Thr Ala Ala Ala Ala Val Thr Ile Ser Ile Pro Lys
1 5 10 15

Lys Pro Val Phe Ile Arg Arg Pro Arg Leu Arg
20 25

<210> 17

<211> 33

<212> PRT

<213> Digitaria sanguinalis

<400> 17

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

Tyr Arg Ser Ala Ser Ala Arg Ala Ser Ser Thr Arg Leu Arg Pro Val
 20 25 30

Leu

<210> 18

<211> 49

<212> PRT

<213> *Amaranthus palmeri*

<400> 18

Met Val Ile Gln Ser Ile Thr His Leu Ser Pro Lys Leu Ala Leu Pro
 1 5 10 15

Ser Pro Leu Ser Ile Ser Ala Lys Asn Tyr Pro Val Ala Val Met Gly
 20 25 30

Asn Ile Ser Glu Arg Glu Glu Pro Thr Ser Ala Lys Arg Val Ala Val
 35 40 45

Val

<210> 19

<211> 48

<212> PRT

<213> *Spinacia oleracea*

<400> 19

Met Val Ile Leu Pro Val Ser Gln Leu Ser Thr Asn Leu Gly Leu Ser
 1 5 10 15

Leu Val Ser Pro Thr Lys Asn Asn Pro Val Met Gly Asn Val Ser Glu
 20 25 30

Arg Asn Gln Val Asn Gln Pro Ile Ser Ala Lys Arg Val Ala Val Val
 35 40 45

<210> 20

<211> 51

<212> PRT

<213> Sedum album

<400> 20

Met Leu Ser Leu Ser Ser Ser His Ser Ser Ala Thr Thr Tyr Ser Leu
1 5 10 15

Arg Gln Arg Tyr Ser Thr Thr Thr Lys Gly Ser Leu Asn Gln Pro Glu
20 25 30

Met Ala Ser Ala Glu Asn Pro Ser Ser Lys Gly Ser Gly Lys Arg Gly
35 40 45

Ala Val Val
50

<210> 21

<211> 38

<212> PRT

<213> Camelina sativa

<400> 21

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

Phe Ser Lys Pro Asn Leu Arg Leu His Val Tyr Lys Pro Leu Lys Leu
20 25 30

Arg Cys Ser Val Ala Gly
35

<210> 22

<211> 43

<212> PRT

<213> Cucumis melo

<400> 22

Met Ala Thr Gly Ala Thr Leu Leu Thr Asp Leu Pro Phe Arg Arg Pro
1 5 10 15

His Pro Leu Thr Leu Leu Arg Pro Ser Asp Ile Pro Ser Phe Tyr Pro
20 25 30

Leu His Ile Ser Leu Gln Asn Asn Arg Leu Arg
35 40

<210> 23

<211> 30
<212> PRT
<213> *Eragrostis tef*

<400> 23

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

Ala Pro Gln Thr Leu Ala Arg Pro Arg Arg Gly Ser Val Arg
20 25 30

<210> 24
<211> 32
<212> PRT
<213> *Canavalia rosea*

<400> 24

Met Tyr Val Ser Pro Ala Ser Asn Asn Pro Arg Ala Cys Leu Lys Leu
1 5 10 15

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

<210> 25
<211> 42
<212> PRT
<213> *Eragrostis tef*

<400> 25

Met Leu Ser Ser Ala Ala Thr Ala Ser Ser Ala Ser Ala His Pro Tyr
1 5 10 15

Arg Pro Ala Ser Ala Arg Ala Ser Arg Ser Val Leu Ala Met Ala Gly
20 25 30

Ser Asp Asp Thr Arg Ala Ala Pro Ala Arg
35 40

<210> 26
<211> 41
<212> PRT
<213> *Andropogon gerardii*

<400> 26

Met Val Ala Ala Thr Ala Met Ala Thr Ala Ala Ser Ala Ala Ala Pro
1 5 10 15

Leu Leu Asn Gly Thr Arg Arg Pro Ala Arg Leu Arg His Arg Gly Leu
20 25 30

Arg Val Arg Cys Ala Ala Val Ala Gly
35 40

<210> 27
<211> 24
<212> PRT
<213> *Digitaria sanguinalis*

<400> 27

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

Tyr Arg Ser Ala Ser Ala Arg Ala
20

<210> 28
<211> 23
<212> PRT
<213> *Eragrostis tef*

<400> 28

Met Leu Ser Ser Ala Ala Thr Ala Ser Ser Ala Ser Ala His Pro Tyr
1 5 10 15

Arg Pro Ala Ser Ala Arg Ala
20

<210> 29
<211> 28
<212> PRT
<213> *Amaranthus graecizans*

<400> 29

Met Ser Ala Met Ala Leu Ser Ser Ser Ile Leu Gln Cys Pro Pro His
1 5 10 15

Ser Asp Ile Ser Phe Arg Phe Phe Ala His Thr Arg
20 25

<210> 30
<211> 13
<212> PRT
<213> *Ambrosia trifida*

<400> 30

Met Ala Ser Pro Thr Ile Val Asp Asn Gln Lys Pro Ala
1 5 10

<210> 31
<211> 14
<212> PRT
<213> Brassica napus

<400> 31

Met Ala Ser Asn Ala Ala Ala Asp His Asp Lys Leu Ser Gly
1 5 10

<210> 32
<211> 13
<212> PRT
<213> Rosa hybrida osiana

<400> 32

Met Ala Ser Pro Ser Pro Gly Asp Lys His Ser Ser Val
1 5 10

<210> 33
<211> 55
<212> PRT
<213> Amaranthus cruentus

<400> 33

Met Lys Gly Arg Lys Arg Arg Ile Thr Arg Glu Ser Ala Arg Glu Met
1 5 10 15

Ser Ala Met Ala Leu Ser Ser Ser Ile Leu Gln Cys Pro Pro His Ser
20 25 30

Asp Ile Ser Phe Arg Phe Ser Ala His Ser Pro Thr His Ser Pro Ile
35 40 45

Phe Phe Gly Arg Pro Arg Lys
50 55

<210> 34
<211> 45
<212> PRT
<213> Taraxacum officinale

<400> 34

Met Thr Tyr Leu Thr Asp Val Gly Ser Leu Asn Cys Tyr Arg Ser Trp
1 5 10 15

Pro Ser Leu Pro Ala Pro Gly Thr Val Gly Ala Leu Thr Ser Lys Asn
20 25 30

Pro Arg Tyr Leu Ile Thr Tyr Gly Pro Ala His Arg Lys
35 40 45

<210> 35
<211> 67
<212> PRT
<213> Canavalia rosea

<400> 35

Met Val Ala Val Phe Asn Asp Val Val Phe Pro Pro Ser Gln Thr Leu
1 5 10 15

Leu Arg Pro Ser Phe His Ser Pro Thr Phe Phe Phe Ser Ser Pro Thr
20 25 30

Pro Lys Phe Thr Arg Thr Arg Pro Asn Arg Ile Leu Arg Cys Ser Ile
35 40 45

Ala Gln Glu Ser Thr Thr Ser Pro Ser Gln Ser Arg Glu Ser Ala Pro
50 55 60

Leu Asp Cys
65

<210> 36
<211> 52
<212> PRT
<213> Eragrostis tef

<400> 36

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

Ala Pro Gln Thr Leu Ala Arg Pro Arg Arg Gly Ser Val Arg Cys Ala
20 25 30

Val Val Ser Asp Ala Ala Glu Ala Pro Ala Ala Pro Gly Ala Arg Leu
35 40 45

Ser Ala Asp Cys
50

<210> 37
 <211> 51
 <212> PRT
 <213> *Allium cepa*

<400> 37

Met Ala Thr Thr Thr Ala Ala Ala Ala Val Thr Ile Ser Ile Pro Lys
 1 5 10 15

Lys Pro Val Phe Ile Arg Arg Pro Arg Leu Arg Cys Ser Ala Val Ala
 20 25 30

Ser Asp Ala Ile Ile Ser Asn Glu Ala Pro Thr Gly Thr Thr Ile Ser
 35 40 45

Ala Asp Cys
 50

<210> 38
 <211> 77
 <212> PRT
 <213> *Taraxacum officinale*

<400> 38

Met Thr Tyr Leu Thr Asp Val Gly Ser Leu Asn Cys Tyr Arg Ser Trp
 1 5 10 15

Pro Ser Leu Pro Ala Pro Gly Thr Val Gly Ala Leu Thr Ser Lys Asn
 20 25 30

Pro Arg Tyr Leu Ile Thr Tyr Gly Pro Ala His Arg Lys Cys Asn Ser
 35 40 45

Trp Arg Phe Arg Cys Ser Ile Ala Lys Asp Ser Pro Ile Thr Pro Pro
 50 55 60

Ile Ser Asn Glu Ser Asn Ser Gln Pro Leu Leu Asp Cys
 65 70 75

<210> 39
 <211> 71
 <212> PRT
 <213> *Cucumis melo*

<400> 39

Met Ala Thr Gly Ala Thr Leu Leu Thr Asp Leu Pro Phe Arg Arg Pro

Phe Ser Lys Pro Asn Leu Arg Leu His Val Tyr Lys Pro Leu Lys Leu
20 25 30

Arg Cys Ser Val Ala Gly
35

<210> 42
<211> 45
<212> PRT
<213> Taraxacum officinale

<400> 42

Met Thr Tyr Leu Thr Asp Val Gly Ser Leu Asn Cys Tyr Arg Ser Trp
1 5 10 15

Pro Ser Leu Pro Ala Pro Gly Thr Val Gly Ala Leu Thr Ser Lys Asn
20 25 30

Pro Arg Tyr Leu Ile Thr Tyr Gly Pro Ala His Arg Lys
35 40 45

<210> 43
<211> 65
<212> PRT
<213> Taraxacum officinale

<400> 43

Met Thr Tyr Leu Thr Asp Val Gly Ser Leu Asn Cys Tyr Arg Ser Trp
1 5 10 15

Pro Ser Leu Pro Ala Pro Gly Thr Val Gly Ala Leu Thr Ser Lys Asn
20 25 30

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

Ile Thr Pro Pro Ile Ser Asn Glu Ser Asn Ser Gln Pro Leu Leu Asp
50 55 60

Cys
65

<210> 44
<211> 60
<212> PRT
<213> Senna obtusifolia

<400> 44

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

Pro Thr Gln Thr Leu Val His Ser Asn Thr Ser Phe Ile Ser Ser Arg
20 25 30

Thr Cys Phe Val Cys Pro Ile Ile Pro Phe Pro Ser Arg Ser Gln Leu
35 40 45

Asn Arg Arg Ile Ala Cys Ile Arg Ser Asn Val Arg
50 55 60

<210> 45
<211> 19
<212> PRT
<213> Eragrostis tef

<400> 45

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

Ala Asp Cys

<210> 46
<211> 25
<212> PRT
<213> Allium cepa

<400> 46

Met Ala Thr Thr Thr Ala Ser Asp Ala Ile Ile Ser Asn Glu Ala Pro
1 5 10 15

Thr Gly Thr Thr Ile Ser Ala Asp Cys
20 25

<210> 47
<211> 12
<212> PRT
<213> Allium cepa

<400> 47

Met Ala Thr Thr Gly Thr Thr Ile Ser Ala Asp Cys
1 5 10

<210> 48
<211> 19

<212> PRT
 <213> Cucumis melo

<400> 48

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

Leu Asp Cys

<210> 49
 <211> 32
 <212> PRT
 <213> Xanthium strumarium

<400> 49

Met Ser Ser Leu Thr Asp Leu Pro Ser Leu Asn His Tyr Arg Thr Cys
 1 5 10 15

Ser Pro Pro Ile Ser Ser Asp Ser Thr Ser His Pro Phe Leu Asp Cys
 20 25 30

<210> 50
 <211> 204
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 50
 atggccaccg ccaccactac cgccaccgct gcgttctccg gcgtgggtgag cgtcggcact 60
 gagacgcgca ggatctactc cttcagccac ctccagcctt ctgctgcggt ccccgctaag 120
 ccgtcttcgt tcaagagcct gaagctgaaa cagtccgcac gccttaccgc gcgcctggac 180
 cataggccat tcgttgctcag gtgc 204

<210> 51
 <211> 204
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 51
 atggcgacgg ctacgacgac tgctacggcg gcgttttagtg gtgtagtcag tgtaggaacg 60
 gagactcgaa ggattttattc gttttctcat cttcaacctt ctgcggcctt tccggcggaag 120

cctagttcct tcaaattctct caaattaaag cagagcgcgga ggctcacacg gcggcttgat 180
catcggccgt tcgttggtccg atgt 204

<210> 52
<211> 204
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 52
atggcttcct ccacgacgac tgctacggcg gcgttttagtg gtgtagtcag tgtaggaacg 60
gagactcgaa ggattttattc gttttctcat cttcaacctt ctgcggcttt tccggcgaaag 120
cctagttcct tcaaattctct caaattaaag cagagcgcgga ggctcacacg gcggcttgat 180
catcggccgt tcgttggtccg atgt 204

<210> 53
<211> 105
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 53
atgcttaaca ttgcgccgag ttgtgttttg gccagcggga tctctaagcc cgtgaccaag 60
atggctagca cggagaacaa ggacgaccac agcagcgcca agagg 105

<210> 54
<211> 135
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 54
atgggtggctg tgttcaacga cgtagtggtc cctccttcgc agacccttct tcgcccctcc 60
ttccacagcc cgacgttctt ttttagcagc cccacaccaa agttcacgcg tacgaggccg 120
aatagaatac tgcgg 135

<210> 55
<211> 180
<212> ADN
<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 55

atgccggcga tagcaatggc ttctttaact gatctgcccgt cgttgagccc cacacagacc 60
ctcgttcact cgaacacgag cttcatttca tcgagaacct gcttcgtctg tccgatcatc 120
cccttcccat cgaggtcgca actgaaccgc cgcacgcct gcacaggtc caacgtaagg 180

<210> 56

<211> 165

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 56

atgaccacaa ccccggtagc aaaccacccc aatatcttca ctcaccgaag ccctccgtca 60
tcttcctcgt cctcaccag cgcgtttctg acccgcacct cttttctgcc cttctctagc 120
atctgcaaaa ggaactctgt gaactgcaat gggtggcgaa cccgg 165

<210> 57

<211> 102

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 57

atggacttca gtctccttag gcccgcttcg acgcagccgt tcctctcacc cttctccaat 60
cccttccac ggagtaggcc atacaagcca cttaatctga gg 102

<210> 58

<211> 111

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 58

atggccatct tgattgacct ctccctcctg aggtcctctc cgtcgggtctt ctccttctcc 60
aagccgaacc acaggatacc accgcggatc tacaagccgt tcaagttgag g 111

<210> 59

<211> 141

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 59

atgaccacgc tttccaggct cgctgacctt ccttcttttg ctgccctcc tcctctcttg 60

accacccggc cccctccttc agttttcctg actccgaagc cgacaaagcc gtcacctcca 120

catcacttct ttaaactgcg c 141

<210> 60

<211> 138

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 60

atgtcgtccc taacggacct cccctccctg aatcactata ggacgtgcag cccgcgccca 60

ttccccatct ccaggcagac cagttcatca attaaccxaa acaacttgac gaccagtaac 120

cgttggcgca ggttcagg 138

<210> 61

<211> 105

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 61

atgacgagtc tcaccaactt cacccegtc aagctgacga accccaacta cctcaacacg 60

accaccacct acaaccaccg taagctctcc aacttcgggt tccgc 105

<210> 62

<211> 135

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 62

atgtcggcca tggcgctgtc cagcagcatt ctacagtgcc cgcctcactc agacatatcc 60

ttccgcttct cggcatacac tgccaccgc tcacctttct tcttcgggag gccaaaggaaa 120

ctatcttaca tccac 135

<210> 63

<211> 141

<212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 63
 atgtcggcca tggcattgag ctccaccatg gccctcagcc tgccacaatc tagcatgtcc 60
 ttgagccact gcagacacaa tagaataact attctgatcc cctcgagctc gttacggcga 120
 cggggaggtt cctcgatccg c 141

<210> 64
 <211> 183
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 64
 atgtctgcta tggcgagccc ctccatcatc ccgcagtcgt tcctccagcg aagcccgacc 60
 tccttgcaat ctcgatccaa ctactcgaag aaccacatca tcctctccat cagcaccgcc 120
 tgctctcatg ggaagaacca gcgacgtttc ttgcgaaaga ccaccactt ccgatccatc 180
 cac 183

<210> 65
 <211> 102
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 65
 atggtcgccg ctgcaatggc tacagccct tccgtggag tccctcctct tagagggaca 60
 aggggtccag caaggtttag aatccgggga gtgtcagtgc gt 102

<210> 66
 <211> 81
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 66
 atggccacta ccacagcagc cgcggcggtc accatcagca ttctaaaaa gcctgttttt 60
 atccgccgcc cacgacttcg t 81

<210> 67
 <211> 99
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 67
 atgttgtcta gcactgctac tgcaagttct gcatcctcac accacccta ccgttcagct 60
 tctgcaaggg cttcgtcgac acgtctccgc ccggtcctt 99

<210> 68
 <211> 147
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 68
 atggtcattc agtcaattac gcatctttct cccaagctcg cactgccctc tccgctgtcg 60
 atctcggcta agaactaccc ggtggccgtg atggggaata tcagcgagag ggaggagcca 120
 acttctgcta aaaggggtggc cgtgggtg 147

<210> 69
 <211> 144
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 69
 atggtcattc taccctgtgc ccagctctcg actaatttgg ggctttccct tgtagtcca 60
 acgaagaaca acccgtgat gggcaacgtg tccgagagga accaggtgaa ccagccaatc 120
 tccgccaagc gcgttgctgt cgtg 144

<210> 70
 <211> 153
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 70
 atgtctctcac tgagcagctc cactcatcc gcgacaacgt attctctccg gcaacggtac 60
 tctacaacga ccaaagggttc gttgaaccag cctgagatgg ccagcgccga aaacccttcc 120

agcaagggat caggtaagag aggagcagtg gtg

153

<210> 71
 <211> 114
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 71
 atggagctga gcctcctaag accgtctact cagtcattgc tcccctcgtt cagcaagcct 60
 aatttgcggc tccacgtgta caagcccctt aagctccgat gcagcgtagc cggc 114

<210> 72
 <211> 129
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 72
 atggcgacag gagccaccct gctaacagac ctgccgttcc gtaggccgca cccgcttacg 60
 ctcttacgtc cgagcgatat cccgtccttt taccactac acataagcct acagaacaat 120
 cgtttgagg 129

<210> 73
 <211> 90
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 73
 atgggtggctg ctgcggcaac gatggctacc gccgcaccac cattaagagc gcctcaaact 60
 cttgcacgac cgccaagagg tagtgtgaga 90

<210> 74
 <211> 96
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 74
 atgtatgtgt cgccgcctc gaacaacca cgagcatgcc tcaagctgtc acaggaaatg 60

gcgtcttcag cagcagacgg caaccaaga tccgtt

96

<210> 75
 <211> 126
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 75
 atgttgctta gcgcagcgac agctagcagc gcaagtgctc atccttatcg acctgcttct 60
 gcccgggcga gtaggagcgt gttggctatg gctggatcag acgatactag ggcagctcct 120
 gcccg 126

<210> 76
 <211> 123
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 76
 atggtggctg cgaccgcaat ggccaccgct gcttcggctg ctgcgcctct cctaaacgga 60
 acgagacgac cggcacgatt gagacataga gggtttacgtg ttaggtgtgc tgcagtagca 120
 gga 123

<210> 77
 <211> 72
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 77
 atgctttcta gcactgccac agcttctca gcttctagcc accacccgta tcgttcagct 60
 tcggcacgtg cc 72

<210> 78
 <211> 69
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 78
 atgcttagct cagcagctac ggctctagt gcttctgccc atccataccg tcccgcattct 60

gctcgagca

69

<210> 79
<211> 84
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 79
atgagcgcgga tggcgctttc ttctagcatc ttgcaatgcc ccccccactc tgacatttct 60
ttccgcttct tcgcccacac tcgc 84

<210> 80
<211> 39
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 80
atggcgagtc ccacgatcgt tgacaaccag aagccagcg 39

<210> 81
<211> 42
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 81
atggctagta acgccgtgc tgaccacgat aagctctcgg gt 42

<210> 82
<211> 39
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 82
atggcgctgc cgtccccagg cgacaaacat tcgtctgta 39

<210> 83
<211> 165
<212> ADN
<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 83

atgaaggggc ggaagagacg gatcacgcgg gagtctgcaa gggagatgtc agcgatggca 60

ttgtcttcga gcatactcca gtgccctcct cactccgaca tctctttccg ttttagcgct 120

cactcaccga cacacagccc tatcttcttt gggcgtccca ggaaa 165

<210> 84

<211> 135

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 84

atgacctacc tcactgatgt gggtagtctc aattgctaca ggtcctggcc tagcctaccg 60

gcccctggga cggtcggagc attgacttct aagaaccccc gctacttgat cacatacggt 120

ccggctcacc gaaag 135

<210> 85

<211> 201

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 85

atgggtggctg tgttcaacga cgtagtggtc cctccttcgc agacccttct tcgcccctcc 60

ttccacagcc cgacgttctt ttttagcagc cccacaccaa agttcacgcg tacgaggccg 120

aatagaatac tgcggtgctc gattgcgagc gagtctacaa catcgccgctc gcagtcgcga 180

gagtcagctc cactcgattg t 201

<210> 86

<211> 156

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 86

atgggtggctg ctgcggcaac gatggctacc gccgcaccac cattaagagc gcctcaaact 60

cttgacagac cgcaagagg tagtgtgaga tgtgccgctc ttagcgatgc tgcagaagct 120

ccggctgctc ctggcgctag actctctgca gattgc 156

<210> 87
 <211> 153
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 87
 atggccacta ccacagcagc cgcggcggtc accatcagca ttcctaaaaa gcctgttttt 60
 atccgccgcc cagcacttcg ttgctcggca gttgcatccg acgcaatcat ctccaacgag 120
 gccctacag ggacgacaat ctcggctgac tgt 153

<210> 88
 <211> 231
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 88
 atgacctacc tcactgatgt gggtagtctc aattgctaca ggtcctggcc tagcctaccg 60
 gccctggga cggtcggagc attgacttct aagaaccccc gctacttgat cacatacggc 120
 ccggctcacc gaaagtgcaa cagctggcgc ttccgggtgct ctattgcaaa ggactcccc 180
 atcacgcccc caatttcgaa cgagagcaat tcacagcccc tgctagactg c 231

<210> 89
 <211> 213
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 89
 atggcgacag gagccaccct gctaacagac ctgccgttcc gtaggccgca cccgcttacg 60
 ctcttacgtc cgagcgatat cccgtccttt taccactac acataagcct acagaacaat 120
 cgtttgagga gtcatttcag gtgctcaatc gccgagggct cgacggcact gagcccatct 180
 aacgcatcgt cgcaatcgag tatcttggac tgc 213

<210> 90
 <211> 213
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 90

atgtcgtccc taacggacct cccctccctg aatcactata ggacgtgcag cccgcgccc 60

ttccccatct ccaggcagac cagttcatca attaaccxaa acaacttgac gaccagtaac 120

cgttggcgca ggttcaggtg ctctattgag aacgacaccc cgatcagccc gccgatttcc 180

agcgactcta cttcccaccc tttcttggac tgt 213

<210> 91

<211> 114

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 91

atggagctaa gcctcctaag accgtctact cagtcattgc tcccctcggt cagcaagcct 60

aatttgcggc tccacgtgta caagcccctt aagctccgat gcagcgtagc cggt 114

<210> 92

<211> 135

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 92

atgacctacc tcaactgatgt gggtagtctc aattgctaca ggtcttggcc tagcctaccg 60

gcccctggga cggtcggagc attgacttct aagaaccccc gctacttgat cacatacggt 120

ccggctcacc gaaag 135

<210> 93

<211> 195

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 93

atgacctacc tcaactgatgt gggtagtctc aattgctaca ggtcctggcc tagcctaccg 60

gcccctggga cggtcggagc attgacttct aagaaccccc gctacttgat cacatacggt 120

ccggctcacc gaaaggactc ccccatcacg cccccaattt cgaacgagag caattcacag 180

cccctgctag actgc 195

<210> 94
 <211> 180
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 94
 atgccggcga tagcaatagc ttctttaact gatctgccgt cggtgagccc cacacagacc 60
 ctcgttcact cgaacacgag cttcatttca tcgagaacct gcttcgtctg tccgatcatc 120
 cccttcccat cgaggtcgca actgaaccgc cgcatacgct gcatcaggtc caacgtaagg 180

<210> 95
 <211> 57
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 95
 atggtggctg ctgcggaagc tccggctgct cctggcgcta gactctctgc agattgc 57

<210> 96
 <211> 75
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 96
 atggccacta ccacagcatc cgacgcaatc atctccaacg aggcccctac agggacgaca 60
 atctcggctg actgt 75

<210> 97
 <211> 36
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 97
 atggccacta cagggacgac aatctcggct gactgt 36

<210> 98
 <211> 57
 <212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 98

atggcgacgg cactgagccc atctaacgca tcgtcgcaat cgagtatctt ggactgc

57

<210> 99

<211> 96

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 99

atgtcgtccc taacggacct cccctccctg aatcactata ggacgtgcag cccgccgatt

60

tccagcgact ctacttccca ccctttcttg gactgt

96

<210> 100

<211> 179

<212> PRT

<213> Enterobacter cloacae

<400> 100

Met Lys Ala Leu Val Leu Tyr Ser Thr Arg Asp Gly Gln Thr His Ala
1 5 10 15

Ile Ala Ser Tyr Ile Ala Ser Cys Met Lys Glu Lys Ala Glu Cys Asp
20 25 30

Val Ile Asp Leu Thr His Gly Glu His Val Asn Leu Thr Gln Tyr Asp
35 40 45

Gln Val Leu Ile Gly Ala Ser Ile Arg Tyr Gly His Phe Asn Ala Val
50 55 60

Leu Asp Lys Phe Ile Lys Arg Asn Val Asp Gln Leu Asn Asn Met Pro
65 70 75 80

Ser Ala Phe Phe Cys Val Asn Leu Thr Ala Arg Lys Pro Glu Lys Arg
85 90 95

Thr Pro Gln Thr Asn Pro Tyr Val Arg Lys Phe Leu Leu Ala Thr Pro
100 105 110

Trp Gln Pro Ala Leu Cys Gly Val Phe Ala Gly Ala Leu Arg Tyr Pro

```

115                      120                      125

Arg Tyr Arg Trp Ile Asp Lys Val Met Ile Gln Leu Ile Met Arg Met
130                      135                      140

Thr Gly Gly Glu Thr Asp Thr Ser Lys Glu Val Glu Tyr Thr Asp Trp
145                      150                      155                      160

Glu Gln Val Lys Lys Phe Ala Glu Asp Phe Ala Lys Leu Ser Tyr Lys
165                      170                      175

Lys Ala Leu

<210> 101
<211> 178
<212> PRT
<213> Pantoea ananatis

<400> 101

Met Lys Ala Leu Ile Leu Phe Ser Thr Arg Asp Gly Gln Thr Gln Lys
1      5      10      15

Ile Ala Ser Ala Ile Ala Asp Glu Ile Lys Gly Gln Gln Ser Cys Asp
20      25      30

Val Ile Asn Ile Gln Asp Ala Lys Thr Leu Asp Trp Gln Gln Tyr Asp
35      40      45

Arg Val Leu Ile Gly Ala Ser Ile Arg Tyr Gly His Phe Gln Pro Val
50      55      60

Val Asn Glu Phe Val Lys His Asn Leu Leu Ala Leu Gln Gln Arg Val
65      70      75      80

Ser Gly Phe Phe Ser Val Asn Leu Thr Ala Arg Lys Pro Glu Lys Arg
85      90      95

Ser Pro Glu Thr Asn Ala Tyr Thr Val Lys Phe Leu Ala Gln Ser Pro
100     105     110

Trp Gln Pro Asp Cys Cys Ala Val Phe Ala Gly Ala Leu Tyr Tyr Pro
115     120     125

Arg Tyr Arg Trp Phe Asp Arg Val Met Ile Gln Phe Ile Met Arg Met

```

130 135 140
 Thr Gly Gly Glu Thr Asp Ala Ser Lys Glu Val Glu Tyr Thr Asp Trp
 145 150 155 160
 Gln Gln Val Gln Arg Phe Ala Arg Asp Phe Ala Gln Leu Pro Gly Lys
 165 170 175
 Ser Tyr
 <210> 102
 <211> 177
 <212> PRT
 <213> Pantoea stewardii
 <400> 102
 Met Lys Ala Leu Ile Leu Tyr Ser Thr Arg Asp Gly Gln Thr Arg Lys
 1 5 10 15
 Ile Ala Ser Ser Ile Ala Asp Val Ile Arg Gln Gln Gln Gln Cys Asp
 20 25 30
 Val Leu Asn Ile Lys Asp Ala Ser Leu Pro Asp Trp Ala Gln Tyr Asp
 35 40 45
 Arg Val Leu Ile Gly Ala Ser Ile Arg Tyr Gly His Phe Gln Pro Val
 50 55 60
 Val Asp Lys Phe Val Lys Gln His Leu His Glu Leu Gln Gln Arg Thr
 65 70 75 80
 Ser Gly Phe Phe Ser Val Asn Leu Thr Ala Arg Lys Pro Glu Lys Arg
 85 90 95
 Ser Pro Glu Thr Asn Ala Tyr Thr Gln Lys Phe Leu Ala His Ser Pro
 100 105 110
 Trp Gln Pro Asp Cys Cys Ala Val Phe Ala Gly Ala Leu Tyr Tyr Pro
 115 120 125
 Arg Tyr Arg Trp Phe Asp Arg Val Met Ile Gln Leu Ile Met Arg Met
 130 135 140
 Thr Gly Gly Glu Thr Asp Ser Thr Lys Glu Val Glu Tyr Thr Asp Trp

Ser

<400> 103

Trp Glu Gln Val Ala Asn Phe Ala Arg Glu Ile Ala His Leu Thr Asp

165

170

175

Lys Pro Thr Leu Lys
180

<210> 104
<211> 178
<212> PRT
<213> Erwinia toletana

<400> 104

Met Lys Ala Leu Ile Leu Phe Ser Ser Arg Glu Gly Gln Thr Arg Glu
1 5 10 15

Ile Ala Ser Tyr Ile Ala Asn Ser Ile Lys Glu Glu Met Glu Cys Asp
20 25 30

Val Phe Asn Ile Leu Arg Val Glu Gln Ile Asp Trp Ser Gln Tyr Asp
35 40 45

Arg Val Leu Ile Gly Gly Ser Ile His Tyr Gly His Phe His Pro Ala
50 55 60

Val Ala Lys Phe Val Lys Arg His Leu His Glu Leu Gln Gln Arg Ser
65 70 75 80

Ser Gly Phe Phe Cys Val Asn Leu Thr Ala Arg Lys Ala Asp Lys Arg
85 90 95

Thr Pro Gln Thr Asn Ala Tyr Met Arg Lys Phe Leu Leu Gln Ser Pro
100 105 110

Trp Gln Pro Asp Cys Cys Ala Val Phe Ala Gly Ala Leu Arg Tyr Thr
115 120 125

Arg Tyr Arg Trp Phe Asp Arg Val Met Ile Gln Leu Ile Met Arg Met
130 135 140

Thr Gly Gly Glu Thr Asp Thr Ser Lys Glu Val Glu Tyr Thr Asp Trp
145 150 155 160

Thr Gln Val Ala Arg Phe Ala Gln Glu Phe Ala His Leu Pro Gly Lys
165 170 175

Thr Gln

<210> 105
 <211> 179
 <212> PRT
 <213> Pectobacterium carotovorum

<400> 105

Met Lys Ala Leu Ile Val Phe Ser Ser Arg Asp Gly Gln Thr Arg Ala
 1 5 10 15

Ile Ala Ser Tyr Ile Ala Asn Thr Leu Lys Gly Thr Leu Glu Cys Asp
 20 25 30

Val Val Asn Val Leu Asn Ala Asn Asp Ile Asp Leu Ser Gln Tyr Asp
 35 40 45

Arg Val Ala Ile Gly Ala Ser Ile Arg Tyr Gly Arg Phe His Pro Ala
 50 55 60

Val Asn Gln Phe Ile Arg Lys His Leu Thr Ser Leu Gln Gln Leu Pro
 65 70 75 80

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

Thr Ile Gln Thr Asn Ala Tyr Thr Arg Lys Phe Leu Leu Asn Ser Pro
 100 105 110

Trp Gln Pro Asp Leu Cys Cys Val Phe Ala Gly Ala Leu Arg Tyr Pro
 115 120 125

Arg Tyr Arg Trp Phe Asp Arg Val Met Ile Gln Leu Ile Met Arg Ile
 130 135 140

Thr Gly Gly Glu Thr Asp Ser Thr Lys Glu Ile Glu Tyr Thr Asp Trp
 145 150 155 160

Gln Gln Val Ala Arg Phe Ala Gln Asp Phe Ala Gln Leu Ala Ala Lys
 165 170 175

Asn Pro Ala

<210> 106

<211> 179
 <212> PRT
 <213> Shimwellia blattae

<400> 106

Met Lys Thr Leu Ile Leu Phe Ser Thr Arg Asp Gly Gln Thr His Lys
 1 5 10 15

Ile Ala Arg His Ile Ala Gly Val Leu Glu Glu Gln Gly Lys Ala Cys
 20 25 30

Glu Leu Val Asp Leu Leu Gln Pro Gly Glu Pro Asp Trp Ser Thr Val
 35 40 45

Glu Cys Val Val Leu Gly Ala Ser Ile Arg Tyr Gly His Phe His Lys
 50 55 60

Ser Phe Ile Arg Phe Val Asn Thr His Ala Gln Arg Leu Asn Asn Met
 65 70 75 80

Pro Gly Ala Leu Phe Thr Val Asn Leu Val Ala Arg Lys Pro Glu Lys
 85 90 95

Gln Ser Pro Gln Thr Asn Ser Tyr Thr Arg Lys Phe Leu Ala Ala Ser
 100 105 110

Pro Trp Gln Pro Gln Arg Cys Gln Val Phe Ala Gly Ala Leu Arg Tyr
 115 120 125

Pro Arg Tyr Ser Trp Tyr Asp Arg Met Met Ile Arg Leu Ile Met Lys
 130 135 140

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

Trp Gln Ser Val Thr Arg Phe Ala Arg Glu Ile Ala Gln Leu Pro Gly
 165 170 175

Glu Thr Arg

<210> 107
 <211> 178
 <212> PRT
 <213> Pantoea stewardii

<400> 107

Met Lys Ala Leu Ile Leu Phe Ser Ser Arg Asp Gly Gln Thr Gln Leu
 1 5 10 15

Ile Ala Ser Ser Ile Ala Lys Glu Leu Glu Gly Lys Gln Ala Cys Asp
 20 25 30

Val Leu Asn Ile Leu Asp Thr Thr Asn Val Glu Trp Thr Gln Tyr Asp
 35 40 45

Arg Val Leu Ile Gly Ala Ser Ile Arg Tyr Gly His Phe His Pro Ala
 50 55 60

Val Ala Glu Phe Val Lys Arg His Gln Arg Glu Leu Gln Gln Arg Ser
 65 70 75 80

Ser Gly Phe Phe Ser Val Asn Leu Thr Ala Arg Lys Pro Glu Lys Arg
 85 90 95

Ser Pro Glu Thr Asn Ala Tyr Thr Ala Lys Phe Leu Asn Gln Ser Pro
 100 105 110

Trp Gln Pro Asp Cys Cys Ala Val Phe Ala Gly Ala Leu Arg Tyr Pro
 115 120 125

Arg Tyr Arg Trp Phe Asp Arg Ile Met Ile Gln Leu Ile Met Arg Met
 130 135 140

Thr Gly Gly Glu Thr Asp Ser Ser Lys Glu Val Glu Tyr Thr Asp Trp
 145 150 155 160

Gln Gln Val Thr Arg Phe Ala Gln Glu Phe Ala Arg Leu Pro Gly Lys
 165 170 175

Thr Ser

<210> 108

<211> 180

<212> PRT

<213> Enterobacter cloacae

<400> 108

Met Lys Thr Leu Ile Leu Phe Ser Thr Arg Asp Gly Gln Thr Arg Glu
 1 5 10 15

Ile Ala Ala Phe Leu Ala Ser Glu Leu Lys Glu Gln Gly Ile Tyr Ala
 20 25 30

Asp Val Ile Asn Leu Asn Arg Thr Glu Glu Ile Ala Trp Gln Glu Tyr
 35 40 45

Asp Arg Val Val Ile Gly Ala Ser Ile Arg Tyr Gly His Phe His Pro
 50 55 60

Ala Val Asp Arg Phe Val Lys Lys His Thr Glu Thr Leu Asn Ser Leu
 65 70 75 80

Pro Gly Ala Phe Phe Ser Val Asn Leu Val Ala Arg Lys Ala Glu Lys
 85 90 95

Arg Thr Pro Gln Thr Asn Ser Tyr Thr Arg Lys Phe Leu Leu Asn Ser
 100 105 110

Pro Trp Lys Pro Ala Ala Cys Ala Val Phe Ala Gly Ala Leu Arg Tyr
 115 120 125

Pro Arg Tyr Arg Trp Tyr Asp Arg Phe Met Ile Arg Leu Ile Met Lys
 130 135 140

Met Thr Gly Gly Glu Thr Asp Thr Arg Lys Glu Val Val Tyr Thr Asp
 145 150 155 160

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

Ser Ser Arg Leu
 180

<210> 109
 <211> 177
 <212> PRT
 <213> Enterobacter mori

<400> 109

Met Lys Ile Leu Ile Leu Phe Ser Thr Arg Asp Gly Gln Thr Arg Glu
 1 5 10 15

Ile Ala Ala Ser Leu Ala Ser Glu Leu Lys Glu Gln Ala Phe Asp Val
 20 25 30

Asp Val Val Asn Leu His Arg Ala Glu Asn Ile Ala Trp Glu Glu Tyr
35 40 45

Asp Gly Val Val Ile Gly Ala Ser Ile Arg Tyr Gly His Phe His Ser
50 55 60

Thr Leu Asn Ser Phe Val Lys Lys His Gln Gln Ala Leu Lys Lys Leu
65 70 75 80

Pro Gly Ala Phe Tyr Ser Val Asn Leu Val Ala Arg Lys Pro Glu Lys
85 90 95

Arg Thr Pro Gln Thr Asn Ser Tyr Thr Arg Lys Phe Leu Leu Asp Ser
100 105 110

Pro Trp Gln Pro Asp Leu Ser Ala Val Phe Ala Gly Ala Leu Arg Tyr
115 120 125

Pro Arg Tyr Asn Trp Tyr Asp Arg Ile Met Ile Arg Leu Ile Met Lys
130 135 140

Ile Thr Gly Gly Glu Thr Asp Thr Arg Lys Glu Val Val Tyr Thr Asp
145 150 155 160

Trp Gln Gln Val Thr His Phe Ala His Glu Ile Val Gln Leu Val Arg
165 170 175

Lys

<210> 110
<211> 178
<212> PRT
<213> Enterobacter cloacae

<400> 110

Lys Ala Leu Val Leu Tyr Ser Thr Arg Asp Gly Gln Thr His Ala Ile
1 5 10 15

Ala Ser Tyr Ile Ala Ser Cys Met Lys Glu Lys Ala Glu Cys Asp Val
20 25 30

Ile Asp Leu Thr His Gly Glu His Val Asn Leu Thr Gln Tyr Asp Gln
35 40 45

Val Leu Ile Gly Ala Ser Ile Arg Tyr Gly His Phe Asn Ala Val Leu
50 55 60

Asp Lys Phe Ile Lys Arg Asn Val Asp Gln Leu Asn Asn Met Pro Ser
65 70 75 80

Ala Phe Phe Cys Val Asn Leu Thr Ala Arg Lys Pro Glu Lys Arg Thr
85 90 95

Pro Gln Thr Asn Pro Tyr Val Arg Lys Phe Leu Leu Ala Thr Pro Trp
100 105 110

Gln Pro Ala Leu Cys Gly Val Phe Ala Gly Ala Leu Arg Tyr Pro Arg
115 120 125

Tyr Arg Trp Ile Asp Lys Val Met Ile Gln Leu Ile Met Arg Met Thr
130 135 140

Gly Gly Glu Thr Asp Thr Ser Lys Glu Val Glu Tyr Thr Asp Trp Glu
145 150 155 160

Gln Val Lys Lys Phe Ala Glu Asp Phe Ala Lys Leu Ser Tyr Lys Lys
165 170 175

Ala Leu

<210> 111
<211> 177
<212> PRT
<213> Pantoea ananatis

<400> 111

Lys Ala Leu Ile Leu Phe Ser Thr Arg Asp Gly Gln Thr Gln Lys Ile
1 5 10 15

Ala Ser Ala Ile Ala Asp Glu Ile Lys Gly Gln Gln Ser Cys Asp Val
20 25 30

Ile Asn Ile Gln Asp Ala Lys Thr Leu Asp Trp Gln Gln Tyr Asp Arg
35 40 45

Val Leu Ile Gly Ala Ser Ile Arg Tyr Gly His Phe Gln Pro Val Val
50 55 60

Asn Glu Phe Val Lys His Asn Leu Leu Ala Leu Gln Gln Arg Val Ser
65 70 75 80

Gly Phe Phe Ser Val Asn Leu Thr Ala Arg Lys Pro Glu Lys Arg Ser
85 90 95

Pro Glu Thr Asn Ala Tyr Thr Val Lys Phe Leu Ala Gln Ser Pro Trp
100 105 110

Gln Pro Asp Cys Cys Ala Val Phe Ala Gly Ala Leu Tyr Tyr Pro Arg
115 120 125

Tyr Arg Trp Phe Asp Arg Val Met Ile Gln Phe Ile Met Arg Met Thr
130 135 140

Gly Gly Glu Thr Asp Ala Ser Lys Glu Val Glu Tyr Thr Asp Trp Gln
145 150 155 160

Gln Val Gln Arg Phe Ala Arg Asp Phe Ala Gln Leu Pro Gly Lys Ser
165 170 175

Tyr

<210> 112
<211> 180
<212> PRT
<213> Escherichia coli

<400> 112

Lys Thr Leu Ile Leu Phe Ser Thr Arg Asp Gly Gln Thr Arg Glu Ile
1 5 10 15

Ala Ser Tyr Leu Ala Ser Glu Leu Lys Glu Leu Gly Ile Gln Ala Asp
20 25 30

Val Ala Asn Val His Arg Ile Glu Glu Pro Gln Trp Glu Asn Tyr Asp
35 40 45

Arg Val Val Ile Gly Ala Ser Ile Arg Tyr Gly His Tyr His Ser Ala
50 55 60

Phe Gln Glu Phe Val Lys Lys His Ala Thr Arg Leu Asn Ser Met Pro
65 70 75 80

Ser Ala Phe Tyr Ser Val Asn Leu Val Ala Arg Lys Pro Glu Lys Arg
85 90 95

Thr Pro Gln Thr Asn Ser Tyr Ala Arg Lys Phe Leu Met Asn Ser Gln
100 105 110

Trp Arg Pro Asp Arg Cys Ala Val Ile Ala Gly Ala Leu Arg Tyr Pro
115 120 125

Arg Tyr Arg Trp Tyr Asp Arg Phe Met Ile Lys Leu Ile Met Lys Met
130 135 140

Ser Gly Gly Glu Thr Asp Thr Arg Lys Glu Val Val Tyr Thr Asp Trp
145 150 155 160

Glu Gln Val Ala Asn Phe Ala Arg Glu Ile Ala His Leu Thr Asp Lys
165 170 175

Pro Thr Leu Lys
180

<210> 113

<211> 177

<212> PRT

<213> Erwinia toletana

<400> 113

Lys Ala Leu Ile Leu Phe Ser Ser Arg Glu Gly Gln Thr Arg Glu Ile
1 5 10 15

Ala Ser Tyr Ile Ala Asn Ser Ile Lys Glu Glu Met Glu Cys Asp Val
20 25 30

Phe Asn Ile Leu Arg Val Glu Gln Ile Asp Trp Ser Gln Tyr Asp Arg
35 40 45

Val Leu Ile Gly Gly Ser Ile His Tyr Gly His Phe His Pro Ala Val
50 55 60

Ala Lys Phe Val Lys Arg His Leu His Glu Leu Gln Gln Arg Ser Ser
65 70 75 80

Gly Phe Phe Cys Val Asn Leu Thr Ala Arg Lys Ala Asp Lys Arg Thr
85 90 95

Pro Gln Thr Asn Ala Tyr Met Arg Lys Phe Leu Leu Gln Ser Pro Trp
100 105 110

Gln Pro Asp Cys Cys Ala Val Phe Ala Gly Ala Leu Arg Tyr Thr Arg
115 120 125

Tyr Arg Trp Phe Asp Arg Val Met Ile Gln Leu Ile Met Arg Met Thr
130 135 140

Gly Gly Glu Thr Asp Thr Ser Lys Glu Val Glu Tyr Thr Asp Trp Thr
145 150 155 160

Gln Val Ala Arg Phe Ala Gln Glu Phe Ala His Leu Pro Gly Lys Thr
165 170 175

Gln

<210> 114

<211> 178

<212> PRT

<213> Pectobacterium carotovorum

<400> 114

Lys Ala Leu Ile Val Phe Ser Ser Arg Asp Gly Gln Thr Arg Ala Ile
1 5 10 15

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

Val Asn Val Leu Asn Ala Asn Asp Ile Asp Leu Ser Gln Tyr Asp Arg
35 40 45

Val Ala Ile Gly Ala Ser Ile Arg Tyr Gly Arg Phe His Pro Ala Val
50 55 60

Asn Gln Phe Ile Arg Lys His Leu Thr Ser Leu Gln Gln Leu Pro Ser
65 70 75 80

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

Ile Gln Thr Asn Ala Tyr Thr Arg Lys Phe Leu Leu Asn Ser Pro Trp
100 105 110

Gln Pro Asp Leu Cys Cys Val Phe Ala Gly Ala Leu Arg Tyr Pro Arg
 115 120 125

Tyr Arg Trp Phe Asp Arg Val Met Ile Gln Leu Ile Met Arg Ile Thr
 130 135 140

Gly Gly Glu Thr Asp Ser Thr Lys Glu Ile Glu Tyr Thr Asp Trp Gln
 145 150 155 160

Gln Val Ala Arg Phe Ala Gln Asp Phe Ala Gln Leu Ala Ala Lys Asn
 165 170 175

Pro Ala

<210> 115
 <211> 178
 <212> PRT
 <213> Shimwellia blattae

<400> 115

Lys Thr Leu Ile Leu Phe Ser Thr Arg Asp Gly Gln Thr His Lys Ile
 1 5 10 15

Ala Arg His Ile Ala Gly Val Leu Glu Glu Gln Gly Lys Ala Cys Glu
 20 25 30

Leu Val Asp Leu Leu Gln Pro Gly Glu Pro Asp Trp Ser Thr Val Glu
 35 40 45

Cys Val Val Leu Gly Ala Ser Ile Arg Tyr Gly His Phe His Lys Ser
 50 55 60

Phe Ile Arg Phe Val Asn Thr His Ala Gln Arg Leu Asn Asn Met Pro
 65 70 75 80

Gly Ala Leu Phe Thr Val Asn Leu Val Ala Arg Lys Pro Glu Lys Gln
 85 90 95

Ser Pro Gln Thr Asn Ser Tyr Thr Arg Lys Phe Leu Ala Ala Ser Pro
 100 105 110

Trp Gln Pro Gln Arg Cys Gln Val Phe Ala Gly Ala Leu Arg Tyr Pro
 115 120 125

Arg Tyr Ser Trp Tyr Asp Arg Met Met Ile Arg Leu Ile Met Lys Met
 130 135 140

Ala Gly Gly Glu Thr Asp Thr Arg Lys Glu Val Glu Tyr Thr Asp Trp
 145 150 155 160

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

Thr Arg

<210> 116
 <211> 179
 <212> PRT
 <213> Enterobacter cloacae

<400> 116

Lys Thr Leu Ile Leu Phe Ser Thr Arg Asp Gly Gln Thr Arg Glu Ile
 1 5 10 15

Ala Ala Phe Leu Ala Ser Glu Leu Lys Glu Gln Gly Ile Tyr Ala Asp
 20 25 30

Val Ile Asn Leu Asn Arg Thr Glu Glu Ile Ala Trp Gln Glu Tyr Asp
 35 40 45

Arg Val Val Ile Gly Ala Ser Ile Arg Tyr Gly His Phe His Pro Ala
 50 55 60

Val Asp Arg Phe Val Lys Lys His Thr Glu Thr Leu Asn Ser Leu Pro
 65 70 75 80

Gly Ala Phe Phe Ser Val Asn Leu Val Ala Arg Lys Ala Glu Lys Arg
 85 90 95

Thr Pro Gln Thr Asn Ser Tyr Thr Arg Lys Phe Leu Leu Asn Ser Pro
 100 105 110

Trp Lys Pro Ala Ala Cys Ala Val Phe Ala Gly Ala Leu Arg Tyr Pro
 115 120 125

Arg Tyr Arg Trp Tyr Asp Arg Phe Met Ile Arg Leu Ile Met Lys Met
 130 135 140

Thr Gly Gly Glu Thr Asp Thr Arg Lys Glu Val Val Tyr Thr Asp Trp
145 150 155 160

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

Ser Arg Leu

<210> 117
<211> 178
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 117

Lys Ala Leu Val Leu Tyr Ser Thr Arg Asp Gly Gln Thr His Ala Ile
1 5 10 15

Ala Ser Tyr Ile Ala Ser Cys Met Lys Glu Lys Ala Glu Cys Asp Val
20 25 30

Ile Asp Leu Thr His Gly Glu His Val Asn Leu Thr Gln Tyr Asp Gln
35 40 45

Val Leu Ile Gly Ala Asn Ile Arg Tyr Gly His Phe Asn Ala Val Leu
50 55 60

Asp Lys Phe Ile Lys Arg Asn Val Asp Gln Leu Asn Asn Met Pro Ser
65 70 75 80

Ala Phe Phe Cys Val Asn Leu Thr Ala Arg Lys Pro Glu Lys Arg Thr
85 90 95

Pro Gln Thr Asn Pro Tyr Val Arg Lys Phe Leu Leu Ala Thr Pro Trp
100 105 110

Gln Pro Ala Leu Cys Gly Val Phe Ala Gly Ala Leu Arg Tyr Pro Arg
115 120 125

Tyr Arg Trp Ile Asp Lys Val Met Ile Gln Leu Ile Met Arg Met Thr
130 135 140

Gly Gly Glu Thr Asp Thr Ser Lys Glu Val Glu Tyr Thr Asp Trp Glu
145 150 155 160

Gln Val Lys Lys Phe Ala Glu Asp Phe Ala Lys Leu Ser Tyr Lys Lys
165 170 175

Ala Leu

<210> 118
<211> 172
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 118

Lys Ala Leu Val Leu Tyr Ser Thr Arg Asp Gly Gln Thr His Ala Ile
1 5 10 15

Ala Ser Tyr Ile Ala Ser Cys Met Lys Glu Lys Ala Glu Cys Asp Val
20 25 30

Ile Asp Leu Thr His Gly Glu His Val Asn Leu Thr Gln Tyr Asp Gln
35 40 45

Val Leu Ile Gly Ala Ser Ile Arg Tyr Gly His Phe Asn Ala Val Leu
50 55 60

Asp Lys Phe Ile Lys Arg Asn Val Asp Gln Leu Asn Asn Met Pro Ser
65 70 75 80

Ala Phe Phe Cys Val Asn Leu Thr Ala Arg Lys Pro Glu Lys Arg Thr
85 90 95

Pro Gln Thr Asn Pro Tyr Val Arg Lys Phe Leu Leu Ala Thr Pro Trp
100 105 110

Gln Pro Ala Leu Cys Gly Val Phe Ala Gly Ala Leu Arg Tyr Pro Arg
115 120 125

Tyr Arg Trp Ile Asp Lys Val Met Ile Gln Leu Ile Met Arg Met Thr
130 135 140

Gly Gly Glu Thr Asp Thr Ser Lys Glu Val Glu Tyr Thr Asp Trp Glu
 145 150 155 160

Gln Val Lys Lys Phe Ala Glu Asp Phe Ala Lys Leu
 165 170

<210> 119
 <211> 179
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 119

Lys Thr Leu Ile Leu Phe Ser Thr Arg Asp Gly Gln Thr Arg Glu Ile
 1 5 10 15

Ala Ala Phe Leu Ala Ser Glu Leu Lys Glu Gln Gly Ile Tyr Ala Asp
 20 25 30

Val Ile Asn Leu Asn Arg Thr Glu Glu Ile Ala Trp Gln Glu Tyr Asp
 35 40 45

Arg Val Val Ile Gly Ala Ser Ile Arg Tyr Gly His Phe His Pro Ala
 50 55 60

Val Asp Arg Phe Val Lys Lys His Thr Glu Thr Leu Asn Ser Leu Pro
 65 70 75 80

Gly Ala Phe Phe Ser Val Asn Leu Val Ala Arg Lys Ala Glu Lys Arg
 85 90 95

Thr Pro Gln Thr Asn Ser Tyr Thr Arg Lys Phe Leu Leu Asn Ser Pro
 100 105 110

Trp Lys Pro Ala Ala Cys Ala Val Phe Ala Gly Ala Leu Arg Tyr Pro
 115 120 125

Arg Tyr Arg Trp Tyr Asp Arg Phe Met Ile Arg Leu Ile Met Lys Met
 130 135 140

Thr Gly Gly Glu Thr Asp Thr Arg Lys Glu Val Val Tyr Thr Asp Trp
 145 150 155 160

Ser Gln Ile Ala Ser Phe Ala Arg Glu Ile Val Gln Leu Thr Arg Ser

165

170

175

Ser Arg Leu

<210> 120

<211> 504

<212> PRT

<213> *Amaranthus tuberculatus*

<400> 120

Met Gly Asn Ile Ser Glu Arg Glu Glu Pro Thr Ser Ala Lys Arg Val
 1 5 10 15

Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala Ala Ala Tyr Lys Leu
 20 25 30

Lys Ser His Gly Leu Ser Val Thr Leu Phe Glu Ala Asp Ser Arg Ala
 35 40 45

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

Gly Ala Asn Thr Met Thr Glu Ser Glu Ala Glu Val Ser Ser Leu Ile
 65 70 75 80

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

Lys Arg Tyr Ile Ala Arg Asp Gly Leu Pro Val Leu Leu Pro Ser Asn
 100 105 110

Pro Ala Ala Leu Leu Thr Ser Asn Ile Leu Ser Ala Lys Ser Lys Leu
 115 120 125

Gln Ile Met Leu Glu Pro Phe Leu Trp Arg Lys His Asn Ala Thr Glu
 130 135 140

Leu Ser Asp Glu His Val Gln Glu Ser Val Gly Glu Phe Phe Glu Arg
 145 150 155 160

His Phe Gly Lys Glu Phe Val Asp Tyr Val Ile Asp Pro Phe Val Ala
 165 170 175

Gly Thr Cys Gly Gly Asp Pro Gln Ser Leu Ser Met His His Thr Phe

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

Gln Leu Leu Gly Thr Glu Asp Glu Pro Ser Phe Val Asn His Leu Phe
 420 425 430

Trp Ser Asn Ala Phe Pro Leu Tyr Gly His Asn Tyr Asp Ser Val Leu
 435 440 445

Arg Ala Ile Asp Lys Met Glu Lys Asp Leu Pro Gly Phe Phe Tyr Ala
 450 455 460

Gly Asn His Lys Gly Gly Leu Ser Val Gly Lys Ala Met Ala Ser Gly
 465 470 475 480

Cys Lys Ala Ala Glu Leu Val Ile Ser Tyr Leu Asp Ser His Ile Tyr
 485 490 495

Val Lys Met Asp Glu Lys Thr Ala
 500

<210> 121
 <211> 537
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 121
 atgaaagcgc tgggtgctgta tagcacccgc gatggccaga cccatgcgat tgcgagctat 60
 attgcgagct gcatgaaaga aaaagcggaa tgcgatgtga ttgatctgac ccatggcgaa 120
 catgtgaacc tgaccagta tgatcaggtg ctgattggcg cgagcattcg ctatggccat 180
 tttaacgcgg tgctggataa atttattaaa cgcaacgtgg atcagctgaa caacatgccg 240
 agcgcgtttt tttgcgtgaa cctgaccgcg cgcaaaccgg aaaaacgcac cccgcagacc 300
 aaccgatatg tgcgcaaatt tctgctggcg acccgtggc agccggcgct gtgcggcgctg 360
 tttgcgggcg cgctgcgcta tccgcgctat cgctggattg ataaagtgat gattcagctg 420
 attatgcgca tgaccggcgg cgaaaccgat accagcaaag aagtggaata taccgattgg 480
 gaacaggtga aaaaatttgc ggaagatttt gcgaaactga gctataaaaa agcgcctg 537

<210> 122
 <211> 534
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 122
 atgaaagcgc tgattctgtt tagcaccgc gatggccaga cccagaaaat tgcgagcgcg 60
 attgcggtatg aaattaaagg ccagcagagc tgcgatgtga ttaacattca ggatgcgaaa 120
 accctggatt ggcagcagta tgatcgcgtg ctgattggcg cgagcattcg ctatggccat 180
 tttcagccgg tgggtgaacga atttgtgaaa cataacctgc tggcgctgca gcagcgcgtg 240
 agcggctttt ttagcgtgaa cctgaccgcg cgcaaaccgg aaaaacgcag cccggaaacc 300
 aacgcgtata ccgtgaaatt tctggcgagc agcccgtggc agccggattg ctgcgcggtg 360
 tttgcggggc cgctgtatta tccgcgctat cgctggtttg atcgcgtgat gattcagttt 420
 attatgcgca tgaccggcgg cgaaaccgat gcgagcaaag aagtggaata taccgattgg 480
 cagcaggtgc agcgttttgc gcgcgatttt gcgcagctgc cgggcaaaag ctat 534

<210> 123
 <211> 531
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 123
 atgaaagcgc tgattctgta tagcaccgc gatggccaga cccgcaaaat tgcgagcagc 60
 attgcggtatg tgattcgcca gcagcagcag tgcgatgtgc tgaacattaa agatgcgagc 120
 ctgccggatt gggcgcagta tgatcgcgtg ctgattggcg cgagcattcg ctatggccat 180
 tttcagccgg tgggtggataa atttgtgaaa cagcatctgc atgaactgca gcagcgcacc 240
 agcggctttt ttagcgtgaa cctgaccgcg cgcaaaccgg aaaaacgcag cccggaaacc 300
 aacgcgtata cccagaaatt tctggcgcat agcccgtggc agccggattg ctgcgcggtg 360
 tttgcggggc cgctgtatta tccgcgctat cgctggtttg atcgcgtgat gattcagctg 420
 attatgcgca tgaccggcgg cgaaaccgat agcaccaaag aagtggaata taccgattgg 480
 cagcaggtga gcacctttgc gaacgatttt gcgcagctgc cgggcaaaaag c 531

<210> 124
 <211> 546
 <212> ADN
 <213> Escherichia coli

<400> 124
 gtgaaaacat taattctttt ctcaacaagg gacggacaaa cgcgcgagat tgcctcctac 60
 ctggcttcgg aactgaaaga actggggatc caggcggtatg tcgccaatgt gcaccgcatt 120

gaagaaccac agtgggaaaa ctatgaccgt gtggtcattg gtgcttctat tcgctatggt 180
 cactaccatt cagcggtcca ggaatttgtc aaaaaacatg cgacgcggct gaattcgatg 240
 ccgagcgcct ttactccgt gaatctgggtg gcgcgcaaac cggagaagcg tactccacag 300
 accaacagct acgcgcgcaa gtttctgatg aactcgcaat ggcggtcccg tcgctgcgcg 360
 gtcattgccg gggcgctgcg ttaccacgt tatcgctggg acgaccgttt tatgatcaag 420
 ctgattatga agatgtcagg cggtgaaacg gatacgcgca aagaagttgt ctataccgat 480
 tgggagcagg tggcgaattt cggccgagaa atcgccatt taaccgacaa accgacgctg 540
 aaataa 546

<210> 125
 <211> 534
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 125
 atgaaagcgc tgattctgtt tagcagccgc gaaggccaga cccgcgaaat tgcgagctat 60
 attgcaaca gcattaaaga agaaatggaa tgcgatgtgt ttaacattct gcgcgtggaa 120
 cagattgatt ggagccagta tgatcgctg ctgattggcg gcagcattca ttatggccat 180
 tttcatccgg cgggtggcgaa atttgtgaaa cgccatctgc atgaactgca gcagcgcagc 240
 agcggctttt tttgcgtgaa cctgaccgcg cgcaaagcgg ataaacgcac cccgcagacc 300
 aacgcgtata tgcgcaaatt tctgctgcag agcccgtggc agccggattg ctgcgcggtg 360
 tttgcgggcg cgctgcgcta taccgctat cgctggtttg atcgctgat gattcagctg 420
 attatgcgca tgaccggcgg cgaaaccgat accagcaaag aagtggaata taccgattgg 480
 acccaggtgg cgcgctttgc gcaggaattt gcgcatctgc cgggcaaaac ccag 534

<210> 126
 <211> 537
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 126
 atgaaagcgc tgattgtgtt tagcagccgc gatggccaga cccgcgcgat tgcgagctat 60
 attgcaaca ccctgaaagg caccctggaa tgcgatgtgg tgaacgtgct gaacgcgaac 120
 gatattgatc tgagccagta tgatcgctg gcgattggcg cgagcattcg ctatggccgc 180

tttcatccgg cgggtgaacca gtttattcgc aaacatctga ccagcctgca gcagctgccg 240
 agcgcgtttt ttagcgtgaa cctgaccgcg cgcaaaccgg aaaaacgcac cattcagacc 300
 aacgcgtata cccgcaaatt tctgctgaac agcccgtggc agccggatct gtgctgcgtg 360
 tttgcgggcg cgctgcgcta tccgcgctat cgctggtttg atcgctgat gattcagctg 420
 attatgcgca ttaccggcgg cgaaaccgat agcaccaaag aaattgaata taccgattgg 480
 cagcaggtgg cgcgctttgc gcaggatttt gcgcagctgg cggcgaaaaa cccggcg 537

<210> 127
 <211> 537
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 127
 atgaagacct tgatcctatt ctccaccagg gacggccaaa cacacaagat cgcaaggcac 60
 atcgcaggag tcctcgaaga gcaggggaag gcctgcgagt tggtcgatct gttacagccc 120
 ggccaaccag actggagtag cgttgaatgc gtcgttctag gggccagcat tagatatggt 180
 cacttccata agtctttcat caggttcgta aacactcacg cgcagcgctt gaataatatg 240
 ccaggcgccc ttttcacagt taacttagtc gcccgaaagc ccgagaagca gagtccacag 300
 acgaactctt acaccgcga gtttctcgcc gcctcccctt ggcagccaca gcgatgcaa 360
 gttttcgcgg gcgctttgag gtaccctagg tactcgtggg acgacagaat gatgatacgt 420
 ttgataatga agatggccgg gggcgagact gacacaagga aggaggttga gtacactgac 480
 tggcagtcgg tgactcgggt cgcgaggagg atcgctcagc tgccggggaga gacgcgg 537

<210> 128
 <211> 534
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 128
 atgaaagcgc tgattctggt tagcagccgc gatggccaga cccagctgat tgcgagcagc 60
 attgcgaaag aactggaagg caaacaggcg tgcgatgtgc tgaacattct ggataccacc 120
 aacgtggaat ggaccagta tgatcgctg ctgattggcg cgagcattcg ctatggccat 180
 tttcatccgg cgggtggcga atttgtgaaa cgccatcagc gcgaactgca gcagcgcagc 240

agcggctttt ttagcgtgaa cctgaccgcg cgcaaaccgg aaaaacgcag cccggaaacc 300
 aacgcgtata ccgcgaaatt tctgaaccag agcccggtggc agccggattg ctgcgcggtg 360
 tttgcggggcg cgctgcgcta tccgcgctat cgctggtttg atcg cattat gattcagctg 420
 attatgcgca tgaccggcgg cgaaaccgat agcagcaaag aagtggaata taccgattgg 480
 cagcaggtga cccgctttgc gcaggaattt gcgcgcctgc cgggcaaaac cagc 534

<210> 129
 <211> 540
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 129
 atgaaaaccc tgattctggt tagcaccgcg gatggccaga cccgcgaaat tgcggcggtt 60
 ctggcgagcg aactgaaaga acagggcatt tatgcggatg tgattaacct gaaccgcacc 120
 gaagaaattg cgtggcagga atatgatcgc gtggtgattg gcgcgagcat tcgctatggc 180
 cattttcatc cggcgggtgga tcgctttgtg aaaaaacata ccgaaaccct gaacagcctg 240
 ccgggcgcgt tttttagcgt gaacctggtg gcgcgcaaag cggaaaaacg caccgccgag 300
 accaacagct ataccgcga atttctgctg aacagcccgt ggaaaccggc ggcgtgcgcg 360
 gtgtttgcgg gcgcgctgcg ctatccgcgc tatcgctggt atgatcgctt tatgattcgc 420
 ctgattatga aaatgaccgg cggcgaaacc gataccgcga aagaagtggg gtataccgat 480
 tggagccagg tggcgagctt tgcgcgcgaa attgtgcagc tgaccgcgag cagccgcctg 540

<210> 130
 <211> 531
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 130
 atgaaaattc tgattctggt tagcaccgcg gatggccaga cccgcgaaat tgcggcgagc 60
 ctggcgagcg aactgaaaga acagggcgtt gatgtggatg tggatgaacct gcatcgcgcg 120
 gaaaacattg cgtgggaaga atatgatggc gtggtgattg gcgcgagcat tcgctatggc 180
 cattttcata gcaccctgaa cagctttgtg aaaaaacatc agcaggcgct gaaaaaactg 240
 ccgggcgcgt tttatagcgt gaacctggtg gcgcgcaaac cggaaaaacg caccgccgag 300
 accaacagct ataccgcga atttctgctg gatagcccgt ggcagccgga tctgagcgcg 360

gtgtttgcgg ggcgcgtgcg ctatccgcgc tataactggg atgatcgcat tatgattcgc 420
 ctgattatga aaattaccgg cggcgaaacc gatacccgca aagaagtggg gtataccgat 480
 tggcagcagg tgaccatttt tgcgcatgaa attgtgcagc tgggtgcgcaa a 531

<210> 131
 <211> 540
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 131
 atgaaggcct tgggtactgta ctgcgacgcg gacggccaga cccacgcaat tgcttcatac 60
 atcgccctct gcatgaagga gaaggccgaa tgcgacgtga tcgacctcac ccacggggag 120
 cactgaacc tcaccaata cgatcaggtg ctaatcgggtg cgagtattcg ttacggccac 180
 ttcaacgccg tgcttgacaa gttcatcaag agaaacgtgg atcagctgaa caacatgcca 240
 agcgcgttct tctgcgtaaa cctcacagca aggaagcccg agaagcgtag tccccagaca 300
 aacccttatg tccgaaaatt cttgcttgct accccctggc agcccgcggt gtgcggagtg 360
 ttcgcagggg cccttcggta cccgcgatac cgggtggatcg acaagggtgat gatccagcta 420
 ataatgcgga tgactggggg agagacagac acgagcaagg aggtcgagta cacggattgg 480
 gagcaggtta agaagttcgc ggaggatttt gcaaagctat cgtacaagaa ggccctctag 540

<210> 132
 <211> 537
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 132
 atgaaggcct tgatcctggt ctctacacgc gacggacaga cacagaagat cgcattctgcc 60
 atcgctgatg agataaaggg gcagcaatcg tgcgacgtga ttaacataca ggatgccaaa 120
 accctcgact ggcagcagta cgaccgggta ctcatcggcg cctccattcg ttacggggcat 180
 ttccagcccg ttgtgaatga gtttgtcaag cacaacctct tggccctaca gcagagagtt 240
 tccggattct tctccgtgaa cttgacagcc cgaaagccag agaagcggag ccccgagact 300
 aacgcttata cagtcaaatt cttggcgag tcaccctggc aaccggactg ctgcgctggt 360
 tttgcggggg ccctgtacta cccacggtag cgggtggttcg atagggtgat gatacagttc 420

ataatgcgaa tgacgggggg agagaccgac gcatcgaaag aggtggagta cactgactgg 480
cagcaggtgc agcggttcgc gcgagacttc gcgcagttac cgggtaagtc ctactga 537

<210> 133
<211> 534
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 133
atgaaggcgc tgatcttgta ctcaaccagg gacggtcaga ctcgcaagat tgcaagtagc 60
attgcggacg tcatcaggca gcagcagcag tgcgacgtct taaacattaa agacgcatca 120
cttcctgact gggcccaata tgaccgagtg ctcatcggag ctagcatccg ttacgggcat 180
ttccagcccg ttgtagacaa gttcgtgaag cagcacttgc acgagcttca gcagcggacc 240
tccggcttct tctccgtgaa cctgacggcg aggaagcctg aaaaaaggag ccctgagacc 300
aatgcctaca cccagaaatt cttggcgcac tccccttggc agcccgattg ctgtgccgtt 360
ttcgcggggg ccctttacta cccaggtac cgttggttcg accgggtgat gatccagttg 420
attatgcgca tgactggtgg agagaccgac tctaccaagg aagtggagta cactgactgg 480
cagcaggtga gtaccttcgc caacgatttt gccagcttc caggcaagag ctaa 534

<210> 134
<211> 546
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 134
atgaagacct tgattctatt ctccacaagg gacggccaga ctagggagat cgcttcttac 60
ctggccagcg agctaaagga gcttggcatt caggcagacg tggctaactg gcaccgaatt 120
gaggagccgc agtgggagaa ctacgatcgg gtcgtgatcg gcgccagcat ccggtatgga 180
cactaccaca gcgcgttcca ggagttcgtg aaaaagcacg cgacctgtct gaatagcatg 240
ccatcagcgt tctactcggc caacctcgtg gtcgtaagc ccgagaagcg gacacccag 300
accaactcgt atgccaggaa gttccttatg aactcgcagt ggcgaccgga ccgctgcgcg 360
gtgatcgccg gtgcgctcag gtaccctcgt tataggtggc acgacaggtt tatgattaaa 420
cttataatga aaatgagcgg cggagagacc gacaccagaa aagaggtggc ttacacagac 480
tgggagcagg tagcaaactt cgctagggag attgctcacc tcaccgacaa gccgaccttg 540

aagtaa

546

<210> 135
 <211> 537
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 135
 atgaaggccc ttatactggt cagttccaga gaaggccaga cgaggagat agcgagttac 60
 attgccaact cgataaagga ggaaatggaa tgcgacgtgt tcaacatcct tcgtgtggag 120
 cagatcgact ggtctcaata cgaccgcgtc ctgatcgggg gctcgatata ctacggccat 180
 ttccacccag cgggtggcaaa atttgtcaag aggcacctcc atgagttgca acagaggtct 240
 tccggctttt tctgcgtaaa cctgacggcc aggaaggccg acaagcggac tccccagacc 300
 aatgcctaca tgagaaagtt cttgttgtag tccccatggc aacccgattg ctgcgccgtg 360
 tttgcggggg cccttaggta caccggttac aggtgggttcg acagggtaat gattcagctg 420
 atcatgagga tgacggggcg agagactgac acatcgaagg aggtggagta cacagactgg 480
 acgcaggtcg cccgcttcgc gcaggagttc gccattttgc ccggcaaaac tcagtga 537

<210> 136
 <211> 540
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 136
 atgaaggctc ttatcgtatt ctcttcgagg gatggccaaa cccgagcgat cgcgtcttat 60
 attgctaata ccctcaaagg gaccctagag tgcgacgtcg tcaacgtcct caatgctaac 120
 gacattgatt tgagccagta cgaccgtgtg gccattggcg cctccattcg ctacgggagg 180
 ttccacccag ctgttaacca gtttatccgg aagcacctta cgagcctcca gcagctacca 240
 tctgcgttct tctccgtgaa cctcacagct cggaagcccg agaagaggac tatacaaacc 300
 aacgcgtaca ctaggaagtt tctactgaac tcgccgtggc agccggacct gtgctgctg 360
 ttgcggggag cccttcgcta tccccgttac aggtgggtttg accgagtgat gattcaactc 420
 ataatgcgca taacgggggg cgagacagac tccaccaagg agatcgagta caccgactgg 480
 cagcaggtcg cgcgattcgc ccaggatttt gcacagcttg ccgcaaagaa cccggcatga 540

<210> 137
 <211> 540
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 137
 atgaagacct tgatcctatt ctccaccagg gacggccaaa cacacaagat cgcaaggcac 60
 atcgcaggag tcctcgaaga gcaggggaag gcctgcgagt tggtcgatct gttacagccc 120
 ggccaaccag actggagtag cgttgaatgc gtcgttctag gggccagcat tagatatggt 180
 cacttccata agtctttcat caggttcgta aacactcacg cgcagcgctt gaataatatg 240
 ccaggcgccc ttttcacagt taacttagtc gcccgaaagc ccgagaagca gagtccacag 300
 acgaactctt acaccgcaa gtttctcgcc gcctcccctt ggcagccaca gcgatgccaa 360
 gttttcgcgg gcgctttgag gtaccctagg tactcgtggt acgacagaat gatgatacgt 420
 ttgataatga agatggccgg gggcgagact gacacaagga aggaggttga gtacactgac 480
 tggcagtcgg tgactcgggt cgcgaggag atcgctcagc tgccgggaga gacgcggtag 540

<210> 138
 <211> 537
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 138
 atgaaggccc taattttatt cagtagtagg gacggccaga cccagcttat agcatcgtct 60
 atcgccaagg agctcgaagg gaagcaggcg tgcgacgtgt tgaatatcct cgacacgact 120
 aatgtggagt ggaccagta cgaccgcgtg ctgattggag catccatccg gtacggggcac 180
 tttcaccctg cggtcgccga gttcgtaaag cgtcaccagc gagagctaca gcagagaagt 240
 agtggctttt tctctgtgaa cttgacggcc cgtaagccgg aaaagaggtc ccccgagact 300
 aacgcctata ccgccaagtt ccttaaccaa agtccatggc agcctgactg ttgcgctgtg 360
 ttcgctgggg ctttgcgata ccctcggtac cgctgggttc acagaattat gatccagcta 420
 atcatgcgga tgactggggg tgagacagat tcttcaaagg aggtcgagta caccgactgg 480
 cagcaggtga cccgcttcgc gcaagagttc gccaggcttc cgggaaagac cagttga 537

<210> 139
 <211> 543

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 139

atgaagaccc taatactggt ctctacccgc gacgggcaga caagggagat cgccgcgttc	60
cttgcctcgg agctgaagga gcaggggatt tacgctgacg tcataaacct taaccggacg	120
gaggagatag cttggcagga gtatgataga gtcgtaatcg gggcgtcgat ccgatacggg	180
catttccacc ctgctgtcga ccgcttcgtg aagaagcaca cagagacact caactcactg	240
cccggcgcct ttttctctgt aaaccttggt gcccggaag ccgagaagag aacgccgcag	300
acgaactcat acaccaggaa gttcctatta aacagcccgt ggaagccagc ggccctgcgcg	360
gtctttgctg gggccctccg ctaccctaga taccgctggt acgacagggt catgatacga	420
ctgattatga aaatgacagg cggggagacg gatacccgaa aggaggtagt ctacactgac	480
tggtcgcagg tcgcgtcggt tgccagagag atagtccagt tgaccaggtc atcgcgcttg	540
tga	543

<210> 140

<211> 534

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 140

atgaagatat taatcctttt ctccaccggt gacggccaaa cccgtgagat tgcggcgtcc	60
ttggcgtccg aactcaagga gcaggcattc gacgtggacg tcgtcaacct tcaccggggc	120
gagaacatcg catgggagga gtacgacggt gttgtcatcg gagcgtccat caggtacggc	180
cactttcata gtaccctgaa ctcatctgtc aagaagcatc agcaggctct taagaagctt	240
cccggggctt tctacagcgt gaacctcgtc gcccggaagc ctgagaagcg cacaccgcag	300
accaatagct acaccgcaa gttcctcttg gattccccgt ggcagcccga cttttcagcc	360
gtgttcgccg gggcactcag gtaccctcgg tacaattggt acgaccgtat catgattaga	420
cttatcatga agattacagg cggcgagact gataccagga aggaagtagt ctacacagac	480
tggcagcagg tcactcactt tgctcacgag atcgtccagc tcgtgcggaa gtag	534

<210> 141

<211> 537

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 141

aaggccttg	tactgtactc	gacgcgggac	ggccagaccc	acgcaattgc	ttcatacatc	60
gcctcctgca	tgaaggagaa	ggccgaatgc	gacgtgatcg	acctcaccca	cggggagcac	120
gtgaacctca	ccaatacga	tcaggtgcta	atcggtgcca	gtattcgta	cggccacttc	180
aacgccgtgc	ttgacaagtt	catcaagaga	aacgtggatc	agctgaacaa	catgccaaagc	240
gcgtttcttct	gcgtaaacct	cacagcaagg	aagcccgaga	agcgactacc	ccagacaaac	300
ccttatgtcc	gaaaattctt	gcttgctacc	ccctggcagc	ccgcgttggtg	cggagtgttc	360
gcagggggccc	ttcgggtaccc	gcgataccgg	tggatcgaca	aggtgatgat	ccagctaata	420
atgcggatga	ctggggggaga	gacagacacg	agcaaggagg	tcgagtacac	ggattgggag	480
caggttaaga	agttcgcgga	ggattttgca	aagctatcgt	acaagaaggc	cctctag	537

<210> 142

<211> 534

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 142

aaggccttga	tcctgttctc	tacacgcgac	ggacagacac	agaagatcgc	atctgccatc	60
gctgatgaga	taaaggggca	gcaatcgtgc	gacgtgatta	acatacagga	tgccaaaacc	120
ctcgactggc	agcagtacga	ccgggtactc	atcggcgcct	ccattcgta	cgggcatttc	180
cagcccgttg	tgaatgagtt	tgtcaagcac	aacctcttg	ccctacagca	gagagtttcc	240
ggattcttct	ccgtgaactt	gacagcccga	aagccagaga	agcggagccc	cgagactaac	300
gcttatacag	tcaaattctt	ggcgcagtca	ccctggcaac	cggactgctg	cgctgttttt	360
gcggggggccc	tgtactaccc	acggtaccgg	tggttcgata	gggtgatgat	acagttcata	420
atgcgaatga	cgggggggaga	gaccgacgca	tcgaaagagg	tggagtacac	tgactggcag	480
cagggtgcagc	ggttcgcgcg	agacttcgcg	cagttaccgg	gtaagtccta	ctga	534

<210> 143

<211> 543

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 143

```
aagaccttga ttctattctc cacaagggac ggccagacta gggagatcgc ttcctacctg      60
gccagcgagc taaaggagct tggcattcag gcagacgtgg ctaacgtgca ccgaattgag      120
gagccgcagt gggagaacta cgatcgggtc gtgatcggcg ccagcatccg gtatggacac      180
taccacagcg cgttccagga gttcgtgaaa aagcacgcga cccgtctgaa tagcatgcca      240
tcagcgttct actcgggtcaa cctcgtggct cgtaagcccc agaagcggac accccagacc      300
aactcgtatg ccaggaagtt ccttatgaac tcgcagtggc gaccggaccg ctgctcggtg      360
atcgccggtg cgctcaggta ccctcgttat aggtgggtacg acaggtttat gattaaactt      420
ataatgaaaa tgagcggcgg agagaccgac accagaaaag aggtggttta cacagactgg      480
gagcaggtag caaacttcgc tagggagatt gctcacctca ccgacaagcc gaccttgaag      540
taa                                          543
```

<210> 144

<211> 534

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 144

```
aaggccctta tactgttcag ttccagagaa ggccagacga gggagatagc gagttacatt      60
gccaaactcga taaaggagga aatggaatgc gacgtgttca acatccttcg tgtggagcag      120
atcgactggt ctcaatacga ccgcgtcctg atcgggggct cgatacacta cggccatttc      180
caccagcgcg tggcaaaatt tgtcaagagg cacctccatg agttgcaaca gaggtcttcc      240
ggctttttct gcgtcaacct gacggccagg aaggccgaca agcggactcc ccagaccaat      300
gcctacatga gaaagttctt gttgcagtcc ccatggcaac ccgattgctg cgccgtgttt      360
gcggggggccc ttaggtacac ccgttacagg tggttcgaca gggtaatgat tcagctgac      420
atgaggatga cgggcggaga gactgacaca tcgaaggagg tggagtacac agactggacg      480
caggtcgccc gcttcgcgca ggagttcgcc catttgcccc gcaaaactca gtga          534
```

<210> 145

<211> 537

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 145
aaggctctta tcgtattctc ttcgagggat ggccaaaccc gagcgatcgc gtcttatatt 60
gctaataccc tcaaaggac cctagagtgc gacgtcgtca acgtcctcaa tgctaacgac 120
attgatttga gccagtacga ccgtgtggcc attggcgccct ccattcgcta cgggaggttc 180
caccagctg ttaaccagtt tatccggaag caccttacga gcctccagca gctaccatct 240
gcgttcttct ccgtgaacct cacagctcgg aagcccgaga agaggactat acaaaccaac 300
gcgtacacta ggaagtttct actgaactcg ccgtggcagc cggacctgtg ctgctgtgtc 360
gcgggagccc ttcgctatcc ccgttacagg tggtttgacc gagtgatgat tcaactcata 420
atgcgcataa cggggggcga gacagactcc accaaggaga tcgagtacac cgactggcag 480
caggtcgcgc gattcgccca ggattttgca cagcttgccg caaagaaccc ggcatga 537

<210> 146
<211> 537
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 146
aagaccttga tcctattctc caccagggac ggccaaacac acaagatcgc aaggcacatc 60
gcaggagtcc tcgaagagca ggggaaggcc tgcgagttgg tcgatctggt acagcccggc 120
gaaccagact ggagtaccgt tgaatgcgtc gttctagggg ccagcattag atatggtcac 180
ttccataagt ctttcatcag gttcgtaaac actcacgcgc agcgcttgaa taatatgcca 240
ggcgcccttt tcacagttaa cttagtcgcc cgaaagcccg agaagcagag tccacagacg 300
aactcttaca cccgcaagtt tctcgccgcc tccccttggc agccacagcg atgccaagtt 360
ttcgcgggcg ctttgaggta ccctaggtac tcgtggtacg acagaatgat gatacgtttg 420
ataatgaaga tggccggggg cgagactgac acaaggaagg aggttgagta cactgactgg 480
cagtcggtga ctcggttcgc gagggagatc gctcagctgc cgggagagac gcggtag 537

<210> 147
<211> 540
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 147
aagaccctaa tactgttctc taccgcgcac gggcagacaa gggagatcgc cgcgttcctt 60

gcctcggagc tgaaggagca ggggatttac gctgacgtca taaaccttaa ccggacggag 120
gagatagctt ggcaggagta tgatagagtc gtaatcgggg cgtcgatccg atacgggcat 180
ttccaccctg ctgtcgaccg cttcgtgaag aagcacacag agacactcaa ctactgccc 240
ggcgccctttt tctctgtaaa ccttggtgcc cggaaagccg agaagagaac gccgcagacg 300
aactcataca ccaggaagtt cctattaaac agcccgtgga agccagcggc ctgcgcggtc 360
tttgctgggg cctccgcta ccctagatac cgctggtacg acaggttcat gatacgactg 420
attatgaaaa tgacaggcgg ggagacggat acccgaaagg aggtagtcta cactgactgg 480
tcgcaggtcg cgtcgtttgc cagagagata gtccagttga ccaggtcatc gcgcttgtga 540

<210> 148
<211> 537
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 148
aaggccttgg tactgtactc gacgcgggac ggccagaccc acgcaattgc ttcatacatc 60
gcctcctgca tgaaggagaa ggccgaatgc gacgtgatcg acctcaccca cggggagcac 120
gtgaacctca cccaatacga tcaggtgcta atcggtgca gtattcgta cggccacttc 180
aacgccgtgc ttgacaagtt catcaagaga aacgtggatc agctgaacaa catgccaagc 240
gcgttcttct gcgtaaacct cacggcaagg aagcccgaga agcgtactcc ccagacaaac 300
ccttatgtcc gaaaattctt gcttgctacc ccctggcagc ccgcgttgtg cggagtgttc 360
gcagggggccc ttcggtaccc gcgataccgg tggatcgaca aggtgatgat ccagctaata 420
atgcggatga ctgggggaga gacagacacg agcaaggagg tcgagtacac ggattgggag 480
caggttaaga agttcgcgga ggattttgca aagctatcgt acaagaaggc cctctag 537

<210> 149
<211> 537
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 149
aaggccttgg tactgtactc gacgcgggac ggccagaccc acgcaattgc ttcatacatc 60
gcctcctgca tgaaggagaa ggccgaatgc gacgtgatcg acctcaccca cggggagcac 120

gtgaacctca cccaatacga tcaggtgcta atcgggtgcga atattcgta cggccacttc 180
 aacgccgtgc ttgacaagtt catcaagaga aacgtggatc agctgaacaa catgccaagc 240
 gcgttcttct gcgtaaacct cacagcaagg aagcccgaga agcgtactcc ccagacaaac 300
 ccttatgtcc gaaaattctt gcttgctacc ccctggcagc ccgcgttggtg cggagtgttc 360
 gcagggggccc ttcggtaccc gcgataccgg tggatcgaca aggtgatgat ccagctaata 420
 atgcggatga ctgggggaga gacagacacg agcaaggagg tcgagtacac ggattgggag 480
 caggttaaga agttcgcgga ggattttgca aagctatcgt acaagaaggc cctctag 537

<210> 150
 <211> 537
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 150
 aaggccttgg tactgtactc gacgcgggac ggccagaccc acgcaattgc ttcatacatc 60
 gcctcctgca tgaaggagaa ggccgaatgc gacgtgatcg acctcaccca cggggagcac 120
 gtgaacctca cccaatacga tcaggtgcta atcgggtgcga gtattcgta cggccacttc 180
 aacgccgtgc ttgacaagtt catcaagaga aacgtggatc agctgaacaa catgccaagc 240
 gcgttcttct gcgtaaacct cacagcaagg aagcccgaga agcgtactcc ccagacaaac 300
 ccttatgtcc gaaaattctt gcttgctacc ccctggcagc ccgcgttggtg cggagtgttc 360
 gcagggggccc ttcggtaccc gcgataccgg tggatcgaca aggtgatgat ccagctaata 420
 atgcggatga ctgggggaga gacagacacg agcaaggagg tcgagtacac ggattgggag 480
 caggttaaga agttcgcgga ggattttgca aagctatagt acaagaaggc cctctag 537

<210> 151
 <211> 534
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 151
 aaggccttga tcctgttctc tacacgcgac ggacagacac agaagatcgc atctgccatc 60
 gctgatgaga taaaggggca gcaatcgtgc gacgtgatta acatacagga tgccaaaacc 120
 ctcgactggc agcagtacga ccgggtactc atcgggcgct ccattcgta cgggcatttc 180
 cagcccgttg tgaatgagtt tgtcaagcac aacctcttgg ccctacagca gagagtttcc 240

ggattcttct ccgatgaactt gacagcccga aagccagaga agcggagccc cgagactaac 300
 gcttatacag tcaaattctt ggcgagtc caacctggcaac cggactgctg cgctgttttt 360
 gcgggggccc tgtactacc acggtaccgg tggttcgata gggatgatgat acagttcata 420
 atgcgaatga cgggggggga gaccgacgca tcgaaagagg tggagtacac tgactggcag 480
 caggtgcagc ggttcgcgcg agacttcgcg cagttaccgg gtaagtccta ctga 534

<210> 152
 <211> 540
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 152
 aagaccctaa tactgttctc taccgcgac gggcagacaa gggagatcgc cgcgttcctt 60
 gcctcggagc tgaaggagca ggggatttac gctgacgtca taaaccttaa ccggacggag 120
 gagatagctt ggcaggagta tgatagagtc gtaatcgggg cgtcgatccg atacgggcat 180
 ttccaccctg ctgtcgaccg cttcgtgaag aagcacacag agacactcaa ctactgccc 240
 ggcgcccttt tctctgtaaa ccttgttgcc cggaagccg agaagagaac gccgcagacg 300
 aactcataca ccaggaagtt cctattaaac agcccgtgga agccagcggc ctgcgcggtc 360
 tttgctgggg ccctccgcta ccctagatac cgctgggtacg acaggttcat gatacgactg 420
 attatgaaaa tgacaggcgg ggagacggat acccgaaagg aggtagtcta cactgactgg 480
 tcgcagatcg cgtcgtttgc cagagagata gtccagttga ccaggtcatc gcgcttgtga 540

<210> 153
 <211> 540
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 153
 atgaaggcgc tcgtgctcta cagcacacgc gacggccaga ctcatgcgat cgcctcttac 60
 atcgcgtcct gtatgaagga gaaggccgag tgcgacgtca tcgatctcac gcacggggag 120
 cacgtgaatc ttacgcagta cgaccaagtg ctgataggcg cctctatccg ttacggccat 180
 tttaacgccg tcctcgacaa attcatcaag cgcaatgtag accagctgaa caacatgccc 240
 tccgcgttct tttgctgaa cctgacggct cggaagcctg agaagcgaac acctcagacc 300

aaccatacgc tgcggaaatt cctactcgca acgccatggc agcccgccct gtgcgggggtt 360
 ttcgcagggg cgctacgcta tccgcgttac cgctggatcg ataaggtgat gatccagcta 420
 ataatgcgca tgaccggcgg cgagacagac acatcgaagg aagtcgaata cacagactgg 480
 gaacaggtga agaagtttgc agaggatttc gccaaagctct catacaaaaa ggcattgtga 540

<210> 154
 <211> 537
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 154
 atgaaggcgc ttatactggt ctcgacacgc gacggtcaga cgcagaaaat cgcctcagcc 60
 atcgccgacg agatcaaggg ccagcagagc tgcgatgtga tcaatattca ggacgccaaa 120
 actctcgact ggcagcagta tgaccgcgtg ctcatggcg catcaatccg ctacgggcat 180
 ttccagccag tcgtcaatga gtttgtgaaa cataacctct tggcattgca gcagcgggtg 240
 tctggcttct tctccgtgaa ccttacagct agaaaaccag agaagcggtc gcccgagact 300
 aacgcctaca ccgttaagtt ccttgccgag tcaccgtggc agcctgattg ctgcgcggtc 360
 ttcgccgggg cactgtacta ccctcgatac cggtgggttg atagggtaat gatccagttc 420
 ataatgcgca tgaccggtgg ggagaccgac gcaagtaaag aagttgagta cacggattgg 480
 cagcaggtgc aaaggttcgc acgcgacttc gcgcagctcc cgggcaaagtc ttactga 537

<210> 155
 <211> 534
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 155
 atgaaagccc tgatcctcta ttccaccagg gacggccaga cccgcaagat agcctcctcc 60
 atcgctgatg tcatccgcca gcagcagcag tgcgacgttt taaacattaa ggacgcttca 120
 ctgcctgatt gggcccagta tgaccgcgtc ctgateggcg cgctcgattcg gtacggccac 180
 ttccagcctg tggttgacaa gttcgtcaag cagcacctgc atgagctgca gcagcgaact 240
 agcgggttct tcagtgtgaa cctgacagct agaaagcccg aaaagagatc ccagaaaacc 300
 aacgcctata cgcagaaaatt ccttgctcac tcaccctggc agcctgactg ttgtgccgtc 360
 ttcgcggggc ccttgtacta tccccgtac cgctgggttcg atagggtgat gatccagctg 420

attatgagaa tgacgggagg ggagaccgat tcgaccaagg aggtagagta cactgactgg 480
 caacaggtgt caactttcgc aaacgacttc gcacaactac ccggttaagtc ttga 534

<210> 156
 <211> 546
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 156
 atgaaaaccc taatactggt ctcgacccgc gacggccaga cgcgtgagat tgcgagctac 60
 ctggcctccg agctcaagga gctggggatc caagccgatg tcggaacgt gcaccgcatt 120
 gaggagccgc agtgggagaa ttacgatcgc gttgtgatag gggccagcat ccgctatggc 180
 cactaccact cggcctttca ggagtttgta aagaaacacg ccacaagatt aaactccatg 240
 cctagcgcct tctactccgt caaccttgtc gcgcgcaagc cggagaagcg gacacctcag 300
 acgaactcct acgcgcggaa gttcctgatg aacagccagt ggcggccgga cagatgtgct 360
 gttattgcgg gagccctgag ataccgagg taccggtggt acgatagggt tatgattaaa 420
 cttattatga agatgtctgg tggggagact gacaccagga aggaggtggt atatacagac 480
 tgggagcagg tcgccaattt cgctcgggaa atcgcgcac tcacagacaa gcctacactg 540
 aagtag 546

<210> 157
 <211> 537
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 157
 atgaaggccc tgatcctctt tagctctagg gagggccaga cccgcgagat cgcgtcatat 60
 atcgcgaatt ccataaagga ggagatggag tgcgatgtgt ttaacatcct tagggtggag 120
 caaatagact ggtctcagta tgaccgtgtg tcatagggg ggagcatcca ctacggccac 180
 tttcaccg cggtggcgaa attcgtcaag cgacacctcc acgagcttca gcagcgctcc 240
 tcagggttct tctgcgtcaa cctgacagca agaaaggcag ataaacgcac cccgcagacg 300
 aacgcctaca tgaggaagtt ccttctgcag tctccttggc agcccgattg ctgcgcggtg 360
 ttcgccggtg cactgcgcta tacgcgctat agatggtttg atagagtcatt gattcagctc 420

atcatgcgga tgaccggcgg ggaaacggat actagtaagg aggtggagta cacggactgg 480
 acccaggtgg cacgtttcgc ccaggagttt gcacatcttc ctgggaagac ccaatga 537

<210> 158
 <211> 540
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 158
 atgaaggcgc taattgtggt cagctccagg gatggccaga cgagggctat agcatcctat 60
 atcgccaata ccttgaaagg aacgctcgag tgtgacgtgg tcaacgtctt gaacgccaat 120
 gacattgacc tttcccagta cgaccgagtt gccataggcg cgtcgatccg ctacggggcga 180
 tttcaccctg cagtcaacca gtttatacgg aagcatttga cctcgctgca gcagctcccg 240
 tcagccttct tctctgtgaa tttaaccgcg cggaagcctg agaaacggac gatccaaaca 300
 aacgcctata cccgaaagtt cctcctgaac agcccatggc agccagacct gtgctgtgtc 360
 ttcgccggcg cgttgccggt tccccgctac aggtgggttcg atagagtgat gatccagctc 420
 atcatgagga tcaccggggg agagaccgat agtaccaagg agatcgagta cacggactgg 480
 cagcaggtgg ctcgcttcgc ccaggacttc gctcagttgg ccgcaaagaa tccagcataa 540

<210> 159
 <211> 540
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 159
 atgaagacac tgatcctggt ctcgactcga gatggccaga ctcataaaat tgcgcgccac 60
 attgcggggg tcctggagga gcagggcaaa gcgtgcgagc tcgtggactt actccagccc 120
 ggggagccgg actggagcac ggtggagtgc gtcgttctgg gcgcttccat acgttacggg 180
 cttttccaca aaagtttcat ccggttcgtc aacacccacg ctcaacgggt gaacaacatg 240
 cctggcgcgc tattcactgt taacttagtg gtcgtaagc ccgagaagca gtctccgcag 300
 actaactcct acacaaggaa atttctagca gcaagcccat ggcaaccgca gcggtgccag 360
 gtgttcgctg gagctctgcg ctatcctagg tacagttggg acgacagaat gatgatacgg 420
 ttgattatga agatggcagg cggggagacg gacaccagga aagaggtcga atacactgac 480
 tggcaatcag tcactcgggt tgctagagag atcgcgcaat taccaggtga gacgcggtaa 540

<210> 160
 <211> 537
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 160
 atgaaggctc tcatactggt cagctcgaga gacgggcaga cccagctgat cgcctcctcc 60
 atagcaaagg agctagaggg caagcaagcc tgcgacgtgc tcaatattct cgacacaacc 120
 aacgtggagt ggactcagta cgacagagtc ctaatcggcg cgtccatcag atacggccac 180
 ttccatcccc cgcgcgctga attcgtgaaa cgccaccagc gtgagctcca gcagcgcagc 240
 agcggcttct tcagcgtgaa tcttactgcg agaaagccgg aaaagcggag tcccagagact 300
 aacgcttata cggcaaagtt cctcaaccaa tctccctggc aaccagactg ctgtgccgtg 360
 ttcgctgggg cactgaggtg tccgcgctat cggtggttcg atagaatcat gatacagctg 420
 ataatgcgta tgactggtgg ggagacggat tccagtaaag aggtagagta tactgattgg 480
 cagcaggtca ctaggttcgc gcaggagttt gctaggctgc cgggcaagac atcctga 537

<210> 161
 <211> 543
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 161
 atgaaaacct taatcttgtt cagcacccgc gacggccaga cgcgtgaaat cgcagcgttc 60
 ctcgcttcgg agctcaagga acaggggaatt tacgccgacg tcattaacct aaaccgtacc 120
 gaagagattg cgtggcagga gtatgaccgc gtggtgattg gcgcttctat ccgctatggc 180
 cacttccacc cggctgttga ccggttcgtg aagaagcaca cggagacctt gaactcactg 240
 ccggggggcat tcttttagcgt aaatctggtg gcgcgcaagg ccgagaagcg cccccccag 300
 acgaacagct acacccgcaa atttttactt aactcccat ggaaacctgc ggcttgcgca 360
 gtgttcgcag gagctctccg ctatcctcgc tatcgatggt acgatcgggt catgattcgg 420
 ctgattatga aaatgacggg cggcgagacg gatacgcgaa aggaagttgt ctacactgac 480
 tgggtcccagg tggcctcggt tgcaaggagg atcgtacagc tcaactcgatc tagtaggctc 540
 tga 543

<210> 162
 <211> 534
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 162
 atgaagattc tcattcttatt ttccacccga gacggccaaa cccgcgagat tgcggcggtcc 60
 ctcgcctccg agttgaagga gcaggcggtt gatgtggatg tgggtcaacct ccaccgcgca 120
 gaaaacatag cgtgggagga gtacgatggg gtcgtcatcg gagcgtcaat ccgctacgga 180
 catttccact caacgctgaa ttcatTTgtg aagaagcacc aacaagcgct caagaagctg 240
 cccggagcat tctacagcgt caacctcgtg gctcggaagc cggaagcg caccgccgaa 300
 acaaacagct acacacgcaa gtttctgctc gactcgccct ggcaaccgga cctgagtgcc 360
 gttttcgccg gggcactgcg ctatccccgt tacaactggg acgatcgcat aatgattcga 420
 ctgatcatga agattacagg cggggaaacc gatactcgga aggaggtggg gtatacagac 480
 tggcagcagg ttaccactt cgccacgag atcgtccagc tcgttcgtaa gtga 534

<210> 163
 <211> 442
 <212> PRT
 <213> Xanthomonas campestris

<400> 163

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

Ser Ile Ala Tyr Glu Leu Gln Gln Lys Gly Ile Pro Tyr Glu Ile Met
 20 25 30

Glu Ala Ser Ser Tyr Ala Gly Gly Val Val Lys Ser Leu His Ile Asp
 35 40 45

Gly Tyr Glu Leu Asp Ala Gly Pro Asn Ser Leu Ala Ala Ser Ala Ala
 50 55 60

Phe Met Ala Tyr Ile Asp Gln Leu Gly Leu Gln Asp Gln Val Leu Glu
 65 70 75 80

Ala Ala Ala Ala Ser Lys Asn Arg Phe Leu Val Arg Asn Asp Lys Leu
 85 90 95

His Ala Val Ser Pro His Pro Phe Lys Ile Leu Gln Ser Ala Tyr Ile
 100 105 110

Ser Gly Gly Ala Lys Trp Arg Leu Phe Thr Glu Arg Phe Arg Lys Ala
 115 120 125

Ala Ala Pro Glu Gly Glu Glu Thr Val Ser Ser Phe Val Thr Arg Arg
 130 135 140

Phe Gly Lys Glu Ile Asn Asp Tyr Leu Phe Glu Pro Val Leu Ser Gly
 145 150 155 160

Ile Tyr Ala Gly Asn Pro Asp Leu Met Ser Val Gly Glu Val Leu Pro
 165 170 175

Met Leu Pro Gln Trp Glu Gln Lys Tyr Gly Ser Val Thr Gln Gly Leu
 180 185 190

Leu Lys Asn Lys Gly Ala Met Gly Gly Arg Lys Ile Ile Ala Phe Lys
 195 200 205

Gly Gly Asn Ala Thr Leu Thr Asn Arg Leu Gln Ser Leu Leu Ser Gly
 210 215 220

Lys Ile Arg Phe Asn Cys Ala Val Thr Gly Val Thr Arg Gly Ala Asp
 225 230 235 240

Asp Tyr Ile Val Gln Tyr Thr Glu Asn Gly Asn Thr Ala Met Leu Asn
 245 250 255

Ala Ser Arg Val Ile Phe Thr Thr Pro Ala Tyr Ser Thr Ala Val Ala
 260 265 270

Ile Gln Ala Leu Asp Ala Ser Leu Ala Thr His Leu Ser Asp Val Pro
 275 280 285

Tyr Pro Arg Met Gly Val Leu His Leu Gly Phe Gly Ala Glu Ala Arg
 290 295 300

Gln Lys Ala Pro Ala Gly Phe Gly Phe Leu Val Pro His Ala Ala Gly
 305 310 315 320

Lys His Phe Leu Gly Ala Ile Cys Asn Ser Ala Ile Phe Pro Ser Arg
 325 330 335

Val Pro Thr Gly Lys Val Leu Phe Thr Val Phe Leu Gly Gly Ala Arg
 340 345 350

Gln Glu Gln Leu Phe Asp Gln Leu Gly Pro Glu Lys Leu Gln Gln Thr
 355 360 365

Val Val Lys Glu Leu Met Glu Leu Leu Gly Leu Thr Thr Pro Pro Glu
 370 375 380

Met Gln Arg Phe Ser Glu Trp Asn Arg Ala Ile Pro Gln Leu Asn Val
 385 390 395 400

Gly Tyr Ala Gln Thr Arg Gln Gln Ile Gly Val Phe Glu Gln Arg Tyr
 405 410 415

Pro Gly Ile Arg Leu Ala Gly Asn Tyr Val Thr Gly Val Ala Val Pro
 420 425 430

Ala Ile Ile Gln Ala Ala Lys Gly Tyr Cys
 435 440

<210> 164
 <211> 439
 <212> PRT
 <213> Xanthomonas campestris

<400> 164

Gln Pro Val Ile Ile Ala Gly Ala Gly Ile Ala Gly Leu Ser Ile Ala
 1 5 10 15

Tyr Glu Leu Gln Gln Lys Gly Ile Pro Tyr Glu Ile Met Glu Ala Ser
 20 25 30

Ser Tyr Ala Gly Gly Val Val Lys Ser Leu His Ile Asp Gly Tyr Glu
 35 40 45

Leu Asp Ala Gly Pro Asn Ser Leu Ala Ala Ser Ala Ala Phe Met Ala
 50 55 60

Tyr Ile Asp Gln Leu Gly Leu Gln Asp Gln Val Leu Glu Ala Ala Ala
 65 70 75 80

Ala Ser Lys Asn Arg Phe Leu Val Arg Asn Asp Lys Leu His Ala Val
 85 90 95

Ser Pro His Pro Phe Lys Ile Leu Gln Ser Ala Tyr Ile Ser Gly Gly
 100 105 110

Ala Lys Trp Arg Leu Phe Thr Glu Arg Phe Arg Lys Ala Ala Ala Pro
 115 120 125

Glu Gly Glu Glu Thr Val Ser Ser Phe Val Thr Arg Arg Phe Gly Lys
 130 135 140

Glu Ile Asn Asp Tyr Leu Phe Glu Pro Val Leu Ser Gly Ile Tyr Ala
 145 150 155 160

Gly Asn Pro Asp Leu Met Ser Val Gly Glu Val Leu Pro Met Leu Pro
 165 170 175

Gln Trp Glu Gln Lys Tyr Gly Ser Val Thr Gln Gly Leu Leu Lys Asn
 180 185 190

Lys Gly Ala Met Gly Gly Arg Lys Ile Ile Ala Phe Lys Gly Gly Asn
 195 200 205

Ala Thr Leu Thr Asn Arg Leu Gln Ser Leu Leu Ser Gly Lys Ile Arg
 210 215 220

Phe Asn Cys Ala Val Thr Gly Val Thr Arg Gly Ala Asp Asp Tyr Ile
 225 230 235 240

Val Gln Tyr Thr Glu Asn Gly Asn Thr Ala Met Leu Asn Ala Ser Arg
 245 250 255

Val Ile Phe Thr Thr Pro Ala Tyr Ser Thr Ala Val Ala Ile Gln Ala
 260 265 270

Leu Asp Ala Ser Leu Ala Thr His Leu Ser Asp Val Pro Tyr Pro Arg
 275 280 285

Met Gly Val Leu His Leu Gly Phe Gly Ala Glu Ala Arg Gln Lys Ala
 290 295 300

Pro Ala Gly Phe Gly Phe Leu Val Pro His Ala Ala Gly Lys His Phe
 305 310 315 320

Leu Gly Ala Ile Cys Asn Ser Ala Ile Phe Pro Ser Arg Val Pro Thr

325 330 335
 Gly Lys Val Leu Phe Thr Val Phe Leu Gly Gly Ala Arg Gln Glu Gln
 340 345 350
 Leu Phe Asp Gln Leu Gly Pro Glu Lys Leu Gln Gln Thr Val Val Lys
 355 360 365
 Glu Leu Met Glu Leu Leu Gly Leu Thr Thr Pro Pro Glu Met Gln Arg
 370 375 380
 Phe Ser Glu Trp Asn Arg Ala Ile Pro Gln Leu Asn Val Gly Tyr Ala
 385 390 395 400
 Gln Thr Arg Gln Gln Ile Gly Val Phe Glu Gln Arg Tyr Pro Gly Ile
 405 410 415
 Arg Leu Ala Gly Asn Tyr Val Thr Gly Val Ala Val Pro Ala Ile Ile
 420 425 430
 Gln Ala Ala Lys Gly Tyr Cys
 435
 <210> 165
 <211> 448
 <212> PRT
 <213> Chitinophaga pinensis
 <400> 165
 Met Ser Asp Gln Pro Val Leu Ile Val Gly Ala Gly Leu Ser Gly Leu
 1 5 10 15
 Ser Ile Ala Tyr Glu Leu Gln Lys Leu Gln Val Pro Tyr Gln Val Leu
 20 25 30
 Glu Val Ser Gly His Ser Gly Gly Val Met Lys Ser Leu Arg Lys Asp
 35 40 45
 Gly Phe Glu Leu Asp Ala Gly Ala Asn Thr Ile Ala Ala Ser Pro Glu
 50 55 60
 Ile Leu Ala Tyr Phe Thr Ser Leu Gly Leu Glu Asn Glu Ile Leu Gln
 65 70 75 80
 Ala Thr Ala Ala Ser Lys His Arg Phe Leu Val Arg Arg Arg Gln Leu

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

Met His Phe Leu Gly Ala Ile Cys Asn Ala Ala Ile Phe Pro Asp Lys
325 330 335

Ala Pro Pro Gly Lys Ile Leu Phe Thr Val Phe Leu Gly Gly Ala Arg
340 345 350

Gln Glu Ser Leu Phe Asp Gln Met Thr Pro Glu Ala Leu Gln Gln Gln
355 360 365

Val Val Ser Glu Val Met Ser Leu Leu His Leu Ser Ala Pro Pro Val
370 375 380

Met Gln His Phe Ser Ser Trp Asn Lys Ala Ile Pro Gln Leu Asn Val
385 390 395 400

Gly His Val Lys Leu Arg Arg Ala Val Glu Ala Phe Glu Lys Lys Tyr
405 410 415

Pro Gly Ile His Leu Ser Gly Asn Tyr Leu Gln Gly Val Ala Ile Pro
420 425 430

Ala Leu Leu Gln His Ala Ala Ala Leu Ala Ala Ser Leu Lys Lys Asn
435 440 445

<210> 166

<211> 445

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 166

Gln Pro Val Leu Ile Val Gly Ala Gly Leu Ser Gly Leu Ser Ile Ala
1 5 10 15

Tyr Glu Leu Gln Lys Leu Gln Val Pro Tyr Gln Val Leu Glu Val Ser
20 25 30

Gly His Ser Gly Gly Val Met Lys Ser Leu Arg Lys Asp Gly Phe Glu
35 40 45

Leu Asp Ala Gly Ala Asn Thr Ile Ala Thr Ser Pro Glu Ile Leu Ala
50 55 60

Tyr Phe Thr Ser Leu Gly Leu Glu Asn Glu Ile Leu Gln Ala Thr Ala

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

Leu Asp Gly Phe Gly Phe Leu Val Pro Asn Ala Glu Asn Met His Phe
305 310 315 320

Leu Gly Ala Ile Cys Asn Ala Ala Ile Phe Pro Asp Lys Ala Pro Pro
325 330 335

Gly Lys Ile Leu Phe Thr Val Phe Leu Gly Gly Ala Arg Gln Glu Ser
340 345 350

Leu Phe Asp Gln Met Thr Pro Glu Ala Leu Gln Gln Gln Val Val Ser
355 360 365

Glu Val Met Ser Leu Leu His Leu Ser Ala Pro Pro Val Met Gln His
370 375 380

Phe Ser Ser Trp Asn Lys Ala Ile Pro Gln Leu Asn Val Gly His Val
385 390 395 400

Lys Leu Arg Arg Ala Val Glu Ala Phe Glu Lys Lys Tyr Pro Gly Ile
405 410 415

His Leu Ser Gly Asn Tyr Leu Gln Gly Val Ala Ile Pro Ala Leu Leu
420 425 430

Gln His Ala Ala Ala Leu Ala Ala Ser Leu Lys Lys Asn
435 440 445

<210> 167

<211> 470

<212> PRT

<213> Bacillus subtilis

<400> 167

Met Ser Asp Gly Lys Lys His Val Val Ile Ile Gly Gly Gly Ile Thr
1 5 10 15

Gly Leu Ala Ala Ala Phe Tyr Met Glu Lys Glu Ile Lys Glu Lys Asn
20 25 30

Leu Pro Leu Glu Leu Thr Leu Val Glu Ala Ser Pro Arg Val Gly Gly
35 40 45

Lys Ile Gln Thr Val Lys Lys Asp Gly Tyr Ile Ile Glu Arg Gly Pro
50 55 60

Asp Ser Phe Leu Glu Arg Lys Lys Ser Ala Pro Gln Leu Val Lys Asp
65 70 75 80

Leu Gly Leu Glu His Leu Leu Val Asn Asn Ala Thr Gly Gln Ser Tyr
85 90 95

Val Leu Val Asn Arg Thr Leu His Pro Met Pro Lys Gly Ala Val Met
100 105 110

Gly Ile Pro Thr Lys Ile Ala Pro Phe Val Ser Thr Gly Leu Phe Ser
115 120 125

Leu Ser Gly Lys Ala Arg Ala Ala Met Asp Phe Ile Leu Pro Ala Ser
130 135 140

Lys Thr Lys Asp Asp Gln Ser Leu Gly Glu Phe Phe Arg Arg Arg Val
145 150 155 160

Gly Asp Glu Val Val Glu Asn Leu Ile Glu Pro Leu Leu Ser Gly Ile
165 170 175

Tyr Ala Gly Asp Ile Asp Lys Leu Ser Leu Met Ser Thr Phe Pro Gln
180 185 190

Phe Tyr Gln Thr Glu Gln Lys His Arg Ser Leu Ile Leu Gly Met Lys
195 200 205

Lys Thr Arg Pro Gln Gly Ser Gly Gln Gln Leu Thr Ala Lys Lys Gln
210 215 220

Gly Gln Phe Gln Thr Leu Ser Thr Gly Leu Gln Thr Leu Val Glu Glu
225 230 235 240

Ile Glu Lys Gln Leu Lys Leu Thr Lys Val Tyr Lys Gly Thr Lys Val
245 250 255

Thr Lys Leu Ser His Ser Gly Ser Gly Tyr Ser Leu Glu Leu Asp Asn
260 265 270

Gly Val Thr Leu Asp Ala Asp Ser Val Ile Val Thr Ala Pro His Lys
275 280 285

Ala Ala Ala Gly Met Leu Ser Glu Leu Pro Ala Ile Ser His Leu Lys
290 295 300

Asn Met His Ser Thr Ser Val Ala Asn Val Ala Leu Gly Phe Pro Glu
305 310 315 320

Gly Ser Val Gln Met Glu His Glu Gly Thr Gly Phe Val Ile Ser Arg
325 330 335

Asn Ser Asp Phe Ala Ile Thr Ala Cys Thr Trp Thr Asn Lys Lys Trp
340 345 350

Pro His Ala Ala Pro Glu Gly Lys Thr Leu Leu Arg Ala Tyr Val Gly
355 360 365

Lys Ala Gly Asp Glu Ser Ile Val Asp Leu Ser Asp Asn Asp Ile Ile
370 375 380

Asn Ile Val Leu Glu Asp Leu Lys Lys Val Met Asn Ile Asn Gly Glu
385 390 395 400

Pro Glu Met Thr Cys Val Thr Arg Trp His Glu Ser Met Pro Gln Tyr
405 410 415

His Val Gly His Lys Gln Arg Ile Lys Glu Leu Arg Glu Ala Leu Ala
420 425 430

Ser Ala Tyr Pro Gly Val Tyr Met Thr Gly Ala Ser Phe Glu Gly Val
435 440 445

Gly Ile Pro Asp Cys Ile Asp Gln Gly Lys Ala Ala Val Ser Asp Ala
450 455 460

Leu Thr Tyr Leu Phe Ser
465 470

<210> 168

<211> 470

<212> PRT

<213> Bacillus pumilus

<400> 168

Met His Asp Asn Gln Lys His Leu Val Ile Ile Gly Gly Gly Ile Thr
1 5 10 15

Gly Leu Ala Ala Ala Phe Tyr Leu Glu Lys Glu Val Glu Glu Lys Gly
20 25 30

Leu Pro Ile Gln Ile Ser Leu Ile Glu Ala Ser Pro Arg Leu Gly Gly
 35 40 45

Lys Ile Gln Thr Leu Tyr Lys Asp Gly Tyr Ile Ile Glu Arg Gly Pro
 50 55 60

Asp Ser Phe Leu Glu Arg Lys Val Ser Gly Pro Gln Leu Ala Lys Asp
 65 70 75 80

Val Gly Leu Ser Asp Gln Leu Val Asn Asn Glu Thr Gly Gln Ala Tyr
 85 90 95

Val Leu Val Asn Glu Lys Leu His Pro Met Pro Lys Gly Ala Val Met
 100 105 110

Gly Ile Pro Thr Gln Ile Ser Pro Phe Ile Thr Thr Gly Leu Phe Ser
 115 120 125

Val Ala Gly Lys Ala Arg Ala Ala Met Asp Phe Val Leu Pro Lys Ser
 130 135 140

Lys Gln Thr Glu Asp Gln Ser Leu Gly Glu Phe Phe Arg Arg Arg Val
 145 150 155 160

Gly Asp Glu Val Val Glu Asn Leu Ile Glu Pro Leu Leu Ser Gly Ile
 165 170 175

Tyr Ala Gly Asp Ile Asp Arg Leu Ser Leu Met Ser Thr Phe Pro Gln
 180 185 190

Phe Tyr Gln Thr Glu Gln Gln His Arg Ser Leu Ile Leu Gly Met Lys
 195 200 205

Lys Ser Gln Gln His Ala Lys Ala Gln Gln Val Thr Ala Lys Lys Gln
 210 215 220

Gly Gln Phe Gln Thr Ile Asn Gln Gly Leu Gln Ser Leu Val Glu Ala
 225 230 235 240

Val Glu Gly Lys Leu Lys Leu Thr Thr Val Tyr Lys Gly Thr Lys Val
 245 250 255

Lys Gln Ile Glu Lys Thr Asp Gly Gly Tyr Gly Leu Gln Leu Asp Ser
 260 265 270

Gly Gln Thr Leu Phe Ala Asp Ser Ala Ile Val Thr Thr Pro His Gln
 275 280 285

Ser Ile Tyr Ser Met Phe Pro Lys Glu Ala Gly Leu Glu Tyr Leu His
 290 295 300

Asp Met Thr Ser Thr Ser Val Ala Thr Val Ala Leu Gly Phe Lys Asp
 305 310 315 320

Glu Asp Val His Asn Glu Tyr Asp Gly Thr Gly Phe Val Ile Ser Arg
 325 330 335

Asn Ser Asp Phe Ser Ile Thr Ala Cys Thr Trp Thr Asn Lys Lys Trp
 340 345 350

Pro His Thr Ala Pro Lys Gly Lys Thr Leu Leu Arg Ala Tyr Val Gly
 355 360 365

Lys Ala Gly Asp Glu Ser Ile Val Glu Gln Ser Asp Ser Gln Ile Val
 370 375 380

Ser Ile Val Leu Glu Asp Leu Lys Lys Ile Met Asp Ile Lys Ala Asp
 385 390 395 400

Pro Glu Leu Thr Thr Val Thr Arg Trp Lys Thr Ser Met Pro Gln Tyr
 405 410 415

His Val Gly His Gln Lys Ala Ile Ser Asn Met Arg Glu Thr Phe Lys
 420 425 430

Gln Ser Tyr Pro Gly Val Tyr Ile Thr Gly Ala Ala Phe Glu Gly Val
 435 440 445

Gly Ile Pro Asp Cys Ile Asp Gln Gly Lys Ala Ala Ile Ser Glu Ala
 450 455 460

Val Ser Tyr Leu Phe Ser
 465 470

<210> 169

<211> 470

<212> PRT

<213> Bacillus pumilus

<400> 169

Met His Asp Asn Gln Lys His Leu Val Ile Ile Gly Gly Gly Ile Thr
 1 5 10 15

Gly Leu Ala Ala Ala Phe Tyr Leu Glu Lys Glu Val Glu Glu Lys Gly
 20 25 30

Leu Pro Ile Gln Ile Ser Leu Ile Glu Ala Ser Pro Arg Leu Gly Gly
 35 40 45

Lys Ile Gln Thr Leu Tyr Lys Asp Gly Tyr Ile Ile Glu Arg Gly Pro
 50 55 60

Asp Ser Phe Leu Glu Arg Lys Val Ser Gly Pro Gln Leu Ala Lys Asp
 65 70 75 80

Val Gly Leu Ser Asp Gln Leu Val Asn Asn Glu Thr Gly Gln Ala Tyr
 85 90 95

Val Leu Val Asn Glu Thr Leu His Pro Met Pro Lys Gly Ala Val Met
 100 105 110

Gly Ile Pro Thr Gln Ile Ser Pro Phe Ile Thr Thr Gly Leu Phe Ser
 115 120 125

Val Ala Gly Lys Ala Arg Ala Ala Met Asp Phe Val Leu Pro Lys Ser
 130 135 140

Lys Gln Thr Glu Asp Gln Ser Leu Gly Glu Phe Phe Arg Arg Arg Val
 145 150 155 160

Gly Asp Glu Val Val Glu Asn Leu Ile Glu Pro Leu Leu Ser Gly Ile
 165 170 175

Tyr Ala Gly Asp Ile Asp Arg Leu Ser Leu Met Ser Thr Phe Pro Gln
 180 185 190

Phe Tyr Gln Thr Glu Gln Lys His Arg Ser Leu Ile Leu Gly Met Lys
 195 200 205

Lys Ser Gln Gln His Ala Lys Ala Gln Gln Val Thr Ala Lys Lys Gln
 210 215 220

Gly Gln Phe Gln Thr Ile Asn Gln Gly Leu Gln Ala Leu Val Glu Ala

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

Val Ser Tyr Leu Phe Ser
465 470

<210> 170
<211> 477
<212> PRT
<213> Paenibacillus macerans

<400> 170

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

Val Ala Tyr Tyr Val Arg Lys Leu Leu Arg Glu Gln Gly Val Asn Ala
20 25 30

Gly Val Thr Leu Val Glu Gln Ser Asp Arg Leu Gly Gly Lys Ile Arg
35 40 45

Ser Leu Arg Arg Asp Gly Phe Thr Ile Glu Gln Gly Pro Asp Ser Met
50 55 60

Ile Ala Arg Lys Pro Ala Ala Leu Glu Leu Ile Arg Glu Leu Gly Leu
65 70 75 80

Glu Asp Lys Leu Ala Gly Thr Asn Pro Gln Ala Lys Arg Ser Tyr Ile
85 90 95

Leu His Arg Gly Lys Phe His Pro Met Pro Pro Gly Leu Met Leu Gly
100 105 110

Ile Pro Thr Gln Met Trp Pro Met Val Lys Thr Gly Leu Leu Ser Pro
115 120 125

Ala Gly Lys Leu Arg Ala Ala Met Asp Leu Leu Leu Pro Ala Arg Arg
130 135 140

Gly Gly Gly Asp Glu Ser Leu Gly Gly Phe Ile Arg Arg Arg Leu Gly
145 150 155 160

Arg Glu Val Leu Glu Gln Met Thr Glu Pro Leu Leu Ala Gly Ile Tyr
165 170 175

Ala Gly Asp Thr Glu Gln Leu Ser Leu Lys Ala Thr Phe Pro Gln Phe
180 185 190

Met Glu Met Glu Arg Lys His Arg Ser Leu Ile Leu Gly Leu Leu Ala
 195 200 205

Gly Lys Lys Gln Pro Pro Arg Pro Gly Gly Ser Gln Val Pro Leu Pro
 210 215 220

Lys Ala Ala Gln Thr Ser Met Phe Leu Thr Leu Thr Gly Gly Leu Glu
 225 230 235 240

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

Thr Gly Gln Ala Val Thr Gly Leu Ser Gln Gln Glu Ala Gly Tyr Glu
 260 265 270

Leu Asn Leu Ser Gly Gly Glu Arg Leu Asn Ala Asp Gly Val Ile Leu
 275 280 285

Ala Val Pro Ala Phe Ala Ala Ala Arg Leu Leu Asp Gly Val Pro Glu
 290 295 300

Ala Ala Tyr Leu Glu Arg Ile Arg Tyr Val Ser Val Ala Asn Leu Ala
 305 310 315 320

Phe Ala Tyr Arg Arg Glu Asp Val Pro His Asp Leu Asn Gly Ser Gly
 325 330 335

Val Leu Ile Pro Arg Gly Glu Gly Arg Met Ile Thr Ala Ile Thr Trp
 340 345 350

Val Ser Ser Lys Trp Leu His Ser Ala Pro Gly Asp Lys Ala Leu Leu
 355 360 365

Arg Ala Tyr Ile Gly Arg Leu Gly Asp Glu Ala Trp Thr Ala Met Cys
 370 375 380

Arg Ala Asp Ile Glu Arg Arg Val Ala Ala Glu Leu Arg Asp Leu Leu
 385 390 395 400

Gly Ile Ala Ala Ser Pro Leu Phe Cys Glu Leu Ala Ala Leu Pro Glu
 405 410 415

Ser Met Pro Gln Tyr Pro Val Gly His Val Glu Arg Leu Glu Ala Leu
 420 425 430

Arg Gly Ala Leu Cys Arg Ala Lys Pro Gly Leu Leu Leu Cys Gly Ala
 435 440 445

Gly Tyr Ala Gly Val Gly Ile Pro Asp Cys Ile Arg Gln Gly Lys Glu
 450 455 460

Ala Ala Glu Ser Met Ala Ala Tyr Leu Arg Asp Gly Arg
 465 470 475

<210> 171
 <211> 493
 <212> PRT
 <213> Paenibacillus thiaminolyticus

<400> 171

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

Leu Ser Ala Ala Phe Tyr Ala Leu Lys Gln Ala Asp Glu Glu Gly Gln
 20 25 30

Pro Ile Ser Val Thr Ile Ile Glu Gln Ser Asp Arg Leu Gly Gly Lys
 35 40 45

Ile Gln Thr Leu Arg Lys Glu Gly Cys Val Ile Glu Lys Gly Pro Asp
 50 55 60

Ser Phe Leu Ala Arg Lys Leu Pro Met Ile Asp Leu Ala Arg Asp Leu
 65 70 75 80

Gly Met Asp Ser Glu Leu Val Ala Thr Asn Pro His Ala Lys Lys Thr
 85 90 95

Tyr Ile Leu Arg Arg Gly Lys Leu Tyr Arg Met Pro Pro Gly Leu Val
 100 105 110

Leu Gly Ile Pro Thr Glu Leu Gly Pro Phe Ala Lys Thr Gly Leu Ile
 115 120 125

Ser Pro Trp Gly Lys Leu Arg Ala Ala Met Asp Leu Phe Ile Lys Pro
 130 135 140

His Pro Ala Asp Glu Asp Glu Ser Val Gly Ala Phe Leu Asp Arg Arg
 145 150 155 160

Leu Gly Arg Glu Val Thr Glu His Ile Ala Glu Pro Leu Leu Ala Gly
 165 170 175

Ile Tyr Ala Gly Asp Leu Gln Ala Leu Ser Leu Gln Ala Thr Phe Pro
 180 185 190

Gln Phe Ala Gln Val Glu Arg Lys His Gly Gly Leu Ile Arg Gly Met
 195 200 205

Lys Ala Ser Arg Gln Ala Gly Gln Ser Val Pro Gly Leu Pro Asp Val
 210 215 220

Ala Lys Gly Thr Met Phe Leu Thr Phe Arg Asn Gly Leu Thr Ser Leu
 225 230 235 240

Val Glu Arg Leu Glu Glu Thr Leu Arg Asp Arg Ala Glu Leu Cys Leu
 245 250 255

Gly Ile Gly Ala Glu Gly Phe Glu Lys Arg Glu Asp Gly Thr Tyr Leu
 260 265 270

Val Arg Leu Ser Asp Gly Ser Arg Leu Gln Ala Asp Ala Val Ile Val
 275 280 285

Thr Thr Pro Ser Tyr His Ala Ala Ser Leu Leu Glu Glu His Val Asp
 290 295 300

Ala Ser Ala Leu Gln Ala Ile Arg His Val Ser Val Ala Asn Val Val
 305 310 315 320

Ser Val Phe Asp Arg Lys Gln Val Asn Asn Gln Phe Asp Gly Thr Gly
 325 330 335

Phe Val Ile Ser Arg Arg Glu Gly Arg Ala Ile Thr Ala Cys Thr Trp
 340 345 350

Thr Ser Val Lys Trp Pro His Thr Ser Arg Gly Asp Lys Leu Ile Ile
 355 360 365

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

Asp Glu Ala Leu Lys Arg Thr Val Arg Ser Glu Leu Arg Glu Leu Leu
 385 390 395 400

Asp Ile Asp Ile Asp Pro Glu Phe Val Glu Ile Thr Arg Leu Arg His
 405 410 415

Ser Met Pro Gln Tyr Pro Val Gly His Val Gln Ala Ile Arg Ser Leu
 420 425 430

Arg Asp Glu Val Gly Arg Thr Leu Pro Gly Val Phe Leu Ala Gly Gln
 435 440 445

Pro Tyr Glu Gly Val Gly Met Pro Asp Cys Val Arg Ser Gly Arg Asp
 450 455 460

Ala Ala Glu Ala Ala Val Ser Ala Met Gln Ala Met Ser Thr Glu Pro
 465 470 475 480

Glu Ala Pro Ala Glu Asp Ala Ala Thr Gly Thr Ala Gly
 485 490

<210> 172

<211> 474

<212> PRT

<213> Paenibacillus polymyxa

<400> 172

Met Gly Asp Lys Lys Arg Arg Val Val Val Val Gly Gly Gly Leu Thr
 1 5 10 15

Gly Leu Ser Ala Ala Phe Tyr Ile Arg Lys His Tyr Arg Glu Ala Gly
 20 25 30

Val Glu Pro Val Ile Thr Leu Val Glu Lys Ser Ser Ser Met Gly Gly
 35 40 45

Met Ile Glu Thr Leu His Arg Asp Gly Phe Val Ile Glu Lys Gly Pro
 50 55 60

Asp Ser Phe Leu Ala Arg Lys Thr Ala Met Ile Asp Leu Ala Lys Glu
 65 70 75 80

Leu Glu Ile Asp His Glu Leu Val Ser Gln Asn Pro Glu Ser Lys Lys
 85 90 95

Thr Tyr Ile Met Gln Arg Gly Lys Leu His Pro Met Pro Ala Gly Leu
 100 105 110

Val Leu Gly Ile Pro Thr Glu Leu Arg Pro Phe Leu Arg Ser Gly Leu
115 120 125

Val Ser Pro Ala Gly Lys Leu Arg Ala Leu Met Asp Phe Val Ile Pro
130 135 140

Pro Arg Arg Thr Thr Glu Asp Glu Ser Leu Gly Tyr Met Ile Glu Arg
145 150 155 160

Arg Leu Gly Ala Glu Val Leu Glu Asn Leu Thr Glu Pro Leu Leu Ala
165 170 175

Gly Ile Tyr Ala Gly Asp Met Arg Arg Leu Ser Leu Gln Ala Thr Phe
180 185 190

Pro Gln Phe Gly Glu Val Glu Arg Asp Tyr Gly Ser Leu Ile Arg Gly
195 200 205

Met Met Thr Gly Arg Lys Pro Ala Glu Thr His Thr Gly Thr Lys Arg
210 215 220

Ser Ala Phe Leu Asn Phe Arg Gln Gly Leu Gln Ser Leu Val His Ala
225 230 235 240

Leu Val His Glu Leu Gln Asp Val Asp Gln Arg Leu Asn Thr Ala Val
245 250 255

Lys Ser Leu Gln Arg Leu Asp Gly Ala Gln Thr Arg Tyr Arg Val Glu
260 265 270

Leu Gly Asn Gly Glu Met Leu Glu Ala Asp Asp Val Val Val Thr Val
275 280 285

Pro Thr Tyr Val Ala Ser Glu Leu Leu Lys Pro His Val Asp Thr Ala
290 295 300

Ala Leu Asp Ala Ile Asn Tyr Val Ser Val Ala Asn Val Val Leu Ala
305 310 315 320

Phe Glu Lys Lys Glu Val Glu His Val Phe Asp Gly Ser Gly Phe Leu
325 330 335

Val Pro Arg Lys Glu Gly Arg Asn Ile Thr Ala Cys Thr Trp Thr Ser

340 345 350
 Thr Lys Trp Leu His Thr Ser Pro Asp Asp Lys Val Leu Leu Arg Cys
 355 360 365
 Tyr Val Gly Arg Ser Gly Asp Glu Gln Asn Val Glu Leu Pro Asp Glu
 370 375 380
 Ala Leu Thr Asn Leu Val Leu Lys Asp Leu Arg Glu Thr Met Gly Ile
 385 390 395 400
 Glu Ala Val Pro Ile Phe Ser Glu Ile Thr Arg Leu Arg Lys Ser Met
 405 410 415
 Pro Gln Tyr Pro Val Gly His Leu Gln His Ile Ala Ala Leu Arg Glu
 420 425 430
 Glu Leu Gly Ser Lys Leu Pro Gly Val Tyr Ile Ala Gly Ala Gly Tyr
 435 440 445
 Glu Gly Val Gly Leu Pro Asp Cys Ile Arg Gln Ala Lys Glu Met Ser
 450 455 460
 Val Gln Ala Thr Gln Glu Leu Ala Ala Asp
 465 470
 <210> 173
 <211> 470
 <212> PRT
 <213> Bacillus atrophaeus
 <400> 173
 Met Ser Asp Gly Lys Lys His Leu Val Ile Ile Gly Gly Gly Ile Thr
 1 5 10 15
 Gly Leu Ala Ser Ala Phe Tyr Met Glu Lys Glu Ile Arg Glu Lys Asn
 20 25 30
 Leu Pro Leu Ser Val Thr Leu Val Glu Ala Ser Pro Arg Val Gly Gly
 35 40 45
 Lys Ile Gln Thr Ala Arg Lys Asp Gly Tyr Ile Ile Glu Arg Gly Pro
 50 55 60
 Asp Ser Phe Leu Glu Arg Lys Lys Ser Ala Pro Glu Leu Val Glu Asp

65					70					75					80
Leu	Gly	Leu	Glu	His 85	Leu	Leu	Val	Asn	Asn 90	Ala	Thr	Gly	Gln	Ser 95	Tyr
Val	Leu	Val	Asn 100	Glu	Thr	Leu	His	Pro 105	Met	Pro	Lys	Gly	Ala 110	Val	Met
Gly	Ile	Pro 115	Thr	Lys	Ile	Ala	Pro 120	Phe	Met	Ser	Thr	Gly 125	Leu	Phe	Ser
Phe	Ser 130	Gly	Lys	Ala	Arg	Ala 135	Ala	Met	Asp	Phe	Val 140	Leu	Pro	Ala	Ser
Lys 145	Pro	Lys	Glu	Asp	Gln 150	Ser	Leu	Gly	Glu	Phe 155	Phe	Arg	Arg	Arg	Val 160
Gly	Asp	Glu	Val	Val 165	Glu	Asn	Leu	Ile	Glu 170	Pro	Leu	Leu	Ser	Gly 175	Ile
Tyr	Ala	Gly	Asp 180	Ile	Asp	Arg	Leu	Ser 185	Leu	Met	Ser	Thr	Phe 190	Pro	Gln
Phe	Tyr	Gln 195	Thr	Glu	Gln	Lys	His 200	Arg	Ser	Leu	Ile	Leu	Gly 205	Met	Lys
Lys	Thr 210	Arg	Pro	Gln	Gly	Ser 215	Gly	Gln	Arg	Leu	Thr 220	Ala	Lys	Lys	Gln
Gly 225	Gln	Phe	Gln	Thr	Leu 230	Lys	Thr	Gly	Leu	Gln 235	Thr	Leu	Val	Glu	Glu 240
Leu	Glu	Asn	Gln	Leu 245	Lys	Leu	Thr	Lys	Val 250	Tyr	Lys	Gly	Thr	Lys 255	Val
Thr	Asn	Ile	Ser 260	Arg	Gly	Glu	Lys	Gly 265	Cys	Ser	Ile	Ala	Leu 270	Asp	Asn
Gly	Met	Thr 275	Leu	Asp	Ala	Asp	Ala 280	Ala	Ile	Val	Thr	Ser 285	Pro	His	Lys
Ser	Ala 290	Ala	Gly	Met	Phe	Pro 295	Asp	Leu	Pro	Ala	Val 300	Ser	Gln	Leu	Lys

Asp Met His Ser Thr Ser Val Ala Asn Val Ala Leu Gly Phe Pro Gln
 305 310 315 320

Glu Ala Val Gln Met Glu His Glu Gly Thr Gly Phe Val Ile Ser Arg
 325 330 335

Asn Ser Asp Phe Ser Ile Thr Ala Cys Thr Trp Thr Asn Lys Lys Trp
 340 345 350

Pro His Ser Ala Pro Glu Gly Lys Thr Leu Leu Arg Ala Tyr Val Gly
 355 360 365

Lys Ala Gly Asp Glu Ser Ile Val Glu Leu Ser Asp Asn Glu Ile Ile
 370 375 380

Lys Ile Val Leu Glu Asp Leu Lys Lys Val Met Lys Ile Lys Gly Glu
 385 390 395 400

Pro Glu Met Thr Cys Val Thr Arg Trp Asn Glu Ser Met Pro Gln Tyr
 405 410 415

His Val Gly His Lys Gln Arg Ile Lys Lys Val Arg Glu Ala Leu Ala
 420 425 430

Ala Ser Tyr Pro Gly Val Tyr Met Thr Gly Ala Ser Phe Glu Gly Val
 435 440 445

Gly Ile Pro Asp Cys Ile Asp Gln Gly Lys Ser Ala Val Ser Asp Val
 450 455 460

Leu Ala Tyr Leu Phe Gly
 465 470

<210> 174

<211> 470

<212> PRT

<213> Bacillus atrophaeus

<400> 174

Met Ser Asp Gly Lys Lys His Leu Val Ile Ile Gly Gly Gly Ile Thr
 1 5 10 15

Gly Leu Ala Ser Ala Phe Tyr Met Glu Lys Glu Ile Arg Glu Lys Asn
 20 25 30

Leu Pro Leu Ser Val Thr Leu Val Glu Ala Ser Pro Arg Val Gly Gly
 35 40 45

Lys Ile Gln Thr Ala Arg Lys Asp Gly Tyr Ile Ile Glu Arg Gly Pro
 50 55 60

Asp Ser Phe Leu Glu Arg Lys Lys Ser Ala Pro Glu Leu Val Glu Asp
 65 70 75 80

Leu Gly Leu Glu His Leu Leu Val Asn Asn Ala Thr Gly Gln Ser Tyr
 85 90 95

Val Leu Val Asn Glu Thr Leu His Pro Met Pro Lys Gly Ala Val Met
 100 105 110

Gly Ile Pro Thr Lys Ile Ala Pro Phe Met Ser Thr Arg Leu Phe Ser
 115 120 125

Phe Ser Gly Lys Ala Arg Ala Ala Met Asp Phe Val Leu Pro Ala Ser
 130 135 140

Lys Pro Lys Glu Asp Gln Ser Leu Gly Glu Phe Phe Arg Arg Arg Val
 145 150 155 160

Gly Asp Glu Val Val Glu Asn Leu Ile Glu Pro Leu Leu Ser Gly Ile
 165 170 175

Tyr Ala Gly Asp Ile Asp Arg Leu Ser Leu Met Ser Thr Phe Pro Gln
 180 185 190

Phe Tyr Gln Thr Glu Gln Lys His Arg Ser Leu Ile Leu Gly Met Lys
 195 200 205

Lys Thr Arg Pro Gln Gly Ser Gly Gln Gln Leu Thr Ala Lys Lys Gln
 210 215 220*

Gly Gln Phe Gln Thr Leu Lys Thr Gly Leu Gln Thr Leu Val Glu Glu
 225 230 235 240

Leu Glu Asn Gln Leu Lys Leu Thr Lys Val Tyr Lys Gly Thr Lys Val
 245 250 255

Thr Asn Ile Ser Arg Gly Glu Lys Gly Cys Ser Ile Ala Leu Asp Asn
 260 265 270

Gly Met Thr Leu Asp Ala Asp Ala Ala Ile Val Thr Ser Pro His Lys
275 280 285

Ser Ala Ala Gly Met Phe Pro Asp Leu Pro Ala Val Ser Gln Leu Lys
290 295 300

Asp Met His Ser Thr Ser Val Ala Asn Val Ala Leu Gly Phe Pro Gln
305 310 315 320

Glu Ala Val Gln Met Glu His Glu Gly Thr Gly Phe Val Ile Ser Arg
325 330 335

Asn Ser Asp Phe Ser Ile Thr Ala Cys Thr Trp Thr Asn Lys Lys Trp
340 345 350

Pro His Ser Ala Pro Glu Gly Lys Thr Leu Leu Arg Ala Tyr Val Gly
355 360 365

Lys Ala Gly Asp Glu Ser Ile Val Glu Leu Ser Asp Asn Glu Ile Ile
370 375 380

Lys Ile Val Leu Glu Asp Leu Lys Lys Val Met Lys Ile Lys Gly Glu
385 390 395 400

Pro Glu Met Thr Cys Val Thr Arg Trp Asn Glu Ser Met Pro Gln Tyr
405 410 415

His Val Gly His Lys Gln Arg Ile Lys Lys Val Arg Glu Ala Leu Ala
420 425 430

Ala Ser Tyr Pro Gly Val Tyr Met Thr Gly Ala Ser Phe Glu Gly Val
435 440 445

Gly Ile Pro Asp Cys Ile Asp Gln Gly Lys Ser Ala Val Ser Asp Val
450 455 460

Leu Ala Tyr Leu Phe Glu
465 470

<210> 175

<211> 475

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 175

Lys Lys Ile Ala Val Ile Gly Gly Gly Ile Thr Gly Leu Ser Val Ala
1 5 10 15

Tyr Tyr Val Arg Lys Leu Leu Arg Glu Gln Gly Val Asn Ala Gly Val
20 25 30

Thr Leu Val Glu Gln Ser Asp Arg Leu Gly Gly Lys Ile Arg Ser Leu
35 40 45

Arg Arg Asp Gly Phe Thr Ile Glu Gln Gly Pro Asp Ser Met Ile Ala
50 55 60

Arg Lys Pro Ala Ala Leu Glu Leu Ile Arg Glu Leu Gly Leu Glu Asp
65 70 75 80

Lys Leu Ala Gly Thr Asn Pro Gln Ala Lys Arg Ser Tyr Ile Leu His
85 90 95

Arg Gly Lys Phe His Pro Met Pro Pro Gly Leu Met Leu Gly Ile Pro
100 105 110

Thr Gln Met Trp Pro Met Val Lys Thr Gly Leu Leu Ser Pro Ala Gly
115 120 125

Lys Leu Arg Ala Ala Met Asp Leu Leu Leu Pro Ala Arg Arg Gly Gly
130 135 140

Gly Asp Glu Ser Leu Gly Gly Phe Ile Arg Arg Arg Leu Gly Arg Glu
145 150 155 160

Val Leu Glu Gln Met Thr Glu Pro Leu Leu Ala Gly Ile Tyr Ala Gly
165 170 175

Asp Thr Glu Gln Leu Ser Leu Lys Ala Thr Phe Pro Gln Phe Met Glu
180 185 190

Met Glu Arg Lys His Arg Ser Leu Ile Leu Gly Leu Leu Ala Gly Lys
195 200 205

Lys Gln Pro Pro Arg Pro Gly Gly Ser Gln Val Pro Leu Pro Lys Ala
210 215 220

Ala Gln Thr Ser Met Phe Leu Thr Leu Thr Gly Gly Leu Glu Gly Leu
225 230 235 240

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

Gln Ala Val Thr Gly Leu Ser Gln Gln Glu Ala Gly Tyr Glu Leu Asn
260 265 270

Leu Ser Gly Gly Glu Arg Leu Asn Ala Asp Gly Val Ile Leu Ala Val
275 280 285

Pro Ala Phe Ala Ala Ala Arg Leu Leu Asp Gly Val Pro Glu Ala Ala
290 295 300

Tyr Leu Glu Arg Ile Arg Tyr Val Ser Val Ala Asn Leu Ala Phe Ala
305 310 315 320

Tyr Arg Arg Glu Asp Val Pro His Asp Leu Asn Gly Ser Gly Val Leu
325 330 335

Ile Pro Arg Gly Glu Gly Arg Met Ile Thr Ala Ile Thr Trp Val Ser
340 345 350

Ser Lys Trp Leu His Ser Ala Pro Gly Asp Lys Ala Leu Leu Arg Ala
355 360 365

Tyr Ile Gly Arg Leu Gly Asp Glu Ala Trp Thr Ala Met Cys Arg Ala
370 375 380

Asp Ile Glu Arg Arg Val Ala Ala Glu Leu Arg Asp Leu Leu Gly Ile
385 390 395 400

Ala Ala Ser Pro Leu Phe Cys Glu Leu Ala Ala Leu Pro Glu Ser Met
405 410 415

Pro Gln Tyr Pro Val Gly His Val Glu Arg Leu Glu Ala Leu Arg Gly
420 425 430

Ala Leu Cys Arg Ala Lys Pro Gly Leu Leu Leu Cys Gly Ala Gly Tyr
435 440 445

Ala Gly Val Gly Ile Pro Asp Cys Ile Arg Gln Gly Lys Glu Ala Ala
450 455 460

Glu Ser Met Ala Ala Tyr Leu Arg Asp Gly Arg
465 470 475

<210> 176
<211> 489
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 176

Arg Lys Leu Val Val Ile Gly Gly Gly Ile Thr Gly Leu Ser Ala Ala
1 5 10 15

Phe Tyr Ala Leu Lys Gln Ala Asp Glu Glu Gly Gln Pro Ile Ser Val
20 25 30

Thr Ile Ile Glu Gln Ser Asp Arg Leu Gly Gly Lys Ile Gln Thr Leu
35 40 45

Arg Lys Glu Gly Cys Val Ile Glu Lys Gly Pro Asp Ser Phe Leu Ala
50 55 60

Arg Lys Leu Pro Met Ile Asp Leu Ala Arg Asp Leu Gly Met Asp Ser
65 70 75 80

Glu Leu Val Ala Thr Asn Pro His Ala Lys Lys Thr Tyr Ile Leu Arg
85 90 95

Arg Gly Lys Leu Tyr Arg Met Pro Pro Gly Leu Val Leu Gly Ile Pro
100 105 110

Thr Glu Leu Gly Pro Phe Ala Lys Thr Gly Leu Ile Ser Pro Trp Gly
115 120 125

Lys Leu Arg Ala Ala Met Asp Leu Phe Ile Lys Pro His Pro Ala Asp
130 135 140

Glu Asp Glu Ser Val Gly Ala Phe Leu Asp Arg Arg Leu Gly Arg Glu
145 150 155 160

Val Thr Glu His Ile Ala Glu Pro Leu Leu Ala Gly Ile Tyr Ala Gly
165 170 175

Asp Leu Gln Ala Leu Ser Leu Gln Ala Thr Phe Pro Gln Phe Ala Gln
 180 185 190

Val Glu Arg Lys His Gly Gly Leu Ile Arg Gly Met Lys Ala Ser Arg
 195 200 205

Gln Ala Gly Gln Ser Val Pro Gly Leu Pro Asp Val Ala Lys Gly Thr
 210 215 220

Met Phe Leu Thr Phe Arg Asn Gly Leu Thr Ser Leu Val Glu Arg Leu
 225 230 235 240

Glu Glu Thr Leu Arg Asp Arg Ala Glu Leu Cys Leu Gly Ile Gly Ala
 245 250 255

Glu Gly Phe Glu Lys Arg Glu Asp Gly Thr Tyr Leu Val Arg Leu Ser
 260 265 270

Asp Gly Ser Arg Leu Gln Ala Asp Ala Val Ile Val Thr Thr Pro Ser
 275 280 285

Tyr His Ala Ala Ser Leu Leu Glu Glu His Val Asp Ala Ser Ala Leu
 290 295 300

Gln Ala Ile Arg His Val Ser Val Ala Asn Val Val Ser Val Phe Asp
 305 310 315 320

Arg Lys Gln Val Asn Asn Gln Phe Asp Gly Thr Gly Phe Val Ile Ser
 325 330 335

Arg Arg Glu Gly Arg Ala Ile Thr Ala Cys Thr Trp Thr Ser Val Lys
 340 345 350

Trp Pro His Thr Ser Arg Gly Asp Lys Leu Ile Ile Arg Cys Tyr Ile
 355 360 365

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

Lys Arg Thr Val Arg Ser Glu Leu Arg Glu Leu Leu Asp Ile Asp Ile
 385 390 395 400

Asp Pro Glu Phe Val Glu Ile Thr Arg Leu Arg His Ser Met Pro Gln
 405 410 415

Tyr Pro Val Gly His Val Gln Ala Ile Arg Ser Leu Arg Asp Glu Val
420 425 430

Gly Arg Thr Leu Pro Gly Val Phe Leu Ala Gly Gln Pro Tyr Glu Gly
435 440 445

Val Gly Met Pro Asp Cys Val Arg Ser Gly Arg Asp Ala Ala Glu Ala
450 455 460

Ala Val Ser Ala Met Gln Ala Met Ser Thr Glu Pro Glu Ala Pro Ala
465 470 475 480

Glu Asp Ala Ala Thr Gly Thr Ala Gly
485

<210> 177
<211> 469
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 177

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

Phe Tyr Ile Arg Lys His Tyr Arg Glu Ala Gly Val Glu Pro Val Ile
20 25 30

Thr Leu Val Glu Lys Ser Ser Ser Met Gly Gly Met Ile Glu Thr Leu
35 40 45

His Arg Asp Gly Phe Val Ile Glu Lys Gly Pro Asp Ser Phe Leu Ala
50 55 60

Arg Lys Thr Ala Met Ile Asp Leu Ala Lys Glu Leu Glu Ile Asp His
65 70 75 80

Glu Leu Val Ser Gln Asn Pro Glu Ser Lys Lys Thr Tyr Ile Met Gln
85 90 95

Arg Gly Lys Leu His Pro Met Pro Ala Gly Leu Val Leu Gly Ile Pro
100 105 110

Thr Glu Leu Arg Pro Phe Leu Arg Ser Gly Leu Val Ser Pro Ala Gly
 115 120 125

Lys Leu Arg Ala Leu Met Asp Phe Val Ile Pro Pro Arg Arg Thr Thr
 130 135 140

Glu Asp Glu Ser Leu Gly Tyr Met Ile Glu Arg Arg Leu Gly Ala Glu
 145 150 155 160

Val Leu Glu Asn Leu Thr Glu Pro Leu Leu Ala Gly Ile Tyr Ala Gly
 165 170 175

Asp Met Arg Arg Leu Ser Leu Gln Ala Thr Phe Pro Gln Phe Gly Glu
 180 185 190

Val Glu Arg Asp Tyr Gly Ser Leu Ile Arg Gly Met Met Thr Gly Arg
 195 200 205

Lys Pro Ala Glu Thr His Thr Gly Thr Lys Arg Ser Ala Phe Leu Asn
 210 215 220

Phe Arg Gln Gly Leu Gln Ser Leu Val His Ala Leu Val His Glu Leu
 225 230 235 240

Gln Asp Val Asp Gln Arg Leu Asn Thr Ala Val Lys Ser Leu Gln Arg
 245 250 255

Leu Asp Gly Ala Gln Thr Arg Tyr Arg Val Glu Leu Gly Asn Gly Glu
 260 265 270

Met Leu Glu Ala Asp Asp Val Val Val Thr Val Pro Thr Tyr Val Ala
 275 280 285

Ser Glu Leu Leu Lys Pro His Val Asp Thr Ala Ala Leu Asp Ala Ile
 290 295 300

Asn Tyr Val Ser Val Ala Asn Val Val Leu Ala Phe Glu Lys Lys Glu
 305 310 315 320

Val Glu His Val Phe Asp Gly Ser Gly Phe Leu Val Pro Arg Lys Glu
 325 330 335

Gly Arg Asn Ile Thr Ala Cys Thr Trp Thr Ser Thr Lys Trp Leu His
 340 345 350

Thr Ser Pro Asp Asp Lys Val Leu Leu Arg Cys Tyr Val Gly Arg Ser
355 360 365

Gly Asp Glu Gln Asn Val Glu Leu Pro Asp Glu Ala Leu Thr Asn Leu
370 375 380

Val Leu Lys Asp Leu Arg Glu Thr Met Gly Ile Glu Ala Val Pro Ile
385 390 395 400

Phe Ser Glu Ile Thr Arg Leu Arg Lys Ser Met Pro Gln Tyr Pro Val
405 410 415

Gly His Leu Gln His Ile Ala Ala Leu Arg Glu Glu Leu Gly Ser Lys
420 425 430

Leu Pro Gly Val Tyr Ile Ala Gly Ala Gly Tyr Glu Gly Val Gly Leu
435 440 445

Pro Asp Cys Ile Arg Gln Ala Lys Glu Met Ser Val Gln Ala Thr Gln
450 455 460

Glu Leu Ala Ala Asp
465

<210> 178
<211> 465
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 178

Lys His Leu Val Ile Ile Gly Gly Gly Ile Thr Gly Leu Ala Ser Ala
1 5 10 15

Phe Tyr Met Glu Lys Glu Ile Arg Glu Lys Asn Leu Pro Leu Ser Val
20 25 30

Thr Leu Val Glu Ala Ser Pro Arg Val Gly Gly Lys Ile Gln Thr Ala
35 40 45

Arg Lys Asp Gly Tyr Ile Ile Glu Arg Gly Pro Asp Ser Phe Leu Glu
50 55 60

Arg Lys Lys Ser Ala Pro Glu Leu Val Glu Asp Leu Gly Leu Glu His
65 70 75 80

Leu Leu Val Asn Asn Ala Thr Gly Gln Ser Tyr Val Leu Val Asn Glu
85 90 95

Thr Leu His Pro Met Pro Lys Gly Ala Val Met Gly Ile Pro Thr Lys
100 105 110

Ile Ala Pro Phe Met Ser Thr Arg Leu Phe Ser Phe Ser Gly Lys Ala
115 120 125

Arg Ala Ala Met Asp Phe Val Leu Pro Ala Ser Lys Pro Lys Glu Asp
130 135 140

Gln Ser Leu Gly Glu Phe Phe Arg Arg Arg Val Gly Asp Glu Val Val
145 150 155 160

Glu Asn Leu Ile Glu Pro Leu Leu Ser Gly Ile Tyr Ala Gly Asp Ile
165 170 175

Asp Arg Leu Ser Leu Met Ser Thr Phe Pro Gln Phe Tyr Gln Thr Glu
180 185 190

Gln Lys His Arg Ser Leu Ile Leu Gly Met Lys Lys Thr Arg Pro Gln
195 200 205

Gly Ser Gly Gln Gln Leu Thr Ala Lys Lys Gln Gly Gln Phe Gln Thr
210 215 220

Leu Lys Thr Gly Leu Gln Thr Leu Val Glu Glu Leu Glu Asn Gln Leu
225 230 235 240

Lys Leu Thr Lys Val Tyr Lys Gly Thr Lys Val Thr Asn Ile Ser Arg
245 250 255

Gly Glu Lys Gly Cys Ser Ile Ala Leu Asp Asn Gly Met Thr Leu Asp
260 265 270

Ala Asp Ala Ala Ile Val Thr Ser Pro His Lys Ser Ala Ala Gly Met
275 280 285

Phe Pro Asp Leu Pro Ala Val Ser Gln Leu Lys Asp Met His Ser Thr
290 295 300

Ser Val Ala Asn Val Ala Leu Gly Phe Pro Gln Glu Ala Val Gln Met
305 310 315 320

Glu His Glu Gly Thr Gly Phe Val Ile Ser Arg Asn Ser Asp Phe Ser
325 330 335

Ile Thr Ala Cys Thr Trp Thr Asn Lys Lys Trp Pro His Ser Ala Pro
340 345 350

Glu Gly Lys Thr Leu Leu Arg Ala Tyr Val Gly Lys Ala Gly Asp Glu
355 360 365

Ser Ile Val Glu Leu Ser Asp Asn Glu Ile Ile Lys Ile Val Leu Glu
370 375 380

Asp Leu Lys Lys Val Met Lys Ile Lys Gly Glu Pro Glu Met Thr Cys
385 390 395 400

Val Thr Arg Trp Asn Glu Ser Met Pro Gln Tyr His Val Gly His Lys
405 410 415

Gln Arg Ile Lys Lys Val Arg Glu Ala Leu Ala Ala Ser Tyr Pro Gly
420 425 430

Val Tyr Met Thr Gly Ala Ser Phe Glu Gly Val Gly Ile Pro Asp Cys
435 440 445

Ile Asp Gln Gly Lys Ser Ala Val Ser Asp Val Leu Ala Tyr Leu Phe
450 455 460

Glu
465

<210> 179
<211> 473
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 179

Ile Ala Val Ile Gly Gly Gly Ile Thr Gly Leu Ser Val Ala Tyr Tyr
1 5 10 15

Val Arg Lys Leu Leu Arg Glu Gln Gly Val Asn Ala Gly Val Thr Leu
 20 25 30

Val Glu Gln Ser Asp Arg Leu Gly Gly Lys Ile Arg Ser Leu Arg Arg
 35 40 45

Asp Gly Phe Thr Ile Glu Gln Gly Pro Asp Ser Met Ile Ala Arg Lys
 50 55 60

Pro Ala Ala Leu Glu Leu Ile Arg Glu Leu Gly Leu Glu Asp Lys Leu
 65 70 75 80

Ala Gly Thr Asn Pro Gln Ala Lys Arg Ser Tyr Ile Leu His Arg Gly
 85 90 95

Lys Phe His Pro Met Pro Pro Gly Leu Met Leu Gly Ile Pro Thr Gln
 100 105 110

Met Trp Pro Met Val Lys Thr Gly Leu Leu Ser Pro Ala Gly Lys Leu
 115 120 125

Arg Ala Ala Met Asp Leu Leu Leu Pro Ala Arg Arg Gly Gly Gly Asp
 130 135 140

Glu Ser Leu Gly Gly Phe Ile Arg Arg Arg Leu Gly Arg Glu Val Leu
 145 150 155 160

Glu Gln Met Thr Glu Pro Leu Leu Ala Gly Ile Tyr Ala Gly Asp Thr
 165 170 175

Glu Gln Leu Ser Leu Lys Ala Thr Phe Pro Gln Phe Met Glu Met Glu
 180 185 190

Arg Lys His Arg Ser Leu Ile Leu Gly Leu Leu Ala Gly Lys Lys Gln
 195 200 205

Pro Pro Arg Pro Gly Gly Ser Gln Val Pro Leu Pro Lys Ala Ala Gln
 210 215 220

Thr Ser Met Phe Leu Thr Leu Thr Gly Gly Leu Glu Gly Leu Thr Glu
 225 230 235 240

Ala Leu Glu Glu Ser Leu Ser Glu Glu Lys Ile Ile Thr Gly Gln Ala
 245 250 255

Val Thr Gly Leu Ser Gln Gln Glu Ala Gly Tyr Glu Leu Asn Leu Ser
260 265 270

Gly Gly Glu Arg Leu Asn Ala Asp Gly Val Ile Leu Ala Val Pro Ala
275 280 285

Phe Ala Ala Ala Arg Leu Leu Asp Gly Val Pro Glu Ala Ala Tyr Leu
290 295 300

Glu Arg Ile Arg Tyr Val Ser Val Ala Asn Leu Ala Phe Ala Tyr Arg
305 310 315 320

Arg Glu Asp Val Pro His Asp Leu Asn Gly Ser Gly Val Leu Ile Pro
325 330 335

Arg Gly Glu Gly Arg Met Ile Thr Ala Ile Thr Trp Val Ser Ser Lys
340 345 350

Trp Leu His Ser Ala Pro Gly Asp Lys Ala Leu Leu Arg Ala Tyr Ile
355 360 365

Gly Arg Leu Gly Asp Glu Ala Trp Thr Ala Met Cys Arg Ala Asp Ile
370 375 380

Glu Arg Arg Val Ala Ala Glu Leu Arg Asp Leu Leu Gly Ile Ala Ala
385 390 395 400

Ser Pro Leu Phe Cys Glu Leu Ala Ala Leu Pro Glu Ser Met Pro Gln
405 410 415

Tyr Pro Val Gly His Val Glu Arg Leu Glu Ala Leu Arg Gly Ala Leu
420 425 430

Cys Arg Ala Lys Pro Gly Leu Leu Leu Cys Gly Ala Gly Tyr Ala Gly
435 440 445

Val Gly Ile Pro Asp Cys Ile Arg Gln Gly Lys Glu Ala Ala Glu Ser
450 455 460

Met Ala Ala Tyr Leu Arg Asp Gly Arg
465 470

<210> 180

<211> 487
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 180

Leu Val Val Ile Gly Gly Gly Ile Thr Gly Leu Ser Ala Ala Phe Tyr
 1 5 10 15

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

Ile Glu Gln Ser Asp Arg Leu Gly Gly Lys Ile Gln Thr Leu Arg Lys
 35 40 45

Glu Gly Cys Val Ile Glu Lys Gly Pro Asp Ser Phe Leu Ala Arg Lys
 50 55 60

Leu Pro Met Ile Asp Leu Ala Arg Asp Leu Gly Met Asp Ser Glu Leu
 65 70 75 80

Val Ala Thr Asn Pro His Ala Lys Lys Thr Tyr Ile Leu Arg Arg Gly
 85 90 95

Lys Leu Tyr Arg Met Pro Pro Gly Leu Val Leu Gly Ile Pro Thr Glu
 100 105 110

Leu Gly Pro Phe Ala Lys Thr Gly Leu Ile Ser Pro Trp Gly Lys Leu
 115 120 125

Arg Ala Ala Met Asp Leu Phe Ile Lys Pro His Pro Ala Asp Glu Asp
 130 135 140

Glu Ser Val Gly Ala Phe Leu Asp Arg Arg Leu Gly Arg Glu Val Thr
 145 150 155 160

Glu His Ile Ala Glu Pro Leu Leu Ala Gly Ile Tyr Ala Gly Asp Leu
 165 170 175

Gln Ala Leu Ser Leu Gln Ala Thr Phe Pro Gln Phe Ala Gln Val Glu
 180 185 190

Arg Lys His Gly Gly Leu Ile Arg Gly Met Lys Ala Ser Arg Gln Ala
 195 200 205

Gly Gln Ser Val Pro Gly Leu Pro Asp Val Ala Lys Gly Thr Met Phe
 210 215 220

Leu Thr Phe Arg Asn Gly Leu Thr Ser Leu Val Glu Arg Leu Glu Glu
 225 230 235 240

Thr Leu Arg Asp Arg Ala Glu Leu Cys Leu Gly Ile Gly Ala Glu Gly
 245 250 255

Phe Glu Lys Arg Glu Asp Gly Thr Tyr Leu Val Arg Leu Ser Asp Gly
 260 265 270

Ser Arg Leu Gln Ala Asp Ala Val Ile Val Thr Thr Pro Ser Tyr His
 275 280 285

Ala Ala Ser Leu Leu Glu Glu His Val Asp Ala Ser Ala Leu Gln Ala
 290 295 300

Ile Arg His Val Ser Val Ala Asn Val Val Ser Val Phe Asp Arg Lys
 305 310 315 320

Gln Val Asn Asn Gln Phe Asp Gly Thr Gly Phe Val Ile Ser Arg Arg
 325 330 335

Glu Gly Arg Ala Ile Thr Ala Cys Thr Trp Thr Ser Val Lys Trp Pro
 340 345 350

His Thr Ser Arg Gly Asp Lys Leu Ile Ile Arg Cys Tyr Ile Gly Arg
 355 360 365

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

Thr Val Arg Ser Glu Leu Arg Glu Leu Leu Asp Ile Asp Ile Asp Pro
 385 390 395 400

Glu Phe Val Glu Ile Thr Arg Leu Arg His Ser Met Pro Gln Tyr Pro
 405 410 415

Val Gly His Val Gln Ala Ile Arg Ser Leu Arg Asp Glu Val Gly Arg
 420 425 430

Thr Leu Pro Gly Val Phe Leu Ala Gly Gln Pro Tyr Glu Gly Val Gly

435 440 445

Met Pro Asp Cys Val Arg Ser Gly Arg Asp Ala Ala Glu Ala Ala Val
450 455 460

Ser Ala Met Gln Ala Met Ser Thr Glu Pro Glu Ala Pro Ala Glu Asp
465 470 475 480

Ala Ala Thr Gly Thr Ala Gly
485

<210> 181
<211> 467
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 181

Val Val Val Val Gly Gly Gly Leu Thr Gly Leu Ser Ala Ala Phe Tyr
1 5 10 15

Ile Arg Lys His Tyr Arg Glu Ala Gly Val Glu Pro Val Ile Thr Leu
20 25 30

Val Glu Lys Ser Ser Ser Met Gly Gly Met Ile Glu Thr Leu His Arg
35 40 45

Asp Gly Phe Val Ile Glu Lys Gly Pro Asp Ser Phe Leu Ala Arg Lys
50 55 60

Thr Ala Met Ile Asp Leu Ala Lys Glu Leu Glu Ile Asp His Glu Leu
65 70 75 80

Val Ser Gln Asn Pro Glu Ser Lys Lys Thr Tyr Ile Met Gln Arg Gly
85 90 95

Lys Leu His Pro Met Pro Ala Gly Leu Val Leu Gly Ile Pro Thr Glu
100 105 110

Leu Arg Pro Phe Leu Arg Ser Gly Leu Val Ser Pro Ala Gly Lys Leu
115 120 125

Arg Ala Leu Met Asp Phe Val Ile Pro Pro Arg Arg Thr Thr Glu Asp
130 135 140

Glu Ser Leu Gly Tyr Met Ile Glu Arg Arg Leu Gly Ala Glu Val Leu
145 150 155 160

Glu Asn Leu Thr Glu Pro Leu Leu Ala Gly Ile Tyr Ala Gly Asp Met
165 170 175

Arg Arg Leu Ser Leu Gln Ala Thr Phe Pro Gln Phe Gly Glu Val Glu
180 185 190

Arg Asp Tyr Gly Ser Leu Ile Arg Gly Met Met Thr Gly Arg Lys Pro
195 200 205

Ala Glu Thr His Thr Gly Thr Lys Arg Ser Ala Phe Leu Asn Phe Arg
210 215 220

Gln Gly Leu Gln Ser Leu Val His Ala Leu Val His Glu Leu Gln Asp
225 230 235 240

Val Asp Gln Arg Leu Asn Thr Ala Val Lys Ser Leu Gln Arg Leu Asp
245 250 255

Gly Ala Gln Thr Arg Tyr Arg Val Glu Leu Gly Asn Gly Glu Met Leu
260 265 270

Glu Ala Asp Asp Val Val Val Thr Val Pro Thr Tyr Val Ala Ser Glu
275 280 285

Leu Leu Lys Pro His Val Asp Thr Ala Ala Leu Asp Ala Ile Asn Tyr
290 295 300

Val Ser Val Ala Asn Val Val Leu Ala Phe Glu Lys Lys Glu Val Glu
305 310 315 320

His Val Phe Asp Gly Ser Gly Phe Leu Val Pro Arg Lys Glu Gly Arg
325 330 335

Asn Ile Thr Ala Cys Thr Trp Thr Ser Thr Lys Trp Leu His Thr Ser
340 345 350

Pro Asp Asp Lys Val Leu Leu Arg Cys Tyr Val Gly Arg Ser Gly Asp
355 360 365

Glu Gln Asn Val Glu Leu Pro Asp Glu Ala Leu Thr Asn Leu Val Leu

370 375 380

Lys Asp Leu Arg Glu Thr Met Gly Ile Glu Ala Val Pro Ile Phe Ser
385 390 395 400

Glu Ile Thr Arg Leu Arg Lys Ser Met Pro Gln Tyr Pro Val Gly His
405 410 415

Leu Gln His Ile Ala Ala Leu Arg Glu Glu Leu Gly Ser Lys Leu Pro
420 425 430

Gly Val Tyr Ile Ala Gly Ala Gly Tyr Glu Gly Val Gly Leu Pro Asp
435 440 445

Cys Ile Arg Gln Ala Lys Glu Met Ser Val Gln Ala Thr Gln Glu Leu
450 455 460

Ala Ala Asp
465

<210> 182
<211> 463
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tái tổ hợp

<400> 182

Leu Val Ile Ile Gly Gly Gly Ile Thr Gly Leu Ala Ser Ala Phe Tyr
1 5 10 15

Met Glu Lys Glu Ile Arg Glu Lys Asn Leu Pro Leu Ser Val Thr Leu
20 25 30

Val Glu Ala Ser Pro Arg Val Gly Gly Lys Ile Gln Thr Ala Arg Lys
35 40 45

Asp Gly Tyr Ile Ile Glu Arg Gly Pro Asp Ser Phe Leu Glu Arg Lys
50 55 60

Lys Ser Ala Pro Glu Leu Val Glu Asp Leu Gly Leu Glu His Leu Leu
65 70 75 80

Val Asn Asn Ala Thr Gly Gln Ser Tyr Val Leu Val Asn Glu Thr Leu
85 90 95

His Pro Met Pro Lys Gly Ala Val Met Gly Ile Pro Thr Lys Ile Ala
 100 105 110

Pro Phe Met Ser Thr Arg Leu Phe Ser Phe Ser Gly Lys Ala Arg Ala
 115 120 125

Ala Met Asp Phe Val Leu Pro Ala Ser Lys Pro Lys Glu Asp Gln Ser
 130 135 140

Leu Gly Glu Phe Phe Arg Arg Arg Val Gly Asp Glu Val Val Glu Asn
 145 150 155 160

Leu Ile Glu Pro Leu Leu Ser Gly Ile Tyr Ala Gly Asp Ile Asp Arg
 165 170 175

Leu Ser Leu Met Ser Thr Phe Pro Gln Phe Tyr Gln Thr Glu Gln Lys
 180 185 190

His Arg Ser Leu Ile Leu Gly Met Lys Lys Thr Arg Pro Gln Gly Ser
 195 200 205

Gly Gln Gln Leu Thr Ala Lys Lys Gln Gly Gln Phe Gln Thr Leu Lys
 210 215 220

Thr Gly Leu Gln Thr Leu Val Glu Glu Leu Glu Asn Gln Leu Lys Leu
 225 230 235 240

Thr Lys Val Tyr Lys Gly Thr Lys Val Thr Asn Ile Ser Arg Gly Glu
 245 250 255

Lys Gly Cys Ser Ile Ala Leu Asp Asn Gly Met Thr Leu Asp Ala Asp
 260 265 270

Ala Ala Ile Val Thr Ser Pro His Lys Ser Ala Ala Gly Met Phe Pro
 275 280 285

Asp Leu Pro Ala Val Ser Gln Leu Lys Asp Met His Ser Thr Ser Val
 290 295 300

Ala Asn Val Ala Leu Gly Phe Pro Gln Glu Ala Val Gln Met Glu His
 305 310 315 320

Glu Gly Thr Gly Phe Val Ile Ser Arg Asn Ser Asp Phe Ser Ile Thr

325 330 335
 Ala Cys Thr Trp Thr Asn Lys Lys Trp Pro His Ser Ala Pro Glu Gly
 340 345 350
 Lys Thr Leu Leu Arg Ala Tyr Val Gly Lys Ala Gly Asp Glu Ser Ile
 355 360 365
 Val Glu Leu Ser Asp Asn Glu Ile Ile Lys Ile Val Leu Glu Asp Leu
 370 375 380
 Lys Lys Val Met Lys Ile Lys Gly Glu Pro Glu Met Thr Cys Val Thr
 385 390 395 400
 Arg Trp Asn Glu Ser Met Pro Gln Tyr His Val Gly His Lys Gln Arg
 405 410 415
 Ile Lys Lys Val Arg Glu Ala Leu Ala Ala Ser Tyr Pro Gly Val Tyr
 420 425 430
 Met Thr Gly Ala Ser Phe Glu Gly Val Gly Ile Pro Asp Cys Ile Asp
 435 440 445
 Gln Gly Lys Ser Ala Val Ser Asp Val Leu Ala Tyr Leu Phe Glu
 450 455 460

<210> 183
 <211> 1329
 <212> ADN
 <213> Xanthomonas campestris

<400> 183
 atgcaaacac agcccggttat cattgccggc gccggtattg ccggactaag tatagcttac 60
 gaattacagc agaaaggcat tccctatgaa atcatggagg cctcttccta tgcaggaggc 120
 gttgtgaaat cattacatat tgatgggttat gaactggatg ctggccctaa ttcgctggcc 180
 gcatctgcag cattcatggc ttatatcgat caactgggtt tgcaggacca ggtattggaa 240
 gctgcggttg ccagtaagaa ccgctttctg gtcagaaatg ataaattgca tgcagtatcg 300
 ccacatccct ttaagatact gcagtcagca tatatcagtg gtggcgccaa gtggcgctcg 360
 ttcacagaaa gatttcgaaa agcggccgct ccggagggag aggaaacagt atcttccttt 420
 gtgacccgcc gttttggaaa ggagatcaat gactaccttt ttgaaccgt gctttctggt 480
 atatatgcag gtaatcctga tctgatgtca gttggtgaag tactgcctat gctgccacaa 540

tgaggagcaaa aatacggtag tgttacgcag ggactcctga agaataaagg agctatgggt	600
ggacgtaaga tcattgcctt taaaggaggt aatgcgacac tgacaaacag attgcaatcc	660
ctgcttagcg gtaagataag atttaactgt gccgtaacgg gtgtaacccg tggggcggac	720
gactatattg tacaatatac cgagaatggg aatacagcta tgctgaatgc atcccgtgtg	780
atattcacca cccctgcata cagtacagcc gtagctatac aggcaattga cgcttccctt	840
gctacacatc tcagcgatgt tccctatccc cgtatgggag tactgcacct ggggtttgga	900
gcggaagccc ggcagaaagc accggcaggt tttgggtttcc tgggtgccga tgctgcagga	960
aagcatttcc tgggcgctat ctgtaacagc gctatattcc cttcccgcgt accgacaggt	1020
aaagtgctgt ttacggtgtt cctgggtggc gcgagacaag aacagctgtt tgatcagctg	1080
gggcctgaaa agctacagca gacagtagtg aaagaactga tggaaactgt gggcctgact	1140
acaccaccag aatgcagcg ttttagtgaa tggaacagag cgattccgca actaaatgta	1200
ggttatgcac agacgaggca gcagataggc gtttttgaac agcgttaccg gggcatcaga	1260
ttagcgggta actatgtgac cggagtggct gtaccgcgta tcatacaggc cgcaaaaggg	1320
tactgttga	1329

<210> 184

<211> 1347

<212> ADN

<213> Chitinophaga pinensis

<400> 184

atgtctgatc aaccggtatt gattgtcggc gccggcttat ccggattgag cattgcgtat	60
gaattgcaga aactgcaggt gccttaccag gtactggaag tttcgggtca tagcggcggc	120
gtgatgaaat cattacggaa agatggattt gaactggatg caggcgctaa tacaatcgca	180
gcttctcctg aaatactggc atacttcaca tcaactgggac tggaaaatga gatattgcag	240
gccaccgctg ccagcaagca ccggttctct gtaagacggc ggcagttgca cgctgtttct	300
ccccatcctt tcaagatcat gtcgtctcct tacctgagca ggggcagtaa atggcggttg	360
tttaccgaac gttttcgcaa acctgttgtg gcaagcggag aagaaaccgt caccgatttt	420
ataacaagaa ggtttaaccg ggagatagca gaatatgtgt ttgaccgggt attatccggc	480
atatatgccg gcaatcccgca ccagatgagc atagcgggaag tattacctgc gttgccgcgc	540
tgggagcggg aatatgggag tgttaccaa gggctgatga aagataaagg cgcaatgggc	600
ggccggaaga ttatcagttt taaagggtgt aaccagttgc tcacaaaccg tttgcagcaa	660
ttgctcacta ccccggtgag ctttaattgt aaggtaaccg gtatcaccgc atccaatggc	720

ggctatatattg taagcgctgt agaagatggc gtatcagaaa gttatactgc ttcaaggggtg 780
 atattaacca cacctgctta cagcgcgaggca gcaactatta cgaatcttga tgctgctacc 840
 gctgccttgt taaatgaaat tcattatccc cgtatgggag tgctgcacct gggttttgac 900
 gctactgcgt tgccgcagcc cctggatgga tttggtttcc tggtagcgaa tgctgaaaat 960
 atgcatttcc tgggagcaat ctgcaacgct gcaattttcc cggataaggc gcctccggga 1020
 aaaatcctct ttacggtatt cctgggagga gcaagacagg aaagtttggt tgaccagatg 1080
 acgcccgaag ctctgcaaca gcaggtagtt tcagagggtca tgtctttact gcatttatct 1140
 gcgcccgcgg taatgcagca tttcagtagc tggaataaag cgattccgca gttaaatgtg 1200
 ggtcatgtta agttacggcg tgccgtggaa gcttttgaaa aaaaatatcc cgggtattcac 1260
 ctacgcggga attacctgca aggcgtagct atcccggctt tactgcaaca tgccgcgcgt 1320
 ttggcggctt ccctgaagaa aaattaa 1347

<210> 185
 <211> 1335
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 185
 caaccggtcc tcacgttgg agctggtctc tccgggctct caatcgctta cgaactacag 60
 aagctgcaag tcccttacca agtgctggag gtttctggac attctggtgg agtcatgaag 120
 tcactccgga aggacggatt tgaactcgac gctggtgcca acaccatagc cacgtctccc 180
 gagattcttg cgtactttac ctactaggt cttgagaatg agatcctcca ggcgactgct 240
 acttctaaac accgcttctt ggtgcggcga aggcaactgc acgccgtgag cccgcacccg 300
 ttcaagatca tgtcatcgcc gtacctctgc cgtggctcca aatggagggt ctttactgag 360
 cggtttcgga aaccgctcgt cgcttcgggc gaggagaccg tcaccgattt catcacgagg 420
 agattcaacc gcgaaatagc ggagtatgtg ttcgaccctg ttctaagtgg gatctacgcc 480
 gggaaccgg accaaatgag tattgctgag gtgttgcttg ccttgccctag gtgggaaagg 540
 gagtacggat cagtgaacaa gggccttatg aaggataagg gtgcgatggg aggtcgaaag 600
 atcatcagct ttaagggtgg caaccagcta cttacaaacc gcttacagca gctactcact 660
 actccggtga gattcaattg caaggtgaca gggattacag ccagcaatgg cgggtacatc 720
 gtgagcgctg ttgaggacgg cgtatctgag agctacaccg catctcgtgt gatcttgacc 780
 acaccgctt actcagcagc ggctaccata actaaccttg atgcagccac tgcggcactg 840

ttgaacgaaa tccattatcc acgtatgggc gtgttacact tgggctttga tgcaactgcc	900
ttgccacagc cgctggacgg gttcggatth ctagtgccga acgcggagaa catgcacttc	960
ctgggagcca tctgcaatgc agccatcttc ccggacaagg ctccgcccgg caagatcctg	1020
tttacagtgt tcctcggagg cgcacgccag gagtcgctct tcgatcagat gactcctgag	1080
gctcttcagc agcaagtcgt tagtgagggtg atgagcttgt tgcacttgct agctccaccg	1140
gtgatgcagc acttctcttc ctggaacaag gccatccctc aattgaacgt cgggcacgtg	1200
aagttgcggc gcgcggtaga ggcgttcgag aagaaatacc ctggaatcca tctctcgggc	1260
aactacctcc agggagttgc aataccagct ttactccagc acgccgcagc tttagctgct	1320
tctcttaaga agaac	1335

<210> 186
 <211> 1413
 <212> ADN
 <213> Bacillus subtilis

<400> 186	
atgagtgcg gcaaaaaaca tgtagtcac atcggcgggc gcattaccgg tttagccgcc	60
gccttctata tggaaaaaga aatcaaagaa aagaatctgc cgcttgaact gacgcttggt	120
gaggcaagtc cgagagtcgg cggaaaaatc cagactgtca agaaggacgg ctatatcatc	180
gaaagagggc cagactcatt tctggaacga aagaaaagcg ccccgagct tgttaaagac	240
ttaggtcttg agcatttgct tgtcaacaat gcgaccggac aatcctatgt gcttgtaaac	300
cgcactctgc atccaatgcc gaagggcgct gtaatgggga taccgacaaa aattgcgccg	360
tttgtttcta cgggtctggt ttctttgtct gggaaggcga gagctgctat ggatttcac	420
ctgcctgcta gcaaaacaaa ggatgatcag tcattgggag aattcttccg cagacgtgtc	480
ggagatgaag tggtcgagaa cttaatcgag ccgcttctat cagggatcta cgcaggcgac	540
attgacaagc tcagcctgat gtcgacattc ccgcaatttt atcagacgga acaaaagcat	600
agaagcctga ttctcggcat gaaaaaaca aggcctcaag gtcaggcca gcagctgacg	660
gcaaaaaaac aagggcagtt ccagactctg tcaaccgggt tgcagaccct tgtagaagag	720
atcgaagaag agttaagct gacgaagggt tataaaggca caaaagtgac caaactcagc	780
catagcggct ctggctattc gctcgaactg gataacggcg tcacacttga tgctgattca	840
gtaattgtga ctgctccga taaagcggct gcgggaatgc tttctgagct tcctgccatt	900
tctcatttga aaaatatgca ctccacatcc gtggcaaacg tcgctttagg tttccctgaa	960
ggctccgtcc aaatggagca tgagggcacg ggttttgtca tttcaagaaa cagtgacttt	1020

gcgatcacag cctgtacgtg gacgaataaa aaatggccgc acgcagcgcc ggaaggcaaa 1080
 acgctgcttc gggcatatgt cggaaaagct ggagacgaat ccattgtcga tctatcagat 1140
 aatgacatta tcaacattgt gttagaagac ttaaagaaag tcatgaacat aaacggcgag 1200
 ccggaaatga catgtgtcac ccgatggcat gaaagcatgc cgcagtacca tgtcggccat 1260
 aagcagcgta tcaaggagct gcgtgaagca cttgcatctg cgtatccggg tgtttatatg 1320
 acaggcgctt ctttcgaagg tgtcggcatt cccgactgca ttgatcaagg aaaagctgcc 1380
 gtgtctgacg cgcttaccta tttattcagc taa 1413

<210> 187

<211> 1413

<212> ADN

<213> *Bacillus pumilus*

<400> 187

atgcatgaca atcaaaaaca ccttgtcatc attggcggtg gcatcactgg tttagccgcc 60
 gccttctatt tggaaaagga agtcgaggaa aaaggctctc cgattcaa atcacttatt 120
 gaagcgagcc ctaggctagg tggaaaaata caaacattat ataaagacgg ctacatcatt 180
 gaacgtggac ctgattcatt tttagaaaga aaggctcagt ggccgcagct tgcaaaagat 240
 gtcggtctgt ccgatcagct cgtcaataat gaaactgggc aagcgtatgt actggccaat 300
 gaaaagcttc acccgatgcc aaaagggtgct gttatgggga ttccaactca aatcagccca 360
 tttattacaa ctgggtctttt ttcagttgcg ggaaaagcaa gagcggcgat ggatttcgtg 420
 ttgccaaaaa gcaagcagac ggaagaccag tcgcttggtg aatttttttag aagacgtgtg 480
 ggtgatgagg tcgttgagaa ttaattgag ccgcttctat caggcattta tgcaggggat 540
 attgaccgtc tgagcttaat gtcgaccttc ccgcaatttt atcaaacaga acagcagcat 600
 cgaagtttga ttcttgggat gaaaaaatca cagcagcatg cgaaagcgca gcaagtgact 660
 gcgaaaaaac aaggacagtt ccaaacgac atcaaggat tgcagtcgct tgtggaagca 720
 gtagaaggta agctcaagct gacaacggtc tataaaggga caaaagtcaa acaaattgaa 780
 aaaacggatg gaggctatgg cttacaatta gacagcgttc aaacgctttt tgccgattca 840
 gccattgtca cgactccgca tcaatcgatt tattccatgt ttctaaaga agcagggcta 900
 gagtatttgc atgacatgac ctctacttct gttgcaacag tagcactcgg ttttaaagat 960
 gaggatgttc ataatagaata tgacggcact ggatttgtca tctcaagaaa cagtgatttc 1020
 tctattacgg cctgtacatg gacaaaca aaatggccgc atactgctcc gaaaggaaaa 1080
 acgctattgc gtgcgtatgt agggaaggct ggcgacgaat caattgtcga gcagtcagac 1140

agtcaaatcg tcagcattgt gctagaagat ttaaagaaaa tcatggatat taaagcagat	1200
ccagaattga cgacagtgc tcgctggaag acaagtatgc cgcaatatca cgtcggtcat	1260
cagaaagcca tttcgaacat gcgagaaacg ttttaagcaat catatcctgg tgtttatatt	1320
acaggtgctg cttttgaagg tgtcggaatc cctgattgta ttgatcaagg aaaagccgcc	1380
atctcagagg ctgtatcgta tctattttca taa	1413

<210> 188
 <211> 1413
 <212> ADN
 <213> *Bacillus pumilus*

<400> 188	
atgcatgaca atcaaaaaca cttgtcatc attggcgggtg gcatcactgg tttagccgcc	60
gccttctatt tggaaaagga agtcgaagaa aaaggctctc ccattcaaatt atctcttatt	120
gaagcgagcc ctaggctagg tggaaaaatc caaacattat ataaagacgg ctacatcatt	180
gaacgtgggc ctgattcatt tttagaaaga aaggctcagt gaccgcagct ggcgaaagat	240
gtaggtctat ccgatcagct cgtcaataat gaaacagggc aggcgtatgt actagtcaat	300
gaaacccttc acccgatgcc aaaaggcgct gtcattgggtta ttccaactca aatcagccca	360
ttcatcaciaa ccggtctttt ttcagttgca ggaaaagcga gagccgcaat ggatttcgtc	420
ttgccaaaaa gcaagcaaac agaagatcag tcgctcgggtg aatttttttag aagacgtgtc	480
ggtgatgaag tagttgagaa tttaatcgaa cctcttctat caggcattta tgcaggtgac	540
attgaccgtc tcagcttaat gtccaccttc ccgcagtttt atcaaacaga acaaaagcat	600
cgagttttga ttcttgggat gaaaaaatca cagcagcatg cgaaagcgca gcaagtgaca	660
gcgaaaaaac aagggcagtt ccaaacgac aatcaaggac ttcaagcgct tgttgaagca	720
gtagaaagca agctcaagct gacaacgatt tataaaggga caaaagtga gcaagattgaa	780
aaaacagatg ggggctacgg tgtgcagtta gacagcggtc aaacgctttt ggctgattca	840
gccattgtga caactccgca tcaatcgatc tattccatgt ttccaaaaga agcggggctt	900
gagtacttgc atgatatgac atctacttct gttgcaacgg ttgcactcgg ttttaaagaa	960
gaggatgttc ataatagaata tgacggtact ggttttgtca tctcaagaaa cagtgatttc	1020
tctattacag cttgtacgtg gacgaacaaa aaatggccgc atacagctcc taaaggaaaa	1080
acattattgc gtgcttatgt agggaaggct ggcgacgaat caattgtcga acagtcagac	1140
catcaaactcg tcagcattgt actggaggat ttgaagaaaa ttatggatat taaagcagat	1200
ccagaactga caacagtgc tcgctggaag acgagcatgc cgcaatatca cgtcggtcat	1260

caaaaagcca tttcgaacat gcgagaaacg ttttaagcaat catatcctgg tgtttatatc	1320
acaggtgctg cttttgaagg tgtcggaatc cctgattgta ttgatcaagg aaaagctgcc	1380
atttcagagg ctgtatctta tctattttca taa	1413

<210> 189
 <211> 1329
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 189	
atgcaaactc agcccgatcat aatagcgggc gctggcattg cgggcctttc tatcgcatatc	60
gagctgcaac agaagggcat tccttacgaa attatggaag cctcgtccta cgccggaggc	120
gtggtcaagt cccttcacat tgatggctac gaactagacg ccggacctaa ttcacttgcc	180
gcgtccgctg ccttcatggc ctacatcgac caactcggac tccaagatca agtgcttgaa	240
gccgccgcag catccaagaa ccgcttcctc gtaagaaacg acaagctcca tgcagtctcg	300
ccgcacccgt ttaagatcct ccagtcggcc tacatcagtg gcggcgctaa gtggagattg	360
tttaccgaaa ggttccgcaa agctgcggct ccagaggggtg aggagacagt gagcagcttc	420
gtgacgagga ggtttggcaa ggagatcaac gactacctgt ttgaaccgt cttgtccggg	480
atctacgcgg gcaaccgga tttgatgagt gttggcgagg ttctgccgat gcttcctcaa	540
tgggagcaga agtacggcag cgttacacaa ggcttggtga agaataaggg cgcaatgggc	600
ggccgaaaga taatcgcttt caagggcggg aatgccacac tgaccaaccg tcttcagtca	660
ctgctctcag gaaagatccg cttcaattgc gccgtgacgg gtgtcacacg aggcgcagac	720
gactacattg ttcagtacac tgagaatggc aataccgcaa tgttgaatgc aagccgcgtg	780
atcttcacaa caccgctta ctcaactgct gttgccatcc aggcgttgga cgccagcttg	840
gccactcacc tctctgatgt accctatcct cgcattgggtg tggtgcactt gggcttcggt	900
gctgaggcaa ggcagaaggc tcctgcgggc tttgggttct tgggtcccaca cgagccgga	960
aagcatttcc tgggagcaat ctgtaactcc gctatcttcc cttcgcgggt gccactggc	1020
aaggtgttat tcaccgtgtt cttgggcggg gccagacagg agcaactgtt tgaccagcta	1080
ggccctgaga agttacaaca gacagtgggtg aaggagctta tggaattgct gggcctaact	1140
acgccgccgg agatgcaacg attctctgag tggaatcgcg caataccgca acttaattgtt	1200
ggctacgcc agactcgtca gcagattggc gtattcgagc agcgtaccc tggcatccgc	1260

ttggccggga actatgtaac tggagtggcg gtgcccgccca ttatccaagc tgcaaagggc 1320
 tattgctaa 1329

 <210> 190
 <211> 1320
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Tái tổ hợp

 <400> 190
 cagcccgtca taatagcggg cgctggcatt gcgggccttt ctatcgcata cgagctgcaa 60
 cagaagggca ttccttacga aattatggaa gcctcgtcct acgccggagg cgtgggtcaag 120
 tcccttcaca ttgatggcta cgaactagac gccggaccta attcacttgc cgcgtccgct 180
 gccttcatgg cctacatcga ccaactcgga ctccaagatc aagtgcttga agccgccgca 240
 gcatccaaga accgcttcct cgtaagaaac gacaagctcc atgcagtctc gccgcacccg 300
 ttttaagatcc tccagtcggc ctacatcagt ggcggcgcta agtggagatt gtttaccgaa 360
 aggttccgca aagctgcggc tccagagggg gaggagacag tgagcagctt cgtgacgagg 420
 aggtttggca aggagatcaa cgactacctg tttgaacccg tcttgtccgg gatctacgcg 480
 ggcaacccgg atttgatgag tgttggcgag gttctgccga tgcttcctca atgggagcag 540
 aagtacggca gcgttacaca aggcttggtg aagaataagg gcgcaatggg cggccgaaag 600
 ataatcgctt tcaagggcgg gaatgccaca ctgaccaacc gtcttcagtc actgctctca 660
 ggaaagatcc gcttcaattg cgccgtgacg ggtgtcacac gaggcgcaga cgactacatt 720
 gttcagtaca ctgagaatgg caataccgca atgttgaatg caagccgcgt gatcttcaca 780
 acacccgctt actcaactgc tgttgccatc caggcggttg acgccagctt ggccactcac 840
 ctctctgatg taccctatcc tcgcatgggt gtgttgcaact tgggcttcgg tgctgaggca 900
 aggcagaagg ctctgcggg ctttgggttc ttggtccac acgcagccgg aaagcacttc 960
 ctgggagcaa tctgtaactc cgctatcttc ccttcgcggg tgcccactgg caaggtgtta 1020
 ttcaccgtgt tcttgggcgg tgccagacag gagcaactgt ttgaccagct aggccctgag 1080
 aagttacaac agacagtggg gaaggagctt atggaattgc tgggcctaac tacgccgccg 1140
 gagatgcaac gattctctga gtggaatcgc gcaataccgc aacttaattg tggctacgcc 1200
 cagactcgtc agcagattgg cgtattcgag cagcgctacc ctggcatccg cttggccggg 1260
 aactatgtaa ctggagtggc ggtgcccgcc attatccaag ctgcaaaggg ctattgctaa 1320

<210> 191
 <211> 1347
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 191
 atgagcgacc agcccgctcct catcggttga gctgggtctct ccgggctctc aatcgcttac 60
 gaactacaga agctgcaagt cccttaccaa gtgctggagg tttctggaca ttctgggtgga 120
 gtcatagaagt cactccggaa ggacggattt gaactcgacg ctgggtgcaa caccatagcc 180
 gcgtctcccg agattcttgc gtactttacc tcactagggtc ttgagaatga gatcctccag 240
 gcgactgctg cttctaaaca ccgcttcttg gtgcggcgaa ggcaactgca cgccgtgagc 300
 ccgcacccgt tcaagatcat gtcacgcgcg tacctcagcc gtggctccaa atggcggtctc 360
 ttactgagc gggttcggaa gcccgctcgc gcttcgggagc aggagaccgt caccgatttc 420
 atcacgagga gattcaaccg cgaaatagcg gagtatgtgt tcgaccctgt tctaagcggg 480
 atctacgcgc ggaacccgga ccaaattgagt attgctgagg tgttgccctgc cttgcctagg 540
 tgggaaaggg agtacggatc agtgaccaag ggccttatga aggataaggg tgcgatggga 600
 ggtcgaaaga tcatcagctt taagggtggc aaccagctac ttacaaaccg cttacagcag 660
 ctactcacta ctccggtgag attcaattgc aagggtgacag ggattacagc cagcaatggc 720
 ggggtacatcg tgagcgctgt tgaggacggc gtatctgaga gctacaccgc atctcgtgtg 780
 atcttgacca caccgctta ctcagcagcg gctaccataa ctaaccttga tgcagccact 840
 gcggcactgt tgaacgaaat ccattatcca cgtatgggagc tggttacactt gggctttgat 900
 gcaactgcct tgccacagcc gctggacggg ttcggatttc tagtgccgaa cgcggagaa 960
 atgcacttcc tgggagccat ctgcaatgca gccatcttcc cggacaaggc tccgcccggc 1020
 aagatcctgt ttacagtgtt cctcggaggc gcacgccagg agtcgctctt cgatcagatg 1080
 actcctgagg ctcttcagca gcaagtcgtt agtgaggtga tgagcttggt gcaactgtca 1140
 gctccaccgg tgatgcagca cttctcctcc tggaacaagg ccatccctca attgaacgtc 1200
 gggcacgtga agttgcggcg cgcggtagag gcgttcgaga agaaataccc tggaatccat 1260
 ctctcgggca actacctcca gggagttgca ataccagctt tactccagca cgccgcagct 1320
 ttagctgctt ctcttaagaa gaactga 1347

<210> 192
 <211> 1413
 <212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 192

atgtcggatg gcaagaagca cgtcgtcatc ataggcgggtg ggatcactgg cttggccgct	60
gcattctaca tggagaagga gattaaggag aagaacctcc cacttgagct gacgctagtt	120
gaggccagtc ccagggtcgg cggcaagatc cagacgggtca agaaggacgg gtacataatt	180
gaacgcggcc ctgacagctt cttagagcgc aagaaatcgg ctccgcagct agttaaggac	240
ttgggacttg agcacctgct cgtcaacaac gcgaccggac agtcgtacgt gctcgtgaac	300
cggacgctcc acccgatgcc gaagggcgct gtgatgggca ttccgaccaa gatagcacca	360
ttcgtgagta ccggcctatt cagcctttcc ggcaaggcaa gggctgcat ggacttcac	420
ttgcctgcct ctaagactaa ggacgatcag tccttgggcg agttcttccg ccgccgggtg	480
ggatgatgagg tgggtggagaa cttaattgag ccgctcctat ctggaatcta cgctggtgac	540
atcgacaaac tgtctctgat gtccaccttt ccgcagttct accaaactga gcagaagcac	600
cgttcactta tcttgggaat gaagaagact agacctcaag gttcgggtca gcaactgacg	660
gccaagaaac agggtcagtt ccagacgcta agcaccgggc ttcagacact cgtggaggag	720
attgagaaac agctcaaact tactaagggtg tacaagggca cgaagggtgac aaagttatcc	780
cactccggca gcgggtactc cctggagttg gacaatggcg taacgttggga cgccgactca	840
gttatcgtga cagcgccgca taaggctgct gccgggatgt tgtcagaact cccggcgatt	900
tcccatctca agaacatgca cagtacctcg gttgccaacg tcgccctcgg attcccggaa	960
ggaagtgttc aaatggagca cgaaggcacg ggtttcgtaa tttccaggaa ctccgacttt	1020
gccatcaccg cttgtacttg gaccaacaag aagtggcctc atgctgcgcc ggagggcaag	1080
acattgctca gagcttacgt cgggaaggcg ggcgacgagt caatcgtcga tcttagcgac	1140
aacgacatca ttaacattgt gctggaggac ttgaagaagg ttatgaacat caatggcgag	1200
ccagagatga cctgcgtgac ccgatggcac gagtctatgc cgcagtacca cgtcggtcac	1260
aagcagcgca tcaaggagtt gcgcgaggca ctgcctcag cttaccctgg cgtgtacatg	1320
actggcgctt cgtttgaggg cgttggtatt cctgactgca tcgaccaggg aaaggcggcc	1380
gtcagtgacg cgctcaccta cctcttcagt tga	1413

<210> 193

<211> 1413

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 193

atgcacgaca accagaagca cctggtcata atcggaggcg gcataaccgg ccttgctgcg	60
gccttctacc tggagaagga ggtcgaggag aagggctctcc ctatccagat ttcattgatt	120
gaggcttcgc ctcggctggg agggaagatc cagacattgt acaaggacgg gtacatcatc	180
gagcgtggtc cagacagttt cctggagcgg aaggtcagcg gaccgcagct cgccaaggac	240
gtgggactta gcgaccaact ggtgaacaac gagacaggac aggcgtacgt cttggtgaat	300
gagaagttgc acccgatgcc taagggtgcc gtgatgggca tccaacgca aatctcacct	360
ttcatcacca cggactctt ctccgtggcc ggaaaggcac gagctgcaat ggacttcgtt	420
ctgcctaagt cgaaacagac cgaagaccag tctctaggcg agttcttccg ccgccgtgtg	480
ggtgacgagg ttgtggagaa cctcatcgag cctttgttgt ctgggatcta cgcgggcgac	540
atcgacagac ttagtctcat gagtaccttt ccgcaattct atcagacaga acagcagcat	600
cgaagtctca tactcgggat gaagaagtca caacaacatg caaaggccca gcaagttacc	660
gccaagaaac agggccagtt ccaaacgac aaccagggcc tccagagctt ggtggaggca	720
gtggagggaa agttgaagct caccaccgtt tacaaaggga caaagggttaa acagattgag	780
aagacggacg gcggttacgg gttacaattg gactccggac agactctctt cgctgattcc	840
gctatcgtaa ctactcctca ccagagcatc tactctatgt tcccgaagga ggcgggcctg	900
gagtacctgc acgacatgac ttcaacgtct gtcgccaccg tggctttggg cttcaaggac	960
gaggacgtcc acaatgagta tgacgggacg ggattcgtta tcagtaggaa ctccgacttc	1020
agcatcaccg cctgcacgtg gaccaacaag aagtggccac acaccgcgcc caaagggaag	1080
acccttctga gggcatacgt gggcaaggcg ggcgacgaga gcatcgtcga gcaatctgat	1140
tctcagattg tttcaatcgt cctcgaagac ctcaagaaga tcatggacat caaggcagac	1200
ccggaactta ccaccgttac tcgatggaag acctcgatgc ctcagtatca cgtcggggcac	1260
cagaaggcaa tcagcaacat gagggagaca ttcaagcagt cgtatcctgg cgtgtacatt	1320
accggagcag cattcgaagg cgtaggaatc cctgactgca ttgaccaggg caaggctgct	1380
atctcagagg ccgtgtccta tctcttctcg tga	1413

<210> 194

<211> 1413

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 194

atgcacgaca accagaagca cctggtgata attggaggcg ggattaccgg cctagcagcc	60
gcttttctatc tggagaagga ggtggaggag aagggcctcc cgatacagat ttcgctgatt	120
gaagcctctc cgcgctggg cggcaagatc cagacattgt acaaggacgg gtacatcatt	180
gagcgcgggc ctgactcgtt cctggagcgg aaggtctccg gtcctcaact ggccaaagac	240
gtgggtcttt ccgatcagct tgtgaacaat gagaccggtc aggcctacgt cttggtcaac	300
gaaactctgc atcccatgcc taaggagacc gttatgggca ttccaacgca aatctctccg	360
ttcataacga ctgggctggt cagcgttgcg ggcaaagcaa gggctgctat ggacttcgtg	420
ctgccaaaga gtaagcagac cgaggaccag tccctcggcg agttcttccg ccgccgagtg	480
ggcgatgagg tggttgagaa tctaatacga ccgctgttgt cgggcatcta tgcgggcgac	540
atcgacaggc taagtcttat gtccactttc cctcagttct accagacaga gcagaaacac	600
aggagtctca tccttggaat gaagaagtcc cagcagcacg cgaaggctca gcaagtgacc	660
gccaagaagc aaggacagtt ccagaccatc aaccagggcc tacaggccct tgtcgaagcc	720
gttgagtcga agttaaagtt gacgacgac tacaagggca ccaaggtgaa gcagattgag	780
aagactgacg gtggctatgg tgtgcaactc gattcgggcc aaacattgct cgctgactcc	840
gctatcgtca cgacgccaca ccagtcgac tactcgatgt tcccgaagga ggcgggccta	900
gagtaccttc acgacatgac ctccacttcg gtcgccaccg ttgcaactcg ctttaaggag	960
gaggacgttc acaacgagta cgatggcacc ggattcgtga tctccaggaa ctcggaactc	1020
tcgattaccg cgtgcacgtg gacaaataag aagtggccgc acacagcgcc aaagggcaag	1080
acccttctgc gggcgatatgt gggcaaggcc ggtgacgaga gcattgtcga acaatctgac	1140
catcagatcg tttctattgt tcttgaggat ctcaagaaga taatggacat taaggccgac	1200
cctgagctta ccacagtgac gaggtggaag acctcgatgc cgcagtatca cgtagggcac	1260
cagaaggcca tctccaacat gcgggagaca ttcaagcagt cgtaccctgg cgtgtacatt	1320
actggcgctg ctttcgaggg cgttggcatc ccggactgca tcgaccaggg caaggccgca	1380
atctcagagg cagtgtcgta cctgttcagc tag	1413

<210> 195

<211> 1329

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 195

atgcaaacgc aaccagtgat aatcgcgggc gctggcatcg ccggactttc cattgcgtac	60
gagctccagc agaaggggtat cccgtacgag atcatggagg caagttccta tgccggcggc	120
gtcgtcaagt cactgcacat cgacgggtac gagctggacg cgggacccaa cagcctagcc	180
gcttccgctg ccttcatggc gtacatcgac cagctggggc tccaagacca agtcctcgag	240
gctgccgcgg cgagtaaaaa tcgctttctc gtgaggaacg acaagcttca cgctgtgtca	300
ccgcaccctt tcaaaatact tcaaagcgcc tacatctcgg gcggtgctaa gtggcgggta	360
ttcacggagc gttttaggaa ggccgccgct cccgaagggtg aggagactgt ttctcttttc	420
gtcacacgca ggttcggaaa ggagatcaac gattatctct ttgagcctgt tctcagcggg	480
atatacgccg gcaaccaga ccttatgagc gtcggagagg ttctccctat gctgccgcaa	540
tgggagcaaa agtatggttc tgtgacccaa ggcctactga agaataaggg ggcgatgggc	600
ggaagaaaga taattgcatt caaggggggt aatgccaccc ttacaaatcg cctgcaaagc	660
cttttgtcgg gaaaaatccg tttcaattgt gccgtcaccg gtgttacaag aggcgcagat	720
gattacatcg ttcagtacac cgagaacggg aataccgcca tgctaaacgc atctaggggtg	780
attttcacaa ccccggccta ctcaactgcc gtcgccatcc aagccctcga cgccagcctg	840
gccactcatc tcagtgatgt gccttaccct cgtatggggg tattacatct tggcttcggg	900
gccgaagcgc gacagaaagc ccccgctgga tttggcttcc tagtccctca cgccgccggt	960
aaacattttc ttggcgccat ctgtaactcc gcaatcttcc catccagagt gcctactggc	1020
aaggttctgt ttactgtggt cctgggcggg gcccgccagg agcagctatt cgaccaatta	1080
ggcccagaaa agctccaaca aaccgttgtg aaggaaactaa tggagttgct cggactgacg	1140
acaccacccg agatgcagag gttttctgag tggaaccgcg cgattccaca actcaacgtc	1200
gggtacgcc agaccgggca acagataggg gttttcgagc agcgctatcc aggcattcga	1260
cttgctggta attacgtcac aggagtcgct gtgccagcca taatacaagc tgcaaagggg	1320
tattgctga	1329

<210> 196

<211> 1347

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 196

atgtcagacc aaccagtctt gattgttggg gccggcctct ctggcctgtc gattgcctac 60
 gaactgcaga agctccaggt gccgtaccaa gtcctggagg tgtcgggcca tagcggcggt 120
 gtcattgaaat cgctgcgtaa ggacggcttc gagttggacg cgggcgcgaa cacaatcgcg 180
 gctagcccag aaatacttgc ttactttaca agtctgggtc tggagaatga gatcctccag 240
 gctacagccg ctagcaaaca tcgattcctg gtgcgcaggc gacaactgca cgccgtcagt 300
 ccacatccat tcaagataat gtcgagcccc tatttaagcc gcgggtccaa gtggaggctc 360
 ttactgaaa gatttcgaaa accggtcgtc gctagcggag aagaaactgt tacagatttt 420
 attactcgca ggttcaacag ggagattgca gaatatgtct tcgatccagt tctctcagga 480
 atttacgcgg gcaaccaga ccagatgagc atcgctgaag tcctgcccgc gctccctcgg 540
 tgggaacgag aatatggaag cgtcaccaaa ggtctcatga aggacaaggg ggccatgggc 600
 ggtcgggaaga tcatatcggt taaaggcggg aaccagcttc tgactaaccg gctgcaacag 660
 ctgctcacta caccagtgcg gtttaattgc aaagtcacag gtataacggc tagtaatggc 720
 ggctacattg tttcagcggg cgaagatggg gtgagcgagt catacaccgc ctcccgcgtg 780
 atccttacca caccggccta ctcggcggca gctacaatca ccaatcttga cgcggctaca 840
 gccgcattac tcaacgagat tcattatccc aggatggggg tcctccatct gggcttcgac 900
 gcgacagctc ttccccagcc cttggatggc ttcgggtttc tgggtccgaa cgccgaaaac 960
 atgcattttc tcggcgccat ttgcaacgcc gcgatcttcc cggataaggc cccgcctgga 1020
 aaaatattgt tcaactgtctt tcttggcggc gcacgccagg agtccttgtt cgaccaaagt 1080
 accccagagg ctctgcagca gcaggtggc tctgaggtga tgtcacttct gcacctttct 1140
 gcacctccag tgatgcagca cttctcaagc tggataaaag ctatccccc gttgaacgtc 1200
 ggccacgtga agcttcgtag ggcggtcgaa gcgttcgaaa agaagtatcc aggcattcac 1260
 ctgtccggca actatctgca gggcgtcgca atcccggcg tactccagca cgccgctgcg 1320
 ctagccgcgt ctcttaagaa gaattag 1347

<210> 197
 <211> 1413
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 197
 atgagtgcg ggaagaagca cgttgtgata atcggaggcg ggataaccgg cctcgccgcc 60
 gccttctata tggagaagga aattaaggag aaaaacctcc cgctagagct gacgttggtg 120

gaagcgtcac caagggctcgg cggtaagatc cagaccgtca aaaaggatgg ctacatcatc 180
 gagcgcgggcc cggacagctt cctcgagcgg aagaagtccg caccacagtt agtcaaagac 240
 ctcggttgg aacacctttt ggtcaacaac gcgacaggtc agtcctatgt gcttgtgaat 300
 cggacgctgc acccgatgcc taagggcgct gtcattgggtg tccccacgaa gatcgcgccg 360
 ttctgtatcga cgggcctgtt ctccctatca ggtaaggccc gcgctgcat ggactttatc 420
 ctccctgcct cgaaaactaa agacgatcag tctactaggcg agttctttcg gcggcgagtg 480
 ggtgacgagg tgggtggagaa cctcatagaa cccctgctgt ccgggatcta cgctggagac 540
 atcgacaagc tgagcctcat gtctactttt ccgcaatttt atcagaccga gcagaaacac 600
 agatctctta tccttggcat gaagaagacc aggcctcagg ggtcgggtca acagctcaca 660
 gcaaagaagc aagggcagtt ccaaaccctg agcacaggct tgcagaccct ggtcgaagaa 720
 attgagaagc agctgaaatt aacgaagggt tacaaggga ccaagggtcac caaacttagt 780
 cacagcggct cgggctacag cctagagctt gacaacggag tgactctgga cgcagacagc 840
 gtgatcgtga cggcgcccca caaggctgcg gcgggaatgc tgagttagct ccccgccata 900
 agtcatctca agaacatgca ctcgacgtcg gtagccaatg tcgcgttggg gtttcccag 960
 ggtagcgtcc aaatggaaca cgaaggaact ggtttcgtca tatcccggaa ctctgacttc 1020
 gcgatcacag cgtgcacttg gacgaataaa aagtggccgc acgcagcggc tgaggggaag 1080
 acccttcttc gagcgtatgt gggcaaagcg ggcgatgaaa gcattgtgga tttatcggac 1140
 aacgacatta tcaacatcgt actggaagac ctaaagaaag tcatgaacat aaacggcgaa 1200
 ccggagatga catgcgtcac taggtggcac gagagcatgc cgcagtacca cgtggggcac 1260
 aagcagcgca tcaaggaatt gagggaggcc ctcgctagcg cgtaccctgg agtttacatg 1320
 accggcgcca gttttgaggg tgtcggtatc cctgactgta tcgaccaggg taaggccgcg 1380
 gtaagcgacg cattgacgta cctgtttctca tga 1413

<210> 198
 <211> 1413
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 198
 atgcacgaca accagaagca tctgggtcatc attggcgggg gcatcacggg cttggcagcc 60
 gccttctacc tggagaagga ggtcgaggag aagggccttc cgattcaaatt atctctgatt 120

gaggcgtctc cccgactcgg cgggaagatc cagaccctct ataaagacgg ctatataatt 180
 gagcggggac cagattcttt cctggagcga aaggtctcgg gccacagtt ggcgaaagat 240
 gtcggcctct cccgatcaact cgtgaacaac gagaccgggc aggcctatgt tctggtgaac 300
 gagaaattgc atcctatgcc taagggggcc gtcattggaa taccaacca aatatctccc 360
 ttcataacaa ccgactgtt ctcggttgcc ggtaaggcca gggccgcgat ggacttcgtc 420
 ctgccaaagt ctaagcagac ggaggaccag tccctcgggg aatttttccg ccgccgggtc 480
 ggcgacgagg ttgtggaaaa cctgattgag ccgttgctgt ctggcatcta cgcaggcgat 540
 atcgacaggc tgagccttat gtctacgttc ccgcaatfff atcagaccga gcagcagcac 600
 cggctctctga tacttggcat gaagaaatca caacagcacg ccaaagcaca acaggttact 660
 gctaagaagc aaggacaatt ccagacaatc aaccaagggt tgcagtccct cgtggaggcg 720
 gtagaaggca aattgaaact caccaccgtc tacaagggca cgaaagttaa gcagatcgag 780
 aaaacggatg gcggggtacgg tctccagctc gatagcggcc agacactggt cgccgactca 840
 gcgatcgtca ccacccccca ccagtccatc tacagcatgt tccctaagga ggcgggggtta 900
 gaatacttac atgacatgac ctccacctcc gtcgccacag tagctctcgg cttcaaggac 960
 gaggacgtgc acaacgaata cgacggtacc gggttcgtga tctcgcggaa ttcggacttc 1020
 agtattactg cctgcacctg gacgaacaag aagtggccac acacagcacc caaaggtaag 1080
 accttgctga gggcttatgt gggtaaggcg ggggacgaga gcatagtgga gcagtctgac 1140
 tcgcagatcg tcagcatcgt actggaagac ctgaagaaga tcatggacat caaggccgac 1200
 ccggagttga ccaccgtcac acggtggaaa acctcaatgc cacaatatca tgtcggacat 1260
 cagaaggcca tctccaacat gcgcgagacc ttcaagcagt cttaccggg cgtgtatatt 1320
 accggagcgg ctttcgaggg ggtcggcatc cctgactgca tagaccaggg gaaggcggcc 1380
 atcagcgagg ctgtgtcgta ctttttctcg tga 1413

<210> 199
 <211> 1413
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 199
 atgcatgaca accagaagca cctggttatc attggcgggtg gcataaccgg gctcgcgcc 60
 gccttttacc tggagaagga ggtggaggaa aaaggcctcc caatccagat cagtttgata 120
 gaggcgagtc cgcgcttggg gggcaagatc cagactctgt ataaagatgg atacattatc 180

gagagggggtc cagacagctt cctggagcgc aaggtctccg ggcctcagtt ggcgaaggat 240
 gttgggttgt cagatcagct cgtgaacaac gaaacgggcc aggcgtatgt gttagtcaat 300
 gaaactctgc accccatgcc caagggcgcg gtgatgggga taccaccca gatcagtccc 360
 ttcatcacia ccggtctgtt ctcggtcgca gggaaggccc gagcggcgat ggattttgtc 420
 ctgcccagt cgaagcagac cgaggaccag agcctcgggg agtttttcag gcgcagagtt 480
 ggcgatgagg tcgtcgagaa cctcattgag ccgcttctca gcgggattta tgcgggagac 540
 atcgacaggc tctccctgat gtcaactttt ccgcagttct accagacgga gcaaaagcac 600
 aggagcctga ttctgggaat gaagaagtca caacaacatg ctaaagccca gcaggtaact 660
 gcaaagaagc agggtcagtt ccaaacaatc aatcaaggtc tccaggcact cgtcgaggcc 720
 gtggagtcaa agctaaagct gaccaccata tacaagggtta ccaaagtga acaaatcgag 780
 aagacagacg gcggttacgg agtgacgctt gactccggcc agaccctcct cgccgactct 840
 gcgatcgtga ccacgcgcga ccagtccatc tactctatgt tcccaagga ggccgggctc 900
 gaatatttgc acgatatgac cagcaccagc gtcgctacgg tagcactcgg gttcaaggag 960
 gaggacgtcc acaacgagta cgatggcact ggcttcgtga tcagccgtaa ctctgatttc 1020
 agcatcactg catgcacatg gactaataag aaatggcccc aactgcacc caagggcaag 1080
 acgctgctgc gagcctacgt cgggaaggcc ggggacgagt ctattgtaga gcagagcgat 1140
 caccagattg tgagtatcgt actggaggac ctgaaaaaga tcatggatat aaaggcggac 1200
 ccagagctga ctaccgtgac ccgctggaaa acatccatgc cgcaatacca tgtggggccac 1260
 caaaaagcga tctccaacat gcgggagacg ttcaagcaat cttatcccg cggtgtacatc 1320
 acgggagccg cgttcgaggg cgtgggcac ccggattgca tcgatcaggg taaggctgcg 1380
 atatcgagg ctgtcagtta cctgttttct tag 1413

<210> 200

<211> 1434

<212> ADN

<213> *Paenibacillus macerans*

<400> 200

gtgagcaaaa aaatcgccgt catcggcgga ggcataaccg ggtaagcgt ggcttattac 60
 gtgcgtaaat tgctgcgtga acagggggta aacgctgggg ttaccctcgt ggaacagtcc 120
 gatcggctgg gcggcaaaat ccgttcccta cgacgtgacg gctttacgat agaacagggc 180
 ccggattcaa tgatcgcgcg caagcccgcc gcgctggaat tgatccggga actcgggctg 240
 gaggataagc tggcggaac gaatccgcag gcgaagcgaa gttatatatt gcatcgcggc 300

aaattccatc ccatgccgcc ggggctgatg ctccggcatac cgacgcaa at gtggccgatg 360
 gtcaagacgg ggctgctctc tccggccggc aagctgcggg ccgcgatgga tctgctgctt 420
 cccgcgcggc gcggcgggcg cgacgaatcg ctccggcggt tcatccgccg ccggctcggc 480
 agagaagtgc tggagcagat gacggagccg cttctagccg gcatatatgc cggggacacc 540
 gaacagctta gcttgaaagc gacgtttccg cagtttatgg agatggagcg caagcaccgc 600
 agcctgatcc ttgggctgct ggccggcaaa aagcagccgc cgcggccggg gggaagccag 660
 gtcccgcctgc cgaaggccgc gcaaaccagc atgtttctga cgttgacggg cggtttgag 720
 ggactgacgg aagcgctgga ggaatcgcta agcgaagaga aaataattac cggccaggcg 780
 gtaaccggac tgtcgcagca agaggcgggt tatgagctta acttaagcg gggcgagcgt 840
 ttgaacgcgg acggagtc at tttggcagtt cctgcttttg ctgcggcccg gctattggat 900
 ggcgttcccg aagccgctta cctggagcgg atccgttatg tgtccgtggc caatttagcc 960
 ttcgcctacc ggcggaaga cgttccgcac gatttgaacg gctccggcgt gcttatcccg 1020
 cgcggggagg ggcgaaatgat tacggccatt acctgggttt cttcgaaatg gctgcattcg 1080
 gctcccggcg ataaagcgct gctgcgagcc tatatcggcc gcctgggcca cgaggcatgg 1140
 accgcgatgt gcagggccga catcgagcgc cgggtggccg ccgagctgcg cgatttgctg 1200
 ggcatcgccg ccagcccgt gttttgcgag ctgcccgctt tgccggagtc gatgccccaa 1260
 tatccggtcg ggcatgtcga gcggcttgag gcgctgcgcg gggcattgtg ccgggcgaag 1320
 ccggggctgc tgctgtgcgg cgcgggatat gccggcgtag gcattcccga ctgcatccgg 1380
 cagggcaagg aagccgctga aagcatggcg gcttatattga gggatggacg gtga 1434

<210> 201
 <211> 1482
 <212> ADN
 <213> *Paenibacillus thiaminolyticus*

<400> 201
 atgaaagctc tgcggaaaact tgtcgttatc ggtggcggaa ttacgggatt gagcgcggcg 60
 ttctatgcgc tgaagcaggc ggatgaagag gggcagcca tctccgttac catcatagag 120
 caatcggacc gtctcggcgg gaagatacag accctgcgga aggaaggggtg tgtcattgag 180
 aaaggcccgg actccttctt cgcccgggaag ctgccgatga tcgatttggc gcgcgacctc 240
 ggaatggatt ctgaattggt cgccacgaat ccgcatgcca aaaaaacata tatattgcgc 300
 cggggcaagc tgtaccggat gccgcccggc ctctgtctgg gcatcccgaac ggagctgggg 360
 ccgttcgcga agacagggt catctccccg tgggggaagc tgcgcgcggc tatggatctg 420

ttcatcaagc cgcacccggc ggatgaagat gaatccgttg gcgcgttcct ggacagacgg	480
ctcggacgcg aagtgcgga gcatattgcc gagccgctgc ttgccggcat ttatgccgga	540
gatttgcagg cgctgagcct gcaggccacc ttcccgcagt tcgcgcaggt ggagcggaag	600
cacggtggcc tgatacgcg aatgaaggcg agccgccaaag caggccaatc ggtaccgggg	660
ctgccggatg tcgccaaagg aacgatgttc ctgacattcc gcaacggctt gacctcgctc	720
gtcgaacggc tggaggagac gctgcgggac cgggccgaat tgtgccttgg catcggcgcg	780
gaaggattcg agaagcggga ggacggaacg tatctggtgc gcttgagcga tgggagcagg	840
ctgcaggcgg atgccgtcat cgtgacgacg ccttcgtatc atgcggcatc cttgctcgag	900
gagcatgtcg atgcgagcgc cttgcaggcg atccgtcatg tatccgtcgc gaatgtcgtc	960
agcgtgttcg atcgcaagca ggtcaataat cagttcgacg gcacagggtt cgtcatctcg	1020
cgccgggaag gccgggcatg tacggcctgc acgtggacct cggatgaagtg gccgcatacg	1080
agccgcgggg acaagcttat tatccgctgc tacattggcc gggccggtga cgaggaacgg	1140
gtggactggc cggacgaggc gctcaagcgg acggtgcgca gcgagctgcg ggagctgctt	1200
gatatcgata tcgacccgga gttcgtcgag attacgcgcc ttcgccactc gatgccccag	1260
tatccggtcg gccatgtgca ggcgatccgc tcgctgaggg acgaggtggg gcgcacgctc	1320
ccaggcgtgt tcctggcagg acagccgtac gaaggggtcg gcatgcccga ttgcgttcgc	1380
agcggccgcg atgcggcgga agccgcgggt agcgcgatgc aggccatgag tacggagcca	1440
gaggcgccag ccgaggatgc cgctactgga acggcggggt aa	1482

<210> 202

<211> 1425

<212> ADN

<213> Paenibacillus polymyxa

<400> 202

atgggtgata agaaacgccg tgttgttggt gtcggcgggtg gccttaccgg cctcagcgcg	60
gcattttata tccgcaagca ttaccgggaa gcaggagttg aacctgtgat tactttggtc	120
gagaaaagct cgtccatggg aggcattgatt gagacactgc accgggatgg atttgtgatt	180
gaaaaagggc ccgattcgtt cctggctcgc aaaacggcaa tgattgatct ggccaaagaa	240
ttggagatcg atcatgagct ggtaagtcag aatccggagt cgaagaaaac gtatatcatg	300
cagcgtggca agcttcatcc tatgccagca ggacttgctc tcggtattcc gacagaacta	360
agaccattct tgagaagtgg tttggtttct ccggcaggca aactgcgggc gttgatggat	420
tttgtcatcc cgccgcgtcg tacaacagag gatgaatcgc tcggttatat gattgaacgc	480

cgtcttggag cagaagtgct ggagaacttg acggaaccac tgctcgcagg aatctatgca 540
 ggtgatatgc ggcgattgag cctccaggct accttcccg c agttcggaga agtagagcgc 600
 gattacggca gcttgatccg gggcatgatg acggggcgca aaccggctga gacgcatacc 660
 ggaacaaaac ggagcgcttt tttgaacttt cgccagggac ttcagagcct tgttcatgca 720
 ctcgccatg agttgcagga tgtggatcaa cgtctgaaca ctgcggtgaa atcgctgcaa 780
 cgccttgatg gagcgcagac cagataccgt gttgaacttg gtaatggcga aatgcttgaa 840
 gccgatgatg tagtggttac tgtgccgaca tatgtcgcgt cggagctggt gaagcctcac 900
 gtggacacag cggcactgga tgcgattaac tatgtgtctg tagccaatgt agtgctcgt 960
 tttgagaaaa aagaggtgga gcatgtattc gacggatcgg gtttcctcgt tccgcggaaa 1020
 gagggctcga atattacggc ttgcacgtgg acatcgacga aatggctgca taccagccc 1080
 gatgataaag tactgcttcg ctgttatgtt ggtcgcctcc gtgacgaaca gaacgtagag 1140
 cttccggatg aagcgtgac gaatctcgtt ctcaaagatc tgagagagac gatgggtatt 1200
 gaagcagtgc cgatcttctc cgagattaca aggcttcgta aatccatgcc acagtatccg 1260
 gtgggacacc ttcaacatat tgccgctctc cgtgaggagc ttggcagcaa attaccgggt 1320
 gtgtacattg caggtgcagg ttatgagggc gtaggcttgc ctgattgcat cagacaagcg 1380
 aaggaaatgt ctgttcaggc tacacaagag cttgcagcag attaa 1425

<210> 203

<211> 1413

<212> ADN

<213> *Bacillus atrophaeus*

<400> 203

atgagtgacg gcaaaaagca tcttgtcatc atcggcggcg gcatcacggg attggcctcc 60
 gccttctata tgaaaaaga aatcagagag aaaaatttgc ctctttctgt gacgttagtc 120
 gaagcaagcc cgagagttgg cgggaaaatt caaacggccc gcaaggacgg ttatattatt 180
 gaaagagggc cggactcatt tttagaaaga aaaaaagcg caccggagct tgtcgaagat 240
 ttaggccttg agcatttgct tgtcaacaat gcgacggggc agtcttatgt gctggttaac 300
 gaaacgcttc acccgatgcc aaagggcgct gttatgggca tacctactaa aatagcgcca 360
 tttatgtcta ccggcttatt ttcattttcc ggcaaagcgc gcgcggctat ggatttcgtt 420
 ttgcccgcaa gcaagccgaa ggaagatcag tccctgggtg aattcttccg caggcgtgtc 480
 ggtgacgaag ttgttgaaaa tttgattgag ccgctattat ccggcattta tgcgggtgac 540
 attgacaggc tcagcctgat gtcgacgttc ccgcagtttt atcagaccga acaaaagcac 600

agaagcttga tcctcggcat gaaaaaaaca aggcctcagg gctccggaca gcggttaacg 660
 gctaaaaaac aagggcaatt ccaaacctta aagaccggct tgcagacact cgtcgaagag 720
 ctggaaaacc agctgaagct gacgaaggta tacaagggtta caaaagtaac caatatcagc 780
 cgcggggaaa agggctgctc catcgctctt gataacggga tgacgctgga tgccgatgca 840
 gcgattgtaa cctcaccgca caaatcggct gccggaatgt ttccggatct gccagctgctc 900
 agtcagttaa aagacatgca ctctacctct gtggcgaatg tcgcgcttgg ctttccacaa 960
 gaggctgtcc aaatggaaca tgaaggaacg ggttttgtca tctcaagaaa cagtgatttt 1020
 tcaataacgg cctgtacttg gacgaataaa aaatggccgc actctgctcc ggaaggcaaa 1080
 acgctcctca gggcttatgt cggaagcg ggtgatgaat caatcgtcga actgtctgat 1140
 aatgagatta tcaaaattgt attagaagac ctaaagaaag tcatgaaaat caaaggcgaa 1200
 cctgaaatga cgtgcgtcac acgctggaat gagagtatgc cccaatatca tgtcggccac 1260
 aaacagcgta taaaaaaagt gcgcgaagca ctggctgctt cctatccggg agtttacatg 1320
 acgggcgctt cattcgaagg cgttgggatt ccggactgta tcgatcaagg gaaaagcgcc 1380
 gtttcagacg tacttgctta tttattcggg tga 1413

<210> 204

<211> 1413

<212> ADN

<213> *Bacillus atrophaeus*

<400> 204

atgagtgacg gcaaaaagca tcttgtcatc atcggcggcg gcatcacggg attggcctcc 60
 gccttctata tgaaaaaga aatcagagag aaaaatttgc ctctttctgt gacgttagtc 120
 gaagcaagcc cgagagttag cgggaaaatt caaacggccc gcaaggacgg ttatattatt 180
 gaaagagggc cggactcatt tttagaaaga aaaaaagcg caccggagct tgtcgaagat 240
 ttaggacttg agcatttgct tgtcaacaat gcgacggggc agtcttatgt gctggttaac 300
 gaaacgcttc acccgatgcc aaagggcgct gttatgggca tacctactaa aatagcgcca 360
 tttatgtcta cccgcttatt ttcattttcc ggcaaagcgc gcgcggctat ggatttcggt 420
 ttgcccgcaa gcaagccgaa ggaagatcag tccttgggtg aattcttccg caggcgtgctc 480
 ggtgacgaag ttgttgaaaa tttgattgag ccgctattat ccggcattta tgcgggtgac 540
 attgacagac tcagcctgat gtcgacgttc ccgcagtttt atcagaccga acaaaagcac 600
 agaagcttga tcctcggcat gaaaaaaaca aggcctcagg gctccggaca gcagtttaacg 660
 gctaaaaaac aagggcaatt ccaaacctta aagaccggct tgcagacact cgtcgaagag 720

ctggaaaacc agctgaaact gacgaaggta tacaagggta caaaagtaac caatatcagc 780
 cgcggggaaa agggctgctc catcgctctt gataacggga tgacgctgga tgccgatgcc 840
 gcgattgtga cctcaccgca caaatcggtt gccggaatgt ttccggatct gccagctgtc 900
 agccagttaa aagacatgca ctctacctct gtggcgaatg tcgcgcttgg ctttccacaa 960
 gaggctgtcc aatggaaca tgaaggaacg ggttttgtca tctcaagaaa cagtgttttt 1020
 tcaataacgg cctgtacttg gacgaataaa aaatggccgc actctgtctc ggaaggcaaa 1080
 acgctcctca gggcttatgt cggaaaagcg ggtgatgaat caatcgctga actgtctgat 1140
 aatgagatta tcaaaattgt attagaagac ctaaagaaaag tcatgaaaat caaaggcgaa 1200
 cctgaaatga cgtgcgtcac acgctggaat gagagtatgc cccaatatca tgtcggccac 1260
 aaacagcgta taaaaaagt gcgcgaagca ctggctgctt cctatccggg agtttacatg 1320
 acgggcgctt cattcgaagg cgttgggatt ccggactgta tcgaccaagg gaaaagcgcc 1380
 gtttcagacg tacttgctta ttattcgaa tga 1413

<210> 205
 <211> 1434
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 205
 atgtcaaaga agattgcagt cattgggtggg gggataacag gggtgtccgt ggcctactac 60
 gtgaggaagc tgcttcggga gcaaggcggt aatgcggggc ttaccctcgt cgagcaatcc 120
 gaccgcctcg gcgggaagat tagatccttg agacgagacg gctttaccat tgagcaaggc 180
 cctgactcta tgattgcacg taagcccgca gctctcgaac ttatccgtga gcttggctctg 240
 gaggacaagt tggcgggcac aaacctcaa gccaaacgct cctacatact gcaccgtggc 300
 aagtttcatc cgatgccacc tgggctgatg ctggggatcc ccaactcaaat gtggccaatg 360
 gtcaagaccg ggctgctatc tccggccgga aagctacggg ctgcgatgga cctacttctt 420
 cctgcaaggc gcggaggcgg cgacgaatca cttgggtggg ttatccggag gcggcttgga 480
 cgtgaggtgt tggagcagat gaccgaacca ctcttgctg gaatctatgc tggcgacaca 540
 gaacagcttt cacttaaagc gacctttcct caattcatgg agatggaaaag gaaacatcgc 600
 agtctcatcc ttggactatt ggctgggaag aaacagccac cgcgtcccgg tggtagccaa 660
 gtgccgctcc caaaggccgc tcagaccagt atgttcttga cactcaccgg cgggttgga 720

ggtctgaccg aagcactaga ggaaagccta tcagaggaga agataattac tggccaagca 780
 gttaccggac tttcgcagca agaggccggg tatgagttaa atctctctgg cggagagaga 840
 cttaatgcag acggagtgat cctcgcagtc ccagcgttcg ctgccgcccg acttcttgac 900
 ggcgtagcctg aggccgccta cctagagcgc atccgctatg tcagtgttgc taatttggcg 960
 ttcgcttaca ggcgtagga cgtgcctcat gatctgaatg ggtccggcgt gttaatccct 1020
 agaggtgaag ggaggatgat tacggccata acttgggttt cgtccaaatg gttgcattca 1080
 gcacccggtg acaaggcact gctgagagcg tacattgggc gactaggtga tgaggcttgg 1140
 acagccatgt gtagggccga catcgagcgt agagtcgccg ctgaactccg cgatctacta 1200
 ggaattgccg ctagtctttt gttctgtgaa ctagccgcac tcccagaatc tatgccgcag 1260
 tatccagtgg gtcacgtcga acgactcgaa gccttgcgag gagcattgtg tcgcgctaaa 1320
 ccagggttgt tgttgtgtgg tgccgggtac gctggcgttg gcattccaga ctgcattcgg 1380
 caaggcaaag aagccgctga gtcgatggcg gcttatttga gggacggacg ctag 1434

<210> 206
 <211> 1482
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 206
 atgaaggctc tgaggaaact tgtggtcacg gccggaggga tcaactgggct ttcggccgcc 60
 ttctatgcac taaagcaagc cgatgaggaa gggcagccca tctcggtcac cataattgaa 120
 cagagcgata ggctcggcgg aaagatccag acaactccgca aggagggctg cgtaattgag 180
 aagggcccgg attccttcct cgctaggaag ttgccgatga ttgatctagc tcgggatcct 240
 ggcattggact ccgaattggg gccgactaat ccgcacgcaa agaagactta catcttgagg 300
 cgcggaagc tctaccggat gcctccaggc ttagtgcttg gcatacctac ggaactagga 360
 ccattcgcta agacagggct cattagccct tggggcaaac tccgcgccgc tatggatttg 420
 ttcattaagc ctcatccagc cgatgaagac gaaagtgttg gcgctttcct ggacagacgt 480
 ctcggtaggg aagtgaccga gcacattgcy gaacctttat tggcgggcat ctacgcgggc 540
 gacttgcaag ccttaagcct tcaagccact ttcccacagt ttgcacaagt agagcgcaag 600
 cacggagggc tgatacgcg tatgaaggcc agcagacagg ccggtcagtc cgtgcctggg 660
 ctgccggacg tcgccaaggg tacgatgttc cttacctttc gcaacgggct taccagctta 720
 gttgaaaggt tggaggaaac tctcagagac agggctgaac tctgtctggg catcggcgca 780

gaaggggtttg agaaacgtga agatggaaca taccttggtc gactaagcga tggttcgagg 840
 ctccaggccg acgcagtaat tgtcactacg ccgagctatc atgcggcatc cctgttgagg 900
 gagcatgtgg atgcttcggc cctccaggcc attcgatcatg taagcgttgc aaatgtcgtt 960
 agcgtcttcg accgaaagca agtgaataac cagttcgacg gcacaggggtt tggtatctca 1020
 cggcgagaag gtcgcgcaat caccgcctgt acctggacat ccgtgaaatg gccgcatact 1080
 tcgcgcgggc acaaactgat tatccgggtg tacatcggta gggctggcga cgaggagcga 1140
 gtggattggc ccgatgaagc tctcaagcgt actgtaagat cagaactgcg tgagttgctg 1200
 gacattgaca ttgatccgga atttgtggag attacacgac tcaggcactc tatgcctcaa 1260
 taccagtcg gccacgtcca ggctatccgc tctttgaggg acgaggtcgg taggacttta 1320
 ccgggctgtg tccttgctgg gcaaccctac gaaggtgtgg gaatgcctga ctgtgtgagg 1380
 tccggccggg atgccgccga agcagcagta agtgctatgc aagcaatgag tacagaacca 1440
 gaagcaccgg cagaggacgc cgctactgga acggcggggtt ga 1482

<210> 207
 <211> 1425
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 207
 atgggagaca agaagcggag agttgttgtt gttggcggcg gcttgactgg cctaagcgcc 60
 gccttctaca tccggaaaca ttatcgagaa gctggagttg agcccgtcat cacgcttgtt 120
 gagaaatcta gctcgatggg agggatgatt gagacccttc atagggacgg gtttgtcatc 180
 gagaagggcc cggacagttt cttggcacgg aagaccgcaa tgattgatct ggcgaaagag 240
 ctggagattg accacgagtt ggtcagccag aatccagaat cgaagaagac ctacataatg 300
 caacgtggaa agctgcaccc tatgccagcg ggacttggtc tgggcattcc caccgaattg 360
 cgtccctttc tccggagcgg gcttgtctca cccgctggga agttgcgggc gctgatggac 420
 ttcgtaatac cgccacgaag gacgaccgaa gatgagtcac tcgggtacat gatcgagcgc 480
 cgactgggtg ccgaggtgtt ggagaacctc acagagccgt tgctcgctgg aatctacgct 540
 ggcgacatga gaagattgtc cctccaggct acgtttccgc agttcgggtga ggtggagcgc 600
 gactacgggt ccttaatcag aggaatgatg accggacgta agcctgcgga gacacacaca 660
 gggaccaaga ggtctgcctt tctcaatttc agacagggtc tgcaatcact ggttcacgcc 720

ttagtccatg aactccagga tgtagatcag aggttaaata ctgcggtgaa gtcgcttcag 780
 aggcttgacg ggcacaaaac ccgttatcgc gttgaactcg gcaatggcga aatgcttgag 840
 gctgacgacg tgggtggttac tgtaccaacc tacgtggcga gcgagcttct taagccgcac 900
 gtggacacgg cggcgttaga cgctattaac tatgtgtcgg tggctaattgt agttcttgca 960
 ttcgagaaga aggaagtaga gcacgtcttc gatggatcgg gcttcttggt gcctcggaag 1020
 gaggggaagga acataaccgc ctgcacctgg acttcgacca agtggctcca cacatcacca 1080
 gatgacaagg ttctgttacg ttgttacgtg ggcagaagtg gagatgagca gaatgtggaa 1140
 ctcccggatg aggcactcac taatctggtg cttaaggatc tgagagagac gatgggcatc 1200
 gaggcggttc caatcttctc agagattacc cggctccgca agtcaatgcc gcagtacca 1260
 gtaggacatc tccagcacat cgccgcattg cgcgaggaac tcggctctaa gctaccagga 1320
 gtgtacatcg ccggagcggg ctacgagggc gttggtcttc cggattgcat tcgccaggcc 1380
 aaagaaatgt cagtccaggc aacgcaagaa ctcgctgccg actga 1425

<210> 208
 <211> 1413
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 208
 atgagtgacg ggaagaagca cttgggttata atcggggggg gaataaccgg cctggccagc 60
 gctttctata tggagaagga gatccgggag aagaacttac ctctcagcgt gaccttggtg 120
 gaggcattccc cgcggttagg ggggaagatc cagactgctc gaaaggacgg ctatatcata 180
 gagcggggcc cggacagctt cctggagcgc aagaagtcgg cgcccagatt agtcgaggac 240
 ctcggtctcg agcacttact cgtaaacaac gctacagggc agtcttacgt cctcgtcaac 300
 gaaacactgc acccgatgcc caaaggcgcg gtgatgggaa tccccactaa gattgcacct 360
 ttcatgtcga ctggcctttt cagcttcagt gggaaggcga gggcggaat ggacttcgtc 420
 ctgcccgcgt ccaagccgaa ggaggatcag agcctcggcg agtttttccg caggcgagtt 480
 ggggatgagg tcgtggaaaa cctcattgag cccttgctat ccggaatcta tgccggagac 540
 atcgacaggc tcagccttat gtctactttc cccagttct accagacaga gcaaaagcac 600
 cgaagtgttg tctcgggat gaagaagacg cgtcctcagg gttctggtca gaggctaaca 660
 gcaaagaagc aggggtcaatt ccagacgctt aaaacagggc ttcaaact tgtggaggaa 720
 ctcgagaatc agcttaaact aaccaaagtg tacaagggca cgaaggtaac taacatcagc 780

cgcggtgaaa agggctgcag catcgcaactt gacaacggga tgacactgga cgcggaacgca 840
 gcaatcgtca cgagccccc caaatcagcg gcgggaatgt tccccgacct tccggcggtc 900
 agccagctga aagacatgca ctccacctcc gtcgcaaacg tcgcgctcgg cttcccgag 960
 gaggtgtcc agatggagca tgaggggact ggcttcgtta tcagcagaaa ttcggacttc 1020
 agtatcacag cgtgcacttg gacaaacaag aaatggcctc acagcgcacc tgaggggaag 1080
 acacttttgc gagcgtacgt ggggaaagct ggggacgagt ccatagttga actaagcgac 1140
 aacgagataa ttaagatcgt gcttgaggac cttaagaaag tgatgaagat aaagggcgag 1200
 cccgaaatga catgcgtaac tagatggaat gagtccatgc cacagtacca cgtcgggcac 1260
 aagcagcgta tcaaaaaggt cagggaggct ttggcggcct catacccggt cgtatacatg 1320
 accggtgcat ccttcgaggg ggtggggata ccagactgca tcgaccaagg caaatccgca 1380
 gtctcagacg ttttggcata cttgttcggc tag 1413

<210> 209
 <211> 1413
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 209
 atgtcggatg gcaagaagca cctcgtcatc atcggcgggg gtatcacagg acttgcgtcc 60
 gcgttctaca tggagaagga gatcaggag aagaacttgc ccctctcagt gaccctgggtg 120
 gaggcctcgc cccgtgttgg tggttaagatc cagacagcgc gaaaagacgg ctacattatc 180
 gagcgggggc ccgactcctt cctcgagagg aagaagtctg cccccgagct tgtggaggac 240
 ttggggcttg agcacctcct cgtgaacaat gcgaccgggc agagctacgt tttggtgaac 300
 gagaccctgc acccgatgcc caagggagcc gtgatgggga tccctaccaa gatcgcgcct 360
 ttcattagca ctcgactttt ttcattcagc ggcaaggcca gagccgctat ggactttgtt 420
 ctcccggctt ctaagcctaa ggaagaccag agtctaggcg aattcttcag gcgaagagtc 480
 ggcatgagg ttgttgagaa cttatagag ccattattgt caggtatata cgcaggagac 540
 attgacaggc tgtctctcat gactaccttc cctcaattct accagacgga gcagaaacac 600
 aggagcctca tattggggat gaagaagacg cgtcctcaag gaagcggaca gcagttgacg 660
 gccagaagc agggccagtt ccaaacgctc aagaccggac ttcagaccct cgtcgaggag 720
 cttgagaacc agctaaagtt gacgaagggt tacaagggca ctaaggtcac aaacatctcg 780

aggggcgaga agggatgcag catcgcggtta gacaacggga tgaccctaga cgctgacgca 840
 gctattgtga ctagcccca taagtccgca gccggcatgt ttccagactt gccggccggt 900
 agccagttga aggacatgca ctcgaccagc gtggcaaacg tcgcattggg cttcccacag 960
 gaggcggtgc agatggagca tgaggggacc ggattcgtga tctcaaggaa ttccgatttc 1020
 tccattacgg catgtacctg gacaaacaaa aaatggcccc acagcgcccc agaagggaaa 1080
 acactcctac gcgcttatgt tggcaaggcc ggcgatgagt caattgtgga gctctccgac 1140
 aatgagatca tcaaaatcgt tcttgaagat cttagaagg taatgaagat taagggggaa 1200
 ccggaaatga cgtgtgtgac aagggtggaac gagagtatgc cccaatatca cgtgggccac 1260
 aagcagagga taaagaagggt gagggaggcg ttggcggcgt cttaccccg cgtgtacatg 1320
 acgggggctt cattcgaggg ggtgggcctc ccgactgca ttgaccaagg caaaagcgcg 1380
 gtgtctgacg tgctcgcgta cctgttcgag tag 1413

<210> 210
 <211> 1413
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 210
 atgtccgacg ggaagaagca cctggtaatc atcgggtggtg ggatcacccg tctggcttca 60
 gcgttctaca tggaaaagga gatccgggag aagaacttgc ccctttcggg gactctagt 120
 gaggcctctc cacgggtggg gggcaagatt cagaccgcgc gcaaggatgg ctacatcata 180
 gagcgaggac cagactcatt cctagagcgt aagaagtccg cccagagct cgctgaggat 240
 ctcggtctag agcacttgct agtgaataac gctacaggac agtcctacgt gctcgtgaac 300
 gagacactac acccgatgcc taagggggct gtcattgggta taccgaccaa gatcgccccg 360
 ttcatgtcca ctgcctttt ctcgttctcg ggcaaagctc gggccgctat ggatttcgtc 420
 ttgcctgcct cgaaaccgaa ggaggaccag tccttaggag agttcttccg ccggagggtc 480
 ggcgacgagg tgggtggagaa cttaatcgaa cccttgctct cggggatcta cgctggagac 540
 attgatcgac tatcgcttat gtctacgttt cctcaatttt accagacgga gcagaagcac 600
 cgtagcctca ttttgggtat gaagaagaca cggcctcaag gttcggggca gcagcttact 660
 gccagaagc agggccaatt ccagacactc aagaccggct tgcagactct agtggaggag 720
 ctggagaatc aattgaagct gacaaaggtc tacaagggta ccaagggtgac aaacatatcg 780
 cgtggcgaaa agggatgctc cattgccctc gacaacggta tgaccctcga cgccgacgca 840

gcgattgtga cgagcccaca caagagcgcc gcgggcatgt tcccggactt gcctgcagtg 900
 tcacagctga aagacatgca ttctacatcc gtcgccaacg tcgccctggg ctttccccag 960
 gaggctgtgc agatggagca cgaggggacg ggcttcgtta tcagccgcaa ctccgacttt 1020
 tctattaccg cgtgcacatg gaccaacaag aagtggccgc acagcgctcc ggaggggaaa 1080
 acacttctcc gagcatacgt aggcaaggcc ggggacgagt caattgttga gctctccgac 1140
 aatgaaatca ttaaaatagt tctggaggat cttaagaagg taatgaagat aaagggggaa 1200
 cctgaaatga cgtgtgttac ccgctggaat gagtcaatgc ccagtagca tgtgggacac 1260
 aagcagagga taaagaaggt gagggaggcg ctgctgcgt cctaccaggg ggtctacatg 1320
 acaggagcga gttttgaggg ggtgggtatt cccgactgta tcgaccaggg taagtcggca 1380
 gtgtctgacg tgctcgctta cctattcgag tag 1413

<210> 211
 <211> 1434
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 211
 atgtcgaaga agatcgccgt tatcgggtgga ggcattacag ggctctcggt cgcctactac 60
 gtgcgtaagc tgcttcgtga gcaaggcgtc aacgctgggtg tgacgctgggt tgagcagtct 120
 gatcgcctcg gtggaaaaat ccgtagcctt cgcagagacg ggttcacgat tgaacaagga 180
 ccagattcca tgatcgcgcg caaaccgcg gcgttgagc taattcgaga actcggactc 240
 gaggacaagc tcgccggcac taaccacag gcaaagcggc cgtacatcct tcaccgcggg 300
 aagttccacc cgatgcccc aggctgatg ctcggcaccc cgaccagat gtggccgatg 360
 gtcaagaccg ggctcctgtc tccgcgggg aaactaaggg ccgctatgga cctcctcctc 420
 cctgctcgga ggggcggcgg tgatgagagt ctgggggat ttatcaggcg gagattaggg 480
 cgcgaggtag ttgagcaaat gaccgaacca ctgctcgag gtatctatgc aggcgatacg 540
 gaacaactgt ccttgaaagc aacatttcca caattcatgg agatggaaaag aaaacatagg 600
 tccctcatac tcggtcttct tgctggaaaa aagcaacctc cgagaccggg tggttcacia 660
 gtgcctctgc ctaaagcggc gcaaacttca atgttctga ctctgacagg cgggctcgaa 720
 ggccttaccg aagctctaga ggaatccttg tctgaggaaa aaataatcac cggccaggct 780
 gttaccgggc ttagccaaca ggaagccggg tatgaactga acctttcagg tggagagagg 840

ttgaacgccg atgggggtcat attggctgta ccggcggttcg ccgcggtcgc cctgctggac	900
ggcgcccttg aggccgcgta tttggagcgc atacgctatg tttctgttgc gaacctcgct	960
tttgcatata gacgggaaga tgtgccccat gatcttaatg gttccggagt gttgatccca	1020
cgcgggggagg gtcgaatgat aacggcaatt acttggggttt ccagcaagtg gttacattcg	1080
gctcctgggg ataaagctct tttgcgggca tacatcggac gtctcggcga cgaagcctgg	1140
acggccatgt gcagagccga cattgagcga cgggtcgcgc gagagctgag agacttggtg	1200
ggcatagctg catctccatt gttctgcgag ctggctgcat tgcctgaaag catgccgcaa	1260
tatccagtag ggcatgtgga gcgcctcgaa gctctccgag gcgcgttggt tagggcgaaa	1320
cctggactgc tgctctgcgg tgccggctat gcaggtgtgg gaattcctga ctgtatcagg	1380
caaggtaaag aagcggcaga gtccatggcc gcttacctta gggatgggcg ctag	1434

<210> 212
 <211> 1482
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 212	
atgaaggcgc tgcggaagct ggtggtaatc ggggggggga tcacggggct gtcggccgcg	60
ttctacgcac tcaagcaggc cgacgaagag ggtcagccaa tttccgtaac gattatcgag	120
caatccgata gacttggcgg caagatacag accctgagaa aggagggatg cgtcattgaa	180
aagggaccag attcatttct ggcgaggaag ctcccatga tcgatctggc gagagactta	240
ggcatggact cggagctggg gccacaaaat cctcatgcaa aaaagactta catcctacgg	300
cgcggtaatg tgtaccgcgc gccaccgggc ctgggtgttg ggattcctac cgagttagga	360
cccttcgcga aaaccggact catcagcccc tgggggaaac ttcgagccgc gatggacctt	420
ttcatcaaac cacatccagc cgatgaagat gagtctgtgg gagctttttt agatagacgt	480
ttaggtcgcg aggtgacgga gcacatcgca gagccgctgc tcgccgggat atacgcaggc	540
gatcttcaag ctttgtcctt gcaagctacg ttccctcagt tcgcgcaagt ggaacgcaaa	600
cacggaggtc tcatcagagg tatgaaagcg tctcgccaag ctggacagtc agtcccaggg	660
ctcccagatg tggccaaggg taccatgttt cttactttca gaaatggttt gactagcctg	720
gtggagcgtc tcgaagaaac ccttcgagat agagccgagc tctgtctggg tatcgggtgca	780
gaggggtttg aaaaacggga agacggcacg taccttgttc gattatctga tggctccaga	840
ttgcaagccg acgccgttat agttaccaca ccatcatacc atgccgcctc cctactggag	900

gagcacgtcg acgccagcgc gttacaggct atccgccacg tatctgtagc caacgtggtg 960
 agcgtttttcg ataggaagca ggttaacaat cagtttgatg ggacagggtt tgttatctca 1020
 agacgcgaag gcagggctat cactgcttgc acttggaact cagttaagtg gccgcatacc 1080
 agccggggggg ataagttgat aatccggtgt tacattggtc gtgcaggaga tgaggagcgc 1140
 gtggattggc cagacgaagc gctaaagcgg accgtgagaa gtgagcttcg cgagctgtta 1200
 gacatagaca tagatcccga attcgtggaa attacacggt tgaggcactc tatgccacaa 1260
 taccctgttg gtcattgtgca agctatacgg tccctgcgcg acgaagtagg ccggaccttg 1320
 ccgggcgtgt ttcttgccgg tcagccgtat gagggggttg ggatgccaga ttgtgtgcgt 1380
 tctggccgcg acgcggcaga ggctgccgta tcagccatgc aagccatgac gacagaaccc 1440
 gaagccccgg cggaagatgc agcgacagga actgcagggt ag 1482

<210> 213
 <211> 1425
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 213
 atgggggata agaagaggag ggtcgttgtc gtgggtgggg gactgaccgg actatcagcc 60
 gcgttctaca ttagaaagca ctaccgagag gccggcgtgg agccggtgat cacgctggtc 120
 gagaagtcga gttcgtatgg cggaatgatc gagaccctac acagggacgg ctttgtgatt 180
 gagaaggagc cagatagctt ccttgcacgc aagacagcca tgatcgatct cgaaaagag 240
 ctcgagatag accacgaact ggtgtctcag aaccgggagt ccaagaagac atatatcatg 300
 cagagaggta aactacaccc catgccagcc ggggttggtc taggaatacc taccgagctc 360
 cgcccgtttt tgcgtagcgg tctcgtgagc ccgcccggga agctgcgtgc gctaattggac 420
 ttcgtgatcc cgctcggcg aacgaccgaa gacgaatcgc tgggatacat gattgaacgg 480
 cgattgggag ctgaggtgct tgaaaatctt acggagcctc tgcttgcagg gatttatgag 540
 ggtgatatga ggcggttgtc tctccaggca acgttccac agttcgggtga ggtagaacgc 600
 gattacggct cactgatacg gggcatgatg accggtcgca agcctgccga gacacacacc 660
 ggtacaaaaa ggtcagcctt tcttaatttc cggcaagggt tacagtcact tgttcatgca 720
 cttgtacacg aattgcagga cgtcgatcaa agacttaata ccgcagtga gagcctgcag 780
 cgcttgatg gggcccaaac taggtaccgt gtggaattag gcaatggaga gatgctggag 840

gccgatgacg tgggtgtcac cgtcccaacg tacgtagctt ctgagctcct caagccccac 900
 gttgacaccg cagctctgga tgcaatcaat tatgtgagcg tggctaattgt cgtcctggcc 960
 tttgagaaga aggaagtgga gcatgtgttc gacggatcag ggttcttggt tccgagaaaa 1020
 gagggcagga atatcacggc gtgcacttgg acttcgacaa aatggctcca cacctccccg 1080
 gatgacaaaag tacttctgcg atgctatgtg ggccgaagtg gtgatgagca gaatgtagag 1140
 ctccccgacg aggcactgac caacctcgtc ctcaaggacc taaggagagac tatgggcatt 1200
 gagggcgtgc caattttctc tgaaataaca cgcttgcgca agtccatgcc ccaataccct 1260
 gtggggccatc ttcaacacat tgcggccctg cggaagaac ttgggtctaa gctgccgggc 1320
 gtgtacatag cgggcgcccg ttacgagggg gtcgggttgc ctgactgtat tagacaggca 1380
 aaggaaatgt ccgtgcaagc aaccaagaa cttgctgctg actga 1425

<210> 214
 <211> 1413
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 214
 atgagtgacg gtaagaagca tttggtcacc atcggcggcg gcatcacccg cttagcctcc 60
 gccttctaca tggaaaagga gattcgggag aagaaccttc cttgtcagt taccctgggtg 120
 gaggcctcgc cacgggtcgg gggtaaaatc cagacggccc ggaaggatgg ttatattatc 180
 gagcgcggac ccgactcgtt cctcgagcgc aagaagagcg caccgaact cgttgaggac 240
 cttggcctcg aacatctcct cgttaacaat gcaactggtc agtcgtacgt cctggtcaac 300
 gagacactcc atcccatgcc caagggcggc gtgatgggca ttccgacgaa gattgccctt 360
 tttatgtcga ctggcctttt cagcttctcg ggaaaggccc gtgccgctat ggacttcgtc 420
 ctccctgcct cgaaaccgaa ggaggaccag tctcttgag aatttttttag gcgcagagtg 480
 ggggacgagg ttgtggagaa tctgatcgaa ccgcttctga gcggaatcta tgcgggcgac 540
 attgaccgcc tctcactcat gagcaccttc ccacaattct accagacgga gcagaagcat 600
 cggtcactca tcctggggat gaaaaaaacc cggcctcaag gatcaggaca aaggcttaca 660
 gctaaaaagc aggggcagtt tcaaactctc aagacgggcc tgcagactct agtcgaggag 720
 ttagaaaacc agttgaagtt gaccaagggtg tacaagggca cgaaagtgac aaacatcagc 780
 cggggcgaaa agggttgttc aatcgcggtg gacaacggca tgaccctgga cgcagacgca 840
 gcaatcgtga catcgcccca caagagtgtc gcgggcatgt tccctgatct gccggcggtc 900

agccagctta aggatatgca ctcaacctcg gtggctaacg tggccttggg cttccctcag 960
 gaggccgtcc aatggagca cgaaggaacc ggctttgtta tcagccgtaa cagtgacttc 1020
 tcgattaccg cttgtacctg gacgaacaag aagtggcctc acagcgcgcc agaagggaag 1080
 accctcctgc gagcctacgt cggcaaggct ggtgacgagt cgatcgttga gttgtctgac 1140
 aacgagatta tcaagatcgt acttgaagat ctcaagaagg tcatgaagat aaagggtgaa 1200
 cccgagatga cttgcgttac tagatggaac gagtctatgc ctacgtatca cgtggggcac 1260
 aagcagagga tcaagaaggc ccgggaggcc ttggctgcct cgtatccggg agtctacatg 1320
 accggggcct catttgaggg agtcggtatc cccgactgca tcgaccaagg aaagtccgcc 1380
 gtctctgacg tgttggctta tctattcggc tag 1413

<210> 215
 <211> 1413
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 215
 atgagcgacg gaaagaaaca tctcgtgatc atcgggggcg gaataacagg cctagcctcg 60
 gcattctaca tggagaagga gatcagagag aaaaacctcc cgctctctgt gaccctgggtg 120
 gaggcttcac cgagagtggg cgggaagata cagacggcgc gcaaggatgg ctacataata 180
 gagcggggcc cagattcttt cctggagaga aaaaaagcg ccccggaatt ggtggaggac 240
 ctcggcctcg aacacctcct ggtgaataac gcaacagggc aaagctacgt actcgttaat 300
 gagactctcc accccatgcc aaaagggggc gtgatgggaa tccccacaaa gatcgctcca 360
 ttcatgagca ccaggttatt ctctttctct ggtaaagcta gggcagccat ggacttcgtc 420
 ctgccagcct ccaaaccgaa agaagacca agcctcgggg aattcttccg ccggaggggtg 480
 ggcgacgagg tggttgagaa tttaattgaa cctctcctct caggtatata cgcaggggac 540
 atcgaccgct tgtcgtgat gagcaccttt ccgcagttct accagacgga gcagaagcat 600
 cgctcactca ttcttggtat gaagaagact cgtccgcaag ggtctggcca gcagctgaca 660
 gccaaagaaac aggggcagtt ccaaactcct aagaccggcc tacagactct ggtggaggag 720
 ctcgagaacc agctgaagct cacaaagggt tacaagggca caaaggtgac aaacatctca 780
 aggggggaga agggttgctc catcgcgctc gataacggca tgacactcga tgctgatgcg 840
 gcgatagtaa ctagcccgca caagtcggcc gcgggaatgt tccccgacct ccccgcggtc 900

tcgcaactga aggcacatgca ttccaccagc gtcgccaacg tagctctagg ctttcctcag	960
gaggcagtcc aaatggaaca cgagggcacg ggtttcgtaa tctcccgcaa cagcgacttc	1020
tcaatcactg cttgcacgtg gactaacaag aagtggccgc attcggcccc cgagggcaag	1080
acgcttcttc gagcatacgt gggtaaggct ggtgatgaga gtatcgtcga gctctcggac	1140
aacgagatca ttaagatcgt gttggaggac ttgaagaagg tgatgaaaat caagggggag	1200
ccggaaatga cttgcgtgac tcgctggaac gagagcatgc cgcagtacca cgttgggcat	1260
aagcagagga taaagaaagt tcgcgaagcg ctggccgcgt cttaccctgg agtgtatatg	1320
acgggagcct cctttgaggg tgtggggatc ccggactgca tcgaccaggg aaagtcagct	1380
gtctccgacg tgctggccta cttattcgag tga	1413

<210> 216
 <211> 1428
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 216	
aagaagattg cagtcattgg tgggtgggata acagggttgt ccgtggccta ctacgtgagg	60
aagctgcttc gggagcaagg cgттаатгсг ggсгттaccс tcgtcgagca atccgaccgc	120
ctcggcgggga agattagatc cttgagacga gacggcttta ccattgagca aggccctgac	180
tctatgattg cacgтааgсс cgcagctctc gaacttatcc gtgagcttgg tctggaggac	240
aagttggcgg gcacaaaccc tcaagccaaa cgctcctaca tactgcaccg tggcaagttt	300
catccgatgc cacctgggct gatgctcggg attcccactc aaatgtggcc aatggtcaag	360
accgggctgc tatctccggc cggaaagcta cgggctgcga tggacctact tcttcctgca	420
aggcgcgag gcggcgacga atcacttggт gggtttatcc ggaggcggct tggacgtgag	480
gtgttgagc agatgaccga accactcctt gctggaatct atgctggcga cacagaacag	540
ctttcactta aagcgacctt tcctcaattc atggagatgg aaaggaaaca tcgcagtctc	600
atccttgгac tattggctgg gaagaaacag ccaccgcgtc ccggtggtag ccaagtgccg	660
ctcccaaagg ccgtcagac cagtatgttc ttgacactca ccggcgggtt ggaaggctctg	720
accgaagcac tagaggaaag cctatcagag gagaagataa ttactggcca agcagttacc	780
ggactttcgc agcaagaggc cgggtatgag ttaaатctct ctggcggaga gagacttaat	840
gcagacggag tgatcctcgc agtcccagcg ttcgctgccg cccgacttct tgacggcgtg	900
cctgaggccg cctacctaga gcgcатccgc tatgtcagtg ttgctaattt ggcgttcgct	960

tacaggcgtg aggacgtgcc tcatgatctg aatgggtccg gcgtgttaat ccctagaggt	1020
gaagggagga tgattacggc cataacttgg gtttcgtcca aatggttgca ttcagcaccc	1080
ggtgacaagg cactgctgag agcgtacatt gggcgactag gtgatgaggc ttggacagcc	1140
atgtgtaggg ccgacatcga gcgtagagtc gccgctgaac tccgcatct actaggaatt	1200
gccgctagtc ctttgttctg tgaactagcc gcactcccag aatctatgcc gcagtatcca	1260
gtgggtcacg tcgaacgact cgaagccttg cgaggagcat tgtgtcgcgc taaaccaggg	1320
ttgttgttgt gtggtgccgg gtacgctggc gttggcattc cagactgcat tcggcaaggc	1380
aaagaagccg ctgagtcgat ggcggcttat ttgagggacg gacgctag	1428

<210> 217
 <211> 1470
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 217	
aggaaacttg tggatcatcg cggagggatc actgggcttt cggccgcctt ctatgcacta	60
aagcaagccg atgaggaagg gcagcccatc tcggtcacca taattgaaca gagcgatagg	120
ctcggcggaag agatccagac actccgcaag gagggctgcg taattgagaa gggcccggat	180
tccttcctcg ctaggaagtt gccgatgatt gatctagctc gggatcttgg catggactcc	240
gaattggtgg cgactaatcc gcacgcaaag aagacttaca tcttgaggcg cggaaagctc	300
taccggatgc ctccaggctt agtgcttggc atacctacgg aactaggacc attcgctaag	360
acagggctca ttagcccttg gggcaaactc cgcgccgcta tggatttgtt cattaagcct	420
catccagccg atgaagacga aagtgttggc gctttcctgg acagacgtct cggtagggaa	480
gtgaccgagc acattgcgga acctttattg gcgggcatct acgcgggcga cttgcaagcc	540
ttaagccttc aagccacttt cccacagttt gcacaagtag agcgcaagca cggagggctg	600
atacgcggtg tgaaggccag cagacaggcc ggtcagtcgg tgcttgggct gccggacgtc	660
gccaaagggtg cgatgttctt tacctttcgc aacgggctta ccagcttagt tgaaagggtg	720
gaggaaactc tcagagacag ggctgaaactc tgtctgggca tcggcgcaga agggtttgag	780
aaacgtgaag atggaacata ccttggtcga ctaagcgatg gttcgaggct ccaggccgac	840
gcagtaattg tcactacgcc gagctatcat gcggcatccc tggttgagga gcatgtggat	900
gcttcggccc tccaggccat tcgtcatgta agcgttgcaa atgtcgttag cgtcttcgac	960

cgaaagcaag tgaataacca gttcgacggc acagggtttg ttatctcacg gcgagaaggt 1020
 cgcgcaatca ccgcctgtac ctggacatcc gtgaaatggc cgcatacttc gcgcggcgac 1080
 aaactgatta tccggtgcta catcggtagg gctggcgacg aggagcgagt ggattggccc 1140
 gatgaagctc tcaagcgtac tgtaagatca gaactgcgtg agttgctgga cattgacatt 1200
 gatccggaat ttgtggagat tacacgactc aggcactcta tgcctcaata ccagtcggc 1260
 cacgtccagg ctatccgctc tttgagggac gaggtcggta ggactttacc gggcgtgttc 1320
 cttgctgggc aaccctacga aggtgtggga atgcctgact gtgtgaggtc cggccgggat 1380
 gccgccgaag cagcagtaag tgctatgcaa gcaatgagta cagaaccaga agcaccggca 1440
 gaggacgccg ctactggaac ggcgggttga 1470

<210> 218
 <211> 1410
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 218
 cggagagttg ttgttggttg cggcggcttg actggcctaa gcgcgcctt ctacatccgg 60
 aaacattatc gagaagcttg agttgagccc gtcacacgc ttgttgagaa atctagctcg 120
 atgggagggga tgattgagac cttcatagg gacgggtttg tcatcgagaa gggcccggac 180
 agtttcttgg cacggaagac cgcaatgatt gatctggcga aagagctgga gattgaccac 240
 gagttggtca gccagaatcc agaatcgaag aagacctaca taatgcaacg tggaaagctg 300
 caccctatgc cagcgggact tgttctgggc attcccaccg aattgcgtcc ctttctccgg 360
 agcgggcttg tctcaccgc tgggaagttg cgggcgctga tggacttcgt aataccgcc 420
 cgaaggacga ccgaagatga gtcactcggg tacatgatcg agcgcgact gggtgccgag 480
 gtgttgagaga acctcacaga gccgttgctc gctggaatct acgctggcga catgagaaga 540
 ttgtccctcc aggtacggtt tccgcagttc ggtgaggttg agcgcgacta cggctcctta 600
 atcagaggaa tgatgaccgg acgtaagcct gcggagacac acacagggac caagaggtct 660
 gcctttctca atttcagaca gggctctgcaa tctactggtc acgccttagt ccatgaactc 720
 caggatgtag atcagaggtt aaatactgcg gtgaagtcgc ttcagaggct tgacggcgca 780
 caaaccggtt atcgcgttga actcggcaat ggcgaaatgc ttgaggctga cgacgtggtg 840
 gttactgtac caacctacgt ggcgagcgag cttcttaagc cgcacgtgga cacggcggcg 900
 ttagacgcta ttaactatgt gtcggtggct aatgtagttc ttgcattcga gaagaaggaa 960

gtagagcacg tcttcgatgg atcgggcttc ttggtgcctc ggaaggaggg aaggaacata 1020
 accgcctgca cctggacttc gaccaagtgg ctccacacat caccagatga caaggttctg 1080
 ttacgttggt acgtgggcag aagtggagat gacgagaatg tggaactccc ggatgaggca 1140
 ctactaatc tgggtgcttaa ggatctgaga gagacgatgg gcatcgaggc ggttccaatc 1200
 ttctcagaga ttacccggct ccgcaagtca atgccgcagt acccagtagg acatctccag 1260
 cacatcgccg cattgcgcga ggaactcggc tctaagctac caggagtgtg catcgccgga 1320
 gcgggctacg agggcgttgg tcttccggat tgcattcgcc aggccaaaga aatgtcagtc 1380
 caggcaacgc aagaactcgc tgccgactga 1410

<210> 219

<211> 1398

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tái tổ hợp

<400> 219

aagcacctgg taatcatcgg tgggtgggac accggtctgg cttcagcggt ctacatggaa 60
 aaggagatcc gggagaagaa cttgcccctt tcggtgactc tagtggaggc ctctccacgg 120
 gtggggggca agattcagac cgcgcgcaag gatggctaca tcatagagcg aggaccagac 180
 tcattcctag agcgtaagaa gtccgcccc gagctcgctg aggatctcgg tctagagcac 240
 ttgctagtga ataacgctac aggacagtcc tacgtgctcg tgaacgagac actacacccg 300
 atgcctaagg gggctgtcat ggggtataccg accaagatcg ccccgttcat gtccactcgc 360
 cttttctcgt tctcgggcaa agctcggggc gctatggatt tcgtcttgcc tgacctgaaa 420
 ccgaaggagg accagtcctt aggagagttc ttccgccgga gggtcggcga cgaggtggtg 480
 gagaacttaa tcgaaccctt gctctcgggg atctacgtg gagacattga tcgactatcg 540
 cttatgtcta cgtttcctca attttaccag acggagcaga agcaccgtag cctcattttg 600
 ggtatgaaga agacacggcc tcaagggttcg gggcagcagc ttactgcaa gaagcagggc 660
 caattccaga cactcaagac cggcttgacg actctagtgg aggagctgga gaatcaattg 720
 aagctgacaa aggtctacaa ggggtaccaag gtgacaaaca tatcgctgg cgaaaaggga 780
 tgctccattg ccctcgacaa cggtatgacc ctcgacgcc acgcagcgat tgtgacgagc 840
 ccacacaaga gcgccgagg catgttccc gacttgctg cagtgtcaca gctgaaagac 900
 atgcattcta catccgtcgc caacgtcgcc ctgggctttc cccaggaggc tgtgcagatg 960

gagcacgagg ggacgggctt cgttatcagc cgcaactccg acttttctat taccgcgtgc 1020
 acatggacca acaagaagtg gccgcacagc gtcgaggagg ggaaaacact tctccgagca 1080
 tacgtaggca aggccgggga cgagtcaatt gttgagctct ccgacaatga aatcattaaa 1140
 atagttcttg aggatcttaa gaaggtaatg aagataaagg gggaacctga aatgacgtgt 1200
 gttacccgct ggaatgagtc aatgccccag taccatgttg gacacaagca gaggataaag 1260
 aaggtgaggg aggcgctcgc tgcgtcctac ccaggggtct acatgacagg agcgagtttt 1320
 gagggggttg gtattcccga ctgtatcgac cagggttaagt cggcagtgtc tgacgtgctc 1380
 gcttacctat tcgagtag 1398

<210> 220
 <211> 1422
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 220
 attgcagtca ttggtggttg gataacaggg ttgtccgttg cctactacgt gaggaagctg 60
 cttcgggagc aaggcgtaa tgcgggctt accctcgtcg agcaatccga ccgcctcggc 120
 gggaagatta gatccttgag acgagacggc ttaccattg agcaaggccc tgactctatg 180
 attgcacgta agcccgagc tctcgaactt atccgtgagc ttggtctgga ggacaagttg 240
 gcgggcacaa accctcaagc caaacgctcc tacatactgc accgtggcaa gtttcatccg 300
 atgccacctg ggctgatgct cgggattccc actcaaagt ggccaatgg caagaccggg 360
 ctgctatctc cgccggaaa gctacgggct gcgatggacc tacttcttcc tgcaaggcgc 420
 ggaggcggcg acgaatcact tgggtgggtt atccggaggc ggcttgagc tgaggtgttg 480
 gagcagatga ccgaaccact ccttgctgga atctatgctg gcgacacaga acagctttca 540
 cttaaagcga ctttcttca attcatggag atggaaagga aacatcgag tctcatcctt 600
 ggactatttg ctgggaagaa acagccaccg cgtcccgtg gtagccaagt gccgctccca 660
 aaggccgctc agaccagtat gttcttgaca ctaccggcg ggttggaagg tctgaccgaa 720
 gcactagagg aaagcctatc agaggagaag ataattactg gccaagcagt taccggactt 780
 tcgcagcaag aggccgggta tgagttaa atctcttggcg gagagagact taatgcagac 840
 ggagtgatcc tcgcagtccc agcgttcgt gccgcccagc ttcttgacgg cgtgcctgag 900
 gccgcctacc tagagcgcat ccgctatgtc agtggtgcta atttggcggt cgcttacagg 960
 cgtgaggacg tgcctcatga tctgaatggg tccggcgtgt taatccctag aggtgaaggg 1020

aggatgatta cggccataac ttgggtttcg tccaaatggt tgcattcagc acccggtgac	1080
aaggcactgc tgagagcgta cattggggcga ctaggtgatg aggcttggac agccatgtgt	1140
agggccgaca tcgagcgtag agtcgccgct gaactccgcg atctactagg aattgccgct	1200
agtcctttgt tctgtgaact agccgcactc ccagaatcta tgccgcagta tccagtgggt	1260
cacgtcgaac gactcgaagc cttgcgagga gcattgtgtc gcgctaaacc agggttgttg	1320
ttgtgtggtg ccgggtacgc tggcgttggc attccagact gcattcggca aggcaaagaa	1380
gccgctgagt cgatggcggc ttatttgagg gacggacgct ag	1422

<210> 221
 <211> 1464
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 221	
cttgtggtca tcggcggagg gatcactggg ctttcggccg ctttctatgc actaaagcaa	60
gccgatgagg aagggcagcc catctcggtc accataattg aacagagcga taggctcggc	120
ggaaagatcc agacactccg caaggagggc tgcgtaattg agaagggccc ggattccttc	180
ctcgctagga agttgccgat gattgatcta gtcgggatc ttggcatgga ctccgaattg	240
gtggcgacta atccgcacgc aaagaagact tacatcttga ggcgcgaaa gctctaccgg	300
atgcctccag gcttagtgct tggcatacct acggaactag gaccattcgc taagacaggg	360
ctcattagcc cttggggcaa actccgcgcc gctatggatt tgttcattaa gcctcatcca	420
gccgatgaag acgaaagtgt tggcgctttc ctggacagac gtctcggtag ggaagtgacc	480
gagcacattg cggaaccttt attggcgggc atctacgcgg gcgacttgca agccttaagc	540
cttcaagcca ctttcccaca gtttgacaaa gtagagcgca agcacggagg gctgatacgc	600
ggtatgaagg ccagcagaca ggccggtcag tccgtgcctg ggctgccgga cgtcgccaag	660
ggtacgatgt tccttacctt tcgcaacggg cttaccagct tagttgaaag gttggaggaa	720
actctcagag acagggctga actctgtctg ggcacggcg cagaaggggt tgagaaacgt	780
gaagatggaa cataccttgt tcgactaagc gatggttcga ggctccaggc cgacgcagta	840
attgtcacta cgccgagcta tcatgcggca tccctgttgg aggagcatgt ggatgcttcg	900
gccctccagg ccattcgtca tgtaagcggt gcaaagtgcg ttagcgtctt cgaccgaaag	960
caagtgaata accagttcga cggcacaggg tttgttatct cacggcgaga aggtcgcgca	1020

atcaccgcct gtacctggac atccgtgaaa tggccgcata cttcgcgcg cgacaaactg	1080
attatccggt gctacatcgg tagggctggc gacgaggagc gagtggattg gcccgatgaa	1140
gctctcaagc gtactgtaag atcagaactg cgtgagttgc tggacattga cattgatccg	1200
gaatttgtgg agattacacg actcaggcac tctatgcctc aatacccagt cggccacgtc	1260
caggctatcc gctctttgag ggacgaggtc ggtaggactt taccgggctt gttccttgct	1320
gggcaaccct acgaaggtgt gggaatgcct gactgtgtga ggtccggccg ggatgccgcc	1380
gaagcagcag taagtgctat gcaagcaatg agtacagaac cagaagcacc ggcagaggac	1440
gccgctactg gaacggcggg ttga	1464

<210> 222
 <211> 1404
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 222	
gttgttgttg ttggcggcgg cttgactggc ctaagcgccg cttctacat ccggaacat	60
tatcgagaag ctggagttga gcccgtcatc acgcttggtg agaaatctag ctcgatggga	120
gggatgattg agacccttca tagggacggg tttgtcatcg agaagggccc ggacagtttc	180
ttggcacgga agaccgcaat gattgatctg gcgaaagagc tggagattga ccacgagttg	240
gtcagccaga atccagaatc gaagaagacc tacataatgc aacgtggaaa gctgcaccct	300
atgccagcgg gacttgttct gggcattccc accgaattgc gtccctttct ccggagcggg	360
cttgtctcac ccgctgggaa gttgcgggcg ctgatggact tcgtaatacc gccacgaagg	420
acgaccgaag atgagtcact cgggtacatg atcgagcgcc gactgggtgc cgaggtgttg	480
gagaacctca cagagccgtt gctcgctgga atctacgctg gcgacatgag aagattgtcc	540
ctccaggcta cgtttccgca gttcggtgag gtggagcgcg actacggctc cttaatcaga	600
ggaatgatga ccggacgtaa gcctgcggag acacacacag ggaccaagag gtctgccttt	660
ctcaatttca gacaggttct gcaatcactg gttcacgcct tagtccatga actccaggat	720
gtagatcaga ggttaaatac tgcggtgaag tcgcttcaga ggcttgacgg cgcacaaacc	780
cgttatcgcg ttgaactcgg caatggcgaa atgcttgagg ctgacgacgt ggtggttact	840
gtaccaacct acgtggcgag cgagcttctt aagccgcacg tggacacggc ggcgttagac	900
gctattaact atgtgtcggg ggctaatagt gttcttgcac tcgagaagaa ggaagtagag	960
cacgtcttcg atggatcggg cttcttggtg cctcggaagg agggaaggaa cataaccgcc	1020

tgcacctgga cttcgaccaa gtggctccac acatcaccag atgacaaggt tctgttacgt	1080
tgttacgtgg gcagaagtgg agatgagcag aatgtggaac tcccggatga ggcaactcact	1140
aatctggtgc ttaaggatct gagagagacg atgggcatcg aggcggttcc aatctttctca	1200
gagattaccc ggctccgcaa gtcaatgccg cagtaccag taggacatct ccagcacatc	1260
gccgcattgc gcgaggaact cggctctaag ctaccaggag tgtacatcgc cggagcgggc	1320
tacgagggcg ttggtcttcc ggattgcatt cgccaggcca aagaaatgtc agtccaggca	1380
acgcaagaac tcgctgccga ctga	1404

<210> 223
 <211> 1398
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 223	
aagcacctgg taatcatcgg tgggtggatc accggtctgg cttcagcgtt ctacatggaa	60
aaggagatcc gggagaagaa cttgcccctt tcggtgactc tagtgagggc ctctccacgg	120
gtggggggca agattcagac cgcgcgcaag gatggctaca tcatagagcg aggaccagac	180
tcattcctag agcgtaagaa gtccgcccca gagctcgtcg aggatctcgg tctagagcac	240
ttgctagtga ataacgctac aggacagtcc tacgtgctcg tgaacgagac actacacccg	300
atgcctaagg gggctgtcat gggatataccg accaagatcg ccccgttcat gtccactcgc	360
cttttctcgt tctcgggcaa agctcgggcc gctatggatt tcgtcttgcc tgctcgaaa	420
ccgaaggagg accagtcctt aggagagttc ttccgccgga gggtcggcga cgaggtgggtg	480
gagaacttaa tcgaaccctt gctctcgggg atctacgctg gagacattga tcgactatcg	540
cttatgtcta cgtttcctca attttaccag acggagcaga agcaccgtag cctcattttg	600
ggtatgaaga agacacggcc tcaaggttcg gggcagcagc ttactgcca gaagcagggc	660
caattccaga cactcaagac cggcttgacg actctagtgg aggagctgga gaatcaattg	720
aagctgacaa aggtctacaa gggataccaag gtgacaaaca tatcgcgtag cgaaaaggga	780
tgctccattg ccctcgacaa cggatatgacc ctcgacgccc acgcagcgat tgtgacgagc	840
ccacacaaga ggcgcgggg catgttcccc gacttgctcg cagtgtcaca gctgaaagac	900
atgcattcta catccgtcgc caacgtcgcc ctgggctttc cccaggaggc tgtgcagatg	960
gagcacgagg ggacgggctt cgttatcagc cgcaactccg acttttctat taccgcgtgc	1020

acatggacca acaagaagtg gccgcacagc gtcctggagg ggaaaacact tctccgagca 1080
 tacgtaggca aggccgggga cgagtcaatt gttgagctct ccgacaatga aatcattaaa 1140
 atagttcttg aggatcttaa gaaggtaatg aagataaagg gggaacctga aatgacgtgt 1200
 gttacccgct ggaatgagtc aatgccccag taccatgtgg gacacaagca gaggataaag 1260
 aaggtgaggg aggcgctcgc tgcgtcctac ccaggggtct acatgacagg agcgagtttt 1320
 gagggggtgg gtattcccga ctgtatcgac cagggttaagt cggcagtgtc tgacgtgctc 1380
 gcttacctat tcgagtag 1398

<210> 224
 <211> 445
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 224

Gln Pro Val Leu Ile Val Gly Ala Gly Leu Ser Gly Leu Ser Ile Ala
 1 5 10 15

Tyr Glu Leu Gln Lys Leu Gln Val Pro Tyr Gln Val Leu Glu Val Ser
 20 25 30

Gly His Ser Gly Gly Val Met Lys Ser Leu Arg Lys Asp Gly Phe Glu
 35 40 45

Leu Asp Ala Gly Ala Asn Thr Ile Ala Ala Ser Pro Glu Ile Leu Ala
 50 55 60

Tyr Phe Thr Ser Leu Gly Leu Glu Asn Glu Ile Leu Gln Ala Thr Ala
 65 70 75 80

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

Ser Pro His Pro Phe Lys Ile Met Ser Ser Pro Tyr Leu Ser Arg Gly
 100 105 110

Ser Lys Trp Arg Leu Phe Thr Glu Arg Phe Arg Lys Pro Val Val Ala
 115 120 125

Ser Gly Glu Glu Thr Val Thr Asp Phe Ile Thr Arg Arg Phe Asn Arg
 130 135 140

Glu Ile Ala Glu Tyr Val Phe Asp Pro Val Leu Ser Gly Ile Tyr Ala
145 150 155 160

Gly Asn Pro Asp Gln Met Ser Ile Ala Glu Val Leu Pro Ala Leu Pro
165 170 175

Arg Trp Glu Arg Glu Tyr Gly Ser Val Thr Lys Gly Leu Met Lys Asp
180 185 190

Lys Gly Ala Met Gly Gly Arg Lys Ile Ile Ser Phe Lys Gly Gly Asn
195 200 205

Gln Leu Leu Thr Asn Arg Leu Gln Gln Leu Leu Thr Thr Pro Val Arg
210 215 220

Phe Asn Cys Lys Val Thr Gly Ile Thr Ala Ser Asn Gly Gly Tyr Ile
225 230 235 240

Val Ser Ala Val Glu Asp Gly Val Ser Glu Ser Tyr Thr Ala Ser Arg
245 250 255

Val Ile Leu Thr Thr Pro Ala Tyr Ser Ala Ala Ala Thr Ile Thr Asn
260 265 270

Leu Asp Ala Ala Thr Ala Ala Leu Leu Asn Glu Ile His Tyr Pro Arg
275 280 285

Met Gly Val Leu His Leu Gly Phe Asp Ala Thr Ala Leu Pro Gln Pro
290 295 300

Leu Asp Gly Phe Gly Phe Leu Val Pro Asn Ala Glu Asn Met His Phe
305 310 315 320

Leu Gly Ala Ile Cys Asn Ala Ala Ile Phe Pro Asp Lys Ala Pro Pro
325 330 335

Gly Lys Ile Leu Phe Thr Val Phe Leu Gly Gly Ala Arg Gln Glu Ser
340 345 350

Leu Phe Asp Gln Met Thr Pro Glu Ala Leu Gln Gln Gln Val Val Ser
355 360 365

Glu Val Met Ser Leu Leu His Leu Ser Ala Pro Pro Val Met Gln His

370 375 380
 Phe Ser Ser Trp Asn Lys Ala Ile Pro Gln Leu Asn Val Gly His Val
 385 390 395 400
 Lys Leu Arg Arg Ala Val Glu Ala Phe Glu Lys Lys Tyr Pro Gly Ile
 405 410 415
 His Leu Ser Gly Asn Tyr Leu Gln Gly Val Ala Ile Pro Ala Leu Leu
 420 425 430
 Gln His Ala Ala Ala Leu Ala Ala Ser Leu Lys Lys Asn
 435 440 445
 <210> 225
 <211> 448
 <212> PRT
 <213> Trình tự nhân tạo
 <220>
 <223> Tái tổ hợp
 <400> 225
 Met Ser Asp Gln Pro Val Leu Ile Val Gly Ala Gly Leu Ser Gly Leu
 1 5 10 15
 Ser Ile Ala Tyr Glu Leu Gln Lys Leu Gln Val Pro Tyr Gln Val Leu
 20 25 30
 Glu Val Ser Gly His Ser Gly Gly Val Met Lys Ser Leu Arg Lys Asp
 35 40 45
 Gly Phe Glu Leu Asp Ala Gly Ala Asn Thr Ile Ala Thr Ser Pro Glu
 50 55 60
 Ile Leu Ala Tyr Phe Thr Ser Leu Gly Leu Glu Asn Glu Ile Leu Gln
 65 70 75 80
 Ala Thr Ala Thr Ser Lys His Arg Phe Leu Val Arg Arg Arg Gln Leu
 85 90 95
 His Ala Val Ser Pro His Pro Phe Lys Ile Met Ser Ser Pro Tyr Leu
 100 105 110
 Cys Arg Gly Ser Lys Trp Arg Leu Phe Thr Glu Arg Phe Arg Lys Pro
 115 120 125

Val Val Ala Ser Gly Glu Glu Thr Val Thr Asp Phe Ile Thr Arg Arg
 130 135 140

Phe Asn Arg Glu Ile Ala Glu Tyr Val Phe Asp Pro Val Leu Ser Gly
 145 150 155 160

Ile Tyr Ala Gly Asn Pro Asp Gln Met Ser Ile Ala Glu Val Leu Pro
 165 170 175

Ala Leu Pro Arg Trp Glu Arg Glu Tyr Gly Ser Val Thr Lys Gly Leu
 180 185 190

Met Lys Asp Lys Gly Ala Met Gly Gly Arg Lys Ile Ile Ser Phe Lys
 195 200 205

Gly Gly Asn Gln Leu Leu Thr Asn Arg Leu Gln Gln Leu Leu Thr Thr
 210 215 220

Pro Val Arg Phe Asn Cys Lys Val Thr Gly Ile Thr Ala Ser Asn Gly
 225 230 235 240

Gly Tyr Ile Val Ser Ala Val Glu Asp Gly Val Ser Glu Ser Tyr Thr
 245 250 255

Ala Ser Arg Val Ile Leu Thr Thr Pro Ala Tyr Ser Ala Ala Ala Thr
 260 265 270

Ile Thr Asn Leu Asp Ala Ala Thr Ala Ala Leu Leu Asn Glu Ile His
 275 280 285

Tyr Pro Arg Met Gly Val Leu His Leu Gly Phe Asp Ala Thr Ala Leu
 290 295 300

Pro Gln Pro Leu Asp Gly Phe Gly Phe Leu Val Pro Asn Ala Glu Asn
 305 310 315 320

Met His Phe Leu Gly Ala Ile Cys Asn Ala Ala Ile Phe Pro Asp Lys
 325 330 335

Ala Pro Pro Gly Lys Ile Leu Phe Thr Val Phe Leu Gly Gly Ala Arg
 340 345 350

Gln Glu Ser Leu Phe Asp Gln Met Thr Pro Glu Ala Leu Gln Gln Gln

355 360 365
 Val Val Ser Glu Val Met Ser Leu Leu His Leu Ser Ala Pro Pro Val
 370 375 380
 Met Gln His Phe Ser Ser Trp Asn Lys Ala Ile Pro Gln Leu Asn Val
 385 390 395 400
 Gly His Val Lys Leu Arg Arg Ala Val Glu Ala Phe Glu Lys Lys Tyr
 405 410 415
 Pro Gly Ile His Leu Ser Gly Asn Tyr Leu Gln Gly Val Ala Ile Pro
 420 425 430
 Ala Leu Leu Gln His Ala Ala Ala Leu Ala Ala Ser Leu Lys Lys Asn
 435 440 445
 <210> 226
 <211> 466
 <212> PRT
 <213> Trình tự nhân tạo
 <220>
 <223> Tái tổ hợp
 <400> 226
 Lys Lys His Val Val Ile Ile Gly Gly Gly Ile Thr Gly Leu Ala Ala
 1 5 10 15
 Ala Phe Tyr Met Glu Lys Glu Ile Lys Glu Lys Asn Leu Pro Leu Glu
 20 25 30
 Leu Thr Leu Val Glu Ala Ser Pro Arg Val Gly Gly Lys Ile Gln Thr
 35 40 45
 Val Lys Lys Asp Gly Tyr Ile Ile Glu Arg Gly Pro Asp Ser Phe Leu
 50 55 60
 Glu Arg Lys Lys Ser Ala Pro Gln Leu Val Lys Asp Leu Gly Leu Glu
 65 70 75 80
 His Leu Leu Val Asn Asn Ala Thr Gly Gln Ser Tyr Val Leu Val Asn
 85 90 95
 Arg Thr Leu His Pro Met Pro Lys Gly Ala Val Met Gly Ile Pro Thr
 100 105 110

Lys Ile Ala Pro Phe Val Ser Thr Gly Leu Phe Ser Leu Ser Gly Lys
 115 120 125

Ala Arg Ala Ala Met Asp Phe Ile Leu Pro Ala Ser Lys Thr Lys Asp
 130 135 140

Asp Gln Ser Leu Gly Glu Phe Phe Arg Arg Arg Val Gly Asp Glu Val
 145 150 155 160

Val Glu Asn Leu Ile Glu Pro Leu Leu Ser Gly Ile Tyr Ala Gly Asp
 165 170 175

Ile Asp Lys Leu Ser Leu Met Ser Thr Phe Pro Gln Phe Tyr Gln Thr
 180 185 190

Glu Gln Lys His Arg Ser Leu Ile Leu Gly Met Lys Lys Thr Arg Pro
 195 200 205

Gln Gly Ser Gly Gln Gln Leu Thr Ala Lys Lys Gln Gly Gln Phe Gln
 210 215 220

Thr Leu Ser Thr Gly Leu Gln Thr Leu Val Glu Glu Ile Glu Lys Gln
 225 230 235 240

Leu Lys Leu Thr Lys Val Tyr Lys Gly Thr Lys Val Thr Lys Leu Ser
 245 250 255

His Ser Gly Ser Gly Tyr Ser Leu Glu Leu Asp Asn Gly Val Thr Leu
 260 265 270

Asp Ala Asp Ser Val Ile Val Thr Ala Pro His Lys Ala Ala Ala Gly
 275 280 285

Met Leu Ser Glu Leu Pro Ala Ile Ser His Leu Lys Asn Met His Ser
 290 295 300

Thr Ser Val Ala Asn Val Ala Leu Gly Phe Pro Glu Gly Ser Val Gln
 305 310 315 320

Met Glu His Glu Gly Thr Gly Phe Val Ile Ser Arg Asn Ser Asp Phe
 325 330 335

Ala Ile Thr Ala Cys Thr Trp Thr Asn Lys Lys Trp Pro His Ala Ala

340 345 350
 Pro Glu Gly Lys Thr Leu Leu Arg Ala Tyr Val Gly Lys Ala Gly Asp
 355 360 365
 Glu Ser Ile Val Asp Leu Ser Asp Asn Asp Ile Ile Asn Ile Val Leu
 370 375 380
 Glu Asp Leu Lys Lys Val Met Asn Ile Asn Gly Glu Pro Glu Met Thr
 385 390 395 400
 Cys Val Thr Arg Trp His Glu Ser Met Pro Gln Tyr His Val Gly His
 405 410 415
 Lys Gln Arg Ile Lys Glu Leu Arg Glu Ala Leu Ala Ser Ala Tyr Pro
 420 425 430
 Gly Val Tyr Met Thr Gly Ala Ser Phe Glu Gly Val Gly Ile Pro Asp
 435 440 445
 Cys Ile Asp Gln Gly Lys Ala Ala Val Ser Asp Ala Leu Thr Tyr Leu
 450 455 460
 Phe Ser
 465
 <210> 227
 <211> 465
 <212> PRT
 <213> Trình tự nhân tạo
 <220>
 <223> Tái tổ hợp
 <400> 227
 Lys His Leu Val Ile Ile Gly Gly Gly Ile Thr Gly Leu Ala Ala Ala
 1 5 10 15
 Phe Tyr Leu Glu Lys Glu Val Glu Glu Lys Gly Leu Pro Ile Gln Ile
 20 25 30
 Ser Leu Ile Glu Ala Ser Pro Arg Leu Gly Gly Lys Ile Gln Thr Leu
 35 40 45
 Tyr Lys Asp Gly Tyr Ile Ile Glu Arg Gly Pro Asp Ser Phe Leu Glu
 50 55 60

Arg Lys Val Ser Gly Pro Gln Leu Ala Lys Asp Val Gly Leu Ser Asp
65 70 75 80

Gln Leu Val Asn Asn Glu Thr Gly Gln Ala Tyr Val Leu Val Asn Glu
85 90 95

Lys Leu His Pro Met Pro Lys Gly Ala Val Met Gly Ile Pro Thr Gln
100 105 110

Ile Ser Pro Phe Ile Thr Thr Gly Leu Phe Ser Val Ala Gly Lys Ala
115 120 125

Arg Ala Ala Met Asp Phe Val Leu Pro Lys Ser Lys Gln Thr Glu Asp
130 135 140

Gln Ser Leu Gly Glu Phe Phe Arg Arg Arg Val Gly Asp Glu Val Val
145 150 155 160

Glu Asn Leu Ile Glu Pro Leu Leu Ser Gly Ile Tyr Ala Gly Asp Ile
165 170 175

Asp Arg Leu Ser Leu Met Ser Thr Phe Pro Gln Phe Tyr Gln Thr Glu
180 185 190

Gln Gln His Arg Ser Leu Ile Leu Gly Met Lys Lys Ser Gln Gln His
195 200 205

Ala Lys Ala Gln Gln Val Thr Ala Lys Lys Gln Gly Gln Phe Gln Thr
210 215 220

Ile Asn Gln Gly Leu Gln Ser Leu Val Glu Ala Val Glu Gly Lys Leu
225 230 235 240

Lys Leu Thr Thr Val Tyr Lys Gly Thr Lys Val Lys Gln Ile Glu Lys
245 250 255

Thr Asp Gly Gly Tyr Gly Leu Gln Leu Asp Ser Gly Gln Thr Leu Phe
260 265 270

Ala Asp Ser Ala Ile Val Thr Thr Pro His Gln Ser Ile Tyr Ser Met
275 280 285

Phe Pro Lys Glu Ala Gly Leu Glu Tyr Leu His Asp Met Thr Ser Thr

290 295 300
 Ser Val Ala Thr Val Ala Leu Gly Phe Lys Asp Glu Asp Val His Asn
 305 310 315 320
 Glu Tyr Asp Gly Thr Gly Phe Val Ile Ser Arg Asn Ser Asp Phe Ser
 325 330 335
 Ile Thr Ala Cys Thr Trp Thr Asn Lys Lys Trp Pro His Thr Ala Pro
 340 345 350
 Lys Gly Lys Thr Leu Leu Arg Ala Tyr Val Gly Lys Ala Gly Asp Glu
 355 360 365
 Ser Ile Val Glu Gln Ser Asp Ser Gln Ile Val Ser Ile Val Leu Glu
 370 375 380
 Asp Leu Lys Lys Ile Met Asp Ile Lys Ala Asp Pro Glu Leu Thr Thr
 385 390 395 400
 Val Thr Arg Trp Lys Thr Ser Met Pro Gln Tyr His Val Gly His Gln
 405 410 415
 Lys Ala Ile Ser Asn Met Arg Glu Thr Phe Lys Gln Ser Tyr Pro Gly
 420 425 430
 Val Tyr Ile Thr Gly Ala Ala Phe Glu Gly Val Gly Ile Pro Asp Cys
 435 440 445
 Ile Asp Gln Gly Lys Ala Ala Ile Ser Glu Ala Val Ser Tyr Leu Phe
 450 455 460
 Ser
 465
 <210> 228
 <211> 465
 <212> PRT
 <213> Trình tự nhân tạo
 <220>
 <223> Tái tổ hợp
 <400> 228
 Lys His Leu Val Ile Ile Gly Gly Gly Ile Thr Gly Leu Ala Ala Ala
 1 5 10 15

Phe Tyr Leu Glu Lys Glu Val Glu Glu Lys Gly Leu Pro Ile Gln Ile
 20 25 30

Ser Leu Ile Glu Ala Ser Pro Arg Leu Gly Gly Lys Ile Gln Thr Leu
 35 40 45

Tyr Lys Asp Gly Tyr Ile Ile Glu Arg Gly Pro Asp Ser Phe Leu Glu
 50 55 60

Arg Lys Val Ser Gly Pro Gln Leu Ala Lys Asp Val Gly Leu Ser Asp
 65 70 75 80

Gln Leu Val Asn Asn Glu Thr Gly Gln Ala Tyr Val Leu Val Asn Glu
 85 90 95

Thr Leu His Pro Met Pro Lys Gly Ala Val Met Gly Ile Pro Thr Gln
 100 105 110

Ile Ser Pro Phe Ile Thr Thr Gly Leu Phe Ser Val Ala Gly Lys Ala
 115 120 125

Arg Ala Ala Met Asp Phe Val Leu Pro Lys Ser Lys Gln Thr Glu Asp
 130 135 140

Gln Ser Leu Gly Glu Phe Phe Arg Arg Arg Val Gly Asp Glu Val Val
 145 150 155 160

Glu Asn Leu Ile Glu Pro Leu Leu Ser Gly Ile Tyr Ala Gly Asp Ile
 165 170 175

Asp Arg Leu Ser Leu Met Ser Thr Phe Pro Gln Phe Tyr Gln Thr Glu
 180 185 190

Gln Lys His Arg Ser Leu Ile Leu Gly Met Lys Lys Ser Gln Gln His
 195 200 205

Ala Lys Ala Gln Gln Val Thr Ala Lys Lys Gln Gly Gln Phe Gln Thr
 210 215 220

Ile Asn Gln Gly Leu Gln Ala Leu Val Glu Ala Val Glu Ser Lys Leu
 225 230 235 240

Lys Leu Thr Thr Ile Tyr Lys Gly Thr Lys Val Lys Gln Ile Glu Lys

[illegible]

<210> 229
 <211> 537
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 229
 aaggccttgg tactgtactc gacgcgggac ggccagaccc acgcaattgc ttcatacatc 60
 gcctcctgca tgaaggagaa ggccgaatgc gacgtgatcg acctcaccca cggggagcac 120
 gtgaacctca cccaatacga tcaggtgcta atcgggtgcga gtattcgta cggccacttc 180
 aacgccgtgc ttgacaagtt catcaagaga aacgtggatc aactgaacaa catgccaaagc 240
 gcgttcttct gcgtaaacct cacagcaagg aagcccgaga agcgtactcc ccagacaaac 300
 ccttatgtcc gaaaattctt gcttgctacc ccctggcagc ccgcgttggtg cggagtgttc 360
 gcagggggccc ttcggtaccc gcgataccgg tggatcgaca aggtgatgat ccagctaata 420
 atgcggatga ctgggggaga gacagacacg agcaaggagg tcgagtacac ggattgggag 480
 caggttaaga agttcgcgga ggattttgca aagctatcgt acaagaaggc cctctag 537

<210> 230
 <211> 1338
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 230
 cagcccgtcc tcatcgttgg agctggtctc tccgggctct caatcgctta cgaactacag 60
 aagctgcaag tcccttacca agtgctggag gtttctggac attctggtgg agtcatgaag 120
 tcactccgga aggacggatt tgaactcgac gctggtgcca acaccatagc cgcgtctccc 180
 gagattcttg cgtactttac ctactaggt cttgagaatg agatcctcca ggcgactgct 240
 gcttctaaac accgcttctt ggtgcggcga aggcaactgc acgcgtgag cccgcacccg 300
 ttcaagatca tgtcatcgcc gtacctcagc cgtgggtcca aatggcggtt ctttactgag 360
 cggtttcgga agcccgctgt cgcttcgggc gaggagaccg tcaccgattt catcacgagg 420
 agattcaacc gcgaaatagc ggagtatgtg ttcgaccctg ttctaagcgg gatctacgcc 480
 gggaacccgg accaaatgag tattgctgag gtgttgcttg ccttgccctag gtgggaaagg 540
 gagtacggat cagtgaccaa gggccttatg aaggataagg gtgcgatggg aggtcgaaag 600
 atcatcagct ttaagggtgg caaccagcta cttacaaacc gcttacagca gctactcact 660

actccggtga gattcaattg caaggtgaca gggattacag ccagcaatgg cgggtacatc 720
 gtgagcgctg ttgaggacgg cgtatctgag agctacaccg catctcgtgt gatcttgacc 780
 acaccgcgtt actcagcagc ggctaccata actaaccttg atgcagccac tgcggcactg 840
 ttgaacgaaa tccattatcc acgtatgggc gtgttacact tgggctttga tgcaactgcc 900
 ttgccacagc cgctggacgg gttcggattt ctagtgccga acgcggagaa catgcacttc 960
 ctgggagcca tctgcaatgc agccatcttc ccggacaagg ctccgcccgg caagatcctg 1020
 tttacagtgt tcctcggagg cgcacgccag gagtcgctct tcgatcagat gactcctgag 1080
 gctcttcagc agcaagtcgt tagtgaggtg atgagcttgt tgcacttgtc agctccaccg 1140
 gtgatgcagc acttctcttc ctggaacaag gccatccctc aattgaacgt cgggcacgtg 1200
 aagttgcggc gcgcggtaga ggcgttcgag aagaaatacc ctggaatcca tctctcgggc 1260
 aactacctcc agggagttgc aataccagct ttactccagc acgccgcagc tttagctgct 1320
 tctcttaaga agaactga 1338

<210> 231
 <211> 1344
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 231
 atgagcgacc aaccgcctt catcggttga gctggtctct ccgggctctc aatcgcttac 60
 gaactacaga agctgcaagt cccttaccaa gtgctggagg tttctggaca ttctgggtga 120
 gtcatgaagt cactccggaa ggacggattt gaactcgacg ctgggtgccaa caccatagcc 180
 acgtctcccg agattcttgc gtactttacc tcaactaggtc ttgagaatga gatcctccag 240
 gcgactgcta cttctaaaca ccgcttcttg gtgcggcgaa ggcaactgca cgccgtgagc 300
 ccgcaccctg tcaagatcat gtcacgcgag tacctctgcc gtgggtccaa atggaggctc 360
 tttactgagc ggtttcggaa acccgctcgtc gttcggggcg aggagaccgt caccgatttc 420
 atcacgagga gattcaaccg cgaaatagcg gagtatgtgt tcgaccctgt tctaagtggg 480
 atctacgccg ggaaccggga ccaaataagc attgctgagg tgttgccctgc cttgcctagg 540
 tgggaaaggg agtacggatc agtgaccaag ggccttatga aggataaggg tgcgatggga 600
 ggtcgaaaga tcacagctt taagggtggc aaccagctac ttacaaaccg cttacagcag 660
 ctactcacta ctccggtgag attcaattgc aaggtgacag ggattacagc cagcaatggc 720
 ggggtacatc tgagcgctgt tgaggacggc gtatctgaga gctacaccgc atctcgtgtg 780

atcttgacca caccgctta ctcagcagcg gctaccataa ctaaccttga tgcagccact	840
gcggcactgt tgaacgaaat ccattatcca cgtatgggcg tgttacactt gggctttgat	900
gcaactgcct tgccacagcc gctggacggg ttcggatttc tagtgccgaa cgcggagAAC	960
atgcacttcc tgggagccat ctgcaatgca gccatcttcc cggacaaggc tccgcccggc	1020
aagatcctgt ttacagtgtt cctcggaggc gcacgccagg agtcgctctt cgatcagatg	1080
actcctgagg ctcttcagca gcaagtcgtt agtgagggtga tgagcttggt gcacttggtca	1140
gctccaccgg tgatgcagca cttctcctcc tggaacaagg ccatccctca attgaacgtc	1200
gggcacgtga agttgcgggc cgcggtagag gcgttcgaga agaaataccc tggaatccat	1260
ctctcgggca actacctcca gggagttgca ataccagctt tactccagca cgccgcagct	1320
ttagctgctt ctcttaagaa gaac	1344

<210> 232
 <211> 1338
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 232	
cagcccgtcc tcacgttgg agctggtctc tccgggctct caatcgctta cgaactacag	60
aagctgcaag tcccttacca agtgctggag gtttctggac attctggtgg agtcatgaag	120
tcactccgga aggacggatt tgaactcgac gctggtgcca acaccatagc cacgtctccc	180
gagattcttg cgtactttac ctactaggt cttgagaatg agatcctcca ggcgactgct	240
acttctaaac accgcttctt ggtgcggcga aggcaactgc acgcggtgag cccgcacccg	300
ttcaagatca tgtcatcgcc gtacctctgc cgtggctcca aatggaggct ctttactgag	360
cggtttcgga aaccgctcgt cgcttcgggc gaggagaccg tcaccgattt catcacgagg	420
agattcaacc gcgaaatagc ggagtatgtg ttcgaccctg ttctaagtgg gatctacgcc	480
gggaacccgg accaaatgag tattgctgag gtgttgcttg ccttgccatg gtgggaaagg	540
gagtacggat cagtgaccaa gggccttatg aaggataagg gtgcgatggg aggtcgaaag	600
atcatcagct ttaaggtgg caaccagcta cttacaaacc gcttacagca gctactcact	660
actccggtga gattcaattg caaggtgaca gggattacag ccagcaatgg cgggtacatc	720
gtgagcgctg ttgaggacgg cgtatctgag agctacaccg catctcgtgt gatcttgacc	780
acacccgctt actcagcagc ggctaccata actaaccttg atgcagccac tgcggcactg	840

ttgaacgaaa tccattatcc acgtatgggc gtgttacact tgggctttga tgcaactgcc	900
ttgccacagc cgctggacgg gttcggatth ctagtgccga acgcggagaa catgcacttc	960
ctgggagcca tctgcaatgc agccatcttc ccggacaagg ctccgcccgg caagatcctg	1020
tttacagtgt tcctcggagg cgcacgccag gagtcgctct tcgatcagat gactcctgag	1080
gctcttcagc agcaagtcgt tagtgagggt atgagcttgt tgcacttgct agctccaccg	1140
gtgatgcagc acttctcttc ctggaacaag gccatccctc aattgaacgt cgggcacgtg	1200
aagttgcggc gcgcggtaga ggcgttcgag aagaaatacc ctggaatcca tctctcgggc	1260
aactacctcc agggagttgc aataccagct ttactccagc acgccgcagc tttagctgct	1320
tctcttaaga agaactga	1338

<210> 233
 <211> 1401
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 233	
aagaagcacg tcgtcatcat aggcgggtggg atcactggct tggccgctgc attctacatg	60
gagaaggaga ttaaggagaa gaacctcca cttgagctga cgctagttga ggccagtccc	120
agggtcggcg gcaagatcca gacgggtcaag aaggacgggt acataattga acgcggccct	180
gacagcttct tagagcgcaa gaaatcggct ccgcagctag ttaaggactt gggacttgag	240
cacctgctcg tcaacaacgc gaccggacag tcgtacgtgc tcgtgaaccg gacgctccac	300
ccgatgccga agggcgctgt gatgggcatt ccgaccaaga tagcaccatt cgtgagtacc	360
ggcctattca gcctttccgg caaggcaagg gctgcgatgg acttcatctt gcctgcctct	420
aagactaagg acgatcagtc cttgggagag ttcttccgcc gccgggtggg tgatgagggtg	480
gtggagaact taattgagcc gctcctatct ggaatctacg ctggtgacat cgacaaactg	540
tctctgatgt ccacctttcc gcagttctac caaactgagc agaagcaccg ttcacttacc	600
ttgggaatga agaagactag acctcaaggt tcgggtcagc aactgacggc caagaaacag	660
ggtcagttcc agacgctaag caccgggctt cagacactcg tggaggagat tgagaaacag	720
ctcaaactta ctaaggtgta caagggcacg aaggtgacaa agttatccca ctccggcagc	780
gggtactccc tggagttgga caatggcgta acgttgagc ccgactcagt tatcgtgaca	840
gcgcgcgata aggtctgtgc cgggatgttg tcagaactcc cggcgatttc ccatctcaag	900
aacatgcaca gtacctcggg tgccaacgtc gccctcggat tcccgggaagg aagtgttcaa	960

atggagcacg aaggcacggg ttctgtaatt tccaggaact cgcactttgc catcaccgct 1020
 tgtacttggg ccaacaagaa gtggcctcat gctgcgccgg agggcaagac attgctcaga 1080
 gcttacgtcg ggaaggcggg cgacgagtca atcgctcgatc ttagcgacaa cgacatcatt 1140
 aacattgtgc tggaggactt gaagaaggtt atgaacatca atggcgagcc agagatgacc 1200
 tgcgtgaccc gatggcacga gtctatgccg cagtaccacg tcggtcacia gcagcgcac 1260
 aaggagttgc gcgaggcact cgcctcagct taccctggcg tgtacatgac tggcgcttcg 1320
 tttagaggcg ttggtattcc tgactgcacg gaccagggaa aggcggccgt cagtgcgcg 1380
 ctcacctacc tcttcagttg a 1401

<210> 234
 <211> 1398
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 234
 aagcacctgg tcataatcgg aggcggcata accggccttg ctgcggcctt ctacctggag 60
 aaggaggtcg aggagaaggg tctccctatc cagatttcat tgattgaggc ttcgcctcgg 120
 ctgggagggg agatccagac attgtacaag gacgggtaca tcatcgagcg tggtcagac 180
 agtttctctg agcgggaagg cagcggaccg cagctcgcca aggacgtggg acttagcgac 240
 caactggtga acaacgagac aggacaggcg tacgtcttgg tgaatgagaa gttgcacccg 300
 atgcctaagg gtgccgtgat gggcatccca acgcaaactt cacctttcat caccaccgga 360
 ctcttctccg tggccggaaa ggcacgagct gcaatggact tcgttctgcc taagtcgaaa 420
 cagaccgaag accagtctct aggcgagttc ttccgccgcc gtgtgggtga cgaggttgtg 480
 gagaacctca tcgagccttt gttgtctggg atctacgcgg ggcacatcga cagacttagt 540
 ctcatgagta cctttccgca attctatcag acagaacagc agcatcgaag tctcactc 600
 gggatgaaga agtcacaaca acatgcaaag gccagcaag ttaccgcca gaaacagggc 660
 cagttccaaa cgatcaacca gggcctccag agcttgggtg aggcagtggg gggaaagttg 720
 aagctcacca ccgtttacaa agggacaaa gttaaacaga ttgagaagac ggacggcggt 780
 tacgggttac aattggactc cggacagact ctcttcgctg attccgctat cgtaactact 840
 cctcaccaga gcatctactc tatgttcccc aaggaggcgg gcctggagta cctgcacgac 900
 atgacttcaa cgtctgtcgc caccgtggct ttgggcttca aggacgagga cgtccacaat 960

gagtatgacg ggacgggatt cgttatcagt aggaactccg acttcagcat caccgcctgc 1020
 acgtggacca acaagaagtg gccacacacc gcgccc aaag ggaagaccct tctgagggca 1080
 tacgtgggca aggcgggcca cgagagcatc gtcgagcaat ctgattctca gattgtttca 1140
 atcgtcctcg aagacctcaa gaagatcatg gacatcaagg cagacccgga acttaccacc 1200
 gttactcgat ggaagacctc gatgcctcag tatcacgtcg ggcaccagaa ggcaatcagc 1260
 aacatgaggg agacattcaa gcagtcgtat cctggcgtgt acattaccgg agcagcattc 1320
 gaaggcgtag gaatccctga ctgcattgac cagggcaagg ctgctatctc agaggccgtg 1380
 tcctatctct tctcgtga 1398

<210> 235
 <211> 1398
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 235
 aagcacctgg tgataattgg aggcgggatt accggcctag cagccgcttt ctatctggag 60
 aaggaggtgg aggagaaggg cctcccgata cagatttcgc tgattgaagc ctctccgcgc 120
 ctgggcggca agatccagac attgtacaag gacgggtaca tcattgagcg cgggcctgac 180
 tcgttcctgg agcgggaagg ctccggtcct caactggcca aagacgtggg tctttccgat 240
 cagcttgtga acaatgagac cggtcaggct tacgtcttgg tcaacgaaac tctgcatccc 300
 atgcctaagg gagccgttat gggcattcca acgcaaactc ctccgttcat aacgactggg 360
 ctgttcagcg ttgcgggcaa agcaagggt gctatggact tcgtgctgcc aaagagtaag 420
 cagaccgagg accagtccct cggcgagttc ttccgccgcc gagtgggcca tgaggtggtt 480
 gagaatctaa tcgaaccgct gttgtcgggc atctatgcgg gcgacatcga caggctaagt 540
 cttatgtcca ctttccctca gttctaccag acagagcaga aacacaggag tctcatcctt 600
 ggaatgaaga agtcccagca gcacgcgaag gctcagcaag tgaccgcca gaagcaagga 660
 cagttccaga ccatcaacca gggcctacag gcccttgctg aagccgttga gtcgaagtta 720
 aagttgacga cgatctacaa gggcaccaag gtgaagcaga ttgagaagac tgacggtggc 780
 tatggtgtgc aactcgattc gggccaaaca ttgctcgtg actccgctat cgtcacgacg 840
 ccacaccagt cgatctactc gatgttcccg aaggaggcgg gcctagagta ctttcacgac 900
 atgacctcca cttcggtcgc caccgttgca ctcggttcta aggaggagga cgttcacaac 960
 gagtacgatg gcaccggatt cgtgatctcc aggaactcgg acttctcgat taccgcgtgc 1020

acgtggacaa ataagaagtg gccgcacaca gcgccaaagg gcaagaccct tctgcgggcg 1080
 tatgtgggca aggccggtga cgagagcatt gtcgaacaat ctgaccatca gatcgtttct 1140
 attgttcttg aggatctcaa gaagataatg gacattaagg ccgaccctga gcttaccaca 1200
 gtgacgaggt ggaagacctc gatgccgcag tatcacgtag ggcaccagaa ggccatctcc 1260
 aacatgcggg agacattcaa gcagtcgtac cctggcgtgt acattactgg cgctgctttc 1320
 gagggcggtg gcatcccgga ctgcatcgac cagggcaagg ccgcaatctc agaggcagtg 1380
 tcgtacctgt tcagctag 1398

<210> 236
 <211> 68
 <212> PRT
 <213> Glycine max

<400> 236

Met Ala Thr Ala Thr Thr Thr Ala Thr Ala Ala Phe Ser Gly Val Val
 1 5 10 15

Ser Val Gly Thr Glu Thr Arg Arg Ile Tyr Ser Phe Ser His Leu Gln
 20 25 30

Pro Ser Ala Ala Phe Pro Ala Lys Pro Ser Ser Phe Lys Ser Leu Lys
 35 40 45

Leu Lys Gln Ser Ala Arg Leu Thr Arg Arg Leu Asp His Arg Pro Phe
 50 55 60

Val Val Arg Cys
 65

<210> 237
 <211> 56
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 237

Met Ala Thr Thr Thr Ala Ala Ala Ala Val Thr Ile Ser Ile Pro Lys
 1 5 10 15

Lys Pro Val Phe Ile Arg Arg Pro Arg Leu Arg Gly Pro Val Asp Cys
 20 25 30

Leu Arg Met Ala Arg Arg Gly Ser Gly Leu Ala Pro Val Val Gly Arg
50 55 60

Arg Pro Arg Arg Pro Pro Leu Ser Val Arg Cys Asp Ala Thr Ser Arg
65 70 75 80

<210> 240
<211> 84
<212> PRT
<213> *Acalypha ostryifolia*

<400> 240

Met Ala Thr Thr Thr Ala Thr Thr Ser Phe Ser Gly Val Ser Ile Cys
1 5 10 15

Pro Pro His Gln Thr Asn Arg Thr Ser Leu Phe Pro Pro Gln Ser Leu
20 25 30

Ser Phe Pro Ser Ser Lys His Gly Ser Leu Val Asn Ser Val Gln Phe
35 40 45

Asn Arg Ser Arg Arg Ala Arg Arg Asn His Phe Ser Leu Thr Ser Ile
50 55 60

Thr Asn Ala Pro Arg Arg Lys Arg Leu Leu Ser Val Arg Cys Asp Ala
65 70 75 80

Ser Ala Thr Ser

<210> 241
<211> 78
<212> PRT
<213> *Adansonia digitata*

<400> 241

Met Ala Ala Ser Ser Ser Ser Val Val Ser Phe Ser Gly Ile Ser Leu
1 5 10 15

Cys Ser Thr His Ser Ile Ser Asn Lys Thr Tyr Leu Phe Ser Ala His
20 25 30

Pro Arg Ile Ser Val Ser Phe Pro Ser Lys Pro Asn Ser Leu Lys Ser
35 40 45

Phe Lys Gln Leu Gln Leu Lys Lys Asn Gly Leu Phe Glu Lys Phe Ser
50 55 60

Arg Thr Ser Ser Arg Ser Phe Val Val Arg Cys Asp Ala Ser
65 70 75

<210> 242
<211> 82
<212> PRT
<213> Taraxacum officinale

<400> 242

Met Ala Thr Thr Ala Ser Phe Ser Gly Val Arg Ile His Ala Pro Ser
1 5 10 15

Ser Thr Cys Ile Asp Arg Thr Thr Leu Phe Ala Gln Pro Ser Val Ser
20 25 30

Phe Ser Ser Phe Ser Lys Pro Arg Arg Thr Thr Leu Arg Ser Leu Lys
35 40 45

Leu Arg Ser Arg Ser Asn Asp Val Leu Leu Arg Thr Arg Thr Gly Asp
50 55 60

Arg Phe Gly Gly Lys Ser Ser Arg Ser Phe Val Val Arg Cys Asp Ala
65 70 75 80

Ser Ser

<210> 243
<211> 77
<212> PRT
<213> Amaranthus cruentus

<400> 243

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

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Gln Ser Ile Val Ser
20 25 30

Ile Ala Arg Asn Ser Arg Lys Pro Lys Ser Leu Lys Ser Leu Lys Leu
35 40 45

Ser Thr Asn Ser Phe Asn Phe Gly Leu His Lys Ser Cys Arg Lys Gly
50 55 60

Ser Lys Ser Gly Ser Phe Val Val Arg Cys Asp Ala Ala

65

70

75

<210> 244
 <211> 80
 <212> PRT
 <213> *Amaranthus cruentus*

<400> 244

Met Ala Ile Ala Thr Thr Ser Phe Pro Gly Thr Tyr Leu Arg Val Pro
 1 5 10 15

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Arg Ser Val Val Ser
 20 25 30

Asn Gly Val Asn Ser Arg Lys Pro Asn Ser Leu Glu Ser Leu Lys Ser
 35 40 45

Ser Arg Asn Ser Ser Asn Val Cys Leu Ser Thr Ser Phe Gly His Tyr
 50 55 60

Arg Lys Ser Ser Lys Ser Gly Ser Phe Phe Val Arg Cys Asn Ala Ala
 65 70 75 80

<210> 245
 <211> 80
 <212> PRT
 <213> *Amaranthus hypochondriacus*

<400> 245

Met Ala Ile Ala Thr Thr Ser Phe Pro Gly Thr Tyr Leu Arg Val Pro
 1 5 10 15

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Arg Ser Val Val Ser
 20 25 30

Asn Gly Val Asn Ser Arg Lys Pro Asn Ser Leu Glu Ser Leu Lys Ser
 35 40 45

Ser Arg Asn Ser Ser Asn Val Cys Leu Ser Thr Ser Phe Gly His Tyr
 50 55 60

Arg Lys Ser Ser Lys Ser Gly Ser Phe Phe Val Arg Cys Asn Ala Ala
 65 70 75 80

<210> 246
 <211> 77

<212> PRT

<213> *Amaranthus palmeri*

<400> 246

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

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Gln Ser Ile Val Ser
 20 25 30

Ile Ala Leu Asn Ser Arg Lys Pro Lys Ser Phe Lys Ser Leu Lys Ser
 35 40 45

Ser Ala Asn Ser Cys Asn Phe Gly Leu His Lys Ser Tyr Arg Lys Gly
 50 55 60

Ser Lys Ser Gly Ser Phe Val Val Arg Cys Asp Ala Ala
 65 70 75

<210> 247

<211> 80

<212> PRT

<213> *Amaranthus palmeri*

<400> 247

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

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Arg Ser Val Val Ser
 20 25 30

Asn Gly Val Asn Ser Arg Lys Pro Asn Ser Leu Lys Ser Leu Lys Leu
 35 40 45

Ser Arg Asn Ser Ser Asn Val Cys Leu Tyr Thr Ser Phe Gly His Tyr
 50 55 60

Arg Lys Ser Ser Lys Ser Gly Ser Phe Ile Ile Arg Cys Asn Ala Ala
 65 70 75 80

<210> 248

<211> 77

<212> PRT

<213> *Amaranthus palmeri*

<400> 248

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

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Gln Ser Ile Val Ser
 20 25 30

Ile Ala Leu Asn Ser Arg Lys Pro Lys Ser Phe Lys Ser Leu Lys Ser
 35 40 45

Ser Ala Asn Ser Cys Asn Phe Gly Leu His Lys Ser Tyr Arg Lys Gly
 50 55 60

Ser Lys Ser Gly Ser Phe Val Val Arg Cys Asp Ala Ala
 65 70 75

<210> 249
 <211> 77
 <212> PRT
 <213> *Amaranthus palmeri*

<400> 249

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

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Gln Ser Ile Val Ser
 20 25 30

Ile Ala Leu Asn Ser Arg Lys Pro Lys Ser Phe Lys Ser Leu Lys Ser
 35 40 45

Ser Ala Asn Ser Cys Asn Phe Gly Leu His Lys Ser Tyr Arg Lys Gly
 50 55 60

Ser Lys Ser Gly Ser Phe Val Val Arg Cys Asp Ala Ala
 65 70 75

<210> 250
 <211> 80
 <212> PRT
 <213> *Amaranthus palmeri*

<400> 250

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

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Arg Ser Val Val Ser

20 25 30
 Asn Gly Val Asn Ser Arg Lys Pro Lys Ser Leu Lys Ser Leu Lys Ser
 35 40 45
 Ser Arg Asn Ser Ser Asn Val Cys Leu Tyr Thr Ser Phe Gly His Tyr
 50 55 60
 Arg Lys Ser Ser Lys Ser Gly Ser Phe Ile Ile Arg Cys Asn Ala Ala
 65 70 75 80

<210> 251
 <211> 80
 <212> PRT
 <213> *Amaranthus rudis*

<400> 251

Met Ala Thr Ala Thr Thr Ser Phe Pro Gly Thr Tyr Leu Arg Val Pro
 1 5 10 15
 Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Arg Ser Val Val Ser
 20 25 30
 Asn Gly Val Asp Ser Arg Lys Pro Asn Ser Leu Lys Ser Met Lys Leu
 35 40 45
 Ser Arg Asn Ser Ser Asn Val Cys Leu Tyr Thr Ser Phe Gly His Tyr
 50 55 60
 Arg Lys Ser Ser Lys Ser Gly Ser Phe Ile Val Arg Cys Asn Ala Ala
 65 70 75 80

<210> 252
 <211> 80
 <212> PRT
 <213> *Amaranthus rudis*

<400> 252

Met Ala Thr Ala Thr Thr Ser Phe Pro Gly Thr Tyr Leu Arg Val Pro
 1 5 10 15
 Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Arg Ser Val Val Ser
 20 25 30
 Asn Gly Val Asp Ser Arg Lys Pro Asn Ser Leu Lys Ser Met Lys Leu
 35 40 45

Ser Arg Asn Ser Ser Asn Val Cys Leu Tyr Thr Ser Phe Gly His Tyr
 50 55 60

Arg Lys Ser Ser Lys Ser Gly Ser Phe Ile Val Arg Cys Asn Ala Ala
 65 70 75 80

<210> 253

<211> 80

<212> PRT

<213> *Amaranthus rudis*

<400> 253

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

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Arg Ser Val Val Ser
 20 25 30

Asn Gly Val Asp Ser Arg Lys Pro Asn Ser Leu Lys Ser Met Lys Leu
 35 40 45

Ser Arg Asn Ser Ser Asn Val Cys Leu Tyr Thr Ser Phe Gly His Tyr
 50 55 60

Arg Lys Ser Ser Lys Ser Gly Ser Phe Ile Val Arg Cys Asn Ala Ala
 65 70 75 80

<210> 254

<211> 80

<212> PRT

<213> *Amaranthus viridis*

<400> 254

Met Ala Ile Ala Thr Thr Ser Phe Pro Gly Thr Tyr Leu Arg Val Pro
 1 5 10 15

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Arg Ser Val Val Ser
 20 25 30

Asn Gly Val Asn Ser Arg Lys Pro Asn Ser Leu Glu Ser Leu Lys Ser
 35 40 45

Ser Arg Asn Ser Ser Asn Val Cys Leu Ser Thr Ser Phe Gly His Tyr
 50 55 60

Arg Lys Gly Ser Lys Ser Gly Ser Phe Val Val Arg Cys Asp Ala Ala
 65 70 75 80

<210> 255
 <211> 77
 <212> PRT
 <213> *Amaranthus viridis*

<400> 255

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

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Gln Ser Ile Val Ser
 20 25 30

Ile Ala Arg Asn Ser Arg Lys Pro Lys Ser Leu Lys Ser Leu Lys Leu
 35 40 45

Ser Thr Asn Ser Phe Asn Phe Gly Leu His Lys Ser Cys Arg Lys Gly
 50 55 60

Ser Lys Ser Gly Ser Phe Val Val Arg Cys Asp Ala Ala
 65 70 75

<210> 256
 <211> 71
 <212> PRT
 <213> *Ambrosia trifida*

<400> 256

Met Ser Thr Met Ser Thr Leu Phe His Leu Pro Ser Ser Leu Cys Thr
 1 5 10 15

Asp Arg Thr Ile Thr Ser Ser Phe Ala Gln Pro Ser Val Ser Val Asn
 20 25 30

Ser Phe Ser Lys Pro Arg Arg Val Ala Leu Arg Ser Leu Lys Leu Lys
 35 40 45

Thr Arg Ser Asn Asp Val Leu Leu Arg Lys Ser Ser Arg Ser Leu Val
 50 55 60

Val Arg Cys Asp Ala Ser Ser
 65 70

<210> 257
 <211> 87
 <212> PRT
 <213> *Conyza canadensis*

<400> 257

Met Ala Thr Ala Ala Phe Ser Gly Val Pro Cys Ile Asp Arg Thr Ser
 1 5 10 15

Leu Leu Ser Ala Gln Pro Ser Ser Ser Ser Ser Ser Val Val Val
 20 25 30

Cys Tyr Ser Ser Phe Ser Lys Pro Gly Thr Thr Leu Leu Pro Ser Leu
 35 40 45

Lys Leu Lys Ser Ser Arg Asn Asn Asn Asn Ser Asn Val Phe Leu Phe
 50 55 60

Gly Asn Thr Arg Lys Thr Ser Arg Leu Ser Phe Leu Val Arg Cys Asp
 65 70 75 80

Ser Ser Ser Ser Ser Ser Ser
 85

<210> 258
 <211> 71
 <212> PRT
 <213> *Cucumis melo*

<400> 258

Met Ala Phe Ser Thr Ala Pro Phe Tyr Ala Ile Gly Ile Arg Phe Pro
 1 5 10 15

Ser His Ser Ser Ser Ile Ser Ser Thr Thr Asn Ala Leu Ile Leu Lys
 20 25 30

Ser Pro Leu Ala Leu Ala Leu Thr Ala Lys Pro Lys Ser Pro Leu Leu
 35 40 45

Leu Lys Arg Asn Val Gly Cys Gln Arg Phe Gly Arg Asn Ser Arg Phe
 50 55 60

Val Val Arg Cys Asp Ala Ser
 65 70

<210> 259

<211> 78
<212> PRT
<213> Kochia Scoparia

<400> 259

Met Ala Thr Ala Thr Thr Ser Phe Pro Gly Ala Tyr Leu His Leu Pro
1 5 10 15

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Gln Pro Ile Cys Ser
20 25 30

Ser Asn Leu Asn Leu Lys Lys Pro Asn Ser Leu Lys Ser Val Lys Leu
35 40 45

Ser Arg Ser Ser Gly Asn Ala Leu Phe Tyr Lys Asn Ala Lys Lys Asn
50 55 60

Ser Lys Phe Gly Ser Leu Val Val Arg Cys Asp Ala Ala Gly
65 70 75

<210> 260
<211> 78
<212> PRT
<213> Kochia Scoparia

<400> 260

Met Ala Thr Ala Thr Thr Ser Phe Pro Gly Ala Tyr Leu His Leu Pro
1 5 10 15

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Gln Pro Ile Cys Ser
20 25 30

Ser Asn Leu Asn Leu Lys Lys Pro Asn Ser Leu Lys Ser Val Lys Leu
35 40 45

Ser Arg Ser Ser Gly Asn Ala Leu Phe Tyr Lys Asn Ala Lys Lys Asn
50 55 60

Ser Lys Phe Gly Ser Leu Val Val Arg Cys Asp Ala Ala Gly
65 70 75

<210> 261
<211> 78
<212> PRT
<213> Kochia Scoparia

<400> 261

Met Ala Thr Ala Thr Thr Ser Phe Pro Gly Ala Tyr Leu His Leu Pro
 1 5 10 15

Pro Lys Asn Gly Val Arg Asn Ala Leu Phe Ser Gln Pro Ile Cys Ser
 20 25 30

Ser Asn Leu Asn Leu Lys Lys Pro Asn Ser Leu Lys Ser Val Lys Leu
 35 40 45

Ser Arg Ser Ser Gly Asn Ala Leu Phe Tyr Lys Asn Ala Lys Lys Asn
 50 55 60

Ser Lys Phe Gly Ser Leu Val Val Arg Cys Asp Ala Ala Gly
 65 70 75

<210> 262
 <211> 84
 <212> PRT
 <213> Rosa hybrid cultivar

<400> 262

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

Leu Pro Gln Ser Val Ser Thr Lys Cys Cys Ser Lys Ala Ser Leu Phe
 20 25 30

Pro His Pro Thr Leu Ser Leu Thr Phe His Ala Arg Pro Gln Phe Phe
 35 40 45

Arg Gly Leu Ala Ser Arg Gln Phe Asn Pro Asn Gly Ala Phe Gly Thr
 50 55 60

Gly Ser Gly Arg Leu Gly Arg Thr Pro Asn Pro Phe Val Val Arg Ser
 65 70 75 80

Glu Ala Ser Ser

<210> 263
 <211> 80
 <212> PRT
 <213> Sedum album

<400> 263

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

Pro Lys Pro Ser Ser Phe Ser Ser Ser Pro Ser Val Thr Val Pro Arg
20 25 30

Phe Ser Val Ser Cys Ser Ala Ile Pro Arg Pro His Lys Asn Pro Cys
35 40 45

Ser Leu Lys Phe Arg Val Lys Asp Ser Arg Phe Asn Gly Ile Val Lys
50 55 60

Lys Arg Ser Asn Ser Asn Ser Phe Val Val Arg Cys Asp Thr Ser Ser
65 70 75 80

<210> 264
<211> 79
<212> PRT
<213> Sedum album

<400> 264

Met Ala Ala Asp Ala Ala Thr Ile Thr Ala Gly Ile Thr Leu Thr Thr
1 5 10 15

Ala Arg Arg Ser Ser Ser Ser Ile Ala Pro Gln Phe Ser Val Cys Cys
20 25 30

Ser Ala Ile Thr Asn Thr Gln Lys Asn Leu Ser Phe Leu Lys Leu Arg
35 40 45

Val Lys Asp Ala Thr Leu Thr Thr Arg Ile Glu Gly Ile Gln Lys Lys
50 55 60

Arg Tyr Asn Ser Ala Ser Phe Val Val Arg Cys Asp Ala Ser Ser
65 70 75

<210> 265
<211> 81
<212> PRT
<213> Spinacia oleracea

<400> 265

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

Pro Asn Asn Gly Val Arg Asn Ala Leu Phe Ser Gln Pro Phe Leu Ser

20 25 30
 Leu Arg Ile Lys Ser Lys Arg Thr Lys Ser Leu Asn Ser Leu Lys Phe
 35 40 45
 Thr Gly Asp Ser Ser Lys Ile Leu Leu Phe Lys Cys Ser Arg Pro Phe
 50 55 60
 Glu Lys Gly Leu Lys Ser Gly Ser Phe Val Val Arg Cys Asp Ala Ala
 65 70 75 80

Gly

<210> 266
 <211> 79
 <212> PRT
 <213> Allium cepa

<400> 266

Met Ala Ala Thr Ser Ser Ala Thr Thr His Leu Pro Phe Phe Ser Pro
 1 5 10 15

His Thr Lys His Ala Lys Thr Asn Ser Phe Phe Ala Ser Leu Pro Val
 20 25 30

Ser Ala Tyr Ser Thr Lys Asn Ser Ile Ser Phe Lys Ala Leu Lys Ala
 35 40 45

Val Arg Trp Ser Glu Thr Phe Gly Gln Ser Lys Lys Ala Asn Gly Phe
 50 55 60

Ala Lys Arg Lys Gln Phe Ala Val Val Arg Cys Asp Ala Ser Ser
 65 70 75

<210> 267
 <211> 204
 <212> ADN
 <213> Glycine max

<400> 267
 atggctactg ctactaccac agctaccgct gcattctctg gtgttgtag tgttggaacc 60
 gagacacgta gaatttactc tttctcacac ttgcaaccta gcgcagcctt ccctgccaag 120
 ccatcatcct ttaagtcctt gaagctgaaa cagtcggcga ggcttacgag gcgcctcgat 180
 catagaccct ttgtggtccg atgc 204

<210> 268
 <211> 168
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tái tổ hợp

<400> 268
 atggccacta ccacagcagc cgcggcggtc accatcagca ttcctaaaaa gcctgttttt 60
 atccgccgcc cagcacttcg tgggcccgtc gactgcagag gcctgcatgc atccgacgca 120
 atcatctcca acgaggcccc tacagggacg acaatctcgg ctgactgt 168

<210> 269
 <211> 228
 <212> ADN
 <213> *Eragrostis tef*

<400> 269
 atggcagccg cacctcccct agcagccgac atgggtgttac catcccatg ccctgccgcg 60
 gttgcaccta cccagtggt tgcagctgct tgggggtgcag cccgagctgg atctgttaga 120
 tgtaaagcga cccaacttcg aatgatgaga actggggggc ctgttgctcc agttgccggg 180
 agacgacgac gagctccatt gagtgtacgt tgtgatgctt cctccaga 228

<210> 270
 <211> 240
 <212> ADN
 <213> *Setaria italica*

<400> 270
 atggctgccg ctccctcccct ctctgcagat gcactatcat tcctaccatc cgccgccgct 60
 ccggcagccg ctgcaccaac acctgttgta gctgcggcat ggggagccgc acgagctgca 120
 gggtcagtta gaggtaaagc tgctttgcgt atggctcgaa ggggtagtgg actggctcca 180
 gtgggttgaa gaagacctcg acgacctcct ctttcagtta gatgtgacgc aacatctcgt 240

<210> 271
 <211> 252
 <212> ADN
 <213> *Acalypha ostryifolia*

<400> 271
 atggctacaa ccaccgcgac gacgtctttc tcgggcgtct cgatctgccc acctcaccag 60
 acgaatcgca cctctttggt tccgccccag tccttgtctt tcccctccag taagcatggc 120
 agtctttgtga actctgtgca attcaaccgt tcgcgacgcg ctagacgtaa tcacttcagc 180

ctcacttcca ttaccaatgc accgaggcgc aaaaggttac tatctgtccg gtgcgacgcg 240
 agtgccacat ct 252

<210> 272
 <211> 234
 <212> ADN
 <213> *Adansonia digitata*

<400> 272
 atggcgcgct catcttcgtc cgtcgtgagc ttctcgggca tctcgttggt cagtactcac 60
 tcgatctcca acaagaccta tctattctcc gccacccgc gcatttcggt gtcgttcccc 120
 agtaagccca atagtttgaa gtccttcaag cagctccagc tgaagaagaa cggactcttt 180
 gagaagttct ctcgtacctc cagtcggagc ttcgtggtga ggtgcgacgc gtcg 234

<210> 273
 <211> 246
 <212> ADN
 <213> *Taraxacum officinale*

<400> 273
 atggctacaa ccgcgagctt ctcgggtggt cgtattcacg cgccttcctc cacatgtatc 60
 gaccggacca ctttattcgc ccagccttcg gtgagctttt cttccttttc caagccgagg 120
 cgaacgacct tgaggtcgct gaagctaagg tcgagggtcca acgatgtggt gcttcgcacc 180
 cgcacaggtg acagattcgg cggaaagagc tcacgttcat ttgttggtgcg ctgcgacgca 240
 tcttct 246

<210> 274
 <211> 231
 <212> ADN
 <213> *Amaranthus cruentus*

<400> 274
 atggcgaccg cgacgacctc gtttcccggc gcgtacctgc gcgtgccgcc caagaacggg 60
 gttcgtaacg ccctcttttag ccagtctatc gtgtcaatag cgcgcaactc tcggaaaccc 120
 aaatcgctca aatcccttaa actatctacc aactccttta acttcggtct gcacaagtct 180
 tgtcgaaagg gaagcaaadc cgggtcggtt gtagtgcggt gtgacgcggc c 231

<210> 275
 <211> 240
 <212> ADN
 <213> *Amaranthus cruentus*

<400> 275

atggccatcg ccaccacgag ctttcggga acgtacctcc gggtgccgcc caagaacggc	60
gtccgaaacg ccctattcag tcgctccgtc gtgtctaata ggggtgaactc aaggaagccg	120
aactcgctgg agtcgcttaa atcgctcgagg aatagctcga acgtctgctt gactacgtcg	180
ttcgggcatt accggaatc gagtaagtcg ggctcgttct tcgttcgggtg taacgccgcc	240

<210> 276
 <211> 240
 <212> ADN
 <213> *Amaranthus hypochondriacus*

<400> 276	
atggccatcg ccacgacctc gttccccggc acgtacctgc gagtgccgcc caagaacggg	60
gtccggaacg cgctgtttct tcgggtccgtg gtcagcaatg gtgttaattc acgaaagccg	120
aacagtctgg aatctctcaa gagcagtcga aactcctcca acgtctgcct ttcgaccagc	180
ttcggtcact accggaagtc tagtaagagc gggtcgttct ttgtccgggtg taatgctgcc	240

<210> 277
 <211> 231
 <212> ADN
 <213> *Amaranthus palmeri*

<400> 277	
atggcgacgg ccaccacctc gttccccggc gcataacctcc gcgtgccgcc caagaacggc	60
gtccgcaacg cactcttttag ccagagcatc gtcagtatcg cccttaacag tcgcaagcct	120
aaatcgttca agtcactaaa gtcaagcgct aattcggtgca actttggact tcacaagtcc	180
taccggaag gcagcaagtc tggcagcttc gtcgttcgtt gtgatgccgc c	231

<210> 278
 <211> 240
 <212> ADN
 <213> *Amaranthus palmeri*

<400> 278	
atggccaccg ccaccacctc gttccccggc acgtacctgc gcgtgccgcc caagaacggc	60
gtcaggaacg cgctgttttag tcgctcagtc gtgtccaacg ggggtgaactc acgaaagccc	120
aacagtttaa agagcttaaa actgtcgagg aactctagta atgtctgcct ctacacctcc	180
ttcggacact atcgaaagtc cagtaagtcg ggctcattca tcatccgggtg caatgcggcc	240

<210> 279
 <211> 231
 <212> ADN
 <213> *Amaranthus palmeri*

<400> 279
 atggccaccg caacaaccag cttccctggc gcgtaccttc gtgtgccgcc caagaacggc 60
 gtccgcaatg cgctgttttag tcagtccatc gtgagtatcg ctctcaattc ccggaaacct 120
 aagagcttta agagcctaaa gtcgagcgct aattcttgca acttcgggct tcacaagagc 180
 tatcggaag ggtctaagag cgggtcattc gtcgtgcggg gcgacgcggc c 231

<210> 280
 <211> 231
 <212> ADN
 <213> *Amaranthus palmeri*

<400> 280
 atggccacgg ccacgacctc gtttccgggt gcgtacctgc gagttccgcc caagaacggt 60
 gtacggaacg ccttgttctc ccaatccatc gtgagcatcg ccctcaacag tcgcaaaccg 120
 aagtcattca aatccctgaa aagttcggcc aatagctgta acttcgggct gcataagagt 180
 taccgcaagg ggtcgaaatc cgggtcgttc gtcgtccggg gcgacgctgc g 231

<210> 281
 <211> 240
 <212> ADN
 <213> *Amaranthus palmeri*

<400> 281
 atggcaactg ccacgacgtc ctttccggga acgtatctcc gcgtgccgcc caagaacggc 60
 gtccgcaacg ccctgttctc acgatccgtc gttagcaatg gcgtcaatag ccgcaagcct 120
 aagtcctga aatcgctcaa gtcgtcgcgc aactctagta atgtctgtct ctacacatcg 180
 ttcggacatt accgcaaatc atccaaatcc gggtcgttca taatccgggtg caatgcggct 240

<210> 282
 <211> 240
 <212> ADN
 <213> *Amaranthus rudis*

<400> 282
 atggcgacag cgaccacatc cttccctggc acttacctga gagtgccgcc caagaatggg 60
 gtgagaaaacg ccttgttcag ccgcagcgta gtctctaatt ggggtggatag tcgcaaaccg 120
 aatagcctca agagtatgaa gtcagccgc aacagctcaa atgtctgcct ctacacgagc 180
 tttggccact accgaaagtc ctccaagtct gggtcgttca tcgtgcgctg taacgccgcg 240

<210> 283
 <211> 240
 <212> ADN
 <213> *Amaranthus rudis*

<400> 283
 atggccacgg ccaccacctc ctttcctggc acatacctcc gcgtccctcc caagaatggg 60
 gtgcgaaacg cactcttttag tagatcggtc gtttccaatg gtgtcgattc ccgcaagccg 120
 aactccctca agtcgatgaa gctgtcccg c aactcatcga acgtttgcct ctatacctcg 180
 tttgggcact accgcaagtc gagcaaatacg ggctcgttca ttgtccggtg taatgcagcc 240

<210> 284
 <211> 240
 <212> ADN
 <213> *Amaranthus rudis*

<400> 284
 atggcgacgg cgacaacctc gtttccggga acgtacctgc gcgtgccgcc caagaacggg 60
 gtgcggaacg ccctgttcag ccgctccgtc gtgtccaatg gcgtcgattc gaggaagcct 120
 aactcattga aatctatgaa gttgtctcgt aattccagca acgtttgcct ctacacctcg 180
 ttcgggcatt accgcaagtc aagcaagtcc ggatcgttta tcgtgcggtg caacgctgcg 240

<210> 285
 <211> 240
 <212> ADN
 <213> *Amaranthus viridis*

<400> 285
 atggccattg ccaccacgtc gtttcccggc acgtacctga gggttccgcc caagaacgga 60
 gtccgcaacg cactgttttag tcgctccgtg gtgagtaacg gggtaactc cagaaaacct 120
 aattcgctgg agtcccttaa atcgagccgg aacagctcga acgtctgctt gtcaacctcc 180
 tttggccact accggaaggg ctccaagtcg ggctcattcg tcgtgcggtg cgatgcggcg 240

<210> 286
 <211> 231
 <212> ADN
 <213> *Amaranthus viridis*

<400> 286
 atggccaccg ccacgacgtc ctttcccggg gcgtatctgc gaggcctcc caagaacggc 60
 gtccggaacg cgctgttcag ccagtccatc gtgagcatcg cgcggaatag tcggaaacct 120
 aagtcgctca aatccttgaa actgtcaacg aactctttca atttcgggtt gcataagtcc 180
 tgccgaaagg gtagcaaata cgggtctttc gttgtgcggt gcgacgcggc c 231

<210> 287
 <211> 213
 <212> ADN

<213> *Ambrosia trifida*

<400> 287

atgagtacga tgtcaaccct atttcacctc ccgtctagcc tgtgtaccga caggacgatac 60
accagcagct tcgcacaacc gagcgtttcg gtcaactcgt tctcgaagcc gcgccgcgctc 120
gcgctccggt ccttaaagct caaaacgcga agtaatgacg tcctgctgcg gaaatcttca 180
cgttcgctag tcgtgcgttg cgacgccagc agc 213

<210> 288

<211> 261

<212> ADN

<213> *Conyza canadensis*

<400> 288

atggcgacgg ccgccttctc gggcgttccg tgcattgacc ggacatcact cctctccgcc 60
cagccatcgt cctcctcttc cagtagcgtc gtggtctgct actcctcctt tagcaagccg 120
ggcacgaccc tattgccgtc gttgaagctc aaaagcagcc gcaacaacaa caattcaaac 180
gtattcctct tcgggaacac caggaaaaca tcccgtctgt cattcctagt gcgctgcgat 240
tcctcatctt caagctctag c 261

<210> 289

<211> 213

<212> ADN

<213> *Cucumis melo*

<400> 289

atggcggtta gcaccgcacc cttctacgca attggatatca gatttcccag ccatagctca 60
tcaatctcaa gcaccactaa cgccctcatc cttaaaagtc cactggcggtt agccctaacc 120
gctaagccga agtctcccct actcctcaag cgcaacgttg gctgccagcg attcgggcga 180
aactcccgtc tcgtcgtgcg ctgcgatgcg tcc 213

<210> 290

<211> 234

<212> ADN

<213> *Kochia Scoparia*

<400> 290

atggcgaccg ccacgacctc cttccccggc gcgtaacctc atctcccgcc caagaacggg 60
gtccgcaacg ccttgttctc tcaaccctac tgttcatcca acctcaacct caagaaacct 120
aactctctca aatcggtgaa gctgtcccgc agttccggca atgccctatt ctacaagaac 180
gccaaagaaga atagtaagtt cggcagctctg gtcgtgcggt gcgacgcggc ggga 234

<210> 291
 <211> 234
 <212> ADN
 <213> Kochia Scoparia

<400> 291
 atggccaccg cgacgaccag ctttcccggc gcgtatctgc acctcccgcc caagaacggc 60
 gtgagaaacg cgctgttcag tcaaccgata tgctcgtcca atctcaacct caagaaaccc 120
 aattctctga aaagcgtcaa actgtcgcgt agtagcggca atgcgctggt ctacaagaac 180
 gccaaaga atagcaagtt cgggtcgcgc gtggtgcgct gcgacgcggc gggc 234

<210> 292
 <211> 234
 <212> ADN
 <213> Kochia Scoparia

<400> 292
 atggccacag ccaccacgtc cttccctggg gcctacctac atctcccgcc caagaatggc 60
 gtgcgaaaacg cgctgttcag tcagcctata tgcagcagta atcttaacct caagaagcct 120
 aattccctca agtcagtga actgagccgg tctagcggga acgcgctggt ctacaagaac 180
 gccaaaaaga atagcaagtt cggctcgcgc gtggtccggt gcgacgcggc gggc 234

<210> 293
 <211> 252
 <212> ADN
 <213> Rosa hybrid cultivar

<400> 293
 atggcgtcat cgaccacttc gttcgccgcc agtggagttg gattgcggct ccctcagtcc 60
 gtgagcacga agtgctgctc taaagcgtca ttgttccac accccacact atcgttgacc 120
 ttccacgcta ggccacagtt ctttagaggc ttggcgtctc gccagttcaa tccaaacgga 180
 gcgtttggga cgggctccgg acggctgggc cggacaccaa atccgtttgt cgtcagaagc 240
 gaagcgagtt ct 252

<210> 294
 <211> 240
 <212> ADN
 <213> Sedum album

<400> 294
 atggcggcga gcgcggctac gatcacctcc agcatatcgg cgattacccc gaagccgtcg 60
 tccttctcaa gcagcccttc ggtcaccgtg ccccgattct ctgtgtcgtg cagcgcgata 120
 ccgcgtccac acaagaatcc ctgctcgttg aagttccggg tgaaggactc acggtttaac 180

ggaattgtca agaagcgagcag taacagcaac tcattcgtag tacgttgtga cacttcctcg 240

<210> 295
<211> 237
<212> ADN
<213> Sedum album

<400> 295
atggccgccc acgcagctac cattacggcg ggtatcactc tcacgacggc ccgcccgtcc 60
tcctccagta ttgcgcccga gttctcggcg tggtgctcag cgattaccaa cacgcaaaaa 120
aatctgagct tcctcaagtt gcgcgtgaaa gacgccacct tgactacacg gattgagggt 180
attcagaaga agcgggtacaa ctccgcgctc ttcgctgtca gatgcgacgc gagcagc 237

<210> 296
<211> 243
<212> ADN
<213> Spinacia oleracea

<400> 296
atggccactg ccacgacgtc ctttctcggc gcttacttgc gggcgccgcc caacaatggc 60
gtgaggaatg cgctgttcag tcaaccgttc ctgtcgcttc gcattaagtc caaacgcact 120
aagagcctca actcgttgaa attcacagga gactcaagta agattctgct gtttaagtgc 180
tcccggccgt ttgagaaggg gcttaaattc ggctcgcttcg tggcgcgctg cgacgcggcc 240
ggt 243

<210> 297
<211> 237
<212> ADN
<213> Allium cepa

<400> 297
atggcggcaa cgagctccgc gaccactcac ctcccttttt tcagcccgc caacaaacac 60
gcaaagacaa actctttctt cgcgtccctt ccggtcagcg cctactccac gaaaaactct 120
atcagtttca aggcgtctaa ggccgtgcga tggagcgaga ctttcgggca atcgaagaag 180
gccaatggtt ttgccaaaag gaagcaattt gccgtcgtgc ggtgcgatgc gaggttca 237