



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0089631
(43) 공개일자 2012년08월13일

- (51) 국제특허분류(Int. Cl.)
C12P 7/06 (2006.01) C12N 15/52 (2006.01)
C12N 15/87 (2006.01) C12N 1/15 (2006.01)
(21) 출원번호 10-2012-7002889
(22) 출원일자(국제) 2010년07월26일
심사청구일자 없음
(85) 번역문제출일자 2012년02월01일
(86) 국제출원번호 PCT/US2010/043279
(87) 국제공개번호 WO 2011/011796
국제공개일자 2011년01월27일
(30) 우선권주장
61/271,833 2009년07월24일 미국(US)
61/285,526 2009년12월10일 미국(US)

- (71) 출원인
더 리전트 오브 더 유니버시티 오브 캘리포니아
미국, 캘리포니아 94607-5200, 오클랜드, 12층, 1111
프랭클린 스트리트, 오피스 오브 테크놀로지 트랜스퍼
비피 코퍼레이션 노쓰 아메리카 인코포레이티드
미국 60555 일리노이주 와렌빌 원펠드 로드 4101
더 보드 오브 트러스티즈 오브 더 유니버시티 오브 일리노이
미국 일리노이주 61801 어바나 에스. 라이트 506
헨리 어드미니스트레이션 빌딩 352
(72) 발명자
글라스, 앤., 루이스
미국 캘리포니아 94563 오란다 시닉 드라이브 115
덴, 차우구앙
중국 100053 베이징 니우지에 둥리 세컨드 블록
빌딩 1 아파트먼트 606
(뒷면에 계속)
(74) 대리인
특허법인무한

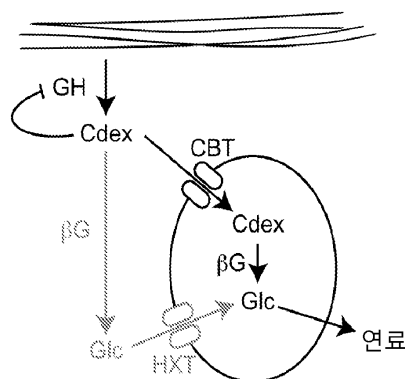
전체 청구항 수 : 총 119 항

(54) 발명의 명칭 개선된 당 수송, 혼합된 당 발효 및 바이오 연료의 생산을 위한 방법 및 조성물

(57) 요약

본 발명은 세포 내로 셀로덱스트린을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포에 관한 것이다. 본 발명은 추가로 세포 내로 셀로덱스트린의 수송을 증가시키는 방법, 셀로덱스트린을 포함하는 배지에서 세포의 성장을 증가시키는 방법, 셀룰로스-유도 및 헤미셀룰로스-유도 당을 공동-발효시키는 방법, 및 세포 내로 셀로덱스트린을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 준비하여 탄화수소 또는 탄화수소 유도체를 제조하는 방법에 관한 것이다. 본 발명은 세포 내로 오탄당을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포, 세포 내로 오탄당의 수송을 증가시키는 방법, 오탄당 당을 포함하는 배지에서 세포의 성장을 증가시키는 방법 및 세포 내로 오탄당을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 준비하여 탄화수소 또는 탄화수소 유도체를 제조하는 방법에 관한 것이다.

대표도 - 도25



(72) 발명자

비슨, 윌리엄, 티. 4세

미국 캘리포니아 94618 오클랜드 스트리트 아파트
먼트 #224 클리프턴 225

짜오, 후이민

미국 일리노이 61820 샴페인 스코츠데일 드라이브
2109

두, 징

미국 일리노이 61820 샴페인 아파트먼트 13 608 에
스. 스테이트 스트리트

최,진,호

미국 일리노이 61801 어바나 아파트먼트 22 미시간
애비뉴 이스트 504

도나 케이트, 제인, 에이치.

미국 캘리포니아 94705 버클리 빈센트 로드 164

갈라즈카, 조나단, 엠.

미국 캘리포니아 94705 버클리 아파트먼트 에이 디
킨 스트리트, 3018

하,숙,진

미국 일리노이 61820 샴페인 벨리 로드 1676

김,수,린

미국 일리노이 61801 어바나 아파트먼트 304 그릭
스 스트리트 웨스트 502

진,용-수

미국 일리노이 61822 샴페인 샌드힐 린 3005

리, 시진

미국 일리노이 61801 어배너 아파트먼트 324 멜로
즈 벨리 시티 1602

순, 지안펑

미국 캘리포니아 94709 버클리 아파트먼트 14 허스
트 애비뉴 2283

양, 샤오민

미국 캘리포니아 94706 알바니 아파트먼트 702 잭
슨 스트리트 1070

특허청구의 범위

청구항 1

하기 단계를 포함하는, 세포 내로 셀로덱스트린(cellodextrin) 수송(transport)을 증가시키는 방법:

- 숙주 세포(host cell)를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인(transmembrane) α-헬릭스 1(α-helix 1), α-헬릭스 2, α-헬릭스 3, α-헬릭스 4, α-헬릭스 5, α-헬릭스 6, α-헬릭스 7, α-헬릭스 8, α-헬릭스 9, α-헬릭스 10, α-헬릭스 11, α-헬릭스 12 를 포함하는 폴리펩티드(polypeptide)를 코딩하는(encoding) 재조합 폴리뉴클레오티드(recombinant polynucleotide)를 포함하고, 트랜스멤브레인 α-헬릭스 1 은 SEQ ID NO : 1 을 포함한다); 및

- 재조합 폴리뉴클레오티드가 발현되는 배지(medium)에 세포를 배양하는 단계

(재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 셀로덱스트린의 수송이 증가되었다).

청구항 2

하기 단계를 포함하는, 세포 내로 셀로덱스트린의 수송을 증가시키는 방법:

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α-헬릭스 1, α-헬릭스 2, α-헬릭스 3, α-헬릭스 4, α-헬릭스 5, α-헬릭스 6, α-헬릭스 7, α-헬릭스 8, α-헬릭스 9, α-헬릭스 10, α-헬릭스 11, α-헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α-헬릭스 2 은 SEQ ID NO : 2 을 포함한다); 및

- 재조합 폴리뉴클레오티드가 발현되는 배지에 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 재조합 폴리뉴클레오티드의 발현으로 세포 내로 셀로덱스트린의 수송이 증가되었다).

청구항 3

하기 단계를 포함하는, 세포 내로 셀로덱스트린 수송을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α-헬릭스 1, α-헬릭스 2, α-헬릭스 3, α-헬릭스 4, α-헬릭스 5, α-헬릭스 6, α-헬릭스 7, α-헬릭스 8, α-헬릭스 9, α-헬릭스 10, α-헬릭스 11, α-헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α-헬릭스 2 및 트랜스멤브레인 α-헬릭스 3 을 연결하는 루프(loop)는 SEQ ID NO : 3 을 포함한다); 및

- 재조합 폴리뉴클레오티드가 발현되는 배지에 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 셀로덱스트린의 수송이 증가되었다).

청구항 4

하기 단계를 포함하는, 세포 내로 셀로덱스트린 수송을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO : 4 를 포함한다) ; 및

- 재조합 폴리뉴클레오티드가 발현되는 배지에 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 셀로텍스트린의 수송이 증가되었다).

청구항 5

하기 단계를 포함하는, 세포 내로 셀로텍스트린 수송을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 6 는 SEQ ID NO : 5 를 포함한다) ; 및

- 재조합 폴리뉴클레오티드가 발현되는 배지에 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 셀로텍스트린의 수송이 증가되었다).

청구항 6

하기 단계를 포함하는, 세포 내로 셀로텍스트린 수송을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 6 및 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO : 6 을 포함한다) ; 및

- 재조합 폴리뉴클레오티드가 발현되는 배지에 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 셀로텍스트린의 수송이 증가되었다).

청구항 7

하기 단계를 포함하는, 세포 내로 셀로텍스트린 수송을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO : 7 을 포함한다) ; 및

- 재조합 폴리뉴클레오티드가 발현되는 배지에 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 셀로텍스트린의 수송이 증가되었다).

청구항 8

하기 단계를 포함하는, 세포 내로 셀로렉스트린 수송을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 10, 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO : 8 을 포함한다) ; 및

- 재조합 폴리뉴클레오티드가 발현되는 배지에 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 셀로렉스트린의 수송이 증가되었다).

청구항 9

제 1 항 내지 제 8 항 중 어느 한 항에 있어서,

상기 폴리펩티드는 NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 % 또는 적어도 100 % 의 아미노산 동일성(amino acid identity)을 갖는 방법.

청구항 10

제 1 항 내지 제 9 항 중 어느 한 항에 있어서,

상기 숙주 세포는 β -글루코시다아제(β -glucosidase)의 적어도 촉매 도메인(catalytic domain)을 코딩하는 두 번째 재조합 폴리뉴클레오티드를 더 포함하는 방법.

청구항 11

제 10 항에 있어서,

상기 β -글루코시다아제는 *네우로스포라 크라사*(*Neurospora crassa*)로부터 온 것인 방법.

청구항 12

제 11 항에 있어서,

상기 β -글루코시다아제는 NCU00130 에 의해 코드되는 방법.

청구항 13

제 1 항 내지 제 12 항 중 어느 한 항에 있어서,

상기 숙주 세포는 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 더 포함하고, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오탄당 이용(pentose utilization)에 포함되는 하나 또는 그 이상의 효소를 코드하는 방법.

청구항 14

제 13 항에 있어서,

상기 하나 또는 그 이상의 효소는, L-아라비노스 이성질화효소(L-arabinose isomerase), L-리블로인산화효소(L-ribulokinase), L-리블로스-5-P 4 에피머화효소(L-ribulose-5-P 4 epimerase), 자일로스 이성질화효소(xylose isomerase), 자일루로키나아제(xylulokinase), 알도스 환원효소(aldose reductase), L-아라비니톨 4-탈수소효소(L-arabinitol 4-dehydrogenase), L-자일루로스 환원효소(L-xylulose reductase) 및 자일리톨 탈수소효소로 이루어진 하나 또는 그 이상의 군으로부터 선택된 것인 방법.

청구항 15

제 1 항 내지 제 14 항 중 어느 한 항에 있어서,

상기 숙주 세포는 세 번째 재조합 폴리뉴클레오티드를 더 포함하고, 상기 세 번째 재조합 폴리뉴클레오티드는 오탄당 운반체(pentose transporter)를 코딩하는 방법.

청구항 16

제 15 항에 있어서,

상기 오탄당 운반체는 NCU00821, NCU04963, NCU06138, STL12/XUT6, SUT2, SUT3, XUT1 및 XUT3 으로 이루어진 군으로부터 선택된 방법.

청구항 17

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO : 1 을 포함하고, 상기 폴리펩티드는 셀로텍스트린 운반체이다) ; 및

- 셀로텍스트린을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하였다).

청구항 18

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO : 2 를 포함하고, 상기 폴리펩티드는 셀로텍스트린 운반체이다) ; 및

- 셀로텍스트린을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하였다).

청구항 19

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 2 및 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프는 SEQ ID NO : 3 을 포함하고, 상기 폴리펩티드는 셀로텍스트린 운반체이다) ; 및

- 셀로텍스트린을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하였다).

청구항 20

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO : 4 를 포함하고, 상기 폴리펩티드는 셀로텍스트린 운반체이다) ; 및

- 셀로텍스트린을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하였다).

청구항 21

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO : 5 를 포함하고, 상기 폴리펩티드는 셀로텍스트린 운반체이다) ; 및

- 셀로텍스트린을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하였다).

청구항 22

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 6 및 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO : 6 을 포함하고, 상기 폴리펩티드는 셀로텍스트린 운반체이다) ; 및

- 셀로텍스트린을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하였다).

청구항 23

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO : 7 을 포함하고, 상기 폴리펩티드는 셀로텍스트린 운반체이다) ; 및

- 셀로텍스트린을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하였다).

청구항 24

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 10 및 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO : 8 을 포함하고, 상기 폴리펩티드는 셀로텍스트린 운반체이다) ; 및

- 셀로텍스트린을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하였다).

청구항 25

제 17 항 내지 제 24 항 중 어느 한 항에 있어서,

상기 폴리펩티드는, NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 % 또는 적어도 100 % 의 아미노산 동일성을 갖는 방법.

청구항 26

제 17 항 내지 제 25 항 중 어느 한 항에 있어서,

상기 숙주 세포는 내인성(endogenous) 또는 두 번째 재조합 폴리뉴클레오티드를 더 포함하고, 상기 폴리뉴클레오티드는 β -글루코시다아제의 적어도 촉매 도메인(catalytic domain)을 코딩하는 방법.

청구항 27

제 26 항에 있어서,

상기 β -글루코시다아제는 *네우로스포라 크라사*(*Neurospora crassa*)로부터 온 것인 방법.

청구항 28

제 27 항에 있어서,

상기 β -글루코시다아제는 NCU00130 에 의해 코드되는 방법.

청구항 29

하기 단계를 포함하는, 셀룰로스-유도(cellulose-derived) 및 헤미셀룰로스-유도 당(hemicelluloses-derived sugars)을 공동 발효시키는(co-fermenting) 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 셀로렉스트린 운반체를 코딩하는 첫 번째 재조합 폴리뉴클레오티드 및 β -글루코시다아제의 촉매 도메인을 코딩하는 두 번째 재조합 폴리뉴클레오티드를 포함한다) ; 및

- 셀룰로스-유도 당 및 헤미셀룰로스-유도 당을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 재조합 폴리뉴클레오티드의 발현은 셀룰로스-유도 당 및 헤미셀룰로스-유도 당을 공동 발효시킬 수 있다).

청구항 30

제 29 항에 있어서,

상기 첫 번째 재조합 폴리뉴클레오티드는, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO : 1 을 포함하는 방법.

청구항 31

제 29 항에 있어서,

상기 첫 번째 재조합 폴리뉴클레오티드는, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO : 2 를 포함하는 방법.

청구항 32

제 29 항에 있어서,

상기 첫 번째 재조합 폴리뉴클레오티드는, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α -헬릭스 2 및 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프는 SEQ ID NO : 3 을 포함하는 방법.

청구항 33

제 29 항에 있어서,

상기 첫 번째 재조합 폴리뉴클레오티드는, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO : 4 를 포함하는 방법.

청구항 34

제 29 항에 있어서,

상기 첫 번째 재조합 폴리뉴클레오티드는, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO : 5 를 포함하는 방법.

청구항 35

제 29 항에 있어서,

상기 첫 번째 재조합 폴리뉴클레오티드는, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α -헬릭스 6 및 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO : 6 을 포함하는 방법.

청구항 36

제 29 항에 있어서,

상기 첫 번째 재조합 폴리뉴클레오티드는, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO : 7 을 포함하는 방법.

청구항 37

제 29 항에 있어서,

상기 첫 번째 재조합 폴리뉴클레오티드는, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α -헬릭스 10, 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO : 8 을 포함하는 방법.

청구항 38

제 30 항 내지 제 37 항 중 어느 한 항에 있어서,

상기 폴리펩티드는 NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 % 또는 적어도 100 % 의 아미노산 동일성을 갖는 방법.

청구항 39

제 29 항 내지 제 38 항 중 어느 한 항에 있어서,

상기 β -글루코시다아제는 *네투로스포라 크라사(Neurospora crassa)*로부터 온 것인 방법.

청구항 40

제 39 항에 있어서,

상기 β -글루코시다아제는 NCU00130 에 의해 코드되는 방법.

청구항 41

제 29 항 내지 제 40 항 중 어느 한 항에 있어서,

상기 숙주 세포는 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 더 포함하고, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용(pentose utilization)에 포함되는 하나 또는 그 이상의 효소를 코드하는 방법.

청구항 42

제 41 항에 있어서,

상기 하나 또는 그 이상의 효소는, L-아라비노스 이성질화효소, L-리블로인산화효소, L-리블로스-5-P 4 에피머화효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비톨 4-탈수소효소, L-자일루로스 환원효소 및 자일리톨 탈수소효소로 이루어진 하나 또는 그 이상의 군으로부터 선택된 것인 방법.

청구항 43

제 29 항 내지 제 42 항 중 어느 한 항에 있어서,

상기 숙주 세포는 세 번째 재조합 폴리뉴클레오티드를 더 포함하고, 상기 폴리뉴클레오티드는 오타당 운반체(pentose transporter)를 코드하는 방법.

청구항 44

제 43 항에 있어서,

상기 오타당 운반체는 NCU00821, NCU04963, NCU06138, STL12/XUT6, SUT2, SUT3, XUT1 및 XUT3 으로 이루어진 군으로부터 선택된 방법.

청구항 45

제 29 항 내지 제 44 항 중 어느 한 항에 있어서,

상기 셀룰로스-유도 당은 셀로비오스(cellobiose), 셀로트리오스(cellobiose) 및 셀로테트라오스(cellobiose)로 이루어진 군으로부터 선택된 것이고, 헤미셀룰로스-유도 당은 자일로스인 방법.

청구항 46

하기 단계를 포함하는, 숙주 세포에 의한 탄화수소(hydrocarbon) 또는 탄화수소 유도체(hydrocarbon derivatives)의 합성을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO : 1 을 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내의 셀로텍스트린을 수송한다) ; 및

- 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원(source)을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 세포 내로 셀로텍스트린의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 47

하기 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO : 2 를 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송한다) ; 및

- 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 세포 내로 셀로텍스트린의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 48

하기 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 2 및 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프는 SEQ ID NO : 3 을 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송한다) ; 및

- 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 세포 내로 셀로텍스트린의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 49

하기 단계를 포함하는, 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO : 4 를 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송한다) ; 및

- 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 세포 내로 셀로텍스트린의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 50

하기 단계를 포함하는, 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO : 5 를 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송한다) ; 및

- 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 세포 내로 셀로텍스트린의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 51

하기 단계를 포함하는, 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 6 및 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO : 6 을 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송한다) ; 및

- 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 세포 내로 셀로텍스트린의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 52

하기 단계를 포함하는, 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO : 7 을 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송한다) ; 및

- 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 세포 내로 셀로텍스트린의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 53

하기 단계를 포함하는, 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 10 및 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO : 8 을 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송한다) ; 및

- 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계

(상기 세포 내로의 셀로텍스트린의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 54

제 46 항 내지 제 53 항 중 어느 한 항에 있어서,

상기 폴리펩티드는, NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 % 또는 적어도 100 % 의 아미노산 동일성을 갖는 방법.

청구항 55

제 46 항 내지 제 54 항 중 어느 한 항에 있어서,

상기 숙주 세포는 두 번째 재조합 폴리뉴클레오티드를 더 포함하고, 상기 폴리뉴클레오티드는 β -글루코시다아제(β -glucosidase)의 적어도 촉매 도메인을 코드하는 방법.

청구항 56

제 55 항에 있어서,

상기 β -글루코시다아제는 *네우로스포라 크라사*(*Neurospora crassa*)로부터 온 것인 방법.

청구항 57

제 56 항에 있어서,

상기 β -글루코시다아제는 NCU00130 에 의해 코드되는 방법.

청구항 58

제 46 항 내지 제 57 항 중 어느 한 항에 있어서,

상기 셀로텍스트린의 근원은 셀룰로스를 포함하는 방법.

청구항 59

제 46 항 내지 제 58 항 중 어느 한 항에 있어서,

상기 탄화수소 또는 탄화수소 유도체는 연료(fuel)로서 사용될 수 있는 방법.

청구항 60

제 59 항에 있어서,

상기 탄화수소 또는 탄화수소 유도체는 에탄올을 포함하는 방법.

청구항 61

제 59 항에 있어서,

상기 탄화수소 또는 탄화수소 유도체는 부탄올을 포함하는 방법.

청구항 62

제 1 항 내지 제 61 항 중 어느 한 항에 있어서,

상기 배지는 변경된 유기체(altered organism)로부터 셀룰라아제-함유 효소 혼합물을 포함하고, 변경되지 않은 유기체(unaltered organism)로부터 셀룰라아제-함유 혼합물과 비교하여, 상기 셀룰라아제-함유 혼합물은 감소된 β -글루코시다아제 활성도(β -glucosidase activity)를 갖는 방법.

청구항 63

제 1 항 내지 제 62 항 중 어느 한 항에 있어서,

상기 숙주 세포는, 사카로마이스(*Saccharomyces* sp.), 사카로미세스 세레비시아(*Saccharomyces cerevisiae*), 사카로미세스 모나센시스(*Saccharomyces monacensis*), 사카로미세스 바야누스(*Saccharomyces bayanus*), 사카로미세스 파스토리아누스(*Saccharomyces pastorianus*), 사카로미세스 칼스베르겐시스(*Saccharomyces carlsbergensis*), 사카로미세스 폼베(*Saccharomyces pombe*), 클루이베르마이세스(*Kluyveromyces* sp.), 클루이베르마이세스 마르시아무스(*Kluyveromyces marxianus*), 클루이베르마이세스 락티스(*Kluyveromyces lactis*), 클루이베르마이세스 프라길리스(*Kluyveromyces fragilis*), 피치아 스티피티스(*Pichia stipitis*), 스포로트리쿰 서머파일(*Sporotrichum thermophile*), 칸디다 쉬하테(*Candida shehatae*), 칸디다 트로피칼리스(*Candida tropicalis*), 네우로스포라 크라사(*Neurospora crassa*), 자이모나스 모빌리스(*Zymomonas mobilis*), 클로스트리듐(*Clostridium* sp.), 클로스트리듐 파이토퍼멘탄스(*Clostridium phytofermentans*), 클로스트리듐 써모셀럼(*Clostridium thermocellum*), 클로스트리듐 베이저린키이(*Clostridium beijerinckii*), 클로스트리듐 아세토부틸리쿰(*Clostridium acetobutylicum*), 무렐라 서모아세티카(*Moorella thermoacetica*), 대장균(*Escherichia coli*), 크레브시엘라 옥시토카(*Klebsiella oxytoca*), 서모안아에로박테리움 사카로라티쿰(*Thermoanaerobacterium saccharolyticum*) 및 고초균(*Bacillus subtilis*)으로 이루어진 군으로부터 선택한 방법.

청구항 64

제 1 항 내지 제 63 항 어느 한 항에 있어서,

상기 셀로텍스트린은, 셀로비오스(cellobiose), 셀로트리오스(cellotriose) 및 셀로테트라오스(cellotetraose)로 이루어진 하나 또는 그 이상의 군으로부터 선택된 방법.

청구항 65

NCU00821 및 STL12/XUT6 로 이루어진 군으로부터 선택한 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포로서, 상기 폴리펩티드는 세포 내로 자일로스를 수송하는 숙주 세포.

청구항 66

XUT1 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포로서, 상기 폴리펩티드는 세포 내로 아라비노스(arabinose)를 수송하는 숙주 세포.

청구항 67

NCU06138 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포로서, 상기 폴리펩티드는 세포 내로 아라비노스 및 글루코스를 수송하는 숙주 세포.

청구항 68

SUT2, SUT3 및 XUT3 으로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포로서, 상기 폴리펩티드는 세포 내로 자일로스와 글루코스를 수송하는 숙주 세포.

청구항 69

NCU04963 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포로서, 상기 폴리펩티드는 세포 내로 자일로스, 아라비노스 및 글루코스를 수송하는 숙주 세포.

청구항 70

제 65 항 내지 제 69 항 중 어느 한 항에 있어서,

하나 또는 그 이상의 재조합 폴리뉴클레오티드를 더 포함하고, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코드하는 숙주 세포.

청구항 71

제 70 항에 있어서,

상기 하나 또는 그 이상의 효소는, L-아라비노스 이성질화효소, L-리블로인산화효소, L-리블로스-5-P 4 에피머화효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비니톨 4-탈수소효소, L-자일루로스 환원효소 및 자일리톨 탈수소효소로 이루어진 하나 또는 그 이상의 군으로부터 선택된 것인 방법.

청구항 72

제 65 항 내지 제 71 항 중 어느 한 항에 있어서,

상기 숙주 세포는, 사카로마이시스(*Saccharomyces* sp.), 사카로미세스 세레비시아(*Saccharomyces cerevisiae*), 사카로미세스 모나센시스(*Saccharomyces monacensis*), 사카로미세스 바야누스(*Saccharomyces bayanus*), 사카로미세스 파스토리아누스(*Saccharomyces pastorianus*), 사카로미세스 칼스베르겐시스(*Saccharomyces carlsbergensis*), 사카로미세스 폼베(*Saccharomyces pombe*), 클루이베르마이세스(*Kluyveromyces* sp.), 클루이베르마이세스 마르시아우스(*Kluyveromyces marxianus*), 클루이베르마이세스 락티스(*Kluyveromyces lactis*), 클루이베르마이세스 프라길리스(*Kluyveromyces fragilis*), 피치아 스티피티스(*Pichia stipitis*), 스포로트리쿰 서머파일(*Sporotrichum thermophile*), 칸디다 쉬하테(*Candida shehatae*), 칸디다 트로피칼리스(*Candida tropicalis*), 네우로스포라 크라사(*Neurospora crassa*), 자이모나스 모빌리스(*Zymomonas mobilis*), 클로스트리듐(*Clostridium* sp.), 클로스트리듐 파이토퍼멘탄스(*Clostridium phytofermentans*), 클로스트리듐 써모셀럼(*Clostridium thermocellum*), 클로스트리듐 베이저린키이(*Clostridium beijerinckii*), 클로스트리듐 아세토부

틸리쿰(*Clostridium acetobutylicum*), 무렐라 서모아세티카(*Moorella thermoacetica*), 대장균(*Escherichia coli*), 크레브시엘라 옥시토카(*Klebsiella oxytoca*), 서모안아에로박테리움 사카로락티쿰(*Thermoanaerobacterium saccharolyticum*) 및 고초균(*Bacillus subtilis*)으로 이루어진 군으로부터 선택된 방법.

청구항 73

하기 단계를 포함하는, 세포 내로 자일로스의 수송을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는, NCU00821 및 STL12/XUT6 로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함한다) ; 및

- 재조합 폴리뉴클레오티드가 발현되는 세포를 배양하는 단계

(재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 자일로스의 수송이 증가되었다).

청구항 74

하기 단계를 포함하는, 세포 내로 아라비노스의 수송을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 XUT1 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함한다) ; 및

- 재조합 폴리뉴클레오티드가 발현되는 세포를 배양하는 단계

(재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 아라비노스의 수송이 증가되었다).

청구항 75

하기 단계를 포함하는, 세포 내로 아라비노스 또는 글루코스의 수송을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 NCU06138 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함한다) ; 및

- 재조합 폴리뉴클레오티드가 발현되는 세포를 배양하는 단계

(재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 아라비노스 또는 글루코스의 수송이 증가되었다).

청구항 76

하기 단계를 포함하는, 세포 내로 자일로스 또는 글루코스의 수송을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 SUT2, SUT3 및 XUT3 으로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함한다) ; 및

- 재조합 폴리뉴클레오티드가 발현되는 세포를 배양하는 단계

(재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 자일로스 또는 글루코스의 수송이 증가되었다).

청구항 77

하기 단계를 포함하는, 세포 내로 자일로스, 아라비노스 또는 글루코스의 수송을 증가시키는 방법 :

- 숙주 세포를 준비하는 단계

(상기 숙주 세포는 NCU04963 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함한다) ; 및

- 재조합 폴리뉴클레오티드가 발현되는 세포를 배양하는 단계

(재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 자일로스, 아라비노스 또는 글루코스의 수송이 증가되었다).

청구항 78

제 73 항 내지 제 77 항 중 어느 한 항에 있어서,

하나 또는 그 이상의 재조합 폴리뉴클레오티드를 더 포함하고, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용에서 포함되는 하나 또는 그 이상의 효소를 코딩하는 방법.

청구항 79

제 78 항에 있어서,

상기 하나 또는 그 이상의 효소는 L-아라비노스 이성질화효소, L-리블로인산화효소, L-리블로스-5-P 4 에피머화 효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비니톨 4-탈수소효소, L-자일루로스 환원효소 및 자일리톨 탈수소효소로 이루어진 하나 또는 그 이상 군으로부터 선택된 것인 방법.

청구항 80

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리뉴클레오티드는 NCU00821 및 STL12/XUT6 로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하고, 상기 폴리펩티드는 자일로스를 수송한다) ; 및

- 자일로스를 포함하는 배지에서 숙주 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 상기 숙주 세포는 배지에서 더 빠른 비율로 성장 하였다).

청구항 81

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리뉴클레오티드는 XUT1 폴리펩티드를 코딩하고, 상기 폴리펩티드는 아라비노스를 수송한다) ; 및

- 아라비노스를 포함하는 배지에서 숙주 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 상기 숙주 세포는 배지에서 더 빠른 비율로 성장 하였다).

청구항 82

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리뉴클레오티드는 NCU06138 폴리펩티드를 코드하고, 상기 폴리펩티드는 아라비노스 및 글루코스를 수송한다) ; 및

- 아라비노스 또는 글루코스를 포함하는 배지에서 숙주 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 상기 숙주 세포는 배지에서 더 빠른 비율로 성장하였다).

청구항 83

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리뉴클레오티드는 SUT2, SUT3 및 XUT3 으로 이루어진 군으로부터 선택된 폴리펩티드를 코드하고, 상기 폴리펩티드는 자일로스 및 글루코스를 수송한다) ; 및

- 자일로스 또는 글루코스를 포함하는 배지에서 숙주 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 상기 숙주 세포는 배지에서 더 빠른 비율로 성장하였다).

청구항 84

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리뉴클레오티드는 NCU04963 폴리펩티드를 코드하고, 상기 폴리펩티드는 자일로스, 아라비노스 및 글루코스를 수송한다) ; 및

- 자일로스, 아라비노스 또는 글루코스를 포함하는 배지에서 숙주 세포를 배양하는 단계

(재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여, 상기 숙주 세포는 배지에서 더 빠른 비율로 성장하였다).

청구항 85

제 80 항 내지 제 84 항 중 어느 한 항에 있어서,

상기 숙주 세포는, 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코딩하는 하나 또는 그 이상의 내인성(endogenous) 또는 재조합 폴리뉴클레오티드를 더 포함하는 방법.

청구항 86

제 85 항에 있어서,

상기 하나 또는 그 이상의 효소는, L-아라비노스 이성질화효소, L-리블로인산화효소, L-리블로스-5-P 4 에피머화효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비니톨 4-탈수소효소, L-자일루로

스 환원효소 및 자일리톨 탈수소효소로 이루어진 하나 또는 그 이상의 군으로부터 선택된 것인 방법.

청구항 87

하기 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체(hydrocarbon derivatives)의 합성을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는, NCU00821 및 STL12/XUT6 로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 자일로스를 수송한다) ; 및

- 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 자일로스 또는 자일로스 근원을 포함하는 배지에 숙주 세포를 배양하는 단계

(세포 내로 자일로스 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 88

하기 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는, XUT1 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 아라비노스를 수송한다) ; 및

- 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 아라비노스 또는 아라비노스 근원을 포함하는 배지에 숙주 세포를 배양하는 단계

(세포 내로 아라비노스 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 89

하기 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는 NCU06138 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 아라비노스 또는 글루코스를 수송한다) ; 및

- 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 아라비노스, 글루코스, 아라비노스 근원 또는 글루코스 근원을 포함하는 배지에 숙주 세포를 배양하는 단계

(세포 내로 아라비노스 또는 글루코스 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 90

하기 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는, SUT2, SUT3, 및 XUT3 으로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 자일로스 또는 글루코스를 수송한다) ; 및

- 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 자일로스, 글루코스, 자일로스

근원 또는 글루코스 근원을 포함하는 배지에 숙주 세포를 배양하는 단계
(세포 내로 자일로스 또는 글루코스 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 91

하기 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는, NCU04963 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 자일로스, 아라비노스 또는 글루코스를 수송한다) ; 및

- 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 자일로스, 아라비노스, 글루코스, 자일로스 근원, 아라비노스 근원 또는 글루코스 근원을 포함하는 배지에 숙주 세포를 배양하는 단계
(세포 내로 자일로스, 아라비노스 또는 글루코스 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다).

청구항 92

제 89 항 내지 제 91 항 중 어느 한 항에 있어서,
상기 글루코스의 근원은 셀룰로스를 포함하는 방법.

청구항 93

제 87 항 내지 제 92 항 중 어느 한 항에 있어서,
자일로스 또는 아라비노스의 근원은 헤미셀룰로스를 포함하는 방법.

청구항 94

제 87 항 내지 제 93 항 중 어느 한 항에 있어서,
상기 탄화수소 또는 탄화수소 유도체는 연료(fuel)로서 사용될 수 있는 방법.

청구항 95

제 94 항에 있어서,
상기 탄화수소 또는 탄화수소 유도체는 에탄올을 포함하는 방법.

청구항 96

제 81 항에 있어서,
상기 탄화수소 또는 탄화수소 유도체는 부탄올을 포함하는 방법.

청구항 97

제 73 항 내지 제 95 항 중 어느 한 항에 있어서,

상기 숙주 세포는, 사카로마이시스(*Saccharomyces* sp.), 사카로미세스 세레비시아(*Saccharomyces cerevisiae*), 사카로미세스 모나센시스(*Saccharomyces monacensis*), 사카로미세스 바야누스(*Saccharomyces bayanus*), 사카로미세스 파스토리아누스(*Saccharomyces pastorianus*), 사카로미세스 칼스베르겐시스(*Saccharomyces carlsbergensis*), 사카로미세스 폼베(*Saccharomyces pombe*), 클루이베르마이세스(*Kluyveromyces* sp.), 클루이베르마이세스 마르시아무스(*Kluyveromyces marxianus*), 클루이베르마이세스 락티스(*Kluyveromyces lactis*), 클루이베르마이세스 프라길리스(*Kluyveromyces fragilis*), 피치아 스티피티스(*Pichia stipitis*), 스포로트리쿰 서머파일(*Sporotrichum thermophile*), 칸디다 쉬하테(*Candida shehatae*), 칸디다 트웨피칼리스(*Candida tropicalis*), 네우로스포라 크라사(*Neurospora crassa*), 자이모나스 모빌리스(*Zymomonas mobilis*), 클로스트리듐(*Clostridium* sp.), 클로스트리듐 파이토퍼멘탄스(*Clostridium phytofermentans*), 클로스트리듐 써모셀럼(*Clostridium thermocellum*), 클로스트리듐 베이저린키이(*Clostridium beijerinckii*), 클로스트리듐 아세토부틸리쿰(*Clostridium acetobutylicum*), 무렐라 서모아세티카(*Moorella thermoacetica*), 대장균(*Escherichia coli*), 크레브시엘라 옥시토카(*Klebsiella oxytoca*), 서모안아에로박테리움 사카웰락티쿰(*Thermoanaerobacterium sacchawlyticum*) 및 고초균(*Bacillus subtilis*)으로 이루어진 군으로부터 선택된 방법.

청구항 98

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리뉴클레오티드는 NCU07705 폴리펩티드를 코딩한다) ; 및

- 셀룰로스를 포함하는 배지에 세포를 배양하는 단계

(재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 상기 숙주 세포는 배지에서 더 빠른 비율로 성장하였다).

청구항 99

제 98 항에 있어서,

상기 숙주 세포는, 사카로마이시스(*Saccharomyces* sp.), 사카로미세스 세레비시아(*Saccharomyces cerevisiae*), 사카로미세스 모나센시스(*Saccharomyces monacensis*), 사카로미세스 바야누스(*Saccharomyces bayanus*), 사카로미세스 파스토리아누스(*Saccharomyces pastorianus*), 사카로미세스 칼스베르겐시스(*Saccharomyces carlsbergensis*), 사카로미세스 폼베(*Saccharomyces pombe*), 클루이베르마이세스(*Kluyveromyces* sp.), 클루이베르마이세스 마르시아무스(*Kluyveromyces marxianus*), 클루이베르마이세스 락티스(*Kluyveromyces lactis*), 클루이베르마이세스 프라길리스(*Kluyveromyces fragilis*), 피치아 스티피티스(*Pichia stipitis*), 스포로트리쿰 서머파일(*Sporotrichum thermophile*), 칸디다 쉬하테(*Candida shehatae*), 칸디다 트웨피칼리스(*Candida tropicalis*), 네우로스포라 크라사(*Neurospora crassa*), 자이모나스 모빌리스(*Zymomonas mobilis*), 클로스트리듐(*Clostridium* sp.), 클로스트리듐 파이토퍼멘탄스(*Clostridium phytofermentans*), 클로스트리듐 써모셀럼(*Clostridium thermocellum*), 클로스트리듐 베이저린키이(*Clostridium beijerinckii*), 클로스트리듐 아세토부틸리쿰(*Clostridium acetobutylicum*), 무렐라 서모아세티카(*Moorella thermoacetica*), 대장균(*Escherichia coli*), 크레브시엘라 옥시토카(*Klebsiella oxytoca*), 서모안아에로박테리움 사카웰락티쿰(*Thermoanaerobacterium sacchawlyticum*) 및 고초균(*Bacillus subtilis*)으로 이루어진 군으로부터 선택된 방법.

청구항 100

제 98 항에 있어서,

상기 숙주 세포는, 재조합 폴리뉴클레오티드에 작동가능하게 연결된(operably linked) 유도성 프로모터

(inducible promoter)를 더 포함하는 방법.

청구항 101

제 98 항에 있어서,

셀룰라아제(cellulases)의 발현은 재조합 폴리뉴클레오티드의 발현 상에서 숙주 세포에서 증가되는 방법.

청구항 102

하기 단계를 포함하는, 바이오매스 중합체(biomass polymer)에서 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는 내인성 폴리뉴클레오티드를 포함하고, 상기 폴리뉴클레오티드는 NCU05137 폴리펩티드를 코드한다) ;

- 내인성 폴리뉴클레오티드의 발현을 저해하는 단계 ; 및

- 바이오매스 중합체를 포함하는 배지에서 세포를 배양하는 단계

(내인성 폴리뉴클레오티드의 발현을 저해하지 않는 세포보다, 상기 숙주 세포는 배지에서 더 빠른 비율로 성장하였다).

청구항 103

제 102 항에 있어서,

상기 숙주 세포는, 사카로마이시스(*Saccharomyces sp.*), 사카로미세스 세레비시아(*Saccharomyces cerevisiae*), 사카로미세스 모나센시스(*Saccharomyces monacensis*), 사카로미세스 바야누스(*Saccharomyces bayanus*), 사카로미세스 파스토리아누스(*Saccharomyces pastorianus*), 사카로미세스 칼스베르겐시스(*Saccharomyces carlsbergensis*), 사카로미세스 폼베(*Saccharomyces pombe*), 클루이베르마이세스(*Kluyveromyces sp.*), 클루이베르마이세스 마르시아무스(*Kluyveromyces marxianus*), 클루이베르마이세스 락티스(*Kluyveromyces lactis*), 클루이베르마이세스 프라길리스(*Kluyveromyces fragilis*), 피치아 스티피티스(*Pichia stipitis*), 스포로트리쿰 서머파일(*Sporotrichum thermophile*), 칸디다 쉬하테(*Candida shehatae*), 칸디다 트웨피칼리스(*Candida tropicalis*), 네우로스포라 크라사(*Neurospora crassa*), 자이모나스 모빌리스(*Zymomonas mobilis*), 클로스트리듐(*Clostridium sp.*), 클로스트리듐 파이토퍼멘탄스(*Clostridium phytofermentans*), 클로스트리듐 써모셀럼(*Clostridium thermocellum*), 클로스트리듐 베이저린키이(*Clostridium beijerinckii*), 클로스트리듐 아세토부틸리쿰(*Clostridium acetobutylicum*), 무렐라 서모아세티카(*Moorella thermoacetica*), 대장균(*Escherichia coli*), 크레브시엘라 옥시토카(*Klebsiella oxytoca*), 서모아나에로박테리움 사카로락티쿰(*Thermoanaerobacterium saccharolyticum*) 및 고초균(*Bacillus subtilis*)으로 이루어진 군으로부터 선택된 방법.

청구항 104

제 102 항에 있어서,

상기 숙주 세포의 셀룰라아제 활성도(cellulose activity)는 내인성 폴리뉴클레오티드 발현의 저해(inhibiting) 상에서 증가되는 방법.

청구항 105

제 102 항에 있어서,

상기 숙주 세포의 헤미셀룰라아제 활성도(hemicellulase activity)는 내인성 폴리뉴클레오티드 발현의 저해 상에서 증가되는 방법.

청구항 106

제 102 항에 있어서,

내인성 폴리뉴클레오티드 발현의 저해는 내인성 폴리뉴클레오티드를 포함하는 유전자(gene)의 돌연변이(mutating) 또는 결실(deleting)을 포함하는 방법.

청구항 107

제 102 항에 있어서,

상기 바이오매스 중합체는 셀룰로스인 방법.

청구항 108

제 102 항에 있어서,

상기 바이오매스 중합체는 헤미셀룰로스인 방법.

청구항 109

하기 단계를 포함하는, 세포의 성장을 증가시키는 방법 :

- 숙주 세포를 제공하는 단계

(상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리뉴클레오티드는 NCU01517, NCU09133 및 NCU10040 으로 이루어진 군으로부터 선택된 폴리펩티드를 코드한다) ; 및

- 헤미셀룰로스를 포함하는 배지에 세포를 배양하는 단계

(재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 상기 숙주 세포는 배지에서 더 빠른 비율로 성장하였다).

청구항 110

제 109 항에 있어서,

상기 숙주 세포는, 사카로마이스(Saccharomyces sp.), 사카로미세스 세레비시아(Saccharomyces cerevisiae), 사카로미세스 모나센시스(Saccharomyces monacensis), 사카로미세스 바야누스(Saccharomyces bayanus), 사카로미세스 파스토리아누스(Saccharomyces pastorianus), 사카로미세스 칼스베르겐시스(Saccharomyces carlsbergensis), 사카로미세스 폼베(Saccharomyces pombe), 클루이베르마이세스(Kluyveromyces sp.), 클루이베르마이세스 마르시아무스(Kluyveromyces marxianus), 클루이베르마이세스 락티스(Kluyveromyces lactis), 클루이베르마이세스 프라길리스(Kluyveromyces fragilis), 피치아 스티피티스(Pichia stipitis), 스포로트리쿰 서머파일(Sporotrichum thermophile), 칸디다 쉬하테(Candida shehatae), 칸디다 트웨피칼리스(Candida tropicalis), 네우로스포라 크라사(Neurospora crassa), 자이모나스 모빌리스(Zymomonas mobilis), 클로스트리듐(Clostridium sp.), 클로스트리듐 파이토퍼멘탄스(Clostridium phytofermentans), 클로스트리듐 써모셀럼(Clostridium thermocellum), 클로스트리듐 베이저린키이(Clostridium beijerinckii), 클로스트리듐 아세토부틸리쿰(Clostridium acetobutylicum), 무렐라 서모아세티카(Moorella thermoacetica), 대장균(Escherichia coli), 크레브시엘라 옥시토카(Klebsiella oxytoca), 서모안아에로박테리움 사카로락티쿰

(*Thermoanaerobacterium saccharolyticum*) 및 고초균(*Bacillus subtilis*)으로 이루어진 균으로부터 선택된 방법.

청구항 111

제 109 항에 있어서,

상기 숙주 세포는, 재조합 폴리뉴클레오티드에 작동가능하게 연결된 유도성 프로모터(inducible promoter)를 더 포함하는 방법.

청구항 112

제 109 항에 있어서,

상기 숙주 세포의 헤미셀룰라아제 활성도는 재조합 폴리뉴클레오티드의 발현 상에서 증가되는 방법.

청구항 113

하기 단계를 포함하는, 셀룰로스를 분해하는(degrading) 방법 :

- 셀룰로스를 포함하는 조성물(composition)을 제공하는 단계 ; 및
- 변경된 유기체로부터 셀룰라아제-함유 효소 혼합물(cellulase-containing enzyme mixture)을 상기 조성물과 접촉시키는 단계

(변경되지 않는 유기체로부터 셀룰라아제-함유 혼합물을 비교하여, 상기 셀룰라아제-함유 혼합물은 감소된 β -글루코시다아제 활성도를 갖고, 상기 셀룰로스는 셀룰라아제-함유 혼합물에 의해 분해되었다).

청구항 114

제 113 항에 있어서,

상기 유기체는 β -글루코시다아제를 코딩하는 유전자의 돌연변이(mutation)에 의해 변형되는(altered) 방법.

청구항 115

제 113 항에 있어서,

상기 유기체는 β -글루코시다아제 발현의 감소에 의해 변형되는 방법.

청구항 116

제 113 항 내지 제 115 항 중 어느 한 항에 있어서,

상기 유기체는 균류(fungus) 및 세균(bacterium)으로 이루어진 균으로부터 선택된 방법.

청구항 117

제 116 항에 있어서,

상기 유기체는 사상균(filamentous fungus)인 방법.

청구항 118

제 113 항 내지 제 117 항 중 어느 한 항에 있어서,
상기 셀룰로스는 식물 물질(plant material)로부터 온 것인 방법.

청구항 119

제 118 항에 있어서,
상기 식물 물질은, 지팡이풀(switchgrass), *미스칸투스*(*Miscanthus*), 왕겨(rice hull), 버개스(bagasse), 아마(flax), 대나무(bamboo), 사이잘(sisal), 마닐라삼섬유(abaca), 짚(straw), 잎(leaves), 잔디 부스러기(grass clippings), 옥수수대(corn stover), 옥수수 속대(corn cobs), 증류립(distillers grains), 콩과 식물(legume plants), 수수(sorghum), 사탕수수(sugar cane), 사탕무 펄프(sugar beet pulp), 우드칩(wood chips), 톱밥(sawdust) 및 바이오매스 크롭(biomass crops)으로 이루어진 군으로부터 선택된 방법.

명세서

기술 분야

[0001] 관련된 출원서(RELATED APPLICATIONS)에 대한 참고문헌(CROSS-REFERENCE)

[0002] 본 출원서는, 2009년 12월 10일 미국 가출원 번호 61/285,526 및 2009년 7월 24일 출원된 미국 가출원 번호 61/271,833 의 내용을 청구하고, 이들 모두는 전체가 본원의 참고문헌으로 인용된다.

[0003] 본 발명의 분야

[0004] 본 발명은, 세포 내에 당의 수송(transport)을 증가시키기 위한, 세포의 성장을 증가시키기 위한, 탄화수소 및 탄화수소 유도체(hydrocarbon derivatives)의 합성을 증가시키기 위한, 셀룰로스-유도 당(cellulose-derived sugars) 및 헤미셀룰로스-유도 당(hemicellulose-derived sugars)을 공동 발효시키기 위한(co-fermenting) 방법 및 조성물에 관한 것이다.

배경 기술

[0005] 바이오 연료(biofuels)는, 에너지 안전 보장(energy security), 지속가능성(sustainability), 및 지구 기후 변화(global climate change)에 대한 증가된 우려(concerns)때문에 집중적인 조사 하에 있다(Lynd 등의 1991). 식물-유도 리그노셀룰로스를 물질(plant-derived lignocellulosic material)을 바이오 연료로의 생물학적 변환(bioconversion)은 화석연료(fossil fuel)의 화학적인 생산에 대한 흥미로운 대안으로서 간주되고 있다(Lynd 등의 2008 ; Hahn-Hagerdal 등의 2006). 리그노셀룰로스 바이오매스(lignocellulose biomass)는 셀룰로스(cellulose), 헤미셀룰로스(hemicellulose) 및 리그닌(lignin)으로 구성된다.

[0006] 리그노셀룰로스를 바이오매스를 에탄올로의 전환을 유효하게 실행하기 위해 미생물 공학은 바이오 연료 분야(biofuels field)의 주요한 목표로 여전히 남아있다. 대부분의 조사는, 셀룰라아제(cellulases) 및 그 밖의 효소의 발현에 의해 단순한 당(sugars)이 알코올로 자연적으로 발효되는 유전적으로 조작된 미생물에 집중되어 있고, 이들은 리그노셀룰로스를 바이오매스 중합체(lignocellulosic biomass polymers)를 분해하여 하나의 세포 내에서 에탄올을 발생시킨다. 그러나, 연구가 잘 되지 않는 영역은 당 운반체(sugar transport) 영역이다. 당-흡수 능력(sugar-uptake ability)을 개선시키기 위한 당 운반체의 조절 및 미생물의 유전공학의 이해는 매우 효율을 증진시킬 것이다(Stephanopoulos 2007). 더욱이, 셀룰라아제 발현 및 활성의 조절에서 포함되는 그 밖

의 유형의 단백질은 연구할 부분이 많이 남아있다.

[0007] 빵 효모(baker's yeast)로 또한 알려진 *사카로미세스 세레비시아*(*Saccharomyces cerevisiae*)는 수 천년 동안 6 탄당(hexose sugars)을 에탄올로 생물전환(bioconversion)에 사용된다. 이는 또한 D-글루코스에서 에탄올로의 대규모 발효 공업에 대해서 미생물을 매우 넓게 사용되고 있다. *에스. 세레비시아*(*S. cerevisiae*)는 리그노셀룰로스로부터 바이오매스를 바이오연료(biofuel)로의 생물전환의 매우 적절한 후보물질이다(van Maris 등의, 2006). 이는 유전적이고 생리학적인 배경기술, 충분한 유전적 도구 및 리그노셀룰로스로부터 가수분해(lignocellulosic hydrolysates)에서 나타낸 높은 에탄올 농도 및 저해제에 대한 높은 내성(tolerance)에 대해서는 잘 연구되어 있다(Jeffries 2006). *에스. 세레비시아*의 낮은 발효 pH 는 발효하는 동안에 세균 오염(bacterial contamination)에 또한 제시될 수 있다.

[0008] 유감스럽게도, 야생형 *에스. 세레비시아*는 오타당 당(pentose sugars)에서 사용할 수 없다(Hector 등의, 2008). 이러한 한계를 극복하기 위해, 오타당-소화하는 유기체(pentose-assimilating organisms)를 *에스. 세레비시아*에 도입하였고, 이는 D-자일로스(D-xylose) 및 L-아라비노스(L-arabinose)의 발효를 가능하게 한다(Hahn-Hagerdal 등의, 2007; Brat 등의, 2009; Wisselink 등의, 2007, 2009; Wiedemann and Boles 2008; Karhumma 등의, 2006). 그러나, 오타당을 바이오연료로의 효율적인 전환은, 세포내 산화환원 반응의 불균형(cellular redox imbalance), 5 탄당 인산경로(pentose phosphate pathway)의 낮은 유입(low influx) 및 세포 내로의 효율적인 5 탄당 운송의 결핍을 포함하는 다양한 문제에 의해 제한되고 있다(Hector 등의, 2008).

[0009] 게다가, 자연적인 미생물 및 유전자 조작된 미생물 모두는, 글루코스 발효와 비교하여 자일로스 발효하는 동안에 감소된 에탄올 내성(ethanol tolerance)을 나타내었다(Jeffries and Jin 2000). 더 낮은 발효 비율과 결합하여, 자일로스 발효 동안의 감소된 에탄올 내성은 셀룰로스 가수분해물(cellulosic hydrolysates)에 있는 높은 농도의 글루코스(~ 70 - 100 g/L) 및 자일로스(~ 40 - 60 g/L)를 함유하는 당 혼합물의 발효에서 중요한 문제로 제시되고 있다. 미생물이 우선적으로 글루코스를 사용하기 때문에, 글루코스가 고갈된 시기(세포가 자일로스를 사용하기 시작할 때)에, 에탄올 농도는 자일로스 발효 비율을 추가적으로 감소시키기에 이미 충분히 높다(~ 35-45 g/L). 결과적으로, "글루코스의 억압(glucose repression)" 때문에 글루코스가 고갈된 후의 자일로스의 순차적인 이용은, 셀룰로스 가수분해물(cellulosic hydrolysates)에서 혼합된 당을 성공적으로 사용하기 위하여 극복하기 위한 중요한 도전이다.

[0010] 따라서, 리그노셀룰로스의 분해에 중요한 부가적인 유전자의 식별, 및 리그노셀룰로스에서의 개선된 성장을 위한 미생물의 공학(engineering)에서의 이들의 사용 및 리그노셀룰로스 분해의 결과로 초래된 화합물의 활용을 위한 필요성이 존재한다. 또한 5 탄당에서 바이오연료로의 효과적인 전환을 위한 개선된 방법 및 바이오연료를 생산하기 위한 혼합된 당의 발효를 위한 개선된 방법이 추가적으로 필요하다.

발명의 내용

[0011] 본 발명의 간단한 요약

[0012] 이러한 필요성을 충족시키기 위해, 본원에 기재된 본 발명은 세포 내로의 셀로텍스트린의 수송(transport)을 증가시키는 방법, 셀로텍스트린을 함유하는 배지에서 세포의 성장을 증가시키는 방법, 셀룰로스-유도(cellulose-derived) 및 헤미셀룰로스-유도 당(hemicellulose-derived sugars)을 공동-발효시키는(co-fermenting) 방법 및 세포 내로 셀로텍스트린을 수송하는 폴리펩티드를 코딩하는(encoding) 재조합 폴리뉴클레오티드(recombinant polynucleotide)를 함유하는 숙주 세포(host cell)를 제공하여 탄화수소(hydrocarbons) 또는 탄화수소 유도체(hydrocarbon derivatives)를 생성하는 방법을 제공하고자 한다. 추가적으로 기재된 숙주 세포는, 세포 내로 셀로텍스트린을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 함유한다. 추가적으로 본 발명은, 세포 내로 오타당을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 함유하는 숙주 세포, 세포 내로 오타당의 수송을 증가시키는 방법, 오타당을 함유하는 배지에서 세포의 성장을 증가시키는 방법, 및 세포 내

로 오탄당을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 함유하는 숙주 세포를 제공하여 탄화수소 또는 탄화수소 유도체를 제조하는 방법에 관한 것이다.

- [0013] 본원에서 사용된 바와 같이, 셀로텍스트린은, 셀로비오스(cellobiose)(2 글루코스 단량체), 셀로트로스(cellotriose)(3 글루코스 단량체), 셀로테트라오스(cellotetraose)(4 글루코스 단량체), 셀로펜타오스(cellopentaose)(5 글루코스 단량체) 및 셀로헥사오스(cellohexaose)(6 글루코스 단량체)를 포함하는, 다양한 길이(length)의 글루코스 중합체에 관한 것이다.
- [0014] 따라서, 한 양상에 있어서, 트랜스멤브레인(transmembrane) α -헬릭스 1(α -helix 1), α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드(polypeptide)를 코딩하는(encoding) 재조합 폴리뉴클레오티드(recombinant polynucleotide)를 포함하는, 상기 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO : 1 을 포함하는 숙주 세포를 준비하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로의 셀로텍스트린의 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현되는 배지(medium)에서 세포를 배양하는 것을 포함하는, 세포 내로 셀로텍스트린 수송을 증가시키는 방법을 포함한다.
- [0015] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 상기 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO : 2 을 포함하는 숙주 세포를 준비하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로의 셀로텍스트린의 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현되는 배지(medium)에서 세포를 배양하는 것을 포함하는, 세포 내로 셀로텍스트린 수송을 증가시키는 방법을 포함한다.
- [0016] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 상기 트랜스멤브레인 α -헬릭스 2 및 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프(loop)는 SEQ ID NO : 3 을 포함하는 숙주 세포를 준비하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로의 셀로텍스트린의 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현되는 배지(medium)에서 세포를 배양하는 것을 포함하는, 세포 내로 셀로텍스트린 수송을 증가시키는 방법을 포함한다.
- [0017] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 상기 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO : 4 을 포함하는 숙주 세포를 준비하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로의 셀로텍스트린의 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현되는 배지(medium)에서 세포를 배양하는 것을 포함하는, 세포 내로 셀로텍스트린 수송을 증가시키는 방법을 포함한다.
- [0018] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 상기 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO : 5 을 포함하는 숙주 세포를 준비하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로의 셀로텍스트린의 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현되는 배지(medium)에서 세포를 배양하는 것을 포함하는, 세포 내로 셀로텍스트린 수송을 증가시키는 방법을 포함한다.

한다.

[0019] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 상기 트랜스멤브레인 α -헬릭스 6 및 상기 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO : 6 을 포함하는 숙주 세포를 준비하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로의 셀로텍스트린의 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현되는 배지(medium)에서 세포를 배양하는 것을 포함하는, 세포 내로 셀로텍스트린 수송을 증가시키는 방법을 포함한다.

[0020] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 상기 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO : 7 을 포함하는 숙주 세포를 준비하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로의 셀로텍스트린의 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현되는 배지에서 세포를 배양하는 것을 포함하는, 세포 내로 셀로텍스트린 수송을 증가시키는 방법을 포함한다.

[0021] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 상기 트랜스멤브레인 α -헬릭스 10, 상기 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO : 8 을 포함하는 숙주 세포를 준비하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로의 셀로텍스트린의 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현되는 배지에서 세포를 배양하는 것을 포함하는, 세포 내로 셀로텍스트린 수송을 증가시키는 방법을 포함한다.

[0022] 전술한 양상 중 어떠한 것과 결합할 수 있는 특정 실시형태에 있어서, 상기 폴리펩티드는 NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 % 또는 적어도 100 % 동일한 아미노산을 갖는다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에 있어서, 숙주 세포는 β -글루코시다아제(β -glucosidase)의 적어도 촉매 도메인(catalytic domain)을 코딩하는(encodes) 두 번째 재조합 폴리뉴클레오티드를 포함한다. β -글루코시다아제의 촉매 도메인을 코딩하는 두 번째 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에 있어서, β -글루코시다아제는 *네투로스포라 크라사*(*Neurospora crassa*)로부터 온 것이다. *네투로스포라 크라사*로부터의 β -글루코시다아제의 적어도 촉매 도메인을 코딩하는 두 번째 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에 있어서, β -글루코시다아제는 NCU00130 에 의해 코딩되었다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에 있어서, 숙주 세포는 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 더 포함하고, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용(utilization)에 포함되는 하나 또는 그 이상의 효소를 코딩한다. 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코딩하는 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 더 함유하는 숙주 세포를 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에 있어서, 하나 또는 그 이상의 효소는 L-아라비노스 이성질화효소(L-arabinose isomerase), L-리블로인산화효소(L- ribulokinase), L-리블로스-5-P 4 에피머화효소(L-ribulose-5-P 4 epimerase), 자일로스 이성질화효소(xylose isomerase), 자일루로키나아제(xylulokinase), 알도스 환원효소(aldose reductase), L-아라비니톨 4-탈수소효소(L-arabinitol 4-dehydrogenase), L-자일루로스 환원효소(L- xylulose reductase) 및 자일리톨 탈수소효소로 이루어진 하나 또는 그 이상의 군으로부터 선택된 것이다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에 있어서, 숙주 세포는 오타당 운반체를 코딩하는 세 번째 재조합 폴리뉴클레오티드를 더 포함한다. 오타당 운반체를 코딩하는 세 번째 재조합 폴리뉴클레오티드를 더 포함하는 숙주 세포를 갖는 전술한 실시형태와 결합할 수

있는 특정 실시형태에 있어서, 오타당 운반체는 NCU00821, NCU04963, NCU06138, STL12/XUT6, SUT2, SUT3, XUT1 및 XUT3 으로 이루어진 군으로부터 선택된다.

[0023] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO : 1 을 포함하며, 상기 폴리펩티드는 셀로텍스트린 운반체인 숙주 세포를 준비하고, 숙주 세포가 재조합 폴리뉴클레오티드를 포함하지 않는 세포 보다, 배지에서 더 빠른 비율로 성장하는, 셀로텍스트린을 함유하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다.

[0024] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 2 은 SEQ ID NO : 2 를 포함하며, 상기 폴리펩티드는 셀로텍스트린 운반체인 숙주 세포를 준비하고, 숙주 세포가 재조합 폴리뉴클레오티드를 포함하지 않는 세포 보다, 배지에서 더 빠른 비율로 성장하는, 셀로텍스트린을 함유하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다.

[0025] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 2 및 상기 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프(loop)은 SEQ ID NO : 3 를 포함하며, 상기 폴리펩티드는 셀로텍스트린 운반체인 숙주 세포를 준비하고, 숙주 세포가 재조합 폴리뉴클레오티드를 포함하지 않는 세포 보다, 배지에서 더 빠른 비율로 성장하는, 셀로텍스트린을 함유하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다.

[0026] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 5 은 SEQ ID NO : 4 를 포함하며, 상기 폴리펩티드는 셀로텍스트린 운반체인 숙주 세포를 준비하고, 숙주 세포가 재조합 폴리뉴클레오티드를 포함하지 않는 세포 보다, 배지에서 더 빠른 비율로 성장하는, 셀로텍스트린을 함유하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다.

[0027] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO : 5 를 포함하며, 상기 폴리펩티드는 셀로텍스트린 운반체인 숙주 세포를 준비하고, 숙주 세포가 재조합 폴리뉴클레오티드를 포함하지 않는 세포 보다, 배지에서 더 빠른 비율로 성장하는, 셀로텍스트린을 함유하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다.

[0028] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 6 및 상기 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO : 6 를 포함하며, 상기 폴리펩티드는 셀로텍스트린 운반체인 숙주 세포를 준비하고, 숙주 세포가 재조합 폴리뉴클레오티드를 포함하지 않는 세포 보다, 배지에서 더 빠른 비율로 성장하는, 셀로텍스트린을 함유하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는

는 방법을 포함한다.

[0029] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO : 7 를 포함하며, 상기 폴리펩티드는 셀로텍스트린 운반체인 숙주 세포를 준비하고, 숙주 세포가 재조합 폴리뉴클레오티드를 포함하지 않는 세포 보다, 배지에서 더 빠른 비율로 성장하는, 셀로텍스트린을 함유하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다.

[0030] 또 다른 양상에 있어서, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 10, 상기 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO : 8 을 포함하며, 상기 폴리펩티드는 셀로텍스트린 운반체인 숙주 세포를 준비하고, 숙주 세포가 재조합 폴리뉴클레오티드를 포함하지 않는 세포 보다, 배지에서 더 빠른 비율로 성장하는, 셀로텍스트린을 함유하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다.

[0031] 세포의 성장을 증가시키는 전술한 양상 중 어떠한 것과 결합할 수 있는 특정 실시형태에 있어서, 상기 폴리펩티드는 NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 % 또는 적어도 100 % 의 아미노산 동일성을 갖는다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에 있어서, 숙주 세포는 내인성 또는 두 번째 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리뉴클레오티드는 β -글루코시다아제(β -glucosidase)의 적어도 촉매 도메인(catalytic domain)을 코드한다. β -글루코시다아제의 적어도 촉매 도메인을 코드하는, 내인성 또는 두 번째 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에 있어서, β -글루코시다아제는 *네투로스포라 크라사*(*Neurospora crassa*)로부터 온 것이다. *네투로스포라 크라사*로부터의 β -글루코시다아제의 적어도 하나의 촉매 도메인을 코드하는, 내인성 또는 두 번째 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에 있어서, β -글루코시다아제는 NCU00130 에 의해 코드되었다.

[0032] 또 다른 양상에 있어서, 본 발명은, 셀로텍스트린 운반체를 코딩하는(encoding) 첫 번째 재조합 폴리뉴클레오티드(first recombinant polynucleotide) 및 β -글루코시다아제의 촉매 도메인(catalytic domain)을 코딩하는 두 번째 재조합 폴리뉴클레오티드(second recombinant polynucleotide)를 포함하는 숙주 세포를 제공하고, 재조합 폴리뉴클레오티드의 발현이 셀룰로스-유도 당(cellulose-derived sugar) 및 헤미셀룰로스-유도 당(hemicellulose-derived sugar)의 공동-발효(co-fermentation)를 가능하게 하는, 셀룰로스-유도 당 및 헤미셀룰로스-유도 당을 포함하는 배지에 상기 숙주 세포를 배양시킴을 포함하는, 셀룰로스-유도 및 헤미셀룰로스-유도 당을 공동-발효시키는 방법을 포함한다. 특정 실시형태에 있어서, 첫 번째 재조합 폴리뉴클레오티드는 트랜스 멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO: 1 을 포함한다. 특정 실시형태에 있어서, 첫 번째 재조합 폴리뉴클레오티드는 트랜스 멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α -헬릭스 2 은 SEQ ID NO: 2 를 포함한다. 특정 실시형태에 있어서, 첫 번째 재조합 폴리뉴클레오티드는 트랜스 멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α -헬릭스 2 및 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프는 SEQ ID NO: 3 을 포함한다. 특정 실시형태에 있어서, 첫 번째 재조합 폴리뉴클레오티드는 트랜스 멤브레인 α -헬릭스

스 1, α-헬릭스 2, α-헬릭스 3, α-헬릭스 4, α-헬릭스 5, α-헬릭스 6, α-헬릭스 7, α-헬릭스 8, α-헬릭스 9, α-헬릭스 10, α-헬릭스 11, α-헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α-헬릭스 5 는 SEQ ID NO: 4 를 포함한다. 특정 실시형태에 있어서, 첫 번째 재조합 폴리뉴클레오티드는 트랜스 멤브레인 α-헬릭스 1, α-헬릭스 2, α-헬릭스 3, α-헬릭스 4, α-헬릭스 5, α-헬릭스 6, α-헬릭스 7, α-헬릭스 8, α-헬릭스 9, α-헬릭스 10, α-헬릭스 11, α-헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α-헬릭스 6 는 SEQ ID NO: 5 를 포함한다. 특정 실시형태에 있어서, 첫 번째 재조합 폴리뉴클레오티드는 트랜스 멤브레인 α-헬릭스 1, α-헬릭스 2, α-헬릭스 3, α-헬릭스 4, α-헬릭스 5, α-헬릭스 6, α-헬릭스 7, α-헬릭스 8, α-헬릭스 9, α-헬릭스 10, α-헬릭스 11, α-헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α-헬릭스 6 및 트랜스멤브레인 α-헬릭스 7 사이의 서열은 SEQ ID NO: 6 을 포함한다. 특정 실시형태에 있어서, 첫 번째 재조합 폴리뉴클레오티드는 트랜스 멤브레인 α-헬릭스 1, α-헬릭스 2, α-헬릭스 3, α-헬릭스 4, α-헬릭스 5, α-헬릭스 6, α-헬릭스 7, α-헬릭스 8, α-헬릭스 9, α-헬릭스 10, α-헬릭스 11, α-헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α-헬릭스 7 은 SEQ ID NO: 7 을 포함한다. 특정 실시형태에 있어서, 첫 번째 재조합 폴리뉴클레오티드는 트랜스 멤브레인 α-헬릭스 1, α-헬릭스 2, α-헬릭스 3, α-헬릭스 4, α-헬릭스 5, α-헬릭스 6, α-헬릭스 7, α-헬릭스 8, α-헬릭스 9, α-헬릭스 10, α-헬릭스 11, α-헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 트랜스멤브레인 α-헬릭스 10, 트랜스멤브레인 α-헬릭스 11 및 이들 사이의 서열은 SEQ ID NO: 8 을 포함한다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에 있어서, 상기 폴리펩티드는 NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 % 또는 적어도 100 % 의 아미노산 동일성을 갖는다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에 있어서, β-글루코시다아제는 *네푸로스포라 크라사(Neurospora crassa)*로부터 온 것이다. *네푸로스포라 크라사*로부터의 β-글루코시다아제의 촉매 도메인을 코딩하는 두 번째 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, β-글루코시다아제는 NCU00130 에 의해 코드되었다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에 있어서, 숙주 세포는 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 추가로 포함하고, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용(pentose utilization)에 포함되는 하나 또는 그 이상의 효소를 코드한다. 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 추가로 포함하는 숙주 세포를 갖되, 상기 하나 또는 그 이상의 재조합 폴리뉴클레오티드는 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코드하는, 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 하나 또는 그 이상의 효소는, L-아라비노스 이성질화효소(L-arabinose isomerase), L-리블로인산화효소(L-ribulokinase), L-리블로스-5-P 4 에피머화효소(L-ribulose-5-P 4 epimerase), 자일로스 이성질화효소(xylose isomerase), 자일루로키나아제(xylulokinase), 알도스 환원효소(aldose reductase), L-아라비니톨 4-탈수소효소(L-arabinitol 4- dehydrogenase), L-자일루로스 환원효소(L- xylulose reductase) 및 자일리톨 탈수소효소로 이루어진 하나 또는 그 이상의 군으로부터 선택한 것이다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에 있어서, 숙주 세포는 세 번째 재조합 폴리뉴클레오티드를 추가로 포함하되, 상기 세 번째 재조합 폴리뉴클레오티드는 오타당 운반체를 코드한다. 숙주 세포가 추가적으로 세 번째 재조합 폴리뉴클레오티드를 포함하되, 상기 세 번째 재조합 폴리뉴클레오티드는 오타당 운반체를 코드하는 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 오타당 운반체는 NCU00821, NCU04963, NCU06138, STL12/XUT6, SUT2, SUT3, XUT1 및 XUT3 으로 이루어진 군으로부터 선택한 것이다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 상기 셀룰로스-유도 당은 셀로비오스, 셀로트리오스 및 셀로테트라오스로 이루어진 군으로부터 선택된 것이고, 헤미셀룰로스-유도 당은 자일로스이다.

[0033] 또 다른 양상에 있어서, 본 발명은, 숙주 세포가 트랜스 멤브레인 α-헬릭스 1, α-헬릭스 2, α-헬릭스 3, α-헬릭스 4, α-헬릭스 5, α-헬릭스 6, α-헬릭스 7, α-헬릭스 8, α-헬릭스 9, α-헬릭스 10, α-헬릭스 11, α-헬릭스 12 를 포함하는 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α-헬릭스 1 은 SEQ ID NO: 1 을 포함하고, 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송하는, 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드의 발현으로 상기 세포 내로 셀로텍스트린의 수송이 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원(source)을 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 숙주 세포에 의해 탄화수소(hydrocarbons) 또는 탄화수소 유도체(hydrocarbon derivatives)의 합성을 증가시키는 방법을 포함한다.

[0034] 또 다른 양상에 있어서, 본 발명은, 숙주 세포가 트랜스 멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO: 2 를 포함하고, 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송하는, 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드의 발현으로 상기 세포 내로 셀로텍스트린의 수송이 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원(source)을 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 숙주 세포에 의해 탄화수소(hydrocarbons) 또는 탄화수소 유도체(hydrocarbon derivative s)의 합성을 증가시키는 방법을 포함한다.

[0035] 또 다른 양상에 있어서, 본 발명은, 숙주 세포가 트랜스 멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 2 및 상기 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프는 SEQ ID NO: 3 을 포함하고, 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송하는, 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드의 발현으로 상기 세포 내로 셀로텍스트린의 수송이 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원(source)을 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 숙주 세포에 의해 탄화수소(hydrocarbons) 또는 탄화수소 유도체(hydrocarbon derivatives)의 합성을 증가시키는 방법을 포함한다.

[0036] 또 다른 양상에 있어서, 본 발명은, 숙주 세포가 트랜스 멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO: 4 를 포함하고, 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송하는, 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드의 발현으로 상기 세포 내로 셀로텍스트린의 수송이 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원(source)을 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 숙주 세포에 의해 탄화수소(hydrocarbons) 또는 탄화수소 유도체(hydrocarbon derivative s)의 합성을 증가시키는 방법을 포함한다.

[0037] 또 다른 양상에 있어서, 본 발명은, 숙주 세포가 트랜스 멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO: 5 를 포함하고, 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송하는, 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드의 발현으로 상기 세포 내로 셀로텍스트린의 수송이 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원(source)을 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 숙주 세포에 의해 탄화수소(hydrocarbons) 또는 탄화수소 유도체(hydrocarbon derivative s)의 합성을 증가시키는 방법을 포함한다.

[0038] 또 다른 양상에 있어서, 본 발명은, 숙주 세포가 트랜스 멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 6 및 상기 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO: 6 을 포함하고, 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송하는, 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드의 발현으로 상기 세포 내로 셀로텍스트린의 수송이 증가되는, 숙주 세

포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원(source)을 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 숙주 세포에 의해 탄화수소(hydrocarbons) 또는 탄화수소 유도체(hydrocarbon derivatives)의 합성을 증가시키는 방법을 포함한다.

[0039] 또 다른 양상에 있어서, 본 발명은, 숙주 세포가 트랜스 멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO: 7 을 포함하고, 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송하는, 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드의 발현으로 상기 세포 내로 셀로텍스트린의 수송이 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원(source)을 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 숙주 세포에 의해 탄화수소(hydrocarbons) 또는 탄화수소 유도체(hydrocarbon derivatives)의 합성을 증가시키는 방법을 포함한다.

[0040] 또 다른 양상에 있어서, 본 발명은, 숙주 세포가 트랜스 멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 10, 상기 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO: 8 을 포함하고, 폴리펩티드는 탄화수소 또는 탄화수소 유도체를 합성하기 위해 숙주 세포 내로 셀로텍스트린을 수송하는, 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드의 발현으로 상기 세포 내로 셀로텍스트린의 수송이 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원(source)을 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 숙주 세포에 의해 탄화수소(hydrocarbons) 또는 탄화수소 유도체(hydrocarbon derivatives)의 합성을 증가시키는 방법을 포함한다.

[0041] 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 전술한 양상 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 상기 폴리펩티드는 NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 % 또는 적어도 100 % 의 아미노산 동일성을 갖는다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 숙주 세포는 추가로 두 번째 재조합 폴리뉴클레오티드를 포함하되, 상기 폴리뉴클레오티드는 β -글루코시다아제의 적어도 촉매 도메인을 코딩한다. 숙주 세포가 추가로 두 번째 재조합 폴리뉴클레오티드를 포함하되, 상기 폴리뉴클레오티드가 β -글루코시다아제의 적어도 촉매 도메인을 코딩하는 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, β -글루코시다아제는 *네투로스포라 크라사(Neurospora crassa)*로부터 온 것이다. 숙주 세포가 추가로 두 번째 재조합 폴리뉴클레오티드를 포함하되, 상기 폴리뉴클레오티드가 *네투로스포라 크라사*로부터의 β -글루코시다아제의 적어도 촉매 도메인을 코딩하는, 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, β -글루코시다아제는 NCU00130 에 의해 코딩되었다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 셀로텍스트린 근원은 셀룰로스를 포함한다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 탄화수소 또는 탄화수소 유도체는 연료(fuel)로서 사용될 수 있다. 탄화수소 또는 탄화수소 유도체는 연료로서 사용될 수 있는 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 탄화수소 또는 탄화수소 유도체는 에탄올을 포함한다. 탄화수소 또는 탄화수소 유도체는 연료로서 사용될 수 있는 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 탄화수소 또는 탄화수소 유도체는 부탄올(butanol)을 포함한다.

[0042] 전술한 양상 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 상기 배지는 변경된 유기체(altered organism)로부터 셀룰라아제-함유 효소 혼합물(cellulase-containing enzyme mixture)을 포함하되, 상기 셀룰라아제-함유 혼합물(cellulase-containing mixture)은 변경되지 않는 유기체(unaltered organism)로부터의 셀룰라아제-함유 혼합물과 비교하여 감소된 β -글루코시다아제 활성도를 갖는다. 전술한 양상 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 상기 숙주 세포는, *사카로마이시스(Saccharomyces sp.)*, *사카로미세스 세레비시아*

(*Saccharomyces cerevisiae*), 사카로미세스 모나센시스(*Saccharomyces monacensis*), 사카로미세스 바야누스(*Saccharomyces bayanus*), 사카로미세스 파스토리아누스(*Saccharomyces pastorianus*), 사카로미세스 칼스베르겐시스(*Saccharomyces carlsbergensis*), 사카로미세스 폼베(*Saccharomyces pombe*), 클루이베르마이세스(*Kluyveromyces sp.*), 클루이베르마이세스 마르시아무스(*Kluyveromyces marxianus*), 클루이베르마이세스 락티스(*Kluyveromyces lactis*), 클루이베르마이세스 프라길리스(*Kluyveromyces fragilis*), 피치아 스티피티스(*Pichia stipitis*), 스포로트리쿰 서머파일(*Sporotrichum thermophile*), 칸디다 쉬하테(*Candida shehatae*), 칸디다 트로피칼리스(*Candida tropicalis*), 네우로스포라 크라사(*Neurospora crassa*), 자이모나스 모빌리스(*Zymomonas mobilis*), 클로스트리듐(*Clostridium sp.*), 클로스트리듐 파이토퍼멘탄스(*Clostridium phytofermentans*), 클로스트리듐 써모셀럼(*Clostridium thermocellum*), 클로스트리듐 베이저린키이(*Clostridium beijerinckii*), 클로스트리듐 아세토부틸리쿰(*Clostridium acetobutylicum*), 무렐라 서모아세티카(*Moorella thermoacetica*), 대장균(*Escherichia coli*), 크레브시엘라 옥시토키아(*Klebsiella oxytoca*), 서모안아에로박테리움 사카로락티쿰(*Thermoanaerobacterium saccharolyticum*) 및 고초균(*Bacillus subtilis*)으로 이루어진 군으로부터 선택한 것이다. 전술한 양상 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 상기 셀로텍스트린은, 셀로비오스, 셀로트리오스 및 셀로테트라오스로 이루어진 하나 또는 그 이상의 군으로부터 선택된 것이다.

[0043]

또 다른 양상은 트랜스 멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12, 세포 내의 N-말단 (intracellular N-terminus), 세포 내의 C-말단, 및 트랜스 멤브레인 α -헬릭스 1 에서의 SEQ ID NO : 1, 트랜스 멤브레인 α -헬릭스 2 에서의 SEQ ID NO : 2, 트랜스 멤브레인 α -헬릭스 2 및 트랜스 멤브레인 α -헬릭스 3 을 연결하는 루프에서의 SEQ ID NO : 3, 트랜스 멤브레인 α -헬릭스 5 에서의 SEQ ID NO : 4, 트랜스 멤브레인 α -헬릭스 6 에서의 SEQ ID NO : 5, 트랜스 멤브레인 α -헬릭스 6 및 트랜스 멤브레인 α -헬릭스 7 사이의 서열에서 SEQ ID NO : 6, 트랜스 멤브레인 α -헬릭스 7 에서의 SEQ ID NO : 7 및 트랜스 멤브레인 α -헬릭스 10, 트랜스 멤브레인 α -헬릭스 11 및 이들 사이의 서열에서 SEQ ID NO : 8 로 이루어진 군으로부터 선택된 서열을 갖는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 포함하되, 상기 폴리펩티드는 셀로텍스트린 운반체이다. 특정 실시형태에 있어서, 상기 폴리펩티드는 NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 % 또는 적어도 100 % 의 아미노산 동일성을 갖는다. 전술한 실시형태 중 어떤 것과 결합할 수 있는 특정 실시형태에서, 상기 숙주 세포는 추가로 두 번째 재조합 폴리뉴클레오티드를 포함하되, 상기 두 번째 재조합 폴리뉴클레오티드는 β -글루코시다아제의 촉매 도메인을 코드한다. 상기 두 번째 재조합 폴리뉴클레오티드는 β -글루코시다아제의 촉매 도메인을 코드하는, 두 번째 재조합 폴리뉴클레오티드를 추가로 포함하는 숙주 세포를 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, β -글루코시다아제는 *네우로스포라 크라사*(*Neurospora crassa*)로부터 온 것이다. *네우로스포라 크라사*로부터의 β -글루코시다아제의 촉매 도메인을 코드하는 두 번째 재조합 폴리뉴클레오티드를 추가로 포함하는 숙주세포를 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, β -글루코시다아제는 NCU00130 에 의해 코드되었다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 상기 숙주 세포는 추가로 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 포함하되, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코드한다. 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코드하고, 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 추가로 포함하는 숙주 세포를 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 하나 또는 그 이상의 효소는, L-아라비노스 이성질화효소, L-리블로인산화효소, L-리블로스-5-P 4 에피머화효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비니톨 4-탈수소효소, L-자일루로스 환원효소 및 자일리톨 탈수소효소로 이루어진 하나 또는 그 이상의 군으로부터 선택된 것이다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 상기 숙주 세포는 추가로 세 번째 재조합 폴리뉴클레오티드를 포함하되, 상기 세 번째 재조합 폴리뉴클레오티드는 오타당 운반체를 코드한다. 오타당 운반체를 코드하는 세 번째 재조합 폴리뉴클레오티드를 추가로 포함하는 숙주 세포를 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 오타당 운반체는 NCU00821, NCU04963, NCU06138, STL12/XUT6, SUT2, SUT3, XUT1 및 XUT3 으로 이루어진 군으로부터 선택된 것이다.

- [0044] 전술한 양상 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 숙주 세포는 하나 또는 그 이상의 재조합 폴리뉴클레오티드에 작동가능하게 연결된(operably linked) 하나 또는 그 이상의 유도성 프로모터(inducible promoters)를 추가로 포함한다.
- [0045] 또 다른 양상은 NCU00821 및 STL12/XUT6 으로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 포함하되, 상기 폴리펩티드는 세포 내로 자일로스를 수송한다.
- [0046] 또 다른 양상은 XUT1 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 포함하되, 상기 폴리펩티드는 세포 내로 아라비노스를 수송한다.
- [0047] 또 다른 양상은 NCU06138 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 포함하되, 상기 폴리펩티드는 세포 내로 아라비노스 및 글루코스를 수송한다.
- [0048] 또 다른 양상은 SUT2, SUT3 및 XUT3 로 이루어진 군으로부터 선택된 폴리뉴클레오티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 포함하되, 상기 폴리펩티드는 세포 내로 자일로스 및 글루코스를 수송한다.
- [0049] 또 다른 양상은 NCU04963 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 포함하되, 상기 폴리펩티드는 세포 내로 자일로스, 아라비노스 및 글루코스를 수송한다.
- [0050] 오타당 운반체를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 갖는 전술한 양상 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 상기 숙주 세포는 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 추가로 포함하되, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코드한다. 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코드하는 하나 또는 그 이상의 폴리뉴클레오티드를 추가로 포함하는 숙주 세포를 갖는 전술한 실시형태와 결합될 수 있는 특정 실시형태에서, 하나 또는 그 이상의 효소는, L-아라비노스 이성질화효소, L-리블로인산화효소, L-리블로스-5-P 4 에피머화효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비니톨 4-탈수소효소, L-자일루로스 환원효소 및 자일리톨 탈수소효소로 이루어진 하나 또는 그 이상의 군으로부터 선택된 것이다.
- [0051] 또 다른 양상에서 본 발명은, NCU00821 및 STL12/XUT6 로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 자일로스 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현될 정도까지 세포를 배양하는 것을 포함하는, 세포 내로 자일로스의 수송을 증가시키는 방법을 포함한다.
- [0052] 또 다른 양상에서 본 발명은, XUT1 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 아라비노스 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현될 정도까지 세포를 배양하는 것을 포함하는, 세포 내로 아라비노스의 수송을 증가시키는 방법을 포함한다.
- [0053] 또 다른 양상에서 본 발명은, NCU06138 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 아라비노스 또는 글루코스 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현될

정도까지 세포를 배양하는 것을 포함하는, 세포 내로 아라비노스 또는 글루코스의 수송을 증가시키는 방법을 포함한다.

[0054] 또 다른 양상에서 본 발명은, SUT2, SUT3 및 XUT3 으로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 자일로스 또는 글루코스 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현될 정도까지 세포를 배양하는 것을 포함하는, 세포 내로 자일로스 또는 글루코스의 수송을 증가시키는 방법을 포함한다.

[0055] 또 다른 양상에서 본 발명은, NCU04963 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 것을 포함하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 자일로스, 아라비노스 또는 글루코스 수송이 증가되는, 재조합 폴리뉴클레오티드가 발현될 정도까지 세포를 배양하는 것을 포함하는, 세포 내로 자일로스, 아라비노스 또는 글루코스의 수송을 증가시키는 방법을 포함한다.

[0056] 세포 내로 자일로스, 아라비노스 또는 글루코스의 수송을 증가시키는 전술한 양상 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 방법은 추가적으로 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 포함하되, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코딩한다. 하나 또는 그 이상의 폴리뉴클레오티드가 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코딩하는, 추가로 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 포함하는 방법을 갖는 전술한 실시형태와 결합될 수 있는 특정 실시형태에서, 하나 또는 그 이상의 효소는, L-아라비노스 이성질화효소, L-리블로인산화효소, L-리블로스-5-P 4 에피머화효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비니톨 4-탈수소효소, L-자일루로스 환원효소 및 자일리톨 탈수소효소로 이루어진 하나 또는 그 이상 군으로부터 선택된 것이다.

[0057] 또 다른 양상은, 숙주 세포는 NCU00821 및 STL12/XUT6 으로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 자일로스를 수송하는, 상기 숙주 세포를 제공하는 것을 포함하며, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하는, 자일로스를 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다.

[0058] 또 다른 양상은, 숙주 세포는 XUT1 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 아라비노스를 수송하는, 상기 숙주 세포를 제공하는 것을 포함하며, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하는, 아라비노스를 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다.

[0059] 또 다른 양상은, 숙주 세포는 NCU06138 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 아라비노스 및 글루코스를 수송하는, 상기 숙주 세포를 제공하는 것을 포함하며, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하는, 아라비노스 또는 글루코스를 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다.

[0060] 또 다른 양상은, 숙주 세포는 SUT2, SUT3 및 XUT3 으로 이루어진 군으로부터 선택한 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 자일로스 및 글루코스를 수송하는, 상기 숙주 세포를 제공하는 것을 포함하며, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하는, 자일로스 또는 글루코스를 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다.

- [0061] 또 다른 양상은, 숙주 세포는 NCU04963 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 자일로스, 아라비노스 및 글루코스를 수송하는, 상기 숙주 세포를 제공하는 것을 포함하며, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하는, 자일로스, 아라비노스 또는 글루코스를 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하는, 세포의 성장을 증가키는 방법을 포함한다.
- [0062] 자일로스 및/또는 아라비노스 및/또는 글루코스를 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 배양하여 세포의 성장을 증가시키는 전술한 양상과 결합될 수 있는 특정 실시형태에서, 숙주 세포는, 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코딩하는 하나 또는 그 이상의 내인성 또는 재조합 폴리뉴클레오티드를 추가로 포함한다. 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코딩하는 하나 또는 그 이상의 내인성 또는 재조합 폴리뉴클레오티드를 추가로 포함하는 숙주 세포를 갖는 전술한 실시형태와 결합될 수 있는 특정 실시형태에서, 하나 또는 그 이상의 효소는, L-아라비노스 이성질화효소, L-리블로인 산화효소, L-리블로스-5-P 4 에피머화효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비니톨 4-탈수소효소, L-자일루로스 환원효소 및 자일리톨 탈수소효소로 이루어진 하나 또는 그 이상 군으로부터 선택된 것이다.
- [0063] 또 다른 양상은, NCU00821 및 STL12/XUT6 로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하되, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체의 합성을 위해 숙주 세포 내로 자일로스를 수송하며, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 자일로스 또는 자일로스 근원을 포함하는 배지에서 숙주 세포를 배양하되, 상기 세포 내로의 자일로스의 수송은 재조합 폴리뉴클레오티드의 발현으로 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함한다.
- [0064] 또 다른 양상은, XUT1 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하되, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체의 합성을 위해 숙주 세포 내로 아라비노스를 수송하며, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 아라비노스 또는 아라비노스 근원을 포함하는 배지에서 숙주 세포를 배양하되, 상기 세포 내로의 아라비노스의 수송은 재조합 폴리뉴클레오티드의 발현으로 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함한다.
- [0065] 또 다른 양상은, NCU06138 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하되, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체의 합성을 위해 숙주 세포 내로 아라비노스 또는 글루코스를 수송하며, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 아라비노스 또는 글루코스 또는 아라비노스 또는 글루코스의 근원을 포함하는 배지에서 숙주 세포를 배양하되, 상기 세포 내로의 아라비노스 또는 글루코스의 수송은 재조합 폴리뉴클레오티드의 발현으로 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함한다.
- [0066] 또 다른 양상은, SUT2, SUT3 및 XUT3 으로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하되, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체의 합성을 위해 숙주 세포 내로 자일로스 또는 글루코스를 수송하며, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 자일로스 또는 글루코스 또는 자일로스 또는 글루코스의 근원을 포함하는 배지에서 숙주 세포를 배양하되, 상기 세포 내로의 자일로스 또는 글루코스의 수송은 재조합 폴리뉴클레오티드의 발현으로 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함한다.
- [0067] 또 다른 양상은, NCU04963 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하되,

상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체의 합성을 위해 숙주 세포 내로 자일로스, 아라비노스 또는 글루코스를 수송하며, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 자일로스, 아라비노스 또는 글루코스 또는 자일로스, 아라비노스 또는 글루코스의 근원을 포함하는 배지에서 숙주 세포를 배양하되, 상기 세포 내로의 자일로스, 아라비노스 또는 글루코스의 수송은 재조합 폴리뉴클레오티드의 발현으로 증가되는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함한다.

[0068]

글루코스를 수송하는 펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 배양하여 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 전술한 양상 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 글루코스의 근원은 셀룰로스를 포함한다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 자일로스 또는 아라비노스의 근원은 헤미셀룰로스를 포함한다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 탄화수소 또는 탄화수소 유도체는 연료로서 사용될 수 있다. 탄화수소 또는 탄화수소 유도체는 연료로서 사용될 수 있는 것을 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 탄화수소 또는 탄화수소 유도체는 에탄올을 포함한다. 탄화수소 또는 탄화수소 유도체는 연료로서 사용될 수 있는 것을 갖는 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 탄화수소 또는 탄화수소 유도체는 부탄올을 포함한다.

[0069]

전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 상기 숙주 세포는, *사카로마이시스* (*Saccharomyces* sp.), *사카로미세스 세레비시아* (*Saccharomyces cerevisiae*), *사카로미세스 모나센시스* (*Saccharomyces monacensis*), *사카로미세스 바야누스* (*Saccharomyces bayanus*), *사카로미세스 파스토리아누스* (*Saccharomyces pastorianus*), *사카로미세스 칼스베르겐시스* (*Saccharomyces carlsbergensis*), *사카로미세스 폼베* (*Saccharomyces pombe*), *클루이베르마이세스* (*Kluyveromyces* sp.), *클루이베르마이세스 마르시아무스* (*Kluyveromyces marxianus*), *클루이베르마이세스 락티스* (*Kluyveromyces lactis*), *클루이베르마이세스 프라길리스* (*Kluyveromyces fragilis*), *피치아 스티피티스* (*Pichia stipitis*), *스포로트리쿰 서머파일* (*Sporotrichum thermophile*), *칸디다 쉬하테* (*Candida shehatae*), *칸디다 트로피칼리스* (*Candida tropicalis*), *네우로스포라 크라사* (*Neurospora crassa*), *자이모나스 모빌리스* (*Zymomonas mobilis*), *클로스트리듐* (*Clostridium* sp.), *클로스트리듐 파이토퍼멘탄스* (*Clostridium phytofermentans*), *클로스트리듐 써모셀럼* (*Clostridium thermocellum*), *클로스트리듐 베이저린키이* (*Clostridium beijerinckii*), *클로스트리듐 아세토부틸리쿰* (*Clostridium acetobutylicum*), *무렐라 서모아세티카* (*Moorella thermoacetica*), *대장균* (*Escherichia coli*), *크레브시엘라 옥시토카* (*Klebsiella oxytoca*), *서모안아에로박테리움 사카로락티쿰* (*Thermoanaerobacterium saccharolyticum*) 및 *고초균* (*Bacillus subtilis*)으로 이루어진 군으로부터 선택한 것이다.

[0070]

또 다른 양상은, NCU07705 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 것을 포함하고, 숙주 세포가 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하고, 셀로비오스를 포함하는 배지에서 세포를 배양하는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다. 특정 실시형태에서, 상기 숙주 세포는, *사카로마이시스* (*Saccharomyces* sp.), *사카로미세스 세레비시아* (*Saccharomyces cerevisiae*), *사카로미세스 모나센시스* (*Saccharomyces monacensis*), *사카로미세스 바야누스* (*Saccharomyces bayanus*), *사카로미세스 파스토리아누스* (*Saccharomyces pastorianus*), *사카로미세스 칼스베르겐시스* (*Saccharomyces carlsbergensis*), *사카로미세스 폼베* (*Saccharomyces pombe*), *클루이베르마이세스* (*Kluyveromyces* sp.), *클루이베르마이세스 마르시아무스* (*Kluyveromyces marxianus*), *클루이베르마이세스 락티스* (*Kluyveromyces lactis*), *클루이베르마이세스 프라길리스* (*Kluyveromyces fragilis*), *피치아 스티피티스* (*Pichia stipitis*), *스포로트리쿰 서머파일* (*Sporotrichum thermophile*), *칸디다 쉬하테* (*Candida shehatae*), *칸디다 트로피칼리스* (*Candida tropicalis*), *네우로스포라 크라사* (*Neurospora crassa*), *자이모나스 모빌리스* (*Zymomonas mobilis*), *클로스트리듐* (*Clostridium* sp.), *클로스트리듐 파이토퍼멘탄스* (*Clostridium phytofermentans*), *클로스트리듐 써모셀럼* (*Clostridium thermocellum*), *클로스트리듐 베이저린키이* (*Clostridium beijerinckii*), *클로스트리듐 아세토부틸리쿰* (*Clostridium acetobutylicum*), *무렐라 서모아세티카* (*Moorella thermoacetica*), *대장균* (*Escherichia coli*), *크레브시엘라 옥시토카* (*Klebsiella oxytoca*), *서모안아에로박테리움 사카로락티쿰* (*Thermoanaerobacterium saccharolyticum*) 및 *고초균* (*Bacillus subtilis*)으로 이루어진 군으로부터 선택한 것이다. 특정 실시형태에서, 숙주 세포는 재조합 폴리뉴클레오티드에 작동가능하게

연결된 유도성 프로모터를 추가로 포함한다. 특정 실시형태에서, 셀룰라아제의 발현은 재조합 폴리뉴클레오티드의 발현으로 숙주 세포에서 증가되었다.

[0071]

또 다른 양상은, NCU05137 폴리펩티드를 코드하는 내인성 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하고, 내인성 폴리뉴클레오티드의 발현을 저해하며, 상기 숙주 세포는 상기 내인성 폴리뉴클레오티드의 발현을 저해하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하는, 바이오매스 중합체를 포함하는 배지에서 세포를 배양하는 것을 포함하는, 바이오매스 중합체(biomass polymer)에서 세포의 성장을 증가시키는 방법을 포함한다. 특정 실시형태에서, 상기 숙주 세포는, *사카로마이스(Saccharomyces sp.)*, *사카로미세스 세레비시아(Saccharomyces cerevisiae)*, *사카로미세스 모나센시스(Saccharomyces monacensis)*, *사카로미세스 바야누스(Saccharomyces bayanus)*, *사카로미세스 파스토리아누스(Saccharomyces pastorianus)*, *사카로미세스 칼스베르겐시스(Saccharomyces carlsbergensis)*, *사카로미세스 폼베(Saccharomyces pombe)*, *클루이베르마이세스(Kluyveromyces sp.)*, *클루이베르마이세스 마르시암우스(Kluyveromyces marxianus)*, *클루이베르마이세스 락티스(Kluyveromyces lactis)*, *클루이베르마이세스 프라길리스(Kluyveromyces fragilis)*, *피치아 스티피티스(Pichia stipitis)*, *스포로트리쿰 서머파일(Sporotrichum thermophile)*, *칸디다 쉬하테(Candida shehatae)*, *칸디다 트로피칼리스(Candida tropicalis)*, *네우로스포라 크라사(Neurospora crassa)*, *자이모나스 모빌리스(Zymomonas mobilis)*, *클로스트리듐(Clostridium sp.)*, *클로스트리듐 파이토퍼멘탄스(Clostridium phytofermentans)*, *클로스트리듐 써모셀럼(Clostridium thermocellum)*, *클로스트리듐 베이저린키이(Clostridium beijerinckii)*, *클로스트리듐 아세토부틸리쿰(Clostridium acetobutylicum)*, *무렐라 서모아세티카(Moorella thermoacetica)*, *대장균(Escherichia coli)*, *크레브시엘라 옥시토카(Klebsiella oxytoca)*, *서모안아에로박테리움 사카로락티쿰(Thermoanaerobacterium saccharolyticum)* 및 *고초균(Bacillus subtilis)*으로 이루어진 군으로부터 선택된 것이다. 특정 실시형태에서, 숙주 세포의 셀룰라아제 활성도는 내인성 폴리뉴클레오티드의 발현 저해로 증가되었다. 특정 실시형태에서, 숙주 세포의 헤미셀룰라아제 활성도는 내인성 폴리뉴클레오티드의 발현 저해로 증가되었다. 특정 실시형태에서, 내인성 폴리뉴클레오티드의 발현을 저해하는 것은 내인성 폴리뉴클레오티드를 포함하는 유전자의 돌연변이(mutating) 또는 결실(deleting)을 포함한다. 특정 실시형태에서, 바이오매스 중합체는 셀룰로스이다. 특정 실시형태에서, 바이오매스 중합체(biomass polymer)는 헤미셀룰로스이다.

[0072]

또 다른 양상은, NCU01517, NCU09133 및 NCU10040 으로 이루어진 군으로부터 선택된 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하고, 숙주 세포가 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하는, 헤미셀룰로스를 포함하는 배지에서 세포를 배양시키는 것을 포함하는, 세포의 성장을 증가시키는 방법을 포함한다. 특정 실시형태에서, 상기 숙주 세포는, *사카로마이스(Saccharomyces sp.)*, *사카로미세스 세레비시아(Saccharomyces cerevisiae)*, *사카로미세스 모나센시스(Saccharomyces monacensis)*, *사카로미세스 바야누스(Saccharomyces bayanus)*, *사카로미세스 파스토리아누스(Saccharomyces pastorianus)*, *사카로미세스 칼스베르겐시스(Saccharomyces carlsbergensis)*, *사카로미세스 폼베(Saccharomyces pombe)*, *클루이베르마이세스(Kluyveromyces sp.)*, *클루이베르마이세스 마르시암우스(Kluyveromyces marxianus)*, *클루이베르마이세스 락티스(Kluyveromyces lactis)*, *클루이베르마이세스 프라길리스(Kluyveromyces fragilis)*, *피치아 스티피티스(Pichia stipitis)*, *스포로트리쿰 서머파일(Sporotrichum thermophile)*, *칸디다 쉬하테(Candida shehatae)*, *칸디다 트로피칼리스(Candida tropicalis)*, *네우로스포라 크라사(Neurospora crassa)*, *자이모나스 모빌리스(Zymomonas mobilis)*, *클로스트리듐(Clostridium sp.)*, *클로스트리듐 파이토퍼멘탄스(Clostridium phytofermentans)*, *클로스트리듐 써모셀럼(Clostridium thermocellum)*, *클로스트리듐 베이저린키이(Clostridium beijerinckii)*, *클로스트리듐 아세토부틸리쿰(Clostridium acetobutylicum)*, *무렐라 서모아세티카(Moorella thermoacetica)*, *대장균(Escherichia coli)*, *크레브시엘라 옥시토카(Klebsiella oxytoca)*, *서모안아에로박테리움 사카로락티쿰(Thermoanaerobacterium saccharolyticum)* 및 *고초균(Bacillus subtilis)*으로 이루어진 군으로부터 선택된 것이다. 특정 실시형태에서, 숙주 세포는 재조합 폴리뉴클레오티드에 작동가능하게 연결된 유도성 프로모터를 더 포함한다. 특정 실시형태에서, 숙주 세포의 헤미셀룰라아제 활성도는 재조합 폴리뉴클레오티드의 발현으로 증가되었다.

[0073]

또 다른 양상은, 셀룰로스를 포함하는 조성물을 제공하고, 변경된 유기체(altered organism)로부터의 셀룰라아

제-함유 효소 혼합물(cellulase-containing enzyme mixture)을 갖는 조성물(composition)과 접촉하되, 상기 셀룰라아제-함유 혼합물은, 변경되지 않는 유기체(unaltered organism)로부터의 셀룰라아제-함유 혼합물과 비교하여, 감소된 β -글루코시다아제 활성도를 갖고, 상기 셀룰로스는 셀룰라아제-함유 혼합물에 의해 분해되는, 셀룰로스 분해(degrading) 방법을 포함한다. 특정 실시형태에서, 상기 유기체는 β -글루코시다아제를 코딩하는 유전자의 돌연변이에 의해 변경되었다(altered). 특정 실시형태에서, 상기 유기체는 β -글루코시다아제의 발현의 감소에 의해 변경되었다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 상기 유기체는 균류(fungus) 또는 세균(bacterium) 으로 이루어진 군으로부터 선택된 것이다. 균류(fungus) 또는 세균(bacterium) 으로 이루어진 군으로부터 선택된 상기 유기체를 갖는 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 상기 유기체는 사상균(filamentous fungus)이다. 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 셀룰로스는 식물 물질(plant material)로부터 온 것이다. 식물 물질(plant material)로부터 온 셀룰로스를 갖는 전술한 실시형태 중 어떠한 것과 결합할 수 있는 특정 실시형태에서, 식물 물질은 지팡이풀(switchgrass), 미스칸투스(Miscanthus), 왕겨(rice hull), 버개스(bagasse), 아마(flax), 대나무(bamboo), 사이잘(sisal), 마닐라삼섬유(abaca), 짚(straw), 잎(leaves), 잔디 부스러기(grass clippings), 옥수수대(corn stover), 옥수수 속대(corn cobs), 증류립(distillers grains), 콩과 식물(legume plants), 수수(sorghum), 사탕수수(sugar cane), 사탕무 펄프(sugar beet pulp), 우드칩(wood chips), 톱밥(sawdust) 및 바이오매스 크롭(biomass crops)으로 이루어진 군으로부터 선택된 것이다.

[0074]

또 다른 양상은, NCU00801, NCU00988, NCU01231, NCU04963, NCU05519, NCU05853, NCU05897, NCU06138, NCU00809, NCU08114, NCU10021 및 표 15 에 나타난 유전자 중의 어떠한 것으로 이루어진 군으로부터 선택된 서열에 의해 코드된 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하고, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 화합물(compound)의 근원을 포함하는 배지에서 숙주 세포를 배양하는 것을 포함하되, 상기 화합물은 탄화수소 또는 탄화수소 유도체의 합성을 위한 기질(substrate)이고, 세포 내로 화합물의 수송은 재조합 폴리뉴클레오티드의 발현으로 증가된, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함한다. 특정 실시형태에서, 숙주 세포는, *사카로마이시스*(*Saccharomyces sp.*), *대장균*(*Escherichia coli*), *자이모나스 모빌리스*(*Zymomonas mobilis*), *네우로스포라 크라사*(*Neurospora crassa*), *칸디다 쉬하타*(*Candida shehatae*), *클로스트리듐*(*Clostridium sp.*), *클로스트리듐 파이토퍼멘탄스*(*Clostridium phytofermentans*), *클로스트리듐 써모셀럼*(*Clostridium thermocellum*), *무렐라 서모아세티카*(*Moorella thermoacetica*), *서모아나어로박테리움 사카로락티쿰*(*Thermoanaerobacterium saccharolyticum*), *클레브시엘라 옥시토카*(*Klebsiella oxytoca*) 및 *피치아 스티피티스*(*Pichia stipitis*)으로 이루어진 군으로부터 선택된 것이다. 특정 실시형태에서, 숙주 세포는 재조합 폴리뉴클레오티드에 작동가능하게 연결된 유도성 프로모터를 더 포함한다. 특정 실시형태에서, 재조합 폴리뉴클레오티드는, NCU00801, NCU00988, NCU01231, NCU04963, NCU05519, NCU05853, NCU05897, NCU06138, NCU00809, NCU08114, NCU10021, 및 표 15 에 나타난 유전자의 어떠한 것으로 이루어진 군으로부터 선택된 서열을 코드하는 폴리펩티드와 적어도 50 % 의 아미노산 동일성을 갖는 폴리펩티드를 코드한다. 몇몇 실시형태에서, 탄화수소 또는 탄화수소 유도체는 연료로서 사용될 수 있다. 특정 실시형태에서, 배지는 셀룰로스를 포함한다. 다른 실시형태에서, 상기 배지는 헤미셀룰로스를 포함한다. 특정 실시형태에서, 상기 화합물은 당이다. 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 상기 당은 오타당이다. 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 상기 당은 옥탄당이다. 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 상기 당은 이당류(disaccharide)이다. 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 상기 당은 올리고당(oligosaccharide)이다. 다른 실시형태에서, 상기 화합물은 식물 페놀(plant phenol)이다. 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 상기 식물 페놀은 퀴산(quinic acid)이다. 전술한 실시형태와 결합할 수 있는 특정 실시형태에서, 상기 식물 페놀은 니코틴아미드(nicotinamide)이다. 다른 실시형태에서, 상기 화합물은 피루브산염(pyruvate) 또는 유산염(lactate)이다.

[0075]

또 다른 양상은, 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하되, 상기 폴리뉴클레오티드는, 내용이 본원의 참조문헌으로 인용된 Supplemental Data, Dataset SI, page 3 in Tian *et al.*, *PNAS*, 2009, vol. 106, no. 52, 22157- 22162, 표 10, 표 15 에 나타난 *네우로스포라* 또는 *피치아 스티피티스* 유전자 중 어떠한 것, 또는 NCU01517, NCU09133 또는 NCU10040 에 의해 코드되는 폴리펩티드를 코드하며, 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하는, 바이오매스 중합체를 포함하는 배지에

서 세포를 배양하는 것을 포함하는 바이오매스 중합체에서 세포의 성장을 증가시키는 방법으로 포함한다. 특정 실시형태에서, 상기 폴리뉴클레오티드는, 서열 NCU00130.2, NCU00248.2, NCU00326.2, NCU00762.2, NCU00810.2, NCU00890.2, NCU03328.2, NCU03415.2, NCU03731.2, NCU03753.2, NCU04197.2, NCU04249.2, NCU04287.2, NCU04349.2, NCU04475.2, NCU04997.2, NCU05057.2, NCU05159.2, NCU05493.2, NCU05751.2, NCU05770.2, NCU05932.2, NCU06009.2, NCU06490.2, NCU07340.2, NCU07853.2, NCU07997.2, NCU08744.2, NCU08746.2, NCU08760.2, NCU09108.2, NCU09495.2, NCU09680.2 또는 NCU10045.2 중 어떠한 것을 코드되는 폴리펩티드를 코드한다. 특정 실시형태에서, 상기 폴리뉴클레오티드는 NCU07705 에 의해 코드되는 폴리펩티드를 코드한다. 특정 실시형태에서, 상기 재조합 폴리뉴클레오티드는, 표 10, Supplemental Data, Dataset SI, page 3 in Tian *et al.*, 2009, 또는 표 15 에 나타난 *네우로스포라* 또는 *피치아 스티피티스* 유전자 중 어떠한 것을 코드하는 폴리펩티드와 적어도 50 % 의 아미노산 동일성을 갖는 폴리펩티드를 코드한다. 특정 실시형태에서, 상기 폴리뉴클레오티드는, 서열 NCU00130.2, NCU00248.2, NCU00326.2, NCU00762.2, NCU00810.2, NCU00890.2, NCU03328.2, NCU03415.2, NCU03731.2, NCU03753.2, NCU04197.2, NCU04249.2, NCU04287.2, NCU04349.2, NCU04475.2, NCU04997.2, NCU05057.2, NCU05159.2, NCU05493.2, NCU05751.2, NCU05770.2, NCU05932.2, NCU06009.2, NCU06490.2, NCU07340.2, NCU07853.2, NCU07997.2, NCU08744.2, NCU08746.2, NCU08760.2, NCU09108.2, NCU09495.2, NCU09680.2 또는 NCU10045.2 중 어떠한 것에 의해 코드되는 폴리펩티드와 적어도 50 % 의 아미노산 동일성을 갖는 폴리펩티드를 코드한다. 특정 실시형태에서, 상기 재조합 폴리뉴클레오티드는 NCU07705 에 의해 코드되는 폴리펩티드와 적어도 50 % 의 아미노산 동일성을 갖는 폴리펩티드를 코드한다. 특정 실시형태에서, 상기 바이오매스 중합체는 셀룰로스이다. 다른 실시형태에서, 상기 바이오매스 중합체는 헤미셀룰로스이다. 특정 실시형태에서, 숙주 세포는, *사카로미세스 세레비시아*(*Saccharomyces cerevisiae*), *대장균*(*Escherichia coli*), *자이모나스 모빌리스*(*Zymomonas mobilis*), *네우로스포라 크라사*(*Neurospora crassa*), *칸디다 쉬하타*(*Candida shehatae*), *클로스트리듐*(*Clostridium sp.*), *클로스트리듐 파이토퍼멘탄스*(*Clostridium phytofermentans*), *클로스트리듐 써모셀럼*(*Clostridium thermocellum*), *무렐라 서모아세티카*(*Moorella thermoacetica*), *서모아나에로박테리움 사카로락티쿰*(*Thermoanaerobacterium saccharolyticum*), *클레브시엘라 옥시토카*(*Klebsiella oxytoca*) 및 *피치아 스티피티스*(*Pichia stipitis*)으로 이루어진 군으로부터 선택된 것이다. 특정 실시형태에서, 숙주 세포는 재조합 폴리뉴클레오티드에 작동가능하게 연결된 유도성 프로모터를 더 포함한다. 특정 실시형태에서, 셀룰라아제의 발현은 재조합 폴리뉴클레오티드의 발현으로 숙주 세포에서 증가되었다. 다른 실시형태에서, 헤미셀룰라아제의 발현은 재조합 폴리뉴클레오티드의 발현으로 숙주 세포에서 증가되었다.

[0076]

또 다른 양상은, 내인성 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하되, 상기 폴리뉴클레오티드는, 표 10, 내용이 본원의 참조문헌으로 인용된 Supplemental Data, Dataset SI, page 3 in Tian *et al.*, *PNAS*, 2009, vol. 106, no. 52, 22157- 22162, 또는 표 15 에 나타난 *네우로스포라* 또는 *피치아 스티피티스* 유전자 중 어떠한 것에 의해 코드되는 폴리펩티드를 코드하며, 또는 내인성 폴리뉴클레오티드의 발현을 저해하고, 및 숙주 세포는 내인성 폴리뉴클레오티드의 발현을 저해하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하는, 바이오매스 중합체를 포함하는 배지에서 세포를 배양하는 것을 포함하는, 바이오매스 중합체에서 세포의 성장을 증가시키는 방법을 포함한다. 특정 실시형태에서, 상기 내인성 폴리뉴클레오티드는 서열 NCU00130.2, NCU00248.2, NCU00326.2, NCU00762.2, NCU00810.2, NCU00890.2, NCU03328.2, NCU03415.2, NCU03731.2, NCU03753.2, NCU04197.2, NCU04249.2, NCU04287.2, NCU04349.2, NCU04475.2, NCU04997.2, NCU05057.2, NCU05159.2, NCU05493.2, NCU05751.2, NCU05770.2, NCU05932.2, NCU06009.2, NCU06490.2, NCU07340.2, NCU07853.2, NCU07997.2, NCU08744.2, NCU08746.2, NCU08760.2, NCU09108.2, NCU09495.2, NCU09680.2 또는 NCU10045.2 의 어떠한 것에 의해 코드되는 폴리펩티드를 코드한다. 특정 실시형태에서, 상기 내인성 폴리뉴클레오티드는 NCU05137 에 의해 코드되는 폴리펩티드를 코드한다. 특정 실시형태에서, 상기 내인성 폴리뉴클레오티드는, 표 10, Supplemental Data, Dataset SI, page 3 in Tian *et al.*, 2009 또는 표 15 에 나타난 *네우로스포라* 또는 *피치아 스티피티스* 유전자 중 어떠한 것에 의해 코드되는 폴리펩티드와 적어도 50 % 의 아미노산 동일성 동일성을 갖는 폴리펩티드를 코드한다. 특정 실시형태에서, 상기 내인성 폴리뉴클레오티드는, 서열 NCU00130.2, NCU00248.2, NCU00326.2, NCU00762.2, NCU00810.2, NCU00890.2, NCU03328.2, NCU03415.2, NCU03731.2, NCU03753.2, NCU04197.2, NCU04249.2, NCU04287.2, NCU04349.2, NCU04475.2, NCU04997.2, NCU05057.2, NCU05159.2, NCU05493.2, NCU05751.2, NCU05770.2, NCU05932.2, NCU06009.2, NCU06490.2, NCU07340.2, NCU07853.2, NCU07997.2, NCU08744.2, NCU08746.2, NCU08760.2, NCU09108.2, NCU09495.2, NCU09680.2, 또는 NCU10045.2 의 어떠한 것에 의해 코드되는 폴리펩티드와 적어도 50 % 동일한 아미노산을 갖는 폴리펩티드를 코

드한다. 특정 실시형태에서, 상기 내인성 폴리뉴클레오티드는 NCU05137 에 의해 코드되는 폴리펩티드와 적어도 50 % 동일한 아이노산을 갖는 폴리펩티드를 코드한다. 특정 실시형태에서, 숙주 세포는, *사카로미세스 세레비시아*(*Saccharomyces cerevisiae*), *대장균*(*Escherichia coli*), *자이모나스 모빌리스*(*Zymomonas mobilis*), *네우로스포라 크라사*(*Neurospora crassa*), *칸디다 쉬하타*(*Candida shehatae*), *클로스트리듐*(*Clostridium sp.*), *클로스트리듐 파이토퍼멘탄스*(*Clostridium phytofermentans*), *클로스트리듐 쉐모셀럼*(*Clostridium thermocellum*), *무렐라 서모아세티카*(*Moorella thermoacetica*), *서모안아에로박테리움 사카로락티쿰*(*Thermoanaerobacterium saccharolyticum*), *클레브시엘라 옥시토카*(*Klebsiella oxytoca*) 및 *피치아 스티피티스*(*Pichia stipitis*)로 이루어진 군으로부터 선택된 것이다. 특정 실시형태에서, 상기 바이오매스 중합체는 셀룰로스이다. 다른 실시형태에서, 바이오매스 중합체는 헤미셀룰로스이다. 특정 실시형태에서, 숙주 세포의 셀룰라아제 활성도는 내인성 폴리뉴클레오티드의 발현의 저해로 증가되었다. 다른 실시형태에서, 숙주 세포의 헤미셀룰라아제 활성도는 내인성 폴리뉴클레오티드의 발현의 저해로 증가되었다. 특정 실시형태에서, 내인성 폴리뉴클레오티드의 발현의 저해는 내인성 폴리뉴클레오티드를 포함하는 유전자의 돌연변이(mutating) 또는 결실(deleting)을 포함한다.

도면의 간단한 설명

도 1 은 NCU07705 에 의해 코드되는 폴리펩티드의 도메인 구조를 나타낸다.

도 2 는 NCU05137 의 계통분석(phylogenetic analysis)을 나타낸다. 엔. 크라사 NCU05137 의 예측된 오솔로그(predicted orthologs)는 BLAST 에 의한 중요한 유사성을 나타내는 아미노산 서열을 기초로 한 NCBI 및 JGI 로부터 검색하였다. 모든 확인된 실모양의 진균의 오솔로그(filamentous fungal orthologs)를 나타낸다 ; NCBI E 수치는 9e-175 인 *B. 푸켈리아나*(*B. fuckeliana*)를 제외한 0.0 이다. NCU05137 의 동족체(Homologs)는 수많은 박테리아에서 또한 확인된다(E 값 ~e-30). 폴라로모나스 나프탈렌이보란스(*Polaromonas naphthalenivorans*)(베타-프로테오박테리움, beta-proteobacterium)로부터의 YP 981875 는 외집단(outgroup)으로서 사용되었다. A. = *아스페르길루스*(*Aspergillus*) ; N. = *네오사토리아*(*Neosartorya*) ; P. 키로소게눔(*P. chyrosogenum*) = *페니실리움*(*Penicillium*) ; S. = *스클레로티니아*(*Sclerotinia*) ; B. = *보트리오티니아*(*Botryotinia*) ; P. = *피레노포라*(*Pyrenophora*) ; C. = *코칠레오볼루스*(*Cochliobolus*) ; N. *하마토코카*(*N. haematococca*) = *넥트리아*(*Nectria*) ; P. *안세리나*(*P. anserina*) = *포도스포라*(*Podospora*) ; N. = *네우로스포라*(*Neurospora*). 나무형(tree)은 MEGA3, NJ 에 의해 생성되었다. 아미노산 사이트 당 바(bar) = 0.2 치환(substitutions).

도 3 은 단일 탄소원(sole carbon source)으로서의 *미스칸투스*(*Miscanthus*) 및 *아비셀*(*Avicel*)에서 성장하였을 때의 *N. 크라사* FGSC2489 및 *T. 레세이*(*T. reesei*) QM9414 엔도글루카나제 활성도(endoglucanase activity)의 분석을 나타낸 것이다. *N. 크라사* WT 계통(strain) FGSC2489 및 *T. 레세이*(*T. reesei*) QM9414 의 배양 여과물에서 엔도글루카나제 활성도. *N. 크라사*는 25 °C 에서 단일 탄소원으로서의 2 % 의 *아비셀* 또는 *역세 파우더*를 함유하는 Vogel's 최소배지에서 성장하였다. *T. 레세이* 계통은 25 °C 에서 단일 탄소원으로서의 1 % 의 *아비셀* 또는 *역세 파우더*를 갖는 MA 배지에서 접종하였다(inoculate). 둘 다의 계통을 동일한 양의 분생자(conidia)[100 mL 배양물(culture)에 1 x 10⁶ /mL]로 접종하였다. 배양물에서 상이한 시점에서의 엔도글루카나제 활성도를, 제조자의 지시(manufacturer's instructions)(아일랜드에 위치한 Megazyme)에 따라 기질(substrate)로서 Azo-CM-셀룰로스를 사용하여 pH 4.5 에서 측정하였다.

도 4 는 *미스칸투스* 및 *아비셀*에서 성장한 *N. 크라사*의 전사 프로파일링(transcriptional profiling)을 나타낸 것이다. (A) *미스칸투스* 배양에서 발현의 차이점으로 나타내는 769 유전자의 계층적 군집 분석(Hierarchical clustering analysis). 어두운 셰이딩(Dark shading)은 상대적으로 보다 높은 발현(expression)을 나타내고, 밝은 셰이딩(light shading)은 상대적으로 낮은 발현을 나타낸다. 라인 1(Lane 1) : 16 시간 동안의 Vogel's 최소 배지 *N. 크라사* 배양물(Vogel 1956)의 발현 프로파일(Expression profile). 라인 2 : 16 시간 동안의 단일 탄소원으로서의 *미스칸투스*에서 성장한 배양물의 발현 프로파일. 라인 3, 4, 5 : 40 시간, 5 일 및 10 일 동안의 *미스칸투스*에서 성장한 배양물로부터의 발현 프로파일. 세 가지 군집(clusters)을 C1, C2 및 C3 으로 나타내었다. 대부분의 셀룰라아제 및 헤미셀룰라아제 유전자의 증가된 발현 수준을 나타낸 군집은 상자 모양을 나타내었다(boxed)(C3 군집). (B) *N. 크라사* *미스칸투스* 대 *아비셀* 성장한 배양물 사이의 발현 프로파일에서 공통부분(overlap)의 분석. 단백질의 분석 및 공통부분은 이중 질량 분석기(tandem mass spectrometry)(Bottom)로 *미스칸투스* 및 *아비셀*에서 성장한 *N. 크라사*의 배양 여과물에서 측정하였다. (C) 미

스칸터스 배양물에서 상대적인 발현 수준에서 증가됨을 나타낸 231 유전자의 기능범주(Functional category, FunCat) 다양한 분석(enrichment analysis)(Ruepp 2004). 분류되지 않은 그룹(the unclassified group)을 포함한 중요한 풍부함(enrichment)($p < 0.001$)을 나타낸 기능범주를 나타낸다.

도 5 는 최소배지(MM)에서 성장하는 동안 및 16 시간(16 hr), 40 시간(40 hr), 5 일(5 days) 및 10 일(10 days) 동안 미스칸터스(Miscanthus)에서 성장하는 동안의 셀룰라아제(A) 및 헤미셀룰라아제(B)를 코딩하는 *N. 크라사* 유전자의 상대적인 발현 수준을 나타낸다.

도 6 은 MS 에 의해 확인된 분비된 단백질을 코딩하는 유전자의 삭제를 포함하는 스트레인(strains)으로부터의 배양 상청액(culture supernatants)의 단백질 프로파일(protein profile) 및 효소 활성도(enzymatic activity)를 나타낸 것이다. (A) 7 일 동안 아비셀에서 성장한 야생형(wild type)과 비교한 16 결실 스트레인(deletion strains)의 배양 여과물(culture filtrates)에 나타난 단백질의 SDS-PAGE. 삭제 스트레인(strain)은 미스칸터스 및 아비셀 배양 여과물 둘 다에서의 질량 분석법(mass spectrometry)에 의한 단백질의 식별(identification)을 기반으로 선택되었다. 스트레인은 유전자 NCU 번호를 기초로 정리하였고, 야생형 스트레인(wild-type strain)은 FGSC 2489 이다. 삭제된 유전자에 해당하는 누락된 단백질 밴드(Missing protein bands)는 상자모양으로 나타났다(boxed). (B) 총 분비된 단백질, 아조-CMCase(azo-CMCase) 및 β -글루코시다아제 활성도 검정(실시에 5 참고)을 (A)로부터의 동일한 샘플을 사용하여 16 삭제 스트레인 및 야생형 부모 스트레인(wild-type parental strain)(FGSC 2489)에서 실행하였다. 활성도 및 단백질 농도를 야생형 농도와 비교하여 정규화시키고, 3 번의 생물학적인 측정(triplicate biological measurements)의 평균을 나타내었다. (C) (A) 에서의 동일한 샘플을 사용한 16 삭제 스트레인로부터의 배양 여과물의 셀룰라아제 활성도. 아비셀라아제 활성도(Avicelase activity)를 측정하기 위해 배양 여과물(Culture filtrates) 10 배 희석하였고 5 mg/mL 아비셀과 혼합하였다(실시에 5 참고). 40 °C 에서 배양하여 8 시간 후에 글루코스(흑색) 및 셀로비오스(cellobiose)(백색)를 측정하였다.

도 7 은 단일 탄소원으로서 아비셀에서 성장한 배양물로부터의 돌연변이체 분석을 기반으로 한 *N. 크라사* 분비된 단백질의 정체(identity)를 나타낸 것이다. 25 °C 에서 7 일 동안 100 mL 의 진탕 플라스크(shake flask s)에서 2 % 아비셀에서 성장한 WT *N. 크라사*(FGSC 2489)로부터의 분비된 단백질의 SDS-PAGE. 15 μ L 의 농축시키지 않는 배양 여과물을 크라이테리언 4-15 % 26-웰 겔 내에 로딩하였다(loaded). 내셔널 진단(National Diagnostics)으로부터는 포토 블루 세이프(Proto Blue Safe)(Coomassie)를 겔을 착색시키는데(stain) 사용하였다. 단백질 밴드를, 결실 스트레인(deletion strains)에서 분비된 단백질의 분석을 기반으로 한 도 6A 에 나타낸 바와 같은 이러한 연구에서 확인하였다.

도 8 은 Δ NCU04952 및 Δ NCU05137 에서 *cbh-1*(NCU07340) 및 *gh6-2* (NCU09680; CBHII)의 발현 및 분비된 단백질 프로파일을 나타낸 것이다. (A) WT, Δ NCU04952, 및 Δ NCU05137 에서 총 분비된 단백질의 SDS-PAGE. 배양물을 분생자(conidia)로부터의 아비셀에서 배양물을 성장시키고 30 시간, 2 일(48 시간) 및 3 일(72 시간)에 채취하였다(harvested)(실시에 5 참고). 라인 1-3 은 각각의 WT, Δ NCU04952 및 Δ NCU05137 스트레인(strains)로부터의 아비셀에서 30 시간 성장 후의 20x 농축시킨 배양 여과물이다. 라인 4-6 은 각각의 WT, Δ NCU04952 및 Δ NCU05137 스트레인로부터의 아비셀에서 2 일 성장 후의 농축시키지 않은 배양 여과물이다. 라인 7-9 는 각각의 WT, Δ NCU04952 및 Δ NCU05137 스트레인으로부터의 3 일후의 농축시키지 않은 배양 여과물이다. (B) 아비셀에서 성장하는 동안의 WT, Δ NCU04952 및 Δ NCU05137 스트레인에서 *cbh-1*(NCU07340; CBHI) 및 *gh6-2*(NCU09680; CBHII)의 RT-PCR. WT 및 삭제 스트레인은 분생자로부터의 아비셀에서 성장하였고, 48 시간 및 72 시간에 채취하였다(실시에 5 참고). 단일 탄소원으로서의 수크로스를 갖는 최소배지(MM) 배양물(Vogel 1956)을 16 시간 동안 성장시켰다[유사한 발달 시점(developmental time point)]. *cbh-1* 및 *gh6-2* 의 폴드 인덕션(fold induction)은, 모든 샘플의 대조군으로서 사용된 액틴 유전자 발현(actin gene expression)을 갖는 MM 조건 하에서 이러한 유전자의 발현과 관련되어 있다.

도 9 는 *N. 크라사*에서 식물 세포 벽 디컨스트럭션(deconstruction)의 모델을 나타낸 것이다. 도입(induction) : 낮은 수준에서 발현된 세포의 효소(Extracellular enzymes)는 대부분 분비된 식물 세포 벽 분해효소(plant cell wall degrading enzymes)를 코딩하는 유전자의 발현 수준을 극적으로 증가시키기 위해 단일 *N. 크라사*의 2 차 대사물질(secondary metabolites)를 생성한다. 이용(utilization) : 세포 벽 분해산물(cell wall degradation products)의 전좌(translocation)에 특이적인 세포의 효소 및 운반체(transporters)는 *N. 크라사*가 성장에 대한 식물 세포 물질을 활용할 수 있게 한다. 몇몇의 세포의 단백질(NCU05137, NCU05057, 및 NCU04952)은 활용 단계에서 셀룰라아제(cellulases) 및 헤미셀룰라아제(hemicellulase)의 유전자 발현을 조절하는 대사물질(metabolites)을 생성할 수 있다 ; 이중의 육각형(double hexagon)(셀로비오스), 이중의 오각형

(double pentagon)(자이로비오스, xylobiose), 육각형(글루코스) 및 오각형(자일로스). 그려진 식물 세포 벽-분해 효소는 CBH(I), CBH(II), EG2, EG1, EG6 및 자일라나아제(xylanase)를 포함한다. 추가적인 셀룰로이스 분해효소(cellulolytic enzymes)를 나타내지 않는다. 화살표의 두께는 반응의 상대적인 강도를 나타낸다.

도 10 은 *S. 서머피일* 단백질 서열(*S. thermophile* protein sequences)의 데이터베이스에 대한 *N. 크라사* 추정상의 운반체(putative transporters)의 서열 또는 *N. 크라사* 단백질 서열의 데이터베이스에 대한 *S. 호열성* 추정상의 운반체의 서열을 BLAST 에서 검색한 결과를 나타낸 것이다.

도 11 은 NCU08114 가 결핍된 *N. 크라사* 스트레인의 성장 표현형(growth phenotype)을 나타낸 것이다. (A) 탄소원으로서 결정질 셀룰로스(crystalline cellulose)를 갖는 성장의 3 일 후의 WT(왼쪽) 및 Δ NCU08114(오른쪽) *N. 크라사* 스트레인의 진탕 플라스크(Shaker flasks). (B) 각각의, 16 시간 또는 28 시간 동안, 탄소원으로서 수크로스 또는 결정질 셀룰로스를 이용하여 성장한 *N. 크라사* 배양물로부터의 평균 알라마블루[®] 형광(Mean Alamar Blue[®] fluorescence). 형광을 100 % 로 WT 에 맞추어 정규화시켰다(normalize). 에러 바(Error bars)를 세 번의 생물학적 반복(biological replicates)으로부터의 측정(measurements) 사이의 표준 편차(standard deviation)이다. NCU00801이 결핍된 *N. 크라사*는, 단당류 운반체에 의해 순차적으로 계속되는 셀로덱스트린에서 글루코스로 가수분해하는(Scarborough 1973) *N. 크라사* 분비된 β -글루코시다아제의 명백한 표현형을 갖지 않는다(Tian 등의, 2009). 이러한 소비(consumption)의 우회 루트(alternate route)는 이러한 삭제 라인(deletion lines)에서 셀로덱스트린 운반체의 결핍의 과소평가(underestimate)로 이어진다.

도 12 는 (A) NCU00130 와 함께 NCU00801, NCU05853 또는 NCU08114 가 발현된 *S. 세레비시아* 스트레인(*S. cerevisiae* strains)에 대한 셀로비오스(cellobiose)의 소비 ; (B) NCU00130 와 함께 NCU00801, NCU05853 또는 NCU08114 가 발현된 *S. 세레비시아* 스트레인에 대한 셀로트리오스(cellobiose)의 소비 ; (C) NCU00130 와 함께 NCU00801, NCU05853 또는 NCU08114 가 발현된 *S. 세레비시아* 스트레인에 대한 셀로테트라오스(cellobiose)의 소비 ; 및 (D) NCU00130 와 함께 NCU00801, NCU05853 또는 NCU08114 가 발현된 *S. 세레비시아* 스트레인에 대한 셀로헥사오스(cellobiose)의 소비를 나타낸 것이다.

도 13 은 NCU08114 또는 NCU00801 이 결핍된 *N. 크라사* 스트레인에 의한 셀로덱스트린의 소비를 나타낸 것이다. 나타낸 *N. 크라사* 스트레인을 15 분 동안 90 μ L 의 각각의 당으로 배양하였다. 바(Bar)는 두 개의 독립적인 실험으로부터의 배양 후에 상청액에 남아있는 당의 평균 농도를 나타낸 것이다. 에러 바는 이들 실험 사이의 표준 편차이다.

도 14 는 NCU00801/*cbt1* 이 발현된 *S. 세레비시아* 스트레인(*S. cerevisiae* strain)에 의한 셀로비오스 수송을 나타낸 것이다. CBT 1 이 있을 때(○) 또는 없을 때(●)에 의한 셀로비오스 수송을 나타낸 것이다. 스트레인 둘 다는 세포 내의 β -글루코시다아제, NCU00130 가 발현되었다. 셀로비오스의 초기 농도는 50 μ M 이다. 모든 값은, 이러한 측정 사이의 표준 편차를 나타낸 에러 바와 함께 두 측정 사이의 평균이다.

도 15 는 CBT1 및 CBT2 에 융합시킨(fused) GFP 의 위치(localization) 및 정량(quantification)을 나타낸 것이다. (A) 이들의 C-말단에 GFP 에 융합시킨 *cbt1*(왼쪽) 또는 *cbt2*(오른쪽)에 발현된 *S. 세레비시아* 스트레인의 이미지. (B) 셀로비오스 운반체가 없는 또는 이들의 C-말단에 GFP 에 융합시킨 *cbt1* 또는 *cbt2* 가 발현된 효모 스트레인의 GFP 형광(fluorescence). 값은 세 번의 생물학적 반복으로부터의 평균이고, 에러 바는 이들의 반복 사이의 표준 편차를 나타낸다.

도 16 은 *S. 세레비시아*에서 발현된 *N. 크라사* 수송 시스템에 의한 셀로덱스트린 수송을 나타낸 것이다. (A) 유전자 NCU00801(*cbt1* 으로 명명, ○), NCU08114 (*cbt2* 로 명명, ▼)가 발현되거나, 또는 운반체가 없는 경우(●)의 효모 스트레인의 셀로비오스-매개 성장(Cellobiose-mediated growth). 모든 스트레인은 또한 세포 내의 β -글루코시다아제, NCU00130 를 발현하였다. 대표적인 실험을 나타내었다. 세 가지의 독립적인 실험으로부터의 성장비율을 하기와 같다 : *cbt1*, $0.0341 \pm 0.0010 \text{ hr}^{-1}$; *cbt2*, $0.0131 \pm 0.0008 \text{ hr}^{-1}$; 운반체가 없는 경우, $0.0026 + 0.0001 \text{ hr}^{-1}$. (B) 셀로트리오스 및 셀로테트라오스에서 효모 스트레인의 성장. 범례(legend)에 열거된 운반체 뿐만 아니라 세포내의 β -글루코시다아제, NCU00130 가 발현된 스트레인은 단일 탄소원으로서 제공된 0.5% (w/v)의 셀로트리오스(G3) 또는 셀로테트라오스(G4)와 함께 성장시켰다. 대표적인 실험을 나타내었다. 세 가지 독립적인 실험으로부터의 성장율(Growth rates)은 하기와 같다 : *cbt1* 셀로트리오스, $0.0332 + 0.0004 \text{ hr}^{-1}$; *cbt1* 셀로테트라오스 $0.0263 + 0.0020 \text{ hr}^{-1}$; 운반체가 없는 셀로트리오스, $0.0043 + 0.0015 \text{ hr}^{-1}$; *cbt2* 셀로트리오스, $0.0178 + 0.0005 \text{ hr}^{-1}$; *cbt2* 셀로테트라오스 $0.0041 + 0.0003 \text{ hr}^{-1}$; 운반체가 없는 셀로테트

라오스, $0.0031 + 0.0008 \text{ hr}^{-1}$. (C) 글루코스는 정제된 NCU00130 에 의해 셀로비오스(G2), 셀로트리오스(G3) 및 셀로테트라오스(G4) 가수분해로부터 생산되었다. 세 가지의 독립적인 측정의 평균 및 표준 편차를 나타내었다. 효소(2 nmol)가 없는 배양에서의 잔류 글루코스는 나타낸 값으로부터 뺐다.

도 17 은 글루코스에서 운반체가 없거나(●) 또는 *cbt1*(○), *cbt2*(▼)가 발현된 *S. 세레비시아* 스트레인의 성장을 나타낸 것이다. 모든 스트레인은 β -글루코시다아제, NCU00130 이 발현되었다. 대표적인 실험을 나타내었다.

도 18 은 250 mL 의 플라스크에서 *S. 세레비시아* 스트레인의 셀로비오스-매개 성장을 나타낸 것이다. 값은 β -글루코시다아제, NCU00130, *cbt1* 또는 *cbt2* 가 발현된 효모 스트레인, 또는 어떠한 운반체도 결핍된 스트레인 발현 NCU00130 의 두 번의 복제 배양물(two replicate cultures) 사이의 평균 OD 를 나타낸 것이다. 에러 바는 복제 사이의 표준 편차를 나타낸다.

도 19 는 CBT1 및 CBT2 에 의한 셀로비오스 수송의 반응속도(kinetics)를 나타낸 것이다. 셀로비오스 수송의 비율은 *cbt1* 또는 *cbt2* 가 발현된 효모 스트레인에 의한 셀로비오스 농도의 작용으로서 측정하였다. 수송 비율(transport rate)은 풍부한 운반체(transporter abundance)에 대해서 정규화시켰다(normalize).

도 20 은, 셀로비오스, 셀로트리오스 또는 셀로테트라오스에서 성장한 x-축에 나타낸 *네우라스포라* 유전자(*Neurospora genes*)의 조합을 발현된 *S. 세레비시아*의 능력을 나타낸 것이다.

도 21 은 *cbt1* 또는 *cbt2* 를 운반하는 스트레인에서 셀로비오스 수송을 위한 셀로텍스트린의 경쟁(competition)을 나타낸 것이다. 각각의 분류되지 않은 당(unlabeled sugar)의 5-폴드 초과량(5-fold excess)은 [^3H]-셀로비오스 수송의 검정 동안에 포함되었다. CBT1 또는 CBT2 의 기질(Substrate)은 결합을 위한 경쟁에 의해 [^3H]-셀로비오스 수송 비율을 감소시켰다. 바는 세 번의 반복으로부터의 평균을 나타낸다. 에러 바는 이러한 반복 사이의 표준 편차를 나타낸다. 값은 100 까지의 경쟁 당 없이 [^3H]-셀로비오스 수송 비율에 맞추어 정규화시켰다.

도 22 는 정제된 NCU00130 의 SDS-PAGE 겔을 나타낸 것이다. 라인 1, kDa 에서의 단백질 분자량 기준. 라인 2, 니켈-NTA 레진(nickel-NTA resin) 상에서 여과한 후의 NCU00130. kDa 의 분자량은 왼쪽에 나타내었다.

도 23 은 셀로비오스 운반체 NCU08114 및 NCU00801 의 최고 가능성을 갖는 계통발생론적 분석(maximum likelihood phylogenetic analysis)을 나타낸 것이다. *S. 세레비시아* HXT1 및 *K. 락티스* LACP 를 예외로 하는, 나타낸 단백질을 코딩하는 모든 유전자는, 균류(fungus)가 식물 세포 벽 물질 또는 셀로비오스와 접촉하여 쓰이기 시작할 때, 발현 수준을 증가시키기 위한 것으로 알려졌다(Tian 등의, 2009; Noguchi 등의, 2009; Wymelenberg 등의, 2010; Martin 등의, 2010). *S. 세레비시아* HXT1, 낮은 친화도(low affinity)의 글루코스 운반체(Reifenberger 등의, 1997)는 외집단(outgroup)으로서 사용되었다.

도 24 는 셀로비오스의 발효(fermentation), *N. 크라사*로부터의 셀로비오스 수송 시스템이 발현된 *S. 세레비시아*에 의한 동시의 당화(saccharification) 및 셀룰로스의 발효를 나타낸 것이다. (A) 에탄올에 대한 셀로비오스 발효. 에탄올은 CBT1 을 갖거나(●) 또는 CBT1 을 갖지 않은(○) 효모 스트레인에 의해 생산되었다. CBT1 을 갖거나(▼) 또는 CBT1 을 갖지 않은(△) 효모 스트레인을 사용한 발효 반응 동안에 셀로비오스 농도. (B) CBT1 을 갖거나, 갖지 않은 효모 스트레인을 사용한 SSF. CBT1 을 갖는 스트레인의 존재 하에서 셀로비오스(●) 및 글루코스(▼)의 농도, 및 CBT1 이 결핍된 스트레인의 존재 하에서 셀로비오스(○) 및 글루코스(△)의 농도. 참고로 0.1 mg/mL 셀로비오스 = $292 \text{ }\mu\text{M}$. (C) CBT1 을 갖는(●) 또는 CBT1 을 갖지 않는(○) 스트레인을 사용한 SSF 동안 생산된 에탄올. 모든 패널(panels)에서, 값은 세 번의 생물학적인 반복의 평균이다. 에러 바는 이러한 반복 사이의 표준 편차이다. 모든 스트레인은 세포 내의 β -글루코시다아제, NCU00130 을 발현시켰다.

도 25 는 효모에 의한 셀룰로스의 동시의 당화(simultaneous saccharification) 및 발효 동안에 곰팡이(filamentous fungi)로부터의 셀로텍스트린 수송 경로(transport pathways)의 사용을 나타낸 것이다. 셀로텍스트린(Cdes) 수송 경로(흑색)은 셀로텍스트린 운반체(CBT) 및 세포 내의 β -글루코시다아제(βG)를 포함한다. 표준 효모(standard yeast)에 나타낸 당 이화 경로(sugar catabolism pathway)는 육탄당 운반체(hexose transporters, HXT)를 포함한다. SSF 에서 셀룰라아제(GH) 및 세포외 β -글루코시다아제(βG) 둘 다는 사용될 수 있다.

도 26 은 기능적으로 중요한 NCU00801 및 NCU0811 에서의 잔여물(residues)을 나타낸 것이다. (A) *cbt1*/NCU00801 의 Ala-스캔(Ala-scan). (B) *cbt1*/NCU00801 의 폴리펩티드 서열[표시된 중요한 잔기(important residues marked)]. (C) *cbt2*/NCU08114 의 폴리펩티드 서열(표시된 중요한 잔기).

도 27 은 *P. 스티피티스*(*P. stipitis*)로부터의 셀로비오스 운반체가 발현된 *S. 세레비시아* 스트레인의 비교를 나타낸 것이다. (A) β -글루코시다아제가 발현된 *S. 세레비시아* 스트레인 및 셀로비오스 운반체 NCU00801, NCU08114 및 NCU05853 의 오솔로그(orthologs)의 세포 성장. (B) *P. 스티피티스*로부터의 셀로비오스 운반체의 비교 : β -글루코시다아제 및 셀로비오스 운반체가 발현된 *S. 세레비시아* 스트레인의 세포 성장. (C) *P. 스티피티스*로부터의 셀로비오스 운반체의 비교 : β -글루코시다아제 및 셀로비오스 운반체가 발현된 *S. 세레비시아* 스트레인에 의한 자일로스 소비 및 에탄올 생산.

도 28 은 셀로비오스 운반체 오솔로그의 배열을 나타낸 것이다. (A) 테스트된 조건 하에서 운반체 기능을 갖는 것으로 나타나지 않은 하나를 포함하는 셀로비오스 운반체 오솔로그의 배열. (B) 운반 기능을 갖는 셀로비오스 운반체 오솔로그의 배열. (C) NCU00801 및 NCU08114 의 배열.

도 29 는 NCU00801 및 NCU08114 의 상동성 모델(homology models)에서 표시된 기능적으로 중요한 모티프(motifs)를 나타낸 것이다. (A) NCU00801 의 상동성 모델에서 셀로비오스 운반체 모티프의 위치. 모티프는 흑색으로 강조되었다. 모티프 [LIVM]-Y-[FL]-x(13)-[YF]-D 를 패널(panels) a 및 b 에 나타내었다. 모티프 [YF]-x(2)-G-x(5)-[PVF]-x(6)-[DQ]를 패널 c 및 d 에 나타내었다. 모티프 G-R-[RK] 를 패널 e 및 f 에 나타내었다. 모티프 R-x(6)-[YF]-N 를 패널 g 및 h 에 나타내었다. 모티프 WR-[IVLA]-P-x(3)-Q 를 패널 i 및 j 에 나타내었다. 모티프 P-E-S-P-R-x-L-x(8)-A-x(3)-x(2)-Y-H 를 패널 k 및 l 에 나타내었다. 모티프 F-[GST]-Q-x-S-G-N-x-[LIV]를 패널 m 및 n 에 나타내었다. 모티프 L-x(3)-[YIV]-x(2)-E-x-L-x(4)-R-[GA]-K-G 를 패널 o 및 p 에 나타내었다. 참고 : 이러한 이미지에서, 트랜스멤브레인 헬릭스 6 및 7 을 연결하는 몇몇의 잔기는, 그들이 구멍을 가렸기 때문에(as they occlude the pore) 명확하게 제거되었다. (B) NCU08114 상동성 모델에서 셀로비오스 운반체 모티프의 위치. 모티프는 흑색으로 강조되었다. 모티프 [LIVM]-Y-[FL]-x(13)-[YF]-D 를 패널 a 및 b 에 나타내었다. 모티프 [YF]-x(2)-G-x(5)-[PVF]-x(6)-[DQ]를 패널 c 및 d 에 나타내었다. 모티프 G-R-[RK]를 패널 e 및 f 에 나타내었다. 모티프 R-x(6)-[YF]-N 을 패널 g 및 h 에 나타내었다. 모티프 WR-[IVLA]-P-x(3)-Q 를 패널 i 및 j 에 나타내었다. 모티프 P-E-S-P-R-x-L-x(8)-A-x(3)-x(2)-Y-H 를 패널 k 및 l 에 나타내었다. 모티프 F-[GST]-Q-x-S-G-N-x-[LIV]를 패널 m 및 n 에 나타내었다. 모티프 L-x(3)-[YIV]-x(2)-E-x-L-x(4)-R-[GA]-K-G 를 패널 o 및 p 에 나타내었다. 참고 : 이러한 이미지에서, 트랜스멤브레인 헬릭스 6 및 7 을 연결하는 몇몇의 잔기는, 그들이 구멍을 가렸기 때문에 명확하게 제거되었다.

도 30 은 플라스미드 발현의 구성에서 사용된 클로닝 과정(cloning process)을 나타낸 것이다 : (A) 추정상의 운반체(putative transporters) 및 (B) 운반체-GFP 융합 단백질(transporter-GFP fusion proteins).

도 31 은 글루코스-흡수 활성도(glucose-uptake activity)를 갖는 것으로 확인된 추정상의 운반체(putative transporters)의 오탄당 수송 활성도를 나타낸 것이다.

도 32 는 글루코스-흡수 활성도(glucose-uptake activity)를 갖지 않는 것으로 확인된 추정상의 운반체의 오탄당 수송 활성도를 나타낸 것이다.

도 33 는 NCU00821(AN25), STL12/XUT6(Xyp29) 및 XUT1(Xyp32)의 오탄당 흡수를 나타낸 것이다. A 는 자일로스 흡수를 나타내고 B 는 아라비노스 흡수를 나타낸 것이다.

도 34 는 STL12 XUT6 가 발현된 *S. 세레비시아*에 의한 ¹⁴C-표지된(labeled) 당 흡수를 나타낸 것이다(Xyp29).

도 35 는 GFP 형광에 의해 관찰된 바와 같은 *S. 세레비시아* 세포에서 발현된 운반체의 위치(localizations)를 나타낸 것이다. 첫 번째 줄의 왼쪽에서 오른쪽 : NCU00821-GFP 형광, NCU00821 핵(nuclei) ; 두 번째 줄의 왼쪽에서 오른쪽 : STL12 XUT6-GFP 형광, STL12/XUT6 핵.

도 36 은 (A) NCU00821(AN25), (B) STL12/XUT6(Xyp29) 및 (C) XUT1(Xyp32)으로 발현된 완충되지 않은 세포 현탁액에 말토오스의 첨가 상에서 pH 에서의 효과를 나타낸 것이다. 검정색 화살표는 말토오스를 첨가하였을 때의 시점(time points)을 나타낸 것이다.

도 37 은 NCU00821, STL12/XUT6 및 XUT1 의 공수송체 검정(syraporter assay)의 결과를 나타낸 것이다. (A) 자일로스에 대한 NCU00821. (B) 아라비노스에 대한 NCU00821. (C) 아라비노스에 대한 XUT1. (D) 자일로스에 대한 NCU00821. (E) 아라비노스에 대한 STL12 XUT6. (F) 자일로스에 대한 STL12/XUT6. 검정색 화살표는 말

토오스를 첨가하였을 때의 시점을 나타낸 것이다.

도 38 은 운반체의 과잉발현(overexpression)의 표현형 분석(phenotypic analyses)를 나타낸 것이다. (A) 는 OD 를 나타낸 것이고, (B) 는 자일로스 농도를 나타낸 것이고, (C) 는 0.5 % 자일로스-함유 배지에서의 자일로스 소비를 나타낸 것이다. (D) 는 OD 를 나타낸 것이고, (E) 는 자일로스 농도를 나타낸 것이고, (F) 는 0.5 % 자일로스-함유 배지에서의 자일로스 소비를 나타낸 것이다. (G) 는 다복제 플라스미드(multicopy plasmid), pRS424 에서 도입된 오타당 운반체를 함유하는 *S. 세레비시아*의 성장 곡선을 나타낸 것이다.

도 39 는 이중 운반체(heterologous transporters)의 클로닝(cloning)에 사용된 플라스미드의 맵(map)을 나타낸 것이다.

도 40 은 오타당 운반체 오솔로그가 발현된 *S. 세레비시아* 스트레인에 의한 당-흡수 검정의 결과를 나타낸 것이다.

도 41 은 Clustal W (1.81)에 의한 오타당 운반체 오솔로그의 서열 정렬(sequence alignments)을 나타낸 것이다. (A) 자일로스 운반체 오솔로그의 정렬. (B) 아라비노스 운반체의 정렬. (C) 자일로스 및 아라비노스 운반체의 정렬. 컨센서스 키(Consensus key) : * - 단일, 완전히 보존된 잔기(single, fully conserved residue) ; : - 강한 그룹(strong groups)의 보존(conservation) ; - 약한 그룹(weak groups)의 보존.

도 42 는 자일로스-이용 효소(xylose-utilizing enzymes)를 발현시키기 위한 제작한(engineered) 상이한 *S. 세레비시아* 스트레인을 나타낸 것이다.

도 43 은 자일로스 이용 경로 효소(xylose utilization pathway enzymes)를 함유하는 동일한 카세트(cassettes)가 발현된 상이한 백그라운드(backgrounds)의 세 가지 *S. 세레비시아* 스트레인의 자일로스 대사(xylose metabolism)(자일로스 소비 및 에탄올 생산 등과 같이 모니터 한 바와 같이)를 나타낸 것이다.

도 44 는 자일로스 이용 경로 효소를 함유하는 동일한 카세트를 발현하는 상이한 백그라운드의 세 가지 *S. 세레비시아* 스트레인의 자일로스-흡수 비율 및 대사물 수득률(metabolite yield)을 나타낸 것이다.

도 45 는 다양한 조건 하에서 *S. 세레비시아* 스트레인 DA24 에 의한 자일로스 발효(xylose fermentation)를 나타낸 것이다. (A) 진탕 플라스크에서 40 g/L 의 자일로스, (B) 진탕 플라스크에서 80 g/L 의 자일로스 및 (C) 바이오리액터(bioreactor)에서 80 g/L 의 자일로스. 부호 : 자일로스(■), 에탄올(◆) 및 OD₆₀₀(●).

도 46 은 (A) *S. 세레비시아* DA24 및 (B) *P. 스티피티스*(*P. stipitis*) 사이의 자일로스 소비 및 에탄올 생산의 비교를 나타낸 것이다. 부호 : 자일로스(■), 에탄올(◆) 및 OD₆₀₀(●).

도 47 은 유전자 조작된(engineered) *S. 세레비시아*에서의 자일로스 대사 상에서 XYL2 과잉-발현 수준(over-expression levels)를 테스트하는데 사용된 실험 설계(experimental design)를 나타낸 것이다.

도 48 은 유전자 조작된(engineered) 자일로스-발효 *S. 세레비시아* 내로의 추가적인 XYL2 통합(integration)(즉, 증가된 XYL2 발현 수준)의 효과를 나타낸 것이다.

도 49 는 유전자 조작된 *S. 세레비시아* 에 의한 자일로스에서 XYL2 및 XYL3추가적인 동시의 과잉-발현(additional simultaneous over-expression)의 효과를 나타낸 것이다.

도 50 은 상이한 수준의 자일로스-발효 효소(xylose-fermenting enzymes)가 발현된 *S. 세레비시아* 스트레인을 나타낸 것이다.

도 51 은 유전자 조작된 *S. 세레비시아* 에 의한 차별된 XYL1 발현(differential XYL1 expression)의 효과를 나타낸 것이다.

도 52 는 상이한 환원효소(reductases)를 갖는 동일한 XYL2 및 XYL3 과잉-발현(over-express)을 위해 유전자 조작된 *S. 세레비시아* 스트레인을 나타낸 것이다(XYL1 대 GRE3).

도 53 은 40 g/L 의 자일로스에서 성장한 유전자 조작된 *S. 세레비시아* 에 의한 자일로스 발현에서 과잉-발현된(over-expressing) XYL1 대 GRE3 의 효과를 나타낸 것이다.

도 54 은 80 g/L 의 자일로스에서 성장한 유전자 조작된 *S. 세레비시아* 에 의한 자일로스 발현에서 과잉-발현된(over-expressing) XYL1 대 GRE3 의 효과를 나타낸 것이다.

도 55 는 상이한 야생형 LAD 효소의 열 특성(thermal property) 및 pH-의존적인 특성(pH-dependent

properties)을 나타낸 것이다 : LAD(■), tLAD(◆) 및 pcLAD(●). (A) 온도-의존적인 촉매 활성도 (Temperature-dependent catalytic activities), (B) 오버타임(over time) 동안의 50 °C 에서의 열의 불활성화 (Thermal inactivation) 및 (C) pH-의존적인 촉매 활성도. 에러바(Error bars)는 평균의 표준 오차(standard error)를 나타낸다(n = 3).

도 56 은 *N. 크라사* 및 *P. 스티피티스*(psXDH)로부터의 XDH 의 정렬(alignment)을 나타낸 것이다.

도 57 은 *N. 크라사* LAD 및 XDH 의 pH 비율 프로파일(pH rate profiles)의 비교를 나타낸 것이다. LAD 의 특성으로부터의 획득한 데이터는 일반적인 완충용액 MES/트리스/글리신에서 수행하였고, 낮은 pH 값에 대해 일반적인 완충용액 아세트산/MES/트리스에서 수행한 ncXDH[폐쇄된 삼각형(closed triangles)] 및 ncLAD[폐쇄된 원형(closed circles)] 에 대한 데이터는 중복되었다.

도 58 은 *박테리오이드 스테코리스*(*Bacterioides stercoris*, *BtXI*), *비피도박테리움 롱검*(*Bifidobacterium longum*, *BfXI*) 및 코돈-최적화된 *BtXI*(codon-optimized *BtXI*)에 대한 *BtXI*O 로부터의 자일로스 이성질화효소(xylose isomerase enzyme)를 갖는 형질전환된(transformed) *S. 세레비시아* 스트레인 L2612 에 의한 에탄올 생산을 나타낸 것이다. XI 유전자는 pRs424TEF 벡터 내로 클론되었다(clone).

도 59 는 벡터 pRS403TEF 에 의해 이의 게놈(genome) 내로 통합된(integrated) *BtXI* 를 갖는, *S. 세레비시아* 스트레인 D452-2 에 의한 자일로스 소비 및 에탄올 생산을 나타낸 것이다. 플라스미드로부터 *BtXI* 가 발현된 *S. 세레비시아* 스트레인 L2612 에 의한 자일로스-발효와 비교하였다.

도 60 은 통합된 *BtXI* 및 발현된 *XYL2* 또는 *XYL3* 또는 *XYL2* 및 *XYL3* 을 함유하는 *S. 세레비시아* 스트레인에 의한 자일로스 발효를 나타낸 것이다.

도 61 은 자일로스의 효율적인 대사작용을 위해 오탄당 인산경로(pentose phosphate pathway)에서 포함된 *GND1* 과 같은 과잉-발현된 효소를 유전자 조작한 *S. 세레비시아* 에서 *XYL3* 발현의 필요성을 나타낸 것이다.

도 62 는 발효 파라미터(fermentation parameters)에서 *대장균*, *S. 세레비시아* 및 *P. 스티피티스*에서 NCU09705 상동성의 과잉-발현의 효과를 나타낸 것이다. (A) 셀로비오스 소비, 성장 및 에탄올 생산 ; 및 (B) 에탄올 수득률 및 생산성에서 *galM*, *GALIO-Sc*, *GALIO-PS*, *YHR210C* 및 *YNR071C* 의 과잉-발현.

도 63 은 곰팡이(*N. 크라사*)로부터의 셀로텍스트린 동화 경로(assimilation pathway) 및 *S. 세레비시아* 내의 자일로스-발효 효모 *P. 스티피티스*로부터의 변형된 자일로스 대사경로의 통합을 통해서 글루코스 리프레션(repression)없이 셀로비오스 및 자일로스의 동시적인 공동-발효(co-fermentation)를 가능하게 하는 실험 설계를 나타낸 것이다. (A) 두 개의 대사작용이 가능하지 않는 당(non-metabolizable sugars)(셀로비오스 및 자일로스)을 발효시킬 수 있는 효모 스트레인을 유전자 조작하기 위한 스트레인 개선 전략(strain improvement strategy). 셀로텍스트린 동화 경로는 *N. 크라사*로부터의 셀로텍스트린 운반체(NCU00801) 및 세포 내의 β -글루코시다아제(NCU00130)으로 이루어져 있다. 변형된 자일로스 대사 경로는 자일로스 환원효소 동질효소(xylose reductase isozymes)(야생형 XR 및 돌연변이체 XR R.276H), 자일리톨 탈수소효소(*XYL2*) 및 자일루로키나아제(xylulokinase)(*XKS1*)를 사용하였다. (B) 유전자 조작된 *S. 세레비시아*에 의한 글루코스 및 자일로스를 함유하는 당 혼합물의 발효 프로파일(Fermentation profile)은 이러한 연구에서 진전되었다(develop). 글루코스가 고갈된 후에만 자일로스 발효가 시작되기 때문에 글루코스 발효는 자일로스 발효를 억제한다. (C) 유전자 조작된 *S. 세레비시아*에 의한 셀로비오스 및 자일로스를 함유하는 당 혼합물의 발효 프로파일은 이러한 실험에서 진전되었다. 탄소원(carbon source)이 다른 물질의 소비를 억제하지 않음에 따라 셀로비오스 및 자일로스는 동시에 사용되었다.

도 64 는 플라스미드 구성(plasmid construction)에 대한 도식(scheme)을 나타낸 것이다. pRS425 셔틀벡터(shuttle vector)는 DNA 조립 방법(DNA assembler method)(Shao 등의, 2009)를 사용한 셀로비오스 운반체 및 β -글루코시다아제 유전자를 조립한 다음에 선 모양으로 만들었다(linearized).

도 65 는 시간의 함수(function)으로서의 *S. 세레비시아* 스트레인 (A) SL01, (B) SL04, (C) SL02, (D) SL05, (E) SL03, (F) SL06 및 (G) SL00 에 의한 4 % 셀로비오스 및 5 % D-자일로스의 공동-발효 동안의 셀로비오스(■), 글루코스(●), D-자일로스(▲), 에탄올(▼), 및 바이오매스(□)의 농도의 변화를 나타낸 것이다.

도 66 은 시간의 함수로서 도표를 나타낸, 진탕-플라스크(a,b) 또는 바이오리액터(bioreactors)(c, d)에서 셀로비오스-자일로스 혼합물에서 성장한 *S. 세레비시아* 스트레인 SL01(a, b) 및 SL00(b, d)에서의 셀로비오스(■), 글루코스(●), D-자일로스(▲), 에탄올(▼), 및 바이오매스(□)의 농도의 변화를 나타낸 것이다.

도 67 은 시간의 함수로서 도표로 나타낸, 바이오리액터에서 5 g/L 글루코스-40 g/L 셀로비오스-50 g/L 자일로스 혼합물(a, b) 또는 10 g/L 글루코스-40 g/L 셀로비오스-50 g/L 자일로스 혼합물(c, d)를 함유하는 배지에서 성장한 *S. 세레비시아* 스트레인 SL01(a, c) 및 SL00(b, d)에서의 셀로비오스(■), 글루코스(●), D-자일로스(▲), 에탄올(▼), 및 바이오매스(□)의 농도의 변화를 나타낸 것이다.

도 68 은 (A) NCU00801, (B) NCU00809 및 (C) NCU08114 가 발현된 β -글루코시다아제(NCU00130)-함유 *S. 세레비시아* 스트레인에 의한 셀로비오스 활용의 비교를 나타낸 것이다. 부호 : 셀로비오스(■), 에탄올(◆) 및 OD₆₀₀(●).

도 69 는 두 가지의 당의 다양한 농도를 함유하는 혼합물에서 성장한 *S. 세레비시아* 스트레인 DA24-16BT3 에 의한 셀로비오스 및 자일로스의 공동-발효를 나타낸 것이다 : (A) 20 g/L(각각)의 셀로비오스 및 자일로스, (B) 30 g/L(각각)의 셀로비오스 및 자일로스, 및 (C) 40 g/L(각각)의 셀로비오스 및 자일로스. 부호 : 셀로비오스(▲), 자일로스(■), 에탄올(◆) 및 OD₆₀₀(●).

도 70 은 *S. 세레비시아* 스트레인 DA24-16BT3 에 의한 셀로비오스 및 자일로스의 공동-발효의 상승효과(synergistic effect)를 나타낸 것이다. 부호 : 셀로비오스(▲), 자일로스(■), 에탄올(◆) 및 OD₆₀₀(●) : (A) 40 g/L 셀로비오스, (B) 40 g/L(각각)의 셀로비오스 및 자일로스, 및 (C) 40 g/L 자일로스.

도 71 은 *S. 세레비시아* 스트레인 DA24-16BT3 및 야생형 *P. 스티피티스* 스트레인에 의한 글루코스, 셀로비오스 및 자일로스의 공동-발효를 나타낸 것이다. 부호 : 셀로비오스(▲), 자일로스(■), 에탄올(◆), OD₆₀₀(●) 및 글루코스(▼). (A) DA24-16BT3 및 (B) *P. 스티피티스*.

도 72 는 *S. 세레비시아* 스트레인 DA24-16BT3 에 의한 셀로비오스 및 자일로스의 c-발효 동안 셀로트리오스 및 셀로테트라오스 축적(accumulation)을 시사하는, 각각의 시점으로부터의 HPLC 크로마토그램(HPLC chromatograms)을 나타낸 것이다.

도 73 은 셀로비오스 및 자일로스의 공동-발효 동안의 *S. 세레비시아* 스트레인 DA24-16BT3 에 의한 22 시간 발효 후의 발효 배지에서 셀로텍스트린 축적을 보여주는 HPAEC 분석을 나타낸 것이다. [G1 : 클루코스, G2 : 셀로비오스, G3 : 셀로트리오스, G4 : 셀로테트라오스, 및 G5 : 셀로펜타오스(cellopentaose)].

도 74 는 40 g/L(각각)의 셀로비오스 및 자일로스의 공동-발효 동안에, 다중-복제 플라스미드(multi-copy plasmid)에서의 (A) NCU00801 의 통합된 복제(integrated copy) 및 (B) NCU00801 이 발현된 *S. 세레비시아* 형질전환체(*S. cerevisiae* transformants)에 의한 당 이용의 비교를 나타낸 것이다. 부호 : 셀로비오스(▲), 자일로스(■), 에탄올(◆), 및 OD₆₀₀(●).

도 75 는 (A) 연구에서 사용된 두 가지 상이한 *S. 세레비시아* 스트레인 : DA24-16 및 D452BT, 두 가지 상이한 효모 스트레인의 배양(cultivation)에 의한 에탄올 생산을 나타낸 것이다. 자일로스 분자는 오각형으로 나타내었고 셀로비오스 분자는 두 개의 육각형으로서 나타내었다 ; 및 (B) 자일로스-발효 스트레인 및 셀로비오스-발효 스트레인의 혼합된 배양물(Mixed cultures).

도 76 은 *N. 크라사*에서 354 자일란-유도 유전자(354 xylan-induced genes)의 목록을 나타낸 것이다.

도 77 은 다양한 *N. 크라사* 녹-아웃 스트레인(knock-out strains)에 대한 분비된 단백질 수준(secreted protein levels), 감소된 당(reducing sugar) 및 아조-자일라나아제 활성도(azo-xylanase activity)를 나타낸 것이다. 분비된 단백질 수준은 모든 스트레인에 대해서 상대적인 상수(relatively constant)이다.

도 78 은 (A) 야생형, Δ NCU05137 및 Δ NCU05137/ Δ NCU05137-GFP *네우로스포라* 스트레인에 대한 총 분비된 단백질 및 CMC-활성도, 및 (B) 세 가지 상이한 스트레인의 배양물(cultures)의 상청액(supernatant)에서의 총 단백질의 쿠마시 염색(Coomassie stain)을 나타낸 것이다.

도 79 는 분생자(conidia)에서 NCU05137-GFP 의 위치(localization)를 나타낸 것이다.

도 80 은 균사 팁(hypha tip)에서 NCU05137-GFP 의 위치를 나타낸 것이다.

발명을 실시하기 위한 구체적인 내용

발명의 상세한 설명

[0078]

- [0079] 본 발명은, 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11 및 α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포에 관한 것이되, 하기의 하나 또는 그 이상은 사실이다(true) : 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO: 1 을 포함하고, 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO: 2 를 포함하고, the loop connecting 트랜스멤브레인 α -헬릭스 2 와 트랜스멤브레인 α -헬릭스 3 를 연결하는 루프는 SEQ ID NO: 3 을 포함하고, 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO: 4 를 포함하고, 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO: 5 를 포함하고, 트랜스멤브레인 α -헬릭스 6 과 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO: 6 을 포함하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO: 7 를 포함하고, 및 트랜스멤브레인 α -헬릭스 10 및 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO: 8 을 포함하고, 및 상기 폴리펩티드는 세포 내로의 셀로텍스트린을 수송한다. 본원은, 세포 내로 셀로텍스트린의 수송을 증가시키는 방법, 셀로텍스트린을 포함하는 배지에 세포의 성장을 증가시키는 방법, 셀룰로스-유도 및 헤미셀룰로스-유도 당을 공동-발효시키는 방법 및 숙주 세포를 이용한 탄화수소 또는 탄화수소 유도체를 생성하는 방법을 추가적으로 기재하고 있다. 본원은, 세포 내로 오타당을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포, 세포 내로 오타당의 수송을 증가시키는 방법, 오타당 당을 포함하는 배지에서 세포의 성장을 증가시키는 방법, 및 세포 내로 오타당을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 탄화수소 또는 탄화수소 유도체를 생성하는 방법을 추가적으로 기재하고 있다.
- [0080] 본원에서 사용된 바와 같이, 셀로텍스트린은 다양한 길이(varying length)의 글루코스 중합체를 나타내고, 이로 제한되지 않는, 셀로비오스(2 글루코스 단량체), 셀로트리오스(cellotriose)(3 글루코스 단량체), 셀로테트라오스(cellotetraose)(4 글루코스 단량체), 셀로펜타오스(cellopentaose)(5 글루코스 단량체) 및 셀로헥사오스(cellohexaose)(6 글루코스 단량체)를 포함한다.
- [0081] 본원에서 사용된 바와 같이, 당은 단당류(monosaccharides)[예를 들어, 글루코스, 과당(fructose), 갈락토스, 자일로스, 아라비노스], 이당류(예를 들어, 셀로비오스, 수크로스, 락토오스, 말토오스), 및 올리고당(일반적으로 3 내지 10 성분(component)의 단당류)을 나타낸다.
- [0082] **본 발명의 폴리뉴클레오티드**
- [0083] 본원의 본 발명은 숙주 세포 및 이러한 숙주 세포를 이용한 방법에 관한 것이되, 상기 숙주 세포는 다양한 당을 수송할 수 있는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함한다.
- [0084] 본원에서 사용된 바와 같이, 용어 "폴리뉴클레오티드", "핵산 서열", "핵산의 서열" 및 이의 변형은 폴리데옥시리보뉴클레오티드(2-데옥시-D-리보오스를 포함), 폴리리보뉴클레오티드(D-리보오스를 포함), 퓨린 또는 피리미딘 염기의 N-글리코시드인 기타 다른 종류의 폴리뉴클레오티드, 및 DNA 및 RNA 에서 발견된 바와 같이, 상기 중합체가 염기 짝짓기(base pairing) 및 염기중첩(base stacking)을 가능하게 하는 입체배치(configuration)에서의 핵염기(nucleobases)를 포함한다면, 비-뉴클레오티드 백본(non-nucleotidic backbones)을 포함하는 그 밖의 중합체를 총칭한다. 따라서, 이러한 용어는 알려진 종류의 핵산 서열 변형물, 예를 들어 유사물(analog)을 갖는 하나 또는 그 이상의 자연적으로 발생하는 뉴클레오티드의 치환 ; 예를 들어, 전하되지 않은 결합(uncharged linkages)을 갖는 것들[예를 들어, 메틸 포스페이트(methyl phosphonates), 포스포트리에스테르(phosphotriesters), 포스포르아미데이트(phosphoramidates), 카르바메이트(carbamates) 등], 음전하된 결합(negatively charged linkages)을 갖는 것들[예를 들어, 포스포로티오에이트(phosphorothioates), 포스포로디티오에이트(phosphorodithioates) 등], 및 양전하된 결합(positively charged linkages)을 갖는 것들[예를 들어, 아미노알킬포스포르아미데이트(aminoalkylphosphoramidates), 아미노알킬포스포트리에스테르(aminoalkylphosphotriesters) 등]와 같은 인터-뉴클레오티드 변형물(inter-nucleotide modifications) ; 예를 들어 단백질[뉴클레아제(nucleases), 독소(toxins), 항체, 단일 펩티드, 폴리-L-리신 등을 포함]과 같은 펜던트의 모이어티(pendant moieties)를 포함하는 것들 ; 인터칼레이터(intercalators)를 갖는 것들[예를 들어 아크리딘(acridine), 소랄렌(psoralen) 등] 및 킬레이터를 포함하는 것들[예를 들어, 금속, 방사성 금속, 붕소, 산화

금속(oxidative metals) 등]을 포함한다. 본원에 사용된 바와 같이, 뉴클레오타이드 및 폴리뉴클레오타이드에 대한 부호(symbols)는 IUPAC-IUB Commission of Biochemical Nomenclature 에 의해 권장된 것이다(Biochem. 9:4022, 1970).

[0085] 본원에서 사용된 바와 같이, "폴리펩티드"는 다수의 연이은 중합 아미노산 잔기(a plurality of consecutive polymerized amino acid residues)(예를 들어, 적어도 약 15 가지의 연이은 중합 아미노산 잔기, 선택적으로 적어도 약 30 가지의 연이은 중합 아미노산 잔기, 적어도 약 50 가지의 연이은 중합 아미노산 잔기)를 포함하는 아미노산 서열이다. 많은 예에서, 폴리펩티드는, 운반체, 전사인자(transcription factor), 알려지지 않은 기능(unknown function)의 예측 단백질(predicted protein), 또는 이의 도메인 또는 단백질 또는 단편인 중합 아미노산 잔기 서열(polymerized amino acid residue sequence)을 포함한다. 운반체는 이온, 작은 분자, 또는 생체막(biological membrane)을 가로지르는 탄수화물과 같은 고분자(macromolecules)를 포함하는 것이다. 전사인자는 유전자 발현을 조절할 수 있고, 숙주 세포에서 유전자 발현을 증가시키거나 감소시킬 수 있다. 폴리펩티드는, 선택적으로 변형된 아미노산 잔기(modified amino acid residues), 코돈(codon)에 의해 코드되지 않는 자연적으로 발생한 아미노산 잔기(naturally occurring amino acid residues) 및 비-자연적으로 발생한 아미노산 잔기를 포함한다.

[0086] 본원에서 사용된 바와 같이, "단백질"은 아미노산 서열, 올리고펩티드, 펩티드, 폴리펩티드 또는 자연적으로 발생하거나 합성된(naturally occurring or synthetic) 이의 부분(portions)에 관한 것이다.

[0087] 본 발명의 재조합 폴리뉴클레오타이드는 표 10, Supplemental Data, Dataset SI, page 3 in Tian et al., 2009 ; 표 14, 15, 16, 29 ; 또는 도 76 에 나타난 유전자 중 어떠한 것에 의해 코드되는 폴리펩티드를 코드하는 어떠한 폴리뉴클레오타이드를 포함한다. 바람직한 실시형태에서, 본 발명의 폴리뉴클레오타이드는, 서열 NCU00801, NCU00809, NCU08114, NCU00130, NCU00821, NCU04963, NCU06138, STL12/XUT6, SUT2, SUT3, XUT1, XUT3, NCU07705, NCU05137, NCU01517, NCU09133, or NCU10040 중 어떤 것에 의해 코드되는 폴리펩티드를 코드하는 어떠한 폴리뉴클레오타이드를 포함한다.

[0088] [표 1]

하기에 기재된 서열 또는 하기에 기재된 폴리펩티드를 코딩하는 서열을 포함하는 본 발명의 폴리뉴클레오티드

유전자 이름/ 로커스(Locus)	대체명	NCBI 참조서열/ 젠뱅크 수납번호 (GenBank Accession Number)	유기체 (Organism)
NCU00801	<i>cbl</i>	XP_963801.1/EAA34565	<i>N. 크라사</i>
NCU00809		XP_964302.1/EAA35116.1	<i>N. 크라사</i>
NCU00821	AN25	XP_964364.2/EAA35128.2	<i>N. 크라사</i>
NCU00988	Xy33	XP_963898.1/EAA34662.1	<i>N. 크라사</i>
NCU01231		XP_961597.2/EAA32361.2	<i>N. 크라사</i>
NCU01494	AN49	XP_955927.2/EAA26691.2	<i>N. 크라사</i>
NCU02188	AN28-3	XP_959582.2/EAA30346.2	<i>N. 크라사</i>
NCU04537	Xy50	XP_955977.1/EAA26741.1	<i>N. 크라사</i>
NCU04963	AN29-2	XP_959411.2/EAA30175.2	<i>N. 크라사</i>
NCU05519		XP_960481.1/EAA31245.1	<i>N. 크라사</i>
NCU05853		XP_959844.1/EAA30608.1	<i>N. 크라사</i>
NCU05897		XP_959888.1/EAA30652.1	<i>N. 크라사</i>
NCU06138	Xy31	XP_960000.1/EAA30764.1	<i>N. 크라사</i>
NCU08114	<i>cbl</i>	XP_963873.1/EAA34637.1	<i>N. 크라사</i>
NCU09287	AN41	XP_958139.1/EAA28903.1	<i>N. 크라사</i>
NCU10021		XP_958069.2/EAA28833.2	<i>N. 크라사</i>
XP_001387242	Ap26	XP_001387242	<i>P. 스티피티스</i>
HGT3	Xyp30-1	XP_001386715.1/ABN68686.1	<i>P. 스티피티스</i>
STL1	Xyp30	XP_001383774.1/ABN65745.1	<i>P. 스티피티스</i>
STL12/XUT6	Xyp29	XP_001386589.1/ABN68560.1	<i>P. 스티피티스</i>
SUT2	Ap31	XP_001384295.2/ABN66266.2	<i>P. 스티피티스</i>
SUT3	Xyp37	XP_001386019.2/ABN67990.2	<i>P. 스티피티스</i>
XUT1	Xyp32	XP_001385583.1/ABN67554.1	<i>P. 스티피티스</i>
XUT2	Xyp31	XP_001387242.1/EAZ63219.2	<i>P. 스티피티스</i>
XUT3	Xyp33	XP_001387138.1/EAZ63115.1	<i>P. 스티피티스</i>
XUT7	Xyp28	XP_001387067.1/EAZ63044.1	<i>P. 스티피티스</i>
NCU07705	<i>cdr-1</i>	XP_962291.1/EAA33055	<i>N. 크라사</i>
NCU05137		XP_956635.1/EAA27399	<i>N. 크라사</i>
NCU01517		XP_956966.1/EAA27730	<i>N. 크라사</i>
NCU09133		XP_958905.1/EAA29669	<i>N. 크라사</i>
NCU10040			<i>N. 크라사</i>

[0089]

[0090] 특정 실시형태에서, 본 발명의 재조합 폴리뉴클레오티드는, 표 10, Supplemental Data, Dataset SI, page 3 in Tian et al., 2009 ; 표 14, 15, 16, 29 ; 또는 도 76 에 나타난 유전자 중 어떠한 것에 의해 코딩되는 폴리펩티드와 적어도 약 20 %, 적어도 약 29 %, 적어도 약 30 %, 적어도 약 40 %, 적어도 약 50 %, 적어도 약 55 %, 적어도 약 60 %, 적어도 약 65 %, 적어도 약 70 %, 적어도 약 75 %, 적어도 약 80 %, 적어도 약 85 %, 적어도 약 90 %, 적어도 약 92 %, 적어도 약 94 %, 적어도 약 96 %, 적어도 약 98 %, 적어도 약 99 % 또는 적어도 약 100 % 의 아미노산 동일성 잔기 서열(amino acid residue sequence)을 갖는 폴리펩티드를 코딩한다. 바람직한 실시형태에서, 본 발명의 폴리뉴클레오티드는, 서열 NCU00801, NCU00809, NCU08114, NCU00130, NCU00821, NCU04963, NCU06138, STL12/XUT6, SUT2, SUT3, XUT1, XUT3, NCU07705, NCU05137, NCU01517, NCU09133 또는 NCU10040 중 어떠한 것에 의해 코딩되는 폴리펩티드와 적어도 약 20 %, 적어도 약 29 %, 적어도 약 30 %, 적어도 약 40 %, 적어도 약 50 %, 적어도 약 55 %, 적어도 약 60 %, 적어도 약 65 %, 적어도 약 70 %, 적어도 약 75 %, 적어도 약 80 %, 적어도 약 85 %, 적어도 약 90 %, 적어도 약 92 %, 적어도 약 94 %, 적어도 약 96 %, 적어도 약 98 %, 적어도 약 99 % 또는 적어도 약 100 % 의 아미노산 동일성 잔기 서열을 갖는 폴리펩티드를 코딩한다.

[0091] 본 발명의 폴리뉴클레오티드는 상기 나타난 유전자에 의해 코딩되는 폴리펩티드의 보존적으로 변형된 변이체(conservatively modified variants)를 코딩하는 폴리뉴클레오티드를 추가로 포함한다. 본원에 사용된 바와 같은 "보존적으로 변형된 변이체(Conservatively modified variants)" 는 화학적으로 유사한 아미노산을 갖는 아미노산의 치환을 야기하는, 폴리펩티드 서열에 개별적인 치환(substitutions), 결실(deletions) 또는 첨가를 포

함한다. 기능적으로 유사한 아미노산을 제공하는 보존적 치환 표(Conservative substitution tables)는 본 분야에서 널리 공지되어 있다. 이러한 보존적으로 변형된 변이는 다형성 변이체(polymorphic variants), 이종간의 상동체(interspecies homologs) 및 디스클로저(disclosure)의 대립유전자(alleles)를 제외하지 않고 이를 포함한다. 하기의 8 가지 그룹은 서로에 대한 보존적 치환인 아미노산을 포함한다 : 1) 알라닌(A), 글리신(G); 2) 아스파르트산(D), 글루탐산(E); 3) 아스파라진(N), 글루타민(Q); 4) 아르기닌(R), 리신(K); 5) 이소류신(I), 류신(L), 메티오닌(M), 발린(V); 6) 페닐알라닌(F), 티로신(Y), 트립토판(W); 7) 세린(S), 트레오닌(T); 및 8) 시스테인(C), 메티오닌(M)[예를 들어 Creighton, Proteins(1984) 참조].

[0092] 본 발명의 폴리뉴클레오티드는, 표 10, Supplemental Data, Dataset SI, page 3 in Tian et al., 2009 ; 표 14, 15, 16, 29 ; 또는 도 76 에 나타난 유전자중 어떠한 것에 의해 코드되는 폴리펩티드의 오솔로그(orthologs) 또는 상동체(homologs)를 코드하는 폴리뉴클레오티드를 추가로 포함한다. 본원에 사용된 바와 같은 "상동체(Homology)" 는 참조서열(reference sequence) 및 두 번째 서열의 적어도 하나의 단편 사이의 서열 유사성(sequence similarity)에 관한 것이다. 상동체는 본 분야에서 공지된 어떠한 방법, 바람직하게 참조 서열과 단일 두 번째 서열(single second sequence) 또는 서열의 단편, 또는 서열의 데이터베이스와 비교하는 BLAST 도구(BLAST tool)를 사용하여 확인할 수 있다. 하기에 기재된 바와 같이, BLAST 는 퍼센트 동일성(identity) 및 유사성(similarity)을 기반으로 한 서열과 비교할 수 있다. 본원에서 사용된 바와 같은, "오솔로그(Orthology)"는 공통 조상 유전자(common ancestor gene)로부터 유래된 상이한 종(different species)에서의 유전자에 관한 것이다.

[0093] 두 가지 또는 그 이상의 핵산 또는 폴리펩티드 서열의 맥락에서 용어 "동일한(identical)" 또는 퍼센트 "동일성(identity)"은 동일한 두 가지 또는 그 이상의 서열 또는 서브서열(subsequences)에 관한 것이다. 비교 윈도우(comparison window), 또는 다음의 서열 비교 알고리즘 또는 매뉴얼 정렬(manual alignment)을 사용하거나 또는 육안검사(visual inspection)에 의해 측정된 바와 같은 지정된 영역(designated region) 상에서 최대한의 관련성(maximum correspondence)에 대해서 비교하고 정렬하였을 때, 만약 두 가지 서열이 의 아미노산 동일성 잔기 또는 뉴클레오티드의 명시된 퍼센트(예를 들어, 명시된 영역, 또는 명시되지 않은 전체 서열 상에서 29 % 동일성, 선택적으로 30 %, 40 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 99 % 또는 100 % 동일성)를 갖는다면, 두 서열은 "상당히 동일하다(substantially identical)". 선택적으로, 동일성은, 길이(length)에서 적어도 약 50 개의 뉴클레오티드(또는 10 개의 아미노산)인 영역, 또는 보다 바람직하게 길이에서 100 내지 500 개 또는 1000 개 또는 그 이상의 뉴클레오티드(또는 20, 50, 200 개 또는 그 이상의 아미노산)인 영역 상에서 존재한다.

[0094] 서열 비교(sequence comparison)에서, 일반적으로 하나의 서열은, 실험 서열과 비교하기 위해, 참조 서열(reference sequence)로서 작용한다. 서열 비교 알고리즘(sequence comparison algorithm)을 사용할 때, 실험 및 참조 서열을 컴퓨터에 입력하고, 만약 필요하다면 서열 코디네이트(subsequence coordinates)를 지정하고, 서열 알고리즘 프로그램 파라미터(sequence algorithm program parameters)를 지정한다. 디폴트 프로그램 파라미터(Default program parameters)를 사용할 수 있고, 또는 선택적인 파라미터를 지정할 수 있다. 그리고 난 다음에 서열 비교 알고리즘은, 프로그램 파라미터를 기반으로 한 참조 서열에 대한 상대적인 테스트 서열에 대한 퍼센트 서열 동일성(percent sequence identities)을 계산한다. 동일성에 대해서 두 가지 서열을 비교하였을 때, 서열이 근접할 필요는 없지만, 갭(gap)은 전체적인 퍼센트 동일성을 감소시킬 수 있는 페널티(penalty)을 줄 수 있다. blastn 에 대해서, 디폴트 파라미터(default parameters)는 갭 오픈링 페널티(Gap opening penalty) = 5 및 갭 오픈링 페널티 = 2 이다. blastp 에 대해서, 디폴트 파라미터는 갭 오픈링 페널티 = 11 및 갭 익스텐션 페널티(Gap extension penalty) = 1 이다.

[0095] 본원에 사용된 바와 같은 "비교 윈도우(comparison window)"는 인접한 위치의 수 중 어떠한 것의 부분(segment)과 관련된 것을 포함하지만, 서열은 두 가지의 서열이 최적으로 정렬된 후에 인접한 위치의 동일한 번호의 참조서열과 비교하여, 20 내지 60, 일반적으로 약 50 내지 약 200, 보다 일반적으로 약 100 내지 약 150 로 제한되지 않는다. 비교를 위한 서열 정렬의 방법은 본 분야에 공지되어 있다. 비교를 위한 서열의 최적의 정렬은,

Smith and Waterman (1981)의 일부의 상동성 알고리즘(local homology algorithm), Needleman and Wunsch (1970) J Mol Biol 48(3):443-453 의 상동성 정렬 알고리즘(homology alignment algorithm), Pearson and Lipman (1988) Proc Natl Acad Sci USA 85(8):2444-2448 의 유사성 방법에 대한 검색, 이러한 알고리즘(GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI)의 컴퓨터화된 이행(computerized implementations) 또는 메뉴얼 정렬 및 육안 검사 [예를 들어, Brent *et al.*, (2003) Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (Ringbou Ed) 참조]에 의해 할 수 있다.

[0096] 퍼센트 서열 동일성 및 서열 유사성에 대한 적절한 알고리즘의 두 가지 예는, 각각 Altschul 등의. (1997) Nucleic Acids Res 25(17):3389-3402 및 Altschul 등의. (1990) J. Mol Biol 215(3)-403-410 에 기재된 BLAST 및 BLAST 2.0 알고리즘이다. BLAST 분석을 수행하기 위한 소프트웨어는 National Center for Biotechnology Information 을 통해 공개적으로 입수가 가능하다. 이러한 알고리즘은 데이터베이스 서열에서 동일한 길이의 마디(word)와 함께 정렬하였을 때, 몇몇 포지티브-값의 한계 점수 T (positive-valued threshold score T)를 맞추거나 또는 충족시키는, 질의 서열(query sequence)에서 길이 W 의 짧은 단어(short words)를 확인하여 첫 번째로 확인된 높은 점수의 서열 짝(high scoring sequence pairs, HSPs)을 포함한다. T 는 인접한 단어 점수 한계(neighborhood word score threshold)로서 언급된 것이다(Altschul et al., supra). 이러한 초기의 인접한 단어 히트(initial neighborhood word hits)는 이들을 포함하는 보다 긴 HSPs 를 발견하기 위해 초기 검색에 대한 시드(seeds)로서의 역할을 한다. 단어 히트(word hits)는 누적된 정렬 점수(cumulative alignment score)가 증가될 수 있을 때까지 각각의 서열을 따라 양 방향(both directions)으로 연장되었다. 누적된 점수는 뉴클레오티드 서열, 파라미터 M[어울리는 잔기(matching residues)의 짝(pair)에 대한 보상 점수(reward score) ; 항상 > 0] 및 N[어울리지 않는 잔기(mismatching residues)에 대한 페널티 점수(penalty score) ; 항상 < 0]에 대해 사용되어 계산되었다. 아미노산 서열에서, 스코어링 매트릭스(scoring matrix)는 누적된 점수를 계산하기 위해 사용되었다. 각각의 방향에서 단어 히트의 연장(Extension)은, 누적된 정렬 점수(cumulative alignment score)가 이의 최대한의 성취된 값(achieved value)으로부터의 다수의 X(quantity X)에 의해 떨어졌을 때 ; 누적된 점수가, 하나 또는 그 이상의 네가티브-스코어링 잔기 정렬(negative-scoring residue alignments)의 축적으로 인하여 0 이하로 갔을 때 ; 또는 어느 하나의 서열이 말단으로 도달하였을 때에 정지한다. BLAST 알고리즘 파라미터 W, T 및 X 은 민감성(sensitivity) 및 정렬의 속도를 결정한다. BLASTN 프로그램(뉴클레오티드 서열에 대한)은 디폴트(defaults) 11 의 워드랜스(wordlength)(W), 기대값(expectation)(E) 또는 10, M=5, N=-4, 및 둘 다의 가닥의 비교로서 사용되었다. 아미노산 서열에 대해서, BLASTP 프로그램은 디폴트 3 의 워드랜스(W), 10 의 기대값(E) 및 BLOSUM62 스코어링 매트릭스[Henikoff and Henikoff, (1992) Proc Natl Acad Sci USA 89(22):10915-10919 참조], 50 의 정렬(alignments)(B), 10 의 기대값, M=5, N=-4 및 둘 다의 가닥의 비교로서 사용되었다.

[0097] BLAST 알고리즘은 또한 두 가지 서열 사이의 유사성의 통계적인 분석(statistical analysis)을 위해 또한 실시하였다(예를 들어, Karlin and Altschul, (1993) Proc Natl Acad Sci USA 90(12):5873-5877 참조). BLAST 알고리즘에 의해 제공된 유사성의 하나의 척도(measure)는, 두 가지의 뉴클레오티드 또는 아미노산 서열 사이의 매치(match)가 우연히 발생할 수 있는 가능성의 표시(indication)를 제공하는, 가장 작은 합계 확률(sum probability)(P(N))이다. 예를 들어, 만약 테스트 핵산과 참조 핵산의 비교에서 가장 작은 합계 확률이 약 0.2 미만, 보다 바람직하게 약 0.01 미만 및 가장 바람직하게 약 0.001 미만이라면, 핵산은 참조서열에 유사한 것으로 간주된다.

[0098] 상기에 나타난 서열 동일성 퍼센트 이외의, 두 가지 핵산 서열 또는 폴리펩티드가 상당히 동일한 또 다른 표시(indication)는, 하기에 나타난 바와 같이, 첫 번째 핵산에 의해 코드된 폴리펩티드가 두 번째 핵산에 의해 코드된 폴리펩티드에 대하여 증가된 항체와 면역학적으로 교차-반응성(immunologically cross-reactive)을 갖는 경우이다. 따라서, 예를 들어 두 가지 펩티드가 보존적 치환에 의해서만 다른 경우에, 폴리펩티드는 두 번째 폴리펩티드와 일반적으로 상당히 동일하다. 두 가지 핵산 서열이 상당히 동일한 또 다른 표시는, 하기에 나타난 바와 같이, 두 가지 분자 또는 이들의 보체(complements)가 엄격한 조건(stringent conditions) 하에서 서로가 혼성화(hybridize)되는 경우이다. 두 가지 핵산 서열이 상당히 동일한 또 다른 표시는, 동일한 프라이머를

서열을 증폭시키기 위해 사용할 수 있는 경우이다.

- [0099] 본원에 나타난 바와 같이, 본 발명의 폴리뉴클레오티드는, NCU00988, NCU10021, NCU04963, NCU06138, NCU00801, NCU08114 및 NCU05853 를 포함하는 주요한 촉진 상과 당 운반체 패밀리(Major Facilitator Superfamily sugar transporter family)의 멤버(members)를 포함한다. 운반체의 주요한 촉진 상과(Major Facilitator Superfamily, MFS)(운반체 유형 # 2.A.1)의 멤버(Members)는, 세포내 N- 및 C-말단을 갖는 12 가지의 트랜스멤브레인 α -헬릭스로 항상 구성되어있고[S. S. Pao, I. T. Paulsen, M. H. Saier, Jr., Microbiol Mol Biol Rev 62, 1 (Mar, 1998)], 반면에 MFS 운반체의 일차 서열(primary sequence)은 폭넓게 다양하고, 모두는 *E. coli* 락토오스 투과효소(lactose permease)(LacY)[J. Abramson *et al.*, *Science* 301, 610 (Aug 1, 2003)] 및 *E. coli* Pi/글리세롤-3-포스페이트(GlpT)[Y. Huang, M. J. Lemieux, J. Song, M. Auer, D. N. Wang, *Science* 301, 616 (Aug 1, 2003)]의 삼차구조를 공유한다. 이러한 예에서, 6 가지의 N- 및 C- 말단 헬릭스는 헬릭스 6 및 7 사이의 긴 세포질 루프(long cytoplasmic loop)에 의해 연결된 두 가지의 별개의 도메인을 형성한다. 이러한 대칭(symmetry)은 MFS 에 대해 주어진 증가(given rise)를 갖는 것을 생각하는 복제 사건(duplication event)에 해당한다. 친수성 캐비티(hydrophilic cavity) 내의 기질 결합은 N-말단 도메인의 헬릭스 1, 2, 4 및 5, 및 C-말단 도메인의 헬릭스 7, 8, 10 및 11 에 의해 형성된다. 이러한 캐비티는 헬릭스 3, 6, 9 및 12 에 의해 안정된다.
- [0100] MFS 의 당 운반체 패밀리(운반체 분류 # 2.A.1.1)는 트랜스멤브레인 헬릭스 6 및 12[PESPR (SEQ ID NO: 9)/PETK (SEQ ID NO: 10)], 및 루프 2 및 8(GRR/GRK)[M. C. Maiden, E. O. Davis, S. A. Baldwin, D. C. Moore, P. J. Henderson, *Nature* 325, 641 (Feb 12-18, 1987)]에서 발견된 모티프(motifs)로 규정된다. 이러한 패밀리에 대해서 전체적인 Hidden Markov Model(HMM)는 pfam.janelia.org/family/PF00083#tabview=tab3 에서 볼 수 있다. PROSITE[N. Hulo *et al.*, *Nucleic Acids Res* 34, D227 (Jan 1, 2006)]은 이러한 패밀리의 멤버를 확인하기 위해 두 가지의 모티프를 사용하였다. 첫 번째는 [LIVMSTAG] - [LIVMFSAG] - {SH} - {RDE} - [LIVMSA] - [DE] - {TD} - [LIVMFYWA] - G - R - [RK] - x(4,6) - [GSTA](SEQ ID NO: 11)이다. 두 번째는 [LIVMF] - x - G - [LIVMFA] - {V} - x - G - {KP} - x(7) - [LIFY] - x(2) - [EQ] - x(6) - [RK] (SEQ ID NO: 12)이다. PROSITE 모티프를 해독하는 방법을 예로서 하기의 모티프 [AC]-x-V-x(4)-{ED}는 [Ala 또는 Cys]-어떠한 것(any)-Val-어떠한 것-어떠한 것-어떠한 것-어떠한 것-{Glu 또는 Asp 외의 어떠한 것}({any but Glu or Asp}) (SEQ ID NO: 13)로서 번역된다.
- [0101] 본원에 기재된 바와 같이, NCU00801, NCU00809, NCU08114, XP_001268541.1 및 LAC2 은 셀로텍스트린을 수송하는 폴리펩티드를 코드하는 것으로 발견되었다. 게다가, 알라닌 스캐닝 실험 및 서열 분석은 12 가지의 트랜스멤브레인 α -헬릭스를 포함하는 제조합 폴리펩티드를 결정하기 위해 사용되었고, SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, and SEQ ID NO: 8 으로 이루어진 하나 또는 그 이상의 서열은 셀로텍스트린을 수송하는 폴리펩티드를 코드한다.
- [0102] 따라서, 하나의 양상에서, 본 발명은 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 제공하고, 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO: 1 을 포함한다. 또 다른 양상에서, 본 발명은 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 제공하고, 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO: 2 를 포함한다. 또 다른 양상에서, 본 발명은 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 제공하고, 트랜스멤브레인 α -헬릭스 2 및 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프는 SEQ ID NO: 3 을 포함한다. 또 다른 양상에서, 본 발명은 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를

코딩하는 폴리뉴클레오티드를 제공하고, 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO: 4 를 포함한다. 또 다른 양상에서, 본 발명은 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 제공하고, 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO: 5 를 포함한다. 또 다른 양상에서, 본 발명은 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 제공하고, 트랜스멤브레인 α -헬릭스 6 및 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO: 6 을 포함한다. 또 다른 양상에서, 본 발명은 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 제공하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO: 7 을 포함한다. 또 다른 양상에서, 본 발명은 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 제공하고, 트랜스멤브레인 α -헬릭스 10 및 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO: 8 을 포함한다.

[0103] 상기한 양상의 각각은 어떠한 수의 조합(combinations)으로 결합될 수 있다. 이러한 양상에 따른 폴리뉴클레오티드는 SEQ ID NOs: 1-8 중 어떠한 것의 1, 2, 3, 4, 5, 6 또는 7 을 포함하는 폴리펩티드를 코딩할 수 있거나, 폴리펩티드는 SEQ ID NOs: 1-8 의 모두를 포함할 수 있다. 예를 들어, 폴리뉴클레오티드는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩할 수 있고, 상기 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO: 1 을 포함하고, 트랜스멤브레인 α -헬릭스 2 와 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프는 SEQ ID NO: 3 을 포함하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO: 7 을 포함한다. 또는 다른 예에서, 폴리뉴클레오티드는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩할 수 있고, 상기 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO: 2 를 포함하고, 상기 트랜스멤브레인 α -헬릭스 3 은 SEQ ID NO: 3 을 포함하고, 상기 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO: 5 를 포함하고, 트랜스멤브레인 α -헬릭스 10 과 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO: 8 을 포함한다.

[0104] 상기한 양상의 특정 실시형태에서, 상기 폴리펩티드는 NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 %, 적어도 100 % 의 의 아미노산 동일성을 갖는다.

[0105] 본원에 추가로 기재된 바와 같이, NCU08221 및 STL12/XUT6 은 자일로스를 수송하는 폴리펩티드를 코딩하는 것으로 발견되었다. XUT1 은 아라비노스를 수송하는 폴리펩티드를 코딩하는 것으로 발견되었다. NCU06138 은 아라비노스 또는 글루코스를 수송하는 폴리펩티드를 코딩하는 것으로 발견되었다. SUT2, SUT3, 및 XUT3 은 자일로스 또는 글루코스를 수송하는 폴리펩티드를 코딩하는 것으로 발견되었다. NCU04963 은 자일로스, 아라비노스 또는 글루코스를 수송하는 폴리펩티드를 코딩하는 것으로 발견되었다. 바람직한 실시형태에서, 본 발명의 폴리뉴클레오티드는 NCU08221 또는 STL12/XUT6 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 자일로스를 수송한다. 다른 바람직한 실시형태에서, 본 발명의 폴리뉴클레오티드는 XUT1 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 아라비노스를 수송한다. 다른 바람직한 실시형태에서, 본 발명의 폴리뉴클레오티드는 NCU06138 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 아라비노스 또는 글루코스를 수송한다. 다른 바람직한 실시형태에서, 본 발명의 폴리뉴클레오티드는 SUT2, SUT3 또는 XUT3 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 자일로스 또는 글루코스를 수송한다. 다른 바람직한 실시형태에서, 본 발명의 폴리뉴클레오티드는 NCU04963 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리펩티드는 자일로스, 아라비

노스 또는 글루코스를 수송한다.

- [0106] NCU07705 에 의해 코드되는 폴리펩티드를 코드하는 본 발명의 폴리뉴클레오티드는, 분류되지 않는 단백질(unclassified protein)을 코드하기 위해 FunCat (Ruepp, 2004; webpage broad.mit.edu/annotation/genome/neurospora/Home.html)에 의해 예측되었다. 그러나, NCU07705 에 의해 코드되는 폴리펩티드의 BLAST 분석은, 전사인자(transcription factors)를 포함하는 많은 C6 징크 핑거 도메인(zinc finger domain)과 높은 유사성을 갖는 것으로 나타났다[도 1 참조 ; 모범적인 상동성(exemplary homologs)의 목록은 관련된 U.S. Appl. No. 61/271,833 의 도 23 에서 발견될 수 있다]. 본 발명의 폴리뉴클레오티드는, NCU07705 에 의해 코드되는 폴리펩티드의 이러한 상동체(homologs) 또는 본 분야에서 공지된 어떠한 방법으로 확인된 그 밖의 상동체를 코드하는 폴리뉴클레오티드를 포함한다.
- [0107] 본 발명의 또 다른 양상에서, 본 발명의 폴리뉴클레오티드는 NCU05137 에 의해 코드되는 폴리펩티드를 코드하는 이러한 폴리뉴클레오티드를 포함한다. FunCat 은 분류되지 않는 단백질(unclassified protein)인 NCU05137 에 의해 코드되는 폴리펩티드를 분류한다. 그러나, NCU05137 은 많은 실모양의 자낭기 균류(filamentous ascomycete fungi)의 게놈에 매우 잘 보존되었다(도 2 참조). 본 발명의 폴리뉴클레오티드는, NCU05137 에 의해 코드되는 폴리펩티드의 이러한 상동체 또는 본 분야에서 공지된 모든 방법으로 확인된 그 밖의 상동체를 코드하는 폴리뉴클레오티드를 포함한다.
- [0108] 본 발명의 또 다른 양상에서, 본 발명의 폴리뉴클레오티드는, NCU01517, NCU09133 또는 NCU10040 에 의해 코드되는 폴리펩티드를 코드하는 이러한 폴리뉴클레오티드를 포함한다. FunCat 은 글루코아밀라아제 전구체(glucoamylase precursor)가 될 예정인 NCU01517 에 의해 코드되는 폴리펩티드를 분류한다. FunCat 은 분류되지 않는 단백질이 될 예정인 NCU09133 및 NCU10040 에 의해 코드되는 폴리펩티드를 분류한다. 본 발명의 폴리뉴클레오티드는, NCU01517, NCU09133 또는 NCU10040 에 의해 코드되는 폴리펩티드의 이러한 상동체 또는 본 분야에서 공지된 어떠한 방법으로 확인된 그 밖의 상동체를 코드하는 폴리뉴클레오티드를 포함한다.
- [0109] 이러한 폴리펩티드의 예측된 기능은 폴리뉴클레오티드 및 이의 코드된 단백질의 기능적인 분석(functional analyses)을 실시하여 확인할 수 있다. 이러한 분석은, 예를 들어 폴리뉴클레오티드의 결실(deletions)을 포함하는 스트레인의 표현형 분석(phenotypic analysis), 유전적인 상보성 실험(genetic complementation experiments), 폴리뉴클레오티드의 야생형 복제(wild-type copy)가 발현된 스트레인의 표현형 분석, 폴리펩티드의 재조합 형태의 발현 및 정제, 및 재조합 폴리펩티드의 생화학적인 특성 및 활성도의 차후의 특성을 포함할 수 있다.
- [0110] 본 발명의 폴리뉴클레오티드의 서열은, 예를 들어 직접적인 화학적 합성(direct chemical synthesis) 또는 클로닝(cloning)을 포함하는, 본 분야의 숙련자에게 공지된 적절한 방법으로 제조할 수 있다. 직접적인 화학적인 합성에 대해서, 핵산의 중합체의 형성은, 성장하는 뉴클레오티드 사슬(growing nucleotide chain)의 말단 5'-히드록실기에 3'-블록된(blocked) 및 5'-블록된 뉴클레오티드 단량체(nucleotide monomers)의 순차적인 첨가(sequential addition)를 일반적으로 포함하고, 포스포트리에스테르(phosphotriester), 포스포르아미디트(phosphoramidite)와 같은 일반적인 인 유도체(phosphorus derivative)인, 상기 각각의 첨가는, 첨가된 단량체의 3'-위치에서 성장하는 사슬의 말단 5'-히드록실기의 핵친화성 공격(nucleophilic attack)에 의해 영향받는다. 이러한 방법론(methodology)은 본 분야에서 일상적인 기술로 알려져 있고, 적절한 문맥 및 문헌에 기재되어 있다[예를 들어, in Matteucci *et al.*, (1980) Tetrahedron Lett 21:719-722; U.S. Pat. Nos. 4,500,707 ; 5,436,327 ; 및 5,700,637]. 게다가, 목적하는 서열은, 적절한 제한 효소를 사용한 DNA 분리(splitting), 겔 전기영동을 사용한 단편의 분리, 및 그 후에 중합효소 연쇄 반응(polymerase chain reactions, PCR; 예를 들어, U.S. Pat. No. 4,683,195)의 이용과 같은 본 분야의 공지된 기술을 통한 겔(gel)로부터의 목적하는 핵산 서열의 회복(recovering)에 의한 천연 자원(natural sources)으로부터 분리될 수 있다.

- [0111] 본 발명의 각각의 뉴클레오티드는 발현 벡터(expression vector) 내로 포함될 수 있다. "발현 벡터" 또는 "벡터"는 숙주 세포에 형질도입되거나(transduces), 형질전환되거나(transforms) 또는 침투한(infect) 화합물(compound) 및/또는 조성물(composition)을 나타낸 것이고, 이에 의해서 세포에 천연 이외의 또는 세포에 천연이 아닌 방식으로 핵산 및/또는 단백질이 발현되는 세포가 야기된다. "발현 벡터"는 숙주 세포에 의해 발현될 핵산의 서열(보통 RNA 또는 DNA)을 포함한다. 선택적으로, 벡터의 발현은, 바이러스, 리포솜(liposome), 단백질 코팅(protein coating) 또는 그 밖에 유사한 것과 같은 숙주 세포 내로의 핵산의 유입을 달성하는데 이를 돕는 물질을 또한 포함한다. 본 발명에 사용에 예상된 발현 벡터는, 바람직하거나 또는 요구되는 요소와 함께 핵산 서열을 삽입할 수 있는 이러한 것들을 포함한다. 추가로, 발현 벡터는, 숙주 세포 내로 이동될 수 있고 그 안에 복제될 수 있는 것이어야 한다. 바람직한 발현 벡터는 플라스미드, 특히 핵산 서열의 전사에 바람직하거나 요구되는 사용할 준비가 갖추어진 요소(operational elements)를 포함하고, 잘 기록된 제한효소 절단위치(restriction sites)를 갖는 이러한 것들이다. 그 밖의 발현 벡터 뿐만 아니라 이러한 플라스미드는 본 분야에 널리 공지되어 있다.
- [0112] 개별적인 폴리뉴클레오티드의 혼합은, 발현벡터, 예를 들어 플라스미드에서의 특정한 부위(specific sites)를 분해하기 위해서, 예를 들어 제한효소(BamR1, EcoR1, Hha1, Xho1, Xma1, 및 등과 같은)의 사용을 포함하는 공지된 방법을 통해서 실시할 수 있다. 제한효소(restriction enzyme)는 분리된 발현벡터의 말단에 대한 서열 상보성을 갖는 말단을 갖기 위해 합성하거나 또는 이를 갖는 폴리뉴클레오티드에 어닐링(annealed)될 수 있는 단일 가닥 말단(single stranded ends)을 생산한다. 어닐링(Annealing)은 적절한 효소, 예를 들어 DNA 리가아제(ligase)를 사용하여 실시하였다. 본 분야에서의 보통의 기술의 이러한 것들에 의해 예측될 수 있는 바와 같이, 발현 벡터 및 목적하는 폴리뉴클레오티드 둘 다는 동일한 제한 효소로 보통 분리되었고, 이에 의해서 발현 벡터의 말단 및 폴리뉴클레오티드의 말단은 서로 상보적임을 확인할 수 있다. 게다가, DNA 링커는 발현 벡터 내로 핵산 서열의 연결(linking)을 가능하게 하는데 사용될 수 있다.
- [0113] 개별적인 폴리뉴클레오티드는 본 분야에서의 기술 분야에서 공지된 방법을 사용하여 또한 결합할 수 있다(예를 들어, U.S. Pat. No. 4,683,195).
- [0114] 예를 들어, 각각의 목적하는 폴리뉴클레오티드는 분리된 PCR(separate PCR)에서 초기에 생성될 수 있다. 그 후에, 특정한 프라이머는, PCR 산물의 말단은 상보적인 서열(complementary sequences)을 포함하는 것과 같이 설계되었다(design). PCR 산물이 혼합되고, 변성되었고(denatured), 다시 어닐되었을 때(annealed), 이들의 3' 말단에 매칭 서열(matching sequences)을 갖는 가닥이 겹쳐지고(overlap) 서로에 대한 프라이머로서 작용할 수 있다. DNA 중합효소에 의한 이러한 오버랩(overlap)의 확대는, 원본 서열이 서로 "이어지는(spliced)" 분자를 생성한다. 이러한 방식에서, 개별적인 폴리뉴클레오티드의 시리즈(series)는 서로 "이어질 수 있고", 동시에 숙주 세포 내로 형질도입되었다(transduced). 따라서, 각각의 많은 수의 폴리뉴클레오티드의 발현에 영향을 미친다.
- [0115] 각각의 폴리뉴클레오티드 또는 "이어진(spliced)" 폴리뉴클레오티드는 발현벡터 내로 포함된다. 본 발명은 폴리뉴클레오티드가 발현 벡터 내로 포함되는 과정에 대해서 제한하지 않는다. 본 분야에서의 기술 분야는 발현 벡터 내로 폴리뉴클레오티드를 포함하는 필수단계와 익숙하다. 일반적인 발현 벡터는, 리보솜 결합 부위(ribosome binding site)와 함께, 하나 또는 그 이상의 조절영역(regulatory regions)에 의해 앞서는(preceded) 목적하는 폴리뉴클레오티드, 예를 들어 *E. coli*에서 길지에서 3-9 뉴클레오티드이고, 개시코돈(initiation codon)의 3-11 뉴클레오티드의 업스트림(upstream)에 위치한 뉴클레오티드 서열을 포함한다. Shine and Dalgarno (1975) Nature 254(5495):34-38 and Steitz (1979) Biological Regulation and Development (ed. Goldberger, R. F.), 1:349-399 (Plenum, New York)을 참조하라.
- [0116] 본원에서 사용된 바와 같은 용어 "작동가능하게 연결된(operably linked)"은, 조절 서열(control sequence)이

폴리펩티드 발현을 지시하는 것과 같이, 조절 서열이 DNA 서열 또는 폴리뉴클레오티드의 코딩 서열(coding sequence)에 관하여 적절한 위치에 놓인, 배열(configuration)을 나타낸 것이다.

[0117] 조절 영역(Regulatory regions)은 예를 들어, 프로모터 및 작동유전자(operator)를 포함하는 이러한 영역을 포함한다. 프로모터는 목적하는 폴리뉴클레오티드에 작동가능하게 연결되었고, 이에 의해서 RNA 폴리메라제 효소(polymerase enzyme)를 통해 폴리뉴클레오티드의 전사(transcription)를 개시할 수 있다. 작동유전자는 프로모터에 인접한 핵산 서열이고, 이는 억제 단백질(repressor protein)이 결합할 수 있는 단백질-결합 도메인을 포함한다. 억제 단백질의 부재에서, 전사는 프로모터를 통해서 개시한다. 나타났을 때, 억제 단백질은 작동유전자에 결합하는 작동유전자의 단백질-결합 도메인에 특이적이고, 이에 의해 전사는 저해된다. 이러한 방식에서, 전사의 조절은, 사용된 특정한 조절 영역 및 이에 대응하는 억제 단백질(repressor protein)의 존재 또는 부재를 기반으로, 성취되었다. 예는 락토오스 프로모터[Lad 억제 단백질은 락토오스와 결합하였을 때 입체구조(conformation)가 변하고 이에 의해서 작동유전자에 결합으로부터 Lad 억제 단백질을 막을 수 있다] 및 트립토판 프로모터(트립토판과 복합되었을 때, TrpR 억제 단백질은 작동유전자에 결합하는 입체구조를 갖는다 ; 트립토판의 부재에서 TrpR 억제 단백질은 작동유전자에 결합하지 않는 입체구조를 갖는다)를 포함한다. 또 다른 예는 tac 프로모터이다(de Boer *et al.*, (1983) Proc Natl Acad Sci USA 80(1):21-25 를 참조). 본 분야에서의 기술 분야에서 인정된 바와 같이, 이러한 및 그 밖의 발현 벡터는 본 발명에서 사용될 수 있고, 본 발명은 이러한 측면으로 한정하지 않는다.

[0118] 어떠한 적절한 발현 벡터는 목적하는 서열을 포함하는데 사용될 수 있을지라도, 쉽게 입수가 가능한 발현 벡터는 pSC101, pBR322, pBBR1MCS-3, pUR, pEX, pMR100, pCR4, pBAD24, pUC19 와 같은 플라스미드 ; M1 3 파지(M1 3) 및 X 파지와 같은 박테리오파지를 포함하지만 이로 제한되지 않는다. 물론, 이러한 발현 벡터는 특정한 숙주 세포에 대해서만 적절할 수 있다. 그러나 본 분야의 기술분야 중 하나는, 어떠한 특정 발현 벡터가 어떠한 주어진 숙주 세포에 적절한지 여부를 일상적인 실험을 통해 쉽게 결정할 수 있다. 예를 들어, 발현 벡터를 숙주 세포 내로 도입할 수 있고, 이는 벡터에 포함된 서열의 발현 및 실행가능성(viability)에 대해서 모니터링하였다. 게다가, 참조문헌은 관련된 문맥 및 문헌을 포함하고, 이는 어떠한 특정한 숙주 세포에 대한 발현 벡터 및 이들의 적합성(suitability)을 기재하였다.

[0119] 본 발명의 숙주 세포

[0120] 본원의 본 발명은 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포에 관한 것이고, 상기 폴리펩티드는 세포 내로 셀로텍스트린 또는 오타당을 수송한다. 본원에 기재된 추가적인 발명은 숙주 세포 내로 셀로텍스트린 수송을 증가시키는 방법, 셀로텍스트린을 포함하는 배지에서 숙주 세포의 성장을 증가시키는 방법, 셀룰로스-유도 및 헤미셀룰로스-유도 당을 공동-발효시키는 방법 및 세포 내로 셀로텍스트린을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하여 탄화수소 또는 탄화수소 유도체를 생성하는 방법에 관한 것이다. 본원에 기재된 추가적인 발명은 숙주 세포 내로 오타당의 수송을 증가시키는 방법, 오타당 당을 포함하는 배지에 숙주 세포의 성장을 증가시키는 방법, 세포 내로 오타당을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하여 탄화수소 또는 탄화수소 유도체를 생성하는 방법에 관한 것이다.

[0121] "숙주 세포" 및 "숙주 미생물(host microorganism)"은, 재조합 DNA 또는 RNA 의 삽입을 통해 형질전환시킬 수 있는 살아있는 생물학적 세포(living biological cell)에 관한 것에 대해서 본원에서 교환할 수 있게(interchangeably) 사용되었다. 이러한 재조합 DNA 또는 RNA 는 발현 벡터 일 수 있다. 따라서, 본원에 기재된 바와 같은 숙주 유기체 또는 세포는 원핵 유기체(prokaryotic organism)[예를 들어, 세균계(kingdom Eubacteria)의 유기체] 또는 진핵 세포(eukaryotic cell)이다. 본 분야의 기술 분야에 의해 예측될 수 있는 바와 같이, 원핵세포는 막-결합 핵(membrane-bound nucleus)이 결핍되어 있는 반면에, 진핵세포는 막-결합 핵을 갖는다.

- [0122] 어떠한 원핵 또는 진핵 숙주 세포는, 핵산의 서열로 형질전환된(transformed) 후에 실행 가능하게 남아있기만 하면 본 발명에서 사용될 수 있다. 바람직하게, 숙주 세포는 필수적인 핵산 서열의 형질도입(transduction), 그 다음의 단백질의 발현(예를 들어 운반체), 또는 결과적으로 생성된 중간물질(intermediates)에 의해 불리하게 영향을 받지 않는다. 적절한 진핵 세포는 균류(fungal), 식물, 곤충 또는 포유류의 세포를 포함하지만 이로 제한되지 않는다.
- [0123] 바람직한 실시형태에서, 숙주 세포는 균류 스프레인(fungal strain)이다. 본원에 사용된 바와 같은 "균류(fungi)"은 문(phyla) 자낭균류(Ascomycota), 담자균류(Basidiomycota), 호상균류(Chytridiomycota) 및 접합균류(Zygomycota)(as defined by Hawksworth *et al.*, In, Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK), 뿐만 아니라 난균강(Oomycota)(Hawksworth *et al.*, 1995, supra, page 171 에 인용된) 및 모든 불완전균류(mitosporic fungi)(Hawksworth *et al.*, 1995, supra)를 포함한다
- [0124] 특정 실시형태에서, 균류 숙주(fungal host)는 효소 스트레인이다. 본원에서 사용된 바와 같은 "효모"은 자낭포자형성효모(ascosporogenous yeast)[자낭균 효모(Endomycetales)], 담자포자효모(basidiosporogenous yeast), 및 불완전균류(Fungi Imperfecti)[불완전효모강(Blastomycetes)]에 속하는 효모를 포함한다. 효모의 분류는 앞으로도 변화될 수 있기 때문에, 본 발명의 목적에서, 효모는 Biology and Activities of Yeast 에 기재된 바와 같이 규정될 수 있다(Skinner, F. A., Passmore, S. M., and Davenport, R. R., eds, Soc. App. Bacteriol. Symposium Series No. 9, 1980).
- [0125] 보다 바람직한 실시형태에서, 효모 숙주는 칸디다(*Candida*), 한제놀라(*Hansenula*), 클루이베르마이세스(*Kluyveromyces*), 피치아(*Pichia*), 사카로미세스(*Saccharomyces*), 스킴조사카로미세스(*Schizosaccharomyces*) 또는 야로이아(*Yarrowia*) 스트레인이다.
- [0126] 특정 실시형태에서, 효모 숙주는 사카로미세스 칼스베르겐시스(*Saccharomyces carlsbergensis*)(Todkar, 2010), 사카로미세스 세레비시아(*Saccharomyces cerevisiae*)(Duarte et al., 2009), 사카로미세스 다아스타티쿠스(*Saccharomyces diastaticus*), 사카로미세스 도우글라시아(*Saccharomyces douglasii*), 사카로미세스 클루이베리(*Saccharomyces kluyveri*), 사카로미세스 노르벤시스(*Saccharomyces norbensis*), 사카로미세스 모나센시스(*Saccharomyces monacensis*)(GB-Analysts Reports, 2008), 사카로미세스 바야누스(*Saccharomyces bayanus*)(Kristen Publicover, 2010), 사카로미세스 파스토리아누스(*Saccharomyces pastorianus*)(Nakao et al., 2007), 사카로미세스 폼베(*Saccharomyces pombe*)(Mousdale, 2008), 또는 사카로미세스 오비포르미스(*Saccharomyces oviformis*) 스트레인이다. 그 밖의 바람직한 실시형태에서, 효모 숙주는 클루이베르마이세스 락티스(*Kluyveromyces lactis*)(O.W. Merten, 2001), 클루이베르마이세스 프라길리스(*Kluyveromyces fragilis*)(Pestal et al., 2006; Siso, 1996), 클루이베르마이세스 마르시아무스(*Kluyveromyces marxianus*)(K. Kourkoutas et al., 2008), 피치아 스티피티스(*Pichia stipitis*)(Almeida et al., 2008), 칸디다 쉬하테(*Candida shehatae*)(Ayhan Demirbas, 2003) 또는 칸디다 트로피칼리스(*Candida tropicalis*)(Jamai et al., 2006)이다. 다른 실시형태에서, 효모 숙주는 야로이아 리폴리티카(*Yarrowia lipolytica*)(Biryukova E.N., 2009), 브레타노미세스 쿠스테르시(*Brettanomyces custersii*)(Spindler D.D. et al., 1992) 또는 지고삭카로마이세스 루시(*Zygosaccharomyces rouxii*)(Chaabane et al., 2006)일 수 있다.
- [0127] 또 다른 특정한 실시형태에서, 균류 숙주는 곰팡이 스트레인(filamentous fungal strain)이다. "곰팡이(Filamentous fungi)"는 세분(subdivision) 진균문(Eumycota) 및 난균강(Oomycota)의 모든 실 모양의 형태(all filamentous forms)를 포함한다(Hawksworth *et al.*, 1995, supra 에 의해 정의된 바와 같은). 곰팡이는, 키틴, 셀룰로스, 글루칸, 키토산(chitosan), 만난 및 그 밖의 복합 다당류(complex polysaccharides)로 이루어진 균사체의 웰(mycelial wall)로 일반적으로 특징지어진다. 영양 성장(Vegetative growth)은 균사의 연장(hyphal elongation)에 의하고, 탄소 이화작용(carbon catabolism)은 절대 호기성(obligately aerobic)이다.

그에 반하여, *사카로미세스 세레비시아*와 같은 효소에 의한 영양 성장은 단세포체(unicellular thallus)의 출아(budding)에 의한 것이고, 탄소 이화작용은 발효에 의한 것이다(fermentative).

[0128] 바람직한 실시형태에서, 곰팡이 숙주는, *아크레모늄(Acremonium)*, *아스페르길루스(Aspergillus)*, *푸사리움(Fusarium)*, *후미콜라(Humicola)*, *무코르(Mucor)*, *마이셀로프로오라(Myceliophthora)*, *네우로스포라(Neurospora)*, *페니실리움(Penicillium)*, *스시탈리디움(Scytalidium)*, *티엘라비아(Thielavia)*, *톨리포클라디움(Tolypocladium)* 또는 *트리코테리마(Trichoderma)* 스트레인이지만, 이로 한정되지 않는다.

[0129] 특정 실시형태에서, 곰팡이 숙주는, *아스페르길루스 아와모리(Aspergillus awamori)*, *아스페르길루스 포에티두스(Aspergillus foetidus)*, *아스페르길루스 자포니쿠스(Aspergillus japonicus)*, *아스페르길루스 니둘란스(Aspergillus nidulans)*, *아스페르길루스 니가(Aspergillus niger)*, 또는 *아스페르길루스 오리제(Aspergillus oryzae)* 스트레인이다. 다른 실시형태에서, 곰팡이 숙주는 *푸사리움 박트리디오이데스(Fusarium bactridioides)*, *푸사리움 세렐리스(Fusarium cerealis)*, *푸사리움 크룩웰렌스(Fusarium crookwellense)*, *푸사리움 쿨모룸(Fusarium culmorum)*, *푸사리움 그래미니아룸(Fusarium graminearum)*, *푸사리움 그래미눔(Fusarium graminum)*, *푸사리움 헤테로스포룸(Fusarium heterosporum)*, *푸사리움 네군디(Fusarium negundi)*, *푸사리움 옥시스포룸(Fusarium oxysporum)*, *푸사리움 레티쿨라툼(Fusarium reticulatum)*, *푸사리움 로세움(Fusarium roseum)*, *푸사리움 삼부시눔(Fusarium sambucinum)*, *푸사리움 사로크로움(Fusarium sarcochroum)*, *푸사리움 스포로트리치오이데스(Fusarium sporotrichioides)*, *푸사리움 숄푸레움(Fusarium sulphureum)*, *푸사리움 토룰로숨(Fusarium torulosum)*, *푸사리움 트리초테시오이데스(Fusarium trichothecioides)*, 또는 *푸사리움 베네나툼(Fusarium venenatum)* 스트레인이다. 또 다른 바람직한 실시형태에서, 곰팡이 숙주는 *후미콜라 인솔렌스(Humicola insolens)*, *후미콜라 라누기노사(Humicola lanuginosa)*, *무코르 미에헤이(Mucor miehei)*, *마이셀로프로토라 서모필라(Myceliophthora thermophila)*, *네우로스포라 크라사(Neurospora crassa)*, *페니실리움 푸르푸로게눔(Penicillium purpurogenum)*, *스시탈리디움 서모필리움(Scytalidium thermophilum)*, *스포로트리쿰 서모필(Sporotrichum thermophile)(Topakas et al., 2003)*, 또는 *티엘라비아 테레스트리스(Thielavia terrestris)* 스트레인이다. 추가적인 실시형태에서, 곰팡이 숙주는 *트리코텔마 하리지아눔(Trichoderma harzianum)*, *트리코테르마 코닝기(Trichoderma koningii)*, *트리코테르마 롱기브라치아툼(Trichoderma longibrachiatum)*, *트리코테르마 레세이(Trichoderma reesei)* 또는 *트리코테르마 비리드(Trichoderma viride)* 스트레인이다.

[0130] 그 밖의 바람직한 실시형태에서, 숙주 세포는 원핵(prokaryotic)이고, 특정 실시형태에서, 원핵생물은 *E. coli*(Dien, B.S. et al., 2003; Yomano, L. P. et al., 1998; Moniruzzaman et al., 1996), *코초균(Bacillus subtilis)*(Susana Romero et al., 2007), *자이모모나스 모빌리스(Zymomonas mobilis)*(B. S. Dien et al., 2003; Weuster Botz, 1993; Alterthum and Ingram, 1989), *클로스트리듐(Clostridium sp.)*(Zeikus, 1980; Lynd et al., 2002; Demain et al., 2005), *클로스트리듐 파이토퍼멘탄스(Clostridium phytofermentans)*(Leschine S., 2010), *클로스트리듐 서모셀룸(Clostridium thermocellum)*(Lynd et al., 2002), *클로스트리듐 베이저린키이(Clostridium beijerinckii)*(Giles Clark, 2008), *클로스트리듐 아세토부틸리쿰(Clostridium acetobutylicum)*[*무렐라 서모아세티카(Moorella thermoacetica)*](Huang W.C. et al., 2004; Dominik et al., 2007), *서모아네로박테리움 사카로락티쿰(Thermoanaerobacterium saccharolyticum)*(Marietta Smith, 2009) 또는 *켈브시엘라 옥시토카(Klebsiella oxytoca)*(Dien, B.S. et al., 2003; Zhou et al., 2001; Brooks and Ingram, 1995)이다. 또 다른 실시형태에서, 원핵 숙주 세포는 *카르복시도셀라(Carboxydocella sp.)*(Dominik et al., 2007), *코리네박테리움 글루타미쿰(Corynebacterium glutamicum)*(Masayuki Inui, et al., 2004), *엔테로박테리아세아에(Enterobacteriaceae)*(Ingram et al., 1995), *어르위니아 크리산테미(Erwinia chrysanthemi)*(Zhou and Ingram, 2000; Zhou et al., 2001), *락토바실리우스(Lactobacillus sp.)*(McCasky, T.A., et al., 1994), *페디오코쿠스 아시디락티시(Pediococcus acidilactici)*(Zhou, S. et al., 2003), *로도프세우도모나스 캡슐라타(Rhodopseudomonas capsulata)*(X.Y. Shi et al., 2004), *스트렙토코쿠스 락토스(Streptococcus lactis)*(J.C. Tang et al., 1988), *비브리오 푸르니시(Vibrio furnissii)*(L.P. Wackett, 2010), *비브리오 푸르니시 M1(Vibrio furnissii M1)*(Park et al., 2001), *칼디셀룰로시루프트로 사카로락티쿠스(Caldicellulosiruptor saccharolyticus)*(Z. Kadar et al., 2004), 또는 *잔토모나스 캄페스트리스(Xanthomonas campestris)*(S.T. Yang et al., 1987)이다. 또 다른 실시형태에서, 숙주 세포는 시아노박테리아

(cyanobacteria)이다. 박테리아 숙주 세포의 추가적인 예는 에세리키아(*Escherichia*), 엔테로박터(*Enterobacter*), 아조토박터(*Azotobacter*), 에르위니아(*Erwinia*), 바실루스(*Bacillus*), 프세우도모나스(*Pseudomonas*), 클렙시엘라(*Klebsiella*), 프로테우스(*Proteus*), 살모넬라(*Salmonella*), 세라티아(*Serratia*), 쉬겔라(*Shigella*), 리조비아(*Rhizobia*), 비트레오스실라(*Vitreoscilla*), 시네초코커스(*Synechococcus*), 시네초시스티스(*Synechocystis*), 및 파라코커스(*Paracoccus*) 탁솔노미칼 클래스(taxonomical classes)를 포함하지만 이로 제한되지 않는다.

[0131] 본 발명의 특히 바람직한 실시형태에서, 상기 숙주 세포는 사카로마이시스(*Saccharomyces* sp.), 사카로미세스 세레비시아(*Saccharomyces cerevisiae*), 사카로미세스 모나센시스(*Saccharomyces monacensis*), 사카로미세스 바야누스(*Saccharomyces bayanus*), 사카로미세스 파스토리아누스(*Saccharomyces pastorianus*), 사카로미세스 칼스베르겐시스(*Saccharomyces carlsbergensis*), 사카로미세스 폼베(*Saccharomyces pombe*), 클루이베르마이세스(*Kluyveromyces* sp.), 클루이베르마이세스 마르시아무스(*Kluyveromyces marxianus*), 클루이베르마이세스 락티스(*Kluyveromyces lactis*), 클루이베르마이세스 프라길리스(*Kluyveromyces fragilis*), 피치아 스티피티스(*Pichia stipitis*), 스포로트리쿰 서머파일(*Sporotrichum thermophile*), 칸디다 쉬하타(*Candida shehatae*), 칸디다 트로피칼리스(*Candida tropicalis*), 네우로스포라 크라사(*Neurospora crassa*), 자이모나스 모빌리스(*Zymomonas mobilis*), 클로스트리듐(*Clostridium* sp.), 클로스트리듐 파이토퍼멘탄스(*Clostridium phytofermentans*), 클로스트리듐 써모셀럼(*Clostridium thermocellum*), 클로스트리듐 베이저린키이(*Clostridium beijerinckii*), 클로스트리듐 아세토부틸리쿰(*Clostridium acetobutylicum*), 무렐라 서모아세티카(*Moorella thermoacetica*), 대장균(*Escherichia coli*), 크렙시엘라 옥시토카(*Klebsiella oxytoca*), 서모안아에로박테리움 사카로락티쿰(*Thermoanaerobacterium saccharolyticum*) 및 고초균(*Bacillus subtilis*)이다. 사카로마이시스는 산업적인(Industrial) 사카로마이시스 스트레인을 포함할 수 있다. Argueso et al. 은 바이오에탄올 생산(bioethanol production)에 중요한 특정한 유전자 다형성(specific gene polymorphisms) 뿐만 아니라 바이오에탄올 생산에 일반적으로 사용되는 산업적인 사카로마이시스 스트레인의 게놈 구조를 논의하였다(*Genome Research*, 19: 2258-2270, 2009).

[0132] 본 발명의 숙주 세포는, 재조합 핵산을 숙주 세포 내로 도입하는, 유전적으로 형질전환될 수 있고, 유전적으로 형질전환된 숙주 세포는 자연적으로 발생하지 않는다. 적절한 숙주 세포는 상이한 기능에 대해서 하나 또는 그 이상의 단백질을 코딩하는 하나 또는 그 이상의 핵산 구조(nucleic acid constructs)가 발현될 수 있는 것이다.

[0133] 본원에서 사용된 바와 같은 "재조합 핵산" 또는 "이형 핵산(heterologous nucleic acid)" 또는 "재조합 폴리뉴클레오티드"는 핵산 중합체에 관한 것이고, 하기의 적어도 하나는 사실(true)이다 : (a) 핵산의 서열이 주어진 숙주 세포와 이질적이다(foreign)(즉, 숙주 세포에서 천연적으로 발견되지 않는다) ; (b) 서열이 주어진 숙주 세포에서 천연적으로 발견되지만, 천연적이지 않는 양(unnatural amount)(예를 들어, 예상했던 양보다 많이)으로 발견되었다 ; 또는 (c) 핵산의 서열이, 천연적으로(in nature) 서로에 대해 동일한 관계(same relationship)로 발견되지 않는 둘 또는 그 이상의 서브서열(subsequences)을 포함한다. 예를 들어, 경우 (c)에 관하여, 재조합 핵산 서열은, 신규한 기능적인 핵산을 생산하기 위해 배열된 관련되지 않은 유전자(unrelated genes)로부터 둘 또는 그 이상의 서열을 갖는다. 명확하게, 본 발명은 숙주 세포 내로 발현 벡터의 도입(introduction)을 기재하였고, 상기 발현 벡터는, 숙주 세포에서 일반적으로 발견되지 않는 단백질을 코딩하는 핵산 서열을 포함하거나, 또는 숙주 세포에서 일반적으로 발견되지만 상이한 조절 서열(regulatory sequences)의 조절 하에 있는 단백질을 코딩하는 핵산을 포함한다. 숙주 세포의 게놈과 관련하여, 단백질을 코딩하는 상기 핵산 서열은 재조합되었다(recombinant).

[0134] 몇몇 실시형태에서, 상기 숙주 세포는 본 발명의 폴리뉴클레오티드에 의해 코딩되는 모든 단백질을 천연적으로 생산한다. 목적하는 단백질을 코딩하는 유전자는 숙주 세포에 대해서 이종 기원(heterologous)일 수 있거나, 또는 이러한 유전자는 숙주 세포에 대해서 내인적(endogenous)일 수 있지만, 이종 기원의 프로모터(heterologous promoters)에 작동가능하게 연결되고/연결되거나 숙주 세포에서 유전자의 보다 높은 발현을 야기

하는 영역을 조절한다. 다른 실시형태에서, 숙주 세포는 목적하는 단백질을 천연적으로 생산하지 않고, 이러한 분자를 생산하기 위해 필요한 하나 또는 그 이상의 유전자를 발현시킬 수 있는 이중 기원 핵산 구조물(heterologous nucleic acid constructs)을 포함한다.

[0135] 핵산 분자 또는 폴리펩티드 및 특정한 세포 또는 미생물과 관련하여 본원에서 사용된 바와 같은 "내인성(Endogenous)" 은 재조합 유전자 조작 기술(recombinant engineering techniques)을 사용한 세포 내로 도입되지 않고, 세포 내에 있는 핵산 서열 또는 펩티드에 관한 것이다 ; 예를 들어 세포를 천연으로부터 본래 분리하였을 때, 세포에 존재하는 유전자.

[0136] "유전적으로 조작된(Genetically engineered)" 또는 "유전적으로 변형된(genetically modified)"은, 높은 수준(elevated levels), 낮은 수준(lowered levels)으로 또는 돌연변이된 형태(mutated form)로 단백질을 발현시키는 원핵 또는 진핵 숙주 세포를 생산하는데 사용된 모든 재조합 DNA 또는 RNA 방법에 관한 것이다. 다시 말해서, 숙주 세포를 재조합 폴리뉴클레오티드 분자로 감염시키거나(transfected), 형질전환시키거나(transformed), 또는 형질 도입시킨 것으로(transduced), 이것 때문에 목적하는 단백질의 발현을 변경시키는 세포가 발생하여 변경되었다. 유전적으로 조작된 숙주 세포에 대한 방법 및 벡터는 본 분야에서 널리 공지되어 있다 ; 예를 들어 다양한 기술은 Molecular Biology, Ausubel *et al.*, eds. (Wiley & Sons, New York, 1988, and quarterly updates)에서의 현재의 프로토콜에 기재되어 있다. 유전적으로 조작된 기술(Genetically engineering techniques)은, 발현 벡터, 목표하는 상동 재조합(targeted homologous recombination) 및 유전자 활성화(gene activation)(예를 들어, U.S. Pat. No. 5,272,071 to Chappel 참조) 및 유전자 조작된 전사 인자에 의한 트랜스- 활성화(trans- activation)[예를 들어, Segal *et al.*, (1999) Proc Natl Acad Sci USA 96(6):2758-2763 참조]를 포함하지만 이로 제한되지 않는다.

[0137] 유전자 발현 또는 기능에서 증가를 야기하는 유전적인 변형(Genetic modifications)은, 유전자의 증폭(amplification), 과잉생산(overproduction), 과잉발현(overexpression), 활성화(activation), 증대(enhancement), 첨가 또는 상향-조절(up-regulation)로 언급될 수 있다. 보다 명확하게, 본원에 기재된 효소 또는 그 밖의 단백질의 작용(action)(또는 활성도)를 증가시키는 것과 관련하여, 효소 또는 단백질의 증가된 발현 및/또는 기능성(생물학적 활성도)을 야기하는 숙주 세포의 어떠한 유전적인 변형을 일반적으로 언급하고, 단백질의 보다 높은 활성도 또는 작용(예를 들어, 특이적인 활성도 또는 생체 내 효소적 활성도), 감소된 저해 또는 단백질의 분해(degradation), 및 단백질의 과잉발현을 포함한다. 예를 들어, 유전자 복제수(gene copy number)는 증가될 수 있고, 발현 수준은 천연의 프로모터의 발현 수준보다 더 높은 발현 수준을 나타내는 프로모터의 사용에 의해 증가될 수 있거나, 또는 유전자가 단백질의 작용 또는 효소의 생물학적 활성도를 증가시키기 위해 유전적으로 조작되거나 또는 일반적인 돌연변이 생성으로 변경될 수 있다. 이러한 변형의 몇몇의 조합은 또한 가능하다.

[0138] 유전자 발현, 유전자의 기능 또는 유전자 산물(즉, 유전자에 의해 코드되는 단백질)의 기능에서 감소를 야기하는 유전적인 변형(genetic modification)은, 유전자 발현의 불활성화(inactivation)(완전한 또는 부분적인), 결실, 중단(interruption), 장애(blockage), 침묵(silencing), 하향-조절(down-regulation), 감쇠(attenuation)로 언급될 수 있다. 예를 들어, 이러한 유전자에 의해 코드되는 단백질의 기능에서의 감소를 야기하는 유전자에서의 유전적인 변형은, 유전자의 완전한 결실(즉, 유전자가 존재하지 않고, 따라서 단백질이 존재하지 않는다), 단백질의 불완전하거나 또는 번역이 없는 것을 야기하는 유전자에서의 돌연변이(예를 들어, 단백질이 발현되지 않는다), 또는 단백질의 천연적인 기능을 감소시키거나 폐지하는(abolish) 유전자에서의 돌연변이(예를 들어, 효소적인 활성도 또는 작용이 없거나 감소된 단백질이 발현된다)의 결과일 수 있다. 보다 명확하게, 본원에 기재된 단백질의 작용을 감소시키는 언급은, 숙주 세포에서 모든 유전적인 변형에 대해서 일반적으로 나타내었고, 단백질의 감소된 발현 및/또는 기능성(생물학적 활성도)을 야기하고, 단백질의 감소된 활성도(예를 들어, 감소된 수송), 단백질의 증가된 저해 또는 분해 뿐만 아니라 단백질의 발현의 감소 또는 저해를 포함한다. 예를 들어, 본 발명의 단백질의 작용 또는 활성도는 단백질 생산의 저지(blocking) 또는 감소, 단백질 작용의 감소, 또는 단백질 작용의 저해에 의해 감소될 수 있다. 이러한 변형의 몇몇을 또한 조합할 수

있다. 단백질 생산의 저지 또는 감소는, 성장 배지에서 유도된 화합물의 존재를 필요로 하는 프로모터의 조절 하에서 단백질을 코딩하는 유전자를 배치하는 것을 포함할 수 있다. 유도물질(inducer)이 배지로부터 대폭 감소되는 것과 같은 조건을 설정함으로써, 단백질을 코딩하는 유전자의 발현(및 이에 따른 단백질의 합성)은 정지될 수 있다(turned off). 단백질의 작용을 저지하거나 감소시키는 것은 본원에서 참조문헌으로 인용된 U.S. Pat. No. 4,743,546 에 기재된 것과 유사한 제거 기술 접근법(excision technology approach)을 사용하는 것을 또한 포함할 수 있다. 이러한 접근법을 사용하기 위해, 흥미로운 단백질을 코딩하는 유전자는, 계놈으로부터 유전자가 특이적이고, 조절된 제거를 가능하게 하는 특이적인 유전적인 서열(specific genetic sequences) 사이에 클론되었다(cloned). 제거(Excision)는, 예를 들어 U.S. Pat. No. 4,743,546 에 기재된 배양물의 배양 온도에서의 시프트(shift), 또는 몇몇 다른 물리적이거나(physical) 또는 영양상의 신호(nutritional signal)에 의해 유도된다(prompted).

[0139]

일반적으로, 본 발명에 따른, 돌연변이체(mutant) 또는 변형된 단백질(modified protein)의 주어진 특성에서의 증가 또는 감소(예를 들어, 효소 활성도, 수송 화합물의 능력)는, 동일한 또는 동등한 조건(equivalent conditions) 하에서 측정되거나 또는 확립된(established), 동일한 유기체[동일한 근원(same source) 또는 부모 서열(parent sequence)로부터]로부터 유도된 야생형(즉, 보통, 변형되지 않은) 단백질의 동일한 특성과 관련하여 행해졌다. 유사하게, 유전적으로 변형된 숙주 세포의 특성의 증가 또는 감소(예를 들어, 단백질의 발현 및/또는 생물학적 활성도, 또는 산물의 생산)는, 동일하거나 또는 동등한 조건 하에서, 동일한 종 및 바람직하게 동일한 스트레인의 야생형 숙주 세포의 동일한 특성과 관련하여 행해졌다. 이러한 조건은, 단백질의 활성도(예를 들어, 발현 또는 생물학적 활성도) 또는 숙주 세포의 그 밖의 특성을 측정하는 것 뿐만 아니라 검정 유형의 사용, 숙주 세포를 측정 하에서, 검정 또는 배양 조건(예를 들어, 배지 조성, 온도, pH 등)을 포함한다. 상기에 나타난 바와 같이, 동등한 조건은, 유사하지만 동일하지 않고[예를 들어, 조건에서의 몇몇의 보존적인 변화(conservative changes)는 용인될 수 있다], 동일한 조건 하에서 만들어진 비교(comparison)와 비교함으로써 세포 성장 또는 효소 발현 또는 생물학적 활성도에서의 효과를 실질적으로 변화시키지 않는 조건(예를 들어, 배양 조건)이다.

[0140]

바람직하게, 주어진 단백질(예를 들어, 운반체, 효소)의 활성도를 증가시키거나 감소시키는 유전적인 변형을 갖는 유전적으로 변형된 숙주 세포는, 적어도 약 5 %, 및 보다 바람직하게 적어도 약 10 %, 및 보다 바람직하게 적어도 약 15 %, 및 보다 바람직하게 적어도 약 20 %, 및 보다 바람직하게 적어도 약 25 %, 및 보다 바람직하게 적어도 약 30 %, 및 보다 바람직하게 적어도 약 35 %, 및 보다 바람직하게 적어도 약 40 %, 및 보다 바람직하게 적어도 약 45 %, 및 보다 바람직하게 적어도 약 50 %, 및 보다 바람직하게 적어도 약 55 %, 및 보다 바람직하게 적어도 약 60 %, 및 보다 바람직하게 적어도 약 65 %, 및 보다 바람직하게 적어도 약 70 %, 및 보다 바람직하게 적어도 약 75 %, 및 보다 바람직하게 적어도 약 80 %, 및 보다 바람직하게 적어도 약 85 %, 및 보다 바람직하게 적어도 약 90 %, 및 보다 바람직하게 적어도 약 95 %, 및 5 % 에서 100 % 사이의 전체 정수(예를 들어, 6 %, 7 %, 8 % 등)인 모든 퍼센트의 야생형 숙주 세포에서의 야생형 단백질의 활성도와 비교한 바와 같이, 각각 단백질의 활성도 또는 작용(예를 들어, 발현, 생산 및/또는 생물학적 활성도)에서 증가되거나 또는 감소시킨다. 분리된 야생형 핵산 분자 또는 단백질에 대해 직접적으로 분리된 변형된 핵산 분자 또는 단백질과 비교하였을 때(예를 들어, 만약 생체 내(in vivo)와 비교한 생체 외(in vitro)에서 비교한다면), 별로 다를 거 없는 것(same differences)이 바람직하다.

[0141]

본 발명의 또 다른 양상에서, 주어진 단백질(예를 들어, 운반체 및 효소)의 활성도가 증가되거나 또는 감소된 유전적인 변형을 갖는 유전적으로 변형된 숙주 세포는, 적어도 약 2-배(fold), 및 보다 바람직하게 적어도 약 5-배, 및 보다 바람직하게 적어도 약 10-배, 및 보다 바람직하게 적어도 약 20-배, 및 보다 바람직하게 적어도 약 30-배, 및 보다 바람직하게 적어도 약 40-배, 및 보다 바람직하게 적어도 약 50-배, 및 보다 바람직하게 적어도 약 75-배, 및 보다 바람직하게 적어도 약 100-배, 및 보다 바람직하게 적어도 약 125-배, 및 보다 바람직하게 적어도 약 150-배, 또는 적어도 약 2-배로부터 시작하는 전체 정수 증가(integer increment)(예를 들어, 3-배, 4-배, 5-배, 6-배 등)의 야생형 숙주 세포에서의 야생형 단백질의 활성도와 비교한 바와 같이, 각각 단백질의 활성도 또는 작용(예를 들어, 발현, 생산 및/또는 생물학적 활성도)에서 증가되거나 감소되었다.

[0142] 숙주 세포 성분

[0143] 하나의 양상에서, 본 발명의 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO : 1 을 포함한다. 또 다른 양상에서, 본 발명의 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO : 2 를 포함한다. 또 다른 양상에서, 본 발명의 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 2 및 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프는 SEQ ID NO : 3 을 포함한다. 또 다른 양상에서, 본 발명의 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO : 4 를 포함한다. 또 다른 양상에서, 본 발명의 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO : 5 를 포함한다. 또 다른 양상에서, 본 발명의 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하고, 상기 트랜스멤브레인 α -헬릭스 6 및 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO : 6 을 포함한다. 또 다른 양상에서, 본 발명의 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO : 7 을 포함한다. 또 다른 양상에서, 본 발명의 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 10 및 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO : 8 을 포함한다.

[0144] 각각의 상기한 양상은 많은 수의 조합으로 조합될 수 있다. 숙주 세포는 SEQ ID NOs: 1-8 중의 1, 2, 3, 4, 5, 6, 또는 7 을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함할 수 있거나, 폴리펩티드는 SEQ ID NOs: 1-8 의 모두를 포함할 수 있다. 예를 들어, 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO: 1 을 포함하고, 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO : 2 를 포함하고, 트랜스멤브레인 α -헬릭스 2 및 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프는 SEQ ID NO : 3 을 포함하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO : 7 을 포함한다. 또는, 또 다른 예에서, 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하고, 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO: 2 를 포함하고, 트랜스멤브레인 α -헬릭스 3 은 SEQ ID NO: 3 을 포함하고, 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO: 5 를 포함하고, 트랜스멤브레인 α -헬릭스 10 및 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO : 8 을 포함한다.

[0145] 상기한 양상의 특정 실시형태에서, 상기 폴리펩티드는 NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 % 또는 적어도 100 % 의 아미노산 동일성을 갖는다.

[0146] 바람직한 실시형태에서, 숙주 세포는 추가로 폴리뉴클레오티드를 포함하고, 상기 폴리뉴클레오티드는 β -글루코시다아제의 촉매 도메인(catalytic domain)을 코딩한다. 본원에 사용된 바와 같이, β -글루코시다아제는 β -D-글루코시드 글루코히드롤라아제(glucohydrolase)(E.C. 3.2.1.21)을 나타내고, β -D-글루코스를 방출하는 말단 비-환원(terminal non-reducing) β -D-글루코스 잔기의 가수분해를 촉매한다. β -글루코시다아제의 촉매 도메인은, 예를 들어, Venturi et al., 2002 에 의해 기재된 기본적인 절차(basic procedure)에 따라 확인된 바와 같은 β -글루코시다아제 활성도를 갖는다. β -글루코시다아제의 촉매 도메인은 글루코스를 방출하는 β -D-글루코시드에서 말단의 비-환원 잔기의 가수분해를 촉매하는 모든 도메인이다. 바람직한 실시형태에서, β -글루코시다아제는 글리코실 가수분해효소 패밀리 1 멤버(glycosyl hydrolase family 1 member)이다. 이러한 그룹의 멤버는 모티프, [LIVMFSTC] - [LIVFYS] - [LIV] - [LIVMST] - E - N - G - [LIVMFAR] - [CSAGN] (SEQ ID NO: 14)에 의해 확인할 수 있다. 여기에서, E 는 촉매작용의 글루타메이트이다(웹페이지 expasy.org/cgi-bin/prosite-search-ac?PDOC00495). 특정 실시형태에서, β -글루코시다아제의 촉매 도메인을 코딩하는 폴리뉴클레오티드는 숙주 세포에 대해서 이종기원이다. 바람직한 실시형태에서, β -글루코시다아제의 촉매 도메인은 숙주 세포에서 세포 내에 위치한다. 바람직한 실시형태에서, β -글루코시다아제는 *N. 크라사*로부터 온 것이고, 특히 바람직한 실시형태에서, β -글루코시다아제는 NCU00130 이다. 특정 실시형태에서, β -글루코시다아제는 NCU00130 의 오솔로그일 수 있다. NCU00130 의 오솔로그는, *T. 멜라노스포룸(T. melanosporum)*, CAZ82985.1; *A. 오리제(A. oryzae)*, BAE57671.1 ; *P. 플라센타(P. placenta)*, EED81359.1; *P. 크라이소스포리움(P. chrysosporium)*, BAE87009.1; *클루이베로마이세스 락티스(Kluyveromyces lactis)*, CAG99696.1; *락카리아 비칼라(Laccaria bicolor)*, EDR09330; *칼라비스포라 루시타니아(Clavispora lusitaniae)*, EEQ37997.1; 및 *피치아 스티피티스(Pichia stipitis)*, ABN67130.1 를 포함하지만 이로 한정하지 않는다. 다른 β -글루코시다아제는 글리코실 가수분해 패밀리 3(glycosyl hydrolase family 3)으로부터 온 것을 사용할 수 있다. 이러한 β -글루코시다아제는 PROSITE 에 따른 하기의 모티프로부터 확인할 수 있다 : [LIVM](2) - [KR] - x - [EQKRD] - x(4) - G - [LIVMFTC] - [LIVT] - [LIVMF] - [ST] - D - x(2) - [SGADNIT] (SEQ ID NO: 15). 여기에서 D 는 촉매의 아스파르트산염(catalytic aspartate)이다. 일반적으로, 모든 β -글루코시다아제는 NCBI 서열 COG2723 에서 발견된 β -글루코시다아제/6-포스포- β -글루코시다아제/ β -갈락토시다아제의 보존 도메인(conserved domain)을 포함한 것을 사용할 수 있다. 특정한 β -글루코시다아제로부터의 촉매 도메인은 숙주 세포에서 포함된 셀로텍스트린 운반체에 따라 선호될 수 있다.

[0147] 특정 실시형태에서, 숙주 세포는 하나 또는 그 이상의 폴리뉴클레오티드를 포함하고, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용(pentose utilization)에서 포함되는 하나 또는 그 이상의 효소를 코딩한다. 하나 또는 그 이상의 폴리뉴클레오티드는 숙주 세포에 대해서 내인성(endogenous) 또는 이종 기원(heterologous)일 수 있다. 본원에서 사용된 바와 같은 오타당은 5 가지 탄소 원자를 갖는 모든 단당류를 나타낸다. 오타당의 예는 자일로스, 아라비노스, 만노스, 갈락토스 및 람노오스(rhamnose)를 포함하지만 이로 한정되지 않는다. 오타당 이용에 포함된 하나 또는 그 이상의 효소는, 예를 들어 어떠한 결합으로 L-아라비노스 이성질화효소, L-리블로인산화효소, L-리블로스-5-P 4 에피머화효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비니톨 4-탈수소효소, L-자일루로스 환원효소 및 자일리톨 탈수소효소를 포함할 수 있다. 이러한 효소는 오타당 당을 천연적으로 대사작용하는 모든 유기체로부터 올 수 있다. 이러한 유기체의 예는, 예를 들어 *클루이베로마이세스(Kluyveromyces sp.)*, *자이모모나스(Zymomonas sp.)*, *E. coli*, *클로스트리듐(Clostridium sp.)* 및 *피치아(Pichia sp.)*를 포함한다.

[0148] 실시예 12-15 는, 숙주 세포는 개선될 수 있거나, 보다 효율적으로 만들어질 수 있는, 오타당 이용 경로에서의 방법을 기재한 것이다. 숙주 세포의 스트레인 백그라운드(Strain background)는 이의 오타당 이용 경로(pentose utilization pathway)의 효율성(efficiency)에 영향을 줄 수 있다. 숙주 세포가 *사카로미세스(Saccharomyces sp.)*인 본 발명의 실시형태에서, 바람직한 오타당 이용은 DA24-16(실시예 13 참조) 및 L2612(실시예 16 참조)을 포함한다. 오타당 이용에 포함되는 효소를 코딩하는 폴리뉴클레오티드를 포함하는 그 밖의 숙주 세포는 DuPont *자이모모나스(Zymomonas)* 스트레인(WO 2009/058927) 및 *사카로미세스(Saccharomyces)* 스트레인(US 5,789,210)을 포함한다.

[0149] 본 발명의 특정 실시형태에서, 숙주 세포는 오탄당 운반체를 코딩하는 재조합 폴리뉴클레오티드를 포함한다. 특정 실시형태에서, 오탄당 운반체는 본원에서 발견되고 기재된, NCU00821, NCU04963, NCU06138, STL12/XUT6, SUT2, SUT3, XUT1 및 XUT3(실시예 11 을 참조)을 포함하는 이러한 운반체를 포함한다. 다른 실시형태에서, 오탄당 운반체는, *C. 인터메디아*(*C. intermedia*)로부터의 Gxsl, *P. 스티피티스*로부터의 Aut1, *D. 한세니*(*D. hansenii*)로부터의 Xylhp(Nobre *et al.*, 1999), *K. 마르시아우스*(*K. marxianus*)로부터의 자일로스 운반체(Stambuk *et al.*, 2003), *암브로시오즈마 모노스포라*(*Ambrosiozyma monospora*)로부터의 LAT1 and LAT2(EMBL AY923868 and AY923869, respectively, R. Verho *et al.*), *C. 아라비노페멘탄스*(*C. arabinofementans*)로부터의 ART1(Fonseca *et al.*, 2007), *K. 마르시아우스*(*K. marxianus*)로부터의 KmLAT1(Knoshaug *et al.*, 2007), *P. 윌리엄슨디*(*P. guilliermondii*)로부터의 PgLAT2(Knoshaug *et al.*, 2007), 및 *P. 스티피티스*(*P. stipitis*)로부터의 araT(Boles & Keller, 2008)을 포함할 수 있다.

[0150] 본 발명의 숙주 세포를 생산하고 배양하는 방법

[0151] 본원에서 본 발명은 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포에 관한 것이고, 상기 폴리펩티드는 세포 내로 셀로렉스트린 또는 오탄당을 수송한다. 본원에 기재된 추가적인 본 발명은 숙주 세포 내로 셀로렉스트린의 수송을 증가시키는 방법, 셀로렉스트린을 포함하는 배지에서 숙주 세포의 성장을 증가시키는 방법, 셀로렉스트린-유도 및 헤미셀로렉스트린 유도 당을 공동-발효시키는 방법, 세포 내로 셀로렉스트린을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 준비하여 탄화수소 또는 탄화수소 유도체를 생성하는 방법에 관한 것이다. 본원에 기재된 추가적인 본 발명은 숙주 세포 내로 오탄당의 수송을 증가시키는 방법, 오탄당 당을 포함하는 배지에서 숙주 세포의 성장을 증가시키는 방법, 및 세포 내로의 오탄당을 수송하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 준비하여 탄화수소 또는 탄화수소 유도체를 생성하는 방법에 관한 것이다.

[0152] 본 발명의 숙주 세포를 생성하고 배양하는 방법은 숙주 세포 내로 본 발명의 재조합 폴리뉴클레오티드를 포함하는 발현 벡터의 도입 또는 수송을 포함할 수 있다. 숙주 세포 내로의 발현 벡터를 수송하는 이러한 방법은 본 분야에서 기술분야에 널리 공지되어 있다. 예를 들어, 발현 벡터를 갖는 *E. coli* 를 변형시키는 방법은 염화칼슘 처리를 포함하고, 상기 발현 벡터는 칼슘 침전물(calcium precipitate)을 통해 도입된다. 그 밖의 염, 예를 들어 인산칼슘은 하기의 유사한 절차를 또한 사용할 수 있다. 게다가, 일렉트로포레이션(electroporation)(즉, 핵산 서열에 대한 세포의 투과성(permeability)을 증가시키기 위해 전류의 적용)을 숙주 세포를 감염시키는데(transfect) 사용될 수 있다. 또한, 핵산 서열의 미세주입법(microinjection)은 숙주 세포를 감염시킬 수 있는 능력을 제공한다. 지질 복합체(lipid complexes), 리포솜(liposomes) 및 덴드리머(dendrimers)와 같은 그 밖의 방법을 또한 사용할 수 있다. 본 분야에서의 기술분야의 이러한 것들은 이러한 또는 그 밖의 방법을 사용하여 목적하는 서열을 갖는 숙주 세포를 감염시킬 수 있다.

[0153] 벡터는 염색체 복제[예를 들어, 플라스미드, 염색체 외 요소(extrachromosomal element), 미니염색체(minichromosome), 또는 인공적인 염색체]에 독립적인 복제, 염색체 외 독립체(extrachromosomal entity)로서 존재하는 독자적으로 복제하는 벡터(autonomously replicating vector)일 수 있다. 벡터는 자기-복제(self-replication) 확인하기 위한 어떠한 수단을 포함할 수 있다. 그 대신에, 숙주 내로 도입되었을 때, 벡터는 게놈 내에 통합되고, 이것이 통합된 것 내로 염색체와 함께 복제되었다. 더욱이, 숙주의 게놈 내로 도입될 수 있는 전체 DNA 또는 전이인자(transposon)를 함께 포함하는, 단일 벡터 또는 플라스미드 또는 둘 또는 그 이상의 벡터 또는 플라스미드를 사용할 수 있다.

[0154] 벡터는 형질전환된 숙주(transformed hosts)의 쉬운 선택(easy selection)을 허용하는 하나 또는 그 이상의 선택가능한 마커(selectable markers)를 바람직하게 포함한다. 선택가능한 마커는, 예를 들어 살생제(biocide) 또는 바이러스성 저항(viral resistance), 중금속(heavy metals)에 대한 저항, 영양요구체(auxotrophs)에 대한 프로토포트로피(protoptrophy) 및 그 밖의 것과 같은 것의 제공에 의한 유전자 생성물이다. 박테리아 세포의 선택은 amp, gpt, neo 및 hyg 유전자와 같은 유전자에 의해 부여되는 길항미생물 저항성(antimicrobial resistance)을 기반으로 할 수 있다.

- [0155] 효모 숙주에 대한 적절한 마커(markers)는 예를 들어, ADE2, HIS3, LEU2, LYS2, MET3, TRP1 및 URA3 이다. 곰팡이 숙주(filamentous fungal host)에 사용되기 위한 선택적인 마커는, *amdS*[아세타미다아제(acetamidase)], *argB*[오르니틴 카르바모일트랜스페라아제(ornithine carbamoyltransferase)], *bar*[포스리노리친 아세틸트랜스페라아제(phosphinothricin acetyltransferase)], *hph*[히그로마이신 포스포트랜스페라아제(hygromycin phosphotransferase)], *niaD*[니트레이트 환원효소(nitrate reductase)], *pyrG*[오로티딘-5'-포스페이트 디카르복실라아제(orotidine-5'-phosphate decarboxylase)], *sC*[sul페이트 아데닐트랜스페라아제(sulfate adenyltransferase)], 및 *trpC*[안트라닐산 생성효소(anthranilate synthase)]를 포함하지만, 이로 제한되지 않는다.
- [0156] 벡터는, 계놈의 독립된 세포에서 벡터의 자율증식(autonomous replication) 또는 숙주의 계놈 내로 벡터의 통합(integration)을 허용하는 요소를 바람직하게 포함한다.
- [0157] 숙주 계놈 내로의 통합을 위해, 벡터는 상동의(homologous) 또는 비상동의 재조합(nonhomologous recombination)에 의해 계놈 내로 벡터의 통합을 위한 벡터의 모든 그 밖의 요소(element) 또는 유전자의 서열을 필요로 할 수 있다. 그 대신에, 벡터는 숙주의 계놈 내로 상동의 재조합에 의해 통합을 지시하기 위해 추가적인 뉴클레오티드 서열을 포함할 수 있다. 추가적인 뉴클레오티드 서열은 벡터가 염색체에서 정확한 위치(precise location)에서의 숙주 계놈 내로 통합될 수 있게 한다. 정확한 위치에서의 통합의 가능성(likelihood)을 증가시키기 위해, 통합의 요소(integrational elements)는 바람직하게 충분한 수의 핵산, 100 내지 10,000 개의 염기쌍, 바람직하게 400 내지 10,000 개의 염기쌍 및 가장 바람직하게 800 내지 10,000 개의 염기쌍을 바람직하게 포함하고, 이는 상동의 재조합(homologous recombination)의 개연성(probability)을 향상시키기 위해서, 이에 상응하는 목표 서열(corresponding target sequence)을 갖는 높은 상동성을 갖는다. 통합의 요소(integrational elements)는 숙주의 계놈에서 목표 서열을 갖는 상동적인 모든 서열일 수 있다. 게다가, 통합의 요소는 비-코딩(non-encoding) 또는 코딩 뉴클레오티드 서열일 수 있다. 반면에, 벡터는 비-상동성(non-homologous) 재조합에 의해 숙주의 계놈 내로 통합될 수 있다.
- [0158] 자율증식(autonomous replication)에 대해서, 벡터는 논의되고 있는 숙주에서 벡터가 자체적으로 복제할 수 있는 복제 원점(origin of replication)을 추가로 포함할 수 있다. 복제 원점은 세포에서의 작용하는 자율증식을 중재하는(mediating) 모든 플라스미드 복제인자(plasmid replicator) 일 수 있다. 용어 "복제 원점" 또는 "플라스미드 복제인자" 는 플라스미드 또는 벡터가 *생체 내에서* 복제할 수 있는 서열로서 나타난 것이다. 효모 숙주에서 사용을 위한 복제 원점의 예는 복제의 2 미크론 오리진(micron origin), ARS1, ARS4, ARS1 및 CEN3의 조합 및 ARS4 및 CEN6의 조합이다. 곰팡이 세포에서 유용한 복제 원점의 예는 AMA1 및 ANSI 이다(Gems *et al.*, 1991; Cullen *et al.*, 1987; WO 00/24883). AMA1 유전자 및 유전자를 포함하는 플라스미드 또는 벡터의 구조(construction)의 분리는 WO 00/24883에 나타난 방법에 따라 성취할 수 있다.
- [0159] 그 밖의 숙주에 대해서, 형질전환의 절차는, 예를 들어 *클루이베르마이세스*에 대해서 in Jeremiah D. Read, *et al.*, *Applied and Environmental Microbiology*, Aug. 2007, p. 5088-5096, *자모모나스*에 대해서 in Osvaldo Delgado, *et al.*, *FEMS Microbiology Letters* 132, 1995, 23-26, *피치아 스티피트스*에 대해서 US 7,501,275 및 *클로스트리디움*에 대해서 WO 2008/040387에 기재되어 있다.
- [0160] 유전자의 하나 보다 많은 복제는, 유전자 산물의 생산을 증가시키기 위해 숙주 내로 삽입할 수 있다. 유전자의 복제 수(copy number)의 증가는 숙주 계놈 내로 유전자의 적어도 하나의 추가적인 복제를 통합시키거나, 또는 세포가 선택가능한 마커의 증폭된 복제를 포함하는 경우 뉴클레오티드 서열을 갖는 증폭가능한 선택가능한 마커 유전자(amplifiable selectable marker gene)를 포함시켜서 수득할 수 있고, 이에 의해서 유전자의 추가적인 복제(additional copies)는 적절한 선택가능한 유전자의 존재 하에서 세포를 배양시켜 선택할 수 있다.

- [0161] 본 발명의 재조합 발현 벡터를 구성하기 위해 상기한 요소를 묶는데(ligate) 사용하는 절차는 본 분야에서의 숙련자에서 널리 공지되어 있다(예를 들어, Sambrook *et al.*, 1989, *supra* 참조).
- [0162] 숙주 세포는 적어도 하나의 발현 벡터와 함께 변형된다. 오직 하나의 발현벡터를 사용하였을 때[중간물질(intermediate)의 첨가 없이], 벡터는 필요한 핵산 서열의 모두를 포함할 수 있다.
- [0163] 숙주 세포를 발현 벡터로 변형시키자마자, 숙주 세포는 성장이 가능하다. 본 발명의 방법은 세포에서 재조합 핵산이 발현된 것과 같이 숙주 세포를 배양하는 것을 포함할 수 있다. 미생물의 숙주에 대해서, 이러한 과정은 적절한 배지에서 세포의 배양을 수반한다. 일반적인 세포는 적절한 배지에서 35 °C 에서 성장한다. 본 발명의 바람직한 성장 배지는, 예를 들어 Luria Bertani(LB) 브로스(broth), Sabouraud Dextrose (SD) 브로스 또는 효모 배지(Yeast medium, YM) 브로스와 같은 상업적으로 바람직한 배지를 포함한다. 그 밖의 특정되거나 합성된 성장 배지를 또한 사용할 수 있고, 특정 숙주 세포의 성장을 위한 적절한 배지는 미생물 또는 발효 과학의 분야에서의 숙련자에게 공지될 것이다. 성장을 위한 적절한 온도 범위 및 그 밖의 조건은 본 분야에 공지되어 있다(예를 들어, Bailey and Ollis 1986 참조).
- [0164] 본 발명의 몇몇의 양상에 따라, 배양 배지는 숙주 세포에 대한 탄소원(carbon source)을 포함한다. 이러한 "탄소원"은 원핵 또는 단순한 진핵 세포의 성장을 위한 탄소원으로서 사용된 적절한 화합물 또는 기질을 일반적으로 나타낸다. 탄소원은, 중합체, 탄수화물, 산, 알코올, 알데히드, 케톤, 아미노산, 펩티드 등을 포함하지 이로 제한되지 않는 다양한 형태일 수 있다. 이러한 것은, 글루코스, 올리고당, 다당류와 같은 다양한 단량체, 셀룰로스 또는 헤미셀룰로스와 같은 바이오매스, 자일로스, 아라비노스, 수크로스와 같은 이당류, 포화 또는 불포화 지방산, 숙시네이트(succinate), 락테이트(lactate), 아세테이트(acetate), 에탄올 등을 포함한다. 탄소원은 글루코스를 포함하지만 이로 한정되지 않는 추가적인 광합성(photosynthesis) 산물일 수 있다.
- [0165] 바람직한 실시형태에서, 탄소원은 셀룰로스 또는 헤미셀룰로스와 같은 바이오매스 중합체이다. 본원에 기재된 바와 같은 "바이오매스 중합체" 는 생물학적인 물질(biological material)을 포함하는 모든 중합체이다. 생물학적인 물질은 살아있거나 죽은 것일 수 있다. 바이오매스 중합체는, 예를 들어 셀룰로스, 자일란, 자일로스, 헤미셀룰로스, 리그닌, 만난 및 바이오매스에서 일반적으로 발견되는 그 밖의 물질을 포함한다. 바이오매스 중합체의 근원(sources)의 비-제한적인 예는, 풀(grasses)[예를 들어, 지팡이풀(switchgrass), *미스칸투스*(*Miscanthus*)], 왕겨(rice hull), 버게스(bagasse), cotton, jute, hemp, 아마(flax), 대나무(bamboo), 사이잘(sisal), 마닐라삼섬유(abaca), 짚(straw), 잎(leaves), 잔디 부스러기(grass clippings), 옥수수대(corn stover), 옥수수 속대(corn cobs), 증류립(distillers grains), 콩과 식물(legume plants), 수수(sorghum), 사탕수수(sugar cane), 사탕무 펄프(sugar beet pulp), 우드칩(wood chips), 톱밥(sawdust) 및 바이오매스 크롭(biomass crops)[예를 들어, *크람베*(*Crambe*)]를 포함한다.
- [0166] 적절한 탄소원에 더하여, 배지는 본 분야에서 숙련자에게 공지되고, 배양물의 성장 및 다양한 당의 발효에 필수적인 효소적인 경로(enzymatic pathways)의 프로모션(promotion), 및 탄화수소 및 탄화수소 유도체의 생산에 적절한, 무기물, 염, 보조인자, 완충용액 및 그 밖의 성분을 포함해야한다. 반응물은, 호기성(aerobic), 산소 결핍(anoxic) 또는 혐기성의 조건(anaerobic conditions)은 미생물의 요구를 기반으로 한 바람직한, 호기성 또는 혐기성 조건 하에서 반응이 실시될 수 있다. 숙주 세포가 성장하고/성장하거나 증가됨에 따라, 다양한 당 또는 바이오매스 중합체에서의 성장, 당 발효, 또는 탄화수소 또는 탄화수소 유도체의 합성에 필요한 효소, 운반체 또는 그 밖의 단백질의 발현은 영향을 받는다.
- [0167] **세포 내로 당의 수송을 증가시키는 방법**
- [0168] 본 발명은 세포 내로 당의 수송을 증가시키는 방법을 제공한다. 하나의 양상에서, 본 발명은, 숙주 세포를 준

비하는 첫 번째 단계를 포함하는 세포 내로 셀로텍스트린을 수송하는 방법을 제공하고, 상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 하기의 하나 또는 그 이상은 사실이다 : 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO: 1 을 포함하고, 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO : 2 를 포함하고, 트랜스멤브레인 α -헬릭스 2 및 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프는 SEQ ID NO : 3 을 포함하고, 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO : 4 를 포함하고, 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO : 5 를 포함하고, 트랜스멤브레인 α -헬릭스 6 및 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO : 6 을 포함하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO : 7 을 포함하고, 트랜스멤브레인 α -헬릭스 10 및 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO : 8 을 포함한다. 본 방법은 재조합 폴리뉴클레오티드가 발현되는 것과 같은 세포를 배양하는 두 번째 단계를 포함하고, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 셀로텍스트린의 수송이 증가되었다. 세포 내로의 셀로텍스트린의 수송은 셀로비오스가 단일 탄소원(sole carbon source)인 경우, *S. 세레비시아* 숙주 세포의 능력을 측정하거나 또는 세포 내로 [3 H]-셀로비오스의 흡수를 측정하는 것과 같은 실시예 9 에서 기재된 이러한 방법을 포함하는, 본 분야에서 숙련자에게 알려진 모든 방법을 측정할 수 있다. 일반적으로, 재조합 폴리뉴클레오티드를 포함하는 숙주 세포 및 재조합 폴리뉴클레오티드를 포함하지 않는 숙주 세포는, 그외의 유전적인 백그라운드(genetic background)에서 동일하다.

[0169] 특정 실시형태에서, 폴리펩티드는 NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 100 % 의 의 아미노산 동일성을 갖는다. 특정 실시형태에서, 숙주 세포는 β -글루코시다아제의 촉매 도메인을 코딩하는 재조합 폴리뉴클레오티드를 또한 포함한다. 이러한 실시형태는 셀로텍스트린을 사용하기 위한 내인성 능력(endogenous ability)이 결핍된 숙주 세포에 유용하다. 바람직하게, β -글루코시다아제의 촉매 도메인은 세포 내이다. 바람직한 실시형태에서, β -글루코시다아제는 *내우로스포라 크라사*로부터 온 것이다. 특히 바람직한 실시형태에서, β -글루코시다아제는 NCU00130 에 의해 코드된다.

[0170] 세포 내로 셀로텍스트린의 수송을 증가시키는 방법에 있어서, 상기 세포는 변경된 유기체(altered organism)로부터 셀룰라아제-함유 효소 혼합물(cellulase-containing enzyme mixture)을 포함하는 배지에서 배양시킬 수 있고, 상기 혼합물은 변경된 유기체로부터 셀룰라아제-함유 혼합물과 비교하여 감소된 β -글루코시다아제 활성도를 갖는다. 유기체는, β -글루코시다아제를 코딩하는 유전자의 돌연변이 또는 목표된 RNA 간섭(targeted RNA interference) 또는 이와 같은 것과 같은 β -글루코시다아제의 발현을 감소시키기 위해 변경시킬 수 있다.

[0171] 또 다른 양상에서, 본 발명은, 하기의 단계를 포함하는 세포 내로 자일로스의 수송을 증가시키는 방법을 제공하고, NCU00821 또는 STL12/XUT6 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 준비하는 단계, 및 재조합 폴리뉴클레오티드가 발현된 것과 같은 세포를 배양하는 단계를 포함하되, 상기 재조합 폴리뉴클레오티드의 발현으로, 재조합 폴리뉴클레오티드가 포함되지 않는 세포와 비교하여 세포 내로 자일로스의 수송을 증가시킨다. 또 다른 양상에서, 본 발명은 하기의 단계를 포함하는, 세포 내로 아라비노스의 수송을 증가시키는 방법을 제공하고, XUT1 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 단계, 및 재조합 폴리뉴클레오티드가 발현되는 세포를 배양하는 단계를 포함하되, 상기 재조합 폴리뉴클레오티드의 발현으로, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여 세포 내로 아라비노스의 수송을 증가시킨다. 또 다른 양상에서, 본 발명은 하기의 단계를 포함하는, 세포 내로 아라비노스 또는 글루코스의 수송을 증가시키는 방법을 제공하고, NCU06138 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 준비하는 단계, 및 재조합 폴리뉴클레오티드가 발현되는 세포를 배양하는 단계를 포함하되, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 아라비노스 또는 글루코스의 수송을 증가시킨다. 또 다른 양상에서, 본 발명은 다음 단계를 포함하는, 세포 내로 자일로스 또는 글루코스의 수송을 증가시키는 방법을 제공하고, UT2, SUT3 또는 XUT3 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 준비하는 단계, 및 재조합 폴리뉴클레오티드가 발현되는 세포를 배양

하는 단계를 포함하되, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 자일로스 또는 글루코스의 수송을 증가시킨다. 또 다른 양상에서, 본 발명은 다음 단계를 포함하는, 세포 내로 자일로스, 아라비노스 또는 글루코스의 수송을 증가시키는 방법을 제공하고, NCU04963 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 준비하는 단계, 및 재조합 폴리뉴클레오티드가 발현되는 세포를 배양하는 단계를 포함하되, 재조합 폴리뉴클레오티드를 포함하지 않는 세포와 비교하여, 상기 재조합 폴리뉴클레오티드의 발현으로 세포 내로 자일로스, 아라비노스 또는 글루코스의 수송을 증가시킨다.

[0172] 세포 내로의 자일로스, 아라비노스 또는 글루코스의 수송은 실시예 11 에 기재된 방법을 포함하는 본 분야의 숙련자에게 공지된 모든 방법으로 측정될 수 있다. 이러한 방법은, 예를 들어 삼투압(osmosis) 및 고성능 액체 크로마토그래피(high performance liquid chromatography)를 사용한 분석으로 숙주 세포로부터 측정된 D-자일로스 및 자일리톨 또는 L-아라비노스 및 아라비니톨의 추출에 의한 D-자일로스 또는 L-아라비노스를 측정하고, 단일 탄소원으로서 글루코스에서 성장할 수 있는 능력이 결핍된(lacking) 숙주 세포를 사용한 글루코스 수송을 측정하는 것을 포함한다. 일반적으로, 재조합 폴리뉴클레오티드를 포함하는 숙주 세포 및 재조합 폴리뉴클레오티드를 포함하지 않는 숙주 세포는, 그외의 유전적인 백그라운드(genetic background)에서 동일하다.

[0173] 특정 실시형태에서, 숙주 세포는, 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 또한 포함하되, 상기 하나 또는 그 이상의 재조합 폴리뉴클레오티드는 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코딩한다. 하나 또는 그 이상은, 예를 들어 L-아라비노스 이성질화효소, L-리블로인산화효소, L-리블로스-5-P 4 에피머화효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비니톨 4-탈수소효소, L-자일루로스 환원효소 및 자일리톨 탈수소효소 또는 본 분야에서 공지된 그 밖의 오타당 이용 효소일 수 있다.

[0174] 세포의 성장을 증가시키는 방법

[0175] 본 발명은 세포의 성장을 증가시키는 방법을 제공한다. 하나의 양상에서, 본 발명은, 숙주 세포를 준비하는 첫 번째 단계를 포함하는, 세포의 성장을 증가시키는 방법을 제공하고, 상기 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11, α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 하기의 하나 또는 그 이상은 사실이다 : 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO : 1 을 포함하고, 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO : 2 를 포함하고, 트랜스멤브레인 α -헬릭스 2 및 트랜스멤브레인 α -헬릭스 3 을 연결하는 루프는 SEQ ID NO : 3 을 포함하고, 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO : 4 를 포함하고, 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO : 5 를 포함하고, 트랜스멤브레인 α -헬릭스 6 및 트랜스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO : 6 을 포함하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO : 7 을 포함하고, 트랜스멤브레인 α -헬릭스 10 및 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO : 8 을 포함하고, 폴리펩티드는 셀로텍스트린을 수송한다. 본 방법은 셀로텍스트린을 포함하는 배지에서의 숙주 세포를 배양하는 두 번째 단계를 포함하고, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 보다 빠른 비율로 성장한다. 숙주 세포의 성장 비율은 본 분야에서의 숙련자에게 공지된 방법에 의해 측정될 수 있다. 일반적으로, 세포의 성장 비율은 광학 밀도(optical density)에 의한 현탁액(suspension)에서의 세포의 농도를 평가하여서 측정할 수 있다. 바람직하게, 재조합 폴리뉴클레오티드를 포함하는 숙주 세포 및 재조합 폴리뉴클레오티드를 포함하지 않는 숙주 세포는, 그외의 유전적인 백그라운드(genetic background)에서 동일하다. 셀로텍스트린을 포함하는 배지는 셀룰로스와 같은 바이오매스 중합체의 효소적인 처리로 인한다.

[0176] 특정 실시형태에서, 폴리펩티드는 NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 %, 적어도 100 % 의 아미노산 동일성을 갖는다. 특정 실시형태에서, 숙주 세포는 β -글루코시다아제의 촉매 도메인을 코딩하는 재조합 폴리뉴클레오티드를 또한 포함한다. 이러한 실시형태는 셀로텍스트린을 이용하기 위해 내인성 능력(endogenous ability)이 결핍된

(lacking) 숙주 세포에 유용하다. 바람직하게, β -글루코시다아제의 촉매 도메인은 세포내이다. 바람직한 실시형태에서, β -글루코시다아제는 *네투로스포라 크라사*로부터 온 것이다. 특히 바람직한 실시형태에서, β -글루코시다아제는 NCU00130 에 의해 코드된다.

[0177] 세포의 성장을 증가시키는 방법에 있어서, 배양하는 배지는 변경된 유기체로부터 셀룰라아제-함유 효소 혼합물을 포함하고, 상기 혼합물을 변경되지 않는 유기체로부터의 셀룰라아제-함유 혼합물과 비교하여 감소된 β -글루코시다아제 활성도를 갖는다. 유기체는 목표하는 RNA 간섭(targeted RNA interference) 또는 이와 같은 것에 의하거나 또는 β -글루코시다아제를 코딩하는 유전자의 돌연변이와 같은, β -글루코시다아제의 발현을 감소시키기 위해 변경시켰다.

[0178] 또 다른 양상에서, 본 발명은 하기의 단계를 포함하는, 세포의 성장을 증가시키는 방법을 제공하고, 자일로스를 수송하는 NCU00821 또는 STL12/XUT6 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 단계, 및 자일로스를 포함하는 배지에서 숙주 세포를 배양하는 단계를 포함하되, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장한다. 또 다른 양상에서, 본 발명은 하기의 단계를 포함하는, 세포의 성장을 증가시키는 방법을 제공하고, 아라비노스를 수송하는 XUT1 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 단계, 및 아라비노스를 포함하는 배지에서 숙주 세포를 배양하는 단계를 포함하되, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장한다. 또 다른 양상에서, 본 발명은 하기의 단계를 포함하는, 세포의 성장을 증가시키는 방법을 제공하고, 아라비노스 및 글루코스를 수송하는 NCU06138 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 단계, 및 아라비노스 또는 글루코스를 포함하는 배지에서 숙주 세포를 배양하는 단계를 포함하되, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장한다. 또 다른 양상에서, 본 발명은 하기의 단계를 포함하는, 세포의 성장을 증가시키는 방법을 제공하고, 자일로스 및 글루코스를 수송하는 SUT2, SUT3 또는 XUT3 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 단계, 및 자일로스 또는 글루코스를 포함하는 배지에서 숙주 세포를 배양하는 단계를 포함하되, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장한다. 또 다른 양상에서, 본 발명은 하기의 단계를 포함하는, 세포의 성장을 증가시키는 방법을 제공하고, 자일로스, 아라비노스 및 글루코스를 수송하는 NCU04963 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 단계, 및 자일로스, 아라비노스 또는 글루코스를 포함하는 배지에서 숙주 세포를 배양하는 단계를 포함하되, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장한다.

[0179] 숙주 세포의 성장 비율은 본 분야의 숙련자에서 공지된 모든 방법으로 측정될 수 있다. 일반적으로, 세포의 성장 비율은 공학 밀도에 의한 현탁액에서의 세포 농도를 평가하여 측정할 수 있다. 바람직하게, 재조합 폴리뉴클레오티드를 포함하는 숙주 세포 및 재조합 폴리뉴클레오티드를 포함하지 않는 숙주 세포는, 그외의 유전적인 백그라운드(genetic background)에서 동일하다. 자일로스 또는 아라비노스를 포함하는 배지는 헤미셀룰로스나 같은 바이오매스 중합체의 산 처리에로부터 기인한다. 글루코스를 포함하는 배지는 셀룰로스와 같은 바이오매스 중합체의 효소적 처리(enzymatic treatment)로부터 기인한다.

[0180] 특정 실시형태에서, 숙주 세포는 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 또한 포함하되, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코드한다. 상기 하나 또는 그 이상의 효소는, 예를 들어 L-아라비노스 이성질화효소, L-리블로인산화효소, L-리블로스-5-P 4 에피머화효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비니톨 4-탈수소효소, L-자일루로스 환원효소 및 자일리톨 탈수소효소 또는 본 분야에서 공지된 그 밖의 오타당 이용 효소일 수 있다.

[0181] 하나의 양상에서, 본 발명은 바이오매스 중합체에서 세포의 성장을 증가시키는 방법을 제공한다. 바람직한 실시형태에서, 바이오매스 중합체는 셀룰로스이다. 또 다른 바람직한 실시형태에서, 바이오매스 중합체는 헤미셀룰로스이다. 본 발명의 하나의 양상에 따라, 본 발명은 NCU07705 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를

티드를 포함하는 숙주 세포를 준비하는 것을 포함한다. 본 발명의 또 다른 양상에 따라, 본 방법은 바이오매스 중합체를 포함하는 배지에서 세포를 배양하는 것을 포함하고, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포 보다 배지에서 더 빠른 비율로 성장하였다.

[0182] 본 발명의 또 다른 양상에서, 본 발명은, 하기의 단계를 포함하는, 세포의 성장을 증가시키는 방법을 제공하고, NCU01517, NCU09133 또는 NCU10040 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 단계, 및 헤미셀룰로스를 포함하는 배지에서 세포를 배양하는 단계를 포함하되, 상기 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포보다, 배지에서 더 빠른 비율로 성장하였다.

[0183] 본 발명의 또 다른 양상에 따라, 본 방법은 NCU05137 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 것을 포함한다. 본 발명의 또 다른 양상에 따라, 본 발명은 내인성 폴리뉴클레오티드의 발현을 저해하고, 바이오매스 중합체를 포함하는 배지에서 세포를 배양하는 것을 포함하되, 상기 숙주 세포는, 내인성 폴리뉴클레오티드의 발현이 저해되지 않는, 세포보다 배지에서 더 빠른 비율로 성장하였다.

[0184] 본 발명의 방법은 세포에서 재조합 핵산이 발현되는 숙주 세포를 배양하는 것을 포함한다. 미생물 숙주로부터, 상기 과정은 적절한 배지에서 세포를 배양하는 것을 수분한다. 일반적인 세포는 적절한 배지에서 35 °C 에서 성장시켰다. 본 발명에서 바람직한 성장 배지는, 예를 들어 Luria Bertani (LB) 브로스(broth), Sabouraud Dextrose (SD) 브로스 또는 효소 배지(YM) 브로스와 같은 일반적으로 상업적으로 제조된 배지를 포함한다. 그 밖의 특정되거나 합성된 성장 배지를 또한 사용할 수 있고, 특정한 숙주 세포의 성장에 적절한 배지는, 미생물학 또는 발효과학의 분야에서 숙련자에게 널리 공지되어 있다. 성장에 적절한 온도 범위 및 그 밖의 조건은 본 분야에서 공지되어 있다(예를 들어, Bailey and Ollis 1986 참조)

[0185] 배지에서의 중합체의 근원은, 예를 들어 풀(grasses)[예를 들어, 지팡이풀(switchgrass), *미스칸투스*(*Miscanthus*)], 왕겨(rice hull), 버개스(bagasse), 무명(cotton), 주트(jute), 대마(hemp), 아마(flax), 대나무(bamboo), 사이잘(sisal), 마닐라삼섬유(abaca), 짚(straw), 잎(leaves), 잔디 부스러기(grass clippings), 옥수수대(corn stover), 옥수수 속대(corn cobs), 증류립(distillers grains), 콩과 식물(legume plants), 수수(sorghum), 사탕수수(sugar cane), 사탕무 펄프(sugar beet pulp), 우드칩(wood chips), 톱밥(sawdust) 및 바이오매스 크롭(biomass crops)[예를 들어, *크람베*(*Crambe*)]를 포함한다. 바이오매스 중합체에 추가적으로, 배지는, 배양물의 성장에 적절한, 무기질, 염, 공동인자(cofactors), 완충용액 및 본 분야의 숙련자에게 널리 알려진 그 밖의 성분을 포함해야한다. 숙주 세포의 성장 비율은 본 분야에서의 숙련자에서 공지된 모든 방법으로 측정할 수 있다.

[0186] 본 발명의 특정 실시형태에서, 셀룰라아제의 발현은 재조합 폴리뉴클레오티드의 발현 하에서 숙주 세포에서 증가되었다. 본원에서 사용된 "셀룰라아제"는, 셀룰로스 중합체를 보다 더 짧은 셀로-올리고당 올리고머(shorter cello-oligosaccharide oligomers), 셀로비오스 및/또는 당으로 가수분해할 수 있는 효소의 카테고리(category)를 나타낸 것이다. 셀룰라아제는 엑소글루카나아제(exoglucanases), 엑소셀로비오히드롤라아제(exocellobiohydrolases), 엔도글루카나아제(endoglucanases) 및 글루코시다아제를 포함하지만 이로 한정되지 않는다. 셀룰라아제의 발현은 본 분야에서 공지된 RP-PCR 또는 그 밖의 방법으로 측정될 수 있다.

[0187] 본 발명의 특정 실시형태에서, 헤미셀룰라아제의 발현은 재조합 폴리뉴클레오티드의 발현 하에서 숙주 세포에서 증가되었다. 본원에 기재된 "헤미셀룰라아제"는 헤미셀룰로스 중합체를 가수분해하는 효소의 카테고리를 나타낸 것이다. 헤미셀룰라아제는, 자일라나아제(xylanases), 만나나아제(mannanases), 아라비나아제(arabinases)[둘 다 엔도 및 엑소 종류(endo and exo kinds)] 및 이들과 상응하는 글리코시다아제를 포함하지만 이로 제한되지 않는다. 헤미셀룰라아제의 발현은 RT-PCR 또는 본 분야에서 공지된 그 밖의 방법으로 측정할 수 있다.

[0188] 내인성 폴리뉴클레오티드의 발현의 저해는, 예를 들어 유전자의 기능 또는 유전자의 산물(예를 들어, 유전자에 의해 코드되는 단백질)의 기능에서의 감소로 인한 유전적 변형에 의해 성취될 수 있고, 이는 유전자의 발현의 불활성화(완전한 또는 부분적인), 결실, 중단(interruption), 장애(blockage), 침묵(silencing), 하향-조절(down-regulation) 또는 감쇠(attenuation)로서 나타낸 것일 수 있다. 예를 들어, 이러한 유전자에 의해 코드되는 단백질의 기능에서의 감소를 야기하는 유전자에서의 유전적인 변형은, 유전자의 완전한 결실(즉, 유전자가 존재하지 않고, 따라서 단백질이 존재하지 않는다), 단백질의 불완전하거나 또는 번역이 없는 것을 야기하는 유전자에서의 돌연변이(예를 들어, 단백질이 발현되지 않는다), 또는 단백질의 천연적인 기능을 감소시키거나 폐지하는(abolish) 유전자에서의 돌연변이(예를 들어, 효소적인 활성도 또는 작용이 없거나 감소된 단백질이 발현된다)의 결과일 수 있다. 보다 명확하게, 본원에 기재된 단백질의 작용을 감소시키는 언급은, 숙주 세포에서 모든 유전적인 변형에 대해서 일반적으로 나타내었고, 단백질의 감소된 발현 및/또는 기능성(생물학적 활성도)을 야기하고, 단백질의 감소된 활성도(예를 들어, 감소된 수송), 단백질의 증가된 저해 또는 분해 뿐만 아니라 단백질의 발현의 감소 또는 저해를 포함하는 숙주 세포에서의 어떠한 유전적인 변형을 나타낸 것이다. 예를 들어, 본 발명의 단백질의 작용 또는 활성도는 단백질 생산의 저지(blocking) 또는 감소, 단백질 작용의 감소, 또는 단백질 작용의 저해에 의해 감소될 수 있다. 이러한 변형의 몇몇을 또한 조합할 수 있다. 단백질 생산의 저지(Blocking) 또는 감소는, 성장 배지에서 유도된 화합물의 존재를 필요로 하는 프로모터의 조절 하에서 단백질을 코딩하는 유전자를 배치하는 것을 포함할 수 있다. 유도물질(inducer)이 배지로부터 대폭 감소되는 것과 같은 조건을 설정함으로써, 단백질을 코딩하는 유전자의 발현(및 이에 따른 단백질의 합성)은 정지될 수 있다(turned off). 단백질의 작용을 저지하거나 감소시키는 것은 본원에서 참조문헌으로 인용된 U.S. Pat. No. 4,743,546 에 기재된 것과 유사한 제거 기술 접근법(excision technology approach)을 사용하는 것을 또한 포함할 수 있다. 이러한 접근법을 사용하기 위해, 흥미로운 단백질을 코딩하는 유전자는, 계능으로부터 유전자가 특이적이고(specific), 조절된 제거(controlled excision)를 가능하게 하는 특이적인 유전적인 서열(specific genetic sequences) 사이에 클론되었다(cloned). 제거(Excision)는, 예를 들어 U.S. Pat. No. 4,743,546 에 기재된, 배양물의 배양 온도에서의 시프트(shift), 또는 몇몇 다른 물리적이거나(physical) 또는 영양상의 신호(nutritional signal)에 의해 유도된다(prompted).

[0189] 본 발명의 특정 실시형태에서, 숙주 세포의 셀룰라아제 활성도는 내인성 폴리뉴클레오티드의 발현의 저해 상에서 증가되었다. 셀룰라아제 활성도는 실시예 5 에 기재된 바와 같은 방법 및 본 분야에서 알려진 그 밖의 방법으로 측정될 수 있다.

[0190] 본 발명의 특정 실시형태에서, 숙주 세포의 헤미셀룰라아제 활성도는 내인성 폴리뉴클레오티드의 발현의 저해 상에서 증가되었다. 헤미셀룰라아제 활성도는 실시예 17 에 기재된 바와 같은 방법 및 본 분야에서 알려진 그 밖의 방법으로 측정될 수 있다.

[0191] 공동-발효의 방법

[0192] 본 발명의 하나의 양상은 셀룰로스-유도(derived) 및 헤미셀룰로스-유도 당을 공동-발효시키는(co-fermenting) 방법을 제공한다. 본원에서 사용된 바와 같이, 공동-발효(co-fermentation)는 동일한 베셀(vessel)에서 하나 이상의 당이 숙주 세포에 의해 동시에 이용되는 것을 나타낸 것이다. 본 발명은, 셀로텍스트린 운반체를 코딩하는 첫 번째 재조합 폴리뉴클레오티드 및 β -글루코시다아제의 촉매 도메인을 코딩하는 두 번째 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 단계, 및 셀룰로스-유도 당 및 헤미셀룰로스-유도 당을 포함하는 배지에서 숙주 세포를 배양하는 단계를 포함하되, 상기 재조합 폴리뉴클레오티드의 발현은 셀룰로스-유도 당 및 헤미셀룰로스-유도 당을 공동-발효시킬 수 있다.

[0193] 첫 번째 재조합 폴리뉴클레오티드는 세포 내로 셀로텍스트린을 수송할 수 있는 어떠한 폴리펩티드를 코드할 수 있다. 셀로텍스트린 수송은, 실시예 9 에 기재된 방법을 포함하는, 본 분야의 숙련자에게 공지된 방법으로 측정될 수 있다. 바람직한 실시형태에서, 첫 번째 재조합 폴리뉴클레오티드는 트랜스멤브레인 α -헬릭스 1, α -

헬릭스 2, α-헬릭스 3, α-헬릭스 4, α-헬릭스 5, α-헬릭스 6, α-헬릭스 7, α-헬릭스 8, α-헬릭스 9, α-헬릭스 10, α-헬릭스 11, α-헬릭스 12 를 포함하는 폴리펩티드를 코드하고, 하기의 하나 또는 그 이상은 사실이다 : 트랜스멤브레인 α-헬릭스 1 은 SEQ ID NO: 1 을 포함하고, 트랜스멤브레인 α-헬릭스 2 는 SEQ ID NO: 2 를 포함하고, 트랜스멤브레인 α-헬릭스 2 와 트랜스멤브레인 α-헬릭스 3 를 연결하는 루프는 SEQ ID NO: 3 을 포함하고, 트랜스멤브레인 α-헬릭스 5 는 SEQ ID NO: 4 를 포함하고, 트랜스멤브레인 α-헬릭스 6 은 SEQ ID NO: 5 를 포함하고, 트랜스멤브레인 α-헬릭스 6 과 트랜스멤브레인 α-헬릭스 7 사이의 서열은 SEQ ID NO: 6 을 포함하고, 트랜스멤브레인 α-헬릭스 7 은 SEQ ID NO: 7 를 포함하고, 및 트랜스멤브레인 α-헬릭스 10 및 트랜스멤브레인 α-헬릭스 11 및 이들 사이의 서열은 SEQ ID NO: 8 을 포함한다. 이러한 폴리펩티드의 예는 NCU00801, NCU00809, NCU08114, XP_001268541.1 및 LAC2 를 포함한다. 바람직한 실시형태에서, 첫 번째 재조합 폴리펩티드는 NCU00801 를 코드한다.

[0194] 두 번째 재조합 폴리뉴클레오티드는, 글루코스를 방출하면서 β-D-글루코시드에서 말단의 비-환원된 잔기(terminal non-reducing residues)를 가수분해를 촉매하는 능력을 갖는 어떠한 촉매 도메인(any catalytic domain)을 코드할 수 있다. 바람직하게, β-글루코시다아제 촉매 도메인은 숙주 세포에서 세포 내에 위치한다. 특정 실시형태에서, β-글루코시다아제의 근원은 *N. 크라사* β-글루코시다아제이다. 바람직한 실시형태에서, β-글루코시다아제 도메인의 근원은 NCU00130 이다. 상이한 근원으로부터의 촉매 도메인이 상이한 셀로렉스트린 운반체를 갖는 것이 최상으로 작업할 수 있다.

[0195] 특정 실시형태에서, 숙주 세포는 하나 또는 그 이상의 재조합 폴리뉴클레오티드를 또한 포함하고, 상기 하나 또는 그 이상의 폴리뉴클레오티드는 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코드한다. 선택적으로, 오타당 이용에 포함되는 하나 또는 그 이상의 효소를 코딩하는 하나 또는 그 이상의 폴리뉴클레오티드는 숙주 세포에 대해서 내인성(endogenous)일 수 있다. 하나 또는 그 이상의 효소는, 예를 들어 L-아라비노스 이성질화 효소, L-리블로인산화효소, L-리블로스-5-P 4 에피머화효소, 자일로스 이성질화효소, 자일루로키나아제, 알도스 환원효소, L-아라비니톨 4-탈수소효소, L-자일루로스 환원효소 및 자일리톨 탈수소효소 및 본 분야의 숙련자에게 공지된 그 밖의 오타당-이용 효소를 포함한다.

[0196] 특정 실시형태에서, 숙주 세포는 세 번째 재조합 폴리뉴클레오티드를 포함하고, 상기 폴리뉴클레오티드는 오타당 운반체를 코드한다. 선택적으로, 숙주 세포는 오타당 운반체를 코딩하는 내인성 폴리뉴클레오티드를 포함할 수 있다. 특정 실시형태에서, 오타당 운반체는 세포 내로 자일로스 및/또는 아라비노스를 수송한다. 특정 실시형태에서, 세 번째 재조합 폴리뉴클레오티드는 NCU00821, NCU04963, NCU06138, STL12/XUT6, SUT2, SUT3, XUT1 또는 XUT3 과 같은 폴리펩티드를 코드한다. 숙주 세포에서 오타당 운반체의 발현은, 만약 오타당 당 중에 존재하는 글루코스가 성장 배지라면, 공동-발효의 효율성은 증진될 것이다.

[0197] 본원에 기재된 바와 같은 공동-발효의 방법에서, 셀룰로스-유도 당은 바람직하게 셀로비오스, 셀로트리오스 및 셀로테트라오스를 포함하고, 헤미셀룰로스-유도 당은 바람직하게 자일로스 및 아라비노스를 포함한다. 일반적으로, 숙주 세포에 의한 공동-발효를 위한 셀룰로스-유도 당 및 헤미셀룰로스-유도 당을 제조하기 위해, 리그노셀룰로스가 바이오매스(lignocellulosic biomass)는, 이의 구조를 변경시키기 위해 첫 번째로 예비처리하였고(pretreated), 셀룰로스의 보다 나은 효소적 가수분해(better enzymatic hydrolysis)를 가능하게 한다. 예비처리는, 예를 들어 암모니아 파이버/냉동 폭발(ammonia fiber/freeze explosion), 칼슘 또는 수산화나트륨을 기반으로 한 라임방법(lime method), 및 산 촉매와 함께 또는 산 촉매 없는 증기 폭발(steam explosion)을 포함하는, 물리적이고 화학적인 방법을 포함할 수 있다. 산 처리는 리그노셀룰로스가 바이오매스의 헤미셀룰로스 성분으로부터의 자일로스 및 아라비노스를 방출할 것이다. 그 후에, 바람직하게, 예비처리된 바이오매스의 셀룰로스 성분은 셀룰라아제의 혼합물에 의해 가수분해되었다. 상업적으로 입수가 가능한 셀룰라아제 혼합물의 예는 Celluclast 1.5L®(Novozymes), Spezyme CP®(Genencor)(Scott W. Pryor, 2010, Appl Biochem Biotechnol) 및 Cellulyve 50L(Lyven)를 포함한다.

- [0198] 셀룰라아제 혼합물은 엔도글루카나아제, 엑소글루카나아제 및 β -글루코시다아제를 일반적으로 포함한다. 본원에 기재된 바와 같은 공동-발효의 방법에 있어서, 셀룰라아제 혼합물에서 β -글루코시다아제 활성도의 양은 가능한 많이 최소화되어야 한다. 예를 들어, 배양된 배지는 변경된 유기체로부터의 셀룰라아제-함유 효소를 포함할 수 있고, 상기 혼합물은 변경되지 않는 유기체로부터 셀룰라아제-함유 혼합물과 비교하여, 감소된 β -글루코시다아제 활성도를 갖는다. 유기체는, β -글루코시다아제를 코딩하는 유전자의 돌연변이 또는 목표하는 RNA 간섭(targeted RNA interference) 또는 이와 유사한 것과 같은 것에 의해 β -글루코시다아제의 발현을 감소시키도록 변경할 수 있다.
- [0199] 놀랍게도 실시예 17 에 기재된 바와 같이, 본 발명의 방법에 의한 셀로비오스 및 자일로스의 공동-발효는 숙주 세포에 의한 당 소비 및 에탄올 생산에 상승효과(synergistic effect)를 나타내었다.
- [0200] **탄화수소 또는 탄화수소 유도체의 합성 방법**
- [0201] 본 발명의 하나의 양상은 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 제공한다.
- [0202] 본원에서 사용된 바와 같은 "탄화수소"는 전부 수소 및 탄소로 이루어진 유기 화합물이다. 탄화수소는, 메탄, 에탄, 에텐, 에틴(ethyne), 프로판, 프로펜, 프로핀(propyne), 시클로프로판, 알렌(allene), 부탄, 이소부텐(isobutene), 부텐, 부틴, 시클로부탄, 메틸시클로프로판, 부타디엔, 펜탄, 이소펜탄, 네오펜탄, 펜텐, 펜틴(pentyne), 시클로펜탄, 메틸시클로부탄, 에틸시클로프로판, 펜타디엔, 이소프레네, 헥산, 헥센, 헥신(hexyne), 시클로헥산, 메틸시클로펜탄, 에틸시클로부탄, 프로필시클로프로판, 헥사디엔, 헵테인(heptane), 헵텐(heptene), 헵틴(heptyne), 시클로헵테인, 메틸시클로헥산, 헵타디엔, 옥탄, 옥텐(octene), 옥틴(octyne), 시클로옥탄(cyclooctane), 옥타디엔(octadiene), 노난(nonane), 노넨(nonene), 노니엔(nonyne), 시클로노난(cyclononane), 노난디엔(nonadiene), 데칸, 데켄(decene), 데킨(decyne), 시클로데칸 및 데카디엔(decadiene)을 포함하지만 이로 제한되지 않는다.
- [0203] 본원에 사용된 바와 같은, "탄화수소 유도체"는 수소가 아닌 적어도 하나의 그 밖의 성분 및 탄소의 유기 화합물이다. 탄화수소 유도체는, 알코올(예를 들어, 아라비니톨, 부탄올, 에탄올, 글리세롤, 메탄올, 1,3-프로판디올, 소르비톨 및 자일리톨) ; 유기산(예를 들어, 아세트산, 아디프산, 아스코르브산, 시트르산, 2,5-디케토-D-글루콘산, 포름산, 푸마르산, 글루카르산, 글루콘산, 글루쿠론산(glucuronic acid), 글루타르산(glutaric acid), 3-히드록시프로피오닉산(hydroxypropionic acid), 이타콘산, 젖산, 말산, 말론산, 옥살산, 프로피온산, 숙신산 및 크실론산) ; 에스테르 ; 케톤(예를 들어, 아세톤) ; 알데히드(예를 들어, 푸르푸랄) ; 아미노산(예를 들어, 아스파르트산, 글루탐산, 글리신, 리신, 세린 및 트레오닌) ; 및 기체(gases)(예를 들어, 이산화탄소 및 일산화탄소)를 포함하지만, 이로 제한되지 않는다.
- [0204] 바람직한 실시형태에서, 탄화수소 또는 탄화수소 유도체는 연료로서 사용될 수 있다. 특히 바람직한 실시형태에서, 탄화수소 또는 탄화수소 유도체는 에탄올 또는 부탄올이다.
- [0205] 본 발명의 하나의 양상에 따라, 숙주 세포를 준비하는 첫 번째 단계를 포함하는 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함하고, 숙주 세포는 트랜스멤브레인 α -헬릭스 1, α -헬릭스 2, α -헬릭스 3, α -헬릭스 4, α -헬릭스 5, α -헬릭스 6, α -헬릭스 7, α -헬릭스 8, α -헬릭스 9, α -헬릭스 10, α -헬릭스 11 및 α -헬릭스 12 를 포함하는 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하고, 상기 하나 또는 그 이상의 하기는 사실이다 : 트랜스멤브레인 α -헬릭스 1 은 SEQ ID NO: 1 을 포함하고, 트랜스멤브레인 α -헬릭스 2 는 SEQ ID NO: 2 를 포함하고, 트랜스멤브레인 α -헬릭스 2 와 트랜스멤브레인 α -헬릭스 3 를 연결하는 루프는 SEQ ID NO: 3 을 포함하고, 트랜스멤브레인 α -헬릭스 5 는 SEQ ID NO: 4 를 포함하고, 트랜스멤브레인 α -헬릭스 6 은 SEQ ID NO: 5 를 포함하고, 트랜스멤브레인 α -헬릭스 6 과 트랜

스멤브레인 α -헬릭스 7 사이의 서열은 SEQ ID NO: 6 을 포함하고, 트랜스멤브레인 α -헬릭스 7 은 SEQ ID NO: 7 를 포함하고, 및 트랜스멤브레인 α -헬릭스 10 및 트랜스멤브레인 α -헬릭스 11 및 이들 사이의 서열은 SEQ ID NO: 8 을 포함하고, 상기 폴리펩티드는 탄화수소 또는 탄화수소 유도체의 합성을 위해 숙주 세포 내로 셀로텍스트린을 수송한다. 본 발명은 숙주 세포에 의해 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 셀로텍스트린 또는 셀로텍스트린의 근원을 포함하는 배지에서 숙주 세포를 배양하는 두 번째 단계를 포함하고, 세포 내로 셀로텍스트린의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다. 특정 실시형태에서, 폴리펩티드는, NCU00801 또는 NCU08114 와 적어도 29 %, 적어도 30 %, 적어도 35 %, 적어도 40 %, 적어도 45 %, 적어도 50 %, 적어도 55 %, 적어도 60 %, 적어도 65 %, 적어도 70 %, 적어도 75 %, 적어도 80 %, 적어도 85 %, 적어도 90 %, 적어도 95 %, 적어도 99 %, 적어도 100 % 의 아미노산 동일성을 갖는다. 특정 실시형태에서, 숙주 세포는 β -글루코시다아제의 촉매 도메인을 코딩하는 재조합 폴리뉴클레오티드를 또한 포함한다. 이러한 실시형태는 셀로텍스트린을 사용하기 위한 내인성 능력(endogenous ability)이 결핍된 숙주 세포에 대해서 유용하다. 바람직하게, β -글루코시다아제의 촉매 도메인은 세포 내이다. 바람직한 실시형태에서, β -글루코시다아제는 *네우로스포라 크라사*로부터 온 것이다. 특히 바람직한 실시형태에서, β -글루코시다아제는 NCU00130 에 의해 코드된다. 세포 내로 셀로텍스트린의 수송은 실시예 9 에 기재된 방법을 포함하는, 본 분야에서의 숙련자에게 공지된 어떠한 방법에 의해 측정될 수 있다. 일반적으로 셀로텍스트린의 근원은 셀룰로스이다.

[0206] 배양 배지는 변경된 유기체로부터 셀룰라아제-포함 효소 혼합물을 포함할 수 있고, 상기 혼합물은 변경되지 않는 유기체로부터 셀룰라아제-포함 혼합물과 비교하여 감소된 β -글루코시다아제 활성도를 갖는다. 유기체는 β -글루코시다아제를 코딩하는 유전자의 돌연변이 또는 목표하는 RNA 간섭 또는 이와 유사한 것과 같은, β -글루코시다아제의 발현을 저해하기 위해 변경될 수 있다.

[0207] 본 발명의 또 다른 양상에 따라, 하기의 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함하고, 탄화수소 또는 탄화수소 유도체의 합성을 위한 숙주 세포 내로 자일로스를 수송하는 NCU00821 또는 STL12/XUT6 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 숙주 세포를 제공하는 단계, 및 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 자일로스 또는 자일로스의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계를 포함하되, 상기 세포 내로 자일로스의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다.

[0208] 본 발명의 또 다른 양상에 따라, 하기의 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함하고, 탄화수소 또는 탄화수소 유도체의 합성을 위한 숙주 세포 내로 아라비노스를 수송하는 XUT1 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 숙주 세포를 제공하는 단계, 및 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 아라비노스 또는 아라비노스의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계를 포함하되, 상기 세포 내로 아라비노스의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다.

[0209] 본 발명의 또 다른 양상에 따라, 하기의 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함하고, 탄화수소 또는 탄화수소 유도체의 합성을 위한 숙주 세포 내로 아라비노스 또는 글루코스를 수송하는 NCU06138 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 숙주 세포를 제공하는 단계, 및 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 아라비노스 또는 글루코스 또는 아라비노스 또는 글루코스의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계를 포함하되, 상기 세포 내로 아라비노스 또는 글루코스의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다.

[0210] 본 발명의 또 다른 양상에 따라, 하기의 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함하고, 탄화수소 또는 탄화수소 유도체의 합성을 위한 숙주 세포 내로 자일로스 또는 글루코스를 수송하는 SUT2, SUT3 또는 XUT3 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 숙주 세포를 제공하는 단계, 및 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 자

일로스 또는 글루코스 또는 자일로스 또는 글루코스의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계를 포함하되, 상기 세포 내로 자일로스 또는 글루코스의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다.

[0211] 본 발명의 또 다른 양상에 따라, 하기의 단계를 포함하는, 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키는 방법을 포함하고, 탄화수소 또는 탄화수소 유도체의 합성을 위한 숙주 세포 내로 자일로스, 아라비노스 또는 글루코스를 수송하는 NCU04963 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는, 숙주 세포를 제공하는 단계, 및 숙주 세포에 의한 탄화수소 또는 탄화수소 유도체의 합성을 증가시키기 위해 자일로스, 아라비노스 또는 글루코스 또는 자일로스, 아라비노스 또는 글루코스의 근원을 포함하는 배지에서 숙주 세포를 배양하는 단계를 포함하되, 상기 세포 내로 자일로스, 아라비노스 또는 글루코스의 수송은 재조합 폴리뉴클레오티드의 발현 상에서 증가되었다.

[0212] 세포 내로 자일로스, 아라비노스 또는 글루코스의 수송은, 실시예 11 에 기재된 방법을 포함하는 본 분야의 숙련자에게 공지된 어떠한 방법으로 측정될 수 있다. 일반적으로 글루코스의 근원은 셀룰로스이고, 자일로스 및 아라비노스의 근원은 헤미셀룰로스이다.

[0213] 셀룰로스 분해(degrading) 방법

[0214] 본 발명의 하나의 양상은 셀룰로스 분해 방법을 제공한다. 본 방법은 셀룰로스를 포함하는 조성물을 준비하는 첫 번째 단계를 포함한다. 셀룰로스는 바람직하게, 지팡이풀(switchgrass), *미스칸투스(Miscanthus)*, 왕겨(rice hull), 버개스(bagasse), 아마(flax), 대나무(bamboo), 사이잘(sisal), 마닐라삼섬유(abaca), 짚(straw), 잎(leaves), 잔디 부스러기(grass clippings), 옥수수대(corn stover), 옥수수 속대(corn cobs), 증류립(distillers grains), 콩과 식물(legume plants), 수수(sorghum), 사탕수수(sugar cane), 사탕무 펄프(sugar beet pulp), 우드칩(wood chips), 톱밥(sawdust) 및 바이오매스 크롭(biomass crops)과 같은 식물물로부터 온 것이다.

[0215] 본 방법은 변경된 유기체로부터 셀룰라아제-함유 효소 혼합물로부터 조성물을 접촉하는 두 번째 단계를 포함하되, 상기 셀룰라아제-함유 효소 혼합물은, 변경되지 않는 유기체로부터 셀룰라아제-함유 혼합물과 비교하여 감소된 β -글루코시다아제 활성도를 갖는다. 셀룰로스는 셀룰라아제-함유 혼합물에 의해 분해된다. 유기체는 β -글루코시다아제를 코딩하는 유전자 또는 RNA 간섭(RNA interference)과 같은 기술로 β -글루코시다아제의 발현을 감소시키는 유전자의 돌연변이에 의해 변경될 수 있다. 유기체는 균류 또는 세균(bacterium)일 수 있다. 바람직한 실시형태에서, 유기체는 *T. 레세이*와 같은 곰팡이(filamentous fungus)이다.

[0216] 선택적으로, 방법은 이의 β -글루코시다아제 활성도를 감소시키기 위해 변경된, 셀룰라아제-함유 효소 혼합물을 갖는 조성물과 접촉하는 두 번째 단계를 포함한다. 예를 들어, 셀룰라아제-함유 효소 혼합물은 친화성 크로마토그래피로 변경될 수 있고, 상기 β -글루코시다아제 효소는 크로마토그래피 동안에 잡힐 수 있고(captured), 이에 의해 혼합물로부터 제거되었다. 또 다른 예에서, 셀룰라아제-함유 효소 혼합물은 저해제(inhibitor)와 함께 혼합물에서 β -글루코시다아제 효소의 불활성화에 의해 변경되었다. 상업적으로 입수가능한 셀룰라아제 혼합물의 예는 Celluclast 1.5L[®](Novozymes), Spezyme CP[®](Genencor)(Scott W. Pryor, 2010, Appl Biochem Biotechnol) 및 Cellulite 50L(Lyven)이다.

[0217] 본 발명은 이의 바람직한 특정한 실시형태와 함께 기재된 반면에, 앞서 말한 상세한 설명은 설명하려는 의도일 뿐, 이로 본 발명의 범위를 제한하려는 의도는 아님을 이해할 수 있을 것이다. 본 발명의 범위 내의 또 다른 양상, 이점(advantages) 및 변형은, 본 발명에 적용됨을 본 분야의 숙련자는 명백히 알 수 있을 것이다.

[0218] 기재된 본 발명 및 하기의 실시예는 설명을 통해 본 발명의 주제를 설명하기 위해 제시하였을 뿐, 이로 제한하려는 의도는 아니다.

실시예 1

[0219] **미스칸투스(*Miscanthus*) 및 아비셀(Avicel)에서 성장한 엔. 크라사(*N. crassa*)의 전사체 해석(transcriptome analysis)**

[0220] 본 실시예에서, *N. 크라사* 게놈(genome)의 발현 프로파일(expression profile)은 엽세 또는 아비셀에서 성장하는 동안 조사하였다. 유일 탄소원(sole carbon source)은 *T. 레세이*(*T. reesei*)(QM9414)(도 3)과 유사한 것으로서, 엔. 크라사(FGSC 2489)의 성장 및 셀룰라아제 활성도(cellulase activity)를, 결정형 셀룰로스(crystalline cellulose)(Avicel로부터 입수)를 갖는 Vogel's 최소배지(minimal media)에서 배양하였다 ; *N. 크라사*는 대략 4 일에 완전하게 아비셀을 분해하였다. *N. 크라사*는, 기능적인 셀룰라아제 및 헤미셀룰라아제 변성 능력(degradative capacity)을 제안하는, 그라운드 미스칸투스 줄기(ground *miscanthus* stem)에서 빠르게 또한 성장하였다. *N. 크라사*에서 식물 세포 벽 해체(deconstruction)와 관련된 전사체(transcriptome)를 측정하기 위해, 우리는 완전한 게놈 마이크로어레이(full genome microarrays)를 사용하여(Kasuga and Glass 2008; Tian 등의, 2007; Kasuga 등의, 2005), 그라운드 미스칸투스 줄기에서 *N. 크라사*의 성장 동안에 유전자 발현을 모니터링하였다. RNA 는 수크로스에서 16 시간 성장한 후에 표본 조사하였고(sample), 16 시간, 40 시간, 5 일 및 10 일에 미스칸투스 배지에서 성장한 *N. 크라사*로부터 분리된 RNA 와 비교하였다(도 4 ; Supplemental Data, Dataset SI, page 1 in Tian 등의, 2009 또한 참조 ; bioinfo.townsend.yale.edu/browse.jsp, Experiment IDs 52 and 53 에서 또한 발견할 수 있음).

[0221] 총 769 *N. 크라사* 유전자는, 수크로스 샘플과 비교한 바와 같이 4 개의 미스칸투스 샘플 중에서 상대적인 발현 수준에서 통계학적으로 중요한 차이점을 나타내었다(Supplemental Data, Dataset SI, page 3 in Tian 등의, 2009 참조). 계층적 군집(Hierarchical clustering)은 이러한 유전자가 세 가지 주요한 군집(three main clusters) 내에 속함을 나타낸 것이다(도 4 의 A). 유전자의 첫 번째 군집(C1; 300 유전자)은 탄소원(carbon source)으로서의 수크로스를 갖는 최소 배지에서 가장 높은 발현 수준을 나타낸 것이다. 이러한 유전자는 기능적인 카테고리 분석[Functional category(FunCat) analysis](Ruepp 2004)은, 호흡(respiration), 전자 전달 및 아미노산 대사와 같은 일차 물질대사(primary metabolism)와 연관된 리보솜단백질(ribosomal proteins) 및 그 밖의 기능적인 카테고리(functional categories)가 풍부함을 나타낸다(Supplemental Data, Dataset SI, page 4 in Tian 등의, 2009 참조). 두 번째 군집(C2)는 나중 시점(40 시간 내지 10 일)에서 미스칸투스 배양물에서 가장 높은 발현 수준을 나타내는 327 유전자를 포함한다(도 4 의 A). 이러한 그룹 내에서, 모든 시점에서 미스칸투스 배양물에서의 높은 상대적인 발현 수준을 나타낸 것은 89 유전자이다. 추가적인 분석에서, 이러한 89 유전자는 16 시간 미스칸투스 배양물로부터 가장 높은 발현 수준을 나타낸 유전자의 군집에 첨가하였다(C3 군집, 하기를 참조). 하나의 기능적인 카테고리(functional category)(C-화합물 및 탄수화물 대사)에서 나타낸 남아 있는 238 유전자의 FunCat 분석(Ruepp 2004)은 약간 강화되었다(enriched)(Supplemental Data, Dataset SI, page 5 in Tian 등의, 2009 참조).

[0222] 142 유전자의 세 번째 군집은 미스칸투스에서 *N. 크라사*의 16 시간 성장 후에 상대적으로 가장 높은 발현을 나타내었다(C3, 도 4 의 A). 모든 시점에서 미스칸투스 배양물에서의 높은 발현 수준을 나타낸 이러한 142 유전자 뿐만 아니라 89 유전자의 FunCat 분석(Ruepp 2004)(C3+ 군집 ; 총 231 유전자)은, 예측된 셀룰라아제 및 헤미셀룰라아제를 포함하는 탄소 대사와 연관된 단백질이 풍부함을 나타낸다(도 4 의 C ; 또한 Supplemental Data, Dataset SI, page 6 in Tian 등의, 2009 참조). *N. 크라사* 게놈에서 23 개의 예측된 셀룰라아제 유전자 중의 18 개는 미스칸투스에서 성장하는 동안, 특히 16 시간 시점에서 발현 수준이 현저한 증가를 나타내었다(도 5)(표 1 in Tian et al., 2009 참조). 5 개의 유전자는 200-폴드 상에서 발현 수준이 증가됨을 나타내었다 [cbh-1(CBH(I); NCU07340, gh6-2(CBH(II)-유사 유전자 ; NCU09680), gh6-3(NCU07190) 및 두 개의 GH61 유전자 (gh61-4 ; NCU01050 및 NCU07898)].

[0223] 식물 세포 벽은 셀룰로스 미세섬유, 헤미셀룰로스, 리그린(lignin), 펙틴(pectin), 각피질(cutin) 및 단백질로 이루어진 복합 구조물(complex structures)이다. 따라서, 우리는 *미스칸투스*에서 성장한 *N. 크라사*의 발현 프로파일과 아비셀에서 성장한 *N. 크라사*의 발현 프로파일, 순수한 형태(pure form)의 결정형 셀룰로스(crystalline cellulose)을 비교하였다(Tian 등의, 2009, Supplemental Data, Dataset S I, page 2 를 참조 ; data can also be found at bioinfo.townsend.yale.edu/browse.jsp, Experiment IDs 52 and 53 에서 또한 데이터를 발견할 수 있음). 187 이상 유전자는 아비셀에서 *N. 크라사*의 성장 동안에 상대적인 발현 수준에서 현저한 증가를 나타내었다. 이러한 유전자 중에 114 는 C3+ 군집에서 231 유전자가 중복되었다(도 4 의 B 참조). 114-오버랩 유전자 세트(the 114-overlap gene set)의 FunCat 분석은 탄소대사에서 포함될 수 있는 것으로 예측된 유전자에 대해서 분명한 풍부함을 나타내었다(Supplemental Data, Dataset S I, page 6 in Tian et al., PNAS, 2009 를 참조). 이러한 유전자 세트내에, 분비된 단백질에 대한 추가적인 풍부함도 있었다 ; 114 개의 유전자 산물 중의 53 개 유전자 산물은 분비될 것으로 예측되었다. 이러한 53 개의 유전자 중에서 32 개의 유전자는 식물 세포 벽 분해(degradation)에서 역할을 제시하는 아노테이션(annotation)을 갖는 것으로 예측된 단백질을 코드하는 반면에 16 유전자는 어떠한 기능적인 예측이 결핍된 것으로 추정되거나(putative) 또는 가상적인 단백질(hypothetical proteins)를 코드한다. 남아있는 61 개의 유전자는 주요한 촉진자 상과 운반체로 예측된 10 개(NCU00801, NCU00988, NCU01231, NCU04963, NCU05519, NCU05853, NCU05897, NCU06138, NCU08114 및 NCU 10021) 및 23 개의 추정되거나 또는 가상적인 단백질을 포함하는, 예측된 세포 내의 단백질을 코드한다.

[0224] *미스칸투스*-특이적인 군집(Miscanthus-specific cluster) 내의 117 개의 유전자 중에 37 개(도 4의 B)는 분비될 것으로 예측되는 단백질을 코드한다. 헤미셀룰로스의 저하(degradation)와 관련된 9 개의 예측된 헤미셀룰라아제 또는 효소는 확인되었다(NCU00710, NCU04265, NCU04870, NCU05751, NCU05965, NCU09170, NCU09775, NCU09923 및 NCU09976)(Tian 등의, 2009 - 표 2). 남아있는 80 개의 *미스칸투스*-특이적인 유전자는, 오타당(예를 들어, NCU00891, 자일리톨 탈수소효소 및 NCU00643, 예측된 아라비니톨 탈수소효소)의 대사에 포함되는 유전자를 포함하는 예측된 세포내 단백질, 예측된 당 운반체(NCU01132) 및 알려지지 않은 기능(unknown function)의 48 개의 단백질을 코드한다.

실시예 2

[0225] 실시예 2 - *미스칸투스* 및 아비셀에서 성장한 *N. 크라사*의 세크레토믹 분석(Secretome Analysis)

[0226] 진균(fungi)에 의한 리그노셀룰로스 저하(Lignocellulose degradation)는 세포 외에서 일어나고 세포 벽 구성성분(cell wall constituents)의 디폴리머라이제이션(depolymerization)과 연관된 단백질 분비를 필요로 한다(Lynd 등의, 2002). *N. 크라사*가 *미스칸투스* 또는 아비셀에서 성장하였을 때, 유전자 코딩하는 것을 예측된 셀룰라아제, 헤미셀룰라아제 및 기타 분비된 단백질은 발현 수준이 증가됨을 나타낸 전사 프로파일링 데이터(transcriptional profiling data)와 비교하기 위해, 우리는 샷건 프로테오믹스 처리 방법(shotgun proteomics approach)을 사용한 *N. 크라사*의 세크레토믹(secretome)을 분석하였다(도 4 의 B). 7 일 된 *미스칸투스* 및 아비셀 배양물로부터의 상청액을 트리핀으로 소화시키고, 액체 크로마토그래피 나노-전자분사 이온화 이중 질량분석기(nano-electrospray ionization tandem mass spectrometry)(MS ; 실시예 5 참조)에 의해 분석하였다. 인산스웰린 셀룰로스(phosphoric acid swollen cellulose, PASC)에 결합하는 분비된 단백질은 풍부하고(enriched) 또한 MS 에 의해 분석하였다.

[0227] 총 50 개의 단백질은 이중 MS 에 의해 확실하게 확인할 수 있었다(표 2 및 표 3). *N. 크라사* 배양물에서 성장한 *미스칸투스*에서 34 개의 단백질이 발견된 반면에, 배양물에서 성장한 아비셀로부터 38 개의 단백질이 확인되었다 ; 22 개의 단백질은 둘 다의 샘플에서 발견되었다. 이러한 22 개의 단백질 중에, 21 개는 컴퓨터 분석(computational analyses)을 기반으로 분비되는 것을 예측되었고, 19 개는 배양물에서 성장한 *미스칸투스* 및 아비셀 둘 다에서 발현 수준이 증가됨을 나타내었다(표 2). 오버랩 데이터세트(overlap dataset)는 *N. 크라사*에서 23 개의 예측된(predicted) 셀룰라아제의 8 개를 포함한다(표 3). 5 개의 예측된 헤미셀룰라아제, 예측된 P-글루코시다아제(*gh3-4*; NCU04952), 탄수화물에서 예측된 활성도를 갖는 5 개의 단백질 및 알려지지 않는 기능을 갖는 2 개의 단백질(NCU07143 and NCU05137)이 또한 있었다(표 4-5).

[0228] 표 2 에 대해서, 아노테이션(annotation)은 웹사이트 broad.mit.edu/ annotation/genome/neurospora/Home.html 에서의 Broad Institute에 의해 생성되었다. "발견된 샘플(sample detected)"은 특정한 단백질에 대해 발견된 펩티드에서의 샘플이다. 펩티드는 점검 메뉴얼(manual inspection)에 의해 승인되었다. 만약 적어도 1 개의 펩티드가 각각의 생물학적 반복(biological repeat)에서 발견된다면 단백질은 존재하는 것으로 결정되었다. "전체(TOTAL)"은 모든 세포외 단백질의 트립신 소화로부터 발견된 펩티드를 언급한다. "PASC BOUND" 는 인산 스웰런 셀룰로스(phosphoric acid swollen cellulose, PASC)에 결합한 단백질이 풍부하게 된(enrichment) 후에 발견된 펩티드를 언급한다. "언바운드(UNBOUND)" 은 PASC 결합 단백질이 제거된 후의 용액에 남아있는 단백질을 언급한다.

[0229] [표 2]

네우로스포라 크라사에서 성장한 아비셀의 배양 여과액(culture filtrates)에서 LC-MS/MS에 의한 확인된 단백질

유전자 ID	아노테이션(ANNOTATION)	발견된 샘플
NCU00206	셀로비오스 탈수소효소 830 nt 와 비슷한 네우로스포라 크라사 가설 단백질(hypothetical protein)	전체 (TOTAL)
NCU00762	네우로스포라 크라사 엔도글루카나아제(endoglucanase) 3 전구체 (precursor) 391 nt	전체
NCU01050	엔도글루카나아제 II 239 nt 와 유사한 네우로스포라 크라사 가설단백질	전체
NCU02343	알파 L 아라비노푸라노시다아제 A 668 nt 와 유사한 네우로스포라 크라사 가설단백질	전체
NCU04870	아세틸 자일란 에스테라아제(acetyl xylan esterase) 313 nt 와 유사한 네우로스포라 크라사 가설단백질	전체
NCU04952	베타 D 글루코시드 글루코히드로라아제(beta D glucoside glucohydrolase) 736 nt 와 유사한 네우로스포라 크라사 가설단백질	전체
NCU05137	네우로스포라 크라사 보존(conserved) 가설단백질 692 nt	전체
NCU05159	네우로스포라 크라사 아세틸자일란 에스테라아제 전구체 301 nt	전체
NCU05924	네우로스포라 크라사 엔도 1 4 베타 자일라나아제(endo 1 4 beta xylanase) 330 nt	전체
NCU07143	네우로스포라 크라사 예측 단백질(Neurospora crassa predicted protein) 391 nt	전체
NCU07190	네우로스포라 크라사 보존 가설단백질 384 nt	전체
NCU07225	네우로스포라 크라사 엔도 1 4 베타 자일라나아제 2 전구체 255 nt	전체
NCU07326	네우로스포라 크라사 예측 단백질 327 nt	전체
NCU07340	네우로스포라 크라사 엑소글루카나아제(exoglucanase) 1 전구체 522 nt	전체
NCU07898	네우로스포라 크라사 예측 단백질 242 nt	전체
NCU08189	엔도 1 4 베타 자일라나아제 385 nt 와 유사한 네우로스포라 크라사 가설단백질	전체
NCU08398	네우로스포라 크라사 보존 가설단백질 391 nt	전체
NCU08760	네우로스포라 크라사 예측 단백질 343 nt	전체
NCU08785	네우로스포라 크라사 보존 가설단백질 291 nt	전체
NCU09491	네우로스포라 크라사 페루로일 에스테라아제(feruloyl esterase) B 전구체 293 nt	전체
NCU09680	네우로스포라 크라사 엑소글루카나아제 2 전구체 485 nt	전체
NCU09923	베타 자일로시다아제(beta xylosidase) 775 nt 와 유사한 네우로스포라 크라사 가설단백질	전체

[0230]

NCU00206	셀로비오스 탈수소효소 830 nt 와 유사한 네우로스포라 크라사 가설 단백질	PASC BOUND
NCU00762	네우로스포라 크라사 엔도글루카나아제 3 전구체 391 nt	PASC BOUND
NCU05159	네우로스포라 크라사 아세틸자일란 에스테라아제 전구체 301 nt	PASC BOUND
NCU05955	Cel74a 862 nt 과 유사한 네우로스포라 크라사 가설단백질	PASC BOUND
NCU07225	네우로스포라 크라사 엔도 1 4 베타 자일라나아제 2 전구체 255 nt	PASC BOUND
NCU07340	네우로스포라 크라사 엑소글루카나아제 1 전구체 522 nt	PASC BOUND
NCU08760	네우로스포라 크라사 예측 단백질 343 nt	PASC BOUND
NCU09680	네우로스포라 크라사 엑소글루카나아제 2 전구체 485 nt	PASC BOUND
NCU09708	네우로스포라 크라사 보존 가설단백질 465 nt	PASC BOUND
NCU00762	네우로스포라 크라사 엔도글루카나아제 3 전구체 391 nt	언바운드 (UNBOUND)
NCU01651	네우로스포라 크라사 보존 가설단백질 783 nt	언바운드
NCU02343	알파 L 아라비노푸라노시다아제 A 668 nt 와 유사한 네우로스포라 크라사 가설단백질	언바운드
NCU04202	네우로스포라 크라사 뉴클레오시드 디포스페이트 키나아제 1 153 nt	언바운드
NCU04870	아세틸 자일란 에스테라아제 313 nt 와 유사한 네우로스포라 크라사 가설단백질	언바운드
NCU04952	베타 D 글루코시드 글루코히드로라아제(beta D glucoside glucohydrolase) 736 nt 와 유사한 네우로스포라 크라사 가설단백질	언바운드
NCU05057	네우로스포라 크라사 엔도글루카나아제 EG 1 전구체 439 nt	언바운드
NCU05137	네우로스포라 크라사 보존 가설단백질 692 nt	언바운드
NCU05751	네우로스포라 크라사 보존 가설단백질 242 nt	언바운드
NCU05924	네우로스포라 크라사 엔도 1 4 베타 자일라나아제 330 nt	언바운드
NCU06239	네우로스포라 크라사 보전 가설단백질 514 nt	언바운드
NCU07143	네우로스포라 크라사 예측 단백질 391 nt	언바운드
NCU07190	네우로스포라 크라사 보전 가설단백질 384 nt	언바운드
NCU07225	네우로스포라 크라사 엔도 1 4 베타 자일라나아제 2 전구체 255 nt	언바운드
NCU07326	네우로스포라 크라사 예측 단백질 327 nt	언바운드
NCU07898	네우로스포라 크라사 예측 단백질 242 nt	언바운드
NCU08189	엔도 1 4 베타 자일라나아제 385 nt 와 유사한 네우로스포라 크라사 가설단백질	언바운드
NCU08398	네우로스포라 크라사 보전 가설단백질 391 nt	언바운드
NCU08412	네우로스포라 크라사 보전 가설단백질 401 nt	언바운드
NCU08760	네우로스포라 크라사 예측 단백질 343 nt	언바운드

[0231]

NCU08785	네우로스포라 크라사 보전 가설단백질 291 nt	언바운드
NCU09024	네우로스포라 크라사 보전 가설단백질 625 nt	언바운드
NCU09175	네우로스포라 크라사 보전 가설단백질 411 nt	언바운드
NCU09267	네우로스포라 크라사 보전 가설단백질 1048 nt	언바운드
NCU09491	네우로스포라 크라사 페루로일 에스테라아제 B 전구체 293 nt	언바운드
NCU09775	알파 L 아라비노푸라노시다아제 343 nt 와 유사한 네우로스포라 크라사 가설단백질	언바운드
NCU09923	베타 자이로시다아제 775 nt 와 유사한 네우로스포라 크라사 가설단백질	언바운드

[0232]

[0233]

[표 3]

미스칸터스 및 아비셀 배양물 둘 다에서 발견된 22 개의 분비된 단백질

유전자 이름	유전자 아노테이션 (Gene annotation)	프로파일링 (Profiling)	kos	CBM1	신호 P
NCU00206.2	CBDH	둘 다	heter	예(yes)	예
NCU00762.2	개연성있는 셀룰라아제 전구체 (probable cellulase precursor)	둘 다	16747	예	예
NCU01050.2	cel 1 단백질 전구체와 관련 (related to cel1 protein precursor)	둘 다	16543	아니요 (no)	예
NCU04952.2	개연성있는 베타-D-글루코시드 글루코하이드롤라아제 (probable beta-D-glucoside glucohydrolase)	둘 다	13732	아니요	예
NCU05057.2	개연성있는 엔도-1,4-베타-글루카나아제 (probable endo-1,4-beta-glucanase)	둘 다	13343	아니요	예
NCU05137.2	보존 가설단백질 (conserved hypothetical protein)	둘 다	11682	아니요	예
NCU05924.2	개연성있는 엔도-베타-1,4-D-자일라나아제 (probable endo-beta-1,4-D-xylanase)	둘 다	15626	아니요	예
NCU05955.2	개연성있는 엔도글루카나아제 C	둘 다	13535	예	예
NCU07143.2	가상(hypothetical)	둘 다	아니요	아니요	예
NCU07190.2	CBHII 동족체(homolog)	둘 다	19315	아니요	예
NCU07225.2	개연성있는 엔도-1,4-베타-자일라나아제 B 전구체	둘 다	heter	예	예
NCU07326.2	추정상의 아라비나아제와 관련 (related to putative arabinase)	둘 다	19534	아니요	예
NCU07340.2	CBHI	둘 다	15630	예	예
NCU07898.2	cel 1 단백질 전구체와 관련	둘 다	19600	아니요	예
NCU08189.2	엔도-1,4-베타 자일라나아제와 유사 (similar to endo-1,4-beta xylanase)	둘 다	19861	아니요	예
NCU08398.2	알도스 1-에피머화 효소와 관련 (related to aldose 1-epimerase)	둘 다	20310	아니요	예
NCU08412.2	가설단백질 401 nt	없음 (NONE)	아니요 (No)	아니요	아니요
NCU08760.2	패밀리 61 엔도글루카나아제와 관련 (related to family 61 endoglucanase)	둘 다	15664	예	예
NCU09024.2	가설단백질 625 nt	없음 (NONE)	아니요	아니요	예
NCU09175.2	글루칸 1,3-베타-글루코시다아제 전구체	둘 다	11750	아니요	예
NCU09491.2	페루로일 에스테라아제 B 전구체 (feruloyl esterase B precursor)	mis	아니요	아니요	예
NCU09680.2	CBHII	둘 다	15633	예	예

[0234]

[0235] [표 4]

네우로스포라 크라사에서 예측된 셀룰라아제 유전자

유전자	GH ¹ 패밀리 (family)	CBM1 ²	SP ³	MS ⁴	EL ⁵ 미스칸터스	EL ⁵ 아비셀
NCU00762	5	예(yes)	예(yes)	둘 다 (both)	29.6	31.5
NCU03996	6	아니요(no)	아니요	ND ⁶	ND	ND
NCU07190	6	아니요	예	둘 다	526.0	119
NCU09680	6	예	예	둘 다	230.9	251.3
NCU04854	7	아니요	예	ND	32.9	10.8
NCU05057	7	아니요	예	둘 다	8.7	7.4
NCU05104	7	아니요	예	ND	11.6	NC ⁷
NCU07340	7	예	예	둘 다	426.4	382.2
NCU05121	45	예	예	avi	8.6	17.2
NCU00836	61	예	예	ND	91.2	31
NCU01050	61	아니요	예	둘 다	206.7	382.1
NCU01867	61	예	예	ND	2.2	NC
NCU02240	61	예	예	avi	193.5	84
NCU02344	61	아니요	예	ND	8.1	4.1
NCU02916	61	예	예	ND	85.2	17.7
NCU03000	61	아니요	예	ND	NC	ND
NCU03328	61	아니요	예	ND	26.4	23.8
NCU05969	61	아니요	예	ND	ND	12.7
NCU07520	61	아니요	예	ND	ND	ND
NCU07760	61	예	예	ND	3.7	NC
NCU07898	61	아니요	예	둘 다	376.3	230
NCU07974	61	아니요	예	ND	NC	NC
NCU08760	61	예	예	둘 다	107.5	44.7

GH 패밀리 - 글루코실

¹글루코시드 가수분해효소 ; ²CBM1, 탄수화물 결합 모듈(module) ; ³신호펩티드 예측(신호 P = 웹페이지 cbs.dtu.dk/services/SignalP/) ; ⁴MS, 질량 스펙트럼 분석 ; ⁵EL, 상대적인 발현 수준 ; ⁶ND, 발견되지 않음 ; ⁷NC, 발현 수준 대 최소배지에서 변화가 없음.

[0236]

[0237]

[표 5]

LC-MS/MS에 의해 확인된 셀룰라아제 및 헤미셀룰라아제

a. 네우로스포라 크라사의 계능에서 예측된 셀룰라아제

유전자 ID	GH 패밀리	AV MS	MIS MS	AV 어레이 (ARRAY)	MIS 어레이
NCU00762	5	+	+	31.5	29.6
NCU00836	61	-	-	31	91.2
NCU01050	61	+	+	382.1	206.7
NCU01867	61	-	-	1	1
NCU02240	61	+	-	84	193.5
NCU02344	61	-	-	4.1	8.1
NCU02916	61	-	-	17.7	85.2
NCU03000	61	-	-	1	1
NCU03328	61	-	-	23.8	26.4
NCU03996	6	-	-	2.5	6.3
NCU04854	7	-	-	10.8	32.9
NCU05057	7	+	+	7.4	8.7
NCU05104	7	-	-	1	1
NCU05121	45	+	-	17.2	8.6
NCU05969	61	-	-	12.7	12.3
NCU07190	6	+	+	119	526
NCU07340	7	+	+	382.2	426.4
NCU07520	61	-	-	1	1
NCU07760	61	-	-	1	3.4
NCU07898	61	+	+	230.5	376.3
NCU07974	61	-	-	1	1
NCU08760	61	+	+	44.7	107.5
NCU09680	6	+	+	251.3	230.9

[0238]

b. 네우로스포라 크라사의 게놈에서 예측된 헤미셀룰라아제

유전자 ID	GH 패밀리	AV MS	MIS MS	AV 어레이 (ARRAY)	MIS 어레이
NCU00852	43	-	-	1	1
NCU00972	53	-	-	9.03	15.6
NCU01900	43	-	-	10.03	26
NCU02343	51	-	+	6.63	174.6
NCU02855	11	+	-	10.2	364
NCU04997	10	-	-	1	25.6
NCU05924	10	+	+	55.9	149.3
NCU05955	74	+	+	19.9	50.5
NCU05965	43	-	-	1	5.4
NCU06861	43	-	-	1	1
NCU07130	10	-	-	1	1
NCU07225	11	+	+	11.43	33.5
NCU07326	43	+	+	104.5	426.6
NCU07351	67	-	-	1	1
NCU08087	26	-	-	1	1
NCU08189	10	+	+	39.8	94.4
NCU09170	43	-	-	1	16.7
NCU09652	43	-	-	12.2	95.4
NCU09775	54	-	+	1	48.3

GH 패밀리 - 글루코실 가수분해효소 패밀리(Glycosyl Hydrolase Family) ;

AV MS - 네우로스포라 크라사에서 성장한 아비셀의 배양 여과액에서 LC-MS/MS에 의해 발견된 단백질. (+) 발견됨, (-) 발견되지 않음 ;

MIS MS - 네우로스포라 크라사에서 성장한 미스칸터스의 배양 여과액에서 LC-MS/MS에 의해 발견된 단백질. (+) 발견됨, (-) 발견되지 않음 ;

AV 어레이(ARRAY) - 프로파일링 데이터(profiling data)로부터의 수크로스에서 16 시간 성장에 비례한(relative to) 아비셀에서 30 시간 성장한 후의 폴드 상향 조절(Fold upregulation) ;

MIS 어레이 - 프로파일링 데이터(profiling data)로부터의 수크로스에서 16 시간 성장에 비례한 미스칸터스에서 16 시간 성장한 후의 폴드 상향 조절. 펩티드는 미스칸터스 배양 여과액에서만 발견되었다.

[0239]

[0240]

아비셀 배양물에서만 16 개의 단백질이 확실하게 확인되었고, 이들 중 14 개는, 두 개의 예측된 셀룰라아제(*gh61-1*; NCU02240 및 *gh45-1*; NCU05121), 하나의 자일라나아제(xylanase)(*gh11-1*; NCU02855), 하나의 예측된 프로테아제(protease)(NCU04205), 탄수화물에서 예측된 활성을 갖는 다른 세 개의 단백질[NCU08909, NCU05974 및 *gh30-1* (NCU04395)], 알려지지 않는 기능의 세 개의 *네우로스포라*-특이적인 단백질, 및 셀룰로스 결합 도메인을 갖는 하나의 단백질을 포함하는 4 개의 보존 가설단백질(conserved hypothetical proteins)(NCU09764)을 포함하고 이들은 분비되는 것으로 예측되었다(표 6). 12 개의 단백질은 *미스칸터스* 배양물의 배양 여과액에 대해 특이적이고, 이들 중 7개는 분비되는 것으로 예측된다(표 3). 5 개의 예측된 세포 내 단백질 중의 3 개는 보존 가설단백질이다. 남은 2 개의 단백질은 예측된 글리옥살 산화효소(glyoxal oxidase)[NCU09267, *N. 크라사 미스칸터스* 전사체(transcriptome)로부터 확인됨] 및 뉴클레오시드 이인산염 키나아제(nucleoside diphosphate kinase)[*ndk-1*; NCU04202, *N. 크라사 미스칸터스* 전사체로부터 확인되지 않음]를 포함한다. 7 개의 단백질은, 3 개의 예측된 에스테라아제(NCU04870, NCU05159 및 NCU08785), 두 개의 예측된 자일라나아제(GH51 ; NCU02343 및 GH54 ; NCU09775), 예측된 P-자일로시다아제(*gh3-7*; NCU09923) 및 보존 가설단백질(NCU05751)을 포함하고, 분비되는 것으로 예측되었다.

[0241]

[표 6]

네우로스포라 크라사에서 성장한 아비셀의 배양 여과액에서 LC-MS/MS에 의해 확인된 단백질

유전자 ID	아노테이션(ANNOTATION)	발견된 샘플
NCU00206	셀로비오스 탈수소효소 830 nt 와 유사한 네우로스포라 크라사 가설 단백질 (Neurospora crassa hypothetical protein similar to cellobiose dehydrogenase 830 nt)	전체 (TOTAL)
NCU00762	네오로스포라 크라사 엔도글루카나아제 3 전구체 391 nt (Neurospora crassa endoglucanase 3 precursor 391 nt)	전체
NCU00798	네오로스포라 크라사 예측 단백질 313 nt (Neurospora crassa predicted protein 313 nt)	전체
NCU01050	엔도글루카나아제 II 239 nt 와 유사한 네오로스포라 크라사 가설 단백질 (Neurospora crassa hypothetical protein similar to endoglucanase II 239 nt)	전체
NCU01595	네우로스포라 크라사 단백질 SOF1 446 nt	전체
NCU02240	엔도글루카나아제 II 323 nt 와 유사한 네오로스포라 크라사 가설 단백질	전체
NCU02696	DEAD DEAH 박스 RNA 헬리카제 1195 nt 와 유사한 네우로스포라 크라사 가설단백질 (Neurospora crassa hypothetical protein similar to DEAD DEAH box RNA helicase 1195 nt)	전체
NCU02855	네우로스포라 크라사 엔도 1 4 베타 자일라나아제 A 전구체 221 nt (Neurospora crassa endo 1 4 beta xylanase A precursor 221 nt)	전체
NCU04952	베타 D 글루코시드 글루코히드롤라아제 736 nt 와 유사한 네우로스포라 크라사 가설단백질 (Neurospora crassa hypothetical protein similar to beta D glucoside glucohydrolase 736 nt)	전체
NCU05057	네우로스포라 크라사 엔도글루카나아제 EG 1 전구체 439 nt	전체
NCU05137	네우로스포라 크라사 보존 (conserved) 가설단백질 692 nt	전체
NCU05924	네우로스포라 크라사 엔도 1 4 베타 자일라나아제 330 nt	전체
NCU05955	Cel74a 862 nt 와 유사한 네우로스포라 크라사 가설단백질	전체
NCU07143	네우로스포라 크라사 예측 단백질 391 nt	전체
NCU07190	네우로스포라 크라사 보존 가설단백질 384 nt	전체
NCU07225	네우로스포라 크라사 엔도 1 4 베타 자일라나아제 2 전구체 255 nt	전체
NCU07326	네우로스포라 크라사 예측 단백질 327 nt	전체
NCU07340	네우로스포라 크라사 엑소글루카나아제 1 전구체 522 nt	전체
NCU07898	네우로스포라 크라사 예측 단백질 242 nt	전체
NCU08171	네우로스포라 크라사 예측 단백질 382 nt	전체
NCU08412	네우로스포라 크라사 보존 가설단백질 401 nt	전체
NCU08760	네우로스포라 크라사 예측 단백질 343 nt	전체
NCU09491	네우로스포라 크라사 페루로일 에스테라아제 (feruloyl esterase) B 전구체 293 nt	전체

[0242]

NCU09680	네우로스포라 크라사 엑소글루카나아제(exoglucanase) 2 전구체 485 nt	전체
NCU09764	네우로스포라 크라사 보전 가설단백질 406 nt	전체
NCU00206	셀로비오스 탈수소효소 830 nt 와 유사한 네우로스포라 크라사 가설 단백질	PASC BOUND
NCU00762	네우로스포라 크라사 엔도글루카나아제 3 전구체 391 nt	PASC BOUND
NCU05121	네우로스포라 크라사 엔도글루카나아제 V 294 nt	PASC BOUND
NCU05955	Cel74a 862 nt 와 유사한 네우로스포라 크라사 가설단백질	PASC BOUND
NCU07225	네우로스포라 크라사 엔도 1 4 베타 자일라나아제 2 전구체 255 nt	PASC BOUND
NCU07340	네우로스포라 크라사 엑소글루카나아제 1 전구체 522 nt	PASC BOUND
NCU08760	네우로스포라 크라사 예측 단백질 343 nt	PASC BOUND
NCU09680	네우로스포라 크라사 엑소글루카나아제 2 전구체 485 nt	PASC BOUND
NCU00206	셀로비오스 탈수소효소 830 nt 와 유사한 네우로스포라 크라사 가설 단백질	언바운드 (UNBOUND)
NCU00762	네우로스포라 크라사 엔도글루카나아제 3 전구체 391 nt	언바운드
NCU00798	네우로스포라 크라사 예측 단백질 313 nt	언바운드
NCU01050	엔도글루카나아제 II 239 nt 와 유사한 네우로스포라 크라사 가설단백질	언바운드
NCU04205	네우로스포라 크라사 예측 단백질 346 nt	언바운드
NCU04395	네우로스포라 크라사 엔도 1 6 베타 D 글루카나아제 전구체 481 nt	언바운드
NCU04952	베타 D 글루코시드 글루코히드롤라아제 736 nt 와 유사한 네우로스포라 크라사 가설단백질	언바운드
NCU05057	네우로스포라 크라사 엔도글루카나아제 EG 1 전구체 439 nt	언바운드
NCU05134	네우로스포라 크라사 가설단백질 124 nt	언바운드
NCU05137	네우로스포라 크라사 보존 가설단백질 692 nt	언바운드
NCU05852	네우로스포라 크라사 보존 가설단백질 254 nt	언바운드
NCU05924	네우로스포라 크라사 엔도 1 4 베타 자일라나아제 330 nt	언바운드
NCU05974	세포 벽 글루카노실트랜스페라아제(cell wall glucanoyltransferase) Mwg1 365 nt 와 유사한 네우로스포라 크라사 가설단백질	언바운드
NCU07143	네우로스포라 크라사 예측 단백질 391 nt	언바운드
NCU07190	네우로스포라 크라사 보전 가설단백질 384 nt	언바운드
NCU07225	네우로스포라 크라사 엔도 1 4 베타 자일라나아제 2 전구체 255 nt	언바운드
NCU07326	네우로스포라 크라사 예측 단백질 327 nt	언바운드
NCU07340	네우로스포라 크라사 엑소글루카나아제 1 전구체 522 nt	언바운드

[0243]

NCU07898	네우로스포라 크라사 예측 단백질 242 nt	언바운드
NCU08171	네우로스포라 크라사 예측 단백질 382 nt	언바운드
NCU08189	엔도 1 4 베타 자일라나아제 385 nt 와 유사한 네우로스포라 크라사 가설단백질	언바운드
NCU08398	네우로스포라 크라사 보존 가설단백질 391 nt	언바운드
NCU08412	네우로스포라 크라사 보존 가설단백질 401 nt	언바운드
NCU08760	네우로스포라 크라사 예측 단백질 343 nt	언바운드
NCU08909	베타 1 3 글루카노실트랜스페라아제 543 nt 와 유사한 네우로스포라 크라사 가설단백질	언바운드
NCU08936	네우로스포라 크라사 클락 조절된 유전자(clock controlled gene) 15 412 nt	언바운드
NCU09024	네우로스포라 크라사 보존 가설단백질 625 nt	언바운드
NCU09046	네우로스포라 크라사 예측 단백질 187 nt	언바운드
NCU09175	네우로스포라 크라사 보존 가설단백질 411 nt	언바운드
NCU09491	네우로스포라 크라사 페루로일 에스테라아제 B 전구체 293 nt	언바운드

아노테이션(ANNOTATION) - Broad Institute 에 의해 생성(웹페이지 broad.mit.edu/annotation/genome/neurospora/Home.html) ;

발견된 샘플 - 특정한 단백질에 대해 발견된 펩티드에서의 샘플. 펩티드는 점검 매뉴얼(manual inspection)에 따라 승인되었다. 만약 적어도 1 개의 펩티드가 각각의 생물학적 반복에서 발견된다면 단백질은 존재하는 것으로 결정되었다.

전체(TOTAL) - 모든 세포 외 단백질의 트립신 소화(tryptic digest)로부터 발견된 펩티드 ;

PASC BOUND - 인산 스웰런 셀룰로오스(phosphoric acid swollen cellulose)에 결합한 단백질에 대해 풍부하게 된 후에 발견된 펩티드 ;

언바운드(UNBOUND) - PASC 결합 단백질의 제거 후에 용액에서 남아있는 단백질.

[0244]

[0245]

많은 식물 세포벽 분해효소(plant cell wall degrading enzymes)는, 기질에 효소를 부가하는 셀룰로스-결합 모듈(cellulose-binding module, CBM)을 함유한다(Linder and Teeri 1996). N. 크라사 게놈 내의, 19 가지 유전자를 코드하는 단백질은 CBM1 도메인을 함유하는 것을 예상하였다(Cantarel 등의, 2009). 이러한 19 가지 유전자 중에, 16 가지는 미스칸터스-성장한 배양물에서 상대적인 유전자 발현에 있어서 증가됨을 나타내었다(표 7).

[0246] [표 7]

N. 크라사 유전자 발현에서의 *미스칸터스* 및 *아비셀*의 효과

유전자 이름	CBM 예측	아노테이션 (Annotation)	Mis 어레이 (Array)	아비셀 어레이	MS
NCU00206	cazy 및 mips	개연성있는 셀로비오스 탈수소효소	164	12	둘 다 (both)
NCU00710	cazy 및 mips	아세틸자일란 에스테라아제	30	발견되지 않음 (no detect)	없음 (none)
NCU00762	cazy 및 mips	EG2	29	31	둘 다
NCU00836	cazy 및 mips	EG, GH61	91	30	없음
NCU01867	cazy 및 mips	EG, GH61	2.2-dl0	차이없음 (no Difference)	없음
NCU02240	cazy 및 mips	EG, GH61	193	84	avi
NCU02916	cazy 및 mips	EG, GH61	85	17	없음
NCU04500	cazy 및 mips	키틴나아제 4 (chitinase 4)와 유사	발견되지 않음	발견되지 않음	없음
NCU04997	cazy 및 mips	자일라나아제와 유사	발견되지 않음	발견되지 않음	없음
NCU05121	cazy 및 mips	EG, GH45	8.5	17	avi
NCU05159	cazy 및 mips	아세틸자일란 에스테라아제 전구체	34	10	mis
NCU05955	cazy 및 mips	GH74	50	19	둘 다
NCU07225	cazy 및 mips	자일라나아제	33	11	둘 다
NCU07340	cazy 및 mips	CBH1	426	382	둘 다
NCU07760	cazy 및 mips	EG,GH61	3.7	차이없음	없음
NCU08760	cazy 및 mips	EG,GH61	107	44	둘 다
NCU09416	cazy 및 mips	가상(hypothetical)	발견되지 않음	27	없음
NCU09680	cazy 및 mips	CBH2	230	251	둘 다
NCU09764	cazy 및 mips	가상	18	16.6	avi

[0247]

[0248] MS 에 의해 확인된 50 개의 단백질로부터, 11 개는 CBM1 도메인을 함유한다. PASC 는 셀룰로스에 결합하는 단백질을 풍부하게 하기 위해(enrich) 사용되었다(방법에 대해서 실시예 4 에 참조). 9 개의 단백질은 *N. 크라사* 배양물에서 성장한 *미스칸터스*의 상청액으로부터의 PASC 에 결합하는 MS 에 의해 확인한 반면에, *아비셀* 상청액으로부터의 8 개의 단백질을 확인하였다 ; 7 개의 셀룰로스 결합 단백질은 둘 다에서 확인되었다(표 2, 3 및 8). 이들은 NCU00206, 예측 셀로비오스 탈수소효소(cellobiose dehydrogenase) ; *gh5-1*(NCU00762), 예측 엔도글루카나아제(predicted endoglucanase) ; NCU05955, 예측 GH74 자일로글루카나아제(xyloglucanase) ; *gh11-2*(NCU07225), 예측 엔도자일라나아제(predicted endoxylanase) ; *cbh-1*(NCU07340) ; *gh61-5*(NCU08760), 예측 엔도글루카나아제 ; 및 *gh6-2*(NCU09680), 예측 셀로비오히드롤라아제 2 전구체(predicted cellobiohydrolase 2 precursor)를 포함한다.

[0249]

[표 8]

아비셀-성장한 *네우로스포라 크라사*의 배양 여과액에서 LC-MS/MS에 의해 확인된 단백질

유전자 ID	아노테이션(ANNOTATION)	배양물 (CULTURE)
NCU00206	셀로비오스 탈수소효소 830 nt 와 유사한 <i>네우로스포라 크라사</i> 가설 단백질(Neurospora crassa hypothetical protein similar to cellobiose dehydrogenase 830 nt)	둘 다 (BOTH)
NCU00762	<i>네우로스포라 크라사</i> 엔도글루카나아제(endoglucanase) 3 전구체 (precursor) 391 nt	둘 다
NCU01050	엔도글루카나아제 II 239 nt 와 유사한 <i>네우로스포라 크라사</i> 가설단백질	둘 다
NCU04952	베타 D 글루코시드 글루코히드로라아제(beta D glucoside glucohydrolase) 736 nt 와 유사한 <i>네우로스포라 크라사</i> 가설단백질	둘 다
NCU05057	<i>네우로스포라 크라사</i> 엔도글루카나아제(endoglucanase) EG 1 전구체 439 nt	둘 다
NCU05137	<i>네우로스포라 크라사</i> 보존(conserved) 가설단백질 692 nt	둘 다
NCU05924	<i>네우로스포라 크라사</i> 엔도 1 4 베타 자일라나아제(endo 1 4 beta xylanase) 330 nt	둘 다
NCU05955	Cel74a 862 nt 와 유사한 <i>네우로스포라 크라사</i> 가설단백질	둘 다
NCU07143	<i>네우로스포라 크라사</i> 예측 단백질(Neurospora crassa predicted protein) 391 nt	둘 다
NCU07190	<i>네우로스포라 크라사</i> 보존 가설단백질 384 nt	둘 다
NCU07225	<i>네우로스포라 크라사</i> 엔도 1 4 베타 자일라나아제 2 전구체 255 nt	둘 다
NCU07326	<i>네우로스포라 크라사</i> 예측 단백질 327 nt	둘 다
NCU07340	<i>네우로스포라 크라사</i> 엑소글루카나아제(exoglucanase) 1 전구체 522 nt	둘 다
NCU07898	<i>네우로스포라 크라사</i> 예측 단백질 242 nt	둘 다
NCU08189	엔도 1 4 베타 자일라나아제 385 nt 와 유사한 <i>네우로스포라 크라사</i> 가설단백질	둘 다
NCU08398	<i>네우로스포라 크라사</i> 보존 가설단백질 391 nt	둘 다
NCU08412	<i>네우로스포라 크라사</i> 보존 가설단백질 401 nt	둘 다
NCU08760	<i>네우로스포라 크라사</i> 예측 단백질 343 nt	둘 다
NCU09024	<i>네우로스포라 크라사</i> 보존 가설단백질 625 nt	둘 다
NCU09175	<i>네우로스포라 크라사</i> 보존 가설단백질 411 nt	둘 다
NCU09491	<i>네우로스포라 크라사</i> 페루로일 에스테라아제(feruloyl esterase) B 전구체 293 nt	둘 다
NCU09680	<i>네우로스포라 크라사</i> 엑소글루카나아제 2 전구체 485 nt	둘 다
NCU00798	<i>네우로스포라 크라사</i> 예측 단백질 313 nt	AV

[0250]

NCU01595	네우로스포라 크라사 단백질 SOF1 446 nt	AV
NCU02240	엔도글루카나아제 II 323 nt 와 유사한 네우로스포라 크라사 가설단백질	AV
NCU02696	DEAD DEAH 박스 RNA 헬리카제(DEAD DEAH box RNA helicase) 1195 nt 와 유사한 네우로스포라 크라사 가설단백질	AV
NCU02855	네우로스포라 크라사 엔도 1 4 베타 자일라나아제 A 전구체 221 nt	AV
NCU04205	네우로스포라 크라사 예측 단백질 346 nt	AV
NCU04395	네우로스포라 크라사 엔도 1 6 베타 D 글루카나아제 전구체 481 nt	AV
NCU05121	네우로스포라 크라사 엔도글루카나아제 V 294 nt	AV
NCU05134	네우로스포라 크라사 가설단백질 124 nt	AV
NCU05852	네우로스포라 크라사 보존 가설단백질 254 nt	AV
NCU05974	세포 벽 글루카노실트랜스페라아제(glucanoyltransferase) Mwg1 365 nt 와 유사한 네우로스포라 크라사 가설단백질	AV
NCU08171	네우로스포라 크라사 예측 단백질 382 nt	AV
NCU08909	베타 1 3 글루카노실트랜스페라아제 543 nt 와 유사한 네우로스포라 크라사 가설단백질	AV
NCU08936	네우로스포라 크라사 클락 조절된 유전자(clock controlled gene) 15 412 nt	AV
NCU09046	네우로스포라 크라사 예측 단백질 187 nt	AV
NCU09764	네우로스포라 크라사 보존 가설단백질 406 nt	AV
NCU01651	네우로스포라 크라사 보존 가설단백질 783 nt	MIS
NCU02343	알파 L 아라비노푸라노시다아제(arabinofuranosidase) A 668 nt 와 유사한 네우로스포라 크라사 가설단백질	MIS
NCU04202	네우로스포라 크라사 뉴클레오시드 디포스페이트 키나아제 1 153 nt	MIS
NCU04870	아세틸 자일란 에스테라아제 313 nt 와 유사한 네우로스포라 크라사 가설단백질	MIS
NCU05159	네우로스포라 크라사 아세틸자일란 에스테라아제(acetylxyloxy esterase) 전구체 301 nt	MIS
NCU05751	네우로스포라 크라사 보존 가설단백질 242 nt	MIS
NCU06239	네우로스포라 크라사 보존 가설단백질 514 nt	MIS
NCU08785	네우로스포라 크라사 보존 가설단백질 291 nt	MIS
NCU09267	네우로스포라 크라사 보존 가설단백질 1048 nt	MIS
NCU09708	네우로스포라 크라사 보존 가설단백질 465 nt	MIS
NCU09775	알파 L 아라비노푸라노시다아제 343 nt 와 유사한 네우로스포라 크라사 가설단백질	MIS
NCU09923	베타 자이로시다아제 775 nt 와 유사한 네우로스포라 크라사 가설단백질	MIS

[0251]

아노테이션(ANNOTATION) – Broad Institute 에 의해 생성(웹페이지 broad.mit.edu/annotation/genome/neurospora/Home.html) ;

배양물(CULTURE) – 특정 단백질에 대해서 발견된 펩티드에서의 배양물 ;

둘 다(BOTH) – 아비셀 및 미스칸터스 배양 여과액 둘 다에서 발견된 펩티드 ;

AV – 아비셀 배양 여과액에서만 발견된 펩티드 ;

MIS – 미스칸터스 배양 여과액에서만 발견된 펩티드.

[0252]

실시예 3

[0253]

전사체/세크레토믹 데이터세트(Transcriptome/Secretome Datasets)의 오버랩(Overlap)에서 확인된 유전자에서의 결실(Deletions)을 함유하는 스트레인에서 세포외 단백질(Extracellular Proteins) 및 셀룰라아제 활성도(Cellulase Activity)의 특성

[0254]

배양물에서 성장한 미스칸터스 및 아비셀 둘 다에서 관찰된 22 가지의 세포외 단백질 중에서, 이들 세포외 단백질 중 16 가지를 코딩하는 유전자에서 결실을 함유하는 동종핵형(homokaryotic) 스트레인은 일반적으로 입수할 수 있다(Dunlap 등의, 2007). 이들 16 가지의 결실 스트레인은 *N. 크라사*에서 식물세포벽 또는 셀룰로스 분해에서 이들의 영향에 대해서 이전에 특징지어지지 않았다. 16 가지의 결실 스트레인은 바람직한 탄소원(carbon

source)으로서 수크로스 또는 아비셀을 함유하는 배지에서 모두 성장하였다. 모든 스트레인은 수크로스에서 야생형 성장 형질형(wild type growth phenotype)을 나타내었다. 아비셀을 함유하는 배지에서, 16 가지 결실 스트레인의 대부분의 성장을 7 일 동안 모니터하였다. 7 일 후에, 총 분비된 단백질(secreted protein), 엔도글루카나아제 활성도(endoglucanase activity), β -글루코시다아제 활성도 및 배양 여과액(culture filtrates)의 총 아비셀라아제 활성도(aggregate Avicelase activity)를 측정하였고, 유도된 모든 돌연변이체(mutants)로부터의 야생형 스트레인과 비교하였다(도 6). SDS-PAGE 는 분비된 단백질의 상대 빈도(relative abundance)를 조사하기 위해 농축시키지 않은 배양 상청액(unconcentrated culture supernatants)에서 또한 실시하였다.

[0255] 두 가지의 예측 엑소글루카나아제(*cbh-1* ; NCU07340 및 *gh6-2* ; NCU09680) 및 예측 β -글루코시다아제(*gh3-4* ; NCU04952)의 결실을 포함하는 스트레인에 대해서 아비셀에서의 성장 결핍(growth deficiencies)이 있었다. *cbh-1* 돌연변이체는 가장 심각하다 ; 7 일 후에 대부분의 아비셀은 남아있는 반면에, 야생형 스트레인의 모든 아비셀은 이 시점에 분해되었다. 16 개의 결실 스트레인 중 10 개에 대해서, 분비된 단백질의 SDS-PAGE 분석은 단일 밴드가 사라진 곳에서 변화된 세포외 단백질 프로파일(altered extracellular protein profile)을 나타내었고, 따라서 예측 유전자에 대한 특정 단백질 밴드(protein band)의 배치(assignment)를 가능하게 한다(도 6A, 박스(boxes) ; 도 7). 이들은 NCU00762(*gh5-1*), NCU04952(*gh3-4*), NCU05057(*gh7-1*), NCU05137, NCU05924(*gh10-1*), NCU05955, NCU07190(*gh6-3*), NCU07326, NCU07340(*cbh-1*) 및 NCU09680(*gh6-2*)을 포함한다.

[0256] 결실 스트레인의 대부분에 대해서, 전체 분비된 단백질, 엔도글루카나아제, β -글루코시다아제 및 배양 상청액의 아비셀라아제 활성도(Avicelase activities)는 야생형(wild type)과 유사하였다(도 6B, C 및 표 9).

[표 9]

결실 스트레인(deletion strains)의 효소 활성도

유전자 이름	아비셀에서 성장	[분비된 단백질] (% of WT)	Azo-CMCase (% of WT)	Bgl (% of WT)	[CB] (mM)	[GLC] (mM)
NCU00762	***	113 ± 8	33 ± 2	102 ± 2	0.9 ± 0.0	2.6 ± 0.1
NCU01050	***	98 ± 12	92 ± 8	88 ± 5	0.8 ± 0.2	2.9 ± 0.3
NCU04952	***	146 ± 6	124 ± 5	1 ± 0.3	2.24 ± 0.2	0.6 ± 0.0
NCU05057	***	143 ± 10	98 ± 3	100 ± 10	1.7 ± 0.1	3.6 ± 0.1
NCU05137	***	154 ± 12	156 ± 10	178 ± 3	1.0 ± 0.0	3.8 ± 0.1
NCU05924	***	108 ± 3	108 ± 5	101 ± 4	1.1 ± 0.1	2.6 ± 0.2
NCU05955	***	92 ± 10	94 ± 8	98 ± 7	0.9 ± 0.1	2.3 ± 0.1
NCU07190	***	111 ± 7	136 ± 6	92 ± 1	1.1 ± 0.0	2.6 ± 0.0
NCU07326	***	105 ± 4	114 ± 17	85 ± 11	1.0 ± 0.0	2.3 ± 0.0
NCU07340	*	41 ± 2.2	43 ± 9	56 ± 9	0.1 ± 0.0	0.5 ± 0.1
NCU07898	***	84 ± 7	86 ± 1.5	59 ± 15	0.5 ± 0.3	2.3 ± 0.5
NCU08189	***	83 ± 12	80 ± 8	69 ± 15	0.5 ± 0.1	2.3 ± 0.4
NCU08398	***	95 ± 11	107 ± 7	97 ± 3	0.6 ± 0.1	1.8 ± 0.0
NCU08760	***	115 ± 3	126 ± 6	115 ± 8	0.9 ± 0.1	2.6 ± 0.1
NCU09175	***	96 ± 7	115 ± 0	101 ± 8	0.7 ± 0.0	1.9 ± 0.1
NCU09680	**	118 ± 7	165 ± 7	150 ± 1	0.23 ± 0.1	1.7 ± 0.1
WT	***	100 ± 7	100 ± 12	100 ± 6	0.97 ± 0.0	2.4 ± 0.1

[0258]

[0259] 이러한 경향(trend)으로부터의 편차(Deviations)는 $\Delta gh5-1$ (NCU00762), $\Delta gh3-4$ (NCU04952), $\Delta NCU05137$, $\Delta cbh-1$ (NCU07340), 및 $\Delta gh6-2$ (NCU09680) 돌연변이체(mutants)에서 볼 수 있다. $\Delta gh5-1$ (NCU00762), $\Delta gh3-4$ (NCU04952) 및 $\Delta cbh-1$ (NCU07340)에서, 아비셀라아제(Avicelase), 엔도글루카나아제(endoglucanase) 또는 Δ -글루코시다아제 활성도는 야생형의 활성도에 상응하는 것 보다 낮았다. 특히, NCU04952의 결실(deletion)은, 아비셀라아제 효소 검정(Avicelase enzyme assays)에서 PNPGase 활성도 및 높은 수준의 셀로비오스 및 낮은 수준의 글루코스에 의해 입증된 바와 같이, 배양 상청액으로부터의 모든 β -글루코시다아제 활성도를 제거하였다(도 6B 및 C). 낮은 엔도글루카나아제 활성도에도 불구하고, $\Delta gh5-1$ (NCU00762)로부터의 배양 여과액(culture filtrate)은 야생형 스트레인에 관하여 아비셀라아제 활성도에서 현저한 결핍을 나타내지 않았다(도 6C). 예상대로, $cbh-1$ (NCU07340)에서의 돌연변이는 좋지 못한 성장 때문에, 엔도글루카나아제 및 아비셀라아제 활성도가 더 낮아졌다. CBH(II)-유사 단백질($gh6-2$)을 코딩하는 NCU09680의 결실을 포함하는 스트레인은 $\Delta cbh-1$ 돌연변이체(mutant)에서 관찰된 바와 같이, 감소된 셀로비오스 축적(accumulation)을 또한 나타내었다(도 6C).

[0260] 세 가지의 스트레인에서 돌연변이로 인하여 특히 CBH(I)(도 6A); $gh3-4$ (NCU04952), $gh7-1$ (NCU05057) 및 가설 단백질 유전자(hypothetical protein gene) (NCU05137)의 수준이 증가되었다. 분비된 단백질의 수준이 증가될 뿐 아니라, $\Delta NCU05137$ 돌연변이체는 증가된 엔도글루카나아제, β -글루코시다아제 및 아비셀라아제 활성도가 증가됨을 나타내었다(도 6B 및 C). NCU05137은 그 밖의 섬유소 분해성 균류(cellulolytic fungi)을 포함하는 수많은 곰팡이 자낭기(filamentous ascomycete fungi)의 계통에서 매우 잘 보존되어 있지만 *T. 레세이*(*T. reesei*)에는 현저히 오솔로그(ortholog)를 갖지 않았다(도 2). 증가된 분비, 단백질 안정성 또는 그렇지 않으면 $cbh-1$ 의 발현에서의 증가를 결과로 나타내는 피드백으로 인하여 $\Delta gh3-4$, $\Delta gh7-1$, 및 $\Delta NCU05137$ 에서 관찰된 CBH(I) 수준이 증가될 수 있다. 이러한 가능성의 구별을 위해서, 정량적 RT-PCR에 의한 검정에 의하면, $\Delta NCU05137$ 및 $\Delta gh3-4$ (NCU04952)에 의해 생산된 세포외 단백질(extracellular proteins)의 프로파일(profile)은 $cbh-1$ (NCU07340) 및 $gh6-2$ (CBH(II); NCU09680)의 유전자 발현 수준과 비교하였다(도 8). 스트레인 $\Delta NCU05137$ 및 $\Delta gh3-4$ 은 아비셀-성장 배양물에서 이미 2일에 CBH(I) 단백질의 수준이 더 높아졌다. 둘 다의 유전자가 성장 2일 후에 야생형 및 $\Delta NCU05137$ 및 $\Delta gh3-4$ 돌연변이체에서 높은 발현 수준을 나타냄을, 아비셀-성장 배양물로부터의 $cbh-1$ 및 $gh6-2$ 의 정량적 RT-PCR에서 나타내었다. 그러나, 이러한 유전자 모두의 발현이 야생형 스트레인에서 3일째에 현저하게 감소되었음에도 불구하고, $cbh-1$ 및 $gh6-2$ 둘 다에서 발현 수준은 $\Delta NCU05137$ 돌연변이체에서 증가되었고, $\Delta gh3-4$ 에 야생형보다 덜 감소되었다(도 8). $\Delta NCU05137$ 및 $\Delta gh3-4$ 돌연변이체에서의 $cbh-1$ and $gh6-2$ 유전자의 지속적인 발현(Sustained expression)은 CBH(I) 및 CBH(II) 단백질 수준에서 관찰된 증가에 대한 책임이 있을 수 있다.

실시예 4

[0261] 전사체(Transcriptome) 및 세크레톰(Secretome) 연구에 대한 물질 및 방법

[0262] 스트레인

[0263] Fungal Genetics Stock Center로부터 모든 *네우로스포라 크라사* 스트레인을 입수하였다(FGSC; 웹사이트 fgsc.net)(Supplemental Data, Dataset SI, page 1 in Tian 등의, 2009). 유전자 결실 스트레인은 N. 크라사 기능 유전학 프로젝트(functional genomics project)로부터 입수하였다(Dunlap 등의, 2007). *트리코데르마 레세이*(*Trichoderma reesei*) QM9414은 Dr. Monika Schmoll로부터의 선물이다(비엔나 기술공과대학). 2%(w/v) 탄 소원을 갖는 Vogel's 염에서 스트레인을 성장시켰다[미스칸투스, 수크로스 또는 아비셀(Sigma)]. *미스칸투스 x 기간테우스*(*Miscanthus x giganteus*)[-0.1 mm에 대한 밀드 스템(milled stem)]은 일리노이노 대학으로부터의 선물이다.

[0264] 효소 활성도 측정

[0265] Bio-Rad DC 단백질 분석 키트[Bio-Rad DC Protein Assay kit(Bio-Rad)]를 사용하여 전체 세포의 단백질 함량을 측정하였다. 배양 상청액에서 엔도글루카나아제 활성도를 azo-CMC 키트(Megazyme SCMCL)로 측정하였다. 10 분

동안 40 °C 에서, pH 5.0 의 50 mM 아세트산 나트륨 완충용액에서 500 μM 4-니트로페닐 β-D-글루코피라노시드(4-nitrophenyl P-D-glucopyranoside)와 10 배 희석한 배양 상청액(10-fold diluted culture supernatant)을 혼합하여 β-글리코시다아제 활성도를 측정하였다. 반응물을 5 % w/v 탄산나트륨으로 킨칭하였고 quenched), 400 nm 에서 흡광도(absorbance)측정하였다. 40 °C 에서 50 mM 아세트산나트륨 및 5 mg/mL 의 아비셀과 2 배 희석한 배양 상청액을 혼합하여 아비셀라아제 활성도를 측정하였다. 상청액(Supernatants)을, 글루코스 산화효소/페록시다아제(glucose oxidase/peroxidase)로 결합된 효소 검정(coupled enzyme assay)을 사용하여 글루코스 함량(glucose content)에 대해 분석하였다. 50 μL 의 아비셀라아제 반응물(avicelase reaction)을, pH 5.0 의 100 mM 의 아세트산 나트륨, 10 U/mL 의 겨자무과산화효소(horseradish peroxidase), 10 U/mL 의 글루코스 산화효소 및 1 mM o-디아니시딘(o-dianisidine)을 함유하는 150 μL 의 글루코스 검출시약(glucose detection reagent)에 이동시켰다. 30 분 후에, 540 nm 에서 흡수(absorption)를 측정하였다. 셀로비오스 농도는 스포로트리쿰 서모필(Sporotrichum thermophil)로부터의 셀로비오스 탈수소효소(cellobiose dehydrogenase, CDH)와 함께 결합된 효소 검정(coupled enzyme assay)을 사용하여 측정하였다. CDH 는 이전의 보고와 유사한 S. 서모필로부터 분리하였다(isolate)(Canevascini 1988). 50 μL 의 아비셀라아제 반응물을, pH 5.0 의 125 mM 아세트산 나트륨, 250 μM 디클로로페놀 인도페놀(dichlorophenol indophenol) 및 0.03 mg/mL CDH 을 함유하는 250 μL 의 셀로비오스 검출시약(cellobiose detection reagent)에 이동시켰다. 10 분 후에, 530 nm 에서 흡수(absorption)를 측정하였다.

[0266] RNA 분리, 마이크로어레이 분석(Microarray Analysis) 및 신호 펩티드 예측

[0267] 균사체(Mycelia)를 여과 및 액체질소에서 냉동시킨 플래시(flash frozen)로 채취하였다(harvested). 전체 RNA 를 트리졸(trizol)을 사용하여 분리하였다(Tian 등의, 2007 ; Kasuga 등의., 2005). 마이크로어레이 혼성(Microarray hybridization) 및 데이터 분석을 상기한 바와 같이 하였다(Tian 등의, 2007). 정규화된 발현값을, 각각의 실험 시점에서 각각의 유전자에 대한 상대적인 유전자 발현 수준(relative gene expression levels) 및 믿을 수 있는 인터벌(credible interval)을 추론할 수 있는 BAGEL[유전자 발현 수준의 베이저안 분석(Bayesian analysis of gene expression levels)]을 사용하여 분석하였다(Townsend 2004; Townsend and Hartl 2002). 신호펩티드는 신호P3 프로그램(signalP3 program)을 갖는 각각의 예측 단백질의 N-터미널(terminal) 70 개의 아미노산 영역을 사용하여 예측할 수 있다(웹페이지 cbs.dtu.dk/services/SignalP-3.0/). 원본 프로파일링 데이터를 (웹페이지 yale.edu/townsend/Links/ffdatabase/)에서 입수하였다.

[0268] 단백질 겔 전기이동(Protein Gel Electrophoresis)

[0269] 다른 방법으로 언급한 경우를 제외하고, 기준 4-15 % 트리스-HCl 폴리아크릴아미드 겔(Criterion 4-15% Tris-HCl polyacrylamide gels) 내로 로딩하기 전에, 농축시키지 않은 배양 상청액을 5x SDS 로딩 염료(loading dye)로 처리하고, 5 분 동안 끓였다(boiled). 쿠마시 염료(Coomassie dye)를 염색(staining)에 사용하였다.

[0270] 세크레톰 분석에 대한 트립신 분해 펩티드(Tryptic Peptides)의 제조

[0271] 배양 상청액을 10 kDa MWCO PES 스핀 농축기(spin concentrators)로 농축시켰다. 셀룰로스 결합 단백질을 인산 스웰런 셀룰로스(PASC)의 첨가에 의해 배양 상청액으로부터 분리하였다. 10 mg/mL PASC 의 5 mL 의 현탁액(suspension)을 10 mL 의 배양 상청액에 첨가하였다. 5 분 동안 4 °C 에서 배양한 후에, 혼합물을 원심분리하고 펠렛된(pelleted) PASC 를 pH 5.0 의 100 mM 아세트산나트륨의 20 펠렛 부피(pellet volumes)로 세척하였다. PASC 로 처리한 후에 상청액을 결합하지 않은 분획으로 저장하고 농축시켰다. 36 mg 의 요소, 5 μL 의 1 M 트리스 PH 8.5 및 5 μL 의 100 mM DTT 를 100 μL 의 농축된 배양 상청액 또는 단백질-결합 PASC 에 첨가하였고, 혼합물을 1 시간 동안 60 °C 에서 가열하였다. 700 μL 의 25 mM 중탄산암모늄 및 140 μL 의 메탄올을 상기 용액에 첨가한 다음에 pH 5.0 의 50 mM 아세트산 나트륨에 용해시킨 50 μL 의 100 μg/mL 트립신으로 처리하였다. PASC 결합 단백질에 대해서, PASC 를 가열한 후에 원심분리하여 제거한 다음에, 상청액을 트립신으로 처리하였다. 트립신을 37 °C 에서 밤새 반응시키도록 두었다. 분해(digestion)시킨 후에, speedvac 으로 부피(volume)를 줄였고, 3 번 MilliQ 물로 세척하였다. 샘플(sample)에서 잔류염(Residual salt)을 제조자의 지시(manufacturer's instructions)에 따라 OMIX 마이크로추출법 피펫팁(OMIX microextraction pipette tips)을 사용하여 제거하였다.

[0272] 트립신 분해 펩티드(Tryptic Peptides)의 액체 크로마토그래피

[0273] 트립신-분해 단백질을 고성능 액체크로마토그래피(ultraperformance liquid chromatography, UPLC)가 인-라인(in-line)으로 연결된 이중질량분석기(tandem mass spectrometer)를 사용하여 분석하였다. C18 트래핑(trapping)(180 μm x 20 mm) 및 분석(100 μm x 100 mm) 컬럼 및 10 μL 샘플 루프가 갖추어진 nanoAcquity UPLC(미국 매사추세츠 밀퍼드에 위치한 Waters)을 사용하여 펩티드를 분리하였다. 용매 A 를 0.1 % 의 포름산 /99.9 % 의 물이고, 용매 B 는 0.1 % 의 포름산/99.9 % 의 아세토니트릴(v/v)이다. 셉타 캡스(septa caps)로 밀봉된 0.3 mL 의 폴리프로필렌 스냅-톱 유리병(snap-top vials)에 포함된 샘플 용액은 분석 전에 nanoAcquity 오토샘플러(autosampler) 내에 로딩하였다. 샘플 주입[2 μL , 부분적인 루프(partial loop)] 후에, 트래핑(trapping)을 3 μL /분의 유량(flow rate)에서 100 % A 를 가지고 5 분 동안 실시하였다. 샘플 사이에 교차-오염(cross-contamination)을 방지하기 위해 주입 후에 각각의 용액 A 및 B 의 750 μL 로 주입 바늘(injection needle)을 세척하였다. 용출 프로그램은 500 nL/분의 유량(flow rate)으로, 55 분 이상 동안 25 % 에서 30 % B 까지의 선형 구배(linear gradient), 20 분 이상 40 % B 까지의 선형구배, 0.33 분 이상 95 % B 까지의 선형 구배, 11.67 분 동안의 95 % 에서의 등용리 조건(isocratic conditions), 0.33 분 상의 1 % B 까지의 선형 구배 및 11.67 분 동안의 1% B 에서의 등용리 조건으로 구성되었다. 분석 컬럼 및 샘플 컴파트먼트(sample compartment)를 각각 35 $^{\circ}\text{C}$ 및 8 $^{\circ}\text{C}$ 에서 유지시켰다.

[0274] 질량 분석법(Mass Spectrometry)

[0275] 컬럼을 사중극 비행 시간 분석기(quadrupole time-of-flight mass spectrometer)(Q-ToF Premier, Waters)의 나노 유속 이온원(nanoflow ion source)에 설치한 NanoEase 나노일렉트로 스프레이 이온화(nanoESI) 에미터[NanoEase nanoelectro spray ionization (nanoESI) emitter]에 컬럼을 연결하였다. nanoESI 소스 파라미터(source parameters)는 다음과 같다 ; nanoESI 모세관 전압(capillary voltage) 2.3 kV, 분무 가스(nebulizing gas)(질소) 압력 0.15 mbar, 샘플 콘 전압(sample cone voltage) 30 V, 추출액 콘 전압(extraction cone voltage) 5 V, 이온 가이드 전압(ion guide voltage) 3 V, 및 소스 블록 온도(source block temperature) 80 $^{\circ}\text{C}$. 어떠한 콘 가스(cone gas)도 사용하지 않았다. 충돌 셀(collision cell)은 8 x 10⁻³ mbar 의 압력으로 아르곤 가스를 함유한다. ToF 분석기(analyzer)는 "V" 모드(mode)에서 작동하였다. 이러한 조건 하에서, 질량 분해능1(mass resolving power)의 1.0 x 10⁴(m/z = 771 에서 측정)은 일상적으로 달성하였고, 이는 이러한 연구에서 측정된 1가 및 다가 펩티드 이온(singly and multiply charged peptide ions)의 동위 원소 분포(isotopic distributions)를 해결하기(resolve) 충분하다. 따라서, 이온의 질량 및 전하(charge)는 독립적으로 결정될 수 있고, 즉 이온 전하는 m/z 스펙트럼에서 인접한 동위원소 피크(adjacent isotope peaks) 사이의 상호간의 간격(reciprocal of the spacing)으로부터 결정된다. 포름산 나트륨(sodium formate)의 용액을 사용하여 외부의 질량 검정(External mass calibration)은 분석 전에 즉시 실시하였다. 조사 스캔(Survey scans)을, 0.95s 스캔 인테그레이션(scan integration) 및 0.05s 인터스캔 지연(interscan delay)을 사용한 범위 m/z = 450-1800 상에서 양성 이온 모드에서 획득하였다. 데이터-의존적인 모드에서, 35 카운트/초(cps)의 강도 한계점(intensity threshold)을 초과한 5 까지의 전구이온(precursor ions)은 이중 질량 분석(MS/MS)에 대한 각각의 조사 스캔(survey scan)으로부터 선택하였다. 리얼-타임 데이소토피ング(Real-time deisotoping) 및 전하 상태 인식(charge state recognition)은 MS/MS 에 대한 2+, 3+, 4+, 5+ 및 6+ 전하 상태 전구이온(charge state precursor ions)을 선택하기 위해 사용되었다. 충돌적인 활성화된 해리(collisionally activated dissociation, CAD)에 대한 충돌 에너지(Collision energies)는 주어진 전구이온(precursor ion)의 질량 및 전하 상태를 기반으로 자동적으로 선택된다. MS/MS 스펙트럼은 0.95s 스캔 인테그레이션(scan integration) 및 0.05s 인터스캔 지연(interscan delay)을 사용한 범위 m/z = 50-2500 상에서 획득하였다. 이온을, 최대 3 초 동안 누적(cumulative) MS/MS 스펙트럼에서 30,000 cps 의 최소한의 총 이온전류(ion current)(TIC)를 성취하기 위해 해체하였다(fragment). 불필요한 MS/MS 측정의 발생을 피하기 위해서, 리얼 타임 배제(real time exclusion)는 180 초 동안의 + 0.25 m/z 단위의 배제 너비(exclusion width) 상에서 사전에 분석한 전구이온의 재-선택(re-selection)을 불가능하게 하기 위해 사용되었다.

[0276] 질량 분석 데이터 분석

[0277] 트립신-분해 단백질의 LC-MS/MS 분석 결과의 데이터는 ProteinLynx Global Server software (version 2.3, Waters)을 사용하여 처리하였고, 이는 질량 스펙트럼 및 MS/MS 스펙트럼의 배경 차감[한계점(threshold) 35% 및 5차 다항식(fifth order polynomial), 스무딩(smoothing)[Savitzky-Golay 2 10 타임즈(times), 세 채널 상에서(over three channels)] 및 센트로딩(centroiding)[반 높이 네 채널(half height four channels)에서의 각각의 피크(peak)의 톱(top) 80 % 및 최소 피크 너비(minimum peak width)]에서 수행하였다. 가공 데이터(processed data)는 하기의 기준(criteria)을 사용하여 *N. 크라사* 데이터베이스와 비교하여 검색하였다 ; 5 까지의 손실된 분열(missed cleavages)를 갖는 트립신 분해 단편(tryptic fragments), 전구이온 질량 허용오차(precursor ion mass tolerance) 500 ppm, 단편 이온 질량 오차 0.1 Da, 그리고 하기의 가변적인 번역후의 변형물(post-translational modifications) : N-말단의 카르바밀화(carbamylation) 및 Lys 잔기(side chains), Met 산화 및 Ser/Thr 탈수(dehydration). 동일한 시리즈로부터 적어도 세 개의 연이은 단편 이온의 식별, 즉 b 또는 y-타입 단편 이온은 MS/MS 스펙트럼에 대한 펩티드의 배치(assignment)를 요구한다. MS/MS 스펙트럼은 상기 펩티드를 유일하게 확인하기 위한 단편 이온의 존재를 확인하기 위해 매뉴얼적으로(manually) 점검하였다.

[0278] 정량적인(Quantitative) RP-PCR

[0279] RT-PCR 을 Qiagen 로부터의 시약으로 AB17300 에서 수행하였다[SYBR-green RT- PCR kit (Cat No. 204243)]. CBHI(NCU07340)에 대한 프라이머(primers)는 포워드(forward) 5'-ATCTGGGAAGCGAACAAG-3'(SEQ ID NO: 16) 및 리버스(reverse) 5'-TAGCGGTCGTCGGAATAG-3'(SEQ ID NO: 17)이다. CBHII(NCU09680)에 대한 프라이머는 포워드 5'-CCCATCACCCTACTACC- 3'(SEQ ID NO: 18) 및 리버스 5'-CCAGCCCTGAACACCAAG-3'(SEQ ID NO: 19)이다. 정규화(normalization)에 대한 대조군(control)으로서 액틴(Actin)을 사용하였다. 액틴에 대한 프라이머는 포워드 5'-TGA TCT TAC CGA CTA CCT-3'(SEQ ID NO: 20) 및 리버스 5'-CAG AGC TTC TCC TTG ATG -3'(SEQ ID NO: 21)이다. 정량적인 RT-PCR 은 Dementhon 등의(2006)에 따라 수행하였다.

실시예 5

[0280] 전사체 및 세크레토믹 연구의 논의

[0281] 식물 바이오매스의 분해는 많은 상이한 효소의 활성도의 생산을 요구하고, 이는 이용할 수 있는 식물 물질(available plant material)의 유형(type) 및 복잡성(complexity)에 의해 조절된다(도 9)(Bouws 등의, 2008). 섬유소 분해 균류(cellulolytic fungus)에 의한 식물 세포 벽 분해(plant cell wall degradation)의 첫 번째 체계적인 분석(systematic analyses)은 본원에 기재되어 있고, 이는 전사체(transcriptome), 세크레토믹(secretome) 및 돌연변이체(mutant) 분석을 포함한다. 프로파일링 데이터(Profiling data)는, *미스칸투스* 또는 아비셀에서의 성장에 의한 도전을 할 때, *N. 크라사*는 세포외(extracellular) 및 세포내(intracellular) 단백질의 숙주(host)를 대등하게 발현시킴을 나타내었다(도 9). 셀룰로스 기질(cellulosic substrates)에서 성장하는 동안에 가장 많이 발현된 유전자의 대부분은 식물 세포 벽 다당류의 물질대사(metabolism)에 포함되는 것으로 예측된 단백질을 코드하고, 이들 중의 다수는 MS 분석에 의해 확인되었다. 곰팡이(filamentous fungi)의 계통 비교는 *N. 크라사*와 밀접하게 덩-분해 종(dung-degrading species), *T. 레세이*(Martinez 등의., 2008)에서 10 부터 *포도스포라 에스리나*(Espagne 등의., 2008)에서 60 까지의 많은 수의 예측된 셀룰라아제와 많은 수의 글리코실 가수분해 효소(glycosyl hydrolases)(~200)를 나타낸다. 백색 부유 담자균(white rot basidiomycete fungus), *파네로케이트 크리소스포리움*(*Phanerochaete chrysosporium*)(Wymelenberg 등의, 2009)에서 최근 전사체/세크레토믹 연구(transcriptome/secretome study) 및 이들의 결과 사이의 비교는, 둘 다의 종(species)이 순수한 셀룰로스(pure cellulose)에서 성장하였을 때, 조절된 유전자(regulated genes)(18 유전자) 및 분비된 단백질(2 단백질)에서 약간의 중복(overlap)을 나타내었다. 이러한 데이터는 식물 세포벽 분해를 위한 상이한 유전자 세트(gene sets)를 사용하는 것을 제안한다. 그러나, 연구 모두가 공동으로 갖는 하나의 측면은 셀룰로스 분해(cellulose degradation)와 관련된 많은 수의 특징지어지지 않은(uncharacterized) 유전자/단백질이다. *P. 크리소스포리움*(*P. chrysosporium*)을 포함하는 그 밖의 섬유소분해성 균류(cellulolytic fungi)는 *N. 크라사*와 함께 손쉽게 응용할 수 있는 유전적이고 분자로 구성된 기술(genetic and molecular tools)을 갖지 않는다. *N. 크라사*와 함께 이용가능한 기능적인 유전적인 기술(functional genomic tools)을 사용하여, 식물 세포벽 분해 효소 시스템의 기능 및 불필요한 반복(redundancy) 둘 다는 리그노셀룰로스 바이오매스(lignocellulose biomass)로부터 액체 연료(liquid fuels)의 산업적인 생산을 위한 최적의 효소 혼합물

(optimal enzyme mixtures)을 생산하기 위해 연구할 수 있다.

- [0282] 이러한 연구에서, *N. 크라사*에서의 셀로비오하이드로라아제(cellobiohydrolase)(I)(CBHI)가 아비셀 또는 미스칸터스에서 성장하는 동안에 가장 많이 생산된 세포외 단백질(extracellular protein)임이 밝혀졌고, 이러한 유전자의 결실은 셀룰로스 기질(cellulosic substrates)에서 가장 심각한 성장 결핍증(deficiencies)을 일으켰다. 이러한 결과는 *T. 레세이*에서 보고된 결과와 유사하다(Suominen 등, 1993, Seiboth 등, 1997). 셀로비오하이드로라아제(II)의 결실은 셀룰로스에 기질(cellulosic substrates)에서의 성장 결핍증을 또한 유발하지만 CBH(I) 보다는 더 적은 정도이고, 이는 *N. 크라사*에서의 엑소글루카나아제 활성도(exoglucanase activity)는 CBH(I)으로부터가 대부분이고, 셀룰라아제 및 그 밖의 CBHs는 CBH(I)의 손실(loss)에 대해서 보상하지 않는 점을 나타내었다. 여기에서, 셀룰로스에 기질에서 성장하는 동안 가장 많이 생산된 세 가지 엔도글루카나아제는 NCU05057, NCU00762, 및 NCU07190 에 의해 코드됨을 나타내었다. 이러한 단백질은 각각 엔도글루카나아제 EG1, EG2 및 EG6 와 상동(homology)이다. 분비된 단백질 수준 및 엔도글루카나아제 활성도에서의 상이한 점은 관찰되지 않았지만, 이러한 유전자의 결실은 아비셀에서의 성장에 영향을 주지 않는다. 예상외로, Δ NCU05057 스트레인에서, 세포의 단백질 수준은 특히 CBH(I) 에서 보다 높았고, 이는 결정질 셀룰로스(crystalline cellulose)에서 야생형 성장 표현형(wild-type growth phenotype)을 유지시키기 위해, 돌연변이체(mutant)는 그 밖의 셀룰라아제의 생산을 증가시키기 위해 집중되거나(force), 또는 NCU05057 촉매의 생산은 셀룰라제 생산을 억제할 수 있음을 나타낸다. *N. 크라사*에서 하나의 엔도글루카나아제는 결정질 셀룰로스에서의 성장을 필요로 하지 않고 상이한 엔도글루카나아제는 중복된 기질 특이성(overlapping substrate specificities)을 갖음을 결론으로 내렸다.
- [0283] 글리코시드 히드로라아제 패밀리(family) 61 효소는 *T. 레세이*와 비교하여 *N. 크라사*에서 크게 확대되었다(Martinez 등, 2008). 이러한 효소는 저조하게 규정된 생물학적 기능(biological function)을 갖지만, 섬유소분해성 균류(cellulolytic fungi)에서의 이들의 일반적인 보존(general conservation) 및 풍부(abundance)는 식물 세포벽 물질대사에서 중요한 역할을 함을 암시한다. 여기에서, 14 가지의 GH61 효소 중의 10 가지에 대한 유전자는 *N. 크라사* 전사체(transcriptome)에서 확인되었고, 이는 이러한 효소는 셀룰로스에 바이오매스(cellulosic biomass)에서 성장하는 동안 사용됨을 암시한다. 실험된 4 가지 GH61 결실 스트레인은 분비된 단백질 수준, 엔도글루카나아제 및 전체 셀룰라아제 활성도에서 야생형과 비교하여 오직 작은 차이점을 나타내었다. 그러나, 유성교차(sexual crosses)을 통해 *N. 크라사*에서의 다양한 돌연변이를 포함하는 스트레인을 생산하기 위해 추가적인 GH61 돌연변이체 및 능력의 분석은 불필요한 중복(redundancy)을 다룰 것이고, 이러한 패밀리에서 기능적인 분석(functional analysis)을 더 신속히 처리할 수 있을 것이다.
- [0284] 예측된 셀룰라아제 유전자 뿐 아니라, 탄수화물에서 활성도를 갖는 것으로 예측된 헤미셀룰라아제, 탄수화물 에스테라아제(carbohydrate esterases), β -글루코시다아제, β -자일로시다아제(xylosidases) 및 그 밖의 단백질은 *미스칸터스* 및 아비셀 둘 다로부터 *N. 크라사* 전사체에서 확인되었다. 아비셀이 헤미셀룰로스 성분을 함유하지 않다는 사실은, 셀룰로스가 아마도 *N. 크라사*에서 식물 세포벽 분해 효소를 코딩하는 유전자의 주된 유도물질(primary inducer)임을 암시한다. 그러나, 몇몇의 헤미셀룰라아제 및 탄수화물 에스테라아제를 코딩하는 유전자는 *미스칸터스*에서 성장하는 동안만 발현되었다. *T. 레세이* 및 *아스페르길루스 니가(Aspergillus niger)*와 같은 그 밖의 섬유소분해성 균주에서 유사하게, 몇몇의 셀룰라아제 및 헤미셀룰라아제를 코딩하는(encoding) 유전자는 대등하게 조절된 반면에, 그 밖에는 별도로 조절되었다(Strieker 등, 2008). 예상대로, NCU05137 및 *gh3-4* 을 제외하고, 비-셀룰라아제 유전자의 삭제는 아비셀 또는 셀룰라아제 활성도에서의 성장에 작은 영향을 준다. 보다 많은 단백질을 분비하는 Δ NCU05137 스트레인은 보다 높은 셀룰라아제 활성도를 갖고, 이는 야생형보다 *cbh-I*(CBH(I)) 및 *gh6-2* (CBH(II))의 보다 높은 발현을 나타내었다. NCU05137 는 알려진 기능(known function)의 단백질과 상동적이지 않은 분비된 가설단백질(secreted hypothetical protein)을 코드하지만, 다른 섬유소분해성 균류에서 매우 잘 보존되었다(도 2 | E 값 0.0). NCU05137 은 보다 거리가 있는 상동성(more distant homologs)을 갖지만, 수많은 균종(bacterial species)에서 알려지지 않은 기능 또한 갖는다. NCU05137 의 단백질 산물은 셀룰라제 유전자 발현 *N. 크라사*의 유도와 관련된 신호전달 과정(signaling processes)에 개입할 수 있다(도 9). 유사하게, *gh3-4*(NCU04952)에서의 돌연변이는 CBH(I) 활성도를 또한 증가시킨다. 이러한 유전자의 결실은 PNPase 활성도를 완전하게 제거하고, 셀로비오스는, Δ gh3-4 배

양 여과액(culture filtrates)을 사용한 *생체외(in vitro)* 셀룰라아제 검정에서 측정되었다. 모든 데이터는, NCU04952 가 *N. 크라사*에서 주된 세포의 β -글루코시다아제를 코딩함을 제시한다. 이러한 데이터는 글루코스에 의한 셀룰라아제 생산의 분해대사물억제(catabolite repression)와 일치한다.

[0285] 셀룰로스 및 헤미셀룰로스의 세포외 분해(Extracellular degradation)로 세포 내로 순차적으로 수송되는 가용성 탄수화물(soluble carbohydrates)이 형성되었다(도 9). 이러한 연구에서, 투과효소(permeases)/운반체(transporters)를 코딩하는 10 가지 유전자는, *N. 크라사*가 *미스칸티스* 또는 아비셀에서 성장하였을 때 현저하게 증가된 발현을 나타낸 것으로 확인되었고, 이는 식물 세포벽 분해 산물(plant cell wall degradation products)이 세포 내로의 수송에 대한 이들의 관여를 암시한다. *생체 외*에서의 셀룰라아제 및 헤미셀룰라아제에 의한 주요한 분해 산물(degradation products)은 셀로비오스, 글루코스, 자일로비오스(xylobiose) 및 자일로스이다. 이러한 운반체의 몇몇은 기능적으로 불필요할 수 있거나, 올리고당(oligosaccharides)을 수송할 수 있다. 이러한 추정상의 운반체(putative transporters)의 기능을 보다 더 분석하였다(실시예 7-9 참조). *N. 크라사* 운반체의 이형 발현(heterologous expression)에 의한 올리고당을 수송할 수 있는 다운스트림 과정(downstream processing strains)의 스트레인의 구성(Construction)은 바이오매스 가수분해 산물(biomass hydrolysis products)의 산업상의 발효(industrial fermentation)를 개선시킬 수 있다. 이러한 운반체의 중 어떠한 것 또는 그들이 수송할 수 있는 어떠한 것도 어느 곰팡이(any filamentous fungi)에서의 분자 또는 기능적인 수준(molecular or functional level)으로 특징지을 수 없다.

[0286] *미스칸티스* 및 아비셀에서 성장하는 동안 증가된 발현수준을 나타내는 많은 유전자는 그 밖의 섬유소분해성 균주(other cellulolytic fungi)에서 보존된 알려지지 않은 기능의 단백질을 코딩한다. 16 가지의 스트레인의 표현형 만을 평가함으로써, 셀룰라아제 활성도에 현저한 영향을 주는 알려지지 않은 기능(unknown function)의 단백질을 코딩하는 유전자에서의 돌연변이체가 확인되었다. *N. 크라사*에서 게놈의 자원(genomic resources)의 잘 알려진 유전학(well-understood genetics) 및 효용성(availability)은, 이러한 단백질의 생물학적 기능(biological function)뿐만 아니라 셀룰라아제 및 헤미셀룰라아제 생산의 조절적인 측면(regulatory aspects)을 결정하고, 식물 세포벽의 분해에 포함되는 세포의 효소 사이의 불필요한 중복 및 시너지(synergies)를 분석하기 위한 이상적인 유기체의 모델(ideal model organism)을 만들 수 있고,

실시예 6

[0287] *미스칸티스*에서 성장하는 동안의 **상향조절된(Upregulate) 유전자의 돌연변이체의 스크리닝(Screening)**

[0288] 전사 프로파일링 실험(transcriptional profiling experiment)에서 확인된 추가적인 유전자를 분석하기 위해, 16 시간 동안 *미스칸티스*에서 성장한 *네우로스포라*에서 상향조절된 188 가지 유전자의 돌연변이체 표현형(phenotypes)을 분석하였다(실시예 1 참조). 각각의 유전자의 녹아웃 돌연변이체(knockout mutant)를 최소 Vogel's 배지에서 10-14 일 동안 성장시켰다. 분생자(Conidia)를 2 mL ddH₂O 와 함께 채취하였고 mL 당 10⁶ 분생자의 농도로 250 mL 의 플라스크에 100 mL 배지 내에서 접종하였다(inoculated). 상이한 탄소원 중의 하나를 각각의 플라스크에 첨가하였다 : 2 % 수크로스, 2 % 아비셀 또는 2 % *미스칸티스*(미국 캘리포니아 버클리의 캘리포니아 대학의 Calvin Laboratory로부터 입수한 1mm 입자). 배양물(Cultures)을 4 일 동안 250 rpm 으로 교반시키면서 25 °C 에서 성장시켰다.

[0289] 표 10 은 돌연변이체의 표현형을 리스트로 작성하였고, 이는 야생형과 비교하여 아비셀 또는 *미스칸티스*에서의 셀룰라아제 활성도 및 성장에서 중요한 차이점을 나타낸다. 아비셀 또는 *미스칸티스*에서의 성장은 "+" 점수 체계(scoring system)로 평가하였다. 야생형의 성장은 "++" 로 나타내었다. 배양 상청액에서 전체 단백질을 Bradford 검정으로 측정하였고(900 μ l Bradford dye 에 대한 100 μ l 상청액), 엔도글루카나아제 활성도는 Megazyme 의 엔도글루카나아제 활성도의 퍼센트로서 표 10 에 나타내었다. 전체 셀룰라아제 활성도는 실시예 4 에 나타낸 바와 같은 상청액에서 셀로비오스 수준의 발견에 의해 측정하였다. 결과를 야생형의 퍼센트로서 표 10 에 나타내었다.

[0290]

[표 10]

돌연변이체 스크리닝 데이터 (mutant screening data)

NUC#	FGSC#	브로드 아노테이션 (도메인)	Pfam*	상향조절 (Up- Regulation)	성장 (Avi, Mis)	%WT 브레드포드 (Avi)	%WT 브레드포드 (Mis)	%WT 엔도 (Avi)	%WT 엔도 (Mis)	%WT 셀로비오스 (Avi)	%WT 셀로비오스 (Mis)
NCU0013 0.2	FGSC 11823	베타- 글루코시 다아제 (GH1)	글리코실 가수분해 효소 1 (2.5e-196)	394.6	+++	203.24779 47	118.39879 72	152.2858578	129.35474 94	n/a	n/a
NCU0024 8.2	FGSC 12214	예측 단백질	중요한 히트(hit) 없음	9.74	+++	86.960132 89	86.044718 58	30.39187506	156.50501 44	93.05143 946	89.2769821 9
NCU0032 6.2	FGSC 15868	보존 가상 (SMP-30/ 글루코노 락토나아제)	SMP-30/ 글루코노 라코나아제 (Glucono laconase)/ LRE-유사 영역 (3.5e-82)	7.7	+++	33.028792 91	144.12104 86	39.91568458	227.03668 09	89.76672 415	79.0515463 9
NCU0076 2.2	FGSC 16747	엔도글루 카나아제-3 전구체 또는 (GH5, CB01)	셀룰라아제 (1.4e-69), 곰팡이 셀룰로오스 결합 도메인 (9.2e-14)	29.6	+++	104.35044 11	84.570569 44	26.92790756	39.268900 58	n/a	n/a
NCU0081 0.2	FGSC 11285	글리코실 가수분해 효소와 유사(GH2, 베타- 갈락토시 다아제)	글리코실 가수분해 효소 패밀리 2 (1.7e-145)	5.3	+++	163.80504 7	123.55647 57	161.2908993	159.49837 44	102.2745 211	91.7334566 4
NCU0089 0.2	FGSC 16749	베타- 마노시 다아제 (manosi dase)와 유사 (GH2)	글리코실 가수분해 효소 패밀리 2 (4.1e-06)	20.45	++	47.574178 03	101.59744 41	43.25546345	164.08197 18	n/a	n/a
NCU0332 8.2	FGSC 16589	보존 가상 (GH61)	글리코실 가수분해 효소 패밀리 61 (2.3e-10)	26.4	+++	100.17528 48	109.96672 48	142.6962073	167.00754 81	n/a	n/a
NCU0341 5.2	FGSC 12922	알데히드 탈수소 효소	알데히드 탈수소 효소 패밀리 (2.5e-267)	9.8	+++	104.22782 04	96.614353 73	96.4633125	63.455233 29	76.96643 943	103.127398 3
NCU0373 1.2	FGSC 18653	HAD 상과 (superfamily) 가수분해 효소와 유사	할로산 디할로게 나아제 (haloacid dehaloge nase)-유사 가수분해 효소 (9.2e-21)	2.7	+++	131.36911 28	110.58014 46	145.0235135	134.62799 5	230.1450 412	100.417237 5

[0291]

NCU0375 3.2	FGSC 16379	ccg-l (클락 조절된 유전자)	중요한 히트 없음	10.5	++,++	107.67928 92	111.34810 86	74.42402278	129.21967 77	n/a	n/a
NCU0419 7.2	FGSC 17499	보존 가상	중요한 히트 없음	5.04	++,++	103.08681 27	99.083054 14	108.9737808	89.861286 25	75.17285 531	96.0507505 4
NCU0424 9.2	FGSC 18628	가설유전자	중요한 히트 없음	5.3	++,++	93.296823 66	106.10121 67	79.0053469	84.161412 36	64.07989 522	100.124185
NCU0428 7.2	FGSC 14573	예측 단백질	중요한 히트 없음	4.7	++,++	115.51578 59	102.23610 65	125.5086234	127.92825 77	202.5161 29	183.867924 5
NCU0434 9.2	FGSC 18634	미토콘드리아 피루브산염 탈수소효소 키나아제와 유사	BCOHK_A dom3 (4.7e-78), 히스티딘 ATPase_c (6.9e-14)	2.9	++,++	87.877764 65	89.362051 96	71.41381803	145.24158 13	208.4329 349	101.299376 3
NCU0447 5.2	FGSC 15386	예측 단백질	중요한 히트 없음	76.7	+++,++	96.102053 52	122.20348 51	156.3643221	127.06766 92	n/a	n/a
NCU0499 7.2	FGSC 15623	자일라 나아제와 유사(GH10, CB01)	글리코실 가수분해효소 패밀리 10 (3.3e-148), 곰팡이 셀룰로스 결합 도메인 (2.1e-16)	25.6	++,++	105.35201 76	114.58401 84	123.3295466	231.69838 95	136.4189 483	102.540398 3
NCU0505 7.2	FGSC 13342	엔도글루카 나아제 EG-1 전구체 또는 (GH7)	글리코실 가수분해 효소 패밀리 7 (3.3e-189)	8.7	++,++	137.53165 63	95.692206 51	133.5226686	174.26793 56	182.0237 75	97.8133065 7
NCU0515 9.2	FGSC 13439	아세틸자일란 에스테라아제 전구체 또는 (큐티나아제, CBD1)	큐티나아제 (3.4e-110), 곰팡이 셀룰로스 결합 도메인 (7.4e-14)	34.8	+++,++	86.185438 71	39.516582 35	92.2873845	67.117794 49	n/a	n/a
NCU0549 3.2	FGSC 14625	예측 단백질	중요한 히트 없음	4.5	+,++	73.252660 13	104.41025 64	102.3841739	116.89545 93	70.37185 126	99.4283792 9
NCU0551 9.2	FGSC 19924	Tna 1 과 유사(MFA 운반체)	주요한 촉진 (Facilitator) 상과 (superfamily) (3.7e-40)	2.8	++,++	85.311913 21	101.06666 67	118.8447721	87.777191 13	51.68869 31	97.8750165 5

[0292]

NCU0575 1.2	FGSC 15757	보존 가상 (GDSL-유사 리파아제)	GDSL-유사 리파아제 /아실히 드롤라아제 (1.3e-11)	3.9	+,++	97.016482 37	111.40512 82	114.7202911	136.37803 59	87.71492 649	105.592058 3
NCU0577 0.2	FGSC 11532	페록시 다아제/ 카탈라아제 리파아제)	페록시 다아제 (9.4e-195)	11.9	++,++	109.86309 89	86.734120 29	69.1872525	146.21553 88	n/a	n/a
NCU0585 3	FGSC 13771	당 운반체	당 운반체	130.7	+	40.279246 87	24.412597 90	n/a	n/a	n/a	n/a
NCU0589 7.2	FGSC 13717	1-푸코스 투과효소 (fucose permease)와 유사 (MFS 운반체)	주요한 촉진 상과 (3.8e-16)	20.9	+,++	33.784641 42	34.727545 41	26.3266891	88.259541 98	n/a	n/a
NCU0593 2.2	FGSC 19952	예측 단백질	중요한 히트 없음	38.2	++,++	70.898264 28	76.871320 44	80.78910753	117.95968 23	58.07431 478	96.7108463
NCU0600 9.2	FGSC 14922	알도/케토 환원효소와 유사	알도/케토 환원효소 패밀리 (4.8e-63)	6.9	+,++	148.66337 26	74.067844 13	120.602266	99.480757 48	70.89513 625	97.0057324 1
NCU0649 0.2	FGSC 15539	보존 가상	중요한 히트 없음	13.8	+,++	77.461041 43	80.283526 77	76.95289207	79.383017 72	59.91109 168	99.9937138 5
NCU0734 0.2	FGSC 15630	엑소글루 카나아제-1 전구체 또는 CBH1(GH7)	글리코실 가수분해효소 패밀리 7 (1e-999), 곰팡이 셀룰로스 결합 도메인 (4.9e-18)	426.4	+,++	21.096345 51	95.219737 86	35.54661301	96.991344 96	93.62619 808	78.4490255 3
NCU0785 3.2	FGSC 19036	우리카아제 (Uricase)	우리카아제 (1.7e-119)	4.3	+++,,	n/a	n/a	120.9286562	168.23406 48	65.75994 56	99.1465917 7
NCU0799 7.2	FGSC 18273	예측 단백질	중요한 히트 없음	4.5	++,++	n/a	n/a	148.127436	98.119122 26	60.65548 063	93.7870427 1

[0293]

NCU0811 4.2	FGSC 17869	MFS 육탄당 운반체와 유사(MFS 운반체)	당(및 기타) 운반체 (5.1e-88), 주요한 촉진 상과 (3.8e-24)	6.7	+,++	81.692639 05	79.226240 54	85.18187239	92.974954 18	58.83068 556	93.1432252
NCU0874 4.2	FGSC 11387	예측 단백질, 가능한 TF [기본 영역 류신 지퍼 (basic region leucine zipper)]	중요한 히트 없음	2.3	++,++	n/a	n/a	168.8527368	110.76280 04	136.2451 567	97.4413419 7
NCU0874 6.2	FGSC 18358	보존 가상 (녹말 결합 도메인)	녹말 결합 도메인 (5.3e-54)	6	++,++	98.695046 24	79.114101 49	111.0713576	120.25045 82	447.2796 518	100.575366 7
NCU0876 0.2	FGSC 15664	예측 단백질 (CBD1)	곰팡이 셀룰로스 결합 도메인 (1.9e-11), 글리코실 가수분해효소 패밀리 61 (1.3e-9)	107.5	++,++	158.13953 49	86.179645 34	208.2590783	81.000137 38	97.32646 961	84.3425177 4
NCU0910 8.2	FGSC 19207	보존 가상	중요한 히트 없음	4.1	++,++	n/a	n/a	93.22148788	111.80773 25	60.89420 655	97.2451790 6
NCU0949 5.2	FGSC 12411	세트-6, 히스톤 메틸트랜스 페라아제	SET 도메인 (6.9e-5)	26.2	++,++	109.33001 11	122.53276 79	129.9223915	130.89710 13	152.7495 439	92.2521655 4
NCU0968 0.2	FGSC 15633	엑소글루 카나아제-2 전구체 또는 CBH2(GH6, CBD1)	글리코실 가수분해효소 패밀리 6 (1.1e-152), 곰팡이 셀룰로스 결합 도메인 (1.2e-13)	230.9	+,++	102.71317 83	95.200462 61	89.54680464	102.67893 94	94.61873 756	83.8766134 3
NCU1004 5.2	FGSC 18480	펙틴에스 테라아제-2 전구체	펙틴에스 테라아제 (4.4e-22)	10.9	+,++	105.30850 12	101.51387 72	109.8886901	132.52901 65	83.25906 421	103.615164 1

*참고: Pfam에 대한 검색한 모든 서열은 모델(models)이고 히트(hits)는 $e\text{-값} < 0.0001$ 로 인정된다.

[0294]

실시예 7

[0295]

운반체 유전자의 추가적인 분석

[0296]

실시예 1 에 기재된 바와 같이, *네투로스포라가 미스칸터스* 및 아비셀에서 성장하였을 때 예측 당 운반체 단백질을 코딩하는 10 가지 유전자는 발현 수준이 증가됨을 나타내었다 : NCU00801, NCU00988, NCU01231, NCU04963, NCU05519, NCU05853, NCU05897, NCU06138, NCU08114 및 NCU10021. 이러한 유전자 중의 9 가지에 대한 결실 스트레인은 Fungal Genetics Stock Center로부터 입수가 가능하다. NCU10021의 결실 스트레인은 입수 가능하지 않다.

[0297]

NCU05853, NCU05897 또는 NCU08114의 결실 돌연변이는 *미스칸터스* 또는 아비셀에서 성장 결함(growth defect)을 나타내는 스트레인 및/또는 셀룰라아제 효소 결함(cellulase enzyme defect)을 갖는 스트레인을 야기한다(실시예 6 ; 표 10 참조). Δ NCU05853는 야생형과 비교하여 아비셀에서의 감소된 성장을 나타내었고 감소된 엔도글루카나아제 활성도를 나타내었고, Δ NCU08114는 야생형과 비교하여 아비셀에서 감소된 성장을 나타내었고, 감소된 셀로비오스 수준을 나타내었다. 특히, *스포로트리쿰 서모필*(*Sporotrichum thermophile*)의 발현 분석과 비교하였을 때, 다른 사상균(filamentous fungus), NCU05853(ST8454) 및 NCU08114(ST5194)의 상동성(homologs)은, *S. 서모필*은 글루코스와 비교하여 아비셀에서 성장하였을 때, 또한 상향조절되었고(upregulated)(실시예 8, 표 11 참조), 이는 추가적으로 셀룰로스 이용(cellulose utilization)에서 이들의 중

요성을 나타낸다.

[표 11]

S. 호열성 균주(S. thermophile) 발현 데이터

유전자 이름	비슷한 NCU#	유전자 길이	Glu	Avi	Cot	Glu_norm	Avi_norm	Cot_norm	Avi/Glu	Cot/Glu
jgilSpath1108890lestExt_fgenesh1_pg.C_60848	NCU00988	1937	322	370	293	42.97830583	60.2250594	48.07756207	1.149068	0.9099379
jgilSpath148439le_gw1.3.3367.1	NCU01132	1539	113	59	56	15.08244894	9.60345542	9.188885583	0.522124	0.4955752
jgilSpath179030lestExt_Genewise1Plus.C_31624	NCU01231	1776	1171	1206	469	156.2968824	196.30114	76.95691676	1.029889	0.4005124
jgilSpath1116270lestExt_fgenesh1_pm.C_50266	NCU05519	1680	103	78	54	13.74771895	12.6960936	8.860711098	0.757282	0.5242718
jgilSpath184164lestExt_Genewise1Plus.C_62100	NCU05853	1706	2703	20760	14284	360.7775176	3379.11414	2343.822173	7.680355	5.2844987
jgilSpath1102977lfgenes1_pm.5_#_763	NCU05897	1446	1510	546	322	201.5442292	88.8726553	52.8360921	0.361589	0.213245
jgilSpath184305lestExt_Genewise1Plus.C_70023	NCU06138	1605	1131	1330	2376	150.9579624	216.484673	389.8712883	1.17595	2.1007958
jgilSpath1114107lestExt_fgenesh1_pm.C_20669	NCU08114	1945	2246	22423	10779	299.7803568	3649.80137	1768.696388	9.983526	4.7991986
jgilSpath1112305lestExt_fgenesh1_kg.C_60263	NCU10021	2026	6204	5287	5619	828.0664888	860.567268	922.006216	0.852192	0.905706
jgilSpath143941le_gw1.2.4209.1	NCU00801	1614	41	71	159	5.472392979	11.5567006	26.08987157	1.731707	3.8780488
jgilSpath162521lestExt_Genewise1.C_21757	NCU04963	2204	799	1548	641	106.6449266	251.968627	105.1799225	1.937422	0.8022528

각각의 예측된 운반체의 기질의 동일성(identity)을 줄이기 위해서(narrow down), NCU05853 또는 NCU08114의 결실 돌연변이를 포함하는 스트레인은 글루코스, 셀로비오스, 자일란 및 아비셀에서 배양하였다(표 12). 배양 배지는 2 %의 탄소원에 Vogel's 배지를 포함한다. 둘 다의 돌연변이체는 아비셀에서 크게 감소되었지만 자일란, 글루코스, 자일로스 또는 셀로비오스에서는 그렇지 않았다.

[0301] [표 12]

상이한 당에서 결실 돌연변이주(deletion mutants)의 성장

유전자 이름	수크로스 에서 성장	아비셀에서 성장	Mis에서 성장	자일란에서 성장	글루코스에서 성장	자일로스 에서 성장	셀로비오스 에서 성장
NCU00801	***	***	***				
NCU00988	***	***	***				
NCU01231	***	***	***				
NCU04963	***	***	***				
NCU05519	***	***	***				
NCU05853	***	*	**	***	***	***	***
NCU05897	***	*	**				
NCU06138	***	***	***				
NCU08114	***	*	**	***	***	***	***
NCU10021	결실 스트레인 없음						
wt	***	***	***	***	***	***	***

[0302]

[0303]

헤미셀룰로스의 이용에서 이러한 운반체의 역할을 조사하기 위해, 네우로스포라가 자일란에서 성장하였을 때, 10 가지 운반체 유전자의 발현을 실험하였다. 스트레인을 2 % (w/v) 자일란을 갖는 Vogel's 염에서 성장시키는 것을 제외하고 실시예 4 에 기재된 바와 같은 방법을 사용하였다. 모든 10 가지 운반체의 발현은 자일란에서 성장하는 동안 상향조절되었고(upregulated)(표 13), 이는 그들인 헤미셀룰로스 분해로부터 유도된 당(예를 들어, 자일로비오스, 자일로스, 아라비노스, 자일로-올리고당) 뿐만 아니라 셀룰로스 분해로부터 유도된 당(예를 들어, 셀로비오스, 글루코스, 셀로-올리고당)을 운반할 수 있다. 돌연변이체 성장의 결과 및 발현 분석은 적어도 두 가지의 예측된 운반체 NCU05853 및 NCU08114 는 이당류(셀로비오스, 자일로비오스) 및/또는 올리고당(셀로텍스트린)을 운반할 수 있다.

[0304] [표 13]

운반체 유전자의 발현 분석

유전자 이름	wt-자일란 4h	St-아비셀-4h/글루코스-4h에서 폴드 체인지(fold change)
NCU00801	~6	10
NCU00988.2	31.1	변화없음
NCU01231.2	732.1	변화없음
NCU04963.2	96.5	발견되지 않음
NCU05519.2	3.9	변화없음
NCU05853.2	71.2	8.5
NCU05897.2	122.3	변화없음
NCU06138.2	141.0	변화없음
NCU08114.2	10.0	11
NCU10021.2	44.7	변화없음

[0305]

실시예 8

[0306] 다양한 탄소원에서 성장하는 동안의 *N. 크라사* 운반체의 스포로트리콧 서모필 상동성의 발현 분석

[0307] 상이한 사상균(filamentous fungus)으로부터의 상동 유전자의 발현을 비교하기 위해, 스포로트리콧 서모필의 발현 프로파일(expression profile)을, 글루코스, 아비셀 또는 코튼(cotton)에서 성장한 배양물(cultures)로부터 분석하였다. cDNA 는 글루코스, 아비셀 또는 코튼의 탄소원(carbon source)을 갖는 최소배지에서 16-30 시간 동안 성장시켰다.

[0308] 첫째로, *S. 서모필* 게놈(genome)에서 *네우로스포라* 운반체 단백질의 상동성을 확인하기 위해서, 각각의 *네우로스포라* 서열은 BLAST 에서의 *S. 서모필* 단백질의 데이터베이스와 대비하여 비교하였다. 그리고 난 다음에 이러한 방법으로 발견된 *S. 서모필* 단백질의 서열을 BLAST 에서의 *네우로스포라* 단백질의 데이터베이스와 비교하였다. 이러한 결과를 도 10 에 나타내었다. 확인된 추정상의(putative) *네우로스포라* 운반체의 모든 *S. 서모필* 상동성에 대해 아미노산 서열을 SEQ ID NOs: 22-32 에서 나타내었다.

[0309] 다음으로, *S. 서모필* 상동성의 발현 프로파일을 실험하였다. 이러한 데이터를 표 11 에 나타내었다. 첫 번째 행(column)은 Joint Genome Institute *S. 서모필* 어셈블리(assembly)로부터 명명된 *S. 서모필*을 나타낸 것이다. 두 번째 행은 *네우로스포라*에서의 가장 밀접하게 관련된 추정상의 운반체에 대한 NCU 번호를 나타낸 것이다. 세 번째 행은 뉴클레오티드(nucleotides)에서 *S. 서모필* 유전자의 유전자 길이(gene length)를 나타낸 것이다. 네 번째 내지 여섯 번째 행은 탄소원(carbon source)으로서 2 % 의 글루코스, 아비셀 또는 코튼 볼(cotton balls)로 보충된 Vogel's 최소배지에서 성장하는 동안 발현수준[완전 발현 수준과 비교한, 리드(reads)의 번호]을 나타낸 것이다. 7 번째 내지 9 번째 행은 정규화된 발현 데이터를 나타낸 것이다[데이터 세트(dataset)에서 전체 리드(total reads)에 의해 나뉘어진 # 의 리드]. 마지막 두 개의 행은 아비셀/글루코스 또는 코튼/글루코스의 비율로서 각각의 유전자에 대한 상대적인 발현수준 데이터를 나타낸 것이다. NCU5853, NCU8114 및 NCU0801 의 상동(Homologs)은 아비셀 및 코튼 둘 다에서 성장하였을 때 상향조절되었다(upregulated). NCU6138 의 상동은 코튼에서 성장하였을 때 상향조절되었고, NCU4963 의 상동은 아비셀에서 성장하였을 때 상향조절되었다. 이러한 데이터는, 추정상의 운반체 NCU5853, NCU8114, NCU0801, NCU6138 및 NCU4963 은 셀룰로스의 이용(utilization)에서 중요한 역할을 함을 추가로 나타낸다.

실시예 9

[0310] 셀로덱스트린 운반체의 분석 및 확인

[0311] 순수한 셀룰로스에서 성장시켰을 때, *N. 크라사*는 세포내 β -글루코시다아제 뿐만 아니라 7 가지 주요한 촉진상과 당 운반체(Major Facilitator Superfamily sugar transporters)의 전사를 증가시키는 것으로 나타냈다(Ex. 1; also see Supplemental Data, Dataset SI, page 6 in Tian 등의, PNAS, 2009). 특히, 이러한 세트(set)로부터의 각각의 운반체가 결실된 녹아웃 스트레인(knockout strains)은 결정형 셀룰로스에서 보다 서서히 성장하였고, 이는 이들이 셀룰로스를 조건(cellulolytic conditions) 하에서 셀로-올리고당의 흡수(cello-oligosaccharide uptake)에서 직접적인 역할을 함을 암시한다(Ex. 7; 표 10, 12). 예를 들어, NCU08114 의 결실은 매우 늦은 *N. 크라사* 성장 발달(도 11) 및 셀로비오스의 감소된 소비(도 12-13)를 야기하였다. 이러한 예에서, 운반체 유전자 NCU00801/*cbt1* 및 NCU08114/*cbt2* 는 셀로덱스트린의 운반체를 코드하기(encode) 위해 추가적으로 분석하고 확인하였다.

[0312] 개별적으로 각각의 운반체의 기능을 검정하기(assay) 위해, 셀로비오스가 *S. 세레비시아*에 의해 대사 작용으로 분해되지(catabolize) 않고, 이들의 세포질(cytoplasm)에서 축적되지 않은 사실을 이용하였다(도 14). 세포내 β -글루코시다아제와 함께 기능적인 셀로비오스 운반체의 발현은, 셀로비오스가 단일 탄소원(sole carbon source)으로서 나타내었을 때 *S. 세레비시아*의 성장을 가능하게함을 판단할 수 있다. 효모 스트레인은, 녹색 형광 단백질(Green Fluorescent Protein, GFP)와 융합된 운반체 NCU00801 또는 NCU08114, 및 추정상의 세포내 β -글루코시다아제를 발현시키도록 유전자 조작되었다(engineered). 둘 다의 운반체는 발현되었고, 원형질막(plasma membrane)에 정확하게 위치하였다(도 15). NCU00801 또는 NCU08114 가 발현된 스트레인은 효모가 각각 0.0341 hr^{-1} 및 0.0131 hr^{-1} 의 특정한 성장률(growth rates)을 갖으면서 성장을 할 수 있게 한다(도 16A). 이러한 성장률은 글루코스에서의 각각의 30 % 및 12 % 의 성장률과 일치한다(도 17). 성장은 0.0026 hr^{-1} 의 비율로 성장한 추정상의 세포내의 β -글루코시다아제만이 발현되는 스트레인으로서, 셀로비오스에서 글루코스의 세포외 가수분해 후의 수송에 의해 설명되지 않고(도 16A), 대규모의 배양물(large-scale cultures)에서 성장하지 않았다(도 18). 이러한 관찰을 기반으로 CBT1 및 CBT2 로 명명된 NCU00801 및 NCU08114 은 셀로비오스 운반체로서의 기능에 대해 결정하였다.

[0313] 운반체 기능을 직접적으로 검정하기 위해, 효모 세포 내로의 $[^3\text{H}]$ -셀로비오스의 흡수(uptake)를 측정하였다. CBT1 및 CBT2 둘 다는 각각 $4.0 \pm 0.3 \text{ }\mu\text{M}$ 및 $3.2 \pm 0.2 \text{ }\mu\text{M}$ 의 K_m 값을 갖는 높은-친화도(affinity)의 셀로비오스 운반체를 발견하였다(도 19). CBT1 의 발현-정규화된(expression-normalized) V_{max} 는 CBT2 의 2.2 배이고, 효모 성장 검정(yeast growth assays)에서 나타난 차이점을 설명한 사실이다. 특히, 셀로비오스보다 더 길게 셀로덱스트린 분자는 *cbt1* 및 *cbt2* 가 발현된 효모의 성장을 지원하고(support)(도 20 ; 도 16B), 이는 셀로덱스트린 분자가 CBT1 및 CBT2 에 의한 수송됨을 나타낸다. 조화되어, CBT1 및 CBT2 에 의한 셀로비오스 수송은 과량의 셀로비오스에 의해 저해되고, CBT1 활성도는 셀로테트라오스(cellobiose)에 의해 또한 저해된다(도 21). 게다가, 정제(purification) 상에서, β -글루코시다아제, NCU00130(도 22)는 셀로비오스, 셀로트리오스(cellobiose) 및 셀로테트라오스(cellobiose)를 가수분해하는 것으로 발견하였다(도 16C).

[0314] *cbt1* 및 *cbt2* 의 오솔로그(Orthologs)는 균류의 계(fungal kingdom)에 폭넓게 분포됨을 확인하였고, 발견하였다(도 23). 최근의 발현 데이터는 균류(fungi) 및 식물 사이에 다양한 상호작용에서 이들이 중요한 역할을 함을 나타낸다. 예를 들어, 자낭기(ascomycete), *투버 멜라노스포룸(Tuber melanosporum)*, 또는 담자균(basidiomycete), *락카리아 비컬라(Laccaria bicolor)*는 엑토마이코리자(ectomycorrhizas)를 형성하기 위한 루트 팁(root tips)과 공생하여 상호작용하고(interacts), *cbt1* 의 오솔로그는 둘 다에서 상호조절되었다(Martin 등의., 2010). 마찬가지로, 사물기생균(saprophytes), *아스페르길루스 오리제(Aspergillus oryzae)*(Noguchi 등의., 2009), *포스티아 플라센타(Vanden Wymelenberg 등의., 2010)*, 및 *파네로채츠 크라이소스포리움(Phanerochaete chrysosporium)*(Vanden Wymelenberg 등의., 2010)은, 식물 벽 물질과 접촉하였을 때 *cbt2* 의 오

솔로그를 상향조절하였다. *클루이베로마이세스 락티스*(*Kluyveromyces lactis*) 및 *피치아 스티피티스*(*Pichia stipitis*)와 같은 특정 효모는 셀로비오스에서 성장하였고(Freer, 1991; Preez 등의, 1986), 셀로비오스 운반체는 *클라비스포라 루시타니아*(*Clavispora lusitaniae*)에서 보고되었다(Freer and Greene 1990). 이러한 연구에서, 이러한 효모의 모두는 *cbt1*, *cbt2* 의 오솔로그 또는 둘 다를 함유됨을 알아냈다(방법에 대해 하기 참조). 셀로비오스 운반체는 *히포크레아 제코리나*(*Hypocrea jecorina*)(*트리코데르마 레세이*, *Trichoderma reesei*)에서 관찰하였지만, 운반체를 확인하지 않았기 때문에, 만약 이러한 활성도가 *cbt1* 또는 *cbt2* 의 오솔로그에 의할 수 있기 때문에 이는 명확하지 않다(Kubicek 등의., 1993).

[0315] 셀룰로스계 균류(cellulolytic fungi)에 의한 셀로비오스 운반체의 용도(use)는, 그들이 셀룰로스에서 그들의 최적의 성장에 필수적임을 제시한다. 셀로비오스 이화작용(cellobiose catabolism)은 효모 에탄올 생산을 개선시키는지 여부를 테스트하기 위해, 상기에 구성된(constructed) 효모 스트레인은 발효 조건(fermentation conditions) 하에서 성장시켰다. 약간의 최적화에 대해, N. 크라사로부터 복사된(ported) 완전한 셀로비오스 이화작용 경로(complete cellobiose catabolism pathway)를 갖는 효모는, 86 % 의 이론값(Bai 등의, 2008)인 0.47 의 에탄올 수득률을 갖는 셀로비오스에서 에탄올로 유효하게 발효시킴을 나타내었다(도 24A). 이는 90-93 % 의 글루코스로부터 산업적인 수득률(industrial yields)과 비교할 만하다(Basso 등의, 2008). 셀로비오스에 대한 CBT1 및 CBT2 의 높은 친화도(high affinity)는 S. 세레비시아의 육탄당 운반체(hexose transporters)와 비교할 만하고(Reifenberger 등의, 1997), 세포외 β -글루코시다아제가 보고되어 있으며(Chauve 등의, 2010), 이는 셀로비오스/셀로텍스트린 수송 시스템(cellobiose/cellobioextrin transport system)은 SSF 동안에 특히 유용함을 나타낸다. 예를 들어, 셀로비오스/셀로텍스트린 수송은 셀룰로스에서 글루코스로의 완전한 가수분해에 대한 요건을 보다 낮출 수 있고, 셀룰로스계 효소(cellulolytic enzymes)의 셀로비오스-매개 저해(cellobiose-mediated inhibition)를 감소시키며, 글루코스-의존적인 유기체(glucose-dependent organisms)에 의한 오염의 위험을 줄인다. 실제로, 셀로비오스/셀로텍스트린 수송 시스템이 발현된 효소는, 셀로비오스 및 글루코스 둘 다의 정상 상태의 농도(steady state concentration)를 감소시키고 에탄올 생산율을 증가시켜 SSF 반응의 효율을 현저하게 개선시킨다(도 24B, C).

[0316] 셀룰로스로부터의 바이오연료(Biofuel)의 생산은 개선된 숙주 스트레인(host strains)에 의한 연료 생산과 함께 식물 바이오매스(plant biomass)에서 당으로의 효율적이고 경제적인 디폴리머라이제이션(depolymerization)을 필요로 한다(Kumar 등의, 2008). 셀룰로스계 균류는 식물 바이오매스에서의 최적의 성장에 대한 셀로-올리고당 수송 경로를 사용하는 것을 본원에서 보여주고 있다. 게다가, 효모에서 이러한 경로의 재구성(reconstitution)은, 그들이 운반체 및 세포내의 셀로-올리고당 가수분해 효소(cello-oligosaccharide hydrolase)로 구성된 최소한의 경로를 갖는, 셀로비오스 이화작용(cellobiose catabolism)을 증진시키기 위한 모듈식의 방식(modular fashion)으로 복사될 수 있다(ported)(도 25). 효모 및 다른 유기체의 바이오연료-생산 스트레인(biofuel-producing strains)에서 셀로텍스트린 수송의 용도(use)는 보다 경제적이고 실행가능한 셀룰로스계 바이오연료 공정(cellulosic biofuel processes)을 생산하는데 대단히 중요하다.

[0317] 운반체 및 β -글리코시다아제 오솔로그

[0318] 셀로텍스트린 운반체에 대한 GenBank 접근번호(accession numbers) 또는 Joint Genome Institute (JGI) 단백질 ID (PID) 번호는 하기와 같다 : *투버 멜라노스포룸*(*Tuber melanosporum*), CAZ81962.1 ; *피치아 스티피티스*(*Pichia stipitis*), ABN65648.2 ; *락카리아 비컬라*(*Laccaria bicolor*), EDR07962 ; *아스페르길루스 오리제*(*Aspergillus oryzae*), BAE58341.1 ; *파네로카에테 크라이소스포리움*(*Phanerochaete chrysosporium*), PID 136620(JGI)(Martinez 등의, 2004); *포스티아 플라센타*(*Postia placenta*), PID 115604(JGI)(Martinez 등의, 2009). *사카로미세스 세레비시아*(*Saccharomyces cerevisiae*) HXT1 및 *클루이베로마이세스 락티스*(*Kluyveromyces lactis*) LACP 에 대한 GenBank 접근번호는 각각 DAA06789.1 및 CAA30053.1 이다. *P. 크라이소스포리움* 및 *P. 플라센타* 게놈은 각각 genome.jgi-psf.org/Phchr1/Phchr1.home.html 및 genome.jgi-psf.org/Posp11/Posp11.home.html 에서 접속할 수 있다.

[0319] NCU00130 의 오솔로그인 셀로텍스트린 가수분해 효소에 대한 GenBank 접근번호는 하기와 같다 : *T. 멜라노스포*

리움(*T. melanosporum*), CAZ82985.1 ; *A. 오리제*(*A. oryzae*), BAE57671.1; *P. 플라센타*(*P. placenta*), EED81359.1 ; 및 *P. 크라이소스포리움*(*P. chrysosporium*), BAE87009.1. 셀로덱스트린 운반체 오솔로그를 포함하는 그 밖의 유기체는 GH3 패밀리에서의 유전자를 세포내 β -글리코시다아제인 것으로 예측된 GH3 패밀리에서의 유전자를 포함하며(Bendtsen 등의, 2004 ; Cantarel 등의, 2009), 하기와 같다 : *클루이베로마이세스 락티스*(*Kluyveromyces lactis*), CAG99696.1 ; *락카리아 비컬라*(*Laccaria bicolor*), EDR09330 ; *클라비스포라 루시타니아*(*Clavispora lusitaniae*), EEQ37997.1 ; 및 *피치아 스티피티스*(*Pichia stipitis*), ABN67130.1.

[0320] 스트레인 및 배지(Media)

[0321] 본 연구에서 사용된 효모 스트레인은 하기와 같은 유전자형(genotype)을 갖는 YPH499 이다(Sikorski 등의, 1989) : *MATa ura3-52 lys2-801_amber ade2-101_ochre trp1- Δ 63 his3- Δ 200 leu2- Δ 1*. 100 mg/L 아데닌 헤미설페이트(adenine hemisulfate)로 보충된 YPD 배지에서 성장시켰다. 변형된 스트레인(Becker 등의, 2001)은, 100 mg/L 아데닌 헤미설페이트(adenine hemisulfate)로 보충된 적절한 완전한 최소 드롭아웃 배지(complete minimal dropout media)에서 성장시켰다. 본 연구에서 사용된 *네우로스포라 크라사* 스트레인은 Fungal Genetics Stock Center로부터 입수하였고(McCluskey 2004), WT(FGSC 2489) 및 두 가지의 셀로비오스 운반체 결실 스트레인[FGSC 16575, ANCU00801.2 및 FGSC 17868, ANCU08114.2 (Colot 등의, 2006)]을 포함한다.

[0322] 플라스미드 및 클로닝

[0323] 프라이머(primers)-ATATATGAGCTCGTGAGTAAGGAAAGAGTGAGGAACATATC(SEQ ID NO : 53) 및 ATATATACTAGTTGTTTATATTGTTGTAAAAAGTAGATAATTACTTCC(SEQ ID NO : 54)-를 사용한 *SacI* 및 *SpeI* 사이에 삽입된 *S. 세레비시아* PGK1 프로모터를 포함한 변형을 갖는 2 μ 의 플라스미드, pRS426 내로 운반체를 클론하였다(cloned). [상기와 하기의 모든 프라이머에서, 제한 부위(restriction sites)는 밑줄을 그었다]. C-말단 Myc-tag 및 최적화된 Kozak 서열을 갖는 NCU00801 (Miyasaka 1999)을, 프라이머-ATGGATCCAAAAATGTCGTCTCACGGCTCC(SEQ ID NO: 55) 및 ATGAATTCCTACAAATCTTCTTCAGAAATCAATTTTGTTCAGCAACGATAGCTTCGGAC(SEQ ID NO: 56) -를 사용하여 *BamHI* 및 *EcoRI* 사이에 삽입하였으며, C-말단 Myc-tag 및 최적화된 Kozak 서열을 갖는 NCU08114 을, 프라이머-ATACTAGTAAAAATGGGCATCTTCAACAAGAAGC(SEQ ID NO : 57) 및 GCATATCGATCTACAAATCTTCTTCAGAAATCAATTTTGTTCAGCAACAGACTTGCCCTCA TG(SEQ ID NO: 58)-를 사용하여 *SpeI* 및 *ClaI* 사이에 삽입하였다. GFP 융합(fusions)을 생산하기 위해, Gly-Ser-Gly-Ser 의 N-말단 링커(linker)를 갖는 슈퍼폴더 GFP(superfolder GFP)(Pedelacq 등의, 2006)를, 프라이머-TATTAATCGATGGTAGTGGTAGTGAGCAAGGGC GAGGAG(SEQ ID NO: 59) 및 TATTAAGTCGACCTACTTGTACAGCTCGTCCATGCC(SEQ ID NO: 60)-를 갖는 PGK1 프로모터- 함유하는 pRS426 플라스미드(PGK1 promoter- containing pRS426 plasmid)를 *ClaI* and *SaiI* 사이에 첫 번째로 삽입하였다. 운반체는 GFP 와 융합된 다음에 하기와 같다 : 프라이머-NCU00801 은 CATGGATCC ATGTCGTCTCACGGCTCC(SEQ ID NO: 61) 및 TATAATGAATTCAGCAACGATAGCTTCGG AC(SEQ ID NO: 62)-를 사용하여 *BamHI* 및 *EcoRI* 사이에 삽입되었으며, NCU08114 은 TATTAAGTACTAGTATGGGCATCTTCAACAAGAAGC(SEQ ID NO: 63) 및 TTATAAGAATTCAGCAAC AGACTTGCCCTCATG(SEQ ID NO: 64) 프라이머를 사용한 *SpeI* 및 *EcoRI* 사이에 삽입되었다.

[0324] β -글리코시다아제, NCU00130 을 상기한 바와 같이 PGK1 프로모터를 포함한 변형된, 2 μ 의 플라스미드, pRS425 내로 클론시켰다. 최적화된 Kozak 서열 및 C-말단 6xHis tag 를 갖는 NCU00130 을, 프라이머-GCATACTAGTAAAAATGTCTCTT CCTAAGGATTTCCTCT (SEQ ID NO: 65) 및 ATACTGCAGTTAATGATGATGATGATGATGGTCTTCTTGATCAAAGAGTCAAAG (SEQ ID NO: 66)-을 사용한 *SpeI* 및 *PstI* 사이에 삽입시켰다. 모든 구조물은 *XhoI* 및 *KpnI* 사이의 *Cyc* 전사 종결자(transcriptional terminator)를 포함한다. 모든 *N. 크라사* 유전자는, 단일 탄소원(sole carbon source)으로서 순수한 셀룰로스(아비셀)를 갖는 최소배지(minimal media)에서 배양된 *N. 크라사*(FGSC 2489)로부터 분리된 mRNA로부터 합성된 cDNA로부터의 PCR 에 의해 증폭되었다.

[0325] 효모 성장 검증

[0326] 셀로-올리고당(cello-oligosaccharides)에서의 성장을 모니터링하기 위해, 유전자조작된 스트레인을 적절한 드롭아웃(dropout)을 갖는 5 mL 의 완전한 최소배지에서 밤새 성장시켰다. 이러한 스타터 배양물(starter cultures)을 25 mL 의 ddH₂O 로 3 번 세척하였고, 적절한 완전 보충 배지(Complete Supplemental Media, CSM) 및 1 % (w/v)의 셀로비오스 또는 0.5 % (w/v)의 셀로트리오스 또는 셀로테트라오스를 더한 효모 질소 염기(Yeast Nitrogen Base, YNB)에 0.1 의 OD(600 nm 에서)에 재현탁시켰다(resuspended). 30 °C 에서 최대한의 진폭으로 일정하게 교반시키고 0.4 mL 의 최종 검정 부피를 갖는 Bioscreen CTM 에서 검정을 실시하였다. OD 에서의 변화는 600 nm 에서 또는 450 - 580 nm 의 광대역필터(wideband filter)를 사용하여 측정하였다. 성장률을 각각의 성장 곡선의 선형 부분(linear portion)으로부터 획득하였고, 세 가지 독립적인 실험의 평균 ± 실험들 사이의 표준편차로서 나타내었다. 셀로트리오스 및 셀로테트라오스를 Seikagaku Biobusiness Corporation (일본 도쿄)로부터 입수하였다.

[0327] NCU00130 의 정제 및 이의 활성도의 검정

[0328] *cbt1* 및 NCU00130 이 발현된 *S. 세레비시아*의 1L 배양물을 완전 최소 배지에서 2.0 의 OD 에서 성장시켰다. 세포를 원심분리하여 채취하였고, 30 mL 의 세포용혈 완충용액(lysis buffer)[50 mM NaH₂PO₄(pH 8.0), 300 mM NaCl, 10 mM 이미다졸, 2 mM p-ME, CompleteTM Mini, EDTA 유리된 분해효소 저해제 혼합물(free protease inhibitor cocktail)]에 재현탁시켰다. 세포를 음파처리(sonication)하여 용해하였고, 용해물(lysate)을 30 분 동안 15,000 g 에서 원심분리하여 투명하게 하였다(cleared). 용해물을 중력 유동(gravity flow)으로 1 mL 의 니켈-NTA 수지(nickel-NTA resin)에 결합시켰고, 25 mL 세척 완충용액(wash buffer)(세포용혈 완충용액과 동일하지만, 20 mM 의 이미다졸을 갖음)으로 세 번 세척하였다. NCU00130 은 5 mL 의 용리 완충용액(elution buffer)으로 용출시켰고(eluted), 적절한 분획(fractions)을 모으고(pooled), 저장 완충용액(storage buffer) [인산 완충 식염수(PBS), 2 mM DTT, 10 % 글리세롤] 내로 교환시키고(exchanged), 알리쿼트하고(aliquoted), 액체 질소에 냉동시키고, -80 °C 로 저장하였다. 순도(Purity)는 SDS-PAGE 에 의해 측정되고(도 22), 단백질 농도는 108,750 M⁻¹cm⁻¹ 의 흡광계수(extinction coefficient)를 사용하여 280 nm 에서의 흡광도로부터 측정하였다.

[0329] 정제된 NCU00130 을, 상이한 셀로텍스트린 기질을 갖는 가수분해 활성도(hydrolysis activity)로부터 검정하였다. 3 mM DTT 가 더해진 150 µL PBS 에서 500 µM 의 각각의 당과 함께 5 pmol 효소를 배양하여 측정하였다. 100 µL 를 제거하기 전에 30 °C 에서 40 분 동안 반응을 진행시키고, 400 µL 의 0.1 M NaOH 로 퀸칭하였다 quenched. CarboPac PA200 컬럼으로, Dionex ICS-3000 를 갖춘 이온 크로마토그래피로 결과를 분석하였다. 피크(Peak)를 전기화학 검출기(electrochemical detector)로 발견하였다.

[0330] 운반체 오솔로그(Orthologs)의 계통 분석(Phylogenetic Analysis)

[0331] CBT1 및 CBT2 의 오솔로그의 아미노산 서열을 온라인 데이터베이스(online databases)로부터 획득하였다. 다수의 서열 정렬(Multiple sequence alignments)을 T-Coffee 를 사용하여 수행하였다(Notredame 등의, 2000). 최대한의 가능성을 갖는 계통발생론(maximum likelihood phylogeny)을 100 부트스트랩(Bootstraps)을 갖는 PhyML 버전 3.0(Guindon and Gascuel 2003)을 사용하여 측정하였다. 둘 다의 프로그램을 Phylogeny.fr(웹페이지 phylogeny.fr/)을 통해 접속하였다. 결과적으로 생성된 트리(tree)는 FigTree v. 1.2.1 으로 나타내었다(웹페이지 tree.bio.ed.ac.uk/).

[0332] 발효(fermentation) 및 SSF

[0333] 발효 및 SSF 실험에서, NCU00130 가 발현된 효모, 및 Myc-태그된(tagged) *cbt1* 또는 운반체가 없는 것 사이에 비교를 실시하였다. 이러한 스트레인은 완전 최소 배지에서 밤새 산소를 필요로하면서 성장하였고, 25 mL 의 물로 세 번 세척하였으며, 밀봉된 혈청 플라스크(sealed serum flasks)에서 적절한 CSM, 및 2% (w/v) 셀로비오스 또는 3% (w/v) 순수한 셀룰로스(아비셀) 둘 중의 하나를 더한 50 mL 의 YNB 에서 2.0 의 최종 OD 에 재현탁

시켰다. SSF 반응물은 β -글루코시다아제 보충(supplementation) 없이 필터-멸균된(filter-sterilized) Celluclast(Sigma C2730)의 50 필터페이퍼 유닛(Filter Paper Units)/g 셀룰로스를 또한 포함한다. 반응물은 교반시키면서 30 °C 에서 산소결핍에 의해(anaerobically) 실행하였다. 나타낸 시점에서, 1 mL 의 샘플을 제거하였고, 0.2 μ m 의 시린지 필터(syringe filter)를 통해 여과하였다. 여과액(filtrate)에서의 에탄올, 글루코스 및 셀로비오스 농도는 Aminex HPX-87H 컬럼 및 굴절률 검출기(refractive index detection)을 갖는 HPLC 에 의해 측정하였다.

[0334] N. 크라사 성장 및 알라마 블루® 검정(Alamar Blue Assays)

[0335] WT N.크라사(FGSC 2489) 및 동종핵성(homokaryotic) NCU08114(FGSC 17868) (Colot 등의, 2006)는 Fungal Genetics Research Center(McCluskey 2003)로부터 입수가가능하고, 250 mL 의 언배플드 플라스크(unbaffled flask)에 2 % 의 수크로스 또는 순수한 셀룰로스(아비셀)을 더한 50 mL 의 Vogel's 염에서 25 °C 에서 성장시켰다. 16 시간 또는 28 시간 후에, 각각의 100 μ L 의 알라마블루®를 첨가하고, 배양물을 20 분 동안 실온에서 배양하였다. 이러한 시간에서 1 mL 의 샘플을 제거하고, 잔해(debris)를 펠렛화시키고(pelleted), 100 μ L 의 상청액의 형광(fluorescence)을, Beckman Coulter Paradigm plate reader 에서 535/595 nm 의 여기/방출(excitation/emission) 파장으로 측정하였다.

[0336] N. 크라사 셀로비오스 수송 검정

[0337] WT N. 크라사(FGSC 2489), 및 NCU00801(FGSC 16575) 및 NCU08114(FGSC 17868)의 동종핵성 결실 라인(homokaryotic deletion lines)을 Fungal Genetics Stock Center 로부터 획득하였고(McCluskey 2003), 10^6 분생자(conidia)/mL 의 접종물(inoculum)을 시작으로, 25 °C 에서 2 % (w/v) 수크로스를 더한 50 mL 의 Vogel's 염에서 16 시간 동안 성장시켰다. 균사체(Mycelia)는 원심분리에 의해 채취하였고, Vogel's 염으로 세 번 세척하였고, 운반체 발현을 유도하기 위해 4 시간 동안 0.5% (w/v) 순수한 셀룰로스(아비셀)을 더한 Vogel's 염에 이동시켰다(transferred). 10 mL 의 배양물을 원심분리로 채취하고, Vogel's 염으로 세 번 세척하고, 시클로헥사미드(100 μ g/mL) 및 90 μ M 의 각각의 셀로텍스트린(셀로비오스, 셀로트리오스 또는 셀로테트라오스)을 더한 1 mL 의 ddH₂O 에서 재현탁시켰다. 셀로텍스트린 소비를 측정하기 위해, 100 μ L 를 15 분 후에 제거하고, 원심분리로 투명하게 만들고(clarified), 900 μ L 의 0.1 M NaOH 내로 이동시켰다. 상청액에 남아있는 당의 양은 CarboPac PA200 컬럼을 사용한 Dionex ICS-3000 을 갖는 HPLC 로 측정하였다. 피크를 전기화학 검출기(electrochemical detector)로 감지하였다(detected).

[0338] GFP 형광 및 공초점 형광 현미경(Confocal Fluorescence Microscopy)

[0339] 벌크-셀 GFP 형광(Bulk-cell GFP fluorescence)을 485/535 nm 의 여기/방출 파장으로 Beckman Coulter Paradigm plate reader 에서 측정하였다. 공초점 형광 현미경은, Metamorph 소프트웨어로 조절되고 488 nm 레이저를 갖는 Yokogawa CSU-X1 스피닝 디스크 헤드(spining disc head)에 부착된 Leica SD6000 현미경에서 100x 1.4 NA 유침 대물렌즈(oil immersion objective)를 사용한 0.8 - 1.2 의 OD (600 nm 에서)에서의 세포를 가지고 실시하였다. Z 시리즈는 200 nm 스텝 사이즈(step size)로 기록하였고 Image J 를 사용하여 분석하였다.

[0340] [³H] 셀로비오스 수송 검정(Transport assays) 및 동역학적 파라미터(Kinetic Parameters)

[0341] 수송 검정을 오일-스톱 방법(oil-stop method)의 변형(modification)을 사용하여 실시하였다(Arendt 등의, 2007). GFP 에 융합된 cbt1 또는 cbt2 가 발현된 효모 스트레인은 OD(선택 배지에서 1.5-3.0 의 600 nm 에서)에 대해 성장되었고, 차가운 검정 완충용액[30 mM MES-NaOH(pH 5.6) 및 50 mM 에탄올]으로 세 번 세척하였으며, 20 의 OD 에 재현탁시켰다. 수송 반응을 시작하기 위해, 50 μ L 의 세포를, 100 μ L 의 실리콘유(silicone oil)(Sigma 85419) 상에서 층으로 놓여진(layered) 50 μ L 의 [³H] 셀로비오스에 첨가하였다. 반응물(Reactions)을 17,000 에서 1 분 동안 오일을 통해 세포를 회전시켜 중단시키고, 튜브를 에탄올/건조 아이스에

서 냉동시키고, 세포-펠렛(cell-pellets)을 함유하는 튜브-바닥(tube-bottoms)을 1 mL 의 0.5 M NaOH 내로 잘라내었다(clipped off). 펠렛(pellets)을 밤새 용해시키고, 5 mL 의 Ultima Gold 신틸레이션 유체(scintillation fluid)를 첨가하고, CPM 을 Tri-Carb 2900TR 신틸레이션 계수기에서 측정하였다. [³H] 셀로비오스는 Moravsek Biochemicals, Inc.로부터 구매하였고, 4 Ci/mmol 의 고유활성도(specific activity) 및 순도 > 99 % 을 갖는다. 동역학적 파라미터는 셀로비오스 농도의 범위에 대한 3 분 이상 [³H] 셀로비오스 흡수의 선형 비율(linear rate)로 측정하여 결정하였다. V_{max} 및 K_m 값은 SigmaPlot[®] 에 의한 비선형회귀(non-linear regression)에 의한 셀로비오스 농도 대 비율의 구성(plot)에 대한 단일 사각형, 2-파라미터의 쌍곡선 함수(single rectangular, 2-parameter hyperbolic function)에 적용시켜 결정하였다. V_{max} 값은 수송 검정을 시작하기 전에 즉시 OD 200 에서 100 μ L 의 세포로부터의 GFP 형광(fluorescence)을 측정하여 운반체의 존재비(abundance)에서의 차이점에 대해 정규화시킨 것이다. 본문에서 나타난 동역학적 파라미터는 세 가지 분리 실험으로부터의 평균 \pm 표준편차이다. 비교 검정은 250 μ M 의 각각의 경쟁자(competitors)의 존재 하에서 20 초 이상 50 μ M [³H]-셀로비오스의 수송을 측정하기 위해 실시하였다.

[0342] 대규모 효모 성장(Large Scale Yeast Growth)

[0343] 상이한 탄소원에서의 성장을 모니터링하기 위해, 유전자 조작된 스트레인을 밤새 적절한 드롭아웃(dropouts)으로 5 mL 의 완전 최소 배지(complete minimal media)에서 성장시켰다. 이러한 스타터 배양물(starter cultures)을 25 mL의 ddH₂O 로 세 번 세척하고, 적절한 완전 보충 배지(Complete Supplemental Media, CSM) 및 2 % (w/v) 셀로비오스를 더하여 50 mL 효모 질소 염기(Yeast Nitrogen Base, YNB)에 0.1 의 OD(600 nm 에서)에 재현탁시켰다(resuspended). 배양물을 200 rpm 으로 교반시키면서 30 °C 에서 250 mL 의 저해시키지 않는 플라스크에서 성장시켰다. OD(600 nm 에서)에서의 변화를 정기적으로 샘플을 제거하면서 모니터링하였다.

실시예 10

[0344] 셀로텍스트린 운반체 기능에 대한 중요한 잔기(Critical Residues)의 확인

[0345] 본 실시예에서, 서열 분석 및 돌연변이유발(mutagenesis) 연구를 셀로텍스트린 운반체에서 보존되고(conserved) 기능적으로 중요한 잔기를 확인하기 위해 사용하였다. 게다가, 추가적인 셀로텍스트린 운반체를 확인하였다.

[0346] 셀로텍스트린 운반체 NCU00801(*cbt1*) 또는 NCU08114(*cbt2*) 및 야생형 β -글루코시다아제 NCU00130 의 다양한 돌연변이체(mutants)가 발현된 효모 스트레인의 성장률을 탄소원으로서 셀로비오스와 함께 성장하였다. NCU00801 의 96 위치 및 NCU08114 의 96 위치에서의 아미노산 잔기는 생산자의 지시(manufacturer's instructions)에 따라 QuickChange[®] II 사이트-직접적인 돌연변이유발 키트(Site-directed Mutagenesis Kit) (미국 캘리포니아 라즐라 Stratagene 에서 입수)를 사용하여 알라닌(alanine)에 대해서 개별적으로 돌연변이시켰다(mutated). 스트레인을 2 % 셀로비오스를 갖는 합성 정의된 배지(synthetic defined media)-ura -leu 100 mg/L 아데닌에서 성장하였다. 배양물을 두 가지의 독립된 군집(colony)로부터 시작하였다.

[0347] 도 26(A, B)에서 나타난 결과가 나타낸 바와 같이, W66, L73, Y74, N87, Y89, D90, Q 104, F107, G113, F120, Y123, D139, G142, K144, M147, G150, Q169, F170, G173, R174, G178, G180, P189, Y191, E194, P198, R201, Y208, W235, R236, Q242, ²⁵⁷PESPRF²⁶² (SEQ ID NO: 67), Y279, G283, E296, D307, K308, W310, D312, R325, G336, Y345, N369, D385, F462, P468, E476, T480, 또는 G486 에서의 치환(substitutions)을 갖는 NCU00801 이 발현된 돌연변이 스트레인은 야생형 스트레인과 비교하여 적어도 25 % 의 성장 결함(growth defect)을 나타내었다.

- [0348] NCU08114 에서 알라닌 스캐닝 실험(alanine scanning experiment)은 기능적으로 중요한 바와 같은 하기의 잔기를 나타낸다 : L38, Y39, G54, D56, F73, G91, P100, D104, G107, R108, M1 18, R139, F144, Q150, P154, E159, P163, H165, R166, Y173, N174, W199, Q214, ²²²PESP²²⁵ (SEQ ID NO: 68), Y244, H245, D249, E258, E268, Q302, W303, S304, N306, Y312, F359, L360, F402, Y403, S404, Y414, E417, P420, Y421, K426, N442, N446, P447, W459, K460, E482, T483, L488, E489, E490, D496, and G497 (도 26b).
- [0349] 특히, 모티프(motifs) ⁷³LYF⁷⁵, ²⁵⁷PESP²⁶⁰ (SEQ ID NO: 69) 및 ²⁷⁸KYH²⁸⁰ [NCU00801 의 잔기 번호(residue numbering)]는 둘 다의 운반체[NCU00801의 잔기 ²⁵⁷PESP²⁶⁰ (SEQ ID NO: 69) 및 NCU08114 의 잔기 ²²²PESP²²⁵ (SEQ ID NO: 68)]에서 기능적으로 중요하게 보이는 것 같고, 이는 29 % 의 아미노산 서열 동일성을 갖는다(도 26B, C). 보통 운반체 또는 특히 (이중-밀줄쳐진) β-링크된 운반체에서 보존된 몇몇의 잔기(도 26B, C 에 이텔릭체)는 예를 들어 D90(NCU00801), D56(NCU08114), 및 L73(NCU00801), L38(NCU08114), 운반체 기능(밀줄쳐진)에 대해 중요함을 실험적으로 나타낸 것이다. 돌연변이 유발 실험(mutagenesis experiment)의 결과는 기능적으로 중요한, 예를 들어 Q168(NCU00801) 및 Q214(NCU08114)와 같은 NCU00801/NCU08114 계통군(clade)[캡드(capped)]에서 보존된 잔기를 또한 보여준다. 게다가, 본 실험에서 기능적으로 중요한 것으로 알려진 다수의 잔기는 S. 세레비시아 당 운반체(Hxt1/Hxt3), 예를 들어 L73(NCU00801) 및 L38(NCU08114)에서 보존되는 것으로 사전에 나타내었다.
- [0350] 상이한 유기체로부터의 *N. 크라사* 셀로텍스트린 운반체의 오솔로그(Orthologs)는 또한 연구되어있다(도 27). 대표적인 오솔로그는 Genescript 에 의해 합성하였고, 발현 벡터, 사이트(sites) BamHI 및 HindIII 를 사용하 Cup1 프로모터를 포함하는 pRS426 내로 클론되었다(cloned). 이러한 구성물은 세포내 β-글루코시다아제, NCU00130 와 함께 효모 스트레인 YPH499 내로 형질전환되었다(transformed). 셀로비오스가 단일 탄소원으로서 존재할 때, 이러한 스트레인의 성장률을 측정하여 운반체 활성도를 측정하였다.
- [0351] 그 대신에, 추정상의 오솔로그를 함유하는 상이한 진균(fungal) 스트레인은 셀로비오스로 보충된 리치 배지(rich media)에서 재배하였다. 전체 RNA 는 분리되었고, cDNA 내로 역전사되었다(reverse transcribed). 중합효소 연쇄 반응(PCR)은 cDNA 로부터 직접적으로 추정상의 운반체 유전자를 증폭시키는데 사용하였다. 그러나, 조절 메카니즘(regulation mechanism) 및 발현 패턴(expression pattern)은 균종(fungal species)에서 셀로텍스트린 운반체에 대해 잘 알려지지 않았기 때문에, 배양 조건(cultivation condition)의 변경에도 불구하고 추정상의 운반체를 코딩하는 cDNAs 를 항상 수득할 수 있는 것은 아니다. 이러한 경우에, 프라이머를 GenBank 로부터의 이에 상응하는 cDNA 서열에 따라 설계하였고(designed), 주형(template)으로서 게놈의(genomic) DNA 를 사용하여 엑손(exons)을 증폭시키는데 사용된다. 오버랩-연장 PCR(Overlap-extension PCR)은 전체 길이의 유전자(full-length genes) 내로 엑손을 모으는데 사용된다. DNA 조립 방법(DNA assembler method)을 사용한 HXT7 프라이머 및 HXT7 종결자(terminator)를 포함하는 pRS424 셔틀벡터(shuttle vector) 내로 결과적으로 생성된 PCR 산물을 클론하였다. 형질전환체(transformant)로부터 분리된 효모 플라스마는 대장균 DH5a 내로 재형질전환되었고(retransformed), 분리된 대장균 플라스마는, 원본(original) 운반체 유전자를 증폭시키는데 사용되는 프라이머를 사용한 진단 PCR(diagnostic PCR)에 의해 첫 번째로 점검되었다(chchecked). 전체적인 오픈리딩프레임(open reading frames)을 플라스마의 정확한 구성을 확인하기 위한 배열 순서를 밝히기 위해(sequencing) 제출되었다. *P. 스티피티스*로부터의 오솔로그 LAC2, HXT2.1 및 HXT2.6 에서, 하나 또는 그 이상의 선택가능한 코돈(alternative codons)(CUG)은 Leu 를 Ser 로 치환하였다. 효모 상동의 제조합 매개 DNA 조립 방법(yeast homologous recombination mediated DNA assembler method)을 사용하여 대부분의 클로닝 작업(cloning work)을 수행하였다. pRS424-HXT7-GFP 플라스미드를 추정상의 셀로텍스트린 운반체의 클로닝(cloning)에 사용하였다. 이러한 플라스미드에서, HXT7 프로모터, 둘 다의 말단(both ends)에서 EcoRI 사이트를 갖는 측면에 배치된(flanked) GFP 유전자 및 HXT7 종결자(terminator)는 ClaI 및 BamHI 에 의해 선 모양으로 만들어진 pRS424 셔틀 벡터(New England Biolabs) 내로 조립되었다. HXT7 프로모터 및 종결자(terminator)와 서열 동일성(sequence identity)을 공유하는(sharing) DNA 단편과 함께 측면에 배치된(flanked) 추정상의 운반체의 PCR 산물은 표준 리튬 아세테이트 방법(standard lithium acetate method)을 사용한 pRS424-HXT7-GFP 소화된(digest) EcoRI 를 갖는 CEN.PK2-1C 내로 공동으로 이동되었다(co-transferred). 결과적으로 생성된 형

질전환 혼합물(transformation mixture)을 형질전환체(transformants)를 회복하기 위해(recover) 2 % D-글루코스로 보충된 SC-Trp 플레이트에 도말하였다(plated). 추정상의 셀로덱스트린 운반체 오솔로그가 발현된 효모 및 NCU00130 를 단일 탄소원으로써 셀로비오스에서의 성장을 테스트하였다.

[0352] 추정상의 셀로덱스트린 운반체 오솔로그의 목록 및 본 연구로부터 수득한 결과를 표 14 에 나타내었다.

[0353] [표 14]

추정상의(putative) 셀로덱스트린 운반체 오솔로그(orthologs) 및 결과 요약의 목록

N. 크로사 오솔로그	종(Species)	NCBI 참조서열 /NCBI GI 번호/ JGI 번호 ♣	Aver. 성장률	성장률 오류	Seq 결과 *
NCU00809	체토미움 글로보숨 (<i>Chaetomium globosum</i>) CBS148.51	XP_001220480	-	-	OK
NCU00809	포도스포라 앤서리나 (<i>Podospora anserine</i>)	XP_001912722	-	-	-
NCU00809	넥트리아 해마토코카 (<i>Nectria haematococca</i>) mpVI77-13-4	EEU41662	-	-	-
NCU00809	아스페르길루스 니돌란스 (<i>Aspergillus nidulans</i>) FGSC A4	XP_660803	-	-	1 인트론 및 50 bp 삽입
NCU00809	아스페르길루스 테레우스 (<i>Aspergillus terreus</i>) NIH2624	XP_001218592	-	-	-
NCU00809	탈라로미케스 스티피타투스 (<i>Talaromyces stipitatus</i>) ATCC 10500	XP_002341594	-	-	-
NCU00809	아스페르길루스 니가 (<i>Aspergillus niger</i>)	XP_001395979	-	-	Ala>Val
NCU00809	아스페르길루스 푸미가투스 (<i>Aspergillus fumigates</i>) Af293	XP_747891	-	-	-
NCU00809	아스페르길루스 테레우스 (<i>Aspergillus terreus</i>) NIH2624	XP_00120996	-	-	-
NCU00809	아스페르길루스 오리재 (<i>Aspergillus oryzae</i>) RIB40	XP_001817400	-	-	OK
NCU08114	포도스포라 앤서리나 (<i>Podospora anserine</i>)	XP_001908539	-	-	N/A
NCU08114	페니실리움 크리스게움 위스콘신 (<i>Penicillium chrysogenum</i> Wisconsin) 54-1255	XP_002568019	-	-	N/A
NCU08114	아스페르길루스 테레우스 (<i>Aspergillus terreus</i>) NIH2624	XP_001209810	-	-	틀림 (Wrong)
NCU08114	아스페르길루스 오리재 (<i>Aspergillus oryzae</i>) RIB40	XP_001820343	-	-	OK
NCU08114	아스페르길루스 테레우스 (<i>Aspergillus terreus</i>) NIH2624	XP_001210859	-	-	N/A
NCU08114	네우로스포라 크라사 (<i>Neurospora crassa</i>) OR74A	XP_001728155	-	-	N/A
NCU08114	아스페르길루스 오리재 (<i>Aspergillus oryzae</i>) RIB40	XP_001826848	-	-	N/A
NCU08114	아스페르길루스 니돌란스 (<i>Aspergillus nidulans</i>) FGSC A4	XP_657617	-	-	OK

[0354]

NCU08114	탈라로미세스 스티피타투스 (<i>Talaromyces stipitatus</i>) ATCC 10500	XP_002487579	-	-	N/A
NCU08114	체토미움 글로보숨 (<i>Chaetomium globosum</i>) CBS 148.51	XP_001227497	-	-	틀림 (Wrong)
NCU08114	트리코데르마 아트로비리데 (<i>Trichoderma atroviridae</i>)	215408	0.000836 364	0.0006487 1	I,D
NCU08114	체토미움 글로보숨 (<i>Chaetomium globosum</i>)	XP_001220290 .1	0.004036 364	0.0004716 8	OK
NCU08114	아스페르길루스 니둘란스 (<i>Aspergillus nidulans</i>)	ANID_08347	0.011109 091	0.0000727 27	기타 (Other)
NCU08114	플루로투스 오스트레아투스 (<i>Pleurotus ostreatus</i>)	51322	0.003903 03	0.0001821 2	-
NCU08114	스포로트리쿰 서모필 (<i>Sporotrichum thermophile</i>)	114107	0.009569 697	0.0021636 6	-
NCU00801	아스페르길루스 니둘란스 (<i>Aspergillus nidulans</i>)	XP_660418.1	0.000860 606	0.000438	P
NCU00801	마그나포스 그리시아 (<i>Magnaporthe grisea</i>)	XP_364883.1	00509090 9	0.0013831 3	OK
NCU00801	아스페르길루스 푸미가투스 (<i>Aspergillus fumigates</i>)	XP_753099.1	0.003975 758	0.0021195 1	OK
NCU00801	트리코데르마 아트로비리데 (<i>Trichoderma atroviridae</i>)	211304	0.002678 788	0.0003119 3	D
NCU00801	체토미움 글로보숨 (<i>Chaetomium globosum</i>)	XP_001220469 .1	0.005890 909	0.0001028 5	OK
NCU00801	트렐라 메센테리카 (<i>Tremella mesenterica</i>)	63529	0.004381 818	0.0011575 1	D
NCU00801	헤테로바시디온 안노숨 (<i>Heterobasidion. annosum</i>)	105952	0.002751 515	0.0006876 3	D
NCU00801	크레폰넥트리아 파라시티카 (<i>Cryphonectria parasitica</i>)	252427	0.022503 03	0.0002169 2	D
NCU00801	트리코데르마 레세이 (<i>Trichoderma ressei</i>)	67752	0.003672 727	0.0006623 3	D
NCU00801	아스페르길루스 클라바투스 (<i>Aspergillus clavatus</i>)	XP_001268541 .1	0.014381 818	0.0005961 3	OK
NCU00801	네우로스포라 디스크레타 (<i>Neurospora discreta</i>)	77429	0.007060 606	0.0011056 6	D
NCU00801	트리코데르마 레세이 (<i>Trichoderma ressei</i>)	3405	0.003264 646	0.0010339 98	D
NCU00801	스포로트리쿰 서모필 (<i>Sporotrichum thermophile</i>)	43941	0.013654 545	0.0043153 4	-
NCU00801	네우로스포라 크라사 (<i>Neurospora crassa</i>)	XP_963801.1	0.048754 872	0.0035401 7	-
NCU05853	체토미움 글로보숨 (<i>Chaetomium globosum</i>)	XP_001226269 .1	0.003593 939	0.0006230 6	OK
NCU05853	트리코데르마 레세이 (<i>Trichoderma ressei</i>)	46819	0.002042 424	0.0000859 24	D

[0355]

NCU05853	<i>마이코스페렐라 그라미니콜라</i> (<i>Mycosphaerella graminicola</i>)	68287	0.002901 01	0.0006012 3	D
NCU05853	<i>아스페르길루스 플라부스</i> (<i>Aspergillus flavus</i>)	AFLA_000820 A	0.003078 788	0.0020913 2	-
-	없음	-	0.0026	0.0001	-
NCU00809	<i>피치아 스티피티스</i> (<i>Pichia stipitis</i>) CBS6054 (LAC1)	XP_001383110 .1 /GI:12613317 0	도 27 참조	-	-
NCU00809	<i>피치아 스티피티스</i> (<i>Pichia stipitis</i>) CBS6054 (LAC2)	XP_001387231 .1 /GI: 12627633 7	도 27 참조	-	-
NCU00809	<i>피치아 스티피티스</i> (<i>Pichia stipitis</i>) CBS6054 (LAC3)	XP_001383677 .2 /GI: 15086472 7	도 27 참조	-	-
NCU08114	<i>피치아 스티피티스</i> (<i>Pichia stipitis</i>) CBS6054 (HXT2.1)	XP_001386873 .1 /GI: 12627557 1	도 27 참조	-	-
NCU05853	<i>피치아 스티피티스</i> (<i>Pichia stipitis</i>) CBS6054 (HXT2.3)	XP_001382754 .1 /GI:12613245 8	도 27 참조	-	-
NCU08114	<i>피치아 스티피티스</i> (<i>Pichia stipitis</i>) CBS6054 (HXT2.4)	XP_001387757 .1 /GI: 12627393 9	도 27 참조	-	-
NCU08114	<i>피치아 스티피티스</i> (<i>Pichia stipitis</i>) CBS6054 (HXT2.5)	XP_001385684 .1 /GI:12613832 2	도 27 참조	-	-
NCU08114	<i>피치아 스티피티스</i> (<i>Pichia stipitis</i>) CBS6054 (HXT2.6)	XP_001384653 .2 /GI: 15086543	도 27 참조	-	-

* 틀림(Wrong) = 실험한 서열 및 NCBI 또는 JGI 데이터베이스에서의 서열 사이의 차이 ;

I = 실험된 서열에서의 삽입 ; D = 실험된 서열에서의 결실(deletion) ; P = 실험된 서열에서 점 돌연변이(point mutation) ; OK = 실험한 서열 및 NCBI 또는 JGI 데이터베이스에서의 서열 사이에 차이가 없음 ; 기타(Other) = 실험된 서열에서 삽입, 결실 및 점 돌연변이를 제외한 시퀀싱(sequencing)에서의 기타 문제 ; “-“ = 아직까지 이용가능하지 않은 결과[진행중인 연구(study in progress)].

¥ 접근 번호(accession numbers)가 가능하지 않을 때, JGI 번호를 사용하였다. JGI 번호는 이러한 유기체에 대한 JGI 게놈 포털(portal)을 통해 유전자 서열에 접근을 가능하게 한다(하기의 페이지로부터 접근 가능 : genome.jgi-psf.org/programs/fungi/index.jsf). A. *플라부스*(*flavus*) 및 A. *니둘란스*(*nidulans*) 식별자(identifiers)는 각각의 웹페이지 cadre-genomes.org.uk/ 및 웹페이지 broadinstitute.org/annotation/genome/aspergillus_group/MultiHome.html 에서 이들의 게놈 포털을 통해 유전자에 접근을 가능하게 한다.

[0356]

[0357]

특정한 경우에, 클론된 오솔로그의 서열은 정확하게 결정되었고, 이러한 클론이 발현된 효모는 셀로비오스를 활용할 수 있었다. 따라서, 이러한 클론, *피치아 스티피티스*로부터의 LAC2 및 *아스페르길루스 클라바투스*로부터의 XP_001268541.1 을 기능적인 셀로비오스 운반체임을 확인해주었다. 그 밖의 클론의 셀로비오스 수송 기능의 테스트는 여전히 진행 중이다(in progress). 데이터베이스에서 발행된 서열(the published sequences)과 상이한 서열을 갖는 클론된 오솔로그(Cloned orthologs)(예를 들어, 삽입, 결실 등을 갖는 하나)(표 14)는, 재-클론되고(re-cloned), 다시 배열 순서를 밝히고(re-sequences), *S. 세레비시아*에서 이들이 발현되고 성장물을 모니터링하여 셀로비오스 운반체 활성도에 대한 유사하게 테스트되었다.

[0358]

NCU00801, NUC08114 의 정렬(alignment) 및 이러한 운반체의 기능적인 오솔로그를 도 28 에 나타내었다. 도 28A 에서의 정렬은 추정적이고 확인된 셀로텍스트린 운반체를 둘 다를 포함하는 반면에, 도 28B 에서의 정렬은 오직 확인된 셀로텍스트린 운반체만을 포함한다. 게다가, 도 28C 는 NCU00801 및 NCU08114 의 정렬을 나타낸다. 두 개의 운반체는 29 % 의 아미노산 서열 동일성(identity)을 공유한다(share).

- [0359] 셀로텍스트린 운반체 기능에 대한 중요한 모티프(Motifs)는 당과 운반체 사이의 다수의 서열 정렬(multiple sequence alignments)의 육안검사(visual inspection)에 의해 확인되었다. 명확하게, 셀로텍스트린 운반체에 공통되는 모티프는 추정상의 셀로텍스트린 운반체 오솔로그 및 확인된 셀로텍스트린 운반체 사이의 T-COFFEE 에서 생성된 다수의 서열 정렬로부터 확인되었다. 이러한 모티프가 셀로텍스트린 운반체에 대해 대체로 특유한지를 확인하기 위해, *S. 세레비시아*의 육탄당 운반체, 인간 글루코스 운반체, Glut1 및 T-COFFEE 에서 생성된 두 가지의 *N. 크라사* 단당류 운반체 사이의 다수의 서열 정렬로부터 이들의 부재를 확인하였다.
- [0360] 확인된 모티프는 하기에 기재되었다. 모티프에서, NCU00801 의 기능에 중요한 것으로 발견된 잔기를 밑줄을 그었다. NCU08114 의 기능에 중요한 잔기를 상첨자(superscript) “⁺” 로 표시하였다. 둘 다의 운반체의 기능에 중요한 잔기를 상첨자 “*” 로 표시하였다. 모든 모티프를 PROSITE 표기법(notation)을 사용하여 나타내었다. PROSITE 모티프를 읽는 방법의 예로서, 하기의 모티프 [AC]-x-V-x(4)-{ED}는 [Ala 또는 Cys]-어떤 것(any)-Val-어떤 것-어떤 것-어떤 것-어떤 것-{Glu 또는 Asp 외의 어떠한 것}({any but Glu or Asp}) (SEQ ID NO: 13)을 의미한다.
- [0361] 모든 당 운반체와 같은 셀로텍스트린 운반체는 12 개의 트랜스멤브레인 α-헬릭스(transmembrane α-helices)를 갖는다. 셀로텍스트린 운반체의 N- 및 C-말단은 모두 세포내(both intracellular)이다.
- [0362] 트랜스멤브레인 헬릭스 1 앞의 서열은 현저한 특징(distinguishing features)을 가지지 않았다.
- [0363] 트랜스멤브레인 헬릭스 1 은 모티프 [L⁺IVM]-Y⁺-[FL]-x(13)-[YF]-D⁺ (SEQ ID NO: 1)을 포함한다.
- [0364] 트랜스멤브레인 헬릭스 2 는 모티프 [YF]-x(2)-G⁺-x(5)-[PVF]-x(6)-[DQ]⁺ (SEQ ID NO: 2)을 포함한다.
- [0365] 트랜스멤브레인 헬릭스 2 및 트랜스멤브레인 헬릭스 3 을 연결하는 루프는 모티프 G⁺-R⁺-[RK]⁺ (SEQ ID NO: 3)을 포함한다.
- [0366] 트랜스멤브레인 헬릭스 3 은 현저한 특징을 가지지 않는다.
- [0367] 트랜스멤브레인 헬릭스 4 는 현저한 특징을 가지지 않는다.
- [0368] 트랜스멤브레인 헬릭스 5 는 모티프 R⁺-x(6)-[YF]⁺-N⁺ (SEQ ID NO: 4)을 포함한다.
- [0369] 트랜스멤브레인 헬릭스 6 은 모티프 W⁺R-[IVLA]-P-x(3)-Q (SEQ ID NO: 5)을 포함한다.
- [0370] 트랜스멤브레인 헬릭스 6 및 트랜스멤브레인 헬릭스 7 사이의 서열은 모티프 P⁺-E⁺-S⁺-P⁺-R-x-L-x(8)-A-x(3)-L-x(2)-Y⁺-H⁺ (SEQ ID NO: 6)을 포함한다.

- [0371] 트랜스멤브레인 헬릭스 7 은 모티프 $F^{\dagger}-[GST]-Q^{\ast}-x-S^{\dagger}-G-N^{\dagger}-x-[LIV]$ (SEQ ID NO: 7)을 포함한다.
- [0372] 트랜스멤브레인 헬릭스 8 는 현저한 특징을 가지지 않는다.
- [0373] 트랜스멤브레인 헬릭스 9 는 현저한 특징을 가지지 않는다.
- [0374] 트랜스멤브레인 헬릭스 10 및 트랜스멤브레인 헬릭스 11 및 이들 사이의 서열은 모티프 $L-x(3)-[YIV]^{\dagger}-x(2)-E^{\ast}-x-L-x(4)-R-[GA]-K^{\dagger}-G$ (SEQ ID NO: 8)을 포함한다.
- [0375] 트랜스멤브레인 헬릭스 12 는 현저한 특징을 가지지 않는다.
- [0376] 트랜스멤브레인 헬릭스 12 뒤의 서열은 현저한 특징을 가지지 않는다.
- [0377] NCU00801 및 NCU08114 의 상동 모델(Homology models)은, zhanglab.ccmb.med.umich.edu/I-TASSER/(Roy 등의, 2010)에서 I-TASSER 서버를 사용한 NCU00801 and NCU08114 의 일차아미노산 서열(primary amino acid sequences)로부터 생성되었다. I-TASSER 에 의해 생산된 최고의 구조적인 모델(top structural models)은 PYMOL(웹페이지 pymol.org/)에서 확인하였다. 모티프의 매핑(Mapping)을 PYMOL 에서 또한 실시하였다. 표시된 셀로텍스트린 운반체 모티프를 갖는 NCU00801 및 NCU08114 의 상동 모델(homology models)을 도 29(A,B)에 나타내었다. 도 29C 는 NCU00801 및 NCU08114 의 예측된 2 차 구조(secondary structures)를 나타낸 것이다.

실시예 11

- [0378] **사카로미세스 세레비시아에서 네우로스포라 크라사 및 피치아 스티피티스로부터의 신규한 오타당-특이적인 운반체의 특징**
- [0379] 본 실시예에서, 생물정보학적 접근(bioinformatics approach)은 *N. 크라사* 및 *P. 스티피티스*에서 신규한 오타당-특이적인 운반체를 확인하기 위해 사용되었다.
- [0380] 오타당-특이적인 운반체의 게놈 발굴(genome mining)
- [0381] 생물정보학적 연구
- [0382] 신규한 D-자일로스-특이적인 운반체를 발견하기 위해서, *C. 인티메디아*로부터의 D-글루코스/D-자일로스 공수송체 Gxsl(Leandro 등의, 2006) 및 *P. 스티피티스*로부터의 특징지어지지 않은 추정상의 L-아라비노스-양성자 공수송체(symporter) Aut1[로커스 태그(locus tag) PICST_87108]를 코딩하는 유전자를 두 가지의 효율적인 자일로스-활용 종(xylose-utilizing species), *N. 크라사* 및 *P. 스티피티스*의 배열순서가 밝혀진 게놈(sequenced genomes)에 대해서 BLAST 검색(웹페이지 ncbi.nlm.nih.gov/)에서의 프로브(probes)로서 사용하였다(Galagan 등의, 2003; Jeffries 등의, 2007). 알려진 D-글루코스 운반체 활성화도 또는 당 수송 외의 활성도를 갖는 모든 단백질은 분석에서 제거되었다. 25 % 최소의 서열 동일성의 차단(cut-off)을 사용하여, *P. 스티피티스*로부터의 *AUT1* 뿐만 아니라 17 가지의 추정상의 오타당 운반체 유전자를 확인하였다(표 15). 이러한 추정상의 오타당 운반체 유전자는 *C. 인티메디아*로부터의 *GXS1* 또는 *P. 스티피티스*로부터의 *AUT1* 와 20-25 % 동일성을 공유한다. 모든 17 가지 추정상의 오타당 운반체는 당-수송 단백질 또는 알려지지 않는 활성화도(unknown activity)를 갖는 가설 단백질(hypothetical proteins)로서 주석을 달았다(annotated). *P. 스티피티스*로부터의 D- 글루코스 운반체 유전자 *SUT1* 및 *SUT2* 는 비교를 위해 또한 클론하였다.

[0383]

[표 15]

(a) 프로브(probe)로서 *P. 스티피티스*로부터의 *AUT1* 및 (b) 프로브로서 *C. 인터메디아* (*intermedia*)로부터의 *GSX1*를 사용하여 BLAST로부터 획득한 추정상의 오타당 운반체

a. 프로브로서 *AUT1*을 사용한 BLAST 검색 결과

이름	기원 (Origin)	AUT1와의 유사성 (identity) %	NCBT 으로부터의 아노테이션(Annotation)	길이 (cDNA)	로커스 태그 (Locus Tag)
Ap31/SUT2	<i>P. 스티피티스</i> (<i>P. stipitis</i>)	31	당 흡수(잠정적인)	1653	ABN6626.6
Ap26/XP_001387242	<i>P. 스티피티스</i>	26	당 운반체	1404	XP_001387242
AN49/NCU01494	<i>N. 크라사</i> (<i>N. crassa</i>)	49	MFS 당 운반체와 유사한, 가설단백질 NCU01494	2025	EAA2669I
AN41/NCU09287	<i>N. 크라사</i>	41	갈락토스-양성자 동반 수송체(symporter)와 유사한, 가설단백질 NCU09287	1968	EAA28903
AN29-2/NCU04963	<i>N. 크라사</i>	29	MFS 단당류 운반체와 유사한 가설단백질 NCU04963	1584	EAA30175
AN28-3/NCU02188	<i>N. 크라사</i>	28	가설단백질 NCU02188, 보존 가설단백질	1458	EAA30346
AN25/NCU00821	<i>N. 크라사</i>	25	당 운반체	1689	EAA35128

[0384]

b. 프로브로서 GSXI 를 사용한 BLAST 검색 결과

이름	기원 (Origin)	AUT1 와의 유사성 (identity) %	NCBT 으로부터의 아노테이션(Annotation)	길이 (cDNA)	로커스 태그 (Locus Tag)
Xy50/NCU 04537	<i>N. 크라사</i> (<i>N. crassa</i>)	50	단당류 운반체와 유사한 가설단백질 NCU04537	1626	EAA26741
Xy31/NCU 06138	<i>N. 크라사</i>	31	MFS 단당류 운반체와 유사한 가설단백질 NCU06138	1757	EAA30764
Xy33/NCU 00988	<i>N. 크라사</i>	33	MFS 퀴네이트(quininate) 운반체와 유사한, 가설단백질 NCU00988	1614	EAA34662
Xyp37/SUT 3	<i>P. 스테피티스</i> (<i>P. stipitis</i>)	37	당 흡수(잠정적인)	1653	ABN6799 0
Xyp33/XUT 3	<i>P. 스테피티스</i>	33	당 운반체, 추정상의 자일로스 흡수(잠정적인); 예측 운반체 [주요한 촉진 상과(major facilitator superfamily)]	1656	EAZ63115
Xyp32/XUT 1	<i>P. 스테피티스</i>	32	당 운반체, 높은 친화도, 추정상의; 자일로스 흡수(잠정적인)	1701	ABN6755 4
Xyp30/STL1	<i>P. 스테피티스</i>	30	당 운반체, 강하게 보존 (strongly conserved)	1590	ABN6574 5
Xyp31/XUT 2	<i>P. 스테피티스</i>	31	당 운반체, GXSI (STL1)와 유사한 자일로스 운반체(잠정적인)	1407	AAVQOI 000002
Xyp29/STL1 2/XUT6	<i>P. 스테피티스</i>	29	당 운반체, 추정상의 (STL12); 자일로스 흡수 (잠정적인)	1641	ABN6856 0
Xyp30- 1/HGT3	<i>P. 스테피티스</i>	30	높은 친화도 자일로스 운반체(추정상의), 자일로스 흡수(잠정적인)	1587	ABN6868 6
Xyp28/XUT 7	<i>P. 스테피티스</i>	28	자일로스 운반체, 높은 친화도, STL13 과의 추정상의 유사성(similarity), 높은 친화도 당 운반체	1257	EAZ63044

[0385]

[0386]

추정상의 오탄당 운반체의 클로닝(cloning)

[0387]

N. 크라사 및 *P. 스테피티스*를 탄소원으로 D-자일로스 또는 L-아라비노스 둘 중 하나로 보충된 리치 배지 (rich media)에서 배양하였다(cultivated). 전체 RNA 를 분리하였고, cDNA 내로 역전사하였다(reverse transcribed). 중합효소 연쇄반응(PCR)은 cDNA 로부터 직접적으로 추정상의 운반체 유전자를 증폭시키는데 사용하였다. 그러나 균종(fungal species)에서의 오탄당 운반체의 조절 메카니즘 및 발현 패턴은 잘 알려지지 않았기 때문에, 배양 조건(cultivation conditions)에서 변경(alteration)에도 불구하고 추정상의 오탄당 운반체를 코딩하는 cDNA 를 항상 획득할 수 있는 것은 아니다. 이러한 경우에서, 프라이머를 GenBank 로부터의 이에 상응하는 cDNA 서열에 따라 설계하였고(designed), 주형(template)으로서 게놈의(genomic) DNA 를 갖는 엑손(exons)을 증폭시키는데 사용된다. 오버랩-연장 PCR(Overlap-extension PCR)은 전체 길이의 유전자(full-length genes) 내로 엑손을 모으는데 사용된다. 효모 상동의 재조합-매개 DNA 조립 방법을 사용한 pRS424-HXT7-GFP 서플렉터 내로 결과적으로 생성된 PCR 산물을 클론하였다(Shao 등의, 2009). 이러한 플라스미드에서, HXT7 프로모터, 둘 다의 말단에서의 EcoRI 사이트를 갖는 측면에 배치된 GFP 유전자 및 HXT7 종결자를 *ClaI* 및 *BamHI* 에 의해 선모양으로 만든(linearized) pRS424 서플렉터(New England Biolabs) 내로 조립하였다. HXT7 프로모터 및 종결자와 서열 동일성(sequence identity)을 공유하는 DNA 단편과 함께 측면 배치된 추정상의 오탄당 운반체의 PCR 산물(도 30A)을, 표준 리튬 아세테이트 방법을 사용한 pRS424-HXT7-GFP 소화된 EcoRI 를 갖는 Euroscarf (독일 프랑크푸르트에 위치한)로부터 입수한 *S. 세레비시아* CEN.PK2-1C 스트레인(MAT α *leu2-3,112 ura3-52*, *trp1-289*, *his3- Δ 1 MAL2-8c*) 내로 공동으로 이동되었다(co-transferred). 결과적으로 생성된 형질 전환 혼합물(transformation mixture)을 2 % 글루코스로 보충된 SC-Trp 플레이트에 도말하였다(plated).

[0388]

Zymoprep Yeast Plasmid Miniprep II(미국 캘리포니아 오렌지에 위치한 Zymo Research)를 사용한 형질전환체로

부터 분리된 효모 플라스마를 대장균 DH5a 세포 (미국 일리노이주 어배너에 위치한 일리노이드 대학의 어배너-삼페인, Cell Media Facility)내로 재-이동시켰다. 플라스미드를 QIAprep Spin Miniprep Kit(미국 캘리포니아 발렌시아에 위치한 QIAGEN)을 사용하여 분리하였고, 그리고 난 다음에 원본 운반체 유전자를 증폭시키는데 사용된 프라이머를 갖는 진단 PCR(diagnostic PCR)에 의해 체크하였다. 전체적인 오픈리딩프레임(entire open reading frames)을 정확한 구성(미국 일리노이드 어배너에 위치한 일리노이드 대학의 어배너-삼페인, Core Sequencing Facility)을 확인하기 위한 DNA 시퀀싱(DNA sequencing) 또한 제출하였다. DNA 시퀀싱 결과(DNA sequencing results)를 Sequencher 4.7(미국 미시간주 앤 아버에 위치한 Gene Codes Corporation)를 사용한 데이터베이스에서의 유전자 서열과 비교하였다. 클론된 추정상의 운반체의 모든 서열은 SEQ ID NOs: 33-52 에 작성하였다.

[0389] 플라스미드를 유지하기 위해 합성 드롭아웃 배지(synthetic dropout media)에서 효모 스트레인을 배양하였다 [0.17% 아미노산 및 황산암모늄이 없는 Difco 효모 질소 염기(yeast nitrogen base), 0.5% 황산암모늄, 0.05% 아미노산 드롭아웃 믹스(amino acid drop out mix)]. 2 % 의 당으로 보충된 YPA 배지를 플라스미드를 함유하지 (harboring) 않는 효모 스트레인을 성장시키는데 사용하였다[1 % 효모 추출물, 2 % 펩톤, 0.01 % 아데닌 헤미술페이트(adenine hemisulfate)]. *S. 세레비시아* 스트레인을 호기성 성장(aerobic growth)에서 30 °C 및 250 rpm 으로 배양시켰고, 산소-제한된 조건에서 30 °C 및 100 rpm 으로 배양시켰다. 효모 스트레인을 다른 방식으로 특정하지 않았으면 세포 조작(cell manipulation)을 위한 호기성 조건 하에서 성장시켰다. 대장균 스트레인을 Luria broth(LB)(미국 펜실베이니아 피치버그에 위치한 Fisher Scientific)에서 37 °C 및 250 rpm 으로 배양하였다. 모든 제한 효소를 New England Biolabs(미국 매사추세츠 입스위치에 위치)로부터 입수하였다. 모든 화학물질을 Sigma Aldrich(미국 미주리주 세인트루이스에 위치) 또는 Fisher Scientific 로부터 입수하였다.

[0390] 클론된 추정상의 운반체에 대한 운반체 활성화도 검정

[0391] 오탄당의 세포 내의 축적

[0392] 클론된 추정상의 오탄당 운반체를 *S. 세레비시아* 당 운반체 결실 스트레인에서 과잉-발현되었고(over-expressed), 오탄당의 흡수를 측정하였다. 추정상의 오탄당 운반체의 D-자일로스-흡수 능력을 세포 내의 D-자일로스 및 자일리톨 농도의 요약(summation)으로 결정하였다. *S. 세레비시아* 세포 내에 축적된 D-자일로스가 내인성 알도스 환원효소(endogenous aldose reductase)의 존재로 인하여 부분적으로 자일리톨로 전환될 수 있다. D-자일로스 및 자일리톨 둘 다는 삼투현상을 사용하여 추출하였고, 고성능 액체 크로마토그래피(high performance liquid chromatography, HPLC)를 사용하여 분석하였다.

[0393] 교수 E. Boles' laboratory 로부터의 선물(Institut für Mikrobiologie, Heinrich-Heine-Universität, Universitätsstr. 1, Geb. 26.12.01, D-40225 Düsseldorf, Germany)인 당 운반체 녹아웃(knock-out) *S. 세레비시아* 스트레인 EBY.VW4000(CEN.PK2-1c $\Delta hxt1$ -17, $\Delta stl1$, $\Delta agt1$, $\Delta ydl247w$, $\Delta yjr160c$, $\Delta gal2$)은 *HXT1*-17 및 *GAL2* 를 포함하는 20 가지 보다 많은 당 운반체의 공존하는 녹아웃(concurrent knock-outs) 및 감지기(sensors)를 갖는다. 단일 탄소원으로서 D-글루코스에서의 성장은 이러한 스트레인에서 완전하게 중단된(abolished) 반면에, 상이한 당 수송 시스템(sugar transport system)을 통한 말토스의 흡수는 유지되었다. EBY.VW4000 스트레인은, 재조합 D-자일로스 흡수를 테스트하기 위해 이를 적절한 숙주로 만드는, HPLC 검정 조건 하에서 최소의 오탄당-흡수를 또한 나타내었다. 클론된 추정상의 오탄당 운반체 유전자가 과잉-발현된(over-expressing) 플라스미드를, 표준 리튬 아세테이트 방법을 사용한 EBY.VW4000 스트레인 내로 이동시켰고, 단일 콜로니(single colonies)를 당 흡수 활성화도를 측정하기 위해 사용하였다.

[0394] 세포를 2 % 말토스로 보충된 2 mL SC-Trp 배지에서 첫 번째로 배양하였다. 그리고 난 다음에 종균배양(Seed culture)을 250 mL 의 플라스크에 50 mL 배양물을 배양하는데 사용하였다. 세포를 24 시간 성장한 후에 원심분리하여 채취하였고, 10 의 최종 OD600 에 대한 2 % D-자일로스 또는 L-아라비노스로 보충된 YPA 배지에서 재-현탁시켰다(re-suspended). 30 분, 60 분, 120 분 및 24 시간에서, 5 mL 배양물을 세포내 당 농도를 측정하기 위해 가져왔다. 배양물 샘플을 얼음처럼 차가운 물로 2 번 세척하고 3 mL 의 탈이온수에서 재-현탁시켰다. 세포 현탁액을, 세포내의 당을 추출하기 위해 2 일 동안 37 °C 에서 250 rpm 으로 교반시키면서 배양하였다. 결

과적으로 생성된 세포 현탁액을 HPLC 분석 전에 0.22 μ m PES 필터(미국 매사추세츠주 로웰에 위치한 Corning)를 통해 여과하였다. 당 및 이에 해당하는 당 알코올(하기에 기재된)의 농도를, 제조업자의 프로토콜(manufacturer's protocol)에 따른 Shimadzu ELSD-LTII 저온도-증발에 의한 광산란 검출기(low temperature-evaporative light scattering detector)(Shimadzu) 및 BioRad HPX-87C 컬럼(미국 캘리포니아주 허큘러스에 위치한 BioRad Laboratories)이 갖추어진 Shimadzu HPLC 을 사용하여 결정하였다. 당-흡수 활성도를 OD ~ 10 에서, 세포 배양물의 mL 에 대한 삼투압을 통해 추출된 당의 mg 으로서 계산하였다.

[0395] 몇몇의 추정상의 오타당 운반체는, D-글루코스 또는 D-자일로스를, 또는 둘 다의 흡수에서 유효한 것을 확인되었다. D-글루코스는 효모 내에서 한번 대사작용을 할 수 있기 때문에(metabolized), D-글루코스 수송 활성도는 세포내의 D-글루코스 농도를 측정하여 결정할 수 없다. 그러나, EBY.VW4000 스트레인은 단일 탄소원으로서 D-글루코스를 포함하는 배지에서 성장할 수 없기 때문에, D-글루코스에서 추정상의 오타당 운반체로 형질전환된(transformed) 스트레인의 성장은, 추정상의 오타당 운반체가 D-글루코스 수송 활성도를 갖음을 나타낸다.

[0396] SUT3(Xyp37), XUT3(Xyp33), SUT2(Ap31), NCU04963(An29-2) 및 NCU06138(Xy31)의 도입(Introduction)은 D-글루코스에서 EBY.VW4000 스트레인의 성장을 회복시키고(restored), 따라서 글루코스 수송 활성을 가능하게 한다. SUT3, XUT3, SUT2 및 NCU04963 은 자일로스 수송 활성도를 또한 가지는 반면에, NCU04963 및 NCU06138 은 아라비노스 수송 활성도를 나타내었다(도 31). 추정상의 운반체의 나머지는 D-글루코스에서 성장을 가능하지 못하게 하고, 이들 중 대부분은 어떠한 오타당 수송 활성도를 나타내지 않았다. 그러나 NCU00821 및 STL12/XUT6 은 자일로스 수송 활성도를 나타내지만, XUT1 은 아라비노스 수송 활성도를 나타내었고, 이는 이들인 오타당에 대해 특이적인 당 운반체임을 나타낸다(도 32).

[0397] *P. 스티피티스*로부터의 STL12/XUT6 및 XUT1, 및 *N. 크라사*로부터의 NCU00821 가 D-글루코스-흡수 활성도를 갖지 않는 실제로 오타당-특이적인 운반체임을 추가적으로 확인하기 위해, 당-흡수 검정을, 기질로서 14 C-표지된 D-글루코스, D-자일로스 및 L-아라비노스를 사용하여 실행하였다. STL12/XUT6 및 NCU00821 만이 과잉-발현된(over-expressing) EBY.VW4000 스트레인의 D-글루코스- 및 L-아라비노스-흡수 활성도는, 둘 다의 운반체의 D-자일로스-흡수 반응속도를 측정하기 위해 사용된 검정 조건 하에서 측정에 대해서 매우 낮은 것으로 발견되었다.

[0398] 14 C-표지된 D-글루코스, L-아라비노스 및 D-자일로스를 90 % 에탄올에서의 용액으로서 American Radiolabeled Chemicals(미국 미주리주 세인트루이스에 위치)로부터 입수하였다. 방사능 표지된 당(Radiolabeled sugars)을 화학적 후드(chemical hood)에서 첫 번째로 건조시키고 난 다음에 물에 재-현탁시켰다. 대략 40,000 dpm/ μ L의 특이적인 방사능(radioactivity)을 갖는 1.33 M 및 1 M 의 농도 및 약 20,000 dpm/ μ L 의 특이적인 방사능(radioactivity)을 갖는 500 mM, 350 mM, 250 mM, 100 mM 및 50 mM 의 농도의 당 용액(Sugar solutions)을 당-흡수 검정에서 사용하였다. 지수기(exponential phase)에서 세포 배양물을 채취하였고, 얼음처럼 차가운 물로 두 번 세척하고, pH 5 의 100 mM 트리스-시트레이트 완충용액(Tris-citrate buffer)에 대해 mL 당 약 60 mg 의 세포 건조무(dry cell weight, DCW)에 재현탁시켰다. 160 μ L 의 세포 현탁액의 세 개의 앨리퀀트(aliquots)를 DCW 를 측정하기 위해 24 시간 동안 65 $^{\circ}$ C 에서 건조시켰다. 세포 현탁액의 나머지를 사용하기 전에 얼음에 보관하였다. 당-흡수 검정에 대해서, 세포 현탁액을 검정하기 전에 5 분 동안 30 $^{\circ}$ C 에서 평형을 유지시켰다. 50 mL 의 원추형 튜브(conical tube)에서, 160 μ L 의 세포 현탁액을 40 또는 60 초(정확한 시간) 동안 방사능-표지된 당 용액(radio-labeled sugar solution)을 혼합하였다. 주사기(syringe)로 운반된(delivered) 얼음처럼 차가운 10 mL 의 물을 첨가하여 반응을 정지시켰다. 제로 시점 샘플(zero-time-point sample)을 얼음처럼 차가운 물 및 방사능-표지된 용액을 함유하는 배양 튜브(culture tube)에 동시에 첨가하여 수득하였다. 혼합물을, 40 % 의 당 용액에 미리-흡수 젖은(pre-soaked) Whatman GF/C 필터(미국 뉴저지 플로함 파크에 위치한 Whatman)를 통해 즉시 여과하였고, 15 mL 의 얼음처럼 차가운 물(ice-cold water)로 세척하였다. 필터를 3 mL 의 Econo I 섬광 각테일(scintillation cocktail)(Fisher Scientific)에 두었고, 1 분 동안 Beckman LS6500 신틸레이션 계수기(scintillation counter)(미국 캘리포니아주 브레아에 위치한 Beckman Coulter)를 사용하여 계수하였다. 모든 측정점(data points)는 세 가지의 독립적인 실험으로 측정하였다. 당-흡수 비율을 mmol 의 수송

된 당/시간/그램의 세포건조무(dry cell weight)로 계산하였다.

[0399] STL12/XUT6, NCU00821 또는 XUT1 이 과잉-발현된 EBY.VW4000 스트레인에서 D-자일로스와 L-아라비노스 둘 다의 세포내 축적(Intracellular accumulation)은 HPLC 를 사용하여 또한 측정하였다. 30 분, 60 분, 120 분 및 24 시간 동안의 오탄당 당(pentose sugars)과 함께 배양한 세포 배양물을 HPLC 로 분석하였다. STL12/XUT6 또는 NCU00821 가 과잉-발현된 EBY.VW4000 스트레인은 D-자일로스 흡수 활성도를 나타낸 반면에, XUT1 과잉-발현된 스트레인은 24 시간 배양 후에 L-아라비노스-흡수 활성도를 나타내었다(도 33).

[0400] 세포내 당 축적의 HPLC 분석을 포함하여 ¹⁴C-표지된 당 흡수 검정(sugar uptake assay)은, 리그노셀룰로직계 가수분해물(lignocellulosic hydrolysates), D-글루코스, D-자일로스 및 L-아라비노스에서 세 가지 가장 풍부한 단당류 중에서, STL12/XUT6 및 NCU00821 은 D-자일로스 흡수에 대한 책임이 있고, XUT1 은 L-아라비노스 흡수에 대해 책임이 있다. 중요한 것은, D-자일로스 흡수에 대한 효모에서 연구된 대부분 당 운반체는 D-자일로스 보다는 D-글루코스에 대해 보다 높은 흡수 활성도를 갖는다. 적응진화(adaptive evolution) 후에 *트리코테르마 레세이*로부터의 Trx1t1 만이 D-자일로스-특이적인 흡수 활성도를 나타내었다(Saloheimo 등의, 2007). 이러한 데이터는, *P. 스티피티스*로부터의 STL12/XUT6 및 *N. 크라사*로부터의 NCU00821 는 *S. 세레비시아* 내로 도입된(introduced) 첫 번째 두 가지의 실험적으로 확인된 자연 발생적으로(naturally)-발생한(occurring) D-자일로스-특이적인 운반체임을 나타낸다. 마찬가지로, *P. 스티피티스*로부터의 XUT1 은, *S. 세레비시아* 내로 도입된(introduced) 첫 번째 실험적으로 확인된 자연 발생적으로(naturally)-발생한(occurring) L-아라비노스-특이적인 운반체임을 나타낸다.

[0401] **동역학적 파라미터(Kinetic parameters)**

[0402] ¹⁴C-표지된 당-흡수 검정을 사용하여, NCU00921, STL12/XUT6 및 XUT1 를 통한 D-자일로스 수송의 동역학적 파라미터를 확인하였다. 검정 조건 하에서, 당 흡수는 첫 번째 60 초 동안의 직선 범위(linear range) 내임이 관찰되었다(도 34). NCU00821, STL12/XUT6 또는 XUT1 가 과잉-발현된 EBY.VW4000 스트레인은 40 또는 60 초 동안 표지된 D-자일로스 또는 L-아라비노스와 함께 배양한 다음에, 추가적인 당 흡수를 정지시키기 위해 얼음처럼 차가운 물을 첨가하였다. 액체 섬광 계수기(liquid scintillation counter)를 사용하여 측정하기 전에 반응 혼합물을 여과한 다음에, 세척하였다. 당-흡수 비율 및 기질 농도를, Origin 소프트웨어(미국 매사추세츠주 노샘프턴에 위치한 OriginLab Corporation)를 사용한 비선형회귀(non-linear regression)에 의한 미하엘리스-멘텐 방정식(Michaelis-Menten equation)에 적용시켰다(fitted). NCU00821 또는 STL12/XUT6 만을 포함하는(harboring) EBY.VW4000 스트레인에 의해 D-자일로스 흡수에 대한 K_m 값은 각각 175.7 ± 21.4 mM 및 56.0 ± 9.4 mM 이다. 이에 해당하는 V_{max} 값은 각각 36.7 ± 2.9 및 41.5 ± 2.3 $\mu\text{mol/h/gram DCW}$ 이다. 유사하게, XUT1 을 포함하는 EBY.VW4000 스트레인에 의한 L-아라비노스 흡수에 대한 K_m 및 V_{max} 값은 각각 48.0 ± 13.2 mM 및 5.6 ± 1.6 $\mu\text{mol/h/gram DCW}$ 이다.

[0403] 자연 발생적으로-발생한 D-자일로스-동화 균종(assimilating fungal species)에서의, 높은 친화도의 D-자일로스-양성자 공동수송 시스템(symport system) 및 낮은 친화도의 D-자일로스 촉진 확산 시스템(facilitated diffusion system)이 존재한다. 이러한 두 가지 시스템의 K_m 값은 공동 수송 시스템에 대한 0.4-4 mM 및 촉진 확산 시스템에 대한 약 140 mM 인 것으로 결정되었다(Leandro 등의, 2006 ; Stambuk 등의, 2003). 이러한 값은 *S. 세레비시아*에서 D-글루코스-흡수 시스템의 친화도와 가깝고(close), 이는 높은 친화도 시스템에 대한 1.5 mM 의 K_m 및 낮은 친화도 시스템에 대한 20 mM 의 K_m 을 갖는다(Lang and Cirillo 1987 ; Ramos 등, 1988). 불행하게도, 야생형 *S. 세레비시아*의 D-자일로스 흡수 친화도는 D-글루코스에 대한 이의 친화도 보다는 10 의 두 제곱(two orders of magnitude)이나 낮았다. *S. 세레비시아*에서 D-자일로스 흡수에 대한 K_m 값은 높은 친화도 시스템에 대해 오직 190 mM 이고, 낮은 친화도에 대해 1.5 M 이다(Kotter and Ciriacy, 1993). 자연 발생적으로 발생한 D-자일로스-동화(assimilating) 효모에서 높은 친화도의 D-자일로스-흡수 시스템과 비교하였을 때, 새로이 발견된 D-자일로스-특이적인 운반체의 친화도는 보다 낮았다. 그러나, 야생형 *S. 세레비시아*에서 D-자

일로스-흡수 시스템과 비교하였을 때, NCU00821 및 STL12/XUT6 은 D-자일로스에 대한 보다 높은 친화도를 나타내었다. 특히, STL12/XUT6 및 XUT1 에 의한 D-자일로의 K_m 은 야생형 *S. 세레비시아*에서 운반체에 의한 자일로스 흡수 K_m 의 오직 4 분의 1 (one-fourth of K_m)이다. D-자일로스-특이적인 운반체의 K_m 값은 Gxf1(K_m 88 mM) 및 Sut1(K_m 145 mM)의 이들에 가깝고, 이는 재조합 *S. 세레비시아*에서 D-자일로스 발효를 증진시키는 것으로 나타났다(Runquist 등의, 2009 ; Katahira 등의, 2008). 따라서, D-자일로스 발효는, *S. 세레비시아* 내에 이러한 최근에 발견된 D-자일로스-특이적인 운반체를 도입하여 증진시킬 수 있다.

[0404] 당 운반체의 세포내 위치(cellular localization)

[0405] 당 운반체는 트랜스멤브레인 단백질이고, 세포막(cell membrane)에서의 정확한 폴딩(correct folding) 및 위치(localization)는 기능적일 수 있는 이들을 필요로 한다. 추정상의 오타당 운반체가 클론되었을 때, 신호 펩티드가 특이적으로 첨가되지 않았기 때문에, D-자일로스-특이적인 운반체가 세포막(cell membrane)에 정확하게 위치하는 것을 보장하기 위해 이는 중요하다. 이는 곰팡이 *N. 크라사*로부터 클론된 추정상의 오타당 운반체 유사 NCU00821 에 대해서 특히 적용되었고(true), 이는 *S. 세레비시아*와 비교한 매우 상이한 생리(physiology)를 나타낸다. *S. 세레비시아*에서의 D-자일로스-특이적인 운반체의 세포내 위치를 연구하기 위해, NCU00821, STL12/XUT6 및 XUT1 을 링커(linkers)를 통해 C-말단에서 녹색 형광 단백질(Green Fluorescent Protein, GFP)로 융합하였고, 이들의 위치(their localization)를 형광 영상(fluorescent imaging)으로 모니터링하였다.

[0406] C-말단에서 GFP 를 갖는 오타당-특이적인 운반체의 융합 단백질(fusion proteins)은 운반체 위치 연구(transporter localization study)를 위해 구성하였다. GS-링커[Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Ser (SEQ ID NO: 70)]를 운반체 및 GFP 사이에 도입하였다. GS-링커를 PCR 프라이머에 의한 GFP 오픈 리딩 프레임(open reading frame)의 N-말단에 첨가하였고, 이로 인하여 5'-말단에서의 운반체 및 3'-말단에서의 HXT7 종결자(terminator)에 상동인 뉴클레오티드 서열을 측면에 둔(flanked with) GS-링커-GFP 의 PCR 산물이 생성되었다. 운반체 유전자는, 5'-말단에서의 HXT7 프로모터 및 3'-말단에서의 GS-링커-GFP 와 동일한 뉴클레오티드 서열을 측면에 둔(flanked with) 운반체의 DNA 단편을 생성하기 위해 원본의 pRS424-HXT7-운반체 구조물(constructs)로부터 증폭시켰다. 그리고 난 다음에 이러한 두 가지 단편을, *EcoRI* 로 소화된(digested) pRS424-HXT7-GFP 로 *S. 세레비시아* 스트레인 CEN.PK2-1C 내로 공동-이동시켰다(co-transferred)(도 30B). 결과적으로 생성된 형질전환 혼합물을 2 % D-글루코스로 보충된 SC-Trp 플레이트에 도말하였다(plated).

[0407] 단일 콜로니를 2 % 말토스로 보충된 2 mL 의 SC-Trp 액체 배지 내로 접종하였다(inoculated). 세포 배양물을 지수기(exponential phase)에서 채취하였다. 원심관(centrifuge tube)에서, 250 μ L 의 세포 배양물을 10 분 동안 실온에서 10 μ L 의 Hoechst 33342 핵 염료(nuclei dye)(미국 캘리포니아주 칼즈바드에 위치한 Invitrogen)로 염색하였다(stained). 세포 배양물의 소량의 물방울(small droplet)을 한 장의 커버 글라스에 이동시킨 다음에, 형광 이미지(fluorescent images)를 Andor Technology Revolution System 회전원판식 공초점 현미경(Spinning Disk Confocal Microscope)(미국 일리노이주 어배너 일리노이대학교 어배너-샘페인의 Institute for Genomic Biology 에 위치한 Core facilities)을 사용하여 찍었다. 이미지를 Imaris 화상해석(image analysis) 및 시각화 소프트웨어(visualization software)(미국 미네소타주 세인트 폴에 위치한 Bitplane)을 사용하여 진행하였다.

[0408] 오타당-특이적인 운반체가 과잉-발현된(over-expressing) 효모 스트레인을 세포 주변부(cell periphery)에서 독특한 형광 헤일로(fluorescent halo)를 나타내었다(도 35). NCU00821 및 XUT1 에 대해서, 대부분의 모든 GFP 형광은 세포막에서 나타난 반면에, STL12/XUT6-과잉-발현된 세포에서의 형광의 대부분은 세포질(cytoplasm)에 잔존하였다. 이는 막 단백질(membrane protein)의 상승된 발현 때문에 STL12/XUT6 운반체의 비효율적인 유출(export)을 나타낸다. 모든 세포가 형광을 나타내지 않았고 이는 운반체의 발현이 최적이지 않음을 나타내는 점에 또한 주목하였다. 운반체 발현의 추가적인 개선은 발현 수준의 변경 및/또는 재조합 *S. 세레비시아*의 게놈 내로 운반체 유전자의 통합을 통해 성취될 수 있다.

[0409] **오탄당 운반체의 유형의 확인(Determination)**

[0410] *S. 세레비시아*에서의 당 운반체의 두 가지 유형, 공수송체(symporters) 및 촉진제(facilitators)가 있다. 공수송체에 대해서, 당 흡수는 양성자 흡수와 결합되어있다. 당 공수송체는 당에 대해서 일반적으로 높은 친화도를 나타내었다. 그 동안에, 촉진제를 통한 당 흡수는 양성자와 결합되지 않았지만, 촉진제는 낮은 당-흡수 친화도를 일반적으로 나타내었다(Leandro et al., 2006). 공수송체 검정은 EBY.VW4000 스트레인에서 발현된 NCU00821, STL12/XUT6 및 XUT1 에 대해서 실시하였다.

[0411] 운반체의 유형을 결정하기 위해, EBY.VW4000 과잉-발현된 오탄당-특이적인 운반체의 pH 변화를, USB 통신장치(communication module) 및 Direct pH 소프트웨어(미국 오하이오주 콜럼버스에 위치한 Mettler Toledo)가 갖추어진 Seven Multi pH 미터(meter)를 사용한 D-자일로스, L-아라비노스 또는 말토스를 포함하는 완충되지 않은(un-buffered) 세포 현탁액에서 측정하였다. 오탄당-특이적인 운반체를 코딩하는 플라스미드를 EBY.VW4000 스트레인 내로 이동한 후에, 2 % 말토스로 보충된 SC-Trp 플레이트에 도말하였다. 단일 콜로니를 2 % 말토스로 보충된 2 mL SC-Trp 배지에서 배양하였다. 2 L 의 플라스크에 400 mL 배양물을 접종하기 위해 종균 배양(Seed culture)을 사용하였다. 배양물을 OD ~ 1 에서 채취하였고, 얼음처럼 차가운 물(ice-cold water)에 두 번 세척하였다. 세포 펠렛(Cell pellets)을 4 mL 의 물에 재-현탁시키고 사용하기 전에 얼음에 두었다. 공수송체 검정을 위해, pH 전극을 25 °C 에서 유지시킨 50 mL 용량(capacity)의 물-자켓 비커(water-jacketed beaker)에 담그었고, 마그네틱 교반(magnetic stirring)을 제공하였다. 비커에, 25 °C 에서 평형시킨(equilibrated) 23 mL 의 탈이온수(deionized water) 및 1 mL 의 세포 현탁액을 첨가하였다. pH 를 5 에 조정하였고, 베이스라인(base line)을 획득하였다. pH 변화를 pH 5 에서 1 mL 의 50 % 당 용액을 첨가하여 기록하였다.

[0412] 도 36 은 말토스의 첨가 후에 완충되지 않은(un-buffered) 세포 현탁액에서의 pH 변화를 나타낸 것이다. 기록된 바와 같이, 완충되지 않은 *S. 세레비시아* 세포 현탁액에서 pH 는 말토스를 첨가하면서 상승되었다. pH 기록 시스템(pH recording system)이 가동되는 것을 확인하기 위해, 1 mL 의 50 % 말토스 용액을 첨가하였다. 모든 샘플에서 관찰된 pH 증가(elevation)는, pH 기록 시스템이 실험적인 세팅(experimental setting)에서 일시적인 pH 변화를 모니터링 할 수 있음을 나타낸다.

[0413] 완충되지 않은 세포 현탁액에서 pH 증가가 없는 것은 오탄당-특이적인 운반체의 어떠한 것에 대해서 관찰되었고, 이는 이러한 운반체를 통한 오탄당 흡수가 양성자 수송과 결합되어 있지 않음을 나타낸다(도 37). 따라서, NCU00821, STL12/XUT6 및 XUT1 을 오탄당 촉진제일 수 있는 것으로 결정되었다.

[0414] 이러한 결과는, NCU00821 및 STL12/XUT6 의 동역학적 파라미터는 자연적으로-발생한 D-자일로스-동화 효모(assimilating yeasts)에서 낮은 친화도 D-자일로스를 가능하게 하는 확산 시스템(low affinity D-xylose facilitated diffusion system)의 이러한 것과 유사한 사실과 일관된다. 공수송체는 D-자일로스에 대해 보다 높은 친화도를 갖는 사실에도 불구하고, 공수송체의 과잉-발현은, 양성자 구배(proton gradient)를 생성하기 위한 ATP 필요성 때문에 D-자일로스-동화 스트레인에 의한 당 이용을 항상 용이하게 할 수 있는 것은 아니다. 사실은, D-자일로스 발효에 유익한 것으로 나타난 대부분의 운반체는 촉진제이다(Runquist et al., 2009; Katahira et al., 2008).

[0415] **D-자일로스-특이적인 운반체의 이중 기원 과잉-발현(Heterologous Over-Expression)**

[0416] D-자일로스 이용 경로를 포함하는 *S. 세레비시아* 스트레인에서 활성(active) 이중 기원 D-자일로스-특이적인 운반체의 과잉-발현은, 이들의 과잉-발현이 자일로스 이용을 개선시키는지 여부를 확인하기 위해 또한 조사하였다. 자일로스 이용은 호기성 조건 하에서 셰이크플라스크(shake-flask)를 사용하여 연구하였다. 자일로스 운반체 NCU00821, NCU04963, XUT1, STL12/XUT6 및 Hxt7 이 발현된 플라스미드를, 스트레인 HZE63 내로 도입하였다[CEN.PK2 ura3::자일로스 이용 경로(xylose utilization pathway)]. 이러한 스트레인은 염색체 내로 URA3 부위(site) 내로 통합된 자일로스 이용 경로를 갖는다. 이는, *N. 크라사*로부터 자일루로스 환원효소(xylulose reductase, XR) 및 자일리톨 탈수소효소(xylitol dehydrogenase, XDH) 및 *P. 스티피티스*로부터 자일

루로키나아제(XKS)를 포함하는, 이전의 작업으로부터 플라스미드를 사용하여 구성하였다. 스트레인 HZE63 를 수득하기 위해 이러한 플라스미드를 *ApaI* 로 절단하고, 효모 스트레인 CEN.PK2 내로 형질전환시켰다(transformed).

[0417] 자일로스 운반체-코딩 플라스미드로 형질전환시킨(transformed) HZE63 스트레인을 2 % 글루코스로 보충된 SC-Ura 플레이트에 도말하여 선별하였다. 변형된 스트레인을 2 % 글루코스를 갖는 SC-Trp-Ura 에 예비-배양한 다음에, 초기의 OD₆₀₀ = 1.0 에 대한 0.5 % 또는 5 % 의 자일로스로 보충된 SC-Trp-Ura 내에 접종하였다. 세포 배양물을, 250 rpm 으로 30 °C 에서 50 mL 의 배양물을 포함하는 125 mL 진탕-플라스크(shake-flask)에서 배양하였다(도 38).

[0418] 형질전환체(transformant)의 효모 플라스미드를 *E.coli* DH5 α 세포 내로 형질전환시켰다(transformed). 정확한 구조물(correct construction)을 확인하기 위해 플라스미드를 분리한 다음에, 진단 PCR(diagnostic PCR)로 체크하였고, 시퀀싱(sequencing)을 위해 제출하였다. 플라스미드 맵(Plasmid maps)을 도 39 에서 확인할 수 있다.

[0419] 불행하게도, 오타당-특이적인 운반체의 과잉-발현의 장점은 발현 전략의 변경, 배양 조건의 변경 및 D-자일로스 이용의 선택에도 불구하고 관찰할 수 없었다. 몇 가지 가능한 이유가 있었다. 첫 번째로, 당 운반체와 같은 막 단백질의 과잉-발현은 세포 막의 온전함(integrity)에 영향을 미칠 수 있고, 그 결과에 따라 세포의 성장을 방해할 수 있다(Wagner et al., 2006). D-글루코스를 탄소원으로 사용하였을 때, 운반체 과잉-발현 스트레인은 보다 느린 성장 비율을 나타내는 것으로 관찰되었다. 글루코스-함유 SC-ura 배지에서 성장한 운반체를 전달하는 스트레인의 2-일 배양물의 최종 OD 는 오직 4 인 반면에, 음성 대조군(negative control)의 OD 는 약 6 이다. 두 번째로, 육탄당 운반체(hexose transporters)를 통한 야생형 *S. 세레비시아*의 D-자일로스-흡수 활성도는, 육탄당 운반체 녹아웃(knockout) 스트레인에서 과잉-발현된 특정 D-자일로스 운반체의 D-자일로스-흡수 활성도보다, 보다 높았다. 최근에 발견된 D-자일로스-특이적인 운반체의 낮은 당 수송 활성도는 당 흡수 능력의 개선을 관찰하기 어렵게 만들고 있다. 세 번째로, 신규한 D-자일로스-특이적인 운반체의 도입이 *S. 세레비시아* 세포 내로 D-자일로스의 흡수를 증가시키는데도 불구하고, D-자일로스 이용 경로가 당-흡수 제한 단계(sugar-uptake the limiting step)를 만들기엔 충분하게 효율적일 때, D-자일로스 이용의 효과를 또한 관찰할 수 있다. 이는, 당 운반체의 과잉-발현의 효과가 스트레인의 백그라운드(strain background) 및 배양 조건에 의존하는 것으로 나타났다(Runquist et al., 2010). 하기의 실시예 12-15 는 효모에서 자일로스 이용 경로의 최적화(optimization)를 기재하였다.

[0420] 추가적인 오타당-특이적인 운반체의 클로닝

[0421] NCU00821, STL12/XUT6 및 XUT1 의 오솔로그를 클로닝하고(cloned) 오타당 흡수를 위해 테스트하였다. 상이한 균류 스트레인(fungal strains)을 글루코스 또는 오타당으로 보충한 리치 배지(rich media)에서 배양하였다. 전체 RNA 를 분리하였고, cDNA 내로 역전사되었다. 중합효소 연쇄 반응(Polymerase chain reaction, PCR)을 cDNA 로부터 직접적으로 추정상의 운반체 유전자(putative transporter genes)를 증폭시키는데 사용하였다. 그러나, 조절 매커니즘(regulation mechanism) 및 발현 패턴(expression pattern)이 균종(fungal species)에서의 오타당 운반체에 대해서 알려지지 않았기 때문에, 추정상의 오타당 운반체를 코딩하는 cDNAs 는 배양 조건의 변경에도 불구하고 항상 수득가능하지 않다. 이러한 경우에, 프라이머는 GenBank로부터 이에 상응하는 cDNA 서열에 따라 설계되었고, 주형(template)으로서 게놈 DNA 를 사용한 엑손을 증폭시키기 위해 사용하였다. 오버랩-연장 PCR(Overlap-extension PCR)을 사용하여 전장 유전자(full-length genes) 내로 엑손을 조립하였다. 결과적으로 생성된 PCR 산물을 DNA 조립 방법(DNA assembler method)을 사용하여 HXT7 프로모터 및 HXT7 종결자(terminator)을 포함하는 pRS424 셔틀 벡터 내로 클론하였다. 형질전환체(transformants)로부터 분리된 효모 플라스미드를 *E. coli* DH5a 내로 재형질전환시키고(retransformed), 분리된 *E. coli* 플라스미드를 원본 운반체 유전자(original transporter genes)를 증폭시키는데 사용하는 프라이머를 사용하여 진단 PCR 로 처음으로 체크하였다. 전체 오픈 리딩 프레임(entire open reading frames)을, 플라스미드의 정확한 구조(correct

construction)를 확인하기 위해 시퀀싱(sequencing)을 위해 제출하였다.

[0422] 대부분의 클로닝 작업을 효소 이중 기원 재조합 매개 DNA 조립 방법(yeast homologous recombination mediated DNA assembler method)을 사용하여 실시하였다. pRS424-HXT7-GFP 플라스미드를 추정상의 오타당 운반체의 클로닝(cloning)을 위해 사용하였다. 이러한 플라스미드에서, HXT7 프로모터, 두 말단에 EcoRI 사이트를 측면에 둔 GFP 유전자 및 HXT7 말단을, ClaI 및 BamHI 에 의해 선모양으로 만든 pRS424 셔틀 벡터(New England Biolabs) 내로 조립하였다. HXT7 프로모터 및 종결자에 대해 서열 동일성을 공유하는(sharing) DNA 단편을 측면에 둔 추정상의 오타당 운반체의 PCR 산물은, 표준 리튬 아세테이트 방법(standard lithium acetate method)을 사용하여 pRS424-HXT7-GFP 로 절단한(digested) EcoRI 로 CEN.PK2-1C 내로 공동-이동시켰다(co-transferred). 결과적으로 생성된 형질전환(transformation) 혼합물을 2 % D-글루코스로 보충된 SC-Trp 플레이트에 도말하였다(plated). 형질전환체를 오타당 수송 활성도에 대해서 테스트하였다.

[0423] 결과를 도 40 및 표 16 에 나타내었다. 8 가지의 추정상의 오타당 특이적인 운반체[XP_960000 (NC52), CAG88709 (DH48), XP_457508 (DH61), XP_681669 (32-10), XP_001487429 (29-6), XP_001727326 (29-9), XP_657854 (32-8), XP_720384 (29-4)] 중에서, NC52 만이 글루코스 플레이트에서 세포 성장할 수 있었고, 이는 그 밖의 7 가지 운반체가 오타당-특이적이거나 불활성일수 있음을 암시한다. HPLC-기반 오타당 흡수 검정을 사용하여, XP_457508(DH61), XP_001727326(29-9), XP_720384(29-4) 및 XP_681669(32-10)를 포함하는 4 가지 자일로스-특이적인 운반체를 발견하였다. 게다가, 하나의 아라비노스-특이적인 운반체, XP_657854 (32-8)를 확인하였다(도 40 ; Top). 5 가지의 추가적인 추정상의 오타당 특이적인 운반체(XP_002488227, AB070824.1, XP_001389300, XP_002488227, EEQ43601.1)를 또한 테스트하였고, 이들 중 글루코스 플레이트에서 세포 성장할 수 있는 것은 없었다. 추가적인 오타당 흡수 검정은, XP_002488227 및 AB070824.1 이 자일로스 특이적인 운반체임을 나타낸다(도 40 ; Bottom). 이러한 결과의 요약은 표 16D 에 나타내었다.

[0424]

[표 16]

자일로스-특이적인 운반체 NCU00821 오솔로그(orthologs)의 클로닝(Cloning)

NCBI 참조서열	기원 (Origin)	서열(Sequence) 결과*	흡수 측정 (Uptake Assay)	상태 (Status)
XP_002488227	탈라로미세스 스티피타투스 (<i>Talaromyces stipitatus</i>)	맞음(Correct)	예(Yes)	클론 (Cloned)
XP_001400900	아스페르길루스 니가 (<i>Aspergillus niger</i>)	맞음(Correct)	예	클론
XP_001220481	체토미움 글로보숨 (<i>Chaetomium globosum</i>) CBS 148.51	없음(No)	아니요(No)	시퀀스(Sequenced), 하나의 인트론
XP_001912725	포도스포라 앤서리나 (<i>Podospora anserine</i>)	없음(No)	아니요	OE-PCR, PCR 산물 없음 (no PCR product)
XP_660079	아스페르길루스 니둘란스 (<i>Aspergillus nidulans</i>) FGSC A4	맞음(Correct)	예	클론
AAL89823	아스페르길루스 니가 (<i>Aspergillus niger</i>)	맞음(Correct)	예	클론
XP_002382573	아스페르길루스 플라부스 (<i>Aspergillus flavus</i>) NRRL3357	틀림(Wrong)	예	클론
XP_459386	데바리오마이세스 한센니 (<i>Debaryomyces hansenii</i>) CBS767	없음(No)	아니요	게놈 DNA (Genomic DNA), PCR 산물 없음
XP_001825132	아스페르길루스 오리재 (<i>Aspergillus oryzae</i>) RIB40	맞음(Correct)	예	클론
XP_001389300	아스페르길루스 니가 (<i>Aspergillus niger</i>)	맞음(Correct)	예	클론

* “맞음(Correct)” = 데이터베이스에서 클론(clone)의 서열이 서열과 일치한다 ;

“틀림(Wrong)” = 데이터베이스에서 클론의 서열이 서열과 일치하지 않는다 ;

“아니요(No)” = 이용가능하지 않은 결과[진행중인 작업(work in progress)].

[0425]

자일로스-특이적인 운반체 STL12/XUT6 오솔로그의 클로닝

NCBI 참조서열	기원 (Origin)	서열(Sequence) 결과*	흡수 측정 (Uptake Assay)	상태 (Status)
XP_457508 (DH61)	<i>데바리오마이세스 한센니</i> (<i>Debaryomyces hansenii</i>) CBS767	맞음(Correct)	아니요(No)	클론 (Cloned)
XP_002551364	<i>칸디다 트로피칼리스</i> (<i>Candida tropicalis</i>) MYA- 3404	틀림(Wrong)	아니요	없음(No)
XP_001523322	<i>로더로마이세스 엘롱기스포루스</i> (<i>Lodderomyces elongisporus</i>) NRRL	틀림	아니요	없음
XP_720384 (29-4)	<i>칸디다 알비칸스</i> (<i>Candida albicans</i>) SC5314	맞음	아니요	클론
XP_456868	<i>데바리오마이세스 한센니</i> (<i>Debaryomyces hansenii</i>) CBS767	틀림	아니요	없음
XP_001487429 (29-6)	<i>피치아 굴리에르몽디</i> (<i>Pichia guilliermondii</i>) ATCC 6260	틀림	아니요	클론
XP_961039	<i>네우로스포라 크라사</i> (<i>Neurospora crassa</i>)	틀림	아니요	없음
CAG88709 (DH48)	<i>데바리오마이세스 한센니</i> (<i>Debaryomyces hansenii</i>) CBS767	맞음	아니요	클론
XP_001727326 (29-9)	<i>아스페르길루스 오리재</i> (<i>Aspergillus oryzae</i>)	맞음	아니요	클론
XP_001816757	<i>아스페르길루스 오리재</i> (<i>Aspergillus oryzae</i>)	맞음	아니요	클론

* “맞음(Correct)” = 클론(clone)의 서열이 데이터베이스에서의 서열과 일치한다 ;

“틀림(Wrong)” = 클론의 서열이 데이터베이스에서의 서열과 일치하지 않는다 ;

“없음(No)” = 이용가능하지 않은 결과[진행중인 작업(work in progress)].

[0426]

아라비노스-특이적인 운반체 XUT1 오솔로그의 클로닝

NCBI 참조서열	기원 (Origin)	서열(Sequence) 결과*	흡수 측정 (Uptake Assay)	상태 (Status)
XP_002545773	칸디다 트로피칼리스 (<i>Candida tropicalis</i>) MYA-3404	맞음(Correct)	예(Yes)	클론(Cloned)
EEQ43601	칸디다 알비칸스 (<i>Candida albicans</i>) WO-1	맞음	예	클론
XP_001818631	아스페르길루스 오리재 (<i>Aspergillus oryzae</i>) RIB40	없음(No)	아니요(No)	PCR 산물 없음 (No PCR product)
XP_002558275	페니실리움 크리스게움 위스콘신(<i>Penicillium chrysogenum Wisconsin</i>) 54-1255	틀림(Wrong)	예	클론
XP_001390883	아스페르길루스 니가 (<i>Aspergillus niger</i>)	없음	아니요	PCR 산물 없음
XP_750103	아스페르길루스 푸미게이트 (<i>Aspergillus fumigates</i>) Af293	틀림	아니요	없음(No)
XP_960000 (NC52)	네우로스포라 크라사 (<i>Neurospora crassa</i>) OR74A	틀림	아니요	클론
XP_657854(32- 8)	아스페르길루스 니둘란스 (<i>Aspergillus nidulans</i>) FGSCA4	맞음	아니요	클론
XP_001825068	아스페르길루스 오리재 (<i>Aspergillus oryzae</i>) RIB40	맞음	아니요	클론
XP_681669(32- 10)	아스페르길루스 니둘란스 (<i>Aspergillus nidulans</i>) FGSC	맞음	아니요	클론

* “맞음(Correct)” = 클론(clone)의 서열이 데이터베이스에서의 서열과 일치한다 ;

“틀림(Wrong)” = 클론의 서열이 데이터베이스에서의 서열과 일치하지 않는다

(예를 들어, 클론에서의 돌연변이 때문에) ;

“없음(No)” = 이용가능하지 않은 결과[진행중인 작업(work in progress)].

[0427]

신규한 자일로스-특이적인 운반체 및 하나의 아라비노스-특이적인 운반체의 목록

NCBI 참조서열	기원 (Origin)	자일로스- 특이적인	아라비노스- 특이적인
XP_457508 (DH61)	데바리오마이세스 한센니 (<i>Debaryomyces hansenii</i>) CBS767	예(Yes)	
XP_001727326 (29-9)	아스페르길루스 오리재 (<i>Aspergillus oryzae</i>)	예	
XP_720384 (29-4)	칸디다 알비칸스 (<i>Candida albicans</i>) SC5314	예	
XP_681669 (32- 10)	아스페르길루스 니둘란스 (<i>Aspergillus nidulans</i>) FGSC A4	예	
XP_657854 (32-8)	아스페르길루스 니둘란스 (<i>Aspergillus nidulans</i>) FGSC A4		예
XP_002488227	탈라로미세스 스티피타투스 (<i>Talaromyces stipitatus</i>)	예	
AB070824.1	아스페르길루스 오리재 (<i>Aspergillus oryzae</i>)	예	

[0428]

[0429]

데이터베이스에서 서열과 불일치하는 서열을 갖는 오솔로그(예를 들어, 돌연변이를 갖는 것)을 재-클론하였고, 배열순서를 밝히고(sequenced), 효모 스트레인에서 발현시키고, 당 흡수 기능에 대해서 테스트하였다. 유사하게, 시퀀싱(sequencing) 결과가 없는 것에 대한 오솔로그 또한 운반체 기능에 대해서 테스트할 것이다.

[0430] 오타당 운반체 오솔로그의 서열 정렬을 보존 잔기를 확인하기 위해 분석하였고, 이는 운반체 기능에서 잠재적인 역할을 할 수 있다. 자일로스 운반체(NCU0821, STL12/XUT6, XP_002488227.1 및 XP_002382573.1) 및 아라비노스 운반체(XUT1 및 EEQ43601.1)의 샘플의 정렬을 도 41(A, B)에 각각 나타내었다. 몇몇의 잔기는 자일로스 운반체에서 특이적으로 보존된 반면에, 다른 것은 아라비노스 운반체에서 특이적으로 보존되었다. 이러한 잔기는 특이적인 오타당을 수송하는데 중요한 역할을 가질 것이다. 자일로스 및 아라비노스 운반체의 서열의 전체적인 비교(도 41C)는, 둘 다의 유형의 오타당 운반체에서 보존된 잔기가 있음을 나타내고, 이는 일반적인 오타당의 흡수에서 기능적인 역할을 한다.

[0431] 실시예 12-15 는 효모에서 자일로스 활용 경로의 최적화와 관련된 것이다.

실시예 12

[0432] **오타당-이용 *S. 세레비시아* 스트레인의 유전자 조작(Engineering)**

[0433] 효율적인 자일로스 대사경로는 동종효소(isoenzymes)의 개념을 이용하여 재구성하였다(reconstituted). 동종효소는 상이한 운동 특성 또는 조절 특성을 갖는 동일한 화학적 반응을 촉진시키고(catalyze), 세포질 환경(cytosolic environment)에서 동적인 변화(dynamic changes)에 대응하여 물질대사의 흐름(metabolic fluxes)의 미세-조정 조절(fine-tuned control)을 부여하기 위한 것으로 알려져있다. 그러나, 사전의 물질대사의 유전자 조작된 접근(metabolic engineering approaches) 없이, 관심을 갖는 흐름(fluxes of interest)을 증가시키기 위해 동종효소를 사용하였다. 이러한 연구는 야생형 및 돌연변이 자일루로스 환원효소(XR) 동종효소 둘 다의 동시의 발현은 자일리톨 축적을 감소시킬 수 있고, 전체적인 자일로스 발효 비율을 증가시킬 수 있음을 나타낸다.

[0434] 살아있는 시스템(living systems)에서 동종효소의 출현율(prevalence)에 의한 영감에 의해, *S. 세레비시아*에서 기능적인 자일로스 대사 경로를 구성하기 위해, 자일리톨 탈수소효소(XDH) 및 자일루로키나아제(XR)와 함께 *S. 세레비시아*에서 야생형 XR 및 돌연변이체 XR(R276H)을 공동-발현시켰다(co-expressed). XR 돌연변이체는 NADH 상에서 NADPH 에 대해 보다 낮은 선호도(preference)를 나타낸 반면에, 야생형 XR 은 NADH 상에서 NADPH 에 대해 116 2-배(two-fold) 보다 높은 선호도를 나타내는 것이 보고되어 있다(Watanabe 등의, 2007).

[0435] *P. 스티피티스*로부터의 자일로스-대사작용하는 유전자(야생형 *XYL1*, 2, 및 3 및 돌연변이체 *XYL1*)를, 발현 카세트(expression cassettes)를 구성하기 위해 PCR-증폭시키고, 구성요소 프로모토(constitutive promoters)(PGK1 and TDH3)의 조절 하에 두었다. 이러한 통합 카세트(integration cassettes)를 D452-2 스트레인의 게놈 내로 통합시켰다(integrated).

[0436] 자일로스 대사 경로의 구성에 대한 발현 카세트의 형질전환(Transformation)은 효모 EZ-형질전환 키트(미국 캘리포니아주 비스타에 위치한 BIO 101)를 사용하여 실시하였다. 아미노산 영양요구 마커(auxotrophic marker)를 사용한 형질전환체(transformants)를 선별하기 위해, 효모 합성 완전한 배지[yeast synthetic complete (YSC) medium]를 사용하였고, 이는 6.7 g/리터(liter) 효모 질소 염기에 20 g/리터 글루코스, 20 g/리터 한천(agar) 및 CSM-Leu-Trp-Ura(BIO 101)을 더한 것을 함유하고, 이는 적절한 뉴클레오타이드 및 아미노산을 공급한다. 효모 스트레인은 20 g/리터 글루코스를 갖는 YP 배지 234[10 g/리터 효모 추출물, 10 g/리터의 박토 펩톤(Bacto peptone)]에서 30 °C 로 일상적으로 배양하였다.

[0437] 자일로스-대사작용 효율(xylose-metabolizing efficiency)에서의 *S. 세레비시아* 스트레인 백그라운드(background)의 효과는 몇몇의 상이한 효모 스트레인에서의 최적화된 자일로스 활용 경로 효소를 함유하는 동일한 구조물의 발현에 의해 또한 테스트되었다. 사용된 3 가지 실험실 스트레인은 D452-2(*MATa*, *leu2*, *his3*,

ura3, can1), L2612(*MATa, leu2-3, leu2-112, ura3-52, trp1-298, can1, cyn1, gal+*) 및 CEN.PK 이다. 자일리톨, 아세테이트 및 에탄올의 생산은 자일로스 및 OD₆₀₀ 의 사용과 함께 모니터링하였다. 결과는 D452-2 스트레인이 세 가지 실험한 스트레인 중에서 최고였다(도 42-44). *S. 세레비시아* D452-2 는 효모에서 자일로스-대사 작용하는 효소의 유전자 조작에 사용되었다. 본 연구에서 사용된 스트레인 및 플라스마를 표 17 에 나타내었다.

[표 17]

연구에서 사용된 스트레인(strain) 및 플라스미드

스트레인 또는 플라스미드	상세한 설명(Description)	참조문헌
스트레인		
D452-2	<i>MATa, leu2, his3, ura3, can1</i>	Hosaka 등의, (1992)
D801-130	β-글루코시다아제가 발현하는 D452-2 (NCU00130)(D452-2 expressing β-glucosidase) 및 <i>cbt1</i> (NCU00801)	본원의 연구 (In this study)
D809-130	β-글루코시다아제가 발현하는 D452-2 (NCU00130) 및 NCU00809	본원의 연구
D8114-130	β-글루코시다아제가 발현하는 D452-2 (NCU00130) 및 <i>cbt2</i> (NCU08114)	본원의 연구
DA24	<i>XYL1, mXYL1, XYL2, and XKS1</i> 가 발현하는 D452-2 [<i>leu2::TDH3P-XYL1-TDH3T, ura3::URA3-PGKP-mXYL1-PGKT-PGKP-XYL2-PGKT, Ty3::neo-TDHP-XKS1-TDHT</i> 을 제외한 D452-2 의 동질유전자(Isogenic)]	본원의 연구
DA24-16	자일로스 함유 배지에서 DA24 의 진화된 스트레인(Evolved strain)	본원의 연구
DA24-16BT3	다-복제(multi-copy) 플라스미드에서의 β-글루코시다아제(NCU00130)가 발현하는 DA24-16 및 단일-복제 통합(single-copy integration)을 통한 <i>cbt1</i> (NCU00801)	본원의 연구
DA24-16BT-M	다-복제 플라스미드에서 β-글루코시다아제(NCU00130) 및 <i>cbt1</i> (NCU00801)가 발현하는 DA24-16	본원의 연구
플라스미드		
pRS425	<i>LEU2</i> , 다복제(multi-copy) 플라스미드	Christianson 등의, (1992)
pRS426	<i>URA3</i> , 다복제 플라스미드	Christianson 등의, (1992)
pRS403	<i>HIS3</i> , 통합 플라스미드(integrative plasmid)	Sikorski 등의, (1989)
pRS405	<i>URA3</i> , 통합 플라스미드	Sikorski 등의, (1989)
pRS425-β-글루코시다아제	pRS425 에서 PGK 프로모터의 조절 하에서 β-글루코시다아제(NCU00130)	제출됨(Submitted)
pRS426- <i>cbt1</i>	pRS426 에서 PGK 프로모터의 조절 하에서 <i>cbt1</i>	제출됨
pRS426- <i>cbt2</i>	pRS426 에서 PGK 프로모터의 조절 하에서 <i>cbt2</i>	제출됨
pRS426-NCU00809	pRS426 에서 PGK 프로모터의 조절 하에서 NCU00809	제출됨
pRS403- <i>cbt1</i>	pRS403 에서 PGK 프로모터의 조절 하에서 <i>cbt1</i>	본원의 연구

유전자 조작된 자일로스-발효된 *S. 세레비시아* 스트레인(DA24)는 자일로스를 소비하고 무시해도 될 정도의 양 (negligible amounts)의 자일리톨 축적을 하면서 에탄올을 생산하였다. 40 및 80 g/L 의 자일로스를 단일 탄소 원으로서 사용하였을 때, DA24 스트레인은 둘 다 진탕-플라스크 및 바이오리액터 발효 실험에서 일관된 수득률 (consistent yields)(Y_{에탄올/자일로스} = 0.31- 0.32 g/g)을 갖는 에탄올을 생산하였다(도 45). 그러나, DA24 스트

레인은 자연 발생적으로 존재하는 자일로스-발효 효소, *P. 스티피티스* 보다 자일로스 소비가 보다 더 늦어졌다. DA24 의 자일로스 발효 능력(Xylose fermentation capability)은 진화하는 유전자 조작된 접근법(evolutionary engineering approach)을 사용하여 추가적으로 개선되었다(Sauer 2001). 자일로스-포함된 배지에서 DA24 의 반복된 계대-배양(sub-cultures) 후에 분리된 하나의 스트레인(DA24-16)을, 다양한 배양 조건 하에서 부모 스트레인(parental strain)과 비교하였을 때, 보다 더 빠른 자일로스 발효 비율을 나타내었다(표 18).

[표 18]

상당한 당 조건 하에서 두 개의 *S. 세레비시아* 스트레인 DA24 및 DA24-16 의 발효 파라미터의 비교

탄소원 (Carbon source)	스트레인	생산된 에탄올 (g/L)	당 소비 비율(g/L/h)	수득율 (g/g)	생산성 (g/L-h)
자일로스(80 g/L)	DA24	24	1.16	0.34	0.40
	DA24-16	28	1.32	0.35	0.47
글루코스(70 g/L) 및 자일로스(40 g/L)	DA24	34	1.45	0.39	0.74
	DA24-16	45	1.78	0.42	0.96

흥미롭게도, DA24-16 스트레인은 *P. 스티피티스*, 공지된 가장 빠른 자일로스-발효 효모 만큼 빨리 자일로스를 소비하였다. 그러나, DA24-16 에 의한 에탄올 수득률은 *P. 스티피티스* 보다 약간 더 낮았다(도 46).

스크린(screen)을, 게놈 라이브러리(genomic library)로 형질전환된(transformed) 자일로스-활용 효소(스트레인 YSX3)가 발현된 *S. 세레비시아* 스트레인 L2612 를 사용하여 준비하였다(set up). 형질전환한 다음에, 자일로스 사용에 효율적인 스트레인에 대해서 풍부하게 하기 위해서 산소-제한 조건 하에서 40 g/L 자일로스에서 시리얼 배양 이전(serial culture transfer) 하였다. $OD_{600} = 10$ 에 도달하였을 때, 발효를 산소-제한 조건 하에서 50 mL 의 YPX 배지에서 실행하였고, 0.1 %(50 μ L)의 완전히 성장한 세포 배양물을 그 다음 시리얼 배양물(the next serial culture)에 이동시켰다. 10 번의 시리얼 배양한 후에, 세포를 YPX(40 g/L) 한천 배지에서 단계 희석법(serial dilution)으로 확산시켰다(spread). 5 mL 의 YPX 배지를 사용한 발효 실험을 통해서, 낮은 자일리톨 및 높은 에탄올 생산을 하는 콜로니를 스크린하였다(screened). DNA 시퀀싱(DNA sequencing)은 두 가지의 가장 효율적인 스트레인은 *XYL2* 의 통합된 복사본(integrated copies)을 포함하고, 이는 상동 재조합(homologous recombination)을 통해서 다중-복제 플라스미드(multi-copy plasmid) 내로 클론되었고(cloned), YSX3 세포 내로 형질전환되었다(transformed).

XYL2 유전자는 상이한 강도의 프로모터 조절 하에서 통합 벡터, 예를 들어 TDHp 또는 PGKp 에 두었고, YSX3 세포 내로 형질전환시켰다(도 47). 형질전환된(transformed) 효모 세포에서 자일리톨 및 에탄올 생산에 대한 이러한 플라스미드의 효과를 모니터링하기 위해, 연구를 실시하였다. *XYL2*(PGKp 하에서)의 보다 높은 수준이 발현된 YSX3 세포는 에탄올 생산에 보다 효율적이었고, 게다가 보다 낮은 양의 자일리톨이 생산된 결과를 나타내었다(도 48). 추가적인 *XYL3* 은 이러한 세포(SRI 스트레인으로 명명)에서 발현되었을 때, 생산된 자일리톨의 양은 결과적으로 스트레인 SRu-23 에서 추가적으로 감소되었다(도 49). 따라서, 유전자 조작된 *S. 세레비시아* 스트레인에서 *XYL2* 발현 수준은 자일로스 발효를 시행하기 위한 중요한 요인(key factor)이고, 강한 프로모터 하에서 발현하였을 때, 스트레인은 높은 에탄올 수득할 뿐만 아니라 보다 낮은 자일리톨 축적을 갖는다. *XYL2* 및 *XYL3* 의 동시의 과잉-발현은 자일리톨 축적의 양을 추가적으로 감소시킬 수 있다. 그러나 *XYL1* 은 스트레인 과잉-발현된 *XYL2* 및 *XYL3* 에서 추가적으로 과잉-발현되었을 때, 상당한 자일리톨이 축적되었고, 따라서 자일로스 발효가 감소되었다(도 50-51). 따라서, 효율적인 자일로스 발효를 위한 최적 수준의 *XYL1* 임을 나타내었다.

XYL2 및 *XYL3* 이 발현된 *S. 세레비시아*에서 내인성 GRE3 의 과잉-발현이 자일로스 발효를 촉진할 수 있는지 여부를 테스트 하기 위해 실험을 또한 실시하였다. pRS403-GRE3 의 구성에 대해서, GRE3 유전자를 *S. 세레비시아* D452-2 로부터 증폭시켰고, TDH3 프로모터 및 CYC 종결자를 갖는 pR403 벡터 내로 삽입하였다. pRS403-GRE3

의 리니어라이제이션(linearization) 후에, D452-2 의 게놈 내로 이를 통합시켰다. 자일로스-활용 유전자를 효모 스트레인 D452-2 내로 도입하였고(도 52), 자일로스 발효 파라미터(parameters)를 모니터링하였다. 특히, 세포가, 높은 OD 접종(inoculations)에서 80g/L 의 자일로스에서 성장하였을 때, GRE3 의 과잉-발현은 에탄올 생산 및 자일리톨 축적에서 XYL1 의 과잉-발현만큼의 효과가 있는 결과를 나타내었다(도 53-54).

실시예 13

[0447] 유전자 조작된(engineering) LAD 및 XDH

[0448] NADPH-의존적인 XR 및 NAD⁺-의존적인 XDH 및 LAD 사이의 보조인자 불균형에 의해 기인하는, L-아라비니톨 및 자일리톨 축적은 오타당-이용(pentose-utilizing) 효소가 발현된 유전자 조작된 *S. 세레비시아*에서의 자일로스 발효 동안에 주요한 장애물(major bottleneck)로서 간주된다. XR 과 XDH 사이의 불균형은 역전된 보조인자 선호(reversed cofactor preferences)를 갖는 유전자 조작된 효소에 의해 정정되는 반면에(Watanabe 등의, 2007; Matsushika 등의, 2008; Bengtsson 등의, 2009), 수정된 효소(modified enzymes)가 감소된 특이적인 활성도를 갖는 바와 같이, 이러한 접근은 감소된 플럭스(reduced flux)를 야기한다. NADH 상에서 NADPH 에 대한 보다 낮은 선호를 나타낸 반면에, 야생형 psXR 은 NADPH 에 대한 2-배(two-fold) 보다 높은 선호를 나타냄이 *P. 스티피티스* XR 돌연변이체에서 보고되고 있다(Watanabe 등의, 2007).

[0449] 이러한 연구에서, 보조인자 특이성을 변경하기 위해 *N. 크라사*로부터 L-아라비니톨 4-탈수소효소(LAD) 및 XDH 에서 유사한 연구를 실시하였고, 이런 이유로 유전자 조작된 *S. 세레비시아*에서 자일로스 발효가 개선되었다.

[0450] 추정상의 LAD-코딩 유전자의 확인

[0451] 추정상의 LAD-코딩 유전자의 확인 방법 및 LAD-코딩 유전자 및 추정상의 LAD-코딩 유전자를 클로닝하는 방법을 기재하였다.

[0452] 추정상의 LAD-코딩 유전자의 확인

[0453] 프로브로서 ncLAD(EAA36547.1)을 사용한 단백질 BLAST 검색으로부터, 두 가지의 추정상의 유전자를 각각 *P. 크리소게눔*(*P. chrysogenum*)(XP_002569286.1) 및 *P. 킬리어몬디*(*P. guilliermondii*)(EDK37120.2)에서 확인하였다. ncLAD 을 갖는 이러한 두 가지 단백질의 아미노산 동일성은 각각 71 % 및 46 % 이었다.

[0454] 클로닝 LAD-코딩 및 추정상의 LAD-코딩 유전자

[0455] *A. 니가*(NRRL 326), *P. 킬리어몬디*(NRRL Y2075) 및 *P. 크리소게눔*(NRRL 807)를 United States Department of Agriculture Agricultural Research Service Culture Collection(미국 일리노이주 피오리아에 위치)로부터 입수하였다. *T. 롱기브라키아툼*(*T. longibrachiatum*)(*T. 레세이*, YSM 768)은 German Resource Centre for Biological Material 로부터 입수하였다(DSMZ).

[0456] *A. 니가*, *T. 롱기브라키아툼*, *P. 크리소게눔* 및 *P. 킬리어몬디*를 1 % 효모 추출물, 2 % 펩톤 및 2 % L-아라비노스를 포함하는 한천 플레이트(agar plates) 또는 액체 배지에서 성장시켰다. 세포를 전체 RNA 또는 게놈 DNA 의 분리를 위해 액체 질소에서 세포를 냉동시켰다. cDNA 를 획득하기 위해, 역 전사-PCR(RT-PCR)을 *T. 롱기브라키아툼*, *P. 크리소게눔* 및 *P. 킬리어몬디*로 부터 분리된 mRNA 에서 실행하였고, PCR 을 (추정상의) LADs 를 코딩하는 유전자를 획득하기 위해 사용하였다. *A. 니가*에 대해서, 추정상의 LAD 유전자를 알려지지 않은 이유 때문에 cDNA 로부터 증폭시키지 않았다. 따라서, 중복 확장-PCR(overlap extension-PCR, OE-PCR)은 분리된 게놈 DNA 로부터 이러한 인트론-포함된 유전자를 클론하는데 사용하였다. 참고로 이러한 유전자는 클론하는데 사용된 모든 프라이머 서열을 표 19 에 작성하였다.

[0457] [표 19]

야생형 LADs 의 클로닝에 사용된 프라이머(primers)

[제한효소 사이트(Restriction enzyme sites)는 볼드(bold) 및 이텔릭체(italicized)로 나타내었다.]

	제한효소	프라이머	서열
anLAD	<i>Nde</i> I	Fwd-단편(fragment) 1 ^a	5'-GACATCGATGAC CATATG ^c <i>GCTACCGCAAC</i> -3' SEQ ID NO: 71
		Rev-단편 1	5'-GTGCACGTCGGACCCGCAGATTCC-3' SEQ ID NO: 72
	<i>Bam</i> RI	Fwd-단편 2 ^b	5'-GGAATCTGCGGGTCCGACGTGCAC-3' SEQ ID NO: 73
		Rev-단편 2	5'-CAGAAGATTTAAG GGATC CTGAACGTAGA-3' SEQ ID NO: 74
tlLAD	<i>Nde</i> I	Fwd	5'-GACATCAGTGAC CATATG <i>TGCGCCTTCC</i> -3' SEQ ID NO: 75
	<i>Bam</i> RI	Rev	5'-CCTGGATTGAG GGATC CTGAACGTATA-3' SEQ ID NO: 76
pcLAD	<i>Nde</i> I	For	5'-GACATCGATGAC CATATG <i>GCTTCCGCAAC</i> -3' SEQ ID NO: 77
	<i>Eco</i> RI	Rev	5'-CCAGAAGTATTGAG AATTCT GAACGTAGA-3' SEQ ID NO: 78
pgLAD	<i>Nde</i> I	Fwd	5'-GACATCGATGAC CATATG <i>GCGACTCTGC</i> -3' SEQ ID NO: 79
	<i>Bam</i> HI	Rev	5'-GGATACAGAATGAG GGATC CTGAACGTAGA-3' SEQ ID NO: 80

· ^{a,b} 단편 1 및 2 는 인트론 옆에 배치된 업스트림(upstream) 및 다운스트림(downstream) 엑손을 나타낸다.

· 볼드(이텔릭체)인 ^c서열은 제한효소 사이트를 나타낸다.

[0458]

[0459]

PCR 산물을 pET-28a 벡터 내로 서브클론하였고(subcloned), 구조물(constructs)을, 각각 클로닝 및 발현을 위한 전기천공법(electroporation)에 의해서 두 가지의 *E. coli* 스트레인, DH5a 및 BL21(DE3) 내로 형질전환시키는데 사용하였다. *Nde*I/*Bam*HI 제한 부위(restriction sites)를 *A. 니가*, *T. 롱기브라키아툼* 및 *P. 켈리어몬 디*로부터의 예측된 유전자의 서브클로닝(subcloning)에 대해서 사용하였고, *Nde*I/*Eco*RI 사이트는 *P. 크리소게눔*에 대해서 사용하였다. 구조물(constructs)을 N-말단 His₆-태그된(tagged) 융합물(fusions)로서의 (추정상의) LADs 를 코드하였다. 플라스미드를, 일리노이 대학 어배너-샴페인에서의 Biotechnology Center 에서의 BIGDYE(TM) 종결자 시퀀싱 방법(Terminator sequencing method)을 사용하여 배열 순서를 밝히고(sequenced), 3730xL Genetic Analyzer(미국 캘리포니아주 포스터 시티에 위치한 Applied Biosystems)로 분석하였다.

[0460]

단백질 발현 및 정제

[0461]

pcLAD(XP_002569286.1), pgLAD(EDK37120.2), anLAD(CAH69383.1) 및 tlLAD(AAL08944.1)을 코딩하는 유전자를 pET-28a 벡터 내로 클론하였고, *E. coli* BL21(DE3)에 발현시켰다. LAD 유전자를 포함하는 *E. coli* BL21(DE3)를, 250 rpm 으로 회전식 진탕기(rotary shaker)에서 30 °C 에서 밤새 성장시켰다. 밤새 배양물(50 μL)를 신선한 배양물(5 mL)을 배양하는데 사용하였고, 이를 600 nm(OD₆₀₀)에 최적의 밀도가 0.6-1.0 에 도달할 때까지, 30 °C 에서 250 rpm 으로 교반시켜 성장시켰다. 그리고 난 다음에 배양물을 3-4 시간 동안 30 °C 로 또는 20 시간 동안 18 °C 로 0.3 mM IPTG 로 배양시켰다.

[0462]

유도된 세포(1 mL)를, 1 mg/mL 의 라이소자임(lysozyme)을 갖는 1 mL 의 50 mM 인산칼륨 완충용액(pH 7.0)에 이들을 재-현탁시키고, 30 °C 에서 30 분 동안 250 rpm 으로 교반시키면서 용해시켰다(lysed). 세포를 밤새 -80 °C 에서 유지시키고, 실온에서 용해시켰다(thawed). 결과적으로 생성된 세포 분해물(cell lysates)을 15 분

동안 13,200 rpm 으로 원심분리하고, 상청액 및 침전물을 도데실 황산 나트륨-폴리아크릴아미드 겔 전기영동(sodium dodecyl sulfate-polyacrylamide gel electrophoresis, SDS-PAGE)에 의한 단백질 발현에 대해서 분석하였다.

[0463]

단백질 정제에 대해서, 유도된 세포(induced cells)(400 mL)를 1 mg/mL 의 라이소자임(lysozyme)을 갖는 15 mL 의 완충용액 A (20 mM Tris, 0.5 M NaCl, 20% 글리세롤, pH 7.6)로 처리하고, 30 °C 에서 30 분 250 rpm 으로 교반시켰다. 냉동-융해 사이클(freeze-thaw cycle) 후에, 결과적으로 생성된 생산물을 추가적으로 음파처리(sonication)로 용해시킨 후에, 세포 잔유물(cell debris)을 제거하기 위해 20 분 동안 12,000 rpm 으로 원심분리하였다. 제조자의 지시에 따라서 His₆-태그된 단백질을 정제하기 위해 Co²⁺-고정화(immobilized) 금속 친화도 크로마토그래피 수지로 가득찬 컬럼에 상청액을 적용시켰다. 정제된 단백질을 초미세여과(ultrafiltration)(미국 매사추세츠주 빌러리카에 위치한 Millipore, Amicon Ultra)으로 탈염하였고(desalted), 150 mM NaCl 및 15 % 의 글리세롤을 포함하는 HEPES 완충용액(pH 7.0)으로 세척하였고, - 20 °C 에서 유지시켰다. 단백질 농도를 제조자의 프로토콜에 따라서 브래드퍼드법(Bradford method)(Bradford 1976)에 따라 결정되었다.

[0464]

LAD 단백질의 특성

[0465]

LAD 효소의 정상-상태 반응속도(steady-state kinetics), 분자량, 4 차 구조(quaternary structure), 온도 의존성, pH 의존성, L-아라비니톨 탈수소효소 활성도 및 금속 함량을 분석하였다.

[0466]

L-아라비니톨 탈수소효소 활성도

[0467]

P. 크리스게눔, *P. 퀴리어몬디*, *A. 니가* 및 *T. 룽기브라키아툼*으로부터의 LAD 발현된 숙주 세포로부터 용해물(Lysate)을 제조하였다. 10 마이크로리터의 세포 용해물(cell lysate)을 50 mM 의 인산칼륨 완충용액(pH 7.0)에서의 기질로서 200 mM L-아라비니톨 및 2 mM NAD⁺ 과 함께 활성도 검정에 사용하였다. NADH 생산을 Cary 300 Bio UV-vis 분광 광도계(미국 노스캐롤라이나주 캐리에 위치한 Varian)를 사용하여 340 nm($\epsilon = 6.22 \text{ mM}^{-1} \text{ cm}^{-1}$)에서 흡광도를 측정하여 모니터링하였다.

[0468]

정상-상태 반응속도(Steady-State Kinetics)

[0469]

상이한 LAD 효소의 동역학적 파라미터를 측정하였다. 초기의 비율을 50 mM 인산칼륨 완충용액(pH 7.0)에 실온에서 UV-vis 분광광도계(spectrophotometer)를 사용한 340 nm 에서 흡광도 변화(absorbance change)를 측정하여 결정하였다. 초기의 비율을, 다양한 농도의 기질(L-아라비니톨) 및 보조인자(NAD⁺/NADP⁺)(L-아라비니톨에 대해서 5 내지 320 mM, 보조인자에 대해서 0.5 내지 3.2 mM)에서 측정하였다. 기질 및 보조인자의 효소 반응속도를 미하엘리스-멘텐 반응속도론(Michaelis-Menten kinetics)을 사용하여 분석하였고, 동역학적 파라미터(kinetic parameters)를 라인웨버-버크도면(Lineweaver-Burk plot)에 데이터를 맞추어 결정하였다. 기질에 대한 파라미터를 포화 보조인자 농도(3.2 mM)에서 초기의 비율을 측정하여 결정하였고, 보조인자에 대한 이러한 비율을 포화 기질 농도(320 mM)에서 결정하였다. 검정을 3 번 실시하였다.

[0470]

클론된 LADs 를 L-아라비니톨에 대해서 상이한 결합 친화도(binding affinities) 및 촉매 활성도(catalytic activities)를 나타내었다 : K_m 은 2 배 달랐고, K_{cat} 는 LADs 중에서 약 3 배나 달랐다. L-아라비니톨에 대해서, anLAD, tILAD, 및 pcLAD 의 K_m 값은 각각 25 ± 1 , 18 ± 1 및 $37 \pm 2 \text{ mM}$ 이고, K_{cat} 값은 각각 507 ± 22 , 346 ± 41 및 $1085 \pm 71 \text{ min}^{-1}$ 이다(표 20). tILAD 효소가 가장 낮은 K_m 을 가진 반면에, pcLAD 는 가장 높은 K_m 값을 가짐에도 불구하고, 가장 높은 촉매 활성도(k_{cat}) 및 효율성(k_{cat}/K_m)을 나타내었다(표 20). 보조인

자 NAD^+ 속도에 대해서, 클론된 LADs 는 0.2-0.3 mM 의 범위에서 K_m 값 및 2526 내지 3460 $\text{mM}^{-1}\text{min}^{-1}$ 의 범위에
서 촉매 효율성(catalytic efficiencies)을 나타내었다(표 21). 모든 클론된 LADs 는 NADP^+ 에 대해서 최소한
의 활성도를 나타내었다(표 20, 21). 초기의 비율은, 광범위한(large) K_m 때문에, 가장 높은 기질 및 보조인자
농도(L-아라비니톨에 대해서 320 mM 및 NADP^+ 에 대해서 3.2 mM)에서 포화되지 않았다. 따라서, 효소의 촉매
효율성만이 NADP^+ 에 대해서 0.1 또는 0.2 mM 및 L-아라비니톨에 대해서 10 또는 20 mM ($K_m \gg [S]$)으로 결정되
었다(표 20, 21).

[0471]

[표 20]

**포화된 보조인자(cofactor) 농도에서 L-아라비니톨에 대한 LADs 의 동역학 파라미터
(Kinetic parameters)**

		고유활성도 (Specific activity) (U/mg 단백질)	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m ($\text{mM}^{-1}\cdot\text{min}^{-1}$)
anLAD	NAD^+	11.7 ± 0.3^a	25 ± 1	507 ± 22	20.0 ± 0.8
	NADP^+	- ^b	-	-	0.04 ± 0.01
tlLAD	NAD^+	8.7 ± 0.1	18 ± 1	346 ± 41	19.0 ± 0.8
	NADP^+	-	-	-	0.13 ± 0.02
pcLAD	NAD^+	25.3 ± 1.4	37 ± 2	1085 ± 71	29 ± 1
	NADP^+	-	-	-	0.04 ± 0.02

^a 오차(error)는 평균으로부터의 표준편차를 나타내었고, n = 3

^b 대쉬(dash)는 나타낸 보조인자에 대한 높은 K_m 때문에 발견되지 않음을 나타낸다.

[0472]

[0473]

[표 21]

**포화된 L-아라비니톨 농도에서 NAD^+ 및 NADP^+ 에 대한 LADs 의 동역학 파라미터
(Kinetic parameters)**

		K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m ($\text{mM}^{-1}\cdot\text{min}^{-1}$)
anLAD	NAD^+	0.20 ± 0.01^a	494 ± 11	2526 ± 83
	NADP^+	- ^b	-	20 ± 9
tlLAD	NAD^+	0.2 ± 0.1	436 ± 96	2689 ± 646
	NADP^+	-	-	17 ± 9
pcLAD	NAD^+	0.3 ± 0.1	1039 ± 165	3460 ± 505
	NADP^+	-	-	15 ± 4

^a 오차(error)는 평균으로부터의 표준편차를 나타내었고, n = 3

^b 대쉬(dash)는 나타낸 보조인자에 대한 높은 K_m 때문에 발견되지 않음을 나타낸다.

[0474]

[0475]

분자량 및 4 차 구조

[0476]

4 가지의 단백질의 서브유닛(subunits)의 계산된 분자량은 43 kDa (anLAD), 41 kDa (tlLAD), 42 kDa (pcLAD),
및 42 kDa (pgLAD)이다. 단백질의 분자량을 Shimadzu HPLC 시스템(일본 교토에 위치한 Shimadzu)에서 Bio-Sil
SEC-250 컬럼(300x7.8 mm, 미국 캘리포니아 허클러스에 위치한 Bio-Rad)을 사용하여 결정하였다. 이동상(mobile
phase)은 50 mM Na_2HPO_4 , 50 mM NaH_2PO_4 , 150 mM NaCl 및 10 mM NaN_3 (pH 6.8)으로 구성되어 있고, 유량(flow
rate)은 1.0 mL/min 이다. 분자량을 표준 단백질 분자량의 보존시간(retention times)과 비교하여 측정하였다.

[0477] 4 차 구조를 HPLC 및 단량체 서브유닛의 분자량에 의해 관찰된 분자량을 기반으로 측정하였고, 이를 SDS-PAGE 분석으로 결정하였다. an-, tl-, 및 pcLAD 의 분자량은 각각 178, 194 및 173 kDa 으로 측정되었다. SDS-PAGE 로 결정된 서브유닛(subunits)의 분자량을 비교하여, LADs 는 이들의 천연형태(native forms)에서 비-공유 결합으로 연결된 사합체(tetramers)이다.

[0478] 온도 및 pH 의존성

[0479] 단백질의 최적의 온도를 10 내지 70 °C 범위의 온도에서 효소 활성도를 검정하여 결정하였다. 열 불활성을, 인산 완충용액에서 50 °C 에서 다양한 배양 시간 후에 효소 활성도(enzyme activity)를 측정하여 결정하였다. 효소 활성도를 2 mM NAD^+ 및 200 mM L-아라비니톨로 측정하였다. 효소 활성도의 반감기(Half-life)를 첫 번째-오더 지수 붕괴 함수(first-order exponential decay function)를 사용하여 측정하였다. 온도를 UV-vis 분광 광도계에 연결된 Cary 온도 컨트롤러(Cary temperature controller)(미국 노스캐롤라이나주 캐리에 위치한 Varian)에 의해 조절하였다. pH-의존적 효소 활성도는, 일반적인 완충용액(50 mM 모르폴린에탄술폰산/50 mM 트리스/50 mM 글리신)에서 NAD^+ (2 mM) 및 L-아라비니톨(200 mM)의 포화 농도에서 5.0 과 11.0 사이의 pH 에서의 활성도를 측정하여 결정하였다(Ellis and Morrison 1982).

[0480] anLAD 및 pcLAD 의 최적 온도는 40 에서 50 °C 사이인 반면에, tlLAD 은 55 에서 65 °C 사이의 보다 높은 최적 온도를 나타내었다(도 55A). LADs 의 촉매 활성도는 50 °C 에서 배양 시간의 길이(length)에 따라 기하급수적으로 감소되었고, 100 분 후에 거의 대부분 완전하게 불활성화되었다(deactivated)(도 55B). tlLAD 는 50 °C 에서 20 분의 반감기를 가지면서 열에 의해 가장 안정되었고, anLAD 는 50 °C 에서 5 분 미만의 반감기를 가지면서 가장 낮게 안정되었다. 모든 특성화된 LADs 는 약 pH 9.4 에서 최대의 활성도를 갖는 pH 범위 7 내지 11 에서 활성(activity)을 나타내었고(도 55C). pH 범위 9 내지 10 외에서, 활성도는 현저하게 감소되었고, 대략 20 % 의 활성도가 pH 7.0 에서 남아있었다(도 55C). pH 5.0 이하에서는 어떠한 활성도도 발견되지 않았다.

[0481] 금속 분석(Metal analysis)

[0482] 금속 분석에 대한 사본 샘플(Duplicate samples)을 완충용액 교환 및 동결건조(lyophilization)에 의한 인산완충식염수(phosphate buffered saline, PBS)에서 제조하였다. 각각의 샘플은 1 mL 의 완충용액에 1-2 mg 의 단백질 샘플을 포함한다. 금속의 함량 및 정체(identity)를 Microanalytical Laboratory at the University of Illinois at Urbana-Champaign(미국 일리노이드주 어배나에 위치)에서의 고주파유도결합형 플라즈마 발광 분석법(inductively coupled plasma atomic emission spectrometry)(OES Optima 2000 DV, 미국 매사추세츠주 보스턴에 위치한 Perkin Elmer)으로 분석하였다.

[0483] Zn^{2+} 의 분자량 퍼센트는 1 : 1 분자비(molar ratio)를 기반으로 계산한 것에 가까웠다(표 22).

[0484] [표 22]

계산된(calculated) 및 측정된(measured) Zn^{2+} 함유량

	계산된 중량 ^a (%) (Calculated Weight)	측정된 중량(%) (Measured weight)
anLAD	0.027	0.027 ± 0.003^b
tILAD	0.047	0.048 ± 0.003
pcLAD	0.048	0.061 ± 0.013

^a계산된 분자량(Calculated molecular weights)은 완충용액 조성물, 단백질 농도 및 LAD 단량체 서브 유닛(subunit) 및 Zn^{2+} 의 1 : 1 분자비(molar ratio)에 근거하여 측정하였다. 완충용액(1 L)는 NaCl(8 g), KCl(0.2 g), Na_2HPO_4 (1.44 g), and KH_2PO_4 (0.24 g)을 함유한다.

^b모든 샘플은 중복하여 분석하였으며, 오차는 표준편차이다.

[0485]

변경된 보조인자 특이성(Altered Cofactor Specificity)을 갖는 LAD 효소의 유전자 조작(engineering)

[0486]

LADs 의 변경된 보조인자 특이성의 방법을 결정하였고, 돌연변이된 LADs 는 변경된 보조인자 특이성 및 그 밖의 특성에 대해서 분석하였다.

[0487]

변경된 보조인자 특이성을 갖는 LADs 의 개발(Development)

[0488]

NAD^+ 에서 $NADP^+$ 까지의 anLAD, tILAD 및 pcLAD 의 보조인자 특이성을 변경하기 위해 부위 특이적 돌연변이 (Site-directed mutagenesis)를 실시하였다. 변경된 보조인자 특이성을 갖는 tILAD 를 생성하기 위해, 자연적으로 발생한 tILAD 의 아미노산 번호 224, 225 및 362 을 각각 세린, 아르기닌 및 트레오닌으로 치환하였다. 클론된 anLAD 및 pcLAD 의 아미노산 서열은 *T. 룡기브라키아* tILAD (tILAD) 서열과 정렬하였고(aligned), tILAD 아미노산 번호 224, 225 및 362 에 부합하는 아미노산이 돌연변이 되었다(mutated). 변경된 보조인자 특이성을 갖는 모든 LADs 에 대해서, 조효소 결합 도메인의 β - α - β 모티프(motif) 내의 두 개의 아미노산 잔기가 세린 및 아르기닌으로 각각 교체되었고 ; anLAD 에 대해서 D213 및 1214, tILAD 에 대해서 D224 및 1225, 및 pcLAD 에 대해서 D212 및 1213(Korkhin 등의, 1998; Pauly 등의, 2003; Watanabe 등의, 2005), 세 번째 돌연변이를 anLAD 에 대해서 A359, tILAD 에 대해서 A362 및 pcLAD 에 대해서 S358 에 도입되었고, 트레오닌으로 교체되었다(프라이머 서열에 대해서, 표 23 참조). 메가프라이머(Megaprimer) PCR 방법을 주형(templates)으로서 야생형 LAD 구조물(constructs)을 사용한 부위-특이적인 돌연변이(site-specific mutations)를 도입하여 사용하였다(Sarkar and Sommer 1990). 정확한 돌연변이를 DNA 서열 분석(DNA sequence analysis)으로 확인하였다.

[0489]

[0490]

[표 23]

메가프라이머(megaprimer) PCR 방법^a에 의한 부위 지정 돌연변이(site directed mutagenesis)에 사용된 프라이머

	Fwd-T7-pro Rev-T7-term	5'-TAATACGACTCACTATAGGG-3' <i>SEQ ID NO: 81</i> 5'-GCTAGTTATTGCTCAGCGG-3' <i>SEQ ID NO: 82</i>
anLAD	Fwd-D213S/I214R	5'-CCTATCGTCATTACCTCACGT ^b GACGAGGGCGGCTG-3' <i>SEQ ID NO: 83</i>
	Rev-D213S/I214R	5'-CAGCCGCCCTCGTCACGTGAGGTAATGACGATAGG-3' <i>SEQ ID NO: 84</i>
	Fwd-A359T	5'- CCT TCGAAACGGCTACAAACCCCAAGACG -3 <i>SEQ ID NO: 85</i>
tILAD	Fwd-D214S/I215R	5' -GCTTGTCAATC ACATCACGTTTCAGAGAGCCGCTG-3' <i>SEQ ID NO: 86</i>
	Rev-D214S/I215R	5' -CAGACGGCTCTCTGAACGTGATGTGATGACAAGC-3' <i>SEQ ID NO: 87</i>
	Fwd-S362T	5' -GCATTTGAGACGTCAACAGATCCCAAGAGC-3' <i>SEQ ID NO: 88</i>
pcLAD	Fwd-D212S/I213R	5' -CCTATTGTCATCACTTCACGTGACGAGGGCCGCTTG-3' <i>SEQ ID NO: 89</i>
	Rev-D212S/I213R	5' -CAAGCGGCCCTCGTCACGTGAAGTGATGACAAT AGG-3' <i>SEQ ID NO: 90</i>
	Fwd-S358T	5' -CCTTTGAGACTGCC ACAAACCTAAGACCGGTG-3' <i>SEQ ID NO: 91</i>

^a돌연변이체 LADs를 형성시키기 위해, 각각 Fwd-T7-pro와 Rev-D213S/I214R 및 Fwd-A359T와 Rev-T7-term 프라이머를 사용하여 단편 1 및 2 를 증폭시켰다. Fwd-D123S/I214R 및 단편 2[Rev 메가프라이머(megaprimer)]를 사용하여 단편 3 을 증폭시켰다. 단편 1 및 3 의 오버랩 익스텐션(overlap extension)에 의해 완전한 돌연변이체 유전자(Full mutant genes)를 증폭시켰다. 주형(template) DNA 는 pET-28a 플라스미드이다.

^b밑줄친 서열은 돌연변이 부위(mutation sites)이다.

[0491]

[0492]

변경된 보조인자 특이성을 갖는 LADs 의 동력학적 분석(Kinetic Analysis)

[0493]

본원의 실시예에서, "tILAD 돌연변이체(mutant)"는 돌연변이 D224S / I225R / A362T 를 갖는 tILAD 로서 규정하였고 ; "anLAD 돌연변이체"는 돌연변이 D213S / I214R / A359T 를 갖는 anLAD 로서 규정하였고 ; "pcLAD 돌연변이체"는 돌연변이 D212S / I213R / S358T 를 갖는 pcLAD 로서 규정하였다. tILAD 돌연변이체는 NAD^+ 에서 NADP^+ 까지의 현저하게 변경된 보조인자 특이성을 나타내었다. 이는 가장 높은 촉매 활성도를 또한 나타내었다. NADP^+ 를 갖는 L-아라비니톨에 대한 tILAD 돌연변이체의 K_m 및 K_{cat} 는 각각 $46 \pm 4 \text{ mM}$ 및 $170 \pm 9 \text{ min}^{-1}$ 이다 (표 24). 포화된 NAD^+ 를 갖는 tILAD 돌연변이체를 포함하는 모든 검정에서, 반응 속도의 정체기(plateau)는 실험된 농도 범위에서는 관찰되지 않았고, 따라서 촉매 효율성(catalytic efficiencies)은 NAD^+ 에 대해서 0.8 mM 및 L-아라비니톨에 대해서 80 mM 으로 측정되었다(표 24, 25). 보조인자에 대해서, anLAD 및 tILAD 돌연변이체는 NAD^+ 상에서 NADP^+ 에 대해 현저하게 보다 높은 선호도를 나타내었다(표 25). anLAD 및 tILAD 돌연변이체의 K_m 값은 각각 0.46 ± 0.09 및 $0.10 \pm 0.01 \text{ mM}$ 이고, K_{cat} 값은 각각 55.7 ± 6.4 및 $90.5 \pm 9.2 \text{ min}^{-1}$ 이다(표 25). anLAD 및 tILAD 돌연변이체의 촉매 효율성(catalytic efficiencies)은 130 ± 32 및 $934 \pm 72 \text{ mM}^{-1}\text{min}^{-1}$ 이고, NAD^+ 에 대한 NADP^+ 을 갖는 촉매 효율성의 비율은 각각 100 및 161 이다. tILAD 돌연변이체에 대해서, NAD^+ 에 대한 NADP^+ 을 갖는 촉매 효율성의 비율은 2.5×10^4 배 증가되었다(표 21, 25). pcLAD 돌연변이체는 NAD^+ 와 함께 어떠한 활성도를 나타내지 않았다.

[0494]

[표 24]

포화된 보조인자(cofactor) 농도에서 L-아라비니톨에 대한 LAD 돌연변이체의 동역학 파라미터(Kinetic parameters)

		고유활성도 (Specific activity) (U/mg 단백질)	K_m (mM)	k_{cat} (min ⁻¹)	k_{cat}/K_m (mM ⁻¹ ·min ⁻¹)
anLAD 돌연변이체	NAD ⁺	- ^a	-	-	0.010 ± 0.002 ^b
	NADP ⁺	-	-	-	0.45 ± 0.20
tlLAD 돌연변이체	NAD ⁺	-	-	-	0.050 ± 0.007
	NADP ⁺	3.9 ± 0.2	46 ± 4	170 ± 9	3.7 ± 0.2
pcLAD 돌연변이체	NAD ⁺	-	-	-	-
	NADP ⁺	-	-	-	0.02 ± 0.02

^a 대쉬(dash)는 나타낸 보조인자에 대한 높은 K_m 때문에 발견되지 않음을 나타낸다.

^b 오차(error)는 평균으로부터의 표준편차를 나타내었고, n = 3.

[0495]

[0496]

[표 25]

포화된 L-아라비니톨 농도에서 NAD⁺ 및 NADP⁺에 대한 LAD 돌연변이체의 동역학 파라미터(Kinetic parameters)

		K_m (mM)	k_{cat} (min ⁻¹)	k_{cat}/K_m (mM ⁻¹ ·min ⁻¹)
anLAD 돌연변이체	NAD ⁺	- ^a	-	1.3 ± 0.3 ^b
	NADP ⁺	0.46 ± 0.09	55.7 ± 6.4	130 ± 32
tlLAD 돌연변이체	NAD ⁺	-	-	5.8 ± 0.8
	NADP ⁺	0.097 ± 0.011	90.5 ± 9.2	934 ± 72
pcLAD 돌연변이체	NAD ⁺	-	-	-
	NADP ⁺	-	-	3.6 ± 1.0

^a 대쉬(dash)는 나타낸 보조인자에 대한 높은 K_m 때문에 발견되지 않음을 나타낸다.

^b 오차(error)는 평균으로부터의 표준편차를 나타내었고, n = 3.

[0497]

변경된 보조인자 특이성을 갖는 N. 크라사 XDH(ncXDH)의 유전자 조작(Engineering)

[0498]

추정상의(putative) ncXDH 의 클로닝(Cloning) 및 특성

[0499]

[0500]

추정상의 N. 크라사 자일리톨 탈수소효소(ncXDH) 서열을, 의문 서열(query sequence)로서 P. 스티피티스 자일리톨 탈수소효소(psXDH)를 사용하여 National Center for Biotechnology Information 웹사이트(웹사이트 ncbi.nlm.nih.gov)에서 단백질 BLAST 검색을 사용하여 발견하였다. 두 가지의 효소를 ClustalW 알고리즘(algorithm)을 사용하여 완전히 정렬하였고, 44 %의 동일성(fully) 및 60 %의 유사성(similarity)을 공유하는 것으로 나타났다(도 56). 네우로스포라 크라사의 전체-게놈 서열은 발표되었고(Galagan 등의., 2003), 이는 추정상의 자일리톨 탈수소효소(XDH) 유전자의 클로닝에 대한 프라이머를 디자인하는데(design) 사용되었다.

[0501]

D-자일로스-유도된 N. 크라사 10333 으로부터 분리된 전체 RNA 에서 실행한 RT-PCR 은 예측된 크기의 유전자 산물(-1.1 kb)을 나타내었다. RT-PCR 산물은 *NdeI* 및 *SacI* 제한효소 부위(restriction sites)를 사용한 pET-28a 벡터 내로 클론되었고, *E. coli* BL21 (DE3) 내로 형질전환되었다(transformed). 이러한 구조물(construct)(pET-28a ncXDH)은 트롬빈 분해 사이트(thrombin cleavage site)를 갖는 N-말단 His6-태그된 융합물(fusion)로서의 ncXDH 가 발현되었다. 이러한 세포의 IPTG-유도된 배양물의 세포 용해물(Cell lysates)을 제조하였고, SDS-PAGE 로 분석하였고, XDH 활성도에 대해서 검정하였다. 제조자의 프로토콜에 따라서, Talon® Co2+ 슈퍼플로우 수지(Superflow resin)(Clontech, Mountain View, CA)를 사용한 고정화 금속 이온 친화도 크

로마토그래피(immobilized metal ion affinity chromatography, IMAC)로 XDH 를 정제하였다. 정제된 단백질을 50 mM 4-(2-히드록시에틸)피페라진-1-에탄술폰산(HEPES) 완충용액(pH 7.25) + 15 % 글리세롤을 여러 번 세척하면서 초미세여과(ultrafiltration)하여 탈염시키고(desalted), - 80 °C 에서 냉동시켜 저장하였다. 단백질을 브래드퍼드법(Bradford method)으로 결정하였다(Bradford 1976).

[0502] ncXDH 는 절대적으로 NAD^+ -선호하는 효소이다. ncXDH 는 높은 안전성(50 °C 에서 ~200 분의 반감기) 및 발현을 또한 나타내었다. Watanabe 등의(2005b)에 의한 이전의 작업은 psXDH 의 보조인자 특이성을 역전시키는 것을 목표로한다.

[0503] **변경된 보조인자 특이성을 갖는 ncXDH 의 개발(development)**

[0504] 서열 정렬을 통해서, ncXDH 의 잔기 D204,1205 및 V206 은, ncXDH-ARS 를 생성하기 위해 각각 알라닌, 아르기닌 및 세린에 대한 부위 특이적인 돌연변이를 목표로 한다. 표 26 은, 현재 선호하는 NADP^+ , ncXDH-ARS 가 완전하게 역전된 보조인자 특이성을 갖음을 나타낸다. 기질 자일리톨에 대한 친화도(affinity)는 보조-인자에 대한 친화도-변화로부터 매우 많이 악화되지 않았다.

[0505] [표 26]

포화된 자일리톨 농도에서 니코틴아미드 보조인자 NAD^+ 및 NADP^+ 를 갖는 *N. 크라사* 및 *P. 스티피티스* XDH 및 XDH-ARS 에 대한 동역학 파라미터(Kinetic parameters)

NAD^+				NADP^+			
효소	k_{cat} (min^{-1})	K_m (mM)	k_{cat}/K_m ($\text{mM}^{-1}\text{min}^{-1}$)	k_{Cat} (min^{-1})	K_m (mM)	$k_{cat}K_m$ ($\text{mM}^{-1}\text{min}^{-1}$)	출처 (Source)
ncXDH- wt	2160	0.127	17000	-a	~5.6	~68	본원의 연구
ncXDH- ARS	-a	~3.5	~165	2080	0.325	6400	본원의 연구
psXDH	1050	0.381	2760	110	170	0.65	Watanabe 등의. (2005b)
psXDH- ARS	240	1.3	181	2500	0.897	2790	Watanabe 등의. (2005b)

a 발견되지 않음. 보조인자 포화가 도달하지 않음.

모든 검정을 50 mM Tris, pH 8.0 에서 25 °C 에서 실시하였다.

[0506]

[0507] **ncXDH 돌연변이체의 동역학적 분석(kinetic analysis)**

[0508] 돌연변이체 ncXDH 는 보조인자 특이성의 극적인 전환(dramatic reversal)을 갖는다. NADP^+ 에 대한 돌연변이체 ncXDH 의 K_m 은, NAD^+ 에 대한 야생형 ncXDH 의 K_m 보다 약 2.5-배 보다 높은 반면에, k_{cat} 값은 유사하였다(표 27).

[0509] [표 27]

기질 자일리톨에 대한 ncXDH 돌연변이체의 동역학 파라미터

효소	k_{cat} (min^{-1})	K_m (mM)	k_{cat}/K_m ($\text{mM}^{-1}\cdot\text{min}^{-1}$)
ncXDH-wt	2170 ± 135	6.6 ± 2.0	330
ncXDH-ARS	2090 ± 35	4.3 ± 0.3	490

a 발견되지 않음. 보조인자 포화가 도달하지 않음.

모든 검정을 50 mM Tris, pH 8.0 에서 25 °C 에서 실시하였다.

모든 효소는 정제하였고, N-His₆-tag로 특징지었다.

[0510]

[0511] 도 57 에 나타낸 바와 같이, XDH 활성도는 pH 4.0 에 이르기까지 확장된 활성도를 갖는 보다 산성인 조건에 대한 보다 높은 내성(tolerance)을 나타낸 반면에, LAD 활성도는 생체외에서 활성도 검정에서 pH 5.0 에서 정지하였다(abolished).

실시예 14

[0512] *S. 세레비시아*에서 박테로이드 스테로코리스(*Bacteroides stercoris*)로부터의 자일로스 이성질화효소(xylose isomerase)의 발현

[0513] 박테리아 자일로스 이성질화효소(XI)는 자일로스에서 자일루로스의 전환을 포함한다. 최근에는, 혐기성 곰팡이의 두 가지 종[피로마이세스종(*Piromyces* sp.) 및 오르피노마이세스종(*Orpinomyces* sp.)] 및 혐기성 박테리아[클로스트리움 파이토퍼먼트탄스(*Clostridium phytofermentans*)]로부터의 활성형 XI 가 발현된 세 가지의 성공적인 경우가 보고되고 있다. 피로마이세스 종 E2 으로부터의 균류에 의한 XYLA 유전자는 *S. 세레비시아*에서 기능적으로 발현되었고, XI 활성도의 최대한의 1.1 U/mg-단백질을 30 °C 에서 수득하였다(Kuyper 등의, 2003). 피로마이세스 종으로부터의 유전자와 94 % 동일성을 갖는, 오르피노마이세스로부터의 두 번째 균류에 의한 XYLA 유전자는 *S. 세레비시아*에서 또한 기능적으로 발현되었다(Madhavan 등의, 2009). 최근에는, 클로스트리움 파이토퍼먼트탄스로부터의 첫 번째 원핵세포의 xylA 유전자가 *S. 세레비시아*에서 기능적으로 발현되었다(Brat 등의., 2009).

[0514] 혐기성 박테리아 박테로이드 스테로코리스(BtXI)로부터의 이성질화효소(isomerase) xylA 는 피로마이세스 종으로부터의 이성질화효소와 높은 서열 동일성을 공유한다(82 %). BtXI 를 pRS424TEF 벡터 내로 클론시켰고, *S. 세레비시아* L2612 스트레인 내로 형질전환시켰다(transformed). 유전자는 pRS403TEF 벡터를 사용하여 *S. 세레비시아* D452-2 내로 또한 통합되었다. 에탄올 생산을 BtXI (L2612 에서 5g/L 및 D452-2 에서 7.8 g/L)가 발현된 둘 다의 스트레인에서 관찰되었다(도 58-59). 그러나, 생산의 비율은 XYL 유전자가 발현된 유전자조작된 스트레인의 비율과 비교하여 상대적으로 낮았다.

[0515] 낮은 에탄올 생산은 어떠한 축적된 자일리톨(내인성 효소 알도스 환원효소에 의해 자일로스로부터 형성)의 억제적 효과에 기인할 수 있다. 자일리톨 축적을 감소시키기 위해서, XDH 및 XK 는 BTXI-발현된 효모 스트레인에서 발현되었다(DBtXI). 결과적으로 생성된 스트레인은 약간의 개선된 에탄올 수득을 갖고, 감소된 자일리톨 생산을 갖는다(도 60). DBtXI 에서 이러한 두 가지 XYL 유전자의 공동-발현(Co-expression)은 혐기성 조건 하에서도 에탄올 생산을 야기한다.

실시예 15

[0516] **오탄당 인산 경로에서의 효소의 과잉-발현(Over-Expression)**

[0517] PPP 효소 *P. 스티피티스*로부터의 글루코스-6-포스페이트 탈수소효소(*ZWF1*), 6-포스포글루코네이트 탈수소효소(*GND1*), 트랜스알돌라아제(*TAL1*) 및 트랜스케톨라아제(*TKT1*)는 강한 프로모터(strong promoter)(P_{GPD})의 조절 하에서 통합 벡터(pRS406) 내로 클론되었다(cloned). 플라스미드를 효소 *StuI*에 의해 선모양으로 만들고(linearized), *S. 세레비시아*의 염색체(chromosome) 내로 통합되었다.

[0518] 그러나, 과잉-발현된 PPP 효소의 유익한 효과를 얻기 위해서, *XYL3(XK)*의 과잉-발현될 수 있는 것을 또한 갖는 다(도 61). *XYL3* 및 PPP 효소의 발현은 YP-자일루로스 배지에서 에탄올 생산을 또한 증진시킨다.

실시예 16

[0519] **알도스-1-에피머화효소(Epimerase)의 발현**

[0520] β -글루코시다아제에 의한 셀로비오스의 가수분해(Hydrolysis)는 β -D-글루코스를 방출하였다. 그러나 효모 헥소키나아제(hexokinases)는 α -D-글루코스를 선호하고(또는 배타적인 사용), β -D-글루코스에서 α -D-글루코스로의 무타로에이션(mutarotation)의 비율은 대사율(metabolic rate)을 효율적으로 낮출 수 있다. 이러한 전환(conversion)을 증진시키는 하나의 방법은 예측된 알도스-1-에피머화 효소 NCU09705를 과잉-발현시킬 수 있다. 이러한 가설(hypothesis)은 과잉-발현된 NCU09705 상동성(homologs)에 의해 테스트되었다: *S. 세레비시아*에서의 *GAL10*, *YHR210C* 및 *YNR071C*; 및 *P. 스티피티스*에서의 *GAL 10*. 스트레인을 셀로비오스 소비 및 에탄올 생산에 대해서 테스트하였다(도 62). 결과는, *S. 세레비시아*에서 상동성의 과잉-발현(over-expression)은 셀로비오스 소비 및 에탄올 생산에서 약간의 증가됨을 결과는 나타내었다.

실시예 17

[0521] **자일로스 및 셀로비오스의 공동-발효(Co-Fermentation)**

[0522] 이러한 실시예에서, 글루코스의 이합체인, 셀로비오스가 자일로스(오탄당)와 함께 공동-발효된 곳에서 글루코스 억제제를 극복하기 위한 새로운 전략이 사용되었다. 셀로비오스는 셀룰로스의 효소의 가수분해로부터의 중간 생산물(intermediate product)이고, 이는 엑소셀룰라아제(exocellulases), 엔도셀룰라아제(endocellulases) 및 β -글루코시다아제를 포함하는 셀룰라아제의 콕테일(cocktail)에서 β -글루코시다아제에 의해 글루코스로 추가적으로 전환될 수 있는 반면에, 오탄당 당은 헤미셀룰로스(hemicellulose)의 희석된 산 가수분해(dilute acid hydrolysis)의 산물이다. 셀로비오스를 글루코스로 가수분해할 수 있는 β -글루코시다아제 및 셀로비오스 운반체 둘 다 결합되어 있기 때문에 야생형 *S. 세레비시아*는 셀로비오스를 축적할 수 없다. 이런 이유로, 최근에 발견된 셀로텍스트린 운반체 유전자를 실시예 9에 나타내었고, *N. 크라사*로부터의 β -글루코시다아제 유전자가 *S. 세레비시아*에서 공동-발현되었고, 자일로스 및 셀로비오스의 혼합물을 탄소원으로서 사용하였다(도 63). *S. 세레비시아*로부터의 셀로비오스 발효를 가능하게 하기 위해, 유사한 접근은 β -글루코시다아제의 분비(secretion) 또는 세포 표면 전시(cell surface display)를 사용하였다(van Rooyen 등의, 2005; Skory 등의, 1996; Kotaka 등의, 2008; Katahira 등의, 2006). 이러한 경우에, 셀로비오스는 *S. 세레비시아*의 내인성 육탄당 수송 시스템(endogenous hexose transport system)에 의해 수송되기 전에 세포 밖에서(extracellularly) 글루코스로 가수분해하였다. 그에 반해서, 이러한 전략에서, 셀로비오스는 세포 내로 수송된 다음에 가수분해되었다.

[0523] *S. 세레비시아*에서 혼합된 당 발효를 위한 일반적인 방법에서, 리그노셀룰로스(lignocellulose)로부터 유도된 글루코스 및 오탄당 당(pentose sugars)의 혼합물을 사용하였다. 그러나, 이러한 새로운 전략(strategy)에서, 셀로비오스 및 오탄당 당의 혼합물을 사용하였다. 셀로비오스는 이종의 셀로텍스트린 운반체(heterologous celloextrin transporters)를 통해 효모 세포 내로 수송된 반면에, 오탄당 당은 내인성 육탄당 운반체(endogenous hexose transporters)에 의한 효모 세포 내로 수송되었고, 이로 인하여 글루코스 저해에 부분적으로 책임이 있는 페노메논(phenomenon)인, 동일한 운반체에 의한 글루코스 및 오탄당 당 사이의 직접적인 경쟁이

제거되었다. 효모 세포 내에서 일단은, 셀로비오스는 β -글루코시다아제에 의해 글루코스로 전환되고, 효모 세포에 의해 즉시 소비되고, 그 결과로 이는 세포내 글루코스 농도가 낮아지는 결과가 발생하고, 이로 인하여 추가적으로 글루코스 저해가 완화되었다.

[0524] 유전자 조작된(engineered) 자일로스-이용 효모 스트레인 L2612 은 셀로텍스트린 운반체 및 β -글루코시다아제 유전자가 공동-발현되는(co-express) 숙주(host)로서 사용되었다. 이러한 스트레인에서, 자일로스 환원효소 및 자일리톨 탈수소효소 및 *피치아 스티피티스*로부터의 자일루로키나아제(xylulokinase)로 구성된 D-자일로스 이용 경로(utilization pathway)는 염색체 내에 통합되었다. NCU008011, NCU08114 및 NCU00809 를 포함하는 *네우로스포라 크라사*로부터의 셀로텍스트린 운반체, 및 *네우로스포라 크라사*로부터의 하나와 *아스페르길루스 아쿨레아투스*(*Aspergillus aculeatus*)로부터의 다른 하나를 포함하는 두 개의 β -글루코시다아제 유전자를 평가하였다(evaluated).

[0525] 플라스미드를 유지하기 위해, S. 세레비시아 L2612(*MATa*, *leu2-3*, *leu2-112*, *ura3-52*, *trp1-298*, *can1*, *cyn1*, *gal+*)를 합성 드롭아웃 배지(synthetic dropout media)에서 배양하였다(아미노산 및 황산암모늄이 없는 0.17 %의 Difco 효모 질소 염기, 0.5 %의 황산암모늄 및 0.05 %의 아미노산 드롭아웃 혼합물). 2 % 당을 함유한 YPA 배지[1 %의 효모 추출물, 2 % 펩톤, 0.01 %의 아데닌 헤미설페이트(adenine hemisulfate)]는 효모 스트레인을 성장시키는데 사용하였다.

[0526] D-자일로스 환원효소, 자일리톨 탈수소효소 및 *피치아 스티피티스*로부터의 자일루로키나아제(xylulokinase)로 구성된 D-자일로스 이용 경로를 통합하기 위해, 이에 상응하는 유전자를 PCR-증폭하였고, DNA 조립 방법(DNA assembler method)을 사용한 pRS416 플라스미드 내로 클론시켰다(Shao *et al.*, 2009). *Bam*HI 및 *Hind*III 는 D-자일로스 이용 경로를 코딩하는 DNA 단편을 제거하기 위해 사용한 다음에, 동일한 두 가지의 제한 효소(restriction enzymes)에 의해 소화된(digested) pRS406 플라스미드 내로 결합되었다(ligated). 결과적으로 생성된 플라스미드를 *Apa*I 에 의해 선 모양으로 만든 다음에, L2612 의 염색체에서의 *URA3* 로커스(locus) 내로 통합되었다.

[0527] pRS425 플라스미드(미국 매사추세츠주 입스위치에 위치한 New England Biolabs)는 셀로텍스트린 운반체 유전자 및 β -글루코시다아제 유전자가 공동-발현시키는데(co-express) 사용하였다. 도 64 에 나타난 바와 같이, pRS425 플라스미드를 *Bam*HI 및 *Apa*I 로 소화시켰다. *PYK1* 프로모터 및 *ADH1* 종결자는 각각 셀로텍스트린 운반체의 N-말단 및 C-말단에 첨가한 반면에, *TEF1* 프로모터 및 *PGK1* 종결자를 각각 β -글루코시다아제의 N-말단 및 C-말단에 첨가하였다. 이러한 DNA 단편을, DNA 조립 방법을 사용하여 선으로 된(linearized) pRS425 셔틀 벡터 내로 조립하였다(Shao *등의.*, 2009). *네우로스포라 크라사*로부터의 이러한 셀로텍스트린 운반체 유전자 NCU00801(XM_958708), NCU08114(XM_958780) 및 NCU00809(XM_959259), 및 *네우로스포라 크라사*로부터의 두 개의 β -글루코시다아제 유전자 NCU00130 (XM_951090) 및 *아스페르길루스 아쿨레아투스*로부터의 *BGL1*(D64088)를 사용하였다. 하나의 셀로텍스트린 운반체 유전자 및 하나의 β -글루코시다아제 유전자 각각을 갖는 전체로서 6 가지의 조합물(combinations)이 있다.

[0528] 효모 플라스미드를, 재조합 DNA 조작(recombinant DNA manipulation)에 사용된 *E. coli* DH5a 내로 이동시켰다(transferred). 형질전환체(transformants)를 00 mg/L 암피실린(ampicillin)을 포함하는 Luria broth 플레이트에 도말하였다(plate). *E. coli* 형질전환체를, 액체 Luria broth 배지(미국 펜실베이니아주 피츠버그에 위치한 Fisher Scientific) 내에서 배양시킨 다음에 37 °C 및 250 rpm 에서 성장시켰다. 플라스마를 QIAprep Spin Miniprep Kit(QIAGEN)를 사용하여 *E. coli* 로부터 분리하였다. 이러한 플라스미드를 다음의 스트레인을 수득하기 위해 개별적으로 L2612 스트레인 내에 형질전환시켰다(transformed) : SL01[*네우로스포라 크라사*로부터의 NCU00801 셀로텍스트린 운반체 유전자 및 NCU00130 β -글루코시다아제 유전자를 포함하는(harboring) 플라스미드를 함유], SL02(*네우로스포라 크라사*로부터의 NCU00809 셀로텍스트린 운반체 유전자 및 NCU00130 β -글루코시다아제 유전자를 포함하는 플라스미드를 함유), SL03(*네우로스포라 크라사*로부터의 NCU08114 셀로텍스트린 운반체 유전자 및 NCU00130 β -글루코시다아제 유전자를 포함하는 플라스미드를 함유), SL04(*아스페르길루스 아쿨레*

아투스로부터의 NCU00801 셀로텍스트린 운반체 유전자 및 BGL1 유전자를 포함하는 플라스미드를 함유), SL05(*아스페르길루스 아쿨레아투스*로부터의 NCU00809 셀로텍스트린 운반체 유전자 및 BGL1 유전자를 포함하는 플라스미드를 함유), 및 SL06(*아스페르길루스 아쿨레아투스*로부터의 NCU08114 셀로텍스트린 운반체 유전자 및 BGL1 유전자를 포함하는 플라스미드를 함유). SL00 스트레인을 획득하기 위해, 비어있는(empty) pRS425 플라스미드를 L2612 스트레인 내로 형질전환시켰고(transformed), 이를 음성 대조군(negative control)으로서 사용하였다. 효모 형질전환(transformation)을 표준 리튬 아세테이트 방법을 사용하여 실시하였다(Gietz 등의, 1995). 결과적으로 생성된 형질전환 혼합물을 2 % D-글루코스로 보충된 SC-Ura-Leu 배지에 도말하였다(plated).

[0529] DNA 조립 방법을 사용한 플라스미드의 적절한 구조를 확인하기 위해, 플라스미드를 Zymoprep Yeast Plasmid Miniprep II 키트(미국 캘리포니아주 오렌지에 위치한 Zymo Research)를 사용하여 효모 세포로부터 분리한 다음에, *E. coli* DH5a 세포 내로 이동시켰다(transferred). 결과적으로 생성된 세포를 100 mg/L의 아피실린을 포함하는 LB 플레이트에 확산시켰다(spread). 단일 *E. coli* 콜로니를 LB 액체 배지 내로 배양시켰다. 플라스미드를 QIAprep Spin Miniprep 키트(미국 캘리포니아주 발렌시아에 위치한 QIAGEN)를 사용하여 *E. coli*로부터 분리하였고, 진단(diagnostic) PCR 또는 *Cla*I 및 *Hind*III를 사용한 제한 소화(restriction digestion)로 체크하였다. 모든 제한 효소를 New England Biolabs (미국 매사추세츠주 입스위치에 위치)로부터 입수하였다. 모든 화학물질은 Sigma Aldrich 또는 Fisher Scientific로부터 입수하였다.

[0530] 혼합된 당 발효 연구에 대한 충분한 세포를 획득하기 위해, 각각의 효소 스트레인에 대해서, 단일 콜로니를 2 % 글루코스를 더한 2 mL의 SC-Ura-Leu 배지에서 첫 번째로 성장시킨 다음에, 250 mL의 진탕 플라스크에 50 mL의 동일한 배지 내에서 배양시켰다. 하루 성장시킨 후에, 세포를 스피ندا운시키고(spun down), 250 mL의 언배플드 진탕-플라스크(unbaffled shake- flask)에서 4 % 셀로비오스 및 5 % D-자일로스, 또는 4 % 셀로비오스, 5 % 자일로스, 및 0.5 % 글루코스, 또는 4 % 셀로비오스, 5 % 자일로스, 및 1 % 글루코스로 보충된 50 mL의 YPA 배지에서 배양시켰다. 초기의 OD₆₀₀ ~1로부터의 개시하여, 세포 배양물을 산소 제한 조건 하에서 발효를 위해 30 °C에서 100 rpm으로 성장시켰다. OD₆₀₀ 리딩(reading) 및 세포 배양물 샘플을 다양한 시점에서 채취하였다. 당 농도를 HPLC를 사용하여 분석한 반면에, 에탄올 형성을 에탄올 키트(독일 다름슈타트에 위치한 R-biopharm)를 사용하여 분석하였다. 각각의 측정점(data point)에 대해서, 세 번의 샘플을 채취하였다. SL00 내지 SL06까지의 범위의 스트레인에 대해서 혼합된 당 발효 결과를 도 65에 나타내었다. 최고의 스트레인 SL01은 추가적인 특성 해석(characterization)을 위해 선별되었다.

[0531] SL01 내지 SL066까지의 6가지 상이한 스트레인의 전체는, 하나의 셀로텍스트린 운반체 유전자 및 하나의 β-글루코시다아제 유전자를 포함하는 pRS425 플라스미드를 L2612 스트레인 내로 도입시켜서 구성하였다. 각각의 플라스미드에서, 셀로텍스트린 운반체 유전자 및 β-글루코시다아제 유전자는 각각 효모 프로모터 및 종결자에 첨가하였고, DNA 10 조립 방법(DNA 10 assembler method)으로 pRS425 다중-복제 플라스미드(multi-copy plasmid) 내로 조립되었다(Shao 등의, 2009)(도 64). SL00 스트레인을 획득하기 위해 비어있는(empty) pRS425 플라스미드를 L2612 스트레인 내로 도입하였고, 이를 음성대조군으로 사용하였다. 모든 스트레인을 진탕 플라스크에서 40 g/L 셀로비오스 및 50 g/L D-자일로스의 혼합물과 함께 배양하였고, 이들의 당 소비 비율, 세포 성장 비율 및 에탄올 역가(titers)를 결정하였다(도 65). 모든 스트레인 중에서, *넵루로스포라 크라사*로부터의 β-글루코시다아제 및 셀로텍스트린 운반체 NCU00801를 포함하는 SL01 스트레인은 가장 높은 당 소비 및 에탄올 생산성을 나타내었다. 따라서, 이러한 스트레인을 추가적인 특성 해석을 위해 선택하였다.

[0532] SL01 및 SL00 둘 다를 진탕-플라스크 및 바이오리액터(bioreactors) 둘 다에서, 40 g/L 셀로비오스 및 50 g/L D-자일로스의 혼합물을 사용하여 배양하였다(도 66). 진탕-플라스크 배양에서(도 66A-B), 83 % 셀로비오스는 SL100(0.33 g/L/h부터 0.46 g/L/h까지)과 비교하여 평균 D-자일로스 소비율(consumption rate)이 41.2 % 보다 높으면서, SL01에 의한 96시간에 83 % 셀로비오스가 소비되었다. 증진된 당 소비율과 일치하는, 1.32배의 증가된 평균 바이오매스 성장비율이 관찰되었다(0.031 g의 세포건조무/L/h 내지 0.072 g의 세포건조무/L/h까지). 에탄올 생산성은 0.07 g/L/h부터 0.23 g/L/h까지 2.1배 보다 증가되었다. 0.31 g/당 g (0.31 g per g sugar)의 가장 높은 에탄올 수득은 48시간에 이르렀고, SL00 스트레인과 비교하여 23 %의 증가를 나타

낸, 평균 에탄올 수득은 0.28 g/당 g(0.28 g per g sugar)이었다. SL01 배양에서, 글루코스 및 D-자일로스의 공동-발효에서의 글루코스 억압(repression)의 특징(hallmark)인 유도기(lag phase)없이 보다 빠른 D-자일로스 소비율이 관찰되었다. 게다가, 증진된 바이오매스 성장(enhanced biomass growth) 및 에탄올 생산이 또한 관찰되었다.

[0533] Multifors system(스위스 보트민겐에 위치한 Infors-HT)을 바이오리액터에서 혼합된 당 발효를 위해 사용하였다. 각각의 용기(vessel)는 750 mL 의 전체 용량 부피를 가졌다. 각각의 용기에 대해서, pO₂ 감지기(sensor), 공기 스파아저(air sparger), 출구 가스 냉각기(exit gas cooler), 온도 감지기, 배양 포트(inoculation port), 예비 부품(spare port), 침적관(dip tube), 소포 감지기(antifoam sensor), pH 감지기, 구동축(drive shaft), 블럭히터(heater block), 로터미터(rotameter), 연동 펌프 시스템(peristaltic pumps system)의 각각의 세트가 있다. 전체 바이오리액터 시스템은 냉각 시스템 ThermoFlex900(미국 매사추세츠주 윌섬에 위치한 Thermo Scientific)가 갖추어져있다.

[0534] 효모 스트레인의 단일 콜로니를 2 % 글루코스에 더하여 2 mL SC-Ura-Leu 배지에서 첫 번째로 성장시킨 다음에, 혼합된 당 발효 연구를 위해서 충분한 세포를 수득하기 위해 250 mL 의 진탕 플라스크에서 50 mL 의 동일한 배지 내에서 접종하였다. 하루 성장시킨 후에, 10 mL 의 포화 배양물을, 4 % 셀로비오스 및 5 % D-자일로스, 또는 4 % 셀로비오스, 5 % 자일로스, 및 0.5% 글루코스, 또는 4 % 셀로비오스, 5 % 자일로스, 및 1 % 글루코스로 보충된 400 mL 의 YPA 배지에 접종하였다. 온도를 30 °C 로 유지시켰고, pH 는 2 N H₂SO₄ 또는 4 N NaOH 를 첨가하여 pH 5.5 를 유지시켰다. 처음 48 시에, 250 rpm 의 임펠러 속도(impeller speed)로 0.5 L/min 으로 송풍량(air flow rate)을 유지시켰다. 그 뒤에, 산소 제한 조건 하에서 높은 에탄올 생산을 달성하기 위해, 송풍량을 0.2 L/min 으로 조절하였다. 다양한 시점에 세 번의 샘플을 채취하였고, OD₆₀₀, 당 농도 및 에탄올 농도를 상기한 바와 같이 결정하였다.

[0535] 바이오리액터 배양에서(도 66C-D), 44 % 증가된 D-자일로스 소비율(0.47 g/L/h 에서 0.68 g/L/h 까지) 및 1.1-배 증가된 바이오매스 성장율(0.08 g 세포건조무/L/h 에서 0.17 g 세포건조무/L/h 까지)을 나타내면서, 48 시간 내에 대부분 모든 셀로비오스 및 66 % D-자일로스는 소비되었다. 에탄올 생산성은 4.3-배(0.09 g/L/h 에서 0.50 g/L/h 까지)보다 증가 되었고, 에탄올 수득은 0.39 g/당 g 이다. 진탕-플라스크 배양과 비교하여서, 처음 24 시간에서의 당 소비율은 낮았고, 이는 배지 배양의 시작에 사용된 낮은 세포 밀도 때문이다.

[0536] 예상외로, 비록 발효에 글루코스를 첨가하지 않았을지라도, 작은 양의 글루코스가 발견되었다(도 66A-B). 최고의 글루코스 농도가 진탕-플라스크(12.1 g/L) 및 바이오리액터(bioreactors)(17.5 g/L) 둘 다에서 대략 24 시에 도달하였고, 그리고 난 다음에 매우 낮은 수준에 떨어졌다. 그러나 명백한 글루코스 억압이 이러한 글루코스의 존재 하에서조차도 관찰되지 않았다. 글루코스가 SL00 스트레인에서 발견되지 않았기 때문에, β-글루코스가 이의 에피머 α-글루코스로 천천히 전환(slow conversion)되면서 세포외 글루코스가 발생할 수 있고, 글루코스의 주요한 형태(main form)는 해당작용(glycolysis)에서 사용되었다. 일반적으로, 글루코스에서 이의 상대적으로 낮은 농도 때문에, β-글루코스는 효소적으로(enzymatically) 또는 화학적으로(chemically) α-글루코스로 효율적으로 전환될 수 있다(Bouffard 등의, 1994). 그러나, 유전자 조작된 SL01 스트레인에서, β-글루코시다아제에 의한 촉매에 의해서, 과량의 β-글루코스는 세포 내에서 셀로비오스로부터 생산되었고, β-갈락토스와 함께 관찰된 것과 유사한, 작은 분획이 세포 바깥쪽으로 분비될 것이다(Bouffard 등의, 1994).

[0537] 소량의 글루코스(전체 당의 10 % 미만)가 산업현장(industrial settings)에서 리그노셀룰로스를 가수분해물(lignocellulosic hydrolysates)에 일반적으로 존재하기 때문에, 유전자 조작된(engineered) SL01 스트레인의 발효 성과(fermentation performance)는 셀로비오스, D-자일로스 및 글루코스의 혼합물을 사용하여 조사하였다. 글루코스의 두 가지의 농도, 5 g/L 또는 10 g/L 를, 혼합된 탄소원으로서 40 g /L 셀로비오스 및 50 g/L D-자일로스와 함께 결합하였다. 5 g/L 글루코스(도 67A-B)를 사용하여, 67.8 % D-자일로스가 배치 배양(batch cultivation)에서 48 시간에 소비되면서, 81.7 % 의 셀로비오스를 SL01 에 의해 소비되었다. D-자일로스 소비

율(consumption rate)은 0.32 g/L/h 부터 0.69 g/L/h 까지, 1.19-배 증가되었다. 에탄올 생산성은 3.3-배 증가된 반면에(0.11 g/L/h 부터 0.46 g/L/h 까지), 에탄올 수득은 0.26 g/g 당(g per g sugar) 부터 0.33 g/g 당으로 증가되었다. 10 g/L 글루코스(도 67C-D)를 사용하여, 74.7 % D-자일로스가 배치 배양(batch cultivation)에서 48 시간에 소비되면서, 83.8 % 의 셀로비오스를 SL01 에 의해 소비되었다. D-자일로스 소비율(consumption rate)은 0.45 g/L/h 부터 0.76 g/L/h 까지, 68 % 로 증가되었다. 에탄올 생산성은 2.1-배 증가되었고(0.16 g/L/h 부터 0.50 g/L/h 까지), 에탄올 수득은 0.30 g/g 당(g per g sugar) 부터 0.33 g/g 당으로 증가되었다. 예상대로, 유전자 조작된(engineered) SL01 스트레인은, SL00 야생형 스트레인 보다, 보다 높은 효율성의 당 소비 및 보다 높은 비율의 에탄올 생산을 나타내었다. 보다 중요하게, 글루코스가 전체 당의 10 % 까지와 함께, 세 가지 당의 공동-발효(co-fermentation)에서 현저한 글루코스 저해가 없었고(도 67C-D), 이는 이러한 접근이 산업에서의 적용(industrial applications)에 대해서 실행가능할 수 있음을 암시한다.

[0538] 유사한 연구를, 세 가지 *N. 크라사* 셀로텍스트린 운반체 NCU00801, NCU08114 및 NCU00809 를 글루코시다아제 NCU00130 와 함께 도입된, *S. 세레비시아* 스트레인 D452-2 에서 실행하였다. 형질전환체(transformants)를, 세포내 β-글루코시다아제(NCU00130)가 발현된 20 g/리터 셀로비오스를 포함하는 YSC 배지에서 선별하였다. 이러한 작업에서 사용된 스트레인 및 플라스미드를 표 17 에 나타내었다(Ex. 12). 사용된 프라이머를 표 28 에 나타내었다.

[0539] [표 28]

본 연구에서 사용된 합성 올리고뉴클레오티드(synthetic oligonucleotides)

이름	서열
NCU00801-F	ATGGATCCAAAAATGTCGTCTCACGGCTCC <i>SEQ ID NO: 92</i>
NCU00801- R	ATGAATTCCTACAAATCTTCTTCAGAAATCAATTTTGTTC GCAACGATAGCTTCGGAC <i>SEQ ID NO: 93</i>
NCU08114-F	ATACTAGTAAAAATGGGCATCTTCAACAAGAAGC <i>SEQ ID NO: 94</i>
NCU08114- R	GCATATCGATCTACAAATCTTCTTCAGAAATCAATTTTGTTC CAGCAACAGACTTGCCCTCATG <i>SEQ ID NO: 95</i>
NCU00130-F	GCATACTAGTAAAAATGTCTCTTCCTAAGGATTCCTCT <i>SEQ ID NO: 96</i>
NCU00130- R	ATACTGCAGTTAATGATGATGATGATGGTCCTTCTTGA TCAAAGAGTCA AAG <i>SEQ ID NO: 97</i>

[0540]

[0541] 각각 자일로스 또는 셀로비오스 발효 실험에 대해서 접종물(inoculums)을 제조하기 위해, 효모를 20 g/L 의 글루코스 또는 20 g/L 의 셀로비오스를 포함하는 YP 배지에서 성장시켰다. 20 g/L 의 글루코스 또는 셀로비오스를 포함하는 YP 배지로부터의 중기-대수증식기(mid-exponential phase)를 채취하였고, 멸균수(sterilized water)로 2 번 세척한 후에 접종하였다(inoculated). 모든 플라스크 발효 실험은, 산소 제한 조건 하에서 초기의 OD₆₀₀ 의 0.1 에서, 30 °C 에서 250 mL 의 플라스크에 40 g/L 또는 80 g/L 의 자일로스를 포함하는 YP 배지의 50 mL 를 사용하여 실시하였다. 바이오리액터 발효(Bioreactor fermentations)는, 산소 제한된 250 조건 하에서 200 rpm 의 교반 속도(agitation speed)로 30 °C 에서, Sixfors Bioreactors (Appropriate Technical Resources, Inc)를 사용한 적절한 양의 당을 포함하는 400 mL 의 YP 배지에서 실시하였다. 초기 세포 밀도(Initial cell densities)는 OD₆₀₀ = 1.0 을 조절되었다.

[0542] 세포 성장을 UV-가시광선 분광광도계(UV-visible Spectrophotometer)(미국 뉴욕주 서모에 위치한 Biomate 5)를 사용한 600 nm 에서의 광학 밀도(OD)로 모니터하였다. 글루코스, 자일로스, 자일리톨, 글리세롤, 아세테이트

및 에탄올 농도를, 265 a Rezex ROA-유기산 H+ (8%) 컬럼(265 a Rezex ROA-Organic Acid H+ (8%) column)(미국 캘리포니아주 토런스에 위치한 Phenomenex Inc.)을 사용한 시차 굴절률 검출기(refractive index detector)가 갖추어진 고성능 액체 크로마토그래피 265(HPLC, Agilent Technologies 1200 Series)로 결정하였다. 컬럼을 50 °C 에서 0.6 mL/min 의 유량(flow rate)으로 0.005 N 의 H₂SO₄ 으로 용출시켰다.

[0543] 셀로비오스가 단일 탄소원이면, 세 가지 모두의 형질전환체(transformants)는 에탄올을 성장하고 생산할 수 있지만(도 68), 세 가지의 형질 전환체는 상이한 셀로비오스 발효 비율을 나타내었다(NCU00801 > NCU08114 > NCU00809). NCU00801 및 NCU00130 둘 다가 발현된 가장 빠른 셀룰로스-발효 형질전환체(D801-130)는, 16.8 g/L 의 에탄올을 생산하면서, 4 시간 내에 40 g/L 의 셀로비오스를 소비하였다. 셀로비오스 발효의 부피 생산성(volumetric productivity)($P_{\text{에탄올/셀로비오스}} = 0.7 \text{ g/L/h}$)은 글루코스 발효($P_{\text{에탄올/글루코스}} = 1.2 \text{ g/L/h}$) 보다 낮았고, 셀로비오스로부터의 에탄올 수득($Y_{\text{에탄올/셀로비오스}} = 0.42 \text{ g/g}$)은 동일한 배양 조건 하에서, 글루코스으로부터의 에탄올 수득($Y_{\text{에탄올/글루코스}} = 0.43 \text{ g/g}$)과 거의 같다. 그러나, D801-130 에 의한 관찰된 셀로비오스 소비율(cellobiose consumption rate) 및 에탄올 수득은, β -글루코시다아제의 표면 전시(surface display)를 통해 셀로비오스를 발효시키기 위해 유전자 조작된 *S. 세레비시아* 스트레인 상에서 개선되었다(Kotaka 등의, 2008; Nakamura 등의, 2008). 이러한 결과는, *S. 세레비시아*에서 NCU00801 및 NCU00130 의 동시 발현은 효율적인 셀로비오스 발효를 야기할 수 있음을 나타낸다.

[0544] 효율적인 자일로스 발효 스트레인 DA24-16(실시예 13 에 기재)을 성장시킨 후에, 셀로텍스트린 운반체 및 β -글루코시다아제를 코딩하는 유전자(NCU00801 및 NCU00130) 효소를, 이것이 셀로비오스 및 자일로스를 동시에 소비할 수 있는 스트레인 내로 도입시켰다. 이는, 셀로비오스의 세포내 가수분해(intracellular hydrolysis)로 인하여, 자일로스 이용(xylose utilization)의 글루코스 저해는 이러한 스트레인에서 완화되는 가설을 세울 수 있다. NCU00801 유전자는 DA24-16 의 게놈(genome) 내로 통합되었고, NCU00130 은 다중-복제 플라스미드(multi-copy plasmid)로부터 발현되었다. 결과적으로 생성된 형질전환체, DA24-16-BT3 는 단일 탄소원(sole carbon source)으로서 셀로비오스를 함유하는 한천 평판(agar plate)에서 선별되었다.

[0545] 다양한 양의 셀로비오스 및 자일로스를 포함하는 배지에서 성장한 DA24-16- BT3 스트레인은 셀로비오스 및 자일로스를 공동-소비하였고(co-consumed), 실험한 모든 조건에서 0.38-0.39 g/g 의 수득으로 에탄올을 생산하였다(도 69). 공동-발효의 잠재적인 상승 효과(potential synergistic effects)를, 세 가지의 상이한 조건 하에서 DA2416- BT3 를 배양하여 테스트하였다 : 40 g/L 의 셀로비오스, 40 g/L 의 자일로스, 및 40 g/L 의 둘 다의 당(총 80 g/L 의 당). 놀랍게도, DA24-16BT3 은, 별도로 40 g/L 의 셀로비오스 또는 40 g/L 의 자일로스가 소비되기 위해 필요로하는 동일한 기간 내의 80 g/L 의 셀로비오스/자일로스 혼합물을 공동-소비할 수 있었다(co-consume)(도 70). 게다가, DA24-16BT3 은, 단일 당 발효(셀로비오스 또는 자일로스)로부터의 에탄올 수득(0.31 ~ 0.33 g/g)과 비교하여, 셀로비오스 및 자일로스의 혼합물로부터 보다 높은 수득(0.39 g/g)을 갖는 에탄올을 생산하였다. 에탄올 생산성은 또한 공동-발효(co-fermentation) 동안에 0.27 g/L/h 에서 0.65 g/L/h 까지로 대폭 증가되었다. 셀로비오스 및 자일로스의 공동-발효(co-fermentation)는 전체적인 에탄올 수득 및 생산성을 향상시킬 수 있는, 이러한 결과를 나타내었다. 발효 실험(Fermentation experiments)은 셀로비오스 및 자일로스를 유효하게 공동-발효시킬 수 있는, *P. 스티피티스*와 이러한 유전자 조작된 *S. 세레비시아* 스트레인(DA24-16BT3)을 비교하기 위해 또한 실행하였다.

[0546] 에너지케인(energycane)의 조성을 기반으로 한 모조의 가수분해물(simulated hydrolysate)(10 g/L 의 글루코스, 80 g/L 의 셀로비오스, 40 g/L 의 자일로스)를 사용하였다. 상이한 리그노셀룰로스계 식물(lignocellulosic plants)의 조성(composition)은 광범위한 범위로 다양하다. 예를 들어, 에너지 바이오매스 데이터베이스(Energy biomass database)의 US 부(US Department)는 150 이상의 바이오매스 샘플의 조성을 나타내었다(웹페이지 eere.energy.gov/biomass/rn/feedstock_databases.html). 이러한 샘플의 셀룰로스 대 헤미셀룰로스의 비율(cellulose-to-hemicellulose ratios)은 1.4 내지 19 사이이고, 평균은 2.3 이다. 에너지 작물(Energy crops)은 일반적으로 목질의 바이오매스(woody biomass) 보다 높은 헤미셀룰로스 함량(content)을

갖는다. 사탕수수 바가스(sugarcane bagasse), 옥수수 대(corn stover), 수수(sorghum)의 셀룰로스 대 헤미셀룰로스의 평균 비율은 각각 2.0, 1.85 및 2.14 이다. 따라서, 우리는 우리의 모조의 당 실험 설계(our simulated sugar experiment design)에서 2 의 글루칸/자일란 비율을 사용하였다. 유전자 조작된 효모는, 바이오연료 생산을 위한 β -글루코시다아제 활성(activities)에 결핍된, 전통적인 셀룰라아제 카테일과 함께 사용될 것으로 예상될 것이다. 바이오매스 가수분해 과정은, 불완전한 셀룰라아제 카테일(incomplete cellulase cocktails)을 갖는 6-30 % 의 글루칸-대-글루코스의 전환(glucan-to-glucose conversions)이 보고된 바와 같이, 리그노셀룰로스를 가수분해물에서 소량의 글루코스를 야기할 것이다(Medve 등의, 1998). 상기의 모든 요인을 고려하여, 10 g/L 글루코스, 80 g/L 셀로비오스 및 40 g/L 자일로스의 당 조합(combination)을 모조의 당 실험(simulated sugar experiments)에서 선택하였다.

[0547] DA24-16BT3 는 셀로비오스 및 자일로스가 순식간에 공동-소비(co-consuming)되기 전에 글루코스를 첫 번째로 소비하였다. 비록 작은 접종물(inoculums)을 사용하였을지라도($OD_{600} = 1$), 총 130 g/L 의 당을 60 시간 내로 소비하였다. 이와 대조적으로, *P. 스티피티스*는 동일한 배양 조건 하에서 동일한 기간 내에 당 혼합물의 발효를 끝낼 수 없었다(도 71). DA24-16BT3 는 60 시간 내에 48 g/L 의 에탄올을 생산하였다($Y_{\text{에탄올/당}} = 0.37 \text{ g/g}$ 및 $P_{\text{에탄올/당}} = 0.79 \text{ g/L/h}$).

[0548] 셀로비오스 소비 동안의 배지에서 셀로텍스트린의 일시적인 축적(transient accumulation)이 관찰되었다(도 72~73). 축적된 셀로트리오스 및 셀로테트라오스는 셀로비오스의 고갈된 후에 다시 소비되었다. 축적된 셀로텍스트린은 β -글루코시다아제(NCU00130)의 트랜스-글리코실화 활성도(trans-glycosylation activity)(Christakopoulos 등의., 1994)에 의해 생성된 것 같고, 양 방향성(both directions)(세포내 <-> 세포외)으로 셀로텍스트린의 수송을 용이하게 할 수 있는, 셀로텍스트린 운반체(NCU00801)에 의해 분비되었다. 이러한 일시적인 셀로텍스트린의 축적은, 축적된 셀로텍스트린이 유전자 조작된 효모에 의해 결국에 소비되기 때문에, 생산 수확량(product yields)을 아마도 감소시키지 않을 것이다. 그러나, 셀로트리오스 및 셀로테트라오스의 수송 비율(transport rates)이 셀로비오스 보다 더 늦을 수 있기 때문에, 이는 생산성을 감소시킬 수 있다.

[0549] 소량의 글루코스는 공동-발효(co-fermentation) 동안에 배지에서 끊임없이 발견되었다. 훨씬 작은 양의 글루코스 축적은 자일로스를 억제할 수 있기 때문에, 글루코스 수준을 최소한으로 유지시켜야한다. 셀로텍스트린 운반체 및 β -글루코시다아제의 상대적인 발현 수준은 글루코스의 축적에 영향을 미칠 가능성이 있음을 가설을 세울 수 있다. NCU00801 가 효모 계능 내로 통합되었을 때 보다, NCU00801 가 다중-복제 플라스미드(multi-copy plasmid)에 도입시켰을 때 배지에서 보다 많은 글루코스가 축적됨을 관찰된 사실이 이를 지지한다. 잠재적인 이유가 글루코스 저해인(potential reason being glucose repression), 다중-복제 플라스미드에서 NCU00801 및 NCU00130 둘 다를 함유하는 스트레인(DA24-16-BT)은, DA24-16-BT3 에서 관찰된 자일로스 이용 비율 보다, 상대적으로 더 느린 자일로스 이용 비율(xylose utilization rates)을 나타내었다(도 74). 셀로텍스트린 운반체 및 β -글루코시다아제 발현 수준의 추가적인 수정(adjustments) 또는 감소된 트랜스-글리코실화 활성도(trans-glycosylation activities)를 갖는 β -글루코시다아제의 확인은 공동-발효(co-fermentation) 동안에 글루코스 및 셀로텍스트린의 축적을 감소시킬 수 있다.

[0550] 자일로스 및 셀로비오스의 공동-발효는 두 가지의 상이한 효모 스트레인의 혼합된 배양에 의해 또한 성취될 수 있다 : 자일로스-발효(fermenting) DA24-16 스트레인 및 셀로비오스-발효 DA452BT(도 75). 상기에 설명한 바와 같이, 효모 스트레인 DA24-16 는 자일로스-활용 효소 야생형 자일로스 환원효소(XYL1), 돌연변이체 자일로스 환원효소 R276H(*mXYL1*), 자일리톨 탈수소효소(XYL2) 및 자일루로키나아제(XKS1)가 발현되었다(Ex. 12; 표 17). DA452BT 는, 셀로텍스트린 운반체 NCU00801 및 β -글루코시다아제 NCU00130 가 발현될 DA452 가 유전자 조작되어 형성되었다. 혼합된 배양물에서, DA24-16 스트레인은 자일로스를 흡수하고(took up)(도 75A 에 녹색 오각형으로 나타난 자일로스 분자), 효소 XYL1(야생형 및 돌연변이체), XYL2 및 XYL3 을 사용하여 이를 대사 작용하는 반면에, 그 밖의 스트레인 DA452BT 는 운반체 NCU00801 를 사용하여 셀로비오스(도 75A 에 두 개의 빨간 색의 육각형으로서 나타난 셀로비오스 분자)를 흡수할 수 있고, 효소 NCU00130 을 사용하여 셀로비오스를 글루코스로

전환된다. 이런 이유로, 혼합된 배양물은 에탄올을 생산하기 위해 자일로스와 셀로비오스가 공동-발효될 수 있다(도 75 B).

[0551] 이러한 연구는 *S. 세레비시아*에 의한 옥탄당 및 오탄당 당의 공동-발효를 가능하게 하기 위한 새로운 전략을 보여준다. 셀로텍스트린 수송 시스템과 함께 효율적인 자일로스 활용 경로가 결합됨으로써, 글루코스 저해에 의해 초래되는 문제가 극복되었다. 결과적으로, 유전자 조작된 효모는 에탄올에 대한 셀룰로스의 가수 분해물 상승 작용(cellulosic hydrolysates synergistically)에 의해 두 가지의 비-대사 작용 가능한 당(two non-metabolizable sugars)을 공동-발효시켰다. 본원에 기재된 새로운 공동-발효 방법은 당화(saccharification) 및 발효 전선(fermentation fronts) 둘 다에서 리그노셀룰로스계 기술(lignocellulosic technologies)이 증진되었다. 대부분의 전통적인 균류에 의한 셀룰라아제 콕테일은 β -글루코시다아제가 부족하고, 효모에서 유효하게 발효되지 않는 셀로비오스와 함께 셀룰로스 가수분해(cellulose hydrolysis)를 끝낸다. 결과적으로, 추가의 β -글루코시다아제 효소는 셀로비오스를 글루코스로 전환시키기 위해 첨가해야 한다. 셀로비오스/자일로스 공동-발효 효모는, 셀룰로스 당화 과정(cellulose saccharification process)과 관련된 에너지 사용(enzyme usage) 및 비용을 낮추면서 제한된 β -글루코시다아제 활성도를 갖는 이러한 셀룰라아제 콕테일(cellulase cocktails)의 사용을 가능하게 한다. 추가적으로, 셀로비오스 및 자일로스 공동 발효 사이의 시너지 효과(synergy)는 에탄올 생산성(ethanol productivity)을 현저하게 증가시키고, 따라서 발효의 경제성(fermentation economics)을 개선시킨다. 리그노셀룰로스계 물질(lignocellulosic materials)의 예비-처리 및 가수분해로부터 소량의 글루코스의 존재는, 옥탄당 및 오탄당 당 혼합물을 에탄올로 전환시키기 위해 유전자 조작된 효모의 생산 능력에 영향을 주지 않는다.

[0552] 이러한 연구는 식물 바이오매스(plant biomass)로부터 모방된 가수분해물(mimic hydrolysates)인 것으로 여겨지는 다양한 혼합물의 당을 발효시키는 유전자 조작된 *S. 세레비시아* 스트레인의 능력을 측정하는 것을 포함한다. 셀로텍스트린 및 자일로스의 공동-발효를 위한 이러한 스트레인의 능력은 예비-처리된 식물 바이오매스(pre-treated plant biomass)의 동시적인 당화 및 공동-발효(simultaneous saccharification and co-fermentation, SSCF) 동안에 특히 유용하다. SSCF 동안에, 헤미셀룰로스는 산 예비-처리에 의해 처음으로 가수분해되고, 이는 자일로스와 여전히-결정질인 셀룰로스(still-crystalline cellulose)의 형성을 야기한다. 그리고 난 후에, 본원에 기재된 균류에 의한 셀룰라아제(fungal cellulases) 및 효모 스트레인이 첨가되고, 셀룰라아제는 자일로스와 셀로비오스를 에탄올로 공동-전환되게 한다. 이러한 계획에서의 제한된 세포의 글루코스 생산 때문에, 자일로스 활용의 억제(repression)는 감소할 것이고, 공동-발효(co-fermentation)는 신속하고, 상승작용에 의해 진행될 것이다.

[0553] 이러한 연구에서 사용된 *S. 세레비시아*는 실험실 스트레인(laboratory strain)일지라도, 발표된 결과(published results)와 비교하였을 때, 유전자 조작된 스트레인의 발효 수행(fermentation performance)은 매우 인상적이다. 중요한 발효 파라미터(수득 및 생산성)는 발판(platform)으로서의 산업적인 효모 스트레인의 용도에 의해 추가적으로 개선될 수 있다. 이러한 공동-발효 전략(co-fermentation strategy)의 응용(Applications)은 에탄올 생산에 제한될 수 있다. 이는 기초적인 기술이기 때문에, 본원에 나타낸 전략은, 화학적 동반물질(commodity chemicals) 및 증진된 바이오연료를 생산하기 위한, 어떠한 모든 제품 다양화 기술(any other product diversification technologies)과 함께 결합될 수 있다.

실시예 18

[0554] **자일란에서 성장한 *N. 크라사*의 전사체 분석(Transcriptome Analysis)**

[0555] 리그노셀룰로스계 바이오매스는 셀룰로스, 헤미셀룰로스 및 리그닌(lignin)으로 구성되었다. 실시예 1-3 은 *N. 크라사*의 전사체(transcriptome) 및 세크레토믹(secretome) 분석을 통해 셀룰로스에서의 성장에 중요한 유전자의 발견을 기재하였다. 이러한 실시예에서, *N. 크라사* 게놈의 발현 프로파일(expression profile)은, 헤미셀룰로스의 활용에 중요한 유전자를 결정하기 위해 자일란에서의 성장 동안에 조사하였다.

[0556] WT 또는 $\Delta xlnR$ 스트레인의 10 일 된 분생자(Ten day old conidia)는, 100 mL 의 IX Vogel's 염 최소 배지(IX Vogel's salts minimal medium)(2 % 글루코스)에서 10^6 분생자/mL 에 접종하고, 일정한 빛(constant light)으로 25 °C 에서 16 시간 성장시키고, IX Vogel's 배지만(IX Vogel's only medium)으로 세척하였다. 분생자는 배지에서 단일 탄소원(sole carbon source)으로서 2 % 글루코스 또는 2 % Beechwood 자일란을 갖는 100mL IX Vogel's 염 내로 이동시킨 다음에, 4 시간 동안 성장시켰다. 균사체(Mycelia)를 여과에 의해 채취하였고, 액체 질소에서 즉시 급속 냉동시켰다(flash frozen). 전체 RNA 를 제조자의 지시에 따라 TRIzol(Invitrogen)을 이용하여 분리하였고, DNase 로 처리하였다(Turbo DNA-free 키트 ; Ambion) (Kasuga, Townsend 등의, 2005).

[0557] cDNA 합성 및 표시(labeling)를 위해, Pronto kit(카탈로그 번호 제 40076 호 ; Corning)를 사용된 총 RNA 가 10 μ g 당 샘플인 점을 제외하고는 제조자의 상세한 설명(specifications)에 따라 사용하였다.

[0558] 마이크로어레이 혼성(Microarray hybridization) 및 데이터 분석을 사전에 기재한 바와 같이 실시하였다(Tian, Kasuga 등의, 2007). GenePix 4000B 스캐너(scanner)(Axon Instruments)을 사용하여 이미지(images)를 획득하였고, GenePix Pro6 소프트웨어를 사용하여 혼성 신호(hybridization signals)를 정량하고 미가공 데이터(raw data)를 수집하였다. 정규화된 발현 값(Normalized expression values)은 BAGEL(Bayesian analysis of gene expression levels : 유전자 발현 수준의 베이지안 분석) 소프트웨어 프로그램을 사용하여 분석하였다(Townsend and Hartl 2002; Townsend 2004). 354 유전자는 자일란 상에서 성장한 *N. 크라사*에서 2-배 보다 더 유도된 점이 발견되었다. 이러한 목록을 도 76 에 나타내었다.

실시예 19

[0559] **자일란에서 성장한 *N. 크라사*의 세크레토믹스 분석**

[0560] 자일란에서 성장하는 동안의 *N. 크라사*의 세크레토믹스(secretome)은 샷건 프로테오믹스 접근법(shotgun proteomics approach)을 사용하여 분석하였다. 자일란 배양물로부터의 상청액(Supernatants)을 트립신으로 소화시키고, 액체 크로마토그래피 나노-전자분사 이온화 이중 질량분석기(nano-electrospray ionization tandem mass spectrometry)로 분석하였다.

[0561] 질량 분석 샘플(Mass spectrometry samples)을 하기와 같이 제조하였다. *N. 크라사* 야생형 스트레인은 4 일 또는 7 일 동안 2 % 자일란 배지에서 성장시켰다. 배양 상청액을 원심분리로 분리하고, 0.22 μ m 필터를 통해 여과시키고, 10 kDa MWCO PES 스핀 농축기(spin concentrators)로 10 번 농축시켰다. 3.36 mg 의 요소, 5 μ L 의 1 M 트리스 pH 8.5 및 5 μ L 의 100 mM DTT 를 100 μ L 의 농축된 배양 상청액에 더하고, 혼합물을 1 시간 동안 60 °C 로 가열하였다. 가열 후에, 700 μ L 의 25 mM 중탄산 암모늄 및 140 μ L 의 메탄올을 용액에 첨가한 다음에 pH 5.0 의 50 mM 아세트산 나트륨에 50 μ L 의 100 μ g/mL 트립신으로 처리하였다. 트립신을, 기저(basal) pH 에서 약 8-9 시간 동안 인버팅(inverting)하면서 37 °C 에서 밤새 반응시켰다. speedvac 으로 건조시키기 위해 소화 부피(digestion the volume)를 감소시킨 후에, 300 μ L MilliQ 물로 3 번 세척하였다. 최종적인 부피는 100 μ l 이다. TFA 를 0.1 - 0.3 % v/v 으로 첨가하였다. 샘플에서의 잔여 염(Residual salts)을 제조자의 지시에 따라 OMIX 마이크로 추출 피펫 팁(microextraction pipette tips)을 사용하여 제거하였다. 아세토니트릴을 증발(evaporation)로 제거하였다. 샘플 용액은 0.1 % - 1 % TFA 을 갖는 수용액(aqueous solution)이고, 최종 부피는 10 마이크로리터 이상이다.

실시예 20

[0562] **분비된 단백질(secreted proteins)을 코딩하는 것으로 예측된 자일란-유도된 유전자(xylan-induced genes)의 분석**

[0563] 전사체 및 세크레토믹스 분석은 나타난 전체의 71 가지의 유전자 중에서 55 가지가 분비되는 것으로 예측된 결과를 나타내었다. 이러한 유전자를 표 29 에 나타내었다. 결실 스트레인(deletion strains)은 69 가지 유전자 중에

49 가지의 유전자에 대해서 이용할 수 있다. 이러한 46 가지의 유전자 중에서, 6 가지의 스트레인은 이핵 접합체(heterokaryons)이고, 그 외의 남아있는 40 가지의 결실 스트레인은 전체 분비된 단백질, 존재하는 자일로스의 양 및 아조-엔도-자일라나아제 활성도(azo-endo-xy lanase activity)에 대해서 분석하였다. 결과를 도 77 에 나타내었다.

[표 29]

자일란-유도 *N. 크라사* 유전자

유전자 이름	신호 P	데이터	아노테이션(Annotation)
NCU00642	Y	전사 (Transcription)	개연성있는 베타-글루코시다아제
NCU00695	Y	전사	추정상의 단백질 (putative protein)
NCU00798		MS	가설단백질(hypothetical protein)
NCU00937	Y	전사	보존 가설단백질 (conserved hypothetical protein)
NCU01517	Y	전사	글루칸 1,4-알파-글루코시다아제
NCU02136		MS	개연성있는 트랜스알돌라아제(transaldolase)
NCU02252		MS	개연성있는 포스포글리세르산자리움김효소 (phosphoglyceromutase)
NCU02343	Y	전사	알파-L-아라비노푸라노시다아제 A 전구체와 관련
NCU02455	Y	전사	FK506-결합 단백질 2 전구체 (펩티딜-프로릴 cis-트랜스 이성질화효소)
NCU02583	Y	전사	개연성있는 알파-글루코시다아제 전구체 (말타아제)
NCU03013	Y	전사	시토졸 Cu/Zn 과산화물제거효소와 관련 (cytosolic Cu/Zn superoxide dismutase)
NCU03222	Y	전사	추정상의 단백질
NCU03636	Y	전사	
NCU03639	Y	전사	개연성있는 트리아실글리세롤 리파아제 전구체
NCU04202		MS	뉴클레오시드-이인산인산화효소 (nucleoside-diphosphate kinase)
NCU04265	Y	전사	베타-프락토포라노시다아제(beta-fructofuranosidase)와 관련
NCU04388	Y	전사	개연성있는 포스파티딜글리세롤/ 포스파티딜이노시톨 수송 단백질 (phosphatidylglycerol/phosphatidylinositol transfer protein)
NCU04395		MS	베타-1,6-글루카나아제 Negl NEG-1
NCU04415	Y	전사	브레펠딘 A 저항 단백질 (brefeldin A resistance protein)와 관련
NCU04431	Y	MS	엔도-1,3-베타-글루카나아제와 관련
NCU04475	Y	전사	개연성있는 리파아제 B 전구체
NCU04482		MS	가설단백질
NCU04623	Y	전사	베타-갈락토시다아제와 관련
NCU04674	Y	전사	알파-글루코시다아제 b 와 관련
NCU04675	Y	전사	추정상의 단백질
NCU04930	Y	전사	트리아실글리세롤 리파아제와 관련
NCU05137	Y	전사	보존 가설단백질

NCU05143	Y	전사	Rds1 단백질과 관련
NCU05159	Y	전사	개연성있는 아세틸자일란 에스테라아제 전구체
NCU05275		MS	개연성있는 유비퀴틴 융합 단백질 (ubiquitin fusion protein) (유비퀴틴/리보솜 단백질)
NCU05315	Y	전사	가설단백질
NCU05395	Y	전사	보존 가설단백질
NCU05686	Y	MS	개연성있는 세포 벽 단백질 UTR2
NCU05751	Y	전사	아세틸자일란 에스테라아제와 관련
NCU05924	Y	전사	개연성있는 엔도-베타-1,4-D-자일라나아제
NCU05965	Y	전사	추정상의 아라비나아제와 관련
NCU05974		MS	세포 벽 단백질(추정상의 글리코시다아제)과 관련
NCU06364	Y	전사	가설단백질
NCU06380	Y	전사	카테콜아민 업 단백질 (catecholamines up protein)과 관련
NCU06650	Y	전사	보존 가설단백질
NCU06781		MS	개연성있는 베타(1-3) 글루카노실트랜스페라아제(glucanotransferase)
NCU06961	Y	전사	개연성있는 엑소폴리갈락투로나아제(exopolysaccharide)
NCU07067		MS	클래스 I 알파-만노시다아제 IB
NCU07143	Y	전사	
NCU07190	Y	전사	셀룰로스 1,4-베타-셀로비오시다아제 II 전구체와 관련
NCU07200	Y	MS	금속단백질분해효소 MEP1 과 관련
NCU07225	Y	전사	개연성있는 엔도-1,4-베타-자일라나아제 B 전구체
NCU07281		MS	개연성있는 글루코스-6-인산이성질체효소
NCU07787	Y	MS	개연성있는 SnodProt 1 전구체
NCU08131	Y	전사	개연성있는 알파-아밀라아제 전구체
NCU08171	Y	MS	보존 가설단백질
NCU08189	Y	전사	엔도-1,4-베타-자일라나아제와 관련
NCU08384		MS	개연성있는 D-자일로스 환원효소
NCU08418		MS	트리펩티딜-펩티다아제 I 과 관련
NCU08457	Y	전사	히드로포빈 Ccg-2 CCG-2
NCU08516	Y	전사	알도스 1-에피머화효소(epimerase)와 관련
NCU08750	Y	전사	이소아밀 알코올 산화효소와 관련
NCU08752	Y	전사	에스테라아제와 관련
NCU08755	Y	전사	가설단백질
NCU08909	Y	MS	개연성있는 베타(1-3) 글루카노실트랜스페라아제 겔 3p
NCU08936		MS	포자형성(sporulation)-특이적인 유전자 SPS2 와 관련

[0566]

NCU09024	Y	MS	콜린 탈수소효소와 관련
NCU09133	Y	전사	추정상의 단백질
NCU09170	Y	MS	개연성있는 알파-N-아라비노푸라노시다아제 (arabinofuranosidase)
NCU09175	Y	전사	글루칸 1,3-베타-글루코시다아제 전구체와 관련
NCU09267		MS	글리옥살 산화효소(glyoxal oxidase) 전구체와 관련
NCU09491		MS	페루로일 에스테라아제 B 전구체[카르복실산 에스테라아제의 서브클래스(subclass)]
NCU09923	Y	전사	자일란 1,4-베타-자일로시다아제(xylosidase)와 관련
NCU09924	Y	전사	보존 가설단백질
NCU10040	Y	전사	
NCU10045	Y	전사	

[0567]

- [0568] 샘플을 하기와 같이 제조하였다. 10 일 된 분생자(10 day old conidia)를 10^6 분생자/mL 의 100 mL 의 2 % 자일란 Vogel's 배지에서 성장시켰다. 두 가지의 복제물(replicate)을 각각의 스트레인에 대해서 제조하였다. 배양물을 일정한 빛(constant light) 및 220 rpm 으로 25 °C 에서 성장시켰다. 샘플을 4 일에 채취하였고, 상청액을 원심분리로 분리하고 검정에서 사용하였다.
- [0569] 브래드포드 단백질 농도(Bradford protein concentrations)를, 분비된 단백질의 총량을 결정하기 위해 측정하였다. 스톡(Stocks)을 BSA 표준(BSA standards)으로 제조하였다 : 0 µg/mL, 50 µg/mL, 100 µg/mL, 250 µg/mL 및 500 µg/mL. 브래드포드 용액을 1 : 4 로 희석하였다. 멀티 채널 피펫(multichannel pipette)을 96 웰 플레이트 내로 200 µL 의 브래드포드 용액을 피펫으로 재서 옮기기 위해(pipette) 사용하였다. 10 µL 의 샘플 및 10 µL 의 각각의 표준을 첨가하였다. 샘플을 10 분 동안 실온에서 배양하였다. 흡광도를 595 nm 에서 관독하였고, 단백질 농도를 결정하였다.
- [0570] 자일로스를 측정하기 위해 사용된 검정은 Bailey 등의, 1992 (*J Biotech* 23: 257- 270)로부터 수정되었다. 자일로스 표준(Xylose standards)은 H₂O 에서 제조되었다. 농축된 0.8 M 자일로스(10 mL 에서의 1.2 g)에 대해서, 표준은 0 mM, 8 mM(1 : 100 희석 ; 990 µl + 10 µl), 20 mM(1 : 100 희석 ; 975 µl + 25 µl), 40 mM(1 : 100 희석 ; 950 µl + 50 µl), 80 mM(1 : 100 희석 ; 900 µl + 100 µl) 및 160 mM(1 : 100 희석 ; 800 µl + 200 µl). 멀티 채널 피펫(multichannel pipette)을, 딥 웰 96-웰 플레이트(deep well 96-well plate)에 900 µL 의 기질 용액(substrate solution)을 첨가하는데 사용하였다. 기질을 10 분 동안 50 °C 에서 배양시켰다. 100 µL 의 배양 상청액 및 표준을 첨가하고, 5 분 동안 50 °C 에서 배양시켰다. 샘플을 10 분 동안 3,400 rpm 으로 원심분리하였다. 멀티 채널 피펫을, 96-웰 PCR 플레이트 내로 75 µL DNS 용액을 피펫으로 재서 옮기는데(pipette) 사용하였다. 5 µL 의 용액을 반응물로부터 제거하였고, DNS 용액을 포함하는 PCR 플레이트에 첨가하였다. 플레이트를 5 분 동안 PCR 기계에서 99 °C 로 가열하였다. 샘플을 냉각시킨 후에, 이들을 확실히 바닥이 평평한 플레이트(flat-bottomed plates)에 이동시키고, 흡광도를 540 nm 에서 관독하였다. 기질 용액(500 mL)은 beechwood 자일란(5 g; 10 mg/mL), 3M NaOAc, pH 5.0 (8.33 mL; 50 mM) 및 물 (491 mL)을 포함하고, 이를 20 분 동안 오토클레이브하였다(autoclave). DNS 용액(100 mL)은 3, 5-디니트로살리실산(707 mg), NaOH (1.32 g), Rochelle 염[Na K 주석산염(tartrate)](20.4 g), 소듐 메타-바이설페이트(Sodium meta-bisulfate)(553 mg), 페놀(507 µL) 및 물 (94.4 mL)을 포함한다.
- [0571] 아조-엔도-자일라나아제 활성도(Azo-endo-xy lanase activity)를 Megazyme 으로부터의 키트로 측정하였다. 이러한 검정을, 염료와 복잡한 자일란 사슬(xylan chain)로부터의 해방된(liberated) 염료의 양을 분광광도계적으로(spectrophotometrically) 측정한 샘플에서의 엔도-자일라나아제 활성(endo-xy lanase activity) 양을 간접적으로 측정하였다. 보다 많은 효소가 존재할수록, 보다 많은 염료가 방출될 것이다. 모든 상청액 샘플을, 분리된 15 mL 의 Falcon 튜브에서 450 µL 의 Na 아세테이트 완충용액(50 mM, pH 4.5)에 50 µL 의 상청액을 첨가하여 1 : 10 으로 희석하였다. 그 다음에, Falcon 튜브를 약 10 분 동안 예비-향온시켰다(pre-warmed). 기질 용액을 모든 튜브에 대한 모든 샘플(500 µL/샘플)에 첨가하였다. 샘플 및 기질 용액을 이들을 예비-평형시키기 위해(pre-equilibrate) 10 분 동안 40 °C 의 워터 베스(water bath)에 두었다. 500 µL 의 기질 용액을 각각의 1 : 10 희석된 샘플에 첨가하고, 10 분 동안 보텍스하고(vortexed), 10 분 동안 40 °C 에서 배양하였다. 각각의 샘플에 대해서 2.5 mL 의 침전 용액(precipitant solution)(95 % 에탄올)을 첨가하여 반응을 종료하였다. 튜브를 10 분 동안 실온에서 두었다. 튜브를 10 분 동안 보텍스하고(vortexed), 실온에서 10 분 동안 1,000 g 으로 원심분리하였다. 각각의 튜브로부터의 1 mL 의 상청액 용액을 큐벳트(cuvette) 내로 직접적으로 두고, 흡광도를 590 nm 에서 측정하였다. 이러한 공정(procedure)에 대해서 사용된 블랭크(blank)는 2.5 mL 의 침전 용액에 첨가된 500 µL 의 기질 용액으로부터의 상청액이다.
- [0572] 마지막으로, *N. 크라사*에서 헤미셀룰로스의 저하(degradation)에 영향을 미치는 것으로 본원에서 확인된 유전자 의 변형물(modulation)은, 헤미셀룰로스와 같은 식물 세포 벽 성분에서의 성장 및 식물 세포 벽을 분해하는 능

력을 증진시키는 유전자 조작된 스트레인을 용이하게 할 수 있는 것으로 예상된다. 관심있는 유전자(Genes of interest)는 예측된 글루카밀라아제(glucamylase)를 코딩하는 NCU01517 ; 예측된 아라비노푸라노시다아제(arabinofuranosidase)를 코딩하는 NCU02343 ; 보존된 가설단백질(conserved hypothetical protein)을 코딩하는 NCU05137 ; 예측된 아세틸자일란 에스테라아제 전구체(predicted acetylxylnan esterase precursor)를 코딩하는 NCU05159 ; 보존된 가설단백질을 코딩하는 NCU09133 ; 및 가설 단백질을 코딩하는 NCU10040 을 포함한다.

- [0573] 헤미셀룰로스에서 세포의 성장은 NCU01517, NCU09133 또는 NCU10040 에 의해 코딩된 폴리펩티드를 코딩하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하여서 증가될 수 있다. 숙주 세포는, 재조합 폴리뉴클레오티드가 발현된 것과 같은 헤미셀룰로스를 포함하는 배지에서 배양될 것이다. 숙주 세포는 재조합 폴리뉴클레오티드를 포함하지 않는 세포 보다는 이러한 배지에서 보다 빠른 비율로 성장할 것이다.

실시예 21

- [0574] **△NCU05137 스트레인의 추가적인 분석**

- [0575] 실시예 1-3 및 18-20 에 기재된 바와 같이, NCU05137 은, *미스칸터스*, 아비셀 또는 자일란 중 어떠한 것에서 *N. 크라사*의 성장 동안에 과잉발현된(overexpressed) 예측된 분비된 단백질(predicted secreted protein)이다. 아비셀에서 성장한 NCU05137 이 결핍된 *N. 크라사*의 결실 스트레인(deletion strain)은 증가된 엔도글루카나아제(endoglucanase), β -글루코시다아제 및 아비셀라아제 활성도(Avicelase activity)를 나타내었다. 자일란에서 성장한 NCU05137 결실 스트레인은 증가된 아조-엔도-자일라나아제 활성도(azo-endo-xylanase activity)를 나타내었다. 이러한 실시예에서 기재된 바와 같이, △NCU05137 의 상보성(complementation)은, NCU05137 유전자의 손실(loss)로 인한 △NCU05137 에 나타난 표현형(phenotypes)의 다양성을 위해서 수행되었다.

- [0576] *cagl* 프로모터의 조절 하에서 C-말단 GFP 태그(tag)를 갖는 NCU05137 를 포함하는 플라스미드를 생성하였다. *N. 크라사* 분생자는 NCU05137-GFP 구조물(construct)과 함께 형질전환되었다. 실험을 표준 *네우로스포라* 절차에 따라 실시하였다(웹페이지 fgsc.net/Neuro_spora/Neuro_sporaProtocolGuide.htm).

- [0577] 야생형, △NCU05137 및 △NCU05137-NCU05137-GFP 스트레인의 전체 분비된 단백질 및 카복시메틸 셀룰라아제(carboxymethyl cellulase, CMC) 활성도를 측정하였다. 각각의 스트레인의 배양물로부터의 100 μ L 의 상청액을 채취하고, 이를 900 μ L 의 브래드포드 염료(Bradford Dye)에 첨가하고, 595 nm 의 흡광도로 측정하여, 전체 분비된 단백질을 측정하였다. CMC 활성도(CMC activity)를, 각각의 스트레인 배양물 및 아조-CMD 키트(azo-CMC kit)(Megazyme SCMCL)로부터의 20 X 희석된 상청액으로 측정하였다. △NCU05137 녹아웃 스트레인(knockout strains)은 분비된 단백질 및 CMC 활성도의 증가된 수준을 나타내었다. △NCU05137 스트레인 내로의 GFP-태그된 NCU05137 의 도입(Introduction)은 야생형 수준으로 다시 돌아가(back to) 이러한 수준을 감소시켰다(도 78).

- [0578] 게다가, 상보적인 스트레인에서 NCU05137-GFP 의 위치(localization)를 관찰하였다. NCU05137-GFP 는 분생자 및 균사의 팁(hypha tip)의 세포 벽에 위치한다(도 79-80). 이러한 데이터는 GFP-태그된 NCU05137 단백질이 완전히 기능적이고, 이는 이러한 단백질의 생화학적인 활성도를 다루는 정제 및 실험에 사용될 수 있음을 나타낸다.

- [0579] 따라서, NCU05137 의 정상적인 기능(normal function)은 셀룰라아제 및 헤미셀룰라아제 유전자 발현의 도입과 관련된 신호 전달 과정(signaling processes)을 저해할 수 있다. 세포에서의 NCU05137 또는 NCU05137 의 상동(homolog)의 발현의 저해는 저 세포에서의 셀룰라아제 및 헤미셀룰라아제 활성도를 증가시키는 것 같고, 그 결과로 셀룰로스 또는 헤미셀룰로스의 성장을 증가시키는 것 같다. 셀룰로스 또는 헤미셀룰로스에서의 세포의 성장을, NCU05137 을 코딩하는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드(endogenous polynucleotide)를 포함하는

숙주 세포를 제공하여 증가시킬 수 있다. 내인성 폴리펩티드의 발현은 저해될 수 있고, 세포는 셀룰로스 및/또는 헤미셀룰로스를 포함하는 배지에서 배양될 수 있다. 숙주 세포는, 내인성 폴리펩티드의 발현을 저해하지 않는 배지에서의 세포 보다, 상기 배지에서 더 빠른 비율로 성장할 것이다.

실시예 22

NCU07705 의 추가적인 분석

NCU07705 의 발현이 셀룰로스에서 *N. 크라사*의 성장 동안에 상향조절될 수 있음을(upregulated) 발견하였다. NCU07705 를 코드하는 폴리펩티드의 BLAST 분석은, 폴리펩티드가 전사 인자(transcription factors)를 포함하는 많은 C6 징크 핑거 도메인(zinc finger domain)과 높은 유사성(similarity)을 갖는 것으로 나타났다(도 1). 셀룰로스의 이용에서 NCU07705 의 역할을 추가적으로 조사하기 위해, NCU07705 가 결핍된 결실 스트레인(deletion strain)의 표현형(phenotype)은 평가되었다.

△NCU07705 스트레인은 2 % 셀룰로스(아비셀), PASC 또는 단일 탄소원으로서 CMC 에서 성장할 수 없지만(표 30), 수크로스, 자일란 및 자일로스에서 야생형 스트레인에 대해서는 유사한 속도(kinetics)로 성장하였다. NCU07705 가 셀룰라아제의 발현을 조절하는 역할을 하는지 여부를 결정하기 위해, 셀룰라아제 및 헤미셀룰라아제 유전자의 발현을, 셀룰로스에서 △NCU07705 의 성장 동안에 조사하였다. 야생형으로부터의 10 일 된 분생자(FGSC 2489) 및 △NCU07705 스트레인을 Vogel's 액체 MM(2 % 수크로스)(Vogel 1956) 내에 배양시키고 16 시간 동안 성장시켰다. 균사체(Mycelia)를 원심분리하고, 1 x Vogel's 염으로 세척하고 난 다음에, 2 % 수크로스 또는 2 % 아비셀을 갖는 Vogel's 배지 내에 이동시키고, 4 시간 동안 일정한 빛(constant light)에서 성장시켰다. 이들은 여과에 의해 채취하였고, 액체 질소에서 즉시 냉동시켰다. 전체 RNA 를 제조자의 지시에 따라 TRIzol (미국 캘리포니아주 칼즈배드에 위치한 Invitrogen)를 사용하여 분리하였고, DNase(Turbo DNA-free kit, 미국 캘리포니아주 포스터시에 위치한 Ambion/Applied Biosystems)로 처리하였다(Kasuga 등의, 2005). ChipShot™ 간접적으로 라벨된/클린-업 시스템(Indirect Labeling/Clean-Up System)(카달로그 번호 제 Z4000 호, 미국 위스콘신주 매디슨에 위치한 Promega) 및 CyDye 포스트-라벨된 반응성 염료 팩(Post-Labeling Reactive Dye Pack)(카달로그 번호 제 RPN5661 호, 미국 뉴저지주 피스카타웨이에 위치한 GE Healthcare)를, RNA 의 양이 10 µg 을 사용하는 것을 제외한 제조자의 지시에 따라 cDNA 를 합성 및 표지하기(label) 위해 사용하였다. 프로토(Pronto)! 혼성 키트(Hybridization Kit)(카달로그 번호 제 40076 호, 미국 매사추세츠주 로웰에 위치한 Corning)를 제조자의 상세한 설명에 따라 마이크로어레이 혼성화(microarray hybridization)를 위해 사용하였다.

데이터 분석을 사전에 기재된 바와 같이 실시하였다(Tian 등의., 2007). GenePix 4000B 스캐너(미국 캘리포니아주 유니언 시티에 위치한 Axon Instruments)를 사용하여 이미지를 획득하였고, GenePix Pro6 소프트웨어를 혼성 신호(hybridization signals)를 정량하고, 미가공된 데이터(raw data)를 수집하기 위해 사용하였다. 정규화된 발현 값을 BAGEL(Bayesian Analysis of Gene Expression Levels : 유전자 발현 수준의 베이지안 분석)을 사용하여 분석하였다(Townsend and Hartl, 2002). 예측된 셀룰라아제 유전자 중 어떠한 것도 △NCU07705 스트레인에 유도되지 않은 반면에, 예측된 헤미셀룰라아제 유전자의 도입(induction)은 영향받지 않았다(unaffected)(하기의 표 30 참조). 따라서, NCU07705 는 *cdr-1*, 셀룰로스 분해 조절인자 1(cellulose degradation regulator 1)로 명명되었다.

따라서, 셀룰로스에서의 세포의 성장은, NCU07705 에 의해 코드된 폴리펩티드를 코드하는 재조합 폴리뉴클레오티드를 포함하는 숙주 세포의 제공에 의해 증가될 것이다. 숙주 세포는, 재조합 폴리뉴클레오티드가 발현된 것과 같이 셀룰로스를 포함하는 배지에서 배양될 수 있다. 숙주 세포는, 재조합 폴리뉴클레오티드를 포함하지 않는 배지에서의 세포보다, 상기 배지에서 더 빠른 비율로 성장할 것이다.

[0585]

[표 30]

N. 크라사 △NCU07705 스트레인에서의 유전자의 발현 프로파일

7705-스위치 ² (switch)	WT-스위치 ¹	유전자 /로커스 이름	GH 패밀리(Family)	클래스(Class)	up in Avi ³
No	15	NCU00762	5	엔도(endo)-	31.5
No	No	NCU03996	6	CBHII 유사 (like)	
No	168	NCU07190	6	CBHII 유사	119
No	26	NCU09680	6	CBHII	251.3
No	18	NCU04854	7	CBHI 유사	10.8
No	3.8	NCU05057	7	CBHI 유사	7.4
No	No	NCU05104	7	CBHI 유사	
No	93	NCU07340	7	CBHI	382.2
No	2	NCU05121	45	엔도-	17.2
No	5.8	NCU00836	61	엔도-	31
No	3.7	NCU01050	61	엔도-	382.1
No	No	NCU01867	61	엔도-	
No	49	NCU02240	61	엔도-	84
No	No	NCU02344	61	엔도-	4.1
No	6.1	NCU02916	61	엔도-	17.7
No	No	NCU03000	61	엔도-	
No	17	NCU03328	61	엔도-	23.8
No	No	NCU05969	61	엔도-	12.7
No	No	NCU07520	61	엔도-	
No	No	NCU07760	61	엔도-	
No	103	NCU07898	61	엔도-	230
No	No	NCU07974	61	엔도-	
No	25	NCU08760	61	엔도-	44.7

¹ RNA 추출(extraction) 전에 16 시간 동안 Vogel's/수크로스에서 성장시키고, 여과하고 4 시간 동안 Vogel's/아비셀에 재현탁시킨 *N. 크라사*(NCU07705) 배양물로부터의 예측 셀룰라아제 유전자의 발현수준

² RNA 추출(extraction) 전에 16 시간 동안 Vogel's/수크로스에서 성장시키고, 여과하고 4 시간 동안 Vogel's/수크로스에 재현탁시킨 *N. 크라사*(야생형 FGSC 2489) 배양물로부터의 예측 셀룰라아제 유전자의 발현수준

³ 아비셀에서 30 시간동안 성장시킨 야생형(FGSC 2489) 세포의 마이크로어레이 분석(microarray analyses)으로부터 유도된 발현 수준(Tian 등의/ 2009).

[0586]

참조문헌

- Arendt, C. S., K. Ri, *et al.* (2007). "Genetic selection for a highly functional cysteine-less membrane protein using site saturation mutagenesis." *Anal Biochem* 365(2):185-193.
- Bai, F. W., W. A. Anderson, and M. Moo-Young (2008). "Ethanol fermentation technologies from sugar and starch feedstocks." *Biotechnol Adv* 26(1): 89-105.
- Bailey, M. J., P. Biely, and K. Poutanen (1992). "Interlaboratory testing of methods for assay of xylanase activity." *J Biotechnol* 23:257-270.
- Bailey, J. E. and D. F. Ollis (1986). *Biochemical Engineering Fundamentals*, McGraw-Hill Book Company, NY.
- Basso, L. C., H. V. de Amorim, *et al.* (2008). "Yeast selection for fuel ethanol production in Brazil." *FEMS Yeast Res* 8(7):1155-1163.
- Becker, D. M. and V. Lundblad (2001). "Introduction of DNA into yeast cells." *Curr Protoc Mol Biol* Chapter 13, Unit13.7.
- Bendtsen, J. D., H. Nielsen, *et al.* (2004). "Improved prediction of signal peptides: SignalP 3.0." *J. Mol. Biol* 340(4):783-95.
- Bengsston, O., B. Hahn-Hägerdal, and M. F. Gorwa-Grauslund (2009). "Xylose reductase from *Pichia stipitis* with altered coenzyme preference improves ethanolic xylose fermentation by recombinant *Saccharomyces cerevisiae*." *Biotechnol Biofuels* 2:9.
- Bhat, K. M. and R. Maheshwari (1987). "Sporotrichum thermophile Growth, Cellulose Degradation, and Cellulase Activity." *Appl Environ Microbiol* 53(9):2175-2182.
- Bouffard, G. G., K. E. Rudd, and S. L. Adhya (1994). "Dependence of lactose metabolism upon mutarotase encoded in the gal operon in *Escherichia coli*." *J Mol Biol* 244(3):269-278.
- Bouws, H, A. Wattenberg and H. Zorn (2008). "Fungal secretomes--nature's toolbox for white biotechnology." *Appl Microbiol Biotechnol* 80(3):381-388.
- Brat, D., E. Boles, and B. Wiedemann (2009). "Functional Expression of a Bacterial Xylose Isomerase in *Saccharomyces cerevisiae*." *Appl Environ Microbiol* 75(8):2304-2311.
- Canevascini, G. (1988). "Cellobiose dehydrogenase from *Sporotrichum thermophile*." *Method Enzymol* 160:443-448.
- Cantarel, B. L., P. M. Coutinho, *et al.* (2009). "The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics." *Nucleic Acids Res* 37(Database issue):D233-238.
- Chauve, M., H. Mathis, *et al.* (2010). "Comparative kinetic analysis of two fungal beta-glucosidases." *Biotechnol Biofuels* 3(1):3.

[0587]

- Christakopoulos, P., M. K. Bhat, *et al.* (1994). "Enzymatic synthesis of trisaccharides and alkyl beta-D-glucosides by the transglycosylation reaction of beta-glucosidase from *Fusarium oxysporum*." *Int J Biol Macromol* 16(6):331-334.
- Christianson, T.W., R. S. Sikorskim, *et al.*, (1992). "Multifunctional yeast high-copy number shuttle vectors." *Gene*, 110(1):119-122.
- Chu, B. C. and H. Lee (2007). "Genetic improvement of *Saccharomyces cerevisiae* for xylose fermentation." *Biotechnol Adv* 25(5):425-441.
- Colot, H. V., G. Park, *et al.* (2006). "A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors." *Proc Natl Acad Sci USA* 103(27):10352-10357.
- Cullen, D., L. J. Wilson, *et al.* (1987). "Sequence and centromere proximal location of a transformation enhancing fragment *ans1* from *Aspergillus nidulans*." *Nucleic Acids Res* 15(22):9163-9175.
- Davies, G. J., A. M. Brzozowski, *et al.* (2000). "Structure and function of *Humicola insolens* family 6 cellulases: structure of the endoglucanase, Cel6B, at 1.6 Å resolution." *Biochem J* 348 Pt 1:201-207.
- Davis, R. H. (2000). "Neurospora: Contributions of a model organism." New York, Oxford University Press.
- Davis, R. H. and D. D. Perkins (2002). "Timeline: *Neurospora*: a model of model microbes." *Nat Rev Genet* 3(5):397-403.
- de Groot, M. J., W. Prathumpai, *et al.* (2005). "Metabolic control analysis of *Aspergillus niger* L-arabinose catabolism." *Biotechnol Prog* 21(6):1610-1616.
- Dementhon K, G. Iyer, and N. L. Glass (2006). "VIB-1 is required for expression of genes necessary for programmed cell death in *Neurospora crassa*." *Eukaryot Cell* 5(12):2161-2173.
- Doran-Peterson, J., A. Jangid, *et al.* (2009). "Simultaneous saccharification and fermentation and partial saccharification and co-fermentation of lignocellulosic biomass for ethanol production." *Methods Mol Biol* 581:263-280.
- Drissen, R. E. T., R. H. W. Mass, *et al.* (2009). "Modelling ethanol production from cellulose: separate hydrolysis and fermentation versus simultaneous saccharification and fermentation." *Biocatal Biotransfor* 27(1):27-35.
- Dunlap, J. C., K. A. Borkovich, *et al.* (2007). "Enabling a Community to Dissect an Organism: Overview of the *Neurospora* Functional Genomics Project." *Adv Genet* 57:49-96.
- du Preez, J. C., M. Bosch, and B. A. Prior (1986). "The fermentation of hexose and pentose sugars by *Candida shehatae* and *Pichia stipitis*." *Appl Microbiol Biotechnol* 23(3):228-233.
- Eberhart, B. M., R. S. Beck, *et al.* (1977). "Cellulase of *Neurospora crassa*." *J Bacteriol* 130(1):181-186.

[0588]

- Ellis, K. J. and J. F. Morrison (1982). "Buffers of constant ionic strength for studying pH-dependent processes." *Methods Enzymol* 87:405-426.
- Espagne, E., O. Lespinet, *et al.* (2008). "The genome sequence of the model ascomycete fungus *Podospora anserina*." *Genome Biol* 9(5):R77.
- Fonseca, C. R. Romao, *et al.* (2007). "L-Arabinose transport and catabolism in yeast." *FEBS J* 274(14):3589-3600.
- Freer, S. N. (1991). "Fermentation and aerobic metabolism of cellodextrins by yeasts." *Appl Environ Microbiol* 57(3):655-659.
- Freer, S. N. and R. V. Greene (1990). "Transport of glucose and cellobiose by *Candida wickerhamii* and *Clavispora lusitaniae*." *J Biol Chem* 265(22):12864-12868.
- Galagan, J. E., S. E. Calvo, *et al.* (2003). "The genome sequence of the filamentous fungus *Neurospora crassa*." *Nature* 422(6934):859-868.
- Gems, D., I. L. Johnstone, and A. J. Clutterbuck (1991). "An autonomously replicating plasmid transforms *Aspergillus nidulans* at high frequency." *Gene* 98(1):61-67.
- Guindon, S. and O. Gascuel (2003). "A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood." *Syst Biol* 52(5):696-704.
- Hahn-Hägerdal, B., K. Karhumaa, *et al.* (2007). "Towards industrial pentose-fermenting yeast strains." *Appl Microbiol Biotechnol* 74(5):937-953.
- Hector, R. E., N. Qureshi *et al.* (2008). "Expression of a heterologous xylose transporter in a *Saccharomyces cerevisiae* strain engineered to utilize xylose improves aerobic xylose consumption." *Appl Microbiol Biotechnol* 80(4):675-684.
- Himmel, M. E., S. Y. Ding, *et al.* (2007). "Biomass recalcitrance: engineering plants and enzymes for biofuels production." *Science* 315(5813):804-807.
- Hosaka, K., J. Nikawa, *et al.*, (1992). "A dominant mutation that alters the regulation of *INO1* expression in *Saccharomyces cerevisiae*." *J Biochem* 111(3):352-358.
- Jacobson, D. J., A. J. Powell, *et al.* (2004). "Neurospora in temperate forests of western North America." *Mycologia* 96(1):66-74.
- Jeffries, T. W., I. V. Grigoriev, *et al.* (2007). "Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis*." *Nat Biotechnol* 25(3):319-326.
- Kasuga, T. and N. L. Glass (2008). "Dissecting colony development of *Neurospora crassa* using mRNA profiling and comparative genomics approaches." *Eukaryot Cell* 7(9): 1549-64.

[0589]

- Kasuga, T., J. P. Townsend, *et al.* (2005). "Long-oligomer microarray profiling in *Neurospora crassa* reveals the transcriptional program underlying biochemical and physiological events of conidial germination." *Nucleic Acids Res* 33(20): 6469-85.
- Katahira, S., M. Ito, *et al.* (2008). "Improvement of ethanol productivity during xylose and glucose co-fermentation by xylose-assimilating *S. cerevisiae* via expression of glucose transporter Sut1." *Enzyme Microb Tech* 43(2):115-119.
- Katahira, S., A. Mizuike, *et al.* (2006). "Ethanol fermentation from lignocellulosic hydrolysates by a recombinant xylose- and celooligosaccharide-assimilating yeast strain." *Appl Microbiol Biotechnol* 72(6):1136-1143.
- Korkhin, Y., A. J. Kalb(Gilboa), *et al.* (1998). "NADP-dependent bacterial alcohol dehydrogenases: crystal structure, cofactor-binding and cofactor specificity of the ADHs of *Clostridium beijerinckii* and *Thermoanaerobacter brockii*." *J Mol Biol* 278(5):967-981.
- Kotaka, A., H. Bando, *et al.* (2008). "Direct ethanol production from barley beta-glucan by sake yeast displaying *Aspergillus oryzae* beta-glucosidase and endoglucanases." *J Biosci Bioeng* 105(6):622-627.
- Kötter, P. and M. Ciriacy (1993). "Xylose fermentation by *Saccharomyces cerevisiae*." *Appl Microbiol Biotechnol* 38(6):776-783.
- Kubicek, C. P., R. Messner, *et al.* (1993). "Triggering of cellulase biosynthesis by cellulose in *Trichoderma reesei*. Involvement of a constitutive, sophorose-inducible, glucose-inhibited beta-diglucoside permease." *J Biol Chem* 268(26):19364-19368.
- Kumar, R., S. Singh, and O. V. Singh (2008). "Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives." *J Ind Microbiol Biotechnol* 35(5):377-391.
- Kuyper, M, H. R. Harhangi, *et al.* (2003). "High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*? *FEMS Yeast Res* 4(1):69-78.
- Lang, J. M. and V. P. Cirillo (1987). "Glucose transport in a kinaseless *Saccharomyces cerevisiae* mutant." *J Bacteriol* 169(7):2932-2937.
- Leandro, M. J., P. Gonçalves, and I. Spencer-Martins (2006). "Two glucose/xylose transporter genes from the yeast *Candida intermedia*: first molecular characterization of a yeast xylose-H⁺ symporter." *Biochem J* 395(3):543-549.
- Linder, M. and T. T. Teeri (1996). "The cellulose-binding domain of the major cellobiohydrolase of *Trichoderma reesei* exhibits true reversibility and a high exchange rate on crystalline cellulose." *Proc Natl Acad Sci USA* 93(22):12251-12255.
- Lynd, L. R., M. S. Laser, *et al.* (2008). "How biotech can transform biofuels." *Nat. Biotechnol.* 26(2):169-172.

- Lynd, L. R., P. J. Weimer, *et al.* (2002). "Microbial cellulose utilization: fundamentals and biotechnology." *Microbiol Mol Biol Rev* 66(3):506-77, table of contents.
- Madhavan, A., S. Tamalampudi, *et al.* (2009). "Xylose isomerase from polycentric fungus orpinomyces: gene sequencing, cloning, and expression in *Saccharomyces cerevisiae* for bioconversion of xylose to ethanol." *Appl Microbiol Biotechnol* 82(6):1067-1078.
- Martin, F., A. Kohler, *et al.* (2010). "Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis." *Nature* 464(7291):1033-1038.
- Martinez, D., R. M. Berka, *et al.* (2008). "Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*)." *Nat Biotechnol* 26(5):553-60.
- Martinez, D., J. Callacombe, *et al.* (2009). "Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion." *Proc Natl Acad Sci USA* 106(6):1954-1959.
- Martinez, D., L. F. Larrondo, *et al.* (2004). "Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78." *Nat Biotechnol* 22(6):695-700.
- Matsushika, A., S. Watanabe, *et al.* (2008). "Expression of protein engineered NADP+-dependent xylitol dehydrogenase increases ethanol production from xylose in recombinant *Saccharomyces cerevisiae*." *Appl Microbiol Biotechnol* 81(2):243-255.
- Matteucci, M. D. and M. H. Caruthers (1980). *Tetrahedron Lett* 21:719-722.
- McCluskey, K. (2003). "The Fungal Genetics Stock Center: from molds to molecules." *Adv Appl Microbiol* 52:245-262.
- Medve, J., J. Karlsson, *et al.* (1998). "Hydrolysis of microcrystalline cellulose by cellobiohydrolase I and endoglucanase II from *Trichoderma reesei*: adsorption, sugar production pattern, and synergism of the enzymes." *Biotechnol Bioeng* 59(5):621-34.
- Miyasaka, H. (1999). "The positive relationship between codon usage bias and translation initiation AUG context in *Saccharomyces cerevisiae*." *Yeast* 15(8):633-637.
- Nair, N. and H. Zhao (2007). "Biochemical characterization of an L-Xylulose reductase from *Neurospora crassa*." *Appl Environ Microbiol* 73(6):2001-2004.
- Nakamura, N., R. Yamada, *et al.* (2008). "Effective xylose/cellobiose co-fermentation and ethanol production by xylose-assimilating *S. cerevisiae* via expression of β -glucosidase on its cell surface." *Enzyme Microb Tech* 43(3):233-236.
- Noguchi, Y., M. Sano, *et al.* (2009). "Genes regulated by AoXlnR, the xylanolytic and cellulolytic transcriptional regulator, in *Aspergillus oryzae*." *Appl Microbiol Biotechnol* 85(1):141-154.
- Notredame, C., D. G. Higgins, and J. Heringa (2000). "T-Coffee: A novel method for fast and accurate multiple sequence alignment." *J Mol Biol* 302(1):205-217.

[0591]

- Pandit, A. and R. Maheshwari (1996). "Life-history of *Neurospora intermedia* in a sugar cane field." *J Biosci (Bangalore)* 21(1):57-79.
- Pauly, T. A., J. L. Ekstrom, *et al.* (2003). "X-ray crystallographic and kinetic studies of human sorbitol dehydrogenase." *Structure* 11(9):1071-1085.
- Pedelacq, J. D., S. Cabantous, *et al.* (2006). "Engineering and characterization of a superfolder green fluorescent protein." *Nat Biotechnol* 24(1):79-88.
- Perkins, D. D., B. C. Turner, *et al.* (1976). "Strains of *Neurospora* collected from nature." *Evolution* 30: 281-313.
- Ramos, J., K. Szkutnicka, and V. P. Cirillo (1988). "Relationship between low- and high-affinity glucose transport systems of *Saccharomyces cerevisiae*." *J Bacteriol* 170(11):5375-5377.
- Rawat, U. and M. Rao (1997). "Site and significance of cysteine residues in xylose reductase from *Neurospora crassa* as deduced by fluorescence studies." *Biochemical and Biophysical Research Communications* 239(3): 789-93.
- Reifenberger E., E. Boles, and M. Ciriacy (1997). "Kinetic characterization of individual hexose transporters of *Saccharomyces cerevisiae* and their relation to the triggering mechanisms of glucose repression." *Eur J Biochem* 245(2):324-333.
- Romero, M. D., J. Aguado, *et al.* (1999). "Cellulase production by *Neurospora crassa* on wheat straw." *Enzyme Microb Tech* 25: 244-250.
- Roy, A., A. Kucukural, and Y. Zhang (2010). "I-TASSER: a unified platform for automated protein structure and function prediction." *Nat Protoc* 5(4):725-738.
- Rubin E. M. (2008). "Genomics of cellulosic biofuels." *Nature* 454(7206):841-845.
- Ruepp, A., A. Zollner, *et al.* (2004). "The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes." *Nucleic Acids Res* 32(18):5539-5545.
- Runquist, D., C. Fonseca, *et al.* (2009). "Expression of the Gxf1 transporter from *Candida intermedia* improves fermentation performance in recombinant xylose-utilizing *Saccharomyces cerevisiae*." *Appl Microbiol Biotechnol* 82(1):123-130.
- Runquist, D., B. Hahn-Hägerdal, and P. Rådström (2010). "Comparison of heterologous xylose transporters in recombinant *Saccharomyces cerevisiae*." *Biotechnol Biofuels* 3:5.
- Saha, B. C. (2003). "Hemicellulose bioconversion." *J Ind Microbiol Biotechnol* 30:279-291.
- Saloheimo, A., J. Rauta, *et al.* (2007). "Xylose transport studies with xylose-utilizing *Saccharomyces cerevisiae* strains expressing heterologous and homologous permeases." *Appl Microbiol Biotechnol* 74(5):1041-1052.

[0592]

- Sarkar, G. and S. S. Sommer (1990). "The "megaprimer" method of site-directed mutagenesis." *Biotechniques* 8(4):404-407.
- Sauer, U. (2001). "Evolutionary engineering of industrially important microbial phenotypes." *Adv Biochem Eng Biotechnol* 73:129-169.
- Scarborough, G. A. (1973). "Transport in Neurospora." *Int Rev Cytol* 34:103-122.
- Seiboth B, S. Hakola, *et al.* (1997). "Role of four major cellulases in triggering of cellulase gene expression by cellulose in *Trichoderma reesei*." *J Bacteriol* 179(17):5318-5320.
- Seidl, V., C. Gamauf, *et al.* (2008). "The *Hypocrea jecorina* (*Trichoderma reesei*) hypercellulolytic mutant RUT C30 lacks a 85 kb (29 gene-encoding) region of the wild-type genome." *BMC Genomics* 9:327.
- Shao, Z., H. Zhao, and H. Zhao (2009). "DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways." *Nucleic Acids Res.* 37(2):e16.
- Sikorski, R. S. and P. Hieter (1989). "A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*." *Genetics* 122(1):19-27.
- Skory, C. D., S. N. Freer, and R. J. Bothast (1996). "Expression and secretion of the *Candida wickerhamii* extracellular beta-glucosidase gene, *bglB*, in *Saccharomyces cerevisiae*." *Curr Genet* 30(5):417-422.
- Smith, M. L., O. C. Micali, *et al.* (2000). "Vegetative incompatibility in the *het-6* region of *Neurospora crassa* is mediated by two linked genes." *Genetics* 155(3):1095-1104.
- Stambuk, B. U., M. A. Franden, *et al.* (2003). "D-Xylose transport by *Candida succiphila* and *Kluyveromyces marxianus*." *Appl Biochem Biotechnol* 105-108:255-263.
- Stephanopoulos, G. (2007). "Challenges in Engineering Microbes for Biofuels Production." *Science* 315(5813):801-804.
- Stricker A.R., R. L. Mach and L. H. de Graaff (2008). "Regulation of transcription of cellulases- and hemicellulases-encoding genes in *Aspergillus niger* and *Hypocrea jecorina* (*Trichoderma reesei*)." *Appl Microbiol Biotechnol* 78(2):211-220.
- Sullivan, R. and H. Zhao (2007). "Cloning, characterization, and mutational analysis of a highly active and stable L-arabinitol 4-dehydrogenase from *Neurospora crassa*." *Appl Microbiol Biotechnol* 77(4):845-52.
- Sun Y. and J. Cheng (2002). "Hydrolysis of lignocellulosic materials for ethanol production: a review." *Bioresource Technol* 83(1):1-11.
- Suominen P. L., A. L. Mantyla, *et al.* (1993). "High frequency one-step gene replacement in *Trichoderma reesei*. II. Effects of deletions of individual cellulase genes." *Mol Gen Genet* 241(5-6):523-530.

[0593]

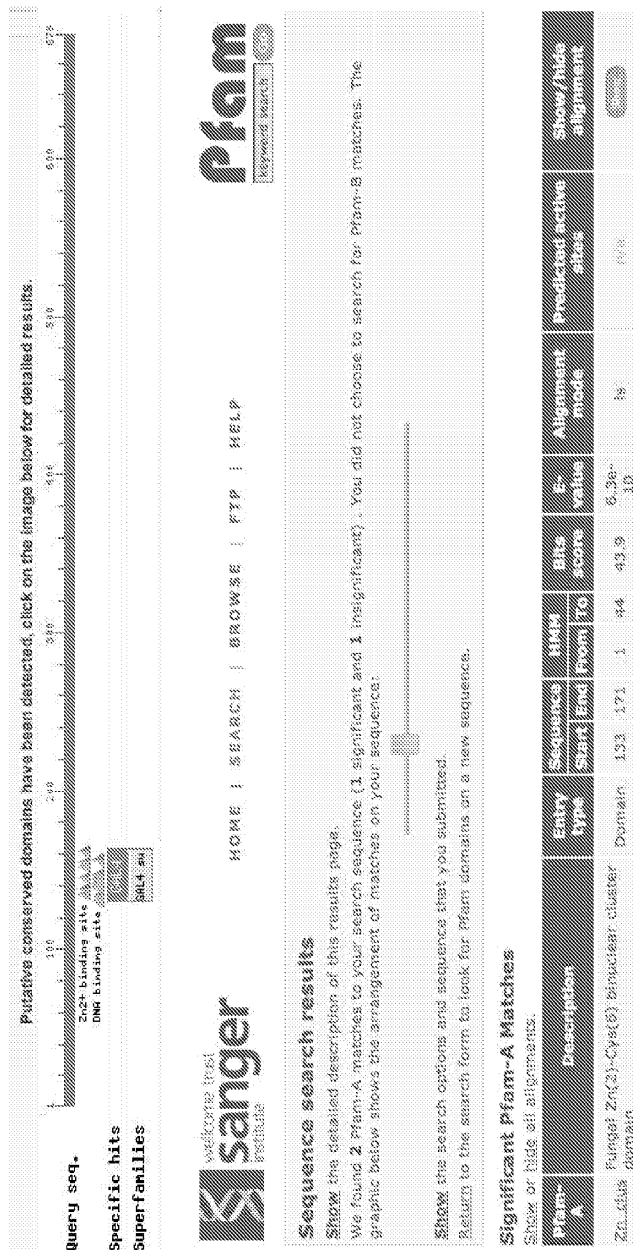
- Tian, C., W. T. Beeson, *et al.* (2009). "Systems analysis of plant cell wall degradation by the model filamentous fungus *Neurospora crassa*." *Proc Natl Acad Sci USA* 106(52):22157-22162.
- Tian, C., T. Kasuga, *et al.* (2007). "Transcriptional profiling of cross pathway control in *Neurospora crassa* and comparative analysis of the *Gcn4* and *CPC1* regulons." *Eukaryot Cell* 6(6): 1018-1029.
- Townsend J. P. (2004). "Resolution of large and small differences in gene expression using models for the Bayesian analysis of gene expression levels and spotted DNA microarrays." *BMC Bioinformatics* 5:54.
- Townsend J. P. and D. L. Hartl (2002). "Bayesian analysis of gene expression levels: statistical quantification of relative mRNA level across multiple strains or treatments." *Genome Biol* 3(12):research0071.1-0071.16.
- USDA, The Economic Feasibility of Ethanol Production from Sugar in the United States, United States Department of Agriculture (2006).
- van Rooyen, R., B. Hahn-Hägerdal, *et al.* (2005). "Construction of cellobiose-growing and fermenting *Saccharomyces cerevisiae* strains." *J Biotechnol* 120(3):284-295.
- Vogel, H. J. (1956). "A convenient growth medium for *Neurospora*." *Microbiol. Genet. Bull.* 13:42-46.
- Wagner, S., M. L. Bader, *et al.* (2006). "Rationalizing membrane protein overexpression." *Trends Biotechnol* 24(8):364-371.
- Watanabe, S., A. Abu Saleh, *et al.* (2007). "Ethanol production from xylose by recombinant *Saccharomyces cerevisiae* expressing protein-engineered NADH-preferring xylose reductase from *Pichia stipitis*." *Microbiology* 153(Pt 9):3044-3054.
- Watanabe, S., T. Kodaki, and K. Makino (2005a). "L-Arabinose 1-dehydrogenase: a novel enzyme involving in bacterial L-arabinose metabolism." *Nucleic Acids Symp Ser (Oxf)* 49:309-310.
- Watanabe, S., T. Kodaki, and K. Makino (2005b). "Complete reversal of coenzyme specificity of xylitol dehydrogenase and increase of thermostability by the introduction of structural zinc." *J Biol Chem* 280(11):10340-10349.
- Wiedemann, B. and E. Boles (2008). "Codon-optimized bacterial genes improve L-arabinose fermentation in recombinant *Saccharomyces cerevisiae*." *Appl Environ Microbiol* 74(7):2043-2050.
- Wisselink, H. W., M. J. Toirkens, *et al.* (2007). "Engineering of *Saccharomyces cerevisiae* for efficient anaerobic alcoholic fermentation of L-arabinose." *Appl Environ Microbiol* 73(15):4881-4891.
- Wisselink, H. W., M. J. Toirkens, *et al.* (2009). "Novel Evolutionary Engineering Approach for Accelerated Utilization of Glucose, Xylose, and Arabinose Mixtures by Engineered *Saccharomyces cerevisiae* Strains." *Appl Environ Microbiol* 75(4):907-914.
- Woodyer, R., M. Simurdiak, *et al.* (2005). "Heterologous expression, purification, and characterization of a highly active xylose reductase from *Neurospora crassa*." *Appl Environ Microbiol* 71(3): 1642-7.
- Wymelenberg, A.V., *et al.* (2009). "Transcriptome and Secretome Analyses of *Phanerochaete chrysosporium* Reveal Complex Patterns of Gene Expression." *Appl Environ Microbiol* 75(12):4058-4068.
- Xin, Z, Q. Yinbo, and G. Peiji. (1993). "Acceleration of ethanol production from paper mill waste fiber by supplementation with β -glucosidase." *Enzyme Microb Tech* 15(1):62.

[0594]

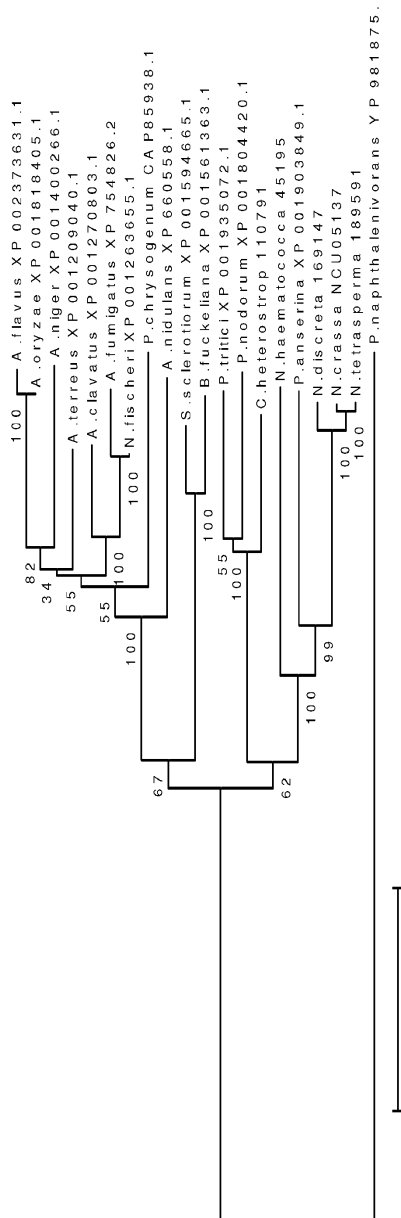
[0595]

도면

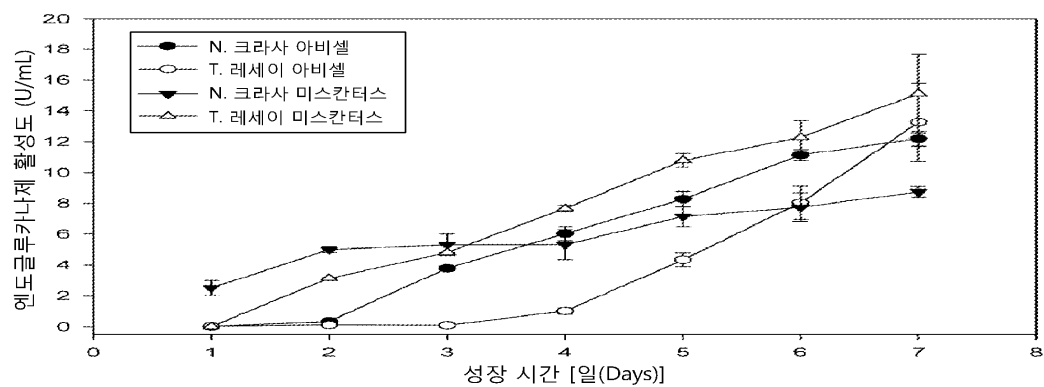
도면1



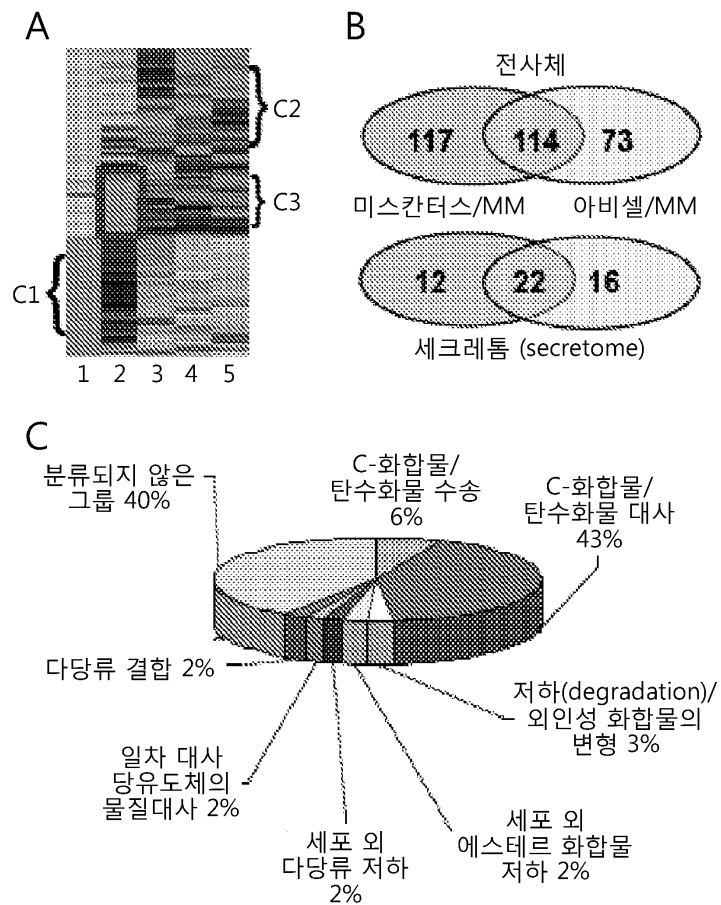
도면2



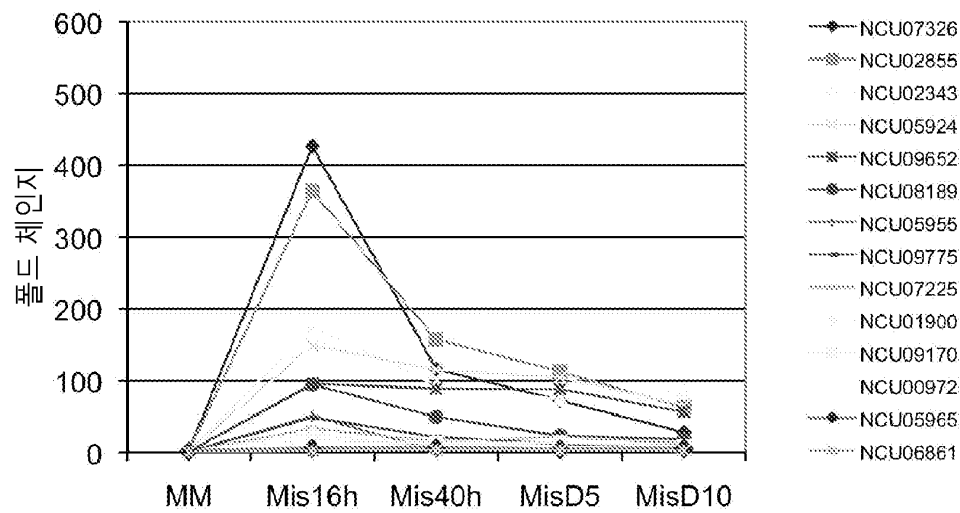
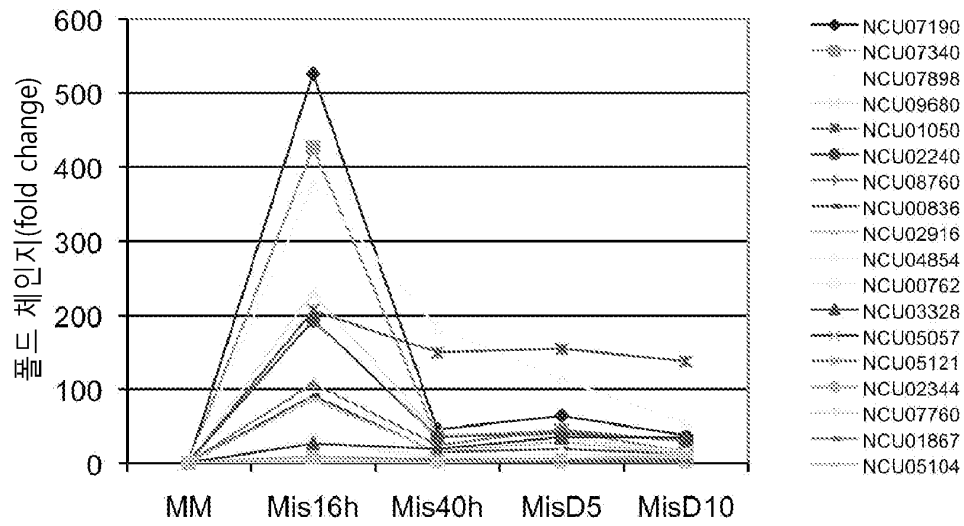
도면3



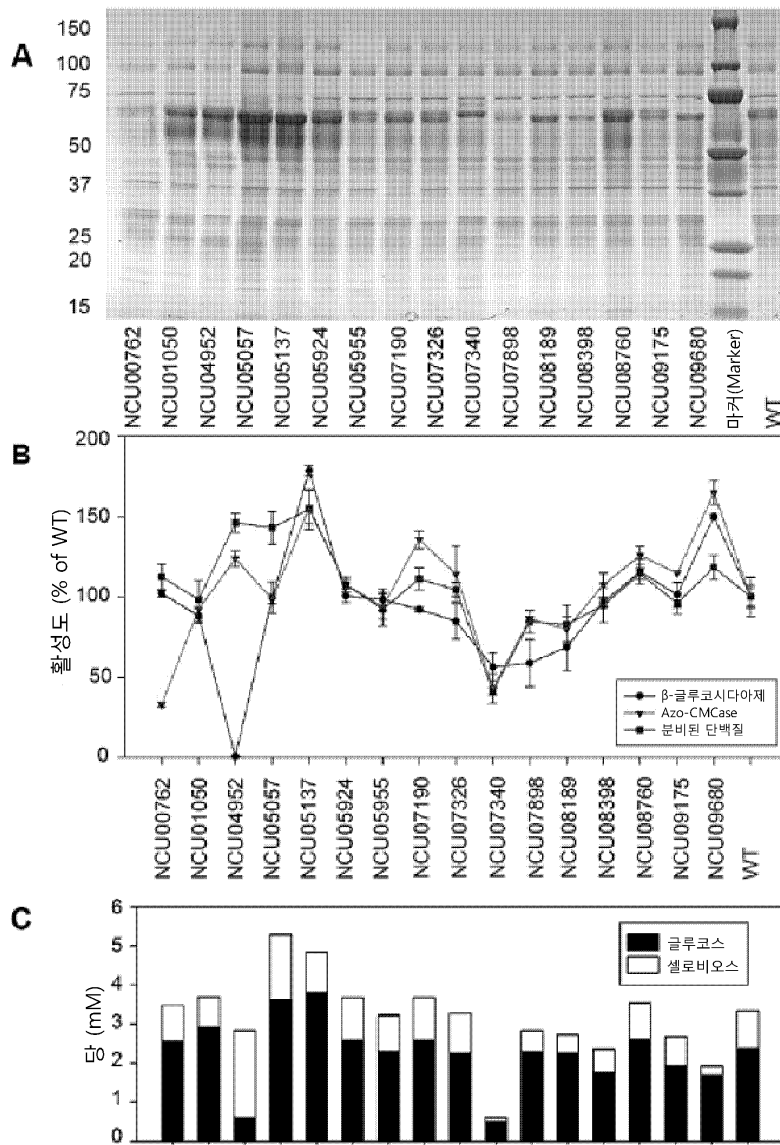
도면4



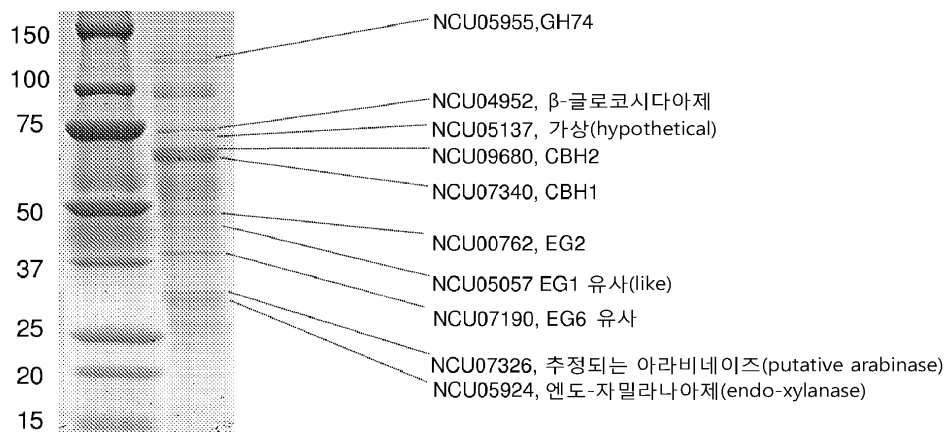
도면5



도면6

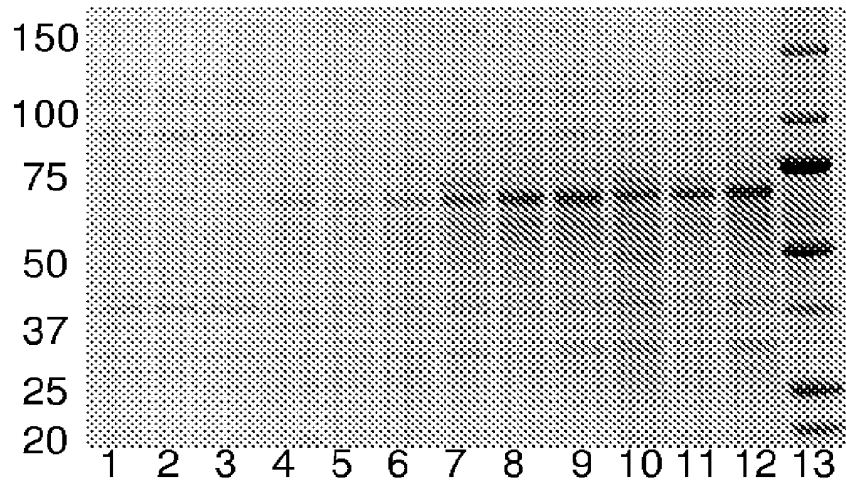


도면7

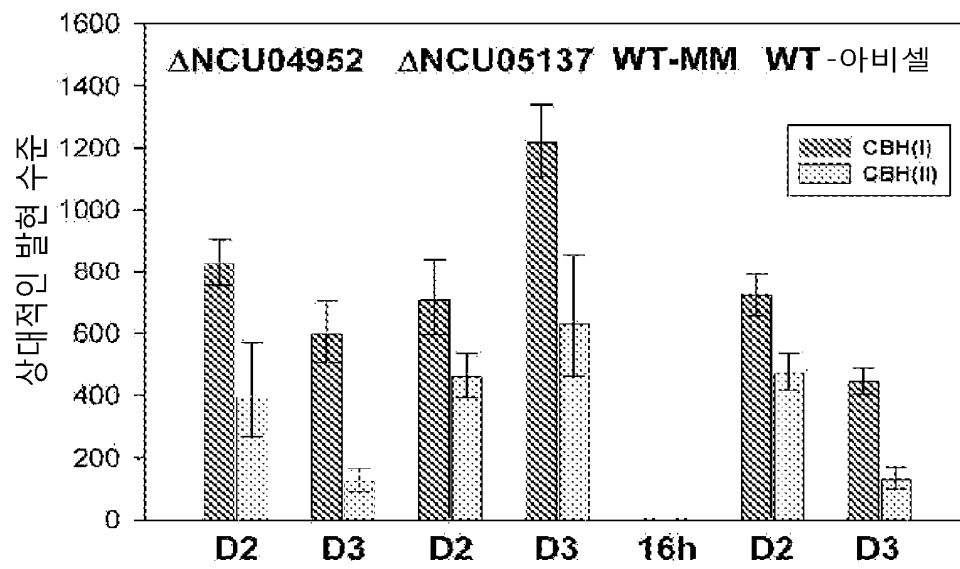


도면8

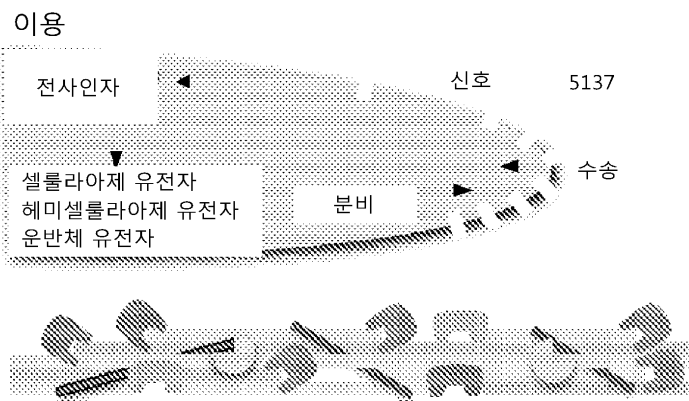
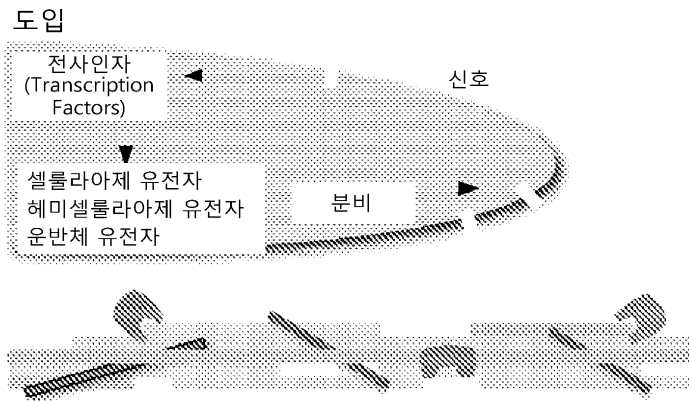
A



B



도면9



도면10a

Query= NCU00801 | neurospora_crassa hypothetical protein similar to MFS lactose permease (580 nt)
(579 letters)

Database: Spoth1_GeneModels_FilteredModels1_aa.fasta
8806 sequences; 4,205,009 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
jgi Spath1 43941 e_gw1.2.4209.1	361	e-100

>jgi|Spath1|43941|e_gw1.2.4209.1
Length = 537

Score = 361 bits (926), Expect = e-100, Method: Compositional matrix adjust.
Identities = 199/534 (37%), Positives = 284/534 (53%), Gaps = 29/534 (5%)

Query: 33 KGHGQTATKPGAQKEVRNAALFAAIKE----SNIKPWSKESIHLYFAIFVAFCCACANG 88
+ +T A + +R A L + + PWS I LY + +++ C C NG
Sbjct: 5 RDEKETVVADHADDALREADLAVQVAHDADGTVYSPWSLRMRILYLVLVLSYLCGCLNG 64

Query: 89 YDGSMLTGIIAMDKFQNGFTGDTGPKVSVIFSLYTVGAMVGAPFAAILSDRFGRKKGMF 148
YDGSML G+ M +Q FH G +IF++Y +G++ F ++D FGR+ GMF
Sbjct: 65 YDGSMLGGGLNGMTSYQRYFHMSTAGSTTGLIFAMYNIGSVAAVFFTGPNVDYFGRRWGMF 124

Query: 149 IGGIFIIVGSIIIVASSSKLAQFVVGRFVLGLGIAIMTVAAPAYSIEIAPPHWRGRCTGFY 208
+G + +IVG+ + A + QF+ GRFVLG G++ V+AP Y E+A P WRG TG Y
Sbjct: 125 VGALLVIVGTCVQAPCTTRGQFLAGRFLVGFVGFCCVSAPCYVSEMAHPKWRGTLTGLY 184

Query: 209 NCGWFGGSIPAACITYGCFIKSNWSWRIPILILQAFTCLIVMSSVFFLPESPRFLFANGR 268
NC W+ SSI A+ + YGC +I + +WRIP+ Q T +V VF+LPESPR+L A R
Sbjct: 185 NCTWYIGSIVASWVVGCSYIDTLDAWRIPWCQMVTSGLVCLGVFWLPESPRWLMQAQDR 244

Query: 269 DAEAVAFVLVKYHNGNDPNSKLVLEETEEMRDGIRIDGVDKVWWDYRPLFMTHSGRWMAQ 328
+A L YHG G + LV L+ +EM + I T+ DK W-DY L+ THS R R+
Sbjct: 245 HDDAARVLATYHGEGRADHPLVKLQMQEMMNQISTEASDKKWDYHELVNTHSARRRLC 304

Query: 329 VLMISIFGQFSGNGL-GYFNTVIFKNIGVTSTSQQLAYNINLSVISAIGALTAVSMTDRM 387
V+ +++FGQ SGN L Y+ + K+ G+T + LA N +N +S +GA+ MTD +
Sbjct: 305 VIGMAVFGQISGNSISSYLVNMLKSAGITEERRVLALNGVNPALSFLGAILGARMTDVV 364

Query: 388 PRRAVLIIGTFMCAAAALATNSGLSATLDKQTQRTQINLNQGMNEQDAKDNAYLHVDSNY 447
RR +L+ + A +G S R N
Sbjct: 365 GRRPLLLYTIVFASVCFVITGTSMATDDPTRTAAAN----- 402

Query: 448 AKGALAAAYFLNVVIFSFTYIPLQGVIPTEALETIRGKGLALSGFIVNAMGFINQFAGPI 507
+A F+F ++FSF +IPLQ + E L T R KG A+ F + I Q+A
Sbjct: 403 --ATIAFIFIFIGVFSFGWIPLQSMYIAETLPTATRAKGTAVGNFSSSVASTILQYASGP 460

Query: 508 ALHNIGYKIFYFVVGWDLIETVAWYFFGVESQGRILEQLEWVYDQPNPVKASLK 561
A IGY + VFV WDLIE YF+ E++ RILE+LE V- PNPVK SL+
Sbjct: 461 AFEIGIGYYFYLVFVFWDLIEGAIMYFYFPETKDRILEELEVFSAPNPVKKSLE 514

도면10b

```

Query= NCU04963 | neurospora_crassa hypothetical protein similar to
MFS monosaccharide transporter (528 nt)
(527 letters)

Database: Spoth1_GeneModels_FilteredModels1_aa.fasta
8806 sequences; 4,205,009 total letters

Searching.....done

Sequences producing significant alignments:

Score      E
(bits) Value

jgi|Spath1|62521|estExt_Genewise1.C_21757          541    e-155

>jgi|Spath1|62521|estExt_Genewise1.C_21757
Length = 566

Score = 541 bits (1395), Expect = e-155, Method: Compositional matrix adjust.
Identities = 263/508 (51%), Positives = 366/508 (72%), Gaps = 10/508 (1%)

Query: 6 KKPEGVPGKSWPAIVIGLFVAFGGVLFYDGTGTIGGILAMPYWDLFSTGYRNPHEHLDV 65
+KP+ V G S PAI++GLFVA GG+LFGYDTG I GILAM -++ F+TGY + + +
Sbjct: 9 QKPDNVAGSSAPAIMVGLFVATGGLLFGYDTGAINGILAMDTFKEDFTTGYTDKQKPG 68

Query: 66 TASQSATIVSILSAGTFFGALGAAPLADWAGRRLLGLLSSFFVIFGVILQTAASVIPLE 125
AS+ + IV+-LSAGT GAL +AP+ D GRRL LI++ VF G I+Q A ++ + +
Sbjct: 69 YASEVSLIVAMLSAGTATGALLSAPMGDRWGRRLSLIVAIGVFCVGATIQVCATNVAMLV 128

Query: 126 AGRFFAGLVGLISATIPLYQSETAPKWIRGVIVGSYQLAITIGLLASIVNNATHNMQN 185
GR AG+GVG++S +PLYQSE APKWIRG +V +YQL+IT GLL A+ VN T+ +-+
Sbjct: 129 VGRTLAGIGVGVSVLVPLVLYQSEMAPKWIRGILVCAYQLSITAGLLAATVNILTYKLKS 188

Query: 186 TGCYRIPIAVQFAWAIILIVGMILPETPRFHIKRDNLPAATRSALILRRLEQNHPAIE 245
YRIPI +Q WA++L +G++ILPETPR+ +KR AA SL+ LRRL+ HPA+IE
Sbjct: 189 AAAYRIPIGLQLTWAIVLALGLVILPETPRYLVRKGLKEAALSLSLRLRLDITHPALIE 248

Query: 246 ELSEIQANHEFEKSLGKATYLDCLKG--NLLKRLLTGCFLQSLQQLTGINFIFYGTQFF 303
EL+EI+ANHE-E +IG TY D + G +L +R LTGC LQ LQQLTG+NFI YYGT FF
Sbjct: 249 ELAIEANHEYEMALGPDYKDIIFGEPHLGRRTLTGCGLQMLQQLTGVNFIMYYGTTF 308

Query: 304 KNSGFSDFSFLISLITNLVNVSTLPLGLYAIKDWGRRPVLLWGAVGMCVCQFIVAILGTT 363
+G ++F -SLI ++N+VST PGL+ ++ WGRR +L+ G-VGM +CQ ++A T +
Sbjct: 309 YGAGIGNAFTVSLIMQVINLVSTFPGLFVVESWGRKLLIVGSVGMAICQLLIASFATAS 368

Query: 364 TSQDASGMIIIVHNLAQAIAFICFYIFFFAASWGPVAVVVGGEIFPLKVRAKSLSITT 423
+ + + I F+ YIFFFAASWGPV WVVT EI+PLKVRAKS+SI-T
Sbjct: 369 GNDNKP-----TQNQILIIIFVATYIFFFAASWGPVWVVGSEIYPLKVRAKSMSIST 420

Query: 424 ASNWLLNWAIAYSTPYLVNYGPGNANLQSKIFFVWGGCCFICIAFVYFMIYETKGLTLEQ 483
ASNW+LN+ IAY TPYLV+ G+ +L S++EFFVWG C + IAFV++M+YET ++LEQ
Sbjct: 421 ASNWVLNFGIAYGTPYLVDTSDGSPDLGSRVFFVWGAFCLISIAFVWYMYETSKISLEQ 480

Query: 484 VDELYEEVSDARKSIGWVPTITFREIRE 511
+DE+YE V+ A S + P+ +F+++R+
Sbjct: 481 IDEMYERVAHAWNSRSFEPSSWSFQQMRD 508

```

도면10c

```

Query= jgi|Spath1|48439|e_gwl.3.3367.1
      (512 letters)

Database: neurospora_crassa_7_proteins_no_asterisks.fas
      9822 sequences; 4,775,003 total letters

Searching.....done

Sequences producing significant alignments:

                                Score      E
                                (bits) Value
NCU01132 | neurospora_crassa hypothetical protein similar to mo...   727    0.0

>NCU01132 | neurospora_crassa hypothetical protein similar to
      monosaccharide transporter (554 nt)
      Length = 553

Score = 727 bits (1876), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 382/548 (69%), Positives = 431/548 (78%), Gaps = 42/548 (7%)

Query: 1 MKKFLGLRGQALNLAVGTIAGCDFLLFGYDQGVGMGGILTLKVFDAFPMPINPEAAGLSHD 60
      MK FLGLRGQ LNLAVG +AGCDFLLFGYDQGVGMGGILTL FL F INP+A GL+
Sbjct: 1 MKPFLGLRGQPLNLAVGAVAGCDFLLFGYDQGVGMGGILTLPEFLGYFEQINPDAPGLTPH 60

Query: 61 ESSRSTYQGAVASYNLGCFLGAIITIFIGNPLGRKRVMILGTSVMVIGAILQASSTTL 120
      ESS RSTYQGAVASYNLGCFLGAIITIFIGNP GRK++I+LGTS+M++GAILQAS+TTL
Sbjct: 61 ESSMRSTYQGAVASYNLGCFLGAIITIFIGNPWGRKKIILLGTSIMIVGAILQASATTL 120

Query: 121 PQFIVGRIITGLNGGNTSTVPTWQSETSKAHKRGKMIF-----CAIILAF----- 167
      FI+GRIITG+NGGNTSTVPTWQSETSKAHKRGKM+ I+L++
Sbjct: 121 GHFIIGRIITGIGNGGNTSTVPTWQSETSKAHKRGKMVMIEGSLVTAGIMLSYWIDLGLS 180

Query: 168 -----IPFLPESPRWLILKGREDEAREVIAALEDTSDR 202
      IPFLPESPRWL+ KGR+ EA+E++AAL D + D
Sbjct: 181 FAPGSVAWRFLAFQIIFCILILIFIPFLPESPRWLVFKGRDAEAKELAALNDVELDDP 240

Query: 203 IVENEFLAIKETVLEMSKGTFRDLFTMDKNRNLHRTLLAYFNQVFQQISGINLITYYAAV 262
      IV+ EF I -TV+EMSKG+F+DLFTMDK+RN HRTLLAY NQVFQQISGINLITYYAAV
Sbjct: 241 IVDTEFHFIHDTVVEMSKGSPKDLFTMDKDRNFHRTLLAYLNQVFQQISGINLITYYAAV 300

Query: 263 IYKGLGMSDFLSRLAALNGTEYFLASWPAVFLVERVGRRLMLFGAVGQAATMAILAGV 322
      IY GLGMSDFL+RLAALNGTEYF+ASWPAVFLVERVGRRLMLFGA+GQAATMAILAGV
Sbjct: 301 IYSGLGMSDFLARLLAALNGTEYF+ASWPAVFLVERVGRRLMLFGAIGQAATMAILAGV 360

Query: 323 NSR-QETGFQAGIVFLFVFNTFFAVGWLGMTWLYPAEIVPLRIRAPANALSTSANWIFN 381
      NSR + +QAGIVFLFVFNT FAVGWLGM+WLYPAEIVPLRIRAPANALSTSANWIFN
Sbjct: 361 NSRPDDKPYQAGIVFLFVFNTVFAVGWLGMWLYPAEIVPLRIRAPANALSTSANWIFN 420

Query: 382 FLVVMITPVAFNNIGYQTYIIFAVINAFMVPCVYFFYPETAYRSLEEMDNIFHKVADGWK 441
      F+VVMITPVAFN I YQTYIIFAVINAF+VP VYFFYPETA RSLEEMD IFHKV DGWK
Sbjct: 421 FMVVMITPVAFNFIKYQTYIIFAVINAFIVPVVYFFYPETACRSLEEMDMIFHKV-DGWK 479

Query: 442 GVFTVVHQAKVEPRWYKNGELL--VDYQQTEHRRHLQQQEGAVTASEKRSVEGAGSGS 499
      G FTVVHQAKVEP+WY K+G+ + D+++T ++ H + K VE +
Sbjct: 480 GYFTVVHQAKVEPKWYDKDQQRIGGADFEKTAGYQSHSIPSSGFEKPTKAHVESPRADD 539

```


도면10d

```

Query: 500 GSGDVKQD 507
      G      D
Sbjct: 540 GITSSSSD 547

Query= jgi|Spath1|112305|estExt_fgenes1_kg.C_60263
      (554 letters)

Database: neurospora_crassa_7_proteins_no_asterisks.fas
      9822 sequences; 4,775,003 total letters

Searching.....done

Sequences producing significant alignments:

Score      E
(bits) Value

NCU10021 | neurospora_crassa high affinity glucose transporter ... 937 0.0
>NCU10021 | neurospora_crassa high affinity glucose transporter
      (554 nt)
      Length = 553

Score = 937 bits (2421), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 474/542 (87%), Positives = 507/542 (93%), Gaps = 7/542 (1%)

Query: 1 MSSSEKEATGPVAAHVGNLATTQDVEKIEAPVTWKAYLICAFASFGGIFFGYDSGYINGV 60
      MSS+ ++ T P AAH+G LA TQDVE+IEAPVTWKAYLICAFASFGGIFFGYDSGYINGV
Sbjct: 1 MSSAHEKETAPTAAHIG-LAHTQDVERIEAPVTWKAYLICAFASFGGIFFGYDSGYINGV 59

Query: 61 LASKLFINAVEGAG--KDAISESHSSLIVSILSCGTFFGALIAAGDLADFIGRKYTVILGC 118
      L S++FI+AVEG +DA+SESH SL+VSILSCGTFFGALIAAGDLAD IGRK+TVILGC
Sbjct: 60 LGSQIFIDAVEGTSPVRDALSESHQSLVVSILSCGTFFGALIAAGDLADMIGRKTIVILGC 119

Query: 119 LIYIIGCVIQ-ITGLGNALGAIVAGRLLIAGIGVGFESAIVILYMSEICPRKVRGALVAGY 178
      LIY+IGCVIQ-ITGLG+ALGAIVAGRLLIAGIGVGFESA+VILYMSEICPRKVRGALVAGY
Sbjct: 120 LIYLIGCVIQMITGLGDALGAIVAGRLLIAGIGVGFESAVVILYMSEICPRKVRGALVAGY 179

Query: 179 QFCITIGLMLASCVVYGTQNRQDTGQYRIPIGIQFIWALILGGGLLCLPDSPRYFVKRGR 238
      QFCITIGLMLASCVVY TQ+R DTG YRIPI IQFIWALIL GGL+CLPDSPRYFVK+G
Sbjct: 180 QFCITIGLMLASCVVYATQDRPDTGAYRIPIAIQFIWALILAGGLMCLPDSPRYFVKKN 239

Query: 239 LADATSALSRLRGQPEDSEYIQVELAEIVANEYERQLIPSTTWFGSWANCFKGSFLKAN 298
      LA ATS+LSRLRGQ +SEYIQVELAEIVANEYERQLIPSTTWFGSWANCFKGS+KAN
Sbjct: 240 LAAATSSLSRLRGQDPNSEYIQVELAEIVANEYERQLIPSTTWFGSWANCFKGSFLKAN 299

Query: 299 SNLRKTIILGTSLQMMQQTGVNFIFYYSTPFLKSTGAIDDPFLMSMVFTIINVSTPISF 358
      SNLRKTIILGTSLQMMQQTGVNFIFYYSTPFLKSTGAI + FL+SMVFTIINVSTPISF
Sbjct: 300 SNLRKTIILGTSLQMMQQTGVNFIFYYSTPFLKSTGAISNTFLISMVFTIINVSTPISF 359

Query: 359 YTVRFGRRTILFWGALGMLICQFLVAIVGVTVGFNHTHPAP*ADDPEATLANNISAVNA 418
      +TVRFGRRTILFWGALGMLICQFLVAI+GVTVGFN TH P + ++ANN+SAVNA
Sbjct: 360 WTVRFGRRTILFWGALGMLICQFLVAIIGVTVGFNTHMGPDGE----SMANNVSAVNA 415

Query: 419 QIAFIAIFIFFFASTWGPAGAWIVIGEIFPLPIRSRGVGLSTASNWLWNTIIAVITPYMVG 478
      QIAFIAIFIFFFASTWGPAGAWI+IGEIFPLPIRSRGVGLSTASNWLWNTIIAVITPYMVG
Sbjct: 416 QIAFIAIFIFFFASTWGPAGAWILIGEIFPLPIRSRGVGLSTASNWLWNTIIAVITPYMVG 475

```

도면10e

```

Query: 479 EDRGNMKSSVFFVWGGCTCAFYVYTYFLVPETKGLSLEQVDKMMEEETTPRTSAKWKPTTT 538
          E RGN+KSSVFFVWGGCTCAFY+YTYFLVPETKGLSLEQVDKMMEEETTPRTSAKWKP TT
Sbjct: 476 EQRGNLKSSVFFVWGGCTCAFIYTYFLVPETKGLSLEQVDKMMEEETTPRTSAKWKPRTT 535

Query: 539 FA 540
          FA
Sbjct: 536 FA 537

Query= jgi|Spath1|114107|estExt_fgenes1_pm.C_20669
      (522 letters)

Database: neurospora_crassa_7_proteins_no_asterisks.fas
        9822 sequences; 4,775,003 total letters

Searching.....done

Sequences producing significant alignments:

Score      E
(bits)     Value

NCU08114 | neurospora_crassa hypothetical protein similar to MF... 832 0.0
>NCU08114 | neurospora_crassa hypothetical protein similar to MFS
          hexose transporter (526 nt)
          Length = 525

Score = 832 bits (2148), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 397/523 (75%), Positives = 452/523 (86%), Gaps = 6/523 (1%)

Query: 1 MGIFAFNKQKPNAEATAVAQ--EEAPQFERVDWKRDPLRLKLYFYAFVLCIASATTGYDG 58
          MGIF +KP A+A + Q EEAPQFERVDWK+DPGLRLKLYFYAF+LCIASATTGYDG
Sbjct: 1 MGIF---NKKPVAQAVDLNLIQEEAPQFERVDWKDPGLRLKLYFYAFILCIASATTGYDG 57

Query: 59 MFFNSVQNFEWENYFNHPTGSKLGVLGALYQIGSLASIPLVPIIADRVGRKPIAIGCV 118
          MFFNSVQNFEW YF P GS+LG+LGALYQIGS+ SIP VP++ D GRK PI IGCV
Sbjct: 58 MFFNSVQNFEWIKYFGDPRGSELGLLGALYQIGSIGSIPFVPLTDFGRKTPIIIGCV 117

Query: 119 IMIVGAVLQAACRNIGTFMGGFRLLGFGNSLAQLCSPLLTELHPQHRRLTTVYNCLW 178
          IMIVGAVLQA +NL TFMGGR +LGFGNSLAQ+ SPMLLTELHPQHR RLTT+YNCLW
Sbjct: 118 IMIVGAVLQATAKNLDTFMGGRTMLGFGNSLAQIASPMLLTELHPQHRRLTTIYNCLW 177

Query: 179 NVGALVVAWVSFGTDYLSKSDWSWRIPALIQAFPSVIQLLFIFWVPESPRYLMAKDKHERA 238
          NVGALVV+W+--FGT+Y+ +DWSWRIPAL+QAFPS+IQLL I+WVPESPR+L+AKDKH+ A
Sbjct: 178 NVGALVVSWLAFGTNYINNDWSWRIPALLQAFPSIIQLLGIWVWVPESPRFLIAKDKHDEA 237

Query: 239 LAILAKYHANGDANHPTVQFEYREIKETLRLEFEASKSSSYLDFVTRTGNRYRLAVLISL 298
          L ILAKYHANGD NHPTVQFE+REIKET+RLE E++K+SSYLDF ++RGNRYRLA+L+SL
Sbjct: 238 LAILAKYHANGDPNHPTVQFEFREIKETIRLEMESTKNSSYLDFFKSRGNRYRLAILLSL 297

Query: 299 GIFSQWSGNAIISNYSSKLYDTAGVTGSTQKLGLSAGQTGLSLIISVTMALLVDKFGRRP 358
          G FSQWSGNAIISNYSSKLY+TAGVT ST KLGLSAGQTGL+LI+SVTMALLVDK GRR
Sbjct: 298 GFFSQWSGNAIISNYSSKLYETAGVTDSTAKLGLSAGQTGLALIVSVTMALLVDKLGRR 357

Query: 359 MFLTSTAGMFCTFIWTLTSGLYEEHNADGARYAMILFIWIHGIFYSISWSGLLVGYAIE 418
          FL ST GM TF+ WTLT+GLY EH GA AMI FIW+ GIFYS++WSGLLVGYAIE
Sbjct: 358 AFLASTGGMCGTFVIWTLTAGLYGEHRLKGADKAMIFFIWIWFGIFYSLAWSGLLVGYAIE 417

```

도면10f

```

Query: 419 VLPYKLRAGLMIMNLTIVQAAITLNTYANPVAFDAF-EGHSWKLYIITYTIWIFLELCFVW 477
      +LPY+LR KGLM+MN++VQ ALTLNTYANPVAFD F HSWKLY+IYT WI E FV+
Sbjct: 418 ILPYRLRGKGLVMNMMSVQCALTNTYANPVAFDYFGPDHSWKLYLIYTCWIAAEFVVFV 477

Query: 478 KMYIETKGPTEELAKIIDGDEAAVAHVDIKQVEKETHINEEK 520
      MY+ETKGPTEELAK+IDGDEA VAH+DI QVEKE I+E -
Sbjct: 478 FMYVETKGPTEELAKVIDGDEADVAHIDIHQVEKEVEIHEHE 520

Query= jgi|Spath1|70023|estExt_Genewise1.C_53218
      (504 letters)

Database: neurospora_crassa_7_proteins_no_asterisks.fas
      9822 sequences; 4,775,003 total letters

Searching.....done

Sequences producing significant alignments:
Score      E
(bits) Value
NCU06138 | neurospora_crassa hypothetical protein similar to MF... 263 6e-71
>NCU06138 | neurospora_crassa hypothetical protein similar to MFS
      monosaccharide transporter (584 nt)
      Length = 583

Score = 263 bits (673), Expect = 6e-71, Method: Compositional matrix adjust.
Identities = 159/488 (32%), Positives = 259/488 (53%), Gaps = 35/488 (7%)

Query: 10 FLVGVFASLGSLLYGYDLGVIAQVIASQSFKSRFSPSD--NEEAA-----VVSVFTGGA 61
      F + FA +G +LYGY+ G+ + V+A +F+ D +E A+ + ++ GA
Sbjct: 39 FSIACFACIGGVLYGYNQGMFSGVLAMPFAFQKHMGEYDPIDENASQTKKGWLTAILLELGA 98

Query: 62 FFGAMAAGPMGDKLGRRWITILCGALVFCLGGALQTGAQALSY--LYAGRSIAGLGVGVLC 119
      + G + +G M + L R++ +L LVF LG +Q + + + + AGR I G+GVG L
Sbjct: 99 WLGTLTSGFMAEVLRSKYGVLVACLVFMLGVVIQATSISGGHETILAGRFITGMGVGSLA 158

Query: 120 MIVPMYQAEALAHPSIRGRITAIQQFMLGIGALAAAWISYGT-YVG---FAPTNDGQWRTS 175
      MI+P+Y +E+A P +RG + ALQQ + G + + WI YGT Y+G +D W
Sbjct: 159 MIPIYNSEVAPPEVRGALVALQQLAICFGIMVSFWIDYGTNYIGGTKLETQSDAAWLVP 218

Query: 176 LGIQVIPAVFLAALILLFPESPRWLIDHGRSEEGRLTLAQLHSHGDVDDAWQAEYQQIR 235
      + +Q+ PA+ L ++ P SPRWLI HGR E + L+ L D V+ E+ +I+
Sbjct: 219 VCLQLAPALILFFGMMFMFSPRWLIHHGREAEARKILSTLRGLPQ-DHELVELEFLEIK 277

Query: 236 ESVEFVRENEAKSYAEL-----FRDRSCFRRFLACAIQGSVQMTG 276
      F + + A+ + EL FR ++ FRR+ +A Q -G
Sbjct: 278 AQSLEKRSIAELFPELREQTAWNTFKLQFVAIEKLFRTKAMFRRVIVATVIMFFQQWSG 337

Query: 277 VSAIQYYSVTIYGLMGIEGDDT-LKYQAISSIIALVAQALCILLIDRLGRRWTLIGGNLG 335
      ++AI YY+ I+ +G+ G+ T L + I+ +A +L IDR+GR+ L G LG
Sbjct: 338 INAILYAPQIFKQLGLSGNTISLLATGVVGIVMFIATVPAVLWIDRVGRKPVLIIGALG 397

Query: 336 NCVIFIIATVMLARYPPGTSNKAAGWGFIVVTWVYNFSFSAICGPLSWIIPAEIFDITKT 395
      II V-+A+ ++KAA W + + W++ F - GP +WII AEI+ T
Sbjct: 398 MATCHIIIVAVVAKNVDQWETHKAAGWAAVAMVWLFVIHFGYSWGPCAWIIVAEIWLPLST 457

```

도면10g

```

Query: 396 RSKGVSIATMTSFAINTMIGQVTGPAMKTVGYRYLLFVVCNFTNALFFWAFIPETAKRP 455
          R GVS+   -++ N ++GQVT   +K + Y Y++F +   - A F W F+PET +
Sbjct: 458 RPYGVSLGASSNWMNFIIVGQVTPDMLKAIPYGTIIFGLLTYMGAAFIWFVPETKRIT 517

Query: 456 LEEMNRLF 463
          LEEM+ +F
Sbjct: 518 LEEMDMIF 525

Query= jgi|Spath1|102977|fgenes1_pm.5_#_763
      (481 letters)

Database: neurospora_crassa_7_proteins_no_asterisks.fas
          9822 sequences; 4,775,003 total letters

Searching.....done

Sequences producing significant alignments:

                                Score   E
                                (bits) Value
NCU05897 | neurospora_crassa hypothetical protein similar to l-...   713   0.0
>NCU05897 | neurospora_crassa hypothetical protein similar to
          l-fucose permease (472 nt)
          Length = 471

Score = 713 bits (1841), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 347/447 (77%), Positives = 385/447 (86%), Gaps = 2/447 (0%)

Query: 1 MLSSGFWKRRSLRVDPNQRTKAAELTLRESLYPLSLVTILFFLWGFSGLLDTLNKHFQ 60
          M S +WKRRSLRV D++ TKAAEL+LRESI PL LVITLFFLWGFSGLLDTLNKHFQ
Sbjct: 1 MFSREWWKRRSLRVDDKVIKAAELSLRESILPLCLVTILFFLWGFSGLLDTLNKHFQ 60

Query: 61 TLGITKTRSSGLQAAYFGAYPLASLGHAAWILRHGYRAVFIWGLFLYGLGALLAIPSIM 120
          TL ITK RS+GLQAAYFGAYPLASLGHA+ILR + YRAVFIWGLFLYGLGALLAIP
Sbjct: 61 TLHITKARSAGLQAAYFGAYPLASLGHAAYILRRFSYRAVFIWGLFLYGLGALLAIPCK 120

Query: 121 HHSFAGFCVCIFIIGNGLGSLETAANPYITVCGPPKFSEIRINVAQAFNGIGTVVAPVLG 180
          SFAGFCVCIFIIGNGLGSLETAANPYITVCGPPK+SEIRIN AQAFNGIGTVVAPVLG
Sbjct: 121 AKSFAGFCVCIFIIGNGLGSLETAANPYITVCGPPKYSEIRINFAQAFNGIGTVVAPVLG 180

Query: 181 SYVFFTFDDQALRNQWVYLAIACFVFLLAGVFFLSVIPEITDADMAFQAAETHAGADD 240
          SYVFF FDD AL+NVQWVYLAIA FV++LA VFFL +PEITDADM QAAETHAG D
Sbjct: 181 SYVFFGDDNLALQNQWVYLAIAVAVVYILAVVFFLIELPEITDADMQHQAETHAGDAD 240

Query: 241 RPFHTQYRLFHAFAQFCYTGQAQVAIAGYFINYATETRPNTDSSLGSKFLAGSQAGFAVG 300
          +PF QYRLFHA+F+QFCYTGQA+AIAGYFINY TETR NTDS+LG++FLAG+Q FAVG
Sbjct: 241 QPFRKQYRLFHASFSQFCYTGQAIAIAGYFINYVTETRNKNTDSALGAQFLAGAQGTFAVG 300

Query: 301 RFGGAAMMQFIKPRKVFALFMTMCIVFSAPAITQRGNAGLSMLYLVMMFFESICFPTIIAL 360
          RF GAA+M F++PRKVF LF+T CI+F AP ITQR N G+S+LY+ +FFESICFPTI+AL
Sbjct: 301 RFAGAAIMHFVRPRKVFLLFLTACIIFVAPTITQRENTGMSLLYVTLFFESICFPTIMAL 360

Query: 361 GMRGLGRHTKRGSGWIVAGVLGGACVPPLMGAAADARGTGFSMLVPLCFFVAAWTYALAV 420
          GMRGLGRHTKRGSG++VAGV GGA VPPLMGA AD T SM+VPL FF AWTYALAV
Sbjct: 361 GMRGLGRHTKRGSGFLVAGVFGGAVVPPLMGAVADMMDTAMSMVVPLAFFAVAWTYALAV 420

```

도면10h

```

Query: 421 NFAPPYRSVVDADFSTTDVGLREKQRED 447
      NF P YR DAF+T ++G+R+ RED
Sbjct: 421 NFWPWRDTCDAFTTAEIGVRD--RED 445

Query= jgi|Spath1|84164|estExt_Genewise1Plus.C_62100
      (546 letters)

Database: neurospora_crassa_7_proteins_no_asterisks.fas
      9822 sequences; 4,775,003 total letters

Searching.....done

Sequences producing significant alignments:

Score      E
(bits) Value

NCU05853 | neurospora_crassa conserved hypothetical protein (54... 586 e-168
>NCU05853 | neurospora_crassa conserved hypothetical protein (542
      nt)
      Length = 541

Score = 586 bits (1511), Expect = e-168, Method: Compositional matrix adjust.
Identities = 276/516 (53%), Positives = 369/516 (71%), Gaps = 11/516 (2%)

Query: 18 ADPIVTRLVQEDKIPWYKKPNLRMYIFLFCMGMVEMTSQFDSQLINTLQYAETPHKYL 77
      +DP+++RL +ED IPWYKKPNLR +Y LF CMG+E+TSGFDSQ+IN LQ E++ +Y
Sbjct: 11 SDPVISRLAKEDPIPWYKKPNLRFLYFMLFPTCMGIELTSGFDSQMINALQILESWIQYF 70

Query: 78 GNGRKDEDGNYAIEPGLLGFVNSCYQLGSIFAVPIAPWFAQRFGRWSIMLGSIMVGG 137
      N + G +++ Y LG+I ++P+ P RFGRRWSI LGS++M+ GA
Sbjct: 71 DNPQG-----ALKGIISAAYSLGAILSLPLVPIINDRFGRWSIALGSIVMIVGA 120

Query: 138 IIQGFQAHVAMYYIARMILGMGILFCIISGAALIGELGHPKERAVLTSLFNSSYFIGQL 197
      IIQGF+QHV MYI+ARMILG GI CI+SG++LIGEL +PKER VLTSLFN SYF+GQ-+
Sbjct: 121 IIQGFQAHVAMYYIARMILGMGILFCIISGAALIGELGHPKERAVLTSLFNVSFYFVQIV 180

Query: 198 ASAITIGTTEMKTNWAWRLPSLLQICPSLLQIVTVFFLPESPRFLISKDRDDAKEVLK 257
      A+AI GT - +NW WR+PSLLQICPSLLQ+ VFF+PESPR+LI+KDR +A ++L K
Sbjct: 181 AAAIVFGTNSIASNWGWRIPLPSLLQICPSLLQLAFVFFIPESPRWLIITKDRSQEAHDILKK 240

Query: 258 YHAEGDASSLLVQAEIVQIRETIRTEME-VSNQSWMELVSTYGMRRRLVITLFIGLFTQL 316
      YH E + V AE Q++ IR E E VS SWM+L+ T GMR+RL+I+ +GLFTQ
Sbjct: 241 YHGEEMERGEFVAAEFAQMQAIRLEYETVSKSSWMDLLRTPGMRKRLLISSMLGLFTQW 300

Query: 317 SGNLTLSYYSGLFEMMGYTEASVKTRINVANACWSLLNATTIAFLVPYFKRRHMFMTSA 376
      SGNLT+SY+ G L +M+G+T+++ +INV+ ACWSL T--+ LV +RR M+M
Sbjct: 301 SGNLTISYFLGDLKLMIGFTDSTFIQKINVSIACWSLFCGVTVSLLVTRIRRRIMYMACT 360

Query: 377 LSMCAVFIATTVSLERTQAAQDAGFKNTAAGISGLFWYFAFAPCYNMGNNALTYTYLVEL 436
      +S+ +IA TVS+ER + G N A I+ LF+ + ++PCYNMG NALTYTY+VE+
Sbjct: 361 ISLLLCYIAWTVSMERAMTGKANGTFNNGANIATLFFIYMYSPCYNMGYNALTYTYMVEV 420

Query: 437 WPYSHRSRGIGVQQIFGKLGFFSTNVNSIALDAIRWKYMAIYCGWIFFEFLIVFFLYPE 496
      WPY+ RSRGI V Q+FG+L GFF+T VN I L + WKY+ YC W+ FE V+F++PE
Sbjct: 421 WPYAERSRGIAVFQIFGRLAGFFTTFVNPIGLKNVWKYKILSYCCWIAFEVCFVYFMFPE 480

```

도면10i

```

Query: 497 TSGRTLEELAFLEFEDASLNEKAAAAVEKQIHYGDEK 532
      T GRTLEEL F+FE L +A AA EK +++ +
Sbjct: 481 TMGRTLEELTFMFEGEDLQRQANAAAEKVVNTEHD 516

Query= jgi|Spath1|116270|estExt_fgenes1_pm.C_50266
      (488 letters)

Database: neurospora_crassa_7_proteins_no_asterisks.fas
      9822 sequences; 4,775,003 total letters

Searching.....done

Sequences producing significant alignments:

Score      E
(bits) Value

NCU05519 | neurospora_crassa hypothetical protein similar to Tn... 700 0.0

>NCU05519 | neurospora_crassa hypothetical protein similar to TnaI
      (520 nt)
      Length = 519

Score = 700 bits (1807), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 338/512 (66%), Positives = 400/512 (78%), Gaps = 25/512 (4%)

Query: 1 MADEKRMGSSSDSKAAVQHSITLPGVSSSTAERGAATDQNGQPIVQFDLKAEARLRRL 60
      M EKR S + E PG + AERG AATD +G P+V+ D AE +LRRK+
Sbjct: 9 MTSEKRRQSLSESDTKEGYFENAPG-AHYRAERGAATDIHGNPLVELDFVAETKLRRKI 67

Query: 61 DLFIVPTVSLLYLFCFIDRANIGNARIAGLEKDLNLTGYDYNALLSVFYISYIVFEIPSN 120
      DL++VPTV++LYLFCFIDRANIGNAR+ LEKDL+L GYDYNALLSVFY+ YIVFEIP-N
Sbjct: 68 DLYVVPVAILLYLFCFIDRANIGNARLDKLEKDLDLHGVDYNALLSVFYVGYIVFEIPAN 127

Query: 121 IACKWIGPGWGFIPASISLGFVSVSLATAFVDNFAQAAGVRFLLGVEAGMMPGIAYYLSRW 180
      I CKW+SPGWF+P SLGFG++S+ AFV+NF+QA GVRFLLGVEAGM+PGIAYYLSRW
Sbjct: 128 IMCKWMGPGWFLPLTSLGFGIMSVCMFVNNFSQACGVRFLLGVEAGMLPGIAYYLSRW 187

Query: 181 YRRAELTFRLSLYIVMAPMAGAFGGLLASGILSLDHVGGVTGWRMIFVVEGIITIGLSVI 240
      YRR+ELT RLSLYIVM+P+AGAFGGLLASGIL LDH G + GWRMIF +EGIIT+GLS-I
Sbjct: 188 YRRSELTILRLSLYIVMSPLAGAFGGLLASGILKLDHFGSLHWRMIFGIEGIITVGLSLI 247

Query: 241 SFITLTDRPETARWLTQEEKDLAIARVKSERVATTEVLDKMDCKKLIQGILSPVILATSF 300
      F+TLTD P TA+WL+QEEKDLAIARVKSER+ TE++D+MD KKL +GI +PV T F
Sbjct: 248 GFLTLTDHPATAKWLSEQEEKDLAIARVKSERIGQTEIIDKMDAKKLKRGIFNPVVEFTGF 307

Query: 301 MFLNNIT-----QLFTVPPYVVGFFFTLALPLLSWYLDNR 336
      FLLNNIT QL TVPPY+VVGFFFT+ +PL+S +LDRR
Sbjct: 308 AFLNNITVQGLAFFAPTIVVATYPTKNIIQKQLLTVPPYIVGGFFTVLMLPLISRWLDNR 367

Query: 337 QIIILLSTPLVIVGYSMFLGTNPSARYGATFLLSSSLFAVGALSNSQVSANVVSdTARS 396
      QIII+ PLV+VGY MFL T N RYGA FL+S+S+FA G L+NSQVSANVVSdTARS
Sbjct: 368 QIIIIACPLVMVGYIMFLATENAHVRYGAAFLVSTSVFAAGPLTNSQVSANVVSdTARS 427

Query: 397 SAIGLNVMMGNVGGIATWSYLPWDGPNYKIGNGLNLAACCTVLILSAVILLWMMKWDNR 456
      SAI NVM+GNVGGI+ATWS+LP+D P Y IGNG+NLAA LI++ L+WMM DN +
Sbjct: 428 SAIAYNVMLGNVGGIATWSFLPFDAPKYHIGNGINLAASGGTLIATCLLIWMMKRDNNK 487

```


도면10j

```

Query: 457 REGRNAAEEELAGMSRQEIQDLDWKHPAFRWRP 488
      RE RNAEEEL GMS+QEIQDLD+KHP FRW+P
Sbjct: 488 RELRNAEEELTGMSQQEIQDLDYKHPGFRWKP 519

Query= jgi|Spath1|79030|estExt_Genewise1Plus.C_31624
      (519 letters)

Database: neurospora_crassa_7_proteins_no_asterisks.fas
      9822 sequences; 4,775,003 total letters

Searching.....done

                                Score      E
Sequences producing significant alignments:          (bits) Value
NCU01231 | neurospora_crassa hypothetical protein similar to ca... 834 0.0
>NCU01231 | neurospora_crassa hypothetical protein similar to
      carboxylic acid transport protein (520 nt)
      Length = 519

Score = 834 bits (2154), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 395/503 (78%), Positives = 441/503 (87%), Gaps = 1/503 (0%)

Query: 1 MESTHEPADP AKGVLATAKQSWHDLFIFKQRVVTVNELGETSTEWARPVPLRNPISLLA 60
      MESTHEPADP AKG+L TA+QSW DLF+KQRVVTVN GET+TEWA+PVPL+NPISLLA
Sbjct: 1 MESTHEPADP VAKGILPTARQSWKDLFIWKQRVVTVN VYGETATEWAKPVPLKNPISLLA 60

Query: 61 QLSARNWLFFLVGFLAWVADAYDFHALSIQQVKLAEFYNTTKTNISTAITLTLLLRSVGA 120
      QLS R+W+ F-VGF AW ADA+DFHALSIQQVKLA +Y +KT++STAITLTLLLR+GA
Sbjct: 61 QLSGRDWICFLVGFCANSAADAFDFHALSIQQVKLAAYYGVSKT+SVSTAITLTLLLR+GA 120

Query: 121 AFFGLAGDKWGRKWPVMANMIVLGVLIQGTIYSVTFSDFLAVRALFGLFMGGVYGNAIAM 180
      A FGLAGD+WGRKWPVM NMIVLG+LQI TIYS T+S FL VRALFGLFMGGVYGNAIAM
Sbjct: 121 AAFGLAGDRWGRKWPVMVNMIVLGILQIATIIYSSTYSQFLGVRLFGLFMGGVYGNAIAM 180

Query: 181 ALENSPPDARGLMGILQQGYSLGYVIAACANLGVGGDNSWKTIVFWIGAGLSIGVGLVR 240
      ALENSP DARGLMGILQQGY+ GYV AACANLGVGG +SWKTIVFWI AGLSIGVG+-R
Sbjct: 181 ALENSPVDARGLMGILQQGYAFGYVCAACANLGVGGDTSWKTIVFWIAAGLSIGVGII R 240

Query: 241 CFFPESQQFLEARAAGKGQASASAFWKETKMMLAQEWKMCVYCIILMTWFNYYSHTSQDS 300
      CFFPES+QFLEAR GK A+ S FW+ETK+ML QEWKMCVYC ILMTWFNYYSHTSQD+
Sbjct: 241 CFFPESKQFLEARKEGKAHANPSQFWRETKVMLRQEWKMCVYCCILMTWFNYYSHTSQDN 300

Query: 301 YTTFMTLQKELDNDGASRASILMKVGACVGGTIIGYISQWFGRRRTIIIVAALISGLIIPA 360
      YTTF+L KE-DN ASRASIMK GACVGGTIIGY+SQ+FGRRRTIIV++LISG +IPA
Sbjct: 301 YTTFVLRKEMDNSAASRASIMKAGACVGGTIIGYLSQYFGRRRTIIVSSLISGCMIPA 360

Query: 361 WILPEGRSLSVTGFFMQFFVQGAWGVIPIHLNELSPPAFRSSFPGLTYQLGNMISSPSA 420
      WILP ER+LS TGFFMQFFVQGAWGVIPIHLNEL+PPAFRSSFPGLTYQ+GNM+SSPSA
Sbjct: 361 WILPNSERALSATGFFMQFFVQGAWGVIPIHLNELAPPAFRSSFPGLTYQVGNMVSSPSA 420

Query: 421 QIVNAIAESHVSSTSKSGKSVNAYGPTMGIAATAIIATGIAVTTALGPEKRGREFEKTLPAG 480
      QIVNA++E + S +GK V AYGPTMGIAATAII GI VTTA GPEKRGREFEK LPAG
Sbjct: 421 QIVNAVSEKIHIVSHTGKLVEAYGPTMGIAATAIIVMGIVVTTAFGPEKRGREFEKALPAG 480

```

도면10k

Query: 481 MNI-IQGGKAVDDLEKGDSDREK 502
 MN+ Q GK VDDLE EK
 Sbjct: 481 MNLQKQHGKQVDDLEMETGHMEK 503

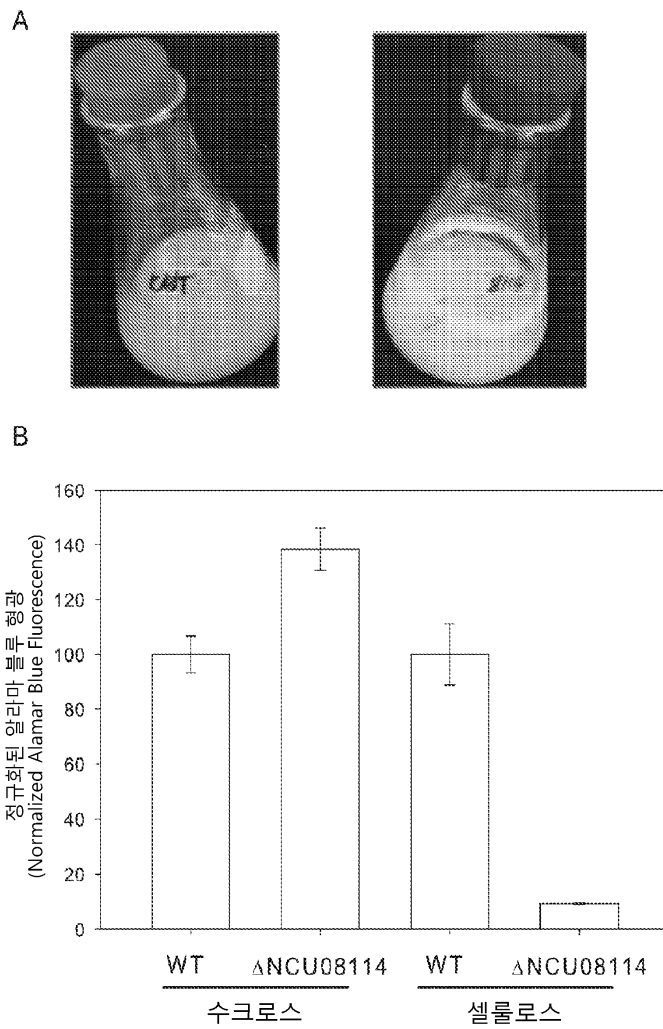
Query= jgi|Spath1|108890|estExt_fgenes1_pg.C_60848
 (533 letters)

Database: neurospora_crassa_7_proteins_no_asterisks.fas
 9822 sequences; 4,775,003 total letters

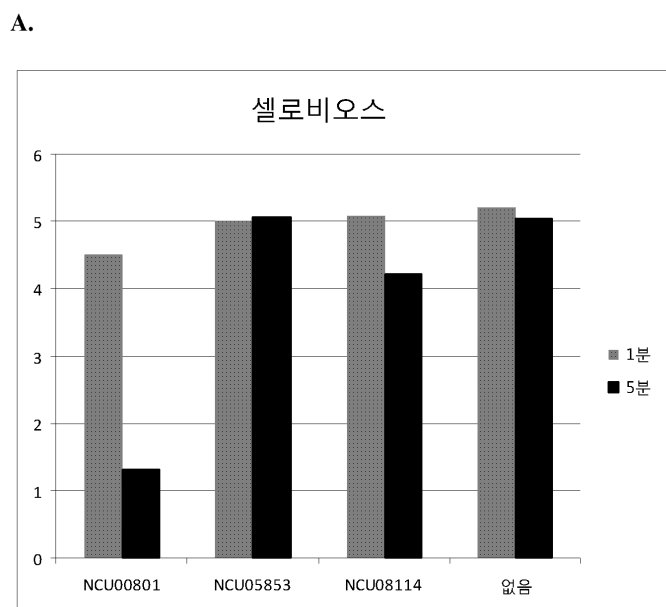
Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
NCU00988 neurospora_crassa hypothetical protein similar to MF...	878	0.0
>NCU00988 neurospora_crassa hypothetical protein similar to MFS quinate transporter (538 nt) Length = 537		
Score = 878 bits (2269), Expect = 0.0, Method: Compositional matrix adjust. Identities = 429/538 (79%), Positives = 471/538 (87%), Gaps = 10/538 (1%)		
Query: 1 MGLS--TKILQKIVRNEAMASDPPEIYGWRVYLLACSAFCGAMSGWDSSVIGGVIVLPP 58 MGLS +IL-KIV+NEAMA DPPEIYGWRVYLLACSAFCGAMSGWDSSVIGGVI L P Sbjct: 1 MGLSIGNRILRKIVRNEAMAEDPPEIYGWRVYLLACSAFCGAMSGWDSSVIGGVIELEP 60		
Query: 59 FIRDEN-LGDPKSQLASANLSANIVSTLQAGCFLGALVASPMTDRFGRKWCLIGVSLIIII 117 F DF +G+ K A ANL ANIVSTLQAGCFLGAL+ASP+TDRFGRKWCLI VSL++II Sbjct: 61 FKHDGFGFIGNDK--AKANLGANIVSTLQAGCFLGALIASPITDRFGRKWCLIAVSLVVI 118		
Query: 118 GIIMQAAASGNLGFMYAGRFIAGAGVGAASTINPIYVSENAPRAIRGLLTGLYQLFIVTG 177 GIIMQAAASGNL PMY GRF+AG GVGAAS INP++VSENAPR+IRGLLTGLYQLFIVTG Sbjct: 119 GIIMQAAASGNLAFPMYIGRFVAGVGGAASCINPVFVSENAPRSIRGLLTGLYQLFIVTG 178		
Query: 178 GMIAFWINYSVSIIHFPETKIMYVFPLAIQALPAALLCLCMLLCQESPRWLARRDRWEDTK 237 GMIAFWINYSVS+HF + K MY+PPLAIQ LPA LLC+CMLLC ESPRWLARRDRWE+ K Sbjct: 179 GMIAFWINYSVSLHF-KGKSMYIFPLAIQGLPAGLLCVCMLLCHESPRWLARRDRWEECK 237		
Query: 238 RVLSRIRNLPDPHPYIQDEFQEIQAQLEHERRLIGDASFVNLRQEMWTIAGNRRRLISI 297 VL+RIRNLPDPHPYI DEF+EI QLE ERRL GDA++W+L R+MWT+AGNR+R LISI Sbjct: 238 SVLARIRNLPDPHPYIVDEFREIQDQLEQERRLQGDATYWDLT+RDMWTVAGNRKRALISI 297		
Query: 298 ILMICQMTGTNAINTYAPTIFKNLGLTGTSTSLFSTGVYGVKVTSCIIIFLLFMADSLG 357 IMICQMTGTNAINTYAPTIFKNLG+TGTSTSLFSTG+YGIVKV SC+IFLLF+ADSLG Sbjct: 298 FLMICQMTGTNAINTYAPTIFKNLGITGTSTSLFSTGIYGVKVVSCVIFLLFLADSLG 357		
Query: 358 RRRSLLWTSIAQGLAMFYIGLYVRIAPPKGESVPPAGYFALVCIFLFAAFFQFGWGPAC 417 RRRSLLWTSIAQGLAMFYIGLYVRI+PP +G+ VPPAGY ALVCIFLFAAFFQFGWGPAC Sbjct: 358 RRRSLLWTSIAQGLAMFYIGLYVRISPPIDGQVPPAGYVALVCIFLFAAFFQFGWGPAC 417		
Query: 418 WIYASEIPAARLRSINVAYAAATQWLFNFVVARVPVPMIVIMGEGGYGYTYLLFGSFCFSM 477 WIYASEIPAARLRSINV+YAAATQWLFNFVVAR VP M+VI+G GYGYTL+FGSFC SM Sbjct: 418 WIYASEIPAARLRSINVSAYAAATQWLFNFVVARVPTMLVTVGPHGYGYTYLIFGSFCFSM 477		
Query: 478 FVFVWFVFPETKGVSLAMDKLFGVTD---ESSKSLTVDEDAKEKEKDGPHARQTEV 531 FVFVWFVFPETKG+SLE MD+LFGVTD E S D+ E K ++ EV Sbjct: 478 FVFVWFVFPETKGISLEHMDLFGVTDGPAAEKSSVHGSDVSGEMGKDGQSKHVEV 535		

도면11

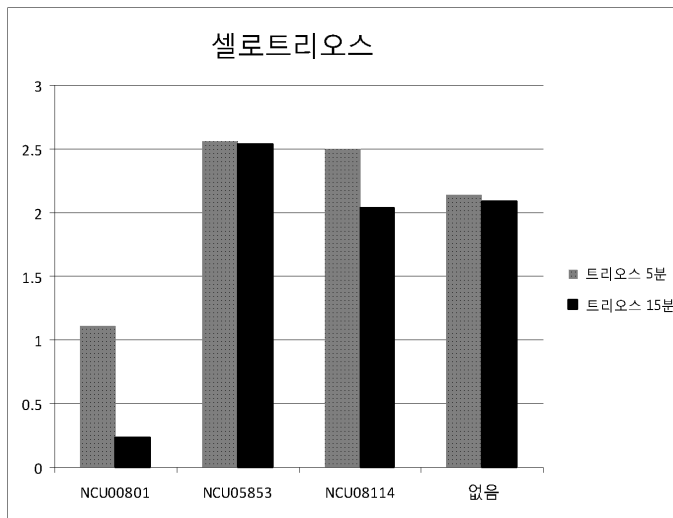


도면12a



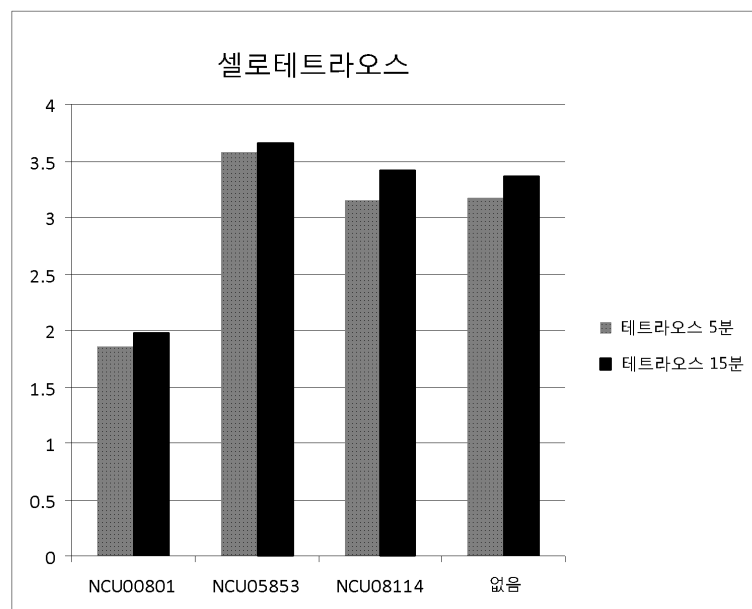
도면12b

B.



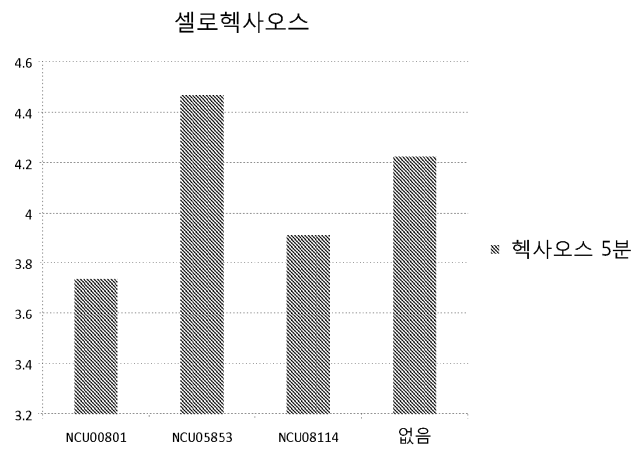
도면12c

C.

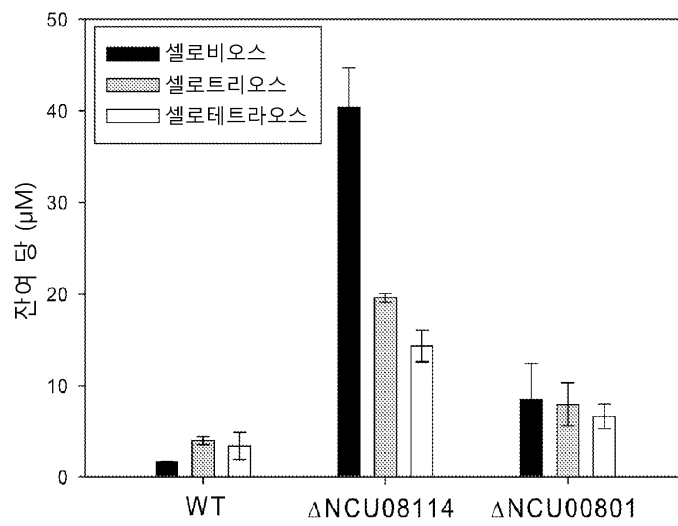


도면12d

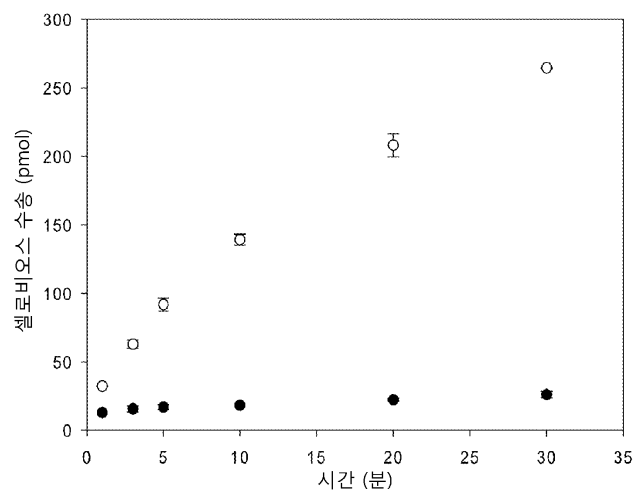
D.



도면13

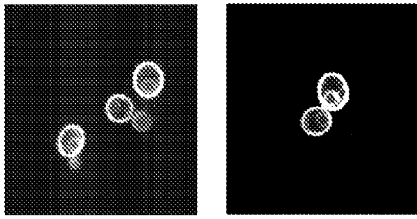


도면14

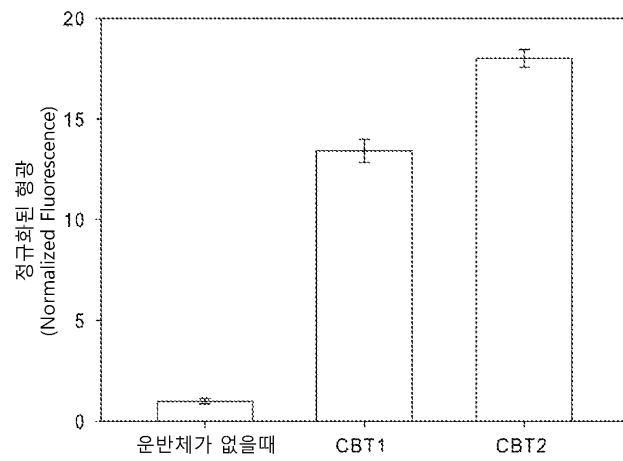


도면15

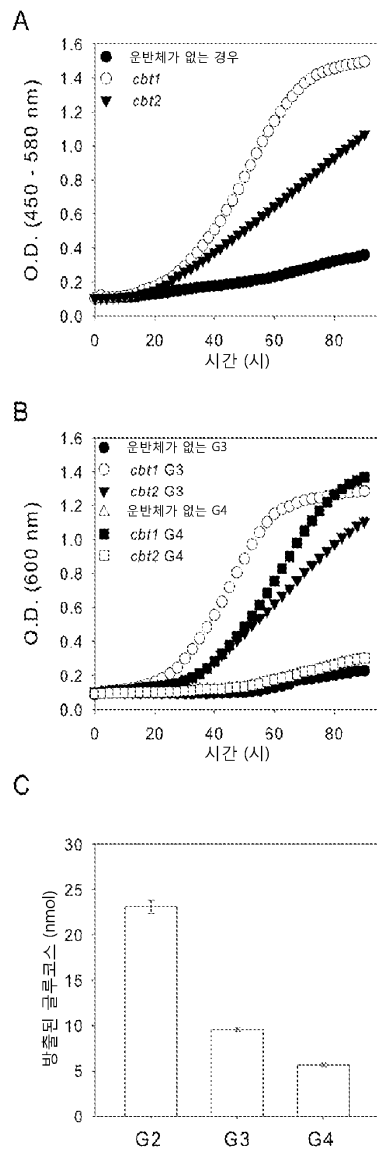
A



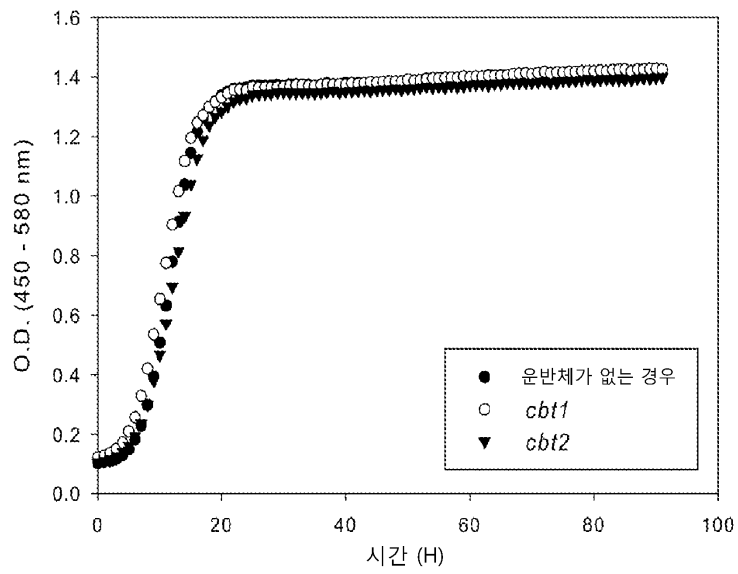
B



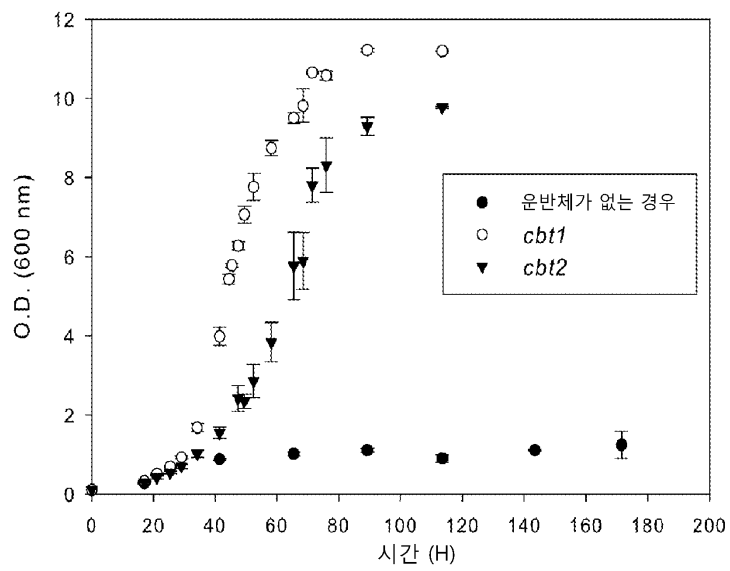
도면16



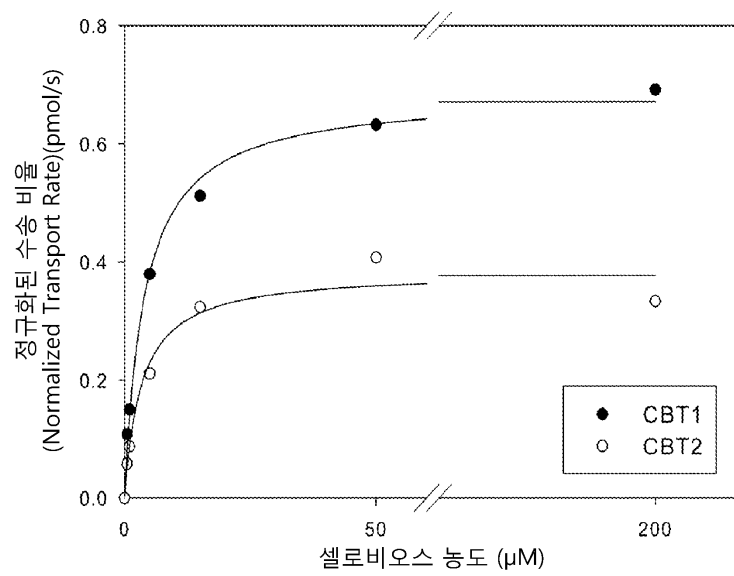
도면17



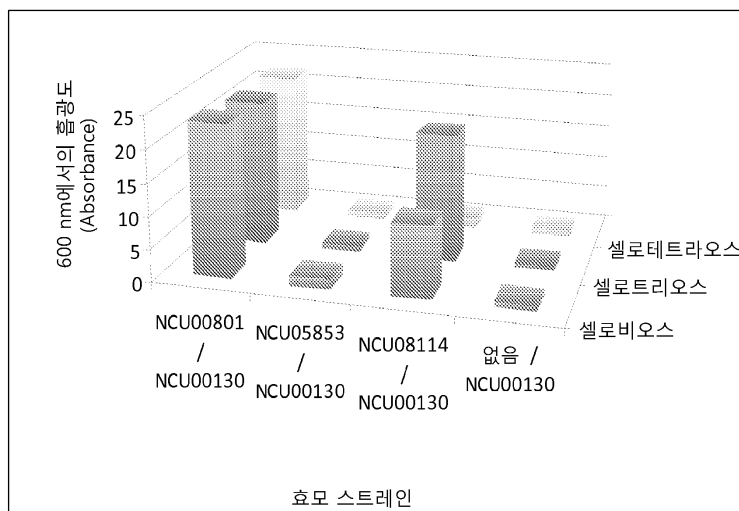
도면18



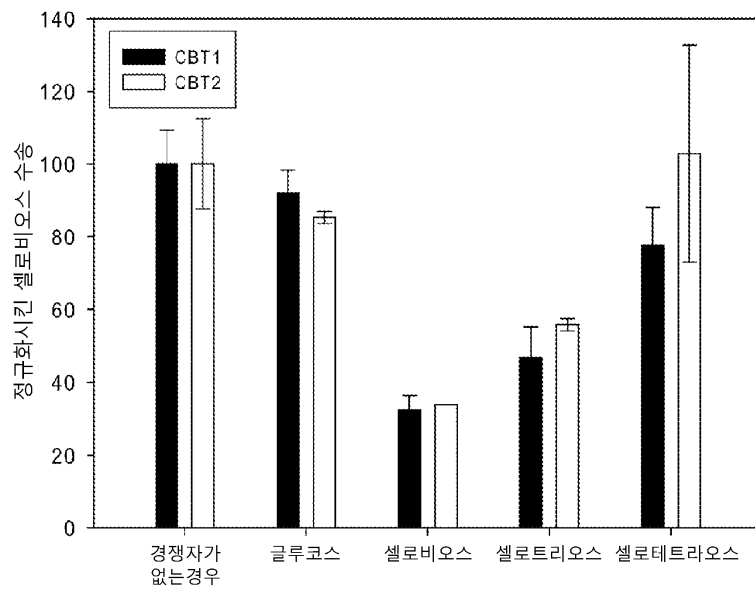
도면19



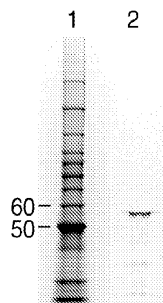
도면20



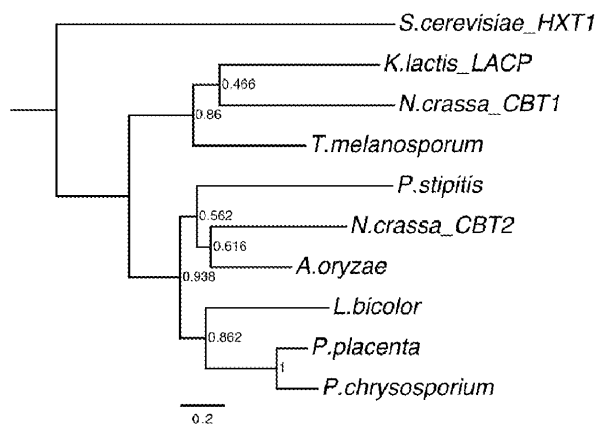
도면21



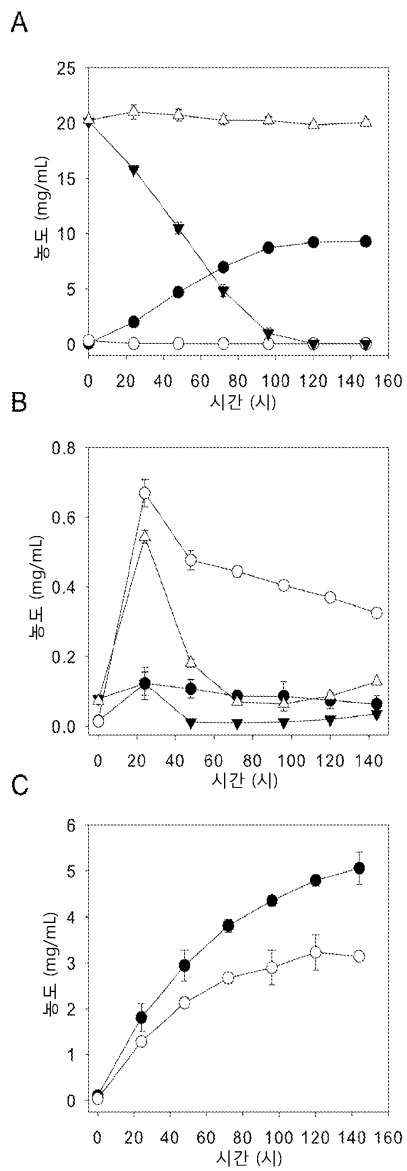
도면22



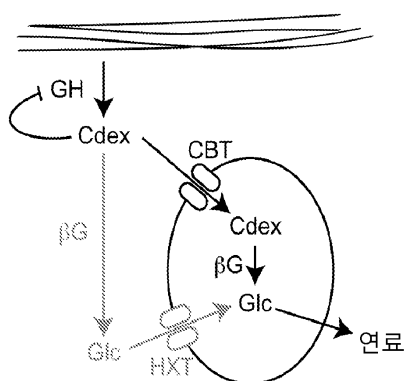
도면23



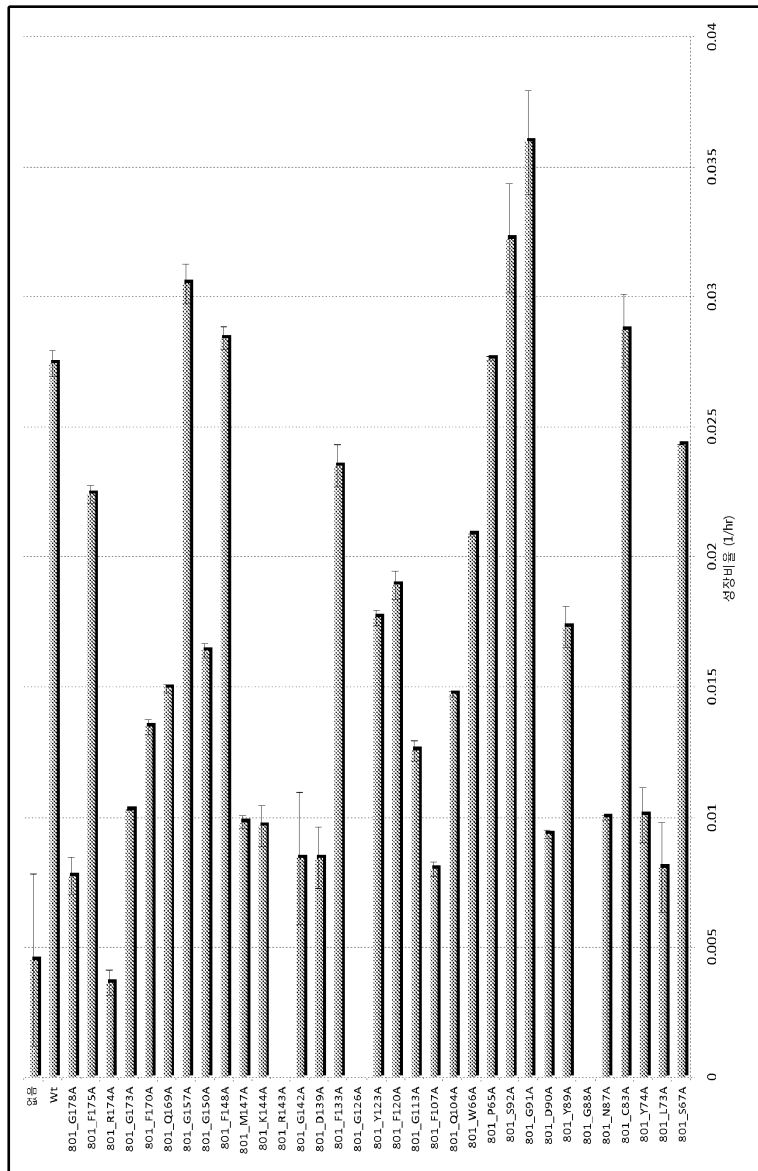
도면24



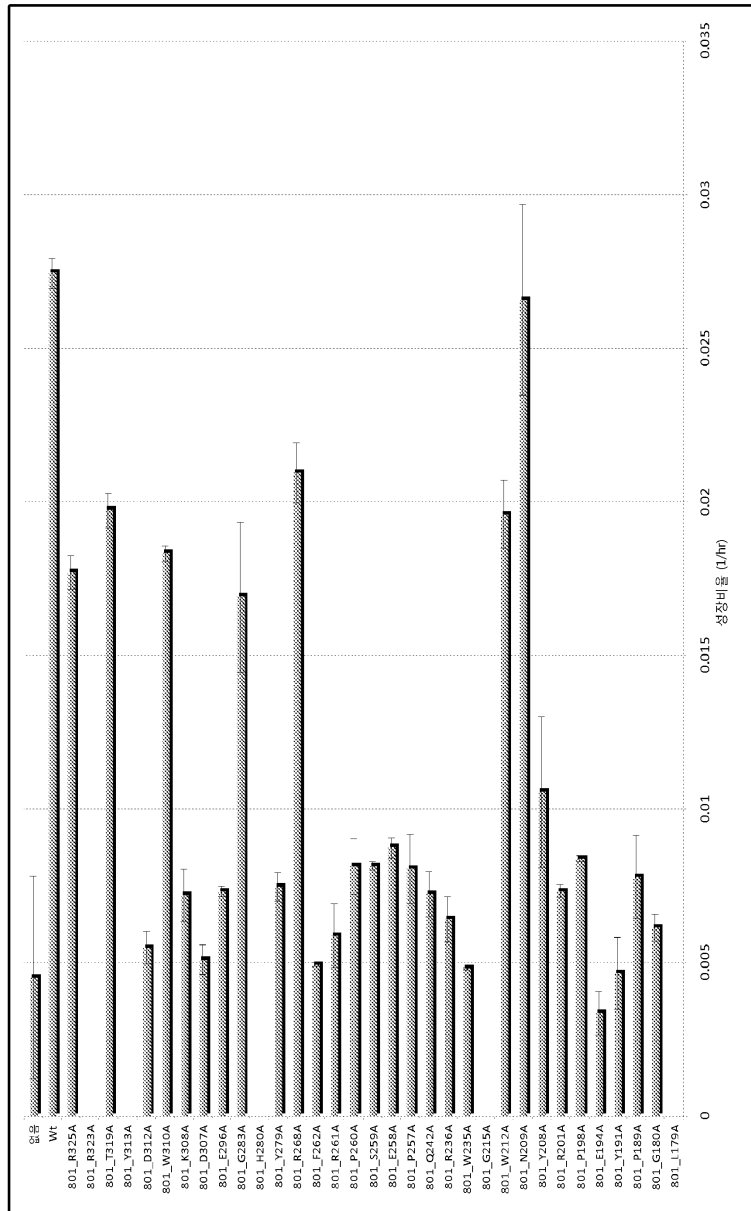
도면25



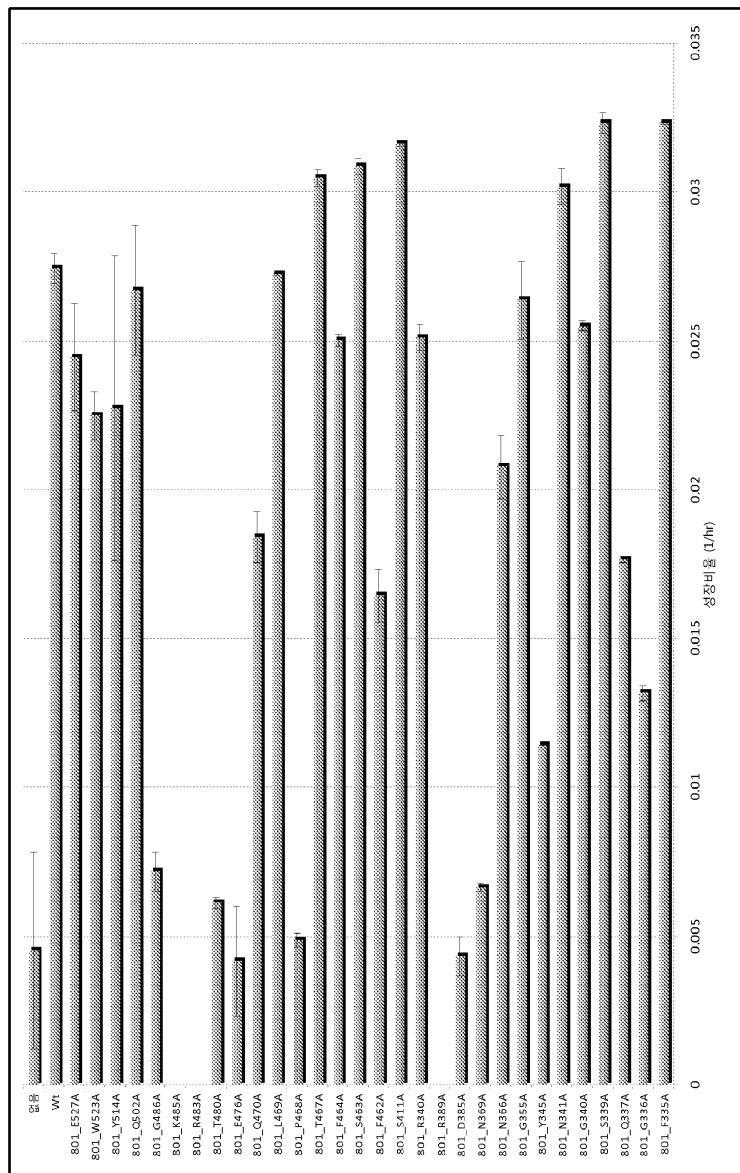
도면26a



도면26b



도면26c



도면26d

B

폴리펩티드 서열 : NCU00801

MSSHGSHDGASTEKHLATHDIAPTHDAIKIVPKGHGQTATKPGAQEKEVR
 NAALFAAIKESNIKPWSKESIHL^YFAIFVAFCCACANG^YD^GSLMTGIAM
 DKFQ^NQFHTGDTGPKVSVIFSL^YTV^GAMVGAPFAA^ILS^DRFGRKKGM^FFIG
 GIFII^VGSII^VASSSKLAQ^FVVGRFVL^GLGIAIMTVAAPAY^SIEIAPP^HW
 RGRCTGFY^NCGWFGGSIPAACITYGCYFIKSN^WSWRIPLILQAFTCLIV^M
 SSVFF^LPESPRFLFANGRD^EAEAVAFLVKYH^GNGDPNSKLV^LLETE^EMRDG
 IRTDGV^DKV^WWD^YRP^LFMTHSGRWRMAQ^VLMISIE^GQFS^GNGLG^YFNTVI
 FKNIGV^ISTSQQ^LAYNILNSVISAIGALTAVSMTDRMPRR^VLIIGTFMC
 AAALATNSGL^SATLDKQTQRTQINLNQGMNEQDAKDNAYLHVDSN^YAKG
 ALAAY^FLENVIF^SF^YTYT^PLQGV^IPT^EAL^ETTIRGKGL^ALSGFIVNAMGFI
 NQFAG^PIALHNIGYKYIFV^FVGWDLI^ETVAWY^FFGVESQGR^TLEQLEW^VY
 DQPNPV^KASLKV^EKVVQADGHVSEAIVA

이탈릭체 : 모든 운반체

이중-밑줄 : β -연계된(linked) 운반체

캡드(capped) : NCU00801 계통군(clade)

볼드체(Bold) : *S. 세레비시아* Hxt1/Hxt3 에서 필수

밑줄친 것 : 알라닌 스캐닝 실험(alanine scanning experiment)에 따른

기능적으로 중요한 것

도면26e

C

폴리펩티드 서열 : NCU08114

MGIFNKKPVAQVDLNQIQEEAPQFERVDWKKDPLRKL^YFYAFILCIAS
 ATTGY^DG^MFFN^SVQ^NFETWIKY^FGD^PRGSEL^GLLGALYQI^GSIGSIP^FV^P
 LLTD^NFGRKTP^IIIGCV^IMIVGA^VLQATAKNLDT^EMGGRTML^GF^GNSLA^Q
 IAS^PMLL^TELAH^FQHRARL^TII^NCL^WNVGA^LVVSWLAFGT^NYINND^WS^W
 RIPALLQAFPS^IIQLLGIW^WV^PESPRFLIAKDKHDEALHILAKYHANGDP
 NHPTV^QFEFREIKETIR^LEMESTKNSSYLDFFK^SRGNRYRLA^ILLSLGFF^E
 S^QWS^GNAI^SNYSSKLYETAGV^IDTSTAKLGLSAGQTGLALIVSVTMALL^V
 DKLGRRLA^FLASTGGMCGTFVIW^TLTAGLYGEHRLKGADKAMIFFIWVFG
 IFY^SLAWSGL^LVG^YAI^ETL^PYRLRGKGLMVNMMSVQCALT^NTYAN^PVAF
 DYFGPDHSWKLYLI^YTC^WIAAE^FV^FV^FFM^YV^ETKGP^TLEELAKVIDGDEA
 DVAHIDIHQVEKEVEIHEHEGKSVA

이탈릭체 : 모든 운반체

이중-밑줄 : β -연계된(linked) 운반체

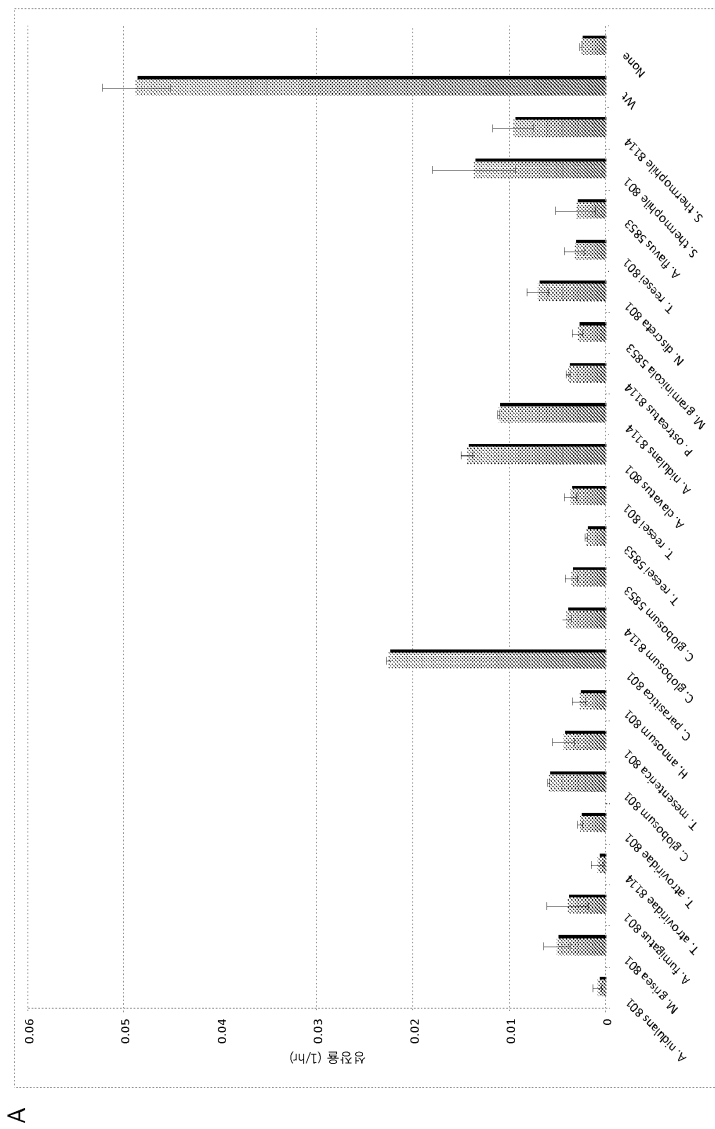
캡드(capped) : NCU00801 계통군(clade)

볼드체(Bold) : *S. 세레비시아* Hxt1/Hxt3

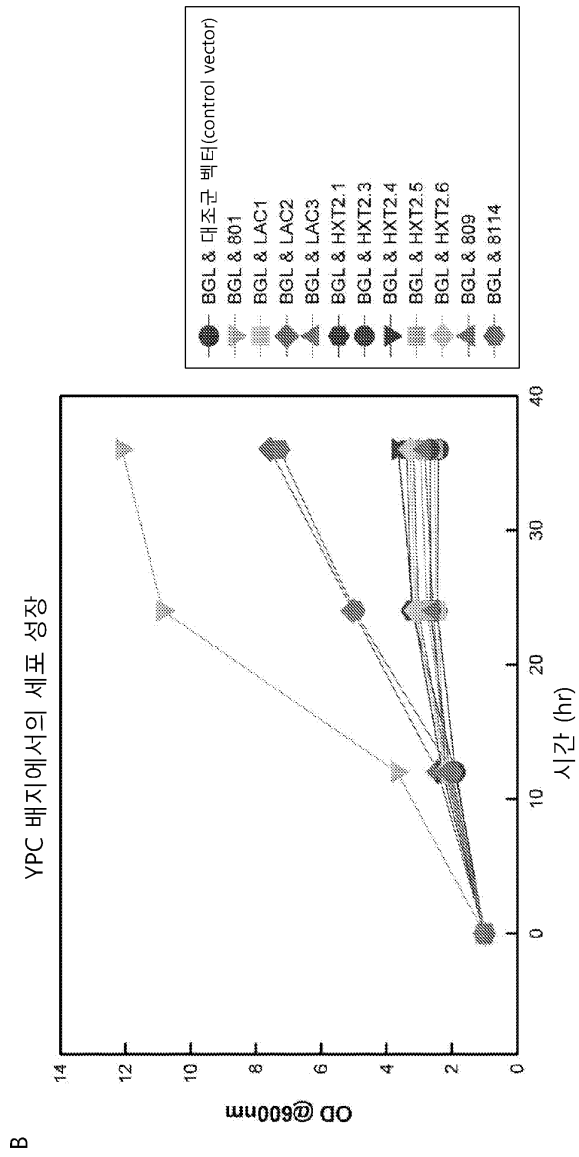
밑줄친 것 : 알라닌 스캐닝 실험(alanine scanning experiment)에 따른

기능적으로 중요한 것

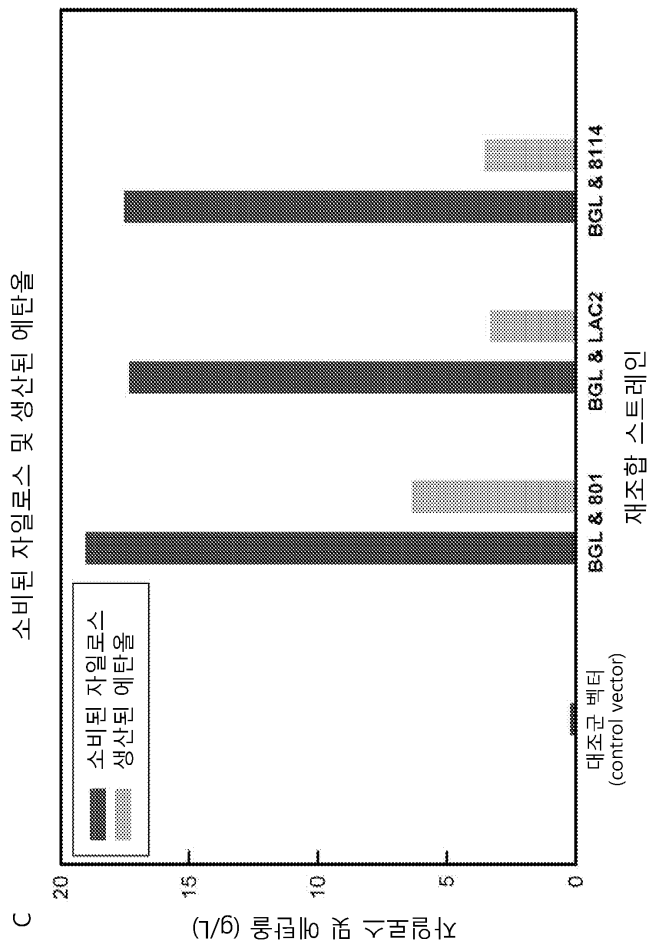
도면27a



도면27b



도면27c



도면28d

B

```

NCU00801      MSHSGSHDGASTEKH-LATHDIAPTHDA--IKIVPKGHGQTATKFGAQEKEVRNAALFAA
NCU008114     -----MGIFNKKFVAQAVDLNQIQEEAP
LAC2          MST-----NSLNDSYNPSSTKEKDIVVQSEALADVAIETAFETDGYK-----KI
A_clavatus    MSS-----TDEKHGISVHHDSPGSDADFKPAVAAGQGETAKVGGAKTRQVHNAELYAA
.

NCU00801      IKESNKPWSKESIHLYFAIFVAFCCACANGYDGSMLTGIIAMDKFQNFHTGDTGPKVS
NCU008114     QFERVDWKKDPGLRKLYFYAPILCIASATTGYDGMFFNSVQNFETWIKYFGD-PRGSELG
LAC2          FQEHFVPRWTKRLRSIYFTCLVYLVSTTNGYDGSLLSSLITMPEFISHLNI-KSASGTG
A_clavatus    IETPTEKWSKESLHMYFAIFVAFCCACANGYDGSMLGAILAMDRHQNTHGTMDGPKVS
      *          : **  : :  : : ****  : :  : :  : :  : :  : :  : :  : :  : :
.

NCU00801      VIFSLYTVGAMVGAPFAAILSDRFGRKKGMFIGGFIIVGSIIVASSSKLAQFVVGRFVL
NCU008114     LLGALYQIGSIGSIPEFVPLLDNFGRKTPIIIGCVIMIVGAVLQATAKNLDFMGGRTML
LAC2          IVFAIFQVGQMVATLF-VVLGDFIGRRAIFIGSVIVCLGAIITSIANNISTFGGRFLL
A_clavatus    LVTSLYTVGSIAATPFSAVLSDKLGRKCMFVGAWVILAGSIIATAKHLEQFVVGRRVL
      : : : : : *  : .  *  *  * : ** : . : : * : : : : : :  *  *  * : *
.

NCU00801      GLGIAIMTVAAPAYSIEIAPPHWRGRCTGFYNCWFGGSIPAACITYGC--YFISKNSWSW
NCU008114     GFENSLAQIASPMLLTELHPQHRARLTTIYNCLWNVGVLSWLAFTG--NYINNDWSW
LAC2          SFGSGSCALSTTYLEITSPDERSALCAIYNSLYYIGSIIATWSSYATSIYSYANSVLSP
A_clavatus    GFGIQMVVSAPAYAVEIAPPHWRGRAVGFYNCWFGGSIPAAAVTYGT--NFIDNDFSFW
      : *  : .  : .  * : : *  *  *  : ** : . : * : : : : : : : . : .  * :
.

NCU00801      RIPLILQAFT-CLIVM-SSVFFLPESPRFLFANGRDAAVAFVVKYHGNQDPNSKLVLE
NCU008114     RIPALLQAFP-SIIQL-LGIWVPEPRFLIAKDKHDEALHILAKYHANGDPNHPTVQFE
LAC2          RIPLWQILCPALVVIGLLVGVAPEPRFYILTGQPKARAFFCKYHANGDEKHPIVEYE
A_clavatus    RIPFILQCFACVIVII-FAVWFIPESPRWQMAHQDEAALAFTRYHGNQDPNARLVLE
      **  *  *  : : : : : :  * : : : : : : : : : : : : : : : : : : : *
.

NCU00801      TEEMRDGI-RIDGVDKV-WWDYRPLEMTHSGRWMAQVLMISIFGQFSGNGL-GYFNNTVI
NCU008114     FREIKETI-RLEMESTK-NSSYLDFFKSRGNRYRLALLSLGFFSQWNGNAIISNYSKL
LAC2          MAQLELSLLEVPKLRVRDYFDARILFKTKSRIYRSLVCIASHAFGLSGNNAVVGYYITNI
A_clavatus    IEEMREGI-RIDGIDKR-WWDYRPFPTHSGRWRFQVMMISVFGQWNGNGL-GYFNNTI
      : :  : .  : .  : : * : : : *  : .  : . *  * : : : : : : : : : :
.

NCU00801      FKNIGVISTSQQLAYNILNSVISAIGALTAVSMIDRMPRAVLIIIGIFMCAAAALATNSGL
NCU008114     YETAGVIDSTAKLGLSAGQTGLALIVSVTMALLVDKLGRRALAFASITGGMCGTFVIWT--
LAC2          FLELGTNPTTRLNLLGVNSILGFIFAMSGSILVGRIGRRPILLYSTTGTFVISFTIIAAC
A_clavatus    YQRLGYTSSSMQLLMNLVNSIVSAIGALTAVALTDRMPRRKVLVWGLACAIAMAINAGV
      :  *  * : : *  .  : : : .  * : : : : : : : : : *  : :  :
.

NCU00801      SATLDKQTRGTQINLNQGMNEQDAKDNYLHVDSNYAKGALAAYFLF-NVIFSFTYTP
NCU008114     -----LTAGLYGE-----HRLKGADKAMIFPIWVF-GIFYSLANSG
LAC2          I-----AAAYTN-----NNNQVAAKVGIPIYIFNNVFFSFGYTP
A_clavatus    SEPMIKQAEIPAGIN-----KTFGQTAVAFYILF-NIIFSFTYTP
      .  :  :  : : *  : : : : : : : : : : : : : : : : : *
.

NCU00801      QGVIPAEALETTIRGKGLALS GFIVNAMGFINQFAGPIALHNIG---YKYIFVFGWDL
NCU008114     LVGYAEILPYRLRGKGLMVNMMSVQCALTINTYANPVAFDYFGEDHSWKLYLYTCWIA
LAC2          QPLYPAEILSSEMRAKGMALFQITQGTASFINTYAAPVAMQNIK---YWYVFFVFWDT
A_clavatus    QGVIPAEALETTIRAKGLALSGLMVSGIGFVSQVAPGLRNIIS---THYFWFVFGWDL
      .  *  *  *  * : : : : : : : : : : : : : : : : : : : : *
.

NCU00801      IETVAWYFFGVESQGRTEQLLEWVYDQPNPVKASLKVEK-----VVVQADGHVSEAIV
NCU008114     AEFVVFVFFVETKCPTEELAKVIDGDEADVAHIDIHQ-----VEKEVEIHEHEGKSV
LAC2          FEVIIYLYFVETKNLTLEEELIFESATPVKTSMIISKPGHAANEELKRLANLKLGNV
A_clavatus    FEALCWYLFGVESQGRTEELWVYQQPNPVKASLQVDK-----VVVQADGGVIEKITD
      *  :  : : : : : : : : : : : : : : : : : : : : : : : : :
.

NCU00801      -A
NCU008114     -A
LAC2          VA
A_clavatus    -A
      *
```

도면28e

C

BLASTP 2.2.24+
Reference: Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Reference for compositional score matrix adjustment: Stephen F. Altschul, John C. Wootton, E. Michael Gertz, Richa Agarwala, Aleksandr Morgulis, Alejandro A. Schaffer, and Yi-Kuo Yu (2005) "Protein database searches using compositionally adjusted substitution matrices", FEBS J. 272:5101-5109.

RID: 4HJKWK6311R

Query= NCU00801
Length=579

		Score
E	Sequences producing significant alignments:	(Bits)
value		
1c1 47823	NCU08114	238
4e-67		

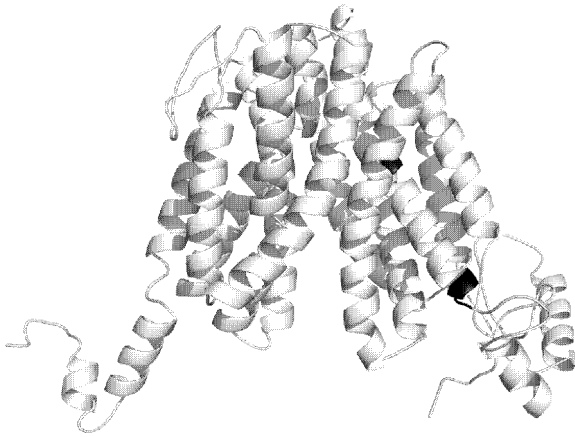
ALIGNMENTS
>1c1|47823 NCU08114
Length=525

Score = 238 bits (608), Expect = 4e-67, Method: Compositional matrix adjust.
Identities = 152/518 (29%), Positives = 257/518 (49%), Gaps = 35/518 (6%)

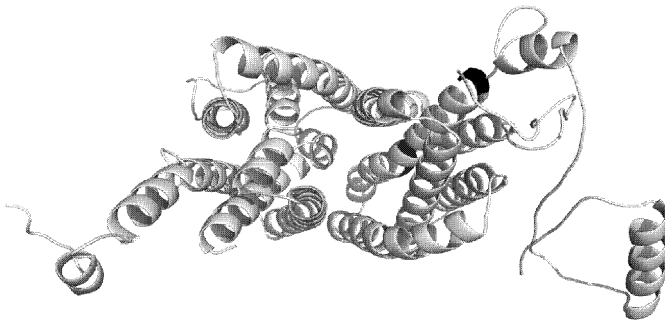
Query	66	WSKES--IHLYFAIFVAFCCACANGYDGLMTGTIAMDKFNQFHTGD-TGPKVSVIFSL	122
		W K+ LYF F+ + GYDG + + + F GD G ++ ++ +L	
Sbjct	30	WKKDPGLKLYFYAFILCTASATTGYDGMFFNSVQNFETWIKYF--GDPRGSELGLL	87
Query	123	YTVGAMVGAPFAAILSDRFGRKKGMFIGGIFIIVGSIIIVASSSKLAQFVVGFRFVLGLGIA	182
		Y +G++ PF +L+D FGRK + IG + +IVG+++ A++ L F+ GR +LG G +	
Sbjct	88	YQIGSIGSIPFVPLLTDFGRKTPIIIGCVIMIVGAVLQATAKNLDTFMGGRTMLGFGNS	147
Query	183	IMTVAAPAYSIEIAPPHWRGCTGFYNCGWFGGSIPAACITYGCFYIKSNWSWRIPLIQ	242
		+ +A+P E+A P R R T YNC W G++ + + +G +I ++WSWRIP +LQ	
Sbjct	148	LAQIASPMLLTALHPQHRARLTTIYNCLWNVGALVVSFLAFGTNYINNDWSWRIPALLQ	207
Query	243	AFTCLIVMSSVFELPESPRFLFANGRDAEAVAFVKYHGNQDPNSKLVLLLETEEMRDGIR	302
		AF +I + ++++PESPRFL A + EA+ L KYH NGDPN V E E+++ IR	
Sbjct	208	AFPSIIQLLGIDWVPESPRFLIAKDKHDEALHILAKYHANGDPNHPVTQFEFREIKETIR	267
Query	303	TDGVDKVWWDYRPLFMTHSGRWRMAQVLMISIFGQFSGNG-LGYFNTVIFKNIGVTSISQ	361
		+ Y F + R+R+A +L + F Q+SGN + +++ +++ GVT ++	
Sbjct	268	LEMESTKNSSYLDFKSRGNRYRLAILLSLGFSSQWSGNAIISNYSSKLYETAGVTDSTA	327
Query	362	QLAYNILNSVISAIGALTAVSMTDRMPRRRAVLIIIGTFMCAALATNSGLSATLKDQTRG	421
		+L + + + + I ++T + D++ RR A LA+ G+ T T	
Sbjct	328	KLGLSAGQTGLALIVSVTMALLVDKLGR-----LAFLASTGGMCGTFVWIW---	374
Query	422	TQINLNQGMNEQDAKDNAYLHVDSNYAKGALAAYFLFNVISFITYTLPQGVIPTEALET	481
		L G+ + H K + ++F + +S ++ L E L	
Sbjct	375	----LTAGLYGE-----HRLKGADKAMIFFIWFVGFIFYSLAWSGLLVGYAIEILPYR	422
Query	482	IRGKGLALSGFIVNAMGFINQFAGPIALHNIG---YKYIFVFGWDLIETVAMVYFVGVE	537
		+RGKGL + V +N +A P+A G +K ++ W E V +F VE	
Sbjct	423	LRGKGLMVNMMSVQCALTNTYANPVAFDYFGPDHWSKLYLIYTCWIAAEFVVFVFFMYVE	482
Query	538	SQGRITLEQLEWVYDQPNPVKASLKVEKVVVQADGHVSE	575
		++G TLE+L V D A + + +V + + H E	
Sbjct	483	TKGPTLEELAKVIDGDEADVAHIDIHQVEKEVEIHEHE	520

도면29a

A
a.



b.

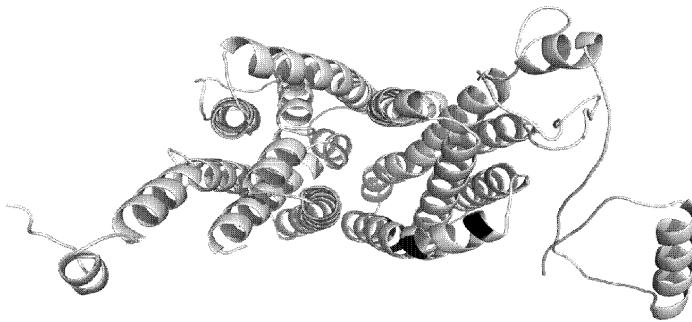


도면29b

c.

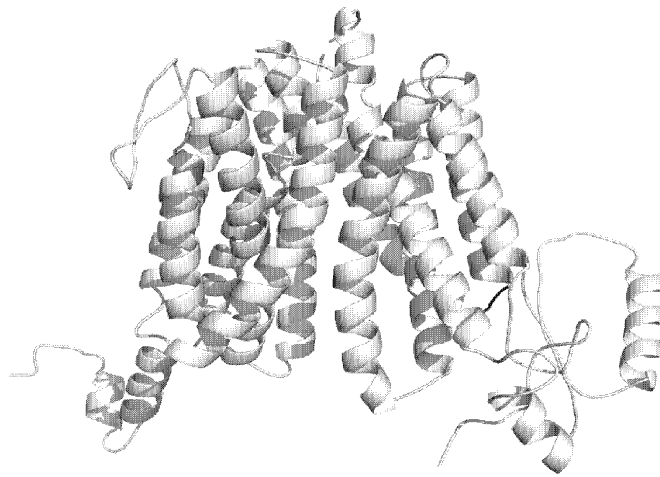


d.

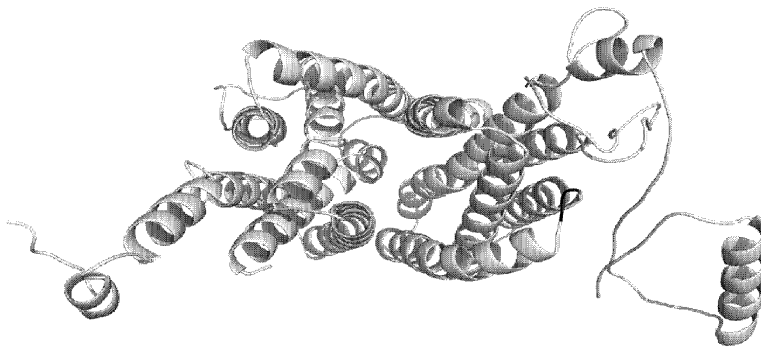


도면29c

e.

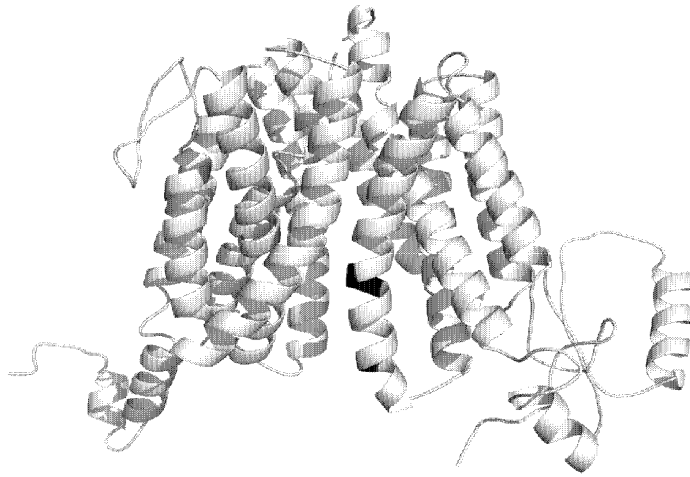


f.

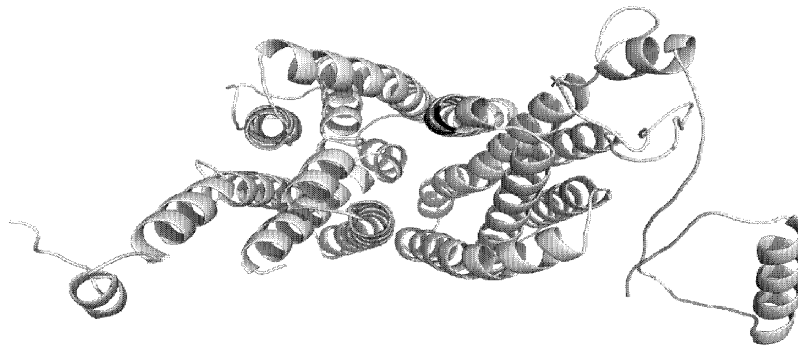


도면29d

g.

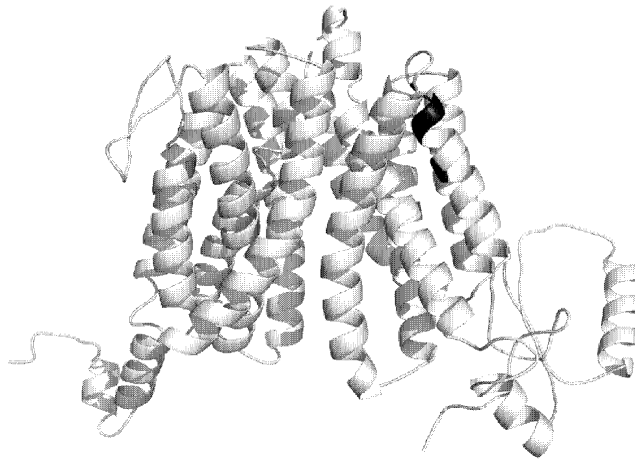


h.

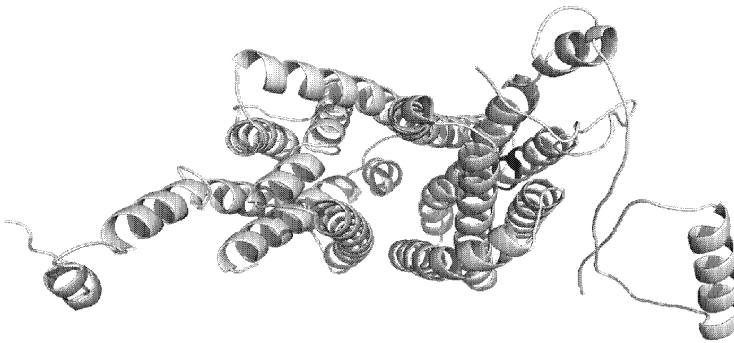


도면29e

i.



j.

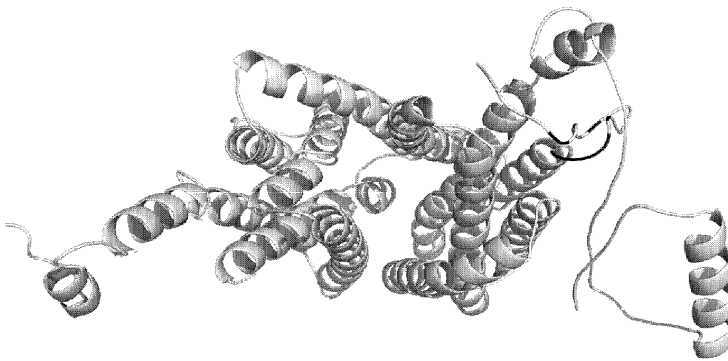


도면29f

k.



l.

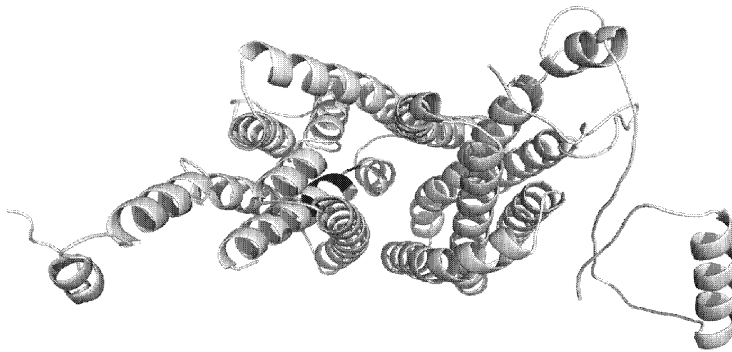


도면29g

m.

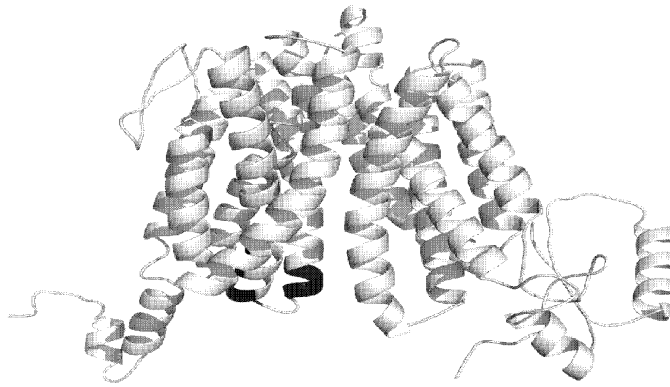


n.

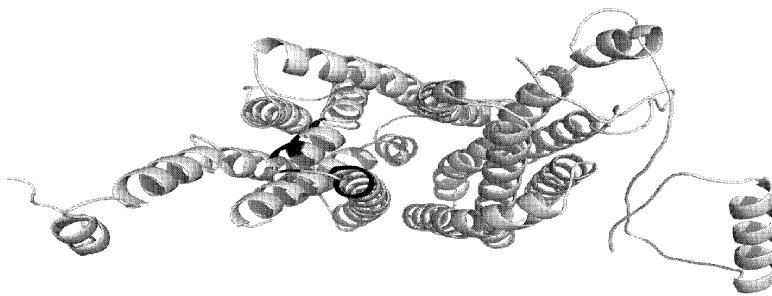


도면29h

o.



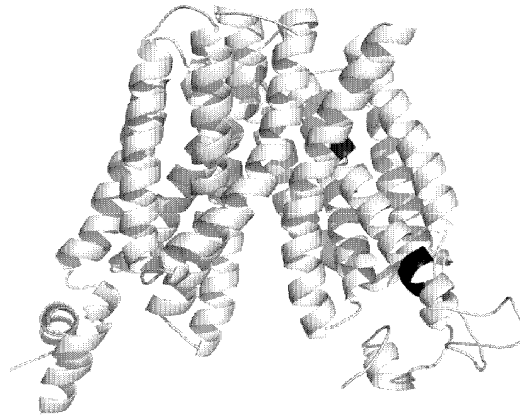
p.



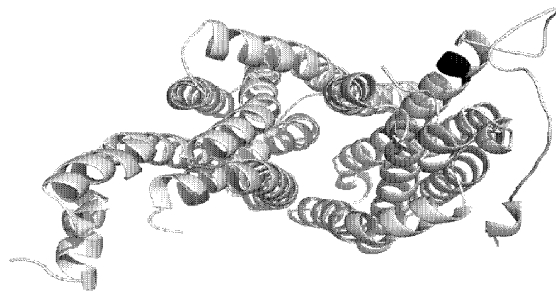
도면29i

B

a.

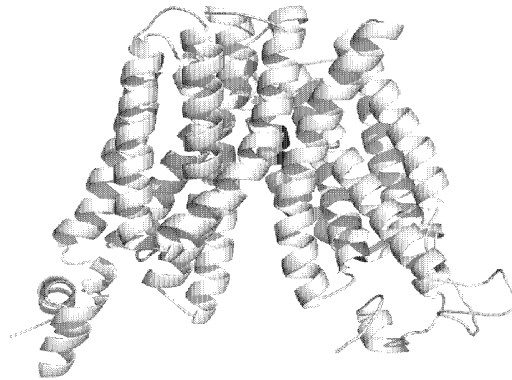


b.

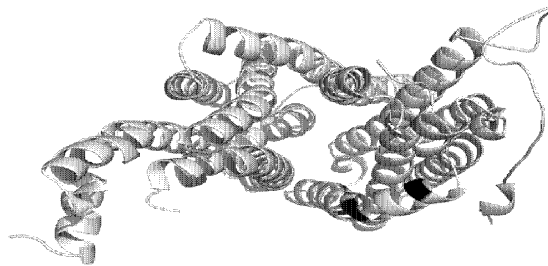


도면29j

c.

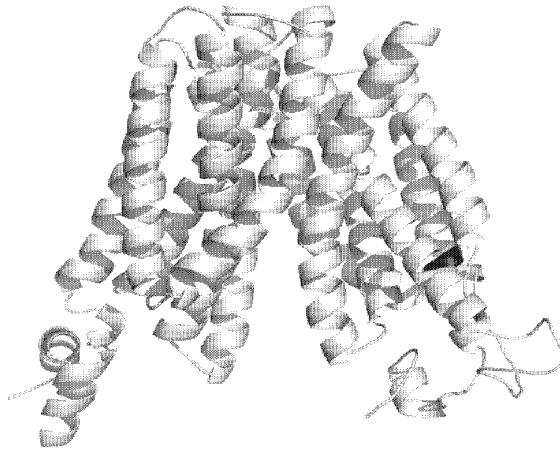


d.

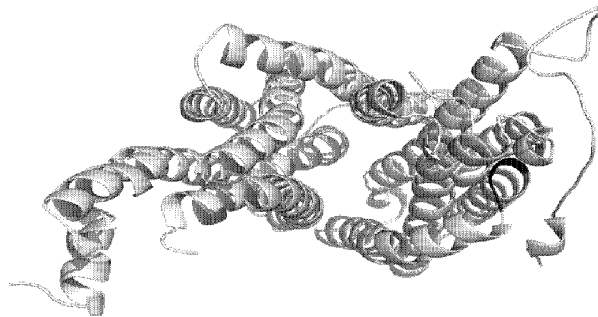


도면29k

e.

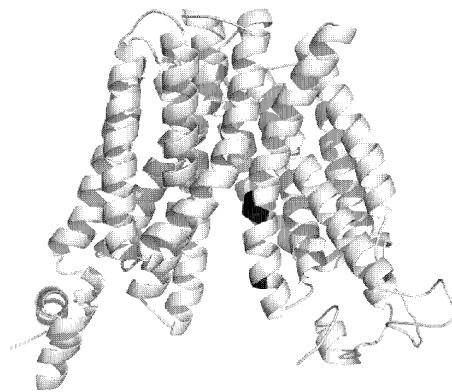


f.

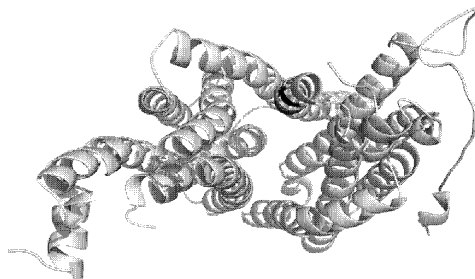


도면291

g.

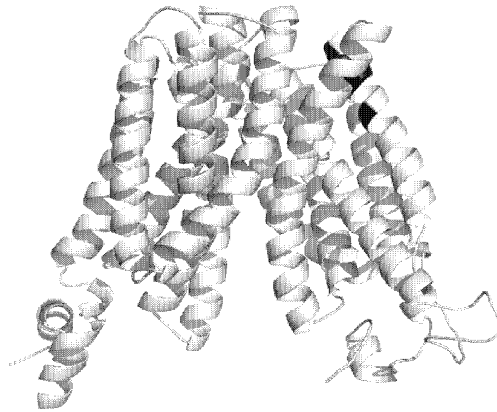


h.

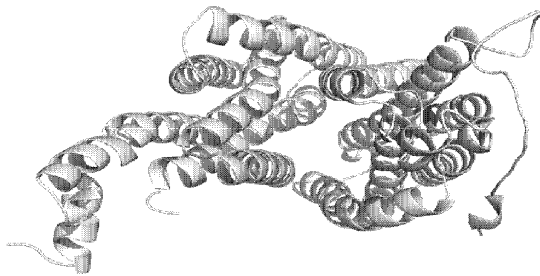


도면29m

i.

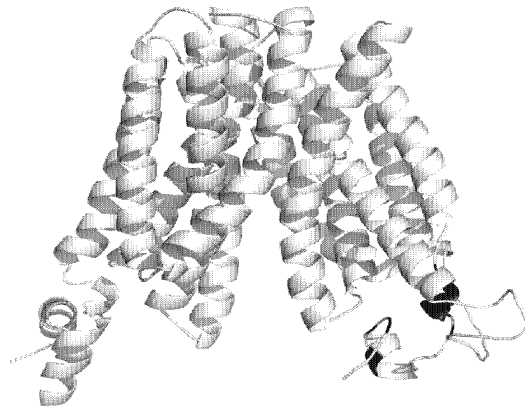


j.

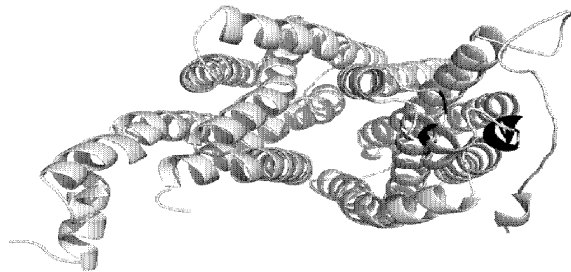


도면29n

k.

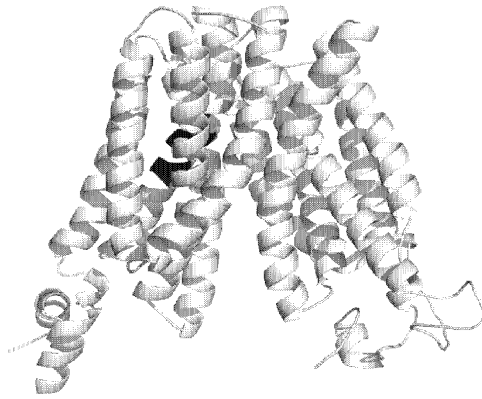


l.

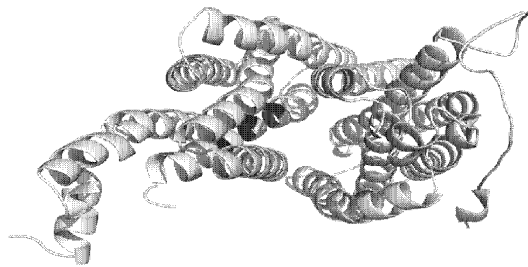


도면290

m.

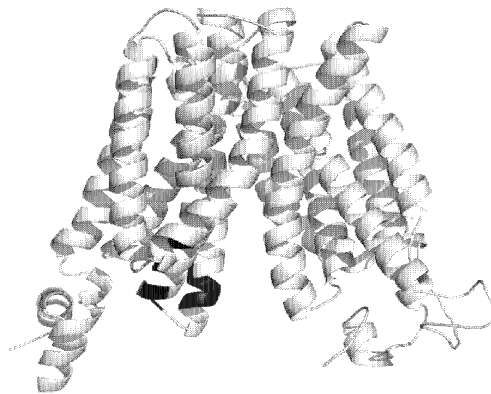


n.

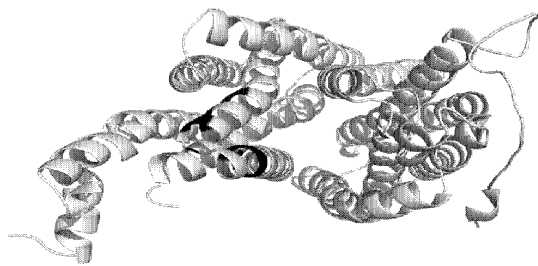


도면29p

o.

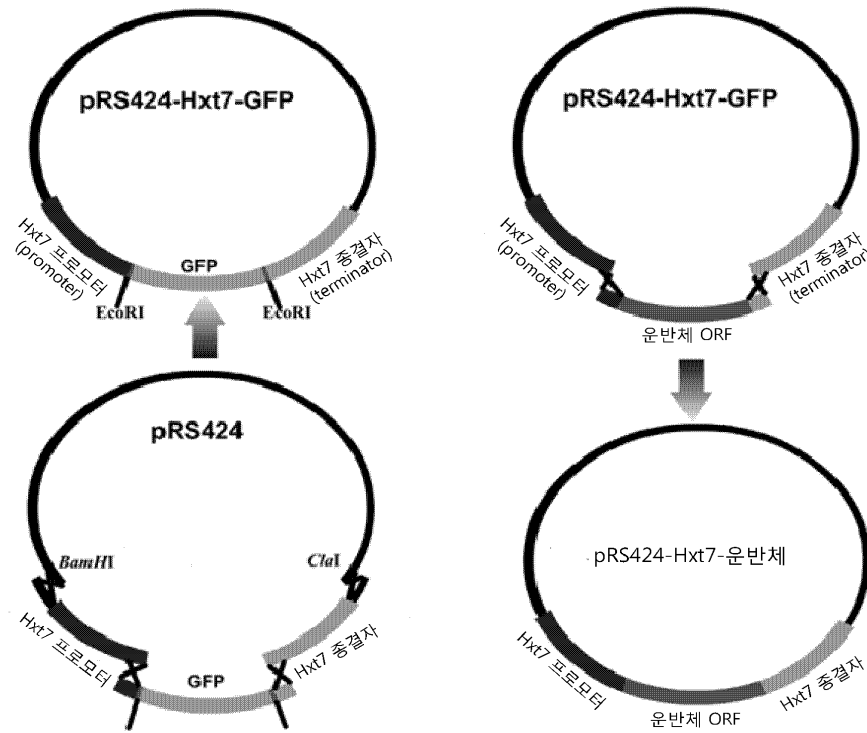


p.

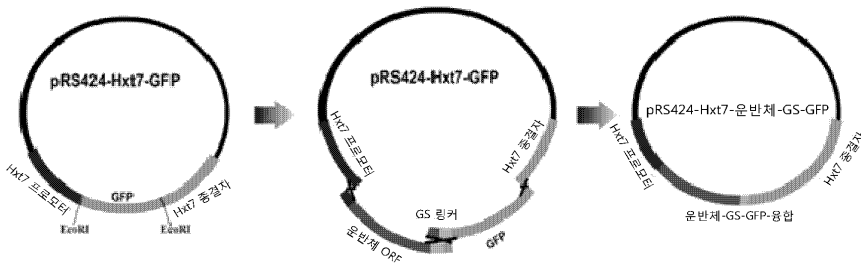


도면30

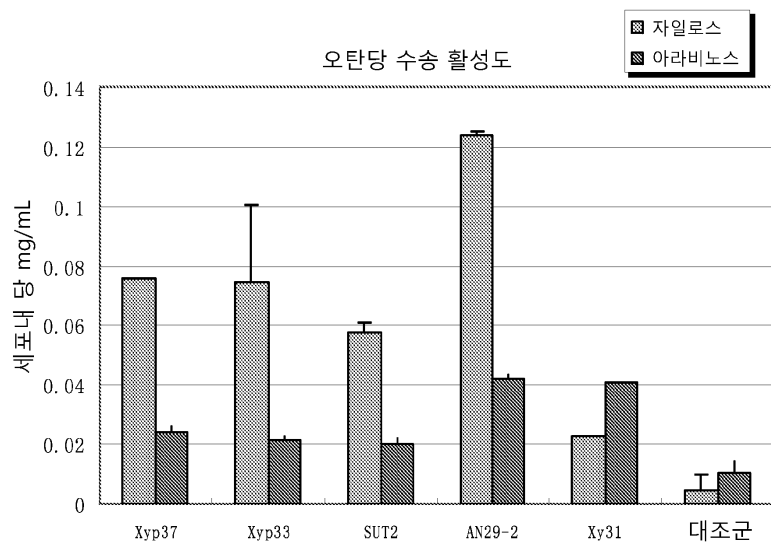
A



B

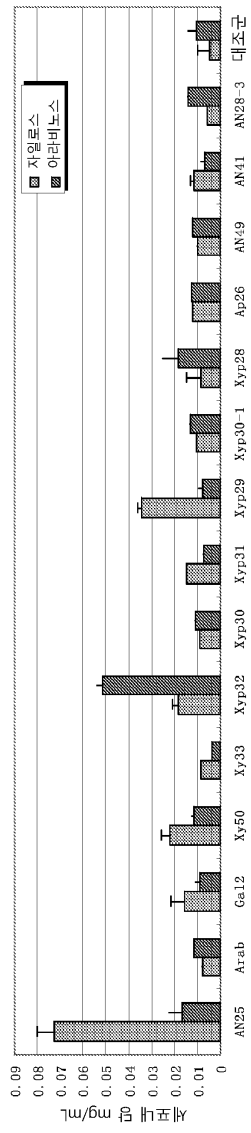


도면31



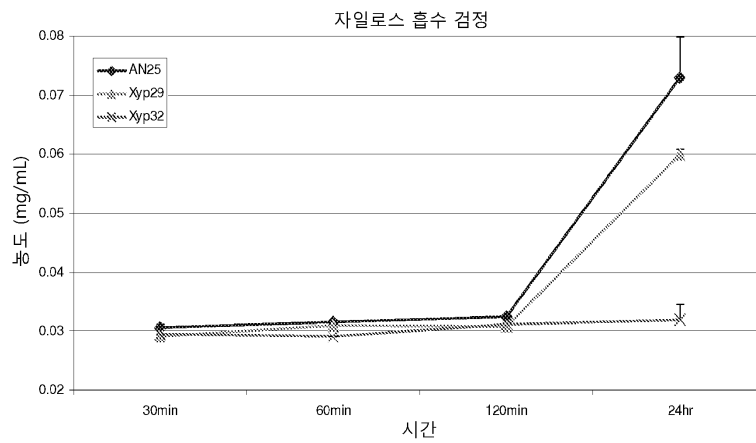
도면32

오탄당 수송 활성도

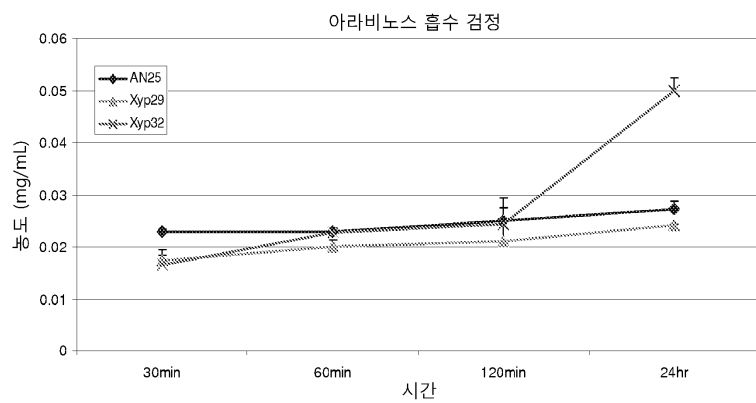


도면33

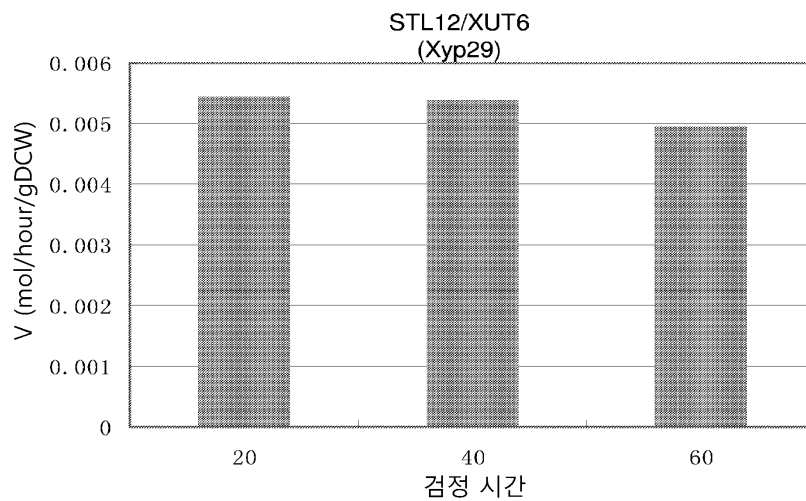
A



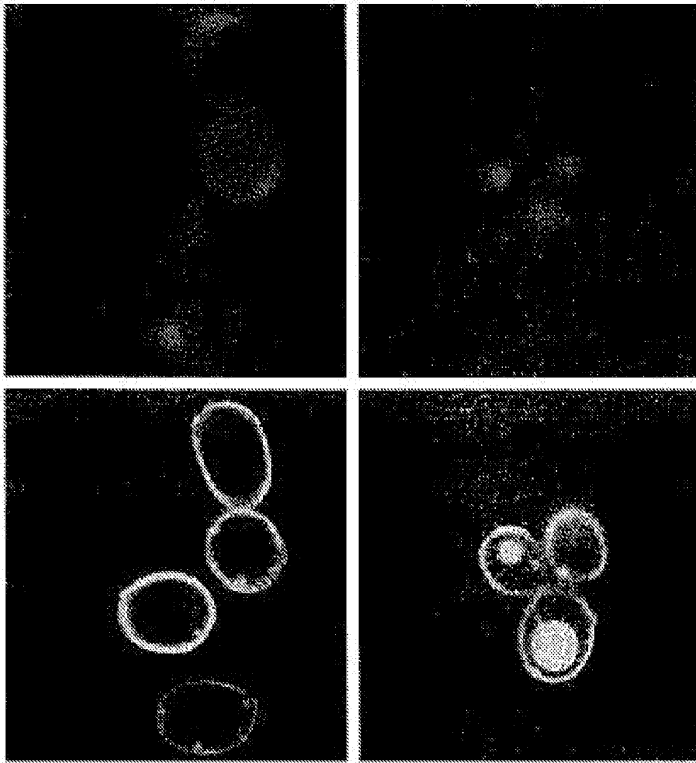
B



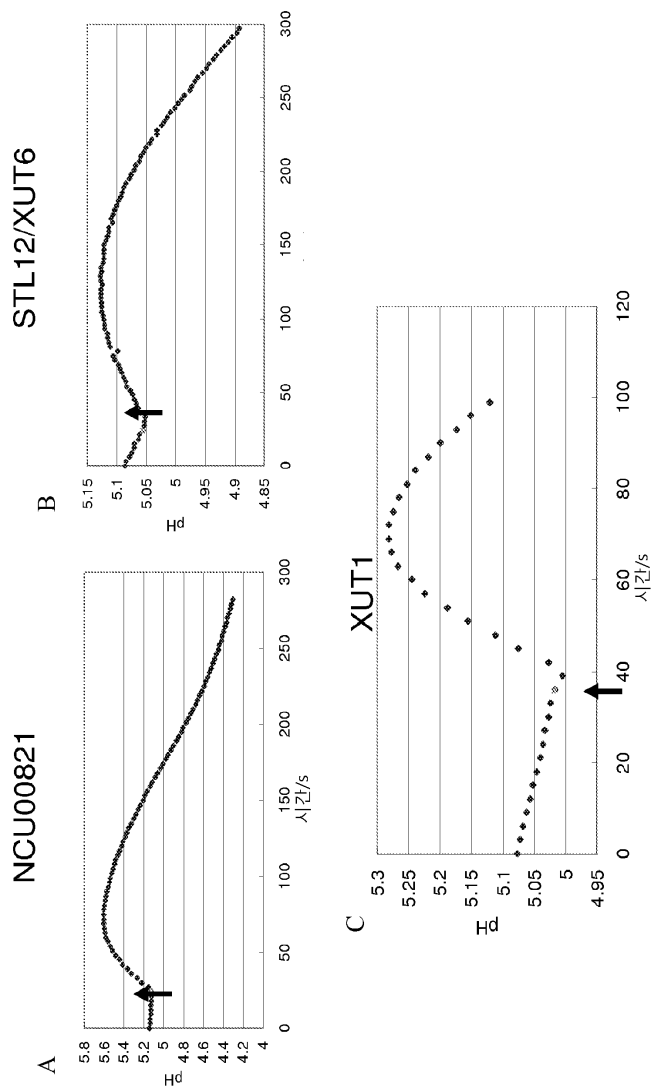
도면34



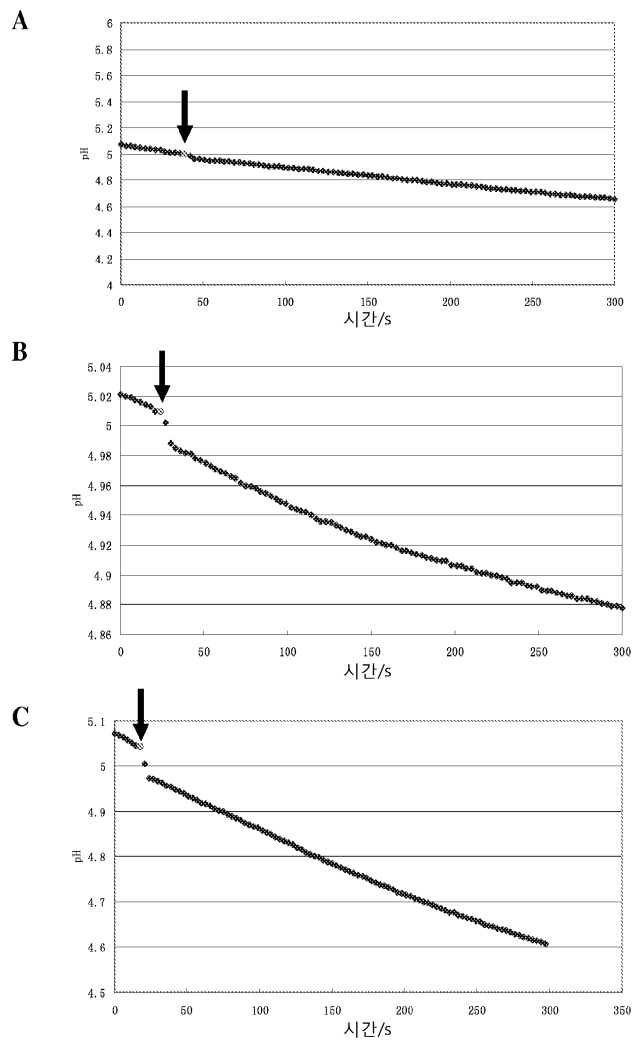
도면35



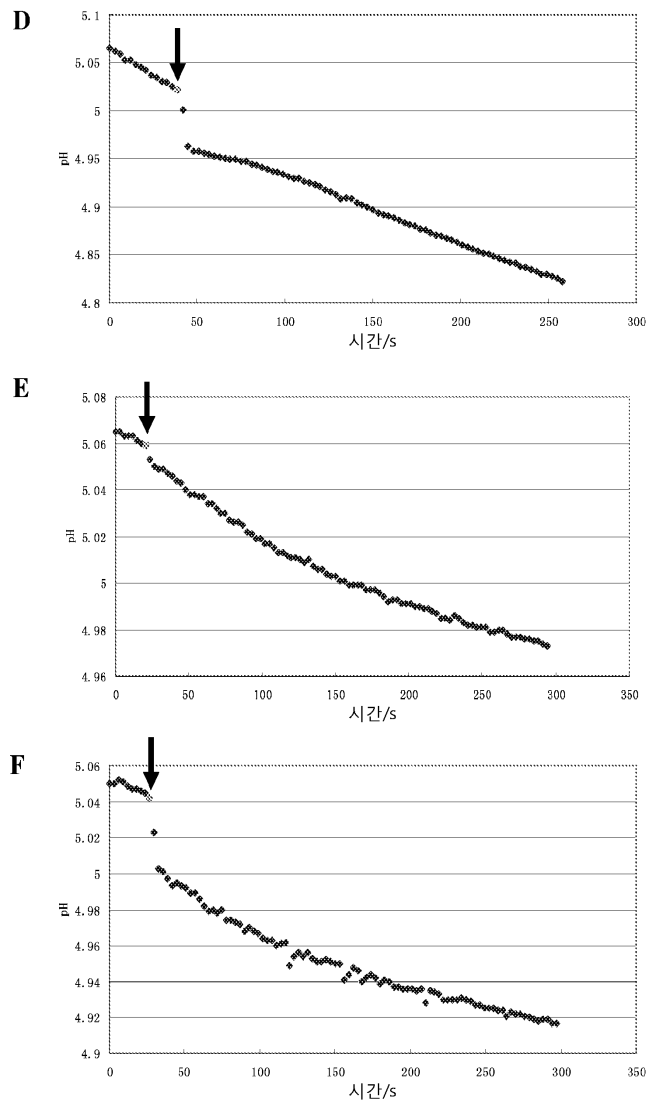
도면36



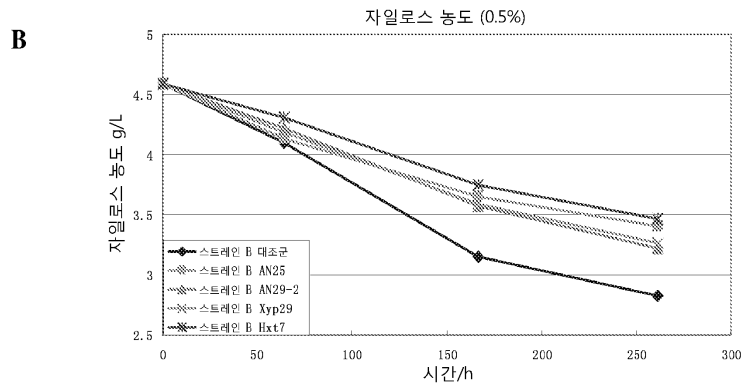
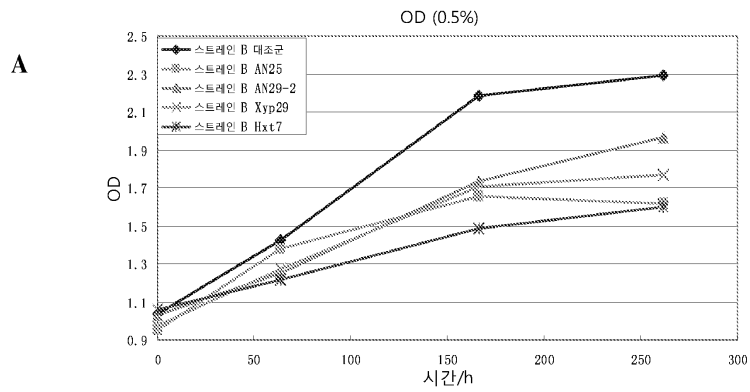
도면37a



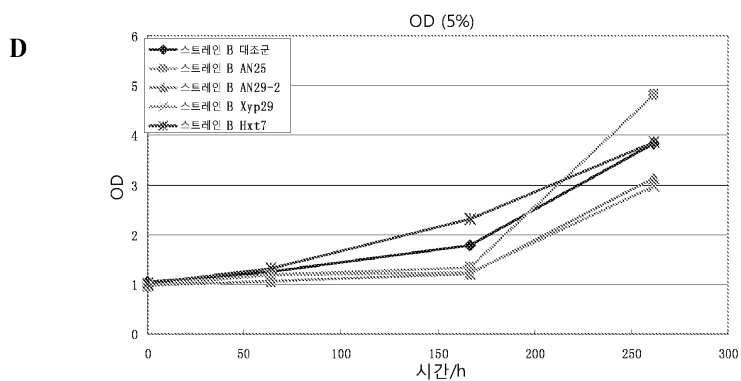
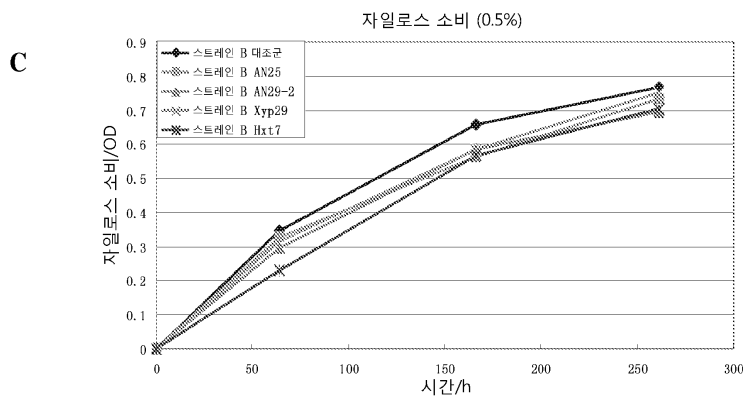
도면37b



도면38a

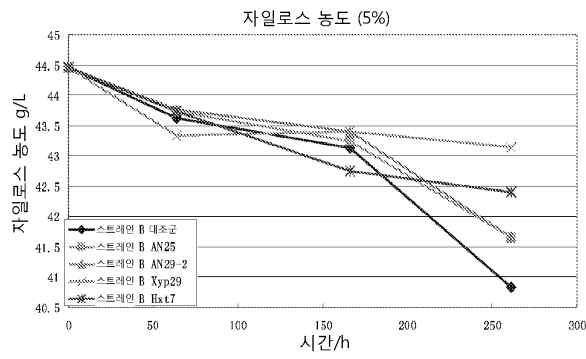


도면38b

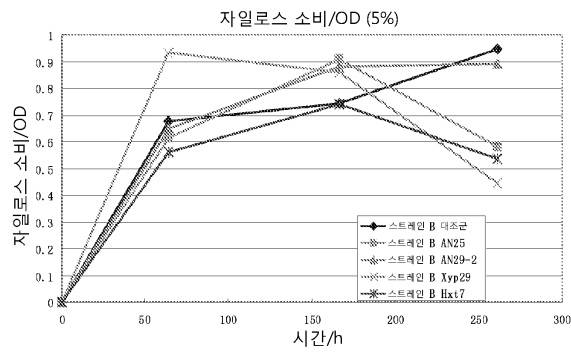


도면38c

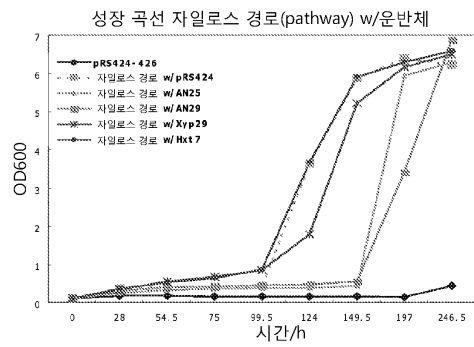
E



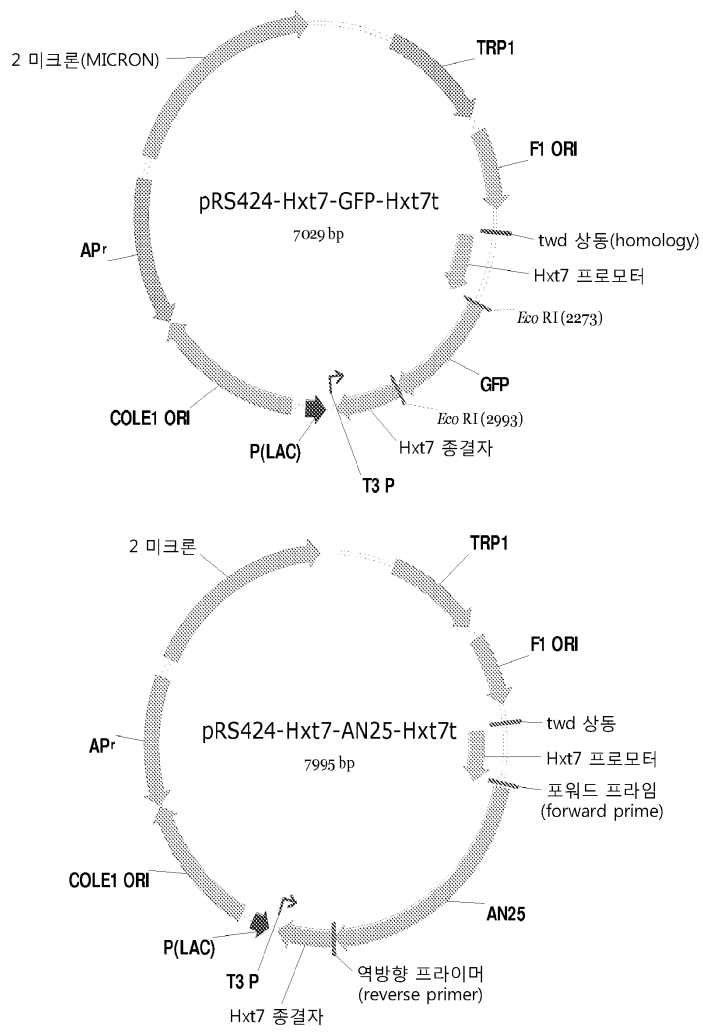
F



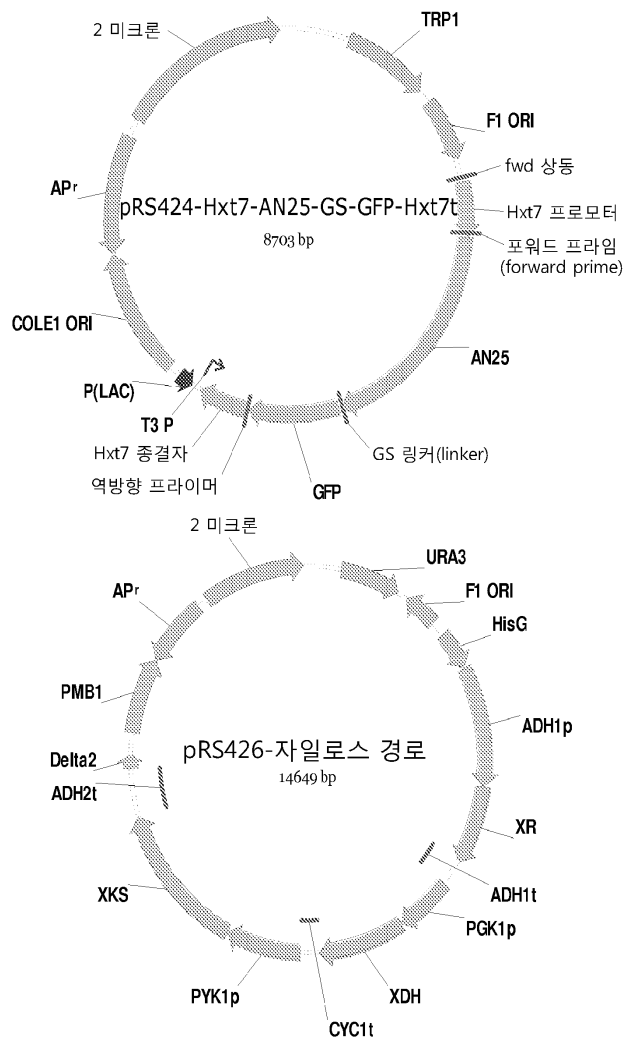
G



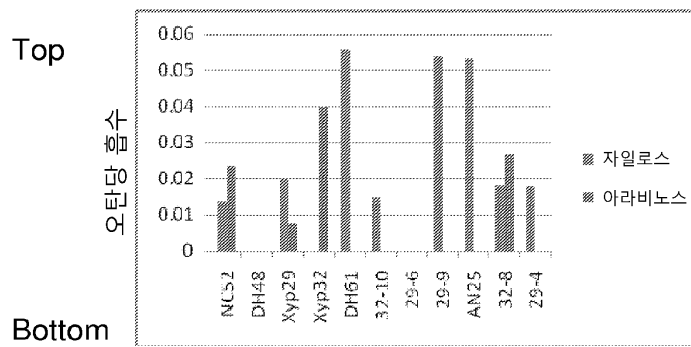
도면39a



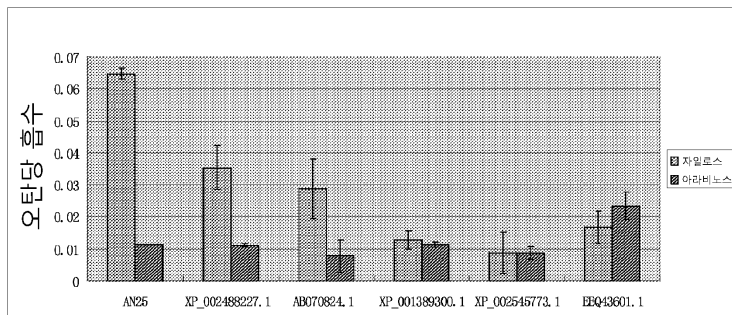
도면39b



도면40



Bottom



도면41a

Consensus key

* - single, fully conserved residue
: - conservation of strong groups
. - conservation of weak groups
- no consensus

CLUSTAL W (1.81) multiple sequence alignment

A

```

NCU00821      -----MAPPKFLGLSGRPLSLAVSTVATTGFLFCYDQGVM
XP_002488227.1 -----MAVPQFAGMSGGRKLSWSYSTIATLGFLLFCYDQGVM
XP_002382573.1 -----MAPTFAGLSGRPLSLAVST-----
XUT6          MEFSSVEKSAETASYTSQVSASGSAKTNSYLGLRGHKLNFAVSCFAGVGFLFCYDQGVM
                  : *: *: *: ::*

NCU00821      SGIITAPAFNNFPTPKD--NSTMQGLITAIYEIGCLIGAMFVLWTGDLGRRRNIMVGA
XP_002488227.1 SGIISAKPFNTVFKATED--DSTMQGTIVTAIYEIGCLFGAMFILWFGDWLGRKSVMLGA
XP_002382573.1 -----
XUT6          GSLLTLPSEFNTFPAMKASNNATLQGAVIALYEIGCMSSSLATIYLGDRLGRLKIMFIGC

NCU00821      FIMALGVIIQVTCQAGSNPFAQLFVGRVVMGIGNMNTSTIPTQAECSTSNRGLLIC
XP_002488227.1 ATMLGVVIQVTAYTGHVPLAQFIVGRVVTGVNGINTSTIPTQAECSTRITNRGLLIC
XP_002382573.1 -----AECSTSNRGLLIC
XUT6          VIVCTGAALQASFT----IAHLTVARIITGLTGFTSTVPVYQSECSPAKRGQLTMM
                  :*** :*,** ** :

NCU00821      EGGVIAFGTLIAYWIDYGASYGPD-----LVWRFPPIAFQLLFAIFICVPMFYLPEsprw
XP_002488227.1 EGGVIAFGTLIAYWIDFGASYGPD-----LTWRFPPIAFQIVFGLIAGMFFLPESPRW
XP_002382573.1 EGGVIAFGTATAYWIDFGAHYGPD-----LVWRFPPIAFQIVFGVITIVGMFFLPDSprY
XUT6          EGSILALGTATSYWIDFGFYFLRNDGLHSSASWRAPIALQCVFAVLISTVFFFPESPRW
                  **:***: *;****:* : :* ** ***:*/*: : :*:***:

NCU00821      LISGRTQEADKVIAALRGYEIDGPETIQERNLIVDSLRA-----SG
XP_002488227.1 LFMREYQEGEAVIAALNEETNSHHVQLQKTLVLDSTRA-----SG
XP_002382573.1 LISDRITQEGEYVLAALGGYEVHDQETQTQKNLYIDSTRAYVCCLEVLNPKSVTNLDNSS
XUT6          LNKGRTEAREVFSAALYDLPADSEKISTQIEETQAIDL-----ER
                  *: :* :*: *;:** . . . : : : :

```

도면41b

[illegible][illegible][illegible]

NCU00821 MYDKIGWGTYLFFAVAGCPLPIYFFPYPETARRSLEEDIFIARG-FVENMSVYTAARE
XP_002488227.1 MISNKGWGYLFFAIPACFLPIYVFWYPETARRSLEEDIFIARG-YTENSIVVRAARE
XP_002382573.1 MIAHIGWGTYLFFAAALFPIPIWLPYPETARRSLEEDIFIARG-YTENSIVVKASRD
XUT6 GFQSIGSYTYLIFAAIALLMAPVIYPLYPETKGRSLEEMDIFNCPVWEPWKVQIARD

```

NCU00821      LPHLTAEETSYANKYGLVDRDSNGGGRNIIDEEKTRDRPDQSDSDSPAIVETDVVDEIIG
XP_002488227.1 LPIYSDEDFERMAITQVGGPAEVPSSG-----EKASAHSEEFATETGTGPKQG
XP_002382573.1 LPKINDEETQKANEYGFGNSTEDPK-----ATAAEYSPSTSE
XUT6          LPTIHSEVLDIEHKNVITKKSRIEIVN-----IS-----
              **      *  *  *  *      *

```

NCU00821 VESGFGDGINTKETR
XP_002488227.1 EEH-----VSKMV-
XP_002382573.1 -----
XUT6 -----

도면41c

B

[illegible]

도면41d

C

NCU00821 -----MAPPKFLG:SGRPLSLAVSTVATTGFLIFGYDQGV
 XP_002488227.1 -----MAVPQFAG:SGRKLWSISTIATLGFLLFGYDQGV
 XP_002382573.1 -----MAPTFAG:SGRPLSLAVST-----
 XUT6 MEFSSEKSAETASYTSQVSASGSAKTSNYLGRHKLNFVSCFAGVGFLLFGYDQGV
 EEQ43601.1 ---MSKGDLEELDIQKLIVEKRLEQSNQSGFAT:KRNKRALGTCLFVSLGGILYGYNQGMF
 XUT1 ---MHGGDGDNDITETIAARRLQTAGKSGVAGLVANSRSFFIAVFASLGGLVYGYNQGMF

: . . .

NCU00821 SGITITAPAFNFFPTPKD---NSTMQGLITAIYETGCLIGAMFVLWTDLLGRRRNIMVGA
 XP_002488227.1 SGTISAKPFNTVFKATED---DSTMGGTVTAIYETGCLFGAMFILWFGDWLGRKSVMLGA
 XP_002382573.1 -----
 XUT6 GSLITLPSFENTFPAMKASNNATLQCAVIALYETGCMSSSLATIYLGDRGLRLKTMFIGC
 EEQ43601.1 GGVSSMHSFGETVIGICKIQDNPTLQGLLTSILELGAWVGVLNMGVADALGRRASVVIGC
 XUT1 GQISGMYSFSKATGVEKIQDNPTLQGLLTSILELGAWVGVLNMGYIADRLGRKSVVVG

NCU00821 FIMALGVITQVTCAGSNPFAQLFVGRVVMGIGNMNTSTIPTYQ:BCSKTSNRGLLICI
 XP_002488227.1 AIMILGVVIQVTAITGHVPLAQFIVGRVVTGVNGINTSTIPTYQ:BCSRITNRGLLICI
 XP_002382573.1 -----ABCSKTSNRGLLICI
 XUT6 VIVICGAALQASFT---IAHLTVARIITGLTGFTITSTVPVYQ:SECPARKRGQLIMM
 EEQ43601.1 ILFNIGVITQAVARDAD---YGYILGGRFVIGLVGVLNMGVADALGRRASVVIGC
 XUT1 FFFFIGVIVQAVARGGN---YDYILGGRFVVGIGVLNMGVADALGRRASVVIGC

:* * . ** :

NCU00821 EGGVIAFGTLIAFYWIDYCASYPDD-----LVWRFPFIAPQLLFAIFICVPMFYLPESPR
 XP_002488227.1 EGGVIAFGTLIAFYWIDYCASYPDD-----LTWRFPFIAPQIVFGIGIAGMFPLPESPR
 XP_002382573.1 EGGVIAFGTLIAFYWIDYCAINYPDD-----LVWRFPFIAPQIVFGVITIVGMFPLPESPR
 XUT6 EGSLIALGTAISYWIDYGFYFLRNDGL-HSSASWRAPIALQCVFAVLITSTVFFPESPR
 EEQ43601.1 YQLSITFGIMISYWITGTNYIGGTGDNQSQASWLVPMCIQAAPAILAVFIYSFPESPR
 XUT1 QQLAITFGIMISYWITGTNYIGGTGSGQSKASWLVFICIQLVPAILLGVGIFFPESPR

:: ***:~* : . * *:~* : : : :~*~*

NCU00821 WLLSHGRTQADKVIAALRGYETDGPETIQERNLIVDSLRA-----S
 XP_002488227.1 WLFMRERYQEGEAVIAALLNEETNSHHVQLGKTLVLDSTRA-----S
 XP_002382573.1 YLISKDRITQEGEYVLAALGGYEVHQBETQTKNLVIDSIRAYVCCLEVLPNKSVYTNLDS
 XUT6 WLLNKGRTQEARVFSALYDLPAISEKISTQIEEIQAAIDL-----E
 EEQ43601.1 WLLNVGQSEKALEVLAALRETEQENVGLQIEFLEMKAQKIFEQTEATEAYP-----H
 XUT1 WLLNEDRENECLSVLSNLSLSKEDTLVQMEFLEMKAQKIFERELSAKYFP-----H

:* : : : ~*~* . . : :

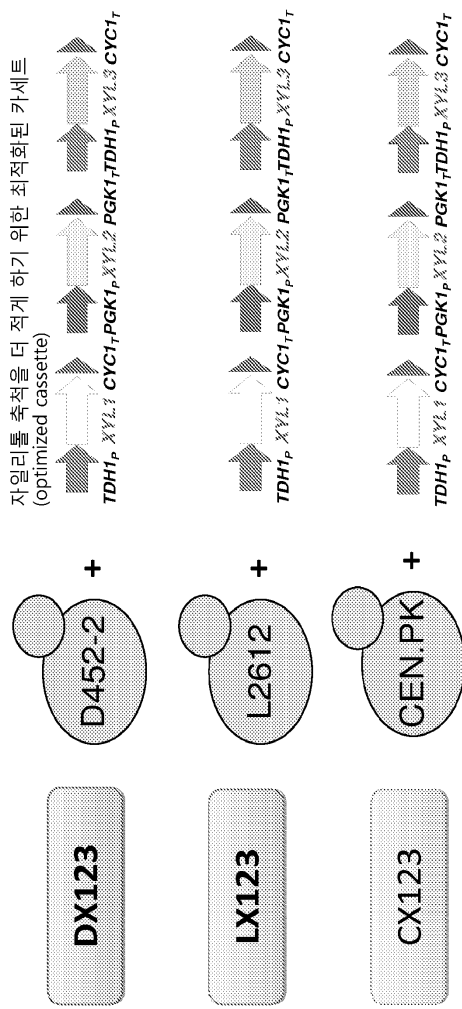
도면41e

NCU00821	GGFGQKSTPFKALFTGCK-----TQHFRRLIGSSSQFQKQVGG-CAAVITYYPTLFEQS
XP_002488227.1	GQMG-KSTPLSAVFTGCK-----TQHFRRLIGSSSQFQKQVGG-CAAVITYYPTLFENS
XP_002382573.1	SGAG-ATTRYRDLITGGR-----SQHFRRLIGSSSQFQKQVGG-CAAVITYYPTLFEQS
XUT6	RQAG-EGFVLKELEFTGCP-----ARNLQRVAISCWSQIMQITG-INITYYAGTIFESY
EEQ43601.1	LQDGTKMSKFKINLQYKSMVTHLPTFKRVSVACLTMVQQWTGAYAFILYYAPPIFASL
XUT1	LQDGSAKSNFLIGFNQKSMITHYPTFKRVAVACLMTQQTG-VAFILYYAPPIFESSL
	* : : : : : ** * * : ** : :
NCU00821	IGESHNMSMLGGINMIVYSIFATMSWFLIERVGRRLFLIGIVQMLSMV-IVFACLIP
XP_002488227.1	IGQTHITMSMLGGVNMIVYSIFATMSWFLIERVGRRLFLIGIVQCGSMI-LTFACLIP
XP_002382573.1	LHQSNDEALLIGGINMIVYAFATPSWFFIERKIGRRKFLIGGSIQTIAMV-YTFACLIP
XUT6	IGMSPFMSRLAALNGTEYFLVSLIAFYTVERRLGRPFLEWGALAMALYMAGLTVTVKLA
EEQ43601.1	GLSGNTTSLASGVGVIVMFLCTIPAVMWVVKYGRKPLISGALVMGLCHFVYAGLGGY
XUT1	GLSGNTISLASGVGVIVMFLATIPAVLWVDRLGKRPVLISGALIMGICHPVYAILGQF
	: :
NCU00821	DD--PMKARG---AAVGLITYIAPFGATWLPPLWYPAEYNPRTKGANAVSTCSNWF
XP_002488227.1	GT--PGPAKG---AAVGLITYIASFGATWLPPLWYPAELSPKTRAKANALSTCTNWF
XP_002382573.1	DD--TQVSKG---AVFGLFLYMAAFGAANLPLPWYPAELSPKTRAKANAVSTCSNWF
XUT6	GEGNTHAGVG---AAVLLEAFNSFFGVSWLGGSWLPPLELSHLRAPGAALSTASNAF
EEQ43601.1	SDNIGSHKAAGWAVVFIFIFAGAPCVSWGPCAWYVAEVFPLGMRAGVSLGSSPNWLM
XUT1	GGNFVNISGAGWAVVFIFIFATGFCVSWGPCAWVLVAEVFPLGLRAKGVSLGASSNWLN
	. : : : : * : : : : * : : : : * : : : : *
NCU00821	NRLIVAVTPIMVDKICWGTYLFFAVNGCFLPTLYFYPETANESLEEDITIFAKG-FVE
XP_002488227.1	NRLIVAVTPVMISNKGWTYLFFAIVNACPLPVLYWYPETANESLEEDITIFAKG-YTE
XP_002382573.1	NFTYVAVTPVMIAHICWGTYLFFAALNALPIPLWLYPETANESLEEDITIFAKG-YTE
XUT6	NFMVVMITPVGPGSGSYTYLFFAATNLLMAPVLYLYPETKGRSLEEDITIFNQCPVWE
EEQ43601.1	NFSYATSTPKFVANAKYGATYPLGLACVIGSMYVYVMVPETKNKTLDELDEYFG-----
XUT1	NFAVAMSTPDFVAKAKPGATYPLGLACIFGAAYVQFCPETKGRTLDELDELFG-----
	** : : * * : : : : : : : : : : : : : : : : : : : *
NCU00821	NMSYVTAARELPHLTAGEIESYANKYGLYDRDSNGEGNRHDEEKTRDRPDQSDSDSPA
XP_002488227.1	NTSYVRAARELPYLSDEYVERMATQYGFPAEVPSDSG-----EKASAHSEEF
XP_002382573.1	NTSYVKASDLPKLNDEHEEQKANEGYFGNSTEDPEK-----ATAAEYSPST
XUT6	PWKVVQIARDLPIHSEVLDHEKNVLIKKSRTIEHVEN-----IS---
EEQ43601.1	--DFTGTSNKESEIREKTLKQVGLVDLLVGSOKELDSF-----RSKPE
XUT1	--DTSGTSMKEKETHEKKEVGLLLQG--EENASESE-----NSKAD
	. : : : : : : : : : : *
NCU00821	VEIDVVDHGVESGFGDGINTKETR
XP_002488227.1	AETTGTPKQGEH-----VSKMV-
XP_002382573.1	SE-----
XUT6	-----
EEQ43601.1	VEYKEKAHSE-----
XUT1	VYHVEK-----

도면42

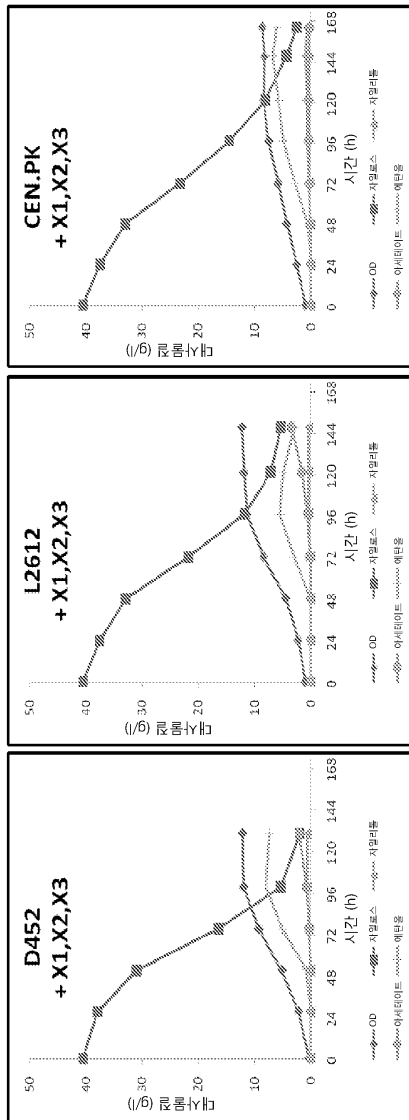
스트레인 백그라운드(backgrounds) vs. 자일로스 발효

- 동일한 구조물(construct) 및 상이한 실험(laboratory) 스트레인을 사용한 유전자 조작된(engineered) 스트레인의 구성(construction)

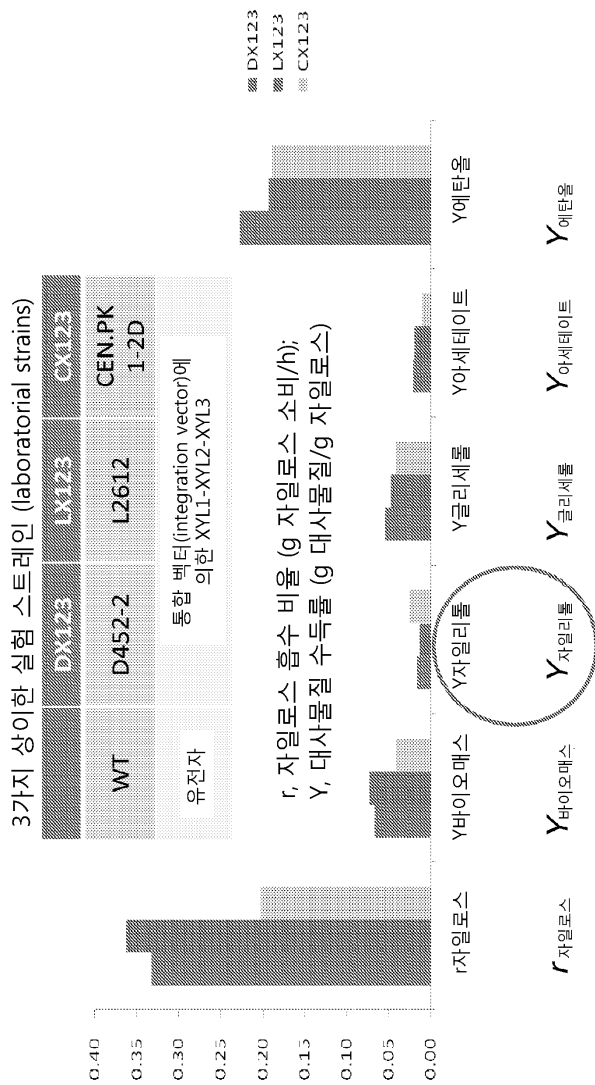


도면43

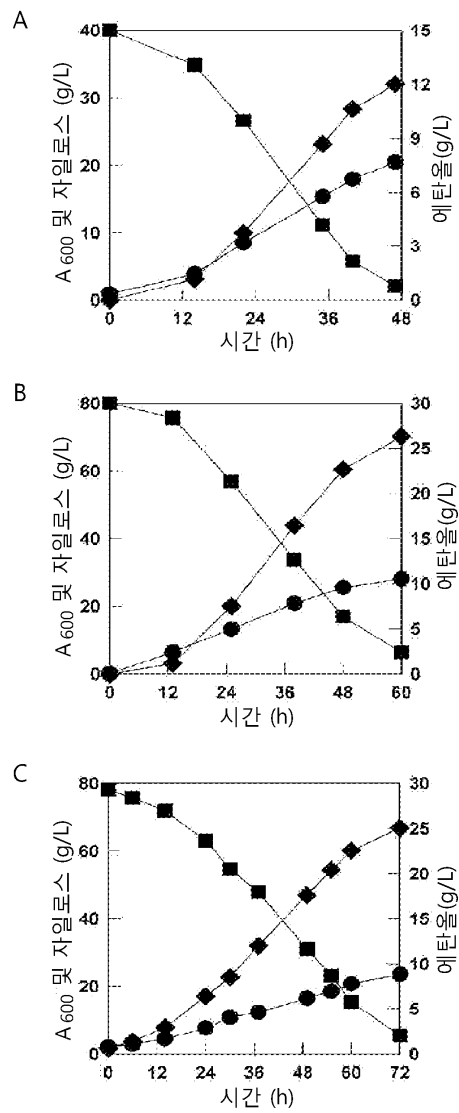
상이한 스트레인 백그라운드(backgrounds) 하에서 동일한 발현 카세트(expression cassette)



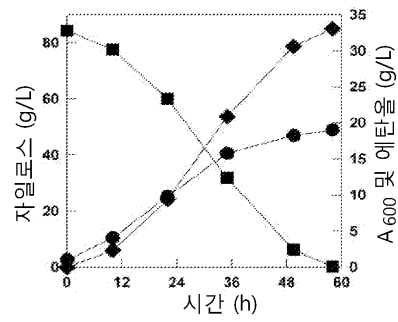
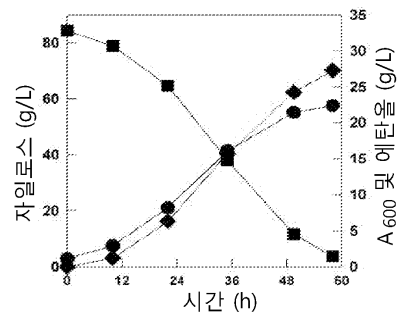
도면44



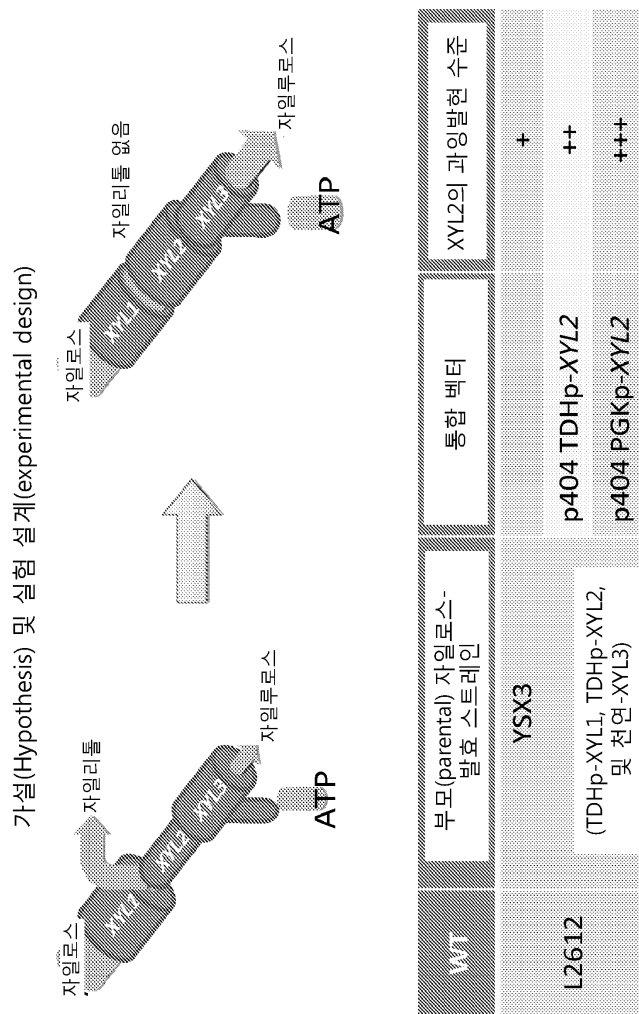
도면45



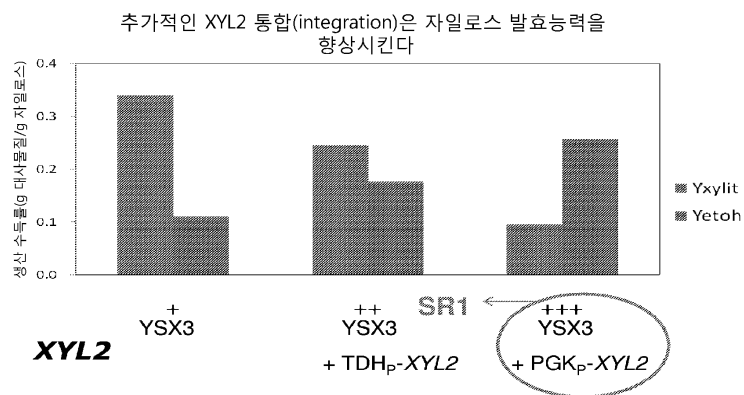
도면46



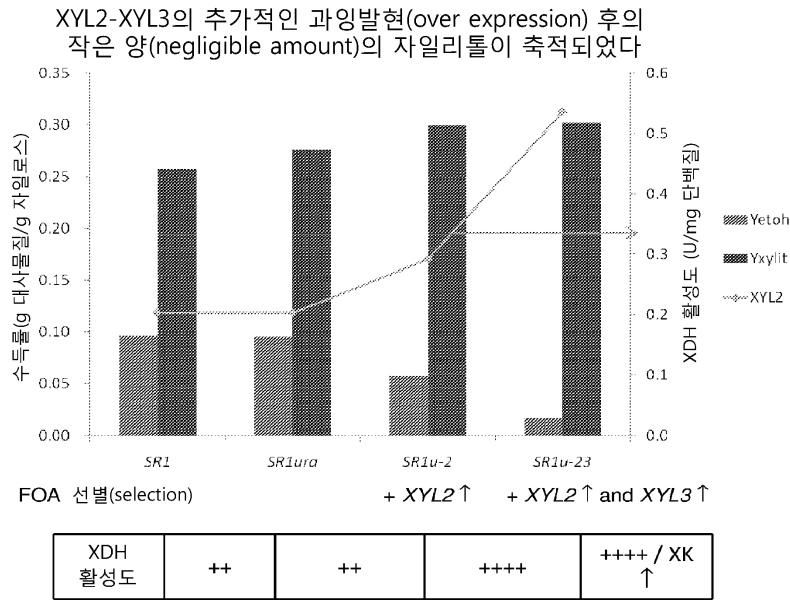
도면47



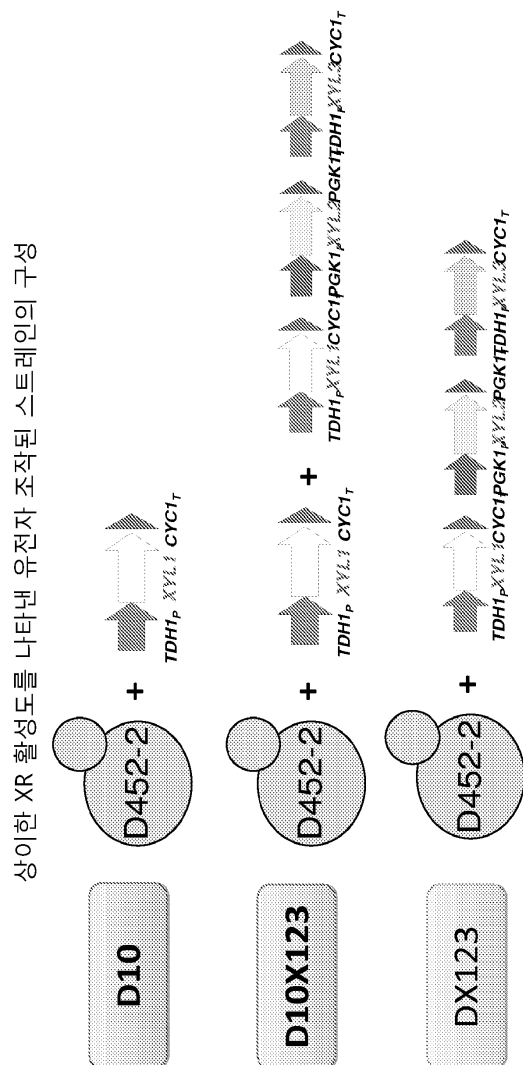
도면48



도면49

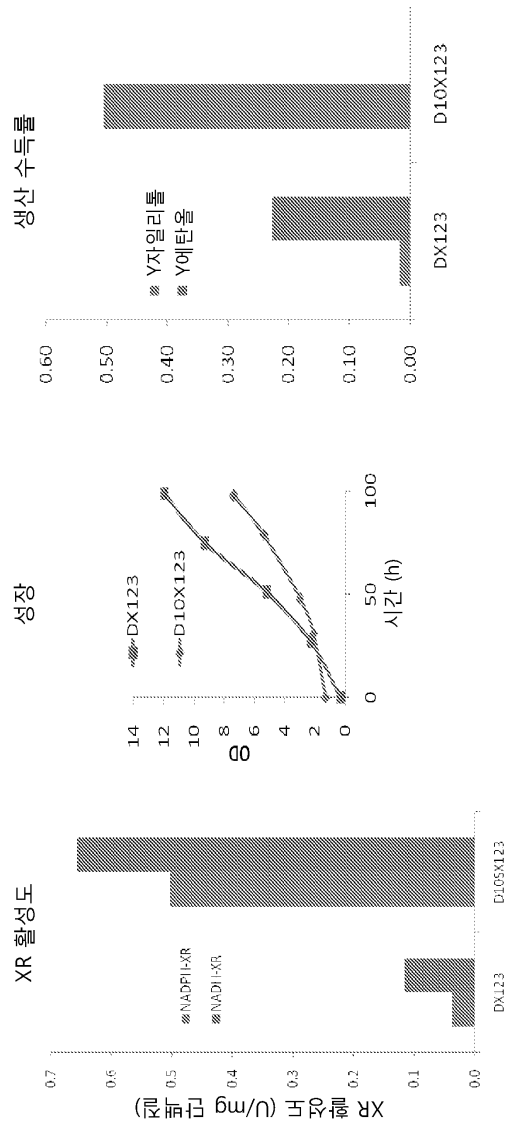


도면50

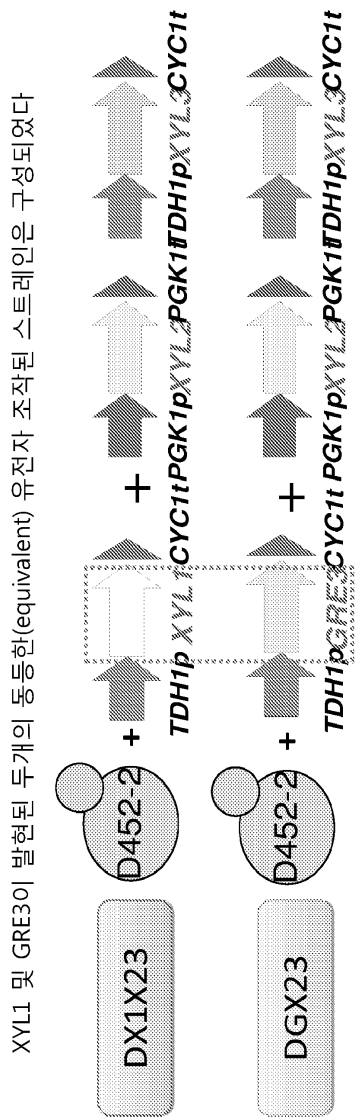


도면51

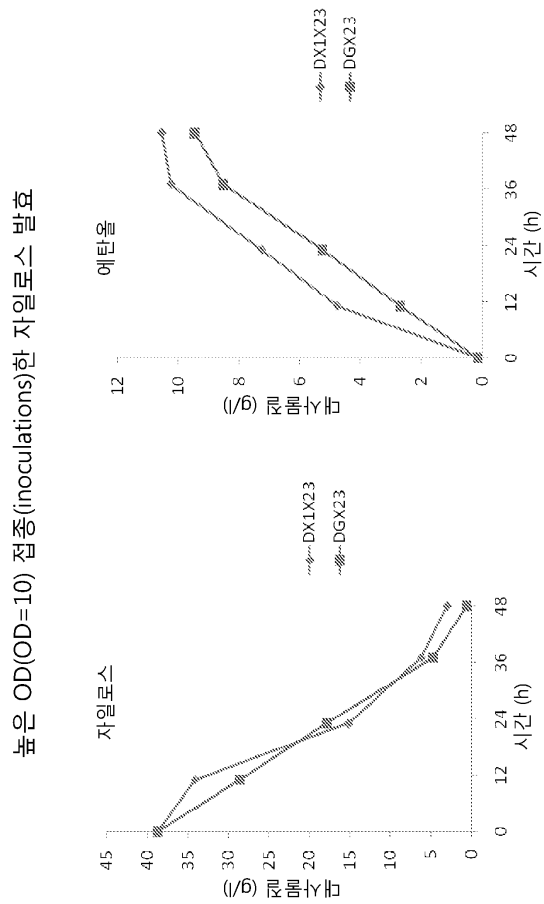
XYL2/XYL3을 갖는 강한(strong) XYL1은 자일리톨 축적을 초래한다



도면52

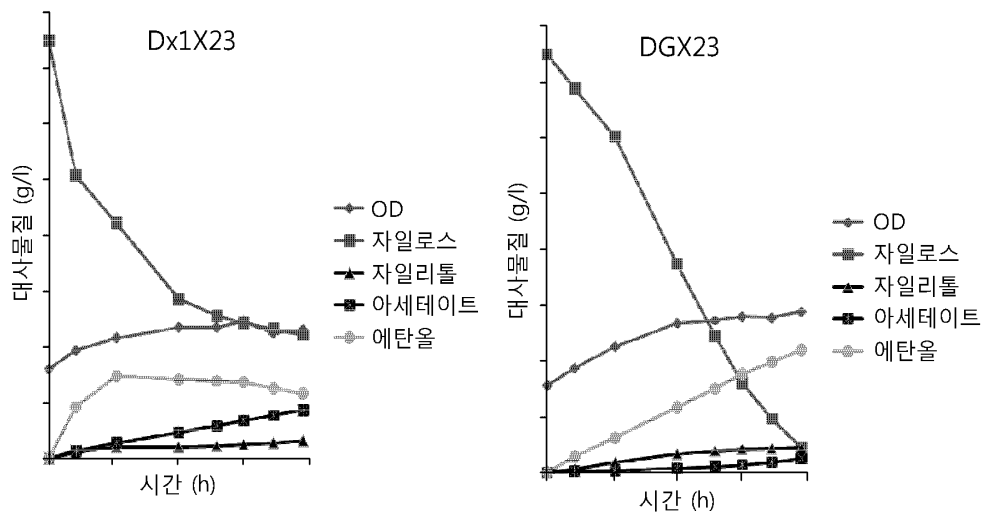


도면53

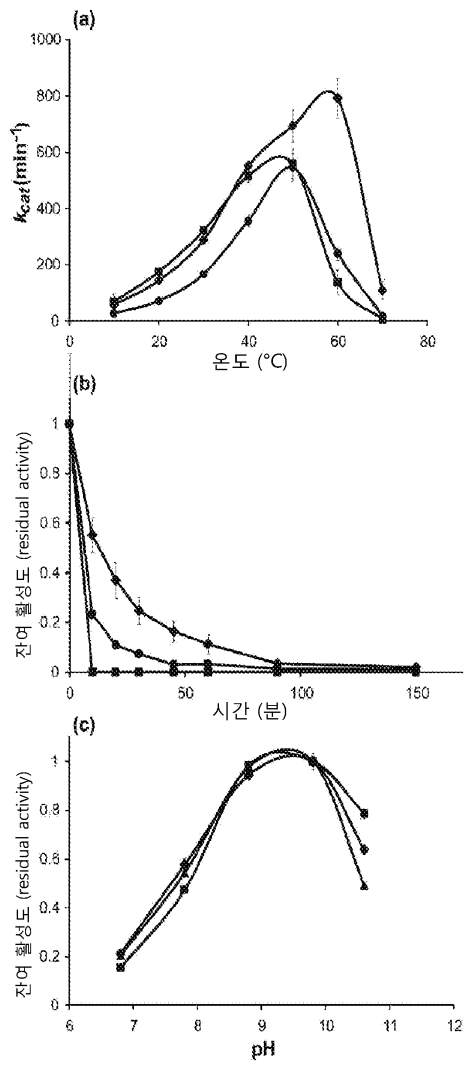


도면54

높은 OD(OD=15) 접종한 80g/L 자일로스 발효



도면55



도면56

Score = 299 bits (765), Expect = 3e-79, Method: Compositional matrix adjust.
Identities = 169/365 (46%), Positives = 222/365 (60%), Gaps = 22/365 (6%)

```

psXDH ANPSVLNKKIDDISFETYDAPEISEPTDVLVQVKKTGICGSDIHFYAHGRIGNFVLTKPM 62
+N S VLNK D- F+ P+I+ P DVLV V TGICGSD+H-+ HG IG+FY+ PM
ncXDH SNLSFVLNKLPLDVCFQDKPVPKINSPhDVLVAVNYTGICGSDVHYWLHGAIGHFVVKDPM 66

psXDH VLGHEsAGTIVVQVGKVTSLKVGDNVAIEPGIPSRFSDEYKSGHYNLCPhMAFAATPNSK 122
VLGHEsAGT+V VG V +L VGD VA+EPG P R SGHYNLCp M FAATP
ncXDH VLGHEsAGTIVAVGDVAVKTLsvGDRVALEPGYPCRRCVHCLSGHYNLCPEMRFAATPPYD 126

psXDH EGEPNPPGTLCKYKSPEDFLVKLPDHVSLELGALVEPLSVGVHASKLGSVAFGDYVAVF 182
GTL -+ +P DF KLP+ VSL+ GAL+EPL+V VH +K + G V V
ncXDH -----GTLTGFWTAPADFCYKLPETVSLQEGALIEPLAVAVHITKQAKIQPGQTVVVM 179

psXDH GAGFVGLLAAAVAKTFGAGKVIVVDIFDNKLKMAKDIGAATHTFNSKTGSSEE-----LI 237
GAGFVGLL AAVAK +GA V+ VDI +KL+ AK AATH+ S+ EE +
ncXDH GAGFVGLLCAAVAKAYGASKVVSVDIVPSKLEFAKSF-AATHYLSQRVSPEENARNIIA 238

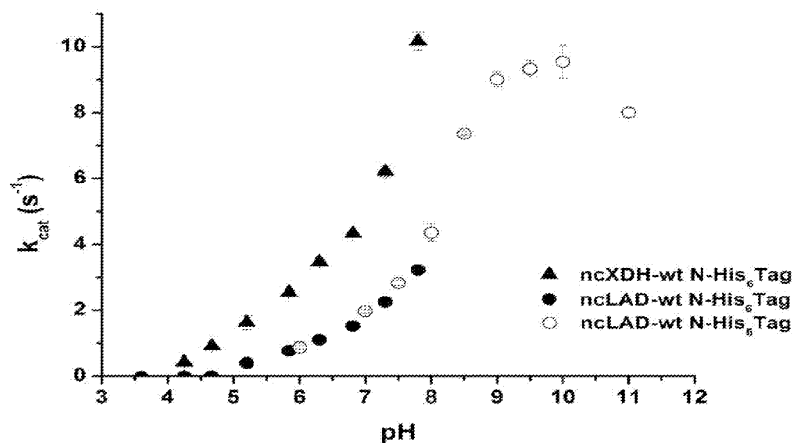
psXDH KAFGGNVPNVVECTGAEPCKLGVDAIAPGGRFVQVGNAAGPVSFPITVFAMKELTLFG 297
A G +V+- +GAEP I+ + + GG +VQ G +-FPI +KE+T G
ncXDH AADLGEgADAVIDASGAEPsIQALHVVVRQGGHYVQGGMGKDNITFPIMALCIKEVTASG 298

psXDH SFRYGFNDYKTAVGIFDTNYQNGRENAPIDFEQLITHRYKFKDAEAYDLVRAGKGAVKC 357
SFRYG DY+ A- + E +D ++L+ FK+A EA+ V+ G+ +K
ncXDH SFRYSGDYRLAQLV-----EQGKVDVKKLVNGVVPFKNAEEAFKKVKEGE-VIKI 349

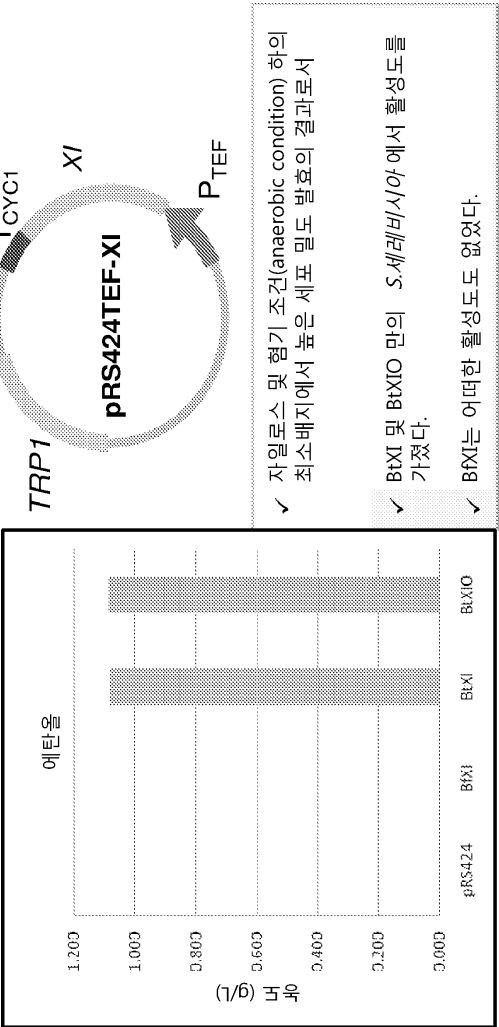
psXDH LIDGP 362
LI GP
ncXDH LIAGP 354

```

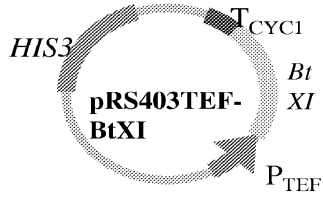
도면57



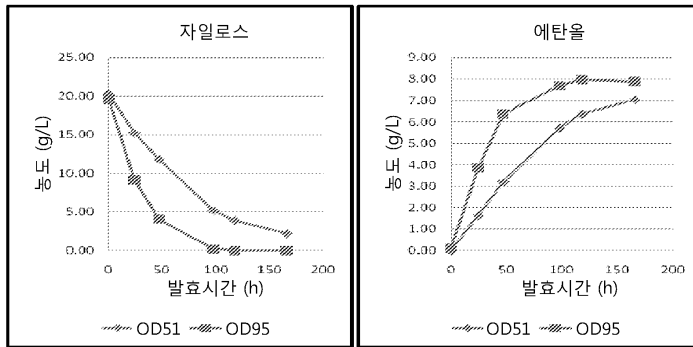
도면58



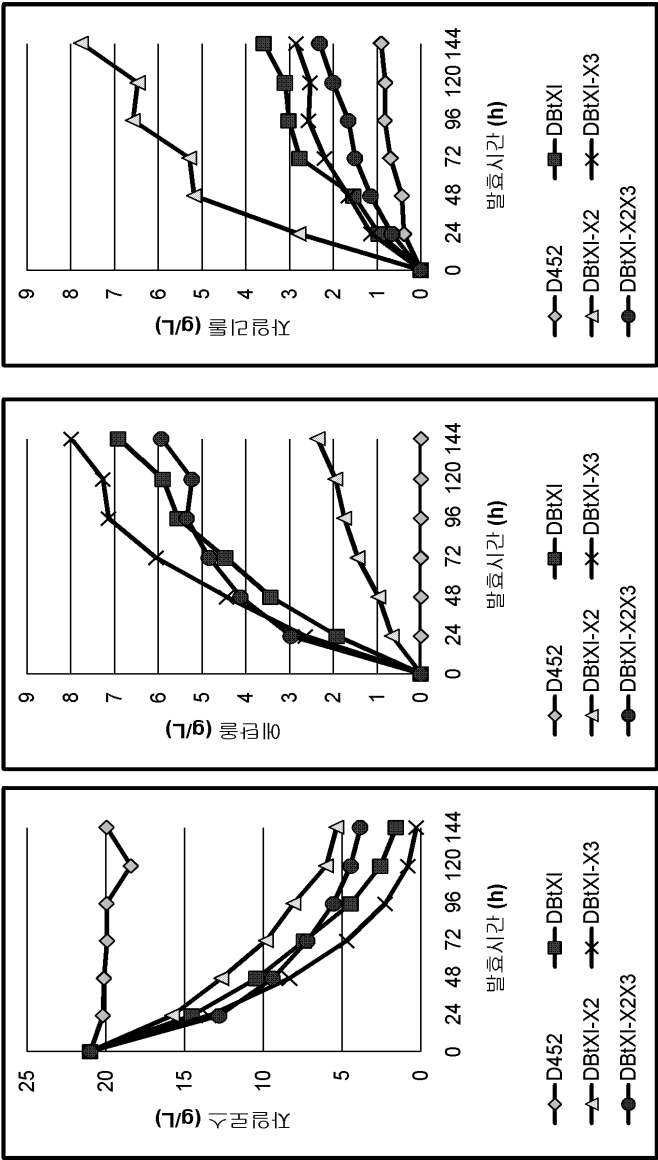
도면59



	플라스미드(L2612)		통합(integration) (D452-2)	
	OD 48	OD 96	OD 51	OD 96
에탄올(g/L)	4.92	5.15	7.67	7.88
수득물(EtOH)	0.35	0.34	0.40	0.40
생산성 (g/L-hr)	0.04	0.04	0.05	0.06
자일리톨(g/L)	3.62	3.61	2.80	2.74
수득물(Xt)	0.26	0.24	0.15	0.14



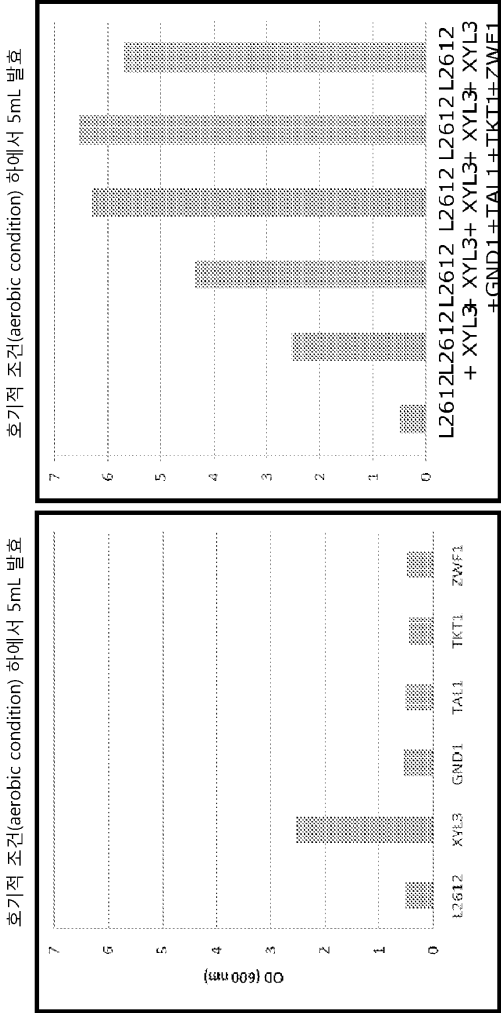
도면60



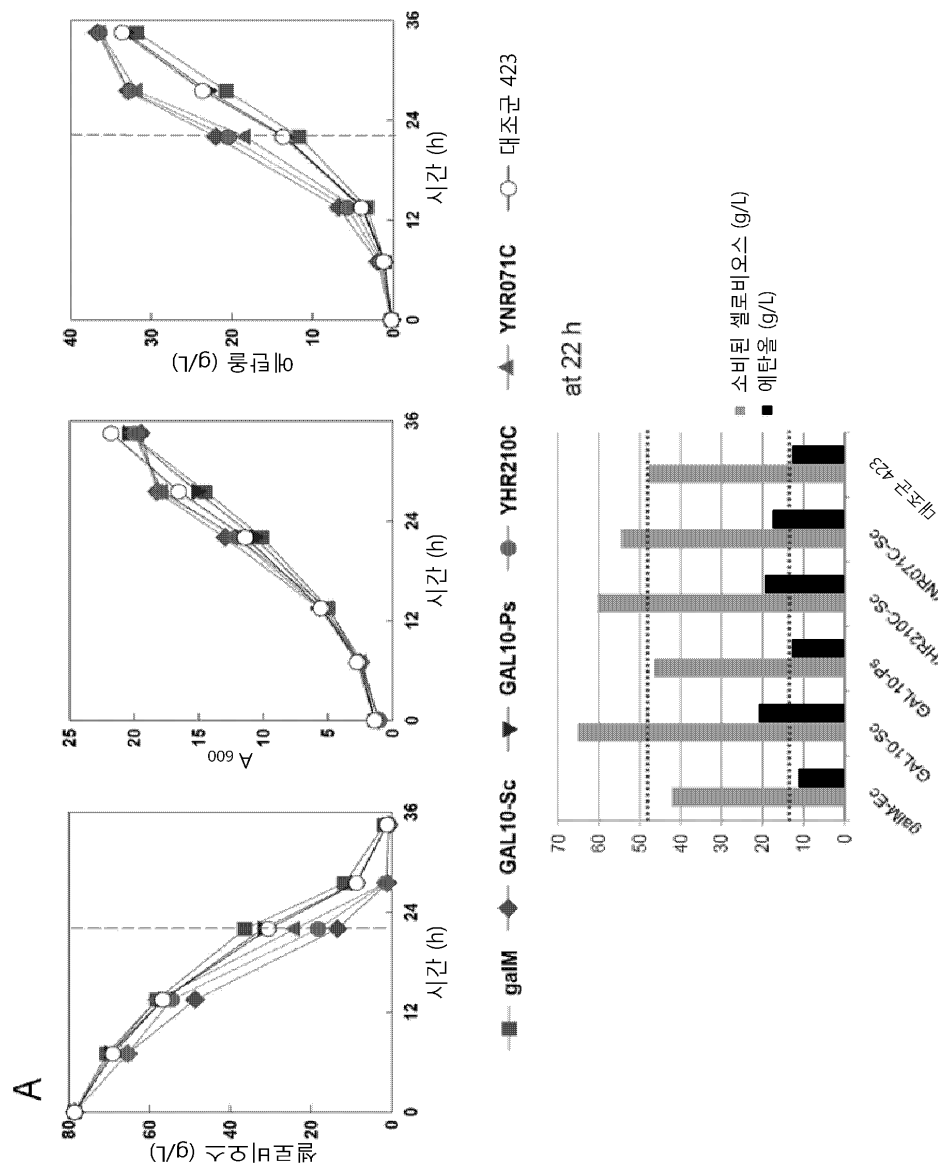
	DBtXI	DBtXI+X2	DBtXI+X3	DBtXI+X23
Y_{EtOH}	0.34	0.15	0.41	0.37
Y_{XOH}	0.18	0.50	0.14	0.12

도면61

자일루로스에서 세포 성장



도면62a

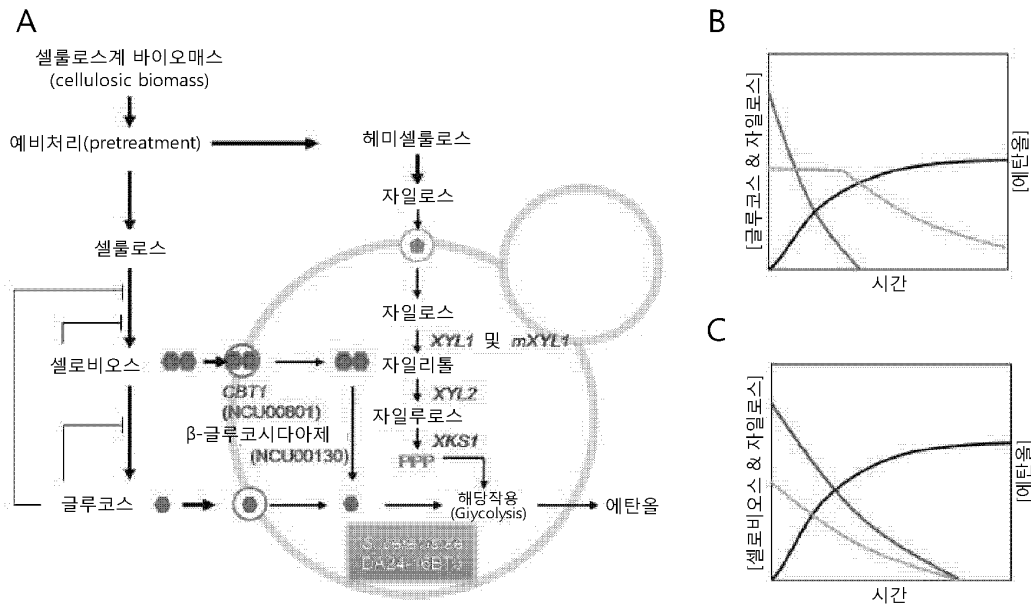


도면62b

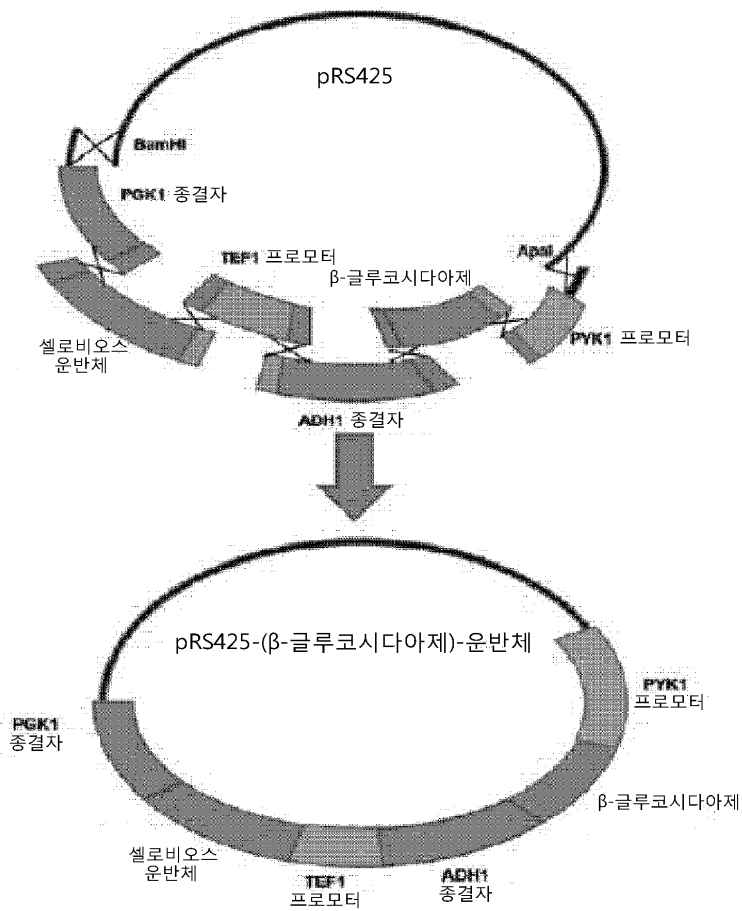
B

	OD (A ₆₀₀)	에탄올 (g/L)	Y _{eth} (g/g)	P _{eth} (g/L-hr)
galM-Ec	20.4	30.3	0.40	0.88
GAL10-Sc	19.5	34.0	0.45	1.01
GAL10-Ps	20.5	31.7	0.41	0.92
YHR210C-Sc	20.0	34.7	0.45	1.01
YNR071C-Sc	20.0	34.6	0.45	1.00
대조군-423	21.8	32.0	0.42	0.93

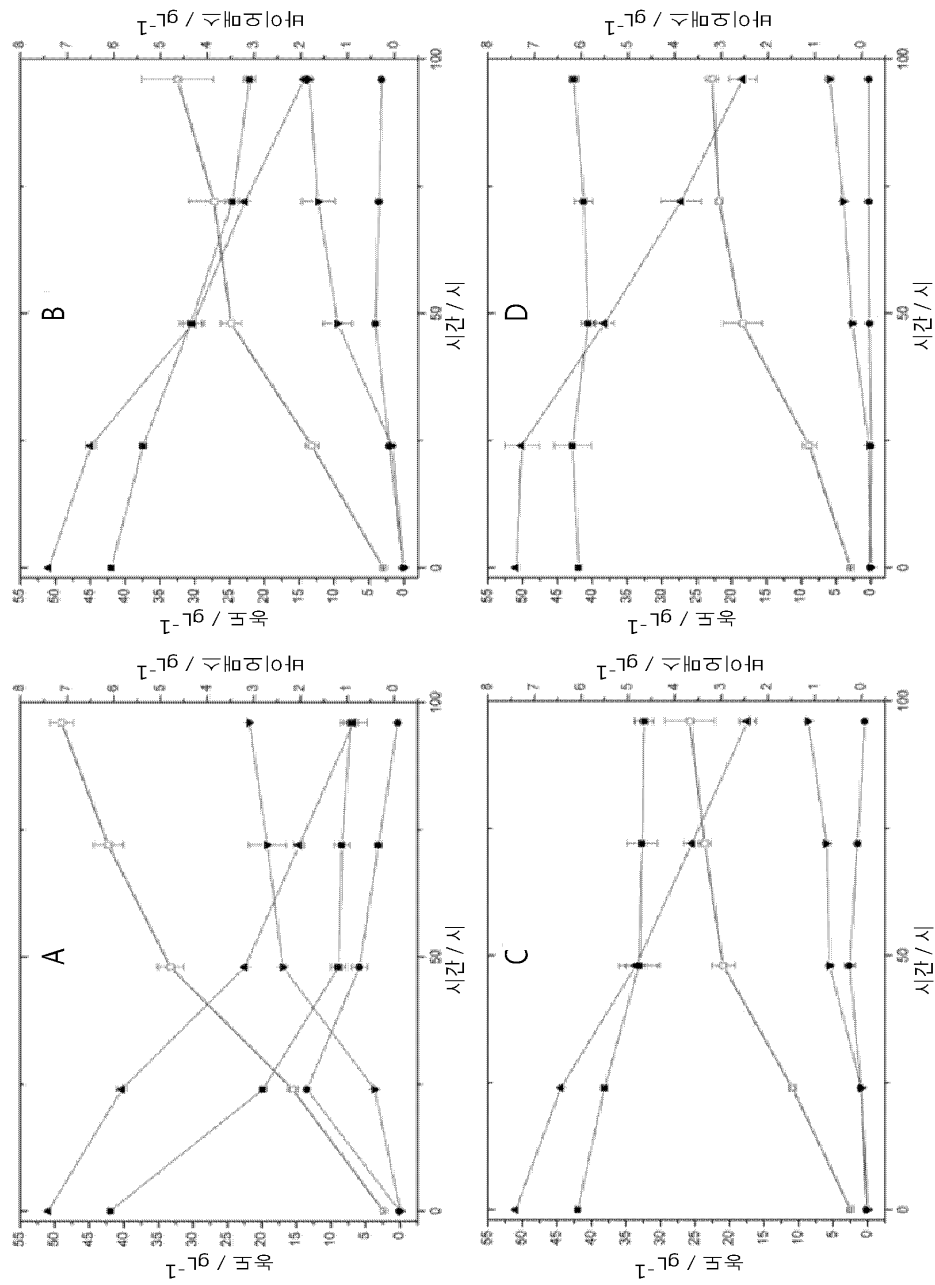
도면63



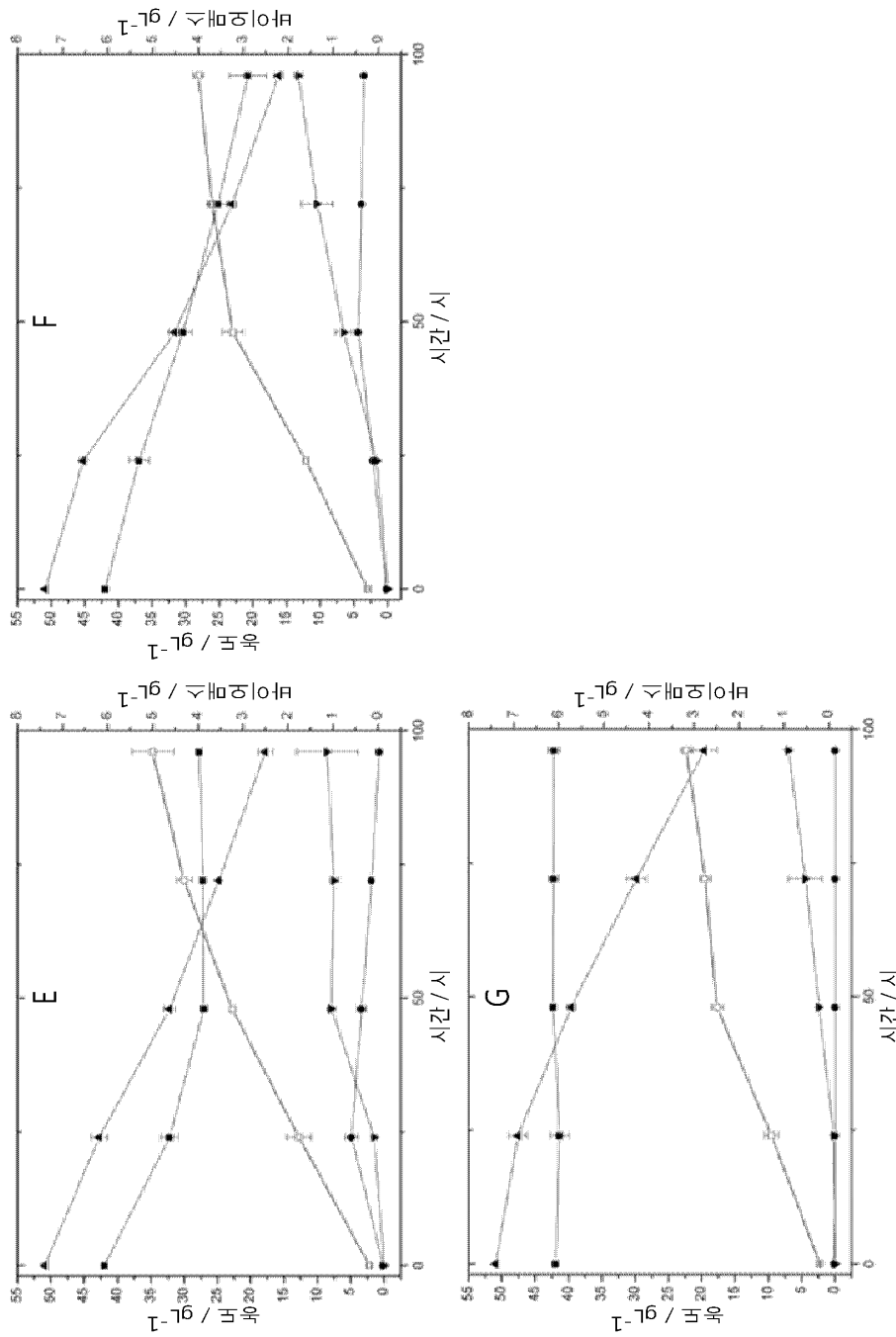
도면64



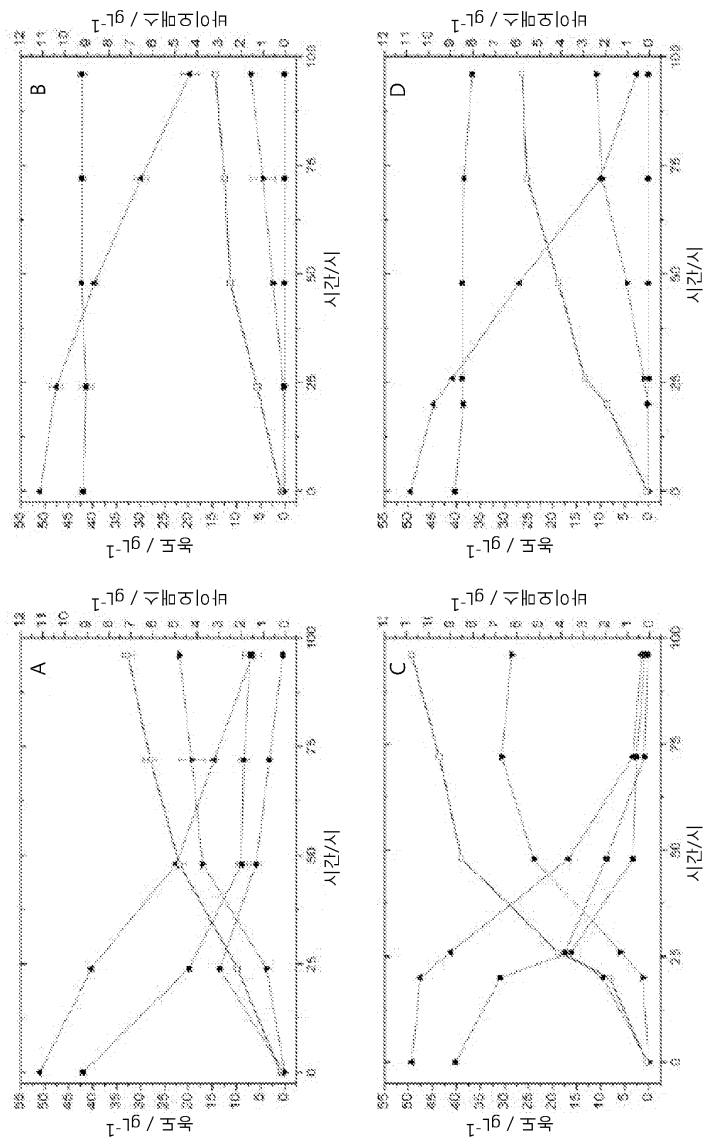
도면65a



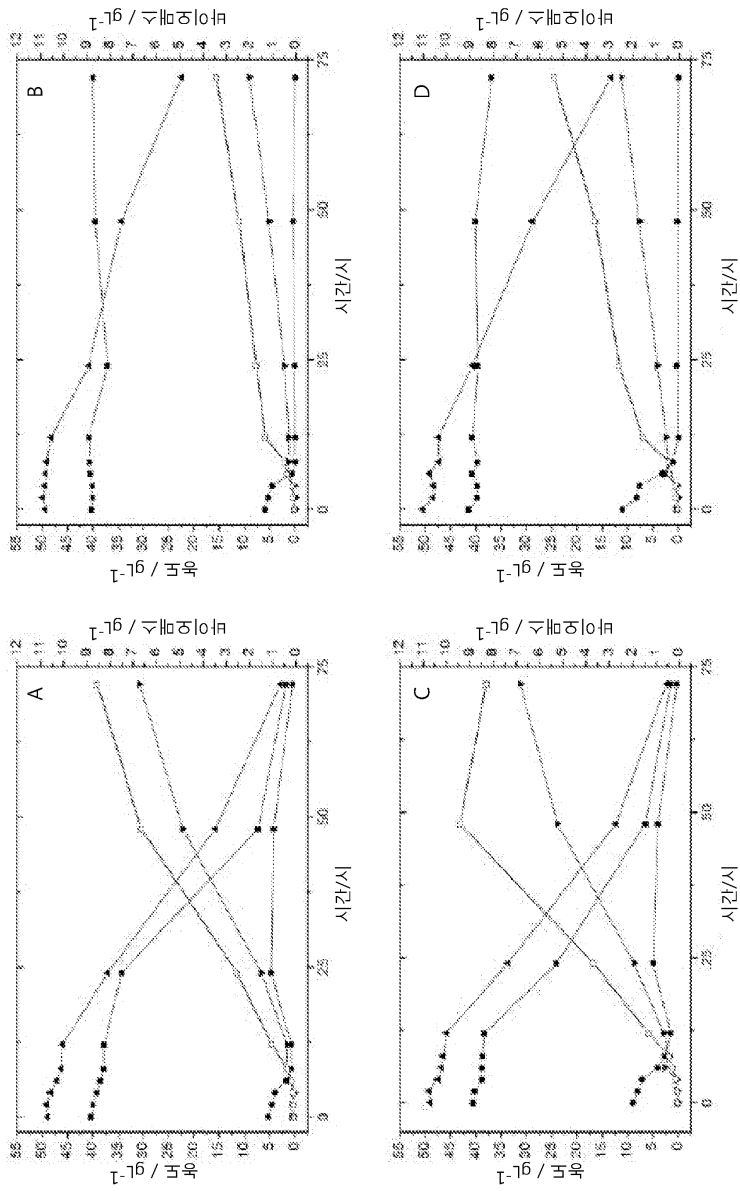
도면65b



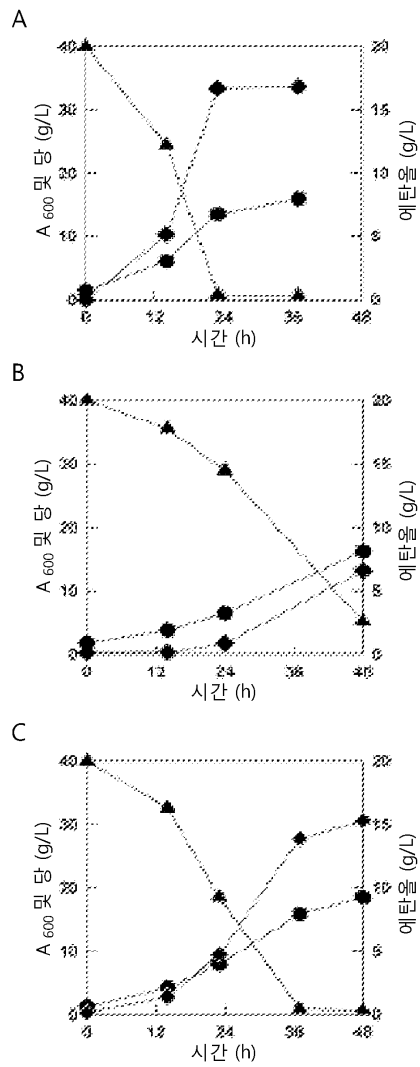
도면66



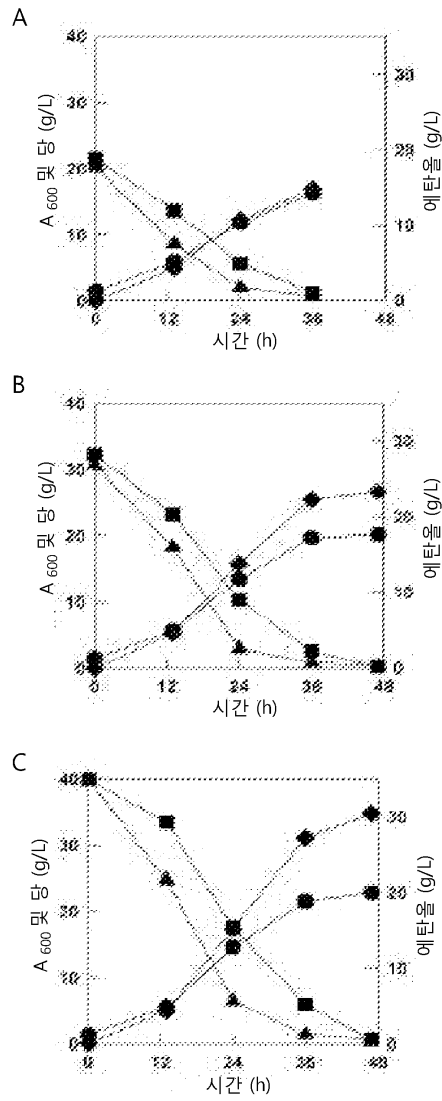
도면67



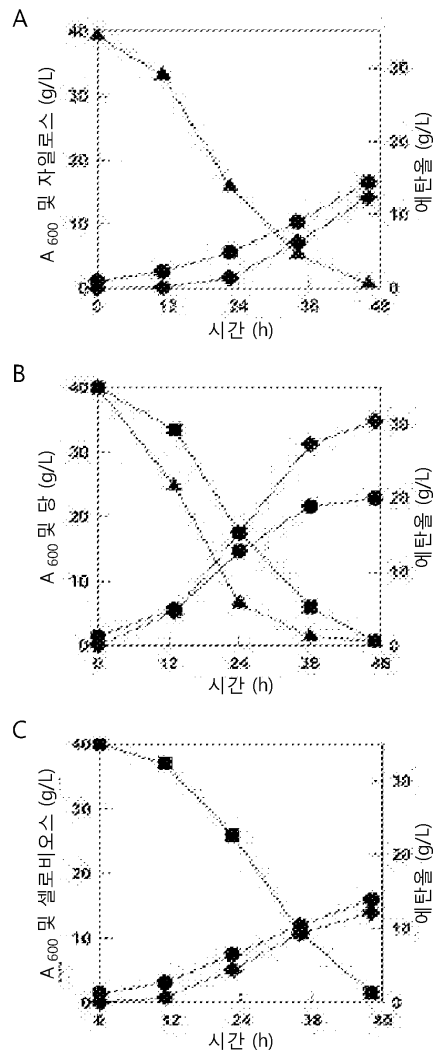
도면68



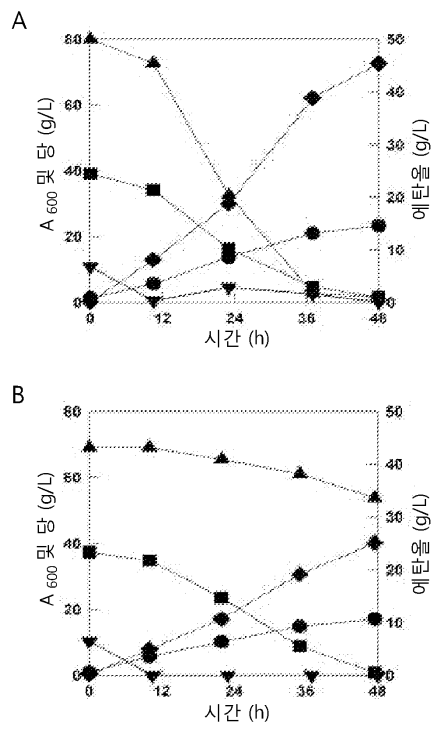
도면69



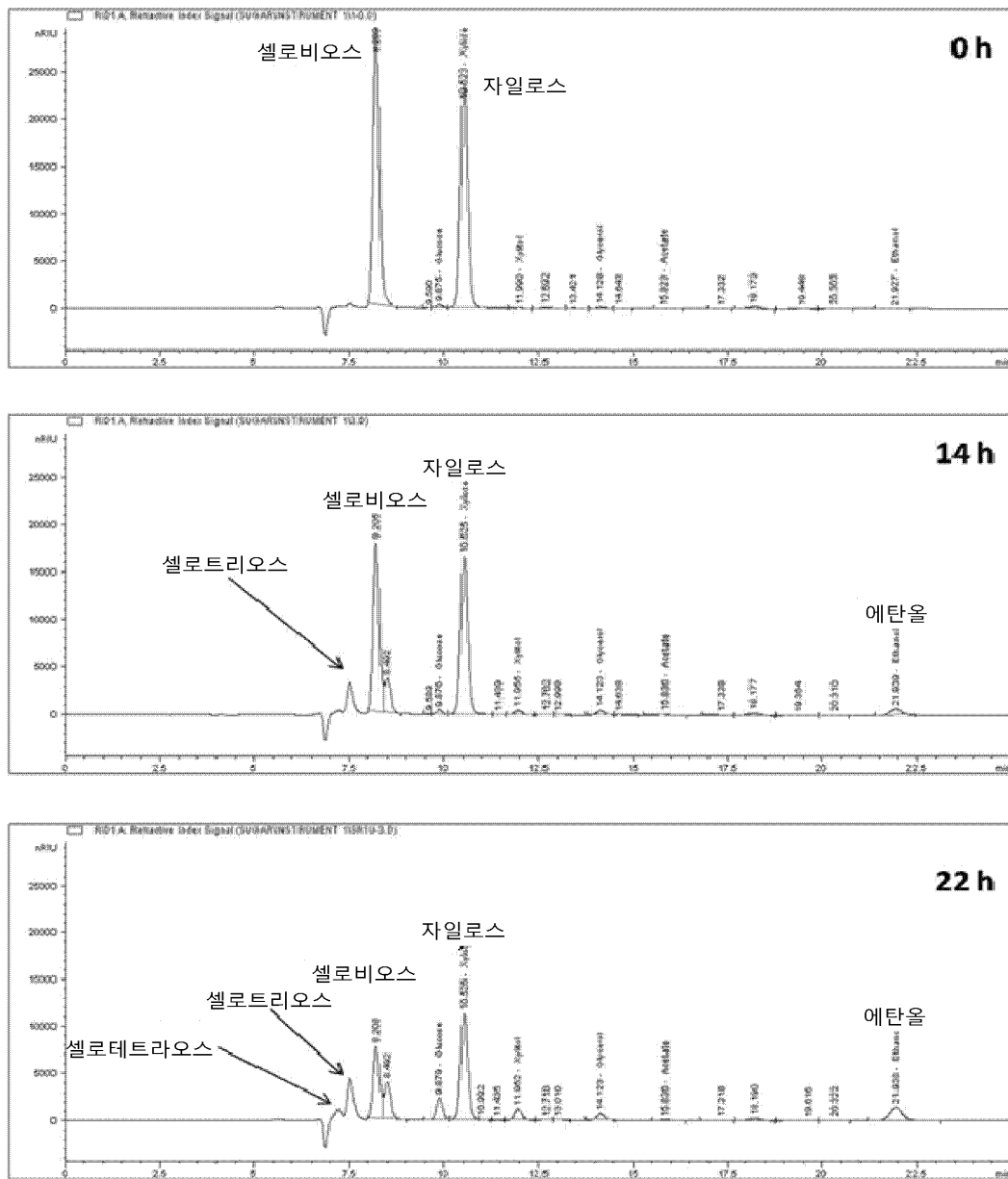
도면70



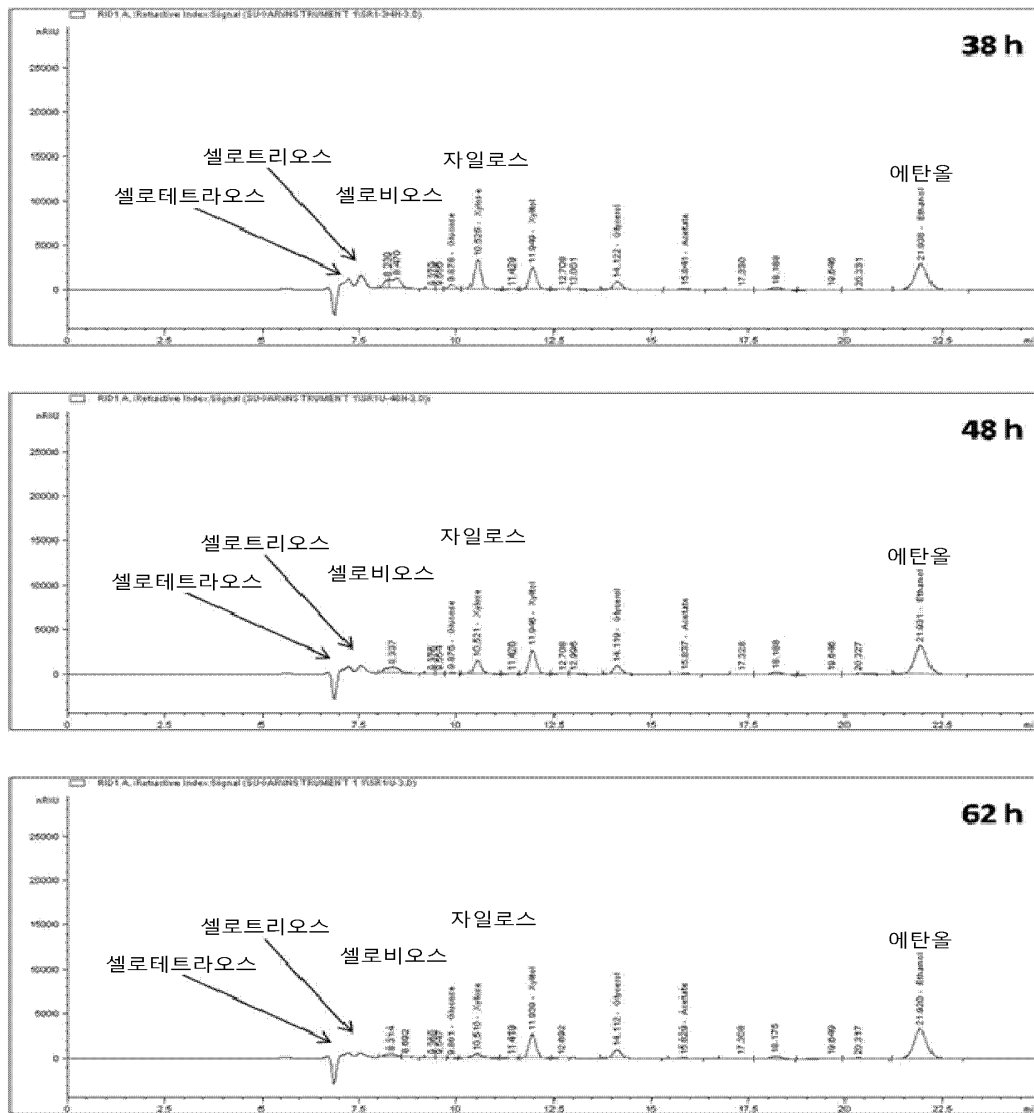
도면71



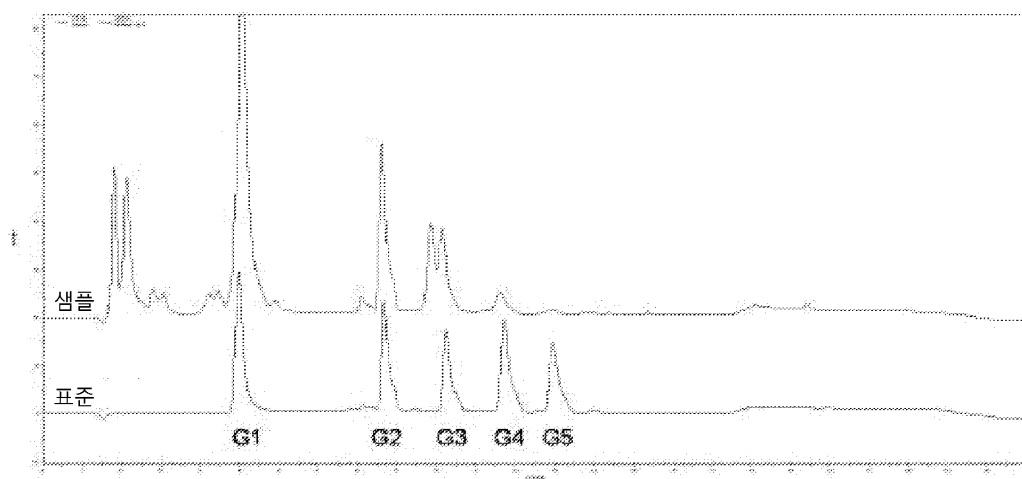
도면72a



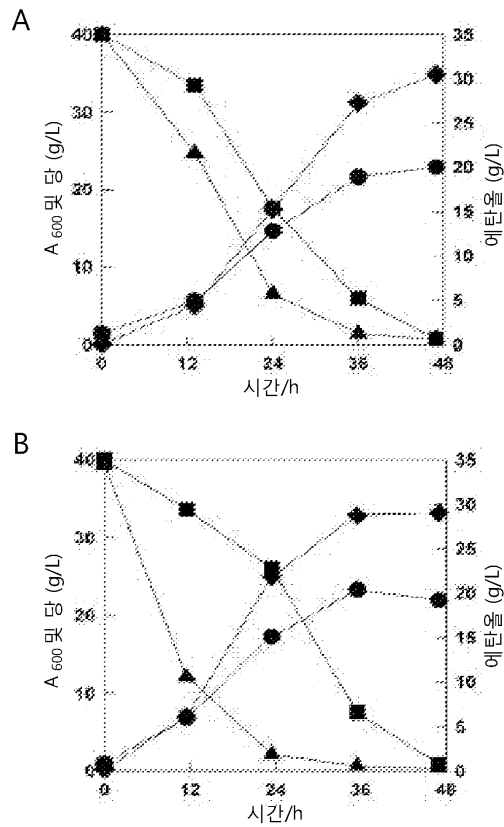
도면72b



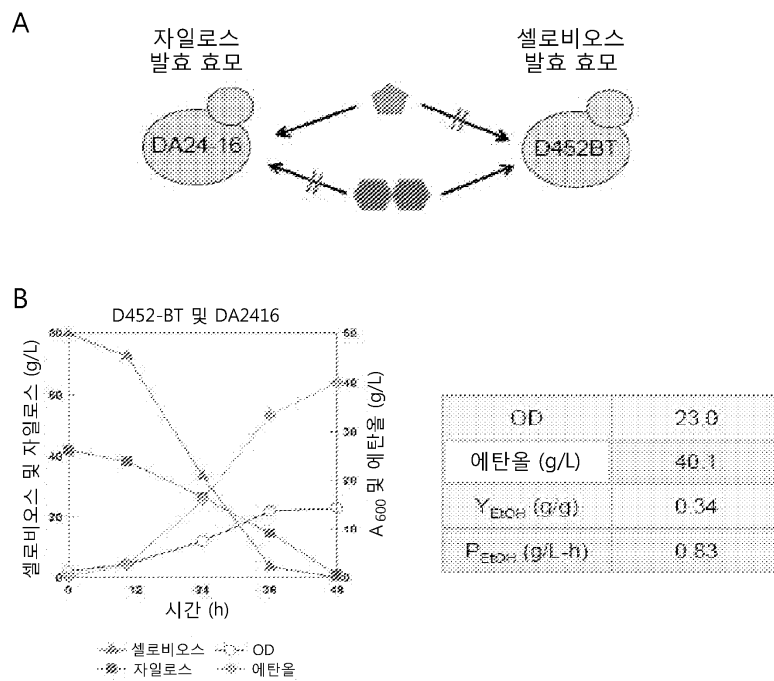
도면73



도면74



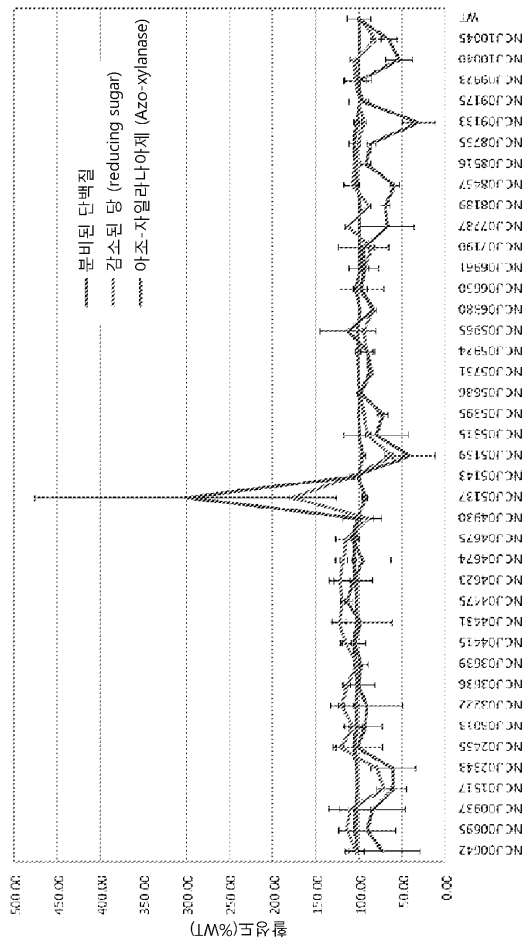
도면75



도면76

NCU00010.2	NCU01906.2	NCU04373.2	NCU06656.2	NCU09051.2b
NCU00028.2	NCU01940.2	NCU04388.2	NCU06681.2	NCU09115.2
NCU00060.2	NCU01970.2	NCU04394.2	NCU06707.2	NCU09133.2
NCU00073.2	NCU01989.2	NCU04400.2	NCU06752.2	NCU09138.2
NCU00102.2	NCU02018.2	NCU04401.2	NCU06803.2	NCU09169.2
NCU00111.2	NCU02027.2	NCU04415.2	NCU06977.2	NCU09175.2
NCU00122.2	NCU02097.2	NCU04417.2	NCU06999.2	NCU09209.2
NCU00124.2	NCU02124.2	NCU04460.2	NCU07027.2	NCU09210.2
NCU00130.2	NCU02136.2	NCU04475.2	NCU07063.2	NCU09266.2
NCU00173.2	NCU02179.2	NCU04491.2b	NCU07064.2b	NCU09267.2
NCU00292.2	NCU02188.2	NCU04510.2	NCU07133.2	NCU09285.2
NCU00299.2	NCU02238.2	NCU04521.2	NCU07143.2	NCU09316.2
NCU00304.2	NCU02307.2	NCU04605.2	NCU07158.2	NCU09427.2
NCU00305.2	NCU02316.2	NCU04623.2	NCU07225.2	NCU09491.2
NCU00378.2	NCU02342.2	NCU04674.2	NCU07267.2	NCU09532.2
NCU00379.2	NCU02343.2	NCU04675.2	NCU07273.2	NCU09533.2
NCU00591.2	NCU02361.2	NCU04676.2	NCU07286.2	NCU09562.2
NCU00608.2	NCU02397.2	NCU04781.2	NCU07287.2b	NCU09682.2
NCU00611.2	NCU02453.1	NCU04801.2	NCU07310.2	NCU09705.2
NCU00642.2	NCU02455.2	NCU04815.2	NCU07311.2	NCU09763.2b
NCU00643.2	NCU02512.2	NCU04905.2	NCU07313.2b	NCU09783.2
NCU00663.2	NCU02582.2	NCU04906.2	NCU07392.2	NCU09821.2
NCU00673.2	NCU02583.2	NCU04908.2a	NCU07453.2	NCU09886.2
NCU00695.2	NCU02596.2	NCU04908.2b	NCU07624.2	NCU09923.2
NCU00709.2	NCU02606.2	NCU04909.2	NCU07705.2	NCU09924.2
NCU00809.2	NCU02637.2	NCU04910.2	NCU07706.2	NCU10020.2
NCU00810.2	NCU02653.2	NCU04918.2	NCU07723.2	NCU10021.2
NCU00821.2	NCU02654.2	NCU04930.2	NCU07771.2	NCU10040.2
NCU00864.2	NCU02701.2	NCU04963.2	NCU07788.2	NCU10045.2
NCU00870.2	NCU02730.2	NCU05134.2	NCU07883.2	NCU10107.2a
NCU00884.2	NCU02887.2	NCU05137.2	NCU07890.2a	NCU10398.2a
NCU00888.2b	NCU03013.2	NCU05143.2	NCU07941.2	NCU10398.2b
NCU00890.2	NCU03043.2	NCU05148.2	NCU07997.2	NCU10547.2a
NCU00891.2	NCU03086.2	NCU05159.2	NCU08017.2	NCU10656.2c
NCU00892.2	NCU03188.2	NCU05164.2	NCU08076.2	NCU10733.2a
NCU00937.2	NCU03216.2	NCU05170.2	NCU08092.2	NCU10966.2a
NCU00972.1	NCU03222.2	NCU05315.2	NCU08114.2	NCU10966.2b
NCU00988.2	NCU03253.2	NCU05395.2	NCU08130.2	NCU10997.2a
NCU00992.2	NCU03263.2	NCU05493.2	NCU08131.2	xmc06_080
NCU00999.1	NCU03281.2	NCU05498.2	NCU08189.2	NCU00132.2
NCU01045.2	NCU03322.2	NCU05627.2	NCU08190.2	NCU00157.2
NCU01068.2	NCU03323.2	NCU05751.2	NCU08224.2	NCU00442.2
NCU01107.2	NCU03350.2	NCU05755.2	NCU08273.2	NCU00987.2
NCU01132.2	NCU03358.2	NCU05770.2	NCU08282.2	NCU01517.2
NCU01140.2	NCU03398.2	NCU05832.2	NCU08330.2	NCU02042.2
NCU01145.2	NCU03398.2	NCU05837.2	NCU08331.2	NCU02128.2
NCU01181.2	NCU03415.2	NCU05850.2	NCU08333.2	NCU03156.2
NCU01230.2	NCU03523.2	NCU05853.2	NCU08351.2	NCU03284.2
NCU01231.2	NCU03605.2	NCU05897.2	NCU08356.2	NCU03518.2b
NCU01233.2	NCU03637.2	NCU05924.2	NCU08384.2	NCU03636.2
NCU01240.2	NCU03639.2	NCU05965.2	NCU08428.2	NCU04053.2
NCU01258.2	NCU03731.2	NCU05977.2	NCU08457.2	NCU04197.2
NCU01275.2	NCU03749.2	NCU05994.2	NCU08516.2	NCU04525.2
NCU01320.2	NCU03753.2	NCU06005.2	NCU08541.2	NCU04626.2
NCU01328.2	NCU03779.2	NCU06043.2	NCU08549.2	NCU05303.2
NCU01417.2	NCU03803.2	NCU06123.2	NCU08554.2	NCU06081.1
NCU01419.2	NCU03813.2	NCU06125.2	NCU08677.2	NCU06961.2
NCU01420.2	NCU03893.2	NCU06138.2	NCU08687.2	NCU07190.2
NCU01424.2	NCU03949.2	NCU06143.2	NCU08750.2	NCU07215.2
NCU01430.2	NCU03965.2	NCU06255.2	NCU08752.2	NCU07325.2
NCU01436.2	NCU04109.2	NCU06261.2	NCU08755.2	NCU09161.2
NCU01555.2	NCU04133.2	NCU06265.2	NCU08771.2	NCU09689.2
NCU01701.2	NCU04264.2	NCU06277.2	NCU08779.2	NCU09698.2
NCU01704.2	NCU04265.2	NCU06305.2	NCU08943.2	NCU09777.2
NCU01740.2	NCU04266.2	NCU06358.2	NCU08949.2	NCU09792.2
NCU01759.2	NCU04272.2	NCU06364.2	NCU08977.2	NCU09883.2
NCU01815.2	NCU04287.2	NCU06380.2	NCU09010.2	NCU09906.2
NCU01847.1	NCU04295.2	NCU06490.2	NCU09013.2	
NCU01866.2	NCU04298.2	NCU06597.2	NCU09027.2	
NCU01900.2	NCU04349.2a	NCU06603.2	NCU09034.2	
NCU01904.2	NCU04369.2	NCU06650.2	NCU09041.2	
NCU01905.2	NCU04371.2	NCU06652.2	NCU09043.2	

도면77



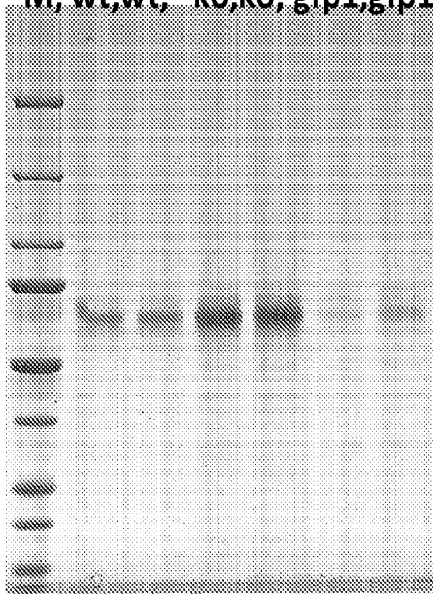
도면78

A

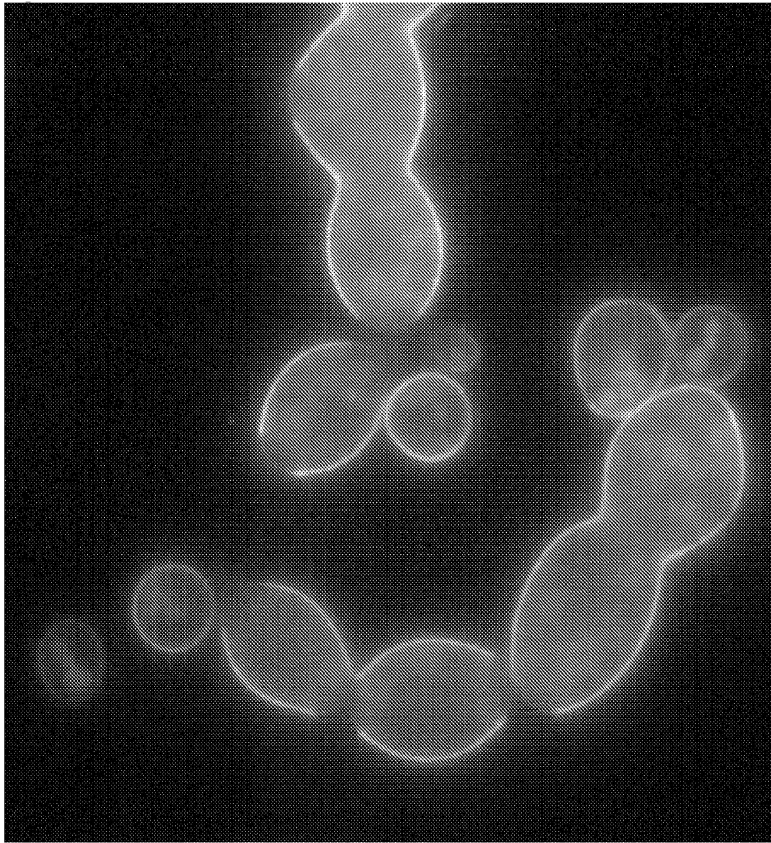
	브래드포드 (Bradford)	CMC 활성도 (activity)
Wt-1	0.57	0.50
Wt-2	0.58	0.53
Ko-1	0.89	0.76
Ko-2	0.89	0.65
GFP	0.56	0.54
GFP	0.60	0.58

B

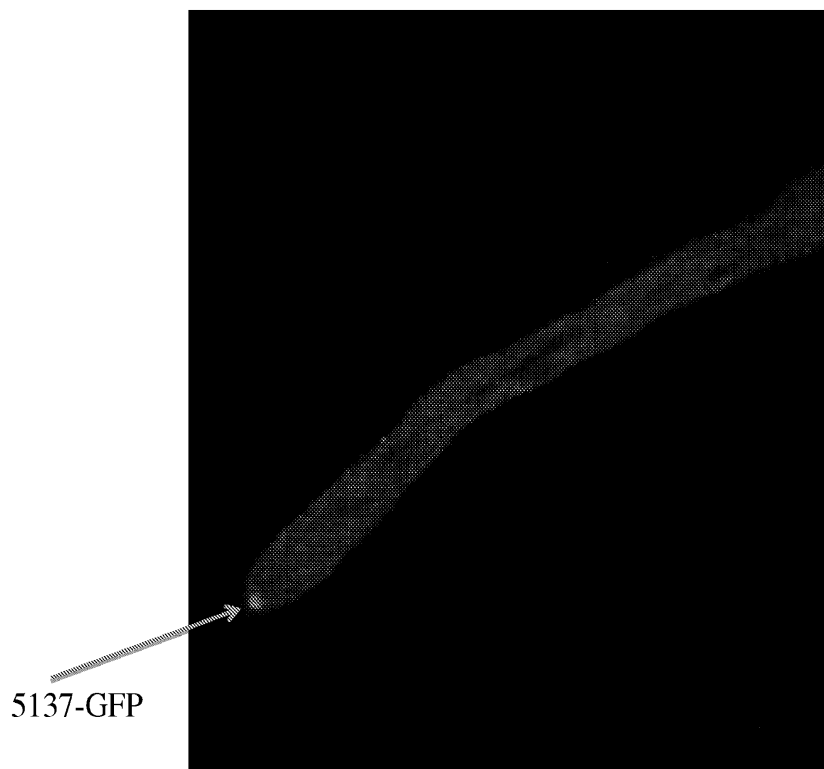
M, wt,wt, ko,ko, gfp1,gfp1



도면79



도면80



서열 목록

SEQUENCE LISTING

<110> The Regents of the University of California
The Board of Trustees of the University of Illinois
British Petroleum Technology Ventures, Inc.
GLASS, N. Louise et al.

<120> METHODS AND COMPOSITIONS FOR IMPROVING SUGAR TRANSPORT,
MIXED SUGAR FERMENTATION, AND PRODUCTION OF BIOFUELS

<130> 67779-2000140

<140> Not Yet Assigned

<141> Concurrently Herewith

<150> US 61/285,526

<151> 2009-12-10

<150> US 61/271,833

<151> 2009-07-24

<160> 97

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<221> VARIANT

<222> 1

<223> Xaa = Leu or Ile or Val or Met

<221> VARIANT

<222> 3

<223> Xaa = Phe or Leu

<221> VARIANT

<222> 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16

<223> Xaa = Any Amino Acid

<221> VARIANT

<222> 17

<223> Xaa = Tyr or Phe

<400> 1

Xaa Tyr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

1	5	10	15
Xaa Asp			

<210> 2

<211> 17

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<221> VARIANT

<222> 1

<223> Xaa = Tyr or Phe

<221> VARIANT

<222> 2, 3, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16

<223> Xaa = Any Amino Acid

<221> VARIANT

<222> 10

<223> Xaa = Pro or Val or Phe

<221> VARIANT

<222> 17

<223> Xaa = Asp or Gln

<400> 2

Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

1	5	10	15
Xaa			

<210> 3

<211> 3

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<221> VARIANT

<222> 3

<223> Xaa = Arg or Lys

<400> 3

Gly Arg Xaa

1

<210> 4

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<221> VARIANT

<222> 2, 3, 4, 5, 6, 7

<223> Xaa = Any Amino Acid

<221> VARIANT

<222> 8

<223> Xaa = Tyr or Phe

<400> 4

Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asn

1

5

<210> 5

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<221> VARIANT

<222> 3

<223> Xaa = Ile or Val or Leu or Ala

<221> VARIANT

<222> 5, 6, 7

<223> Xaa = Any Amino Acid

<400> 5

Trp Arg Xaa Pro Xaa Xaa Xaa Gln

1 5

<210> 6

<211> 24

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<221> VARIANT

<222> 6, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19, 21, 22

<223> Xaa = Any Amino Acid

<400> 6

Pro Glu Ser Pro Arg Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ala

1 5 10 15

Xaa Xaa Xaa Leu Xaa Xaa Tyr His

20

<210> 7

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<221> VARIANT

<222> 2

<223> Xaa = Gly or Ser or Thr

<221> VARIANT

<222> 4, 8

<223> Xaa = Any Amino Acid

<221> VARIANT

<222> 9

<223> Xaa = Leu or Ile or Val

<400> 7

Phe Xaa Gln Xaa Ser Gly Asn Xaa Xaa

1 5

<210> 8

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<221> VARIANT

<222> 2, 3, 4, 6, 7, 9, 11, 12, 13, 14

<223> Xaa = Any Amino Acid

<221> VARIANT

<222> 5

<223> Xaa = Tyr or Ile or Val

<221> VARIANT

<222> 16

<223> Xaa = Gly or Ala

<400> 8

Leu Xaa Xaa Xaa Xaa Xaa Xaa Glu Xaa Leu Xaa Xaa Xaa Xaa Arg Xaa

1 5 10 15

Lys Gly

<210> 9

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Transmembrane helix

<400> 9

Pro Glu Ser Pro Arg

1 5

<210> 10

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Transmembrane helix

<400> 10

Pro Glu Thr Lys

1

<210> 11

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<221> VARIANT

<222> 1

<223> Xaa = Leu or Ile or Val or Met or Ser or Thr or Ala or Gly

<221> VARIANT

<222> 2

<223> Xaa = Leu or Ile or Val or Met or Phe or Ser or Ala or Gly

<221> VARIANT

<222> 3

<223> Xaa = Any Amino Acid But Ser or His

<221> VARIANT

<222> 4

<223> Xaa = Any Amino Acid But Arg or Asp or Glu

<221> VARIANT

<222> 5

<223> Xaa = Leu or Ile or Val or Met or Ser or Ala

<221> VARIANT

<222> 6

<223> Xaa = Asp or Glu

<221> VARIANT

<222> 7

<223> Xaa = Any Amino Acid But Thr or Asp

<221> VARIANT

<

<222> 8

<223> Xaa = Leu or Ile or Val or Met or Phe or Tyr or Trp or Ala

<221> VARIANT

<222> 11

<223> Xaa = Arg or Lys

<221> VARIANT

<222> 12, 13, 14, 15, 16, 17

<223> Xaa = Any Amino Acid and up to two of them can be
present or absent

<221> VARIANT

<222> 18

<223> Xaa = Gly or Ser or Thr or Ala

<400> 11

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Arg Xaa Xaa Xaa Xaa Xaa Xaa

1 5 10 15

Xaa Xaa

<210> 12

<211> 26

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<221> VARIANT

<222> 1

<223> Xaa = Leu or Ile or Val or Met or Phe

<221> VARIANT

<222> 2, 6, 9, 10, 11, 12, 13, 14, 15, 17, 18, 20, 21, 22, 23, 24, 25

<223> Xaa = Any Amino Acid

<221> VARIANT

<222> 4

<223> Xaa = Leu or Ile or Val or Met or Phe or Ala

<221> VARIANT

<222> 5

<223> Xaa = Any Amino Acid But Val

<221> VARIANT

<222> 8

<223> Xaa = Any Amino Acid But Lys or Pro

<221> VARIANT

<222> 16

<223> Xaa = Leu or Ile or Phe or Tyr

<221> VARIANT

<222> 19

<223> Xaa = Glu or Gln

<221> VARIANT

<222> 26

<223> Xaa = Arg or Lys

<400> 12

Xaa Xaa Gly Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

20 25

<210> 13

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<221> VARIANT

<222> 1

<223> Xaa = Ala or Cys

<221> VARIANT

<222> 2, 4, 5, 6, 7

<223> Xaa = Any Amino Acid

<221> VARIANT

<222> 8

<223> Xaa = Any Amino Acid But Glu or Asp

<400> 13

Xaa Xaa Val Xaa Xaa Xaa Xaa Xaa

1 5

<210> 14

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<220>

<221> VARIANT

<222> 1

<223> Xaa = Leu or Ile or Val or Met or Phe or Ser or Thr or Cys

<220>

<221> VARIANT

<222> 2

<223> Xaa = Leu or Ile or Val or Phe or Tyr or Ser

<220>

<221> VARIANT

<222> 3

<223> Xaa = Leu or Ile or Val

<220>

<221> VARIANT

<222> 4

<223> Xaa = Leu or Ile or Val or Met or Ser or Thr

<220>

<221> VARIANT

<222> 8

<223> Xaa = Leu or Ile or Val or Met or Phe or Ala or Arg

<220>

<221> VARIANT

<222> 9

<223> Xaa = Cys or Ser or Ala or Gly or Asn

<400> 14

Xaa Xaa Xaa Xaa Glu Asn Gly Xaa Xaa

1

5

<210> 15

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence motif

<220>

<221> VARIANT

<222> 1, 2

<223> Xaa = Leu or Ile or Val or Met

<220>

<221> VARIANT

<222> 3

<223> Xaa = Lys or Arg

<220>

<221> VARIANT

<222> 4, 6, 7, 8, 9, 16, 17

<223> Xaa = Any Amino Acid

<220>

<221> VARIANT

<222> 5

<223> Xaa = Glu or Gln or Lys or Arg or Asp

<220>

<221> VARIANT

<222> 11

<223> Xaa = Leu or Ile or Val or Met or Phe or Thr or Cys

<220>

<221> VARIANT

<222> 12

<223> Xaa = Leu or Ile or Val or Thr

<220>

<221> VARIANT

<222> 13

<223> Xaa = Leu or Ile or Val or Met or Phe

<220>

<221> VARIANT

<222> 14

<223> Xaa = Ser or Thr

<220>

<221> VARIANT

<222> 18

<223> Xaa = Ser or Gly or Ala or Asp or Asn or Ile or Thr

<400> 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Asp Xaa

1 5 10 15

Xaa Xaa

<210

> 16

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 16

atctggaag cgaacaaag 19

<210> 17

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 17

tagcggtcgt cggaatag 18

<210> 18

<211> 18
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 18
 cccatcacca ctactacc 18

<210> 19
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 19
 ccagccctga acaccaag 18

<210> 20
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 20
 tgatcttacc gactacct 18

<210> 21
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 21
 cagagcttct ccttgatg 18

<210> 22
 <211> 533

<212> PRT

<213> Sporotrichum thermophile

<400> 22

Met Gly Leu Ser Thr Lys Ile Leu Gln Lys Ile Val Arg Asn Glu Ala

1 5 10 15

Met Ala Ser Asp Pro Pro Glu Ile Tyr Gly Trp Arg Val Tyr Leu Leu

20 25 30

Ala Cys Ser Ala Cys Phe Gly Ala Met Ser Phe Gly Trp Asp Ser Ser

35 40 45

Val Ile Gly Gly Val Ile Val Leu Pro Pro Phe Ile Arg Asp Phe Asn

50 55 60

Leu Gly Asp Pro Lys Ser Gln Ala Ser Ala Asn Leu Ser Ala Asn Ile

65 70 75 80

Val Ser Thr Leu Gln Ala Gly Cys Phe Leu Gly Ala Leu Val Ala Ser

85 90 95

Pro Met Thr Asp Arg Phe Gly Arg Lys Trp Cys Leu Ile Gly Val Ser

100 105 110

Leu Ile Ile Ile Ile Gly Ile Ile Met Gln Ala Ala Ala Ser Gly Asn

115 120 125

Leu Gly Pro Met Tyr Ala Gly Arg Phe Ile Ala Gly Ala Gly Val Gly

130 135 140

Ala Ala Ser Thr Ile Asn Pro Ile Tyr Val Ser Glu Asn Ala Pro Arg

145 150 155 160

Ala Ile Arg Gly Leu Leu Thr Gly Leu Tyr Gln Leu Phe Ile Val Thr

165 170 175

Gly Gly Met Ile Ala Phe Trp Ile Asn Tyr Ser Val Ser Ile His Phe

180 185 190

Pro Glu Thr Lys Ile Met Tyr Val Phe Pro Leu Ala Ile Gln Ala Leu

195 200 205

Pro Ala Ala Leu Leu Cys Leu Cys Met Leu Leu Cys Gln Glu Ser Pro

210 215 220

Arg Trp Leu Ala Arg Arg Asp Arg Trp Glu Asp Thr Lys Arg Val Leu

225 230 235 240
 Ser Arg Ile Arg Asn Leu Pro Pro Asp His Pro Tyr Ile Gln Asp Glu

 245 250 255
 Phe Gln Glu Ile Val Ala Gln Leu Glu His Glu Arg Arg Leu Ile Gly
 260 265 270
 Asp Ala Ser Phe Trp Asn Leu Gln Arg Glu Met Trp Thr Ile Ala Gly
 275 280 285
 Asn Arg Arg Arg Val Leu Ile Ser Ile Ile Leu Met Ile Cys Gln Gln
 290 295 300
 Met Thr Gly Thr Asn Ala Ile Asn Thr Tyr Ala Pro Thr Ile Phe Lys

 305 310 315 320
 Asn Leu Gly Leu Thr Gly Thr Ser Thr Ser Leu Phe Ser Thr Gly Val
 325 330 335
 Tyr Gly Ile Val Lys Val Thr Ser Cys Ile Ile Phe Leu Leu Phe Met
 340 345 350
 Ala Asp Ser Leu Gly Arg Arg Arg Ser Leu Leu Trp Thr Ser Ile Ala
 355 360 365
 Gln Gly Leu Ala Met Phe Tyr Ile Gly Leu Tyr Val Arg Ile Ala Pro

 370 375 380
 Pro Lys Glu Gly Glu Ser Val Pro Pro Ala Gly Tyr Phe Ala Leu Val
 385 390 395 400
 Cys Ile Phe Leu Phe Ala Ala Phe Phe Gln Phe Gly Trp Gly Pro Ala
 405 410 415
 Cys Trp Ile Tyr Ala Ser Glu Ile Pro Ala Ala Arg Leu Arg Ser Leu
 420 425 430
 Asn Val Ala Tyr Ala Ala Ala Thr Gln Trp Leu Phe Asn Phe Val Val

 435 440 445
 Ala Arg Thr Val Pro Val Met Ile Val Thr Met Gly Glu Gly Gly Tyr
 450 455 460
 Gly Thr Tyr Leu Leu Phe Gly Ser Phe Cys Phe Ser Met Phe Val Phe
 465 470 475 480

Val Trp Phe Phe Val Pro Glu Thr Lys Gly Val Ser Leu Glu Ala Met
485 490 495
Asp Lys Leu Phe Gly Val Thr Asp Glu Ser Ser Lys Ser Leu Thr Val
500 505 510
Asp Glu Asp Ala Lys Glu Lys Glu Lys Asp Gly Pro His Ala Arg Gln
515 520 525
Thr Glu Val Val Ala
530
<210> 23
<211> 512
<212> PRT
<213> Sporotrichum thermophile
<400> 23
Met Lys Lys Phe Leu Gly Leu Arg Gly Gln Ala Leu Asn Leu Ala Val
1 5 10 15
Gly Thr Ile Ala Gly Cys Asp Phe Leu Leu Phe Gly Tyr Asp Gln Gly
20 25 30
Val Met Gly Gly Ile Leu Thr Leu Lys Val Phe Leu Asp Ala Phe Pro
35 40 45
Met Ile Asn Pro Glu Ala Ala Gly Leu Ser His Asp Glu Ser Ser Thr
50 55 60
Arg Ser Thr Tyr Gln Gly Ile Ala Val Ala Ser Tyr Asn Leu Gly Cys
65 70 75 80
Phe Leu Gly Ala Ile Ile Thr Ile Phe Ile Gly Asn Pro Leu Gly Arg
85 90 95
Lys Arg Val Ile Met Leu Gly Thr Ser Val Met Val Ile Gly Ala Ile
100 105 110
Leu Gln Ala Ser Ser Thr Thr Leu Pro Gln Phe Ile Val Gly Arg Ile
115 120 125
Ile Thr Gly Leu Gly Asn Gly Gly Asn Thr Ser Thr Val Pro Thr Trp
130 135 140

Gln Ser Glu Thr Ser Lys Ala His Lys Arg Gly Lys Met Ile Phe Phe

145 150 155 160

Cys Ala Ile Ile Leu Ala Phe Ile Pro Phe Leu Pro Glu Ser Pro Arg

165 170 175

Trp Leu Ile Leu Lys Gly Arg Glu Asp Glu Ala Arg Glu Val Ile Ala

180 185 190

Ala Leu Glu Asp Thr Asp Thr Ser Asp Arg Ile Val Glu Asn Glu Phe

195 200 205

Leu Ala Ile Lys Glu Thr Val Leu Glu Met Ser Lys Gly Thr Phe Arg

210 215 220

Asp Leu Phe Thr Met Asp Lys Asn Arg Asn Leu His Arg Thr Leu Leu

225 230 235 240

Ala Tyr Phe Asn Gln Val Phe Gln Gln Ile Ser Gly Ile Asn Leu Ile

245 250 255

Thr Tyr Tyr Ala Ala Val Ile Tyr Lys Gly Leu Gly Met Ser Asp Phe

260 265 270

Leu Ser Arg Leu Leu Ala Ala Leu Asn Gly Thr Glu Tyr Phe Leu Ala

275 280 285

Ser Trp Pro Ala Val Phe Leu Val Glu Arg Val Gly Arg Arg Asn Leu

290 295 300

Met Leu Phe Gly Ala Val Gly Gln Ala Ala Thr Met Ala Ile Leu Ala

305 310 315 320

Gly Val Asn Ser Arg Gln Glu Thr Gly Phe Gln Ile Ala Gly Ile Val

325 330 335

Phe Leu Phe Val Phe Asn Thr Phe Phe Ala Val Gly Trp Leu Gly Met

340 345 350

Thr Trp Leu Tyr Pro Ala Glu Ile Val Pro Leu Arg Ile Arg Ala Pro

355 360 365

Ala Asn Ala Leu Ser Thr Ser Ala Asn Trp Ile Phe Asn Phe Leu Val

370 375 380

Val Met Ile Thr Pro Val Ala Phe Asn Asn Ile Gly Tyr Gln Thr Tyr

385 390 395 400
Ile Ile Phe Ala Val Ile Asn Ala Phe Met Val Pro Cys Val Tyr Phe

 405 410 415
Phe Tyr Pro Glu Thr Ala Tyr Arg Ser Leu Glu Glu Met Asp Asn Ile

 420 425 430
Phe His Lys Val Ala Asp Gly Trp Lys Gly Val Phe Thr Val Val His

 435 440 445
Gln Ala Lys Val Glu Pro Arg Trp Tyr Gly Lys Asn Gly Glu Leu Leu

 450 455 460
Val Asp Tyr Gln Gln Thr Glu Glu His Arg Arg His Leu Gln Gln Gln

465 470 475 480
Glu Gly Ala Val Thr Ala Ser Glu Lys Arg Ser Val Glu Gly Ala Gly

 485 490 495
Ser Gly Ser Gly Ser Gly Asp Val Lys Gln Asp Glu Tyr Lys Asp Cys

 500 505 510

<210> 24

<211> 519

<212> PRT

<213> Sporotrichum thermophile

<400> 24

Met Glu Ser Thr His Glu Pro Ala Asp Pro Ile Ala Lys Gly Val Leu

1 5 10 15

Ala Thr Ala Lys Gln Ser Trp His Asp Leu Phe Ile Phe Lys Gln Arg

20 25 30

Val Val Val Thr Asn Glu Leu Gly Glu Thr Ser Thr Glu Trp Ala Arg

35 40 45

Pro Val Pro Leu Arg Asn Pro Ile Ser Leu Leu Ala Gln Leu Ser Ala

50 55 60

Arg Asn Trp Leu Phe Phe Ile Val Gly Phe Leu Ala Trp Val Ala Asp

65 70 75 80

Ala Tyr Asp Phe His Ala Leu Ser Ile Gln Gln Val Lys Leu Ala Glu

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

Ile Ser Gln Trp Phe Gly Arg Arg Arg Thr Ile Ile Val Ala Ala Leu
340 345 350
Ile Ser Gly Leu Ile Ile Pro Ala Trp Ile Leu Pro Glu Gly Glu Arg
355 360 365
Ser Leu Ser Val Thr Gly Phe Phe Met Gln Phe Phe Val Gln Gly Ala
370 375 380
Trp Gly Val Ile Pro Ile His Leu Asn Glu Leu Ser Pro Pro Ala Phe
385 390 395 400

Arg Ser Ser Phe Pro Gly Leu Thr Tyr Gln Leu Gly Asn Met Ile Ser
405 410 415
Ser Pro Ser Ala Gln Ile Val Asn Ala Ile Ala Glu Ser His Ser Val
420 425 430
Thr Ser Lys Ser Gly Lys Ser Val Asn Ala Tyr Gly Pro Thr Met Gly
435 440 445
Ile Ala Thr Ala Ile Ile Ala Thr Gly Ile Ala Val Thr Thr Ala Leu
450 455 460

Gly Pro Glu Lys Arg Gly Arg Glu Phe Glu Lys Thr Leu Pro Ala Gly
465 470 475 480
Met Asn Ile Ile Gln Gly Gly Lys Ala Val Asp Asp Leu Glu Lys Gly
485 490 495
Asp Ser Arg Asp Glu Lys Pro Val Val Gly Glu Val Glu Gly Gly Asn
500 505 510
Asp Gly Ser Gly Glu Lys Lys
515

<210> 25

<211> 488

<212> PRT

<213> Sporotrichum thermophile

<400> 25

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

20 25 30
 Arg Gly Phe Ala Ala Thr Asp Gln Asn Gly Gln Pro Ile Val Gln Phe
 35 40 45
 Asp Leu Lys Ala Glu Ala Arg Leu Arg Arg Lys Leu Asp Leu Phe Ile
 50 55 60

 Val Pro Thr Val Ser Leu Leu Tyr Leu Phe Cys Phe Ile Asp Arg Ala
 65 70 75 80
 Asn Ile Gly Asn Ala Arg Ile Ala Gly Leu Glu Lys Asp Leu Asn Leu
 85 90 95
 Thr Gly Tyr Asp Tyr Asn Ala Leu Leu Ser Val Phe Tyr Ile Ser Tyr
 100 105 110
 Ile Val Phe Glu Ile Pro Ser Asn Ile Ala Cys Lys Trp Ile Gly Pro
 115 120 125

 Gly Trp Phe Ile Pro Ala Ile Ser Leu Gly Phe Gly Val Val Ser Leu
 130 135 140
 Ala Thr Ala Phe Val Asp Asn Phe Ala Gln Ala Ala Gly Val Arg Phe
 145 150 155 160
 Leu Leu Gly Val Phe Glu Ala Gly Met Met Pro Gly Ile Ala Tyr Tyr
 165 170 175
 Leu Ser Arg Trp Tyr Arg Arg Ala Glu Leu Thr Phe Arg Leu Ser Leu
 180 185 190

 Tyr Ile Val Met Ala Pro Met Ala Gly Ala Phe Gly Gly Leu Leu Ala
 195 200 205
 Ser Gly Ile Leu Ser Leu Asp His Val Gly Gly Val Thr Gly Trp Arg
 210 215 220
 Met Ile Phe Val Val Glu Gly Ile Ile Thr Ile Gly Leu Ser Val Ile
 225 230 235 240
 Ser Phe Ile Thr Leu Thr Asp Arg Pro Glu Thr Ala Arg Trp Leu Thr
 245 250 255

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

Ala Thr Thr Glu Val Leu Asp Arg Met Asp Thr Lys Lys Leu Ile Gln
275 280 285

Gly Ile Leu Ser Pro Val Thr Leu Ala Thr Ser Phe Met Phe Leu Leu
290 295 300

Asn Asn Ile Thr Gln Leu Phe Thr Val Pro Pro Tyr Val Val Gly Gly
305 310 315 320

Phe Phe Thr Leu Ala Leu Pro Leu Leu Ser Trp Tyr Leu Asp Arg Arg
325 330 335

Gln Ile Ile Ile Leu Leu Ser Thr Pro Leu Val Ile Val Gly Tyr Ser
340 345 350

Met Phe Leu Gly Thr Thr Asn Pro Ser Ala Arg Tyr Gly Ala Thr Phe
355 360 365

Leu Leu Ser Ser Ser Leu Phe Ala Val Gly Ala Leu Ser Asn Ser Gln
370 375 380

Val Ser Ala Asn Val Val Ser Asp Thr Ala Arg Ser Ser Ala Ile Gly
385 390 395 400

Leu Asn Val Met Met Gly Asn Val Gly Gly Leu Ile Ala Thr Trp Ser
405 410 415

Tyr Leu Pro Trp Asp Gly Pro Asn Tyr Lys Ile Gly Asn Gly Leu Asn
420 425 430

Leu Ala Ala Cys Cys Thr Val Leu Ile Leu Ser Ala Val Thr Leu Leu
435 440 445

Trp Met Lys Trp Asp Asn Arg Arg Arg Glu Gly Arg Asn Ala Glu Glu
450 455 460

Glu Leu Ala Gly Met Ser Arg Gln Glu Ile Gln Asp Leu Asp Trp Lys
465 470 475 480

His Pro Ala Phe Arg Trp Arg Pro
485

<210> 26

<211> 546

<212> PRT

<213> Sporotrichum thermophile

<400> 26

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

Ser Ala Asp Pro Ile Val Thr Arg Leu Val Gln Glu Asp Lys Ile Pro
20 25 30

Trp Tyr Lys Lys Pro Asn Leu Arg Arg Met Tyr Ile Phe Leu Phe Leu
35 40 45

Cys Cys Met Gly Val Glu Met Thr Ser Gly Phe Asp Ser Gln Leu Ile
50 55 60

Asn Thr Leu Gln Tyr Ala Glu Thr Phe His Lys Tyr Leu Gly Asn Gly
65 70 75 80

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

Phe Val Asn Ser Cys Tyr Gln Leu Gly Ser Ile Phe Ala Val Pro Ile
100 105 110

Ala Pro Trp Phe Ala Gln Arg Phe Gly Arg Arg Trp Ser Ile Met Leu
115 120 125

Gly Ser Leu Ile Met Val Gly Gly Ala Ile Ile Gln Gly Phe Ala Gln
130 135 140

His Val Ala Met Tyr Ile Ile Ala Arg Met Ile Leu Gly Met Gly Ile
145 150 155 160

Leu Phe Cys Ile Ile Ser Gly Ala Ala Leu Ile Gly Glu Leu Gly His
165 170 175

Pro Lys Glu Arg Ala Val Leu Thr Ser Leu Phe Asn Ser Ser Tyr Phe
180 185 190

Ile Gly Gln Ile Leu Ala Ser Ala Ile Thr Ile Gly Thr Thr Glu Met
195 200 205

Lys Thr Asn Trp Ala Trp Arg Leu Pro Ser Leu Leu Gln Ile Cys Pro
210 215 220

Ser Leu Leu Gln Ile Val Thr Val Phe Phe Leu Pro Glu Ser Pro Arg
225 230 235 240

Phe Leu Ile Ser Lys Asp Arg Asp Asp Asp Ala Lys Glu Val Leu Ile
 245 250 255
 Lys Tyr His Ala Glu Gly Asp Ala Ser Ser Leu Leu Val Gln Ala Glu
 260 265 270

 Ile Val Gln Ile Arg Glu Thr Ile Arg Thr Glu Met Glu Val Ser Asn
 275 280 285
 Gln Ser Trp Met Glu Leu Val Ser Thr Tyr Gly Met Arg Arg Arg Leu
 290 295 300
 Val Ile Thr Leu Phe Ile Gly Leu Phe Thr Gln Leu Ser Gly Asn Thr
 305 310 315 320
 Leu Leu Ser Tyr Tyr Ser Gly Lys Leu Phe Glu Met Met Gly Tyr Thr
 325 330 335

 Glu Ala Ser Val Lys Thr Arg Ile Asn Val Ala Asn Ala Cys Trp Ser
 340 345 350
 Leu Leu Asn Ala Thr Thr Ile Ala Phe Leu Val Pro Tyr Phe Lys Arg
 355 360 365
 Arg His Met Phe Met Thr Ser Ala Leu Ser Met Cys Ala Val Phe Ile
 370 375 380
 Ala Ile Thr Val Ser Leu Glu Arg Thr Gln Ala Ala Gln Asp Ala Gly
 385 390 395 400

 Phe Lys Asn Thr Ala Ala Gly Ile Ser Gly Leu Phe Trp Tyr Phe Ala
 405 410 415
 Phe Ala Pro Cys Tyr Asn Met Gly Asn Asn Ala Leu Thr Tyr Thr Tyr
 420 425 430
 Leu Val Glu Leu Trp Pro Tyr Ser His Arg Ser Arg Gly Ile Gly Val
 435 440 445
 Gln Gln Ile Phe Gly Lys Leu Gly Gly Phe Phe Ser Thr Asn Val Asn
 450 455 460

 Ser Ile Ala Leu Asp Ala Ile Arg Trp Lys Tyr Met Ala Ile Tyr Cys
 465 470 475 480
 Gly Trp Ile Phe Phe Glu Phe Leu Ile Val Phe Phe Leu Tyr Pro Glu

485 490 495
 Thr Ser Gly Arg Thr Leu Glu Glu Leu Ala Phe Leu Phe Glu Asp Ala
 500 505 510
 Ser Leu Asn Glu Lys Ala Ala Ala Val Glu Lys Gln Ile His Tyr
 515 520 525

 Gly Asp Glu Lys Val Val His Glu Glu Gly Gln Pro Ala Ala Lys Ser
 530 535 540
 Val Val
 545
 <210> 27
 <211> 481
 <212> PRT
 <213> *Sporotrichum thermophile*
 <400> 27
 Met Leu Ser Ser Gly Phe Trp Lys Arg Arg Ser Leu Arg Val Pro Asp
 1 5 10 15
 Asn Gln Arg Thr Lys Ala Ala Glu Leu Thr Leu Arg Glu Ser Leu Tyr
 20 25 30
 Pro Leu Ser Leu Val Thr Ile Leu Phe Phe Leu Trp Gly Phe Ser Tyr

 35 40 45
 Gly Leu Leu Asp Thr Leu Asn Lys His Phe Gln Asn Thr Leu Gly Ile
 50 55 60
 Thr Lys Thr Arg Ser Ser Gly Leu Gln Ala Ala Tyr Phe Gly Ala Tyr
 65 70 75 80
 Pro Leu Ala Ser Leu Gly His Ala Ala Trp Ile Leu Arg His Tyr Gly
 85 90 95
 Tyr Arg Ala Val Phe Ile Trp Gly Leu Phe Leu Tyr Gly Leu Gly Ala

 100 105 110
 Leu Leu Ala Ile Pro Ser Ile Met His His Ser Phe Ala Gly Phe Cys
 115 120 125
 Val Cys Ile Phe Ile Ile Gly Asn Gly Leu Gly Ser Leu Glu Thr Ala
 130 135 140

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

385 390 395 400
Phe Ser Met Leu Val Pro Leu Cys Phe Phe Val Ala Ala Trp Thr Tyr
 405 410 415
Ala Leu Ala Val Asn Phe Ala Pro Pro Tyr Arg Ser Val Val Asp Ala

 420 425 430
Phe Ser Thr Thr Asp Val Gly Leu Arg Glu Lys Gln Arg Glu Asp Val
 435 440 445
Gly Ala Glu Lys Gly Gly Glu Ala Gly Gly Lys Gly Gly Val Thr Gly
 450 455 460
Pro Glu Asp Ala Ser Glu Asp Lys Pro Asp Val Val Asn Ser Glu Lys
465 470 475 480
Val

<210> 28

<211> 419

<212> PRT

<213> Sporotrichum thermophile

<400> 28

Met Leu Ser Ser Leu Arg Ile Ala Ser Arg Arg Ala Ala Val Ala Arg
1 5 10 15
Asn Phe Ser Ala Val Arg Ala Ala Ser Thr Trp Ala Asn Val Pro Gln
 20 25 30
Gly Pro Pro Val Cys Ile Thr Glu Ala Phe Lys Ala Asp Pro Phe Glu
 35 40 45
Lys Lys Ile Asn Leu Gly Val Gly Ala Tyr Arg Asp Asp Lys Gly Lys
50 55 60

Pro Tyr Val Leu Pro Ser Val Arg Lys Ala Glu Glu Lys Val Ile Ala
65 70 75 80
Ser Arg Leu Asn Lys Glu Tyr Ala Gly Ile Thr Gly Val Pro Glu Phe
 85 90 95
Thr Lys Ala Ala Ala Val Leu Ala Tyr Gly Lys Asp Ser Ser Ala Leu
 100 105 110

Asp Arg Leu Ala Ile Thr Gln Ser Ile Ser Gly Thr Gly Ala Leu Arg
 115 120 125

Ile Gly Ala Ala Phe Leu Ser Arg Phe Tyr Pro Gly Ala Lys Thr Ile
 130 135 140

Tyr Ile Pro Thr Pro Ser Trp Ala Asn His Ala Ala Val Phe Lys Asp
 145 150 155 160

Ser Gly Leu Gln Val Glu Lys Tyr Ala Tyr Tyr Asn Lys Asp Thr Ile
 165 170 175

Arg Leu Asp Phe Glu Gly Met Ile Ala Asp Ile Asn Lys Ala Pro Asn
 180 185 190

Gly Ser Ile Phe Leu Phe His Ala Cys Ala His Asn Pro Thr Gly Val
 195 200 205

Asp Pro Thr Gln Glu Gln Trp Lys Glu Ile Glu Ala Ala Val Lys Ala
 210 215 220

Lys Gly His Phe Ala Phe Phe Asp Met Ala Tyr Gln Gly Phe Ala Ser
 225 230 235 240

Gly Asp Ile His Arg Asp Ala Phe Ala Val Arg Tyr Phe Val Glu Lys
 245 250 255

Gly His Asn Ile Cys Leu Ala Gln Ser Phe Ala Lys Asn Met Gly Leu
 260 265 270

Tyr Gly Glu Arg Thr Gly Ala Phe Ser Ile Val Cys Ala Asp Ala Glu
 275 280 285

Glu Arg Lys Arg Val Asp Ser Gln Ile Lys Ile Leu Val Arg Pro Met
 290 295 300

Tyr Ser Asn Pro Pro Ile His Gly Ala Arg Ile Ala Ala Glu Ile Leu
 305 310 315 320

Asn Thr Pro Glu Leu Tyr Asp Gln Trp Leu Val Glu Val Lys Glu Met
 325 330 335

Ala Asn Arg Ile Ile Thr Met Arg Ala Leu Leu Lys Glu Asn Leu Glu
 340 345 350

Lys Leu Gly Ser Lys His Asp Trp Ser His Ile Thr Ser Gln Ile Gly

355 360 365
Met Phe Ala Tyr Thr Gly Leu Thr Pro Glu Gln Met Glu Lys Leu Ala
370 375 380

Lys Glu His Ser Val Tyr Ala Thr Arg Asp Gly Arg Ile Ser Val Ala
385 390 395 400
Gly Ile Thr Thr Asp Asn Val Gly Arg Leu Ala Glu Ala Ile Phe Lys
405 410 415
Val Lys Gly

<210> 29

<211> 522

<212> PRT

<213> Sporotrichum thermophile

<400> 29

Met Gly Ile Phe Ala Phe Asn Lys Gln Lys Pro Asn Ala Glu Ala Thr
1 5 10 15
Ala Val Ala Gln Glu Glu Ala Pro Gln Phe Glu Arg Val Asp Trp Lys

20 25 30
Arg Asp Pro Gly Leu Arg Lys Leu Tyr Phe Tyr Ala Phe Val Leu Cys
35 40 45
Ile Ala Ser Ala Thr Thr Gly Tyr Asp Gly Met Phe Phe Asn Ser Val
50 55 60
Gln Asn Phe Glu Thr Trp Glu Asn Tyr Phe Asn His Pro Thr Gly Ser
65 70 75 80
Lys Leu Gly Val Leu Gly Ala Leu Tyr Gln Ile Gly Ser Leu Ala Ser

85 90 95
Ile Pro Leu Val Pro Ile Ile Ala Asp Arg Val Gly Arg Lys Ile Pro
100 105 110
Ile Ala Ile Gly Cys Val Ile Met Ile Val Gly Ala Val Leu Gln Ala
115 120 125
Ala Cys Arg Asn Leu Gly Thr Phe Met Gly Gly Arg Phe Leu Leu Gly
130 135 140

Phe Gly Asn Ser Leu Ala Gln Leu Cys Ser Pro Met Leu Leu Thr Glu

145 150 155 160

Leu Ala His Pro Gln His Arg Gly Arg Leu Thr Thr Val Tyr Asn Cys

165 170 175

Leu Trp Asn Val Gly Ala Leu Val Val Ala Trp Val Ser Phe Gly Thr

180 185 190

Asp Tyr Leu Lys Ser Asp Trp Ser Trp Arg Ile Pro Ala Leu Ile Gln

195 200 205

Ala Phe Pro Ser Val Ile Gln Leu Leu Phe Ile Phe Trp Val Pro Glu

210 215 220

Ser Pro Arg Tyr Leu Met Ala Lys Asp Lys His Glu Arg Ala Leu Ala

225 230 235 240

Ile Leu Ala Lys Tyr His Ala Asn Gly Asp Ala Asn His Pro Thr Val

245 250 255

Gln Phe Glu Tyr Arg Glu Ile Lys Glu Thr Leu Arg Leu Glu Phe Glu

260 265 270

Ala Ser Lys Ser Ser Ser Tyr Leu Asp Phe Val Arg Thr Arg Gly Asn

275 280 285

Arg Tyr Arg Leu Ala Val Leu Ile Ser Leu Gly Ile Phe Ser Gln Trp

290 295 300

Ser Gly Asn Ala Ile Ile Ser Asn Tyr Ser Ser Lys Leu Tyr Asp Thr

305 310 315 320

Ala Gly Val Thr Gly Ser Thr Gln Lys Leu Gly Leu Ser Ala Gly Gln

325 330 335

Thr Gly Leu Ser Leu Ile Ile Ser Val Thr Met Ala Leu Leu Val Asp

340 345 350

Lys Phe Gly Arg Arg Pro Met Phe Leu Thr Ser Thr Ala Gly Met Phe

355 360 365

Cys Thr Phe Ile Phe Trp Thr Leu Thr Ser Gly Leu Tyr Glu Glu His

370 375 380

Asn Ala Asp Gly Ala Arg Tyr Ala Met Ile Leu Phe Ile Trp Ile His

385 390 395 400
 Gly Ile Phe Tyr Ser Ile Ser Trp Ser Gly Leu Leu Val Gly Tyr Ala

 405 410 415
 Ile Glu Val Leu Pro Tyr Lys Leu Arg Ala Lys Gly Leu Met Ile Met
 420 425 430
 Asn Leu Thr Val Gln Ala Ala Leu Thr Leu Asn Thr Tyr Ala Asn Pro
 435 440 445
 Val Ala Phe Asp Ala Phe Glu Gly His Ser Trp Lys Leu Tyr Ile Ile
 450 455 460
 Tyr Thr Ile Trp Ile Phe Leu Glu Leu Cys Phe Val Trp Lys Met Tyr

 465 470 475 480
 Ile Glu Thr Lys Gly Pro Thr Leu Glu Glu Leu Ala Lys Ile Ile Asp
 485 490 495
 Gly Asp Glu Ala Ala Val Ala His Val Asp Ile Lys Gln Val Glu Lys
 500 505 510
 Glu Thr His Ile Asn Glu Glu Lys Ser Val
 515 520
 <210> 30
 <211> 554
 <212> PRT
 <213> Sporotrichum thermophile
 <400> 30
 Met Ser Ser Ser Glu Lys Glu Ala Thr Gly Pro Val Ala Ala His Val

 1 5 10 15
 Gly Asn Leu Ala Thr Thr Gln Asp Val Glu Lys Ile Glu Ala Pro Val
 20 25 30
 Thr Trp Lys Ala Tyr Leu Ile Cys Ala Phe Ala Ser Phe Gly Gly Ile
 35 40 45
 Phe Phe Gly Tyr Asp Ser Gly Tyr Ile Asn Gly Val Leu Ala Ser Lys
 50 55 60
 Leu Phe Ile Asn Ala Val Glu Gly Ala Gly Lys Asp Ala Ile Ser Glu

65 70 75 80
 Ser His Ser Ser Leu Ile Val Ser Ile Leu Ser Cys Gly Thr Phe Phe
 85 90 95
 Gly Ala Leu Ile Ala Gly Asp Leu Ala Asp Phe Ile Gly Arg Lys Tyr
 100 105 110
 Thr Val Ile Leu Gly Cys Leu Ile Tyr Ile Ile Gly Cys Val Ile Gln
 115 120 125
 Ile Ile Thr Gly Leu Gly Asn Ala Leu Gly Ala Ile Val Ala Gly Arg

 130 135 140
 Leu Ile Ala Gly Ile Gly Val Gly Phe Glu Ser Ala Ile Val Ile Leu
 145 150 155 160
 Tyr Met Ser Glu Ile Cys Pro Arg Lys Val Arg Gly Ala Leu Val Ala
 165 170 175
 Gly Tyr Gln Phe Cys Ile Thr Ile Gly Leu Met Leu Ala Ser Cys Val
 180 185 190
 Val Tyr Gly Thr Gln Asn Arg Gln Asp Thr Gly Gln Tyr Arg Ile Pro

 195 200 205
 Ile Gly Ile Gln Phe Ile Trp Ala Leu Ile Leu Gly Gly Gly Leu Leu
 210 215 220
 Cys Leu Pro Asp Ser Pro Arg Tyr Phe Val Lys Arg Gly Arg Leu Ala
 225 230 235 240
 Asp Ala Thr Ser Ala Leu Ser Arg Leu Arg Gly Gln Pro Glu Asp Ser
 245 250 255
 Glu Tyr Ile Gln Val Glu Leu Ala Glu Ile Val Ala Asn Glu Glu Tyr

 260 265 270
 Glu Arg Gln Leu Ile Pro Ser Thr Thr Trp Phe Gly Ser Trp Ala Asn
 275 280 285
 Cys Phe Lys Gly Ser Leu Phe Lys Ala Asn Ser Asn Leu Arg Lys Thr
 290 295 300
 Ile Leu Gly Thr Ser Leu Gln Met Met Gln Gln Trp Thr Gly Val Asn
 305 310 315 320

Phe Ile Phe Tyr Tyr Ser Thr Pro Phe Leu Lys Ser Thr Gly Ala Ile

325 330 335

Asp Asp Pro Phe Leu Met Ser Met Val Phe Thr Ile Ile Asn Val Phe

340 345 350

Ser Thr Pro Ile Ser Phe Tyr Thr Val Glu Arg Phe Gly Arg Arg Thr

355 360 365

Ile Leu Phe Trp Gly Ala Leu Gly Met Leu Ile Cys Gln Phe Leu Val

370 375 380

Ala Ile Val Gly Val Thr Val Gly Phe Asn His Thr His Pro Ala Pro

385 390 395 400

Thr Ala Asp Asp Pro Glu Ala Thr Leu Ala Asn Asn Ile Ser Ala Val

405 410 415

Asn Ala Gln Ile Ala Phe Ile Ala Ile Phe Ile Phe Phe Phe Ala Ser

420 425 430

Thr Trp Gly Pro Gly Ala Trp Ile Val Ile Gly Glu Ile Phe Pro Leu

435 440 445

Pro Ile Arg Ser Arg Gly Val Gly Leu Ser Thr Ala Ser Asn Trp Leu

450 455 460

Trp Asn Thr Ile Ile Ala Val Ile Thr Pro Tyr Met Val Gly Glu Asp

465 470 475 480

Arg Gly Asn Met Lys Ser Ser Val Phe Phe Val Trp Gly Gly Leu Cys

485 490 495

Thr Cys Ala Phe Val Tyr Thr Tyr Phe Leu Val Pro Glu Thr Lys Gly

500 505 510

Leu Ser Leu Glu Gln Val Asp Lys Met Met Glu Glu Thr Thr Pro Arg

515 520 525

Thr Ser Ala Lys Trp Lys Pro Thr Thr Thr Phe Ala Ala Ser His Pro

530 535 540

Thr Asp Leu Lys Gln Gly Glu Ala Ala Val

545 550

<210> 31

<211> 537

<212> PRT

<213> Sporotrichum thermophile

<400> 31

Met Gly Thr Ser Arg Asp Glu Lys Glu Thr Val Val Ala Asp His Ala

1 5 10 15

Asp Asp Asp Ala Leu Arg Glu Ala Asp Leu Ala Val Gln Val Ala His

20 25 30

Asp Ala Asp Gly Thr Val Tyr Ser Pro Trp Ser Leu Arg Met Ile Arg

35 40 45

Leu Tyr Leu Val Leu Ser Leu Ser Tyr Leu Cys Gly Cys Leu Asn Gly

50 55 60

Tyr Asp Gly Ser Leu Met Gly Gly Leu Asn Gly Met Thr Ser Tyr Gln

65 70 75 80

Arg Tyr Phe His Met Ser Thr Ala Gly Ser Thr Thr Gly Leu Ile Phe

85 90 95

Ala Met Tyr Asn Ile Gly Ser Val Ala Ala Val Phe Phe Thr Gly Pro

100 105 110

Val Asn Asp Tyr Phe Gly Arg Arg Trp Gly Met Phe Val Gly Ala Leu

115 120 125

Leu Val Ile Val Gly Thr Cys Val Gln Ala Pro Cys Thr Thr Arg Gly

130 135 140

Gln Phe Leu Ala Gly Arg Phe Val Leu Gly Phe Gly Val Ser Phe Cys

145 150 155 160

Cys Val Ser Ala Pro Cys Tyr Val Ser Glu Met Ala His Pro Lys Trp

165 170 175

Arg Gly Thr Leu Thr Gly Leu Tyr Asn Cys Thr Trp Tyr Ile Gly Ser

180 185 190

Ile Val Ala Ser Trp Val Val Tyr Gly Cys Ser Tyr Ile Asp Thr Leu

195 200 205

Asp Ala Trp Arg Ile Pro Ile Trp Cys Gln Met Val Thr Ser Gly Leu

210 215 220
 Val Cys Leu Gly Val Phe Trp Leu Pro Glu Ser Pro Arg Trp Leu Met
 225 230 235 240
 Ala Gln Asp Arg His Asp Asp Ala Ala Arg Val Leu Ala Thr Tyr His
 245 250 255
 Gly Glu Gly Arg Ala Asp His Pro Leu Val Lys Leu Gln Met Gln Glu
 260 265 270
 Met Met Asn Gln Ile Ser Thr Glu Ala Ser Asp Lys Lys Trp Tyr Asp

 275 280 285
 Tyr His Glu Leu Trp Asn Thr His Ser Ala Arg Arg Arg Leu Ile Cys
 290 295 300
 Val Ile Gly Met Ala Val Phe Gly Gln Ile Ser Gly Asn Ser Leu Ser
 305 310 315 320
 Ser Tyr Tyr Leu Val Asn Met Leu Lys Ser Ala Gly Ile Thr Glu Glu
 325 330 335
 Arg Arg Val Leu Ala Leu Asn Gly Val Asn Pro Ala Leu Ser Phe Leu

 340 345 350
 Gly Ala Ile Leu Gly Ala Arg Met Thr Asp Val Val Gly Arg Arg Pro
 355 360 365
 Leu Leu Leu Tyr Thr Ile Val Phe Ala Ser Val Cys Phe Ala Val Ile
 370 375 380
 Thr Gly Thr Ser Lys Met Ala Thr Asp Asp Pro Thr Arg Thr Ala Ala
 385 390 395 400
 Ala Asn Ala Thr Ile Ala Phe Ile Phe Ile Phe Gly Ile Val Phe Ser

 405 410 415
 Phe Gly Trp Thr Pro Leu Gln Ser Met Tyr Ile Ala Glu Thr Leu Pro
 420 425 430
 Thr Ala Thr Arg Ala Lys Gly Thr Ala Val Gly Asn Phe Ser Ser Ser
 435 440 445
 Val Ala Ser Thr Ile Leu Gln Tyr Ala Ser Gly Pro Ala Phe Glu Gly
 450 455 460

Ile Gly Tyr Tyr Phe Tyr Leu Val Phe Val Phe Trp Asp Leu Ile Glu

465 470 475 480

Gly Ala Ile Met Tyr Phe Tyr Phe Pro Glu Thr Lys Asp Arg Thr Leu

485 490 495

Glu Glu Leu Glu Glu Val Phe Ser Ala Pro Asn Pro Val Lys Lys Ser

500 505 510

Leu Glu Lys Arg Ser Ala Gln Thr Val Leu Asn Thr Val Gly Ala Ala

515 520 525

Gln Asn Glu Lys Leu Ala Arg Asp Val

530 535

<210> 32

<211> 566

<212> PRT

<213> Sporotrichum thermophile

<400> 32

Met Ala Val Phe Ala Met Gly Trp Gln Lys Pro Asp Asn Val Ala Gly

1 5 10 15

Ser Ser Ala Pro Ala Ile Met Val Gly Leu Phe Val Ala Thr Gly Gly

20 25 30

Leu Leu Phe Gly Tyr Asp Thr Gly Ala Ile Asn Gly Ile Leu Ala Met

35 40 45

Asp Thr Phe Lys Glu Asp Phe Thr Thr Gly Tyr Thr Asp Lys Gln Gly

50 55 60

Lys Pro Gly Leu Tyr Ala Ser Glu Val Ser Leu Ile Val Ala Met Leu

65 70 75 80

Ser Ala Gly Thr Ala Thr Gly Ala Leu Leu Ser Ala Pro Met Gly Asp

85 90 95

Arg Trp Gly Arg Arg Leu Ser Leu Ile Val Ala Ile Gly Val Phe Cys

100 105 110

Val Gly Ala Ile Ile Gln Val Cys Ala Thr Asn Val Ala Met Leu Val

115 120 125

Val Gly Arg Thr Leu Ala Gly Ile Gly Val Gly Val Val Ser Val Leu
130 135 140
Val Pro Leu Tyr Gln Ser Glu Met Ala Pro Lys Trp Ile Arg Gly Thr
145 150 155 160
Leu Val Cys Ala Tyr Gln Leu Ser Ile Thr Ala Gly Leu Leu Ala Ala
165 170 175
Ala Thr Val Asn Ile Leu Thr Tyr Lys Leu Lys Ser Ala Ala Ala Tyr
180 185 190
Arg Ile Pro Ile Gly Leu Gln Leu Thr Trp Ala Leu Val Leu Ala Leu
195 200 205
Gly Leu Val Ile Leu Pro Glu Thr Pro Arg Tyr Leu Val Lys Arg Gly
210 215 220
Leu Lys Glu Ala Ala Ala Leu Ser Leu Ser Arg Leu Arg Arg Leu Asp
225 230 235 240
Ile Thr His Pro Ala Leu Ile Glu Glu Leu Ala Glu Ile Glu Ala Asn
245 250 255
His Glu Tyr Glu Met Ala Leu Gly Pro Asp Thr Tyr Lys Asp Ile Ile
260 265 270
Phe Gly Glu Pro His Leu Gly Arg Arg Thr Leu Thr Gly Cys Gly Leu
275 280 285
Gln Met Leu Gln Gln Leu Thr Gly Val Asn Phe Ile Met Tyr Tyr Gly
290 295 300
Thr Thr Phe Phe Tyr Gly Ala Gly Ile Gly Asn Ala Phe Thr Val Ser
305 310 315 320
Leu Ile Met Gln Val Ile Asn Leu Val Ser Thr Phe Pro Gly Leu Phe
325 330 335
Val Val Glu Ser Trp Gly Arg Arg Lys Leu Leu Ile Val Gly Ser Val
340 345 350
Gly Met Ala Ile Cys Gln Leu Leu Ile Ala Ser Phe Ala Thr Ala Ser
355 360 365
Gly Asn Asp Asn Lys Pro Thr Gln Asn Gln Ile Leu Ile Ile Phe Val

370 375 380
 Ala Ile Tyr Ile Phe Phe Phe Ala Ala Ser Trp Gly Pro Val Val Trp
 385 390 395 400
 Val Val Thr Ser Glu Ile Tyr Pro Leu Lys Val Arg Ala Lys Ser Met
 405 410 415
 Ser Ile Ser Thr Ala Ser Asn Trp Val Leu Asn Phe Gly Ile Ala Tyr
 420 425 430
 Gly Thr Pro Tyr Leu Val Asp Thr Ser Asp Gly Ser Pro Asp Leu Gly

 435 440 445
 Ser Arg Val Phe Phe Val Trp Gly Ala Phe Cys Ile Leu Ser Ile Ala
 450 455 460
 Phe Val Trp Tyr Met Val Tyr Glu Thr Ser Lys Ile Ser Leu Glu Gln
 465 470 475 480
 Ile Asp Glu Met Tyr Glu Arg Val Ala His Ala Trp Asn Ser Arg Ser
 485 490 495
 Phe Glu Pro Ser Trp Ser Phe Gln Gln Met Arg Asp Phe Gly Phe Ser

 500 505 510
 Asp Ser Gly Ile Pro Pro Ala Glu Pro Gln Leu Glu Leu Gln Gln Ser
 515 520 525
 Asn Ala Ser Thr Ser Gln Ser Asp Thr Gly Gly Ser Ser Ala Thr His
 530 535 540
 Ala Thr Ala Ala Asn Pro Gln Asp Ala Lys Met Val Ser Gln Leu Ala
 545 550 555 560
 Asn Ile Asp Leu Ser Tyr
 565

<210> 33

<211> 1404

<212> DNA

<213> *Pichia stipitis*

<400> 33

atgaagtatt ttcaaatctg gaaatcaggc aaacaagtaa gctacgtgt tacattcact 60
 tgtgaattgg catttattct ttttggatt gaacagggta ttattggtaa tcttattaac 120

aaccaggact tcctaaacac ttttggaac ccacccgta gttatttagg tattatcggt 180
tctatctata ccttaggggtg ttttttgggt tgtgttatga acttcttcat tggatgacga 240
atgggcagaa gaagcaaaat tgccttctca atgacagtta tcacaattgg tgttgctctt 300
caatgtagtt ccttttcagt tgaacaattg atgattggaa gatttatcac tgggcttga 360
actggttggg aaacttctac ttgtccaatg tatcaggcag aactttcacc tccaaaagtt 420

agaggacgtt tgggtgtctc agaagcattg tttgttggag ttggtttaat ctatgcata 480
tggtttgatt atgctcttct tttcacttct ggtcctattg catggagact tcctcttgcc 540
tctcagattg tgttcgctt tgttgttttc tgtttcactt tcacaatacc cgaatccct 600
agatacatgt tttaaaagg agagaaagaa gaagccaaa gaattttatc ttatgtctt 660
ggaaagccag gagatcatcc tgacattctt aaggaaatga atgatattaa tgatgtgtt 720
atgttgaaa cttcagaagg agctttctcg tgggcaaac tttcaagcc cgataaggca 780
agaactggat acagagtctt cttggcatatc atgagcatgt ttgcgcaaca gttgagtgg 840
gttaatgtag ttaattacta tattacattt gttttgatta acagtgttg catcgaagac 900

aacttggccc taattcttgg tgggtgttgc gtcactgtt tcaactgttg ttcattagtt 960
cctactttct ttgtgatag gatgggaaga agattgcctt cagcagttgg agcttttggc 1020
tgttgtgtt gtatgatgt aatttcaatc ttattaagtt ttcaagaca tccaaagttg 1080
aagaagagca gtggagctgg tgcgttggt ttttttttcg tttccaact tgtcttcggc 1140
tccactggta attgtattcc atggctgatg atttcagagc ttatccccct tcatgcacgt 1200
gctaaaggat cttcattagc tacatcaagt aactggcttt ggaatttctt tgttgttgag 1260
atcactccaa ctatcattga aaagttgaag tggaaagcat attgatctt tatgtgtgc 1320
aacttctct tcttaccat gttttacttt tttttcccg agacaaagaa cttacttta 1380

gaagccattg acgatttggt ctca 1404

<210> 34

<211> 1653

<212> DNA

<213> *Pichia stipitis*

<400> 34

atgtctcac aagatttacc ctgggtgct caaacccaa tcatggttc ttccatctc 60
gaagataaag ttgagcaaag ttcttctca aatagccaaa gtgatttagc ttccattcca 120
gcaacaggta tcaaagccta tctcttgggt tgtttctct gcatgttgg tgccttgggt 180
ggcttcgtat tccgtttcga taccgttaca atttccggt tccttaatat gtctgatttc 240
ctttccagat ttgttcaaga tggttctgaa ggaaaatatt tgtctgatat cagagtcggt 300

ttgattgttt ccatttttaa cattggttgt gcaattggtg gtattttcct ttctaagata 360

ggagatgttt acggtagaag aattggtatc atttcagcta tggttgtcta cgtcgtcgtt 420

attatcatcc agatctcgtc ccaagacaag tggtagcaac ttacaattgg acgtggagtt 480

acaggattag ctgttggtag tgtttcagtg ttgtctccaa tgttcattag tgaaagtgtt 540

ccaaagcatt tgagaggtag ttgtgtatc tgttagcaat tatgtatcac cttaggtatt 600

ttcattggtt actgtgtcac ttatggaacc aaagatttaa atgattcaag acaatggaga 660

gttcccttgg gcttatgctt cctttgggct attttcttag ttgtcggtag gttggctatg 720

ccagaatccc caagattctt aattgaaaag aagagaatcg aagaagccaa gaagtcctt 780

gcaagatcca acaagtatc tccggaagat ccaggtgtct aactgaact tcaattgatt 840

caggctggta ttgacagaga agctcgtgca ggttctgctt cgtggatgga attgatcact 900

ggtaagccag ctattttcag aagagttatc atgggaatta tcttcagtc ttgcaacaa 960

ttaactgggt tcaactattt cttctattac ggaactacaa tcttccaagc tgttggtttg 1020

caagattcct tccagacttc catcatctta ggtacagtca actttcttcc tacatttgtt 1080

ggtatttggg ccattgaaag atttggaaga agacaatgtt tgttagtcgg ttctgctggt 1140

atgttcgttt gtttcatcat ttactccgtt attggtacaa ctcatctgtt cattgatgga 1200

gtagtagata acgacaacac ccgtcaactg tctggtaatg ctatgatctt taccacttgt 1260

ttgttcatct tcttctttgc ctgtacatgg gctggagggt tttttacat catttccgaa 1320

tcataccat tgagaatcag atccaaggca atgtctattg ctactgtgc taactggatg 1380

tggggcttct tgatttctt ctgcactcca ttcatgtta atgcatcaa ctccaagttc 1440

ggctttgtgt ttactgggtg ttactcttct tcgttcttct atgtctactt ctttgcagc 1500

gaaaccaaag gtttgcgtt ggaagaagt gatgagttgt acgtgaagg tattgcacca 1560

tggaagtctg gtgcatgggt tcttcttctt gcccaacaac aaatgcaaaa ctccacttat 1620

ggtgccgaag caaaagagca agagcaagtt tag 1653

<210> 35

<211> 2025

<212> DNA

<213> Neurospora crassa

<400> 35

atggcgtcga acccaacgaa caccgcggcc cctacgggtg gccttaccga gaagaagcat 60

gaccgccgtt caacatcgtc cgaatccgtc tcgggaaccg gggttgcgga acatgcagac 120

cgcaccggca cttttaacca gaacgtcga ctagaggctt caaaaagat agcgaatcct 180

ttggccggtc taagccctca gcgtctcgag gccatgggag aagaatatgc aatgatggcc 240
 ggtctcacca gcgaggagga catcagggcc ttctgactcg gagccagaat cgccggcgat 300
 gagagcaact acgacctcat cccggagctt actgaacggg agaaagaggt gttggtgcgc 360
 gaaacaactc acaagtggtc taaccacccc atgctttact gggttgttgt catttgcctc 420
 ctatgcgcgc ccgtccaagg aatggacgag acggctcgta acggcgccca gctcttctac 480
 aaggacaagt ttggcattgg tactgatagc cagagagaca cttggcttct gggtctcgtc 540

 aactcagcgc cctacctttg ctgtgccttt atcggtctgt ggctcactga accgatgaac 600
 agaactcttg gcagacgagg caccatcttt gtttcttgca tcatctcagc cgtagcttgc 660
 ttccaccagg cctttaccaa cacgtgggtg cacatgttca tcgccgctt ctacctcggc 720
 cttggcatcg gtcccaagtc agccaccacc cccatcttcg ccgccgaatg ctccccctcc 780
 aagctccgcg gtgcgtggt catgcaatgg cagatgtgga ccgccttcgg tatcatggtc 840
 ggctacattg ccgatctcgc tttctacttc gttcccgatc acggcatcgg cttgggtctg 900
 aactggcgtc tgatgatggg ctccgccatg attcccgccg tcatcgctgt ctgcctcgcc 960
 ttctctgccc cggagtcgcc ccgttggtag ctgagcaagg gccgacacca agacgccttc 1020

 ggggcgctct gccgcctcgc ttctgaaaag gtccaagccg cccgcgacct cttctacacc 1080
 cacaccctcc tagaagccga gaagcaagcc atgtcgggcg tcaagaagg taaccgcttt 1140
 aaggagctct tcaccgtcgc tcgtaaccgc aacgcggtca ttgcctcgtc gggactcatg 1200
 ttcatgcagc agttctcggc cgtaacatc atcgctact actcctcggc ggtcttccga 1260
 gacgccggtc tcagcgacgt ctgagcactg gccgcctcgc tcgggttttg cgtcgtcaac 1320
 tggtgtttg ccatcccggc catgtacacc atcgacactt tcggccggcg caacctgctg 1380
 ctgaccacct tcccgtcat gtccctcttc ctcttcttca ccggcttcag cttttggatc 1440
 cccgaggact ccaaagccca catcggtcgc atcgcgctag gcatttactt gttcgcatg 1500

 gtctactccc ccggtgaagg gccggtgccg ttacttact cggccgaggc ctaccgctg 1560
 tacatccggc caatcgcat gtccctcgcc acggcgacta cctggttctt caatttcatt 1620
 ctttccatca cctggcctag gatggtcacg gccttcaagc cgcagggcgc gtttggctgg 1680
 tatgcagggt ggaatatcat tgggtttctc ttaccctgt tcttggctcc cgagaccaag 1740
 ggcaagacgc tggaggagct cgatcacgtg ttgacgtgc cgttgaagaa gttggtcaga 1800
 tacggggcgg atcagagctt gtggtttttc cacaggggaa agaattgaaa tggaatgagg 1860
 ccgacggcgc ctagtcgga gatgtatcat ggggatgcgg agcgatgaa cgaggtggtt 1920
 agcgggcagc agcttgggga gggtagaggg gagaagaggt ggaacaagga acaagagagg 1980

 gaagggggga ttatgggacg aggggatgct gctgggaagg tgtag 2025

<210> 36

<211> 1458

<212> DNA

<213> Neurospora crassa

<400> 36

atgtcgccea tcgtcgtgac cgaccaatac ctcacctact tcaacaaccc ccatgatata 60
atccaaggag ccatcggtc tgcccttgct gctggctccg tcgtcggttc cgccatcgcc 120
ggtctctttt ccgacaagat cggtcgtcgt gactccatct ttttcgcctg cttcttcttg 180
ctcattggta cctccgtcca gggtgcctgc aagaactatg gccagctcat cgccggccgt 240
gtgctcaacg gccttacgt cggtcact tctcccagg ttcctgtga ccttgccgag 300
atcgccaagg cagagaagcg tggttccttg gtcacatcc agcaactcgc catcgagttt 360

ggtatcttga tcatgtactt tatcggtac ggctgtcgt cgatcgaggg ccctgcttcg 420
ttccggaccg cttggggcat tcagttatc ccttgctttt tctcatggt cggcttccc 480
ttcttgcta ggtcgccag atggctggcc aaggctgga gggaccagga ggccattgct 540
gtcctggta acatccaggc tgatggcaac gttgatgacc cgagagtcgt tgctgagtgg 600
gaggagattg tcaccgttat gaacccgag cgtgaggccg gtaagggatg gaggaagttt 660
gtcaagaacg gcatgtgaa gcgaacctg gctggcatga ctgtacaggc ttggcagcaa 720
ctcgccggcg ccaacgtaat cgtctactac ctaacctaca tcgcccacat ggccggactc 780
acaggcaacg tcgcatggt gacctcgggc atccaatacg ccgttttcat catcttcacc 840

ggcgtcatgt ggtcttcat cgacaagacc ggtcgtcgca ccttttagt ttacggcgcc 900
ttgggaatgg cttctgcca ctttgcgtc ggcgcgctca tggcgcgca ccacgacaac 960
gttccggacg gcgtcggcgg caacccaac attgtcatta gcgtgcacaa gggcgcgccc 1020
gccaacacgg tcatcctgtt ctgtacctg ctattgtcg tctacgctt gacgctcgt 1080
cccgtctgct ggatctacg cgccgaggtc tggctgttg gactcgcgc tacgggcatg 1140
tccatggctg ccatgtcaa ctgggtgtt aactttgcgc tgggcatgtt cagccgccc 1200
gcgtttgtca atattacgtg gaagctgtt atcattttcg ggggtctttg cgtcacggcg 1260
gcggtctggt tctggtgtt ttaccggag acgtgtgga agacgtgga ggagattgag 1320

atcctgtttg gtgatcagg tcctaagccg tggaagacaa agaagggcga gtcgagactt 1380
acggcgga ttgagctgt caagcgagg aagacgttg agcacgagat tgaggtgcat 1440
gagcatgaga aggtttag 1458

<210> 37

<211> 1407

<212> DNA

<213> *Pichia stipitis*

<400> 37

atgaagtatt ttcaaatctg gaaatcaggc aaacaagtaa gctacgtgt tacattcact 60
tgtgaattgg catttattct ttttggatt gaacagggtt ttattggtta tcttattaac 120
aaccaggact tctaaacac ttttggaaac cccaccggta gttatttagg tattatcggt 180
tctatctata ccttaggggtg ttttttgggt tgggttatga acttcttcat tggatgatga 240

atgggcagaa gaagcaaat tgcctcctca atgacagtta tcacaattgg tgttgcctct 300
caatgtagtt ccttttcagt tgaacaattg atgattggaa gatttatcac tgggcttggg 360
actggttggg aaacttctac ttgtccaatg tatcaggcag aactttcacc tccaaaagtt 420
agaggacgtt tgggtgtgctc agaagcattg tttgttggag ttggtttaat ctatgcata 480
tggtttgatt atgctcttct tttcacttct ggtcctattg catggagact tctcttggc 540
tctcagattg tgttcgcctt tgttgttttc tgtttcactt tcacaatacc cgaatccct 600
agatacatgt tttacaaagg agagaaagaa gaagccaaaa gaattttatc ttatgtcttt 660
ggaaagccag gagatcatcc tgacattctt aaggaatgga atgatattaa tgatgctgtt 720

attttggaaa cttcagaagg agctttctcg tgggcaaac ttttcaagcc cgataaggca 780
agaactggat acagagtctt cttggcatac atgagcatgt ttgcgcaaca gttgagtgg 840
gttaatgtag ttaattacta tattacattt gttttgatta acagtgttgg catcgaagac 900
aacttggccc taattcttgg tgggtttgcc gtcactgtt tcaactgttgg ttcattagtt 960
cctactttct ttgtgatag gatgggaaga agattgcctt cagcagttgg agcttttggc 1020
tgtggtgttt gtatgatgt aatttcaatc ttattaagtt ttcaagaca tccaaagttg 1080
aagaagagca gtggagctgg tgcgttggct tttttttcg ttttccaact tgtcttcggc 1140
tccactggta attgtattcc atggctgatg atttcagagc ttatccccct tcatgcacgt 1200

gctaaaggat cttcattagc tacatcaagt aactggcttt ggaatttctt tgttgttgag 1260
atcactccaa ctatcattga aaagtgaag tggaaagcat attgatctt tatgtgctgc 1320
aacttctct tcttaccat gttttacttt ttttttccg agacaaagaa cttacttta 1380
gaagccattg acgatttgtt ctcataa 1407

<210> 38

<211> 1653

<212> DNA

<213> *Pichia stipitis*

<400> 38

atgtcctcac aagatttacc ctcggtgct caaaccccaa tcatggttc tccatcctc 60
gaagataaag ttgagcaag ttgctcctca aatagccaaa gtgatttagc ttccattcca 120
gcaacaggta tcaaagccta tctcttggtt tgtttcttct gcatgttggt tgcctttggt 180

ggattcgtat tgggtttcga taccggtaca atttccggtt tccitaatat gtctgatttc 240
ctttccagat ttgttcaaga tgggtctgaa ggaaaatatt tgtctgatat cagagtcggt 300
ttgattgttt ccatTTTTaa cattggttgt gcaattggtg gtattttcct ttctaagata 360
ggagatgttt acggtagaag aattggtatc atttcagcta tggttgtcta cgtcgtcgg 420
attatcatcc agatctcgtc ccaagacaag tggtagcaac ttacaattgg acgtggagtt 480
acaggattag ctgttggtac tgtttcagtg ttgtctccaa tgttcattag tgaaagtgt 540
ccaaagcatt tgagaggtag tttggtatc tgttagcaat tatgtatcac cttaggtatt 600
ttcattggtt actgtgtcac ttatggaacc aaagatttaa atgattcaag acaatggaga 660

gttcccttgg gtttatgttt cctttgggct attttcttag ttgtcggat gttggctatg 720
cctgaatccc caagattctt aattgaaaag aagagaatcg aagaagccaa gaagtcctt 780
gcaagatcca acaagtatc tccagaagat ccagggtgtc aactgaagt tcaattgatt 840
caggctggta ttgacagaga agctcgtgca ggttctgctt catggatgga attgatcact 900
ggtaagccag ctattttcag aagagttatc atgggaatta tcttacagtc ttgcaacaa 960
ttaactgggt tcaactattt ctctattac ggaactacaa tcttcaagc tgttggtttg 1020
caagattcct tccagacttc catcatctta ggtacagtca actttcttc tacatttgtt 1080
ggtatttggg ccattgaaag atttgaaga agacaatgtt tgttagtcgg tctcgtggt 1140

atgttcgttt gttcatcat ttactccatt attggtacaa ctcatTTgtt cattgatgga 1200
gtagtagata acgacaacac ccgtcaactg tctggtaatg ctatgatctt tatcatttgt 1260
ttgttcatct tcttctttgc ctgtacatgg gctggagggtg tttttacat catttccgaa 1320
tcataccat tgagaatcag atccaaggca atgtctattg ctactgctgc taactggatg 1380
tggggcttct tgatttctt ctgcactcca ttcatTTta atgcatcaa cttaagttc 1440
ggcttttgtt ttactggttg ttactcttt tcgttcttct atgtctactt ctttgcagc 1500
gaaaccaaag gtttgtcgtt ggaagaagt gatgagttgt acgtgaagg tattgcacca 1560
tggaagtctg gtgcatgggt tcttcttct gcccaacaac aaatgcaaaa ctccattat 1620

ggtgccgaag caaaagagca agagcaagtt tag 1653

<210> 39

<211> 1641

<212> DNA

<213> *Pichia stipitis*

<400> 39

atggaattct ccagtgttga aaaaagtgt gaaactgttt cctatacgtc gcaggtcagc 60
gcaagcggct ctgcaaagac caacagctac ctggcctca gaggccacaa acttaatttt 120
gtgtgtctctt gttttgtcgg tgttggtttc ttacttttcg gttacgatca aggtgtcatg 180
ggttcattgt tgaccttgcc atccttcgaa aacactttcc cggccatgaa ggctagcaac 240
aacgtctacct tacaaggcgc cgttattgca ctttatgaaa tcggttgat gtcttcttct 300
ttagcaacca ttaccttgg tgacagattg ggtagattga agatcatgtt tattggctgt 360

gtaattgtct gtattggcgc tgctttgcaa gcttctgctt tcaactattgc tcacttgact 420
gttgctagaa ttatcactgg ttaggtaca ggtttcatca cttctactgt tccagtttac 480
caatcggagt gctctccagc caagaaaaga ggacagtga tcatgatgga aggttctctt 540
atcgcccttg gcattgccat ctcatctgg attgactttg gattttactt ttgagaaac 600
gatggtttgc actcctcggc ttcttggaga gcacctatcg cgcttcaatg tgtcttcgct 660
gtcttgttga tttccacagt cttcttcttc ccagaatctc caagatggtt gctcaacaaa 720
ggtaggaccg aagaagctag agaagttttt tctgtctttt acgacttgcc agccgactct 780
gaaaagattt ctattcaat tgaagaaatt caagctgcta tagatttaga aagacaagcc 840

ggagaagggt tcgtacttaa ggaattgttc actcagggcc cagccagaaa ctgcagcgt 900
gtggccttgt catgttggtc tcaataatg caacaaatca ctggtattaa cattattacg 960
tactatgtcg gtactatttt tgaatcatac attggatga gtccatttat gtcaagaatc 1020
ttggctgcct tgaacggtag tgaatatttc cttgtctctc ttattgcttt ctacaccgtc 1080
gaaagattag gtagaagatt cttttgttc tggggtgcca tcgccatggc tcttgtcatg 1140
gttggtttta ctgttaccgt taaacttgcc ggtgaaggca acacccatgc tgggtgcggt 1200
gtgtgtgttc ttttgtttgc attcaactca ttcttcggcg tctcctggtt aggtggatcc 1260
tggttgttac cacctgaatt gttgtctttg aaattgagag ctcttggtgc tgctttgtcg 1320

accgcttcta actgggcttt taacttcatg gttgtcatga tcaactcctg cggtttccaa 1380
agtattgggt cctacaccta cttatcttt gctgcatca atttgttgat ggctccggtc 1440
atctacttct tgtatccga aaccaagggt agatcgttgg aagaaatgga tatcattttc 1500
aaccaatgtc ctgtttggga gccatggaag gttgtccaaa ttgccagaga cctccctatt 1560
atgcactcag aagtcttga ccacgaaaag aatgtcatta ttaaaaaatc tagaatagag 1620
catgtcgaaa acatcagcta a 1641

<210> 40

<211> 1701

<212> DNA

<213> *Pichia stipitis*

<400> 40

atgcacggtg gtggtgacgg taacgataac acagaaatta ttgcagccag acgtctccag 60

atcgctggta agtctgggtg ggctgggtta gtcgcaaact caagatcttt cttcatcgca 120

gtctttgcat ctcttgggtg attggtctac ggttacaact aaggtatgtt cggtaaaatt 180

tccggtatgt atcattcttc caaagctatt ggtgttgaaa agattcaaga caatcctact 240

ttgcaaggtt tgttgacttc tattcttgaa ctgggtgcct gggttgggtg cttgatgaac 300

ggttacattg ctgatagatt gggtcgtaag aagtcagttg ttgtcgggtg tttcttcttc 360

ttcatcgggt tcattgtaca agctgttgct cgtgggtgta actacgacta catcttaggt 420

ggtagatttg tcgtcggtat tgggtgggtt attcttctta tgggtgtgcc attgtacaat 480

gtcgaaattt ctccaccaga aattcgtggt tctttggttg ctttgaaca attggtatt 540

actttcggta ttatgatttc ttactggatt acctacggta ccaactacat tgggtgtact 600

ggctctggtc aaagtaaage ttcttgggtg gttctatatt gtatccaatt ggttcagct 660

ttgctcttgg gtgttggat ctcttctcatg cctgagcttc caagatggtt gatgaacga 720

gacagagaag acgaatgttt gtccgttctt tccaacttgc gttccttgag taaggaagat 780

actcttgttc aaatggaatt ctttgaatg aaggcacaaa agttgttcga aagagaactt 840

tctgcaaagt acttccctca cctccaagac ggttctgcca agagcaactt cttgattggt 900

ttcaaccaat acaagtccat gattactcac tacccaacct tcaagcgtgt tgcagttgcc 960

tgtttaatta tgaccttcca acaatggact ggtgttaact tcattcttga ctatgtcca 1020

ttcatcttca gttcttttagg ttgtctgga aacaccattt ctcttttagc ttctggtgtt 1080

gtcggtatcg tcatgttctt tgctaccatt ccagctgttc tttgggtcga cagacttgg 1140

agaaagccag ttttgatttc cggtgccatt atcatgggtt tttgtcactt tgttgtggct 1200

gcaatcttag gtcagttcgg tggtaacttt gtcaaccact ccggtgctgg ttgggttgc 1260

gttgtcttcg ttggatttt cgctatcgggt ttcggttact cttgggttcc atgtgcttgg 1320

gtccttgttg ccgaagtctt cccattgggt ttgcgtgcta aggtgttctc tatcgggtgcc 1380

tcttctaact ggttgaacaa ctctcgtgtc gccatgtcta cccagattt tgttgctaag 1440

gctaagttcg gtgcttacat tttcttaggt ttgatgtgta ttttcggtgc cgcatacgtt 1500

caattcttct gtccagaaac taagggtcgt accttgaag aaattgatga acttttcggt 1560

gacacctctg gtacttccaa gatggaaaag gaaatccatg agcaaaagct taaggaagtt 1620

ggtttgcttc aattgctcgg tgaagaaaat gcttctgaat ccgaaaacag caaggctgat 1680

gtctaccacg ttgaaaaata a 1701

<210> 41

<211> 1656

<212> DNA

<213> *Pichia stipitis*

<400> 41

atgagagaag ttggtattct tgatgttgcc catggcaacg ttgtaactat aatgatgaaa 60

gatccagtag tatttttggg gatattttt gcatcccttg gaggtttgct ttttggttat 120

gatcaagggg ttattagtgg cattgtcaca atggaatctt ttggtgcaaa attccccaga 180

atttttatgg atgccgatta caagggttgg ttgtgtcta cttttttgct atgcgcatgg 240

tttggtctta ttattaatac tccaattgtt gataggtttg gaagacgtga ttctatcaca 300

atctcttggt ttatttttgg cattggttct gcgttccaat gtgctggcat taatacaagt 360

atgttatattg gtgggcgtgc tgttctgggt ctgcagtcg gtcaattaac catggtagtt 420

ccaatgtaca tgtcggaatt ggctcctcca tcggtgagag gtgggttggg tgtaattcag 480

caactttcga ttacaattgg tatcatgatt tcctattggg tggattatgg cactcatttt 540

attggaggta ctagatgtgc tcctagtcac ccataccaag gtgaaacttt taaccctaatt 600

tggtgatgtc ctccagggtg ctgctatggg caaagtgatg ccagtggag aattcctttt 660

gggtttcaga ttgtccagc agtgtgttg ggtattggaa tgatattttt cccaagatct 720

cccagatggt tactctctaa aggtcgcgac gaagaagctt ggagctcttt gaaatatctc 780

agaagaaaga gtcatgagga tcaagtcgaa agagagtttg ctgaaattaa ggcagaggtc 840

gtttatgaag acaagtacaa ggaaaagaga ttccctggta agactggagt tgctttaaca 900

cttactggat actgggatat tcttactact aaatctcact tcaagagagt ttttattgga 960

tcagctgtca tgttcttcca acaattcatt ggctgcaatg caataattta ttacgcacct 1020

acaattttca cacaattggg aatgaactct acaactactt ccttgcttgg tactggtctt 1080

tatggtattg ttaattgtct ttccaccctt ccagcagtgt tcttgatcga tagatgtgga 1140

agaaagactt tgtaaatggc aggtgctatt ggaactttta tttccttggg tattgtcggc 1200

gcaatcgttg gcaagtatgg cgatcgttta tctgaattca agacagcagg gagaactgca 1260

attgctttca ttttcattta tgatgtgaat ttctcgtaca gttgggctcc aattggatgg 1320

gttttaccct cagagatttt cccaatcggc atcagatcca atgcatctc cataactacc 1380

tcactacttt ggatgaataa ttttattatt ggcttgggtc ctccacatat gttagaaaca 1440

atgaagtggg gcacttacat tttttttgca gcgtttgcta ttattgcgtt ctttttccact 1500

tggtttatca tcccggaaac caagggagtt ccattggaag aaatggatgc cgtgtttggc 1560

gatactgcag cattgcagga aaagaatttg gttaccatta cgtcagtttc tgaatctgac 1620

gccaaaggatc gcaactcgat tgaaatgtca gaataa 1656

<210> 42

<211> 1590

<212> DNA

<213> *Pichia stipitis*

<400> 42

atggcatatc ttgattgggtt aacagctaga accaacaactt tcgggttgag gggcaagaag 60
 ttgagagcct tcatcactgt agtggctgtc actggtttct cattattcgg atatgatcaa 120
 gggttgatgt ccggaattat tactgctgat caattcaact ctgagtttcc cgccactaga 180
 aataacagta ctatccaagg tgccgtcacc tctgtttacg agcttggttg tttcttttgt 240
 gctgtgtttg ccttggttaag aggtgaaaga attggaagaa gacctcttgt gctttgtggc 300
 tcgcttatta tcatcttggg aacagttatt tctgtaaccg ctttccatcc aactgggtca 360

ttaggtcagt ttgttattgg tagagtatc actggtattg gtaatggtat gaatactgcc 420
 accattccag ttggcaatc ggaaatgtca agagctgaaa acagaggaag attggtcaac 480
 ttggaagggtt ccgttgtcgc tgtgggtaca tgtattgcct actggttgga tttcggtttg 540
 tcttatgtcg acaattcagt ttcttgagga tttccagttg ctttccaaat agtgtttgct 600
 tccgttttat ttgtgggaat gttgcaattg cccgactctc caagatggtt ggttgctaac 660
 cacagaagag cagaggctct tcaagtgttg tctgctttga aagacttgcc cgaagacgac 720
 gaagaaatc ttaatgaagc tgaagtattt caggaaagtg tagacaagtt tgctggacat 780
 gcttccgtca aggaagtgtt tactggtggt aagaccagc actggcaaag aatggttatt 840

ggatccagca cccaattctt tcagcagttc actggttgta acgtgccat ttactattcc 900
 actgtcttgt ttcaagacac tattggttta gaaagaagaa tggcattgat tatcgttggt 960
 gttttcgcaa ctgtctacgc cattttcaca attccttctt tcttcttggt cgatactctt 1020
 ggacgtagaa actgtttctt gattggtgct atgggacaag gtattgcatt cactatcacc 1080
 tttgcctgtt tgattgacga tactgaaaac aacccaagg gtgccgcagt tggtttattc 1140
 ttgtttattt gtttcttcgc cttcaccatc ttgccattgc catgggtata cccaccagaa 1200
 atcaatcctt tgagaactag aactatagct tctgcaattt ccaactgtac caactggatc 1260
 tgaactttg ctgttgttat gttcaccctt gtctttgtca ctaacaccag atggggagcc 1320

tatcttttct ttgctgtgat gaacttctt ttcgttcccta ttattttctt cttctacceca 1380
 gaaacagctg gaagatcgtt ggaagaaatc gatatcatct ttgcgaaggc attcgttgac 1440

aaaagacagc catggagagt tgctgcaacc atgccaaagt tgtccaacca cgaaattgaa 1500
gacgaagcca acagattggg cttgtttgac gatggtacat tgcacaagga agcatttgaa 1560
accaaagaaa acgcatccag cagctcttaa 1590

<210> 43

<211> 1689

<212> DNA

<213> Neurospora crassa

<400> 43

atggcgctc caaagttcct gggcctctca ggccgaccgc tctctctagc tgtctcgact 60
gtagccacca cgggcttctt tctcttcggc tatgaccaag gtgtcatgag cggcatcatt 120

accgccccg ctttcaaca cttcttcaca ccaaccaaag acaactcgac catgcagggt 180
ctcatcactg ccatctacga aattggatgc ttgattgggt ccatgttcgt cctctggacc 240
ggcgatttgt tgggtagacg caggaacatc atggtgggcg ctttcattat ggctctcggt 300
gtcattattc aggttacctg tcagctgga tccaaccctt ttgctcagct gtctctcggc 360
agagtctgca tgggtattgg caacggcatg aacacttcga ccattccac ttatcaagcc 420
gaatgtcaa agacatcgaa ccgcggctct ttgatctgca ttgaaggcgg tgcattgcc 480
tttggtaact tgattgctta ttggatcgac tatggtgcat cttacggtcc cgatgacctc 540
gtttggcgct tccccatgc tttccagctt ctcttcgcca tcttcactg cgtcccatg 600

ttttacctc ccgagtcgcc cagatggctc ctccagccatg gccggacca agaagctgac 660
aaggtcattg ctgccccg tggtacgag atcgatggtc ccgagacat ccaagagcgc 720
aacctcattg ttgactcctt gcgtgcctct ggaggtttcg gccaaaagag cactccctc 780
aaggccctct tcaactggcg caagaccag catctccgct gtctcttgc cggttccagc 840
tcccagttca tgcagcaagt tgggtggtgc aacgccgtca tctactactt cccattctg 900
ttccaggatt ctattggcga gtccacaac atgtccatgt tgctggcggt tatcaacatg 960
atcgtctact ccatcttcgc taccgtttcc tggttcgcca ttgagcgtgt cggctcgtct 1020
cgtctgttct tgatcggcac cgttggccag atgctctcca tggcatcgt cttcgccctg 1080

ttgatcccc acgacctat gaaggccgc ggtgccgagg tcggtctctt cacttacatt 1140
gcctttttcg gtgccattg gcttccctc cctggctct acccgccga ggttaacccc 1200
atccgcacac gtggaaaggc taacgccgtc tccacctgct ccaactggat gttcaactc 1260
ctcatcgta tggtcacccc catcatggtc gacaagattg gctggggaac ttacctctc 1320
ttcgcggtca tgaacggctg ctctcttccc atcatttact tctctaccc cgagactgcg 1380
aaccgctcg tcgaggagat cgacatcatc ttcgccaagg gcttcgtcga gaacatgctg 1440

tacgtcactg ccgccaagga gctgcctcac ctcaactgccg aggagatcga gtcctatgcc 1500
aacaagtatg gcctcgtcga ccgcgattcc aacggcgagg gcggcaaccg ccatgacgag 1560

gagaagacgc gcgaccgccc cgaccagagt gacagcgact cccccgtca cgtcgagatt 1620
gatgttgtcg acgagcacgg tgtcgagtcc ggcttcggtg atggtattaa caccaaggaa 1680
acacgttaa 1689

<210> 44

<211> 1626

<212> DNA

<213> *Neurospora crassa*

<400> 44

atggaattcg agcacgatca ctccgcctcc gacattgaga aggaggccgt cactgtggcc 60
cggccacagg gcgatgtcac ccgcgttgag gctcccgtaa cctcaaggc gtacatgatg 120
tgcgtctttg ccgctttcgg cggatatctt tttggctacg attcaggtaa catctctggt 180
gtcatgggca tgaagtactt tatcgaaacc atcaacggac ccggcgccac ctctctgcca 240

tccaaggaaa agtcgctcat caccitccatt ctctctgccg gaaccttctt tggcgccctc 300
atgggcggtg atctcgtcga ctgggttggc cgtcgtccta ccatcatctt cggtgcctc 360
gtcttcacg tcggtgttgt tctccagact gcctccaga gcttgggtct cattgtggcc 420
ggccgtctcg tcgttggtt cgggtgcggt ttcgtctcgg ccattatcat cctgtacatg 480
tctgagatcg cggcccgcaa ggtccgcggt gctatggtgt cgggctacca gtcttgcac 540
tgcttgggtc tgcctctggc ctctgtcgtt gactacggca cccagaaccg caccgacagc 600
ggctcttaca gaatcccgat tgggtctccag atggcctggg cctcattct tgctactggt 660
atctttttcc ttctgaatc cctcgcctt ttcgtcaaga agggcaagct cgacaaggcc 720

gccggcgtgc tctccgcct gcgcgaccag ccgctcgatt ccgactacgt cagggacgaa 780
cttgccgaga tcgttgccaa ccacgaattc gaaatgaccg tcgtcccta cggcaactac 840
ttccagcagt gggccaactg ctccgcggc tccatctggc aggttgggtc ttacctccgc 900
cgcaccattc tcggcacttc gatgcagatg atgcagcagt ggacgggaat caactttatc 960
ttttactttg gaaccacctt ctccagcag ctggcacca ttgacaacce ctctctgatg 1020
tctctggtca ctactctgt caacgtctgc tccacccca tctcttcta caccatggag 1080
aagctcggc gtcgtacct cctcatctgg ggcgtctcg gcatgctgat ctgcgagttt 1140
atcgtcgcca tcgttggta ctgcaggccg gatgatacca tggccatcaa ggccatgctc 1200

gccttcatct gcatctacat cttcttctt gctaccacct ggggccctgc ttcttgggtc 1260

gtcatcggcg aggttttccc tcttccatt cgtgccaagg gtgttgcct tccaccgcc 1320
tccaactggc tctgtaactg catcatcgcc gtcactc cctacatggt cgacgaggac 1380
aagggaacc tgggcccac ggtgttctac atctggggtg gcctctgcac ctgctgcttc 1440
atctacgcct acctgcttgt gcccagacc aagggcctca cgctcgagca ggtcgaccaa 1500
atgttttccg agtctacccc ccgacactcg accaagtga agcctcacac cacttatgct 1560
gtgagatgg gcatgaccga gaagactgtt gctggccacg ctgagaaccg cagcgatagc 1620
gagtaa 1626

<210> 45

<211> 1614

<212> DNA

<213> Neurospora crassa

<400> 45

atgggtcttt cgataggaaa taggatcctc cggaaaattg tcaaaatga ggccatggca 60
gaagatcccc cagagatcta tggctggcgt gtctatctcc tagcgtgctc tgctgcttc 120
ggcgccatgt ctttcggctg ggattcctcc gtcacggcg gcgtcatcga actcgaaccc 180
tttaaacacg actttggctt catcggcaac gataaagcca aggccaacct gggcgccaat 240
atcgtctcta cctccaage cggctgcttc ctgggtgcgc tgatcgctc acctataacc 300
gatcgcttcg gccgaagtgt gtgtctcatc gctgtctccc tggctgctcat catcggtatc 360
atcatgcaag ccgccgcctc aggcaacctc gcacccatgt acattggccg tttcgtcgcc 420

ggcgtggcg tcggcgccgc cagctgcac aacccgtct ttgtgtctga gaacgtccc 480
cgctcgatcc ggggtctgtt gacgggcctc taccaactct tcattgtcac cggcgcatg 540
atcgcatttt ggatcaacta ctccgtctct ctgcacttca agggcaaacc catgtacac 600
tccccgtcg ccatcaagg tcttccgcc ggcttttgt gcgtctgcat gtcctctgc 660
cacgaaagcc cgcgtggct ggcccgtcgt gaccgatggg aagaatgca gtctgtgctg 720
gcgcgcatcc gcaacctccc ccagaccac ccgtacatcg tcgacgagtt ccgcgagatc 780
caggaccagc tcgaacagga gcgtcgtctc caggcgacg ccacttactg ggacttgacc 840
cgcatatgt ggaccgtcgc cggcaaccgc aagcgcgcc tgattagtat tttcttgatg 900

atctgccagc aaatgacggg caccaacgcc atcaacacgt acgcgcctac catcttcaag 960
aacttgggta tcaccggcac gtcgactagc ttgttttagta ccggcatcta tggatattgtc 1020
aaggtcgtta gctgcgtcat tttcttctg ttttggccg actcgtggg tcgtagacgt 1080
tcgtgctgt ggacgtcgat tgcgcagggt cttgctatgt tttatattgg cttttatgtc 1140
cgcatctcgc cgccgattga tggccagccg gtgccgcctg cgggttatgt agcgttgggtg 1200

tgcattttt tgtttgccgc tttcttccaa ttggctggg gtcctgcctg ctggatctac 1260
gcctcgaaa tccccgccgc ccgctgcgc tcctcaacg tgcctacgc cggcgacg 1320
cagtggctgt tcaatttctg cgtggccgc gccgtgccta ctatgtgtt caggtcggc 1380

ccccacggtt acggcaccta cctcatctt ggagcttct gcctcagcat gttgtcttt 1440
gtctggttct tcgtgccga gacaaagggt atctcgcttg agcacatgga tgagctgttt 1500
ggcgttactg atgggcctgc cgctgagaag tcgtcggtgc atggtggaga tgatgtcggg 1560
tcggagatgg ggaaggggga tcagaagtcg aagcatgttg aggtttatgt ttaa 1614

<210> 46

<211> 1587

<212> DNA

<213> *Pichia stipitis*

<400> 46

atgtcttctg tattgactaa cgaatactt aaagactact accacaacc gactcctgtt 60
gaagtgggta ctatgattgc tatcttagag atcgccgcac tttttctc cttcatagct 120
ggaagagtag gtgacatctg tggcagaaga agaaccatta gatacgggtc tttcatTTTT 180

gtagtaggcg gtctgtaca agctacttcg gtcaatattg tcaatctctc actaggaaga 240
ttgattgccg gtattgccat tggctttttg acaaccatca tccatgcta ccagtctgaa 300
atcagcccc cagacgatag aggtttctat gcctgttttg agttcaccgg aaatatcatt 360
ggatatgcta gtagtatttg ggtagactac gggttttcat ttttagaca tgatttcagc 420
tggaggagcc cattgtatgt tcaggttgtt attggctcca tgttatttat tggttcattc 480
cttattgtag aaacccttag atggctcttg gatcacaacc atgatatcga aggcattgatt 540
gtcatttctg acttgatgc agatgggtgat gtggaagacg atgatgctat tgctgagtac 600
agaaacataa aggaaagtgt cttgatagcc agagtgaag gcggagagag atcgtaccag 660

tatttgttca ccaaatatac caagagactt tctgtggcat gcttttcga aatgtttgcc 720
cagatgaatg gtataacat ggtatcttac tatgtccta tgatcttga atctgtggc 780
tgggttggta gacaagctat cttgatgact ggtatcaact ccattatcta catcttagt 840
accattctc catggtactt agttgattct tggggcagaa aacctttgct tttatctgga 900
tctgtgctca tgggtgttcc gctcttaacc attgcttgtt cgttattctt aaacaacaca 960
tacacaccg gggttgttgt tggcagtgt atcgtattca atgtgtctt tggatacagt 1020
tgggttccaa ttccttggt catgagcgaa gtgttccta actcagttag atcaaaaggt 1080
gtgccatgt ctactgcaac caactggctc ttaacttta ttgttgaga gatgacacct 1140

atTTTgtTgg atacaattac ctggagaact tacttgatcc cggcaacttc gtgtgtatta 1200
 tcgtTTTTtg ctgttgatt tttatttcca gagaccaagg gtttagcatt ggaggatatg 1260
 ggctccgatat tcgatgataa ttctgtaata ttttcatatc actcaacttc ttccactggg 1320
 tatggTgcga ccgagtctaa cagtaatgcc aggagagcaa gtgtcatctc ttcagaaaac 1380
 taccaggata gtttgcatca gacagcggct tcattggcta gaaatccttc aagcatgagg 1440
 cctgattacg atggcataat cacaggagct gctacccttt cgccagtacc accattaaaa 1500
 ccaataaaca tticcagcaa tattccgcag gaaattgaac caccaacctt tgatgaaatc 1560
 tttaagtaca agttgaatga gatggaa 1587

<210> 47

<211> 1257

<212> DNA

<213> *Pichia stipitis*

<400> 47

atgactTTtg cagttaactt gtatgtgttt gcagttggta gagtgcTTtc tggggTgggt 60
 glaggagtTc tatcgactat ggtgccgtcc tatcaatgcg aaattagtcc cagcgaagaa 120
 agaggcaagt tggTgtTgg agagTtcacg gaaatatca ctggttatgc tctcagtTga 180
 tgggccgatt acttctgcta ctttatTcaa gatataggTg atgcaaggga gaagcctcat 240
 agcttctTTtg cccactTgtc ctggcgattg cctctattca tccaggTggT gatagcggct 300
 gttctctTTg ttgggggatt ttttattgtc gagtccctc gttggttatt agatgtagac 360
 caggaccaac aaggattcca tgtattagcg ttgctctatg attcacatct agatgataac 420

aaaccacgtg aagagtTctt tatgatcaaa aactccatct tgtagaaaag agaaactaca 480
 cctaagagcg aacgaactTg gaaacatatg ttcaagaact acatgacccg agtgcTTata 540
 gctTgtTcag cactTggctt tgcacagtTc aacggcataa atatcattTc gtactatgcc 600
 cccatggtat ttgaagaagc aggttTcaac aactccaagg ctttacttat gacaggcatc 660
 aactctatag tatattggTt cagtacgatt cctccgtggT ttctcgtTga tcattggggT 720
 agaaagccaa ttttgatatc cgggggtTta tctatgggaa tatgtattgg ttTgattgcg 780
 gtggtaattc tactagacaa gtcgtTcaca ccgtctatgg ttgcggtatt ggtgataatc 840
 tacaatgcat cttttggcta cagtTggggT cctatcggat tcttgatccc gccggaggTg 900

atgccattgg cagttagatc gaaaggTgtt tctatttcta cggctacaaa ctggttTgcc 960
 aattttgtTg tgggtcagat gacgccaatt ctacagcaga gattgggctg gggaacttat 1020
 ctattcccgg ctggtagtTg tatcatctcg gtgatagtgg tgattttctt ctatccagag 1080
 acaaagggtg cagagctaga ggatatggac tctgtgtTcg agagctTtta caactacaag 1140

tctccgttca agatttcacg aaagagacac cagaatgatg gccaggcgta ccaaagggtg 1200
gagaacgata tccgccacaa cgatgtagaa atggacgatt tggacgattt ggactaa 1257
<210> 48
<211> 1757
<212> DNA
<213> Neurospora crassa
<400> 48
atggaattcg gtggcggagg cggctccggc gcagctggtt tttatgatgc ggctcttcag 60

aggcgtgagg cagtgatggg gaagagtggc cctgcagcac ttgtcaagaa cttccgggtc 120
ttttctattg catgcttcgc atgtatcggg ggtgtgctct atgggtacaa ccaagggatg 180
ttttcgggtg tctcgcgat gccagccttt cagaaacaca tgggcgaata cgatccgata 240
gacgagaacg cgagtcagac aaagaagggc tggctaaccg caattcttga gctcgggtgt 300
tggcttggta cgcttctgtc tgggttcatg gcagaggttc tctcgagaaa gtacgggtgtg 360
ctagtggcgt gcttggtttt catgctgggt gtggtcatcc aagccacgtc tatctctgga 420
ggacatgaga ccattcttgc cggacgggtt atcacgggta tgggtgtcgg atccttagcc 480
atgatcattc ccatttaca ctcggaagt gcaccacctg aggtccgtgg agctcttgtt 540

gctctccagc agttggctat ctgcttcggt atcatggtca gcttctggat tgactacgga 600
accaactata tggcggcgc caagctcgag acccaatccg acgccgcctg gcttgtacct 660
gtctgcctgc aactcgcgcc tgcctcatt ctgttttctg gcatgatgtt catgcccttc 720
tccccacgtt ggctcatcca ccatggccgc gaggcggaag ctcgaaagat cctctccacc 780
cttcgcggtc taccccaaga ccacgagctt gtcgagctcg agttcctcga aataaaggct 840
cagtctctct tcgaaaaacg cagcattgcc gattgtttc ccgaattgcg cgagcagact 900
gcctggaata cctttaagct ccagtttgc gccatagaga agcttttccg gacaaaggca 960
atgttccgac gcgttgtcgt ggcaaccgta accatgttct tccagcagtg gtccggcatc 1020

aatgcgattc tctactacgc cccgaaatc ttcaagcagc ttggactgag cggtaacaca 1080
acctcactcc tggctacggg ttagtaggac atcgtcatgt tcatcgcaac ggctcctgcc 1140
gtgctgtgga tcgaccgtgt tggctgcaag cccgtgttta ctatcggtgc cctcgggatg 1200
gtacctgcc atatcatcat cgctgtcatt gttgccaaga acgtggacca atgggagact 1260
cataaggctg ctggatgggc tgctgtagcc atggtctggc tattcgtcat tcactttgga 1320
tattcatggg gtccatgtgc ctggatcatt gttgctgaga tctggccgtt gactacgagg 1380
ccatatggtg tctctctagg agcttcgagc aactggatga acaactttat cgtcggtcag 1440

gtcacgccgg atatgttaaa ggcgatcccg tacggaacgt atatcatctt cgggttggtg 1500

actatatggg tgccgccttt atttggttct ttgtgccgga aacgaagaga ttaaccttgg 1560

aagagatgga catgatcttc ggatccgaag gcactgcaca agccgacaat gagcgcatgg 1620

aggagatcaa tgctgagatt ggtcttacc gattcctgca aggtggtagt ggtgcaaacc 1680

aaggtgctgc tgatggaagc gatactggtt atgatgcgga gaagggaag agcgaacact 1740

attctcagca tgtctaa 1757

<210> 49

<211> 1584

<212> DNA

<213> Neurospora crassa

<400> 49

atgggcttgt cactcaagaa gcctgaaggt gtgccgggca agtcatggcc cgccattgtc 60

attggcttgt ttgtgcctt tgggtgtgta ctctttgggt atgacactgg cactattggc 120

ggatatccttg ctatgccta ttggcaagat ttgttttcga caggttacag aaaccagag 180

catcacttgg acgttacgc gtgcagctct gccactatcg tctccattct gtctgctgga 240

accttctttg gcgtcttgg tgccgtccc ctggccgact gggctggacg acgttgggc 300

cttattctgt cgtgcttgg gtttatcttc ggtgtcatcc tgcagaccgc agccgtcagc 360

attctctttt ttctggctgg ccgattcttt gctggattgg gagttggtct catatcggca 420

accatcccc tctatcaatc cgagactgcc ccgaaatgga ttcgtggtgt catcgtcggg 480

tcctatcagc tagccattac catcggtctt ctctttgcct ccattgtcaa caatgccacg 540

cataacatgc agaacaccgg ctgctatcgc atcccatag ctgtccaatt tgcattggcg 600

atcatcctga tcgttggcat gatcattctt ccgaaactc cacgtttca tatcaagaga 660

gacaatctcc cagccgccac taggtctcta gctatctcc gccgtctgga gcagaacct 720

ccagcgatca tcgaagagct ttccgagatc caagccaatc atgaattga gaagagctc 780

gggaaggcga cctacttga ctgcctcaag ggcaatttac tcaagcggt ccttactggc 840

tgttttctcc agagcctgca gcagttgact ggcatcaact ttatcttcta ctacggcaca 900

cagttcttca aaaactccgg attctcagac tcgtttctga tacccttgat cactaatctt 960

gtcaatgtcg tgcgacct tcccgactc tacgcatcg acaaatgggg ccggaggcct 1020

gttttactct ggggagctgt tgggatgtgt gtctgccagt tcacgttgc tattcttggg 1080

acaacaacga caagtcaaga tgcaagcgga atgatattg tgcataatct cgccgcacag 1140

aaagcagcta ttgcattcat ctgcttctac atctttttct tcgctgcac ttgggttcca 1200

gttgctggg tcgttacagg cgagatcttc ccccttaaag tccgcgcaa gtcgctctcc 1260
ataactacag cgtcgaattg gctgctcaac tgggccattg cttacagcac accttacctt 1320
gtcaactacg gccctggcaa tgccaacctg cagtccaaga tcttcttcgt ctggggcgga 1380
tgctgcttca tctgcatcgc attcgtttac ttcattgatct atgagacaaa aggtctcaca 1440
ctggagcagg ttgacgagct atatgaagag gtctcggatg ccaggaagag tattggttgg 1500
gtgccgacca tcactttccg ggagatccgg gaggaaaaga aagtaaggga tccagttgtt 1560

gatatcactg aagaggcagc ttga 1584

<210> 50

<211> 1968

<212> DNA

<213> Neurospora crassa

<400> 50

atgggccaca atccagacct ggacagtagc ggcaactgccg gagaacccaa aggtgtcacc 60
ggctcgaca ttgaacaaac ctctccaac ctggaagcca acatcaacct cgaagccaag 120
ctcaagaacc cgctcgacgg cctttccgc gtcgagctcc tgtcacgctg cgagaccttt 180
tgcgccgaaa agaacctaac cgagcacctc cctcttttcc gtaaggagc actcatcgcc 240
cagtcgccg acagctatgc gtccatctcg ggcccgaag ccttggacga tgaggagaag 300
gcagtacttt tgaaggaggt cgaacacaag tggcggctgc cggcaagact gtccctgacg 360

attgctactt gctcgatcgg tgctgctgtc caaggttggg atcagacggg cacgaatggc 420
gcgaatatct tctttccaa ctattacggt atcggaggcg aactgcgag ggagaagtgg 480
cttgtcggat tgatcaatgc tgggcctat attgggagcg cattcatcgg ttgctggctt 540
tctgatccca tcaacaactg gattggtcgt cgtggtgtta tcttgtctc tgctcacttc 600
tgtatctggc ccgtcatcgg ttctgctttc tgcacacat ggcccagca actggcctgc 660
cgtctgctga tgggtatcgg tatgggtgtg aaggcatcaa cggtgccgat ctatgccg 720
gaaaactcgc ctgcttctat tcgaggtgcg ctggtcatgt catggcagat gtggacagcc 780
ttcggcactt tcttgggcac tgcctttaac ctgcccgtct tccacgccag ctccaacgtt 840

aactggcgcc tcatgctcgg tgcccccttc attcccgcg taccctgct tctgctcacc 900
tatctttgcc ccgagtcgcc gcgctggtag atgaagaagg gccgtaccc agaagcctgg 960
aatccatgg tcaagctcgc caaccacccc atccaagtgg cccgcgacat gttctacatc 1020
cactcgcaat tggaagtcca gcaccagctc ctgccgggt ccaactatgc caagcgttc 1080
gtcgagctct tcaccgtccc tcgtgttcgc cgcgccaccc tcgccgttt caccgtcatg 1140
attgcccagc agatgtgcgg aatcaacatc atcgctttt acagcaccac catcttcaag 1200

gattccggct ccaccgaatt ccaagccctg ctttctctct tcggcttcgg tctagticaat 1260
 tggctctttg ctttccccgc cttctggact atcgacactt ttggccggcg ctctctgctt 1320

ctttttacct tcccgcaaat gatgtggacc ctgctagcag ccggcctctt caccttgctc 1380
 gacatgggtc ccgcccggac cgggctcgtc gccttattcg tcttctctt cggcgcgttc 1440
 tactcaccgc gtgaaggtec tgtcccttc acctactcgg ccgaagtctt cccctctct 1500
 cacagagaag taggcatggg cttcgccgtc gccacctgcc tcttctgggc atctgtttt 1560
 ggtattacct tccccctctt gcttgactct ctcggcaccg tcggcgcctt tggctctgtac 1620
 gcggtcttca acctagtggc gtttattgcc atcttcttgg tcgtgccgga gacgaagcag 1680
 aagacgctcg aggagtggga ttatgtcttt gctgtgaaga cgagcaagtt catgtcgtat 1740
 cagtgcacca aggcgctgcc gtgggtcatc aagaggtggg tgttttgga gaggaatgca 1800

aagctggagc cactgtatga gtttgatcgg atcaaggagg cggagaagga gaggagagca 1860
 gaggaggaga gaagggcaaa ggagacggga acgatacct ctactgtac aggagctgag 1920
 ttggatgaga agaagggact gactcatgtt aatgtccta attcttag 1968

<210> 51

<211> 1725

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 51

atggcagttg aggagaacaa tatgcctgtt gtttcacagc aacccaagc tggatgaagac 60
 gtgatctctt cactcagtaa agattcccat ttaagcgac aatctcaaaa gtattctaata 120
 gatgaattga aagccggatg gtcagggtct gaaggctccc aaagtgttcc tatagagata 180

ccaagaagc ccatgtctga atatgttacc gtttcttgc tttgtttgtg tgttgccttc 240
 ggcggcttca tgtttggctg ggataccggt actatttctg ggtttgttgt ccaaacagac 300
 tttttgagaa ggtttggat gaaacataag gatggtaccc actatttgc aaacgtcaga 360
 acaggtttaa tcgtgccat tttcaatatt ggctgtgcct ttggtggat tatactttcc 420
 aaaggtggag atatgtatgg ccgtaaaaag ggtctttcga ttgtcgtctc ggtttatata 480
 gttggtatta tcattcaat tgcctctatc aacaagtgg accaatattt cattggtaga 540
 atcatactg gtttgggtgt cggcggcatc gccgtcttat gtcctatgtt gatctctgaa 600
 attgctcaa agcacttgag aggcacacta gtttcttgtt atcagctgat gattactgca 660

ggtatctttt tgggctactg tactaattac ggtacaaaga gctattcgaa ctcagttcaa 720
 tggagagttc cattagggt atgtttcgct tggtcattat ttatgattgg cgctttgacg 780

ttagttcctg aatccccacg ttatttatgt gaggtgaata aggtagaaga cgccaagcgt 840
 tccattgcta agtctaaca ggtgtcacca gaggatcctg cgtccaggc agagtttagat 900
 ctgatcatgg ccggtataga agctgaaaaa ctggctggca atgcgtcctg gggggaatta 960
 tttccacca agaccaaagt atttcaacgt ttgttgatgg gtgtgtttgt tcaaagtgc 1020
 caacaattaa ccggtacaa ttattttttc tactacggta ccgttatatt caagtcagtt 1080
 ggcttgatg attcctttga aacatccatt gtcattgggt tagtcaactt tgcctccact 1140

ttcttttagtt tgggactgt cgaaaacttg ggacatcgta aatgtttact tttgggcgt 1200
 gccactatga tggcttgat ggtcatctac gcctctgttg gtgttactag attatattcct 1260
 caccgtaaaa gccagccatc ttctaaggt gccgtaact gtatgattgt ctttacctgt 1320
 ttttatattt tctgttatgc cacaacctgg gcgccagtgt cctgggtcat cacagcagaa 1380
 tcattccac tgagagtcaa gtcgaaatgt atggcgttgg cctctgttc caattgggta 1440
 tgggggttct tgattgcatt ttaccacca ttcacacat ctgccattaa ctctactac 1500
 ggttatgtct tcatgggctg ttgggttgc atgtttttt atgtctttt ctttgttcca 1560
 gaaactaaag gcctatcgtt agaagaaatt caagaattat gggaagaagg tgttttacct 1620

tggaaatctg aaggctggat tccttcatcc agaagaggta ataattacga tttagaggat 1680
 ttacaacatg acgacaaacc gtggtacaag gccatgctag aataa 1725

<210> 52

<211> 1908

<212> DNA

<213> *Pichia stipitis*

<400> 52

atgagtgtg acgaaaaagt cgctgtgcc ggccaggacg gcttgtttga acacaacagt 60
 tccacttga gcatcgagga caagaagccc tccaagagct ccgatgtcga ttccgtgaac 120
 tcgaattag tagacaactc ggtagagggc aacatcttgt ccagtagac cgaaagtcag 180
 gtgatgcaga tgggtagaag ctatgccacc aagcacggct tggaccaga attgttcgcc 240
 aaggcagctg ctgttgccag aactcctctt ggtttcaact ccatgccctt ctgacagag 300

gaagagaagg ttggtttgaa tgccgaagcc actaataagt ggcacattcc accagattg 360
 atcggggtta ttgccttggg ttctatggcc gctgtgtgc aggtatgga cgaatcggtc 420
 attaacggtg ccaacttgtt ctacccaag gctttcggag tcgacacat gcacaattcg 480
 gacttgattg aaggtttgat caatggtgct ccttacctt gctgtggtat tctttcctgt 540
 tggttgtctg acgtttgtaa ccgtcgtctt ggtagaaaat ggaccatttt ctggtgttgt 600
 gtcatttctg ccatcacctg tgtctggcaa ggtcttgtca acaactggta ccatttgttc 660

attgctcggt tcttccttgg atttgggtgt ggatcaagt ccgccactgt tctgcctac 720
tctgccgaat gtactcctaa acacatcaga ggttcgttag tcatgtttgt gcaattcttc 780

acagctgttg gtattatgtt tggttatgtt gcttccttgg ctttctacaa tgcggagat 840
agaggaatcc attacgggtt gaactggaga ttgatgcttg gttcggccgc tattcctgct 900
gtcatcatct tgttccaaat tcttttcgct cctgaatctc cacgttgggt aatgggtaag 960
gacagacacc ttgaagcctt tgagtccttg aagcaattga gatacgaaga acttgcctgct 1020
gctcgtgact gtttctacca gtacgtcttg ttagctgaag aaggttctta caagatccca 1080
acctcacca gatttaagga aatgttcacc aagagaagaa acagaaacgg tgccatcggt 1140
gcatttattg tcatgttcat gcaacagttc tgtggatca acgtcattgc ttactactct 1200
tcgtctatct ttgtccaatc tggtttctct caaacttctg ctttgatcgc ttcttgggggt 1260

ttcggtatgc ttaacttcac ctttgccatt cctgccttct tcacaatcga tcgtttcgggt 1320
agaagatcct tattgttgggt taccttcccc ttgatggcta ttttcttatt gattgccgggt 1380
ttcggtttct tgataaacga agaaacaaac tccaaggga gattgggaat gatcatcatc 1440
ggatctata tgttcacat ctgttactct tccggtgaag gtccagttcc ttacacctac 1500
tctgccgaag ccttccatt gtacatcaga gacttgggta tgtcttttgc tactgccacc 1560
tgttggactt tcaacttcat cttggccttc acctggaaca gatttgtcaa tgcattcaca 1620
tctactgggt ccttcggctt ctacgtgct tggaacatca ttggtttctt cttgggtctta 1680
tggttcttgc cagaaaccaa gggcttgacc ttggaagaat tggacgaagt cttcgccgtt 1740

tccgcgtec aacacgcaa gtacaaacc aagagtttga tcaacttcat ccaaagatac 1800
gttttacgtt ccaaggtggc tccattgcct ccattgtacg accaccagag attggctgtc 1860
accaaccag aatggaacga caagccagaa gtctcttatg tcgagtag 1908

<210> 53

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 53

atatatgagc tcgtgagtaa ggaaagagtg aggaactatc 40

<210> 54

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 54

atataacta gttgttttat atttgttgta aaaagtagat aattacttcc 50

<210> 55

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 55

atggatcaa aaatgtcgtc tcacggctcc 30

<210> 56

<211> 60

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 56

atgaattcct acaaattctt ttcaaaaac aatttttgtt cagcaacgat agcttcggac 60

<210> 57

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223>

Synthesized Construct

<400> 57

atactagtaa aaatgggcat cttcaacaag aagc 34

<210> 58

<211> 64

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 58

gcataatcgat ctacaaatct tcttcagaaa tcaatttttg ttcagcaaca gacttgcct 60
catg 64

<210> 59

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 59

tattaaatcg atggtagtgg tagtgtgagc aaggcgagg ag 42

<210> 60

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 60

tattaaatcg acctacttgt acagtcgtc catgcc 36

<210> 61

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 61

gcatggatcc atgtcgtctc acggtcc 28

<210> 62

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct
 <400> 62
 tataatgaat tcagcaacga tagcttcgga c 31

<210> 63
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 63
 tattaaacta gtatgggcat cttcaacaag aagc 34

<210> 64
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 64
 ttataagaat tcagcaacag acttgccctc atg 33

<210> 65
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 65
 gcatactagt aaaaatgtct cttcctaagg atttcctct 39

<210> 66
 <211> 55
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct

<400> 66

atactgcagt taatgatgat gatgatgatg gtccttcttg atcaaagagt caaag 55

<210> 67

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 67

Pro Glu Ser Pro Arg Phe

1 5

<210> 68

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 68

Pro Glu Ser Pro

1

<210> 69

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 69

Pro Glu Ser Pro

1

<210> 70

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> GS-linker

<400> 70

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10

<210> 71

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 71

gacatcgatg acatatgcgc taccgcaac

29

<210> 72

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 72

gtgcacgtcg gacccgcaga ttcc

24

<210> 73

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 73

ggaatctgcg ggtccgacgt gcac

24

<210> 74

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 74	
cagaagattt aaggatcctg aacgtaga	28
<210> 75	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthesized Construct	
<400> 75	
gacatcagtg acatatgtcg ccttcc	26
<210> 76	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthesized Construct	
<400> 76	
cctggattga ggatcctgaa cgtata	26
<210> 77	
<211> 28	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthesized Construct	
<400> 77	
gacatcgatg acatatggct tccgcaac	28
<210> 78	
<211> 29	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthesized Construct	
<400> 78	

ccagaagtat tgagaattct gaacgtaga	29
<210> 79	
<211> 27	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthesized Construct	
<400> 79	
gacatcgatg acatatggcg actctgc	27
<210> 80	
<211> 29	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthesized Construct	
<400> 80	
ggatacagaa tgaggatcct gaacgtaga	29
<210> 81	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthesized Construct	
<400> 81	
taatacgact cactatagg	20
<210> 82	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthesized Construct	
<400> 82	
gctagttatt gctcagcgg	19
<210> 83	

<211> 36
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 83
 cctatcgtca ttacctcacg tgacgagggg cggctg 36

<210> 84
 <211> 36
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 84
 cagccgcccc tcgtcacgtg aggtaatgac gatagg 36

<210> 85
 <211> 29
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 85
 ccttcgaaac ggctacaaac cccaagacg 29

<210> 86
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 86
 gcttgtcatc acatcacgtt cagagagccg tctg 34

<210> 87
 <211> 34

<212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 87
 cagacggctc tctgaacgtg atgtgatgac aagc 34
 <210> 88
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 88
 gcatttgaga cgtaacaga tccaagagc 30
 <210> 89
 <211> 36
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 89
 cctattgtca tcacttcacg tgacgagggc cgcttg 36
 <210> 90
 <211> 36
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 90
 caagcggccc tcgtcacgtg aagtgatgac aatagg 36
 <210> 91
 <211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthesized Construct
 <400> 91
 cctttgagac tgccacaaac cctaagaccg gtg 33
 <210> 92
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 92
 atggatccaa aaatgtcgtc tcacggctcc 30
 <210> 93
 <211> 60
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 93
 atgaattcct acaaatcttc ttcagaaatc aatttttgtt cagcaacgat agcttcggac 60
 <210> 94
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct
 <400> 94
 atactagtaa aaatgggcat cttcaacaag aagc 34
 <210> 95
 <211> 64
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized Construct

<400> 95

gcataatcgat ctacaaatct tcttcagaaa tcaatttttg ttcagcaaca gacttgcct 60

catg

64

<210> 96

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 96

gcataactagt aaaaatgtct cttcctaagg atttcctct 39

<210> 97

<211> 55

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct

<400> 97

atactgcagt taatgatgat gatgatgatg gtccttcttg atcaaagagt caaag 55