



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0070151  
(43) 공개일자 2015년06월24일

- (51) 국제특허분류(Int. C1.)  
*C12N 15/63* (2006.01) *C12N 9/88* (2006.01)  
*C12N 9/90* (2006.01) *C12N 9/96* (2006.01)  
*C12P 7/06* (2006.01)
- (52) CPC특허분류  
*C12N 15/63* (2013.01)  
*C12N 9/88* (2013.01)
- (21) 출원번호 10-2015-7009586
- (22) 출원일자(국제) 2013년10월15일  
 심사청구일자 없음
- (85) 번역문제출일자 2015년04월14일
- (86) 국제출원번호 PCT/EP2013/071462
- (87) 국제공개번호 WO 2014/060377  
 국제공개일자 2014년04월24일
- (30) 우선권주장  
 12188715.2 2012년10월16일  
 유럽특허청(EPO)(EP)  
 13166959.0 2013년05월08일  
 유럽특허청(EPO)(EP)
- (71) 출원인  
 디에스엠 아이피 어셋츠 비.브이.  
 네덜란드 앤엘-6411 티이 헤트렌 헤트 오버룬 1
- (72) 발명자  
 온케르스 마리아 베르네디나 엘리자베스  
 네덜란드 앤엘-6100 아아 에흐트 피오 박스 4  
 클라아센 폴  
 네덜란드 앤엘-6100 아아 에흐트 피오 박스 4  
 (뒷면에 계속)
- (74) 대리인  
 제일특허법인

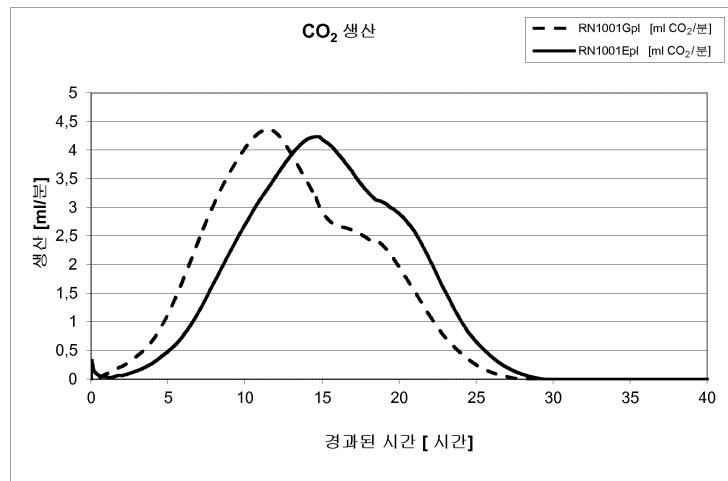
전체 청구항 수 : 총 17 항

(54) 발명의 명칭 5탄당 전환에 개선된 세포

**(57) 요 약**

본 발명은 서열 번호 20으로 제시된 아미노산 서열을 갖는 하나 이상의 이종성 또는 상동성 폴리펩티드, 또는 서열 번호 20과 45% 이상의 동일성을 갖는 이의 변형 폴리펩티드를 구성적으로 발현하는, 1종 이상의 5탄당 당분 및 1종 이상의 6탄당 당분을 발효 생성물로 전환시킬 수 있는 세포에 관한 것이다. 하나의 실시태양에서 이종성 폴리펩티드는 글리وك랄라제 활성을 갖는다.

**대 표 도** - 도10



(52) CPC특허분류  
*C12N 9/90* (2013.01)  
*C12N 9/96* (2013.01)  
*C12P 7/06* (2013.01)  
*C12Y 202/01001* (2013.01)  
*C12Y 202/01002* (2013.01)  
*C12Y 207/01017* (2013.01)  
*C12Y 404/01005* (2013.01)  
*C12Y 501/03004* (2013.01)  
*C12Y 503/01006* (2013.01)

(72) 발명자  
퇴니센 알로이시우스 윌헬무스 루돌푸스 후베르투스  
네덜란드 앤엘-6100 아아 에흐트 피오 박스 4

---

드 용 르네 마르셀  
네덜란드 앤엘-6100 아아 에흐트 피오 박스 4

## 명세서

### 청구범위

#### 청구항 1

서열 번호 20으로 제시된 아미노산 서열을 갖는 하나 이상의 이종성 또는 상동성 폴리펩티드, 또는 서열 번호 20과 45% 이상의 동일성을 갖는 이의 변형 폴리펩티드를 구성적으로(constitutionally) 발현하는, 1종 이상의 5탄당 당분 및 1종 이상의 6탄당 당분을 발효 생성물로 전환시킬 수 있는 세포.

#### 청구항 2

제1항에 있어서,

이종성 폴리펩티드가 글리옥살라아제(glyoxalase) 활성, 바람직하게는 글리옥살라아제 I 활성을 갖는 세포.

#### 청구항 3

- (a) 서열 번호 27의 뉴클레오티드 서열;
- (b) 서열 번호 27의 뉴클레오티드 서열과 약 50% 이상의 서열 동일성을 갖는 뉴클레오티드 서열;
- (c) 100개 이상의 뉴클레오티드를 갖는 (a), (b) 또는 (c)에 정의된 뉴클레오티드 서열의 단편;
- (d) (a), (b) 또는 (c)에 정의된 서열에 대한 유전자 코드의 결과로서 축퇴된 서열;
- (e) (a), (b), (c) 또는 (d)에 정의된 뉴클레오티드 서열의 역방향 보체인 뉴클레오티드 서열

을 포함하는, 구성적으로 발현된 이종성 또는 상동성 폴리뉴클레오티드를 포함하는, 1종 이상의 5탄당 당분 및/ 또는 1종 이상의 6탄당 당분을 발효 생성물로 전환시킬 수 있는 세포.

#### 청구항 4

제1항 내지 제3항중 어느 한 항에 있어서,

자일로스 이소머라아제(isomerase)를 코딩하는 뉴클레오티드 서열을 포함하는 세포.

#### 청구항 5

제1항 내지 제4항중 어느 한 항에 있어서,

- (a) 세포에서의 자일로스 수송의 증가;
- (b) 자일룰로스 키나아제(kinase) 활성의 증가;
- (c) 5탄당 포스페이트 경로를 통한 플럭스(flux)의 증가;
- (d) 알도스 환원효소 활성의 감소;
- (e) 이화산물 억제에 대한 감수성의 감소;
- (f) 에탄올, 오스몰농도(osmolarity) 또는 유기산에 대한 내성의 증가; 또는
- (g) 부산물 생산의 감소

를 초래하는 하나 이상의 유전자 변형을 포함하는 세포.

#### 청구항 6

제5항에 있어서,

하나 이상의 유전자 변형이 5탄당 포스페이트 경로의 비산화적 부분의 효소를 코딩하는 하나 이상의 유전자의 과발현을 초래하는 세포.

### 청구항 7

제6항에 있어서,

유전자가 리볼로스-5-포스페이트 이소머라아제, 리볼로스-5-포스페이트 에피머라아제(epimerase), 트랜스케톨라아제(transketolase) 또는 트랜스알돌라아제(transaldolase)를 코딩하는 유전자인 세포.

### 청구항 8

제5항 내지 제7항중 어느 한 항에 있어서,

하나 이상의 유전자 변형이 자일룰로스 키나아제를 코딩하는 유전자의 과발현을 초래하는 세포.

### 청구항 9

제5항 내지 제8항중 어느 한 항에 있어서,

하나 이상의 유전자 변형이 세포에서 알도스 환원효소 활성의 감소를 초래하는 세포.

### 청구항 10

제5항 내지 제9항중 어느 한 항에 있어서,

유전자 TAL1, TKL1, RPE1 및 RKI1이 과발현되는, L-아라비노스를 사용하는 능력을 갖는 세포.

### 청구항 11

제1항 내지 제10항중 어느 한 항에 있어서,

GRE3-유전자의 코딩 영역이, 코딩 영역을 유전자 TAL1, TKL1, RPE1 및 RKI1이 포함된 뉴클레오티드 서열과 대체 함으로써 불활성화되는 세포.

### 청구항 12

제1항 내지 제11항중 어느 한 항에 있어서,

유전자 araA, araB 및 araD가 발현되는 세포.

### 청구항 13

제1항 내지 제11항중 어느 한 항에 있어서,

하나 이상의 구성적으로 발현되거나 구성적으로 과발현된 유전자가 세포의 계놈내로 안정하게 통합되는 세포.

### 청구항 14

제1항 내지 제13항중 어느 한 항에 따른 세포에 의해 자일로스 및/또는 아라비노스의 공급원이 함유된 배지를 발효시켜서 세포가 자일로스 및/또는 아라비노스를 발효 생성물로 발효시키는 단계를 포함하는, 발효 생성물의 제조 방법.

### 청구항 15

제14항에 있어서,

발효 생성물이 에탄올, 부탄올, 락트산, 3-하이드록시-프로피온산, 아크릴산, 아세트산, 석신산, 시트르산, 말산, 푸마르산, 이타콘산, 아미노산, 1,3-프로판-디올, 에틸렌, 글리세롤, 부탄올,  $\beta$ -락탐 항생물질 및 세팔로스포린(cephalosporin)을 포함하는 군으로부터 선택되는 제조 방법.

### 청구항 16

제15항에 있어서,

발효 생성물이 에탄올인 제조 방법.

## 청구항 17

제14항 내지 제16항 중 어느 한 항에 있어서,  
협기성인 제조 방법.

### 발명의 설명

#### 기술 분야

[0001] 본 발명은 5탄당 전환이 개선된 세포에 관한 것이다. 더욱 구체적으로 본 발명은 자일로스 및/또는 L-아라비노스의 전환이 개선된 세포에 관한 것이다. 세포는 발효 생성물의 생산, 예를 들어 리그노셀룰로스성 물질로부터 유래된 당분으로부터의 에탄올의 생산에 유용하다.

#### 배경기술

[0002] 최근 수십년간 전통적인 화석 연료(석유계 연료)의 대규모 소비는 높은 수준의 오염을 일으켰다. 이는, 화석 연료의 세계적 비축물이 제한되지 않는다는 현실 및 증가되는 환경적 의식과 함께, 리터 기준당 무연 가솔린에 비해  $\text{CO}_2$ 를 더 적게 배출하는 미립자-부채 연소 연료 공급원인 대체 연료, 예컨대 에탄올의 실행가능성을 조사하고자 하는 새로운 계획을 자극하였다.

[0003] 생물자원(biomass)-유래된 에탄올이 많은 상이한 공급원들로부터 수득되는 6탄당의 발효에 의해 생산될 수 있지만, 연료 알코올의 상업적 규모의 생산에 전형적으로 사용되는 기질, 예컨대 사탕수수 당분 및 옥수수 전분은 고가이다. 따라서 연료 에탄올의 생산을 증가시키기 위해서는 더 저렴한 공급원료의 사용이 필요할 것이다.

[0004] 현재, 식물 생물자원으로부터 유래된 리그노셀룰로스성(lignocellulosic) 공급원료만이 에탄올 생산을 위해 현재 사용되는 작물을 대체하기에 충분한 양으로 이용가능하다. 대부분의 리그노셀룰로스성 물질에서, 글루코스 다음으로 두 번째로 가장 흔한 당분은 자일로스이다. 또한 L-아라비노스는 몇몇 리그노셀룰로스성 물질로부터 유래된 당분이다. 또한 이와 같이, 경제적으로 실행가능한 연료 생산 공정을 위해서는, 에탄올을 형성하기 위해 6탄당 및 5탄당이 발효되어야 한다. 효모 사카로마이세스 세레비지아(*Saccharomyces cerevisiae*)는 강하고, 에탄올 생산을 위해 잘 적응되지만, 탄소원으로서 자일로스를 사용하여 에탄올을 생산하는 것은 불가능하다. 따라서, 리그노셀룰로스성 공급원료로부터 에탄올의 상업적으로 실현될 수 있는 생산을 가능하게 하기 위해 이들 특성을 갖는 유기체가 요구된다. 협기성 균류(fungus) 피로마이세스(*Piromyces*) Sp.E2로부터의 자일로스 이소머라이제(isomerase)가 사카로마이세스 세레비지아에 도입되었고, 높은 수준의 효소 활성이 관찰되었는데, 이는 이러한 균주가 협기성으로 성장하고 자일로스로부터 에탄올을 생산할 수 있도록 하였다(국제특허공개 제W003/062340호 및 제W006/009434호). 이러한 효모 균주는 상업적 규모에서의 에탄올 생산과 상용가능한 자일로스 소비 및 에탄올 형성의 특이적 속도를 최초로 제공하였다.

[0005] 그러나, 5탄당 전환을 개선시키고 발효 시간을 감소시키려는 요구는 여전히 존재한다.

#### 발명의 내용

[0006] 본 발명의 목적은 5탄당 전환이 개선된 세포를 제공하는 것이다. 본 발명의 또 다른 목적은 자일로스 및/또는 L-아라비노스의 전환이 개선된 재조합 균주를 제공하는 것이다. 또 다른 목적은 글루코스의 존재하에 자일로스 및/또는 L-아라비노스의 전환이 개선된 균주를 제공하는 것이다. 또 다른 목적은 5탄당 및 6탄당을 포함하는 당분 혼합물의 발효에 있어서 발효 시간을 감소시키는 것이다.

[0007] 이들 목적중 하나 이상은 본 발명에 따라 달성된다. 본 발명에 따라서, 서열 번호 20의 아미노산 서열, 또는 서열 번호 27의 뉴클레오티드 서열에 의해 코딩되는 아미노산 서열을 갖는 이종성 또는 상동성 폴리펩티드, 또는 서열 번호 20과 45% 이상의 서열 동일성을 갖는 아미노산 서열을 갖는 이의 변형체를 구성적으로(consitutively) 발현하는 세포가 제공된다.

[0008] 실시예에 따른 세포는 5탄당 전환이 개선되고 5탄당 및 6탄당을 포함하는 당분 혼합물의 발효에 있어서 발효 시간을 감소시킨다. 하나의 실시태양에서 폴리펩티드는 글리وك살라아제(glyoxalase) 활성을 갖는다.

[0009] 본 발명은 또한

[0010] - 자일로스 및/또는 L-아라비노스의 공급원을 함유하는 배지를 본 발명의 세포에 의해 발효시켜 세포가 자일로스 및/또는 L-아라비노스를 발효 생성물로 발효시키도록 하는 단계를 포함하는 발효 생성물의 제조 방법;

[0011] - 발효 생성물의 제조 방법에 있어서의 본 발명의 세포의 용도를 제공한다.

### 도면의 간단한 설명

[0012] 도 1: PGK1 프로모터 및 PGI1 종결자를 갖는 GL01 발현 카세트(cassette)의 그래프.

도 2: 플라스미드 pRN228(실시예 참조)의 지도.

도 3: 플라스미드 pRN935(실시예 참조)의 지도.

도 4: 플라스미드 pRN685(실시예 참조)의 지도.

도 5: 플라스미드 pRN1048(실시예 참조)의 지도.

도 6: 플라스미드 pRN1129(실시예 참조)의 지도.

도 7: 효모 셔틀 벡터(shuttle vector) pRN599(실시예 참조)의 지도.

도 8: 플라스미드 pRN1049의 지도.

도 9: 유전자 발현 데이터. RN1001Ep1 및 RN1001Gp1에 대한 정규화된 발현 배수.

도 10: 균주 RN1001Gp1(상부) 및 RN1001Ep1(하부)의 성장에 대한 CO<sub>2</sub> 생산 곡선.

도 11: 당분 소비를 보여주는 균주 RN1001Ep1 및 RN1001Gp1의 성장 곡선.

도 12: 플라스미드 pRN324의 지도.

도 13: 플라스미드 pRN1142의 지도.

도 14: 사카로마이세스 세레비지아 계놈에서의 통합 자리를 보여주는 지도: 통합 자리 1(Chr XV, 458319 bp에서 459320 bp로 조정함). 특정화된 지도는 2개의 수직 점선 사이의 영역에 의해 지시된다.

도 15: PCR 단편(His 또는 His-Glo)의 사카로마이세스 세레비지아 계놈내로의 통합에 대한 기작을 보여주는 개략도.

도 16: 유전자 발현 데이터. 균주 RN1041(상부) 및 RN1216(하부)의 GL01에 대한 정규화된 발현 배수.

도 17: 균주 RN1041H, RN1041HG-1 및 RN1041HG-2의 성장에 대한 CO<sub>2</sub> 생산 곡선.

도 18: 균주 RN1041H(상부), RN 1041HG-1(중간부) 및 RN1041HG-2(하부)의 당분 소비 곡선.

도 19: 균주 RN1216H, RN1216HG-1 및 RN1216HG-2의 성장에 대한 CO<sub>2</sub> 생산 곡선.

도 20: 당분 소비를 보여주는 균주 RN1216H(상부), RN1216HG-1(중간부) 및 RN1216HG-2(하부)의 당분 소비 곡선.

도 21: 플라스미드 pDB1175(실시예 참조)의 지도.

도 22: 플라스미드 pDB1176(실시예 참조)의 지도.

도 23: 플라스미드 pDB1177(실시예 참조)의 지도.

도 24: 플라스미드 pDB1178(실시예 참조)의 지도.

도 25: 플라스미드 pRN1179(실시예 참조)의 지도.

도 26: PCR 단편의 사카로마이세스 세레비지아 계놈으로의 통합에 대한 기작을 보여주는 개략도.

도 27: 균주 RN1216 ScG\_H, RN1216 CglaG\_H, RN1216 ZrouG\_H, RN1216 KIG\_H, RN1216 CmagG\_H 및 RN1216H의 성

장에 대한  $\text{CO}_2$  생산 곡선.

- 도 28: 균주 RN1216 ScG\_H의 당분 소비 곡선.
- 도 29: 균주 RN1216 Cg1aG\_H의 당분 소비 곡선.
- 도 30: 균주 RN1216 ZrouG\_H의 당분 소비 곡선.
- 도 31: 균주 RN1216 KIG\_H의 당분 소비 곡선.
- 도 32: 균주 RN1216 CmagG\_H의 당분 소비 곡선.
- 도 33: 균주 R1216H의 당분 소비 곡선.

#### 서열 목록에 대한 간단한 설명

- 서열 번호 1: 순방향 프라이머 PGK1 프로모터;
- 서열 번호 2: 역방향 프라이머 PGI1 종결자;
- 서열 번호 3: 순방향 프라이머 ACT1 유전자;
- 서열 번호 4: 역방향 프라이머 ACT1 유전자;
- 서열 번호 5: 순방향 프라이머 GL01 유전자, Q-PCR;
- 서열 번호 6: 역방향 프라이머 GL01 유전자, Q-PCR;
- 서열 번호 7: 순방향 프라이머 ALG9 유전자, Q-PCR;
- 서열 번호 8: 역방향 프라이머 ALG9 유전자, Q-PCR;
- 서열 번호 9: 순방향 프라이머 UBC6 유전자, Q-PCR;
- 서열 번호 10: 역방향 프라이머 UBC6 유전자, Q-PCR;
- 서열 번호 11: 순방향 프라이머 HIS3 카세트, 5' 플랭크(flank) INT1;
- 서열 번호 12: 역방향 프라이머 GL01 카세트, 3' 플랭크 INT1;
- 서열 번호 13: 역방향 프라이머 HIS3 카세트, 3' 플랭크 INT1;
- 서열 번호 14: 순방향 프라이머 5' 플랭크 INT1;
- 서열 번호 15: 역방향 프라이머 5' 플랭크 INT1;
- 서열 번호 16: 순방향 프라이머 3' 플랭크 INT1;
- 서열 번호 17: 역방향 프라이머 3' 플랭크 INT1;
- 서열 번호 18: 통합 자리 INT1의 뉴클레오티드 서열 5' 플랭크;
- 서열 번호 19: 통합 자리 INT1의 뉴클레오티드 서열 3' 플랭크;
- 서열 번호 20: 사카로마이세스 세레비지아로부터의 GL01의 단백질 서열;
- 서열 번호 21: 칸디다 글라브라타(*Candida glabrata*)로부터의 GL01의 단백질 서열;
- 서열 번호 22: 자이고사카로마이세스 록시(*Zygosaccharomyces rouxii*)로부터의 GL01의 단백질 서열;
- 서열 번호 23: 클루이베로마이세스 락티스(*Kluyveromyces lactis*)로부터의 GL01의 단백질 서열;
- 서열 번호 24: 칸디다 마그놀리아(*Candida magnolia*)로부터의 GL01의 단백질 서열;
- 서열 번호 25: GL01 ORF의 순방향 프라이머;
- 서열 번호 26: GL01 ORF의 역방향 프라이머;
- 서열 번호 27: 사카로마이세스 세레비지아로부터의 GL01의 뉴클레오티드 서열;

서열 번호 28: 상동성 GL01 발현 카세트(프로모터 및 종결자 포함)의 순방향 프라이머;

서열 번호 29: 상동성 GL01 발현 카세트(프로모터 및 종결자 포함)의 역방향 프라이머;

서열 번호 30: 상동성 GL01 발현 카세트의 3'-말단의 50개 뉴클레오티드와 동일한, 5'-말단 상의 20개의 뉴클레오티드 및 50개 뉴클레오티드의 테일로 구성된, HIS3 카세트의 순방향 프라이머;

서열 번호 31: 3' 500 bp INT1 플랭크의 5'-말단의 50개 뉴클레오티드와 동일한, 5'-말단 상의 21개의 뉴클레오티드 및 50개 뉴클레오티드의 테일로 구성된, HIS3 카세트의 역방향 프라이머;

서열 번호 32: 상동성 GL01 발현 카세트의 5'-말단의 50개 뉴클레오티드와 동일한, 5'-말단 상의 23개의 뉴클레오티드 및 50개 뉴클레오티드의 테일로 구성된, 5' 500 bp INT1 플랭크의 역방향 프라이머;

서열 번호 33: HIS3 발현 카세트의 3'-말단의 50개 뉴클레오티드와 동일한, 5'-말단 상의 24개의 뉴클레오티드 및 50개 뉴클레오티드의 테일로 구성된, 3' 500 bp INT1 플랭크의 순방향 프라이머;

서열 번호 34: 사카로마이세스 세레비지아로부터의 GL01의 코돈 쌍 최적화된 뉴클레오티드 서열;

서열 번호 35: 칸디다 글라브라타로부터의 GL01의 코돈 쌍 최적화된 뉴클레오티드 서열;

서열 번호 36: 칸디다 마그놀리아로부터의 GL01의 코돈 쌍 최적화된 뉴클레오티드 서열;

서열 번호 37: 클루이베로마이세스 락티스로부터의 GL01의 코돈 쌍 최적화된 뉴클레오티드 서열;

서열 번호 38: 자이고사카로마이세스 룩시로부터의 GL01의 코돈 쌍 최적화된 뉴클레오티드 서열.

### 발명을 실시하기 위한 구체적인 내용

[0013] 본 명세서 및 첨부된 특허청구범위 전체를 통해, 단어 "포함하다" 및 "포함되다" 및 변형어, 예컨대 "포함하는", "포함되는"은 포괄적으로 해석되어야 한다. 즉, 이들 단어는, 내용상 허용되는 경우, 구체적으로 언급되지 않은 다른 요소 또는 정수의 가능한 포함을 전달하고자 한다.

[0014] 본 발명에 따라서, 세포는 글리옥살라아제 활성, 바람직하게는 글리옥살라아제 I 활성을 포함할 수 있다. 글리옥살라아제1(GL01)은 메틸글리옥살 이화작용에 관여하는 글리옥살라아제 I(EC 4.4.1.5)을 코딩하는 유전자이다 (2). 메틸글리옥살은 당분해작용의 부산물로서 형성되는 독성 화합물이다.

[0015] 글리옥살라아제의 다른 명칭은 예를 들어 락토일글루타티온 리아제(lyase), 알도케토뮤타아제(aldoketomutase), 케톤-알데히드 뮤타아제, 메틸글리옥살라아제, S-D-락토일글루타티온 메틸글리옥살 리아제이다.

[0016] 메틸글리옥살 이화작용의 한 방법은 글리옥살라아제 시스템을 포함하고, 여기서 메틸글리옥살은 글루타티온과 Glo1p에 의해 축합되어 S-D-락토일글루타티온을 생성한다. 이러한 글루타티온 티올에스테르는 이어서 글리옥살라아제 II(Glo2p 및 Glo4p)에 의해 락트산 및 글루타티온으로 가수분해된다. GL01 발현은 메틸글리옥살에 의해 유도되고, 높은 오스몰농도(osmolarity)의 글리세롤(Hog7p)-미토겐-활성화된 단백질(MAP) 키나아제(kinase)-의 존적 방식으로 삼투 스트레스(osmotic stress)에 의해 특이적으로 유도된다(1).

[0017] GL01의 감소는 메틸글리옥살에 대해 과민감성을 초래한다. 사카로마이세스 세레비지아에서, 글리옥살라아제 I(Glo1p)은 고유적이고 단량체로 발생된다. 이러한 시스템은 내생성 독소를 쳐분하는 효소의 전형적인 많은 특징을 보여준다. 우선적으로, 생체이물 대사작용에 관여하는 많은 효소의 엄청난 기질 범위와 대조적으로, 이는 좁은 기질 특이성을 보여준다. 두번째로, 세포내 티올이 그의 효소적 기작의 일부로서 요구되고, 세번째로, 시스템은 반응성 대사물질을 세포 대사작용에 유용할 수 있는 형태로 다시 재순환시키는 작용을 한다.

[0018] 하나의 실시태양에서, 본 발명은 글리옥살라아제를 발현하는 세포에 관한 것으로, 여기서 글리옥살라아제의 아미노산 서열은 서열 번호 20으로 제시된 아미노산 서열과 45% 이상의 동일성을 갖고, 뉴클레오티드 서열은 세포에 상동성 또는 이종성으로 구성적으로 통합된다.

### 아미노산/폴리뉴클레오티드 서열

[0020] 본 발명의 세포는 서열 번호 20의 아미노산 서열 또는 이와 45% 이상의 서열 동일성을 갖는 서열을 갖는 단백질에 관하여 규정된다. 하나의 실시태양에서, 단백질은 서열 번호 20의 아미노산 서열과 약 50% 이상, 약 55% 이상, 약 60% 이상, 약 65% 이상, 약 70% 이상, 약 80% 이상, 약 90% 이상, 약 94% 이상, 약 95% 이상, 약 96% 이상, 약 97% 이상, 약 98% 이상 또는 약 99% 이상의 서열 동일성을 갖는다.

- [0021] 하나의 실시태양에서, 세포는 서열 번호 20에서 위치 H25, E89, H269 및 E318 및/또는 H185, E242, H117 및 E163에 상응하는 아미노산 H, E, H 및 E를 갖는 폴리펩티드를 포함하고/하거나 발현한다. 하나의 실시태양에서, 세포는 서열 번호 20에서 단편 E242-L243-X-H245-N246에 상응하는 단편 (E,s,d)-L-X-(H,Y)-(N,s)를 갖는 폴리펩티드를 포함하고/하거나 발현한다. 하나의 실시태양에서 세포는 서열 번호 20에서 단편 G266-Y267-G268-H269에 상응하는 단편 G-(F,Y)-G-H를 갖는 폴리펩티드를 포함하고/하거나 발현한다. 하나의 실시태양에서 세포는 서열 번호 20에서 단편 G301-X(6)-F308-X(2)-D311-X(3)-Y315에 상응하는 단편 G-X(6)-(F,i)-X(2,3)-D-X(3)-Y를 갖는 폴리펩티드를 포함하고/하거나 발현한다.
- [0022] 상기에서, 아미노산은 1 문자 코드로 지시된다. X는 임의의 아미노산이고; (X,Y)는 아미노산 X 또는 Y이고; X(y)는 y개의 아미노산 X이고 X(y,z)에서 y 또는 z는 아미노산 X의 수이다. 소문자 코드는 소량으로 출현하는 아미노산을 지시한다.
- [0023] 하나의 실시태양에서 세포는 글리옥살라아제 활성을 갖는 폴리펩티드를 포함하고/하거나 발현한다.
- [0024] 본 발명에 따른 세포는 서열 번호 27의 뉴클레오티드 서열, 또는 서열 번호 27에 제시된 핵산 서열과 약 50% 이상, 약 55% 이상, 약 60% 이상, 약 65% 이상, 약 70% 이상, 약 75% 이상, 바람직하게는 약 80% 이상, 약 85% 이상, 약 90% 이상, 약 92% 이상, 약 93% 이상, 약 94% 이상, 약 95% 이상, 약 96% 이상, 약 97% 이상, 약 98% 이상 또는 약 99% 이상 또는 75% 이상, 80% 이상, 85% 이상, 90% 이상, 92% 이상, 93% 이상, 94% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상 또는 99% 이상의 서열 동일성을 갖는 서열을 갖는 글리옥살라아제를 코딩하는 뉴클레오티드 서열을 포함한다.
- [0025] 따라서, 본 발명은 글리옥살라아제 폴리펩티드를 코딩하는 유전자를 포함하는 폴리뉴클레오티드 서열을 갖는 세포, 뿐만 아니라 이의 코딩 서열을 제공한다.
- [0026] 본 발명의 폴리뉴클레오티드는 단리되거나 합성될 수 있다. 글리옥살라아제 폴리펩티드 및 글리옥살라아제 폴리뉴클레오티드는 본원에서 합성 폴리펩티드, 개별적으로 폴리뉴클레오티드일 수 있다. 합성 폴리뉴클레오티드는, 바람직하게는 본원에 참고로 인용된 국제특허공개 제WO2006/077258호 및/또는 제PCT/EP2007/055943호에 기재된 방법에 따라서, 코돈 사용에서 최적화될 수 있다. 제PCT/EP2007/055943호는 코돈-쌍 최적화를 개시한다.
- [0027] 이 용어는 단일 가닥 또는 이중 가닥의 리보핵산(RNA) 또는 데옥시리보핵산(DNA) 분자인 폴리뉴클레오티드 분자를 지칭한다. 폴리뉴클레오티드는 단리된 형태로 존재하거나, 재조합 핵산 분자 또는 벡터내에 포함되거나, 숙주 세포에 포함될 수 있다.
- [0028] 단어 "폴리펩티드"는 7개 이상의 아미노산 잔기를 함유하는 쇄에 대해 본원에서 사용된다. 모든 올리고펩티드 및 폴리펩티드 화학식 또는 서열은 본원에서 왼쪽에서 오른쪽으로, 아미노 말단에서 카복시 말단 방향으로 기입된다. 본원에 사용된 아미노산의 1 문자 코드는 당분야에 통상적으로 공지되어 있다.
- [0029] "단리된" 폴리펩티드 또는 단백질은 그의 고유의 환경으로부터 제거된 폴리펩티드 또는 단백질을 의미한다. 예를 들면, 숙주 세포에서 발현되는 재조합적으로 생산된 폴리펩티드 및 단백질은, 임의의 적합한 기법, 예컨대, 예를 들면, 스미쓰(Smith) 및 존슨(Johnson)의 문헌[Gene 67:31-40 (1988)]에 개시된 단일-단계 정제 방법에 의해 실질적으로 정제된 재조합 또는 고유의 폴리펩티드이므로, 본 발명의 목적을 위해 단리된 것으로 고려된다.
- [0030] 본 발명의 폴리뉴클레오티드, 예컨대 글리옥살라아제 폴리펩티드를 코딩하는 폴리뉴클레오티드는 본원에 제공된 표준 분자 생물학 기법 및 서열 정보를 사용하여 단리되거나 합성될 수 있다.
- [0031] 본 발명의 글리옥살라아제 폴리펩티드를 코딩하는 폴리뉴클레오티드는 cDNA, mRNA 또는 다른게는, 계놈성 DNA를, 주형 및 적절한 올리고뉴클레오티드 프라이머로서 사용하여 표준 PCR 증폭 기법에 따라 증폭될 수 있다. 이렇게 증폭된 핵산은 적절한 벡터내로 클로닝되고 DNA 서열 분석에 의해 특징화될 수 있다.
- [0032] **효소 반응속도론(kinetic)**
- [0033] 효소는, 모든 촉매와 마찬가지로, 공정에서 고갈되지 않고 화학 반응의 속도를 증진시키는 단백질 촉매이다.
- [0034] 이들은 기질에 일시적으로 결합함으로써 이들 효과를 달성하고, 그렇게 하는 동안, 기질을 생성물로 전환시키는데 필요한 활성화 에너지를 저하시킨다.
- [0035] 효소가 작용하는 속도는 수 개의 인자들, 예를 들어 기질 분자의 농도(지정된 [S], 몰농도의 단위로 표시됨), 온도, 저해제의 존재(기질과 동일한 자리에 결합하는 경쟁적 저해제 또는 그의 촉매력을 감소시키는 저해제) 등에 의해 결정된다.

또 다른 자리에 결합하는 비경쟁적 저해제), pH 등에 의해 영향받는다.

[0036] 효소 반응속도론 연구는 효소가 작용하는 속도를 연구하고 설명한다. 많은 측정 방법들이 존재한다. 효소는, 반응 시작 이후 짧은 기간 동안 대략 선형인 초기 속도에서 기질을 생성물로 전환시킨다. 반응이 진행되고 기질이 소비될 때, 기질은 여전히 포화 수준으로 존재하므로, 속도는 연속적으로 둔화된다. 초기(및 최대) 속도를 측정하기 위해, 반응이 완전한 완료에 대해 단지 소수의 퍼센트까지 진행되는 동안 효소 분석이 전형적으로 실행된다. 초기 속도 기간의 길이는 무엇보다도 분석 조건에 의존한다.

[0037] 효소 반응속도론에 대한 연구는 2가지 기본적인 이유로 인해 중요하다. 먼저, 이는 효소가 어떻게 작용하는지를 설명하는 것을 돋고, 두번째로, 이는 효소가 살아있는 유기체에서 어떻게 행동할지를 예측하는 것을 돋는다. 반응속도 상수,  $K_m$  및  $V_{\text{최대값}}$ 은 효소가 함께 대사작용을 제어하기 위해 어떻게 작용하는지를 이해하기 위해 중요하다.

[0038] 미카엘리스(Michaelis) 상수  $K_m$ 은 효소 반응의 속도가  $V_{\text{최대값}}$ 의 절반인 농도로서 실험적으로 규정된다.  $V_{\text{최대값}}$ 은 효소가 반응을 촉매화하는 최대 속도이다:  $[S]$ 가 더 높아질 수록, 효소는 기질로 포화되고 속도는  $V_{\text{최대값}}$ , 즉 효소의 최대 속도에 도달한다.

[0039]  $K_m$ 은 효소 및 그의 기질 사이의 결합 친화도 또는 강도의 (대략적인) 역 측정값이다.  $K_m$  값(mM)이 낮을 수록, 친화도는 더 크고, 이에 따라 소정의 속도를 달성하기 위해 필요한 기질의 농도가 더 낮다.

#### 숙주 세포

[0040] 숙주 세포는 유용한 생성물의 생산에 적합한 임의의 숙주 세포일 수 있다. 숙주 세포는 임의의 적합한 세포, 예컨대 원핵 세포, 예컨대 세균, 또는 진핵 세포일 수 있다. 전형적으로, 세포는 진핵 세포, 예를 들면 효모 또는 사상균류일 수 있다. 숙주 세포는 하나 이상의 5탄당 당분, 예컨대 L-아라비노스 또는 자일로스를 발효 생성물로 전환시킬 수 있다.

[0042] 숙주 세포는 세균일 수 있다. 하나의 실시태양에서 세균은 5탄당 전환을 위해 유전자 조작된다. 하나의 실시태양에서 세균은 에스케리키아 콜라이(*Escherichia coli*)가 아니다. 본 발명의 적용을 위해 적합한 숙주인 세균의 일례는 유전자 조작된 자이모모나스 모빌리스(*Zymomonas mobilis*)이다. 이러한 균주는 야나(Yanna) 등의 문헌[*Appl Microbiol Biotechnol*. 2012 Jun; 94(6):1667-78]에 기재되어 있다.

[0043] 본 발명을 위한 숙주는 (유전자 조작된) 효모일 수 있다. 유전자 조작은 이후 상세히 설명된다. 효모는 본원에서 진핵 미생물로서 규정되고, 주로 단세포 형태로 성장하는 하위부류 유마이코티나(*Eumycotina*)의 모든 종을 포함한다[알렉소풀로스(Alexopoulos, C. J.)의 문헌 "1962, In : Introductory Mycology, John Wiley & Sons, Inc., New York"].

[0044] 효모는 단세포 염상체의 출아에 의해 성장할 수 있거나, 유기체의 핵분열에 의해 성장할 수 있다. 형질전환된 숙주 세포로서 바람직한 효모는 사카로마이세스, 클루이베로마이세스(*Kluyveromyces*), 칸디다(*Candida*), 피키아(*Pichia*), 스키조사카로마이세스(*Schizosaccharomyces*), 한세놀라(*Hansenula*), 클로엑케라(*Kloeckera*), 슈와니오마이세스(*Schwanniomyces*), 및 야로위아(*Yarrowia*) 속에 속한다. 바람직하게는 효모는 협기성 발효, 더 바람직하게는 협기성 알코올성 발효를 가능하게 한다.

[0045] 한 실시태양에서, 숙주 세포는 효모이다.

[0046] 바람직하게는 숙주는 산업용 숙주, 더 바람직하게는 산업용 효모이다. 산업용 숙주 및 산업용 효모 세포는 다음과 같이 규정될 수 있다. 산업 공정에서 효모 세포가 살아있는 환경은 실험실의 환경과 상당히 상이하다. 산업용 효모 세포는 공정 동안 달라질 수 있는 복합적 환경 조건하에 잘 수행할 수 있어야 한다. 이러한 변수로는 영양 공급원, pH, 에탄올 농도, 온도, 산소 농도 등에서의 변화가 포함되고, 이들은 함께 사카로마이세스 세레비지아의 세포 성장 및 에탄올 생산에 잠재적인 영향을 갖는다. 불리한 산업 조건하에, 내환경성 균주는 강건한 성장 및 생산을 허용해야 한다. 산업용 효모 균주는 일반적으로 이들이 사용되는 적용분야, 예컨대 제빵 산업, 양조 산업, 와인 제조 및 에탄올 산업에서 초래될 수 있는 환경 조건에서의 이들 변화에 대해 더 강건하다. 산업용 효모의 예(사카로마이세스 세레비지아)는 에탄올 레드(Ethanol Red: 등록상표)[페르멘티스(Fermentis)], 페르미올(Fermiol: 등록상표)(DSM) 및 씨모사크(Thermosacc: 등록상표)[랄레만드(Lallemand)]이다.

[0047] 하나의 실시태양에서 숙주는 저해제 내성이다. 저해제 내성 숙주 세포는, 예컨대 카다(Kadar) 등의 문헌[*Appl*.

Biochem. Biotechnol. (2007), Vol. 136-140, 847-858]에 예시된 물질을 함유하는 저해제 상에서 성장하는데 대하여 균주를 선별함으로써 선택될 수 있고, 여기서 저해제 내성 사카로마이세스 세레비지아 균주 ATCC 26602가 선택되었다.

[0048] **형질전환**

본 발명에 따른 폴리뉴클레오티드는 적합한 숙주에서 발현될 수 있다. 따라서 표준 형질전환 기법이 사용될 수 있다.

[0050] 본 발명은 추가로 앞서 기재된 바와 같은 폴리뉴클레오티드를 포함하는 핵산 구조물, 예를 들어 벡터에 관한 것이다.

[0051] 이와 같이 본 발명의 또 다른 양태는 글리옥살라아제 폴리펩티드 단백질을 코딩하는 본 발명의 폴리뉴클레오티드 또는 이의 작용성 등가물을 포함하는 클로닝 및 발현 벡터를 비롯한 벡터, 이러한 벡터를 적합한 숙주 세포에서, 예를 들면 본 발명의 글리옥살라아제의 발현이 초래되는 조건하에 성장시키거나, 형질전환시키거나, 형질감염시키는 방법에 관련된다. 본원에 사용될 경우, 용어 "벡터"는 이것이 연결된 또 다른 핵산을 수송할 수 있는 핵산 분자를 지칭한다.

[0052] 본 발명의 폴리뉴클레오티드는 재조합 복제가능 벡터, 예를 들면 클로닝 또는 발현 벡터내로 흡입될 수 있다. 벡터는 상용성 숙주 세포에서 핵산을 복제하기 위해 사용될 수 있다. 그러므로 추가의 실시태양에서, 본 발명은 본 발명의 폴리뉴클레오티드를 복제가능 벡터내로 도입하고, 벡터를 상용성 숙주 세포내로 도입하고, 숙주 세포를 벡터의 복제를 가져오는 조건하에 성장시킴으로써 본 발명의 폴리뉴클레오티드를 제조하는 방법을 제공한다. 벡터는 숙주 세포로부터 회수될 수 있다. 적합한 숙주 세포는 이후 기재된다.

[0053] 당분야의 숙련가라면, 발현 벡터의 고안이 형질전환되는 숙주 세포의 선택, 원하는 단백질의 발현 수준 등과 같은 인자들에 의해 좌우될 수 있음을 인식할 것이다. 본 발명의 벡터, 예컨대 발현 벡터는 숙주 세포내로 도입됨으로써, 본원에 기재된 바와 같은 핵산에 의해 코딩되는 단백질 또는 웨პ티드를 생산할 수 있다. 본 발명의 벡터, 예컨대 재조합 발현 벡터는 원핵 또는 진핵 세포에서의 글리옥살라아제 폴리펩티드 단백질의 발현을 위해 고안될 수 있다.

[0054] 예를 들면, 글리옥살라아제 폴리펩티드는 세균 세포, 예컨대 에스카리키아 콜라이, 곤충 세포[바클로바이러스 (baculovirus) 발현 벡터 사용], 사상균류, 효모 세포 또는 포유동물 세포에서 발현될 수 있다. 적합한 숙주 세포는 추가로 고에델(Goeddel)의 문헌[*Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990)]에 논의되어 있다. 적절한 숙주의 대표적인 예는 이후 기재된다.

[0055] 상기 기재된 숙주 세포에 대한 적절한 배양 배지 및 조건은 당분야에 공지되어 있다.

[0056] 대부분의 사상균류 및 효모의 경우, 벡터 또는 발현 구조물은 바람직하게는 안정한 형질전환체를 수득하기 위해 숙주 세포의 계놈내로 통합된다. 그러나, 특정 효모의 경우 또한 적합한 에피솜성(episomal) 벡터가 안정하고 높은 수준의 발현을 위해 발현 구조물이 흡입되어지는데 이용가능하고, 이의 예로는 각각 사카로마이세스 및 클루이베로마이세스의 2  $\mu$  및 pKD1 플라스미드로부터 유래된 벡터, 또는 AMA 서열을 함유한 벡터[예를 들어, 아스페르길러스(Aspergillus)로부터의 AMA1]가 포함된다. 발현 구조물이 숙주 세포 계놈내로 통합되는 경우, 구조물은 계놈에서 랜덤한 유전자자리에, 또는 상동성 재조합을 사용하여 예정된 표적 유전자자리에 통합되고, 이러한 경우에 표적 유전자자는 바람직하게는 고도로 발현된 유전자를 포함한다.

[0057] 따라서, 본 발명에 유용한 발현 벡터로는 염색체성-, 에피솜성- 및 바이러스-유래된 벡터, 예를 들어, 세균 플라스미드, 박테리오파아지(bacteriophage), 효모 에피솜, 효모 염색체 요소, 바이러스, 예컨대 바클로바이러스, 파포바(papova) 바이러스, 백시나(vaccinia) 바이러스, 아데노바이러스(adenovirus), 계두(fowl pox) 바이러스, 가성광견병(pseudorabies) 바이러스 및 레트로바이러스(retrovirus)로부터 유래된 벡터, 및 이의 조합으로부터 유래된 벡터, 예컨대 플라스미드 및 박테리오파아지 유전적 유소, 예컨대 코스미드(cosmid) 및 파지미드(phagemid)로부터 유래된 벡터가 포함된다.

[0058] 벡터는 진핵생물 계놈 서열 또는 바이러스 계놈 서열에 상동성인 서열을 포함하는 RNA를 일으키는 폴리뉴클레오티드를 플랭킹(flanking)하는 서열을 추가로 포함할 수 있다. 이는 본 발명의 폴리뉴클레오티드가 숙주 세포의 계놈내로 도입될 수 있도록 허용할 것이다.

[0059] 통합적 클로닝 벡터는, 이것이 통합되어지는 숙주 세포의 염색체(들)에서 예정된 표적 유전자자리에 또는 랜덤

하게 통합될 수 있다.

[0060] 벡터 시스템은 단일 벡터, 예컨대 단일 플라스미드, 또는 숙주 세포의 계놈내로 도입되는 전체 DNA를 함께 함유하는 둘 이상의 벡터, 예컨대 둘 이상의 플라스미드일 수 있다.

[0061] 벡터는 안티센스(antisense) RNA의 생산을 제공하는 안티센스 방향으로 배향된 본 발명의 폴리뉴클레오티드를 포함할 수 있다.

[0062] 벡터 DNA는 종래의 형질전환 또는 형질감염 기법을 통해 원핵 또는 진핵 세포내로 도입될 수 있다. 본원에 사용될 경우, 용어 "형질전환" 및 "형질감염"은 외래 핵산(예를 들어, DNA)을 숙주 세포내로 도입하기 위한 당분야에 인식된 다양한 기법, 예컨대 인산 칼슘 또는 염화 칼슘 공-침전, DEAE-덱스트란-중재된 형질감염, 형질도입, 감염, 리포펙션(lipofection), 양이온성 지질-중재된 형질감염 또는 전기천공을 지칭하려는 것이다. 숙주 세포를 형질전환 또는 형질감염하기 위한 적합한 방법은 샘브룩(Sambrook) 등의 문헌[*Molecular Cloning: A Laboratory Manual, 2nd, ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989*], 데이비스(Davis) 등의 문헌[*Basic Methods in Molecular Biology* (1986)] 및 다른 실험실 매뉴얼에서 찾아볼 수 있다.

[0063] 본 발명에 따른 폴리뉴클레오티드는 구성적으로 발현된다. 본원에서 용어 "구성적"은, 대개 숙주 세포와 관련하여, 폴리뉴클레오티드가 이것이 숙주에 천연적으로 존재하지 않을 경우(또는 존재하지 않을 때만) 존재함을 의미한다. 이러한 의미에서, 이는 또한 폴리뉴클레오티드가 인간 개입에 의해 세포내로 옮겨지는 것으로 설명될 수 있다. 구성적으로 발현된 유전자의 하나의 특징은 이러한 유전자가 성장 조건, 예컨대 탄소 공급원 또는 세포의 성장 상과 무관하게 연속적으로 전사된다는 것이다. 폴리뉴클레오티드는 숙주 세포의 계놈에 이종성일 수 있다. 용어 "이종성"은, 대개 숙주 세포와 관련하여, 폴리뉴클레오티드가 숙주 세포의 계놈에서 천연적으로 발생되지 않거나 폴리펩티드가 세포에 의해 천연적으로 생산되지 않음을 의미한다.

[0064] 폴리뉴클레오티드는 숙주 세포의 계놈에 상동성일 수 있다. 용어 "상동성"은, 대개 숙주 세포와 관련하여, 폴리뉴클레오티드가 숙주 세포의 계놈에서 천연적으로 발생되거나 폴리펩티드가 세포에 의해 천연적으로 생산됨을 의미한다.

[0065] 또 다른 실시태양에서, 본 발명은 본 발명에 의해 포괄되는 핵산을 함유하는 세포, 예를 들어, 형질전환된 숙주 세포 또는 재조합 숙주 세포를 특징으로 한다. "형질전환된 세포" 또는 "재조합 세포"는 재조합 DNA 기법의 수단에 의해, 본 발명에 따른 핵산이 도입되어지는 세포(또는 이의 이전 세대)이다. 원핵 및 진핵 세포 둘 다가 포함되고, 예를 들어, 세균, 균류, 효모 등이 포함되고, 특별히 효모 세포, 예를 들어 사카로마이세스, 예를 들면 사카로마이세스 세레비지아가 바람직하다.

[0066] 삽입된 서열의 발현을 조절하거나, 특정한 원하는 양식으로 유전자 생성물을 변형 및 가공하는 숙주 세포가 선택될 수 있다. 단백질 생성물의 이러한 변형(예를 들어, 글리코실화) 및 가공(예를 들어 분할)은 단백질의 최적의 기능을 촉진시킬 수 있다.

[0067] 다양한 숙주 세포는 단백질 및 유전자 생성물의 번역후 가공 및 변형을 위한 특징적 및 특이적 기작을 갖는다. 분자 생물학 및/또는 미생물학 분야의 숙련가에게 친숙한 적절한 세포주 또는 숙주 시스템은 발현되는 외래 단백질의 원하는 정확한 변형 및 가공을 확보하기 위해 선택될 수 있다. 이를 위해, 1차 전사체의 적절한 가공, 글리코실화, 및 유전자 생성물의 인산화를 위한 세포 기작을 소유한 진핵 숙주 세포가 사용될 수 있다. 이러한 숙주 세포는 당분야에 잘 공지되어 있다.

[0068] 원한다면, 상기 기재된 바와 같은 세포는 본 발명에 따른 폴리펩티드의 제조에 사용될 수 있다. 이러한 방법은 전형적으로 숙주 세포(예를 들어, 상기 기재된 바와 같은 발현 벡터에 의해 형질전환되거나 형질감염됨)를, 폴리펩티드를 코딩하는 코딩 서열의 발현(벡터에 의해)을 제공하는 조건하에 배양하는 단계, 및 선택적으로 발현된 폴리펩티드를 회수하는 단계를 포함한다. 본 발명의 폴리뉴클레오티드는 복제가능한 재조합 벡터, 예를 들어 발현 벡터내로 혼입될 수 있다. 벡터는 상용성 숙주 세포에서 핵산을 복제하기 위해 사용될 수 있다. 그러므로 추가의 실시태양에서, 본 발명은 본 발명의 폴리뉴클레오티드를 복제가능한 벡터내로 도입하고, 벡터를 상용성 숙주 세포내로 도입하고, 벡터의 복제를 가져오는 조건하에 숙주 세포를 성장시킴으로써 본 발명의 폴리뉴클레오티드를 제조하는 방법을 제공한다. 벡터는 숙주 세포로부터 회수될 수 있다.

[0069] 벡터는 상기 기재된 바와 같이 적합한 숙주 세포내로 형질전환되거나 형질감염되어 본 발명의 폴리펩티드의 발현을 제공할 수 있다. 이러한 공정은 폴리펩티드를 코딩하는 코딩 서열의 벡터에 의한 발현을 제공하는 조건하

에 상기 기재된 바와 같은 발현 베타로 형질전환된 숙주 세포를 배양하는 단계를 포함할 수 있다.

[0070] 본원에서 표준 단리, 하이브리드화(hybridization), 형질전환 및 클로닝 기법이 사용된다[예를 들어, 샘브룩(Sambrook, J.), 프리취(Fritsh, E. F.), 및 마니아티스(Maniatis, T.)의 문헌 "*Molecular Cloning: A Laboratory Manual*.2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989"].

#### [0071] 상동성 및 동일성

[0072] 아미노산 또는 뉴클레오티드 서열은 특정 수준의 유사성을 나타낼 때 상동성인 것으로 일컬어 진다. 상동성인 2종의 서열은 공통의 진화 기원을 나타낸다. 2종의 상동성 서열이 가깝게 관련되는지 더욱 멀리 관련되는지의 여부는 "동일성 퍼센트" 또는 "유사성 퍼센트"에 의해 지시되고, 이는 각각 높거나 낮다. 이의 제기될 수 있지만, "동일성 퍼센트" 또는 "유사성 퍼센트", "상동성 수준" 또는 "상동성 퍼센트"는 종종 상호교환적으로 사용된다.

[0073] 서열의 비교 및 두 서열 간의 동일성 퍼센트의 결정은 수학적 알고리즘(algorithm)을 사용하여 달성될 수 있다. 숙련가라면 수 개의 상이한 컴퓨터 프로그램이 2개의 서열을 정렬하고 2개의 서열 사이의 상동성을 결정할 수 있다는 사실을 인식할 것이다[크루스칼(Kruskal, J. B)의 문헌 "(1983) An overview of sequence comparison In D. Sankoff and J. B. Kruskal, (ed.), Time warps, string edits and macromolecules: the theory and practice of sequence comparison, pp. 1-44 Addison Wesley"]. 2개의 아미노산 서열 사이의 동일성 퍼센트는 2개의 서열의 정렬을 위해 니들만(Needleman) 및 운쉬(Wunsch) 알고리즘을 사용하여 결정될 수 있다[니들만(Needleman, S. B.) 및 운쉬(Wunsch, C. D.)의 문헌 "(1970) J. Mol. Biol. 48, 443-453"]. 이 알고리즘은 아미노산 서열 뿐만 아니라 뉴클레오티드 서열을 정렬한다. 니들만-운쉬 알고리즘은 컴퓨터 프로그램 NEEDLE에서 실행되어 왔다. 본 발명의 목적을 위해, EMBOSS 패키지로부터의 NEEDLE 프로그램이 사용되었다(버전 2.8.0 또는 그 이상, EMBOSS: The European Molecular Biology Open Software Suite (2000) Rice, P. LongdenJ. and BleasbyA Trends in Genetics 16, (6) pp276-277, <http://emboss.bioinformatics.nl/>). 단백질 서열의 경우, EBLOSUM62가 치환 행렬을 위해 사용된다. 뉴클레오티드 서열의 경우, EDNAFULL이 사용된다. 다른 행렬이 특정화될 수 있다. 아미노산 서열의 정렬을 위해 사용되는 선택적 매개변수는 10의 캡-개방 폐널티(gap-open penalty) 및 0.5의 캡 연장 폐널티이다. 숙련가라면, 모든 이들 상이한 매개변수가 약간 다른 결과를 산출하지만, 2개의 서열의 전체적인 동일성 백분율은 상이한 알고리즘을 사용할 때 그다지 변경되지 않음을 인식할 것이다.

#### [0074] 전반적 상동성 규정

[0075] 상동성 또는 동일성은 임의의 캡 또는 연장부를 포함하는 전체 정렬된 영역에 대한 2개의 완전한 서열 사이의 동일한 부합성에 대한 백분율이다. 2개의 정렬된 서열 사이의 상동성 또는 동일성은 다음과 같이 계산될 수 있다: 캡을 포함하는 정렬부의 총 길이로 나누어진, 두 서열에서의 동일한 아미노산을 나타내는 정렬부 중의 상응하는 위치의 수. 본원에서와 같이 규정된 동일성은 NEEDLE로부터 수득될 수 있고, "동일성"으로서 프로그램의 출력물에 표지될 수 있다.

#### [0076] 가장 긴 동일성 규정

[0077] 2개의 정렬된 서열 사이의 상동성 및 동일성은 다음과 같이 계산된다: 정렬부에서 캡의 총 수가 차감된 후 정렬부의 총 길이로 나눠진, 두 서열중의 동일한 아미노산을 나타내는 정렬부 중의 상응하는 위치의 수. 본원에서와 같이 규정된 동일성은 NOBRIEF 옵션을 사용하여 NEEDLE로부터 수득될 수 있고, "가장 긴-동일성"으로서 프로그램의 출력물에 표지될 수 있다.

[0078] 상기 본원에 정의된 서열 번호 20에 대한 동일성과 함께, 본 발명에 따른 세포에서 활성인 GL01 아미노산에 대한 개요가 표 1에 제공된다.

## 표 1

본 발명에 따른 세포(실시예 참조)에서 활성인 GL01 아미노산 및 사카로마이

세스 세레비지아의 GL01(서열 번호 20)에 대한 동일성에 대한 개요.

GL01의 공급원	서열 번호 20에 대한 동일성%	서열 번호
사카로마이세스 세레비지아	100	20
칸디다 글라브라타	69	35
클루이베로마이세스 락티스	61	37
자이고사카로마이세스 룩시	61	38
칸디다 마그놀리아	45	36

[0079]

[0080] 따라서, 하나의 실시태양에서 GL01 뉴클레오티드 또는 아미노산은 사카로마이세스 세레비지아, 칸디다 글라브라타, 클루이베로마이세스 락티스, 자이고사카로마이세스 룩시 및 칸디다 마그놀리아로 구성된 군에서 선택된 미생물로부터 유래된다.

[0081]

[0081] 하나의 실시태양에서 GL01 뉴클레오티드 또는 아미노산은 사카로마이세스 세레비지아, 칸디다 글라브라타, 클루이베로마이세스 락티스 및 칸디다 마그놀리아로 구성된 군에서 선택된 미생물로부터 유래된다.

[0082]

[0082] 하나의 실시태양에서, GLO 뉴클레오티드는 코돈 쌍 최적화된 뉴클레오티드 서열이고, 여기서 서열은 서열 번호 34, 서열 번호 35, 서열 번호 36, 서열 번호 37, 서열 번호 38이다.

[0083]

[0083] 하나의 실시태양에서, GLO 뉴클레오티드는 코돈 쌍 최적화된 뉴클레오티드 서열이고, 여기서 서열은 서열 번호 34, 서열 번호 35, 서열 번호 36, 또는 서열 번호 37이다.

[0084]

[0084] 본원에 기재된 본 발명의 다양한 실시태양은 교차-조합될 수 있다. 본 발명은 청구항 1에 따른 세포에 관한 것이다. 이러한 세포는 본원에서 또한 형질전환된 숙주 세포로서 지정된다. 이의 실시태양이 이제 기재된다.

[0085]

[0085] 세포는 L-아라비노스 및 자일로스를 사용할 수 있다. 하나의 실시태양에서, 세포는 L-아라비노스를 L-리불로스 및/또는 자일룰로스 5-포스페이트로, 및/또는 원하는 발효 생성물로, 예를 들면 본원에 언급된 생성물들중 하나로 전환시킬 수 있다.

[0086]

[0086] L-아라비노스로부터 에탄올을 생산할 수 있는 유기체, 예를 들면 사카로마이세스 세레비지아 균주는 적합한 공급원으로부터의 araA(L-아라비노스 이소머라아제), araB(L-리불로키나아제) 및 araD(L-리불로스-5-P4-에피머라아제) 유전자를 도입하도록 숙주 세포를 변형시킴으로써 생산될 수 있다. 이러한 유전자는 아라비노스를 사용할 수 있도록 숙주 세포내로 도입될 수 있다. 제공된 이러한 접근법은 국제특허공개 제WO2003/095627호에 기재되어 있다. 락토바실러스 플란타룸(*Lactobacillus plantarum*)으로부터의 araA, araB 및 araD 유전자가 사용될 수 있고, 국제특허공개 제WO2008/041840호에 개시되어 있다. 바실러스 서브틸리스(*Bacillus subtilis*)로부터의 araA 유전자 및 에스캐리키아 콜라이로부터의 araB 및 araD 유전자가 사용될 수 있고, 제EP1499708호에 개시되어 있다. 또 다른 실시태양에서, araA, araB 및 araD 유전자는, 국제특허공개 제WO2009/011591호에 개시된 바와 같이, 클라비박터(*Clavibacter*), 아르트로박터(*Arthrobacter*) 및/또는 그라멜라(*Gramella*) 속중 1종 이상, 특별히 클라비박터 미키가넨시스(*Clavibacter michiganensis*), 아르트로박터 아우레센스(*Arthrobacter aurescens*), 및/또는 그라멜라 포르세티(*Gramella forsetii*)중 하나로부터 유래될 수 있다.

[0087]

[0087] 하나의 실시태양에서, 형질전환된 숙주 세포는 자일로스 이소머라아제 유전자의 하나 이상의 복사물 및/또는 자일로스 환원효소 및/또는 자일리톨 탈수소효소의 하나 이상의 복사물을 포함할 수도 있다.

[0088]

[0088] 복사물의 수는 임의의 공지된 방법에 의해 숙련가에 의해 결정될 수 있다. 하나의 실시태양에서, 형질전환된 숙주 세포는 글루코스, 아라비노스, 자일로스 및 갈락토스를 발효시킬 수 있다.

[0089]

[0089] 하나의 실시태양에서, 세포는 이용가능한 글루코스, 자일로스 아라비노스, 갈락토스 및 만노스를 90% 이상 발효 생성물로 전환시킬 수 있다. 하나의 실시태양에서, 세포는 이용가능한 모든 글루코스, 자일로스 아라비노스, 갈락토스 및 만노스의 91% 이상, 92% 이상, 94% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상 또는 100%를 발효 생성물로 전환시킬 수 있다.

- [0090] 본 발명의 한 실시태양에서, 형질전환된 숙주 세포는 하나 이상의 추가의 당분, 바람직하게는 C5 및/또는 C6 당분, 예를 들어 만노스를 발효시킬 수 있다. 본 발명의 하나의 실시태양에서, 형질전환된 숙주 세포는; 형질전환된 숙주 세포가 자일로스를 발효시키도록 허용하는 xylA-유전자, XYL1 유전자 및 XYL2 유전자 및/또는 XKS1-유전자중 하나 이상을 포함하고; 알도스 환원효소(GRE3) 유전자가 결실되고; PPP-유전자 TAL1, TKL1, RPE1 및 RKI1가 과발현되어 세포에서 5탄당 포스페이트 경로를 통해 플러스(flux)의 증가를 허용한다.
- [0091] 하나의 실시태양에서, 형질전환된 숙주 세포는 산업용 세포, 더 바람직하게는 산업용 효모이다(숙주 부분에서 상기 규정된 바와 같음).
- [0092] 하나의 실시태양에서 형질전환된 숙주 세포는 저해제 내성이다. 저해제 내성은 저해 화합물에 대한 저항성이다. 리그노셀룰로스에서 저해 화합물의 존재 및 수준은 공급원료의 변화, 예비처리 방법, 가수분해 공정에 의해 매우 달라질 것이다. 저해제 범주의 예는 카복실산, 푸란 및/또는 폐놀계 화합물이다. 카복실산의 예는 락트산, 아세트산 또는 포름산이다. 푸란의 예는 푸르푸랄 및 하이드록시-메틸푸르푸랄이다. 폐놀계 화합물의 예는 바닐린(vanillin), 시링산(syringic acid), 폐롤산 및 쿠마르산이다. 저해제의 전형적인 양은 카복실산의 경우: 공급원료, 예비처리 및 가수분해 조건에 따라 리터 당 수 g 내지 리터 당 20 g까지, 또는 그 이상이다. 푸란의 경우: 공급원료, 예비처리 및 가수분해 조건에 따라 리터 당 수 백 mg 내지 리터 당 수 g까지이다.
- [0093] 폐놀계 화합물의 경우, 공급원료, 예비처리 및 가수분해 조건에 따라 리터 당 수십 mg 내지 리터 당 1 g까지이다.
- [0094] 본 발명에 따른 형질전환된 숙주 세포는 저해제 내성일 수 있고, 즉 이들은 이들이 전형적으로 통상의 예비처리 및 가수분해 조건에 의한 수준에서 통상의 저해제를 견뎌내어, 형질전환된 숙주 세포가 광범위한 적용범위를 가질 수 있고, 즉 이는 상이한 공급원료, 상이한 예비처리 방법 및 상이한 가수분해 조건에 대하여 높은 적용성을 가진다.
- [0095] 한 실시태양에서, 산업용 형질전환된 숙주 세포는 저해제 내성 숙주 세포에 기초하여 작성되고, 여기서 작성은 이후 본원에 기재된 바와 같이 이루어 진다. 저해제 내성 숙주 세포는, 예컨대 카다(Kadar) 등의 문헌[Appl. Biochem. Biotechnol. (2007), Vol. 136-140, 847-858]에 예시된 물질을 함유하는 저해제 상에서 성장하는데 대하여 균주를 선별함으로써 선택될 수 있고, 여기서 저해제 내성 사카로마이세스 세레비지아 균주 ATCC 26602 가 선택되었다.
- [0096] 하나의 실시태양에서, 형질전환된 숙주 세포는 마커(marker)-부재이다. 본원에 사용될 경우, 용어 "마커"는, 마커를 함유하는 숙주 세포의 선택 또는 선별을 허용하는 특질 또는 표현형을 코딩하는 유전자를 지칭한다. 마커-부재는, 마커가 형질전환된 숙주 세포에 본질적으로 존재하지 않음을 의미한다. 마커-부재는 항생성 마커가 형질전환된 숙주 세포의 작성시 사용되고 이후 제거될 때 특히 유리하다. 마커의 제거는 임의의 적합한 선행 기술 기법, 예를 들어 문자내 재조합을 사용하여 실행될 수 있다. 마커 제거의 적합한 방법이 실시예에 예시된다.
- [0097] 형질전환된 숙주 세포는 식물 생물자원, 셀룰로스, 헤미셀룰로스, 페틴, 전분, 전분 유도체를, 예를 들면 발효 가능한 당분으로 전환시킬 수 있다. 따라서, 형질전환된 숙주 세포는 하나 이상의 효소, 예컨대 셀룰로스를 글루코스 단량체로 전환시키고 헤미셀룰로스를 자일로스 및 아라비노스 단량체로 전환시키는데 필수적인 셀룰라아제(cellulase)(엔도셀룰라아제 또는 엑소셀룰라아제), 헤미셀룰라아제(엔도- 또는 엑소-자일라나아제(xylanase) 또는 아라비나아제(arabinase)), 페틴을 글루코론산 및 갈락토론산으로 전환시킬 수 있는 페티나아제(pectinase) 또는 전분을 글루코스 단량체로 전환시키는 아밀라아제(amylose)를 발현할 수 있다.
- [0098] 형질전환된 숙주 세포는 추가로 당분을 원하는 발효 생성물, 예컨대 에탄올, 부탄올, 락트산, 3-하이드록시-프로피온산, 아크릴산, 아세트산, 석신산, 시트르산, 푸마르산, 말산, 이타콘산, 아미노산, 1,3-프로판-디올, 에틸렌, 글리세롤,  $\beta$ -락탐 항생물질 또는 세팔로스포린(cephalosporin)으로 전환시키는데 필요한 이들 효소 활성을 포함할 수 있다.
- [0099] 하나의 실시태양에서, 형질전환된 숙주 세포는 알코올성 발효, 바람직하게는, 혐기성 알코올성 발효를 천연적으로 가능하게 하는 세포이다. 형질전환된 숙주 세포는 바람직하게는 에탄올에 대한 높은 내성, 낮은 pH에 대한 높은 내성(즉, 약 5, 약 4, 약 3, 또는 약 2.5 보다 낮은 pH에서 성장 가능함) 및 유기산에 대한 높은 내성 및/또는 상승된 온도에 대한 높은 내성을 갖는다.

[0100] 형질전환된 숙주 세포의 임의의 상기 특징 또는 활성은 세포에 천연적으로 존재할 수 있거나, 유전자 변형에 의해 도입되거나 변형될 수 있다.

형질전환된 숙주 세포의 작성

[0102] 하나의 실시태양에 따라서, 유전자는, a) 내지 e)중 하나 이상을 숙주 세포내로 도입하고 순응적으로 진화시킴으로써 숙주 세포에 도입되어, 형질전환된 숙주 세포를 생산할 수 있다:

a) 강한 구성적 프로모터(들)의 제어하에서의 유전자 araA, araB 및 araD;

b) 선택적으로 하나 이상의 강한 구성적 프로모터의 제어하에서의 PPP-유전자 TAL1, TKL1, RPE1 및 RKI1;

c) 알도스 환원효소 유전자의 결합;

d) 강한 구성적 프로모터(들)의 제어하에서의 xylA-유전자 및 XKS1-유전자;

e) 강한 구성적 프로모터의 제어하에서의 xylA 유전자(이는 다수의 유전자자리 상에서 게놈내로 삽입하는 능력을 가짐).

[0108] 상기 세포는 재조합 발현 기법을 사용하여 작성될 수 있다.

재조합 발현

[0110] 형질전환된 숙주 세포는 재조합 세포이다. 말하자면, 형질전환된 숙주 세포는 해당 세포에서 천연적으로 초래되지 않는 뉴클레오티드 서열을 포함하거나, 이에 의해 형질전환되거나, 이에 의해 유전적으로 변형된다.

[0111] 세포에서 효소의 재조합 발현을 위한 기법, 뿐만 아니라 형질전환된 숙주 세포의 추가적 유전자 변형을 위한 기법은 당분야의 숙련가에게 잘 공지되어 있다. 전형적으로 이러한 기법은 관련 서열을 포함하는 핵산 구조물을 위한 세포의 형질전환을 포함한다. 이러한 방법은, 예를 들면, 표준 편람, 예컨대 샘브룩(Sambrook) 및 러셀(Rusell)의 문헌[(2001) "Molecular Cloning: A Laboratory Manual(3rd edition), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, or F. Ausubel et al., eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987)]으로부터 공지되어 있다. 숙주 세포의 형질전환 및 유전자 변형을 위한 방법은 예를 들어 제EP-A-0635574호, 국제특허공개 제W098/46772호, 제W099/60102호, 제W000/37671호, 제W090/14423호, 제EP-A-0481008호, 제EP-A-0635574호 및 제US6,265,186호로부터 공지되어 있다.

[0112] 전형적으로, 핵산 구조물은 플라스미드, 예를 들어 저 복사 플라스미드 또는 고 복사 플라스미드일 수 있다. 본 발명에 따른 세포는, 예를 들어 뉴클레오티드 구조물의 다수의 복사물에 의해, 또는 효소 서열의 다수의 복사물을 갖는 구조물의 사용에 의해, 효소를 코딩하는 뉴클레오티드 서열의 1개의 복사물 또는 다수의 복사물을 포함할 수 있다.

[0113] 핵산 구조물은 에피솜성으로 유지될 수 있고, 그러므로 자율적 복제를 위한 서열, 예컨대 상염색체 복제 서열을 포함한다. 적합한 에피솜성 핵산 구조물은 예를 들어 효모  $2\ \mu$  또는 pkD1 플라스미드[글리어(Gleer) 등의 문헌 "1991, Biotechnology 9: 968-975"], 또는 AMA 플라스미드[피에로(Fierro) 등의 문헌 "1995, Curr Genet. 29:482-489"]에 기초할 수 있다. 다르게는, 각각의 핵산 구조물은 세포의 게놈내로 하나 이상의 복사물로 통합될 수 있다. 세포의 게놈내로의 통합은 비-상동성 재조합에 의해 랜덤하게 일어나지만, 바람직하게는, 핵산 구조물은 당분야에 공지된 바와 같이 상동성 재조합에 의해 세포의 게놈내로 통합될 수 있다(예를 들어 국제특허공개 제W090/14423호, 제EP-A-0481008호, 제EP-A-0635574호 및 제US6,265,186호 참조).

[0114] 대부분의 에피솜성 또는  $2\ \mu$  플라스미드는 효모에서 비교적 불안정하고, 각각의 세대 이후 대략  $10^{-2}$  이상의 세포에서 소실된다. 선택적 성장의 조건하에서 조차도, 세포의 단지 60% 내지 95%가 에피솜성 플라스미드를 보유한다. 대부분의 에피솜성 플라스미드의 복사물 수는  $cir^+$  숙주의 세포 당 20 내지 100개의 범위이다. 그러나, 플라스미드는 세포들 사이에 동일하지 않게 분포되고, 개체군에서 세포에 대한 복사물 수에 있어서 높은 편차가 존재한다. 통합적 플라스미드에 의해 형질전환된 균주는, 선택적 압력의 부재하에서 조차도 극히 안정적이다. 그러나, 플라스미드 손실은 나란히 반복된 DNA 사이의 상동성 재조합에 의해 대략  $10^{-3}$  내지  $10^{-4}$  빈도로 일어나서, 벡터 서열 밖으로 루핑(looping)되도록 유도한다. 바람직하게는, 안정한 통합의 경우 벡터 고안은 그러므로 선택 마커 유전자의 손실시(이는 또한 분자내 상동성 재조합에 의해 초래됨) 통합된 구조물 밖으로의 루핑은 더 이상 가능하지 않게 된다. 바람직하게는 유전자는 그러므로 안정하게 통합된다. 안정한 통합

은 본원에서 게놈내로의 통합으로서 규정되고, 여기서 통합된 구조물 밖으로의 투평은 더이상 가능하지 않다. 바람직하게는 선택 마커는 존재하지 않는다. 전형적으로, 효소 코딩 서열은 효소 서열의 전사 및/또는 번역을 제공하거나 보조할 수 있는 하나 이상의 핵산 서열에 작동적으로 연결될 것이다.

[0115] 용어 "작동적으로 연결된"은 기재된 성분들이 이들의 의도된 방식으로 작용하도록 이들을 허용하는 관계로 존재하는 병렬을 지칭한다. 예를 들어, 프로모터 또는 증폭자(enhancer)는 코딩 서열에 작동적으로 연결되고, 상기 프로모터 또는 증폭자는 코딩 서열의 전사에 영향을 준다.

[0116] 본원에 사용될 경우, 용어 "프로모터"는 하나 이상의 유전자의 전사를 제어하는 작용을 하고, 유전자의 전사 개시 자리의 전사 방향에 대하여 상류에 위치하며, DNA-의존적 RNA 폴리머라아제(polymerase)를 위한 결합 자리, 전사 개시 자리 및 당분야의 숙련가에게 공지된 임의의 다른 DNA 서열의 존재에 의해 구조적으로 식별되는 핵산 단편을 지칭한다. "구성적(constitutive)" 프로모터는 대부분의 환경적 및 발달학적 조건하에 활성인 프로모터이다. "유도가능한(inducible)" 프로모터는 환경적 또는 발달학적 규제하에 활성인 프로모터이다.

[0117] 본 발명에 따른 효소를 코딩하는 뉴클레오티드 서열의 발현을 달성하기 위해 사용될 수 있는 프로모터는 발현될 효소를 코딩하는 뉴클레오티드 서열에 고유적이지 않을 수 있고, 즉 이것이 작동적으로 연결된 뉴클레오티드 서열(코딩 서열)에 이종성인 프로모터일 수 있다. 그러나, 프로모터는 상동성, 즉 숙주 세포에 내생성일 수 있다.

[0118] 프로모터는 광범위하게 이용가능하고 숙련가에게 공지되어 있다. 이러한 프로모터의 적합한 예로는, 예를 들어 당분해 유전자로부터의 프로모터, 예컨대 효모 또는 사상균류로부터의 포스포프루토키나아제(PFK: phosphofructokinase), 트리오스 포스페이트 이소미라아제(TPI: triose phosphate isomerase), 글리세를데하이드-3-포스페이트 탈수소효소(GPD, TDH3 또는 GAPDH), 피루베이트 키나아제(PYK: pyruvate kinase), 포스포글리세레이트 키나아제(PGK: phosphoglycerate kinase) 프로모터가 포함되고; 효모로부터의 이러한 프로모터에 대한 추가의 세부사항을 찾아볼 수 있다(국제특허공개 제W093/03159호). 다른 유용한 프로모터는 리보솜성 단백질 코딩 유전자 프로모터, 락타아제(lactase) 유전자 프로모터(LAC4), 알코올 탈수소효소 프로모터(ADH1, ADH4 등) 및 에놀라아제(enolase) 프로모터(ENO)이다. 다른 프로모터(구성적 및 유도가능한 프로모터 둘다), 및 증폭자 또는 상류 활성화 서열은 당분야의 숙련가에게 공지될 것이다. 본 발명의 숙주 세포에 사용되는 프로모터는, 원활 경우, 이들의 제어 특징에 영향을 주도록 변형될 수 있다. 이와 관련하여 적합한 프로모터는 구성적 및 유도가능한 천연 프로모터 둘 다 뿐만 아니라 조작된 프로모터를 포함하고, 이는 당분야의 숙련가에게 잘 공지되어 있다. 진핵 숙주 세포에서 적합한 프로모터는 GAL7, GAL10, 또는 GAL1, CYC1, HIS3, ADH1, PGL, PH05, GAPDH, ADC1, TRP1, URA3, LEU2, EN01, TPI1, 및 AOX1일 수 있다. 다른 적합한 프로모터로는 PDC1, GPD1, PGK1, TEF1, 및 TDH3이 포함된다.

[0119] 형질전환된 숙주 세포에서, 효소를 코딩하는 뉴클레오티드 산 서열의 3'-말단은 바람직하게는 전사 종결자 서열에 작동적으로 연결된다. 바람직하게는 종결자 서열은 선택된 숙주 세포, 예컨대 예를 들어 선택된 효모 종에서 작동가능하다. 종결자의 선택이 중요하지 않은 임의의 경우에; 이는, 예를 들어 임의의 효모 유전자로부터 일 수 있지만, 종결자는 때때로 비-효모 진핵 유전자로부터인 경우에 작용할 수 있다. 대개 효소를 코딩하는 뉴클레오티드 서열은 종결자를 포함한다. 바람직하게는, 이러한 종결자는 형질전환된 숙주 세포에서 넌센스(nonsense) 중재된 mRNA 분해를 방지하는 돌연변이와 조합된다[예를 들면: 셜리(Shirley) 등의 문헌 "2002, Genetics 161 :1465-1482"을 참조한다].

[0120] 전사 종결 서열은 추가로 바람직하게는 폴리아데닐화 신호를 포함한다.

[0121] 선택적으로, 선택가능한 마커는 본 발명에 사용하기에 적합한 핵산 구조물에 존재할 수 있다. 본원에 사용될 경우, 용어 "마커"는 마커 함유 숙주 세포의 선택 또는 선별을 허용하는 특질 또는 표현형을 코딩하는 유전자를 지칭한다. 마커 유전자는 내항생제성 유전자일 수 있고, 이로써 적절한 항생제가 형질전환되지 않은 세포들 중에서 형질전환된 세포를 선택하기 위해 사용될 수 있다. 적합한 항생제 내성 마커의 예로는, 예를 들어 다이하이드로폴리에이트(dihydrofolate) 환원효소, 하이그로마이신(hygromycin)-B-인산전달효소, 3'-0-인산전달효소 II [카나마이신(kanamycin), 네오마이신(neomycin) 및 G418 내성]가 포함된다. 항생제 내성 마커는 배수체 숙주 세포의 형질전환을 위해 가장 편리할 것이다. 또한 비-항생제 내성 마커가 사용될 수 있고, 예컨대 영양요구성 마커(URA3, TRP1, LEU2) 또는 애스. 품베(*S. pombe*) TPI 유전자[러셀(Russell P R)의 문헌 "1985, Gene 40: 125-130"에 기재됨]이다. 바람직한 실시태양에서, 핵산 구조물로 형질전환된 숙주 세포는 마커 유전자 부재이다. 재조합 마커 유전자 부재 미생물 숙주 세포를 작성하기 위한 방법은 제EP-A-0635 574호에 개시되고, 2방향 성 마커, 예컨대 애이. 니들란스(*A. nidulans*) amdS[아세트아미다아제(acetamidase)] 유전자 또는 효모 URA3

및 LYS2 유전자의 사용에 기초한다. 다르게는, 선별가능한 마커, 예컨대 녹색 형광 단백질(Green Fluorescent Protein), lacL, 루시페라아제(luciferase), 클로람페니콜(chloramphenicol) 아세틸전달효소, 베타-글루쿠로니다아제(glucuronidase)는 본 발명의 핵산 구조물내로 혼입되어 형질전환된 세포에 대한 선별을 허용할 수 있다.

[0122] 본 발명에 사용하기에 적합한 핵산 구조물에 존재할 수 있는 선택적인 추가의 요소로는, 제한되지 않지만, 하나 이상의 리더(leader) 서열, 증폭자, 통합 인자, 및/또는 리포터(reporter) 유전자, 인트론(intron) 서열, 중립 심(centromer), 텔로머(telomeric) 및/또는 매트릭스 결합(matrix attachment)(MAR) 서열이 포함된다. 본 발명의 핵산 구조물은 추가로 자율 복제를 위한 서열, 예컨대 ARS 서열을 포함할 수 있다.

[0123] 그러므로, 재조합 공정은 공지된 재조합 기법에 의해 실행될 수 있다. 형질전환된 숙주 세포에서 효소의 발현 및 과발현을 위한 다양한 수단이 당분야의 숙련가에게 공지되어 있다. 특별히, 효소는 숙주 세포에서 효소를 코딩하는 유전자의 복사물 수를 증가시킴으로써, 예를 들어 숙주 세포의 계놈에서 유전자의 추가의 복사물을 통합함으로써, 에피솜성 다중복사 발현 벡터로부터 유전자를 발현시킴으로써, 또는 유전자의 다수의 복사물을 포함하는 에피솜성 발현 벡터를 도입함으로써 과발현될 수 있다.

[0124] 다르게는, 본 발명의 숙주 세포중의 효소의 과발현은, 과발현될 효소를 코딩하는 서열에 고유적이지 않은 프로모터, 즉 이것이 작동적으로 연결된 코딩 서열에 이종성인 프로모터를 사용함으로써 달성될 수 있다. 프로모터는 바람직하게는 이것이 작동적으로 연결되는 코딩 서열에 이종성일지라도, 또한 프로모터가 상동성, 즉 숙주 세포에 내생성인 것이 바람직하다. 바람직하게는 이종성 프로모터는 코딩 서열에 고유적인 프로모터에 비하여, 코딩 서열을 포함하는 전사체의 더 높은 정지 상태 수준을 생산할 수 있다(또는 시간 단위 당, 더 많은 전사 분자, 즉 mRNA 분자를 생산할 수 있다). 이와 관련하여 적합한 프로모터는 구성적 및 유도가능한 천연 프로모터 둘 다 뿐만 아니라 조작된 프로모터를 포함한다.

[0125] 하나의 실시태양에서, 형질전환된 숙주 세포는 마커 부재이고, 이는 영양요구성 또는 우성 마커, 특별히 항생제 내성 마커가 계놈내에 또는 염색체외에 존재하지 않음을 의미한다.

[0126] 상기 언급된 효소의 과발현을 위해 사용되는 코딩 서열은 바람직하게는 숙주 세포에 상동성일 수 있다. 그러나, 숙주에 이종성인 코딩 서열이 사용될 수 있다.

[0127] 효소의 과발현은, 유전자 변형된 세포에서 효소의 생산을 지칭할 경우, 효소가 동일한 조건하에 변형되지 않은 숙주 세포에 비하여 특이적 효소 활성의 더 높은 수준에서 생산됨을 의미한다. 대개, 이는 효소적으로 활성인 단백질(또는 다중-서브유닛 효소의 경우에 단백질들)이 동일한 조건하에 변형되지 않은 숙주 세포에 비하여 더 많은 양으로, 또는 더 높은 정지 상태 수준에서 생산됨을 의미한다. 유사하게, 이는 대개 효소적으로 활성인 단백질을 코딩하는 mRNA가 동일한 조건하에 변형되지 않은 숙주 세포에 비하여 더 많은 양으로 또는 더 높은 정지 상태 수준에서 생산됨을 의미한다. 바람직하게는 숙주에서, 과발현된 효소는, 과발현을 일으키는 유전자 변형을 제외하고 유전적으로 동일한 균주에 비하여, 적어도 약 1.1배, 약 1.2배, 약 1.5배, 약 2배, 약 5배, 약 10배 또는 약 20배 과발현된다. 이러한 과발현 수준은 효소의 활성의 정지 상태 수준, 효소의 단백질의 정지 상태 수준, 뿐만 아니라 효소를 코딩하는 전사체의 정지 상태 수준에 적용될 수 있다.

[0128] 바람직하게는, 글리وك랄라아제는 사이토졸(cytosol)에서 발현된다.

#### 적응

[0130] 적응은 진화 과정이고, 이로써 개체군은 그의 서식지에 더 잘 순응(적응)된다. 이러한 과정은 몇몇 내지 많은 세대에 걸쳐 발생되고, 생물학의 기본 현상중 하나이다.

[0131] 용어 "적응"은 또한 특별히 유기체의 생존을 위해 중요한 특징을 지칭할 수 있다. 이러한 적응은 자연 선택에 의해, 더욱 성공적으로 번식하는 더 잘 순응된 형태에 의해 가변성 개체군에서 생성된다.

[0132] 환경 조건의 변화는 자연 선택의 결과를 바꾸어, 새로운 조건하에서 유기체의 적합성을 개선시키는 후속적인 적응의 선택적 이점에 영향을 준다. 극한 환경 변화의 경우, 유리한 적응의 출현 및 고정은 생존을 위해 필수적일 수 있다. 다수의 상이한 인자들, 예컨대 예를 들어 영양분 이용가능성, 온도, 산소의 이용가능성 등은 적응 진화를 구동시킬 수 있다.

#### 적합성

[0134] 적응성(유기체가 소정의 서식지에서 생존하고 번식할 수 있는 정도) 및 적합성 사이에 분명한 관계가 존재한다. 적합성은 자연 선택의 비율에 대한 예측자 및 추정값이다. 자연 선택의 적용에 의해, 대체 표현형의 상대 빈도

는, 이들이 유전가능하다면, 시간에 따라 달라질 것이다.

[0135] 유전적 변화

[0136] 자연 선택이 개체군의 유전적 다양성에 대하여 작용할 경우, 유전적 변화는 근본적인 기작이다. 이로 인해, 개체군은 유전적으로 그의 환경에 적응한다. 유전적 변화는 변화된 서식지에 순응하는 방식으로 유기체의 생리학적 활성을 조정할 수 있거나, 시각적 구조체를 생성할 수 있다.

[0137] 서식지가 빈번히 변화될 수 있다. 따라서, 적응의 과정은 결코 최종적으로 완료되지 않는다. 시간내에, 환경 변화가 점진적으로 일어나고, 종은 그의 주변에 더욱 더 적합하게 될 수 있다. 한편, 환경에서의 변화가 비교적 신속히 초래되면, 종은 잘 적응하지 못하게 된다. 적응은 유전적 과정이고, 이는 또한 개체군이 서식지 또는 환경을 바꾸지 않을 경우에도 어느 정도 항상 진행된다.

[0138] 적응 진화

[0139] 형질전환된 숙주 세포는 이들의 제조시 적응 진화될 수 있다. 형질전환된 숙주 세포는 자발적인 또는 유도된 (예를 들어, 방사선 또는 화학물질에 의함) 돌연변이체의 선택에 의해, 원하는 당분 상에서의 성장을 위해, 바람직하게는 유일한 탄소원으로서, 바람직하게는 협기성 조건하에 당분을 활용하도록 적응될 수 있다. 돌연변이체의 선택은, 예를 들어 쿠이퍼(Kuyper) 등의 문헌[2004, FEMS Yeast Res. 4: 655-664]에 기재된 바와 같은 배양물의 연속적 전달을 비롯한 기법에 의해, 또는 물질환경조절장치(chemostat) 배양에서 선택적 압력하에서의 배양에 의해 수행될 수 있다. 예를 들어 바람직한 숙주 세포에서, 상기 기재된 하나 이상의 유전자 변형(돌연변이체의 선택에 의해 수득된 변형 포함)은 탄소 공급원으로서, 바람직하게는 유일한 탄소 공급원으로서의 자일로스 상에서, 바람직하게는 협기성 조건하에 성장할 수 있는 능력을 숙주 세포에 부여한다. XI가 자일로스를 전환시키기 위한 유전자로서 사용될 경우, 바람직하게는 세포는 본질적으로 자일리톨을 생산하지 않고, 예를 들어 생산된 자일리톨은 검출 한계 미만, 또는 예를 들어 물 기준으로 소비된 탄소의 약 5, 약 2, 약 1, 약 0.5, 또는 약 0.3% 미만이다.

[0140] 적응 진화는 또한, 예를 들어 위셀린크(Wisselink H.W.) 등의 문헌[Applied and Environmental Microbiology Aug. 2007, p. 4881-4891]에 기재되어 있다.

[0141] 적응 진화의 한 실시태양에서, 상이한 배지에서 연속 성장의 반복된 주기에 의한 반복된 배치(batch)식 배양으로 구성된 섭생이 적용되고, 예를 들어 상이한 조성을 갖는 3개의 배지이다(글루코스, 자일로스, 및 아라비노스; 자일로스 및 아라비노스). 위셀린크(Wisselink) 등의 문헌[(2009) Applied and Environmental Microbiology, Feb. 2009, p. 907-914]을 참조한다.

[0142] 효모 형질전환 및 유전적 안정성

[0143] 유전자 조작, 즉 재조합 DNA에 의한 효소 세포의 형질전환은 1978년에 최초로 실행가능하게 되었다[베그스(Beggs), 1978; 히넨(Hinnen) 등, 1978]. 이후 효모에서 재조합 DNA 기술은 그 자체로 확립되었다. 다수의 상이한 벡터 구조물이 이용가능하다. 일반적으로, 셔틀 벡터로 지칭되는 이들 플라스미드 벡터는 복제 기원 및 선택가능한 마커[종종  $\beta$ -락타마아제(lactamase) 유전자, ampR]로 구성된 에스케리키아 콜라이 벡터로부터 유래된 유전 물질을 함유하고, 이는 효모 세포로 형질전환되기 이전에 에스케리키아 콜라이에서 이들을 증식되게 할 수 있다. 추가적으로, 셔틀 벡터는 효모에서의 선택을 위해 선택가능한 마커를 함유한다. 마커는 특별한 아미노산 또는 뉴클레오티드의 합성을 위한 효소를 코딩하는 유전자일 수 있어서, 상응하는 게놈 결실(돌연변이)을 수반하는 세포는 영양요구성 또는 독립영양에 상보성이다. 다르게는, 이들 벡터는 이종성 우성 내성 마커를 함유하고, 이는 G418[제네티신(geneticin)], 하이그로마이신 B 또는 플레오마이신(phleomycin)과 같은 특정 항생제에 대한 내성을 재조합 효모 세포(즉, DNA를 수용하고 마커 유전자를 발현하는 세포)에 제공한다. 또한, 이들 벡터는 외래 DNA를 해당 자리에 클로닝할 수 있게 허용할 (조합된) 제한 자리(복합 클로닝 자리 또는 MCS)의 서열을 함유할 수 있지만, 대체 방법이 또한 존재한다.

[0144] 전통적으로, 4개 유형의 셔틀 벡터는 추가의 유전 요소의 부재 또는 존재에 의해 구별될 수 있다:

[0145] ◎ 제한 효소에 의해 개방되고 선형화된 DNA가 효모 세포의 형질전환을 위해 사용될 경우, 상동성 재조합에 의해 마커의 유전자 자리 또는 또 다른 유전자에서 숙주 게놈내로 통합되는 통합적 플라스미드(YIp). 이는 일반적으로 게놈에서 이러한 특별한 자리에 삽입된 외래 DNA의 하나의 복사물의 존재를 초래한다.

[0146] ◎ 효모 세포에서 자율 복제에 필수적인 2  $\mu$  플라스미드 DNA 서열의 일부를 수반하는 에피솜성 플라스미드

(YE<sub>p</sub>). 형질전환된 플라스미드의 다수의 복사물이 효모 세포에서 증식하고 에피솜으로서 유지된다.

[0147] ◎ 형질전환된 플라스미드가 수 백배 증식하도록 허용하는 복제의 효모 기원[ARS(autonomously replicated sequence), 자율 복제된 서열]을 운반하는 자율 복제 플라스미드(YR<sub>p</sub>).

[0148] ◎ ARS 서열에 추가로, 정상적으로 안정한 유사분열 분리를 보장하고 대개 자가-복제된 플라스미드의 복사물 수를 단지 1개로 감소시키는 중립심성 서열(핵 염색체중 하나로부터 유래됨)을 운반하는 CEN 플라스미드(YC<sub>p</sub>).

[0149] 이들 플라스미드는 형질전환에 의해 효소 세포내로 도입된다. 효모 세포의 형질전환은 몇몇 상이한 기법, 예컨대 아세트산 리튬에 의한 세포의 투과성화[이토(Ito) 등, 1983] 및 전기천공 방법에 의해 달성된다

[0150] 재조합 미생물의 상업적 적용에서, 플라스미드 불안정성은 가장 중요한 문제이다. 불안정성은 형질전환된 세포가 플라스미드에 대한 변화 또는 이의 손실로 인하여 이들의 조작된 특성을 손실하는 경향이다. 이러한 문제는 장(Zhang) 등의 문헌["재조합 사카로마이세스 세레비지아에서의 플라스미드 안정성", Biotechnology Advances, Vol. 14, No. 4, pp. 401 -435, 1996"]에 상세히 논의되어 있다. 통합적 플라스미드에 의해 형질전환된 균주는, 선택적 압력의 부재하에서 조차도 극히 안정하다[쉐르만(Sherman, F.) [http://dbb.urmc.rochester.edu/labs/sherman\\_f/yeast/9.html](http://dbb.urmc.rochester.edu/labs/sherman_f/yeast/9.html) 및 이의 참고문헌].

[0151] 이종성 DNA는 대개 유기체내로 염색체외 플라스미드(YE<sub>p</sub>, YC<sub>p</sub> 및 YR<sub>p</sub>)의 형태로 도입된다. 불행하게도, 특히 선택 압력이 연속적으로 적용되지 않는 경우, 세균 및 효모 둘다에 의해 신규한 특징이 보유되지 않는 것으로 밝혀졌다. 이는 재조합 세포가 장기간 동안 성장할 때 하이브리드 플라스미드의 분리 불안정성에 기인한다. 이는 개체군 이종성 및 클론 다양성을 이끌어 내고, 결국 대부분의 세포가 형질전환에 의해 도입된 특성을 손실하는 세포 개체군을 유도한다. 영양요구성 마커를 갖는 벡터가 사용된다면, 풍부한 배지중에서의 배양은 종종 벡터의 신속한 손실을 유도하는데, 이는 벡터가 단지 최소 배지에서 보유되기 때문이다. 우성 항생제 내성 마커의 대안의 사용은 종종 생산 공정과 비상용성이다. 항생제의 사용은 등록 관점(흔적량의 항생제가 최종 생성물에 존재할 가능성) 또는 경제적 이유(산업용 규모에서의 항생제의 사용 비용)로 바람직하지 않다.

[0152] 벡터의 손실은 대규모 생산 상황에서 문제를 유발한다. DNA의 도입을 위한 대안의 방법이 효모에 대해 존재하고, 예컨대 통합 플라스미드(YIp)의 사용이다. DNA를 재조합에 의해 숙주 계놈내로 통합하여, 높은 안정성을 야기시켰다[카운트(Caunt, P.)의 문헌 "효모에서의 재조합 플라스미드의 안정성(Stability of recombinant plasmids in yeast)", Journal of Biotechnology 9(1988) 173 - 192]. 본 발명자들은 숙주 트랜스포손(transposon)을 사용하는 통합 방법이 양호한 대안임을 발견하였다. 하나의 실시태양에서 유전자는 형질전환된 숙주 세포 계놈내로 통합될 수 있다. 다수의 복사물의 초기 도입(즉, 적용 진화 이전)은 유전자의 도입을 이끄는 당분야에 공지된 임의의 방식으로 실행된다. 하나의 실시태양에서, 이는 숙주 세포의 반복된 서열에 상동성인 부분을 갖는 벡터를 사용하여 달성될 수 있다. 숙주 세포가 효모 세포인 경우, 적합한 반복된 서열은 멜타 서열로서 공지된, Ty 요소의 긴 말단 반복부(LTR: long terminal repeat)이다. Ty 요소는 Ty1 및 Ty2로 지칭되는 2개의 다소 유사한 하위계열로 나뉜다. 이를 요소는 약 6 킬로베이스(kb)의 길이이고, 긴 말단 반복부(LTR)에 의해 경계지워지며, 약 335개 염기 쌍의 서열이다[보에케(Boeke JD) 등의 문헌 "사카로마이세스 세레비지아 계놈은 트랜스포손 Ty1의 작용성 및 비작용성 복사물을 함유한다(The *Saccharomyces cerevisiae* Genome Contains Functional and Nonfunctional Copies of Transposon Ty1). Molecular and Cellular Biology, Apr. 1988, p. 1432-1442 Vol. 8, No. 4"]. 완전히 서열분석된 사카로마이세스 세레비지아 균주 S288c에서, 가장 풍부한 트랜스포손은 Ty1(31개 복사물) 및 Ty2(13개 복사물)이다[가브리엘(Gabriel A), 다프리치(Dapprich J), 쿤켈(Kunkel M), 그레셤(Gresham D), 프라트(Pratt SC) 등의 문헌 "(2006) 트랜스포손 위치의 전반적 지도화(Global mapping of transposon location). PLoS Genet 2(12): e212.doi:10.1371/journal.pgen.0020212"]. 이들 트랜스포손은 2개의 중첩 개방 판독 프레임(ORF: open reading frame)으로 구성되고, 이들 각각은 수 개의 단백질을 코딩한다. 코딩 영역은 전술된 거의 동일한 LTR에 의해 플랭크화된다. 사카로마이세스 세레비지아에서 그 외의, 덜 풍부하고 더 뚜렷한 Ty 요소는 Ty3, Ty4 및 Ty5이다. 전장(full-length) Ty 요소의 각각의 계열에 대하여, 계놈을 통해 분산된 10배 더 많은 솔로 LTR 요소들이 존재한다. 이들은 전장 요소의 LTR-LTR 재조합에 의해 야기되고, 내부 단백질 코딩 영역 밖으로 투평된다.

[0153] Ty 레트로트랜스포손(retrotransposon)의 레트로트랜스포지션(retrotransposition) 기작은 계놈 전체를 통해 다수의 복사물을 통합하기 위해 연구되었다[보에케(Boeke) 등, 1988; 제이콥스(Jacobs) 등, 1988]. 멜타 서열로서 공지된, Ty 요소의 긴 말단 반복부(LTR)는 또한 상동성 재조합에 의한 통합에 대한 우수한 표적인데, 이들이 Ty 회합되거나 솔로 자리인 약 150 내지 200개의 복사물로 존재하기 때문이다[보에케(Boeke), 1989; 킹스맨(Kingsman) 및 킹스맨(Kingsman), 1988; 파레크(Parekh R.N.)의 문헌 "(1996). 사카로마이세스 세레비지아의

분산된 Ty 델타 자리로의 조정가능한, 고 복사의 안정한 통합을 위한 통합 벡터(An Integrating Vector for Tunable, High Copy, Stable Integration into the Dispersed Ty DELTA Sites of *Saccharomyces cerevisiae*). *Biotechnol. Prog.* 1996, 12, 16-21".]. 적응 진화에 의해, 복사물의 수는 변할 수 있다.

[0154] araA, arab 및 araD 유전자

형질전환된 숙주 세포는 아라비노스를 사용할 수 있다. 형질전환된 숙주 세포는 따라서, L-아라비노스를 L-리불로스 및/또는 자일룰로스 5-포스페이트로 및/또는 원하는 발효 생성물, 예를 들면 본원에 언급된 것들로 전환시킬 수 있다.

[0156] 에탄올을 L-아라비노스로부터 생산할 수 있는 유기체, 예를 들면 사카로마이세스 세레비지아 균주는 적합한 공급원으로부터 araA(L-아라비노스 이소머라아제), araB(L-리불로키나아제) 및 araD(L-리불로스-5-P4-에피머라아제) 유전자를 도입하도록 세포를 변형시킴으로써 생산될 수 있다. 형질전환된 숙주 세포내로 도입될 수 있는 이러한 유전자는 아라비노스를 사용할 수 있음을 목적으로 한다. 이러한 접근법은 국제특허공개 제WO2003/095627호에 제공된다. 락토바실러스 플란타룸으로부터의 araA, araB 및 araD 유전자가 사용될 수 있고 국제특허공개 제WO2008/041840호에 개시되어 있다. 바실러스 서브틸리스로부터의 araA 유전자 및 에스캐리키아 콜라이로부터의 araB 및 araD 유전자는 제EP1499708호에 개시되어 있다. 또 다른 실시태양에서, araA, araB 및 araD 유전자는, 국제특허공개 제WO2009/011591호에 개시된 바와 같이, 클라비박터, 아르트로박터 및/또는 그라멜라 속중 1종 이상, 특별히 클라비박터 미키가넨시스, 아르트로박터 아우레센스, 및/또는 그라멜라 포르세티중 하나로부터 유래될 수 있다.

[0157] PPP-유전자

[0158] 형질전환된 숙주 세포는 5탄당 포스페이트 경로(PPP: pentose phosphate pathway)의 플럭스를 증가시키는 하나 이상의 유전자 변형을 포함할 수 있다. 특별히, 유전자 변형(들)은 5탄당 포스페이트 경로의 비산화적 부분을 통해 증가된 플럭스를 유도할 수 있다. 5탄당 포스페이트 경로의 비산화적 부분의 증가된 플럭스를 야기시키는 유전자 변형은, 본원에서 증가된 플럭스를 야기시키는 유전자 변형을 제외하고 유전적으로 동일한 균주에서의 플럭스에 비해 적어도 약 1.1배, 약 1.2배, 약 1.5배, 약 2배, 약 5배, 약 10배 또는 약 20배 플럭스를 증가시키는 변형을 의미하는 것으로 이해된다. 5탄당 포스페이트 경로의 비산화적 부분의 플럭스는 유일한 탄소원으로서의 자일로스 상에서 변형된 숙주를 성장시키고, 특정 자일로스 소비 속도를 결정하고, 만일 자일리톨이 생산된다면 특정 자일로스 생산 속도를 특정 자일로스 소비 속도로부터 차감함으로써 측정될 수 있다. 그러나, 5탄당 포스페이트 경로의 비산화적 부분의 플럭스는 유일한 탄소원으로서의 자일로스 상에서의 성장 속도와 비례하고, 바람직하게는 유일한 탄소원으로서의 자일로스 상에서 혐기성 성장 속도와 비례한다. 유일한 탄소원으로서의 자일로스 상에서의 성장 속도( $\mu_{\text{최대값}}$ ) 및 5탄당 포스페이트 경로의 비산화적 부분의 플럭스 사이에 선형 관계가 존재한다. 특정 자일로스 소비 속도( $Q_s$ )는 하기 식에 따라 비성장 속도(specific growth rate)( $\mu$ ) 및 당분에 대한 생물자원의 수율( $Y_{xs}$ )에 관련된다:

$$Q_s = m_s + \mu / Y_{xs}^{\text{최대값}} \quad \text{또는}$$

$$1/Y_{xs} = m_s / \mu + 1/Y_{xs}^{\text{최대값}}$$

[0161] 따라서,  $\mu$ 이 상수라면, 5탄당 포스페이트 경로의 비산화적 부분의 증가된 플럭스는 수송이 없는 한(선택이 제한됨) 이들 조건하에 최대 성장 속도의 증가로부터 추론될 수 있다.

[0162] 5탄당 포스페이트 경로의 플럭스를 증가시키는 하나 이상의 유전자 변형은 다양한 방식으로 숙주 세포에 도입될 수 있다. 이들은, 예를 들면 자일룰로스 키나아제 및/또는 5탄당 포스페이트 경로의 비산화적 부분의 하나 이상의 효소의 더 높은 정지 상태 활성 수준, 및/또는 비특이적 알도스 환원효소 활성의 감소된 정지 상태 수준을 탈성함을 포함한다. 정지 상태 활성 수준에서의 이러한 변화는 돌연변이체(자발적 또는 화학물질이나 방사선에 의해 유도됨)의 선택에 의해, 및/또는 재조합 DNA 기술, 예를 들어 각각 효소를 코딩하는 유전자 또는 이들 유전자를 조절하는 인자의 과발현 또는 불활성화에 의해 실행될 수 있다.

[0163] 하나의 바람직한 숙주 세포에서, 유전자 변형은 5탄당 포스페이트 경로(비산화적 부분)의 하나 이상의 효소의 과발현을 포함한다. 바람직하게는 효소는 리불로스-5-포스페이트 이소머라아제, 리불로스-5-포스페이트 에피머라아제, 트랜스케톨라아제(transketolase) 및 트랜스알돌라아제(transaldolase)를 코딩하는 효소로 구성된 군에서 선택된다. 5탄당 포스페이트 경로(비산화적 부분)의 효소의 다양한 조합물이 과발현될 수 있다. 예를

들어, 과발현되는 효소는 적어도 효소 리불로스-5-포스페이트 이소머라아제 및 리불로스-5-포스페이트 에피머라아제; 또는 적어도 효소 리불로스-5-포스페이트 이소머라아제 및 트랜스케톨라아제; 또는 적어도 효소 리불로스-5-포스페이트 에피머라아제 및 트랜스케톨라아제; 또는 적어도 효소 리불로스-5-포스페이트 에피머라아제 및 트랜스알돌라아제; 또는 적어도 효소 리불로스-5-포스페이트 에피머라아제, 트랜스케톨라아제 및 트랜스알돌라아제; 또는 적어도 효소 리불로스-5-포스페이트 이소머라아제, 리불로스-5-포스페이트 에피머라아제, 및 트랜스알돌라아제; 또는 적어도 효소 리불로스-5-포스페이트 이소머라아제, 리불로스-5-포스페이트 에피머라아제, 및 트랜스케톨라아제일 수 있다. 본 발명의 한 실시태양에서, 효소 리불로스-5-포스페이트 이소머라아제, 리불로스-5-포스페이트 에피머라아제, 트랜스케톨라아제 및 트랜스알돌라아제는 각각 숙주 세포에서 과발현된다. 유전자 변형이 적어도 효소 트랜스케톨라아제 및 트랜스알돌라아제를 과발현하는 자일로스 전환 숙주 세포가 더 바람직한데, 이러한 숙주 세포가 이미 자일로스 상에서 협기성 성장을 할 수 있기 때문이다. 사실, 몇몇 조건에서, 단지 트랜스케톨라아제 및 트랜스알돌라아제를 과발현하는 자일로스 전환 숙주 세포는, 자일로스 상에서 모든 4개의 효소, 즉 리불로스-5-포스페이트 이소머라아제, 리불로스-5-포스페이트 에피머라아제, 트랜스케톨라아제 및 트랜스알돌라아제를 과발현하는 숙주 세포와 동일한 협기성 성장 속도를 이미 갖는다. 게다가, 효소 리불로스-5-포스페이트 이소머라아제 및 리불로스-5-포스페이트 에피머라아제 둘 다를 과발현하는 숙주 세포는 단지 이소머라아제만 또는 단지 에피머라아제만 과발현하는 숙주 세포에 비해 바람직한데, 이들 효소 중 하나만의 과발현이 대사적 불균형을 만들 수 있기 때문이다.

[0164] 효소 "리불로스 5-포스페이트 에피머라아제"(EC 5.1.3.1)는 D-자일룰로스 5-포스페이트의 D-리불로스 5-포스페이트로의 에피머화 및 이의 반대 과정을 촉매화하는 효소로서 본원에서 규정된다. 이 효소는 또한 포스포리불로스 에피머라아제; 에리트로스(erythrose)-4-포스페이트 이소머라아제; 포스포케토펜토스 3-에피머라아제; 자일룰로스 포스페이트 3-에피머라아제; 포스포케토펜토스 에피머라아제; 리불로스 5-포스페이트 3-에피머라아제; D-리불로스 포스페이트-3-에피머라아제; D-리불로스 5-포스페이트 에피머라아제; D-리불로스-5-P 3-에피머라아제; D-자일룰로스-5-포스페이트 3-에피머라아제; 펜토스-5-포스페이트 3-에피머라아제; 또는 D-리불로스-5-포스페이트 3-에피머라아제로서도 공지되어 있다. 리불로스 5-포스페이트 에피머라아제는 추가로 그의 아미노산 서열에 의해 규정될 수 있다. 유사하게 리불로스 5-포스페이트 에피머라아제는 효소를 코딩하는 뉴클레오티드 서열에 의해, 뿐만 아니라 리불로스 5-포스페이트 에피머라아제를 코딩하는 기준 뉴클레오티드 서열에 하이브리드화하는 뉴클레오티드 서열에 의해 규정될 수 있다. 리불로스 5-포스페이트 에피머라아제를 코딩하는 뉴클레오티드 서열은 본원에서 RPE1로 지정된다.

[0165] 효소 "리불로스 5-포스페이트 이소머라아제"(EC 5.3.1.6)는 본원에서 D-리보스 5-포스페이트의 D-리불로스 5-포스페이트로의 직접적 이성화 및 이의 반대 과정을 촉매화하는 효소로서 본원에서 규정된다. 이 효소는 또한 포스포펜토스이소머라아제; 포스포리보이소머라아제; 리보스 포스페이트 이소머라아제; 5-포스포리보스 이소머라아제; D-리보스 5-포스페이트 이소머라아제; D-리보스-5-포스페이트 케톨-이소머라아제; 또는 D-리보스-5-포스페이트 알도스-케토스-이소머라아제로서도 공지되어 있다. 리불로스 5-포스페이트 이소머라아제는 추가로 그의 아미노산 서열에 의해 규정된다. 유사하게 리불로스 5-포스페이트 이소머라아제는 그 효소를 코딩하는 뉴클레오티드 서열에 의해, 뿐만 아니라 리불로스 5-포스페이트 이소머라아제를 코딩하는 기준 뉴클레오티드 서열에 하이브리드화하는 뉴클레오티드 서열에 의해 규정될 수 있다. 리불로스 5-포스페이트 이소머라아제를 코딩하는 뉴클레오티드 서열은 본원에서 RPI1로 지정된다.

[0166] 효소 "트랜스케톨라아제"(EC 2.2.1.1)는 D-리보스 5-포스페이트 + D-자일룰로스 5-포스페이트 ↔ 세도헵툴로스 (sedoheptulose) 7-포스페이트 + D-글리세르알데히드 3-포스페이트 반응 및 이의 반대 과정을 촉매화하는 효소로서 본원에서 규정된다. 이 효소는 또한 글리코알데히드전달효소 또는 세도헵툴로스-7-포스페이트; D-글리세르알데히드-3-포스페이트 글리코알데히드전달효소로서도 공지되어 있다. 트랜스케톨라아제는 추가로 그의 아미노산에 의해 규정된다. 유사하게 트랜스케톨라아제는 그 효소를 코딩하는 뉴클레오티드 서열에 의해, 뿐만 아니라 트랜스케톨라아제를 코딩하는 기준 뉴클레오티드 서열에 하이브리드화하는 뉴클레오티드 서열에 의해 규정될 수 있다. 트랜스케톨라아제를 코딩하는 뉴클레오티드 서열은 본원에서 TKL1로 지정된다.

[0167] 효소 "트랜스알돌라아제"(EC 2.2.1.2)는 세도헵툴로스 7-포스페이트 + D-글리세르알데히드 3-포스페이트 ↔ D-에리트로스 4-포스페이트 + D-프리토스 6-포스페이트 반응 및 이의 반대 과정을 촉매화하는 효소로서 본원에서 규정된다. 이 효소는 또한 다이하이드록시아세톤 전달효소; 다이하이드록시아세톤 신타아제(synthase); 포름알데히드 트랜스케톨라아제; 또는 세도헵툴로스-7-포스페이트:D-글리세르알데히드-3-포스페이트 글리세론 전달효

소로서도 공지되어 있다. 트랜스알돌라아제는 추가로 그의 아미노산 서열에 의해 규정될 수 있다. 유사하게 트랜스알돌라아제는 그 효소를 코딩하는 뉴클레오티드 서열에 의해, 뿐만 아니라 트랜스알돌라아제를 코딩하는 기준 뉴클레오티드 서열에 하이브리드화하는 뉴클레오티드 서열에 의해 규정될 수 있다. 트랜스알돌라아제를 코딩하는 뉴클레오티드 서열은 본원에서 TAL1로 지정된다.

[0168] 자일로스 이소머라아제 또는 자일로스 환원효소 유전자

본 발명에 따라서, 하나 이상의 자일로스 이소머라아제 유전자 및/또는 하나 이상의 자일로스 환원효소 및 자일리톨 탈수효소의 하나 이상의 복사물은 숙주 세포의 게놈내로 도입된다. 이들 유전 요소의 존재는 이성화 또는 환원에 의해 자일로스를 전환시키는 능력을 세포에 부여한다.

한 실시태양에서, 하나 이상의 자일로스 이소머라아제 유전자의 하나 이상의 복사물은 숙주 세포의 게놈내로 도입된다.

"자일로스 이소머라아제"(EC 5.3.1.5)는 D-자일로스의 D-자일룰로스로의 직접적 이성화 및/또는 이의 반대 과정을 촉매하는 효소로서 본원에서 규정된다. 이 효소는 D-자일로스 케토이소머라아제로서도 공지되어 있다. 자일로스 이소머라아제는 본원에서 또한 D-글루코스 및 D-프럭토스 사이의 전환을 촉매화할 수 있다(따라서, 글루코스 이소머라아제로도 지칭될 수 있음). 자일로스 이소머라아제는 본원에서 2가 양이온, 예컨대 마그네슘, 망간 또는 코발트를 보조인자로서 필요로 할 수 있다.

따라서, 이러한 형질전환된 숙주 세포는 자일로스를 자일룰로스로 이성화할 수 있다. 자일로스를 자일룰로스로 이성화하는 능력은 규정된 자일로스 이소머라아제를 코딩하는 뉴클레오티드 서열을 포함하는 핵산 구조물로 숙주 세포를 형질전환함으로써 숙주 세포 상에 부여된다. 형질전환된 숙주 세포는 자일로스의 자일룰로스로의 직접적 이성화에 의해 자일로스를 자일룰로스로 이성화시킨다.

자일로스 이소머라아제 유전자는 다양한 기원을 가질 수 있고, 예컨대 예를 들면 국제특허공개 제WO2006/009434 호에 개시된 바와 같이 피로마이세스 종이다. 다른 적합한 기원은 박테로이데스(*Bacteroides*), 특별히 제PCT/EP2009/52623호에 기재된 바와 같이 박테로이데스 유니포미스(*Bacteroides uniformis*)이다. 또 다른 실시태양에서, 하나 이상의 자일로스 환원효소 및 자일리톨 탈수효소 유전자의 하나 이상의 복사물이 숙주 세포의 게놈내로 도입된다. 이러한 실시태양에서, 자일로스의 전환은 자일리톨 중간체를 경유하는 자일로스의 자일룰로스로의 2-단계 전환으로, 각각 자일로스 환원효소 및 자일리톨 탈수효소에 의해 촉매화된다. 이의 하나의 실시태양에서, 국제특허공개 제WO2004/085627호에 개시된 바와 같이, 자일로스 환원효소(XR), 자일리톨 탈수효소(XDH), 및 자일룰로키나아제(xylulokinase)(XK)는 과발현될 수 있고, 선택적으로 NADPH 생산 효소를 코딩하는 하나 이상의 유전자는 상향조절되고 NADH 소비 효소를 코딩하는 하나 이상의 유전자는 상향조절된다.

[0174] XKS1 유전자

형질전환된 숙주 세포는 특이적 자일룰로스 키나아제 활성을 증가시키는 하나 이상의 유전자 변형을 포함할 수 있다. 바람직하게는 유전자 변형 또는 변형은, 예를 들어 자일룰로키나아제를 코딩하는 뉴클레오티드 서열의 과발현에 의해 자일룰로키나아제의 과발현을 초래한다. 자일룰로키나아제를 코딩하는 유전자는 숙주 세포에 내생성이거나 숙주 세포에 이종성인 자일룰로키나아제일 수 있다. 숙주 세포에서 자일룰로키나아제의 과발현을 위해 사용되는 뉴클레오티드 서열은 자일룰로키나아제 활성을 갖는 폴리펩티드를 코딩하는 뉴클레오티드 서열이다.

효소 "자일룰로키나아제"(EC 2.7.1.17)는  $ATP + D\text{-자일룰로스} = ADP + D\text{-자일룰로스 5-포스페이트}$  반응을 촉매하는 효소로서 본원에서 규정된다. 이 효소는 또한 포스포릴화 자일룰로키나아제, D-자일룰로키나아제 또는 ATP:D-자일룰로스 5-인산전달효소로서 공지되어 있다. 본 발명에 사용되는 자일룰로키나아제는 추가로 그의 아미노산 서열에 의해 규정된다. 유사하게 자일룰로키나아제는 그 효소를 코딩하는 뉴클레오티드 서열에 의해, 뿐만 아니라 자일룰로키나아제를 코딩하는 기준 뉴클레오티드 서열에 하이브리드화하는 뉴클레오티드 서열에 의해 규정될 수 있다.

형질전환된 숙주 세포에서, 특이적 자일룰로키나아제 활성을 증가시키는 유전자 변형 또는 변형들은 상기 기재된 바와 같은 5탄당 포스페이트 경로의 플럭스를 증가시키는 임의의 변형과 조합될 수 있다. 그러나, 이는 필수적이지 않다.

본 발명의 숙주 세포에서, 과발현되는 자일룰로키나아제는 과발현을 일으키는 유전자 변형(들)을 제외하고 유전적으로 동일한 균주에 비하여, 적어도 약 1.1배, 약 1.2배, 약 1.5배, 약 2배, 약 5배, 약 10배 또는 약 20배

과발현된다. 이러한 과발현 수준은 효소의 활성의 정지 상태 수준, 효소의 단백질의 정지 상태 수준, 뿐만 아니라 효소를 코딩하는 전사체의 정지 상태 수준에 적용될 수 있음을 알아야 한다.

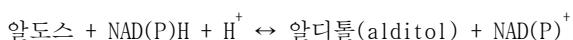
**[0179] 알도스 환원효소(GRE3) 유전자 결실**

XI가 자일로스를 전환시키는 유전자로서 사용되는 실시태양에서, 이는 알도스 환원효소 활성을 감소시키는데 유리할 수 있다. 따라서, 형질전환된 숙주 세포는 숙주 세포에서 알도스 환원효소 활성을 감소시키는 하나 이상의 유전자 변형을 포함할 수 있다. 바람직하게는, 비특이적 알도스 환원효소 활성은 비특이적 알도스 환원효소를 코딩하는 유전자의 발현을 감소시키거나 이를 불활성화시키는 하나 이상의 유전자 변형에 의해 숙주 세포에서 감소된다. 바람직하게는, 유전자 변형(들)은 숙주 세포에서 알도스 환원효소를 코딩하는 유전자의 각각의 내생성 복사물의 발현을 감소시키거나 불활성화시키고, 하나의 실시태양에서, GRE3 알도스 환원효소 결실(본원에서 GRE3 결실로 지칭됨)이다. 형질전환된 숙주 세포는 이배성, 배수성, 또는 이수성(aneuploidy)의 결과로서 비특이적 알도스 환원효소를 코딩하는 유전자의 다수의 복사물을 포함하고/하거나, 숙주 세포는 상이한 유전자에 의해 각각 코딩되고 아미노산 서열이 상이한 알도스 환원효소 활성을 갖는 몇몇 상이한 (이소)효소를 함유할 수 있다. 또한 이러한 경우에 바람직하게는 비특이적 알도스 환원효소를 코딩하는 각각의 유전자의 발현은 감소되거나 불활성화된다. 바람직하게는, 유전자는 유전자의 적어도 일부의 결실 또는 유전자의 파괴에 의해 불활성화되고, 이로써 이와 관련하여 용어 유전자는 또한 코딩 서열의 상류 또는 하류에 임의의 비-코딩 서열을 포함하고, 이의 (부분적) 결실 또는 불활성화는 숙주 세포에서 비특이적 알도스 환원효소 활성의 발현을 감소시킨다.

숙주 세포에서 활성이 감소되어질 알도스 환원효소를 코딩하는 뉴클레오티드 서열은 알도스 환원효소 활성을 갖는 폴리펩티드를 코딩하는 뉴클레오티드 서열이다.

그러므로, 비특이적 알도스 환원효소 활성을 숙주 세포에서 감소시키는 유전자 변형 또는 변형들만을 포함하는 숙주 세포가 특별히 본 발명에 포함된다.

효소 "알도스 환원효소"(EC 1.1.1.21)는 자일로스 또는 자일룰로스를 자일리톨로 환원시킬 수 있는 임의의 효소로서 본원에서 규정된다. 본 발명과 관련하여, 알도스 환원효소는 본 발명의 숙주 세포에 고유적(내생성)이고 자일로스 또는 자일룰로스를 자일리톨로 환원시킬 수 있는 임의의 비특이적 알도스 환원효소일 수 있다. 비특이적 알도스 환원효소는 하기의 반응을 촉매화한다:



효소는 광범위한 특이성을 갖고 알도스 환원효소; 폴리올 탈수소효소(NADP<sup>+</sup>); 알디톨:NADP 옥시도환원효소; 알디톨:NADP<sup>+</sup> 1-옥시도환원효소; NADPH-알도펜토스 환원효소; 또는 NADPH-알도스 환원효소로서도 공지되어 있다.

이의 특별한 예는 사카로마이세스 세레비지아에 내생성이고 GRE3 유전자에 의해 코딩되는 비특이적 알도스 환원효소이다[트래프(Traff) 등의 문헌 "2001, Appl. Environ. Microbiol. 67: 5668-74"]. 그러므로, 본 발명의 알도스 환원효소는 추가로 그의 아미노산 서열에 의해 규정된다. 유사하게 알도스 환원효소는 그 효소를 코딩하는 뉴클레오티드 서열에 의해, 뿐만 아니라 알도스 환원효소를 코딩하는 기준 뉴클레오티드 서열에 하이브리드화하는 뉴클레오티드 서열에 의해 규정될 수 있다.

**바이오제품 생산**

작물 당분으로부터 바이오-에탄올을 생산하기 위한 다양한 유기체의 도입에 대해 수년에 걸쳐 제안되어 왔다. 그러나, 실제로, 모든 주요 바이오-에탄올 생산 공정은 계속해서 사카로마이세스 속의 효모를 에탄올 생산자로서 사용하여 왔다. 이는 산업용 공정을 위한 사카로마이세스 종의 많은 유리한 특징, 즉 높은 내산성, 내에탄올성 및 내삼투성, 혐기성 성장 능력, 및 물론 이의 높은 알코올 발효 능력에 기인한다. 바람직한 효모 종으로는 사카로마이세스 세레비지아, 사카로마이세스 불데리(*S. bulderi*), 사카로마이세스 바르네티(*S. barnetti*), 사카로마이세스 엑시구어스(*S. exiguum*), 사카로마이세스 우바람(*S. uvarum*), 사카로마이세스 디아스타티커스(*S. diastaticus*), 클루이베로마이세스 락티스, 클루이베로마이세스 마륵시아누스(*K. marxianus*) 또는 클루이베로마이세스 프라길리스(*K. fragilis*)가 포함된다.

형질전환된 숙주 세포는 에탄올의 생산에 적합한 세포일 수 있다. 그러나, 형질전환된 숙주 세포는 에탄올 이외의 발효 생성물의 생산에 적합할 수 있다.

이러한 비-에탄올성 발효 생성물은 원칙적으로 진핵 미생물, 예컨대 효모 또는 사상균류에 의해 생산가능한 임

의의 별크형 또는 미세 화학물질을 포함한다.

[0191] 비-에탄올성 발효 생성물의 생산을 위해 사용될 수 있는 형질전환된 숙주 세포는 감소된 알코올 탈수소효소 활성을 초래하는 유전자 변형을 함유하는 숙주 세포이다.

[0192] 하나의 실시태양에서 형질전환된 숙주 세포는 리그노셀룰로스로부터 기원된 당분이 에탄올로 전환되는 공정에서 사용될 수 있다.

### 리그노셀룰로스

[0194] 잠재적 재생성 공급원료로서 고려될 수 있는 리그노셀룰로스는, 일반적으로 다당류 셀룰로스(글루칸) 및 헤미셀룰로스(자일란, 헤테로자일란 및 자일로글루칸)를 포함한다. 또한, 몇몇 헤미셀룰로스는 글루코만나난으로서, 예를 들면 목재-유래된 공급원료에 존재할 수 있다. 이들 다당류의 단량체 및 다량체 둘다를 비롯한 가용성 당분, 예를 들면 글루코스, 셀로비오스, 자일로스, 아라비노스, 갈락토스, 프리토스, 만노스, 람노스, 리보스, 갈락투론산, 글루쿠론산 및 다른 6탄당 및 5탄당으로의 효소적 가수분해는 일제히 작용하는 상이한 효소들의 작용 하에 초래된다.

[0195] 또한, 페틴 및 다른 페틴 물질, 예컨대 아라비난은 전형적으로 비-목재 식물 조직의 세포 벽의 건조 질량의 상당한 비율을 구성할 수 있다(건조 질량의 약 1/4 내지 절반이 페틴일 수 있다).

### 예비처리

[0197] 효소 처리 이전에, 리그노셀룰로스성 물질은 예비처리될 수 있다. 예비처리는 리그노셀룰로스성 물질을 산, 염기, 용매, 열, 과산화물, 오존, 기계적 슈레딩(shredding), 연마, 제분 또는 신속한 감압(depressurization), 또는 이의 임의의 2종 이상의 조합에 노출시킴을 포함한다. 이러한 화학적 예비처리는 종종, 예를 들어 150 내지 220°C에서 1 내지 30 분 동안의 열-예비처리와 조합된다.

### 효소적 가수분해

[0199] 예비처리된 물질은 통상적으로 효과적으로 가수분해되어 본 발명에 따라 발효될 수 있는 당분을 방출할 수 있다. 이는 종래의 방법, 예를 들어 셀룰라아제(cellulase), 예를 들어 셀로비오하이드롤라아제(celllobiohydrolase)(들), 엔도글루카나아제(endoglucanase)(들), 베타-글루코시다아제(beta-glucosidase)(들) 및 선택적으로 그 외의 효소와 접촉시킴으로써 실행될 수 있다. 셀룰라아제에 의한 전환은 주변 온도 또는 더 높은 온도에서 당분(들)의 충분한 양이 방출되는 반응 시간에서 실행될 수 있다. 효소적 가수분해의 결과는, 본원에서 당분 조성으로서 지정된 C5/C6 당분을 포함하는 가수분해 생성물이다.

### 당분 조성

[0201] 본 발명에 따른 당분 조성은 글루코스, 아라비노스 및 자일로스를 포함한다. 임의의 당분 조성은 이들 범주를 충족시키는 본 발명에서 사용될 수 있다. 당분 조성에서 선택적 당분은 갈락토스 및 만노스이다. 한 바람직한 실시태양에서, 당분 조성은 하나 이상의 리그노셀룰로스성 물질의 가수분해물이다. 리그노셀룰로스는 본원에서 생물자원의 헤미셀룰로스 부분 및 헤미셀룰로스를 포함한다. 또한 리그노셀룰로스는 생물자원의 리그노셀룰로스성 단편을 포함한다. 적합한 리그노셀룰로스성 물질은 하기 목록에서 발견될 수 있다: 과일 당액, 수풀, 제분 폐기물(mill waste), 도심 폐재(urban wood waste), 도시 폐기물(municipal waste), 벌목 폐기물, 산림 간벌, 단기-회전 목재 작물, 산업용 폐기물, 밀집, 귀리짚, 벼짚, 보리짚, 호밀짚, 아마대, 콩 깨질, 왕겨, 벼짚, 옥수수 글루텐 사료, 귀리 깨질, 사탕수수, 옥수수 대, 옥수수 줄기, 옥수수 심, 옥수수 깨질, 수수속 풀, 얹새, 단수수, 카놀라 줄기, 대두 줄기, 대초원 목초, 포아풀과 식물, 강아지풀; 사탕무 펄프, 시트러스 과일 펄프, 씨 깨질, 셀룰로스성 동물 폐기물, 잔디 깨은 잔해, 목화, 해조류, 나무, 연질 목재, 경질 목재, 포플리, 솔, 관목, 목초, 밀, 밀짚, 사탕수수 바가스(bagasse), 옥수수, 옥수수 깨질, 옥수수 흡(hob), 옥수수 알, 알맹이로부터의 섬유, 곡물의 습식 또는 건식 제분으로부터의 생성물 및 부산물, 도시 고형 폐기물, 폐지, 정원 쓰레기, 초본(herbaceous) 물질, 농사 잔여물, 삼림관리 잔여물, 도시 고형 폐기물, 폐지, 펄프, 제지 잔여물, 가지, 덤불, 줄기, 옥수수, 옥수수 깨질, 에너지 작물, 쟁, 과일, 꽃, 곡물, 목초, 초본 작물, 잎, 나무껍질, 솔잎, 통나무, 뿌리, 묘목, 관목, 수수속 풀, 나무, 야채, 과일 깨질, 덩쿨, 사탕무 펄프, 밀 미들링(midling), 귀리 깨질, 경질 또는 연질 목재, 농사 과정으로부터 생성된 유기 폐기물질, 삼림 목재 폐기물, 또는 이의 임의의 둘 이상의 조합물.

[0202] 리그노셀룰로스로부터 유래된 몇몇 적합한 당분 조성 및 이들의 가수분해물의 당분 조성에 대한 개요가 표 2에 제공된다. 열거된 리그노셀룰로스는 옥수수 심, 옥수수 섬유, 왕겨, 멜론 깨질, 사탕무 펄프, 밀짚, 사탕수수

바가스, 목재, 목초 및 올리브 프레싱(olive pressing)이 포함된다.

## 표 2

리그노셀룰로스성 물질로부터의 당분 조성에 대한 개요. Gal=갈락토스,

Xyl=자일로스, Ara=아라비노스, Man=만노스, Glu=글루코스, Rham=람노스. 갈락토

스 백분율(% Gal)이 제공된다.

리그노셀룰로스성 물질	Gal	Xyl	Ara	Man	Glu	Rham	합계	%. Gal.
옥수수 심 a	10	286	36		227	11	570	1,7
옥수수 심 b	131	228	160		144		663	19,8
왕겨 a	9	122	24	18	234	10	417	2,2
왕겨 b	8	120	28		209	12	378	2,2
멜론 껌질	6	120	11		208	16	361	1,7
사탕무 펄프	51	17	209	11	211	24	523	9,8
밀짚 아이다호 (Idaho)	15	249	36		396		696	2,2
옥수수 썸유	36	176	113		372		697	5,2
사탕수수 바가스	14	180	24	5	391		614	2,3
옥수수 대	19	209	29		370		626	
아텔 (목재)	5	118	7	3	493		625	0,7
유칼립투스 (목재)	22	105	8	3	445		583	3,8
CWR (목초)	8	165	33		340		546	1,4
JTW (목초)	7	169	28		311		515	1,3
MSW	4	24	5	20	440		493	0,9
리아드 카나리 그래스 (Read Canary Grass) 야채	16	117	30	6	209	1	379	4,2
리아드 카나리 그래스 씨	13	163	28	6	265	1	476	2,7
올리브 프레싱 잔여물	15	111	24	8	329		487	3,1

[0203]

표 2로부터, 이를 리그노셀룰로스에 다양한 당분이 글루코스, 자일로스, 아라비노스 및 갈락토스의 형태로 존재하는 것이 분명하다. 글루코스, 자일로스, 아라비노스 및 갈락토스의 발효 생성물로의 전환은 그러므로 경제적으로 매우 중요하다. 또한 만노스는 몇몇 리그노셀룰로스 물질중에 이전에 언급된 당분에 비해 대개 더 적은 양으로 존재한다. 유리하게는 따라서 만노스 또한 형질전환된 숙주 세포에 의해 전환된다.

[0204] 발효

발효 공정은 호기성 또는 혐기성 발효 공정일 수 있다. 혐기성 발효 공정은 본원에서 산소의 부재하에 운행되거나 실질적으로 산소가 소비되지 않고, 바람직하게는 약 5, 약 2.5 또는 약 1 밀리몰/ℓ/시간 미만, 더 바람직하게는 0 밀리몰/ℓ/시간이 소비되고(즉, 산소 소비는 겹출될 수 없음) 유기 분자가 전자 공여체(donor) 및 전자 수용체(acceptor) 둘 다로서 작용하는 발효 공정으로서 규정된다. 산소의 부재하에, 당분해작용 및 생물자원 형성시에 생산되는 NADH는 산화적 인산화에 의해 산화될 수 없다. 이러한 문제를 해결하기 위해, 많은 미생물은 피루베이트 또는 그의 유도체중 하나를 전자 및 수소 수용체로서 사용하여 NAD<sup>+</sup>를 재생한다.

[0205] 그러므로, 바람직한 혐기성 발효에서 5탄당은 발효 생성물, 예컨대 에탄올, 부탄올, 락트산, 3-하이드록시-프로

피온산, 아크릴산, 아세트산, 석신산, 시트르산, 밀산, 푸마르산, 아미노산, 1,3-프로판-디올, 에틸렌, 글리세롤,  $\beta$ -락탐 항생물질 또는 세팔로스포린으로 전환된다.

[0208] 발효 공정은 바람직하게는 세포를 위해 최적인 온도에서 운행된다. 그러므로, 대부분의 효모 또는 균류 숙주 세포의 경우, 발효 공정은 약 42°C 미만, 바람직하게는 약 38°C 미만의 온도에서 수행된다. 효모 또는 사상균류 숙주 세포의 경우, 발효 공정은 바람직하게는 약 35°C, 약 33°C, 약 30°C 또는 약 28°C 보다 낮은 온도에서, 및 약 20°C, 약 22°C, 또는 약 25°C 보다 높은 온도에서 수행된다.

[0209] 이 공정에서 자일로스 및/또는 글루코스에 대한 에탄올 수율은 바람직하게는 적어도 약 50, 약 60, 약 70, 약 80, 약 90, 약 95 또는 약 98%이다. 에탄올 수율은 본원에서 이론적 최대 수율의 백분율로서 규정된다.

[0210] 본 발명은 발효 생성물의 제조 방법에 관한 것이다.

[0211] 본 발명에 따른 발효 공정은 호기성 및 혐기성 조건하에 운행될 수 있다. 하나의 실시태양에서, 이 공정은 미호기성(micro-aerophilic) 또는 산소 제한된 조건하에 실행된다.

[0212] 혐기성 발효 공정은 산소의 부재하에 운행되거나 실질적으로 산소가 소비되지 않는, 바람직하게는 약 5, 약 2.5 또는 약 1 밀리몰/ℓ/시간 미만이고, 유기 분자가 전자 공여체 및 전자 수용체로서 작용하는 발효 공정으로서 규정된다.

[0213] 한 바람직한 공정에서, 세포는 자일로스 및 글루코스 둘 다를 발효시키고, 글루코스 억제에 비감수성인 세포가 사용되는 경우 동시적인 것이 바람직하다. 탄소원으로서의 자일로스(및 글루코스)의 공급원에 더하여, 발효 배지는 추가로 세포의 성장을 위해 필요한 적절한 구성성분을 포함할 것이다. 미생물, 예컨대 효모의 성장을 위한 발효 배지의 조성은 당분야에 잘 공지되어 있다.

[0214] 발효 공정은 배치식, 공급-배치식(fed-batch) 또는 연속 방식으로 실행될 수 있다. 별도의 가수분해 및 발효 (SHF: separate hydrolysis and fermentation) 공정 또는 동시적 사카린화 및 발효(SSF: simultaneous saccharification and fermentation) 공정이 또한 적용될 수 있다. 이들 발효 공정 방식의 조합이 또한 최적 생산성을 위해 가능할 수 있다. 이들 공정은 이후 더욱 상세히 기재된다.

#### SSF 방식

[0215] 동시적 사카린화 및 발효(SSF) 방식을 위해, 액화/가수분해 또는 예비사카린화 단계를 위한 반응 시간은, 원하는 수율, 즉 셀룰로스의 글루코스로의 전환 수율을 실현하기 위한 시간에 좌우된다. 이러한 수율은 바람직하게는 가능한 한 높고, 바람직하게는 60% 이상, 65% 이상, 70% 이상, 75% 이상, 80% 이상, 85% 이상, 90% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상, 99% 이상, 더욱이 99.5% 이상 또는 99.9% 이상이다.

[0216] 본 발명에 따라서 SHF 방식에서 매우 높은 당분 농도 및 SSF 방식에서 매우 높은 생성물 농도(예를 들어, 에탄올)가 실현된다. SHF 조작시, 글루코스 농도는 25 g/ℓ 이상, 30 g/ℓ 이상, 35 g/ℓ 이상, 40 g/ℓ 이상, 45 g/ℓ 이상, 50 g/ℓ 이상, 55 g/ℓ 이상, 60 g/ℓ 이상, 65 g/ℓ 이상, 70 g/ℓ 이상, 75 g/ℓ 이상, 80 g/ℓ 이상, 85 g/ℓ 이상, 90 g/ℓ 이상, 95 g/ℓ 이상, 100 g/ℓ 이상, 110 g/ℓ 이상, 120 g/ℓ 이상이거나, 예를 들어 25 g/ℓ 내지 250 g/ℓ, 30 g/ℓ 내지 200 g/ℓ, 40 g/ℓ 내지 200 g/ℓ, 50 g/ℓ 내지 200 g/ℓ, 60 g/ℓ 내지 200 g/ℓ, 70 g/ℓ 내지 200 g/ℓ, 80 g/ℓ 내지 200 g/ℓ, 90 g/ℓ, 80 g/ℓ 내지 200 g/ℓ 일 수 있다.

#### SSF 방식에서의 생성물 농도

[0217] SSF 조작시, 생성물 농도(g/ℓ)는 생산된 글루코스의 양에 의존하지만, 이는 시각적인 것은 아닌데, 당분이 SSF에서 생성물로 전환되기 때문이고, 생성물 농도는 이론적 최대 수율을 곱함으로써 기본 글루코스 농도에 관련될 수 있다(글루코스 1 g당 생성물(g)중 Yps 최대값).

[0218] 발효 생성물의 이론적 최대 수율(글루코스 1 g당 생성물(g)중 Yps 최대값)은 생화학 교재로부터 유도될 수 있다. 에탄올의 경우, 효모에서 정상적인 당분해작용 발효 경로에 따라, 1 몰의 글루코스(180 g)는 2 몰의 에탄올을 수득한다( $= 2 \times 46 = 92$  g 에탄올). 글루코스에 대한 에탄올의 이론적 최대 수율은 따라서  $92/180 = 0.51$  g 에탄올/g 글루코스이다.

[0219] 부탄올(MW 74 g/몰) 또는 이소부탄올의 경우, 이론적 최대 수율은 글루코스 1 몰당 부탄올 1 몰이다. 따라서 (이소-)부탄올에 대한 Yps 최대값 =  $74/180 = 0.41$  g (이소-)부탄올/g 글루코스이다.

[0222] 젖산의 경우, 동종젖산 발효에 대한 발효 수율은 글루코스 1 몰당 젖산 2몰(MW = 90 g/몰)이다. 화학양론적 양에 따라서, Yps 최대값 = 1 g 젖산/g 글루코스이다.

[0223] 다른 발효 생성물의 경우, 유사한 계산이 이루어질 수 있다.

#### SSF 방식

[0225] SSF 조작시, 생성물 농도는  $25g * Yps g/\ell / \ell$  이상,  $30 * Yps g/\ell / \ell$  이상,  $35g * Yps g/\ell / \ell$  이상,  $40 * Yps g/\ell / \ell$  이상,  $45 * Yps g/\ell / \ell$  이상,  $50 * Yps g/\ell / \ell$  이상,  $55 * Yps g/\ell / \ell$  이상,  $60 * Yps g/\ell / \ell$  이상,  $65 * Yps g/\ell / \ell$  이상,  $70 * Yps g/\ell / \ell$  이상,  $75 * Yps g/\ell / \ell$  이상,  $80 * Yps g/\ell / \ell$  이상,  $85 * Yps g/\ell / \ell$  이상,  $90 * Yps g/\ell / \ell$  이상,  $95 * Yps g/\ell / \ell$  이상,  $100 * Yps g/\ell / \ell$  이상,  $110 * Yps g/\ell / \ell$  이상,  $120 g/\ell * Yps$  이상이거나, 예를 들어  $25 * Yps g/\ell / \ell$  내지  $250 * Yps g/\ell / \ell$ ,  $30 * Yps g/\ell / \ell$  내지  $200 * Yps g/\ell / \ell$ ,  $40 * Yps g/\ell / \ell$  내지  $200 * Yps g/\ell / \ell$ ,  $50 * Yps g/\ell / \ell$  내지  $200 * Yps g/\ell / \ell$ ,  $60 * Yps g/\ell / \ell$  내지  $200 * Yps g/\ell / \ell$ ,  $70 * Yps g/\ell / \ell$  내지  $200 * Yps g/\ell / \ell$ ,  $80 * Yps g/\ell / \ell$  내지  $200 * Yps g/\ell / \ell$ ,  $90 * Yps g/\ell / \ell$ ,  $80 * Yps g/\ell / \ell$  내지  $200 * Yps g/\ell / \ell$  일 수 있다.

[0226] 따라서, 본 발명은 발효 생성물의 제조 방법을 제공하고, 이러한 방법은:

[0227] a. 본원에 기재된 바와 같은 방법을 사용하여 리그노셀룰로스를 분해하는 단계; 및

[0228] b. 생성된 물질을 발효시키는 단계를 포함하고, 이로써 발효 생성물을 제조한다.

#### 발효 생성물

[0230] 본 발명의 발효 생성물은 임의의 유용한 생성물일 수 있다. 한 실시태양에서, 이는 에탄올, n-부탄올, 이소부탄올, 락트산, 3-하이드록시-프로피온산, 아크릴산, 아세트산, 석신산, 푸마르산, 말산, 이타콘산, 말레산, 시트르산, 아디프산, 아미노산, 예컨대 라이신, 메티오닌, 트립토판, 트레오닌, 및 아스파르트산, 1,3-프로판-디올, 에틸렌, 글리세롤,  $\beta$ -락탐 항생물질 또는 세팔로스포린, 비타민, 약제, 동물 사료 보충물, 특수 화학물질, 화학 공급원료, 플라스틱, 용매, 연료, 예컨대 바이오연료 또는 바이오가스 또는 유기 중합체, 및 산업용 효소, 예컨대 단백질분해효소, 셀룰라아제, 아밀라아제, 글루카나아제, 락타아제, 리파아제, 리아제, 옥시도환원효소, 전달효소 또는 자일라나아제로 구성된 군에서 선택된 생성물이다. 예를 들면, 발효 생성물은 선행 기술의 세포 제조 방법 및 발효 공정을 수행하여 본 발명에 따른 세포에 의해 생성될 수 있지만, 이러한 예는 본원에서 제한되는 것으로 해석되어서는 안된다. n-부탄올은 국제특허공개 제WO2008/121701호 또는 제WO2008/086124호에 기재된 바와 같은 세포에 의해 생산될 수 있고; 젖산은 제US2011/053231호 또는 제US2010/137551호에 기재된 바와 같은 세포에 의해 생산될 수 있고; 3-하이드록시-프로피온산은 국제특허공개 제WO2010/010291호에 기재된 바와 같은 세포에 의해 생산될 수 있고; 아크릴산은 국제특허공개 제WO2009/153047호에 기재된 바와 같은 세포에 의해 생산될 수 있다.

#### 발효 생성물의 회수

[0232] 발효 생성물을 회수하기 위해, 현준하는 기술이 사용된다. 상이한 발효 생성물에 대해, 상이한 회수 공정이 적절하다. 에탄올을 수성 혼합물로부터 회수하는 현준하는 방법은 흔히 분별화 및 흡착 기법을 사용한다. 예를 들면, 맥주는 여전히 발효된 생성물(이는 수성 혼합물에 에탄올을 함유함)을 가공하기 위해 사용되어 풍부한 에탄올-함유 혼합물을 생산하고, 이는 이어서 분별화된다(예를 들어, 분별 증류 또는 다른 유사 기법). 그런 다음, 가장 높은 농도의 에탄올을 포함하는 분별물은 흡착제를 통해 통과되어, 전부는 아니지만, 대부분의 잔여수를 에탄올로부터 제거한다.

[0233] 하기 실시예는 본 발명을 예시한다:

#### 실시예

##### 일반적 분자 생물학 기법

[0236] 달리 지시되지 않지만, 사용되는 방법은 표준 생화학 기법이다. 적합한 일반적 방법론 교재의 예로는 샘브록(Sambrook) 등의 문헌[Molecular Cloning, a Laboratory Manual (1989)] 및 어수벨(Ausubel) 등의 문헌[Current Protocols in Molecular Biology (1995), John Wiley & Sons, Inc.]이 포함된다.

#### 배지

[0238] 실험에 사용된 배지는 실시예에 지시된 바와 같이 당분이 보충된 YEP-배지(10 g/ℓ의 효모 추출물, 20 g/ℓ의

펩톤) 또는 고체 YNB-배지(6.7 g/ℓ의 효모 질소 베이스, 15 g/ℓ의 아가)였다. 고체 YEP 배지의 경우, 15 g/ℓ의 아가를 멸균 전에 액체 배지에 첨가하였다.

[0239] AFM 실험에서, 광물 배지를 사용하였다. 광물 배지의 조성은 베르듀인(Verduyn) 등의 문헌[Yeast (1992), Volume 8, 501-517]에 기재되어 있고, 실시예에 지시된 바와 같이 2.325 g/ℓ의 우레아 및 당분으로 보충되었다.

#### 효모 세포의 형질전환

[0241] 효모 형질전환을 쉬에스틀(Schiestl) 및 기에츠(Gietz)의 문헌[Current Genetics (1989), Volume 16, 339-346]에 기재된 방법에 따라 수행하였다.

#### 콜로니 PCR

[0243] 루케(Looke) 등의 문헌[BioTechniques (2011), Volume 50, 325-328]에 의해 기재된 방법에 따라 PCR을 위해 단일 효모 콜로니로부터 계놈성 DNA를 추출하였다.

#### AFM 절차

[0245] 알코올 발효 모니터[AFM: Alcohol Fermentation Monitor, 할로텍(Halotec), 네델란드 비넨탈]는 6가지 동시적 협기성 발효에 대한 탄소 전환 속도 및 수율의 정확한 비교를 허용하는 튼튼하고 사용자 친화적인 실험실 병행 생물반응기이다.

[0246] AFM 실험의 출발 배양액은 50 mg의 효모(건조 중량)를 함유하였다. 이를 결정하기 위해, 보정 곡선을 OD<sub>700</sub>에 대한 생물자원의 RN1001 균주로 제조하였다. 이러한 보정 곡선을 실험에 사용하여 50 mg의 효모(건조 중량)에 필요한 세포 배양액의 부피를 결정하였다.

[0247] AFM 실험을 시작하기 이전에, 예비배양액을 실시예에 지시된 바와 같이 성장시켰다. 각각의 균주에 대해, OD<sub>700</sub>을 측정하였고, 50 mg의 효모(건조 중량)를 실시예에 지시된 바와 같이 2.325 g/ℓ의 우레아 및 당분이 보충된 400 mL의 광물 배지[베르듀인(Verduyn) 등의 문헌 "Yeast (1992), Volume 8, 501-517"]에 접종하였다.

#### RNA 단리

[0249] 약간 변경된 제조업체의 프로토콜에 의해 뉴클레오스핀(Nucleospin) RNA II 키트[마쉐리-나겔(Machery-Nagel), 게엠베하 운트 코 카게(GmbH & Co. KG), 독일 뒤렌]를 사용하여 RNA를 단리하였다. 프로토콜의 시작 이전에, 신선하게 성장한 효모 세포를 0.2 mg의 라이티카제(lyticase)/mL, 1 M 소르비톨 및 0.1 M EDTA 완충액 중에서 30 분 동안 30°C에서 처리하였다. 이러한 단계 이후에, 단계 6 및 7(실리카 막의 탈염 및 DNA 분해효소 처리)을 제외하고 제조업체 프로토콜을 실행하였다. 계놈성 DNA를 제거하기 위해, DNA 분해효소 처리를 칼럼으로부터 RNA가 용리된 후 실행하였다. 이를 위해 10 μL의 RNA 용액을 7 μL의 DNA 분해효소 및 RNA 분해효소 제거수, 2 μL의 10× 완충액 및 1 μL의 DNA 분해효소[페르멘타스(Fermentas), 독일 68789 세인트 레온-로트]와 혼합하였다. 이러한 혼합물을 30 분 동안 37°C에서 항온처리하였다. 효소의 불활성화를 위해, 1 μL의 25mM EDTA(페르멘타스, 독일 68789 세인트 레온-로트)를 첨가하고 10 분 동안 65°C에서 항온처리하였다.

#### cDNA 합성

[0251] 레버트에이드(RevertAid: 등록상표) H 마이너스 퍼스트 스트랜드(Minus First Strand) cDNA 합성 키트(페르멘타스, 독일 68789 세인트 레온-로트)를 사용하여 cDNA 합성을 RNA 상에서 수행하였다.

#### Q-PCR

[0253] Q-PCR 실험을 위해, 다이나모 컬러플래쉬(DyNAmo Colorflash) SYBR 그린(Green) qPCR 키트[핀자이메스(Finnzymes), 핀란드 01620 반타]를 사용하였다. DNA 혼합물을 1 μL의 cDNA, 39 μL의 H<sub>2</sub>O 및 10 μL의 황색 샘플 염료로 제조하였다. 각각의 반응에 대하여, 5 μL의 DNA 혼합물을 10 μL의 2× 다이나모 마스터믹스(mastermix), 7 μL의 H<sub>2</sub>O, 1 μL의 10 μM 순방향 프라이머 및 1 μL의 10 μM 역방향 프라이머(프라이머 수는 실시예에 지시됨)와 함께 사용하였다. 표 3에 지시된 바와 같은 프로그램을 사용하여 바이오-래드(Bio-Rad) CFX96 실시간 시스템(바이오-래드, 미국 캘리포니아주 헤르큘레스) 상에서 Q-PCR을 운행하였다.

표 3

## Q-PCR 프로그램

프로그램
5:00 동안 95.0°C
0:20 동안 95.0°C
0:20 동안 67.0°C
0:20 동안 72.0°C
+ 플레이트 판독
GOTO 2, 49회 이상
용융 곡선 65.0 내지 95.0°C, 증분 0.5°C, 0:05 + 플레이트 판독
종료

[0254]

## 균주

[0255] 실험에서 사용된 균주는 RN1001, RN1041 및 RN1216이다. RN1041은 국제특허공개 제WO2012/067510호에 기재되어 있다. 이러한 균주는 하기 유전자형을 갖는다: MAT a, ura3-52, leu2-112, his3::loxP, gre3::loxP, loxP-pTPI1::TAL1, loxP-pTPI1::RKI1, loxP-pTPI1-TKL1, loxP-pTPI1-RPE1, 텔타::pADH1-XKS1-tCYC1-LEU2, 텔타::URA3-pTPI1-xy1A-tCYC1

[0256] MAT a = 교배형 a

[0257] ura3-52, leu2-112, HIS3::loxP 돌연변이는 각각 URA3, LEU2 및 HIS3 유전자에서 발생된다. ura3-52 돌연변이는 xy1A 과발현 구조를 상의 URA3 유전자에 의해 보완되고; leu2-112 돌연변이는 XKS1 과발현 구조를 상에서 LEU2 유전자에 의해 보완된다. HIS3-유전자의 결실은 히스티딘 영양요구성을 초래한다. 이러한 이유로, RN1041은 성장을 위해 배지중에 히스티딘을 필요로 한다.

[0258] gre3::loxP는 알도스 환원효소를 코딩하는 GRE3 유전자의 결실이다. loxP 자리는 마커 제거 후 계놈중에 뒤에 남는다.

[0259] loxP-pTPI1은 본원에 기재된 실험에서 TPI1 유전자의 프로모터가 고유 프로모터를 대체함에 의한 비산화적 5탄당 포스페이트 경로의 유전자의 과발현을 지정한다. 강한 구성적 TPI1 프로모터 상류의 loxP 자리는 마커 제거 후 계놈에 남는다[쿠이퍼(Kuyper) 등의 문헌 "FEMS Yeast Research 5 (2005) 925-934"].

[0260] 텔타::은 Ty1 레트로트랜스포손의 긴 말단 반복부 상에서의 재조합 이후의 구조물의 염색체 통합을 의미한다.

[0261] 균주 RN1001은 균주 RN1041의 모 균주, 즉 HIS3-유전자의 결실 이전의 균주이다.

[0262] 균주 RN1216은 균주 RN1041과 동일한 유전자형을 갖는다.

실시예 1글리وك살라아제 과발현 플라스미드의 작성

[0263] [0264] 프라이머 서열 번호 25 및 서열 번호 26을 사용하여 GL01 ORF를 CEN.PK113-7D로부터 수득하였다. 후속적으로, 제로 블런트(Zero Blunt: 등록상표) TOPO(등록상표) PCR 클로닝 키트[인비트로젠(Invitrogen), 미국 캘리포니아주 칼스바드]를 사용하여 PCR 증폭된 GL01 ORF를 TOPO 블런트(blunt) 벡터에서 클로닝하고, 원 샷(One Shot: 등록상표) TOP10 화학적 컴페턴트(competent) 에스케리키아 콜라이(인비트로젠, 미국 캘리포니아주 칼스바드)로 형질전환시켰다. 미니프렙(Miniprep) DNA 단리를 몇몇 콜로니 상에서 수행하였다. 수득된 플라스미드를 제한효소 분석에 의해 검토하였다. 정확한 플라스미드를 pRN935로 지칭하였다.

[0265] GL01의 과발현을 2  $\mu$  의 플라스미드 상에서 실행하였다. 따라서 GL01 발현 카세트를 갖는 2  $\mu$  의 플라스미드를 작성하였다. PGK1 프로모터 및 PGI1 종결자를 갖는 GL01 발현 카세트의 작성은 도 1에 제시된다. 플라스미드 pRN228, pRN935 및 pRN685로 제시된, 이들 3개 요소를 함유하는 플라스미드의 물리 지도는 도 2, 3 및 4에 각각 제공된다. 후속적으로 작성된 GL01 발현 카세트를 프라이머 서열 번호 1(pPGK1-5'-F; 도 1) 및 서열 번호 2(tPGI1-3'-R; 도 1)에 의해 PCR 증폭시켰다. PCR 생성물을 진제트(GeneJET: 등록상표) 젤 추출 키트(페르멘타스, 독일 68789 세인트 레온-로트)에 의해 칼럼 상에서 경제하고, 제로 블런트(등록상표) TOPO(등록상표) PCR 클로닝 키트(인비트로젠, 미국 캘리포니아주 칼스바드)를 사용하여 TOPO 블런트 벡터내로 클로닝하고, 원 샷(등

록상표) TOP10 화학적 컴피턴트 에스케리키아 콜라이(인비트로겐, 미국 캘리포니아주 칼스바드)로 형질전환시켰다. 미니프렙 DNA 단리를 몇몇 콜로니 상에서 수행하였다. 수득된 플라스미드를 제한 효소 분석으로 검토하였다. 2종의 상이한 플라스미드를 수득하였는데, 하나의 플라스미드는 역방향 배향으로 삽입된 GL01 발현 카세트를 갖고(pRN1048, 도 5) 하나의 플라스미드는 순방향 배향으로 삽입된 GL01 발현 카세트를 갖는다(pRN1129, 도 6). 후속적으로 pRN1048로부터의 GL01 발현 카세트를 효모 셔틀 벡터 pRN599(도 7)내로 KpnI 및 ApaI 자리 사이에 클로닝하였다. 생성된 플라스미드를 원 샷(등록상표) TOP10 화학적 컴피턴트 에스케리키아 콜라이(인비트로겐, 미국 캘리포니아주 칼스바드)로 형질전환시켰다. 미니프렙 DNA 단리를 몇몇 콜로니 상에서 수행하였다. 수득된 플라스미드를 제한 효소 분석으로 검토하였다. 정확한 플라스미드를 pRN1049로 지칭하였다(도 8).

[0268] 플라스미드 pRN228은 다음의 관련 요소들을 함유한다: 제한 자리 SpeI 및 PstI에 의해 플랭크화된 PGK1 프로모터 및 카나마이신 내성 마커.

[0269] 플라스미드 pRN935는 다음의 관련 요소들을 함유한다: 제한 자리 PstI 및 SalI에 의해 플랭크화된 GL01 ORF 및 카나마이신 내성 마커.

[0270] 플라스미드 pRN685는 다음의 관련 요소들을 함유한다: 제한 자리 XbaI 및 HindIII에 의해 플랭크화된 PGI1 종결자 및 카나마이신 내성 마커.

[0271] 플라스미드 pRN599는 다음의 관련 요소들을 함유한다: kanMX 마커[아쉬바이아 고시피(*Ashbya gossypii*) TEF1 프로모터, KanMX 내성 유전자 및 아쉬바이아 고시피 TEF1 종결자로 구성됨] 및 앰피실린(ampicillin) 내성 마커가 수반되는 복제의 2  $\mu$  기원.

[0272] 프라이머 서열 번호 1은 PGK1 프로모터의 순방향 프라이머이다.

[0273] 프라이머 서열 번호 2는 PGI1 종결자의 역방향 프라이머이다.

[0274] 프라이머 서열 번호 25는 GL01 유전자의 증폭을 위한 순방향 프라이머이다.

[0275] 프라이머 서열 번호 26은 GL01 유전자의 증폭을 위한 역방향 프라이머이다.

## 실시예 2

### 과발현 플라스미드에 의한 효모 균주의 형질전환

[0276] 작성된 GL01 과발현 플라스미드 pRN1049를 사용하여 효모 균주 RN1001을 형질전환시켰다. 또한 효모 균주 RN1001을 빈 플라스미드 pRN599로 형질전환시킴으로써 대조 균주를 생성하였다. 정확한 형질전환체를 2% 글루코스 및 200  $\mu$ g/ml의 G418이 함유된 YEP 플레이트 상에서 선택하였다. pRN1049를 함유하는 RN1001 균주를 RN1001Gp1로 지칭한다. 빈 플라스미드 pRN599를 함유하는 대조 균주 RN1001을 RN 1001Ep1로 지칭한다.

[0277] RN1001Gp1에서 GL01 발현의 수준을 Q-PCR로 결정하였다. 대조 균주 RN1001Ep1을 기준으로 사용하였다. 이러한 실험을 위해, 양쪽 균주를 2% 글루코스 및 200  $\mu$ g/ml의 G418이 함유된 YEP 배지에서 하룻밤 성장시켰다. 후속적으로 RNA를 배양액으로부터 단리하고, 프라이머 서열 번호 3 및 서열 번호 4를 사용하는 PCR에 의해 게놈성 DNA 오염에 대해 검토하였다. 오염은 발견되지 않았다. 이어서 리버트에이드 키트(페르멘타스, 독일 68789 세인트 레온-로트)를 사용하여 RNA 상에서 cDNA 합성을 수행하였다. 그런 다음 Q-PCR 실험을 프라이머 서열 번호 5 및 서열 번호 6에 의해 실행하였다. 2개의 항존 유전자를 기준으로 사용하였는데, 프라이머 서열 번호 7 및 서열 번호 8을 사용하는 ALG9 및 프라이머 서열 번호 9 및 서열 번호 10을 사용하는 UBC6이었다. 최적 2배(best duplo) CT-값을 갖는 항존 유전자, 이 경우 ALG9에 대해 GL01 발현 데이터를 정규화하였다. RN1001Gp1 균주는 RN1001Ep1에 비하여 GL01의 더 높은 발현을 보여주었다(도 9). 도 9에 지시된 바와 같이, GL01-유전자의 정규화된 발현은 양쪽 균주에서 동일하였다.

[0278] 프라이머 서열 번호 3은 ACT1 유전자의 서열과 동일하였다.

[0279] 프라이머 서열 번호 4는 ACT1 유전자의 서열과 동일하였다.

[0280] 프라이머 서열 번호 5는 GL01 유전자의 서열과 동일하였다.

[0281] 프라이머 서열 번호 6은 GL01 유전자의 서열과 동일하였다.

[0282] 프라이머 서열 번호 7은 ALG9 유전자의 서열과 동일하였다.

[0285] 프라이머 서열 번호 8은 ALG9 유전자의 서열과 동일하였다.

[0286] 프라이머 서열 번호 9는 UBC6 유전자의 서열과 동일하였다.

[0287] 프라이머 서열 번호 10은 UBC6 유전자의 서열과 동일하였다.

실시예 3

AFM 실험

[0289] Q-PCR에 의해 대조 균주 RN1001Ep1과 비교하여 RN1001Gp1에서 더 높은 GL01 발현을 입증한 후(실시예 2), 에탄올 및 CO<sub>2</sub>가 동물량으로 생산되므로, 실험 동안 CO<sub>2</sub> 생성을 측정함으로써 제공된 당분의 에탄올로의 전환 속도를 결정하기 위해 AFM 실험을 시작하였다. 실험 동안, 글루코스 및 자일로스 소비 속도 및 에탄올 생산 속도를 결정하기 위해 배양액의 HPLC 샘플을 상이한 시점에 취하였다.

[0291] 아가 플레이트로부터의 일부 세포 물질을 2% 글루코스 및 200 µg/ml의 G418이 함유된 100 ml의 YEP에 접종함으로써 두 균주로 예비배양액을 제조하였다. 다음날, 2% 글루코스 및 2% 자일로스를 함유한 400 ml의 광물 배지에 예비배양액으로부터의 세포 물질을 접종하였다. AFM 실험에서 G418을 광물 배지에 첨가하지 않았다.

[0292] CO<sub>2</sub> 생산 곡선(도 10)은 RN1001Ep1에 비하여 RN1001Gp1에서 더 신속하고 더 높은 CO<sub>2</sub> 생산 속도를 나타내고, 이는 대조 균주 RN1001Ep1에 비하여 RN1001Gp1에서 더 신속한 글루코스 및 자일로스 소비를 지시한다. HPLC 데이터(표 4 및 5, 및 도 11)는 RN1001Ep1 균주에 비하여 RN1001Gp1 균주에서 더 신속한 자일로스 소비 속도를 확인하였다. 글루코스는 제2 시점에서 이미 소비된 반면(16.5 시간), 이는 균주 RN1001Ep1에서는 그렇지 않았다.

**표 4**

HPLC 데이터 RN1001Gp1

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
<b>0</b>	21,4	20,2	0,0
<b>16,5</b>	0,1	8,8	20,0
<b>18,5</b>	0,0	5,2	16,1
<b>21</b>	0,0	1,9	16,6
<b>22,5</b>	0,0	0,8	17,0
<b>24,5</b>	0,0	0,3	17,3
<b>40</b>	0,0	0,0	17,6

[0293]

표 5

HPLC 데이터 RN1001Epl

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
0	21,4	20,3	0,0
16,5	0,3	13,4	17,7
18,5	0,0	9,1	14,4
21	0,0	4,5	15,8
22,5	0,0	2,0	16,8
24,5	0,0	0,6	17,5
40	0,0	0,1	17,9

[0294]

G418이 AFM 실험에 참가되지 않았으므로, 이는 얼마나 많은 세포가 여전히 이들 플라스미드를 함유하는지를 결정하였다. 따라서 마지막 AFM 실험의 종료시, 동일한 양의 세포 배양액을 YEPD 아가 플레이트 및 G418이 함유된 YEPD 아가 플레이트 상에 풀레이팅하였다. G418이 없는 아가 플레이트 상에서 4배 많은 콜로니가 수득되었다. 이들 결과로부터 세포의 대략 25%가 여전히 이들 플라스미드를 함유함을 알 수 있고, 이는 RN1001Gp1 및 RN1001Epl 둘 다에서 플라스미드가 손실됨을 지시한다. 따라서, 계놈에 안정하게 통합된, GL01 유전자의 외부의 구성적으로 발현된 복사물을 함유하는 신규한 형질전환체가 작성되었다.

[0296]

#### 실시예 4

[0297]

##### 통합적 단편의 구조화

[0298]

AFM 실험 동안 플라스미드 손실의 문제를 극복하기 위해, GL01 발현 카세트를 정확한 형질전환체의 선택을 위해 HIS3과 함께 효모 균주의 계놈내로 통합하였다. 이를 위해, GL01 발현 카세트 및 HIS3 발현 카세트를 함유하는 플라스미드를 작성하였다. GL01 발현 카세트(pRN1129에 존재함) 및 HIS3 발현 카세트(pRN324에 존재함)를 제한 효소 Pvull 및 SphI 또는 ApaLI에 의해 플라스미드로부터 추출하였다(도 6 및 12). 후속적으로 양쪽 단편을 Pvull 자리 상에서 함께 결찰시켰다. 프라이머 서열 번호 11 및 서열 번호 12를 사용하여 PCR에 의해 생성된 단편을 증폭시키고, 제로 블런트(등록상표) TOPO(등록상표) PCR 클로닝 키트(인비트로젠, 미국 캘리포니아주 칼스바드)를 사용하여 TOPO 블런트 벡터내로 클로닝하고, 원 샷(등록상표) TOP10 화학적 컴파인트 에스케리키아 콜라이(인비트로젠, 미국 캘리포니아주 칼스바드)로 형질전환시켰다. 미니프렙 DNA 단리를 몇몇 콜로니 상에서 수행하였다. 수득된 플라스미드를 제한 효소 분석으로 검토하였다. 정확한 플라스미드를 pRN1142로 지칭하였다(도 13).

[0299]

후속적으로 프라이머 서열 번호 11 및 서열 번호 12 및 pRN1142를 주형으로서 사용하여 GL01 및 HIS3 단편을 PCR 증폭시켰다. 대조 균주를 생성하기 위해, 프라이머 서열 번호 11 및 서열 번호 13 및 pRN324를 주형으로서 사용하여 HIS3 발현 카세트를 PCR 증폭시켰다. 계놈에서의 통합을 위해, 2개의 500 bp 플랭크가 통합 자리 1(INT1)을 수득하기 위해 필요하고, 통합 자리의 지도는 도 14에 제공된다(<http://www.yeastgenome.org/>로부터 수득됨). INT1로의 통합을 위해, 5' 플랭크에 대해 프라이머 서열 번호 14 및 서열 번호 15를 사용하고 3' 플랭크에 대해 프라이머 서열 번호 16 및 서열 번호 17을 사용하여 5' 및 3' 500bp 플랭크를 PCR 증폭시켰다. 주형으로서 RN1001로부터의 계놈성 DNA를 사용하였다. 모든 PCR 반응물을 진제트(등록상표) 젤 추출 키트(페르멘타스, 독일 68789 세인트 레온-로트)에 의해 칼럼 상에서 정제하였다. 5' 플랭크의 뉴클레오티드 서열은 서열 번호 18로서, 3' 플랭크는 서열 번호 19로서 포함되었다.

[0300]

플라스미드 pRN324는 다음의 관련 요소들을 함유한다: 제한 자리 ApaLI 및 Pvull에 의해 플랭크화된 HIS3 발현 카세트, 및 앰피실린 내성 마커.

[0301]

플라스미드 pRN1129는 다음의 관련 요소들을 함유한다: 제한 자리 Pvull 및 SphI에 의해 플랭크화된 GL01 발현

카세트(PGK1 프로모터, GL01 ORF 및 PGI1 종결자로 구성됨), 및 카나마이신 내성 마커.

[0302] 프라이머 서열 번호 11은 5' 500 bp INT1 플랭크의 3'-말단의 50개의 뉴클레오티드와 동일한, 5'-말단 상의 25개의 뉴클레오티드 및 50개 뉴클레오티드의 테일로 구성된, HIS3 카세트의 순방향 프라이머이다.

[0303] 프라이머 서열 번호 12는 3' 500 bp INT1 플랭크의 5'-말단의 50개의 뉴클레오티드와 동일한, 5'-말단 상의 25개의 뉴클레오티드 및 50개 뉴클레오티드의 테일로 구성된, GL01 카세트의 역방향 프라이머이다.

[0304] 프라이머 서열 번호 13은 3' 500 bp INT1 플랭크의 5'-말단의 50개의 뉴클레오티드와 동일한, 5'-말단 상의 25개의 뉴클레오티드 및 50개 뉴클레오티드의 테일로 구성된, HIS3 카세트의 역방향 프라이머이다.

[0305] 프라이머 서열 번호 14는 5' 500 bp INT1 플랭크의 순방향 프라이머이다.

[0306] 프라이머 서열 번호 15는 5' 500 bp INT1 플랭크의 역방향 프라이머이다.

[0307] 프라이머 서열 번호 16은 3' 500 bp INT1 플랭크의 순방향 프라이머이다.

[0308] 프라이머 서열 번호 17은 3' 500 bp INT1 플랭크의 역방향 프라이머이다.

[0309] 상기 언급된 모든 프라이머 서열은 또한 도 15에 지시된다.

#### 실시예 5

##### 통합적 단편에 의한 효모 균주의 형질전환

[0312] PCR 단편의 개념으로의 통합 기작은 도 15에 개략적으로 제시된다. 수득된 PCR 단편(실시예 4)을 사용하여 균주 RN1041 및 RN1216을 형질전환시켰고, 이를 둘다 히스티딘에 대해 영양요구성을 갖는다. 정확한 형질전환체를 2% 글루코스가 함유된 YNB 플레이트 상에서 선택하였다. 프라이머 서열 번호 14 및 서열 번호 17을 사용하여, 발현 카세트의 개념내로의 통합에 대해 형질전환체를 콜로니 PCR에 의해 검토하였다. HIS3 및 GL01 발현 카세트를 함유하는 RN1041 균주를 RN1041HG로 지칭하고, HIS3 발현 카세트를 함유하는 RN1041 대조 균주를 RN1041H로 지칭한다. HIS3 및 GL01 발현 카세트를 함유하는 RN1216 균주를 RN1216HG로 지칭하고, HIS3 발현 카세트를 함유하는 RN1216 대조 균주를 RN1216H로 지칭한다.

[0313] 두 대조 균주, RN1041H 및 RN1216H로부터, 각각 하나의 콜로니를 Q-PCR을 위해 선택하였다. RN1041HG 및 RN1216HG로부터, 각각 2개의 개별적 콜로니를 Q-PCR을 위해 선택하였다(클론 1 및 클론 2).

[0314] 형질전환체중 GL01-유전자의 발현을 Q-PCR 실험에서 검토하였다. Q-PCR 실험을 위해, 모든 선택된 균주를 2% 글루코스가 함유된 YEP 배지에서 하룻밤 성장시켰다. 후속적으로 RNA를 배양액으로부터 단리하고, 프라이머 서열 번호 3 및 서열 번호 4를 사용하여 PCR에 의해 개념성 DNA 오염에 대해 검토하였다. 오염은 발견되지 않았다. 이어서 리버트에이드 키트(페르멘타스, 독일 68789 세인트 레온-로트)를 사용하여 RNA 상에서 cDNA 합성을 수행하였다. 그런 다음 Q-PCR 실험을 프라이머 서열 번호 5 및 서열 번호 6에 의해 실행하였다. 2개의 항준 유전자를 기준으로 사용하였는데, 프라이머 서열 번호 7 및 서열 번호 8을 사용하는 ALG9 및 프라이머 서열 번호 9 및 서열 번호 10을 사용하는 UBC6이었다. 최적 2배 CT-값을 갖는 항준 유전자, 이 경우 UBC6에 대해 GL01 발현 데이터를 정규화하였다. HIS3 및 GL01을 함유하는 모든 4종의 균주는 이들의 대조 균주에 비하여 GL01의 더 높은 정규화된 발현을 보여주었고(도 16), ALG9에 비하여 RN1041 배경에서 약 4 내지 5배 더 높고, RN1216 배경에서 약 2 내지 3 배 더 높았다.

#### 실시예 6

##### AFM 실험

[0317] HIS3 및 GL01을 함유하는 모든 4종의 균주에서 Q-PCR에 의해 이들의 대조 균주에 비해 더 높은 GL01 발현을 입증한 후(실시예 5), 에탄올 및 CO<sub>2</sub>가 동물량으로 생산되므로, 실험 동안 CO<sub>2</sub> 생산을 측정함으로써 제공된 당분의 에탄올로의 전환 속도를 결정하기 위해 AFM 실험을 시작하였다. 실험 동안, 당분 소비 속도 및 에탄올 생산 속도를 결정하기 위해 배양액의 HPLC 샘플을 상이한 시점에 취하였다. GL01 발현의 결정을 위해 Q-PCR 실험에서 사용된 것과 동일한 6종의 균주를 이러한 AFM 실험에서 시험하였다.

[0318] 플레이트로부터의 일부 세포 물질을 2% 글루코스가 함유된 100 ml의 YEP에 접종함으로써 모든 3종의 RN1041 유래된 균주(RN1041H, RN1041HG-1 및 RN1041HG-2)로 예비배양액을 제조하였다. 다음날 2% 글루코스 및 2% 자일로스가 함유된 400 ml의 광물 배지에 예비배양액으로부터의 세포 물질을 접종하였다.

[0319]

RN1041 유래된 균주의  $\text{CO}_2$  생산 곡선(도 17)은 RN1041H에 비해 RN1041HG(클론 1 및 2)에서 더 신속하고 더 높은  $\text{CO}_2$  생산 속도를 보여주었고, 이는 대조 균주 RN1041H에 비해 RN1041HG(클론 1 및 2)에서 더 신속한 글루코스 및 자일로스 소비를 지시한다. HPLC 데이터(표 6, 7 및 8, 및 도 18)는 실제로 RN1041H 균주에 비해 RN1041HG(클론 1 및 2)에서 더 신속한 자일로스 소비를 확인하고, 글루코스는 이미 제2 시점에 소비되었다(15 시간).

표 6

HPLC 데이터 RN1041H

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
0	20,9	20,7	0,0
15	0,3	16,1	14,2
17	0,0	14,5	12,5
19	0,0	12,3	12,6
21	0,0	9,6	13,4
23	0,0	7,1	14,5
39	0,0	0,5	17,5

[0320]

표 7

HPLC 데이터 RN1041HG-1

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
0	20,8	20,8	0,0
15	0,1	15,4	14,0
17	0,0	13,4	12,9
19	0,0	10,6	13,1
21	0,0	7,5	14,2
23	0,0	4,6	15,6
39	0,0	0,4	17,6

[0321]

표 8

HPLC 데이터 RN1041HG-2

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
0	21,3	21,3	0,0
15	0,1	15,3	13,3
17	0,0	12,5	13,6
19	0,0	9,4	14,2
21	0,0	6,2	15,6
23	0,0	3,3	16,6
39	0,0	0,2	18,2

[0322]

[0323] 이를 결과는, GL01-유전자의 과발현이 생물자원 당분의 더 신속하고 더 효율적인 발효를 유도하여 발효 시간을 감소시킴을 명백히 보여주었다. 이는 당분이 더욱 효율적으로 에탄올로 전환됨을 의미한다.

[0324]

신규한 AFM 실험을 RN1216 유래된 균주로 시작하였다. 아가 플레이트로부터의 일부 세포 물질을 2% 글루코스가 함유된 100 ml의 YEP에 접종함으로써 모두 3종의 균주(RN1216H, RN1216HG-1 및 RN1216HG-2)로 예비배양액을 제조하였다. 다음날, 5% 글루코스 및 5% 자일로스를 함유한 400 ml의 광물 배지에 예비배양액으로부터의 세포 물질을 접종하였다.

[0325]

RN1216의 CO<sub>2</sub> 생산 곡선(도 19)은 RN1216H에 비하여 RN1216HG(클론 1 및 2)에서 더 신속하고 더 높은 CO<sub>2</sub> 생산 속도를 나타내었고, 이는 대조 균주 RN1216H에 비하여 RN1216HG(클론 1 및 2)에서 더 신속하고 더 효율적인 글루코스 및 자일로스 소비를 지시한다. HPLC 데이터(표 9, 10 및 11, 및 도 20)는 RN1216H 균주에 비하여 RN1216HG(클론 1 및 2) 균주에서 더 신속한 글루코스 및 자일로스 소비를 확인하였다.

표 9

HPLC 데이터 RN1216H

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
0	52,5	51,4	0,0
4	49,8	50,5	0,7
8	45,9	50,3	2,2
12	39,1	49,6	5,1
16	28,5	48,6	10,0
20	14,7	47,0	16,8
24,5	1,4	44,2	28,5
39,5	0,2	27,5	35,1
43,5	0,0	21,0	36,5
47,5	0,0	16,1	38,5
112	0,0	2,1	45,4

[0326]

**표 10**

HPLC 테이터 RN1216HG-1

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
0	52,1	51,3	0,0
4	49,6	50,4	0,7
8	45,5	50,1	2,3
12	38,1	49,4	5,5
16	26,8	48,3	10,8
20	12,0	46,2	18,1
24,5	0,2	42,1	28,9
39,5	0,0	22,7	36,5
43,5	0,0	15,2	38,8
47,5	0,0	10,4	40,7
112	0,0	1,1	45,5

[0327]

**표 11**

HPLC 테이터 RN1216HG-2

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
0	52,2	51,2	0,0
4	49,6	50,4	0,8
8	44,9	50,0	2,6
12	37,2	49,3	6,0
16	25,9	47,9	11,3
20	11,5	45,8	18,6
24,5	0,2	41,0	29,1
39,5	0,0	17,5	38,5
43,5	0,0	12,4	40,2
47,5	0,0	7,3	42,4
112	0,0	0,8	45,7

[0328]

[0329] 이들 결과는 상기 결과를 재확인하였다: GL01-유전자의 과발현이 생물자원 당분, 이 경우 자일로스 및 글루코스의 더 신속하고 더 효율적인 발효를 유도하여 전체 발효 시간을 감소시키며, 전술된 당분을 더욱 효율적으로 에탄올로 전환시킨다.

[0330] 실시예 7[0331] GL01 변형체

[0332] 사카로마이세스 세레비지아 이외의 유기체로부터의 GL01 상동체를 이전 실시예에 기재된 바와 동일한 양식으로 발현시킨다. 이러한 목적을 위해서, 서열 번호 20 내지 서열 번호 24로서 열거된 단백질 서열에 기초하여 유전자 서열을 합성하였다. 개방 판독 프레임의 서열을 국제특허공개 제W02008/000632호에 기재된 방법에 의해 생성하였다. 단백질 서열은 사카로마이세스 세레비지아, 칸디다 글라브라타, 자이고사카로마이세스 륙시, 클루이 베로마이세스 락티스 및 칸디다 마그놀리아로부터 유래된다. 기준으로서, 사카로마이세스 세레비지아로부터의 야생형 GL01 유전자(이전 실시예 참조)가 사용된다.

[0333] 이렇게 수득된 구조물을 사용하여 이전 실시예에 기재된 바와 같이 균주 RN 1216을 형질전환시켰다.

[0334] 모든 형질전환체는 AFM 실험에서 시험될 경우 증진된 당분 소비 속도를 보여주고(또한 실시예 10 참조), 둘 다 실험 동안 이산화탄소 프로파일 및 실제 당분 농도에 기초한다.

[0335] 하나의 실시태양에서 몇몇 활성 GL01 형태로 존재하는 수반되는 특징은, 특별히 혼합된 당분 기질의 공동-발효를 증진시키는데 있어서, 다음과 같다:

## 표 12

GL01 특징 패턴

서열	사카로마이세스 세레비지아에서의 위치(서열 번호 20)
(E,s,d)-L-X-(H,Y)-(N,s)	E242-L243-X-H245-N246
G-(F,Y)-G-H	G266-Y267-G268-H269
G-X(6)-(F,i)-X(2,3)-D-X(3)-Y	G301-X(6)-F308-X(2)-D311-X(3)-Y315

[0336] [0337] 표 12에서, X는 아미노산(즉 임의의 아미노산)을 지정한다.

[0338] 2개의 아미노산 잔기가 괄호안에 언급된다면, 하나를 적용한다(예를 들어 (F,Y)).

[0339] 소문자는 소수의 변형체를 지정하고, 예를 들어 (F,i)는 대부분의 경우 아미노산 F가 관찰되지만, 소수의 경우 아미노산 I(이소로이신)이 관찰된다.

## 실시예 8

### 통합적 단편의 작성

[0340] [0341] GL01 상동체를 정확한 형질전환체의 선택을 위해 HIS3과 함께 효모 균주 RN1216의 계놈내로 통합시켰다. 이를 위해, 플라스미드를 작성하였고, 각각 상이한 GL01 상동체를 PGK1 프로모터 및 PGI1 종결자와 함께 함유한다. 플라스미드를 제PCT/EP2013/056623호에 기재된 방법에 따라 작성하였다. 플라스미드를 pDB1175(도 21), pDB1176(도 22), pDB1177(도 23), pDB1178(도 24) 및 pDB1179(도 25)로 지칭하였다.

[0342] [0343] 후속적으로 프라이머 서열 번호 28 및 서열 번호 29 및 pDB1175, pDB1176, pDB1177, pDB1178 및 pDB1179를 주형으로서 사용하여 상이한 GL01 발현 카세트를 PCR 증폭시켰다. 서열 번호 30 및 서열 번호 31 및 pRN324(도 12)를 주형으로서 사용하여 HIS3 발현 카세트를 PCR 증폭시켰다. 계놈에서의 통합을 위해, 2개의 500 bp 플랭크가 통합 자리 1(INT1)을 수득하기 위해 필요하였다(실시예 4에 기재됨). INT1로의 통합을 위해, 5' 플랭크에 대해 프라이머 서열 번호 14 및 서열 번호 32를 사용하고 3' 플랭크에 대해 프라이머 서열 번호 33 및 서열 번호 17을 사용하여 5' 및 3' 500 bp 플랭크를 PCR 증폭시켰다. RN1001로부터의 계놈성 DNA를 주형으로서 사용하였다. 모든 PCR 반응을 진제트(등록상표) 웰 추출 키트(페르멘타스, 독일 68789 세인트 레온-로트)에 의해 칼럼 상에서 정제하였다.

[0344] 플라스미드 pDB1175는 다음의 관련 요소들을 함유한다: Sc\_GL01 발현 카세트(PGK1 프로모터, 사카로마이세스 세레비지아로부터의 코돈 쌍 최적화된 GL01 ORF 및 PGI1 종결자로 구성됨), 및 카나마이신 내성 마커.

[0345] 플라스미드 pDB1176은 다음의 관련 요소들을 함유한다: Cgla\_GL01 발현 카세트(PGK1 프로모터, 칸디다 글라브라타로부터의 코돈 쌍 최적화된 GL01 ORF 및 PGI1 종결자로 구성됨), 및 카나마이신 내성 마커.

[0346] 플라스미드 pDB1177은 다음의 관련 요소들을 함유한다: Zrou\_GL01 발현 카세트(PGK1 프로모터, 자이고사카로마이세스 륙시로부터의 코돈 쌍 최적화된 GL01 ORF 및 PGI1 종결자로 구성됨), 및 카나마이신 내성 마커.

[0347] 플라스미드 pDB1178은 다음의 관련 요소들을 함유한다: KI,GL01 발현 카세트(PGK1 프로모터, 클루이베로마이세스 락티스로부터의 코돈 쌍 최적화된 GL01 ORF 및 PGI1 종결자로 구성됨), 및 카나마이신 내성 마커.

[0348] 플라스미드 pDB1179는 다음의 관련 요소들을 함유한다: Cmag,GL01 발현 카세트(PGK1 프로모터, 칸디다 마그놀리아로부터의 코돈 쌍 최적화된 GL01 ORF 및 PGI1 종결자로 구성됨), 및 카나마이신 내성 마커.

[0349] 플라스미드 pRN324: 실시예 4에 기재됨

[0350] 프라이머 서열 번호 28은 상동성 GL01 발현 카세트(프로모터 및 종결자를 포함함)의 순방향 프라이머이다. 프라이머 서열 번호 29는 상동성 GL01 발현 카세트(프로모터 및 종결자를 포함함)의 역방향 프라이머이다. 프라이머 서열 번호 30은 상동성 GL01 발현 카세트의 3'-말단의 50개의 뉴클레오티드와 동일한, 5'-말단 상의 20개의 뉴클레오티드 및 50개 뉴클레오티드의 테일로 구성된, HIS3 카세트의 순방향 프라이머이다. 프라이머 서열 번호 31은 3' 500 bp INT1 플랭크의 5'-말단의 50개의 뉴클레오티드와 동일한, 5'-말단 상의 21개의 뉴클레오티드 및 50개 뉴클레오티드의 테일로 구성된, HIS3 카세트의 역방향 프라이머이다. 프라이머 서열 번호 32는 상동성 GL01 발현 카세트의 5'-말단의 50개의 뉴클레오티드와 동일한, 5'-말단 상의 23개의 뉴클레오티드 및 50개 뉴클레오티드의 테일로 구성된, 5' 500 bp INT1 플랭크의 역방향 프라이머이다. 프라이머 서열 번호 33은 HIS3 발현 카세트의 3'-말단의 50개의 뉴클레오티드와 동일한, 5'-말단 상의 24개의 뉴클레오티드 및 50개 뉴클레오티드의 테일로 구성된, 3' 500 bp INT1 플랭크의 순방향 프라이머이다.

[0351] 상기 언급된 모든 프라이머 서열은 또한 도 26에 지시되어 있다.

### 실시예 9

#### 통합적 단편에 의한 효모 균주의 형질전환

[0354] PCR 단편의 계놈으로의 통합 기작은 도 26에 개략적으로 도시되어 있다. 수득된 PCR 단편(실시예 8)을 사용하여 균주 RN1216을 형질전환시켰고, 이는 히스티딘에 대한 영양요구성을 함유한다. 정확한 형질전환체를 2% 글루코스가 함유된 YNB 플레이트 상에서 선택하였다. 발현 카세트의 계놈으로의 통합을 위해 프라이머 서열 번호 14 및 서열 번호 17을 사용하여 콜로니 PCR에 의해 형질전환체를 검토하였다. 정확한 균주를 다음과 같이 지칭하였다: RN1216 ScG\_H, RN1216 CglaG\_H, RN1216 ZrouG\_H, RN1216 KIG\_H 및 RN1216 CmagG\_H.

[0355] 균주 RN1216 ScG\_H는 Sc\_GL01 발현 카세트 및 HIS3 발현 카세트를 함유한다.

[0356] 균주 RN1216 CglaG\_H는 Cgla\_GL01 발현 카세트 및 HIS3 발현 카세트를 함유한다.

[0357] 균주 RN1216 ZrouG\_H는 Zrou\_GL01 발현 카세트 및 HIS3 발현 카세트를 함유한다.

[0358] 균주 RN1216 KIG\_H는 KI,GL01 발현 카세트 및 HIS3 발현 카세트를 함유한다.

[0359] 균주 RN1216 CmagG\_H는 Cmag,GL01 발현 카세트 및 HIS3 발현 카세트를 함유한다.

### 실시예 10

#### AFM 실험

[0362] 에탄올 및 CO<sub>2</sub>가 동물량으로 생산되므로, 실험 동안 CO<sub>2</sub> 생산을 측정함으로써 제공된 당분의 에탄올로의 전환 속도를 결정하기 위해 AFM 실험을 시작하였다. 실험 동안, 당분 소비 속도 및 에탄올 생산 속도를 결정하기 위해 배양액의 HPLC 샘플을 상이한 시점에 취하였다. 이러한 AFM 실험에서 시험된 균주는 실시예 9에 기재된 5종의 작성된 균주이고, 대조 균주로서 실시예 5에 기재된 RN1216H 균주를 사용하였다.

[0363] 플레이트로부터의 일부 세포 물질을 2% 글루코스가 함유된 100 ml의 YEP에 접종함으로써, 예비배양액을 모든 6종의 균주로 제조하였다. 다음날 5% 글루코스 및 5% 자일로스가 함유된 400 ml의 광물 배지에 예비배양액으로부터의 세포 물질을 접종하였다.

[0364] RN1216 유래된 균주의 CO<sub>2</sub> 생산 곡선(도 27)은 RN1216H에 비해 RN1216 ScG\_H, RN1216 CglaG\_H, RN1216 KIG\_H 및 RN1216 CmagG\_H에서 더 신속하고 더 높은 CO<sub>2</sub> 생산을 보여주었고, 이는 대조 균주 RN1216H에 비해 이들 4종의 GL01 변형 균주에서 더 신속한 글루코스 및 자일로스 소비를 지시한다. 균주 RN1216 ZrouG\_H는 RN1216H에 비해 실험의 최초 20 시간 이내에 더 신속하고 더 높은 CO<sub>2</sub> 생산을 보여주었고, 이는 이러한 AFM 실험에서 시험된 5종의 다른 균주에 비해 더 신속한 글루코스 소비 및 더 느린 자일로스 소비를 지시한다. HPLC 데이터(표

13, 14, 15, 16, 17, 18 및 도 28, 29, 30, 31, 32, 33)는 실제로 RN1216H 균주에 비해 RN1216 ScG\_H, RN1216 CglaG\_H, RN1216 KIG\_H 및 RN1216 CmagG\_H 균주에서 더 신속한 글루코스 및 자일로스 소비를 확인한다. HPLC 데이터는 또한 이러한 AFM 실험에서 시험된 5종의 다른 균주에 비해 RN1216 ZrouG\_H 균주에서 더 신속한 글루코스 소비를 확인한다.

표 13

HPLC 데이터 RN1216 ScG\_H

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
<b>0</b>	50,2	52,9	0,0
<b>8</b>	46,8	51,5	2,5
<b>12</b>	40,5	52,4	5,5
<b>16</b>	28,5	53,6	10,9
<b>20</b>	13,0	48,6	17,6
<b>24</b>	3,3	46,2	22,9
<b>28</b>	0,2	40,2	26,0
<b>32</b>	0,0	33,0	28,5
<b>36</b>	0,0	19,7	35,1
<b>40</b>	0,0	7,9	41,2
<b>44</b>	0,0	2,6	44,0
<b>48</b>	0,0	1,1	44,5

[0365]

표 14

HPLC 데이터 RN1216 CglaG\_H

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
<b>0</b>	50,4	53,2	0,0
<b>4</b>	49,6	53,1	1,3 <sup>20</sup>
<b>8</b>	49,0	55,1	2,4
<b>16</b>	29,1	53,4	10,5
<b>20</b>	9,6	46,5	19,5
<b>24</b>	1,2	43,6	24,1
<b>28</b>	0,0	35,2	28, <sup>95</sup>
<b>32</b>	0,0	24,8	32,7
<b>36</b>	0,0	15,0	37,6
<b>40</b>	0,0	7,7	41,7
<b>44</b>	0,0	3,1	44,1
<b>48</b>	0,0	1,2	44, <sup>80</sup>

[0366]

**표 15**

HPLC 테이터 RN1216 ZrouG\_H

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
<b>0</b>	50,3	53,0	0,0
<b>4</b>	48,6	53,1	0,7
<b>12</b>	38,4	51,2	5,1
<b>16</b>	26,1	53,1	11,3
<b>20</b>	3,5	46,8	21,8
<b>24</b>	0,4	44,0	24,1
<b>28</b>	0,0	36,8	28,2
<b>32</b>	0,0	29,4	30,6
<b>36</b>	0,0	21,6	34,5
<b>40</b>	0,0	15,4	37,9
<b>44</b>	0,0	10,3	40,7
<b>48</b>	0,0	6,7	42,3

[0367]

**표 16**

HPLC 테이터 RN1216 KIG\_H

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
<b>0</b>	50,5	53,0	0,0
<b>8</b>	47,8	53,8	1,7
<b>16</b>	30,8	53,3	9,6
<b>20</b>	11,8	45,5	18,5
<b>24</b>	1,1	41,0	24,6
<b>28</b>	0,2	33,9	25,6
<b>32</b>	0,0	27,7	31,0
<b>36</b>	0,0	15,8	37,4
<b>40</b>	0,0	7,9	41,6
<b>44</b>	0,0	3,3	44,3
<b>48</b>	0,0	1,4	44,8

[0368]

**표 17**

HPLC 데이터 RN1216 CmagG\_H

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
<b>0</b>	50,5	52,8	0,0
<b>4</b>	46,5	50,3	3,2
<b>8</b>	40,4	49,4	7,1
<b>12</b>	33,3	49,6	9,0
<b>16</b>	31,0	52,3	9,9
<b>24</b>	2,3	43,6	23,8
<b>28</b>	0,0	35,4	28,9
<b>32</b>	0,0	24,6	32,7
<b>36</b>	0,0	14,5	37,8
<b>40</b>	0,0	6,8	41,6
<b>44</b>	0,0	2,5	44,1
<b>48</b>	0,0	0,9	44,8

[0369]

**표 18**

HPLC 데이터 RN1216H

시간 (시간)	글루코스 (g/l)	자일로스 (g/l)	에탄올 (g/l)
<b>0</b>	50,5	52,9	0,0
<b>8</b>	45,7	51,1	4,2 <sup>10</sup>
<b>12</b>	34,6	51,0	7,1
<b>16</b>	27,8	50,0	9,9
<b>20</b>	16,3	50,7	15,8
<b>24</b>	6,8	47,8	20,7 <sub>15</sub>
<b>28</b>	1,1	44,6	24,1
<b>32</b>	0,0	39,2	25,8
<b>36</b>	0,0	28,0	31,6
<b>40</b>	0,0	12,8	30,8
<b>44</b>	0,0	8,9	41,5 <sub>0</sub>
<b>48</b>	0,0	5,3	42,7

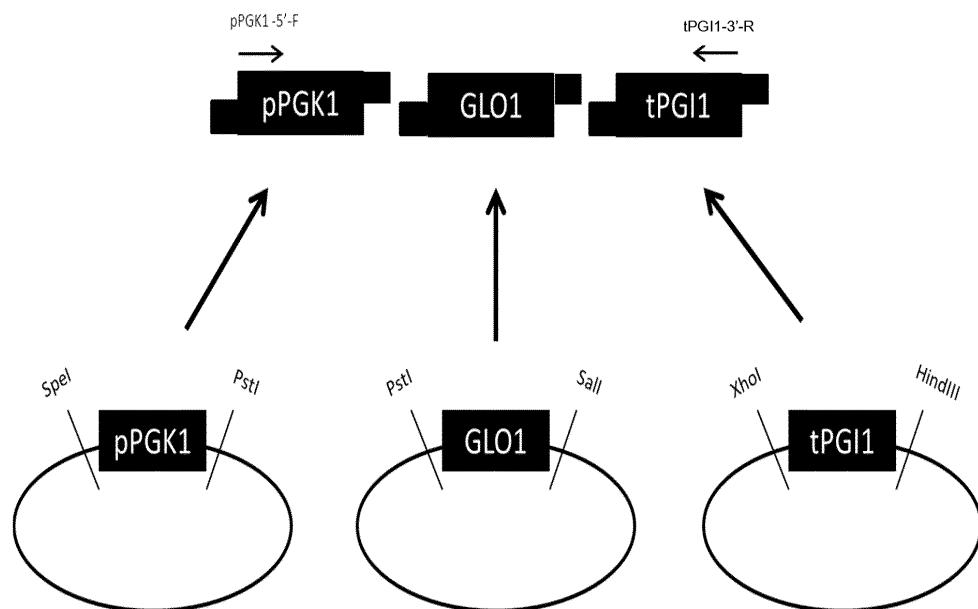
[0370]

[0371] 이들 결과로부터, 균주 사카로마이세스 세레비지아, 칸디다 글라브라타, 클루이베로마이세스 락티스 또는 칸디다 마그놀리아로부터 유래된 GL01-유전자의 과발현이 생물자원 당분, 이 경우 자일로스 및 글루코스의 더 신속하고 더 효율적인 발효를 유도하여, 감소된 전체 발효 시간 및 전술된 당분의 에탄올로의 더 효율적인 전환을

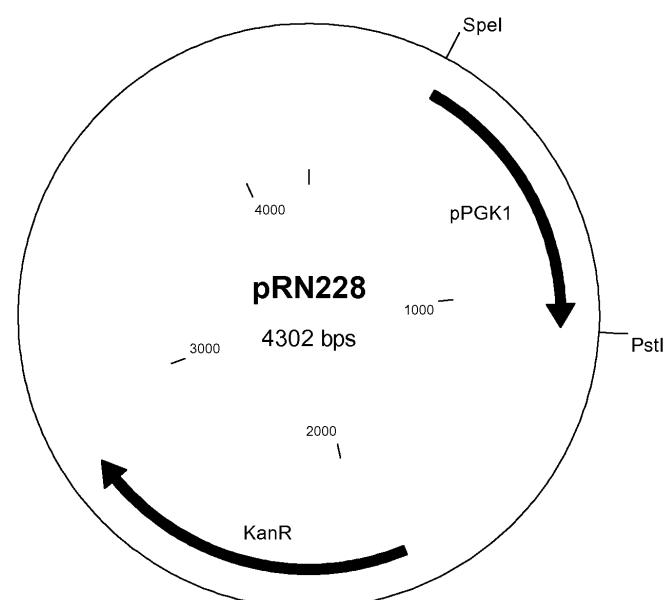
이끌어 냅을 알 수 있다. 하나의 예외는 균주 자이고사카로마이세스 륙시로부터 유래된 GL01-유전자의 과발현인데, 이는 이 실험에서 시험된 다른 균주에 비해 더 신속한 글루코스 소비를 나타내지만, 이 균주의 전체 성능은 더 불량하다.

## 도면

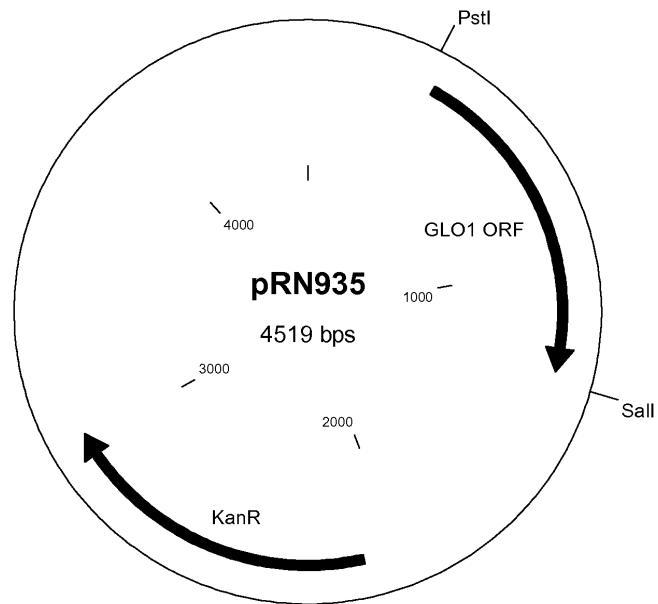
### 도면1



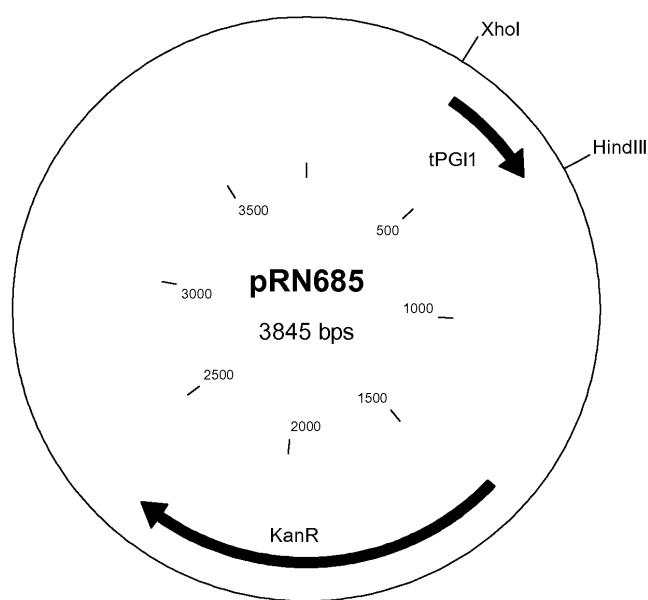
### 도면2



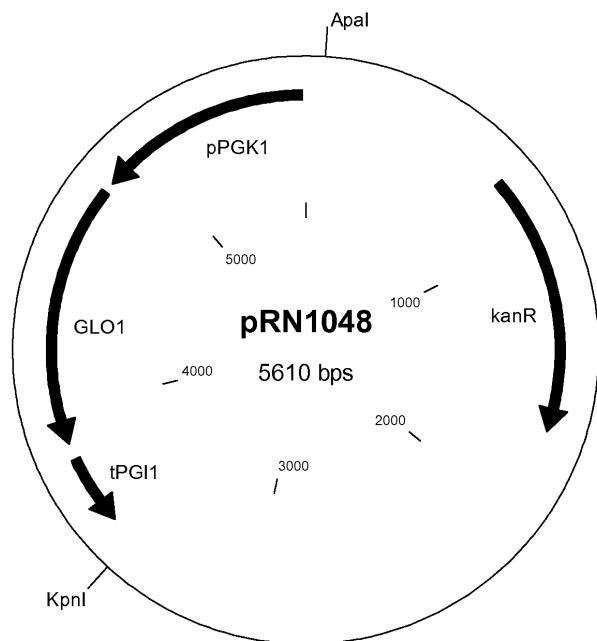
도면3



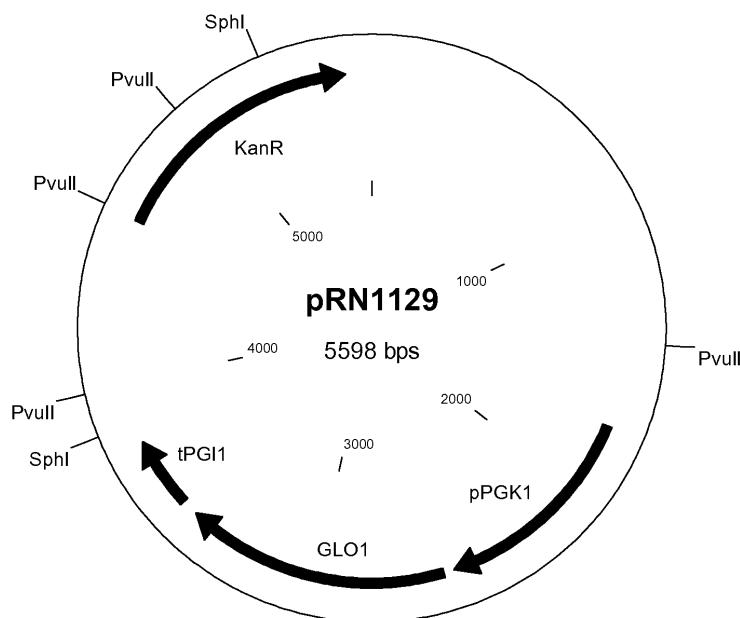
도면4



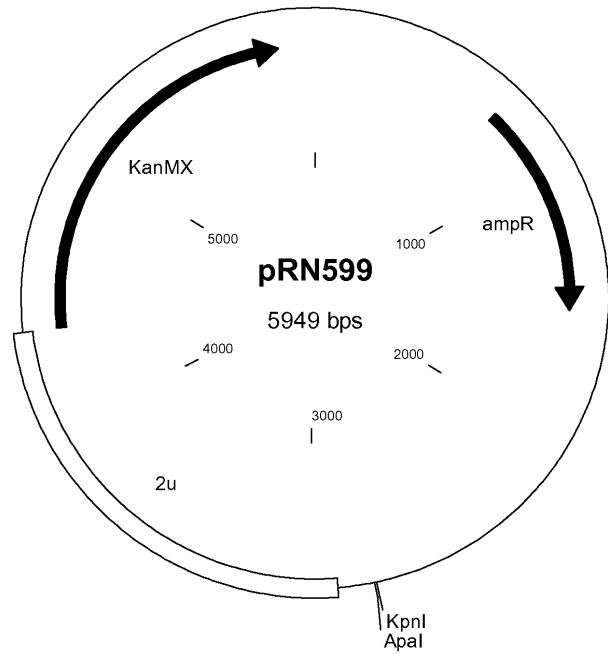
도면5



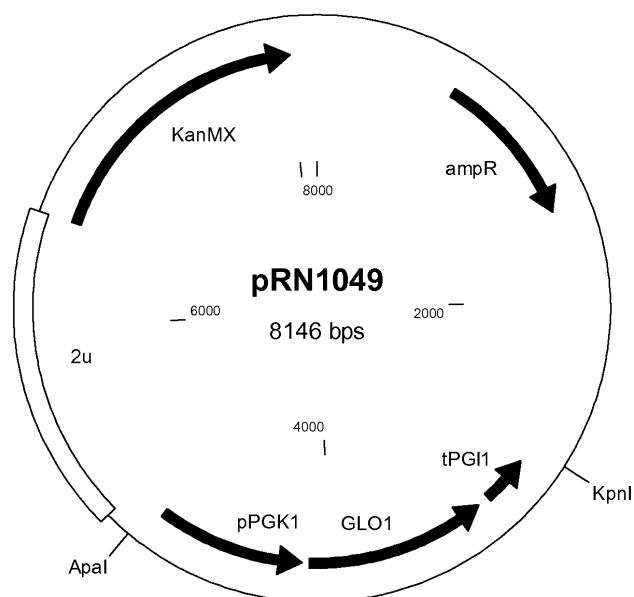
도면6



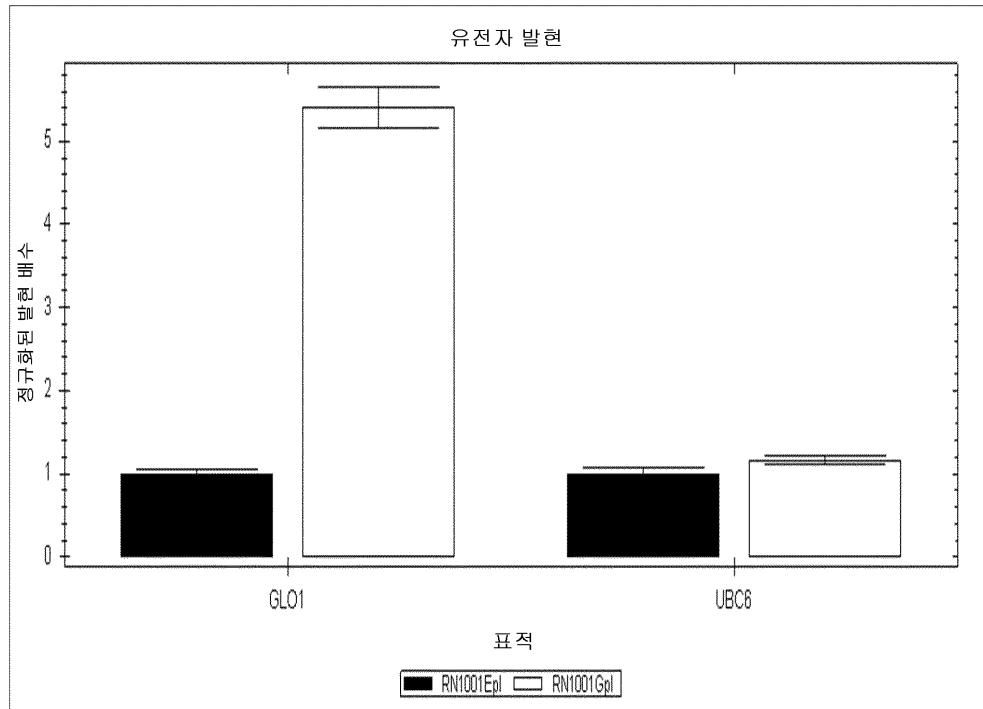
도면7



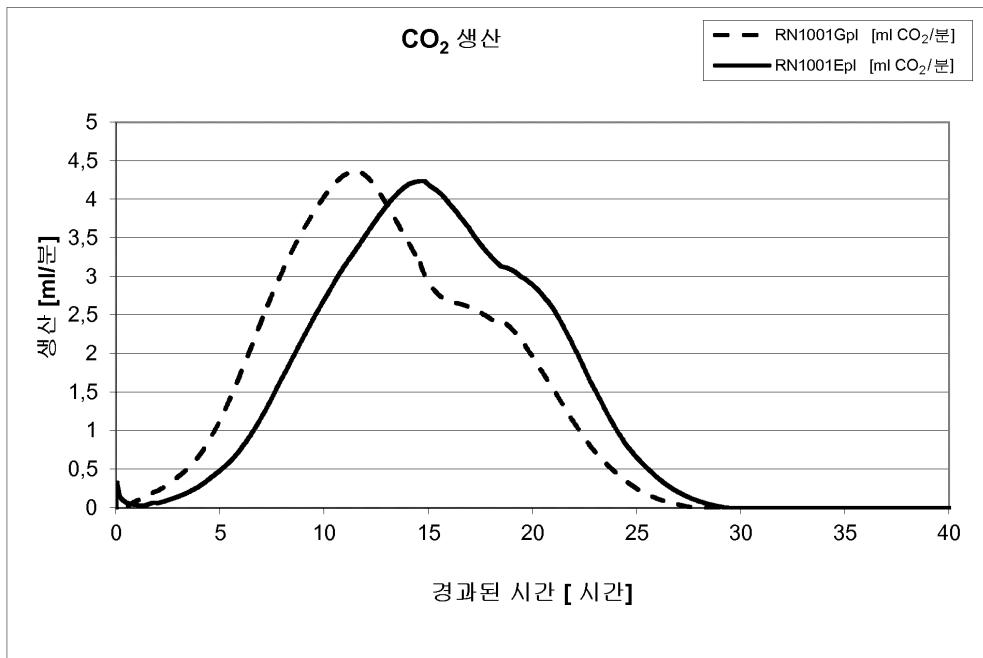
도면8



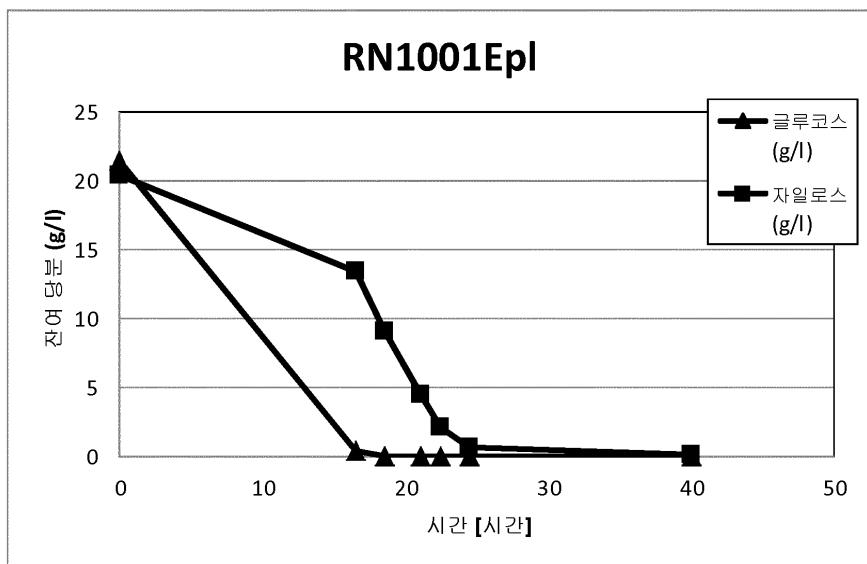
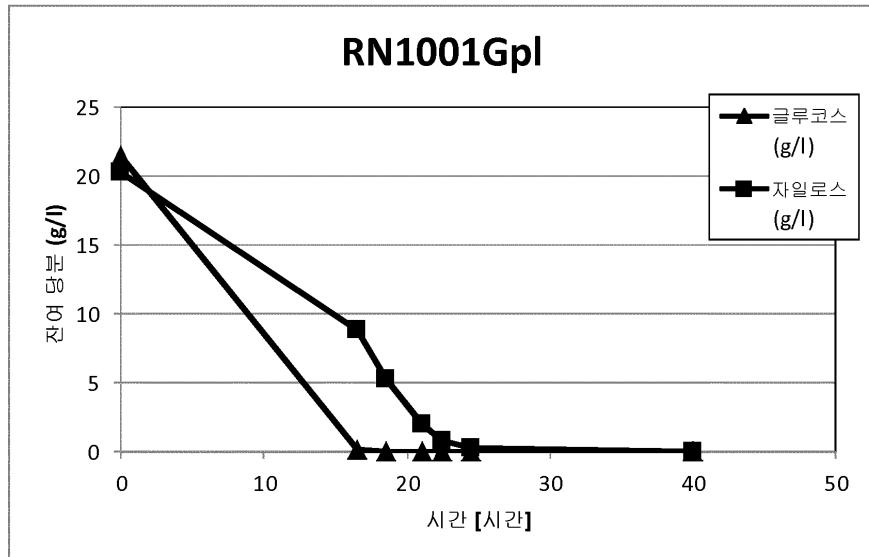
도면9



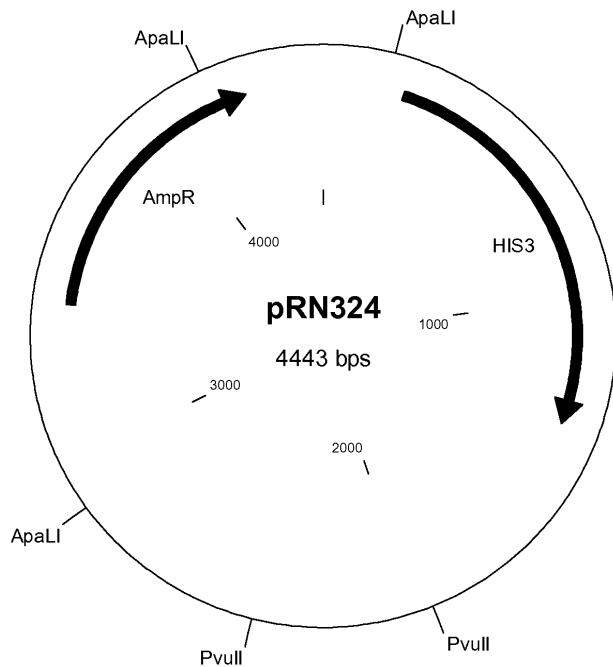
도면10



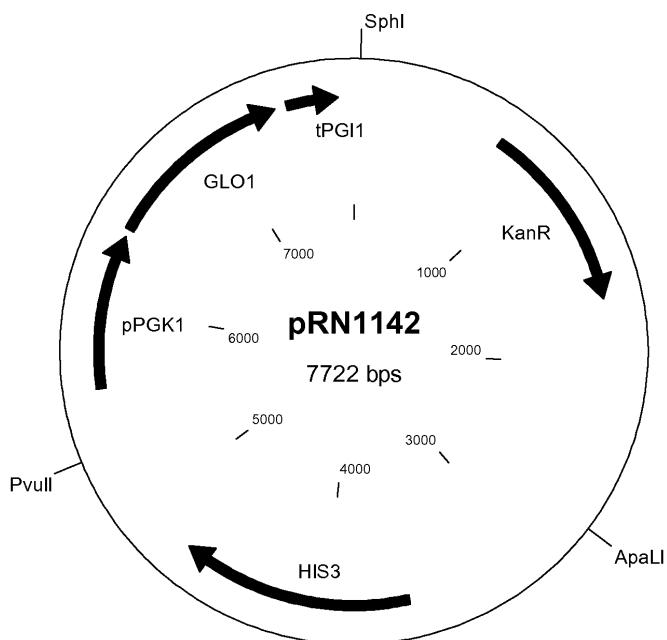
도면11



도면12



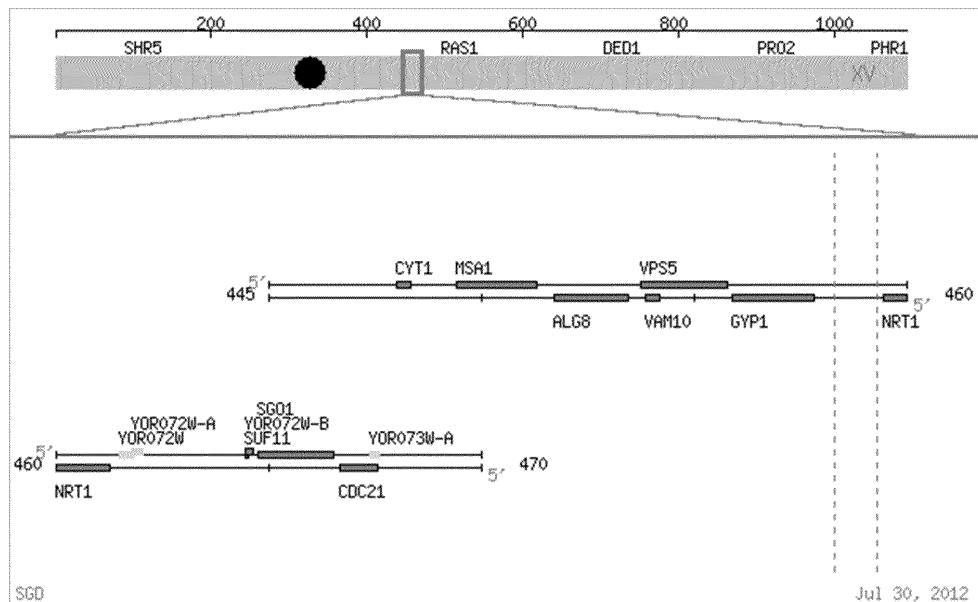
도면13



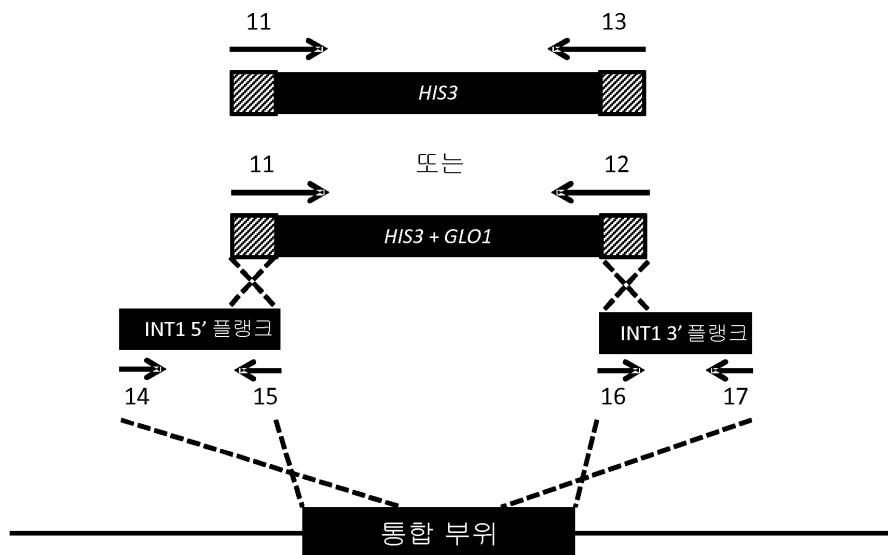
## 도면14

통합 부위 1 (Chr XV, 458319 bp에서 459320 bp로 조정)

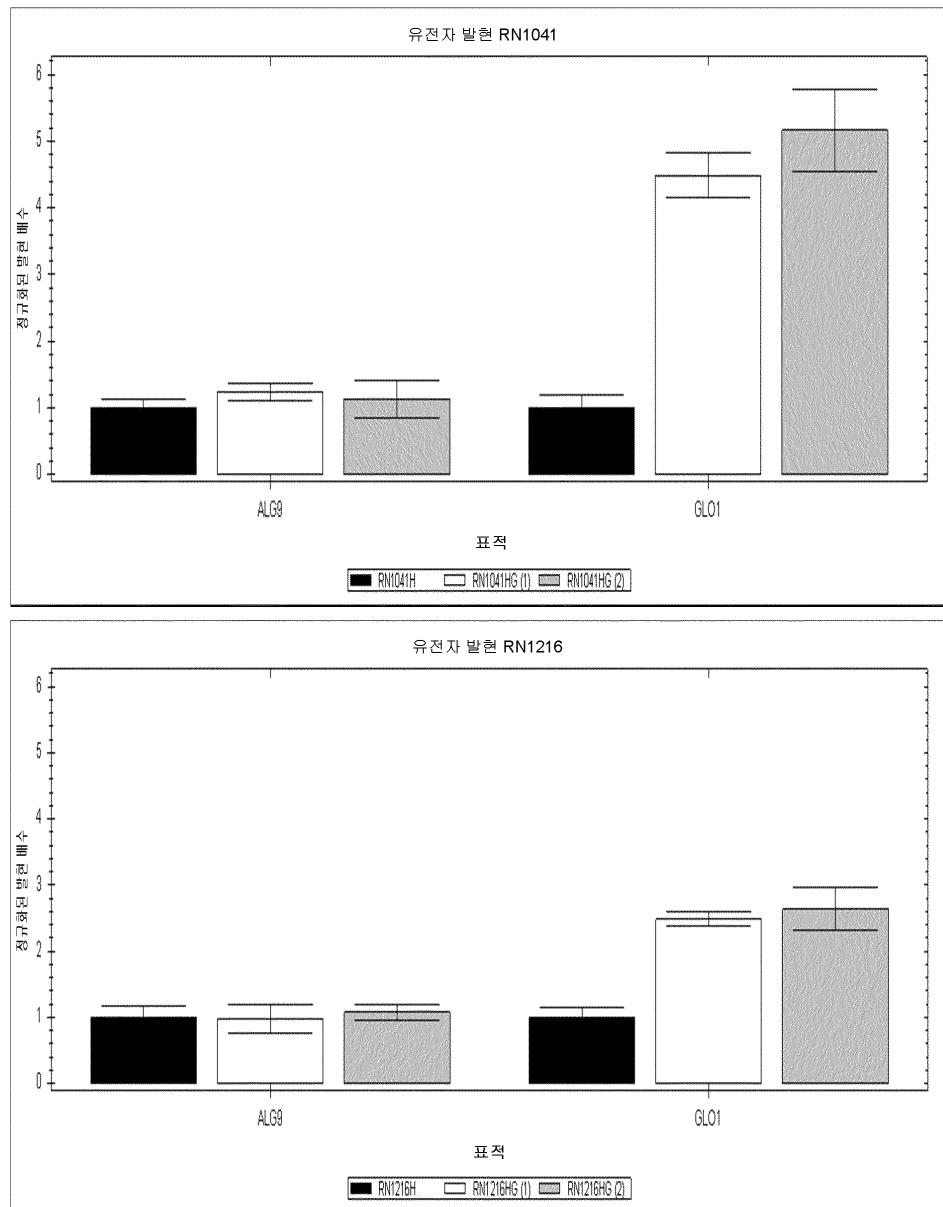
특정화된 지도는 2개의 수직 점선 사이의 영역에 의해 지시됨



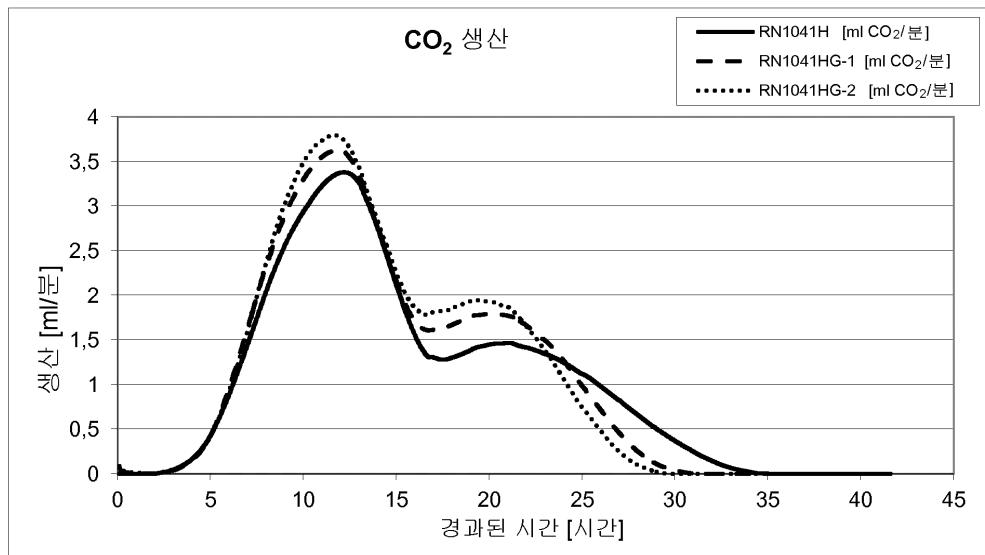
## 도면15



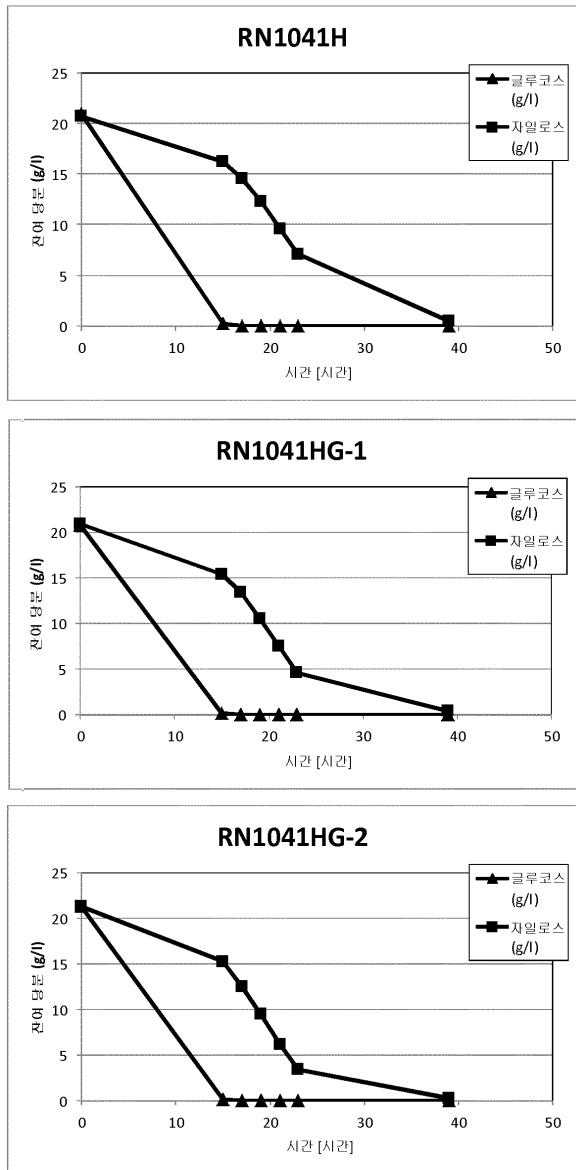
## 도면16



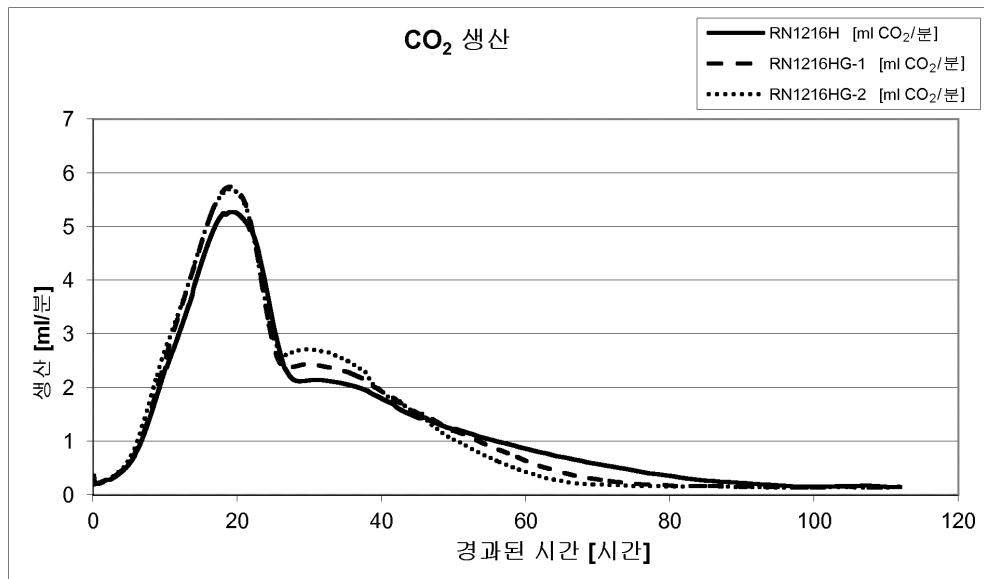
도면17



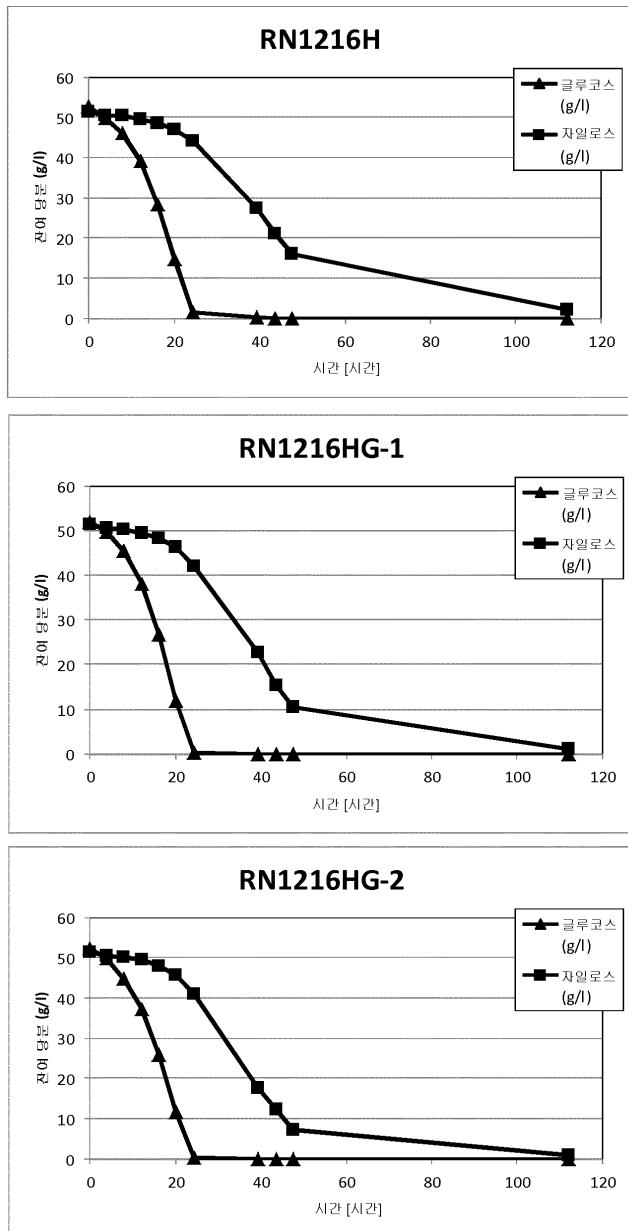
## 도면18



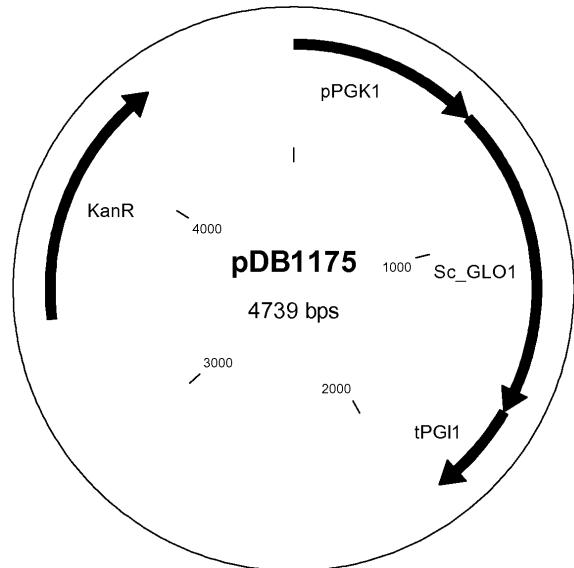
도면19



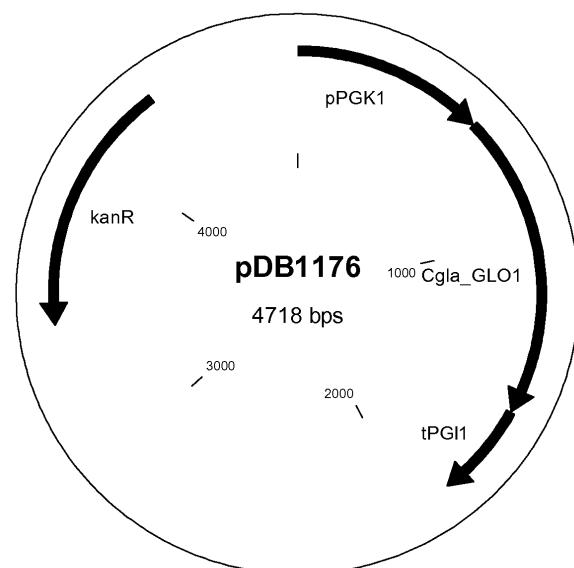
## 도면20



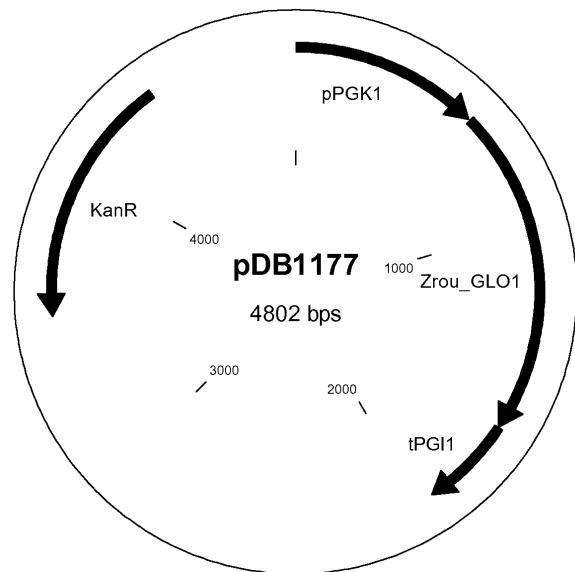
도면21



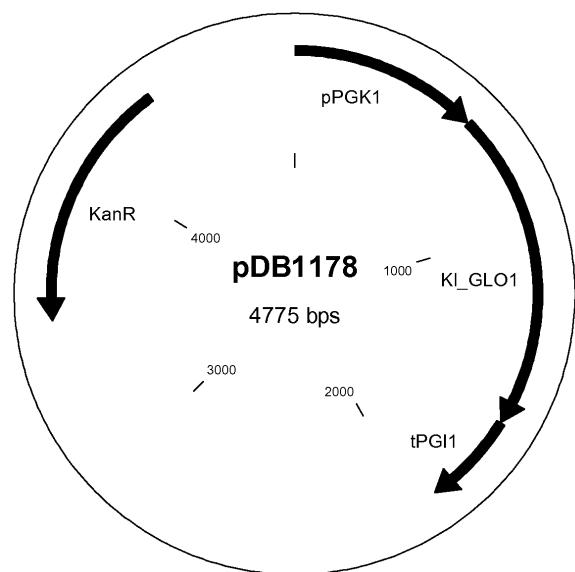
도면22



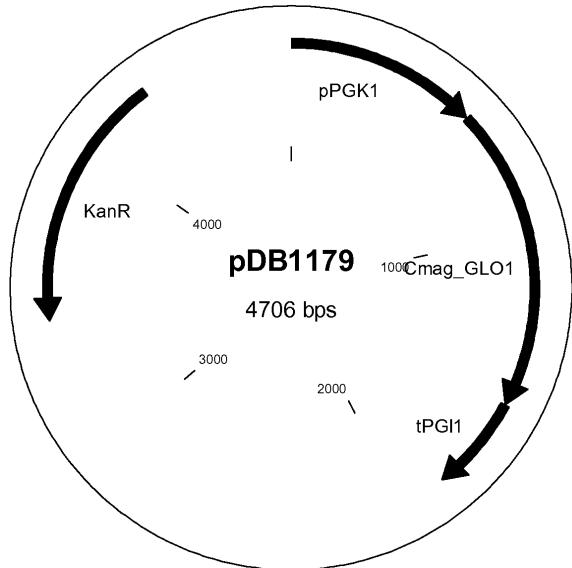
도면23



도면24

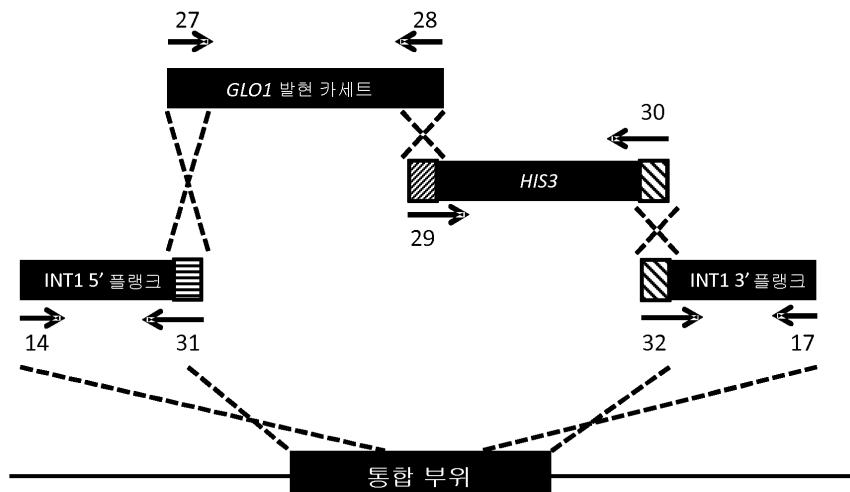


도면25

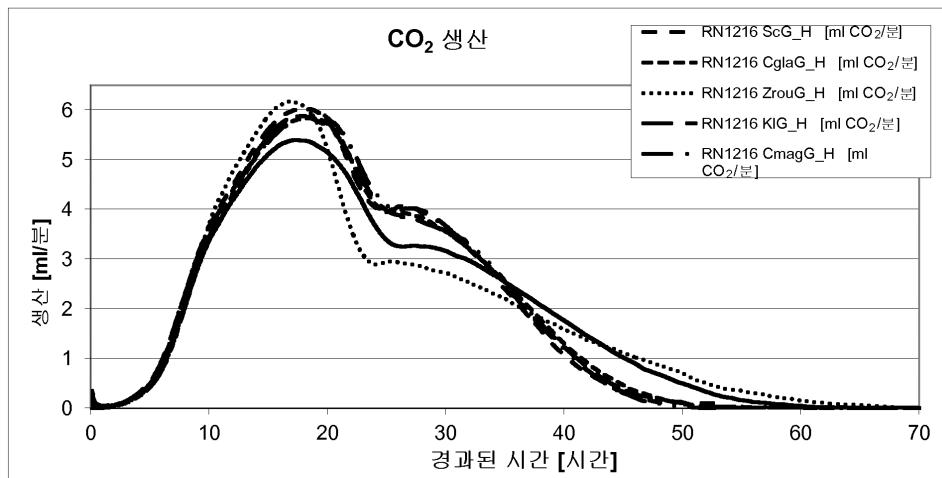


도면26

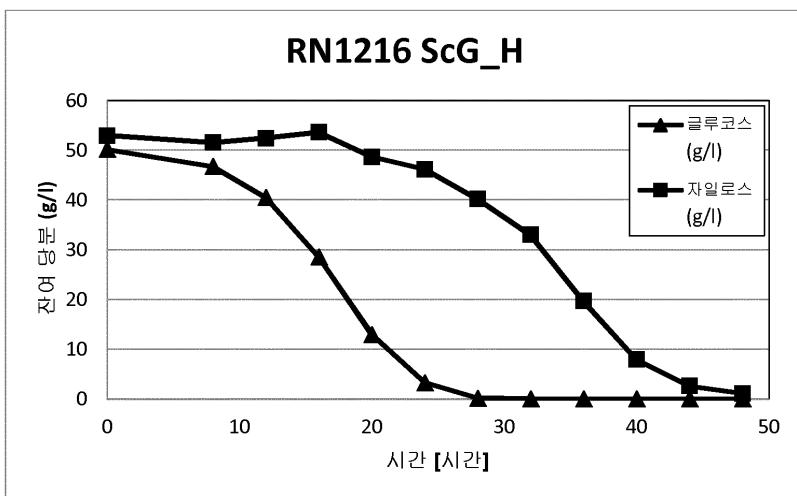
26/30



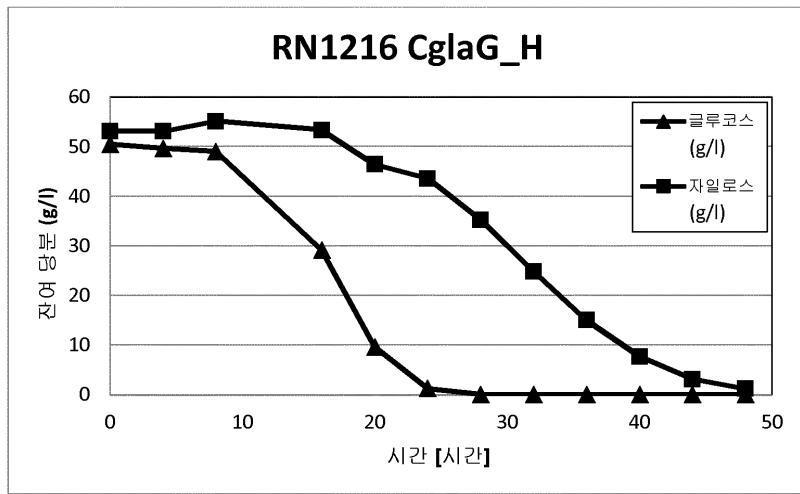
도면27



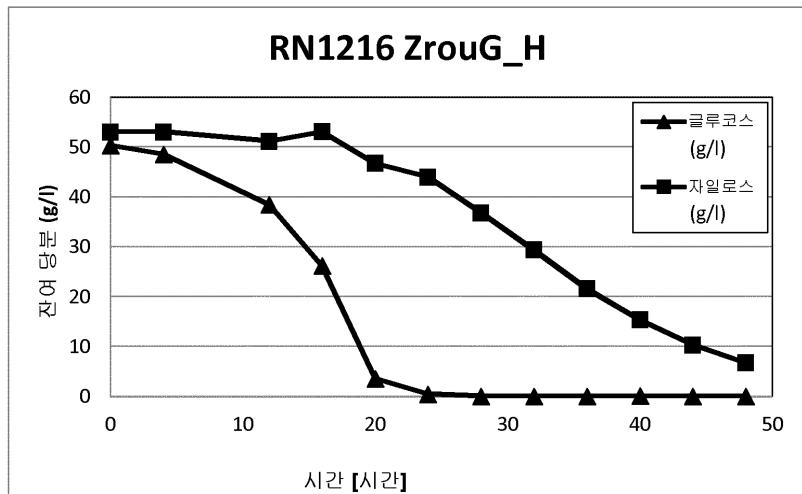
도면28



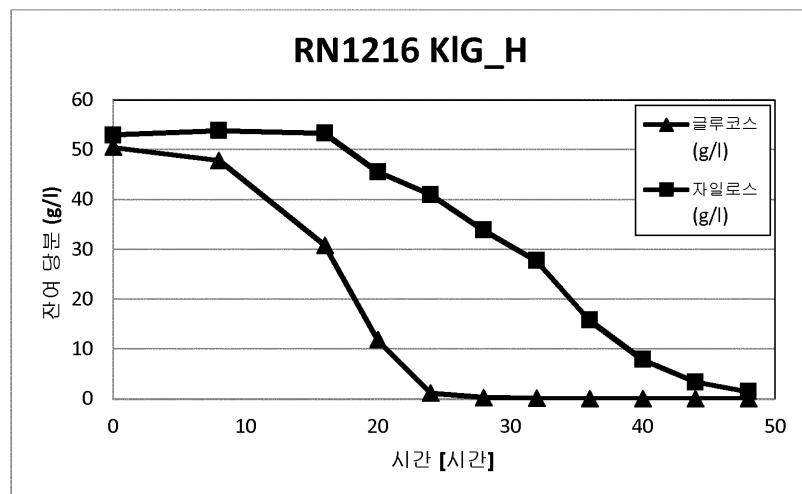
도면29



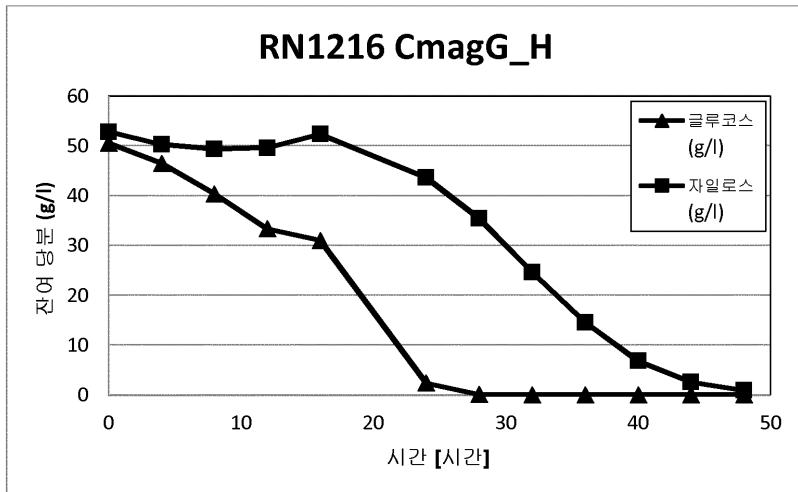
도면30



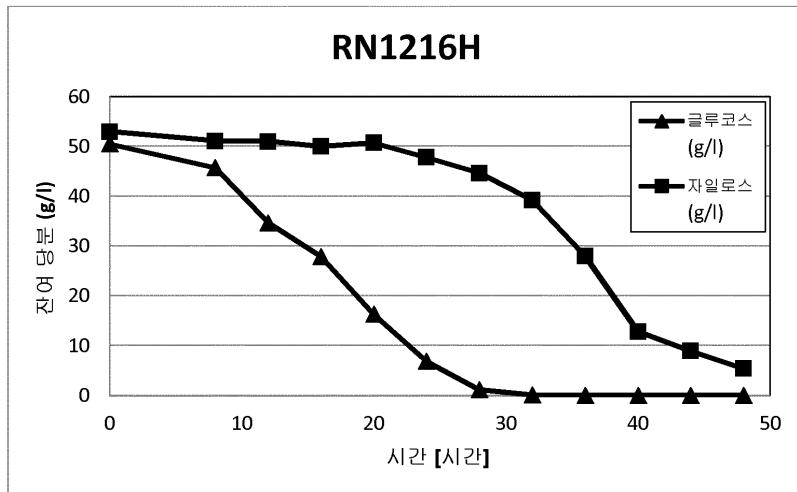
도면31



## 도면32



## 도면33



## 서 열 목록

## SEQUENCE LISTING

- <110> DSM IP Assets B.V.
- <120> CELLS WITH IMPROVED PENTOSE CONVERSION
- <130> 29095-WO-PCT
- <140> PCT/EP2013/071462
- <141> 2013-10-15
- <150> EP 12188715.2
- <151> 2012-10-16
- <150> EP 13166959.0
- <151> 2013-05-08

<160> 38  
 <170> BiSSAP 1.2  
 <210> 1  
 <211> 32  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <221> source  
 <222> 1..32  
 <223> /organism="Artificial Sequence"  
     /note="Forward primer PGK1 promoter"  
     /mol\_type="unassigned DNA"  
 <400> 1

actagtctcg agctcttcaa ctcaagacgc ac 32

<210> 2  
 <211> 36  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <221> source  
 <222> 1..36  
 <223> /organism="Artificial Sequence"  
     /note="Reverse primer PGI1 terminator"  
     /mol\_type="unassigned DNA"

<400> 2  
 ttaagcttcg tacgtttaa acagttgatg agaacc 36  
 <210> 3  
 <211> 26  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <221> source  
 <222> 1..26

<223> /organism="Artificial Sequence"

/note="Forward primer ACT1 gene"  
 /mol\_type="unassigned DNA"  
 <400> 3  
 ccatttttag aatcgattt gccggt 26  
 <210> 4  
 <211> 28  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <221> source  
 <222> 1..28  
 <223> /organism="Artificial Sequence"  
 /note="Reverse primer ACT1 gene"  
 /mol\_type="unassigned DNA"  
 <400> 4  
 ggtgatttcc tttgcattc tttcgcca 28  
 <210> 5  
 <211> 24  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <221> source  
 <222> 1..24  
 <223> /organism="Artificial Sequence"  
 /note="Forward primer GLO1 gene, Q-PCR"  
 /mol\_type="unassigned DNA"  
 <400> 5  
 ggagagcctg atgttttag cgca 24  
 <210> 6  
 <211> 23  
 <212> DNA  
 <213> Artificial sequence

<220>

<221> source

<222> 1..23

<223> /organism="Artificial Sequence"  
   /note="Reverse primer GL01 gene, Q-PCR"  
   /mol\_type="unassigned DNA"

<400> 6

tcaatccagt atccatcagg gcc 23

<210> 7

<211> 29

<212> DNA

<213> Artificial sequence

<220>

<221> source

<222> 1..29

<223> /organism="Artificial Sequence"  
   /note="Forward primer ALG9 gene, Q-PCR"  
   /mol\_type="unassigned DNA"

<400> 7

cacggatagt ggctttggtg aacaattac 29

<210> 8

<211> 32

<212> DNA

<213> Artificial sequence

<220>

<221> source

<222> 1..32

<223> /organism="Artificial Sequence"  
   /note="Reverse primer ALG9 gene, Q-PCR"  
   /mol\_type="unassigned DNA"

<400> 8

tatgattatc tggcagcagg aaagaacttg gg 32

<210> 9

<211> 29

<212> DNA

<213> Artificial sequence

<220>

<221> source

<222> 1..29

<223> /organism="Artificial Sequence"

/note="Forward primer UBC6 gene, Q-PCR"

/mol\_type="unassigned DNA"

<400> 9

gatacttggaa atcctggctg gtctgtctc 29

<210> 10

<211> 32

<212> DNA

<213> Artificial sequence

<220>

<221> source

<222> 1..32

<223> /organism="Artificial Sequence"

/note="Reverse primer UBC6 gene, Q-PCR"

/mol\_type="unassigned DNA"

<400> 10

aaagggtctt ctgtttcatc acctgtatcc gc 32

<210> 11

<211> 75

<212> DNA

<213> Artificial sequence

<220>

<221> source

<222> 1..75

<223> /organism="Artificial Sequence"

/note="Forward primer HIS3 cassette, 5' flank INT1"

/mol\_type="unassigned DNA"

<400> 11

cagtttaaa aagttagaga atgttagaga gtagatggatct ttgaaaccct ttggtgagcg	60
ctaggagtca ctgcc	75

<210> 12	
<211> 75	
<212> DNA	
<213> Artificial sequence	
<220>	
<221> source	
<222> 1..75	
<223> /organism="Artificial Sequence"	
/note="Reverse primer GL01 cassette, 3' flank INT1"	
/mol_type="unassigned DNA"	
<400> 12	
atcttacata gtgtcgaa caggtcattc taaaaaaagt aaaataaaat acgtttaaa	60
cagttgatga gaacc	75
<210> 13	
<211> 75	
<212> DNA	
<213> Artificial sequence	
<220>	
<221> source	
<222> 1..75	
<223> /organism="Artificial Sequence"	
/note="Reverse primer HIS3 cassette, 3' flank INT1"	
/mol_type="unassigned DNA"	
<400> 13	
cataatatgt taaaagctag atcttacata gtgtcgaa caggtcattc ttccacaccgc	60
atatgatccg tcgag	75
<210> 14	
<211> 19	
<212> DNA	
<213> Artificial sequence	
<220>	

```

<221> source
<222> 1..19
<223> /organism="Artificial Sequence"
      /note="Forward primer 5' flank INT1"
      /mol_type="unassigned DNA"

<400> 14
cggcattatt gtgtatggc                                19

<210> 15
<211> 23
<212> DNA
<213> Artificial sequence
<220>
<221> source
<222> 1..23
<223> /organism="Artificial Sequence"
      /note="Reverse primer 5' flank INT1"
      /mol_type="unassigned DNA"

<400> 15
agggtttcaa agatccatac ttc                                23
<210> 16
<211> 24
<212> DNA
<213> Artificial sequence
<220>
<221> source
<222> 1..24
<223> /organism="Artificial Sequence"
      /note="Forward primer 3' flank INT1"
      /mol_type="unassigned DNA"

<400> 16
gaatgacctg ttcccgacac tatg                                24
<210> 17
<211> 24

```

<212> DNA

<213> Artificial sequence

<220>

<221> source

<222> 1..24

<223> /organism="Artificial Sequence"

/note="Reverse primer 3' flank INT1"

/mol\_type="unassigned DNA"

<400> 17

cacaagctta ttcttccaaa aatc	24
----------------------------	----

<210> 18

<211> 500

<212> DNA

<213> *Saccharomyces cerevisiae*

<220>

<221> source

<222> 1..500

<223> /organism="Saccharomyces cerevisiae"

/mol\_type="unassigned DNA"

<400> 18

cgccattatt gtgtatggct caataattt ataaaaaaag gaactattgg ttcttagtat	60
tttcttgcta gaagacatat tcttaccaat ctttcataa gctaattatg ccatccatat	120
agcaagagaa tccgggtggg gcgcctatgcc tatccggcg caacattatt actctggat	180
acgggcgtaa ctccataata tgccaccact taccttaac atgttcatgg tagtacccc	240
acccagccat aaggaaattt tcaaaggcgt tggatcaaaa aataggcctt tatttcatcg	300
cgtgatttag gagcataaca tgtttagtga aggttcttt tgaaaactt cagtcgtca	360
ttattagaac caggaggagtc caggcttgc tggtgggaga gaaagcttat gaagctgggg	420
ttgcagattt gtcgatttgtt cgccagtaca cagtttaaa aagttagaga atgttagagaa	480
gtatggatct ttgaaacctt	500

<210> 19

<211> 478

<212> DNA

<213> *Saccharomyces cerevisiae*

<220>

<221> source

<222> 1..478

<223> /organism="Saccharomyces cerevisiae"  
 /mol\_type="unassigned DNA"

<400> 19

gaatgacctg ttcccgacac tatgtaagat ctagcttta acatattatg gaaacctgaa 60  
 atgtaaaatc tgaattttg tatatgtgtt tatatttggg tagttcttt gagggaaagca 120  
 tgcatacact tgctgtacga actttatgtg actttagtgc acgcgtttc atgagacttt 180

agcccttga acatattatc atatctcagc ttgaaatact atagattac tttgcagcc 240  
 atttcttggt gctccaaggt tgtgcgtatc tattacttaa tttctgtcct tgccaagttt 300  
 tgcagcaggg cggtcacaag actcctctgc cgtcattcct tagccttcg ggaacacact 360  
 tatttatgtt tttgtattct acaattctac ggtgcacaag ggttggcac tggtagctc 420  
 agcacgcac tattgctggc atgaagataa gattgattt tggaagaata agcttgc 478

<210> 20

<211> 326

<212> PRT

<213> Saccharomyces cerevisiae

<400> 20

Met Ser Thr Asp Ser Thr Arg Tyr Pro Ile Gln Ile Glu Lys Ala Ser

1	5	10	15
Asn Asp Pro Thr Leu Leu Asn His Thr Cys Leu Arg Val Lys Asp			
20	25	30	
Pro Ala Arg Thr Val Lys Phe Tyr Thr Glu His Phe Gly Met Lys Leu			
35	40	45	
Leu Ser Arg Lys Asp Phe Glu Ala Lys Phe Ser Leu Tyr Phe Leu			
50	55	60	
Ser Phe Pro Lys Asp Asp Ile Pro Lys Asn Lys Asn Gly Glu Pro Asp			
65	70	75	80
Val Phe Ser Ala His Gly Val Leu Glu Leu Thr His Asn Trp Gly Thr			
85	90	95	
Glu Lys Asn Pro Asp Tyr Lys Ile Asn Asn Gly Asn Glu Glu Pro His			

100	105	110	
Arg Gly Phe Gly His Ile Cys Phe Ser Val Ser Asp Ile Asn Lys Thr			
115	120	125	
Cys Glu Glu Leu Glu Ser Gln Gly Val Lys Phe Lys Lys Arg Leu Ser			
130	135	140	
Glu Gly Arg Gln Lys Asp Ile Ala Phe Ala Leu Gly Pro Asp Gly Tyr			
145	150	155	160
Trp Ile Glu Leu Ile Thr Tyr Ser Arg Glu Gly Gln Glu Tyr Pro Lys			
165	170	175	
Gly Ser Val Gly Asn Lys Phe Asn His Thr Met Ile Arg Ile Lys Asn			
180	185	190	
Pro Thr Arg Ser Leu Glu Phe Tyr Gln Asn Val Leu Gly Met Lys Leu			
195	200	205	
Leu Arg Thr Ser Glu His Glu Ser Ala Lys Phe Thr Leu Tyr Phe Leu			
210	215	220	
Gly Tyr Gly Val Pro Lys Thr Asp Ser Val Phe Ser Cys Glu Ser Val			
225	230	235	240
Leu Glu Leu Thr His Asn Trp Gly Thr Glu Asn Asp Pro Asn Phe His			
245	250	255	
Tyr His Asn Gly Asn Ser Glu Pro Gln Gly Tyr Gly His Ile Cys Ile			
260	265	270	
Ser Cys Asp Asp Ala Gly Ala Leu Cys Lys Glu Ile Glu Val Lys Tyr			
275	280	285	
Gly Asp Lys Ile Gln Trp Ser Pro Lys Phe Asn Gln Gly Arg Met Lys			
290	295	300	
Asn Ile Ala Phe Leu Lys Asp Pro Asp Gly Tyr Ser Ile Glu Val Val			
305	310	315	320
Pro His Gly Leu Ile Ala			
325			

&lt;210&gt; 21

&lt;211&gt; 319

&lt;212&gt; PRT

&lt;213&gt; Candida glabrata

&lt;400&gt; 21

Met Ser Tyr Pro His Lys Ile Ala Ala Ala His Asp Asp Pro Thr Leu

1 5 10 15

Met Phe Asn His Thr Cys Leu Arg Ile Lys Asp Pro Ala Lys Ser Ile

20 25 30

Pro Phe Tyr Gln Lys His Phe Gly Met Glu Leu Leu Asn Lys Leu Asp

35 40 45

Phe Pro Glu Met Lys Phe Ser Leu Phe Phe Leu Ser Phe Pro Lys Asp

50 55 60

Asn Val Ala Lys Asn Ser Glu Gly Lys Asn Asp Val Phe Ser Thr Ser

65 70 75 80

Gly Ile Leu Glu Leu Thr His Asn Trp Gly Ser Glu Asn Asp Ala Asp

85 90 95

Phe Lys Ile Cys Asn Gly Asn Glu Glu Pro His Arg Gly Phe Gly His

100 105 110

Ile Cys Phe Ser Tyr Ala Asp Ile Asn Ala Ala Cys Ser Lys Leu Glu

115 120 125

Ala Glu Gly Val Ser Phe Lys Lys Arg Leu Thr Asp Gly Arg Met Lys

130 135 140

Asp Ile Ala Phe Ala Leu Asp Pro Asp Gly Tyr Trp Ile Glu Leu Ile

145 150 155 160

Arg Tyr Asp Arg Glu Asn Ser Pro Lys Lys Asp Val Gly Ser Arg Phe

165 170 175

Asn His Thr Met Val Arg Val Lys Asp Pro Lys Ala Ser Leu Glu Phe

180 185 190

Tyr Gln Asn Val Leu Gly Met Lys Leu Leu Arg Thr Ser Glu His Glu

195 200 205

Ala Ala Lys Phe Thr Leu Tyr Phe Leu Gly Tyr Lys Val Ser Ser Glu

210 215 220

Asp Asn Glu Phe Ser His Glu Gly Val Leu Glu Leu Thr His Asn Trp

225 230 235 240

Gly Thr Glu Asn Glu Ala Asp Phe Lys Tyr His Asn Gly Asn Asp Lys

245 250 255

Pro Gln Gly Tyr Gly His Ile Cys Val Ser Cys Lys Asp Pro Ala Lys

260 265 270

Leu Cys Asn Glu Ile Glu Gln Thr Tyr Gly Asp Lys Ile Gln Trp Ala

275 280 285

Pro Lys Phe Asn Gln Gly Lys Leu Lys Asn Ile Ala Phe Leu Lys Asp

290 295 300

Pro Asp Gly Tyr Ser Ile Glu Val Val Pro His Gly Leu Ile Val

305 310 315

<210> 22

<211> 347

<212> PRT

<213> Zygosaccharomyces rouxii

<400> 22

Met Phe Ser Arg Val Phe Ser Arg Leu Gly Leu Ile Lys Gln His Ile

1 5 10 15

Arg Thr Met Ser Thr Lys Thr Ser Glu Ala Gln Tyr Tyr Thr Lys Lys

20 25 30

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

35 40 45

Leu Arg Ile Lys Asp Pro Ser Ala Ser Val Glu Phe Tyr Lys His

50 55 60

Phe Asn Met Thr Leu Leu Ser Lys Lys Asp Phe Pro Asp Met Lys Phe

65 70 75 80

Ser Leu Tyr Phe Leu Val Met Thr Lys Glu Asn Leu Pro Lys Asn Glu

85 90 95

Lys Gly Glu Asn Leu Val Phe Ala Asn Arg Gly Ile Leu Glu Leu Thr

100 105 110

His Asn Trp Gly Thr Glu Ala Asp Pro Glu Tyr Lys Val Asn Asn Gly

115	120	125
Asn Val Glu Pro His Arg Gly Phe Gly His Ile Cys Phe Ser Val Ala		
130	135	140
Asn Val Glu Ser Thr Cys Gln Arg Leu Glu Ser Glu Gly Val Lys Phe		
145	150	155
Gln Lys Arg Leu Val Asp Gly Arg Gln Lys Asn Ile Ala Phe Ala Leu		
165	170	175
Asp Pro Asp Gly Tyr Trp Ile Glu Leu Ile Gln Tyr Ile Asn Glu Ser		
180	185	190
Gly Glu Gly Pro Lys Thr Asp Leu Gly Asn Arg Phe Asn His Thr Met		
195	200	205
Val Arg Val Lys Asp Pro Val Lys Ser Leu Glu Phe Tyr Gln Asn Val		
210	215	220
Leu Gly Met Thr Leu His Arg Val Ser Glu His Ala Asn Ala Lys Phe		
225	230	235
Asn Val Glu Pro Asp Gly Tyr Trp Ile Glu Leu Ile Gln Tyr Ile Asn Glu Ser		
240		
Thr Leu Tyr Phe Leu Gly Tyr Asp Ile Pro Gln Gly Asp Ser Thr Gly		
245	250	255
Ser Ala Glu Thr Leu Leu Glu Leu Thr His Asn Trp Gly Thr Glu Asn		
260	265	270
Asp Pro Asp Phe His Tyr His Asn Gly Asn Ala Gln Pro Gln Gly Tyr		
275	280	285
Gly His Ile Cys Ile Thr Cys Lys Asp Pro Gly Ala Leu Cys Glu Glu		
290	295	300
Ile Glu Lys Lys Tyr Asn Glu Gln Val Val Trp Ser Pro Lys Trp Asn		
305	310	315
His Gly Lys Met Lys Asn Leu Ala Phe Ile Lys Asp Pro Asp Gly Tyr		
325	330	335
Ser Ile Glu Ile Val Pro Ala Glu Leu Val Leu		
340	345	
<210> 23		
<211> 338		
<212> PRT		

&lt;213&gt; Kluyveromyces lactis

&lt;400&gt; 23

Met Phe Lys Glu Arg Leu Leu Asn Ile Leu Lys His Ile Arg Pro Met

1 5 10 15

Ser Thr Glu Thr Thr Ala Lys Tyr Tyr Pro Lys Ile Val Glu Ser Ala

20 25 30

Gln Ala Asp Gln Ser Leu Lys Leu Asn His Thr Cys Phe Arg Val Lys

35 40 45

Asp Pro Lys Val Thr Val Ala Phe Tyr Gln Glu Gln Phe Gly Met Lys

50 55 60

Leu Leu Asp His Lys Lys Phe Pro Asp Met Lys Phe Asp Leu Tyr Phe

65 70 75 80

Leu Ser Phe Pro Asn Lys Gln Phe Ser Asn Asn Ser Gln Gly Ala Ile

85 90 95

Asp Val Phe Arg Glu Asn Gly Ile Leu Glu Leu Thr His Asn Tyr Gly

100 105 110

Thr Glu Ser Asp Pro Ala Tyr Lys Val Asn Asn Gly Asn Glu Glu Pro

115 120 125

His Arg Gly Phe Gly His Ile Cys Phe Ser Val Ser Asn Leu Glu Ala

130 135 140

Glu Cys Glu Arg Leu Glu Ser Asn Gly Val Lys Phe Lys Lys Arg Leu

145 150 155 160

Thr Asp Gly Ser Gln Arg Asn Ile Ala Phe Ala Leu Asp Pro Asn Gly

165 170 175

Tyr Trp Ile Glu Leu Ile Gln Asn Asn Glu Ser Gly Glu Gly Asn Asn

180 185 190

Tyr Lys Phe Asn His Thr Met Val Arg Val Lys Asp Pro Ile Lys Ser

195 200 205

Leu Glu Phe Tyr Gln Asn Val Leu Gly Met Lys Ile Leu Asp Val Ser

210 215 220

Asp His Ser Asn Ala Lys Phe Thr Leu Tyr Phe Leu Gly Tyr Glu Asn

225	230	235	240
Asp Gln Lys Gly Ile Ala Arg Gly Ser Arg Glu Ser Ile Leu Glu Leu			
245	250	255	
Thr His Asn Trp Gly Thr Glu Asn Asp Pro Phe Ala Tyr His Thr			
260	265	270	
Gly Asn Thr Glu Pro Gln Gly Tyr His Ile Cys Ile Ser Asn Lys			
275	280	285	
Asp Pro Ala Thr Leu Cys Ala Glu Ile Glu Lys Leu Tyr Pro Asp Ile			
290	295	300	
Gln Trp Ser Pro Lys Phe Asn Gln Gly Lys Met Lys Asn Leu Ala Phe			
305	310	315	320
Ile Lys Asp Pro Asp Gly Tyr Ser Ile Glu Val Val Pro Tyr Gly Leu			
325	330	335	
Gly Val			

<210> 24			
<211> 315			
<212> PRT			
<213> Candida magnoliae			
<400> 24			
Met Leu Gly Lys Val Ala Gln Lys Phe Leu Asn His Thr Cys Ile Arg			
1	5	10	15
Ile Ala Asp Pro Ala Arg Ser Leu Ala Phe Tyr Glu Lys Asn Phe Gly			
20	25	30	
Met Lys Leu Val Thr Gln Leu Asp Val Lys Glu Val Gly Phe Thr Leu			
35	40	45	
Tyr Tyr Leu Gly Phe Thr Gly Pro Lys Ser Leu Tyr Lys Asp Thr Pro			
50	55	60	
Trp Tyr Lys Arg Gly Gly Leu Leu Glu Leu Thr His Asn His Gly Ala			
65	70	75	80
Thr Pro Glu Asn Phe Glu Ala Asn Asn Gly Asn Lys Glu Pro His Arg			
85	90	95	

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

&lt;210&gt; 25

&lt;211&gt; 32

&lt;212&gt; DNA

<213> Artificial sequence

<220>

<221> source

<222> 1..32

<223> /organism="Artificial Sequence"

/note="Forward primer of the GL01 ORF"

/mol\_type="unassigned DNA"

<400> 25

agctgcagaa aatgtccact gatagtacac gc	32
-------------------------------------	----

<210> 26

<211> 35

<212> DNA

<213> Artificial sequence

<220>

<221> source

<222> 1..35

<223> /organism="Artificial Sequence"

/note="Reverse primer of the GL01 ORF"

/mol\_type="unassigned DNA"

<400> 26

aagtgcactt aggcaatcaa accatgagga acgac	35
--	----

<210> 27

<211> 981

<212> DNA

<213> *Saccharomyces cerevisiae*

<220>

<221> source

<222> 1..981

<223> /organism="*Saccharomyces cerevisiae*"

/mol\_type="unassigned DNA"

<400> 27

atgtccactg atagtagacg ctatccaatt cagattgaga aaggcctgaa tgatccaacc	60
cttctgttta atcacacatg tttaagagtc aaggatccag caaggaccgt taagttctac	120
accgaacact tcggtatgaa gctattaagc agaaaggatt ttgaagaagc aaaatttac	180

ttgtacttt taagcttcc aaaagacgac atacccaaa ataagaatgg agagcctgat 240

gttttttagcg cacacggtgt cttagaacta actcacaatt ggggtactga aaaaaaccca 300  
 gactacaaga tcaacaacgg gaatgaggaa cctcatcgtg gattggca catctgttt 360  
 tctgtatccg atatcaataa aacctgcga gagctagaat ctcagggtgt caaatcaag 420  
 aagagactct ctgaaggaag acagaaggac attgcgttg cttaggccc ttagggatac 480  
 tggattgagt ttagtacata ttcttagagag ggtcaggaat acccaaaggc ctcagtaggt 540  
 aacaagttaa atcataaccat gattcgatt aaaaaccca cccggcttt agaattctac 600  
 cagaatgtgt tggcatgaa attattaaga actagtgagc acgaaagtgc aaaatttacg 660

ttatacttc ttggttatgg cttccaaag accgacagcg ttttcatg taaaagtgt 720  
 ttggagttaa ctcataattg gggactgag aatgatccaa acttccacta tcataacgg 780  
 aactcagagc cccagggtta tggcacatc tgcataagtt gtatgacgc tggccctt 840  
 tggaaagaaa ttgaagtgaa atacggcgat aagatccaaat ggttcctaa attaacc 900  
 ggcagaatga agaatattgc cttttgaag gatcctgatg gttattccat tgaagtgc 960  
 ctcatggtt tggatgccta a 981

<210> 28

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<222> 1..22

<223> /organism="Artificial Sequence"

/note="Forward primer homologous GL01 expression cassettes"

/mol\_type="unassigned DNA"

<400> 28

aagcgacttc caatcgctt gc 22

<210> 29

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<222> 1..21  
<223> /organism="Artificial Sequence"  
/note="Reverse primer homologous GL01 expression cassettes"  
/mol\_type="unassigned DNA"  
  
<400  
> 29  
  
aaagcaaagg aaggagagaa c 21  
  
<210> 30  
  
<211> 70  
  
<212> DNA  
  
<213> Artificial Sequence  
  
<220>  
  
<221> source  
  
<222> 1..70  
  
<223> /organism="Artificial Sequence"  
/note="Forward primer HIS3 cassette, homologous GL01 expression  
cassettes"  
/mol\_type="unassigned DNA"  
  
<400> 30  
  
ttgccccatcg aacgtacaag tactccctg ttctctcctt ccttgcttt taactatgcg 60  
gcatcagagc 70  
  
<210>  
31  
  
<211> 71  
  
<212> DNA  
  
<213> Artificial Sequence  
  
<220>  
  
<221> source  
  
<222> 1..71  
  
<223> /organism="Artificial Sequence"  
/note="Reverse primer HIS3 cassette, 3a? flank INT1"  
/mol\_type="unassigned DNA"  
  
<400> 31  
  
acttagtatg gtctgttgaa aaggattgtg gcttcgcata caggcttct tcctgatgcg 60  
gtatttctc c 71

<210> 32

<211> 73

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<222> 1..73

<223> /organism="Artificial Sequence"

/note="Reverse primer INT1 5a? flank, homologous GL01 expression

cassettes"

/mol\_type="unassigned DNA"

<400> 32

aaacgcctgt gggtgtggta ctggatatgc aaagcgattg gaagtcgctt agggtttcaa 60

agatccatac ttc 73

<210> 33

<211> 74

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<222> 1..74

<223> /organism="Artificial Sequence"

/note="Forward primer INT1 3a? flank, HIS3 cassette"

/mol\_type="unassigned DNA"

<400> 33

agaaagcctg tatgcgaagc cacaatcctt tccaaacagac catactaagt gaatgacctg 60

ttcccgacac tatg 74

<210> 34

<211> 981

<212> DNA

<213> Artificial Sequence

<220>

<221> source

&lt;222&gt; 1..981

&lt;223&gt; /organism="Artificial Sequence"

/note="GL01 S. cerevisiae codon-pair optimized sequence"

/mol\_type="unassigned DNA"

&lt;400&gt; 34

atgtccactg actctaccag atacccaatt caaatgaaa aggctccaa cgacccaact	60
ctattattga accacacccgt tttgagagtc aaggacccag ctagaactgt caaattctac	120

actgaacact tcggtatgaa attgttgtcc agaaaggact ttgaagaagc taagttctct	180
ttgtacttct tgtctttccc aaaggatgac attccaaaga acaagaacgg tgaaccagat	240
gtttctctg ctcacgggtgt ctggaaatta acccacaact ggggcactga aaagaaccca	300
gactacaaga tcaacaacgg taacgaagaa cctcaccgtg gttcggtca tatctgttcc	360
tctgtttctg acatcaacaa gacctgtgaa gaatttggaaat ctcaagggtgt caagtcaag	420
aagagattat ctgaaggctcg tcaaaggat atcgcccttg ctggggtcc agatggttac	480
tggattgaat tgatcaccta ctccagagaa ggtcaagaat acccaaggg ttccgttggt	540

aacaaattca accacaccat gatcagaatc aagaacccaa ccagatctt ggaattctac	600
caaaacgtt tgggtatgaa gtgttgaga acttctgaac acgaatctgc caagtctact	660
ttatacttct tgggttacgg tggccaaag accgattctg tctttctg tgaatccgtt	720
ttggaaattga cccataactg gggtaactgaa aatgacccaa acttccacta ccacaatgg	780
aactctgaac ctcaagggtta cggtcacatc tgtatctt gtgtatgtgc tgggtcttg	840
tgtaaaggaaa ttgaagtcaa atacggtgac aagattcaat ggtccccaaa gttcaaccaa	900
ggtagaatga agaacattgc ttcttgaaa gaccagacg gttactccat cgaagttgtt	960

ccacacgggt tgatcgctta a	981
-------------------------	-----

&lt;210&gt; 35

&lt;211&gt; 960

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; source

&lt;222&gt; 1..960

&lt;223&gt; /organism="Artificial Sequence"

/note="GL01 C. glabrata codon-pair optimized sequence"

/mol\_type="unassigned DNA"

&lt;400&gt; 35

atgtcttacc cacacaagat tgctgctgcc cacatgacc caacttgat gttcaaccac	60
acctgttga gaatcaagga cccagctaag tccattccat ttaccaaaa gcattcggt	120
atgaaattgt tgaacaaatt ggacttccca gaaatgaaat ttccttgc ttcttatct	180

ttcccaaagg acaacgttgc caagaactct gaaggtaaga acgtgttctt ctccacttct	240
ggtatcttgg aattgaccca caactgggt tctgaaaacg atgtgattt caagatctgt	300
aacggtaacg aagaaccaca ccgtggttc ggtcacatct gttcttta cgctgatatc	360
aacgctgctt gttccaaatt ggaagctgaa ggtgttctt tcaagaagag attgaccgat	420
ggtagaatga agatatcgc ctttgcttg gaccagacg gttactggat cgaattaatc	480
agatacgaca gagaaaactc tccaaagaag gacgtcggtt ccagattcaa ccataccatg	540
gtcagagtta aggacccaaa ggcttcttg gaattcttacc aaaacgttct aggtatgaaa	600

ttgttgagaa cctctgaaca cgaagctgcc aagttcactt tgtaattctt aggttacaag	660
gttcctctg aagacaacga attctctcac gaagggtct tagaattgac tcataactgg	720
ggtaactgaaa atgaagctga cttaataac cacaacggta atgacaagcc acaagggtac	780
ggtcacattt gtgtctctg taaggatcca gccaagttgt gtaacgaaat tgaacaacc	840
tacggtgaca agatccaatg ggctccaaag ttcaaccaag gtaaattgaa gaacattgct	900
ttcttgaagg accctgtatgg ttactccatc gaagttgttc cacacggttt gattgtataa	960

&lt;210&gt; 36

&lt;211&gt; 948

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; source

&lt;222&gt; 1..948

&lt;223&gt; /organism="Artificial Sequence"

/note="GL01 C. magnoliae codon-pair optimized sequence"

/mol\_type="unassigned DNA"

&lt;400&gt; 36

atgttggta aggttgctca aaagttcttg aaccacacct gtatcagaat tgctgaccca	60
gctcgttctt tggcttcta cggaaagaac ttggatga aattggttac tcaattggat	120
gtcaaggaag tcggttacac ttgtactac ttgggttca cgggtccaaa gtcttatac	180
aaggacactc catggtacaa gagaggtggt ctatggaaat tgactcacaa ccatggtgcc	240

actccagaaa actttgaagc taacaatggt aacaaggaac cacacagagg tttggtcac	300
atctgttctt ctgtttctga ttggaaaag acctgtaaa aattggagg taacgggtc ggtttccaaa agagattgac cgatggtaga caaaagaaca ttgccttgc ttggaccca gacggttact ggattgaatt gatcagaac ggttaacgaag gtgtgaatc tccagaaacc tgttaccacca gattcaacca ctccatgatc agagtttaagg acaaggacgc tgcttggat ttctacacca acaaattggg tatgactttg gttgacacct ctgacttccc agaagccaag ttcactttat tttcttgc ttgcaccca actccgtca agggaaatgc cagaggtggt actgaagggtt taatcgaatt gacctacaac tacgggtctg aacaagatgt caacttgcac	360 420 480 540 600 660 720
taccacaacg gtaacactga cccacaagg ttcggtcact tcgggtcac cggtccagat gctaaggcct tcttgcga attggaatcc aagggtgtcc gtgtcacca gcaattgact gaaggtaaga tgaatttcat gggttcgtt tctgacccag atggttactt gattgaagtt ttgcctcaaa gagatttccc aaaggactta ttcttccat cgctctaa	780 840 900 948
<210> 37	
<211> 1017	
<212> DNA	
<213> Artificial Sequence	
<220>	
<221> source	
<222> 1..1017	
<223> /organism="Artificial Sequence" /note="GL01 K. lactis codon-pair optimized sequence" /mol_type="unassigned DNA"	
<400> 37	
atgttcaagg aaagatttattt gaacatctt aagcacatca gaccaatgtc tactgaaacc accgccaagt actacccaaa gattgtcgaa tctgctcaag ctgaccaatc ttgaaattt aaccataacct gttcagagt caaggatcca aaggttaccc ttgtttctt ccaagaacaa tttggatgtt aattgttggc ccacaagaag ttccagaca tgaagttcga ctgttactt ttgtctttcc caaacaagca attctccaac aactctcaag gtgcatttga tttttcaga gaaaacggta tcttggattt gacccacaac tacggtaactg aatctgaccc agtttacaag gtcaacaacg gtaacgaaga accacacccgt gtttggcgtt acatctgtttt ctccgttcc	60 120 180 240 300 360 420
aacttggaaag ctgaatgtga aagattggaa tccaaacgggtg tcaagttcaa gaagagattt	480

accgatggtt ctcaaagaaa cattgcttgc gcttggacc caaacggta ctggattgaa	540
ttgatccaaa acaacgaatc tggtaaggt aacaactaca aattcaacca caccatggtt	600
cgtgttaagg acccaatcaa atcttgaa ttctacccaa acgtcttggg tatgaagatc	660
ttggatgtct ctgaccattc caatgctaag ttcaacttat acttcttggg ttacgaaaat	720
gacccaaaagg gtattgccag aggttccaga gaatccatct tagaattgac tcacaactgg	780
ggtaactgaaa acgaccaga ttttgcttac cacaccgta acactgaacc tcaaggttac	840
ggtcacattt gtatctctaa caaagaccca gctactctat gtgctgaaat tgaaaagttg	900
tacccagata tccaatggtc cccaaagttc aaccaaggta agatgaagaa ctggccctt	960
atcaaggacc cagatggta ctctattgaa gttgttccat acggtttagg ggtgtaa	1017
<210> 38	
<211> 1044	
<212> DNA	
<213> Artificial Sequence	
<220>	
<221> source	
<222> 1..1044	
<223> /organism="Artificial Sequence"	
/note="GL01 Z. rouxii codon-pair optimized sequence"	
/mol_type="unassigned DNA"	
<400> 38	
atgttctccc gtgtttctc cagatttagt ttgatcaagc aacacatcag aaccatgtcc	60
accaggactt ctgaagccca atactacacc aagaagattt cttctgctgt cgggtgaccca	120
tctttgcgtt tcaaccacac ctgtttgaga atcaaggacc catctgcttc cgtcgaattc	180
tacaagaagc acttcaacat gactttttt tccaagaagg acttccaga catgaaattc	240
tctctatact ttttggttat gaccaaggaa aacttgccaa agaacgaaaa gggtgaaaac	300
ttggttttt ttaacacaggat tatcttgaa ttaacccaca actggggcac tgaagctgac	360
ccagaataaca aggtcaacaa tggtaacgtt gaaacctaca gaggtttcgg tcacatttg	420
ttctctgttg ctaacgtcga atccacttgt caaagattgg aatctgaagg tggtaagttc	480
caaaagagat tagtcgatgg tagacaaaag aacattgcct ttgttttggg cccagatgg	540
tactggatcg aattgattca atacatcaac gaatctggtg aaggccaaa gaccgatttgc	600
ggtaacagat tcaaccatac catggtaga gtcaaggatc cagttaaatc ttggaaattc	660
tacccaaaatg tcttaggtat gactttgcac agagttctg aacacgctaa cgccaaattc	720

actttgtact tcttgggta cgatatccca caaggtgact ctactggttc tgctgaaact	780
ttgttggaat tgacccataa ctggggtaact gaaaacgacc cagatttcca ctaccacaac	840
ggtaacgccc aaccacaagg ttacggtcac atctgtatca cctgtaagga cccaggtgct	900
ttgtgtgaag aaattgaaaa gaaatacaac gaacaagttg tctggtcccc aaaatggaac	960
cacggtaaga tgaagaactt ggctttcatc aaagatccag acggttactc cattgaaatt	1020
gttccagctg aattggtctt ataa	1044