



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2021년10월18일  
(11) 등록번호 10-2314885  
(24) 등록일자 2021년10월13일

(51) 국제특허분류(Int. Cl.)  
C07K 14/34 (2006.01) C12N 15/77 (2006.01)  
C12P 13/08 (2006.01)  
(52) CPC특허분류  
C07K 14/34 (2013.01)  
C12N 15/77 (2013.01)  
(21) 출원번호 10-2021-0047408  
(22) 출원일자 2021년04월12일  
심사청구일자 2021년04월12일  
(56) 선행기술조사문헌  
KR1020150000773 A  
KR1020090077738 A

(73) 특허권자  
씨제이제일제당 (주)  
서울특별시 중구 동호로 330 (쌍림동)  
(72) 발명자  
장진숙  
서울특별시 중구 동호로 330(쌍림동)  
이한형  
서울특별시 중구 동호로 330(쌍림동)  
(뒤편에 계속)  
(74) 대리인  
유미특허법인

전체 청구항 수 : 총 5 항

심사관 : 문명순

(54) 발명의 명칭 **신규한 단백질 변이체 및 이를 이용한 L-라이신 생산 방법**

**(57) 요약**

본 출원은 신규한 단백질 변이체, 상기 변이체를 포함하는 코리네박테리움 글루타미쿰 균주 및 상기 균주를 이용한 L-라이신 생산 방법에 관한 것이다.

(52) CPC특허분류  
*C12P 13/08* (2013.01)

(72) 발명자  
**김혜미**  
서울특별시 중구 동호로 330(쌍림동)

**박소정**  
서울특별시 중구 동호로 330(쌍림동)

**김병수**  
서울특별시 중구 동호로 330(쌍림동)

**명세서**

**청구범위**

**청구항 1**

서열번호 3의 676번째 위치에 상응하는 아미노산인 아스파르트산이 타이로신으로 치환된, 서열번호 1로 기재된 아미노산 서열로 이루어진 단백질 변이체.

**청구항 2**

제1항의 변이체를 코딩하는 폴리뉴클레오티드.

**청구항 3**

제1항의 변이체 또는 상기 변이체를 코딩하는 폴리뉴클레오티드를 포함하는, 코리네박테리움 글루타미쿰 균주.

**청구항 4**

제3항에 있어서, 상기 균주는 서열번호 3의 폴리펩티드 또는 이를 코딩하는 폴리뉴클레오티드를 포함하는 코리네박테리움 글루타미쿰과 비교하여 L-라이신 생산능이 증가된, 균주.

**청구항 5**

제1항의 변이체 또는 상기 변이체를 코딩하는 폴리뉴클레오티드를 포함하는 코리네박테리움 글루타미쿰 균주를 배지에서 배양하는 단계를 포함하는, L-라이신 생산 방법.

**발명의 설명**

**기술분야**

[0001] 본 출원은 신규한 단백질 변이체, 상기 변이체를 포함하는 코리네박테리움 글루타미쿰 균주 및 상기 균주를 이용한 L-라이신 생산 방법에 관한 것이다.

**배경기술**

[0003] L-아미노산 및 기타 유용물질을 생산하기 위하여, 고효율 생산 미생물 및 발효공정기술 개발을 위한 다양한 연구들이 수행되고 있다. 예를 들어, L-라이신 생합성에 관여하는 효소를 코딩하는 유전자의 발현을 증가시키거나 또는 생합성에 불필요한 유전자를 제거하는 것과 같은 목적 물질 특이적 접근 방법이 주로 이용되고 있다 (W02008-082181 A1).

[0004] 다만, L-라이신의 수요 증가에 따라 효과적인 L-라이신의 생산능 증가를 위한 연구가 여전히 필요한 실정이다.

**발명의 내용**

**해결하려는 과제**

[0006] 본 출원의 하나의 목적은 서열번호 3의 아미노산 서열의 676번째 위치에 상응하는 아미노산인 아스파르트산 (Aspartic acid, Asp, 또는 D)이 타이로신(Tyrosine, Tyr, 또는 Y)으로 치환된, 서열번호 1로 기재된 아미노산 서열로 이루어진, 단백질 변이체를 제공하는 것이다.

[0007] 본 출원의 다른 하나의 목적은 본 출원의 변이체를 코딩하는 폴리뉴클레오티드를 제공하는 것이다.

[0008] 본 출원의 또 다른 하나의 목적은 본 출원의 변이체 또는 상기 변이체를 코딩하는 폴리뉴클레오티드를

포함하고, L-라이신 생산능을 가진, 코리네박테리움 글루타미쿰(*Corynebacterium glutamicum*) 균주를 제공하는 것이다.

[0009] 본 출원의 또 다른 하나의 목적은 변이체 또는 상기 변이체를 코딩하는 폴리뉴클레오티드를 포함하고, L-라이신 생산능을 가진, 코리네박테리움 글루타미쿰 균주를 배지에서 배양하는 단계를 포함하는, L-라이신 생산 방법을 제공하는 것이다.

**과제의 해결 수단**

[0011] 이를 구체적으로 설명하면 다음과 같다. 한편, 본 출원에서 개시된 각각의 설명 및 실시형태는 각각의 다른 설명 및 실시 형태에도 적용될 수 있다. 즉, 본 출원에서 개시된 다양한 요소들의 모든 조합이 본 출원의 범주에 속한다. 또한, 하기 기술된 구체적인 서술에 의하여 본 출원의 범주가 제한된다고 볼 수 없다. 또한, 본 명세서 전체에 걸쳐 다수의 논문 및 특허문헌이 참조되고 그 인용이 표시되어 있다. 인용된 논문 및 특허문헌의 개시 내용은 그 전체로서 본 명세서에 참조로 삽입되어 본 발명이 속하는 기술 분야의 수준 및 본 발명의 내용이 보다 명확하게 설명된다.

[0013] 본 출원의 하나의 양태는 서열번호 3의 아미노산 서열의 676번째 위치에 상응하는 아미노산인 아스파르트산(Aspartic acid, Asp, 또는 D)이 타이로신(Tyrosine, Tyr, 또는 Y)으로 치환된, 서열번호 1로 기재된 아미노산 서열로 이루어진, 단백질 변이체를 제공한다.

[0014] 본 출원의 변이체는 서열번호 1로 기재된 아미노산 서열을 가지거나 포함하거나, 상기 아미노산 서열로 필수적으로 이루어질(essentially consisting of) 수 있다.

[0016] 또한, 본 출원의 변이체는 상기 서열번호 1로 기재된 아미노산 서열에서 서열번호 3의 아미노산 서열을 기준으로 676번 위치에 상응하는 아미노산은 타이로신(Tyrosine)이고, 상기 서열번호 1로 기재된 아미노산 서열과 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.7% 또는 99.9% 이상의 상동성 또는 동일성을 가지는 아미노산 서열을 포함할 수 있다. 또한, 이러한 상동성 또는 동일성을 가지며 본 출원의 변이체에 상응하는 효능을 나타내는 아미노산 서열이라면, 일부 서열이 결실, 변형, 치환, 보존적 치환 또는 부가된 아미노산 서열을 갖는 변이체도 본 출원의 범위 내에 포함됨은 자명하다.

[0017] 예를 들어, 상기 아미노산 서열 N-말단, C-말단 그리고/또는 내부에 본 출원의 변이체의 기능을 변경하지 않는 서열 추가 또는 결실, 자연적으로 발생할 수 있는 돌연변이, 잠재성 돌연변이(silent mutation) 또는 보존적 치환을 가지는 경우이다.

[0018] 상기 “보존적 치환(conservative substitution)”은 한 아미노산을 유사한 구조적 및/또는 화학적 성질을 갖는 또 다른 아미노산으로 치환시키는 것을 의미한다. 이러한 아미노산 치환은 일반적으로 잔기의 극성, 전하, 용해도, 소수성, 친수성 및/또는 양친매성(amphipathic nature)에서의 유사성에 근거하여 발생할 수 있다. 통상적으로, 보존적 치환은 단백질 또는 폴리펩티드의 활성에 거의 영향을 미치지 않거나 또는 영향을 미치지 않을 수 있다.

[0019] 본 출원에서 용어, “변이체(variant)”는 하나 이상의 아미노산이 보존적 치환(conservative substitution) 및/또는 변형(modification)되어 상기 변이체의 변이 전 아미노산 서열과 상이하나 기능(functions) 또는 특성(properties)이 유지되는 폴리펩티드를 지칭한다. 이러한 변이체는 일반적으로 상기 폴리펩티드의 아미노산 서열 중 하나 이상의 아미노산을 변형하고, 상기 변형된 폴리펩티드의 특성을 평가하여 동정(identify)될 수 있다. 즉, 변이체의 능력은 변이 전 폴리펩티드에 비하여 증가되거나, 변하지 않거나, 또는 감소될 수 있다. 또한, 일부 변이체는 N-말단 리더 서열 또는 막전이 도메인(transmembrane domain)과 같은 하나 이상의 부분이 제거된 변이체를 포함할 수 있다. 다른 변이체는 성숙 단백질(mature protein)의 N- 및/또는 C-말단으로부터 일부 부분이 제거된 변이체를 포함할 수 있다. 상기 용어 “변이체”는 변이형, 변형, 변이형 폴리펩티드, 변이된 단백질, 변이 및 변이체 등의 용어(영문 표현으로는 modification, modified polypeptide, modified protein, mutant, mutein, divergent 등)가 혼용되어 사용될 수 있으며, 변이된 의미로 사용되는 용어라면 이에 제한되지 않는다. 본 출원의 목적상 상기 변이체는 서열번호 3의 아미노산 서열의 676번째 위치에 상응하는 아미노산인 아스파르트산(Aspartic acid, Asp, 또는 D)이 타이로신(Tyrosine, Tyr, 또는 Y)으로 치환된, 서열번호 1로 기재된 아미노산 서열을 포함하는 폴리펩티드일 수 있다.

[0020] 또한, 변이체는 폴리펩티드의 특성과 2차 구조에 최소한의 영향을 갖는 아미노산들의 결실 또는 부가를 포함할 수 있다. 예를 들면 변이체의 N-말단에는 번역-동시에(co-translationally) 또는 번역-후에(post-

translationally) 단백질의 이동(translocation)에 관여하는 시그널(또는 리더) 서열이 컨주게이트 될 수 있다. 또한 상기 변이체는 확인, 정제, 또는 합성할 수 있도록 다른 서열 또는 링커와 컨주게이트 될 수 있다.

- [0022] 본 출원에서 용어, '상동성(homology)' 또는 '동일성(identity)' 은 두 개의 주어진 아미노산 서열 또는 염기 서열 상호간 유사한 정도를 의미하며 백분율로 표시될 수 있다. 용어 상동성 및 동일성은 종종 상호교환적으로 이용될 수 있다.
- [0023] 보존된(conserved) 폴리뉴클레오티드 또는 폴리펩티드의 서열 상동성 또는 동일성은 표준 배열 알고리즘에 의해 결정되며, 사용되는 프로그램에 의해 확립된 디폴트 갭 페널티가 함께 이용될 수 있다. 실질적으로, 상동성을 갖거나(homologous) 또는 동일한(identical) 서열은 일반적으로 서열 전체 또는 일부분과 중간 또는 높은 엄격한 조건(stringent conditions)에서 하이브리드할 수 있다. 하이브리드화는 폴리뉴클레오티드에서 일반 코돈 또는 코돈 축퇴성을 고려한 코돈을 함유하는 폴리뉴클레오티드와의 하이브리드화 역시 포함됨이 자명하다.
- [0024] 임의의 두 폴리뉴클레오티드 또는 폴리펩티드 서열이 상동성, 유사성 또는 동일성을 갖는지 여부는, 예를 들어, Pearson et al(1988) [Proc. Natl. Acad. Sci. USA 85]: 2444에서와 같은 디폴트 파라미터를 이용하여 "FASTA" 프로그램과 같은 공지의 컴퓨터 알고리즘을 이용하여 결정될 수 있다. 또는, EMBOSS 패키지의 니들만 프로그램(EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Trends Genet. 16: 276-277)(버전 5.0.0 또는 이후 버전)에서 수행되는 바와 같은, 니들만-운치(Needleman-Wunsch) 알고리즘(Needleman and Wunsch, 1970, J. Mol. Biol. 48: 443-453)이 사용되어 결정될 수 있다(GCG 프로그램 패키지(Devereux, J., et al, Nucleic Acids Research 12: 387(1984)), BLASTP, BLASTN, FASTA (Atschul, [S.] [F.,] [ET AL, J MOLEC BIOL 215]: 403(1990); Guide to Huge Computers, Martin J. Bishop, [ED.,] Academic Press, San Diego,1994, 및 [CARILLO ETA/.](1988) SIAM J Applied Math 48: 1073을 포함한다). 예를 들어, 국립 생물공학 정보 데이터베이스 센터의 BLAST, 또는 ClustalW를 이용하여 상동성, 유사성 또는 동일성을 결정할 수 있다.
- [0025] 폴리뉴클레오티드 또는 폴리펩티드의 상동성, 유사성 또는 동일성은, 예를 들어, Smith and Waterman, Adv. Appl. Math(1981) 2:482 에 공지된 대로, 예를 들면, Needleman et al.(1970), J Mol Biol. 48:443과 같은 GAP 컴퓨터 프로그램을 이용하여 서열 정보를 비교함으로써 결정될 수 있다. 요약하면, GAP 프로그램은 두 서열 중 더 짧은 것에서의 기호의 전체 수로, 유사한 배열된 기호(즉, 뉴클레오티드 또는 아미노산)의 수를 나눈 값으로 정의할 수 있다. GAP 프로그램을 위한 디폴트 파라미터는 (1) 이진법 비교 매트릭스(동일성을 위해 1 그리고 비-동일성을 위해 0의 값을 함유함) 및 Schwartz and Dayhoff, eds., Atlas Of Protein Sequence And Structure, National Biomedical Research Foundation, pp. 353-358 (1979)에 의해 개시된 대로, Gribskov et al(1986) Nucl. Acids Res. 14: 6745의 가중된 비교 매트릭스(또는 EDNAFULL (NCBI NUC4.4의 EMBOSS 버전) 치환 매트릭스); (2) 각 갭을 위한 3.0의 페널티 및 각 갭에서 각 기호를 위한 추가의 0.10 페널티 (또는 갭 개방 페널티 10, 갭 연장 페널티 0.5); 및 (3) 말단 갭을 위한 무 페널티를 포함할 수 있다.
- [0026] 본 출원의 일 예로, 본 출원의 변이체는 야생형 폴리펩티드에 비해 L-라이신 생산능이 증가되도록 하는 활성을 가질 수 있다.
- [0027] 상기 서열번호 3의 아미노산 서열을 포함하는 폴리펩티드 및/또는 본 출원의 변이체는 NCg10752 유전자에 의해 코딩되는 단백질 활성을 갖는 폴리펩티드일 수 있으나 이에 제한되지 않는다.
- [0028] 본 출원에서 용어, "상응하는(corresponding to)"은, 폴리펩티드에서 열거되는 위치의 아미노산 잔기이거나, 또는 폴리펩티드에서 열거되는 잔기와 유사하거나 동일하거나 상동한 아미노산 잔기를 지칭한다. 상응하는 위치의 아미노산을 확인하는 것은 특정 서열을 참조하는 서열의 특정 아미노산을 결정하는 것일 수 있다. 본 출원에 사용된 "상응 영역"은 일반적으로 관련 단백질 또는 참조(reference) 단백질에서의 유사하거나 대응되는 위치를 지칭한다.
- [0029] 예를 들어, 임의의 아미노산 서열을 서열번호 3과 정렬(align)하고, 이를 토대로 상기 아미노산 서열의 각 아미노산 잔기는 서열번호 3의 아미노산 잔기와 상응하는 아미노산 잔기의 숫자 위치를 참조하여 넘버링 할 수 있다. 예를 들어, 본 출원에 기재된 것과 같은 서열 정렬 알고리즘은, 쿼리 시퀀스("참조 서열"이라고도 함)와 비교하여 아미노산의 위치, 또는 치환, 삽입 또는 결실 등의 변형이 발생하는 위치를 확인할 수 있다.
- [0030] 이러한 정렬에는 예를 들어 Needleman-Wunsch 알고리즘(Needleman 및 Wunsch, 1970, J. Mol. Biol. 48: 443-453), EMBOSS 패키지의 Needle 프로그램(EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000), Trends Genet. 16: 276-277) 등을 이용할 수 있으나, 이에 제한되지 않고 당업계에 알려진 서

열 정렬 프로그램, 쌍 서열(pairwise sequence) 비교 알고리즘 등을 적절히 사용할 수 있다.

- [0031] 본 출원의 다른 하나의 양태는 본 출원의 변이체를 코딩하는 폴리뉴클레오티드를 제공하는 것이다.
- [0032] 본 출원에서 용어, "폴리뉴클레오티드"는 뉴클레오티드 단위체(monomer)가 공유결합에 의해 길게 사슬모양으로 이어진 뉴클레오티드의 중합체(polymer)로 일정한 길이 이상의 DNA 또는 RNA 가닥으로서, 보다 구체적으로는 상기 변이체를 코딩하는 폴리뉴클레오티드 단편을 의미한다.
- [0033] 본 출원의 변이체를 코딩하는 폴리뉴클레오티드는 서열번호 1로 기재된 아미노산 서열을 코딩하는 염기서열을 포함할 수 있다. 본 출원의 일 예로, 본 출원의 폴리뉴클레오티드는 서열번호 2의 서열을 가지거나 포함할 수 있다. 또한, 본 출원의 폴리뉴클레오티드는 서열번호 2의 서열로 이루어지거나, 필수적으로 구성될 수 있다. 또 다른 예에서, 본 출원의 폴리뉴클레오티드는 상기 서열번호 2로 기재된 핵산 서열에서 서열번호 4의 핵산 서열을 기준으로 2026번 위치에 상응하는 염기는 T이고, 상기 서열번호 2로 기재된 핵산 서열과 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.7% 또는 99.9% 이상의 상동성 또는 동일성을 가지는 핵산 서열을 포함할 수 있다. 또한, 이러한 상동성 또는 동일성을 가지며 본 출원의 변이체에 상응하는 효능을 나타내는 폴리펩티드나 단백질을 암호화하는 서열이라면, 일부 서열이 결실, 변형, 치환, 보존적 치환 또는 부가된 핵산 서열을 갖는 폴리뉴클레오티드도 본 출원의 범위 내에 포함됨은 자명하다.
- [0034] 본 출원의 폴리뉴클레오티드는 코돈의 축퇴성(degeneracy) 또는 본 출원의 변이체를 발현시키고자 하는 생물에서 선호되는 코돈을 고려하여, 본 출원의 변이체의 아미노산 서열을 변화시키지 않는 범위 내에서 코딩 영역에 다양한 변형이 이루어질 수 있다. 구체적으로, 본 출원의 폴리뉴클레오티드는 서열번호 2의 서열과 상동성 또는 동일성이 70% 이상, 75% 이상, 80% 이상, 85% 이상, 90% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상, 및 100% 미만인 염기서열을 가지거나 포함하거나, 또는 서열번호 2의 서열과 상동성 또는 동일성이 70% 이상, 75% 이상, 80% 이상, 85% 이상, 90% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상, 및 100% 미만인 염기서열로 이루어지거나 필수적으로 이루어질 수 있으나, 이에 제한되지 않는다. 이때, 상기 상동성 또는 동일성을 갖는 서열에서, 서열번호 1의 676번째 위치에 상응하는 아미노산을 코딩하는 코돈은, 타이로신을 코딩하는 코돈 중 하나일 수 있다.
- [0035] 또한, 본 출원의 폴리뉴클레오티드는 공지의 유전자 서열로부터 제조될 수 있는 프로브, 예를 들면, 본 출원의 폴리뉴클레오티드 서열의 전체 또는 일부에 대한 상보 서열과 엄격한 조건 하에 하이브리드화할 수 있는 서열이라면 제한없이 포함될 수 있다. 상기 "엄격한 조건(stringent condition)"이란 폴리뉴클레오티드 간의 특이적 혼성화를 가능하게 하는 조건을 의미한다. 이러한 조건은 문헌(J. Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory press, Cold Spring Harbor, New York, 1989; F.M. Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, Inc., New York, 9.50-9.51, 11.7-11.8 참조)에 구체적으로 기재되어 있다. 예를 들어, 상동성 또는 동일성이 높은 폴리뉴클레오티드끼리, 70% 이상, 75% 이상, 80% 이상, 85% 이상, 90% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상, 또는 99% 이상의 상동성 또는 동일성을 갖는 폴리뉴클레오티드끼리 하이브리드화하고, 그보다 상동성 또는 동일성이 낮은 폴리뉴클레오티드끼리 하이브리드화하지 않는 조건, 또는 통상의 써던 하이브리드화(southern hybridization)의 세척 조건인 60℃, 1×SSC, 0.1% SDS, 구체적으로 60℃, 0.1×SSC, 0.1% SDS, 보다 구체적으로 68℃, 0.1×SSC, 0.1% SDS에 상당하는 염 농도 및 온도에서, 1회, 구체적으로 2회 내지 3회 세정하는 조건을 열거할 수 있다.
- [0036] 혼성화는 비록 혼성화의 엄격도에 따라 염기 간의 미스매치(mismatch)가 가능할지라도, 두 개의 핵산이 상보적 서열을 가질 것을 요구한다. 용어, "상보적"은 서로 혼성화가 가능한 뉴클레오티드 염기 간의 관계를 기술하는데 사용된다. 예를 들면, DNA에 관하여, 아데닌은 티민에 상보적이며 시토신은 구아닌에 상보적이다. 따라서, 본 출원의 폴리뉴클레오티드는 또한 실질적으로 유사한 핵산 서열뿐만 아니라 전체 서열에 상보적인 단리된 핵산 단편을 포함할 수 있다.
- [0037] 구체적으로, 본 출원의 폴리뉴클레오티드와 상동성 또는 동일성을 가지는 폴리뉴클레오티드는 55℃의 T<sub>m</sub> 값에서 혼성화 단계를 포함하는 혼성화 조건을 사용하고 상술한 조건을 사용하여 탐지할 수 있다. 또한, 상기 T<sub>m</sub> 값은 60℃, 63℃ 또는 65℃일 수 있으나, 이에 제한되는 것은 아니고 그 목적에 따라 당업자에 의해 적절히 조절될 수 있다.
- [0038] 상기 폴리뉴클레오티드를 혼성화하는 적절한 엄격도는 폴리뉴클레오티드의 길이 및 상보성 정도에 의존하고 변하는 해당기술분야에 잘 알려져 있다(예컨대, J.Sambrook et al., Molecular Cloning, A Laboratory Manual,

2nd Edition, Cold Spring Harbor Laboratory press, Cold Spring Harbor, New York, 1989; F.M. Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, Inc., New York, 9.50-9.51, 11.7-11.8 참조).

- [0040] 본 출원의 또 다른 하나의 양태는 본 출원의 폴리뉴클레오티드를 포함하는 벡터를 제공하는 것이다. 상기 벡터는 상기 폴리뉴클레오티드를 숙주세포에서 발현시키기 위한 발현 벡터일 수 있으나, 이에 제한되지 않는다.
- [0041] 본 출원에서, "벡터"는 적합한 숙주 내에서 목적 폴리펩티드를 발현시킬 수 있도록 적합한 발현조절영역(또는 발현조절서열)에 작동 가능하게 연결된 상기 목적 폴리펩티드를 코딩하는 폴리뉴클레오티드의 염기서열을 포함하는 DNA 제조물을 포함할 수 있다. 상기 발현조절영역은 전사를 개시할 수 있는 프로모터, 그러한 전사를 조절하기 위한 임의의 오퍼레이터 서열, 적합한 mRNA 리보솜 결합부위를 코딩하는 서열, 및 전사 및 해독의 종결을 조절하는 서열을 포함할 수 있다. 벡터는 적당한 숙주세포 내로 형질전환된 후, 숙주 계놈과 무관하게 복제되거나 기능할 수 있으며, 계놈 그 자체에 통합될 수 있다.
- [0042] 본 출원에서 사용되는 벡터는 특별히 한정되지 않으며, 당업계에 알려진 임의의 벡터를 이용할 수 있다. 통상 사용되는 벡터의 예로는 천연 상태이거나 재조합된 상태의 플라스미드, 코스미드, 바이러스 및 박테리오파지를 들 수 있다. 예를 들어, 파지 벡터 또는 코스미드 벡터로서 pWE15, M13, MBL3, MBL4, IXII, ASHII, APII, t10, t11, Charon4A, 및 Charon21A 등을 사용할 수 있으며, 플라스미드 벡터로서 pDZ계, pBR계, pUC계, pBluescriptII계, pGEM계, pTZ계, pCL계 및 pET계 등을 사용할 수 있다. 구체적으로는 pDZ, pDC, pDCM2, pACYC177, pACYC184, pCL, pECCG117, pUC19, pBR322, pMW118, pCC1BAC 벡터 등을 사용할 수 있다.
- [0043] 일례로 세포 내 염색체 삽입용 벡터를 통해 목적 폴리펩티드를 코딩하는 폴리뉴클레오티드를 염색체 내로 삽입할 수 있다. 상기 폴리뉴클레오티드의 염색체 내로의 삽입은 당업계에 알려진 임의의 방법, 예를 들면, 상동재조합(homologous recombination)에 의하여 이루어질 수 있으나, 이에 한정되지는 않는다. 상기 염색체 삽입 여부를 확인하기 위한 선별 마커(selection marker)를 추가로 포함할 수 있다. 상기 선별 마커는 벡터로 형질전환된 세포를 선별, 즉 목적 핵산 분자의 삽입 여부를 확인하기 위한 것으로, 약물 내성, 영양 요구성, 세포 독성제에 대한 내성 또는 표면 폴리펩티드의 발현과 같은 선택가능 표현형을 부여하는 마커들이 사용될 수 있다. 선택제(selective agent)가 처리된 환경에서는 선별 마커를 발현하는 세포만 생존하거나 다른 표현 형질을 나타내므로, 형질전환된 세포를 선별할 수 있다.
- [0044] 본 출원에서 용어 "형질전환"은 표적 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 벡터를 숙주세포 혹은 미생물 내에 도입하여 숙주세포 내에서 상기 폴리뉴클레오티드가 코딩하는 폴리펩티드가 발현할 수 있도록 하는 것을 의미한다. 형질전환된 폴리뉴클레오티드는 숙주세포 내에서 발현될 수 있지만 한다면, 숙주세포의 염색체 내에 삽입되어 위치하거나 염색체 외에 위치하거나 상관없이 이들 모두를 포함할 수 있다. 또한, 상기 폴리뉴클레오티드는 목적 폴리펩티드를 코딩하는 DNA 및/또는 RNA를 포함한다. 상기 폴리뉴클레오티드는 숙주세포 내로 도입되어 발현될 수 있는 것이면, 어떠한 형태로도 도입될 수 있다. 예를 들면, 상기 폴리뉴클레오티드는 자체적으로 발현되는데 필요한 모든 요소를 포함하는 유전자 구조체인 발현 카세트(expression cassette)의 형태로 숙주세포에 도입될 수 있다. 상기 발현 카세트는 통상 상기 폴리뉴클레오티드에 작동 가능하게 연결되어 있는 프로모터(promoter), 전사 종결신호, 리보솜 결합부위 및 번역 종결신호를 포함할 수 있다. 상기 발현 카세트는 자체 복제가 가능한 발현 벡터 형태일 수 있다. 또한, 상기 폴리뉴클레오티드는 그 자체의 형태로 숙주세포에 도입되어 숙주세포에서 발현에 필요한 서열과 작동 가능하게 연결되어 있는 것일 수도 있으며, 이에 제한되지 않는다.
- [0045] 또한, 상기에서 용어 "작동 가능하게 연결"된 것이란 본 출원의 목적 변이체를 코딩하는 폴리뉴클레오티드의 전사를 개시 및 매개하도록 하는 프로모터 서열과 상기 폴리뉴클레오티드 서열이 기능적으로 연결되어 있는 것을 의미한다.
- [0047] 본 출원의 또 다른 하나의 양태는 본 출원의 변이체 또는 본 출원의 폴리뉴클레오티드를 포함하는, 코리네박테리움 글루타미쿰(*Corynebacterium glutamicum*) 균주를 제공하는 것이다.
- [0048] 본 출원의 균주는 본 출원의 변이형 폴리펩티드, 상기 폴리펩티드를 암호화하는 폴리뉴클레오티드, 또는 본 출원의 폴리뉴클레오티드를 포함하는 벡터를 포함할 수 있다.
- [0049] 본 출원에서 용어, "균주(또는, 미생물)"는 야생형 미생물이나 자연적 또는 인위적으로 유전적 변형이 일어난 미생물을 모두 포함하며, 외부 유전자가 삽입되거나 내재적 유전자의 활성이 강화되거나 불활성화되는 등의 원인으로 인해서 특정 기작이 약화되거나 강화된 미생물로서, 목적하는 폴리펩티드, 단백질 또는 산물의 생산을

위하여 유전적 변형(modification)을 포함하는 미생물일 수 있다.

- [0050] 본 출원의 균주는 본 출원의 변이체, 본 출원의 폴리뉴클레오티드 및 본 출원의 폴리뉴클레오티드를 포함하는 벡터 중 어느 하나 이상을 포함하는 균주; 본 출원의 변이체 또는 본 출원의 폴리뉴클레오티드를 발현하도록 변형된 균주; 본 출원의 변이체, 또는 본 출원의 폴리뉴클레오티드를 발현하는 균주(예컨대, 재조합 균주); 또는 본 출원의 변이체 활성을 갖는 균주(예컨대, 재조합 균주)일 수 있으나, 이에 제한되지 않는다.
- [0051] 본 출원의 균주는 L-라이신 생산능을 가진 균주일 수 있다.
- [0052] 본 출원의 균주는 자연적으로 서열번호 3의 야생형 단백질을 발현하고/발현하거나 L-라이신 생산능을 가지고 있는 미생물, 또는 서열번호 3의 야생형 단백질을 발현하지 않고/않거나 L-라이신 생산능이 없는 모균주에 본 출원의 변이체 또는 이를 코딩하는 폴리뉴클레오티드(또는 상기 폴리뉴클레오티드를 포함하는 벡터)가 도입되거나 및/또는 L-라이신 생산능이 부여된 미생물일 수 있으나 이에 제한되지 않는다.
- [0053] 일 예로, 본 출원의 균주는 본 출원의 폴리뉴클레오티드 또는 본 출원의 변이체를 코딩하는 폴리뉴클레오티드를 포함하는 벡터로 형질전환되어, 본 출원의 변이체를 발현하는 세포 또는 미생물로서, 본 출원의 목적상 본 출원의 균주는 본 출원의 변이체를 포함하여 L-라이신을 생산할 수 있는 미생물을 모두 포함할 수 있다. 예를 들어, 본 출원의 균주는 천연의 야생형 미생물 또는 L-라이신을 생산하는 미생물에 본 출원의 변이체를 코딩하는 폴리뉴클레오티드가 도입됨으로써 L-라이신 생산능이 증가된 재조합 균주일 수 있다. 상기 L-아미노산 생산능이 증가된 재조합 균주는, 천연의 야생형 미생물 또는 비변형 미생물(즉, 야생형 단백질(예를 들면, 서열번호 3의 아미노산 서열을 포함하는 단백질)을 발현하는 미생물 또는 변이형 단백질(예를 들면, 서열번호 1의 아미노산 서열을 포함하는 단백질)을 발현하지 않는 미생물)에 비하여 L-라이신 생산능이 증가된 미생물일 수 있으나, 이에 제한되는 것은 아니다. 그 예로, 상기 L-라이신 생산능의 증가 여부를 비교하는 대상 균주인, 단백질 비변형 미생물은 ATCC13032 균주 및/또는 코리네박테리움 글루타미쿰 CJ3P(US 9556463 B2; 문헌 전체가 본 명세서에 참조로서 포함됨)일 수 있으나, 이에 제한되지 않는다.
- [0054] 일 예로, 상기 생산능이 증가된 재조합 균주는 변이 전 모균주 또는 비변형 미생물에 비하여, L-라이신 생산능(또는 생산량)이 약 1% 이상, 약 2.5% 이상, 약 5% 이상, 약 6% 이상, 약 7% 이상, 약 8% 이상, 약 9% 이상, 약 10% 이상, 약 10.5% 이상, 약 11% 이상, 약 11.5% 이상, 약 12% 이상, 약 12.5% 이상, 약 13% 이상, 약 13.5% 이상, 약 14% 이상, 약 14.5% 이상, 약 15% 이상, 약 15.5% 이상, 약 16% 이상, 약 16.5% 이상, 약 17% 이상, 약 17.5% 이상, 약 18% 이상, 약 18.5% 이상, 약 18.9% 이상, 약 19% 이상, 약 19.5% 이상, 약 20% 이상, 약 20.5% 이상, 약 21% 이상, 약 21.5% 이상, 약 22% 이상, 약 22.5% 이상, 약 23% 이상, 약 23.5% 이상, 약 24% 이상, 약 24.5% 이상, 약 25% 이상, 약 25.5% 이상, 약 26% 이상, 약 26.5% 이상, 약 27% 이상, 약 27.5% 이상, 약 28% 이상, 약 28.5% 이상, 약 29% 이상, 약 29.5% 이상, 약 30% 이상, 약 31% 이상, 약 32% 이상, 약 33% 이상, 약 34% 이상, 또는 약 35% 이상(상한값은 특별한 제한은 없으며, 예컨대, 약 200% 이하, 약 150% 이하, 약 100% 이하, 약 50% 이하, 약 45% 이하, 약 40% 이하, 또는 약 35% 이하일 수 있음) 증가된 것일 수 있다. 다른 예에서, 상기 생산능(또는 생산량)이 증가된 재조합 균주는 변이 전 모균주 또는 비변형 미생물에 비하여, L-라이신 생산능(또는 생산량)이 약 1.1배 이상, 약 1.12배 이상, 약 1.13배 이상, 1.15배 이상, 1.16배 이상, 1.17배 이상, 1.18배 이상, 1.19배 이상, 약 1.2배 이상, 약 1.21배 이상, 1.25배 이상, 또는 약 1.3배 이상(상한값은 특별한 제한은 없으며, 예컨대, 약 10배 이하, 약 5배 이하, 약 3배 이하, 또는 약 2배 이하일 수 있음) 증가된 것일 수 있다. 보다 구체적으로는, 상기 생산능(또는 생산량)이 증가된 재조합 균주는 변이 전 모균주 또는 비변형 미생물에 비하여, L-라이신 생산능(또는 생산량)이 약 18.9%(또는 약 1.19배) 증가된 것일 수 있으나, 이에 제한되지 않는다. 상기 용어 “약(about)”은  $\pm 0.5$ ,  $\pm 0.4$ ,  $\pm 0.3$ ,  $\pm 0.2$ ,  $\pm 0.1$  등을 모두 포함하는 범위로, 약 이란 용어 뒤에 나오는 수치와 동등하거나 유사한 범위의 수치를 모두 포함하나, 이에 제한되지 않는다.
- [0055] 본 출원에서 용어, "비변형 미생물"은 미생물에 자연적으로 발생할 수 있는 돌연변이를 포함하는 균주를 제외하는 것이 아니며, 야생형 균주 또는 천연형 균주 자체이거나, 자연적 또는 인위적 요인에 의한 유전적 변이로 형질이 변화되기 전 균주를 의미할 수 있다. 예를 들어, 상기 비변형 미생물은 본 명세서에 기재된 단백질 변이체가 도입되지 않거나 도입되기 전의 균주를 의미할 수 있다. 상기 "비변형 미생물"은 “변형 전 균주”, “변형 전 미생물”, “비변이 균주”, “비변형 균주”, “비변이 미생물” 또는 “기준 미생물”과 혼용될 수 있다.
- [0056] 본 출원의 또 다른 일 예로, 본 출원의 미생물은 코리네박테리움 글루타미쿰(*Corynebacterium glutamicum*), 코리네박테리움 크루디락티스(*Corynebacterium crudilactis*), 코리네박테리움 데세르티(*Corynebacterium deserti*), 코리네박테리움 이피시엔스(*Corynebacterium efficiens*), 코리네박테리움 칼루네(*Corynebacterium*

*callunae*), 코리네박테리움 스테셔니스(*Corynebacterium stationis*), 코리네박테리움 싱굴라레(*Corynebacterium singulare*), 코리네박테리움 할로톨레란스(*Corynebacterium halotolerans*), 코리네박테리움 스트리아툼(*Corynebacterium striatum*), 코리네박테리움 암모니아게네스(*Corynebacterium ammoniagenes*), 코리네박테리움 폴루티솔리(*Corynebacterium pollutisoli*), 코리네박테리움 이미탄스(*Corynebacterium imitans*), 코리네박테리움 테스투디노리스(*Corynebacterium testudinoris*) 또는 코리네박테리움 플라베스센스(*Corynebacterium flavescens*)일 수 있다.

- [0057] 본 출원에서 용어, 폴리펩티드의 “약화”는 내재적 활성에 비하여 활성이 감소되거나 또는 활성이 없는 것을 모두 포함하는 개념이다. 상기 약화는 불활성화(inactivation), 결핍(deficiency), 하향조절(down-regulation), 감소(decrease), 저하(reduce), 감쇠(attenuation) 등의 용어와 혼용될 수 있다.
- [0058] 상기 약화는 상기 폴리펩티드를 코딩하는 폴리뉴클레오티드의 변이 등으로 폴리펩티드 자체의 활성이 본래 미생물이 가지고 있는 폴리펩티드의 활성에 비해 감소 또는 제거된 경우, 이를 코딩하는 폴리뉴클레오티드의 유전자의 발현 저해 또는 폴리펩티드로의 번역(translation) 저해 등으로 세포 내에서 전체적인 폴리펩티드 활성 정도 및/또는 농도(발현량)가 천연형 균주에 비하여 낮은 경우, 상기 폴리뉴클레오티드의 발현이 전혀 이루어지지 않은 경우, 및/또는 폴리뉴클레오티드의 발현이 되더라도 폴리펩티드의 활성이 없는 경우 역시 포함할 수 있다. 상기 “내재적 활성”은 자연적 또는 인위적 요인에 의한 유전적 변이로 형질이 변화하는 경우, 형질 변화 전 모균주, 야생형 또는 비변형 미생물이 본래 가지고 있던 특정 폴리펩티드의 활성을 의미한다. 이는 “변형 전 활성”과 혼용되어 사용될 수 있다. 폴리펩티드의 활성이 내재적 활성에 비하여 “불활성화, 결핍, 감소, 하향 조절, 저하, 감쇠” 한다는 것은, 형질 변화 전 모균주 또는 비변형 미생물이 본래 가지고 있던 특정 폴리펩티드의 활성에 비하여 낮아진 것을 의미한다.
- [0059] 이러한 폴리펩티드의 활성의 약화는, 당업계에 알려진 임의의 방법에 의하여 수행될 수 있으나 이로 제한되는 것은 아니며, 당해 분야에 잘 알려진 다양한 방법의 적용으로 달성될 수 있다(예컨대, Nakashima N et al., Bacterial cellular engineering by genome editing and gene silencing. Int J Mol Sci. 2014;15(2):2773-2793, Sambrook et al. Molecular Cloning 2012 등).
- [0061] 구체적으로, 본 출원의 폴리펩티드의 약화는
  - [0062] 1) 폴리펩티드를 코딩하는 유전자 전체 또는 일부의 결손;
  - [0063] 2) 폴리펩티드를 코딩하는 유전자의 발현이 감소하도록 발현조절영역(또는 발현조절서열)의 변형;
  - [0064] 3) 폴리펩티드의 활성이 제거 또는 약화되도록 상기 폴리펩티드를 구성하는 아미노산 서열의 변형(예컨대, 아미노산 서열 상의 1 이상의 아미노산의 삭제/치환/부가);
  - [0065] 4) 폴리펩티드의 활성이 제거 또는 약화되도록 상기 폴리펩티드를 코딩하는 유전자 서열의 변형(예를 들어, 폴리펩티드의 활성이 제거 또는 약화되도록 변형된 폴리펩티드를 코딩하도록 상기 폴리펩티드 유전자의 핵산염기 서열 상의 1 이상의 핵산염기의 삭제/치환/부가);
  - [0066] 5) 폴리펩티드를 코딩하는 유전자 전사체의 개시코돈 또는 5'-UTR 지역을 코딩하는 염기서열의 변형;
  - [0067] 6) 폴리펩티드를 코딩하는 상기 유전자의 전사체에 상보적으로 결합하는 안티센스 올리고뉴클레오티드(예컨대, 안티센스 RNA)의 도입;
  - [0068] 7) 리보솜(ribosome)의 부착이 불가능한 2차 구조물을 형성시키기 위하여 폴리펩티드를 코딩하는 유전자의 사인-달가르노(Shine-Dalgarno) 서열 앞단에 사인-달가르노 서열과 상보적인 서열의 부가;
  - [0069] 8) 폴리펩티드를 코딩하는 유전자 서열의 ORF(open reading frame)의 3' 말단에 반대 방향으로 전사되는 프로모터의 부가(Reverse transcription engineering, RTE); 또는
  - [0070] 9) 상기 1) 내지 8) 중 선택된 2 이상의 조합일 수 있으나, 이에, 특별히 제한되는 것은 아니다.
- [0071] 예컨대,
  - [0072] 상기 1) 폴리펩티드를 코딩하는 상기 유전자 일부 또는 전체의 결손은, 염색체 내 내재적 목적 폴리펩티드를 코딩하는 폴리뉴클레오티드 전체의 제거, 일부 뉴클레오티드가 결실된 폴리뉴클레오티드로의 교체 또는 마커 유전자로 교체될 수 있다.
  - [0073] 또한, 상기 2) 발현조절영역(또는 발현조절서열)의 변형은, 결실, 삽입, 비보존적 또는 보존적 치환 또는 이들

의 조합으로 발현조절영역(또는 발현조절서열) 상의 변이 발생, 또는 더욱 약한 활성을 갖는 서열로의 교체일 수 있다. 상기 발현조절영역에는 프로모터, 오퍼레이터 서열, 리보솜 결합부위를 코딩하는 서열, 및 전사와 해독의 종결을 조절하는 서열을 포함하나, 이에 한정되는 것은 아니다.

- [0074] 또한, 상기 3) 폴리펩티드를 코딩하는 유전자 전사체의 개시코돈 또는 5'-UTR 지역을 코딩하는 염기서열 변형은, 예를 들면, 내재적 개시코돈에 비해 폴리펩티드 발현율이 더 낮은 다른 개시코돈을 코딩하는 염기서열로 치환하는 것일 수 있으나, 이에 제한되지 않는다.
- [0075] 또한, 상기 4) 및 5)의 아미노산 서열 또는 폴리뉴클레오티드 서열의 변형은 폴리펩티드의 활성을 약화하도록 상기 폴리펩티드의 아미노산 서열 또는 상기 폴리펩티드를 코딩하는 폴리뉴클레오티드 서열을 결실, 삽입, 비보존적 또는 보존적 치환 또는 이들의 조합으로 서열상의 변이 발생, 또는 더욱 약한 활성을 갖도록 개량된 아미노산 서열 또는 폴리뉴클레오티드 서열 또는 활성이 없도록 개량된 아미노산 서열 또는 폴리뉴클레오티드 서열로의 교체일 수 있으나, 이에 한정되는 것은 아니다. 예를 들면, 폴리뉴클레오티드 서열 내 변이를 도입하여 종결 코돈을 형성시킴으로써, 유전자의 발현을 저해하거나 약화시킬 수 있으나, 이에 제한되지 않는다.
- [0076] 상기 6) 폴리펩티드를 코딩하는 상기 유전자의 전사체에 상보적으로 결합하는 안티센스 올리고뉴클레오티드(예컨대, 안티센스 RNA)의 도입은 예를 들어 문헌 [Weintraub, H. et al., Antisense-RNA as a molecular tool for genetic analysis, Reviews - Trends in Genetics, Vol. 1(1) 1986]을 참고할 수 있다.
- [0077] 상기 7) 리보솜(ribosome)의 부착이 불가능한 2차 구조물을 형성시키기 위하여 폴리펩티드를 코딩하는 유전자의 사인-달가르노(Shine-Dalgarno) 서열 앞단에 사인-달가르노 서열과 상보적인 서열의 부가는 mRNA 번역을 불가능하게 하거나 속도를 저하시키는 것일 수 있다.
- [0078] 상기 8) 폴리펩티드를 코딩하는 유전자서열의 ORF(open reading frame)의 3' 말단에 반대 방향으로 전사되는 프로모터의 부가(Reverse transcription engineering, RTE)는 상기 폴리펩티드를 코딩하는 유전자의 전사체에 상보적인 안티센스 뉴클레오티드를 만들어 활성을 약화하는 것일 수 있다.
- [0079] 본 출원에서 용어, 폴리펩티드 활성의 “강화”는, 폴리펩티드의 활성이 내재적 활성에 비하여 증가되는 것을 의미한다. 상기 강화는 활성화(activation), 상향조절(up-regulation), 과발현(overexpression), 증가(increase) 등의 용어와 혼용될 수 있다. 여기서 활성화, 강화, 상향조절, 과발현, 증가는 본래 가지고 있지 않았던 활성을 나타내게 되는 것, 또는 내재적 활성 또는 변형 전 활성에 비하여 향상된 활성을 나타내게 되는 것을 모두 포함할 수 있다. 상기 “내재적 활성”은 자연적 또는 인위적 요인에 의한 유전적 변이로 형질이 변화하는 경우, 형질 변화 전 모균주 또는 비변형 미생물이 본래 가지고 있던 특정 폴리펩티드의 활성을 의미한다. 이는 “변형 전 활성”과 혼용되어 사용될 수 있다. 폴리펩티드의 활성이 내재적 활성에 비하여 “강화”, “상향조절”, “과발현” 또는 “증가”한다는 것은, 형질 변화 전 모균주 또는 비변형 미생물이 본래 가지고 있던 특정 폴리펩티드의 활성 및/또는 농도(발현량)에 비하여 향상된 것을 의미한다.
- [0080] 상기 강화는 외래의 폴리펩티드를 도입하거나, 내재적인 폴리펩티드의 활성 강화 및/또는 농도(발현량)를 통해 달성할 수 있다. 상기 폴리펩티드의 활성의 강화 여부는 해당 폴리펩티드의 활성 정도, 발현량 또는 해당 폴리펩티드로부터 배출되는 산물의 양의 증가로부터 확인할 수 있다.
- [0081] 상기 폴리펩티드의 활성의 강화는 당해 분야에 잘 알려진 다양한 방법의 적용이 가능하며, 목적 폴리펩티드의 활성을 변형전 미생물보다 강화시킬 수 있는 한, 제한되지 않는다. 구체적으로, 분자생물학의 일상적 방법인 당업계의 통상의 기술자에게 잘 알려진 유전자 공학 및/또는 단백질 공학을 이용한 것일 수 있으나, 이로 제한되지 않는다(예컨대, Sitnicka et al. Functional Analysis of Genes. Advances in Cell Biology. 2010, Vol. 2. 1-16, Sambrook et al. Molecular Cloning 2012 등).
- [0082] 구체적으로, 본 출원의 폴리펩티드의 강화는
  - [0083] 1) 폴리펩티드를 코딩하는 폴리뉴클레오티드의 세포 내 카피수 증가;
  - [0084] 2) 폴리펩티드를 코딩하는 염색체상의 유전자 발현조절영역을 활성이 강력한 서열로 교체;
  - [0085] 3) 폴리펩티드를 코딩하는 유전자 전사체의 개시코돈 또는 5'-UTR 지역을 코딩하는 염기서열의 변형;
  - [0086] 4) 폴리펩티드 활성이 강화되도록 상기 폴리펩티드의 아미노산 서열의 변형;
  - [0087] 5) 폴리펩티드 활성이 강화되도록 상기 폴리펩티드를 코딩하는 폴리뉴클레오티드 서열의 변형(예를 들어, 폴리펩티드의 활성이 강화되도록 변형된 폴리펩티드를 코딩하도록 상기 폴리펩티드 유전자의 폴리뉴클레오티드 서열의

변형);

- [0088] 6) 폴리펩티드의 활성을 나타내는 외래 폴리펩티드 또는 이를 코딩하는 외래 폴리뉴클레오티드의 도입;
- [0089] 7) 폴리펩티드를 암호화하는 폴리뉴클레오티드의 코돈 최적화;
- [0090] 8) 폴리펩티드의 삼차구조를 분석하여 노출 부위를 선택하여 변형하거나 화학적으로 수식; 또는
- [0091] 9) 상기 1) 내지 8) 중 선택된 2 이상의 조합일 수 있으나, 이에, 특별히 제한되는 것은 아니다.
- [0092] 보다 구체적으로,
- [0093] 상기 1) 폴리펩티드를 코딩하는 폴리뉴클레오티드의 세포 내 카피수 증가는, 해당 폴리펩티드를 코딩하는 폴리뉴클레오티드가 작동가능하게 연결된, 숙주와 무관하게 복제되고 기능할 수 있는 벡터의 숙주세포 내로의 도입에 의해 달성되는 것일 수 있다. 또는, 해당 폴리펩티드를 코딩하는 폴리뉴클레오티드가 숙주세포 내의 염색체 내에 1 카피 또는 2 카피 이상 도입에 의해 달성되는 것일 수 있다. 상기 염색체 내에 도입은 숙주세포 내의 염색체 내로 상기 폴리뉴클레오티드를 삽입시킬 수 있는 벡터가 숙주세포 내에 도입됨으로써 수행될 수 있으나, 이에 제한되지 않는다. 상기 벡터는 전술한 바와 같다.
- [0094] 상기 2) 폴리펩티드를 코딩하는 염색체상의 유전자 발현조절영역(또는 발현조절서열)을 활성이 강력한 서열로 교체는, 예를 들면, 상기 발현조절영역의 활성을 더욱 강화하도록 결실, 삽입, 비보존적 또는 보존적 치환 또는 이들의 조합으로 서열상의 변이 발생, 또는 더욱 강한 활성을 가지는 서열로의 교체일 수 있다. 상기 발현조절영역은, 특별히 이에 제한되지 않으나 프로모터, 오피레이터 서열, 리보솜 결합 부위를 코딩하는 서열, 그리고 전사 및 해독의 종결을 조절하는 서열 등을 포함할 수 있다. 일 예로, 본래의 프로모터를 강력한 프로모터로 교체시키는 것일 수 있으나, 이에 제한되지 않는다.
- [0095] 공지된 강력한 프로모터의 예에는 CJ1 내지 CJ7 프로모터(미국등록특허 US 7662943 B2), lac 프로모터, trp 프로모터, trc 프로모터, tac 프로모터, 람다 파아지 PR 프로모터, PL 프로모터, tet 프로모터, gapA 프로모터, SPL7 프로모터, SPL13(sm3) 프로모터(미국등록특허 US 10584338 B2), O2 프로모터(미국등록특허 US 10273491 B2), tkt 프로모터, yccA 프로모터 등이 있으나, 이에 제한되지 않는다.
- [0096] 상기 3) 폴리펩티드를 코딩하는 유전자 전사체의 개시코돈 또는 5'-UTR 지역을 코딩하는 염기서열 변형은, 예를 들면, 내재적 개시코돈에 비해 폴리펩티드 발현율이 더 높은 다른 개시코돈을 코딩하는 염기 서열로 치환하는 것일 수 있으나, 이에 제한되지 않는다.
- [0097] 상기 4) 및 5)의 아미노산 서열 또는 폴리뉴클레오티드 서열의 변형은, 폴리펩티드의 활성을 강화하도록 상기 폴리펩티드의 아미노산 서열 또는 상기 폴리펩티드를 코딩하는 폴리뉴클레오티드 서열을 결실, 삽입, 비보존적 또는 보존적 치환 또는 이들의 조합으로 서열상의 변이 발생, 또는 더욱 강한 활성을 갖도록 개량된 아미노산 서열 또는 폴리뉴클레오티드 서열 또는 활성이 증가하도록 개량된 아미노산 서열 또는 폴리뉴클레오티드 서열로의 교체일 수 있으나, 이에 한정되는 것은 아니다. 상기 교체는 구체적으로 상동재조합에 의하여 폴리뉴클레오티드를 염색체내로 삽입함으로써 수행될 수 있으나, 이에 제한되지 않는다. 이때 사용되는 벡터는 염색체 삽입 여부를 확인하기 위한 선별 마커 (selection marker)를 추가로 포함할 수 있다. 상기 선별 마커는 전술한 바와 같다.
- [0098] 상기 6) 폴리펩티드의 활성을 나타내는 외래 폴리뉴클레오티드의 도입은, 상기 폴리펩티드와 동일/유사한 활성을 나타내는 폴리펩티드를 코딩하는 외래 폴리뉴클레오티드의 숙주세포 내 도입일 수 있다. 상기 외래 폴리뉴클레오티드는 상기 폴리펩티드와 동일/유사한 활성을 나타내는 한 그 유래나 서열에 제한이 없다. 상기 도입에 이용되는 방법은 공지된 형질전환 방법을 당업자가 적절히 선택하여 수행될 수 있으며, 숙주 세포 내에서 상기 도입된 폴리뉴클레오티드가 발현됨으로써 폴리펩티드가 생성되어 그 활성이 증가될 수 있다.
- [0099] 상기 7) 폴리펩티드를 암호화하는 폴리뉴클레오티드의 코돈 최적화는, 내재 폴리뉴클레오티드가 숙주세포 내에서 전사 또는 번역이 증가하도록 코돈 최적화한 것이거나, 또는 외래 폴리뉴클레오티드가 숙주세포 내에서 최적화된 전사, 번역이 이루어지도록 이의 코돈을 최적화한 것일 수 있다.
- [0100] 상기 8) 폴리펩티드의 삼차구조를 분석하여 노출 부위를 선택하여 변형하거나 화학적으로 수식하는 것은, 예를 들어 분석하고자 하는 폴리펩티드의 서열정보를 기지 단백질들의 서열정보가 저장된 데이터베이스와 비교함으로써 서열의 유사성 정도에 따라 주형 단백질 후보를 결정하고 이를 토대로 구조를 확인하여, 변형하거나 화학적으로 수식할 노출 부위를 선택하여 변형 또는 수식하는 것일 수 있다.

- [0101] 이와 같은 폴리펩티드 활성의 강화는, 상응하는 폴리펩티드의 활성 또는 농도 발현량이 야생형이나 변형 전 미생물 균주에서 발현된 폴리펩티드의 활성 또는 농도를 기준으로 하여 증가되거나, 해당 폴리펩티드로부터 생산되는 산물의 양의 증가되는 것일 수 있으나, 이에 제한되는 것은 아니다.
- [0103] 본 출원의 미생물에서 폴리뉴클레오티드의 일부 또는 전체의 변형(예컨대, 상술한 단백질 변이체를 코딩하기 위한 변형)은 (a) 미생물 내 염색체 삽입용 벡터를 이용한 상동 재조합 또는 유전자가위(engineered nuclease, e.g., CRISPR-Cas9)을 이용한 유전체 교정 및/또는 (b) 자외선 및 방사선 등과 같은 빛 및/또는 화학물질 처리에 의해 유도될 수 있으나 이에 제한되지 않는다. 상기 유전자 일부 또는 전체의 변형 방법에는 DNA 재조합 기술에 의한 방법이 포함될 수 있다. 예를 들면, 목적 유전자와 상동성이 있는 뉴클레오티드 서열을 포함하는 뉴클레오티드 서열 또는 벡터를 상기 미생물에 주입하여 상동 재조합(homologous recombination)이 일어나게 함으로써 유전자 일부 또는 전체의 결손이 이루어질 수 있다. 상기 주입되는 뉴클레오티드 서열 또는 벡터는 우성 선별 마커를 포함할 수 있으나, 이에 제한되는 것은 아니다.
- [0105] 본 출원의 미생물에서, 변이체, 폴리뉴클레오티드 및 L-라이신 등은 상기 다른 양태에서 기재한 바와 같다.
- [0107] 본 출원의 또 다른 하나의 양태는 본 출원의 변이체 또는 본 출원의 폴리뉴클레오티드를 포함하는 코리네박테리움 글루타미쿰 균주를 배지에서 배양하는 단계를 포함하는, L-아미노산 생산방법을 제공한다.
- [0108] 본 출원의 L-아미노산 생산방법은 본 출원의 변이체 또는 본 출원의 폴리뉴클레오티드 또는 본 출원의 벡터를 포함하는 코리네박테리움 글루타미쿰 균주를 배지에서 배양하는 단계를 포함할 수 있다.
- [0109] 더불어, 본 출원의 L-아미노산은 L- 라이신일 수 있다.
- [0110] 본 출원에서, 용어 "배양"은 본 출원의 코리네박테리움 글루타미쿰 균주를 적당히 조절된 환경 조건에서 생육시키는 것을 의미한다. 본 출원의 배양과정은 당업계에 알려진 적당한 배지와 배양조건에 따라 이루어질 수 있다. 이러한 배양 과정은 선택되는 균주에 따라 당업자가 용이하게 조정하여 사용할 수 있다. 구체적으로 상기 배양은 회분식, 연속식 및/또는 유가식일 수 있으나, 이에 제한되는 것은 아니다.
- [0111] 본 출원에서 용어, "배지"는 본 출원의 코리네박테리움 글루타미쿰 균주를 배양하기 위해 필요로 하는 영양물질을 주성분으로 혼합한 물질을 의미하며, 생존 및 발육에 불가결한 물을 비롯하여 영양물질 및 발육인자 등을 공급한다. 구체적으로, 본 출원의 코리네박테리움 글루타미쿰 균주의 배양에 사용되는 배지 및 기타 배양 조건은 통상의 미생물의 배양에 사용되는 배지라면 특별한 제한 없이 어느 것이나 사용할 수 있으나, 본 출원의 코리네박테리움 글루타미쿰 균주를 적당한 탄소원, 질소원, 인원, 무기화합물, 아미노산 및/또는 비타민 등을 함유한 통상의 배지 내에서 호기성 조건 하에서 온도, pH 등을 조절하면서 배양할 수 있다.
- [0112] 구체적으로, 코리네박테리움 속 균주에 대한 배양 배지는 문헌["Manual of Methods for General Bacteriology" by the American Society for Bacteriology (Washington D.C., USA, 1981)]에서 찾아 볼 수 있다.
- [0113] 본 출원에서 상기 탄소원으로는 글루코오스, 사카로오스, 락토오스, 프룩토오스, 수크로오스, 말토오스 등과 같은 탄수화물; 만니톨, 소르비톨 등과 같은 당 알코올, 피루브산, 락트산, 시트르산 등과 같은 유기산; 글루탐산, 메티오닌, 라이신 등과 같은 아미노산 등이 포함될 수 있다. 또한, 전분 가수분해물, 당밀, 블랙스트랩 당밀, 쌀겨울, 카사버, 사탕수수 찌꺼기 및 옥수수 침지액 같은 천연의 유기 영양원을 사용할 수 있으며, 구체적으로는 글루코오스 및 살균된 전처리 당밀(즉, 환원당으로 전환된 당밀) 등과 같은 탄수화물이 사용될 수 있으며, 그 외의 적정량의 탄소원을 제한 없이 다양하게 이용할 수 있다. 이들 탄소원은 단독으로 사용되거나 2 종 이상이 조합되어 사용될 수 있으며, 이에 한정되는 것은 아니다.
- [0114] 상기 질소원으로는 암모니아, 황산암모늄, 염화암모늄, 초산암모늄, 인산암모늄, 탄산암모늄, 질산암모늄 등과 같은 무기질소원; 글루탐산, 메티오닌, 글루타민 등과 같은 아미노산, 펩톤, NZ-아민, 육류 추출물, 효모 추출물, 맥아 추출물, 옥수수 침지액, 카제인 가수분해물, 어류 또는 그의 분해생성물, 탈지 대두 케이크 또는 그의 분해 생성물 등과 같은 유기 질소원이 사용될 수 있다. 이들 질소원은 단독으로 사용되거나 2 종 이상이 조합되어 사용될 수 있으며, 이에 한정되는 것은 아니다.
- [0115] 상기 인원으로는 인산 제1칼륨, 인산 제2칼륨, 또는 이에 대응되는 소듐-함유 염 등이 포함될 수 있다. 무기 화합물로는 염화나트륨, 염화칼슘, 염화철, 황산마그네슘, 황산철, 황산망간, 탄산칼슘 등이 사용될 수 있으며, 그 외에 아미노산, 비타민 및/또는 적절한 전구체 등이 포함될 수 있다. 이들 구성성분 또는 전구체는 배지에 회분식 또는 연속식으로 첨가될 수 있다. 그러나, 이에 한정되는 것은 아니다.
- [0116] 또한, 본 출원의 코리네박테리움 글루타미쿰 균주의 배양 중에 수산화암모늄, 수산화칼륨, 암모니아, 인산, 황

산 등과 같은 화합물을 배지에 적절한 방식으로 첨가하여, 배지의 pH를 조정할 수 있다. 또한, 배양 중에는 지방산 폴리글리콜 에스테르와 같은 소포제를 사용하여 기포 생성을 억제할 수 있다. 또한, 배지의 호기 상태를 유지하기 위하여, 배지 내로 산소 또는 산소 함유 기체를 주입하거나 혐기 및 미호기 상태를 유지하기 위해 기체의 주입 없이 혹은 질소, 수소 또는 이산화탄소 가스를 주입할 수 있으며, 이에 한정되는 것은 아니다.

[0117] 본 출원의 배양에서 배양온도는 20 내지 45℃ 구체적으로는 25 내지 40℃ 를 유지할 수 있고, 약 10 내지 160 시간 동안 배양할 수 있으나, 이에 한정되는 것은 아니다.

[0118] 본 출원의 배양에 의하여 생산된 L-아미노산은 배지 중으로 분비되거나 세포 내에 잔류할 수 있다.

[0120] 본 출원의 L-아미노산 생산방법은, 본 출원의 코리네박테리움 글루타미쿰 균주를 준비하는 단계, 상기 균주를 배양하기 위한 배지를 준비하는 단계, 또는 이들의 조합(순서에 무관, in any order)을, 예를 들어, 상기 배양하는 단계 이전에, 추가로 포함할 수 있다.

[0121] 본 출원의 L-아미노산 생산방법은, 상기 배양에 따른 배지(배양이 수행된 배지) 또는 코리네박테리움 글루타미쿰 균주로부터 L-아미노산을 회수하는 단계를 추가로 포함할 수 있다. 상기 회수하는 단계는 상기 배양하는 단계 이후에 추가로 포함될 수 있다.

[0122] 상기 회수는 본 출원의 미생물의 배양 방법, 예를 들어 회분식, 연속식 또는 유가식 배양 방법 등에 따라 당해 기술 분야에 공지된 적합한 방법을 이용하여 목적하는 L-아미노산을 수집(collect)하는 것일 수 있다. 예를 들어, 원심분리, 여과, 결정화 단백질 침전제에 의한 처리(염석법), 추출, 초음파 파쇄, 한외여과, 투석법, 분자체 크로마토그래피(겔여과), 흡착크로마토그래피, 이온교환 크로마토그래피, 친화도 크로마토그래피 등의 각종 크로마토그래피, HPLC 또는 이들의 방법을 조합하여 사용될 수 있으며, 당해 분야에 공지된 적합한 방법을 이용하여 배지 또는 미생물로부터 목적하는 L-아미노산을 회수할 수 있다.

[0123] 또한, 본 출원의 L-아미노산 생산방법은, 추가적으로 정제 단계를 포함할 수 있다. 상기 정제는 당해 기술분야에 공지된 적합한 방법을 이용하여, 수행할 수 있다. 일 예에서, 본 출원의 L-아미노산 생산방법이 회수 단계와 정제 단계를 모두 포함하는 경우, 상기 회수 단계와 정제 단계는 순서에 상관없이 연속적 또는 비연속적으로 수행되거나, 동시에 또는 하나의 단계로 통합되어 수행될 수 있으나, 이에 제한되는 것은 아니다.

[0124] 본 출원의 방법에서, 변이체, 폴리뉴클레오티드, 벡터 및 균주 등은 상기 다른 양태에서 기재한 바와 같다.

[0126] 본 출원의 또 다른 하나의 양태는 본 출원의 변이체, 상기 변이체를 코딩하는 폴리뉴클레오티드, 상기 폴리뉴클레오티드를 포함하는 벡터 또는 본 출원의 폴리뉴클레오티드를 포함하는 코리네박테리움 글루타미쿰 균주; 이를 배양한 배지; 또는 이들 중 2 이상의 조합을 포함하는 L-아미노산 생산용 조성물을 제공하는 것이다.

[0127] 본 출원의 조성물은 아미노산 생산용 조성물에 통상 사용되는 임의의 적합한 부형제를 추가로 포함할 수 있으며, 이러한 부형제는, 예를 들어 보존제, 습윤제, 분산제, 현탁화제, 완충제, 안정화제 또는 등장화제 등일 수 있으나, 이에 한정되는 것은 아니다.

[0128] 본 출원의 조성물에서, 변이체, 폴리뉴클레오티드, 벡터, 균주, 배지 및 L-아미노산 등은 상기 다른 양태에서 기재한 바와 같다.

**발명의 효과**

[0130] 본 출원의 신규한 단백질 변이체를 포함하는, 코리네박테리움 글루타미쿰 균주를 배양하는 경우, 기존 비변형 폴리펩티드를 갖는 미생물에 비해 고수율의 L-라이신 생산이 가능하다.

**발명을 실시하기 위한 구체적인 내용**

[0132] 이하 본 출원을 실시예에 의해 보다 상세하게 설명한다. 그러나 하기 실시예는 본 출원을 예시하기 위한 바람직한 실시양태에 불과한 것이며 따라서, 본 출원의 권리범위를 이에 한정하는 것으로 의도되지는 않는다. 한편, 본 명세서에 기재되지 않은 기술적 사항들은 본 출원의 기술 분야 또는 유사 기술 분야에서 숙련된 통상의 기술자이면 충분히 이해하고 용이하게 실시할 수 있다.

**[0134] 실시예 1: 미생물내 단백질 변이체 발현을 위한 벡터 제작**

[0136] 아미노산 서열(서열번호 3)의 676번째 위치의 아스파르트산(Asp)이 타이로신(Tyr) 아미노산으로 치환된 변이체(D676Y; 서열번호 1)가 L-라이신 생산에 미치는 영향을 확인하고자 이의 발현 균주 제작을 위한 벡터를 코리네

박테리움 염색체 내 유전자의 삽입 및 교체를 위한 플라스미드 pDCM2(대한민국 공개번호 제 10-2020-0136813 호)를 이용하여 하기와 같이 제작하였다.

[0137] 야생형 코리네박테리움 글루타미쿰 ATCC13032의 gDNA(genomic DNA)를 주형으로 서열번호 5 및 6의 서열의 프라이머 쌍과 서열번호 7 및 8의 서열의 프라이머 쌍을 이용하여 각각 PCR을 수행하였다. 상기에서 얻어진 두 단편의 혼합물을 주형으로 서열번호 5 및 서열번호 8의 서열의 프라이머 쌍을 이용하여 다시 오버랩핑(overlapping) PCR을 수행하여 단편을 수득하였다. PCR은 94℃에서 5분간 변성 후, 94℃에서 30초, 55℃에서 30초, 72℃에서 1분 30초를 30회 반복한 후, 72℃에서 5분간 수행하였다. pDCM2 플라스미드는 SmaI을 처리하고 상기에서 수득한 PCR 산물을 퓨전 클로닝하였다. 퓨전 클로닝은 In-Fusion® HD 클로닝 키트(Clontech)를 사용하였다. 결과로 얻은 벡터를 pDCM2-NCg10752(D676Y)라 명명하였다. 본 실시예에서 사용한 프라이머의 서열은 하기 표 1에 기재하였다.

표 1

명칭	서열 (5' ->3' )	서열번호
NCg10752_1F	TGAATTCGAGCTCGGTACCCCGTGAGGAGCTTGGCATC	서열번호 5
NCg10752_2R	GCCTTCAGAACAGGATaTTCGGTGTGCAATTCG	서열번호 6
NCg10752_3F	CGAATTCGACACCGAAtATCCTGTTCTGAAGGC	서열번호 7
NCg10752_4R	GTCGACTCTAGAGATCCCCGATCCTCTCGGGCTG	서열번호 8

[0140] 실시예 2: 단백질 변이체를 발현하는 미생물의 L-라이신 생산능 평가

[0141] 2-1. 단백질 변이체 발현 균주 제작

[0142] 상기 실시예 1에서 제작한 벡터를 코리네박테리움 글루타미쿰 CJ3P(US 9556463 B2)에 형질전환 하였다.

[0143] 형질 전환된 균주들에서 서열번호 9와 10의 서열의 프라이머 쌍을 이용하여 상동성 재조합이 일어난 균주를 선별하고 CJ3P\_NCg10752\_D676Y로 명명하였다. 본 실시예에서 사용한 프라이머의 서열은 하기 표 2에 기재하였다.

[0144]

표 2

명칭	서열 (5' ->3' )	서열번호
NCg10752_5F	GGATCCACTCGTGGTGG	서열번호 9
NCg10752_6R	GATGCGGTGTCTCTAGGC	서열번호 10

[0146] 2-2. 단백질 변이체 발현 균주의 L-라이신 생산능 비교

[0147] 상기 2-1에서 제작된 균주 및 대조군 모균주의 플라스크 발효역가 평가를 통해 L-라이신 생산능을 분석하였다.

[0148] 먼저, 종 배지 25 ml을 함유하는 250 ml 코너-바플 플라스크에 각 균주를 접종하고, 30℃에서 20시간 동안, 200 rpm으로 진탕 배양하였다. 생산 배지 24 ml을 함유하는 250 ml 코너-바플 플라스크에 1 ml의 종 배양액을 접종하고 30℃에서 72시간 동안, 200 rpm으로 진탕 배양하였다. 배양 종료 후 HPLC로 L-라이신의 생산능을 측정하였다. 실험한 각 균주에 대한 배양액 중의 라이신 농도 및 농도 증가율은 하기 표 3과 같다.

[0150] <종배지 (pH 7.0)>

[0151] 포도당 20 g, 펩톤 10 g, 효모추출물 5 g, 요소 1.5 g, KH<sub>2</sub>PO<sub>4</sub> 4 g, K<sub>2</sub>HPO<sub>4</sub> 8g, MgSO<sub>4</sub> · 7H<sub>2</sub>O 0.5 g, 바이오틴 0.1 mg, 티아민 HCl 1 mg, 칼슘-판토텐산 2 mg, 니코틴아미드 2 mg (증류수 1 리터 기준)

[0153] <생산배지 (pH 7.0)>

[0154] 포도당 100 g, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 40 g, 대두 단백질 2.5 g, 옥수수 침지 고형분(Corn Steep Solids) 5 g, 요소 3 g, KH<sub>2</sub>PO<sub>4</sub> 1 g, MgSO<sub>4</sub> · 7H<sub>2</sub>O 0.5 g, 바이오틴 100 μg, 티아민 염산염 1000 μg, 칼슘-판토텐산 2000 μg, 니코틴아미드 3000 μg, CaCO<sub>3</sub> 30 g (증류수 1리터 기준).

[0156] 상기 실험은 3번 반복하였으며, 그 분석 결과의 평균값을 아래 표 3에 나타내었다.

표 3

[0157]

균 주	L-라이신 농도 (g/L)	라이신 농도 증가율 (%)
CJ3P	8.04	-
CJ3P_NCg10752_D676Y	9.56	18.9

[0158]

표 3과 같이, CJ3P\_NCg10752\_D676Y 균주는 대조균에 비해 증가된 L-라이신 생산량을 나타내었다.

[0160]

이상의 설명으로부터, 본 출원이 속하는 기술분야의 당업자는 본 출원이 그 기술적 사상이나 필수적 특징을 변경하지 않고서 다른 구체적인 형태로 실시될 수 있다는 것을 이해할 수 있을 것이다. 이와 관련하여, 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적인 것이 아닌 것으로 이해해야만 한다. 본 출원의 범위는 상기 상세한 설명보다는 후술하는 특허 청구범위의 의미 및 범위 그리고 그 등가 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 출원의 범위에 포함되는 것으로 해석되어야 한다.

서 열 목 록

- <110> CJ CheilJedang Corporation
- <120> Novel protein variant and a method for producing L-lysine using the same
- <130> DPP20211201KR
- <160> 10
- <170> KoPatent In 3.0
- <210> 1
- <211> 985
- <212> PRT
- <213> Artificial Sequence
- <220><223> (D676Y)\_NCg10752\_amino acid
- <400> 1

```

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

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

Ala Tyr Gly Ile Thr Asp Glu Asp Val Thr Tyr Glu Glu Asn Trp Gly  
 340 345 350

Ala Gly Glu Thr Thr Asn Glu Glu Val Ala Ala Asp Ser Ala Thr Ile  
 355 360 365

Ser Asn Ile Arg Leu Leu Asp Pro Gln Ile Leu Ser Pro Thr Phe Thr  
 370 375 380

Gln Gln Gln Gln Leu Arg Asn Phe Tyr Gly Phe Pro Asp Gln Leu Ala  
 385 390 395 400

Met Asp Arg Phe Glu Val Asp Gly Lys Leu Arg Asp Phe Val Val Ala  
 405 410 415

Ala Arg Glu Leu Asp Pro Asn Ala Leu Gln Gln Asn Gln Gln Asp Trp  
 420 425 430

Ile Asn Arg His Thr Val Tyr Thr His Gly Asn Gly Phe Ile Ala Ala  
 435 440 445

Gln Ala Asn Gln Val Asp Glu Val Ala Arg Asp Val Gly Ser Thr Arg  
 450 455 460

Gly Gly Tyr Pro Val Tyr Thr Val Ser Asp Leu Gln Ser Asn Ala Arg  
 465 470 475 480

Ala Ala Glu Ser Glu Asp Ala Glu Glu Leu Gly Ile Lys Val Asp Glu  
 485 490 495

Pro Arg Val Tyr Tyr Gly Pro Leu Ile Ala Ser Ala Thr Asp Gly Ala  
 500 505 510

Asp Tyr Ala Ile Val Gly Asp Thr Gly Asp Gly Pro Val Glu Tyr Asp  
 515 520 525

Thr Asp Thr Ser Ser Tyr Thr Tyr Glu Gly Ala Gly Gly Val Asp Ile  
 530 535 540

Gly Asn Met Val Asn Arg Ala Met Phe Ala Leu Arg Tyr Gln Glu Met  
 545 550 555 560

Asn Met Leu Leu Ser Asp Arg Val Gly Ser Glu Ser Lys Ile Leu Phe  
 565 570 575

Glu Arg Asp Pro Arg Ser Arg Val Glu Lys Val Ala Pro Trp Leu Thr  
 580 585 590

Thr Asp Ser Lys Thr Tyr Pro Thr Val Ile Asp Gly Arg Ile Lys Trp  
 595 600 605  
 Ile Val Asp Gly Tyr Thr Thr Leu Asp Ser Leu Pro Tyr Ser Thr Arg  
 610 615 620  
 Thr Ser Leu Thr Glu Ala Thr Gln Asp Ala Val Met Pro Asp Gly Thr  
 625 630 635 640  
 Pro Gln Pro Leu Ile Thr Asp Arg Val Gly Tyr Ile Arg Asn Ser Val  
 645 650 655  
 Lys Ala Val Val Asp Ala Tyr Asp Gly Thr Val Glu Leu Tyr Glu Phe  
 660 665 670  
  
 Asp Thr Glu Tyr Pro Val Leu Lys Ala Trp Arg Gly Val Phe Pro Asp  
 675 680 685  
 Thr Val Lys Asp Gly Ser Glu Ile Ser Asp Glu Leu Arg Ala His Leu  
 690 695 700  
 Arg Tyr Pro Glu Asp Leu Phe Lys Val Gln Arg Asp Met Leu Ala Lys  
 705 710 715 720  
 Tyr Asn Val Asp Asp Ser Gly Thr Phe Phe Thr Asn Asp Ala Phe Trp  
 725 730 735  
 Ser Val Pro Gly Asp Pro Thr Ala Ala Glu Gly Arg Gln Glu Leu Lys  
 740 745 750  
 Gln Pro Pro Tyr Tyr Val Val Ala Ala Asp Pro Glu Thr Gly Glu Ser  
 755 760 765  
 Ser Phe Gln Leu Ile Thr Pro Phe Arg Gly Leu Gln Arg Glu Tyr Leu  
 770 775 780  
 Ser Ala His Met Ser Ala Ser Ser Asp Pro Val Thr Tyr Gly Glu Ile  
 785 790 795 800  
 Thr Val Arg Val Leu Pro Thr Asp Ser Val Thr Gln Gly Pro Lys Gln  
 805 810 815  
  
 Ala Gln Asp Ala Met Met Ser Ser Asp Gln Val Ala Gln Asp Gln Thr  
 820 825 830  
 Leu Trp Arg Gly Ser Asn Asp Leu His Asn Gly Asn Leu Leu Thr Leu



tggctgaatg gccaggactt tggcgtttcg gatcagcagt tcggcttggga ttatggattc	480
tacgcgtttg atcttccgat gcttcgcctc atcgtgact cactgtcgat gatgttgatc	540
gttgctttcc tgatcgcaat cgttggccat tacctcatgg gtggcattcg cgctggaaac	600
cagatgacgg gccagaagtc ctttgtatcc cgtgggtcgc gcaactcagct tgcggttaact	660
gctggctctgt ggatgcttgt taaggtcgct ggctactggc tggatcgcta tgacctgctg	720
actaaggaaa actcaacctt cacaggtgca agctacaccg acatcaatgc acagctgcca	780
gcgaagatca tcctgctggt tatcgattg ttcgtggcga tcgcgttctt ctctcgatc	840
ttctgaagg acctccgat cccaggtctt gccgtagttc tgatgctgct gagctctgtg	900
atcattggtg ctgcttggcc tttgatgctg gagcgtttct ccgttcagcc aaacctgca	960
gaaaaagaag ccgagtacat ttctcgaac attgaatcga ctcgctatgc ctatggcatc	1020
accgatgaag atgtcaccta cgaggaaaac tggggagcag gggagaccac caatgaagaa	1080
gtcgcgcgctg attctgcaac gatttctaac attcgtctgc tggatccaca gatcctgtca	1140
ccaaccttca cccagcaaca gcagctcgca aacttctac gtttcccaga ccagctggcg	1200
atggaccgct ttgaagtaga tggcaaacctc cgcgactttg ttgtggcagc acgtgagctc	1260
gatccaaacg ccctgcagca aaaccagcag gactggatta accgtcacac tgtttatacc	1320
cacggcaacg gcttcattgc agctcaagca aaccaggtgg atgaggtcgc ccgcgacgtc	1380
ggatccactc gtggtgggta ccctgtctac accgtctctg atttgcagtc gaatgctcgt	1440
gctgcagaaa gcgaagatgc tgaggagctt ggcatcaagg ttgatgagcc tcgtgtgtac	1500
tacggaccac tgattgcttc tgcgactgat ggtgctgact acgcaattgt cggtgacacc	1560
ggcgatggcc cagtcgagta cgacactgac acctccagct acacctacga aggtgctggc	1620
ggcgtggaca ttggaaacat ggtcaaccgt gcgatgtttg cattgcgcta ccaggaatg	1680
aacatgctcc tgtctgatcg tgttggttcc gaatccaaga tcctatttga gcgcatcct	1740
cgttcccgtg tggaaaaggt tgcaccttgg ttgacctg actccaagac ctaccaact	1800
gtgattgatg gtcgcatcaa gtggatcgtc gatggctaca ccaccttga tagtcttccg	1860
tactccacgc gcacctcact gacggaagcg actcaggatg ctgtcatgcc tgacggcacc	1920
ccacagccac tgatcacaga tagggtcggf tacatccga actccgtgaa ggctgttgtt	1980
gatgcgtacg acggaactgt tgaactctac gaattcgaca ccgaatatcc tgttctgaag	2040
gcatggcgtg gcgtgttccc agacaccgtg aaggacgggt cggagatttc cgatgagctt	2100
cgcgcacacc tgcgttacc agaagattg ttcaaggtcc agcgtgacat gctggccaag	2160
tacaacgttg atgattctgg aacattcttc accaacgatg cgttctggtc tgtcccaggt	2220

gaccaactg cagcggaggg ccgccaggaa cttaagcagc ctccttacta cgtgggtggca 2280

gcagaccag agaccggiga gtccagcttc cagctgatca ccccgttccg tggacttcag 2340

cgcgagtacc tctctgcaca catgtctgcg tctgtgatc cagttaccta cgggaaatc 2400

actgttcgtg tgctgcctac cgattctgtg acccagggtc caaagcagc ccaggatgcg 2460

atgatgtcat ctgaccagt tgctcaggac caaacactgt ggcgtggatc gaacgatctg 2520

cacaacggaa acctgttgac cttgccagt ggtggcggag agatcctcta cgttgagccg 2580

attactcgc agcgaagga tcaggcatcg gccttccga agcttctgcg cgtgctggtc 2640

ttctacaagg gtcaggttgg ttacgcacca acgatcgtg aagccctatc gcaggtcggc 2700

attgatccga aggaagcga ggacatcga gaggtagatg gcaccgtac gacgccatcg 2760

actgatgaga ctgacactga cactgatcag cctgcaaccg aaacccaac tgcaccagtg 2820

agtgaggcgg aaggaatcgc ggccatcaac gatgcgttga gcaaccttga agctgctcgc 2880

gatagctctt tcgaagagta tggctgtgca ctgatcgcg ttgatcgtgc cgtcgatagc 2940

taccagtccg cacagtag 2958

<210> 3

<211> 985

<212> PRT

<213> Artificial Sequence

<220><223> (WT)\_NCg10752\_amino acid

<400

> 3

Met Ser Thr Gly Leu Thr Pro Pro Pro Gln Pro Ile Lys Arg Pro Pro

1 5 10 15

Lys Ala Val Thr Trp Ile Phe Ala Ile Ile Ala Leu Val Ile Leu Ile

20 25 30

Ala Pro Met Ser Val Gly Phe Tyr Thr Asp Trp Leu Trp Phe Gly Glu

35 40 45

Val Asp Phe Arg Gly Val Phe Ser Lys Val Ile Val Thr Arg Ile Val

50 55 60

Leu Phe Val Ile Phe Ala Leu Ile Ala Gly Phe Val Thr Trp Leu Ala

65 70 75 80

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

85 90 95

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

Ala Gly Glu Thr Thr Asn Glu Glu Val Ala Ala Asp Ser Ala Thr Ile  
 355 360 365  
 Ser Asn Ile Arg Leu Leu Asp Pro Gln Ile Leu Ser Pro Thr Phe Thr  
 370 375 380  
 Gln Gln Gln Gln Leu Arg Asn Phe Tyr Gly Phe Pro Asp Gln Leu Ala  
 385 390 395 400  
 Met Asp Arg Phe Glu Val Asp Gly Lys Leu Arg Asp Phe Val Val Ala  
 405 410 415  
 Ala Arg Glu Leu Asp Pro Asn Ala Leu Gln Gln Asn Gln Gln Asp Trp  
 420 425 430  
 Ile Asn Arg His Thr Val Tyr Thr His Gly Asn Gly Phe Ile Ala Ala  
 435 440 445  
 Gln Ala Asn Gln Val Asp Glu Val Ala Arg Asp Val Gly Ser Thr Arg  
 450 455 460  
 Gly Gly Tyr Pro Val Tyr Thr Val Ser Asp Leu Gln Ser Asn Ala Arg  
 465 470 475 480  
 Ala Ala Glu Ser Glu Asp Ala Glu Glu Leu Gly Ile Lys Val Asp Glu  
 485 490 495  
 Pro Arg Val Tyr Tyr Gly Pro Leu Ile Ala Ser Ala Thr Asp Gly Ala  
 500 505 510  
 Asp Tyr Ala Ile Val Gly Asp Thr Gly Asp Gly Pro Val Glu Tyr Asp  
 515 520 525  
 Thr Asp Thr Ser Ser Tyr Thr Tyr Glu Gly Ala Gly Gly Val Asp Ile  
 530 535 540  
 Gly Asn Met Val Asn Arg Ala Met Phe Ala Leu Arg Tyr Gln Glu Met  
 545 550 555 560  
 Asn Met Leu Leu Ser Asp Arg Val Gly Ser Glu Ser Lys Ile Leu Phe  
 565 570 575  
 Glu Arg Asp Pro Arg Ser Arg Val Glu Lys Val Ala Pro Trp Leu Thr  
 580 585 590  
 Thr Asp Ser Lys Thr Tyr Pro Thr Val Ile Asp Gly Arg Ile Lys Trp





tacgcgtttg atcttccgat gcttcgcctc atcgtgact cactgtcgat gatgttgatc 540  
 gttgctttcc tgatcgact cgttggccat tacctcatgg gtggcattcg cgctggaaac 600  
 cagatgacgg gccagaagtc ctttgtatcc cgtggtgccg gcactcagct tgcgtaact 660  
 gctggctgtg ggatgcttgt taaggctgct ggctactggc tggatcgcta tgacctgctg 720  
 actaaggaaa actcaacctt cacaggtgca agctacaccg acatcaatgc acagctgcca 780  
 gcgaagatca tctgctggt tategcattg ttcgtggcga tgcgttctt ctctcgatc 840

ttcctgaagg acctccgat cccaggtctt gccgtagttc tgatgctgct gagctctgtg 900  
 atcattggtg ctgcttggcc tttgatgctg gagcgtttct cgttcagcc aaaccgtgca 960  
 gaaaaagaag ccgagtacat ttctcgaac attgaatcga ctcgctatgc ctatggcatc 1020  
 accgatgaag atgtcaccta cgaggaaaac tggggagcag gggagaccac caatgaagaa 1080  
 gtgcgctgtg attctgcaac gatttctaac attcgtctgc tggatccaca gatcctgtca 1140  
 ccaaccttca cccagcaaca gcagctgcga aacttctac gtttccaga ccagctggcg 1200  
 atggaccgct ttgaagtaga tggcaaacctc cgcgactttg ttgtggcagc acgtgagctc 1260

gatccaaacg cctgcagca aaaccagcag gactggatta accgtcacac tgtttatacc 1320  
 cacggcaacg gcttcattgc agctcaagca aaccaggtgg atgaggtcgc ccgcgacgtc 1380  
 ggatccactc gtggtggtta cctgtctac accgtctctg atttgcagtc gaatgctcgt 1440  
 gctgcagaaa gcgaagatgc tgaggagctt ggcatcaagg ttgatgagcc tegtgtgtac 1500  
 tacggaccac tgattgcttc tgcgactgat ggtgctgact acgcaattgt cggtgacacc 1560  
 ggcgatggcc cagtgcagta cgacactgac acctccagct acacctacga aggtgctggc 1620  
 ggcgtggaca ttggaaacat ggtcaaccgt gcgatgtttg cattgcgcta ccaggaaatg 1680

aacatgctcc tgtctgatcg tgttggttcc gaatccaaga tcctatttga gcgcgatcct 1740  
 cgttcccgtg tggaaaaggt tgcaccttgg ttgacctg actccaagac ctaccecaact 1800  
 gtgattgatg gtcgatcaa gtggatcgtc gatggctaca ccaccttga tagtcttccg 1860  
 tactccagc gcacctcact gacggaagcg actcaggatg ctgtcatgcc tgacggcacc 1920  
 ccacagccac tgatcacaga tagggtcggg tacatccgca actccgtgaa ggctgttgtt 1980  
 gatgcttacg acggaactgt tgaactctac gaattcgaca ccgaagatcc tgttctgaag 2040  
 gcatggcgtg gcgtgttccc agacaccgtg aaggacgggt cggagatttc cgatgagctt 2100

cgcgcacacc tgcgttacc agaagatttg ttcaaggtcc agcgtgacat gctggccaag 2160  
 tacaacttg atgattctgg aacattcttc accaacgatg cgttctggtc tgtcccaggt 2220  
 gaccaactg cagcggaggg ccgccaggaa cttaaagcagc ctcttacta cgtgggtgca 2280  
 gcagaccag agaccggtga gtccagcttc cagctgatca ccccgttccg tggacttcag 2340

cgcgagtacc tcctgcaca catgctcgcg tcgtctgac cagttaccta cggatgaaatc 2400  
 actgttcgtg tgctgcctac cgattctgtg acccagggtc caaagcaggc ccaggatgcg 2460  
 atgatgtcat ctgaccaggt tgctcaggac caaacactgt ggcgtggatc gaacgatctg 2520  
  
 cacaacggaa acctgttgac cttgccagtt ggtggcggag agatcctcta cgttgagccg 2580  
 atttactcgc agcgaagga tcaggcatcg gccttcccga agcttctcgc cgtgctggtc 2640  
 ttctacaagg gtcaggttgg ttacgacca acgatcgtg aagccctatc gcaggtcggc 2700  
 attgatccga aggaagcga ggacatcga gaggtagatg gcaccgctac gacgccatcg 2760  
 actgatgaga ctgacctga cactgatcag cctgcaaccg aaacccaac tgcaccagtg 2820  
 agtgaggcgg aaggaatcgc ggccatcaac gatgcgttga gcaaccttga agctgctcgc 2880  
 gatagctctt tcgaagagta tggctgtgca ctcgatgcgc ttgatcgtgc cgtcgatagc 2940  
  
 taccagtccg cacagtag 2958  
  
 <210> 5  
 <211> 38  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> NCg10752\_1F  
 <400> 5  
 tgaattcgag ctcggtaccc gctgaggagc ttggcatc 38  
  
 <210> 6  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> NCg10752\_2R  
 <400> 6  
 gccttcagaa caggatattc ggtgtcgaat tcg 33  
  
 <210> 7  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> NCg10752\_3F  
 <400> 7  
 cgaattcgac accgaatc ctgttctgaa ggc 33

<210> 8  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> NCg10752\_4R  
 <400> 8  
 gtcgactcta gaggatcccc cgatcctctc gggctg 36  
 <210> 9  
 <211> 17  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> NCg10752\_5F  
 <400> 9  
 ggatccactc gtggtgg 17  
 <210> 10  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> NCg10752\_6R  
 <400> 10  
 gatgcggtgt ctctaggc 18