

명세서

발명의 명칭: 신규한 베타-카로틴 15,15'-옥시게네이즈 변이체 및 이를 이용한 레티노이드 생산방법

기술분야

- [1] 본 출원은 신규한 베타-카로틴 15,15'-옥시게네이즈 변이체, 상기 변이체를 포함하는 미생물, 상기 미생물을 이용한 레티노이드 생산방법, 레티노이드 생산용 조성물, 및 상기 변이체 또는 미생물의 레티노이드 생산 용도에 관한 것이다.

[2]

배경기술

- [3] 지용성 비타민 중 하나로서 필수 영양소인 비타민 A 및 이를 포함하는 개념인 레티노이드는 현재 화학합성법으로 생산 및 판매되고 있으나, 원자재 수급 등에 대한 이유로 미생물 발효 기반의 레티노이드를 생산하고자 하는 연구 즉, 고효율 생산 미생물 및 발효공정기술 개발을 위한 다양한 연구들이 수행되고 있다. 예를 들어, 레티노이드 생합성에 관여하는 효소를 코딩하는 유전자의 발현을 증가시키거나 또는 생합성에 불필요한 유전자를 제거하는 것과 같은 목적 물질 특이적 접근 방법이 그것이다. 상기 레티노이드 생합성에 관여하는 효소로는 일 예로, 베타-카로틴 15,15'-옥시게네이즈 등이 있다(US 2020-0277644 A1).
- [4] 그러나, 베타-카로틴 15,15'-옥시게네이즈에 대한 연구는 미미하고 비타민 A의 수요 증가에 따라 효과적인 비타민 A의 생산능 증가를 위한 연구가 여전히 필요한 실정이다.

[5]

발명의 상세한 설명

기술적 과제

- [6] 본 출원의 해결하고자 하는 과제는 베타-카로틴 15,15'-옥시게네이즈 변이체 및 이를 이용한 레티노이드 생산 방법을 제공하는 것이다.

[7]

과제 해결 수단

- [8] 본 출원의 하나의 목적은 베타-카로틴 15,15'-옥시게네이즈 변이체를 제공하는 것이다.
- [9] 본 출원의 다른 하나의 목적은 본 출원의 변이체를 코딩하는 폴리뉴클레오티드를 제공하는 것이다.
- [10] 본 출원의 또 다른 하나의 목적은 본 출원의 변이체 또는 상기 변이체를 코딩하는 폴리뉴클레오티드를 포함하는 미생물을 제공하는 것이다.
- [11] 본 출원의 또 다른 하나의 목적은 본 출원의 변이체 또는 상기 변이체를 코딩하는 폴리뉴클레오티드를 포함하는 미생물을 배지에서 배양하는 단계를

포함하는, 레티노이드 생산방법을 제공하는 것이다.

[12] 본 출원의 또 다른 하나의 목적은 본 출원의 변이체, 상기 변이체를 코딩하는 폴리뉴클레오티드를 포함하는 미생물, 또는 이의 배양물을 포함하는 레티노이드 생산용 조성물을 제공하는 것이다.

[13] 본 출원의 또 다른 하나의 목적은 본 출원의 변이체, 상기 변이체를 코딩하는 폴리뉴클레오티드, 또는 이를 포함하는 미생물 중 어느 하나 이상의 레티노이드 생산 용도를 제공하는 것이다.

[14]

발명의 효과

[15] 본 출원의 베타-카로틴 15,15'-옥시게네이즈 변이체를 포함하는 미생물을 배양하는 경우, 기존 비변형 폴리펩티드를 갖는 미생물에 비해 고수율의 레티노이드 생산이 가능하다.

[16]

발명의 실시를 위한 최선의 형태

[17] 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.

[18]

[19] 본 출원의 하나의 양태는 서열번호 1의 아미노산 서열의 N-말단으로부터 43번 위치에 상응하는 아미노산인 알라닌, 58번 위치에 상응하는 아미노산인 류신, 및 74번 위치에 상응하는 아미노산인 이소류신 중 어느 하나 이상의 아미노산이 다른 아미노산으로 치환된, 베타-카로틴 15,15'-옥시게네이즈 변이체를 제공한다.

[20]

상기 베타-카로틴 15,15'-옥시게네이즈 변이체는 베타-카로틴 15,15'-옥시게네이즈 활성을 갖는 폴리펩티드 또는, 베타-카로틴 15,15'-옥시게네이즈에서 서열번호 1의 N-말단으로부터 43번, 58번, 및 74번 위치에 상응하는 아미노산 중 어느 하나 이상의 아미노산이 다른 아미노산으로 치환된 변이체일 수 있다. 구체적으로, 상기 43번, 58번, 및 74번 위치 중 1 이상, 2 이상, 또는 3 이상의 위치의 아미노산이 치환된 것일 수 있으나, 이에 제한되지 않는다.

[21]

상기 변이체의 모체가 되는 베타-카로틴 15,15'-옥시게네이즈 활성을 갖는 폴리펩티드 또는 베타-카로틴 15,15'-옥시게네이즈에 있어서, 서열번호 1의

아미노산 서열의 N-말단으로부터 43번 위치에 상응하는 아미노산은 알라닌; 58번 위치에 상응하는 아미노산은 류신; 및/또는 74번 위치에 상응하는 아미노산은 이소류신일 수 있으나, 이에 제한되지 않는다.

- [22] 일 구현 예로, 상기 변이체는 서열번호 1의 아미노산 서열의 N-말단으로부터 43번 위치에 상응하는 아미노산이 알라닌 외의 아미노산으로 치환; 58번 위치에 상응하는 아미노산이 류신 외의 아미노산으로 치환; 74번 위치에 상응하는 아미노산이 이소류신 외의 아미노산으로 치환 중 어느 하나 이상의 치환을 포함하는 것일 수 있으나, 이에 제한되지 않는다.
- [23] 상기 "다른 아미노산"은 치환 전 아미노산과 치환 전의 아미노산과 다른 아미노산이면 제한되지 않는다. 한편, 본 출원에서 '특정 아미노산이 치환되었다'고 표현하는 경우, 다른 아미노산으로 치환되었다고 별도로 표기하지 않더라도 치환 전의 아미노산과 다른 아미노산으로 치환되는 것임은 자명하다.
- [24] 상기 다른 아미노산으로의 치환은 비극성(nonpolar) 아미노산, 극성(polar) 아미노산, 또는 양으로 하전된(염기성) 아미노산으로의 치환일 수 있다. 구체적으로, 상기 비극성(nonpolar) 아미노산은 글리신, 알라닌, 발린, 류신, 이소류신, 페티오닌, 및 프롤린으로 이루어지는 군에서 선택되는 것일 수 있으나, 이에 제한되지 않는다. 상기 극성(polar) 아미노산은 친수성(hydrophilic) 아미노산과 혼용될 수 있으며, 세린, 트레오닌, 시스테인, 티로신, 아스파라긴 및 글루타민으로 이루어진 군에서 선택되는 것일 수 있고, 상기 양으로 하전된(염기성) 아미노산은 아르기닌, 라이신, 및 히스티딘으로 이루어진 군에서 선택되는 것일 수 있으나, 이에 제한되지 않는다.
- [25] 본 출원의 변이체는 참조(reference) 단백질인 서열번호 1의 아미노산 서열의 43번, 58번, 및 74번 위치에 상응하는 아미노산 중 어느 하나 이상의 아미노산이 비극성 아미노산, 극성 아미노산, 또는 양으로 하전된(염기성) 아미노산 중 치환 전 아미노산과 다른 아미노산으로 치환된 변이체일 수 있으나, 이에 제한되지 않는다.
- [26] 일 구현 예로, 상기 43번 위치에 상응하는 아미노산의 다른 아미노산으로의 치환은 비극성 아미노산; 상기 58번 위치에 상응하는 아미노산의 다른 아미노산으로의 치환은 극성 아미노산; 및 74번 위치에 상응하는 아미노산의 다른 아미노산으로의 치환은 비극성 아미노산으로의 치환일 수 있으나, 이에 제한되지 않는다.
- [27] 상기 비극성 아미노산은 발린일 수 있고, 상기 극성 아미노산은 글루타민일 수 있으나, 이에 제한되지 않는다.
- [28] 상기 다른 아미노산은 발린 또는 글루타민일 수 있으나, 이에 제한되지 않는다.
- [29] 일 구현 예로, 본 출원의 변이체는 서열번호 1의 아미노산 서열의 43번 위치에 상응하는 아미노산이 발린으로 치환; 58번 위치에 상응하는 아미노산이 글루타민으로 치환; 74번 위치에 상응하는 아미노산이 발린으로 치환; 및 이들의

조합으로 구성되는 군에서 선택되어 치환된 것일 수 있으나, 이에 제한되지 않는다.

- [30] 일 구현 예로, 본 출원의 변이체는 서열번호 1의 아미노산 서열의 N-말단으로부터 i) 43번 위치에 상응하는 아미노산인 알라닌이 발린으로 치환; ii) 58번 위치에 상응하는 아미노산인 류신이 글루타민으로 치환; iii) 74번 위치에 상응하는 아미노산인 이소류신이 발린으로 치환; iv) 43번 위치에 상응하는 아미노산인 알라닌이 발린으로 치환되고 58번 위치에 상응하는 아미노산인 류신이 글루타민으로 치환; v) 58번 위치에 상응하는 아미노산인 류신이 글루타민으로 치환되고 74번 위치에 상응하는 아미노산인 이소류신이 발린으로 치환; vi) 43번 위치에 상응하는 아미노산인 알라닌이 발린으로 치환되고 74번 위치에 상응하는 아미노산인 이소류신이 발린으로 치환; 또는 vii) 43번 위치에 상응하는 아미노산인 알라닌이 발린으로 치환, 58번 위치에 상응하는 아미노산인 류신이 글루타민으로 치환, 및 74번 위치에 상응하는 아미노산인 이소류신이 발린으로 치환된 것일 수 있으나, 이에 제한되지 않는다.
- [31] 일 구현 예로, 본 출원의 변이체는 서열번호 1의 아미노산 서열의 43번 위치에 상응하는 아미노산인 알라닌이 발린으로 치환, 58번 위치에 상응하는 아미노산인 류신이 글루타민으로 치환, 및 74번 위치에 상응하는 아미노산인 이소류신이 발린으로 치환된, 서열번호 3으로 기재된 아미노산 서열을 포함하는 폴리펩티드일 수 있으나, 이에 제한되지 않는다.
- [32]
- [33] 본 출원의 변이체는 서열번호 3으로 기재된 아미노산 서열을 가지거나, 포함하거나, 이루어지거나, 상기 아미노산 서열로 필수적으로 이루어질(essentially consisting of) 수 있다.
- [34]
- [35] 또한, 본 출원의 변이체는 상기 서열번호 3으로 기재된 아미노산 서열에서 서열번호 1의 아미노산 서열을 기준으로 43번 위치에 상응하는 아미노산은 발린, 58번 위치에 상응하는 아미노산은 글루타민, 및 74번 위치에 상응하는 아미노산은 발린으로 고정되고, 상기 서열번호 3으로 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 각각의 서열의 고정된 위치 외에 상기 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.
- [36] 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.

- [37] 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 활성화에 거의 영향을 미치지 않거나 또는 영향을 미치지 않을 수 있다.
- [38]
- [39] 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 일부분이 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.
- [40] 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 컨쥬게이트 될 수 있다. 또한 상기 변이체는 확인, 정제, 또는 합성할 수 있도록 다른 서열 또는 링커와 컨쥬게이트 될 수 있다.
- [41]
- [42] 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.
- [43] 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.

- [44] 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.
- [45] 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.
- [46]
- [47] 본 출원의 일 예로, 본 출원의 변이체는 베타-카로틴 15,15'-옥시게네이스 효소 활성을 가질 수 있다. 또한, 본 출원의 변이체는 베타-카로틴 15,15'-옥시게네이스 활성을 갖는 야생형 폴리펩티드에 비해 레티노이드(retinoid) 생산능이 증가되도록 하는 활성을 가질 수 있다.
- [48] 본 출원에서 용어, "베타-카로틴 15,15'-옥시게네이스(beta-carotene 15,15'-oxygenase)"는 한 분자의 베타-카로틴(beta-carotene)을 절단하여 두 분자의 레티날(retinal)로 전환시키는 작용을 촉매하는 기능을 갖는 것이다.
- [49] 본 출원의 베타-카로틴 15,15'-옥시게네이스는 본 출원에서 제공하는

- 베타-카로틴 15,15'-옥시게네이즈 변이체를 제조하기 위해 변형이 가해지는 베타-카로틴 15,15'-옥시게네이즈 또는 베타-카로틴 15,15'-옥시게네이즈 활성을 갖는 폴리펩티드일 수 있다. 구체적으로, 자연적으로 발생하는 폴리펩티드 또는 야생형 폴리펩티드일 수 있고, 이의 성숙 폴리펩티드일 수 있으며, 이의 변이체 또는 기능적 단편을 포함할 수 있으나 본 출원의 베타-카로틴 15,15'-옥시게네이즈 변이체의 모체(parent)가 될 수 있는 한 제한 없이 포함된다.
- [50] 본 출원에서 상기 베타-카로틴 15,15'-옥시게네이즈는 이에 제한되지는 않으나, 서열번호 1의 폴리펩티드일 수 있다. 또한, 서열번호 1의 폴리펩티드와 약 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 폴리펩티드일 수 있으며, 서열번호 1의 아미노산 서열로 이루어진 폴리펩티드와 동일 혹은 상응하는 활성을 가지는 경우라면 제한 없이 베타-카로틴 15,15'-옥시게네이즈 범위에 포함된다.
- [51] 구체적으로, 본 출원의 베타-카로틴 15,15'-옥시게네이즈는 BCO 또는 Blh로 혼용하여 사용될 수 있다. 본 출원에서 상기 베타-카로틴 15,15'-옥시게네이즈는 공지의 데이터 베이스인 NCBI의 GenBank에서 그 서열을 얻을 수 있다. 구체적으로 *blh*에 의해 코딩되는 베타-카로틴 15,15'-옥시게네이즈 활성을 갖는 폴리펩티드일 수 있으나, 이에 제한되지 않는다.
- [52]
- [53] 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.
- [54] 예를 들어, 임의의 아미노산 서열을 서열번호 1과 정렬(align)하고, 이를 토대로 상기 아미노산 서열의 각 아미노산 잔기는 서열번호 1의 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.
- [55] 이러한 정렬에는 예를 들어 Needleman-Wunsch 알고리즘(Needleman 및 Wunsch, 1970, J. Mol. Biol. 48: 443-453), EMBOSS 패키지의 Needleman 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 서열 정렬 프로그램, 쌍 서열(pairwise sequence) 비교 알고리즘 등을 적절히 사용할 수 있다.
- [56]
- [57] 본 출원의 다른 하나의 양태는 본 출원의 변이체를 코딩하는

폴리뉴클레오티드를 제공한다.

- [58] 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.
- [59] 본 출원의 변이체를 코딩하는 폴리뉴클레오티드는 일 예로 서열번호 3으로 기재된 아미노산 서열을 코딩하는 염기서열을 포함할 수 있다. 본 출원의 일 예로, 본 출원의 폴리뉴클레오티드는 서열번호 4의 서열을 가지거나 포함할 수 있다. 또한, 본 출원의 폴리뉴클레오티드는 서열번호 4의 서열로 이루어지거나, 필수적으로 구성될 수 있다.
- [60] 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 4와 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 4의 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.
- [61] 이때, 상기 상동성 또는 동일성을 갖는 서열에서, 서열번호 3의 43번 위치에 상응하는 아미노산을 코딩하는 코돈은, 발린을 코딩하는 코돈 중 하나이고, 58번 위치에 상응하는 아미노산을 코딩하는 코돈은, 글루타민을 코딩하는 코돈 중 하나이며, 74번 위치에 상응하는 아미노산을 코딩하는 코돈은, 발린을 코딩하는 코돈 중 하나일 수 있다.
- [62] 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°C, 1XSSC, 0.1% SDS, 구체적으로 60°C, 0.1XSSC, 0.1% SDS, 보다 구체적으로 68°C 0.1XSSC, 0.1% SDS에 상당하는 염 농도 및 온도에서, 1회, 구체적으로 2회 내지 3회 세정하는 조건을 열거할 수 있다.
- [63] 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.
- [64] 구체적으로, 본 출원의 폴리뉴클레오티드와 상동성 또는 동일성을 가지는 폴리뉴클레오티드는 55°C의 T_m 값에서 혼성화 단계를 포함하는 혼성화 조건을 사용하고 상술한 조건을 사용하여 탐지할 수 있다. 또한, 상기 T_m 값은 60°C, 63°C 또는 65°C일 수 있으나, 이에 제한되는 것은 아니고 그 목적에 따라 당업자에 의해 적절히 조절될 수 있다.
- [65] 상기 폴리뉴클레오티드를 혼성화하는 적절한 엄격도는 폴리뉴클레오티드의 길이 및 상보성 정도에 의존하고 변수는 해당기술분야에 잘 알려져 있다(예컨대, J. Sambrook et al., 상동).
- [66]
- [67] 본 출원의 또 다른 하나의 양태는 본 출원의 폴리뉴클레오티드를 포함하는 벡터를 제공한다. 상기 벡터는 상기 폴리뉴클레오티드를 숙주세포에서 발현시키기 위한 발현 벡터일 수 있으나, 이에 제한되지 않는다.
- [68] 본 출원의 벡터는 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 세포와 무관하게 복제되거나 기능할 수 있으며, 게놈 그 자체에 통합될 수 있다.
- [69] 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(대한민국 공개특허공보 제10-2020-0136813호), pACYC177, pACYC184, pCL, pECCG117, pUC19, pBR322, pMW118, pCC1BAC, pIMR53 벡터 등을 사용할 수 있다.

- [70] 일레로 세포 내 염색체 삽입용 벡터를 통해 목적 폴리펩티드를 코딩하는 폴리뉴클레오티드를 염색체 내로 삽입할 수 있다. 상기 폴리뉴클레오티드의 염색체 내로의 삽입은 당업계 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.
- [71] 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.
- [72] 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.
- [73]
- [74] 본 출원의 또 다른 하나의 양태는 본 출원의 변이체 또는 본 출원의 폴리뉴클레오티드를 포함하는 미생물을 제공한다.

- [75] 본 출원의 미생물은 본 출원의 변이체, 상기 변이체를 코딩하는 폴리뉴클레오티드, 및 상기 폴리뉴클레오티드를 포함하는 벡터 중 어느 하나 이상을 포함하는 것일 수 있다. 본 출원의 미생물은 본 출원의 변이형 폴리펩티드, 상기 폴리펩티드를 암호화하는 폴리뉴클레오티드, 또는 본 출원의 폴리뉴클레오티드를 포함하는 벡터를 포함할 수 있다.
- [76] 본 출원에서 용어, "균주" 또는 "미생물"은 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.
- [77] 본 출원의 미생물은 본 출원의 변이체, 본 출원의 폴리뉴클레오티드 및 본 출원의 폴리뉴클레오티드를 포함하는 벡터 중 어느 하나 이상을 포함하는 미생물; 본 출원의 변이체 또는 본 출원의 폴리뉴클레오티드를 발현하도록 변형된 미생물; 본 출원의 변이체 또는 본 출원의 폴리뉴클레오티드를 발현하는 미생물(예컨대, 재조합 미생물); 또는 본 출원의 변이체 활성을 갖는 미생물(예컨대, 재조합 미생물)일 수 있으나, 이에 제한되지 않는다. 상기 미생물은 라이코펜 사이클라제/파이토엔 신타아제(lycopene cyclase/phytoene synthase, crtYB) 및 파이토엔 디세튜라아제(phytoene desaturase, crtI) 단백질을 코딩하는 폴리뉴클레오티드를 내재적으로 포함하거나 상기 단백질을 코딩하는 폴리뉴클레오티드를 더 포함하도록 변형되어 이들 단백질 활성을 나타내거나 활성이 강화된 미생물일 수 있다. 또한 상기 미생물은 라이코펜 사이클라제/파이토엔 신타아제(lycopene cyclase/phytoene synthase, crtYB) 및 파이토엔 디세튜라아제(phytoene desaturase, crtI) 단백질의 활성이 없는 모균주가 상기 단백질을 코딩하는 폴리뉴클레오티드를 포함하도록 변형되어, 이들 단백질 활성을 나타내는 미생물일 수 있다. 상기 라이코펜 사이클라제/파이토엔 신타아제 또는 파이토엔 디세튜라아제는 크산토피로마이세스 덴드로하우스(*Xanthophyllomyces dendrorhous*) 유래의 단백질일 수 있으나 이에 제한되지 않는다. 일 구체 예로 상기 라이코펜 사이클라제/파이토엔 신타아제 또는 파이토엔 디세튜라아제를 코딩하는 폴리뉴클레오티드는 각각 서열번호 13 또는 서열번호 14의 서열을 가지거나 포함하는 것일 수 있다. 상기 폴리뉴클레오티드는 코돈의 축퇴성(degeneracy) 또는 본 출원의 변이체를 발현시키고자 하는 미생물에서 선호되는 코돈을 고려하여, 아미노산 서열을 변화시키지 않는 범위 내에서 코딩 영역에 다양한 변형이 이루어질 수 있다. 구체적으로, 상기 폴리뉴클레오티드는 서열번호 13 또는 서열번호 14의 서열과 상동성 또는 동일성이 70% 이상, 80% 이상, 85% 이상, 90% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상, 및 100% 미만인 염기서열을 가지거나 포함하거나, 또는 서열번호 13 또는 서열번호 14의 서열과 상동성 또는 동일성이 70% 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.

[78]

[79] 본 출원의 미생물은 레티노이드 생산능을 가진 것일 수 있고, 레티놀 생산능을 가진 것일 수 있다.

[80]

본 출원에서 "레티노이드"는 비타민 A로서의 작용을 할 수 있거나, 이와 관련된 물질을 의미한다. 구체적으로, 상기 레티노이드는 레티놀(retinol), 레티날(retinal), 레티닐 에스터(retinyl ester), 및 레티노산(retinoic acid)으로 구성되는 군에서 선택되는 어느 하나 이상일 수 있다. 본 출원의 미생물은 레티날 및 레티놀 생산능을 가진 미생물일 수 있으며, 레티노산은 레티놀로부터 전환되어 생산될 수 있다.

[81]

본 출원의 미생물은 자연적으로 베타-카로틴 15,15'-옥시게네이즈 또는 레티노이드 생산능을 가지고 있는 미생물, 또는 베타-카로틴 15,15'-옥시게네이즈 또는 레티노이드 생산능이 없는 모균주에 본 출원의 변이체 또는 이를 코딩하는 폴리뉴클레오티드 (또는 상기 폴리뉴클레오티드를 포함하는 벡터)가 도입되거나 및/또는 레티노이드 생산능이 부여된 미생물일 수 있으나 이에 제한되지 않는다.

[82]

일 예로, 본 출원의 미생물은 본 출원의 폴리뉴클레오티드 또는 본 출원의 변이체를 코딩하는 폴리뉴클레오티드를 포함하는 벡터로 형질전환되어, 본 출원의 변이체를 발현하는 세포 또는 미생물로서, 본 출원의 목적상 본 출원의 미생물은 본 출원의 변이체를 포함하여 레티노이드를 생산할 수 있는 미생물을 모두 포함할 수 있다. 예를 들어, 본 출원의 미생물은 천연의 야생형 미생물 또는 레티노이드를 생산하는 미생물에 본 출원의 변이체를 코딩하는 폴리뉴클레오티드가 도입됨으로써 베타-카로틴 15,15'-옥시게네이즈 변이체가 발현되어, 레티노이드 생산능이 증가된 재조합 미생물일 수 있다. 상기 레티노이드 생산능이 증가된 재조합 미생물은, 천연의 야생형 미생물 또는 이중기능성 베타-카로틴 15,15'-옥시게네이즈 비변형 미생물(즉, 야생형 베타-카로틴 15,15'-옥시게네이즈 (서열번호 1)를 발현하는 미생물 또는 변이형 (서열번호 3 또는 상기에서 설명한 베타-카로틴 15,15'-옥시게네이즈 변이체) 단백질을 발현하지 않는 미생물)에 비하여 레티노이드 생산능이 증가된 미생물일 수 있으나, 이에 제한되는 것은 아니다. 그 예로, 상기 레티노이드 생산능의 증가 여부를 비교하는 대상 미생물인, 베타-카로틴 15,15'-옥시게네이즈 비변형 미생물은 CC08-1023 균주일 수 있으나, 이에 제한되지 않는다.

[83]

일 예로, 상기 생산능이 증가된 재조합 미생물은 변이 전 모균주 또는 비변형 미생물의 레티노이드 생산능에 비하여 약 1% 이상, 구체적으로는 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 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 20% 이상, 약 22% 이상, 약 24% 이상, 약 26% 이상, 약 28% 이상, 약 30% 이상, 약 32% 이상, 약 34% 이상, 약 36% 이상, 약 37% 이상, 약 38% 이상, 약 39% 이상, 약 40% 이상, 약 41% 이상, 약 41.5% 이상, 약 42% 이상, 약 42.5% 이상, 약 43% 이상, 약 43.5% 이상, 약 44% 이상, 약 44.5% 이상, 약 45% 이상, 약 45.5% 이상, 또는 약 46% 이상 (상한값은 특별한 제한은 없으며, 예컨대, 약 200% 이하, 약 150% 이하, 약 100% 이하, 또는 약 50% 이하일 수 있음) 증가된 것일 수 있으나, 변이 전 모균주 또는 비변형 미생물의 생산능에 비해 +값의 증가량을 갖는 한, 이에 제한되지 않는다. 다른 예에서, 상기 생산능이 증가된 재조합 미생물은 변이 전 모균주 또는 비변형 미생물에 비하여, 레티노이드 생산능이 약 1.01배 이상, 약 1.02배 이상, 약 1.03배 이상, 약 1.05배 이상, 약 1.06배 이상, 약 1.07배 이상, 약 1.08배 이상, 약 1.09배 이상, 약 1.10배 이상, 약 1.11배 이상, 약 1.12배 이상, 약 1.13배 이상, 약 1.14배 이상, 약 1.15배 이상, 약 1.16배 이상, 약 1.17배 이상, 약 1.18배 이상, 약 1.20배 이상, 약 1.22배 이상, 약 1.24배 이상, 약 1.26배 이상, 약 1.28배 이상, 약 1.30배 이상, 약 1.32배 이상, 약 1.34배 이상, 약 1.36배 이상, 약 1.38배 이상, 약 1.40배 이상, 약 1.41배 이상, 약 1.42배 이상, 약 1.43배 이상, 약 1.44배 이상, 약 1.45배 이상, 또는 약 1.46배 이상 (상한값은 특별한 제한은 없으며, 예컨대, 약 10배 이하, 약 5배 이하, 약 3배 이하, 또는 약 2배 이하일 수 있음) 증가된 것일 수 있으나, 이에 제한되지 않는다. 상기 용어 “약(about)”은 ± 0.5 , ± 0.4 , ± 0.3 , ± 0.2 , ± 0.1 등을 모두 포함하는 범위로, 약 이란 용어 뒤에 나오는 수치와 동등하거나 유사한 범위의 수치를 모두 포함하나, 이에 제한되지 않는다.

- [84] 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 베타-카로틴 15,15'-옥시게네이즈 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.
- [85] 본 출원의 또 다른 일 예로, 본 출원의 미생물은 야로위아(*Yarrowia sp.*) 속 미생물일 수 있고, 구체적으로 야로위아 리폴리티카(*Yarrowia lipolytica*)일 수 있으나, 이에 제한되지 않는다.
- [86]
- [87] 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.

- [88] 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.
- [89] 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).
- [90] 구체적으로, 본 출원의 폴리펩티드 활성의 강화는
- [91] 1) 폴리펩티드를 코딩하는 폴리뉴클레오티드의 세포 내 카피수 증가;
- [92] 2) 폴리펩티드를 코딩하는 염색체상의 유전자 발현조절영역을 활성이 강력한 서열로 교체;
- [93] 3) 폴리펩티드를 코딩하는 유전자 전사체의 개시코돈 또는 5'-UTR 지역을 코딩하는 염기서열의 변형;
- [94] 4) 폴리펩티드 활성이 강화되도록 상기 폴리펩티드의 아미노산 서열의 변형;
- [95] 5) 폴리펩티드 활성이 강화되도록 상기 폴리펩티드를 코딩하는 폴리뉴클레오티드 서열의 변형(예를 들어, 폴리펩티드의 활성이 강화되도록 변형된 폴리펩티드를 코딩하도록 상기 폴리펩티드 유전자의 폴리뉴클레오티드 서열의 변형);
- [96] 6) 폴리펩티드의 활성을 나타내는 외래 폴리펩티드 또는 이를 코딩하는 외래 폴리뉴클레오티드의 도입;
- [97] 7) 폴리펩티드를 암호화하는 폴리뉴클레오티드의 코돈 최적화;
- [98] 8) 폴리펩티드의 삼차구조를 분석하여 노출 부위를 선택하여 변형하거나 화학적으로 수식; 또는
- [99] 9) 상기 1) 내지 8) 중 선택된 2 이상의 조합일 수 있으나, 이에, 특별히 제한되는 것은 아니다.

- [100] 보다 구체적으로,
- [101] 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.
- [102] 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, 오퍼레이터 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.
- [103] 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, TEFInt 프로모터, EXP1 프로모터, TEF1 프로모터 등이 있으나, 이에 제한되지 않는다.
- [104] 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.
- [105] 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 상동제조합에 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.

- [106] 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.
- [107] 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.
- [108] 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.
- [109] 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.
- [110]
- [111] 본 출원의 미생물에서 폴리뉴클레오티드의 일부 또는 전체의 변형은 (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.
- [112]
- [113] 본 출원의 미생물에서, 변이체, 폴리뉴클레오티드 및 레티노이드 등은 상기 다른 양태에서 기재한 바와 같다.
- [114]

- [115] 본 출원의 또 다른 하나의 양태는 본 출원의 변이체 또는 본 출원의 폴리뉴클레오티드를 포함하는 미생물을 배지에서 배양하는 단계를 포함하는, 레티노이드 생산방법을 제공한다.
- [116] 본 출원의 레티노이드 생산방법은 본 출원의 변이체 또는 본 출원의 폴리뉴클레오티드 또는 본 출원의 벡터를 포함하는 미생물을 배지에서 배양하는 단계를 포함할 수 있다.
- [117] 본 출원에서, 용어 "배양"은 본 출원의 미생물을 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.
- [118] 본 출원에서 용어, "배지"는 본 출원의 미생물을 배양하기 위해 필요로 하는 영양물질을 주성분으로 혼합한 물질을 의미하며, 생존 및 발육에 불가결한 물을 비롯하여 영양물질 및 발육인자 등을 공급한다. 구체적으로, 본 출원의 미생물의 배양에 사용되는 배지 및 기타 배양 조건은 통상의 미생물의 배양에 사용되는 배지라면 특별한 제한 없이 어느 것이나 사용할 수 있으나, 본 출원의 미생물을 적당한 탄소원, 질소원, 인원, 무기화합물, 아미노산 및/또는 비타민 등을 함유한 통상의 배지 내에서 호기성 조건 하에서 온도, pH 등을 조절하면서 배양할 수 있다.
- [119] 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, 리신 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.
- [120] 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, 카세인 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.
- [121] 상기 인원으로는 인산 제1칼륨, 인산 제2칼륨, 또는 이에 대응되는 소듐-함유 염 등이 포함될 수 있다. 무기화합물로는 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.

- [122] 또한, 본 출원의 미생물의 배양 중에 수산화암모늄, 수산화칼륨, 암모니아, 인산, 황산 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.
- [123] 본 출원의 배양에서 배양온도는 20 내지 45°C 구체적으로는 25 내지 40°C, 20 내지 35°C, 또는 25 내지 35°C를 유지할 수 있고, 약 10 내지 160 시간 동안 배양할 수 있으나, 이에 한정되는 것은 아니다.
- [124] 본 출원의 배양에 의하여 생산된 레티노이드는 배지 중으로 분비되거나 세포 내에 잔류할 수 있다.
- [125]
- [126] 본 출원의 레티노이드 생산방법은, 본 출원의 미생물을 준비하는 단계, 상기 미생물을 배양하기 위한 배지를 준비하는 단계, 또는 이들의 조합(순서에 무관, in any order)을, 예를 들어, 상기 배양하는 단계 이전에, 추가로 포함할 수 있다.
- [127] 본 출원의 레티노이드 생산방법은, 상기 배양에 따른 배지(배양이 수행된 배지) 또는 본 출원의 미생물로부터 레티노이드를 회수하는 단계를 추가로 포함할 수 있다. 상기 회수하는 단계는 상기 배양하는 단계 이후에 추가로 포함될 수 있다.
- [128] 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 레티노이드를 수집(collect)하는 것일 수 있다. 예를 들어, 원심분리, 여과, 결정화 단백질 침전제에 의한 처리(염석법), 추출, 세포 파쇄, 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 레티노이드를 회수할 수 있다.
- [129] 또한, 본 출원의 레티노이드 생산방법은, 추가적으로 정제 단계를 포함할 수 있다. 상기 정제는 당해 기술분야에 공지된 적합한 방법을 이용하여, 수행할 수 있다. 일 예에서, 본 출원의 레티노이드 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.

- [130] 본 출원에서 레티노이드의 일종인 레티놀 유도체 생산방법은, 상기 배양하는 단계는 레티놀을 레티놀 유도체로 전환하는 단계를 추가로 포함할 수 있다. 본 출원의 레티놀 유도체 생산방법에 있어서, 상기 전환하는 단계는 상기 배양하는 단계 또는 상기 회수하는 단계 이후에 추가로 포함될 수 있다. 상기 전환하는 단계는 당해 기술 분야에 공지된 적합한 방법을 이용하여 수행할 수 있다(Jang et al. Microbial Cell Factories 2011, 10:59, Choi et al. antioxidants. 2020).
- [131]
- [132] 본 출원의 방법에서, 변이체, 폴리뉴클레오티드, 벡터 및 미생물 등은 상기 다른 양태에서 기재한 바와 같다.
- [133]
- [134] 본 출원의 또 다른 하나의 양태는 본 출원의 변이체, 상기 변이체를 코딩하는 폴리뉴클레오타이드, 상기 폴리뉴클레오타이드를 포함하는 벡터 또는 본 출원의 폴리뉴클레오티드를 포함하는 미생물; 이를 배양한 배지; 또는 이들 중 2 이상의 조합을 포함하는 레티노이드 생산용 조성물을 제공한다.
- [135] 본 출원의 조성물은 레티노이드 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.
- [136] 본 출원의 조성물에서, 변이체, 폴리뉴클레오티드, 벡터, 미생물, 배지 및 레티노이드 등은 상기 다른 양태에서 기재한 바와 같다.
- [137]
- [138] 본 출원의 또 다른 하나의 양태는 본 출원의 변이체, 폴리뉴클레오티드, 벡터, 및 미생물 중 어느 하나 이상의 레티노이드 생산 용도를 제공한다.
- [139] 상기 변이체, 폴리뉴클레오티드, 벡터, 미생물, 및 레티노이드 등은 상기 다른 양태에서 기재한 바와 같다.
- [140]
- 발명의 실시를 위한 형태**
- [141] 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 기술적인 사항들은 본 출원의 기술 분야 또는 유사 기술 분야에서 숙련된 통상의 기술자이면 충분히 이해하고 용이하게 실시할 수 있다.
- [142]
- [143] **실시예 1. BCO 변이 벡터 라이브러리 구축**
- [144] 해양세균 66A03(Uncultured marine bacterium 66A03) 유래의 BCO 유전자를 증폭하기 위해 UniProtKB (UniProt Knowledgebase)에 등록되어 있는 아미노산 서열(Q4PNI0)에 근거하여 서열번호 1의 아미노산 서열 및 서열번호 2의

폴리뉴클레오티드 서열을 확보하였다. Error-prone PCR을 통해 random 변이를 유도하기 위해 diversify PCR random mutagenesis kit(Takara社)를 사용하였다. PCR 조건 및 방법은 최적화를 통해 user manual의 buffer 조건 1 및 2로 각각 진행하였다. 구체적인 Error-prone PCR 조건은 하기 표 1에 나타내었다.

- [145] 마크로젠社에 합성 의뢰한 해양세균 66A03유래의 야생형 BCO 유전자(서열번호 1)를 주형으로 하고, 서열번호 5와 서열번호 6의 프라이머를 이용하여 PCR을 수행하였다. PCR 조건은 변성 95°C, 30초; 어닐링 55°C 30초; 및 중합반응 68°C 1분을 25회 반복 수행하였다. PCR로 증폭된 2 가지 조건의 BCO PCR products를 혼합하여 pIMR53-TEFINtp-CYC1t 벡터 기반 BCO 변이 벡터 라이브러리를 제작하였다.

[146]

[147] [표1]

	Buffer 조건 1	Buffer 조건 2
PCR grade water	40 μ l	39 μ l
10X Titanium Taq buffer	5 μ l	5 μ l
MnSO ₄	0 μ l	1 μ l
dGTP	1 μ l	1 μ l
50X Diversify dNTP Mix	1 μ l	1 μ l
Primer mix	1 μ l	1 μ l
Template DNA	1 μ l	1 μ l
Titanium Taq polymerase	1 μ l	1 μ l

- [148] pIMR53-TEFINtp-CYC1t 벡터는 하기와 같은 방법으로 제작되었다.

- [149] TEFINtp를 증폭하기 위해 야로위아 리폴리티카 (*Yarrowia lipolytica*) PO1f 게놈 DNA를 주형으로 하고 서열번호 7과 서열번호 8의 프라이머를 이용하여 PCR을 수행하였고, CYC1 터미네이터를 증폭하기 위해 사카로마이세스 세레비지에 (*Saccharomyces cerevisiae*) 게놈 DNA를 주형으로 하고 서열번호 9와 서열번호 10의 프라이머를 이용하여 PCR을 수행하였다. PCR 조건으로, 변성 95°C, 1분; 어닐링 55°C, 1분; 및 중합반응 68°C, 1분을 32회 반복 수행하였다. 추가적으로 pIMR53 벡터를 주형으로 하고 서열번호 11과 서열번호 12의 프라이머를 이용하여 PCR을 수행하였다. PCR 조건으로, 변성 95°C, 1분; 어닐링 55°C, 1분; 및 중합반응 68°C, 5분을 32회 반복 수행하였다.

- [150] 그 결과 6,774bp의 linearized pIMR53, 531bp의 TEFINtp, 그리고 248bp의 CYC1t를 얻었다. PCR로 증폭된 상기 TEFINtp 및 CYC1t DNA 단편은 Infusion 클로닝 키트(Invitrogen)를 사용하여 pIMR53 벡터에 접합한 후 대장균 DH5 α 에 형질전환하고, 30 mg/L의 앰피실린이 포함된 LB 고체배지에 도말하였다.

- [151] PCR을 통해 목적인 TEFINt 프로모터와 CYC1 터미네이터가 삽입된 플라스미드로 형질 전환된 콜로니를 선별한 후, 통상적으로 알려진 플라스미드 추출법을 이용하여 플라스미드를 획득하였고, 이 플라스미드를 pIMR53-TEFINtp-CYC1t이라 명명하였다.
- [152]
- [153] 실시예 2. 레티노이드 생산용 *Yarrowia lipolytica* 플랫폼 균주 제작
- [154] 실시예 2-1. *X. dendrorhous* 유래 crtYB-crtI 삽입주 제작
- [155] 레티노이드 생산을 위한 플랫폼 균주 제작을 위해 고지방 효모인 *Yarrowia lipolytica* CC08-0125 (기탁번호 KCCM12972P) 균주의 게놈에 *Xanthophyllomyces dendrorhous* 유래 lycopene cyclase/phytoene synthase (crtYB)와 phytoene desaturase(crtI) 유전자를 삽입하였다. crtYB는 NCBI(National Center for Biotechnology Information Search database)에 등록되어 있는 염기서열(GenBank: AY177204.1)에 근거하여 서열번호 13의 폴리뉴클레오티드를 확보하였고, crtI는 NCBI에 등록되어 있는 염기서열(GenBank: AY177424.1)에 근거하여 서열번호 14의 폴리뉴클레오티드를 확보하였다. crtYB와 crtI의 폴리뉴클레오티드 서열은 마크로젠사를 통해 TEFINtp-crtYB-CYC1t(서열번호 15), TEFINtp-crtI-CYC1t(서열번호 16)의 형태로 유전자를 합성하였으며, 선별 마커로는 *Y. lipolytica*의 URA3 유전자(서열번호 17)를 이용하여 MHY1(YALI0B21582g) 유전자 위치에 삽입되는 카세트를 디자인 하였다. 합성된 crtYB, crtI 유전자 및 KCCM12972P 게놈 DNA를 주형으로 하고 서열번호 18 및 서열번호 19, 서열번호 20 및 서열번호 21, 서열번호 22 및 서열번호 23, 서열번호 24 및 서열번호 25, 서열번호 26 및 서열번호 27, 및 서열번호 28 및 서열번호 29의 프라이머를 이용하여 각각의 PCR을 수행하였다. PCR 조건은 변성 95°C, 1분; 어닐링 55°C, 1분; 및 중합반응 72°C, 3분 30초를 35회 반복 수행하였다. 그 결과로 얻어진 5개의 DNA 단편은 overlap extension PCR을 통해 하나의 카세트로 제작되었다.
- [156] 이렇게 제작된 카세트를 열충격법 (D.-C. Chen et al., *Appl Microbiol Biotechnol*, 1997) 으로 KCCM12972P 균주에 도입한 후, uracil이 포함되지 않은 고체배지(YLMM1)에서 형성된 콜로니를 획득하였다. 서열번호 30 및 서열번호 31의 프라이머를 이용하여 게놈 내에 카세트 삽입이 확인된 콜로니들을 5-FOA 고체배지에 spotting하여 30°C에서 3일간 배양하였고, 5-FOA 고체 배지에서 자란 콜로니를 획득함으로써 URA3 마커를 회수하였다.
- [157]
- [158] 실시예 2-2. HMGR 강화주 제작
- [159] 앞서 실시예 2-1에서 제작된 crtYB 및 crtI 유전자가 삽입된 균주의 하이드록시메틸글루타릴 리덕테이스(3-hydroxy-3-methylglutaryl-CoA reductase, HMGR) 유전자의 native 프로모터(서열번호 32) 부위를 TEFINt 프로모터로 교체하기 위한 카세트를 디자인하였고, KCCM12972P 게놈 DNA를 주형으로

하고 서열번호 33 및 서열번호 34, 서열번호 35 및 서열번호 36, 서열번호 37 및 서열번호 38, 서열번호 39 및 서열번호 40, 및 서열번호 41 및 서열번호 42의 프라이머를 이용하여 각각의 PCR을 수행하였다. PCR 조건은 변성 95°C, 1분; 어닐링 55°C, 1분; 및 중합반응 72°C, 1분 30초를 35회 반복 수행하였다. 그 결과로 얻어진 5개의 DNA 단편은 overlap extension PCR을 통해 하나의 카세트로 제작되었다.

- [160] 이렇게 제작된 카세트를 열충격법으로 실시예 2-1에서 제작한 균주에 도입한 후, uracil이 포함되지 않은 고체배지(YLMM1)에서 형성된 콜로니를 획득하였다. 서열번호 43 및 서열번호 44의 프라이머를 이용하여 카세트 삽입이 확인된 콜로니들을 5-FOA 고체배지에 spotting하여 30°C에서 3일간 배양하였고, 5-FOA 고체 배지에서 자란 콜로니를 획득함으로써 URA3 마커를 회수하였다. 최종적으로 제작된 레티노이드 생산 플랫폼 균주를 CC08-1023으로 명명하였다.

[161]

[162] < Yarrowia lipolytica minimal media1 (YLMM1)>

- [163] 포도당 20 g/L, 아미노산을 포함하지 않는 효모 질소 염기(Yeast nitrogen base without amino acids) 6.7 g/L, 우라실을 포함하지 않는 효모 합성 드롭 아웃 배지 보충물 (Yeast Synthetic Drop-out Medium Supplements without uracil) 2 g/L, 한천(agar) 15 g/L

[164]

[165] <5-Fluoroorotic Acid (5-FOA)>

- [166] 포도당 20 g/L, 아미노산을 포함하지 않는 효모 질소 염기(Yeast nitrogen base without amino acids) 6.7 g/L, 우라실을 포함하지 않는 효모 합성 드롭 아웃 배지 보충물 (Yeast Synthetic Drop-out Medium Supplements without uracil) 2 g/L, 우라실(Uracil) 50 µg/mL, 5-플루오로 오로틴산(5-Fluoroorotic acid; 5-FOA) 1 g/L, 한천(agar) 15 g/L

[167]

[168] 실시예 3. BCO 활성 강화주 1차 선별

- [169] 활성이 강화된 BCO 선별을 위해 레티노이드 생산 플랫폼 균주인 CC08-1023 균주에 pIMR53(negative control, NC) 벡터, pIMR53-TEFINtp-BCO(야생형, 서열번호 2)-CYC1t 벡터(positive control, PC), 및 실시예 1에서 구축된 BCO random 변이체 발현 벡터 라이브러리를 열 충격법(heat-shock)으로 형질 전환하였다. 이후, uracil이 포함되지 않은 고체배지에서 형성된 콜로니를 획득하였다.

- [170] 96 deep-well plate 31판에 우라실이 포함되지 않은 YLMM2(Yarrowia lipolytica minimal media2) 액체배지를 350 µl씩 분주하고, 획득한 콜로니들을 접종하여 30°C에서 48시간 동안 배양하였다. 이후, 우라실이 포함되지 않은 YLMM1 고체배지에 spotting하여 30°C에서 3일간 배양하였다. 배양 종료 후, 배양된 총 3,000 콜로니들 중에서 positive control

(CC08-1023/pIMR53-TEFIN_{tp}-BCO-CYC1_t) 대비 색깔이 옅은 균주 17종을 1차로 선별하였다. 상기 YLMM2의 제조방법 및 조성은 하기와 같다.

[171]

[172] <YLMM2 (pH 7.0)>

[173] 포도당 40 g/L, 아미노산 및 암모늄 설페이트가 없는 효모 질소 염기(Yeast nitrogen base without amino acids and ammonium sulfate) 1.7 g/L, 우라실(Uracil) 0.5 g/L, 암모늄 설페이트(Ammonium sulfate) 2.5 g/L를 0.1 M 포스페이트 완충액(sodium phosphate buffer)(pH 7.0)에 녹인다.

[174]

[175] 실시예 4. BCO 활성 강화주 2차 선별

[176] 실시예 3에서 1차로 선별된 17종에 대해 플라스크 평가를 진행하였다.

[177] 우라실이 포함되지 않은 YLMM2(*Yarrowia lipolytica* minimal media 2) 50 ml을 함유하는 250 ml 코너-바플 플라스크에, 상기 17종의 균주 및 PC(positive control) 벡터로 형질 전환된 균주를 초기 OD가 1이 되도록 접종하고, 30°C에서 48시간 동안, 250 rpm으로 진탕 배양하였다. 배양 종료 후, 배양액 1 ml을 원심 분리하여 상등액을 제거하였다. 다음으로, DMSO(Dimethyl sulfoxide, sigma社) 0.5 ml을 첨가하고 55°C에서 10분동안 2,000rpm에서 진탕하여 셀을 파쇄하였다. 추가로, 아세톤(sigma社) 0.5 ml을 첨가하고 45°C에서 15분동안 2,000rpm에서 진탕하여 레티놀 및 레티날을 추출하였고, HPLC 설비로 농도를 분석하였다. 분석된 OD, 레티놀 및 레티날 농도 결과는 표 2에 표시하였다.

[178]

[179] [표2]

Number	OD(A600)	Retinol(mg/L)	Retinal(mg/L)	Retinol(%)
NC	65	ND	ND	-
PC	61	8.2	1.4	100
1	64	12.0	1.9	146
2	42	7.0	1.2	85
3	32	4.9	0.8	59
4	66	9.7	1.5	118
5	57	9.3	1.5	113
6	44	6.1	1.0	74
7	47	6.0	1.0	72
8	38	5.9	1.0	72
9	56	7.1	1.2	86
10	47	7.0	1.1	85
11	58	7.7	1.3	94
12	52	8.3	1.3	101
13	46	7.1	1.2	86
14	46	7.1	1.2	86
15	54	6.9	1.1	84
16	58	8.7	1.4	106
17	54	7.7	1.3	94

[180]

[181] 이후, 평가 시료로부터 추출된 레티놀 농도를 2차 선별을 위한 기준으로 이용하였다. 평가 균주 중 5종(number1,4,5,12,16)이 PC(positive control) 대비 레티놀 농도(Retinol(%))가 증가한 것을 확인하였다. 특히나 1번 균주의 경우 PC(positive control) 대비 레티놀 농도가 약 46% 증가함을 확인하였다. 따라서 레티놀 농도 상승폭이 가장 큰 1번 균주에서 형질 전환된 플라스미드를 추출하여 BCO ORF에 위치한 random 변이를 확인하기 위한 서열분석을 진행하였다. Zymoprep yeast plasmid miniprep kit2(Zymo research社)를 사용하여 선별된 균주로부터 형질 전환된 BCO 플라스미드를 추출하고 서열분석의 신뢰도를 높이기 위해 서열번호 45와 서열번호 46을 사용하여 양방향으로 서열분석을 진행하였다. 그 결과, BCO ORF내에서 돌연변이 수는 3개로

확인되었으며, 확인된 돌연변이는 표 3에 나타내었다.

[182]

[183] [표3]

변이주 Number	Amino acid seq (Nucleotide seq)
1	A43V(GCC→GTC)
	L58Q(CTG→CAG)
	I74V(ATC→GTC)

[184]

[185] 상기 결과로부터 본 출원의 BCO 변이체를 도입할 경우 레티노이드 생산량을 증가시킬 수 있음을 확인하였다.

[186]

[187] 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.

[188]

[189] 서열목록

[190] 본 출원의 서열은 하기 표 4 내지 6과 같다.

[191] [표4]

서열번호	명칭	서열	타입
서열번호 1	Blh_W T	MGLMLIDWCA LALVVFGLP HGALDAAISF SMISSAKRIA RLAGILLIYL LLATAFFLIW 60YQLPAFSLLI FLLISIIHFG MADFNASPSK LKWPHIIAHG GVVTVWLPLI QKNEVTKLFS 120 ILTNGPTPIL WDILLIFFLC WSIGVCLHTY ETLRSKHYNI AFELIGLIFL AWYAPPLVTF 180 ATYFCFIHSR RHFSFVWKQL QHMSSKKMMI GSAILSCTS WLIGGGIYFF LNSKMIASEA 240 ALQTVFIGLA ALTVPHMILI DFIFRPHSSR IKIKN	단백질
서열번호 2	blh_W T	atgggcttga tgtaattga ttggtgtgct ttagcattgg ttgtgtttat tggtttgcca 60catggtgcct tagatgctgc tatttctttt tcaatgattt cttcagcaaa gagaattgct 120 agattagcag gaatactatt aatttacctg ttgtagcaa ccgcattttt tttaatttgg 180 tatcaattac cagcatttcc tcttcttatt ttcttttga taagcataat ccatttttga 240 atggctgatt tcaatgcac cccaagtaaa ctttaagtggc ctcatattat tgcacatggc 300 ggcgttgta ctgtttggtt gccgcttacc caaaaaaatg aagttacgaa gctattttca 360 atattaacaa atggccaac tcccatttta tgggacatac tattgatatt tttttatgt 420 tggagcatag gagtatgtct tcatacctat gaaactttac gttctaaca ttataatata 480 gcctttgaac ttattggatt aatttttcta gcctggtag caccceact cgttactttt 540 gccacatact tctgctttat ccacagcaga cgtcacttta gtttgtttg gaaacagta 600 cagcatatga gttcaaaaaa aatgatgata ggtagtgcca ttattttatc ttgtacgagc 660 tggttgatag gcggaggaa atatttttcc ctcaattcga	DNA

		aaatgattgc cagtgaagct 720 gctttacaaa ctgtctttat tggcttgca gctttaacag ttcctcacat gatacttate 780 gactttatat ttagaccaca ctctccaga attaaaatca aaaattga 828	
서열번호 3	Blh_V T	MGLMLIDWCA LALVVFIGLP HGALDAAISF SMISSAKRIA RLVGILLIYL LLATAFFQIW 60YQLPAFSLLI FLLVSIHFG MADFNASPSK LKWPPIIAHG GVVTVWLPLI QKNEVTKLFS 120 ILTNGPTPIL WDILLIFFLC WSIGVCLHTY ETLRSKHYN I AFELIGLIFL AWYAPPLVTF 180 ATYFCFIHSR RHFSFVWKQL QHMSSKKMMI GSAILSCTS WLIGGGIYFF LNSKMIASEA 240 ALQTVFIGLA ALTVPHMILI DFIFRPHSSR IKIKN	단백질
서열번호 4	blh_V T	atgggcctga tgctgatcga ctggtgtgcc ctggccctgg tgggttcat cgccctgcc 60cacggcgccc tggacgcgc catctcttc tctatgatct ctctgccaa gcaatcgcc 120 cgactggctg gcactctgct gatctacctg ctgctggcca ccgcttctt ccagatctgg 180 taccagctgc ccgcttctc tctgctgac ttctgctgg tctctatcat ccaactcggc 240 atggccgact tcaacgctc tcccttaag ctgaagtggc cccacatcat cggccacggc 300 ggctgtgtga ccgtgtggct gccctgac cagaagaacg aggtgaccaa gctgttctct 360 atctgacca acggcccccac cccatcctg tgggacatcc tgctgatctt ctctctgtgt 420 tggctatcg gcgtgtgtc gcacacctac gagacctgc gatctaagca ctacaacac 480 gccttcgagc tgatcgccct gatcttctg gcctggtacg cccccccct ggtgaccttc 540 gccacctact tctgttctat ccaactctga cgacactct ctctctgtg gaagcagctg 600	DNA

		cagcacatgt ctctaagaa gatgatgac ggctctgcca tcatectgtc ttgtacctct 660 tggctgacgc gcggcgcat ctacttctc ctgaactcta agatgatgc ctctgaggcc 720 gccctgcaga ccgtgttcat cggcctggcc gccctgaccg tgccccacat gatcctgac 780 gacttcactc tccgaccca ctctctcga atcaagatca agaactag 828	
서열번호 5	BCO-F	ttttgcagta ctaaccgcag ggcctgatgc tgatcgactg	DNA
서열번호 6	BCO-R	tgacataact aattacatga ctagtcttg atcttgattc	DNA
서열번호 7	TEFIN tp-F	cccttcgtct tcaaggaatt agagaccggg ttggcggcgc	DNA
서열번호 8	TEFIN tp-R	cgtagacataa ctaattacat gactgcgggt agtactgcaa aaag	DNA
서열번호 9	CYC1t -F	cttttgcag tactaaccgc agtcatgtaa ttagttatgt cacg	DNA
서열번호 10	CYC1t -R	tcgtcgcggc cgcggcgcgc cggcaaatta aagccttcga gcgt	DNA
서열번호 11	pIMR5 3-F	gcgccgcaa cccggtctct aattccttga agacgaaagg	DNA
서열번호 12	pIMR5 3-R	acgctcgaag gctttaattt gccggcgcgc cgcggccgcg acga	DNA
서열번호 13	X. dendro rhous crtYB	atgacggctc tcgcatatta ccagatccat ctgatctata ctctcccaat tcttggtctt 60ctcggcctgc tcaactcccc gattttgaca aaatttgaca tctacaaaat atcgatcctc 120 gtatttattg cgtttagtgc aaccacacca tgggactcat ggatcatcag aaatggcgca 180 tggacatac catcagcgga gagtggccaa ggcgtgtttg gaacgtttct agatgttcca 240 tatgaagagt acgctttctt tgcattcaa accgtaatca ccggcttggt ctacgtcttg 300 gcaactaggc acctctccc atctctcgcg ctcccaaga ctagatcgtc cgcctttct 360 ctcgcgctca aggcgctcat ccctctgccc attatctacc	DNA

	tatttaccgc tcacccagc 420 ccatcgcccg acccgctcgt gacagatcac tacttctaca tgcgggcact ctccttactc 480 atcacccac ctaccatgct cttggcagca ttatcaggcg aatatgcttt cgattggaaa 540 agtggccgag caaagtcaac tattgcagca atcatgatcc cgacgggtga tctgatttgg 600 gtagattatg ttgctgtcgg tcaagactct tggtcgatca acgatgagaa gattgtaggg 660 tggaggcttg gaggtgtact acccattgag gaagctatgt tcttcttact gacgaatcta 720 atgattgttc tgggtctgtc tgcctgcgat catactcagg ccctatacct gctacacggt 780 cgaactatft atggcaacaa aaagatgcca tcttcatttc ccctcattac accgcctgtg 840 ctctccctgt tttttagcag ccgaccatac tcttctcage caaaacgtga cttggaactg 900 gcagtcaagt tgttgaggga aaagagccgg agcttttttg ttgcctcggc tggatttctc 960 agcgaagtta gggagaggct ggttgacta tacgcattct gccgggtgac tgatgatctt 1020	
	atcgactctc ctgaagtatc ttccaaccg catgccacaa ttgacatggt ctccgatttt 1080 cttaccctac tatttgggcc cccgctacac ccttcgcaac ctgacaagat cctttcttcg 1140 cctttacttc ctcttcgca ccttcccgga cccacgggaa tgtatccct cccgcctcct 1200 ccttcgctct cgcctgccga gctcgttcaa ttccttaccg aaagggttcc cgttcaatac 1260 catttcgcct tcagggtgct cgctaagttg caagggtga tccctcgata cccactcgac 1320 gaactcctta gaggatacac cactgatctt atcttccct tatcgacaga ggcagtccag 1380 gctcggaaga cgcctatcga gaccacagct gacttgctgg actatggtct atgtgtagca 1440 ggctcagtcg ccgagctatt ggtctatgtc tcttgggcaa gtgcaccaag tcaggtccct 1500	

		gccaccatag aagaaagaga agctgtgtta gtggcaagcc gagagatggg aactgccctt 1560 cagttggtga acattgctag ggacattaaa ggggacgcaa cagaaggagg attttaccta 1620 ccactctcat tctttggtct tcgggatgaa tcaaagcttg cgatcccgac tgattggacg 1680 gaacctcggc ctcaagattt cgacaaactc ctcaagtctat ctcttcgctc cacattacca 1740 tcttcaaacg cctcagaaag cttccggttc gaatggaaga cgtactcgct tcattagtc 1800 gcctacgcag aggatcttgc caaacattct tataagggaa ttgaccgact tctaccgag 1860 gttcaagcgg gaatgcgagc ggcttgcgcg agctacctac tgatcgcccg agagatcaaa 1920 gtcgtttgga aaggagacgt cggagagaga aggacagttg ccggatggag gagagtacgg 1980 aaagtcttga gtgtggtcat gagcggatgg gaagggcag 2019	
서열번호 14	X. dendro rhous crtI	atgggaaaag aacaagatca ggataaaccc acagctatca tcgtgggatg tggatcggt 60ggaatcgcca ctgccgctcg tcttgctaaa gaaggtttcc aggtcacggt gttcgagaag 120 aacgactact ccggaggctg atgctcttta atcgagcgag atggttatcg attcgatcag 180 gggcccagtt tgctgctctt gccagatctc ttcaagcaga cattcgaaga ttgggagag 240 aagatggaag attgggtcga tctcatcaag tgtgaaccca actatgtttg ccactccac 300 gatgaagaga ctttactctt ttcaaccgac atggcgttgc tcaagcggga agtcgagcgt 360 tttgaaggca aagatggatt tgatcgggtc ttgtcgttta tccaagaagc ccacagacat 420 tacgagcttg ctgtcgttca cgtcctgcag aagaacttcc ctggcttcgc agcattctta 480 cggctacagt tcattggcca aatcctggct cttaccacct tcgagtctat ctggacaaga 540 gtttgtcgat atttcaagac cgacagatta cgaagagtct tctcgtttgc agtgatgtac 600 atgggtcaaa gccatcacag tgcgcccgga acatattcct	DNA

	tgcaccaata caccgaattg 660 accgaggcca tctggtatcc gagaggaggc tttggcagg ttctaatac tcttctcag 720 atcgtaagc gcaacaatcc ctcagccaag ttcaattca acgtccagt ttcccaggtt 780 cttctctc ctgccaagga ccgagcgact ggtgttcgac ttgaatccgg cgaggaacat 840 cagccgatg ttgtgattgt caatgctgac ctcgtttacg cctccgagca ctgattcct 900 gacgatgcca gaaacaagat tggccaactg ggtgaagtca agagaagttg gtgggctgac 960 ttagttggtg gaaagaagct caagggaagt tgcagtagtt tgagcttcta ctggagcatg 1020	
	gaccgaatcg tggacggtct gggcggacac aatatctct tggccgagga ctcaaggga 1080 tcattcgaca caatcttga ggagtgggt ctcccagccg atccttctt ttacgtgaac 1140 gttccctcgc gaatcgatcc ttctccgct cccgaaggca aagatgctat cgtcattctt 1200 gtgccgtgtg gccatatga cgttcgaac cctcaagatt acaacaagct tgtgctcgg 1260 gcaagggaagt ttgtgatcca cacgctttcc gccaagcttg gacttcccga cttgaaaaa 1320 atgattgtgg cagagaaggt tcacgatgct ccctcttggg agaaagaatt caacctcaag 1380 gacggaagca tctgggact ggctcacaac ttatgcaag ttcttggtt caggccgagc 1440 accagacatc ccaagtatga caagttgtc tttgtcgggg cttegactca tcccgaact 1500 ggggttcca tcgtcttggc tggagccaag ttaactgcca accaagtct cgaatcctt 1560 gaccgatccc cagctccaga tccaatatg tcactctcg taccatatgg aaaacctctc 1620 aaatcaaatg gaacgggtat cgattctcag gtccagctga agttcatgga ttggagaga 1680 tgggtatacc tttgtgtt gttgattggg gccgtgatcg ctgatccgt tgggttctt 1740	

		gctttctgaa gcaagacaac gatcctttct tagagtttta ttttagtctc ttcctgtgtt 1800 ctctctatat acatactctg ctctgtctgt ctcttctcga gggttcctct ttactttgtg 1860 tcagagtcac acccggtctc tcaacgtccg ttgagggtct agacaattgt tagtctcgaa 1920 atctccatca cctcaagtct gatgttcac atctttttta ttcggt 1966	
서열번호 15	TEFIN tp-crtY B-CY C1t	agagaccggg ttggcggcgc atttgtgtcc caaaaaacag cccccaattgc cccaattgac 60cccccaattga ccagtagcg ggcccaaccc cggcgagagc ccccttctcc ccacatatca 120 aacctcccc ggttcccaca ctgcccgtta agggcgtagg gtactgcagt ctggaatcta 180 cgttgttca gactttgtac tagtttctt gtctggccat ccgggtaacc catgccggac 240 gcaaaataga ctactgaaaa ttttttgcct ttgtggttgg gactttagcc aagggtataa 300 aagaccaccg tccccgaatt acctttctc ttctttctc tctctcctg tcaactcaca 360 cccgaatcg ttaagcattt ccttctgagt ataagaatca ttcaaaatgg tgagtttcag 420 aggcagcagc aattgccacg ggctttgagc acacggccgg gtgtgggtccc attcccatcg 480 acacaagacg ccacgtcac cgaccagcac ttttgcagt actaaccgca gacggctctc 540 gcatattacc agatccatct gatctatact ctcccaatc ttggtctct cggctctgctc 600 acttccccga ttttgacaaa atttgacatc tacaaaatat cgatcctcgt atttattgcg 660 tttagtgcaa ccacaccatg ggactcatgg atcatcagaa atggcgcatg gacatatcca 720 tcagcggaga gtggccaagg cgtgtttgga acgtttctag atgttcata tgaagagtac 780 gctttcttgg tcatlcaaac cgtaatcacc ggcttggctc acgtcttggc aactaggcac 840 cttctcccat ctctcgcgt tcccaagact agatcgctcg ccctttctct cgcgtcaag 900	DNA

		gcgtcatcc ctctgccc atctaccta ttaccgctc accccagccc atcgcccgac 960 ccgctcgtga cagatcacta cttctacatg cgggcactct ccttactcat caccacac 1020	
		accatgctct tggcagcatt atcaggcgaa tatgcttgc attggaaaag tggccgagca 1080 aagtcaacta ttgcagcaat catgatcccg acgggtgatac tgattgggt agattatgtt 1140 gctgtcggtc aagactctg gtcgatcaac gatgagaaga ttgtagggtg gaggcttga 1200 ggtgtactac ccattgagga agctatgttc ttctactga cgaatctaat gattgttctg 1260 ggtctgtctg cctgcgatca tactcaggcc ctatacctgc tacacggtcg aactattat 1320 ggcaacaaaa agatgccatc ttcatfccc ctcattacac cgctgtgct ctccctgtt 1380 tttagcagcc gaccatactc ttctagcca aaacgtgact tggaactggc agtcaagttg 1440 ttggaggaaa agagccggag ctttttgtt gcctcggctg gatttctag cgaagttagg 1500 gagaggctgg ttggactata cgcattctgc cgggtgactg atgatctat cgactctct 1560 gaagtatctt ccaaccgca tgccacaatt gacatggtct ccgattttct taccctacta 1620 tttgggcccc cgtacaccc ttgcacac 1680 tttcttccc ttacttct 1740 ccttcgacc ctcccgacc cacgggaatg tatccctcc cgctctctc ttgctctcg 1800 cctgcggagc tcgttcaatt cctaccgaa aggggtcccg ttcaatacca ttgccttc 1860 aggttgctcg ctaagttgca agggctgac cctcgatacc cactcgacga actccttaga 1920 ggatacacca ctgatcttat cttccttta tcgacagagg cagtcagggc tcggaagacg 1980 cctatcgaga ccacagctga cttgctggac tatggtctat gtgtagcagg ctacgtgcc 2040 gagctattgg tctatgtctc ttgggcaagt gcaccaagtc	

		aggtccctgc caccatagaa 2040 gaaagagaag ctgtgttagt ggcaagccga gagatgggaa ctgcccttca gttggtgaac 2100 attgctaggg acattaaagg ggacgcaaca gaaggagat ttacctacc actctcattc 2160 ttggtcttc gggatgaatc aaagcttgcg atccccgactg attggacgga acctcgccct 2220 caagatttcg acaaactcct cagtctatct ccttcgtcca cattaccatc ttcaaacgcc 2280 tcagaaagct tccggttcga atggaagacg tactcgcttc cattagtcgc ctacgcagag 2340 gatcttgcca aacattctta taagggaatt gaccgacttc ctaccgaggt tcaagcggga 2400 atgcgagcgg ctgctcgag ctacctactg atcgcccgag agatcaaagt cgtttgaaa 2460 ggagacgtcg gagagagaag gacagtgtcc ggatggagga gagtacggaa agtcttgagt 2520 gtggtcatga gcggatggga agggcagtaa ctcgagcat gtaattagtt atgtcacgct 2580 tacattcacg cctccccccc acatccgctc taaccgaaaa ggaaggagtt agacaacctg 2640 aagtctaggt ccctatttat tttttatag ttatgttagt attaagaacg ttatttatat 2700 ttcaaatttt tcttttttt ctgtacagac gcgtgtacgc atgtaacatt atactgaaaa 2760 ccttgcttga gaaggtttg ggacgctcga aggctttaat ttgcggcgcg cggagccct 2820 ttcgtcttca 2830	
서열번호 16	TEFIN tp-crtI- CYC1t	agagaccggg ttggcggcgc atttgtgtcc caaaaaacag ccccaattgc cccaattgac 60cccaaattga cccagtagcg ggcccaaccc cggcgagagc cccctctcc ccacatatca 120 aacctcccc ggttccaca ctgcccgtta agggcgtagg gtactgcagt ctggaatcta 180 cgcttgttca gactttgtac tagtttcttt gtctggccat ccgggtaacc catgccggac 240 gcaaaataga ctactgaaaa ttttttgct ttgtggttgg gactttagcc aagggtataa 300	DNA

	aagaccaccg tccccgaatt accttccctc ttctttctc tctctccttg tcaactcaca 360 cccgaatcg ttaagcattt ccttctgagt ataagaatca ttcaaatgg tgagtttcag 420 aggcagcagc aattgccacg ggctttgagc acacggccgg gtgtggctcc attcccatcg 480 acacaagacg ccacgtcatc cgaccagcac tttttgcagt actaaccgca gggaaaagaa 540 caagatcagg ataaaccac agctatcatc gtgggatgtg gtatcggtgg aatcgccact 600 gccgctcgtc ttgctaaaga aggtttccag gtcacgggtg tcgagaagaa cgactactcc 660 ggaggtcgat gctctttaat cgagcgagat ggttatcgat tcgatcaggg gccagtttg 720 ctgctcttgc cagatctctt caagcagaca ttcgaagatt tgaggagagaa gatggaagat 780 tggttcgac tcataagtg tgaaccaac tatgtttgcc acttcacga tgaagagact 840 ttcactctt caaccgacat ggctgtgctc aagcgggaag tcgagcgtt tgaaggcaaa 900 gatggatttg atcggttctt gtcgtttatc caagaagccc acagacatta cgagcttgct 960 gtcgttcacg tctgcagaa gaacttcctt ggcttcgcag cattcttacg gctacagttc 1020	
	attggccaaa tcttggtctt tcaccccttc gagtctatct ggacaagagt ttgtcgatat 1080 ttcaagaccg acagattacg aagagtcttc tcgtttgcag tgatgtacat gggtaaagc 1140 ccatacagtg cggccggaac atattccttg ctccaatata ccgaattgac cgagggcac 1200 tggtatccga gaggaggctt ttggcagggt cctaatactc ttcttcagat cgtcaagcgc 1260 aacaatccct cagccaagtt caatttcaac gctccagttt cccaggttct tctctctct 1320 gccaaggacc gagcgactgg tgttcgactt gaatccggcg aggaacatca cgccgatgtt 1380 gtgattgtca atgctgacct cgtttacgcc tccgagcact	

	tgattcctga cgatgccaga 1440 aacaagattg gccaaactggg tgaagtcaag agaagtgtgt gggctgactt agttggtgga 1500 aagaagctca aggggaagttg cagtagtttg agcttctact ggagcatgga ccgaatcgtg 1560 gacggtctgg gcggacacaa tatcttcttg gccgaggact tcaagggatc attcgacaca 1620 atcttcgagg agttgggtct cccagccgat ccttcctttt acgtgaacgt tccctcgcga 1680 atcgatcctt ctgccgctcc cgaaggcaaa gatgctatcg tcattcttgt gccgtgtggc 1740 catatcgacg cttcgaacct tcaagattac aacaagcttg ttgctcgggc aaggaagttt 1800 gtgatccaca cgtttccgc caagcttggga cttcccgact ttgaaaaaat gattgtggca 1860 gagaagggtc acgatgctcc ctctggggag aaagaattca acctcaagga cggaagcatc 1920 ttgggactgg ctcaaacct tatgcaagtt cttggtttca ggccgagcac cagacatccc 1980 aagtatgaca agttgttctt tgcggggct tcgactcatc ccggaactgg ggttcccatc 2040 gtcttggtg gagccaagtt aactgccaac caagttctcg aatcctttga ccgatcccca 2100 gctccagatc ccaatatgtc actctccgta ccatatggaa aacctctcaa atcaaatgga 2160 acgggtatcg attctcaggt ccagctgaag ttcattggatt tggagagatg ggtatacctt 2220 ttggtattgt tgattggggc cgtgatcgct cgatccgttg gtgttcttg tttctgactc 2280 gagtcatgta attagttagt tcacgcttac attcacgcc tccccccaca tccgtctaa 2340 ccgaaaagga aggagttaga caacctgaag tctaggtccc tatttatttt ttatagtta 2400 tgtagtatt aagaacgtta ttatatttc aaattttct tttttctg tacagacgcg 2460 tgtacgcatg taacattata ctgaaaacct tgcttgagaa ggttttggga cgctgaagg 2520 cttaatttgcggccgcga cgagtatctg tctga 2555	
--	---	--

서열번호 17	Y. lipolyti ca URA3	tgcctcctgt ctgactcgtc attgccgcct ttggagtacg actccaacta tgagtgtgct 60tggatcactt tgacgataca ttcttcgttg gaggctgtgg gtctgacagc tgcgttttcg 120 gcgcgggttg ccgacaacaa tatcagctgc aacgtcattg ctggctttca tcatgatcac 180 atTTTTgtcg gcaaaggcga cggccagaga gccattgacg ttctttctaa ttggaccga 240 tagccgtata gtccagtcta tctataagtt caactaactc gtaactatta ccataacata 300 tacttcactg cccagataa gggtccgata aaaagttctg cagactaaat ttatttcagt 360 ctcctcttca ccacaaaat gccctctac gaagctcgag ctaacgtcca caagtcggcc 420 tttggcgtc gagtgctcaa gctcgtggca gccaaagaaa ccaacctgtg tgcttctctg 480 gatgttacca ccaccaagga gctcattgag ctggccgata aggtcggacc ttatgtgtgc 540 atgatcaaga cccatatcga catcattgac gacttcacct acgccggcac tgtgtccccc 600 ctcaaggaac ttgctcttaa gcacggtttc ttctgttcg aggacagaaa gtgcgcagat 660 attggcaaca ctgtcaagca ccagtacaag aacggtgtct accgaatcgc cgagtgtgcc 720 gatatcacca acgcccacgg tgtaccggga accggaatca ttgctggcct gcgagctggt 780 gccgaggaaa ctgtctctga acagaagaag gaggacgtct ctgactacga gaactcccag 840 tacaaggagt tctgggtccc ctctcccaac gagaagctgg ccagaggtct gctcatgctg 900 gccgagctgt cttgcaaggg ctctctggcc actggcgagt actccaagca gaccattgag 960 cttggccgat ccgacccga gtttgtggtt ggcttcattg cccagaaccg acctaagggc 1020 gactctgagg actggcttat tctgaccccc ggggtgggtc ttgacgacaa gggagacgct 1080 ctcggacagc agtaccgaac tgttgaggat gtcattgcta ccggaacgga tatcataatt 1140 gtcggccgag gtctgtacgg ccagaaccga gatcctattg	DNA
---------	------------------------------	---	-----

		aggaggccaa gcgataccag 1200 aaggctggct gggaggctta ccagaagatt aactgttaga ggtagacta tggatatgc 1260 atttaactgt gtatatagag agcgtgcaag tatggagcgc ttgtcagct tgtatgatgg 1320 tcagacgacc tgtctgatcg agtatgtatg atactgcaca acctgtgtat ccgcatgac 1380 tgtccaatgg ggcatgtgt tgtgtttctc gatacggaga tgctgggtac aagtagctaa 1440 tacgattgaa ctacttatac ttatatgagg ctgaagaaa gctgacttgt gtatgactta 1500 ttctcaactacatccccagtcacaatacca cca 1533	
서열번호 18	Left arm1-F	gtgcgcttct ctcgtctcgg taaccctgtc	DNA
서열번호 19	Left arm1- R	atgcgccgcc aaccggctct ctgggggtgtg gtggatgggg tgtg	DNA
서열번호 20	TEFIN tp-crtY B-CY C1t-F	cacaccccat ccaccacacc ccagagaccg ggttgccggc gcat	DNA

[192] [표5]

서열번호 21	TEFIN tp-crtY B-CY C1t-R	cgccgccaac ccggtctctt gaagacgaaa gggcctccg	DNA
서열번호 22	TEFIN tp-crtI- CYC1t -F	cggaggccct ttcgtcttca agagaccggg ttggcggcg	DNA
서열번호 23	TEFIN tp-crtI- CYC1t -R	gacgagtcag acaggaggca tcagacagat actcgtcgcg	DNA
서열번호 24	Repeat -URA3 -F	cgcgacgagt atctgtctga tgcctctctg ctgactcgc	DNA
서열번호 25	Repeat -URA3 -R	atgacgagtc agacaggagg catggtgga ttgtgactgg ggat	DNA
서열번호 26	Repeat -F	atccccagtc acaataccac catgcctct gtctgactcg teat	DNA
서열번호 27	Repeat -R	cggegtcctt ctgtagtcc gcttttggtg gtgaagagga gact	DNA
서열번호 28	Right arm1-F	agtctcctt tcaccacaa aagcggacta cgagaaggac gccg	DNA
서열번호 29	Right arm1- R	ccactcgtca ccaacagtgc cgtgtgttgc	DNA
서열번호 30	MHY1 -CF	tcgtacgtct ataccaacag atgg	DNA
서열번호 31	MHY1 -CR	cgcatacaca cacactgccg gggg	DNA
서열번호 32	HMG R	tccacacgtc gttctttttt ccttagcctt tttgcagtg cgcgtgtccc aaaccccagc 60tctacacacc agcacaaca aagttaagct	DNA

	native promoter region	cagggtgtc gttgaggtcg cttactgtag 120 tcagtgtcg tatggttcgt tcaatttcg ccaaaaatcg ttttgcctttgtatcttggg 180 aataacatca actgtggttc ttcaacaggc ctaaggaacg aaacaagccg gaccaagatc 240 agggtcaagg tgagtactga gaaggaatag aaggcctaaa ggcgcaaacc gacaggtggc 300 aacagctcca caccgaccac gaaggccacg aaatcaagggtcctaaagt tagtctttgt 360 ggcctcgacg gtcagcgaaa acgcgagacc acaacgcgatcagaaccagg acctaaacaa 420 cacaggacgg ggtcacaata ggcttgaaca gcaagtacaagctgtgatct ctctatattt 480 gatttcaaa ccacccctga ctacttcagc gcctctgtgacacagcccc ctatcatccg 540 actaacacag	
서열번호 33	Left arm2-F	gacaatgcct cgaggagggtt taaaagtaac t	DNA
서열번호 34	Left arm2-R	gegccgceaa cccggtctct ctgtgttagt cggatgatag g	DNA
서열번호 35	TEFIN tp-F	cctatcatcc gactaacaca gagagaccgg gttggcggcg c	DNA
서열번호 36	TEFIN tp-R	gacgagtcag acaggaggca ctgcggtag tactgcaaaa ag	DNA
서열번호 37	URA3-F	cttttgcag tactaaccgc agtgccctct gtctgactcg tc	DNA
서열번호 38	URA3-R	atgcgccgcc aacccggtct ctggttgga ttgtgactgg ggat	DNA
서열번호 39	Repeat -F	atccccagtc acaataccac caagagaccg ggttggcggc gcat	DNA
서열번호 40	Repeat -R	ctttccaata gctgcttgta gctgcggta gtactgcaaa a	DNA

[194] [표6]

서열번호 41	Right arm2-F	ttttgcagta ctaaccgcag ctacaagcag ctattggaaa g	DNA
서열번호 42	Right arm2-R	gcttaatgtg attgatctca aacttgatag	DNA
서열번호 43	HMGR -CF	gctgtctctg cgagagcacg tga	DNA
서열번호 44	HMGR -CR	ggttgcaca acttctcggg tggc	DNA
서열번호 45	BCO-s q1	cttctgagta taagaatcat tcaa	DNA
서열번호 46	BCO-s q2	tttgaaatat aaataacgtt ctta	DNA

[195]

[196]

특허출원을 위한 부다페스트 국제 조약 하의 미생물 수탁 증명

국 제 양 식

수신: 씨제이제일제당 주식회사,
대한민국 서울시 중구 동호로 330 CJ 제일제당센터 (우) 100-400

국제기탁기관에 의해 규칙 7.1에 따라 발행된 원기탁에 대한 수탁증

I. 미생물의 표시	
기탁자가 첨부한 미생물 식별에 대한 표시: <i>Yarrowia lipolytica</i> CC08-0125	국제 기탁기관이 부여한 수탁번호: KCCM12972P
II. 과학적 성질의 설명 및/또는 분류학상의 위치	
항목 I에 표시된 미생물에 대하여 다음 사항이 포함되어 있다: ☐ 과학적 성질의 설명 ☐ 제안된 분류학상의 위치 (적용시 x 표시)	
III. 기탁 및 수탁	
본 국제기탁기관은 2021년 4월 9일에 기탁된 항목 I에 표시된 미생물을 수탁하였다.	
IV. 전환 요청의 수령	
상기 항목 I에 표시된 미생물은 에 본 국제기탁기관에 수탁되었고, 원기탁을 부다페스트 조약 하의 기탁으로 전환하기 위한 요청을 에 받았다.	
V. 국제기탁기관	
명칭: 한국미생물보존센터 주소: 대한민국 서울시 서대문구 홍제내 2가길 45 유림빌딩 (우) 03641	국제기탁기관을 대표하는 권한을 가진 자 또는 권한을 부여받은 공무원 서명: 서명일: 2021. 4. 9.

상기 번역은 원문의 내용과 상위 있음을 증명함.

2022년 7월 12일

특허법인 한일 인



청구범위

- [청구항 1] 서열번호 1의 아미노산 서열의 N-말단으로부터 43번 위치에 상응하는 아미노산인 알라닌, 58번 위치에 상응하는 아미노산인 류신, 및 74번 위치에 상응하는 아미노산인 이소류신이 다른 아미노산으로 치환된, 베타-카로틴 15,15'-옥시게네이즈 변이체.
- [청구항 2] 제1항에 있어서, 상기 변이체는 서열번호 1의 아미노산 서열의 N-말단으로부터 43번 위치에 상응하는 아미노산이 알라닌 외의 아미노산으로 치환; 58번 위치에 상응하는 아미노산이 류신 외의 아미노산으로 치환; 및 74번 위치에 상응하는 아미노산이 이소류신 외의 아미노산으로 치환을 포함하는, 베타-카로틴 15,15'-옥시게네이즈 변이체.
- [청구항 3] 제1항에 있어서, 상기 43번 위치에 상응하는 아미노산의 다른 아미노산으로의 치환은 비극성 아미노산; 상기 58번 위치에 상응하는 아미노산의 다른 아미노산으로의 치환은 극성 아미노산; 및 74번 위치에 상응하는 아미노산의 다른 아미노산으로의 치환은 비극성 아미노산으로의 치환인 것인, 베타-카로틴 15,15'-옥시게네이즈 변이체.
- [청구항 4] 제1항에 있어서, 상기 다른 아미노산은 발린 또는 글루타민인 것인, 베타-카로틴 15,15'-옥시게네이즈 변이체.
- [청구항 5] 제1항의 변이체를 코딩하는 폴리뉴클레오티드.
- [청구항 6] 제1항의 변이체 및 상기 변이체를 코딩하는 폴리뉴클레오티드 중 어느 하나 이상을 포함하는, 미생물.
- [청구항 7] 제6항에 있어서, 상기 미생물은 레티노이드 생산능을 갖는 것인, 미생물.
- [청구항 8] 제7항에 있어서, 상기 레티노이드는 레티날, 레티놀, 레티닐 에스터, 및 레티노산으로 구성되는 군에서 선택되는 어느 하나 이상인, 미생물.
- [청구항 9] 제6항에 있어서, 상기 미생물은 야로위아 속(*Yarrowia sp.*)인, 미생물.
- [청구항 10] 제6항에 있어서, 상기 미생물은 야로위아 리폴리티카 (*Yarrowia lipolytica*)인, 미생물.
- [청구항 11] 제1항의 변이체 또는 상기 변이체를 코딩하는 폴리뉴클레오티드를 포함하는 미생물을 배지에서 배양하는 단계를 포함하는 레티노이드 생산방법.
- [청구항 12] 제11항에 있어서, 상기 방법은 상기 배지 또는 미생물로부터 레티노이드를 회수하는 단계를 추가로 포함하는 레티노이드 생산방법.
- [청구항 13] 제11항에 있어서, 상기 미생물은 야로위아 속(*Yarrowia sp.*)인, 레티노이드 생산방법.
- [청구항 14] 제11항에 있어서, 상기 레티노이드는 레티날, 레티놀, 레티닐 에스터, 및 레티노산으로 구성되는 군에서 선택되는 어느 하나 이상인, 레티노이드 생산방법.
- [청구항 15] 제1항의 변이체, 제8항의 미생물, 또는 상기 미생물의 배양물을 포함하는

레티노이드 생산용 조성물.

- [청구항 16] 서열번호 1의 아미노산 서열의 N-말단으로부터 43번 위치에 상응하는 아미노산인 알라닌, 58번 위치에 상응하는 아미노산인 류신, 및 74번 위치에 상응하는 아미노산인 이소류신이 다른 아미노산으로 치환된, 베타-카로틴 15,15'-옥시게네이즈 변이체; 및 상기 변이체를 코딩하는 폴리뉴클레오티드 중 어느 하나 이상을 포함하는, 미생물의 레티노이드 생산 용도.