



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2009-0097056
(43) 공개일자 2009년09월15일

(51) Int. Cl.

C12Q 1/68 (2006.01) C12N 15/63 (2006.01)
C12N 15/09 (2006.01)

(21) 출원번호 10-2008-0022217

(22) 출원일자 2008년03월10일

심사청구일자 2008년03월10일

(71) 출원인

충남대학교산학협력단

대전광역시 유성구 공동 220번지 충남대학교

(72) 발명자

김동명

대전광역시 서구 삼천동 청솔 아파트 7동 303호

안진호

서울시 종로구 삼청동 35-137

금정원

서울시 서초구 반포본동 주공아파트 111동 105호

(74) 대리인

권오식, 박창희, 김종관

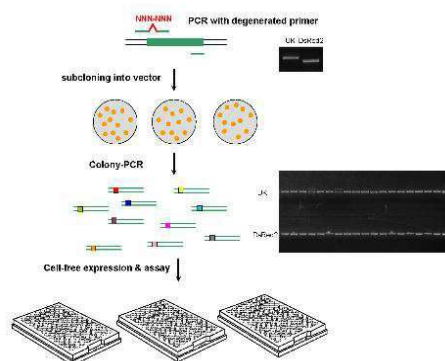
전체 청구항 수 : 총 9 항

(54) 목적하는 단백질 발현 수준을 제공하는 초기 코돈의 발굴방법, 및 이를 이용한 재조합 단백질의 발현 조절 방법 및 생산 방법

(57) 요약

본 발명은 유전자의 초기 코돈(initial codon), 즉 +2 내지 +13 코돈 중 하나 이상, 특히 +2 및 +3 코돈을 무작위로 변화시킨 후 이들을 세포 내로 도입하여 개별 유전자로 분리하고 콜로니 PCR에 의해 증폭한 후, 무세포 단백질 합성 시스템을 이용하여 각각의 유전자를 단백질로 발현시킴으로써, 목적하는 단백질 발현 수준을 제공하는 초기 코돈을 발굴하는 방법, 및 이를 이용한 재조합 단백질의 발현 조절 방법 및 생산 방법에 관한 것으로서, 단백질 생산에 있어 체계적이고 정확한 조절 수단이 될 수 있으며, 활성 단백질을 원하는 수준으로 고속으로 발현할 수 있다.

대표도 - 도1



특허청구의 범위

청구항 1

- 1) 목적 유전자를 +2 내지 +13 코돈 중 하나 이상의 위치에서 무작위로 변화시키고;
 - 2) 단계 1)의 무작위로 변화된 유전자를 각각 세포 내로 도입하여 개별 유전자로 분리한 후, 콜로니 PCR(polymerase chain reaction)에 의해 증폭하고;
 - 3) 단계 2)의 유전자로부터 무세포 단백질 합성(cell-free protein synthesis)을 수행하고;
 - 4) 단계 3)에서의 단백질 발현효율을 조사하여 목적하는 단백질 발현 수준을 제공하는 코돈을 선별하는:
- 단계를 포함하는, 목적하는 단백질 발현 수준을 제공하는 초기 코돈의 발굴방법.

청구항 2

제1항에 있어서, 단계 1)에서 유전자를 +2, +3, 또는 +2 및 +3 코돈 위치에서 변화시키는 방법.

청구항 3

제1항에 있어서, 단계 1)에서 유전자를 +2 내지 +13 코돈 중 하나 이상의 위치에서 무작위로 변화된 축퇴 정방향 프라이머(degenerate forward primer)를 이용하여 PCR(polymerase chain reaction)에 의해 증폭시키는 방법.

청구항 4

제1항에 있어서, 단계 2)에서 세포가 세균 세포인 방법.

청구항 5

제4항에 있어서, 세균 세포가 대장균(*E. coli*) 세포인 방법.

청구항 6

제1항에 있어서, 초기 코돈이 증가된 단백질 발현 수준을 제공하는 것인 방법.

청구항 7

제6항에 있어서, 초기 코돈이 A 및 T 풍부(A- and T-rich) 서열인 방법.

청구항 8

제1항 내지 제7항 중 어느 한 항에 따른 방법에 의해 발굴된 초기 코돈을 갖는 유전자로부터 단백질을 발현시키는 단계를 포함하는, 재조합 단백질의 발현 조절 방법.

청구항 9

제1항 내지 제7항 중 어느 한 항에 따른 방법에 의해 발굴된 초기 코돈을 갖는 유전자로부터 단백질을 발현하는 단계를 포함하는, 재조합 단백질의 생산 방법.

명세서

발명의 상세한 설명

기술 분야

<1> 본 발명은 재조합 단백질의 발현 조절에 관한 것으로서, 보다 상세하게는 목적하는 단백질 발현 수준을 제공하는 초기 코돈의 발굴방법, 및 이를 이용한 재조합 단백질의 발현 조절 방법 및 생산 방법에 관한 것이다.

배경 기술

<2> 프로테오믹스(proteomics)의 원리는 복잡한 시료 혼합물로부터 단백질을 발굴하거나 특정 단백질의 발현 또는 번역 후 수식(post translational modification, PTM) 과정의 정도를 규명하는데 널리 이용되고 있다. 그러나 이러한 포워드 프로테오믹스(forward proteomics) 방법만으로는 생물학적 구성요소의 조화된 기능을 통합적으로 이해하는 데는 충분하지 못하다. 대신에, 포워드 프로테오믹스를 통해 얻은 결과를 바탕으로 확인된 단백질의 추가적인 기능 규명이 요구되는데, 이를 위하여 재조합 단백질 합성 방법을 통해 이미 확인된 유전자의 발현과 분석이 필요하게 된다. 리버스 프로테오믹스(reverse proteomics)라 불리는 이러한 방법에서는, 재조합 단백질의 합성에 있어서 정확한 발현 조절이 필요하다. 대부분의 경우 목적 단백질을 높은 수율로 발현시키는 것이 일반적으로 요구되지만, 세포의 독성과 비정상 단백질 접합을 피하기 위해 경우에 따라서는 발현되는 단백질의 발현 수준을 조절해야 할 필요가 있다. 단백질 합성 과정은 세포 내에서 여러 단계를 거치는 복잡한 과정이며, mRNA의 번역 효율은 주로 단백질 합성 과정의 초기 단계와 이른 신장 단계에서 주로 조절되는 것으로 알려져 있다. 또한 잘 알려진 샤인-달가르노(Shine-Dalgarno) 서열 이외에도, 통계적 분석을 통해 개시 코돈(initiation codon)과 종결 코돈(termination codon) 주위의 염기 서열이 특정한 분포를 가진다는 사실이 알려져 있다. 특히, 개시 코돈 주위의 초기 염기 서열(downstream box)은 mRNA의 발현 효율에 있어서 매우 중요한 영향을 미치는 것으로 알려져 있다. 예를 들어, *LacZ* 유전자의 발현 효율은 두 번째 코돈의 종류에 따라 약 20배 차이가 나타나는 것으로 보고되어 있다(문헌 [Stenstrom, C. M. et al., *Gene* 2002, 288(1-2), 1-8]). 한편, 유전자의 초기 코돈 중에 NGG(CGG, AGG, GGG, UGG) 코돈이 존재할 때 mRNA의 번역 효율이 심각하게 저해되는 사실도 알려져 있다.

<3> 상기와 같은 사실로부터, 목적 유전자의 초기 코돈을 변화시킴으로써 재조합 단백질의 발현 효율을 조절할 수 있으리라는 것을 예상할 수 있다. 일반적으로 이러한 실험을 하기 위해서는 특정 유전자의 코돈을 가능한 61개 코돈으로 모두 변화시켜 이들에 대한 발현량을 조사하여야 한다. 그러나 이러한 방법은 많은 시간과 장비가 요구되며, 특히 기존의 세포를 이용한 단백질 발현 방법을 사용할 경우 큰 제약이 따른다. 모든 가능한 코돈을 변화시키기 위해서는 3721(61×61)개의 변이체를 만들고 이로부터 단백질을 발현 및 분석하여야 하는데, 기존의 유전자 합성 및 단백질 발현 방법으로는 현실적으로 불가능하다.

발명의 내용

해결 하고자하는 과제

<4> 이에, 본 발명자들은 특정 유전자의 시작 부분에 있는 발현 촉진 및 억제 코돈의 조합을 통하여 목적 유전자의 발현 수준을 조절할 수 있으리라는 사실에 착안하여, 유전자가 원하는 수준의 발현량을 나타내도록 하기 위하여 무작위로 변화된 초기 염기 서열을 제작하고, 세포와 무세포 방법을 동시에 사용하여 무작위로 변화된 개별 유전자를 각각 분리, 발현 및 분석하여 여러 코돈 조합 중에서 특정 발현수준을 보이는 코돈 조합을 선별하는 방법을 안출해내고, 본 발명을 완성하기에 이르렀다.

<5> 따라서 본 발명의 목적은 목적하는 단백질 발현 수준을 제공하는 초기 코돈의 발굴방법을 제공하기 위한 것이다.

<6> 본 발명의 다른 목적은 상기 방법에 의해 발굴된 초기 코돈을 이용하여 재조합 단백질 발현을 조절하는 방법을 제공하기 위한 것이다.

<7> 본 발명의 또 다른 목적은 상기 방법에 의해 발굴된 초기 코돈을 이용하여 재조합 단백질을 생산하는 방법을 제공하기 위한 것이다.

과제 해결수단

<8> 본 발명의 제1면은

<9> 1) 목적 유전자를 +2 내지 +13 코돈 중 하나 이상의 위치에서 무작위로 변화시키고;

<10> 2) 단계 1)의 무작위로 변화된 유전자를 각각 세포 내로 도입하여 개별 유전자로 분리한 후, 콜로니 PCR(polymerase chain reaction)에 의해 증폭하고;

<11> 3) 단계 2)의 유전자로부터 무세포 단백질 합성(cell-free protein synthesis)을 수행하고;

<12> 4) 단계 3)에서의 단백질 발현효율을 조사하여 목적하는 단백질 발현 수준을 제공하는 코돈을 선별하는;

- <13> 단계를 포함하는, 목적하는 단백질 발현 수준을 제공하는 초기 코돈의 발굴방법에 관한 것이다.
- <14> 본 발명의 제2면은 상기 방법에 의해 발굴된 초기 코돈을 갖는 유전자로부터 단백질을 발현시키는 단계를 포함하는, 재조합 단백질의 발현 조절 방법에 관한 것이다.
- <15> 본 발명의 제3면은 상기 방법에 의해 발굴된 초기 코돈을 갖는 유전자로부터 단백질을 발현하는 단계를 포함하는, 재조합 단백질의 생산 방법에 관한 것이다.

효 과

- <16> 본 발명은 단백질 생산에 있어 체계적이고 정확한 조절 수단이 될 수 있으며, 활성 단백질을 원하는 수준으로 고속으로 발현함으로써, 복잡한 단백질 네트워크에 대한 보다 자세한 이해와 포워드 프로테오믹스 성과물의 보충을 가능하게 한다. 특히, 신호 펩타이드(signal peptide)의 발현 강도를 조절하여 다양한 발현 강도를 갖는 신호 펩타이드를 합성할 수 있으며, 결과적으로 단백질 생산을 극대화하는 데에도 적용될 수 있다. 또한 단백질 발현 조절뿐만 아니라, 고속으로 유전자를 변화시키고 이를 빠르게 단백질로 전환함으로써 다양한 분석 기법과 함께 사용 시 단백질의 개량이나 진화에도 이용될 수 있다.

발명의 실시를 위한 구체적인 내용

- <17> 이하, 본 발명을 상세히 설명한다.
- <18> 본 명세서에서, '+n 코돈'은 번역 시작 위치로부터 하류방향(downstream)으로 n번째 위치에 존재하는 코돈을 가리키는 바, 예를 들어 +1, +2, +3, +4, +5 및 +6 코돈은 각각 번역 시작 위치로부터 하류방향으로 첫 번째(시작 코돈(start codon, 예: AT(U)G), 두 번째, 세 번째, 네 번째, 다섯 번째 및 여섯 번째 코돈을 가리킨다.
- <19> 본 발명에서는, 목적 단백질의 초기 코돈의 무작위 돌연변이를 통하여 재조합 단백질의 발현 효율을 정교하게 조절하고자 하였다. 즉, +2 내지 +13 코돈, 특히 +2 및 +3 코돈을 무작위로 변화시키고 변화된 코돈과 단백질 발현 수준간의 상관관계를 분석하였다. 본 발명에서는, 무작위로 변화시킨 초기 코돈을 가진 유전자 라이브러리로부터 발현량을 조사하기 위하여 세포와 무세포 기법을 조합하는 전략을 사용하는 바, 구체적인 과정은 아래와 같다.
- <20> **단계 1)**
- <21> 목적 유전자의 +2 내지 +13 코돈 중 하나 이상, 예를 들어, +2 및 +3 코돈을 무작위로 변화시킨다. 이를 위해, 유전자를 +2 내지 +13 코돈 중 하나 이상, 예를 들어, +2 및 +3 코돈 위치에서 무작위로 변화된 축퇴 정방향 프라이머(degenerate forward primer)를 이용하여 중합효소연쇄반응(polymerase chain reaction, PCR)에 의해 증폭시킨다.
- <22> **단계 2)**
- <23> 단계 1)에서 얻어진 변이체들을 세포 내로 도입하여 개별 유전자로 분리한다. 예를 들어, 변이체들을 벡터에 클로닝하여 세포, 특히 세균 세포, 보다 특히 대장균(*E. coli*) 세포에 형질전환시킨다. 이때 각각의 변이체들은 하나의 세포로 들어가 형질전환된다. 각각의 변이체들을, 예를 들어, 형질전환된 콜로니를 주형으로 사용하여 콜로니 PCR(colony PCR)을 수행함으로써, 증폭한다.
- <24> **단계 3)**
- <25> 단계 2)에서 얻어진 각각의 PCR 산물을 무세포 단백질 합성 시스템을 사용하여 발현한다. 통상적으로, 무세포 단백질 합성을 위한 반응액은 목적 단백질의 합성을 위해 필요한 세포 소기관 및 인자를 포함하는 세포 파쇄액(cell lysate), 아미노산 혼합물, 단백질 합성 에너지원, 유전자 정보원 및 완충용액을 포함한다. 세포 파쇄액을 제조하기 위한 세포로는 대장균, 고초균, 밀배아, 쌀배아, 보리 배아, CHO 세포, 하이브리도마 세포 또는 망상적혈구 세포를 사용할 수 있고, 세포 파쇄액은 문헌 [Ahn, J.H. et al., Nucleic Acids Res. 2007, 35(4), e21] 및 [Ahn, J.H. et al., Biochem. Biophys. Res. Commun. 2005, 338(3), 1346-1352]에 기재된 방법에 따라 제조된 S30 추출액(S30 extract), 한국등록특허 제0733712호(2007. 6. 29 공고)에 기재된 방법에 따라 제조된 S12 추출액 등일 수 있다. 아미노산 혼합물은 글리신, 알라닌, 발린, 류신, 이소류신, 프롤린, 페닐알라닌, 티로신, 트립토판, 시스테인, 메티오닌, 세린, 트레오닌, 리신, 아르기닌, 히스티딘, 아스파테이트, 글루타메이트, 아스파라긴 또는 글루타민으로부터 선택되는 L-아미노산의 혼합물이며, 단백질 합성 에너지원은 ATP, CTP, GTP, TTP 또는 UTP로부터 선택되는 하나 이상이고, 유전자 정보원은 목적 단백질을 코딩하는 DNA 또

는 mRNA일 수 있다. 예를 들어, 무세포 단백질 발현은 멀티-웰 플레이트, 특히 96-웰 플레이트 상, 27~37 °C, 특히 37 °C 항온배양기 내에서 1~5 시간, 특히 3 시간 동안 이루어질 수 있다.

<26> **단계 4)**

<27> 단계 3)에서의 단백질 발현효율을 조사한다. 단백질 발현효율을 조사하기 위해서는 당업계에 알려진 통상의 방법 중 어느 것이라도 사용할 수 있으며, 예를 들어, 방사성 계수(radioactivity counting), SDS-PAGE(sodium dodecyl sulfate polyacrylamide gel electrophoresis) 등을 사용할 수 있다.

<28> 도 1에 본 발명에 따른 방법의 일례를 도식적으로 나타내었다. 간략하게는, 유전자의 +2 및 +3 코돈을 NNN으로 합성한 프라이머를 이용하여 PCR로 무작위로 변화시킨 후, 이렇게 합성된 PCR 산물 혼합물을 플라스미드로 클로닝하여 이를 *E. coli* DH5α에 형질전환시킨다. 아가 배지에서 형질전환된 세포를 배양한 후, 콜로니 PCR을 통하여 개별 콜로니에 들어 있는 유전자를 증폭하여 각각의 PCR 산물을 합성한다. 이 PCR 산물들을 무세포 단백질 합성 반응액에 첨가하여 단백질로 전환시킨다. 합성된 단백질의 발현량을 동위원소로 표지된 류신(leucine)을 측정함으로써 정량화한다.

<29> 본 발명에서는, 4 종류의 단백질에 대해 수 백 개의 코돈 변이체들의 발현량을 조사하였다. 그 결과, 약 70배의 발현량 차이가 나타나는 연속적인 상대 발현량을 보였으며, 원하는 수준의 발현을 나타내는 코돈 조합을 선별해낼 수 있었다. 이는 특정 단백질의 초기 코돈, 특히 +2 및 +3 코돈의 변화를 통해 단백질 발현을 정교하게 조절할 수 있음을 보여주는 것이다. 따라서 본 발명은 단백질 발현수준을 최적화하는 매우 효율적인 방법이라 할 수 있다.

<30> 문헌 [de Valdivia, E.I.G et al., *Nucleic Acids Res.* 2004, 32(17), 5198-5205] 및 [de Valdivia, E.I.G et al., *FEBS J.* 2005, 272(20), 5306-5316]은 초기 코돈의 영향은 발현 진행 중인 라이보솜(ribosome) 복합체에서 펩티딜-tRNA(peptidyl-tRNA)가 이탈되는 빈도와 관련이 있다고 보고하였다. 발현 중인 라이보솜 복합체는 발현 초기 단계에서는 mRNA와 덜 안정하게 결합되어 있는 것으로 알려져 있는데, 코돈에 따라서 펩티딜-tRNA의 이탈 정도가 달라지고 다양한 발현 효율 분포를 나타내는 것으로 판단된다. 그러나, 본 발명에 따른 결과를 보면 초기의 몇 개의 코돈이 다운스트림 박스(downstream box)의 발현 효율을 완전히 결정하는 것은 아니고, 유전자의 다른 서열과 관련된 요소들과 관련하여 작용하는 것으로 판단된다. 또한 이러한 결과는 초기 코돈의 적합한 배열을 탐색하는 고속 발굴 방법의 필요성을 시사하는 것이라 할 수 있다.

<31> 단백질 발현에 대한 초기 코돈의 영향은 무세포 단백질 합성 시스템에 국한되는 것이 아니고, 세포 내에서(*in vivo*) 단백질을 발현할 경우에도 같은 경향을 가진다. 무세포 시스템에서 발현이 잘 되는 코돈 조합은 세포에서의 발현에서도 발현이 증가하는 것으로 확인되었고, 중간 및 낮은 발현을 보이는 클론도 무세포 시스템과 세포 발현 결과가 일치하였다. 따라서, 본 발명에 따른 방법은 단백질 생산에 있어 체계적이고 정확한 조절 수단이 될 수 있다. 또한, 본 발명의 방법은 신호 펩타이드의 염기서열에도 적용하여 신호 펩타이드의 발현 강도를 조절함으로써 다양한 발현 강도를 갖는 신호 펩타이드를 합성할 수 있으며, 결과적으로 단백질 생산을 극대화하는 데에도 적용될 수 있다.

실시예

<32> 이하, 본 발명을 실시예에 의해 보다 구체적으로 설명하나, 이는 본 발명의 이해를 돕기 위한 것일 뿐 본 발명의 범위를 어떤 식으로든 제한하고자 하는 것이 아니다.

<33> **실시예 1: 클로닝 및 콜로니 PCR**

<34> 4 종류의 단백질, DsRed2(AAV97910), EPO(human erythropoietin, CAA26095), UK(serine protease domain of murine urokinase, EDL01484) 및 EGFP(enhanced green fluorescent protein, AAB02572)의 발현을 조절하기 위하여, 각각의 유전자의 AUG 시작 코돈 바로 다음의 +2 및 +3 코돈을 변화시키기로 하였다. 각각의 목적 유전자에 대해 +2 및 +3 코돈 부위가 무작위로 변화된 프라이머를 이용한 PCR 방법으로 목적 유전자를 증폭하였다. 사용된 프라이머는 아래 표 1a에 나타낸 바와 같다.

표 1a

프라이머 명칭	염기서열
F-NNN-DsRed2	AAGAAGGAGATATACATATGNNNNNTCCGAGAACGTCATCAC(서열번호 1)

F-NNN-EPO	AAGAAGGAGATATACATATGNNNNNNCCACGCCTCATCTGTGA(서열번호 2)
F-NNN-UK	TAGAAGGAGATATACATATGNNNNNNCAACAAGGCTTCCAGTG(서열번호 3)
F-NNN-EGFP	AAGGAGATATACATATGNNNNNNAAGGGCGAGGAGCTGTTC(서열번호 4)

<36> (밑줄: 제한 위치, 굵은 체: 시작 코돈)

<37> 증폭된 PCR 산물을 플라스미드(DsRED2, EPO 및 EGFP의 경우 pK7(문헌 [Kim et al., *Eur J Biochem*, 1996, 239(3) 881-886]), UK의 경우 pIVEX2.3d(로슈 어플라이드 사이언스(Roche Applied Science)))에 클로닝하고 대장균 DH5 α에 형질전환하였다. LB-아가 플레이트에서 하루 동안 배양한 후 각각의 콜로니를 찍어서 아래 표 1b에 나타난 T715up과 GTB 프라이머가 있는 PCR 반응액에 옮기고 콜로니 PCR을 수행하였다. 콜로니 PCR의 반응조건은 아래와 같았다: 95 °C에서 5 분, 95 °C에서 30 초, 55 °C에서 1 분 및 72 °C에서 1 분의 30 주기에 이어서 7 분간 신장.

표 1b

프라이머 명칭	염기서열
T715up	TCGATCCCGCGAAATTAATACGACTCACTATAGG(서열번호 5)
GTB	CAAAAACCCCTCAAGACCGTTTA(서열번호 6)

<39> 실시예 2: 콜로니 PCR로 증폭된 유전자의 무세포 단백질 발현

<40> 각 유전자의 +2 및 +3 코돈 위치에서 무작위로 변화된 300 EPO, 222 DsRed2, 300 UK 및 120 EGFP PCR 산물을 무세포 단백질 합성을 위한 발현주형으로 직접 사용하였다. 15 μl의 부피로 이루어진 무세포 단백질 합성 시스템은 다음과 같은 성분을 갖는다: 57 mM Hepes-KOH(pH 8.2), 1.2 mM ATP, 0.85 mM CTP, GTP, UTP, 2 mM DTT, 0.64 mM cAMP, 90 mM 포타슘글루타메이트, 80 mM 암모늄아세테이트, 12 mM 마그네슘아세테이트, 34 μg/ml L-5-포밀-5,6,7,8-테트라하이드로폴산(폴린산), 1 mM 20 아미노산, 0.17 mg/ml *E. coli* 총 tRNA 혼합물(MRE600 균주 유래), 2% PEG(8000), 67 mM 크레아틴포스페이트(CP), 5.6 μg/ml 크레아틴 키나제, 10 M L-[U-¹⁴C]류신(11.3 GBq/mmol, 아머삼 바이오사이언시즈(Amersham Biosciences)), 4μl S30 세포 파쇄액, 1 μl 콜로니 PCR 산물.

<41> S30 추출물은 *E. coli* BL21-Star™(DE3)(인비트로젠(Invitrogen))으로부터 문헌 [Ahn, J.H. et al., *Nucleic Acids Res.* 2007, 35(4), e21] 및 [Ahn, J.H. et al., *Biochem. Biophys. Res. Commun.* 2005, 338(3), 1346-1352]에 기재된 방법에 따라 제조하였다. 3 시간 동안 발현시킨 후, 15 μl의 반응샘플을 반응혼합물로부터 채취하여, 발현된 단백질을 TCA(trichloroacetic acid)에 침전되는 방사성을 액체 섬광계수기(Liquid Scintillation Counter)로 측정하여 정량하였다.

<42> 그 결과를 도 2에 나타내었다. 도 2에 나타난 바와 같이, 4 가지 단백질 모두에서 발현량은 넓고 연속적인 분포를 나타내었다. 예를 들면, 야생의 EPO 유전자는 72 μg/ml의 발현량을 보이는데 비해, +2 및 +3 코돈 변이체들은 8 μg/ml에서 509 μg/ml까지의 발현량을 보였다(도 2A). 이 결과는 초기 코돈이 단백질 발현 효율에 중요한 영향을 미친다는 것을 재차 확인시켜주는 것이다. 도 2B 내지 2D에 나타난 바와 같이, 다른 세 종류의 단백질에서도 비슷한 패턴이 관찰되었다. 콜로니 PCR 산물을 발현한 결과 각각의 야생형에 비해서 EPO의 경우는 11~706%, UK는 12~422%, DsRed2는 12~834%, EGFP는 31~767%의 발현량 분포를 보였다.

<43> [¹⁴C]류신 도입 결과에 기초하여, 높은, 중간 또는 낮은 발현수준을 나타내는 반응혼합물 2 μl를 13% 트리신-SDS-폴리아크릴아마이드 젤 상에서 분석하였다. 발현된 단백질을 쿠마시 블루 염색에 의해 가시화하였다. 그 결과를 도 3(화살표: 발현된 단백질)에 나타내었다.

<44> 높은 수준의 발현을 나타내는 5개의 클론, 중간 수준의 발현을 나타내는 3개의 클론 및 낮은 발현을 나타내는 3개의 클론을 서열분석하고, 그 결과를 아래 표 2에 나타내었다(높은 수준의 발현을 나타내는 클론: 적색, 중간 수준의 발현을 나타내는 클론: 청색, 낮은 발현을 나타내는 클론: 흑색, +2 및 +3 코돈은 굵은체로 나타냄). 발현수준은 세 독립적인 실험의 중간값으로, 모든 경우 표준오차는 10% 미만이었다.

표 2a

초기 코돈의 실제와 번역효율 간의 상관관계

유전자	클론	발현수준 ($\mu\text{g}/\text{ml}$)	서열분석된 +2/+3 코돈
DsRed2	R134	고	726 ATG TGCA ATCCGAG
	R85		715 ATG TTTGA ATCCGAG
	R100		662 ATG TTGAT ATCCGAG
	R123		653 ATG AGAAT GTCGAG
	R112		644 ATG TATGA ATCCGAG
	R33	중	322 ATG GGAAG GTCGAG
	R65		302 ATG GATAG TCGAG
	R75		285 ATG TATAG ATCCGAG
	R56	저	55 ATG GGGAG GTCGAG
	R69		30 ATG AGGAG GTCGAG
	R152		11 ATG AGGTA ATCCGAG
	야생형		87 ATG GCTCT CCGAG
EPO	E171	고	509 ATG TTGCA TCCACGC
	E37		494 ATG TTGAC CCACGC
	E47		493 ATG CTTA TCCACGC
	E40		491 ATG GCGAT TCCACGC
	E50		491 ATG CAAT TCCACGC
	E296	중	351 ATG AADGG ACCACGC
	E180		343 ATG GGA CGCCACGC
	E185		229 ATG CGA CCACGC
	E74	저	51 ATG GCTGT CCACGC
	E36		42 ATG CGGTG CCACGC
	E227		8 ATG CGATG CCACGC
	야생형		72 ATG GCCCC ACCACGC
UK	U171	고	351 ATG TTGAT ACAACAA
	U44		345 ATG AGATG CAACAA
	U174		339 ATG AAAA CGCAACAA
	U220		336 ATG AAATTG CAACAA
	U108		332 ATG AAATT TACAACAA
	U276	중	243 ATG TCAGT CAACAA
	U154		228 ATG GAA CAGCAACAA
	U64		211 ATG AAGCG CAACAA

표 2b

	U107	저	30	ATGGGGTGACAACAA
	U295		10	ATGAAATAACAACAA
	U188		49	ATGCGGTAGCAACAA
	야생형		83	ATGGTGGACCAACAA
EGFP	G004	고	629	ATGAAGGTAAGGGC
	G006		610	ATGACTTATAAGGGC
	G002		583	ATGTTGATGAAGGGC
	G068		578	ATGTCAATAAAGGGC
	G069		577	ATGTCTGACAAAGGGC
	G046	중	407	ATGATTCTTAAGGGC
	G010		291	ATGAAGGCTAAGGGC
	G085		154	ATGATGTTCAAGGGC
	G041	저	87	ATGTCTGGTAAGGGC
	G084		53	ATGTCAGATAAGGGC
	G089		26	ATGTGGTGTAAAGGGC
	야생형		82	ATGGTGACAAAGGGC

<46>

<47> 표 2에 나타난 바와 같이, 염기서열 분석에서 단백질 합성을 증가시키거나 억제시키는 공통적인 코돈을 확인하는 어려웠지만, 발현이 잘 되는 유전자의 초기 서열은 대부분의 경우 A와 T 염기를 많이 가지고 있었다.

<48> 실시예 3: mRNA 분석

<49> 무세포 단백질 합성의 반응물에서 5 μ l를 회수하여 같은 부피의 RNA 프로텍트™ 세균 시약(RNAProtect™ bacterial reagent)(키아젠(Qiagen))을 첨가한 후 혼합하였다. 각각의 시료에서 RNA를 추출한 후 RNeasy 미니 칼럼을 통하여 분리한 후, 1.2% 포름아마이드 아가로스 겔을 통해 mRNA의 농도를 분석하였다.

<50> 그 결과를 도 4에 나타내었다. 도 4에 나타난 바와 같이, 각기 다른 발현수준을 보이는 유전자들 중에서 mRNA의 상대적인 농도는 거의 차이가 없었다. 이로부터 코돈에 따른 발현량의 변화는 mRNA의 합성이나 분해로 인한 것이라기보다는 단백질 합성 단계에서의 효율의 차이로 인해 유발되는 것으로 판단된다.

<51> 또한, Mfold 컴퓨터 프로그램을 이용한 mRNA의 2차 구조 분석을 수행하였다. 그 결과를 도 5에 나타내었다. 도 5에 나타난 바와 같이, +2 및 +3 코돈의 변화로 mRNA의 구조적인 차이와 발현량 간의 상관관계가 없는 것을 확인하였다.

<52> 실시예 5: EGFP 활성 분석

<53> 발현된 EGFP의 생물학적 활성을 분석하기 위하여, 증폭된 EGFP 유전자들(96 EGFP 클론)을 96 웰 플레이트에서 37 °C에서 3 시간 동안 무세포 시스템에서 발현시켰다. 이 과정을 형광분석기 내에서 수행하여 발현된 EGFP의 형광의 세기를 실시간으로 측정하였다(여기 488 nm; 방출 555 nm). EGFP의 형광 세기는 37 °C에서 한 시간 동안 측정하였다. 발현된 EGFP의 형광 이미지는 UV를 조사하여 사진을 얻었다.

<54> 그 결과를 도 6에 나타내었다. 도 6에 나타난 바와 같이, +2 및 +3 코돈 변이체의 PCR 산물을 발현하였을 때, 반응 시간에 따라 단백질이 합성되면서 형광이 함께 증가하는 것을 확인하였다. 각각의 반응 웰 마다 형광 증가 속도의 차이를 보였으며 상대적인 형광 세기는 동위원소를 이용한 측정법으로 얻은 결과와 유사하였다(도 2D). 통계학적 분석 결과, [¹⁴C]Leu 도입과 EGFP 형광이 좋은 상관관계(R^2 값 = 0.971)를 나타내었다. 이는 +2 및 +3 코돈의 무작위 변이체가 합성된 단백질의 활성에는 큰 영향을 주지 않음을 의미하는 것이다.

<55> 실시예 5: 세포 내 단백질 발현

<56> 무세포 단백질 합성 시스템에서 관찰된 초기 코돈의 영향이 세포를 이용한 단백질 발현에서도 비슷한 경향을 가지고 재현되는지 조사하였다. 콜로니 PCR 산물을 pET24ma 벡터(문헌 [Yun et al., *App Environ Microbiol*, 2005, 71(8) 4220-4224])에 클로닝하여 *E. coli* BL21-Star™(DE3)(인비트로젠)에 형질전환한 후 5 ml의 LB 배지에서 배양하였다. OD₆₀₀이 0.6이 되었을 때 1 mM IPTG(isopropyl β -D-thiogalactopyranoside)를 첨가하여 4

시간 추가 배양하였다. 세포 배양액을 회수하여 발현된 단백질을 13% 트리신-SDS-PAGE를 통하여 분석하였다.

<57> 그 결과를 도 7에 나타내었다. 도 7에 나타낸 바와 같이, 무세포 단백질 합성 시스템에서 단백질 합성 효율이 세포 내 단백질 발현 실험에서도 비슷한 경향을 가지는 것을 확인하였다.

<58> **실시예 6: 목적 유전자에 대한 초기 코돈의 특이성 확인**

<59> 특정 유전자에서 높은 발현량을 보였던 +2 및 +3 코돈이 다른 유전자에도 비슷한 발현 영향을 주는지를 확인해 보기 위하여, 가장 발현이 잘되는 +2 및 +3 코돈 조합을 각기 다른 단백질에 옮겨서 발현 결과를 조사해보았다. 그 결과를 아래 표 3에 나타내었다.

표 3

코돈 교환에 의한 상대적 발현효율

유전자	번역 수준 ($\mu\text{g}/\text{ml}$)				
	+2 및 +3 코돈				
	TGC-AAT	TTG-CAT	TTG-ATA	AAG-AGT	야생형 대조군
DsRed2	726 (8.3)	155 (1.8)	210 (2.4)	137 (1.6)	87 (1.0)
EPO	308 (4.3)	509 (7.1)	220 (3.1)	209 (2.9)	78 (1.0)
UK	229 (2.8)	252 (3.0)	351 (4.2)	178 (2.1)	83 (1.0)
EGFP	135 (1.1)	95 (0.8)	179 (1.5)	629 (5.1)	82 (1.0)

<60>

<61> (괄호 내 숫자: 야생형 유전자와 비교한 상대적 발현효율, 굵은체: 4가지 돌연변이 유전자 중에서 가장 높은 발현수준을 나타내는 코돈 배열)

<62> 표 3으로부터 알 수 있는 바와 같이, 모든 단백질은 서로 다른 코돈 조합에서 최대 발현량을 보였으며, 개별 코돈의 교차 발현 증가 효과는 거의 없음을 확인하였다.

도면의 간단한 설명

<63> 도 1은 본 발명에 따른 방법의 일례를 도식적으로 보여주는 도면이고;

<64> 도 2는 콜로니 PCR 산물의 무세포 단백질 합성 결과를 보여주는 그래프이며;

<65> 도 3은 무세포 합성된 단백질의 전기영동 분석결과를 보여주는 사진이고;

<66> 도 4는 정상-상태(steady-state) mRNA 수준 분석결과를 보여주는 사진이며;

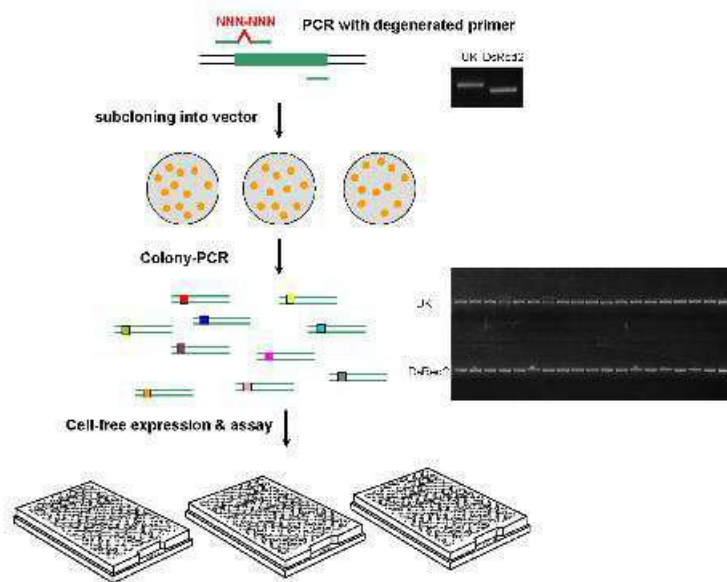
<67> 도 5는 mRNA 이차구조 예측결과를 보여주는 그래프이고;

<68> 도 6은 EGFP 변이체의 발현결과를 보여주는 도면이며;

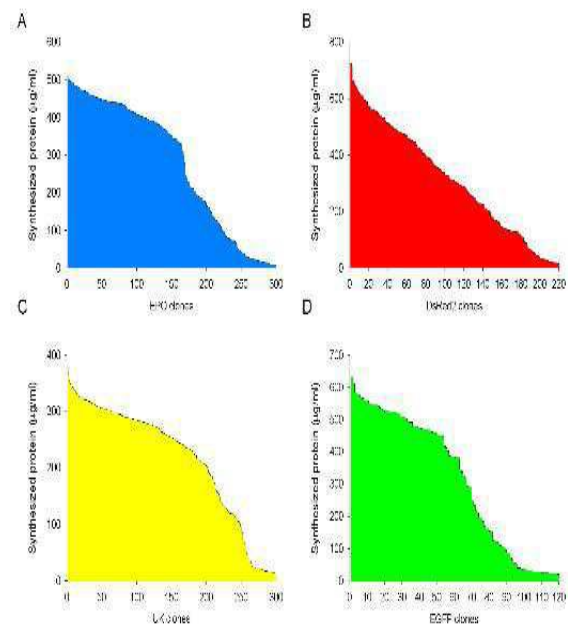
<69> 도 7은 선별된 클론의 세포 내 발현결과를 보여주는 사진이다.

도면

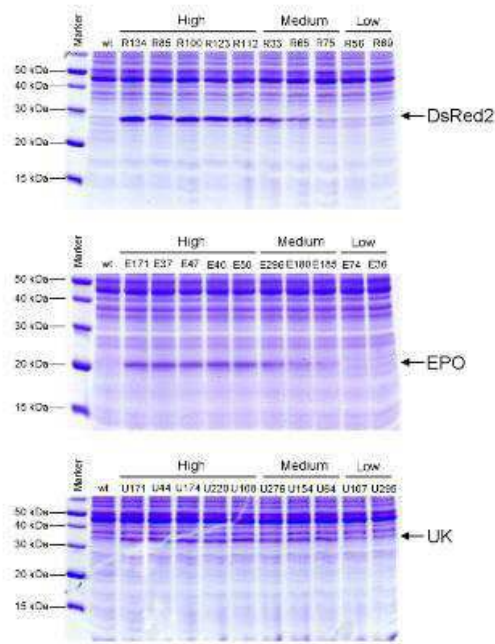
도면1



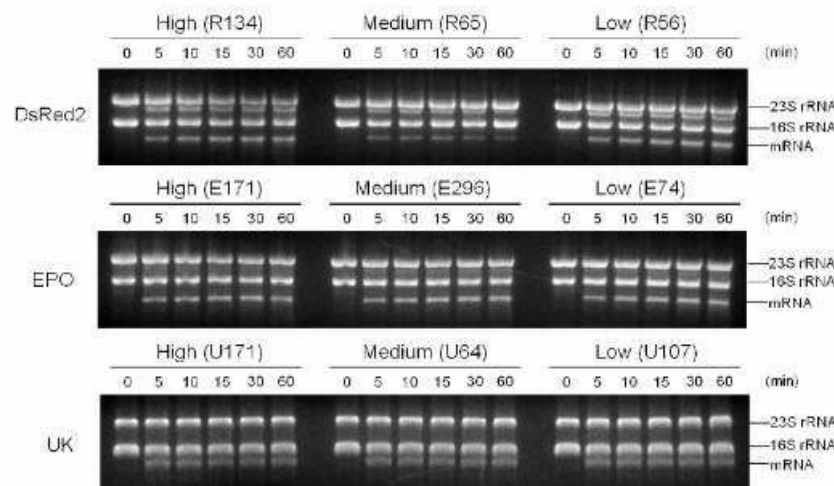
도면2



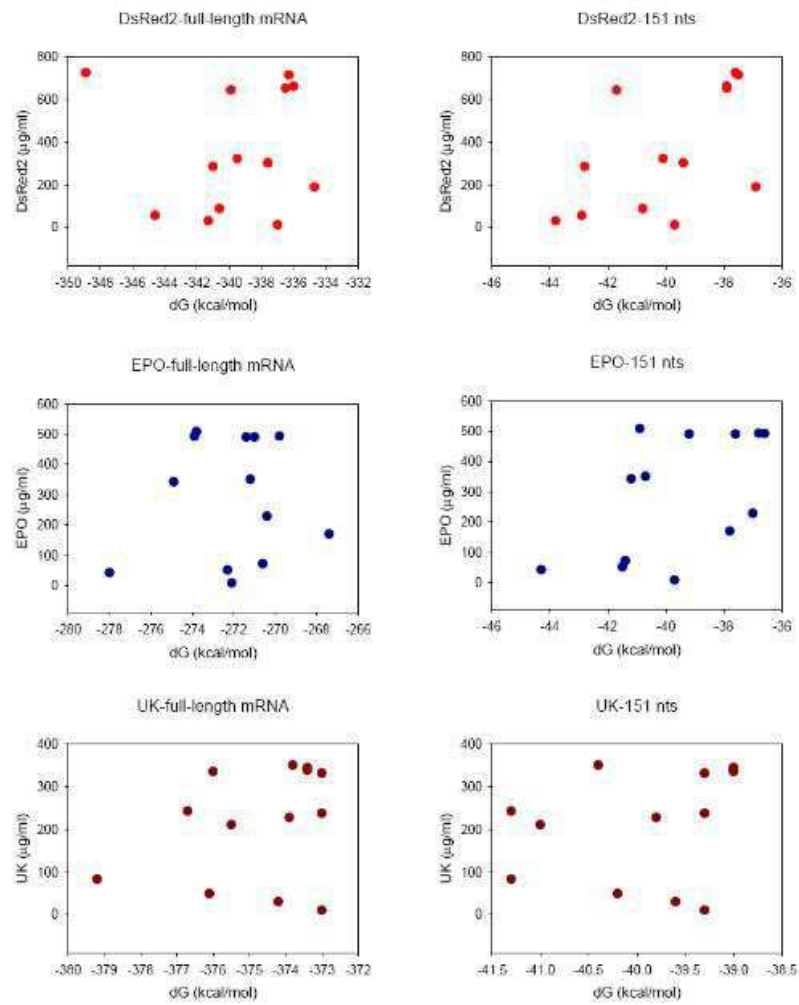
도면3



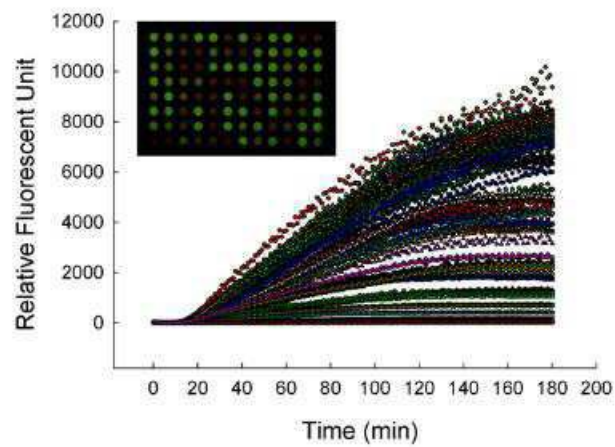
도면4



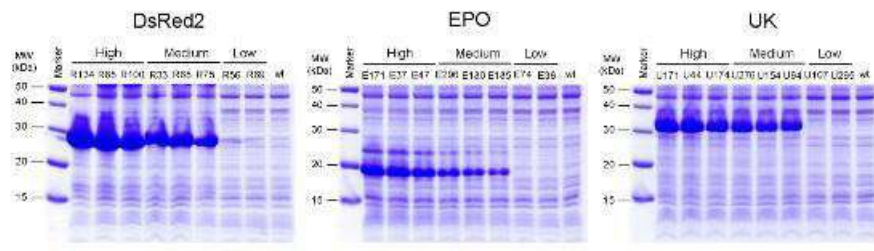
도면5



도면6



도면7



서열 목록

<110> The Industry & Academic Cooperation in Chungnam National University (IAC)
 <120> Methods for screening initial codons providing desired expression levels of proteins, and method for tuning expression and production of recombinant proteins

<160> 6

<170> KopatentIn 1.71

<210> 1

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer F-NNN-DsRed2

<400> 1

aagaaggaga tatacatatg nnnnnntccg agaacgtcat cac

43

<210> 2

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer F-NNN-EPO

<400> 2

aagaaggaga tatacatatg nnnnnnccac gcctcatctg tga 43

<210> 3
 <211> 43
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer F-NNN-UK

<400> 3
 tagaaggaga tatacatatg nnnnnncaac aaggcttcca gtg 43

<210> 4
 <211> 41
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer F-NNN-EGFP

<400> 4
 aaggagatat acatatgnnn nnnaagggcg aggagctgtt c 41

<210> 5
 <211> 34
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer T715up

<400> 5
 tcgatccgc gaaattaata cgactcacta tagg 34

<210> 6
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Primer GTB

<400> 6

caaaaaaccc ctcaagaccc gttta

25