

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0017239

(43) 공개일자 2021년02월17일

(51) 국제특허분류(Int. Cl.)

C12N 7/00 (2006.01) C07K 14/005 (2006.01)

C12N 15/70 (2006.01) C12P 19/34 (2006.01)

(52) CPC특허분류

C12N 7/00 (2013.01)

C07K 14/005 (2013.01)

(21) 출원번호 10-2019-0096197

(22) 출원일자 2019년08월07일

심사청구일자 2019년08월07일

(71) 출원인

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

오민규

경기도 성남시 분당구 정자일로 248, 602동 1601호(정자동, 파크뷰)

안동준

서울특별시 서초구 효령로68길 81, 102동 403호(서초동, 서초자이아파트)

(뒷면에 계속)

(74) 대리인

특허법인리체

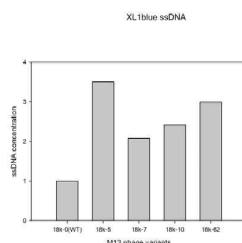
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 DNA 생산량을 증가시키기 위한 M13 변이체

(57) 요약

본 발명은 프로모터, 리보솜 결합 부위 (ribosome binding site; RBS), 및 단백질 5 (p5) 코딩 영역을 포함하는 M13 p5 발현 카세트를 포함하는 유전자에 있어서, 상기 RBS와 상기 p5 코딩 영역 사이의 서열 중 적어도 하나의 염기가 변이된, 유전자에 관한 것으로, 이를 통하여 단일 가닥 DNA의 생산량을 증진시킬 수 있다.

대표도 - 도3



(52) CPC특허분류

C12N 15/70 (2013.01)

C12P 19/34 (2013.01)

(72) 발명자

이승우

서울특별시 중랑구 속선웅주로 6-9, 101동 1101호
(목동, 목동자이아파트1단지)

이보영

서울특별시 성동구 난계로 100, 104동 1505호(하왕
십리동, 왕십리자이)

이준민

서울특별시 성북구 안암로9가길 9-27, 102호 (안암
동4가)

이재원

서울특별시 성북구 고려대로 39, 402호 (안암동2
가)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711092886
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	미래소재디스커버리지원(R&D)
연구과제명	결빙제어 나노소재의 대량생산 및 응용 기술 개발
기 여 율	1/1
과제수행기관명	서울시립대학교
연구기간	2017.09.01 ~ 2023.08.31

명세서

청구범위

청구항 1

프로모터, 리보솜 결합 부위 (ribosome binding site; RBS), 및 단백질 5 (p5) 코딩 영역을 포함하는 M13 p5 발현 카세트를 포함하는 유전자에 있어서, 상기 RBS와 상기 p5 코딩 영역 사이의 서열 중 적어도 하나의 염기가 변이된, 유전자.

청구항 2

제1항에 있어서, 상기 RBS와 상기 p5 코딩 영역 사이의 서열이 서열번호 6 내지 서열번호 9 중 어느 하나의 서열로 치환된, 유전자.

청구항 3

제1항에 있어서, 상기 RBS 서열은 서열번호 3으로 표시되고, 상기 p5 코딩 영역은 서열번호 4로 표시되는, 유전자.

청구항 4

제1항에 있어서, 상기 p5 발현 카세트는 서열번호 10 내지 서열번호 13 중 어느 하나의 서열로 표시되는 서열을 포함하는, 유전자.

청구항 5

제1항의 유전자를 포함하는, M13 박테리오파지.

청구항 6

제1항의 유전자를 포함하는, 숙주 세포.

청구항 7

제5항의 M13 박테리오파지에 감염된, 숙주 세포.

청구항 8

제6항 또는 제7항에 있어서, 상기 숙주 세포는 F 섬모를 발현하는, 숙주 세포.

청구항 9

M13 박테리오파지 내 프로모터, 리보솜 결합 부위 (ribosome binding site; RBS), 및 단백질 5 (p5) 코딩 영역을 포함하는 M13 p5 발현 카세트에서, 상기 RBS와 상기 p5 코딩 영역 사이의 서열 중 적어도 하나의 염기를 변이시키는 단계; 및

상기 변이된 M13 박테리오파지를 숙주 세포에 감염시키는 단계를 포함하는, DNA 생산량을 증가시키는 방법.

청구항 10

제9항에 있어서, 상기 RBS와 상기 p5 코딩 영역 사이의 서열이 서열번호 6 내지 서열번호 9 중 어느 하나의 서열로 치환된, 방법.

청구항 11

제9항에 있어서, 상기 RBS 서열은 서열번호 3으로 표시되고, 상기 p5 코딩 영역은 서열번호 4로 표시되는, 방법.

청구항 12

제9항에 있어서, 상기 p5 발현 카세트는 서열번호 10 내지 서열번호 13 중 어느 하나의 서열로 표시되는 서열을 포함하는, 방법.

청구항 13

제9항에 있어서, 상기 숙주 세포는 F 섬모를 발현하는, 방법.

발명의 설명

기술 분야

[0001] 본 발명은 DNA 생산량을 증가시키기 위한 M13 변이체에 관한 것이다.

배경 기술

[0003] DNA 오리가미를 생산하기 위해서는 그람(gram) 단위 정도의 대량의 단일 가닥 DNA 분자가 필요하다. 그러나 PCR, 스트랩타빈 비드를 이용한 비오틴화된 가닥을 분리하는 방법, 하나의 인산화된 가닥을 램다 뉴클레아제로 분해하는 방법, 우레아-변성 폴리아크릴아마이드 겔 전기영동을 통해 길이에 따라 분리하는 방법, 환형 주형의 회전복제, 또는 플라스미드의 한 가닥을 자르고 순차적으로 분해시키는 것과 같은 통상적인 DNA 합성 방법은 μg 내지 mg 수준의 DNA를 생산할 수 있을 뿐이다.

[0004] 또한, 비용균성 박테리오파지 M13을 이용하여 DNA를 생산할 수 있는데, 이 경우, 1 내지 14mg/L 의 단일 가닥 DNA를 제조할 수 있는 것으로 알려져 있다 (Sambrook, J.; Fritsch, E. F.; Maniatis, T. Molecular cloning. A laboratory manual, 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001.; Bellot, G.; McClintock, M. A.; Chou, J. J.; Shih, W. M. Nat. Protoc 2013, 8 (4), 755-770 DOI: 10.1038/nprot.2013.037.). 이러한 단일 가닥 DNA의 생산량의 한계를 극복하기 위하여 Benjamin Kick, et al., Efficient Production of Single-Stranded Phage DNA as Scaffolds for DNA Origami, Nano letters, 2015, 15, 4672-4676에서는 pH, 용존 산소, 및 기질 공급과 같은 배양 조건을 조절하였으며, 이를 통해 종래에 비하여 2배 가깝게 DNA 생산량을 증가시킬 수 있음을 개시하였다.

[0005] 그러나 여전히 더 높은 수준의 DNA를 효율적으로 생산할 수 있는 방법에 관한 요구가 존재한다.

선행기술문헌

비특허문헌

[0007] (비특허문헌 0001) Benjamin Kick, et al., Efficient Production of Single-Stranded Phage DNA as Scaffolds for DNA Origami, Nano letters, 2015, 15, 4672-4676

발명의 내용

해결하려는 과제

[0008] 본 발명은 M13 박테리오파지의 유전자를 변형시킴으로써 DNA 생산량을 증가시키는 방법을 제공하는 것을 목적으로 한다.

[0009] 본 발명의 다른 목적은 DNA 생산량을 증가시킬 수 있는 박테리오파지, 또는 이를 포함하는 유전자 또는 벡터를 제공하는 것이다.

[0010] 본 발명의 또 다른 목적은, DNA 생산량이 증진된 숙주세포를 제공하는 것이다.

과제의 해결 수단

- [0012] 1. 프로모터, 리보솜 결합 부위 (ribosome binding site; RBS), 및 단백질 5 (p5) 코딩 영역을 포함하는 M13 p5 발현 카세트를 포함하는 유전자에 있어서, 상기 RBS와 상기 p5 코딩 영역 사이의 서열 중 적어도 하나의 염기가 변이된, 유전자.
- [0013] 2. 항목 1에 있어서, 상기 RBS와 상기 p5 코딩 영역 사이의 서열이 서열번호 6 내지 서열번호 9 중 어느 하나의 서열로 치환된, 유전자.
- [0014] 3. 항목 1에 있어서, 상기 RBS 서열은 서열번호 3으로 표시되고, 상기 p5 코딩 영역은 서열번호 4로 표시되는, 유전자.
- [0015] 4. 항목 1에 있어서, 상기 p5 발현 카세트는 서열번호 10 내지 서열번호 13 중 어느 하나의 서열로 표시되는 서열을 포함하는, 유전자.
- [0016] 5. 항목 1의 유전자를 포함하는, M13 박테리오파지.
- [0017] 6. 항목 1의 유전자를 포함하는, 숙주 세포.
- [0018] 7. 항목 5의 M13 박테리오파지에 감염된, 숙주 세포.
- [0019] 8. 항목 6 또는 7에 있어서, 상기 숙주 세포는 F 섬모를 발현하는, 숙주 세포.
- [0020] 9. M13 박테리오파지 내 프로모터, 리보솜 결합 부위 (ribosome binding site; RBS), 및 단백질 5 (p5) 코딩 영역을 포함하는 M13 p5 발현 카세트에서, 상기 RBS와 상기 p5 코딩 영역 사이의 서열 중 적어도 하나의 염기를 변이시키는 단계; 및 상기 변이된 M13 박테리오파지를 숙주 세포에 감염시키는 단계를 포함하는, DNA 생산량을 증가시키는 방법.
- [0021] 10. 항목 9에 있어서, 상기 RBS와 상기 p5 코딩 영역 사이의 서열이 서열번호 6 내지 서열번호 9 중 어느 하나의 서열로 치환된, 방법.
- [0022] 11. 항목 9에 있어서, 상기 RBS 서열은 서열번호 3으로 표시되고, 상기 p5 코딩 영역은 서열번호 4로 표시되는, 방법.
- [0023] 12. 항목 9에 있어서, 상기 p5 발현 카세트는 서열번호 10 내지 서열번호 13 중 어느 하나의 서열로 표시되는 서열을 포함하는, 방법.
- [0024] 13. 항목 9에 있어서, 상기 숙주 세포는 F 섬모를 발현하는, 방법.

발명의 효과

- [0026] 본 발명에 따른 방법을 이용함으로써 종래에 비하여 DNA 생산량을 증진시킬 수 있다.
- [0027] 본 발명의 박테리오파지, 또는 이를 포함하는 유전자 또는 벡터를 숙주 세포에 도입함으로써 종래에 비하여 DNA 생산량을 증진시킬 수 있다.

도면의 간단한 설명

- [0029] 도 1은 본 발명의 일 실시예에 따른, DNA 서열 변이에 따른 박테리오파지의 생산량을 야생형에 대한 비율로 나타낸 그래프이다.
- 도 2는 본 발명의 일 실시예에 따른, 변이된 스페이서 서열이 도입된 XL1blue 균주에서 야생형 대비 박테리오파지의 생산량을 나타낸 그래프이다. 왼쪽부터 순서대로 18k-0, 18k-5, 18k-7, 18k-10, 및 18k-62.
- 도 3은 본 발명의 일 실시예에 따른, 변이된 스페이서 서열이 도입된 XL1blue 균주에서 야생형 대비 단일 가닥

DNA의 생산량을 나타낸 그래프이다. 왼쪽부터 순서대로 18k-0, 18k-5, 18k-7, 18k-10, 및 18k-62.

도 4는 본 발명의 일 실시예에 따른, 변이된 스페이서 서열이 도입된 Jude1 DH10B F' 균주에서 야생형 대비 박테리오파지의 생산량을 나타낸 그래프이다. 왼쪽부터 순서대로 18k-0, 18k-5, 18k-7, 18k-10, 및 18k-62.

도 5는 본 발명의 일 실시예에 따른, 야생형 M13mp18k (18k-0)의 p5 단백질 발현 카세트 DNA 서열을 나타낸 것이다.

도 6은 본 발명의 일 실시예에 따른, 스페이서 서열이 CTGAC로 변이된 M13mp18k 변이체 (18k-5)의 p5 단백질 발현 카세트 DNA 서열을 나타낸 것이다.

도 7은 본 발명의 일 실시예에 따른, 스페이서 서열이 TCGTA로 변이된 M13mp18k 변이체 (18k-7)의 p5 단백질 발현 카세트 DNA 서열을 나타낸 것이다.

도 8은 본 발명의 일 실시예에 따른, 스페이서 서열이 TCTTC로 변이된 M13mp18k 변이체 (18k-10)의 p5 단백질 발현 카세트 DNA 서열을 나타낸 것이다.

도 9는 본 발명의 일 실시예에 따른, 스페이서 서열이 GAGGT로 변이된 M13mp18k 변이체 (18k-62)의 p5 단백질 발현 카세트 DNA 서열을 나타낸 것이다.

도 10은 본 발명의 일 실시예에 따른, 야생형 M13mp18k의 전장 DNA 서열을 나타낸 것이다.

도 11은 본 발명의 일 실시예에 따른, 야생형 M13mp18k (18k-0)의 스페이서 서열, 및 스페이서 변이 서열을 나타낸 것이다.

발명을 실시하기 위한 구체적인 내용

[0030] 이제 본 발명은 첨부된 도면을 참조로 하기에서 더욱 충분히 기술될 것이며, 그러나 본 발명의 전부가 아닌 단지 일부의 구체예가 예시된다. 실제로, 이들 발명은 많은 다양한 형태로 구체화될 수 있으며, 본원에 제시된 구체예로 제한되는 것으로 해석되어서는 안된다. 본 명세서 및 첨부된 청구범위에 사용되는 단수 형태는 달리 명확하게 지시하지 않는 한 복수한 대상을 포함한다.

[0031] M13 파지는 F 섬모를 통해 *E. coli*를 감염시키는 섬유상 파지이다. M13 파지가 F 섬모를 갖는 *E. coli*에 들어가면, M13 DNA는 숙주 세포의 DNA 중합효소, RNA 중합효소 및 리가아제와 같은 효소를 이용하여 스스로 복제한다. 이러한 과정은 회전환 복제(rolling circle replication)라고 알려져 있다.

[0032] 회전환 복제는 3 단계로 구성된다. 단계 1은 감염 및 RF (replication form) 형성 단계이다. M13 주형 (+) 가닥이 세포질로 들어가면, 이의 상보 (-) 가닥이 형성된다. 그 후, *E. coli*의 DNA 자이라아제(gyrase)가 슈퍼코일된 이중 가닥 M13 DNA (RF)를 형성시킨다. 단계 2에서, 파지 단백질 p2가 RF의 (+) 가닥을 자른다. 3' 수산기가 DNA 중합효소의 프라이머로서 작용하고 새로운 (-) 가닥을 주형으로 하여 (+) 가닥을 합성한다. 완전한 (+) 가닥이 형성되면, p2는 (+) 가닥을 자르고 (+) 가닥을 분리한다. 이로써 하나의 RF DNA 및 하나의 단일 (+) 가닥이 형성된다. 이 단계에서, 단백질 5(p5)가 조절자로서 작용한다. p5의 농도가 높으면 단계 3으로 진행되고, p5의 농도가 낮으면 단계 2를 반복한다. 마지막으로 단계 3은 방출 단계이다. p5 농도가 충분히 높아지면, 수천 개의 p5가 단계 2에서 제조된 단일 (+) 가닥에 결합하여 (+) 가닥-p5 이량체를 형성한다. 상기 이량체는 방출 신호로서 작용하고 파지 단백질 p1, p4, 및 p11로 구성된 수송체를 통해 방출된다. 상기 이량체가 수송체를 통과하면, p5는 제거되고 수개의 복제된 p3, p6, p7, p8 및 p9가 (+) 단일 가닥 DNA와 조합되어 파지 캡시드를 형성한다.

[0033] 복제 속도가 일정하다는 가정 하에서, 단계 2에서, 단일 가닥 DNA (ssDNA)는 피보나치(Fibonacci) 수열로서 증가되고, 단계 3에서는 산술적(arithmetic)으로 증가된다:

[0034] 단계 2의 DNA 증가 수열

$$\frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right)$$

[0035]

[0036] 단계 3의 DNA 증가 수열

$$\frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right) * x$$

[0037]

[0038] 상기 식에서 n은 사이클의 반복 횟수이며, x는 단계 2에서 만들어진 (+) 가닥 DNA 수 이다.

[0040] 본 발명은 단일 가닥 DNA 생산을 증폭시키기 위하여, 단계 2를 오랫동안 유지하는 것이 바람직하다는 가정을 기초로 개발되었다. 회전환 복제에서, p5의 농도는 단계 2와 단계 3의 분기점이 되며, p5가 빠르게 축적된다면, 단계 3으로 빠르게 넘어갈 것이지만, p5가 느리게 축적된다면, 단계 3으로는 천천히 넘어가고, 더 많은 단일 가닥 DNA가 생산될 것이다.

[0041] 본 발명자는 p5가 축적되는 것을 번역 수준에서 조절하고자 하였으며, 이를 위하여 RBS, 스페이서 서열 및 N-말단 코딩 서열을 변이시킴으로써 p5의 발현 수준을 조절하고자 하였다. 상기 “스페이서 서열”은 RBS와 개시 코돈 사이에 있는 염기 서열을 의미한다. 본 발명의 발명자는 p5 발현 카세트 중 개시 코돈 상류에서 총 71개의 유전자 변이체의 p5 발현 수준을 예측하고, 이로부터 4개의 발현 카세트를 제작하였으며, 스페이서 서열이 변이되는 경우, 파지 생산량 및 DNA 생산량이 증가함을 확인함으로써 본 발명을 완성하였다.

[0042] 야생형 스페이서 서열은 5개의 핵산 (TCACA, 서열번호 5)으로 이루어지며, 이 서열은 당해 기술분야에서 잘 알려진 방법들에 따라 변이시킬 수 있다. 예를 들어, 위치지정 변이, 또는 무작위 변이를 이용할 수 있으며, 무작위 변위로서, 예를 들어, 에러-프론 PCR 방법, 롤링 서클 에러-프론 PCR 방법, 돌연변이 유발 균주 이용 방법, 상시적 돌연변이 유발 균주 이용 방법, 삽입 돌연변이 생성 방법, 에틸 메탄설포네이트 이용 방법, 아질산 이용 방법, DNA 서플링 방법을 이용, dNTP-돌연변이 생성 방법을 이용할 수 있으나, 이에 제한되는 것은 아니다.

[0043] 스페이서 서열은 1개의 뉴클레오타이드, 2개의 뉴클레오타이드, 3개의 뉴클레오타이드, 4개의 뉴클레오타이드, 또는 5개의 뉴클레오타이드가 변이될 수 있으며, 2개 내지 4개의 뉴클레오타이드가 변이되는 경우, 변이된 서열은 연속적인 뉴클레오타이드 또는 부분적으로 불연속적인 뉴클레오타이드, 또는 모두 불연속적인 뉴클레오타이드일 수 있다. 예를 들어, 상기 변이된 스페이서 서열은 GTCTT, TATTA, CTCTA, TGGAT, CTGAC, CATTC, TCGTA, GTTCT, GCATT, TCTTC, CGGT, CGGC, TTATT, GAGT, TATTT, ACACA, GATTG, AAGTC, CGTAA, CTGGC, ATGGA, CAGGC, CGGGC, TTGTA, TCGTT, CCAAC, TGTAT, GAGTA, AGTAT, TTGTC, GTTTA, CCGGC, ATCGA, GTTGT, TTTTA, TGTCT, TTAA, GTGTC, AAGGA, AAGTA, TGCTA, ATGGT, CGTTA, GTGAA, GAAAT, AGGTT, GTGTA, GACTC, ATTCG, TAATT, CATGA, TCTTA, TTTT, GCTAA, GTTTT, ATGCT, CAGTT, TACTA, AGCCA, CCCTC, CACTA, GAGGT, AACAA, TTAGG, GAGTT, CGATC, CACAG, TTTTG, TTTGA, CGGTA, 또는 GATTT로 변이될 수 있고, 바람직하게는 CTGAC, TCGTA, TCTTC, TTGTA, CCAAC, TGTAT, GAGTA, AGTAT, TTGTC, GTTTA, ATCGA, GTTGT, TTTTA, TGTCT, TTAA, GTGTC, AAGGA, AAGTA, TGCTA, ATGGT, CGTTA, GTGAA, GAAAT, AGGTT, GTGTA, GACTC, ATTCG, TAATT, CATGA, GCTAA, GTTTT, ATGCT, CAGTT, TACTA, AGCCA, CCCTC, CACTA, GAGGT, GAGTT, CGATC, CACAG, TTTTG, TTTGA, CGGTA, 또는 GATTT로 변이될 수 있으며, 더욱 바람직하게는 CTGAC, TCGTA, TCTTC, 또는 GAGGT로 변이될 수 있다.

[0044] 본 발명에서 사용하는 용어 “조절 영역(regulatory region)”은 핵산 서열의 발현을 조절하는 핵산 서열을 지칭한다. 조절 영역은 특정 핵산을 발현하는데 선천적으로 책임을 지고 있는 서열(동종 영역)을 포함할 수 있으며, 상이한 단백질이나 나아가 합성 단백질을 발현하는 책임을 지고 있는 유래가 다른 서열(이종 영역)을 포함할 수 있다. 특히, 상기 서열들은 특정 또는 비특정 방식 및 유도성 또는 비유도성 방식으로 유전자의 전사를 촉진 또는 억제하는, 원핵, 진핵, 또는 바이러스 유전자의 서열 또는 이로부터 유래한 서열일 수 있다. 조절 영역은 복제 원점, RNA 스플라이스 부위, 프로모터, 인핸서, 전사 종결 서열, 및/또는 폴리펩티드의 표적 세포의 분비 경로로 유도하는 분비 신호를 포함한다.

[0045] 본 발명에서 사용하는 용어 “프로모터” 및 “프로모터 서열”은 상호교환적으로 사용되며, 암호화 서열 또는 기능적 RNA의 발현을 조절할 수 있는 DNA 서열을 지칭한다. 일반적으로, 암호화 서열은 프로모터 서열의 3'쪽에 위치한다. 프로모터는 그 전체가 천연형 유전자로부터 유래되거나, 자연에서 발견되는 상이한 프로모터로부터 유래한 상이한 요소들로 구성되거나, 합성 DNA 절편들을 포함할 수 있다. 상이한 프로모터들이 상이한 조직 또는 세포형 내의 유전자, 또는 상이한 발전 단계에 있는 유전자, 또는 상이한 환경적 또는 생리적 조건에 반응하는 유전자의 발현을 유도할 수 있다. 대부분의 경우 조절 서열의 정확한 경계가 완벽히 나타나 있지 않기 때문에

상이한 길이의 DNA 단편이 동일한 프로모터 활성을 가질 수 있음이 인식된다.

- [0046] 일반적으로 전사 개시 부위는 프로모터 서열 하류(downstream)에 작동 가능하게 연결되며, 상류(5' 방향)로 확장되어 백그라운드 이상의 검출가능한 수준으로 전사를 개시하는데 필요한 최소한의 개수의 염기 또는 요소를 포함한다. 상기 프로모터 서열 내에는, RNA 중합효소의 결합을 책임지는 단백질 결합 도메인(공통 서열(consensus sequence)) 뿐만 아니라 전사 개시 부위(예를 들어 뉴클레아제 S1을 이용한 맵핑(mapping)에 의해 간편하게 정의됨)가 발견될 것이다.
- [0047] 본 발명에서 사용하는 용어 "상류(upstream)"는 언급된 핵산 서열의 5'쪽에 위치하는 뉴클레오티드 서열을 지칭한다. 특히, 상류 뉴클레오티드 서열은 일반적으로 암호화 서열 또는 전사 출발점의 5'쪽에 위치한 서열에 관한 것이다. 예를 들어, 대부분의 프로모터들은 개시 코돈의 상류에 위치한다. 본 발명에서 사용하는 용어 "하류(downstream)"는 언급된 핵산 서열의 3'쪽에 위치하는 뉴클레오티드 서열을 지칭한다.
- [0048] 본 발명에서 사용하는 용어 "작동가능하게 연결된(operably linked)"은 하나의 기능이 다른 것에 의해 영향을 받도록 하기 위하여 단일 핵산 단편 상의 핵산 서열의 회합(association)을 지칭한다. 즉, 핵산 발현 조절 서열(예: 프로모터, 시그널 서열, 또는 전사조절인자 결합 위치의 어레이)과 다른 핵산 서열 사이의 기능적인 결합을 의미하며, 이에 의해 상기 조절 서열은 상기 다른 핵산 서열의 전사 및/또는 해독을 조절하게 된다. 예를 들어, 프로모터는 암호화 서열의 발현에 영향을 줄 수 있을 때(즉, 암호화 서열이 프로모터의 전사 조절 하에 있는 것), 암호화 서열에 작동 가능하게 연결된다. 암호화 서열은 센스 또는 안티센스 배향으로 조절 서열에 작동 가능하게 연결될 수 있다.
- [0049] 본 발명에서 사용하는 용어 "리보솜 결합 부위(ribosome binding site; RBS)"는 단백질 번역 중 리보솜이 결합되는 mRNA 전사체의 개시 코돈의 상류(upstream)에 존재하는 뉴클레오티드 서열이다. 진핵 생물 또는 진핵 생물을 감염시키는 바이러스의 mRNA에서 발견되는 내부 리보솜 유입 영역(internal ribosome entry sites; IRES)도 있지만, 대부분 RBS는 박테리아 서열을 의미한다.
- [0050] M13 박테리오파지는 *E. coli*를 감염시킬 수 있는 바이러스로서, 주로 약 2700개의 껍질 단백질 P8 (pVIII)로 둘러 쌓인 얇고 유연한 튜브형 단일 가닥 DNA로 구성되며 유연한 튜브 형태이며, 추가의 껍질 단백질 p3, p6, p7, p9을 더 포함한다. 감염은 p3이 박테리아의 F 섬모의 말단에 수용체가 도달하면서 개시된다. 세포질에 존재하는 파지 단백질은 p2, p10, 및 p5이고, 수 천개의 p5는 새로 합성된 단일 가닥 DNA에 결합하여 RF DNA가 되는 것을 방지한다. p5-DNA 구조체는 길이가 약 800nm, 지름은 약 8nm이다.
- [0051] 본 발명에서 사용하는 용어 "코딩 영역(coding region)"은 아미노산으로 번역되는 코돈으로 구성된 핵산의 일부이다. "정지 코돈(stop codon)(TAG, TGA 또는 TAA)"은 아미노산으로 번역되지 않아도 코딩 영역의 일부로 간주될 수 있으나, 임의의 플랭킹 서열(flanking sequence), 예를 들면, 프로모터, 리보솜 결합 부위, 전사 종결자 등은 코딩 영역의 일부로 간주되지 않는다. 임의의 핵산 또는 핵산 단편은 단일 폴리펩티드 또는 단편, 유도체 또는 이의 변이체를 코딩할 수 있거나, 또는 하나 이상의 폴리펩티드를 코딩할 수 있으며, 예를 들면, 하나의 핵산이 둘 이상의 폴리펩티드를 코딩할 수 있다. 또한, 핵산은 프로모터와 같은 조절 요소, 리보솜 결합 부위 또는 전사 종결자를 포함할 수 있거나, 또는 하나의 단백질 코딩 영역에 융합된 이중의 코딩 영역, 예를 들면, 분비성 신호 펩티드 또는 이중 기능성 도메인과 같은 특화된 요소 또는 모티프를 코딩할 수 있다.
- [0052] 본 발명에서 사용하는 용어 "카세트(cassette)", "발현 카세트" 및 "유전자 발현 카세트"는 특정 제한효소 부위에서 또는 동종 재조합에 의해 핵산 또는 폴리뉴클레오티드로 삽입될 수 있는 DNA의 절편을 지칭한다. 상기 DNA의 절편은 원하는 폴리펩티드를 암호화하는 폴리뉴클레오티드를 포함하며, 상기 카세트 및 제한효소 부위는 전사 및 번역을 위한 적절한 리딩 프레임(reading frame)에서 카세트의 삽입을 보장하도록 고안된다. "형질전환 카세트"는 원하는 폴리펩티드를 암호화하는 폴리뉴클레오티드를 포함하며, 특정 숙주 세포의 형질전환을 용이하게 하는 폴리뉴클레오티드 외의 요소를 갖는 특정 벡터를 지칭한다. 카세트, 발현 카세트, 유전자 발현 카세트 및 형질전환 카세트는 숙주 내 원하는 폴리펩티드를 암호화하는 폴리뉴클레오티드의 발현을 조절하는 요소를 포함할 수도 있다. 이러한 요소들에는 프로모터, 인핸서, 반응 인자, 종결자 서열, 폴리아데닐레이션 서열 등이 포함되나, 이에 제한되지 않는다.
- [0053] 본 발명에서 사용하는 용어 "유전자"는 전사에 의해서만 생성되는 기능성 분자(예컨대, 생활성 RNA종) 또는 전사 및 번역에 의해 생성되는 기능성 분자(예컨대, 폴리펩티드)를 포함하는 기능성 분자를 암호화하는 뉴클레오티드를 포함하는 폴리뉴클레오티드를 지칭한다. 상기 용어 "유전자"는 cDNA 및 게놈 DNA 핵산을 포함한다. "유전자"는 암호화 영역에 선행하는 조절 서열(5'-비암호화 서열) 및 후행하는 조절 서열(3'-비암호화 서열)을 포

합하는, 특정 RNA, 단백질 또는 폴리펩티드를 발현하는 핵산 단편을 지칭하기도 한다. "야생형(native) 유전자"는 자기 자신의 조절 서열을 갖는, 자연에서 발견되는 유전자를 지칭한다. "키메라 유전자"는 자연에서는 함께 발견되지 않는 조절 및/또는 암호화 서열을 포함하는, 천연 유전자가 아닌 임의의 유전자를 지칭한다. 따라서, 키메라 유전자는 다른 공급원으로부터 유래된 조절 서열 및 암호화 서열을 포함하거나, 동일 공급원으로부터 유래되었지만, 자연에서 발견되는 것과는 상이한 방식으로 배열된 조절 서열 및 암호화 서열을 포함할 수 있다. 키메라 유전자는 다른 공급원으로부터 유래된 암호화 서열 및/또는 다른 공급원으로부터 유래된 조절 서열을 포함할 수 있다. "내생적(endogenous) 유전자"는 생물의 게놈 내에 원래의 위치에 있는 천연형 유전자를 지칭한다. "외래(foreign)" 유전자 또는 "이종(heterologous)" 유전자는 정상적으로는 숙주 생물에서 발견되지 않지만, 유전자 전달에 의해 숙주 생물로 전달된 유전자를 지칭한다. 외래 유전자는 비천연형 생물 내로 삽입된 천연형 유전자, 또는 키메라 유전자를 포함할 수 있다. "이식 유전자(transgene)" 또는 "전이 유전자"는 유전자 전달 과정에 의해 세포 내로 도입된 유전자이다.

[0054] 본 발명에서 사용하는 용어 “변이”는 하나 또는 복수의 핵산의 치환, 결손, 부가 또는 대체를 의미할 수 있다.

[0055] 본 발명에서 사용하는 용어 "발현"은 코딩 서열에 의해 코딩된 생성물의 생물학적 생산을 지칭한다. 대부분의 경우에서, 코딩 서열을 포함하는 DNA 서열은 전사되어 메신저-RNA(mRNA)를 형성한다. 이어서, 상기 메신저 RNA는 번역되어 관련 생물학적 활성을 갖는 폴리펩티드 생성물을 형성한다. 또한, 발현 과정은 RNA 전사 생성물에 대한 추가 가공 단계(예를 들면, 인트론을 제거하기 위한 스플라이싱), 및/또는 폴리펩티드 생성물의 번역-후가공을 포함할 수 있다.

[0056] 본 발명에서 사용하는 용어 “숙주 세포”는 원핵(예를 들어, 박테리아) 및 진핵(예를 들어, 곰팡이, 효모, 동물, 곤충, 식물) 세포를 포함하며 DNA 생산에 적절한 임의의 세포일 수 있고, F 섬모를 갖는 세포인 것이 바람직하다. 적절한 원핵 숙주 세포는 F 섬모를 갖는 E.coli로서, 예를 들어, DH5 α, XL1blue, 또는 Jude1 DH10B를 포함하나 이에 제한되지 않는다.

[0057] 본 발명에서 사용하는 용어 "프라이머"는 표적 핵산과 혼성화하여 적절한 조건하에서 DNA 합성을 위한 개시점으로 작용할 수 있는 이중 가닥의 핵산 영역을 형성하는 올리고뉴클레오티드를 지칭한다. 이러한 프라이머들은 중합효소 연쇄반응에서 사용되거나 DNA 서열분석을 위해 사용될 수 있다.

[0058] 이하에서는 본 발명에 따른 방법에 관하여 자세히 설명한다.

[0059] M13 박테리오파지 내 프로모터, 리보솜 결합 부위 (ribosome binding site; RBS), 및 단백질 5 (p5) 코딩 영역을 포함하는 M13 p5 발현 카세트에서, 상기 RBS와 상기 p5 코딩 영역 사이의 서열 중 적어도 하나의 염기, 적어도 2개의 연속 또는 불연속 염기, 적어도 3개의 연속 또는 불연속 염기, 적어도 4개의 연속 또는 불연속 염기, 또는 5개의 염기를 변이시키고, 상기 변이된 M13 박테리오파지를 숙주 세포에 감염시킨 후 상기 숙주 세포를 배양함으로써 DNA 생산량을 증가시킬 수 있다.

[0060] 상기 배양은 당해 기술 분야에서 잘 알려진 적합한 배양 조건 하에서 적합한 배양 배지를 이용하여 수행될 수 있다. 배양 배지 및 배양 조건은 당해 기술 분야에서 통상의 지식을 가진 자에 의해 변경될 수 있다. 예를 들어, 배양 배지는 액체 배지일 수 있으나, 이에 제한되는 것은 아니다. 배양 방법은, 예를 들어, 배치 배양(batch culture), 연속 배양(continuous culture), 페드-배치 배양(fed-batch culture), 또는 이의 조합일 수 있으나, 이에 제한되는 것은 아니다.

[0061] 배양 배지는 특정 균주를 위한 적합한 조건이 필요하고, 이는 통상의 기술자에 의해 적합하게 변경될 수 있다. 배양 배지는 다양한 탄소원, 질소원 및 소량 원소를 포함할 수 있다. 배양 배지에 이용 가능한 탄소원은, 이에 제한되는 것은 아니지만, 예를 들어 글루코오스, 수크로오스, 락토오스, 프룩토오스, 말토오스, 전분 및 셀룰로오스와 같은 당이나 탄수화물; 예를 들어 콩기름, 해바라기유, 피마자유, 코코넛유와 같은 오일 및 지방; 예를 들어 팔미트산, 스테아르산, 및 리놀레산과 같은 지방산; 글리콜, 및 에탄올과 같은 알코올; 아세트산과 같은 유기산을 포함할 수 있고, 이러한 탄소원은 개별적으로 또는 조합하여 사용할 수 있다. 배양 배지로 사용 가능한 질소원은, 이에 제한되는 것은 아니나, 펩톤, 효모 추출물, 고기 추출물, 몰트 추출물, 옥수수 침지액, 콩가루 및 우레아와 같은 유기물; 및 암모늄 셀레이트, 암모늄 클로라이드, 암모늄 포스페이트, 암모늄 카보네이트, 및 암모늄 나이트레이트와 같은 비유기물을 포함할 수 있고, 상기 질소원은 개별적으로 또는 조합하여 사용될 수 있다. 배양 배지로서 이용 가능한 인원은, 이에 제한되는 것은 아니나, 포타슘 디하이드로젠 포스페이트, 디포타슘 하이드로젠 포스페이트 및 이에 상응하는 나트륨 함유 염을 포함할 수 있다. 일부 구체예에서, 배양 배

지는 마그네슘 설페이트 또는 철 설페이트와 같은 성장에 필수적인 금속 염을 포함할 수 있으나, 이에 제한되는 것은 아니다. 일부 구체예에서, 배양 배지는 상기 성분들에 추가하여, 성장에 필수적인 아미노산 및 비타민을 더 포함할 수 있다. 일부 구체예에서, 배양 배지는 적합한 전구체를 또한 포함할 수 있다. 배양 배지 또는 개별 성분은 적합한 방식으로, 예를 들어, 배치 또는 연속적인 방식으로 배양 용액에 첨가될 수 있으나, 이에 제한되는 것은 아니다.

[0062] 일부 구체예에서, 배양 과정 중에, 암모늄 하이드록사이드, 포타슘 하이드록사이드, 암모니아, 인산, 또는 황산과 같은 화합물을 배양액에 적합한 방식으로 첨가함으로써 배양액의 pH를 선택된 미생물에 맞게 조절될 수 있다. 일부 구체예에서, 배양하는 동안, 배양액에서 거품이 생기는 것은 지방산 폴리글리콜 에스테르와 같은 소포제를 사용함으로써 억제될 수 있다. 배양액을 호기성 조건으로 유지하기 위하여, 산소 또는 산소 함유 기체(예를 들어, 공기)를 배양액에 공급할 수 있다. 예를 들어, 배양액의 온도는 약 20℃ 내지 약 45℃의 범위에서 유지될 수 있고, 일부 구체예에서는 약 25℃ 내지 약 40℃ 범위에서 유지될 수 있다. 예를 들어, 희망하는 DNA의 양에 도달될 때까지, 예를 들어, 약 8시간 내지 약 160시간 동안 배양할 수 있다.

[0064] 이하에서는 구체적인 실시예로서 본 발명을 설명한다.

[0066] 실시예 1. p5 발현 카세트 변이체 및 이에 따른 발현 양의 예측

[0067] p5를 발현하기 위한 p5 발현 카세트는 프로모터, RBS, 및 p5 코딩 서열이 포함되고, 상기 RBS 및 p5 코딩 서열 사이에는 수개의 염기서열이 존재하는데, 본 발명에서는 이를 스페이서라고 명명한다.

[0068] 본 발명자는 New England Biolabs에서 구입한 M13mp18 벡터를 열 충격(heat shock) 방법을 이용하여 E.coli XL1blue에 도입하여 이중 가닥 DNA(dsDNA)를 생산하고 Exprep™ Plasmid SV(GeneA11, 한국)으로 정제하였다. 상기 생산된 dsDNA에 분자 마커로서 카나마이신 저항성 유전자 KanR을 도입하였으며, 이를 M13mp18k (서열번호 18)로 명명하였다.

[0069] UTR designer (https://sbi.postech.ac.kr/utr_designer/)를 이용하여 프로모터 영역 및 리보솜 결합 부위(RBS)를 확인함으로써, RBS 서열과 p5 코딩 영역 사이의 스페이서 서열을 확인하였으며, 프로모터, RBS, 스페이서 및 p5 코딩 영역을 포함하는 야생형 p5 발현 카세트 DNA 서열을 서열번호 1로 표시하였다.

[0070] 스페이서 서열을 무작위 변이시키기 위하여, M13mp18k 벡터를 두개의 단편으로 분리시킨 후, 첫번째 단편을 M13mp18pV TCACA random mut. 1F (ttaaatcgcataagtgtaatNNNNNATGATTAAAGTTGAAATTAACCATC, 서열번호 14), 및 M13mp18pV TCACA random mut. 1R (attttgacgcTCAATCGTCTGAAATGGATTATTAC, 서열번호 15)으로, 두번째 단편을 M13mp18pV TCACA random mut. 2F (agacgattgaGCGTCAAAATGTAGGTATTTTC, 서열번호 16), 및 M13mp18pV TCACA random mut. 2R (ATTACCTTATGCGATTTAAGAAC, 서열번호 17)를 프라이머로 이용하여 각각 58℃ 및 59℃를 어닐링 온도로 하여 PCR하고 접합시킴으로써 p5 단백질의 발현을 감소시킬 수 있는 변이체 후보 71가지 (18k-1 내지 18k-71, 도 11)를 제작하였다.

[0071] 제작된 각각의 변이체들을 E.coli XL1blue 균주에 도입하고 15ml 코니칼 튜브(conical tube)에 담긴 5ml LB 배지에 접종하여 배양하였으며, 이로부터 M13 파지의 생산량을 측정하고, 이를 야생형 M13을 동일 조건으로 배양하여 생산된 생산량에 대한 비로써 나타내었다 (도 1).

[0072] 이 중 야생형 M13에 비하여 상대적 파지 생산량이 1 이상인 변이체에 대하여 250ml 플라스크 내 25ml M9 2x 배지에서 파지를 생산하여 야생형에 대한 상대적 파지 생산량을 확인하였으며, 일관성 있는 결과를 나타낸 18k-5, 18k-7, 18k-10, 및 18k-62 변이체를 선별하였다.

표 1

파지	스페이서 서열
18k-0(WT) (서열번호 1)	TCACA (서열번호 5)
18k-5 (서열번호 10)	CTGAC (서열번호 6)
18k-7 (서열번호 11)	TCGTA (서열번호 7)
18k-10 (서열번호 12)	TCTTC (서열번호 8)
18k-62 (서열번호 13)	GAGGT (서열번호 9)

[0074] 상기 선별된 변이체들의 프로모터-RBS-스페이서-p5 코딩 서열을 UTR designer (https://sbi.postech.ac.kr/utr_designer/)에 입력하여 RBS 결합 강도를 예측하였으며, 이를 통해 p5 단백질의 발현 수준을 통해 예측하였다 (표 2).

표 2

스페이서 서열	예측된 p5 발현 수준
18k-0	314958.9
18k-5	102908.7
18k-7	453043.1
18k-10	251821.9
18k-62	87012.5

[0076] 실시예 2. 단일가닥 DNA 생산량 측정

[0077] 상기 실시예 1로부터 도출된, p5의 발현량을 감소시킬 것으로 예상된 18k-5, 18k-7, 18k-10, 및 18k-62 파지를 F 섬모를 발현하는 숙주 세포인 XL1blue *endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB⁺ lacI^q Δ(lacZ)M15] hsdR17(r_K⁻m_K⁺)* 균주에 감염시키고, 배양 배지 (2XM9(Na₂HPO₄ 12g/L, KH₂PO₄ 6g/L, NaCl 1g/L, NH₄Cl 2g/L)+citrate 1.7g/L+thiamine 25mg/L) 25ml에서 27시간동안 배양하였다. 대조군으로서 야생형 파지 18k-0을 사용하였다. 파지는 PEG 침전법을 사용하여 수집하고 230nm-400nm에서 흡광도를 측정한 후 아래의 식을 이용하여 농도를 측정하였다.

[0078]
$$(A_{269}-A_{320}) \times (6 \times 10^{16}) / \text{number of bases/virion} = \text{PFU/ml}$$

[0079] 그 결과는 18k-0를 접종했을 때의 파지 생산량에 대한 각 파지 변이체를 접종했을 때의 파지 생산량의 비율로서 나타내었다 (표 3, 도 2).

표 3

파지	상대적 파지 생산량
18k-0(WT)	1
18k-5	2.432817
18k-7	2.376569
18k-10	2.270974
18k-62	2.129119

[0081] 또한, 단일가닥 DNA농도는 나노분광계를 이용하여 260nm에서 흡광도를 측정한 후 아래의 식을 이용하여 농도를 산출하였다

[0082]
$$A_{260} \times 33 = \text{ssDNA concentration (ng/ul)}$$

[0083] 그 결과는 하기 표 4 및 도 3에 나타내었다.

표 4

파지	ssDNA 농도 (260/280)	상대적 ssDNA 생산량
18k-0(WT)	380	1
18k-5	1330	3.5
18k-7	790	2.07895
18k-10	920	2.42105
18k-62	1138	2.99474

[0085] 실시예 3. p5의 생산량 측정

[0086] 상기 18k-0, 18k-5, 18k-7, 18k-10, 및 18k-62 파지를 F 섬모를 발현하는 숙주 세포인 Jude1 DH10B F' [proAB lacI^Q lacZ ΔM15 Tn10(Tet^R)] 균주에 감염시키고, 배양 배지(2XM9(Na₂HPO₄ 12g/L, KH₂PO₄ 6g/L, NaCl 1g/L, NH₄Cl 2g/L)+citrate 1.7g/L+thiamine 25mg/L+leucine 40mg/L) 25ml에서 27시간동안 배양하였다. 대조군으로서 야생형 파지 18k-0을 사용하였다. 생산된 파지의 양은 상기 실시예 2와 동일한 방법으로 측정하였으며, 그 결과를 18k-0에 대한 상대적 양으로 나타내었다 (표 5, 도 4).

표 5

파지	상대적 파지 생산량
18k-0	1
18k-5	1.114822
18k-7	1.212842
18k-10	1.207001
18k-62	1.36875

[0089] 서열 목록

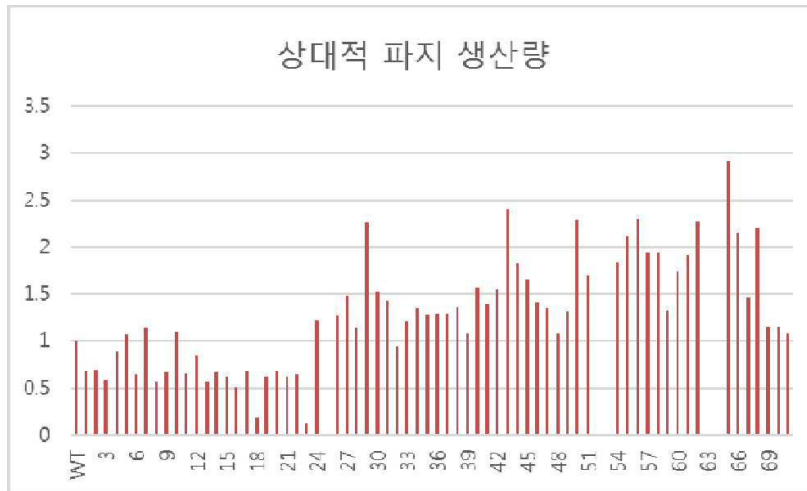
표 6

서열번호	설명
1	야생형 p5 단백질 발현 카세트 서열
2	야생형 p5 단백질 발현 카세트의 프로모터 서열
3	야생형 p5 단백질 발현 카세트의 RBS 서열
4	야생형 p5 단백질 발현 카세트의 p5 코딩 영역 서열
5	18k-0의 스페이서 서열
6	18k-5의 스페이서 서열
7	18k-7의 스페이서 서열
8	18k-10의 스페이서 서열
9	18k-62의 스페이서 서열
10	18k-5의 p5 단백질 발현 카세트 서열
11	18k-7의 p5 단백질 발현 카세트 서열
12	18k-10의 p5 단백질 발현 카세트 서열
13	18k-62의 p5 단백질 발현 카세트 서열
14	M13mp18pV TCACA random mut. 1F 서열
15	M13mp18pV TCACA random mut. 1R 서열
16	M13mp18pV TCACA random mut. 2F 서열
17	M13mp18pV TCACA random mut. 2R 서열
18	M13mp18k 전장 서열

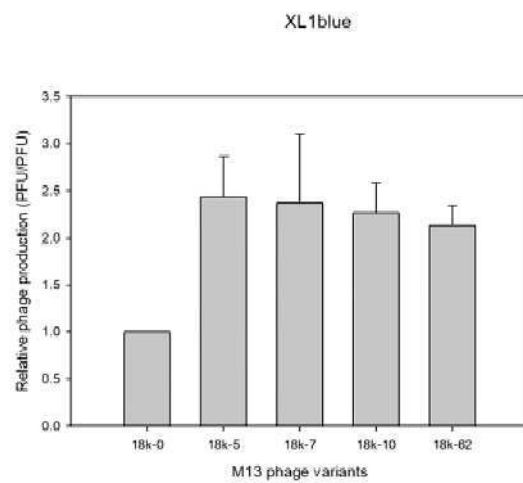
[0091] 예를 들어, 청구항 구성 목적을 위해, 이하 기재되는 청구항은 어떤 식으로든 이의 문자 그대로의 언어보다 좁게 해석되어서는 안 되고, 따라서 명세서로부터의 예시적 구현예가 청구항으로 읽혀서는 안 된다. 따라서, 본 발명은 예시로서 기재되었고, 청구항의 범위에 대한 제한이 아님이 이해되어야 한다. 따라서, 본 발명은 하기 청구항에 의해서만 제한된다. 본 출원에 인용된 모든 간행물, 발행된 특허, 특허 출원, 서적 및 저널 논문은 이들의 전체내용이 참조로서 본원에 각각 포함된다.

도면

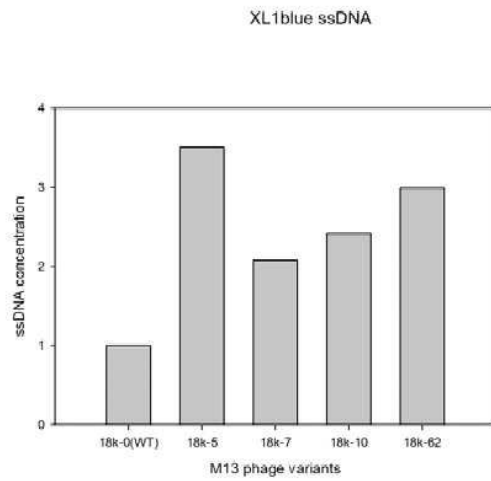
도면1



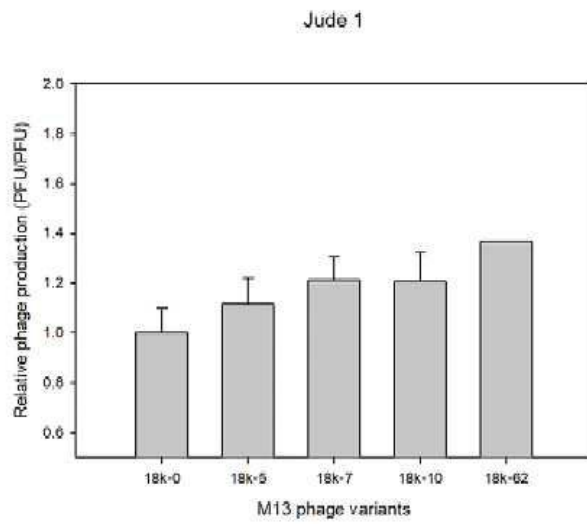
도면2



도면3



도면4



도면5

p5 발현 카세트_18k-0

```
TAACGTAGATTTTCTTCCCAACGTCCTGACTGGTATAATGAGCCAGTTCTTAAAATCGCATAAGGTAATCACA
ATGATTAAAGTTGAAATTAAACCATCTCAAGCCCAATTTACTACTCGTTCTGGTGTTTCTCGTCAGGGCAAGCCT
TATTCAGTGAATGAGCAGCTTTGTTACGTTGATTTGGGTAATGAATATCCGGTCTTGTCAAGATTACTCTTGATG
AAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACCGTTCATCTGTCCTCTTCAAAGTTGGTCAGTTCGGTTCC
CTTATGATTGACCGTCTGCGCCTCGTTCGGGCTAAGTAA
```

Promoter region

RBS

Spacer

P5 Coding region

도면6

p5 발현 카세트_18k-5

```
TAACGTAGATTTTCTTCCCAACGTCCTGACTGGTATAATGAGCCAGTTCTTAAAATCGCATAAGGTAATCTGAC
ATGATTAAAGTTGAAATTAAACCATCTCAAGCCCAATTTACTACTCGTTCTGGTGTTTCTCGTCAGGGCAAGCCT
TATTCAGTGAATGAGCAGCTTTGTTACGTTGATTTGGGTAATGAATATCCGGTCTTGTCAAGATTACTCTTGATG
AAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACCGTTCATCTGTCCTCTTCAAAGTTGGTCAGTTCGGTTCC
CTTATGATTGACCGTCTGCGCCTCGTTCGGGCTAAGTAA
```

Promoter region

RBS

Spacer

P5 Coding region

도면7

p5 발현 카세트_18k-7

TAACGTAGATTTTCTTCCCAACGTCCTGACTGGTATAATGAGCCAGTCTTAAAATCGCATAAGGTAATTCGTA
ATGATTAAAGTTGAAATTAAACCATCTCAAGCCCAATTACTACTCGTTCGGTGTTCCTCGTCAGGGCAAGCCT
TATTCAGTGAATGAGCAGCTTTGTTACGTTGATTGGGTAATGAATATCCGGTCTTGTCAGATTACTCTTGATG
AAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACCGTTCATCTGTCCTCTTCAAAGTTGGTCAGTTCGGTTCC
CTTATGATTGACCGTCTGCGCCTCGTTCGGGCTAAGTAA

Promoter region

RBS

Spacer

P5 Coding region

도면8

p5 발현 카세트_18k-10

TAACGTAGATTTTCTTCCCAACGTCCTGACTGGTATAATGAGCCAGTCTTAAAATCGCATAAGGTAATCTTTC
ATGATTAAAGTTGAAATTAAACCATCTCAAGCCCAATTACTACTCGTTCGGTGTTCCTCGTCAGGGCAAGCCT
TATTCAGTGAATGAGCAGCTTTGTTACGTTGATTGGGTAATGAATATCCGGTCTTGTCAGATTACTCTTGATG
AAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACCGTTCATCTGTCCTCTTCAAAGTTGGTCAGTTCGGTTCC
CTTATGATTGACCGTCTGCGCCTCGTTCGGGCTAAGTAA

Promoter region

RBS

Spacer

P5 Coding region

도면9

p5 발현 카세트_18k-62

TAACGTAGATTTTCTTCCCAACGTCCTGACTGGTATAATGAGCCAGTTC TTAAAATCGCATAAGGTAATGAGGT
 ATGATTAAAGTTGAAATTAAACCATCTCAAGCCCAATTTACTACTCGTTCTGGTGTCTCTCGTCAGGGCAAGCCT
 TATTCAGTGAATGAGCAGCTTTGTTACGTTGATTGGGTAATGAATATCCGGTCTTGTCAAGATTACTCTTGATG
 AAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACCGTTCATCTGTCTCTTCAAAGTTGGTCAGTTCGGTTCC
 CTTATGATTGACCGTCTGCGCCTCGTTCGGGCTAAGTAA

Promoter region

RBS

Spacer

P5 Coding region

도면 10a

M13mp18k

[illegible]

도면11

파지	스페이서 서열
18k-0	TCACA
18k-1	GTCTT
18k-2	TATTA
18k-3	CTCTA
18k-4	TGGAT
18k-5	CTGAC
18k-6	CATTC
18k-7	TCGTA
18k-8	GTTCT
18k-9	GCATT
18k-10	TCTTC
18k-11	CGGT
18k-12	CGGC
18k-13	TTATT
18k-14	GCAGT
18k-15	TATTT
18k-16	ACACA
18k-17	GATTG
18k-18	AAGTC
18k-19	CGTAA
18k-20	CTGGC
18k-21	ATGGA
18k-22	CAGGC
18k-23	CGGGC
18k-24	TTGTA
18k-25	TCGTT
18k-26	CCAAC
18k-27	TGTAT
18k-28	GAGTA
18k-29	AGTAT
18k-30	TTGTC
18k-31	GTTTA
18k-32	CCGGC
18k-33	ATCGA
18k-34	GTTGT
18k-35	TTTTA
18k-36	TGTCT

파지	스페이서 서열
18k-37	TTTAA
18k-38	GTGTC
18k-39	AAGGA
18k-40	AAGTA
18k-41	TGCTA
18k-42	ATGGT
18k-43	CGTTA
18k-44	GTGAA
18k-45	GAAAT
18k-46	AGGTT
18k-47	GTGTA
18k-48	GACTC
18k-49	ATTCG
18k-50	TAATT
18k-51	CATGA
18k-52	TCTTA
18k-53	TTTTT
18k-54	GCTAA
18k-55	GTTTT
18k-56	ATGCT
18k-57	CAGTT
18k-58	TACTA
18k-59	AGCCA
18k-60	CCCTC
18k-61	CACTA
18k-62	GAGGT
18k-63	AACAA
18k-64	TTAGG
18k-65	GAGTT
18k-66	CGATC
18k-67	CACAG
18k-68	TTTIG
18k-69	TTTGA
18k-70	CGGTA
18k-71	GAITT

서열 목록

- <110> Korea University Research & Business Foundation
- <120> M13 mutant for increasing the production of DNA
- <130> 19P05040
- <160> 18
- <170> KoPatent In 3.0
- <210> 1
- <211> 339
- <212> DNA

<213> Artificial Sequence

<220><223> M13mp18k

<220><221> gene

<222> (1)..(339)

<223> wild type p5-expressing cassette

<400> 1

taacgtagat ttttcttccc aacgtcctga ctggtataat gagccagttc ttaaaatcgc 60

ataaggaat tcacaatgat taaagttgaa attaaacat ctcaagccca atttactact 120

cgttctgggtg ttctctgtca gggcaagcct tattcactga atgagcagct ttgttacgtt 180

gatttgggta atgaatatcc ggttcttgtc aagattactc ttgatgaagg tcagccagcc 240

tatgcgcctg gtctgtacac cgttcatctg tcctctttca aagttgggtca gttcgggttc 300

cttatgattg accgtctgcg cctcgttccg gctaagtaa 339

<210> 2

<211> 50

<212> DNA

<213> Artificial Sequence

<220><223> M13mp18k

<220><221> promoter

<222> (1)..(50)

<223> promoter of wild type p5-expressing cassette

<400> 2

taacgtagat ttttcttccc aacgtcctga ctggtataat gagccagttc 50

<210> 3

<211> 6

<212> DNA

<213> Artificial Sequence

<220><223> M13mp18k

<220><221> RBS

<222> (1)..(6)

<223> RBS of wild type p5-expressing cassette

<400> 3

ggtaat 6

<210> 4

<211> 264
 <212> DNA
 <213> Artificial Sequence
 <220><223> M13mp18k
 <220><221> gene
 <222> (1)..(264)

 <223> p5 coding sequence
 <400> 4
 atgattaaag ttgaaattaa accatctcaa gcccaattta ctactcgttc tgggtgttct 60
 cgtcaggga agccttattc actgaatgag cagctttgtt acgttgattt gggtaatgaa 120
 tatccggttc ttgtcaagat tactcttgat gaaggtcagc cagcctatgc gcctggtctg 180
 tacaccgttc atctgtcctc tttcaaagtt ggtcagttcg gttcccttat gattgaccgt 240
 ctgcgcctcg ttccggctaa gtaa 264
 <210> 5
 <211> 5
 <212> DNA
 <213> Artificial Sequence
 <220>
 ><223> M13mp18k
 <220><221> gene
 <222> (1)..(5)
 <223> spacer sequence of 18k-0 p5-expressing cassette
 <400> 5
 tcaca 5
 <210> 6
 <211> 5
 <212> DNA
 <213> Artificial Sequence
 <220><223> M13mp18k
 <220><221> gene
 <222> (1)..(5)
 <223> spacer sequence of 18k-5 p5-expressing cassette
 <400> 6
 ctgac 5

<210> 7

<211> 5

<212> DNA

<213> Artificial Sequence

<220><223> M13mp18k

<220><221> gene

<222> (1)..(5)

<223> spacer sequence of 18k-7 p5-expressing cassette

<400> 7

tcgta

5

<210> 8

<211> 5

<212> DNA

<213> Artificial Sequence

<220><223> M13mp18k

<220><221> gene

<222> (1)..(5)

<223> spacer sequence of 18k-10 p5-expressing cassette

<400> 8

tcttc

5

<210> 9

<211> 5

<212> DNA

<213> Artificial Sequence

<220><223> M13mp18k

<220><221> gene

<222> (1)..(5)

<223> spacer sequence of 18k-62 p5-expressing cassette

<400> 9

gaggt

5

<210> 10

<211> 339

<212> DNA

<213> Artificial Sequence

<220><223> M13mp18k

<220><221> gene

<222> (1)..(339)

<223> 18k-5 p5-expressing cassette

<400> 10

taacgtagat ttttcttccc aacgtcctga ctggtataat gagccagttc ttaaaatcgc 60

ataaggtaat ctgacatgat taaagttgaa attaaacccat ctcaagccca atttactact 120

cgttctgggtg tttctcgtca gggcaagcct tattcactga atgagcagct ttgttacgtt 180

gatttgggta atgaatatcc ggttcttctg aagattactc ttgatgaagg tcagccagcc 240

tatgcgcctg gtctgtacac cgttcatctg tcctctttca aagtttgtca gttcggttcc 300

cttatgattg accgtctgcg cctcgttccg gctaagtaa 339

<210> 11

<211> 339

<212> DNA

<213> Artificial Sequence

<220><223> M13mp18k

<220><221> gene

<222> (1)..(339)

<223> 18k-7 p5-expressing cassette

<400> 11

taacgtagat ttttcttccc aacgtcctga ctggtataat gagccagttc ttaaaatcgc 60

ataaggtaat tcgtaatgat taaagttgaa attaaacccat ctcaagccca atttactact 120

cgttctgggtg tttctcgtca gggcaagcct tattcactga atgagcagct ttgttacgtt 180

gatttgggta atgaatatcc ggttcttctg aagattactc ttgatgaagg tcagccagcc 240

tatgcgcctg gtctgtacac cgttcatctg tcctctttca aagtttgtca gttcggttcc 300

cttatgattg accgtctgcg cctcgttccg gctaagtaa 339

<210> 12

<211> 339

<212> DNA

<213> Artificial Sequence

<220><223> M13mp18k

<220><221> gene

<222> (1)..(339)

<223> 18k-10 p5-expressing cassette

<400> 12

taacgtagat ttttcttccc aacgtcctga ctggtataat gagccagttc ttaaaatcgc 60

ataaggtaat tcttcatgat taaagttgaa attaaacat ctcaagccca atttactact 120

cgttctggtg tttctcgtca gggcaagcct ttttactga atgagcagct ttgttacgtt 180

gatttgggta atgaatatcc ggttcttctc aagattactc ttgatgaagg tcagccagcc 240

tatgcgcctg gtctgtacac cggtcatctg tctcttttca aagttggtca gttcgggttc 300

cttatgattg accgtctcgc cctcgttccg gctaagtaa 339

<210> 13

<211> 339

<212> DNA

<213> Artificial Sequence

<220><223> M13mp18k

<220><221> gene

<222> (1)..(339)

<223> 18k-62 p5-expressing cassette

<400> 13

taacgtagat ttttcttccc aacgtcctga ctggtataat gagccagttc ttaaaatcgc 60

ataaggtaat gaggtatgat taaagttgaa attaaacat ctcaagccca atttactact 120

cgttctggtg tttctcgtca gggcaagcct ttttactga atgagcagct ttgttacgtt 180

gatttgggta atgaatatcc ggttcttctc aagattactc ttgatgaagg tcagccagcc 240

tatgcgcctg gtctgtacac cggtcatctg tctcttttca aagttggtca gttcgggttc 300

cttatgattg accgtctcgc cctcgttccg gctaagtaa 339

<210> 14

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> M13mp18pV TCACA random mut. 1F primer

<400> 14

ttaaaatcgc ataaggtaat nnnnatgat taaagttgaa attaaacat c 51

<210>	15	
<211>	36	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M13mp18pV TCACA random mut. 1R primer	
<400>	15	
	attttgacgc tcaatcgctc gaaatggatt atttac	36
<210>	16	
<211>	31	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M13mp18pV TCACA random mut. 2F primer	
<400>	16	
	agacgattga gcgtcaaat gtaggtattt c	31
<210>	17	
<211>	24	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M13mp18pV TCACA random mut. 2R primer	
<400>	17	
	attaccttat gcgattttaa gaac	24
<210>	18	
<211>	8582	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M13mp18k full sequence	
<400>	18	
	gtaccggggg atcctctaga gtcgacctgc aggcattgcaa gcttggcact ggccgtcggt	60
	ttacaacgtc gtgactggga aaacctggc gttaccaac ttaatcgct tgcagacat	120
	ccccctttcg ccagctggcg taatagcgaa gaggcccgca ccgatcgccc ttccaacag	180
	ttgcgcagcc tgaatggcga atggcgcttt gcttggttc cggcaccaga agcggtgccg	240
	gaaagctggc tggagtgcga tcttctgag gccgatactg tcgtcgtccc ctcaactgg	300
	cagatgcacg gttacgatgc gccatctac accaacgtga cctatcccat tacgggtcaat	360

ccgccgtttg ttcccacgga gaatccgacg ggttggttact cgctcacatt taatgttgat 420
gaaagctggc tacaggaagg ccagacgcga attatTTTTg atggcggtcc tattggttaa 480
aaaatgagct gatttaacaa aaatttaatg cgaattttaa caaaatatta acgtttacaa 540
tttaaatatt tgcttataca atcttcctgt ttttggggct tttctgatta tcaaccgggg 600

tacatatgat tgacatgcta gttttacgat taccgttcat cgattctctt gtttgctcca 660
gactctcagg caatgacctg atagcctttg tagatctctc aaaaatagct accctctccg 720
gcattaattt atcagctaga acggttgaat atcataatga tggtgatttg actgtctccg 780
gcctttctca ccttttgaa tctttaccta cacattactc aggcattgca tttaaaatat 840
atgagggttc taaaaatttt tatccttgcg ttgaaataaa ggcttctccc gcaaaagtat 900
tacagggtca taatgttttt ggtacaaccg atttagcttt atgctctgag gctttattgc 960
ttaattttgc taattctttg ccttgcctgt atgatttatt ggatgttaat gctactacta 1020

ttagtagaat tgatgccacc ttttcagctc gcgccccaaa tgaaaatata gctaaacagg 1080
ttattgacca ttgcgaaat gtatctaata gtcaaaactaa atctactcgt tcgcagaatt 1140
gggaatcaac tgttatatgg aatgaaactt ccagacaccg tactttagtt gcatatttaa 1200
aacatgttga gctacagcat tataatcagc aattaagctc taagccatcc gcaaaaatga 1260
cctcttatca aaaggagcaa ttaaaggtag tctctaatec tgacctgttg gagtttgctt 1320
ccggtctggg tcgctttgaa gctcgaatta aaacgcgata tttgaagtct ttcgggcttc 1380
ctcttaatct ttttgatgca atccgctttg ctcttgacta taatagtcag ggtaaagacc 1440

tgatttttga tttatggta tttctgtttt ctgaactgtt taaagcattt gagggggatt 1500
caatgaatat ttatgacgat tccgcagtat tggacgtat ccagtcctaa cattttacta 1560
ttacccctc tgcaaaaact tcttttgcaa aagcctctcg ctattttggt ttttatcgtc 1620
gtctggtaaa cgagggttat gatagtgttg ctcttactat gcctcgtaat tccttttggc 1680
gttatgtatc tgcatagtt gaatgtggtt ttctaaatc tcaactgatg aatctttcta 1740
cctgtaataa tgttggtccg ttagttcgtt ttattaacgt agatttttct tcccaacgtc 1800
ctgactggta taatgagcca gttcttaaaa tcgcataagg taattcacia tgattaaagt 1860

tgaaattaaa ccatctcaag cccaatttac tactcgttct ggtgtttctc gtcagggcaa 1920
gccttattca ctgaatgagc agctttgtta cgttgatttg ggtaatgaat atccggttct 1980
tgtcaagatt acitctgatg aaggtcagcc agcctatgcg cctggctctgt acaccgttca 2040
tctgtcctct ttcaaagtg gtcagttcgg ttcccttatg attgaccgtc tgcgcctcgt 2100
tccggctaag taacatggag caggtcgagg atttcgacac aatttatcag gcgatgatac 2160
aaatctccgt tgiactttgt ttcgcgcttg gtataatcgc tgggggtcaa agatgagtgt 2220

tttagtgat tcttttgcct ctttcgtttt aggttgggtgc cttcgtatg gcattacgta	2280
ttttaccgt ttaatggaaa cttcctcatg aaaaagtctt tagtcctcaa agcctctgta	2340
gccgttgcta cctcgttcc gatgctgtct ttcgtgctg agggtagcga tcccgcacaaa	2400
gcggccttta acicccctgca agcctcagcg accgaatata tcggttatgc gtgggcgatg	2460
gttgttgtca ttgtcggcgc aactatcggg atcaagctgt ttaagaaatt cacctcgaaa	2520
gcaagctgat aaaccgatac aattaaaggc tccttttggg gccttttttt tggagatfff	2580
caacgtgaaa aaattattat tcgcaattcc tttagttgtt ctttctatt ctcactccgc	2640
tgaaactgtt gaaagtgtt tagcaaaatc ccatacagaa aattcattta ctaacgtctg	2700
gaaagacgac aaaactttag atcggtacgc taactatgag ggctgtctgt ggaatgctac	2760
aggcgttgta gtttgtactg gtgacgaaac tcagtgttac ggtacatggg ttcctattgg	2820
gcttgcctatc cctgaaaatg aggggtgggtg cctgagggt ggcggttctg aggggtggcg	2880
ttctgagggt ggcggtacta aacctcctga gtacgggat acacctattc cgggctatac	2940
ttatatcaac cctctcgacg gcacttatcc gcctgggtact gagcaaaacc ccgctaacc	3000
taatccttct cttgaggagt ctcagcctct taatacttcc atgtttcaga ataataggtt	3060
ccgaaatagg cagggggcat taactgttta tacgggcact gttactcaag gcactgaccc	3120
cgttaaaact tattaccagt acactcctgt atcatcaaaa gccatgtatg acgcttactg	3180
gaacggtaaa ttacagact gcgctttcca ttctggcttt aatgaggatt tatttgtttg	3240
tgaatatcaa ggccaatcgt ctgacctgcc tcaacctcct gtcaatgctg gcggcggtc	3300
tggtaggtgt tctggtggcg gctctgaggg tggtaggtct gagggtagcg gtctgaggg	3360
tggcggtctt gagggaggcg gtccgggtgg tggctctggt tccggtgatt ttgattatga	3420
aaagatggca aacgctaata agggggctat gaccgaaaat gccgatgaaa acgcgtaca	3480
gtctgacgt aaaggcaaac ttgattctgt cgtactgat tacggtgctg ctatcgatgg	3540
tttcattggt gacgtttccg gccttgctaa tggtaatggt gctactggtg attttgctgg	3600
ctctaattcc caaatggctc aagtcggtga cggtagaat tcacctttaa tgaataattt	3660
ccgtcaatat ttaccttccc tccctcaatc ggttgaatgt cgcccttttg tctttggcgc	3720
tggtaaacca tatgaatttt ctattgattg tgacaaaata aacttattcc gtggtgtctt	3780
tgcgtttctt ttatatgttg ccacctttat gtagtatatt tctacgtttg ctaacatact	3840
gcgtaataag gagtcttaat catgccagtt cttttgggta ttccgttatt attgcgtttc	3900
ctcggtttcc ttctggtaac ttgttcggc tatctgctta cttttcttaa aaagggttc	3960

ggtaagatag ctattgctat ttcatgtttt ctgtctctta ttattgggct taactcaatt	4020
cttgtgggtt atctctctga tattagcgct caattaccct ctgactttgt tcagggtgtt	4080
cagttaattc tcccgctctaa tgcgcttccc tgtttttatg ttattctctc tgtaaaggct	4140
gctattttca tttttgacgt taaacaaaaa atcgtttctt atttgattg ggataaataa	4200
tatggctgtt tattttgtaa ctggcaaatt aggcctctga aagacgctcg ttagcgttgg	4260
taagattcag gataaaattg tagctgggtg caaaatagca actaatcttg atttaaggct	4320
tcaaacctc ccgcaagtcg ggaggttcgc taaaacgcct cgcgttctta gaataccgga	4380
taagccttct atatctgatt tgcctgctat tgggcgcggt aatgattcct acgatgaaaa	4440
taaaaacggc ttgcttgctc tcgatgagtg cggctacttg tttaataccc gttcttgga	4500
tgataaggaa agacagccga ttattgattg gtttctacat gctcgtaaat taggatggga	4560
tattattttt ctgttcagg acttatctat tgttgataaa caggcgcgtt ctgcattagc	4620
tgaacatgtt gtttattgtc gtcgtctgga cagaattact ttacctttg tcggtacttt	4680
atattctctt attactggct cgaaaatgcc tctgcctaaa ttacatgttg gcgttgtaa	4740
atatggcgat tctcaattaa gccctactgt tgagcgttgg ctttatactg gtaagaattt	4800
glataacgca taigatacta aacaggcttt ttctagtaat tatgattccg gtgtttattc	4860
ttatttaacg cttattttat cacacggctc gtatttcaaa ccattaaatt taggtcagaa	4920
gatgaaatta actaaaatat atttgaaaaa gttttctcgc gttctttgtc ttgcgattgg	4980
atttgcatca gcatttacct atagtatat aacccaacct aagccggagg ttaaaaagg	5040
agtctctcag acctatgatt ttgataaatt cactattgac tcttctcagc gtcttaactc	5100
aagctatcgc tatgttttca aggattctaa gggaaaatta attaatagcg acgatttaca	5160
gaagcaaggt tattcactca catatattga tttatgtact gtttcatta aaaaaggtaa	5220
ttcaaatgaa attgttaaat gtaattaatt ttgttttctt gatgtttgtt tcatcatctt	5280
cttttgctca ggtaattgaa atgaataatt cgcctctcgc cgattttgta acttggtatt	5340
caaagcaatc aggcgaatcc gttattgttt ctcccgatgt aaaagggtact gttactgtat	5400
attcatctga cgtaaacct gaaaatctac gcaatttctt tatttctgtt ttacgtgcaa	5460
ataattttga tatggtaggt tctaaccctt ccattattca gaagtataat ccaaacaatc	5520
aggattatat tgatgaattg ccattcatctg ataatacagga atatgatgat aattccgctc	5580
cttctgggtg tttctttgtt ccgcaaaatg ataatgttac tcaaactttt aaaattaata	5640
acgttcgggc aaaggattta atacgagttg tcgaattgtt tgtaaagtct aatacttcta	5700
aatcctcaaa tgtattatct attgacggct ctaatctatt agttgttagt gctcctaaag	5760
atattttaga taaccttctt caattccttt caactgttga ttgccaact gaccagatat	5820

tgattgaggg ttigatatit gaggttcagc aaggtgatgc tttagatttt tcatttgctg	5880
ctggctctca gcgtggcact gttgcaggcg gtgttaatac tgaccgcctc acctctgttt	5940
tatcttctgc tgggtggttcg ttccgtattt ttaatggcga tgttttaggg ctatcagttc	6000
gcgcattaaa gactaatagc cattcaaaaa tatgtctgt gccacgtatt cttacgcttt	6060
caggtcagaa gggttctatc tctgttggcc agaattgtccc ttttattact ggtcgtgtga	6120
ctggtgaatc tgccaatgta aataatccat ttcagacgat tgagcgtcaa aatgtaggta	6180
tttccatgag cgtttttctt gttgcaatgg ctggcggtaa tattgttctg gatattacca	6240
gcaagccga tagtttgagt tcttctactc aggcaagtga tgttattact aatcaaagaa	6300
gtattgctac aacggttaat ttgcgtgatg gacagactct tttactcggt ggcctcactg	6360
attataaaaa cacttctcag gattctggcg taccgttctt gtctaaaatc cctttaatcg	6420
gcctcctgtt tagctcccg cctgattcta acgaggaaag cacgttatac gtgctcgta	6480
aagcaacat agtacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg	6540
cgcagcgtga ccgtacact tgccagcgcc ctagcggccg ctcctttcgc tttcttccct	6600
tcctttctcg ccacgttcgc cggctttccc cgtcaagctc taaatcgggg gctcccttta	6660
gggttccgat ttagtgcttt acggcacctc gaccccaaaa aacttgattt gggatgatgt	6720
tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg	6780
ttctttaata gtggactctt gttccaaact ggaacaacac tcaacctat ctcgggctat	6840
tcttttgatt tataagggat ttgtccgatt tcggaaccac catcaaacag gattttcgcc	6900
tgctggggca aaccagcgtg gaccgcttgc tgcaactctc tcagggccag gcggtgaagg	6960
gcaatcagct gtigcccgtc tcactgggtga aaagaaaaac caccctggcg cccaatcgc	7020
aaaccgcctc tccccgcgcg ttggccgatt cattaatgca gctggcacga caggtttccc	7080
gactggaaag cgggcagtga gcgcaacgca attaatgtga gttagctcac tcattagga	7140
ccccaggctt tacatttat gcttcggct cgtatgttgt gtggaattgt gagcggataa	7200
caatttcaca caggaaacag ctatgacat gattacgaag atcccctcac gctgccgcaa	7260
gcactcaggg cgcaagggt gctaaaggaa gcggaacacg tagaaagcca gtccgcagaa	7320
acggtgctga ccccgaatga atgtcagcta ctgggctatc tggacaaggg aaaacgcaag	7380
cgcaaagaga aagcaggtag cttgcagtgg gcttacatgg cgatagctag actggcggt	7440
tttatggaca gcaagcgaac cggaattgcc agctggggcg cctcttggtt aggttgggaa	7500
gccctgcaaa gtaaaactgga tggctttctt gccccaagg atctgatggc gcaggggatc	7560
aagatctgat caagagacag gatgaggatc gtttcgatg attgaacaag atggattgca	7620
cgcaggttct ccggccgctt ggggtggagag gctattcggc tatgactggg cacaacagac	7680

aatcggctgc tctgatgccg ccgtgttccg gctgtcagcg caggggcgcc cggttctttt	7740
tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag gacgaggcag cgcggctatc	7800
gtggctggcc acgacgggcg ttccttgccg agctgtgctc gacgttgtca ctgaagcggg	7860
aagggaactgg ctgctattgg gcgaagtgcc ggggcaggat ctctgtcat ctacattgc	7920
tcctgccgag aaagtatcca tcatggctga tgcaatgcgg cggctgcata cgttgatcc	7980
ggctacctgc ccattcgacc accaagcgaa acatcgcatc gagcgagcac gtactcgat	8040
ggaagccggt ctgtcgatc aggatgatct ggacgaagag catcaggggc tcgcgccagc	8100
cgaactgttc gccaggctca aggcgcgcat gcccgacggc gaggatctcg tcgtgacca	8160
tggcgatgcc tgcttgccga atatcatggt ggaaaatggc cgcttttctg gattcatcga	8220
ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata gcgttgcta cccgtgatat	8280
tgctgaagag cttggcggcg aatgggctga ccgcttcctc gtgctttacg gtatgccgc	8340
tcccgattcg cagcgcatcg cttctatcg cttcttgac gagttcttct gagcgggact	8400
ctggggttcg aaatgaccga ccaagcgacg cccaacctgc catcacgaga ttctgattcc	8460
accgccgct tctatgaaag gttgggcttc ggaatcgttt tccgggacgc cggctggatg	8520
atcctccagc gcggggatct catgctggag ttcttcgccc accccagctt caaaagcgt	8580
ct	8582