(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0035084
(43) 공개일자 2016년03월30일

(51) 국제특허분류(Int. Cl.)
C12N 15/85 (2006.01) *C07K 16/32* (2006.01)
C07K 16/46 (2006.01) *C12P 21/02* (2006.01)
(52) CPC특허분류
C12N 15/85 (2013.01)
C07K 16/32 (2013.01)
(21) 출원번호 10-2016-7006036
(22) 출원일자(국제) 2014년08월05일
심사청구일자 없음
(85) 번역문제출일자 2016년03월07일
(86) 국제출원번호 PCT/EP2014/066826
(87) 국제공개번호 WO 2015/018832
국제공개일자 2015년02월12일
(30) 우선권주장
13179375.4 2013년08월06일
유럽특허청(EPO)(EP)

(71) 출원인
그렌마크 파머수티칼스 에스. 아.
스위스, 2300 라 쇼드퐁, 슈멍 드 라 꼼베타 5
(72) 발명자
애비셔-구미, 크리스텔
스위스, 2300 라 쇼드퐁, 슈멍 드 라 꼼베타 5, 씨/오 그렌마크 파머수티칼스 에스. 아.
베르츠신저, 마틴
스위스, 2300 라 쇼드퐁, 슈멍 드 라 꼼베타 5, 씨/오 그렌마크 파머수티칼스 에스. 아.
모레티, 피에르
스위스, 2300 라 쇼드퐁, 슈멍 드 라 꼼베타 5, 씨/오 그렌마크 파머수티칼스 에스. 아.
(74) 대리인
이원희

전체 청구항 수 : 총 27 항

(54) 발명의 명칭 **진핵세포의 폴리펩티드 발현 방법 및 발현 구조체**

(57) 요 약

본 발명은 일반적으로 재조합 폴리펩티드 제조에서 제품 품질을 최적화하기 위해 및 증가된 발현을 얻기 위해 사용될 수 있는 발현 벡터 및 발현 구조체와 같은 발현 시스템에 관한 것이다. 본 발명에서 설명된 바와 같은 발현 구조체를 사용하여, 높은 과도 및 안정한 역가를 얻을 수 있고, 과도 발현에서 과거, 종래 기술 연구에서 관찰된 과도 역가와 비교시 60배 까지 높은 것으로 나타났다.

(52) CPC특허분류
C07K 16/468 (2013.01)
C12P 21/02 (2013.01)
C07K 2317/14 (2013.01)
C07K 2317/52 (2013.01)
C07K 2317/622 (2013.01)
C07K 2317/64 (2013.01)
C12N 2840/20 (2013.01)
C12N 2840/445 (2013.01)

명 세 서

청구범위

청구항 1

5'에서 3' 방향으로 하기를 포함하는 발현 구조체:

프로모터;

첫번째 선택형 스플라이스 공여 자리;

첫번째 측면 인트론;

첫번째 스플라이스 수용 자리;

첫번째 폴리펩티드가 코딩된 첫번째 엑손;

두번재 선택형 스플라이스 공여 자리;

두번째 측면 인트론;

두번째 스플라이스 수용 자리; 및

두번째 폴리펩티드가 코딩된 두번째 엑손,

(여기서, 숙주 세포에 진입시, 상기 첫번째 엑손의 전사는 상기 첫번째 폴리펩티드의 발현을 야기하고 및/또는 상기 두번째 엑손의 전사는 상기 두번째 폴리펩티드의 발현을 야기한다).

청구항 2

제1항에 있어서,

상기 첫번째 및 두번째 측면 인트론은 cTNT(chicken troponin) 인트론 4, cTNT 인트론 5 및 사람의 EFI알파 유전자의 첫번째 인트론으로 이루어진 군으로부터 선택되는 것을 특징으로 하는 발현 구조체.

청구항 3

제1항 또는 제2항에 있어서,

상기 첫번째 및 두번째 측면 인트론은 최소 50개의 뉴클리오티드에 대하여 최소 80%의 핵산 서열 상동성을 가지는 것을 특징으로 하는 발현 구조체.

청구항 4

제1항 또는 제2항에 있어서,

상기 첫번째 및 두번째 측면 인트론은 최소 50개의 뉴클리오티드에 대하여 최소 95%의 핵산 서열 상동성을 가지는 것을 특징으로 하는 발현 구조체.

청구항 5

제1항 또는 제2항에 있어서,

상기 첫번째 및 두번째 측면 인트론은 최소 450개의 뉴클리오티드에 대하여 최소 95%의 핵산 서열 상동성을 가지는 것을 특징으로 하는 발현 구조체.

청구항 6

제1, 2 또는 3항 중 어느 한 항에 있어서,

최소 하나의 폴리피리미딘(poly(Y)) 관을 더 포함하는 것을 특징으로 하는 발현 구조체.

청구항 7

제6항에 있어서,

상기 poly(Y) 관은 상기 첫번째 엑손의 윗 줄기에 있는 것을 특징으로 하는 발현 구조체.

청구항 8

제6항에 있어서,

상기 poly(Y) 관은 상기 첫번째 엑손의 아래 줄기에 있는 것을 특징으로 하는 발현 구조체.

청구항 9

제6, 7 또는 8항 중 어느 한 항에 있어서,

상기 poly(Y) 관은 30보다 적은 수의 피리미딘 염기를 포함하는 것을 특징으로 하는 발현 구조체.

청구항 10

제6, 7 또는 8항 중 어느 한 항에 있어서,

상기 poly(Y) 관은 10 이하의 수로 피리미딘 염기를 포함하는 것을 특징으로 하는 발현 구조체.

청구항 11

이전 항 중 어느 한 항에 있어서,

상기 발현 구조체는 두번째 스플라이스 공여 자리가 없는 것을 특징으로 하는 발현 구조체.

청구항 12

이전 항 중 어느 한 항에 있어서,

상기 발현 구조체는 상기 프로모터의 아래 줄기에 세번째 스플라이스 공여 자리, 인트론 및 세번째 스플라이스 수용 자리를 더 포함하는 것을 특징으로 하는 발현 구조체.

청구항 13

제12항에 있어서,

상기 스플라이스 공여 자리, 인트론 및 스플라이스 수용 자리로 구성된 것을 특징으로 하는 발현 구조체.

청구항 14

제12항에 있어서,

상기 세번째 스플라이스 공여 자리는 5'UTR의 뒤에 및/또는 상기 세번째 수용 자리는 5'UTR의 앞에 오는 것을 특징으로 하는 발현 구조체.

청구항 15

이전 항 중 어느 한 항에 있어서,

상기 측면 인트론의 배열은 SEQ ID Nos: 129내지 175로 이루어진 군으로부터 선택되는 것을 특징으로 하는 발현 구조체.

청구항 16

이전 항 중 어느 한 항에 있어서,

상기 첫번째 폴리펩티드는 항체 중쇄 또는 이의 절편이고, 상기 두번째 폴리펩티드는 항체 경쇄 또는 이의 절편이거나, 상기 첫번째 폴리펩티드는 항체 경쇄 또는 이의 절편이고, 상기 두번째 폴리펩티드는 항체 중쇄 또는 이의 절편인 것을 특징으로 하는 발현 구조체.

청구항 17

이전 항 중 어느 한 항에 있어서,

상기 첫번째 폴리펩티드는 항체 중쇄이고 상기 두번째 폴리펩티드는 Fc-scFv이거나, 상기 첫번째 폴리펩티드는 Fc-scFv이고 상기 두번째 폴리펩티드는 항체 중쇄인 것을 특징으로 하는 발현 구조체.

청구항 18

이전 항 중 어느 한 항에 따른 발현 카세트(cassette)가 코딩된 폴리뉴클리오티드.

청구항 19

제18항의 하나 또는 그 이상의 폴리뉴클리오티드를 포함하는 복제 또는 발현 벡터.

청구항 20

제19항의 하나 또는 그 이상의 복제 또는 발현 벡터를 포함하는 숙주 세포.

청구항 21

제20항에 있어서,

항체 경쇄 또는 중쇄의 발현이 코딩된 폴리뉴클리오티드를 포함하는 제17항의 발현 구조체가 코딩된 폴리뉴클리오티드를 포함하는, 발현 벡터를 포함하는 숙주 세포.

청구항 22

제20항 또는 21항에 있어서,

상기 발현 벡터는 숙주세포로 안정적으로 감염된 것을 특징으로 하는 숙주 세포.

청구항 23

제20 내지 22항 중 어느 한 항에 있어서,

상기 숙주 세포는 포유류의 세포, 곤충 세포 및 효모 세포로 이루어진 군으로부터 선택되는 것을 특징으로 하는 숙주 세포.

청구항 24

제20 내지 23항 중 어느 한 항의 숙주 세포의 배양 및 상기 배양으로부터 발현된 폴리펩티드의 분리를 포함하는 폴리펩티드의 제조방법.

청구항 25

제21항의 숙주 세포의 배양 및 상기 배양으로부터 발현된 폴리펩티드의 분리를 포함하는 이중 특이성 항체의 제조방법.

청구항 26

하기를 포함하는, 제1항에 따른 하나 또는 그 이상의 발현 카세트가 코딩된 목적 단백질의 발현 수준을 최적화하는 방법:

(i) 최소 50개의 뉴클리오티드에 대하여 최소 80%의 핵산 서열 상동성을 가지는 첫번째 및 두번째 측면 인트론을 사용하는 단계(단계 i);

(ii) 첫번째 엑손의 윗줄기 폴리(Y) 관에서 피리미딘 염기 수를 감소시키거나 첫번째 엑손의 아래줄기 폴리(Y) 관에서 피리미딘 염기 수를 증가시키는 단계(단계 ii); 및/또는

(iii) 두번째 측면 인트론의 윗줄기 스플라이스 공여 자리를 삭제 시키는 단계(단계 iii).

청구항 27

하기를 포함하는, 제1항에 따른 하나 또는 그 이상의 발현 카세트가 코딩된 목적 단백질의 이형 이합체(heterodimerisation) 수준을 최적화하는 방법:

(i) 최소 50개의 뉴클리오티드에 대하여 최소 80%의 핵산 서열 상동성을 가지는 첫번째 및 두번째 측면 인트론을 사용하는 단계(단계 i);

(ii) 첫번째 엑손의 윗줄기 폴리(Y) 관에서 피리미딘 염기 수를 감소시키거나, 첫번째 엑손의 아래줄기 폴리(Y) 관에서 피리미딘 염기 수를 증가시키는 단계(단계 ii); 및/또는

(iii) 두번째 측면 인트론의 윗줄기 스플라이스 공여 자리를 삭제시키는 단계(단계 iii).

발명의 설명

기 술 분 야

[0001] 본 발명은 대체 접합을 사용하여 진핵세포의 폴리펩티드 및/또는 폴피펩티드 다량체 발현 방법 및 발현 구조체에 관한 것이다. 상기 구조체를 포함하는 숙주 세포의 제조방법, 효율적인 단백질 제조를 위한 상기 구조체 및 이로부터 발현된 폴리펩티드의 사용에 관한 것이다.

배 경 기 술

[0002] 진핵세포의 단백질을 제조하기 위해서, 상기 단백질의 DNA 코딩이 단백질로 차례로 번역되는 메신저 RNA(mRNA)에 전사되어야 한다. 상기 mRNA는 핵에서 인트론과 엑손을 함유하는, 전-mRNA로써 처음 전사된다. 전-mRNA에서 성숙 mRNA로의 성숙화 과정 동안, 상기 인트론은 이어 맞추기 복합체(spliceosome)라 불리는 단백질 장치에 의해 잘려("접합")진다. 상기 엑손은 함께 융합되고 상기 mRNA는 5' 말단에서의 소위 CAP 및 3' 말단에서의 폴리(A) 꼬리의 첨가에 의해 변형된다. 상기 성숙 mRNA는 세포질로 내보내지고, mRNA에 코딩된 단백질의 번역을 위한 주형의 역할을 한다.

[0003] 대체 접합은 상기 동일한 전-mRNA 전사가 다른 양상으로 접합되어 다른 성숙 mRNA들로 되는 것과 몇몇의 경우 다른 단배질로 되는 것을 설명한 용어이다. 상기 메카니즘은 자연에서 단백질의 발현 수준을 변화하거나 발달 동안의 특정 단백질의 활성을 변형하기 위해 사용된다(Cooper TA & Ordahl CP(1985), J Biol Chem, 260(20): 11140-8). 대체 접합은 대게 많은 인자들의 복합 상호작용에 의해 조절된다(Orengo JP et al., (2006) Nucleic Acids Res, 34(22):e148).

[0004] 비록 접합은 상기 문헌에서 잘 알려져 있고 컨센서스 서열은 사람 세포의 접합 용으로 공개되었지만, 대체 접합건들의 정확한 결과는 접합에 영향을 줄 수 있는 다양한 요인들로 인해 예상하기 쉽지 않다. 접합에 영향을 주는 요인들은 상기 컨센서스 서열의 분기 점, 접합 공여 및 접합 수용 부위, 엑손 및 인트론의 크기, 및 접합의 증가 또는 감소를 야기하는 조절 단백질의 결합 자리을 포함한다(Alberts B et al(2002) Molecular Biology of the Cell, 4th edition, New York: Garland Sience).

[0005] 대체 접합은 폴리펩티드, 특히, 다량 단백질, 예를들어 항체의 발현 수준을 증가 시키기위해 사용될 수 있다. 항체 발현의 수준은 경쇄 발현에 대한 중쇄의 비율에 의존적이다. 비록 다음 문헌에서는 중쇄 보다 더 많은 경쇄의 발현이 더 선호됨을 제시하고 있으나(Dorai H et al.,(2006) Hybridoma(Larchmt), 25(1): 1-9), 본 출원자들은 최대 발현을 야기하는 중쇄에 대한 경쇄의 최적화 비율이 크게 항체에 의존적임을 확인하였다. 이중 특이성 항체에서도 동일한 결과를 확인하였고, 본 발명자들은 상기 항체 발현 수준이 이중 특이성 항체를 형성하는 각기 다른 쇄의 비율에 의존적임을 보였다.

[0006] 대체 접합을 사용한 숙주 세포의 폴리펩티드 발현 방법은 종래 기술에 제시되어 있다. 예를들어, Prentice(WO200589285)는 대체 접합을 가능하게 하는 접합 자리를 가지는 발현 카세트에서 단일 프로모터의 조절 하에 둘 또는 그 이상의 발현 카세트를 포함하는 발현 벡터를 설명하고 있다. 상기 구조체에서, 폴리아데닐화(폴리(A)) 자리는 각각의 개방 판독 프레임 후에 포함된다. 유사하게, Fallot et al(WO2007135515)도 이후의 폴리펩티드 발현에 대한 둘 또는 그 이상의 mRNAf로 접합될 수 있는 전-mRNA의 전사를 유도하는 단일 프로모터를 사용한 숙주 세포에서 발현될 수 있는 발현 카세트를 설명하고 있다.

[0007] 대체 접합은 가장 높은 역가를 야기시키는 비율에 항체에 필요한 단위체를 발현시키기 위해서 사용될 수 있다. 예를들어, 중쇄 및 경쇄는 동일한 구조체에 복제될 수 있다. 접합은 상기 중쇄 또는 경쇄를 발현하는 mRNA의 특이 비율을 야기시킬 것이다. 상기 비율은 최종 항체의 발현을 위한 최적의 값에 가깝게 조정될 수 있다. 이중 특이성 분자의 제조에서 상기 비율은 발현 수준 뿐만 아니라, 제품 품질에 영향을 준다. 상기 최적의 비율은 목적 제조 종들의 최대 발현을 관찰함으로써 식별될 수 있다.

선행기술문헌

비특허문헌

[0008] (비특허문헌 0001) Orengo JP et al., (2006) Nucleic Acids Res, 34(22):e148

발명의 내용

해결하려는 과제

[0009] 본 발명은 일반적으로 재조합 폴리펩티드 제조에서 제품 품질을 최적화하기 위해 및 증가된 발현을 얻기 위해 사용될 수 있는 발현 벡터 및 발현 구조체와 같은 발현 시스템에 관한 것이다. 본 발명에서 설명된 바와 같은 발현 구조체를 사용하여, 높은 과도 및 안정한 역가를 얻을 수 있고, 과도 발현에서 과거, 종래 기술 연구에서

관찰된 과도 역가와 비교시 60배 까지 높은 것으로 나타났다.

과제의 해결 수단

[0010] 첫번째 측면에서, 본 발명은 폴리펩티드의 효율적인 발현을 위해 사용될 수 있는 발현 구조체에 관한 것이다. 바람직하게, 상기 발현 구조체는 5' 에서 3'의 방향으로 다음을 포함한다:

[0011] 프로모터;

[0012] 첫번째 선택형 스플라이스 공여 자리;

[0013] 첫번째 측면 인트론;

[0014] 첫번째 스플라이스 수용 자리;

[0015] 첫번째 폴리펩티드가 코딩된 첫번째 엑손;

[0016] 두번재 선택형 스플라이스 공여 자리;

[0017] 두번째 측면 인트론;

[0018] 두번째 스플라이스 수용 자리; 및

[0019] 두번째 폴리펩티드가 코딩된 두번째 엑손,

[0020] 여기서, 숙주 세포에 진입시, 상기 첫번째 엑손의 전사는 상기 첫번째 폴리펩티드의 발현을 야기하고 및/또는 상기 두번째 엑손의 전사는 상기 두번째 폴리펩티드의 발현을 야기한다.

[0021] 본 발명의 발명자들은 상기 첫번째 엑손 전 및 후에 측면 인트론 또는 이의 조각들의 사용과 이들이 서로 최소 80%의 핵산 서열 상동성을 공유한다는 것과, 폴리펩티드의 발현 수준에 상당한 영향을 준다는 것을 알아냈다. 본 발명의 한 실시예에서, 상기 첫번째 엑손 측면 인트론은 대체 접합된 자연적으로 발생하는 인트론, 및 구조적으로 접합된 인트론으로부터 유도될 수 있다. 바람직하게, 상기 인트론은 하기로 이루어진 군으로부터 선택될 수 있다: 치킨 트로포닌(cTNT) 인트론 4, cTNT 인트론 5 및 사람 EF1 알파 유전자의 인트론, 바람직하게 사람 EF1 알파 유전자의 첫번째 인트론. 보다 바람직하게, 상기 첫번째 엑손의 측면 인트론은 치킨 트로포닌 인트론 5(cTNT-14). 바람직하게, 상기 측면 인트론은 80%의 핵산 서열 상동성, 보다 바람직하게 90%의 핵산 서열 상동성, 가장 바람직하게 95%의 핵산 서열 상동성을 공유한다. 본 발명의 보다 바람직한 한 실시예에서, 상기 측면 인트론은 98%의 핵산 서열 상동성을 공유한다. 본 발명의 가장 바람직한 한 실시예에서, 상기 측면 인트론은 100%의 핵산 서열 상동성을 공유하고 동일한 핵산 서열을 갖는다. 상기 측면 인트론 서열들 간의 서열 상동성의 백분율은 폴리(Y) 관 서열을 배재한 핵산 구간을 비교함으로써 결정되는 것이다.

[0022] 바람직하게, 상기 측면 인트론은 길이에서 최소 50개의 뉴클리오티드의 핵산 구간에 대하여 상동성을 공유한다. 바람직하게 상기 측면 인트론은 길이에 최소 50 내지 100개의 뉴클리오티드, 바람직하게 길이에 최소 50 내지 150개의 뉴클리오티드, 바람직하게 길이에 최소 50 내지 200개의 뉴클리오티드, 바람직하게 길이에 최소 50 내지 250개의 뉴클리오티드, 보다 바람직하게 길이에 최소 50 내지 300개의 뉴클리오티드, 보다 바람직하게 길이에 최소 50 내지 350개의 뉴클리오티드, 한층 보다 바람직하게 길이에 최소 50 내지 400개의 뉴클리오티드 및 가장 바람직하게 길이에 최소 50 내지 450개의 뉴클리오티드의 핵산 구간을 따라 상동성을 공유한다. 본 발명의 한 실시예에서, 상기 측면 인트론의 최대 길이는 450 뉴클리오티드이다.

[0023] 본 발명의 한 양상에서, 상기 발현 구조체는 최소 하나의 폴리피리미딘(폴리(Y)) 관을 포함한다. 이것은 상기 분기점 및 스플리이스 수용체 간에 위치할 수 있고, 상기 첫번째 엑손의 윗줄기에 위치할 수 있다. 한 실시예로, 폴리(Y) 관에 피리미딘 염기의 수를 감소시키는 것은 상기 두번째 액손으로부터 두번째 폴리펩티드의 발현의 증가를 야기시킨다. 상기 폴리(Y) 관에 존재하는 피리미딘 염기의 수는 30 또는 그 이하일 수 있고, 바람직하게 20 또는 그 이하, 보다 바람직하게 10 또는 그 이하, 한층 더 바람직하게 7 또는 그 이하 및 가장 바람직하게 5 또는 그 이하일 수 있다. 선택적으로 상기 폴리(Y) 관은 상기 첫번째 엑손의 아래줄기에 위치할 수 있다.

[0024] 본 발명의 다른 양상에서, 상기 두번째 스플라이스 공여 자리는 제거된다. 바람직한 실시예에서, 상기 두번째 스플라이스 공여 자리의 제거는 상기 첫번째 엑손의 윗줄기 폴리(Y) 관에 피리미딘 염기의 수의 감소와 수반된다.

[0025] 본 발명의 또 다른 실시예에서, 상기 발현 구조체는 5'UTR, 세번째 스플라이스, 인트론, 세번째 스플라이스 수용 자리 및 추가 5'UTR을 더 포함할 수 있다. 바람직하게, 상기 스플라이스 공여 자리, 인트론 및 스플라이스 수용 자리는 상기 인트론이 성숙 mRNA에서 구조적으로 접합된 것과 같은 구성이다. 바람직하게 상기 구조적인 구성요소들은 상기 프로모터와 상기 첫번째 측면 인트론의 앞의 상기 스플라이스 공여 자리 사이에 위치한다.

[0026] 본 발명의 한 바람직한 실시예에서, 폴리아데닐화(폴리(A)) 자리는 상기 발현 구조체 안에 존재하지 않는다. 바람직하게 폴리(A) 자리는 상기 발현 구조체의 말단에 존재할 것이다.

[0027] 본 발명에서 생성된, 상기 분기점부터 다음 엑손의 시작으로 시작되는 측면 인트론 서열은 모두 독특한 인공 서열이다. 바람직하게, 상기 인공 서열은 SEQ ID Nos: 38 내지 128로 이루어진 군으로부터 선택된다. 보다 바람직하게, 상기 인공 서열은 상기 분기점부터 상기 첫번째 엑손의 시작까지로 시작되는 서열을 가지고 SEQ ID Nos: 129 내지 175로 이루어진 군으로부터 선택된다.

[0028] 본 발명의 한 양상에서, 상기 첫번째 및 두번째 엑손에 의해 코딩된 상기 폴리펩티드는 단백질 다량체 i.e 재결합 항체 또는 이의 조각들과 같은 이종 다량의 폴리펩티드. 상기 항체 조각은 하기로 이루어진 목록으로부터 선택된다: Fab, Fd, Fv, dAb, $F(ab')_2$ 및 scFv. 한 실시예로, 상기 발현 구조체에 의해 발현된 상기 첫번째 폴리펩티드는 항체 중쇄 또는 항체 경쇄 또는 이의 조각일 수 있다. 상기 발현된 첫번째 폴리펩티드는 항체 중쇄일 경우, 상기 발현 구조체에 의해 발현된 상기 두번째 폴리펩티드는 항체 경쇄이다. 선택적으로, 상기 발현된 첫번째 폴리펩티드는 항체 경쇄일 경우, 상기 두번째 폴리펩티드는 항체 중쇄이다.

[0029] 본 발명의 다른 양상에서, 상기 발현 구조체는 숙주 세포에서 이중 특이성 항체의 발현을 위해 사용될 수 있다. 한 실시예로, 상기 발현된 첫번째 폴리펩티드는 항체 중쇄이고 상기 발현된 두번째 폴리펩티드는 항체 Fc 부위에 연결된 항체의 조각이다. 상기 항체 조각은 하기로 이루어진 목록으로부터 선택될 것이다: Fab, Fd, Fv, dAb, $F(ab')_2$ 및 scFv. 바람직하게 상기 항체 조각은 Fab 또는 scFv이다. 보다 바람직하게 상기 항체 조각은 scFv이다.

[0030] 또한, 분리 발현 구조체는 숙주 세포에서 항체 경쇄의 발현을 위해 제공된다. 상기 숙주 세포에서 항체 경쇄를 코딩한 발현 구조체와 항체 조각-Fc 및 항체 중쇄를 코딩한 발현 구조체의 보조-발현은 상기 이중 특이성 항체의 발현의 결과를 야기한다. 본 발명의 보다 바람직한 실시예로 상기 항체 중쇄의 Fc 부위 및 상기 첫번째 및 두번째 폴리펩티드에의해 발현된 항체 조각에 연결된 Fc 부위는 상기의 Fc 부위의 상호작용이 강화된 것과 같은 변형을 포함한다.

[0031] 또한, 상기 Fc 부위에 대한 변형은 상기 이중 특이성 항체의 증가된 안정성의 결과를 야기한다.

도면의 간단한 설명

[0032] 도 1a: 본 발명의 대체 접합 구조체의 개략도이다. 상기 구조체는 네개의 엑손을 포함하고 있다. 상기 엑손 1 및 엑손 2는 첫번째 인트론(AS 인트론#1)에 의해 분리되고, 이것은 세포의 스플라이스 장치에 의해 구조적으로 잘려진 것이다. 엑손 3("대체 엑손"으로 언급된)도 포함되거나 잘려진 것이다. 이것은 dsRED를 코딩한 첫번째 개방 판독 프레임을 포함한다. 상기 엑손은 AS 인트론 #2에 의해 측면 윗줄기에 있고, 이것은(기본적인 구조체에서) 치킨 트로포닌 인트론 4(cTNT-I4)로부터 유도되고 (기본적인 구조체에서) 치킨 트로포닌 인트론 5(cTNT-I5)로부터 유도되는 AS 인트론 #3에 의해 아래줄기에 있다. 엑손 4는 구조적으로 mRNA에 포함된다. 비록 GFP를 코딩한 개방 판독 프레임이 성숙 mRNA의 첫번째 개방 판독 프레임이라 하더라도, GFP를 코딩한 상기 개방 판독 프레임만이 발현된다. 따라서, 대체 엑손 3이 구조체에 포함된다면, 오직 엑손 3에 코딩된 dsRED가 번역될 것이다(도의 상단에 위치). 만약 엑손 3이 접합에서 제외된다면, 엑손 4는 상기 mRNA의 첫번째 개방 판독 프레임을 포함하고 GFP는 발현될 것이다(도의 하단에 위치).

도 1b: FACS 결과 분석에 적용되는 게이팅 예: 오직 감염된 세포만을 고려하고 네개의 집단으로 분리하였다: $dsRED^{-}GFP^{+}$, $dsRED^{+}GFP^{++}$, $dsRED^{++}GFP^{+}$ 및 $dsRED^{+}GFP^{-}$. 상기 집단의 감염된 세포의 백분율은 결과 분석에 고려되었다.

도 2: 접합 구조체의 세부 묘사. (2a) dsRED에 대한 개방 판독 프레임을 포함하는 대체 엑손의 스플라이스 수용체 자리에서 변경. 상기 변경은 분기점 및 폴리(Y) 관이라 불리는 인트론-엑손 컨센서스 부위 사이의 피리미딘의 수(Ys; the bases C and T), 인트론-엑손 컨센서스 서열의 변경 및 분기점의 변경을 포함한다. (2b) GFP를 코딩한 엑손의 윗줄기 두번째 스플라이스 수용체의 상기 폴리(Y) 관에서 변경. 원본의 구조체에서 cTNT-I5가 사

용되었다. 상기 폴리(Y) 관은 Y에서 농축되었다. 원본 구조체와 비교시 (I5), Y의 양은 거의 3배 증가하였다. (2c) 대체 엑손의 아래줄기에 위치한 cTNT-I4의 스플라이스 공여 자리의 제거. 원본 I4 배열 및 단축된 I4(sh)의 배열을 나타냈고, 엑손-인트론 컨센서스 배열이 결여되어 있다.

도 3: HEK293의 일시적 형질 감염 (3a) 또는 폴리(Y) 관에 변경을 수반한 대체 접합 구조체의 CHO-S (3b) 세포. 게이팅(Gating)이 도 1에 설명된 바와 같이 수행되었다. 숫자는 감염된 세포의 각각의 집단($dsRED^{-}GFP^{+}$, $dsRED^{+}GFP^{++}$, $dsRED^{++}GFP^{+}$ 및 $dsRED^{+}GFP^{-}$)의 백분율을 나타낸다. GSC2250의 기초 구조는 GFP 보다(엑손 #4에서- 도 1) dsRED의 발현에 강한 선호를 나타낸다(엑손 #3, 대체 엑손에서- 도 1). AS 인트론 #2의 폴리(Y) 관에서 Y의 함량은 dsRED를 코딩한 엑손의 스플라이스 수용 자리를 약화하기위해 감소되었고 AS 인트론 #3의 폴리(Y) 관에서 Y의 함량은 GFP를 코딩한 엑손의 스플라이스 수용 자리를 강화하기위해 증가되었다. dsRED를 코딩한 엑손의 스플라이스 수용체 자리의 감소, 특히, 5Y-5, 5Y누드 및 0Y 구조체를 나타내는 크지만, 완만한 이동이 관찰되었다. GFP를 코딩한 엑손의 스플라이스 수용체 자리의 증가를 나타내는 어떤 효과도 관찰되지 않았다. 일반적인 경향은 CHO-S 및 HEK293 세포에 대해 같은 것으로 나타났다. 양성 대조군으로, 세포를 GFP 또는 dsRED만으로 형질 감염시켰다.

도 4: 분기점 부위 및 인트론-엑손 컨센서스 배열에서(각각, 4a 및 4b의 상단 열) 및 HEK293 세포(4a) 및 CHO-S 세포 (4b)에 대한 인트론 배열(각각, 4a 및 4b의 중간 열)의 변경. (4a) 및 (4b)의 하단 열, 각각: 양성 대조군으로 세포를 dsRED 또는 GFP만으로 형질 감염시켰다. 상기 GSC2250 구조체는 (cTNT-I4/cTNT-I5) 기초 구조체의 스플라이스 비율에 대한 기준으로 포함되었다. 숫자는 감염된 세포의 각각의 집단($dsRED^{-}GFP^{+}$, $dsRED^{+}GFP^{++}$, $dsRED^{++}GFP^{+}$ 및 $dsRED^{+}GFP^{-}$)의 백분율을 나타낸다. 게이팅은 도 1에 설명된 바와 같이 수행되었다.

도 5: cTNT-I4/cTNT-I4 구조체의 폴리(Y) 관에서 Y의 감소 및 분기점 부위의 서열 변경. (5a) HEK293 세포의 형질감염. 상단 열: 상기 폴리 (Y)에서 Y의 양의 감소는 GFP의 발현에 주요한 영향을 준다. 중간 열: 분기점 부위에서 변경. GFP의 발현에서 주요한 증가는 식별되지 않았다. 하단 열: 세포를 dsRED 또는 GFP 만으로 형질 감염시켰다. GSC 2250 구조체는 기초 구조체의 스플라이스 비율에 대한 기준으로 포함되었다. (5b) CHO-S 세포의 감염. 실험의 설정은 (5a)의 상단 및 하단 열에 동등하게 하였고 결과는 유사하였다. 숫자는 감염된 세포의 각각의 집단($dsRED^{-}GFP^{+}$, $dsRED^{+}GFP^{++}$, $dsRED^{++}GFP^{+}$ 및 $dsRED^{+}GFP^{-}$)의 백분율을 나타낸다. 게이팅은 도 1에 설명된 바와 같이 수행되었다.

도 6: 두번째 스플라이스 공여 자리의 제거는 대체 접합 비율을 더욱 이동시켰다. 감염은 CHO-S 세포에서 이뤄졌다. 몇몇 구조체에서, 두번째 스플라이스 공여 자리의 제거는 첫번째 엑손의 측면 부위에서 폴리(Y) 관의 감소와 함께 수반되었다. 이때 두번째 개방 판독 프레임에 대한 대체 접합의 이동은 한층 더 뚜렷해졌다. dsRED 및 GFP는 각각 세포에서 감염시켰고 대조군으로 사용되었다. cTNT-I4/cTNT-I4 기초 구조체는 이전 구조체의 스플라이스 비율에 대한 대조군으로 사용하기위해 포함되었다. 숫자는 감염된 세포의 각각의 집단($dsRED^{-}GFP^{+}$, $dsRED^{+}GFP^{++}$, $dsRED^{++}GFP^{+}$ 및 $dsRED^{+}GFP^{-}$)의 백분율을 나타낸다. 게이팅은 도 1에 설명된 바와 같이 수행되었다.

도 7: GFP 발현 대 dsRED 발현의 계략도. 대체 접합 건은 구조체에 따라 다른 평형을 가진다. dsRED의 대부분, dsRED 및 GFP의 중간양, 또는 GFP의 대부분으로 발현하는 구조체가 제조되었다.

도 8: 전형적인 GFP 및 dsRED 발현의 여덟개의 무작위로 선택된 클론.

도 9: 구조체의 서열 정렬.

도 10: pGLEX3 중추에서 안티-HER2 항체를 발현하는 구조체의 발현 결과. 상기 구조체는 대체 엑손의 순서로 첫번째 및 구조체에서 폴리(Y)의 감소하는 순서로 두번째의 순서로 나열되어 있다.

상기 최대 발현하는 두 구조체는 LC-HC: I4(0Y)-I4 방향 및 HC-LC: I4(7Ynude)-I4sh 방향을 향해있다.

도 11: 인트론-엑손 컨센서스 부위 변경 및 분기점 변이를 사용한 안티-HER2 항체 대체 접합 카세트의 미세 조정. 12 웰 플레이트 규모에서 표 7에 열거된 구조체의 사전 선택 후에(데이터 미도시), 선택된 구조체는 튜브스핀 규모에서 재평가되었다. 옥텟 장치(Fortebio, Melo Park, CA)를 사용하여 감염시킨 후 6일째에 역가가 결정되었다.

도 12: 대체 엑손의 윗줄기 및 아래줄기의 동일한 인트론은 보다 높은 발현으로 이어진다. 동일한 인트론이 상기 대체 엑손의 앞과 뒤에 사용될 경우, 두개의 다른 방향에 대하여 최대 발현이 관찰되었다. 상기 대체 엑손의

측면 cTNT-I4 인트론을 사용하는 경우, 발현 수준이 가장 높은 것으로 관찰되었다.

도 13: 진탕 생물 반응기에 37℃, 5% CO_2, 및 80%의 습도에서 2주 보충 배치의 끝에 튜브스핀 50 ml 생물 반응기에서 72 미니풀의 발현 수준. 클론들은 발현 수준이 감소하는 순서로 랭크되었다.

도 14: 진탕 생물 반응기에 37℃, 5% CO_2, 및 80%의 습도에서 2주 보충 배치의 끝에 튜브스핀 50 ml 생물 반응기에서, 각각 #68, 164 및 184 모 미니풀에 대한 최고 23 클론 및 #148 모 미니풀에 대한 최고 25 클론의 발현 수준. 상기 모 미니풀의 발현 수준은 개방형 막대로 나타내었고, 각각의 미니풀에서 유도된 클론의 발현은 폐쇄형 막대로 나타내었다.

도 15: 다른 비율에서 경쇄로 보조-감염된 대체 접합 구조체의 발현 수준.

발명을 실시하기 위한 구체적인 내용

[0033] 본 발명은 대체 접합을 사용하여 숙주 세포에서 이중 특이성 항체 또는 재조합 항체 또는 이의 조각과 같은 폴리펩티드, 특히 이종 다량 폴리펩티드의 발현 방법 및 발현 구조체를 제공한다. 본 발명은 각기 다른 폴리펩티드로 후속 번역과 둘 또는 그 이상의 mRNA로 접합될 수 있는 전-mRNA의 전사를 유도하는 단일 프로모터를 사용하여 숙주 세포에서 발현될 수 있는 구조체를 제공한다.

[0034] 여기서 상호 교환적으로 사용되는 "발현 구조체" 또는 "구조체" 용어는 발현되는 폴리펩티드가 코딩된 폴리뉴클리오티드 서열 및 프로모터와 cis-작용 전사적 제어 요소의 조합을 포함하여, 선택적으로 강화된 서열과 같은 그것의 발현을 조절하는 서열을 포함한다.

[0035] 상기 유전자의 발현을, i.e. 이의 전사 및 전사 제조물의 번역, 조절하는 서열은 일반적으로 규제 단위라고 한다. 규제 단위의 대부분은 유전자의 코딩 서열의 윗줄기에 위치해 있고 작동 가능하게 이에 연결되어 있다. 상기 발현 구조체는 폴리아데닐화 자리를 포함하는 아래줄기 3' 비번역 부위를 함유한다. 상기 본 발명의 규제 단위는 상기 유전자에, i.e. 전사 단위, 작동 가능하게 연결되어 발현되거나 또는 예를들어 이종 유전자의 5'-비번역 부위(5'UTR)에 의한 것과 같은 개재 DNA에 의해 이로부터 분리된다. 바람직하게 상기 발현 구조체는 벡터 및/또는 벡터로부터 이의 절단으로 발현 구조체의 삽입을 가능하게 하기 위해 하나 또는 그 이상의 적합한 제한 부위에 의해 측면으로 접해있다. 따라서, 본 발명에 따른 상기 발현 구조체는 발현 벡터, 특히 포유류의 발현 벡터의 구축을 위해 사용될 수 있다.

[0036] 여기서 사용된 "폴리펩티드가 코딩된 폴리뉴클리오티드 서열" 용어는 유전자에 대한 DNA 코딩, 바람직하게 폴리펩티드를 발현하는 이종 유전자를 포함한다. "이종 코딩 서열", "이종 유전자 서열", "이종 유전자", "재결합 유전자" 또는 "유전자"는 상호 교환적으로 사용되었다. 상기의 용어들은 재결합 유전자, 특히 숙주 세포, 바람직하게 포요류 세포에서 발현되고 수확되는 것으로 보여지는 재결합 이종 단백질 산물을 코딩하는 DNA 서열을 말한다. 상기 유전자의 산물은 폴리펩티드가 될 수 있다. 상기 이종 유전자 서열은 숙주 세포에서 자연적으로 존재하지 않고 다른 종 및 일반적으로 변형된 종 또는 같은 종의 유기체로부터 유도된다.

[0037] "단백질" 및 "폴리펩티드" 용어는 상호 교환적으로 사용되고 인접한 잔기의 카르복시기와 알파-아미노 사이의 펩티드 결합에 의해 다른 아미노산과 연결된 아미노산 잔기 계열을 포함한다.

[0038] 여기서 사용된 "프로모터" 용어는 일반적으로 RNA 중합효소가 DNA에 결합하도록 유도하고 RNA 합성을 시작함으로써 전사의 개시를 매개하는 유전자의 윗줄기에 위치하는 규제 DNA 서열로 정의된다. 본 발명에서 사용하기 위한 프로모터는, 예를들어, 높은 수준의 발현을 제공하는 바이러스, 포유류, 곤충 및 효모 프로모터, e.g. 포유류 시토메갈로바이러스(거대세포바이러스) 또는 CMV 프로모터, SV40 프로모터, 또는 진핵 세포에서 발현에 적합한 기술 분야에서 알려진 어느 프로모터를 포함한다.

[0039] "5' 비번역 부위(5'UTR)" 용어는 전-mRNA 또는 성숙 mRNA의 5' 말단에 비번역 구간을 말한다. 성숙 mRNA에서, 상기 5'UTR은 전형적으로 그것의 5' 말단에 7-메틸구아노신 캡을 품고 있고 접합, 폴리아데닐화 반응, 세포질로 mRNA 배출, 번역 장치에 의한 mRNA의 5' 말단의 식별 및 분해에 대한 mRNA의 보호와 같은 많은 과정에 관여한다.

[0040] "인트론" 용어는 전-mRNA에 존재하고 전사되지만 각각 인트론의 5' 및 3' 말단에서 수용 스플라이스 자리 및 공여 스플라이스 자리의 서열에 기초하여 접합 장치에 의해 제거되고, 이에 성숙 mRNA 전사에는 존재하지 않는 핵산 비-코딩 서열의 부분을 말한다. 전형적으로 인트론은 3' 스플라이스 자리의 윗줄기 20 내지 50 뉴클리오티드 사이에 위치하는 ,분기점이라 불리는, 내부 자리를 갖는다. 본 발명에서 사용되는 인트론의 길이는 50 내지 450

뉴클리오티드 사이의 길이일 수 있다. 단축된 인트론은 50 또는 그 이상의 뉴클리오티드를 포함한다. 완전체 길이 인트론은 450 뉴클리오티드까지 포함할 수 있다.

[0041] "엑손" 용어는 mRNA로 전사되는 핵산 서열의 부위를 말한다.

[0042] "스플라이스 자리" 용어는 해당 스플라이스 자리에 결찰 및/또는 잘리기에 접합한 진핵 세포의 접합 장치에 의해 식별될 수 있는 특정한 핵산 서열을 말한다.

[0043] 스플라이스 자리는 전-mRNA 전사에 존재하는 인트론의 절단을 허용한다. 전형적으로 스플라이스 자리의 5' 부분을 스플라이스 공여 자리라고 말하고 3' 해당하는 스플라이스 자리는 수용 스플라이스 자리라고 말한다. 스플라이스 자리는, 예를들어, 자연적으로 발생하는 스플라이스 자리, 조작된 스플라이스 자리, 예를들어, 합성 스플라이스 자리, 정형 또는 컨센서스 스플라이스 자리, 및/또는 비-정형 스플라이스 자리, 예를들어, 애매한(cryptic) 스플라이스 자리를 포함한다.

[0044] "폴리(Y) 관" 용어는 인트론-엑손 경계 및 분기점 사이에서 발견되는 핵산의 구간을 말한다 (도 2a 또는 2b에 도시되어 있다). 상기 핵산의 구간은 폴리피리미딘(Ys)을 풍부하게 갖고 있고, 피리미딘 염기 C 또는 T가 풍부함을 뜻한다.

[0045] "3' 비번역 부위 (3'UTR)" 용어는 전-mRNA 또는 성숙 mRNA의 3' 말단에 비번역 구간을 말한다. 성숙 mRNA에서 이 부위는 폴리(A) 꼬리를 내포하고 있고 mRNA 안정성, 번역 개시 및 mRNA 배출에서 많은 역할을 갖는 것으로 알려져있다.

[0046] 여기서 사용된 "증강" 용어는 유전자의 식별, 유전자 관련된 서열의 위치 또는 상기 서열의 방향과 무관한 유전자의 전사를 강화시키는 역할을 하는 뉴클리오티드 서열로 정의된다. 본 발명의 벡터는 선택적으로 증강을 포함한다.

[0047] "폴리아데닐화 신호" 용어는 mRNA 전사에 존재하는 핵산 서열을 말한고, 전사를 가능하게 하며, 폴리(A) 중합효소 존재하에서, 상기 폴리(A) 신호 후속의 10 내지 30 염기가 위치한 폴리아데닐화 자리에서 폴리아데닐화 되도록 한다. 많은 폴리아데닐화 신호들은 당해 분야의 기술에 공지되어 있고 본 발명에서 유용할 수 있다. 예로는 사람의 변형 성장 호르몬 폴리아데닐화 신호, SV40 후속 폴리아데닐화 신호 및 소 성장 호르몬 폴리아데닐화 신호를 포함한다. "기능적으로 연결된" 및 "작동 가능하게 연결된" 용어는 상호 교환적으로 사용되었고 둘 또는 그 이상의 DNA 부위, 특히 발현되는 유전자 서열 및 이의 발현을 조절하는 하기 서열 사이의 기능적인 관계를 말한다. 예를들어, cis-작동 전사적 조절 인자의 가능한 조합을 포함하는, 프로모터 및/또는 증진 서열은 이것이 적절한 숙주 세포 또는 여타의 발현 시스템에서 코딩 서열의 전사를 자극 또는 변조한다면 작동 가능하게 코딩 서열과 연결되어 있다. 전사된 유전자 서열에 작동 가능하게 연결되어 있는 프로모터 규제 서열들은 전사된 서열에 물리적으로 인접해 있다. "방향"은 주어진 DNA 서열에서 뉴클리오티드의 순서를 말한다. 예를들어, 또 다른 DNA 서열에 관계에서 반대 방향의 DNA 서열의 방향은 상기 서열이 제조되는 DNA에서 기준의 지점과 비교시 또 다른 서열에 관계 에서 서열의 5'에서 3'로 순서가 반전되는 방향이다. 이러한 기준점들은 서열을 함유하는 복제 벡터의 복제의 방향 및/또는 원본 DNA에서 다른 특정화된 DNA 서열의 전사의 방향을 포함할 수 있다.

[0048] 여기서 사용된 "핵산 서열 상동성" 또는 "뉴클리오티드 서열 상동성" 용어는 서열을 정렬하고 차이를 도입한 후에, 필요하다면, 최대 백분율 서열 동일성을 달성하기 위해 비교 서열의 뉴클리오티드 서열과 동등한 후보 서열에서 뉴클리오티드의 백분율 e.g. 두번째 측면 인트론의 뉴클리오티드 서열과 동일한 첫번째 인트론에서 뉴클리오티드의 백분율을 포함한다. 따라서 서열 동일성은 두 뉴클리오티드 서열의 뉴클리오티드의 위치에서 유사성을 비교하기 위해 일반적으로 사용되는 표준 방법으로 결정될 수 있다. 대게 상호에 상기 측면 서열의 상기 핵산 서열 상동성은 최소 80%, 바람직하게 최소 85%, 더 바람직하게 최소 90%, 및 가장 바람직하게 최소 95%, 특별히 96%, 더 특별히 97%, 및 한층 더 특별히 98%, 가장 특별히 99%이고, 예를들어, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 및 100%를 포함한다.

[0049] 여기서 사용된 "발현 벡터" 용어는 숙주 세포 내의 재조합 유전자 산물의 높은-수준 발현을 제공하는 적절한 숙주 세포로 감염에 따른 분리 및 정제된 DNA 분자를 포함한다. 재조합 또는 유전자 산물에 대한 DNA 서열을 코딩하는 것 외에도 상기 발현 벡터는 숙주 세포주에서 단백질로 mRNA의 효율적인 번역 및 mRNA로 DNA 코딩 서열의 효율적인 전사를 위해 요구되는 규제 DNA서열을 포함한다.

[0050] 핵산 서열의 길이와 관련해서 여기서 사용되는 '대략' 용어는 ± 50%의 최대 편차, 바람직하게 명시된 값의 ± 10%의 최대 편차를 포함한다 e.g. 대략 50 뉴클리오티드는 25 내지 75 뉴클리오티드, 바람직하게 45 내지 55 뉴클리오티드의 값을 포함하고, 대략 450 뉴클리오티드는 225 내지 675 뉴클리오티드, 바람직하게 405 내지 495

뉴클리오티드의 값을 포함한다. 여기서 사용된 "숙주 세포" 또는 "숙주 세포주" 용어는 어느 세포든, 특히 포유류 세포를 포함한다, 이것들은 배양에서 성장할 수 있고 원하는 재조합 산물 단백질을 발현할 수 있다.

[0051] 재조합 폴리펩티드 및 단백질은 원핵 (e.g. E.coli), 진핵 (e.g. 효모, 곤충, 척추동물, 포유류), 및 시험관 내 발현 시스템과 같은 다양한 발현 시스템에서 제조될 수 있다. 가장 흔하게 사용되는 단백질-기초 생물학의 대량 생산 방법은 DNA 벡터의 감염에 의한 숙주 세포로 유전 물질의 도입에 의존한다. 폴리펩티드의 일과성 발현은 숙주 세포의 일과성 감염과 함께 달성될 수 있다. 상기 숙주 세포 유전자로 벡터 DNA의 통합은 안정적으로 감염된 세포주의 결과로 나타나고 이러한 안정한 세포주의 증식은 폴리펩티드 및 단백질의 대량 생산에 사용될 수 있다. 과거 개시된 대체 접합 접근법과 다르게, 본 출원자들은 발현 구조체에서 다중 공여 및 수용 자리의 사용을 통해 원하는 비율에서 폴리펩티드의 발현에 대한 대체 접합 접근법을 설계했다. 이러한 접근법은 종래 접근법에서 제조된 것과 비교하여 60 배까지 더 높은 일과성 역가로, 높은 일과성이고 안정한 역가의 폴리펩티드가 제조될 수 있도록 만든다. 예를들어, 본 발명의 발현 구조체를 사용하여 일과성 감염에 따라 항체의 15 μg/ml 까지의 역가가, 예를들어, WO200589285 상의 표 1에서 관찰된 0.25 μg/ml의 수준과 비교하여, 관찰되었다. 안정적으로 감염된 세포주에 대해, 항체의 200 μg/ml까지의 역가가 배치 배양에서 관찰되었고 (도 13), 이것은 2차 제한 희석에 따라 250 μg/ml 까지 증가되었다 (예 4). WO200589285와 비교에서, 상기문헌, 안정한 풀의 특정 생산성의 최대 역가가 377 ng/ml인 것으로 관찰되었고 (상기문헌, WO200589285, 도 4에 개시), 본 출원자들에 의해 관찰된 상기 역가 수준은 650배 더 높고, 종래 기술에서 관찰된 값 보다 광대한 증가를 얻었다. 본 발명의 발현 구조체는, 폴리펩티드가 코딩된, 두개의 대체 엑손을 포함한다. 스플라이스 공여 자리는 첫번째 엑손의 윗줄기 및 아래줄기 둘다에 포함된다. 또한, 스플라이스 수용 자리는 첫번째 엑손의 윗줄기 및 아래줄기 둘다에 포함된다. 본 발명의 바람직한 실시예에서, 상기 첫번째 엑손은 동일한 인토론의 기능적 복사에 의해 측면에 위치해 있다. 스플라이스 발생 동안, 이러한 동일한 인트론 서열은 잘려지게 되고 성숙 mRNA에 존재하지 않는다. 이러한 구조체는 기능적으로 자연 발생 대체 엑손과 유사하다. 본 발명의 발현 구조체에서 사용하기에 접합한 인트론은 하기로 이루어진 군으로부터 선택될 수 있다: β-글로빈/IgG 키메릭(chimeric) 인트론, β- 글로빈 인트론, IgG 인트론, 마우스 CMV 첫번째 인트론, 랫트 CMV 첫번째 인트론, 사람 CMV 첫번째 인트론, Ig 다양한 부위 인트론 및 스플라이스 수용체 서열 (Bothwell et al., (1981) Cell, 24: 625-637; US5, 024,939), introns of the 치킨 TNT 유전자의 인트론 및 EF1알파의 인트론, 바람직하게 EFlalpha의 첫번째 인트론. 바람직한 한 실시예로, 상기 첫번째 엑손의 측면 인트론은 cTNT 인트론 4번(cTNT-I4), cTNT 인트론 5번(cTNT-I5) 또는 상기 EFlalpha 첫번째 인트론. 보다 바람직한 한 실시예로, 상기 첫번째 엑손의 측면 인트론은 cTNT-14이다.

[0052] 상기 첫번째 및 두번째 엑손 사이의 발현의 비율을 조정하기 위해서, 상기 첫번째 엑손의 윗줄기 인트론에서 작은 변화가 도입될 수 있다. 이러한 변화는 상기 첫번째 엑손의 윗줄이에 위치한 폴리피리미딘 (폴리(Y)) 관에 피리미딘 염기의 수의 변경을 포함한다. 실시예 2에 설명된 바와같이, 폴리(Y) 관에 피리미딘 염기의 수를 변경하는 것은 상기 첫번째 및 두번째 엑손의 발현에 주요한 영향을 줄 수 있다. 예를들어, 폴리(Y) 관에 피리미딘 염기의 수를 증가시키는 것은 두번째 폴리펩티드를 코딩하는 상기 두번째 엑손의 스플라이스 수용 자리를 강화한다. 선택적으로, 폴리(Y) 관애 피리미딘 염기의 수를 감소시키는 것은 첫번째 폴리펩티드를 코딩하는 상기 첫번째 엑손의 스플라이스 수용 자리를 약화시킨다. 상기 첫번째 엑손의 윗줄기 스플라이스 수용 자리의 강도를 감소시키는 것은 첫번째 엑손의 배제를 야기하도록 하고 따라서 상기 두번째 엑손으로부터 더 높은 발현이 나타난다. 본 발명의 한 실시예에서, 상기 발현 구조체는 상기 첫번째 엑손의 윗줄기 폴리(Y) 관을 포함한다. 폴리(Y) 관에 피리미딘 염기의 수는 0 내지 30 염기 사이를 포함한다.

[0053] 바람직하게 상기 폴리(Y) 관은 28, 27, 26, 25 및 24 염기로 이루어진 군으로부터 선택되는 피리미딘 염기의 수를 포함한다.

[0054] 더 바람직하게, 상기 폴리(Y) 관은 10 피리미딘 염기 또는 그 이하, 한층 더 바람직하게 7 염기 또는 그 이하, 가장 바람직하게 5 염기 또는 그 이하를 포함한다. 본 발명의 한 실시예에서, 상기 폴리(Y) 관은 상기 발현 구조체에서 결여되어 있다.

[0055] 본 발명의 또 다른 실시예에서, 상기 첫번째 엑손부터 두번째 엑손까지 발현의 비율을 이동시키기 위해서, 상기 두번째 엑손의 윗줄기 상기 두번째 스플라이스 공여 자리는 삭제될 수 있다. 이러한 삭제는 상기 두번째 스플라이스 부위의 윗줄기 전체적인 인트론 및 엑손-인트론 컨센서스 부위를 삭제함으로써 달성될 수 있다. 이러한 삭제는 상기 첫번째 폴리펩티드의 발현에서 상기 두번째 폴리펩티드의 발현으로 이동을 증가시켰다. 바람직한 실시예에서, 상기 두번째 스플라이스 공여 자리의 제거는 상기 발현 구조체의 첫번째 엑손의 윗줄기 폴리(Y) 관에 피리미딘 염기의 수에서 감소와 함께 수반될 수 있다. 실시예 1에서 설명된 바와 같이, 상기와 같은 두 양태의 조합은 상기 두번째 엑손 및 이로부터 상기 두번째 폴리펩티드의 거의 우세한 발현을 야기 시킨다. 본 발명의

한 양상에서, 상기 첫번째 및 두번째 엑손 사이의 발현의 비율은 상기 첫번째 엑손 측면에 위치한 동일한 서열의 인트론을 사용, 상기 폴리(Y) 관에 피리미딘 염기의 수를 변경 및/또는 상기 두번째 측면 인트론의 윗줄기 스플라이스 공여 자리를 제거함으로써 변경될 수 있다. 본 발명의 또 다른 실시예에서, 상기 발현 구조체는 상기발현 구조체의 5' 말단에 프로모토 부위의 아래줄기 인트론을 측면에 위치 시키는 스플라이스 수용 자리 및 스플라이스 공여 자리를 더 포함할 수 있다. 상기의 구조의 인트론, 스플라이스 공여 및 스플라이스 수용 자리는 전-mRNA가 성숙 mRNA로 성숙동안 구조적으로 접합된다. 상기 발현 구조체의 이러한 구조적 요소들은 5' 비번역 부위에 의해 상기 첫번째 엑손의 윗줄기 인트론으로부터 분리된다. 본 발명의 또 다른 실시예에서, 폴리아데닐화 자리는 상기 구조체의 3' 말단에 상기 두번째 엑손의 아래줄기에 위치한다.

[0056] 본 발명의 한 양상에서, 상기 발현 구조체는 둘 또는 그 이상의 폴리펩티드, 특히 폴리펩티드 다량체 예를들어 항체 또는 이의 조각을 발현 시키는데 적합하다.

[0057] 여기서 언급된 "항체" 용어는 모든 항체 및 어느 항원 결합 조각 또는 이의 단일쇄를 포함한다. "항체"는 디설파이드 결합 또는 이의 항원 결합 조각에 의해 상호-연결된 최소 두 중(H)쇄 및 두 경(L)쇄를 포함하는 글리코단백질(glycoprotein)을 말한다. 각각의 중쇄는 중쇄 가변 부위(여기서 약칭 VH) 및 중쇄 고정 부위가 포함된다. 상기 중쇄 고정 부위는 CHI, CH2 및 CH3, 세개의 영역이 포함된다. 각각의 경쇄는 경쇄 가변 부위(여기서 약칭 VL) 및 경쇄 고정 부위가 포함된다. 상기 경쇄 고정 부위는 CL, 하나의 영역이 포함된다. 상기 VH 및 VL 영역은, 서열에서 초 가변적인 상보 결정 영역(CDR)으로 불리는 및/또는 항원 인식에 포함되는 초가변성의 영역으로 더 세분화할 수 있고, 및/또는 프레임작업 영역 (FR or FW)라 불리는, 더 보존 되어진 영역에 분포하는, 일반적으로 구조적으로 정의된 루프를 형성한다. 각각의 VH 및 VL은 세개의 CDR 및 네개의 FW로 구성되어 있고, 아미노- 말단 에서 카복시- 말단으로 하기의 순서로 배열되어 있다: FW1, CDR1, FW2, CDR2, FW3, CDR3, FW4. 상기 FW1, FW2, FW3 및 FW4 아미노산 서열 모두 여기서 언급된 VH 또는 VL의 "비-CDR 영역' 또는 "비-확장 CDR 영역"을 구성한다.

[0058] 상기 중쇄 및 경쇄의 가변 영역은 항원에 상호작용하는 결합 영역을 포함한다. 상기 항체의 고정 영역은, 면역 시스템 (e.g., 작동 세포(effector cells))의 다양한 세포 및 전형적인 보체계의 첫번째 성분 (C 1q)을 포함하는, 조직 또는 인자들을 호스팅하는 면역글로불린의 결합을 매개한다.

[0059] 상기 고정 영역에 의해 유전적으로 결정되는 항체는, 또한 동기준 표본이라 불리는, 종류별로 분류된다. 사람 고정 경쇄는 카파(CK) 및 람다(CX) 경쇄로 분류된다. 중쇄는 뮤 (μ), 델타 (δ), 감마 (γ), 알파 (a), 또는 입실론 (ε)로 분류되고, 및 항체의 동기준 표본이 각각 IgM, IgD, IgG, IgA, 및 IgE으로 정의된다. 상기 IgG 종류는 치료의 목적으로 가장 흔하게 사용된다. 사람에서 상기 종류는 IgG1, IgG2, IgG3 및 IgG4 하위종류를 포함한다.

[0060] 여기서 사용된 "Fab" 또는 "Fab 영역" 용어는 VH, CHI, VL 및 CL 면역글로불린 영역을 포함하는 폴리펩티드를 포함한다. Fab는 완전한 길이의 항체 또는 항체 조각의 맥락의 상기 영역 또는 별개의 상기 영역을 말합니다.

[0061] 여기서 사용된 "Fc" 또는 "Fc 영역" 용어는 면역글로불린 부위 첫번째 고정 영역을 제외한 항체의 고정 영역을 포함하는 폴리펩티드를 포함한다. 이에 Fc는 IgA, IgD 및 IgG의 마지막 두개의 고정 영역 면역글로불린 부위, IgM 및 IgA의 마지막 세개의 고정 영역 면역글로불린 부위, 및 상기 부위에 유연한 힌지(hinge) N-말단을 말합니다. IgA 및 IgM에 대해, Fc는 J 쇄를 포함한다. IgG에 대해, Fc는 면역글로불린 부의 C 감마 2 및 C 감마 3 (Cy2 및 Cy3) 및 C 감마 1 (Cy1)과 C 감마 2 (Cy2) 사이에 힌지를 포함한다. Fc의 경계가 다양할 지라도, 상기 사람 IgG 중쇄 Fc 영역은 이의 카복시-말단에 일반적으로 C226 또는 P230 잔기를 포함하는 것으로 정의되고, 여기서 번호는 EU 번호체계에 따른다. 사람 IgG1에 대한 Fc 영역은 여기서 이의 카복시-말단에 P232 잔기를 포함하는 것으로 정의 되고, 여기서 번호는 EU 번호체계에 따른다. (Edelman GM et al, (1969) Proc Natl Acad Sci USA, 63(1): 78-85). Fc는 Fc 폴리펩티드의, 예를들어 항체, 맥락에서 상기 영역 또는 별개의 상기 영역을 말한다.

[0062] 여기서 사용된 "전체 길이 항체" 용어는, 가변 및 고정 영역을 포함하는, 항체의 자연 생물학적 형태를 구성하는 구조를 포함한다. 예를들어, 사람 및 쥐를 포함함, 대부분의 포유류에서, 상기 IgG 종류의 전체 길이 항체는 사량체이고 두개의 면역글로불린 쇄의 두개의 동일한 쌍으로 구성되어 있고, 각가의 쌍은 하나의 경쇄 또는 하나의 중쇄를 가지고, 각각의 경쇄는 면역글로불린 부위 VL 및 CL을 포함하고, 및 각각의 중쇄는 VH, CHI (Cy1), CH2 (Cy2), 및 CH3(Cy3) 면역글로불린 부위를 포함한다. 몇몇의 포유류에서, 예를들어 낙타와 라마에서, IgG 항체는 오직 두개의 중쇄들로 구성되어 있고, 각각의 중쇄는 Fc 영역에 붙어있는 다양한 부위를 포함한다.

[0063] 항체 조각은, 이에 제한되지 않으나, (i) Fab' 및 Fab'-SH를 포함하는, VL, VH, CL 및 CHI 부위로 이루어진 Fab 조각, (ii) VH 및 CHI 부위로 이루어진 Fd 조각, (iii) 단일 항체의 VL 및 VH 부위로 이루어진 Fv 조각; (iv) 단일 가변으로 구성되는 (Ward ES et al, (1989) Nature, 341 : 544-546) dAb 조각, (v) F(ab')2 조각, 두개의 연결된 Fab 조각을 포함하는 이가의 조각, (vi) 단일 쇄 Fv 분자 (scFv), 여기서 VH 부위 및 VL 부위는 항원 결합 자리를 형성하는 것과 연관되는 두개의 부위가 가능한 펩티드 링커에 의해 연결되어 있고 (Bird RE et al, (1988) Science 242: 423-426; Huston JS et al, (1988) Proc. Natl. Acad. Sci. USA, 85: 5879-83), (vii) 이중 특이성 단일쇄 Fv 이량체(PCT/US92/09965), (viii) "이가항체(diabodies)" or "삼가항체(triabodies)", 유전자 융합에 의해 구성된 다가 또는 다중 특이성 조각 (Tomlinson I & Hollinger P (2000) Methods Enzymol. 326: 461-79; WO94/13804; Holliger P et al, (1993) Proc. Natl. Acad. Sci. USA, 90: 6444-48) 및 (ix) 동일한 것 또는 다른 항체에 유전적으로 융합된 scFv (Coloma MJ & Morrison SL (1997) Nature Biotechnology, 15(2): 159-163)을 포함한다. 여기서 개시된 발현 구조체에 의해 발현될 수 있는 항체 및 이의 조각은 하기로 이루어진 목록으로부터 선택되는 항원에 결합할 수 있다: AXL, Bcl2, HER2, HER3, EGF, EGFR, VEGF, VEGFR, IGFR, PD-1, PD-1L, BTLA, CTLA-4, GITR, mTOR, CS1, CD3, CD16, CD16a, CD19, CD20, CD22, CD25, CD27, CD28, CD30, CD32b, CD33, CD38, CD40, CD52, CD64, CD79, CD89, CD137, CD138, CA125, cMet, CCR6, MUCI, PEM antigen, Ep-CAM, EphA2, 17- la, CEA, AFP, HLA class II, HLA-DR, HSG, IgE, IL-12, IL-17a, IL-18, IL-23, IL-lalpha, IL-lbeta, GD2-ganglioside, MCSP, NG2, SK-I antigen, Lag3, PAR2, PDGFR, PSMA, Tim3, TF, CTLA4, TL1A, TIGIT, SIRPa, ICOS, Treml2, NCR3, HVEM, OX40, VLA-2 및 4-1BB.

[0064] 이중 특이성 또는 이종 이량의 항체는 수년 동안 당해 분야에서 사용할 수 있었다. 그러나 이와 같은 항체의 생성은, 원하는 이중 특이성 항체의 생성 수율을 상당히 감소시키는, 짝지어지지 못한 부산물의 존재와 종종 연관되고 제품의 균일성을 달성하기 위해 복잡한 정제 과정을 요구한다. 상기 면역글로불린 중쇄의 짝지어지지 못함은 여러 합리적인 설계 전략을 사용하여 줄일 수 있고, 상기 대부분의 설계 전략은 CH 부위 동종 이량체의 두개의 하위단위 사이의 사람이-만든 상보적인 이종 이량의 인터페이스의 설계를 통해 이종 이량화를 위한 항체 중쇄를 조작하는 것이다. CH 이종 이량의 부위 쌍이 조작된 최초의 보고는 이종-이량의 Fc 부분을 생성하기 위한 "융기된 부분-으로-공동(protuberance-into-cavity)"을 제시하고 있는 Carter 등에 의해 만들어졌다 (US5,807,706; 'knobs-into-holes'; Merchant AM et al, (1998) Nat Biotechnol, 16(7):677-81). 대안적인 설계가 최근 개발되었고 WO2007110205에서 개시된 바와 같이, 핵심 조성을 변경함으로써 신규한 CH3 모듈 쌍의 설계 또는 WO2007147901 또는 WO2009089004에 개시된 바와 같이, 모듈간의 상보적인 염 다리의 설계가 관련되어 있다. 상기 CH3 조작 전략의 단점은 상기의 기술들은 여전히 상당한 양의 원치 않는 동종-이량체를 생성한다는 것이다. 주로 이종 이량체가 생산되는 이중 특이성 항체를 생성하기 위한 더욱 바람직한 기술을 WO2012131555에서 개시하고 있다. 이중 특이성 항체는 많은 대상, 예를들어, 종양 세포에 위치한 대상 및/또는 작동 세포에 위치한 대상에서 생성될 수 있다. 바람직하게, 이중 특이성 항체는 하기로 이루어진 군으로 부터 선택되는 두개의 대상에 결합할 수 있다: AXL, Bcl2, HER2, HER3, EGF, EGFR, VEGF, VEGFR, IGFR, PD-1, PD-1L, BTLA, CTLA-4, GITR, mTOR, CS1, CD3, CD16, CD16a, CD19, CD20, CD22, CD25, CD27, CD28, CD30, CD32b, CD33, CD38, CD40, CD52, CD64, CD79, CD89, CD137, CD138, CA125, cMet, CCR6, MUCI, PEM 항원, Ep-CAM, EphA2, 17- la, CEA, AFP, HLA 분류 II, HLA-DR, HSG, IgE, IL-12, IL-17a, IL-18, IL-23, IL-1알파, IL-1베타, GD2-강글리오사이드, MCSP, NG2, SK-I 항원, Lag3, PAR2, PDGFR, PSMA, Tim3, TF, CTLA4, TL1A, TIGIT, SIRPa, ICOS, Treml2, NCR3, HVEM, OX40, VLA-2 및 4-1BB.

[0065] 다른 관점에서, 본 발명은 앞서 설명된 바와 같이 발현 벡터 또는 발현 구조체를 포함하는 숙주 세포를 제공한다. 상기 숙주 세포는 사람 또는 비-사람 세포일 수 있다. 바람직한 숙주 세포는 포유류의 세포이다. 바람직한 포유류 숙주 세포의 예로는, 이에 한정되지 않고, 사람 배아 신장 세포 (Graham FL et al., (1977) J. Gen. Virol. 36: 59-74), MRC5 사람 섬유아세포, 983M 사람 흑색종 세포, MDCK 개과 신장 세포, 웅성 흰쥐(Sprague-Dawley rats)로부터 분리된 RF 배양된 쥐의 폐 섬유아세포, B16BL6 쥐과 흑색종 세포, P815 쥐과 비만 세포종 세포, MT1 A2 쥐과 유선암종 세포, PER:C6 세포 (Leiden, Netherlands) 및 중국 햄스터 난소(CHO) 세포 또는 세포주 (Puck TT et al., (1958), J. Exp. Med. 108: 945-955).

[0066] 특히 바람직한 실시 예에서, 숙주 세포는 중국 햄스터 난소(CHO) 세포 또는 세포주이다. 적합한 CHO 세포주는 e.g. CHO-S (Invitrogen, Carlsbad, CA, USA), CHO K1 (ATCC CCL-61), CHO pro3-, CHO DG44, CHO P12 또는 dhfr- CHO 세포주 DUK-BII (Urlaub G & Chasin LA (1980) PNAS 77(7): 4216-4220), DUXBI 1 (Simonsen CC & Levinson AD (1983) PNAS 80(9): 2495-2499), 또는 CHO-K1SV (Lonza, Basel, Switzerland)을 포함한다.

[0067] 본 발명의 바람직한 양상에서, 상기 두번째 폴리펩티드에 대한 상기 첫번째 폴리펩티드의 발현의 최적의 비율은

일과성 감염 실험으로 결정된다. 접합의 비율은 일과성 및 안정한 세포주에서 유사하게 남아있다. 최적 접합 비율의 구조체는, 예를들어, 최적의 비율에서 항체 중쇄 및 경쇄 (또는 이중 특이성 분자의 하위 단위)를 발현하는 세포주를 야기하는, 안정한 세포주 발생에서 사용될 수 있다. 본 발명의 한 실시예에서, 상기 발현 구조체는, 실시예 2에서 보이는 것과 같이, 다중 발생을 위한 변경되지 않은 비율에서 안정한 발현을 가능하게 한다. 또한, 원하는 비율에서 안정한 발현을 유지하기 위해 선택 압력의 사용을 필요로 하지 않는다.

[0068] 한 측면에서, 최적의 발현을 위한 경쇄에 대한 상기 항체 중쇄의 접합 비율은 1 : 1이다. 바람직하게 최적의 발현을 위한 경쇄에 대한 상기 항체 중쇄의 접합 비율은 1 : 2 또는 1 : 3 또는 2 : 3이다. 선택적으로, 최적의 발현을 위한 경쇄에 대한 상기 항체 중쇄의 접합 비율은 2 : 1 또는 3 : 1 또는 3 : 2일 수 있다. 최적 발현을 위한 상기의 비율은 각각의 항체에 의존적일 것이다. 또 다른 관점에서, 이중 특이성 항체의 최적 발현을 위한 각기 다른 하위단위들은 대체 접합을 사용하는 각 다른 비율에서 발현될 것이다. 본 발명의 바람직한 이중 특이성 항체는 중쇄, 경쇄 및 Fc-scFv의 하위단위를 포함한다. 이중 특이성 항체에 대하여, 본 발명에서 보여진 바와 같이, Fc-scFv 발현에 대한 중쇄의 비율은 가장 중요한 변수인 것으로 나타났다. 따라서 최적의 발현을 위한 Fc-scFv에 대한 중쇄의 접합 비율은 1 : 1이다. 바람직하게 최적의 발현을 위한 Fc-scFv에 대한 중쇄의 접합 비율은 1 : 2 또는 1 : 3 또는 2 : 3이다. 선택적으로, 최적의 발현을 위한 Fc-scFv에 대한 중쇄의 접합 비율은 2 : 1 또는 3 : 1 또는 3 : 2이다. 최적의 발현을 위한 상기의 비율은 각각의 항체에 의존적이다. 다른 관점에서, 본 발명 발명은, 상기에서 설명된 바와 같이 발현 벡터 또는 발현 구조체로 숙주 세포를 감염, 상기 숙주 세포를 배양 및 폴리펩티드를 회수하는 것을 포함하는 폴리펩티드의 발현을 위한 시험관 내(in vitro) 방법을 제공한다. 상기 폴리펩티드는 바람직하게 이종성이고, 보다 바람직하게 사람 폴리펩티드이다. 만약 주어진 숙주 세포 유형에 적절하다면 본 발명에 따른 숙주 세포에 발현 구조체 또는 발현 벡터를 감염시키기 위해 어떠한 감염 기술도, e. g. 전기 천공법, 인산 칼슘 공침법, DEAE-덱스트란 감염, 리포펙션법, 적용될 수 있다. 본 발명의 발현 벡터 또는 발현 구조체로 감염된 상기 숙주 세포는 일시적으로 또는 안정적으로 감염된 세포주로 해석되어야 함에 주목 해야한다. 따라서, 본 발명에 따른 본 발현 구조체 또는 발현 벡터는 유전자 부체 같이(episomally) 유지될 수 있다 i.e.일시적으로 감염되었고 또는 상기 숙주 세포의 게놈에 안정적으로 통합될 수 있다 i.e. 안정적으로 감염되었다. 일과성 감염은 선발표지 유전자 부담 벡터에 대한 모든 선택 압력의 비-기기(non-appliance)에 의한 것으로 특정지어진다. 일반적으로 감염 후 2 내지 10일까지 지속되는 일시적인 발현 실험에서, 상기 감염된 별현 구조체 또는 발현 벡터는 에피손 요소로 유지되고 아직 상기 게놈에 통합되지 않았다. 즉 상기 감염된 DNA는 상기 숙주 세포 게놈에 보통 통합되지 않는다. 상기 숙주 세포는 상기 감염된 DNA를 잃고 일시적으로 감염된 세포 풀의 배양에 따른 집단에서 감염된 세포를 과성장시키는 경향이 있다. 따라서 발현은 즉시 감염에 따른 시기에 가장 강하고 시간이 지남과 함께 감소한다. 바람직하게, 본 발명에 따른 일시적인 감염체는 감염 후 2 내지 10일의 시간까지 선택 압력의 부재하에 세포 배양에서 유지되는 세포로 이해된다.

[0069] 본 발명의 바람직한 실시예에서 상기 숙주 세포 e.g. CHO 숙주 세포는 본 발명의 상기 발현 구조체 또는 발현 벡터로 안정하게 감염된다. 안정한 감염은 벡터 DNA와 같은 새롭게 도입된 외부 DNA가, 일반적으로 무작위로, 비-동성 재조합인 건에 의해 게놈 DNA에 통합되어지는 것을 뜻한다. 상기 벡터 DNA의 복제 수와 부수적으로 유전자 생성물의 양은 상기 숙주 세포의 DNA로 통합 후에 상기 멕터 서열이 증폭되어진 세포주를 선택함으로써 증가되어질 수 있다. 따라서, 상기와 같은 안정한 통합은 CHO 세포에서 미량의 염색체를 두배로, 유전자 증폭을 위한 선택 압력에서 추가 증가에 노출시, 상승을 주는 것이 가능하다. 또한, 안정한 감염은 재조합 유전자 산물, e.g. 박테리아 복제수 조절 영역을 유전적 통합상에 불필요하게 만드는 것과 같은, 재조합 유전자 산물의 발현과 직접적으로 연관되지 않은 벡터의 서열 부분의 손실을 야기한다. 따라서, 감염된 숙주 세포는 게놈으로 발현 벡터 또는 발현 구조체의 다른 부분 또는 최소 부위에 합쳐진다.

[0070]

[0071] 또 다른 관점에서, 상기 본 발명은 포유류 숙주 세포로부터 이종 폴리펩티드의 발현을 위한 상기에서 설명된 바와 같은 상기 발현 구조체 또는 상기 발현 벡터의 용도, 특히 포유류 숙주 세포로부터 이종 폴리펩티드의 시험관 발현의 용도를 제공한다.

[0072] 본 발명에서 설명된 바와 같이 발현 구조체는 목적하는 단백질의 발현 수준을 최적화의 방법에서 사용될 수 있다. 예를들어, 상기 목적하는 단백질이 항체인 경우, 숙주 세포에 발현될 때 항체의 최적 발현 수준을 달성하기 위해, 중쇄에 대한 경쇄의 발현 비율 또는 반대인 경우의 비율은 변경될 수 있다. 5' 에서 3' 방향으로 하기를 포함하는 발현 구조체를 사용하는 단계:

[0073] 프로모터; 첫번째 선택형 스플라이스 공여 자리;

[0074] 첫번째 측면 인트론;

[0075] 스플라이스 수용 자리;

[0076] 첫번째 폴리펩티드가 코딩된 첫번째 엑손;

[0077] 두번재 선택형 스플라이스 공여 자리;

[0078] 두번째 측면 인트론;

[0079] 두번째 스플라이스 수용 자리; 및

[0080] 두번째 폴리펩티드가 코딩된 두번째 엑손,

[0081] 상기 목적하는 단백질의 발현 수준은 하기의 단계를 포함하는 방법으로 최적화된다:

[0082] 단백질의 발현 수준을 최적화하는 방법:

[0083] (i) 최소 50개의 뉴클리오티드의 핵산의 구간에 대하여 최소 80%의 핵산 서열 상동성을 가지는 첫번째 및 두번째 측면 인트론을 사용하는 단계(단계 i);

[0084] (ii) 첫번째 엑손의 윗줄기에 위치한 폴리(Y) 관에서 피리미딘 염기 수를 감소시키거나 첫번째 엑손의 아래줄기에 위치한 폴리(Y) 관에서 피리미딘 염기 수를 증가시키는 단계(단계 ii); 및/또는

[0085] (iii) 두번째 측면 인트론의 윗줄기 스플라이스 공여 자리를 삭제시키는 단계(단계 iii).

[0086] 또한, 본 발명에서 설명하는 바와 같은 발현 구조체는 목적하는 단백질의 이종이량화 수준을 최적화 하는 방법에서 사용될 수 있다. 예를들어, 목적하는 단백질이 이중 특이성 항체인 경우, 상기와 같은 이중 특이성 항체는 본 발명에 따른 하나 또는 그 이상의 발현 구조체에 의해 코딩될 수 있고, 이것은 중쇄, 경쇄 및 Fc-scFvr가 코딩 되어 있다. 여기서 설명된 바와 같이 대체 접합의 방법을 사용함으로써, Fv-scFv에 대한 중쇄의 발현 비율 또는 이의 반대인 경우의 비율, 예를들어, 숙주 세포에서 발현시 이중 특이성 항체의 최적 발현 수준을 달성하기 위해 변경될 수 있다.

[0087] 5'에서 3' 방향으로 하기를 포함하는 발현 구조체의 사용:

[0088] 프로모터;

[0089] 첫번째 선택형 스플라이스 공여 자리;

[0090] 첫번째 측면 인트론;

[0091] 스플라이스 공여 자리;

[0092] 첫번째 폴리펩티드가 코딩된 첫번째 엑손;

[0093] 두번재 선택형 스플라이스 공여 자리;

[0094] 두번째 측면 인트론;

[0095] 두번째 스플라이스 수용 자리; 및 두번째 폴리펩티드가 코딩된 두번째 엑손,

[0096] 상기 목적하는 단백질의 이종이량화 수준은 하기의 단계를 포함하는 방법으로 최적화된다:

[0097] 단백질의 발현 수준을 최적화하는 방법:

[0098] (i) 최소 50개의 뉴클리오티드의 핵산의 구간에 대하여 최소 80%의 핵산 서열 상동성을 가지는 첫번째 및 두번째 측면 인트론을 사용하는 단계(단계 i);

[0099] (ii) 첫번째 엑손의 윗줄기 폴리(Y) 관에서 피리미딘 염기 수를 감소시키거나 첫번째 엑손의 아래줄기 폴리(Y) 관에서 피리미딘 염기 수를 증가시키는 단계(단계 ii); 및/또는

[0100] (iii) 두번째 측면 인트론의 윗줄기 스플라이스 공여 자리를 삭제시키는 단계(단계 iii).

[0101] 상기 단백질의 발현 및 회수는 당해 분야의 당업자에 알려진 방법에 따라 수행될 수 있다.

[0102] 또 다른 관점에서, 본 발명은 질환의 치료를 위한 의약의 제조를 위해 위에서 설명된 바와 같이 상기 발현 벡터 또는 상기 발현 구조체의 용도를 제공한다.

[0103] 다른 관점에서, 본 발명은 질환의 치료를 위한 의약으로 사용하기 위한 위에서 설명한 바와 같이 상기 발현 구조체 또는 상기 발현 벡터를 제공한다.

[0104] 또 다른 관점에서, 본 발명은 유전자 치료에 사용하기 위한 위에서 설명한 바와 같이 상기 발현 구조체 또는 상기 발현 벡터를 제공한다.

[0105] 실시예

[0106] 실시예 1

[0107] 재료 및 방법

[0108] LB 배양 플레이트

[0109] 500 ml의 물을 16 gdml LB 한천과 혼합하고 끓여주었다. (Invitrogen, Carlsbad, CA, USA) (1 리터의 LB는 10 g 트립톤(tryptone), 5 g 효모 추출물 및 10 g NaCl을 함유한다). 냉각 후, 배양 접시에 분포된 상기 용액에 각각의 항생제를 첨가하였다 (100 μg/ml 암피실린(ampicilin) 플레이트 및 50 μg/ml 카나마이신 플레이트).

[0110] 중합효소 연쇄 반응 (PCR)

[0111] 모든 PCR은 50 μl의 최종 부피에서 1 μl의 dNTPs (10 mM for each dNTP; Invitrogen, Carlsbad, CA, USA), 2 단위의 Phusion® DNA 중합효소 (Finnzymes Oy, Espoo, Finland), 25 nmol의 프라이머 A (Mycrosynth, Balgach, Switzerland), 25 nmol의 프라이머 B (Mycrosynth, Balgach, Switzerland), 10 μl 의 5X HF 버퍼 (7.5 mM MgC12, Finnzymes, Espoo, Finland), 1.5 μl의 디메틸 설폭사이드 (DMSO, Finnzymes, Espoo, Finland) 및 1-3 μl의 주형 (10-20ng)를 사용하여 수

[0112] 행되었다.

[0113] 상기 PCR은 3분 동안 98℃에 개시 변성에 의해 시작되었고, 이어 98℃에 30 초 변성의 35회, 프라이머-특정 온도에서 30초 가열냉각 (CG 함량에 따라) 및 72℃에서 신장 (30 sec/kB의 주형). 냉각 및 4℃ 유지하기 전 10분 동안 최종 72℃에서 신장이 수행되었다. 상기 실시예에 대한 모든 프라이머는 하기 표 1에 열거되어 있다.

[0114] 표 1 : 복제를 위해 사용된 모든 프라이머 목록

표 1

[0115]

프라이머	서열 ID 번호:	서열
Glnpr991	001	GGTCATTTCGAATCATTACTTGTACAGCTCGT
Glnpr1095	002	CGCTGGCTAGCGTTTAAACTTAAG
Glnpr1096	003	ATCGTTCGAATATGGGCCCTCTCGCACACCGGTCTCCTCTTCCTCCTC
Glnpr1097	004	TATAGGGCCCTGTGAGCAAGGGCGAGGAG
Glnpr1098	005	GCGCTTCGAATCATTACTTGTACAGCTCGTC
Glnpr1099	006	TATAGGGCCCTCTACAGGAACAGGTGGTG
Glnpr1100	007	ATTAACCGGTGCCTCCTCCGAGGACGTC
Glnpr1138	008	AATTAAGCTAGCGTTTAAACTTAAGCTTCCTTGGATTACAAGGATGACGAT
Glnpr1139	009	GTGGCGATATCGCCTGGATCCTGAG
Glnpr1140	010	CCAGGCGATATCGCCACCATGGGTGCCTCCTCCGAGGA
Glnpr1141	011	CTACCTGAATTCTTCCGTTACTACAGGAACAGGTGGTGGCGGC
Glnpr1142	012	GAGGAGACCGGTGCCACCATGGAGCAAGGGCGAGGAGCTGT
Glnpr1158	013	AATTAAGCTAGCGTTTAAACTTAAGCTTCCTTGGAGGACCCAGTACCCGGATCTAGAGGTAGG
Glnpr1180	014	AATTAAACCGGTGCCACCATGGTGAGCAAGGGCGAGGAGC
Glnpr1181	015	GCGCGGCTAGCGTTTAAACTTAAGC
Glnpr1182	016	TTGTGATATCGCCTGGATCCTGTGCAATAAGGACAGGGTTAGCCAGGTGCCTTAAAGCTGTG
Glnpr1183	017	AGCAGGATATCGCCTGGATCCTGAGACAGGGAGGAGG
Glnpr1184	018	ATATGATATCGCCTGGATCCTGAGCCAGGGAGCAGGCAAGGCAAGAAGCGCAGAGGTTAGCC
Glnpr1185	019	AGTCGATATCGCCTGGATCCTGAGCCAGGTAGCAGGGAAGGGAAG

Glnprl186	020	GATGGATATCGCCTGGATCCTGAGCCAGGGAGGAGGGAAGGCAACAAGCGCAGAGGTTAGCC
Glnprl187	021	GCGCGAATTCAGGTAGTTACTGCAC
Glnprl189	022	TATAACCGGTCTCCTCTTCCTCCTCGTCCTCCTGATCCTCCTGACCTGAGCCAGGGAGGAGGGAAG
Glnprl190	023	TAATACCGGTCTCCTCTTCCTCCTCGTCCTCCTGATCCTCCTGACCTGAGCCAGGGAGCAGGCAAGGCAAGAAG
Glnprl191	024	ATATACCGGTCTCCTCTTCCTCCTCGTCCTCCTGATCCTCCTGACCTGAGACAGGGAGGAGGGAAG
Glnprl192	025	ATATACCGGTCTCCTCTTCCTCCTCGTCCTCCTGATCCTCCTGACCTGAGCCAGGGAGGAGGGAAG
Glnprl193	026	ATATACCGGTCTCCTCTTCCTCCTCGTCCTCCTGATCCTCCTGACCTGAGCCAGGTAGCAGGGAAGGGAAGAAG
Glnpr1237	027	GGCGGCTAGCGTTTAAACTTAAGCTTCCTTGGAGGACCCAGTACCCGGATCTAGAGTAGTTACTGCACCTTTCTTTG
Glnpr1238	028	ATCGGATATCGCCTGGATCCTGTGCAATAAGGACAGGGTC
Glnpr1239	029	GTGGCGATATCGCCTGGATCCTHTGCAATAAGGAC
Glnpr1240	030	TGGCGATATCGCCTGGATCCTGTGCAATAAGGACAGCCTTAGCCAGGTGCCTTAAAG
Glnpr1241	031	TGGCGATATCGCCTGGATCCTGTGCAATAAGGACAGGGTTCTCCAGGTGCCTTAAAG
Glnpr1242	032	TGGCGATATCGCCTGGATCCTGTGCAATAAGGACAGGGCAAGCCAGGTGCCTTAAAG
Glnpr1243	033	TGGCGATATCGCCTGGATCCTGTGCAATAAGGACAGCGTAGGCCAGGTGCCTTAAAG
Glnpr1244	034	GCGATATCGCCTGGATCCTGTCCCCTAAGGACTCGGTTAGCCAGGTGCCTTAAAGCTGTG
Glnpr1245	035	GCGATATCGCCTGGATCCTGTGCAATCCTCCCAGGGTTAGCCAGGTGCCTTAAAGCTGTG
Glnpr1246	036	GCGATATCGCCTGGATCCTGTTCCCTCCTCCCTCGGTTAGCCAGGTGCCTTAAAGCTGTG
Glnpr1285	037	CGGAAGAATTCAGCCACAGCTTTAAGGCACCTGGCTAAC

[0116] 제한효소 분해(Restriction digest)

[0117] 모든 제한효소 분해에 대하여 1 μg의 플라스미드 DNA를 (나노 드롭로 정량화된) was mixed to 10-20 단위의 각각의 효소units of each enzyme, 4 μl에 상응하는 10X NEBuffer (NEB, Ipswich, MA, USA)에 혼합하였고, 및 살균한 H_2O로 부피가 40 μl가 되도록 만들었다. Without further indication, digestions were incubated 1 hour at 37° C. After each preparative digestion of backbone, 1 unit of Calf Intestinal Alkaline Phosphatase (CIP; NEB, Ipswich, MA, USA) was added and the mix was incubated 30 min at 37° C.

[0118] PCR 정제 및 아가로오스 겔(Gel Agarose) 전기영동

[0119] 분해를 하기 위해 Macherey Nagel 핵스핀 추출 II 키트(Macherey Nagel, Oensingen, Switzerland)를 사용하여 제조사의 메뉴얼에 따라 PCR 조각들을 제한효소 분해에 앞서 씻어 주었다. DNA 샘플의 버퍼를 교환해주기 위해 이러한 프로토콜 또한 사용되었다. 겔 전기영동을 하기 위해, UltraPureTM 아가로오스 (Invitrogen, Carlsbad, CA, USA) 및 50X 트리스 아세트 산 EDTA 버퍼 (TAE, pH 8.3; Bio RAD, Munich, Germany)를 사용하여 1% 겔이 준비되었다. DNA의 염색을 하기 위해 1 μl의 겔 붉은 염료 겔 (Biotum, Hayward, CA, USA)이 100 ml의 아가로오스 겔에 첨가되었다. 사이즈 마커로써 2 μg의 1 kb DNA 레더 (NEB, Ipswich, MA, USA)가 사용되었다. 전기영동이 125볼트에서 1시간 동안 수행되었다.

[0120] 제조사의 메뉴얼에 따라, 아가로오스 겔로부터 목적하는 띠(bands)를 잘라내고 핵산 추출 II 키트 (Macherey-Nagel, Oensingen, Switzerland)을 사용하여 정제하였다.

[0121] 결찰(Ligation)

[0122] 각각의 결찰을 하기 위해, 4 μl의 삽입물이 10 μl의 부피에 1 μl의 벡터, 400 단위의 연결효소 (T4 DNA ligase, NEB, Ipswich, MA, USA), 1 μl 의 10X 연결효소 버퍼 (T4 DNA ligase buffer; NEB, Ipswich, MA, USA)에 혼합되었다. 상기 혼합물을 1-2시간 동안 실온에서 배양하였다.

[0123] 25-50 μl의 수용 박테리아(competent bacteria)를 (One Shot® TOP 10 Competent E. coli; Invitrogen, Carlsbad, CA, USA) 얼음에서 5분 동안 해동하였다. 5 μl의 결찰 생성물을 수용박테리아 첨가하고 42℃에서 1분 동안 열에의한 충격(thermic shock) 전에 얼음에서 20-30분 동안 배양하였다. 이후, 튜브당 500 μl 의

S.O.C 배양액을 (Invitrogen, Carlsbad, CA, USA) 첨가해 주고 열교반기(thermoshaker)에서 600rpm으로 교반하에 37℃에서 1시간 동안 배양하였다. 최종적으로, 상기 박테리아는 암피실린 (Sigma- Aldrich, St. Louis, MO, USA) 또는 카나마이신과 LB 플레이트에 넣어주고 37℃에서 밤새도록 배양하였다.

[0124] 소형 (mini) 및 중형 (midi) 크기의 플라스미드 제조

[0125] 소형-제조를 하기 위해, 형질 전환된 박테리아 군을 200 rpm, 37℃에서 암피실린 또는 카나마이신 및 LB의 2.5 ml에서 6-16시간 동안 성장시켰다. 제공된 메뉴얼에 따라, DNA가 E.co/z용 플라스미드 정제 키트로(NucleoSpin QuickPure or NucleoSpin Plasmid (No Lid), Macherey Nagel, Oensingen, Switzerland) 추출되었다.

[0126] 중형-제조를 하기 위해, 형질 전환된 박테리아를 37℃에 암피실린(또는 카나마이신) 및 LB 200 ml에서 밤새도록 성장시켰다. 이후, 제공된 메뉴얼에 따라, DNA가 E.co/z용 플라스미드 정제 키트로(NucleoSpin QuickPure or NucleoSpin Plasmid (No Lid), Macherey Nagel, Oensingen, Switzerland) 추출되었다. 상기 배양된 것을 725 g에 20분 동안 원심분리하고 상기 플라스미드는 제조사의 메뉴얼에서 제공한 프로토콜에 따라 구입한 키트를 (NucleoBond Xtra Midi; Macherey Nagel, Oensingen, Switzerland) 사용하여 정제되었다. 중형-제조로부터 플라스미드-DNA는 나노 드롭 ND- 1000 분광광도계로 세회 정량하고 제한효소 분해에 의해 확인되었으며, 및 최종적으로 서열(sequencing)하기 위해 보내졌다 (Fasteris SA, Geneva, Switzerland).

[0127] 세포의 배양 및 감염

[0128] 일반적인 계대배양방법(passaging)으로 세포들은 100 ml 성장 배양액에서 배양되었다 (PowerCHO2 (Lonza, Venders, Belgium), 4mM Gin for CHO-S cells and Ex-cell293 (Sigma- Aldrich, St. Louis, MO), 4mM Gin for HEK293 cells). 세포들은 0.5E6 cells/ml 일주일에 두번 심었고 진탕 배양기에서 80& 습도 및 5% CO_2의 대기에서 배양되었다.

[0129] 구조체들은 CHO-S 세포 및 HEK293 세포에서 감염시켰다. 감염을 하기 위해, 상기 세포들은 하루 동안의 감염에 앞서 1E6 cells/ml의 밀도로 심어주었다. 감염의 하루, 상기 세포들은 Optimem (CHO-S) 또는 RPMI (HEK293)에 재현탁 하고 제조사의 메뉴얼에 따라 JetPEI™ (Polyplus-transfection, Strasbourg, France)으로 감염시켰다. 5시간 경과 후, 각 성장 배양액의 하나의 볼륨이 첨가되었다 (HEK293 세포에 대해, 플루로닉 F68이 보충되었다). GFP 및 dsRED 발현을 위해 FACS로 감염 3 내지 5일 후 세포들을 분석하였다. 상기 감염은 각각 2 ml 또는 1 ml의 최종 부피를 사용하여 12 또는 24 웰 플레이트 (TPP, Trasadingen, Switzerland)에서, 또는 10 ml의 최종 배양액 부피를 사용하여 50 ml 생물 반응기 튜브 ("Tubespins", TPP)에서 수행되었다.

[0130] FACS 분석

[0131] 세포들은 앞쪽과 옆의 스케터를 사용하여 살아있는 세포에서 선별하였다. dsRED 및 GFP 발현 세포의 비율의 분석을 하기 위해, dsRED 감염 세포 및 GFP 감염세포를 사용하여 보정이 수행되었다. dsRED 에서 GFP 발현 세포로 이동의 평가를 하기 위해, 비-감염 세포는 게이트롤 추가함으로써 배재되었다.

[0132] 결과

[0133] 구조체 및 복제 단계의 설계

[0134] 동일 일차 전사의 두개의 다른 엑손에 위치한 두개의 대체 개방 판독 프레임의 발현을 가시화될 수 있도록 하기 위해, 형광 마커 GFP 및 dsRED가 사용되었다. 둘 모두의 단백질은 세포내에서 높은 수준으로 발현될 수 있었고, 세포에 잘 적용되었고 FACS 분석에서 또는 형광 현미경하에서 쉽게 구별되었다. 형광 마커의 단점은 측정된 형광이 쉽게 단백질의 양에 기인할 수 없었고 따라서 오직 또 다른 대상에 대한 한 단백질의 상대적인 발현 수준 상에서 결론만이 가능하다. 이에 이러한 초기 실험 상태에서, 다른 구좇체들이 엑손 1 및 2로부터 상대적으로 다른 발현 수준의 범위를 얻기 위해 만들어졌다 (도 1a의 스킴 참조).

[0135] 대체 접합 구조체는 대체 cTNT 엑손 5를 둘러싼 치킨 트로포닌 (cTNT) 인트론 4 및 5에 기초하여 만들어졌다. 트로포닌은 심장 근육 및 배아 골격 근육에 독점적으로 발현된다. 초기 배아 심장 및 골격 근육에서 mRNA의 90% 이상은 상기 엑손을 포함하고, 반면 성인에서 mRNA의 >95%은 상기 엑손을 배제한다 (Cooper & Ordahl (1985) JBC 260(20): 11140-8). 본 발명의 구조체에서, 상기 cTNT 인트론은 일차 전사의 두번째 및 세번째 인트론으로 복제되었다. 상기 첫번째 인트론은 mCMV 또는 hCMV 프로모터와 조합에 사용된 구성 인트론이다. 이번 실시예에서 사용된 상기 cTNT 인트론 이름은 인트론 서열을 지정하고 구조체에서 인트론의 위치를 지정하지 않는다 (cTNT 인트론 4는 상기 구조체에서 인트론 번호 2 또는 3이다).

[0136] 혼돈을 피하기 위해서 상기 cTNT 인트론 4는 cTNT-I4로 단축 시킬 수 있고 상기 cTNT 인트론 5는 cTNT-I5로 단축 시킬 수 있다, 각각의 상기 구조체에서 인트론의 위치는 AS 인트론 수를 사용하여 계산하였다 (예를들어 기본 구조체에서, cTNT-I4는 AS 인트론 #2 위치에 복제되었다). 기초 구조에서 (GSC2250), 상기 인트론 서열 cTNT-I4 (AS 인트론 #2) 및 cTNT-I5 (AS 인트론 #3)은 dsRED를 코딩하는 상기 개방 판독 프레임을 함유하는 변경된 대체 엑손을 측면에 위치시켰다. AS 인트론 #3는 (기초 구조체 cTNT-I5) GFP의 상기 개방 판독을 포함하는 상기 엑손을 뒤따른다 (도 1a 계략도 참조).

[0137] Orengo 등에 의해 개시된 벡터의 복제

[0138] 본 발명의 상기 대체 접합 구조체는 Orengo에 의해 개시된 구조체에 기초하였다 (Orengo JR et al, (2006) Nucleic Acids Res. 2006; 34(22): e148). 상기 구조체에서, 상기 발현 카세트의 시작 코돈은 dsRED 및 GFP를 코딩하는 개방 판독 프레임 사이에 공유되었고, 뒤이어 플레그 태그 및 짧은 핵 편재 서열이 왔다. 상기 치킨 트로포닌 인트론 4 및 5에 의해 측면에 위치된 상기 매우 짧은 대체 엑손은 저자에 의해 길이가 조절되어 대략 50%에서 제외되었다. 제외된 경우, 상기 dsRED의 개방 판독은 개시 코돈과 프레임 안에 있고 오직 dsRED만이 발현되었다. 작은 대체 액손의 포함은 상기 판독 프레임으로 프레임이동을 도입시킨다.

[0139] 상기 dsRED의 개방 판독 프레임은 두번째 프레임에서 판독되어 (dsRED의 상기 프레임에는 종결 코돈이 존재하지 않는다) leading to a fusion protein of dsRED (두번째 프레임에서 판독) 및 GFP의 융합 단백질을 야기시킨다. 상기 기술의 단점은 많이 있다. 첫째로, 단백질 중 하나는 필수적으로 첫번째 단백질 및 두번째 단백질의 상기 두번째 프레임의 융합 단백질이다. 두번째, 많지 않은 단백질이 종결 코돈 없이 두번째 개방 판독 프레임을 갖고 아주 적은 단백질만이 N-말단에 융합된 의미없는 단백질과 생물학적 활성을 보일 것이다. 또한, 상기 기술은, 펼쳐진 융합 단백질의 면역성 가능성 때문에, 치료 상황에서의 용도로 부적합하고, 따라서 상기 구조체는 GFP 및 dsRED의 대체 발현을 위한 조절자로써 및 추가적인 그리고 최적화된 구조체를 위한 기초로써 사용되었다. 상기 DNA 구조체는 GeneArt로부터 주문되었다 (Regensburg, Germany, now Life Technologies).

[0140] GeneArt로부터 동결건조된 플라스미드 DNA는 GeneArt의 규격에 따라 재현탁 되었고 GlnPrl095 및 GlnPrl096 프라이머를 사용하여 PCR 증폭을 위한 주형으로 사용되었다. 이것은 Nhel 자리를 5' 말단에 첨가하였다. 3' 말단의 SacII 제한 자리는 Apal에 의해 대체되었고 추가적인 BstBI 자리는 3' 말단에 첨가되었다. 제한 효소 Nhel 및 BstBI로 상기 조각의 분해는 pGLEX3HM-MCS의 중추로 결찰을 가능하게 하고, 상기 동일한 효소 및 CIPed를 사용하여 개방되어진다. 상기 pGLEX3HM-MCS 벡터는 hCMV 프로모터의 조절하에 발현 카세트를 포함한다. 상기 pGLEX3HM-MCS 중추에 GeneArt 조각과 새로운 벡터는 pGLEX3-ASC으로 불린다.

[0141] EGFP는 GlnPrl097 및 GlnPrl098을 사용하여 pGLEX3으로부터 증폭되었다 (이전에 실내에서 복제된 EGFP(짧게: GFP)를 코딩하는 개방 판독 프레임을 함유하는 벡터는 pEGFP-N1 (Clontech) 플라스미드로부터 유도되었다). 상기 증폭은 GFP의 상기 개방 판독 프레임으로부터 개시 코돈 ATG를 제거시키고 5' 말단에 Apal 자리 및 3' 말단에 BstBI를 첨가시킨다. 제한 효소 Apal, BstBI를 사용하여 증폭물(amplicon)의 분해 및, 상기 동일한 효소로 개방된, pGLEX3-ASC로 결찰은 pGLEX3- ASC-GFP벡터를 야기시킨다.

[0142] 사익 dsRED 개방 판독 프레임은 GlnPrl099 및 GlnPrl 100 프라이머를 사용하여 플라스미드 pdsRED-Express 1 (Clontech)으로부터 증폭되었다. 상기의 프라이머는 개시 코돈 ATG를 5' 말단으로부터 제거하고 5' 말단에 Agel 제한 자리 및 3' 말단에 Apal 자리를 첨가한다. 상기 증폭물은 제한 효소 Agel 및 Apal을 사용하여 분해되고, 상기 동일한 효소 및 CIPed를 사용하여 분해된, pGLEX3-ASC-GFP에 결찰된다. 이것은 pGLEX3-ASC-dsRED-GFP 플라스미드를 발생시켰다. 상기 벡터는 상기 Orengo 등의 문헌으로부터 만들어진 구조체를 포함한다.

[0143] 벡터 pGLEX3-ASC-dsRED-GFP-woFLA Gcorr의 복제

[0144] 대체 접합 구조체의 변경은 PCR을 변경함으로써 수행되었다. 첫번째 PCR은 GlnPrl 142 및 GlnPr991 프라이머 및 pGLEX3-ASC- dsRED-EGFP 주형을 사용하여 수행되었다. 상기 PCR 산물은 Agel 및 BstBI 제한효소를 사용하여 잘려지고 상기 동일한 효소 및 CIPed를 사용하여 개방된 pGLEX-ASC-dsRED-GFP로 복제되어, 중간 구조체 pGLEX-ASC-dsRED-GFP-interm를 야기시킨다. 주형으로써 pGLEX3- ASC-dsRED-EGFP 플라스미드를 사용하고, 두번째 증폭물이 GlnPrl 138 및 GlnPrl 139 프라이머를 사용하여 제조되었고 세번째는 GlnPrl 140 및 GlnPrl 141 프라이머를 사용하였다. 상기 두 증폭물은 이후 GlnPrl 138 및 GlnPrl 141 프라이머를 사용하여 융합 PCR을 위한 주형으로 사용되었다. 이 융합 생성물은 Nhel 및 EcoRI 제한 효소를 사용하여 잘려지고 최종 구조체 pGLEX3-ASC-dsRED-GFP-sep를 제조하기 위해 상기 동일한 효소로 개방 및 CIP 처리된 벡터 pGLEX-ASC-dsRED-GFP-interm로 복제된다. 상기 벡터는 GSD634로 번호가 주어졌다.

[0145] 플래그 태그는 여전히 번역 시작점 (개시 코돈) 으로 사용되는 서열 모티프 ATG를 함유하는 pGLEX3-ASC-dsRED-GFP-sep에 존재한다. 삭제는, GlnPr1 158 및 1139 프라이머 및 주형으로 GSD634 플라스미드를 사용하여, PCR을 변경함으로써 수행되었다. 상기 PCR 생성물은 Nhe1 및 EcoRV 제한 효소를 사용하여 분해되었고, 상기 동일한 효소로 개방되고 뒤이어 재-원형화를 최소화하기 위해 CIP 처리해준 GSD634로 복제되었다. 상기 결과 플라스미드는 배치 번호 GSC2223 (SEQ ID No: 110)로 pGLEX3-ASC-dsRED-GFP- sepwoFLAG로 불린다. 상기 결과 중형 제조의 플라스미드는 GSD679 배치 번호를 부여 받고 GSC2223과 동일한 서열을 갖는다. GFP의 두개의 뉴클리오티드가 표준 GFP 서열과 비교시 다른 것으로 관찰되었다. 이것은 전방 프라이머의 설계때문이다. GlnPr991 및 1180 프라이머 및 주형 pGLEX3을 사용하여, GFP 조각은 정확한 서열로 재-증폭되었다. 상기 조각은 Age1 효소를 사용하여 분해되고. Age1을 사용하여 개방되고 및 이후 CIP한, GSD679의 중추 벡터로 복제되어 벡터 pGLEX3 - ASC-dsRED-GFP-woFLAGcorr를 야기시킨다. pGLEX3 - ASC-dsRED-GFP-woFLAGcorr의 소형 제조는 배치 번호 GSC2246으로 주어졌고 및 소형 제조, 배치 번호 GSC2250 (SEQ ID No: 38)로 주어졌고, 따라서 상기 두개의 구조체는 같은 서열을 갖는다.

[0146] 대체 접합 패턴으로 구조체의 복제

[0147] 상기 GSC2250 구조체는 대체 접합의 다른 비율로 구조체를 제조하기 위해 추가적으로 변경되어, 구조체에 첫번째에서 두번째 개방 판독 프레임으로 발현에서 이동을 야기하였다. 상기 변경은 변경된 프라이머를 사용하여 치킨 트로포닌 인트론 4 또는 5의 증폭으로 도입되었다. 상기 증폭물은 이후 AS 인트론 #3의 위체에 복제하기 위해 Age1 및 EcoRI 제한 효소 및 AS 인트론 #2의 위치에 복제하기 위해 Nhe1 및 EcoRV 제한 효소를 사용하여 유사한 플라스미드 또는 GSC2250의 중추에 재복사 되었다 (도 1 방향 참조). 다음 표 2 및 표 3은 AS 인트론 #2 및 #3, 각각에인트론의 필수적인 복제 단계에 사용되는 프라이머 및 주형을 요약한 것이다. 표 4는 복제된 모든 조합들을 보여준다.

[0148] 표 2: AS 인트론 #2의 변경을 위해 사용된 프라이머 및 주형.

표 2

[0149]

구조체의 이름	전방 사용 프라이머	후방 사용 프라이머	증폭에 사용된 주형
14(22+1)	GlnPr1181	GlnPr1183	GSC2246(소형제조)
14(15Y-5')	GlnPr1181	GlnPr1186	GSC2246(소형제조)
14(15Y-3')	GlnPr1181	GlnPr1185	GSC2246(소형제조)
14(22Y-3)	GlnPr1181	GlnPr1184	GSC2246(소형제조)
14(5Y)	GlnPr1181	GlnPr1182	GSC2246(소형제조)
14(5Y-5)	GlnPr1181	GlnPr1245	GSC2238
14(0Y)	GlnPr1181	GlnPr1246	GSC2238
14(5Ynude)	GlnPr1181	GlnPr1244	GSC2238
14(5Y,b-2)	GlnPr1181	GlnPr1243	GSC2238
14(5Y,b-a)	GlnPr1181	GlnPr1242	GSC2238
14(5Y,b-c)	GlnPr1181	GlnPr1241	GSC2238
14(5Y,b-y)	GlnPr1181	GlnPr1240	GSC2238
14(5Y-G)	GlnPr1181	GlnPr1239	GSC2238
cTNT-I5	GlnPr1237	GlnPr1238	GSC2250

[0150] 표 3: AS 인트론 #3의 변경을 위해 사용된 프라이머 및 주형

표 3

[0151]

구조체의 이름	전방 사용 프라이머	후방 사용 프리이머	증폭에 사용된 주형
I5 (22Y+1)	GlnPr1187	GlnPr1191	GSC2246 상 (소형제조) 증폭물(Amplicon) 1187/1188
I5 (22Y-3)	GlnPr1187	GlnPr1190	GSC2246 상 (소형제조) 증폭물1187/1188

I5 (22Y)	GlnPr1187	GlnPr1189	GSC2246 상 (소형제조) 증폭물1187/1188
I5 (15Y-3')	GlnPr1187	GlnPr1193	GSC2246 상 (소형제조) 증폭물1187/1188
I5 (15Y-5')	GlnPr1187	GlnPr1192	GSC2246 상 (소형제조) 증폭물1187/1188
I4 (sh)	GlnPr1285	GlnPr991	GSC2741

[0152] GFP 및 dsRED를 사용한 일시적인 대체 접합 구조체의 검사

[0153] 각기 다른 구조체는 표 4에 열거되 있는 조합으로 복제되었고, 중형으로 제조되고 철저하게 서열에 의해 확인되었다 (Fasteris, Plan-les-Ouates, Switzerland). 모든 도입된 변형의 정렬은 도 2에 나타내었다. 상기 플라스미드는 CHO-S 세포 및 HEK293 세포에 감염되었다. 양성 대조군으로, dsRED (GSD636는, dsRED 유전자를 발현한 pGLEX3에 기초한 자체 벡터, pDsRED-Express 1 (Clontech)으로부터 유도되었다) 및 GFP (pEGFP-N1, Clontech) 만을 발현한 벡터가 숙주 세포로 각각 감염되었다. 분석은, 적절한 필터를 사용하여 형광 현미경에 의해 지원되는, 유동 세포 계측법에 의해 이루어졌다.

[0154] 상기 감염은 HEK293 및 CHO-S 세포를 사용하여 재료 및 방법 파트에 설명된 바와 같이 12 웰 플레이트 규모에서 수행되었다.

[0155] 상기 감염이 견고할 지라도, 감염 효율에서 변화는 각각의 구조체의 절대적인 발현 수준에 결론을 가늠할 수 없다.

[0156] 표 4: 상기 첫번째 엑손 (dsRED 발현)으로부터 상기 두번째 엑손 (GFP 발현)으로 이동시키기 위해 사용된 구조체의 목록. 자체 플라스미드 배치 번호 및 SEQ ID 열거에 의해 가능한 복제를 예상하였다. 상기 SEQ ID는, 상기 첫번째 엑손의 뉴클리오티드로부터 SV 40 폴리(A) 자리의 말단까지, 전체 mRNA을 포함한다.

표 4

[0157]

			대체 엑손의 아래줄기에 사용된 인트론 구조체 (AS 인트론 #3 위치)						
		구조체의 이름	cTNT-I5	I5 (22Y+1)	I5 (22Y-3)	I5 (22Y)	I5 (15Y-3')	cTNT-I4	I4 (sh)

대체 엑손의 윗줄기에 사용된 인트론 구조체 (AS 인트론 #2 위치)	폴리(Y) 관 변경	cTNT-I4	GSC2250 SeqID38	GSC2329 SeqID39		GSC2330 SeqID40	GSC2323 SeqID41	GSC2619 SeqID42	GSC2781 SeqID43
		I4 (22Y+1)	GSC2342 SeqID44		GSC2328 SeqID45	GSC2321 SeqID46	GSC2324 SeqID47		
		I4 (15Y-5')	GSC2339 SeqID48			GSC2334 SeqID49	GSC2336 SeqID50		
		I4 (15Y-3')	GSC2340 SeqID51	GSC2331SeqID52	GSC2453 SeqID53	GSC2325 SeqID54	GSC2332 SeqID55		
		I4 (22Y-3)	GSC2341 SeqID56		GSC2326 SeqID57	GSC2454 SeqID58	GSC2327 SeqID59		
		I4 (5Y)	GSC2338 SeqID60	GSC2335 SeqID61	GSC2333 SeqID62	GSC2337 SeqID63	GSC2322 SeqID64		
		I4 (5Y-5)				GSC2617 SeqID65		GSC2739 SeqID66	GSC2782 SeqID67
		I4 (0Y)				GSC2621 SeqID68		GSC2740 SeqID69	GSC2783 SeqID70
		I4 (5Y누드)				GSC2622 SeqID71		GSC2742 SeqID72	GSC2784 SeqID73
	분기점 변이	I4 (5Y,b-2)				GSC2620 SeqID74		GSC2737 SeqID75	
		I4 (5Y,b-a)						GSC2743 SeqID77	
		I4 (5Y,b-ct)				GSC2615 SeqID76		GSC2738 SeqID78	
		I4 (5Y,b-y)				GSC2618 SeqID79		GSC2975 SeqID80	
	인트론-엑손 컨센서스	I4 (5Y,G)				GSC2613 SeqID81			
	인트론 스위치	cTNT-I5	GSC2614 SeqID82					GSC2741 SeqID83	GSC2780 SeqID84

[0158] 폴리(Y) 관에서 변경으로 구조체의 발현

[0159] 기본 구조체 GSC2250은 AS 인트론 # 2로 변경되지 않은 cTNT-I4 서열 및 AS 인트론 #3로 변경도지 않은 cTNT-I5 서열에 의해 측면에 위치한 dsRED의 개방 판독 프레임을 코딩한 대체 엑손을 함유하고, 뒤이어 GFP (짧은 cTNT-I4|cTNT-I5의 배향)의 개방 판독 프레임을 코딩하는 엑손이 온다. 감염된 CHO-S 또는 HEK293 세포에서, 상기 구조체는 dsRED 및 GFP의 발현을 보여준다 (도 3 참조). 이것은 상기 구조체가 대체 접함을 야기한다는 것을 확인시켜 준다. 그럼에도, dsRED 발현은 크게 GFP 발현보다 선호되어진다 (도 3a 및 b 참조). dsRED를 코딩하는 상기 대체 엑손의 스플라이스 수용 자리는 GFP를 코딩하는 엑손의 상기 두번째 스플라이스 수용자리와 경합한다. 인트론-엑손 경계(일명 폴리(Y) 관) 및 분기점 사이의 Y(피리미딘 염기 C 또는 T)의 충족은 스플라이스 수용 자리의 강도에 중요하다 (예를들어, Dominiski & Kole (1992) Mol Cell Biol 12(5): 2108-14 참조). Y의 양을 감소함으로써 상기 스플라이스 수용 강도의 감소는 dsRED를 코딩하는 대체 엑손의 바람직한 배제를 야기할 것으로 기대되고 따라서 결론적으로 GFP의 더 많은 발현을 할것이다.

[0160] AS 인트론 #2 위치에 cTNT- I4의 폴리(Y) 관에서 (도 2a 배열 참조) Y의 양을 감소함으로(cTNT-I4 기초 구조체의 변경된 버전에서 28으로 부터 0으로) 각기 다른 구조체는 CHO-S 및 HEK293 세포에서 감염되었다. 3-6일 경과 후 상기 세포는 유동 세포 계측법을 사용하여 분석되었다. 폴리(Y) 관에서 Y의 양의 감소는 dsRED 및 GFP에 대해 양의 두배인 세포의 군집에서 완만한 증가를 야기시킨다 (도 3 참조). GFP의 가장 높은 상대 비율을 발현하는 구조체는 변화되지 않은 cTNT-I4 (27 Y)와 비교시 폴리(Y) 관에 상당히 적은 Y (0 내지 5 사이)를 함유하는 구조체 14 (0Y), 14 (5Y-5) 및 14 (5Ynude)이었다. 이것은 AS 인트론 #2 위치에 스플라이스 수용의 강도에서 감소가 GS 엑손 #3 (dsRED를 코딩하는)의 배제를 야기하고 따라서 GS 엑손 #4 (GFP를 코딩하는)로부터 더 놓은

발현을 한다.

[0161] 상기 초기 구조체의 발현으로부터, 새로운 구조체의 기본 발현 수준은 dsRED 발현에서 더욱 선호된다는 점은 명백하다. 치킨 트로포닌 대체 엑손에 대해 상기 엑손의 크기는 상기 대체 접합 건의 주요 요소라는 것은 설명되었다. Xu et al., 1993 (Mol Cell Biol, 13(6): 3660-74)은 49 뉴클리오티드보다 더 작은 인위적인 엑손은 만약 상기 엑손들이 (발명의 구조체에서 존재하지 않는) 스플라이스 강화 요소가 결핍된다면 스플라이스 장치에 의해 인식되지 못한다는 것을 설명한다. 반면 앞선 저자들은 49 내지 119 뉴클리오티드 크기의 엑손은 선택적으로 접합된다는 것을 보였다. dsRED와 엑손은 718 뉴클리오티드의 크기를 갖고 (상기 Xu et al에 의해 분석된 최대 엑손 크기의 6배) 주로 포함되어 있다. 따라서 상기 첫번째 엑손의 발현에 대한 이동은 간단하게 상기 엑손의 크기 때문일 것이다.

[0162] 폴리(Y)에서 변경에 의해 dsRED로부터 GFP로 발현에서 이동의 변화는 (예를들어, Fallot et al, 2009 (Nucleic Acids Res, 37(20):e134)에 개신된 것과 비교시) 상기 문헌에 개시된 데이터와 비교시 실망스럽다. 명백하게 대체 접합은 단순히 대체 엑손의 윗줄기 인트론의 폴리(Y) 함량을 감소시킴으로써 제조될수 없었다.

[0163] 대체 엑손 (AS 인트론#3)의 아랫줄기에 복제된, 상기 인트론 cTNT-I5는 오직 10 Y를 함유하는 더욱 감소된 폴리(Y) 관을 갖는다. AS 인트론 #2에 (스플라이스 수용 강도의 약화를 야기하는) Y의 수의 감소는 GFP 발현으로 이동을 선호하기 때문에, AS 인트론#3에서 Y의 함량에서 증가는 스플라이스 수용 강화를 야기할 것이고 따라서 dsRED로부터 GFP 발현으로 이동을 야기할 것이라는 것은 생각되었다. 28 Y까지 함유하는 변경된 cTNT-I5 인트론 서열은 (원본 구조체에서 존재했던 10과 비교시) AS 인트론#3 위치에 복제되었다 (도 2b 서열 참조). 그럼에도 GFP 발현에 큰 이동은 관찰되지 않았다 (도 3). 따라서 원본 cTNT-I5 서열은 분기점 및 인트론-엑손 컨센서스 부위의 변경의 효과를 분석하기 위해 사용되었다.

[0164] 인트론-엑손 경계에서 및 분기점에서 변경과 구조체의 감염

[0165] GFP 발현의 선호 스플라이스 비율을 더 이동시키기 위해, 서열 변경이, 대체 엑손의 윗줄기 (도 1a에 엑손 #3) AS 인트론 #2의 분기점에서 및 인트론-엑손 컨센서스 부위에서 도입되다. 상기의 변경은 스플라이스 수용 영역의 강도를 더 감소시키는 것으로 생각되었다. 도입된 수정의 세부사항은 도 2b에 배열에 나타내었다. 상기 변경 중 어느 것도 dsRED에서 GFP 발현으로 상당한 이동을 야기하지 못했다 (도 4, 윗열). 상기의 변경이 대체 접합에 지대한 영향을 보인 것은 놀라운 것이었다 (예를 들어 Fallot 등의 문헌).

[0166] 게대가, 상기 인트론 cTNT-I4 및 cTNT-I5는 다른 방법으로 재배열되었다. 첫째로, 인트론 cTNT-I4 및 cTNT-I-5가 교환되었고, dsRED를 발현하는 대체 엑손은 AS 인트론 #2의 cTNT-I5 및 AS 인트론 #3 위치에 cTNT-I4의 측면에 위치되었다. 이후, 서열 cTNT-I4는 AS 인트론#2 및 #3으로 사용되었다. 인트론 서열 cTNT-I5를 사용하요 동일하게 수행하였다. 두개의 동일한 인트론으로 대체 엑손을 측면에 위치시키는 것은 상당하게 양의 두배 (dsRED 및 GFP) 집단을 증가시켰다. HEK293 및 CHO-S에서 최고 구조체는 (GSC2614; cTNT-I5|cTNT-I5) 상당하게 두배 양의 집단을 증가시켰다 (도 4, 중간열 참조). cTNT-I4|cTNT-I4 방향을 가지는, GSC2619 구조체는 HEK293 및 CHO-S 세포에서 상당한 두배 양성의 양의 세포 증가를 보였고 추가 구조체로 사용되었다. 대체 엑손을 측면에 위치하는 인트론의 유서성이 접합 비율에 대한 영향을 가진다는 것을 제시한 어떠한 문헌도 없었기 때문에, 이것은 상당히 놀라운 것이다. 그럼에도 우리의 데이터가 엑손을 측면에 위치시키는 두개의 동일한 인트론이 엑손의 대체 접합을 야기한다는 것을 지지한다. 이것은 치킨 트로포닌 인트론 4, 치킨 트로포닌 인트론 5 및 사람 EF1 알파 유전자의 첫번째 인트론을 구조적으로 자른 것에서 나타났다 (실시예3 참조). 이전의 실험에서 상당하지만, 미량의 GFP로 이동, cTNT-I4₩cTNT-I4의 분기점 및 폴리(Y) 조합의 변경은 대체 엑손 (cTNT-I4|cTNT-I4 또는 cTNT-I5|cTNT-I5 배향)을 측면에 위치시키는 동일한 인트론을 가지는 구조체의 및 폴리(Y) 관에서 Y의 감소된 함량으로 구조체에 대해 관찰될 수 있다. 상기 변경들을 결합이 GFP의 발현으로 추가의 이동을 야기하였는지 분석하기 위해, AS 인트론#2의 분기점 및 폴리(A) 관의 변경이 대체 엑손 (cTNT-I4|cTNT-I4 배향)의 위 및 아래 줄기에 cTNT-I4 인트론을 함유하는 구조체 GSC2619에 도입되었다. 상기 실험에 대해 GFP 발현에 가장 높은 이동을 보인 폴리(Y) 변경이 사용되었다 (14(5 Y- 5), I4(0Y), I4(5Ynude)). 상기 구조체 GSC2250 (cTNT-I4|cTNT-I5)은 기초 구조체의 접합 비율에대한 기준으로 포함되었다. 폴리(Y) 관의 감소 및 cTNT-I4|cTNT-I4 구성의 조합은 HEK293 및 CHO-S 세포에서 세개의 구조체에 대해 GFP 발현으로 상당한 이동을 보였다 (도 5a 중간열 및 도 5b 상위열). 흥미롭게, 동일한 인트론 (여기서 cTNT-I4)의 사용 및 폴리(Y) 관의 수반된 감소의 조합은 상기 두번째 개방 판독 프레임으로 접합 비율의 이동에 상승효과가 있었다. 반면, 분기점에서 변경 및 I4(5Y)|cTNT-I4 구조체를 사용한 폴리(Y) 관의 감소의 조합은 dsRED에서 GFP로 상당한 이동을 보이지 못했다 (도 5a 상위열).

[0167] 스플라이스 공여 자리의 제거

[0168] dsRED를 발현하는 첫번째 엑손으로부터 GFP를 발현하는 두번째 엑손으로 한층 더 접합 비율을 이동시키기 위해서, AS 인트론 #3 위치에 cTNT-I4의 스플라이스 공여 자리는 제거되었다 (도 2c 배열 참조). 이것은 AS 인트론 #3의 스플라이스 수용 영역 (분기점, 폴리(Y) 및 인트론-엑손 컨센서스는 변경되지 않았다)의 전체 인트론 윗줄기 (5') 및 엑손-인트론 컨센서스 부위를 제거함으로써 수행되었다. 상기 스플라이스 공여의 제거는 dsRED 발현에서 GFP 발현으로 이동을 더욱 증가시켰다. 폴리(Y) 관에 Y의 감소와 조합에서, 이것은 거의 GFP 발현의 우세를 야기한다 (도 6).

[0169] GFP-dsRED 발현 실험의 요약

[0170] 대체 접합 구조체의 다른 설계가 인트론의 측면에 위치한 cTNT 대체 엑손 5에 기초하여 실험되었다. 기초 구조체(cTNT-I4|cTNT-I5)는 대체 엑손의 포함에 대한 선호를 나타내고 주로 dsRED를 발현시켰고, 보고 단백질이 첫번째 개방 판독 프레임에서 발현 되었다. 대체 엑손의 크기는 대체 엑손의 포함 (더 큰 엑손의 경우) 또는 배제 (작은 엑손일 경우)에 주요한 영향을 준다는 것은 문헌에서 보여져 왔다. 폴리(Y) 관에 Y의 양의 감소 및 대체 엑손의 위 및 아래줄기에 동일 인트론, 특히 cTNT-I4의 사용은 dsRED 발현 (대체 엑손 상)에서 GFP의 발현으로 (두번째 개방 판독 프레임에서 발현되는) 상당한 이동을 야기시키는 것을 보였다. 상기 이동은 폴리(Y) 관 감소 및 대체 엑손의 위 및 아래줄기 cTNT-I4의 결합함으로써 더욱 증가될 수 있다. 이것은. 현재 문헌에서 엑손의 위 및 아래줄기에 동일한 인트론 서열의 사용이 측면배치된 엑손의 배제로 이동을 야기한다는 것을 제시하지 않던 것이어서, 놀라운 발견이다. 게다가 더 놀랍게도, 이것의 효과가 EF1 알파 첫번째 인트론을 사용해서 확인될 수 있다. 상기 인트론은 대게 대체 접합의 대상이 아니다. 이것은 대체 접합을 야기하는 일반적인 메카니즘을 설명한다. 최종적으로, 대체 엑손 (AS 인트론 #3)의 아래줄기 스플라이스 도너 자리의 제거가 대체 엑손의 추가적인 배제를 야기한다. 상기 세 구조체로 감염된 세포는 GFP를 주로 발현하는 것으로 보였다. 최종 대체 접합 구조체는 대체 접합의 양 극단 (주로 우세한 GFP 발현을 야기하는 대체 엑손의 배제에 우세한 dsRED 발현을 야기하는 대체 엑손의 포함)뿐 아니라 중간체 비율을 밝혀냈다 (도 7 계략도 참조).

[0171] 상기에서 언급한 바와 같이, 단백질당 형광 신호, 사용된 두 보고 단백질의 생산 효율 및 수준의 감지가 상당히 다르다는 것을 완전히 배제할 수는 없다. 그럼에도, 상기에서 밝혀진 세 조건 (대체 엑손의 앞과 뒤에 동일 인트론의 사용, 폴리(Y) 관의 Y의 양 감소, 스플라이스 공여 자리의 제고)은 대체 접합을 사용하여 발현된 다른 단백질에 대해서도 유효하다.

[0172] 표 5 : 구조체의 목록

표 5

Name of plasmid	SEQ ID No.		Name of plasmid	SEQ ID No.
GSC 2250	38		GSC 2333	62
GSC 2329	39		GSC 2337	63
GSC 2330	40		GSC 2322	64
GSC 2323	41		GSC 2617	65
GSC 2619	42		GSC 2739	66
GSC 2781	43		GSC 2782	67
GSC 2342	44		GSC 2621	68
GSC 2328	45		GSC 2740	69
GSC 2321	46		GSC 2783	70
GSC 2324	47		GSC 2622	71
GSC 2339	48		GSC 2742	72
GSC 2334	49		GSC 2784	73
GSC 2336	50		GSC 2620	74
GSC 2340	51		GSC 2737	75
GSC 2331	52		GSC 2615	76
GSC 2453	53		GSC 2743	77
GSC 2325	54		GSC 2738	78
GSC 2332	55		GSC 2618	79
GSC 2341	56		GSC 2975	80
GSC 2326	57		GSC 2613	81
GSC 2454	58		GSC 2614	82
GSC 2327	59		GSC 2741	83
GSC 2338	60		GSC 2780	84
GSC 2335	61			

[0173]

[0174] 실시예 2: dsRED 및 GFP를 발현하는 안정한 세포

[0175] 재료 및 방법

[0176] 실시예 2의 재료 및 방법은 상기 실시예 1의 재료 및 방법에서 설명한 바와 동일하다.

[0177] 결과

[0178] 발현 구조체의 복제

[0179] GFP 및 dsRED의 발현을 야기하는 전-mRNA의 대체 접합을 위한 다른 구조체는 실시예 1에서 설명하였다. 상기 구조체의 하나는 안정한 CHO 세포주의 개발을 위하여 선택되었다. pGLEX3 벡터 중추는 HEK293 세포에서 일시적인 발현에 가장 적합하기 때문에, 선택된 GSC 2739 구조체의 대체 접합 카세트는 전용 pGLEX41 발현 벡터에 삽입되었다 (배치 번호 GSC281). 상기 벡터에서 상기 대체 접합 카세트는 CHO 세포에서 안정한 발현에 잘 부합하는 mCMV 프로모터에 의해 유도된다. 상기 발현 카세트는 Nhel 및 BstBI 효소를 사용하여 잘려지고 동일 효소로 개방되고 CIP 처리된 pGLEX41의 중추에 복제된다. 상기 결과 벡터는 pGLEX41-ASC-cTNT-I4(5Y-5)|cTNT-I4- dsRED-GFP로 불리고 배치 번호 GSC3166 (SEQ ID NO: 111)을 부여받았다. 항생물질 퓨로마이신에 대한 저항 유전자를 부여 받은 상기 벡터는 pSEL3, pGL3(Promega, Madison, WI) 유도벡터였다. 상기 벡터의 퓨로마이신 저항은 SV40 프로모터의 조절 하에 있다.

[0180] 안정한 감염

[0181] 일반적인 세포 배양 및 CHO-S의 감염은 실시예 1에서 설명하였다. 안정한 세포주를 아기하는 이번 감염에 사용된 DNA 칵테일은 95%의 pGLEX41 및 5%의 pSEL3의 혼합물 이었다(몰비). 감염 후, 상기 세포는 하루 동안 궤도 교반기에 배양되었다. 다음날, 상기 세포는 선택 압력하에 다른 희석으로 96 웰 플레이트에 주입되었다. 선별에 사용된 퓨로마이신의 농도는, 다른 안정한 통합의 혼합이 될 수 있기 때문에, 오히려 복제군 보다 "미니풀"이라 하는 안정한 개체군을 산출한다. 일주일 후 선택 압력은 재생해 주었다. 미니풀을 함유하는 웰의 검사는 엘리사 플레이트 판독기를 사용하여 2주 후에 수행되었다. 높은 형광신호를 보이는 세포는 24 웰 플레이트 규모로 확장하고 FACS로 분석하였다. 복제군을 얻기위해, 하나의 미니풀이 제한 희석의 두번째 회차를 하기위해 선택되었다. 이것을 위해 다른 농도로 세포를 희석하고 96웰 플레이트에 주입하였다. 복제군은 플레이트에 자라난 복제의 양 및 웰의 다중 성장 중심의 부재에 기초하여 선별되고 확장되었다. 24 웰로 확장 후, 상기 복제군의 dsRED 및 GFP 발현이 FACS에 의해 평가되었다. 제한 희석 2후 제조된 복제의 GFP 및 dsRED의 상대적인 발현 수준의 비교는 대부분의 복제에 대한 GFP 발현에 dsRED의 아주 유사한 비율을 보였고, 전체적인 발현 수준은 다른 복제 간에 다양할지라도, 모든 복사가 GFP 및 dsRED에대해 양의 두배로 나타났다. 오직 GFP 또는 dsRED만을 발현한 복제는 관찰되지 않았다. 도 8은 무작위로 선택된 8개의 대표적인 복제의 GFP 및 dsRED 발현을 보여준다.

[0182] 동일한 모 미니풀로부터 유도된 각 다른 복제의 상기 유사 접합 비율은 상기 접합 비율이 두 엑손 중 하나로 이동 없이, 다중 발생에 대해 안정하게 유지되었다는 것을 보여주었다. 이것은, 모든 복제가 대체 엑손 (deRED 발현에 대한 GFP의 비율에서 미세한 차이를 야기하는)에 대해 약간씩 다른 접합 비율을 가질지라도, 상기 대체 접합 비율이 대부분 DNA 구조체에 의해 정의된다는 것을 예상할 수 있게 한다. 또한 이것은 재조합 단백질의 발현에 대한 대체 접합의 사용에 대해 강한 선택 압력이 없고, 반면 많은 복제들이 발현을 잃을 것이란 것을 예상할 수 있게 한다.

[0183] 요약해서, 이 실시예에서 복제군은 본 발명의 상기 대체 접합 구조체가 선택 압력의 사용 없이 다중 발생에 대해 변화없는 비율에서 안정한 발현을 가능케 한다는 것을 보여준다.

[0184] 실시예 3: 항체의 일과성 발현

[0185] 재료 및 방법

[0186] 구조체의 복제

[0187] 항-HER2 항체가 보고 구조체의 제조에 사용되었다. 상기 항-HER2 항체의 중쇄 및 경쇄는 CHO 세포에서 발현하기 위한 최적화-코돈이다. 유전자는 상기 실시예 1에서 설명된 멕터의 dsRED 및 GFP의 위치에 가능한 두개의 조합으로 복제되었다. 선택된 구조체는 추가적인 분석을 위해 pGLEX41 플라스미드에서 복제되었다. 상기 벡터에서

대체 접합 구조체의 발현은 마우스 CMV 프로모터에 의해 조절되었다.

[0188] 분비된 항-HER2 항체의 세포의 감염 및 정량

[0189] 상기 구조체는 상기 실시예 1 내지 2에서 설명된 바와 같이 50 ml의 생물 반응기 포멧 또는 24 웰 포멧에서 CHO-S 세포 및 HEK293 세포에서 감염되었다. 감염 후 상기 세포는 37℃, 5% CO_2 및 80% 습도에서 진탕 플렛폼상에서 배양되었다. 상기 분비된 항체는 감염 3 내지 6일 후 제조사의 규격에 따라 Protein A bioprobe로 옥텟 QK 시스템 (Fortebio)을 사용하여 정량화 되었다. 검정 곡선은 상기 정제된 항-HER2 항체를 사용하여 완료되었다.

[0190] 대체 접합 구조체를 사용한 항-HER2의 일시적인 발현

[0191] 상기 항-HER2 항체는 대체 접합을 사용하여 항체의 발현을 위한 견본 단백질로 사용되었다. 상기 항체는 잘 별현되었고 생산 단계 동안에 배양 상청액에서 안정하였다. 중쇄가 경쇄 보다 두배의 몰 과량으로 감염되었다면, 상기 항-HER2 항체는 더욱 잘 발현된다는 것을 이전의 보조-감염 실험에서 보여주었다. 상기 비율은 각각의 항체에 의존적인 것으로 나타났다. 따라서 본 연구에 있어 최고의 구조체는 문제의 항-HER2 항체만의 높은 발현을 보여준다. 다른 항체들은 경쇄에 대한 중쇄의 다른 최적 비율을 가질 것이고 다른 접합 구조체를 필요로 할 것이다.

[0192] 항-HER2 항체 중쇄 및 경쇄를 코딩한 상기 개방 판독 프레임은 상기 실시예 1의 dsRED 및 GFP 두개의 형광 마커의 위치에서 두개의 다른 배양으로 복제되었다

[0193] (배향 1 : 첫번째 경쇄, 이후 중쇄;

[0194] 배향 2 : 첫번째 중쇄, 이후 경쇄)

[0195] 상기 실시예 1에서 설명된 바와 같이, 상기 첫번째 인트론 (AS 인트론 #1)은 모든 구조체에 존재하는 구조적으로 접합된 인트론 서열이다. 상기 두번째 인트론 (AS 인트론 #2)은 대체 엑손의 윗줄기에 위치해 있고, 이것은 상기 두개의 개방 판독 프레임의 첫번째를 함유한다. 상기 세번째 인트론 (AS 인트론 #3)은 상기 대체 엑손의 아래줄기에 있다. 상기 인트론은 두번째 개방 판독 프레임을 함유하는 엑손의 윗줄기에 있다. 접합 건에 의존하여 최종 성숙 mRNA는 대체 엑손에 개방 판독 프레임 1 또는 개방 판독 프레임 2를 코딩할 것이다 (대체 접합 건의 도 1a 계략도 참조).

[0196] 다양한 양의 폴리(Y)와 발현 구조체는 상기 첫번째 (dsRED)에서 상기 두번째 개방 판독 프레임 (GFP)으로 발현에서 이동 및 절대 발현 수준에 기초한 dsRED 및 GFP (도 1 참조)를 사용한 예비 연구로부터 선택되었다. 상기의 구조체들은 상기 두번째 개방 판독 프레임의 효율적인 발현을 야기하는 것으로 나타난 전체 길이 AS 인트론 #3 또는 단축 형태("sh")로 결합 되었다. 구조체가 상기 항-HER2 항체의 발현 수준의 영향을 줄 수 있는 GFP에 대한 dsRED 비율에서 미세한 이동만이 보이는지 확인하기 위해, 분명한 효과를 보이지 않는 구조체의 몇몇이 (분기점 변경 및 인트론-엑손 컨센서스 부위 변경) 보고 단백질로써 항-HER2 항체를 사용하여 재평가 되었고 폴리(Y) 관의 영향이 세부적으로 분석되었다 (표 6 모든 구조체 및 서열 정보에 대한 도 9 정렬).

[0197] 항체의 발현을 하기 위해, 중쇄 및 경쇄 둘 모두 상당한 수준으로 발현되어야만 하고, 항-HER2 항체에 대해, 두배 과량의 HC 발현이 일시적인 감염에서 항체 분산을 하기 위해 바람직하다는 것이 나타났다. 폴리(Y) 관에 Y의 각기 다른 양으로 구조체가 복제되었고 CHO-S 세포에서 감염되었다. 여섯째 날 상층액에서 축적된 항-HER2 항체의 양은 옥텟으로 정량화 되었다.

[0198] LC-HC 배향 및 HC-LC 배향으로 구조체의 발현 수준은 도 10에 나타내었다. 전체적인 발현 수준은, 대체 (첫번째) 엑손상 경쇄 및 완전한 길이의 두번째 인트론과, LC-HC 배향에서 가장 높았다. 제조된 역가는 경쇄에 대한 중쇄의 최적 비율을 사용하여 보조-감염 조절의 60%까지였다. 이것은 항체의 발현에 대한 대체 접합의 가능성을 보여주었다.

[0199] 모든 구조체의 발현 수준은 폴리(Y) 관에서 Y의 양 감소를 수반하여 증가하였다 (HC-LC 방향으로 1414 시리즈의 발현과). 상기 첫번째 인트론에서 더 적은 Y는 상기 접합 비율은 우세하게 발현된 첫번째 엑손으로부터 상기 두번째 대체 엑손으로 이동 시키고 이에 보다 높은 상대적인 개방 판독 프레임의 발현이 상기 두번째 대체 엑손에 존재한다. 상기 항체는 성공적인 조립 및 분비를 하기 위해 중쇄 및 경쇄의 발현이 필요하기 때문에, 이것은 전체 항체의 발현에 유리하다. 상기 발현 수준은 폴리(Y) 관이 7 또는 그 이하의 Y를 가지면 크게 증가하기 시작하는 것이 관찰 되었다. 이것은 대체 접합이 두개의 대체 엑손의 약 동등한 몰의 발현으로 이동되었을 때이다 (효과가 양 배향에서 I4I4sh 구조체에 대해 관찰되었기 때문에). 놀랍게도, 상기 AS 인트론 #3의 단축은 최대의

발현을 야기하는 폴리(Y) 관에 Y의 양에 대한 효과가 거의 없다. 이것은, 상대적인 넓이 범위의 HC:LC 비율이 가능한 보고 시스템의 둔감때문일 것이다.

[0200] 표 6: 항-HER2 항체 발현을 위해 만들어진 pGLEX3에 기초한 구조체의 목록. SEQ ID Nos: 85 내지 102는 첫번째 개방 판독 프레임의 개시 코돈 (ATG)까지 mRNA의 첫번째 엑손을 포함한다. SEQ ID 103 내지 108 상기 첫번째 개방 판독 프레임의 종결 코돈과 시작되고 상기 두번째 개방 판독 프레임의 개시 코돈으로 종결된다.

표 6

[0201]

	Ys in construct	LC-HC			HC-LC		
		cTNT-I4 SEQ ID No: 103	cTNT-I5 SEQ ID No: 104	I4(sh) SEQ ID No: 105	cTNT-I4 SEQ ID No: 106	cTNT-I5 SEQ ID No: 107	I4(sh) SEQ ID No: 108
Ys in construct	-						
cTNT-I4	27	GSC2821	GSC2822	GSC3164	GSC2816	GSC2819	GSC3170
I4 (5Y) SEQ ID No: 085	10	GSC4218		GSC4228	GSC4222		GSC4225
I4 (9Ynude) SEQ ID No: 086	9	GSC4344		GSC4339	GSC4335		GSC4336
I4 (7Ynude) SEQ ID No: 087	7	GSC4345		GSC4355	GSC4337		GSC4341
I4 (5Y-5) SEQ ID No: 088	5	GSC2820		GSC4226	GSC4217		GSC4221
I4 (5Ynude) SEQ ID No: 089	5	GSC4220		GSC4215	GSC2823		GSC4223
I4 (3Ynude) SEQ ID No: 090	3	GSC4340		GSC4354			GSC4333
I4 (1Ynude) SEQ ID No: 091	1	GSC4332		GSC4407	GSC4331		GSC4405
I4 (0Y) SEQ ID No: 092	0	GSC2818		GSC4224	GSC3151		GSC4214
I4(5Y, b-ct) SEQ ID No: 093		GSC2977			GSC3154		
I4 (5Y, b-y) SEQ ID No: 094		GSC3182					
I4(5Y, b-2) SEQ ID No: 095		GSC2985		GSC3155	GSC2984		GSC3147
I4(5Y, b-a) SEQ ID No: 096		GSC2986					
I4(5Y-A) SEQ ID No: 097		GSC2976			GSC3158		
I4(5Y-5, G) SEQ ID No: 098		GSC3085					
I4 (5Ynude, A) SEQ ID No: 099		GSC3089					
I4(5Ynude, b-2) SEQ ID No: 100		GSC3184					
I4(5Ynude, A) SEQ ID No: 101					GSC3153		
I4(5Y-5, G) SEQ ID No: 102					GSC3160		

[0202] LC-HC 배향 구조체에 대해, 상기 구조체 3Y누드 및 1Y누드는 폴리(Y) 관에 적은 (0Y) 또는 Y (5Y누드)의 구조체와 비교하여 적은 발현을 보인다. 이것은 서열에서 작은 변화는 접합 비율에 영향을 주고 상기 폴리(Y) 관에 Y의 수 및 엑손의 크기가 접합 효율에 영향을주는 유일한 요소들이 아님을 보이고 있다.

[0203] 이와 반대로, HC-LC 배향의 상기 1414-구조체는 폴리(Y) 함량에 독립적으로 상대적인 높은 발현 수준을 보인다. 대체 엑손의 길이를 증가시키는 것은 대체 (첫번째) 엑손에 대한 (따라서 개방 판독 프레임 1) 접합 비율을 이동시킨다는 것은 문헌에 개시되어 있었다. 단축 AS 인트론 #3을 사용하여, 상기 폴리(Y) 함량은 실험된 항-HER2 항체의 발현, 따라서 상기 접합 비율에 영향을 준다. 상기의 실험 결과의 하나의 설명은 상기 첫번째 위치에 중쇄의 개방 판독 프레임을 코딩하는 큰 엑손이 접합 비율에 대한 폴리(Y) 관의 영향을 약화시킨 다는 것이고, 두 접합 변형의 고정된 비율을 야기한다. 오직 접합 건이 상기 두번째 인트론을 단축시키는 것 및 상기 두번째 인트론의 스플라이스 공여의 제거에 의해 더욱 안정화되지 않을때, 상기 폴리(Y) 관은 접합 비율에 영향을 줄 것이다.

[0204] 상기에서 설명된 검출에서, 5Y-5, 5 Y누드 및 0Y 구조체는 LC-HC 배향에 대해 가장 높은 일과성 발현 결과를 주는 구조체로 식별되었다. 상기 발현 구조체는 안정한 세포주 개발을 위해 사용되는 발현 벡터로 복제 되었다. 전-접합 R A 구조체는 바뀌지 않는 것으로 유지도기 때문에 (프로모터만이 변한다) 이 복제 단계는 접합 비율에 상당한 차이를 야기할 것으로 기대되지 않는다. 보고 단백질로 GFP 및 dsRED를 사용하여, 인트론-엑손 컨센서스 변경 또는 분기점 변경의 효과가 관찰될 수 없었다 (도 1 참조). 그러나, 접합 비율에서 사소한 변화는

GFP/dsRED 보고 시스템을 사용하여 관찰되기 어려울 것이다. 인트론-엑손 변경 또는 분기점 변경이 항체 발현을 위한 메세 접합 비율 조절로 유용한지 확인하기위해, 새로운 구조체가 pGLEX41에서 5Y-5, 5Y누드 및 0Y 구조체에 기초하여 복제되었다 (표 7 구조체의 완전한 목록 및 도 11 0Y 구조체의 발현 결과 참조).

[0205] 표 7 : pGLEX41 최종 벡터에 경쇄 발현에 대한 중쇄의 미세 조정을 사용한 구조체의 목록. 하기 열거된 SEQ ID Nos: 88, 89, 92, 99, 100, 102 및 112 내지 128은, 상기 첫번째 개방 판독 프레임의 개시 코돈 (ATG)까지 mRNA의 첫번째 엑손을 포함한다. SEQ ID No: 103 상가 첫번째 개방 판독 프레임의 종결코돈과 개시되고 상기 두번째 개방 판독 프레임의 개시 코돈과 종결된다.

표 7

	Intron constructs used downstream of the alternate exon (position AS intron #3) cTNT-I4 (SEQ ID No: 103)	
Intron constructs used upstream of the alternate exon (position AS intron #2)	GS number	
	LC-HC	HC-LC
I4(0Y) SEQ ID No: 92	GSC3157	GSC3151 GSC4219
I4(0Y, b-a) SEQ ID No: 112	GSC3436	GSC3466
I4(0Y, b-ct) SEQ ID No: 113	GSC3432	GSC3470
I4(0Y, b-y) SEQ ID No: 114	GSC3439	GSC3465
I4(0Y, b-2) SEQ ID No: 115	GSC3462	GSC3465
I4(0Y, A) SEQ ID No: 116	GSC3447	GSC3442
I4(0Y, T) SEQ ID No: 117	GSC3453	GSC3430
I4(0Y, G) SEQ ID No: 118	GSC3434	GSC3446
I4(5Ynude) SEQ ID No: 89	GSC3162	GSC3169
I4(5Ynude, b-a) SEQ ID No: 119	GSC3460	GSC3441 GSC3449
I4(5Ynude, b-ct) SEQ ID No: 120		GSC3461
I4(5Ynude, b-y) SEQ ID No: 121	GSC3451	GSC3444
I4(5Ynude, b-2) SEQ ID No:100	GSC3464	GSC3433
I4(5Ynude, A) SEQ ID No: 99	GSC3448	GSC3458
I4(5Ynude, T) SEQ ID No: 122	GSC3457	GSC3450
I4(5Y-5) SEQ ID No: 88	GSC3150	
I4(5Y-5, b-a) SEQ ID No: 123	GSC3455	GSC3463
I4(5Y-5, b-ct) SEQ ID No: 124	GSC3431	
I4(5Y-5, b-y) SEQ ID No: 125	GSC3467	GSC3429
I4(5Y-5, b-2) SEQ ID No: 126	GSC3454	
I4(5Y-5, A) SEQ ID No: 127	GSC3456	
I4(5Y-5, T) SEQ ID No: 128		GSC3459 GSC3468
I4(5Y-5, G) SEQ ID No: 102	GSC3452	GSC3437

[0206]

[0207] 도 11에 보이는 바와 같이, 분기점 변경 또는 인트론-엑손 컨센서스 부위는 일시적인 감염에서 얻어진 항-HER2 항체 역가에서 상당한 증가를 나타낸다. 상기의 변경들은 발현에 대해 중립 (ATG) 또는 음성 (예를들어 b-y)인 것으로 보인다.

[0208] 오직 작은 차이만이 분기점 및 인트론-엑손 변경의 발현 수준에서 관찰되고, 안정한 세포주 개발을 위한 두개의 구조체가 편리성 및 가용성에서 선택되었다. 두 구조체 모두 비슷한 발현 수준을 보였다: I4(0Y)-I4 및 I4(0Y,

b-2)-I4.

[0209] 대체 접합은 대체 엑손이 비슷한 인트론에 측면에 위치한 경우 증폭되었다.

[0210] 이전의 실험(실시예 1)에서 대체 엑손의 위 및 아래줄기에 동일한 인트론을 사용하는 것은 (cTNT intron #4 또는 cTNT intron #5) 두번째 개방 판독 프레임의 더 높은 발현을 야기한다. 이것이 대체 접합이 관여하는 자연적으로 인트론에 대한 유일한 사실인지 분석하기 위해, 사람 EF1 알파 유전자로부터 구조적인 인트론이 항-HER2 항체의 발현을 위해 사용되었다. 상기 EF1 알파 인트론은 대체 엑손의 위 및 아래줄기에서 복제되었다. 첫번째 인트론으로 EF1 알파 및 두번째 인트론으로 cTNT-I4와 중간 구조체도 역시 복제되었다.

[0211] 상기 결과는 도 12에 나타내었다. 대체 엑손 위 및 아래줄기 측면에 위치하는 동일한 인트론과 구조체는, 항-HER2 항체의 중쇄 또는 경쇄가 대체 엑손상에서 발현되었는지 여부와 무관하게, 다른 인트론을 가지는 구조체와 비교시 더 놓은 발현을 나타낸다.

[0212] cTNT 인트론을 사용하여 발현 수준이, 비록 사람 EF1 알파 인트론은 강화된 활성을 가짐에도, EF1 알파 인트론과 비교시 더 높았다. 상기와 같은 놀라운 결과는 대체 접합과 자연적으로 관련된 인트론을 사용하는 것은 두번째 엑손의 더 놓은 발현을 야기하고 이에 항체와 같은 다중 단백질의 더 좋은 발현을 야기한다. 대체 엑손의 측면에 위치한 동일한 인트론을 사용한 또 다른 예는 상기 실시예 1에 cTNT-인트론 5와 나타났다. 여기서도 역시 동일한 인트론의 사용은 두 대체 엑손의 더욱 평형된 발현을 야기하였다.

[0213] 실시예 4: 항-HER2 항체를 발현하는 안정한 세포주의 제조

[0214] CHO-S 세포에 보고 항-HER2 항체의 안정한 발현을 얻기 위해, 상기 실시예 3에서 설명된 대체 접합 구조체 I4(0Y)I4-anti-HER2-LC-HC가 마우스 CMV 프로모터 및 Ig 가변 부위 인트론 및 스플라이스 수용 서열의 조절하에 pGLEX41 발현 벡터에서 복제 되었다 (Bothwell et al., supra). 상기 복제 단계는 벡터 pGLEX41-ASC-I4(0Y)I4-anti-HER2-LC-HC를 야기한다.

[0215] 두개의 추가적인 벡터가 퓨로마이신 및 네오마이신에 대한 저항 유전자를 운반한다. 두개의 저항 유전자 모두 SV40 프로모터 조절하에 있다.

[0216] 상기 세포는 제조사에서 추천한 과정에 따라 JetPEI™을 사용하여 감염되었다 (Polyplus-transfections, Strasbourg, France). 생성 유전자를 운반하는 상기 발현 벡터 및 선택에 (퓨로마이신 및 제네티신) 사용되는 항생체에 저항에 대한 유전자를 제공하는 두개의 벡터는 선형화 되었고 CHO-S(cGNP banked) 숙주 세포로 보조-감염되었다. 상기 플라스미드는 CHO-S 세포 주의 게놈에 무작위 통합 자리에 도입되었다. 우리의 손에, 상기 과정은 빠르게 고도로 재현되고 안정하고 높은 발현 세포주를 효과적으로 발생시켰다.

[0217] 상기 세포의 이후 배양뿐 아니라 감염은 배양액 없이 구성을 유도하는 동물에서 수행되었다. 감염 후 그 날에, 세포들은 선택된 배양액에 (퓨로마이신 및 제네티신을 함유한 성장 배양액) 96 웰 플리에트에 다른 세포 밀도로 주입되었다. 둘 항생제 모두는 효과적인 단백질 합성의 저해제이다. 두배의 선택 때문에 높은 선택 압력은 효과적으로 비 감염된 세포뿐만 아니라 비- 및 낮은-생성 복제를 제거하였다. 37℃, 5% CO_2, 및 80% 습도에서 배양 일주일 후, 상기 선택 압력은 1 볼륨의 선택 배양액을 세포로 추가함으로써 갱신되었다. 정적 배양의 또 다른 일주일 후 성장을 보이는 웰의 30% 보다 적은 희석 수율이 확인되었다. 상기 성장을 보인 웰의 상등액은 옥텟을 사용하여 축적된 항-HER2 항체에 대해 분석되었다 (Fortebio, Manlo Park, CA). 가장 높은 발현을 보인 72 미니풀은 첫번째로 24 웰 플레이트로 확장되었고, 이후 튜브스핀 규모로 현탁액에 및 튜브스핀 50 ml 생물반응기에 보충된 14일 배치에서 평가되었다. 배치 배양의 끝에 가장 높은 역가는 197 μg/ml이었다 (도 13 참조).

[0218] 복제 군을 제조하기 위해, 네개의 최고의 발현 미니풀은 150-197 μg/ml로부터 발현 범위로 선택되어 두번째 제한 희석의 차수가 진행되었다. 이것은 상기 세포를 96 웰 플레이트에 성장 배양액에 다른 희석에 플레이팅 함으로써 수행되었다. 이주 후에 다른 희석에서 키워진 복제의 수를 평가하였다. 복제 군은 처음 24 웰 플레이트로 확장되었고 이후 50 ml 생물 반응기 튜브 규모로 하였다. 상기 규모에서 가장 높은 역가는 배양액 10 ml를 사용한 50 ml의 생물 반응기 튜브에서 비-최적화된 보충 배치에서 250 μg/ml로 얻어졌다 (도 14 참조).

[0219] 대체 접합과 얻어진 동일한 항체 최대 역가와 이 단계에서 얻어진 일반적인 역가와 비교시 3배 작은 것으로 나타났다. 그럼에도 상기의 역가는 대체 접합 기술에 기초한 항체를 생산하는 안정한 세포주의 첫번째 산업적으로 상당한 생산 수준의 것이다.

[0220] 실시예 5: 대체 접합 구조체를 사용한 이중 특이성 항체의 발현

[0221] 이중 특이성 항체는 두개의 다른 항원 결정기를 인식하기 위해 인위적으로 조절된 항체이다. 치료 용도의 이중 특이성 항체의 개발에 주요 문제점은 산업적인 큰 규모의 생산이다. 따라서 이중 특이성의 더 높은 발현 또는 높은 순도로 이중 특이성 항체의 생산을 가능하게하는 기술의 개발은 (생성물에 의한 이중 특이성 항체의 낮은 부산물로) 매우 중요하다.

[0222] 이중 특이성 항체는 여러개의 하위단위들로 구성되어 있다. 발현에 필요한 하위단위들의 수는 선태되는 포멧에 의존적이다. 본 발명의 한 관점에서, 이중 특이성 항체 구조체는 경쇄, 중쇄 및 Fc-scFv를 코딩하는 세개의 다른 하위단위로 구성되어 있다. 일반 항체와 비슷한 중쇄 및 경쇄가 최적의 비율로 감염될 필요에서, 이중 특이성 구조체는 세개의 하위단위의 특정 비율에서 가장 잘 발현되었다. 상기 비율은 이중 특이성 항체에 의존적이고 하나의 포멧에서 또 다른 것으로 변화될 수 있다. 실시예 1-3에서 개발된 대체 접합 발현 카세트는 두개의 다른 단백질 (GFP 또는 dsRED)의 동시 발현 또는 고정된 비율에서 동일 단백질 (항체의 중쇄 및 경쇄)의 하위단위들의 동시 발현을 가능하게 한다. 정확한 몰 비에서 이중 특이성 항체의 하위단위들을 발현하는 것이 바람직하기 때문에, 상기 대체 접합 구조체는 가장 높은 발현 또는 가장 낮은 오염과 부산물을 야기하는 비율에서 두 하위단위의 발현시키기 위해 유용한 것임을 증명하였다. 자체적으로 생성된 이중 특이성 항체는 세개의 다른 하위단위들로 구성된다: 중쇄, 경쇄 및 Fc-scFv. 정확하게 구성된 생성물의 최적 발현을 위해서, Fc-scFv에 대한 중쇄의 비율이 일시적인 보조-감염 실험에서 가장 중요한 변수인 것으로 나타났다. 경쇄의 상대적인 비율은 덜 중요했다.

[0223] 상기 관찰에 기초하여, 실시예 3에서 설명된 바와 같이, 중쇄 및 Fc-scFv는 eoc 접합 구조체 I4(7Y)I4h로 복제되었고, GSC5642 (배향: HC-scFv), GSC5643 (배향: scFv-HC) 및 경쇄의 발현을 위한 GSC5641 벡터를 야기하였다.

[0224] 대체 접합 구조체와 벡터 및 경쇄에 대한 벡터는 대체 접합 구조체 및 경쇄를 코딩한 벡터의 다른 비율을 사용하여 CHO-S 세포에 보조-감염되었다. 결과 항체의 발현 수준은 도 15에 나타내었다.

[0225] 일반적으로, 상기 발현 수준은 경쇄 구조체보다 대체 접합 구조체의 비율의 증가와 함께 둘 모두의 구조체에 대해 증가한다. 경쇄의 보다 높은 발현은 상등액의 항체의 양을 감소시킨다. 가장 높은 발현 수준은 세배 몰 과량에서 관찰되었다. 정체가 관찰되지 않았기 때문에, 진정한 최적화는 한층 더 높은 몰 과량일 수 있다. 이중 특이성의 발현 수준 또는 폴리(Y)의 양의 변화를 사용하여 분산 단백질에서 부산물의 수준을 최적화 하기 위한 실험이 수행되지 않았다. 따라서 더높은 발현 또는 사용되는 구조체에서 더 낮은 부산물 오염을 위한 가능성이 존재할 것이다.

[0226] 이중 특이성 항체의 존재는 ELISA에 의해 확인 되었다 (이중 특이성 항체의 두개의 할에 대한 특정). 대체 접합 구조체 I4(7Y)I4sh를 사용한 이중 특이성 항체의 성공적인 발현은 대체 접합이 일반 항체뿐 아니라 두 종류 하위단위보다 많은 이중 특이성 항체의 성공적인 발현을 위해 사용될 수 있다는 것을 설명해 준다. 최적 비율에서 발현은 보조-감염에 의해 달성될 수 있다 (최적 비율의 확인을 위해 수행되기 때문에). 그럼에도 대체 접합 카세트를 사용하는 것의 주요 이점은 안정한 세포 포멧에서 최적의 비율을 직접 번역할 수 있는 가능성이다.

도면

도면1a

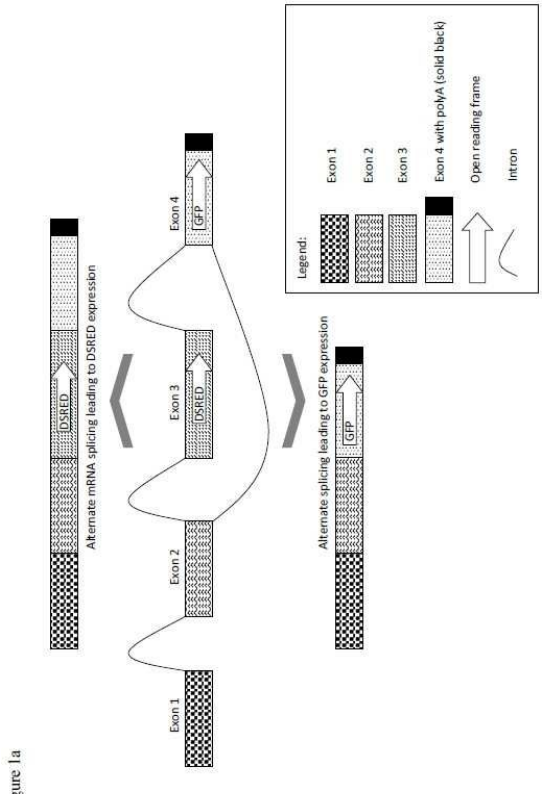


Figure 1a

도면1b

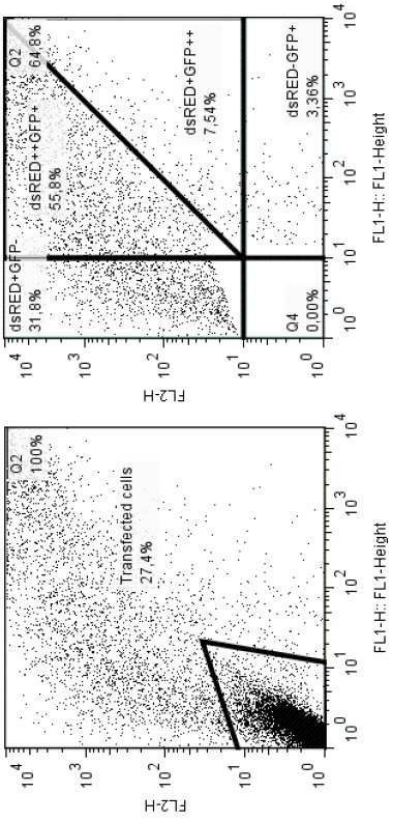
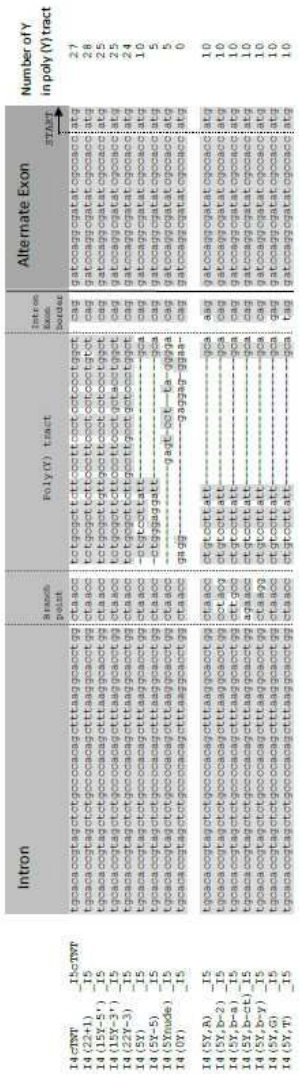


Figure 1b

도면2a

Figure 2a



도면2b

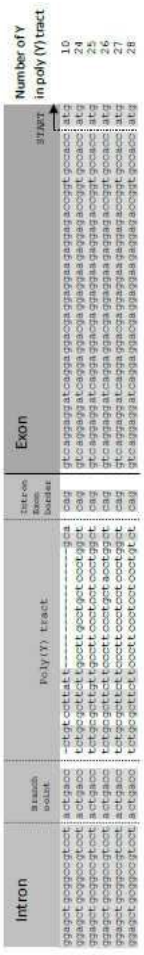


Figure 2b

도면2c

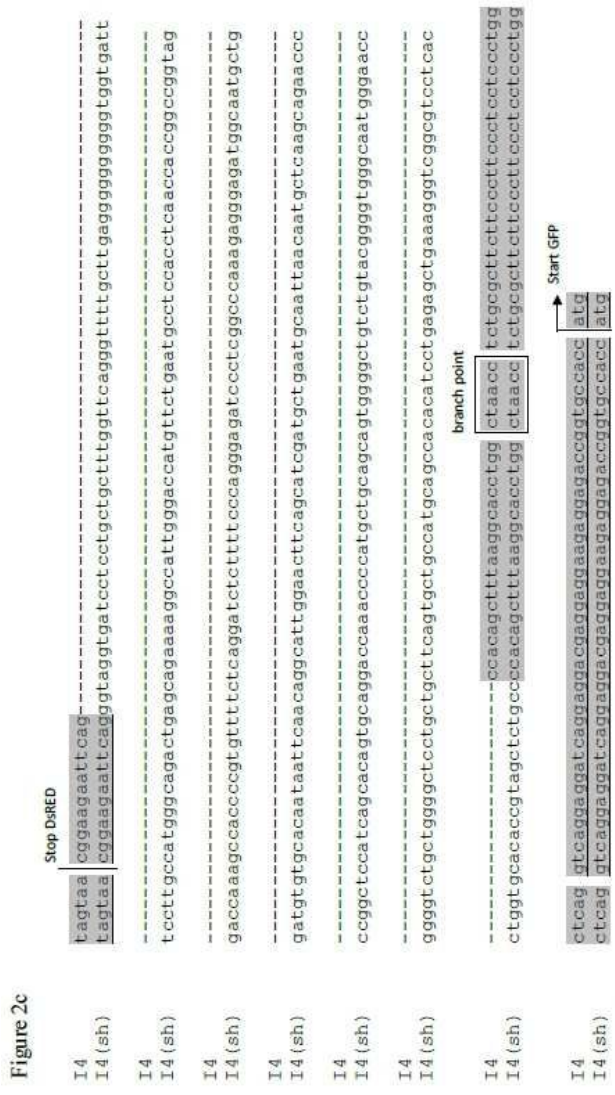


Figure 2c
Stop DsRED
I4 tagtaa cggaagaattcag----------
I4(sh) tagtaa cggaagaattcaggtaggtgatcctcctgctgctttggttcaggggttttgcttgaggggggggggtggtgatt
I4 ----------
I4(sh) tccttgccatgggcagactgagcagaaaaggccattgggaccatgtttctgaatgcctccacctcaaccaccggccggtag
I4 ----------
I4(sh) gaccaaagccaccccgtgtttttctcaggatctcttttcccagggagatcccctcggcccaaagagggagatggcaatgctg
I4 ----------
I4(sh) gatgtgtgcacaatattcaacaggcattggaacttcagcatcgatgctgaatgcaattaacaatgctcaagcagaaccc
I4 ----------
I4(sh) ccggctccatcagcacagtgcaggaccaaacccccatgctgcagcagtgggggctgtctgtacgggggtgggcaatggaacc
I4 ----------
I4(sh) ggggtctgctgggggctcctgctgcttcagtgctgccatgcagccacacatcctgagagctgaaagggtcggcgtcctcac
branch point
I4 ----------ccacagctttaaggcacctgg ctaacc tctgcgcttctttccctttccctcctccctgg
I4(sh) ctggtgcacaccgtagctctgcccccacagctttaaggcacctgg ctaacc tctgcgcttctttccctttccctcctccctgg
Start GFP
I4 ctcag gtcaggaggatcaggaggacgaggaggaagaggaggagaccggtgccacc atg
I4(sh) ctcag gtcaggaggatcaggaggacgaggaggaagaggaggagaccggtgccacc atg
Underlined: exon
Not underlined: intron

도면3a

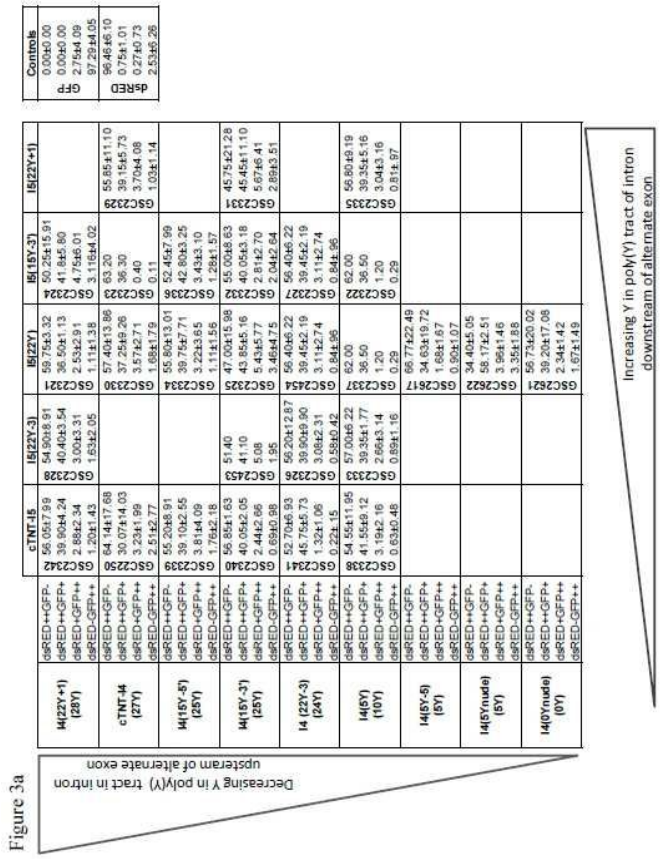


Figure 3a
Decreasing Y in poly(Y) tract in intron upstream of alternate exon
Increasing Y in poly(Y) tract of intron downstream of alternate exon
I4(22Y+1) (28Y)
cTNT-I4 (27Y)
I4(15Y-5') (25Y)
I4(15Y-3') (25Y)
I4 (22Y-3') (24Y)
I4(5Y) (10Y)
I4(5Y-5') (5Y)
I4(5Ynude) (5Y)
I4(0Ynude) (0Y)
dsRED++/GFP-
dsRED++/GFP+
dsRED+/GFP++
dsRED-/GFP++
cTNT-I5
I5(22Y-3')
I5(22Y)
I5(15Y-3')
I5(22Y+1)
GSC2342
GSC2250
GSC2339
GSC2340
GSC2341
GSC2338
GSC2337
GSC2617
GSC2622
GSC2621
GSC2328
GSC2330
GSC2334
GSC2325
GSC2454
GSC2453
GSC2336
GSC2332
GSC2327
GSC2329
GSC2331
GSC2335
dsRED
GFP
Controls
0.00±0.00
0.00±0.00
2.75±4.09
97.29±4.05
99.46±6.10
0.75±1.01
0.27±0.73
2.53±6.28

도면3b

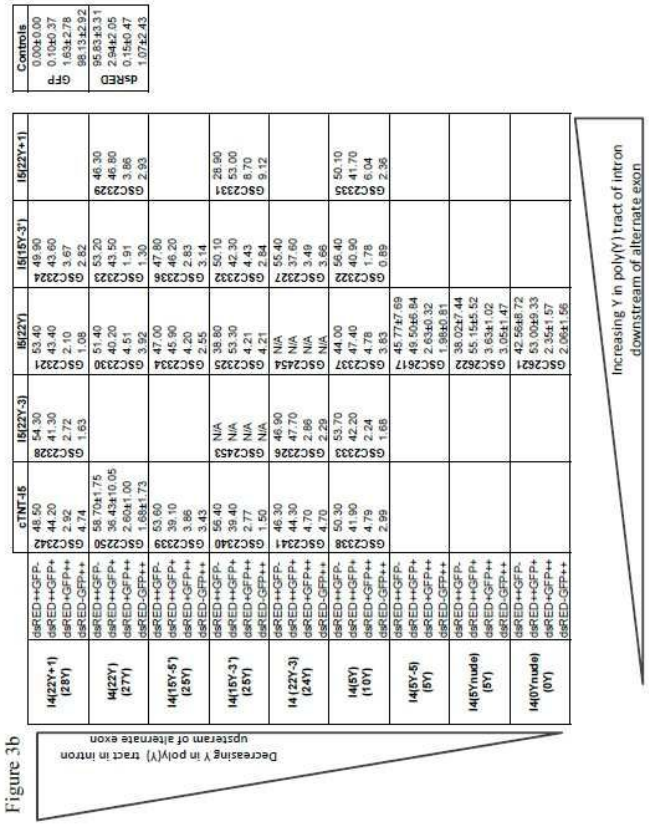


Figure 3b
Decreasing Y in poly(Y) tract in intron upstream of alternate exon
Increasing Y in poly(Y) tract of intron downstream of alternate exon
I4(22Y+1) (28Y)
I4(22Y) (27Y)
I4(15Y-5') (25Y)
I4(15Y-3') (25Y)
I4 (22Y-3) (24Y)
I4(5Y) (10Y)
I4(5Y-5) (5Y)
I4(0Ynude) (0Y)
dsRED+GFP-
dsRED++GFP+
dsRED+GFP++
dsRED-GFP++
cTNT-I5
I5(22Y-3)
I5(22Y)
I5(16Y-3')
I5(22Y+1)
GSC2342
GSC2328
GSC2321
GSC2329
GSC2250
GSC2330
GSC2323
GSC2339
GSC2334
GSC2336
GSC2340
GSC2453
GSC2325
GSC2332
GSC2331
GSC2341
GSC2326
GSC2454
GSC2327
GSC2338
GSC2333
GSC2337
GSC2335
GSC2617
GSC2622
GSC2621
Controls
GFP
dsRED

도면4a

Branch point and intron-exon border sequence	dsRED++GFP-	95.40±1.13	95.00	94.80±1.41
	dsRED++GFP+	4.45±1.11	4.82	4.98±1.51
	dsRED+GFP++	0.14±0.02	0.11	0.15±0.05
	dsRED-GFP++	0.024±0.00	0.06	0.06±0.02
		I4(5Y,G) || cTNT-I5(22Y)	I4(5Y,b-ct) || cTNT-I5(22Y)	I4(5Y,b-y) || cTNT-I5(22Y)
Introns arrangements	dsRED++GFP-	63.98±21.66	78.35±0.92	64.13±17.68
	dsRED++GFP+	28.30±14.09	19.30±0.42	30.07±14.03
	dsRED+GFP++	5.77±5.76	1.31±0.18	3.23±1.99
	dsRED-GFP++	1.78±1.63	0.94±0.57	2.51±2.77
		cTNT-I5 || cTNT-I5	cTNT-I5 || cTNT-I4	cTNT-I4 || cTNT-I5
Controls	dsRED++GFP-	0.00±0.00	96.46±6.10	
	dsRED++GFP+	0.00±0.00	0.75±1.01	
	dsRED+GFP++	2.75±4.09	0.27±0.73	
	dsRED-GFP++	97.29±4.05	2.53±6.26	
		GFP	dsRED	

GSC2614 GSC2613 GSC2741 GSC2615 GSC2250 GSC2618 GSC2619

Figure 4a

도면4b

Branch point and intron-exon border sequence		GSC2613	GSC2615	GSC2618	GSC2619
	dsRED++GFP-	75.73±3.17	71.95±2.78	76.37±1.72	
	dsRED++GFP+	23.25±3.21	26.35±2.83	22.77±2.48	
	dsRED+GFP++	0.75±0.20	1.16±0.30	0.64±0.56	
	dsRED-GFP++	0.14±0.13	0.26±0.28	0.18±0.31	
		I4(5Y,G) || cTNT-I5(22Y)	I4(5Y,b-ct) || cTNT-I5(22Y)	I4(5Y,b-y) || cTNT-I5(22Y)	45.28±7.36
Intron arrangements	dsRED++GFP-	43.63±9.67	61.00±1.76	58.70±1.75	45.90±4.81
	dsRED++GFP+	46.12±8.92	36.00±1.57	36.43±10.05	5.47±2.40
	dsRED+GFP++	8.48±3.46	1.76±0.04	2.60±1.00	1.90±1.39
	dsRED-GFP++	1.14±0.99	0.80±0.51	1.68±1.73	
		GSC2614	GSC2741	GSC2250	GSC2619
		cTNT-I5 || cTNT-I5	cTNT-I5 || cTNT-I4	cTNT-I4 || cTNT-I5	cTNT-I4 || cTNT-I4
Controls	dsRED++GFP-	0.00±0.00	95.83±3.31		
	dsRED++GFP+	0.10±0.37	2.94±2.05		
	dsRED+GFP++	1.63±2.78	0.15±0.47		
	dsRED-GFP++	98.13±2.92	1.07±2.43		
		GFP	dsRED		

Figure 4b

도면5a

		GSC26I9	GSC2739	GSC2742	GSC2740
Reduction in poly(Y)	dsRED++GFP-	58.58±19.74	35.75±28.07	8.44±0.52	24.80±3.82
	dsRED++GFP+	33.52±14.50	49.45±12.80	55.20±3.25	60.90±4.38
	dsRED+GFP++	4.20±2.74	7.99±7.93	19.40±0.71	8.46±0.33
	dsRED-GFP++	3.50±3.26	6.67±7.25	16.40±3.11	5.61±0.43
		cTNT-I4 | cTNT-I4	I4(5Y-5) | cTNT-I4	I4(5Ynude) | cTNT-I4	I4(0Y) | cTNT-I4
Branch Point Region	dsRED++GFP-	51.15±20.50	54.70±15.56	40.88±8.86	48.53±19.69
	dsRED++GFP+	39.13±13.38	38.65±11.81	47.30±4.14	40.73±10.84
	dsRED+GFP++	4.88±3.95	3.86±2.26	5.81±2.24	6.43±5.88
	dsRED-GFP++	4.48±4.19	2.70±1.44	5.84±3.20	4.10±4.22
		I4(5Y,b-y) | cTNT-I4	I4(5Y,b-2) | cTNT-I4	I4(5Y,b-ct) | cTNT-I4	I4(5Y,b-a) | cTNT-I4
Controls	dsRED++GFP-	64.13±17.68	0.00±0.00	96.46±6.10	
	dsRED++GFP+	30.07±14.03	0.00±0.00	0.75±1.01	
	dsRED+GFP++	3.23±1.99	2.75±4.09	0.27±0.73	
	dsRED-GFP++	2.51±2.77	97.29±4.05	2.53±6.26	
		cTNT-I4 | cTNT-I5	GFP	dsRED	

Figure 5a

도면5b

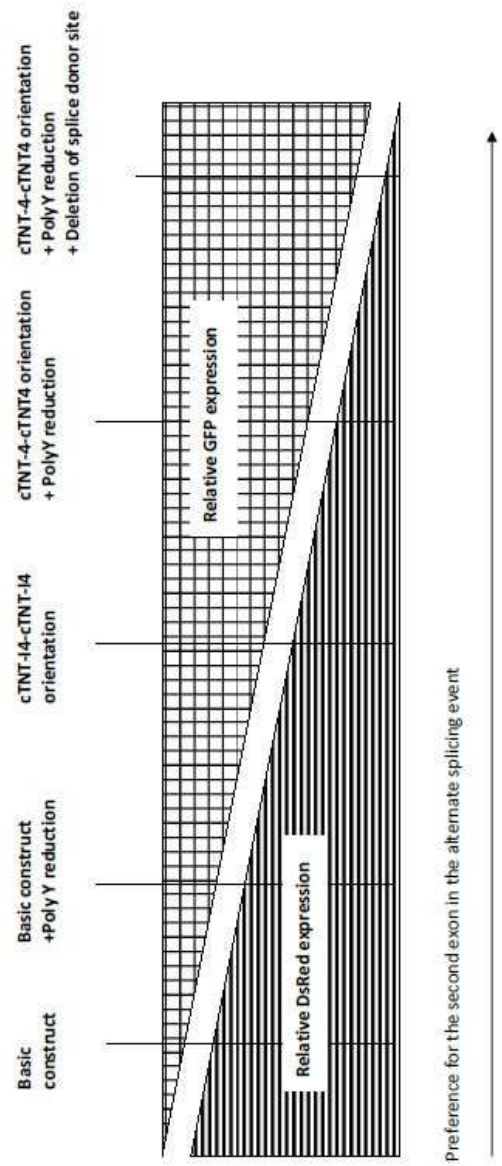
		GSC2619	GSC2739	GSC2740	GSC2742
Reduction in poly(Y)	dsRED++GFP-	45.28±7.36	20.35±3.53	22.13±2.39	14.87±2.78
	dsRED++GFP+	45.90±4.81	59.88±7.94	57.57±6.68	55.45±9.92
	dsRED+GFP++	5.47±2.40	12.52±4.27	12.20±3.78	17.68±5.79
	dsRED-GFP++	1.90±1.39	6.30±0.84	7.44±1.29	11.14±1.40
		cTNT-I4 | cTNT-I4	I4(5Y-5) | cTNT-I4	I4(0Y) | cTNT-I4	I4(5Ynude) | cTNT-I4
Controls	dsRED++GFP-	58.70±1.75	0.00±0.00	95.83±3.31	
	dsRED++GFP+	36.43±10.05	0.10±0.37	2.94±2.05	
	dsRED+GFP++	2.60±1.00	1.63±2.78	0.15±0.47	
	dsRED-GFP++	1.68±1.73	98.13±2.92	1.07±2.43	
		cTNT-I4 | cTNT-I5	GFP	dsRED	

Figure 5b

도면6

		GSC2781	GSC2782	GSC2784	GSC2783
Shortening of second intron	dsRED++GFP-	0.18±0.13	0.09±0.08	0.01±0.01	0.06±0.06
	dsRED++GFP+	0.44±0.18	0.20±0.13	0.07±0.09	0.23±0.16
	dsRED+GFP++	53.12±2.22	48.35±2.11	38.27±3.48	47.77±2.41
	dsRED-GFP++	41.50±2.66	46.35±2.38	54.58±1.81	46.40±2.98
		cTNT-I4 | I4(sh)	I4(5Y-5) | I4(sh)	I4(5Ynude) | I4(sh)	I4(0Y) | I4(sh)
		GSC2619	GSC2250	GFP	dsRED
Controls	dsRED++GFP-	45.28±7.36	58.70±1.75	0.00±0.00	95.83±3.31
	dsRED++GFP+	45.90±4.81	36.43±10.05	0.10±0.37	2.94±2.05
	dsRED+GFP++	5.47±2.40	2.60±1.00	1.63±2.78	0.15±0.47
	dsRED-GFP++	1.90±1.39	1.68±1.73	98.13±2.92	1.07±2.43
		cTNT-I4 | cTNT-I4	cTNT-I4 | cTNT-I5		

도면7


Basic construct
Basic construct +PolyY reduction
cTNT-14-cTNT-I4 orientation
cTNT-4-cTNT4 orientation + PolyY reduction
cTNT-4-cTNT4 orientation + PolyY reduction + Deletion of splice donor site
Relative DsRed expression
Relative GFP expression
Preference for the second exon in the alternate splicing event

도면8

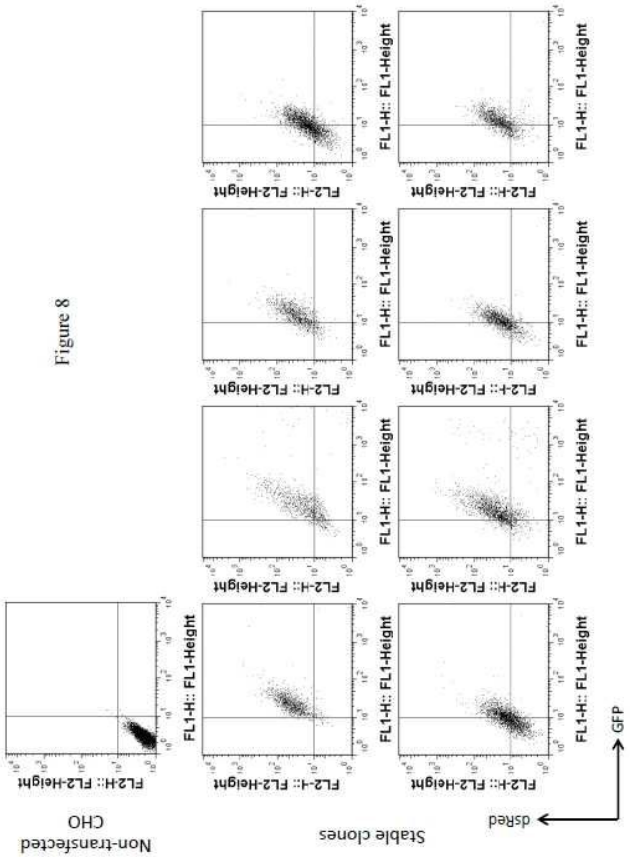
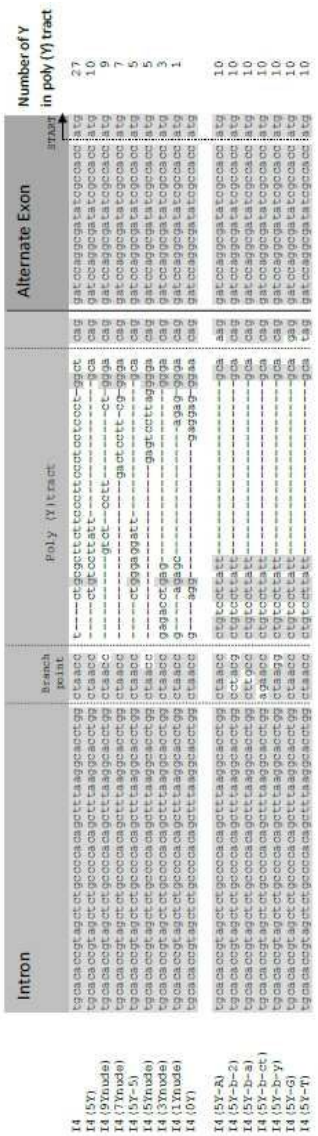
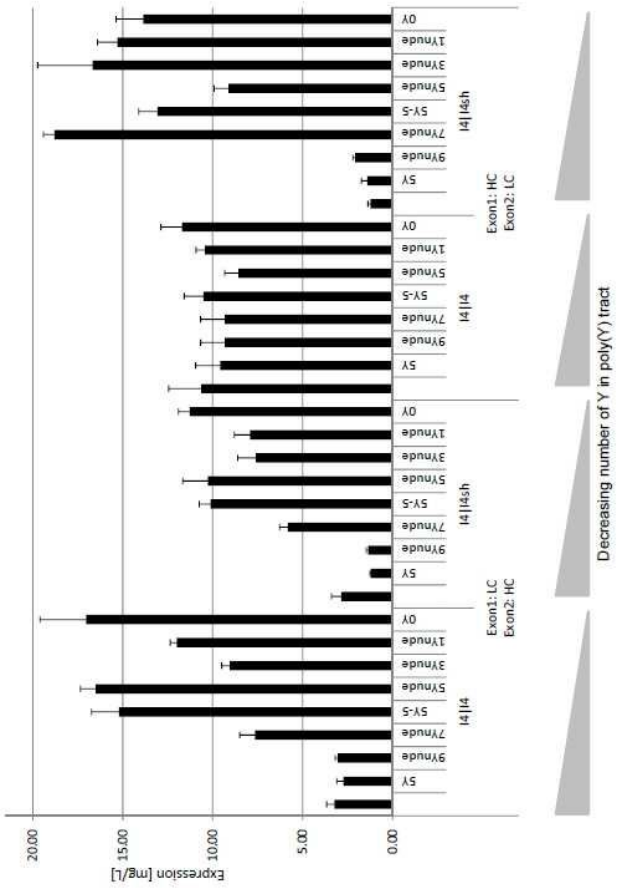


Figure 8
Non-transfected CHO
Stable clones
dsRed
GFP
FL1-H:: FL1-Height
FL2-H:: FL2-Height

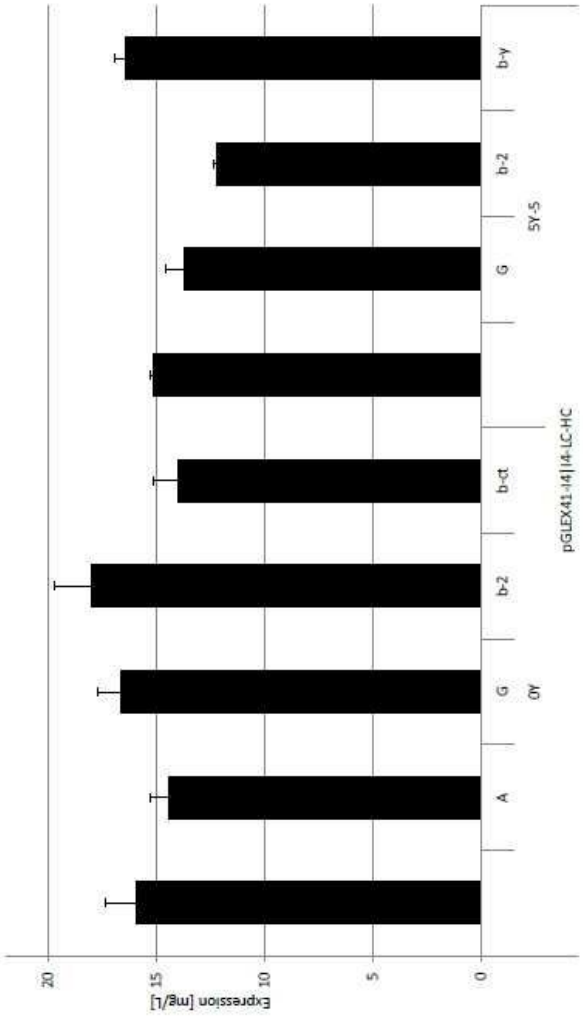
도면9


Intron
Branch point
Poly (Y)tract
Alternate Exon
start
Number of Y in poly (Y) tract
I4
I4(5Y)
I4(9Ynode)
I4(7Ynode)
I4(5Y-5)
I4(5Ynode)
I4(3Ynode)
I4(1Ynode)
I4(0Y)
I4(5Y-A)
I4(5Y-b-2)
I4(5Y-b-a)
I4(5Y-b-ct)
I4(5Y-b-y)
I4(5Y-G)
I4(5Y-T)
27
10
9
7
5
5
3
1
10
10
10
10
10
10
10

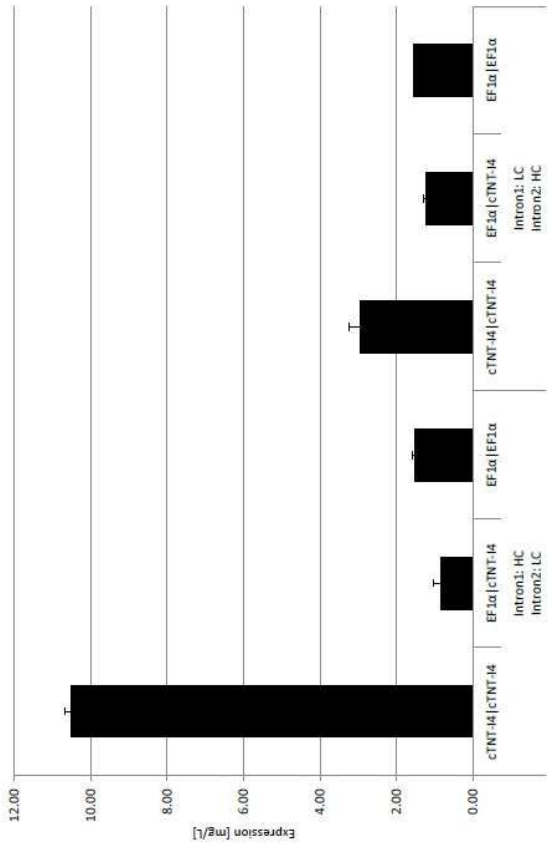
도면10


Expression [mg/L]
20.00
15.00
10.00
5.00
0.00
I4|I4sh
I4|I4
I4|I4sh
I4|I4
Exon1: HC
Exon2: LC
Exon1: LC
Exon2: HC
0Y
1Ynude
3Ynude
5Ynude
5Y-S
7Ynude
9Ynude
5Y
Decreasing number of Y in poly(Y) tract

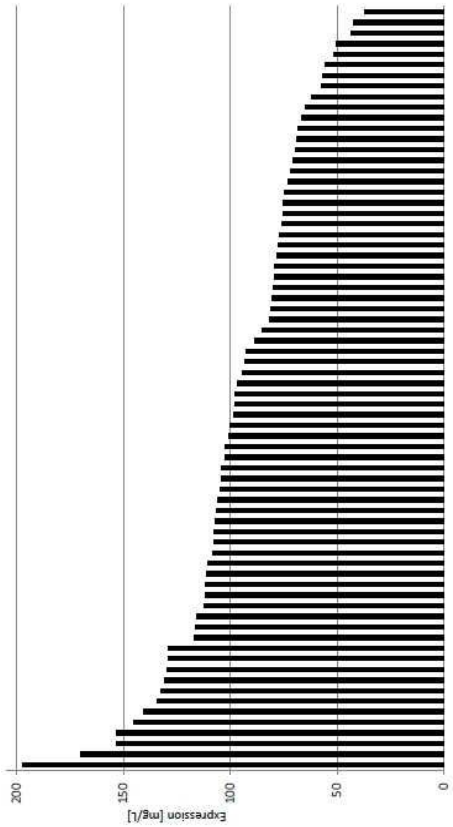
도면11


Expression [mg/L]
20
15
10
5
0
A
G
b-2
b-ct
G
b-2
b-y
0Y
5Y-5
pGLEX41-I4|I4-LC-HC

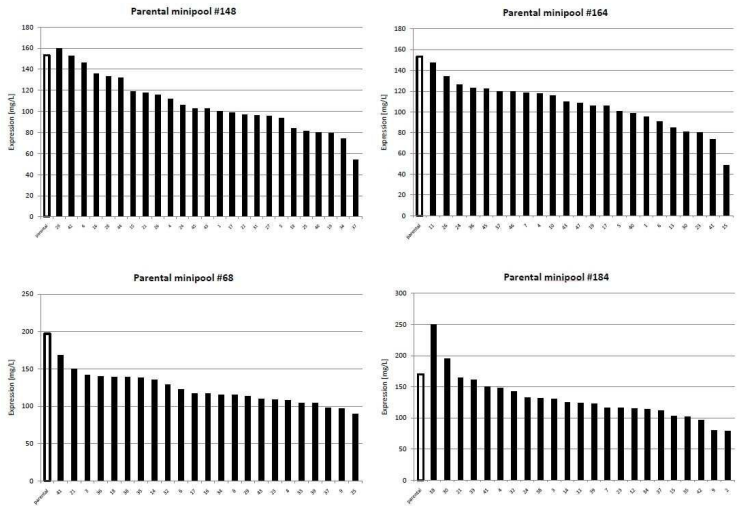
도면12


Expression [mg/L]
12.00
10.00
8.00
6.00
4.00
2.00
0.00
cTNT-I4|cTNT-I4
EF1α|cTNT-I4
Intron1: HC
Intron2: LC
EF1α|EF1α
cTNT-I4|cTNT-I4
EF1α|cTNT-I4
Intron1: LC
Intron2: HC
EF1α|EF1α

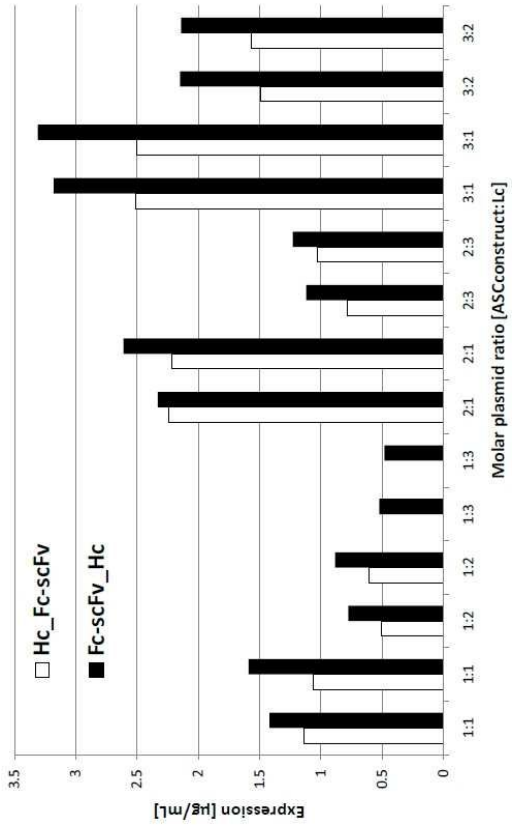
도면13


200
150
100
50
0
Expression [mg/L]

도면14


Parental minipool #148
Parental minipool #164
Parental minipool #68
Parental minipool #184
Expression [mg/L]

도면15



서 열 목 록

<110> Glenmark Pharmaceuticals SA

<120> Expression constructs and methods for expressing polypeptides in eukaryotic cells

<130> 2016FPI-02-007/CH

<150> EP 13179375.4

<151> 2013-08-06

<150> PCT/EP 2014/066826

<151> 2014-08-05

<160> 175

<170> KopatentIn 2.0

<210> 1

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr991_Primer

<400> 1

ggtcatttcg aatcattact tgtacagctc gt 32

<210

> 2

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1095_Primer

<400> 2

cgctggctag cgtttaaact taag 24

<210> 3

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1096_Primer

<400> 3

atcgttcgaa tatgggccct ctcgcacacc ggtctcctct tcctcctc 48

<210> 4

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1097_Primer

<400> 4

tatagggccc tgtgagcaag ggcgaggag 29

<210> 5

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1098_Primer

<400> 5

gcgcttcgaa tcattacttg tacagctcgt c 31

<210> 6

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1099_Primer

<400> 6

tatagggccc tctacaggaa caggtggtg 29

<210> 7

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1100_Primer

<400> 7

attaaccggt gcctcctccg aggacgtc 28

<210> 8

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1138_Primer

<400> 8

aattaagcta gcgtttaaac ttaagcttcc ttggattaca aggatgacga t 51

<210> 9

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1139_Primer

<400> 9

gtggcgatat cgcctggatc ctgag 25

<210> 10

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1140_Primer

<400> 10

ccaggcgata tcgccaccat gggtgcctcc tccgagga 38

<210> 11

<211> 43

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1141_Primer

<400> 11

ctacctgaat tcttccgtta ctacaggaac aggtggtggc ggc 43

<210> 12

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1142_Primer

<400> 12

gaggagaccg gtgccaccat ggagcaaggg cgaggagctg t 41

<210> 13

<211> 63

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1158_Primer

<400> 13

aattaagcta gcgtttaaac ttaagcttcc ttggaggacc cagtacccgg atctagaggt 60

agg 63

<210> 14

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1180_Primer

<400> 14

aattaaaccg gtgccaccat ggtgagcaag ggcgaggagc 40

<210> 15

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1181_Primer

<400> 15

gcgcggctag cgtttaaact taagc 25

<210> 16

<211> 62

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1182_Primer

<400> 16

ttgtgatatc gcctggatcc tgtgcaataa ggacagggtt agccaggtgc cttaaagctg 60

tg 62

<210> 17

<211> 37

<212> DNA

<213> Artificial Sequence

<220

><223> Glnpr1183_Primer

<400> 17

agcaggatat cgcctggatc ctgagacagg gaggagg 37

<210> 18

<211> 62

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1184_Primer

<400> 18

atatgatatc gcctggatcc tgagccaggg agcaggcaag gcaagaagcg cagaggttag 60

cc 62

<210> 19

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1185_Primer

<400>

19

agtcgatatc gcctggatcc tgagccaggt agcagggaag ggaag 45

<210> 20

<211> 62

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1186_Primer

<400> 20

gatggatatc gcctggatcc tgagccaggg aggagggaag gcaacaagcg cagaggttag 60

cc 62

<210> 21

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1187_Primer

<400> 21

gcgcgaattc aggtagttac tgcac 25

<210> 22

<211> 66

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1189_Primer

<400> 22

tataaccggt ctcctcttcc tcctcgtcct cctgatcctc ctgacctgag ccagggagga 60

gggaag 66

<210> 23

<211> 74

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1190_Primer

<400> 23

taataccggt ctcctcttcc tcctcgtcct cctgatcctc ctgacctgag ccagggagca 60

ggcaaggcaa gaag 74

<210> 24

<211> 66

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1191_Primer

<400> 24

atataccggt ctcctcttcc tcctcgtcct cctgatcctc ctgacctgag acagggagga 60

gggaag 66

<210> 25

<211> 66

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1192_Primer

<400> 25

atataccggt ctcctcttcc tcctcgtcct cctgatcctc ctgacctgag ccagggagga 60

gggaag 66

<210> 26

<211> 74

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1193_Primer

<400> 26

atataccggt ctcctcttcc tcctcgtcct cctgatcctc ctgacctgag ccaggtagca 60

gggaagggaa gaag 74

<210> 27

<211> 77

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1237_Primer

<400> 27

ggcggctagc gtttaaactt aagcttcctt ggaggaccca gtacccggat ctagagtagt 60

tactgcacct ttctttg 77

<210> 28

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1238_Primer

<400> 28

atcggatatc gcctggatcc tgtgcaataa ggacagggtc 40

<210> 29

<211> 35

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1239_Primer

<400> 29

gtggcgatat cgcctggatc cthtgcaata aggac 35

<210> 30

<211> 57

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1240_Primer

<400> 30

tggcgatatc gcctggatcc tgtgcaataa ggacagcctt agccaggtgc cttaaag 57

<210> 31

<211> 57

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1241_Primer

<400> 31

tggcgatatc gcctggatcc tgtgcaataa ggacagggtt ctccaggtgc cttaaag 57

<210> 32

<211> 57

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1242_Primer

<400> 32

tggcgatatc gcctggatcc tgtgcaataa ggacagggca agccaggtgc cttaaag 57

<210> 33

<211> 57

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1243_Primer

<400> 33

tggcgatatc gcctggatcc tgtgcaataa ggacagcgta ggccaggtgc cttaaag 57

<210> 34

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1244_Primer

<400> 34

gcgatatcgc ctggatcctg tcccctaagg actcggttag ccaggtgcct taaagctgtg 60

60

<210> 35

<211> 60

<212> DNA

<213> Artificial Sequence

<220

><223> Glnpr1245_Primer

<400> 35

gcgatatcgc ctggatcctg tgcaatcctc ccagggttag ccaggtgcct taaagctgtg 60

60

<210> 36

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> Glnpr1246_Primer

<400> 36

gcgatatcgc ctggatcctg ttccctcctc cctcggttag ccaggtgcct taaagctgtg 60

60

<210> 37

<211> 39

<212>

DNA

<213> Artificial Sequence

<220><223> Glnpr1285_Primer

<400> 37

cggaagaatt cagccacagc tttaaggcac ctggctaac 39

<210> 38

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2250/GSC2246_Construct

<400> 38

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gaccctgtcc ttattgcaca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 39

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2329_Construct

<400> 39

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gcccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320
acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380
cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440
actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500
actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560
agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620
gtagttactg caccttttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740
tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800
tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860
ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920
tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980
agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040
ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160
ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220
gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280
gacctctgcg cttcttccct tccctcctcc ctgtctcagg tcaggaggat caggaggacg 2340
aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400
tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460
gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580
tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640
gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700
aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760
aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820
atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880
tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 40

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2330_Construct

<400> 40

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140
acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200
agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260
ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320
acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380
cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440
actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560
agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620
gtagttactg caccttttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680
cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740
tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800
tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860
ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980
agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040
ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100
agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160
ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220
gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280
gacctctgcg cttcttccct tccctcctcc ctggctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400
tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460
gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520
gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580
tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640
gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700
aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 41

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2323_Construct

<400> 41

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gacctctgcg cttcttccct tccctgctac ctggctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400

tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460

gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640

gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700

aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 42

<211> 3227

<212> DNA

<213> Artificial Sequence

<220><223> GSC2619_Construct

<400> 42

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

ggtaggtgat cctcctgctg ctttggttca gggttttgct tgagggggggg gggtggtgat 1680

ttccttgcca tgggcagact gagcagaaaa ggccattggg accatgttct gaatgcctcc 1740

acctcaacca ccggccggta ggaccaaagc caccccgtgt tttctcagga tctcttttcc 1800

cagggagatc cctcggccca aagagggaga tggcaatgct ggatgtgtgc acaataattc 1860

aacaggcatt ggaacttcag catcgatgct gaatgcaatt aacaatgctc aagcagaacc 1920

cccggctcca tcagcacagt gcaggaccaa accccatgct gcagcagtgg ggctgtctgt 1980

acggggtggg caatgggaac cggggtctgc tggggctcct gctgcttcag tgctgccatg 2040

cagccacaca tcctgagagc tgaaagggtc ggcgtcctca cctggtgcac accgtagctc 2100

tgccccacag ctttaaggca cctggctaac ctctgcgctt cttcccttcc ctcctccctg 2160

gctcaggtca ggaggatcag gaggacgagg aggaagagga gaccggtgcc accatggtga 2220

gcaagggcga ggagctgttc accggggtgg tgcccatcct ggtcgagctg gacggcgacg 2280

taaacggcca caagttcagc gtgtccggcg agggcgaggg cgatgccacc tacggcaagc 2340

tgaccctgaa gttcatctgc accaccggca agctgcccgt gccctggccc accctcgtga 2400

ccaccctgac ctacggcgtg cagtgcttca gccgctaccc cgaccacatg aagcagcacg 2460

acttcttcaa gtccgccatg cccgaaggct acgtccagga gcgcaccatc ttcttcaagg 2520

acgacggcaa ctacaagacc cgcgccgagg tgaagttcga gggcgacacc ctggtgaacc 2580

gcatcgagct gaagggcatc gacttcaagg aggacggcaa catcctgggg cacaagctgg 2640

agtacaacta caacagccac aacgtctata tcatggccga caagcagaag aacggcatca 2700

aggtgaactt caagatccgc cacaacatcg aggacggcag cgtgcagctc gccgaccact 2760

accagcagaa cacccccatc ggcgacggcc ccgtgctgct gcccgacaac cactacctga 2820

gcacccagtc cgccctgagc aaagacccca acgagaagcg cgatcacatg gtcctgctgg 2880

agttcgtgac cgccgccggg atcactctcg gcatggacga gctgtacaag taatgattcg 2940

aaatgaccga ccaagcgacg cccaacctgc catcacgaga tttcgattcc accgccgcct 3000

tctatgaaag gttgggcttc ggaatcgttt tccgggacgc cggctggatg atcctccagc 3060

gcggggatct catgctggag ttcttcgccc accccaactt gtttattgca gcttataatg 3120

gttacaaata aagcaatagc atcacaaatt tcacaaataa agcatttttt tcactgcatt 3180

ctagttgtgg tttgtccaaa ctcatcaatg tatcttatca tgtctgt 3227

<210> 43

<211> 2743

<212> DNA

<213> Artificial Sequence

<220><223> GSC2781_Construct

<400> 43

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctggctc 1680

aggtcaggag gatcaggagg acgaggagga agaggagacc ggtgccacca tggtgagcaa 1740

gggcgaggag ctgttcaccg gggtggtgcc catcctggtc gagctggacg gcgacgtaaa 1800

cggccacaag ttcagcgtgt ccggcgaggg cgagggcgat gccacctacg gcaagctgac 1860

cctgaagttc atctgcacca ccggcaagct gcccgtgccc tggcccaccc tcgtgaccac 1920

cctgacctac ggcgtgcagt gcttcagccg ctaccccgac cacatgaagc agcacgactt 1980

cttcaagtcc gccatgcccg aaggctacgt ccaggagcgc accatcttct tcaaggacga 2040

cggcaactac aagacccgcg ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat 2100

cgagctgaag ggcatcgact tcaaggagga cggcaacatc ctggggcaca agctggagta 2160

caactacaac agccacaacg tctatatcat ggccgacaag cagaagaacg gcatcaaggt 2220

gaacttcaag atccgccaca acatcgagga cggcagcgtg cagctcgccg accactacca 2280

gcagaacacc cccatcggcg acggccccgt gctgctgccc gacaaccact acctgagcac 2340

ccagtccgcc ctgagcaaag accccaacga gaagcgcgat cacatggtcc tgctggagtt 2400

cgtgaccgcc gccgggatca ctctcggcat ggacgagctg tacaagtaat gattcgaaat 2460

gaccgaccaa gcgacgccca acctgccatc acgagatttc gattccaccg ccgccttcta 2520

tgaaaggttg ggcttcggaa tcgttttccg ggacgccggc tggatgatcc tccagcgcgg 2580

ggatctcatg ctggagttct tcgcccaccc caacttgttt attgcagctt ataatggtta 2640

caaataaagc aatagcatca caaatttcac aaataaagca ttttttttcac tgcattctag 2700

ttgtggtttg tccaaactca tcaatgtatc ttatcatgtc tgt 2743

<210> 44

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2342_Construct

<400> 44

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctgtctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gaccctgtcc ttattgcaca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

tttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 45

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2328_Construct

<400> 45

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctgtctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gacctctgcg cttcttgcct tgcctgctcc ctggctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400

tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460

gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640

gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700

aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 46

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2321_Construct

<400> 46

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctgtctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg caccttttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gacctctgcg cttcttccct tccctcctcc ctggctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400

tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460

gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640

gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700

aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 47

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2324_Construct

<400> 47

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctgtctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gacctctgcg cttcttccct tccctgctac ctggctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400

tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460

gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640

gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700

aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 48

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2339_Construct

<400> 48

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttgttg ccttccctcc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg caccttttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gaccctgtcc ttattgcaca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 49

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2334_Construct

<400> 49

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttgttg ccttccctcc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

cccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gacctctgcg cttcttccct tccctcctcc ctggctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400

tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460

gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640

gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700

aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 50

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2336_Construct

<400> 50

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag ggggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttgttg ccttccctcc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gacctctgcg cttcttccct tccctgctac ctggctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400

tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460

gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640

gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700

aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 51

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2340_Construct

<400> 51

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctgc tacctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gcccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gaccctgtcc ttattgcaca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 52

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2331_Construct

<400> 52

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctgc tacctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gacctctgcg cttcttccct tccctcctcc ctgtctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400

tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460

gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640

gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700

aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 53

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2453_Construct

<400> 53

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctgc tacctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740
tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800
tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860
ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920
tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980
agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040
ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160
ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220
gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280
gacctctgcg cttcttgcct tgcctgctcc ctggctcagg tcaggaggat caggaggacg 2340
aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400
tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460
gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580
tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640
gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700
aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760
aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820
atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880
tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000
ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060
tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120
tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180
ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240
cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300
atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 54

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2325_Construct

<400> 54

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctgc tacctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560
agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620
gtagttactg caccttttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680
cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740
tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800
tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860
ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980
agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040
ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100
agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160
ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220
gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280
gacctctgcg cttcttccct tccctcctcc ctggctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400
tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460
gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520
gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580
tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640
gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700
aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820
atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880
tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940
gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000
ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060
tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120
tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 55

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2332_Construct

<400> 55

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctgc tacctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gcccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380
cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440
actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500
actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560
agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620
gtagttactg caccttttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680
cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800
tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860
ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920
tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980
agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040
ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100
agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220
gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280
gacctctgcg cttcttccct tccctgctac ctggctcagg tcaggaggat caggaggacg 2340
aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400
tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460
gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520
gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640
gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700
aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760
aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820
atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880
tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940
gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 56

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2341_Construct

<400> 56

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttg ccttgcctgc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200
agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260
ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320
acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380
cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440
actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500
actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620
gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680
cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740
tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800
tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860
ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920
tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040
ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100
agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160
ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220
gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280
gaccctgtcc ttattgcaca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340
gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460
ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520
ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580
acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640
ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700
acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760
tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 57

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2326_Construct

<400> 57

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttg ccttgcctgc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg caccttttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gacctctgcg cttcttgcct tgcctgctcc ctggctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400

tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460

gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640

gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700

aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 58

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2454_Construct

<400> 58

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttg ccttgcctgc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gacctctgcg cttcttccct tccctcctcc ctggctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400

tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460

gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640

gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700

aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 59

<211> 3380

<212> DNA

<213> Artificial Sequence

<220><223> GSC2327_Construct

<400> 59

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttg ccttgcctgc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gacctctgcg cttcttccct tccctgctac ctggctcagg tcaggaggat caggaggacg 2340

aggaggaaga ggagaccggt gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg 2400

tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 2460

gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2520

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2580

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2640

gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2700

aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2760

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2820

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2880

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2940

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 3000

ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3060

tcggcatgga cgagctgtac aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc 3120

tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 3180

ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 3240

cccaccccaa cttgtttatt gcagcttata atggttacaa ataaagcaat agcatcacaa 3300

atttcacaaa taaagcattt ttttcactgc attctagttg tggtttgtcc aaactcatca 3360

atgtatctta tcatgtctgt 3380

<210> 60

<211> 3344

<212> DNA

<213> Artificial Sequence

<220><223> GSC2338_Construct

<400> 60

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag ggggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740

ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800

acccctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860

cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920

aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980

atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg agggggcac 2040

agctggggag ggaacaaggg acaaaaccag gaggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160

tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220

gggggtctgt agggagccag caggagctgc ggccgtccta ctgaccctgt ccttattgca 2280

caggtcagga ggatcaggag gacgaggagg aagaggagac cggtgccacc atggtgagca 2340

agggcgagga gctgttcacc ggggtggtgc ccatcctggt cgagctggac ggcgacgtaa 2400

acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac ggcaagctga 2460

ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc ctcgtgacca 2520

ccctgaccta cggcgtgcag tgcttcagcc gctaccccga ccacatgaag cagcacgact 2580

tcttcaagtc cgccatgccc gaaggctacg tccaggagcg caccatcttc ttcaaggacg 2640

acggcaacta caagacccgc gccgaggtga agttcgaggg cgacaccctg gtgaaccgca 2700

tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac aagctggagt 2760

acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac ggcatcaagg 2820

tgaacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc gaccactacc 2880

agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac tacctgagca 2940

cccagtccgc cctgagcaaa gaccccaacg agaagcgcga tcacatggtc ctgctggagt 3000

tcgtgaccgc cgccgggatc actctcggca tggacgagct gtacaagtaa tgattcgaaa 3060

tgaccgacca agcgacgccc aacctgccat cacgagattt cgattccacc gccgccttct 3120

atgaaaggtt gggcttcgga atcgttttcc gggacgccgg ctggatgatc ctccagcgcg 3180

gggatctcat gctggagttc ttcgcccacc ccaacttgtt tattgcagct tataatggtt 3240

acaaataaag caatagcatc acaaatttca caaataaagc attttttca ctgcattcta 3300

gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctgt 3344

<210> 61

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2335_Construct

<400> 61

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740

ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800

acccctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860

cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920

aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980

atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg agggggggcac 2040

agctggggag ggaacaaggg acaaaaccag gagggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160

tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220

gggggtctgt agggagccag caggagctgc ggccgtccta ctgacctctg cgcttcttcc 2280

cttccctcct ccctgtctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 62

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2333_Construct

<400> 62

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag ggggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgccccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740

ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800

accctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860

cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920

aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980

atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg agggggcac 2040

agctggggag ggaacaaggg acaaaaccag gaggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160

tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220

gggggtctgt agggagccag caggagctgc ggccgtccta ctgacctctg cgcttcttgc 2280

cttgcctgct ccctggctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 63

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2337_Construct

<400> 63

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740

ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800

acccctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860

cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920

aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980

atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg agggggcac 2040

agctggggag ggaacaaggg acaaaaccag gaggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160

tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220

gggggtctgt agggagccag caggagctgc ggccgtccta ctgacctctg cgcttcttcc 2280

cttccctcct ccctggctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 64

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2322_Construct

<400> 64

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740

ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800

acccctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860

cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920

aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980

atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg agggggcac 2040

agctggggag ggaacaaggg acaaaaccag gaggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160

tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220

gggggtctgt agggagccag caggagctgc ggccgtccta ctgacctctg cgcttcttcc 2280

cttccctgct acctggctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 65

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2617_Construct

<400> 65

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg ggaggattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740
ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800
acccctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860
cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920
aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980
atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg agggggcac 2040
agctggggag ggaacaaggg acaaaaccag gaggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160
tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220
gggggtctgt agggagccag caggagctgc ggccgtccta ctgacctctg cgcttcttcc 2280
cttccctcct ccctggctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340
gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400
agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460
ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580
acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640
ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700
acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760
tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820
agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880
agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000
acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060
acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120
attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180
ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240
ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300
ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 66

<211> 3209

<212> DNA

<213> Artificial Sequence

<220><223> GSC2739_Construct

<400> 66

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg ggaggattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560
agggccgcca ccacctgttc ctgtagtaac ggaagaattc agggtaggtg atcctcctgc 1620
tgctttggtt cagggttttg cttgaggggg gggggtggtg atttccttgc catgggcaga 1680
ctgagcagaa aaggccattg ggaccatgtt ctgaatgcct ccacctcaac caccggccgg 1740
taggaccaaa gccacccccgt gttttctcag gatctctttt cccagggaga tccctcggcc 1800
caaagaggga gatggcaatg ctggatgtgt gcacaataat tcaacaggca ttggaacttc 1860
agcatcgatg ctgaatgcaa ttaacaatgc tcaagcagaa cccccggctc catcagcaca 1920

gtgcaggacc aaaccccatg ctgcagcagt ggggctgtct gtacggggtg ggcaatggga 1980
accggggtct gctggggctc ctgctgcttc agtgctgcca tgcagccaca catcctgaga 2040
gctgaaaggg tcggcgtcct cacctggtgc acaccgtagc tctgccccac agctttaagg 2100
cacctggcta acctctgcgc ttcttccctt ccctcctccc tggctcaggt caggaggatc 2160
aggaggacga ggaggaagag gagaccggtg ccaccatggt gagcaagggc gaggagctgt 2220
tcaccggggt ggtgcccatc ctggtcgagc tggacggcga cgtaaacggc cacaagttca 2280
gcgtgtccgg cgagggcgag ggcgatgcca cctacggcaa gctgaccctg aagttcatct 2340

gcaccaccgg caagctgccc gtgccctggc ccaccctcgt gaccaccctg acctacggcg 2400
tgcagtgctt cagccgctac cccgaccaca tgaagcagca cgacttcttc aagtccgcca 2460
tgcccgaagg ctacgtccag gagcgcacca tcttcttcaa ggacgacggc aactacaaga 2520
cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa ccgcatcgag ctgaagggca 2580
tcgacttcaa ggaggacggc aacatcctgg ggcacaagct ggagtacaac tacaacagcc 2640
acaacgtcta tatcatggcc gacaagcaga agaacggcat caaggtgaac ttcaagatcc 2700
gccacaacat cgaggacggc agcgtgcagc tcgccgacca ctaccagcag aacacccccа 2760

tcggcgacgg ccccgtgctg ctgcccgaca accactacct gagcacccag tccgccctga 2820
gcaaagaccc caacgagaag cgcgatcaca tggtcctgct ggagttcgtg accgccgccg 2880
ggatcactct cggcatggac gagctgtaca agtaatgatt cgaaatgacc gaccaagcga 2940
cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa aggttgggct 3000
tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat ctcatgctgg 3060
agttcttcgc ccaccccaac ttgtttattg cagcttataa tggttacaaa taaagcaata 3120
gcatcacaaa tttcacaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca 3180

aactcatcaa tgtatcttat catgtctgt 3209

<210> 67

<211> 2725

<212> DNA

<213> Artificial Sequence

<220><223> GSC2782_Construct

<400> 67

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg ggaggattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgccccc 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc agccacagct ttaaggcacc 1620

tggctaacct ctgcgcttct tcccttccct cctccctggc tcaggtcagg aggatcagga 1680

ggacgaggag gaagaggaga ccggtgccac catggtgagc aagggcgagg agctgttcac 1740

cggggtggtg cccatcctgg tcgagctgga cggcgacgta aacggccaca agttcagcgt 1800

gtccggcgag ggcgagggcg atgccaccta cggcaagctg accctgaagt tcatctgcac 1860

caccggcaag ctgcccgtgc cctggcccac cctcgtgacc accctgacct acggcgtgca 1920

gtgcttcagc cgctaccccg accacatgaa gcagcacgac ttcttcaagt ccgccatgcc 1980

cgaaggctac gtccaggagc gcaccatctt cttcaaggac gacggcaact acaagacccg 2040

cgccgaggtg aagttcgagg gcgacaccct ggtgaaccgc atcgagctga agggcatcga 2100

cttcaaggag gacggcaaca tcctggggca caagctggag tacaactaca acagccacaa 2160

cgtctatatc atggccgaca agcagaagaa cggcatcaag gtgaacttca agatccgcca 2220

caacatcgag gacggcagcg tgcagctcgc cgaccactac cagcagaaca cccccatcgg 2280

cgacggcccc gtgctgctgc ccgacaacca ctacctgagc acccagtccg ccctgagcaa 2340

agaccccaac gagaagcgcg atcacatggt cctgctggag ttcgtgaccg ccgccgggat 2400

cactctcggc atggacgagc tgtacaagta atgattcgaa atgaccgacc aagcgacgcc 2460

caacctgcca tcacgagatt tcgattccac cgccgccttc tatgaaaggt tgggcttcgg 2520

aatcgttttc cgggacgccg gctggatgat cctccagcgc ggggatctca tgctggagtt 2580

cttcgcccac cccaacttgt ttattgcagc ttataatggt tacaaataaa gcaatagcat 2640

cacaaatttc acaaataaag catttttttc actgcattct agttgtggtt tgtccaaact 2700

catcaatgta tcttatcatg tctgt 2725

<210> 68

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2621_Construct

<400> 68

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag ggaggaggga acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740

ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800

acccctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860

cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920

aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980

atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg agggggggcac 2040

agctggggag ggaacaaggg acaaaaccag gagggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160

tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220

gggggtctgt agggagccag caggagctgc ggccgtccta ctgacctctg cgcttcttcc 2280

cttccctcct ccctggctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 69

<211> 3209

<212> DNA

<213> Artificial Sequence

<220><223> GSC2740_Construct

<400> 69

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag ggaggaggga acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc agggtaggtg atcctcctgc 1620

tgctttggtt cagggttttg cttgaggggg gggggtggtg atttccttgc catgggcaga 1680

ctgagcagaa aaggccattg ggaccatgtt ctgaatgcct ccacctcaac caccggccgg 1740

taggaccaaa gccaccccgt gttttctcag gatctctttt cccagggaga tccctcggcc 1800

caaagaggga gatggcaatg ctggatgtgt gcacaataat tcaacaggca ttggaacttc 1860

agcatcgatg ctgaatgcaa ttaacaatgc tcaagcagaa ccccccggctc catcagcaca 1920

gtgcaggacc aaaccccatg ctgcagcagt ggggctgtct gtacggggtg ggcaatggga 1980

accggggtct gctggggctc ctgctgcttc agtgctgcca tgcagccaca catcctgaga 2040

gctgaaaggg tcggcgtcct cacctggtgc acaccgtagc tctgccccac agctttaagg 2100

cacctggcta acctctgcgc ttcttccctt ccctcctccc tggctcaggt caggaggatc 2160

aggaggacga ggaggaagag gagaccggtg ccaccatggt gagcaagggc gaggagctgt 2220

tcaccggggt ggtgcccatc ctggtcgagc tggacggcga cgtaaacggc cacaagttca 2280

gcgtgtccgg cgagggcgag ggcgatgcca cctacggcaa gctgaccctg aagttcatct 2340

gcaccaccgg caagctgccc gtgccctggc ccaccctcgt gaccaccctg acctacggcg 2400

tgcagtgctt cagccgctac cccgaccaca tgaagcagca cgacttcttc aagtccgcca 2460

tgcccgaagg ctacgtccag gagcgcacca tcttcttcaa ggacgacggc aactacaaga 2520

cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa ccgcatcgag ctgaagggca 2580

tcgacttcaa ggaggacggc aacatcctgg ggcacaagct ggagtacaac tacaacagcc 2640

acaacgtcta tatcatggcc gacaagcaga agaacggcat caaggtgaac ttcaagatcc 2700

gccacaacat cgaggacggc agcgtgcagc tcgccgacca ctaccagcag aacaccccca 2760

tcggcgacgg ccccgtgctg ctgcccgaca accactacct gagcacccag tccgccctga 2820

gcaaagaccc caacgagaag cgcgatcaca tggtcctgct ggagttcgtg accgccgccg 2880

ggatcactct cggcatggac gagctgtaca agtaatgatt cgaaatgacc gaccaagcga 2940

cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa aggttgggct 3000

tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat ctcatgctgg 3060

agttcttcgc ccaccccaac ttgtttattg cagcttataa tggttacaaa taaagcaata 3120

gcatcacaaa tttcacaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca 3180

aactcatcaa tgtatcttat catgtctgt 3209

<210> 70

<211> 2725

<212> DNA

<213> Artificial Sequence

<220><223> GSC2783_Construct

<400> 70

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag ggaggaggga acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc agccacagct ttaaggcacc 1620

tggctaacct ctgcgcttct tcccttccct cctccctggc tcaggtcagg aggatcagga 1680

ggacgaggag gaagaggaga ccggtgccac catggtgagc aagggcgagg agctgttcac 1740

cggggtggtg cccatcctgg tcgagctgga cggcgacgta aacggccaca agttcagcgt 1800

gtccggcgag ggcgagggcg atgccaccta cggcaagctg accctgaagt tcatctgcac 1860
caccggcaag ctgcccgtgc cctggcccac cctcgtgacc accctgacct acggcgtgca 1920
gtgcttcagc cgctaccccg accacatgaa gcagcacgac ttcttcaagt ccgccatgcc 1980
cgaaggctac gtccaggagc gcaccatctt cttcaaggac gacggcaact acaagacccg 2040
cgccgaggtg aagttcgagg gcgacaccct ggtgaaccgc atcgagctga agggcatcga 2100
cttcaaggag gacggcaaca tcctggggca caagctggag tacaactaca acagccacaa 2160
cgtctatatc atggccgaca agcagaagaa cggcatcaag gtgaacttca agatccgcca 2220

caacatcgag gacggcagcg tgcagctcgc cgaccactac cagcagaaca cccccatcgg 2280
cgacggcccc gtgctgctgc ccgacaacca ctacctgagc acccagtccg ccctgagcaa 2340
agaccccaac gagaagcgcg atcacatggt cctgctggag ttcgtgaccg ccgccgggat 2400
cactctcggc atggacgagc tgtacaagta atgattcgaa atgaccgacc aagcgacgcc 2460
caacctgcca tcacgagatt tcgattccac cgccgccttc tatgaaaggt tgggcttcgg 2520
aatcgttttc cgggacgccg gctggatgat cctccagcgc ggggatctca tgctggagtt 2580
cttcgcccac cccaacttgt ttattgcagc ttataatggt tacaaataaa gcaatagcat 2640

cacaaatttc acaaataaag catttttttc actgcattct agttgtggtt tgtccaaact 2700
catcaatgta tcttatcatg tctgt 2725

<210> 71
<211> 3362
<212> DNA
<213> Artificial Sequence
<220><223> GSC2622_Construct
<400> 71

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60
gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120
cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240
ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300
attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360
ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420
ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag tccttagggg acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgccccc 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740

ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800

acccctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860

cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920

aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980

atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg agggggggcac 2040

agctggggag ggaacaaggg acaaaaccag gaggggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160

tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220

gggggtctgt agggagccag caggagctgc ggccgtccta ctgacctctg cgcttcttcc 2280

cttccctcct ccctggctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 72

<211> 3209

<212> DNA

<213> Artificial Sequence

<220><223> GSC2742_Construct

<400> 72

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag tccttagggg acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgccccc 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc agggtaggtg atcctcctgc 1620

tgctttggtt cagggttttg cttgaggggg gggggtggtg atttccttgc catgggcaga 1680

ctgagcagaa aaggccattg ggaccatgtt ctgaatgcct ccacctcaac caccggccgg 1740

taggaccaaa gccaccccgt gttttctcag gatctctttt cccagggaga tccctcggcc 1800

caaagaggga gatggcaatg ctggatgtgt gcacaataat tcaacaggca ttggaacttc 1860

agcatcgatg ctgaatgcaa ttaacaatgc tcaagcagaa cccccggctc catcagcaca 1920

gtgcaggacc aaaccccatg ctgcagcagt ggggctgtct gtacggggtg ggcaatggga 1980

accggggtct gctggggctc ctgctgcttc agtgctgcca tgcagccaca catcctgaga 2040

gctgaaaggg tcggcgtcct cacctggtgc acaccgtagc tctgccccac agctttaagg 2100

cacctggcta acctctgcgc ttcttccctt ccctcctccc tggctcaggt caggaggatc 2160

aggaggacga ggaggaagag gagaccggtg ccaccatggt gagcaagggc gaggagctgt 2220

tcaccggggt ggtgcccatc ctggtcgagc tggacggcga cgtaaacggc cacaagttca 2280

gcgtgtccgg cgagggcgag ggcgatgcca cctacggcaa gctgaccctg aagttcatct 2340

gcaccaccgg caagctgccc gtgccctggc ccaccctcgt gaccaccctg acctacggcg 2400

tgcagtgctt cagccgctac cccgaccaca tgaagcagca cgacttcttc aagtccgcca 2460

tgcccgaagg ctacgtccag gagcgcacca tcttcttcaa ggacgacggc aactacaaga 2520

cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa ccgcatcgag ctgaagggca 2580

tcgacttcaa ggaggacggc aacatcctgg ggcacaagct ggagtacaac tacaacagcc 2640

acaacgtcta tatcatggcc gacaagcaga agaacggcat caaggtgaac ttcaagatcc 2700

gccacaacat cgaggacggc agcgtgcagc tcgccgacca ctaccagcag aacacccccа 2760

tcggcgacgg ccccgtgctg ctgcccgaca accactacct gagcacccag tccgccctga 2820

gcaaagaccc caacgagaag cgcgatcaca tggtcctgct ggagttcgtg accgccgccg 2880

ggatcactct cggcatggac gagctgtaca agtaatgatt cgaaatgacc gaccaagcga 2940

cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa aggttgggct 3000

tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat ctcatgctgg 3060

agttcttcgc ccaccccaac ttgtttattg cagcttataa tggttacaaa taaagcaata 3120

gcatcacaaa tttcacaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca 3180

aactcatcaa tgtatcttat catgtctgt 3209

<210> 73

<211> 2725

<212> DNA

<213> Artificial Sequence

<220><223> GSC2784_Construct

<400> 73

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag tccttagggg acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgccccc 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc agccacagct ttaaggcacc 1620

tggctaacct ctgcgcttct tcccttccct cctccctggc tcaggtcagg aggatcagga 1680

ggacgaggag gaagaggaga ccggtgccac catggtgagc aagggcgagg agctgttcac 1740

cggggtggtg cccatcctgg tcgagctgga cggcgacgta aacggccaca agttcagcgt 1800

gtccggcgag ggcgagggcg atgccaccta cggcaagctg accctgaagt tcatctgcac 1860

caccggcaag ctgcccgtgc cctggcccac cctcgtgacc accctgacct acggcgtgca 1920

gtgcttcagc cgctaccccg accacatgaa gcagcacgac ttcttcaagt ccgccatgcc 1980

cgaaggctac gtccaggagc gcaccatctt cttcaaggac gacggcaact acaagacccg 2040

cgccgaggtg aagttcgagg gcgacaccct ggtgaaccgc atcgagctga agggcatcga 2100

cttcaaggag gacggcaaca tcctggggca caagctggag tacaactaca acagccacaa 2160

cgtctatatc atggccgaca agcagaagaa cggcatcaag gtgaacttca agatccgcca 2220

caacatcgag gacggcagcg tgcagctcgc cgaccactac cagcagaaca cccccatcgg 2280

cgacggcccc gtgctgctgc ccgacaacca ctacctgagc acccagtccg ccctgagcaa 2340

agaccccaac gagaagcgcg atcacatggt cctgctggag ttcgtgaccg ccgccgggat 2400

cactctcggc atggacgagc tgtacaagta atgattcgaa atgaccgacc aagcgacgcc 2460

caacctgcca tcacgagatt tcgattccac cgccgccttc tatgaaaggt tgggcttcgg 2520

aatcgttttc cgggacgccg gctggatgat cctccagcgc ggggatctca tgctggagtt 2580

cttcgcccac cccaacttgt ttattgcagc ttataatggt tacaaataaa gcaatagcat 2640

cacaaatttc acaaataaag catttttttc actgcattct agttgtggtt tgtccaaact 2700

catcaatgta tcttatcatg tctgt 2725

<210> 74

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2620_Construct

<400> 74

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gcctacgctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740

ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800

acccctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860

cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920

aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980

atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg aggggggcac 2040

agctggggag ggaacaaggg acaaaaccag gagggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160

tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220

gggggtctgt agggagccag caggagctgc ggccgtccta ctgacctctg cgcttcttcc 2280

cttccctcct ccctggctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640
ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700
acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760
tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820
agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880
agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940
acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060
acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120
attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180
ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240
ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300
ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360
gt 3362

<210> 75
<211> 3209
<212> DNA
<213> Artificial Sequence
<220><223> GSC2737_Construct
<400> 75
ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60
gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120
cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180
tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240
ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360
ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420
ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480
caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540
gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gcctacgctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc agggtaggtg atcctcctgc 1620

tgctttggtt cagggttttg cttgaggggg gggggtggtg atttccttgc catgggcaga 1680

ctgagcagaa aaggccattg ggaccatgtt ctgaatgcct ccacctcaac caccggccgg 1740

taggaccaaa gccacccccgt gttttctcag gatctctttt cccagggaga tccctcggcc 1800

caaagaggga gatggcaatg ctggatgtgt gcacaataat tcaacaggca ttggaacttc 1860

agcatcgatg ctgaatgcaa ttaacaatgc tcaagcagaa cccccggctc catcagcaca 1920

gtgcaggacc aaaccccatg ctgcagcagt ggggctgtct gtacggggtg ggcaatggga 1980

accggggtct gctggggctc ctgctgcttc agtgctgcca tgcagccaca catcctgaga 2040

gctgaaaggg tcggcgtcct cacctggtgc acaccgtagc tctgccccac agctttaagg 2100

cacctggcta acctctgcgc ttcttccctt ccctcctccc tggctcaggt caggaggatc 2160

aggaggacga ggaggaagag gagaccggtg ccaccatggt gagcaagggc gaggagctgt 2220

tcaccggggt ggtgcccatc ctggtcgagc tggacggcga cgtaaacggc cacaagttca 2280

gcgtgtccgg cgagggcgag ggcgatgcca cctacggcaa gctgaccctg aagttcatct 2340

gcaccaccgg caagctgccc gtgccctggc ccaccctcgt gaccaccctg acctacggcg 2400

tgcagtgctt cagccgctac cccgaccaca tgaagcagca cgacttcttc aagtccgcca 2460
tgcccgaagg ctacgtccag gagcgcacca tcttcttcaa ggacgacggc aactacaaga 2520
cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa ccgcatcgag ctgaagggca 2580
tcgacttcaa ggaggacggc aacatcctgg ggcacaagct ggagtacaac tacaacagcc 2640
acaacgtcta tatcatggcc gacaagcaga agaacggcat caaggtgaac ttcaagatcc 2700
gccacaacat cgaggacggc agcgtgcagc tcgccgacca ctaccagcag aacaccccca 2760
tcggcgacgg ccccgtgctg ctgcccgaca accactacct gagcacccag tccgccctga 2820

gcaaagaccc caacgagaag cgcgatcaca tggtcctgct ggagttcgtg accgccgccg 2880
ggatcactct cggcatggac gagctgtaca agtaatgatt cgaaatgacc gaccaagcga 2940
cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa aggttgggct 3000
tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat ctcatgctgg 3060
agttcttcgc ccaccccaac ttgtttattg cagcttataa tggttacaaa taaagcaata 3120
gcatcacaaa tttcacaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca 3180
aactcatcaa tgtatcttat catgtctgt 3209

<210> 76
<211> 3362
<212> DNA
<213> Artificial Sequence
<220><223> GSC2615_Construct
<400> 76
ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60
gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120
cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180
tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240
ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360
ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420
ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480
caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540
gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gagaaccctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740

ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800

acccctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860

cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920

aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980

atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg agggggcac 2040

agctggggag ggaacaaggg acaaaaccag gaggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160

tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220

gggggtctgt agggagccag caggagctgc ggccgtccta ctgacctctg cgcttcttcc 2280

cttccctcct ccctggctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460
ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520
ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580
acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640
ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700
acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760
tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880
agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940
acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000
acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060
acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120
attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180
ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300
ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360
gt 3362

<210> 77
<211> 3209
<212> DNA
<213> Artificial Sequence
<220><223> GSC2743_Construct
<400> 77

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60
gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180
tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240
ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300
attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360
ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gcttgccctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc agggtaggtg atcctcctgc 1620

tgctttggtt cagggttttg cttgaggggg gggggtggtg atttccttgc catgggcaga 1680

ctgagcagaa aaggccattg ggaccatgtt ctgaatgcct ccacctcaac caccggccgg 1740

taggaccaaa gccaccccgt gttttctcag gatctctttt cccagggaga tccctcggcc 1800

caaagaggga gatggcaatg ctggatgtgt gcacaataat tcaacaggca ttggaacttc 1860

agcatcgatg ctgaatgcaa ttaacaatgc tcaagcagaa cccccggctc catcagcaca 1920

gtgcaggacc aaaccccatg ctgcagcagt ggggctgtct gtacggggtg ggcaatggga 1980

accggggtct gctggggctc ctgctgcttc agtgctgcca tgcagccaca catcctgaga 2040

gctgaaaggg tcggcgtcct cacctggtgc acaccgtagc tctgccccac agctttaagg 2100

cacctggcta acctctgcgc ttcttccctt ccctcctccc tggctcaggt caggaggatc 2160

aggaggacga ggaggaagag gagaccggtg ccaccatggt gagcaagggc gaggagctgt 2220

tcaccggggt ggtgcccatc ctggtcgagc tggacggcga cgtaaacggc cacaagttca 2280
gcgtgtccgg cgagggcgag ggcgatgcca cctacggcaa gctgaccctg aagttcatct 2340
gcaccaccgg caagctgccc gtgccctggc ccaccctcgt gaccaccctg acctacggcg 2400
tgcagtgctt cagccgctac cccgaccaca tgaagcagca cgacttcttc aagtccgcca 2460
tgcccgaagg ctacgtccag gagcgcacca tcttcttcaa ggacgacggc aactacaaga 2520
cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa ccgcatcgag ctgaagggca 2580
tcgacttcaa ggaggacggc aacatcctgg ggcacaagct ggagtacaac tacaacagcc 2640

acaacgtcta tatcatggcc gacaagcaga agaacggcat caaggtgaac ttcaagatcc 2700
gccacaacat cgaggacggc agcgtgcagc tcgccgacca ctaccagcag aacaccccca 2760
tcggcgacgg ccccgtgctg ctgcccgaca accactacct gagcacccag tccgccctga 2820
gcaaagaccc caacgagaag cgcgatcaca tggtcctgct ggagttcgtg accgccgccg 2880
ggatcactct cggcatggac gagctgtaca agtaatgatt cgaaatgacc gaccaagcga 2940
cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa aggttgggct 3000
tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat ctcatgctgg 3060

agttcttcgc ccaccccaac ttgtttattg cagcttataa tggttacaaa taaagcaata 3120
gcatcacaaa tttcacaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca 3180
aactcatcaa tgtatcttat catgtctgt 3209

<210> 78
<211> 3209
<212> DNA
<213> Artificial Sequence
<220><223> GSC2738_Construct
<400> 78

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60
gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180
tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240
ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300
attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360
ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gagaaccctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc agggtaggtg atcctcctgc 1620

tgctttggtt cagggttttg cttgaggggg gggggtggtg atttccttgc catgggcaga 1680

ctgagcagaa aaggccattg ggaccatgtt ctgaatgcct ccacctcaac caccggccgg 1740

taggaccaaa gccaccccgt gttttctcag gatctctttt cccagggaga tccctcggcc 1800

caaagaggga gatggcaatg ctggatgtgt gcacaataat tcaacaggca ttggaacttc 1860

agcatcgatg ctgaatgcaa ttaacaatgc tcaagcagaa cccccggctc catcagcaca 1920

gtgcaggacc aaaccccatg ctgcagcagt ggggctgtct gtacggggtg ggcaatggga 1980

accggggtct gctggggctc ctgctgcttc agtgctgcca tgcagccaca catcctgaga 2040

gctgaaaggg tcggcgtcct cacctggtgc acaccgtagc tctgccccac agctttaagg 2100

cacctggcta acctctgcgc ttcttccctt ccctcctccc tggctcaggt caggaggatc 2160

aggaggacga ggaggaagag gagaccggtg ccaccatggt gagcaagggc gaggagctgt 2220

tcaccggggt ggtgcccatc ctggtcgagc tggacggcga cgtaaacggc cacaagttca 2280
gcgtgtccgg cgagggcgag ggcgatgcca cctacggcaa gctgaccctg aagttcatct 2340
gcaccaccgg caagctgccc gtgccctggc ccaccctcgt gaccaccctg acctacggcg 2400
tgcagtgctt cagccgctac cccgaccaca tgaagcagca cgacttcttc aagtccgcca 2460
tgcccgaagg ctacgtccag gagcgcacca tcttcttcaa ggacgacggc aactacaaga 2520
cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa ccgcatcgag ctgaagggca 2580
tcgacttcaa ggaggacggc aacatcctgg ggcacaagct ggagtacaac tacaacagcc 2640

acaacgtcta tatcatggcc gacaagcaga agaacggcat caaggtgaac ttcaagatcc 2700
gccacaacat cgaggacggc agcgtgcagc tcgccgacca ctaccagcag aacaccccca 2760
tcggcgacgg ccccgtgctg ctgcccgaca accactacct gagcacccag tccgccctga 2820
gcaaagaccc caacgagaag cgcgatcaca tggtcctgct ggagttcgtg accgccgccg 2880
ggatcactct cggcatggac gagctgtaca agtaatgatt cgaaatgacc gaccaagcga 2940
cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa aggttgggct 3000
tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat ctcatgctgg 3060

agttcttcgc ccaccccaac ttgtttattg cagcttataa tggttacaaa taaagcaata 3120
gcatcacaaa tttcacaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca 3180
aactcatcaa tgtatcttat catgtctgt 3209

<210> 79
<211> 3362
<212> DNA
<213> Artificial Sequence
<220><223> GSC2618_Construct
<400> 79

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60
gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180
tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240
ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300
attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360
ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaggctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740

ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800

acccctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860

cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920

aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980

atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg agggggcac 2040

agctggggag ggaacaaggg acaaaaccag gaggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160

tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220

gggggtctgt agggagccag caggagctgc ggccgtccta ctgacctctg cgcttcttcc 2280

cttccctcct ccctggctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 80

<211> 3209

<212> DNA

<213> Artificial Sequence

<220><223> GSC2975_Construct

<400> 80

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaggctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc agggtaggtg atcctcctgc 1620

tgctttggtt cagggttttg cttgaggggg gggggtggtg atttccttgc catgggcaga 1680

ctgagcagaa aaggccattg ggaccatgtt ctgaatgcct ccacctcaac caccggccgg 1740

taggaccaaa gccaccccgt gttttctcag gatctctttt cccagggaga tccctcggcc 1800

caaagaggga gatggcaatg ctggatgtgt gcacaataat tcaacaggca ttggaacttc 1860

agcatcgatg ctgaatgcaa ttaacaatgc tcaagcagaa cccccggctc catcagcaca 1920

gtgcaggacc aaaccccatg ctgcagcagt ggggctgtct gtacggggtg ggcaatggga 1980

accggggtct gctggggctc ctgctgcttc agtgctgcca tgcagccaca catcctgaga 2040

gctgaaaggg tcggcgtcct cacctggtgc acaccgtagc tctgccccac agctttaagg 2100

cacctggcta acctctgcgc ttcttccctt ccctcctccc tggctcaggt caggaggatc 2160

aggaggacga ggaggaagag gagaccggtg ccaccatggt gagcaagggc gaggagctgt 2220

tcaccggggt ggtgcccatc ctggtcgagc tggacggcga cgtaaacggc cacaagttca 2280

gcgtgtccgg cgagggcgag ggcgatgcca cctacggcaa gctgaccctg aagttcatct 2340

gcaccaccgg caagctgccc gtgccctggc ccaccctcgt gaccaccctg acctacggcg 2400

tgcagtgctt cagccgctac cccgaccaca tgaagcagca cgacttcttc aagtccgcca 2460

tgcccgaagg ctacgtccag gagcgcacca tcttcttcaa ggacgacggc aactacaaga 2520

cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa ccgcatcgag ctgaagggca 2580

tcgacttcaa ggaggacggc aacatcctgg ggcacaagct ggagtacaac tacaacagcc 2640

acaacgtcta tatcatggcc gacaagcaga agaacggcat caaggtgaac ttcaagatcc 2700

gccacaacat cgaggacggc agcgtgcagc tcgccgacca ctaccagcag aacacccccа 2760

tcggcgacgg ccccgtgctg ctgcccgaca accactacct gagcacccag tccgccctga 2820

gcaaagaccc caacgagaag cgcgatcaca tggtcctgct ggagttcgtg accgccgccg 2880

ggatcactct cggcatggac gagctgtaca agtaatgatt cgaaatgacc gaccaagcga 2940

cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa aggttgggct 3000

tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat ctcatgctgg 3060

agttcttcgc ccaccccaac ttgtttattg cagcttataa tggttacaaa taaagcaata 3120

gcatcacaaa tttcacaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca 3180

aactcatcaa tgtatcttat catgtctgt 3209

<210> 81

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2613_Construct

<400> 81

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg tccttattgc agaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc aggtagttac tgcacctttc 1620

tttgttccat ctctccacct ctgctgtgaa taaatcgcgg gtcggtgtgt cctgtgcctt 1680

tccctgcttg ggaaacgctt tcctttcatt ctttcacttc tctgctgctt tttgcgctct 1740

ccccatcctg ctgtgccaac ctgctctcag ttctgtgctt tctgtcttcc atcccaacac 1800

acccctgggt tgctgtcttc tttctccttt cttcctctct tgctgtggga ccaaacgtct 1860

cctgcaggac ctgcgggctc tgacagagga ctctcgtggg ggtactgctc cctccagtgg 1920

aaaaatgctc cagcagtgtc atgcaggaga tttatgccat acagttttgc tctctgctgc 1980

atggagggga gcagcagaag tcgatctccc ccactctggg gtccccctcg agggggggcac 2040

agctggggag ggaacaaggg acaaaaccag gaggggggctc cgagtccttg gatttattcc 2100

ccctcatcca tgccttacct tcaggtaagg gcctgaacag agccctttac ttcctgcttc 2160

tttctcccat agctccctct cttcgggtct cctggactca gtgccacggt tgtccattct 2220

gggggtctgt agggagccag caggagctgc ggccgtccta ctgacctctg cgcttcttcc 2280

cttccctcct ccctggctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460

ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520

ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580

acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640

ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700

acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760

tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880

agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940

acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000

acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060

acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120

attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180

ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 82

<211> 3497

<212> DNA

<213> Artificial Sequence

<220><223> GSC2614_Construct

<400> 82

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagagtag 360

ttactgcacc tttctttgtt ccatctctcc acctctgctg tgaataaatc gcgggtcggt 420

gtgtcctgtg cctttccctg cttgggaaac gctttccttt cattctttca cttctctgct 480

gctttttgcg ctctccccat cctgctgtgc caacctgctc tcagttctgt gctttctgtc 540

ttccatccca acacacccct gggttgctgt cttctttctc ctttcttcct ctcttgctgt 600

gggaccaaac gtctcctgca ggacctgcgg gctctgacag aggactctcg tgggggtact 660

gctccctcca gtggaaaaat gctccagcag tgtcatgcag gagatttatg ccatacagtt 720

ttgctctctg ctgcatggag gggagcagca gaagtcgatc tcccccactc tggggtcccc 780

ctcgaggggg gcacagctgg ggagggaaca agggacaaaa ccaggagggg gctccgagtc 840

cttggattta ttccccctca tccatgcctt accttcaggt aagggcctga acagagccct 900

ttacttcctg cttctttctc ccatagctcc ctctcttcgg gtctcctgga ctcagtgcca 960

cggttgtcca ttctgggggt ctgtagggag ccagcaggag ctgcggccgt cctactgacc 1020

ctgtccttat tgcacaggat ccaggcgata tcgccaccat gggtgcctcc tccgaggacg 1080

tcatcaagga gttcatgcgc ttcaaggtgc gcatggaggg ctccgtgaac ggccacgagt 1140

tcgagatcga gggcgagggc gagggccgcc cctacgaggg cacccagacc gccaagctga 1200

aggtgaccaa gggcggcccc ctgcccttcg cctgggacat cctgtccccc cagttccagt 1260

acggctccaa ggtgtacgtg aagcaccccg ccgacatccc cgactacaag aagctgtcct 1320

tccccgaggg cttcaagtgg gagcgcgtga tgaacttcga ggacggcggc gtggtgaccg 1380

tgacccagga ctcctccctg caggacggct ccttcatcta caaggtgaag ttcatcggcg 1440

tgaacttccc ctccgacggc cccgtaatgc agaagaagac tatgggctgg gaggcctcca 1500

ccgagcgcct gtacccccgc gacggcgtgc tgaagggcga gatccacaag gccctgaagc 1560

tgaaggacgg cggccactac ctggtggagt tcaagtccat ctacatggcc aagaagcccg 1620

tgcagctgcc cggctactac tacgtggact ccaagctgga catcacctcc cacaacgagg 1680

actacaccat cgtggagcag tacgagcgcg ccgagggccg ccaccacctg ttcctgtagt 1740

aacggaagaa ttcaggtagt tactgcacct ttctttgttc catctctcca cctctgctgt 1800

gaataaatcg cgggtcggtg tgtcctgtgc ctttccctgc ttgggaaacg ctttcctttc 1860

attctttcac ttctctgctg ctttttgcgc tctccccatc ctgctgtgcc aacctgctct 1920
cagttctgtg ctttctgtct tccatcccaa cacacccctg ggttgctgtc ttctttctcc 1980
tttcttcctc tcttgctgtg ggaccaaacg tctcctgcag gacctgcggg ctctgacaga 2040
ggactctcgt gggggtactg ctccctccag tggaaaaatg ctccagcagt gtcatgcagg 2100
agatttatgc catacagttt tgctctctgc tgcatggagg ggagcagcag aagtcgatct 2160
cccccactct ggggtccccc tcgaggggggg cacagctggg gagggaacaa gggacaaaac 2220
caggaggggg ctccgagtcc ttggatttat tccccctcat ccatgcctta ccttcaggta 2280

agggcctgaa cagagccctt tacttcctgc ttctttctcc catagctccc tctcttcggg 2340
tctcctggac tcagtgccac ggttgtccat tctgggggtc tgtagggagc cagcaggagc 2400
tgcggccgtc ctactgaccc tgtccttatt gcacaggtca ggaggatcag gaggacgagg 2460
aggaagagga gaccggtgcc accatggtga gcaagggcga ggagctgttc accggggtgg 2520
tgcccatcct ggtcgagctg gacggcgacg taaacggcca caagttcagc gtgtccggcg 2580
agggcgaggg cgatgccacc tacggcaagc tgaccctgaa gttcatctgc accaccggca 2640
agctgcccgt gccctggccc accctcgtga ccaccctgac ctacggcgtg cagtgcttca 2700

gccgctaccc cgaccacatg aagcagcacg acttcttcaa gtccgccatg cccgaaggct 2760
acgtccagga gcgcaccatc ttcttcaagg acgacggcaa ctacaagacc cgcgccgagg 2820
tgaagttcga gggcgacacc ctggtgaacc gcatcgagct gaagggcatc gacttcaagg 2880
aggacggcaa catcctgggg cacaagctgg agtacaacta caacagccac aacgtctata 2940
tcatggccga caagcagaag aacggcatca aggtgaactt caagatccgc cacaacatcg 3000
aggacggcag cgtgcagctc gccgaccact accagcagaa cacccccatc ggcgacggcc 3060
ccgtgctgct gcccgacaac cactacctga gcacccagtc cgccctgagc aaagacccca 3120

acgagaagcg cgatcacatg gtcctgctgg agttcgtgac cgccgccggg atcactctcg 3180
gcatggacga gctgtacaag taatgattcg aaatgaccga ccaagcgacg cccaacctgc 3240
catcacgaga tttcgattcc accgccgcct tctatgaaag gttgggcttc ggaatcgttt 3300
tccgggacgc cggctggatg atcctccagc gcggggatct catgctggag ttcttcgccc 3360
accccaactt gtttattgca gcttataatg gttacaaata aagcaatagc atcacaaatt 3420
tcacaaataa agcatttttt tcactgcatt ctagttgtgg tttgtccaaa ctcatcaatg 3480
tatcttatca tgtctgt 3497

<210> 83

<211> 3362

<212> DNA

<213> Artificial Sequence

<220><223> GSC2741_Construct

<400> 83

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagagtag 360

ttactgcacc tttctttgtt ccatctctcc acctctgctg tgaataaatc gcgggtcggt 420

gtgtcctgtg cctttccctg cttgggaaac gctttccttt cattctttca cttctctgct 480

gctttttgcg ctctccccat cctgctgtgc caacctgctc tcagttctgt gctttctgtc 540

ttccatccca acacacccct gggttgctgt cttctttctc ctttcttcct ctcttgctgt 600

gggaccaaac gtctcctgca ggacctgcgg gctctgacag aggactctcg tgggggtact 660

gctccctcca gtggaaaaat gctccagcag tgtcatgcag gagatttatg ccatacagtt 720

ttgctctctg ctgcatggag gggagcagca gaagtcgatc tcccccactc tggggtcccc 780

ctcgaggggg gcacagctgg ggagggaaca agggacaaaa ccaggagggg gctccgagtc 840

cttggattta ttccccctca tccatgcctt accttcaggt aagggcctga acagagccct 900

ttacttcctg cttctttctc ccatagctcc ctctcttcgg gtctcctgga ctcagtgcca 960

cggttgtcca ttctgggggt ctgtagggag ccagcaggag ctgcggccgt cctactgacc 1020

ctgtccttat tgcacaggat ccaggcgata tcgccaccat gggtgcctcc tccgaggacg 1080

tcatcaagga gttcatgcgc ttcaaggtgc gcatggaggg ctccgtgaac ggccacgagt 1140

tcgagatcga gggcgagggc gagggccgcc cctacgaggg cacccagacc gccaagctga 1200

aggtgaccaa gggcggcccc ctgcccttcg cctgggacat cctgtccccc cagttccagt 1260

acggctccaa ggtgtacgtg aagcaccccg ccgacatccc cgactacaag aagctgtcct 1320

tccccgaggg cttcaagtgg gagcgcgtga tgaacttcga ggacggcggc gtggtgaccg 1380

tgacccagga ctcctccctg caggacggct ccttcatcta caaggtgaag ttcatcggcg 1440

tgaacttccc ctccgacggc cccgtaatgc agaagaagac tatgggctgg gaggcctcca 1500

ccgagcgcct gtacccccgc gacggcgtgc tgaagggcga gatccacaag gccctgaagc 1560

tgaaggacgg cggccactac ctggtggagt tcaagtccat ctacatggcc aagaagcccg 1620
tgcagctgcc cggctactac tacgtggact ccaagctgga catcacctcc cacaacgagg 1680
actacaccat cgtggagcag tacgagcgcg ccgagggccg ccaccacctg ttcctgtagt 1740
aacggaagaa ttcagggtag gtgatcctcc tgctgctttg gttcagggtt ttgcttgagg 1800
ggggggggtg gtgatttcct tgccatgggc agactgagca gaaaaggcca ttgggaccat 1860
gttctgaatg cctccacctc aaccaccggc cggtaggacc aaagccaccc cgtgttttct 1920
caggatctct tttcccaggg agatccctcg gcccaaagag ggagatggca atgctggatg 1980

tgtgcacaat aattcaacag gcattggaac ttcagcatcg atgctgaatg caattaacaa 2040
tgctcaagca gaaccccccgg ctccatcagc acagtgcagg accaaacccc atgctgcagc 2100
agtggggctg tctgtacggg gtgggcaatg ggaaccgggg tctgctgggg ctcctgctgc 2160
ttcagtgctg ccatgcagcc acacatcctg agagctgaaa gggtcggcgt cctcacctgg 2220
tgcacaccgt agctctgccc cacagcttta aggcacctgg ctaacctctg cgcttcttcc 2280
cttccctcct ccctggctca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340
gtgccaccat ggtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg 2400

agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg 2460
ccacctacgg caagctgacc ctgaagttca tctgcaccac cggcaagctg cccgtgccct 2520
ggcccaccct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccccgacc 2580
acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca 2640
ccatcttctt caaggacgac ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg 2700
acaccctggt gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc 2760
tggggcacaa gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc 2820

agaagaacgg catcaaggtg aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 2880
agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgcccg 2940
acaaccacta cctgagcacc cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc 3000
acatggtcct gctggagttc gtgaccgccg ccgggatcac tctcggcatg gacgagctgt 3060
acaagtaatg attcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg 3120
attccaccgc cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct 3180
ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 3240

ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 3300

ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 3360

gt 3362

<210> 84

<211> 2743

<212> DNA

<213> Artificial Sequence

<220><223> GSC2780_Construct

<400> 84

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctggctc 900

aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960

tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gcccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctggctc 1680

aggtcaggag gatcaggagg acgaggagga agaggagacc ggtgccacca tggtgagcaa 1740

gggcgaggag ctgttcaccg gggtggtgcc catcctggtc gagctggacg gcgacgtaaa 1800

cggccacaag ttcagcgtgt ccggcgaggg cgagggcgat gccacctacg gcaagctgac 1860

cctgaagttc atctgcacca ccggcaagct gcccgtgccc tggcccaccc tcgtgaccac 1920

cctgacctac ggcgtgcagt gcttcagccg ctaccccgac cacatgaagc agcacgactt 1980

cttcaagtcc gccatgcccg aaggctacgt ccaggagcgc accatcttct tcaaggacga 2040

cggcaactac aagacccgcg ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat 2100

cgagctgaag ggcatcgact tcaaggagga cggcaacatc ctggggcaca agctggagta 2160

caactacaac agccacaacg tctatatcat ggccgacaag cagaagaacg gcatcaaggt 2220

gaacttcaag atccgccaca acatcgagga cggcagcgtg cagctcgccg accactacca 2280

gcagaacacc cccatcggcg acggccccgt gctgctgccc gacaaccact acctgagcac 2340

ccagtccgcc ctgagcaaag accccaacga gaagcgcgat cacatggtcc tgctggagtt 2400

cgtgaccgcc gccgggatca ctctcggcat ggacgagctg tacaagtaat gattcgaaat 2460

gaccgaccaa gcgacgccca acctgccatc acgagatttc gattccaccg ccgccttcta 2520

tgaaaggttg ggcttcggaa tcgttttccg ggacgccggc tggatgatcc tccagcgcgg 2580

ggatctcatg ctggagttct tcgcccaccc caacttgttt attgcagctt ataatggtta 2640

caaataaagc aatagcatca caaatttcac aaataaagca ttttttttcac tgcattctag 2700

ttgtggtttg tccaaactca tcaatgtatc ttatcatgtc tgt 2743

<210> 85

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y)_Construct

<400> 85

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 86

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(9Y nude)_Construct

<400> 86

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgtc tccttctggg acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 87

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(7Y nude)_Construct

<400> 87

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgac tccttcgggg acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 88

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5)_Construct

<400> 88

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg ggaggattgc acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 89

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y nude)_Construct

<400> 89

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540
gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600
ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660
gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720
ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780
cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840
ccacagcttt aaggcacctg gctaaccgag tccttagggg acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 90
<211> 908
<212> DNA
<213> Artificial Sequence
<220><223> I4(3Y nude)_Construct
<400> 90

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60
gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120
cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180
tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300
attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360
ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420
ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480
caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540
gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600
ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720
ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780
cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840
ccacagcttt aaggcacctg gctaaccgag acctgagggg acaggatcca ggcgatatcg 900
ccaccatg 908

<210> 91

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(1Y nude)_Construct

<400>

91

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag agcagagggg acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 92

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y)_Construct

<400> 92

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag ggaggaggga acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 93

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-b-ct)_Construct

<400> 93

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gagaaccctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 94

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-b-y)_Construct

<400> 94

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaggctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 95

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-b-2)_Construct

<400> 95

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gcctacgctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 96

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-b-a)_Construct

<400> 96

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gcttgccctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 97

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-A)_Construct

<400> 97

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg tccttattgc aaaggatcca ggcgatatcg 900

ccaccatg 908

<210> 98

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5,G)_Construct

<400> 98

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg ggaggattgc agaggatcca ggcgatatcg 900

ccaccatg 908

<210> 99

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude,A)_Construct

<400> 99

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag tccttagggg aaaggatcca ggcgatatcg 900

ccaccatg 908

<210> 100

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude,b-2)_Construct

<400> 100

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gcctacggag tccttagggg acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 101

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude,A)_Construct

<400> 101

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag tccttagggg aaaggatcca ggcgatatcg 900

ccaccatg 908

<210> 102

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5,G)_Construct

<400> 102

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg ggaggattgc agaggatcca ggcgatatcg 900

ccaccatg 908

<210> 103

<211> 612

<212> DNA

<213> Artificial Sequence

<220><223> cTNT-I4_LC-HC_Construct

<400> 103

taacggaaga attcagggta ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag 60

gggggggggt ggtgatttcc ttgccatggg cagactgagc agaaaaggcc attgggacca 120

tgttctgaat gcctccacct caaccaccgg ccggtaggac caaagccacc ccgtgttttc 180

tcaggatctc ttttcccagg gagatccctc ggcccaaaga gggagatggc aatgctggat 240

gtgtgcacaa taattcaaca ggcattggaa cttcagcatc gatgctgaat gcaattaaca 300

atgctcaagc agaacccccg gctccatcag cacagtgcag gaccaaaccc catgctgcag 360

cagtggggct gtctgtacgg ggtgggcaat gggaaccggg gtctgctggg gctcctgctg 420

cttcagtgct gccatgcagc cacacatcct gagagctgaa agggtcggcg tcctcacctg 480

gtgcacaccg tagctctgcc ccacagcttt aaggcacctg gctaacctct gcgcttcttc 540

ccttccctcc tccctggctc aggtcaggag gatcaggagg acgaggagga agaggagacc 600

ggtgccacca tg 612

<210> 104

<211> 747

<212> DNA

<213> Artificial Sequence

<220><223> cTNT-I5_LC-HC_Construct

<400> 104

taacggaaga attcaggtag ttactgcacc tttctttgtt ccatctctcc acctctgctg 60

tgaataaatc gcgggtcggt gtgtcctgtg cctttccctg cttgggaaac gctttccttt 120

cattctttca cttctctgct gctttttgcg ctctccccat cctgctgtgc caacctgctc 180

tcagttctgt gctttctgtc ttccatccca acacacccct gggttgctgt cttctttctc 240

ctttcttcct ctcttgctgt gggaccaaac gtctcctgca ggacctgcgg gctctgacag 300

aggactctcg tgggggtact gctccctcca gtggaaaaat gctccagcag tgtcatgcag 360

gagatttatg ccatacagtt ttgctctctg ctgcatggag gggagcagca gaagtcgatc 420

tcccccactc tggggtcccc ctcgaggggg gcacagctgg ggagggaaca agggacaaaa 480

ccaggagggg gctccgagtc cttggattta ttccccctca tccatgcctt accttcaggt 540

aagggcctga acagagccct ttacttcctg cttctttctc ccatagctcc ctctcttcgg 600

gtctcctgga ctcagtgcca cggttgtcca ttctgggggt ctgtagggag ccagcaggag 660

ctgcggccgt cctactgacc ctgtccttat tgcacaggtc aggaggatca ggaggacgag 720

gaggaagagg agaccggtgc caccatg 747

<210> 105

<211> 128

<212> DNA

<213> Artificial Sequence

<220><223> I4(sh)_LC-HC_Construct

<400> 105

taacggaaga attcagccac agctttaagg cacctggcta acctctgcgc ttcttccctt 60

ccctcctccc tggctcaggt caggaggatc aggaggacga ggaggaagag gagaccggtg 120

ccaccatg 128

<210> 106

<211> 612

<212> DNA

<213> Artificial Sequence

<220><223> cTNT-I4_HC-LC_Construct

<400> 106

taacggaaga attcagggta ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag 60

gggggggggt ggtgatttcc ttgccatggg cagactgagc agaaaaggcc attgggacca 120

tgttctgaat gcctccacct caaccaccgg ccggtaggac caaagccacc ccgtgttttc 180

tcaggatctc ttttcccagg gagatccctc ggcccaaaga gggagatggc aatgctggat 240

gtgtgcacaa taattcaaca ggcattggaa cttcagcatc gatgctgaat gcaattaaca 300

atgctcaagc agaaccccg gctccatcag cacagtgcag gaccaaaccc catgctgcag 360

cagtggggct gtctgtacgg ggtgggcaat gggaaccggg gtctgctggg gctcctgctg 420

cttcagtgct gccatgcagc cacacatcct gagagctgaa agggtcggcg tcctcacctg 480

gtgcacaccg tagctctgcc ccacagcttt aaggcacctg gctaacctct gcgcttcttc 540

ccttccctcc tccctggctc aggtcaggag gatcaggagg acgaggagga agaggagacc 600

ggtgccacca tg 612

<210> 107

<211> 747

<212> DNA

<213> Artificial Sequence

<220><223> cTNT-I5_HC-LC_Construct

<400> 107

taacggaaga attcaggtag ttactgcacc tttctttgtt ccatctctcc acctctgctg 60

tgaataaatc gcgggtcggt gtgtcctgtg cctttccctg cttgggaaac gctttccttt 120

cattctttca cttctctgct gctttttgcg ctctccccat cctgctgtgc caacctgctc 180

tcagttctgt gctttctgtc ttccatccca acacacccct gggttgctgt cttctttctc 240

ctttcttcct ctcttgctgt gggaccaaac gtctcctgca ggacctgcgg gctctgacag 300

aggactctcg tgggggtact gctccctcca gtggaaaaat gctccagcag tgtcatgcag 360

gagatttatg ccatacagtt ttgctctctg ctgcatggag gggagcagca gaagtcgatc 420

tcccccactc tggggtcccc ctcgaggggg gcacagctgg ggagggaaca agggacaaaa 480

ccaggagggg gctccgagtc cttggattta ttccccctca tccatgcctt accttcaggt 540

aagggcctga acagagccct ttacttcctg cttctttctc ccatagctcc ctctcttcgg 600

gtctcctgga ctcagtgcca cggttgtcca ttctgggggt ctgtagggag ccagcaggag 660

ctgcggccgt cctactgacc ctgtccttat tgcacaggtc aggaggatca ggaggacgag 720

gaggaagagg agaccggtgc caccatg 747

<210> 108

<211> 128

<212> DNA

<213> Artificial Sequence

<220><223> I4(sh)_HC-LC_Construct

<400> 108

taacggaaga attcagccac agctttaagg cacctggcta acctctgcgc ttcttccctt 60

ccctcctccc tggctcaggt caggaggatc aggaggacga ggaggaagag gagaccggtg 120

ccaccatg 128

<210> 109

<211> 3209

<212> DNA

<213> Artificial Sequence

<220><223> GSC2975_Construct

<400> 109

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaggctg tccttattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc agggtaggtg atcctcctgc 1620

tgctttggtt cagggttttg cttgaggggg gggggtggtg atttccttgc catgggcaga 1680

ctgagcagaa aaggccattg ggaccatgtt ctgaatgcct ccacctcaac caccggccgg 1740

taggaccaaa gccaccccgt gttttctcag gatctctttt cccagggaga tccctcggcc 1800

caaagaggga gatggcaatg ctggatgtgt gcacaataat tcaacaggca ttggaacttc 1860

agcatcgatg ctgaatgcaa ttaacaatgc tcaagcagaa cccccggctc catcagcaca 1920

gtgcaggacc aaaccccatg ctgcagcagt ggggctgtct gtacggggtg ggcaatggga 1980

accggggtct gctggggctc ctgctgcttc agtgctgcca tgcagccaca catcctgaga 2040

gctgaaaggg tcggcgtcct cacctggtgc acaccgtagc tctgccccac agctttaagg 2100

cacctggcta acctctgcgc ttcttccctt ccctcctccc tggctcaggt caggaggatc 2160

aggaggacga ggaggaagag gagaccggtg ccaccatggt gagcaagggc gaggagctgt 2220

tcaccggggt ggtgcccatc ctggtcgagc tggacggcga cgtaaacggc cacaagttca 2280

gcgtgtccgg cgagggcgag ggcgatgcca cctacggcaa gctgaccctg aagttcatct 2340

gcaccaccgg caagctgccc gtgccctggc ccaccctcgt gaccaccctg acctacggcg 2400

tgcagtgctt cagccgctac cccgaccaca tgaagcagca cgacttcttc aagtccgcca 2460

tgcccgaagg ctacgtccag gagcgcacca tcttcttcaa ggacgacggc aactacaaga 2520

cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa ccgcatcgag ctgaagggca 2580

tcgacttcaa ggaggacggc aacatcctgg ggcacaagct ggagtacaac tacaacagcc 2640

acaacgtcta tatcatggcc gacaagcaga agaacggcat caaggtgaac ttcaagatcc 2700

gccacaacat cgaggacggc agcgtgcagc tcgccgacca ctaccagcag aacaccccca 2760

tcggcgacgg ccccgtgctg ctgcccgaca accactacct gagcacccag tccgccctga 2820

gcaaagaccc caacgagaag cgcgatcaca tggtcctgct ggagttcgtg accgccgccg 2880
ggatcactct cggcatggac gagctgtaca agtaatgatt cgaaatgacc gaccaagcga 2940
cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa aggttgggct 3000
tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat ctcatgctgg 3060
agttcttcgc ccaccccaac ttgtttattg cagcttataa tggttacaaa taaagcaata 3120
gcatcacaaa tttcacaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca 3180
aactcatcaa tgtatcttat catgtctgt 3209

<210> 110
<211> 3360
<212> DNA
<213> Artificial Sequence
<220><223> GSC2223_Construct
<400> 110
ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60
gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120
cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180
tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240
ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360
ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420
ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480
caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540
gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600
ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660
gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780
cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840
ccacagcttt aaggcacctg gctaacctct gcgcttcttc ccttccctcc tccctggctc 900
aggatccagg cgatatcgcc accatgggtg cctcctccga ggacgtcatc aaggagttca 960
tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagttcgag atcgagggcg 1020

agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg accaagggcg 1080

gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc tccaaggtgt 1140

acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 1200

agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct 1260

ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg 1320

acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgcctgtacc 1380

cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc 1440

actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtgcag ctgcccggct 1500

actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 1560

agcagtacga gcgcgccgag ggccgccacc acctgttcct gtagtaacgg aagaattcag 1620

gtagttactg cacctttctt tgttccatct ctccacctct gctgtgaata aatcgcgggt 1680

cggtgtgtcc tgtgcctttc cctgcttggg aaacgctttc ctttcattct ttcacttctc 1740

tgctgctttt tgcgctctcc ccatcctgct gtgccaacct gctctcagtt ctgtgctttc 1800

tgtcttccat cccaacacac ccctgggttg ctgtcttctt tctcctttct tcctctcttg 1860

ctgtgggacc aaacgtctcc tgcaggacct gcgggctctg acagaggact ctcgtggggg 1920

tactgctccc tccagtggaa aaatgctcca gcagtgtcat gcaggagatt tatgccatac 1980

agttttgctc tctgctgcat ggaggggagc agcagaagtc gatctccccc actctggggt 2040

ccccctcgag gggggcacag ctggggaggg aacaagggac aaaaccagga gggggctccg 2100

agtccttgga tttattcccc ctcatccatg ccttaccttc aggtaagggc ctgaacagag 2160

ccctttactt cctgcttctt tctcccatag ctccctctct tcgggtctcc tggactcagt 2220

gccacggttg tccattctgg gggtctgtag ggagccagca ggagctgcgg ccgtcctact 2280

gaccctgtcc ttattgcaca ggtcaggagg atcaggagga cgaggaggaa gaggagaccg 2340

gtgccaccat ggagcaaggg cgaggagctg ttcaccgggg tggtgcccat cctggtcgag 2400

ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg gcgagggcga gggcgatgcc 2460

acctacggca agctgaccct gaagttcatc tgcaccaccg gcaagctgcc cgtgccctgg 2520

cccaccctcg tgaccaccct gacctacggc gtgcagtgct tcagccgcta ccccgaccac 2580

atgaagcagc acgacttctt caagtccgcc atgcccgaag gctacgtcca ggagcgcacc 2640

atcttcttca aggacgacgg caactacaag acccgcgccg aggtgaagtt cgagggcgac 2700

accctggtga accgcatcga gctgaagggc atcgacttca aggaggacgg caacatcctg 2760

gggcacaagc tggagtacaa ctacaacagc cacaacgtct atatcatggc cgacaagcag 2820

aagaacggca tcaaggtgaa cttcaagatc cgccacaaca tcgaggacgg cagcgtgcag 2880
ctcgccgacc actaccagca gaacaccccc atcggcgacg gccccgtgct gctgcccgac 2940
aaccactacc tgagcaccca gtccgccctg agcaaagacc ccaacgagaa gcgcgatcac 3000
atggtcctgc tggagttcgt gaccgccgcc gggatcactc tcggcatgga cgagctgtac 3060
aagtaatgat tcgaaatgac cgaccaagcg acgcccaacc tgccatcacg agatttcgat 3120
tccaccgccg ccttctatga aaggttgggc ttcggaatcg ttttccggga cgccggctgg 3180
atgatcctcc agcgcgggga tctcatgctg gagttcttcg cccaccccaa cttgtttatt 3240

gcagcttata atggttacaa ataaagcaat agcatcacaa atttcacaaa taaagcattt 3300
ttttcactgc attctagttg tggtttgtcc aaactcatca atgtatctta tcatgtctgt 3360
3360

<210> 111
<211> 3209
<212> DNA
<213> Artificial Sequence
<220><223> GSC3166_Construct
<400> 111

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60
gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180
tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240
ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300
attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360
ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420
ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480
caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600
ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660
gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720
ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780
cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg ggaggattgc acaggatcca ggcgatatcg 900

ccaccatggg tgcctcctcc gaggacgtca tcaaggagtt catgcgcttc aaggtgcgca 960

tggagggctc cgtgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 1020

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cggccccctg cccttcgcct 1080

gggacatcct gtcccccag ttccagtacg gctccaaggt gtacgtgaag caccccgccg 1140

acatccccga ctacaagaag ctgtccttcc ccgagggctt caagtgggag cgcgtgatga 1200

acttcgagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag gacggctcct 1260

tcatctacaa ggtgaagttc atcggcgtga acttcccctc cgacggcccc gtaatgcaga 1320

agaagactat gggctgggag gcctccaccg agcgcctgta cccccgcgac ggcgtgctga 1380

agggcgagat ccacaaggcc ctgaagctga aggacggcgg ccactacctg gtggagttca 1440

agtccatcta catggccaag aagcccgtgc agctgcccgg ctactactac gtggactcca 1500

agctggacat cacctcccac aacgaggact acaccatcgt ggagcagtac gagcgcgccg 1560

agggccgcca ccacctgttc ctgtagtaac ggaagaattc agggtaggtg atcctcctgc 1620

tgctttggtt cagggttttg cttgaggggg gggggtggtg atttccttgc catgggcaga 1680

ctgagcagaa aaggccattg ggaccatgtt ctgaatgcct ccacctcaac caccggccgg 1740

taggaccaaa gccaccccgt gttttctcag gatctctttt cccagggaga tccctcggcc 1800

caaagaggga gatggcaatg ctggatgtgt gcacaataat tcaacaggca ttggaacttc 1860

agcatcgatg ctgaatgcaa ttaacaatgc tcaagcagaa cccccggctc catcagcaca 1920

gtgcaggacc aaaccccatg ctgcagcagt ggggctgtct gtacggggtg ggcaatggga 1980

accggggtct gctggggctc ctgctgcttc agtgctgcca tgcagccaca catcctgaga 2040

gctgaaaggg tcggcgtcct cacctggtgc acaccgtagc tctgccccac agctttaagg 2100

cacctggcta acctctgcgc ttcttccctt ccctcctccc tggctcaggt caggaggatc 2160

aggaggacga ggaggaagag gagaccggtg ccaccatggt gagcaagggc gaggagctgt 2220

tcaccggggt ggtgcccatc ctggtcgagc tggacggcga cgtaaacggc cacaagttca 2280

gcgtgtccgg cgagggcgag ggcgatgcca cctacggcaa gctgaccctg aagttcatct 2340

gcaccaccgg caagctgccc gtgccctggc ccaccctcgt gaccaccctg acctacggcg 2400

tgcagtgctt cagccgctac cccgaccaca tgaagcagca cgacttcttc aagtccgcca 2460

tgcccgaagg ctacgtccag gagcgcacca tcttcttcaa ggacgacggc aactacaaga 2520

cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa ccgcatcgag ctgaagggca 2580

tcgacttcaa ggaggacggc aacatcctgg ggcacaagct ggagtacaac tacaacagcc 2640

acaacgtcta tatcatggcc gacaagcaga agaacggcat caaggtgaac ttcaagatcc 2700

gccacaacat cgaggacggc agcgtgcagc tcgccgacca ctaccagcag aacacccccа 2760

tcggcgacgg ccccgtgctg ctgcccgaca accactacct gagcacccag tccgccctga 2820

gcaaagaccc caacgagaag cgcgatcaca tggtcctgct ggagttcgtg accgccgccg 2880

ggatcactct cggcatggac gagctgtaca agtaatgatt cgaaatgacc gaccaagcga 2940

cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa aggttgggct 3000

tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat ctcatgctgg 3060

agttcttcgc ccaccccaac ttgtttattg cagcttataa tggttacaaa taaagcaata 3120

gcatcacaaa tttcacaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca 3180

aactcatcaa tgtatcttat catgtctgt 3209

<210> 112

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y; b-a)_Construct

<400> 112

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gcttgccgag ggaggaggga acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 113

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y; b-ct) _Construct

<400> 113

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gagaaccgag ggaggaggga acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 114

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y; b-y) _Construct

<400> 114

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaagggag ggaggaggga acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 115

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y, b-2) _Construct

<400> 115

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gcctacggag ggaggaggga acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 116

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y, A) _Construct

<400>

116

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaacccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag ggaggaggga aaaggatcca ggcgatatcg 900

ccaccatg 908

<210> 117

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y, T) _Construct

<400> 117

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag ggaggaggga ataggatcca ggcgatatcg 900

ccaccatg 908

<210> 118

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y, G) _Construct

<400> 118

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag ggaggaggga agaggatcca ggcgatatcg 900

ccaccatg 908

<210> 119

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude; b-a) _Construct

<400> 119

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gcttgccgag tccttagggg acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 120

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude; b-ct) _Construct

<400> 120

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaagggag tccttagggg acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 121

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude; b-y) _Construct

<400> 121

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaagggag tccttagggg acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 122

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude, T) _Construct

<400> 122

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccgag tccttagggg ataggatcca ggcgatatcg 900

ccaccatg 908

<210> 123

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5, b-a) _Construct

<

400> 123

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gcttgccctg ggaggattgc acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 124

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5, b-ct) _Construct

<400> 124

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gagaaccctg ggaggattgc acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 125

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5;b-y) _Construct

<400> 125

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaggctg ggaggattgc acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 126

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5, b-2) _Construct

<400> 126

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gcctacgctg ggaggattgc acaggatcca ggcgatatcg 900

ccaccatg 908

<210> 127

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5, A) _Construct

<400> 127

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg ggaggattgc aaaggatcca ggcgatatcg 900

ccaccatg 908

<210> 128

<211> 908

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5, T) _Construct

<400> 128

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc 60

gcggccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtgac gtaagtaccg 120

cctatagagt ctataggccc acccccttgg cttcttatgc gacggatccc gtactaagct 180

tgaggtgtgg caggcttgag atctggccat acacttgagt gacaatgaca tccactttgc 240

ctttctctcc acaggtgtcc actcccacgt ccaactgcag ctcggttcga tcgataatta 300

attaagctag cgtttaaact taagcttcct tggaggaccc agtacccgga tctagaggta 360

ggtgatcctc ctgctgcttt ggttcagggt tttgcttgag gggggggggt ggtgatttcc 420

ttgccatggg cagactgagc agaaaaggcc attgggacca tgttctgaat gcctccacct 480

caaccaccgg ccggtaggac caaagccacc ccgtgttttc tcaggatctc ttttcccagg 540

gagatccctc ggcccaaaga gggagatggc aatgctggat gtgtgcacaa taattcaaca 600

ggcattggaa cttcagcatc gatgctgaat gcaattaaca atgctcaagc agaaccccg 660

gctccatcag cacagtgcag gaccaaaccc catgctgcag cagtggggct gtctgtacgg 720

ggtgggcaat gggaaccggg gtctgctggg gctcctgctg cttcagtgct gccatgcagc 780

cacacatcct gagagctgaa agggtcggcg tcctcacctg gtgcacaccg tagctctgcc 840

ccacagcttt aaggcacctg gctaaccctg ggaggattgc ataggatcca ggcgatatcg 900

ccaccatg 908

<210> 129

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> I4_Flanking intron

<400> 129

ctaacctctg cgcttcttcc cttccctcct ccctggctca g 41

<210> 130

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> I4(22Y+1) _Flanking intron

<400> 130

ctaacctctg cgcttcttcc cttccctcct ccctgtctca g 41

<210> 131

<211

> 41

<212> DNA

<213> Artificial Sequence

<220><223> I4(15Y-5')_Flanking intron

<400> 131

ctaacctctg cgcttgttgc cttccctcct ccctggctca g 41

<210> 132

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> I4(15Y-3')_Flanking intron

<400> 132

ctaacctctg cgcttcttcc cttccctgct acctggctca g 41

<210> 133

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> I4(22Y-3) _Flanking intron

<400> 133

ctaacctctg cgcttcttgc cttgcctgct ccctggctca g 41

<210> 134

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y)_Flanking intron

<400> 134

ctaaccctgt ccttattgca cag 23

<210> 135

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5) _Flanking intron

<400> 135

ctaaccctgg gaggattgca cag 23

<210> 136

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y) _Flanking intron

<400> 136

ctaaccgagg gaggagggaa cag 23

<210> 137

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude) _Flanking intron

<400> 137

ctaaccgagt ccttagggga cag 23

<210> 138

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-b-2) _Flanking intron

<400> 138

cctacgctgt ccttattgca cag 23

<210> 139

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-b-a) _Flanking intron

<400> 139

cttgccctgt ccttattgca cag 23

<210> 140

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-b-ct) _Flanking intron

<400> 140

agaaccctgt ccttattgca cag 23

<210> 141

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-b-y) _Flanking intron

<400> 141

ctaaggctgt ccttattgca cag 23

<210> 142

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-G) _Flanking intron

<400> 142

ctaaccctgt ccttattgca gag 23

<210> 143

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-A) _Flanking intron

<400> 143

ctaaccctgt ccttattgca aag 23

<210> 144

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5-G) _Flanking intron

<400> 144

ctaaccctgg gaggattgca gag 23

<210> 145

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude-A) _Flanking intron

<400> 145

ctaaccgagt ccttagggga aag 23

<210> 146

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude-b-2) _Flanking intron

<400> 146

cctacggagt ccttagggga cag 23

<210> 147

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(9Ynude) _Flanking intron

<400> 147

ctaaccgtct ccttctggga cag 23

<210> 148

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223

> I4(7Ynude) _Flanking intron

<400> 148

ctaaccgact ccttcgggga cag 23

<210> 149

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude-b-a) _Flanking intron

<400> 149

cttgccgagt ccttagggga cag 23

<210> 150

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(3Ynude) _Flanking intron

<400> 150

ctaaccgaga cctgagggga cag 23

<210> 151

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(1Ynude) _Flanking intron

<400> 151

ctaaccgaga gcagagggga cag 23

<210> 152

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-T) _Flanking intron

<400> 152

ctaaccctgt ccttattgca tag 23

<210> 153

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> I4sh_Flanking intron

<400> 153

ctaacctctg cgcttcttcc cttccctcct ccctggctca g 41

<210> 154

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> I5_Flanking intron

<400> 154

actgaccctg tccttattgc acag 24

<210> 155

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> I5(22Y) _Flanking intron

<400> 155

actgacctct gcgcttcttc ccttccctcc tccctggctc ag 42

<210>

156

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> I5(22Y+1) _Flanking intron

<400> 156

actgacctct gcgcttcttc ccttccctcc tccctgtctc ag 42

<210> 157

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> I5(22Y-3) _Flanking intron

<400> 157

actgacctct gcgcttcttg ccttgcctgc tccctggctc ag 42

<210> 158

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> I5(15Y-3') _Flanking intron

<400> 158

actgacctct gcgcttcttc ccttccctgc tacctggctc ag 42

<210> 159

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> I5(15T-5') _Flanking intron

<400> 159

actgacctct gcgcttgttg ccttccctcc tccctggctc ag 42

<210> 160

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y; b-a) _Flanking intron

<400> 160

cttgccgagg gaggagggaa cag 23

<210> 161

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y; b-ct) _Flanking intron

<400> 161

agaaccgagg gaggagggaa cag 23

<210> 162

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y; b-y) _Flanking intron

<400> 162

ctaagggagg gaggagggaa cag 23

<210> 163

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223

> I4(0Y, b-2) _Flanking intron

<400> 163

cctacggagg gaggagggaa cag 23

<210> 164

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y, A) _Flanking intron

<400> 164

ctaaccgagg gaggagggaa cag 23

<210> 165

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y, T) _Flanking intron

<400> 165

ctaaccgagg gaggagggaa cag 23

<210> 166

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(0Y, G) _Flanking intron

<400> 166

ctaaccgagg gaggagggaa cag 23

<210> 167

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude; b-ct) _Flanking intron

<400> 167

ctaagggagt ccttagggga cag 23

<210> 168

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223

> I4(5Ynude; b-y) _Flanking intron

<400> 168

ctaagggagt ccttagggga cag 23

<210> 169

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Ynude, T) _Flanking intron

<400> 169

ctaaccgagt ccttagggga cag 23

<210> 170

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5, b-a) _Flanking intron

<400> 170

cttgccctgg gaggattgca cag 23

<210> 171

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5, b-ct) _Flanking intron

<400> 171

agaaccctgg gaggattgca cag 23

<210> 172

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5;b-y) _Flanking intron

<400> 172

ctaaggctgg gaggattgca cag 23

<210> 173

<211> 23

<212> DNA

<213> Artificial Sequence

<220><

223> I4(5Y-5, b-2) _Flanking intron

<400> 173

cctacgctgg gaggattgca cag 23

<210> 174

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5, A) _Flanking intron

<400> 174

ctaaccctgg gaggattgca cag 23

<210> 175

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> I4(5Y-5, T) _Flanking intron

<400> 175

ctaaccctgg gaggattgca cag 23