



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0143684
(43) 공개일자 2022년10월25일

(51) 국제특허분류(Int. Cl.)
A61K 39/215 (2006.01) A61K 39/00 (2006.01)
A61P 31/14 (2006.01)
(52) CPC특허분류
A61K 39/215 (2013.01)
A61P 31/14 (2018.01)
(21) 출원번호 10-2022-7030556
(22) 출원일자(국제) 2021년02월08일
심사청구일자 없음
(85) 번역문제출일자 2022년09월02일
(86) 국제출원번호 PCT/US2021/017149
(87) 국제공개번호 WO 2021/159118
국제공개일자 2021년08월12일
(30) 우선권주장
62/971,834 2020년02월07일 미국(US)
(뒷면에 계속)

(71) 출원인
알엔에이이뮤, 인코포레이티드
미국, 메릴랜드 20879, 게이더스버그, 스위트
280, 401 프로페셔널 드라이브
(72) 발명자
탕 성가오
미국, 메릴랜드 20879, 게이더스버그, 스위트
280, 401 프로페셔널 드라이브, 씨/오 알엔에이이
뮤, 인코포레이티드
센 동
미국, 메릴랜드 20879, 게이더스버그, 스위트
280, 401 프로페셔널 드라이브, 씨/오 알엔에이이
뮤, 인코포레이티드
(뒷면에 계속)
(74) 대리인
특허법인한얼

전체 청구항 수 : 총 34 항

(54) 발명의 명칭 **신종 코로나바이러스 감염에 대한 mRNA 백신의 조성물 및 방법**

(57) 요약

2019 신종 코로나바이러스(2019-nCoV) 감염의 예방 및 치료를 위한 강력한 mRNA 백신을 위한 조성물 및 방법이 제공된다. 조성물은 돌연변이된 에피토프를 포함하는 스파이크 단백질 에피토프를 암호화하는 하나 이상의 mRNA 분자, 또는 약제학적으로 허용되는 중합체성 나노입자 담체 및 리포솜 나노입자 담체와 함께 중요한 바이러스 유전자를 암호화하는 mRNA 카테일을 함유하는 약제학적 조성물을 포함한다. 피하, 복강내 및 근육내 주사를 포함하여 시스템 면역 반응 및 치료를 자극하는 방법이 제공된다.

(52) CPC특허분류

A61K 2039/53 (2013.01)
A61K 2039/55555 (2013.01)
C12N 2770/20034 (2013.01)

(30) 우선권주장

63/058,463 2020년07월29일 미국(US)
63/130,581 2020년12월24일 미국(US)

(72) 발명자

루 춘

미국, 메릴랜드 20879, 게이더스버그, 스위트 280,
401 프로페셔널 드라이브, 씨/오 알엔에이이퓨, 인
코포레이티드

헤 지양

미국, 메릴랜드 20879, 게이더스버그, 스위트 280,
401 프로페셔널 드라이브, 씨/오 알엔에이이퓨, 인
코포레이티드

헤 지아시

미국, 메릴랜드 20879, 게이더스버그, 스위트 280,
401 프로페셔널 드라이브, 씨/오 알엔에이이퓨, 인
코포레이티드

루 패트릭 와이.

미국, 메릴랜드 20879, 게이더스버그, 스위트 280,
401 프로페셔널 드라이브, 씨/오 알엔에이이퓨, 인
코포레이티드

명세서

청구범위

청구항 1

약제학적으로 허용되는 담체에 제형화된 2019-nCoV 단백질 또는 단백질 단편을 암호화(encoding)하는 개방 판독 프레임(open reading frame; ORF)을 포함하는 메신저 리보핵산(mRNA)의 유효량을 함유하는 조성물로서, 여기서 상기 mRNA 및 상기 담체가 중합체성 나노입자(polymeric nanoparticle)를 형성하는 조성물.

청구항 2

제1항에 있어서, 2019-nCoV 단백질이 스파이크(spike; S) 단백질 또는 스파이크(S) 서브유닛(subunit)인 조성물.

청구항 3

제2항에 있어서, ORF가 서열 번호 14 내지 31로부터 선택된 펩티드를 암호화하는 mRNA 분자로 이루어진 그룹으로부터 선택되는 조성물.

청구항 4

제2항에 있어서, 상기 ORF가 서열 번호 32 내지 49로 이루어진 그룹으로부터 선택되는 조성물.

청구항 5

제2항에 있어서, ORF가 서열 번호 50 내지 55로부터 선택된 펩티드를 암호화하는 mRNA를 포함하는 조성물.

청구항 6

제4항에 있어서, 상기 ORF가 서열 번호 91 내지 96으로 이루어진 그룹으로부터 선택되는 조성물.

청구항 7

제2항에 있어서, ORF가 서열 번호 56 내지 66으로부터 선택된 펩티드를 암호화하는 mRNA를 포함하는 조성물.

청구항 8

제6항에 있어서, 상기 ORF가 서열 번호 97 내지 107로 이루어진 그룹으로부터 선택되는 조성물.

청구항 9

제2항에 있어서, ORF가 서열 번호 11, 13, 68, 70, 72, 74 및 90으로부터 선택된 아미노산 서열을 암호화하는 mRNA 분자로 이루어진 그룹으로부터 선택되는 조성물.

청구항 10

제8항에 있어서, 상기 ORF가 서열 번호 10, 12, 67, 69, 71, 73 및 89로 이루어진 그룹으로부터 선택되는 조성물.

청구항 11

제2항에 있어서, 상기 ORF가 서열 번호 1, 2, 3, 및 4로 이루어진 그룹으로부터 선택되는 조성물.

청구항 12

제1항 내지 제11항 중 어느 한 항에 있어서, orf1a 및/또는 orf1b의 전부 또는 일부를 포함하는 mRNA를 추가로 포함하고, 여기서 상기 mRNA가 2019-nCoV 비구조적(non-structural) 단백질 폴리머라제 및/또는 RdRp를 암호화하는 조성물.

청구항 13

제12항에 있어서, 상기 ORF가 서열 번호 5 및 6으로 이루어진 그룹으로부터 선택된 mRNA인 조성물.

청구항 14

제1항에 있어서, 상기 ORF가 외피(envelope; E), 막(M) 및/또는 뉴클레오캡시드(N) 단백질을 암호화하는 2019 nCoV의 전부 또는 일부를 암호화하는 조성물.

청구항 15

제1항 내지 제14항 중 어느 한 항에 있어서, 외피(E), 막(M) 및/또는 뉴클레오캡시드(N) 단백질을 암호화하는 2019 nCoV의 전부 또는 일부를 암호화하는 하나 이상의 mRNA 서열을 추가로 포함하는 조성물.

청구항 16

제14항 또는 제15항에 있어서, 상기 ORF가 서열 번호 7, 8, 및 9로 이루어진 그룹으로부터 선택되는 조성물.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서, mRNA가 하나 이상의 N1-메틸-슈도우리딘 및/또는 슈도우리딘 잔기를 포함하는 화학적으로 변형된 mRNA 서열인 조성물.

청구항 18

제17항에 있어서, mRNA의 모든 우리딘 잔기가 N1-메틸 슈도우리딘 또는 슈도우리딘인 조성물.

청구항 19

제1항 내지 제18항 중 어느 한 항에 있어서, mRNA 분자가 3'-UTR, 5'-UTR을 추가로 포함하고, 임의로 (a) 조성물을 안정화하고 (b) 상기 ORF에 의해 암호화된 폴리펩티드의 발현을 향상시키는 추가 요소를 추가로 포함하는 조성물.

청구항 20

제19항에 있어서, 5'-UTR이 m7G 캡 구조(cap structure) 및 개시 코돈(start codon)을 포함하는 조성물.

청구항 21

제19항 또는 제20항에 있어서, 3'-UTR이 정지 코돈 및 폴리A 꼬리(polyA tail)를 포함하는 조성물.

청구항 22

선형 시험관내 전사(IVT) 시스템에서 제1항 내지 제21항 중 어느 한 항에 따른 개방 관독 프레임(ORF)을 발현함을 포함하여, 제1항 내지 제21항 중 어느 한 항에 따른 조성물을 제조하는 방법.

청구항 23

플라스미드 DNA(pDNA) 벡터 전달 시스템을 사용하여 제1항 내지 제21항 중 어느 한 항에 따른 개방 관독 프레임(ORF)을 발현함을 포함하여, 제1항 내지 제21항 중 어느 한 항에 따른 조성물을 제조하는 방법.

청구항 24

제1항 내지 제21항 중 어느 한 항에 있어서, 조성물이 하나 이상의 ORF를 포함하고 적어도 두 개의 폴리펩티드를 암호화하는 조성물.

청구항 25

제1항 내지 제21항 또는 제24항 중 어느 한 항에 있어서, 조성물이 히스티딘-리신 공중합체(HKP)를 포함하는 중합체성 나노입자 담체를 추가로 포함하는 조성물.

청구항 26

제25항에 있어서, HKP가 서열 번호 75-88로부터 선택된 측쇄를 갖는 중합체로 이루어진 그룹으로부터 선택되는

조성물.

청구항 27

제25항 또는 제26항에 따른 조성물을 제조하는 방법으로서, 상기 ORF를 포함하는 조성물을 HKP와 접촉시킴을 포함하고, 상기 mRNA 및 상기 HKP가 나노입자로 자가-조립되는 방법.

청구항 28

제1항 내지 제21항 또는 제24항 내지 제26항 중 어느 한 항에 있어서, 상기 나노입자가 지질을 추가로 포함하고, 여기서 상기 지질이 임의로 양이온성 지질인 조성물.

청구항 29

약제학적 유효량의 제1항 내지 제21항, 제24항 내지 제26항 또는 제28항 중 어느 한 항에 따른 조성물을 2019-nCoV 감염을 갖는 대상체(subject)에게 투여함을 포함하여, 2019-nCoV 감염을 갖는 대상체를 치료하는 방법.

청구항 30

제29항에 있어서, 조성물이 투여 전 또는 투여 동안 분무기 흡입 시스템에 의해 분무되는 방법.

청구항 31

제30항에 있어서, 분무기 시스템이 전체 호흡기 약물 전달을 위한 휴대용 분무기인 방법.

청구항 32

제29항에 있어서, 조성물이 피하 주사에 의해 투여되는 방법.

청구항 33

제29항에 있어서, 조성물이 근육내 주사에 의해 투여되는 방법.

청구항 34

제29항에 있어서, 조성물이 복강내 주사(i.p)에 의해 투여되는 방법.

발명의 설명

기술 분야

[0001] 본 출원은 2020년 2월 7일에 출원된 미국 가출원 제62/971,834호, 2020년 7월 29일에 출원된 제63/058,463호 및 2020년 12월 24일에 출원된 제63/130,581호에 대한 우선권을 주장하며, 이들 각각의 내용은 전문이 본원에 참고로 포함된다.

[0002] **발명의 분야**

[0003] 2019-nCoV 감염의 예방접종, 예방 및 치료를 위한 예방제 및 치료제가 이들의 사용 방법과 함께 제공된다.

배경 기술

[0004] **신종 코로나바이러스(2019-nCoV) 감염: 생물학 및 병리학**

[0005] 코로나바이러스(CoV)는 반복적으로 종의 장벽을 넘어섰고 일부는 중요한 인간 병원체로 등장하였다. 지난 20년 동안, 동물을 감염시키는 두 가지 코로나바이러스가 진화하였고 인간에서 발병을 유발하였다: SARS-CoV(2002, 베타코로나바이러스, 아속 사베코바이러스), 및 MERS-CoV(2012, 베타코로나바이러스, 아속 펠베코바이러스 [Drosten *et al.*, *New Engl J Med.* 2003; 348: 1967-1976; Zaki *et al.*, *New Engl J Med.* 2012; 367: 1814-1820]. 2019년 12월 31일, 중국 후베이성 우한시의 우한시 보건위원회는 7건의 중증 사례를 포함하여 병인을 알 수 없는 27건의 폐렴 사례 집단을 보고하였으며, 우한의 화난 수산물 도매시장(수산물 도매 및 다른 동물 종을 판매하는 살아있는 동물 시장)과 연관성이 있다고 일반적으로 보고하였다[Wuhan City Health Committee (WCHC). Wuhan Municipal Health and Health Commission's briefing on the current pneumonia epidemic

situation in our city 2019 (updated 31 December 2019/14 January 2020). Available from: <http://wjw.wuhan.gov.cn/front/web/showDetail/2019123108989>. 2019-nCoV는 이전에 인간에서 확인되지 않은 코로나바이러스의 새로운 균주이다[Zhu *et al.*, *N Engl J Med.* 382:727-733 (2020)]. 발병은 빠르게 확산되어 중국의 다른 지역과 해외에 영향을 미쳤다. 현재 아시아의 여러 국가 뿐만 아니라 호주, 유럽 및 북미에서도 사례가 감지되었다. 2019-nCoV는 호흡기 비밀을 통한 사람-사람 접촉을 통해 확산될 수 있는 것으로 보인다[Li *et al.*, *N Engl J Med.* 2020; 382:1199-1207].

[0006] 유전자 분석으로 2019-nCoV가 SARS-CoV와 밀접하게 관련되어 있으며 베타코로나바이러스 속에 유전적으로 군집하여 2개의 박쥐-유래 SARS-유사 균주와 함께 아속 사베코바이러스의 계통 B의 별개의 분기군을 형성한다는 것이 밝혀졌다. 바이러스 계통은 코로나바이러스에 공통적인 6개의 주요 개방 관독 프레임(open reading frame; ORF) 및 다수의 다른 보조 유전자로 구성된다. 추가 분석은 2019-nCoV 유전자 중 일부가 SARS-CoV와 80% 미만의 서열 동일성을 공유하였음을 나타낸다. 그러나, CoV 중 분류에 사용된 ORF1ab의 7개 보존 복제효소(RdRP) 도메인은 2019-nCoV와 SARS-CoV 간에 94.6% AA 서열이 동일하며, 이는 두 가지가 동일한 종에 속한다는 것을 의미한다. 막 엑소펩티다제인 안지오텐신 전환 효소 II(ACE2)는 SARS-CoV에 대한 세포 수용체로 알려져 있었다. 이 수용체는 또한 인간 세포에 진입하기 위해 2019-nCoV에 의해 사용되는 것으로 알려졌다. 도메인 266-330(성숙 단백질의 서열을 기반으로 번호매김)은 2019-nCoV와 ACE2의 상호작용에 필수적인 것으로 나타났다[Veljkovic *et al.*, *F1000Research.* 2020; 9: 52.].

[0007] **신종 코로나바이러스(2019-nCoV)에 대한 현재 예방 및 치료**

[0008] 현재, 2019-nCoV의 전체 계통 287,543개 서열이 GenBank(수탁: MN908947.3) [<http://www.ncbi.nlm.nih.gov/genbank/>.]에 공개된 것을 포함하여 중국 CDC 등에 의해 GISAID에 제출되었다[GISAID, Newly discovered betacoronavirus, Wuhan 2019-2020 (2019); <https://www.gisaid.org/>.]. 현재, 2019-nCoV 감염을 포함하여 코로나바이러스 감염을 예방하기 위한 긴급 사용 승인된 백신은 두 가지뿐이며 승인된 백신은 없다.

[0009] **mRNA 백신의 이점**

[0010] 예방접종은 질병 예방 및 통제에 대한 가장 성공적인 의학적 접근법이다. 백신의 성공적인 개발과 사용은 수천 명의 생명을 구하고 많은 돈을 절감하였다. RNA 백신의 주요 이점은 계란 또는 다른 포유류 세포의 사용이 필요할 수 있는 기존 백신 생산보다 덜 고가이고 빠르게 실험실에서 쉽게 이용 가능한 있는 재료를 사용하여 DNA 주형으로부터 RNA를 생산할 수 있다는 것이다. 또한, mRNA 백신은 백신 발견 및 개발을 간소화하고 신종 전염병에 대한 신속한 반응을 촉진할 가능성이 있다[Maruggi *et al.*, *Mol Ther.* 2019; 27(4): 757-772].

발명의 내용

해결하려는 과제

[0011] 전임상 및 임상 시험에서는 mRNA 백신이 동물 모델 및 인간에서 안전하고 오래 지속되는 면역 반응을 제공하는 것으로 나타났다. 전염병에 대한 mRNA 백신은 예방적 또는 치료적 치료법으로서 개발될 수 있다. 감염성 병원체의 항원을 발현하는 mRNA 백신은 강력한 T 세포 및 체액성 면역 반응을 유도하는 것으로 나타났다[Pardi *et al.*, *Nat Rev Drug Discov.* 2018; 17: 261-279.]. mRNA 백신을 생성하는 생산 절차는 전체 미생물, 생 약독화, 서브유닛 백신의 생산에 비해 무세포이고 간단하고 신속하다. 이러한 신속하고 간단한 제조 공정은 mRNA를 잠재적으로 신종 전염병과 효과적인 백신에 대한 절박한 필요성 사이의 격차를 메울 수 있는 유망한 바이오-제품으로 만든다.

과제의 해결 수단

[0012] 약제학적으로 허용되는 담체에 제형화된 2019-nCoV 단백질 또는 단백질 단편을 암호화(encoding)하는 개방 관독 프레임(ORF)을 함유하는 메신저 리보핵산(mRNA)의 유효량을 함유하는 조성물이 제공되며, 여기서 mRNA 및 담체는 중합체성 나노입자(polymeric nanoparticle)를 형성한다. 2019-nCoV 단백질은 스파이크(spike; S) 단백질 또는 스파이크(S) 서브유닛(subunit)일 수 있다. ORF는 서열 번호 14 내지 31, 또는 50 내지 66으로부터 선택된 서열을 갖는 펩티드와 같은 펩티드를 암호화하는 mRNA일 수 있다. ORF는, 예를 들면, 서열 번호 32 내지 49 또는 서열 번호 91 내지 107에 나타난 바와 같은 서열을 갖는 mRNA를 함유할 수 있다. ORF는 서열 번호 11, 13, 68, 70, 72, 74 및 90으로부터 선택된 아미노산 서열을 암호화하는 mRNA 분자로부터 선택될 수 있다. ORF

는 서열 번호 1-4, 10, 12, 67, 69, 71, 73 또는 89와 같은 서열을 가질 수 있다.

- [0013] 이들 조성물은 또한 orf1a 및/또는 orf1b의 전부 또는 일부를 함유하는 mRNA를 함유할 수 있으며, 여기서 mRNA는 2019-nCoV 비구조적(non-structural) 단백질 폴리머라제 및/또는 RdRp를 암호화한다. ORF는 서열 번호 5 또는 6과 같은 서열을 갖는 mRNA일 수 있다. ORF는 외피(envelope; E), 막(M) 및/또는 뉴클레오캡시드(N) 단백질을 암호화하는 2019 nCoV의 전부 또는 일부를 암호화할 수 있다. 이러한 조성물은 또한 외피(E), 막(M) 및/또는 뉴클레오캡시드(N) 단백질을 암호화하는 2019 nCoV의 전부 또는 일부를 암호화하는 하나 이상의 mRNA 서열을 함유할 수 있다. ORF는 서열 번호 7, 8 또는 9에 나타난 서열을 갖는 mRNA 분자로부터 선택될 수 있다.
- [0014] 이들 조성물 중 임의의 것에서, mRNA는 하나 이상의 N1-메틸-슈도우리딘 및/또는 슈도우리딘 잔기로 화학적으로 변형될 수 있다. mRNA의 우리딘 잔기는 모두 N1-메틸 슈도우리딘 또는 슈도우리딘일 수 있다.
- [0015] 이들 조성물 중 임의의 것에서 mRNA 분자는 3'-UTR, 5'-UTR을 추가로 함유할 수 있으며 임의로 (a) 조성물을 안정화하고 (b) ORF에 의해 암호화된 폴리펩티드의 발현을 향상시키는 추가 요소를 추가로 함유할 수 있다. 5'-UTR은 m7G 캡 구조(cap structure) 및 개시 코돈(start codon)을 함유할 수 있고/있거나 3'-UTR은 정지 코돈 및 폴리A 꼬리(polyA tail)를 함유할 수 있다.
- [0016] 조성물은 하나 이상의 ORF를 함유할 수 있고 적어도 두 개의 폴리펩티드를 암호화할 수 있다.
- [0017] 조성물은 히스티딘-리신 공중합체(HKP)를 함유하는 중합체성 나노입자 담체를 함유할 수 있다. 적합한 HKP의 예는 서열 번호 75-88로부터 선택된 측쇄를 갖는 중합체를 포함하지만, 이에 제한되지 않는다.
- [0018] 또한 선형 시험관내 전사(IVT) 시스템에서 또는 플라스미드 DNA(pDNA) 벡터 전달 시스템에 의해 ORF를 발현함으로써 상기 기술된 바와 같은 조성물을 제조하는 방법이 제공된다.
- [0019] 상기 기술된 바와 같은 조성물은 ORF(들)를 함유하는 조성물을 HKP와 접촉시킴으로써 제조될 수 있으며, 여기서 mRNA 및 HKP는 나노입자로 자가-조립된다. 조성물은 지질을 추가로 함유할 수 있고, 여기서 지질은 임의로 양이온성 지질일 수 있다.
- [0020] 또한, 약제학적 유효량의 상기 기술된 바와 같은 조성물을 대상체(subject)에게 투여함으로써 2019-nCoV 감염을 앓고 있는 대상체를 치료하는 방법이 제공된다. 조성물은 투여 전 또는 투여 동안 분무기 흡입 시스템에 의해 분무될 수 있다. 분무기 시스템은 전체 호흡기 약물 전달을 위한 휴대용 분무기일 수 있다. 조성물은 피하 주사, 근육내 주사, 또는 복강내 주사(i.p)에 의해 투여될 수 있다.

도면의 간단한 설명

- [0021] 도 1은 최적화된 mRNA 백신 발현 구조의 개략적 도식을 보여준다. 이 구조는 선형 시험관내 전사(IVT) 발현 시스템 또는 플라스미드 DNA 전달 벡터에 포함된다.
- 도 2는 mRNA 일시적 형질감염 후 스파이크 단백질 발현의 면역형광 분석 결과를 보여준다.
- 도 3은 mRNA 일시적 형질감염 후 스파이크 단백질 발현의 웨스턴 블롯 결과를 보여준다.
- 도 4는 스파이크 단백질 ORF mRNA 서열을 보여준다.
- 도 5는 D614G 돌연변이된 스파이크 단백질 ORF mRNA 및 단백질 서열을 보여준다.
- 도 6은 SARS RBD mRNA 서열, 및 D614G 돌연변이된 스파이크 단백질 ORF에 대한 아미노산 서열을 보여준다.
- 도 7은 2019-nCoV에 대한 폴리펩티드 에피토프에 대한 mRNA 및 아미노산 서열을 보여준다.
- 도 8은 D614G_S6P_DNA 및 아미노산 서열을 보여준다.
- 도 9는 RBD 삼량체 DNA 및 아미노산 서열을 보여준다.
- 도 10은 VUI-202012 DNA 및 아미노산 서열을 보여준다.
- 도 11은 HKP 펩티드 서열을 보여준다.
- 도 12는 D614G S2P DNA 서열을 보여준다.
- 도 13은 D614G S2P 아미노산 서열을 보여준다.

도 14는 2019-nCoV의 펩티드 에피토프를 암호화하는 mRNA 분자의 서열을 보여준다.

발명을 실시하기 위한 구체적인 내용

- [0022] 2019-nCoV 감염의 예방접종, 예방 및 치료를 위한 예방제 및 치료제가 제공된다. 약제학적 부형제와 함께 제형화된 mRNA 서열을 함유하는 약제학적 조성물이 제공된다. mRNA-기반 약제학적 제형의 조성물은 코로나바이러스 감염 및 복제에 필요한 바이러스 단백질을 코딩하는 mRNA 서열 및 특정 담체를 함유한다. 적합한 담체는 히스티딘-리신 공중합체(HKP), 및/또는 HKP + 지질, 예를 들어 양이온성 지질을 포함한다. 치료 방법, 공정 개발, 및 특정 전달 경로 및 섭생을 포함하는 이러한 약제학적 조성물을 사용하는 방법이 또한 제공된다.
- [0023] **2019-nCoV에 대한 mRNA 백신의 표적 선택 및 설계**
- [0024] 단백질 수준과 관련하여, 2019-nCoV의 예측 유전자를 식별하고 주석을 달았다. SARS-CoV-유사 코로나바이러스에서 검증된 7개의 기능성 단백질(pp1a, E2 당단백질 전구체, 가설 단백질 sars3a, 기질 단백질, 가설 단백질 sars6, 가설 단백질 sars7a, 뉴클레오캡시드 단백질)을 포함하여 모든 6개의 2019-nCoV 게놈의 주석 결과가 동일한 것으로 밝혀졌다[B. Buchfink, *et al.*, *Nat Methods*. 2015; 12: 59-60]. 이러한 결과는 2019-nCoV와 SARS-CoV 또는 MERS-CoV-유사 코로나바이러스 간의 유사점과 차이점을 예시한다.
- [0025] 2019-nCoV에 대한 우세한 인간 수용체는 인간 안지오텐신-전환 효소 2(ACE2)인 것으로 여겨진다. 2019-nCoV로부터의 S1 단백질의 도메인 288-330은 유망한 치료 및/또는 백신 표적을 나타낸다[Veljkovic *et al.*, *supra*]. 일반적으로, 2019-nCoV를 포함한 코로나바이러스의 스파이크(S) 단백질은 중화 항체의 주요 유도인자이며, S 단백질의 S1 서브유닛에 있는 수용체-결합 도메인(RBD)은 다중 형태적 중화 에피토프를 함유한다[He *et al.*, *J Immunol*. 2005; 174(8): 4908-4915]. 이는 RBD를 함유하는 재조합 단백질 및 RBD 서열을 암호화하는 mRNA를 포함하는 벡터가 안전하고 효과적인 2019-nCoV 백신을 개발하는데 사용될 수 있음을 시사한다.
- [0026] 본원에 기술된 바와 같이, 새로운 유형의 mRNA 백신(mRNA-1273)은 2019-nCoV의 하나 이상의 단백질로부터의 하나 이상의 에피토프를 암호화하는 개방 판독 프레임(ORF)을 함유하는 메신저 리보핵산(mRNA)을 함유하는 면역원성 조성물을 사용하여 개발되었다. mRNA는 중합체성 나노입자 및/또는 리포솜 나노입자를 함유하는 약제학적으로 허용되는 담체에서 제형화된다. 조성물은 대상체에서 2019-nCoV에 대한 특이 면역 반응을 유도하기에 효과적인 양으로 대상체에게 투여될 수 있다. 조성물은 하나 이상의 유형의 mRNA 분자를 함유할 수 있다. 각각의 추가 카피는 조성물의 면역 자극 활성을 향상시킨다. 모든 mRNA는 매우 면역원성이고 2019-nCoV 감염의 첫 번째 단계인 2019-nCoV가 상피 세포에 들어가는데 필요한 항원을 암호화한다.
- [0027] 하나의 실시양태에서, 조성물은 2019-nCoV의 전체 스파이크(S) 단백질을 암호화하는 하나 이상의 mRNA를 함유한다(도 3). S 단백질은 바이러스가 숙주 세포 수용체에 부착하는데 사용된다. 스파이크 단백질-코딩 mRNA 서열은 길며, 따라서 mRNA 발현 시스템 제조의 효율성을 개선하기 위해, 4개의 영역으로 나누었으며, 여기서 이들 영역 중 하나, 일부 또는 전부는 스파이크 단백질 도메인에 대한 프레임으로 발현될 때 스파이크 단백질 구조(서열 번호 1 내지 4)에 대한 mRNA 항원(들)으로서 사용될 수 있다. 잠재적 에피토프인 24개의 짧은 펩티드 단편이 표적으로 선택되었다(서열 번호 14 내지 31). 선택된 짧은 펩티드는 20개 미만의 아미노산, 유리하게는 15개 미만의 아미노산이다. 짧은 펩티드 서열을 암호화하는 mRNA 서열은 당업계에 잘 알려진 방법을 사용하여 설계하였다(서열 번호 32 내지 49).
- [0028] 또한, 다른 코로나바이러스와 비교하여, 2019-nCoV는 스파이크 단백질에 몇 가지 고유한 삽입을 갖는다. 2019-nCoV의 S 단백질의 삽입 부위의 아미노산 잔기는 인간 면역결핍 바이러스 1형(HIV-1)의 복제 단백질 gp120 또는 Gag의 아미노산 잔기와 동일하거나 유사하다. 또 다른 실시양태에서, 인간 면역계에 의해 인식되는, 2019-nCoV와 HIV-1 사이에 동일한 아미노산 잔기를 암호화하는 일련의 mRNA 서열이 설계되었다(서열 번호 50 내지 66).
- [0029] 또 다른 실시양태에서, mRNA는 스파이크 단백질의 ORF의 D614G 변이체를 암호화한다. SARS-CoV-2로 인한 코로나바이러스 질병 2019(COVID-19) 발병이 전 세계적인 대유행으로 급속히 확대되었다. 바이러스 스파이크(S) 단백질의 D614G 돌연변이를 암호화하는 SARS-CoV-2 단리물은 시간이 지남에 따라 이것이 발견된 위치에서 우세하며, 이는 이러한 변화가 바이러스 전파를 향상시킨다는 것을 의미한다. 잔기 614에 아스파르트산(D614) 및 글리신(G614)을 갖는 S 단백질은 상이한 기능적 특성을 갖는다. 특히, G614로 위형화된 레트로바이러스는 야생형(D614)보다 ACE2-발현 세포를 훨씬 더 효율적으로 감염시키는 것으로 밝혀졌다. 이러한 보다 큰 감염성은 더 적은 S1 웨딩(shedding) 및 슈도비리온으로의 S 단백질의 더 큰 통합과 상관관계가 있었다. SARS-CoV-2 M, N, E 및 S 단백질로 생성된 바이러스-유사 입자를 사용하여 유사한 결과가 획득되었다. 상관 분석에서, D614G 돌연변이를 지닌 스파이크 단백질은 사망률과 유의한 양의 상관관계를 보였다($r = 0.43$, $P = 0.022$). 이러한 결

과는 D614G 돌연변이를 가진 스파이크 단백질이 야생형보다 안정하다는 것을 나타내며, 이는 스파이크 D614G 변이를 가진 바이러스가 더 효율적으로 전염시킨다는 역학 데이터와 일치한다(서열 번호 10). D614G 돌연변이를 포함하는 ORF를 갖는 백신을 생산하였으며, 그 이유는 이것이 Sars-Cov-2의 우성 균주를 보다 효과적으로 보호할 것이기 때문이다.

[0030] 또 다른 실시양태에서, mRNA는 D614G 스파이크 단백질 전구체 S2P를 암호화하는 서열 번호 89를 갖는 ORF를 암호화한다.

[0031] 또 다른 실시양태에서, mRNA는 서열 번호 67을 갖는 ORF를 암호화하며, 이것은 확인된 스파이크 단백질의 다른 돌연변이를 포함한다. 이들은 HV 69-70 결실된 스파이크 단백질, Y144 결실된 스파이크 단백질, N501Y 돌연변이된 스파이크 단백질, A570D 돌연변이된 스파이크 단백질, P681H 돌연변이된 스파이크 단백질, T716I 돌연변이된 스파이크 단백질, F817P 변형된 스파이크 단백질, A892P 변형된 스파이크 단백질, A899P 변형된 스파이크 단백질, A942P 변형된 스파이크 단백질, S982A 돌연변이된 스파이크 단백질, 및 D1118H 돌연변이된 스파이크 단백질을 포함한다. 이들 공지된 돌연변이 중 하나 이상을 암호화하는 mRNA를 함유하는 백신 조성물은 본원에 기술된 방법을 사용하여 제조될 수 있다.

[0032] 또 다른 실시양태에서, mRNA는 외피(E), 막(M) 또는 뉴클레오캡시드(N) 구조 단백질을 포함하는 하나 이상의 다른 SARS-CoV-2 바이러스 단백질 코딩 서열을 암호화할 수 있다. 2019-nCoV 계통은 5'-복제효소(orf1/ab)-구조 단백질[스�파이크(S)-외피(E)-막(M)-뉴클레오캡시드(N)]-3'의 순서로 정렬되고(Zulma et al. 2016, Nature Reviews Drug Discovery 15:327-347.), Aβ-CoV 계통에서 특징적으로 발견되는 헤마글루티닌-에스테라제 유전자가 없다. 따라서, 2019-nCoV 코딩 mRNA의 S 단백질 이외에 또는 이에 추가하여 다양한 바이러스 단백질이 mRNA 기반 예방접종 표적으로 적용될 수 있다. 실시양태에서, 바이러스 표면 단백질은 유리하게는 mRNA 백신의 표적이다. ORF1a(서열 번호 5) 영역에서 바이러스 폴리머라제 및 ORF1b(서열 번호 6) 영역에서 복제(RdRp)를 담당하는 바이러스 기능성 단백질이 또한 백신 조성물을 제조하는데 사용될 수 있다.

[0033] 다른 바이러스 단백질도 mRNA 백신의 표적이 될 수 있다. 예를 들면, 코로나바이러스 막(M) 단백질은 비리온 조립의 핵심 역할을 한다. 이의 기능 중 하나는 스파이크가 바이러스 외피에 통합되는 것을 매개하는 것이다. 뉴클레오캡시드 N-단백질은 전형적으로 고도로 보존된 매우 면역원성인 인단백질이다. 코로나바이러스 N-단백질은 종종 진단 분석에서 마커로서 사용된다. 코로나바이러스(CoV) 외피(E) 단백질은 바이러스 형태형성 및 발병을 포함하여 감염 동안 여러 역할을 한다. 또 다른 실시양태에서, 특히 항원으로서 하나 이상의 바이러스 M 단백질, N 단백질 및 E 단백질의 전부 또는 일부를 코딩하는 mRNA는 항원성 백신 조성물(서열 번호 7,8,9)의 성분으로서 사용하기 위해 선택될 수 있다.

[0034] 백신 조성물을 제조하기 위해 mRNA 서열에 변형이 이루어질 수 있다. 예를 들면, 5' 캡 구조가 mRNA 서열에 추가될 수 있다. 하나의 실시양태에서, 5'CleanCap이 추가된다. 이 구조는 개시 캡이 있는 삼량체를 사용하여 자연 발생 5' 캡 구조를 생성한다. 유리하게는, mRNA 안정성을 증가시키기 위해 폴리-A 꼬리가 서열에 추가된다. AU 또는 GU-풍부 서열은 3' UTR에서 선택될 수 있는 반면, 무작위 서열은 숙주 유전자 서열에 대한 유사성을 감소시키기 위해 5' UTR에서 사용될 수 있다. mRNA 서열의 코돈 최적화는 또한 번역 효율을 개선하기 위해 사용될 수 있다. 구체적으로, 더 낮은 존재비의 tRNA 종에 상응하여 더 낮은 단백질 발현을 유도하는 코돈이 더 높은 수준의 단백질 발현을 제공하는 더 높은 존재비의 tRNA에 상응하는 코돈으로 대체된다(Brown, et al. Protein Expr Purif, 2008, 59:94-102; Maertens et al., 2010, Protein Sci, 9:1312-1326.; Plotkin et al., 2011, Nat Rev Genet, 12:32-42.; Kudla et al., 2009, Science, 324:255-258.)

[0035] **mRNA 백신 발현 효율 개선**

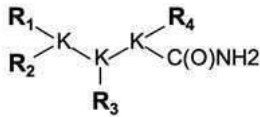
[0036] 포유류 세포에서 mRNA 발현 효율을 개선하기 위해, mRNA 안정성은 부분적인 화학적 변형에 의해 향상될 수 있다. 번역 효율을 더욱 증가시키기 위해, 비정상적인 RNA 폴리머라제 활성으로부터 유래되는 짧은 가닥 및 이중 가닥 RNA는, 예를 들면, FPLC, HPLC, 또는 셀룰로스를 매트릭스로 사용하는 크로마토그래피에 의해 제거된다. mRNA 백신의 효능을 개선하기 위해, 슈도우리딘(ϕ), 5-메틸시티딘(5mC), Cap-1 구조 및 최적화 코돈과 같은 변형된 뉴클레오시드의 사용과 함께 서열 최적화가 사용될 수 있으며, 이것이 결국 번역 효율을 개선시킨다. [변형된 뉴클레오시드의 사용에 대한 추가 정보를 갖는가? 아니면 그것이 상식인가? 이 시점에서 코돈 최적화는 상식이라고 생각된다] mRNA의 시험관내 전사 동안, 미성숙 mRNA는 선천 면역 활성화 자극을 통해 번역을 억제하는 오염물질로서 생성될 수 있다(Kariko et al, Mol Ther. 2008 16(11):1833-40.; Andries et al, J Control Release. 2015 Nov 10;217:337-44.). FPLC 및 HPLC 정제가 이러한 오염물질을 제거하기 위해 사용될 수 있다.

- [0037] 본원에 기술된 조성물에서, mRNA의 *시험관내* 전사를 위한 주형은 5'에서 3'로의 5개의 시스-작용성 구조 요소를 함유한다: (i) 최적화된 캡 구조, (ii) 최적화된 5' 비번역 영역(UTR), (iii) 코돈 최적화된 코딩 서열, (iv) 최적화된 3' UTR 및 (v) 반복된 아데닌 뉴클레오티드의 스트레치(polyA 꼬리)(도 18). (Murray et al, *Mol. Cell. Biol.* 2007, 27, 2791-2799. ; Louis et al, *Curr. Opin. Genet. Dev.* 2011, 21, 444-451. ; Ferizi et al, *Lab Chip* 2015, 15, 3561-3571. ; Von Niessen et al, *Mol. Ther.* 2019, 27, 824-836.) 이러한 시스-작용성 구조 요소는 반감기 및 발현 수준을 포함하여 보다 양호한 mRNA 특성을 위한 노력에서 더욱 최적화된다. 본원에서 제공되는 5'-UTR은 시작을 포함하지만 폴리펩티드를 암호화하지 않는다(즉, 비-코딩이다). 일부 실시양태에서, 본 개시내용의 5'-UTR은 7-메틸구아노신(7mG) 서열을 갖는 캡 구조를 포함한다. 3'-UTR은 정지 코돈(종료 신호를 나타내는 mRNA 전사체의 코돈)으로부터 바로 하류(3')에 있고 폴리펩티드를 암호화하지 않는다(비-코딩이다). 폴리A 꼬리는 3'-UTR의 하류에 있고 다중 연속 아데노신 모노포스페이트를 함유하는 mRNA의 특수 영역이다; 이것은 효율적인 mRNA 생산을 위해 필요하다. mRNA *시험관내* 전사 주형은 mRNA 생산의 효율을 위해서 뿐만 아니라 후속 단백질 또는 펩티드 생산에도 사용된다.
- [0038] 일부 실시양태에서, mRNA는 RNA 폴리머라제(T7, T3 또는 SP6) 및 상이한 뉴클레오티드의 혼합물을 사용함으로써 박테리오파지 프로모터, 최적화된 UTR 및 코돈 최적화된 서열을 함유하는 선형 DNA 주형으로부터 *시험관내* 전사(IVT)에 의해 생성된다. 또 다른 실시양태에서, 선형 DNA 주형은 전달 벡터로서 플라스미드 DNA(pDNA)에 클로닝될 수 있다. 플라스미드 벡터는 mRNA 백신 생산에 적합할 수 있다. 일반적으로 사용되는 플라스미드는 pSFV1, pcDNA3 및 pTK126을 포함하며, 이들은 모두 상업적으로 이용 가능하다. 하나의 독특한 mRNA 발현 시스템은 pEVL이다(참조; Grier et al. *Mol Ther Nucleic Acids.* 19;5:e306 (2016)).
- [0039] 일부 실시양태에서, mRNA는 *시험관내* 전사 동안 모든 우리딘 잔기가 슈도우리딘으로 대체되는 변형된 mRNA이다. 이러한 변형은 세포에서 효소 분해에 대해 mRNA를 안정화시켜, mRNA의 번역 효율을 향상시킨다. 사용된 슈도우리딘은 N1-메틸-슈도우리딘, 또는 N6-메틸아데노신(m6A), 이노신, 슈도우리딘, 5-메틸시티딘(m5C), 5-하이드록시메틸시티딘(hm5C) 및 N1-메틸아데노신(m1A)과 같이 당업계에 잘 알려진 다른 변형일 수 있다. 변형은 임의로 ORF, 5'UTR 및 3'UTR을 포함하는 전체 mRNA에 걸쳐 이루어진다. 숙련가는 다른 변형된 RNA 잔기가 단백질 3차원 구조를 안정화하고 단백질 번역을 증가시키기 위해 사용될 수 있음을 인식할 것이다.
- [0040] 효율의 수준을 시험하기 위해, 실시예 3, 4 및 5에 기술된 실험을 수행하였다.
- [0041] **mRNA 백신 전달 시스템으로서의 히스티딘-리신(HK) 폴리펩티드**
- [0042] 핵산을 표적 세포로 전달하는데 효과적인 수단은 기초 연구 환경 및 임상 적용 모두에서 중요한 도구이다. 특정 담체의 효과는 형질전환되는 핵산의 특징에 따라 좌우되기 때문에 현재 핵산 담체의 다양한 어레이가 필요하다[Blakney et al. *Biomacromolecules* 2018, **19**: 2870-2879. Goncalves et al. *Mol Pharm* 2016; **13**: 3153-3163. Kauffman et al. *Biomacromolecules* 2018; **19**: 3861-3873. Peng et al. *Biomacromolecules* 2019; **20**: 3613-3626. Scholz et al. *J Control Release* 2012; **161**: 554-565.]. 다양한 담체 중, 비-바이러스 전달 시스템이 개발되었으며 여러 측면에서 바이러스 전달 시스템보다 더 유리한 것으로 보고되었다[Brito et al. *Adv Genet.* 2015; 89: 179-233]. 예를 들면, 고분자량 분지형 폴리에틸렌민(PEI, 25kDa)은 플라스미드 DNA에 대한 우수한 담체이지만 mRNA에 대해서는 그렇지 않다. 그러나, PEI의 분자량을 2kDa로 감소시킴으로써, mRNA의 보다 효과적인 담체가 된다[Bettinger et al. *Nucleic Acids Res* 2001; **29**: 3882-3891.].
- [0043] 4-분지된 히스티딘-리신(HK) 펩티드 중합체 H2K4b는 고분자량 DNA 플라스미드의 우수한 담체이지만[Leng et al. *Nucleic Acids Res* 2005; **33**: e40.] 비교적 저분자량 siRNA의 불량한 담체인 것으로 나타났다[Leng et al. *J Gene Med* 2005; **7**: 977-986.]. H2K4b의 2개의 히스티딘-풍부 펩티드 유사체, 즉 H3K4b 및 H3K(+H)4b가 siRNA의 효과적인 담체인 것으로 나타났지만[Leng et al. *J Gene Med* 2005; **7**: 977-986. Chou et al. *Biomaterials* 2014; **35**: 846-855.], H3K(+H)4b가 약간 더 효과적인 것으로 보였다[Leng et al. *Mol Ther* 2012; **20**: 2282-2290.]. 더욱이, siRNA의 H3K4b 담체는 H3K(+H)4b siRNA 다중체보다 *시험관내* 및 *생체내에서* 유의하게 더 큰 정도로 사이토카인을 유도하였다[Leng et al. *Mol Ther* 2012; **20**: 2282-2290.]. 적합한 HK 폴리펩티드는 제 WO/2001/047496호, 제 WO/2003/090719호, 및 제 WO/2006/060182호에 기술되어 있으며, 이들 각각의 내용은 전문이 본원에 포함된다. 이러한 폴리펩티드는 리신 측쇄 ε-아미노 그룹과 N-말단이 다양한 HK 서열에 결합된 리신 골격(3개의 라이신 잔기)을 갖는다. HK 폴리펩티드 담체는, 예를 들면, 고체상 합성을 포함하는 당업계에 잘 알려진 방법에 의해 합성될 수 있다.
- [0044] 이러한 히스티딘-리신 펩티드 중합체("HK 중합체")는 놀랍게도 mRNA 담체로서 효과적이며, 이들은 단독으로 또는 리포솜과 조합하여 사용되어 표적 세포로의 mRNA의 효과적인 전달을 제공할 수 있는 것으로 밝혀졌다. PEI

및 기타 담체와 유사하게, 초기 결과는 HK 중합체가 핵산을 운반하고 방출하는 능력에 있어 상이하다는 것을 시사하였다. 그러나, HK 중합체는 펩티드 합성기에서 재현 가능하게 만들 수 있기 때문에, 이들의 아미노산 서열을 쉽게 변경할 수 있어 mRNA의 결합과 방출 뿐만 아니라 HK 중합체와 mRNA를 함유하는 다중체의 안정성을 미세하게 제어할 수 있다[Chou et al. *Biomaterials* 2014; **35**: 846-855. Midoux et al. *Bioconjug Chem* 1999; **10**: 406-411. Henig et al. *Journal of American Chemical Society* 1999; **121**: 5123-5126.]. mRNA 분자가 하나 이상의 HKP 담체와 혼합될 때, 성분들은 나노입자로 자가-조립된다.

[0045] 본원에 기술된 바와 같이, 유리하게는 HK 중합체는 3개-리신 아미노산 코어에 연결된 4개의 짧은 펩티드 분지를 포함한다. 펩티드 분지는 상이한 구성으로 히스티딘 및 리신 아미노산으로 이루어진다. 이들 히스티딘-리신 펩티드 중합체(HK 중합체)의 일반적인 구조는 화학식 I에 나타나 있으며, 여기서 R은 펩티드 분지를 나타내고 K는 아미노산 L-리신을 나타낸다.

[0046] [화학식 I]



[0047]

[0048] 화학식 I에서 K는 L-리신이고 각각의 R₁, R₂, R₃ 및 R₄는 독립적으로 히스티딘-리신 펩티드이다. R₁₋₄ 분지는 본 발명의 HK 중합체에서 동일하거나 상이할 수 있다. R 분지가 "상이"할 때, 그 분지의 아미노산 서열은 중합체의 다른 R 분자 각각과는 상이하다. 화학식 I에 나타낸 본 발명의 HK 중합체에 사용되는 적합한 R 분지는 하기 R 분지 R_A - R_J를 포함하지만, 이에 제한되지 않는다:

[0049] R_A = KHKHKHKHKHKHKHKHKHK- (서열 번호 79)

[0050] R_B = KHKHKHKHKHKHKHKHK- (서열 번호 80)

[0051] R_C = KHKHKHKHKHKHKHKHK- (서열 번호 81)

[0052] R_D = kHKHKHKHKHKHKHKHK- (서열 번호 82)

[0053] R_E = HKHKHKHKHKHKHKHK- (서열 번호 83)

[0054] R_F = HKHKHKHKHKHKHKHK- (서열 번호 84)

[0055] R_G = KHKHKHKHKHKHKHKHK- (서열 번호 85)

[0056] R_H = KHKHKHKHKHKHKHKHK- (서열 번호 86)

[0057] R_I = KHKHKHKHKHKHKHKHK- (서열 번호 87)

[0058] R_J = KHKHKHKHKHKHKHKHK- (서열 번호 88)

[0059] mRNA 조성물에 사용될 수 있는 특정 HK 중합체는 R₁, R₂, R₃ 및 R₄가 각각 동일하고 R_A - R_J로부터 선택되는 HK 중합체를 포함하지만, 이에 제한되지 않는다(표 1). 이러한 HK 중합체를 각각 H2K4b, H3K4b, H3K(+H)4b, H3k(+H)4b, H-H3K(+H)4b, HH-H3K(+H)4b, H4K4b, H3K(1+H)4b, H3K(3+H)4b 및 H3K(1,3+H)4b라고 한다. 이러한 10개의 예 각각에서, 대문자 "K"는 L-리신을 나타내고, 소문자 "k"는 D-리신을 나타낸다. H3K4b와 비교하여 추가 히스티딘 잔기는 분지 서열 내에서 밀줄이 그어져 있다. HK 중합체의 명명법은 다음과 같다:

- [0060] 1) H3K4b의 경우, 분지의 우세한 반복 서열은 -HHHK-이며, 따라서 "H3K"가 이름의 일부이며; "4b"는 분지의 수를 지칭하고;
- [0061] 2) H3K4b 및 유사체의 각 분지에는 4개의 -HHHK- 모티프가 있고; 첫 번째 -HHHK- 모티프("1")는 리신 코어에 가장 가까우며;
- [0062] 3) H3K(+H)4b는 H3K4b의 두 번째 -HHHK- 모티프(모티프 2)에 하나의 추가 히스티딘이 삽입된 H3K4b의 유사체이

고;

[0063] 4) H3K(1+H)4b 및 H3K(3+H)4b 펩티드의 경우, 첫 번째(모티프 1) 및 세 번째(모티프 3) 모티프에 각각 추가 히스티딘이 있고;

[0064] 5) H3K(1,3+H)4b의 경우, 분지의 첫 번째 및 세 번째 모티프 모두에 2개의 추가 히스티딘이 있다.

[0065] 표 1

중합체	분지 서열	서열 식별자
H2K4b	$R_A = \text{KHKHHKHHKHHKHHKHHKHHK-}$	(서열 번호 79)
	$\begin{array}{cccc} & \underline{4} & \underline{3} & \underline{2} & \underline{1} \\ \text{---} & & & & \end{array}$	
H3K4b	$R_B = \text{KHHHKHHHKHHHKHHHK-}$	(서열 번호 80)
H3K(+H)4b	$R_C = \text{KHHHKHHHKHHHKKHHHK-}$	(서열 번호 81)
H3k(+H)4b	$R_D = \text{kHHHkHHHkHHHhkHHHk-}$	(서열 번호 82)
H-H3K(+H)4b	$R_E = \text{HKHHHKHHHKHHHKKHHHK-}$	(서열 번호 83)
HH-H3K(+H)4b	$R_F = \text{HHKHHHKHHHKHHHKKHHHK-}$	(서열 번호 84)
H4K4b	$R_G = \text{KHHHKKHHHKKHHHKKHHHKK-}$	(서열 번호 85)
H3K(1+H)4b	$R_H = \text{KHHHKHHHKHHHKKHHHKK-}$	(서열 번호 86)
H3K(3+H)4b	$R_I = \text{KHHHKHHHKKHHHKKHHHKK-}$	(서열 번호 87)
H3K(1,3+H)4b	$R_J = \text{KHHHKHHHKKHHHKKHHHKK-}$	(서열 번호 88)

[0066]

[0067] 겔 지연 분석, 헤파린 치환 분석 및 유세포 분석을 포함하는 당업계에 잘 알려진 방법을 수행하여 mRNA를 성공적으로 전달하는데 있어서 HK 중합체 + 리포솜을 함유하는 다른 제형의 성능을 평가할 수 있다. 적합한 방법이, 예를 들면, 문헌[Gujrati *et al.*, *Mol. Pharmaceutics* 11:2734-2744 (2014), Parnaste *et al.*, *Mol Ther Nucleic Acids*, 7: 1-10 (2017)]에 기술되어 있다.

[0068] 세포로의 mRNA 흡수의 검출은 또한 SmartFlare® 기술(Millipore Sigma)을 사용하여 달성할 수 있다. 이러한 스마트 플레어는, 세포에서 RNA 서열을 인식할 때, 형광 현미경으로 분석할 수 있는 형광의 증가를 야기하는 서열이 부착된 비드이다.

[0069] 또 다른 방법은 mRNA로부터 단백질 발현을 측정하는 것을 포함한다 - 예를 들면, 루시페라제를 암호화하는 mRNA는 당업계에 잘 알려진 방법을 사용하여 형질감염의 효율을 측정하는데 사용될 수 있다. 예를 들면, 이들은 HKP 및 리포솜 제형을 사용하여 mRNA를 전달하는 효율을 입증하기 위해 최신 간행물에서 루시페라제 mRNA로 수행되었다(He *et al.*, *J Gene Med.* 2021 Feb;23(2):e3295).

[0070] mRNA 전달에서 DOTAP 및 HK 담체의 상승 활성

[0071] H3K(+H)4b와 DOTAP(양이온성 지질)의 조합은 놀랍게도 MDA-MB-231 세포로 mRNA를 운반하는 능력에서 상승적이었다(H3K(+H)4b/리포솜 vs 리포솜, $P < 0.0001$). 조합은 중합체 단독 및 양이온성 지질 담체보다 mRNA의 담체로서 각각 약 3배 및 8배 더 효과적이었다. 모든 HK 펩티드가 DOTAP 지질과의 상승 작용을 입증하는 것은 아니다. 예를 들면, H3K4b와 DOTAP의 조합은 루시페라제 mRNA의 담체로서 DOTAP 리포솜보다 덜 효과적이었다. DOTAP 외에도, HK 펩티드와 함께 사용될 수 있는 다른 양이온성 지질은 리포펙틴(ThermoFisher), 리포펙타민(ThermoFisher) 및 DOSPER을 포함한다(도 22).

[0072] 분지의 L-리신이 D-리신으로 대체된 H3k(+H)4b의 D-이성체가 가장 효과적인 중합체 담체였다(H3k(+H)4b vs. H3K(+H)4b, $P < 0.05$). mRNA의 D-이성체/리포솜 담체는 H3k(+H)4b 단독 및 리포솜 담체보다 각각 거의 4배 및

10배 더 효과적이었다. D-H3k(+H)4b/지질 조합이 L-H3K(+H)4b/지질 조합보다 약간 더 효과적이었지만, 이 비교는 통계적으로 다르지 않았다(도 23).

[0073] **HK 중합체 중 mRNA에 대한 담체로서의 H3K(+H)4b**

[0074] H3K4b 및 H3K(+H)4b 둘 다는 시험관내에서 핵산의 담체로서 사용될 수 있다[Leng *et al.* *J Gene Med* 2005; 7: 977-986. Chou *et al.*, *Cancer Gene Ther* 2011; 18: 707-716.]. 이러한 이전의 발견에도 불구하고, H3K(+H)4b는 다른 유사한 유사체에 비해 mRNA의 담체로서 현저히 더 우수하였다(표 2).

[0075] **표 2**

중합체	비율(wt:wt; mRNA:중합체)	RLU/ug-단백질
H3K(+H)4b	1:4	1532.9±122.9
	1:8	1656.3±202.5
	1:12	1033.4±197
H3k(+H)4b	1:4	1851.6±138.3
	1:8	1787.2±195.2
	1:12	1982.3±210.7
H3K4b	1:4	156.8±41.8
	1:8	62.1±13.2
	1:12	18.1±4.0
H3K(3+H)4b	1:4	61.7±5.7
	1:8	68.7±3.1
	1:12	59.0±7.5
H3K(1+H)4b	1:4	24.3±4.5
	1:8	15.0±3.6
	1:12	7.3±2.5
H-H3K(+H)4b	1:4	1107.5±140.4
	1:8	874.6±65.2
	1:12	676.4±25.7

[0076]

HH-H3K(+H)4b	1:4	1101.9±106.6
	1:8	832.2±75.3
	1:12	739.8±105.4
H4K4b	1:4	896.4±112.6
	1:8	821.8±115.6
	1:12	522.4±69.2
H3K(1,3+H)4b	1:4	518.3±134.7
	1:8	427.7±18.1
	1:12	378±5.2
H2K4b	1:4	546.7±70.1
	1:8	132.3±58.5
	1:12	194.7±18.4

[0077]

- [0078] 특히, 이것은 다양한 중량:중량 (HK:mRNA) 비율에서 H3K4b보다 더 높은 mRNA 형질감염 효율을 갖는다. 4:1 비율에서, 루시퍼라제 발현은 상당한 세포독성 없이 MDA-MB-231 세포에서 H3K4b보다 H3K(+H)4b로 10배 더 높았다. 또한, H3K4b 및 H3K(+H)4b에서 히스티딘의 백분율(중량 기준)이 각각 68.9% 및 70.6%이기 때문에 완충 능력이 이들의 형질감염 차이의 필수 요소인 것으로 보이지 않는다.
- [0079] 겔 지연 분석은 mRNA의 전기영동 이동성이 HK 중합체에 의해 지연되었음을 보여준다. 지연 효과는 펩티드 대 mRNA 중량 비율이 높을수록 증가하였다. 그러나, mRNA는 H3K(+H)4b의 2:1 비율에서 완전히 지연되었지만 H3K4b에 의해서는 완전히 지연되지 않았다. 이것은 H3K(+H)4b가 보다 안정적인 다중체를 형성할 수 있음을 시사하였으며, 이는 mRNA 전달에 적합한 담체가 되는 능력에 유리하였다.
- [0080] H3K(+H)4b 펩티드가 mRNA에 더 단단히 결합한다는 추가 확인은 헤파린-전위 분석(heparin-displacement assay)으로 입증되었다. 다양한 농도의 헤파린이 mRNA 및 HK로 형성된 다중체에 첨가되었으며, 특히 낮은 농도의 헤파린에서, mRNA가 H3K(+H)4b 중합체보다 더 쉽게 H3K4b 중합체에 의해 방출되는 것으로 관찰되었다. 이러한 데이터는 H3K(+H)4b가 mRNA에 결합하여 H3K4b보다 더 안정적인 다중체를 형성할 수 있음을 시사한다.
- [0081] 시아닌-5로 표지된 mRNA를 사용하여, H3K4b 및 H3K(+H)4b 다중체의 MDA-MB-231 세포로의 흡수를 유세포 분석을 사용하여 비교하였다. 상이한 시점(1, 2 및 4시간)에서, H3K(+H)4b 다중체는 H3K4b 다중체보다 더 효율적으로 세포로 들어왔다(데이터는 표시되지 않음). 이러한 결과와 유사하게, 형광 현미경 검사법은 H3K(+H)4b 다중체가 H3K4b 다중체보다 훨씬 더 많은 산성 엔도솜 소포 내에 편재됨을 나타내었다(H3K4b vs. H3K(+H)4b, $P < 0.001$). 흥미롭게도, 식작용 소포와 겹치지 않는 불규칙한 형상의 H3K4b 다중체는 아마도 세포외였으며, H3K(+H)4b 다중체에서는 관찰되지 않았다.
- [0082] HK 중합체와 양이온성 지질(즉, DOTAP)은 플라스미드로의 형질감염을 상당히 독립적으로 증가시키는 것으로 알려져 있다[Chen et al. *Gene Ther* 2000; **7**: 1698-1705.]. 결과적으로, 이들 지질이 HK 중합체와 함께 mRNA 형질감염을 향상시켰는지 여부를 조사하였다. 특히, H3K(+H)4b 및 H3k(+H)4b 담체는 DOTAP 리포솜보다 mRNA의 훨씬 더 나은 담체였다. H3K(+H)4b와 DOTAP 지질의 조합은 MDA-MB-231 세포로 mRNA를 운반하는 능력에서 상승적이었다. 이 조합은 중합체 단독 및 리포솜 담체보다 mRNA의 담체로서 각각 약 3배 및 8배 더 효과적이었다(H3K(+H)4b/lipid vs. 리포솜 또는 H3K(+H)4b). 특히, 모든 HK 펩티드가 DOTAP 지질과 함께 개선된 활성을 입증한 것은 아니다. H3K4b와 DOTAP 담체의 조합은 루시퍼라제 mRNA의 담체로서 DOTAP 리포솜보다 덜 효과적이었다. DOTAP와 H3K(+H)4b 담체의 조합은 mRNA를 세포로 운반하는 능력에서 상승적인 것으로 밝혀졌다[He et al. *J Gene Med.* 2020 Nov 10:e3295].
- [0083] **mRNA 백신의 투여**
- [0084] 본원에 기술된 백신 제형은 백신을 투여하기 위해 통상적으로 사용되는 임의의 투여 방식에 의해 인간 대상체를 포함한 대상체에게 투여될 수 있다. 따라서 조성물은 에어로졸, 분산액, 용액 또는 현탁액의 형태일 수 있고 흡입, 근육내, 경구, 설하, 협측, 비경구, 비강, 피하, 피내 또는 국소 투여용으로 제형화될 수 있다. 본원에 사용된 용어 비경구는 경피, 피하, 혈관내(예를 들어, 정맥내), 근육내, 또는 척추강내 주사 또는 주입 기술 등을 포함한다.
- [0085] 본원에 사용된 바와 같이, 백신의 유효량은 백신이 투여되는 대상체에서 보호 면역 반응을 생성하는데 필요한 용량이다. 본 맥락에서 보호 면역 반응은 2019-nCoV로 시험감염된 대상체에서 질병을 예방하거나 개선하는 것이다. 백신이 보호 면역 반응을 생성하는지 여부를 결정하기 위해 다수의 대상체에서 임상 시험을 실행하는 방법은 당업계에 잘 알려져 있다.
- [0086] 백신은 1회 이상 투여될 수 있다. 백신에 대한 면역 반응의 초기 측정은 백신을 제공받는 대상체에서 항체의 생산을 측정함으로써 이루어질 수 있다. 이러한 방식으로 항체 생산을 측정하는 방법은 또한 당업계에 잘 알려져 있다. 이것은 질병 상태의 예방, 발생 억제 또는 치료(증상을 어느 정도 완화, 바람직하게는 모든 증상 완화)에 필요한 용량이다. 약제학적 유효량은 질병의 유형, 사용된 조성물, 투여 경로, 치료되는 포유동물의 유형, 고려 중인 특정 포유동물의 신체적 특성, 동시 투여 약물, 및 기타 의학 분야의 숙련가가 인지하는 요인에 따라 좌우된다. 일반적으로, 제형화된 조성물의 효능에 따라 0.1mg/kg 내지 100mg/kg 체중/일의 활성 성분의 양이 투여된다.
- [0087] 그러나, 이들의 적용은 최근까지 mRNA의 불안정성 및 비효율적인 생체내 전달에 의해 제한되었다. 본원에 기술된 방법은 폴리펩티드 전달 시스템 및 폴리펩티드 나노입자 전달 시스템을 갖는 mRNA 백신을 제조하고 사용하는 방법을 제공한다.

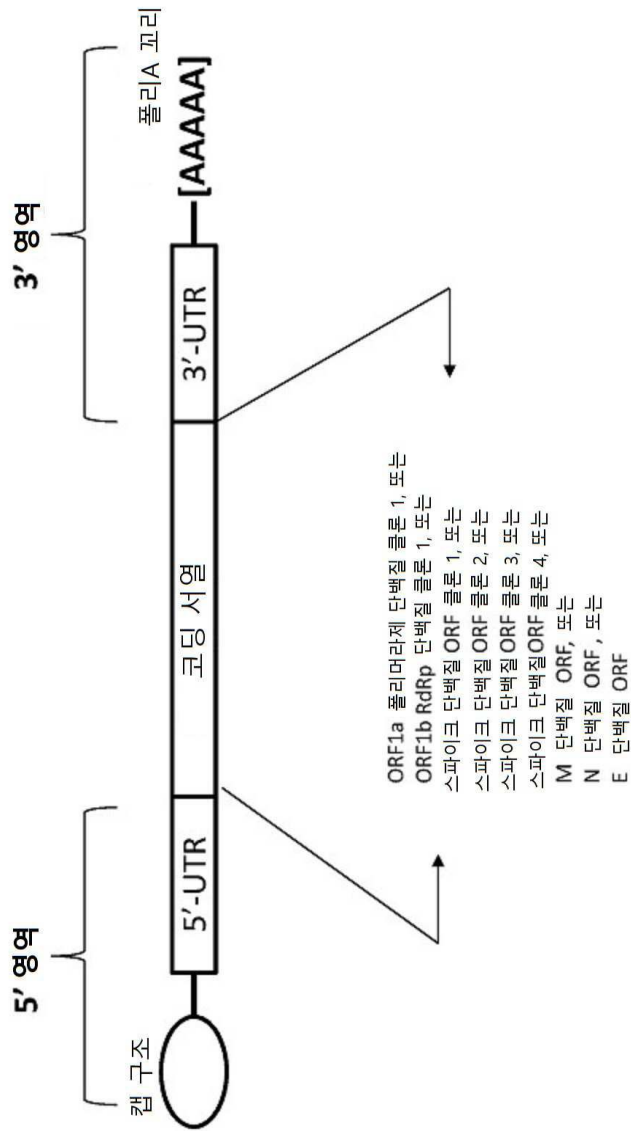
- [0088] 본원에 기술된 방법은 다양한 질병, 특히 감염성 질병에 대한 예방 및 치료 백신을 포함하는 mRNA의 임상 적용에 사용될 수 있다.
- [0089] **실시예 1: mRNA 합성:**
- [0090] CleanCap SARS-COV 2 스파이크 단백질(S1P 및 S2P) 및 CleanCap 강화된 녹색 형광 단백질(EGFP) 발현 mRNA를 합성하였으며 CleanCap SARS-COV 2 스파이크 단백질(S6P)은 TriLink BioTechnologies, Inc.(San Diego, CA)에 의해 제조하였고 에탄올-함유 완충액에서 셀룰로스에 dsRNA를 선택적으로 결합시켜 정제하였다. 10mM HEPES(pH 7.2), 0.1mM EDTA, 125mM NaCl, 16% 에탄올 및 셀룰로스 섬유를 함유하는 크로마토그래피 완충액을 mRNA 생성물과 함께 마이크로원심분리 스피ن 컬럼에 첨가하고 원심분리하였다. 이 절차 후에 dsRNA의 거의 90%를 제거할 수 있다(Baiersdorfer et.al, 2019). FPLC 및 HPLC를 사용하여 오염 물질을 또한 제거할 수 있다(Kariko et.al, 2011).
- [0091] **실시예 2: 펩티드 (HK 중합체) 제조**
- [0092] HK 펩티드 중합체는 메릴랜드 대학의 바이오폴리머 코어 시설에 의해 Rainin Voyager 합성기(Tucson, AZ)에서 합성하였다.
- [0093] **실시예 3: mRNA의 시험관내 형질감염**
- [0094] mRNA, EGFP mRNA 및 SARS-COV 2 SP의 적절한 단백질 발현을 확인하기 위해, mRNA를 인간 배아 신장 293T 세포(293T 세포)에 일시적으로 형질감염시켰다. 간단히 말해서, 4.8×10^5 개 세포를 2ml의 DMEM(10% 소 태아 혈청 및 1% 페니실린-스트렙토마이신(ThermoFisher Scientific))을 함유하는 6웰 플레이트에 플레이팅하였다. 24시간 후, 세포가 70-90% 융합되었을 때, mRNA를 제공된 프로토콜에 따라 리포펙타민 MessengerMAX 형질감염 시약(ThermoFisher Scientific)을 사용하여 293T 세포로 형질감염시켰다. **시험관내** 단백질 발현 측정을 위해 293T 세포를 2일 동안 배양하였다.
- [0095] 양이온성 리포솜 DOTAP(1,2-디올레오일-3-트리메틸암모늄-프로판, Roche), MC3(DLin-MC3-DMA, ChemScene) 및 PLA(폴리락트산)와 관련된 두 개의 HK 기반 중합체를 EGFP mRNA 및 SARS-COV 2 SP mRNA를 인간 배아 신장 293 세포(HEK 293 세포)로 운반하는 능력에 대해 실험하였다. 간단히 말해서, 4.8×10^5 개 세포를 2ml의 DMEM(10% 소 태아 혈청 및 페니실린-스트렙토마이신(ThermoFisher Scientific))을 함유하는 6웰 플레이트에 플레이팅하였다. 24시간 후, 세포가 70-90% 융합되었을 때, EGFP 및 SP mRNA를 갖는 두 개의 HK 기반 중합체 제형을 각 웰에 첨가하였다. 293T 세포를 시험관내 단백질 발현 측정을 위해 2일 동안 배양하였다.
- [0096] **실시예 4: 시험관내 단백질 발현 측정**
- [0097] **면역형광 분석:** 2일간의 형질감염 후, Cytation5 세포 이미징 멀티-모드 판독기(Biotek, Winooski, VT)를 사용하여 면역형광 이미징에 의해 단백질 발현을 측정하였다.
- [0098] **세포 용해물 제조:** 2일간의 형질감염 후, 배양 배지를 흡인하고 세포를 빙냉 PBS로 얼음 위에서 세척하였다. 프로테아제 억제제(ThermoFisher Scientific)를 갖는 빙냉 용해 완충액(RIPA, ThermoFisher Scientific)를 첨가하고 세포를 4°C에서 30분 동안 배양하였다. 세포 스크레이퍼를 사용하여 세포를 수확하고 초음파 처리에 의해 용해시켰다. 4°C에서 20분간 10,000g에서 원심분리하여 세포 파편을 펠렛화하고, 상청액을 새로운 미세원심분리 튜브로 옮겼다. 용해물의 단백질 농도를 웨스턴 블롯을 위한 Bradford 또는 BCA 단백질 분석에 의해 결정하였다.
- [0099] **웨스턴 블롯:** 간단히 말해서, 겔의 각 웰에서, 20-50ug의 단백질을 25 μ l/웰의 총 부하 용적으로 4X SDS 샘플 완충액(ThermoFisher Scientific), 10X 환원 완충액(ThermoFisher Scientific) 및 추가 ddH₂O(ThermoFisher Scientific)와 혼합하였다. 혼합물을 95°C에서 5분간 가열하여 변성시키고 실온으로 냉각시키며, NuPAGE™ 4 내지 12%, Bis-Tris 겔(ThermoFisher Scientific)에 부하하기 전에 원심분리하였다. 전기영동 분리 후, 겔을 카세트로부터 제거하고 iBlot™ 2 Dry Blotting 시스템(ThermoFisher Scientific)을 사용하여 옮겼다. TBST 중의 5% 무지방 분유로 전달 막을 RT에서 1시간 동안 차단하고, 1차 항체와 함께 4°C에서 밤새 배양하고, TBST(TBS 중의 0.05% Tween20) 완충액으로 3회 세척하고, HRP 접합된 마우스 IgG (H+L) 2차 항체(ThermoFisher Scientific, A24512)인 2차 항체와 실온에서 1시간 동안 배양하였다. 전달 막을 Pierce ECL 웨스턴 블롯팅 기질(Thermo Fisher Scientific)에 의해 전개시키고 화학발광 이미징 시스템을 사용하여 이미징하였다.
- [0100] **실시예 5: 생체내 제형 제조**

- [0101] **1. PNI-Genvoy LNP 제형:** 지질 나노입자는 시험관내 및 생체내 분석 모두에서 PNI NanoAssemblr(Precision NanoSystems, Vancouver, British Columbia, Canada)을 양성 대조군으로 하여 GenVoy 플랫폼을 사용하여 제형화하였다.
- [0102] **2. HKP(+H) 제형:** HKP(+H) 스톱 용액(10mg/mL)을 뉴클레아제가 없는 물에서 제조하였다. 농축된 스톱 용액을 물에서 4mg/mL로 희석하였다. mRNA 작업 용액(1mg/mL)을 1mM 시트레이트 완충액(pH 6.0)에서 제조하였다. mRNA/HKP(+H) 다중체는 동일한 용적의 4mg/mL HKP(+H) 및 1mg/mL mRNA를 혼합하여 형성하였다. HKP(+H) 대 mRNA의 질량비는 4:1이었다. mRNA/HKP(+H) 다중체를 사용 전에 실온에서 30분 동안 배양하였다. 각 캡티드-기반 다중체의 크기는 형질감염 또는 주입 전에 Zetasizer(Malvern Panalytical)로 결정하였다.
- [0103] **3. HKP(+H)/DOTAP 제형(후-혼합된 DOTAP):** HKP(+H) 스톱 용액(10 mg/mL)을 뉴클레아제가 없는 물에서 제조하였다. 농축된 스톱 용액을 물에서 4mg/mL로 희석하였다. DOTAP(Sigma-Aldrich)는 수성 완충 용액 중의 1mg/mL이다. mRNA 작업 용액(1mg/mL)을 1mM 시트레이트 완충액(pH 6.0)에서 제조하였다. 먼저 mRNA/HKP(+H) 다중체는 동일 용적의 4mg/mL HKP(+H) 및 1mg/mL mRNA를 혼합하여 형성하였다. mRNA/HKP(+H) 다중체를 실온에서 30분 동안 배양하였다. 다음으로, HKP(+H) 용액에 대해 동일 용적의 DOTAP를 mRNA/HKP(+H) 다중체에 첨가하였다. HKP(+H)/DOTAP 대 mRNA의 질량비는 4:1:1이었다. mRNA/HKP(+H)/DOTAP 나노입자를 사용 전에 실온에서 30분 동안 배양하였다.
- [0104] **4. HKP(+H)/MC3 또는 HKP(+H)/DOTAP 제형(사전-혼합된 MC3 또는 DOTAP):** HKP(+H) 스톱 용액(10mg/mL)을 뉴클레아제가 없는 물에서 제조하였다. 농축된 스톱 용액을 물에서 4mg/mL로 희석하였다. DOTAP 또는 MC3는 수성 완충 용액 중의 1mg/mL이다. mRNA 작업 용액(1mg/mL)을 1mM 시트레이트 완충액(pH 6.0)에서 제조하였다. 동일 용적의 HKP(+H) 및 MC3를 4:1 질량 비율로 사전-혼합하고 HKP(+H) 용액에 대한 동일 용적의 mRNA를 사전-혼합된 HKP(+H)/MC3에 첨가하였다. mRNA/HKP(+H)/MC3 나노입자는 사전-혼합된 4mg/mL HKP(+H)/1mg/mL MC3 및 1mg/mL mRNA를 혼합하여 형성하였다. HKP(+H)/MC3 대 mRNA의 질량비는 4:1:1이다. mRNA/HKP(+H)/MC3 나노입자를 사용 전에 실온에서 30분 동안 배양하였다.
- [0105] **5. HKP(+H)/PLA NP 제형:** HKP(+H) 스톱 용액(10mg/mL)을 뉴클레아제가 없는 물에서 제조하였다. 농축된 스톱 용액을 물에서 4 mg/mL로 희석하였다. 폴리-L-락트산(PLA) 나노입자(5mg/mL)를 물에서 제조하였다. mRNA 작업 용액(1mg/mL)을 1mM 시트레이트 완충액(pH 6.0)에서 제조하였다. 동일 용적의 HKP(+H) 및 mRNA를 4:1 질량 비율로 혼합하였다. mRNA/HKP(+H) 다중체를 실온에서 30분 동안 배양한 다음 HKP(+H) 용액에 대한 동일 용적의 PLA 나노입자를 mRNA/HKP(+H) 나노입자에 첨가하여 mRNA/HKP(+H) 다중체를 PLA 나노입자 표면에 흡착시켰다. HKP(+H)/PLA 대 mRNA의 질량 비율은 4:5:1이었다. mRNA/HKP(+H)/PLA 나노입자를 사용 전에 실온에서 30분 동안 배양하였다.
- [0106] **실시예 6: 생체내 동물 모델 및 주입**
- [0107] 생체내 연구는 Antibody and Immunoassay Consultants(AIC) Inc(Gaithersburg, MD)에서 수행하였다. 간단히 말해서, 8마리의 6-8주령 암컷 BALB/C를 각 그룹당 4마리 마우스로 하여 각 그룹으로 무작위배정하고 상이한 제형으로 30 µg의 EGFP를 오른쪽 옆구리에 근육내 주사하였다. 상기 언급한 동일한 제형으로, CleanCap SARS-COV 2 스파이크 단백질(S6P) 다중체를 생체내 분석 및 항체 역가 측정 및 결합을 위해 제조하였다.
- [0108] 28일째에, 부스팅을 위한 2차 주사를 하고 35일째에 혈청을 채취하여 항체 역가 측정을 위해 면역분석법(ELISA)으로 분석하였다.
- [0109] **결과**
- [0110] **면역형광 분석**
- [0111] 제조한지 1주일 후, 제형을 HEK 293T 세포에 전달하였다(도 2). 모든 HKP(H) 그룹은 시험관내에서 mRNA를 세포로 운반할 가능성을 입증하였다. HKP(H)군 중에서, HKP(H) 단독 그룹이 다른 그룹 및 PNI Genvoy(양성대조군)보다 높은 EGFP 발현을 보였다.
- [0112] **웨스턴 블롯 분석**
- [0113] 다음으로 웨스턴 블롯 분석을 수행하였으며 결과(도 3)는 S1P D614G mRNA 전달 그룹으로부터 세포 표면에 발현된 스파이크 단백질이 세포막에 국소화 후 웨딩을 겪는 것으로 나타났다. 스파이크 단백질의 S2 도메인만 세포 표면에 남아 있었다. 웨스턴 블롯의 밴드 크기(70-80 KDa)가 이 가설을 뒷받침하였다. 그러나 S2P D614G로부

터의 스파이크 단백질은 스파이크 단백질의 온전한 크기(180 kDa)를 보였다.

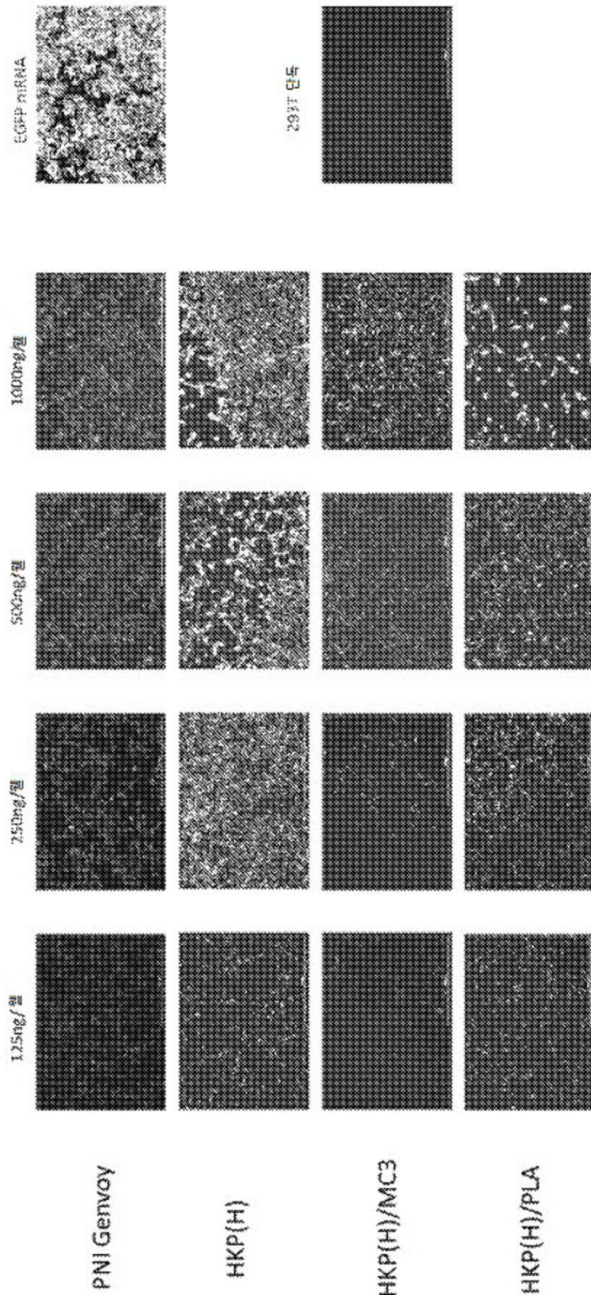
도면

도면1



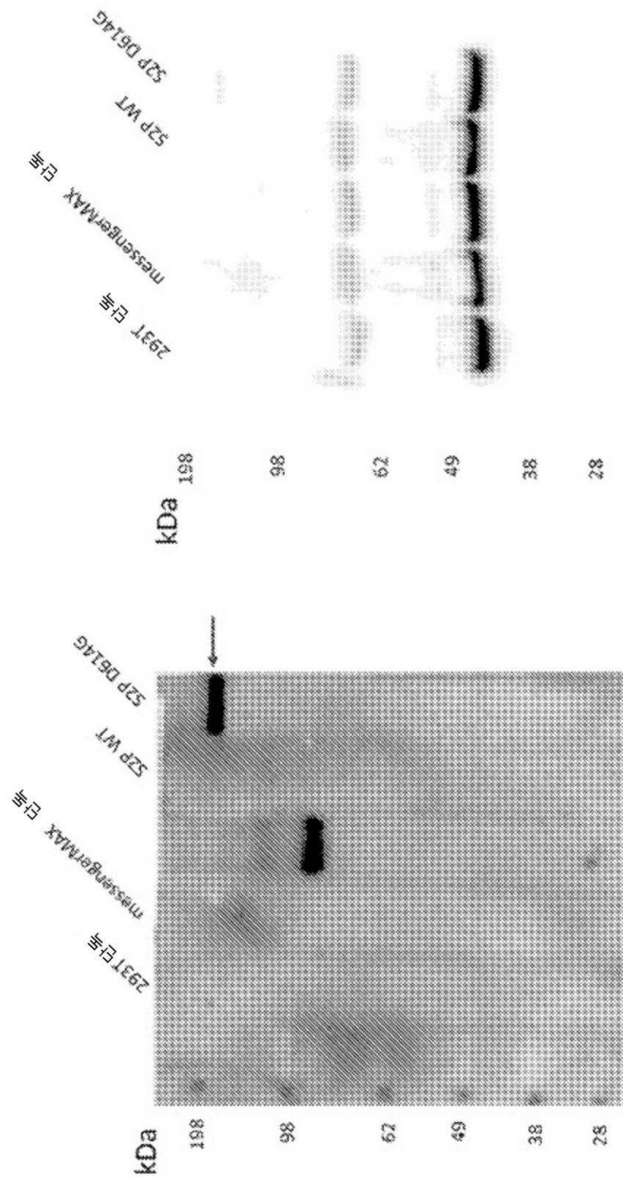
최적화된 mRNA 백신 발현 구조의 개략적 도식

도면2



mRNA 일시적 형질감염 후 스파이크 단백질 발현의 면역형광 분석 결과. 형질감염 조건:
 ThermoFisher Scientific Lipofectamine MessengerMAX 형질감염 시약
 프로토콜 (Pin. No. MAN0010803 Rev D.0)

도면3



웨스턴 블롯

SARS-CoV/SARS-CoV-2 스파이크 단백질 S2
단클론 항체 (Product # MA5-35946)

항- β -액틴 항체
(Product # ab49900)

mRNA 일시적 형질감염 후 스파이크 단백질 발현의 웨스턴 블롯 결과. 형질감염 조건 :
: ThermoFisher Scientific Lipofectamine MessengerMAX 형질감염
시약 프로토콜 (Pin. No. MAN0010803 Rev D.0). 1차 항체 :
SARS-CoV/SARS-CoV-2 스파이크 단백질 S2 단클론 항체 (ThermoFisher Scientific, MA5-35946). 2차 항체 : 마우스 IgG (H+L) 2차 항체
(ThermoFisher Scientific, A24512). 부하 대조군 : HRP 항-베타 액틴 항체
(abcam, ab49900)

도면4a

스파이크 단백질 ORF 클론 1 (MN996531: viral cDNA: 21523-23040) SEQ ID NO:1

atgttctt gttacaact 21541
 aaacgaacaa tgtttgttt tcttgttta ttgccactag tctctagtca gtgtgttaat cttacaacca gaactcaatt
 accccctgca tacactaatt cttccacacg tgggtgttat taccctgaca aagittctag atcctcagtt ttacatfca
 ctgaggacti gttcttacct tcttttcca atgttacttg gtccatgct atacatgtct ctgggaccaa tggfactaag
 aggtttgata accctgtcct accattaat gatgggtgtt attttcttc cactgagaag tctaacataa taaggaggctg
 gatttttgg actactitag atcgaagac ccagtcacct cttattgta ataacgctac taatgttgtt ataaagtct
 gtgaatttca atttgtlaat gatccatttt tgggtgttta ttaccacaaa aacaacaaaa gtggatgga aagtgagttc
 agagttatt ctagtgcgaa taattgcact ttgaaatag tctctcagcc ttttctatg gacctgaaag gaaaacaggg
 taattcaaa aatcttaggg aatttgtgt taagaatatt gatgggtatt taaaatata ttctaagcac acgcctatta
 atttagtgcg tgaatccct caggggtttt cggctttaga accattggta gattgcca taggtatfaa catcactagg
 ttcaaacit tacttcttt acatagaagt tatttgactc ctgggtactc tcttcagggt tggacagctg gtgctgcagc
 ttattatgtg ggttatctc aacctaggac ttcttatta aaatataatg aaaatggaac cattacagat gctgtagat
 gtgcactga ccctctctca gaaacaaagt gtacgtgaa atcctcact gtagaaaaag gaatctatca aacttcaac
 tttagagtc aaccaacaga atctattgtt agatttcta atattacaaa ctgtgcccct ttgggtgaaag ttttaacgc
 caccagatt gcaatcttt atgctggaa caggaagaga atcagcaact gtgtgctga ttattctgtc ctataaatt
 ccgcatcatt ttccatttt aaggttatg gagtctctcc tactaaatta aatgatctct gctttactaa tgtctatgca
 gattcattg taattagagg tgaatgaagc agacaaacg ctccagggca aactggaag atgtctgatt ataattataa
 attaccagat gatttacag gctgcgttat agcttggat tctaacatc ttgattctaa ggttgggtgt aattataat
 acctgtatag atgttttagg aagtctaac tcaaccttt tgaagagat attcaactg aaatctatca ggccggtagc
 acacctgta atgggttga aggttttaatt tttactttc cttacaac atattgtttc

스파이크 단백질 ORF 클론 2, (MN996531: viral cDNA: 23041-23740) SEQ ID NO:2

23041
 caaccacta atgggttgg ttaccaacca tacagagtag tagtacttc tttgaacll ctacatgac cagcaactgt
 ttgtggacct aaaaagtcta ctaatttgg taaaaacaaa tgtgtcaatt tcaactcaa tggtttaaca ggcacaggtg
 ttcttactga gtctaacaaa aagtttctgc cttccaaca atttggcaga gacattgctg acaclactga tgcgtccgt
 gatccacaga cacttgagat tcttgacatt acaccatgtt ctttgggtgg tgcagtggt ataacaccag gaacaaatc
 ttctaaccag gtgtctgttc ttatcagga tgttaactgc acagaagtc ctgttctat tcatgcagat caacttactc
 ctacttggcg tgtttattc acaggttcta atgttttca aacacgtgca ggtgtttfaa tagggctga acatgtcaac
 aactcatag agtgtgacat accatttgg gcaggtatat gcgctagta tcagactcag actaattctc
 ctccggggcg acgtatgta gctagtcaat ccatcattgc ctacactatg tcacttgggt cagaaaaac agttgcttac
 tctaataact ctattgccat accacaaat ttactatta gtgttaccac agaaattcta ccagtgtcta

도면4b

스파이크 단백질 ORF 클론 3 (MN996531: viral cDNA: 23740-24540) SEQ ID NO:3

23740

tgaccaagac atcagtagat tgfacaatgt acallftggtg tgattcaact gaatgcagca atctlttftt gcaatatggc
 aglttttgia cacaaftaaa cegtgtcttta actggaatag ctgttgaaca agacaaaaac acccaagaag tttttgcaca
 aglcaaaaaa attlcaaaaa caccaccaat laaagalltt ggtgttttta attttcaca aatattacca gatccatcaa
 aaccaagcaa gaggtcalft attgaagalc tacitftcaa caaagtgaca ctgtcagatg ctggcttcat caacaatat
 ggtgaltgcc ttggtagat tgctgtclaga gacctcattt glgcacaaaa gtttaacggc ctactgttt tgccaccttt
 gctcacagat gaaatgattg ctcaalacac ttctgcactg ttagcgggia caatcactc tggttggacc ttgggtgcag
 gtgtgtcalt acaaatacca ttgtctatgc aaatggccta taggtltaat ggtattggag ttacacagaa tttctctat
 gagaacaaa aattgatgc caaccaattt aatagtgtca ttggcaaat tcaagactca ctftctcca cagcaagtgc
 acttggaaaa ctcaagatg tgttcaaca aaatgcaca gcttlaaca cgtttgttaa acaacttagc tccaatttg
 gtcaatttc aagtgtttta aatgataacc ttacagctc tgacaaagt gaggtgaag tgcaaatga taggtgac

스파이크 단백질 ORF 클론 4 (MN996531: viral cDNA: 24540-25371) SEQ ID NO:4

24540

acaggcagac tcaaaagt ttacacatat gtgactcaac aattaattag agctgcagaa atcagagctt
 ctgttaact tgctgtact aaaatgtcag agtgtgtact tggacaatca aaaagagttg atttttgg aaagggtat
 catctatgt ccttccctca gtcagacct catggttag tcttctgca tgtacttat gtccctgcac aagaaaagaa
 ctcaacaact gctcctgcca ttgtcatga tggaaaagca cacttccctc gtgaaggtgt ctftgttca aatggcacac
 actggttftg aacacaaagg aattttatg aaccacaaat cactactaca gacaacacat ttgtgtctgg taactgtat
 gttgtaalag gaaatgtcaa caacacagtt fatgactct tgaacctga attagactca tcaaggagg agttagataa
 atattttaa aatcatacat caccagatgt tgatttagt gacatctctg gcaatattc ttcaattgia aacattcaa
 aagaattga cgcctcaat gaggttcca agaattaaa tgaatctc atcactcc aagaactgg
 aaagtatgag cagtatataa aatggccatg gtacattgg ctaggttta tagctggctt gatgccala gtaattgga
 caattatgct ttgtgtatg accagtgtct gtagttgtc caagggtctg ttctctgtg gatcctgctg caaattgat
 gaagacgact ctgagccagt gctcaaaagga gtcaaatfac attacacata a

ORF1a 폴리머라제 단백질 클론 1 (MN996531: viral cDNA: 6253-7140) SEQ ID NO:5

6253

gccacgta taaaccaat acctggtgta tacgtgtct ttggagcaca aaaccagtg aaacatcaa ttcgtttgat
 gfactgaagt cagaggacgc gcagggaaatg gataacttg cctgcgaaga tclaaacca gtctctgaag
 aagtatgga aaatctacc atacagaag acgttctga gtgtaatgtg aaaactaccg aagtgtagg
 agacattata ctlaaccag caaataatag tttaaaaat acagaagagg ttgccacac agatctaag gctgctatg
 tagacaatc tagcttact attaagaac claatgaat atctagatg ttaggttga aaaccttgc tactcatgt
 ttactgtctg ttaatagtgt ccttgggat actatagctia attatgctaa gccitttctt aacaaagtg ttagtacaac
 tactaacata gttacacggt gtttaaccg ttgttact aattatagc ctatttctt tacttattg ctacaattgt
 gfactttac tagaagtaca aattctagaa ttaaagcacc tatgccgact actatagcaa agaactactg taagagtgc
 ggtlaatttt gtctagaggc tcaattat fattgaagt cacctaatt ttctaaactg ataatatta taatttgggt

도면4c

tttactatta agtgtttgcc taggtcttt aatctactca accgctgctt taggtgtttt aatgtctaat ttaggcacgc
cttcttactg tactgttac agagaaggct attgaacac factaatgac actatfgcaa cctactgtac tggttctata
cctttagtg ttgtcttag tggtttagat

ORF1b RdRp 단백질 클론 1 (MN996531: viral cDNA: 15253-16153) SEQ ID NO:6

15253

tagaaaa cctcacctta tgggttggga ttactctaaa tgtgatagag ccatgcctaa catgctttaga attatggcct
cactgttct tgcicgcaaa catacaacgt gtgttagctt gtcacaccgt ttctatagal tagctaatga gtgtgctcaa
gtattgagtg aaatgtcat gtgtggcggg tcactatag taaaccagg tggaaacctca tcaggagatg
ccacaactgc ttatgctaat agtgtttta acattgtca agctgtcacg gccaatgta atgcactttt atctactgat
ggtaacaaaa ttgccgataa gtagtccgc aatttacaac acagacttta ttagtgtctc tatagaata gagatgtga
cacagacttt gtgaatgagt ttacgcata ttgcgtaaa cafttctcaa tgatgalact ctctgacga gctgtgtgt
gtttcaatag cacttatgca tctcaaggct tagtggctag cataaagaac ttaagtcag ttcttatta tcaaaacaat
gtttttatgt ctgaagcaaa atgttggact gagactgacc ttactaaagg acctatgaa ttgtctctc aacatacaat
gctagttaaa cagggtgatg attatgtgta ccttcttac ccagatccat caagaatcct aggggcccggc tgttttag
atgatatcgt aaaaacagat ggtacactta tgaattgaacg gttcgtgtct ttactatag atgcttacc acttactaaa
catctaate agggatgac ttagtcttt cafttgtact tacaatacat aagaaagcta catgatgagt taacaggaca
catgttagac atgtattctg ttatgcttac taa

M 단백질 ORF (MN996531: viral cDNA: 26510-27178) SEQ ID NO:7

a tggcagattc 26521 caacggact attaccgtt aagagcttaa aaagctcctt gaacaatgga acctagfaat
aggtttctta ttcttacct ggalttgtct tctacaatt gcctatgcca acaggaatag gttttgtat ataattaagt
taatttctct ctggctgta tggccagtaa ctttagctt tttgtgctt gctgctgttt acagaataaa ttggatcacc
gggtggaatg ctatgcaat ggcitgtct gttagcttga tgtggctcag ctactcatt gcttcttca gactgttgc
gctgtaccgt tccatgtgtt caftcaatc agaaactaac attcttctca acgtgccact ccatggcact attctgacca
gaccgctct agaaagtga ctcgtaatcg gagctgtgat cttcgtggga cafttctgta ttgctggaca
ccatclagga cgtgtgaca tcaaggacct gcctaaagaa atcactgtt gctatcaccg aacgcttct tattacaat
tgggagcttc gacgctgta gcagggtact caggtttgc tgcatacag cgtctacagga ttggcaacta
taaattaac acagaccatt ccagtagcag tgacaatatt gctttgctg tacagtaa

N 단백질 ORF (MN996531: viral cDNA: 28261-29520) SEQ ID NO:8

28261

atgtctgata atggaccca aaatcagcga aatgcacccc gcattacgtt tgggtggacc tcagattcaa
ctggcagtaa ccagaatgga gaacgcagtg gggcgcgac aaaacaactt cggccccaag gtttaccctaa
taatactgcg tcttgggtca ccgctctcac tcaacatggc aagggaagacc ttaaattccc tcgaggacaa
ggcgttccaa taaacacaa tagcagtcga gatgacaaa ttggctacta ccgaagagct accagacgaa
ttcgtgtgg tgacggtaaa atgaaagatc tcagccaag atggtatttc tactacctag gaactgggcc
agaagctgga ctccctatg gtgctaacaa agacggcacc atatgggttg caactgaggg agccttgaat

도면4d

acacaaaag atcaccattgg caccgcgaat cctgctaaca atgctgcaal cgtgctacaa ctctcfaag
gaacaacaff gccaaaaggc ttctacgcag aaggggagcag aggcggcagc caagcctct ctcgtctc
atcacgtagt cgcacagat caagaaatc aactccaggc agcagtaggg gaactctcc tgcagaatg
gctggcaatg cgggtgatgc tgccttgc tctgctgc ttgacagat gaaccagctf gagagcaaaa
tgcctgtaa aggccaa caacaaggcc aactgtcac taagaaatct gctgctgagg ctctaagaa
gcctggcaa aaacgactg ccaclaaagc atacaalgta acacaagctf tggcagacg tggccagaa
caaaccaag gaattttgg gggaccagaa ctaatcagac aaggaaactga ttacaacat tggccgcaaa
ttgcacaaff tgcgccagc gcttcagcgt tcttcggaa gtcgcgcall ggcatggaag tcaaccttc
gggaacgtgg ttgacctaca cagggtccat caattggat gacaagatc caaaitcaa agatcaagtc
atlltgciga ataagcalat tgacgcatal aaaacatcc caccaacaga gcclaaaag gacaaaaaga
agaaggctga tgaactcaa gcttaccgc agagacagaa gaaacagcaa actgtgactc ttctctgc
tgcagatttg gatgallct caaaacaatl gcaacaatcc atgagcagtg ctgactcaac tcaggcclaa

E 단백질 ORF (MN996531: viral cDNA: 26232-26459) SEQ ID NO:9

26232

atgtactca ttcgtttcgg aagagacagg tacgttaata gtaatagcg tacttcttt tctgcttc gtggtattct
tgcctaglac actagccatc ctactgcgc ttcgattgtg tgcgtactgc tgcfaatfg ttaacgtgag tcttgaaaa
ccttctttt acgtttactc tctgttaaa aatctgaatt cttctagat tctctatct ctggtctaa

도면5a

D614G 돌연변이된 스파이크 단백질 ORF SEQ ID NO:10

atgtttgtttctgttttattgccactagctctagtcagtggttaatcttacaaccagaactcaatfacccttcatacactaat
ctttcacacgtggtgtttatfaccctgacaaaagtttcagatcctcagttttacatcaactcaggactfcttctaccttctttcc
aatgttacttgggttcacatgctatacatgctctcggaccatggctactaagaggtttgataacctgctctaccatfataatggtg
tgtttatttgcctcactgagaagctcaatafaaaggcctggatlltggctactctfagatcgaagaccagtcctact
tattgtaaacgctactaatgtttattaaagctgtgaattcaattttgaaatcattttgggtttattaccacaaaa
caacaaaagtggalggaagtgagttcagagtttctagtgccgaaatggcaatttgaatatgctcagccfctctat
ggaccftgaagaaaacagggtaattcaaaaatcttagggaaatttggtttaagaatattgaggtttatftaaaataattctaa
gcacacgcttaatttagtgcgtgatctccctcagggttttccggcttagaaccttggtagalltgcctaaaggatlaaca
tcaactaggttcaacttctctcttcaatagaagttatggactcctgggtatctctcaggtggacagctgggctgag
cttatttgggttattctcaactaggtcttctatataaataatgaaaatggaaacctatagctgtagactgctgac
ttgacctctctcagaacaaagtgtaggttgaatctcactgtagaaaaggaaatctcaacttcaacttttaggtcca
accacagaatctattgttagatttctaatatacaacttggccctttgggtgaagllttaaagccaccagatttgcactgtt
atgcttggaaacagagagaaatcagcaactggttgcctgatttctgctataataatccgcatcatttccacttttaagttt
atggagtgctcctactaaataatgactctcttactaaagctctatgagatcttgaattagaggtgatgaagtcagac
aatcgtccaggcacaactggaaagattgctgattataatataaataaccagatgatttaccaggtcgtttatagctggaa
ttcaacaatctgattctaaagttgggttaattataaataactgtagatgtaggaaagctcaatctcaaacctlltagaga
gatatttcaactgaaatctacagccggtagcacactgtfaatgggttgaaggttttaattgttactcttcaatcatatg
gtttcaaccactaatggtgttgggttacaaccatacagagtagtagcttctttggaacttctacatgaccagcaactgtt
gtggacttaaaaagctactaaattgggttaaaaacaatgtgtcaattcaactcaatggtttaaagggcacaggtgttctact
gagctcaacaaaagttctgcttccaacaatttggcagagacattgctacactactgatgctgctgctgacacagaca
ctgagattctgacattacacatgttctttgggtgtgctggttataaacacaggaacaaatcttcaaccaggttctgtt
ctttatcagggtgttactgcacagaggtccctgttctatctcagatcaactactcacttggcgtgtttattctacaggt
tclaatgttttcaaacagctgcaggctgttataaggggctgaactgtcaacaactcatatgaggtgacataccattggg
cagggtatgctgctagttatcagactcagactaatctctcctggcggcagctaggttagctagctaatcctcattgcttaca
ctatgctacttgggtcagaaaatcagttgcttactctaatactcttctccatacccaaaaatttactattaggtttacacag
aaatctaccagtgctatgaccaagacatcagtagattgtacaatgtacatttgggtgattcaactgaatgagcaatctttg
ttgcaatggcagttttgtacacaataaacctgtcttactggaaatagctgttgaacagacaaaacaccaagaagttt
ttgcacaagtcaaaaatftacaacacaccacaaftaaagatttgggttttaattttcaacaatattacagatccatcaa
aaccaagcaagaggtcttattgaaatctacttcaacaaggtacacttgcagatgctgctctcaacaacaatattgggt
attgcttgggtatgctgctgtagaacctctatttgcacaaaagttaacggccttactgttttccacccttggctcagat
gaaatgattgtcaatacaactctgcaactgttagcgggtacaacactctggttggaccttgggtgaggtgctgctatcaaaa
taccattgctatgcaaatgcttataggttttaaggttgggttacacagaatgttctctatgagaaccaaaaattgattcc
aaccaattaatagtgctattggcaaaatcaagactcacttctccacagcaaggtgcaacttgaaaactcaagatgtgtca
accaaaatgcacagctttaaaccagctgtttaaacaacttagctccaatttgggtcaatttcaaggtttttaaagatattctt
cagctctgacaaaagttgaggtgaaatgcaaatgataagttgacacagcagacttcaaaattggagacataatgact
caacaatfaatagagctgcagaatcagagcttctgtaactctgctactaataatgtagaggtgtgacttggacaatca
aaaagattgatttttggaaaggctctatctatgcttctcctcagtcagcaccctatgggtgagcttcttctgcatgact
tatgctcctgcacaagaaaagaacttcaacaactcctgctccatttgcagatggaagaacacacttctcctggaaggtgtc
ttgtttcaaatggcacacactggtttglaacacaaaagaaattttatgaaccacaaatcattactacagacaacacatttgtc

도면5b

tggtaactgtagtggtaataagggaattgcaacaacacagttatgatccttgcacacctgaattagactcattcaaggaggag
 ttagataaatafttaagantacacacaccagatgttgatttaggtgacatctctggcaltaatgcttcagftgaaacattcaa
 aaagaaattgaccgacctcaatgagggtgccaagaattaaatgaatctctcatcgatccaagaacctggaaagtatgagca
 gtatataaaatggccatggacalltggctagggtttatagctggcttgattgcatagtaatggtgacaallatgcttggctgat
 gaccagttgctgtagttgctcaaggctgtgttcttgggatcctgctgcaaatgtagaagacgactctgagccagtgctc
 aaaggagtcacaaattacattacacataa

D614G 돌연변이된 스파이크 단백질 아미노산 서열 SEQ ID NO:11

MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVVYYPDKVFRSSVLHSTQD
 LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFG
 TTLDSKTQSLLVNNAITNVVIKVFCEFCNDPFLGVYHKNKSWMESEFRV
 YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
 VRDL PQGFSALEPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAY
 YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFVTEKGIYQTSNF
 RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
 ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYNYK
 LPDDFTGCVIAWNSNLDKVGGNYNLYRLFRKSNLKPFERDISTEIYQAGS
 TPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKK
 STNLVKNKCVNFNENGLTGTGVLTESNKKFLPFQFGRDIADTTDAVRDPQT
 LEILDITPCSFQGVSVITPGTNTSNQVAVLYQGVNCTEVPV AIHADQLTPTWR
 VYSTGSNVFQTRAGCLIGAEHVNNSECDIPIGAGICASYQTQTSNPRRARSV
 ASQSHIAYTMSLGAENSVAYSNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYI
 CGDSTECSNLLLYGSGFCTQLNRALTGIAVEQDKNTQEVFAQVKIYKTPPIK
 DFGGFNFSQILPDPSPKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLI
 CAQKFNGLTVLPLLTDemiaQYTSALLAGTITSGWTFGAGAALQIPFAMQM
 AYRFNGIGVTONVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQ
 NAQALNTLVKQLSSNFGAISSVLDILSRDKVEAEVQIDRLITGRLQSLQTYV
 TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGV
 VFLHVTVYVPAQEKNFTTAPAICHGDKAHFPREGVFSNGTHWFVTQRNFYEP
 QIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVL
 GDISGINASVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWL
 FIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVKGVKLVH
 T

도면6

SARS RBD mRNA 클론 # 2874,

Cleancap 유사체 및 N1-메틸-슈도우리딘-5'-트리포스페이트로 변형된 SARS2 수용체 결합 도메인 (RBD)-foldon mRNA, 인간 SEQ ID NO: 12에 대해 코돈 최적화된

2874

```

ATTAATACGACTCACTATAAGGAGGCCACCATGTTCTGTTTCTGGTGCTG
CTGCCTCTGGTGTCCAGCCAGCGGGTGCAGCCCACCGAATCCATCGTGCG
GTTCCCAATATACCAATCTGTGCCCTTCGGCGAGGTGTTCAATGCCAC
CAGATTCGCCTCTGTGTACGCCTGGAACCGGAAGCGGATCAGCAATTGCG
TGGCCGACTACTCCGTGTGTACAACCTCCGCCAGCTTCAGCACCTCAAGT
GCTACGGCGTGTCCCTACCAAGCTGAACGACCTGTGCTTACAAAACGTG
TACGCCGACAGCTTCGTGATCCGGGGAGATGAAGTGCGGCAGATTGCCCC
TGGACAGACAGGCAAGATCGCCGACTACAACCTACAAGCTGCCCGACGACT
TCACCGGCTGTGTGATTGCCCTGGAACAGCAACAACCTGGACTCCAAAGTC
GGCGGCAACTACAATTACCTGTACCGGCTGTTCCGGAAAGTCCAATCTGAA
GCCCTTCGAGCGGGACATCTCCACCGAGATCTATCAGGCCGGCAGCACCC
CTTGTAACGGCGTGGAAGGCTTCAACTGCTACTTCCACTGCAGTCCTACG
GCTTTCAGCCCAAAATGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTG
CTGAGCTTCGAACTGCTGCATGCCCTGCCACAGTGTGCGGCCCTAAGAA
AAGACCAATCTCGTGAAGAACAATGCGTGAACCTCGGATCCGGAGGTG
GATACATCCCGGAGGCCCTAGGGACGGTCAAGCTTACGTGAGAAAGGA
CGGCGAATGGGTTCTGCTGTGACCTTCTTGGGATAAGCAAGCTGCTTAC
ATAGAACTCGCGGCGATTGGCATGCCGCCTAAAATTTTTATTTATTTTTT
CTTTCTTTTCCGAATCGGATTTTGTTTTAAATTTTCAAAAAAAAAAAAAA
AAAAAAAAAAAAACGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
    
```

D614G 돌연변이된 스파이크 단백질 ORF SEQ ID NO:13 에 대한 단백질 서열
 MFVFLVLLPLVSSQRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRI
 SNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQA
 PGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGNGYNYLYRFRKSNLK
 PFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSF
 ELLHAPATVCGPKKSTNLVKNKCVNFGSGGYPEAPRDGQAYVRKDGWV
 LLST- FLG

도면7a

2019-nCoV 스파이크 단백질에 대한 18개 폴리펩티드 에피토프

SEQ ID	개시	정지	펩티드
14	576	589	RDIADTTDAVRDPQ
15	416	429	VRQIAPGQTGKIAD
16	684	696	QTQTNPPRRARSV
17	258	270	LTPGDSSSGWTAG
18	1146	1157	VYDPLQPELDSF
19	814	825	ILPDPSKPSKRS
20	504	515	YGFQPTNGVGYQ
21	1166	1176	KNHTSPDVDLG
22	482	492	YQAGSTPCNGV
23	80	90	SGTNGTKRFDN
24	30	40	RTQLPPAYTNS
25	1265	1274	FDEDDSEPVL
26	797	806	IYKTPPIKDF
27	606	615	VITPGTNTSN
28	532	541	TVCGPKKSTN
29	1078	1086	PAQEKNFTT
30	781	789	VEQDKNTQE
31	448	456	NNLDSKVG

SEQ ID NO:32 - 49

2019-nCoV 스파이크 단백질의 18개 에피토프에 상응하는 mRNA 서열

SEQ ID	개시	정지	뉴클레오타이드 서열
32	576	589	AGGGACATCGCCGACACCACCGACCCGTGAGGGACCCCCAG
33	416	429	GTGAGGCAGATCGCCCCGGCCAGACCGGCAAGATCGCCGAC
34	684	696	CAGACCCAGACCAACAGCCCCAGGAGGCCAGGAGCGTG
35	258	270	CTGACCCCCGGCGACAGCAGCAGCGGCTGGACCGCCGGC
36	1146	1157	GTGTACGACCCCTGCAGCCCGAGCTGGACAGCTTC
37	814	825	ATCCTGCCCCGACCCAGCAAGCCAGCAAGAGGAGC
38	504	515	TACGGCTTCCAGCCACCAACGGCGTGGGCTACCAG
39	1166	1176	AAGAACCACACAGCCCCGACGTGGACCTGGGC
40	482	492	TACCAGGCCGGCAGCACCCCTGCAACGGCGTG

도면7b

41	80	90	AGCGGCACCAACGGCACCAAGAGGTTTCGACAAC
42	30	40	AGGACCCAGCTGCCCCCGCTACACCAACAGC
43	1265	1274	TTCGACGAGGACGACAGCGAGCCCGTGCTG
44	797	806	ATCTACAAGACCCCCCATCAAGGACTTC
45	606	615	GTGATCACCCCCGGCACCAACACCAGCAAC
46	532	541	ACCGTGTGCGGCCCAAGAAGAGCACCAAC
47	1078	1086	CCCCCCAGGAGAAGAACTTACCACC
48	781	789	GTGGAGCAGGACAAGAACACCCAGGAG
49	448	456	AACAACCTGGACAGCAAGGTGGGCGGC

SEQ ID NO: 50-55

인간 면역계에 의해 인식되는 다른 코로나바이러스에 대해 정렬함으로써 확인된 6개의 고유한 삽입

SEQ ID	개시	정지	펩티드	뉴클레오티드
50	153	167	YYHKNNKSWM ESEFR	TACTACCACAAGAACAACAAGAGCTGGATG GAGAGCGAGTTCAGG
51	76	87	AIHVS GTNGTK R	GCCATCCACGTGAGCGGCACCAACGGCACC AAGAGG
52	252	266	ALHRSYLTPGD SSSG	GCCCTGCACAGGAGCTACCTGACCCCGGC GACAGCAGCAGCGGC
53	447	457	NNLDSKVGGNY	AACAACCTGGACAGCAAGGTGGGCGCAAC TAC
54	479	495	TEIYQAGSTPCN GVEGF	ACCGAGATCTACCAGGCCGGCAGCACCCCC TGCAACGGCGTGAGGGCTTC
55	683	692	QTQTNSPRRA	CAGACCCAGACCAACAGCCCCAGGAGGGCC

SEQ ID NO: 56 - 66

2019-nCoV 바이러스 감염에 대한 예방접종을 위해 가능한 항원로서 역할을 하는 다양한 바이러스 펩티드에 대해 암호화된 RNA 서열

SEQ ID	개시	정지	펩티드 서열	뉴클레오티드 서열
56	606	615	VITPGTNTSN	GUGAUCACCCCGGCACCAACACCAGCAA C
57	532	541	TVCGPKKSTN	ACCGUGUGCGGCCCAAGAAGAGCACCAA C

도면7c

58	1078	1086	PAQEKNF TT	CCCCCCAGGAGAAGAACUUCACCACC
59	781	789	VEQDKNTQE	GUGGAGCAGGACAAGAACACCCAGGAG
60	448	456	NNLDSKVGG	AACAACCUUGGACAGCAAGGUGGGCGGC
61	153	167	YYHKNNKSWM ESEFR	UACUACCACAAGAACAACAGAGCUGGAU GGAGAGCGAGUUCAGG
62	76	87	AIHVS GTNGTK R	GCCAUCCACGUGAGCGGCACCAACGGCAC CAAGAGG
63	252	266	ALHRSYLTPGD SSSG	GCCUGCACAGGAGCUACCUGACCCCGG CGACAGCAGCAGCGGC
64	447	457	NNLDSKVGGNY	AACAACCUUGGACAGCAAGGUGGGCGCAA CUAC
65	479	495	TEIYQAGSTPCN GVEGF	ACCGAGAUCUACCAGGCCGGCAGCACCC CUGCAACGGCGUGGAGGGUUC
66	683	692	QTQTNSPRRA	CAGACCCAGACCAACAGCCCCAGGAGGGC C

도면 8a

SEQ ID No. 67

D614G_S6P_DNA_서열

atgtttgtttttctgttttallgccactagctcttagtcagtggttaactctacaaccagaactcaattaccctctgatacactaat
 tclttcacacgtgggttttaaacctgacaaggfttcagatcctcagttttacattcaactcaggactgttcttaccctttctttcc
 aatgttacttgggtccatctatacatgctctctggaccaatgggtactaagagggtttgataaccctgctaccattaatgatgg
 tgtttatttctccactgagaagctacaataaagaaggctggattttggtactactttagatcgaagaccagctccctact
 taattgtaataacgctactaaltgttttaaaagctctgaafttcaatttgaatgacccattttgggtgtttattaccacaaaa
 caacaaaagtggatggnaagtgagttcagagtttattctagtcgcaataatgacactttgataatgctctcagcctttcttat
 ggacctgaaagaaaaacagggaatttcaaaaatcttagggaafttgggttaagaataatgatggftatttfaaaatattctaa
 gcacacgccattaafttagtgcgtatcctccctcagggttttcggcttagaacctatgtagatgtgccaataggatataaca
 tcaactagttcaaaccttactcttllacatagaagttatttgactcctgggtgattctctcaggttgacagctgggtctcag
 ctllattatgggtttattctcaacclaggactttlctlaaaaalataatgaaaatggaaacattacagatgctgtagactgctac
 ttgacctctctcagaacaagaagtgacgttgaatcctcactgtagaaaaaggaaatctatcaacttcaactttagagctca
 accaacagaatctattgttagatctcctaattacaacattggtccctttgggtgaagttttaacgccaccagatttgcactgttt
 atgcttggaaacaggaaagaaatcagcaactgtgtgctgattattctgctctataatccgcatctttccacttttaagtgtt
 atggagtgctcctactaaalaaatgatctctgctttactaatgctatgtagcagatcattgttaatttagagggtgatgaagtcagac
 aaatcctccaggggcaaacctgaaagattgctgattataattataaattaccagatgattttacagcctgctgtatagctgaa
 ttcaacaatctgattcaaggttgggtgtaataatataatctctatagatgttttaggaagctcaatctcaaaccttttagaga
 gataattcaactgaaatctcagggcggtagcacactgtatgggttgaaggttttaattgttactttctttacaatcatatg
 gttccaaccctaatgggttgggttacaaccatacagagtagtagacttctttgaaacttctacatgaccacagcaactgttt
 gtagcctaaaaagctactaatttggtaaaaaacaatgtgtcaattcaactcaatggtttaacaggccacaggtgttctact
 gactcaacaaaaagtttctccttccaacaattggcagagacattgctgacactactgatgctgctccggtatccacagaca
 ctggagattctgacattacacatgttcttttgggtgctcagttataacaccaggaaacaatacttcaacagggttgetgtt
 ctttatcaggggttgaactgacagaaagtcctgtgctatfctagcagatcaacttactcacttggcgttttattctacaggt
 tctaattgtttcaaacacgtgaggctgttttaataggggctgaacatgtcaacaactatagtgtagacafaccatfctgtg
 cagggtatgctgctagttatcagactcagactaatctcctctcgggacaggtagtgtagctagctaatccatcttgcctaca
 ctatgtcacttgggtcagaaaatcagttgcttacttaataactctattgccataccacaaafttactattaggtttaccacag
 aaattcaccagtgctatgaccaagacatcagtagatgtacaatgfacatttgggtgattcaactgaaatgacgcaatctttg
 ttgcaataggcagttttgtacacaatlaaacctgcttlaactggaatagctgtttgaacagacaaaaacaccaagaagttt
 ttgcacaagcaacaanaatttacaanaaccaccaaftaaagattttgggttttaattttcacaanaatfaccagatccatcaa
 aaccaagcaagaggtcactattgaagatcactttcaacaaagtgacacttgacagatctggcttcalcaacaatattggt
 gattgctgtgataattgctgtagagacctcatttggcacaaaagtttaacggccttactgttttggccactttgtcacaga
 tgaatgattgctcaatacactctgcaactgttagcgggtacaatcacttctggttggacctttgggtcaggtcctgcatlaca
 ataccattcctatgcaaatggcttatagggttaattgggtattggagttacacagaatgttctctatggaacaaaaatgattgc
 caaccaattlaatagtgctatgtgcaaaattcaagactcactttctccacaccaagtgcaacttggaaaacttcaagatgtggc
 aaccaaaatgcacaagctttaaaccagctttaaacaacttagctccaatttgggtgcaatttcaaggttttaaatgatalcctt
 lcacgctgaccaccctgaggctgaaagtcaaaatgafaggtttatcacaggcagacttcaaggtttgacagacatagtgac
 tcaacaatlaattagagctcagaatcagagcttctgctaatctgctgctactaaaatgtcagagtgctgacttggacaatca
 aaaagagttgallttttggaaagggtcatalcctttagtccctcctcagcagaccctatgggtgtagtcttctgcatgtgact
 latgtccctgcacaagaaaaaacctcaaacctgctcctgctccatttgcagatggaaaaagcacactttctctggaaggtgct
 ttgtttcaaatggcacacactggtttgaacacaaaggaaatttttgaaccacaaatcattactacagacaacatttgtgct

도면8b

tggttaactgtgatgtgtaataaggaattgcaacaacacagttatgatccttgcaacctgaattagactcattcaaggaggag
ttagataaafatfittaagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgctcagttgtaacattcaa
aaagaaatgaccgacctcaatgaggtgccaagaattaaatgaatctctcatcgaatccaagaacttggaaagtatgagca
gtatataaaaatggccatggtlacatftggctaggtttatagctggctgattgcatagtaatggtgacaattatgctttgctgat
gaccagttgctgtagttgctcaagggtgctgttcttctggtgatcctgctgcaaatgtagaagacgactctgagccagtgctc
aaaggagtcaaaftacattacataa

D614G_S6P_아미노산_서열 SEQ ID No. 68

MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFG
TTLDSKTQSLLLIVNNATNVVIKVEFCQFCNDPFLGVYYHKNNKSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFNIDGYFKIYKHTPINL
VRDLPQGFSALEPLVDLPIGINITRFQTLALHRSYLTGDSGSSGWTAGAAAY
YVGYLQPRFTLLKYNENGTITDAVDCALDPLSEKCTLKSFTVEKGIYQTSNF
RVQPTEIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFRDISTEIQAGS
TPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKK
STNLVKNKCVNFNGLTGTGVLTESNKKFLPFQFGRDIADTTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNNSEYCDIPIGAGICASYQTQTNPSRAGSV
ASQIIAYTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYI
CGDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSPKRSPIEDLLFNKVTADAGFIKQYGDCLGDIAARDLI
CAQKFNGLTVLPPLLDEMIQAQYTSALLAGTITSGWTFGAGPALQIPFPMQMA
YRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTPSALGKLQDVVNQNA
QALNTLVKQLSSNFGAISSVLDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQ
QLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVF
LHVTVVPAQEKNFTTAPAICHHDGKAHFPREGVFSNGTHWFVTQRNFYEPQII
TTDNFTVSGNCDVIGIVNNTVYDPLQPELDSFKEELDKYFNHTSPDVLG
DISGINASVNNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFI
AGLIAIVMVTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSEPVKGVKLYHT

도면9

RBD_삼량체_DNA_서열 SEQ ID No. 69

Gccacatgttctgtttctgggtgctgctcctctgggtccagccagcgggtgcagcccaccgaalccatcgtgcggttcc
ccaatataccaatctgtgccccctcggcgaggtgttcaatgccaccagattcctctgtgtacgcctggaaccggaagcg
gatcagcaatgctgctggccgactactcctgctgtacaaactcgcagcttcagcacttcaagtgctacggcgtgtcccct
accaagctgaacgacctgtcttcacaaactgtacgacagcctctgtatccggggagatgaagtgcggcagattgc
ccctggacagacagcaagatcggcactacaactacaagctgcccagcactcaccgctgtgtattgcttgaaca
gcaacaacctggactcgaagtgcggcgaactacaactacactgtaccggctgttccggaagtccaatcgaagcccttcg
agcgggacatctccaccgagatctatcagccggcagcacccttgaacggcgtggaaggcttcaactgctacttcccac
tgcagtctacggcttccagcccacaatggcgtggctatcagccctacagagtggtggtgctgagcttcaactgctgca
tgcccctgccacagtgctggccctaagaaaagaccaatctctgtaagaacaaatgctggaactcggatccggagggtg
gatacatccggaggcccctagggacggtcaagcttacgtgagaaggacggcgaatgggttctgctgtcagaccttctg
ggataa

RBD_삼량체_아미노산_서열 SEQ ID No. 70

ATMFVFLVLLPLVSSQVRVQPTEIVRFPNITNLCPFGEVFNATRFASVYAWNR
KRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVR
QIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKS
NLKPFRDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVV
LSFELLHAPATVCGPKKSTNLVKNKCVNFGSGGGYIPEAPRDGQAYVRKDGE
WVLLSTFLG

도면 10a

VUI-202012_01_S6P_DNA_서열 SEQ ID No. 71

atgfttfttttcttcttatttaccactagctctctagtcagtgfttaattcacaaccagaactcaaltacccccgcatacactaat
tctttcacacggtggttttattaccctgacaaagfttccagatcctcagttttacattcaactcaggactgtcttaccfttcttcc
aatgttacttgggtccatgctatacatgctctgggaccaatggtaclaaagaggttgataaccctgctctaccattaatgafgg
tgtttatttggctccactgagaagctaaacataaagaagcctggatfttggfactactttagatcgaagaccagtcctact
tattgftaataacgctactaattgttattaaagctctgtaattcaatttgaatgatccatttgggtgtttattaccacaaaa
caacaaaagtggatggaagtgagttcagagtttattctagtcggaataattgcactttgaaatgctctcagccfttctfat
ggacctggaagaaaaacagggttaattcaaaaatcttagggaafttggttfaagaataattgagttatftaaatataattctaa
gcacacgctataaatttagtgcgtatcctccctcagggtttccggctttagaaccttggtagatftgccaataggllataaca
tcactaggtttcaacttacttcttaccatagaagttattgacccctgggtgattcttctcaggttgacagctggctgctgag
cttattatggtggttacttcaacttaggacttttctataaaaataatgaaaatggaaacattacagatgctgtagactgctgac
ttgacctctctcagaacaaagtgtagctgtaaaccttctactgtagaaaaggaatctcaaaccttaacttttagagcca
accaacagaatctattgttagalltctaataatacaaaacttggccctttgggtgaagttttaaaccgccaccagatttgcactgttt
atgcttggaaacaggaagagaatcagcaactggttctgctgattattctgctctafataaattccgcatctttccacttttaaggtt
atggaggtgctcctactaaalltaaatgctctgcttactaaagctctatgtagatcatttgaatttagagggtgaggaagtcagac
aaatcctccaggcgaactggaagattgctgattataataaattaccagatgattttacaggtcgtcttagctgtagaa
ttcaacaacttgaattcaaggttgggtgtaataataaattaccgtatagatgtttaggaagctaaatctcaaaccttttagaga
gatattcaactgaatctatcagccggtagcaccctgtatgggttgaagtttgaatttacttcttacaacataatg
gtttcaaccactatggttgggttaccacacagagtagtagtacttttgaactctacatgaccagcaactgttt
gtggacctaaaaagctactaatttggtaaaaaacaatgtgcaattcaactcaatggtttaaaccagcagcaggtgttctact
gagctaaacaaaagtttctgctttccaacaattggcagagacattgatgacactactgatgctgctcgtgattccagagaca
cttgagatcttgacattacacatgcttgggtggtgctagttataacaccggaacaaacttcaaccagggttctgctgt
ctttatcagggtttaaactgcacagagtcctgttctatc atgtagatcaacttactctacttggcgtgtttattctacaggt
tctaattgtttcaaacagctgtaggctgtttaaaggggctgaacatgcaacaactataggtgtagafaccattgggtg
caggtatagcgtgtagttatcagactcagactaattctcattcggggcaggtagttagctagtaactcaatcattgcttaca
ctatgctacttgggtcagaaaattcagttgcttacttaataacttattgcccataccataaatttactatagtttaccacaga
aattctaccaggtctatgaccaagacatcagtagattgtacatgtacatttgggtgattcaactgaatgcagcaactctttgtt
gcaatfatggcagttttgtacacaaftaaaccgtgcttfaactggaatagctgttgaacaagacaaaaacccaagaagtftt
gcacaagtcaacaanafttcaaaaacaccacaataaagatttgggtgtttaaattttcaacaanaftaccagatccatcaaa
accaagcaagaggtcacctfatgaagatctactttcaacaanaagtgacacttgcagatgctgcttcalcaaacanaftatgggtg
attgcttgggtgatattgctgtagagacctcatlgtgcaaaaagtttaacggccttactgtttgcccacttggctcacagat
gaaatgattgctcaalacacttctgctgctttagcgggtacaacacttctgggttggaccttgggtcaggtctgctgattacaaa
taccattctatgcaaatggcttataggttfaatggfatgggttacacagaatgttctctatgagaacaaaaattgattgcc
aaccaalltaatagtgctalltggcaaaatcaagactcacttcttccacaccaagtgcaactggaanaactcaagatgggtca
accaaaatgcacaagcttfaaacacgctgttfaaacactttagctccaatttgggtgcaatttcaaggttttaaatgatactctg
cagcttggaccacctgaggtgaaagctgcaaatgtaggtttagatcacagggcagacttcaaaagttgtagacatattgtagct
caacaalltaatagagctgcagaatcagagcttctgctaacttctgctgctactaaaatgtagaggtgtagtacttggcaatca
aaaagagttgatttttgggaaagggctatctctatgctcctcctcagtcagcaccctatgggtgtagcttctgcatgtgact
tagtccctgcacaagaaaagaactcacaactgctcctcctcatttgcattggaagacacacttccctcgtgaaaggtgct
ttgttcaaatggcacacactggttgaacacaaaggaattttatgaaccacaaatcattactacacacacacatttgtct

도면 10b

ggtaactgtagtggtaataaggaaatgcaacaacacagttatgatccttgcacacctgaatagactcattcaaggaggagt
tagataaataatftaagaalcatacatcaccagatgtagttaggtgacatctctggcattaatgctcagttgaaacattcaaa
aagaaatgaccgctcaatgaggtgccaagaatftaaatgaatctctatcagatccaagaactggaaagtatgagcag
latalaaaatggccatggtacattggctaggtttatagctggcttgattgcatagtaatggtagacaattatgcttgcgtatg
accagttgctgtagttgctcaaggctgtgttcttggatcctgctgcaaatftagalgaagacgactctgagccaagtgcia
aaggagtcaaatatcacataa

VUI-202012_01_S6P_아미노산_서열 SEQ ID No. 72

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVVYDPKVFRRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIRGWIFG
TTLDSKTQSLLIIVNNATNVVIKVFCEFCNDPFLGVYHKNKNSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNREFVFNIDGYFKIYKHTPINL
VRDL PQGFSALEPLVDLPIGINITRFQTLALHRSYLPDGDSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRRKISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGNYNYLYRLFRKSNLKPFRDISTEIYQAGS
TPCNGVEGFNCYFPLQSYGFQPTYGVGYQPYRVVLSFELLHAPATVCGPKK
STNLVKNKCVNENFNGLTGTGVLTESNKKFLPFQFGRDIDDTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNSYECDIPIGAGICASYQTQTNHSHRAGSV
ASQSIHAYTMSLGAENSVAYSNNIAIPINFITISVTEILPVSMTKTSVDCMTYIC
GDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKD
FGGFNSQILPDPSPKRSPIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLIC
AQKFNGLTVLPPLL TDEMIAQYTSALLAGTITSGWTFGAGPALQIPFPMQMA
YRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTPSALGKLQDVVNQNA
QALNLTVKQLSSNFGAISSVLNDILARLDPPEAEVQIDRLITGRLQSLQTYVTQ
QLIRAAEIRASANLAATKMSECVLGGQSKRVDFCGKGYHLSFPQSAHPGVVF
LHVTVYVPAQEKNFAPAICHGDKAHFPREGVVFVSNGTHWFVTVQRNFYEPQII
TTHNTFVSGNCDVVIIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG
DISGINASVVNIQKEIDRLNEVAKNLESIDLQELGKYEQYIKWPWYIWLGF
AGLIAIVMVTIMLCMTSCCCLKGCCSCGSCCKFDEDDSEPVKGVKLHYT

VUI-202012_del_S6P_DNA_서열 SEQ ID No. 73

atgtttgttttctgttttattgccactagctctagtcagtggttaattacaaccagaactcaattacccccgcatacactaat
ctttcacacgtgggtttattaccctgacaaagtttcagatcctcagttttacattcaactcaggactgttcttacccttctttcc
aatgttactgggtccatctatatactgggaccaatggtactaagagggttgataaccctgtctaccattaatgatgggtttatt
ttgcttccactgagaagtcaacataalaagaggctggtttttggtactacttagattcgaagaccagtccttactattgta
ataacgctactaatgtttatlaaagctctgtgaattcaattttgtaatgacattttgggtgtttaccacaaaacaacaaag
ttggatgaaagttagttagagtttattctagtcgaataatgcactttgaaatgctctcagcctttcttatggacctgaa

도면10c

ggaaaacagggttaattcaaaaactctagggatftgttgaagaatftgatggftatftaaaatataatctaaagcacagcct
atfaatttagtgcgtgatccctcagggttttctggcctfagaaccalltggatgattgccaataggatfaaacacactaggttcc
aaacttacttgccttaccatagaagttafttgactcctgggtgatctctcagggtggacagctgggtctgagccttattatgtggg
ttatctcaactaggaclttctatfaaaatataatgaaaatggaaacattacagatgctgtagactgtgacttggccctctc
agaacaagaagtgtacgtfagaactctcactgtagaaaaaggaaatctatacaacttcaactttagagtccaaccaacagaat
ctatgttagatllcctaatafacaacttggccctllgggtgaagllttaaaggccaccagatfagcactgllatgctgggaac
aggaaagagaatcagcaactgtgtgctgattatctgctataataatccgcaatcatttccactlltaaggttattggagtgtc
ctactaaatlaaatgatcttgccttactaatgctatgacagatcattgtaatfagagggtatgaagtcagacaalctgctcca
gggcaaacggaaagattgctgattataataatfaccagatgattllacaggctgctgattagcctggaaattcaacaact
tgaatctaaagggttgggttaattataatfaccctgataagattttaggaagctaaatcaaacctlltagagagatattcaact
gaaactatcagggcggtagcacacttgaatgtgtgaaggtltaattgltactllcttacaatcatatggttccaacc
acttatgggttgggttaccacacacagagtagtagactllctllgaaactlclacatgaccagcaactglltggaccta
aaagtctactaatttggftaaaacaatgtgcaactcaactcaatggtllaacaggcacaggttcttactgagtcaaca
aaaagttctgctllccaacaatttggcagagacattgacactactgctgctccgtagatccacagacactttagattct
tgcattacaccatgctlltgggtgctcaggttataacaccaggaacaacttcaaccagggtgctgcttcttaccagg
gtgttaactgcacagaagctcctgttctattcatgacatcaacttacttacttggcgtgttattctacaggttctaattttt
caaacactgctcaggctgttfaataggggctgaacatgtcaacaactcatatgagtgtgacafaccattgggtcagggtatg
cgtagttatcagactcagactaattctcactcctggcaggtagtgttagctagcaatccatcattgctcactacatgctact
ggtagcaaaaatcagttgcttacttaataacttattgccaataccataaatttactatfagttaccacagaaatctacca
gtgtctatgaccaagacatcagtagattgacaatgtacatttgggtgattcaactgaatgcagcaacttttggctcaatagg
cagttttgtacacaatfaaacctgctttaaactggaaatgctgttgaacaagacaaaaacaccaagaagtttggcacaagt
caaacaaatfacaacaaccaccaatfaaagatttgggtggtttaaattttcaacaatfaccagatccatcaaaccaagca
agaggtcactattgaaatctacttftcaacaagtgacactgacagatgctgcttcatcaacaatattggtgattgcttgg
tgaattgtctgtagagacctatttgcacaaaagttaacggccttacttfttccaccttggctcagatgaaatgattg
ctcaatcacttctgactgttagcgggtacaactctctggttggaccttgggtcaggctcctgcaatcaaanfaccattct
atgcaaatggctatagggttfaatgggtattgggttacacagaatgtctctatgagaacaaaaatgattgccaacaaftaa
tagtctalltggcaaaatcaagactcacttctccacacaaagtgacattggaaaactcaagatgtggcaacaaaatgc
acaagcttfaaacacactggttfaaacacacttagctccaatfllgggtcaatfcaaggttfaatgatactctgacgcttga
cccacctgaggtgaaatgcaaatgtagaggtttagatcacagggcagactcaaaagttgacagacatattgactcaacaalaa
ltagagctgcagaaatcagagcttctctaatcttctgctactaaaatgtcagaggtgtacttggcaaatcaaaaagggtg
atlltgggaaagggtatcattatgcttccctcagtcagcacctcaggtgtgacttllgcaatgacttattgctcctgc
acaagaaaagaacttcaacaactgctcctgccaatgctatgagaaaagcacacttccctgtagaaggtgcttfttcaaat
ggcacacactggttgaacacaaaggaaatlltagaacacaaatcattactacacacacacattggtctggtgaactgtg
atgttgaataggaaatgtcaacaacacagtttatgatcctlltgaacactgaatlagactcattcaaggagggttagataaat
ltaagaatcatalaccagatgttgaatlaggtgacalctctgcatfaatgcttcaagttgaaacttcaaaaagaatfga
ccgctcaatgaggttccaagaattfaaatgaatctctcactgatctccaagaacttggaaagtatgagcagtatataaatg
gccatggtacatttggctagggtttagactgctgctgattgccaatgtaatggtgacaatfatcctllgctgtatgaccagttgctg
tagttgtcicaagggtgttcttctgtagatcctgctgcaatfllgatgaagacgactctgagccagttgctcaaaagggtcaa
attacattacataaa

도면10d

VUI-202012_del_S6P_아미노산_서열 SEQ ID No. 74

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
 LFLPFFSNVTWFHAISGNTGKRFDNVLPFNDGVYFASTEKSNIIRGWIFGTT
 LDSKTQSLIVNNAATNVVIVKVEFQFCNDPFLGVYHKNNKSWMESEFRVYSS
 ANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
 LPQGFSALEPLVDLPIGINITRFQTLALHRSYLTPGDSSSGWTAGAAAYYVG
 YLQPRFTLLKYNENGTITDAVDCALDPLSETKCTLSFTVEKGIYQTSNFRVQ
 PTESIVRFPNITNLCPFGEVFNATRFASVYAWNRRKISNCVADYSVLYNSASF
 TFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPD
 DFTGCVIAWNSNLDKVGNYNYLYRLFRKSNLKPFERDISTEIQAGSTPC
 NGVEGFNCYFPLQSYGFQPTYGVGYQPYRVVLSFELLHAPATVCGPKKSTN
 LVKNKCVNFNGLTGTGVLTESNKKFLPFQFGRDIDDTDAVRDPQTLEIL
 DITPCSFQGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYST
 GSNVFQTRAGCLIGAEHVNNSECDPIGAGICASYQTQTNHSHSRAGSVASQSI
 IAYTMSLGAENSVAYSNNIAIPNFTISVTEILPVSMTKTSVDCTMYICGDST
 ECSNLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF
 NFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQK
 FNGLTVLPLLTDEMIAQYTSALLAGTITSGWTFGAGPALQIPFPMQMAYRFN
 GIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTPSALGKLQDVVNQNAQALN
 TLVKQLSSNFGAISVLDILARLDPPAEVQIDRLITGRLQSLQTYVTQQLIRA
 AEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSMFPQSAPHGVVFLHVT
 YVPAQEKNFTTAPAICHGDKAHFPREGVFSVNGTHWVFTQRNFYEPQIITHN
 TVVSGNCDVIVIGVNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGI
 NASVVNIQKEIDRLNEVAKNLSLIDLQELGKYEYQIKWPWYIWLGFIAGLI
 AIVMVTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSEPVKGVKLYHT

도면11

SEQ ID No. 75 - 88

중합체	분지 서열	SEQ ID No.
-	R = K _n K _n K _n K _n K _n K-	75
-	R = H _n K _n K _n K _n K _n K _n K-	76
-	R = K _n K _n K _n K _n K _n K _n -	77
-	R = H _n K _n K _n K _n K _n K _n K _n -	78
H2K4b	R _A = K _n K _n K _n K _n K _n K _n K _n K _n K _n K-	79
H3K4b	R _B = K _n K _n K _n K _n K _n K _n K _n K _n K-	80
H3K(+H)4b	R _C = K _n K _n K _n K _n K _n K _n K _n K _n K-	81
H3k(+H)4b	R _D = k _n K _n K _n K _n K _n K _n K _n K _n K-	82
H-H3K(+H)4b	R _E = H _n K _n K _n K _n K _n K _n K _n K _n K-	83
HH- H3K(+H)4b	R _F = HH _n K _n K _n K _n K _n K _n K _n K _n K-	84
H4K4b	R _G = K _n HH _n K _n K _n K _n K _n K _n K _n K-	85
H3K(1+H)4b	R _H = K _n HH _n K _n K _n K _n K _n K _n K _n K-	86
H3K(3+H)4b	R _I = K _n HH _n K _n K _n K _n K _n K _n K _n K-	87
H3K(1,3+H)4b	R _J = K _n HH _n K _n K _n K _n K _n K _n K _n K-	88

도면13

D614G S2P 아미노산 서열 SEQ ID No.90

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVVYYPDKVFRSSVLHSTQD
 LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFG
 TTLDSKTQSLNATNVVIVKVEFCNDPFLGVYYHKNNKSWMESEFRV
 YSSANNCTFEYVVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
 VRDLPQGFSALEPLVDLPIGINITRFQTLALHRSYLTPGDSSSGWTAGAAAY
 YVGYLQPRFTLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
 RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
 ASFSTFKCYGVSPTKLNLDLCTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYK
 LPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEIQAGS
 TPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKK
 STNLVKNKCVNFNENGLTGTGVLTESNKKFLPFQFGRDIADTTDAVRDPQT
 LEILDITPCSFQGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
 VYSTGNSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNPSRAGSV
 ASQSHIAYTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYI
 CGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
 DFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLI
 CAQKFNGLTVLPPLLTDEMIAYTSALLAGTITSGWTFGAGAALQIPFAMQM
 AYRFNGIGVTVQNVLYENQKLIANQFNQSAIGKIQDSLSSTASALGKLQDVVNQ
 NAQALNTLVKQLSSNFGAISSVLNDILSRLDPPPEAEVQIDRLITGRLQSLQTYV
 TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGV
 VFLHVTYVPAQEKNFAPAICHHDGKAHFPREGVFSNGTHWVFTQRNFYEP
 QIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDIDL
 GDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEYIKWPWYIWLG
 FIAGLIAIVMTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSEPVKGVKLHY
 T

도면14a

SEQ ID No.90 내지 95

SEQ ID	개시	정지	뉴클레오티드
90	153	167	TACTACCACAAGAACAACAAGAGCTGGATG GAGAGCGAGTTCAGG
91	76	87	GCCATCCACGTGAGCGGCACCAACGGCACC AAGAGG
92	252	266	GCCCTGCACAGGAGCTACCTGACCCCCGGC GACAGCAGCAGCGGC
93	447	457	AACAACCTGGACAGCAAGGTGGGCGGCAAC TAC
94	479	495	ACCGAGATCTACCAGGCCGGCAGCACCCCC TGCAACGGCGTGGAGGGCTTC
95	683	692	CAGACCCAGACCAACAGCCCCAGGAGGGCC

SEQ ID No. 96 내지 106

SEQ ID	개시	정지	뉴클레오티드 서열
96	606	615	GUGAUCACCCCCGGCACCAACACCAGCAA C
97	532	541	ACCGUGUGCGGCCCAAGAAGAGCACCAA C
98	1078	1086	CCCGCCAGGAGAAGAACUUCACCACC
99	781	789	GUGGAGCAGGACAAGAACACCCAGGAG
100	448	456	AACAACCUGGACAGCAAGGUGGGCGGC
101	153	167	UACUACCACAAGAACAACAAGAGCUGGAU GGAGAGCGAGUUCAGG
102	76	87	GCCAUCCACGUGAGCGGCACCAACGGCAC CAAGAGG
103	252	266	GCCUGCACAGGAGCUACCUGACCCCCGG CGACAGCAGCAGCGGC
104	447	457	AACAACCUGGACAGCAAGGUGGGCGGCAA CUAC

도면14b

105	479	495	ACCGAGAUCUACCAGGCCGGCAGCACCCC CUGCAACGGCGUGGAGGGCUUC
106	683	692	CAGACCCAGACCAACAGCCCCAGGAGGGC C