



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년10월28일
(11) 등록번호 10-2319508
(24) 등록일자 2021년10월25일

(51) 국제특허분류(Int. Cl.)
G01N 33/569 (2017.01) C12Q 1/70 (2006.01)
G01N 33/68 (2006.01)
(52) CPC특허분류
G01N 33/56983 (2013.01)
C12Q 1/701 (2013.01)
(21) 출원번호 10-2020-0093218
(22) 출원일자 2020년07월27일
심사청구일자 2020년07월27일
(56) 선행기술조사문헌
KR1020160099099 A
W02005035712 A2
Jeong Yoon Lee et al, J Microbiol Biotechnol
(2019), vol 29(8), pp 1316-1323.

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
오종원
경기도 고양시 일산동구 위시티4로 79, 311동 90
2호(식사동, 위시티블루밍3단지아파트)
김민우
서울특별시 서대문구 연희로 38-20, 105동 2002
호(연희동, 연희 대우 아파트)
(74) 대리인
특허법인(유한)아이시스

전체 청구항 수 : 총 19 항

심사관 : 이수진

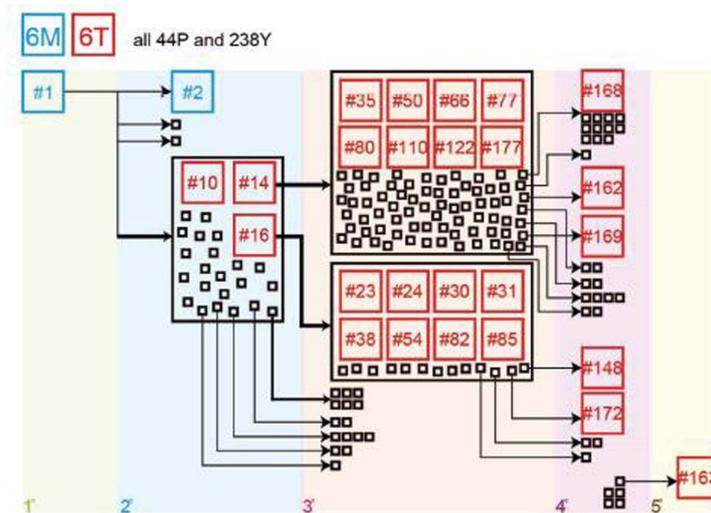
(54) 발명의 명칭 인체 감염형 닥타 유래 메르스코로나바이러스 검출용 바이오마커 및 이의 용도

(57) 요약

본 발명은 메르스코로나바이러스를 식별 또는 검출하기 위한 바이오마커에 관한 것으로서, 보다 구체적으로 인간 감염이 가능한 메르스코로나바이러스를 식별 또는 검출하기 위한 바이오마커로 메르스코로나바이러스의 특정 단백질 변이체와 이 변이체를 코딩하는 유전자서열을 이용하는 기술에 관한 것이다.

본 발명에서 발굴한 메르스코로나바이러스의 ORF4b 단백질 변이체를 마커로 이용할 경우, 인체 감염이 용이한 메르스코로나바이러스를 식별 또는 검출할 수 있으며, 이를 통해 감염을 사전에 차단할 수 있다.

대표도 - 도2



(52) CPC특허분류

G01N 33/68 (2013.01)

C12Q 2600/158 (2013.01)

G01N 2333/165 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	2018-11-0138
과제번호	2018R1AB6000985
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	중견연구자지원사업
연구과제명	tRNA 유래 small RNA에 의한 RNA 바이러스 유전자발현 조절(1/1)
기여율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2018.03.01 ~ 2019.02.28

명세서

청구범위

청구항 1

서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 치환된 단백질 변이체 또는 상기 단백질 변이체를 암호화하는 돌연변이 유전자를 포함하는, 인체 감염형 코로나바이러스 검출용 마커 조성물.

청구항 2

제 1항에 있어서,

상기 단백질 변이체는 6번째 아미노산이 쓰레오닌으로 치환된 것을 특징으로 하는, 마커 조성물.

청구항 3

제 2항에 있어서,

상기 단백질 변이체는 서열번호 2의 아미노산 서열로 이루어진 것을 특징으로 하는, 마커 조성물.

청구항 4

제 1항에 있어서,

상기 돌연변이 유전자는 서열번호 3의 염기서열에서 17번째 염기가 치환된 것을 특징으로 하는, 마커 조성물.

청구항 5

제 4항에 있어서,

상기 돌연변이 유전자는 17번째 염기가 사이토신으로 치환된 것을 특징으로 하는, 마커 조성물.

청구항 6

제 5항에 있어서,

상기 돌연변이 유전자는 서열번호 4의 염기서열로 이루어진 것을 특징으로 하는, 마커 조성물.

청구항 7

제 1항에 있어서,

상기 코로나바이러스는 메르스코로나바이러스인 것을 특징으로 하는, 마커 조성물.

청구항 8

서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 치환된 단백질 변이체 또는 이를 암호화하는 돌연변이 유전자를 검출하는 제제를 포함하는, 인체 감염형 코로나바이러스 검출용 조성물.

청구항 9

제 8항에 있어서,

상기 단백질 변이체는 6번째 아미노산이 쓰레오닌으로 치환된 것을 특징으로 하는, 검출용 조성물.

청구항 10

제 8항에 있어서,

상기 돌연변이 유전자는 서열번호 3의 염기서열에서 17번째 염기가 사이토신으로 치환된 것을 특징으로 하는, 검출용 조성물.

청구항 11

제 8항에 있어서,

상기 단백질 변이체를 검출하는 제제는 상기 단백질 변이체에 특이적으로 결합하는 항체인 것을 특징으로 하는, 검출용 조성물.

청구항 12

제 8항에 있어서,

상기 돌연변이 유전자를 검출하는 제제는 상기 돌연변이 유전자의 mRNA에 상보적으로 결합하는 센스 및 안티센스 프라이머, 또는 프로브(probe)인 것을 특징으로 하는, 검출용 조성물.

청구항 13

제 8항에 있어서,

상기 코로나바이러스는 메르스코로나바이러스인 것을 특징으로 하는, 검출용 조성물.

청구항 14

제 8항의 검출용 조성물을 포함하는, 인간 감염형 코로나바이러스 검출용 키트.

청구항 15

서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 치환된 단백질 변이체 또는 이를 암호화하는 돌연변이 유전자를 검출하는 단계를 포함하는, 인체 감염형 코로나바이러스 검출 방법.

청구항 16

제15항에 있어서,

상기 단백질 변이체는 6번째 아미노산이 쓰레오닌으로 치환된 것을 특징으로 하는, 검출 방법.

청구항 17

제 15항에 있어서,

상기 돌연변이 유전자는 서열번호 3의 염기서열에서 17번째 염기가 사이토신으로 치환된 것을 특징으로 하는, 검출 방법.

청구항 18

제 15항에 있어서,

상기 돌연변이 유전자는 중합효소연쇄반응(polymerase chain reaction), 형광상관분석법(fluorescence correlation spectroscopy), 마이크로어레이(microarray) 및 칩어세이(chip-assay)으로 이루어진 군 중에서 선택되는 적어도 하나 이상의 방법을 통해 검출되는 것을 특징으로 하는, 검출 방법.

청구항 19

제 15항에 있어서,

상기 코로나바이러스는 메르스코로나바이러스인 것을 특징으로 하는, 검출 방법.

발명의 설명

기술 분야

[0001] 본 발명은 메르스코로나바이러스를 식별 또는 검출하기 위한 바이오마커에 관한 것으로서, 보다 구체적으로 인간 감염이 용이한 낙타 유래 메르스코로나바이러스를 식별 또는 검출하기 위한 바이오마커로 메르스코로나바이러스의 특정 단백질 변이체를 이용하는 기술에 관한 것이다.

배경 기술

[0002] 메르스코로나바이러스는 2012년 사우디아라비아에서 처음 발견된 뒤 중동 지역에서 집중적으로 발생한 중동호흡기증후군(Middle East Respiratory Syndrome, MERS; 메르스)의 원인 바이러스다. 초기에는 신종 코로나바이러스로 불렸지만 이후 사우디를 비롯한 요르단·카타르·아랍에미리트(UAE) 등 중동 지역에서 환자가 집중적으로 발생하면서 '메르스코로나바이러스(MERS coronavirus; MERS-CoV)'로 명명됐다.

[0003] 메르스코로나바이러스에 감염되면 2~14일가량의 잠복기를 거친 뒤 섭씨 38도 이상의 고열, 기침, 호흡곤란 등 심한 호흡기 증상이 나타난다. 여기에 설사, 변비 등의 소화기 증상을 보이는 경우도 있으며 만성질환 또는 면역저하자의 경우 폐렴·급성신부전 등의 합병증이 동반되면서 사망에 이르기도 한다. 특히 메르스는 사스보다 치사율이 6배가량 높다는 조사 결과가 나오기도 하는 등 사스보다 더 치명적인 것으로 알려져 있다.

[0004] 메르스코로나바이러스의 감염 사례는 2012년 4월 요르단 자르카의 병원에서 시작되었으며, 2019년 11월까지 7년 동안 27개 국가에서 총 858명의 사망자가 발생하였으며, 사망률은 34.4 %에 달한다.

[0005] 메르스코로나바이러스는 중동 지역의 낙타에서 자주 발견되며, 최초의 기원은 박쥐일 것으로 추정되고 있다. 1983년 수단 및 소말리아에서 수집된 낙타 혈청 샘플이 항체 양성인 것으로 측정된 점에 미루어, 메르스코로나바이러스는 최소 36년 전에 확산되었을 것으로 추정되고 있다. 여러 나라에서 낙타에 대한 누적 혈청 역학 조사 결과는, 메르스코로나바이러스가 중동, 북부 및 동부 아프리카에서 널리 유행하며 파키스탄 및 방글라데시까지도 확산되어 있는 것을 보여주고 있다. 낙타에서 인간으로의 전염사례들과 인간과 낙타에서 분리된 메르스코로나바이러스들에 대한 축적된 유전자정보의 계통발생학적 분석 결과들은 대부분의 인간 감염이 낙타에서 유래한 것을 보여주고 있다. 특히, 중동 지역의 낙타에서 분리된 메르스코로나바이러스는 다양한 유전자 변형을 보이고 있으며, 낙타에서 인간으로의 일방성 감염이 주 감염경로임을 보여주고 있다.

[0006] 그럼에도 불구하고 최근 들어서는, 메르스코로나바이러스의 사람 간 감염이 지속적으로 보고되고 있다. 한국의 경우 2015년 중동에서 온 한 명의 환자에 의해 4개 병원에서 사람 간 감염이 시작되었으며, 2015년 5월 20일 처음으로 감염이 확인된 이후 반년 동안 186명이 감염되었고 38명이 사망하였다. 다만, 한국에서의 발병은 사우디

아라비아에서의 전염과 다른 몇 가지의 차이점을 보인 바 있다. 특징적으로, 한국에서의 감염은 사우디아라비아의 초기 감염(2012년 - 2013년)에 비해 사망률이 낮아 한국의 분리체와 EMC 바이러스 주를 포함한 여러 사우디아라비아의 초기 분리체 간 병원성과 전염성의 차이가 있을 가능성이 시사된 바 있다.

[0007] 한편, 메르스코로나바이러스의 진단은 검체(가래, 기관 흡인물, 기관지 폐포 세척액, 구인두 도말, 비인두 도말, 비인두 흡인물, 비강 흡인물 등)에서 두 가지 이상의 특이 유전자를 검출하거나, 특이 유전자 검출 및 염기서열 확인의 방법으로 이루어지고 있다. 이와 같이 메르스코로나바이러스를 진단하기 위한 마커는 이미 알려져 있으나, 낙타로부터 인간, 나아가 인간 간 전파 및 감염을 용이하게 하는데 관여하는 바이러스 유전자 변이에 대해 현재까지 확인된 바는 없다(공개특허 10-2020-0000045). 따라서, 낙타 또는 인간에 존재하는 메르스코로나바이러스의 어떠한 유전자 변이가 낙타-인간 간 그리고 인간 간 전파와 감염을 촉진시키는지 확인하여, 낙타 등에서 분리된 검체가 이러한 변이를 가지고 있음을 사전에 검출해 인간 감염을 차단할 필요성이 있다.

발명의 내용

해결하려는 과제

[0008] 이에, 본 발명자들은 인간 감염이 용이한 메르스코로나바이러스를 식별 또는 검출하기 위한 마커를 발굴하기 위해 연구 노력한 결과, 메르스코로나바이러스의 ORF4b 단백질 변이체가 주요 마커임을 확인하고, 이에 기초하여 본 발명을 완성하였다.

[0009] 이에, 본 발명은 서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 치환된 단백질 변이체 또는 상기 단백질 변이체를 암호화하는 돌연변이 유전자를 포함하는, 인체 감염형 코로나바이러스 검출용 마커 조성물을 제공하는 것을 목적으로 한다.

[0010] 또한, 본 발명은 서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 치환된 단백질 변이체 또는 이를 암호화하는 돌연변이 유전자를 검출하는 체제를 포함하는, 인간 감염형 메르스코로나바이러스 검출용 조성물을 제공하는 것을 다른 목적으로 한다.

[0011] 또한, 본 발명은 상기 검출용 조성물을 포함하는, 인간 감염형 메르스코로나바이러스 검출용 키트를 제공하는 것을 또 다른 목적으로 한다.

[0012] 또한, 본 발명은 서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 치환된 단백질 변이체 또는 이를 암호화하는 돌연변이 유전자를 검출하는 단계를 포함하는 인체 감염형 코로나바이러스 검출 방법을 제공하는 것을 또 다른 목적으로 한다.

[0013] 그러나 본 발명이 이루고자 하는 기술적 과제는 이상에서 언급한 과제에 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0014] 상기와 같은 본 발명의 목적을 달성하기 위하여, 본 발명은 서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 치환된 단백질 변이체 또는 상기 단백질 변이체를 암호화하는 돌연변이 유전자를 포함하는, 인체 감염형 코로나바이러스 검출용 마커 조성물을 제공한다.

[0015] 본 발명의 일 구현예로, 상기 유전자 변이체는 ORF4b 단백질의 6번째 아미노산을 쓰레오닌으로 치환된 것을 특징으로 할 수 있다.

[0016] 본 발명의 다른 구현예로, 상기 유전자 변이 서열은 서열번호 2의 변이 아미노산 서열로 이루어진 것을 특징으로 할 수 있다.

[0017] 본 발명의 또 다른 구현 예로, 상기 돌연변이 유전자는 서열번호 3의 염기서열에서 17번째 염기가 치환된 것을 특징으로 할 수 있다.

[0018] 본 발명의 또 다른 구현 예로, 상기 돌연변이 유전자는 17번째 염기가 사이토신으로 치환된 것을 특징으로 할 수 있다.

[0019] 본 발명의 또 다른 구현 예로, 상기 돌연변이 유전자는 서열번호 4의 염기서열로 이루어진 것을 특징으로 할 수 있다.

- [0020] 본 발명의 또 다른 구현 예로, 상기 코로나바이러스는 메르스코로나바이러스일 수 있다.
- [0021] 또한, 본 발명은 서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 치환된 단백질 변이체 또는 이를 암호화하는 돌연변이 유전자를 검출하는 체제를 포함하는, 인체 감염형 코로나바이러스 검출용 조성물 및 상기 조성물을 포함하는 키트를 제공한다.
- [0022] 본 발명의 일 구현 예로, 상기 변이체는 ORF4b 6번째 아미노산이 쓰레오닌으로 치환된 것을 특징으로 할 수 있다.
- [0023] 본 발명의 다른 구현 예로, 상기 유전자 변이서열은 17번째 염기가 사이토신으로 치환된 것을 특징으로 할 수 있다.
- [0024] 본 발명의 또 다른 구현 예로, 상기 단백질 변이체를 검출하는 체제는 상기 단백질 변이체에 특이적으로 결합하는 항체인 것을 특징으로 할 수 있다.
- [0025] 본 발명의 또 다른 구현 예로, 상기 돌연변이 유전자를 검출하는 체제는 상기 돌연변이 유전자의 mRNA에 상보적으로 결합하는 센스 및 안티센스 프라이머, 또는 프로브(probe)인 것을 특징으로 할 수 있다.
- [0026] 본 발명의 또 다른 구현 예로, 상기 코로나바이러스는 메르스코로나바이러스일 수 있다.
- [0027] 또한, 본 발명은 서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 치환된 단백질 변이체 또는 이를 암호화하는 돌연변이 유전자를 검출하는 단계를 포함하는, 인체 감염형 코로나바이러스 검출 방법을 제공한다.
- [0028] 본 발명의 일 구현예로, 상기 유전자 변이체는 6번째 아미노산을 쓰레오닌으로 치환시키는 바이러스 유전자(코돈) 변이체이며 단백질 변이체는 6번째 아미노산이 쓰레오닌으로 치환된 것을 특징으로 할 수 있다.
- [0029] 본 발명의 다른 구현예로, 상기 돌연변이 유전자는 17번째 염기가 사이토신으로 치환된 것을 특징으로 할 수 있다.
- [0030] 본 발명의 또 다른 구현 예로, 상기 돌연변이 유전자는 중합효소연쇄반응(polymerase chain reaction), 형광상관분석법(fluorescence correlation spectroscopy), 마이크로어레이(microarray) 및 칩어세이(chip-assay)으로 이루어진 군 중에서 선택되는 적어도 하나 이상의 방법을 통해 검출될 수 있다.
- [0031] 본 발명의 또 다른 구현 예로, 상기 코로나바이러스는 메르스코로나바이러스일 수 있다.

발명의 효과

- [0032] 본 발명에서 발굴한 메르스코로나바이러스의 ORF4b 단백질 변이체 또는 상기 단백질 변이체를 암호화하는 돌연변이 유전자를 마커로 이용할 경우, 낙타에서 인체로 전파될 수 있는 메르스코로나바이러스를 정확히 식별 또는 검출할 수 있으며, 이를 통해 상기 낙타-인간 간 바이러스의 감염을 사전에 차단할 수 있다.
- [0033] 단, 본 발명의 효과는 상기 효과와 진단 방법으로만 한정되는 것은 아니며, 본 발명의 상세한 설명 또는 청구범위에 기재된 발명의 구성으로부터 추론 가능한 모든 진단기술 및 효과를 포함하는 것으로 이해되어야 한다.

도면의 간단한 설명

- [0034] 도 1은 538개 메르스코로나바이러스 분리주의 아미노산 서열 중 ORF4b 단백질에 대한 아미노산 변이를 분석한 결과이다. 박스 안에는 4종류의 대표적 메르스코로나바이러스 분리주의 ORF4b 6번째와 44번째 서열의 차이를 보여주고 있다.

도 2는 2015년 한국에서 발생한 메르스코로나바이러스 감염 환자 간 감염 경로를 나타낸 것이다. 각 네모박스는 감염환자 개체를 나타내며 번호가 들어가 있는 큰 박스는 환자 바이러스 유전자 서열이 보고된 개체이다. 파란색 박스와 적색 박스는 6번째 아미노산 위치에 각각 메티오닌과 쓰레오닌 아미노산을 지닌 ORF4b를 발현하는 바이러스가 분리된 개체이다. 국내에 바이러스 전파를 시작한 1번 환자로부터 5단계 전파과정을 거쳐 163번 환자에 이르는 각 전파과정은 각각 다른 색 바탕으로 나타내었다. 바이러스 유전자 분석이 된 국내 환자로부터 분리한 메르스코로나바이러스의 ORF4b 모두에서 6번째 아미노산과 달리 44번째와 238번째 아미노산 위치에 있는 프롤린과 타이로신은 서열이 잘 보존되어 있었다.

도 3은 2012년 3월부터 2019년 3월 사이 인간에서 분리된 메르스코로나바이러스 유전자의 GC 베이스 함량을 바

이러스 분리 시점에 따라 프로팅 한 분포도이다. 각 서열에서 ORF4b 내 6번째 아미노산 중 메티오닌(M6)을 지닌 것과 쓰레오닌(T6)을 지닌 것은 각각 파란색과 적색으로 표시하였다. 상기 도면으로부터 2015년 초 이전에는 인간에서 분리된 모든 메르스코로나바이러스가 ORF4b M6를 지니고 있으며 GC 함량도 41.18% 이상인 것으로 보여주고 있다. 반면 그 이후 인간에서 분리된 바이러스 들은 거의 모두 ORF4b T6를 지니고 있으며 GC 함량도 41.18% 이하인 것을 볼 수 있다.

도 4는 ORF4b의 메티오닌이 쓰레오닌으로 치환된 낙타 유래 메르스코로나바이러스의 분리 시점에 따른 GC 함량 변화를 나타낸 결과이다. 총 225개의 낙타에서 분리된 메르스코로나바이러스 유전자 서열 중 ORF4b T6를(적색) 지닌 것은 2014년 12월 전후 2015년 3월에 출현빈도가 높아지는 것을 볼 수 있다(아래 박스는 이 기간 중 중첩된 데이터 포인트를 상세히 볼 수 있도록 확대한 그림이다. 각 포인트에 표시한 숫자는 한 포인트에 중첩된 데이터 숫자이다). 주목할 것은 그 이전에는 낙타 유래 바이러스는 예외 없이 GC함량이 41.18% 이상이며 모두 ORF4b M6 (파란색)를 지니고 있음을 볼 수 있다. 도 3과 도 4의 인간 및 낙타 유래 바이러스 유전자 분석 결과는 낙타 내 바이러스 중 일부가 T6로 치환된 ORF4b를 발현하는 유전자를 가지게 됨에 따라 2015년 3월 이후에 이들 변이 바이러스가 인간에서 전파되고 있는 것을 확연하게 보여주고 있다(도 3과 도 4에서 적색 포인트의 분포 참고).

도 5는 2012년부터 2019년 사이 인간에서 분리된 메르스코로나바이러스 분리주의 GC 함량 변화를 나타낸 결과이다. 상기 도면으로부터 2015년 가을을 전후로 하여 바이러스 유전자의 GC 함량의 변화 감소 속도에서 변화가 있음을 확인하였다. 이를 기점으로 바이러스 유전자 GC함량 감소 속도가 감소하는 경향을 보이는, 2단계 변화양상이 나타나고 있다. 도3 및 도5는 2015년 9월 이전에 GC 함량이 급하게 감소되며 동시에 이 기간 중 ORF4b 6번째 아미노산이 메티오닌에서 쓰레오닌으로 변이가 나타나고 이후 인간 분리 바이러스에서 이 서열이 고정되는 양상을 보여주고 있다.

도 6은 메르스코로나바이러스 분리주인 EMC 및 KNIH002의 염기서열 변이 경향을 나타낸 결과이다. 좌측 그래프는 제시한 3종류의 메르스코로나바이러스 분리주간에 유전자 서열 상에서 N (N은 G, A, C, U 베이스 중 하나) 뉴클레오타이드가 W (약한 베이스 결합을 하는 A 혹은 U) 혹은 S (강한 베이스 결합을 하는 G 혹은 C) 로 치환되는 빈도를 보여주고 있다. 전반적으로 EMC 분리주(2012년 분리)와 TSVi 분리주(2018년 분리: 도 5 참조) 유전자를 비교해 보면 N --> W로 치환되는 빈도가 N --> S로 치환되는 빈도 대비 TSVi에서 1.92배 증감하는 것을 볼 수 있다. 또한 이와 같은 염기서열의 비균등 치환 경향을 통해 아미노산 서열도 N --> W 염기서열의 경우 다른 아미노산으로 치환되는 빈도가 높음(2.10)을 볼 수 있다. 이상의 분석 결과는 바이러스 유전자의 N --> W 염기서열 선호적 치환과정을 통해 전체 유전자 GC 함량의 감소를 갖게되고 이 과정 중 일어나는 다양한 ORFs에서 보이는 아미노산 치환이 바이러스의 특성을 변화시키는 역할을 하였을 개연성을 시사하고 있다.

도 7은 메르스코로나바이러스 KNIH002 바이러스 분리주가 감염된 Huh7 세포에서 KNIH002 유전자 유래 small RNA의 유전자 위치별 coverage depth를 보여주고 있다.

도 8은 KNIH002 분리 바이러스주가 감염된 Huh7 세포에서 총 12종류의 가능한 핵산 치환들에 대해 보고된 KNIH002 유전자 서열대비 변화 정도를 전체 치환 건수에 대한 백분율로 계산하여 표시한 파이 차트이다. 이 결과는 도 7에 도시한 한국 분리주 바이러스 KNIH002의 전장 유전자 서열을 커버하는 small RNA들에 대한 염기서열을 분석하여 얻은 결과이다. 이 도에서 볼 수 있듯이 감염된 세포에서도 바이러스 RNA 복제효소는 N --> W 변이, 즉 약한 베이스 결합을 하는 방향으로 치우친 변이를 유도하고 있음을 확인할 수 있었다.

발명을 실시하기 위한 구체적인 내용

[0035] 본 발명자들은 낙타에 존재하는 메르스코로나바이러스 중 인간에 전파될 수 있는 바이러스를 식별하기 위한 바이오마커를 발굴하기 위해 인간 및 낙타 유래 메르스코로나바이러스에 대한 유전자 정보분석을 진행하였다. 그 결과, 메르스코로나바이러스의 ORF4b 변이체를 마커로서 발굴하였다. 보다 구체적으로, ORF4b의 아미노산 서열에서 6번째 아미노산이 쓰레오닌으로 치환(M6T 치환)될 경우, 상기 변이가 낙타에서 인간, 나아가 인간 간 감염에 기여할 수 있는 유전자 변이 서열 정보를 찾아 본 발명을 완성하였다.

[0036] 이하, 본 발명을 상세히 설명한다.

[0037] 본 발명은 서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 M6T로 치환된 단백질 변이체 또는 상기 단백질 변이체를 암호화 하는 돌연변이 유전자를 포함하는, 인체 감염형 코로나바이러스 검출용 마커 조성물을 제공한다.

- [0038] 본 발명에서 사용하는 용어 “코로나바이러스”란 코로나바이러스 과에 속하는 바이러스들을 지칭하며, 외막으로 둘러싸인 RNA 계놈을 가진 바이러스로서 사람, 포유동물, 및 조류 사이에 발견되는 바이러스를 의미한다.
- [0039] 본 발명에 있어서 상기 코로나바이러스는 전파 매개체인 낙타 유래 메르스코로나바이러스일 수 있다. 더욱 바람직하게는, 상기 낙타에서 분리되는 ORF4b M6 잔기를 지닌 메르스코로나바이러스는 RNA 유전체의 GC 함량이 41.18% 이상일 수 있다. 또한 ORF4b M6T 치환을 보이는 인간 분리 메르스코로나바이러스 유전체 함량은 GC 함량이 41.18% 이하일 수 있다.
- [0040] 본 발명에서 사용하는 용어 “변이체”란 유전자를 이루는 염기서열의 변화로 유전정보가 변하면서 유전형질이 달라지는 변이현상을 의미한다. 코돈의 염기쌍 순서 조합에 따라 20개의 아미노산이 만들어지고, 아미노산의 조합에 따라 단백질이 형성되므로, 염기쌍의 순서가 원래와 달라지면 생성되는 아미노산이 달라지며 결과적으로 형성되는 단백질에 변화가 생긴다.
- [0041] 이에 따라, 본 발명에 있어서 “단백질 변이체”는 유전자 염기서열 변화로 인해 형성되는 단백질에 변화가 생김을 의미한다. 단백질 돌연변이는 원래 단백질과 비교해 구조 또는 기능에 변화가 일어나 원래 기능을 수행할 수 없거나 전혀 다른 기능을 수행할 수 있다.
- [0042] 본 발명에 있어서 상기 단백질 변이체는 ORF4b의 특정 위치가 치환되는 것일 수 있으며, 바람직하게는 서열번호 1의 6번째 아미노산이 쓰레오닌으로 치환되는 것일 수 있다.
- [0043] 본 발명에 있어서 상기 단백질 변이체는 서열번호 2의 아미노산 서열로 이루어질 수 있다.
- [0044] 본 발명에 있어서 상기 돌연변이 유전자는 서열번호 3의 염기서열에서 17번째 염기가 치환된 것일 수 있으며, 상기 17번째 염기가 사이토신으로 치환된 것일 수 있다.
- [0045] 본 발명에 있어서 상기 돌연변이 유전자는 서열번호 4의 염기서열로 이루어진 것일 수 있다.
- [0046] 본 발명에서 사용하는 용어 “감염”은 동물이나 식물의 생체조직이나 체액에 침입하여 증식하거나 해로운 물질을 만들어내는 것을 의미한다. 감염의 대상은 인간, 낙타, 박쥐를 포함한 포유류와 그 외에 메르스코로나 바이러스의 증식이 가능한 생명체 일 수 있다.
- [0047] 이에 따라, 본 발명에 있어서 “인체 감염형 코로나바이러스”란 인간을 대상으로 체내에 침입하여 병을 일으킬 수 있는 코로나바이러스를 의미한다. 또한 “낙타 감염형 혹은 낙타 유래 코로나바이러스”는 단봉 낙타를 포함한 메르스코로나 바이러스가 증식할 수 있는 낙타에 침입하여 증식하는 코로나바이러스를 의미한다.
- [0048] 본 발명자들은 구체적인 실시예를 통해 본 발명의 마커 조성물이 인간 감염형 메르스코로나바이러스의 식별이 가능함을 확인하였다.
- [0049] 구체적으로, 본 발명의 일 실시예에서는, 한국에서 전파된 메르스코로나바이러스의 ORF4b의 아미노산 서열을 감염 차수에 따라 비교하였다. 그 결과, 1차 감염자 및 타인에게 바이러스를 전파하지 않는 2차 감염자는 ORF4b의 6번째 아미노산이 메티오닌으로 보존되나, 이를 제외한 2차, 3차, 4차 및 5차 감염자는 ORF4b의 6번째 아미노산이 메티오닌에서 쓰레오닌으로 치환된 것을 확인하였다. 이를 통해 메르스코로나바이러스의 ORF4b의 M6T 치환이 인간 감염이 용이한 메르스코로나바이러스를 식별하는 마커로 활용할 수 있음을 확인하였다(실시예 2).
- [0050] 본 발명의 다른 실시 예에서는, 메르스코로나바이러스의 ORF4b의 M6T 치환과 연관된 분리된 여러 바이러스의 특성 변화를 파악하기 위하여, RNA 계놈 분석을 통해 GC 함량을 결정하였다. 그 결과, M6T 치환이 발생한 메르스코로나바이러스의 경우 GC 함량이 41.18% 이하로 낮게 분포하는 것을 확인하였다(실시예 3).
- [0051] 본 발명의 또 다른 실시 예에서는, 메르스코로나바이러스의 전파에 따라 연도별로 출현하는 바이러스들의 GC 함량이 특정 시점 이전에 출현한 바이러스에 비해 감소한다는 사실에 기초하여, 메르스코로나바이러스가 수 차례 감염을 거쳐 약염기쌍 편향적 돌연변이를 일으켜 인간 숙주에 적응함을 확인하였다(실시예 4).
- [0052] 상기 실시 예의 결과로부터, ORF4b 단백질의 6번째 아미노산이 치환된 단백질, 바람직하게는 쓰레오닌으로 치환된 ORF4b 단백질 변이체 또는 이를 암호화하는 돌연변이 유전자, 바람직하게는 17번째 염기가 사이토신으로 치환된 돌연변이 유전자를 인체 감염형 메르스코로나바이러스 검출용 마커로 이용할 수 있으며, 상기 변이체 또는 돌연변이 유전자를 측정하여 인체 감염형 메르스코로나바이러스를 검출할 수 있다.
- [0053] 또한, 본 발명의 다른 양태로서, 본 발명은 서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 치환된 단백질 변이체 또는 이를 암호화하는 돌연변이 유전자를 검출하는 체제를 포함하는 인체 감염형

코로나 바이러스 검출용 조성물 및 상기 조성물을 포함하는 인체 감염형 코로나 바이러스 검출용 키트를 제공한다.

- [0054] 본 발명에 있어서, 상기 단백질의 변이체를 검출하는 제제는 상기 단백질 변이체에 특이적으로 결합하는 항체일 수 있으나, 이에 제한되는 것은 아니다.
- [0055] 상기 돌연변이 유전자를 검출하는 제제는 상기 돌연변이 유전자의 mRNA에 상보적으로 결합하는 센스 및 안티센스 프라이머, 또는 프로브(probe)일 수 있으며, 바람직하게는 프로브(probe)를 이용하여 검출할 수 있다.
- [0056] 본 발명에서 사용되는 용어, “항체”는 면역학적으로 특정 항원과 반응성을 갖는 면역글로불린 분자를 포함하며, 단클론(monoclonal) 항체 및 다클론(polyclonal) 항체를 모두 포함한다. 또한, 상기 항체는 키메라성 항체(예를 들면, 인간화 무린 항체) 및 이중결합항체(예를 들면, 양특이성 항체)와 같은 유전공학에 의해 생산된 형태를 포함한다.
- [0057] 본 발명에서 사용되는 용어, “프로브”란 효소 화학적인 분리정제 또는 합성과정을 거쳐 제작된 수 염기 내지 수백 염기길이의 mRNA(혹은 mRNA 서열에 해당하는 바이러스 유전자 PCR 증폭 산물 DNA)와 특이적으로 결합할 수 있는 핵산을 의미한다. 방사성 동위원소, 효소, 또는 형광체 등을 표지하여 mRNA의 존재 유무를 확인할 수 있으며, 공지된 방법으로 디자인하고 변형시켜 사용할 수 있다.
- [0058] 본 발명에 따른 인체 감염형 코로나바이러스 검출용 키트는 상기 검출용 조성물을 동결 건조된 형태로 플라스틱 튜브에 담아 제조될 수 있으며, 상기 조성물 외에 인간 감염형 코로나바이러스 식별에 사용되는 적절한 시약 및 도구를 더 포함할 수 있다. 상기 더 포함할 수 있는 시약 및 도구로는 적합한 담체, 검출 가능한 신호를 생성할 수 있는 표지, 용해제, 세정제 등을 예로 들 수 있으나, 이에 제한되는 것은 아니다.
- [0059] 또한, 본 발명의 또 다른 양태로서, 본 발명은 서열번호 1의 아미노산 서열로 이루어진 ORF4b 단백질의 6번째 아미노산이 치환된 단백질 변이체 또는 이를 암호화하는 돌연변이유전자를 검출하는 단계를 포함하는, 인체 감염형 코로나바이러스 검출 방법을 제공한다.

[0060] 상기 돌연변이 유전자는 당업계에서 알려진 통상적인 방법에 따라 중합효소연쇄반응(polymerase chain reaction), 형광상관분석법(fluorescence correlation spectroscopy), 마이크로어레이(microarray) 및 칩어세이(chip-assay)으로 이루어진 군 중에서 선택되는 적어도 하나 이상의 방법을 통해 검출될 수 있으며, 바람직하게는 칩어세이(chip-assay)를 이용하여 검출될 수 있으나 이에 제한되는 것은 아니다,

[0061] 이하, 본 발명의 이해를 돕기 위하여 바람직한 실시예를 제시한다. 그러나 하기의 실시예는 본 발명을 보다 쉽게 이해하기 위하여 제공되는 것일 뿐, 하기 실시예에 의해 본 발명의 내용이 한정되는 것은 아니다.

[0063] **[실시예]**

[0064] **실시예 1. 실험재료 및 방법**

[0065] **1-1. 재료**

[0066] 메르스코로나바이러스의 KNIH002 바이러스주는 질병관리본부에서 제공받았다. KNIH002 바이러스 스타크는 2% fetal bovine serum (FBS)을 첨가한 최소 영양 배지[minimal essential medium (MEM)]에서 자란 Vero E6 세포에서 증식시켰다.

[0067] Huh7 cell은 American Type Culture Collection (ATCC, Rockville, MD, USA)에서 구매하였다. 구매한 Huh7 cell은 10% fetal bovine serum (FBS), 100 U/ml 페니실린 및 100 µg/ml 스트렙토마이신을 첨가한 100% Dulbecco's modified Eagle's medium (DMEM)에서 5% CO₂ 존재 하에 37°C에서 배양하였다.

[0068] **1-2. 메르스코로나바이러스 게놈 서열**

[0069] GeneBank 데이터베이스에 저장된 538개의 메르스코로나바이러스 게놈 서열을 확보하였으며, Clustal Omega를 이용해 서열을 배열하였다. 그 후 Jalview 2.11.0을 이용하여 그래프를 시각화하였다.

[0070] **1-3. 바이러스 감염**

[0071] Huh7 세포를 T25 플라스크에 씨딩 후 밤새 배양하였으며, 그 후 37°C에서 1시간 동안 MOI 0.001의 바이러스와 배양하여 감염시켰다. 감염된 세포는 세척 후 일정기간 동안 2% FBS를 첨가한 배지에서 배양하였다. 감염된 세포 및 배양 상등액은 바이러스 농도, 세포 내 바이러스 RNA 농도 및 바이러스 단백질 발현 수준을 분석하기 위

해 보관되었다.

[0072] 1-4. Small RNA 라이브러리 구축

[0073] Total RNA는 메르스코로나바이러스가 감염된 Huh7 세포에 2일간 감염 후 회수한 세포로부터 TRIzol reagent (Invitrogen)을 이용하여 추출하였고, 제조자의 시험법에 따라 정제하였다. Small RNAcDNA 라이브러리는 TrueSeq Small RNA sample preparation kit (Illumina, San Diego, CA, USA)를 사용하여 제작하였다.

[0074] 1-5. Small RNA deep sequencing 데이터세트의 바이오인포메틱 분석

[0075] Small RNA 서열 관독은 Cutadapt (version 1.9.1)을 이용하여 어댑터를 제거한 후 진행하였다. 16 뉴클레오타이드 길이 이상의 small RNA 서열 관독은 BWA 프로그램 (version 0.7.12)를 이용해 2 뉴클레오타이드의 미스매치를 허용하면서 메르스코로나바이러스 게놈과 안티게놈 서열과 매칭이 되는 것을 선별하였다. 정제된 관독은 메르스코로나바이러스 돌연변이 분석을 위해 사용하였다.

[0077] **실시예 2. 메르스코로나바이러스의 ORF4b에서의 M6T 치환 확인**

[0078] 인간 감염형 메르스코로나바이러스를 검출하기 위한 마커를 발굴하기 위해, 하기와 같은 실험을 수행하였다.

[0079] 먼저, 메르스코로나바이러스가 전파되는 과정에서 단백질의 돌연변이가 일어나는지 확인하기 위해, 메르스코로나바이러스에 감염된 538명의 분리주의 아미노산 서열을 비교하였다.

[0080] 그 결과, 도 1에서 나타난 바와 같이, 538개의 분리주의 ORF4b에서 6, 44 및 238번째 아미노산의 뚜렷한 변화가 나타남을 확인할 수 있었다. 이를 통해, ORF4b를 인간 감염형 메르스코로나바이러스 검출을 위한 마커 후보로 선정하였다.

[0081] 다음으로, ORF4b의 아미노산 서열의 6번째 위치에서 메티오닌이 쓰레오닌으로 치환된 돌연변이의 마커로서의 사용 가능성을 확인하기 위해, 2015년 한국에서 6개월 동안 분리된 메르스코로나바이러스의 게놈 서열이 상호 비교되었다. 표 1은 상기 메르스코로나바이러스의 게놈 서열 리스트를 나타낸다.

표 1

[0082]

GenBank #	환자 번호	나이, 성별	감염 차수	ORF4b 6번째 아미노산
KT326819	1	68, M	1°	M
KT029139	2	63, F	2°	M
KT006149	10	44, M	2°	T
KT374052, KT374053	14b	35, M	2°	T
KT868865	16	40, M	2°	T
KT868866	23	73, M	3°	T
KT868867	24	78, M	3°	T
KT868868	30	60, M	3°	T
KT868869	31	69, M	3°	T
KU308549, KT374054, KT374055	35b	38, M	3°	T
KT868870	38	49, M	3°	T
KX034094	50	81, F	3°	T
KT868871	54	63, F	3°	T
KX034095	66	42, F	3°	T
KX034096	77	63, M	3°	T
KX034097	80	35, M	3°	T
KT868872	82	83, F	3°	T
KT868873	85	66, F	3°	T
KT868874	110	57, F	3°	T
KT868875	122	55, F	3°	T
KX034100	177	50, F	3°	T
KT868876	148	39, F	4°	T
KX034098	162	33, M	4°	T
KT374056, KT374057	168b	36, M	4°	T

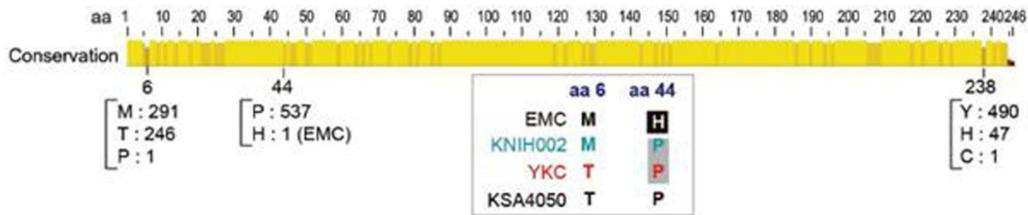
KX034099	169	34, M	4°	T
KT868877	172	61, F	4°	T
KT374051, KT374050	163b	53, F	5°	T

- [0083] 한국에 전파된 메르스코로나바이러스의 첫 번째 환자 및 그 배우자인 두 번째 환자는 ORF4b의 6번째 아미노산이 메티오닌으로 확인되었으나, 이를 제외한 나머지 2차, 3차, 4차 및 5차 감염 환자는 모두 ORF4b의 6번째 아미노산이 메티오닌에서 쓰레오닌으로 치환되었음을 확인할 수 있었다. 도 2에서 나타난 바와 같이, 이러한 ORF4b 6 번째 아미노산의 치환은 메르스코로나바이러스의 한국 전파 초기에 이루어졌다는 것을 확인할 수 있었다.
- [0084] 2차 감염자이자 ORF4b의 6번째 아미노산이 쓰레오닌으로 치환된 돌연변이를 가지는 14번 및 16번 환자는 슈퍼전파자인 것으로 확인되었다. 반면 ORF4b의 6번째 아미노산이 메티오닌으로 유지된 2번 환자는 타인에게 바이러스를 전파하지 않는 것으로 확인되었다. 또한 6번째 아미노산이 쓰레오닌으로 치환된 ORF4b는 사람 간 5차 감염이 진행될 때까지 다른 아미노산으로 치환되는 현상이 발생하지 않았다.
- [0085] 이를 통해 6번째 아미노산이 메티오닌에서 쓰레오닌으로 치환된 ORF4b가 인간 감염이 용이한 메르스코로나바이러스를 식별하는 마커로 이용될 수 있음을 확인하였다.
- [0087] **실시예 3. ORF4b의 M6T 치환과 GC 함량 간 연관 관계 확인**
- [0088] ORF4b의 M6T 치환이 발생한 메르스코로나바이러스 분리주들이 공유하고 있는 유전자 특성을 파악하기 위하여, 메르스코로나바이러스에 감염된 환자들의 RNA 게놈을 분석하였다.
- [0089] 2019년 8월 22일까지 축적된 시퀀스 중에서, 뉴클레오타이드 52-30,073 (게놈 상 삭제된 뉴클레오타이드가 21 이하이며 불명 시퀀스 없음) 부분을 커버하는 총 206개의 시퀀스를 선별하였다. 그 후 선별한 시퀀스의 GC 함량을 바이러스가 분리된 샘플링 날짜와 맞춰 배열하였다.
- [0090] 그 결과, 도 3에 나타난 바와 같이, ORF4b의 6번째 아미노산이 메티오닌으로 유지된 EMC 주(2012년 6월 13일 수집 됨)는 41.27%의 GC 함량을 나타내었다. 반면 M6T 치환이 발생한 분리주 들의 경우 GC 함량이 41.18% 이하인 것으로 확인되었다. 또한, 도 4에 나타난 바와 같이, 낙타에서 분리한 225개의 메르스코로나바이러스에 대한 유전자 분석을 수행한 결과, M6T 치환을 지닌 바이러스주의 경우 GC 함량이 41.18% 이하로 낮게 분포하는 것이 확인되었다. 특히 이러한 바이러스 변이체는 낙타에서 2014년 말부터 2015년 초에 집중적 발생하고 이후는 출현빈도가 낮아지는 것을 볼 수 있었다.
- [0091] 이를 통해, 인간 및 낙타에서 분리된 메르스코로나바이러스 중 GC 함량이 41.18% 이하인 바이러스 분리주에서 특징적으로 ORF4b의 M6T 치환이 보여지는 것을 알 수 있었다.
- [0092] 또한, 도 5에 나타난 바와 같이, 한국에서 2012년부터 2019년까지 인간에서 분리된 메르스코로나바이러스 분리주의 GC 함량이 꾸준히 감소하며, 특히 한국에서 인간 간 메르스코로나바이러스의 전파와 발병이 가장 활발했던 2015년 봄 이전까지 가파른 감소가 나타나는 것을 확인하였다. 이러한 사실은 ORF4b의 M6T 치환 및 41.18% 이하의 GC 함량 특성을 지닌 메르스코로나바이러스에 의한 인간 감염이 용이하다는 것을 시사한다.
- [0094] **실시예 4. 메르스코로나바이러스의 약염기쌍 편향적 돌연변이에 의한 인간 숙주 적응 확인**
- [0095] 본 발명의 또 다른 실시 예에서는, 메르스코로나바이러스의 전파에 따라 연도별로 출연하는 바이러스 분리주들의 GC 함량이 이전에 출연한 것들에 비해 감소한다는 사실에 기초하여, 메르스코로나바이러스가 수 차례 감염을 거쳐 약염기쌍 편향적 돌연변이를 일으켜 인간 숙주에 적응하는지 여부를 확인하기 위해 하기의 분석으로 수행하였다.
- [0096] 먼저, 메르스코로나바이러스 EMC주(2012년 6월 수집), KNIH002주(2015년 5월 수집) 및 TSVi주(2018년 9월 수집)의 염기서열을 비교하였다.
- [0097] 그 결과, 도 6에 나타난 바와 같이, EMC주에 비해 KNIH002주에서 약염기(A 또는 U)로의 돌연변이 비율이 강염기(C 또는 G)로의 돌연변이 비율보다 더 높았으며, KNIH002 균주에 비해 TSVi 균주의 약염기(A 또는 U)로의 돌연변이 비율이 더 높았다.
- [0098] 이를 통해 메르스코로나바이러스가 인간 숙주에 여러 차례 감염됨에 따라 약염기쌍 편향적 돌연변이가 발생한다는 것을 확인할 수 있다.

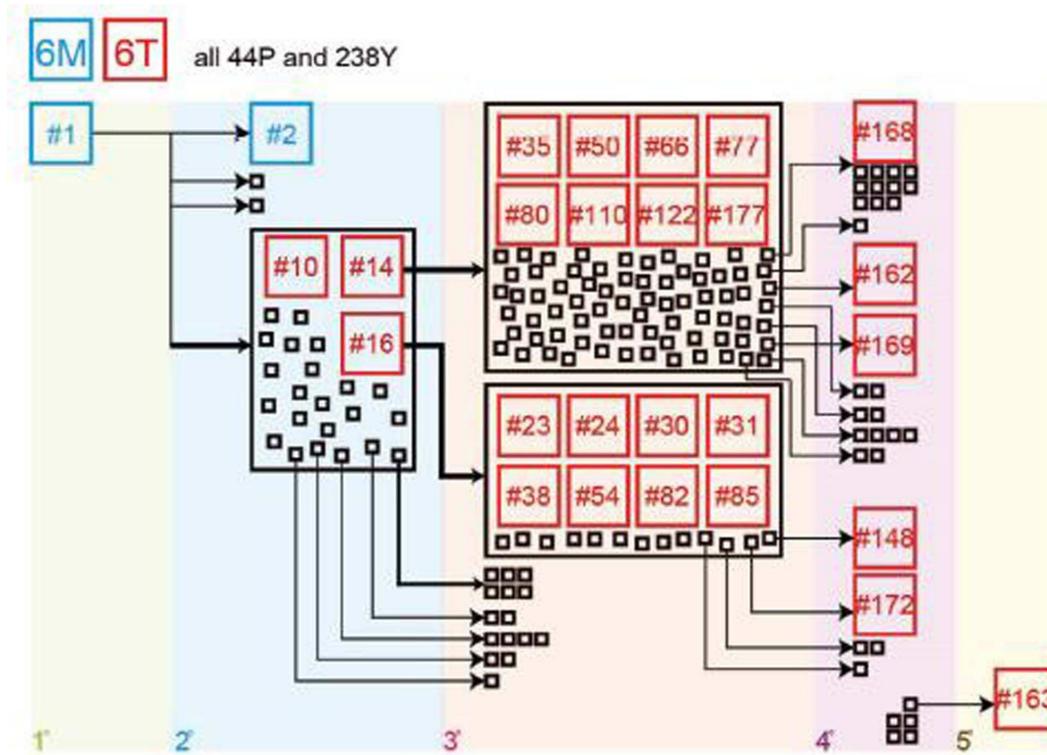
- [0099] 또한, 메르스코로나바이러스의 감염이 반복될수록 GC 함량 감소 경향이 나타나는지 확인하기 위하여, 메르스코로나바이러스가 감염된 세포의 바이러스 게놈에 나타난 돌연변이를 딥 시퀀싱에 의해 확인하는 실험을 수행하였다.
- [0100] 먼저, 도 7에 나타난 바와 같이, KNIH002주로 감염시킨 Huh7 세포에서 KNIH002주의 게놈에서 유래한 small RNA가 확인되었다. 2회의 독립적 서열 분석으로 평균 89,003개의 리드에 대한 돌연변이 프로파일 분석한 결과, 전체 small RNA 중 대략 4.95%가 돌연변이를 일으킨 것으로 나타났다. RNA의 돌연변이를 바이러스 게놈을 구성하는 총 30108개의 염기수, 즉 A는 7,890(26.2%), U는 9,830(32.6%), G는 6,301(20.9%), C는 6,087 (20.2%)로 노말라이제이션 하여 약염기와 강염기로 변화되는 정도를 비교하였다.
- [0101] 그 결과, 도 8에 나타난 바와 같이 약염기로의 돌연변이는 강염기로의 돌연변이 보다 1.6배 더 많은 비율로 나타남을 확인하였다.
- [0102] 보다 구체적으로, 강염기에서 약염기로의 돌연변이는 약염기에서 강염기로의 돌연변이보다 1.32배 더 많은 비율로 나타났다.
- [0103] 이를 통해, 메르스코로나바이러스가 세포에 감염됨에 따라 게놈 상 GC 함량이 감소하는 경향을 나타냄을 확인하였으며, 이는 GC 함량의 감소가 지속적인 감염에 의한 인간 숙주에의 적응에 기여함을 시사한다.
- [0105] 진술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야 한다.

도면

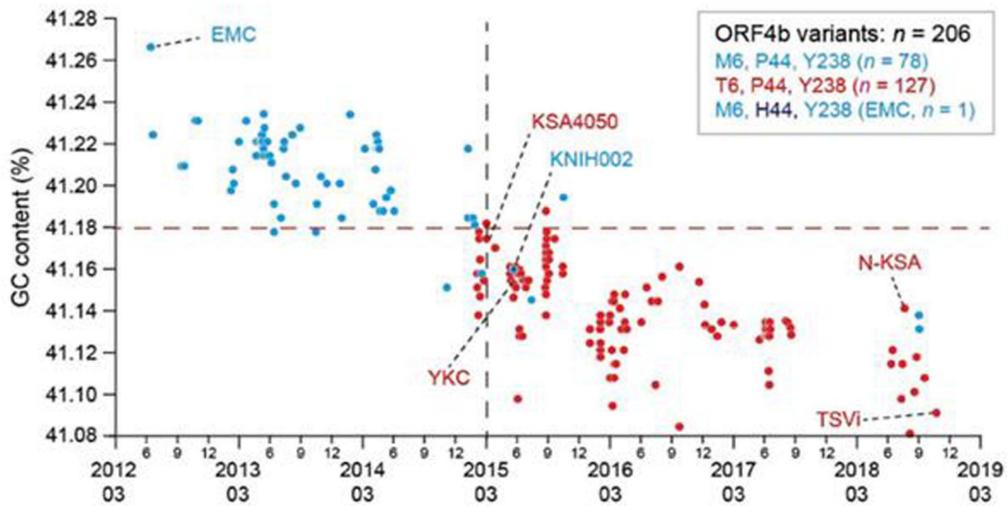
도면1



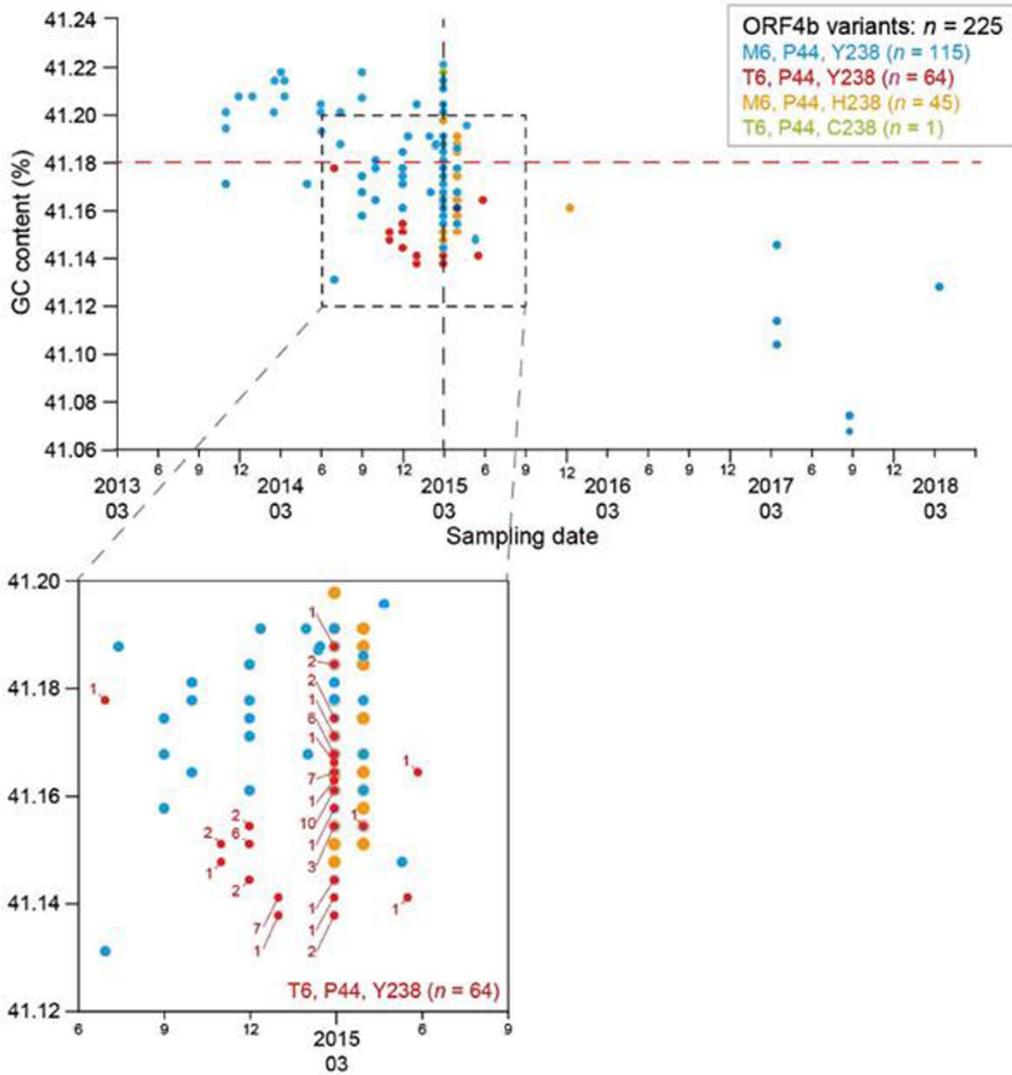
도면2



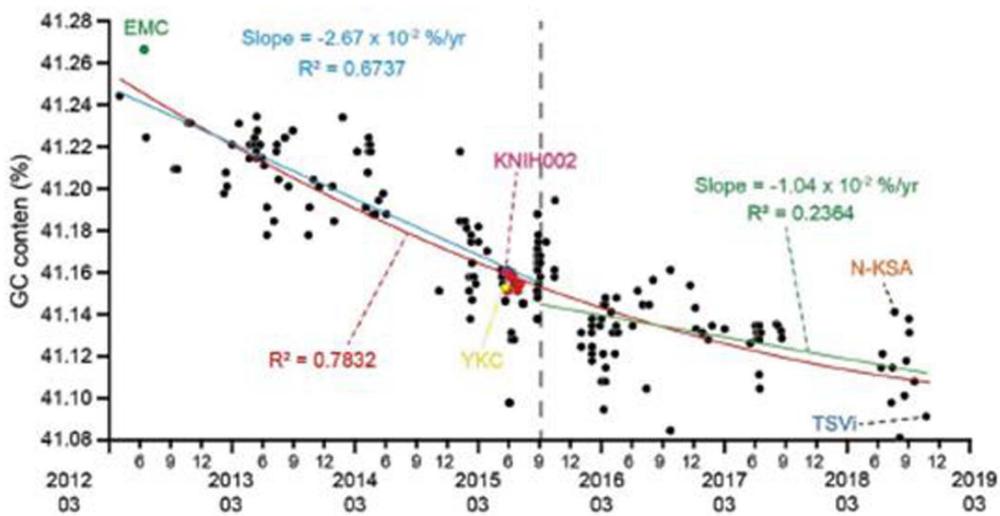
도면3



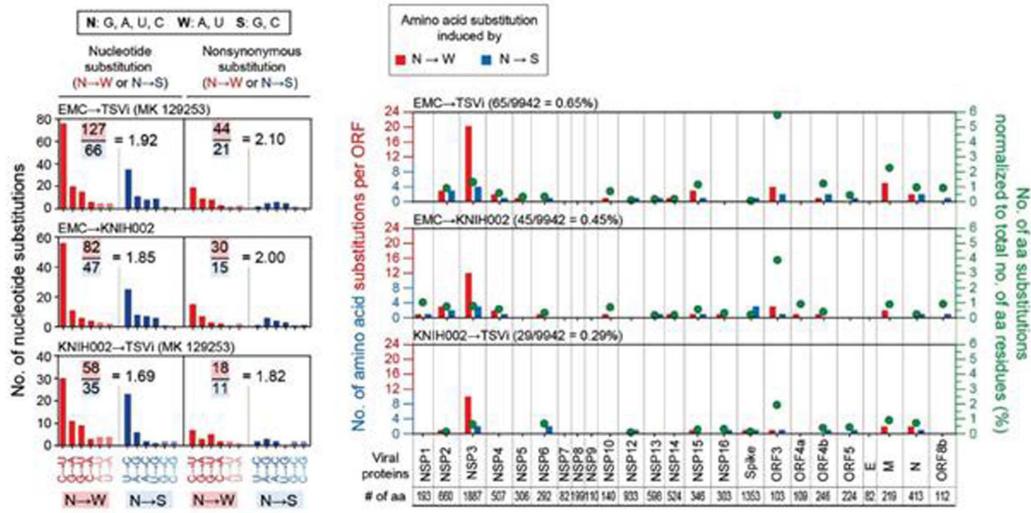
도면4



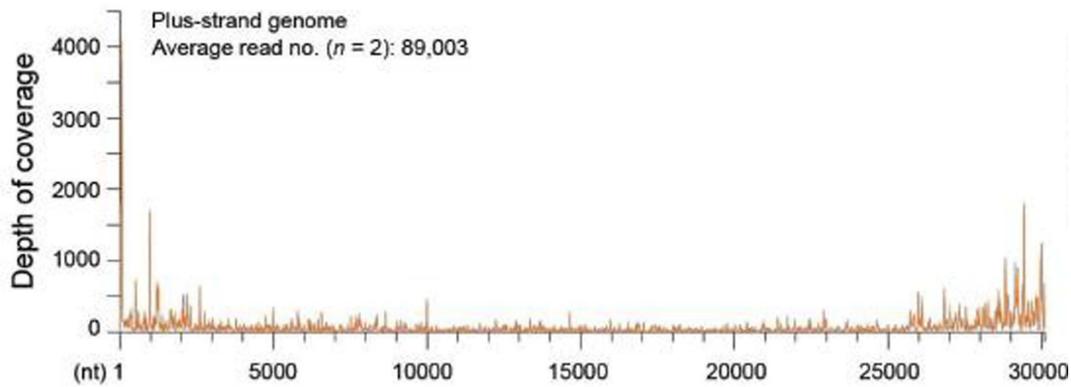
도면5



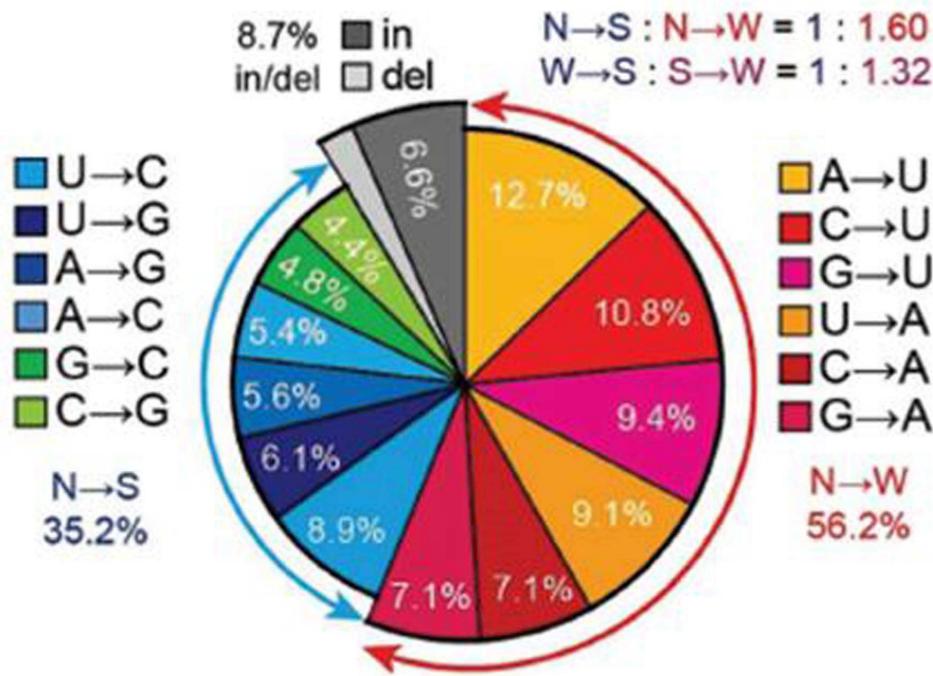
도면6



도면7



도면8



서열 목록

- <110> Industry-Academic Cooperation Foundation, Yonsei University
- <120> Biomarker for Identification of Camel-origin MERS Coronavirus
Species Prone to Spreading in Human Networks and uses thereof
- <130> PD20-133
- <160> 4
- <170> KoPatentIn 3.0
- <210> 1
- <211> 246
- <212> PRT
- <213> Middle East respiratory syndrome coronavirus
- <400> 1

Met Glu Glu Ser Leu Met Asp Val Pro Ser Thr Ser Gly Thr Gln Val
1 5 10 15

Tyr Ser Arg Lys Ala Arg Lys Arg Ser His Ser Pro Thr Lys Lys Leu
20 25 30

Arg Tyr Val Lys Arg Arg Phe Ser Leu Leu Arg Pro Glu Asp Leu Ser
35 40 45

Val Ile Val Gln Pro Thr His Tyr Val Arg Val Thr Phe Ser Asp Pro

Tyr Ser Arg Lys Ala Arg Lys Arg Ser His Ser Pro Thr Lys Lys Leu
 20 25 30
 Arg Tyr Val Lys Arg Arg Phe Ser Leu Leu Arg Pro Glu Asp Leu Ser
 35 40 45
 Val Ile Val Gln Pro Thr His Tyr Val Arg Val Thr Phe Ser Asp Pro
 50 55 60
 Asn Met Trp Tyr Leu Arg Ser Gly His His Leu His Ser Val His Asn
 65 70 75 80
 Trp Leu Lys Pro Tyr Gly Gly Gln Pro Val Ser Glu Tyr His Ile Thr
 85 90 95
 Leu Ala Leu Leu Asn Leu Thr Asp Glu Asp Leu Ala Arg Asp Phe Ser
 100 105 110

 Pro Ile Ala Leu Phe Leu Arg Asn Val Arg Phe Glu Leu His Glu Phe
 115 120 125
 Ala Leu Leu Arg Lys Thr Leu Val Leu Asn Ala Ser Glu Ile Tyr Cys
 130 135 140
 Ala Asn Ile His Arg Phe Lys Pro Val Tyr Arg Val Asn Thr Ala Ile
 145 150 155 160
 Pro Thr Ile Lys Asp Trp Leu Leu Val Gln Gly Phe Ser Leu Tyr His
 165 170 175
 Ser Gly Leu Pro Leu His Met Ser Ile Ser Lys Leu His Ala Leu Asp
 180 185 190
 Asp Val Thr Arg Asn Tyr Ile Ile Thr Met Pro Cys Phe Arg Thr Tyr
 195 200 205
 Pro Gln Gln Met Phe Val Thr Pro Leu Ala Val Asp Val Val Ser Ile
 210 215 220
 Arg Ser Ser Asn Gln Gly Asn Lys Gln Ile Val His Ser Tyr Pro Ile
 225 230 235 240
 Leu His His Pro Gly Phe
 245
 <210> 3
 <211> 741

<212> DNA

<213> Middle East respiratory syndrome coronavirus

<400> 3

```

atggaggaat ccctgatgga tgttcctct acctccggca ctcaagtta ttcgcgcaaa      60
gcgaggaaga ggagtcattc tccaactaag aaactgcgct acgttaagcg tagattttct      120
cttctgcgcc ctgaagacct tagtgttatt gtccaaccaa cacactatgt cagggttaca      180
ttttcagacc ccaacatgtg gtatctacgt tcgggtcadc atttacctc agttcacaat      240
tggcttaaac cttatggcgg ccaacctgtt tctgagtacc atattactct agctttgcta      300
aatctcactg atgaagattt agctagagat tttcaccca ttgcgctctt ttgcaaat      360

gtcagatttg agctacatga gttcgccttg ctgcgcaaaa ctcttgttct taatgcatca      420
gagatctact gtgctaacat acatagattt aagcctgtgt atagagtaa cacggcaatc      480
cctactatta aggattggct tctcgttcag ggattttccc tttaccatag tggcctccct      540
ttacatatgt caatctctaa attgcatgca ctggatgatg ttactcgcaa ttacatcatt      600
acaatgccaat gctttagaac ttatcctcaa caaatgtttg ttactccttt ggccgtagat      660
gttgtctcca tacggtcttc caatcagggt aataaacaaa ttgttcattc ttacccatt      720
ttacatcacc caggatttta a                                     741
    
```

<210> 4

<211> 741

<212> DNA

<213> Middle East respiratory syndrome coronavirus mutant

<400> 4

```

atggaggaat ccctgacgga tgttcctct acctccggca ctcaagtta ttcgcgcaaa      60
gcgaggaaga ggagtcattc tccaactaag aaactgcgct acgttaagcg tagattttct      120
cttctgcgcc ctgaagacct tagtgttatt gtccaaccaa cacactatgt cagggttaca      180
ttttcagacc ccaacatgtg gtatctacgt tcgggtcadc atttacctc agttcacaat      240
tggcttaaac cttatggcgg ccaacctgtt tctgagtacc atattactct agctttgcta      300

aatctcactg atgaagattt agctagagat tttcaccca ttgcgctctt ttgcaaat      360
gtcagatttg agctacatga gttcgccttg ctgcgcaaaa ctcttgttct taatgcatca      420
gagatctact gtgctaacat acatagattt aagcctgtgt atagagtaa cacggcaatc      480
cctactatta aggattggct tctcgttcag ggattttccc tttaccatag tggcctccct      540
ttacatatgt caatctctaa attgcatgca ctggatgatg ttactcgcaa ttacatcatt      600
    
```

acaatgcat gcittagaac ttatctcaa caaatgttg ttactcttt gccctagat	660
gtgtctcca tacgtcttc caatcaggt aataaaciaa ttgtcattc ttaccatt	720
ttacatcacc caggatttta a	741