

# 표적화 차세대염기서열분석법을 이용한 인수공통 바이러스의 유전체 역학과 예찰

이성현<sup>1†</sup> · 백승환<sup>1†</sup> · 라조리아 시바니<sup>1</sup> · 푸스파레니 사라<sup>1</sup> · 김원근<sup>1,2\*</sup>

한림대학교 의과대학 미생물학교실<sup>1</sup>, 한림대학교 의과대학 의과학연구소 미생물학과<sup>2</sup>

## Genomic epidemiology and surveillance of zoonotic viruses using targeted next-generation sequencing

Seonghyeon Lee<sup>1†</sup>, Seung-Hwan Baek<sup>1†</sup>, Shivani Rajoriya<sup>1</sup>, Sara Pusporeni<sup>1</sup>, Won-Keun Kim<sup>1,2\*</sup>

<sup>1</sup>Department of Microbiology, College of Medicine, Hallym University, Chuncheon 24252, Korea

<sup>2</sup>Department of Microbiology, Institute of Medical Science, College of Medicine, Hallym University, Chuncheon 24252, Korea

Received December 1, 2022  
Revised December 16, 2022  
Accepted December 20, 2022

Corresponding author:

Won-Keun Kim

E-mail: wkim1061@hallym.ac.kr

https://orcid.org/0000-0003-3178-8569

<sup>†</sup>These first two authors contributed equally to this work.

Emerging and re-emerging zoonotic viruses become critical public health, economic, societal, and cultural burdens. The Coronavirus disease-19 (COVID-19) pandemic reveals needs for effective preparedness and responsiveness against the emergence of variants and the next virus outbreak. The targeted next-generation sequencing (NGS) significantly contributes to the acquisition of viral genome sequences directly from clinical specimens. Using this advanced NGS technology, the genomic epidemiology and surveillance play a critical role in identifying of infectious source and origin, tracking of transmission chains and virus evolution, and characterizing the virulence and developing of vaccines during the outbreak. In this review, we highlight the platforms and preparation of targeted NGS for the viral genomics. We also demonstrate the application of this strategy to take advantage of the responsiveness and prevention of emerging zoonotic viruses. This article provides broad and deep insights into the preparedness and responsiveness for the next zoonotic virus outbreak.

**Key Words:** Genomic epidemiology, Genomic surveillance, Zoonotic viruses, Targeted next-generation sequencing

### 차세대염기서열분석법 기반 유전체 역학 및 예찰 연구의 필요성

최근 다양한 국내 가축 및 인수공통 바이러스에 의한 질병이 창궐하고 있다. 대표적인 예로 2002년 박쥐에서 사향 고양이를 거쳐 사람에게 전파된 Severe acute respiratory syndrome coronavirus (SARS-CoV) (Zhong 등, 2003), 2012년 박쥐에서 낙타로 전파되어 사람에게 전파된 Middle East respiratory syndrome-related coronavirus (MERS-CoV) (Zaki 등, 2012), 2019년 12월 처음 발견되어 현재까지 팬데믹인 Severe

acute respiratory syndrome coronavirus 2 (SARS-CoV-2)가 있다(Khan 등, 2020). 또한 전 세계적으로 구제역(Bo 등 2019), 조류인플루엔자(WHO, 2022)와 같은 바이러스성 가축 전염병도 빈번하게 발생하고 있다. 2018년에 발생한 아프리카 돼지 열병 역시 야생 멧돼지에서 가축으로 빠르게 확산되며, 축산업과 인류 보건 건강에 큰 위협이 되고 있다(Zhou 등, 2018).

인수공통 감염병 출현 증가의 원인으로는 야생동물 사육의 증가, 동물성 단백질에 대한 인간의 수요증가 및 도시화와 토지이용 변화에 따른 인간과 동물 사이의 서식지 공유 등이 제시되었다. 또한 우리나라와 같은 밀집 사육 형태 축산업의 발달은 감염



병이 가축-가축 간 또는 인간-가축 간에 쉽게 전파되는 환경을 만들었다. 그 외에도 지구 온난화는 병원체를 전파하는 위생 해충의 개체수를 증가시켰다. 해외여행의 증가와 국가 간 무역의 자유화로 인하여 국가 간의 교류가 증가하며 한 지역에 기존에 없던 감염병의 병원체가 새롭게 유입되거나, 해외에서 국내에 보고되지 않은 감염병에 걸린 상태로 귀국하는 사례도 증가하고 있다. 가축전염병은 한번 발생하면 발원지와 인근에 존재하고 있는 가축을 살처분하기 때문에 경제 및 국가적 손실을 크게 발생시킨다(Hobbs 등, 2021).

가축 및 인수공통 바이러스 출현에 대응하기 위해서는 원인 바이러스의 유전체 염기서열 기반의 유전체 역학 및 능동적 역할이 중요한 역할을 하고 있다. 우선 표적 바이러스의 유전체 서열정보를 획득함으로써 바이러스의 종류를 확인하는 조기 진단이 가능하다. 전장 유전체 염기서열을 비교하여 실시간으로 유행하는 바이러스의 유전자형을 파악할 수 있다. 이미 SARS-CoV-2의 변이주 연구를 위하여 환자로부터 바이러스를 분리하여 바이러스 유전체 서열을 확보하는 연구가 활발히 진행되었다(Paraskevis 등, 2020). 국내에서도 가축 및 인수공통 바이러스에 대한 유전체 염기서열 확보를 위한 다양한 연구가 진행되었다. 질병관리본부에서는 SARS-CoV-2에 감염된 국내 환자에서 바이러스의 분리주를 획득하였고(Kim 등, 2020), 농림축산검역본부에서는 차세대염기서열분석법(Next-generation sequencing)을 이용하여 Fowlpox virus의 전장유전체 분석 기술을 개발하였으며, 개발 기술을 이용하여 바이러스를 분리에 성공하였다. 유전체 역학 연구를 통해 바이러스의 계통지리학적인 특징을 분석할 수 있다. 유전자 서열을 기반으로 하는 분석 결과를 토대로 신·변종 바이러스의 발원지를 추적할 수 있고, 변이 양상을 탐색하여 바이러스의 지역별 바이러스 전염에 대한 전파 형태를 예측할 수 있다. 관련된 국내 연구로는 신증후출혈열 환자로부터 표적 차세대염기서열분석법을 이용하여 바이러스 전장 유전체 염기서열 정보를 확보한 연구가 있다(Kim 등, 2020). 원인 바이러스의 전장 유전체 서열을 기반으로, 계통지리학적 분석을 통해 환자가 바이러스에 감염된 장소를 밝혀냈다. 해당 연구를 통해 국내 Hantaan virus의 전장 유전체 서열 확보와 질병 발생 지역을 대상으로 진행한 역학조사 및 표적 채집을 통해 신·변종 바이러스 출현에 대하여 차세대염기서열분석법이 원인 바이러스의 추적 및 감시연구에 효과적으로 적용될 수 있음을 보여주었다(Kim 등, 2020). 최근 전세계를 위협하고 있는 SARS-CoV-2의 전장 유전체를 확보해서 환자로부터 얻어진 바이러스 분리주 사이의 유전학적 변이 양상을 역추적하여 바이러스의 발원지가 중국 우한일 것이라고 예상하는 연

구가 보고되었다(Li 등 2020). 또한 고병원성 influenza virus의 유전체와 발병 지역 내에서 바이러스 유전체의 변이 양상을 분석해서 influenza virus의 지역별 전파 양상을 추적하기도 했다(Yang 등, 2020). 마지막으로, 차세대염기서열분석법을 통해 바이러스 유전체 간의 차이에 따라 병원성을 규명할 수 있고, 더 나아가 염기서열 정보를 바탕으로 신·변종 바이러스 단백질의 3차원 구조를 예측하여, 아미노산 서열에 따라 얻어진 구조 예측 모델을 기반으로 백신 개발에 이용할 수 있다. 실제로 유전체 서열 정보를 기반으로 바이러스의 병원성을 분석하는 연구가 진행되었으며, 표적 바이러스인 SARS-CoV-2의 유전체 서열을 기반으로 바이러스의 병원성 분석 시스템이 개발되기도 하였다(Ayal 등 2020). 또한 최근 각광 받는 머신러닝을 사용하여 SARS-CoV-2의 항원 예측 모델을 개발된 바 있다(Rhiane 등, 2021). 이런 경우에서와 같이 표적 차세대염기서열분석법을 통해 바이러스 유전체 서열을 식별하고 특성화하는 것은 바이러스의 진단 및 대응, 제어 전략 개발에 매우 중요하다.

신·변종 바이러스성 전염병에 빠르게 대응하기 위한 유전체 역학 및 감시를 위해서는 바이러스의 분리 없이 임상 시료로부터 바로 바이러스의 유전체를 얻어 효율적으로 유전체 염기서열 정보를 확보하는 것이다. 전통적인 바이러스 유전체 염기서열 분석법인 Sanger 시퀀싱은 많은 노동력과 오랜 시간이 필요하다는 단점이 있다. 1990년에 시작된 인간 게놈 프로젝트를 통해 기존의 Sanger 시퀀싱으로 전장 유전체의 염기서열을 분석하는데 한계점이 드러났고, 그에 따라 다양한 염기서열 분석법이 개발되었다. 다양한 노력을 통하여 2000년대 중반에 들어서며, 여러 DNA의 염기서열을 동시에 분석하는 차세대염기서열분석법이 개발되었다(McCombie 등, 2019). 차세대염기서열분석법이 발전함에 따라 전장 유전체 염기서열 분석에 대한 관심과 차세대염기서열분석법의 응용이 증가하였다. 현재 보편적으로 상용화 되어 있는 차세대염기서열분석법에는 크게 표적 유전체를 짧은 길이의 라이브러리로 절편화하여 증폭하는 방식과 표적 유전체에 인위적인 절편화 과정 없이 염기서열을 분석하는 플랫폼으로 나눌 수 있다. 우선 절편화된 유전체의 증폭 과정에서 각각의 핵산마다 표지된 형광 신호를 감지하여 염기서열을 분석하는 방식인 영국 Solexa사에서 개발되어 현재는 미국의 Illumina사에서 제공하는 '합성 기반의 시퀀싱(Sequencing by synthesis, SBS) 플랫폼'과 미국의 Life Technologies사에서 개발한 반도체 기반에서 증폭 과정 동안 발생하는 전기적 신호를 분석하는 이온 토렌트(Ion Torrent) 플랫폼이 있다. 그 외에 가장 작은 크기로 표적 유전체의 인위적인 절편화 과정을 생략하고 표적 유전체 그대로 염기서열을 분석하는 영국의 Oxford Nanopore

Technology사에서 제공하는 나노포어(Nanopore) 플랫폼이 대표적이다.

최근까지 발전된 차세대염기서열분석법을 활용하여 다양한 분야에서 유전체 기반 역학 및 예찰 연구가 진행되었다(Barzon 등, 2011). 신·변종 바이러스의 발견에 표적화 차세대염기서열분석법을 이용해서 헨니파-유사 바이러스인 Gamakvirus (Lee 등, 2021)와 Langyavirus (Zhang 등, 2022)가 밝혀졌다. SARS-CoV-2의 자연 숙주를 파악하고 감염의 기원을 규명하기 위해 진행된 연구(Chen 등, 2021)에도 표적화 차세대염기서열분석법이 이용되었다. 전염병의 원인 바이러스의 감염 경로 및 질병의 전파에 대한 추적 연구에도 다양한 표적화 차세대염기서열분석법이 적용되었다. Ebola virus의 전파가 성 매개 감염으로도 일어날 수 있음을 밝힌 연구(Mate 등, 2015; Quick 등, 2016)와 미국에서 발생한 Zika virus에 대한 감염 경로를 밝힌 연구(Grubaugh 등, 2017), 사람에서 신증후출혈열을 일으키는 Hantaan virus의 감염 경로에 대한 유전체 추적 연구(Kim 등, 2020)와 나이지리아에서 발병한 Lassa virus에 대한 유전체 분석을 통해 질병의 감염 경로가 사람 대 사람이 아닌 지역 내 설치류 숙주로부터 유래된 것임을 밝힌 연구(Anderson 등, 2015)에서도 표적화 차세대염기서열분석법을 이용하였다. 가축에서 발생하는 바이러스성 전염병의 유전체 역학에도 표적화 차세대염기서열분석법이 도입되어 연구가 진행되었다. 양돈 농가에서 발생한 Porcine epidemic diarrhea virus와 Porcine kobuvirus에 동시 감염된 포유돈에 대한 추적 조사를 한 연구(Garcia-Hernandez 등, 2021)가 멕시코에서 진행된 바 있으며, 2018년부터 2019년까지 아프리카에서 발생한 African swine fever virus의 전장 유전체 서열을 분석하여 유전형을 밝혀 아프리카 동부에서 유행한 바이러스의 유전체 진화와 전파의 다양성을 밝히기 위해 표적화 차세대염기서열분석법이 이용되었다(Hakizimana 등, 2021). 또한 소에 감염되어 설사증을 일으키지만, 바이러스에 감염되더라도 증상이 나타나지 않아 진단에 어려움을 겪는 것으로 알려진 Bovine astrovirus에 대한 축산 농가에서의 감염 빈도(prevalance)와 임상 증상과 관련된 유전자를 밝히기 위해서 표적화 차세대염기서열분석법이 응용되었다(Zhu 등, 2021).

본 논문에서는 기존에 발생하는 다양한 가축전염병과 발생이 예측되는 새로운 전염병에 대한 유전체 역학 및 예찰 연구를 위해 적용될 수 있는 다양한 표적 차세대염기서열분석법을 소개하고, 최근 급변하는 환경의 변화에 따른 가축 감염 및 인수공통 감염병에 대한 대비(preparedness)와 대응(responsiveness)에 있어 표적 차세대염기서열분석법을 이용한 유전체 감시 및

예찰 연구의 응용과 그 예를 기술할 것이다.

## 유전체 역학과 예찰에 이용되는 차세대염기서열분석법의 플랫폼

### 나노포어 시스템(MinION sequencing)

2005년에 Oxford Nanopore Technology (ONT)에서 개발한 차세대염기서열분석 플랫폼으로 나노포어 시퀀싱이 개발되었다. 나노포어 시퀀싱은 플로우셀(flow cell)의 막(membrane)에 존재하는 채널 단백질에 의해 만들어지는 나노포어에 선행의 단일 가닥 DNA가 통과하며 각각의 핵산에 따라 발생하는 전압과 전류의 패턴을 분석하고, 알고리즘을 통하여 염기서열을 알아내는 방식이다(van Dijk 등, 2018). 시퀀싱 과정에서 추가적인 증폭 과정이 없어서 상대적으로 적은 염기서열분석 결과를 출력하는 것과 채널 단백질을 복수의 핵산이 통과하면서 발생하는 시퀀싱 방식의 특성으로 인하여 증폭을 기반으로 하는 SBS 방식이나 이온 토렌트 방식에 비하여 결과의 오류율이 높은 것이 나노포어 시퀀싱의 한계점이었다(Jain 등, 2017). 그러나 시퀀싱 과정에서 발생하는 오류를 개선하기 위해 이중가닥 DNA를 시퀀싱할 수 있는 어댑터 서열이 개발되었고, 염기서열에 대한 정보를 분석하는 알고리즘을 개선하여, 플랫폼 개발 초기에는 85% 내외에 불과하던 정확도는 97~98%까지 향상되었다(Logsdon 등, 2020; Wang 등, 2021). 나노포어 플랫폼은 한번에 수용할 수 있는 플로우 셀의 수에 따라서 나뉜다. 이 중 1개의 플로우 셀을 사용하는 MinION 장비인 MinION Mk1C은 가로 15 cm, 세로 11.5 cm의 작은 크기를 갖고 있어서 기존 차세대염기서열분석 장비와 달리 휴대성과 편리성이 높다. Mk1C 장비와 같이 높은 휴대성의 장비를 이용한다면, 신·변종바이러스에 의한 새로운 질병이 발생한 현장에서 바이러스의 분리를 위한 세포 배양 없이 병원성 바이러스의 유전체 서열을 확보하고, 질병의 원인 바이러스에 대한 유전체 역학 및 예찰 연구가 가능할 것이다. 현재까지 나노포어 시퀀싱 플랫폼을 이용한 유전체 역학 및 예찰 연구에는 다음과 같은 연구가 보고되었다. 인수공통바이러스 중에서 설치류를 자연 숙주로 하며 인간에서 신증후출혈열을 일으키는 Hantaan virus의 전장 유전체 염기서열(Lee 등, 2022) 분석에 차세대염기서열분석법이 이용되었다. 그리고 신고 전염병에 해당하는 렘피스킨병을 유발하며 기존의 진단을 위해서는 소의 정소에서 분리한 세포 배양을 통한 진단만이 가능했던 Capripoxvirus를 바이러스의 분리 없이 현장에서 확보한 분변 시료로부터 바이러스 유전체를 확보하여 차세대

염기서열분석법을 통해 전장 유전체 서열을 확보하고, 같은 과에 속하는 Poxvirus와의 구별을 위한 연구(Eltom 등, 2021)가 진행되었다. 또한 인도에서 최초로 오리 전염성 장염의 원인 바이러스 전체 유전체의 특성을 규명한 연구(Aasdev 등, 2021)에서 역시 나노포어 플랫폼을 이용했다. 그 외에도 케냐, 탄자니아, 필리핀의 공동 연구를 통해 상용화된 백신이 있지만 여전히 공중 보건에 큰 위협을 가하는 광견병 바이러스의 예찰 연구에서 나노포어 플랫폼을 통해 기존 유전체 예찰 연구보다 더 빠르고 적은 비용으로 광견병 바이러스 유전체 서열을 확보한 연구가 보고되기도 했다(Brunker 등, 2020).

### 합성 기반 시퀀싱 시스템(Illumina sequencing)

물리-화학적 방식을 통해 일정한 길이로 표적 유전체를 절편화한 뒤, 절편화된 DNA 양 말단에 어댑터 서열을 붙여준 라이브러리를 증폭하는 과정에서 발생하는 신호 분자를 감지한 이미징센서를 통해 염기서열을 분석하는 시퀀싱(Sequencing by synthesis) (Mardis 등, 2013) 방식이다. 현재 대표적으로 미국의 Illumina사에서 제공하며, 염기서열 분석에서 최소 1.2 Gb (iSeq)에서 최대 3 Tb (NovaSeq)의 데이터를 얻을 수 있고, 데이터 양에 따라 iSeq과 MiniSeq, MiSeq, HiSeq, NextSeq, NovaSeq 이 있다. 이러한 SBS 방식은 표적 DNA의 증폭을 통한 시퀀싱으로 다른 시퀀싱 플랫폼에 비해 상대적으로 높은 정확도를 가지는 장점이 있다. 그러나 짧은 시퀀싱 라이브러리를 이용하는 점 때문에 유전체의 구조 변이를 탐지하는 데에는 부적합하며, 특히 반복 서열이 있거나 높은 GC 비율을 보이는 동원체, 텔로미어와 같이 다양한 돌연변이가 생기는 구간의 유전체 연구에는 부적합하다(Logsdon 등, 2020). 그 외에도 장비의 가격이 다른 플랫폼에 비하여 고가인 경향이 있고, 휴대형 장비가 개발되지 않아서 현장 적용형 신·변종 바이러스 유전체 역학 및 예찰 연구에 대한 접근 장벽이 높다.

### 이온 토렌트 시스템(Ion Torrent)

이온 토렌트 시스템은 절편화된 표적 DNA를 비드에 부착하고 순서대로 한 종류의 염기만을 이용하여 증폭하며, 해당 염기가 표적 DNA에 상보적으로 결합하며 방출되는 수소이온에 의한 pH 변화를 반도체 칩에서 전기 신호로 얻어 염기서열을 분석하는 방식이다. 이 방식은 형광을 이용하지 않으므로, 관련된 부품이 필요하지 않으며 Illumina 사에서 제공하는 분석기기보다

장비의 크기가 작고 염기서열 분석이 수 시간 내에 끝난다는 장점이 있다. 그러나 다른 플랫폼에서는 한 번의 염기서열분석을 통해 적게는 수 기가바이트에서 테라바이트 단위의 유전체 정보를 얻을 수 있는 것과 달리 이온 토렌트 시스템은 사용하는 반도체에 따라 1.2 Gb (Ion318 chip)와 15 Gb (Ion540 chip), 50 Gb (Ion550 chip)의 유전체 정보를 얻게 된다. 또한 절편화된 짧은 리드에 대한 염기서열을 분석해야 하는 한계점을 보여주고 있다(Quail 등, 2012).

나노포어 플랫폼의 경우, 표적 유전체의 절편화 과정 없이 긴 리드의 시퀀싱 라이브러리를 이용하므로 특정 염기서열의 반복이 나타나는 것과 같은 다양한 유전적 구조를 나타내는 유전체 연구에 적합하다. 그러나 나노포어 플랫폼은 하나의 채널 단백질에 동시에 4개의 염기가 통과하며 발생하는 정보를 분석하므로, 단일 염기에서 유래하는 신호를 분석하는 다른 플랫폼에 비하여 높은 어려움을 가지고, 시퀀싱 동안 추가적인 증폭과정이 생략되어 상대적으로 낮은 민감도를 가진다는 단점이 있다. 표적 유전체를 절편화해서 처리하고 증폭 과정에서 얻어지는 신호 분자를 감지하는 합성 기반 시퀀싱 플랫폼은 전통적으로 이용된 Sanger 시퀀싱 대비 낮은 비용으로 시퀀싱을 가능하게 하였고, 짧은 리드를 최대한 많이 증폭해서 염기서열을 분석하는 플랫폼 특성상 적은 양의 시료로 정확한 유전체 서열을 확보할 수 있다는 장점이 있다. 그러나 짧은 리드는 반복적인 서열이 다수 등장하는 유전체 서열을 밝히는 연구에는 부적합하며, 장비의 비용이 다른 플랫폼보다 고가라는 점에서 현장 적용형 유전체 연구에는 적합하지 않다. 기존 합성 기반 시퀀싱 시스템 플랫폼의 신호 분자 이용을 증폭 과정에서 발생하는 전기 신호로 대체한 이온 토렌트 플랫폼은 다른 플랫폼에 비해 빠른 시퀀싱이 가능하지만, 반복 서열에 대하여 정확도가 낮고 현재 개발된 휴대형 시퀀싱 장비가 없으므로, 합성 기반 시퀀싱 플랫폼과 마찬가지로 현장 적용형 신·변종 바이러스의 유전체 역학을 위한 연구에는 적합하지 않다. 지금까지 소개한 세 종류의 시퀀싱 플랫폼 기술의 장점과 단점을 Table 1을 통해 비교하였다.

실시간으로 유전체의 정보를 분석하고 MinION Mk1C와 같은 휴대형 시퀀싱 장비가 개발된 나노포어 시퀀싱 플랫폼은 미래 대응 신·변종 바이러스 감염병이 발생할 경우 발병 현장에서 바로 원인 바이러스에 대한 시퀀싱이 가능한 현장 적용형 유전체 역학 및 예찰 연구에 적합한 시스템이라고 할 수 있다. 앞으로는 나노포어 시스템을 이용한 시퀀싱을 위한 표적 유전체 강화 방법에 대하여 논하고, 각각의 방법에 따른 선행 연구를 소개하고자 한다.

**Table 1.** Strengths and limitation of various NGS platform

NGS sequencing platform	Strengths	Limitations
Nanopore sequencing system (Oxford Nanopore Technology)	Long read Genome sequencing in real-time High-throughput Portable sequencing instrument Superior detection of tandem repeats and structural variants	High error rate Low sensitivity Large amounts of templates
Sequencing by synthesis system (Illumina sequencing)	Lower costs of whole-genome sequencing High accuracy and sensitivity	Short read Limited throughput Expensive sequencing equipment Short read length Sequencing error due to PCR amplification
Ion torrent system (Thermo Fisher system)	Fast sequencing	Low accuracy for repeat sequences

### 바이러스 유전체 정보 강화(Enrichment) 기술

현장이나 임상시료에서 직접적으로 바이러스 유전체 염기서열 정보를 확보하기 위한 나노포어 플랫폼은 시퀀싱 과정 동안 시료의 증폭 과정이 없어서 시료 전처리 과정에서 진행되는 표적 바이러스의 유전체에 대한 강화(Target enrichment) 과정이 중요하다. 표적 유전체의 강화 방법에 따라 기존에 서열이 알려진 표적 미생물에 대한 유전체를 증폭할 것인지, 알려지지 않은 새로운 미생물에 대한 유전체를 증폭하는지 선택하게 된다. 이에 따라 실제 응용에서는 연구 목적에 맞는 적합한 표적 유전체 강화 전략의 선정이 중요하다.

#### SISPA-based NGS

분변이나 혈액 또는 병변에서 추출한 시료와 같이 정상미생물총(Microbiome)과 바이러스, 숙주의 유전체와 같은 다양한 생물체에 대한 정보가 모두 포함된 높은 유전체의 다양성을 갖는 시료에서 비선택적으로 모든 유전자를 증폭시켜 다양한 유전체 중에서 질병의 원인이 되는 표적 유전체를 찾을 수 있는 차세대염기서열분석방식으로는 시퀀스-독립 단일-프라이머 증폭(SISPA, Sequence-Independent, Single-Primer Amplification) 방식이 있다. SISPA 방식은 표적 유전체 강화를 위한 증폭 과정에서 비특이적인 프라이머를 이용하기 때문에 원하는 표적 유전체에 대한 정보가 없거나, 미지의 감염성 질환의 원인 균을 규명하고자 할 때 이용될 수 있다. 또한 강화 과정에서 증합효소연쇄반응(Polymerase chain reaction, PCR)을 이용하기 때문에 시료의 양이 소량일 때 유리하다(Kozarewa 등, 2015; Leomar 등, 2016; Hess 등, 2020). 선정한 비특이적 프

라이머와 검체에서 확보한 DNA를 주형으로 이용하여 증합효소연쇄반응을 통해 시료에 있는 유전체를 증폭한다. SISPA 방식을 통해 증폭된 유전체를 바탕으로 차세대염기서열분석법을 통해 유전체 연구에 이용된 예로는 인간의 분변 시료로부터 SISPA 방식으로 유전체를 증폭하여 Norovirus GII (Jothikumar 등, 2022)와 설치류로부터 분리된 신증후출혈열의 원인 바이러스인 Hantaan virus의 전장 유전체 염기서열을 밝히는(Song 등, 2017) 바이러스의 유전체에 대한 예찰과 질병 대응에 대한 연구가 진행되었다.

#### Target capture-based NGS

질병과 관련된 특정 유전자들의 집합체 또는 관심 있는 특정 유전자 뿐만 아니라 특정 병원체의 전장 유전체 서열 전체를 강화할 수 있는 표적 시퀀싱(Targeted sequencing)은 시료 내에서 선택된 유전자 서열만을 증폭하여 시퀀싱하므로, 염기서열 분석 결과를 바탕으로 진행되는 생물 정보학적 분석이 다른 방식에 비하여 상대적으로 용이하다. 또한 표적 유전자를 집중적으로 증폭하므로 더 높은 수준의 시퀀싱 정도(Sequencing depth)를 얻을 수 있다. 표적 시퀀싱에는 표적 라이브러리의 풍부화 단계가 필요하다. 라이브러리 풍부화에는 혼성화 방식, 트랜스포존 매개 단편화 방식 등이 있다. 먼저 혼성화 방식은 고상 형태와 액상 형태의 두 가지 방식으로 개발되었다. 고상 혼성화는 유전체의 관심 영역(Region of interest)이나 표적 유전자에 대하여 상보적인 서열로 구성된 고농도의 올리고 핵산이 이용된다. 표적 유전자가 정해져 있으므로 선택 영역에 대한 높은 특이성을 유지하며 유전체의 관심 영역에 대하여 정확한 염기서열 분석을 가능하게 한다(Leomar 등, 2016). 액상 혼성화는

고농도의 미끼 DNA (bait DNA)를 이용한다. 미끼 DNA란 표적 유전자에 상보적인 DNA에 바이오틴을 결합시킨 것으로, 스트렙타비딘 비드를 이용하게 되면 표적 DNA와의 혼성화 단계(Hybridization) 이후, 비특이적인 DNA 조각들을 쉽게 제거할 수 있다. 이런 방법은 전체 엑솜의 풍부화와 같은 다수의 표적 유전체가 있을 때 사용하기 유리하다(Kozarewa 등, 2015; Leomar 등, 2016). 트랜스포존-매개 단편화 방식은 혼성화 방식과 유사한 프로브(Probe)를 사용하지만, 표적 유전자에 상보적인 올리고 핵산을 이용하는 혼성화 기반 방식과는 달리 효소인 트랜스포제이스(Transposase)를 이용한다. 따라서 트랜스포존-매개 단편화 방식은 DNA 절편화와 어댑터 서열의 결합이 동시에 일어난다(Kozarewa 등, 2015). 다양한 표적 풍부화 방식으로 확보한 라이브러리를 이용한 차세대염기서열분석법 기반으로 중국에서는 대장암 환자에서 체세포 돌연변이의 프로파일링을 진행했고(Dong 등, 2019), 튀니지의 자궁경부암 환자 시료로부터 원인 바이러스인 Human papillomavirus의 유전형 분석 연구(Ardhaoui 등, 2021)가 진행된 바 있다.

**Multiplex PCR (Amplicon-based) NGS**

중합효소연쇄반응(Polymerase Chain Reaction, PCR) 역시 차세대염기서열분석법에서 표적 유전자를 강화하는 방식으로 유용하게 이용된다. 특히 동시에 많은 프라이머를 이용해서 진행되는 다중 PCR (multiplex PCR)은 정해진 시료 내에서 PCR 반응 수를 증가시킴에 따라 발생하는 오류를 감소시키고, PCR의 처리량을 증가하므로 표적 유전체 강화에 효과적인 방법으로 사용된다. 다른 표적 유전체 강화 방식에 비해 다중 PCR

을 이용한 표적 바이러스 유전자를 증폭하는 방법의 장점은 빠르고 경제적이고 효율적인 방식으로 원하는 서열 영역만을 특이적으로 증폭할 수 있다는 것이다. 앞서 소개한 두 가지 방식과 달리 다중 PCR을 이용한 표적 강화 방식에는 하나의 표적 유전체에 특이적으로 적용할 수 있어서 전장 유전체 서열의 확보가 용이하다. 따라서 바이러스 돌연변이에 대한 정보를 얻기 쉬우며, 유전적 진화 과정과 역학 연구에 이용되기에 적합할 것이다. 또한 다른 풍부화 과정에 비해 풍부화 과정에서의 소요 시간이 짧고 더 적은 양의 DNA가 필요하다는 것이 장점이다. 그러나 다중 PCR을 위한 프라이머를 제작하는 데에 이미 알려진 전장 유전체 서열이 필요하다는 점과 단일 반응에 여러 개의 프라이머를 사용하는 것은 비특이적 증폭 산물의 확률이 높고 한 번에 여러 프라이머가 반응하게 되므로 프라이머 간의 간섭으로 인해 증폭의 효율이 떨어질 수 있는 단점이 있기도 하다. 다중 PCR을 통한 표적 강화 방식으로 증폭된 유전체를 기반으로 차세대염기서열분석법을 이용한 바이러스 유전체 연구에는 환자 시료로부터 바이러스의 전장 유전체를 확보하고 염기서열을 기반으로 유전형 분석해서 고위험 Hepatitis A virus 감염에 대한 분자역학 연구(Lee 등, 2022)와 호주에서 발견된 Ross River virus의 숙주가 되는 모기에 대한 자연 감염 여부를 통해 지역 내 바이러스의 분포를 추적한 연구(Batovska 등, 2018)가 있다.

다양한 바이러스를 대상으로 하는 유전체 역학 분야에서 이용되는 표적 유전체 강화 기술의 차이를 비교한 Table 2를 통해 표적 바이러스의 특성과 연구의 목적에 따라 적합한 차세대염기서열분석법에 이용할 유전체 강화 기술의 선별을 이해할 수 있다.

비특이적 프라이머를 이용하여 시료 전체의 DNA를 증폭하여

**Table 2.** Comparison of enrichment method for targeted sequencing

Enrichment method	Characteristic	Strength	Limitations
SISPA-based NGS	• Enrichment of entire genome sequences	• Small amount of a sample required	• Hardness to get genomic sequences in a sample containing ultra-low viral titer
Target capture-based NGS	Solid-phase hybridization	• Easy to perform than in-solution hybridization • Large target sequence available	• Large amount of a sample (3~10 µg) required • High specified instrument required • Vulnerability in high GC/AT genetic areas • Vulnerability in high GC/AT genetic areas
	In-solution hybridization	• Long target sequence	• Small amount of a sample (<1~3 µg) required
Amplicon-based NGS	• Short target sequence	• Fast performance • Small amount of a sample required • Easy to perform than other enrichment methods • Comparable acquisition and depth of whole-genome sequences	• Non-specific amplicon • Primer-prime dimers

표적 유전체를 강화하는 SISPA 방식은 적은 양의 DNA 시료에서도 표적 유전체를 강화할 수 있지만, 시료 내에 전염병의 원인 바이러스의 양이 극소량인 경우, 숙주나 정상미생물총의 유전체의 증폭 속도를 따라오지 못할 수 있다는 한계점이 있다. 정해진 유전자를 특이적으로 강화하는 표적 시퀀싱 방식은 다른 방식에 비하여 길이가 긴 유전자를 표적으로 할 수 있어 전체 엑솜과 같은 길이가 긴 관심 영역 유전자 연구에 적합할 수 있다. 하지만 염기서열에서 GC 또는 AT의 비율이 높은 경우 혼성화 확률이 떨어져, 표적 유전체의 특성에 따라 유전체 연구에는 적합하지 않을 수도 있다. 마지막으로 다중 PCR을 기반으로 하는 표적 유전체 강화 방식은 강화하는 유전체의 부분이 짧으며, SISPA 방식과 마찬가지로 PCR 과정으로 표적 유전체를 증폭하므로 적은 양의 시료로도 표적 유전체를 강화할 수 있는 장점이 있다. 또한 다른 강화 방식에 비해 실험 방법이 간단하고, 일정한 길이로 표적 유전체를 대상으로 하는 프라이머를 이용하므로 표적 유전체 전체에 대하여 비슷한 수준의 염기서열 정보를 확보할 수 있다. 하지만, 다중 프라이머를 이용하기 때문에 비특이적인 증폭 산물이 발생할 수 있으며, 프라이머와 프라이머 이합체가 형성될 수 있다.

## 표적 차세대염기서열분석법 기반 유전체 역학 및 예찰 연구의 현황

조류인플루엔자, 구제역, 아프리카돼지열병과 같은 악성 가축 전염병 뿐만 아니라 인수공통전염병은 현재 치료제나 백신이 거의 개발되어있지 않아 축산업뿐만 아니라 인류를 계속해서 위협할 것이다. 병원체는 변화된 환경에 적응하여 새로운 모습으로 나타난다. 이에 효율적으로 대응하기 위해서는 국가적인 가축 방역체계와 함께 가축 감염 병원균의 유전체 정보를 통해 축산 현장에서 차단방역과 능동 예찰 체계가 중요하다.

### 출현 바이러스 감염병의 원인균 진단 및 확인(Identification of infectious source and origin)

바이러스의 생활사 특성상 변이가 빠르고 숙주에 감염되어도 무증상 보균상태인 경우가 대부분이므로, 환자는 유증상 상태인 경우를 제외하고는 바이러스 감염 여부에 대한 정확한 진단 가능성이 매우 낮다. 이러한 한계를 극복하기 위해 비침윤성 바이러스 검출 방법을 사용한 선제적인 진단 기술의 필요성이 증가하였다. 이를 위해서는 바이러스 유전체 염기서열 정보의 확보가 선행되어야 한다. 다양한 또한 신증후출혈열 환자의 소변

에서 다중 PCR 기반-차세대염기서열분석법(Multiplex PCR-based NGS)을 통해 극소량의 바이러스 감염 시료를 이용해서 바이러스에 대한 전체 유전체 염기 서열을 확보하는 것에 성공하였다(Cho 등, 2021). 다중 PCR 기반 차세대염기서열분석법을 통해 Hepatitis A virus의 전장 유전체를 획득하여 조기진단을 위한 VP3 유전자 마커를 제시할 수 있었다(Lee 등, 2022). 이런 결과를 기반으로 실시간 qPCR을 통한 정량화로 A형 간염의 신속한 분자 진단과 유행하는 유전형 확인, 출현 바이러스 감시 모니터링이 가능하게 되었다. SARS-CoV-2 유전자 분석은 차세대염기서열분석법을 포함한 다양한 기술이 개발되었고, 이를 기반으로 핵산기반 분자 진단 기술이 발전하고 있다. 실시간으로 나노포어 시퀀싱을 이용한 SARS-CoV-2의 분자 진단법은 바이러스의 E gene을 표적으로 최대 64개의 시료를 동시 감별 진단이 가능하다는 장점이 있다(François 등, 2021).

동물 감염 및 인수공통 바이러스의 조기 진단 및 다양한 변이주에 대한 감시를 위한 전략은 현재 높은 고비용, 복잡한 연구 단계가 진입장벽이라고 할 수 있다. 따라서 앞으로 개발될 감염성 바이러스를 진단하기 위한 핵산 기반 조기 진단 체계는 단순하고 명확한 연구 단계와 비용을 줄이기 위한 노력이 필요하다. 또한 유전체 역학과 감시 기술을 이용하여 효과적으로 핵산기반 분자진단과 출현 바이러스의 동정을 위해서는 다양한 임상 시료 확보를 위한 국내외 네트워크 구축도 필요하다.

### 출현 바이러스 감염병의 전염 경로 확인(Identification of transmission chain)

현재 각 국가는 팬데믹 코로나 이후로 다양한 바이러스에 대해 막대한 비용과 시간을 투자해서 바이러스의 변이 양상을 모니터링(www.nextstrain.org) 하고 있으며, 이에 따라 유전체 기반 생물정보분석 수요가 증가하고 있다. SARS-CoV-2 감염 환자에게서 직접 채취한 바이러스의 차세대염기서열분석 데이터만 있으면 신속하고 정확하게 글로벌 데이터와 비교·분석한 결과를 확인할 수 있는 기술이 개발되었다. 이 기술은 수 만개의 SARS-CoV-2 유전체 데이터와 비교해 바이러스의 변이와 진화 양상을 분자유전학으로 파악할 수 있으며 또 다른 신종 감염병 유입 시에도 신속한 대응이 가능할 것이다. 'Avian influenza (AI)'로 알려진 H5N1과 같은 3종의 Influenza virus 감염의 발생 패턴이 중국 내 가금류의 지역별 유통 패턴과 일치한다는 사실을 발견한 연구가 보고되었다(Qiqi 등, 2020). 사람이 감염될 경우, 50~60%의 치사율을 보이는 고병원성 Influenza A virus H5N1과 H7N9, H5N6 바이러스의 유전체를 획득 및 분

석해서 중국 내에서의 바이러스 유전체 변이 양상과 지역별 전파 양상을 추적했다. 이를 통해 생물학적 통계 및 네트워크 모형을 이용해서 조류 독감의 원인 바이러스를 보균하고 있는 가금류의 이동에 의해 중국 내 고병원성 Influenza A virus가 전파된다는 것을 알 수 있었다. Hantaan virus에 감염된 신증후출혈열 환자의 감염 예상 지역에서 채집된 설치류에서 확보한 바이러스의 유전체를 분석해서 환자의 감염 지역을 확인하는 연구가 진행된 바 있다(Kim 등, 2020). 이 연구는 쥐에서 확보된 바이러스 유전체를 분석해서 환자의 감염 지역을 추적하고, 바이러스 감염 위험이 높은 지역과 시기를 찾아내서 인수공통 바이러스 감염병의 출현을 예방 및 관리하는 능동적 예방 연구의 중요성을 보여주고 있다. 또한 아프리카에서 높은 치사율을 보이는 Ebola virus는 이전에는 동물-사람, 사람-사람 혹은 바이러스가 존재하고 있는 음식을 섭취함으로써 감염이 된다고 알려져 있었다. 그러나 차세대염기서열분석법을 이용한 유전체 역학 연구를 통해 타액 및 혈액과 같은 다양한 체액에서 바이러스가 배출되며, 성접촉 매개에 의한 새로운 전파 경로를 제시하고, 이를 막기 위한 위생학적 방안을 제시하기도 하였다(Ashma 등, 2018).

### 출현 바이러스의 진화 양상 및 변이주 확인(Identification of evolution pattern)

지금까지 축적된 바이러스 유전체의 데이터베이스는 출현 바이러스에 대응하는 하나의 핵심 플랫폼으로 자리 잡았다. 그러나 바이러스의 유전적 변이 속도는 현재까지 데이터베이스에 유전체 정보가 추가되는 속도보다 빠르다. 변이 바이러스에 대한 모니터링을 강화하기 위하여 발생한 돌연변이의 위험성을 기준으로 주요 변이 바이러스(Variant of Concern)를 분류하고, 변이 바이러스에 대한 감시 강화를 권고하고 있다. 이에 따라 바이러스의 계통군을 분석하기 위해 전장유전체 염기서열 획득에 대한 중요성이 증대되고 있다. GISAID 데이터베이스에서 470만 개의 SARS-CoV-2 유전체를 분석하여 다양한 변이주에 대한 변이 패턴을 분석한 보고가 있다(Monika 등, 2022). 다양한 VOC에 대한 출현과 범지구적으로 바이러스 변이의 발생, 전파에 대해 분석했으며, 특히 VOC에 속하는 변이주에 대한 계통도를 보여 주었고 새로운 변이체 'Omicron'의 출현에 대한 근거를 제시했다. 국내에서도 표적화된 차세대염기서열분석법을 이용해서 COVID-19의 1차 및 2차 대유행 바이러스의 변이주에 대한 분석을 통해 SARS-CoV-2의 진화 양상과 전파 경로를 밝힌 바 있다(Shilpa 등, 2022).

### 출현 바이러스의 병원성 확인(Identification of pathogenicity)

2003년 발생한 SARS-CoV 이후로, 신출현 및 재출현 바이러스 창궐로 인해 가축 감염 및 인수공통 감염 바이러스에 대한 병원성 평가 체계 확립의 필요성과 중요성이 꾸준히 제시되었다. 국립보건연구원에서는 가금류 및 야생 철새 분변에서 분리된 H5N6형 AI 바이러스 유전체를 획득하여 지금까지 중국, 베트남, 라오스 및 홍콩과 같은 다양한 지역에서 분리된 바이러스의 유전체와 비교했을 때 H5N6형 AI 바이러스가 인체감염 및 병원성이 증가했다는 결과를 보고했다(KCDC, 2016; Cho 등, 2021). 약 30년간 채취한 Human immunodeficiency virus (HIV)의 유전자와 세계유전자은행에 등록된 HIV의 전장유전체 서열을 분석한 결과, 시간이 지남에 따라 HIV 병원성이 점점 약화되고 있음을 보고한 연구 결과도 있다(Cho 등, 2021). 또한 SARS-CoV-2의 VOC의 확보한 전장유전체 서열을 토대로 각 변이주의 특성을 분석한 연구를 통하여 Spike 단백질의 변이가 VOC의 감염력을 조절하고, Nucleocapsid 단백질의 변이가 병원성을 조절한다는 것을 밝혔다(Ayal 등, 2020). 우리나라 농림축산검역본부는 Fowlpox virus의 전장유전체 분석 기술을 개발하였으며, Poxvirus 21주의 전체 유전자 염기서열을 분석해서 지역과 변이에 따라 유전형 차이가 있음을 세계 최초로 규명했으며, 서열 변이에 따라 바이러스의 병원성도 조절된다는 것을 밝혔다(Kim 등, 2022).

### 출현 바이러스 감염병의 백신 개발(Development of vaccines)

바이러스의 염기서열 데이터 분석은 신·변종 변이 병원체에 대한 효과적인 백신 후보를 개발에 도움이 된다. 현재 COVID-19에 대한 백신으로 각광받게 된 mRNA 백신의 경우, 차세대염기서열분석 기반으로 획득한 SARS-CoV-2 유전체 데이터를 기반으로 제작되었다. 만약 더 많은 SARS-CoV-2에 대한 유전체 분석이 가능하게 되면, 과학자들은 더 신속히 변이 발생 양상을 알 수 있게 된다. 이를 기반으로 백신 개발 연구들은 더 효과적인 백신을 신속하게 업데이트해 개발을 진행할 수 있다. 전염성을 높일 수 있는 '돌연변이 후보군'을 사전에 예측하고, 그것의 유발 인자를 분석해서 백신 개발에 선제적으로 대비하는 것이다(Moore 등, 2021). 유전체 기반 mRNA 백신은 표적 바이러스의 유전체 염기서열 정보를 안다면 신·변종 병원체가 등장해도 빠르게 백신을 설계하고 생산할 수 있기 때문에 가장 먼



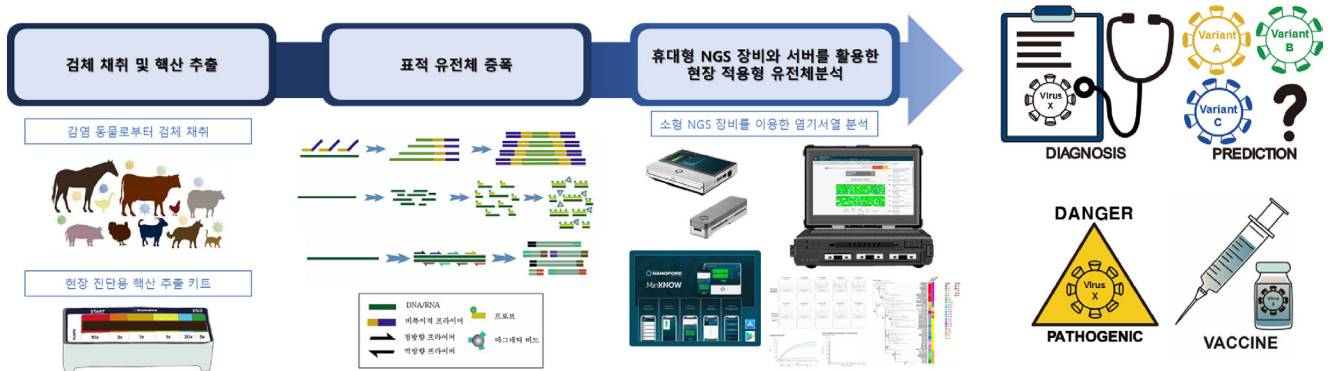
결론

저 차세대염기서열분석법 기반으로 상용화된 mRNA 백신은 SARS-CoV-2 단일 항원뿐만 아니라 Influenza virus까지 예방하는 복합 백신을 위한 개발을 시작했다(Rui 등, 2022).

한편, 차세대염기서열분석 기술의 발전은 대용량 유전체 데이터를 이용한 알고리즘 기반의 항원 후보 예측 기술을 제공하였다. 최근에는 심층 돌연변이 스캐닝(Deep Mutational Scanning, 이하 DMS)이라 불리는 첨단 기술이 개발되었다. 이와 같이 급부상하고 있는 예측 기반 에피토프(Epitope) 선정 기술은 Coronaviridae의 알려진 에피토프 서열을 활용해서 SARS-CoV-2 유전체 염기서열 정보로부터 효과적인 항원 후보물질을 제시하기도 했다. SARS-CoV-2 유전체 전반에 걸친 에피토프 후보를 추려 가장 일반적으로 발견되는 에피토프를 탐색하였고, 단백질 수준에서 에피토프 후보 물질의 특성을 밝힌 것이다(Akshay 등, 2022). 또한 서로 다른 유전자 영역의 돌연변이 빈도에 따라 SARS-CoV-2와 SARS-CoV 사이에 112개의 B 세포와 279개의 T 세포 에피토프가 있는 것이 확인되었다. 바이러스의 주요 항원부위 유전정보는 동적(Dynamic)으로 변화하기 때문에, 이러한 유전자 변이를 반영할 수 있는 유전체 역학과 감시 연구는 빠르게 변이하는 바이러스 감염병에 대응할 수 있는 차세대 백신 개발을 위해 매우 중요하다.

현재까지 보고된 현장 진단형 유전체 연구를 바탕으로 그림 1과 같이 미래 발생이 예상되는 인수공통 감염병의 원인 바이러스 X에 대하여 표적화된 차세대염기서열분석법을 이용한 유전체 감시 및 예찰 연구의 대략적인 흐름을 Figure 1에 정리하였다.

앞으로 자연환경의 변화와 파괴에 대한 영향으로 가축 감염 바이러스와 인수공통 바이러스 감염병은 지속적으로 출현하며 축산업과 인류 보건 건강에 위협이 될 것이다. COVID-19 팬데믹 이후에도 미래 출현 신·변종 바이러스 감염병에 대한 대비와 대응 체계 구축을 위한 지속적인 준비가 필요하다. 다양한 전략의 차세대염기서열분석 기술 발전과 함께 바이러스 유전체 역학 및 예찰 기술은 출현 바이러스에 대한 감염원 또는 감염 지역의 신속한 확인, 감염 및 진화 양상 확인, 병원성 확인을 위한 정보를 제공한다. 또한 이러한 기술은 출현 바이러스에 대한 효과적인 백신 생산뿐만 아니라 치료제 개발에도 이용되고 방역 정책 제시와 결정에 도움을 준다. 그러므로 표적화된 차세대염기서열분석법을 기반으로 미생물 유전체 역학 및 예찰 기술의 연구 개발과 임상 연구가 지속적으로 증대되어야 한다. 이는 미래 신·변종 가축 감염 및 인수공통 바이러스에 대한 신속하고 정확한 대비 및 대응을 가능하게 하고, 가축 산업의 피해를 줄이며 인류 보건 건강의 증대와 사회 안전망을 제공할 것이다.



**Fig. 1.** A schematic diagram of genomic epidemiology and surveillance on emerging virus X using targeted next-generation sequencing. The laboratory investigation should be promptly carried out from extracting nucleic acids to identifying the pathogen at the occurrence of a new or re-emerging infectious disease. To perform the next-generation sequencing (NGS) for the clinical samples, the library preparation is critical to enrich the genome sequences due to the ultra-low copy of viral genomes: When an etiological agent of the disease is anticipated, the multiplex PCR-based NGS is implemented to generate an effective library preparation. The SISPA- or target capture-based NGS should be carried out for the library preparation when the target is unclear. Subsequently, the library sample is directly sequenced in the field by using a portable next-generation sequencer such as MinION Mk1C or Flongle. The genome sequence of the unknown etiological agent is identified in real-time by the bioinformatic pipeline installed in a mobile server. Finally, the genome sequence of the pathogen allows to the genomic epidemiology and surveillance by investigating the infectious source and origin, tracking of transmission chain, and risk assessment. This platform provides preventive and mitigated strategies for emerging zoonotic virus outbreaks in the future.

## ACKNOWLEDGEMENTS AND FUNDING SOURCES

This research was supported by Korea Institute of Marine Science & Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries, Korea (20210466).

## DISCLOSURE

The authors declare no competing financial interests.

## CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

## ORCID

Seonghyeon Lee, <https://orcid.org/0000-0003-4963-6440>  
Seung-Hwan Baek, <https://orcid.org/0000-0002-7861-7049>  
Shivani Rajoriya, <https://orcid.org/0000-0001-6873-8552>  
Sara Puspareni, <https://orcid.org/0000-0003-4989-3973>  
Won-Keun Kim, <https://orcid.org/0000-0003-3178-8569>

## REFERENCES

- Aasdev A, Pawar SD, Mishra A, Dubey CK, Patil SS, Gogoi SM, Bora DP, Barman NN, Raut AA. 2021. First complete genome characterization of duck plaque virus from India. *Virusdisease* 32: 789-796.
- Akshay A, Kristen LB, Sara C, Mark K, Gowri N, Edward S, Gandhar M, Simone B, Vandana M, James HK. 2022. Predicting Epitope Candidates for SARS-CoV-2. *Viruses* 14: 1837.
- Anderson KG, Shapiro BJ, Matranga CB, Sealfon R, Lin AE, Moses LM, Folarin OA, Goba A, Odia I, Ehiane PE, Momoh M, England EM, Winnicki S, Branco LM, Gire SK, Phelan E, Tariyal R, Tewhey R, Omoniwa O, Fullah M, Fonnier R, Fonnier M, Kanneh L, Jalloh S, Gbakie M, Saffa S, Karbo k, Gladden AD, Qu J, Stremlau M, Nekoui M, Finucane HK, Tabrizi S, Vitti JJ, Birren B, Fitzgerald M, McCowan C, Ireland A, Berlin AM, Bochicchio J, Tazon-Vega B, Lennon NJ, Ryan EM, Bjornson Z, Milner Jr DA, Lukens AK, Broodie N, Rowland M, Heinrich M, Akdag M, Schieffelin JS, Levy D, Akpan H, Bausch DG, Rubins K, McCormick JB, Lander ES, Gunther S, Hensley L, Okogbenin S, VHFC, Schaffner SF, Okokhere PO, Khan SH, Grant DS, Akpede GO, Asogun DA, Gnirke A, Levin JZ, Happi CT, Garry RF, Sabeti PC. 2015. Clinical Sequencing Unravels Origins and Evolution of Lassa Virus. *Cell* 162: 738-750.
- Ardhaoui M, Ennaifer E, Salim ACDM, Gomez FM, Laasili T, Boubaker MS, Guizani I. 2021. Nested PCR followed by NGS: Validation and application for HPV genotyping of Tunisian cervical samples. *PLoS One* 16: e0255914.
- Ashma S, Yogesh SM, Ashutosh P, Ruchi S. 2018. Ebola virus: An emerging sexually transmissible infection pathogen. *Indian J Sex Transm Dis AIDS* 39: 65-67.
- Ayal BG, Noam A, Guihem F, Yuri IW, Feng Z, Eugene VK. 2020. Genomic determinants of pathogenicity in SARS-CoV-2 and other human coronaviruses. *Proc Natl Acad Sci U S A* 117: 15193-15199.
- Barzon L, Lavezzo E, Militello V, Toppo S, Palu G. 2011. Applications of next-generation sequencing technologies to diagnostic virology. *Int J Mol Sci* 12: 7861-7884.
- Batovska J, Lynch SE, Cogan NOI, Brown K, Darbro JM, Kho EA, Blacket MJ. 2018. Effective mosquito and arbovirus surveillance using metabarcoding. *Mol Ecol Resour* 18: 32-40.
- Bo LL, Lwin KS, Ungvanijban S, Knowles NJ, Wadsworth J, King DP, Abila R, Qiu Y. 2019. Foot-and-mouth disease outbreaks due to an exotic serotype Asia 1 virus in Myanmar in 2017. *Transbound Emerg Dis* 66: 1067-1072.
- Brunker K, Jaswant G, Thumbi SM, Lushasi K, Lugelo A, Czupryna AM, Ade F, Wambura G, Chuchu V, Steenson R, Ngeleja C, Bautista C, Manalo DL, Gomez MRR, Chu MYJV, Miranda ME, Kamat M, Rysa-

- va K, Espineda J, Silo EAV, Aringo AM, Bernales RP, Adonay FF, Tildesley MJ, Marston DA, Jennings DL, Fooks AR, Zhu W, Meredith LW, Hill SC, Poplawski R, Gifford RJ, Singer JB, Maturi M, Mwatondo A, Biek R, Hampson K. 2020. Rapid in-country sequencing of whole virus genomes to inform rabies elimination programmes. *Wellcome Open Res* 5: 1-30.
- Chen X, Kang Y, Luo J, Pang K, Xu X, Wu J, Li X, Jin S. 2021. Next-Generation Sequencing Reveals the Progression of COVID-19. *Front Cell Infect Microbiol* 11: 632490.
- Chen YC, Liu T, Yu CH, Chiang TY, Hwang CC. 2013. Effects of GC bias in next-generation-sequencing data on de novo genome assembly. *PLoS One* 8: e62856.
- Cho S, Kim WK, No JS, Lee SH, Jung J, Yi Y, Park HC, Lee GY, Park K, Kim JA, Kim J, Lee J, Lee D, Song DH, Gu SH, Jeong ST, Song JW. 2021. Urinary genome detection and tracking of Hantaan virus from hemorrhagic fever with renal syndrome patients using multiplex PCR-based next-generation sequencing. *PLoS Negl Trop Dis* 15: e0009707.
- Cho YK, Kim JE, Foley BT. 2021. Sequence Length of HIV-1 Subtype B Increases over Time: Analysis of a Cohort of Patients with Hemophilia over 30 Years. *Viruses* 13: 806.
- Dong Z, Kong L, Wan Z, Zhu F, Zhong M, Lv Y, Zhao P, Shi H. 2019. Somatic mutation profiling and HER2 status in KRAS-positive Chinese colorectal cancer patients. *Sci Rep* 9: 16894.
- Eltom KH, Althoff AC, Hansen S, Bohlken-Fascher S, Yousif A, El-Sheikh HA, ElWakeel A, Elgamal MA, Mossa HM, Aboul-Soud EA, Wolff J, Korthase C, Hoffmann B, Adam NM, Abdelaziz SA, Shalaby MA, Wahed AAE. 2021. Differentiation of Capripox Viruses by Nanopore Sequencing. *Vaccines (Basel)* 9: 351-361.
- François S, Jean-Louis P, Stefan E, Marco AMP. 2021. Real-time SARS-CoV-2 diagnostic and variants tracking over multiple candidates using nanopore DNA sequencing. *Sci Rep* 11: 15869.
- Garcia-Hernandez ME, Trujillo-Ortega ME, Alcaraz-Estrada SL, Lozano-Aguirre-Beltran L, Sandoval-Jaime C, Taboada-Ramirez BI, Sarmiento-Silva RE. 2021. Molecular Detection and Characterization of Porcine Epidemic Diarrhea Virus and Porcine Aichivirus C Coinfection in México. *Viruses* 13: 738-749.
- Grubaugh ND, Ladner JT, Kraemer MUG, Dudas H, Tan AL, Gangavarapu K, Wiley MR, White S, Theze J, Manani DM, Prieto, Reyes D, Bingham Am, Paul LM, Robles-Sikisaka R, Oliveira G, Pronty D, Barcellona CM, Metsky HC, Baniecki ML, Barnes KG, Chak B, Freije CA, Gladden-Young A, Gnirke A, Luo C, McClinnis B, Matranga CB, Park DJ, Qu J, Schaffner SF, Tomkins-Tinch C, West KL, Winnicki SM, Wohl S, Yozwiak NL, Quick J, Fauver JR, Khan K, Brent SE, Reiner Jr RC, Lichtenberger PN, Ricciardi MJ, Bailey VK, Watkins DI, Cone MR, Kopp IV EW, Hogan KN, Cannos AC, Jean R, Monaghan AJ, Garry RF, Loman NJ, Faria NR, Porcelli MC, Vasquez C, Nagle ER, Cummings DAT, Stanek D, Rambaut A, Sanchez-Lockhart M, Sabeti PC, Gillis LD, Michael SF, Bedford T, Pybus OG, Isern S, Palacios G, Anderson KG. 2017. Genomic epidemiology reveals multiple introductions of Zika virus into the United States. *Nature* 546: 401-405.
- Hakizimana JN, Ntirandekura JB, Yona C, Nyabongo L, Kamwendo G, Chulu JLC, Ntakirutimana D, Kamana O, Nauwynck H, Misinzo H. 2021. Complete genome analysis of African swine fever virus responsible for outbreaks in domestic pigs in 2018 in Burundi and 2019 in Malawi. *Trop Anim Health Prod* 53: 438-447.
- Hess JF, Kohl TA, Kotrova M, Ronsch K, Paprotka T, Mohr V, Hutzenlaub T, Bruggemann M, Zengerle R, Niemann S, Paust N. 2020. Library Preparation for Next generation Sequencing: A Review of Automation Strategies. *Biotechnol Adv* 41: 107537-10750.
- Hobbs EC, Colling A, Gurung RB, Allen J. 2021. The potential of diagnostic point-of-care tests (POCTs)

- for infectious and zoonotic animal diseases in developing countries: Technical, regulatory and sociocultural considerations. *Transbound Emerg Dis* 68: 1835-1849.
- Jain M, Tyson JR, Loose M, Ip C, Eccles DA, O'Grady J, Malla S, Leggett RM, Wallerman O, Jansen HJ, Zalunin V, Birney E, Brown BL, Snutch T, Olsen H, MinION Analysis and Reference Consortium. 2017. MinION Analysis and Reference Consortium: Phase 2 data release and analysis of R9.0 chemistry. *F1000Res* 6: 760-778.
- Jothikumar N, Cromeans T, Shivajothi J, Vinjé J, Murphy J. 2022. Development and evaluation of a ligation-free sequence-independent, single-primer amplification (LF-SISPA) assay for whole genome characterization of viruses. *J Virol Methods* 299: 114346.
- Khan S, Nabi G, Han G, Siddique R, Lian S, Shi H, Bashir N, Ali A, Shereen MA. 2020. Novel coronavirus: how the things are in Wuhan. *Clin Microbiol Infect* 26: 399-400.
- Kim HR, Jang I, Song HS, Kim SH, Kim HS, Kwon YK. 2022. Genetic Diversity of Fowlpox Virus and Putative Genes Involved in Its Pathogenicity. *Microbiol Spectr* 10: e0141522.
- Kim JM, Chung YS, Jo HJ, Lee NJ, Kim MS, Woo SH, Park S, Kim JW, Kim HM, Han MG. 2020. Identification of Coronavirus Isolated from a Patient in Korea with COVID-19. *Osong Public Health Res Perspect* 11: 3-7.
- Kim WK, No JS, Lee D, Jung J, Park H, Yi Y, Kim JA, Lee SH, Kim Y, Park S, Cho S, Lee GY, Song DH, Gu SH, Park K, Kim HC, Wiley MR, Chain PSG, Jeong ST, Klein TA, Palacios G, Song JW. 2020. Active Targeted Surveillance to Identify Sites of Emergence of Hantavirus. *Clin Infect Dis* 70: 464-473.
- Kozarewa I, Armisen, J, Gardner, A.F, Slatko B.E, Hendrickson, C.L. 2015. Overview of target enrichment strategies. *Curr Protoc Mol Biol* 112:721-723.
- Lee GY, Kim WK, Cho SC, Park KM, Kim JW, Lee SH, Lee JY, Lee YS, Kim JH, Byun KW, Song JW. 2022. Genotyping and Molecular Diagnosis of Hepatitis A Virus in Human Clinical Samples Using Multiplex PCR-Based Next-Generation Sequencing. *Microorganisms* 10: 100.
- Lee JY, Park KM, Kim JW, Lee SH, Lee GY, Cho SC, Kim HC, Klein TA, Kim JA, Choi JW, Park JW, Song DH, Gu SH, Yun HS, Kim JE, Lee DS, Hur GH, Jeong ST, Hwang IU, Kim WK, Song JW. 2022. Whole-genome sequencing and genetic diversity of severe fever with thrombocytopenia syndrome virus using multiplex PCR-based nanopore sequencing, Republic of Korea. *PLoS Negl Trop Dis* 16: e0010763.
- Lee SH, Kim KJ, Kim JW, No JS, Park KM, Budhathoki S, Lee SH, Lee JY, Cho SH, Cho SC, Lee GY, Hwang JS, Kim HC, Klein TA, Uhm CS, Kim WK, Song JW. 2021. Discovery and Genetic Characterization of Novel Paramyxoviruses Related to the Genus Henipavirus in *Crocidura* Species in the Republic of Korea. *Viruses* 13: 2020-2035.
- Leomar Y.B, Rajyalakshmi L, Rashmi KS, Rajesh R.S. 2016. Advances in clinical next-generation sequencing: target enrichment and sequencing technologies. *Expert Review of Molecular Diagnostics* 16:3, 357-372.
- Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Liu M, Tu W, Chen C, Jin L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JT, Gao GF, Cowling BJ, Yang B, Leung GM, Feng Z. 2020. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med* 382: 1199-1207.
- Logsdon GA, Vollger MR, Eichler EE. 2020. Long-read human genome sequencing and its applications. *Nat Rev Genet* 21: 597-614.
- Mardis ER. 2013. Next-generation sequencing platforms. *Annu Rev Anal Chem (Palo Alto Calif)* 6: 287-303.
- Mate SE, Kugelman JR, Nyenswah TG, Ladner JT, Wiley MR, Cordier-Lassalle T, Cristie A, Schroth GP, Gross

- SM, Davies-Wayne GJ, Shinde SA, Murugan R, Sieh SB, Badio M, Fakoli L, Taweh F, Wit E, Doremalen N, Munster VJ, Pettitt J, Prieto K, Humrighouse BW, Stroher Ute, DiClaro JW, Hensley LE, Schoepp RJ, Safronetz D, Fair J, Kuhn JH, Blackley DJ, Laney AS, Williams DE, Lo T, Gasasira A, Nichol ST, Formenty P, Kateh FN, Cock KMD, Bolay F, Sanchez-Lockhart M, Palacios G. 2015. Molecular Evidence of Sexual Transmission of Ebola Virus. *N Engl J Med* 373: 2448-2454.
- McCombie WR, McPherson JD, Mardis ER. 2019. Next-Generation Sequencing Technologies. *Cold Spring Harb Perspect Med* 9: a036798.
- Monika KK, Roman J, Pawel K, Marek K. 2022. Genomic Analysis of SARS-CoV-2 Alpha, Beta and Delta Variants of Concern Uncovers Signatures of Neutral and Non-Neutral Evolution. *Viruses* 14: 2375.
- Moore JP, Offit PA. 2021. SARS-CoV-2 Vaccines and the Growing Threat of Viral Variants. *JAMA* 325: 821-822.
- Paraskevis D, Kostaki EG, Magiorkins G, Panayiotakopoulos, Sourvinos G, Tsiodras S. 2020. Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event. *Infect Genet Evol* 79: 104212-104215.
- Qiqi Y, Xiang Z, Philippe L, Marc AS, Yuhai B, Weifeng S, Di L, Wenbao Q, Guogang Z, Nils CS, Oliver GP, Huaiyu Tian. 2020. Assessing the role of live poultry trade in community-structured transmission of avian influenza in China. *Proc Natl Acad Sci U S A* 117: 5949-5954.
- Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13: 341-353.
- Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, Bore JA, Koundouno R, Dudas G, Mikhail A, Ouedraogo N, Afrough B, Bah A, Baum JH, Becker-Ziaja B, Boettcher JP, Cabeza-Cabrerizo M, Caamino-Sanchez A, Carter LL, Doerrbecker J, Enkirch T, Garcia-Dorival I, Hetzelt N, Hinzmann J, Holm T, Kafetzopoulou LE, Koropogui M, Kosgey A, Kuisma E, Logue CH, Mazzarelli A, Meisel S, Mertens M, Michel J, Ngabo D, Nitzsche K, Pallasch E, Patrono LV, Portmann J, Repits JG, Rickett NY, Sachse A, Singethan K, Vitoriano I, Yemanaberhan RL, Zekeng EG, Racine T, Bello A, Sall AA, Faye O, Faye O, Magassouba N, Williams CV, Amburgey V, Winona L, Davis E, Gerlach J, Washington F, Monteil V, Jourdain M, Bererd M, Camara A, Somlare H, Camara A, Gerard M, Bado G, Baillet B, Delaune D, Nebie KY, Diarra A, Savane Y, Pallawo RB, Gutierrez GJ, Mihano N, Roger I, Williams CJ, Yattara F, Lewandowski K, Taylor J, Rachwal P, Turner DJ, Pollakis G, Hiscox JA, Matthews DA, O'Shea AD, Wolfel R, Stoecker K, Fleishmann E, Gabriel M, Weller SA, Koivogui L, Diallo B, Keita S, Rambaut A, Formenty P, Gunther S, Carroll MW. 2016. Real-time, portable genome sequencing for Ebola surveillance. *Nature* 530: 228-232.
- Rui S, Jiawei Z, Ling X, Fengz W, Xiaomin D, Yue W, Zheng W, Dandan Y, Qingrui H, Yong-Gang Yao, Jinghua Y. 2022. A combination vaccine against SARS-CoV-2 and H1N1 influenza based on receptor binding domain trimerized by six-helix bundle fusion core. *EbioMedicine* 85: 104297.
- Sanger F, Coulson AR. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* 94: 441-448.
- Song DH, Kim WK, Hu SH, Lee DS, Kim JA, No JS, Lee SH, Wiley MR, Palacios G, Song JW, Jeong ST. 2017. Sequence-Independent, Single-Primer Amplification Next-Generation Sequencing of Hantaan Virus Cell Culture-Based Isolates. *Am J Trop Med Hyg* 96: 389-394.
- Shilpa C, Kim CM, Lee YM, Seo JW, Kim DY, Yun NR, Kim DM. 2022. Whole-genome analysis and mutation pattern of SARS-CoV-2 during first and second wave outbreak in Gwangju, Republic of Korea. *Sci*

- Rep 12: 11354.
- van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C. 2018. The Third Revolution in Sequencing Technology. *Trends Genet* 34: 666-681.
- Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. 2021. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol* 39: 1348-1365.
- Yang Q, Zhao X, Lemey P, Suchard MA, Bi Y, Shi W, Liu D, Qi W, Zhang G, Stenseth NC, Pybus OG, Tian H. 2020. Assessing the role of live poultry trade in community-structured transmission of avian influenza in China. *Proc Natl Acad Sci U S A* 117: 5949-5954.
- Zaki AM, Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* 367: 1814-1820.
- Zhang XA, Li H, Jiang FC, Zhu F, Zhang YF, Chen JJ, Tan CW, Anderson DE, Fan H, Dong LY, Li C, Zhang PH, Li Y, Ding H, Fang LQ, Wang LF, Liu W. 2022. A Zoonotic Henipavirus in Febrile Patients in China. *N Engl J Med* 387: 470-472.
- Zhong NS, Zheng BJ, Li YM, Poon, Xie ZH, Chan KH, Li PH, Tan SY, Chang Q, Xie JP, Liu XQ, Xu J, Li DX, Yuen KY, Guan PY. 2003. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *Lancet* 362: 1353-1358.
- Zhou X, Li N, Luo Y, Liu Y, Miao F, Chen T, Zhang S, Cao P, Li X, Tian K, Qiu HJ, Hu R. 2018. Emergence of African Swine Fever in China, 2018. *Transbound Emerg Dis* 65: 1482-1484.
- Zhu J, Qi M, Jiang C, Peng Y, Peng Q, Chen Y, Hu C, Chen J, Chen X, Chen H, Guo A. 2021. Prevalence of bovine astroviruses and their genotypes in sampled Chinese calves with and without diarrhoea. *J Gen Virol* 102:001640.