

# 인플루엔자 바이러스의 진화유전학

김유섭<sup>1\*</sup>

요약: 인플루엔자 바이러스는 사람의 심각한 호흡기 전파 감염병을 일으키며 숙주 면역 반응을 회피하는 항원 구조의 매우 빠른 진화를 겪는다. 따라서 감염병 예방과 대처를 위해 인플루엔자 바이러스의 진화를 연구하는 것이 매우 중요하며, 또한 이러한 연구는 다른 생명 시스템에서 접할 수 없는 이점으로 인해 진화생물학의 발전에 큰 기여를 할 수 있다. A형 인플루엔자 바이러스는 야생 조류를 자연 숙주로 오랜 기간 진화해 왔으나 비교적 최근에 집단 크기가 늘어난 사람과 가축을 새 숙주로 삼아 새로운 감염병 유행을 일으켰다. 따라서 이러한 숙주 이동과 적응에 수반되는 진화적 변화를 분석하는 것이 중요한 연구 주제이다. 또한 바이러스의 유전적 다양성과 진화 속도를 결정하는 핵심 요인들(돌연변이, 재조합, 항원소진화, 유효집단크기, 빈도 반비례 선택 등)을 파악하기 위해 기존 다른 생물을 대상으로 개발된 집단유전학적 분석 방법을 다양하게 적용하고 불규칙 시계열로 확보되는 바이러스 염기서열 표본에 맞춘 새로운 분석 방법을 개발하는 것이 중요하다.

키워드: 인플루엔자 바이러스, 진화, 돌연변이, 자연선택, 재조합

서울시 서대문구 이화여대길 52 이화여자대학교 에코과학부

\*Corresponding author: [yuseob@ewha.ac.kr](mailto:yuseob@ewha.ac.kr)

## 서론: 인플루엔자 바이러스의 특성과 전파

인플루엔자(influenza) 바이러스는 보통 유행성 독감이라고 부르는 호흡기 질환의 원인 바이러스이며 한 해에 전 세계적으로 수십 만 명을 사망에 이르게 하고 있다. 단일 가닥 (negative-sense) RNA를 유전 물질로 가지는 Orthomyxoviridae에 속하며 A, B, C, D 네 형(type)으로 분류된다. 여기서 인플루엔자 A형과 B형 바이러스만이 사람의 계절적 인플루엔자 유행(epidemic)을 일으키는 것으로 확인된다 (Krammer, et al. 2018). 바이러스 입자는 막 단백질들을 포함하는 지질막으로 싸여 있으며 8개의 RNA 절편이 nucleoprotein 및 polymerase 단백질과 함께 내부를 구성한다. 막 단백질인 hemagglutinin(HA)과 neuraminidase(NA)는 바이러스가 숙주 세포 내로 침입하고 빠져 나가는 과정에 중요한 역할을 하고 있으며, 바이러스 표면에 노출되었기 때문에 숙주 면역계에 의해 항원으로 인식된다. 따라서 HA와 NA의 단백질 구조는 숙주의 적응면역을 회피해야 하는 선택압에 놓이게 되어 아미노산 서열이 매우 빠르게 변하는 적응 진화를 겪게 된다 (Nelson and Holmes 2007). 이러한 HA와 NA의 빠른 항원 구조 진화를 항원소진화(antigenic drift)라고 부른다. 그 결과 A형 인플루엔자 바이러스(influenza A virus; IAV)의 HA와 NA 염기서열은 각각 크게 차이 나는 여러 계통으로 분화되었으며, 이에 따라 HA와 NA 유전자의 아형(subtype)이 결정된다. 지금까지 HA는 총 18개의 아형(H1, ..., H18)으로, NA는 총 11개의 아형(N1, ..., N11)으로 이루어진 것으로 관찰되었다(<https://www.cdc.gov/flu/about/viruses/types.htm>). 결과적으로 IAV는 두 단백질의

조성에 따라 H1N1, H3N2 등의 아형으로 분류된다.

IAV는 오리를 포함하는 거위목(Anseriiformes)과 물새류를 포함하는 도요목(Charadriiformes) 등 매우 높은 밀도의 개체군을 이루어 살아가는 조류를 자연 숙주(natural reservoir)로 삼아 오랫동안 전파되어 왔던 것으로 추정된다 (Webster, et al. 1992; Olsen, et al. 2006). 자연 숙주는 바이러스에 감염되어도 증상이 심각하지 않은 것으로 알려져 있다 (Hinshaw, et al. 1980; Kuiken 2013). 이를 토대로 자연 숙주와 IAV는 오랜 시간에 걸친 공진화의 결과 서로에 대해 적응한 평형 상태에 도달했다는 가설이 제기되었다 (Webster et al. 1992; Suarez 2000). 한편, 현대 문명에 의해 사람과 가축(닭, 돼지 등)의 밀도가 급격히 증가하며 IAV가 감염을 지속하기에 적합한 숙주가 되었다. 특히 사람은 교통 수단의 발달에 따른 활발한 지역/대륙간 이동을 함으로써 이 바이러스가 전 세계에 걸친 전파를 지속할 수 있는 조건을 제공하였다. 특히 H3N2 아형의 전 세계적 전파를 이어가는 데 있어 아열대 동아시아-동남아시아 지역이 주요한 역할을 하는 것으로 분석되었다 (Russell, et al. 2008).

조류에서 포유류로 숙주를 옮겨 가는 데에는 바이러스의 특정 유전자 염기서열의 치환, 즉 간단하지 않은 진화적 적응이 필요한 것으로 밝혀지고 있다 (Long, et al. 2019). 그럼에도 불구하고 포유류 여러 중, 특히 사람으로의 숙주 이동(host switching)은 반복적으로 일어났다. 예를 들어, 1918년 조류에서 사람으로 H1N1 아형 바이러스가 건너와 대유행을 일으켰고 그 후 사람 간 전파가 지속되는 계절성 감염병으로 자리잡았다. 1968년에는 기존에 사람을 숙주로 삼던 바이러스(H2N2)와 조류에서 기인한 바이러스 사이의 재조합이 일어나며 H3N2 아형의 시초가 되었고, 이후 매년 계절적 유행을 일으키는 바이러스가 되었다. 2009년에는 기존에 유행했던 H1N1 바이러스가 돼지와 조류에서 유래한 바이러스들과 재조합을 겪어 새로운 아형(pH1N1)이 만들어지면서 대유행을 일으켰고 그 후 아직까지 이 아형의 계절적 유행이 이어지고 있다.

## 본론: 인플루엔자 바이러스 진화의 분석

### 유전적 다양성의 생성: 돌연변이, 염기서열 치환, 재조합

인플루엔자 바이러스는 다른 RNA 바이러스와 마찬가지로 매우 빈번한 돌연변이(mutation), 즉 염기서열 복제 과정에서의 오류를 겪는다 (Elena and Sanjuan 2005; Duffy, et al. 2008). 이는 바이러스 RNA 복제를 담당하는 중합효소 복합체(polymerase complex)가 바이러스 자체 유전자(PA, PB1, PB2)에서 만들어지며 여기에는 복제 오류를 교정(proof-read)하는 기능이 결여되어 있기 때문이다 (Steinhauer, et al. 1992). 돌연변이는 바이러스 집단 내의 유전적 변이를 일차적으로 생성하며 이는 바이러스 진화의 원동력이다. 돌연변이는 어떤 숙주 내 환경에서 어떤 염기서열 위치에 발생하느냐에 따라 바이러스의 성공적인 복제와 전파능력, 즉 바이러스 적응도(fitness)를 감소시킬 수도, 증가시킬 수도 있고 아무런 영향을 끼치지 않을 수도 있다. 이들을 각각 해로운(deleterious), 이로운(beneficial), 중립적인(neutral) 돌연변이라고 할 때, 이로운 변이와 중립적인 변이만이 바이러스 집단 내에서 높은 빈도로 증가하고 고정(fixation)에 이르러 시간에 따른 바이러스의 진화적 변화로 관찰된다. 여기서 중립 돌연변이에 의한 변화의 속도는 바이러스 집단의 크기, 자연선택 유무, 전파 패턴 등에 상관없이 돌연변이율(mutation rate), 즉 복제 과정의 오류 빈도에 의해서만 결정된다 (Birky and Walsh 1988; Charlesworth and Charlesworth 2010).

지난 50여 년간 수집된 H3N2 바이러스 염기서열의 중립적 변화를 단백질 암호화 서열의

동의(synonymous) 치환으로부터 관찰했을 때 시간에 따라 일정한 속도의 치환이 일어나는, 즉 분자시계(molecular clock)를 따르는 진화가 일어났음을 알 수 있다 (Croze and Kim 2021). 이는 바이러스의 돌연변이가 한 숙주 내에서는 시간에 따라 일정한 확률로 일어남을 알려준다. 최근에는 숙주 분류군에 따라 바이러스의 단위시간당 중립변이 치환율, 즉 돌연변이율이 다르다는 결과가 조류 IAV의 염기서열 분석에 의해 얻어졌다 (Kim, et al. 2022). 특히 닭을 숙주로 삼는 IAV가 야생 또는 가축 오리를 숙주로 하는 IAV에 비해 빠른 돌연변이를 겪는 것이 밝혀졌다. 이는 복제과정에서 오류가 일어날 확률의 차이보다는 세포 내 복제 환경 차이에 의한 단위시간당 복제 속도의 차이일 가능성이 크다 (Long, et al. 2019). 오리가 IAV의 자연 숙주(natural reservoir)이고 닭은 최근에 가축화에 의한 숫자와 밀도 증가로 바이러스의 새로운 숙주가 되었다는 것을 고려할 때, 새로운 숙주는 바이러스와 공진화를 거치지 못해 바이러스의 빠른 복제를 허용한다는 가설을 제시할 수 있다 (Suarez 2000).

돌연변이와 함께 인플루엔자 바이러스의 유전적 다양성을 결정하는 요인으로 재조합(reassortment) 과정이 있다 (Worobey and Holmes 1999; McDonald, et al. 2016). 한 숙주가 유전적으로 상이한 두 인플루엔자 바이러스에 동시 감염(coinfection)되어 숙주 세포 내에서 두 바이러스가 함께 복제되면, 새로 조립되는 바이러스 입자에 포함되는 각 RNA 절편(segment)이 어느 부모 바이러스에서 복제되었는지는 확률적으로 결정된다. 따라서 유전적으로 새로운 조합의 절편 구성을 가지는 바이러스가 등장하게 된다 (Simonsen, et al. 2007). 이는 유성생식을 하는 생물이 감수분열에 의한 재조합(recombination)을 통해 새로운 유전적 다양성을 획득하는 것과 같은 효과를 가진다. 이러한 바이러스 유전정보의 재조합 현상은 여러 절편으로 RNA 유전체가 나뉘어진 인플루엔자 바이러스에만 일어나는 것이 아니고, 절편으로 나뉘어지지 않은 RNA 유전체를 지니는 바이러스, 예를 들어 코로나바이러스나 HIV에서도 상동 재조합(homologous recombination)에 의해서 일어나고 있다 (Bentley and Evans 2018). 이는 지구상에 존재하는 거의 모든 진핵생물이 유성생식을 하도록 압력을 가하는 재조합의 진화적 이점이 바이러스에도 적용됨을 암시한다 (Chao 1992). 특별히 인플루엔자 바이러스와 같이 여러 절편으로 나뉘어진 유전체를 갖게 되면 바이러스 입자를 조립할 때 오류가 일어나기 매우 쉽기 때문에 과연 얼마나 큰 진화적 압력, 즉 재조합의 이점이 이러한 바이러스 구조를 생성하고 유지하게 만드는지에 대해 심화된 이론적 분석이 요구된다.

바이러스 재조합은 IAV가 다른 동물로부터 사람에게 유입되어 대유행(pandemic)을 개시할 때 매우 중요한 역할을 한다 (Nelson and Holmes 2007). 기존 사람 사이에 전파되던 IAV 아형이 재조합 과정을 통해 다른 동물 사이에 전파되던 바이러스의 항원유전자 절편(HA와 NA)을 획득하면서 H3N2와 pH1N1의 대유행이 시작되었다. 이러한 재조합을 통한 새로운 IAV 아형의 등장을 항원대변이(antigenic shift)라고 부른다. 또한 한 아형 내에서도 바이러스 간의 재조합은 끊임없이 일어나고 있다 (Neverov, et al. 2015). 재조합이 일어난다는 것은 바이러스가 복제를 거치며 전파되어 갈 때 절편들이 전달되는 계보(lineage)가 서로 엇갈린다는 것을 뜻한다. 따라서 어떤 바이러스 표본들을 모아 계통수(evolutionary tree)를 추정하게 되면 절편마다 다른 계통수가 얻어진다. 즉, IAV의 각 절편 염기서열이 겪는 진화 과정은 다른 절편의 진화 과정과 독립적이므로 바이러스 유전체 전체를 아우르는 계통수는 원칙적으로 얻을 수 없다.

### 집단유전학적 구조와 적응 진화

인플루엔자 바이러스 염기서열이 시간과 공간에 따라 어떤 변이를 나타내는지는 바이러스가 전파되는

양상과 자연선택의 빈도와 세기에 달려 있다. 사람의 계절적 독감을 가져오는 H3N2 아형은 그 염기서열 변이의 결정 요인이 가장 잘 분석된 바이러스 집단이다. 먼저, 다른 아형과 마찬가지로 H3N2 바이러스들은 높은 돌연변이율과 매우 큰 숫자로 추정되는 전 세계 바이러스 총 개수(census population size)에 비해 낮은 수준의 염기서열 변이를 나타낸다 (Grenfell, et al. 2004; Nelson and Holmes 2007). 이것을 다르게 표현하면, H3N2의 유효집단크기(effective population size;  $N_e$ )가 매우 작다고 할 수 있다. 유효집단크기는 실제 집단을 가장 단순한 방식에 따라 생식이 일어나는 기준 모델(Wright-Fisher 모델)로 근사했을 때 모델 집단의 크기를 뜻한다 (Charlesworth and Charlesworth 2010). H3N2 바이러스는 한 숙주에서 다른 숙주로 감염이 이어지는데 걸리는 시간이 평균 4.5일 정도라고 알려져 있고 이때 숙주 내 증식하는 바이러스의 숫자에 관계없이 한 줄기의 바이러스 계보(바이러스 한 개체의 전파)가 이어진다는 가정을 할 때, 이 바이러스 집단을 4.5일에 한 세대를 교체하는 기준 모델로 근사하는 경우 HA 절편은  $N_e \approx 500$ , 그 외의 절편은  $N_e \approx 1000$  정도의 값을 얻는다 (Kim and Kim 2016; Croze and Kim 2021). 즉, 단순한 생식 과정을 가정할 때 주어진 세대에 지구 전체에 겨우 1000여개의 바이러스만 계보만이 복제로 이어지는 것과 같은 수준의 변이가 관찰된다. 그 결과 10년 이상에 걸쳐 채집된 바이러스 염기서열의 계통수를 그리면 하나의 긴 줄기에 짧은 여러 가지가 달려있는 (“cactus-like”) 모양을 얻는다 (Buonagurio, et al. 1986; Grenfell, et al. 2004).

이렇게 낮은 값의 유효집단크기를 가지는 이유로 여러 가설을 제시할 수 있다. 첫째, 많은 연구자들은 새로운 항원 구조를 가지는 HA와 NA 변이가 반복적으로 나타나 자연선택에 의해 집단 내에 고정되는 과정, 즉 항원소진화(antigenic drift)가 IAV의 계통수 모양을 결정한다고 가정하였다 (Fitch, et al. 1991; Grenfell, et al. 2004; Bedford, et al. 2011). 어떤 변이가 처음 나타난 특정 염기서열이 낮은 빈도에서 높은 빈도로 빠르게 증가하면서 집단의 염기서열 다양성을 크게 감소시키는 과정을 선택일소(selective sweep)이라고 한다 (Maynard Smith and Haigh 1974; Stephan 2019). 따라서 IAV의 낮은 유전적 다양성은 반복적 선택일소(recurrent selective sweep)의 결과라는 가설이다. 그러나, 직접 바이러스의 가상집단을 만들어 시뮬레이션을 하였을 때 반복적 선택일소만으로는 낮은 유효집단크기와 다른 통계량으로 기술되는 염기서열 변이의 양상을 설명할 수 없다는 것이 밝혀졌다 (Kim and Kim 2016).

항원 변이에 작용하는 양(positive)의 자연선택 외에 유효집단크기 감소에 기여하는 요인으로 바이러스 집단의 인구학적 구조와 음(negative)의 자연선택을 들 수 있다. 먼저, 계절성 독감을 일으키는 H3N2 바이러스는 세계 각 지역에 존재하는 숙주(사람)의 분포와 이동에 의해 시공간적 구조가 결정된다. 열대 지방을 제외한 북반구 및 남반구 각 지역에서는 매년 겨울철에 바이러스가 유행하고 여름이 가까워지면서 바이러스가 소멸한다. 따라서 바이러스의 전파가 지속되기 위해서는 북반구와 남반구 사이를 순차적으로 건너뛰는 이동이 필요하다. 즉, 하나의 지역은 바이러스의 유입과 정착에 이은 소멸 (colonization & extinction) 과정을 반복적으로 겪고, 이러한 소집단이 서로 연결되어있는 메타집단 (metapopulation) 구조가 형성된다. 여기서 소수의 바이러스가 지역 사이를 건너뛰며 전파의 고리를 이어나가게 되면 그 과정에서 유전적 다양성이 감소하는 집단 병목현상(population bottleneck)이 발생한다. 대다수의 소집단이 이러한 병목현상을 반복적으로 겪게 되면 전체 집단의 유전적 다양성, 즉 유효집단크기도 감소하게 된다. 음의 자연선택(negative selection)은 개체의 적응도를 낮추는, 즉 해로운 돌연변이가 집단으로부터 제거되는 과정을 뜻한다. 자연선택에 의해 해로운 돌연변이 하나가 제거될 때 그 돌연변이를 포함하는 염기서열도 같이 제거가

된다. 이 과정이 매우 높은 빈도로 일어나게 되면 결국 집단의 많은 바이러스가 미래에 자신의 복제본을 남기지 못하는, 즉 유효집단크기가 감소하는 일이 발생한다. 이러한 현상을 배경선택(background selection)이라고 한다 (Charlesworth 1994). 바이러스의 복제와 전파를 담당하는 여러 단백질은 지난 오랜 세월에 걸친 바이러스 진화 과정에서 결정된 특정 구조를 가진다. 따라서 매우 높은 빈도로 일어나는 무작위 돌연변이의 상당 부분은 이러한 단백질을 변형시켜 최적화된 구조를 와해하는 해로운 돌연변이가 되고 배경선택을 일으켜 유전변이를 감소시킨다. Kim and Kim (2016)은 H3N2 바이러스의 유효집단크기를 낮추는 요인으로 메타집단 구조와 배경선택이 항원 변이의 자연선택보다 더 크게 작용하고 있음을 밝혔다.

IAV 염기서열이 평균적으로 서로 얼마나 차이 나는가 외에도 염기서열 변이의 다양한 양상이 정량적으로 측정될 수 있다. 소집단 간의 유전적 분화를 측정하는 Wright의  $F_{ST}$ , 유전자 계보의 모양을 간접적으로 알려주는 변이의 빈도 스펙트럼(site frequency spectrum), 두 위치의 변이가 서로 얼마나 독립적인 진화과정을 겪었는지 알려주는 연관불평형(linkage disequilibrium) 등 기존 다른 생물들의 집단유전학적 분석을 위해 개발되었던 통계량 (Charlesworth 2009)이 바이러스 염기서열 변이로부터 계산될 수 있다. 이러한 통계량을 결정하는 요인은 역시 바이러스가 어떤 집단 크기와 공간적 분포의 변동을 겪으며 전파되어 왔는가와 얼마나 강하고 빈번한 자연선택을 겪었는가이다. 따라서 어떤 진화 모델이 실제 관찰값에 가까운 통계량을 산출하는지 분석하여 바이러스의 진화 역사를 추정할 수 있다. 그러나 이러한 방법론을 적용할 때 유의할 점은 기존 여러 생물에 적용되었던 통계량은 염기서열이 동 시간대에 수집된다는 것을 전제로 개발되었다는 것이다. 인플루엔자 바이러스는 매우 불규칙한 시간 간격을 두고 염기서열 결정을 위한 검체가 얻어진다. 따라서 집단 내 변이의 단순한 정량화, 예를 들어 어떤 대립유전자의 빈도 계산을 어떤 시간 간격 내에 포함된 표본들로 수행해야 할 지가 명확하지 않다. IAV와 같이 유효집단 크기가 작아 염기서열들의 공통 조상이 겨우 몇 년 내로 거슬러 올라가면 찾아지는데 반해 바이러스 표본이 얻어지는 시기가 여러 해에 걸쳐있는 경우 표본 시간 차이가 주어진 통계량을 과대 또는 과소 평가하는 오류를 가져올 수 있다. 최근 Croze and Kim (2021)은 이러한 표본 시간 차이를 보정하여  $F_{ST}$ 와 빈도 스펙트럼을 계산하는 방법을 제안하였다. 이 방법에 의해 H3N2 바이러스의 진화 모델을 추정하였을 때 항원 변이에 작용하는 자연선택 과정 없이 메타집단의 진화 만으로는 실제 HA 유전자의 염기서열 변이 양상을 설명할 수 없다는 결론을 얻었다.

## 향후 인플루엔자 진화 연구의 주요 과제

인플루엔자 바이러스에 관한 많은 연구에도 불구하고 아직 이 바이러스의 진화 과정에 대한 많은 부분이 밝혀지지 않고 있다. 예를 들어, 숙주 면역계와 상호작용으로 인플루엔자 바이러스의 항원 구조가 바뀌는 적응적 자연선택이 일어난다는 것은 오랜 연구로 입증되었으나, 얼마나 자주 항원 변이가 자연선택에 의해 치환되는지는 불확실하다. Koel, et al. (2013)은 1968년부터 2003년까지 H3N2 바이러스의 항원 구조가 유의하게 바뀌는 항원 클러스터 (antigenic cluster) 전환이 겨우 10번에 걸쳐 이루어졌으며 그 중 대부분은 하나의 아미노산이 치환되며 일어난 것으로 추정하였다. 이는 H3N2의 헤마글루티닌 유전자에서 매년 관찰할 수 있는 아미노산 치환 중 일부만이 바이러스의 면역 회피에 기여하여 적응 진화를 겪었음을 의미한다. 반면, Bhatt, et al. (2011)은 McDonald-Kreitman 테스트에서 유래한 적응진화의 속도를 추정하는 방법 (McDonald and Kreitman 1991; Smith and Eyre-Walker 2002)을 바이러스 데이터에 적용하여 매년 약 1.3개의 아미노산 치환이 자연선택에

의해 일어나고 있다고 추정하였다. 이러한 연구 결과의 불일치는 좀 더 다양한 집단유전학적 통계량에 대한 진화 모델을 개발하고 그 예측과 비교를 수행하여야 해소될 수 있을 것이다.

또한, 지금까지 양의 자연선택(positive selection) 모델은 새 항원 변이를 지니는 바이러스가 다른 바이러스에 비해 얼마나 더 빨리 복제되는지, 즉 선택계수(selection coefficient)가 시간에 따라 일정하다는 가정을 암묵적으로 하였다. 그러나 새 항원 변이의 빈도가 증가함에 따라 이에 대한 면역력을 획득하는 숙주의 숫자도 함께 증가하면서 이에 작용하는 자연선택의 세기는 점점 감소할 것이라 예상된다. 이러한 빈도-반비례 자연선택은 다른 생물의 경우 여러 적응적 변이가 함께 집단 내에 존재하도록 하는 평형선택(balancing selection)의 한 요인이 된다. 사람의 계절적 감염을 가져오는 인플루엔자 바이러스의 경우에 A/pH1N1, A/H3N2, B/Victoria, B/Yamagata 네 아형이 공존하는 평형상태인 것이 관찰된다. 또한 여러 조류와 포유류 숙주를 망라하여 매우 많은 숫자의 인플루엔자 아형이 존재한다는 것은 바이러스에 작용하는 자연선택의 방향이 새 항원 변이 획득을 통한 상호 배타적 경쟁 하나만 존재하는 것이 아님을 시사한다. 이렇게 숙주 면역계와 바이러스 항원 변이 사이의 복잡한 상호작용에 의해 여러 아형이 분기되어 생성될 수 있음은 Ferguson, et al. (2003)의 수리 모델과 컴퓨터 시뮬레이션으로 제시되었다. 이러한 모델에 기초하여 빈도-반비례 선택이 정확히 어떻게 일어나며 실제 데이터에서 관찰할 수 있는 염기서열 변이의 양상과는 어떤 관련을 맺는지 많은 연구가 필요한 단계이다.

### 결론: 진화유전학적 모델 시스템으로서의 인플루엔자 바이러스

인플루엔자 바이러스는 2020년 코로나바이러스 (SARS-CoV-2) 판데믹이 일어나기 전까지 가장 많은 환자와 사망자를 낳은 호흡기 감염병의 원인이었기에 지난 몇 십 년간 많은 연구가 집중이 되고 데이터가 얻어졌다. 이 바이러스의 진화 양상을 분석하는 것은 감염병 대책 전략을 발전시켜 인류의 복지에 기여하는 일이기도 하지만 진화생물학 발전에도 큰 기여를 할 수 있다. 여러 다양한 생물 또는 생물시스템이 진화생물학의 분석 대상이 되었지만 인플루엔자 바이러스는 다른 시스템이 갖지 못하는 연구 대상으로서의 장점을 지니고 있다. 첫째, 매우 빠르고 지속적으로 항원소변이, 즉 자연선택 과정이 일어나고 있기 때문에 자연선택과 관련된 여러 진화적 현상(선택일소, 중립변이에 미치는 효과, clonal interference, 복잡한 인구구조와 자연선택의 상호작용, 아미노산의 화학적/구조적 성질에 따른 진화, 재조합의 기원 등)의 기존 모델을 검증하고 새로운 가설을 세우는데 매우 유용한 연구 대상이다. 특히, 지난 몇 십 년간 꾸준히 바이러스의 유전체 데이터가 수집되어서 적응적 진화가 일어나는 과정을 거의 실시간으로 관찰할 수 있는 효과가 있다. 이는 긴 시간에 걸쳐 진화가 일어나 현재 관찰되는 염기서열 변이를 가지고 과거의 사건을 추정해야 하는 여러 다른 생물에서는 기대할 수 없는 이점이다. 둘째, 인플루엔자 바이러스는 다른 바이러스와 달리 여러 절편으로 나누어진 유전체를 가지며, 재조합 현상이 절편 내에서는 일어나지 않고 절편 사이에서만 일어나기 때문에 집단유전학적 분석이 쉬워진다. 절편 내에서 재조합이 일어나지 않기 때문에 하나의 절편에 대한 유전적 계보(genealogy)를 계통수 작성 도구로 얻을 수 있고, 따라서 바이러스의 진화 양상을 시각적으로 해석하기 용이하다. 반면 절편 사이의 재조합이 일어남에 따라서 서로 다른 역할을 하는 바이러스 유전자들에 가해지는 특정한 세기와 방향의 자연선택이 절편에 따른 유전적 계보의 차이가 나타날 수 있게 된다. 예를 들어, 항원소진화에 의한 반복적 선택일소를 겪는 HA와 NA 유전자 절편은 다른 유전자 절편에 비해 낮은 염기서열 변이를 가질 것이라는 이론적 예측이

실제로 관찰되었다 (Rambaut et al. 2008; Croze and Kim 2021). 이에 대한 더 정확한 수리적 모델을 세우는 것은 재조합이 무작위적 위치에서 일어나는 대부분의 생물 시스템에서보다 용이할 것으로 예상된다. 이를 통해 정확한 자연선택의 속도와 세기가 추정될 수 있을 것이다. 인플루엔자 바이러스, 더 나아가 여러 감염병 관련 바이러스의 진화 분석에 더 많은 집단유전학 전문가의 참여가 요구되고 있다.

### 감사의글

본 연구는 한국연구재단 중견연구자 지원사업(과제 번호 2020R1A2C1009261)의 지원으로 수행되었다.

### 참고문헌

- Bedford T, Cobey S, Pascual M. 2011. Strength and tempo of selection revealed in viral gene genealogies. *BMC Evol Biol* 11:220.
- Bentley K, Evans DJ. 2018. Mechanisms and consequences of positive-strand RNA virus recombination. *Journal of General Virology* 99:1345-1356.
- Bhatt S, Holmes EC, Pybus OG. 2011. The genomic rate of molecular adaptation of the human influenza A virus. *Molecular biology and evolution* 28:2443-2451.
- Birky CW, Walsh JB. 1988. Effects of linkage on rates of molecular evolution. *Proceedings of the National Academy of Sciences* 85:6414-6418.
- Buonagurio DA, Nakada S, Parvin JD, Krystal M, Palese P, Fitch WM. 1986. Evolution of human influenza A viruses over 50 years: rapid, uniform rate of change in NS gene. *Science* 232:980-982.
- Chao L. 1992. Evolution of sex in RNA viruses. *Trends in ecology & evolution* 7:147-151.
- Charlesworth B. 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genetics Research* 63:213-227.
- Charlesworth B. 2009. Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* 10:195-205.
- Charlesworth B, Charlesworth D. 2010. *Elements of evolutionary genetics*: Roberts and Company.
- Croze M, Kim Y. 2021. Inference of population genetic parameters from an irregular time series of seasonal influenza virus sequences. *Genetics* 217.
- Duffy S, Shackelton LA, Holmes EC. 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet* 9:267-276.
- Elena SF, Sanjuan R. 2005. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. *J Virol* 79:11555-11558.
- Ferguson NM, Galvani AP, Bush RM. 2003. Ecological and immunological determinants of influenza evolution. *Nature* 422:428-433.
- Fitch WM, Leiter J, Li X, Palese P. 1991. Positive Darwinian evolution in human influenza A viruses. *Proceedings of the National Academy of Sciences* 88:4270-4274.
- Grenfell BT, Pybus OG, Gog JR, Wood JL, Daly JM, Mumford JA, Holmes EC. 2004. Unifying the

- epidemiological and evolutionary dynamics of pathogens. *Science* 303:327–332.
- Hinshaw V, Webster R, Bean W, Sriram G. 1980. The ecology of influenza viruses in ducks and analysis of influenza viruses with monoclonal antibodies. *Comparative immunology, microbiology and infectious diseases* 3:155–164.
- Kim G, Shin HM, Kim H-R, Kim Y. 2022. Effects of host and pathogenicity on mutation rates in avian influenza A viruses. *Virus evolution* 8:veac013.
- Kim K, Kim Y. 2016. Population genetic processes affecting the mode of selective sweeps and effective population size in influenza virus H3N2. *BMC evolutionary biology* 16:1–15.
- Koel BF, Burke DF, Bestebroer TM, van der Vliet S, Zondag GCM, Vervaeke G, Skepner E, Lewis NS, Spronken MIJ, Russell CA, et al. 2013. Substitutions Near the Receptor Binding Site Determine Major Antigenic Change During Influenza Virus Evolution. *Science* 342:976–979.
- Krammer F, Smith GJD, Fouchier RAM, Peiris M, Kedzierska K, Doherty PC, Palese P, Shaw ML, Treanor J, Webster RG, et al. 2018. Influenza. *Nature Reviews Disease Primers* 4:1–21.
- Kuiken T. 2013. Is low pathogenic avian influenza virus virulent for wild waterbirds? *Proceedings of the Royal Society B: Biological Sciences* 280:20130990.
- Long JS, Mistry B, Haslam SM, Barclay WS. 2019. Host and viral determinants of influenza A virus species specificity. *Nature Reviews Microbiology* 17:67–81.
- Maynard Smith J, Haigh J. 1974. The hitch-hiking effect of a favourable gene. *Genetics Research* 23:23–35.
- McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* 351:652–654.
- McDonald SM, Nelson MI, Turner PE, Patton JT. 2016. Reassortment in segmented RNA viruses: mechanisms and outcomes. *Nature Reviews Microbiology* 14:448–460.
- Nelson MI, Holmes EC. 2007. The evolution of epidemic influenza. *Nat Rev Genet* 8:196–205.
- Neverov AD, Kryazhimskiy S, Plotkin JB, Bazykin GA. 2015. Coordinated Evolution of Influenza A Surface Proteins. *PLoS Genet* 11:e1005404.
- Olsen B, Munster VJ, Wallensten A, Waldenstrom J, Osterhaus AD, Fouchier RA. 2006. Global patterns of influenza A virus in wild birds. *Science* 312:384–388.
- Rambaut, A., et al. (2008). "The genomic and epidemiological dynamics of human influenza A virus." *Nature* 453: 615–619.
- Russell CA, Jones TC, Barr IG, Cox NJ, Garten RJ, Gregory V, Gust ID, Hampson AW, Hay AJ, Hurt AC. 2008. The global circulation of seasonal influenza A (H3N2) viruses. *Science* 320:340–346.
- Simonsen L, Viboud C, Grenfell BT, Dushoff J, Jennings L, Smit M, Macken C, Hata M, Gog J, Miller MA. 2007. The genesis and spread of reassortment human influenza A/H3N2 viruses conferring adamantane resistance. *Molecular biology and evolution* 24:1811–1820.
- Smith NG, Eyre-Walker A. 2002. Adaptive protein evolution in *Drosophila*. *Nature* 415:1022–1024.
- Steinhauer DA, Domingo E, Holland JJ. 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene* 122:281–288.



Stephan W. 2019. Selective Sweeps. *Genetics* 211:5-13.

Suarez DL. 2000. Evolution of avian influenza viruses. *Veterinary Microbiology* 74:15-27.

Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. 1992. Evolution and ecology of influenza A viruses. *Microbiological reviews* 56:152-179.

Worobey M, Holmes EC. 1999. Evolutionary aspects of recombination in RNA viruses. *Journal of General Virology* 80:2535-2543.

### 영문초록

**Title:** Evolutionary genetics of influenza virus

**Abstract:** Influenza virus causes severe infectious disease in humans and undergoes very rapid evolution in their antigenic structure, which enables them to evade host immunity. Evolutionary analysis of influenza virus is therefore important not only for public health but also for advancing evolutionary biology in general, thanks to several unique advantages that influenza virus provides in research. Type A influenza evolved in its natural hosts, ducks and shorebirds, and spread to humans and livestock, new hosts that recently attained high density suitable for viral transmission. Researches therefore focused on evolutionary changes accompanying such host switches. The methods of population genetic inference developed for other organisms are being applied to the viral population for the investigation of key evolutionary genetic factors (mutation, reassortment, antigenic drift, effective population size, frequency-dependent selection, and so on). It is important to modify such methods for studying influenza virus, particularly to deal with irregular time series of viral samples,

**Authors:** Yuseob Kim\*

**Affiliation:** Division of EcoScience, Ewha Womans University, Seoul, Korea 03760

**Corresponding author:** \*yuseob@ewha.ac.kr