

경북지역 코로나19 진단검사 특이사례 변이 특성 분석

박정주 · 박성익 · 박윤지 · 신지윤 · 차성경 · 김사라 · 이창일

바이러스분석과

Characterization of COVID-19 Diagnosis Genetic Mutation in Gyeongbuk Region

J. J. Park, S. I. Park, Y. J. Park, J. Y. Shin, S. K. Cha, S. R. Kim and C. I. Lee

Viral Disease Division

Abstract

In order to analyze the differences between the results of coronavirus disease 2019 (COVID-19) genomic diagnosis in Gyeongbuk region, we performed full length genome analysis of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) used by Ion Torrent Genexus Integrated Sequencer and dedicated analysis program. Full-length genomic sequence information of the SARS-CoV-2 was obtained from specimen of 24 confirmed cases of COVID-19. As a result of the analysis, the Omicron sub-lineages were BA.1.1 (n=8), BA.2 (n=10), BA.2.3 (n=4), BA.2.3.12(n=1) and BA.2.65(n=1). In detail, R61C or R61H were confirmed in the E gene, and L142W, P141L and 166-178 nucleotide deletions were detected in the ORF1ab NSP3 region. These results suggest that the detailed lineage and genetic mutations of SARS-CoV-2 using whole genome sequencing may be used to improve molecular diagnosis for COVID-19.

Key Words : COVID-19, SARS-CoV-2, Whole Genome Sequencing, Omicron variants, Real-time RT-PCR

1. 서 론

코로나바이러스감염증-19(코로나19) 전 세계적 확산과 유행이 장기화되면서 다양한 유전적 계통의 변이 바이러스들이 출현하고 있다. 세계보건기구(WHO)는 각 변이 바이러스들의 발생 현황과 공중보건에 미치는 영향 등을 평가하여 주요 변이(Variant of concern, VOC)와 기타 변이(Variant of interest, VOI)로 분류하고 이를 주기적으로 업데이트하고 있다.¹⁾

주요 변이는 전파력 증가 혹은 역학적으로 부정적 변화가 확인되고, 병원성 증가 혹은 임상적으로 중증도 변화가 확인되거나 진단, 백신, 치료제 등의 유효성 저하가 확인된 변이로 정의된다. 2020년 12월에 알파와 베타, 2021년 1월 감마, 2021년 5월 델타, 2021년 11월에 오미크론이 주요 변이로 지정되었고, 현재는 오미크론만이 현재 유행 중인 주요 변이(currently circulating VOC) 범주에 포함되어 있다.^{1, 2)}

오미크론은 전 세계적으로 지배적인 변이로 2022년 이후 GISAID(Global Initiative on Sharing All Influenza Data)에 공유되는 바이러스의 98% 이상을 차지하고 있다. 오미크론의 유행과 전파가 지속되고 오미크론 변이 내에서도 다양한 하위 변이가 나타나면서 BA.5는 140개, BA.4는 20개, BA.2는 172개의 세부 계통(sublineages)으로 재분류 되었다('22.10월).^{3, 4)}

코로나19의 원인병원체인 SARS-CoV-2(Severe acute respiratory syndrome coronavirus 2)는 양성 단일가닥의 RNA 바이러스로, DNA 바이러스 등과 비해 비교적 높은 빈도로 변이 발

생이 가능하다. 유전체의 크기는 대략 30 kb이고, 5' 방향으로부터 2/3 정도를 차지하는 ORF1 유전자는 pp1a과 pp1ab 복제효소 다단백질(replicase polyproteins)로 번역되는 ORF 1a와 ORF 1b를 포함하고 있으며, 복제과정 중 16개의 비구조단백질(NSP 1~16)로 분할된다. 나머지 1/3의 유전자는 6개의 보조단백질(accessory proteins)과 스파이크 단백질(spike protein, S), 막 단백질(membrane protein M), 외피(envelope, E), 뉴클레오캡시드(nucleocapsid, N)의 구조단백질을 암호화하고 있다.^{5, 6)}

국내 코로나19 진단검사법은 민감도와 특이도가 높은 실시간 역전사 중합효소연쇄반응(Real-time RT-PCR) 기법이 주로 이용되고 있으며, 2개 이상 다수의 SARS-CoV-2 특이 유전자를 검출하는 검사법을 사용할 것을 권장하고 있다. 식품의약품안전처의 허가를 받은 유전자검출검사법 기반 진단검사 제품의 대부분은 코로나19 유전체내 비교적 잘 보존된 ORF1b 부위의 RdRp(RNA-dependent RNA polymerase), E, N gene 등을 표적 유전자로 선택하여 검출하고 있으나⁸⁾, PCR 증폭을 위해 표적 유전자내 특이 부위에 결합하는 primer의 염기서열은 제조사별로 차이를 보인다.

SARS-CoV-2 유전자검출검사 판정 시 동일한 제품을 사용했을 경우 표적 유전자의 검출값(Ct value)이 일정 범위내에서 비슷한 값으로 검출되는 것이 일반적이나, 최근 경북지역 일부 코로나19 환자의 검체에서 표적 유전자가 불검출되거나 검출값의 차이를 보이는 사례가 확인되었다. 본 연구에서는 전장유전체 분석을 기반으로 각각의 변이 세부계통을 확인하고, 표적 유전자

영역에 대한 염기서열 분석을 통하여 이들의 상관관계를 분석하고자 하였다.

2. 재료 및 방법

2.1. 코로나19 확진자 검체

2022년 2월~4월 동안 코로나19 확진자와 접촉하여 군위군, 의성군, 청도군 및 포항시 일대 선별진료소(보건소)를 내원한 코로나19 의심 환자 24명의 임상 가검물(구·비인두도말)을 확보하였다.

2.2. 코로나19 진단 검사

확보된 각각의 검체 140 μ l를 QIAamp Viral RNA Kit(QIAGEN)에 제시된 방법에 따라 핵산 추출한 뒤 PCR 반응의 주형으로 사용했다. 코로나19 진단 검사는 식약처 허가를 받은 Real-time RT-PCR기법의 국내 2개 제조사 제품을 QuantStudio5 System(Thermo Fisher Scientific)을 활용하여 검사하였다. A사 제품은 ORF1ab(FAM), E(JOE) gene을 표적(형광검출)으로 하고, 50 $^{\circ}$ C 30분(reverse transcription), 95 $^{\circ}$ C 10분(denaturation), 95 $^{\circ}$ C 15초에서 60 $^{\circ}$ C 1분 40회 반복(amplification)하였으며, B사 제품은 RdRp(FAM), E(JOE) gene을 표적으로 50 $^{\circ}$ C 15분(reverse transcription), 95 $^{\circ}$ C 3분(denaturation), 95 $^{\circ}$ C 5초에서 60 $^{\circ}$ C 40초 5회 반복(pre-amplification), 95 $^{\circ}$ C 5초에서 60 $^{\circ}$ C 40초 40회 반복(amplification)한 후 각 제조사 기준에 따라 확진 판정하였다.

2.3. 주요 변이 PCR 검사

코로나19 주요 변이 5종에 대한 추정 감별은 PowerChek SARS-CoV-2 S-gene Mutation Detection Kit Ver.3.0(Kogenebiotech)를 사용하여 Real-time RT-PCR기법(QuantStudio5 System)으로 검사하였다. PCR 반응조건은 50 $^{\circ}$ C 15분(reverse transcription), 95 $^{\circ}$ C 5분(denaturation), 95 $^{\circ}$ C 10초에서 60 $^{\circ}$ C 30초 5회 반복(pre-amplification), 95 $^{\circ}$ C 5초에서 60 $^{\circ}$ C 40초 40회 반복(amplification)한 후 제조사에서 제시한 결과 분석표에 따라 감별하였다.

2.4. 전장유전체 정보 분석

코로나19 확진자로부터 추출한 바이러스 유전자를 TaqMan 2019 nCoV Assay kit v1(Applied biosystems, Thermo Fisher Scientific)을 사용하여 정량하고, Ion Torrent Genexus Integrated Sequencer와 Ion AmpliSeq SARS-CoV-2 Insight Research Assay GX 등 분석시약(Thermo Fisher Scientific)을 사용하여 전장유전체 분석을 실시하였다. 확보한 코로나19 전장유전체는 SARS-CoV-2-Panglin과 CLC Main Workbench(Version 21.0.3, QIAGEN)를 사용하여 세부계통과 유전체 염기서열을 비교 분석하였다.^{3, 4)}

3. 결과

3.1. 검체 정보 및 코로나19 진단 검사

유전자검출검사를 실시한 코로나19 확진자는

I . 연구사업

검체 기간별로 2월 1명, 3월 16명, 4월 7명이었고, 지역은 군위군 18명, 의성군 3명, 포항시 2명, 청도군 1명으로 남성 13명, 여성 11명이었고 연령 분포는 2세~87세로 평균 38세이다. A사 진단시약을 사용하여 검사한 결과 3가지 유형의 특이사례가 확인되었다. 군위군에서 확보한 9명과 포항시에서 확보한 1명(1~10번)은 ORF1ab 유전자의 Ct 값과 비교시 E 유전자가 증폭되지 않거나 Ct 값이 현저하게 증가되었고, 의성군에서 확보한 3명(11~13번)은 ORF1ab 유전자가 증폭되지 않거나 판정기준(Cut-off 38) Ct 값보다 높게 나타났다. 3번째 유형은 A사 진단시약 사용 시 일반적으로 표적유전자인 ORF1ab와 E 유전자의 Ct 값이 비슷하거나 E 유전자의 Ct 값이 1~2 정

도 증가하는 것과는 반대로, 군위군에서 확보한 9명과 포항시, 청도군에서 확보한 각 1명(14~24번)은 E 유전자 대비 ORF1ab 유전자의 Ct 값이 2~3 정도 증가되는 현상을 보였다. 미증폭 또는 양성 판정기준에 부합되지 않는 4명(1, 10~13번)은 재검사에서 동일한 현상이 나타나, RdRp와 E유전자를 표적으로 하는 B사 진단시약으로 검사 후 최종 양성 판정하였다(Table 1).

3.2. 주요 변이 PCR 검사

코로나19 특이 유전자 부위인 N gene과 spike 단백질을 암호화하는 S gene에 존재하는 변이 부위(N501Y, K417N, E484K, P681R, E484A, L452R, T547K)의 PCR 증폭 유무에 따라 주요

Table 1. Characteristics of Ct values of SARS-CoV-2 in Gyeongbuk region

No	Date	Region	Age	Sex	A kit (Value of Ct)		B kit (Value of Ct)		Characteristics
					ORF1ab	E	RdRp	E	
1	2.22.	GW	65	F	18.08	-	17.18	17.10	Not detected E gene or Ct value increasement of E gene
2	3. 4.	GW	55	M	23.1	34.31	16.94	18.88	
3	3. 8.	GW	56	M	23.68	31.37	21.01	20.14	
4	3. 8.	GW	43	F	20.52	26.58	17.80	16.83	
5	3.10.	GW	41	F	20.05	28.65	14.86	14.47	
6	3.10.	GW	9	F	21.24	31.23	16.93	16.90	
7	3.11.	GW	47	M	19.96	32.42	14.96	15.16	
8	3.13.	GW	55	M	20.45	36.04	15.97	16.58	
9	3.26.	GW	87	F	17.77	25.04	15.62	16.39	
10	4. 4.	PH	53	M	21.68	-	18.92	19.21	
11	3.28.	US	12	M	-	20.06	15.20	15.19	Not detected ORF1ab
12	3.28.	US	14	M	38.23	19.01	14.06	14.18	
13	3.28.	US	14	F	38.14	18.91	14.12	13.98	
14	3.20.	GW	38	F	19.15	16.61	14.92	15.50	ORF1ab Ct value increased compared to E gene
15	3.20.	GW	10	M	19.07	16.71	15.13	15.97	
16	3.20.	GW	2	M	22.96	20.30	17.69	18.49	
17	3.23.	GW	39	M	22.06	19.11	16.01	16.35	
18	3.23.	CD	39	F	19.36	16.46	14.56	15.43	
19	4. 1.	PH	64	M	22.83	20.21	15.96	16.93	
20	4. 4.	GW	20	F	26.76	24.43	19.53	20.38	
21	4.16.	GW	38	F	22.66	20.22	16.04	16.72	
22	4.16.	GW	9	M	24.53	21.98	17.95	18.77	
23	4.20.	GW	50	M	24.94	21.84	17.59	18.18	
24	4.25.	GW	47	F	19.09	16.79	14.82	15.12	

*Ct: cycle threshold, -: not detected, GW: Gunwi-gun, PH: Pohang city, CD: Cheongdo-gun, US: Uiseong-gun.

변이 5종(오미크론 BA.1과 BA.2, 알파, 베타, 델타, 감마 변이)을 추정 감별하였으며, 분석 결과 모두 오미크론 BA.1(K417N, E484A, T547K 증폭) 또는 오미크론 BA.2(K417N, E484A 증폭) 변이로 확인되었다(결과 미제시). 각각의 세부계통에 대한 최종 판정은 전장유전체 정보 분석을 통하여 판단하였다.

3.3. 전장유전체 분석

코로나19 전장유전체 정보 분석을 통하여

2022년 2~4월 동안 경북지역에서는 오미크론 세부변이 BA.1 계통과 BA.2 계통이 동시에 유행하는 것을 확인하였다. 진단 검사특이사례에 대한 변이 세부계통은 오미크론 BA.1.1이 8명, BA.2 10명, BA.2.3이 4명, BA.2.65와 BA.2.3.12 이 각각 1명 확인(Pangolin version 4.1.3)되었다. 유형별 유전체 분석 결과, ORF1ab 유전자의 Ct 값과 비교 시 E 유전자가 증폭되지 않거나 Ct 값이 현저하게 증가된 유형에서는 특이적으로 R61C 변이(1번 BA.2.3, 10번 BA.2.65)와 R61H

Table 2. Distribution of SARS-CoV-2 variants and main amino acid mutations

No.	GISAID	Pango Lineage	Main amino acid mutations	
			E gene	NSP3
1	GRA	BA.2.3	T9I, R61C	T24I, G489S
2		BA.1.1	T9I, R61H	K38R, S1265del, L1266I, A1892T
3		BA.1.1		
4		BA.1.1		
5		BA.1.1		
6		BA.1.1		
7		BA.1.1		
8		BA.1.1		
9		BA.1.1		
10		BA.2.65	T9I, R61C	T24I, G489S
11		BA.2.3	T9I	T24I, 166-178 del , G489S
12		BA.2.3		
13		BA.2.3		
14		BA.2	T9I	T24I, L142W , G171V, G489S
15		BA.2		
16		BA.2		T24I, L142W , G489S
17		BA.2		
18		BA.2		T24I, F25L, E26N, D28M, E29K, R30G, I31L, D32I, V34Y, N36M, E37R, K38del, C39del, Y42L, T43P, V44I, E45Q, G47N, T48S, L142W , G171V, G489S
19		BA.2.3.12		T24I, P141L , G489S
20		BA.2	T9I	T24I, L142W , G171V, G489S
21		BA.2		
22		BA.2		
23		BA.2		
24		BA.2		

* GISAID: Global Initiative on Sharing Avian Influenza Data

A; Alanine C; Cysteine D; Aspartic acid E; Glutamic acid F; Phenylalanine G; Glycine H; Histidine I; Isoleucine K; Lysine L; Leucine M; Methionine N; Asparagine P; Proline Q; Glutamine R; Arginine S; Serine T; Threonine V; Valine W; Tryptophan Y; Tyrosine del; deletion NSP3: The multi-domain Non structural protein 3

변이(2~9번 BA.1.1)가 확인되었다. ORF1ab 유전자가 증폭되지 않거나 판정기준 보다 높게 나타난 유형에서는 3명(11~13번) 모두 BA.2.3 계통으로 NSP3 유전자 부위에 연속적인 유전자 결손(166-178)이 확인되었다. E 유전자 대비 ORF1ab 유전자의 Ct값이 증가된 유형(14~24번)에서도 NSP3 유전자 부위에서 특이적인 변이가 확인되었다. 군위군에서 확보된 8명은 모두 BA.2 계통으로 L142W 변이가 공통적으로 나타났고, 청도군 1명(18번)에서는 추가적으로 변이가 다수 확인되었다. 나머지 포항시에서 확보된 1명(19번)은 BA.2.3.12 계통으로 P141L 변이가 확인되었다(Table 2). RdRp 유전자를 포함하는 나머지 15개 NSP 도메인에서는 특이적인 변이가 확인되지 않았다.

4. 결론

2022년 2월~4월 경북지역 코로나19 확진자 검체를 대상으로 전장유전체 분석을 통하여 변이 세부계통을 확인하고 진단 검사에 영향을 미치는 유전자 변이 발생 부위를 분석하였다. 오미크론 BA.1.1, BA.2, BA.2.3, BA.2.3.12, BA.2.65 세부계통의 지역 내 유행을 확인하였고, E 유전자 부위에서 R61C 또는 R61H 변이, ORF1ab NSP3 유전자 부위에서 L142W, P141L 변이와 166-178 영역에서 염기서열이 결손된 특이 변이를 확인하였다.

Real-time RT-PCR기법의 진단 검사에서 민감도와 특이도에 영향을 미치는 중요한 요소

중의 하나는 primer 서열의 디자인이다. 돌연변이가 잘 일어나지 않는, 매우 잘 보존되는 표적 유전자를 선택하고 각각의 유전자마다 특이적으로 인식할 수 있는 최적의 염기서열을 결정한 후, 선택된 primer가 *in silico* 상에서 기능을 하는지 판단하고, dimerization 반응 여부 등의 확인 과정을 거쳐 최적화한다.^{10, 11)} Banko 등의 보고에서 Real-time PCR기법의 상용화된 진단 시약의 성능을 비교 평가하고, 개선을 위한 필수 사항으로 새로운 변이 출현에 따른 신속한 primer 개발을 제시하였다.¹²⁾

본 연구의 전장유전체 분석 정보를 바탕으로 컴퓨터 프로그램을 활용한 인실리코 분석을 실시한 결과, E 유전자 부위의 R61C 또는 R61H 변이는 특정 제조사 진단 시약의 검출 표적 부위가 아니었으나 일부 검체에서 확인된 NSP3 유전자 부위의 L142W 변이와 166-178 영역의 염기서열 결손은 동일한 진단 시약으로 검사 시 유전자 검출에 영향을 주는 부위로 확인되었다.(질병관리청 경북권질병대응센터)

코로나19 유전자검출검사 시 2개 이상의 표적 유전자를 선택하므로, 이 가운데 하나만 양성으로 판정되는 경우 등 예상하지 못한 결과가 나타날 경우 표적이 다른 진단 시약을 선택하여 특정 유전자 변이에 따른 영향을 최소화 할 수 있다.

따라서, 코로나19 전장유전체 분석을 통하여 새로운 변이 바이러스 출현에 대한 감시뿐만 아니라 변이 발생이 유전자 진단에 미치는 영향을 분석하여 검사법 개선에 활용할 수 있을 것으로 기대한다.

참 고 문 헌

1. World Health Organization, <https://who.int/en/activities/tracking-SARS-CoV-2-variants>
2. PANGO NETWORK, Tracking SARS-CoV-2 variants, <https://www.pango.network>
3. GISAIID, [https://gisaid.org/Global Initiative on Sharing Avian Influenza Data](https://gisaid.org/Global%20Initiative%20on%20Sharing%20Avian%20Influenza%20Data)
4. PANGO Lineages, <https://cov-lineages.org>
5. Khailany, R. A., M. Safdar, and M. Ozaslan “Genomic characterization of a novel SARS-CoV-2”. *Gene Rep.* 19: 100682(2020).
6. Eun-Joong Kim, Dongsup Lee, “Coronaviruses: SARS, MERS and COVID-19”, *Korean J Clin Lab Sci.* 52(4): 297-309(2020).
7. 박송희, “COVID-19 진단기법: Primer에 따른 SARS-CoV-2 (COVID-19)의 핵산기반 진단법 분석”. *Korean Society for Biotechnology and Bioengineering Journal* 36(1): 23-29(2021)
8. 질병관리청, “주요변이바이러스가 국내 코로나 19 확진진단에 미치는 영향”, *주간건강과질병* 제 14권 제35호(2021)
9. 이경훈, “코로나바이러스감염증-19 진단검사에 대한 이해”, *Ewha Med J* 44(1):1-10(2021)
10. Li, D., J. Zhang, and J. Li, “Primer design for quantitative real-time PCR for the emerging coronavirus SARS-CoV-2”, *Theranostics.* 10: 7150-7162(2020).
11. Park, M., J. Won, B. Y. Choi, and C. J. Lee, “Optimization of primer sets and detection protocols for SARS-CoV-2 of coronavirus disease 2019 (COVID-19) using PCR and real-time PCR”. *Exp Mol Med.* 52: 963-977(2020).
12. Banko A., Petrovic G., Miljanovic D., et al., “Comparison and sensitivity evaluation of three different commercial real-time quantitative PCR kits for SARS-CoV-2 detection”. *Viruses* 13: 1321-1332(2021).

