

## Article

# Molecular Analysis of SARS-CoV-2 Lineages in Armenia

Diana Avetyan <sup>1,2,\*</sup>, Siras Hakobyan <sup>3,4</sup>, Maria Nikoghosyan <sup>2,3</sup>, Lilit Ghukasyan <sup>1</sup>, Gisane Khachatryan <sup>1,2</sup>, Tamara Sirunyan <sup>1,2</sup>, Nelli Muradyan <sup>1</sup>, Roksana Zakharyan <sup>1,2</sup>, Andranik Chavushyan <sup>1,5</sup>, Varduhi Hayrapetyan <sup>1,2</sup>, Anahit Hovhannisyanyan <sup>2,6</sup>, Shah A. Mohamed Bakhsh <sup>7,8</sup>, Keith R. Jerome <sup>7,8</sup>, Pavitra Roychoudhury <sup>7,8</sup>, Alexander L. Greninger <sup>7,8</sup>, Lyudmila Niazyan <sup>9</sup>, Mher Davidyants <sup>9</sup>, Gayane Melik-Andreasyan <sup>10</sup>, Shushan Sargsyan <sup>10</sup>, Lilit Nersisyan <sup>4,11</sup> and Arsen Arakelyan <sup>1,2,3,4,\*</sup>

- <sup>1</sup> Laboratory of Human Genomics, Institute of Molecular Biology NAS RA, Yerevan 0014, Armenia; l\_ghukasyan@mb.sci.am (L.G.); g\_khachatryan@mb.sci.am (G.K.); t\_sirunyan@mb.sci.am (T.S.); n\_muradyan@mb.sci.am (N.M.); roksana.zakharyan@rau.am (R.Z.); a\_chavushyan@mb.sci.am (A.C.); haivard@mail.ru (V.H.)
- <sup>2</sup> Institute of Biomedicine and Pharmacy, Russian-Armenian University, Yerevan 0051, Armenia; m\_nikoghosyan@mb.sci.am (M.N.); hovhannisyanyanahit19@gmail.com (A.H.)
- <sup>3</sup> Bioinformatics Group, Institute of Molecular Biology NAS RA, Yerevan 0014, Armenia; s\_hakobyan@mb.sci.am
- <sup>4</sup> Armenian Bioinformatics Institute, Yerevan 0014, Armenia; lilit.nersisyan@ki.se
- <sup>5</sup> Davidyants Laboratories, Yerevan 0054, Armenia
- <sup>6</sup> Laboratory of Evolutionary Genomics, Institute of Molecular Biology NAS RA, Yerevan 0014, Armenia
- <sup>7</sup> Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA 98102, USA; shahm4@uw.edu (S.A.M.B.); kjerome@fredhutch.org (K.R.J.); proychou@fredhutch.org (P.R.); agrening@uw.edu (A.L.G.)
- <sup>8</sup> Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
- <sup>9</sup> NORC Infection Clinical Hospital, MoH RA, Yerevan 0047, Armenia; lyudmila.niazyan@gmail.com (L.N.); davidyants@gmail.com (M.D.)
- <sup>10</sup> National Center of Disease Control and Prevention, Ministry of Health RA, Yerevan 0025, Armenia; gayane.melik-andreasyan@ncdc.am (G.M.-A.); premier\_h@yahoo.com (S.S.)
- <sup>11</sup> SciLifeLab, Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, 17177 Solna, Sweden
- \* Correspondence: d\_avetyan@mb.sci.am (D.A.); aarakelyan@sci.am (A.A.)



**Citation:** Avetyan, D.; Hakobyan, S.; Nikoghosyan, M.; Ghukasyan, L.; Khachatryan, G.; Sirunyan, T.; Muradyan, N.; Zakharyan, R.; Chavushyan, A.; Hayrapetyan, V.; et al. Molecular Analysis of SARS-CoV-2 Lineages in Armenia. *Viruses* **2022**, *14*, 1074. <https://doi.org/10.3390/v14051074>

Academic Editors: Judd F. Hultquist, Ramón Lorenzo-Redondo and Egon Anderson Ozer

Received: 4 April 2022

Accepted: 13 May 2022

Published: 17 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** The sequencing of SARS-CoV-2 provides essential information on viral evolution, transmission, and epidemiology. In this paper, we performed the whole-genome sequencing of SARS-CoV-2 using nanopore and Illumina sequencing to describe the circulation of the virus lineages in Armenia. The analysis of 145 full genomes identified six clades (19A, 20A, 20B, 20I, 21J, and 21K) and considerable intra-clade PANGO lineage diversity. Phylodynamic and transmission analysis allowed to attribute specific clades as well as infer their importation routes. Thus, the first two waves of positive case increase were caused by the 20B clade, the third peak caused by the 20I (Alpha), while the last two peaks were caused by the 21J (Delta) and 21K (Omicron) variants. The functional analyses of mutations in sequences largely affected epitopes associated with protective HLA loci and did not cause the loss of the signal in PCR tests targeting ORF1ab and N genes as confirmed by RT-PCR. We also compared the performance of nanopore and Illumina short-read sequencing and showed the utility of nanopore sequencing as an efficient and affordable alternative for large-scale molecular epidemiology research. Thus, our paper describes new data on the genomic diversity of SARS-CoV-2 variants in Armenia in the global context of the virus molecular genomic surveillance.

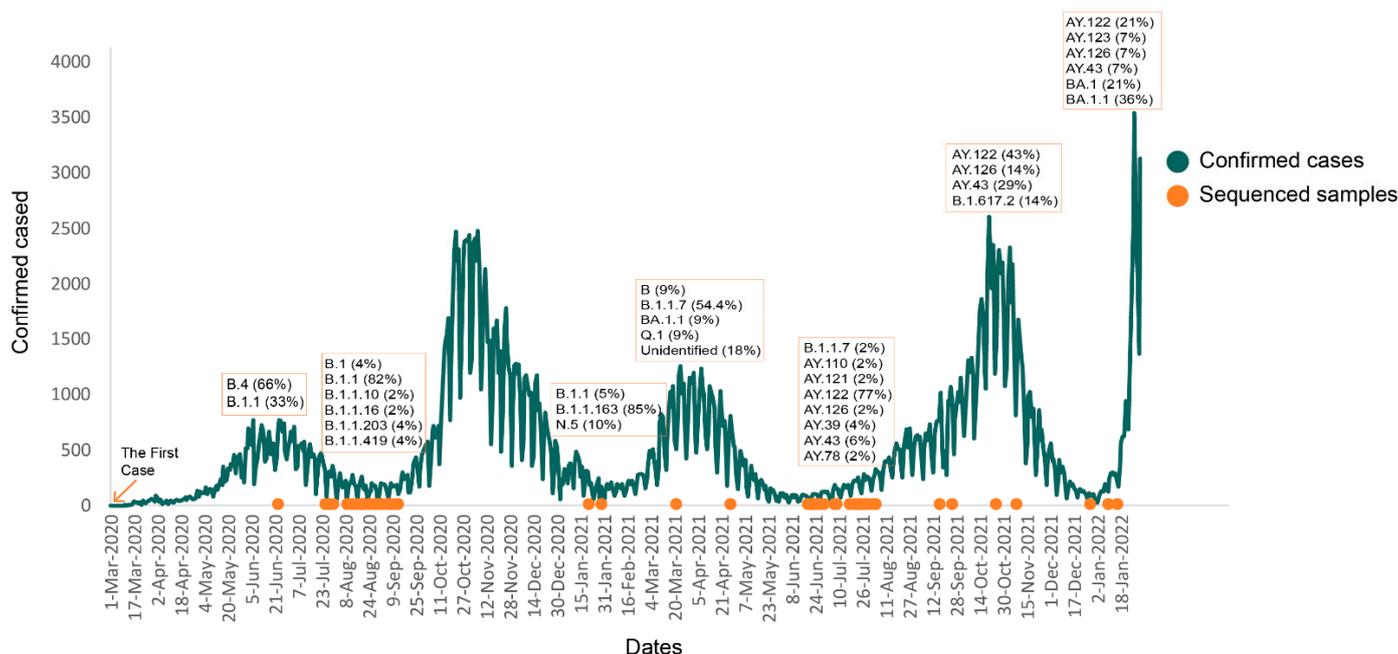
**Keywords:** COVID-19; SARS-CoV-2; coronavirus; nanopore sequencing; Illumina sequencing; whole-genome sequencing; Armenia

## 1. Introduction

Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), which causes the novel coronavirus pneumonia COVID-19 [1], was first identified in China, in the city of

Wuhan, in December 2019. The complete genome sequence of SARS-CoV-2 was published in January 2020 [2–4] and led to the development of real-time reverse transcription polymerase chain reaction (qRT-PCR) assays for SARS-CoV-2 detection that have served as a diagnostic standard during the ongoing COVID-19 pandemic [5]. Since then, whole-genome sequencing has been used for the evolutionary analysis of the virus, monitoring of circulating genetic lineages, and identifying signs of adaptation to hosts, which have important implications for treatment and vaccine development [6–8]. In the last two years, hundreds of studies were published describing country, region-specific and global insights into the dynamics and sources of SARS-CoV-2 importations and transmissions [9–12]. These results were obtained from the analysis of viral sequences, which were continuously sampled throughout the pandemic period.

In Armenia, the first confirmed case was reported on 1 March 2020. Since then, the number of positive cases reached 374,878 (as of February 2022) with several peaks at different time periods (Figure 1) with 8060 deaths and many re-infections [13]. In the absence of sequencing facilities in the country, virtually nothing was known about the transmission histories and epidemiological dynamics of the virus in Armenia. From March–August 2020, only three samples from Armenia obtained in July were sequenced in the Institute of Virology Charité Universitätsmedizin Berlin, which were deposited in the GISAID EpiCov [14,15] (accession ID: EPI\_ISL\_683449; EPI\_ISL\_683450; EPI\_ISL\_683451) in late December 2020. Another set of samples from September–November 2020 was sequenced by our colleagues at the Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center (USA), which became available in May 2021 (GISAID accession IDs are presented in Supplementary Table S1). This delay in analysis of molecular epidemiologic information hampered the informed and timely decision making by health authorities. In early 2021, our laboratory established the SARS-CoV-2 nanopore sequencing protocol and was able to perform almost real-time genomic surveillance by the monthly sequencing of viral samples during 2021 and 2022.



**Figure 1.** Daily incidence of reported confirmed cases for Armenia, sampling dates, and clade distribution of the sequenced samples.

In the present study, we combine all above-mentioned genomic data to report the first molecular analysis of SARS-CoV-2 virus in Armenia in order to (1) understand the emergence and the transmission of the virus, (2) identify the most prevalent lineages

at different time points, and (3) investigate the potential functional consequences of the mutations detected in the sequenced Armenian samples.

## 2. Materials and Methods

### 2.1. Samples

One hundred and ninety-one samples isolated from nasopharyngeal swabs were obtained from the Nork infection clinical hospital and the National Center for Disease Control and Prevention, Ministry of Health RA (NCDC), which served as primary testing sites. These samples were randomly selected from batches of COVID-19 positive samples tested at NCDC (Armenia), immediately after the confirmation of positive status, between June 2020–February 2022. Three additional samples previously sequenced at Charité Universitätsmedizin Berlin, Institute of Virology (Germany) and deposited in the GISAID EpiCov [14,15] were also included in this study (accessions: EPI\_ISL\_683451, EPI\_ISL\_683450, EPI\_ISL\_683449). The total number of samples was 194.

### 2.2. Real-Time PCR Detection of SARS-CoV-2

Automated RNA isolation was performed with Maxwell RSC Instrument using Maxwell RSC Viral Total Nucleic Acid Purification Kit (Promega Corporation Inc, Fitchburg, WI, USA). SARS-CoV-2 PCR testing was performed using Real-Time PCR Detection Kit for COVID-19 Coronavirus CE-IVD kit (Biotech & Biomedicine (Shenyang) Group Ltd., Shenyang, China) targeting ORF1ab and N genes. Samples were selected based on viral RNA load as measured by Ct values between 18–35 (Supplementary Table S2) for both targets.

### 2.3. Sequencing

Samples were sequenced with Oxford Nanopore and Illumina platforms (Supplementary Table S3). Nanopore sequencing of 146 samples was performed at the Institute of Molecular Biology NAS RA. Illumina sequencing of 45 samples was performed at Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. Five nanopore samples were additionally resequenced on the Illumina Nextseq platform to compare genome coverage and consensus level accuracy.

### 2.4. Nanopore Sequencing

Nanopore sequencing was performed according to “nCoV-2019 sequencing protocol v3 (LoCost) V.3” [16] based on ARTIC SARS-CoV-2 sequencing protocol with ARTIC nCoV-2019 V3 PCR panel [17,18].

#### 2.4.1. cDNA Generation

RNA samples were directly used for the first-strand synthesis using the LunaScript RT SuperMix Kit (New England Biolabs, Ipswich, MA, USA) with random hexamer and oligo-dT primers. Briefly, 8 µL RNA were mixed with 2 µL LunaScript RT SuperMix (5X) and were placed in a thermocycler and incubated for 2 min at 25 °C, followed by 10 min at 55 °C and 1 min at 95 °C and cooling to 4 °C. cDNAs were immediately used in subsequent steps.

#### 2.4.2. Amplicon Generation

Primer pairs from the ARTIC V3 primer scheme were used to amplify amplicons in cDNA [19]. Two multiplex PCR reactions were performed with 2.5 µL cDNA, 12.5 µL Q5 Hot Start High-Fidelity 2X Master Mix (New England Biolabs, USA), and 4 µL ARTIC V3 pool 1 (10 µM) or 4 µL ARTIC V3 pool 2 (10 µM). PCR cycling conditions were: 98 °C for 30 s followed by 35 cycles of 98 °C for 15 s, 65 °C for 5 min, and hold at 4 °C. The amplified products were purified with a 0.4x volume of AMPure XP beads (Beckman Coulter, Brea, CA, USA) to exclude small nonspecific fragments.

#### 2.4.3. Barcoding and Library Preparation

The purified PCR amplicons were treated with NEBNext End repair/dA-tailing Module (New England Biolabs, USA) and were barcoded with native barcodes and sequencing adapters (EXP-NBD104 and EXP-NBD114 kits Oxford Nanopore Technologies, Oxford, UK). Twelve or twenty-four samples were multiplexed in each sequencing run.

#### 2.4.4. Nanopore Sequencing

After priming the flow cell, 15 ng of the final sequencing library diluted to a final volume of 75  $\mu$ L was loaded. Following the ligation sequencing kit (SQK-LSK109, Oxford Nanopore Technologies, UK) protocol, MinION Mk1B was used to perform genome sequencing in an FLO-MINSP6 R 9.4.1 flow cell for 3–6 h. The mean genome coverage across runs was  $289 \pm 189$  (Supplementary Materials Figure S1).

#### 2.4.5. Data Preprocessing, Demultiplexing, and Alignment

Base-calling and demultiplexing were performed using Guppy (v4.0.14). Raw FASTQ files were filtered and reads with lengths 400–700 b were selected using the ARTIC pipeline (release 1.1.0) [20]. Downstream analyses were performed using the nanopolish workflow implemented in the ARTIC pipeline [21]. The pipeline includes an alignment to the hCoV-19/Wuhan/WIV04 reference genome with minimap2 (2.17-r941) [22] followed by variant calling and consensus-building. Positions in consensus genomes with coverage lower than 20 were masked with “N” bases.

#### 2.5. Illumina Short-Read Sequencing

The short-read sequencing of 50 samples was performed using the Illumina Nextseq platform following the protocol described in detail elsewhere [23]. In brief, sequencing libraries were prepared using the Swift Biosciences Normalase Amplicon protocol and SARS-CoV-2 amplicon panel, which contains 341 primer pairs spanning nucleotides 200 to 29,741 of the Wuhan reference genome (NC\_045512v2) and produces amplicons ranging in length from 116 to 255 base pairs. The multiplex amplicon libraries were produced following the manufacturer’s recommendations, followed by  $1.0\times$  volumes of AMPure XP beads cleaning. Then, barcoded sequencing adapters were added to the amplicons. The resulting libraries were cleaned using  $0.85\times$  volumes of PEG NaCl and sequenced on the Illumina Nextseq instrument using  $2 \times 150$  reads. Genomes were assembled using a custom pipeline described previously [23,24].

#### 2.6. Phylogenetic and Variant Analysis

To perform preliminary QC of Armenian samples and select contextual samples, we initially screened our sequences with Nextclade [25] and PANGOLIN [26]. As contextual sample sequences GISAID nextregions selections were used (Global, Africa, Asia, Europe, North America, South America, and Oceania collections downloaded on 22 January 2022). In addition, we downloaded sequences for PANGO lineages B.1.1.163, B.1.1.419, B.1.1.528, and BA.1.1 that were detected in Armenian sequences, but were absent in the downloaded nextregions collections. In total, 17,721 contextual sequences were selected for combined analysis with Armenian samples.

Phylogenetic analysis, Nextstrain clade, and PANGO lineage identification was performed using the SARS-CoV-2 genomic epidemiology-specific pipeline implemented in Nextstrain 3.0.6 [25]. After preliminary QC (genome length more than 27,000, number of ambiguous reads < 3000), sequences were aligned to the reference genome (Wuhan/Hu-1/2019) with MAFFT v7.490 [27,28]. For phylogenetic context, we performed contextual subsampling based on genomic proximity [25] to select 10 sequences per country per year per month as representative sequences from the background. The final dataset consisted of 145 Armenian sequences (97 nanopore and 48 Illumina) and 9449 contextual background sequences. A maximum likelihood (ML) phylogenetic tree was constructed using IQ-TREE [29] under the GTR nucleotide substitution model. We used TreeTime [30] trait

reconstruction on the resulting time-labeled tree to infer Armenia-centered trans-country transmissions with a “migration model”. The temporal signal was evaluated by root-to-tip regression with TempEst v.1.5.3 [31].

The Coalescent Bayesian Skyline model of Armenian samples was constructed with BEAST v1.10.4 [32] with previously described parameters [9]. The model parameters were estimated with 40,000,000 Markov Chain Monte Carlo (MCMC) iterations, with 4,000,000 burn-in states and sampling every 1000 states. The MCMC parameter quality was assessed using Tracer v.1.7.1 [33] and were accepted if effective sampling size (ESS) values were higher than 100. The maximum clade credibility (MCC) tree was annotated using Tree Annotator v.1.8.4 [34].

### 2.7. Functional Annotation of SARS-CoV-2 Genomes

The functional annotation of SARS-CoV-2 genomes from Armenia included in the present study was performed using the Coronavirus Genome Analysis Tool (CorGAT) [35], where bioinformatic prediction of potential T-cell epitopes for SARS-CoV-2 were performed according to Kiyotani et al. (2020) [36].

## 3. Results and Discussion

### 3.1. Phylogenetic and Phylogeographic Analysis of Sequences

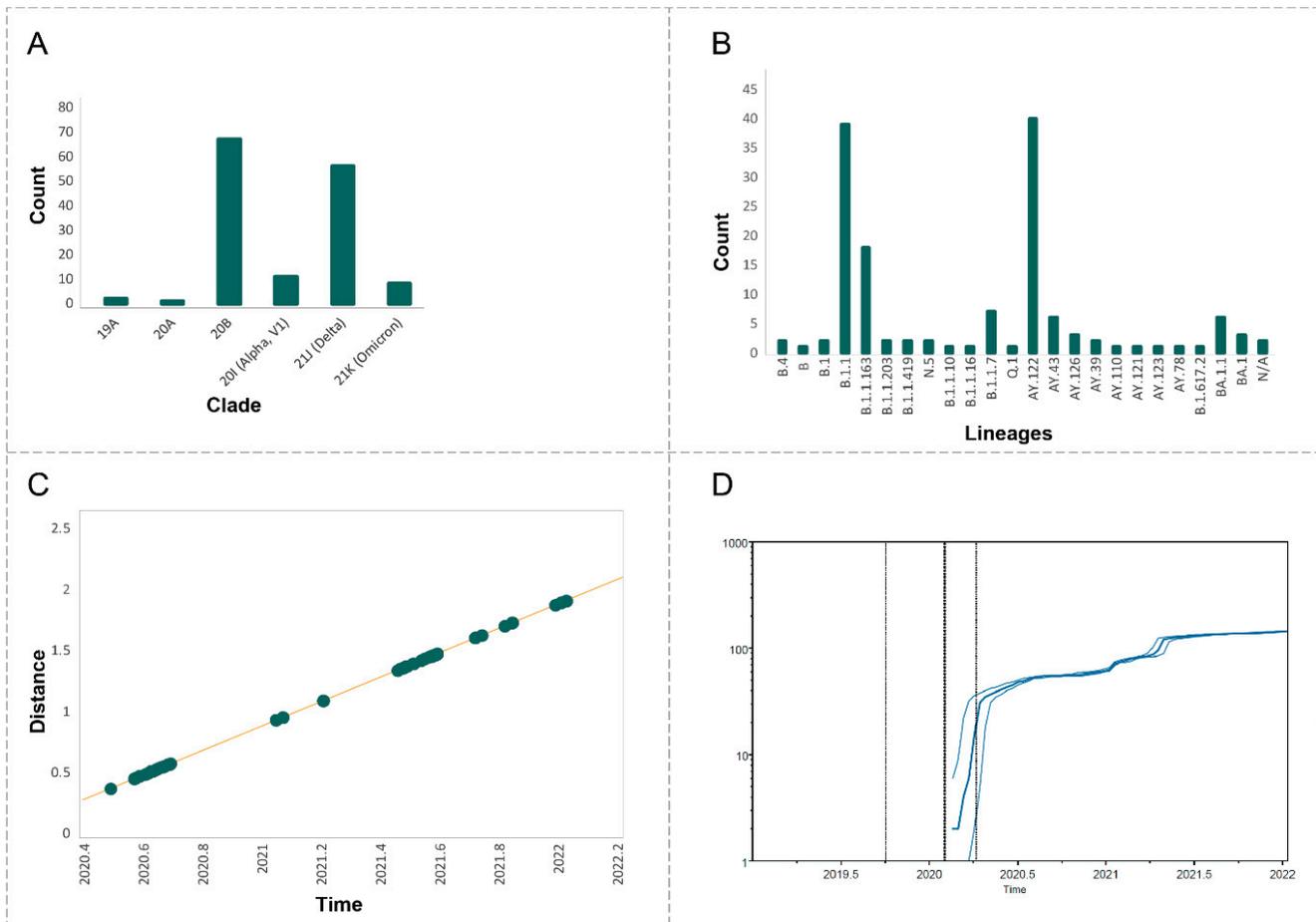
The total of 194 sequences represents 0.04% of 399,727 reported cases in Armenia as of 11 February 2022 (Figure 1).

Of the 194 sequenced samples, 145 (75%) sequences met the quality criteria (genome length of more than 27,000; number of ambiguous reads < 3000) and were included in the analyses. These 145 sequences represented 6 Nextstrain clades and 23 PANGO lineages (Figure 2A,B). The highest genomic diversity was noticed for the clades 21J (Delta) (nine PANGO lineages) and 20B (eight PANGO lineages). The analysis of root-to-tip regression with TempEst demonstrated a very strong temporal signal in our data (adjusted  $R^2 = 1$ ,  $2.9 \times 10^{24}$  on 1 and 143 DE,  $p$ -value:  $< 2.2 \times 10^{-16}$ , Figure 2C).

The analysis of clades in sequencing samples and lineage-through-time plots (Figure 2D) indicated several rounds of clade substitutions within the time of sampling. June 2020–January 2021 samples were mostly represented with sequences belonging to the clade 20B with only two 19A sequences. Samples from March 2021 belonged exclusively to the clade 20I (Alpha), while May–July 2021 samples were in majority represented by 21J (Delta). Finally, late 2021–January 2022 were represented mostly by 21K (Omicron). Further analysis demonstrated inter-clade variability for the time of introduction, transmission routes, and PANGO lineages (Figure 3).

The clade 19A (B.4) was represented by only two sequences. The estimated time for their introduction was late July (21 June 2020, Date Confidence Interval: 9 June 2020–27 June 2020); however, this clade was not detected in later samples suggesting its replacement in the Summer of 2020. The analysis of the transmission routes indicated Iran as the source of introduction for these sequences around early March 2020, which corresponds well with the date of the first positive case of COVID-19 in Armenia identified in a traveler from Iran [37]. Thus, we can speculate that July was the period of substitution of the 19A clade with 20B, which fits with the global domination of clades harvesting Spike D614G [38].

The clade 20B formed three big clusters associated with different introduction events as well as a few single introductions that did not result in large intra-country transmissions. Early introduction sources were Italy (2 March 2020, Date Confidence Interval: 25 February 2020–5 March 2020) and New Zealand (20 April 2020, Date Confidence Interval: 15 March 2020–22 May 2020). Interestingly, the latest introduction (12 October 2020, Date Confidence Interval: 18 September 2020–9 November 2020) observed was almost exclusively represented by the B.1.1.163 lineage imported from Russia that formed a big intra-country cluster. The estimated time of importation coincided with the sharp increase in positive cases in September–November 2020 (Figure 1).

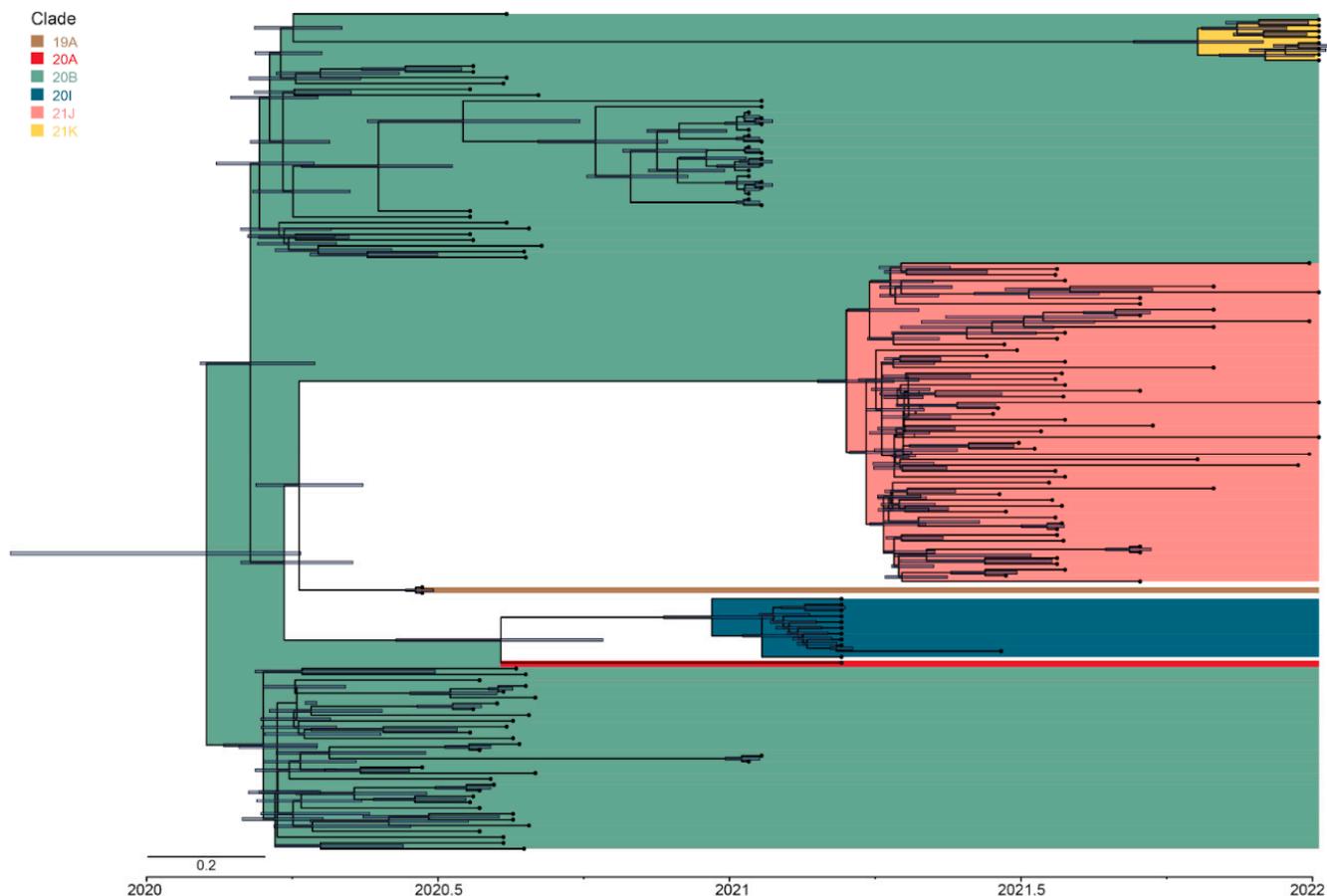


**Figure 2.** (A) Distribution of SARS-CoV-2 lineages according to Nextstrain clade nomenclature. (B) Distribution of SARS-CoV-2 lineages according to PANGO lineages nomenclature. (C) Root-to-tip regression analyses using TempEst. The plot represents linear regression of root-to-tip genetic distance within the ML phylogeny against sampling time. (D) “Lineage through time” plot of the phylogenetic tree of Armenian samples. The solid thick line represents the log median of ancestral lineages present at each time interval. Thin lines represent 95% confidence intervals for the number of lineages.

The clade 20B formed three big clusters associated with different introduction events as well as a few single introductions that did not result in large intra-country transmissions. Early introduction sources were Italy (2 March 2020, Date Confidence Interval 25 February 2020–5 March 2020) and New Zealand (20 April 2020, Date Confidence Interval: 15 March 2020–22 May 2020). Interestingly, the latest introduction (12 October 2020, Date Confidence Interval: 18 September 2020–9 November 2020) observed was almost exclusively represented by the B.1.1.163 lineage imported from Russia that formed a big intra-country cluster. The estimated time of importation coincided with the sharp increase in positive cases in September–November 2020 (Figure 1).

The variant of concern 20I (Alpha) had two introductions in Armenia. According to the temporal analysis, the lineage B.1.1.7 was introduced around 24 December 2020 (Date Confidence Interval: 26 September 2020–10 January 2021) and 24 November 2020 (Date Confidence Interval: 6 September 2020–15 January 2021). The Nextstrain pipeline identified Jordan and Germany as the main transmission route for the 20I (Alpha) clade sequences. The 20I (Alpha) was primarily responsible for the third peak of positive cases in late February–March 2021 (Figure 1). The introduction of this variant in Armenia happened with several months’ delay compared with the West European Countries and resulted in

considerably fewer infections as well. One of the reasons can be the strict travel restrictions and a negative 72 h PCR test requirement for inbound travel [39]. The other reason can be the peak of infection caused by 20B (B.1.1.163) in late 2020.



**Figure 3.** Annotated MCC tree of Armenian sequences. Background colors indicate Nextstrain clades; tip labels indicate PANGO lineages. Bars at nodes represent 95% height posterior density.

The estimated earliest introduction of the clade 20J (Delta) in Armenia was 28 December 2021 from India (B.1.617.2). The majority of the 20J (Delta) sequences, however, were represented by the AY.122 lineage (39 of 56 sequences) forming a single cluster with an estimated date 22 February 2021 (Date Confidence Interval: 8 December 2020–17 March 2021), introduced from Liechtenstein. More recent sequences for this clade have diverse geography (Bahrain, Denmark, Greece, India, Jordan, Portugal, South Africa, Spain, Suriname, and Turkey), but mostly without producing many secondary cases according to the phylogenetic tree.

Finally, the sequences belonging to the 21K (Omicron) clade demonstrated the highest geographical diversity of introduction (Brazil, France, Maldives, Mexico, Netherlands, and Sweden). The earliest inferred date for this clade introduction was estimated at 6 December 2021 (Date Confidence Interval: 11 October 2021–29 December 2021).

Both 20J (Delta) and 21K (Omicron) caused a sharp increase in positive cases compared to previous waves. On the other hand, the deaths accompanying the 21 (Delta) wave were considerably higher compared with the 21K (Omicron) (Supplementary Materials Figure S2), which is in line with observations in other countries [40,41].

Thus, our phylodynamic and phylogeographic analysis of the SARS-CoV-2 Armenian sequences allowed us to identify and characterize virus clades/lineages transmissions to Armenia. The results indicate multiple inter-country importations and their persistence in the country.

### 3.2. Functional Annotation of Variants

We performed the functional annotation of analyzed sequences using Nextclade [25] and CorGAT [35] tools (Supplementary Materials Tables S4 and S5). Besides known effects of clade/lineage signature mutations (for the most comprehensive list see <https://covariants.org/>, accessed on 12 May 2022), we were interested in the functional consequences of “private mutations” (reversions to reference, mutations ascribed to different clades, and mutations that are for reversions or belonging to other clades) as defined by the Nextclade app.

The reversions were identified in six 20I (Alpha) and one 21K (Omicron) samples, which constituted 54% and 12% of clade sequences, respectively. Most private mutations ascribed to other clades were detected in 20B (27 sequences) and 21J (Delta) (12 sequences); however, only seven such mutations were found in more than one sequence. Finally, at least one private unlabeled mutation was identified in all the 20B, 20I (Alpha), 21J, and 21K (Omicron) sequences. Overall, 138 mutations were detected in more than two samples. The functional annotation of all mentioned mutation types is provided in Supplementary Materials Table S5. No specific enrichment for HLA epitopes [42–44], evolutionary selection, or secondary structure elements were observed in private mutations compared to all known mutations (HLA-epitopes  $p_{\text{Fisher exact}} = 1$ ; Selection pressure  $p_{\text{Fisher exact}} = 0.45$ ; Secondary structure  $p_{\text{Fisher exact}} = 0.63$ ); however, the overlap between these categories was observed for some of the mutations (Figure 4).



**Figure 4.** Venn diagram depicting the overlap between the Nextclade private mutations in HLA epitopes, under evolutionary selection and causing changes in secondary structure elements (SSE) observed in the Armenian samples ( $n = 145$ ).

HLA loci association with COVID-19 incidence, risk, or severity has become one of the research focus areas. Many associations have been reported in various countries and populations [43,45]; for example, HLA-A\*02:01 was predicted to have a high binding to the virus epitopes and shown to be protective against COVID-19 severity, while HLA-A\*01:01 was considered a risk factor for the disease [43]. Moreover, another recent study evaluated the association of HLA loci with the side effects of mRNA vaccines [46]. Recently, the study of HLA loci association with COVID-19 has also identified HLA-C\*04:01 as a risk factor for severe disease in Armenia [47]. Previous population-scale studies identified

the common HLA alleles in the Armenian population (HLA-A\*02:01, HLA-A\*01:01, HLA-A\*24:02, HLA-A\*03:01, HLA-B\*51:01, HLA-B\*35:01, and HLA-B\*49:01) [48]. We evaluated the representation of epitopes targeted by these loci in our sequences. We identified epitopes for two protective HLA loci (HLA-A\*02:01 and HLA-A\*24:02) and three risk loci (HLA-A\*01:01, HLA-A\*03:01, and HLA-B\*51:01). Our results demonstrate that the majority of sequences harbor mutations in epitopes with a high-binding affinity to protective HLA loci, while only a few sequences showed the presence of the mutations associated with low-binding HLA alleles (Table 1, Supplementary Material Table S5). No HLA-C\*04:01 locus-related epitopes/mutations were observed; however, this allele was not present in the CorGAT's HLA annotation dataset. Thus, the question of whether mutations in viral sequences may be related to the observed association of HLA-C\*04:01 with disease severity remains open.

**Table 1.** Absolute count of sequences harboring private mutations in HLA epitopes across clades.

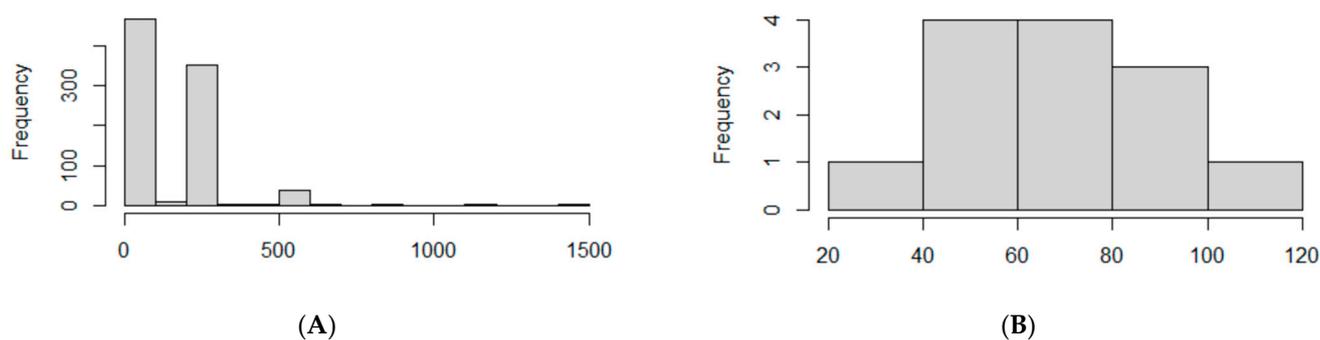
Clade	HLA-A*02:01, HLA-A*24:02 (Protective Alleles)	HLA-A*01:01, HLA-A*03:01, HLA-B*51:01 (Risk Alleles)
19A	3 mutations	2 mutations
20B	228 mutations	128 mutations
20I (Alpha, V1)	68 mutations	24 mutations
21J (Delta)	453 mutations	303 mutations
21K (Omicron)	76 mutations	19 mutations

Polymerase chain reaction (PCR) is the current standard method for COVID-19 clinical diagnosis from clinical samples. Therefore, we conducted a reassessment of published diagnostic PCR assays, including those recommended by the World Health Organization (WHO), through the evaluation of the possible effect of identified mutations on the efficacy of recommended primers and probes used for PCR detection of SARS-CoV-2 with the Nextclade app. In 143 sequences, we observed 39 mutations in the viral genome that did not match RT-PCR primers/probes for SARS-CoV-2 detection (Supplementary Materials Table S6). However, mutations located in template regions for US CDC N3 and China CDC Orf1AB primers and probes did not influence the primer binding since we obtained the N gene PCR signal in all studied samples (Supplementary Materials Table S2).

### 3.3. Comparison of Oxford Nanopore and Illumina Sequencing

In this study, 97 nanopore sequencing samples and 48 Illumina sequencing samples were included, which gave us an opportunity to evaluate the performance of both approaches. First, we assessed the number of missed nucleotides in the Nextclade app analysis, which can indicate gaps in genomes because of insufficient read coverage. Out of the 97 nanopore samples, 86 had missing sites, while in Illumina samples, they were detected only in 5 samples. The length distribution of missing sites was  $146 \pm 176$  nt and  $70 \pm 22$  nt in nanopore and Illumina samples, respectively. The large SD in nanopore samples is caused by a high number of single missing sites as well as supposed amplicon drop-outs (Figure 5).

We also sequenced five samples by two methods, so we compared them to evaluate the correspondence of clade/lineage assignment with nanopore and Illumina short-read sequencing. We compared the Nextstrain clade and PANGO lineage assignment for the consensus these sequences produced (Table 2, Supplementary Materials Table S7).



**Figure 5.** The bar graph represents the size distribution of missing sites for (A) nanopore and (B) Illumina samples. The X-axis represents the size of missing sites; Y-axis represents the absolute frequency (number) of missing sites per size.

**Table 2.** Comparison of the PANGO lineage and GISAID clade assignment for consensus sequences produced with Oxford Nanopore Technologies and Illumina.

Sample	Oxford Nanopore		Illumina	
	PANGO Lineage	Nextstrain Clade	PANGO Lineage	Nextstrain Clade
IMB1-1/2021	B.1.1.163	20B	B.1.1.163	20B
IMB1-2/2021	B.1.1.163	20B	B.1.1.163	20B
IMB1-5/2021	B.1.1.163	20B	B.1	20A
IMB2-1/2021	B.1.1	20B	B.1.1.7	20I (Alpha, V1)
IMB2-2/2021	B.1.1.163	20B	B.1.1.163	20B

In three cases out of five, the clade and lineage were in agreement between nanopore and Illumina sequencing. IMB2-1/2021 isolate was initially assigned to B.1.1 with nanopore sequencing, while Illumina consensus was identified as B.1.1.7. The analysis of the BAM files for this strain indicated that the amino acid substitutions and deletions characteristic of B.1.1.7 lineage also existed in the nanopore sequencing reads; however, they did not pass the quality check during calling by nanopore variant pipeline (Supplementary Materials Figures S3 and S4). Moreover, the temporal analysis also indicated January as the estimated time for this lineage introduction in Armenia (see Section 3.1). In another case, the Illumina sequence for IMB1-5/2021 was assigned to B.1, while nanopore sequencing assigned the same sequence to B.1.1.163. Overall, our data suggest that Illumina sequencing can produce better consensus sequences than nanopore; the possible reason could be differences in coverage generated in the two approaches and also specific amplicon dropouts described for the ARTIC primer scheme [16,49]. However, nanopore sequencing can serve as an efficient and affordable alternative to Illumina (short-read) next-generation sequencing and be used for the epidemiologic surveillance and molecular-genetic analyses of SARS-CoV-2. This is particularly important in countries with underdeveloped NGS sequencing facilities, such as Armenia.

#### 4. Conclusions

Our study added new data to the global context of genomic epidemiology of SARS-CoV-2 and provided a holistic overview of the emergence, transmission, and diversity of the virus in Armenia. We identified multiple introductions of genomic lineages and their relations with the dynamics of positive cases during 2020–2022. Interestingly, the majority of importations inferred by phylogeographic analyses were through airway travels, while ground transportation played very little or no role, consistent with closed ground borders in Armenia and neighboring countries almost immediately after the first positive case in Armenia [50]. The majority of early importations were from countries with a considerably

large Armenian diaspora (such as Russia and Kazakhstan) as well as touristic destinations (Italy) and much a wider geography for later VOC lineages.

The functional analysis of mutations (both lineage defining and private) identified a considerable number of mutations that affected the binding of predicted viral epitopes to protective HLA loci. Consistent with the previous reports, such mutations were present in the majority of VOC lineages compared to older lineages [51].

Our results also show multiple mutations in regions covered by several primers/probes compared with the reference sequence of the virus. This observation is of particular importance, since mutations may lead to the alteration of the sensitivity of qRT-PCR tests. Diagnostic tests mostly used in Armenia target ORF1ab and N genes, and our results suggest that identified mutations will not influence their accuracy.

The results of the study again emphasize the need for constant sequencing-based surveillance of SARS-CoV-2 strains for public health decision making and health care. Illumina short-read whole-genome sequencing platforms enable accurate sequence determination and are currently the method of choice for SARS-CoV-2 sequencing [52]. However, whole-genome sequencing essential for epidemiological monitoring and surveillance of viral pathogens is still challenging in many countries with limited technical resources. While the superiority of Illumina platforms over nanopore sequencing has been established in several studies [17], the latter still can serve as an efficient and affordable alternative to short-read next-generation sequencing and be used for epidemiologic surveillance and molecular-genetic analyses of SARS-CoV-2. This is particularly important in countries with underdeveloped NGS sequencing facilities, such as Armenia, and can play an important role in shaping local, national, and regional COVID-19 response strategies.

It is also worth noting the limitations of this study. First, the number of genomes sequenced and analyzed was small compared to the total number of positive cases in the country. Moreover, the sample collection and sequencing started in Autumn 2020, which limited our ability to accurately reconstruct events close to the first dates of the epidemic. The geography of trajectories for lineage importations also should be treated with caution since they were inferred from the phylogenetic analyses based on the limited number of background sequences selected. Unfortunately, we did not have access to the travel and contact history information, which definitely would otherwise improve the accuracy of the results.

However, even with these limitations, we believe that this paper is an important contribution in an attempt to fill the knowledge gap and demonstrate the importance of the real-time genomic surveillance of SARS-CoV-2 for informed and timely public health interventions.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/v14051074/s1>, Figure S1: The average sequencing coverage for nanopore sequencing runs. The order of the positions corresponds to the reference genome (MN908947.3); Figure S2: Daily reported deaths from COVID-19 in Armenia, sampling dates, and clade distribution of sequenced samples; Figure S3: B.1.1.7 characteristic mutations in raw nanopore reads of the IMB2-1/2021 isolate; Figure S4: B.1.1.7 characteristic mutations in nanopore and Illumina consensus sequences for the IMB2-1/2021 isolate; Table S1: Accession IDs for SARS-CoV-2 sequenced genomes deposited in the GISAID EpiCoV database; Table S2: The viral RNA load expressed in Ct values performed by qRT-PCR targeting the ORF1ab and N genes in the conserved region of the SARS-CoV-2 genome; Table S3: Sample counts and sequencing scheme; Table S4: Nextclade annotation of the 145 analyzed sequences; Table S5: CorGAT annotation of the 145 analyzed sequences; Table S6: Point mutations in primers and probes for the detection of SARS-CoV-2 of global research institutions; Table S7: Nextclade annotation of nanopore and Illumina paired sequences; Table S8: We gratefully acknowledge the authors from the originating laboratories for obtaining the specimens and the submitting laboratories for data generation and sharing via the GISAID Initiative.

**Author Contributions:** Conceptualization, A.A.; methodology, D.A., R.Z., A.C. and A.A.; replication, S.A.M.B., K.R.J., P.R. and A.L.G.; formal analysis, S.H., M.N. and A.A.; sample sequencing, L.G., G.K., T.S., N.M., V.H., S.A.M.B., K.R.J., P.R., A.L.G. and D.A.; biological samples resources, L.N. (Lyudmila Niazyan), M.D., A.C., G.M.-A. and S.S.; data curation, A.L.G. and A.A.; writing—original draft preparation, D.A., S.H., M.N. and A.A.; writing—review and editing, D.A. and A.A.; visualization, S.H., M.N. and A.A.; supervision, A.A.; project administration, D.A. and A.A.; funding acquisition, R.Z., L.N. (Lilit Nersisyan), A.H. and A.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by SeedingLabs “Instrumental Access” -2017 (A.A.) and -2019 (R.Z.), as well as the Educational-Scientific Center of Excellence for “Genetic engineering, genome editing and 3rd generation sequencing” grant in the frames of the “Competitive Innovation Fund” under “Education Improvement” project supported by World Bank (2019–2021, A.A. and R.Z.), Armenian National Science and Education Fund (NS-molbio-2522 to A.A.) the State Target Program of the Government of the Republic of Armenia under grant agreement № 1-8/20TB project “Creating a Cloud Computing Environment for Solving Scientific and Applied Problems”.

**Institutional Review Board Statement:** The studies involving human participants were reviewed and approved by the Ethics Committee of the Institute of Molecular Biology NAS RA (IRB 00004079, Protocol N3 from 25.06.2020).

**Informed Consent Statement:** Written informed consent was waived by the Ethics Committee since samples arrived at the laboratories fully anonymized.

**Data Availability Statement:** Consensus FASTA files were deposited in the GISAID EpiCoV database (<https://www.gisaid.org/>, accessed on 12 May 2022) (Accessions are available in Supplementary Materials Table S2). Illumina consensus sequences for 5 resequenced samples were deposited in the GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>, accessed on 12 May 2022) (accessions: MZ577122, MZ577123, MZ577124, MZ577125, and MZ577126). Nextstrain configuration files, the auspice JSON file, BEAST output logs, and trees files, and resulting log and tree files, as well as R scripts and data files used in the analyses, are available in Zenodo (<https://doi.org/10.5281/zenodo.6406278>, accessed on 12 May 2022) [53].

**Acknowledgments:** We would like to thank Vicent Pelechano (SciLifeLab, Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Solna, Sweden) for their valuable technical and methodological support. We gratefully acknowledge the authors from the originating laboratories for obtaining the specimens and the submitting laboratories for data generation and sharing via the GISAID Initiative (Supplementary Table S8).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses The Species Severe Acute Respiratory Syndrome-Related Coronavirus: Classifying 2019-NCoV and Naming It SARS-CoV-2. *Nat. Microbiol.* **2020**, *5*, 536–544. [[CrossRef](#)] [[PubMed](#)]
2. Wang, C.; Horby, P.W.; Hayden, F.G.; Gao, G.F. A Novel Coronavirus Outbreak of Global Health Concern. *Lancet* **2020**, *395*, 470–473. [[CrossRef](#)]
3. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.M.; Wang, W.; Song, Z.G.; Hu, Y.; Tao, Z.W.; Tian, J.H.; Pei, Y.Y.; et al. A New Coronavirus Associated with Human Respiratory Disease in China. *Nature* **2020**, *579*, 265–269. [[CrossRef](#)] [[PubMed](#)]
4. Zhou, P.; Yang, X.L.; Wang, X.G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.R.; Zhu, Y.; Li, B.; Huang, C.L.; et al. A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin. *Nature* **2020**, *579*, 270–273. [[CrossRef](#)]
5. van Kasteren, P.B.; van der Veer, B.; van den Brink, S.; Wijsman, L.; de Jonge, J.; van den Brandt, A.; Molenkamp, R.; Reusken, C.B.E.M.; Meijer, A. Comparison of Seven Commercial RT-PCR Diagnostic Kits for COVID-19. *J. Clin. Virol.* **2020**, *128*, 104412. [[CrossRef](#)]
6. Pybus, O.G.; Tatem, A.J.; Lemey, P. Virus Evolution and Transmission in an Ever More Connected World. *Proc. R. Soc. B Biol. Sci.* **2015**, *282*, 20142878. [[CrossRef](#)]
7. Nie, Q.; Li, X.; Chen, W.; Liu, D.; Chen, Y.; Li, H.; Li, D.; Tian, M.; Tan, W.; Zai, J. Phylogenetic and Phylodynamic Analyses of SARS-CoV-2. *Virus Res.* **2020**, *287*, 198098. [[CrossRef](#)]
8. du Plessis, L.; McCrone, J.T.; Zarebski, A.E.; Hill, V.; Ruis, C.; Gutierrez, B.; Raghwani, J.; Ashworth, J.; Colquhoun, R.; Connor, T.R.; et al. Establishment and Lineage Dynamics of the SARS-CoV-2 Epidemic in the UK. *Science* **2021**, *371*, 708–712. [[CrossRef](#)]

9. Nemira, A.; Adeniyi, A.E.; Gasich, E.L.; Bulda, K.Y.; Valentovich, L.N.; Krasko, A.G.; Glebova, O.; Kirpich, A.; Skums, P. SARS-CoV-2 Transmission Dynamics in Belarus in 2020 Revealed by Genomic and Incidence Data Analysis. *Commun. Med.* **2021**, *1*, 31. [[CrossRef](#)]
10. Gutierrez, B.; Márquez, S.; Prado-Vivar, B.; Becerra-Wong, M.; Guadalupe, J.J.; da Silva Candido, D.; Fernandez-Cadena, J.C.; Morey-Leon, G.; Armas-Gonzalez, R.; Andrade-Molina, D.M.; et al. Genomic Epidemiology of SARS-CoV-2 Transmission Lineages in Ecuador. *Virus Evol.* **2021**, *7*, veab051. [[CrossRef](#)]
11. Gankin, Y.; Nemira, A.; Koniukhovskii, V.; Chowell, G.; Weppelmann, T.A.; Skums, P.; Kirpich, A. Investigating the First Stage of the COVID-19 Pandemic in Ukraine Using Epidemiological and Genomic Data. *Infect. Genet. Evol.* **2021**, *95*, 105087. [[CrossRef](#)]
12. Kaleta, T.; Kern, L.; Hong, S.L.; Hölzer, M.; Kochs, G.; Beer, J.; Schnepf, D.; Schwemmle, M.; Bollen, N.; Kolb, P.; et al. Antibody Escape and Global Spread of SARS-CoV-2 Lineage A.27. *Nat. Commun.* **2022**, *13*, 1152. [[CrossRef](#)]
13. Home—Covid. Available online: <https://covid.ncdc.am/en/home> (accessed on 31 March 2022).
14. Elbe, S.; Buckland-Merrett, G. Data, Disease and Diplomacy: GISAID’s Innovative Contribution to Global Health. *Glob. Chall.* **2017**, *1*, 33–46. [[CrossRef](#)] [[PubMed](#)]
15. Shu, Y.; McCauley, J. GISAID: Global Initiative on Sharing All Influenza Data—From Vision to Reality. *Eurosurveillance* **2017**, *22*, 30494. [[CrossRef](#)]
16. Tyson, J.R.; James, P.; Stoddart, D.; Sparks, N.; Wickenhagen, A.; Hall, G.; Choi, J.H.; Lapointe, H.; Kamelian, K.; Smith, A.D.; et al. Improvements to the ARTIC Multiplex PCR Method for SARS-CoV-2 Genome Sequencing Using Nanopore. *bioRxiv* **2020**, *3*, 1. [[CrossRef](#)]
17. Bull, R.A.; Adikari, T.N.; Ferguson, J.M.; Hammond, J.M.; Stevanovski, I.; Beukers, A.G.; Naing, Z.; Yeang, M.; Verich, A.; Gamaarachchi, H.; et al. Analytical Validity of Nanopore Sequencing for Rapid SARS-CoV-2 Genome Analysis. *Nat. Commun.* **2020**, *11*, 6272. [[CrossRef](#)] [[PubMed](#)]
18. Li, J.; Wang, H.; Mao, L.; Yu, H.; Yu, X.; Sun, Z.; Qian, X.; Cheng, S.; Chen, S.; Chen, J.; et al. Rapid Genomic Characterization of SARS-CoV-2 Viruses from Clinical Specimens Using Nanopore Sequencing. *Sci. Rep.* **2020**, *10*, 17492. [[CrossRef](#)]
19. Gohl, D.M.; Garbe, J.; Grady, P.; Daniel, J.; Watson, R.H.B.; Auch, B.; Nelson, A.; Yohe, S.; Beckman, K.B. A Rapid, Cost-Effective Tailed Amplicon Method for Sequencing SARS-CoV-2. *BMC Genom.* **2020**, *21*, 863. [[CrossRef](#)]
20. Artic Network. Available online: <https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html> (accessed on 31 March 2022).
21. Loman, N.J.; Quick, J.; Simpson, J.T. A Complete Bacterial Genome Assembled de Novo Using Only Nanopore Sequencing Data. *Nat. Methods* **2015**, *12*, 733–735. [[CrossRef](#)]
22. Li, H. Minimap2: Pairwise Alignment for Nucleotide Sequences. *Bioinformatics* **2018**, *34*, 3094–3100. [[CrossRef](#)]
23. Addetia, A.; Lin, M.J.; Peddu, V.; Roychoudhury, P.; Jerome, K.R.; Greninger, A.L. Sensitive Recovery of Complete SARS-CoV-2 Genomes from Clinical Samples by Use of Swift Biosciences’ SARS-CoV-2 Multiplex Amplicon Sequencing Panel. *J. Clin. Microbiol.* **2021**, *59*, e02226-20. [[CrossRef](#)] [[PubMed](#)]
24. Lin, M.J.; Rachleff, V.M.; Xie, H.; Shrestha, L.; Lieberman, N.A.P.; Peddu, V.; Addetia, A.; Casto, A.M.; Breit, N.; Mathias, P.C.; et al. Host-Pathogen Dynamics in Longitudinal Clinical Specimens from Patients with COVID-19. *Sci Rep* **2022**, *12*, 5856. [[CrossRef](#)] [[PubMed](#)]
25. Hadfield, J.; Megill, C.; Bell, S.M.; Huddleston, J.; Potter, B.; Callender, C.; Sagulenko, P.; Bedford, T.; Neher, R.A. Nextstrain: Real-Time Tracking of Pathogen Evolution. *Bioinformatics* **2018**, *34*, 4121–4123. [[CrossRef](#)] [[PubMed](#)]
26. Rambaut, A.; Holmes, E.C.; O’Toole, Á.; Hill, V.; McCrone, J.T.; Ruis, C.; du Plessis, L.; Pybus, O.G. A Dynamic Nomenclature Proposal for SARS-CoV-2 Lineages to Assist Genomic Epidemiology. *Nat. Microbiol.* **2020**, *5*, 1403–1407. [[CrossRef](#)]
27. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
28. Katoh, K.; Misawa, K.; Kuma, K.I.; Miyata, T. MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)]
29. Nguyen, L.T.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)] [[PubMed](#)]
30. Sagulenko, P.; Puller, V.; Neher, R.A. TreeTime: Maximum-Likelihood Phylodynamic Analysis. *Virus Evol.* **2018**, *4*, vex042. [[CrossRef](#)]
31. Rambaut, A.; Lam, T.T.; Carvalho, L.M.; Pybus, O.G. Exploring the Temporal Structure of Heterochronous Sequences Using TempEst (Formerly Path-O-Gen). *Virus Evol.* **2016**, *2*, vew007. [[CrossRef](#)] [[PubMed](#)]
32. Drummond, A.J.; Rambaut, A. BEAST: Bayesian Evolutionary Analysis by Sampling Trees. *BMC Evol. Biol.* **2007**, *7*, 214. [[CrossRef](#)]
33. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [[CrossRef](#)] [[PubMed](#)]
34. Heled, J.; Bouckaert, R.R. Looking for Trees in the Forest: Summary Tree from Posterior Samples. *BMC Evol. Biol.* **2013**, *13*, 221. [[CrossRef](#)] [[PubMed](#)]
35. Chiara, M.; Zambelli, F.; Tangaro, M.A.; Mandreoli, P.; Horner, D.S.; Pesole, G. CorGAT: A Tool for the Functional Annotation of SARS-CoV-2 Genomes. *Bioinformatics* **2020**, *36*, 5522–5523. [[CrossRef](#)]
36. Kiyotani, K.; Toyoshima, Y.; Nemoto, K.; Nakamura, Y. Bioinformatic Prediction of Potential T Cell Epitopes for SARS-Cov-2. *J. Hum. Genet.* **2020**, *65*, 569–575. [[CrossRef](#)] [[PubMed](#)]

37. RFE/RL's Armenian Service—Ազատություն ռ/լ. Available online: <https://www.azatutyun.am/a/30462197.html> (accessed on 1 April 2022).
38. Korber, B.; Fischer, W.M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E.E.; Bhattacharya, T.; Foley, B.; et al. Tracking Changes in SARS-CoV-2 Spike: Evidence That D614G Increases Infectivity of the COVID-19 Virus. *Cell* **2020**, *182*, 812–827.e19. [[CrossRef](#)] [[PubMed](#)]
39. COVID-19 Travel Restrictions—The Government of the Republic of Armenia. Available online: <https://www.gov.am/en/covid-travel-restrictions/> (accessed on 1 April 2022).
40. Nyberg, T.; Ferguson, N.M.; Nash, S.G.; Webster, H.H.; Flaxman, S.; Andrews, N.; Hinsley, W.; Bernal, J.L.; Kall, M.; Bhatt, S.; et al. Comparative Analysis of the Risks of Hospitalisation and Death Associated with SARS-CoV-2 Omicron (B.1.1.529) and Delta (B.1.617.2) Variants in England: A Cohort Study. *Lancet* **2022**, *399*, 1303–1312. [[CrossRef](#)]
41. Taylor, C.A.; Whitaker, M.; Anglin, O.; Milucky, J.; Patel, K.; Pham, H.; Chai, S.J.; Alden, N.B.; Yousey-Hindes, K.; Anderson, E.J.; et al. COVID-19–Associated Hospitalizations among Adults During SARS-CoV-2 Delta and Omicron Variant Predominance, by Race/Ethnicity and Vaccination Status—COVID-NET, 14 States, July 2021–January 2022. *Morb. Mortal. Wkly. Rep.* **2022**, *71*, 466–473. [[CrossRef](#)]
42. Migliorini, F.; Torsiello, E.; Spiezia, F.; Oliva, F.; Tingart, M.; Maffulli, N. Association between HLA Genotypes and COVID-19 Susceptibility, Severity and Progression: A Comprehensive Review of the Literature. *Eur. J. Med. Res.* **2021**, *26*, 84. [[CrossRef](#)]
43. Shkurnikov, M.; Nersisyan, S.; Jankevic, T.; Galatenko, A.; Gordeev, I.; Vechorko, V.; Tonevitsky, A. Association of HLA Class I Genotypes With Severity of Coronavirus Disease-19. *Front. Immunol.* **2021**, *12*, 423. [[CrossRef](#)]
44. Douillard, V.; Castelli, E.C.; Mack, S.J.; Hollenbach, J.A.; Gourraud, P.A.; Vince, N.; Limou, S. Current HLA Investigations on SARS-CoV-2 and Perspectives. *Front. Genet.* **2021**, *12*, 774922. [[CrossRef](#)]
45. Pisanti, S.; Deelen, J.; Gallina, A.M.; Caputo, M.; Citro, M.; Abate, M.; Sacchi, N.; Vecchione, C.; Martinelli, R. Correlation of the Two Most Frequent HLA Haplotypes in the Italian Population to the Differential Regional Incidence of COVID-19. *J. Transl. Med.* **2020**, *18*, 84. [[CrossRef](#)] [[PubMed](#)]
46. Bolze, A.; Neveux, I.; Schiabor Barrett, K.M.; White, S.; Isaksson, M.; Dabe, S.; Lee, W.; Grzymiski, J.J.; Washington, N.L.; Cirulli, E.T. HLA-A\*03:01 Is Associated with Increased Risk of Fever, Chills, and Stronger Side Effects from Pfizer-BioNTech COVID-19 Vaccination. *Hum. Genet. Genom. Adv.* **2022**, *3*, 100084. [[CrossRef](#)] [[PubMed](#)]
47. Hovhannisyann, A.; Madelian, V.; Avagyan, S.; Nazaretyan, M.; Hyussyan, A.; Sirunyan, A.; Arakelyan, R.; Manukyan, Z.; Yepiskoposyan, L.; Mayilyan, K.R.; et al. HLA-C\*04:01 Affects HLA Class I Heterozygosity and Predicted Affinity to SARS-CoV-2 Peptides, and in Combination With Age and Sex of Armenian Patients Contributes to COVID-19 Severity. *Front. Immunol.* **2022**, *13*, 769900. [[CrossRef](#)] [[PubMed](#)]
48. Matevosyan, L.; Chattopadhyay, S.; Madelian, V.; Avagyan, S.; Nazaretyan, M.; Hyussian, A.; Vardapetyan, E.; Arutunyan, R.; Jordan, F. HLA-A, HLA-B, and HLA-DRB1 Allele Distribution in a Large Armenian Population Sample. *Tissue Antigens* **2011**, *78*, 21–30. [[CrossRef](#)]
49. Itokawa, K.; Sekizuka, T.; Hashino, M.; Tanaka, R.; Kuroda, M. A Proposal of Alternative Primers for the ARTIC Network's Multiplex PCR to Improve Coverage of SARS-CoV-2 Genome Sequencing. *PLoS One* **2020**, *15*, e0239403. [[CrossRef](#)]
50. Armenia Extends Closure of Border with Iran Over Coronavirus Fears. Available online: <https://www.rferl.org/a/armenia-extends-iran-border-closure-coronavirus-fears/30465130.html> (accessed on 1 April 2022).
51. Thye, A.Y.K.; Law, J.W.F.; Pusparajah, P.; Letchumanan, V.; Chan, K.G.; Lee, L.H. Emerging SARS-CoV-2 Variants of Concern (VOCs): An Impending Global Crisis. *Biomedicines* **2021**, *9*, 1303. [[CrossRef](#)]
52. Muttineni, R.; Kammili, N.; Bingi, T.C.; Raja Rao, M.; Putty, K.; Dholaniya, P.S.; Puli, R.K.; Pakalapati, S.; Doodipala, M.R.; Upadhyay, A.A.; et al. Clinical and Whole Genome Characterization of SARS-CoV-2 in India. *PLoS ONE* **2021**, *16*, e0246173. [[CrossRef](#)]
53. Arakelyan, A.; Avetyan, D. Molecular Genetic Analysis of SARS-CoV-2 Lineages in Armenia—additional data [Data set]. *Zenodo* **2022**. [[CrossRef](#)]