

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Diagnostic Applications for RNA-Seq Technology and Transcriptome Analyses in Human Diseases Caused by RNA Viruses

*Irina Vlasova-St. Louis, Andrew Gorzalski and Mark Pandori*

## Abstract

Human diseases caused by single-stranded, positive-sense RNA viruses, are among the deadliest of the 21st Century. In particular, there are two notable stand-outs: human immunodeficiency virus (HIV) and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Detection of these disease-causing viral transcripts, by next-generation RNA sequencing (RNA-Seq), represents the most immediate opportunity for advances in diagnostic, therapeutic, and preventive applicability in infectious diseases (e.g., AIDS and COVID-19). Moreover, RNA-Seq technologies add significant value to public health studies by first, providing real-time surveillance of known viral strains, and second, by the augmentation of epidemiological databases, construction of annotations and classifications of novel sequence variants. This chapter intends to recapitulate the current knowledge of HIV and SARS-CoV-2 transcriptome architecture, pathogenicity, and some features of the host immune response. Additionally, it provides an overview of recent advances in diagnostic sequencing methodologies and discusses the future challenges and prospects on the utilization of RNA-Seq technologies.

**Keywords:** Next generation RNA Sequencing, RNA-Seq, RNA viruses, human immunodeficiency virus, severe acute respiratory syndrome coronavirus, HIV transcriptome, SARS-CoV-2 transcriptome

## 1. Introduction

Next-generation RNA sequencing is rapidly replacing cDNA microarrays and quantitative PCR in clinics and in the public health laboratories, due to higher sensitivity and precision, as well as its cost-effective ability to identify novel RNA species. High-throughput sequencing has become widely adopted in infectious diseases. Supporting bioinformatics services have improved substantially, which aids in genotyping causative mutations for antimicrobial resistance, pathogens' virulence factors, and global epidemiological surveillance to monitor molecular dynamics of pathogen diversity [1, 2]. Integration of pathogen genome sequencing into infectious disease surveillance was established by United States Centers for Disease Control and Prevention (CDC) under Advanced Molecular Detection program in 2016 [2]. As described below, RNA sequencing became particularly

important to identify newly emerging HIV and SARS-CoV strains. RNA-Seq can be used for many different purposes, such as improve diagnostics, characterization of novel strains, investigation of disease epidemiology, spatiotemporal spread and transmission routes, assessment of evolutionary rates, and the development of countermeasures and public health policies.

The standard practice of viral RNA detection is composed of a two-step procedure - reverse transcription of RNA into cDNA and cDNA amplification. The reverse transcription reaction is prone to recombination errors leading to potential alterations in the amplicon library and inadequate sequence representation of the original sample. Although both the nucleic acid amplification test (NAAT) and real-time reverse transcriptase - polymerase chain reaction (RT-PCR) test are cost-effective, the occurrence of amplicon failures needs to be monitored for substitutions in primer binding sites. In order to circumvent possible mutations in the primer regions, it is recommended multiple primer designs targeting the same genes, when designing RT-PCR assays [3, 4]. The RNA-Seq method is capable to unbiasedly identify variants across thousands of target regions with single-base resolution, in cases where NAAT or qRT-PCR produce false-negative results.

Illumina's technology appears to be at the forefront of viral sequencing, however others such as IonTorrent and Oxford Nanopore technologies are quickly closing the gaps [5]. Illumina deploys sequencing-by-synthesis chemistry, which provides high accuracy and deep coverage of the viral genome. A limitation of Illumina is that short read length (< 400 nt) fragmented sequences should be re-assembled computationally, during which the haplotype information can be lost. Short-read sequencing data requires computational pipelines for trimming of low-quality reads, removal of optical duplicates, detection of reads orientation, and alignment to reconstruct full-length viral sequences. Short read RNA-Seq technologies allow to pull multiple samples per lane to reduce the cost, however, demultiplexing of reads from different samples can be bioinformatically challenging.

The recent shift toward emerging long-read technologies, enabled direct sequencing of individual RNA molecules as opposed to classical indirect approaches [6]. Long-read RNA-seq technologies, such as that provided by Ion Torrent or Oxford Nanopore, allow mapping of novel insertions, deletions, or substitutions in natural variants [7]. Thus, the technology offers advantages over established sequencing technologies, as it has simplified procedures for library preparation and bioinformatics analysis. Although direct RNA sequencing possesses low accuracy, than Illumina's short read method, it enables easier full-length sequencing, rapid viral genome annotation and analysis, which would be particularly useful for understanding changes in transcriptome architectures [8]. Additionally, long-read technologies are more cost-effective, portable, and provide robust reproducibility of the results, when compared to short-read RNA-seq [9]. A heterozygous variants can be detected when filtered by bioinformatic pipelines, which may identify co-infections and estimate the risk of superinfection [10]. Detailed descriptions and of long- and short- read methodologies, as well as analysis workflows, are provided in other chapters of this book.

## **2. Sequencing of human immunodeficiency virus (HIV)**

The human immunodeficiency virus (HIV) is organized in the genus Lentivirus, within the family of Retroviridae, subfamily Orthoretrovirinae. HIV is a human retrovirus with an RNA genome that is composed of copies of a single stranded positive sense RNA [11]. The RNA genome encodes viral enzymes and is packed into nucleocapsid proteins (the viral capsid). After infecting the cell, the viral RNA

genome is reverse transcribed into double stranded viral DNA, which is subsequently integrated into cellular host DNA with the aid of viral and host cofactors. The HIV life cycle has several life-cycle phases; a prolonged latent phase of host genome integration, and a phase of active transcription of new viral RNA in the infected cell [11]. The of HIV replication process is characterized by high rates of nucleotide misincorporations, insertions, deletions, and recombinations. Twentieth century research studies, in 1980s and 1990s, uncovered only a partial viral genome. Determination of the HIV partial genome sequence provided the basis for the development of next-generation sequencing and real-time reverse transcriptase polymerase chain reaction (RT-PCR) methods.

Next-generation RNA-Sequencing enabled the first successful pan-HIV-1 sequencing, including subtype identification and phylodynamic diversification during the course of AIDS pandemic [12]. The current HIV nomenclature includes two types: HIV-1 and HIV-2. The HIV-1 is phylogenetically classified into of 4 groups (M, N, O, and P) [13]. Phylodynamic analyses and evolutionary clock models estimated the time of the most recent common ancestor (HIV-1 pandemic group) to be around the late 19th – early 20th centuries, in Central Africa, and where it may have been associated with an epidemic in primates [14]. The major (M) pandemic group, which causes human-to-human transmission, is classified into 9 subtypes. The most prevalent M subtypes are A, B, C, D, and G; while subtypes F, H, J, L and K are collectively responsible for approximately 1% of all HIV infections [14–16]. An important genetic feature of HIV is that it is prone to recombination. The highest levels of inter-subtype re-combinations are found in HIV-1 infected specimens, from patients in the Sub-Saharan region of Africa. Analysis of a whole-genome HIV-1 sequence, from the Congo Basin, identified ancestral HIV species which clustered basal to all major subtypes; many of which underwent purifying selection and are no longer in circulation. However, 72 additional recombinant forms of HIV remain in circulation [17].

Intra-individual mixtures of recombinant genomes have been reported throughout the world. It is hypothesized that inter-subtype recombinant viruses have an advantage of transmissibility over parental strains [18–24]. To achieve full viral genome sequence coverage, several approaches were employed. One such approach was the amplification of two large fragments (“half genomes”), spanning the full HIV-1 sequence, including all critical regions. Using this approach, molecular surveillance studies of circulating and recombinant forms of HIV, has been conducted in Cameroon. Researchers used primers which target two overlapping “half genome” sequencing approach: the 5’ region (*gag* and *pol* genes), the 3’ region (*env* and *nef*), which overlap the accessory genes (*vif*, *vpr*, and *vpu*) [25]. The two-amplicon approach followed by deep sequencing allowed to characterize several novel circulating recombinant forms, CRFs, (CRF02\_AG, \_AE, \_01A1, and F2, CRF\_36cpx and \_37cpx, etc.) and more than a dozen unique recombinant forms (URF, NYU6541\_6, NYU6556\_3, etc.) that clustered separately from their reference sequences, as determined by the maximum likelihood phylogenetic algorithm [26]. The introduction of circulating recombinant forms (CRF01\_AE and CRF02\_AG) into the Asian-Pacific region from Continental Asia was identified using similar “half genome” sequencing approach [27]. Another approach, the “switching mechanism at the 5’ end of RNA transcript” (SMART), leverages the template-switching capability of certain RT enzymes toward full-length template. During RT reaction, three additional nucleotides are added to the 3’ end of the first cDNA strand of viral RNA template, which serve as an anchor for selective amplification of full length template with a set of 5’ adaptor-ligated primer [28]. This method was adopted to capture full-length HIV sequence in clinical samples, in which viral loads were > 5 log<sub>10</sub> copies/mL [29]. Neighbor-joining phylogenetics trees built from this study data revealed all known viral recombinant species from all four groups

(M, N, P, O), and rare M-subtypes (J and H). A successful strategy has been to study HIV species diversity by embedding a degenerate block of nucleotides, within the primer, during first round of cDNA synthesis, to avoid PCR-related errors (and sequencing errors [30, 31]. A unique ID (Primer ID), created by an incorporated primer tag, can correct allelic skewing during sequencing [32]. Using this approach the authors were able to identify minor variants in the HIV-1 protease gene resulting in multiple alleles that conferred drug-resistance [30].

The percentage of people living with HIV drug resistance appears to be increasing [33]. RNA-seq has been key to revealing so-called viral invasive genes and the acquisition of drug-resistance mutations throughout of the 9.7 Kb of viral genome [34]. Achievement of long amplicon length allowed quantification of viral load and multiple drug-resistant mutations using viral RNA and pro-viral DNA as a template [35–38]. RNA-based long-read sequencing is also being explored for surveillance of viral diversification and assessment of HIV quasispecies [39, 40]. The FDA has cleared Sentosa HIV-1 RNA-Seq genotyping SQ assay for clinical use, to detect and monitor antiretroviral drug resistance mutations in the blood of HIV-infected patients. This product helps clinicians to prescribe effective combination of antiretroviral drugs, such as reverse transcriptase, integrase, envelope (gp41), or CCR5/CXCR4 inhibitors [39]. Phylogenetic clustering of genotyped samples from HIV patients who diagnosed between 1999 and 2012 identified strong associations between molecular clusters and epidemiological hotspots for transmitted drug resistance [41].

### **3. Sequencing of SARS-CoV-2**

Viruses of the order Nidovirales - family Coronaviridae - subfamily Coronavirinae are positive-sense, single-stranded sequences of RNAs; ranging from 26 to 32 kilobases in length [42]. The novel SARS-CoV-2 species was identified via metagenomic RNA sequencing and made publicly available, in NCBI Virus, within two months of unknown pneumonia cases outbreak in China (accession number NC\_045512) [43]. The SARS-CoV-2 viral genome (made up of 29,903 nucleotides) is phylogenetically related to genus beta-coronavirus (subgenus sarbecovirus), which has demonstrated the ability to crossover, from the animal kingdom, and subsequently cause infection in humans [44]. A viral RNA genome encodes several structural proteins (spike, S; nucleocapsid, N; transmembrane, M; envelope, E). SARS-CoV-2 is known to produce 16 non-structural proteins, and at least six accessory proteins (encoded primarily by reading frames ORF1 and ORF2) [45]. Each viral transcript has a 5' cap structure and a 3' poly(A) tail [46]. Genomic characterization of SARS-CoV-2 receptor-binding domains, in conjunction with phylogenetic analysis and homology modeling, revealed the sequence for viral port of entry proteins that conferred human-to-human transmission [47]. Direct RNA sequencing via nanopore and DNA nanoball methodologies identified that SARS-CoV-2 generates a tiered series of canonical and non-canonical subgenomic RNAs, through a process involving homology recombinations between transcriptional regulatory sequences [48]. The function of these transcripts is currently unknown [49].

Sequencing-based genomic surveillance has been employed by UK, USA, and Canadian consortia, to coordinate efforts to sequence large numbers of SARS-CoV-2 genomes. An executive summary “Genomic sequencing of SARS-CoV-2” was issued by the WHO, in January 2021. It discusses the implementation of NGS, for “maximum impact on public health” and provide detailed overview of current sequencing methodologies [50].

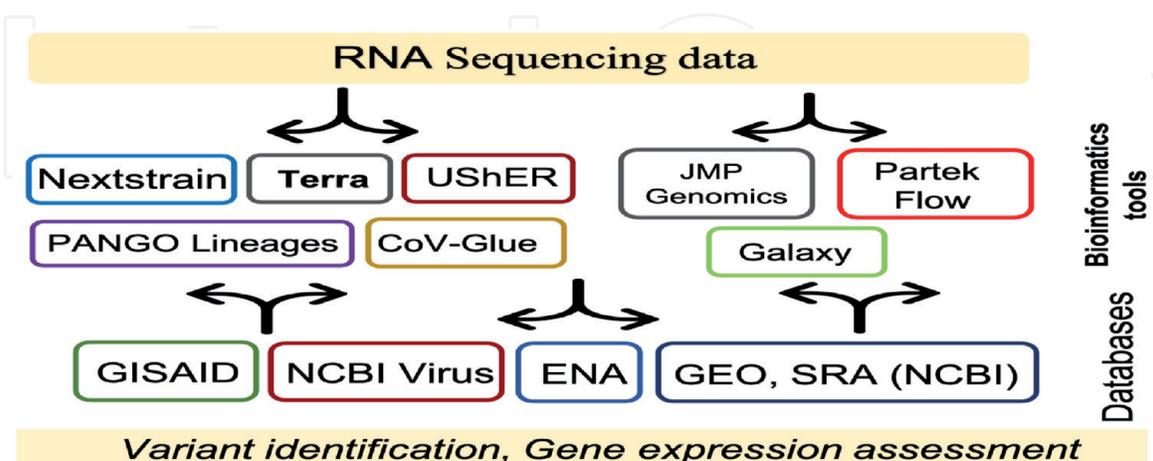
In lockstep with the continuing advances of next-generation sequencing technologies, the bioinformatics community has developed many computational tools

capable of keeping up with big data analysis and interpretation. Data sharing repositories such as GISAID (EpiCoV), NCBI Virus/GenBank, COG-UK, ENA provide free access to sequencing data (**Figure 1**). Bioinformatics workspaces (e.g., Terra [51], UshER [52], NextStrain [53, 54], PANGO Lineages [55], CoV-Glue [56], see **Figure 1**) allow users to assign their own sequences to globally circulating lineages [57].

Genomics researchers who track phylogenetic dynamics of SARS-CoV-2 developed several schemes to describe the accumulation of mutations and detectable divisions within circulating SARS-CoV-2 populations (**Figure 2**). Three popular clade naming conventions are used:

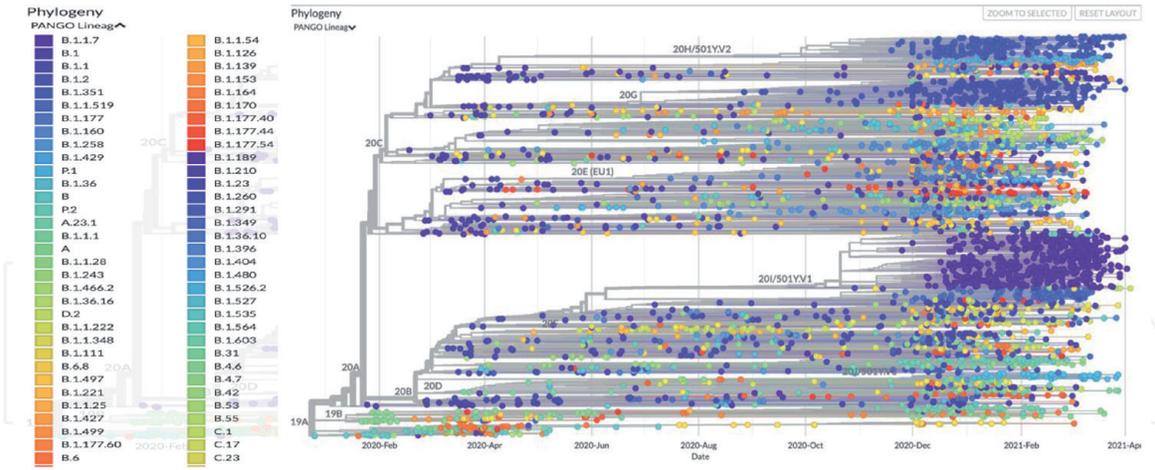
1. PANGO Lineages [55];
2. Clades by NextStrain [58];
3. Clades by GISAID [59].

The PANGO Lineages nomenclature refers to corresponding outbreaks caused by distinct lineages (**Figure 2A**). Lineage A is suggested to be ancestral because it shares more nucleotides similarities with related coronaviruses in animals [60]. Lineages B presumably originated from United Kingdom, and lineages P – from Brazil [61, 62]. The Clades by NextStrain and GISAID nomenclature refers to the divergence of newly emergent clades (see **Figure 2B** and **C**) [57, 59, 63]. A formal naming system for SARS-CoV-2 evolutionary lineages has not been universally adopted. The nomenclatures from PANGO Lineages and GISAID are updated on CDC and WHO web [64, 65]. Novel phylodynamic models for visualization of phylogenomic datasets have been employed in different web formats. For example, Auspice, Augur, NextClade are open-source interactive web tools for visualizing phylogenomic data within NextStrain interface [66]. UshER (Ultrafast Sample placement on Existing tTree) is available in UCSC, and the virus phylogeny is a part of T-BioInfo platform [52]. Novel human coronavirus lineages are now regularly reported, and lineage nomenclature is continually updated [64]. The WHO Virus Evolution Working Group has been working on simplifying the nomenclature, for SARS-CoV-2 variants of interests (VOI) and variants of

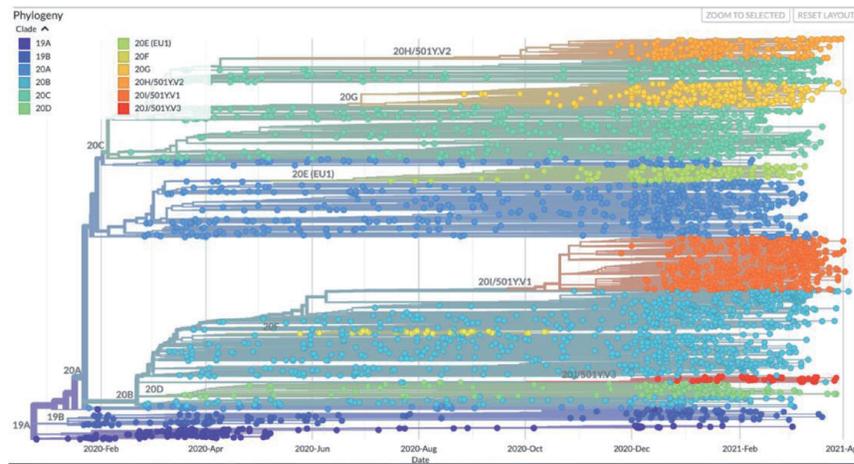


**Figure 1.** Diagnostic applications for RNA-Seq data analyses in human diseases caused by RNA viruses. The upper boxes represent bioinformatics workspaces for data analysis as described in the text below. The lower boxes represent major public repositories/databases commonly used for collecting and sharing genome sequence data (see abbreviations list). Note: Viral sequencing data are deposited in GISAID and NCBI virus; the ENA (EMBL-EBI), GEO, and SRA (NCBI) contain sequences from all domains of life, including human specimens). Double-headed arrows represent interactive nature between bioinformatics tools and data repositories, which allows for variants identification and gene expression assessment.

2A. PANGO Lineages.



2B. Clades by NextStrain.



2C. Clades by GISAID.

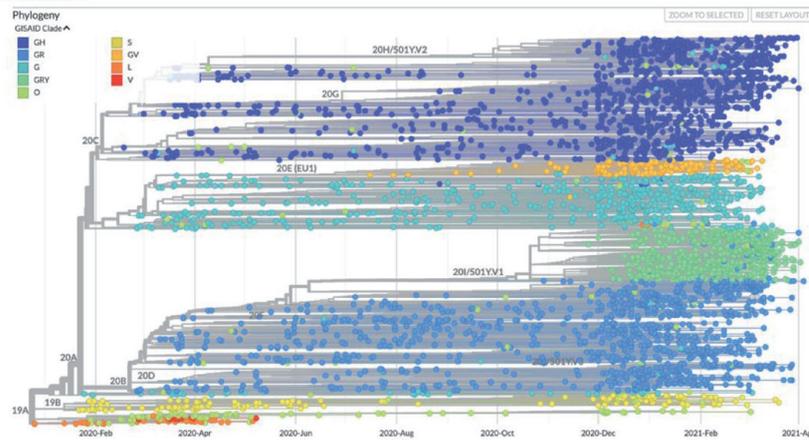


Figure 2.

Genomic epidemiology of novel severe acute respiratory syndrome coronavirus 2 (global subsampling). 2A. PANGO lineages. 2B. Clades by NextStrain. 2C. Clades by GISAID. These phylogenetic trees show time-resolved visualization of evolutionary relationships between SARS-CoV-2 viruses from the ongoing COVID-19 pandemic. Exported from <https://nextstrain.org/ncov/global>. The dots on the trees' branches are color-coded in accordance with phylogeny classification by (2A) PANGO lineages; (2B) by NextStrain clades; (2C) clades by GISAID. Figure legends on the left represent lineages or clades names. Accessed on Apr 20th.

concern (VOC). This group has recommended easy-to-pronounce variant naming based on the Greek alphabet (e.g. Alpha, Beta, Gamma, Delta - for VOC; and Epsilon through Lambda – for VOI) [67].

Genomic epidemiology is a novel discipline that assesses viral genomic diversification, including the acquisition of pathogenic mutations, and molecular dynamics

of a pandemic, at the population level. Real-time viral sequencing helps delineate the origin of imported cases. For example, the assessment of WGS data (in several European countries) enabled a more precise understanding of SARS-CoV-2 transmission patterns, during various phases of the current pandemic. Molecular epidemiology data aided in identifying multiple introduction events, from the community spread in the Netherlands, Brazil, and several other countries [68–71]. WGS analysis leveraged with phylogenetic methodology approximated the source of lineage exportation. For example, the first outbreak of COVID-19 in Taiwan was imported from China. However, the second was from Italy, as strains that caused the second outbreak were phylogenetically more related to GISAID Accession IDs from the Italian outbreak [72]. As determined by the neighbor-joining method using MEGAX software, the circulating Moroccan lineage is more closely related to the South African lineage (20H/501YV2). This is not surprising given that travel within the continent is more predominant than travel between continents [73, 74]. In the United States, epidemiological investigation of sequencing data showed that SARS-CoV-2 varies at the single-nucleotide polymorphism (SNP) resolution, within various States. Additionally, SARS-CoV-2 genomic surveillance identified introduction and community spread of more transmissible strains in the states of California, Ohio, Washington and others (e.g., 20C > 20G clade switch) [75–79]. Genome sequencing in the New York City and Boston areas identified a cryptic spread of multiple simultaneous lineage introductions, from European or Asian travel-related exposures [80–82]. Using Global Epidemic and Mobility Model (GLEAM), Davis et al. estimated the time for the arrival and cryptic phase of the COVID-19 epidemic, which began in early to mid-January on both the East and West coasts of the USA [83]. During the cryptic phase of transmission, monophyletic clades of imported SARS-CoV-2 were spreading until air travel restrictions were dictated [83, 84].

US Center for Disease Controls and Prevention regularly updates the SARS-CoV-2 variant classification, adding novel mutations that have an impact on the immune response or virus virulence factor. Variants are generally classified as a variant of interest, a variant of concern, and a variant of high consequence [64]. Genome-wide sequencing analysis of samples, from various geographic locations, has revealed novel mutations dispersed throughout SARS-CoV-2 genome [85]. Weber et.al has recently detected a number of hotspot mutations, juxtaposing SARS-CoV-2 sequences from various geographic regions, which occurred *de novo*, during the dissemination of infection [60]. These differences in genome variation are not necessarily all pathogenic, but may be important in vaccine design and assessment of immunogenicity [86]. Molecular clock computations estimate that the average evolutionary rate for coronaviruses is roughly  $10^{-4}$  nucleotide substitutions per site, per year, with point mutations that contribute to the diversification of global strains [87, 88]. Although most acquired genetic changes do not substantially affect virulence or transmissibility, those that do cause the corresponding phenotypic changes in SARS-CoV-2 - are of public health importance [67, 85].

#### 4. Advances in diagnostic sequencing methods

Diagnostic sequencing methods have reached a state of capability, which enables the detection of low quantities of viral RNA, with greater sensitivity and specificity. Several amplicon construction approaches have been utilized: targeted capture-based, tiling hexamer-based, and standard amplicon-based [89–91]. A newly commercialized methodology that includes the hybridization capture method yielded near-complete genome coverage, even in samples with relatively low viral loads (cycle thresholds, Ct > 25) [92]. Several studies reported that 93 to 99% of genome

coverage is the samples with Ct values ranging from 26 to 32. The median on-target percentage of genome coverage dropped below 50%, in higher Ct value samples (Ct > 30) [93]. Public Health Laboratories have recently utilized capture-based methods followed by sequencing to evaluate if SARS-CoV-2 is present in wastewater samples. The goal of the study was the detection the meta-transcriptomic differences between SARS-CoV-2 variants, found in wastewater, and compare the differences to locally reported clinical genotypes [77]. An Illumina Flex for Enrichment kit and Illumina Respiratory Virus Oligo Panel were used to enrich the samples for virus cDNA amplicon.

The tiling amplicon-based approach has been adapted for SARS-CoV-2 sequencing. It is based on modified reverse transcription (RT with designed primer pool) and an overlapping long amplicon multiplex PCR strategy [94]. Primers can be designed in a random-hexamer-primed RT fashion, and a not-so-random-hexamer-primed fashion, to eliminate human rRNA during cDNA synthesis. The number of primers can reach up to 400, in metagenomics protocols, to improve genome coverage. However, less than 20 primers are used more widely for amplicon construction [95–97]. The tiling amplicon-based approach has been validated by many laboratories and adapted for nanopore sequencing technologies as well as Illumina and Sanger sequencing [97–99]. A Swift Biosciences' Normalase® tiling amplicon SARS-CoV-2 panel achieved 80% success of partial genome recovery from samples with a Ct value between 32 and 34, and 40% - for samples with a Ct value between 34 and 36 [100]. A 98 non-overlapping primer pair set have been designed by a group of scientists from ARTIC network in early March of 2020. This primer set has been modified on 3 occasions, reducing primer-dimers formation during RT-PCR (a major cause of coverage bias) [101]. Recently, the ARTIC's CoronaHiT platform (Coronavirus High Throughput) has been launched and provided accurate consensus SARS-CoV-2 genomes when at least 20x coverage is obtained. It is important to note that in the low-titer samples, with Ct > 32 (< 100 viral genome copies/uL), sequencing results became less reliable for all sequencing methods (long- and short- read) [96, 102–104].

## **5. Assessment of host response by RNA-Seq**

The host-pathogen interactions have been studied by RNA-Seq in research settings. Transcriptome analysis of hosts' diseased tissues has been an integral part of the discovery of immune response biomarkers. University of Minnesota Division of Infectious Diseases (USA) leads intensive transcriptomic studies for discoveries of novel biomarkers of immune reconstitution inflammatory syndrome (IRIS) in the HIV-infected population [105, 106]. IRIS presents as an exaggerated immune reaction (in a form of cytokine release syndrome) that occurs in immunocompromised patients who have commenced antiretroviral treatments (ART) [106]. Several predictive and diagnostic biomarkers have already been identified in peripheral blood, utilizing Galaxy, JMP Genomics, or Partek Flow software (**Figure 1**). For example, the low expression of type I interferon, interferon-response genes, and components of antiviral defense pathways were identified as a risk factor for subsequent occurrence of non-fatal forms of IRIS [107, 108]. More recently, transcriptomic predictors of fatal IRIS have been delineated. These include the overexpressed genes that encode numerous proinflammatory cytokines (e.g., IL6), chemokines, stress response kinase signaling and production of reactive oxygen species in monocytes pathways [109]. Interestingly, the deficiency in innate antiviral defense genes, and poly-immuno-cytopenia have also been identified as significant contributors to subsequent cytokine release syndrome during COVID-19 [110–112]. The reduced IFN I and IFN III type gene expression and elevated monocyte- and granulocyte- derived chemokines at an early stage of

COVID-19 have been proposed as predictors of subsequent cytokine release syndrome and disease severity [113].

In a small percentage of the human population, the SARS-CoV-2 infection results in a critical illness that begins with uncontrolled overexpression of proinflammatory cytokines (the so-called “cytokine storm”) [114]. A cytokine storm may lead to pathophysiological events such as activation of the coagulation cascades, systemic oxidative stress, and various forms of cell death [115]. Clinically, these events manifest as an acute respiratory distress syndrome (ARDS), or death from multiple organ failure [116, 117]. SARS-CoV-2 is recognized by innate immune cells and, in cases of ARDS, elicits highly abundant transcriptional gene expression, which is somewhat comparable to that of immunocompromised patients who have developed fatal IRIS [109, 118, 119]. The upregulation of inflammasome, Toll-like receptor signaling, HMGB1, and oxidative stress response transcripts, have been recently demonstrated as biomarkers of fatal immune reconstitution inflammatory syndrome [109]. Additionally, NF- $\kappa$ B-associated inflammatory genes and strong neutrophil activation/degranulation signatures differentiate fatal IRIS events from those of survivors. Likewise, fatal ARDS in COVID-19 patients have been characterized by maladapted hyperactivation of the same innate inflammatory pathways described for fatal IRIS [120–123]. Systemic upregulation of cytokines TNF $\alpha$ , IL6, IL1B, and NLRP3 are noteworthy of mention as they are targetable molecules for drugs such as adalimumab, tocilizumab, anakinra, and RAPA-501-Allo, respectively [124–127].

SARS-CoV-2 has the propensity to selectively induce mortality in elderly population (over 65 years of age) and in immunocompromised individuals [128, 129]. The most reasonable explanation for this propensity is the depletion of the thymic population, and the inability to mount adaptive T cell immune responses in timely manner [130]. Innate immune cells become the first line of defense against SARS-CoV-2, however in the absence of proper regulatory feedback – are thrown into an uncontrolled proinflammatory loop. Similarly, the irreparable thymic damage and a waning adaptive immunity in AIDS patients enable the sustained HIV replication, which may explain why serious complications such as fatal IRIS are hinged on responses driven by innate immune cells.

There have been isolated cases of SARS-CoV-2 reinfection in individuals who had no known immune disorders [131–133]. Since SARS-CoV-2 has not diversified significantly, these patients have been re-infected with variants from the same or adjacent clade [132]. Studies have revealed a number of SARS-CoV-2 epitopes are recognized by CD8 $^{+}$  T cells in COVID-19 convalescent subjects (as determined by TruSight HLA sequencing panel) [134]. It is hypothesized that re-infected individuals are unable to mount a sufficient response to primary infection, which result in secondary COVID-19 manifestation. Many of these patients had tested positive for anti-SARS-CoV-2 antibodies after the reinfection, but it is unknown whether they developed antibodies after the first infection at all. It has been hypothesized that susceptibility to re-infection is genetically driven [135]. Several hypothetical models had been proposed [136]. Thus, simultaneous RNA-seq of host and pathogen during viral infection would be helpful to characterize the molecular responses in case of COVID-19 reinfection.

There are now several vaccine types that are being used, in an attempt to reach herd immunity, in various countries [137]. The most successful vaccines are mRNA-based, which have an estimated 92% efficacy [138, 139]. Still, there is an accumulating body of evidence of the so-called ‘vaccine breakthrough’ phenomenon, when SARS-CoV-2 infection occurs in persons who are fully vaccinated [140, 141]. The immune gene variants that contribute to poor T cell memory and undelaying vaccine breakthroughs are not yet known [142]. As vaccination increases, RNA-Seq

may become a useful tool in exploring the consequence of vaccine dose-sparing strategies, the emergence of novel vaccine-escape variants, intra- host variant diversity, as well as in understanding the phenomenon of vaccine breakthrough.

## **6. Conclusion and future directions**

RNA-seq has become an indispensable diagnostic tool of clinically important RNA viruses. It has taken the research community almost 20 years to sequence a complete HIV viral genome, but it took a little over 2 months to sequence a near-complete SARS-CoV-2 genome and identify its origin! With an improved understanding of the genomic structures of virus populations, RNA-seq can be used to track the origins, genetic diversity, and global monitoring of transmission patterns. In a public health context, RNA-Seq is an indispensable tool to be utilized in the assessment of risks and the emergence of future outbreaks. Additionally, assessment of transcriptomic profiles, during host responses to infection improved the understanding of disease pathogenesis and identify patients at risk for severe outcomes. To overcome drug resistance, novel viral genotype-guided agents based on antisense RNAs (e.g., aptamers, peptide nucleic acids oligomers, ribozymes, and RNAi silencing therapies) are being developed to specifically inhibit viral replication. Complex bioinformatics approaches enabled vaccines to be designed with high probabilities of intercepting viral escape [143]. Much work awaits the global scientific community, including the establishment of unified reference standards and the adoption of a single nomenclature for SARS-CoV-2. Additionally, the assessment of long-term vaccination efficacy by RNA-Seq in the clinical laboratory is necessary to fully understand its benefits [144].

## **Abbreviations**

SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
CDC	Center for Disease Controls and Prevention
UCSC	University of California, Santa Cruz
GISAID	Global Initiative on Sharing Avian Flu Data
ENA	European Nucleotide Archive
COG-UK	COVID-19 Genomics UK Consortium
WGS	whole genome sequencing
SNP	single-nucleotide polymorphism
NCBI	The National Center for Biotechnology Information
UShER	Ultrafast Sample placement on Existing tRee
IRIS	immune reconstitution inflammatory syndrome
TNFA	Tumor Necrosis Factor Alfa
HMGB1	High Mobility Group Box 1 gene
NLRP3	NLR (nucleotide-binding domain and leucine-rich repeat containing) Family Pyrin Domain Containing 3
EMBL-EBI	European Molecular Biology Laboratory - European Bioinformatics Institute

IntechOpen

### **Author details**

Irina Vlasova-St. Louis<sup>1,2\*</sup>, Andrew Gorzalski<sup>2</sup> and Mark Pandori<sup>2</sup>

1 Department of Medicine, University of Minnesota, Minneapolis, MN, USA

2 Nevada State Public Health Laboratory, Reno, NV, USA

\*Address all correspondence to: [irinastl@umn.edu](mailto:irinastl@umn.edu) and [istlouis@med.unr.edu](mailto:istlouis@med.unr.edu)

### **IntechOpen**

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Avery, M.; Mills, S.J.; Stephan, E. Real-time monitoring through the use of technology to enhance performances throughout HIV cascades. *Curr. Opin. HIV AIDS* 2017, 12. 10.1097/COH.0000000000000397.
- [2] Canters for Disease Control and Prevention Advanced molecular detection (AMD) Available online: <https://www.cdc.gov/amd/index.html> (accessed on Mar 16, 2021).
- [3] Vogels, C.B.F.; Brito, A.F.; Wyllie, A.L.; Fauver, J.R.; Ott, I.M.; Kalinich, C.C.; Petrone, M.E.; Casanovas-Massana, A.; Catherine Muenker, M.; Moore, A.J.; et al. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT-qPCR primer-probe sets. *Nat. Microbiol.* 2020, 5. 10.1038/s41564-020-0761-6.
- [4] Bustin, S.A.; Nolan, T. RT-QPCR testing of SARS-COV-2: A primer. *Int. J. Mol. Sci.* 2020, 21. 10.3390/ijms21083004.
- [5] 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. 10.1038/s41467-020-20075-6.
- [6] Viehweger, A.; Krautwurst, S.; Lamkiewicz, K.; Madhugiri, R.; Ziebuhr, J.; Hölzer, M.; Marz, M. Direct RNA nanopore sequencing of full-length coronavirus genomes provides novel insights into structural variants and enables modification analysis. *Genome Res.* 2019, 29. 10.1101/gr.247064.118.
- [7] Lee, E.R.; Parkin, N.; Jennings, C.; Brumme, C.J.; Enns, E.; Casadellà, M.; Howison, M.; Coetzer, M.; Avila-Rios, S.; Capina, R.; et al. Performance comparison of next generation sequencing analysis pipelines for HIV-1 drug resistance testing. *Sci. Rep.* 2020, 10. 10.1038/s41598-020-58544-z.
- [8] Gwinn, M.; Maccannell, D.; Armstrong, G.L. Next-Generation Sequencing of Infectious Pathogens. *JAMA - J. Am. Med. Assoc.* 2019, 321. 10.1001/jama.2018.21669.
- [9] Helmy, Y.A.; Fawzy, M.; Elasad, A.; Sobieh, A.; Kenney, S.P.; Shehata, A.A. The COVID-19 Pandemic: A Comprehensive Review of Taxonomy, Genetics, Epidemiology, Diagnosis, Treatment, and Control. *J. Clin. Med.* 2020, 9. 10.3390/jcm9041225.
- [10] Sjaarda, C.P.; Rustom, N.; Evans, G.A.; Huang, D.; Perez-Patrigeon, S.; Hudson, M.L.; Wong, H.; Sun, Z.; Guan, T.H.; Ayub, M.; et al. Phylogenomics reveals viral sources, transmission, and potential superinfection in early-stage COVID-19 patients in Ontario, Canada. *Sci. Rep.* 2021, 11. 10.1038/s41598-021-83355-1.
- [11] Seitz, R. Human Immunodeficiency Virus (HIV). *Transfus. Med. Hemotherapy* 2004, 31. 10.1159/000078043.
- [12] Etienne, L.; Delaporte, E.; Peeters, M. Origin and Emergence of HIV/AIDS. In *Genetics and Evolution of Infectious Diseases*; 2011.10.1016/B978-0-12-384890-1.00026-1.
- [13] Castro-Nallar, E.; Pérez-Losada, M.; Burton, G.F.; Crandall, K.A. The evolution of HIV: Inferences using phylogenetics. *Mol. Phylogenet. Evol.* 2012, 62. 10.1016/j.ympev.2011.11.019.
- [14] Gryseels, S.; Watts, T.D.; Mpolesha, J.M.K.; Larsen, B.B.; Lemey, P.; Muyembe-Tamfum, J.J.; Teuwen, D.E.; Worobey, M. A near full-length HIV-1 genome from 1966 recovered from formalin-fixed paraffin-embedded tissue. *Proc. Natl. Acad. Sci. U. S. A.* 2020, 117. 10.1073/pnas.1913682117.
- [15] LANL HIV database Available online: <https://www.hiv.lanl.gov/content/index> (accessed on Mar 3, 2021).

- [16] Souza, J.S.M.; Silva Júnior, J.J.; Brites, C.; Monteiro-Cunha, J.P. Molecular and geographic characterization of hiv-1 bf recombinant viruses. *Virus Res.* **2019**, *270*. 10.1016/j.virusres.2019.197650.
- [17] Rodgers, M.A.; Wilkinson, E.; Vallari, A.; McArthur, C.; Sthreshley, L.; Brennan, C.A.; Cloherty, G.; de Oliveira, T. Sensitive Next-Generation Sequencing Method Reveals Deep Genetic Diversity of HIV-1 in the Democratic Republic of the Congo. *J. Virol.* **2017**, *91*. 10.1128/jvi.01841-16.
- [18] Sierra, M.; Thomson, M.M.; Ríos, M.; Casado, G.; Ojea-de Castro, R.; Delgado, E.; Echevarría, G.; Muñoz, M.; Colomina, J.; Carmona, R.; et al. The analysis of near full-length genome sequences of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from Chile, Venezuela and Spain reveals their relationship to diverse lineages of recombinant viruses related to CRF12\_BF. *Infect. Genet. Evol.* **2005**, *5*. 10.1016/j.meegid.2004.07.010.
- [19] Bello, G.; Aulicino, P.C.; Ruchansky, D.; Guimarães, M.L.; Lopez-Galindez, C.; Casado, C.; Chiparelli, H.; Rocco, C.; Mangano, A.; Sen, L.; et al. Phylogenetics of HIV-1 Circulating Recombinant Forms 12\_BF and 38\_BF in Argentina and Uruguay. *Retrovirology* **2010**, *7*. 10.1186/1742-4690-7-22.
- [20] Ruchansky, D.; Casado, C.; Russi, J.C.; Arbiza, J.R.; Lopez-Galindez, C. Identification of a new HIV Type 1 circulating recombinant form (CRF38-BF1) in Uruguay. *AIDS Res. Hum. Retroviruses* **2009**, *25*. 10.1089/aid.2008.0248.
- [21] Shcherbakova, N.S.; Shalamova, L.A.; Delgado, E.; Fernández-García, A.; Vega, Y.; Karpenko, L.I.; Ilyichev, A.A.; Sokolov, Y. V.; Shcherbakov, D.N.; Pérez-Álvarez, L.; et al. Short communication: Molecular epidemiology, phylogeny, and phylodynamics of CRF63-02A1, a recently originated HIV-1 circulating recombinant form spreading in Siberia. *AIDS Res. Hum. Retroviruses* **2014**, *30*. 10.1089/aid.2014.0075.
- [22] Takebe, Y.; Liao, H.; Hase, S.; Uenishi, R.; Li, Y.; Li, X.J.; Han, X.; Shang, H.; Kamarulzaman, A.; Yamamoto, N.; et al. Reconstructing the epidemic history of HIV-1 circulating recombinant forms CRF07\_BC and CRF08\_BC in East Asia: The relevance of genetic diversity and phylodynamics for vaccine strategies. *Vaccine* **2010**, *28*. 10.1016/j.vaccine.2009.07.101.
- [23] Pessôa, R.; Loureiro, P.; Lopes, M.E.; Carneiro-Proietti, A.B.F.; Sabino, E.C.; Busch, M.P.; Sanabani, S.S. Ultra-deep sequencing of HIV-1 near full-length and partial proviral genomes reveals high genetic diversity among Brazilian blood donors. *PLoS One* **2016**, *11*. 10.1371/journal.pone.0152499.
- [24] Cañada, J.E.; Delgado, E.; Gil, H.; Sánchez, M.; Benito, S.; García-Bodas, E.; Gómez-González, C.; Canut-Blasco, A.; Portu-Zapirain, J.; Sáez de Adana, E.; et al. Identification of a New HIV-1 BC Intersubtype Circulating Recombinant Form (CRF108\_BC) in Spain. *Viruses* **2021**, *13*. 10.3390/v13010093.
- [25] Banin, A.N.; Tuen, M.; Bimela, J.S.; Tongo, M.; Zappile, P.; Khodadadi-Jamayran, A.; Nanfack, A.J.; Meli, J.; Wang, X.; Mbanya, D.; et al. Development of a versatile, near full genome amplification and sequencing approach for a broad variety of HIV-1 group M variants. *Viruses* **2019**, *11*. 10.3390/v11040317.
- [26] Banin, A.N.; Tuen, M.; Bimela, J.S.; Tongo, M.; Zappile, P.; Khodadadi-Jamayran, A.; Nanfack, A.J.; Okonko, I.O.; Meli, J.; Wang, X.; et al. Near full genome characterization of HIV-1 unique recombinant forms in Cameroon reveals dominant CRF02\_AG and F2 recombination patterns. **2019**. 10.1002/jia2.25362/full.

- [27] Chen, Y.; Hora, B.; DeMarco, T.; Berba, R.; Register, H.; Hood, S.; Carter, M.; Stone, M.; Pappas, A.; Sanchez, A.M.; et al. Increased predominance of HIV-1 CRF01\_AE and its recombinants in the Philippines. *J. Gen. Virol.* **2019**, *100*. 10.1099/jgv.0.001198.
- [28] Zhu, Y.Y.; Machleder, E.M.; Chenchik, A.; Li, R.; Siebert, P.D. Reverse transcriptase template switching: A SMART™ approach for full-length cDNA library construction. *Biotechniques* **2001**, *30*. 10.2144/01304pf02.
- [29] Berg, M.G.; Yamaguchi, J.; Alessandri-Gradt, E.; Tell, R.W.; Plantier, J.C.; Brennan, C.A. A pan-HIV strategy for complete genome sequencing. *J. Clin. Microbiol.* **2016**, *54*. 10.1128/JCM.02479-15.
- [30] Jabara, C.B.; Jones, C.D.; Roach, J.; Anderson, J.A.; Swanstrom, R. Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*. 10.1073/pnas.1110064108.
- [31] Boltz, V.F.; Rausch, J.; Shao, W.; Coomer, C.; Mellors, J.W.; Kearney, M.F.; Coffin, J.M. Analysis of Resistance Haplotypes Using Primer IDs and Next Gen Sequencing of HIV RNA Methods Available online: <http://www.croiconference.org/sites/default/files/posters-2015/593.pdf> (accessed on Jul 19, 2015).
- [32] Dennis, A.M.; Zhou, S.; Sellers, C.J.; Learner, E.; Potempa, M.; Cohen, M.S.; Miller, W.C.; Eron, J.J.; Swanstrom, R. Using primer-ID deep sequencing to detect recent human immunodeficiency virus type 1 infection. *J. Infect. Dis.* **2018**, *218*. 10.1093/infdis/jiy426.
- [33] Porter, D.P.; Daeumer, M.; Thielen, A.; Chang, S.; Martin, R.; Cohen, C.; Miller, M.D.; White, K.L. Emergent HIV-1 drug resistance mutations were not present at low-frequency at baseline in non-nucleoside reverse transcriptase inhibitor-treated subjects in the STaR study. *Viruses* **2015**, *7*. 10.3390/v7122943.
- [34] Ekici, H.; Rao, S.D.; Sönnnerborg, A.; Ramprasad, V.L.; Gupta, R.; Neogi, U. Cost-efficient HIV-1 drug resistance surveillance using multiplexed high-throughput amplicon sequencing: Implications for use in low- and middle-income countries. *J. Antimicrob. Chemother.* **2014**, *69*. 10.1093/jac/dku278.
- [35] Fogel, J.M.; Bonsall, D.; Cummings, V.; Bowden, R.; Golubchik, T.; De Cesare, M.; Wilson, E.A.; Gamble, T.; Del Rio, C.; Batey, D.S.; et al. Performance of a high-throughput next-generation sequencing method for analysis of HIV drug resistance and viral load. *J. Antimicrob. Chemother.* **2020**, *75*. 10.1093/jac/dkaa352.
- [36] Novitsky, V.; Zahralban-Steele, M.; McLane, M.F.; Moyo, S.; Van Widenfelt, E.; Gaseitsiwe, S.; Makhema, J.; Essex, M. Long-range HIV genotyping using viral RNA and proviral DNA for analysis of HIV drug resistance and HIV clustering. *J. Clin. Microbiol.* **2015**, *53*. 10.1128/JCM.00756-15.
- [37] Gall, A.; Ferns, B.; Morris, C.; Watson, S.; Cotten, M.; Robinson, M.; Berry, N.; Pillay, D.; Kellam, P. Universal amplification, next-generation sequencing, and assembly of HIV-1 genomes. *J. Clin. Microbiol.* **2012**, *50*. 10.1128/JCM.01516-12.
- [38] Taylor, T.; Lee, E.R.; Nykoluk, M.; Enns, E.; Liang, B.; Capina, R.; Gauthier, M.K.; Domselaar, G. Van; Sandstrom, P.; Brooks, J.; et al. A MiSeq-HyDRA platform for enhanced HIV drug resistance genotyping and surveillance. *Sci. Rep.* **2019**, *9*. 10.1038/s41598-019-45328-3.
- [39] Sidhu, G.; Schuster, L.; Liu, L.; Tamashiro, R.; Li, E.; Langaee, T.; Wagner, R.; Wang, G.P. A single variant sequencing method for sensitive and quantitative detection of HIV-1

minority variants. *Sci. Rep.* **2020**, *10*.  
10.1038/s41598-020-65085-y.

[40] Kijak, G.H.; Sanders-Buell, E.; Harbolick, E.A.; Pham, P.; Chenine, A.L.; Eller, L.A.; Rono, K.; Robb, M.L.; Michael, N.L.; Kim, J.H.; et al. Targeted deep sequencing of HIV-1 using the IonTorrentPGM platform. *J. Virol. Methods* **2014**, *205*. 10.1016/j.jviromet.2014.04.017.

[41] Chung, Y.S.; Choi, J.Y.; Yoo, M.S.; Seong, J.H.; Choi, B.S.; Kang, C. Phylogenetic transmission clusters among newly diagnosed antiretroviral drug-naïve patients with human immunodeficiency virus-1 in Korea: A study from 1999 to 2012. *PLoS One* **2019**, *14*. 10.1371/journal.pone.0217817.

[42] Pal, M.; Berhanu, G.; Desalegn, C.; Kandi, V. Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update. *Cureus* **2020**. 10.7759/cureus.7423.

[43] NCBI Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome Available online: <https://www.ncbi.nlm.nih.gov/nuccore/1798174254> (accessed on Apr 18, 2021).

[44] Gorbalenya, A.E.; Baker, S.C.; Baric, R.S.; de Groot, R.J.; Drosten, C.; Gulyaeva, A.A.; Haagmans, B.L.; Lauber, C.; Leontovich, A.M.; Neuman, B.W.; et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **2020**, *5*. 10.1038/s41564-020-0695-z.

[45] Michel, C.J.; Mayer, C.; Poch, O.; Thompson, J.D. Characterization of accessory genes in coronavirus genomes. *Virol. J.* **2020**, *17*. 10.1186/s12985-020-01402-1.

[46] de Breyne, S.; Vindry, C.; Guillin, O.; Condé, L.; Mure, F.; Gruffat, H.; Chavatte, L.; Ohlmann, T. Translational

control of coronaviruses. *Nucleic Acids Res.* **2020**, *48*. 10.1093/nar/gkaa1116.

[47] Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **2020**, *395*. 10.1016/S0140-6736(20)30251-8.

[48] Kim, D.; Lee, J.Y.; Yang, J.S.; Kim, J.W.; Kim, V.N.; Chang, H. The Architecture of SARS-CoV-2 Transcriptome. *Cell* **2020**, *181*. 10.1016/j.cell.2020.04.011.

[49] Nomburg, J.; Meyerson, M.; DeCaprio, J.A. Pervasive generation of non-canonical subgenomic RNAs by SARS-CoV-2. *Genome Med.* **2020**, *12*. 10.1186/s13073-020-00802-w.

[50] World Health Organization Genomic sequencing of SARS-CoV-2. A guide to implementation for maximum impact on public health Available online: <https://apps.who.int/iris/bitstream/handle/10665/338480/9789240018440-eng.pdf?sequence=1&isAllowed=y>.

[51] Terra Terra.bio Available online: <https://app.terra.bio/> (accessed on May 5, 2021).

[52] University of California Santa Cruz UShER: Ultrafast Sample placement on Existing tRee Available online: <https://genome.ucsc.edu/cgi-bin/hgPhyloPlace> (accessed on Mar 21, 2021).

[53] 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*. 10.1093/bioinformatics/bty407.

[54] Nextstrain Genomic epidemiology of novel coronavirus - Global subsampling Available online: <https://nextstrain.org/ncov/global?l=radial>.

- [55] PANGO PANGO lineages Available online: <https://cov-lineages.org/> (accessed on Mar 16, 2021).
- [56] Singer, J.; Gifford, R.; Cotten, M.; Robertson, D. CoV-GLUE: A Web Application for Tracking SARS-CoV-2 Genomic Variation. *Preprints* **2020**.
- [57] 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*. 10.1038/s41564-020-0770-5.
- [58] Nextstrain Nextstrain SARS-CoV-2 resources Available online: <https://nextstrain.org/sars-cov-2> (accessed on Mar 28, 2021).
- [59] GISAID Clade and lineage nomenclature Available online: <https://www.gisaid.org/references/statements-clarifications/clade-and-lineage-nomenclature-aids-in-genomic-epidemiology-of-active-hcov-19-viruses/#:~:text=Clade definitions in GISAID are augmented with more detailed lineages,of the pandemic strain causing> (accessed on Mar 28, 2021).
- [60] Zhou, P.; Yang, X. Lou; 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. 10.1038/s41586-020-2012-7.
- [61] Virology.org Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations.
- [62] virology.org Spike E484K mutation in the first SARS-CoV-2 reinfection case confirmed in Brazil, 2020 Available online: <https://virological.org/t/spike-e484k-mutation-in-the-first-sars-cov-2-reinfection-case-confirmed-in-brazil-2020/584> (accessed on Mar 29, 2021).
- [63] Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K.S.M.; Lau, E.H.Y.; Wong, J.Y.; et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* **2020**, *382*. 10.1056/nejmoa2001316.
- [64] Centers for Disease Control and Prevention SARS-CoV-2 Variants Available online: <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance/variant-info.html> (accessed on Mar 18, 2021).
- [65] World Health Organization WHO | SARS-CoV-2 Variants. *Who* **2020**.
- [66] Nextstrain Auspice 2.24.1 Available online: <https://docs.nextstrain.org/projects/auspice/en/latest/index.html#> (accessed on Mar 25, 2021).
- [67] WHO Tracking SARS-CoV-2 variants Available online: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/> (accessed on Jun 17, 2021).
- [68] Oude Munnink, B.B.; Nieuwenhuijse, D.F.; Stein, M.; O'Toole, Á.; Haverkate, M.; Mollers, M.; Kamga, S.K.; Schapendonk, C.; Pronk, M.; Lexmond, P.; et al. Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. *Nat. Med.* **2020**, *26*. 10.1038/s41591-020-0997-y.
- [69] Toovey, O.T.R.; Harvey, K.N.; Bird, P.W.; Tang, J.W.-T.W.-T. Introduction of Brazilian SARS-CoV-2 484K.V2 related variants into the UK. *J. Infect.* **2021**. 10.1016/j.jinf.2021.01.025.
- [70] Volz, E.; Hill, V.; McCrone, J.T.; Price, A.; Jorgensen, D.; O'Toole, Á.; Southgate, J.; Johnson, R.; Jackson, B.; Nascimento, F.F.; et al. Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. *Cell* **2021**, *184*. 10.1016/j.cell.2020.11.020.

- [71] Alteri, C.; Cento, V.; Piralla, A.; Costabile, V.; Tallarita, M.; Colagrossi, L.; Renica, S.; Giardina, F.; Novazzi, F.; Gaiarsa, S.; et al. Genomic epidemiology of SARS-CoV-2 reveals multiple lineages and early spread of SARS-CoV-2 infections in Lombardy, Italy. *Nat. Commun.* **2021**, *12*. 10.1038/s41467-020-20688-x.
- [72] Liu, J.Y.; Chen, T.J.; Hwang, S.J. Analysis of imported cases of covid-19 in taiwan: A nationwide study. *Int. J. Environ. Res. Public Health* **2020**, *17*. 10.3390/ijerph17093311.
- [73] Lemriss, S.; Souiri, A.; Amar, N.; Lemzaoui, N.; Mestoui, O.; Labioui, M.; Ouaraiba, N.; Jibjibe, A.; Yartaoui, M.; Chahmi, M.; et al. Complete Genome Sequence of a 2019 Novel Coronavirus (SARS-CoV-2) Strain Causing a COVID-19 Case in Morocco. *Microbiol. Resour. Announc.* **2020**, *9*. 10.1128/mra.00633-20.
- [74] Pearson, C.A.; Russell, T.W.; Davies, N.G.; Kucharski, A.J.; CMMID COVID-19 working group; Edmunds, W.J.; Eggo, R.M. Estimates of severity and transmissibility of novel South Africa SARS-CoV-2 variant 501YV2. *Preprint* **2021**, *50*.
- [75] Zhang, W.; Davis, B.D.; Chen, S.S.; Sincuir Martinez, J.M.; Plummer, J.T.; Vail, E. Emergence of a Novel SARS-CoV-2 Variant in Southern California. *JAMA* **2021**. 10.1001/jama.2021.1612.
- [76] Zhang, W.; Govindavari, J.P.; Davis, B.; Chen, S.; Kim, J.T.; Song, J.; Lopategui, J.; Plummer, J.T.; Vail, E. Analysis of SARS-CoV-2 genomes from southern California reveals community transmission pathways in the early stage of the US COVID-19 pandemic. *medRxiv* **2020**. 10.1101/2020.06.12.20129999.
- [77] Crits-Christoph, A.; Kantor, R.S.; Olm, M.R.; Whitney, O.N.; Al-Shayeb, B.; Lou, Y.C.; Flamholz, A.; Kennedy, L.C.; Greenwald, H.; Hinkle, A.; et al. Genome sequencing of sewage detects regionally prevalent SARS-CoV-2 variants. *MBio* **2021**, *12*. 10.1128/mBio.02703-20.
- [78] Bedford, T.; Greninger, A.L.; Roychoudhury, P.; Starita, L.M.; Famulare, M.; Huang, M.L.; Nalla, A.; Pepper, G.; Reinhardt, A.; Xie, H.; et al. Cryptic transmission of SARS-CoV-2 in Washington state. *Science (80-. )*. **2020**, *370*. 10.1126/SCIENCE.ABC0523.
- [79] Deng, X.; Gu, W.; Federman, S.; du Plessis, L.; Pybus, O.G.; Faria, N.R.; Wang, C.; Yu, G.; Bushnell, B.; Pan, C.Y.; et al. Genomic surveillance reveals multiple introductions of SARS-CoV-2 into Northern California. *Science (80-. )*. **2020**, *369*. 10.1126/science.abb9263.
- [80] Maurano, M.T.; Ramaswami, S.; Zappile, P.; Dimartino, D.; Boytard, L.; Ribeiro-Dos-Santos, A.M.; Vulpescu, N.A.; Westby, G.; Shen, G.; Feng, X.; et al. Sequencing identifies multiple early introductions of SARS-CoV-2 to the New York City region. *Genome Res.* **2020**, *31*. 10.1101/gr.266676.120.
- [81] Gonzalez-Reiche, A.S.; Hernandez, M.M.; Sullivan, M.J.; Ciferri, B.; Alshammari, H.; Obla, A.; Fabre, S.; Kleiner, G.; Polanco, J.; Khan, Z.; et al. Introductions and early spread of SARS-CoV-2 in the New York City area. *Science (80-. )*. **2020**, *369*. 10.1126/science.abc1917.
- [82] Lemieux, J.E.; Siddle, K.J.; Shaw, B.M.; Loreth, C.; Schaffner, S.F.; Gladden-Young, A.; Adams, G.; Fink, T.; Tomkins-Tinch, C.H.; Krasilnikova, L.A.; et al. Phylogenetic analysis of SARS-CoV-2 in Boston highlights the impact of superspreading events. *Science (80-. )*. **2021**, *371*. 10.1126/science.abe3261.
- [83] Davis, J.T.; Chinazzi, M.; Perra, N.; Mu, K.; Piontti, A.P. y.; Ajelli, M.; Dean, N.E.; Gioannini, C.; Litvinova, M.; Merler, S.; et al. Estimating the establishment of local transmission and the cryptic phase of the COVID-19

pandemic in the USA. *medRxiv* 2020. 10.1101/2020.07.06.20140285.

[84] Deng, X.; Gu, W.; Federman, S.; du Plessis, L.; Pybus, O.G.; Faria, N.; Wang, C.; Yu, G.; Pan, C.Y.; Guevara, H.; et al. A genomic survey of SARS-CoV-2 reveals multiple introductions into Northern California without a predominant lineage. *medRxiv* 2020. 10.1101/2020.03.27.20044925.

[85] Islam, M.R.; Hoque, M.N.; Rahman, M.S.; Alam, A.S.M.R.U.; Akther, M.; Puspo, J.A.; Akter, S.; Sultana, M.; Crandall, K.A.; Hossain, M.A. Genome-wide analysis of SARS-CoV-2 virus strains circulating worldwide implicates heterogeneity. *Sci. Rep.* **2020**, *10*. 10.1038/s41598-020-70812-6.

[86] Weber, S.; Ramirez, C.; Doerfler, W. Signal hotspot mutations in SARS-CoV-2 genomes evolve as the virus spreads and actively replicates in different parts of the world. *Virus Res.* **2020**, *289*. 10.1016/j.virusres.2020.198170.

[87] van Dorp, L.; Acman, M.; Richard, D.; Shaw, L.P.; Ford, C.E.; Ormond, L.; Owen, C.J.; Pang, J.; Tan, C.C.S.; Boshier, F.A.T.; et al. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. *Infect. Genet. Evol.* **2020**, *83*. 10.1016/j.meegid.2020.104351.

[88] Liu, Q.; Zhao, S.; Shi, C.M.; Song, S.; Zhu, S.; Su, Y.; Zhao, W.; Li, M.; Bao, Y.; Xue, Y.; et al. Population Genetics of SARS-CoV-2: Disentangling Effects of Sampling Bias and Infection Clusters. *Genomics, Proteomics Bioinforma.* **2020**. 10.1016/j.gpb.2020.06.001.

[89] Jayamohan, H.; Lambert, C.J.; Sant, H.J.; Jafek, A.; Patel, D.; Feng, H.; Beeman, M.; Mahmood, T.; Nze, U.; Gale, B.K. SARS-CoV-2 pandemic: a review of molecular diagnostic tools including sample collection and commercial response with associated advantages and limitations. *Anal. Bioanal. Chem.* **2021**, *413*. 10.1007/s00216-020-02958-1.

[90] Gorzynski, J.E.; de Jong, H.N.; Amar, D.; Hughes, C.R.; Ioannidis, A.; Bierman, R.; Liu, D.; Tanigawa, Y.; Kistler, A.L.; Kamm, J.; et al. High-throughput SARS-CoV-2 and host genome sequencing from single nasopharyngeal swabs. *medRxiv* 2020. 10.1101/2020.07.27.20163147.

[91] Nasir, J.A.; Kozak, R.A.; Aftanas, P.; Raphenya, A.R.; Smith, K.M.; Maguire, F.; Maan, H.; Alruwaili, M.; Banerjee, A.; Mbareche, H.; et al. A comparison of whole genome sequencing of sars-cov-2 using amplicon-based sequencing, random hexamers, and bait capture. *Viruses* **2020**, *12*. 10.3390/v12080895.

[92] Charre, C.; Ginevra, C.; Sabatier, M.; Regue, H.; Destras, G.; Brun, S.; Burfin, G.; Scholtes, C.; Morfin, F.; Valette, M.; et al. Evaluation of NGS-based approaches for SARS-CoV-2 whole genome characterisation. *Virus Evol.* **2020**, *6*. 10.1093/ve/veaa075.

[93] 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 Genomics* **2020**, *21*. 10.1186/s12864-020-07283-6.

[94] Alessandrini, F.; Caucci, S.; Onofri, V.; Melchionda, F.; Tagliabracci, A.; Bagnarelli, P.; Sante, L. Di; Turchi, C.; Menzo, S. Evaluation of the ion ampliseq SARS-CoV-2 research panel by massive parallel sequencing. *Genes (Basel)*. **2020**, *11*. 10.3390/genes11080929.

[95] Guo, L.; Boocock, J.; Tome, J.M.; Chandrasekaran, S.; Hilt, E.E.; Zhang, Y.; Sathe, L.; Li, X.; Luo, C.; Kosuri, S.; et al. Rapid cost-effective viral genome sequencing by V-seq. *bioRxiv* 2020. 10.1101/2020.08.15.252510.

[96] Resende, P.C.; Motta, F.C.; Roy, S.; Appolinario, L.; Fabri, A.; Xavier, J.; Harris, K.; Matos, A.R.; Caetano, B.; Orgeswalska, M.; et al. SARS-CoV-2 genomes recovered by long amplicon

tiling multiplex approach using nanopore sequencing and applicable to other sequencing platforms. *bioRxiv* 2020. 10.1101/2020.04.30.069039.

[97] 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. *bioRxiv* 2020.

[98] 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. 10.1038/s41598-020-74656-y.

[99] Hourdel, V.; Kwasiborski, A.; Balière, C.; Matheus, S.; Batéjat, C.F.; Manuguerra, J.C.; Vanhomwegen, J.; Caro, V. Rapid Genomic Characterization of SARS-CoV-2 by Direct Amplicon-Based Sequencing Through Comparison of MinION and Illumina iSeq100TM System. *Front. Microbiol.* 2020, 11. 10.3389/fmicb.2020.571328.

[100] Fontenele, R.S.; Kraberger, S.; Hadfield, J.; Driver, E.M.; Bowes, D.; Holland, L.A.; Faleye, T.O.C.; Adhikari, S.; Kumar, R.; Inchausti, R.; et al. High-throughput sequencing of SARS-CoV-2 in wastewater provides insights into circulating variants. *medRxiv Prepr. Serv. Heal. Sci.* 2021. 10.1101/2021.01.22.21250320.

[101] Itokawa, K.; Sekizuka, T.; Hashino, M.; Tanaka, R.; Kuroda, M. Disentangling primer interactions improves SARS-CoV-2 genome sequencing by multiplex tiling PCR. *PLoS One* 2020, 15. 10.1371/journal.pone.0239403.

[102] Baker, D.J.; Aydin, A.; Le-Viet, T.; Kay, G.L.; Rudder, S.; de Oliveira Martins, L.; Tedim, A.P.; Kolyva, A.; Diaz, M.; Alikhan, N.F.; et al. CoronaHiT: high-throughput sequencing

of SARS-CoV-2 genomes. *Genome Med.* 2021, 13. 10.1186/s13073-021-00839-5.

[103] 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. 10.1101/2020.09.04.283077.

[104] Wen, S.; Sun, C.; Zheng, H.; Wang, L.; Zhang, H.; Zou, L.; Liu, Z.; Du, P.; Xu, X.; Liang, L.; et al. High-coverage SARS-CoV-2 genome sequences acquired by target capture sequencing. *J. Med. Virol.* 2020, 92, 2221-2226. 10.1002/jmv.26116.

[105] Boulware, D.R.; Meya, D.B.; Bergemann, T.L.; Williams, D.; Irina, I.A.; Rhein, J.; Staddon, J.; Kambugu, A.; Janoff, E.N.; Bohjanen, P.R. Antiretroviral therapy down-regulates innate antiviral response genes in patients with AIDS in sub-Saharan Africa. *J. Acquir. Immune Defic. Syndr.* 2010. 10.1097/QAI.0b013e3181ef4963.

[106] Brienze, V.M.S.; André, J.C.; Liso, E.; Vlasova-St. Louis, I. Cryptococcal immune reconstitution inflammatory syndrome: From blood and cerebrospinal fluid biomarkers to treatment approaches. *Life* 2021, 11. 10.3390/life11020095.

[107] Vlasova-St. Louis, I.; Chang, C.C.; Shahid, S.; French, M.A.; Bohjanen, P.R. Transcriptomic predictors of paradoxical cryptococcosis-associated immune reconstitution inflammatory syndrome. *Open Forum Infect. Dis.* 2018, 5, 1-10. 10.1093/ofid/ofy157.

[108] Mohei, Hesham; Kellampalli, Usha; Vlasova-St. Louis, I.; Vlasova-St. Louis, I. Immune Reconstitution Disorders: Spotlight on Interferons. *Int. J. Biomed. Investig.* 2019, 2, 1-21. 10.31531/2581-4745.1000119.

- [109] Vlasova-St Louis, I.; Musubire, A.K.; Meya, D.B.; Nabeta, H.W.; Mohei, H.; Boulware, D.R.; Bohjanen, P.R. Transcriptomic biomarker pathways associated with death in HIV-infected patients with cryptococcal meningitis. *BMC Med. Genomics* **2021**, *14*. 10.1186/s12920-021-00914-1.
- [110] Zhou, Y.; Fu, B.; Zheng, X.; Wang, D.; Zhao, C.; Qi, Y.; Sun, R.; Tian, Z.; Xu, X.; Wei, H. Pathogenic T-cells and inflammatory monocytes incite inflammatory storms in severe COVID-19 patients. *Natl. Sci. Rev.* **2020**, *7*. 10.1093/nsr/nwaa041.
- [111] Sawalha, A.H.; Zhao, M.; Coit, P.; Lu, Q. Epigenetic dysregulation of ACE2 and interferon-regulated genes might suggest increased COVID-19 susceptibility and severity in lupus patients. *Clin. Immunol.* **2020**, *215*. 10.1016/j.clim.2020.108410.
- [112] Chen, G.; Wu, D.; Guo, W.; Cao, Y.; Huang, D.; Wang, H.; Wang, T.; Zhang, X.; Chen, H.; Yu, H.; et al. Clinical and immunological features of severe and moderate coronavirus disease 2019. *J. Clin. Invest.* **2020**, *130*. 10.1172/JCI137244.
- [113] Guo, C.; Li, B.; Ma, H.; Wang, X.; Cai, P.; Yu, Q.; Zhu, L.; Jin, L.; Jiang, C.; Fang, J.; et al. Single-cell analysis of two severe COVID-19 patients reveals a monocyte-associated and tocilizumab-responding cytokine storm. *Nat. Commun.* **2020**, *11*. 10.1038/s41467-020-17834-w.
- [114] Leisman, D.E.; Ronner, L.; Pinotti, R.; Taylor, M.D.; Sinha, P.; Calfee, C.S.; Hirayama, A. V.; Mastroiani, F.; Turtle, C.J.; Harhay, M.O.; et al. Cytokine elevation in severe and critical COVID-19: a rapid systematic review, meta-analysis, and comparison with other inflammatory syndromes. *Lancet Respir. Med.* **2020**, *8*. 10.1016/S2213-2600(20)30404-5.
- [115] Crimi, E.; Benincasa, G.; Figueroa-Marrero, N.; Galdiero, M.; Napoli, C. Epigenetic susceptibility to severe respiratory viral infections and its therapeutic implications: a narrative review. *Br. J. Anaesth.* **2020**, *125*. 10.1016/j.bja.2020.06.060.
- [116] Saksena, N.; Bonam, S.R.; Miranda-Saksena, M. Epigenetic Lens to Visualize the Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) Infection in COVID-19 Pandemic. *Front. Genet.* **2021**, *12*. 10.3389/fgene.2021.581726.
- [117] Herrmann, J.; Adam, E.H.; Notz, Q.; Helmer, P.; Sonntagbauer, M.; Ungemach-Papenberg, P.; Sanns, A.; Zausig, Y.; Steinfeldt, T.; Torje, I.; et al. COVID-19 Induced Acute Respiratory Distress Syndrome—A Multicenter Observational Study. *Front. Med.* **2020**, *7*. 10.3389/fmed.2020.599533.
- [118] Aschenbrenner, A.C.; Mouktaroudi, M.; Krämer, B.; Antonakos, N.; Oestreich, M.; Gkizeli, K.; Nuesch-Germano, M.; Saridaki, M.; Bonaguro, L.; Reusch, N.; et al. Disease severity-specific neutrophil signatures in blood transcriptomes stratify COVID-19 patients. *medRxiv* **2020**. 10.1101/2020.07.07.20148395.
- [119] Cheng, L.C.; Kao, T.J.; Phan, N.N.; Chiao, C.C.; Yen, M.C.; Chen, C.F.; Hung, J.H.; Jiang, J.Z.; Sun, Z.; Wang, C.Y.; et al. Novel signaling pathways regulate SARS-CoV and SARS-CoV-2 infectious disease. *Medicine (Baltimore)*. **2021**, *100*. 10.1097/MD.00000000000024321.
- [120] Schultze, J.L.; Aschenbrenner, A.C. COVID-19 and the human innate immune system. *Cell* **2021**. 10.1016/j.cell.2021.02.029.
- [121] Shah, A. Novel Coronavirus-Induced NLRP3 Inflammasome Activation: A Potential Drug Target in the Treatment of COVID-19. *Front. Immunol.* **2020**, *11*. 10.3389/fimmu.2020.01021.
- [122] van den Berg, D.F.; te Velde, A.A. Severe COVID-19: NLRP3 Inflammasome

Dysregulated. *Front. Immunol.* 2020, 11, 10.3389/fimmu.2020.01580.

[123] Haghjooy Javanmard, S.; Vaseghi, G.; Manteghinejad, A.; Nasirian, M. Neutrophil-to-Lymphocyte ratio as a potential biomarker for disease severity in COVID-19 patients. *J. Glob. Antimicrob. Resist.* 2020, 22, 862-863. 10.1016/j.jgar.2020.07.029.

[124] Gianfrancesco, M.; Hyrich, K.L.; Hyrich, K.L.; Al-Adely, S.; Al-Adely, S.; Carmona, L.; Danila, M.I.; Gossec, L.; Gossec, L.; Izadi, Z.; et al. Characteristics associated with hospitalisation for COVID-19 in people with rheumatic disease: Data from the COVID-19 Global Rheumatology Alliance physician-reported registry. *Ann. Rheum. Dis.* 2020, 79. 10.1136/annrheumdis-2020-217871.

[125] Sang, E.R.; Tian, Y.; Miller, L.C.; Sang, Y. Epigenetic evolution of ACE2 and IL-6 genes: Non-canonical interferon-stimulated genes correlate to COVID-19 susceptibility in vertebrates. *Genes (Basel).* 2021, 12. 10.3390/genes12020154.

[126] Yaqinuddin, A.; Kashir, J. Novel therapeutic targets for SARS-CoV-2-induced acute lung injury: Targeting a potential IL-1 $\beta$ /neutrophil extracellular traps feedback loop. *Med. Hypotheses* 2020, 143. 10.1016/j.mehy.2020.109906.

[127] NCT04482699 RAPA-501-Allo Off-the-Shelf Therapy of COVID-19. <https://clinicaltrials.gov/show/NCT04482699> 2020.

[128] Blanco-Melo, D.; Nilsson-Payant, B.E.; Liu, W.C.; Uhl, S.; Hoagland, D.; Møller, R.; Jordan, T.X.; Oishi, K.; Panis, M.; Sachs, D.; et al. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* 2020, 181. 10.1016/j.cell.2020.04.026.

[129] Angioni, R.; Sánchez-Rodríguez, R.; Munari, F.; Bertoldi, N.; Arcidiacono,

D.; Cavinato, S.; Marturano, D.; Zaramella, A.; Realdon, S.; Cattelan, A.; et al. Age-severity matched cytokine profiling reveals specific signatures in Covid-19 patients. *Cell Death Dis.* 2020, 11. 10.1038/s41419-020-03151-z.

[130] Van Elslande, J.; Vermeersch, P.; Vandervoort, K.; Wawina-Bokalanga, T.; Vanmechelen, B.; Wollants, E.; Laenen, L.; André, E.; Van Ranst, M.; Lagrou, K.; et al. Symptomatic Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Reinfection by a Phylogenetically Distinct Strain. *Clin. Infect. Dis.* 2020. 10.1093/cid/ciaa1330.

[131] Tillett, R.L.; Sevinsky, J.R.; Hartley, P.D.; Kerwin, H.; Crawford, N.; Gorzalski, A.; Laverdure, C.; Verma, S.C.; Rossetto, C.C.; Jackson, D.; et al. Genomic evidence for reinfection with SARS-CoV-2: a case study. *Lancet Infect. Dis.* 2021, 21, 52-58. 10.1016/S1473-3099(20)30764-7.

[132] K.K.-W., T.; I.F.-N., H.; J.D., I.; A.W.-H., C.; W.-M., C.; A.R., T.; C.H.-Y., F.; S., Y.; H.-W., T.; A.C.-K., N.; et al. COVID-19 re-infection by a phylogenetically distinct SARS-coronavirus-2 strain confirmed by whole genome sequencing. *Clin. Infect. Dis.* 2020.

[133] Ak, R.; Yilmaz, E.; Seyhan, A.U.; Doganay, F. Recurrence of COVID-19 documented with RT-PCR. *J. Coll. Physicians Surg. Pakistan* 2021, 31. 10.29271/jcsp.2021.Supp1.S26.

[134] Kared, H.; Redd, A.D.; Bloch, E.M.; Bonny, T.S.; Sumatoh, H.R.; Kairi, F.; Carbajo, D.; Abel, B.; Newell, E.W.; Bettinotti, M.; et al. SARS-CoV-2-specific CD8<sup>+</sup> T cell responses in convalescent COVID-19 individuals. *J. Clin. Invest.* 2021. 10.1172/jci145476.

[135] Dearlove, B.; Lewitus, E.; Bai, H.; Li, Y.; Reeves, D.B.; Joyce, M.G.; Scott, P.T.; Amare, M.F.; Vasan, S.; Michael, N.L.; et al. A SARS-CoV-2 vaccine candidate would likely match all

currently circulating variants. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*. 10.1073/pnas.2008281117.

[136] Lipsitch, M.; Grad, Y.H.; Sette, A.; Crotty, S. Cross-reactive memory T cells and herd immunity to SARS-CoV-2. *Nat. Rev. Immunol.* **2020**, *20*. 10.1038/s41577-020-00460-4.

[137] Williams, T.C.; Burgers, W.A. SARS-CoV-2 evolution and vaccines: cause for concern? *Lancet Respir. Med.* **2021**. 10.1016/S2213-2600(21)00075-8.

[138] The Lancet Respiratory Medicine Realising the potential of SARS-CoV-2 vaccines—a long shot? *Lancet Respir. Med.* **2021**, *9*. 10.1016/S2213-2600(21)00045-X.

[139] Hagen Ashley; M.S SARS-CoV-2 Variants vs. Vaccines.

[140] Keehner, J.; Horton, L.E.; Pfeiffer, M.A.; Longhurst, C.A.; Schooley, R.T.; Currier, J.S.; Abeles, S.R.; Torriani, F.J. SARS-CoV-2 Infection after Vaccination in Health Care Workers in California. *N. Engl. J. Med.* **2021**. 10.1056/NEJMc2101927.

[141] Kim, J.H.; Marks, F.; Clemens, J.D. Looking beyond COVID-19 vaccine phase 3 trials. *Nat. Med.* **2021**, *27*. 10.1038/s41591-021-01230-y.

[142] Fergie, J.; Srivastava, A. Immunity to SARS-CoV-2: Lessons Learned. *Front. Immunol.* **2021**, *12*. 10.3389/fimmu.2021.654165.

[143] Priyanka; Choudhary, O.P.; Singh, I. Evolution of SARS-CoV-2: A prediction on the lineages and vaccine effectiveness. *Travel Med. Infect. Dis.* **2021**, *40*. 10.1016/j.tmaid.2021.101983.

[144] Leguia, M.; Vila-Sanjurjo, A.; Chain, P.S.G.; Berry, I.M.; Jarman, R.G.; Pollett, S. Precision Medicine and

Precision Public Health in the Era of Pathogen Next-Generation Sequencing. *J. Infect. Dis.* **2020**, *221*. 10.1093/infdis/jiz424.