viruses

MDPI

Article

Intra-Host Evolution of SARS-CoV-2 During Persistent Infection of Pediatric COVID-19 Patients

Charlie R. Boyle [1,2], Tien Doan [2,3], Estefany Rios-Guzman [2,3], Jessica Maciuch [4], Lacy M. Simons [2,3], Dulce S. Garcia [2,3], David B. Williams [5], Arghavan Alisoltani [2,3], Egon A. Ozer [2,3], Ramon Lorenzo-Redondo [2,3] and Judd F. Hultquist [2,3,*]

1 Division of Pediatric Infectious Diseases, Anne and Robert H. Lurie Children's Hospital, Chicago, IL 60611, USA; cboyle@luriechildrens.org
2 Center for Pathogen Genomics and Microbial Evolution, Northwestern University Havey Institute for Global Health, Chicago, IL 60611, USA; buihanhtien.doan@northwestern.edu (T.D.); estefany.guzman@northwestern.edu (E.R.-G.); lacy.simons@northwestern.edu (L.M.S.); dulce.garcia@northwestern.edu (D.S.G.); e-ozer@northwestern.edu (E.A.O.); ramon.lorenzo@northwestern.edu (R.L.-R.)
3 Division of Infectious Diseases, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA
4 Division of Allergy and Immunology, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA; jessica.maciuch@northwestern.edu
5 Stanley Manne Children's Research Institute, Ann & Robert H. Lurie Children's Hospital of Chicago, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA
* Correspondence: judd.hultquist@northwestern.edu

Abstract

The Coronavirus disease 2019 (COVID-19) pandemic had a profound global impact, yet children exhibited distinct clinical and epidemiological patterns compared to adults. Pediatric cases of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) were generally characterized by milder disease, lower hospitalization rates, and few long-term sequelae. However, a subset of children developed severe complications such as multisystem inflammatory syndrome in children (MIS-C), highlighting the heterogeneity in disease presentation. Differences in immune system maturity and comorbidities likely contribute to the age-dependent manifestation of SARS-CoV-2 and other respiratory viruses. Persistent SARS-CoV-2 infection, particularly in immunocompromised individuals, has been implicated in the emergence of new viral variants with immune escape characteristics due to ongoing viral replication in the presence of selective pressure. While SARS-CoV-2 evolution in persistently infected adults has been well-documented, it is less clear how the virus evolves during persistent infection in the pediatric population. To address this question, we performed viral whole genome sequencing of longitudinal specimens collected from immunocompetent and immunocompromised pediatric COVID-19 patients. Similarly to what has been observed in adult cohorts, mutations associated with enhanced viral fitness and immune escape arose intra-host over time. Intra-host diversity accumulated at similar rates in immunocompetent and immunocompromised children, though more mutations overall were observed in the immunocompromised cohort due to the longer infection time courses. Overall, we identified similar viral evolutionary trends over the course of infection despite clinical differences in pediatric COVID-19 manifestation and severity. This similarity suggests that persistent infection in children may be an additional, but not unique, source of ongoing viral diversification.

check for updates

Academic Editor: Ahmed El-Shamy

Received: 22 August 2025
Revised: 17 September 2025
Accepted: 22 September 2025
Published: 28 September 2025

Citation: Boyle, C.R.; Doan, T.; Rios-Guzman, E.; Maciuch, J.; Simons, L.M.; Garcia, D.S.; Williams, D.B.; Alisoltani, A.; Ozer, E.A.; Lorenzo-Redondo, R.; et al. Intra-Host Evolution of SARS-CoV-2 During Persistent Infection of Pediatric COVID-19 Patients. *Viruses* **2025**, *17*, 1313. https://doi.org/10.3390/v17101313

Copyright: © 2025 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/).

Keywords: Coronavirus disease 2019; COVID-19; severe acute respiratory syndrome coronavirus; SARS-CoV-2; viral evolution; longitudinal cohort; pediatric; intra-host diversity

1. Introduction

The Coronavirus disease 2019 (COVID-19) pandemic had profound global socioeconomical and cultural impacts, but its overall clinical impact on pediatric populations was muted in comparison to adults [1,2]. Most severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in children are asymptomatic to mild, though more severe sequelae may develop in children with underlying health conditions [1–9]. In rare cases, pediatric COVID-19 patients may develop multisystem inflammatory syndrome in children (MIS-C), which can ultimately lead to cardiovascular compromise and shock [10,11]. Nonetheless, clinical metrics of disease severity, hospitalization rates, admission to the intensive care unit (ICU), and rates of long-term sequalae are all decreased in pediatric relative to adult populations [1,6,7]. This pattern of milder disease in children is not unique to SARS-CoV-2, as several respiratory pathogens are known to manifest differently in children verses adults [12,13]. These differences in disease manifestation are thought to be driven by underlying factors such as prior exposure, immune system maturity, and the frequency and nature of other underlying health conditions [14–20].

Immunocompromised or immunosuppressed people are at particularly high risk for developing severe COVID-19 due to an inability to clear the infection [14,18–21]. SARS-CoV-2 can persist in these patients for several months, enabling the intra-host evolution of viral variants with resistance to therapeutic pressures. Longitudinal sequencing of SARS-CoV-2 in adult COVID-19 patients with persistent infection has revealed the intra-host emergence of several Spike mutations that confer resistance to neutralizing antibodies and monoclonal antibody therapeutics, such as E340K and E484K [21–27]. Several of these mutations are shared with those that emerged in assorted variants of concern (VOCs), leading to the theory that these VOCs may have originally evolved in immunocompromised patients with persistent infection [14,16–19,25,28–31]. More recently, longitudinal sequencing in immunocompromised patients has identified the intra-host emergence of resistance mutations to administered antiviral therapies [32–35], though these have not yet emerged in widely circulating strains in the community.

While several studies have explored the evolution of SARS-CoV-2 in adults with persistent COVID-19, much less is known about how the virus evolves and diversifies in a pediatric population. We hypothesized that the unique immune environment of younger pediatric patients would enable greater viral diversification compared to teenagers and adults with fewer differences observed between immunocompromised and immunocompetent individuals. In this single-center, retrospective cohort study, we performed longitudinal viral whole genome sequencing in serially collected specimens from 16 pediatric patients with COVID-19. We assessed viral divergence and diversification over time and in response to administered therapeutics. Similarly to findings in adult cohorts, we observed the intra-host emergence of mutations associated with enhanced viral fitness and immune escape over time. Although rates of intra-host diversity accumulation appeared comparable between immunocompetent and immunocompromised children, the latter group exhibited a greater overall number of mutations, likely reflecting their prolonged infection courses. Taken together, these results suggest that, despite the distinct clinical manifestations and generally milder disease course of pediatric COVID-19, viral evolution within children follows patterns consistent with those seen in adults. Thus, persistent infec-

tion in pediatric populations may represent an additional, though not unique, contributor to ongoing SARS-CoV-2 diversification.

2. Materials and Methods

2.1. Cohort Description

This retrospective cohort study was conducted at Northwestern University's Feinberg School of Medicine (FSM) in collaboration with Ann & Robert H. Lurie Children's Hospital of Chicago (LCH) with approval from the LCH Institutional Review Board (study protocols 2020-3792 and 2025-7507). A waiver of informed consent was granted by the IRB. All residual diagnostic specimens testing positive for SARS-CoV-2 at LCH between 22 March 2020, and 22 March 2023 were de-identified and stored in a biobank at the Center for Pathogen Genomics and Microbial Evolution at FSM (n = 11,783 specimens). Patients younger than 20 years of age who provided at least 3 specimens with no more than a 90-day gap between samples were considered for inclusion in the study. We excluded any patients that did not meet the aforementioned requirements. All specimens included in this study were nasopharyngeal swabs collected in viral transport media.

2.2. Study Data and Definitions

Through manual chart review, specimens were retrospectively matched with demographic and clinical datasets, including information on sex, age, race, ethnicity, hospitalizations, comorbidities, vaccination history, and laboratory values. Vaccine records for the state of Illinois, including SARS-CoV-2 vaccines, were made available through the Illinois Comprehensive Automated Immunization Registry Exchange (I-CARE). Full vaccination status was defined as 14 days after the second dose of Pfizer-BioNTech (Pfizer) or Moderna mRNA vaccine (Moderna). Patients were categorized as partially vaccinated if they only had a single dose of either the Pfizer or Moderna vaccine. The clinical data included comorbid conditions [defined as cardiac disease, respiratory disease, diabetes, asthma, and obesity (defined as body mass index >95th percentile)], hospitalization and/or ICU admission for COVID-19, presence of viral co-infection, and receipt of therapies for COVID-19 (e.g., hydroxychloroquine, remdesivir, tocilizumab, sotrovimab, dexamethasone, and casirivimab/imdevimab). Patients were considered to be immunocompromised if: they were actively receiving or recently received chemotherapy within the past 30 days (e.g., antithymoglobulin, busulfan, fludarabine, or cyclophosphamide); they had a hematologic malignancy, solid organ transplant, or hematopoietic stem cell transplant; or if they received treatment with tacrolimus, sirolimus, or mycophenolate. Additional data, including absolute lymphocyte count (ALC), was collected to confirm immune suppression. Patients were considered asymptomatic if the clinician noted no symptoms or change from clinical baseline.

2.3. Viral RNA Extraction

Residual diagnostic specimens underwent viral RNA extraction using the QIAamp 96 Virus QIAcube HT Kit (Qiagen, Hilden, Germany, cat. no. 57731). SARS-CoV-2 viral load was quantified by quantitative reverse transcription PCR (qRT-PCR) with the CDC 2019-nCoV RT-PCR Diagnostic Panel (CDC, Atlanta, GA, USA, cat 2019-nCoVEUA-01), employing the N1 probe set for SARS-CoV-2 and human RP probes for quality control as previously described (CDC, Atlanta, GA, USA, cat 2019-nCoVEUA-01) [36]. Specimens failing to amplify the RP housekeeping gene were excluded, and those with an N1 probe cycle threshold (Ct) value less than or equal to 35 were considered positive and advanced to whole genome sequencing.

2.4. cDNA Synthesis and Viral Genome Amplification

cDNA was synthesized from 11 µL of extracted viral nucleic acids using SuperScript IV (ThermoFisher, Waltham, MA, USA, cat. no. 18090200) with random hexamers according to manufacturer's specifications. Viral genomic cDNA was then amplified in 2 multiplexed PCR reactions to generate ~400 base pair amplicons tiled across the genome. The multiplex primer set, consisting of two non-overlapping primer pools designed with Primal Scheme, was obtained from Integrated DNA Technologies (IDT) as the Artic V5.3.2 NCOV-2019 Panel (IDT, Coralville, IA, USA, cat. no. 10016495). PCR amplification was performed using Q5 HF DNA Polymerase (NEB, Ipswich, MA, USA, cat. no. M0491L) with 5 µL of cDNA in a 25 µL reaction volume. The thermal cycling program consisted of an initial step of 98 °C for 30 s, then 40 cycles of 98 °C for 15 s and 64 °C for 5 min. Separate reactions were run for each primer pool, with all pre- and post-PCR steps conducted in physically distinct spaces to minimize contamination. Amplicon sets generated for each genome were then pooled prior to sequencing library preparation.

2.5. Illumina Sequencing, Genome Assembly, and Variant Calling

Sequencing libraries were prepared from genome amplicon pools using the SeqWell ExpressPlex 384 kit per manufacturer's instructions (Beverly, MA, USA, product #101221). Pooled libraries of up to 96 genomes were sequenced on the Illumina MiSeq using the V2 500 cycle kit (Illumina, San Diego, CA, USA, product #MS-102-2003). Raw reads were processed with Trimmomatic v0.39 to remove adapters and low-quality sequences, then aligned to the Wuhan-Hu-1 reference genome (accession MN908947.3) using bwa v0.7.18. Alignment pileups were generated from the alignment using samtools v1.21, and consensus sequences were called with iVar v1.4.3 using a minimum depth of 10, a minimum base quality score of 20, and a consensus frequency threshold of 0 (i.e., majority base as the consensus). Genomes with less than 90% coverage were excluded from further analysis. Mutational frequencies were calculated using the iVar "variants" function using a minimum threshold of 3%. Only positions with a mutational frequency higher than the mean error rate ($p \leq 0.05$ by Fisher's exact test by iVar) were kept for subsequent analyses.

2.6. Phylogenetic Analysis

Genome sequences were aligned using MAFFT v7.453. Lineages were assigned using the Pango classification scheme (https://cov-lineages.org/ (accessed on 4 August 2025); PangoLearn version 9/28/2021) and subsequently grouped in variants using scorpio (scorpio v0.3.12) [37]. To examine molecular evolution and phylogenetic relationships among sequences from immunocompetent and immunocompromised patients, we constructed a Maximum Likelihood phylogeny with IQ-Tree v2.0.5 using its ModelFinder function before each analysis to estimate the nucleotide substitution model best-fitted for each dataset by means of Bayesian information criterion (BIC). We assessed the support of the clusters formed by the different lineages both with the Shimodaira–Hasegawa approximate likelihood-ratio test (SH-aLRT) and with ultrafast bootstrap (UFboot) with 1000 replicates each.

2.7. Viral Diversity Analysis

Shannon entropy (Sh) was calculated according to the equation Sh = SUM[−(pi) × log2(pi)], where pi represents the mutational frequency of all nucleotides, including the reference nucleotide, at each position for all reads in a sample. Total Sh across the whole genome or each gene was calculated as the sum of all positions within the specified region and was calculated separately per sample. Only nucleotide substitutions, not insertions or deletions, were considered in this analysis.

2.8. Statistical Modeling

Statistical analysis was performed in R v4.5.1. Differences in Shannon entropy between immunocompetent and immunocompromised patients over time were tested using a linear mixed effects model using lme4 (v 1.1-37) and lmerTest (v 3.1-3) controlling for Ct value, remdesivir administration, and patient-level random effects. Remdesivir administration was accounted for in the model as a binary indicator variable if the date of a given sample was taken during a course of Remdesivir treatment or within 7 days of the end of the most recent treatment. Log transformations were used for total Shannon entropy for whole genome model to preserve assumptions of linearity and normality of residuals.

Sh values per position were imputed as 0 for positions that did not pass the 3% mutational frequency threshold and/or failed Fisher's exact test by iVar, but had a depth greater than or equal to the lower bound to avoid group imbalance in modeling. The lower bound for read depth at each position was calculated as 90% of the lowest depth value greater than or equal to 200 reads. Positions with a depth less than the lower bound were removed from the analysis for all participants across all timepoints. *p*-values were corrected for multiple-hypothesis testing using the Benjamini–Hochberg procedure for FDR correction for per-gene model. All plots were generated with ggplot2 v3.5.2.

2.9. Code and Data Availability

Code, tools, and parameter settings required to reproduce these results, including our genome assembly pipeline and within-host variant analysis, is available on GitHub (version 3.17.6) at: https://github.com/Sakazakii/Intra-Host-Evolution-of-SARS-CoV-2-during-Persistent-Infection-of-Pediatric-COVID-19-Patients. All viral whole genome consensus sequences have been uploaded to GISAID (GISAID accession numbers provided in Supplemental Table S1). The raw reads from each sequencing run have been deposited to the NCBI Sequence Read Archive (SRA accession numbers provided in Supplemental Table S1) and have been uploaded as BioProject PRJNA1295507 (BioSample accession numbers provided in Supplemental Table S1). The individualized clinical data reported in this study cannot be deposited in a public repository due to IRB constraints, but population-level metrics are provided in Table 1.

Table 1. Summary of Demographics, Comorbidities, and Clinical Outcomes of Study Patients.

Clinical Variables	Immunocompetent Patients n = 5	Immunocompromised Patients † n = 11
Age (years)	6 (7.97)	9.2 (5.65)
Sex (male)	3 (60%)	7 (64%)
Race/Ethnicity		
Non-Hispanic, White	2 (40%)	5 (45%)
Non-Hispanic, Black	0 (0%)	1 (9%)
Non-Hispanic, Other	0 (0%)	0 (0%)
Hispanic	3 (60%)	5 (45%)
BMI	19.95 (8.5)	20.18 (6.1)
Length of Hospitalization (days)	46.5 (55.9)	114 (136)
Patients with Comorbid Conditions ‡	2 (40%)	1 (10%)
Number of Patients Who Received Treatment for COVID-19 (any)	1 (20%)	8 (73%)

Table 1. *Cont.*

Clinical Variables	Immunocompetent Patients n = 5	Immunocompromised Patients † n = 11
Remdesivir	0 (0%)	7 (64%)
Dexamethasone	0 (0%)	1 (9%)
Sotrovimab	0 (0%)	1 (9%)
Tocilizumab	1 (20%)	0 (0%)
Casirivimab/imdevimab	0 (0%)	2 (18%)
Hydroxychloroquine	1 (20%)	0 (0%)
Vaccination Status at time of Diagnosis		
Unvaccinated	5 (100%)	8 (73%)
Partially Vaccinated ◊	0 (0%)	1 (9%)
Fully Vaccinated *	0 (0%)	2 (18%)
Absolute Lymphocyte Count Prior to COVID-19 Diagnosis	N/A	1114 (832)
Absolute Lymphocyte Count After COVID-19 Diagnosis	3075 (672)	905 (825)
Maximum WHO Severity Score	2.4 (2.79)	2.2 (2.4)
Outcome		
Asymptomatic	1 (20%)	4 (36%)
Outpatient	2 (40%)	1 (9%)
Inpatient, No ICU	1 (20%)	3 (28%)
Inpatient, ICU Admission	1 (20%)	1 (9%)
Death	0 (0%)	2 (18%)
Maximum Respiratory Requirements		
Room Air	4 (80%)	10 (91%)
Nasal Cannula	0 (0%)	0 (0%)
Positive Pressure Ventilation	0 (0%)	0 (0%)
Intubation	1 (20%)	1 (9%)

Data are represented as mean (standard deviation) or number (%) for continuous or categorical variables, respectively. † Immunocompromising conditions or medications: hematologic malignancy, solid organ transplant, hematopoietic stem cell transplant, tacrolimus, sirolimus, mycophenolate, thymoglobulin, busulfan, fludarabine, cyclophosphamide. ‡ Comorbid conditions: cardiac disease, respiratory disease, diabetes, obesity, asthma. ◊ Partially Vaccinated: one dose of Moderna or Pfizer mRNA vaccines. * Fully Vaccinated: at least two doses of either Moderna or Pfizer mRNA vaccines.

3. Results

3.1. SARS-CoV-2 Whole Genome Sequencing in Longitudinal Specimens from a Cohort of Pediatric COVID-19 Patients

In March of 2020, a biobank of residual diagnostic specimens (primarily nasopharyngeal swabs) was established at the Center for Pathogen Genomics and Microbial Evolution at the Northwestern University Feinberg School of Medicine from patients who tested positive for SARS-CoV-2 by PCR-based clinical testing at Ann & Robert H. Lurie Children's Hospital of Chicago (LCH) or at one of its affiliated drive-thru testing centers, outpatient clinics, or immediate care clinics. Over a three-year period between 22 March 2020, and 22 March 2023, we collected 11,783 SARS-CoV-2 positive specimens. During this timeframe, 25 patients provided at least 3 unique specimens with no more than a 90-day gap

between samples (n = 112 specimens total). These specimens were retrospectively linked to demographic (e.g., sex, age, race, ethnicity, etc.) and clinical (e.g., hospitalization records, co-morbid conditions, vaccination history, clinical lab values, etc.) datasets. 3 of the 25 patients were found to be older than 20 years of age and excluded from further analysis, leaving 99 specimens collected from 22 patients.

RNA was extracted from each specimen and SARS-CoV-2 viral RNA was measured by quantitative reverse transcription polymerase chain reaction (qRT-PCR) (SARS-CoV-2 N1 primer set, CDC) [36]. Specimens with sufficient viral load [cycle threshold (Ct) value < 35] underwent SARS-CoV-2 whole genome sequencing on the Illumina platform using a multiplex PCR amplicon-based strategy per the ARTIC protocol. We assembled near-complete whole genome sequences from 55 specimens (at least 90% coverage, minimum read depth of 10), 5 of which were unique to a patient (i.e., there were no other sequences available) and were dropped from subsequent analyses. This left a total of 50 whole genome sequences longitudinally collected from 16 patients (consensus sequences and sequence alignments deposited in GISAID and the NCBI Sequence Read Archive, respectively; accession numbers available in Supplemental Table S1).

While we initially sought to evaluate differences in SARS-CoV-2 evolution in different age groups, our cohort was too small to achieve sufficient statistical power. We were, however, able to compare viral divergence and diversification in immunocompromised versus immunocompetent patients for comparison to the adult literature. Of the 16 patients with longitudinal sequence information, 11 were immunosuppressed or immunocompromised (defined by the patient's medical record of having a hematologic malignancy, solid organ transplant, hematopoietic stem cell transplant, or having had treatment with tacrolimus, sirolimus, mycophenolate, thymoglobulin, busulfan, fludarabine, or cyclophosphamide) (Table 1; list of immunocompromising conditions in Table 2). The two groups had similar age, sex, and race/ethnicity distributions. All the immunocompetent and 82% of the immunocompromised patients were not vaccinated at the time of the study, which largely predated vaccine rollout in the pediatric population. Only 1 (20%) of the immunocompetent patients received treatment for COVID-19 (including remdesivir, hydroxychloroquine, dexamethasone, and monoclonal antibodies) as opposed to 8 (73%) of the immunocompromised patients. Four patients received monoclonal antibodies for the treatment of COVID-19, including one immunocompetent patient (tocilizumab) and three immunocompromised patients (one received sotrovimab and two received casirivimab/imdevimab).

Absolute lymphocyte count (ALC) data extracted from complete blood counts collected within the 5 days prior to COVID-19 diagnosis in the immunocompromised patients confirmed immunosuppression compared to expected age-matched standards [mean (SD) = 1114 (832); data not available for immunocompetent patients] (Table 1). This was confirmed in the ALC data collected within two days after diagnosis, with immunocompetent patients having significantly higher ALCs (mean of 3075 compared to 905, respectively). Note there was a relative, but not statistically significant, reduction in the mean ALC after COVID-19 diagnosis in the immunocompromised patient population.

A majority of the immunocompetent patients were asymptomatic (20%) or outpatients (40%) while a majority of the immunocompromised patients were admitted to the hospital (28%), required ICU care (9%), or succumbed to their infection (18%) (Table 1). Despite these differences in outcome, disease severity as measured by the maximally recorded WHO severity score during their infection course was not significantly different between these two groups (mean of 2.4 and 2.2 for immunocompetent and immunocompromised patients, respectively). For reference, a WHO severity score of 2 corresponds to symptomatic infection requiring only ambulatory care (i.e., without the need for supplemental oxygen

or hospitalization). This highlights that, on average, both groups experienced relatively mild disease despite the observed differences in clinical care and outcomes.

Table 2. Immunocompromising conditions and medications.

Patient	Immunocompromising Conditions/Medications
1	Osteosarcoma, methotrexate
2	Living donor kidney transplant, mycophenolate mofetil, tacrolimus
3	CAR-T, blinatumomab
4	T-ALL
5	Neuroblastoma, hematopoietic stem cell transplant, thiotepa, cyclophosphamide, eculizumab
6	Living donor kidney transplant, mycophenolate mofetil
7	ALL, CAR-T, hematopoietic stem cell transplant
8	B-ALL, CAR-T, fludarabine, cyclophosphamide
9	Mesenchymal stem cell infusion
10	AML, hematopoietic stem cell transplant, fludarabine, busulfan, antithymocyte globulin
11	B-ALL, hematopoietic stem cell transplant, CAR-T, inotuzimab

For the hospitalized patients, the mean length of stay was longer for the immunocompromised patients (114 days) compared to the immunocompetent patients (46.5 days) (Table 1). This generally correlated with longer time courses over which these patients provided SARS-CoV-2 positive specimens (Figure 1A). Patient specimens were not collected over uniform schedules as testing was dictated by the ordering clinician and the residual specimen was collected retrospectively. As some of the patients were asymptomatic, all collection dates are shown relative to the date of first diagnosis. The total length of specimen collection from individual patients ranged from less than a week to over 200 days. The average length of collection was longer in immunocompromised (48 days) compared to immunocompetent patients (27 days), which is consistent with longer persistence in immunocompromised patients as reported in adult populations [21,37].

As described above, Ct values for SARS-CoV-2 viral RNA in each specimen were calculated using a previously validated qRT-PCR assay as a proxy for a viral load (Figure 1B). In general, viral load decreased over time in each patient with some stochasticity. A handful of immunocompromised patients retained higher viral loads (Ct value less than 30) beyond 25 days since initial diagnosis, though this did not clearly correlate with outcome (1 asymptomatic patient, 2 inpatients not requiring ICU care, and 1 patient who died). One immunocompetent outpatient registered a high viral load (low Ct value) 68 days after initial diagnosis, though this was later found to be a reinfection (refer to next section).

3.2. Phylogenetic Analysis Reveals Minor Changes in Viral Consensus Sequences over Time

To better characterize the consensus level mutations that arose over time, maximum-likelihood phylogenetic trees were constructed for the 35 longitudinal isolates collected from the 11 immunocompromised patients (Figure 2A) and the 15 longitudinal isolates collected from the 5 immunocompetent patients (Figure 2B). Nearly all the sequences collected from each patient were closely related and belonged to the same lineage/clade, confirming that they were all collected over the course of the same infection. The one exception to this was an immunocompetent patient from which two specimens with identical consensus sequences mapping to lineage AY.25.1 (clade 21J, Delta variant) were collected two days

apart and a third specimen mapping to lineage BA.1 (clade 21K, Omicron variant) was collected 66 days later. This third specimen was excluded from further analysis as a reinfection. Longitudinal specimens from the same patient clustered together in a larger subsampling of SARS-CoV-2 whole genome sequences form the US over the time course of this study and clustered with their expected lineage designations (Supplemental Figure S1).

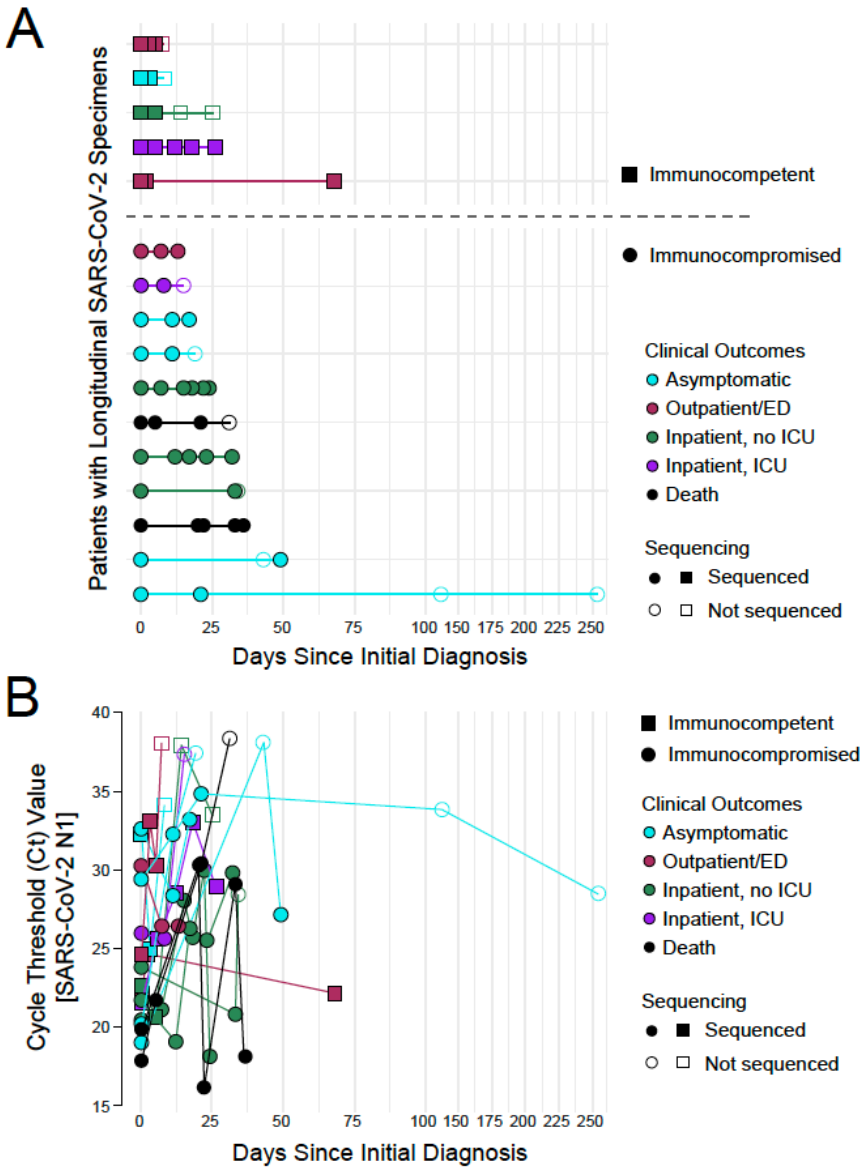


Figure 1. Overview of longitudinal SARS-CoV-2 specimens collected from pediatric patients. (**A**) A graphical representation of the longitudinal specimens included in this study plotted relative to the day of first diagnosis (n = 16 total participants). Participants are grouped by immune status [immunocompetent as squares (n = 5), immunocompromised as circles (n = 11)] and colored by outcome [asymptomatic in blue (n = 5), outpatients in red (n = 3), inpatients never requiring ICU care in green (n = 4), inpatients requiring ICU care in purple (n = 2), and COVID-related deaths in black (n = 2)]. Specimens that yielded a SARS-CoV-2 whole genome sequence are in closed symbols and those that did not are in open symbols. (**B**) Plot of the qRT-PCR Ct values for SARS-CoV-2 N1 in each specimen plotted relative to the day of first diagnosis and visualized using the same schema as in (**A**).

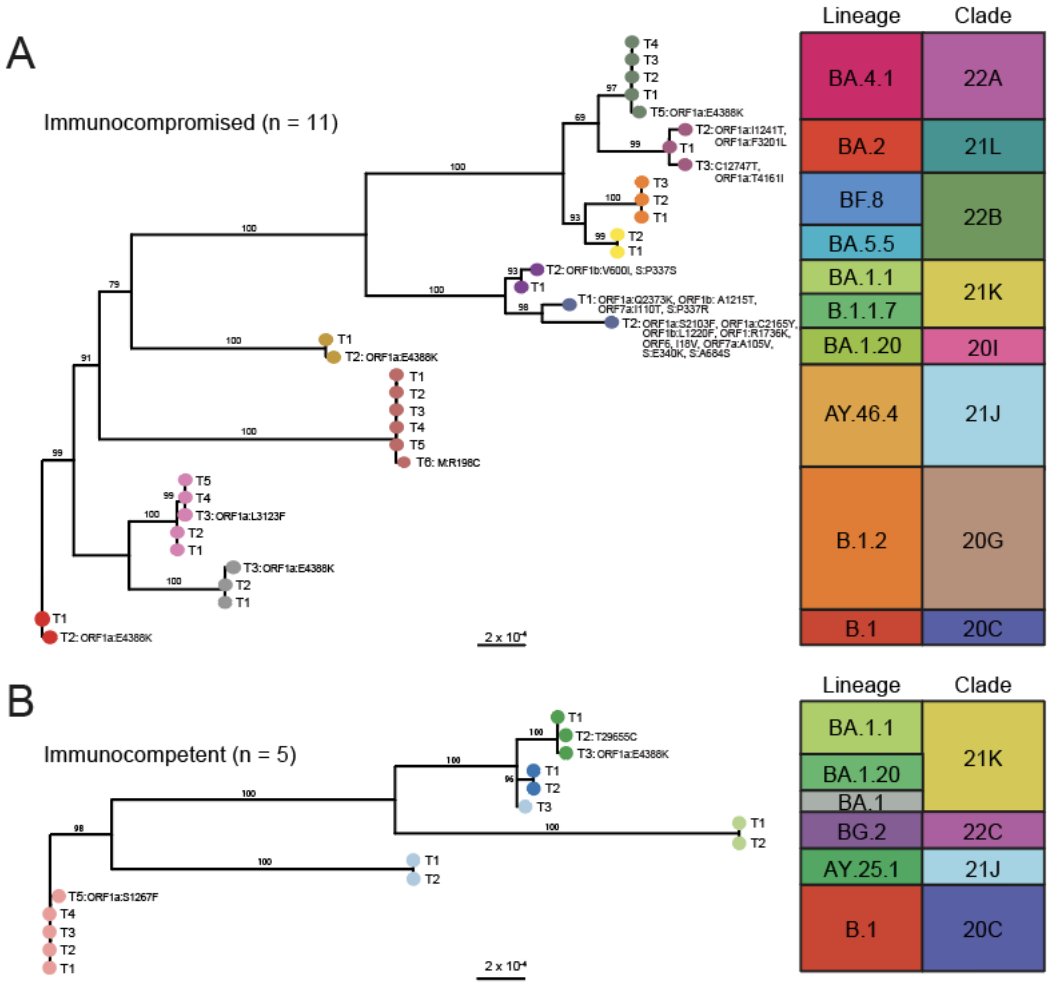


Figure 2. Phylogenetic trees of longitudinally collected SARS-CoV-2 whole genome sequences. Maximum likelihood phylogenetic trees of SARS-CoV-2 whole genome sequences from longitudinally collected clinical specimens from (**A**) immunocompromised (n = 11) and (**B**) immunocompetent (n = 5) patients. Branch tips are colored by patient and labelled by timepoint (T1, T2, etc.). Consensus sequence mutations relative to the nearest branch point are labeled by nucleotide change (synonymous) or open-reading frame and amino acid substitution (non-synonymous). Bootstrap values are given above each branch as a percent of 1000 replicates. The lineage and clade of each sequence is shown on the right.

Over time, the immunocompetent patients saw either no (n = 3, 60%) or minimal (n = 2, 40%) changes in the consensus sequence of the infecting isolate (Figure 2B) with no changes observed in the Spike open reading frame (ORF). On the contrary, only 2 of the 11 (n = 2, 18%) immunocompromised patients had no change in the consensus sequence of the virus over time, with a majority of specimens showing at least one base pair change at later timepoints (Figure 2A). Some patient specimens showed greater divergence, with one set of paired specimens differing by 9 amino acid changes. This included three amino acid substitutions in Spike: P337R (present only at the first timepoint), and E340K and A684S (present at the second timepoint). Notably, this patient was administered sotrovimab prior to collection of the first specimen and variation at positions 337 and 340 in the receptor-binding domain (RBD) have been previously associated with escape from sotrovimab in adult populations [34,38]. A second immunocompromised patient also showed a substitution in Spike at position 337, though to a different amino acid (P337S), and this patient did not receive any monoclonal antibody therapeutics. None of the other patients who received monoclonal antibodies for the treatment of COVID-19 showed any

consensus mutations in Spike. Outside of Spike, several patients evolved an ORF1a:E4388K mutation (n = 1 immunocompetent patient and n = 4 immunocompromised patients). While more mutations were observed in the immunocompromised cohort, direct comparison with the immunocompetent cohort is confounded by the larger cohort size and generally greater lengths of time between samplings.

3.3. The Rate of Change in SARS-CoV-2 Intra-Host Diversity Does Not Differ by Immune Status in Pediatric Populations

To enable a more granular quantitative comparison of SARS-CoV-2 evolution in the immunocompromised versus immunocompetent pediatric cohorts, we next evaluated intra-host genetic diversity of the viral population in each specimen. Variant allele frequencies per sample were used to calculate mutational frequency and Shannon entropy (Sh) at each position across the genome relative to the Wuhan-Hu-1 reference genome (accession MN908947.3). Only positions with a variant allele frequency greater than or equal to 3% and with a significant *p*-value (p-value < 0.05) were included in these analyses. Overlaying all the mutational frequency data from the immunocompetent (left) and immunocompromised (right) patients provides a snapshot of overall positional diversity (Figure 3A). Viruses from both sets of patients have a large number of positions with 100% mutational frequency corresponding with known mutations in the originating lineages. To better visualize this, Sh was calculated as nucleotide-agnostic metric of variability at each position and plotted similarly (Figure 3B, refer to Methods for Sh formula and GitHub for associated code). Diversity was dispersed relatively equally along the genome in both cohorts, though the immunocompromised patients tended to have more sites with higher entropy values (>0.5). Again, however, this is confounded by the longer infection time courses in these patients.

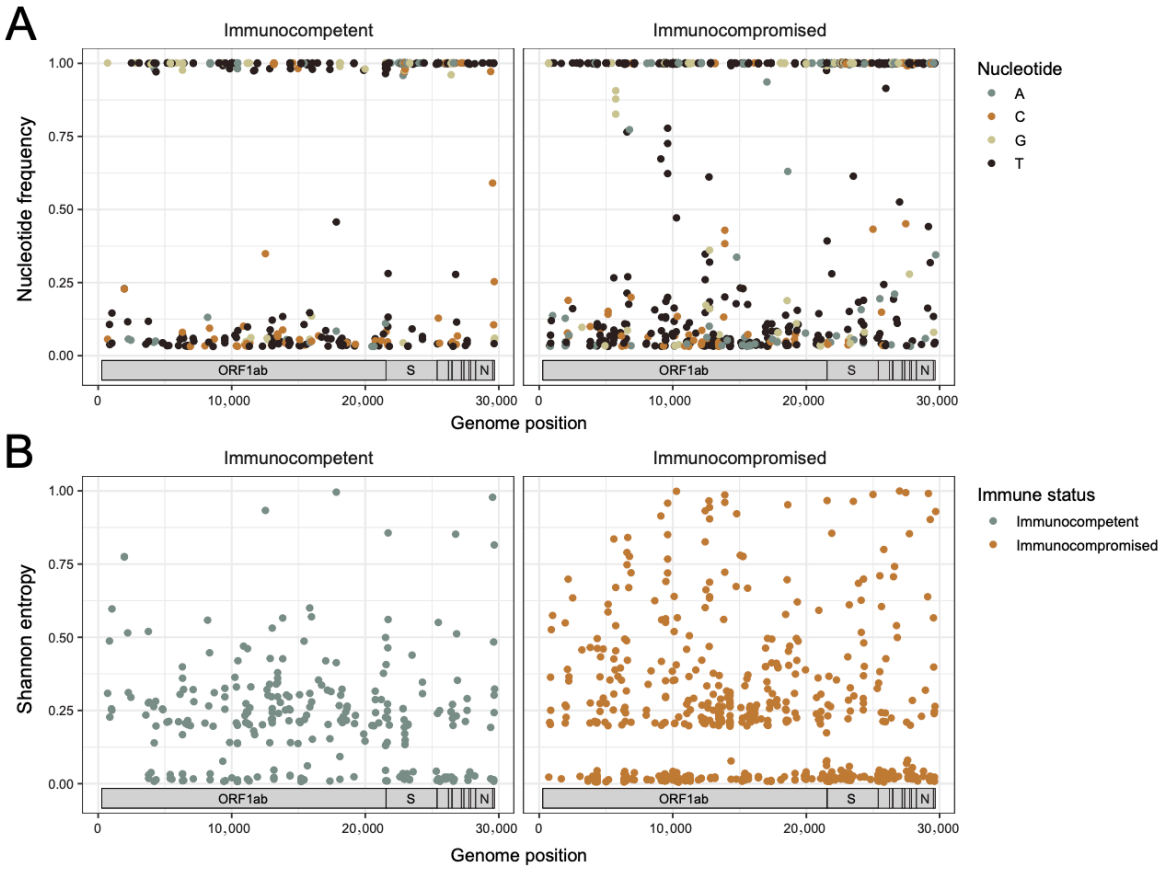


Figure 3. Intra-host diversity across the SARS-CoV-2 genome. (**A**) Dot plots showing mutational frequency per alternative alleles across the genome relative to the Wuhan-Hu-1 reference genome

calculated as a fraction of raw reads for immunocompetent (left, n = 15 specimens) and immunocompromised (right, n = 35 specimens) patients, color-coded by nucleotide. The data from all patients and all timepoints per immune status category are overlaid. Only positions with a variable allele frequency greater than or equal to 3% and sequencing depth > 400 reads were included. (**B**) Dot plots showing Shannon entropy per position across the genome for the sites shown in (**A**) from immunocompetent (left, n = 15 specimens) and immunocompromised (right, n = 35 specimens) patients. The data from all patients and all timepoints per immune status category are overlaid.

To better visualize the change over time, the Sh per genome position was plotted for each specimen in each patient independently (Supplemental Figure S2A). For this analysis, only positions with a read depth greater than or equal to the lower bound across all timepoints within a patient were included (lower bound calculated as 90% of the lowest depth value exceeding 400 reads). If the read depth failed to cross the lower bound in any one specimen, it was dropped from the analysis across all specimens from that patient. In general, Sh increased over time in each patient with more obvious increases observed with longer times between sampling. To better quantify this, the total Sh for each specimen across all positions was calculated and plotted over time (Figure 4A, refer to calculations in Methods). Linear mixed-effects models were independently fit to the immunocompetent and immunocompromised specimens, which revealed a clear correlation between days since initial diagnosis and entropy. Adjusting for Ct value, remdesivir administration, and patient-level effects, there was no significant difference between the immunocompromised and immunocompetent patients over time, with both accumulating total entropy at very similar rates (p-value = 0.401).

While the rate of accumulation was similar, there could be positional differences in where genomic diversity was accumulating in each cohort. To assess this, the total Sh for each viral ORF across all specimens was calculated and compared between the immunocompetent and immunocompromised patients (Supplemental Figure S2B). To ensure direct comparison, the lower bound was calculated as 90% of the smallest depth value within the same position across the dataset that exceeds 200 reads. Positions with a depth less than the lower bound were removed from the analysis for all participants across all timepoints. While some ORFs accumulated more diversity than others (i.e., nsp3, nsp12, Spike, etc.), there were no significant differences between the immunocompetent and immunocompromised groups over time as measured by a linear mixed-effects model adjusting for Ct value, remdesivir administration, and patient-level effects.

While there were no significant differences in intra-host diversity by ORF, it is still possible that specific positions would be preferentially selected for based on immune status. To assess this, we extracted mutational frequency data for all positions in Spike with mutation frequency greater than or equal to 3% (or less than or equal to 97%) as compared to the Wuhan-Hu-1 reference genome (accession MN908947.3) (Supplemental Table S2). A modified heatmap of the mutational frequency by position was plotted relative to the Spike domain structure with diversity in each patient represented by a line above (immunocompetent) or below (immunocompromised) the schematic (Figure 4B). If positional diversity was observed at multiple timepoints, the greatest change in diversity was visualized. Non-synonymous changes at 9 sites were found in 2 or more patients (vertical lines and labels). While only the changes at P337 reached consensus (Figure 2A), several other minority variants previously associated with immune escape were observed in multiple patients, including changes at D405 and K417. Nevertheless, these were generally observed in both the immunocompromised and immunocompetent cohorts. Taken together, these data suggest similar patterns of SARS-CoV-2 intra-host evolution between immunocompro-

mised and immunocompetent pediatric COVID-19 patients, though additional sequencing in larger cohorts will be required to validate these results.

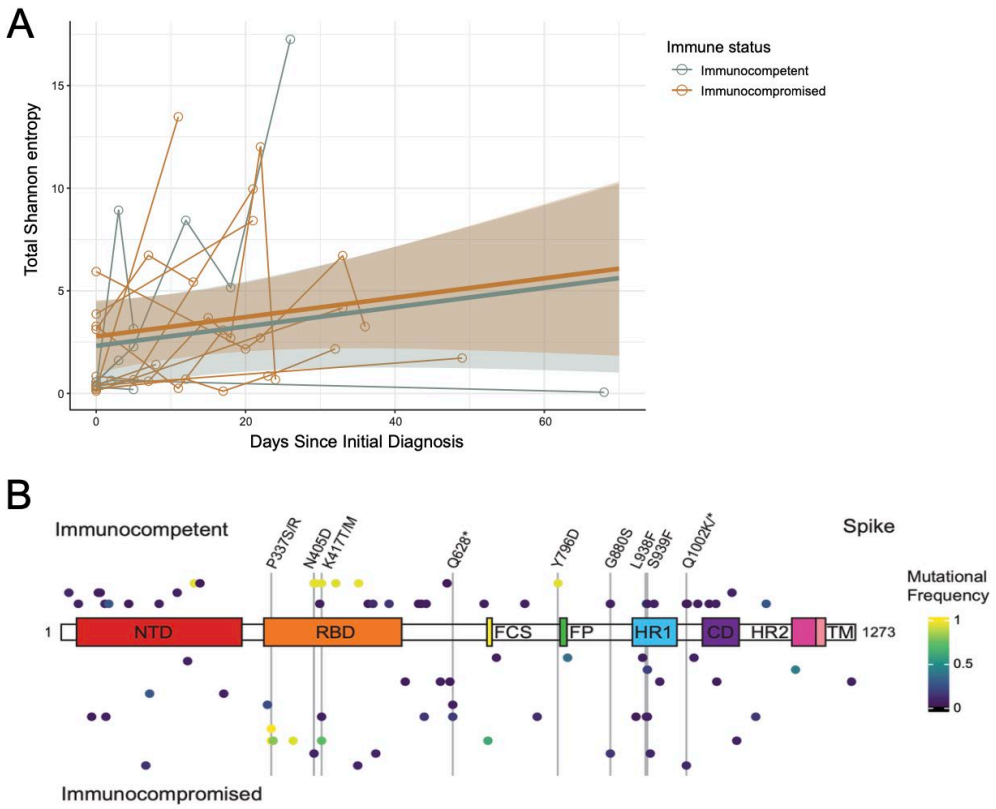


Figure 4. Similar trends in intra-host diversification of SARS-CoV-2 are observed in immunocompromised and immunocompetent pediatric COVID-19 patients. (**A**) Line plot of Shannon entropy [Sh = SUM[−(pi) × log2(pi)], where pi is nucleotide frequency] across the genome of each isolate with longitudinal specimens from the same patient connected by lines. Sh values per position were imputed as 0 for positions that did not pass the 3% mutational frequency threshold and/or failed iVar test ($p \leq 0.05$ by Fisher's exact test by iVar) but have a depth greater than or equal to lower bound to avoid group imbalance in modeling. The lower bound was calculated as 90% of the smallest depth value within the same position across the dataset that exceeds 200 reads. A linear mixed-effects model evaluating the trend in total Sh over time in immunocompetent (blue) and immunocompromised (orange) pediatric patients was conducted controlling for Ct, time between sample acquisition, remdesivir administration, and within-patient variability. (**B**) Dot plot of mutational frequency by position across the Spike open reading frame. Diversified sites in one patient are all equidistant from the Spike domain schematic (immunocompetent above, immunocompromised below). Only positional changes that resulted in a non-synonymous change with a variable allele frequency > 3% or <97% and a sequencing depth > 10 reads in at least 1 timepoint are shown. Common sites of diversification between two or more patients are labelled with vertical lines and the amino acid substitution.

4. Discussion

This study presents a genome-wide longitudinal analysis of SARS-CoV-2 intra-host evolution in pediatric COVID-19 patients, comparing divergence and intra-host diversity between immunocompetent and immunocompromised children. While prior studies have largely focused on adult cases of persistent infection as sources of SARS-CoV-2 variant emergence, our findings demonstrate that pediatric patients, especially those with impaired immune function, are also capable of sustaining prolonged viral replication and developing genetically divergent viral populations. Consistent with previous studies in adult immunocompromised patients [13,16,30], our pediatric cohort demonstrated prolonged infection in immunocompromised children with sustained viral replication evidenced by persistently

low Ct values, extended PCR positivity (up to 256 days), and the accumulation of mutations over time. These findings reinforce the need to view viral persistence in immunocompromised children as a plausible contributor to the ongoing viral genetic landscape.

SARS-CoV-2 accumulated non-synonymous mutations in several immunocompromised patients over time, including several mutations in regions of Spike with known immunologic significance [39–42]. For example, one immunocompromised patient who received the monoclonal antibody sotrovimab prior to specimen collection acquired two mutations in the Spike RBD known to confer sotrovimab resistance: P337R and E340K. Sotrovimab targets a highly conserved epitope on the RBD and its binding is facilitated by a proteoglycan at residue N343 [43]. Mutations at positions P337 and E340 are exceedingly rare (<0.1% global prevalence) and have been noted by multiple studies as hallmark resistance mutations to sotrovimab, which has been confirmed in a number of neutralization assays pseudovirus neutralization assays when exposed to sotrovimab [44–50]. It is notable then that a second pediatric patient in this study evolved a P337S mutation despite not having received sotrovimab or any other monoclonal antibody therapeutic. This may suggest a unique selective pressure in this patient that warrants additional surveillance in similar pediatric patients. That being said, sotrovimab is no longer recommended for clinical use following global emergence of the BA.2.86 lineage and its descendants, all of which encode a Spike K356T mutation that ablates sotrovimab binding through addition of a glycosylation site and epitope masking [51]. Furthermore, the three other patients who received monoclonal antibody therapeutics in our cohort did not evolve any resistance mutations. Regardless, the recurrence of these substitutions in our cohort underscores that selective pressures from therapeutic interventions can shape intra-host viral populations in pediatric patients similarly to adults [14,33,52,53].

Outside of Spike, longitudinal sequencing of SARS-CoV-2 in several pediatric patients identified the convergent evolution of an ORF1a:E4388K mutation. This corresponds to an E135K mutation in nsp10, a virally encoded cofactor that binds to and enhances the activities of the nsp14 exoribonuclease and the nsp16 methyltransferase. Position 135 is near the C-terminal tail of nsp10, which was truncated in available co-crystal structures and is not thought to play a direct role in mediating the nsp14 or nsp16 interaction [54,55]. This mutation was exceedingly rare over the timeframe of this study, but has appeared more frequently in some Omicron sublineages, including in the currently circulating lineages LP.8.1 and XFG [56]. If and how this mutation may alter the function of these enzymatic complexes or alter viral fitness has not yet been explored.

While several non-synonymous mutations were observed at the consensus level in our immunocompromised cohort, this was rare in our immunocompetent cohort. That being said, there were fewer immunocompetent patients that fit our inclusion criteria and they generally had shorter infection time courses, as expected. Therefore, to better compare viral evolution in our immunocompromised versus immunocompetent patients, we conducted an analysis of intra-host variation across the genome over time. A majority of observed variants were present at low to moderate frequencies, consistent with previous reports suggesting that intra-host variation in SARS-CoV-2 is typically limited during acute infection [39–41]. Nevertheless, this diversity accumulated at similar rates in both cohorts, even when controlling for time, viral load, and patient-to-patient variability. Furthermore, immune status was not significantly associated with total Shannon entropy in any particular ORF, suggesting lack of a localized selective pressure in one cohort versus the other. Focusing on variation observed in Spike, there were some positions that varied in multiple patients over time, including several positions in the Spike RBD known to confer neutralizing antibody escape (i.e., N405D, K417T, etc.) [53,57,58]. However, these were largely observed in both immunocompetent and immunocompromised patients.

Taken together, these data suggest that the rate of SARS-CoV-2 intra-host evolution is the same regardless of a patient's immune status, but that longer persistence allows longer times for mutations to arise in immunocompromised patients. Prolonged infection then ultimately provides the diversity required to respond to selective pressures when they arise, including to monoclonal antibody therapeutics and other treatment regimens. These data are consistent with what has been previously reported in adult populations [22,29,59].

There are several important limitations to our study. Our study is limited by a small sample size, which decreased our statistical power and prevented us from comparing viral evolution across different age categories. Our inclusion criteria requiring at least 3 longitudinal specimens from each patient made inclusion of immunocompromised individuals more likely as repeat testing of these patients was standard of care to evaluate risk for dissemination of disease, determine length of therapy for remdesivir, and inform additional clinical decision-making practices. Despite the higher number of immunocompromised individuals in the study, there was substantial diversity in the immunosuppressive conditions and treatments administered (Table 2), which prevented a more granular analysis. Understanding how different immunocompromising conditions—such as a particular chemotherapy, underlying immunodeficiency, immunosuppressive therapy, or oncologic malignancy—may drive unique patterns of viral persistence and evolution remains a major knowledge gap. There is some evidence correlating the use of calcineurin inhibitors with increased development of escape mutations within longitudinally sequenced SARS-CoV-2, but more research in this area is sorely needed [46].

Additionally, this study followed a single-site, retrospective design with non-standardized sampling intervals, making the number of samples per patient and overall duration of sample collection variable. While we strived to account for this by incorporating time into our statistical models, it may skew data interpretation. Finally, these studies were conducted with specimens collected prior to 2023 and it is possible that newly emergent variants of SARS-CoV-2 might have different behavioral patterns. To better assess evolution over patient age, future studies should incorporate larger cohorts that include both pediatric and adult populations for direct comparison, ideally with functional assays of emergent mutations to investigate host–virus immune interactions longitudinally.

Despite these limitations, this study fills an important research gap providing whole genome sequencing analysis of longitudinal SARS-CoV-2 isolates from pediatric COVID-19 patients at a quaternary care hospital. Our work highlights that persistent SARS-CoV-2 infection in immunocompromised pediatric patients can drive intra-host viral evolution and the emergence of critical immune escape variants, similar to what has been observed in adult populations. This suggests that persistent infection in children may be an additional, but not unique, source of ongoing viral diversification and highlights the need for viral surveillance efforts to include pediatric patients in their design. Furthermore, it emphasizes the need tailored clinical management strategies for immunocompromised patients across all age groups to limit the emergence and spread of immune escape variants.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/v17101313/s1, Supplemental Figure S1: Intra-host diversity of SARS-CoV-2 by patient and open reading frame; Supplemental Table S1: GISAID and NCBI Accession numbers for deposited sequences in this study; Supplemental Table S2: Mutational frequency at positions in Spike.

Author Contributions: Conceptualization, C.R.B. and J.F.H.; Methodology, C.R.B., T.D., E.R.-G., J.M., L.M.S., R.L.-R., A.A., E.A.O. and J.F.H.; Software, C.R.B., T.D., E.R.-G., R.L.-R., A.A. and E.A.O.; Validation, C.R.B., T.D., R.L.-R., E.A.O. and J.F.H.; Formal Analysis, C.R.B., T.D., E.R.-G., J.M. and R.L.-R.; Investigation, C.R.B., T.D., E.R.-G., L.M.S., R.L.-R., A.A. and J.F.H.; Resources, D.S.G., D.B.W., L.M.S., E.A.O. and J.F.H.; Data Curation, C.R.B., T.D., E.R.-G., J.M., L.M.S. and A.A.; Writing—Original

Draft, C.R.B. and J.F.H.; Writing—Review & Editing, C.R.B. and J.F.H.; Visualization, C.R.B., T.D., E.R.-G., R.L.-R. and J.F.H.; Supervision, R.L.-R., E.A.O. and J.F.H.; Project Administration, L.M.S., R.L.-R., E.A.O. and J.F.H.; Funding Acquisition, E.A.O., R.L.-R. and J.F.H. All authors have read and agreed to the published version of the manuscript.

Funding: Funding for this work was provided by: the Quantitative Biosciences Institute Coronavirus Research Group (QCRG) Antiviral Drug Discovery (AViDD) center (NIH U19AI171110), the Successful Clinical Response to Pneumonia Therapy (SCRIPT) Center (NIH U19AI135964), other support for respiratory virus research from the NIH (R01AI177498 to J.F.H., S10OD032243 to E.A.O.), and by institutional support for the Center for Pathogen Genomics and Microbial Evolution. E.R.-G. was supported by the Immunology and Molecular Sciences Training Grant at Northwestern University (NIH T32AI007476). C.R.B. was supported by a RAPID Scholars award from the Academic Pediatric Association. The funding sources had no role in the study design, data collection, analysis, interpretation, or writing of the report.

Institutional Review Board Statement: This study was conducted with approval by the LCH Institutional Review Board per study protocols 2020-3792 and 2025-7507. Protocol 2025-7507 received IRB approval as an exempt study on 1 October 2024 (also date of last review). Protocol 2020-3792 received IRB approval as an exempt study on 6 February 2020 (date of last review 5 February 2024).

Informed Consent Statement: A waiver of informed consent was granted by the IRB.

Data Availability Statement: Code, tools, and parameter settings required to reproduce these results, including our genome assembly pipeline and within-host variant analysis, is available on GitHub at: https://github.com/Sakazakii/Intra-Host-Evolution-of-SARS-CoV-2-during-Persistent-Infection-of-Pediatric-COVID-19-Patients. All viral whole genome consensus sequences have been uploaded to GISAID (GISAID accession numbers provided in Supplemental Table S1). The raw reads from each sequencing run have been deposited to the NCBI Sequence Read Archive (SRA accession numbers provided in Supplemental Table S1) and have been uploaded as BioProject PRJNA1295507 (BioSample accession numbers provided in Supplemental Table S1). The individualized clinical data reported in this study cannot be deposited in a public repository due to IRB constraints, but population-level metrics are provided in Table 1.

Acknowledgments: The authors are grateful for the assistance from Cecilia Thompson for her assistance with specimen acquisition. This research was supported in part through the computational resources and staff contributions provided by the Genomics Compute Cluster which is jointly supported by the Feinberg School of Medicine, the Center for Genetic Medicine, and Feinberg's Department of Biochemistry and Molecular Genetics, the Office of the Provost, the Office for Research, and Northwestern Information Technology. The Genomics Compute Cluster is part of Quest, Northwestern University's high-performance computing facility, to advance genomics research. We further gratefully acknowledge all data contributors, i.e., the Authors and their Originating laboratories responsible for obtaining the specimens, and their Submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative, on which this research is based.

Conflicts of Interest: J.F.H. received research support, paid to Northwestern University, from Gilead Sciences and is a paid consultant for Merck and Ridgeback Biotherapeutics. E.R.-G. has previously been a paid consultant for Merck. All other authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
COVID-19	Coronavirus Disease 2019
MIS-C	Multisystem Inflammatory Syndrome in Children
ICU	Intensive Care Unit
VOC	Variant of concern

LCH	Ann and Robert H. Lurie Children's Hospital
ALC	Absolute lymphocyte count
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
ORF	Open reading frame
Sh	Shannon entropy
Ct	Cycle threshold

References

1. Inagaki, K.; Hobbs, C.V. COVID-19: A Pediatric Update in Epidemiology, Management, Prevention, and Long-term Effects. *Pediatr. Rev.* **2023**, *44*, 243–254. [CrossRef]
2. Powell, A.A.; Dowell, A.C.; Moss, P.; Ladhani, S.N.; sKIDS Investigation Team. Current state of COVID-19 in children: 4 years on. *J. Infect.* **2024**, *88*, 106134. [CrossRef] [PubMed]
3. Drouin, O.; Hepburn, C.M.; Farrar, D.S.; Baerg, K.; Chan, K.; Cyr, C.; Donner, E.J.; Embree, J.E.; Farrell, C.; Forgie, S.; et al. Characteristics of children admitted to hospital with acute SARS-CoV-2 infection in Canada in 2020. *Can. Med Assoc. J.—CMAJ* **2021**, *193*, E1483–E1493. [CrossRef] [PubMed]
4. Hoang, A.; Chorath, K.; Moreira, A.; Evans, M.; Burmeister-Morton, F.; Burmeister, F.; Naqvi, R.; Petershack, M.; Moreira, A. COVID-19 in 7780 pediatric patients: A systematic review. *EClinicalMedicine* **2020**, *24*, 100433. [CrossRef] [PubMed]
5. Martin, B.; DeWitt, P.E.; Russell, S.; Anand, A.; Bradwell, K.R.; Bremer, C.; Gabriel, D.; Girvin, A.T.; Hajagos, J.G.; McMurry, J.A.; et al. Characteristics, Outcomes, and Severity Risk Factors Associated with SARS-CoV-2 Infection Among Children in the US National COVID Cohort Collaborative. *JAMA Netw. Open* **2022**, *5*, e2143151. [CrossRef] [PubMed]
6. CDC COVID-19 Response Team. Coronavirus Diseases 2019 in Children—United States, February 12–April 2, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **2020**, *69*, 422–426. [CrossRef]
7. Zimmermann, P.; Curtis, N. Coronavirus Infections in Children Including COVID-19: An Overview of the Epidemiology, Clinical Features, Diagnosis, Treatment and Prevention Options in Children. *Pediatr. Infect. Dis. J.* **2020**, *39*, 355–368. [CrossRef] [PubMed] [PubMed Central]
8. Saydah, S.H.; Campbell, A.P.; Randolph, A.G. Consequences beyond acute SARS-CoV-2 infection in children. *Sci. Transl. Med.* **2024**, *16*, eado2099. [CrossRef] [PubMed] [PubMed Central]
9. Ebrahim, S.; Husain, W.J.; Hubail, Z. Cardiac Manifestations in Post-COVID-19 Multisystem Inflammatory Disease in Children (MIS-C) in Pediatric Patients at a Tertiary Hospital in Bahrain: A Retrospective Study. *Cureus* **2025**, *17*, e83409. [CrossRef] [PubMed] [PubMed Central]
10. Gadiwala, S.; Mistry, A.; Patel, S.; Chaithanya, A.; Pathak, S.; Satnarine, T.; Bekina-Sreenivasan, D.; Bakarr, A.A.; Das, B.B.; Chakinala, R.C.; et al. MIS-C related to SARS-CoV-2 infection: A narrative review of presentation, differential diagnosis, and management. *Infez. Med.* **2022**, *30*, 344–352. [CrossRef] [PubMed] [PubMed Central]
11. Patel, J.M. Multisystem Inflammatory Syndrome in Children (MIS-C). *Curr. Allergy Asthma Rep.* **2022**, *22*, 53–60. [CrossRef] [PubMed] [PubMed Central]
12. Tenenbaum, T.; Liese, J.; Welte, T.; Rademacher, J. Respiratory Syncytial Virus-Associated Respiratory Diseases in Children and Adults. *Dtsch. Arztebl. Int.* **2024**, *121*, 303–312. [CrossRef] [PubMed] [PubMed Central]
13. Hansen, C.L.; Chaves, S.S.; Demont, C.; Viboud, C. Mortality Associated with Influenza and Respiratory Syncytial Virus in the US, 1999–2018. *JAMA Netw. Open* **2022**, *5*, e220527. [CrossRef]
14. Kemp, S.A.; Collier, D.A.; Datir, R.P.; Ferreira, I.A.T.M.; Gayed, S.; Jahun, A.; Hosmillo, M.; Rees-Spear, C.; Mlcochova, P.; Lumb, I.U.; et al. SARS-CoV-2 evolution during treatment of chronic infection. *Nature* **2021**, *592*, 277–282. [CrossRef] [PubMed]
15. Mushegian, A.A.; Long, S.W.; Olsen, R.J.; Christensen, P.A.; Subedi, S.; Chung, M.; Davis, J.; Musser, J.; Ghedin, E. Within-host genetic diversity of SARS-CoV-2 in the context of large-scale hospital-associated genomic surveillance. *arXiv* **2022**, arXiv:2022.08.17.22278898. [CrossRef] [PubMed] [PubMed Central]
16. Rothman, J.A.; Saghir, A.; Zimmer-Faust, A.G.; Langlois, K.; Raygoza, K.; Steele, J.A.; Griffith, J.F.; Whiteson, K.L. Longitudinal Sequencing and Variant Detection of SARS-CoV-2 across Southern California Wastewater. *Appl. Microbiol.* **2024**, *4*, 635–649. [CrossRef]
17. Choi, B.; Choudhary, M.C.; Regan, J.; Sparks, J.A.; Padera, R.F.; Qiu, X.; Solomon, I.H.; Kuo, H.-H.; Boucau, J.; Bowman, K.; et al. Persistence and evolution of SARS-CoV-2 in an immunocompromised host. *N. Engl. J. Med.* **2020**, *383*, 2291–2293. [CrossRef]
18. Truong, T.T.; Ryutov, A.; Pandey, U.; Yee, R.; Goldberg, L.; Bhojwani, D.; Aguayo-Hiraldo, P.; Pinsky, B.A.; Pekosz, A.; Shen, L.; et al. Persistent SARS-CoV-2 infection and increasing viral variants in children and young adults with impaired humoral immunity. *eBioMedicine.* **2021**, *67*, 103355. [CrossRef] [PubMed] [PubMed Central]

19. Wilkinson, S.A.J.; Richter, A.; Casey, A.; Osman, H.; Mirza, J.D.; Stockton, J.; Quick, J.; Ratcliffe, L.; Sparks, N.; Cumley, N.; et al. Recurrent SARS-CoV-2 mutations in immunodeficient patients. *Virus Evol.* **2022**, *8*, veac050. [CrossRef] [PubMed] [PubMed Central]
20. Avanzato, V.A.; Matson, M.J.; Seifert, S.N.; Pryce, R.; Williamson, B.N.; Anzick, S.L.; Barbian, K.; Judson, S.D.; Fischer, E.R.; Martens, C.; et al. infectious SARS-CoV-2 shedding from an asymptomatic immunocompromised individual with cancer. *Cell* **2020**, *183*, 1901–1912.e9. [CrossRef] [PubMed]
21. Hettle, D.; Hutchings, S.; Muir, P.; Moran, E. COVID-19 Genomics UK (COG-UK) consortium. Persistent SARS-CoV-2 infection in immunocompromised patients facilitates rapid viral evolution: Retrospective cohort study and literature review. *Clin. Infect. Pract.* **2022**, *16*, 100210. [CrossRef] [PubMed] [PubMed Central]
22. Choudhary, M.C.; Crain, C.R.; Qiu, X.; Hanage, W.; Li, J.Z. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Sequence Characteristics of Coronavirus Disease 2019 (COVID-19) Persistence and Reinfection. *Clin. Infect. Dis.* **2022**, *74*, 237–245. [CrossRef] [PubMed] [PubMed Central]
23. Cherian, S.; Potdar, V.; Jadhav, S.; Yadav, P.; Gupta, N.; Das, M.; Rakshit, P.; Singh, S.; Abraham, P.; Panda, S.; et al. SARS-CoV-2 Spike Mutations, L452R, T478K, E484Q and P681R, in the Second Wave of COVID-19 in Maharashtra, India. *Microorganisms* **2021**, *9*, 1542. [CrossRef] [PubMed] [PubMed Central]
24. Rockett, R.; Basile, K.; Maddocks, S.; Fong, W.; Agius, J.E.; Johnson-Mackinnon, J.; Arnott, A.; Chandra, S.; Gall, M.; Draper, J.; et al. Resistance Mutations in SARS-CoV-2 Delta Variant after Sotrovimab Use. *N. Engl. J. Med.* **2022**, *386*, 1477–1479. [CrossRef] [PubMed] [PubMed Central]
25. Lohr, B.; Niemann, D.; Verheyen, J. Bamlanivimab Treatment Leads to Rapid Selection of Immune Escape Variant Carrying the E484K Mutation in a B.1.1.7-Infected and Immunosuppressed Patient. *Clin. Infect. Dis.* **2021**, *73*, 2144–2145. [CrossRef] [PubMed]
26. Harari, S.; Miller, D.; Fleishon, S.; Burstein, D.; Stern, A. Using big sequencing data to identify chronic SARS-Coronavirus-2 infections. *Nat. Commun.* **2024**, *15*, 648. [CrossRef]
27. Machkovech, H.M.; Hahn, A.M.; Wang, J.G.; Grubaugh, N.D.; Halfmann, P.J.; Johnson, M.C.; E Lemieux, J.; O'COnnor, D.H.; Piantadosi, A.; Wei, W.; et al. Persistent SARS-CoV-2 Infection: Significance and Implications. *Lancet Infect. Dis.* **2024**, *24*, E453–E462. [CrossRef] [PubMed]
28. Garcia, I.; Bråte, J.; Fossum, E.; Rohringer, A.; Moen, L.V.; Hungnes, O.; Fjære, O.; Zaragkoulias, K.; Bragstad, K. Recombinant SARS-CoV-2 Delta/Omicron BA.5 emerging in an immunocompromised long-term infected COVID-19 patient. *Sci. Rep.* **2024**, *14*, 25790. [CrossRef] [PubMed] [PubMed Central]
29. Destras, G.; Assaad, S.; Bal, A.; Bouscambert-Duchamp, M.; Avrillon, V.; Simon, B.; Valette, M.; Blay, J.Y.; Lina, B.; Frobert, E.; et al. Bamlanivimab as monotherapy in two immunocompromised patients with COVID-19. *Lancet Microbe* **2021**, *2*, e424. [CrossRef] [PubMed] [PubMed Central]
30. Snell, L.B.; Alcolea-Medina, A.; Charalampous, T.; Alder, C.; Williams, T.G.S.; Flaviani, F.; Batra, R.; Bakrania, P.; Thangarajah, R.; Neil, S.J.D.; et al. Real-Time Whole Genome Sequencing to Guide Patient-Tailored Therapy of Severe Acute Respiratory Syndrome Coronavirus 2 Infection. *Clin. Infect. Dis.* **2023**, *76*, 1125–1128. [CrossRef] [PubMed] [PubMed Central]
31. Weigang, S.; Fuchs, J.; Zimmer, G.; Schnepf, D.; Kern, L.; Beer, J.; Luxenburger, H.; Ankerhold, J.; Falcone, V.; Kemming, J.; et al. Within-host evolution of SARS-CoV-2 in an immunosuppressed COVID-19 patient as a source of immune escape variants. *Nat. Commun.* **2021**, *12*, 6405. [CrossRef] [PubMed] [PubMed Central]
32. Ling-Hu, T.; Simons, L.M.; Rios-Guzman, E.; Carvalho, A.M.; Agnes, M.F.R.; Alisoltanidehkordi, A.; Ozer, E.A.; Lorenzo-Redondo, R.; Hultquist, J.F. The impact of remdesivir on SARS-CoV-2 evolution in vivo. *JCI Insight* **2025**, *10*, e182376. [CrossRef] [PubMed] [PubMed Central]
33. Iriyama, C.; Ichikawa, T.; Tamura, T.; Takahata, M.; Ishio, T.; Ibata, M.; Kawai, R.; Iwata, M.; Suzuki, M.; Adachi, H.; et al. Clinical and molecular landscape of prolonged SARS-CoV-2 infection with resistance to remdesivir in immunocompromised patients. *PNAS Nexus* **2025**, *4*, pgaf085. [CrossRef] [PubMed] [PubMed Central]
34. Breuer, J.; Drysdale, M.; Walker, J.; Han, J.; Aylott, A.; Van Dyke, M.K.; Birch, H.J.; McKie, E.; Jordan, W.; Gemzoe, K.; et al. Monitoring the emergence of resistance with sotrovimab in immunocompromised patients with COVID-19: LUNAR study. *J. Infect.* **2025**, *91*, 106510. [CrossRef] [PubMed]
35. Ichikawa, T.; Tamura, T.; Nao, N.; Suzuki, H.; Maruyama, S.; Wada, D.; Tsujino, S.; Yoshinaga, N.; Asagoe, K.; Takahata, M.; et al. Characterization of remdesivir resistance mutations in COVID-19 patients with various immunosuppressive diseases. *Antivir. Res.* **2025**, *242*, 106264. [CrossRef] [PubMed]
36. Centers for Disease Control and Prevention. Research Use Only 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Primers and Probes. 2020. Available online: https://www.cdc.gov/locs/2020/information_about_emergency_use_authorization_for_2019_novel_coronavirus_real_time_rt-pcr_diagnostic_panel.html (accessed on 3 May 2021).
37. O'Toole, Á.; Hill, V.; Pybus, O.G.; Watts, A.; Bogoch, I.I.; Khan, K.; Messina, J.P.; Tegally, H.; Lessells, R.R.; Giandhari, J.; et al. Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2. *Wellcome Open Res.* **2021**. [CrossRef]

38. Birnie, E.; Biemond, J.J.; Appelman, B.; de Bree, G.J.; Jonges, M.; Welkers, M.R.A.; Wiersinga, W.J. Development of Resistance-Associated Mutations After Sotrovimab Administration in High-risk Individuals Infected with the SARS-CoV-2 Omicron Variant. *JAMA* **2022**, *328*, 1104–1107. [CrossRef] [PubMed] [PubMed Central]
39. Braun, K.M.; Moreno, G.K.; Wagner, C.; Accola, M.A.; Rehrauer, W.M.; Baker, D.A.; Koelle, K.; O'Connor, D.H.; Bedford, T.; Friedrich, T.C.; et al. Acute SARS-CoV-2 infections harbor limited within-host diversity and transmit via tight transmission bottlenecks. *PLoS Pathog.* **2021**, *17*, e1009849. [CrossRef] [PubMed] [PubMed Central]
40. Bugatti, A.; Filippini, F.; Messali, S.; Giovanetti, M.; Ravelli, C.; Zani, A.; Ciccozzi, M.; Caruso, A.; Caccuri, F. The D405N Mutation in the Spike Protein of SARS-CoV-2 Omicron BA.5 Inhibits Spike/Integrins Interaction and Viral Infection of Human Lung Microvascular Endothelial Cells. *Viruses* **2023**, *15*, 332. [CrossRef] [PubMed] [PubMed Central]
41. Gómez, S.A.; Rojas-Valencia, N.; Gómez, S.; Cappelli, C.; Restrepo, A. The Role of Spike Protein Mutations in the Infectious Power of SARS-COV-2 Variants: A Molecular Interaction Perspective. *Chembiochem* **2022**, *23*, e202100393. [CrossRef] [PubMed] [PubMed Central]
42. Carabelli, A.M.; Peacock, T.P.; Thorne, L.G.; Harvey, W.T.; Hughes, J.; COVID-19 Genomics UK Consortium; Peacock, S.J.; Barclay, W.S.; de Silva, T.I.; Towers, G.J.; et al. SARS-CoV-2 variant biology: Immune escape, transmission and fitness. *Nat. Rev. Microbiol.* **2023**, *21*, 162–177. [CrossRef]
43. Barton, M.I.; MacGowan, S.A.; Kutuzov, M.A.; Dushek, O.; Barton, G.J.; Van Der Merwe, P.A. Effects of common mutations in the SARS-CoV-2 Spike RBD and its ligand, the human ACE2 receptor on binding affinity and kinetics. *eLife* **2021**, *10*, e70658. [CrossRef]
44. Gliga, S.; Luebke, N.; Killer, A.; Gruell, H.; Walker, A.; Dilthey, A.; Lohr, C.; Flaßhove, C.; Orth, H.; Feldt, T.; et al. Longitudinal Analysis of Escape Mutations in SARS-CoV-2 Omicron Variants after Sotrovimab Use in Immunodeficient Patients. *SSRN Electron. J.* **2022**. [CrossRef]
45. Ragonnet-Cronin, M.; Nutalai, R.; Huo, J.; Dijokaite-Guraliuc, A.; Das, R.; Tuekprakhon, A.; Supasa, P.; Liu, C.; Selvaraj, M.; Groves, N.; et al. Generation of SARS-CoV-2 escape mutations by monoclonal antibody therapy. *Nat. Commun.* **2023**, *14*, 3334. [CrossRef]
46. Gliga, S.; Lübke, N.; Killer, A.; Gruell, H.; Walker, A.; Dilthey, A.T.; Thielen, A.; Lohr, C.; Flaßhove, C.; Krieg, S.; et al. Rapid Selection of Sotrovimab Escape Variants in Severe Acute Respiratory Syndrome Coronavirus 2 Omicron-Infected Immunocompromised Patients. *Clin. Infect. Dis.* **2023**, *76*, 408–415. [CrossRef] [PubMed] [PubMed Central]
47. Cox, M.; Peacock, T.P.; Harvey, W.T.; Hughes, J.; Wright, D.W.; Willett, B.J.; Thomson, E.; Gupta, R.K.; Peacock, S.J.; Robertson, D.L.; et al. SARS-CoV-2 variant evasion of monoclonal antibodies based on in vitro studies. *Nat. Rev. Microbiol.* **2023**, *21*, 112–124. [CrossRef] [PubMed] [PubMed Central]
48. Evans, J.P.; Qu, P.; Zeng, C.; Zheng, Y.M.; Carlin, C.; Bednash, J.S.; Lozanski, G.; Mallampalli, R.K.; Saif, L.J.; Oltz, E.M.; et al. Neutralization of the SARS-CoV-2 Deltacron and BA.3 Variants. *N. Engl. J. Med.* **2022**, *386*, 2340–2342. [CrossRef] [PubMed] [PubMed Central]
49. Kurhade, C.; Zou, J.; Xia, H.; Liu, M.; Yang, Q.; Cutler, M.; Cooper, D.; Muik, A.; Sahin, U.; Jansen, K.U.; et al. Neutralization of Omicron sublineages and Deltacron SARS-CoV-2 by three doses of BNT162b2 vaccine or BA.1 infection. *Emerg. Microbes Infect.* **2022**, *11*, 1828–1832. [CrossRef] [PubMed] [PubMed Central]
50. Duerr, R.; Zhou, H.; Tada, T.; Dimartino, D.; Marier, C.; Zappile, P.; Wang, G.; Plitnick, J.; Griesemer, S.B.; Girardin, R.; et al. Delta-Omicron recombinant escapes therapeutic antibody neutralization. *iScience* **2023**, *26*, 106075. [CrossRef] [PubMed] [PubMed Central]
51. Focosi, D.; Casadevall, A.; Franchini, M.; Maggi, F. Sotrovimab: A Review of Its Efficacy against SARS-CoV-2 Variants. *Viruses* **2024**, *16*, 217. [CrossRef] [PubMed] [PubMed Central]
52. Schoefbaenker, M.; Günther, T.; Lorentzen, E.U.; Romberg, M.L.; Hennies, M.T.; Neddermeyer, R.; Müller, M.M.; Mellmann, A.; Bojarzyn, C.R.; Lenz, G.; et al. Characterisation of the antibody-mediated selective pressure driving intra-host evolution of SARS-CoV-2 in prolonged infection. *PLoS Pathog.* **2024**, *20*, e1012624. [CrossRef] [PubMed] [PubMed Central]
53. Chen, D.; Liu, G.; Du, H.; Jones, B.; Wee, J.; Wang, R.; Chen, J.; Shen, J.; Wei, G.W. Drug Resistance Predictions Based on a Directed Flag Transformer. *arXiv* **2025**, arXiv:2403.02603v2. [CrossRef] [PubMed] [PubMed Central]
54. Liu, C.; Shi, W.; Becker, S.T.; Schatz, D.G.; Liu, B.; Yang, Y. Structural basis of mismatch recognition by a SARS-CoV-2 proofreading enzyme. *Science* **2021**, *373*, 1142–1146. [CrossRef] [PubMed] [PubMed Central]
55. Rosas-Lemus, M.; Minasov, G.; Shuvalova, L.; Inniss, N.L.; Kiryukhina, O.; Brunzelle, J.; Satchell, K.J.F. High-resolution structures of the SARS-CoV-2 2′-O-methyltransferase reveal strategies for structure-based inhibitor design. *Sci. Signal.* **2020**, *13*, eabe1202. [CrossRef] [PubMed] [PubMed Central]
56. The Nextstrain Team. Genomic Epidemiology of SARS-CoV-2 with Subsampling Focused Globally Since Pandemic Start. Available online: https://nextstrain.org/ncov/gisaid/global/all-time?c=gt-ORF1a_4388&gmax=13468&gmin=266 (accessed on 7 July 2025).

57. Sun, C.; Xie, C.; Bu, G.-L.; Zhong, L.-Y.; Zeng, M.-S. Molecular characteristics, immune evasion, and impact of SARS-CoV-2 variants. *Signal Transduct. Target. Ther.* **2022**, *7*, 202. [CrossRef]
58. Tao, K.; Tzou, P.L.; Nouhin, J.; Gupta, R.K.; de Oliveira, T.; Kosakovsky Pond, S.L.; Fera, D.; Shafer, R.W. The biological and clinical significance of emerging SARS-CoV-2 variants. *Nat. Rev. Genet.* **2021**, *22*, 757–773. [CrossRef]
59. Joseph, R.; Marais, G.; Iranzadeh, I.; Alisoltani, A.; Hardie, D.; Davies, M.-A.; Heekes, A.; Chetty, N.; Timmerman, V.; Hsiao, N.-Y.; et al. Intra-host SARS-CoV-2 diversity in immunocompromised people living with HIV provides insight into the evolutionary trajectory of SARS-CoV-2. *J. Virol.* **2025**, e0078025. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.