Developing Methods to Understand Within-Host Evolution and the Effect of Antiviral Drugs on RNA Viruses

Wan Man Juanita Pang

A thesis submitted in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Division of Infection and Immunity

University College London

Oct 2022

Declaration

I, Wan Man Juanita Pang, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Publications

Publications arising from this thesis:

Pang J, Boshier FAT, Alders N, Dixon G, Breuer J. No evidence of viral polymorphisms associated with Paediatric Inflammatory Multisystem Syndrome Temporally Associated With SARS-CoV-2 (PIMS-TS). (2020) *Pediatrics, 146 (6) e2020019844.* https://doi.org/10.1542/peds.2020-019844

Pang J, Slyker JA, Roy S, Bryant J, Atkinson C, Cudini J, Farquhar C, Griffiths P, Kiarie J, Morfopoulou S, Roxby A, Tutil H, Williams R, Gantt S, Goldstein RA, Breuer J. Haplotype reconstruction of cytomegalovirus genomes reveals similarities in viruses transmitted congenitally by HIV-positive mothers. (2020) *eLife* 2020;9:e63199. https://doi.org/10.7554/eLife.63199

Venturini C, **Pang J**, Tamuri AU, Roy S, Breuer J, Goldstein RA. Haplotype assignment of longitudinal viral deep-sequencing data using co-variation of variant frequencies**.** (2022) *Virus Evolution, veac093.* https://doi.org/10.1093/ve/veac093

Boshier, F.A.T., **Pang, J.**, Penner, J., Hughes, J., Parker, M., Shepherd, J., Alders, N., Bamford, A., Grandjean, L., Grunewald, S., Hatcher, J., Best, T., Dalton, C., Bynoe, P.D., Frauenfelder, C., Köeglmeier, J., Myerson, P., Roy, S., Williams, R., The COVID-19 Genomics UK (COG-UK) consortium, Thomson, E.C., de Silva, T.I., Goldstein, R.A., Breuer, J. Evolution of viral variants in remdesivir-treated and untreated SARS-CoV-2-infected pediatrics patients. (2021) *Journal of Medical Virology.* https://doi.org/10.1002/jmv.27285

Sanchez Clemente N, **Pang J**, Rodrigues C et al. Case Report: severe paediatric COVID-19 pneumonitis treated with remdesivir and nitazoxanide. (2021) *Wellcome Open Research.* 6:329. https://doi.org/10.12688/wellcomeopenres.17377.1

Gastine S, **Pang J**, Boshier FAT, Carter S, Lonsdale D, Cortina-Borja M, Hung I, Breuer J, Kloprogge F, Standing JF. Systematic review and patient-level meta-

analysis of SARS-CoV-2 viral dynamics to model response to antiviral therapies. (2021) *Clinical Pharmacology & Therapeutics.* https://doi.org/10.1002/cpt.2223

Kaptein, S.J.F., Jacobs, S., Langendries, L., Seldeslachts, L., Ter Horst, S., Liesenborghs, L., Hens, B., Vergote, V., Heylen, E., Barthelemy, K., Maas, E., De Keyzer, C., Bervoets, L., Rymenants, J., Van Buyten, T., Zhang, X., Abdelnabi, R., **Pang, J**., Williams, R., Thibaut, H.J., Dallmeier, K., Boudewijns, R., Wouters, J., Augustijns, P., Verougstraete, N., Cawthorne, C., Breuer, J., Solas, C., Weynand, B., Annaert, P., Spriet, I., Vande Velde, G., Neyts, J., Rocha-Pereira, J., Delang, L. Favipiravir at high doses has potent antiviral activity in SARS-CoV-2-infected hamsters, whereas hydroxychloroquine lacks activity. (2020) *Proc Natl Acad Sci U S A.* https://doi.org/10.1073/pnas.2014441117

Abdelnabi, R., Foo, C.S., Kaptein, S.J.F., Zhang, X., Langendries, L., Vangeel, L., Breuer, J., **Pang, J.**, Williams, R., Vergote, V., Heylen, E., Leyssen, P., Dallmeier, K., Coelmont, L., Jochmans, D., Chatterjee, A.K., De Jonghe, S., Weynand, B., Neyts, J. The combined treatment of Molnupiravir and Favipiravir results in a marked potentiation of antiviral efficacy in a SARS-CoV-2 hamster infection model. (2021) *eBioMedicine.* https://doi.org/10.1016/j.ebiom.2021.103595

Publication and Preprints not directly related to this thesis:

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., Ortiz, A.T., Balloux, F. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. (2020) *Infection, Genetics and Evolution.* https://doi.org/10.1016/j.meegid.2020.104351

Guerra-Assunção, J.A., Randell, P.A., Boshier, F.A.T., Crone, M.A., **Pang, J.**, Mahungu, T., Freemont, P.S., Breuer, J. Reliability of Spike Gene Target Failure for ascertaining SARS-CoV-2 lineage B.1.1.7 prevalence in a hospital setting. (2020) *medRxiv.* https://doi.org/10.1101/2021.04.12.21255084

van Dorp, L., Nimmo, C., Ortiz, A.T., **Pang, J.**, Acman, M., Tan, C.C.S., Millard, J., Padayatchi, N., Grant, A., O'Donnell, M., Pym, A., Brynildsrud, O.B., Eldholm, V., Grandjean, L., Didelot, X., Balloux, F. Detection of a bedaquiline / clofazimine resistance reservoir in *Mycobacterium tuberculosis* predating the antibiotic era. (2020) *bioRxiv.* https://doi.org/10.1101/2020.10.06.328799

Statement of Work

Chapter 2

Some of the work and text in Chapter 2 was published in Gastine et al., 2020, on which I was a second author. The mathematical model described in Section 2.5.2 was mainly developed by Dr Silke Gastine and Professor Joseph Standing. All other work including the systematic literature search, data processing and analysis was jointly performed by myself and Dr Florencia A.T. Boshier.

Chapter 3

Some of the work and text in Chapter 3 was published in Pang et al., 2020b, on which I was a co-first author. The program HaROLD was written by Dr Asif U. Tamuri and Prof Richard A. Goldstein. All synthetic and clinical data sets for programs evaluation were prepared by myself. The evaluation of HaROLD and other haplotype reconstruction programs was jointly performed by myself and Dr Cristina Venturini. The clinical application of HaROLD on norovirus and cytomegalovirus was done predominantly by myself. Some of the work and text of the clinical application of HaROLD on cytomegalovirus was published in Pang et al., 2020a, on which I was a co-first author.

Chapter 4

Some of the work and text on the remdesivir treated SARS-CoV-2 patients analysis in Section 4.3.7 was published in Boshier et al., 2020a, on which I was a co-first author. This work was jointly done by myself and Dr Florencia A.T. Boshier. Some of the work on the first clinical case on a nitazoxanide and remdesivir-treated SARS-CoV-2 patient presented in Section 4.4.8 was published in Sanchez Clemente et al., 2021, on which I was a co-first author. The viral load and viral genome sequence analysis on this patient was done by myself. The structural analysis of favipiravir associated mutations in Section 4.3.4 was jointly done by myself and Oscar J. Charles. The zebrafish larvae experimental work described in Section 4.3.5 was performed by Dr Joana Duarte Da Rocha Pereira and Emma Roux. Viral sequence analysis of samples taken from the zebrafish larvae was carried out by myself. The Syrian hamsters experimental work described in Section 4.3.9 was performed by Dr Joana Duarte Da Rocha Pereira, Dr Rana Abdelnabi and their research group. Some of the work was published in Abdelnabi et al., 2021; Kaptein et al., 2020a, on which I was a co-author. All clinical details presented in this chapter were collated by clinicians at Great Ormond Street Hospital and Royal Free Hospital unless otherwise stated. All other work was performed by myself.

Acknowledgement

First, I would like to express my deepest gratitude to my supervisors, Prof Judith Breuer and Prof Richard Goldstein, for their support, guidance and expertise throughout my PhD. I have greatly enjoyed my time working with them. Thank you for providing me with the valuable opportunity to work as an undergraduate research intern student 5 years ago. You have introduced me to the world of scientific research and have successfully passed on your enthusiasm for pathogen genomics and evolution.

I would also like to thank Rosetrees Trust for funding my research.

Special thanks to Dr Florencia Tettamanti Boshier for always listening, helping, and encouraging me. I genuinely loved every moment working with you.

Thank you to Dr Sunando Roy and Dr Cristina Venturini for introducing me to bioinformatics during my time as an undergraduate student and generously sharing very useful tips and pipelines in the past few years.

Thank you to my fellow PhD student Oscar Charles as well as Dr Jose Afonso Guerra-Assuncao and Dr Sofia Morfopoulou from the Breuer and Goldstein lab, for supporting me throughout my PhD. You made me laugh during the most frustrating moments.

Last but not least, I would like to thank my family and friends, and especially my partner, Jack, for his constant support and assistance.

Abstract

Viral infections are common and are particularly problematic in immunocompromised individuals. However, other than for HIV, Hepatitis B, Hepatitis C, Influenza, and more recently SARS-CoV-2, there have been few approved drugs available for treating viral infections. Instead, repurposed drugs are often used, especially at the beginning of the current pandemic, for treating SARS-CoV-2. It remains unclear how these repurposed drugs act on the viral population and whether the suppression of viral load we observe is attributed to the drug or the immune response or a combination of both.

The research presented in this thesis primarily focuses on the study of two RNA viruses, SARS-CoV-2 and Norovirus. A mixture of viral load data and viral genomic data were analysed to understand the course of infection within individuals. First, we presented a meta-analysis on SARS-CoV-2 viral load dynamics where we investigated the changes of viral dynamics over time, with and without the presence of antiviral drugs. Then, we presented an evolutionary model used for reconstructing haplotypes in mixed infections. Finally, we demonstrated the use of viral deep sequencing to study the within-host evolution of RNA viruses. We identified mutagenic signatures and consensus level changes associated with antiviral treatments. We developed unique methods to analyse viral sequences which allow us to understand the within-host genomic variations and hence inform our understanding of the heterogeneous efficacy of a drug between patients.

Overall, this thesis provides insights into how the efficacy of a drug can be evaluated by monitoring the within-host viral dynamics and evolution.

Impact Statement

RNA viruses are one of the key pathogen groups responsible for zoonotic disease transmission and they transmit rapidly among humans. They have caused numerous epidemics and pandemics in human history, which led to significant clinical and economic burdens in many countries all over the world. Their ability to replicate and mutate rapidly poses a challenge to vaccine and treatment development. During the COVID-19 pandemic, the unprecedented collaborative effort in the scientific community has advanced our understanding of the dynamics of viral infections. Yet, the within-host variations in most RNA viral infections have not been fully characterised. The efficacy of frequently used broad-spectrum antiviral drugs remain poorly understood.

This thesis focuses on studying the within-host viral variations in two RNA viruses, norovirus and SARS-CoV-2. Beginning with Chapter 2, the SARS-CoV-2 viral load dynamic meta-analysis was conducted during the early stage of the pandemic, in May 2020. The analysis of the viral load dynamic over time provided a fundamental understanding of the infection. It highlighted the characteristics of viral load dynamics in SARS-CoV-2 and provided insights which were useful for setting clinical guidance at the time. The results mostly echoed what was found in later studies.

The use of viral whole genome sequencing was then explored in Chapter 3. The haplotype reconstruction evolutionary model presented allows in-depth study of within-host viral population even in complex infections with mixed viral strains. The wider application of the method

was demonstrated using a norovirus clinical data set and a cytomegalovirus, a DNA virus, clinical data set.

Finally, in Chapter 4, building on the application of the haplotype reconstruction method described in Chapter 3, viral whole genome sequencing was used to deeply characterise norovirus and SARS-CoV-2 infections in both untreated and treated individuals. The drug-associated mutations were identified. The efficacy of commonly used antiviral drugs such as favipiravir and remdesivir was investigated. The results provided insights into the use of antiviral treatments. The methods used in this chapter can potentially be applied to study other RNA viral infections.

Future pandemics are difficult to predict, but there are increasing concerns that RNA viruses might cause more pandemics in the future. The work presented in this thesis provides valuable insights into within-host variations in RNA viral infections which are essential for the preparation of future pandemics.

Table of Contents

List of Figures and Tables

Chapter 1

Chapter 2

Chapter 3

Figure 2. Violin plots showing the accuracy of haplotype reconstruction in the HCMV test set.

Figure 3. Accuracy of sample diversity estimations based on reconstructed haplotypes for norovirus test set (A) and HCMV test set (B).

Figure 4. Pairwise genetic distances calculated as proportion.

Figure 5. Multi-dimensional scaling (MDS) of reconstructed haplotypes from different methods.

Figure 6. Maximum likelihood phylogenetic tree of HaROLD reconstructed haplotypes from a patient infected with norovirus.

Figure 7. Bar plot of estimated abundances over time of HaROLD reconstructed haplotypes from a patient infected with norovirus.

Figure 8. Multi-dimensional scaling (MDS) comparing all sequences obtained with different methods from a patient infected with norovirus.

Figure 9. Maximum likelihood phylogenetic tree of (A) PredictHaplo, (B) CliqueSNV and (C) EVORhA reconstructed haplotypes from a patient infected with norovirus.

Figure 10. HCMV viral loads of longitudinal samples for each family from breast milk (red), baby blood spots (green), and cervix (blue), and HIV viral loads from mother's blood plasma.

Figure 11. Abundance of haplotypes within each sample for breast milk (BM), cervix (CV), and blood spots (BS).

Figure 12. Boxplot showing the number and frequencies of haplotypes reconstructed by HaROLD in relation to read depth. Analysis was performed on a mother's breast milk sample from family 12, which has an average read depth of 780.

Chapter 4

Chapter 1

Introduction

1.1 RNA Viruses

1.1.1 Overview

RNA viruses constitute a diverse class of viruses that infect eukaryotic, prokaryotic and archaeal hosts. They can cause a variety of common human infectious diseases, including COVID-19, common cold, norovirus, influenza, hepatitis, Ebola virus disease, West Nile fever, polio, and measles (Binder et al., 1999; Payne, 2017). Studies have shown that RNA viruses are responsible for up to 44% of emerging infectious diseases (Carrasco-Hernandez et al., 2017). On average two to three novel RNA viruses are discovered each year (Rosenberg, 2015). With the increase in interspecies contact, human population and globalisation, RNA viruses have become major zoonotic agents which cause recurring hazard to the global public health (Carrasco-Hernandez et al., 2017; Reperant and Osterhaus, 2017; Woolhouse et al., 2016).

The populations of RNA viruses harbour abundant genetic variability which is partly explained by four basic properties:

1) A large population size. For example, the peak number of viral particles in SARS-CoV-2 infection might be as high as 10^{11} (Sender et al., 2021).

2) A short viral generation time. A single virion particle can produce an average of 100,000 viral copies in 10 hours (Moya et al., 2004).

3) The lack of proof-reading activity in the genome replication process (except for coronaviruses). The RNA viral replication process is prone to error. They can generate a maximum of one mutation per genome, per replication (Drake, 1993; Malpica et al., 2002; Peck and Lauring, 2018).

4) A small genome size. RNA viruses have a genome size ranging from around 3 to 30 kilobases, with a median size of 9 kilobases (Moya et al., 2004). The genome size is negatively correlated with the mutation rate (Drake, 1991). A small genome size allows maximal adaptability (Bradwell et al., 2013).

As the same with all organisms, the major mechanisms of evolution in RNA viruses include genetic drift and natural selection. With their high mutation rates, strong purifying selection drives rapid adaptive evolution. They can rapidly adapt to environmental changes. Fittest variants will be selected by the host immune pressure and antiviral drugs used (Dolan et al., 2018).

1.1.2 Taxonomy

According to the International Committee on Taxonomy of Viruses, viruses in Groups III, IV and V in the Baltimore classification have a ribonucleic acid genome and are classified as RNA viruses (Figure 1). Based on their genome organisation, they can be further classified into positive (Group IV) or negative-sense (Group V) single-stranded RNA viruses, or double-stranded (Group III) RNA viruses (Poltronieri et al., 2015) (Figure 2). The genomes of positive-sense single-stranded RNA viruses can be read and translated into proteins directly whereas the genomes of negative-sense single-stranded or double-stranded RNA viruses need to be first transcribed before being translated into proteins.

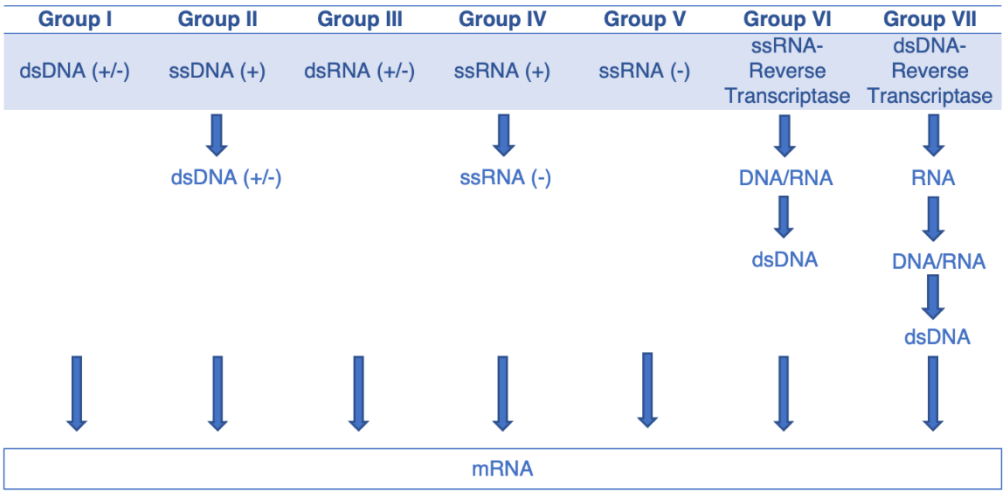


Figure 1. The Baltimore classification, a virus classification scheme based on the viral genetic materials. *Groups III, IV and V are classified as RNA viruses. Group VI viruses, such as HIV, are retroviruses that have RNA viral genomes. Since they use DNA genetic materials during their life cycle, they are not classified as RNA viruses. They will not be discussed in this thesis.*

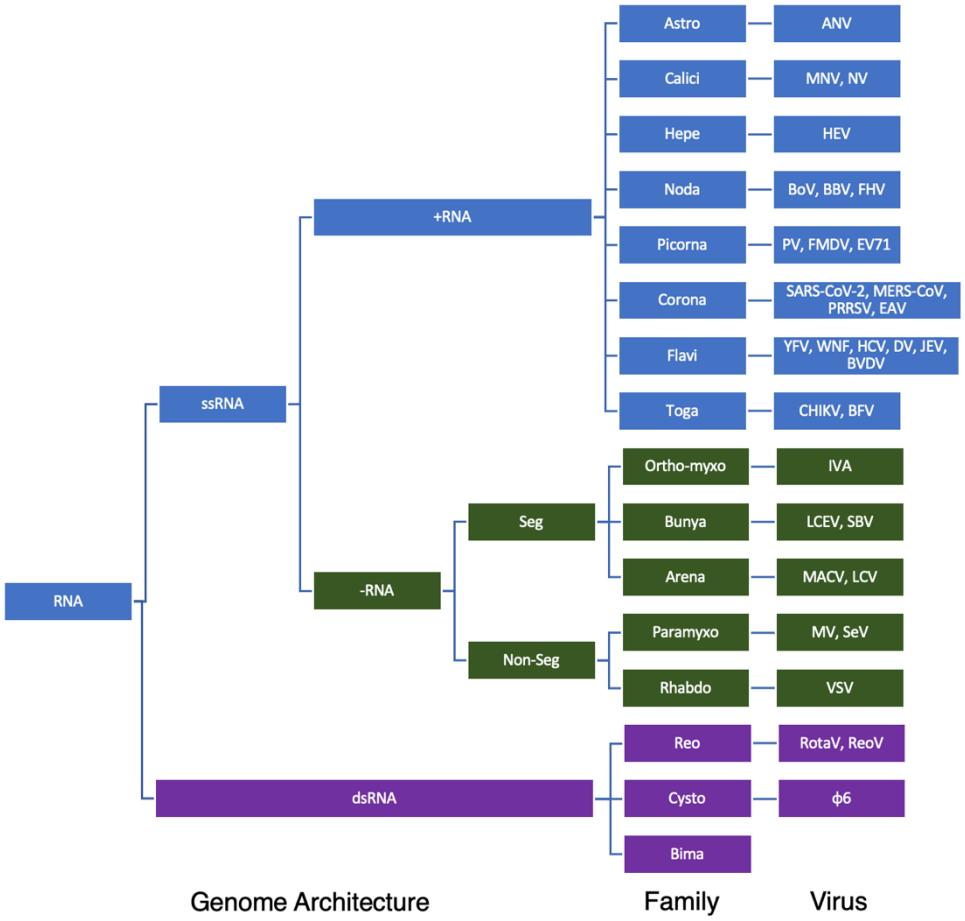


Figure 2. Taxonomy of RNA viruses (Group III, IV and V of Baltimore classification). *Examples of some RNA viral families and species are included. ssRNA = single-stranded RNA viruses, dsRNA = double-stranded RNA viruses (Group III). +RNA = positive-sense single-stranded RNA viruses (Group IV), -RNA = negative-sense single-stranded RNA viruses (Group V). Seg/ non-seg = segmented or non-segmented genome.*

1.1.3 Positive-Sense Single-Stranded RNA viruses

In this thesis, we will discuss two positive-sense single-stranded RNA viruses, namely norovirus and SARS-CoV-2.

As of 2021, 63 viral families are known to have a positive-sense single-stranded RNA genome (International Committee on Taxonomy of Viruses, 2021). Of which, nine major families are known to infect vertebrates (Modrow et al., 2013). Five of the families, *Astroviridae*, *Caliciviridae, Hepeviridae, Nodaviridae* and *Picornaviridae*, have non-enveloped capsids, where another four families, *Arteriviridae, Coronaviridae, Flaviviridae* and *Togaviridae* have enveloped capsids (Modrow et al., 2013) (Figure 3).

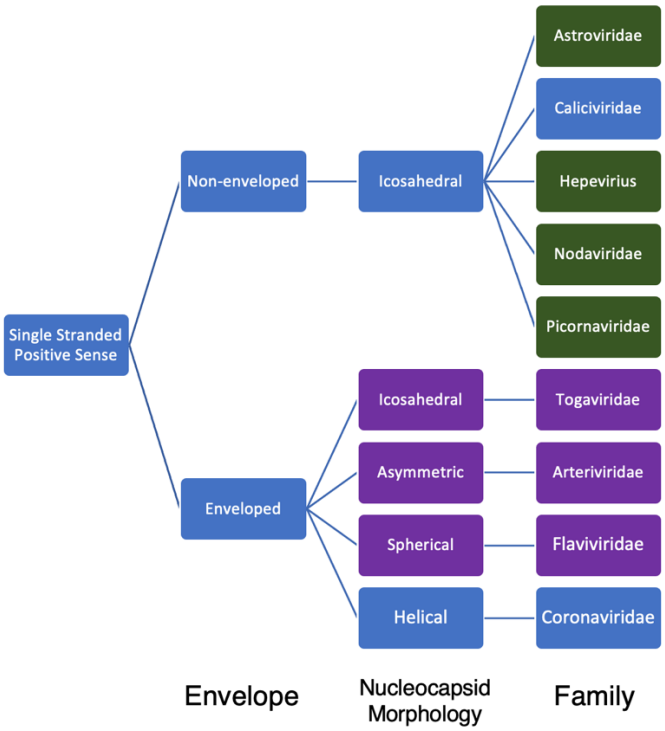


Figure 3. Classification of positive-sense single-stranded RNA viruses. *They can be classified based on their nucleocapsid morphology and whether they are enveloped. In this thesis, we will discuss norovirus and SARS-CoV-2. They belong to the Caliciviridae and Coronaviridae respectively (highlighted in blue).*

The genomes of positive-sense single-stranded RNA viruses function as messenger RNAs (mRNA), which are directly translated into one or more polyproteins. They typically have a viral protein genome-linked 5' end, with an untranslated region (UTR), followed by multiple open reading frames, and a 3' terminus with polyadenosine monophosphate tail (Yuanzhi Liu et al., 2020). Once the virus entered the host cell, the genomic RNA will attach to the host ribosome and subsequently be translated into multiple polyproteins. These polyproteins are cleaved by either viral or cellular enzymes to produce various structural and non-structural proteins. The virus-encoded RNA dependent RNA polymerase (RdRp) will use the positive-strand genomic RNA as a template to synthesise a negative-strand genomic RNA, which is then used to synthesise more positive-strand genomic RNAs (Yuanzhi Liu et al., 2020).

However, for some positive-strand RNA viruses, such as those in the *Coronaviridae* and *Caliciviridae* families, instead of having the full genomic RNA directly translated into both structural and non-structural proteins, only non-structural proteins, including the RdRp, are produced at first. A replication complex which includes all viral encoded replication enzymes is then formed. The genomic RNA is replicated into a whole antigenome and progeny genomes, and subsequently a subgenomic mRNA that encodes for structural proteins will be produced and translated (Long, 2021; Smertina et al., 2019).

Further details on genome structure and the process of genome replication and translation will be discussed separately for norovirus and SARS-CoV-2 in Sections 1.2.5 to 1.2.6 and 1.3.5 to 1.3.6.

1.1.4 RNA-Dependent RNA Polymerase

Structure

In general, all RNA viruses have similar replication mechanisms which rely on the RNA-dependent RNA polymerase (RdRp) (Rampersad and Tennant, 2018). The structure of RdRp, the key enzyme used for viral genome transcription and replication, is largely conserved across different RNA viruses (Ferrerorta et al., 2006; Mönttinen et al., 2014). Although the RdRp of positive and negative-sense RNA viruses have slightly different enzymatic modes of action with the divalent metal ions which coordinate the catalytic aspartates to facilitate the formation of phosphodiester bond between nucleoside triphosphates, (Venkataraman et al., 2018), the catalytic motifs in the RdRp palm and fingers domains are mostly conserved (Jia and Gong, 2019). Therefore, RdRp inhibiting antiviral drugs are often repurposed and used to treat different RNA viral infections (Choudhury et al., 2021; Simonis et al., 2021).

The RdRp has three structural subdomains, thumb, palm, and fingers, which resemble a cupped right hand and an N-terminal subdomain that connects the thumb and the fingers subdomains (Figure 4) (Bruenn, 2003; Venkataraman et al., 2018). The N-terminal subdomain, which is extremely conserved, acts as the active site of the RdRp and supports communication between other subdomains. The fingers subdomain holds the RNA template and facilitates polymerisation while the palm subdomain catalyses the phosphoryl transfer reaction, and the thumb subdomain stabilises the nucleoside triphosphate on the RNA template and supports the conformational changes and translocation of this

template after polymerisation (Ng et al., 2008; Venkataraman et al., 2018).

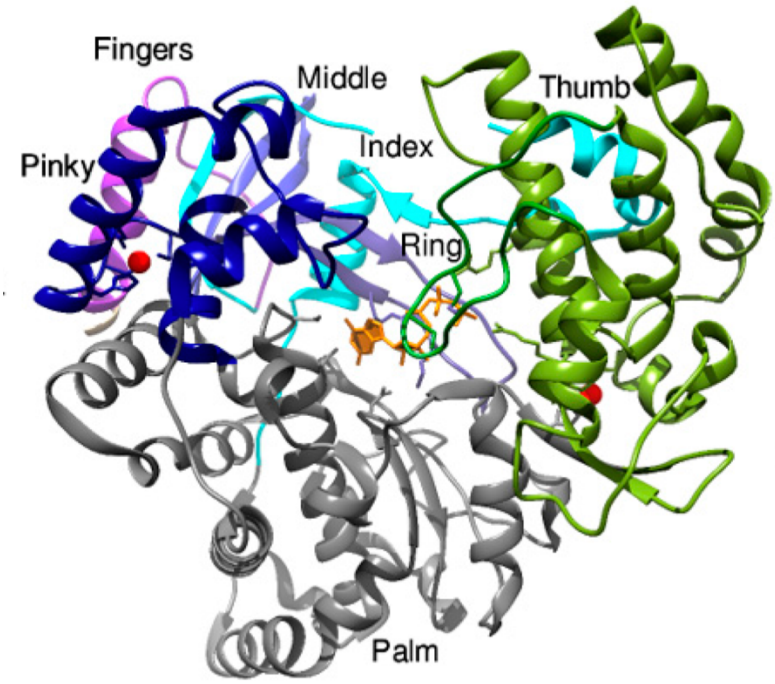


Figure 4. Structure of the RNA-dependent RNA polymerase of Japanese encephalitis virus (a positive-sense single-stranded RNA virus). *The thumb, palm and fingers subdomain are coloured in green, grey and shades of blue and purple respectively. The nucleotide undergoing catalysis in the active site in the palm domain is coloured in orange. Figure adapted from Venkataraman et al., 2018.*

Function

The catalytic mechanisms of most RdRps are similar. The RdRp catalyses the formation of phosphodiester bonds between ribonucleotides in the presence of divalent metal ions (Bruenn, 2003). It governs the synthesis of an RNA strand complementary to a given RNA

template (Venkataraman et al., 2018). RdRp initiates the formation of the replication complex with other non-structural proteins. During replication, it regulates the elongation of the RNA strand. It synthesises the viral RNA from the 3' end of an RNA template using either a primer-dependent or a primer-independent *de novo* mechanism. In primer-dependent synthesis, a short oligonucleotide or a protein that is covalently attached to a nucleotide is used as a primer (Zhu et al., 2020). In *de novo* synthesis, in the absence of a primer, the RdRp directly uses the nucleotides to begin synthesis (Gong, 2021). The replication mechanism will be discussed separately for norovirus and SARS-CoV-2 in Sections 1.2.6 and 1.3.6.

1.1.5 RdRp Inhibitors

As the structure of RdRp is conserved across different RNA viruses, it has become the key target for antiviral drug development. There are two classes of RdRp inhibitors, non-nucleoside analogue inhibitors (NNIs) and nucleoside analogue inhibitors (NIs) (Tian et al., 2021). NNIs bind to the RdRp at allosteric sites whereas NIs bind to the RdRp at the active site to terminate the RNA replication (Mittal et al., 2019).

Nucleoside Analogue Inhibitors

NIs are developed as prodrugs (Picarazzi et al., 2020). They are metabolised *in vivo* by the hepatic enzymes in the liver (Delang et al., 2013; Eltahla et al., 2015). The NIs are then incorporated into the growing RNA chain by the RdRp. It has been proposed that the 3'-hydroxyl group in NIs induce a steric effect which slows down or prevent bonding and eventually leads to termination of RNA synthesis (Clercq, 2007). However, some NIs can also act by inducing lethal mutagenesis (Bull et al., 2007). Examples of NIs include favipiravir, ribavirin, remdesivir and molnupiravir.

Favipiravir

Favipiravir is a NI developed for treating influenza (Baranovich et al., 2013; Furuta et al., 2013), but it also has proven efficacy against other unrelated RNA viruses such as enterovirus (Wang et al., 2016, p. 71), yellow fever virus (Delang et al., 2018; Gowen et al., 2010, p. 1106), Lassa virus (Rosenke et al., 2018), West Nile virus (Morrey et al., 2008) *in vitro*, SARS-CoV-2 (Driouich et al., 2021; Kaptein et al., 2020b) and Ebola (Guedj et al., 2018; Madelain et al., 2015; Oestereich et al., 2014),

both *in vitro* and *in vivo*. However, it is known to cause teratogenic and embryotoxic effects in mice which could lead to delay in the development or death of embryo (Evaluation and Licensing Division, Pharmaceutical and Food Safety Bureau, Japan, 2011). Two COVID-19 clinical trials have also shown that favipiravir treatment did not reduce the time to viral clearance and had no significant benefits in terms of mortality (Bosaeed et al., 2022; Hassanipour et al., 2021).

Favipiravir can act as a chain terminator after being incorporated into the growing RNA strand (Jin et al., 2013; Sangawa et al., 2013). However, more frequently, it acts as a mutagen by incorporating into both positive- and negative-stranded RNA during replication (Arias et al., 2014; Baranovich et al., 2013; Barauskas et al., 2017; de Ávila et al., 2016; Furuta et al., 2009; Goldhill et al., 2019). It is a purine analogue which primarily acts as a guanosine analogue and secondarily as an adenosine analogue (Jin et al., 2017, 2013; Sangawa et al., 2013). Studies have confirmed favipiravir causes a C to T and G to A mutagenic signature (Baranovich et al., 2013; Delang et al., 2014; Goldhill et al., 2019; Guedj et al., 2018; Vanderlinden et al., 2016). These mutations will accumulate over time. The viral population will go into extinction through a process described as lethal mutagenesis when the equilibrium mean fitness times the number of offspring of any non-mutated virus is less than one (Agostini et al., 2019; Bull et al., 2007; Goldhill et al., 2019; Shannon et al., 2020; Sheahan et al., 2020; Urakova et al., 2018; Vignuzzi et al., 2005; Yoon et al., 2018) (Figure 5).

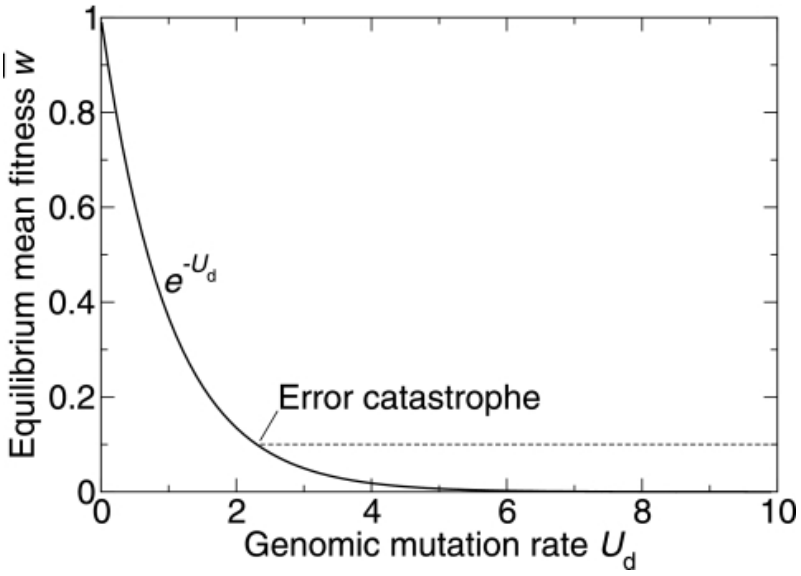


Figure 5. Equilibrium mean fitness value as a function of deleterious mutation rate during the process of lethal mutagenesis. *The dashed line indicates the error catastrophe threshold of the model described in (Bull et al., 2007). At a low mutation rate, fitness declines exponentially with increasing mutation rate, until it reaches the error catastrophe threshold. At a mutation rate above the threshold, the mean fitness remains unchanged because the remaining genotypes are insensitive to mutations. This slows down the extinction of the population. Figure adapted from Bull et al., 2007.*

In a more recent study, which uses Primer ID to measure the mutation rate and mutation bias in influenza polymerase *in vitro*, it has been shown that favipiravir induces not only C to T and G to A mutations, but also a lower rate of T to C and A to G mutations. As a guanosine or adenosine analogue, during the RNA synthesis, favipiravir can bind to either C or T, in place of G or A, on either strand. It then pairs with a T or C in the following synthesis cycle. The study has also shown that at a higher concentration of 100 μM of favipiravir, a reduction of mRNA was

observed, indicating favipiravir might cause a chain termination at a higher concentration (Goldhill et al., 2019).

Ribavirin

Ribavirin is mainly used to treat respiratory syncytial virus and hepatitis C in combination with other antivirals (Thomas et al., 2012). It has also demonstrated some efficacy against Lassa virus (Bausch et al., 2010), rabies (Hemachudha et al., 2013), and poxviruses (Baker et al., 2003). Although ribavirin has been shown to be effective in some SARS-CoV and MERS-CoV patients, it has lower efficacy against SARS-CoV-2 compared to other antiviral drugs (Arabi et al., 2020; Stockman et al., 2006). In addition, toxicity and severe side effects have been associated with ribavirin treatments (Kowdley, 2005).

Ribavirin is a purine analogue which could act as a guanosine or adenosine analogue (Wu et al., 2005). The exact mode of action of ribavirin is unknown (Nyström et al., 2019). However, several mechanisms have been proposed.

Ribavirin is converted into ribavirin monophosphate intracellularly, and subsequently dephosphorylated and triphosphorylated into ribavirin triphosphate (RTP). It has been reported that RTP, as a competitive inhibitor, can be incorporated into the nascent RNA strand which causes premature termination of the nascent RNA and significantly reduces the efficacy of RNA synthesis in influenza and poliovirus and hepatitis C virus (Crotty et al., 2000; Eriksson et al., 1977; Maag et al., 2001). However, RTP has a much lower binding affinity compared to a natural nucleotide (Vo et al., 2003).

It has also been proposed that ribavirin can limit the replication of viral genomes by inhibiting the inosine monophosphate dehydrogenase (IMPDH), an enzyme involved in *de novo* synthesis of guanine nucleotides (Malinoski and Stollar, 1981; Streeter et al., 1973; Wray et al., 1985).

In vitro and *in vivo* studies have demonstrated that ribavirin has an immunomodulatory effect. It can induce a shift from T helper cell 2 to T helper cell 1 immune response, which has been demonstrated to correlate with viral clearance in hepatitis C virus (Lau et al., 2002; Peavy et al., 1981; Rehermann and Nascimbeni, 2005).

Similar to favipiravir, another proposed mechanism of action of ribavirin is mutagenesis. The mutagenic nature of ribavirin has been reported in poliovirus (Crotty et al., 2000), West Nile virus (Day et al., 2005), Hantaan virus (Severson et al., 2003) and hepatitis C virus (Vo et al., 2003). However, it has been suggested that since ribavirin (RTP) is incorporated at a slower rate than natural nucleotide, when using ribavirin monotherapy, the frequency of polymorphisms is not high enough to induce lethal mutagenesis (Te et al., 2007).

Remdesivir

Remdesivir was developed for treating hepatitis C virus (Siegel et al., 2017), but it has a broad-spectrum activity against Ebola (E. Tchesnokov et al., 2019), coronaviruses (Frediansyah et al., 2021), respiratory syncytial virus and several paramyxoviruses (Lo et al., 2017). In a COVID-19 clinical trial, remdesivir has been shown to shorten the recovery time in hospitalised patients (Beigel et al., 2020a). Another more recent COVID-19 clinical trial reported that the early use of

remdesivir resulted in an 87% lower risk of hospitalisation or death (Gottlieb et al., 2022). However, remdesivir is associated with common side effects including respiratory failure (Piscoya et al., 2020), gastrointestinal distress (Li et al., 2020), and organ impairment (Chouchana et al., 2021).

Remdesivir is an adenosine analogue. It is metabolised into remdesivir triphosphate (RTP), which is used as a substrate by the RdRp. This leads to incorporation of the remdesivir monophosphate (RMP) into the nascent RNA strand. In SARS-CoV-2, it has been reported that the RdRp then extends the RNA by three additional nucleotides before it encounters a translocation barrier and stalls (Gordon et al., 2020; Kokic et al., 2021; Q. Wang et al., 2020).

It has also been reported that remdesivir can be incorporated into the copy of the first RNA strand which is subsequently used as a template, leading to reduced efficiency of nucleotide incorporation in the complementary strand (E. Tchesnokov et al., 2019).

In vitro studies have not shown remdesivir induces lethal mutagenesis (E. Tchesnokov et al., 2019). Sequence analysis we did confirmed this finding (details can be found in Chapter 4, Section 4.3.7) (Boshier et al., 2020a).

Molnupiravir

Molnupiravir was originally developed for treating influenza (Kabinger et al., 2021). It has demonstrated antiviral activities in influenza (Toots et al., 2019), Venezuelan equine encephalitis virus (Painter et al., 2019), respiratory syncytial virus (Yoon et al., 2018), Chikungunya virus

(Ehteshami et al., 2017), Ebola virus (Reynard et al., 2015), norovirus (Costantini et al., 2012), and hepatitis C viruses (Stuyver et al., 2003). In more recent clinical trials, molnupiravir has shown to reduce the risk of hospitalisation and death in SARS-CoV-2 infected at-risk adults (Jayk Bernal et al., 2022; Singh et al., 2021) and has been approved for use to treat COVID-19 in the UK since November 2021 (UK Parliament, 2022).

Molnupiravir is a cytidine or uridine analogue. It has an active form of β-D-N^4-hydroxycytidine (NHC) triphosphate, which is used as a substrate by the RdRp, in place of the cytidine triphosphate or uridine triphosphate (Kabinger et al., 2021). NHC monophosphate is incorporated into the nascent RNA strand. However, the RdRp does not stall after the incorporation, meaning molnupiravir does not act by directly inhibiting the RdRp. Instead, a full RNA strand can be synthesised in the presence of NHC. When RdRp uses this RNA strand as a template, NHC directs the incorporation of either G or A, which leads to an increase in transition mutations C to T and G to A, and subsequently triggers the process of lethal mutagenesis (Gordon et al., 2021; Kabinger et al., 2021). Molnupiravir has been shown to cause lethal mutagenesis *in vivo* in both influenza (Toots et al., 2019) and coronaviruses (Sheahan et al., 2020).

4'-Fluorouridine

4'-Fluorouridine is an oral antiviral which has recently been shown to be effective against RSV in mice and SARS-CoV-2 in ferrets and human airway organoids (Sourimant et al., 2022). It has been reported that a daily dose of 4'-Fluorouridine can significantly reduce the viral shedding *in vivo* (Sourimant et al., 2022). Derivatives of 4'-Fluorouridine has also

been shown to possess antiviral properties in HCV *in vitro* (Ivanov et al., 2010).

4'-Fluorouridine is a uridine analogue. Instead of inducing an error catastrophe like favipiravir or molnupiravir, it causes delayed stalling of RNA polymerases in a similar mechanism of action used by remdesivir. In an *in vitro* study, it has been shown that the RdRp recognises and incorporates the bioactive 5'-triphosphate form of 4'-Fluorouridine in place of uridine triphosphate which leads to the stalling of RdRp at the site of incorporation or 3 nucleotides downstream to the site of incorporation, depending on the template sequence (Sourimant et al., 2022).

Non-Nucleoside Analogue Inhibitors

NNIs on the other hand have very diverse structures. They bind to allosteric sites in the thumb and palm subdomain which then change the spatial conformation of the RdRp and lead to an inhibition of the replication activities (Barreca et al., 2011; Shunmugam and Soliman, 2018). Since allosteric sites are more variable compared to the active site of RdRp, resistance to NNIs is developed very rapidly. In addition, a number of NNIs such as pimodivir, developed for influenza, and lomibuvir and tegobuvir, developed for hepatitis C were discontinued during clinical trials due to severe adverse events (M. C. Patel et al., 2021). Due to the structural variability of the allosteric sites between different RNA viruses (Mittal et al., 2019), NNIs have relatively limited targets and mostly cannot be repurposed (Tian et al., 2021). Therefore, NNIs are less frequently used for treating RNA viruses.

In this thesis, we will study two widespread positive-sense single-stranded RNA viruses, norovirus and SARS-CoV-2, as well as some broad-spectrum nucleoside analogues RdRp inhibitors which are used to treat these two viral infections.

1.2 Norovirus

1.2.1 History

Norovirus was first described as the winter vomiting bug by J. Zahorsky in 1929. It was characterised by the sudden onset of diarrhoea and vomiting (L. Adler and Zickl, 1969). However, the infectious agent causing winter vomiting disease was only identified by A.Z. Kapikian in 1972 using immune electron microscopy in stool samples collected from a gastroenteritis outbreak in an elementary school in Norwalk, Ohio, USA (Kapikian et al., 1972). Small, round viral particles of 27 to 30 nm in diameter were found in the samples. This infectious agent was named as Norwalk virus and it became a prototype strain for the related Norwalk-like viruses identified. They were subsequently grouped into the Norovirus genus in the family of Caliciviradae (Robilotti et al., 2015).

1.2.2 Clinical Features

Norovirus is responsible for around 20% of acute gastroenteritis (Hall, 2011, pp. 2004–2005; Scallan et al., 2011) and over 85% of viral gastroenteritis outbreaks (Lee and Pang, 2013). It has been estimated that norovirus causes around 700 million episodes of diarrhoea and 200,000 deaths annually, which costs an economical burden of approximately $60 billion worldwide per year (Kirk et al., 2015; Pires et al., 2015; Lopman et al., 2016).

Symptoms including abdominal cramps, nausea, vomiting and diarrhoea typically appear after an incubation period of 24 to 48 hours (Robilotti et al., 2015). In healthy individuals, human noroviruses infections are self-

limiting, and the acute gastroenteritis associated are normally cleared within days (de Graaf et al., 2016; Karst, 2010). However, in immunocompromised individuals, such as patients with immune disease (Rodríguez-Guillén et al., 2005; Schwartz et al., 2011) or transplant recipients who are on immunosuppressants (Capizzi et al., 2011; Roos-Weil et al., 2011; Schorn et al., 2010), as well as in children (Ludwig et al., 2008; Murata et al., 2007) and elderly (Aoki et al., 2010; Lai et al., 2013), norovirus can cause severe diarrhoea and prolonged viral shedding. Severe weight loss problems have been reported in over 30% of chronic norovirus cases (Petrignani et al., 2018). Chronic diarrhoea can lead to severe malabsorption, renal failure, life-threatening pneumatosis intestinalis and many other complications (Bok and Green, 2012; Brown et al., 2019, 2017; Roddie et al., 2009).

Norovirus is highly contagious (reproductive number $R0 > 2$) (Gaythorpe et al., 2018). It spreads quickly via food, surfaces and water and often leads to outbreaks in schools, hospitals, and care homes (Lopman et al., 2012, 2003; Thornley et al., 2011). It is transmitted via the faecal-oral route (de Graaf et al., 2016). Faeces and vomitus of infected individuals contain a high concentration of norovirus particles, and shedding can continue for weeks even after the individual has recovered (Teunis et al., 2008). Yet, the minimal infectious dose of norovirus is extremely low, where 18 viral particles can already lead an infection and cause acute gastroenteritis (Glass et al., 2009).

1.2.3 Taxonomy

There are five genera in the family of *Caliciviridae*, namely Norovirus, *Sapovirus*, *Lagovirus*, *Nebovirus* and *Vesivirus*. They fall into distinct phylogenetic clades within the family (Green et al., 2000) (Figure 6).

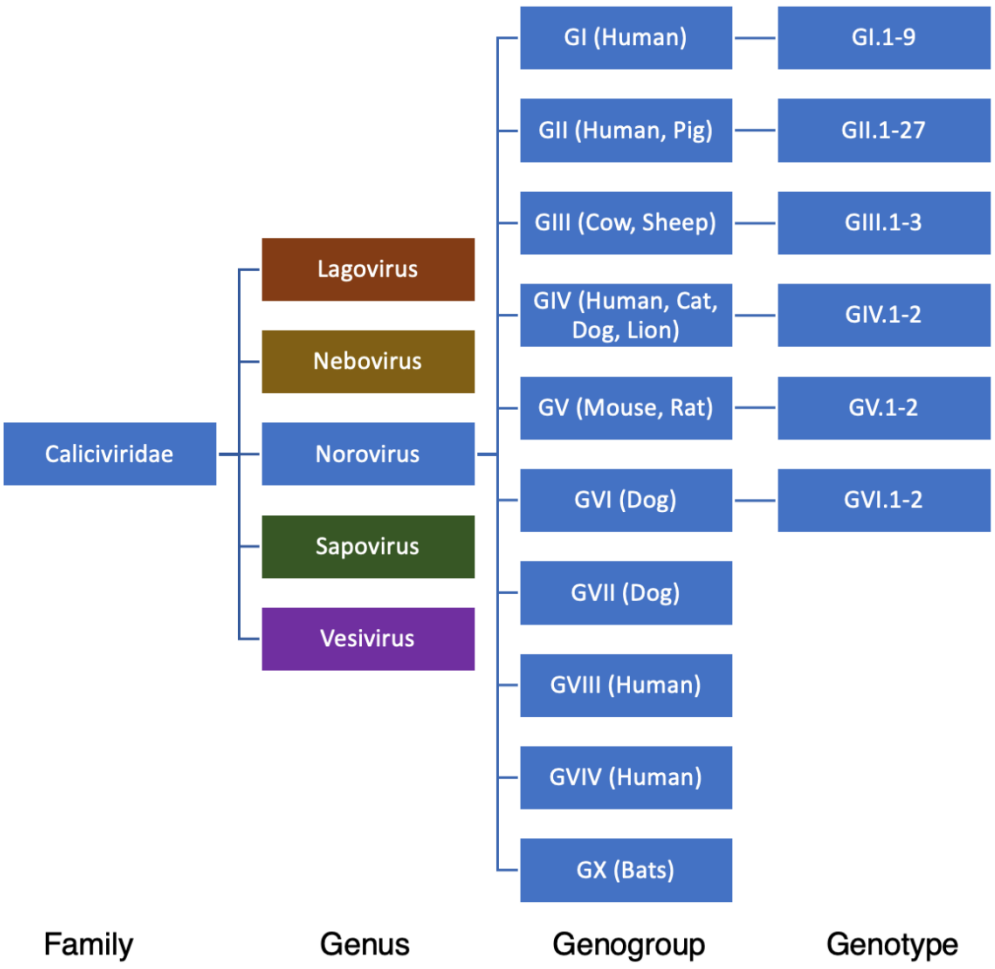


Figure 6. Classification of Caliciviridae. *Norovirus is one of the five genera in the family of Caliciviridae. It can be divided into 10 genogroups. Each genogroup can be further subdivided into different number of genotypes.*

In mid-1990s, noroviruses were organised into genogroups and genotypes based on partial RdRp sequences (Green et al., 1995;

Nakayama et al., 1996; Vinje and Koopmans, 1996). In the 2000s, the classification of noroviruses was based on the similarity of the VP1 amino acid sequences, where a cut off of 20% sequence difference was used as a threshold for a new genotype (Vinjé et al., 2000; Zheng et al., 2006). However, it was shown to be unreliable. Therefore, in 2013, the Norovirus Classification Working Group proposed a universal typing system using phylogenetic analyses of the complete VP1 amino acid sequences (Kroneman et al., 2013). Since recombination happens frequently at the boundary of the first and second open reading frames (Eden et al., 2013), a dual typing nomenclature system, where the RdRp (P-type) and major capsid (VP1) (genotype) are independently classified into genogroups and genotypes, is now being used for norovirus (Chhabra et al., 2019; Kroneman et al., 2013).

The genus of noroviruses is subdivided into ten genogroups (GI to GX) and 49 genotypes (9 in GI, 27 in GII, 3 in GIII, 2 in GIV, 2 in GV, 2 in GVI, and 1 genotype each for GVII, GVIII, GIX and GX) based on the VP1. Of the ten genogroups, only three (GI, GII and GIV) can infect and cause gastroenteritis in humans (Franco and Greenberg, 2012). Three genotypes of GII (GII.11, GII.18 and GII.19), however, are found to infect swine. The genotype GIV.2 has, on the other hand, been only observed in cats and dogs (Figure 6) (Chhabra et al., 2019).

Apart from human cases, noroviruses have also been identified in different animal species. For example, GII with the infection of pigs (Shen et al., 2012; Wang et al., 2005), GIII with the infection of cows and sheep (Wolf et al., 2009), GIV and GVI with the infections of felines (Di Martino et al., 2016), GIV, GVI and GVII with the infections of canines (Lizasoain et al., 2015), GIV and GVI with the infections of dogs (Mesquita et al., 2010), and GV with the infections of mice and rats

(Zhang et al., 2015). A number of unclassified noroviruses have also been found in harbour porpoise (de Graaf et al., 2017), bats (Wu et al., 2016) and sea lions (Teng et al., 2018).

The RdRp clusters are referred to P-groups at the genogroup level, and P-types at the genotype level. Phylogenetic analysis confirmed there are eight P-groups, namely GI.P, GII.P, GIII.P, GIV.P, GV.P, GVI.P, GVII.P and GX.P. RdRp sequences of GI viruses could be divided into 14 P-type; GII, into 37 P-types; GIII, GV GVI, into two P-types; GIV, GVII and GX, into one P-type (Chhabra et al., 2019).

1.2.4 Epidemiology

GII.4 viruses are responsible for 70 to 80% of norovirus outbreaks worldwide since the mid-1990s (Siebenga et al., 2009). Since 2000, there have been eight GII.4 variants circulating globally with each new variant surpassing the previous dominant variant (Parra, 2019). This includes GII.4 Grimsby strain in 1995, GII.4 Farmington Hills in 2002, GII.4 Hunter in 2004, GII.4 Den Haag in 2006, GII.4 New Orleans in 2009 and GII.4 Sydney in 2012 (de Graaf et al., 2016; Vinjé, 2015).

GII.4 strains are usually replaced by a newly emerged divergent GII.4 strain every two to three years (Lindesmith et al., 2011). Grimsby strain in 1995 and GII.4 Sydney in 2012 each dominated for over five years. These lull periods are very similar to a period in the early 1980s when a variant of influenza A virus (H3N2) dominated for eight years until 1987 (Smith et al., 2004). These periods have been described as a “strain lock” where variants are unable to overcome a fitness barrier. This often occurs when the majority of the population has been naturally infected or vaccinated (Pangburn et al., 2008).

GII.4 Sydney has predominated globally since 2012 with millions of reported cases which replaced the previously circulating GII.4 New Orleans (Kroneman et al., 2013). The dominance of GII.4 and the re-emergence of new variants have been studied extensively and it is believed to be caused by the mutations on VP1 (Parra et al., 2017).

The most prevalent genotypes are GII.4 Sydney (Cannon et al., 2021) and New Orleans (Eden et al., 2014). While GII.4 has been the dominant human infecting variant for over two decades across the globe, other genotypes have been spotted to predominate in certain geographical

locations such as GII.17 and GII.2 (Ao et al., 2017; Chan et al., 2015; Matsushima et al., 2015; Niendorf et al., 2017; Parra and Green, 2015).

In the 2014 to 2015 season, the first non-GII.4 pandemic of GII.17 Kawasaki was described as emerging initially in China in late 2014. The viruses were found in 13 countries across 4 continents. It has been widely speculated as to whether GII.17 will eventually replace GII.4 (Xue et al., 2019), although to date that has not happened.

In 2016, there was a decline in the GII.4 Sydney 2012 variant circulating at the time in Sydney and New Zealand and the emergence of two new GII.4 Sydney variants, GII.P4 New Orleans 2009/GII.4 Sydney 2012 and GII.P16/GII.4 Sydney 2012 (J. H. Lun et al., 2018). Compared to the GII.4 Sydney 2012 capsid, the consensus sequences of the GII.P4 New Orleans 2009/GII.4 Sydney 2012 and GII.P16/GII.4 Sydney 2012 have seven and ten amino acid changes in the non-structural region, respectively (J. Lun et al., 2018). The co-existence of two GII.4 variants is uncommon as a single GII.4 pandemic variant usually account for 60-80% of the infection cases at any given time during the past two decades and dominates until being replaced by an incumbent pandemic variant (Siebenga et al., 2009). In mid-late 2016, there was an increase of a third recombinant, GII.P16/GII.2, which accounted for 14 to 42% of all norovirus outbreaks around the globe (J. Lun et al., 2018; J. H. Lun et al., 2018).

From 2016 to 2020, it has been reported by NoroSurv that P16/GII.4 Sydney 2012 or GII.4 Sydney [P31] strains are the most common strains among children under five years old (Cannon et al., 2021). GII.4 Sydney 2012 virus has been circulating for longer than other global strains. From 2014 to 2020, GII.4 Sydney strains showed some changes or drifts

in histo-blood group antigen-binding sites (II) and epitope sites (A, B, E, G, and H). The alterations in residues could potentially be one of the reasons for Sydney 2012 to evade population immunity and dominate over a long period of time (Bull et al., 2010; Zhu et al., 2021).

During the 2021 to 2022 season, in the UK, 92% of samples were identified in the GII group, with 51% identified as GII.4 (UK Health and Security Agency, 2022).

Seasonality

Studies in both Northern and Southern hemisphere showed that norovirus outbreaks peak in winter (da Silva Poló et al., 2016; Eden et al., 2014). Yet, there does not seem to be a seasonal trend observed in more tropical area such as Africa (Mans et al., 2016).

1.2.5 Genome Structure and Organisation

The norovirus genome is approximately 7.5 to 7.7 kilobases (kb) long (Thorne and Goodfellow, 2014) with a viral protein linked at the 5'end and a polyadenylation tail at the 3' end (Chhabra et al., 2019). The genome consists of three open reading frames (Atmar and Estes, 2001; Bull et al., 2005). The first open reading frame, ORF1 encodes for non-structural proteins including p48, nucleoside triphosphatase (NTPase), p22, viral protein genome-linked (VPg), 3-chymotrpsin like protease (3CLpro), and the RNA-dependent RNA polymerase (RdRp), which are processed post-translationally by the 3CLpro (Hardy, 2005; Someya et al., 2000). The second and third open reading frames, ORF2 and ORF3 encode for the major capsid protein VP1 and minor capsid protein VP2 respectively (Hardy, 2005) (Figure 7).

The non-structural protein p48 has around 400 amino acids and a molecular weight of 37 to 48 kDa (Belliot et al., 2003; Donaldson et al., 2008; Liu et al., 1996). It is located at the N-terminus of the polyprotein encoded by the ORF1. It is involved in regulating the docking and fusion apparatus for soluble N-ethylmaleimide sensitive factor attachment protein receptor (SNARE)-mediated vesicle fusion (Campillay-Véliz et al., 2020). NTPase has 366 amino acids and a molecular weight of 40 kDA (Belliot et al., 2003, p. 145). It supports the enzymatic activities during viral replication, including the unwinding of RNA helices, and the annealing and remodelling of RNA structures (Li et al., 2018). P22 has around 180 amino acids and a molecular weight of 20 to 22 kDA (Belliot et al., 2003; Sharp et al., 2010). The function of p22 is not fully understood (Campillay-Véliz et al., 2020) but it is known to interfere with the host cell protein secretion and modification pathways (Sharp et al., 2010). VPg has 132 amino acids and a molecular weight of 15.8 kDa

(Belliot et al., 2003). It acts as a primer for the replication of viral RNA (Goodfellow, 2011). 3CLpro has 180 amino acids and a molecular weight of 19.4 kDA (Campillay-Véliz et al., 2020; Nakamura et al., 2005). It cleaves the polyprotein encoded by ORF1 into multiple mature non-structural proteins which are involved in genome replication and viral pathogenesis (Weerasekara et al., 2016). The RdRp, also referred to as the polymerase (Pol), has 510 amino acids and a molecular weight of 56.8 kDA. Its main function is to replicate the viral genome (Belliot et al., 2003; Campillay-Véliz et al., 2020).

Norovirus has two structural proteins, VP1 and VP2, which form the viral capsid (Hardy, 2005). VP1 has 530 to 555 amino acids and a molecular weight of 58 to 60 kDa (Hardy, 2005). As resolved and described by X-ray crystallography, 180 copies of VP1 self-assemble into an icosahedral virus-like particle which forms into the shell (S) domain and the protruding (P) domain (Hardy, 2005). The P domain is further divided into the P1 and P2 subdomains which interact with each other to provide stability to the capsid (Hardy, 2005). The P2 subdomain is also known to be responsible for receptor binding (Prasad et al., 1999). VP2, the minor capsid protein, has 208 to 269 amino acids and a molecular weight of 22 to 29 kDa (Vongpunsawad et al., 2013). The sequence of VP2 is highly diverse across different genotypes. It stabilises the viral capsid and enhances the expression of VP1 during viral replication. It has been hypothesised that VP2 is involved in RNA binding and genome packaging since the norovirus VP1 does not have a domain which is typically found in other capsidated viruses, however, no experimental data is available to support this hypothesis (Glass et al., 2000).

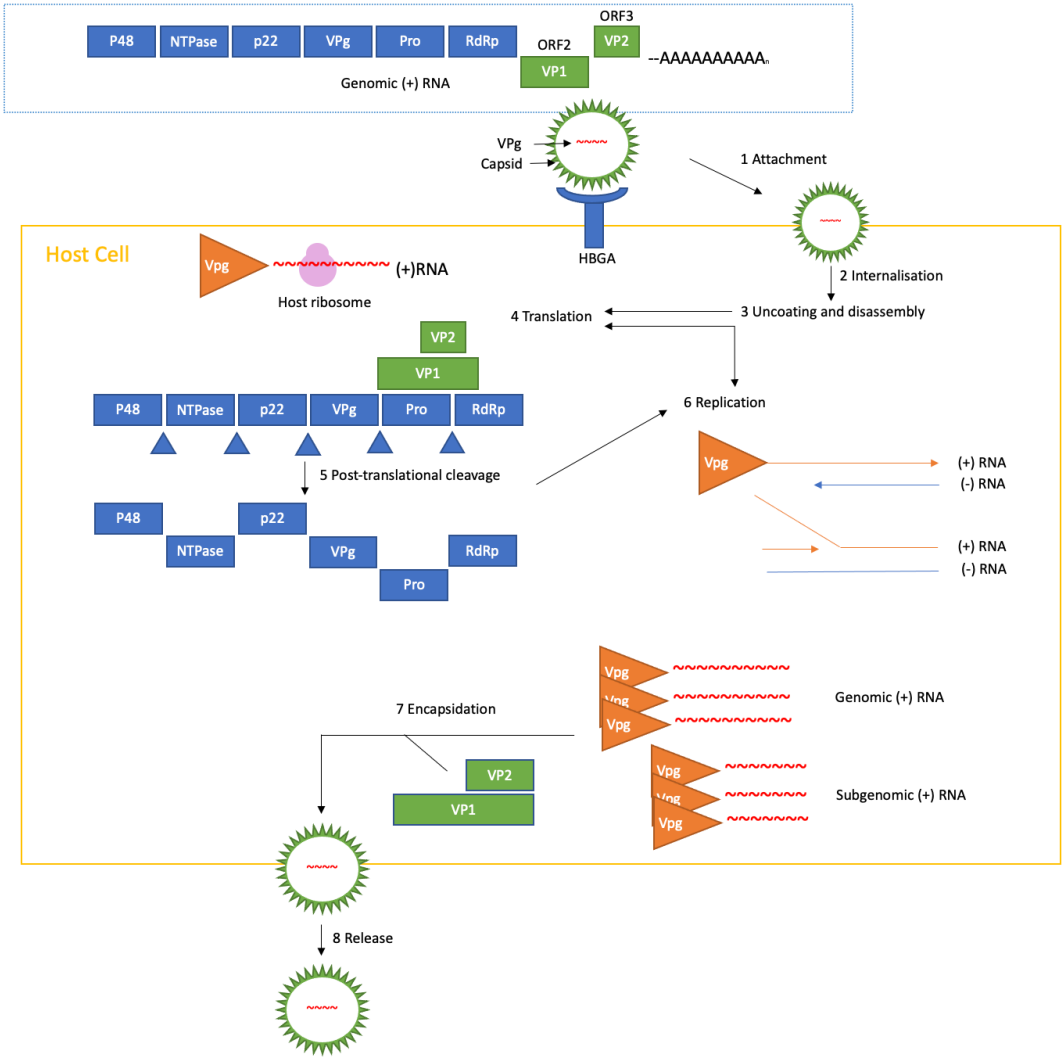


Figure 7. Genome map and schematic representation of the replication cycle of norovirus. *Details of the replication cycle will be described in the text. This figure presents a brief illustration of the replication cycle, which involves eight major steps: viral attachment, entry, uncoating and disassembly, 'pioneer' translation, post-translation cleaving, genome replication with additional rounds of translation, and viral assembly and release.*

1.2.6 Replication Cycle

Cell Entry

The first stage of the norovirus replication cycle is the entry into the target cell. In general viral entry involves binding and interaction of the virus with specific cellular receptors on the surface of a susceptible target cell. Noroviruses bind to the cell using the P2 region of the P domain in the major capsid protein (VP1). The main cellular receptor for human norovirus remains unknown (Lin et al., 2014; Marionneau et al., 2002). However, some cell-binding factors and coreceptors, including the histo-blood group antigens (HBGAs), which are carbohydrates that can be found on the surface of gut epithelial cells, have been described (Lindesmith et al., 2010; Reeck et al., 2010). Precursors of HBGAs are added onto the surface of certain cell types by the alpha(1,2)-fucosyltransferase (FUT2) (de Graaf et al., 2016). H HBGA and subsequently A and B HBGAs are then expressed on the surface of the cells. FUT1 also synthesizes A, B and H HBGAs corresponding to the blood groups A, B and O, for expression on the surface of red blood cell(de Graaf et al., 2016). The binding affinity and specificity of human norovirus capsid to HBGAs differ among different genotypes. Therefore, individuals with no FUT2 are less susceptible to infections of certain genotypes of human noroviruses, such as GII.4 (Le Pendu et al., 2006). It has also been reported that individuals with blood types AB and B are less susceptible to noroviruses of certain genotypes (Liao et al., 2020).

While the exact entry mechanism of norovirus is still poorly understood, a multistep process for cellular entry has been hypothesised (Karst, 2010). Norovirus first interacts with and binds to the HBGAs which are expressed on the surface of intestinal epithelial cell. This triggers the

uptake of the virus into the cell (Donaldson et al., 2010; Hassan and Baldridge, 2019). Following the internalisation of the virus, the viral capsid will be disassembled and the viral RNA will be released into the cytoplasm of the target cell (Daughenbaugh et al., 2006). Once the viral RNA is released, the VPg covalently linked at the 5' end acts as the cap of the cellular mRNA (Royall and Locker, 2016) and interacts with the cellular eukaryotic translation initial factors (eIFs) 3, 4e and 4G and the cap-binding protein, forming a translation complex, which subsequently recruits the ribosomal complex (Chung et al., 2014). ORF1 which encodes for the viral non-structural proteins is then translated into a polyprotein, which is cleaved by the protease co- and post-translationally, generating three protein precursors ProPol, p22VPg and p48NTPase (Todd and Tripp, 2019). The function or activity of the p22VPg and p48NTPase precursors are currently unknown (Campillay-Véliz et al., 2020). However, the ProPol precursor is known to be involved in further cleaving of the three precursors into six individual non-structural proteins, p48, nucleotide triphosphatase (NTPase), p22, VPg, 3C-like protease (3CL-pro), and the RdRp (Campillay-Véliz et al., 2020).

Genome Replication

The viral genome replication is mainly carried out by the RdRp, NTPase and VPg on cellular membranes which are recruited by the p22 and p48 proteins (Deval et al., 2017). Although the exact mechanisms of membrane recruitment are poorly understood, p22 and p48 are known to be involved in localising to the endoplasmic reticulum, Golgi apparatus and endosomes, the key components of the secretory pathway (Sharp et al., 2010, 2012). Two genome replication mechanisms for the initiation of RNA strand synthesis have been demonstrated *in vitro*. The first mechanism involves a *de novo* initiation

of the negative-sense genomic and subgenomic RNA synthesis, where specific loops in the shell domain of the major capsid protein enhance genome replication (Kao et al., 2001; Rohayem et al., 2006). The second mechanism of replication involves nucleotidylation of the VPg, where the RdRp attaches an initiating nucleotide to the tyrosine residue in the VPg, generating the positive-sense genomic RNA (Belliot et al., 2008). Finally, both the genomic and subgenomic RNA are mobilised by the NTPase RNA chaperone activities. After multiple rounds of viral genome replication and protein translation, the major and minor capsid proteins are generated from the subgenomic RNA containing ORF2 and ORF3 (Hardy, 2005).

During the viral replication process, the p22 and p48 proteins interfere with the host immune response signalling pathways NFκB, MAPK, and PI3K-Akt (Ettayebi and Hardy, 2003; Lateef et al., 2017). Both p22 and p48 induce the disassembly of the Golgi apparatus (Sharp et al., 2010). While p48 binds to proteins which are involved in the SNARE regulated vesicular transport system and blocks the host cell protein transport, p22, which has a motif that mimics the export signal of the endoplasmic reticulum, blocks the transport of COPII-coated vesicles (Sharp et al., 2010).

Viral Assembly

The mechanisms of viral encapsidation and the viral exit from host cell are still poorly understood. It has been hypothesised that once all viral proteins are generated, the major capsid protein self assembles into a viral-like particle, and the minor capsid which is located on the inside of the viral particle may recruit the viral genome into the particle (Vongpunsawad et al., 2013). The virion is then fully assembled and

released. Apoptosis of the host cell facilitates the release of norovirus virions from the host cell and inhibition of apoptosis reduces the production of murine norovirus virions (Karst, 2010).

1.2.7 Cell Culture and Cell Tropism

Historically, the lack of robust cell culture system or animal models for human noroviruses has been the greatest challenge for studying the human norovirus life cycle and host-pathogen interactions (Estes et al., 2019). Initial efforts to develop non-human primates as large animal models such as monkeys and chimpanzees have been proven unsuccessful since these animals do not develop gastroenteritis or any other symptoms (Bok et al., 2011; Wyatt et al., 1978). In recent years, Gnotobiotic piglets and calves have been used since they develop diarrhoea upon oral infection, have detectable levels of human norovirus in the intestine and shed virus up to 6 days (Bui et al., 2013, p. 4; Cheetham et al., 2006, 2006; Kocher et al., 2014; Lei et al., 2016; Souza et al., 2008; Takanashi et al., 2011; Todd and Tripp, 2019). However, they are extremely costly and require high capacity in the experimental facility. More recently, a simple and robust zebrafish larvae model has been shown to allow efficient replication of norovirus (Van Dycke et al., 2019). The further application of the zebrafish larvae model will be described in Chapter 4.

While murine noroviruses have been shown to infect and replicate in primary murine macrophages, dendritic cells (Perry et al., 2009; Wobus et al., 2004), T cells and B cells *in vitro* (Hsu et al., 2018), attempts to culture human noroviruses from macrophages and dendritic cells have not been successful (Lay et al., 2010).

For human noroviruses, the BJAB B cell line has been used as a cell culture. Replication of human norovirus within the cell line was stable and reproducible (Jones et al., 2015, 2014). It has been found that the co-culture of human norovirus with enteric bacteria which expresses

histo-blood group antigens (HBGAs) can lead to an increase in human norovirus replication (Jones et al., 2014). However, it is known that even in immunocompromised patients with no B cells, norovirus can still replicate and lead to a high viral load, meaning human norovirus can replicate in other tissue types apart from B cells (Brown et al., 2016). In recent years, the major capsid proteins, VPg and RdRp have been detected in the duodenum and jejunum enterocytes (Karandikar et al., 2016). The major capsid protein was also detected in macrophages, T cells and dendritic cells in the lamina propria, the connective tissues in the intestinal mucosa, in immunocompromised patients (Karandikar et al., 2016). A cell culture system developed from human intestinal glands stem cells was employed to demonstrate the replication of human noroviruses in enterocytes (Ettayebi et al., 2016). Multiple genotypes have successfully demonstrated viral replication in this cell culture system. It has therefore been suggested that enterocytes are the major target cell for human norovirus *in vivo* (Ettayebi et al., 2016; Green et al., 2020).

1.2.8 Immune Response

Innate Immunity

Innate immunity responses, in particular type I Interferon responses, have been shown to be critical for suppressing norovirus infection in both mice and human (Karst et al., 2003; Mboko et al., 2022). Mice with STAT1 or other interferon receptors deficiency developed a lethal infection when challenged with murine norovirus (Karst, 2010; Thackray et al., 2012). It has been demonstrated that robust innate immune response driven by type I and type III interferons can be induced by human noroviruses (Hosmillo et al., 2020; Lin et al., 2020; Sarvestani et al., 2016).

In recent years, human norovirus replication has been successfully demonstrated in the human intestinal enteroid/ human intestinal organoid (HIE/HIO) system. Such a system can be used to study the cellular processes and signalling pathways of human norovirus. Using this system, a recent study has shown that the GII.4 and GII.3 norovirus strains are sensitive to type I (IFNα1, IFNβ1) and type III interferon (IFNλ1, IFNλ2, and IFNλ3). In interferon-receptor-knockout cell lines, the noroviruses showed higher level of replication compared to wild-type cells. When Ruxolitinib, a Janus kinase 1 (JAK1) JAK 2 inhibitor was used to disrupt the interferon signalling pathway, an increase in norovirus replication was observed (Mboko et al., 2022).

Adaptive Immunity

Components of the adaptive immune system, including B cells, CD4+ and CD8+ T cells, have been shown to promote viral clearance from the intestinal tract (Chachu et al., 2008). A study has shown that RAG1 knockout mice with no B cells have failed to clear murine norovirus infections and B cells transfer into these chronic infected mice resulted in immediate viral clearance (Chachu et al., 2008).

Although the role of T cells in norovirus infection has not been well described, it has been suggested that a cellular immunity is required to achieve viral control (Brown et al., 2019, 2017; Lindesmith et al., 2020; Newman et al., 2016; Siebenga et al., 2008). T cell responses are known to be effective against closely related strains of norovirus (Lindesmith et al., 2010, 2008). In one study, individuals infected with GII.2 virus showed a cross-reactive immune response against GI.1 and GII.1 viral-like particles in assays (Lindesmith et al., 2005). GI.1 infected individuals also exhibited an immune response to GI.2, GI.3 and GI.4 variants in a similar experiment (Lindesmith et al., 2010). However, immunity against norovirus tends to be short-lived, meaning individuals can be reinfected by the same strain of norovirus genotypes (Parrino et al., 1977; Simmons et al., 2013).

1.2.9 Treatments and Vaccines

Treatments

Although there are no approved antiviral drugs for treating norovirus, repurposed RdRp inhibitors discussed in Section 1.1.5, such as favipiravir and ribavirin have been used (Netzler et al., 2019). Further details on the application of these antiviral treatments against norovirus will be discussed in Chapter 4.

Vaccines

Norovirus has been labelled by the World Health Organisation as a high priority for vaccine development since 2016 (Giersing et al., 2019). However, there are a few major challenges. First, human noroviruses evolve rapidly and are extremely diverse. Multiple genotypes are often circulating simultaneously (Chhabra et al., 2019; Parra et al., 2017). Secondly, noroviruses do not grow and replicate effectively in cell cultures, which makes live-attenuated vaccine unfeasible (Ettayebi et al., 2016). Finally, the lack of animal models has become the main barrier to our understanding of norovirus and to evaluate the vaccine candidates (Ha et al., 2016; Todd and Tripp, 2019).

Despite all the challenges, four human norovirus vaccine candidates have reached the clinical development stage (Tan, 2021). The most studied candidate is the TAK-214, an adjuvanted viral-like particle-based bivalent vaccine, which was developed by Takeda Pharmaceuticals International (Baehner et al., 2016). It contains two viral-like particles, one from the consensus sequence of three GII.4 genotypes which cause the most disease burden worldwide (Lopman et

al., 2016); the other from the GI.1 Norwalk virus which intends to broaden the protective immunity of the vaccine. The clinical trial conducted in a U.S. Navy training facility with over 4600 participants showed that the vaccine has an effectiveness of over 60% against moderate to severe norovirus induced acute gastroenteritis (Sherwood et al., 2020, p. 4). Another norovirus vaccine candidate VXA-NVV-104, developed by Vaxart Pharmaceutical Inc, contains recombinant adenovirus-based vectors that carry genetic materials of the GI.1 Norwalk virus and GII.4 Sydney major capsid protein, which can be expressed in the epithelial cells in the intestines and induce host immunity (Kim et al., 2018; Scallan et al., 2013). A phase 1 clinical trial shows that the vaccine recipients have a significantly higher antibody titre (Kim et al., 2018). The vaccine is currently undergoing another clinical trial in adults aged between 55 and 80 years old (Vaxart, 2022). The other two vaccine candidates NVSI (National Vaccine and Serum Institute, China, 2021) and Longkoma (Anhui Zhifei Longcom Biologic Pharmacy Co., Ltd., 2020), developed by the National Vaccine and Serum Institute of China and Institut Pasteur of Shanghai respectively, are currently undergoing clinical trial, but no data have been peer-reviewed or formally published to date.

1.3 SARS-CoV-2

1.3.1 History

The novel coronavirus 2019-nCoV was first identified in Wuhan, Hubei Province, China, in December 2019 (N. Chen et al., 2020). It was renamed severe acute respiratory syndrome coronavirus 2, SARS-CoV-2, by the International Committee on Taxonomy of Viruses in February 2020 (Michael Rajnik et al., 2021). The disease it causes was officially named coronavirus disease 2019, COVID-19. The initial cases were epidemiologically linked to the Wuhan Huanan Seafood Wholesale Market where wild animals such as rodents, pangolin and bats were sold (N. Chen et al., 2020). Bats are known to be the largest natural reservoirs for coronaviruses (Banerjee et al., 2019), and viral sequencing data support the hypothesis that the bats are responsible for the zoonotic spillover of SARS-CoV-2 (Andersen et al., 2020; Wacharapluesadee et al., 2021; P. Zhou et al., 2020). Since then, SARS-CoV-2 spread globally on all continents. In March 2020, the World Health Organisation declared COVID-19 a pandemic. As of May 2022, SARS-CoV-2 has caused over 500 million episodes of infections and over 6 million deaths worldwide (Worldometers.info, 2022).

1.3.2 Clinical Features

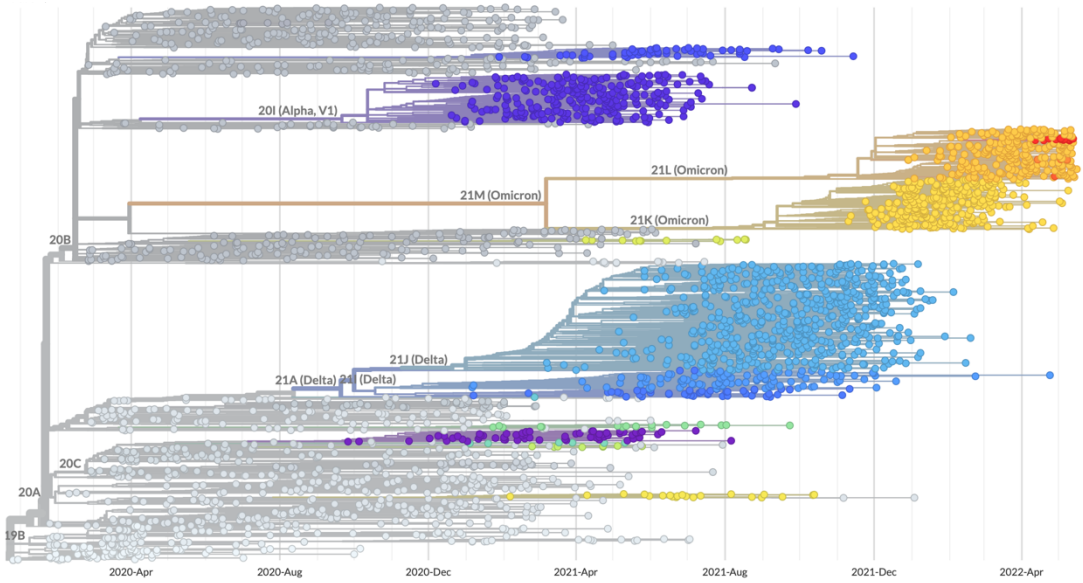
COVID-19 has an incubation period of 2 to 14 days (McAloon et al., 2020). Symptoms include fever, fatigue, loss of sense of taste or smell, body ache, headache, palpitation, diarrhoea, as well as respiratory symptoms such as continuous cough and shortness of breath (Çalıca Utku et al., 2020; Menni et al., 2020; Nehme et al., 2021). Currently, it is estimated that 40 to 45% of infected individuals remain asymptomatic (Oran and Topol, 2020). In mild symptomatic cases, healthy individuals are normally recovered within one to two weeks (J. Chen et al., 2020; Voinsky et al., 2020). However, elderly, pregnant women, individuals with certain heart or respiratory diseases and immunocompromised individuals are at higher risk of developing severe symptoms from COVID-19 (Phoswa and Khaliq, 2020; E. J. Williamson et al., 2020). In severe cases, a significant increase in cytokines such as IL-2, IL-7, IL-10, GSCF, IP10, MCP-1, MIP1A and TNF-α (described as a cytokine storm), has been reported as an important factor in disease progression (C. Huang et al., 2020; Song et al., 2020). Patients can develop pneumonia and organ failure which could lead to death (George et al., 2020).

More recently, the term “Long COVID” or “post-COVID syndrome” has been used to describe the presence of long term symptoms such as fatigue, problems with memory and concertation (also known as “brain fog”), insomnia, breathlessness and palpitation, months after the initial episode of SARS-CoV-2 infection (Raveendran et al., 2021). Currently, 1.8 million people in the UK (2.8% of the population) are experiencing self-reported long COVID (Office for National Statistics, 2022). However, it is unclear what causes long COVID.

1.3.3 Epidemiology

As the pandemic progressed, SARS-CoV-2 evolved into different lineages which cause varying degrees of disease severity. These lineages are identified and named based on a phylogenetic framework (Rambaut et al., 2020). In December 2020, the first variant of concern was identified in the UK as the Alpha variant or the B.1.1.7 lineage (Public Health England, 2020). It has a higher mortality rate compared to other strains of SARS-CoV-2 circulating previously (CMMID COVID-19 Working Group et al., 2021). Around the same time, the Beta variant (B.1.351 lineage), Gamma variant (P.1 lineage), and Delta variant (B.1.617.2 lineage) were identified in South Africa, Brazil and India respectively (Duong, 2021). These three variants are associated with an increased rate of transmission (Lin et al., 2021). In particular, the Delta variant, which is also known to be associated with significantly more severe diseases, has become the dominant strain in many countries in mid-2021. In November 2021, another variant, Omicron (B.1.1.529) was identified in South Africa (Thakur and Ratho, 2022). It was immediately recognised as a variant of concern as the number of positive cases soared in South Africa and viral genome deep sequencing revealed that the omicron strain carries more than 30 amino acid changes in the spike protein (Kumar et al., 2022). Omicron is highly transmissible, but it is associated with mild disease (Abdullah et al., 2022). Currently omicron is the dominant strain circulating worldwide (Figure 8).

(A)



(B)

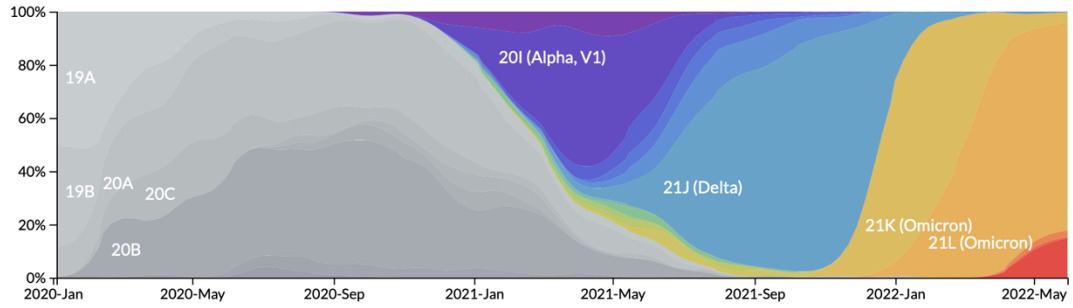


Figure 8. Phylogenetic tree and the relative global frequency of major SARS-CoV-2 clades from January 2020 to the present (May 2022). *The relative global frequency chart in panel B is coloured by the clade in the phylogenetic tree shown in panel A. Currently, the 21L Omicron viruses are dominated in the global population. Figure adapted from nextstrain.org with data provided by GISAID.*

1.3.4 Taxonomy

SARS-CoV-2 belongs to the *Betacoronavirus* genus in the sub-family of *Coronavirinae*, in the family of *Coronaviridae*, within the order of *Nidovirales* (Mousavizadeh and Ghasemi, 2021). *Coronaviridae* has two subfamilies, *Coronavirinae* and *Torovirinae* (Figure 9). Toroviruses infect horse, swine, cattle and fish, but not human (Schütze et al., 2006; Snijder and Horzinek, 1993). The viruses commonly known as coronaviruses are members of the *Coronavirinae*, The *Coronavirinae* subfamily is sub-divided into three genera, alpha, beta and gamma coronaviruses. Most alpha and beta coronaviruses infect mammalian hosts, while gamma and deltacoronaviruses mainly infect avian hosts (Woo et al., 2012).

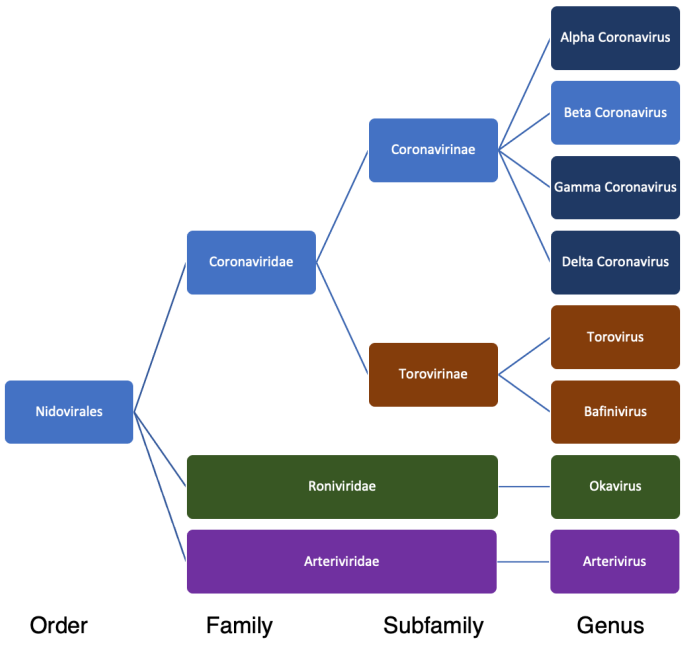


Figure 9. Classification of Nidovirales. *The order of Nidovirales consists of three families, Coronaviridae, Roniviridae and Arteriviridae. The Coronaviridae can be subdivided into the Coronavirinae and Torovirinae. The Coronavirinae can be further subdivided into four genera. SARS-CoV-2 belongs to the Betacoronavirus genus.*

Members of the *Coronavirinae* have been identified in the 1930s as the causative agents for gastroenteritis in pigs, bronchitis in chickens and hepatitis in mice (Cheever et al., 1949). In the 1960s, through electron microscopy, these viruses were found to have shared characteristics of club-shaped spikes projected on their virion surface. Viruses with such unique appearance were grouped into the *Coronavirinae* ("Virology," 1968).

Despite most human coronaviruses (e.g. alphacoronaviruses HCoV-229E and HCoV-NL63; betacoronaviruses HCoV-OC43 and HCoV-HKU1) only causing mild common cold symptoms, three betacoronaviruses have caused deadly epidemics and pandemics, including severe acute respiratory syndrome coronavirus (SARS-CoV), identified in 2003, Middle East respiratory syndrome-related coronavirus (MERS-CoV), identified in 2012, and SARS-CoV-2, which is causing the current COVID-19 pandemic (Artika et al., 2020).

1.3.5 Genome Structure and Organisation

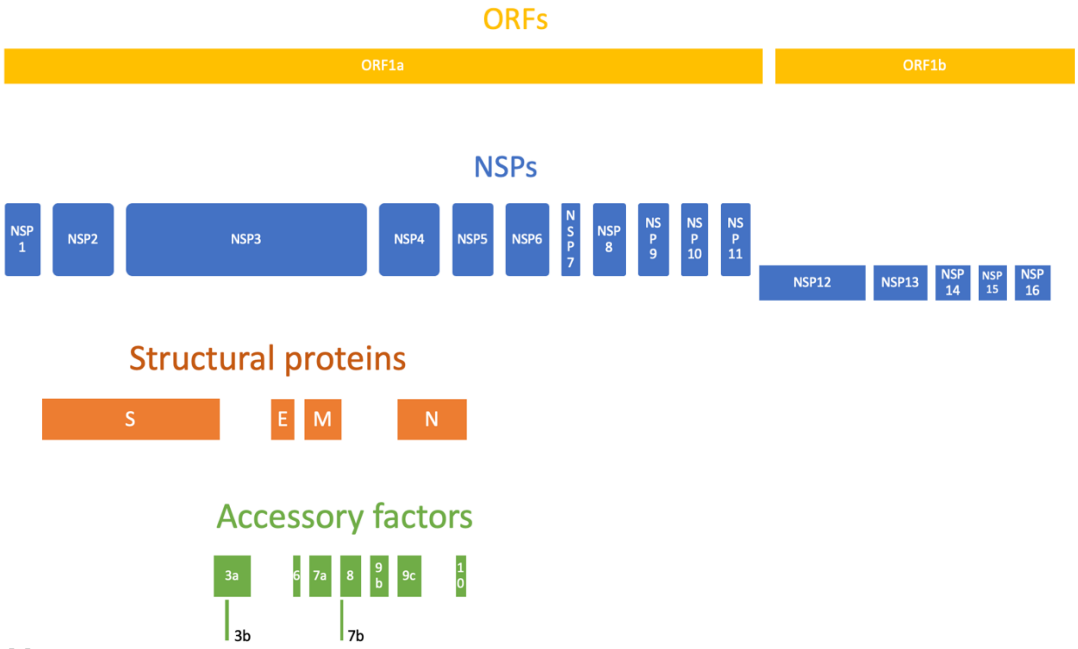
SARS-CoV-2 has a positive-sense RNA genome of 30 kilobases which is associated with a nucleoprotein within the viral capsid (Sahin, 2020). The genome consists of 14 open reading frames (six major open reading frames) which encode 27 different proteins (Wu et al., 2020) (Figure 10). The first two major ORFs, ORF1a and ORF1b encode for the frame-shift non-structural polyprotein which is cleaved into various non-structural proteins including papain-like protease (nsp3), chymotrypsin-like protease (nsp5), the RdRp (nsp12), and helicase (nsp13) (V'kovski et al., 2021, p. 2). These proteins form into a replicase complex that is involved in genome transcription and viral replication. The other major ORFs encode for spike (S), envelope (E), membrane (M) and nucleocapsid (N) proteins (V'kovski et al., 2021, p. 2).

The spike glycoprotein (S) consists of 1273 amino acid residues (Y. Huang et al., 2020) and has a molecular weight of 486 kDa (Herrera et al., 2021). It forms three subunits (S1, S2 and S2') that support the viral attachment to the host cell (Naqvi et al., 2020). The S1 subunit interacts with the human ACE2 receptors and attaches the virion onto the host cell membrane (Hoffmann et al., 2020, p. 2). Upon the viral attachment, S2 promotes the fusion of the virion with the host cell membrane (Duan et al., 2020). The spike protein goes through conformational changes where S2 changes from the pre-fusion native state to the post-fusion hairpin state (Naqvi et al., 2020). During the entry process, the subunit S2' functions as a fusion peptide which supports the rapid formation of intermediates during the fusion pathway (Qing and Gallagher, 2020).

The envelope protein (E), which is essential for viral assembly, consists of 75 amino acid residues (Naqvi et al., 2020) and has a molecular

weight of 12 kDa (Schoeman and Fielding, 2019). The membrane protein (M), which plays a significant role in supporting the shape of the virion and packing the RNA genome, consists of 222 amino acid residues (Naqvi et al., 2020) and has a molecular weight of 25 kDa (Artika et al., 2020). The nucleoprotein (N), interacts with the viral genome and M protein to support viral assembly, consists of 419 amino acid residues and has a molecular weight of 114 kDa (Zeng et al., 2020).

Apart from these major proteins, the SARS-CoV-2 genome encodes for a set of accessory proteins which are mostly known to involve in viral pathogenesis and impairment of the host immune response. In addition, there are two flanking regions of 265 nucleotides at the 5' end and 358 nucleotides at the 3' end (Chan et al., 2020).



...

Figure 10. Genome Map of SARS-CoV-2. *The genome of SARS-CoV-2 has 14 open reading frames (ORFs). The first two ORFs (ORF1a and ORF1b, shown in yellow) encode for non-structural proteins shown in blue. Genes encode for major structural proteins and accessory proteins are shown in orange and green respectively.*

1.3.6 Replication Cycle

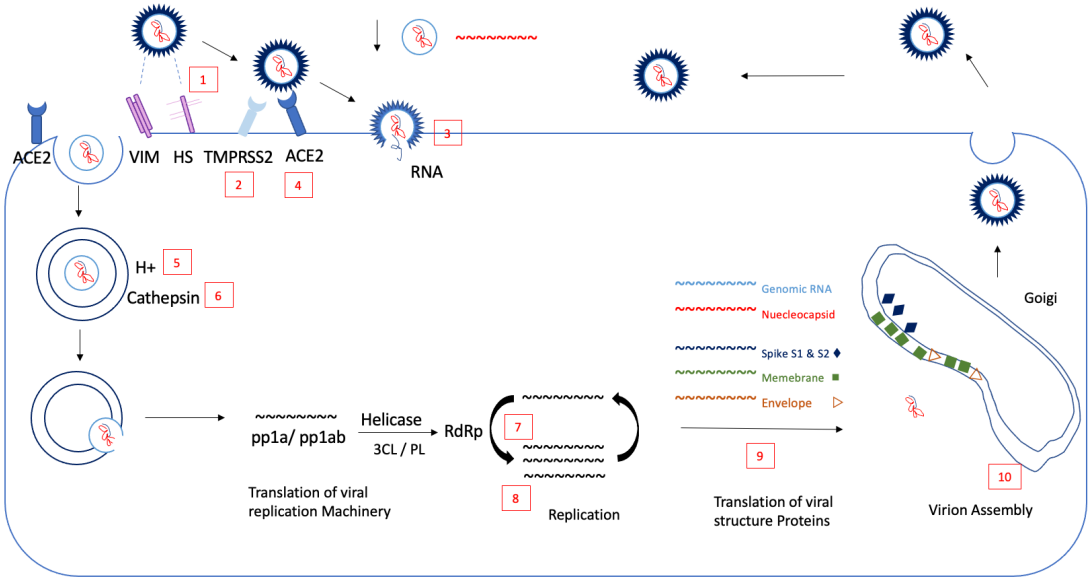


Figure 11. Schematic representation of the replication cycle of SARS-CoV-2. *The detail of the replication cycle will be described in the text. This figure presents a brief illustration of the replication cycle, which shows the key steps including viral attachment, internalisation, uncoating, translation of non-structural proteins, genome replication, translation of structural proteins, virion assembly and release.*

Cell Entry

To enter the host cell, the spike protein on SARS-CoV-2 bind to the angiotensin-converting enzyme 2 (ACE2) receptors expressed on human epithelial cells, endothelial cells and enterocytes (Hoffmann et al., 2020; Lan et al., 2020). Cellular proteases such as the transmembrane serine protease 2 (TMPRSS2) and furin facilitate the membrane fusion and endocytic viral entry into the host cell (Bergmann

and Silverman, 2020; Hoffmann et al., 2020). Upon viral entry, the capsid is degraded, and the viral genome is released.

Genome Replication

The ORF1ab is then translated into polyprotein 1a (pp1a) and pp1ab (Naqvi et al., 2020). The proteases non-structural protein (nsp) 3 and nsp5 inside pp1a and pp1ab cleave the polyproteins into 16 non-structural proteins which are then assembled into the replicase-transcriptase complex in the endoplasmic reticulum (V'kovski et al., 2021, p. 2). The replicase-transcriptase complex, composed of nsp2 to nsp16, is primarily driven by nsp12 (RdRp), nsp13, nsp14 and nsp16. The complex replicates the positive-sense RNA genome into negative-sense RNA copies which acts as the template for further replication into full-length positive-sense RNA genomes (V'kovski et al., 2021). The RdRp synthesises the RNA along with two cofactors nsp7 and nsp8 which have demonstrated adenylytransferase activity (Gao et al., 2020; Perlman and Netland, 2009; Snijder et al., 2016; Tvarogová et al., 2019). During the replication, nsp14 encoded exoribonuclease (ExoN) proofreads the nascent RNA strand and excises the misincorporated nucleotides (Pandey et al., 2020). Although the mechanism has not been fully understood, nsp13 along with the cofactor nsp10 is known to involve in the capping of the RNA (Chen et al., 2011, 2009; Ivanov and Ziebuhr, 2004).

These newly generated genomic RNAs can be translated into additional non-structural proteins or packaged into new virions. Similar to norovirus, a set of subgenomic mRNAs which encode structural proteins and accessory proteins will also be generated within the viral-induced membranous replication organelles (Long, 2021). During the negative-

strand RNA synthesis, the replicase-transcriptase complex interrupts after encountering the transcription regulatory sequence which is located at one-third of the viral genome from the 3’ end. The RNA synthesis then continues at the transcription regulatory sequence at the 5’ end which is adjacent to the leader sequence. This synthesis results in the production of a negative-strand subgenomic RNA, which is used as a template to generate a set of positive-sense subgenomic mRNA (V’kovski et al., 2021).

Viral Assembly

During virion assembly, the genomic RNAs are coated with nucleocapsid proteins in the endoplasmic reticulum which buds off to form the endoplasmic reticulum Golgi intermediate complex (ERGIC). The structure consists of a phospholipid bilayer with the viral spike. The membrane and envelope proteins inside are then formed into a Golgi vesicle which is subsequently released by exocytosis (Bergmann and Silverman, 2020).

1.3.7 Immune Response

SARS-CoV-2 infects cells which express ACE2 receptors. It primarily infects the respiratory system, but it has been reported that SARS-CoV-2 can also infect cells with ACE2 receptors in other organs such as the kidney, small intestines, pancreas (Liu et al., 2021), sweat glands and blood vessels (J. Liu et al., 2020).

Innate Immunity

When the virus enters the target cell, the pathogen recognition receptors, such as Toll-like receptor 2, recognise the surface epitopes on the virus, which triggers the production of type I and type III interferons (Shah et al., 2020). Impaired type I and type III interferon responses have been reported to associate with the risk of developing severe COVID 19 (Galani et al., 2021; Hadjadj et al., 2020).

The infiltration of monocytes, macrophages and neutrophils lead to the production of pro-inflammatory cytokines. A dysregulated release of pro-inflammatory cytokine can however contribute to a cytokine storm which leads to inflammatory cell death (Diamond and Kanneganti, 2022).

Apolipoprotein B mRNA editing catalytic polypeptide-like proteins (APOBEC) also play an important role in antiviral host defence. In particular, APOBEC3A, APOBEC1, and APOBEC3G have been shown to have innate immune functions against SARS-CoV-2 (Kim et al., 2022; Takaori-Kondo, 2006). They catalyse the deamination of C to U on the viral genome which potentially affects the viral fitness and replication (Kim et al., 2022; Sadler et al., 2010). Another two antiviral host defence mechanisms Reactive oxygen species (ROS) and adenosine

deaminase acting on RNA proteins (ADAR) can also leave characteristic mutational signatures on SARS-CoV-2 viral genomes. ROS induces G to U and C to A changes while ADAR introduces A to G changes (Kim et al., 2022; Mourier et al., 2021).

Adaptive Immunity

As the first line of defence, innate immunity restricts the viral replication within the infected individual and triggers an adaptive immune response.

In SARS-CoV-2, CD4+ T cells respond more prominently than CD8+ T cells. CD4+ T cells differentiate into helper and effector cells which instruct B cells, support CD8+ T cells and recruit innate immunity cells (Sette and Crotty, 2021).

In SARS-CoV-2, CD4+ T cells mostly differentiate into Th1 cells which trigger the production of interferon and cytokines, or T follicular helper cells which support B cells to develop neutralising antibodies and long-term humoral immunity. CD8+ T cells can kill the infected cells and are critical for viral clearance (Sette and Crotty, 2021). Although the presence of SARS-CoV-2 specific CD8+ T cells has been associated with better disease outcomes (Peng et al., 2020), they are less consistently observed than CD4+ T cells (Grifoni et al., 2020; Rydyznski Moderbacher et al., 2020; Sekine et al., 2020).

The humoral response to SARS-CoV-2 was shown to be comparable to that of previous coronavirus infections, with the production of the IgG and IgM antibodies. IgA, IgG, and IgM antibodies were detected after the onset of symptoms at different time points in infected patients. IgM and IgA antibodies were detected 5 days after the onset of initial

symptoms, whereas IgG was detected after 14 days (Guo et al., 2020). A persistent level of IgG was detected for a longer period, whereas IgM levels started to decline after 3 months (Li et al., 2008, 2003; Shah et al., 2020).

1.3.8 Vaccines

Vaccines

There are many approved vaccines for SARS-CoV-2. Currently, the following six vaccines have been approved for use in the UK: Pfizer/BioNTech, Moderna, Oxford/AstraZeneca, Janssen, Novavax and Valneva vaccines. Of which, the Pifzer/BioNtech, Moderna, and Oxford/AstraZeneca vaccines are most widely used globally (NHS, 2022).

mRNA vaccines

Pfizer and Moderna are both nucleoside-modified messenger RNA (mRNA) based vaccines. These vaccines contain mRNA which encodes for the spike protein of SARS-CoV-2, which will be translated by the host cells. Some of these spike proteins will be degraded into antigenic peptide epitopes by proteasome, which will then be transported back to the endoplasmic reticulum and subsequently presented in the major histocompatibility complex (MHC) class I on the cell surface (Rijkers et al., 2021). Alternatively, some spike proteins will be secreted and subsequently taken up by antigen-presenting cells and processed and presented in MHC class II (Wadhwa et al., 2020). An immune response will be triggered when the T cells recognise the MHC molecules and the foreign peptides bound to the molecules. B cells will be activated and specific antibodies against the spike proteins of SARS-CoV-2 will be developed (Shah et al., 2020).

Viral Vector Vaccine

The Oxford/AstraZeneca vaccine, on the other hand, is an adenoviral vector vaccine. Adenoviruses were used as a vector since they tend not to cause serious illness and they are easy to grow and replicate in tissue cultures (Rijkers et al., 2021). They were genetically modified to prevent replication in humans and a recombinant viral genome containing the gene encoding SARS-CoV-2 antigen was introduced. Upon administration, the muscle cells will be infected, the gene encoding antigens will be processed. The antigens will be presented, and a similar immune response will be triggered (Rijkers et al., 2021).

Other Vaccine Technologies

As of May 2022, there are 160 vaccine candidates in the clinical phase. Different vaccine technologies including protein subunit (Novavax, using a modified spike subunit), virus-like particle (CoVLP), and inactivated virus (CoronaVac, BBIBP-CorV, Covaxin) have also been used (Kyriakidis et al., 2021; World Health Organization, 2022a).

1.3.9 Treatments

Apart from the RdRp inhibitors described in Section 1.1, there are a few other treatments available for treating COVID-19. Many repurposed drugs have been used, although the majority showed a lack of efficacy. More recently, two effective treatments, Nirmatrelvir/Ritonavir (Paxlovid) and sotrovimab (Xevudy) have been approved in the UK (Gupta et al., 2021; Hammond et al., 2022).

Repurposed Drugs

Over 400 repurposed drugs have been tested against SARS-CoV-2. These drugs used can be mainly divided into five groups (Ashour et al., 2022). Various combinations of these antiviral drugs have also been tested.

1) Reverse Transcriptase Inhibitors (targeting RdRp), e.g. remdesivir, ribavirin, favipiravir, molnupiravir, tenofovir
2) Fusion Inhibitors (targeting spike), e.g. umifenovir, camostat mesylate
3) Neuraminidase Inhibitors, e.g. zanamivir, permivir, oseltamivir
4) Protease Inhibitors, e.g. Lopinavir, ritonavir, danoprevir, darunavir
5) M2 ion-channel protein blockers, e.g. amantadine, rimantadine, adamatane

Although majority of these drugs showed efficacy *in vitro*, they were not reflected in clinical trials (Martinez, 2022; T. K. Patel et al., 2021). Clinical trials of many repurposed drugs showed little clinical efficacy or additional benefits compared to drugs specifically developed for SARS-CoV-2 (Bosaeed et al., 2022; Y.-Q. Huang et al., 2020; T. K. Patel et

al., 2021). For example, clinical trials have shown that the use of hydroxychloroquine, an anti-malaria drug which has been proposed to suppress SARS-CoV-2 infections by inhibiting viral entry via sialic acid receptor binding or preventing the cytokine storm (Satarker et al., 2020), has no clinical benefits in COVID-19 patients (Axfors et al., 2021; Omrani et al., 2020; Reis et al., 2021). RdRp inhibitors we discussed in Section 1.1.5, are some of the more successful repurposed drugs used in SARS-CoV-2 infections. Two of them have been approved for use in the UK for treating SARS-CoV-2 infections (UK Parliament, 2022). The clinical application of these drugs will be discussed in Chapter 4.

Protease Inhibitors

One of the more recently approved drugs, Nirmatrlvir/Ritonavir, is a protease inhibitor which can be taken as an oral tablet. It has been shown in two randomised controlled clinical trials involving over 3000 patients that Nirmatrlvir/Ritonavir can reduce hospitalisation by 85%. Given the great efficacy, ease of administration and fewer side effects and concerns compared to other antiviral drugs such as molnupiravir, on 22nd April 2022, the World Health Organisation made a strong recommendation for the use of Nirmatrlvir/Ritonavir in mild to moderately high-risk COVID-19 patients including older, immunosuppressed, or unvaccinated individuals (World Health Organization, 2022b).

Monoclonal Antibodies

Many monoclonal antibodies have been tested or used for treating COVID-19 (e.g. bamlanivimab/etesevimab, casirivimab/imdevimab). Most of them target the spike protein. However, as the virus evolves over time, especially when the Omicron variant became dominant, the

neutralising efficacy for some of these monoclonal antibodies has significantly reduced (Takashita et al., 2022).

In September 2021, the WHO made its recommendation for Sotrovimab (World Health Organization, 2022c). Sotrovimab is a human-engineered neutralising monoclonal antibody (nMAb) cocktail. It neutralises SARS-CoV-2 by binding to a highly conserved epitope in the receptor-binding domain of the spike protein. Early studies have shown Sotrovimab can significantly reduce hospitalisation and death by 79% (Gupta et al., 2021).

Another monoclonal antibody, tixagevimab/cilgavimab, has been authorised to use as pre-exposure prophylaxis in individuals who cannot be vaccinated or showed no immune response post-vaccination (Ely et al., 2021; Kalil et al., 2021; Marconi et al., 2021).

Anti-inflammatory Drugs

For severe to critical COVID-19 patients, the World Health Organisation has made a strong recommendation for anti-inflammatory drugs in combination with corticosteroids. Interleukin-6 (IL-6) receptor blockers (torcilizumab and sarilumab) slow down inflammations by inhibiting the proinflammatory cytokine IL-6 (Cellina et al., 2020; Michot et al., 2020). Baricitinib on the other hand inhibits the endocytosis of SARS-CoV-2 and the intracellular signalling pathway of cytokine which causes hyperinflammation (Richardson et al., 2020). Clinical trials of these treatments showed great efficacy in treating severe COVID-19 patients (Marconi et al., 2021).

Currently more than 690 drug development programs are still in the planning stages in the US (Center for Drug Evaluation and Research, 2022). Combinations of antiviral drugs and vaccinations with high efficacy will be required to control this pandemic.

1.4 Within-Host Evolution

RNA viruses evolve rapidly on the population level. These evolutions often begin with variants arising within infected individuals. As discussed in Chapter 1.1.1, the populations of RNA viruses harbour abundant genetic variability due to their basic properties such as having a large population size and an error prone viral replication mechanism. During the course of infection, viral variants can arise and eventually transmit from one host to another (Grenfell et al., 2004).

The within-host evolution of viral infections is driven by factors such as host immune defences, tissue compartmentalisation and antiviral drug interactions. In acute infections, the within-host viral population accumulate few variants, but in chronic infections, the viral population can undergo substantial evolution, which allows us to identify patterns of selection and adaptation as well as the transmission bottlenecks which limit the intra-host transmission of genetic diversity. The within-host viral evolution provides a substrate for the global population level evolution and analysing the within-host variations allow us to understand and monitor the emergence of new, and potentially drug resistant, variants in epidemics and pandemics (Xue et al., 2018).

Multiple studies have used viral whole genome sequencing to characterise the within-host evolution in RNA viruses such as influenza (Illingworth et al., 2020; Ks et al., 2018; Lumby et al., 2020a). When analysing within-host evolution using deep sequencing, viral populations are usually summarised as a single consensus sequence which represents the most frequent nucleotide at each position across the genome. Variants are analysed to estimate the within-host genetic diversity and the possibility of having a mixed infection (Poon et al.,

2016). These analyses can help to identify the close contacts and reconstruct the transmission chain in the population (McCrone et al., 2017). In influenza, for example, one study analysing within-host viral sequencing data has reported that mixed infections are observed in approximately half of the patients in the cohort (Poon et al., 2016). It has also been shown that the within-host genetic diversity recapitulates the global genetic changes (Xue et al., 2018). Such analyses were used to determine epidemiological factors which transform within-host variations into global population variations (Xue et al., 2018). In chronic infections, the viruses evolve within an individual over a longer period of time, which provides the opportunities for selections to happen. In SARS-CoV-2, it has been suggested that chronic infections were responsible for the emergence of highly transmission new variants such as Delta and Omicron (Chaguza et al., 2022). In influenza, chronic infection datasets have been used to identify patterns of selection and to estimate the mutations rates (Lumby et al., 2020a; Lythgoe et al., 2021). Haplotypes can also be reconstructed to deconvolute the mixed infections and the structure of the within-host viral population (Eliseev et al., 2020a). These methods have been frequently used in RNA viruses such as influenza, and more recently in SARS-CoV-2 (Lumby et al., 2020a; Lythgoe et al., 2021; Xue et al., 2018). However, many existing methods such as those for haplotype reconstruction and mutation rate estimations have not been extensively validated with standardised datasets, resulting in contradicting results from similar studies (Eliseev et al., 2020a; Lumby et al., 2020a; Zanini et al., 2017). Due to the limited availability of data, most methods have not been applied to norovirus studies. In this thesis, we focus on applying these existing methods for analysing norovirus and SARS-CoV-2 longitudinal data spanning across a couple of days to over a year. We will also present more standardised and accurate methods for analysing longitudinal datasets in Chapter 3 and 4.

1.5 Thesis aims and organisation

Both norovirus and SARS-CoV-2 cause significant clinical and economic burdens in countries all over the world. However, treatments available remain limited and the variations in the within-host population remain poorly understood.

In this thesis, we investigate the within-host viral dynamics and evolution in norovirus and SARS-CoV-2. First, to answer the questions on how viral dynamics change during the course of infection, and what these dynamics can tell us about the mechanism of infection, we investigate the longitudinal viral load dynamics in SARS-CoV-2 infections in Chapter 2. We then move on to study the within-host variability and evolution in both norovirus and SARS-CoV-2. In Chapter 3, we present a model used for reconstructing haplotypes to distinguish different viral populations and possible mixed viral infections in deep sequencing data. In Chapter 4, we investigate the effect of antiviral drugs on the norovirus and SARS-CoV-2 within-host viral population.

Chapter 2

Viral Load Modelling

2.1 Introduction

Overview

Understanding the intra-host viral load dynamics can aid the discovery and development of novel treatment plans, as well as inform public health policies. In the early stages of the COVID-19 pandemic in 2020, little was known about the viral dynamics of SARS-CoV-2 with most studies presenting small cohorts of regional patients with disparate interventions. Herein, we provide a more coherent picture of intra-host viral load dynamics by collating data from these early pandemic studies. To this end, we conducted a systematic review that identified 33 case reports, case series and clinical trial datasets from publications from 1st January 2020 to 7th May 2020. We performed a patient-level association analysis of SARS-CoV-2 dynamics to assess the infection profiles of individuals based on the duration of viral shedding, the peak viral load, and the area under the viraemia curve. The study provided a basic understanding of SARS-CoV-2 viral load dynamics at the time. The use of viral load kinetic ordinary differential equation models was also explored.

The analysis in this chapter highlights the heterogeneous characteristics of viral load dynamics in SARS-CoV-2 and provides insights into intra-

host viral dynamics. The results had implications for developing effective anti-viral treatment, vaccination, and ultimately epidemiological control of COVID-19. Potential markers identified could be significant for future research.

What is viral load?

Viral load is a measure of the quantity of virus in a sample, which is usually expressed in the scale of log10 copies per millilitre (Shenoy, 2021). Quantification of the viral load has been frequently included in routine clinical testing. It allows early detection of infections and continuous monitoring of the state of infections (Lescure et al., 2020). It provides an assessment of risk which can be used to support the implementation of treatments in patients (To et al., 2020). Quantifying the viral load is also cheaper and easier than using other infection monitoring methods such as viral whole genome sequencing. Viral load is usually quantified by reverse transcription polymerase chain reaction (RT-PCR) tests (Vogels et al., 2020). A cycle threshold value, which represents the number of RT-PCR amplification cycles needed for the target gene to exceed the detection threshold, is inversely correlated to viral load, meaning the lower the CT value, the higher the viral load. While sensitivity and efficiency vary in different primer-probe sets and diagnostic assays, most methods can consistently detect SARS-CoV-2 viral load down to 40 to 50 copies per 1 millilitre (Vogels et al., 2020). The quantification gives a single viral load value for each sample, which is easy to understand and compare across samples collected over time.

Course of Infection

A minimum infectious dose of viral particles is required to establish an infection. For SARS-CoV-2, it is around 100 viral particles (Karimzadeh et al., 2021). After the infection has been established, an increase in viral load will be observed, which indicates the virus is replicating at a higher rate than it is being cleared. This could be due to the increased level of viral replication or the failure of viral clearance, which is mainly driven by the host immune response (Chachu et al., 2008; Thimme et al., 2001). In SARS-CoV-2, on average, it takes 9 days on average from the point of infection to reach the viral peak (Li et al., 2022). During this period, even before symptoms appear, the patient is already infectious, and the virus can be transmitted to other individuals. For SARS-CoV-2, a high level of viral shedding has been reported in the pre-symptomatic period (He et al., 2020; S. E. Kim et al., 2020; Lescure et al., 2020; Yan et al., 2021). When the rate of viral clearance is higher than the rate of viral replication, the viral load starts to decline. The symptoms might start to disappear, but this is usually followed by another short period of viral shedding, after the resolution of the symptoms. For SARS-CoV-2, the mean viral shedding time from symptom onset is 16.8 days (Yan et al., 2021). Eventually, in healthy individuals, the immune system will fully suppress the viral replication and the viral load will fall below the limit of detection (Challenger et al., 2022; Contreras et al., 2021). However, immunocompromised individuals might fail to clear the virus and the infection will become chronic (Kemp et al., 2021). One study on post-mortem examinations of SARS-CoV-2 infected patients reported that the virus remains viable with a high viral load for up to 16 days following death (Grassi et al., 2022).

Understanding of SARS-CoV-2 viral dynamics in May 2020

As discussed in Chapter 1, the clinical and epidemiological characteristics of COVID-19 are extremely heterogeneous. Cases can range from asymptomatic or mild to severe, which require hospitalisation and admission to ICU. One of the key biomarkers used for clinical monitoring of COVID-19 disease progression is the viral load in nasal swabs or nasopharyngeal aspirates (The Massachusetts Consortium for Pathogen Readiness et al., 2020).

Since the beginning of the pandemic, a substantial number of case reports have assessed the viral load data of SARS-CoV-2. However, in May 2020, when this analysis was conducted, the studies available had yet to paint a coherent picture. Whereas some studies observe higher viral loads in the lower respiratory tract (LRT) vs upper respiratory tract (URT) others report the opposite or no difference (J. Y. Kim et al., 2020; Lui et al., 2020). Since the majority of studies only consisted of a few patients, at the time, it was unclear whether the viral load was correlated to the age, mortality or disease burden.

At the time our study was conducted, there were only two SARS-CoV-2 viral load dynamics analyses which include a large cohort of patients (He et al., 2020; T. Xu et al., 2020). He et al., 2020 studied the dynamics of viral shedding. The viral shedding patterns were plotted and stratified by age, gender and disease severity. Based on these graphs, no significant differences were observed. The authors estimated the peak of viral shedding in the cohort of 77 patients, but no comments were made on the difference in viral shedding patterns by age, gender or disease severity. T. Xu et al., 2020 compared the viral load dynamics of imported and non-imported patients in China, but no comparisons were

made on viral load between different age groups and cases with different disease severity.

Other viral load dynamic analyses were conducted on a single case or a series of a small number of cases (Han et al., 2020; Kam et al., 2020; J. Y. Kim et al., 2020). Their conclusions contradict one another. With the lack of information on clinical status, comorbidities and viral load assay, the findings cannot be replicated and the results were not comparable across different studies.

Motivation

We aimed to gather a standardised data set that could be used to assess viral load in relation to disease severity and other covariates. At the time, according to the PROSPERO systematic review register, no previous or ongoing review had systematically searched for viral load dynamics data using a meta-analysis. To address this, we conducted a systematic review and meta-analysis to advance our understanding of the dynamics of COVID-19 infections and provide relevant evidence during this public health emergency of international concern.

We searched for case reports, case series and clinical trials which report longitudinal individual patient-level SARS-CoV-2 viral load data. We assessed the infection profiles of individuals based on the duration of viral shedding, the peak viral load, and the area under the viraemia curve. Such viral load metrics have been proposed as possible endpoints in drug trials (Duke et al., 2020; Hudgens et al., 2003a; Natori et al., 2018). Characterisation of these metrics from pre-existing covariate data may provide crucial insight for clinical diagnosis, intervention, and future clinical trials.

2.2 Materials and Methods

2.2.1 Search strategy and selection criteria

We conducted a systematic literature search with no language restrictions on PubMed, Ovid and Embase for papers published between the 1st January 2020 (the day after SARS-CoV-2 was first reported to the WHO) and 7th May 2020 (when the search was conducted) using the search terms ("SARS-CoV-2" OR “COVID” OR “coronavirus” OR “2019 -nCoV”) AND ("viral load" OR “cycle threshold” OR “rtPCR” OR “real-timePCR” OR "viral kinetics" OR “viral dynamics” OR “shedding” OR “detection” OR “clinical trial”). We selected studies which reported SARS-CoV-2 viral load or RT-PCR cycle threshold values (CT values) from longitudinal patient samples. Any duplicates in the search were removed. Two reviewers (Dr Florencia A.T. Boshier and I) independently identified papers for full-text screening. Discrepancies were resolved by a third reviewer (Dr Silke Gastine).

2.2.2 Data extraction and processing

Studies with longitudinal viral load data and clinical meta data were selected. Viral load data reported as numerical values in tables and figures in the identified studies were copied and pasted directly into a comma separated value (csv) format. For studies for which the viral loads are not accessible in the article or supplementary files, we contacted the relevant authors to request said data and relevant clinical meta-data. In the case that no response was received from the authors, we digitally extracted the data available from the figures using WebPlotDigitizer (Rohatgi, Ankit, 2019). The limit of detection and the PCR assay primers used was extracted from the studies. The following patient-level meta data were extracted when they were available: age, sex, disease status, symptoms, presence of fever, requirements for intensive care treatment, requirements for mechanical ventilation, and the treatment used.

To account for the different levels of sensitivity of various RT-PCR assays, we converted the CT values into viral load copies per mL using the calibration curve provided in the study or reported in (Vogels et al., 2020), based on the primers used.

We converted the sampling dates as days since symptoms onset, where day 1 is the day on which the patient first shown any COVID-19 related symptoms. For asymptomatic patients, day 1 is the day on which they first tested positive. We grouped samples collected from various sites into 8 categories: 1) upper respiratory tract (nasopharyngeal swab, oropharyngeal swab, pharyngeal swab, nasal swab, throat swab, oral swab, saliva, endotracheal aspirate, sputum); 2) lower respiratory

tract (bronchoalveolar lavage); 3) upper gastric tract (gastric fluid); 4) lower gastric tract (stool, anal swab, rectal swab); 5) eye (ocular swab, conjunctival swab); 6) blood (serum, plasma); 7) urine; and 8) breastmilk.

As discussed in Chapter 1, SARS-CoV-2 infections can range from asymptomatic to severe. To access the correlation between viral load and the disease severity, we classified the patients into 5 categories based on the clinical observations or symptoms reported in the studies: 1) asymptomatic, where the patient was tested positive but did not show any sign of symptoms; 2) mild, where the patient showed symptoms such as light coughing and fatigue but did not develop pneumonia; 3) moderate, where the patient showed fever and respiratory symptoms, and had developed pneumonia; 4) severe, where the patient had difficulty in breathing and required oxygen supply; and 5) critical, where the patient required mechanical ventilation and intensive care treatment (National Health Commission of the People's Republic of China, 2020; World Health Organization, 2020).

2.2.3 Statistical analysis

We considered three viral load metrics to assess the infection burden on individuals. The first is the length of viral shedding, we define this as the longest time between first positive and first consecutive negatives reported across any site or assay for an individual. The second metric, peak viral load is the maximum viral load observed during the observation period by site and assay for an individual. We should note that this is not the true peak of the infection which was not generally captured in the data since the peak of viral load is often observed on or before the symptoms onset (He et al., 2020). Finally, we consider area under the viraemia curve (AUC) by site and assay for each individual. The AUC gives a measure of the total amount of virus shed over time and can be used as a marker of infectiousness, severity of infection and the speed of viral clearance. We calculated AUC using the auc function from the MESS package in R Studio version 1.4.1106 (Ekstrøm, Claus Thorn, 2019).

Normality of data was assessed visually and using the Kolmogorov-Smirnov (with Lilliefors correction) and the Shapiro-Wilk normality tests. Continuous variables are presented as median (interquartile range, IQR). The three viral metrics described above were compared across each non-parametric covariates using the Mann-Whitney U test, implemented using the wilcox.test in R Studio version 1.4.1106 (R Core Team, 2019). We used Bonferroni correction to adjust for multiple comparisons.

We used Pearson's correlation coefficient and linear regression for all correlation analysis. This was done using the cor and cor.test function in the stats package in R Studio version 1.4.1106 (R Core Team, 2019).

2.3 Results

2.3.1 Data selection

The systematic literature search identified 30 studies for inclusion. There were 3 additional studies identified through other sources, in which two were pre-prints available on medRxiv and one was missed by the database search terms (Figure 1). We have identified a total of 347 patients with longitudinally collected viral loads or CT values data.

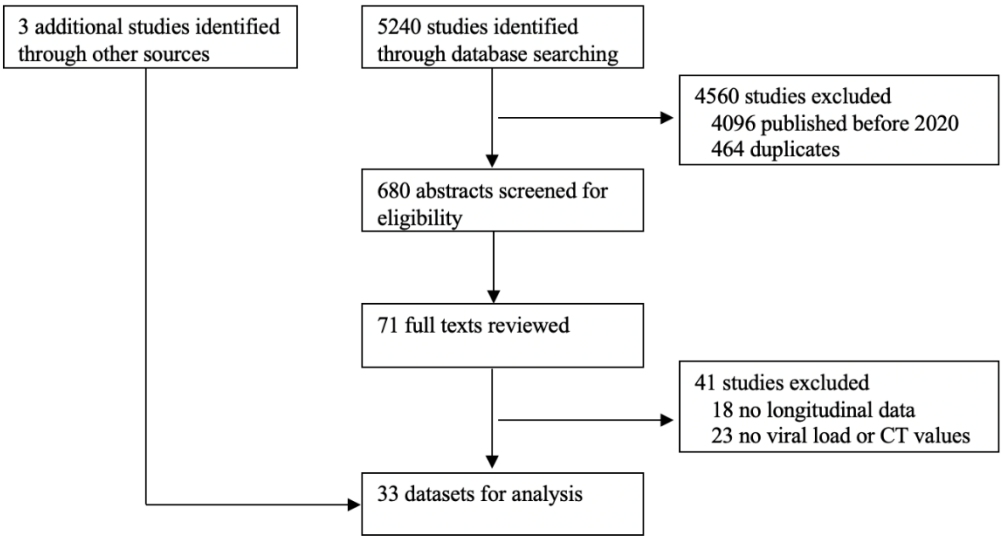


Figure 1. A flow chart of the process of data selection.

Among these 33 studies, 22 reported the individual patient's age and sex, 4 reported the median age and gender distribution of the cohort; 8 reported the individual patient's comorbidities, 5 reported all comorbidities found in the cohort; 23 reported individual patient symptoms, 3 reported all symptoms observed in the cohort; 20 reported treatments received by individual patients, and 1 reported all treatments used by the cohort (Table 1).

Table 1. A summary of the 41 studies identified. *Details including the number of patients, sample types, assay genes, treatments, comorbidities and symptoms are listed. The asterisks indicate only cohort level information is provided, no details on individual patients are available.*

Authors	Number of Patients with Longitudinal Data	Location, Country	Sample Type	Assay Gene	Treatment	Comorbidities	Symptom details	Age median (IQR) {R}	Gender M (F)
Kam, KQ, et al.(Kam et al., 2020)	2	Singapore	Blood, Urine, Stool, Nasopharyngeal, Breastmilk	N, Orf1lab	No treatment used	Not reported	Reported	0.5, Only reported infant's age	1 (1)
Kim, JY, et al.(J. Y. Kim et al., 2020)	2	South Korea	URT, LRT, Serum, Plasma, Urine, Stool	RdRp, E	LPV/r, Ceftriaxone, Moxifloxacin	Not reported	Reported	45 (35 -55)	1 (1)
To, KKW, et al.(To et al., 2020)	23	Hong Kong	Saliva, Endotracheal aspirate, Remnant serum	RdRp	LPV/r, Ribavirin, Interferon	Reported*	Reported* (not longitudinal)	66 {37-75}*	13 (10)*
Lim, J, et al.(Lim et al., 2020)	1	South Korea	Throat swab	RdRp	LPV/r, Azithromycin, Ceftriaxone,Tazobactam, Levofloxaxin	Reported	Reported	54	1 (0)
Pan, Y, et al.(Pan et al., 2020)	2	Beijing, China	Throat swab, Sputum	N	Not reported	Not reported	Not reported	Not reported	Not reported
Zou, L., et al.(Zou et al., 2020)	14	Zhuhai, Guangdong, China	Nasal swab, Throat swab	N, Orf1lab	Not reported	Not reported	Reported (not longitudinal)	59 (50-68.25)	5 (9)
Zhang, W, et al.(Zhang et al., 2020)	16	Wuhan, China	Oral swab, Anal swab, Blood	S	Not reported	Not reported	Not reported	Not reported	Not reported
Gautret, P, et al.(Gautret et al., 2020)	26	South France	Nasopharyngeal swab	RdRP, E	Hydroxychloroquine, Azithromycin	Not reported	Not reported	48 (24-60)	13 (13)
Xu, T, et al.(T. Xu et al., 2020)	49	Changzhou, China	Throat swab	N, Orf1lab	LPV/r, Interferon, Methylprednisolone, Umifenovir, Thymosin	Reported	Reported	43 (28.5-53)	24 (25)
Chen, W, et al.(W. Chen et al., 2020)	6	Guangzhou, China	Pharyngeal swab, Blood, Anal	N, Orf1lab	Not reported	Not reported	Not reported	Not reported	Not reported

He, X, et al.(He et al., 2020)	52	Guangzhou, China	Throat swab	N	Not reported	Not reported	Not reported	46 (33-61)*	47 (47)*
Young, BE, et al.(Young et al., 2020)	18	Singapore	Nasopharyngeal swab	N, S, and Orf1ab	LPV/r	Reported*	Reported	47 {31-73}*	9 (9)*
Shen, C, et al.(Shen et al., 2020, p. 5)	5	Shenzen, China	Nasopharyngeal swab	Not reported	LPV/r, Interferon, Flavipivir, darunavir, Methylprednisolone	Reported	Reported	Age range reported	3 (2)
Wan, R, et al.(Wan et al., 2020)	2	China	Throat swab	Orf1ab	LPV/r, Ribavirin	Not reported	Reported	27.5 (19-36)	2 (0)
Lescure, FX, et al.(Lescure et al., 2020)	5	France	Nasopharyngeal swab, Stool	RdRp, E, RdRp-IP1, GAPDH	Remdesevir	Reported	Reported (not longitudinal)	46 (31-48)	3 (2)
Wölfel, R, et al.(Wölfel et al., 2020)	9	Munich, Germany	Pharyngeal swab, Sputum, Stool	RdRP, E	Not reported	Reported	Reported	40 (33-49)	8 (1)
Xu, Y, et al.(Y. Xu et al., 2020)	9	Guangzhou, China	Rectal swab, Nasal Swab	N, Orf1lab	Not reported	Not reported	Reported	6 (3-13)	6 (3)
Han, MS, et al.(Han et al., 2020)	2	Seoul , Korea	Nasopharyngeal swab, Oropharyngeal swab, Stool, Plasma, Saliva, Urine	E	No treatment used	Not reported	Reported	0.08, Only reported neonate's age	0 (2)
Wyllie, AL, et al.(Wyllie et al., 2020)	19	USA	Saliva, Nasal swab	Not reported	Not reported	Not reported	Not reported	{23-92}, Mean = 61*	23 (21)*
Cheng, CY, et al.(Cheng et al., 2020)	5	Taiwan	Oropharyngeal swab, Sputum	RdRp1, RdRp2, E, N	LPV/r, No treatment	Not reported	Reported	52 (50-53)	2 (3)
Yang Y, et al.(Y. Yang et al., 2020)	13	Guandong, China	URT, LRT	Not reported	Antivirals treatment start date given, no further details	Reported	Reported	36 (34-65)	2 (1), Not reported
Yang, JR, et al.(J.-R. Yang et al., 2020)	1	Wuhan, China	Orophrangyeal swab	Not reported	Oseltamivir, Ganciclovir, Cefoperazone, Tazobartam, Arbidol, Methylprednisolone, Inteferon, Thymalfasin, Chloroquine	Not reported	Reported (not longitudinal)	44	1 (0)
COVID-19 Investigation Team.(COVID-19	12	USA	Nasopharyngeal swab, Oropharyngeal swab	Not reported	Remdesevr, Oseltamovir	Reported	Reported (not longitudinal)	Age range reported	5 (2), Not reported

Investigation Team, 2020)									
Lui, G, et al.(Lui et al., 2020)	4	Hong Kong	Nasal swab, Sputum, Plasma, Stool	Not reported	LPV/r, Ribavirin, Interferon*	Reported*	Reported (not longitudinal)	Not reported	Not reported
Hu, Y, et al.(Hu et al., 2020)	3	Zhejiang, China	Nasopharyngeal swab, Anal swab	ORF1ab, N gene	Inteferon, Thymalfasin, Chloroquine	Not reported	Reported	28 (25-32)	2 (1)
Seah, IYJ, et al.(Seah et al., 2020)	17	Singapore	Nasopharyngeal swab	E	Not reported	Not reported	Reported*	Not reported	Not reported
Liu, WD, et al.(W.-D. Liu et al., 2020)	1	Taiwan	Orophrangyeal swab, Sputum, Throat wash	N, RdRp, E	Not reported	Not reported	Reported (not longitudinal)	50	0 (1)
Xing, YH, et al.(Xing et al., 2020)	1	Qingdao, Shandong, China	Orophrangyeal swab, Fecal sample	Not reported	Interferon and Ribavirin	Not reported	Reported (not longitudinal)	1.5	1 (0)
Qian, GQ, et al.(Qian et al., 2020)	1	Ningbo, China	Orophrangyeal swab, Rectal	N, RdRp, E	LPV/r, Umifenovir, Interferon, Chinese medicine	No comorbidities	Reported	47	1 (0)
Kim, ES, et al.(E. S. Kim et al., 2020)	9	Korea	URT, LRT	E	LPV/r, No treatment	Reported*	Reported*	Not reported	Not reported
Colavita, F, et al.(Colavita et al., 2020)	1	Rome, Italy	Nasal swab, Ocular swab	Not reported	Not reported	Not reported	Reported	65	0 (1)
Hill, KJ., et al.(Hill et al., 2020)	1	Scotland	Nasal swab, Oropharangyeal swab	Not reported	No treatment used	Reported	Reported	65	0 (1)
Huang, Y, et al.(Yongbo Huang et al., 2020)	16	Guangzhou, China	Nasal swab, Oropharangyeal swab, Sputum, Conjutival swab, Serum, Plasma, Urine, Gastric fluid, Faeces, Anal swab	Not reported	Not reported	Reported*	Not reported	Not reported	Not reported

2.3.2 Viral Load Trajectories

While some studies have reported viral load in copies per mL, others only have CT values available. We considered the normalisation of viral load and CT values using a log scale. Along with some studies which provided both viral load and CT values, we found that the CT values and log transformed viral loads are negatively correlated ($r^2 = 0.78$, p-value < 0.01) (Figure 2). This indicates that CT values can be as a proxy of viral load (copies/ml) and either measurement can be used to evaluate the progression of the disease.

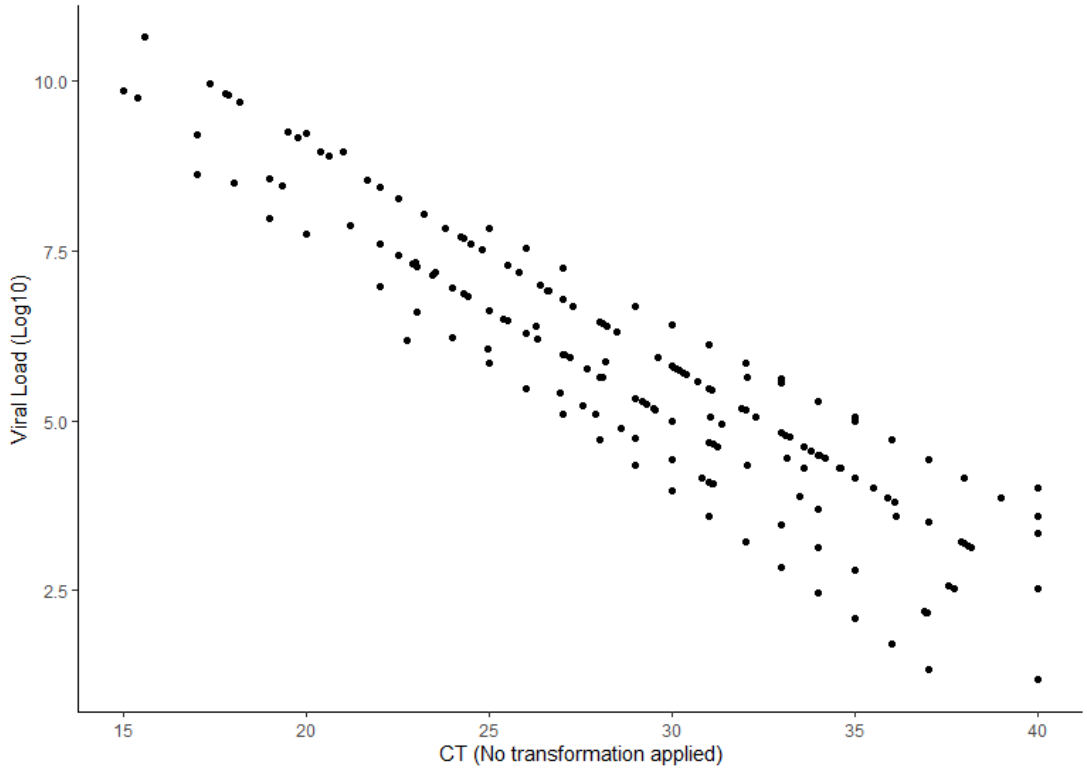


Figure 2. Correlation between CT values and log-normalised viral load values. *The CT and log_{10} viral load values are negatively correlated.*

The viral loads of all individuals identified, across all sites and assays are shown in Figure 3. We observed three major types of viral load kinetics. The viral load kinetics of three randomly selected representative individuals have been highlighted in Figure 3A. The trajectory highlighted in green showed a clear drop in viral load within a seven-day period. The first viral load value recorded from this patient was already lower than the two other patients highlighted in red and blue. This was the most common viral load kinetic observed in our data set. Since SARS-CoV-2 has an incubation period of two to 14 days, the viral load usually peaked before the patient presented with any symptoms (S. E. Kim et al., 2020). Apart from in test and trace studies, most viral load quantifications were carried out only when a patient became symptomatic. It is likely that only the final stage of infection was captured in this patient. The red trajectory showed one of the patients with the longest shedding period of 60 days from symptom onset. The blue trajectory showed a patient with a viral load that fell below the limit of detection in every other sample. Recurrent positives observed in the red and blue patients have been reported in more recent studies. It has been reported that 17% of patients showed positive RT-PCR tests following consecutive negative results. These positives have lower viral load than the initial samples and are less likely to be infectious (Gao et al., 2021).

Across all patients, we found that most of the shedding episodes occur approximately three weeks post symptoms onset, the densest observation window in the figure (0 to 21 days). However, several individuals shed the virus for more than 40 days with one shedding up to day 63 post symptom onsets.

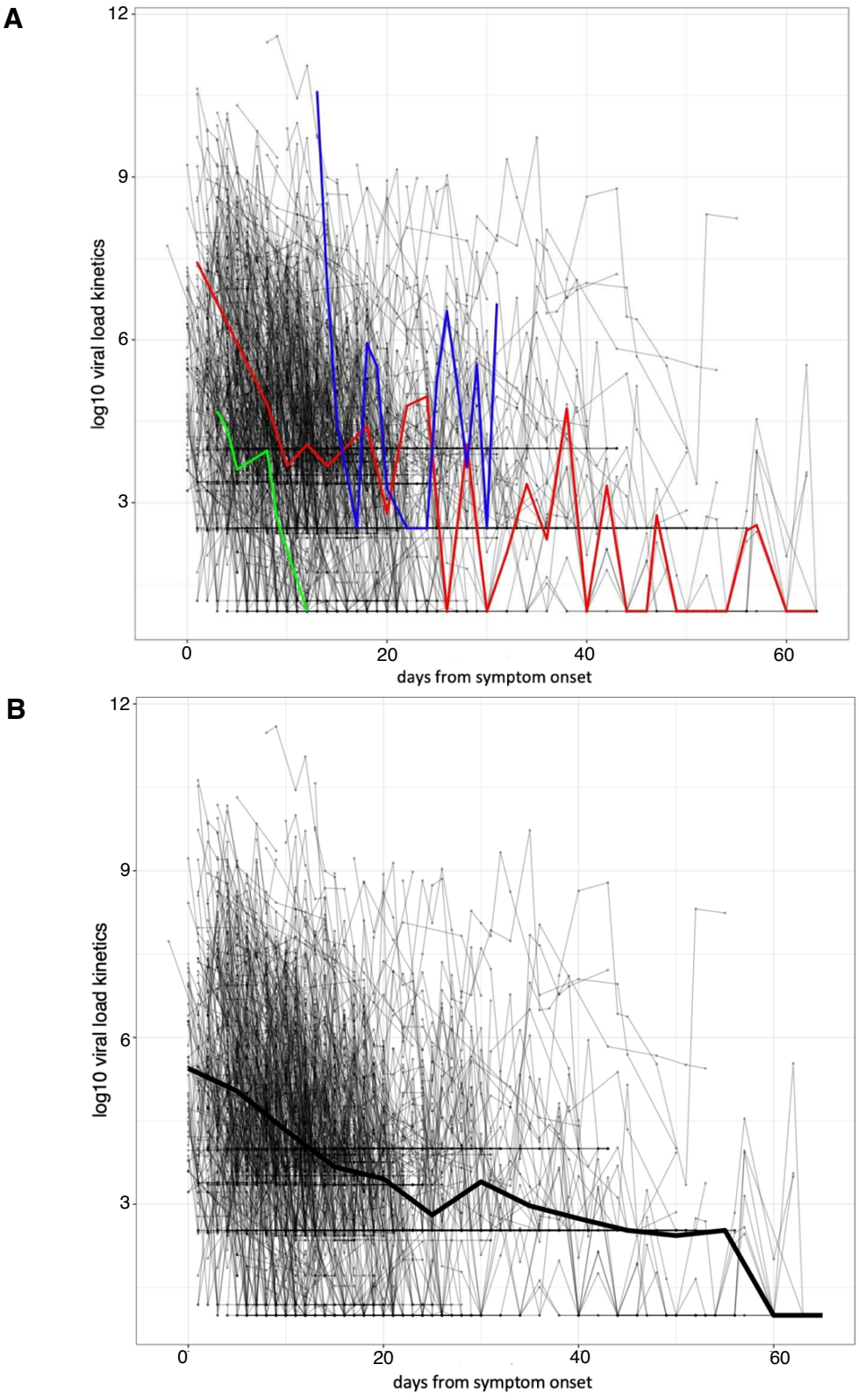


Figure 3. SARS-CoV-2 viral load kinetics of all identified patients. *Y-axis shows the log_{10} viral load values, x-axis shows the days from symptom onset. (A) Three representative individuals highlighted. (B) Median highlighted.*

2.3.3 Statistical analysis

We have identified 7 covariates in the collated data set: patients' age and sex, fever status, drug administration, ICU admission, disease status, sampling site, geographical location. We next analysed whether the three viral load metrics AUC, peak viral load, and length of viral shedding, are significantly different and if covariates can explain any of these differences.

Age and sex

First, we considered demographical covariates. A comparison of the viral load metrics by age and sex did not show any statistical significance (p-value >0.05, Mann-Whitney U test).

Disease severity

We next considered covariates related to disease severity, including the need for ICU admission, drug administration, fever recorded and disease status. Individuals admitted to ICU were found to shed virus for longer and had a lower peak viral load and AUC compared to individuals who did not require intensive care (Figure 4A). The time between symptom onset and the first positive viral load for individuals in ICU was greater (median=12, IQR: 10-18) than that observed in individuals not in ICU (median=4 IQR: 2-9) (p-value <0.05 Mann-Whitney U test) (Figure 4A). The mean and median viral load were higher in the non-ICU patients than ICU patients (2.4, 1.43 and 1.98 and 1 respectively). This suggests that when individuals were admitted to ICU their viral load had already significantly declined, probably indicating that these patients

were observed later in the course of infection. The peak viral load for individuals who were administered antiviral drugs during infection was lower than those who were not given any treatments (Figure 14B). This latter group did not include any patients recorded to have been admitted to ICU, whilst the former group included 9 such individuals. The recording of fever at any point during infection, was associated with higher AUC and increased length of shedding ($p < 0.05$), however there was only a tendency of higher peak viral load for those reporting fever ($p = 0.10$) (Figure 4C). Severe cases of infection were associated with lower peak viral load, longer shedding time and lower AUC ($p<0.05$), similarly to what was found in patients admitted to ICU (Figure 4D).

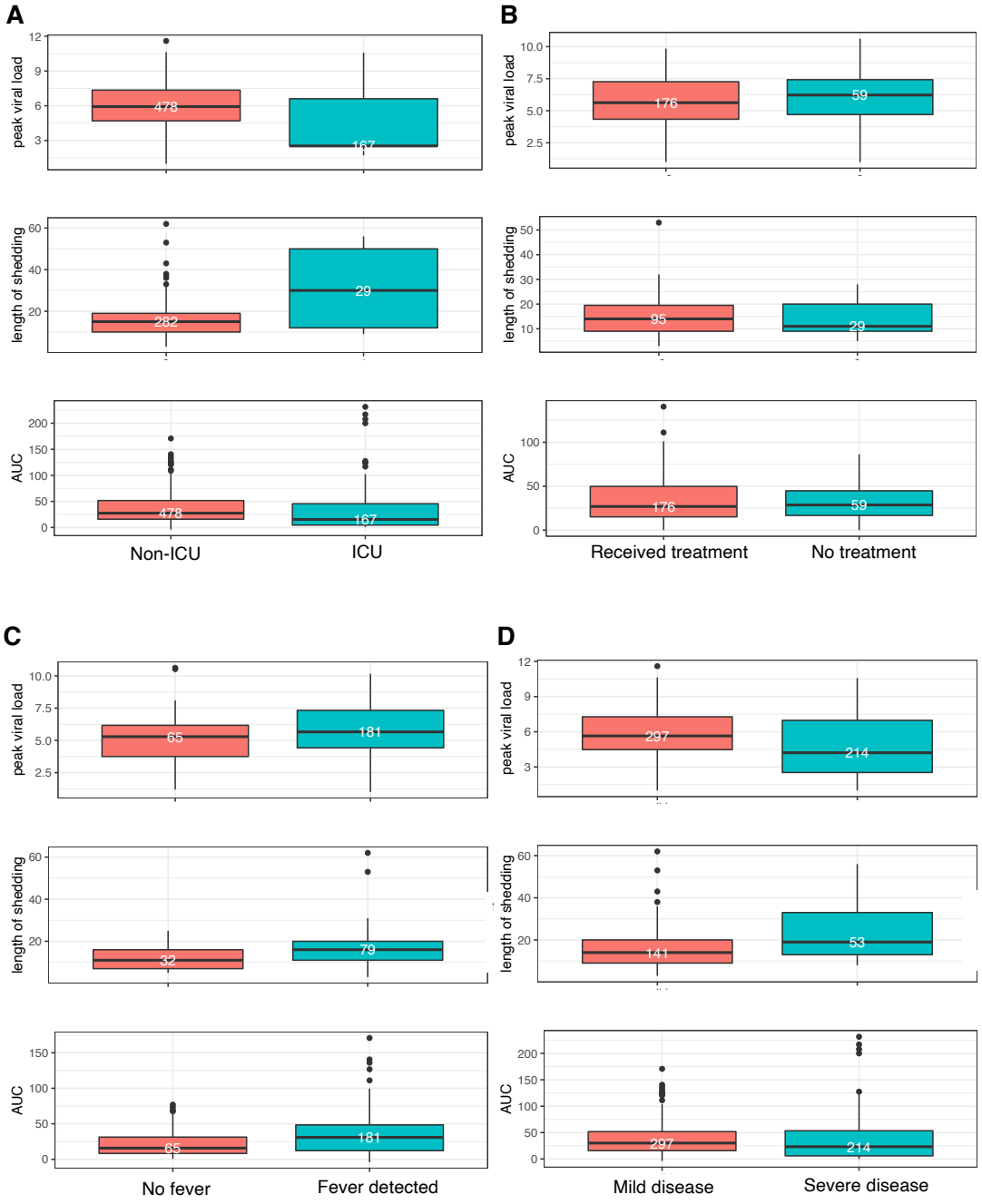


Figure 4. Comparison of three viral load metrics (peak viral load, length of shedding, area under the viraemia curve) by (A) the admission to intensive care unit, (B) the use of drugs, (C) the presence of fever and (D) the disease severity. *The number of samples included in each group is indicated on the boxplots.*

Sampling sites

Next, we evaluated the three metrics in the upper respiratory tract (URT), lower respiratory tract (LRT) and lower gastrointestinal tract (LGT) samples (Figure 5). We found that the LGT presents with lower AUC and peak viral load than URT, however the length of shedding was not found to be different. Since taking samples from LRT involves invasive procedures, there were very few measurements, which reduced the statistical power for formal comparison with the other sample sites.

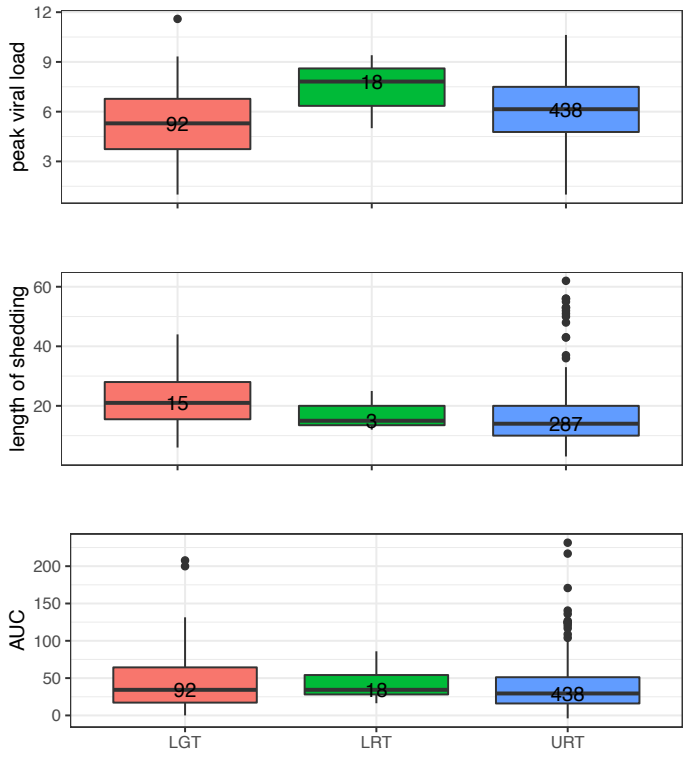


Figure 5. Comparison of three viral load metrics by the sample sites, lower gastric tract, lower respiratory tract or upper respiratory tract. *The boxplots are coloured based on the sampling sites. Red = lower gastric tract, Green = lower respiratory tract, Blue = upper respiratory tract.*

In addition, the three viral metrics were also found to vary by geographical location of the study. Individuals from the US were found to have larger AUC and longer shedding times compared to Europe and Asia ($p < 0.05$). Those from Europe had the lowest length of shedding but the highest peak viral load amongst all ($p < 0.05$). The peak viral load was higher in the US than in Asia ($p < 0.05$).

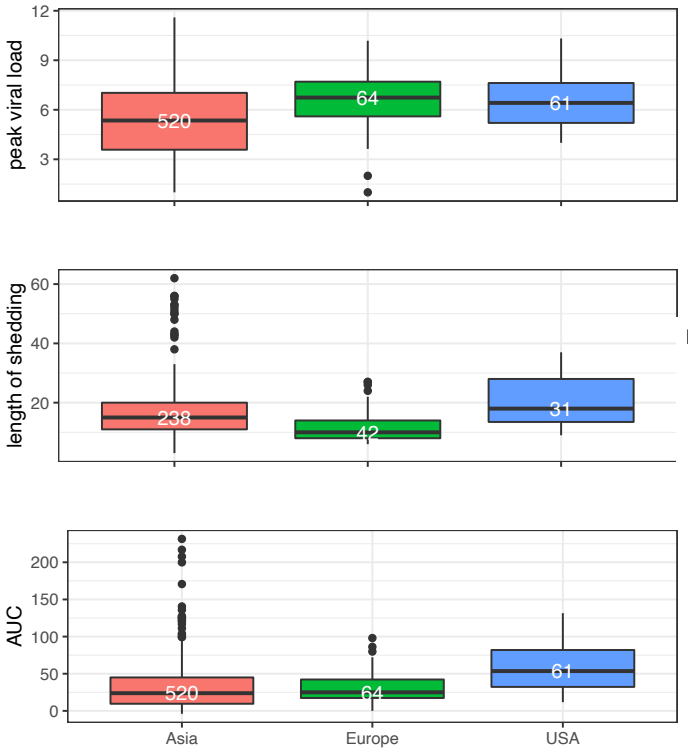


Figure 6. Comparison of the three viral metrics by the geographical location of the patients. *The box plots are coloured based on geographical locations. Red = Asia, green = Europe, blue = the US.*

2.4 Discussion

2.4.1 Limitations

At the time when this analysis was done, it was the first to include data from numerous geographical locations and laboratories. This endeavour had provided a set of novel challenges in collating and aligning the data coherently. For example, for most patients, the sample collection time points were recorded as days since symptom onset. We relied on this information to align the viral load trajectories. However, this information mainly depended on self-reporting, which was not clinically objective, especially when the list of COVID-19 associated symptoms had changed over time. In contrast, in test and trace studies, where the patients would have been continuously monitored before showing any symptoms, the day of symptom onset was more reliable. The inconsistency in the definitions of the point of infection across multiple studies posed a great challenge to align the data in our cohort. Varying scales of disease severity classification had also been reported across the studies, which we tried to mitigate by applying our own scale based on the the metadata information available to us. The difference in sampling techniques and technologies has been proven to create bias in viral load measurements across different studies. Unfortunately, replicate samples were not available because of the limited level of resources and staff support available during the pandemic and the potential to cause ethical concerns, especially in critically ill patients.

In addition, these viral load data were collected at a very early stage of the pandemic. At the time there was no known effective treatment for COVID-19. A wide variety of treatment strategies were in place across

the world. Consequently, treatment varied widely by geographical location, but also evolved during the pandemic as our understanding grew. As we primarily focused on viral load metrics, we considered only the reported antiviral drugs as treatments. However, there were other treatment regimens that may have impacted the disease progression such as the use of traditional Chinese herbal medicine and antibiotics. Despite these limitations, we were able to make some important observations from this work.

2.4.2 Comparison with recent studies

Sampling sites

We found that the viral metrics we considered did not vary widely by sample site. Whilst there was a tendency for the virus to be detected in the LGT at later time points, contrasting with previous findings, this was not found to be statistically significant (Pan et al., 2020). A recent study showed that sputum samples have an extended period of viral shedding compared to nasopharyngeal swab samples, although it has been suggested that suboptimal sampling of nasopharyngeal swabs might have contributed to the discrepancy (Irifune et al., 2021). Another study has reported that viral load can be detected in saliva 4.5 days before being detected in nasal swabs. Nasal swabs, however reached a higher peak viral load (Savela et al., 2022). In our analysis, due to the lack of granularity of data, saliva and nasal swabs were both grouped into the upper respiratory tract samples. Therefore, we were unable to demonstrate such differences.

Geographic location

In the geographic comparisons, we noticed that patients from the European studies have the highest peak viral load compared to the US and Asia. Although this could be due to the different compositions of ethnicities in the population, this observation is more likely to be confounded by the fact that there were different guidelines on sample collection, sample storage and assays employed across studies. PCR protocols from earlier studies might have lower sensitivity and accuracy. It can also be explained by the public health policies adopted in different

countries. Public health policies played a crucial role during the pandemic. They had great influence over the data we collected. For example, most Asian countries had been doing test and trace where patients, even with mild symptoms, tend to be admitted and treated earlier, which was not necessarily the case in the rest of the world, where there was a tendency to only test and admit more severe patients when necessary (Amul et al., 2022; Raoofi et al., 2021). Many of these severely infected patients might have pre-existing medical conditions which we do not have information on and could potentially skew the data. In addition, most East Asian countries had strict protocols for discharging COVID-19 patients. Patients, in general, were only discharged after multiple consecutive negative PCR tests, which means viral shedding was closely monitored and shedding times were likely longer as a result. This was less likely to be observed in other parts of the world. The association between geographical factor and viral load has not been analysed in more recent studies.

Disease severity and demographics

At the time, it was still unclear whether viral load metrics are determinants of severity of disease, however viral load is often used as a marker for surrogate endpoints in clinical trials (Hudgens et al., 2003b; Smith and Stein, 2002). In our study, we found that patients admitted to ICU were typically in a later stage of the disease. They also appeared to have a more prolonged period of viral shedding. Individuals found in ICU were further along in their disease progression, but they still presented with a substantial decline in viral load, in line with less serious cases observed at the same point.

A number of new studies focused on the duration of viral shedding and association of viral dynamics between different age groups, sex and disease severity have been published after our analysis was conducted. While one study, in line with our findings, suggests there is no statistical difference in viral shedding time (the AUC metric we used) between different sex and age group (B. Zhou et al., 2020), other studies suggest that the duration of viral shedding is positively correlated with age (Cevik et al., 2021; X. Chen et al., 2020; Sakurai et al., 2020; Talmy et al., 2021). In one study, viral load from nasal swabs has been found to increase with age, which could explain the correlation of the increased risk of severe disease in elderly individuals (Euser et al., 2021).

Most studies found that the viral load in upper respiratory tract peaks before or at the time of symptom onset and that the viral load dynamics are similar between symptomatic and asymptomatic individuals (Cevik et al., 2021; Wölfel et al., 2020; Wyllie et al., 2020; Zou et al., 2020). At higher viral loads individuals are likely to be more infectious (The Massachusetts Consortium for Pathogen Readiness et al., 2020).

While the majority of studies showed a positive correlation between the peak viral load and mortality (Bryan et al., 2020; Satlin et al., 2021; Soria et al., 2021), three studies suggested there is a negative correlation between the two (Argyropoulos et al., 2020; Carrasquer et al., 2021; Hasanoglu et al., 2021). A recent human challenge has shown that there is no correlation between viral load and symptoms (Killingley et al., 2022).

Drugs

Due to the lack of granular data, we were not able to study the effect on viral load caused by each antiviral drug. This has vastly limited the impact of our results. Since we did not have any information on the drug dosage or a good understanding of the effectiveness of antiviral drugs on different populations or age groups, we could not conclude when and whether the administration of a drug may result in a decrease in viral load such that it will impact disease severity and progression. Our analysis suggested that the outcome of individuals in ICU may not be as greatly affected by the reduction of viral load caused by antiviral drugs.

We found statistical differences between the peak viral load of patients treated with antivirals and patients who did not receive any treatments. However, since different drugs, dosage and treatment period was used across different studies, we cannot conclude that the difference in peak viral load was due to the drugs.

The majority of clinical trials use viral load as a key marker for determining the efficacy of the drugs (Pandey et al., 2020). Pharmacokinetic modelling is used to analyse the data (Kern et al., 2021). As the recent clinical trials have clearer protocols, more standardised cohort of patients and greater granularity of data, most of the limitations we discussed have been addressed.

The statistical analysis in our studies provided preliminary insights on SARS-CoV-2 viral load metrics. The acquisition of further data in the form of clinical trials, in conjunction with mechanistic mathematical models, will capture the full viral load dynamic changes over time and advance our understanding in this regard.

2.4.3 Alternative methods for studying viral load dynamics

In other well-studied RNA viruses, various types of models have been employed to study the viral load dynamics. For example, in (Baccam et al., 2006), the authors used an ordinary differential equation (ODE) model to study the viral load dynamics on influenza A H1N1 viruses in the upper respiratory tracks of experimentally infected adults. The model describes the number of uninfected target cells, productively infected cells, and infectious viral titre over time. The model considers four constant parameters, namely, the infection rate, viral replication rate, virus death rate and virus clearance rate. To make the model more realistic, the authors have also extended the initial model to incorporate the effect of host immune response and the potential delay in viral replication. The model was used to study viral load kinetic data under the use of an antiviral drug, zanamivir, and provide insights into the efficacy of antiviral therapies for influenza A infections. Variations of ODE models taking into consideration pharmacokinetics and pharmacodynamics (PKPD), detailed immune response, shedding and transmission was used to study other viruses, such as respiratory syncytial virus and measles virus (González-Parra and Dobrovolny, 2018; Kombe et al., 2019; Lee et al., 2017; Lin et al., 2012; Patel et al., 2019).

In the context of SARS-CoV-2, (K. S. Kim et al., 2020) employed a similar target cell limited ODE model used in (Baccam et al., 2006) to study the viral load dynamics. The model describes the changes in the number of uninfected target cells and the infectious viral titre over time. These parameters were fitted and estimated using a non-linear mixed

effect model. One of the most interesting findings in this study is that they conclude treatments that block viral replications are only effective if they were administrated before the viral load has peaked.

As part of the collaboration with Professor Joe Standing and his research group, we have also further explored the dataset described in this chapter using mathematical models. First, a Cox proportional hazards model was used to study the viral clearance in the patients. In line with other studies, we found that the viral clearance from the upper respiratory tract was the quickest, while the stool was the slowest (Cevik et al., 2021; Walsh et al., 2020). In addition, we found that older individuals, males, and patients with more severe diseases have a relatively long viral clearance time. The use of drugs such as remdesivir, ribavirin and interferons were found to shorten the viral clearance time and reduce the viral load.

Apart from the above analysis, we have also adapted the non-linear mixed effect model described in (K. S. Kim et al., 2020). It was used to model the individual viral load level over time in NONMEM and Monolix. Various covariates discussed in the study described in this chapter were considered. In addition, the effects of different antiviral drugs and combinations of antiviral drugs were investigated using a simulation. Full detail can be found in Gastine et al., 2020. The major limitation of our study is the small sample size and the lack of detailed patient records in multiple datasets we included. Having clinical trial data would significantly increase the accuracy of our analysis and would verify the conclusions.

2.4.4 Closing remarks

Understanding the intra-host viral dynamics is an essential component to developing effective antiviral treatment, vaccination, and ultimately epidemiological control of COVID-19. Our initial study described for the first time the characteristics of viral load markers across numerous studies. We analysed summary statistics of viral load which highlighted interesting characteristics and potential markers for future research. Our further study in collaboration with Professor Joe Standing and his group demonstrated the use of mathematical models for the understanding of viral load dynamics.

The major limitation of our study is the lack of good quality data and clear clinical information. Inconsistent sampling techniques and poor documentation of drug usage limited the reliability of our study. Yet, our analysis provided a basic understanding of viral load at the beginning of the pandemic. Results from our analysis are comparable to more recent studies (Cevik et al., 2021; Wölfel et al., 2020; Zou et al., 2020).

Most recent studies on viral load have a bigger cohort of patients which are clearly grouped by covariates such as age, sex and comorbidities (Satlin et al., 2021; Soria et al., 2021),. Viral load data from clinical trials are attached with clear information on drug usage, which allow better comparison across multiple studies.

Two years into the pandemic, we now have a better understanding of the viral load dynamics of SARS-CoV-2 infections. It is important to note that as the pandemic progresses and new variants such as omicron arise, we might observe changes in the viral load dynamics over time.

While the viral load dynamics provide a good overview of the infection, sampling and sequencing the actual virus will provide a better picture.

Chapter 3

Haplotype Reconstruction

3.1 Introduction

Overview

Viral whole genome sequencing can provide a deeper understanding of the infections than viral load data. However, the current standard practice for assembling next-generation sequencing (NGS) data involves summarising thousands of reads into a consensus sequence which often fails to capture the true diversity in each sample and the evolution over time.

Viral mixed infections, where multiple variant strains are present in one sample, are not uncommon. The process to obtain different sequences for each viral strain in a sample is called haplotype reconstruction, which involves identifying sets of variants which are co-located on the same genome. However, most currently available haplotype reconstruction programs are developed for human immunodeficiency virus (HIV), a small and fast-evolving virus, which makes these programs not suitable for more slowly evolving viruses and too compute-intensive for larger viruses.

In this chapter, we present a new approach, HaROLD (HAplotype Reconstruction Of Longitudinal Deep sequencing data), which takes

advantage of longitudinal clinical data and performs this reconstruction using a probabilistic framework. We evaluated HaROLD with both RNA and DNA viruses synthetic and clinical Illumina sequencing data and we compared the performance of HaROLD with other top performing haplotype reconstruction methods. Finally, we demonstrated how an accurate haplotype reconstruction can be used to understand the evolution of within-host viral populations.

Benefits of using next generation sequencing

Next Generation sequencing (NGS) can improve our understanding of the fundamental biology of the virus, inter- and intra-host pathogen evolution, as well as the development of drug resistance. Illumina, one of the most commonly used sequencing technology, generates highly accurate whole genome sequence with good coverage (Heather and Chain, 2016; Segerman, 2020; Slatko et al., 2018), which allows us to characterise the genetic diversity of a within-host viral population.

It has become increasing popular to include NGS in routine clinical diagnostic tests as sequencing technologies are more mature and cost-effective. However, in complex clinical cases, statistical models are often required for analysing these data. For example, when a patient has a mixed infection where multiple strains of viruses or haplotypes are present, it is important to identify the differences between these strains and estimate their respective frequencies in the population. A mixed infection can occur through multiple independent infection events or when there are subclones arising from within-host evolution (Ghedin et al., 2009; Ross et al., 2011). This is particularly common in immunocompromised individuals (Yu et al., 2020). The process of

identifying which variants belong to the same viral genome is known as haplotype reconstruction (Schirmer et al., 2014).

Methods for haplotype reconstruction

Haplotype reconstruction is commonly performed by identifying two or more variants which are observed together on the same Illumina reads. If there are enough variants distributed equally across the genome, a whole genome sequence of the haplotype can be reconstructed by stitching together the overlapping reads. Over the past decade, a number of haplotype reconstruction programs have been developed to assemble NGS sequencing reads into whole genome haplotype sequences (Astrovskaya et al., 2011; Beerenwinkel and Zagordi, 2011; Pelizzola et al., 2021; Prabhakaran et al., 2014; Pulido-Tamayo et al., 2015; Töpfer et al., 2014; Yang et al., 2013; Zagordi et al., 2011). The majority of these haplotype reconstruction programs were developed for HIV, which has highly variable genome, rather than other RNA viruses or DNA viruses with a larger and less variable genome.

Two of the top-performing haplotype reconstruction programs evaluated by Eliseev et al., 2020, CliqueSNV (Pelizzola et al., 2021) and PredictHaplo (Prabhakaran et al., 2014), also reconstruct haplotypes with this approach. CliqueSNV, which was evaluated for HIV, constructs a diagram using linkage information between variants and identifies haplotypes by merging the cliques in the graph (Pelizzola et al., 2021). PredictHaplo, which was also specifically developed for identifying haplotypes in HIV, reconstructs haplotypes using a probabilistic cluster (Prabhakaran et al., 2014). Further details of these two programs wiil be discussed in Section 3.3.2.

However, if the variants are not sufficiently dense across the genome, for example, when there are regions in the genome with very few polymorphic sites, these haplotype reconstruction programs will struggle to connect the variants that span across these regions.

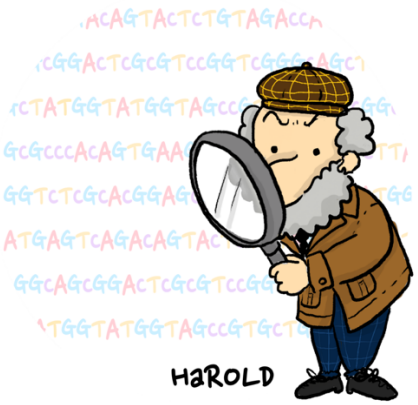
Motivation

In recent years, there has been an increased focus on monitoring within-host evolutionary dynamics by collecting and sequencing longitudinal clinical samples. These samples are collected at multiple time points across a set period of time ranging from a few weeks to years. With the longitudinal data available, we could make use of the co-variation of variant frequencies across multiple samples to provide additional information for the reconstruction process. We could reconstruct whole genome haplotype sequences with high confidence even when the polymorphic sites are far from one another in the genome.

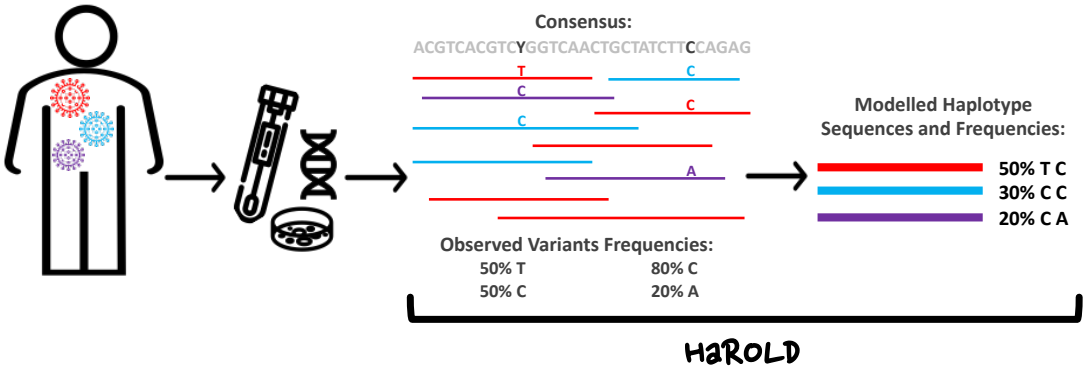
In this chapter, we present a high-performance haplotype reconstruction program, HaROLD. We evaluated HaROLD by comparing its performance with other currently available haplotype reconstruction programs, Clique SNV, PredictHaplo and EVORhA on synthetic NGS data of norovirus, a highly diverse RNA virus studied in this thesis, and human cytomegalovirus (HCMV), a large and slowly evolving DNA virus. We also illustrated the use of HaROLD on real clinical data from an immunocompromised patient infected with norovirus and five mother-infant pairs with vertically transmitted HCMV infections.

3.2 Materials and Methods

3.2.1 HaROLD



HaROLD reconstructs haplotypes from next generation sequencing raw reads data. When multiple strains of viruses or multiple haplotypes (e.g. the red, blue and purple viruses illustrated below) are present in a viral population, a mixture of reads of different haplotypes will be generated when viral whole genome sequencing is performed. Standard practice of consensus calling will be insufficient to separate out the three haplotypes in the viral population. Instead, HaROLD takes into consideration of all the variants and variant frequencies at every position to compute the actual haplotype whole genome sequence for each of the red, blue and purple virus as well as their respective frequencies in the viral population.



HaROLD takes advantage of longitudinal clinical data that contains samples from a patient at different time points, but it can also analyse a single sample. The program involves two steps, first, an initial estimation of haplotype sequences and frequencies, followed by refinement of haplotypes using observed read data and co-localisation information. It outputs the sequences of the haplotypes as the predicted probability of each base at each position, and the frequency of each haplotype in each sample.

Initial estimation

HaROLD assumes the samples share a common set of related haplotypes but allows the haplotypes to be in different proportions in different samples (i.e. some samples might also only contain one of the haplotypes). In this step, the program statistically models the observed base count data through 1) the frequencies of haplotypes in each sample and 2) the sequencing and mapping error rates that are sampled from a Dirichlet distribution. Since the haplotype sequence is unknown, at this step, each site in the alignment is considered separately and the information about the co-occurrence of variants along the reads is disregarded. This is done by summing over all possible bases in each variable position in each haplotype instead of searching over the space for all possible haplotype sequence. This independence approximation avoids the expensive exploration of the different combinations of possible haplotype sequences.

The haplotype frequencies and error rate parameters are optimised iteratively to maximise the likelihood of the read data. This involves the summation over the sequence of the haplotypes and an integration over the error rates, which results in a closed-form as it follows a Dirichlet

distribution. Once the optimal haplotype frequencies and error rate parameters are found, the posterior probability of each base at each site in each haplotype can be derived. This provides a probabilistic representation of the haplotypes.

This procedure is repeated over a range of different initial numbers of haplotypes. The optimal number of haplotypes with the maximum log-likelihood will be chosen.

Refinement process

The initial step assumes the samples share the same set of haplotype sequences ignoring the fact that mutations might arise in between two samples. Also, it does not consider the presence of multiple variants on the same read. The refinement process relaxes the above assumptions and use the information of co-localisation.

In the refinement process, the program considers each sample individually. This step takes in the output file from the initial step and starts with the preliminary estimated haplotype frequencies and the posterior probability of each base at each site in each haplotype.

With these two parameters, each read is to be assigned probabilistically to each of the haplotypes (the probability that a read would be from a particular haplotype). The total number of reads assigned to each haplotype is then used to update the haplotype frequencies. The read reassignment and haplotype frequencies update is then repeated until convergence.

Next, the re-assigned reads are used to update the posterior probability of each base at each site in each haplotype. Similar to the above, the update of read reassignment and base probabilities are repeated until convergence. The two updates, the update of haplotype frequencies and the update of base probabilities, are then repeated iteratively until convergence of the log-likelihood. If requested by the users, HaROLD can consider the recombination of two haplotypes and splitting or merging of haplotypes, which results in changes in the number of haplotypes. After each modification, the haplotype frequencies and base probabilities are readjusted as described above. The final output of haplotype frequencies and sequences with base probabilities at each variable site will then be generated.

Data Availability

Further detail of the statistical model can be found in (Pang et al., 2020b). The software HaROLD is deposited in the GitHub repository https://github.com/ucl-pathgenomics/HAROLD.

3.2.2 Preparation of synthetic data for method validation

To evaluate the ability of HaROLD to reconstruct haplotypes and estimate the relative haplotype frequencies, we created two synthetic sequence datasets using a mixture of whole genome sequences from GenBank (Clark et al., 2016). The first synthetic dataset consists of mixtures of two to four whole genome sequences of norovirus, an RNA virus with an approximate size of genome of 7.5 kb (Table 1). The second set consists of mixtures of two to three whole genome sequences of human cytomegalovirus (HCMV), a DNA virus with an approximate size of genome of 230 kb (Table 2). Norovirus was chosen because it is fast evolving, highly diverse, and mixed infections with multiple haplotypes are frequently observed in clinical settings. HCMV was chosen because it is slow evolving and has a large genome which has been proven challenging for other haplotype reconstruction programs. Nonetheless, HCMCV often presents with high within-host diversity due to multiple HCMV strains (Cudini et al., 2019). The opposite nature of these two viruses can provide a thorough evaluation of the capability of HaROLD.

SimSeq (Benidt and Nettleton, 2015) was used to create 1,000,000 paired end reads of length 250 for each GenBank norovirus reference sequence listed in Table 1, and 100,000 paired end reads for each GenBank CMV reference sequence listed in Table 2. The getErrorProfile module in SimSeq as used to generate an error model for the illumina sequencing simulator. The output SAM files from SimSeq were then converted into Fastq files using Picard version 2.21.1 ‘SamToFastq’ (Broad Institute, 2019a). In order to construct the data sets, Seqtk 1.3

(Shen et al., 2016) was used to mix the reads from each ensemble according to the relative fractions listed in Tables 1 and 2. Reads were then trimmed for adapters using Trim Galore version 0.6.0 (The Babraham Institute, 2019). Duplicate reads were removed using Picard version 2.21.1 'MarkDuplicates'. Reads were mapped to the GII.Pe-GII.4 Sydney 2012 reference strain JX459907 for norovirus, and the Merlin reference strain NC_006273.2 for CMV using BWA version 0.7.17 (Li and Durbin, 2009a). The Makereadcount.jar (https://github.com/ucl-pathgenomics/HaROLD/tree/master/jar) was used to obtain the strand specific nucleotide counts from BAM files. These strand count files were used as the input for HaROLD.

Table 1. Summary of the longitudinal norovirus synthetic data set used to test the accuracy of the haplotype reconstruction methods. *Four synthetic runs with five samples each were created for norovirus mixing GenBank sequences for a total of 20 samples.*

2 haplotypes Low similarity 5 time points

	Original samples	
Sample	**A: KJ19283**	**B: MH218631**
1	0%	100%
2	10%	90%
3	20%	80%
4	40%	60%
5	50%	50%

	A
B	98.6

2 haplotypes High similarity 5 time points

	Original samples	
Sample	**A: KC175323**	**B: KJ196279**
1	0%	100%
2	10%	90%
3	20%	80%
4	40%	60%
5	50%	50%

	A
B	99.7

3 haplotypes 5 time points

	Original samples		
Sample	**A: KC631827**	**B: KJ196283**	**C: MH218631**
1	20%	30%	50%
2	40%	30%	30%
3	60%	30%	10%
4	80%	20%	0%
5	70%	30%	0%

	A	**B**
B	99.3	
C	98.9	98.6

4 haplotypes 5 time points

	Original samples			
Sample	**A: KC176323**	**B: KJ196279**	**C: KJ196283**	**D: MH218631**
1	0%	0%	30%	70%
2	0%	20%	30%	50%
3	0%	30%	30%	40%
4	30%	20%	40%	10%
5	40%	20%	40%	0%

	A	**B**	**C**
B	99.7		
C	99.4	99.0	
D	99.3	98.9	98.6

Table 2. Summary of the longitudinal human cytomegalovirus synthetic data sets used to test the accuracy of the haplotype reconstruction methods. *Four synthetic runs with three samples each were created for HCMV mixing GenBank sequences for a total of 12 samples.*

<table>
<tr><th>Set</th><th>Sample composition</th><th>Similarity between haplotypes (Percentage identity)</th></tr>
<tr><td>2 haplotypes
Low similarity
3 time points</td><td>
<table>
<tr><th></th><th colspan="2">Original samples</th></tr>
<tr><th>Sample</th><th>A: KP745652.1</th><th>B: KP745644.1</th></tr>
<tr><td>1</td><td>0%</td><td>100%</td></tr>
<tr><td>2</td><td>10%</td><td>90%</td></tr>
<tr><td>3</td><td>30%</td><td>80%</td></tr>
</table>
</td><td>
<table>
<tr><th></th><th>A</th></tr>
<tr><td>B</td><td>99.1</td></tr>
</table>
</td></tr>
<tr><td>2 haplotypes
High similarity
3 time points</td><td>
<table>
<tr><th></th><th colspan="2">Original samples</th></tr>
<tr><th>Sample</th><th>A: KU221098.1</th><th>B: KT726952.2</th></tr>
<tr><td>1</td><td>0%</td><td>100%</td></tr>
<tr><td>2</td><td>20%</td><td>80%</td></tr>
<tr><td>3</td><td>40%</td><td>60%</td></tr>
</table>
</td><td>
<table>
<tr><th></th><th>A</th></tr>
<tr><td>B</td><td>99.4</td></tr>
</table>
</td></tr>
<tr><td>3 haplotypes
Low similarity
3 time points</td><td>
<table>
<tr><th></th><th colspan="3">Original samples</th></tr>
<tr><th>Sample</th><th>A: KP745652.1</th><th>B: KP745644.1</th><th>C: KP745670.1</th></tr>
<tr><td>1</td><td>20%</td><td>50%</td><td>30%</td></tr>
<tr><td>2</td><td>40%</td><td>40%</td><td>20%</td></tr>
<tr><td>3</td><td>60%</td><td>30%</td><td>10%</td></tr>
</table>
</td><td>
<table>
<tr><th></th><th>A</th><th>B</th></tr>
<tr><td>B</td><td>99.1</td><td></td></tr>
<tr><td>C</td><td>99.0</td><td>99.0</td></tr>
</table>
</td></tr>
<tr><td>3 haplotypes
High similarity
3 time points</td><td>
<table>
<tr><th>Sample</th><th>A: KU221098.1</th><th>B: KT726952.2</th><th>C: KJ361953.1</th></tr>
<tr><td>1</td><td>80%</td><td>20%</td><td>0%</td></tr>
<tr><td>2</td><td>60%</td><td>30%</td><td>10%</td></tr>
<tr><td>3</td><td>40%</td><td>40%</td><td>20%</td></tr>
</table>
</td><td>
<table>
<tr><th></th><th>A</th><th>B</th></tr>
<tr><td>B</td><td>99.4</td><td></td></tr>
<tr><td>C</td><td>99.4</td><td>99.4</td></tr>
</table>
</td></tr>
</table>

3.2.3 Evaluation of the performance of HaROLD

We evaluated the performance of HaROLD based on the accuracy of reconstructed whole genome sequences and the accuracy of the computed haplotype frequencies in each sample. The accuracy of the reconstructed sequences was calculated as the number of sites that are identical between the GenBank sequences we used to generate the synthetic data, and the reconstructed haplotypes, using the "dist.dna" function in R version 1.4.1106 package ape version 5.4.1. The "raw" model was used to calculate the pairwise distances between the sequences and the results representing the proportion of sites that differ between each pair of sequences were output as a matrix. The accuracy of frequencies was calculated as the difference between the actual haplotype frequency and the computed haplotype frequency in each sample, using the equation
1 – lactual haplotype frequency – computed haplotype frequencyl.

We compared the performance of HaROLD with 3 other commonly used haplotype reconstruction programs, namely, EVORhA (Pulido-Tamayo et al., 2015), CliqueSNV (Pelizzola et al., 2021) and PredictHaplo (Prabhakaran et al., 2014), using the same metrics. All programs were run on a High-Performance Computing cluster with a maximum memory of 50 GB and a run time of 48 hours. EVORhA was run with default parameters. CliqueSNV was run with the -tf option (the minimum threshold for frequency relative to the read's coverage) set to 0.01 (default was 0.05, decreasing the parameters increase the sensitivity of the program) and -cm option (cliques merging algorithm) set as "fast" (recommended by the developers for dataset with large number of SNPs). PredictHaplo was run with default parameters, except for the

entropy threshold, which was set to 0.05, the max gap fraction, set to 0.05, the local window size factor, set to 0.9, the MCMC interaction, set to 100 and deletions were not included.

All plots were generated using the R version 1.4.1106 package ggplot2 version 3.3.5 (R Core Team, 2019; Wickham, 2016).

3.2.4 Application of HaROLD to clinical data

We used HaROLD on a chronic infected norovirus patient data set previously reported in the case report by Ruis et al., 2018a. Further detail of this patient will be discussed in Chapter 4. Samples were sequenced as reported in Ruis et al., 2018a. Fastq files were mapped to the closest GenBank reference sequence FJ537136 using the same pipeline described in section 3.2.2 (Clark et al., 2016).

We have also applied HaROLD on a HCMV mother-infant congenital transmission data set (Pang et al., 2020a). Details of the sequencing method, sample processing and mapping have been described in Pang et al., 2020a.

HaROLD was run with the default settings for both data sets. The reconstructed haplotype sequences were aligned using Mafft version 7.490 (Katoh et al., 2002) and phylogenetic trees were built using RAxML version 8 with the GTRGAMMA model and 1000 bootstraps (Stamatakis, 2014a). Phylogenetic trees were plotted using R version 1.4.1106 package ggtree version 2.4.1 (Yu et al., 2017).

3.3 Results

3.3.1 Validation of HaROLD using norovirus and HCMV synthetic data

Norovirus

The norovirus synthetic data set consists of four independent runs, with a varying number of haplotypes (between one to four) and similarity between the input GenBank sequences. Each run consists of five longitudinal samples with different haplotype frequencies. The summary of this data set is presented in Table 1.

The reconstructed norovirus haplotypes were identical to the input GenBank sequences in every sample, in every run (100% accuracy) (sky blue in Figure 1A). The haplotype frequencies estimated by HaROLD were also highly accurate, with differences between the actual and estimated frequencies less than 0.002 across the whole data set (sky blue in Figure 1B).

HCMV

The HCMV synthetic data set was constructed in a similar format. It consists of four independent runs, with a varying number of haplotypes (between one to three). Three longitudinal samples with varying haplotype frequencies were included in each run. The summary of this data set is presented in Table 2.

The reconstructed HCMV haplotypes were highly similar to the original GenBank sequences, with a sequence accuracy larger than 0.997 (sky blue in Figure 2A). The differences between the actual and computed haplotype frequencies are less than 0.06 across the data set (sky blue in Figure 2B).

Evaluation

The computational time of HaROLD which depends on the number of haplotypes, average read depth and genome length, varied between 40 to 226 seconds for norovirus, and 35 to 39 minutes for HCMV (Table 3).

Although HaROLD was developed to take advantage of the availability of multiple longitudinal samples, very often, the number of clinical samples available is limited. Therefore, we also evaluated the performance of HaROLD on independent single samples (HaROLD-Single). HaROLD-Single gives reconstructed haplotype sequence accuracies between 0.99 to 1 for both the norovirus and HCMV synthetic data set. The accuracy of estimated frequencies varies between 0.93 to 1 and 0.78 to 0.99 for norovirus and HCMV respectively (dark blue in Figures 1 and 2).

The performance of HaROLD on single independent samples was generally not as good as when longitudinal data was available, which highlights the advantage of using serial samples. However, the results were still highly accurate, especially for the norovirus samples, where the genome is smaller, and when there were relatively few haplotypes.

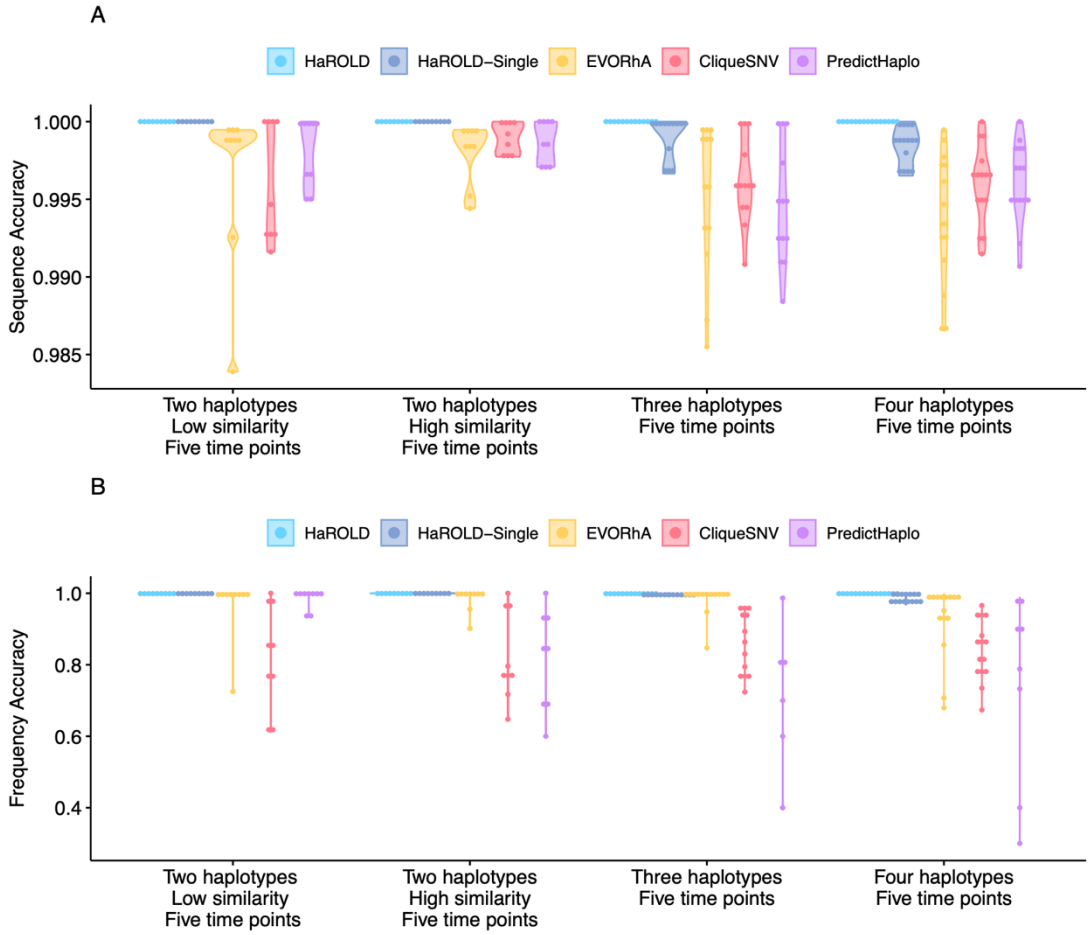


Figure 1. Violin plots showing the accuracy of haplotype reconstruction in the norovirus test set. *(A) The accuracy of reconstructed sequence (pairwise distance between the actual sequence and reconstructed sequence). (B) The accuracy of estimated frequencies. Colours indicate different haplotype reconstruction methods. Each dot represents a sequence from each sample (one to four sequences, for five samples, for four run).*

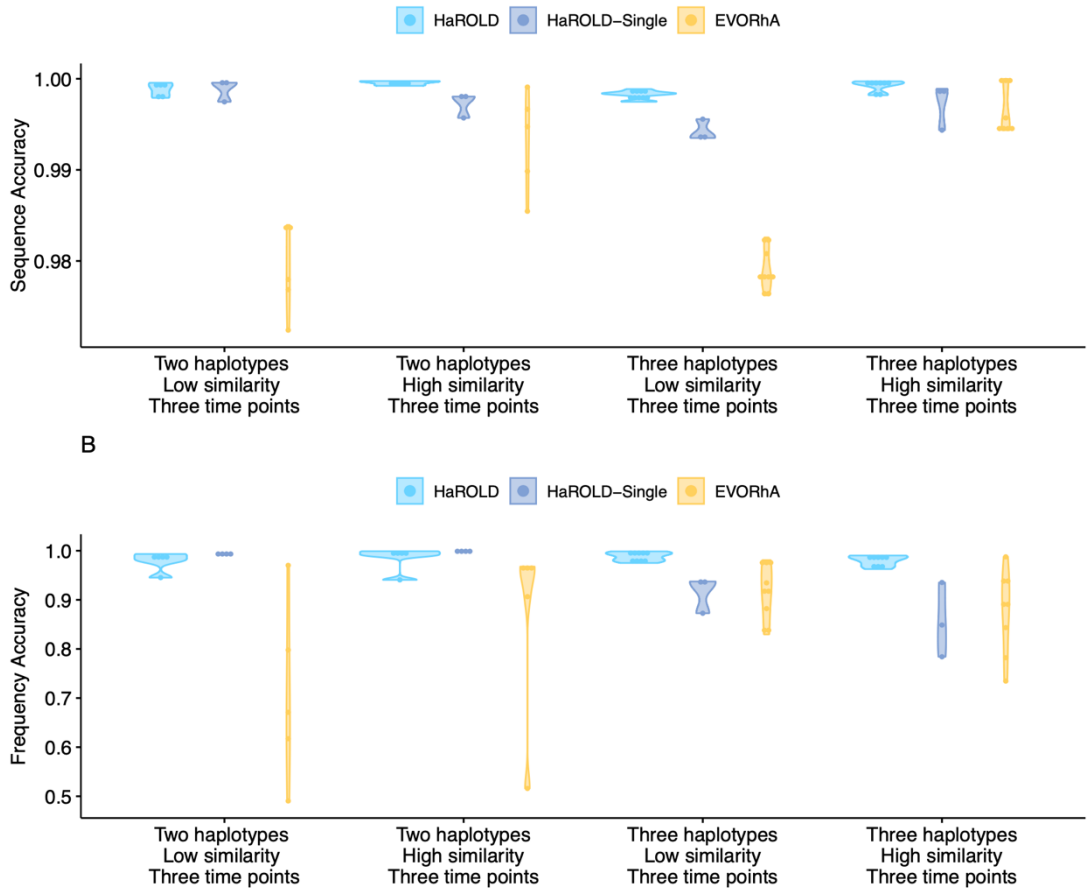


Figure 2. Violin plots showing the accuracy of haplotype reconstruction in the HCMV test set. *(A) The accuracy of reconstructed sequence (pairwise distance between the actual sequence and reconstructed sequence). (B) The accuracy of estimated frequencies. Colours indicate different haplotype reconstruction methods. Each dot represents a sequence from each sample (one to three sequences, for three samples, for four runs).*

Table 3. Computational time for HaROLD and other haplotype reconstruction programs for norovirus and HCMV synthetic data set.

Norovirus				
	2 haplotypes Low similarity 5 time points	**2 haplotypes High similarity 5 time points**	**3 haplotypes 5 time points**	**4 haplotypes 5 time points**
Harold	40 sec	1 min 4 sec	48 sec	3 min 46 sec
CliqueSNV	13 min 27 sec	20 min 24 sec	7 min 7 sec	13 min 43 sec
PredictHaplo	5 h 17 min	6 h 27 min	4 h 40 min	5 h 4 min
EVORhA	18 min	19 min	16 min	20 min
HCMV				
	2 haplotypes Low similarity 3 time points	**2 haplotypes High similarity 3 time points**	**3 haplotypes Low similarity 3 time points**	**3 haplotypes High similarity 3 time points**
Harold	39 min	36 min 28 sec	25 min	26 min
EVORhA	6 min	5 min	8 min	6 min

3.3.2 Comparison with other haplotype reconstruction methods

We compared the performance of HaROLD with two best performed haplotype reconstruction programs reviewed in (Eliseev et al., 2020b), namely CliqueSNV and PredictHaplo. In addition, we compared HaROLD to a third method, EVORhA.

Summary of the methods

CliqueSNV is a reference-based method to reconstruct haplotypes from next-generation sequencing short reads data, which constructs an allele graph based on linkage between variants and identifies true viral variants by merging cliques of that graph via combinatorial optimisation techniques (Knyazev et al., 2020). PredictHaplo implements a fully probabilistic approach to quasispecies reconstruction. Given a set of aligned reads, it uses a Bayesian mixture model with a Dirichlet process prior to estimating the unknown number of underlying haplotypes (Prabhakaran et al., 2014). The third method, EVORhA (Pulido-Tamayo et al., 2015), was developed for bacterial haplotype reconstruction. It combines phasing information in regions of overlapping reads with the estimated frequencies of inferred local haplotypes. This method was chosen because it is one of the very few haplotype reconstruction methods apart from HaROLD which also considers variant frequencies.

In both the Norovirus and HCMV synthetic data sets, EVORhA generally estimated a larger number of haplotypes than the actual number present in the sample (ranging from one to five additional haplotypes). It also consistently yielded haplotypes that most resembled the required input

reference sequence for mapping. The sequence accuracies ranged from 0.983 to 0.999 for Norovirus (Figure 1A, in yellow), and 0.972 to 0.999 for HCMV (Figure 2A, in yellow), which was consistently lower than HaROLD. The performance of EVORhA in estimating the relative haplotype frequencies was uneven and overall worse compared to HaROLD (Figure 1B and 2B).

Evaluation

We were not able to analyse the HCMV data set using CliqueSNV and PredictHaplo due to the memory and computational time constraints. Both programs were developed for small RNA viruses, such as HIV, which has a genome size of 9.2 kb, and were not capable to analyse an HCMV genome with the size of 230 kb using the available computational resources. In the norovirus data set, CliqueSNV yielded more accurate haplotype sequences compared to EVORhA, the frequency accuracy was, however, uneven (Figures 1A and B, in red). PredictHaplo performed similarly to CliqueSNV, with sequence accuracy ranging from 0.988 to 1 (Figure 1A, in purple). The frequency accuracy was again uneven, especially with the samples including four haplotypes (Figure 1B). HaROLD consistently outperformed these programs in both sequence and frequency accuracies, even when longitudinal information was not considered (HaROLD-Single, dark blue in Figure 1 and 2).

In terms of the computational time, HaROLD was more efficient than all other methods for the norovirus data set, although EVORhA was more efficient when analysing the HCMV data set where the average read depth was low (Table 3).

To gain a better understanding of the reconstructed haplotypes, we estimated the within-host diversity (heterozygosity) of various samples based on the reconstructed haplotype sequences by considering the frequency of each base at each position. The haplotypes generated by HaROLD produced highly accurate estimated within-host diversity, especially in the longitudinal runs, in both data sets. CliqueSNV and EVORhA underestimated the within-host diversity in all sample, except for one sample where only two highly similar norovirus haplotypes were present. PredictHaplo generally produced accurate within-host diversity, albeit the accuracy decreased when four haplotypes were present (Figure 3). All in all, HaROLD performed significantly better than other haplotype reconstruction programs.

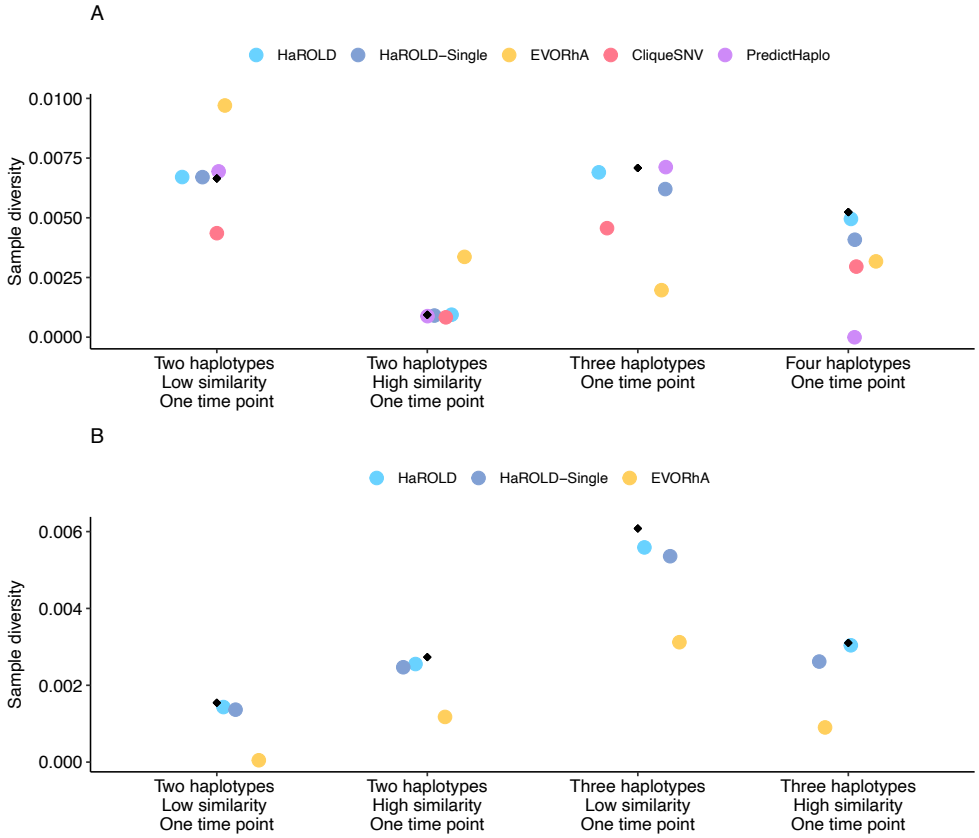


Figure 3. Accuracy of sample diversity estimations based on reconstructed haplotypes for norovirus test set (A) and HCMV test set (B). *True sequence diversity shown in black diamonds.*

3.3.3 Haplotype reconstruction in clinical norovirus data

In addition to synthetic data sets. We used HaROLD to analyse clinical sequencing data. First, we studied the norovirus samples collected from an immunocompromised 48-year-old patient with chronic norovirus infection (Ruis et al., 2018a). We analysed 12 longitudinal samples collected over a year. The patient was symptomatic and had received Favipiravir treatment. Phylogenetic analysis showed evidence for selective pressure in the within-host norovirus population. Further details of this patient will be discussed in Chapter 4, Section 4.3.3.

To better understand whether and how different viral populations evolved over time in response to the antiviral treatment, we reconstructed haplotypes from all samples using HaROLD. Each sample yielded two to five haplotypes which were then used to build a multiple sequences alignment together with the closest GenBank reference sequence (FJ537136). Analysis of pairwise genetic distances showed a clear bimodal distribution (Figure 4) and two main clusters were observed with multi-dimensional scaling (Figure 5). Sequences with more than 8% differences were considered members of different clusters as two phylogenetically distinct viral strain. The two clades were also present in the maximum-likelihood phylogenetic tree (Figure 6). The first viral population (in orange) was present since the first time point and was dominant in almost all samples (Figure 7). However, following treatment, a second viral population (in grey) appeared and became the dominant viral strain in time point 6 and 7, which correspond to the time when the patient received extensive treatment with Favipiravir.

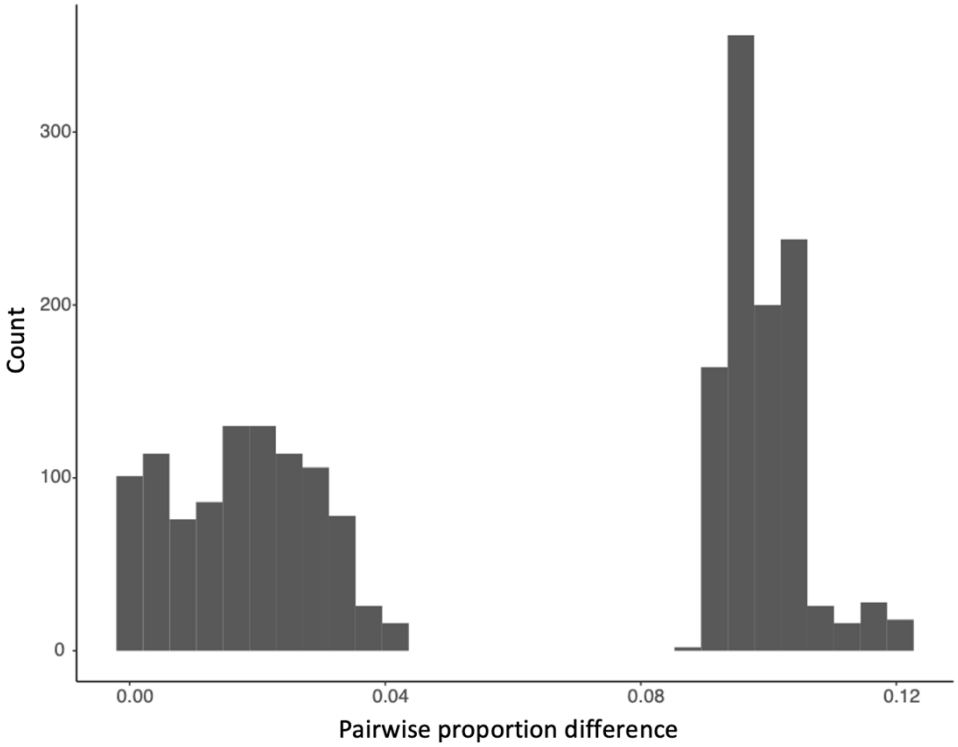


Figure 4. Pairwise genetic distances calculated as proportion. *Genetic distances were calculated for all reconstructed haplotypes obtained with HaROLD from 12 norovirus samples from an immunocompromised patient.*

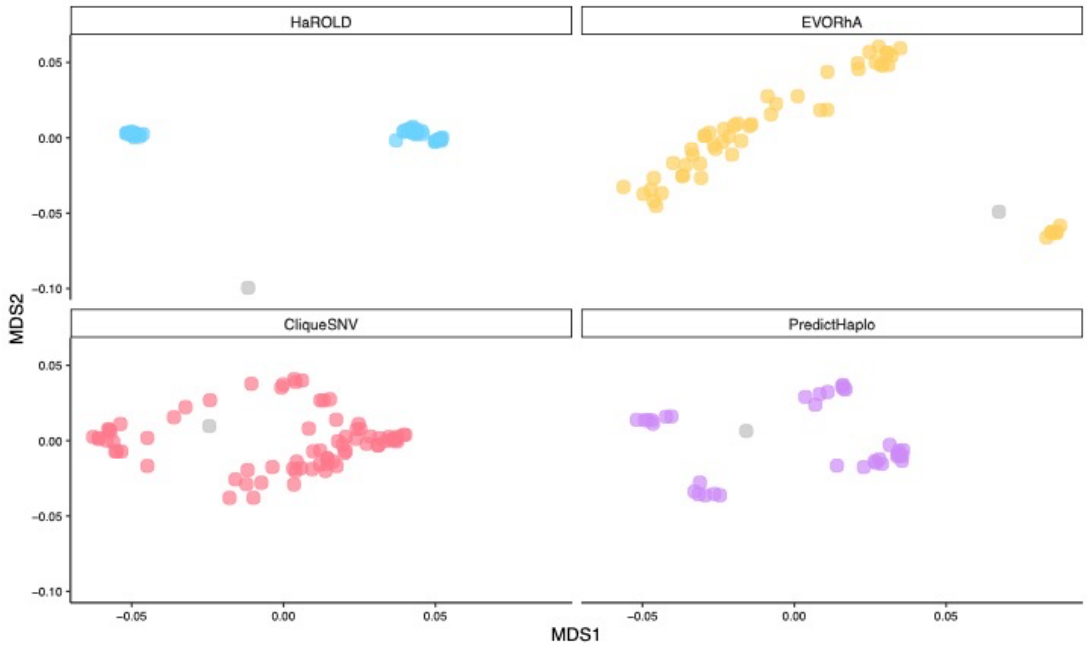


Figure 5. Multi-dimensional scaling (MDS) of reconstructed haplotypes from different methods. *Pairwise differences between haplotypes were calculated and used for MDS clustering. X-axis shows the first component obtained with MDS (MDS1) plotted against the second MDS component (y-axis, MDS2). Reference GenBank strain is coloured in grey.*

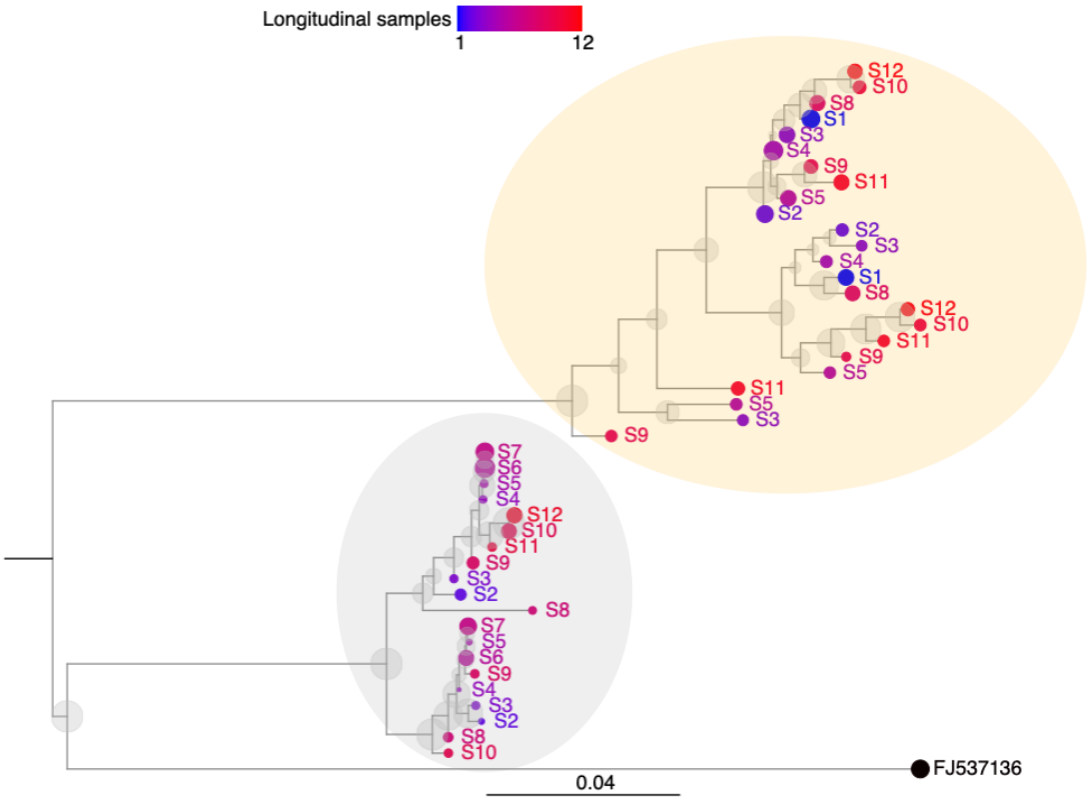


Figure 6. Maximum likelihood phylogenetic tree of HaROLD reconstructed haplotypes from a patient infected with norovirus. *Twelve samples were available for this patient (S1-S12) and were coloured differently using a continuous scale representing time (from blue S1 to red S12). The tips' size indicates the frequency of the haplotype. The black sequence is the Genbank strain used for mapping (tip size set as 50% frequency). Grey transparent circles represent the bootstrap values (1000 bootstraps in total). Two viral populations identified are represented in orange and grey transparent circles.*

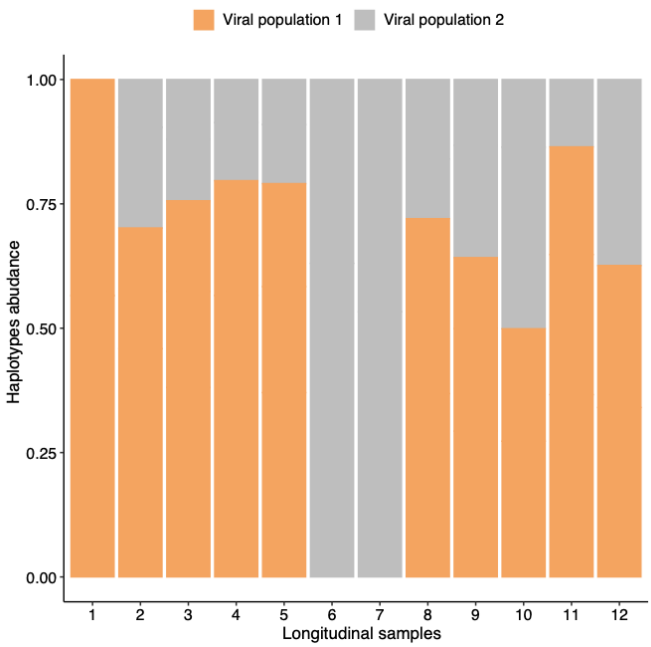


Figure 7. Bar plot of estimated abundances over time of HaROLD reconstructed haplotypes from a patient infected with norovirus. *The two viral populations identified are coloured in orange and grey (same colours as the maximum likelihood phylogenetic tree).*

We then compared these results obtained using other haplotype reconstruction methods. PredictHaplo gave similar results compared to HaROLD. It identified three to eight haplotypes for each sample which generally clustered within two main viral populations. However, they were not as distinct as the haplotypes identified by HaROLD. Each haplotype cluster was further divided into two sub-clusters (Figure 5, 8 and 9A). Even though PredictHaplo performed similarly to HaROLD, we encountered computational issues due to time and memory limits. It did not finish on five out of twelve samples when running on an HPC node with 50 GB memory and 14-day time limit. Both CliqueSNV and EVORhA yielded many haplotypes (EVORhA: two to ten; CliqueSNV: four to eight) that were low frequencies and tended to form diffuse clusters by sample (Figures 5, 8 and 9B to C), which did not give

information about the evolution of viral populations over time. In addition, when sequences from all methods were compared together, the EVORhA haplotypes were genetically very separated from the haplotypes identified by other methods (Figure 8).

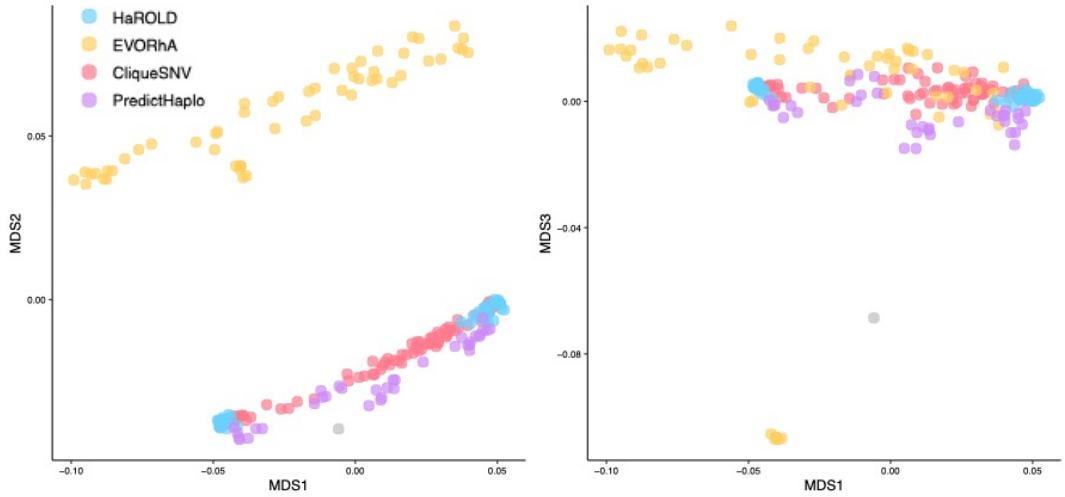


Figure 8. Multi-dimensional scaling (MDS) comparing all sequences obtained with different methods from a patient infected with norovirus. *All haplotypes were aligned to analyse the relationship between sequences retrieved by different methods. Pairwise differences between haplotypes were calculated and used for MDS clustering. Left plot shows the first component obtained with MDS (X-axis, MDS1) plotted against the second MDS component (y-axis, MDS2). Right plot shows the first component obtained with MDS (X-axis, MDS1) plotted against the third MDS component (y-axis, MDS3). Reference GenBank strain is coloured in grey.*

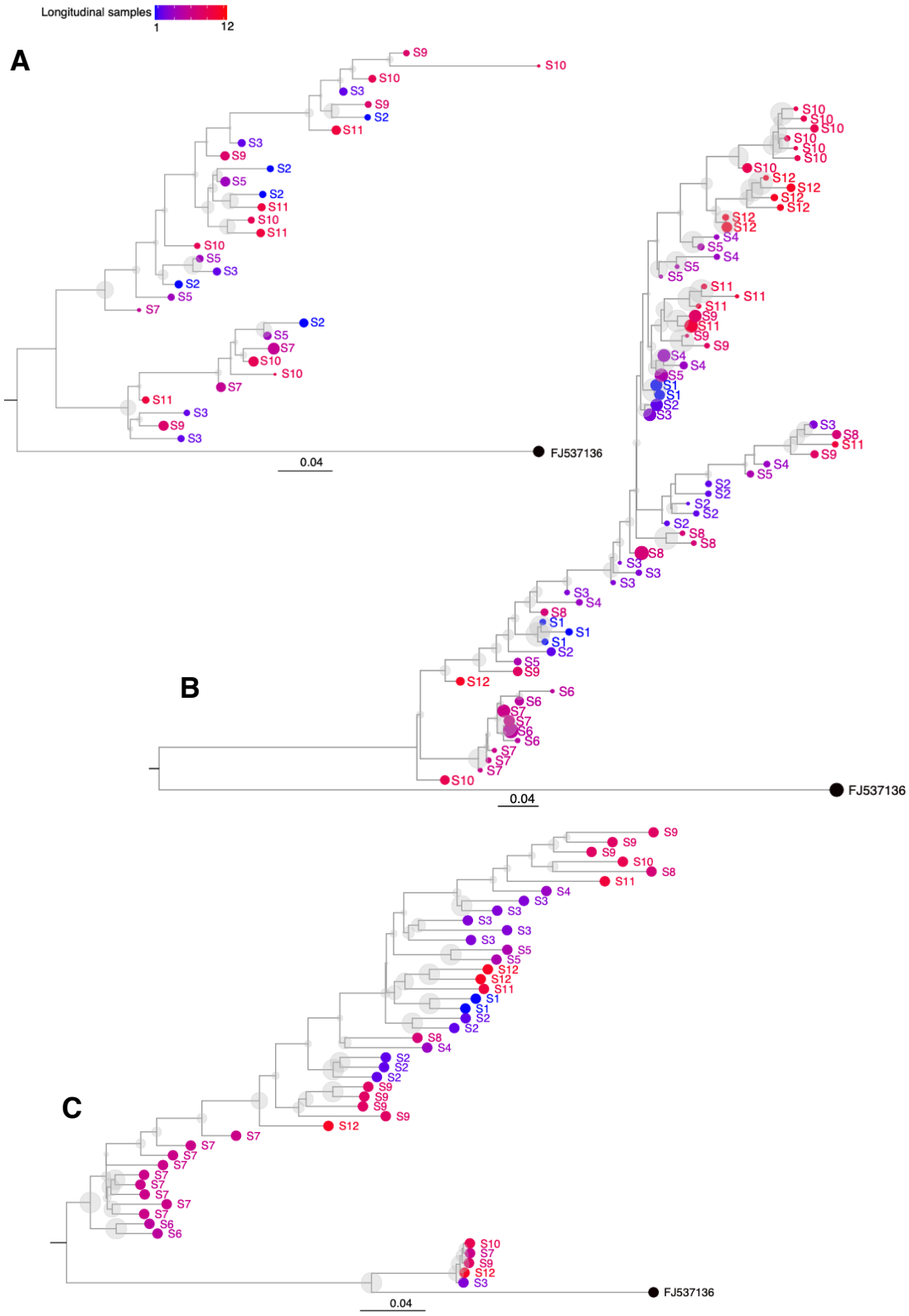


Figure 9. Maximum likelihood phylogenetic tree of (A) PredictHaplo, (B) CliqueSNV and (C) EVORhA reconstructed haplotypes from a patient infected with norovirus. *Twelve*

samples were available from this patient (S1-S12). The samples were coloured differently using a continuous scale representing time (from blue S1 to red S12). The size of the tips indicates the frequency of the haplotype. The black sequence is the Genbank strain used for mapping (tip size set as 50% frequency). Grey transparent circles represent the bootstrap values (1000 bootstraps in total). In panel C, since the haplotypes' frequency was always below 1%, the size of all tips was set to the same.

3.3.4 Haplotype reconstruction in clinical HCMV data

Apart from norovirus, we have also applied HaROLD on HCMV clinical data. Details of the patients can be found in Pang et al., 2020a. We analysed longitudinal HCMV samples from five HIV-infected Kenyan women and their infants (Richardson et al., 2016). By reconstructing whole genome haplotypes from these longitudinal HCMV samples, we identified the specific genotypes that were transmitted in congenital and postnatal infections, and the chronology with which specific HCMV variants were transmitted from mothers to infants.

Samples overview

We had longitudinal HCMV samples collected from mother's breast milk, mother's cervix, and baby's blood spot (Figure 10). Three infants (from family 12, 22, 123) acquired HCMV congenitally, and two (from family 14, 41) acquired HCMV postnatally. We observed high within-sample nucleotide diversity in almost all mother's breast milk samples, a metric which has previously been used as a proxy for the likelihood of mixed strain infections (Cudini et al., 2019). Therefore, we used HaROLD to resolve the individual haplotypes within each sample.

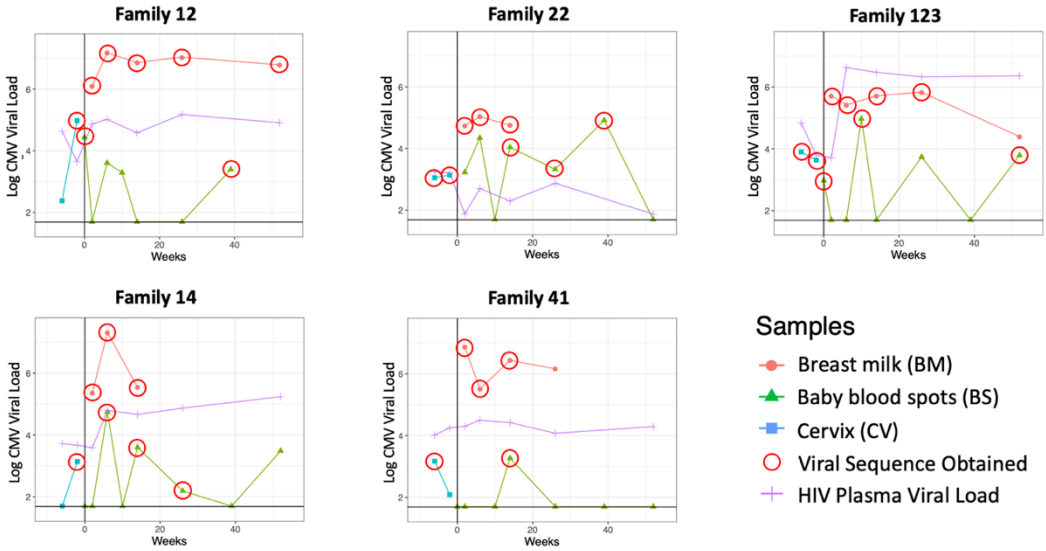


Figure 10. HCMV viral loads of longitudinal samples for each family from breast milk (red), baby blood spots (green), and cervix (blue), and HIV viral loads from mother's blood plasma. *Vertical line indicates date of delivery. Horizontal line indicates minimum threshold of detection. Red circles indicate the samples that were submitted for whole-genome sequencing.*

Haplotype reconstruction

Through HaROLD, we identified the haplotypes shared between the maternal and infant samples (Figure 11). Most breast milk samples (17 out of 18) contain a mixture of viral strains, whereas only one out of seven cervical samples has multiple haplotypes and none of the baby blood spot samples have more than one haplotype (Figure 11). This supports the previously published conclusions on compartmentalisation and transmission bottlenecks (Renzette et al., 2011).

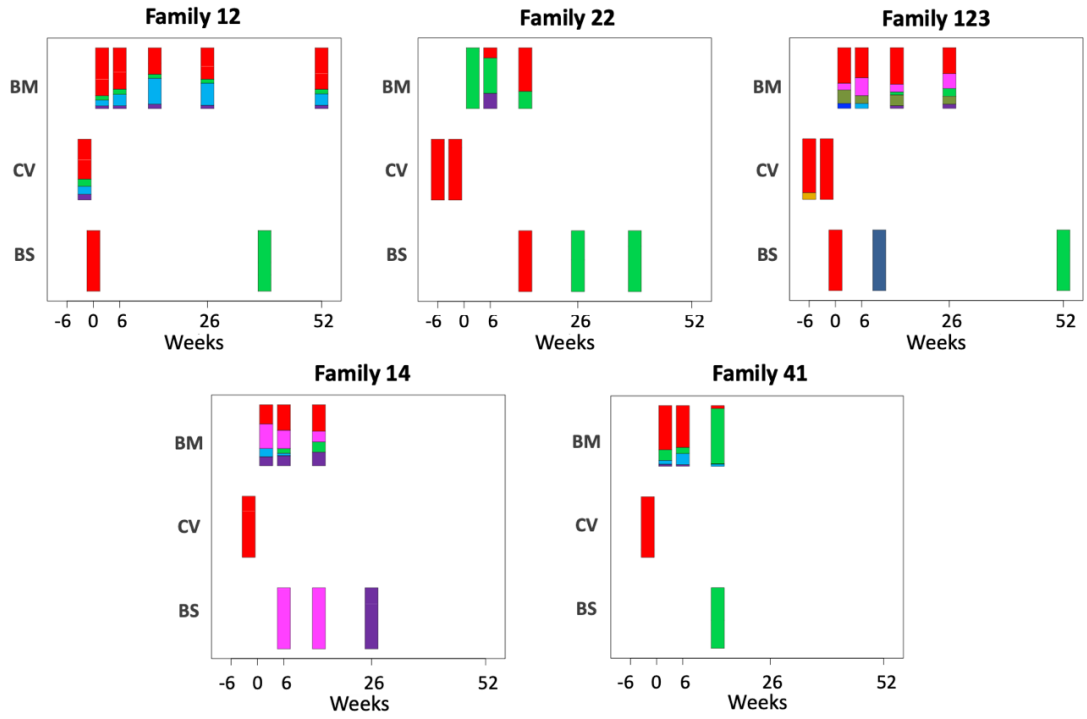


Figure 11. Abundance of haplotypes within each sample for breast milk (BM), cervix (CV), and blood spots (BS). *The timing of sampling is shown along the x-axis. For ease of reference, the genotype containing the most abundant haplotype present in the cervix is coloured red for each family. Thereafter sequences that are genetically closest to the red genotype are coloured magenta. Genotypes that are as distant from the cervical genotype as unrelated GenBank sequences are coloured shades of green, blue, and purple.*

We found that the three infants (from families 12, 22, 123) who acquired HCMV infection congenitally (tested positive at birth or within three weeks after birth, Figure 10) were first infected with the genotype present in the greatest abundance in the cervix and were subsequently re-infected with distinct genotypes present in the mother's breast milk (Figure 11). The two infants (from families 14 and 41) who acquired HCMV through post-natal transmission (first tested positive at six and ten weeks respectively) were infected by the haplotypes detected in the

mother's breast milk, which were genetically different from the haplotype present in the mother's cervix.

Given the observation of possible compartmentalisation and the similarity between the strain in congenital HCMV to the cervix strain, we searched for shared genotyping features which might confer a fitness advantage to establish the initial infection. We compared the haplotypes found in different compartments and looked for common characteristics of haplotypes observed in the same compartment using a fixation index (F_{ST}) analysis. We identified 19 genes that were likely to be contributing to the genetic similarity between congenitally transmitted genotypes from mothers 12, 22, and 123 (Pang et al., 2020a). Most of the identified genes were known to associate with tissue tropism (Stanton et al., 2010) or immunomodulation (Bruno et al., 2016; Cortese et al., 2012; Gabaev et al., 2014; Pérez-Carmona et al., 2018; Stanton et al., 2010; Van Damme and Van Loock, 2014)

Performance in poor quality data

Since some of the samples in this data set had a relatively poor sequencing quality (a low average read depth), as part of the quality check for this analysis, we subsampled the reads of a high-quality sample with an average read depth of 780, down to an average read depth of four. We found that HaROLD can identify all five haplotypes present in the original sample even when we subsampled the reads down to an average read depth of 22. At an average read depth of four, HaROLD still identified all three majority haplotypes (Figure 12).

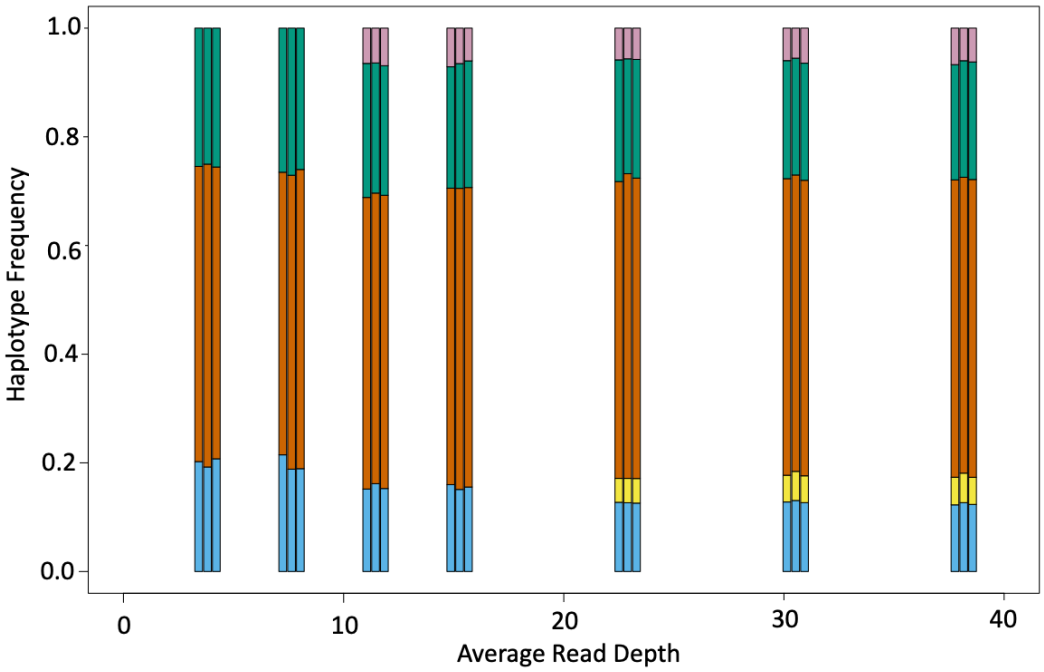


Figure 12. Boxplot showing the number and frequencies of haplotypes reconstructed by HaROLD in relation to read depth. *Analysis was performed on a mother's breast milk sample from family 12, which has an average read depth of 780. Reads were subsampled down to an average read depth of 40 to four. At an average read depth of 22 to 40, all five haplotypes were still present. At an average read depth of 15, the lowest frequency haplotypes were not detected. At an average read depth of four, only the three major haplotypes were present.*

3.4 Closing Remarks

Next generation sequencing (NGS) can detect single nucleotide polymorphisms (SNPs) in a viral population. To understand the within-host viral population structure and provide clinical insights, it is essential to identify which SNPs belong to the same haplotype, through a process called haplotype reconstruction.

Haplotype reconstruction is challenging, and it requires a robust statistical framework. Most programs rely on reads that contain multiple polymorphic sites to reconstruct haplotypes (Knyazev et al., 2018; Prabhakaran et al., 2014). Unfortunately, this is not always the case. For some viruses, such as HCMV, the genetic diversity observed is usually confined to short intervals.

Why is HaROLD better?

Instead, HaROLD considers each site in the alignment independently. It calculates the posterior probability that each possible base occurs at each site in each haplotype and then assigns the base to a haplotype if the probability is sufficiently high. In addition, it can make use of longitudinal clinical samples, which are likely to share closely related haplotypes. The additional information on variant frequencies can improve the accuracy of the haplotypes reconstructed. Based on our evaluation, HaROLD generates highly accurate haplotypes and associated frequencies, with or without the presence of longitudinal samples, and even in samples with a very low average read depth. Since HaROLD is much less computationally demanding compared to other haplotype reconstruction programs which are mainly optimised for small

RNA viruses like HIV, it can analyse sequence data from viruses with a larger genome such as HCMV or samples with a greater read depth.

Limitations of HaROLD

The major limitation of HaROLD is the ability to analyse data from different sequencing technologies. While some programs can analyse data from multiple sequencing platforms including 454/Roche, PacBio, amplicon and illumina sequencing (Astrovskaya et al., 2011; Knyazev et al., 2018; Zagordi et al., 2011), HaROLD can only be used to analyse illumina sequencing data. During the development process, HaROLD was tested on amplicon and nanopore sequencing data, however, results were inconsistent. HaROLD relies heavily on robust variant data and consistent sequencing quality across the genome.

Application of HaROLD

In this chapter, we demonstrated how HaROLD can be used to identify unique haplotypes within a mixed infection. In the norovirus clinical case, HaROLD identified two distinct viral populations. Based on the haplotype frequencies estimated by HaROLD, we noticed the composition of the viral population had completely changed shortly after the start of the antiviral treatment. The minority haplotype in the samples prior to treatment became dominant, which suggests the haplotype might be resistant to the drug.

In the HCMV clinical data set, HaROLD provides further evidence of compartmentalisation of viral strains between breast milk and cervix. By reconstructing the haplotypes, we identified the viral strains which were transmitted from the mothers to their infants, and subsequently, the

genes that might contain determinant of congenital HCMV transmission. HaROLD uncovered detail which would have been missed by standard analysis of consensus sequences and basic variants calling.
In an acute infection of a relatively slow evolving RNA virus such as SARS-CoV-2, the infection might be cleared before the viral strain is evolved into multiple haplotypes with a large set of distinct variants. In such case, HaROLD can still identify stable clones of distinct minority variant genomes within the viral subpopulations in a heterogenous infection. Further details will be discussed in chapter 5.

In a further clinical analysis done by Dr Cristina Venturini (unpublished), HaROLD has identified a sample contamination from a lab strain with high accuracy. HaROLD was used to look for the presence of multiple haplotypes within two samples with high nucleotide diversity. In both samples, two haplotypes were identified by HaROLD, and the minority haplotypes were nearly identical to the Merlin laboratory strain which was present in the sequencing lab. Resequencing of these two samples revealed the ‘haplotypes’ identified by HaROLD in the initial sequencing samples were indeed sample contaminants. This provided a real-world scenario for validation of HaROLD.

In summary, we illustrated how HaROLD, a high accuracy haplotype reconstruction program can be useful for understanding mixed-infections and within-host evolutionary dynamics. In the next two chapters, we will investigate the within-host evolution of norovirus and SARS-CoV-2 using HaROLD along with other bioinformatics pipelines and statistical models.

Chapter 4

Within-Host Variations and the Impact of Antiviral Drugs

4.1 Introduction

Overview

Before the COVID-19 pandemic, there were few approved antiviral drugs available for treating RNA viral infections other than for influenza and blood-borne viruses. This includes infections that can become life-threatening in patients with underlying immunodeficiency. Several repurposed drugs have been proven effective in certain cases; however, it is not clear why only some patients appear to respond to treatment. Since these viral infections usually occur sporadically, unless there were sufficient infected individuals, such as during the COVID-19 pandemic, it is difficult and expensive to conduct clinical trials to investigate the efficacy of such antiviral drugs.

To address this disparity in response to treatment, we must first understand the within-host variations in these infections. In the first half of this chapter, we analyse deep sequenced longitudinal norovirus samples collected from two untreated norovirus patients. We then characterised the clinical and virological response in three immunocompromised norovirus patients treated with favipiravir by analysing the longitudinal deep sequenced samples. We used HaROLD

to reconstruct the haplotypes and we identified the putative drug-associated amino acid changes and analysed the drug-induced mutational signature. In the second half of the chapter, we describe similar methods to study the deep sequenced longitudinal SARS-CoV-2 samples collected from untreated patients and three SARS-CoV-2 patients treated with remdesivir, as well as two patients who received dual therapy of remdesivir and nitazoxanide. To improve our understanding of the *in vivo* antiviral activities, we evaluated the effect of two RdRp inhibitors, favipiravir and molnupiravir, in SARS-CoV-2 infected Syrian hamsters. This served to further our understanding of the mechanism of action of RdRp inhibitors in treatment of the RNA viral infections. The results provided valuable insights to inform the debate surrounding the widespread use of this class of antiviral drugs for treatment of RNA viral infections, including those for SARS-CoV-2 during the ongoing COVID-19 pandemic.

Norovirus

At the time of writing, no antiviral treatment has been licensed for treating human norovirus infections. In immunocompetent individuals, no treatment is usually needed. However, patients with underlying immunodeficiency can develop chronic infections which require treatments. Various types of treatments have been used. As discussed in Chapter 1, Section 1.2.8, adaptive and innate immune mechanisms are both essential for norovirus viral clearance. Therefore, intravenous immunoglobulin (IVIG) is frequently used in immunocompromised individuals with chronic norovirus infection to support viral clearance (Brown et al., 2017). In addition, nitazoxanide and RdRp inhibitors are often used as combination therapies in compassionate access programmes. Since norovirus infections in normal individuals are self-

limiting and require no intervention, the majority of patients we analysed in this chapter had underlying immunodeficiency and chronic norovirus infections. They received treatments via compassionate programmes in Great Ormond Street Hospital or Royal Free Hospital.

Nitazoxanide

Nitazoxanide was initially developed as an antiparasitic drug. Following administration, it is absorbed in the gastrointestinal tract and hydrolysed to form the active metabolite tizoxanide (Stockis et al., 2002). It has been licensed to treat *Cryptosporidium parvum*- and *Giardia intestinalis*-induced diarrhoea in both adults and children (Rossignol, 2014). It has demonstrated a broad spectrum antiviral activity against influenza, hepatitis B virus, hepatitis C virus, HIV, norovirus and rotavirus (La Frazia et al., 2013; Rossignol, 2014). It has also been reported to exhibit significant synergy with remdesivir for treating SARS-CoV-2 (Bobrowski et al., 2020).

In norovirus, nitazoxanide inhibits viral replication by stimulating the overexpression of a subset of interferon stimulated genes such as interferon regulatory factor 1 (IRF-1) and activating the cellular antiviral response (Dang et al., 2018). It has been shown to synergise with ribavirin, an RdRp inhibitor, to inhibit human norovirus infection *in vitro* (Dang et al., 2018). However, clinical experiences on the efficacy of nitazoxanide are mixed. Some studies reported that patients showed clinical improvement following the use of nitazoxanide for treating post haematopoietic stem cell transplantation norovirus gastroenteritis (Dang et al., 2018; Gorgeis et al., 2017; Morris and Morris, 2015). In a study with a cohort of 10 immunocompromised patients, some showed clinical improvement or viral clearance (although subsequently relapsed), while

others symptomatically deteriorated during treatment (Brown et al., 2019). Another study demonstrated the lack of efficacy of nitazoxanide in treating an immunocompromised patient with chronic norovirus infection (Kempf et al., 2017).

RdRp Inhibitors

RdRp inhibitors we discussed in Chapter 1, Section 1.1.5, such as favipiravir, remdesivir, ribavirin and molnupiravir have been frequently used to treat human norovirus infections as a monotherapy or in combination with nitazoxanide (Arias et al., 2014; Day et al., 2005; Ortega-Prieto et al., 2013; Rocha-Pereira et al., 2012; Ruis et al., 2018b; van Kampen et al., 2022; Vignuzzi et al., 2005; Woodward et al., 2012).

As discussed, they can inhibit the action of the RdRp by acting as nucleoside analogues which are incorporated into the viral RNA strand by the RdRp during replication (Furuta et al., 2017), or more frequently, drive mutations occurring in the complementary RNA strand, which leads to lethal mutagenesis (Agostini et al., 2019; Bull et al., 2007; Goldhill et al., 2019; Shannon et al., 2020; Sheahan et al., 2020; Urakova et al., 2018; Vignuzzi et al., 2005; Yoon et al., 2018). They are known to have broad antiviral activities against RNA viruses *in vitro* and *in vivo*, but the experience of their use for clinical treatment is mixed. Previous reports showed contradictory outcomes (Arias et al., 2014; Rocha-Pereira et al., 2016, 2012; Ruis et al., 2018b) and the optimal dosage and timing for using these antivirals remain unclear. Despite increased clinical benefits in some cases, no fall in viral load was observed (Ruis et al., 2018b). It has been suggested the mutagenic signature induced by RdRp inhibitors can be used as a measure of clinical efficacy (Illingworth et al., 2020; Ruis et al., 2018b).

SARS-CoV-2

For SARS-CoV-2, hundreds of repurposed antiviral drugs have been tested for treating COVID-19. As of May 2022, there are almost 700 on-going drug development programs (Center for Drug Evaluation and Research, 2022). As discussed in Chapter 1, Section 1.3.9, three antiviral drugs, Paxlovid (Nirmatrelvir/ ritonavir), Lagevrio (molnupiravir) and remdesivir, and four antibody therapies, ronapreve, sotrovimab (Xevudy), Evusheld and tocilizumab/ sarilumab, have been approved for treating COVID-19 in the UK (UK Parliament, 2022). Currently, paxlovid, a protease inhibitor, is one of the antiviral drugs proven most effective (Hammond et al., 2022).

However, at the beginning of the pandemic, there were no approved antiviral drugs for treating COVID-19. When this study was conducted, repurposed drugs were the only treatments available, and they are still being used in many countries for treating COVID-19.

Aims

To gain a better understanding of the infections, and explain the disparity in patient response to treatments, we aimed to link the viral genetic signals with the clinical effect of the antiviral drugs. In this chapter, we present the within-host variants and drug effect analysis of the following data sets:

Virus	Section	No. of Patients	Treatment	Detail
Norovirus	4.3.1	2	Untreated	analysed the viral dynamics of natural infection and characterised the changes in consensus level variants as compared to the first sample collected from the patients.
	4.3.2	11	Ribavirin or Nitazoxanide	analysed by previous members of the lab
	4.3.3 & 4	3	Favipiravir	analysed stool samples collected at multiple time points to monitor the drug-induced mutagenesis and the emergence of drug-associated mutations
	4.3.5	Zebrafish larvae	Favipiravir	inoculated with stool samples from one of the favipiravir treated patients in section 4.3.3. Our collaborators investigated the in vivo antiviral activity of favipiravir.
SARS-CoV-2	4.3.7	9	Remesivir (3) or Untreated (6)	performed viral deep sequencing, mutational analysis, and evolutionary modelling
	4.3.8	2	Remdesivir and Nitazoxanide	using methods described in section 4.3.7 to study the efficacy of the combination therapy
	4.3.9	Syrian hamsters	Favipiravir or Molnupiravir	evaluated the in vivo efficacy and mutagenic effect of the two RdRp inhibitors

4.2 Materials and Methods

4.2.1 Sample collection, data assembly and sequence processing

Norovirus

For norovirus, stool samples were collected by the clinical teams at Great Ormond Street Hospital, Newcastle Hospitals Trust or Royal Free Hospital. Full-length Norovirus genome sequences were obtained from PCR positive samples using SureSelect^{XT} target enrichment and Illumina sequencing done by the UCL Pathogen Genomics Unit. All sequencing reads were trimmed using Trim Galore version 0.6.0 (The Babraham Institute, 2019) and the sequencing quality was checked using FastQC version 0.11.9 (Andrews, 2020). Duplicated reads were identified and removed using Picard version 2.21.1 'MarkDuplicates' (Broad Institute, 2019b). For each patient, the first sample was genotyped by mapping the reads to a panel of norovirus reference sequences from Genbank using bwa-mem version 0.7.17 (Clark et al., 2016; Li and Durbin, 2009b). The genotypes identified were further validated using the Genome Detective Norovirus Typing Tool version 1.9 (Kroneman et al., 2011). A unique patient reference was generated by remapping the reads of the first sample to the identified genotype reference sequence from GenBank (Clark et al., 2016). Reads from the subsequent samples of the same patient were mapped to this patient reference. The mapping quality was checked using Qualimap version 2.2.1 (Okonechnikov et al., 2015). BAM files were processed using samtools version 1.9 'mpileup' (Danecek et al., 2021). Consensus

sequences were called using QUASR version 6.08 (Watson et al., 2013) and aligned using MAFFT version 7 (Katoh and Standley, 2013). Only genomes with more than 80 percent genome coverage and a mean read depth of 100 or above were included in downstream analysis. Minority allele variants with a frequency of above 2% and a minimum of 2 supporting reads were identified at sites with a read depth of ≥ 5 using VarScan (Koboldt et al., 2012). Maximum likelihood phylogenies of the major capsid (VP1) alignments, including 850 Genbank reference sequences of different human norovirus genotypes, were constructed using RAxML (Stamatakis, 2014b), with the GTR model and 1000 bootstrap replicates.

SARS-CoV-2

For SARS-CoV-2, nasopharyngeal swab samples were collected by clinical teams at Great Ormond Street Hospital. All samples were tested for SARS-CoV-2. Whole genome SARS-CoV-2 sequences were obtained from all polymerase chain reaction (PCR) positive samples using SureSelect^{XT} target enrichment and Illumina sequencing done by the UCL Pathogen Genomics Unit. The protocol used for sequence processing is the same as described above. In summary, for each patient, a unique patient reference was obtained by mapping all reads to the SARS-Cov-2 GenBank reference sequence NC_045512 Wuhan-Hu-1 using bwa-mem (Clark et al., 2016; Li and Durbin, 2009b). Reads from the subsequent samples of the same patients were mapped to this patient reference sequence. Variants were called at sites with a minimum read depth of five, a frequency of 2% and a minimum of four supporting reads. Transient variants which are only observed in a single time point were discarded from downstream analysis. Consensus sequences were extracted and aligned as described. Only genomes with

more than 80 percent genome coverage and a mean read depth of 100 or above were included in downstream analysis. Maximum likelihood phylogenetic trees of the consensus sequence alignment and haplotype sequence alignment were constructed using RAxML (Stamatakis, 2014b), with the GTR model and 1000 bootstrap replicates. Both phylogenies were rooted on the reference sequence NC_045512 Wuhan-Hu-1. Periscope was used to detect subgenomic RNA (sgRNA) in all samples (Parker et al., 2020). sgRNAs are identified based on the detection of the leader sequence at the 5'end (5'-AACCAACTTTCGATCTCTTGTAGATCTGTTCT-3') of the sequence. To optimise the reliability of the analysis, we excluded genomes with less than 90% coverage and lower than 100 mean read depth.

Sequence analysis

For all patients, haplotypes were reconstructed using HAplotype Reconstruction Of Longitudinal Deep Sequences (HaROLD) with default settings (Pang et al., 2020b). Downstream analysis was done in R 3.6.1 using Rstudio 1.2 unless otherwise stated (R Core Team, 2019). In general, data was processed using the tidyverse family of packages version 1.2.1 (Wickham et al., 2019). Pairwise distances were calculated using the dist.dna function with raw model in ape package version 5.6 (Paradis et al., 2004). To compare the number of mutations in treated and untreated samples across all patients, Fisher's exact test was done using the fisher.test function in R stats package version 3.6.2. For correlation analysis, Pearson's and Spearman's rank correlation tests were done using the lm.test and cor.test function in the R stats package version 3.6.2 (R Core Team, 2019). Figures were made using ggplot2 version 3.3.5 (Wickham, 2016).

4.2.2 Mutagenesis analysis

As discussed, Favipiravir, the major drug used to treat the norovirus patients P3, P4 and P5, is a nucleoside analogue which is known to induce lethal mutagenesis (Baranovich et al., 2013; de Ávila et al., 2016; Goldhill et al., 2019; Ruis et al., 2018b). To evaluate the drug effect, we looked for mutagenic signatures by comparing the variants in each subsequent sample to the variant found in the first time point sample, at each position. For this analysis, all variants with a minimum of two supporting reads and 2% of variant frequency were included. We counted the number of polymorphic sites for each type of transition and transversion across the genome. To remove bias of uneven sequencing depth across the genome and across different samples, these counts were adjusted using the correction factor of Watterson Estimator (Ferretti and Ramos-Onsins, 2015; Watterson, 1975). Each polymorphic site increases the count by 1/a, where a is the correction factor:

$$a = \sum_{i=1}^{n-1} \frac{1}{i},$$

and n is the read depth at the site.

To further investigate the drug effect of RdRp inhibitors, we evaluated the mutagenic signature of favipiravir and another RdRp inhibitor, molnupiravir in SARS-CoV-2 infected hamsters by comparing the variants in SARS-CoV-2 viral sequences collected 4 to 5 days post infection with the virus used for inoculation, using the same metrics and correction factor.

4.2.3 Structural biology

To understand the effect of the norovirus variants selected by favipiravir, we placed the mutations onto the structure of the norovirus RNA dependent RNA polymerase (RdRp) protein PDB 4LQ3 and visualised it using ChimeraX 1.2.5 (Goddard et al., 2018; Pettersen et al., 2021). To find out whether the identified variants are present at highly variable or conserved sites, we analysed these polymorphic sites in a data set of 1000 sequences from Genbank (Clark et al., 2016) and our own database created by Dr Christopher Ruis, Dr Sunando Roy and Dr Florencia Tettamanti Boshier as part of another project in the lab. We calculated the percentage of sequences carrying each variant of interest. For the rare mutation from Alanine to Serine at the position 44 in the RdRp, a position which is otherwise highly conserved across all noroviruses and other RNA viruses, we modelled the substitution, fixed the rotamer and visualised the new bond formed using ChimeraX 1.2.5 (Goddard et al., 2018; Pettersen et al., 2021).

4.3 Results

4.3.1 Untreated norovirus patients

To understand the viral variations in natural infections, we first examined the CT value trajectory and longitudinal whole genome sequences from norovirus patients who had not been treated with any antiviral drugs.

Patients 1 and 2

P1 and P2 were paediatric patients diagnosed with Wiskott-Aldrich syndrome (a X-linked recessive disease characterised by a low platelet count, eczema and immunodeficiency) and primary immunodeficiency disorder respectively. Six and seven samples were collected and sequenced from P1 and P2 (Figure 1). Genotyping revealed P1 was infected by the GII.P7 GII.6 norovirus strain and P2 was infected by GII.P2 GII.2 norovirus strain. Neither patient showed a significant change in CT value over time ($p > 0.05$ in t-test) (Figure 1A and B). In both patients, when compared to the first sample, fewer than 10 consensus level changes were found across the genome. In particular, no consensus changes were found in the VPg, 3CLpro or RdRp region in either patient (Figure 2).

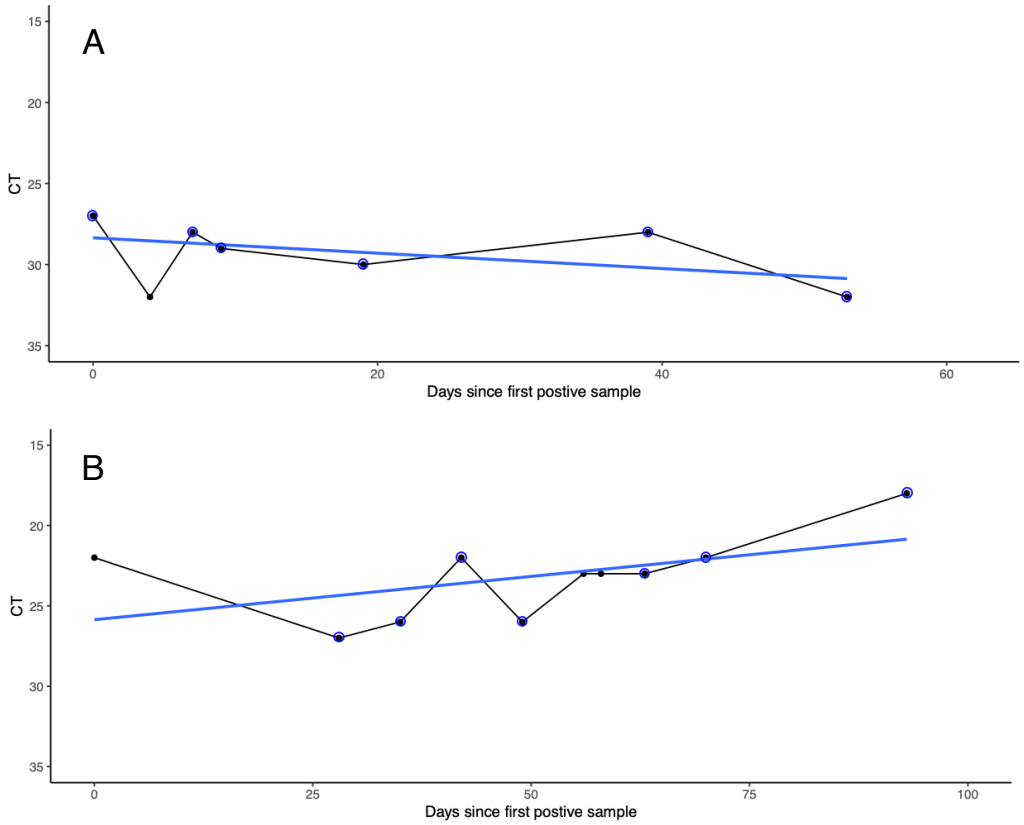


Figure 1. CT values of samples collected from P1 (A) and P2 (B). *Y-axis = inverted CT value, X-axis = days since the first positive sample collected at Great Ormond Street Hospital. Blue circles indicate samples which have been sequenced. The lines of best fit are shown in blue. The changes in CT values over time in both patients were not significant (p values > 0.05 in t-test).*

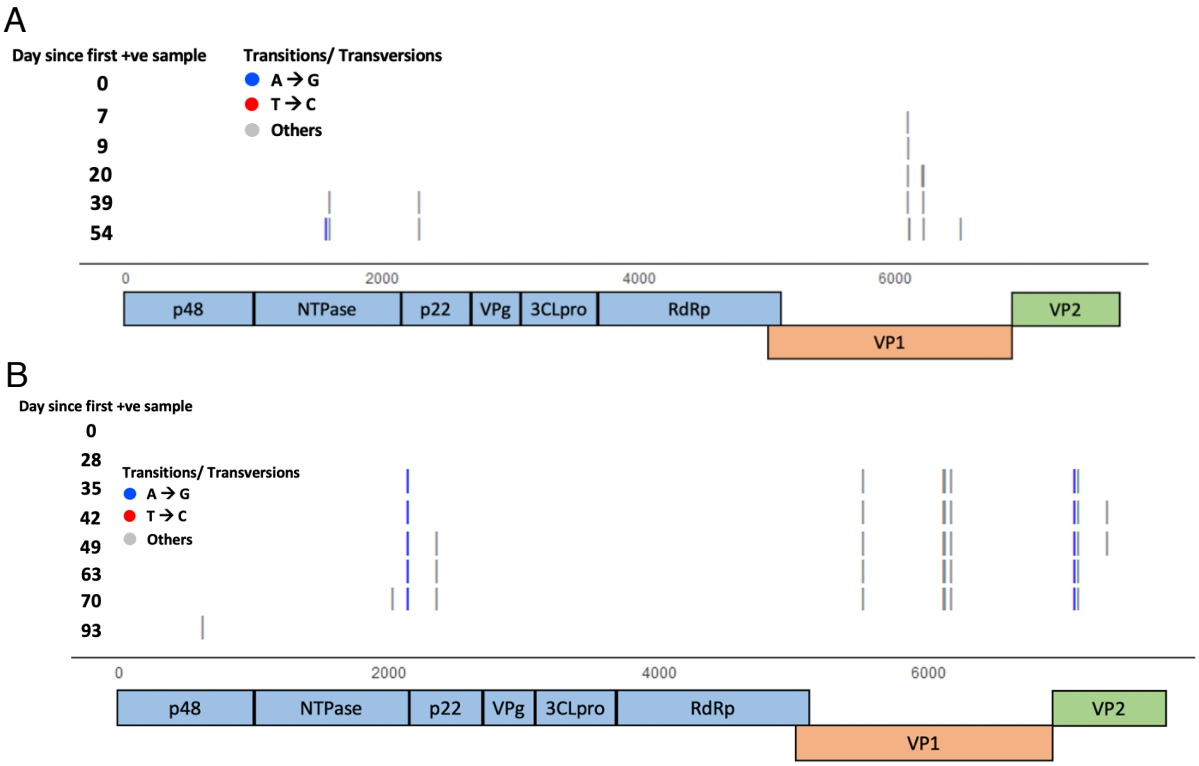


Figure 2. Waterfall plot showing non-synonymous (NS) substitutions at a consensus level (>50%) in each sample from P1 (A) and P2 (B) compared to the patient reference. *X-axis = positions across the norovirus genome, Y-axis (each row) = each sample. Each vertical bar indicates a consensus level non-synonymous substitution compared to the patient reference. The bars are coloured based on the nucleotide change, blue = A to G, red = T to C, grey = others.*

4.3.2 Norovirus patients treated with nitazoxanide and ribavirin

Previous members of the lab have analysed a cohort of 11 norovirus patients treated with monotherapy of two frequently used antiviral drugs nitazoxanide and ribavirin. Five were treated with nitazoxanide and six were treated with ribavirin. Some but not all patients have shown clinical improvements following the use of nitazoxanide or ribavirin (Brown et al., 2019).

.

Longitudinal viral sequences with sufficiently good quality were not available for every individual. Therefore, only three key clinical and virological metrics, namely the CT value, the amount of stool, and heterozygosity, were compared before, during and after the treatments. The heterozygosity was calculated by summing over the square of base frequencies over all possible bases, {A,G,C,T}, and then summing over all positions in the sequence.

$$Var = \frac{1}{\#pos} \sum_{pos} (1 - \sum_{k \in \{A,G,C,T\}} \pi_{k,l}^2)$$

, where the frequency of observing base k at site l is $\pi_{k,l}$.

Despite the clinical improvement in some patients, no significant difference in the three metrics were found when comparing pre and post treatment samples from patients treated with nitazoxanide or ribavirin.

4.3.3 Norovirus patients treated with favipiravir

In recent years, another antiviral drug, favipiravir, has been frequently used in treating norovirus infections due to its proven efficacy *in vitro* and *in vivo* (Arias et al., 2014; Furuta et al., 2013). Here, we characterised the clinical and virological response in three norovirus patients treated with favipiravir.

Patient 3 – clinical features

Patient 3 (P3), a 5-year-old girl, was presented with severe combined immunodeficiency and chronic human norovirus infection at Great Ormond Street Hospital. Based on the reports of clinical efficacy of monotherapy of the favipiravir and nitazoxanide (Brown et al., 2019; Ruis et al., 2018a), she was treated with a combination of the two antiviral drugs. She had a long history of chronic diarrhoea with multiple PCR testing positive for human norovirus infection in stool samples, with cycle threshold (CT) values ranging from 11.6 to 16.2 (Figure 3), corresponding to an extremely high viral load. On admission, she presented with severe muscle atrophy, and she weighted below the 0.1 percentile for her age (Figure 3). She was treated with twice a day, 200 mg of nitazoxanide and three times a day, 200 mg of favipiravir. Both the CT value and her weight increased steadily and her viral load decreased (Figure 3). Her treatment was paused for a week (at day 230 in Figure 3) and restarted shortly after her thymus transplantation.

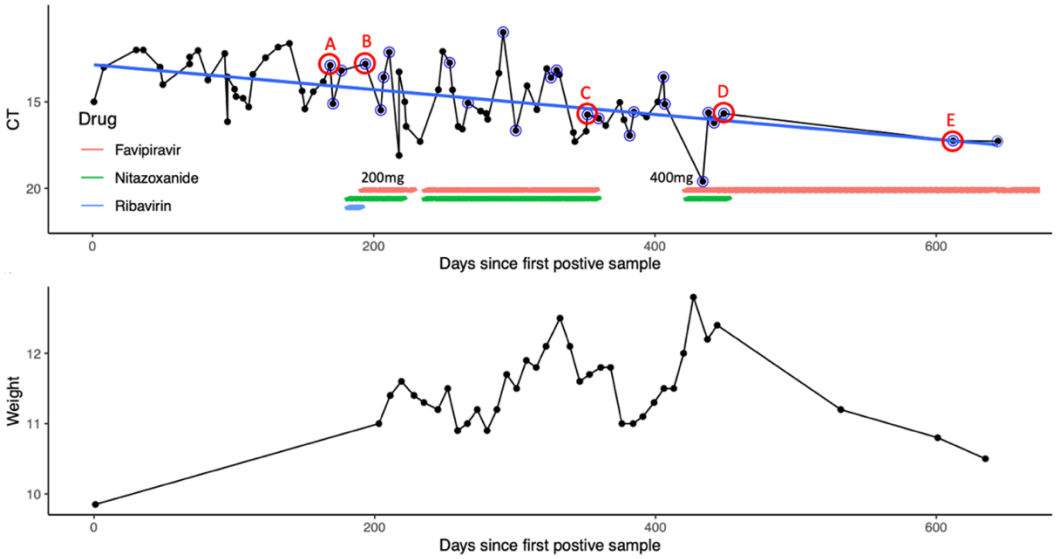


Figure 3. Overview of samples collected from P3. *Top panel: Norovirus cycle threshold (CT) values plotted with a line of best fit. The y-axis is inverted. The increase in CT value (decrease in viral load) observed over time is statistically significant ($p < 0.01$ in t-test). Blue circles indicate samples which have been sequenced. Horizontal bars below the CT values line graph indicate the period in which P3 received treatment. Red = Favipiravir, green = nitazoxanide, blue = ribavirin. Bottom panel: Weight of the patient in kilograms.*

Around 6 months after the start of the treatment, she developed a suspected drug-induced conjugated hyperbilirubinemia. Favipiravir and nitazoxanide were discontinued for two months and were restarted on an increased favipiravir dose of 400 mg once her hyperbilirubinemia had recovered. She showed good clinical improvements with decreased stool output and progressive weight gain. She was discharged home 9 months after the initial administration. Due to supply issues, nitazoxanide was discontinued at the same time, favipiravir monotherapy was continued. Unfortunately, she died of sepsis a year after the thymus transplantation, before achieving immune reconstitution.

Despite the apparent clinical improvement, including the weight gain, the reduction in viral load, although statistically significant, was slow. To better understand the impact of favipiravir, we performed whole genome sequencing on norovirus from 25 stool samples which were collected longitudinally over 16 months (Figure 3). Four samples were collected before the start of the treatment, ten during the first course of treatment and six during the second course of treatment.

Patient 3 – phylogenetic analysis

Viral genotyping confirmed P3 was infected by the GII.4 P15 Sydney 2012 human norovirus strain. All sequences from P3 clustered together in the human norovirus major capsid (VP1) phylogenetic tree (Figure 4).

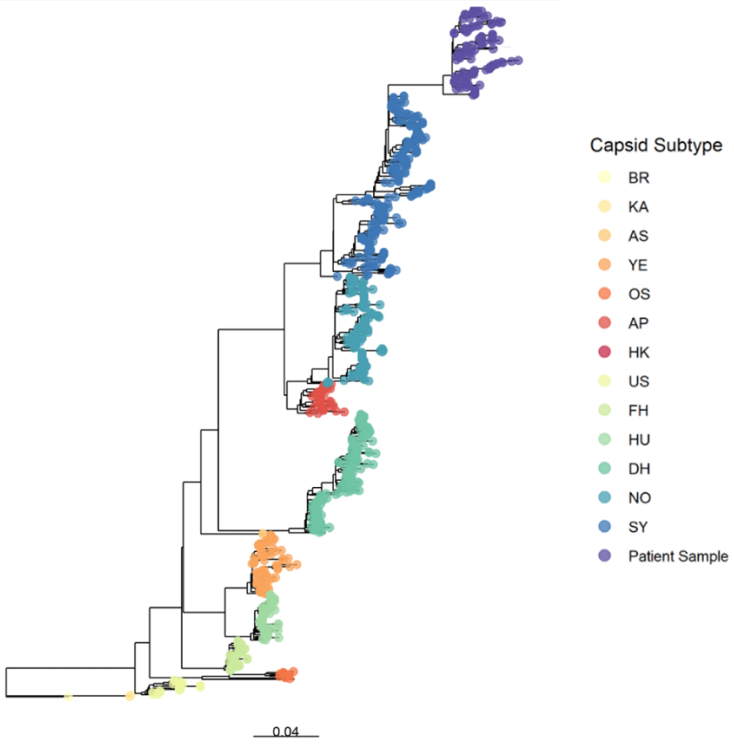


Figure 4. A major capsid (VP1) phylogenetic tree constructed with an alignment of 1000 random human norovirus sequences and all samples from P3. *The phylogenetic tree is rooted on the oldest sample (1994). Tips are coloured based on the capsid subtype. Samples from P1 are indicated in purple.*

Phylogenetic analysis revealed two stable clones which appeared to have evolved within the patient and this was confirmed by haplotype reconstruction using HaROLD (shown in amber and purple in Figure 5). The haplotype abundances varied with time (Figure 5) and the non-synonymous (NS) mutations present in the consensus sequence at each timepoint relative to the first sample also revealed evidence for dominance of one haplotype (the amber haplotype shown in Figure 5) during favipiravir and nitazoxanide treatment and its reduction in abundance when the drugs were stopped (Figure 6). In particular, in the samples taken during favipiravir monotherapy (the last two samples), the amber haplotype had completely taken over and the purple haplotype had disappeared (Figure 5).

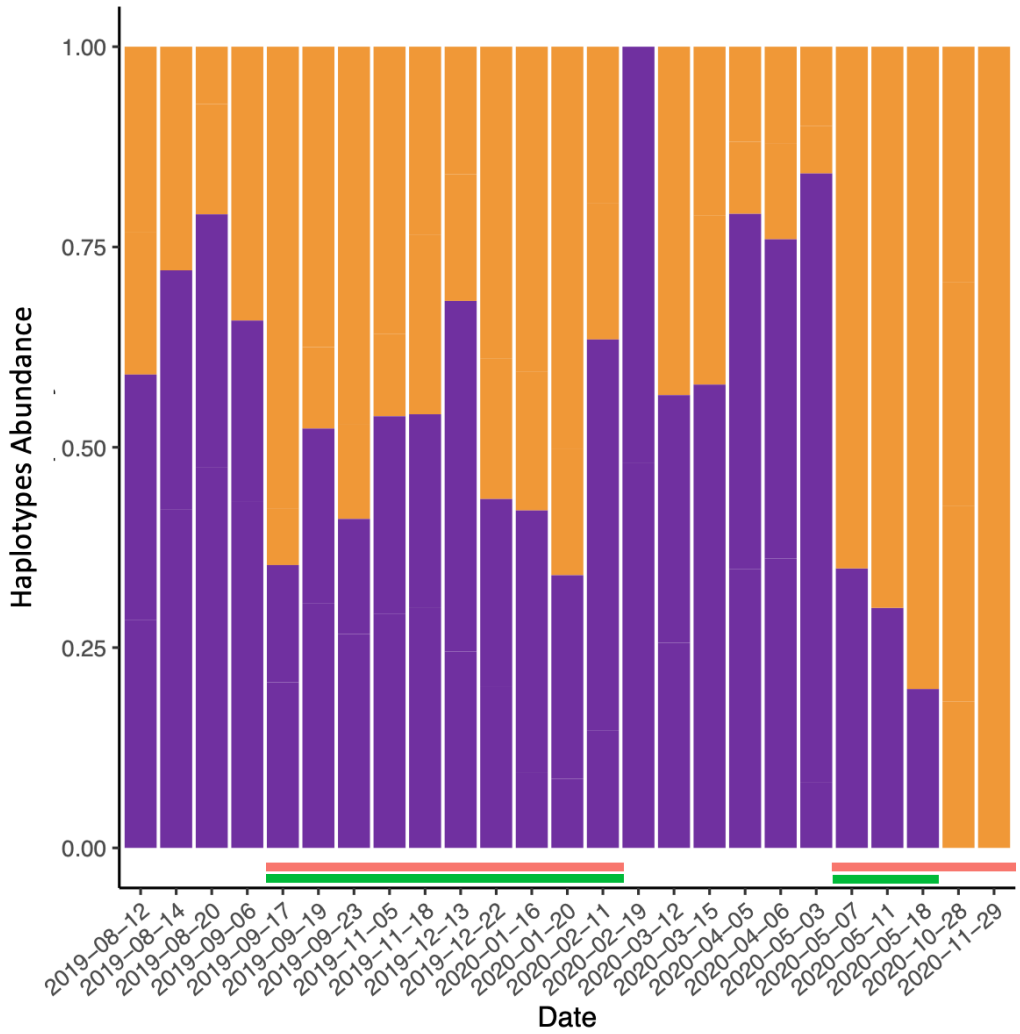


Figure 5. Changes in haplotype abundances overtime. *Y-axis = haplotype frequency shown in percentage, X-axis = samples shown in the order of collection date. Two distinct haplotypes were present. The haplotype selected during the period of drug treatment is coloured in amber. The other haplotype selected during off-treatment period is coloured in purple. Colour bars under the chart indicate samples taken during treatment, red = favipiravir, green = nitazoxanide.*

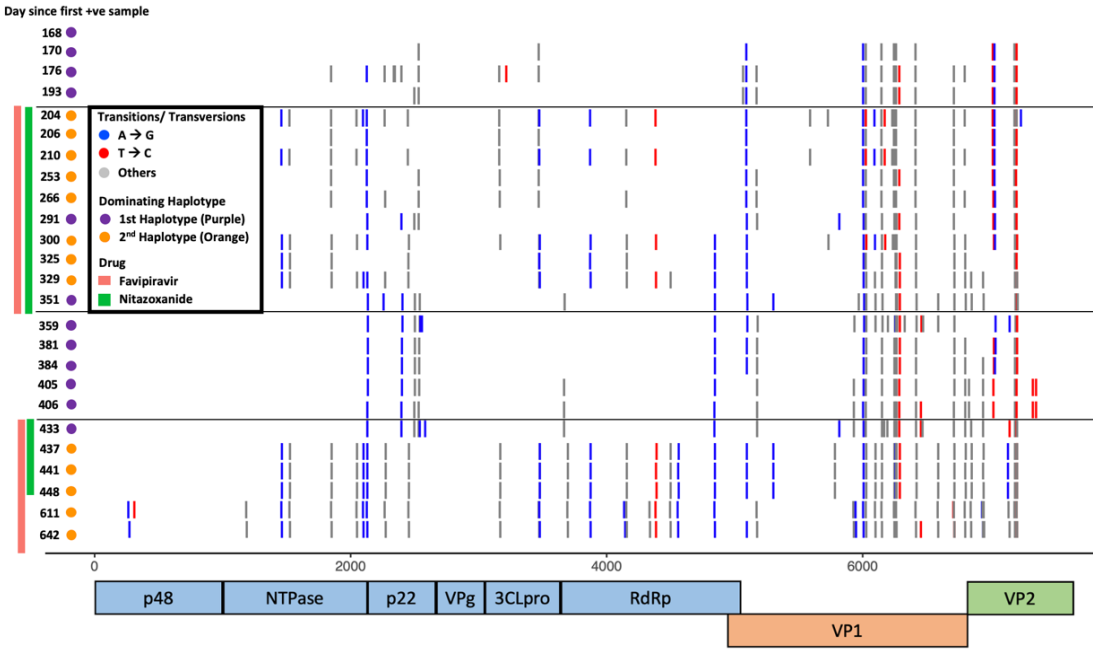


Figure 6. Waterfall plot showing all non-synonymous (NS) substitutions at consensus level (>50%) in each sample from P3 compared to the patient reference. *X-axis = positions across the norovirus genome, Y-axis (each row) = each sample. Vertical bars on the left indicate whether the sample was taken during treatment. Red = Favipiravir, green = nitazoxanide.*

Patient 4

P4 was a teenage patient with common variable immune deficiency (CVID) and chronic norovirus associated enteropathy. She showed a rapid reduction in viral load shortly after receiving favipiravir at 1200 mg, however the treatment had to stop due to drug toxicity (Figure 7). Treatment was restarted at a lower dose two weeks later. She improved clinically, where nasogastric feeding was no longer required, a decrease in stool output and increase in weight were observed. 6 stool samples were collected across 2 months (Figure 7). Deep sequencing revealed she was infected by the GII.P31 GII.4 Sydney strain.

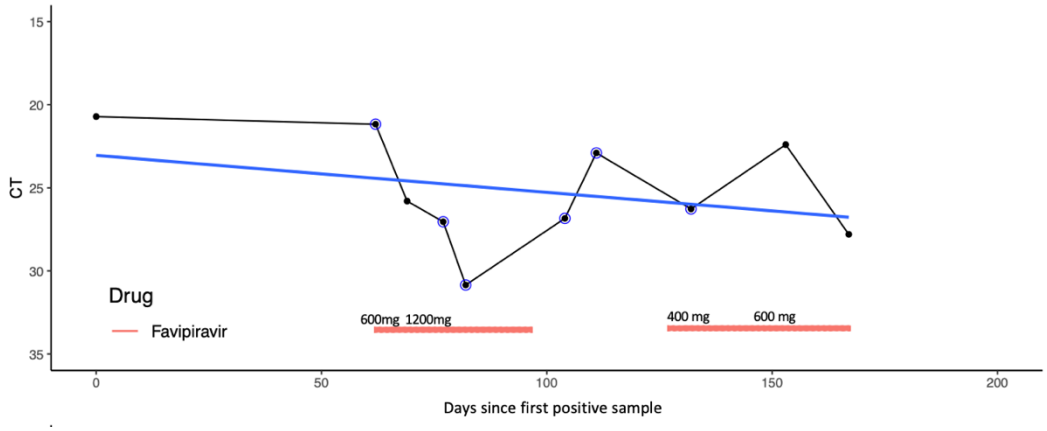


Figure 7. Overview of samples collected from P4. *Norovirus CT values plotted with a line of best fit. Changes in CT value over time is insignificant (p-value > 0.05 in t-test). However, the CT value during on treatment period is significantly lower (p-value < 0.05 in t-test). Blue circles indicate samples which have been sequenced. Horizontal bars below the CT values line graph indicate the period in which P4 received Favipiravir.*

Non-synonymous substitutions were found across the genome in all genes apart from the NTPase and the RdRp, which comprise the norovirus replication complex. Some consensus changes in the VP1 which arose during the first treatment period persisted even when the treatment has stopped (Figure 8).

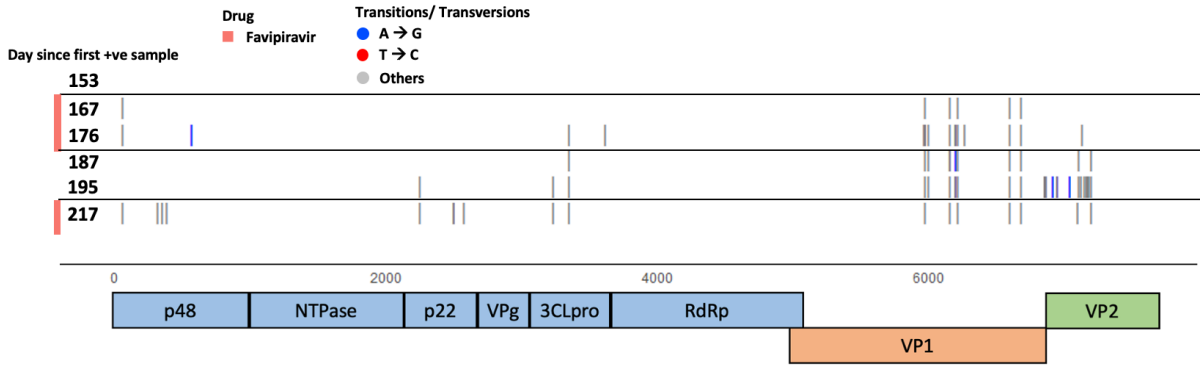


Figure 8. Waterfall plot showing non-synonymous (NS) substitutions at a consensus level (>50%) in each sample from P4 compared to the patient reference. *X-axis = positions across the norovirus genome, Y-axis (each row) = each sample. The bars on each row are coloured based on the nucleotide change, blue = A to G, red = T to C, grey = others. Red vertical bars on the left indicate samples taken during favipiravir treatment.*

Patient 5

P5 was the norovirus clinical case we analysed with HaROLD in Chapter 3. The clinical details have previously been published in Ruis et al., 2018b. P5 was a 48-year-old man with CVID and chronic norovirus infection for more than a decade. Nine stool samples were collected across 9 months (Figure 9). Two distinct haplotypes were identified, where one predominated during treatment (Figure 10 and 11).

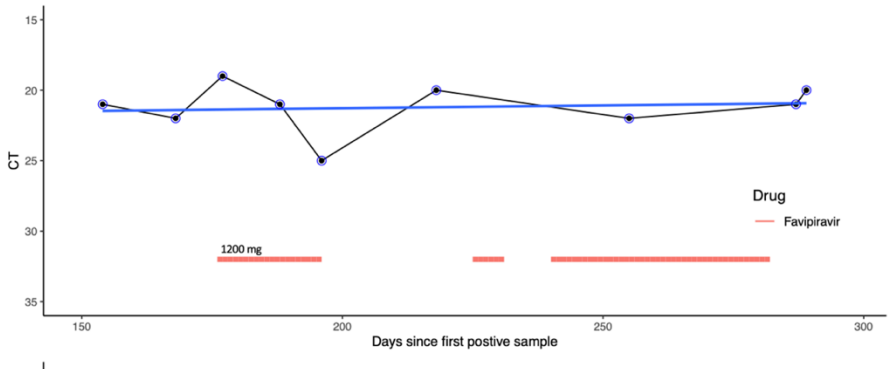


Figure 9. Overview of samples collected from P5. *Norovirus CT values plotted with a line of best fit. The changes in CT values over time, as well as difference of CT values between on and off treatment period, were insignificant (p-value > 0.05 in t-test). Blue circles indicate samples which have been sequenced. Horizontal bars below the CT values line graph indicate the period in which P5 received Favipiravir.*

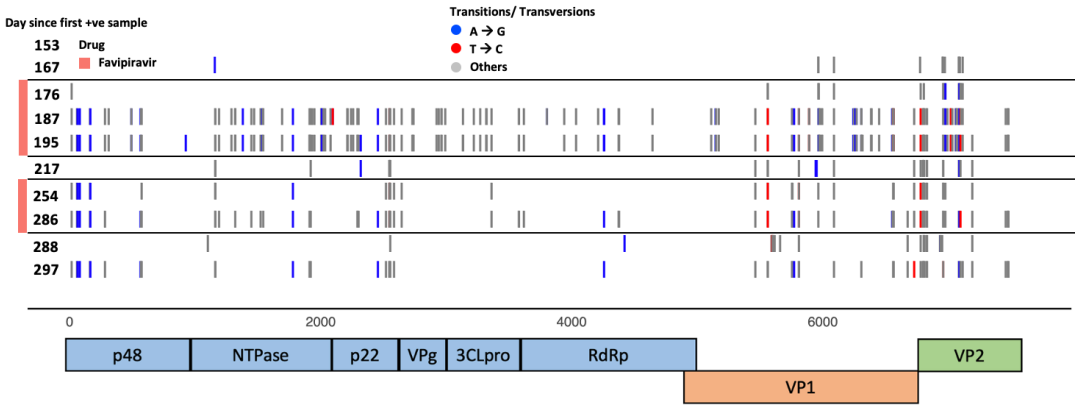


Figure 10. Waterfall plot showing non-synonymous (NS) substitutions at a consensus level (>50%) in each sample from P5 compared to the patient reference. *X-axis = positions across the norovirus genome, Y-axis (each row) = each sample. The bars on each row are coloured based on the nucleotide change, blue = A to G, red = T to C, grey = others. Red vertical bars on the left indicate samples taken during favipiravir treatment.*

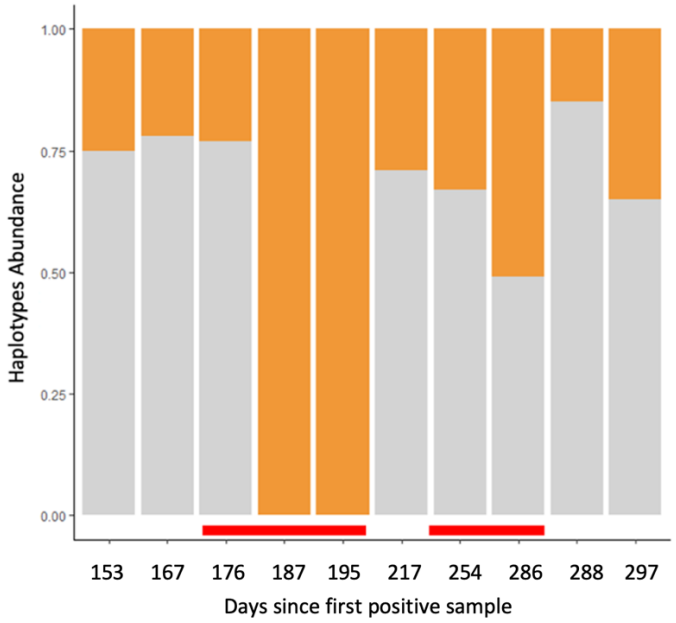


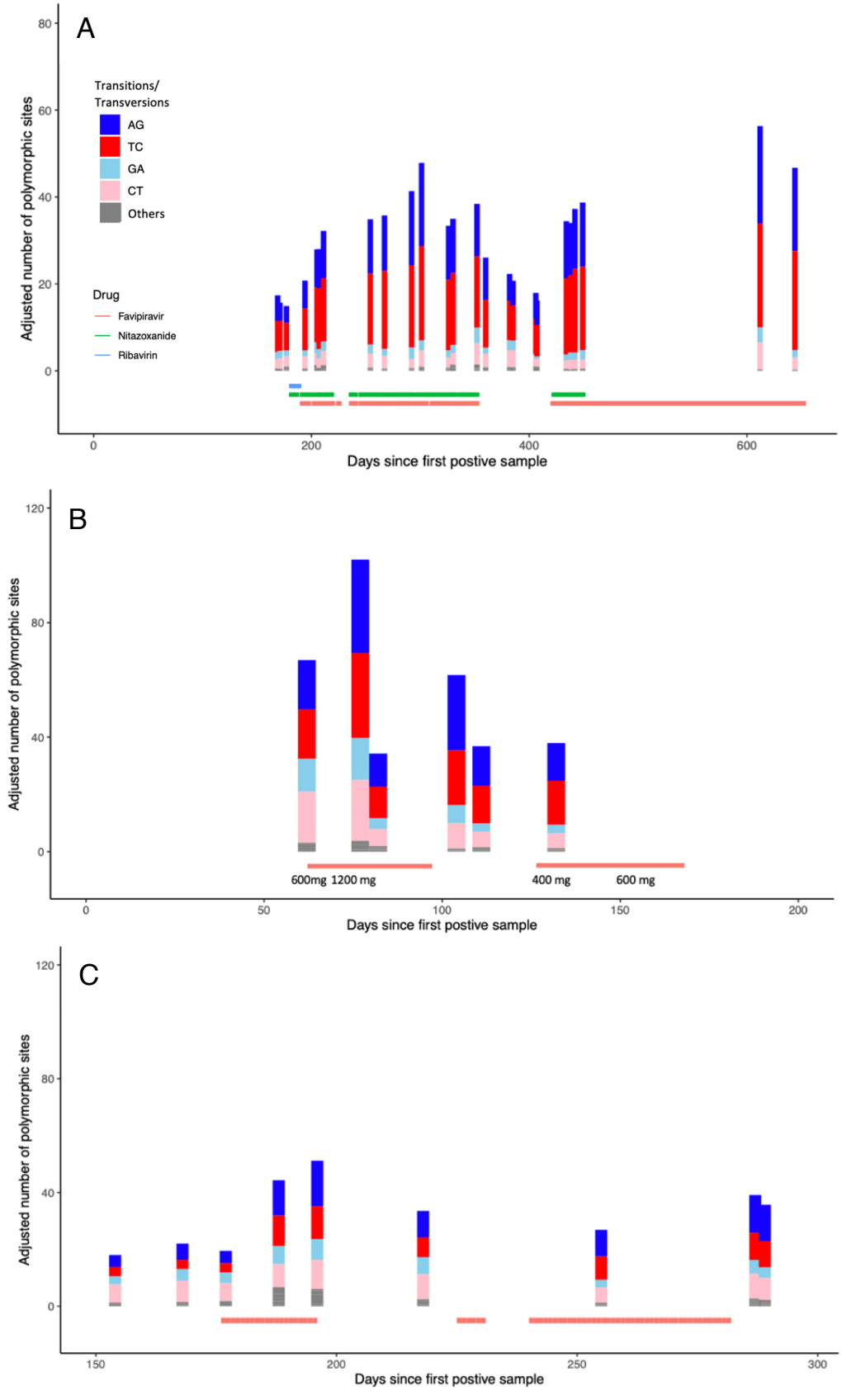
Figure 11. Changes in haplotype abundances overtime. *Y-axis = haplotype frequency shown in percentage, X-axis = samples shown in the order of collection date. Two distinct haplotypes were present. The haplotype selected during the period of drug treatment is coloured in amber. The other haplotype selected during off-treatment period is coloured in grey. Red bars at the bottom indicate samples taken during treatment.*

4.3.4 Analysis of favipiravir associated mutations

Previous studies have suggested that, in addition to directly inhibiting the RdRp, favipiravir inhibits viral replication by inducing lethal mutagenesis, an accumulation of mutations which drives the viral population to extinction (Baranovich et al., 2013; de Ávila et al., 2016; Goldhill et al., 2019; Ruis et al., 2018b). The studies reported that favipiravir primarily acts as a guanosine analogue and induces G to A and C to T polymorphisms in Ebola virus (Guedj et al., 2018), hepatitis C virus (de Ávila et al., 2016), West Nile virus (Escribano-Romero et al., 2017), Zika virus (Bassi et al., 2018) and influenza virus (Baranovich et al., 2013; Goldhill et al., 2019). In influenza, a study has shown that favipiravir can also act as an adenosine analogue and induce a small number of A to G and T to C polymorphisms (Goldhill et al., 2019).

In murine norovirus and human norovirus, it has been shown that favipiravir induces A to G and T to C polymorphisms instead (Arias et al., 2014; Ruis et al., 2018b). Here, we observed a major accumulation of A to G and T to C polymorphisms and a small accumulation of G to A and C to T polymorphisms during treatment in P3 and during the first period of treatment in P5 (Figures 12A and C). This suggests favipiravir might primarily act as an adenosine analogue instead of guanosine analogue in norovirus. Interestingly, in P3, the number of sites with these transition polymorphisms dropped after treatment cessation but rebounded and increased again during the second period of treatment (Figure 12A). Compared to P3 and P5, very few consensus level changes were observed in P4 (Figure 8). Due to the limited number of samples, it was unclear whether we observed an accumulation of A to G and T to C polymorphisms during treatment. There was also a drop in

the number of sites with these two transitions during the initial period of treatment (Figure 12B). P4 is known to have received a higher dose of treatment relative to her weight. It has previously been suggested that favipiravir can cause chain termination at a high concentration (Goldhill et al., 2019). It is possible that when P4 received a high concentration of favipiravir (1200 mg), the viral replication was completely suppressed, which resulted in the rapid fall in viral load (Figure 7). The viruses we sequenced might be sampled from compartments with poor drug penetration, which could explain the minimal mutagenic signature observed (Figure 12B).


A
Transitions/
Transversions
AG
TC
GA
CT
Others
Drug
Favipiravir
Nitazoxanide
Ribavirin
Adjusted number of polymorphic sites
Days since first postive sample
B
600mg 1200 mg
400 mg
600 mg
Adjusted number of polymorphic sites
Days since first postive sample
C
Adjusted number of polymorphic sites
Days since first postive sample

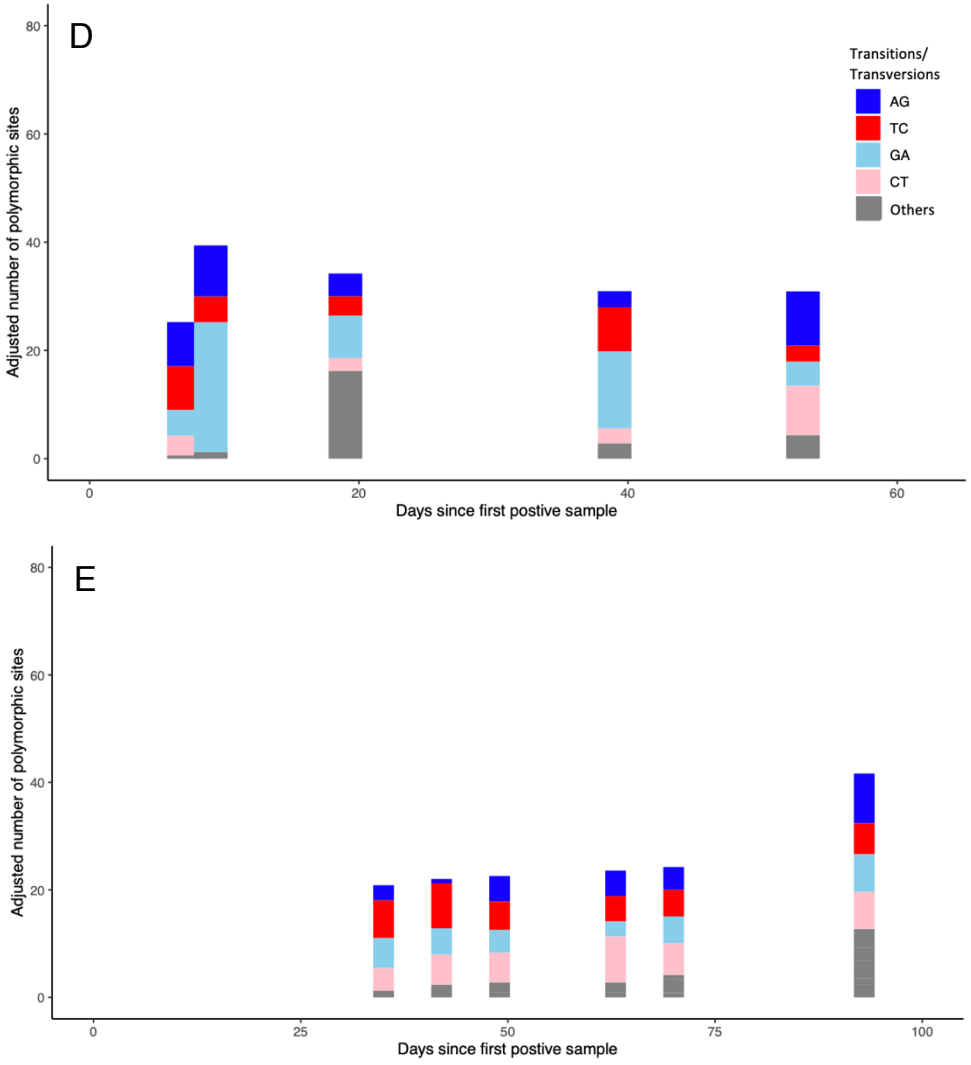


Figure 12. Stacked bar charts showing the adjusted number of polymorphic sites with different types of transitions and transversions (compared to the first patient sample) in treated patients P3 (A), P4 (B) and P5 (C) and untreated patients P1 (D) and P2 (E). *Red = A to G, royal blue = T to C, pink = G to A, sky blue = C to T, grey = other types of transitions or transversions. Only the difference of transitions and transversions count between treated and untreated samples in P3 were statistically significant (p-value > 0.05 in t-test).*

As a control, we also compared the subsequent sequences to the first patient sequence in the untreated patients P1 and P2. In the untreated

patients, there was no clear trend of increase or decrease in the number of sites with the transitions of interest (A to G and T to C) (Figures 12D and 12E). A similar number of sites with A to G, G to A, T to C and C to T polymorphisms was found in P1 and P2 (Figures 12D and 12E). In the treated patients, although a small increase in the number of sites with the reverse transitions G to A and C to T polymorphisms was also observed, during the treatment period, in general, we found a higher number of sites with A to G and C to T polymorphisms (Figures 12A, 12B, 12C and 13). In the untreated patient P1, an increase in other types of transitions and transversions was found in the third sample (Figure 12D). The same was observed in the last sample from P2 (Figure 12E). Only the difference of transitions and transversions count between treated and untreated samples in P3 were statistically significant (p-value > 0.05 in t-test).

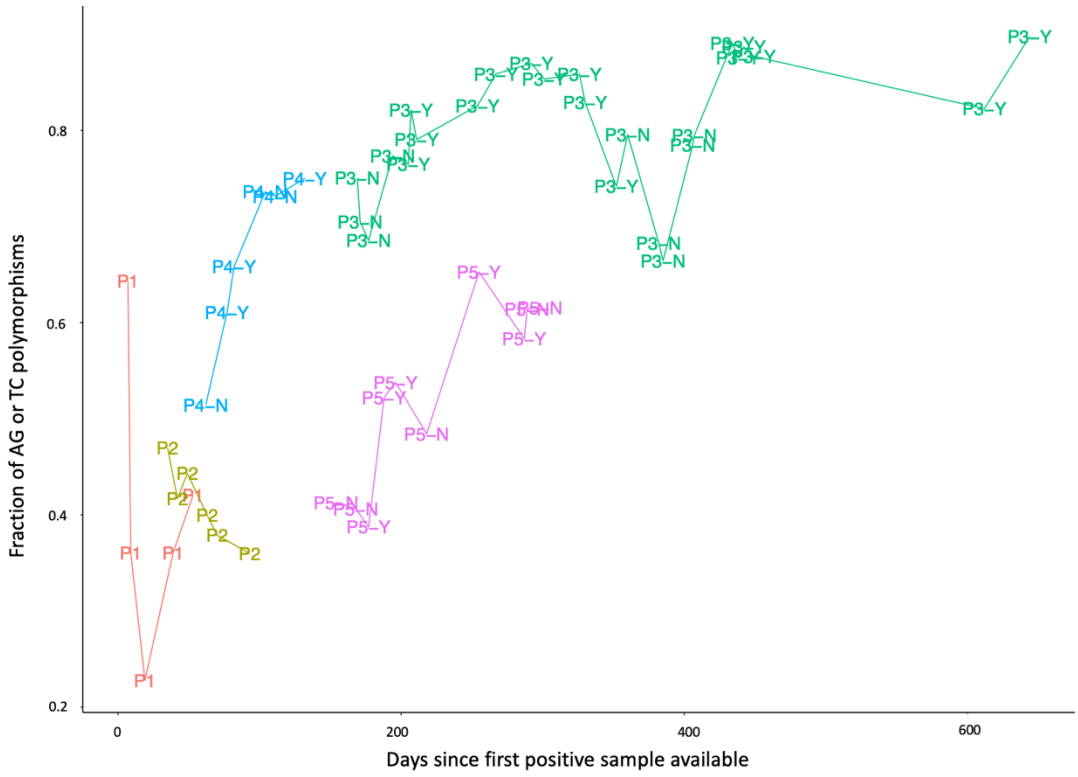


Figure 13. The fraction of AG or TC polymorphisms in samples overtime in all five patients. *P1 to P5 are coloured in*

orange, dark green, bright green, blue and pink respectively. In the sample labels, Y indicates the sample was taking during treatment, N indicates the sample was taking during pre-treatment or off-treatment period. Although the trend is complex, in general, an accumulation of A to G or T to C polymorphisms was found during treatment, especially in P3. However, compared to other patients, P3 had a higher percentage of A to G and T to C polymorphisms even before the start of the treatment. In the untreated patients P1 and P2, no accumulation of A to G or T to C polymorphisms was found.

Treated patients – variants analysis

Relative to the first sequence, all five patients showed consensus level amino acid substitutions over time (Figures 2, 6, 8, 10). However, only patients P3 and P5 developed mutations in the RdRp, the target of favipiravir binding (Figures 6 and 10). RdRp amino acid substitutions arising in the haplotype predominating during favipiravir in patients P3 and P5 are listed in Table 1. The numbers of consensus level changes in untreated patients were significantly lower than for the favipiravir-treated patients P3 and P5 (Figure 2) and the proportion of low-level mutations due to A to G and T to C changes was also lower (Figure 13).

Table 1. Favipiravir associated non-synonymous mutations in the RdRp region in P3 and P5. *Nucleotide changes that correspond to the drug-induced polymorphisms are highlighted in blue (A to G) and red (T to C). Columns 4-8 show the number of database sequences (out of 1000 randomly selected sequences analysed) with the original amino acid, the mutated amino acid, any other amino acid, a stop codon, and missing data (X), at the position of interest.*

Patient	Amino Acid Changes	Nucleotide Changes	Original AA	Mutated AA	Others	Stop Codon	NA
P3	A44S	G-->T	992	0	0	0	8
P3	K103R	A-->G	923	30	5	42	0
P3	I193V	A-->G	892	108	0	0	0
P3	S198A	T-->G	588	412	0	0	0
P3	L239M	T-->A	954	46	0	0	0
P3	I274T	T-->C	735	173	92	0	0
P3	A312T	G-->A	928	14	58	0	0
P3	I332V	A-->G	812	188	0	0	0
P5	L5K	T-->A	0	993	4	3	0
P5	S18N	G-->A	93	0	603	304	0
P5	V125M	G-->A	721	159	120	0	0
P5	S156N	G-->A	7	0	808	185	0
P5	V215I	G-->A	725	275	0	0	0
P5	K231R	A-->G	968	32	0	0	0
P5	H270N	C-->A	216	0	783	1	0
P5	T360S	A-->T	174	0	826	0	0

In P3, four RdRp variants (K103R, S198A, I274T, I332V) were dominant during the first period of low dose favipiravir treatment (shown in orange in Figure 14). They became less abundant during the off-treatment period but rose to fixation when favipiravir was restarted at a higher dose (shown in orange in Figure 14). This suggests the variants might be induced by favipiravir. A further four mutations (A44S, I193V, L239M, A312T) that had been present at below consensus level during the initial, low-dose favipiravir treatment, also rose to fixation six months after high-dose favipiravir was started. During the first period of treatment, A312T rose from below 5 percent to close to 50 percent within a month (with two samples in between). It then followed the same trend of rapid increase in frequency with other variants highlighted in orange in Figure

13. I193V and L239M (shown in red in Figure 14) remained at low frequency (< 10%) during the low dose favipiravir treatment and off-treatment period. After the treatment has restarted on a higher dose, they did not show the same rapid increase in frequency observed in the variants highlighted in orange (A44S, K103R, S198A, I274T, A312T, I332V). They only rose to fixation almost six months after the start of the higher dose favipiravir treatment. Mutation N427S (indicated with an arrow in Figure 14) which became fixed in the population before the second treatment period started may also have acted to stabilise the genomes carrying the putative resistance mutations.

Figure 14. Changes of RdRp variant frequencies over time in samples from P3. *X-axis shows different samples over time, y-axis shows the variant frequencies in percentage. X-axis is not presented on a time scale. The last two samples were taken six*

months after the treatment was restarted. Red bars at the bottom indicate samples which were taken during the treatment period. Variants highlighted in orange are those that were dominant (i.e. at consensus level, > 50%) during the first period of treatment, subsequently suppressed following the pause of treatment, rapidly increased in frequency when the treated was restarted and rose to fixation (100%) six months after the start of the second period of treatment. Variants highlighted in red are those that appeared at low level during the first period of treatment, the off-treatment period, and the initial samples collected during the second period of treatment but rose to fixation six months after the start of the second period of treatment.

A similar picture was evident for the RdRp in P5, with eight NS mutations (Table 1) rising to fixation during favipiravir treatment and falling again when treatment was interrupted (shown orange in Figure 15). Four of these mutations (S18N, S156N, H270N, T360S) were never observed in the 1000 randomly selected Genbank sequences (Table 1).

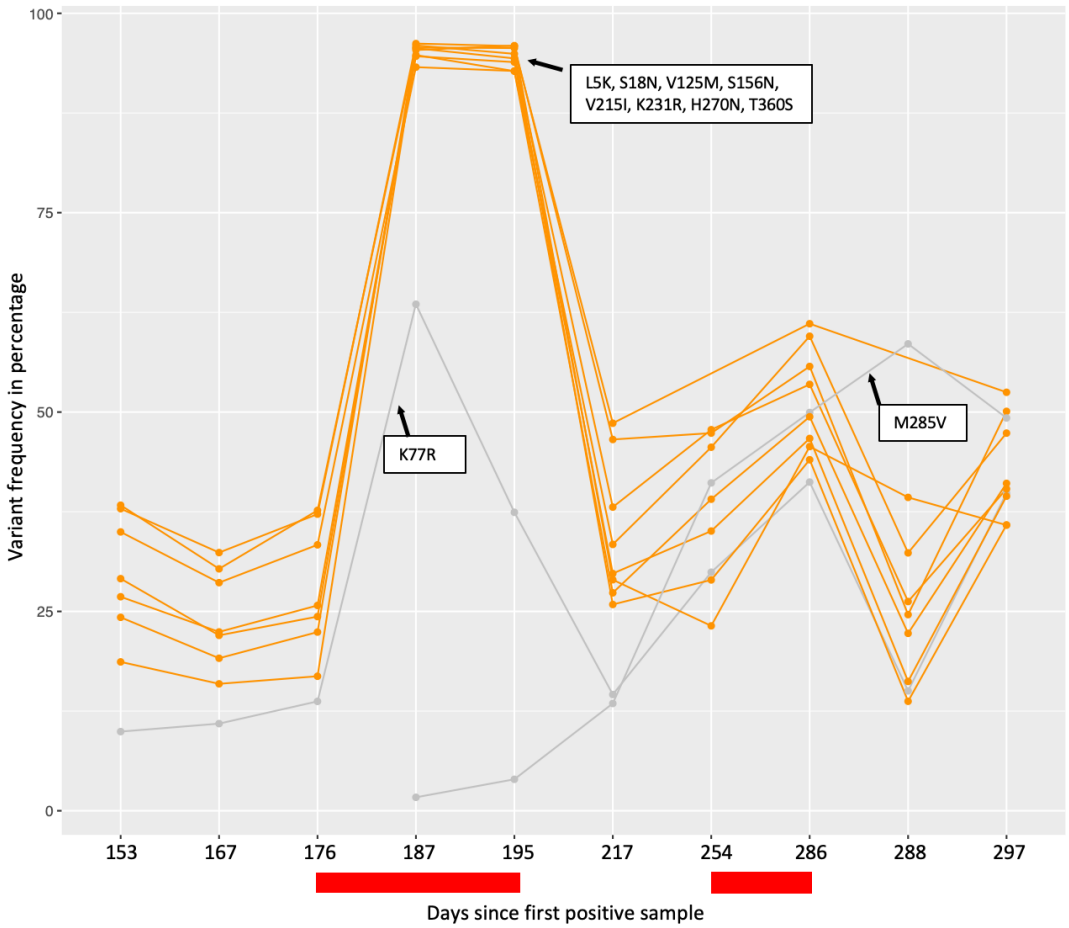
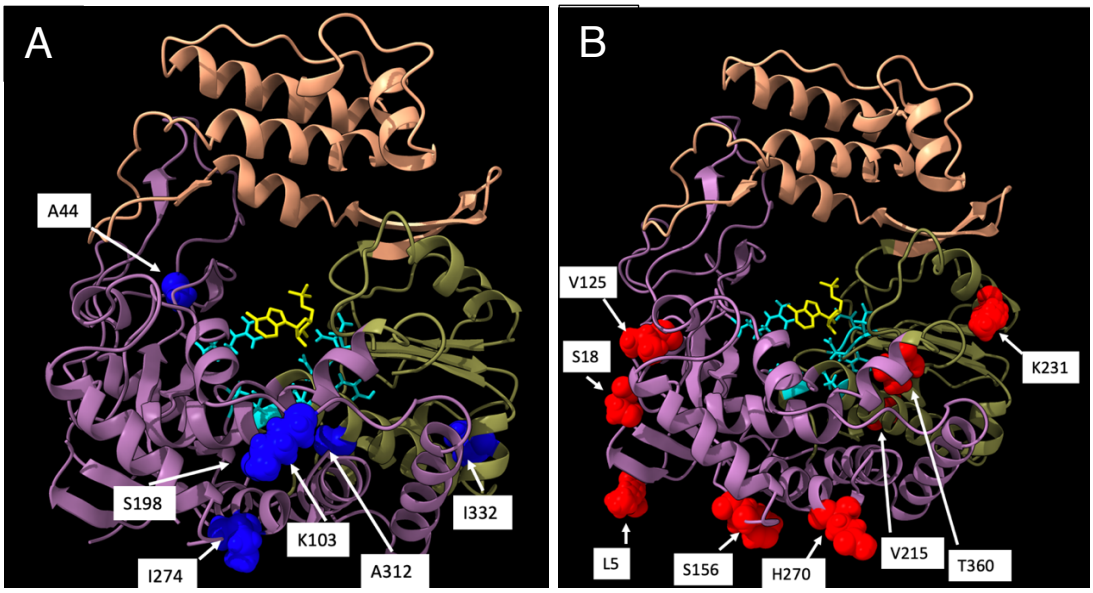


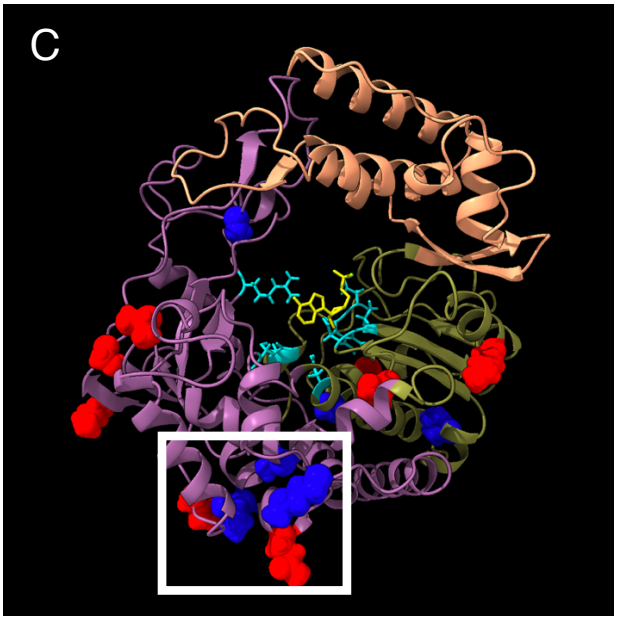
Figure 15. Changes of RdRp variant frequencies over time in samples from P5. *X-axis shows different samples over time, y-axis shows the variant frequencies in percentage. X-axis is not presented on a time scale. Red bars at the bottom indicate samples which were taken during the treatment period.*

Structural analysis

The positions of the P3 and P5 RdRp mutations on the predicted crystal structure are shown in Figures 16A and 16B. None of the substitutions mapped to residues reported to be favipiravir resistant from studies of influenza and chikungunya (Baranovich et al., 2013; Delang et al., 2014). Three of the favipiravir-induced substitutions in P3 (K103R, S198A, I274T) and two in P5 (S156N, H270N) appeared closely cluster together

(Figure 16C). These substitutions identified were found in sites that were variable in the 1000 randomly selected database sequences summarised in Table 1. These substitutions were present in the fingers subdomain of the RdRp, but not in the catalytic region in the palm subdomain. It is unclear whether these substitutions affect the function of the RdRp. Nonetheless, by simulating the distribution of the 16 mutations of interest (Table 1) a thousand times, the probability of five mutations clustering in the same space with the same distance in between was low (p = 0.00102).





D

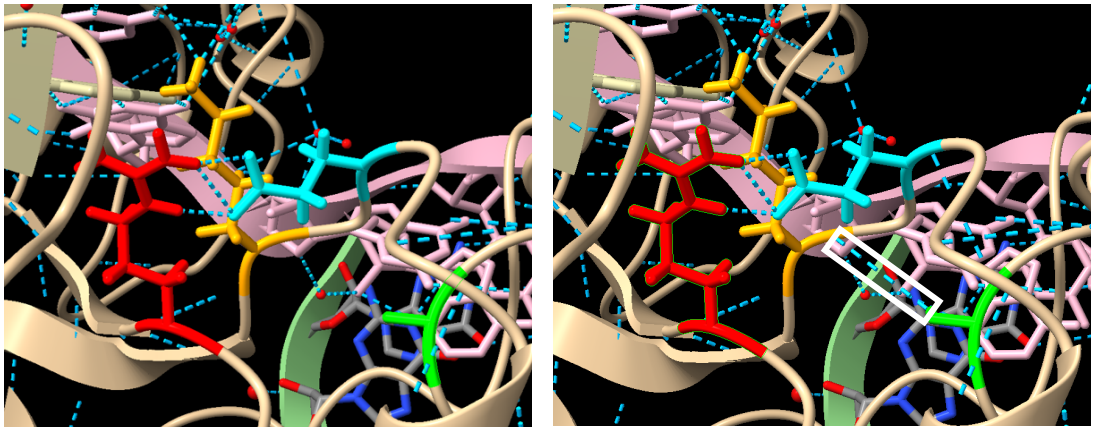


Figure 16. Positions of favipiravir associated mutations on predicted crystal structure of the RdRp. *The protein structure is shown as a ribbon diagram. The RdRp active site is indicated in cyan. The ribbons are coloured based on the subdomains. Dark pink = fingers, green = palm, peach = thumb. No mutations were present in the thumb subdomain. The nucleotide incorporated most recently via catalytic activity is coloured in yellow. (A) Mutations from P3. (B) Mutations from P5. (C) Plotting the mutations from P3 (blue) and P5 (red) together, with the white box indicating the cluster of mutations of interest. The cluster is*

located in the fingers subdomain. (D) Changes in bonds and interactions caused by the mutation from Alanine to Serine at position 44 (highlighted in green). Red = site 182, a key catalytic residue. Orange = K166, based on the homologous structures, this site is associated with favipiravir resistance in influenza (Goldhill et al., 2018)*. Blue = E168, which has strong hydrogen bond with site 182, shifting of this site could alter the binding domain. A substitution from Alanine to Serine at site 44 is predicted to introduce the new hydrogen bond indicated with the white box.*

In contrast, five favipiravir-induced substitutions in P3 (A44S, K103R, I193V, L259M, A312T) were in residues which are normally conserved (Table 1), including one of three sites, A44S, that became fixed in P1 after five months of favipiravir treatment. A44S, which is otherwise invariant in 1000 randomly selected publicly available norovirus genomes (Table 1) is close to the catalytic residue R182 (6.7Å, i.e. these residues can interact directly or indirectly with one another). A44S is also close (6.1Å) to K166 in norovirus which is the homologue of the favipiravir resistance mutation K229R in influenza (Baranovich et al., 2013). In silico substitution of serine at position A44 is predicted to introduce a new hydrogen bond with the amide backbone of K166. (Figure 16D) This could potentially cause resistance through displacing K166 with which it tightly contacts thus effecting an indirect change in the favipirivir-contacting R182. A similar mode of action has been proposed for K229R mediated favipiravir resistance (Baranovich et al., 2013).

4.3.5 *In vivo* activity of favipiravir in norovirus infected zebrafish larvae

Previously, the lack of robust small animal models and *in vitro* systems that support efficient norovirus replication was the main barrier to validating the efficacy of favipiravir (de Graaf et al., 2016; Todd and Tripp, 2019). Here, our collaborator Dr Joana Duarte Da Rocha Pereira and Emma Roux from Katholieke Universiteit Leuven evaluated the *in vivo* activity of the favipiravir in a zebrafish larvae model which has been shown to allow efficient replication of human norovirus in a manner that recapitulates infection of human host (Van Dycke et al., 2019) (Pang et al., manuscript in preparation).

The three-day-old zebrafish larvae were inoculated using a pre-treatment stool sample obtained from P3. The experiment was repeated four times. At each timepoint, 10 zebrafish larvae were harvested. At the peak of replication, 3 days after the inoculation, a single injection of 25 ng of favipiravir (dose selected was based on weight-dependent conversions of the patient dose) resulted in an average of 1.59 $\log_{10}$ reduction in viral RNA copies per zebrafish.

In addition, to evaluate whether prolonged treatment with favipiravir can lead to a loss of norovirus infectivity, five stool samples (~ 1000 genome copies each) from P3 (Samples A to E, Figure 4) were independently injected into five sets of zebrafish larvae. Each of the five experiments (Samples A to E) was repeated five to twelve times and ten zebrafish larvae were harvested at each time point. No antiviral treatment was given to the zebrafish larvae. They observed that the samples taken before the patient received treatment (Sample A) or after a treatment

gap (Sample D) had significantly higher increases in viral RNA copies compared to those observed during and immediately following treatment (Samples B, C and E) (Figure 17A).

They then estimated the number of viral RNA copies per zebrafish needed to achieve a successful infection in 50% of the cases (ID_{50}) by inoculating zebrafish larvae with dilutions of samples A to E in separate experiments. The ID_{50} of samples taken before the patient received treatment (Sample A) or after a treatment gap (Sample D) were significantly lower than the other samples taken during or immediately following treatment (Samples B, C and E) (Figure 17B).

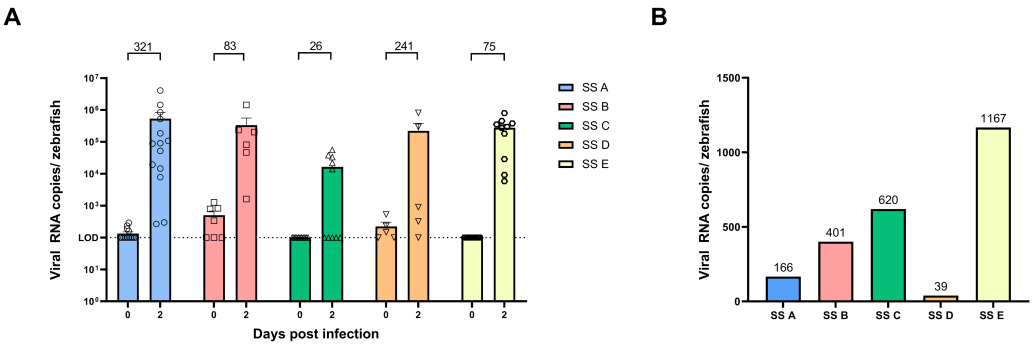


Figure 17. Replication efficacy of norovirus in zebrafish receiving favipiravir. (A) *Bars represent viral RNA copies/zebrafish, quantified by RT-qPCR. The dotted line represents the limit of detection (LOD). In every independent experiment (n=5-14), 10 zebrafish larvae were harvested at each time point. Mean values ± standard errors are presented. An average of 321-fold increase in viral RNA was observed on day 2 post-inoculation for Sample A (pre-treatment). In Samples B and C (collected 1 week and 6 months into treatment) a 83- and 26-fold increase in viral RNA were observed respectively. After a*

treatment gap of two months, the norovirus infectivity (Sample D) increased with a 241-fold increase in viral RNA copies at the peak of replication) and decline once again to a fold increase of 75 (Sample E) 6 months after treatment was re-introduced. ***(B)*** *Bars represent the* ID_{50}*. The* ID_{50} *of Sample A (pre-treatment) was quantified as 166 viral RNA copies/zebrafish, which increases to an average of 401 (Sample B) and 620 (Sample C) viral RNA copies/zebrafish after 1 week and 6 months of treatment, which indicates a decrease in infectivity. After a 2-month treatment gap, viral infectivity again increased with the* ID_{50} *falling to an average of 39 viral RNA copies/zebrafish (Sample D). Reintroduction of favipiravir treatment of the patient again reduced the infectivity with the* ID_{50} *rising to an average of 1167 viral RNA copies/zebrafish for Sample E (Figure 15B). (Figures provided by Emma Roux.)*

4.3.6 Discussion on norovirus analysis

Summary

In Sections 4.3.1 to 4.3.3, we presented data from untreated norovirus patients, ribavirin or nitazoxanide monotherapy treated patients and favipiravir-treated patients. In the untreated patients, there were no significant changes in CT value overtime. In the ribavirin and nitazoxanide monotherapy treated patients, there were no significant changes in the CT value, number of stools, heterozygosity, before, during or after treatment. Of the three favipiravir-treated patients, only one (P3) showed a significant increase in CT value (decrease in viral load) over time. However, in the favipiravir-treated patients, we observed an accumulation of drug-associated mutagenic signatures. In Section 4.3.4, we described some putative drug-resistant mutations in P3 and P5. We identified a region in the RdRp with a cluster of favipiravir-associated mutations which could be further tested in the laboratory. In Section 4.3.5, we presented evidence that the apparent clinical efficacy observed was associated with favipiravir-induced reduction in viral infectivity in a zebrafish larvae model.

The major limitation of our analysis is that all three treated patients received favipiravir through compassionate access programme instead of being part of a well-designed randomised clinical trial. As the treatments were non-standardised and the data were opportunistically collected, it was difficult to compare across different patients and draw conclusions, especially in such a small cohort with mixed demographics and comorbidities.

Efficacy of favipiravir

Despite the lack of reduction in viral load, the favipiravir-treated patients improved clinically. We have demonstrated that favipiravir has clinical efficacy in treating human norovirus. Similar to previous studies (Crotty et al., 2001; Dapp et al., 2012; Day et al., 2005; de Ávila et al., 2016; Díaz-Martínez et al., 2018, 2018; Goldhill et al., 2019; Ortega-Prieto et al., 2013; Perales et al., 2009; Sierra et al., 2000; Vignuzzi et al., 2005), we found that despite minimal changes in the viral load, favipiravir induced mutagenesis in the viral population. In other viruses, favipiravir primarily acts as a guanosine analogue which induces G to A and C to T polymorphisms, and secondarily acts as an adenosine analogue which induces A to G and T to C polymorphisms. In our patients, we observed a major accumulation of A to G and T to C polymorphisms, and a small accumulation of the reverse transitions of G to A and C to T polymorphisms, which suggest favipiravir primarily acts as an adenosine analogue in human norovirus. In P3, we demonstrated that pausing treatment resulted in a decrease in the frequency of these transitions, a partial restoration of *in vitro* infectivity and a clinical deterioration. Reintroduction of favipiravir leads to an increase in mutagenic signature, again a reduction in viral infectivity and a gradual improvement of clinical symptoms. This provides further evidence to support the theory that in addition to directly inhibiting the action of RdRp, favipiravir as well as other nucleoside analogues can act on the virus by inducing mutagenesis which leads to the loss of fitness of the virus (Baranovich et al., 2013; de Ávila et al., 2016; Goldhill et al., 2019; Ruis et al., 2018b).

Favipiravir-associated mutations

In addition, we have identified 16 substitutions from P3 and P5 which emerged with the use of favipiravir, including one which is located near the amino acid homologue of a residue which confers favipiravir resistance in influenza. We have also identified a region in the RdRp with a cluster of favipiravir-associated mutations from P1 and P3, which can be tested in further study in the laboratory (Figure 16C).

During the treatment period, the genomes carrying the favipiravir-associated RdRp mutations (putative drug resistant mutations) were selected. When treatments were paused, the genome without these mutations again dominated the viral population, which suggests that the putative resistant genomes were less fit. This could also be explained by the hitchhiking of deleterious mutations on the putative resistance mutations, or the fitness-reducing pleiotropic effects of these mutations. Due to their loss of fitness, they are unlikely to circulate in the community as a resistant strain or pose a risk to other patients. However, there is a possibility that compensatory changes might arise through processes such as random genetic drift, which could increase the fitness of the virus with these mutations. In addition, they could potentially be selected if there is widespread use of favpiriavir or other RdRp inhibitors in the population.

Evaluation of norovirus infectivity in Zebrafish larvae model

In Section 4.3.5, for the first time, we provided *in vivo* evidence that the norovirus infectivity was reduced following the treatment, even when viral load remained unchanged. Although the zebrafish larvae model allows robust human norovirus replication, norovirus RNA can only be detected for six days post-infection (Van Dycke et al., 2019). We were, therefore, unable to fully recapitulate the chronic infection in the patients or monitor the long-term effect, including any mutagenic signature, caused by the antiviral treatments.

4.3.7 Within host variations in untreated and remdesivir-treated SARS-CoV-2 patients

Similar to norovirus infections, at the start of the COVID-19 pandemic in 2020, knowledge of within-host viral variations in SARS-CoV-2 was limited. Different drugs, including RdRp inhibitors such as favipiravir and remdesivir, have been offered to patients through the compassionate access programme and repurposed to treat the infections. However, the results from clinical trials of these repurposed drugs are mixed. To understand the within-host variations in SARS-CoV-2 infections as well as the heterogeneity in the efficacy of remdesivir, a frequently used repurposed drug in the UK, we applied similar methods used in analysing norovirus within-host variations in Section 4.3.3 to study the longitudinal deep sequencing data of SARS-CoV-2 from the upper respiratory tract in a cohort of six untreated and three remdesivir-treated patients.

Overview of patients

The full clinical detail including age, ethnicity, comorbidities and diagnosis has been published in (Boshier et al., 2020b). In summary, all nine patients are children, with age ranging from 0 to 14 years old at the time of the infection (mean age = 4.7 years old). Multiple longitudinal nasopharyngeal samples were repeatedly taken during the course of infection (minimum = 2 samples, maximum = 13 samples, median = 5 samples). Of the nine patients, four (A, D, G, H) were admitted into the paediatric intensive care unit (PICU) (Figure 18). All patients admitted to the PICU were offered remdesivir treatment through a compassionate access programme. Patient H refused treatment, while patients A, D and

G received eight to ten days of remdesivir treatment. Patients A and G received a 200 mg loading dose of remdesivir, followed by 100 mg of daily dose; patient D, who was an infant, received 10 mg of loading dose (5 mg of drug per kg weight) and a daily dose of 2.5 mg (1.25 mg per kg).

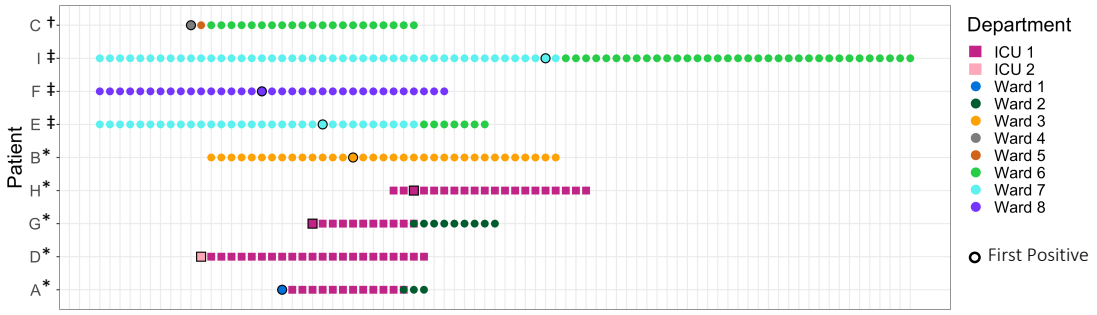


Figure 18. Overview of patients' timeline. *The first positive sample from each patient is outlined in black. The colours indicate the ward to which the patients were admitted to, the shapes indicate the type of ward (general ward or intensive care unit). * = transferred from another hospital, ‡ = long-term in-patient, † = admitted directly from the community.*

CT values

We monitored the longitudinal viral PCR CT values for all nine patients (Figure 19). In line with previous studies (Y. Liu et al., 2020), and similar to our observations described in Chapter 2, the CT values showed considerable day to day variation between 0.16 and 14.4, with a median of 5.5. Viral RNA was detectable for 7 to over 50 days, with a median of 16 days, following the first positive sample available to us (Figure 19).

In untreated patients B, F and I, the CT value falls below the detection limit within ten days since the first tested positive at Great Ormond Street

Hospital. In patients C and E, the CT values showed great fluctuation where multiple negative samples were found in between two positive samples. Although patient H was in PICU and received no treatment, a clear decline in viral load (increase in CT value) was observed. Of the three patients (A, D and G) who received remdesivir, only patient D showed a drug-associated suppression of viral load. The viral load in patient D initially fell below the detectable limit during treatment, but it rebounded after treatment cessation (Figure 19). Patients A and G showed a small gradual decline in viral load, but it is unclear whether these were caused by the antiviral treatment. However, all three patients who treated with remdesivir showed a clinical improvement after starting the drug, a fall in temperature (in A, D and G) and inflammatory markers (in A and G) were observed. Other clinical details have been described in Boshier et al., 2020b.

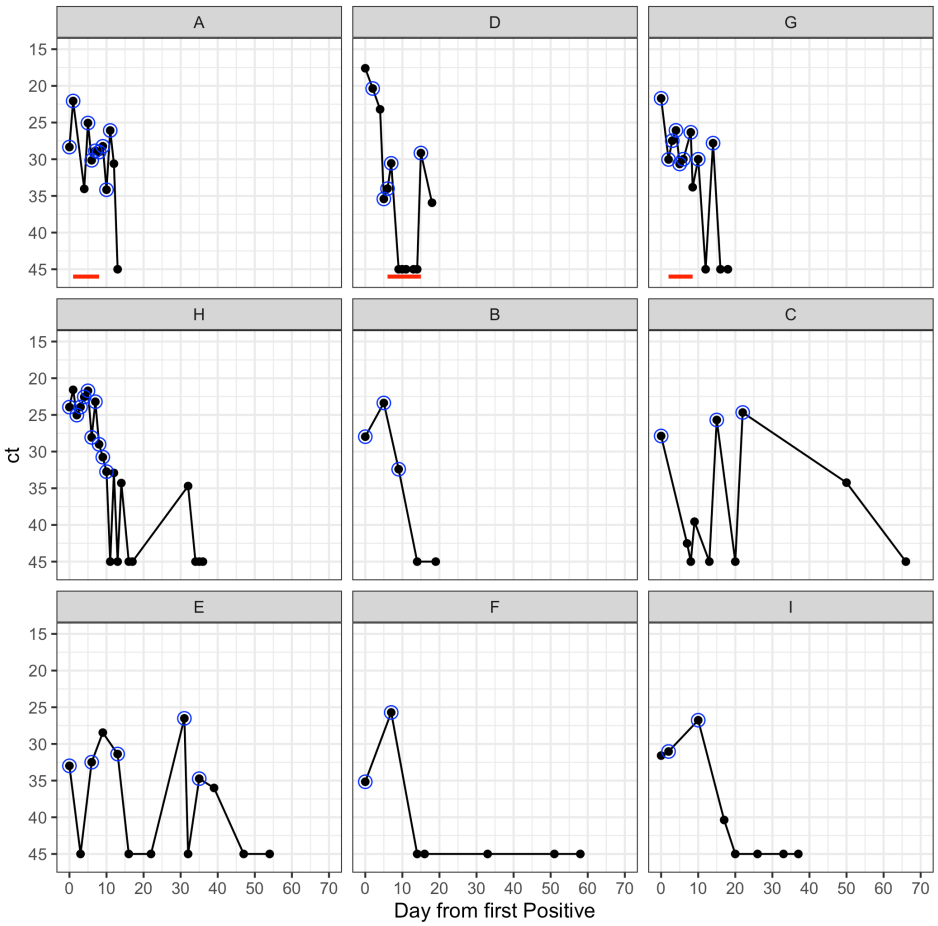


Figure 19. CT trajectories of the nine patients (A to I) from the day which they were first tested positive, to up to 70 days following their first positive. *One panel per patient. The red line indicates the period for which remdesivir was received, the blue circle indicates a sample which was successfully sequenced with good quality.*

Sequence analysis

To further study the viral dynamics in these patients, we deep sequenced all positive samples available at the time in all nine patients (Figure 19). Relative to each patient reference (first sample of each patient), no consensus level polymorphisms were found in the subsequent samples from patients B to G. A total of nine consensus

changes (five non-synonymous, four synonymous) were identified from the viruses from patients A, H and I. Of the five non-synonymous mutations, four were found in the ORF1ab (nsp 1, 3, 4, 5), one was found in the S2 domain of the spike protein (Table 2). None of the identified polymorphisms were at sites known to be susceptible to Illumina sequencing error, identified as common homoplasies, or associated with remdesivir resistance. Interestingly, sequences at samples 6, 7 and 8 from patient A were identical to the sequences from patient D (Figure 20).

Table 2. Summary of consensus level mutations in patients A, H, and I relative to their first available sample. *No consensus level mutations were found for remaining patients (B, C, D, E, F, and G). Nucleotide replacement relative to entire genome, protein replacement site within gene.*

Patient	Day Post-First Positive	Nucleotide Replacement	Protein Replacement	Gene [product]
A	1, 5, 6, 7	T10776C	L241P	Orf1ab [nsp5]
A	6, 7, 8	T9438C	I295T	Orf1ab [nsp4]
H	4,5,6,8,9,10	T3096C	L126S	Orf1ab [nsp3]
H	4,5,6,8,9,10	T16308C	synonymous	Orf1ab [nsp13]
H	4,5,6,8,9,10	G19671A	synonymous	Orf1ab [nsp15]
H	12	C28253T	synonymous	Orf8
I	10	G376T	D37E	Orf1ab [nsp1]
I	10	C23997T	P812L	S [S2 domain]
I	10	C28732T	synonymous	N

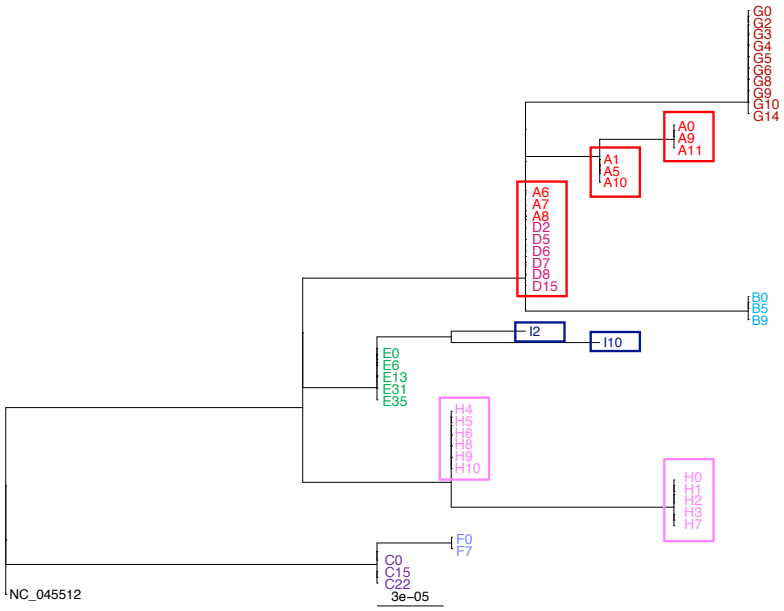


Figure 20. Phylogenetic tree of consensus sequences of all samples from Patients A to I. *The boxes highlight sequences which are exactly identical apart from missing data gaps, found in patients A, H and I over time. All samples are labelled as [Patient][Time (days)].*

Mutagenic signature

In sections 4.1, 4.3.4 and 4.3.5, we discussed lethal mutagenesis being the mechanism of action for some nucleoside analogue RdRp inhibitors. Remdesivir, despite being a nucleoside analogue, have not shown to induce lethal mutagenesis *in vitro* (Kaptein et al., 2020a; E. P. Tchesnokov et al., 2019). To understand the within host variations in the patients and to exclude non-lethal mutagenesis as an explanation for the observed high viral RNAs during treatment, we compared the mutational burden and the count of different types of transitions and transversions in treated and untreated patients. We did not find an increase in mutational burden in the treated patients, nor any mutagenic signature associated with lethal mutagenesis (Figures 21 and 22). We

did not observe any difference between the proportion of different types of transitions and transversions in treated and untreated patients (Fisher T-test, p-value = 0.13) (Figure 22).

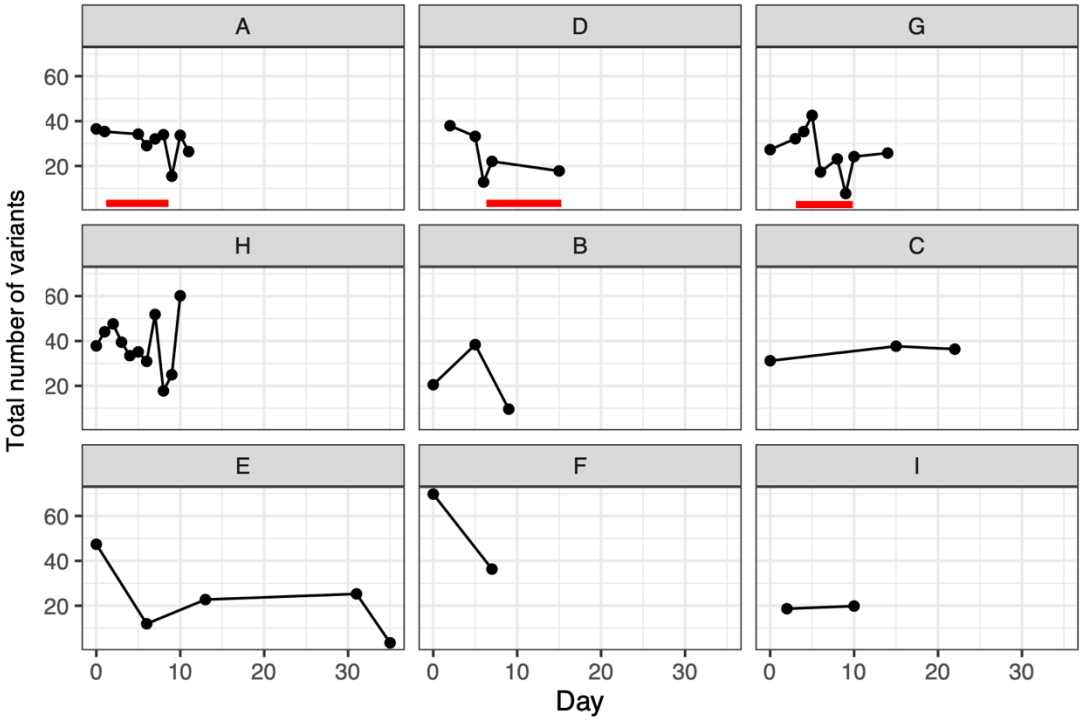


Figure 21. Mutational burden over time in each patient. *Y-axis is the count of polymorphisms with frequency higher than 2%. X-axis is days post-infection. Red line in patients A, D, and G indicate administration of remdesivir. Remdesivir was not found to be associated with a change in mutational burden over time. Mutational burden is relatively stable over time across all patients.*

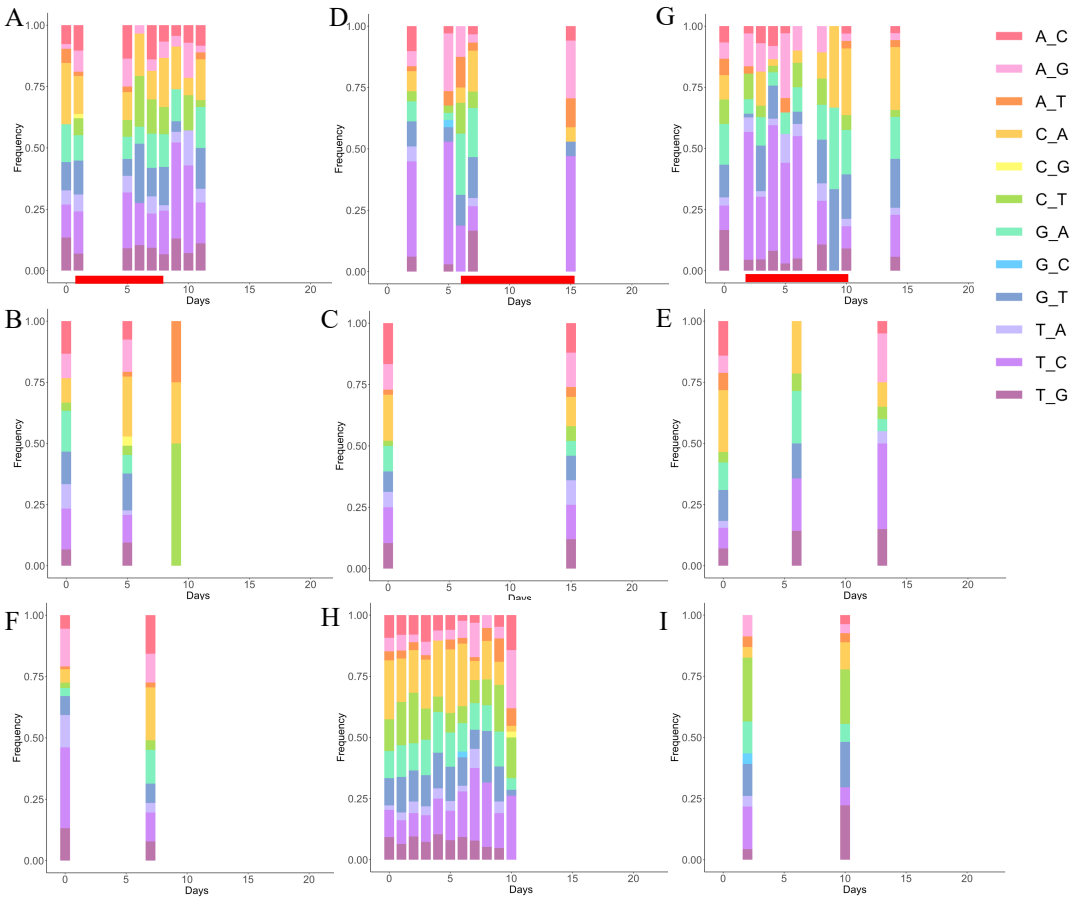


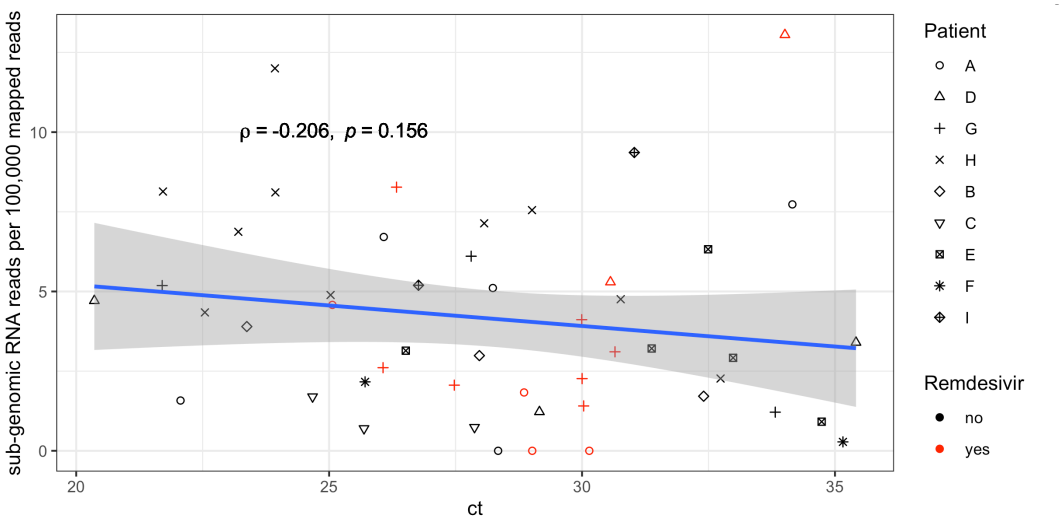
Figure 22. Mutagenic signature for each patient over time. *Stacked bars indicate frequency of transitions and transversions. The red bars indicate the periods which remdesivir treatment was received by the patients. There was no significant difference between the proportion of different types of transitions and transversions in treated and untreated patients.*

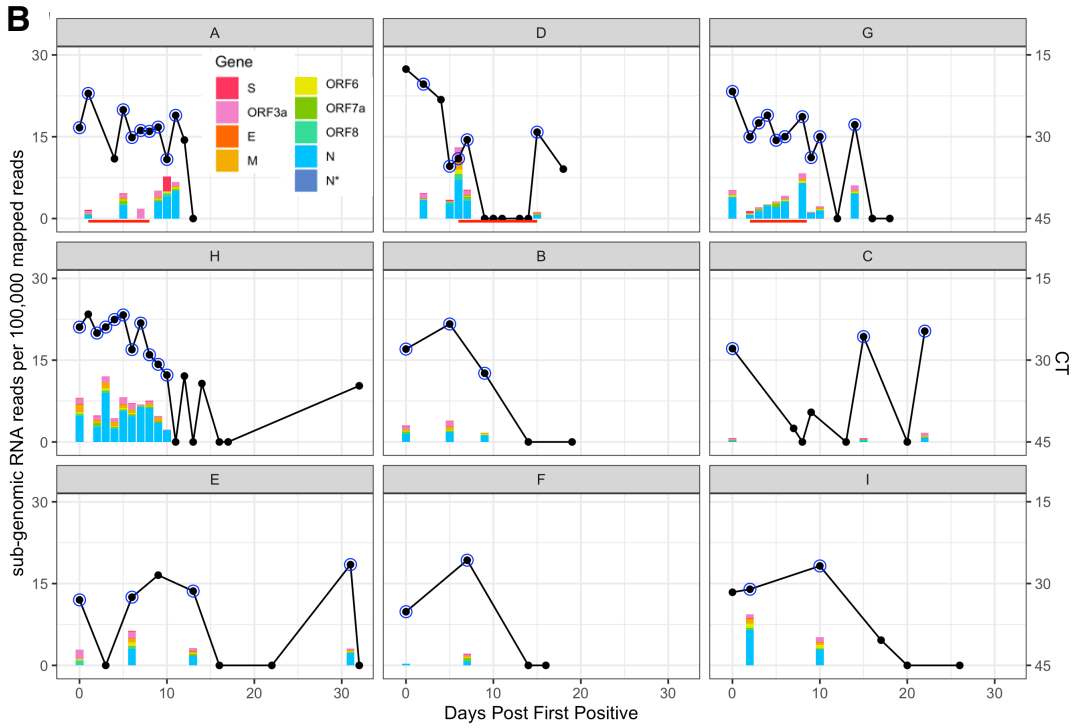
Subgenomic RNA

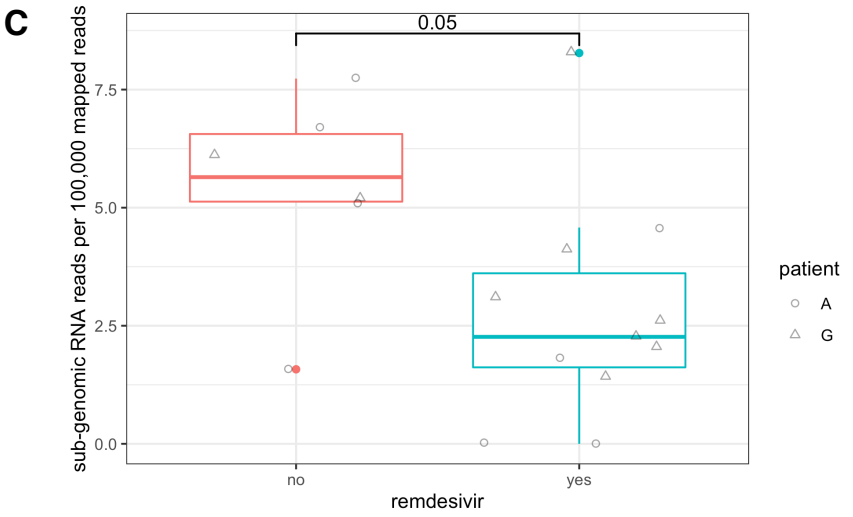
To determine whether remdesivir has an effect on the virus, given the lack of suppression on viral load, we analysed the subgenomic RNA (sgRNA) read count using Periscope (Parker et al., 2020). As discussed in Chapter 1, Section 1.3.6, subgenomic RNAs are generated during the viral replication of SARS-CoV-2. A previous study has shown that subgenomic RNA count can be used as an indicator of active viral

replication (Parker et al., 2020). Unlike previous studies (Chen et al., 2022; Verma et al., 2021), we found no correlation between the CT values and the sgRNA reads per 100,000 mapped reads (sgRPHT) (Figure 23A and B). We compared the sgRNA levels in samples taken during remdesivir treatment and those taken during the off-treatment period in patients A and G (Figure 23B). Patient D was excluded from this analysis because their viral load fell below the detection limit during the treatment period. We found that in patients A and G, the level of sgRNA during treatment was lower (Mann-Whitney-Wilcoxon test, p-value = 0.05) (Figure 23C). When considering all samples across all nine patients, we found a similar tendency towards significance of the sgRPHT levels during remdesivir treatment (Mann-Whitney-Wilcoxon Test p= 0.059) (Figure 23D).

A




B
A
D
G
H
B
C
E
F
I
Gene
S
ORF3a
E
M
ORF6
ORF7a
ORF8
N
N*
sub-genomic RNA reads per 100,000 mapped reads
CT
Days Post First Positive


C
0.05
sub-genomic RNA reads per 100,000 mapped reads
remdesivir
no
yes
patient
A
G

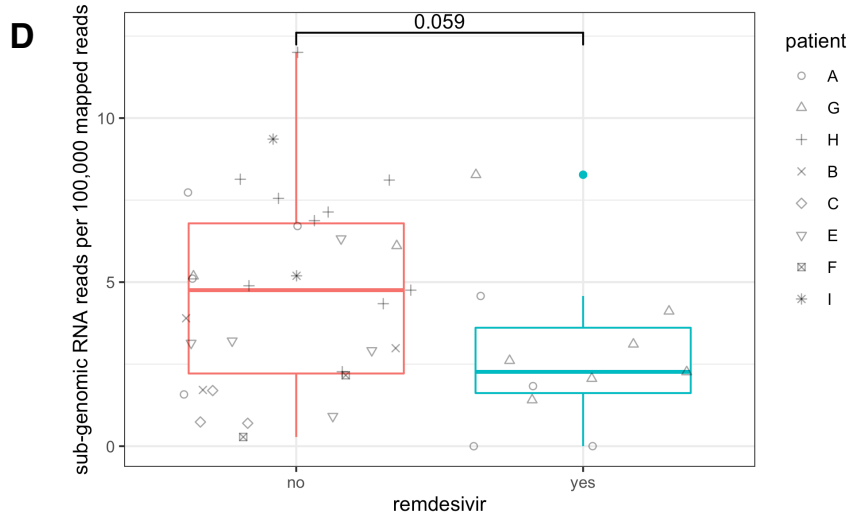

D
0.059
sub-genomic RNA reads per 100,000 mapped reads
remdesivir
no
yes
patient
A
G
H
B
C
E
F
I

Figure 23. Comparison of sgRNA reads per 100,000 mapped reads (sgRPHT) in treated and untreated individuals. *A) Scatter plot of sgRPHT vs CT values for all patients. B) Comparison of CT values and sgRPHT over time by patient. Stacked bars represent sgRPHT values coloured by gene. Black line represents CT values, with blue circles indicate samples we successfully sequenced. Y-axis = number of days since the first positive sample. C) Box plot of sgRPHT on (teal) and off (red) remdesivir for patients A and G. Samples from each individual are identified by their shape. D) Box plot comparing sgRPHT in treated and untreated samples across entire dataset.*

Variants and haplotypes analysis

To further understand the effect of the drug and the within-host variations, we analysed the minority variants found in each patient. Most patients had transient variants which were either detected at low level (less than 20%) with low read count support, or only observed in a single sample. Only patients A, B, H and I had well-supported minority variants which changed in frequency over time (Figure 24). To determine whether these variants fall on the same minority variant genome, we reconstructed haplotypes using HaROLD (Pang et al., 2020b). HaROLD identified three distinct haplotypes in patient A, two in patient B, four in patient H, and two in patient I. We constructed a phylogenetic tree using the haplotype sequence alignment. We found that the haplotype sequences were clustered by patients (Figure 25). Interestingly, the haplotype 1 from patient A cluster with all the consensus sequences from patient D (Figure 26B). We analysed the change in haplotype frequencies over time, but we observed no obvious pattern of change in all four patients (Figure 26A).

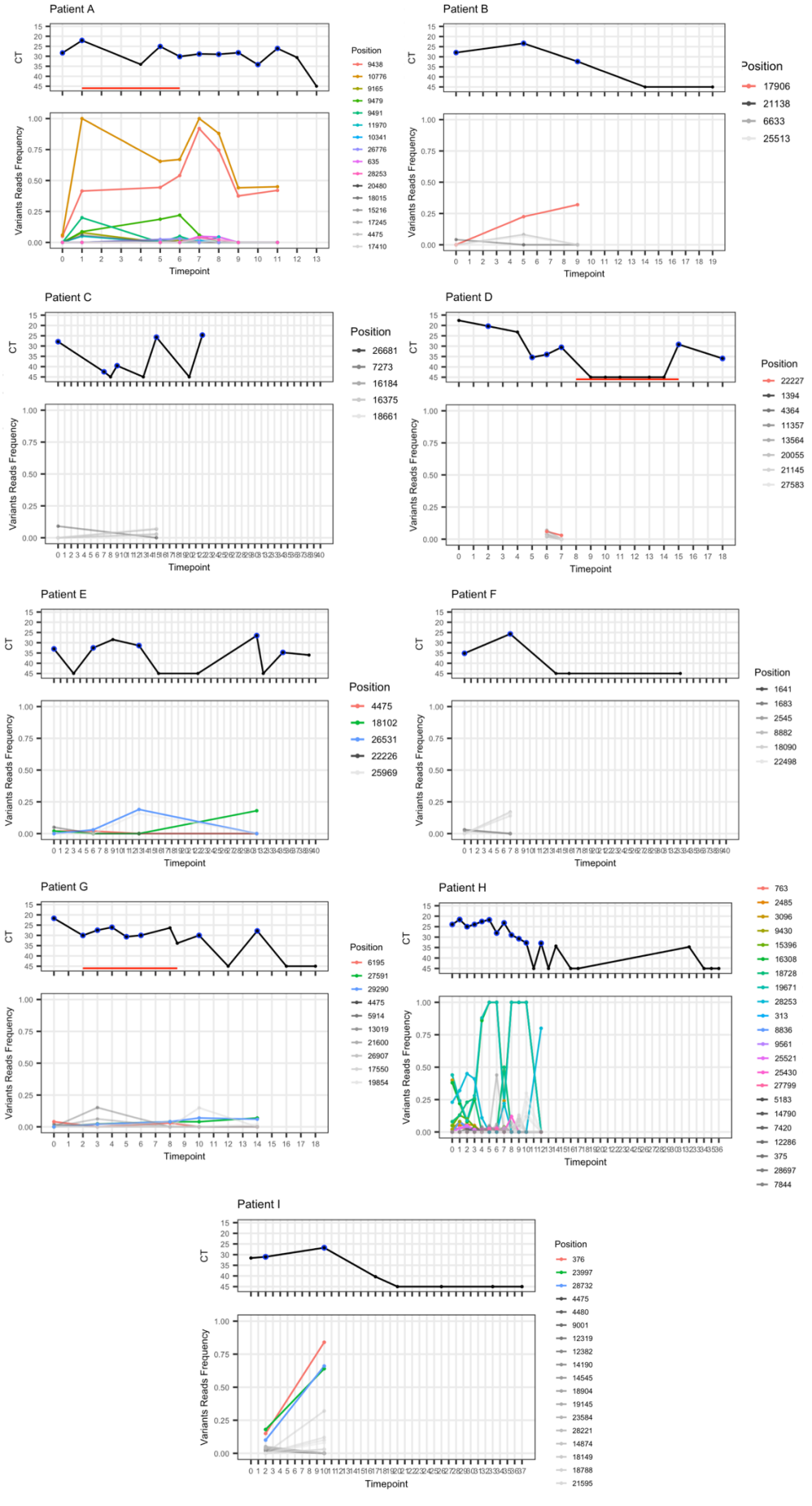

Patient A
CT
Variants Reads Frequency
Timepoint
Position
9438
10776
9165
9479
9491
11970
10341
26776
635
28253
20480
18015
15216
17245
4475
17410
Patient B
CT
Variants Reads Frequency
Timepoint
Position
17906
21138
6633
25513
Patient C
CT
Variants Reads Frequency
Timepoint
Position
26681
7273
16184
16375
18661
Patient D
CT
Variants Reads Frequency
Timepoint
Position
22227
1394
4364
11357
13564
20055
21145
27583
Patient E
CT
Variants Reads Frequency
Timepoint
Position
4475
18102
26531
22226
25969
Patient F
CT
Variants Reads Frequency
Timepoint
Position
1641
1683
2545
8882
18090
22498
Patient G
CT
Variants Reads Frequency
Timepoint
Position
6195
27591
29290
4475
5914
13019
21600
26907
17550
19854
Patient H
CT
Variants Reads Frequency
Timepoint
763
2485
3096
9430
15396
16308
18728
19671
28253
313
8836
9561
25521
25430
27799
5183
14790
7420
12286
375
28697
7844
Patient I
CT
Variants Reads Frequency
Timepoint
Position
376
23997
28732
4475
4480
9001
12319
12382
14190
14545
18904
19145
23584
28221
14874
18149
18788
21595

Figure 24. Polymorphisms trajectories for patients A to I. *Top panel: Black line = CT values. Red line = remdesivir treatment period. Blue circle = sample successfully sequenced. Bottom panel: Variant reads frequencies color coded by site. Lines in grey scale indicate polymorphisms which are transient.*

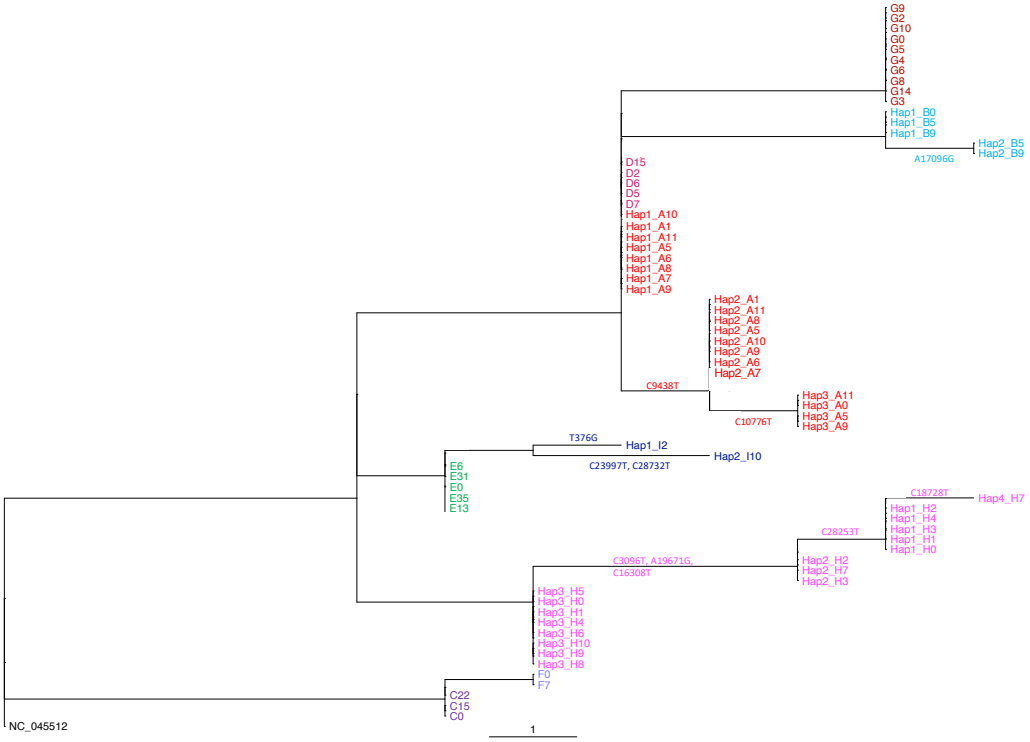


Figure 25. Phylogenetic tree of haplotype sequences from Patients A, B, H, and I and consensus sequences from Patients D-G for which no haplotypes are identified. *Haplotypes defining mutations are shown along the corresponding branches. Samples are labelled as Hap [number]_[Patient][Time (days)].*

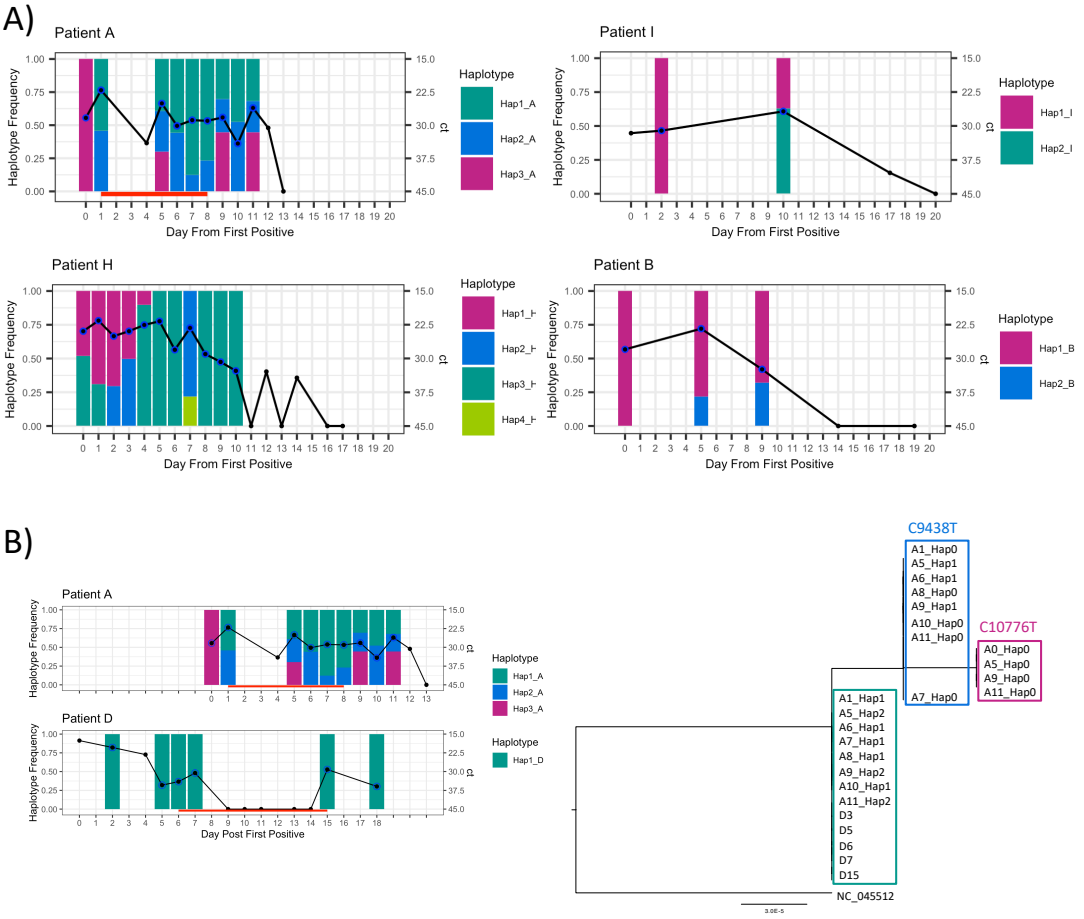


Figure 26. Frequency of identified haplotype over time for individual patients. A) *Haplotype frequency over time for patients A, B, H and I.* ***B)*** *On the right: phylogenetic tree of haplotypes from patients A and D, nucleotide mutation shown above each cluster. On the left: Haplotype frequency over time for patients A and D. Black line is CT value, red line indicates remdesivir received, black dot is sample taken, blue circle indicates sample successfully sequenced. Bars indicate frequency of identified haplotypes.*

To investigate whether the multiple haplotypes or minority variant genomes found within these four patients were due to co-infections or within-host evolution, we used Local Lineage and Monophyly Assessment (LLAMA) (O'Toole, Aine, 2020) to identify samples with the

highest similarity to each haplotype in the global alignment provided by the COVID-19 Genomics UK (COG-UK) consortium (COVID-19 Genomics UK (COG-UK) consortium, 2020) (Figure 27). We found that haplotypes 1 and 2 from patient A; haplotype 3 from patient H; haplotype 1 from patient I; as well as the single genotypes from patients C, E and F were circulating independently in the UK (Figure 27). However, the haplotypes 3 from patient A; 1 and 2 from patient B; 1 and 2 from patient H; as well as the single genotype from patient G was not found in the global alignment with over 60,000 sequences at the time of the search (August 2021), which might reflect incomplete sampling in the population. The clustering of the samples by patient suggests the variations and haplotypes found in each patient were likely due to within-host evolution rather than co-infections with multiple distinct viruses (Lythgoe et al., 2020).

The possible exception to this was patient A. The haplotype 1 from patient A was identical to the virus found in patient D. Given patients A and D were both staying in the intensive care unit at the same time (Figure 18), this could possibly be explained by a superinfection. The haplotype 1 was not detected in the first sequencing sample taken from patient A. It was only detected after patient A has transferred to the intensive care unit, 28 hours after the first sample was taken, and two hours before the start of treatment. However, upon investigation with the infection control team at Great Ormond Street Hospital, we found that no other healthcare associated transmissions have been reported and the local epidemiological set-up in the intensive care unit suggests it is unlikely that a transmission has occurred between patient A and patient D. Another explanation is that haplotypes 1 and 2 from patient A were present in the first sample but remained undetected due to uneven sampling.

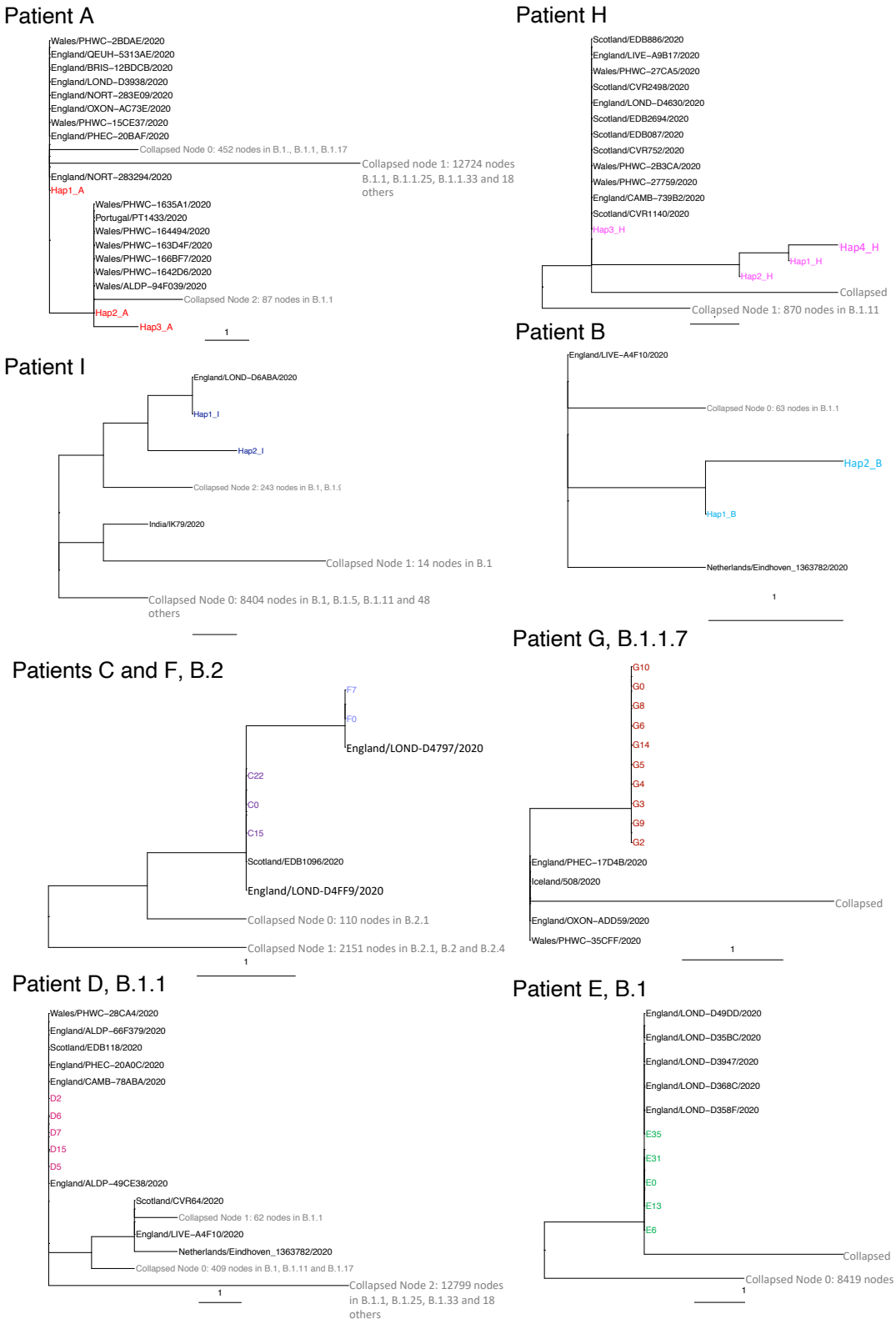


Figure 27. Comparison of haplotypes from patient A, B, H, and I and consensus of C, D, E, G with global sequences. *Trees are grouped by corresponding lineages for patients C, D, E, F, and G.*

4.3.8 Within host variations in nitazoxanide and remdesivir-treated SARS-CoV-2 patients

As reported in other lung infections (Lumby et al., 2020b; Strydom et al., 2019), the heterogeneous response to remdesivir could be due to the poor drug penetration caused by tissue compartmentalisation. It has been suggested that using combination therapies will provide optimal clinical improvements in SARS-CoV-2 infections (Akinbolade et al., 2022). As nitazoxanide has demonstrated synergy with remdesivir, two patients admitted to Great Ormond Street Hospital in 2021 (after the above remdesivir monotherapy study was conducted) were treated with the combination of these two antiviral drugs.

First patient

To understand the effect of combination therapy, we studied two patients who received dual therapy of remdesivir and nitazoxanide. The first patient is a five-year-old girl who developed severe pneumonitis following SARS-CoV-2 infection. The clinical details have been published in (Sanchez Clemente et al., 2021). In summary, she had a critical SARS-CoV-2 infection which required oxygenation and maximum intensive care. During the remdesivir and nitazoxanide dual therapy treatment period, a clear suppression in viral load was observed (Figure 28). Although the viral load rebounded following treatment cessation, the patient has improved clinically and has made a full recovery.

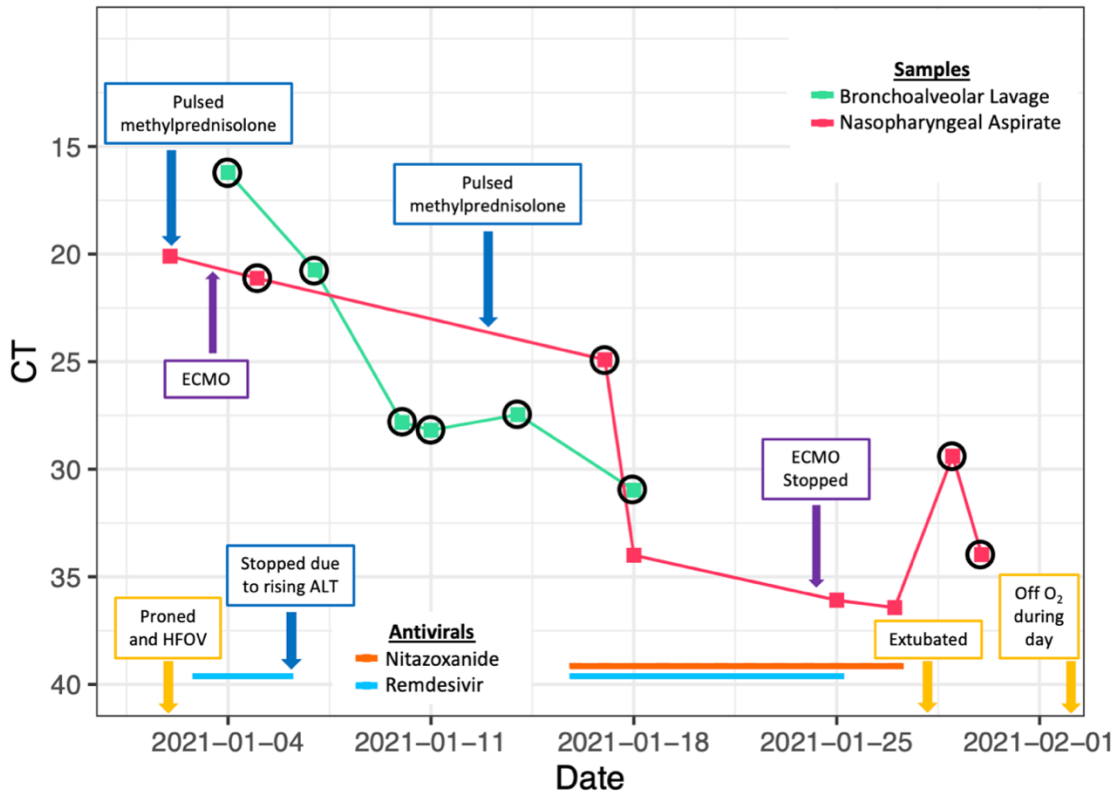


Figure 28. Line graph of CT value trajectories with details of treatments received. *Y-axis = inverted CT values, X-axis = date. Lines are coloured by sample type, green = bronchoalveolar lavage samples, pink = nasopharyngeal aspirate samples. Antiviral received is indicated as coloured bars at the bottom of the graph, orange = nitazoxanide, blue = remdesivir.*

By performing viral deep sequencing on the bronchoalveolar lavage (BAL) and nasopharyngeal aspirate (NPA) samples, we found that the variant profile of samples from the two sites are different, which support our theory of compartmentalisation (Figure 29). None of the variants is known to associate with antiviral treatment. Despite a clear clinical improvement and a decrease in viral load, we found no significant changes in the level of sgRNA (Figure 30). Since the viral load was suppressed during the treatment period, no viral genome sequences have been successfully obtained. Therefore, no sgRNA counts were available during the treatment period.

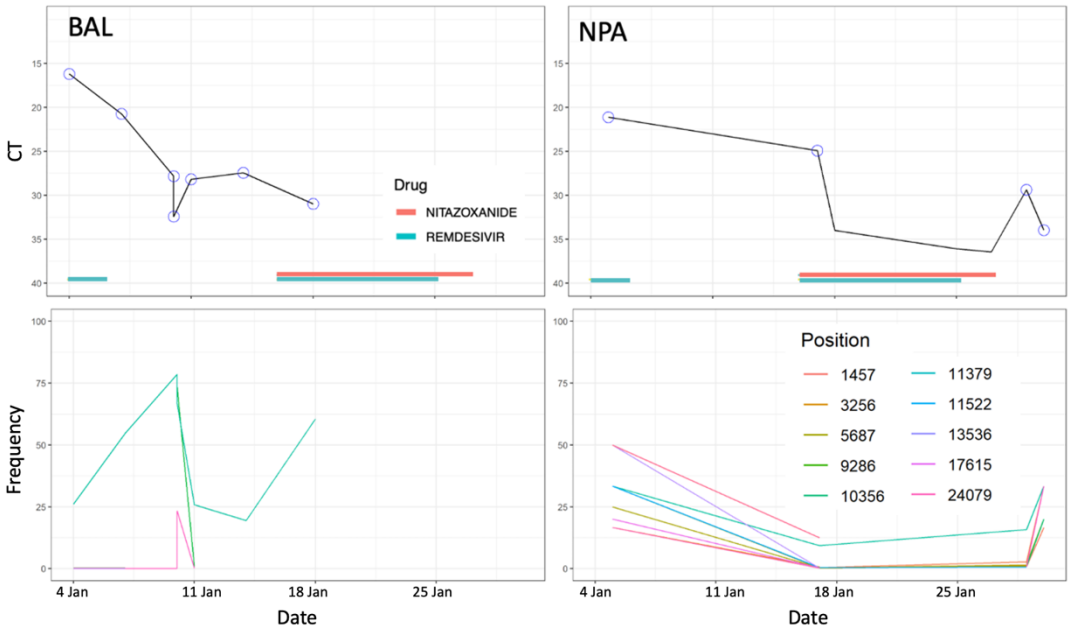


Figure 29. Comparison of variant profiles of BAL and NPA samples. *The top panels show the CT values. The treatment periods are indicated by red (nitazoxanide) and teal (remdesivir) bars. The bottom panels show the changes in variant frequencies over time.*

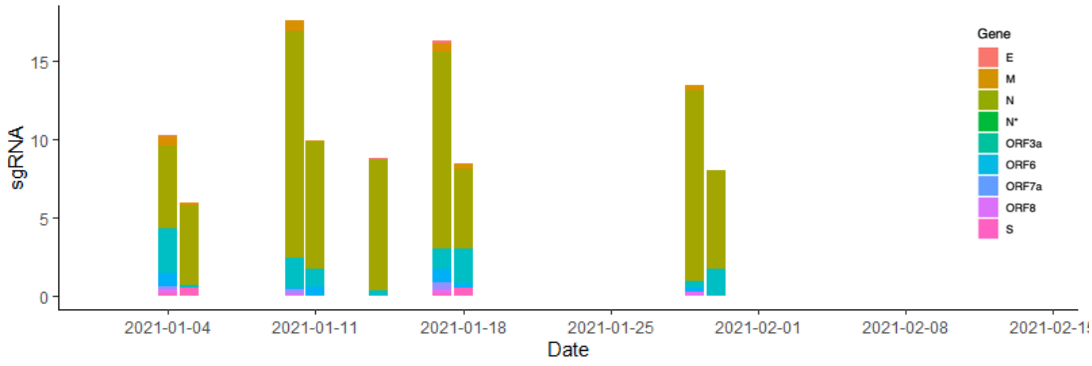


Figure 30. Stacked bar charts showing the number of sgRNA reads per 100,000 mapped reads (sgRPHT). *Y-axis = counts, x-axis = samples collected over time. The bars are coloured by the gene of the sgRNA.*

Second patient

The second patient is a 4-year-old boy who was transferred to Great Ormond Street Hospital following worsening respiratory distress and increasing oxygen requirement. Despite showing no significant changes (p-value > 0.1) in the CT value or sgRNA level (Figure 31) following treatments, he improved clinically and was discharged after a month since he first tested positive for SARS-CoV-2.

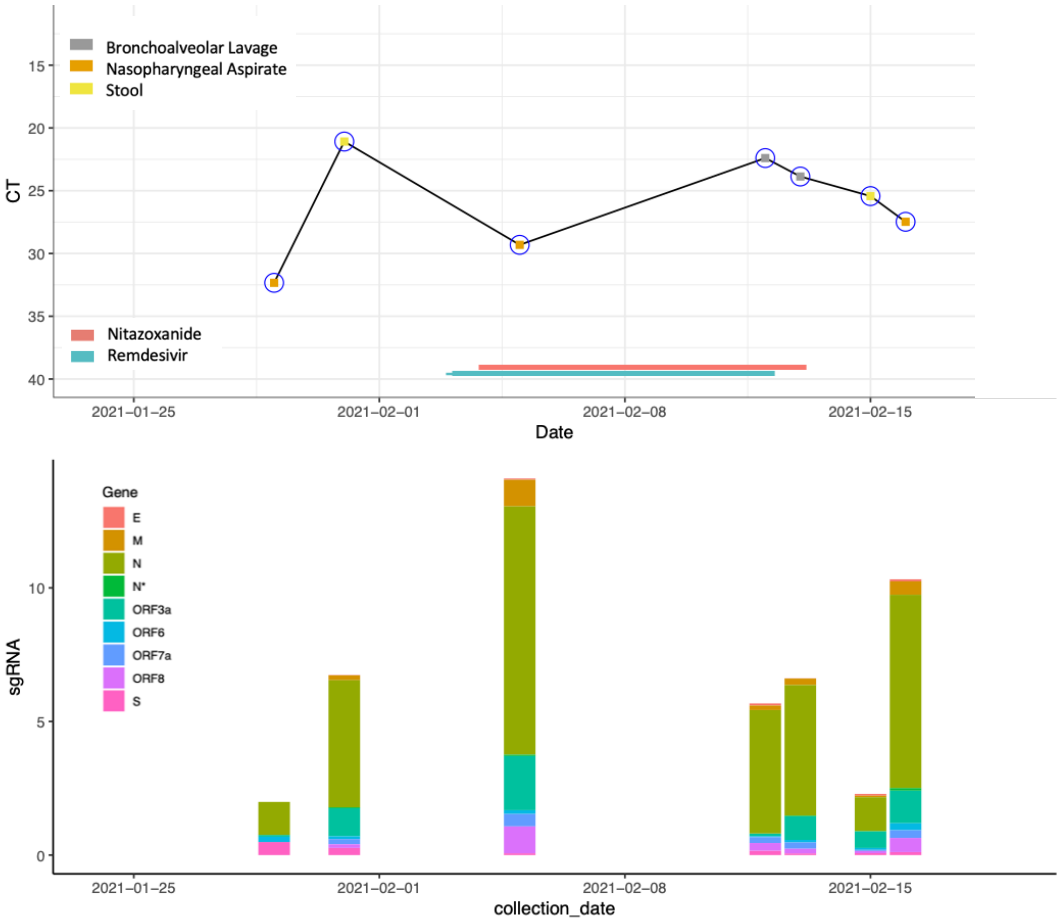


Figure 31. CT values trajectory (top panel) and the number of sgRNA reads per 100,000 mapped reads (sgRPHT) (bottom panel). *In the CT values panel, the data points are coloured by the site of sample collection (BAL, NPA or stool). The treatment periods are indicated with the bars below the CT trajectory, red = nitazoxanide, teal = remdesivir.*

Seven samples were collected and sequenced from this patient. We found that the samples clustered by collection site in the phylogenetic tree (Figure 32). It is interesting to note that although the 6th sample (stool) and the 7th sample (NPA) were collected on the same day, they fell into two separate clusters on the phylogenetic tree (Figure 32). However, all variants identified between the different compartments were synonymous.

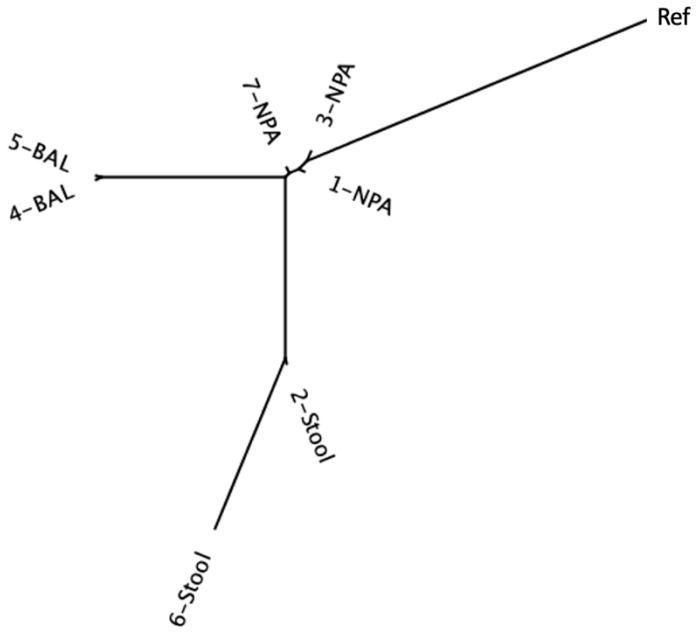


Figure 32. Phylogenetic tree of all samples collected from the 4-year-old male SARS-CoV-2 patients. *Samples are labelled as [sample number]-[site of collection]. NPA = nasopharyngeal aspirate, BAL = bronchoalveolar lavage samples.*

4.3.9 *In vivo* activity of favipiravir and molnupiravir in SARS-CoV-2 infected Syrian hamsters

To further understand the *in vivo* antiviral activity of the two RdRp inhibitors, favpiravir and molnupiravir, we collaborated with Dr Joana Duarte Da Rocha Pereira and Dr Rana Abdelnabi from Katholieke Universiteit Leuven to study the efficacy and the mutagenic effect of these two drugs when used to treat SARS-CoV-2 in Syrian hamsters. Details of the studies have been published in (Abdelnabi et al., 2021; Kaptein et al., 2020a).

In summary, we found that both favipiravir and molnupiravir can reduce viral load in hamsters, in particular the combination therapy of the two drugs led to a reduction of viral load by 5log10 (Abdelnabi et al., 2021). By performing deep sequencing, we found that hamsters which received favipiravir treatment showed an average of 3.2-fold increase in mutational burden (number of variants) compared to untreated hamsters (Figure 35). We also confirmed that both favipiravir and molnupiravir induced dose-dependent C to T and G to A mutagenesis in SARS-CoV-2 (Figure 36). We found that molnupiravir, even at lower dose, induced a higher level of mutational burden and mutagenesis compared to favipiravir (Figures 35 and 36).

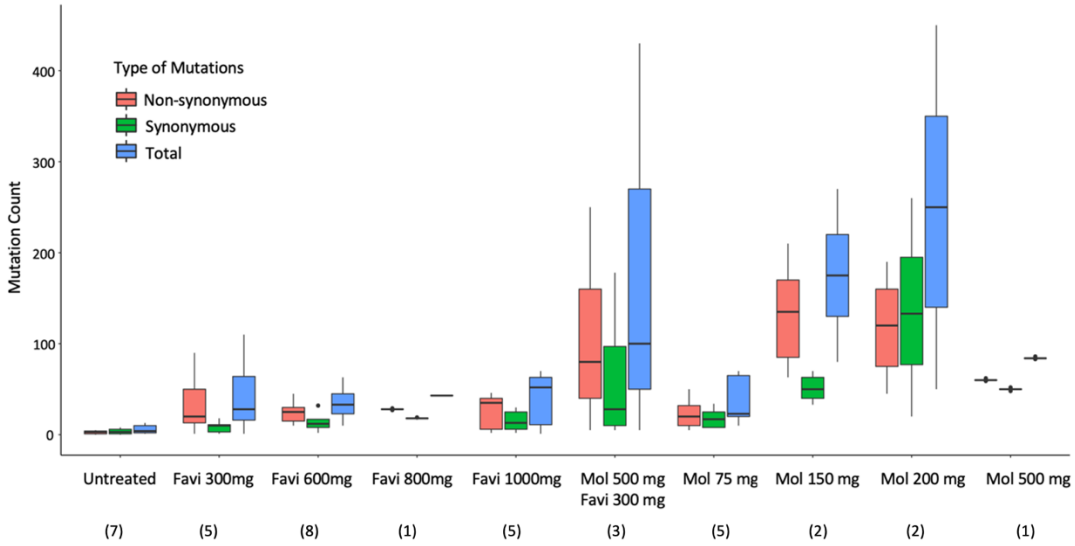


Figure 35. Box plots showing the mutational burden caused by different doses of favipiravir or molnupiravir treatment. *The mutational burden is represented as the adjusted number of polymorphic sites on the y-axis. On the x-axis, each column represents a group of hamsters which are untreated (first column on the left), or treated by favipiravir (300, 600, 800 or 1000 mg), molnupiravir (75, 150, 200, or 500 mg) or a combination of both (150 mg of molnupiravir + 300 mg of favipiravir). The bars are coloured by the type of mutations, red = non-synonymous mutations, green = synonymous, blue = total. Numbers shown in brackets indicate the sample size.*

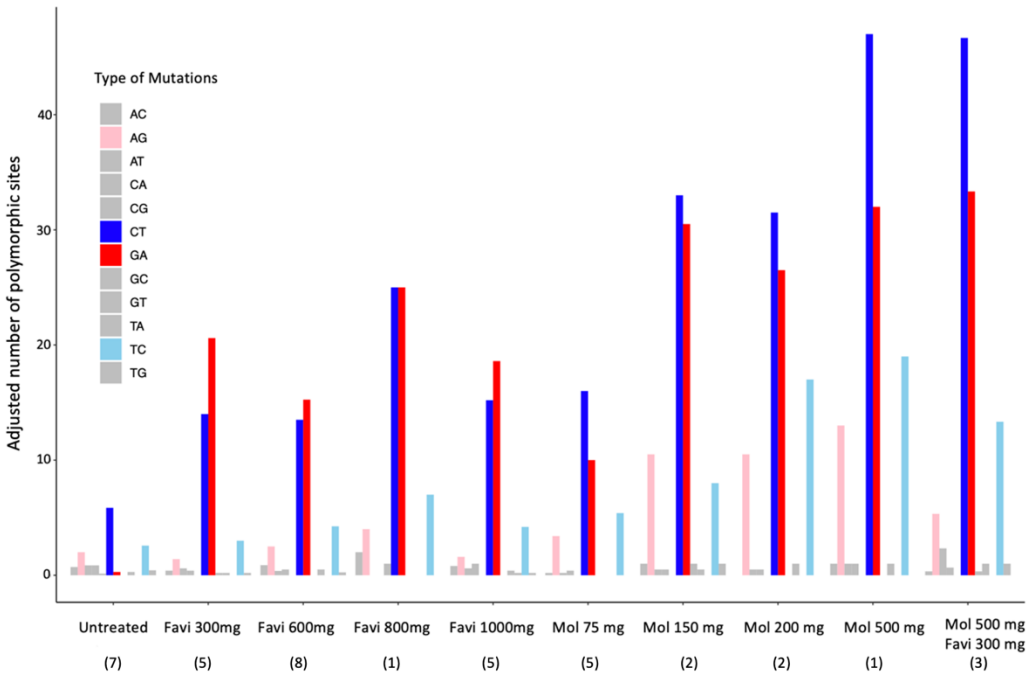


Figure 36. Bar charts showing the normalised number of transitions and transversions induced by different doses of favipiravir or molnupiravir treatment. *Y-axis = normalised count of different type of transition and transversion compared to the virus used for inoculation. X-axis = different groups of hamsters which are either untreated, or treated with different doses of favipiravir or molnupiravir, or a combination therapy of both. The two mutagenic signature mutations (C to T and G to A) are highted in blue and red. Numbers shown in brackets indicate the sample size.*

4.3.10 Discussions on SARS-CoV-2 Analysis

Summary

In Sections 4.3.7 and 4.3.8, we analysed SARS-CoV-2 patients treated with remdesivir monotherapy and nitazoxanide and remdesivir dual therapy. In Section 4.3.9, we presented evidence of dose-dependent mutagenic signature induced by favipiravir and molnupiravir.

By comparing longitudinal samples from six untreated and three remdesivir treated paediatric patients with COVID-19, we found evidence of remdesivir-associated suppression of SARS-CoV-2 subgenomic RNA. We identified the presence of multiple distinct viral haplotypes in four out of nine patients, which have likely emerged through within-host evolution during the early stage of infection and persisted possibly in different compartments in the lung. To alleviate the drug penetration issue possibly caused by tissue compartmentalisation, we combined the use of nitazoxanide and remdesivir in two patients. We demonstrated the clinical efficacy of the combination therapy.

Similar to the norovirus study, the major limitation of this study is that we only have a small cohort of nine patients with heterogeneity patterns in both clinical and virological parameters, which limits some of the conclusions being drawn. In addition, all individuals in our cohort are paediatric patients, for which their clinical characteristics and progression of SARS-CoV-2 are known to differ from adults (Du et al., 2020). However, various studies have described similar patterns of virological and clinical response in remdesivir treated adult patients (Beigel et al., 2020b; Gastine et al., 2021; Y. Wang et al., 2020). Similar

clinical response patterns in adults and children to repurposed antiviral drugs for treating other RNA viral infections have also been described in various studies (Lumby et al., 2020b; Strydom et al., 2019).

Efficacy of remdesivir

In Section 4.3.7, while all three treated patients in our cohort showed clinical improvement, suppression of viral load was only observed in patient D, where the viral RNA remained undetectable during the treatment period and rebounded following treatment cessation. Our results in this chapter mirror findings in norovirus described in section 4.3.3, where the patients showed drug-associated clinical improvement without showing any changes to the viral load. A study on remdesivir treatment in a macaque model also showed that despite no significant changes to viral load, clinical improvement and a decrease in the viability of SARS-CoV-2 in *in vitro* culture were observed following treatment (B. N. Williamson et al., 2020).

In patient A, despite seeing no significant changes in the viral load during the treatment period, the sgRNA levels have reduced during treatment and rebounded after the treatment was stopped. Some samples with good coverage and mean read dept taken during treatment had low sgRNA levels even when the viral load was high (Figure 23B). One study reported that sgRNA levels have a weak association with viral replication in *in vitro* culture, but it remains controversial whether sgRNA is a better indicator of viable virus compared to viral load obtained through PCR (Alexandersen et al., 2020a; Parker et al., 2021). Another study suggested subgenomic RNAs can be detected in diagnostic samples up to 17 days after the initial detection of infection, which suggests subgenomic RNAs might not be a suitable indicator of active viral

replication (Alexandersen et al., 2020b). However, the potential of combining sgRNA count with viral load data together as a biomarker for understanding the drug effect should be further explored.

Within-host minority variants analysis

In Section 4.3.7, we observed consensus level substitutions in three out of nine patients (one treated with remdesivir, two untreated) which can be explained by the presence of multiple distinct viral haplotypes (between two to four) that vary in abundance (Figure 26A). Interesting, while remdesivir suppressed the viral replication in patient D, leading to the viral load falling below the detection limit during the treatment period, the same dose of remdesivir had little effect on the identical viral strain, haplotype 1 from patient A (Figure 26B). One possible explanation for this is tissue compartmentalisation, where the pathogen replicates in distinct niches in different area of the lung (Turner et al., 2014). Viruses could evolve separately in these physically separated tissue compartments and mutations can potentially accumulate independently which would allow us to distinguish populations in different compartments via viral deep sequencing. The observation of sample-to-sample variations in frequencies of multiple haplotypes supports this theory. When poorly mixed viral populations exist, clinical samples often fail to capture the full diversity. As discussed in Chapter 2, this could contribute to the high day-to-day variations in viral load levels we observed. A study on post-mortem sampling of the lung also confirmed SARS-CoV-2 viral loads vary in different parts of the lung (Desai et al., 2020).

Compartmentalisation and the effect on antiviral drugs

Tissue compartmentalisation has been observed in other respiratory tract infections including influenza and *Mycobacterium tuberculosis*, and has been associated with poor drug penetration, reducing the effectiveness of drug and potentially leading to the emergence of drug resistance in some niches (Ks et al., 2018; Lumby et al., 2020b; Strydom et al., 2019). A study which modelled the remdesivir drug levels within different lung tissue in COVID-19 patients also suggested that remdesivir has poor penetration into lung tissue (Wang and Chen, 2020). In patient A, the persistence of distinct viral strains at different frequencies over time supports the theory of tissue compartmentalisation, which could explain the reduced effect of remdesivir on the same viral strain which was also found and suppressed in patient D. The rebound of viral sgRNA in patient A and viral load in patient D following treatment cessation also suggests the duration of remdesivir treatment used was suboptimal. Pharmacokinetic-pharmacodynamic (PKPD) models of SARS-CoV-2 viral dynamics predicts that remdesivir has an effectiveness of 87% in inhibiting viral replication while a drug with greater than 90% of effectiveness is required to completely interrupt the viral replication and successfully clear the infection in patients (Gonçalves et al., 2020). Considering the poor drug penetration and the predicted high drug effectiveness required for clearing the infections, early treatment with combination therapy might be required for an observable antiviral effect. Although most clinical trials focus on a single drug, combination therapies are now frequently used for treating SARS-CoV-2 infections as they tend to optimise the potency of each drug compound and reduce the chance of developing drug resistance (Akinbolade et al., 2022).

Combination therapy

Our data in section 4.3.7 suggested that the within-host variations observed in SARS-CoV-2 were likely due to viral compartmentalisation within different lung tissues, which could be the major cause of poor drug penetration. Therefore, we combined the use of remdesivir and nitazoxanide in the patients described in section 4.3.8. Both patients showed significant clinical improvement during and after the combination therapy of nitazoxanide and remdesivir. In the first patient, the viral load was suppressed during the treatment, which is similar to what we observed in the remdesivir treated patient D in Section 4.3.7. In the second patient, no changes in viral load were observed. Based on our data, it is unclear whether the use of a second antiviral agent provided any additional benefits. However, in the phylogenetic tree of samples collected from the second patient, we showed that the samples cluster by site instead of sample collection date. This provides evidence of the presence of viral compartmentalisation.

Mutagenic signature in Syrian hamsters

In Section 3.3.9, we found that the two RdRp inhibitors, favipiravir and molnupiravir, can induce dose-dependent drug-associated mutagenic signature in Syrian hamsters. Even at a lower dose, molnupiravir can cause more mutations (Figure 35) and induce a stronger mutagenic signature compared to favipiriavir (Figure 36). A combination therapy of 150 mg molnupiravir and 300 mg favipiravir produced a similar level of mutagenic effect as 500 mg of molnupiravir monotherapy. Since mutagenesis can lead to a reduction in fitness, the use of high dose molnupiravir monotherapy or favipiravir and molnupiravir combination therapy should be explored.

4.4 Closing Remarks

Drug Efficacy

In Chapter 1, we discussed that some clinical trials on RdRp inhibitors including favipiravir and molnupiravir showed clinical benefits for COVID-19 patients. However, such clinical trials are not available for norovirus. In clinical trials, a large number of patients are recruited and the patients' demographics and comorbidities are evaluated. Yet, in clinical case series or case reports, only a small number of patients are included and they can have very different demographics and comorbidities. In a small cohort, it is difficult to measure and distinguish the effect of antiviral treatments from chance. For most repurposed drug, there is no clear guidance on the optimal dosage or treatment duration. Treatments and dosage are primarily selected based on previous clinical experiences or clinical reports. With problems such as the insufficiency of dose or duration of treatment, it is difficult to make comparisons across different studies and understand the mixed efficacy of antiviral drugs.

In this chapter, we matched the virological response to the clinical efficacy by analysing the deep sequencing variants and reconstructing haplotypes in a novel way. We identified a drug-induced increase in mutagenic signature in favipiravir-treated norovirus patients and a dose-dependent increase in mutagenic signature in favipiravir and molnupiravir-treated SARS-CoV-2 infected hamsters. For the first time, we presented *in vivo* evidence of a drug-associated loss of infectivity of norovirus in a zebrafish larvae model. We have also identified putative drug-associated mutations induced by favipiravir in norovirus patients,

but we found no evidence that this interferes with the clinical efficacy of the drug.

In our analysis, although nearly all patients did not show an antiviral-associated reduction in viral load, we observed rapid and significant reduction of the disease burden in most patients following the use of treatment. In norovirus, compared to untreated patients, with the presence of drug pressure, more mutations were observed in treated patients. However, apart from some mutations described in Section 4.3.5, we found no potential drug-resistant conferring mutations.

In Sections 4.3.7 and 4.3.8, we described three remdesivir-treated, and two nitazoxanide and remdesivir combination therapy treated SARS-CoV-2 patients who improved clinically. Despite no changes in viral load was observed in most patients, two patients showed suppression of viral load during treatment, followed by a rebound in viral load following cessation of treatment. This rebound of viral load could be due to the suboptimal dose or duration of treatment used.

Although combination therapies, such as the use of ribavirin and nitazoxanide, have been reported to show clinical efficacy in treating norovirus, in the cohort of patients we presented in Section 4.3.2, no significant clinical or virological changes were observed. However, more than half of the patients in the cohort had only received treatment for less than a week due to suspected toxicity. The limited duration of treatment could explain the lack of efficacy observed in the cohort. In the two SARS-CoV-2 patients described in Section 4.3.8, despite observing clinical improvements following the use of remdesivir and nitazoxanide combination therapy, it is unclear whether the use of a second antiviral agent provided any additional benefits.

Mutagenesis

It has been suggested that non-lethal mutagenesis (what we observed) induced by nucleoside analogues can be a potential threat to the emergence of drug resistance (Nelson and Otto, 2021; Sadler et al., 2010). We found mutagenic signature in both favipiravir treated norovirus patients and favipiravir and molnupiravir treated SARS-CoV-2 infected hamsters. Our study in norovirus showed that prolonged use of favpiravir indeed resulted in multiple mutations in the RdRp, which are not seen in the untreated patients. In the norovirus P3 described in section 4.3.4, of the six putative favipiravir-associated amino acid changes, three were indeed caused by nucleotide changes which are associated with the drug-induced mutagenic signature (A to G or T to C). However, it is important to note that even before the start of the treatment, P3 has a higher number of A to G and T to C polymorphisms compared to other patients. In P5, of the eight putative drug-associated mutations in the RdRp, only one was caused by a drug-induced mutagenic signature (A to G). In addition, we showed that when treatment was paused in P3, the non-resistant genome again become the dominant haplotype in the viral population, which suggests the genomes carrying drug-associated mutations tend to be less fit. The virus is less likely to be transmitted, even if drug-resistant variants were generated in the process of lethal mutagenesis. However, if compensatory mutations arise, the fitness of the virus could be significantly improved. These drug-resistant viruses could also be selected in the community if the drug is widely used.

It is difficult to conclude whether the use of mutagenic antiviral drugs would lead to the emergence of drug resistance. As the mutation rate increases, the probability of getting drug-resistant mutations also

increases. Given the large number of COVID-19 infections, the widespread use of mutagenic drug could increase the probability of generating drug-resistant mutations in the community. However, the drug could also increase the background selection by inducing deleterious mutations in other parts of the resistant genome, which reduces its ability to grow and dominate. More study is needed to conclude whether the use of mutagenic drugs such as favipiravir and molnupirivir is beneficial to public health.

Chapter 5

Conclusions and Future Directions

5.1 Application to other RNA viruses

Summary

RNA viruses are responsible for numerous pandemics and epidemics throughout human history (Carrasco-Hernandez et al., 2017; Piret and Boivin, 2021). They are also responsible for non-human diseases which lead to mortality of livestock and crops (Maksimov et al., 2019; Tomley and Shirley, 2009). RNA viruses both directly and indirectly exert a significant cost on the global economy and public health (Bartsch et al., 2020; Courville et al., 2022; Kolahchi et al., 2021; Putri et al., 2018; Richards et al., 2022). Although the high mutation rate of RNA viruses poses a risk to the human population, they allow us to study the evolutionary processes on a shorter timescale (Moya et al., 2000). SARS-CoV-2, which causes the current COVID-19 pandemic, provided a unique opportunity for the scientific community to collaboratively develop analysis methods and treatments which could potentially be used in other RNA viruses (Chacón-Labella et al., 2021).

In this thesis, we demonstrated the use of longitudinal deep sequencing data in addition to clinical data to study and model the intra host evolutionary and population genetics dynamics. In Chapter 2, we explored the viral load dynamics in acute infections. In Chapter 3 and 4,

we used standard measures of sequence variation to monitor the response of viral population to the antiviral treatment and how it is correlated and connected to other patient data. We also demonstrated how statistical models such as HaROLD, described in Chapter 3, can be used to solve more complex problems such as reconstructing accurate haplotypes in mixed infections and subsequently identifying the amino acid substitutions and the selective forces, including antiviral drugs and host immune responses, that are driving the mutations. In Chapter 4, we investigated the viral variations in untreated and treated patients, and animal models. For the first time, we provided evidence that the apparent clinical efficacy is linked to a reduction in viral infectivity. In the absence of randomised clinical trials, we provided preliminary data that antiviral drugs such as favipiravir can be effective and the clinical improvements observed were not only due to chance. Our data support the use of combination treatments and our results can be further validated in clinical trials. The methods presented can be applied to other clinical data sets.

Norovirus and SARS-CoV-2

The two viruses we studied in this thesis, norovirus and SARS-CoV-2, are both positive-sense single-stranded RNA virus. They infect different cell types, but they are both highly contagious. Norovirus has caused many epidemics and SARS-CoV-2 has caused one of the largest pandemics in human history. As RNA viruses, norovirus and SARS-CoV-2 share many similarities. In particular, they have a similar viral replication cycle and have a polymerase with similar structure (Deval et al., 2017; Gao et al., 2020).

The RdRp plays an important role in the RNA viral genome replication, and it has a conserved structure which allows broad-spectrum antiviral drugs to be developed, making it a key therapeutic target for most RNA viruses (Picarazzi et al., 2020). In addition, human cells, as well as other mammalian cells do not have an RdRp or proteins with similar function (Lai, 2005). The inhibition of RdRp is unlikely to cause target-related side effects (Tian et al., 2021). Therefore, RdRp has been a major target in antiviral drug discovery and development.

For SARS-CoV-2, given a large number of patients, hundreds of antiviral drugs have been tested in clinical trials and case studies (Chen et al., 2021). It is the most extensive approach which could guarantee finding the best available treatment. However, such an approach is not always feasible. For other RNA viruses, it is difficult to find a large cohort of patients, and even if there are enough patients, carrying out clinical trials is extremely costly (Martin et al., 2017). Therefore, only antiviral drugs with the best potential could be tested.

Antiviral Drugs for positive-sense single-stranded RNA Viruses

Norovirus and SARS-CoV-2 are both positive-sense single-stranded RNA virus. Before the COVID-19 pandemic, the only positive-sense single-stranded RNA virus which has an effective RdRp inhibiting antiviral is hepatitis C virus (Buonaguro and Buonaguro, 2020). Sofosbuvir is one of the most frequently used antiviral drugs which directly inhibits the hepatitis C virus NS5B RdRp protein by acting as a chain terminator (Koff, 2014). It is mostly used in combination therapy with other antivirals including ribavirin (World Health Organization, 2018).

All RdRp inhibitors we studied in this thesis were developed for negative-sense single-stranded RNA viruses such as influenza and Ebola (Baranovich et al., 2013; Furuta et al., 2013; Toots et al., 2019). Since the RdRp of positive and negative-sense RNA viruses interact with the metal ions slightly differently during catalysis (Chapter 1, Section 1.1.4) (te Velthuis, 2014), an antiviral drug developed for positive-sense single-stranded RNA virus might be the optimal option for treating norovirus and SARS-CoV-2.

Some studies have already shown that sofosbuvir is effective against SARS-CoV-2 (Abbass et al., 2021; Jácome et al., 2020; Jockusch et al., 2020; Zein et al., 2021) as well as other unrelated positive-sense RNA viruses including enterovirus (Sun et al., 2020), yellow fever (Mendes et al., 2019) and hepatitis A virus (Jiang et al., 2018).

However, sofosbuvir is priced at $1000 USD per day, which is the major barrier to accessing the drug. It has never been tested in animal models or human for treating norovirus.

Assessing drug efficacy

We discussed that despite showing an *in vitro* antiviral activity, for many antiviral treatments, there is a lack of virological efficacy in patients. Viral load, the main clinical marker, which is monitored regularly, often show no reduction following the use of antiviral treatments. The lack of impact on viral load makes the evaluation of treatment difficult, which leads to uncertainty about the efficacy of these drugs, even when clinical improvement occurs. Studies have reported the lack of reduction in viral load could be due to the suboptimal dose. However, toxicity has also been reported in patients who received a higher dose.

In Chapter 1, Section 1.1.5 and Chapter 4, we discussed lethal mutagenesis being the key mechanism of action of some RdRp inhibitors such as favipiravir and molnupiravir. In line with previous studies (Baranovich et al., 2013; de Ávila et al., 2016; Goldhill et al., 2019; Ruis et al., 2018b), we identified drug-associated mutagenic signatures in viral genome sequences, even when no virological response was observed. As routine viral whole genome sequencing has become more popular in clinical settings, this important marker should be further explored. A mutagenic signature or a normalised mutagenic level threshold could potentially be used to monitor the individual response to drug and evaluate the drug efficacy. This would provide support to a more personalised approach to antiviral treatment on a regular basis.

5.2 Potential use of simulation models to study intra-host variations

Limitations of clinical studies

One of the main barriers to study the viral variations is the lack of good quality and consistent data. Human challenge studies and clinical trials are extremely costly. For COVID-19, researchers from all over the world come together to study this single virus and tackle the global pandemic with collaborative effort. Numerous SARS-CoV-2 clinical trials have been successfully carried out. They provided a better picture of SARS-CoV-2 viral dynamics which is crucial for the development of effective treatments. However, the scientific community might not share the same priority and interest in other viruses and infectious diseases. Studies available for norovirus for example, are mainly observational reports published on a case-by-case basis. Similar to many RNA viral infections, norovirus infections happen sporadically and clear within days in most individuals. The amount of data available is incomparable to the current COVID-19 pandemic. Chronically infected patients can be infected for years and provide plenty of clinical and virological data. However, these patients have other comorbidities which pose a challenge in aligning the data across multiple patients. The selective pressures and viral dynamics could also be very dissimilar in immunocompromised patients.

Throughout the thesis, we demonstrated the use of viral load as a clinical metric to monitor disease progression. We illustrated the sample-to-sample variability of viral load, potentially due to the inconsistent sampling or storage techniques or uneven sampling caused by viral compartmentalisation. In addition, other than in human challenge

studies and clinical trials, it is difficult to capture the viral dynamics of the full course of infection or determine the point of infection. When patients are present with symptoms, the incubation period has already passed, meaning we will no longer be able to capture the viral variations in the initial period of infection. We have shown that viral load can be a challenging metric to assess infections. Combining viral load data with deep sequencing data will give a better picture of the state of the infections.

Statistical modelling

Some studies use mathematical and statistical models to solve this complex problem. However, these methods are mostly developed for the specific data sets, with only a limited number of clinical cases. They might not be applicable to other data sets, and most have not been validated on any benchmark data sets. With limited replicability, conclusions from different studies can have contradicting results. Important questions such as the viral mutation rate, population dynamics, compartmentalisation remain unanswered (Tisthammer et al., 2020; Zanini et al., 2017). The lack of generalised and reliable mathematical models hinders our understanding of viral dynamics and evolution. With no sufficient data, methods cannot be tested or validated.

Wright-Fisher model simulations on within-host viral population could provide useful data and insights into evolutionary dynamics. The Wright-Fisher model is frequently used in the field of population genetics and evolutionary biology. It describes a population where all members of a generation reproduce and die, where each generation is completely replaced by the offspring from the previous generation. It can be used to study how allele frequency changes over time. A modified version of

Wright-Fisher model has potential for the study of within-host viral population.

Preliminary work conducted by Professor Richard A. Goldstein, Professor David D. Pollock and Dr. Asif U. Tamuri has demonstrated the applications of the Wright-Fisher model (Goldstein and Pollock, 2017; Pollock et al., 2012). These models can produce generations of viral genomic sequences based on an initial input sequence of interest. When applied to virology, the fitness of virus, determined by the protein folding entropy, can also be considered. In simulation programs, every parameter involved in the evolutionary processes such as mutation rate, population size, and sampling frequency can be adjusted. The data generated from the simulations can be used as a benchmark for method validation. Future work on developing a similar type of Wright-Fisher simulation for studying with-host viral genomic sequence variations would be highly valuable.

Machine learning

In recent years machine learning and neural network methods have been applied to a wide range of studies in genomics. They can accurately separate noise and error from actual data and uncover patterns in large complex clinical and genomics data sets. However, a good training data set is required for these models to achieve a good performance and give accurate predictions. Simulation data described above can be employed for training models.

A machine learning or neural network model that can combine clinical and genomic data to identify the approximate time of infection or predict the treatment outcome based on the diversity or variant frequency

spectrum in the genome sequence would be highly valuable. It would provide reliable metrics for decision-making and support a more personalised approach to antiviral treatment. However, it is essential for the users of these algorithms to have basic theoretical and practical knowledge of machine learning to avoid the misinterpretation of results.

References

Abbass, S., Kamal, E., Salama, M., Salman, T., Sabry, A., Abdel-Razek, W., Helmy, S., Abdelgwad, A., Sakr, N., Elgazzar, M., et al., 2021. Efficacy and safety of sofosbuvir plus daclatasvir or ravidasvir in patients with COVID-19: A randomized controlled trial. Journal of Medical Virology 93, 6750–6759. https://doi.org/10.1002/jmv.27264

Abdelnabi, R., Foo, C.S., Kaptein, S.J.F., Zhang, X., Do, T.N.D., Langendries, L., Vangeel, L., Breuer, J., Pang, J., Williams, R., et al., 2021. The combined treatment of Molnupiravir and Favipiravir results in a potentiation of antiviral efficacy in a SARS-CoV-2 hamster infection model. eBioMedicine 72, 103595. https://doi.org/10.1016/j.ebiom.2021.103595

Abdullah, F., Myers, J., Basu, D., Tintinger, G., Ueckermann, V., Mathebula, M., Ramlall, R., Spoor, S., de Villiers, T., Van der Walt, Z., et al., 2022. Decreased severity of disease during the first global omicron variant covid-19 outbreak in a large hospital in tshwane, south africa. International Journal of Infectious Diseases 116, 38–42. https://doi.org/10.1016/j.ijid.2021.12.357

Agostini, M.L., Pruijssers, A.J., Chappell, J.D., Gribble, J., Lu, X., Andres, E.L., Bluemling, G.R., Lockwood, M.A., Sheahan, T.P., Sims, A.C., et al., 2019. Small-Molecule Antiviral β- D - N^4 -Hydroxycytidine Inhibits a Proofreading-Intact Coronavirus with a High Genetic Barrier to Resistance. J Virol 93, e01348-19. https://doi.org/10.1128/JVI.01348-19

Akinbolade, S., Coughlan, D., Fairbairn, R., McConkey, G., Powell, H., Ogunbayo, D., Craig, D., 2022. Combination therapies for COVID-19: An overview of the clinical trials landscape. Brit J Clinical Pharma 88, 1590–1597. https://doi.org/10.1111/bcp.15089

Alexandersen, S., Chamings, A., Bhatta, T.R., 2020a. SARS-CoV-2 genomic and subgenomic RNAs in diagnostic samples are not an indicator of active replication. medRxiv 2020.06.01.20119750. https://doi.org/10.1101/2020.06.01.20119750

Alexandersen, S., Chamings, A., Bhatta, T.R., 2020b. SARS-CoV-2 genomic and subgenomic RNAs in diagnostic samples are not an indicator of active replication. Nat Commun 11, 6059. https://doi.org/10.1038/s41467-020-19883-7

Amul, G.G., Ang, M., Kraybill, D., Ong, S.E., Yoong, J., 2022. Responses to COVID-19 in Southeast Asia: Diverse Paths and Ongoing Challenges. Asian Economic Policy Review 17, 90–110. https://doi.org/10.1111/aepr.12362

Andersen, K.G., Rambaut, A., Lipkin, W.I., Holmes, E.C., Garry, R.F., 2020. The proximal origin of SARS-CoV-2. Nat Med 26, 450–452. https://doi.org/10.1038/s41591-020-0820-9

Andrews, 2020. FastQC: A Quality Control Tool for High Throughput Sequence Data.

Anhui Zhifei Longcom Biologic Pharmacy Co., Ltd., 2020. Phase I/IIa Clinical Trial to Evaluate the Safety, Tolerability and Immunogenicity of Longkoma Quadrivalent Recombinant Norovirus Vaccine (Pichia Pastoris) Implanted in

Populations of 6 Weeks and Older (Clinical trial registration No. NCT04563533). clinicaltrials.gov.

Ao, Y., Wang, J., Ling, H., He, Y., Dong, X., Wang, X., Peng, J., Zhang, H., Jin, M., Duan, Z., 2017. Norovirus GII.P16/GII.2-Associated Gastroenteritis, China, 2016. Emerg Infect Dis 23, 1172–1175. https://doi.org/10.3201/eid2307.170034

Aoki, Y., Suto, A., Mizuta, K., Ahiko, T., Osaka, K., Matsuzaki, Y., 2010. Duration of norovirus excretion and the longitudinal course of viral load in norovirus-infected elderly patients. Journal of Hospital Infection 75, 42–46. https://doi.org/10.1016/j.jhin.2009.12.016

Arabi, Y.M., Shalhoub, S., Mandourah, Y., Al-Hameed, F., Al-Omari, A., Al Qasim, E., Jose, J., Alraddadi, B., Almotairi, A., Al Khatib, K., et al., 2020. Ribavirin and Interferon Therapy for Critically Ill Patients With Middle East Respiratory Syndrome: A Multicenter Observational Study. Clinical Infectious Diseases 70, 1837–1844. https://doi.org/10.1093/cid/ciz544

Argyropoulos, K.V., Serrano, A., Hu, J., Black, M., Feng, X., Shen, G., Call, M., Kim, M.J., Lytle, A., Belovarac, B., et al., 2020. Association of Initial Viral Load in Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Patients with Outcome and Symptoms. The American Journal of Pathology 190, 1881–1887. https://doi.org/10.1016/j.ajpath.2020.07.001

Arias, A., Thorne, L., Goodfellow, I., 2014. Favipiravir elicits antiviral mutagenesis during virus replication in vivo. eLife 3, e03679. https://doi.org/10.7554/eLife.03679

Artika, I.M., Dewantari, A.K., Wiyatno, A., 2020. Molecular biology of coronaviruses: current knowledge. Heliyon 6, e04743. https://doi.org/10.1016/j.heliyon.2020.e04743

Ashour, N.A., Abo Elmaaty, A., Sarhan, A.A., Elkaeed, E.B., Moussa, A.M., Erfan, I.A., Al-Karmalawy, A.A., 2022. A Systematic Review of the Global Intervention for SARS-CoV-2 Combating: From Drugs Repurposing to Molnupiravir Approval. DDDT Volume 16, 685–715. https://doi.org/10.2147/DDDT.S354841

Astrovskaya, I., Tork, B., Mangul, S., Westbrooks, K., Măndoiu, I., Balfe, P., Zelikovsky, A., 2011. Inferring viral quasispecies spectra from 454 pyrosequencing reads. BMC Bioinformatics 12, S1. https://doi.org/10.1186/1471-2105-12-S6-S1

Atmar, R.L., Estes, M.K., 2001. Diagnosis of Noncultivatable Gastroenteritis Viruses, the Human Caliciviruses. Clin Microbiol Rev 14, 15–37. https://doi.org/10.1128/CMR.14.1.15-37.2001

Axfors, C., Schmitt, A.M., Janiaud, P., van't Hooft, J., Abd-Elsalam, S., Abdo, E.F., Abella, B.S., Akram, J., Amaravadi, R.K., Angus, D.C., et al., 2021. Mortality outcomes with hydroxychloroquine and chloroquine in COVID-19 from an international collaborative meta-analysis of randomized trials. Nat Commun 12, 2349. https://doi.org/10.1038/s41467-021-22446-z

Baccam, P., Beauchemin, C., Macken, C.A., Hayden, F.G., Perelson, A.S., 2006. Kinetics of Influenza A Virus Infection in Humans. J Virol 80, 7590–7599. https://doi.org/10.1128/JVI.01623-05

Baehner, F., Bogaerts, H., Goodwin, R., 2016. Vaccines against norovirus: state of the art trials in children and adults. Clinical Microbiology and Infection 22, S136–S139. https://doi.org/10.1016/j.cmi.2015.12.023

Baker, R.O., Bray, M., Huggins, J.W., 2003. Potential antiviral therapeutics for smallpox, monkeypox and other orthopoxvirus infections. Antiviral Research 57, 13–23. https://doi.org/10.1016/S0166-3542(02)00196-1

Banerjee, A., Kulcsar, K., Misra, V., Frieman, M., Mossman, K., 2019. Bats and Coronaviruses. Viruses 11, 41. https://doi.org/10.3390/v11010041

Baranovich, T., Wong, S.-S., Armstrong, J., Marjuki, H., Webby, R.J., Webster, R.G., Govorkova, E.A., 2013. T-705 (Favipiravir) Induces Lethal Mutagenesis in Influenza A H1N1 Viruses *In Vitro*. J Virol 87, 3741–3751. https://doi.org/10.1128/JVI.02346-12

Barauskas, O., Xing, W., Aguayo, E., Willkom, M., Sapre, A., Clarke, M., Birkus, G., Schultz, B.E., Sakowicz, R., Kwon, H., et al., 2017. Biochemical characterization of recombinant influenza A polymerase heterotrimer complex: Polymerase activity and mechanisms of action of nucleotide analogs. PLoS ONE 12, e0185998. https://doi.org/10.1371/journal.pone.0185998

Barreca, M.L., Iraci, N., Manfroni, G., Cecchetti, V., 2011. Allosteric inhibition of the hepatitis C virus NS5B polymerase: *in silico* strategies for drug discovery and development. Future Medicinal Chemistry 3, 1027–1055. https://doi.org/10.4155/fmc.11.53

Bartsch, S.M., O'Shea, K.J., Lee, B.Y., 2020. The Clinical and Economic Burden of Norovirus Gastroenteritis in the United States. The Journal of Infectious Diseases 222, 1910–1919. https://doi.org/10.1093/infdis/jiaa292

Bassi, M.R., Sempere, R.N., Meyn, P., Polacek, C., Arias, A., 2018. Extinction of Zika Virus and Usutu Virus by Lethal Mutagenesis Reveals Different Patterns of Sensitivity to Three Mutagenic Drugs. Antimicrob Agents Chemother 62, e00380-18. https://doi.org/10.1128/AAC.00380-18

Bausch, D.G., Hadi, C.M., Khan, S.H., Lertora, J.J.L., 2010. Review of the Literature and Proposed Guidelines for the Use of Oral Ribavirin as Postexposure Prophylaxis for Lassa Fever. CLIN INFECT DIS 51, 1435–1441. https://doi.org/10.1086/657315

Beerenwinkel, N., Zagordi, O., 2011. Ultra-deep sequencing for the analysis of viral populations. Current Opinion in Virology 1, 413–418. https://doi.org/10.1016/j.coviro.2011.07.008

Beigel, J.H., Tomashek, K.M., Dodd, L.E., Mehta, A.K., Zingman, B.S., Kalil, A.C., Hohmann, E., Chu, H.Y., Luetkemeyer, A., Kline, S., et al., 2020a. Remdesivir for the Treatment of Covid-19 — Final Report. N Engl J Med 383, 1813–1826. https://doi.org/10.1056/NEJMoa2007764

Beigel, J.H., Tomashek, K.M., Dodd, L.E., Mehta, A.K., Zingman, B.S., Kalil, A.C., Hohmann, E., Chu, H.Y., Luetkemeyer, A., Kline, S., et al., 2020b. Remdesivir for the Treatment of Covid-19 — Final Report. New England Journal of Medicine 0, null. https://doi.org/10.1056/NEJMoa2007764

Belliot, G., Sosnovtsev, S.V., Chang, K.-O., McPhie, P., Green, K.Y., 2008. Nucleotidylylation of the VPg protein of a human norovirus by its proteinase-

polymerase precursor protein. Virology 374, 33–49. https://doi.org/10.1016/j.virol.2007.12.028

Belliot, G., Sosnovtsev, S.V., Mitra, T., Hammer, C., Garfield, M., Green, K.Y., 2003. In Vitro Proteolytic Processing of the MD145 Norovirus ORF1 Nonstructural Polyprotein Yields Stable Precursors and Products Similar to Those Detected in Calicivirus-Infected Cells. J Virol 77, 10957–10974. https://doi.org/10.1128/JVI.77.20.10957-10974.2003

Benidt, S., Nettleton, D., 2015. SimSeq: a nonparametric approach to simulation of RNA-sequence datasets. Bioinformatics 31, 2131–2140. https://doi.org/10.1093/bioinformatics/btv124

Bergmann, C.C., Silverman, R.H., 2020. COVID-19: Coronavirus replication, pathogenesis, and therapeutic strategies. CCJM 87, 321–327. https://doi.org/10.3949/ccjm.87a.20047

Binder, S., Levitt, A.M., Sacks, J.J., Hughes, J.M., 1999. Emerging Infectious Diseases: Public Health Issues for the 21st Century. Science 284, 1311–1313. https://doi.org/10.1126/science.284.5418.1311

Bobrowski, T., Chen, L., Eastman, R.T., Itkin, Z., Shinn, P., Chen, C., Guo, H., Zheng, W., Michael, S., Simeonov, A., et al., 2020. Discovery of Synergistic and Antagonistic Drug Combinations against SARS-CoV-2 In Vitro (preprint). Microbiology. https://doi.org/10.1101/2020.06.29.178889

Bok, K., Green, K.Y., 2012. Norovirus Gastroenteritis in Immunocompromised Patients. N Engl J Med 367, 2126–2132. https://doi.org/10.1056/NEJMra1207742

Bok, K., Parra, G.I., Mitra, T., Abente, E., Shaver, C.K., Boon, D., Engle, R., Yu, C., Kapikian, A.Z., Sosnovtsev, S.V., et al., 2011. Chimpanzees as an animal model for human norovirus infection and vaccine development. Proc. Natl. Acad. Sci. U.S.A. 108, 325–330. https://doi.org/10.1073/pnas.1014577107

Bosaeed, M., Alharbi, A., Mahmoud, E., Alrehily, S., Bahlaq, M., Gaifer, Z., Alturkistani, H., Alhagan, K., Alshahrani, S., Tolbah, A., et al., 2022. Efficacy of favipiravir in adults with mild COVID-19: a randomized, double-blind, multicentre, placebo-controlled clinical trial. Clinical Microbiology and Infection 28, 602–608. https://doi.org/10.1016/j.cmi.2021.12.026

Boshier, F.A.T., Pang, J., Penner, J., Hughes, J., Parker, M., Shepherd, J., Alders, N., Bamford, A., Grandjean, L., Grunewald, S., et al., 2020a. Remdesivir induced viral RNA and subgenomic RNA suppression, and evolution of viral variants in SARS-CoV-2 infected patients (preprint). Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2020.11.18.20230599

Boshier, F.A.T., Pang, J., Penner, J., Hughes, J., Parker, M., Shepherd, J., Alders, N., Bamford, A., Grandjean, L., Grunewald, S., et al., 2020b. Remdesivir induced viral RNA and subgenomic RNA suppression, and evolution of viral variants in SARS-CoV-2 infected patients (preprint). Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2020.11.18.20230599

Bradwell, K., Combe, M., Domingo-Calap, P., Sanjuán, R., 2013. Correlation Between Mutation Rate and Genome Size in Riboviruses: Mutation Rate of Bacteriophage Qβ. Genetics 195, 243–251. https://doi.org/10.1534/genetics.113.154963

Broad Institute, 2019a. Picard GitHub Repository [WWW Document]. URL http://broadinstitute.github.io/picard/

Broad Institute, 2019b. Picard GitHub Repository [WWW Document]. URL http://broadinstitute.github.io/picard/

Brown, J.R., Gilmour, K., Breuer, J., 2016. Norovirus Infections Occur in B-Cell–Deficient Patients: Table 1. Clin Infect Dis. 62, 1136–1138. https://doi.org/10.1093/cid/ciw060

Brown, L.-A.K., Clark, I., Brown, J.R., Breuer, J., Lowe, D.M., 2017. Norovirus infection in primary immune deficiency. Rev. Med. Virol. 27, e1926. https://doi.org/10.1002/rmv.1926

Brown, L.-A.K., Ruis, C., Clark, I., Roy, S., Brown, J.R., Albuquerque, A.S., Patel, S.Y., Miller, J., Karim, M.Y., Dervisevic, S., et al., 2019. A comprehensive characterization of chronic norovirus infection in immunodeficient hosts. Journal of Allergy and Clinical Immunology 144, 1450–1453. https://doi.org/10.1016/j.jaci.2019.07.036

Bruenn, J.A., 2003. A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. Nucleic Acids Research 31, 1821–1829. https://doi.org/10.1093/nar/gkg277

Bruno, L., Cortese, M., Monda, G., Gentile, M., Calò, S., Schiavetti, F., Zedda, L., Cattaneo, E., Piccioli, D., Schaefer, M., et al., 2016. Human cytomegalovirus pUL10 interacts with leukocytes and impairs TCR-mediated T-cell activation. Immunol Cell Biol 94, 849–860. https://doi.org/10.1038/icb.2016.49

Bryan, A., Fink, S.L., Gattuso, M.A., Pepper, G., Chaudhary, A., Wener, M.H., Morishima, C., Jerome, K.R., Mathias, P.C., Greninger, A.L., 2020. SARS-CoV-2 Viral Load on Admission Is Associated With 30-Day Mortality. Open Forum Infectious Diseases 7, ofaa535. https://doi.org/10.1093/ofid/ofaa535

Bui, T., Kocher, J., Li, Y., Wen, K., Li, G., Liu, F., Yang, X., LeRoith, T., Tan, M., Xia, M., et al., 2013. Median infectious dose of human norovirus GII.4 in gnotobiotic pigs is decreased by simvastatin treatment and increased by age. Journal of General Virology 94, 2005–2016. https://doi.org/10.1099/vir.0.054080-0

Bull, J.J., Sanjuán, R., Wilke, C.O., 2007. Theory of Lethal Mutagenesis for Viruses. J Virol 81, 2930–2939. https://doi.org/10.1128/JVI.01624-06

Bull, R.A., Eden, J.-S., Rawlinson, W.D., White, P.A., 2010. Rapid Evolution of Pandemic Noroviruses of the GII.4 Lineage. PLoS Pathog 6, e1000831. https://doi.org/10.1371/journal.ppat.1000831

Bull, R.A., Hansman, G.S., Clancy, L.E., Tanaka, M.M., Rawlinson, W.D., White, P.A., 2005. Norovirus Recombination in ORF1/ORF2 Overlap. Emerg. Infect. Dis. 11, 1079–1085. https://doi.org/10.3201/eid1107.041273

Buonaguro, L., Buonaguro, F.M., 2020. Knowledge-based repositioning of the anti-HCV direct antiviral agent Sofosbuvir as SARS-CoV-2 treatment. Infect Agents Cancer 15, 32. https://doi.org/10.1186/s13027-020-00302-x

Çalıca Utku, A., Budak, G., Karabay, O., Güçlü, E., Okan, H.D., Vatan, A., 2020. Main symptoms in patients presenting in the COVID-19 period. Scott Med J 65, 127–132. https://doi.org/10.1177/0036933020949253

Campillay-Véliz, C.P., Carvajal, J.J., Avellaneda, A.M., Escobar, D., Covián, C., Kalergis, A.M., Lay, M.K., 2020. Human Norovirus Proteins: Implications in the Replicative Cycle, Pathogenesis, and the Host Immune Response. Front. Immunol. 11, 961. https://doi.org/10.3389/fimmu.2020.00961

Cannon, J.L., Bonifacio, J., Bucardo, F., Buesa, J., Bruggink, L., Chan, M.C.-W., Fumian, T.M., Giri, S., Gonzalez, M.D., Hewitt, J., et al., 2021. Global Trends in Norovirus Genotype Distribution among Children with Acute Gastroenteritis. Emerg. Infect. Dis. 27, 1438–1445. https://doi.org/10.3201/eid2705.204756

Capizzi, T., Makari-Judson, G., Steingart, R., Mertens, W.C., 2011. Chronic diarrhea associated with persistent norovirus excretion in patients with chronic lymphocytic leukemia: report of two cases. BMC Infect Dis 11, 131. https://doi.org/10.1186/1471-2334-11-131

Carrasco-Hernandez, R., Jácome, R., López Vidal, Y., Ponce de León, S., 2017. Are RNA Viruses Candidate Agents for the Next Global Pandemic? A Review. ILAR Journal 58, 343–358. https://doi.org/10.1093/ilar/ilx026

Carrasquer, A., Peiró, Ó.M., Sanchez-Gimenez, R., Lal-Trehan, N., del-Moral-Ronda, V., Bonet, G., Gutierrez, C., Fort-Gallifa, I., Martin-Grau, C., Benavent, C., et al., 2021. Lack of Association of Initial Viral Load in SARS-CoV-2 Patients with In-Hospital Mortality. The American Journal of Tropical Medicine and Hygiene 104, 540–545. https://doi.org/10.4269/ajtmh.20-1427

Cellina, M., Orsi, M., Bombaci, F., Sala, M., Marino, P., Oliva, G., 2020. Favorable changes of CT findings in a patient with COVID-19 pneumonia after treatment with tocilizumab. Diagnostic and Interventional Imaging 101, 323–324. https://doi.org/10.1016/j.diii.2020.03.010

Center for Drug Evaluation and Research, 2022. Coronavirus Treatment Acceleration Program (CTAP). FDA.

Cevik, M., Tate, M., Lloyd, O., Maraolo, A.E., Schafers, J., Ho, A., 2021. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. The Lancet Microbe 2, e13–e22. https://doi.org/10.1016/S2666-5247(20)30172-5

Chachu, K.A., Strong, D.W., LoBue, A.D., Wobus, C.E., Baric, R.S., Virgin, H.W., 2008. Antibody Is Critical for the Clearance of Murine Norovirus Infection. J Virol 82, 6610–6617. https://doi.org/10.1128/JVI.00141-08

Chacón-Labella, J., Boakye, M., Enquist, B.J., Farfan-Rios, W., Gya, R., Halbritter, A.H., Middleton, S.L., Oppen, J., Pastor-Ploskonka, S., Strydom, T., et al., 2021. From a crisis to an opportunity: Eight insights for doing science in the COVID-19 era and beyond. Ecol. Evol. 11, 3588–3596. https://doi.org/10.1002/ece3.7026

Chaguza, C., Hahn, A.M., Petrone, M.E., Zhou, S., Ferguson, D., Breban, M.I., Pham, K., Peña-Hernández, M.A., Castaldi, C., Hill, V., et al., 2022. Accelerated SARS-CoV-2 intrahost evolution leading to distinct genotypes during chronic infection (preprint). Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2022.06.29.22276868

Challenger, J.D., Foo, C.Y., Wu, Y., Yan, A.W.C., Marjaneh, M.M., Liew, F., Thwaites, R.S., Okell, L.C., Cunnington, A.J., 2022. Modelling upper respiratory viral

load dynamics of SARS-CoV-2. BMC Med 20, 25. https://doi.org/10.1186/s12916-021-02220-0

Chan, J.F.-W., Kok, K.-H., Zhu, Z., Chu, H., To, K.K.-W., Yuan, S., Yuen, K.-Y., 2020. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerging Microbes & Infections 9, 221–236. https://doi.org/10.1080/22221751.2020.1719902

Chan, M.C.W., Lee, N., Hung, T.-N., Kwok, K., Cheung, K., Tin, E.K.Y., Lai, R.W.M., Nelson, E.A.S., Leung, T.F., Chan, P.K.S., 2015. Rapid emergence and predominance of a broadly recognizing and fast-evolving norovirus GII.17 variant in late 2014. Nat Commun 6, 10061. https://doi.org/10.1038/ncomms10061

Cheetham, S., Souza, M., Meulia, T., Grimes, S., Han, M.G., Saif, L.J., 2006. Pathogenesis of a Genogroup II Human Norovirus in Gnotobiotic Pigs. J Virol 80, 10372–10381. https://doi.org/10.1128/JVI.00809-06

Cheever, F.S., Daniels, J.B., Pappenheimer, A.M., Bailey, O.T., 1949. A MURINE VIRUS (JHM) CAUSING DISSEMINATED ENCEPHALOMYELITIS WITH EXTENSIVE DESTRUCTION OF MYELIN. Journal of Experimental Medicine 90, 181–194. https://doi.org/10.1084/jem.90.3.181

Chen, C.Z., Shinn, P., Itkin, Z., Eastman, R.T., Bostwick, R., Rasmussen, L., Huang, R., Shen, M., Hu, X., Wilson, K.M., et al., 2021. Drug Repurposing Screen for Compounds Inhibiting the Cytopathic Effect of SARS-CoV-2. Front. Pharmacol. 11, 592737. https://doi.org/10.3389/fphar.2020.592737

Chen, J., Qi, T., Liu, L., Ling, Y., Qian, Z., Li, T., Li, F., Xu, Q., Zhang, Y., Xu, S., et al., 2020. Clinical progression of patients with COVID-19 in Shanghai, China. Journal of Infection 80, e1–e6. https://doi.org/10.1016/j.jinf.2020.03.004

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., et al., 2020. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. The Lancet 395, 507–513. https://doi.org/10.1016/S0140-6736(20)30211-7

Chen, W., Lan, Y., Yuan, X., Deng, Xilong, Li, Y., Cai, X., Li, L., He, R., Tan, Y., Deng, Xizi, et al., 2020. Detectable 2019-nCoV viral RNA in blood is a strong indicator for the further clinical severity. Emerging Microbes & Infections 9, 469–473. https://doi.org/10.1080/22221751.2020.1732837

Chen, X., Zhu, B., Hong, W., Zeng, J., He, X., Chen, J., Zheng, H., Qiu, S., Deng, Y., Chan, J.C.N., et al., 2020. Associations of clinical characteristics and treatment regimens with the duration of viral RNA shedding in patients with COVID-19. International Journal of Infectious Diseases 98, 252–260. https://doi.org/10.1016/j.ijid.2020.06.091

Chen, Y., Cai, H., Pan, J., Xiang, N., Tien, P., Ahola, T., Guo, D., 2009. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. Proc. Natl. Acad. Sci. U.S.A. 106, 3484–3489. https://doi.org/10.1073/pnas.0808790106

Chen, Y., Su, C., Ke, M., Jin, X., Xu, L., Zhang, Z., Wu, A., Sun, Y., Yang, Z., Tien, P., et al., 2011. Biochemical and Structural Insights into the Mechanisms of SARS Coronavirus RNA Ribose 2′-O-Methylation by nsp16/nsp10 Protein

Complex. PLoS Pathog 7, e1002294. https://doi.org/10.1371/journal.ppat.1002294

Chen, Z., Ng, R.W.Y., Lui, G., Ling, L., Chow, C., Yeung, A.C.M., Boon, S.S., Wang, M.H., Chan, K.C.C., Chan, R.W.Y., et al., 2022. Profiling of SARS-CoV-2 Subgenomic RNAs in Clinical Specimens. Microbiol Spectr 10, e00182-22. https://doi.org/10.1128/spectrum.00182-22

Cheng, C.-Y., Lee, Y.-L., Chen, C.-P., Lin, Y.-C., Liu, C.-E., Liao, C.-H., Cheng, S.-H., 2020. Lopinavir/ritonavir did not shorten the duration of SARS CoV-2 shedding in patients with mild pneumonia in Taiwan. J Microbiol Immunol Infect. https://doi.org/10.1016/j.jmii.2020.03.032

Chhabra, P., de Graaf, M., Parra, G.I., Chan, M.C.-W., Green, K., Martella, V., Wang, Q., White, P.A., Katayama, K., Vennema, H., et al., 2019. Updated classification of norovirus genogroups and genotypes. Journal of General Virology 100, 1393–1406. https://doi.org/10.1099/jgv.0.001318

Chouchana, L., Preta, L.-H., Tisseyre, M., Terrier, B., Treluyer, J.-M., Montastruc, F., 2021. Kidney disorders as serious adverse drug reactions of remdesivir in coronavirus disease 2019: a retrospective case–noncase study. Kidney International 99, 1235–1236. https://doi.org/10.1016/j.kint.2021.02.015

Choudhury, S., Moulick, D., Saikia, P., Mazumder, M.K., 2021. Evaluating the potential of different inhibitors on RNA-dependent RNA polymerase of severe acute respiratory syndrome coronavirus 2: A molecular modeling approach. Medical Journal Armed Forces India 77, S373–S378. https://doi.org/10.1016/j.mjafi.2020.05.005

Chung, L., Bailey, D., Leen, E.N., Emmott, E.P., Chaudhry, Y., Roberts, L.O., Curry, S., Locker, N., Goodfellow, I.G., 2014. Norovirus Translation Requires an Interaction between the C Terminus of the Genome-linked Viral Protein VPg and Eukaryotic Translation Initiation Factor 4G. Journal of Biological Chemistry 289, 21738–21750. https://doi.org/10.1074/jbc.M114.550657

Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2016. GenBank. Nucleic Acids Res 44, D67–D72. https://doi.org/10.1093/nar/gkv1276

Clercq, E.D., 2007. The design of drugs for HIV and HCV. Nat Rev Drug Discov 6, 1001–1018. https://doi.org/10.1038/nrd2424

CMMID COVID-19 Working Group, Davies, N.G., Jarvis, C.I., Edmunds, W.J., Jewell, N.P., Diaz-Ordaz, K., Keogh, R.H., 2021. Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. Nature 593, 270–274. https://doi.org/10.1038/s41586-021-03426-1

Colavita, F., Lapa, D., Carletti, F., Lalle, E., Bordi, L., Marsella, P., Nicastri, E., Bevilacqua, N., Giancola, M.L., Corpolongo, A., et al., 2020. SARS-CoV-2 Isolation From Ocular Secretions of a Patient With COVID-19 in Italy With Prolonged Viral RNA Detection. Ann. Intern. Med. https://doi.org/10.7326/M20-1176

Contreras, C., Newby, J.M., Hillen, T., 2021. Personalized Virus Load Curves for Acute Viral Infections. Viruses 13, 1815. https://doi.org/10.3390/v13091815

Cortese, M., Calò, S., D'Aurizio, R., Lilja, A., Pacchiani, N., Merola, M., 2012. Recombinant Human Cytomegalovirus (HCMV) RL13 Binds Human

Immunoglobulin G Fc. PLoS ONE 7, e50166. https://doi.org/10.1371/journal.pone.0050166

Costantini, V.P., Whitaker, T., Barclay, L., Lee, D., McBrayer, T.R., Schinazi, R.F., Vinjé, J., 2012. Antiviral Activity of Nucleoside Analogues against Norovirus. Antiviral Therapy 17, 981–991. https://doi.org/10.3851/IMP2229

Courville, C., Cadarette, S.M., Wissinger, E., Alvarez, F.P., 2022. The economic burden of influenza among adults aged 18 to 64: A systematic literature review. Influenza Resp Viruses 16, 376–385. https://doi.org/10.1111/irv.12963

COVID-19 Genomics UK (COG-UK) consortium, 2020. An integrated national scale SARS-CoV-2 genomic surveillance network. The Lancet Microbe 1, e99–e100. https://doi.org/10.1016/S2666-5247(20)30054-9

COVID-19 Investigation Team, 2020. Clinical and virologic characteristics of the first 12 patients with coronavirus disease 2019 (COVID-19) in the United States. Nat. Med. https://doi.org/10.1038/s41591-020-0877-5

Crotty, S., Cameron, C.E., Andino, R., 2001. RNA virus error catastrophe: Direct molecular test by using ribavirin. Proceedings of the National Academy of Sciences 98, 6895–6900. https://doi.org/10.1073/pnas.111085598

Crotty, S., Maag, D., Arnold, J.J., Zhong, W., Lau, J.Y.N., Hong, Z., Andino, R., Cameron, C.E., 2000. The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. Nat Med 6, 1375–1379. https://doi.org/10.1038/82191

Cudini, J., Roy, S., Houldcroft, C.J., Bryant, J.M., Depledge, D.P., Tutill, H., Veys, P., Williams, R., Worth, A.J.J., Tamuri, A.U., et al., 2019. Human cytomegalovirus haplotype reconstruction reveals high diversity due to superinfection and evidence of within-host recombination. Proc. Natl. Acad. Sci. U.S.A. 116, 5693–5698. https://doi.org/10.1073/pnas.1818130116

da Silva Poló, T., Peiró, J.R., Mendes, L.C.N., Ludwig, L.F., de Oliveira-Filho, E.F., Bucardo, F., Huynen, P., Melin, P., Thiry, E., Mauroy, A., 2016. Human norovirus infection in Latin America. Journal of Clinical Virology 78, 111–119. https://doi.org/10.1016/j.jcv.2016.03.016

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al., 2021. Twelve years of SAMtools and BCFtools. GigaScience 10, giab008. https://doi.org/10.1093/gigascience/giab008

Dang, W., Xu, L., Ma, B., Chen, S., Yin, Y., Chang, K.-O., Peppelenbosch, M.P., Pan, Q., 2018. Nitazoxanide Inhibits Human Norovirus Replication and Synergizes with Ribavirin by Activation of Cellular Antiviral Response. Antimicrob Agents Chemother 62, e00707-18. https://doi.org/10.1128/AAC.00707-18

Dapp, M.J., Holtz, C.M., Mansky, L.M., 2012. Concomitant Lethal Mutagenesis of Human Immunodeficiency Virus Type 1. Journal of Molecular Biology 419, 158–170. https://doi.org/10.1016/j.jmb.2012.03.003

Daughenbaugh, K.F., Wobus, C.E., Hardy, M.E., 2006. VPg of murine norovirus binds translation initiation factors in infected cells. Virol J 3, 33. https://doi.org/10.1186/1743-422X-3-33

Day, C., Smee, D., Julander, J., Yamshchikov, V., Sidwell, R., Morrey, J., 2005. Error-prone replication of West Nile virus caused by ribavirin. Antiviral Research 67, 38–45. https://doi.org/10.1016/j.antiviral.2005.04.002

de Ávila, A.I., Gallego, I., Soria, M.E., Gregori, J., Quer, J., Esteban, J.I., Rice, C.M., Domingo, E., Perales, C., 2016. Lethal Mutagenesis of Hepatitis C Virus Induced by Favipiravir. PLoS ONE 11, e0164691. https://doi.org/10.1371/journal.pone.0164691

de Graaf, M., Bodewes, R., van Elk, C.E., van de Bildt, M., Getu, S., Aron, G.I., Verjans, G.M.G.M., Osterhaus, A.D.M.E., van den Brand, J.M.A., Kuiken, T., et al., 2017. Norovirus Infection in Harbor Porpoises. Emerg. Infect. Dis. 23, 87–91. https://doi.org/10.3201/eid2301.161081

de Graaf, M., van Beek, J., Koopmans, M.P.G., 2016. Human norovirus transmission and evolution in a changing world. Nat Rev Microbiol 14, 421–433. https://doi.org/10.1038/nrmicro.2016.48

Delang, L., Abdelnabi, R., Neyts, J., 2018. Favipiravir as a potential countermeasure against neglected and emerging RNA viruses. Antiviral Research 153, 85–94. https://doi.org/10.1016/j.antiviral.2018.03.003

Delang, L., Neyts, J., Vliegen, I., Abrignani, S., Neddermann, P., De Francesco, R., 2013. Hepatitis C Virus-Specific Directly Acting Antiviral Drugs, in: Bartenschlager, R. (Ed.), Hepatitis C Virus: From Molecular Virology to Antiviral Therapy, Current Topics in Microbiology and Immunology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 289–320. https://doi.org/10.1007/978-3-642-27340-7_12

Delang, L., Segura Guerrero, N., Tas, A., Quérat, G., Pastorino, B., Froeyen, M., Dallmeier, K., Jochmans, D., Herdewijn, P., Bello, F., et al., 2014. Mutations in the chikungunya virus non-structural proteins cause resistance to favipiravir (T-705), a broad-spectrum antiviral. Journal of Antimicrobial Chemotherapy 69, 2770–2784. https://doi.org/10.1093/jac/dku209

Desai, N., Neyaz, A., Szabolcs, A., Shih, A.R., Chen, J.H., Thapar, V., Nieman, L.T., Solovyov, A., Mehta, A., Lieb, D.J., et al., 2020. Temporal and Spatial Heterogeneity of Host Response to SARS-CoV-2 Pulmonary Infection. medRxiv. https://doi.org/10.1101/2020.07.30.20165241

Deval, J., Jin, Z., Chuang, Y.-C., Kao, C.C., 2017. Structure(s), function(s), and inhibition of the RNA-dependent RNA polymerase of noroviruses. Virus Research 234, 21–33. https://doi.org/10.1016/j.virusres.2016.12.018

Di Martino, B., Di Profio, F., Melegari, I., Sarchese, V., Cafiero, M.A., Robetto, S., Aste, G., Lanave, G., Marsilio, F., Martella, V., 2016. A novel feline norovirus in diarrheic cats. Infection, Genetics and Evolution 38, 132–137. https://doi.org/10.1016/j.meegid.2015.12.019

Diamond, M.S., Kanneganti, T.-D., 2022. Innate immunity: the first line of defense against SARS-CoV-2. Nat Immunol 23, 165–176. https://doi.org/10.1038/s41590-021-01091-0

Díaz-Martínez, L., Brichette-Mieg, I., Pineño-Ramos, A., Domínguez-Huerta, G., Grande-Pérez, A., 2018. Lethal mutagenesis of an RNA plant virus via lethal defection. Sci Rep 8, 1444. https://doi.org/10.1038/s41598-018-19829-6

Dolan, P.T., Whitfield, Z.J., Andino, R., 2018. Mechanisms and Concepts in RNA Virus Population Dynamics and Evolution. Annu. Rev. Virol. 5, 69–92. https://doi.org/10.1146/annurev-virology-101416-041718

Donaldson, E.F., Lindesmith, L.C., LoBue, A.D., Baric, R.S., 2010. Viral shape-shifting: norovirus evasion of the human immune system. Nat Rev Microbiol 8, 231–241. https://doi.org/10.1038/nrmicro2296

Donaldson, E.F., Lindesmith, L.C., Lobue, A.D., Baric, R.S., 2008. Norovirus pathogenesis: mechanisms of persistence and immune evasion in human populations. Immunological Reviews 225, 190–211. https://doi.org/10.1111/j.1600-065X.2008.00680.x

Drake, J.W., 1993. Rates of spontaneous mutation among RNA viruses. Proc. Natl. Acad. Sci. U.S.A. 90, 4171–4175. https://doi.org/10.1073/pnas.90.9.4171

Drake, J.W., 1991. A constant rate of spontaneous mutation in DNA-based microbes. Proc. Natl. Acad. Sci. U.S.A. 88, 7160–7164. https://doi.org/10.1073/pnas.88.16.7160

Driouich, J.-S., Cochin, M., Lingas, G., Moureau, G., Touret, F., Petit, P.-R., Piorkowski, G., Barthélémy, K., Laprie, C., Coutard, B., et al., 2021. Favipiravir antiviral efficacy against SARS-CoV-2 in a hamster model. Nat Commun 12, 1735. https://doi.org/10.1038/s41467-021-21992-w

Du, W., Yu, J., Wang, H., Zhang, X., Zhang, S., Li, Q., Zhang, Z., 2020. Clinical characteristics of COVID-19 in children compared with adults in Shandong Province, China. Infection 48, 445–452. https://doi.org/10.1007/s15010-020-01427-2

Duan, L., Zheng, Q., Zhang, H., Niu, Y., Lou, Y., Wang, H., 2020. The SARS-CoV-2 Spike Glycoprotein Biosynthesis, Structure, Function, and Antigenicity: Implications for the Design of Spike-Based Vaccine Immunogens. Front. Immunol. 11, 576622. https://doi.org/10.3389/fimmu.2020.576622

Duke, E.R., Williamson, B.D., Wychera, C., Cossrow, N., Marks, M.A., Wan, H., Mast, T.C., Huang, M.-L., Jerome, K., Corey, L., et al., 2020. CMV Viral Load Kinetics as Surrogate Endpoints for Antiviral Prophylaxis Trials. Biology of Blood and Marrow Transplantation 26, S327–S328. https://doi.org/10.1016/j.bbmt.2019.12.352

Duong, D., 2021. Alpha, Beta, Delta, Gamma: What's important to know about SARS-CoV-2 variants of concern? CMAJ 193, E1059–E1060. https://doi.org/10.1503/cmaj.1095949

Eden, J.-S., Hewitt, J., Lim, K.L., Boni, M.F., Merif, J., Greening, G., Ratcliff, R.M., Holmes, E.C., Tanaka, M.M., Rawlinson, W.D., et al., 2014. The emergence and evolution of the novel epidemic norovirus GII.4 variant Sydney 2012. Virology 450–451, 106–113. https://doi.org/10.1016/j.virol.2013.12.005

Eden, J.-S., Tanaka, M.M., Boni, M.F., Rawlinson, W.D., White, P.A., 2013. Recombination within the Pandemic Norovirus GII.4 Lineage. J Virol 87, 6270–6282. https://doi.org/10.1128/JVI.03464-12

Ehteshami, M., Tao, S., Zandi, K., Hsiao, H.-M., Jiang, Y., Hammond, E., Amblard, F., Russell, O.O., Merits, A., Schinazi, R.F., 2017. Characterization of β- D -*N* 4 -Hydroxycytidine as a Novel Inhibitor of Chikungunya Virus. Antimicrob Agents Chemother 61, e02395-16. https://doi.org/10.1128/AAC.02395-16

Ekstrøm, Claus Thorn, 2019. MESS: Miscellaneous Esoteric Statistical Scripts. R package version 0.5.6.

Eliseev, A., Gibson, K.M., Avdeyev, P., Novik, D., Bendall, M.L., Pérez-Losada, M., Alexeev, N., Crandall, K.A., 2020a. Evaluation of haplotype callers for next-generation sequencing of viruses. Infection, Genetics and Evolution 82, 104277. https://doi.org/10.1016/j.meegid.2020.104277

Eliseev, A., Gibson, K.M., Avdeyev, P., Novik, D., Bendall, M.L., Pérez-Losada, M., Alexeev, N., Crandall, K.A., 2020b. Evaluation of haplotype callers for next-generation sequencing of viruses. Infection, Genetics and Evolution 82, 104277. https://doi.org/10.1016/j.meegid.2020.104277

Eltahla, A., Luciani, F., White, P., Lloyd, A., Bull, R., 2015. Inhibitors of the Hepatitis C Virus Polymerase; Mode of Action and Resistance. Viruses 7, 5206–5224. https://doi.org/10.3390/v7102868

Ely, E.W., Ramanan, A.V., Kartman, C.E., de Bono, S., Liao, R., Piruzeli, M.L.B., Goldman, J.D., Saraiva, J.F.K., Chakladar, S., Marconi, V.C., 2021. Baricitinib plus Standard of Care for Hospitalised Adults with COVID-19 on Invasive Mechanical Ventilation or Extracorporeal Membrane Oxygenation: Results of a Randomised, Placebo-Controlled Trial (preprint). Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2021.10.11.21263897

Eriksson, B., Helgstrand, E., Johansson, N.G., Larsson, A., Misiorny, A., Noren, J.O., Philipson, L., Stenberg, K., Stening, G., Stridh, S., et al., 1977. Inhibition of Influenza Virus Ribonucleic Acid Polymerase by Ribavirin Triphosphate. Antimicrob Agents Chemother 11, 946–951. https://doi.org/10.1128/AAC.11.6.946

Escribano-Romero, E., Jiménez de Oya, N., Domingo, E., Saiz, J.C., 2017. Extinction of West Nile Virus by Favipiravir through Lethal Mutagenesis. Antimicrob Agents Chemother 61, e01400-17. https://doi.org/10.1128/AAC.01400-17

Estes, M.K., Ettayebi, K., Tenge, V.R., Murakami, K., Karandikar, U., Lin, S.-C., Ayyar, B.V., Cortes-Penfield, N.W., Haga, K., Neill, F.H., et al., 2019. Human Norovirus Cultivation in Nontransformed Stem Cell-Derived Human Intestinal Enteroid Cultures: Success and Challenges. Viruses 11, 638. https://doi.org/10.3390/v11070638

Ettayebi, K., Crawford, S.E., Murakami, K., Broughman, J.R., Karandikar, U., Tenge, V.R., Neill, F.H., Blutt, S.E., Zeng, X.-L., Qu, L., et al., 2016. Replication of human noroviruses in stem cell-derived human enteroids. Science 353, 1387–1393. https://doi.org/10.1126/science.aaf5211

Ettayebi, K., Hardy, M.E., 2003. Norwalk Virus Nonstructural Protein p48 Forms a Complex with the SNARE Regulator VAP-A and Prevents Cell Surface Expression of Vesicular Stomatitis Virus G Protein. J Virol 77, 11790–11797. https://doi.org/10.1128/JVI.77.21.11790-11797.2003

Euser, S., Aronson, S., Manders, I., van Lelyveld, S., Herpers, B., Sinnige, J., Kalpoe, J., van Gemeren, C., Snijders, D., Jansen, R., et al., 2021. SARS-CoV-2 viral load distribution reveals that viral loads increase with age: a retrospective cross-sectional cohort study (preprint). Epidemiology. https://doi.org/10.1101/2021.01.15.21249691

Evaluation and Licensing Division, Pharmaceutical and Food Safety Bureau, Japan, 2011. Report on the deliberation results – avigan.

Ferrerorta, C., Arias, A., Escarmis, C., Verdaguer, N., 2006. A comparison of viral RNA-dependent RNA polymerases. Current Opinion in Structural Biology 16, 27–34. https://doi.org/10.1016/j.sbi.2005.12.002

Ferretti, L., Ramos-Onsins, S.E., 2015. A generalized Watterson estimator for next-generation sequencing: From trios to autopolyploids. Theoretical Population Biology 100, 79–87. https://doi.org/10.1016/j.tpb.2015.01.001

Franco, M.A., Greenberg, H.B., 2012. Rotaviruses, Noroviruses, and Other Gastrointestinal Viruses, in: Goldman's Cecil Medicine. Elsevier, pp. 2144–2147. https://doi.org/10.1016/B978-1-4377-1604-7.00388-2

Frediansyah, A., Nainu, F., Dhama, K., Mudatsir, M., Harapan, H., 2021. Remdesivir and its antiviral activity against COVID-19: A systematic review. Clinical Epidemiology and Global Health 9, 123–127. https://doi.org/10.1016/j.cegh.2020.07.011

Furuta, Y., Gowen, B.B., Takahashi, K., Shiraki, K., Smee, D.F., Barnard, D.L., 2013. Favipiravir (T-705), a novel viral RNA polymerase inhibitor. Antiviral Research 100, 446–454. https://doi.org/10.1016/j.antiviral.2013.09.015

Furuta, Y., Komeno, T., Nakamura, T., 2017. Favipiravir (T-705), a broad spectrum inhibitor of viral RNA polymerase. Proceedings of the Japan Academy. Ser. B: Physical and Biological Sciences 93, 449–463. https://doi.org/10.2183/pjab.93.027

Furuta, Y., Takahashi, K., Shiraki, K., Sakamoto, K., Smee, D.F., Barnard, D.L., Gowen, B.B., Julander, J.G., Morrey, J.D., 2009. T-705 (favipiravir) and related compounds: Novel broad-spectrum inhibitors of RNA viral infections. Antiviral Research 82, 95–102. https://doi.org/10.1016/j.antiviral.2009.02.198

Gabaev, I., Elbasani, E., Ameres, S., Steinbrück, L., Stanton, R., Döring, M., Lenac Rovis, T., Kalinke, U., Jonjic, S., Moosmann, A., et al., 2014. Expression of the Human Cytomegalovirus UL11 Glycoprotein in Viral Infection and Evaluation of Its Effect on Virus-Specific CD8 T Cells. J Virol 88, 14326–14339. https://doi.org/10.1128/JVI.01691-14

Galani, I.-E., Rovina, N., Lampropoulou, V., Triantafyllia, V., Manioudaki, M., Pavlos, E., Koukaki, E., Fragkou, P.C., Panou, V., Rapti, V., et al., 2021. Untuned antiviral immunity in COVID-19 revealed by temporal type I/III interferon patterns and flu comparison. Nat Immunol 22, 32–40. https://doi.org/10.1038/s41590-020-00840-x

Gao, C., Zhu, L., Jin, C.C., Tong, Y.X., Xiao, A.T., Zhang, S., 2021. Prevalence and impact factors of recurrent positive SARS-CoV-2 detection in 599 hospitalized COVID-19 patients. Clinical Microbiology and Infection 27, 785.e1-785.e7. https://doi.org/10.1016/j.cmi.2021.01.028

Gao, Y., Yan, L., Huang, Y., Liu, F., Zhao, Y., Cao, L., Wang, T., Sun, Q., Ming, Z., Zhang, L., et al., 2020. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science 368, 779–782. https://doi.org/10.1126/science.abb7498

Gastine, S., Pang, J., Boshier, F.A.T., Carter, S.J., Lonsdale, D.O., Cortina-Borja, M., Hung, I.F.N., Breuer, J., Kloprogge, F., Standing, J.F., 2021. Systematic review and patient-level meta-analysis of SARS-CoV-2 viral dynamics to

model response to antiviral therapies. Clin Pharmacol Ther. https://doi.org/10.1002/cpt.2223

Gastine, S., Pang, J., Boshier, F.A.T., Carter, S.J., Lonsdale, D.O., Cortina-Borja, M., Hung, I.F.N., Breuer, J., Kloprogge, F., Standing, J.F., 2020. Systematic review and patient-level meta-analysis of SARS-CoV-2 viral dynamics to model response to antiviral therapies (preprint). Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2020.08.20.20178699

Gautret, P., Lagier, J.-C., Parola, P., Hoang, V.T., Meddeb, L., Mailhe, M., Doudier, B., Courjon, J., Giordanengo, V., Vieira, V.E., et al., 2020. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. International Journal of Antimicrobial Agents 105949. https://doi.org/10.1016/j.ijantimicag.2020.105949

Gaythorpe, K.A.M., Trotter, C.L., Lopman, B., Steele, M., Conlan, A.J.K., 2018. Norovirus transmission dynamics: a modelling review. Epidemiol. Infect. 146, 147–158. https://doi.org/10.1017/S0950268817002692

George, P.M., Barratt, S.L., Condliffe, R., Desai, S.R., Devaraj, A., Forrest, I., Gibbons, M.A., Hart, N., Jenkins, R.G., McAuley, D.F., et al., 2020. Respiratory follow-up of patients with COVID-19 pneumonia. Thorax 75, 1009–1016. https://doi.org/10.1136/thoraxjnl-2020-215314

Ghedin, E., Fitch, A., Boyne, A., Griesemer, S., DePasse, J., Bera, J., Zhang, X., Halpin, R.A., Smit, M., Jennings, L., et al., 2009. Mixed Infection and the Genesis of Influenza Virus Diversity. J Virol 83, 8832–8841. https://doi.org/10.1128/JVI.00773-09

Giersing, B.K., Vekemans, J., Nava, S., Kaslow, D.C., Moorthy, V., 2019. Report from the World Health Organization's third Product Development for Vaccines Advisory Committee (PDVAC) meeting, Geneva, 8–10th June 2016. Vaccine 37, 7315–7327. https://doi.org/10.1016/j.vaccine.2016.10.090

Glass, P.J., White, L.J., Ball, J.M., Leparc-Goffart, I., Hardy, M.E., Estes, M.K., 2000. Norwalk Virus Open Reading Frame 3 Encodes a Minor Structural Protein. J Virol 74, 6581–6591. https://doi.org/10.1128/JVI.74.14.6581-6591.2000

Glass, R.I., Parashar, U.D., Estes, M.K., 2009. Norovirus Gastroenteritis. N Engl J Med 361, 1776–1785. https://doi.org/10.1056/NEJMra0804575

Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., Ferrin, T.E., 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis: UCSF ChimeraX Visualization System. Protein Science 27, 14–25. https://doi.org/10.1002/pro.3235

Goldhill, D.H., Langat, P., Xie, H., Galiano, M., Miah, S., Kellam, P., Zambon, M., Lackenby, A., Barclay, W.S., 2019. Determining the Mutation Bias of Favipiravir in Influenza Virus Using Next-Generation Sequencing. J Virol 93, e01217-18. https://doi.org/10.1128/JVI.01217-18

Goldhill, D.H., te Velthuis, A.J.W., Fletcher, R.A., Langat, P., Zambon, M., Lackenby, A., Barclay, W.S., 2018. The mechanism of resistance to favipiravir in influenza. Proc Natl Acad Sci USA 115, 11613–11618. https://doi.org/10.1073/pnas.1811345115

Goldstein, R.A., Pollock, D.D., 2017. Sequence entropy of folding and the absolute rate of amino acid substitutions. Nat Ecol Evol 1, 1923–1930. https://doi.org/10.1038/s41559-017-0338-9

Gonçalves, A., Bertrand, J., Ke, R., Comets, E., Lamballerie, X. de, Malvy, D., Pizzorno, A., Terrier, O., Calatrava, M.R., Mentré, F., et al., 2020. Timing of Antiviral Treatment Initiation is Critical to Reduce SARS-CoV-2 Viral Load. CPT: Pharmacometrics & Systems Pharmacology 9, 509–514. https://doi.org/10.1002/psp4.12543

Gong, P., 2021. Structural basis of viral RNA-dependent RNA polymerase nucleotide addition cycle in picornaviruses, in: The Enzymes. Elsevier, pp. 215–233. https://doi.org/10.1016/bs.enz.2021.06.002

González-Parra, G., Dobrovolny, H.M., 2018. Modeling of fusion inhibitor treatment of RSV in African green monkeys. Journal of Theoretical Biology 456, 62–73. https://doi.org/10.1016/j.jtbi.2018.07.029

Goodfellow, I., 2011. The genome-linked protein VPg of vertebrate viruses — a multifaceted protein. Current Opinion in Virology 1, 355–362. https://doi.org/10.1016/j.coviro.2011.09.003

Gordon, C.J., Tchesnokov, E.P., Feng, J.Y., Porter, D.P., Götte, M., 2020. The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. Journal of Biological Chemistry 295, 4773–4779. https://doi.org/10.1074/jbc.AC120.013056

Gordon, C.J., Tchesnokov, E.P., Schinazi, R.F., Götte, M., 2021. Molnupiravir promotes SARS-CoV-2 mutagenesis via the RNA template. Journal of Biological Chemistry 297, 100770. https://doi.org/10.1016/j.jbc.2021.100770

Gorgeis, J., Sizemore, C., Bashey, A., Holland, H.K., Solomon, S.R., Morris, L.E., Solh, M., 2017. Nitazoxanide Is Effective Therapy for Norovirus Gastroenteritis after Hematopoietic Stem Cell Transplantation. Biology of Blood and Marrow Transplantation 23, S197–S198. https://doi.org/10.1016/j.bbmt.2016.12.381

Gottlieb, R.L., Vaca, C.E., Paredes, R., Mera, J., Webb, B.J., Perez, G., Oguchi, G., Ryan, P., Nielsen, B.U., Brown, M., et al., 2022. Early Remdesivir to Prevent Progression to Severe Covid-19 in Outpatients. N Engl J Med 386, 305–315. https://doi.org/10.1056/NEJMoa2116846

Gowen, B.B., Wong, M.-H., Jung, K.-H., Smee, D.F., Morrey, J.D., Furuta, Y., 2010. Efficacy of favipiravir (T-705) and T-1106 pyrazine derivatives in phlebovirus disease models. Antiviral Research 86, 121–127. https://doi.org/10.1016/j.antiviral.2009.10.015

Grassi, S., Arena, V., Cattani, P., Dell'Aquila, M., Liotti, F.M., Sanguinetti, M., Oliva, A., GEMELLI AGAINST COVID-19 group, 2022. SARS-CoV-2 viral load and replication in postmortem examinations. Int J Legal Med 136, 935–939. https://doi.org/10.1007/s00414-021-02753-2

Green, K.Y., Ando, T., Balayan, M.S., Berke, T., Clarke, I.N., Estes, M.K., Matson, D.O., Nakata, S., Neill, J.D., Studdert, M.J., et al., 2000. Taxonomy of the Caliciviruses. J INFECT DIS 181, S322–S330. https://doi.org/10.1086/315591

Green, K.Y., Kaufman, S.S., Nagata, B.M., Chaimongkol, N., Kim, D.Y., Levenson, E.A., Tin, C.M., Yardley, A.B., Johnson, J.A., Barletta, A.B.F., et al., 2020. Human norovirus targets enteroendocrine epithelial cells in the small intestine. Nat Commun 11, 2759. https://doi.org/10.1038/s41467-020-16491-3

Green, S.M., Lambden, P.R., Owen Caul, E., Ashley, C.R., Clarke, I.N., 1995. Capsid diversity in small round-structured viruses: molecular characterization of an antigenically distinct human enteric calicivirus. Virus Research 37, 271–283. https://doi.org/10.1016/0168-1702(95)00041-N

Grenfell, B.T., Pybus, O.G., Gog, J.R., Wood, J.L.N., Daly, J.M., Mumford, J.A., Holmes, E.C., 2004. Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. Science 303, 327–332. https://doi.org/10.1126/science.1090727

Grifoni, A., Weiskopf, D., Ramirez, S.I., Mateus, J., Dan, J.M., Moderbacher, C.R., Rawlings, S.A., Sutherland, A., Premkumar, L., Jadi, R.S., et al., 2020. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. Cell 181, 1489-1501.e15. https://doi.org/10.1016/j.cell.2020.05.015

Guedj, J., Piorkowski, G., Jacquot, F., Madelain, V., Nguyen, T.H.T., Rodallec, A., Gunther, S., Carbonnelle, C., Mentré, F., Raoul, H., et al., 2018. Antiviral efficacy of favipiravir against Ebola virus: A translational study in cynomolgus macaques. PLoS Med 15, e1002535. https://doi.org/10.1371/journal.pmed.1002535

Guo, L., Ren, L., Yang, S., Xiao, M., Chang, D., Yang, F., Dela Cruz, C.S., Wang, Y., Wu, C., Xiao, Y., et al., 2020. Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19). Clinical Infectious Diseases 71, 778–785. https://doi.org/10.1093/cid/ciaa310

Gupta, A., Gonzalez-Rojas, Y., Juarez, E., Crespo Casal, M., Moya, J., Falci, D.R., Sarkis, E., Solis, J., Zheng, H., Scott, N., et al., 2021. Early Treatment for Covid-19 with SARS-CoV-2 Neutralizing Antibody Sotrovimab. N Engl J Med 385, 1941–1950. https://doi.org/10.1056/NEJMoa2107934

Ha, S., Choi, I.-S., Choi, C., Myoung, J., 2016. Infection models of human norovirus: challenges and recent progress. Arch Virol 161, 779–788. https://doi.org/10.1007/s00705-016-2748-4

Hadjadj, J., Yatim, N., Barnabei, L., Corneau, A., Boussier, J., Smith, N., Péré, H., Charbit, B., Bondet, V., Chenevier-Gobeaux, C., et al., 2020. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science 369, 718–724. https://doi.org/10.1126/science.abc6027

Hall, A., 2011. Incidence of Acute Gastroenteritis and Role of Norovirus, Georgia, USA, 2004-2005. Emerg. Infect. Dis. https://doi.org/10.3201/eid1708.101533

Hammond, J., Leister-Tebbe, H., Gardner, A., Abreu, P., Bao, W., Wisemandle, W., Baniecki, M., Hendrick, V.M., Damle, B., Simón-Campos, A., et al., 2022. Oral Nirmatrelvir for High-Risk, Nonhospitalized Adults with Covid-19. N Engl J Med 386, 1397–1408. https://doi.org/10.1056/NEJMoa2118542

Han, M.S., Seong, M.-W., Heo, E.Y., Park, J.H., Kim, N., Shin, S., Cho, S.I., Park, S.S., Choi, E.H., 2020. Sequential analysis of viral load in a neonate and her

mother infected with SARS-CoV-2. Clin. Infect. Dis. https://doi.org/10.1093/cid/ciaa447

Hardy, M.E., 2005. Norovirus protein structure and function. FEMS Microbiology Letters 253, 1–8. https://doi.org/10.1016/j.femsle.2005.08.031

Hasanoglu, I., Korukluoglu, G., Asilturk, D., Cosgun, Y., Kalem, A.K., Altas, A.B., Kayaaslan, B., Eser, F., Kuzucu, E.A., Guner, R., 2021. Higher viral loads in asymptomatic COVID-19 patients might be the invisible part of the iceberg. Infection 49, 117–126. https://doi.org/10.1007/s15010-020-01548-8

Hassan, E., Baldridge, M.T., 2019. Norovirus encounters in the gut: multifaceted interactions and disease outcomes. Mucosal Immunol 12, 1259–1267. https://doi.org/10.1038/s41385-019-0199-4

Hassanipour, S., Arab-Zozani, M., Amani, B., Heidarzad, F., Fathalipour, M., Martinez-de-Hoyo, R., 2021. The efficacy and safety of Favipiravir in treatment of COVID-19: a systematic review and meta-analysis of clinical trials. Sci Rep 11, 11022. https://doi.org/10.1038/s41598-021-90551-6

He, X., Lau, E.H.Y., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y.C., Wong, J.Y., Guan, Y., Tan, X., et al., 2020. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med. https://doi.org/10.1038/s41591-020-0869-5

Heather, J.M., Chain, B., 2016. The sequence of sequencers: The history of sequencing DNA. Genomics 107, 1–8. https://doi.org/10.1016/j.ygeno.2015.11.003

Hemachudha, T., Ugolini, G., Wacharapluesadee, S., Sungkarat, W., Shuangshoti, S., Laothamatas, J., 2013. Human rabies: neuropathogenesis, diagnosis, and management. The Lancet Neurology 12, 498–513. https://doi.org/10.1016/S1474-4422(13)70038-3

Herrera, N.G., Morano, N.C., Celikgil, A., Georgiev, G.I., Malonis, R.J., Lee, J.H., Tong, K., Vergnolle, O., Massimi, A.B., Yen, L.Y., et al., 2021. Characterization of the SARS-CoV-2 S Protein: Biophysical, Biochemical, Structural, and Antigenic Analysis. ACS Omega 6, 85–102. https://doi.org/10.1021/acsomega.0c03512

Hill, K.J., Russell, C.D., Clifford, S., Templeton, K., Mackintosh, C.L., Koch, O., Sutherland, R.K., 2020. The index case of SARS-CoV-2 in Scotland. J. Infect. https://doi.org/10.1016/j.jinf.2020.03.022

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.-H., Nitsche, A., et al., 2020. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 181, 271-280.e8. https://doi.org/10.1016/j.cell.2020.02.052

Hosmillo, M., Chaudhry, Y., Nayak, K., Sorgeloos, F., Koo, B.-K., Merenda, A., Lillestol, R., Drumright, L., Zilbauer, M., Goodfellow, I., 2020. Norovirus Replication in Human Intestinal Epithelial Cells Is Restricted by the Interferon-Induced JAK/STAT Signaling Pathway and RNA Polymerase II-Mediated Transcriptional Responses. mBio 11, e00215-20. https://doi.org/10.1128/mBio.00215-20

Hsu, C.C., Meeker, S.M., Escobar, S., Brabb, T.L., Paik, J., Park, H., Iritani, B.M., Maggio-Price, L., 2018. Murine norovirus inhibits B cell development in the bone marrow of STAT1-deficient mice. Virology 515, 123–133. https://doi.org/10.1016/j.virol.2017.12.013

Hu, Y., Shen, L., Yao, Y., Xu, Z., Zhou, J., Zhou, H., 2020. A report of three COVID-19 cases with prolonged viral RNA detection in anal swabs. Clin. Microbiol. Infect. https://doi.org/10.1016/j.cmi.2020.04.010

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., et al., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet 395, 497–506. https://doi.org/10.1016/S0140-6736(20)30183-5

Huang, Yongbo, Chen, S., Yang, Z., Guan, W., Liu, D., Lin, Z., Zhang, Y., Xu, Z., Liu, X., Li, Y., 2020. SARS-CoV-2 Viral Load in Clinical Samples of Critically Ill Patients. Am. J. Respir. Crit. Care Med. https://doi.org/10.1164/rccm.202003-0572LE

Huang, Y., Yang, C., Xu, X., Xu, W., Liu, S., 2020. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharmacol Sin 41, 1141–1149. https://doi.org/10.1038/s41401-020-0485-4

Huang, Y.-Q., Tang, S.-Q., Xu, X.-L., Zeng, Y.-M., He, X.-Q., Li, Y., Harypursat, V., Lu, Y.-Q., Wan, Y., Zhang, L., et al., 2020. No Statistically Apparent Difference in Antiviral Effectiveness Observed Among Ribavirin Plus Interferon-Alpha, Lopinavir/Ritonavir Plus Interferon-Alpha, and Ribavirin Plus Lopinavir/Ritonavir Plus Interferon-Alpha in Patients With Mild to Moderate Coronavirus Disease 2019: Results of a Randomized, Open-Labeled Prospective Study. Front. Pharmacol. 11, 1071. https://doi.org/10.3389/fphar.2020.01071

Hudgens, M.G., Hoering, A., Self, S.G., 2003a. On the analysis of viral load endpoints in HIV vaccine trials. Statist. Med. 22, 2281–2298. https://doi.org/10.1002/sim.1394

Hudgens, M.G., Hoering, A., Self, S.G., 2003b. On the analysis of viral load endpoints in HIV vaccine trials. Statist. Med. 22, 2281–2298. https://doi.org/10.1002/sim.1394

Illingworth, C.J.R., Raghwani, J., Serwadda, D., Sewankambo, N.K., Robb, M.L., Eller, M.A., Redd, A.R., Quinn, T.C., Lythgoe, K.A., 2020. A de novo approach to inferring within-host fitness effects during untreated HIV-1 infection. PLoS Pathog 16, e1008171. https://doi.org/10.1371/journal.ppat.1008171

International Committee on Taxonomy of Viruses, 2021. ICTV Master Species List.

Irifune, S., Ashizawa, N., Takazono, T., Mutantu, P., Nabeshima, T., Ngwe Tun, M.M., Ota, K., Hirayama, T., Fujita, A., Tashiro, M., et al., 2021. Discrepancy of SARS-CoV-2 PCR results due to the sample collection sites and possible improper sampling. Journal of Infection and Chemotherapy 27, 1525–1528. https://doi.org/10.1016/j.jiac.2021.07.008

Ivanov, K.A., Ziebuhr, J., 2004. Human Coronavirus 229E Nonstructural Protein 13: Characterization of Duplex-Unwinding, Nucleoside Triphosphatase, and RNA

5′-Triphosphatase Activities. J Virol 78, 7833–7838. https://doi.org/10.1128/JVI.78.14.7833-7838.2004

Ivanov, M.A., Ludva, G.S., Mukovnya, A.V., Kochetkov, S.N., Tunitskaya, V.L., Alexandrova, L.A., 2010. Synthesis and biological properties of pyrimidine 4′-fluoronucleosides and 4′-fluorouridine 5′-O-triphosphate. Russ J Bioorg Chem 36, 488–496. https://doi.org/10.1134/S1068162010040072

Jácome, R., Campillo-Balderas, J.A., Ponce de León, S., Becerra, A., Lazcano, A., 2020. Sofosbuvir as a potential alternative to treat the SARS-CoV-2 epidemic. Sci Rep 10, 9294. https://doi.org/10.1038/s41598-020-66440-9

Jayk Bernal, A., Gomes da Silva, M.M., Musungaie, D.B., Kovalchuk, E., Gonzalez, A., Delos Reyes, V., Martín-Quirós, A., Caraco, Y., Williams-Diaz, A., Brown, M.L., et al., 2022. Molnupiravir for Oral Treatment of Covid-19 in Nonhospitalized Patients. N Engl J Med 386, 509–520. https://doi.org/10.1056/NEJMoa2116044

Jia, H., Gong, P., 2019. A Structure-Function Diversity Survey of the RNA-Dependent RNA Polymerases From the Positive-Strand RNA Viruses. Front. Microbiol. 10, 1945. https://doi.org/10.3389/fmicb.2019.01945

Jiang, W., Muhammad, F., Ma, P., Liu, X., Long, G., 2018. Sofosbuvir inhibits hepatitis A virus replication in vitro assessed by a cell-based fluorescent reporter system. Antiviral Research 154, 51–57. https://doi.org/10.1016/j.antiviral.2018.04.007

Jin, Z., Kinkade, A., Behera, I., Chaudhuri, S., Tucker, K., Dyatkina, N., Rajwanshi, V.K., Wang, G., Jekle, A., Smith, D.B., et al., 2017. Structure-activity relationship analysis of mitochondrial toxicity caused by antiviral ribonucleoside analogs. Antiviral Research 143, 151–161. https://doi.org/10.1016/j.antiviral.2017.04.005

Jin, Z., Smith, L.K., Rajwanshi, V.K., Kim, B., Deval, J., 2013. The Ambiguous Base-Pairing and High Substrate Efficiency of T-705 (Favipiravir) Ribofuranosyl 5′-Triphosphate towards Influenza A Virus Polymerase. PLoS ONE 8, e68347. https://doi.org/10.1371/journal.pone.0068347

Jockusch, S., Tao, C., Li, X., Chien, M., Kumar, S., Morozova, I., Kalachikov, S., Russo, J.J., Ju, J., 2020. Sofosbuvir terminated RNA is more resistant to SARS-CoV-2 proofreader than RNA terminated by Remdesivir. Sci Rep 10, 16577. https://doi.org/10.1038/s41598-020-73641-9

Jones, M.K., Grau, K.R., Costantini, V., Kolawole, A.O., de Graaf, M., Freiden, P., Graves, C.L., Koopmans, M., Wallet, S.M., Tibbetts, S.A., et al., 2015. Human norovirus culture in B cells. Nat Protoc 10, 1939–1947. https://doi.org/10.1038/nprot.2015.121

Jones, M.K., Watanabe, M., Zhu, S., Graves, C.L., Keyes, L.R., Grau, K.R., Gonzalez-Hernandez, M.B., Iovine, N.M., Wobus, C.E., Vinjé, J., et al., 2014. Enteric bacteria promote human and mouse norovirus infection of B cells. Science 346, 755–759. https://doi.org/10.1126/science.1257147

Kabinger, F., Stiller, C., Schmitzová, J., Dienemann, C., Kokic, G., Hillen, H.S., Höbartner, C., Cramer, P., 2021. Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis. Nat Struct Mol Biol 28, 740–746. https://doi.org/10.1038/s41594-021-00651-0

Kalil, A.C., Patterson, T.F., Mehta, A.K., Tomashek, K.M., Wolfe, C.R., Ghazaryan, V., Marconi, V.C., Ruiz-Palacios, G.M., Hsieh, L., Kline, S., et al., 2021. Baricitinib plus Remdesivir for Hospitalized Adults with Covid-19. N Engl J Med 384, 795–807. https://doi.org/10.1056/NEJMoa2031994

Kam, K., Yung, C.F., Cui, L., Tzer Pin Lin, R., Mak, T.M., Maiwald, M., Li, J., Chong, C.Y., Nadua, K., Tan, N.W.H., et al., 2020. A Well Infant With Coronavirus Disease 2019 With High Viral Load. Clinical Infectious Diseases ciaa201. https://doi.org/10.1093/cid/ciaa201

Kao, C.C., Singh, P., Ecker, D.J., 2001. De Novo Initiation of Viral RNA-Dependent RNA Synthesis. Virology 287, 251–260. https://doi.org/10.1006/viro.2001.1039

Kapikian, A.Z., Wyatt, R.G., Dolin, R., Thornhill, T.S., Kalica, A.R., Chanock, R.M., 1972. Visualization by Immune Electron Microscopy of a 27-nm Particle Associated with Acute Infectious Nonbacterial Gastroenteritis. J Virol 10, 1075–1081. https://doi.org/10.1128/jvi.10.5.1075-1081.1972

Kaptein, S.J.F., Jacobs, S., Langendries, L., Seldeslachts, L., Ter Horst, S., Liesenborghs, L., Hens, B., Vergote, V., Heylen, E., Barthelemy, K., et al., 2020a. Favipiravir at high doses has potent antiviral activity in SARS-CoV-2-infected hamsters, whereas hydroxychloroquine lacks activity. Proc Natl Acad Sci U S A. https://doi.org/10.1073/pnas.2014441117

Kaptein, S.J.F., Jacobs, S., Langendries, L., Seldeslachts, L., ter Horst, S., Liesenborghs, L., Hens, B., Vergote, V., Heylen, E., Barthelemy, K., et al., 2020b. Favipiravir at high doses has potent antiviral activity in SARS-CoV-2–infected hamsters, whereas hydroxychloroquine lacks activity. Proc Natl Acad Sci USA 117, 26955–26965. https://doi.org/10.1073/pnas.2014441117

Karandikar, U.C., Crawford, S.E., Ajami, N.J., Murakami, K., Kou, B., Ettayebi, K., Papanicolaou, G.A., Jongwutiwes, U., Perales, M.-A., Shia, J., et al., 2016. Detection of human norovirus in intestinal biopsies from immunocompromised transplant patients. Journal of General Virology 97, 2291–2300. https://doi.org/10.1099/jgv.0.000545

Karimzadeh, S., Bhopal, R., Nguyen Tien, H., 2021. Review of infective dose, routes of transmission and outcome of COVID-19 caused by the SARS-COV-2: comparison with other respiratory viruses. Epidemiol. Infect. 149, e96. https://doi.org/10.1017/S0950268821000790

Karst, S.M., 2010. Pathogenesis of Noroviruses, Emerging RNA Viruses. Viruses 2, 748–781. https://doi.org/10.3390/v2030748

Karst, S.M., Wobus, C.E., Lay, M., Davidson, J., Virgin, H.W., 2003. STAT1-Dependent Innate Immunity to a Norwalk-Like Virus. Science 299, 1575–1578. https://doi.org/10.1126/science.1077905

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30, 3059–3066. https://doi.org/10.1093/nar/gkf436

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010

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., 2021. SARS-CoV-2 evolution during treatment of chronic infection. Nature 592, 277–282. https://doi.org/10.1038/s41586-021-03291-y

Kempf, B., Edgar, J.D., Mc Caughey, C., Devlin, L.A., 2017. Nitazoxanide Is an Ineffective Treatment of Chronic Norovirus in Patients With X-Linked Agammaglobulinemia and May Yield False-Negative Polymerase Chain Reaction Findings in Stool Specimens. The Journal of Infectious Diseases 215, 486–487. https://doi.org/10.1093/infdis/jiw497

Kern, C., Schöning, V., Chaccour, C., Hammann, F., 2021. Modeling of SARS-CoV-2 Treatment Effects for Informed Drug Repurposing. Front. Pharmacol. 12, 625678. https://doi.org/10.3389/fphar.2021.625678

Killingley, B., Mann, A.J., Kalinova, M., Boyers, A., Goonawardane, N., Zhou, J., Lindsell, K., Hare, S.S., Brown, J., Frise, R., et al., 2022. Safety, tolerability and viral kinetics during SARS-CoV-2 human challenge in young adults. Nat Med 28, 1031–1041. https://doi.org/10.1038/s41591-022-01780-9

Kim, E.S., Chin, B.S., Kang, C.K., Kim, N.J., Kang, Y.M., Choi, J.P., Oh, D.H., Kim, J.H., Koh, B., Kim, S.E., et al., 2020. Clinical Course and Outcomes of Patients with Severe Acute Respiratory Syndrome Coronavirus 2 Infection: a Preliminary Report of the First 28 Patients from the Korean Cohort Study on COVID-19. J. Korean Med. Sci. 35, e142. https://doi.org/10.3346/jkms.2020.35.e142

Kim, J.Y., Ko, J.-H., Kim, Y., Kim, Y.-J., Kim, J.-M., Chung, Y.-S., Kim, H.M., Han, M.-G., Kim, S.Y., Chin, B.S., 2020. Viral Load Kinetics of SARS-CoV-2 Infection in First Two Patients in Korea. J Korean Med Sci 35, e86. https://doi.org/10.3346/jkms.2020.35.e86

Kim, K., Calabrese, P., Wang, S., Qin, C., Rao, Y., Feng, P., Chen, X.S., 2022. The roles of APOBEC-mediated RNA editing in SARS-CoV-2 mutations, replication and fitness. Sci Rep 12, 14972. https://doi.org/10.1038/s41598-022-19067-x

Kim, K.S., Ejima, K., Ito, Y., Iwanami, S., Ohashi, H., Koizumi, Y., Asai, Y., Nakaoka, S., Watashi, K., Thompson, R.N., et al., 2020. Modelling SARS-CoV-2 Dynamics: Implications for Therapy (preprint). Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2020.03.23.20040493

Kim, L., Liebowitz, D., Lin, K., Kasparek, K., Pasetti, M.F., Garg, S.J., Gottlieb, K., Trager, G., Tucker, S.N., 2018. Safety and immunogenicity of an oral tablet norovirus vaccine, a phase I randomized, placebo-controlled trial. JCI Insight 3, e121077. https://doi.org/10.1172/jci.insight.121077

Kim, S.E., Jeong, H.S., Yu, Y., Shin, S.U., Kim, S., Oh, T.H., Kim, U.J., Kang, S.-J., Jang, H.-C., Jung, S.-I., et al., 2020. Viral kinetics of SARS-CoV-2 in asymptomatic carriers and presymptomatic patients. International Journal of Infectious Diseases 95, 441–443. https://doi.org/10.1016/j.ijid.2020.04.083

Knyazev, S., Tsyvina, V., Melnyk, A., Artyomenko, A., Malygina, T., Porozov, Y.B., Campbell, E., Switzer, W.M., Skums, P., Zelikovsky, A., 2018. CliqueSNV: An Efficient Noise Reduction Technique for Accurate Assembly of ViralVariants from NGS Data (preprint). Bioinformatics. https://doi.org/10.1101/264242

Knyazev, S., Tsyvina, V., Shankar, A., Melnyk, A., Artyomenko, A., Malygina, T., Porozov, Y.B., Campbell, E.M., Switzer, W.M., Skums, P., et al., 2020. CliqueSNV: An Efficient Noise Reduction Technique for Accurate Assembly of Viral Variants from NGS Data. bioRxiv 264242. https://doi.org/10.1101/264242

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., Wilson, R.K., 2012. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Research 22, 568–576. https://doi.org/10.1101/gr.129684.111

Kocher, J., Bui, T., Giri-Rachman, E., Wen, K., Li, G., Yang, X., Liu, F., Tan, M., Xia, M., Zhong, W., et al., 2014. Intranasal P Particle Vaccine Provided Partial Cross-Variant Protection against Human GII.4 Norovirus Diarrhea in Gnotobiotic Pigs. J Virol 88, 9728–9743. https://doi.org/10.1128/JVI.01249-14

Koff, R.S., 2014. Review article: the efficacy and safety of sofosbuvir, a novel, oral nucleotide NS5B polymerase inhibitor, in the treatment of chronic hepatitis C virus infection. Aliment Pharmacol Ther 39, 478–487. https://doi.org/10.1111/apt.12601

Kokic, G., Hillen, H.S., Tegunov, D., Dienemann, C., Seitz, F., Schmitzova, J., Farnung, L., Siewert, A., Höbartner, C., Cramer, P., 2021. Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. Nat Commun 12, 279. https://doi.org/10.1038/s41467-020-20542-0

Kolahchi, Z., De Domenico, M., Uddin, L.Q., Cauda, V., Grossmann, I., Lacasa, L., Grancini, G., Mahmoudi, M., Rezaei, N., 2021. COVID-19 and Its Global Economic Impact, in: Rezaei, N. (Ed.), Coronavirus Disease - COVID-19, Advances in Experimental Medicine and Biology. Springer International Publishing, Cham, pp. 825–837. https://doi.org/10.1007/978-3-030-63761-3_46

Kombe, I.K., Munywoki, P.K., Baguelin, M., Nokes, D.J., Medley, G.F., 2019. Model-based estimates of transmission of respiratory syncytial virus within households. Epidemics 27, 1–11. https://doi.org/10.1016/j.epidem.2018.12.001

Kowdley, K.V., 2005. Hematologic Side Effects of Interferon and Ribavirin Therapy. Journal of Clinical Gastroenterology 39, S3–S8. https://doi.org/10.1097/01.mcg.0000145494.76305.11

Kroneman, A., Vega, E., Vennema, H., Vinjé, J., White, P.A., Hansman, G., Green, K., Martella, V., Katayama, K., Koopmans, M., 2013. Proposal for a unified norovirus nomenclature and genotyping. Arch Virol 158, 2059–2068. https://doi.org/10.1007/s00705-013-1708-5

Kroneman, A., Vennema, H., Deforche, K., Avoort, H. v. d., Peñaranda, S., Oberste, M.S., Vinjé, J., Koopmans, M., 2011. An automated genotyping tool for enteroviruses and noroviruses. Journal of Clinical Virology 51, 121–125. https://doi.org/10.1016/j.jcv.2011.03.006

Ks, X., Lh, M., T, B., Jd, B., 2018. Within-Host Evolution of Human Influenza Virus [WWW Document]. Trends in microbiology. https://doi.org/10.1016/j.tim.2018.02.007

Kumar, S., Thambiraja, T.S., Karuppanan, K., Subramaniam, G., 2022. Omicron and Delta variant of SARS-CoV-2: A comparative computational study of spike protein. Journal of Medical Virology 94, 1641–1649. https://doi.org/10.1002/jmv.27526

Kyriakidis, N.C., López-Cortés, A., González, E.V., Grimaldos, A.B., Prado, E.O., 2021. SARS-CoV-2 vaccines strategies: a comprehensive review of phase 3 candidates. npj Vaccines 6, 28. https://doi.org/10.1038/s41541-021-00292-w

L. Adler, J., Zickl, R., 1969. Winter Vomiting Disease. The Journal of Infectious Diseases 119, 668–673.

La Frazia, S., Ciucci, A., Arnoldi, F., Coira, M., Gianferretti, P., Angelini, M., Belardo, G., Burrone, O.R., Rossignol, J.-F., Santoro, M.G., 2013. Thiazolides, a New Class of Antiviral Agents Effective against Rotavirus Infection, Target Viral Morphogenesis, Inhibiting Viroplasm Formation. J Virol 87, 11096–11106. https://doi.org/10.1128/JVI.01213-13

Lai, C.-C., Wang, Y.-H., Wu, C.-Y., Hung, C.-H., Jiang, D.D.-S., Wu, F.-T., 2013. A norovirus outbreak in a nursing home: Norovirus shedding time associated with age. Journal of Clinical Virology 56, 96–101. https://doi.org/10.1016/j.jcv.2012.10.011

Lai, M.M.C., 2005. RNA Replication without RNA-Dependent RNA Polymerase: Surprises from Hepatitis Delta Virus. J Virol 79, 7951–7958. https://doi.org/10.1128/JVI.79.13.7951-7958.2005

Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., et al., 2020. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 581, 215–220. https://doi.org/10.1038/s41586-020-2180-5

Lateef, Z., Gimenez, G., Baker, E.S., Ward, V.K., 2017. Transcriptomic analysis of human norovirus NS1-2 protein highlights a multifunctional role in murine monocytes. BMC Genomics 18, 39. https://doi.org/10.1186/s12864-016-3417-4

Lau, J.Y.N., Tam, R.C., Liang, T.J., Hong, Z., 2002. Mechanism of action of ribavirin in the combination treatment of chronic HCV infection: Mechanism of action of ribavirin in the combination treatment of chronic HCV infection. Hepatology 35, 1002–1009. https://doi.org/10.1053/jhep.2002.32672

Lay, M.K., Atmar, R.L., Guix, S., Bharadwaj, U., He, H., Neill, F.H., Sastry, K.J., Yao, Q., Estes, M.K., 2010. Norwalk virus does not replicate in human macrophages or dendritic cells derived from the peripheral blood of susceptible humans. Virology 406, 1–11. https://doi.org/10.1016/j.virol.2010.07.001

Le Pendu, J., Ruvoën-Clouet, N., Kindberg, E., Svensson, L., 2006. Mendelian resistance to human norovirus infections. Seminars in Immunology 18, 375–386. https://doi.org/10.1016/j.smim.2006.07.009

Lee, B.E., Pang, X.-L., 2013. New strains of norovirus and the mystery of viral gastroenteritis epidemics. CMAJ 185, 1381–1382. https://doi.org/10.1503/cmaj.130426

Lee, J., Adler, F.R., Kim, P.S., 2017. A Mathematical Model for the Macrophage Response to Respiratory Viral Infection in Normal and Asthmatic Conditions. Bull Math Biol 79, 1979–1998. https://doi.org/10.1007/s11538-017-0315-0

Lei, S., Ryu, J., Wen, K., Twitchell, E., Bui, T., Ramesh, A., Weiss, M., Li, G., Samuel, H., Clark-Deener, S., et al., 2016. Increased and prolonged human norovirus infection in RAG2/IL2RG deficient gnotobiotic pigs with severe combined immunodeficiency. Sci Rep 6, 25222. https://doi.org/10.1038/srep25222

Lescure, F.-X., Bouadma, L., Nguyen, D., Parisey, M., Wicky, P.-H., Behillil, S., Gaymard, A., Bouscambert-Duchamp, M., Donati, F., Le Hingrat, Q., et al., 2020. Clinical and virological data of the first cases of COVID-19 in Europe: a case series. Lancet Infect Dis. https://doi.org/10.1016/S1473-3099(20)30200-0

Li, B., Deng, A., Li, K., Hu, Y., Li, Z., Shi, Y., Xiong, Q., Liu, Z., Guo, Q., Zou, L., et al., 2022. Viral infection and transmission in a large, well-traced outbreak caused by the SARS-CoV-2 Delta variant. Nat Commun 13, 460. https://doi.org/10.1038/s41467-022-28089-y

Li, C.K., Wu, H., Yan, H., Ma, S., Wang, L., Zhang, M., Tang, X., Temperton, N.J., Weiss, R.A., Brenchley, J.M., et al., 2008. T Cell Responses to Whole SARS Coronavirus in Humans. J Immunol 181, 5490–5500. https://doi.org/10.4049/jimmunol.181.8.5490

Li, G., Chen, X., Xu, A., 2003. Profile of Specific Antibodies to the SARS-Associated Coronavirus. N Engl J Med 349, 508–509. https://doi.org/10.1056/NEJM200307313490520

Li, H., Durbin, R., 2009a. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Durbin, R., 2009b. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, T.-F., Hosmillo, M., Schwanke, H., Shu, T., Wang, Z., Yin, L., Curry, S., Goodfellow, I.G., Zhou, X., 2018. Human Norovirus NS3 Has RNA Helicase and Chaperoning Activities. J Virol 92, e01606-17. https://doi.org/10.1128/JVI.01606-17

Li, Y., Cai, H., Rajabalee, N., Au, X., Friedenberg, F., Wallach, S., 2020. S1027 Hepatotoxicity of Remdesivir for COVID-19: Systematic Review and Meta-Analysis. Am J Gastroenterol 115, S523–S523. https://doi.org/10.14309/01.ajg.0000706156.26271.8a

Liao, Y., Xue, L., Gao, J., Wu, A., Kou, X., 2020. ABO blood group-associated susceptibility to norovirus infection: A systematic review and meta-analysis. Infection, Genetics and Evolution 81, 104245. https://doi.org/10.1016/j.meegid.2020.104245

Lim, J., Jeon, S., Shin, H.-Y., Kim, M.J., Seong, Y.M., Lee, W.J., Choe, K.-W., Kang, Y.M., Lee, B., Park, S.-J., 2020. Case of the Index Patient Who Caused Tertiary Transmission of Coronavirus Disease 2019 in Korea: the Application of Lopinavir/Ritonavir for the Treatment of COVID-19 Pneumonia Monitored

by Quantitative RT-PCR. J Korean Med Sci 35, e79. https://doi.org/10.3346/jkms.2020.35.e79

Lin, L., Liu, Y., Tang, X., He, D., 2021. The Disease Severity and Clinical Outcomes of the SARS-CoV-2 Variants of Concern. Front. Public Health 9, 775224. https://doi.org/10.3389/fpubh.2021.775224

Lin, S.-C., Qu, L., Ettayebi, K., Crawford, S.E., Blutt, S.E., Robertson, M.J., Zeng, X.-L., Tenge, V.R., Ayyar, B.V., Karandikar, U.C., et al., 2020. Human norovirus exhibits strain-specific sensitivity to host interferon pathways in human intestinal enteroids. Proc. Natl. Acad. Sci. U.S.A. 117, 23782–23793. https://doi.org/10.1073/pnas.2010834117

Lin, W.-H.W., Kouyos, R.D., Adams, R.J., Grenfell, B.T., Griffin, D.E., 2012. Prolonged persistence of measles virus RNA is characteristic of primary infection dynamics. Proc. Natl. Acad. Sci. U.S.A. 109, 14989–14994. https://doi.org/10.1073/pnas.1211138109

Lin, Y., Fengling, L., Lianzhu, W., Yuxiu, Z., Yanhua, J., 2014. Function of VP2 protein in the stability of the secondary structure of virus-like particles of genogroup II norovirus at different pH levels: Function of VP2 protein in the stability of NoV VLPs. J Microbiol. 52, 970–975. https://doi.org/10.1007/s12275-014-4323-6

Lindesmith, L., Moe, C., LePendu, J., Frelinger, J.A., Treanor, J., Baric, R.S., 2005. Cellular and Humoral Immunity following Snow Mountain Virus Challenge. J Virol 79, 2900–2909. https://doi.org/10.1128/JVI.79.5.2900-2909.2005

Lindesmith, L.C., Brewer-Jensen, P.D., Mallory, M.L., Jensen, K., Yount, B.L., Costantini, V., Collins, M.H., Edwards, C.E., Sheahan, T.P., Vinjé, J., et al., 2020. Virus–Host Interactions Between Nonsecretors and Human Norovirus. Cellular and Molecular Gastroenterology and Hepatology 10, 245–267. https://doi.org/10.1016/j.jcmgh.2020.03.006

Lindesmith, L.C., Donaldson, E., Leon, J., Moe, C.L., Frelinger, J.A., Johnston, R.E., Weber, D.J., Baric, R.S., 2010. Heterotypic Humoral and Cellular Immune Responses following Norwalk Virus Infection. J Virol 84, 1800–1815. https://doi.org/10.1128/JVI.02179-09

Lindesmith, L.C., Donaldson, E.F., Baric, R.S., 2011. Norovirus GII.4 Strain Antigenic Variation. J Virol 85, 231–242. https://doi.org/10.1128/JVI.01364-10

Lindesmith, L.C., Donaldson, E.F., LoBue, A.D., Cannon, J.L., Zheng, D.-P., Vinje, J., Baric, R.S., 2008. Mechanisms of GII.4 Norovirus Persistence in Human Populations. PLoS Med 5, e31. https://doi.org/10.1371/journal.pmed.0050031

Liu, B., Clarke, I.N., Lambden, P.R., 1996. Polyprotein processing in Southampton virus: identification of 3C-like protease cleavage sites by in vitro mutagenesis. J Virol 70, 2605–2610. https://doi.org/10.1128/jvi.70.4.2605-2610.1996

Liu, J., Li, Y., Liu, L., Hu, X., Wang, X., Hu, H., Hu, Z., Zhou, Y., Wang, M., 2020. Infection of human sweat glands by SARS-CoV-2. Cell Discov 6, 84. https://doi.org/10.1038/s41421-020-00229-y

Liu, J., Li, Y., Liu, Q., Yao, Q., Wang, X., Zhang, H., Chen, R., Ren, L., Min, J., Deng, F., et al., 2021. SARS-CoV-2 cell tropism and multiorgan infection. Cell Discov 7, 17. https://doi.org/10.1038/s41421-021-00249-2

Liu, W.-D., Chang, S.-Y., Wang, J.-T., Tsai, M.-J., Hung, C.-C., Hsu, C.-L., Chang, S.-C., 2020. Prolonged virus shedding even after seroconversion in a patient with COVID-19. J. Infect. https://doi.org/10.1016/j.jinf.2020.03.063

Liu, Y., Yan, L.-M., Wan, L., Xiang, T.-X., Le, A., Liu, J.-M., Peiris, M., Poon, L.L.M., Zhang, W., 2020. Viral dynamics in mild and severe cases of COVID-19. The Lancet Infectious Diseases 20, 656–657. https://doi.org/10.1016/S1473-3099(20)30232-2

Liu, Yuanzhi, Zhang, Y., Wang, M., Cheng, A., Yang, Q., Wu, Y., Jia, R., Liu, M., Zhu, D., Chen, S., et al., 2020. Structures and Functions of the 3′ Untranslated Regions of Positive-Sense Single-Stranded RNA Viruses Infecting Humans and Animals. Front. Cell. Infect. Microbiol. 10, 453. https://doi.org/10.3389/fcimb.2020.00453

Lizasoain, A., Tort, L.F.L., García, M., Gómez, M.M., Leite, J.P.G., Miagostovich, M.P., Cristina, J., Berois, M., Colina, R., Victoria, M., 2015. Sewage surveillance reveals the presence of canine GVII norovirus and canine astrovirus in Uruguay. Arch Virol 160, 2839–2843. https://doi.org/10.1007/s00705-015-2571-3

Lo, M.K., Jordan, R., Arvey, A., Sudhamsu, J., Shrivastava-Ranjan, P., Hotard, A.L., Flint, M., McMullan, L.K., Siegel, D., Clarke, M.O., et al., 2017. GS-5734 and its parent nucleoside analog inhibit Filo-, Pneumo-, and Paramyxoviruses. Sci Rep 7, 43395. https://doi.org/10.1038/srep43395

Long, S., 2021. SARS-CoV-2 Subgenomic RNAs: Characterization, Utility, and Perspectives. Viruses 13, 1923. https://doi.org/10.3390/v13101923

Lopman, B., Gastañaduy, P., Park, G.W., Hall, A.J., Parashar, U.D., Vinjé, J., 2012. Environmental transmission of norovirus gastroenteritis. Current Opinion in Virology 2, 96–102. https://doi.org/10.1016/j.coviro.2011.11.005

Lopman, B.A., Adak, G.K., Reacher, M., Brown, D.W.G., 2003. Two Epidemiologic Patterns of *Norovirus* Outbreaks: Surveillance in England and Wales, 1992–2000. Emerg. Infect. Dis. 9, 71–77. https://doi.org/10.3201/eid0901.020175

Lopman, B.A., Steele, D., Kirkwood, C.D., Parashar, U.D., 2016. The Vast and Varied Global Burden of Norovirus: Prospects for Prevention and Control. PLoS Med 13, e1001999. https://doi.org/10.1371/journal.pmed.1001999

Ludwig, A., Adams, O., Laws, H.-J., Schroten, H., Tenenbaum, T., 2008. Quantitative detection of norovirus excretion in pediatric patients with cancer and prolonged gastroenteritis and shedding of norovirus. J. Med. Virol. 80, 1461–1467. https://doi.org/10.1002/jmv.21217

Lui, G., Ling, L., Lai, C.K., Tso, E.Y., Fung, K.S., Chan, V., Ho, T.H., Luk, F., Chen, Z., Ng, J.K., et al., 2020. Viral dynamics of SARS-CoV-2 across a spectrum of disease severity in COVID-19. J. Infect. https://doi.org/10.1016/j.jinf.2020.04.014

Lumby, C.K., Zhao, L., Breuer, J., Illingworth, C.J., 2020a. A large effective population size for established within-host influenza virus infection. eLife 9, e56915. https://doi.org/10.7554/eLife.56915

Lumby, C.K., Zhao, L., Oporto, M., Best, T., Tutill, H., Shah, D., Veys, P., Williams, R., Worth, A., Illingworth, C.J.R., et al., 2020b. Favipiravir and Zanamivir Cleared Infection with Influenza B in a Severely Immunocompromised Child.

Clinical Infectious Diseases 71, e191–e194. https://doi.org/10.1093/cid/ciaa023

Lun, J., Hewitt, J., Yan, G., Enosi Tuipulotu, D., Rawlinson, W., White, P., 2018. Recombinant GII.P16/GII.4 Sydney 2012 Was the Dominant Norovirus Identified in Australia and New Zealand in 2017. Viruses 10, 548. https://doi.org/10.3390/v10100548

Lun, J.H., Hewitt, J., Sitabkhan, A., Eden, J.-S., Enosi Tuipulotu, D., Netzler, N.E., Morrell, L., Merif, J., Jones, R., Huang, B., et al., 2018. Emerging recombinant noroviruses identified by clinical and waste water screening. Emerging Microbes & Infections 7, 1–14. https://doi.org/10.1038/s41426-018-0047-8

Lythgoe, K.A., Hall, M., Ferretti, L., Cesare, M. de, MacIntyre-Cockett, G., Trebes, A., Andersson, M., Otecko, N., Wise, E.L., Moore, N., et al., 2020. Shared SARS-CoV-2 diversity suggests localised transmission of minority variants. bioRxiv 2020.05.28.118992. https://doi.org/10.1101/2020.05.28.118992

Lythgoe, K.A., Hall, M., Ferretti, L., de Cesare, M., MacIntyre-Cockett, G., Trebes, A., Andersson, M., Otecko, N., Wise, E.L., Moore, N., et al., 2021. SARS-CoV-2 within-host diversity and transmission. Science 372, eabg0821. https://doi.org/10.1126/science.abg0821

Maag, D., Castro, C., Hong, Z., Cameron, C.E., 2001. Hepatitis C Virus RNA-dependent RNA Polymerase (NS5B) as a Mediator of the Antiviral Activity of Ribavirin. Journal of Biological Chemistry 276, 46094–46098. https://doi.org/10.1074/jbc.C100349200

Madelain, V., Oestereich, L., Graw, F., Nguyen, T.H.T., de Lamballerie, X., Mentré, F., Günther, S., Guedj, J., 2015. Ebola virus dynamics in mice treated with favipiravir. Antiviral Research 123, 70–77. https://doi.org/10.1016/j.antiviral.2015.08.015

Maksimov, I., Sorokan, A., Burkhanova, G., Veselova, S., Alekseev, V., Shein, M., Avalbaev, A., Dhaware, P., Mehetre, G., Singh, B., et al., 2019. Mechanisms of Plant Tolerance to RNA Viruses Induced by Plant-Growth-Promoting Microorganisms. Plants 8, 575. https://doi.org/10.3390/plants8120575

Malinoski, F., Stollar, V., 1981. Inhibitors of IMP dehydrogenase prevent sindbis virus replication and reduce GTP levels in Aedes albopictus cells. Virology 110, 281–291. https://doi.org/10.1016/0042-6822(81)90060-X

Malpica, J.M., Fraile, A., Moreno, I., Obies, C.I., Drake, J.W., García-Arenal, F., 2002. The Rate and Character of Spontaneous Mutation in an RNA Virus. Genetics 162, 1505–1511. https://doi.org/10.1093/genetics/162.4.1505

Mans, J., Armah, G.E., Steele, A.D., Taylor, M.B., 2016. Norovirus Epidemiology in Africa: A Review. PLoS ONE 11, e0146280. https://doi.org/10.1371/journal.pone.0146280

Marconi, V.C., Ramanan, A.V., de Bono, S., Kartman, C.E., Krishnan, V., Liao, R., Piruzeli, M.L.B., Goldman, J.D., Alatorre-Alexander, J., de Cassia Pellegrini, R., et al., 2021. Efficacy and safety of baricitinib for the treatment of hospitalised adults with COVID-19 (COV-BARRIER): a randomised, double-blind, parallel-group, placebo-controlled phase 3 trial. The Lancet Respiratory Medicine 9, 1407–1418. https://doi.org/10.1016/S2213-2600(21)00331-3

Marionneau, S., Ruvoën, N., Le Moullac–Vaidye, B., Clement, M., Cailleau–Thomas, A., Ruiz–Palacois, G., Huang, P., Jiang, X., Le Pendu, J., 2002. Norwalk virus binds to histo-blood group antigens present on gastroduodenal epithelial cells of secretor individuals. Gastroenterology 122, 1967–1977. https://doi.org/10.1053/gast.2002.33661

Martin, L., Hutchens, M., Hawkins, C., Radnov, A., 2017. How much do clinical trials cost? Nat Rev Drug Discov 16, 381–382. https://doi.org/10.1038/nrd.2017.70

Martinez, M.A., 2022. What Should Be Learned From Repurposed Antivirals Against SARS-CoV-2? Front. Microbiol. 13, 843587. https://doi.org/10.3389/fmicb.2022.843587

Matsushima, Y., Ishikawa, M., Shimizu, T., Komane, A., Kasuo, S., Shinohara, M., Nagasawa, K., Kimura, H., Ryo, A., Okabe, N., et al., 2015. Genetic analyses of GII.17 norovirus strains in diarrheal disease outbreaks from December 2014 to March 2015 in Japan reveal a novel polymerase sequence and amino acid substitutions in the capsid region. Eurosurveillance 20. https://doi.org/10.2807/1560-7917.ES2015.20.26.21173

Mboko, W.P., Chhabra, P., Valcarce, M.D., Costantini, V., Vinjé, J., 2022. Advances in understanding of the innate immune response to human norovirus infection using organoid models. Journal of General Virology 103. https://doi.org/10.1099/jgv.0.001720

McAloon, C., Collins, Á., Hunt, K., Barber, A., Byrne, A.W., Butler, F., Casey, M., Griffin, J., Lane, E., McEvoy, D., et al., 2020. Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. BMJ Open 10, e039652. https://doi.org/10.1136/bmjopen-2020-039652

McCrone, J.T., Woods, R.J., Martin, E.T., Malosh, R.E., Monto, A.S., Lauring, A.S., 2017. Stochastic processes dominate the within and between host evolution of influenza virus (preprint). Microbiology. https://doi.org/10.1101/176362

Mendes, É.A., Pilger, D.R.B. de, Santos Nastri, A.C. de S., Malta, F. de M., Pascoalino, B. dos S., Carneiro D'Albuquerque, L.A., Balan, A., Freitas, L.H.G. de, Durigon, E.L., Carrilho, F.J., et al., 2019. Sofosbuvir inhibits yellow fever virus in vitro and in patients with acute liver failure. Annals of Hepatology 18, 816–824. https://doi.org/10.1016/j.aohep.2019.09.001

Menni, C., Valdes, A.M., Freidin, M.B., Sudre, C.H., Nguyen, L.H., Drew, D.A., Ganesh, S., Varsavsky, T., Cardoso, M.J., El-Sayed Moustafa, J.S., et al., 2020. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med 26, 1037–1040. https://doi.org/10.1038/s41591-020-0916-2

Mesquita, J.R., Barclay, L., Nascimento, M.S.J., Vinjé, J., 2010. Novel Norovirus in Dogs with Diarrhea. Emerg. Infect. Dis. 16, 980–982. https://doi.org/10.3201/eid1606.091861

Michael Rajnik, Marco Cascella, Arturo Cuomo, Scott C. Dulebohn, Raffaela Di Napoli, 2021. Features, Evaluation, and Treatment of Coronavirus (COVID-19). Uniformed Services University Of The Health Sciences.

Michot, J.-M., Albiges, L., Chaput, N., Saada, V., Pommeret, F., Griscelli, F., Balleyguier, C., Besse, B., Marabelle, A., Netzer, F., et al., 2020. Tocilizumab, an anti-IL-6 receptor antibody, to treat COVID-19-related respiratory failure: a

case report. Annals of Oncology 31, 961–964. https://doi.org/10.1016/j.annonc.2020.03.300

Mittal, L., Kumari, A., Suri, C., Bhattacharya, S., Asthana, S., 2019. Insights into structural dynamics of allosteric binding sites in HCV RNA-dependent RNA polymerase. Journal of Biomolecular Structure and Dynamics 1–14. https://doi.org/10.1080/07391102.2019.1614480

Modrow, S., Falke, D., Truyen, U., Schätzl, H., 2013. Viruses with Single-Stranded, Positive-Sense RNA Genomes, in: Molecular Virology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 185–349. https://doi.org/10.1007/978-3-642-20718-1_14

Mönttinen, H.A.M., Ravantti, J.J., Stuart, D.I., Poranen, M.M., 2014. Automated Structural Comparisons Clarify the Phylogeny of the Right-Hand-Shaped Polymerases. Molecular Biology and Evolution 31, 2741–2752. https://doi.org/10.1093/molbev/msu219

Morrey, J., Taro, B., Siddharthan, V., Wang, H., Smee, D., Christensen, A., Furuta, Y., 2008. Efficacy of orally administered T-705 pyrazine analog on lethal West Nile virus infection in rodents. Antiviral Research 80, 377–379. https://doi.org/10.1016/j.antiviral.2008.07.009

Morris, J., Morris, C., 2015. Nitazoxanide Is Effective Therapy for Norovirus Gastroenteritis after Chemotherapy and Hematopoietic Stem Cell Transplantation (HSCT). Biology of Blood and Marrow Transplantation 21, S255–S256. https://doi.org/10.1016/j.bbmt.2014.11.405

Mourier, T., Sadykov, M., Carr, M.J., Gonzalez, G., Hall, W.W., Pain, A., 2021. Host-directed editing of the SARS-CoV-2 genome. Biochemical and Biophysical Research Communications 538, 35–39. https://doi.org/10.1016/j.bbrc.2020.10.092

Mousavizadeh, L., Ghasemi, S., 2021. Genotype and phenotype of COVID-19: Their roles in pathogenesis. Journal of Microbiology, Immunology and Infection 54, 159–163. https://doi.org/10.1016/j.jmii.2020.03.022

Moya, A., Elena, S.F., Bracho, A., Miralles, R., Barrio, E., 2000. The evolution of RNA viruses: A population genetics view. Proc. Natl. Acad. Sci. U.S.A. 97, 6967–6973. https://doi.org/10.1073/pnas.97.13.6967

Moya, A., Holmes, E.C., González-Candelas, F., 2004. The population genetics and evolutionary epidemiology of RNA viruses. Nat Rev Microbiol 2, 279–288. https://doi.org/10.1038/nrmicro863

Murata, T., Katsushima, N., Mizuta, K., Muraki, Y., Hongo, S., Matsuzaki, Y., 2007. Prolonged Norovirus Shedding in Infants ≤6 Months of Age With Gastroenteritis. Pediatric Infectious Disease Journal 26, 46–49. https://doi.org/10.1097/01.inf.0000247102.04997.e0

Nakamura, K., Someya, Y., Kumasaka, T., Ueno, G., Yamamoto, M., Sato, T., Takeda, N., Miyamura, T., Tanaka, N., 2005. A Norovirus Protease Structure Provides Insights into Active and Substrate Binding Site Integrity. J Virol 79, 13685–13693. https://doi.org/10.1128/JVI.79.21.13685-13693.2005

Nakayama, M., Ueda, Y., Kawamoto, H., Han-jun, Y., Saito, K., Nishio, O., Ushijima, H., 1996. Detection and Sequencing of Norwalk-Like Viruses from Stool Samples in Japan Using Reverse Transcription-Polymerase Chain Reaction

Amplification. Microbiology and Immunology 40, 317–320. https://doi.org/10.1111/j.1348-0421.1996.tb03343.x

Naqvi, A.A.T., Fatima, K., Mohammad, T., Fatima, U., Singh, I.K., Singh, A., Atif, S.M., Hariprasad, G., Hasan, G.M., Hassan, Md.I., 2020. Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease 1866, 165878. https://doi.org/10.1016/j.bbadis.2020.165878

National Health Commission of the People's Republic of China, 2020. Diagnosis and treatment plan of Corona virus disease 2019, 5th Edition. [WWW Document]. URL http://www.nhc.gov.cn/yzygj/s7653p/202002/3b09b894ac9b4204a79db5b8912d4440.shtml

National Vaccine and Serum Institute, China, 2021. A Randomized, Blind, Placebo-controlled Phase I Clinical Trial to Evaluate the Safety and Immunogenicity of Recombinant Norovirus Bivalent (GI. 1 / GII. 4) Vaccine (Hansenulapolymorpha) in Healthy People Aged 6 Months to 59 Years (Clinical trial registration No. NCT04188691). clinicaltrials.gov.

Natori, Y., Alghamdi, A., Tazari, M., Miller, V., Husain, S., Komatsu, T., Griffiths, P., Ljungman, P., Orchanian-Cheff, A., Kumar, D., et al., 2018. Use of Viral Load as a Surrogate Marker in Clinical Studies of Cytomegalovirus in Solid Organ Transplantation: A Systematic Review and Meta-analysis. Clinical Infectious Diseases 66, 617–631. https://doi.org/10.1093/cid/cix793

Nehme, M., Braillard, O., Alcoba, G., Aebischer Perone, S., Courvoisier, D., Chappuis, F., Guessous, I., 2021. COVID-19 Symptoms: Longitudinal Evolution and Persistence in Outpatient Settings. Ann Intern Med 174, 723–725. https://doi.org/10.7326/M20-5926

Nelson, C.W., Otto, S.P., 2021. Mutagenic antivirals: the evolutionary risk of low doses. URL https://virological.org/t/mutagenic-antivirals-the-evolutionary-risk-of-low-doses/768

Netzler, N.E., Enosi Tuipulotu, D., White, P.A., 2019. Norovirus antivirals: Where are we now? Med Res Rev 39, 860–886. https://doi.org/10.1002/med.21545

Newman, K.L., Moe, C.L., Kirby, A.E., Flanders, W.D., Parkos, C.A., Leon, J.S., 2016. Norovirus in symptomatic and asymptomatic individuals: cytokines and viral shedding. Clinical and Experimental Immunology 184, 347–357. https://doi.org/10.1111/cei.12772

Ng, K.K.-S., Arnold, J.J., Cameron, C.E., 2008. Structure-Function Relationships Among RNA-Dependent RNA Polymerases, in: Paddison, P.J., Vogt, P.K. (Eds.), RNA Interference, Current Topics in Microbiology and Immunology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 137–156. https://doi.org/10.1007/978-3-540-75157-1_7

NHS, 2022. Coronavirus (COVID-19) vaccine [WWW Document]. nhs.uk. URL https://www.nhs.uk/conditions/coronavirus-covid-19/coronavirus-vaccination/coronavirus-vaccine/ (accessed 5.22.22).

Niendorf, S., Jacobsen, S., Faber, M., Eis-Hübinger, A.M., Hofmann, J., Zimmermann, O., Höhne, M., Bock, C.T., 2017. Steep rise in norovirus cases and emergence of a new recombinant strain GII.P16-GII.2, Germany, winter 2016.

Euro Surveill 22, 30447. https://doi.org/10.2807/1560-7917.ES.2017.22.4.30447

Nyström, K., Waldenström, J., Tang, K.-W., Lagging, M., 2019. Ribavirin: pharmacology, multiple modes of action and possible future perspectives. Future Virology 14, 153–160. https://doi.org/10.2217/fvl-2018-0166

Oestereich, L., Lüdtke, A., Wurr, S., Rieger, T., Muñoz-Fontela, C., Günther, S., 2014. Successful treatment of advanced Ebola virus infection with T-705 (favipiravir) in a small animal model. Antiviral Research 105, 17–21. https://doi.org/10.1016/j.antiviral.2014.02.014

Office for National Statistics, 2022. Self-reported long COVID after infection with the Omicron variant in the UK: 6 May 2022.

Okonechnikov, K., Conesa, A., García-Alcalde, F., 2015. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics btv566. https://doi.org/10.1093/bioinformatics/btv566

Omrani, A.S., Pathan, S.A., Thomas, S.A., Harris, T.R.E., Coyle, P.V., Thomas, C.E., Qureshi, I., Bhutta, Z.A., Mawlawi, N.A., Kahlout, R.A., et al., 2020. Randomized double-blinded placebo-controlled trial of hydroxychloroquine with or without azithromycin for virologic cure of non-severe Covid-19. EClinicalMedicine 29–30, 100645. https://doi.org/10.1016/j.eclinm.2020.100645

Oran, D.P., Topol, E.J., 2020. Prevalence of Asymptomatic SARS-CoV-2 Infection: A Narrative Review. Annals of Internal Medicine 173, 362–367. https://doi.org/10.7326/M20-3012

Ortega-Prieto, A.M., Sheldon, J., Grande-Pérez, A., Tejero, H., Gregori, J., Quer, J., Esteban, J.I., Domingo, E., Perales, C., 2013. Extinction of Hepatitis C Virus by Ribavirin in Hepatoma Cells Involves Lethal Mutagenesis. PLoS ONE 8, e71039. https://doi.org/10.1371/journal.pone.0071039

O'Toole, Aine, 2020. Local Lineage and Monophyly Assessment.

Painter, G.R., Bowen, R.A., Bluemling, G.R., DeBergh, J., Edpuganti, V., Gruddanti, P.R., Guthrie, D.B., Hager, M., Kuiper, D.L., Lockwood, M.A., et al., 2019. The prophylactic and therapeutic activity of a broadly active ribonucleoside analog in a murine model of intranasal venezuelan equine encephalitis virus infection. Antiviral Research 171, 104597. https://doi.org/10.1016/j.antiviral.2019.104597

Pan, Y., Zhang, D., Yang, P., Poon, L.L.M., Wang, Q., 2020. Viral load of SARS-CoV-2 in clinical samples. The Lancet Infectious Diseases 20, 411–412. https://doi.org/10.1016/S1473-3099(20)30113-4

Pandey, A., Nikam, A.N., Shreya, A.B., Mutalik, S.P., Gopalan, D., Kulkarni, S., Padya, B.S., Fernandes, G., Mutalik, S., Prassl, R., 2020. Potential therapeutic targets for combating SARS-CoV-2: Drug repurposing, clinical trials and recent advancements. Life Sciences 256, 117883. https://doi.org/10.1016/j.lfs.2020.117883

Pang, J., Slyker, J.A., Roy, S., Bryant, J., Atkinson, C., Cudini, J., Farquhar, C., Griffiths, P., Kiarie, J., Morfopoulou, S., et al., 2020a. Mixed cytomegalovirus genotypes in HIV positive mothers show compartmentalization and distinct

patterns of transmission to infants. (preprint). Genetic and Genomic Medicine. https://doi.org/10.1101/2020.09.17.20196790

Pang, J., Venturini, C., Tamuri, A.U., Roy, S., Breuer, J., Goldstein, R.A., 2020b. Haplotype assignment of longitudinal viral deep-sequencing data using co-variation of variant frequencies (preprint). Genomics. https://doi.org/10.1101/444877

Pangburn, M.K., Ferreira, V.P., Cortes, C., 2008. Discrimination between host and pathogens by the complement system. Vaccine 26, I15–I21. https://doi.org/10.1016/j.vaccine.2008.11.023

Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20, 289–290. https://doi.org/10.1093/bioinformatics/btg412

Parker, M.D., Lindsey, B.B., Leary, S., Gaudieri, S., Chopra, A., Wyles, M., Angyal, A., Green, L.R., Parsons, P., Tucker, R.M., et al., 2021. Subgenomic RNA identification in SARS-CoV-2 genomic sequencing data. Genome Res. 31, 645–658. https://doi.org/10.1101/gr.268110.120

Parker, M.D., Lindsey, B.B., Leary, S., Gaudieri, S., Chopra, A., Wyles, M., Angyal, A., Green, L.R., Parsons, P., Tucker, R.M., et al., 2020. periscope: sub-genomic RNA identification in SARS-CoV-2 ARTIC Network Nanopore Sequencing Data (preprint). Bioinformatics. https://doi.org/10.1101/2020.07.01.181867

Parra, G.I., 2019. Emergence of norovirus strains: A tale of two genes. Virus Evol 5, vez048. https://doi.org/10.1093/ve/vez048

Parra, G.I., Green, K.Y., 2015. Genome of Emerging Norovirus GII.17, United States, 2014. Emerg Infect Dis 21, 1477–1479. https://doi.org/10.3201/eid2108.150652

Parra, G.I., Squires, R.B., Karangwa, C.K., Johnson, J.A., Lepore, C.J., Sosnovtsev, S.V., Green, K.Y., 2017. Static and Evolving Norovirus Genotypes: Implications for Epidemiology and Immunity. PLoS Pathog 13, e1006136. https://doi.org/10.1371/journal.ppat.1006136

Parrino, T.A., Schreiber, D.S., Trier, J.S., Kapikian, A.Z., Blacklow, N.R., 1977. Clinical Immunity in Acute Gastroenteritis Caused by Norwalk Agent. N Engl J Med 297, 86–89. https://doi.org/10.1056/NEJM197707142970204

Patel, K., Kirkpatrick, C.M., Nieforth, K.A., Chanda, S., Zhang, Q., McClure, M., Fry, J., Symons, J.A., Blatt, L.M., Beigelman, L., et al., 2019. Respiratory syncytial virus-A dynamics and the effects of lumicitabine, a nucleoside viral replication inhibitor, in experimentally infected humans. Journal of Antimicrobial Chemotherapy 74, 442–452. https://doi.org/10.1093/jac/dky415

Patel, M.C., Chesnokov, A., Jones, J., Mishin, V.P., De La Cruz, J.A., Nguyen, H.T., Zanders, N., Wentworth, D.E., Davis, T.C., Gubareva, L.V., 2021. Susceptibility of widely diverse influenza a viruses to PB2 polymerase inhibitor pimodivir. Antiviral Research 188, 105035. https://doi.org/10.1016/j.antiviral.2021.105035

Patel, T.K., Patel, P.B., Barvaliya, M., Saurabh, M.K., Bhalla, H.L., Khosla, P.P., 2021. Efficacy and safety of lopinavir-ritonavir in COVID-19: A systematic review of

randomized controlled trials. Journal of Infection and Public Health 14, 740–748. https://doi.org/10.1016/j.jiph.2021.03.015

Payne, S., 2017. Introduction to RNA Viruses, in: Viruses. Elsevier, pp. 97–105. https://doi.org/10.1016/B978-0-12-803109-4.00010-6

Peavy, D.L., Powers, C.N., Knight, V., 1981. Inhibition of murine plaque-forming cell responses in vivo by ribavirin. J Immunol 126, 861–864.

Peck, K.M., Lauring, A.S., 2018. Complexities of Viral Mutation Rates. J Virol 92, e01031-17. https://doi.org/10.1128/JVI.01031-17

Pelizzola, M., Behr, M., Li, H., Munk, A., Futschik, A., 2021. Multiple haplotype reconstruction from allele frequency data. Nat Comput Sci 1, 262–271. https://doi.org/10.1038/s43588-021-00056-5

Peng, Y., Mentzer, A.J., Liu, G., Yao, X., Yin, Z., Dong, D., Dejnirattisai, W., Rostron, T., Supasa, P., Liu, C., et al., 2020. Broad and strong memory CD4+ and CD8+ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. Nat Immunol 21, 1336–1345. https://doi.org/10.1038/s41590-020-0782-6

Perales, C., Agudo, R., Domingo, E., 2009. Counteracting Quasispecies Adaptability: Extinction of a Ribavirin-Resistant Virus Mutant by an Alternative Mutagenic Treatment. PLoS ONE 4, e5554. https://doi.org/10.1371/journal.pone.0005554

Pérez-Carmona, N., Martínez-Vicente, P., Farré, D., Gabaev, I., Messerle, M., Engel, P., Angulo, A., 2018. A Prominent Role of the Human Cytomegalovirus UL8 Glycoprotein in Restraining Proinflammatory Cytokine Production by Myeloid Cells at Late Times during Infection. J Virol 92, e02229-17. https://doi.org/10.1128/JVI.02229-17

Perlman, S., Netland, J., 2009. Coronaviruses post-SARS: update on replication and pathogenesis. Nat Rev Microbiol 7, 439–450. https://doi.org/10.1038/nrmicro2147

Perry, J.W., Taube, S., Wobus, C.E., 2009. Murine norovirus-1 entry into permissive macrophages and dendritic cells is pH-independent. Virus Research 143, 125–129. https://doi.org/10.1016/j.virusres.2009.03.002

Petrignani, M., Verhoef, L., de Graaf, M., Richardus, J.H., Koopmans, M., 2018. Chronic sequelae and severe complications of norovirus infection: A systematic review of literature. Journal of Clinical Virology 105, 1–10. https://doi.org/10.1016/j.jcv.2018.05.004

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., Ferrin, T.E., 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Science 30, 70–82. https://doi.org/10.1002/pro.3943

Phoswa, W.N., Khaliq, O.P., 2020. Is pregnancy a risk factor of COVID-19? European Journal of Obstetrics & Gynecology and Reproductive Biology 252, 605–609. https://doi.org/10.1016/j.ejogrb.2020.06.058

Picarazzi, F., Vicenti, I., Saladini, F., Zazzi, M., Mori, M., 2020. Targeting the RdRp of Emerging RNA Viruses: The Structure-Based Drug Design Challenge. Molecules 25, 5695. https://doi.org/10.3390/molecules25235695

Piret, J., Boivin, G., 2021. Pandemics Throughout History. Front. Microbiol. 11, 631736. https://doi.org/10.3389/fmicb.2020.631736

Piscoya, A., Ng-Sueng, L.F., Parra del Riego, A., Cerna-Viacava, R., Pasupuleti, V., Roman, Y.M., Thota, P., White, C.M., Hernandez, A.V., 2020. Efficacy and harms of remdesivir for the treatment of COVID-19: A systematic review and meta-analysis. PLoS ONE 15, e0243705. https://doi.org/10.1371/journal.pone.0243705

Pollock, D.D., Thiltgen, G., Goldstein, R.A., 2012. Amino acid coevolution induces an evolutionary Stokes shift. Proc. Natl. Acad. Sci. U.S.A. 109. https://doi.org/10.1073/pnas.1120084109

Poltronieri, P., Sun, B., Mallardo, M., 2015. RNA Viruses: RNA Roles in Pathogenesis, Coreplication and Viral Load. CG 16, 327–335. https://doi.org/10.2174/1389202916666150707160613

Poon, L.L.M., Song, T., Rosenfeld, R., Lin, X., Rogers, M.B., Zhou, B., Sebra, R., Halpin, R.A., Guan, Y., Twaddle, A., et al., 2016. Quantifying influenza virus diversity and transmission in humans. Nat Genet 48, 195–200. https://doi.org/10.1038/ng.3479

Prabhakaran, S., Rey, M., Zagordi, O., Beerenwinkel, N., Roth, V., 2014. HIV Haplotype Inference Using a Propagating Dirichlet Process Mixture Model. IEEE/ACM Transactions on Computational Biology and Bioinformatics 11, 182–191. https://doi.org/10.1109/TCBB.2013.145

Prasad, B.V.V., Hardy, M.E., Dokland, T., Bella, J., Rossmann, M.G., Estes, M.K., 1999. X-ray Crystallographic Structure of the Norwalk Virus Capsid. Science 286, 287–290. https://doi.org/10.1126/science.286.5438.287

Public Health England, 2020. Investigation of novel SARS-CoV-2 variant Variant of Concern 202012/01. London, United Kingdom.

Pulido-Tamayo, S., Sánchez-Rodríguez, A., Swings, T., Van den Bergh, B., Dubey, A., Steenackers, H., Michiels, J., Fostier, J., Marchal, K., 2015. Frequency-based haplotype reconstruction from deep sequencing data of bacterial populations. Nucleic Acids Res 43, e105–e105. https://doi.org/10.1093/nar/gkv478

Putri, W.C.W.S., Muscatello, D.J., Stockwell, M.S., Newall, A.T., 2018. Economic burden of seasonal influenza in the United States. Vaccine 36, 3960–3966. https://doi.org/10.1016/j.vaccine.2018.05.057

Qian, G.-Q., Chen, X.-Q., Lv, D.-F., Ma, A.H.Y., Wang, L.-P., Yang, N.-B., Chen, X.-M., 2020. Duration of SARS-CoV-2 viral shedding during COVID-19 infection. Infect Dis (Lond) 1–2. https://doi.org/10.1080/23744235.2020.1748705

Qing, E., Gallagher, T., 2020. SARS Coronavirus Redux. Trends in Immunology 41, 271–273. https://doi.org/10.1016/j.it.2020.02.007

R Core Team, 2019. R: A language and environment for statistical computing.

Rambaut, A., Holmes, E.C., O'Toole, Á., Hill, V., McCrone, J.T., Ruis, C., du Plessis, L., Pybus, O.G., 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nature Microbiology 5, 1403–1407. https://doi.org/10.1038/s41564-020-0770-5

Rampersad, S., Tennant, P., 2018. Replication and Expression Strategies of Viruses, in: Viruses. Elsevier, pp. 55–82. https://doi.org/10.1016/B978-0-12-811257-1.00003-6

Raoofi, A., Takian, A., Haghighi, H., Rajizadeh, A., Rezaei, Z., Radmerikhi, S., Olyaeemanesh, A., Akbari Sari, A., 2021. COVID-19 and Comparative Health Policy Learning; the Experience of 10 Countries. Arch Iran Med 24, 260–272. https://doi.org/10.34172/aim.2021.37

Raveendran, A.V., Jayadevan, R., Sashidharan, S., 2021. Long COVID: An overview. Diabetes & Metabolic Syndrome: Clinical Research & Reviews 15, 869–875. https://doi.org/10.1016/j.dsx.2021.04.007

Reeck, A., Kavanagh, O., Estes, M.K., Opekun, A.R., Gilger, M.A., Graham, D.Y., Atmar, R.L., 2010. Serological Correlate of Protection against Norovirus-Induced Gastroenteritis. J INFECT DIS 202, 1212–1218. https://doi.org/10.1086/656364

Rehermann, B., Nascimbeni, M., 2005. Immunology of hepatitis B virus and hepatitis C virus infection. Nat Rev Immunol 5, 215–229. https://doi.org/10.1038/nri1573

Reis, G., Moreira Silva, E.A. dos S., Medeiros Silva, D.C., Thabane, L., Singh, G., Park, J.J.H., Forrest, J.I., Harari, O., Quirino dos Santos, C.V., Guimarães de Almeida, A.P.F., et al., 2021. Effect of Early Treatment With Hydroxychloroquine or Lopinavir and Ritonavir on Risk of Hospitalization Among Patients With COVID-19: The TOGETHER Randomized Clinical Trial. JAMA Netw Open 4, e216468. https://doi.org/10.1001/jamanetworkopen.2021.6468

Renzette, N., Bhattacharjee, B., Jensen, J.D., Gibson, L., Kowalik, T.F., 2011. Extensive Genome-Wide Variability of Human Cytomegalovirus in Congenitally Infected Infants. PLoS Pathog 7, e1001344. https://doi.org/10.1371/journal.ppat.1001344

Reperant, L.A., Osterhaus, A.D.M.E., 2017. AIDS, Avian flu, SARS, MERS, Ebola, Zika… what next? Vaccine 35, 4470–4474. https://doi.org/10.1016/j.vaccine.2017.04.082

Reynard, O., Nguyen, X.-N., Alazard-Dany, N., Barateau, V., Cimarelli, A., Volchkov, V., 2015. Identification of a New Ribonucleoside Inhibitor of Ebola Virus Replication. Viruses 7, 6233–6240. https://doi.org/10.3390/v7122934

Richards, F., Kodjamanova, P., Chen, X., Li, N., Atanasov, P., Bennetts, L., Patterson, B.J., Yektashenas, B., Mesa-Frias, M., Tronczynski, K., et al., 2022. Economic Burden of COVID-19: A Systematic Review. CEOR Volume 14, 293–307. https://doi.org/10.2147/CEOR.S338225

Richardson, B.A., John-Stewart, G., Atkinson, C., Nduati, R., Ásbjörnsdóttir, K., Boeckh, M., Overbaugh, J., Emery, V., Slyker, J.A., 2016. Vertical Cytomegalovirus Transmission From HIV-Infected Women Randomized to Formula-Feed or Breastfeed Their Infants. J Infect Dis. 213, 992–998. https://doi.org/10.1093/infdis/jiv515

Richardson, P., Griffin, I., Tucker, C., Smith, D., Oechsle, O., Phelan, A., Rawling, M., Savory, E., Stebbing, J., 2020. Baricitinib as potential treatment for 2019-

nCoV acute respiratory disease. The Lancet 395, e30–e31. https://doi.org/10.1016/S0140-6736(20)30304-4

Rijkers, G.T., Weterings, N., Obregon-Henao, A., Lepolder, M., Dutt, T.S., van Overveld, F.J., Henao-Tamayo, M., 2021. Antigen Presentation of mRNA-Based and Virus-Vectored SARS-CoV-2 Vaccines. Vaccines 9, 848. https://doi.org/10.3390/vaccines9080848

Robilotti, E., Deresinski, S., Pinsky, B.A., 2015. Norovirus. Clin Microbiol Rev 28, 134–164. https://doi.org/10.1128/CMR.00075-14

Rocha-Pereira, J., Jochmans, D., Dallmeier, K., Leyssen, P., Nascimento, M.S.J., Neyts, J., 2012. Favipiravir (T-705) inhibits in vitro norovirus replication. Biochemical and Biophysical Research Communications 424, 777–780. https://doi.org/10.1016/j.bbrc.2012.07.034

Rocha-Pereira, J., Van Dycke, J., Neyts, J., 2016. Treatment with a Nucleoside Polymerase Inhibitor Reduces Shedding of Murine Norovirus in Stool to Undetectable Levels without Emergence of Drug-Resistant Variants. Antimicrob Agents Chemother 60, 1907–1911. https://doi.org/10.1128/AAC.02198-15

Roddie, C., Paul, J.P.V., Benjamin, R., Gallimore, C.I., Xerry, J., Gray, J.J., Peggs, K.S., Morris, E.C., Thomson, K.J., Ward, K.N., 2009. Allogeneic Hematopoietic Stem Cell Transplantation and Norovirus Gastroenteritis: A Previously Unrecognized Cause of Morbidity. CLIN INFECT DIS 49, 1061–1068. https://doi.org/10.1086/605557

Rodríguez-Guillén, L., Vizzi, E., Alcalá, A.C., Pujol, F.H., Liprandi, F., Ludert, J.E., 2005. Calicivirus infection in human immunodeficiency virus seropositive children and adults. Journal of Clinical Virology 33, 104–109. https://doi.org/10.1016/j.jcv.2004.09.031

Rohatgi, Ankit, 2019. WebPlotDigitizer.

Rohayem, J., Robel, I., Jäger, K., Scheffler, U., Rudolph, W., 2006. Protein-Primed and De Novo Initiation of RNA Synthesis by Norovirus 3D ^{pol}. J Virol 80, 7060–7069. https://doi.org/10.1128/JVI.02195-05

Roos-Weil, D., Ambert-Balay, K., Lanternier, F., Mamzer-Bruneel, M.-F., Nochy, D., Pothier, P., Avettand-Fenoel, V., Anglicheau, D., Snanoudj, R., Bererhi, L., et al., 2011. Impact of Norovirus/Sapovirus-Related Diarrhea in Renal Transplant Recipients Hospitalized for Diarrhea. Transplantation 92, 61–69. https://doi.org/10.1097/TP.0b013e31821c9392

Rosenberg, R., 2015. Detecting the emergence of novel, zoonotic viruses pathogenic to humans. Cell. Mol. Life Sci. 72, 1115–1125. https://doi.org/10.1007/s00018-014-1785-y

Rosenke, K., Feldmann, H., Westover, J.B., Hanley, P.W., Martellaro, C., Feldmann, F., Saturday, G., Lovaglio, J., Scott, D.P., Furuta, Y., et al., 2018. Use of Favipiravir to Treat Lassa Virus Infection in Macaques. Emerg. Infect. Dis. 24, 1696–1699. https://doi.org/10.3201/eid2409.180233

Ross, S.A., Novak, Z., Pati, S., Patro, R.K., Blumenthal, J., Danthuluri, V.R., Ahmed, A., Michaels, M.G., Sánchez, P.J., Bernstein, D.I., et al., 2011. Mixed Infection and Strain Diversity in Congenital Cytomegalovirus Infection. The Journal of Infectious Diseases 204, 1003–1007. https://doi.org/10.1093/infdis/jir457

Rossignol, J.-F., 2014. Nitazoxanide: A first-in-class broad-spectrum antiviral agent. Antiviral Research 110, 94–103. https://doi.org/10.1016/j.antiviral.2014.07.014

Royall, E., Locker, N., 2016. Translational Control during Calicivirus Infection. Viruses 8, 104. https://doi.org/10.3390/v8040104

Ruis, C., Brown, L.-A.K., Roy, S., Atkinson, C., Williams, R., Burns, S.O., Yara-Romero, E., Jacobs, M., Goldstein, R., Breuer, J., et al., 2018a. Mutagenesis in Norovirus in Response to Favipiravir Treatment. New England Journal of Medicine 379, 2173–2176. https://doi.org/10.1056/NEJMc1806941

Ruis, C., Brown, L.-A.K., Roy, S., Atkinson, C., Williams, R., Burns, S.O., Yara-Romero, E., Jacobs, M., Goldstein, R., Breuer, J., et al., 2018b. Mutagenesis in Norovirus in Response to Favipiravir Treatment. New England Journal of Medicine 379, 2173–2176. https://doi.org/10.1056/NEJMc1806941

Rydyznski Moderbacher, C., Ramirez, S.I., Dan, J.M., Grifoni, A., Hastie, K.M., Weiskopf, D., Belanger, S., Abbott, R.K., Kim, Christina, Choi, J., et al., 2020. Antigen-Specific Adaptive Immunity to SARS-CoV-2 in Acute COVID-19 and Associations with Age and Disease Severity. Cell 183, 996-1012.e19. https://doi.org/10.1016/j.cell.2020.09.038

Sadler, H.A., Stenglein, M.D., Harris, R.S., Mansky, L.M., 2010. APOBEC3G Contributes to HIV-1 Variation through Sublethal Mutagenesis. J Virol 84, 7396–7404. https://doi.org/10.1128/JVI.00056-10

Sahin, A.R., 2020. 2019 Novel Coronavirus (COVID-19) Outbreak: A Review of the Current Literature. EJMO. https://doi.org/10.14744/ejmo.2020.12220

Sakurai, A., Sasaki, T., Kato, S., Hayashi, M., Tsuzuki, S., Ishihara, T., Iwata, M., Morise, Z., Doi, Y., 2020. Natural History of Asymptomatic SARS-CoV-2 Infection. N Engl J Med 383, 885–886. https://doi.org/10.1056/NEJMc2013020

Sanchez Clemente, N., Pang, J., Rodrigues, C., Aurora, P., Breuer, J., 2021. Case Report: severe paediatric COVID-19 pneumonitis treated with remdesivir and nitazoxanide. Wellcome Open Res 6, 329. https://doi.org/10.12688/wellcomeopenres.17377.1

Sangawa, H., Komeno, T., Nishikawa, H., Yoshida, A., Takahashi, K., Nomura, N., Furuta, Y., 2013. Mechanism of Action of T-705 Ribosyl Triphosphate against Influenza Virus RNA Polymerase. Antimicrob Agents Chemother 57, 5202–5208. https://doi.org/10.1128/AAC.00649-13

Sarvestani, S.T., Cotton, B., Fritzlar, S., O'Donnell, T.B., Mackenzie, J.M., 2016. Norovirus Infection: Replication, Manipulation of Host, and Interaction with the Host Immune Response. Journal of Interferon & Cytokine Research 36, 215–225. https://doi.org/10.1089/jir.2015.0124

Satarker, S., Ahuja, T., Banerjee, M., E, V.B., Dogra, S., Agarwal, T., Nampoothiri, M., 2020. Hydroxychloroquine in COVID-19: Potential Mechanism of Action Against SARS-CoV-2. Curr Pharmacol Rep 6, 203–211. https://doi.org/10.1007/s40495-020-00231-8

Satlin, M.J., Zucker, J., Baer, B.R., Rajan, M., Hupert, N., Schang, L.M., Pinheiro, L.C., Shen, Y., Sobieszczyk, M.E., Westblade, L.F., et al., 2021. Changes in SARS-CoV-2 viral load and mortality during the initial wave of the pandemic

in New York City. PLoS ONE 16, e0257979. https://doi.org/10.1371/journal.pone.0257979

Savela, E.S., Viloria Winnett, A., Romano, A.E., Porter, M.K., Shelby, N., Akana, R., Ji, J., Cooper, M.M., Schlenker, N.W., Reyes, J.A., et al., 2022. Quantitative SARS-CoV-2 Viral-Load Curves in Paired Saliva Samples and Nasal Swabs Inform Appropriate Respiratory Sampling Site and Analytical Test Sensitivity Required for Earliest Viral Detection. J Clin Microbiol 60, e01785-21. https://doi.org/10.1128/jcm.01785-21

Scallan, C.D., Tingley, D.W., Lindbloom, J.D., Toomey, J.S., Tucker, S.N., 2013. An Adenovirus-Based Vaccine with a Double-Stranded RNA Adjuvant Protects Mice and Ferrets against H5N1 Avian Influenza in Oral Delivery Models. Clin Vaccine Immunol 20, 85–94. https://doi.org/10.1128/CVI.00552-12

Scallan, E., Hoekstra, R.M., Angulo, F.J., Tauxe, R.V., Widdowson, M.-A., Roy, S.L., Jones, J.L., Griffin, P.M., 2011. Foodborne Illness Acquired in the United States—Major Pathogens. Emerg. Infect. Dis. 17, 7–15. https://doi.org/10.3201/eid1701.P11101

Schirmer, M., Sloan, W.T., Quince, C., 2014. Benchmarking of viral haplotype reconstruction programmes: an overview of the capacities and limitations of currently available programmes. Briefings in Bioinformatics 15, 431–442. https://doi.org/10.1093/bib/bbs081

Schoeman, D., Fielding, B.C., 2019. Coronavirus envelope protein: current knowledge. Virol J 16, 69. https://doi.org/10.1186/s12985-019-1182-0

Schorn, R., Höhne, M., Meerbach, A., Bossart, W., Wüthrich, R.P., Schreier, E., Müller, N.J., Fehr, T., 2010. Chronic Norovirus Infection after Kidney Transplantation: Molecular Evidence for Immune-Driven Viral Evolution. CLIN INFECT DIS 51, 307–314. https://doi.org/10.1086/653939

Schwartz, S., Vergoulidou, M., Schreier, E., Loddenkemper, C., Reinwald, M., Schmidt-Hieber, M., Flegel, W.A., Thiel, E., Schneider, T., 2011. Norovirus gastroenteritis causes severe and lethal complications after chemotherapy and hematopoietic stem cell transplantation. Blood 117, 5850–5856. https://doi.org/10.1182/blood-2010-12-325886

Seah, I.Y.J., Anderson, D.E., Kang, A.E.Z., Wang, L., Rao, P., Young, B.E., Lye, D.C., Agrawal, R., 2020. Assessing Viral Shedding and Infectivity of Tears in Coronavirus Disease 2019 (COVID-19) Patients. Ophthalmology. https://doi.org/10.1016/j.ophtha.2020.03.026

Segerman, B., 2020. The Most Frequently Used Sequencing Technologies and Assembly Methods in Different Time Segments of the Bacterial Surveillance and RefSeq Genome Databases. Front. Cell. Infect. Microbiol. 10, 527102. https://doi.org/10.3389/fcimb.2020.527102

Sekine, T., Perez-Potti, A., Rivera-Ballesteros, O., Strålin, K., Gorin, J.-B., Olsson, A., Llewellyn-Lacey, S., Kamal, H., Bogdanovic, G., Muschiol, S., et al., 2020. Robust T Cell Immunity in Convalescent Individuals with Asymptomatic or Mild COVID-19. Cell 183, 158-168.e14. https://doi.org/10.1016/j.cell.2020.08.017

Sender, R., Bar-On, Y.M., Gleizer, S., Bernshtein, B., Flamholz, A., Phillips, R., Milo, R., 2021. The total number and mass of SARS-CoV-2 virions. Proc. Natl. Acad. Sci. U.S.A. 118, e2024815118. https://doi.org/10.1073/pnas.2024815118

Sette, A., Crotty, S., 2021. Adaptive immunity to SARS-CoV-2 and COVID-19. Cell 184, 861–880. https://doi.org/10.1016/j.cell.2021.01.007

Severson, W.E., Schmaljohn, C.S., Javadian, A., Jonsson, C.B., 2003. Ribavirin Causes Error Catastrophe during Hantaan Virus Replication. J Virol 77, 481–488. https://doi.org/10.1128/JVI.77.1.481-488.2003

Shah, V.K., Firmal, P., Alam, A., Ganguly, D., Chattopadhyay, S., 2020. Overview of Immune Response During SARS-CoV-2 Infection: Lessons From the Past. Front. Immunol. 11, 1949. https://doi.org/10.3389/fimmu.2020.01949

Shannon, A., Selisko, B., Le, N.-T.-T., Huchting, J., Touret, F., Piorkowski, G., Fattorini, V., Ferron, F., Decroly, E., Meier, C., et al., 2020. Rapid incorporation of Favipiravir by the fast and permissive viral RNA polymerase complex results in SARS-CoV-2 lethal mutagenesis. Nat Commun 11, 4682. https://doi.org/10.1038/s41467-020-18463-z

Sharp, T.M., Crawford, S.E., Ajami, N.J., Neill, F.H., Atmar, R.L., Katayama, K., Utama, B., Estes, M.K., 2012. Secretory pathway antagonism by calicivirus homologues of Norwalk virus nonstructural protein p22 is restricted to noroviruses. Virol J 9, 181. https://doi.org/10.1186/1743-422X-9-181

Sharp, T.M., Guix, S., Katayama, K., Crawford, S.E., Estes, M.K., 2010. Inhibition of Cellular Protein Secretion by Norwalk Virus Nonstructural Protein p22 Requires a Mimic of an Endoplasmic Reticulum Export Signal. PLoS ONE 5, e13130. https://doi.org/10.1371/journal.pone.0013130

Sheahan, T.P., Sims, A.C., Zhou, S., Graham, R.L., Pruijssers, A.J., Agostini, M.L., Leist, S.R., Schäfer, A., Dinnon, K.H., Stevens, L.J., et al., 2020. An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. Sci. Transl. Med. 12, eabb5883. https://doi.org/10.1126/scitranslmed.abb5883

Shen, C., Wang, Z., Zhao, F., Yang, Yang, Li, J., Yuan, J., Wang, F., Li, D., Yang, M., Xing, L., et al., 2020. Treatment of 5 Critically Ill Patients With COVID-19 With Convalescent Plasma. JAMA. https://doi.org/10.1001/jama.2020.4783

Shen, Q., Zhang, W., Yang, S., Cui, L., Hua, X., 2012. Complete Genome Sequence of a New-Genotype Porcine Norovirus Isolated from Piglets with Diarrhea. J Virol 86, 7015–7016. https://doi.org/10.1128/JVI.00757-12

Shen, W., Le, S., Li, Y., Hu, F., 2016. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. PLOS ONE 11, e0163962. https://doi.org/10.1371/journal.pone.0163962

Shenoy, S., 2021. SARS-CoV-2 (COVID-19), viral load and clinical outcomes; lessons learned one year into the pandemic: A systematic review. WJCCM 10, 132–150. https://doi.org/10.5492/wjccm.v10.i4.132

Sherwood, J., Mendelman, P.M., Lloyd, E., Liu, M., Boslego, J., Borkowski, A., Jackson, A., Faix, D., 2020. Efficacy of an intramuscular bivalent norovirus GI.1/GII.4 virus-like particle vaccine candidate in healthy US adults. Vaccine 38, 6442–6449. https://doi.org/10.1016/j.vaccine.2020.07.069

Shunmugam, L., Soliman, M.E.S., 2018. Targeting HCV polymerase: a structural and dynamic perspective into the mechanism of selective covalent inhibition. RSC Adv. 8, 42210–42222. https://doi.org/10.1039/C8RA07346E

Siebenga, J.J., Beersma, M.F.C., Vennema, H., van Biezen, P., Hartwig, N.J., Koopmans, M., 2008. High Prevalence of Prolonged Norovirus Shedding and Illness among Hospitalized Patients: A Model for In Vivo Molecular Evolution. J INFECT DIS 198, 994–1001. https://doi.org/10.1086/591627

Siebenga, J.J., Vennema, H., Zheng, D., Vinjé, J., Lee, B.E., Pang, X., Ho, E.C.M., Lim, W., Choudekar, A., Broor, S., et al., 2009. Norovirus Illness Is a Global Problem: Emergence and Spread of Norovirus GII.4 Variants, 2001–2007. J INFECT DIS 200, 802–812. https://doi.org/10.1086/605127

Siegel, D., Hui, H.C., Doerffler, E., Clarke, M.O., Chun, K., Zhang, L., Neville, S., Carra, E., Lew, W., Ross, B., et al., 2017. Discovery and Synthesis of a Phosphoramidate Prodrug of a Pyrrolo[2,1- *f*][triazin-4-amino] Adenine *C* -Nucleoside (GS-5734) for the Treatment of Ebola and Emerging Viruses. J. Med. Chem. 60, 1648–1661. https://doi.org/10.1021/acs.jmedchem.6b01594

Sierra, S., Dávila, M., Lowenstein, P.R., Domingo, E., 2000. Response of Foot-and-Mouth Disease Virus to Increased Mutagenesis: Influence of Viral Load and Fitness in Loss of Infectivity. J Virol 74, 8316–8323. https://doi.org/10.1128/JVI.74.18.8316-8323.2000

Simmons, K., Gambhir, M., Leon, J., Lopman, B., 2013. Duration of Immunity to Norovirus Gastroenteritis. Emerg. Infect. Dis. 19, 1260–1267. https://doi.org/10.3201/eid1908.130472

Simonis, A., Theobald, S.J., Fätkenheuer, G., Rybniker, J., Malin, J.J., 2021. A comparative analysis of remdesivir and other repurposed antivirals against SARS-CoV-2. EMBO Mol Med 13. https://doi.org/10.15252/emmm.202013105

Singh, A.K., Singh, A., Singh, R., Misra, A., 2021. Molnupiravir in COVID-19: A systematic review of literature. Diabetes & Metabolic Syndrome: Clinical Research & Reviews 15, 102329. https://doi.org/10.1016/j.dsx.2021.102329

Slatko, B.E., Gardner, A.F., Ausubel, F.M., 2018. Overview of Next-Generation Sequencing Technologies. Current Protocols in Molecular Biology 122. https://doi.org/10.1002/cpmb.59

Smertina, E., Urakova, N., Strive, T., Frese, M., 2019. Calicivirus RNA-Dependent RNA Polymerases: Evolution, Structure, Protein Dynamics, and Function. Front. Microbiol. 10, 1280. https://doi.org/10.3389/fmicb.2019.01280

Smith, C.L., Stein, G.E., 2002. Viral Load as a Surrogate End Point in HIV Disease. Ann Pharmacother 36, 280–287. https://doi.org/10.1345/aph.1A118

Smith, D.J., Lapedes, A.S., de Jong, J.C., Bestebroer, T.M., Rimmelzwaan, G.F., Osterhaus, A.D.M.E., Fouchier, R.A.M., 2004. Mapping the Antigenic and Genetic Evolution of Influenza Virus. Science 305, 371–376. https://doi.org/10.1126/science.1097211

Snijder, E.J., Decroly, E., Ziebuhr, J., 2016. The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing, in: Advances in Virus Research. Elsevier, pp. 59–126. https://doi.org/10.1016/bs.aivir.2016.08.008

Someya, Y., Takeda, N., Miyamura, T., 2000. Complete Nucleotide Sequence of the Chiba Virus Genome and Functional Expression of the 3C-Like Protease in Escherichia coli. Virology 278, 490–500. https://doi.org/10.1006/viro.2000.0672

Song, P., Li, W., Xie, J., Hou, Y., You, C., 2020. Cytokine storm induced by SARS-CoV-2. Clinica Chimica Acta 509, 280–287. https://doi.org/10.1016/j.cca.2020.06.017

Soria, M.E., Cortón, M., Martínez-González, B., Lobo-Vega, R., Vázquez-Sirvent, L., López-Rodríguez, R., Almoguera, B., Mahillo, I., Mínguez, P., Herrero, A., et al., 2021. High SARS-CoV-2 viral load is associated with a worse clinical outcome of COVID-19 disease. Access Microbiology 3. https://doi.org/10.1099/acmi.0.000259

Sourimant, J., Lieber, C.M., Aggarwal, M., Cox, R.M., Wolf, J.D., Yoon, J.-J., Toots, M., Ye, C., Sticher, Z., Kolykhalov, A.A., et al., 2022. 4'-Fluorouridine is an oral antiviral that blocks respiratory syncytial virus and SARS-CoV-2 replication. Science 375, 161–167. https://doi.org/10.1126/science.abj5508

Souza, M., Azevedo, M.S.P., Jung, K., Cheetham, S., Saif, L.J., 2008. Pathogenesis and Immune Responses in Gnotobiotic Calves after Infection with the Genogroup II.4-HS66 Strain of Human Norovirus. J Virol 82, 1777–1786. https://doi.org/10.1128/JVI.01347-07

Stamatakis, A., 2014a. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stamatakis, A., 2014b. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stanton, R.J., Baluchova, K., Dargan, D.J., Cunningham, C., Sheehy, O., Seirafian, S., McSharry, B.P., Neale, M.L., Davies, J.A., Tomasec, P., et al., 2010. Reconstruction of the complete human cytomegalovirus genome in a BAC reveals RL13 to be a potent inhibitor of replication. J. Clin. Invest. 120, 3191–3208. https://doi.org/10.1172/JCI42955

Stockis, A., Bruyn, S.D., Gengler, C., Rosillon, D., 2002. Nitazoxanide pharmacokinetics and tolerability in man during 7 days dosing with 0.5 g and 1 g b.i.d. CP 40, 221–227. https://doi.org/10.5414/CPP40221

Stockman, L.J., Bellamy, R., Garner, P., 2006. SARS: Systematic Review of Treatment Effects. PLoS Med 3, e343. https://doi.org/10.1371/journal.pmed.0030343

Streeter, D.G., Witkowski, J.T., Khare, G.P., Sidwell, R.W., Bauer, R.J., Robins, R.K., Simon, L.N., 1973. Mechanism of Action of 1-β-D-Ribofuranosyl-1,2,4-Triazole-3-Carboxamide (Virazole), A New Broad-Spectrum Antiviral Agent. Proc. Natl. Acad. Sci. U.S.A. 70, 1174–1178. https://doi.org/10.1073/pnas.70.4.1174

Strydom, N., Gupta, S.V., Fox, W.S., Via, L.E., Bang, H., Lee, M., Eum, S., Shim, T., Barry, C.E., Zimmerman, M., et al., 2019. Tuberculosis drugs' distribution and emergence of resistance in patient's lung lesions: A mechanistic model and

tool for regimen and dose optimization. PLoS Med 16, e1002773. https://doi.org/10.1371/journal.pmed.1002773

Stuyver, L.J., Whitaker, T., McBrayer, T.R., Hernandez-Santiago, B.I., Lostia, S., Tharnish, P.M., Ramesh, M., Chu, C.K., Jordan, R., Shi, J., et al., 2003. Ribonucleoside Analogue That Blocks Replication of Bovine Viral Diarrhea and Hepatitis C Viruses in Culture. Antimicrob Agents Chemother 47, 244–254. https://doi.org/10.1128/AAC.47.1.244-254.2003

Sun, J., Yogarajah, T., Lee, R.C.H., Kaur, P., Inoue, M., Tan, Y.W., Chu, J.J.H., 2020. Drug repurposing of pyrimidine analogs as potent antiviral compounds against human enterovirus A71 infection with potential clinical applications. Sci Rep 10, 8159. https://doi.org/10.1038/s41598-020-65152-4

Takanashi, S., Wang, Q., Chen, N., Shen, Q., Jung, K., Zhang, Z., Yokoyama, M., Lindesmith, L.C., Baric, R.S., Saif, L.J., 2011. Characterization of Emerging GII.g/GII.12 Noroviruses from a Gastroenteritis Outbreak in the United States in 2010. J Clin Microbiol 49, 3234–3244. https://doi.org/10.1128/JCM.00305-11

Takaori-Kondo, A., 2006. APOBEC Family Proteins: Novel Antiviral Innate Immunity. International Journal of Hematology 83, 213–216. https://doi.org/10.1532/IJH97.05187

Takashita, E., Kinoshita, N., Yamayoshi, S., Sakai-Tagawa, Y., Fujisaki, S., Ito, M., Iwatsuki-Horimoto, K., Chiba, S., Halfmann, P., Nagai, H., et al., 2022. Efficacy of Antibodies and Antiviral Drugs against Covid-19 Omicron Variant. N Engl J Med 386, 995–998. https://doi.org/10.1056/NEJMc2119407

Talmy, T., Tsur, A., Shabtay, O., 2021. Duration of SARS-CoV-2 detection in Israel Defense Forces soldiers with mild COVID-19. J Med Virol 93, 608–610. https://doi.org/10.1002/jmv.26374

Tan, M., 2021. Norovirus Vaccines: Current Clinical Development and Challenges. Pathogens 10, 1641. https://doi.org/10.3390/pathogens10121641

Tchesnokov, E., Feng, J., Porter, D., Götte, M., 2019. Mechanism of Inhibition of Ebola Virus RNA-Dependent RNA Polymerase by Remdesivir. Viruses 11, 326. https://doi.org/10.3390/v11040326

Tchesnokov, E.P., Feng, J.Y., Porter, D.P., Götte, M., 2019. Mechanism of Inhibition of Ebola Virus RNA-Dependent RNA Polymerase by Remdesivir. Viruses 11. https://doi.org/10.3390/v11040326

Te, H.S., Randall, G., Jensen, D.M., 2007. Mechanism of action of ribavirin in the treatment of chronic hepatitis C. Gastroenterol Hepatol (N Y) 3, 218–225.

te Velthuis, A.J.W., 2014. Common and unique features of viral RNA-dependent polymerases. Cell. Mol. Life Sci. 71, 4403–4420. https://doi.org/10.1007/s00018-014-1695-z

Teng, J.L.L., Martelli, P., Chan, W.-M., Lee, H.H., Hui, S.-W., Lau, C.C.Y., Tse, H., Yuen, K.-Y., Lau, S.K.P., Woo, P.C.Y., 2018. Two novel noroviruses and a novel norovirus genogroup in California sea lions. Journal of General Virology 99, 777–782. https://doi.org/10.1099/jgv.0.001071

Teunis, P.F.M., Moe, C.L., Liu, P., E. Miller, S., Lindesmith, L., Baric, R.S., Le Pendu, J., Calderon, R.L., 2008. Norwalk virus: How infectious is it? J. Med. Virol. 80, 1468–1476. https://doi.org/10.1002/jmv.21237

Thackray, L.B., Duan, E., Lazear, H.M., Kambal, A., Schreiber, R.D., Diamond, M.S., Virgin, H.W., 2012. Critical Role for Interferon Regulatory Factor 3 (IRF-3) and IRF-7 in Type I Interferon-Mediated Control of Murine Norovirus Replication. J Virol 86, 13515–13523. https://doi.org/10.1128/JVI.01824-12

Thakur, V., Ratho, R.K., 2022. OMICRON (B.1.1.529): A new SARS-CoV-2 variant of concern mounting worldwide fear. Journal of Medical Virology 94, 1821–1824. https://doi.org/10.1002/jmv.27541

The Babraham Institute, 2019. TrimGalore [WWW Document]. URL https://github.com/FelixKrueger/TrimGalore

The Massachusetts Consortium for Pathogen Readiness, Fajnzylber, J., Regan, J., Coxen, K., Corry, H., Wong, C., Rosenthal, A., Worrall, D., Giguel, F., Piechocka-Trocha, A., et al., 2020. SARS-CoV-2 viral load is associated with increased disease severity and mortality. Nat Commun 11, 5493. https://doi.org/10.1038/s41467-020-19057-5

Thimme, R., Oldach, D., Chang, K.-M., Steiger, C., Ray, S.C., Chisari, F.V., 2001. Determinants of Viral Clearance and Persistence during Acute Hepatitis C Virus Infection. Journal of Experimental Medicine 194, 1395–1406. https://doi.org/10.1084/jem.194.10.1395

Thomas, E., Ghany, M.G., Liang, T.J., 2012. The Application and Mechanism of Action of Ribavirin in Therapy of Hepatitis C. Antivir Chem Chemother 23, 1–12. https://doi.org/10.3851/IMP2125

Thorne, L.G., Goodfellow, I.G., 2014. Norovirus gene expression and replication. Journal of General Virology 95, 278–291. https://doi.org/10.1099/vir.0.059634-0

Thornley, C.N., Emslie, N.A., Sprott, T.W., Greening, G.E., Rapana, J.P., 2011. Recurring Norovirus Transmission on an Airplane. Clinical Infectious Diseases 53, 515–520. https://doi.org/10.1093/cid/cir465

Tian, L., Qiang, T., Liang, C., Ren, X., Jia, M., Zhang, J., Li, J., Wan, M., YuWen, X., Li, H., et al., 2021. RNA-dependent RNA polymerase (RdRp) inhibitors: The current landscape and repurposing for the COVID-19 pandemic. European Journal of Medicinal Chemistry 213, 113201. https://doi.org/10.1016/j.ejmech.2021.113201

Tisthammer, K.H., Solis, C., Oracles, F., Nzerem, M., Winstead, R., Dong, W., Joy, J.B., Pennings, P.S., 2020. Assessing *in vivo* mutation frequencies and creating a high-resolution genome-wide map of fitness costs of Hepatitis C virus (preprint). Evolutionary Biology. https://doi.org/10.1101/2020.10.01.323253

To, K.K.-W., Tsang, O.T.-Y., Leung, W.-S., Tam, A.R., Wu, T.-C., Lung, D.C., Yip, C.C.-Y., Cai, J.-P., Chan, J.M.-C., Chik, T.S.-H., et al., 2020. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. The Lancet Infectious Diseases S1473309920301961. https://doi.org/10.1016/S1473-3099(20)30196-1

Todd, K., Tripp, R., 2019. Human Norovirus: Experimental Models of Infection. Viruses 11, 151. https://doi.org/10.3390/v11020151

Tomley, F.M., Shirley, M.W., 2009. Livestock infectious diseases and zoonoses. Phil. Trans. R. Soc. B 364, 2637–2642. https://doi.org/10.1098/rstb.2009.0133

Toots, M., Yoon, J.-J., Cox, R.M., Hart, M., Sticher, Z.M., Makhsous, N., Plesker, R., Barrena, A.H., Reddy, P.G., Mitchell, D.G., et al., 2019. Characterization of orally efficacious influenza drug with high resistance barrier in ferrets and human airway epithelia. Sci. Transl. Med. 11, eaax5866. https://doi.org/10.1126/scitranslmed.aax5866

Töpfer, A., Marschall, T., Bull, R.A., Luciani, F., Schönhuth, A., Beerenwinkel, N., 2014. Viral Quasispecies Assembly via Maximal Clique Enumeration. PLOS Computational Biology 10, e1003515. https://doi.org/10.1371/journal.pcbi.1003515

Turner, D.L., Bickham, K.L., Thome, J.J., Kim, C.Y., D'Ovidio, F., Wherry, E.J., Farber, D.L., 2014. Lung niches for the generation and maintenance of tissue-resident memory T cells. Mucosal Immunol 7, 501–510. https://doi.org/10.1038/mi.2013.67

Tvarogová, J., Madhugiri, R., Bylapudi, G., Ferguson, L.J., Karl, N., Ziebuhr, J., 2019. Identification and Characterization of a Human Coronavirus 229E Nonstructural Protein 8-Associated RNA 3′-Terminal Adenylyltransferase Activity. J Virol 93, e00291-19. https://doi.org/10.1128/JVI.00291-19

UK Health and Security Agency, 2022. National norovirus and rotavirus bulletin - Routine norovirus and rotavirus surveillance in England, 2021 to 2022 season.

UK Parliament, 2022. Drug Therapies for COVID-19.

Urakova, N., Kuznetsova, V., Crossman, D.K., Sokratian, A., Guthrie, D.B., Kolykhalov, A.A., Lockwood, M.A., Natchus, M.G., Crowley, M.R., Painter, G.R., et al., 2018. β- D - N^4 -Hydroxycytidine Is a Potent Anti-alphavirus Compound That Induces a High Level of Mutations in the Viral Genome. J Virol 92, e01965-17. https://doi.org/10.1128/JVI.01965-17

Van Damme, E., Van Loock, M., 2014. Functional annotation of human cytomegalovirus gene products: an update. Front. Microbiol. 5. https://doi.org/10.3389/fmicb.2014.00218

Van Dycke, J., Ny, A., Conceição-Neto, N., Maes, J., Hosmillo, M., Cuvry, A., Goodfellow, I., Nogueira, T.C., Verbeken, E., Matthijnssens, J., et al., 2019. A robust human norovirus replication model in zebrafish larvae. PLoS Pathog 15, e1008009. https://doi.org/10.1371/journal.ppat.1008009

van Kampen, J.J.A., Dalm, V.A.S.H., Fraaij, P.L.A., Oude Munnink, B.B., Schapendonk, C.M.E., Izquierdo-Lara, R.W., Villabruna, N., Ettayebi, K., Estes, M.K., Koopmans, M.P.G., et al., 2022. Clinical and In Vitro Evidence Favoring Immunoglobulin Treatment of a Chronic Norovirus Infection in a Patient With Common Variable Immunodeficiency. The Journal of Infectious Diseases jiac085. https://doi.org/10.1093/infdis/jiac085

Vanderlinden, E., Vrancken, B., Van Houdt, J., Rajwanshi, V.K., Gillemot, S., Andrei, G., Lemey, P., Naesens, L., 2016. Distinct Effects of T-705 (Favipiravir) and Ribavirin on Influenza Virus Replication and Viral RNA Synthesis. Antimicrob Agents Chemother 60, 6679–6691. https://doi.org/10.1128/AAC.01156-16

Vaxart, 2022. A Phase 1b, Multicenter, Randomized, Double-blind, Placebo-controlled Study to Determine the Safety and Immunogenicity of an

Adenoviral-vector Based Oral Norovirus Vaccine Expressing GI.1 VP1 Administered Orally to Health Stable Older Adult Volunteers 55-80 Years of Age (Clinical trial registration No. NCT04854746). clinicaltrials.gov.

Venkataraman, S., Prasad, B., Selvarajan, R., 2018. RNA Dependent RNA Polymerases: Insights from Structure, Function and Evolution. Viruses 10, 76. https://doi.org/10.3390/v10020076

Verma, R., Kim, E., Martínez-Colón, G.J., Jagannathan, P., Rustagi, A., Parsonnet, J., Bonilla, H., Khosla, C., Holubar, M., Subramanian, A., et al., 2021. SARS-CoV-2 Subgenomic RNA Kinetics in Longitudinal Clinical Samples. Open Forum Infectious Diseases 8, ofab310. https://doi.org/10.1093/ofid/ofab310

Vignuzzi, M., Stone, J.K., Andino, R., 2005. Ribavirin and lethal mutagenesis of poliovirus: molecular mechanisms, resistance and biological implications. Virus Research 107, 173–181. https://doi.org/10.1016/j.virusres.2004.11.007

Vinjé, J., 2015. Advances in Laboratory Methods for Detection and Typing of Norovirus. J Clin Microbiol 53, 373–381. https://doi.org/10.1128/JCM.01535-14

Vinjé, J., Green, J., Lewis, D.C., Gallimore, C.I., Brown, D.W.G., Koopmans, M.P.G., 2000. Genetic polymorphism across regions of the three open reading frames of “Norwalk-like viruses.” Archives of Virology 145, 223–241. https://doi.org/10.1007/s007050050020

Vinje, J., Koopmans, M.P.G., 1996. Molecular Detection and Epidemiology of Small Round-Structured Viruses in Outbreaks of Gastroenteritis in the Netherlands. Journal of Infectious Diseases 174, 610–615. https://doi.org/10.1093/infdis/174.3.610

Virology: Coronaviruses, 1968. . Nature 220, 650–650. https://doi.org/10.1038/220650b0

V’kovski, P., Kratzel, A., Steiner, S., Stalder, H., Thiel, V., 2021. Coronavirus biology and replication: implications for SARS-CoV-2. Nat Rev Microbiol 19, 155–170. https://doi.org/10.1038/s41579-020-00468-6

Vo, N.V., Young, K.-C., Lai, M.M.C., 2003. Mutagenic and Inhibitory Effects of Ribavirin on Hepatitis C Virus RNA Polymerase. Biochemistry 42, 10462–10471. https://doi.org/10.1021/bi0344681

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., Muenker, M.C., Moore, A.J., et al., 2020. Analytical sensitivity and efficiency comparisons of SARS-COV-2 qRT-PCR primer-probe sets (preprint). Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2020.03.30.20048108

Voinsky, I., Baristaite, G., Gurwitz, D., 2020. Effects of age and sex on recovery from COVID-19: Analysis of 5769 Israeli patients. Journal of Infection 81, e102–e103. https://doi.org/10.1016/j.jinf.2020.05.026

Vongpunsawad, S., Venkataram Prasad, B.V., Estes, M.K., 2013. Norwalk Virus Minor Capsid Protein VP2 Associates within the VP1 Shell Domain. J Virol 87, 4818–4825. https://doi.org/10.1128/JVI.03508-12

Wacharapluesadee, S., Tan, C.W., Maneeorn, P., Duengkae, P., Zhu, F., Joyjinda, Y., Kaewpom, T., Chia, W.N., Ampoot, W., Lim, B.L., et al., 2021. Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in

Southeast Asia. Nat Commun 12, 972. https://doi.org/10.1038/s41467-021-21240-1

Wadhwa, A., Aljabbari, A., Lokras, A., Foged, C., Thakur, A., 2020. Opportunities and Challenges in the Delivery of mRNA-Based Vaccines. Pharmaceutics 12, 102. https://doi.org/10.3390/pharmaceutics12020102

Walsh, K.A., Jordan, K., Clyne, B., Rohde, D., Drummond, L., Byrne, P., Ahern, S., Carty, P.G., O'Brien, K.K., O'Murchu, E., et al., 2020. SARS-CoV-2 detection, viral load and infectivity over the course of an infection. Journal of Infection 81, 357–371. https://doi.org/10.1016/j.jinf.2020.06.067

Wan, R., Mao, Z.-Q., He, L.-Y., Hu, Y.-C., Wei-Chen, null, 2020. Evidence from two cases of asymptomatic infection with SARS-CoV-2: Are 14 days of isolation sufficient? Int. J. Infect. Dis. 95, 174–175. https://doi.org/10.1016/j.ijid.2020.03.041

Wang, Q., Wu, J., Wang, H., Gao, Y., Liu, Q., Mu, A., Ji, W., Yan, L., Zhu, Y., Zhu, C., et al., 2020. Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. Cell 182, 417-428.e13. https://doi.org/10.1016/j.cell.2020.05.034

Wang, Q.-H., Han, M.G., Cheetham, S., Souza, M., Funk, J.A., Saif, L.J., 2005. Porcine Noroviruses Related to Human Noroviruses. Emerg. Infect. Dis. 11, 1874–1881. https://doi.org/10.3201/eid1112.050485

Wang, Y., Chen, L., 2020. Tissue distributions of antiviral drugs affect their capabilities of reducing viral loads in COVID-19 treatment. European Journal of Pharmacology 889, 173634. https://doi.org/10.1016/j.ejphar.2020.173634

Wang, Y., Li, G., Yuan, S., Gao, Q., Lan, K., Altmeyer, R., Zou, G., 2016. *In Vitro* Assessment of Combinations of Enterovirus Inhibitors against Enterovirus 71. Antimicrob Agents Chemother 60, 5357–5367. https://doi.org/10.1128/AAC.01073-16

Wang, Y., Zhang, D., Du, G., Du, R., Zhao, J., Jin, Y., Fu, S., Gao, L., Cheng, Z., Lu, Q., et al., 2020. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. The Lancet 395, 1569–1578. https://doi.org/10.1016/S0140-6736(20)31022-9

Watson, S.J., Welkers, M.R.A., Depledge, D.P., Coulter, E., Breuer, J.M., de Jong, M.D., Kellam, P., 2013. Viral population analysis and minority-variant detection using short read next-generation sequencing. Phil. Trans. R. Soc. B 368, 20120205. https://doi.org/10.1098/rstb.2012.0205

Watterson, G.A., 1975. On the number of segregating sites in genetical models without recombination. Theoretical Population Biology 7, 256–276. https://doi.org/10.1016/0040-5809(75)90020-9

Weerasekara, S., Prior, A.M., Hua, D.H., 2016. Current tools for norovirus drug discovery. Expert Opinion on Drug Discovery 11, 529–541. https://doi.org/10.1080/17460441.2016.1178231

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al., 2019. Welcome to the Tidyverse. JOSS 4, 1686. https://doi.org/10.21105/joss.01686

Williamson, B.N., Feldmann, F., Schwarz, B., Meade-White, K., Porter, D.P., Schulz, J., van Doremalen, N., Leighton, I., Yinda, C.K., Pérez-Pérez, L., et al., 2020. Clinical benefit of remdesivir in rhesus macaques infected with SARS-CoV-2. Nature 585, 273–276. https://doi.org/10.1038/s41586-020-2423-5

Williamson, E.J., Walker, A.J., Bhaskaran, K., Bacon, S., Bates, C., Morton, C.E., Curtis, H.J., Mehrkar, A., Evans, D., Inglesby, P., et al., 2020. Factors associated with COVID-19-related death using OpenSAFELY. Nature 584, 430–436. https://doi.org/10.1038/s41586-020-2521-4

Wobus, C.E., Karst, S.M., Thackray, L.B., Chang, K.-O., Sosnovtsev, S.V., Belliot, G., Krug, A., Mackenzie, J.M., Green, K.Y., Virgin, H.W., 2004. Replication of Norovirus in Cell Culture Reveals a Tropism for Dendritic Cells and Macrophages. PLoS Biol 2, e432. https://doi.org/10.1371/journal.pbio.0020432

Wolf, S., Williamson, W., Hewitt, J., Lin, S., Rivera-Aban, M., Ball, A., Scholes, P., Savill, M., Greening, G.E., 2009. Molecular detection of norovirus in sheep and pigs in New Zealand farms. Veterinary Microbiology 133, 184–189. https://doi.org/10.1016/j.vetmic.2008.06.019

Wölfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Müller, M.A., Niemeyer, D., Jones, T.C., Vollmar, P., Rothe, C., et al., 2020. Virological assessment of hospitalized patients with COVID-2019. Nature. https://doi.org/10.1038/s41586-020-2196-x

Woo, P.C.Y., Lau, S.K.P., Lam, C.S.F., Lau, C.C.Y., Tsang, A.K.L., Lau, J.H.N., Bai, R., Teng, J.L.L., Tsang, C.C.C., Wang, M., et al., 2012. Discovery of Seven Novel Mammalian and Avian Coronaviruses in the Genus Deltacoronavirus Supports Bat Coronaviruses as the Gene Source of Alphacoronavirus and Betacoronavirus and Avian Coronaviruses as the Gene Source of Gammacoronavirus and Deltacoronavirus. J Virol 86, 3995–4008. https://doi.org/10.1128/JVI.06540-11

Woodward, J., Ng, A., Aravinthan, A., Bandoh, B., Liu, H., Davies, S., Stevenson, P., Curran, M., Kumararatne, D., 2012. PTU-160 Successful clearance of chronic noroviral infection by ribavirin in a patient with common variable immunodeficiency-associated enteropathy results in complete symptomatic and histopathological resolution. Gut 61, A251. https://doi.org/10.1136/gutjnl-2012-302514c.160

Woolhouse, M.E.J., Brierley, L., McCaffery, C., Lycett, S., 2016. Assessing the Epidemic Potential of RNA and DNA Viruses. Emerg. Infect. Dis. 22, 2037–2044. https://doi.org/10.3201/eid2212.160123

World Health Organization, 2022a. COVID-19 vaccine tracker and landscape.

World Health Organization, 2022b. WHO recommends highly successful COVID-19 therapy and calls for wide geographical distribution and transparency from originator [WWW Document]. URL https://www.who.int/news/item/22-04-2022-who-recommends-highly-successful-covid-19-therapy-and-calls-for-wide-geographical-distribution-and-transparency-from-originator (accessed 5.22.22).

World Health Organization, 2022c. WHO recommends two new drugs to treat COVID-19 [WWW Document]. URL https://www.who.int/news/item/14-01-2022-who-recommends-two-new-drugs-to-treat-covid-19 (accessed 5.22.22).

World Health Organization, 2020. Clinical management of severe acute respiratory infection (SARI) when COVID-19 disease is suspected.

World Health Organization, 2018. Guidelines for the Care and Treatment of Persons Diagnosed with Chronic Hepatitis C Virus Infection.

Worldometers.info, 2022. COVID-19 Coronavirus Pandemic Counter.

Wray, S.K., Gilbert, B.E., Noall, M.W., Knight, V., 1985. Mode of action of ribavirin: Effect of nucleotide pool alterations on influenza virus ribonucleoprotein synthesis. Antiviral Research 5, 29–37. https://doi.org/10.1016/0166-3542(85)90012-9

Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., Meng, J., Zhu, Z., Zhang, Z., Wang, J., et al., 2020. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. Cell Host & Microbe 27, 325–328. https://doi.org/10.1016/j.chom.2020.02.001

Wu, J.Z., Larson, G., Walker, H., Shim, J.H., Hong, Z., 2005. Phosphorylation of Ribavirin and Viramidine by Adenosine Kinase and Cytosolic 5′-Nucleotidase II: Implications for Ribavirin Metabolism in Erythrocytes. Antimicrob Agents Chemother 49, 2164–2171. https://doi.org/10.1128/AAC.49.6.2164-2171.2005

Wu, Z., Yang, L., Ren, X., He, G., Zhang, J., Yang, J., Qian, Z., Dong, J., Sun, L., Zhu, Y., et al., 2016. Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. ISME J 10, 609–620. https://doi.org/10.1038/ismej.2015.138

Wyatt, R.G., Greenberg, H.B., Dalgard, D.W., Allen, W.P., Sly, D.L., Thornhill, T.S., Chanock, R.M., Kapikian, A.Z., 1978. Experimental infection of chimpanzees with the Norwalk agent of epidemic viral gastroenteritis. J. Med. Virol. 2, 89–96. https://doi.org/10.1002/jmv.1890020203

Wyllie, A.L., Fournier, J., Casanovas-Massana, A., Campbell, M., Tokuyama, M., Vijayakumar, P., Geng, B., Muenker, M.C., Moore, A.J., Vogels, C.B.F., et al., 2020. Saliva is more sensitive for SARS-CoV-2 detection in COVID-19 patients than nasopharyngeal swabs (preprint). Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2020.04.16.20067835

Xing, Y.-H., Ni, W., Wu, Q., Li, W.-J., Li, G.-J., Wang, W.-D., Tong, J.-N., Song, X.-F., Wing-Kin Wong, G., Xing, Q.-S., 2020. Prolonged viral shedding in feces of pediatric patients with coronavirus disease 2019. J Microbiol Immunol Infect. https://doi.org/10.1016/j.jmii.2020.03.021

Xu, T., Chen, Cong, Zhu, Z., Cui, M., Chen, Chunhua, Dai, H., Xue, Y., 2020. Clinical features and dynamics of viral load in imported and non-imported patients with COVID-19. International Journal of Infectious Diseases S1201971220301417. https://doi.org/10.1016/j.ijid.2020.03.022

Xu, Y., Li, X., Zhu, B., Liang, H., Fang, C., Gong, Y., Guo, Q., Sun, X., Zhao, D., Shen, J., et al., 2020. Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. Nat. Med. 26, 502–505. https://doi.org/10.1038/s41591-020-0817-4

Xue, K.S., Moncla, L.H., Bedford, T., Bloom, J.D., 2018. Within-Host Evolution of Human Influenza Virus. Trends in Microbiology 26, 781–793. https://doi.org/10.1016/j.tim.2018.02.007

Xue, L., Cai, W., Gao, J., Zhang, L., Dong, R., Li, Y., Wu, H., Chen, M., Zhang, J., Wang, J., et al., 2019. The resurgence of the norovirus GII.4 variant associated with sporadic gastroenteritis in the post-GII.17 period in South China, 2015 to 2017. BMC Infect Dis 19, 696. https://doi.org/10.1186/s12879-019-4331-6

Yan, D., Zhang, X., Chen, C., Jiang, D., Liu, X., Zhou, Yuqing, Huang, C., Zhou, Yiyi, Guan, Z., Ding, C., et al., 2021. Characteristics of Viral Shedding Time in SARS-CoV-2 Infections: A Systematic Review and Meta-Analysis. Front. Public Health 9, 652842. https://doi.org/10.3389/fpubh.2021.652842

Yang, J.-R., Deng, D.-T., Wu, N., Yang, B., Li, H.-J., Pan, X.-B., 2020. Persistent viral RNA positivity during recovery period of a patient with SARS-CoV-2 infection. J. Med. Virol. https://doi.org/10.1002/jmv.25940

Yang, X., Charlebois, P., Macalalad, A., Henn, M.R., Zody, M.C., 2013. V-Phaser 2: variant inference for viral populations. BMC Genomics 14, 674. https://doi.org/10.1186/1471-2164-14-674

Yang, Y., Yang, M., Shen, C., Wang, F., Yuan, J., Li, Jinxiu, Zhang, M., Wang, Z., Xing, L., Wei, J., et al., 2020. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections (preprint). Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2020.02.11.20021493

Yoon, J.-J., Toots, M., Lee, S., Lee, M.-E., Ludeke, B., Luczo, J.M., Ganti, K., Cox, R.M., Sticher, Z.M., Edpuganti, V., et al., 2018. Orally Efficacious Broad-Spectrum Ribonucleoside Analog Inhibitor of Influenza and Respiratory Syncytial Viruses. Antimicrob Agents Chemother 62, e00766-18. https://doi.org/10.1128/AAC.00766-18

Young, B.E., Ong, S.W.X., Kalimuddin, S., Low, J.G., Tan, S.Y., Loh, J., Ng, O.-T., Marimuthu, K., Ang, L.W., Mak, T.M., et al., 2020. Epidemiologic Features and Clinical Course of Patients Infected With SARS-CoV-2 in Singapore. JAMA. https://doi.org/10.1001/jama.2020.3204

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.-Y., 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8, 28–36. https://doi.org/10.1111/2041-210X.12628

Yu, J., Liang, Z., Guo, K., Sun, X., Zhang, Q., Dong, Y., Duan, Z., 2020. Intra-Host Evolution of Norovirus GII.4 in a Chronic Infected Patient With Hematopoietic Stem Cell Transplantation. Front. Microbiol. 11, 375. https://doi.org/10.3389/fmicb.2020.00375

Zagordi, O., Bhattacharya, A., Eriksson, N., Beerenwinkel, N., 2011. ShoRAH: estimating the genetic diversity of a mixed sample from next-generation sequencing data. BMC Bioinformatics 12, 119. https://doi.org/10.1186/1471-2105-12-119

Zanini, F., Puller, V., Brodin, J., Albert, J., Neher, R.A., 2017. In vivo mutation rates and the landscape of fitness costs of HIV-1. Virus Evolution 3. https://doi.org/10.1093/ve/vex003

Zein, A.F.M.Z., Sulistiyana, C.S., Raffaello, W.M., Wibowo, A., Pranata, R., 2021. Sofosbuvir with daclatasvir and the outcomes of patients with COVID-19: a systematic review and meta-analysis with GRADE assessment. Postgrad Med J postgradmedj-2021-140287. https://doi.org/10.1136/postgradmedj-2021-140287

Zeng, W., Liu, G., Ma, H., Zhao, D., Yang, Yunru, Liu, M., Mohammed, A., Zhao, C., Yang, Yun, Xie, J., et al., 2020. Biochemical characterization of SARS-CoV-2 nucleocapsid protein. Biochemical and Biophysical Research Communications 527, 618–623. https://doi.org/10.1016/j.bbrc.2020.04.136

Zhang, H., Cockrell, S.K., Kolawole, A.O., Rotem, A., Serohijos, A.W.R., Chang, C.B., Tao, Y., Mehoke, T.S., Han, Y., Lin, J.S., et al., 2015. Isolation and Analysis of Rare Norovirus Recombinants from Coinfected Mice Using Drop-Based Microfluidics. J Virol 89, 7722–7734. https://doi.org/10.1128/JVI.01137-15

Zhang, W., Du, R.-H., Li, B., Zheng, X.-S., Yang, X.-L., Hu, B., Wang, Y.-Y., Xiao, G.-F., Yan, B., Shi, Z.-L., et al., 2020. Molecular and serological investigation of 2019-nCoV infected patients: implication of multiple shedding routes. Emerging Microbes & Infections 9, 386–389. https://doi.org/10.1080/22221751.2020.1729071

Zheng, D.-P., Ando, T., Fankhauser, R.L., Beard, R.S., Glass, R.I., Monroe, S.S., 2006. Norovirus classification and proposed strain nomenclature. Virology 346, 312–323. https://doi.org/10.1016/j.virol.2005.11.015

Zhou, B., She, J., Wang, Y., Ma, X., 2020. Duration of Viral Shedding of Discharged Patients With Severe COVID-19. Clinical Infectious Diseases 71, 2240–2242. https://doi.org/10.1093/cid/ciaa451

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., et al., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579, 270–273. https://doi.org/10.1038/s41586-020-2012-7

Zhu, W., Chen, C.Z., Gorshkov, K., Xu, M., Lo, D.C., Zheng, W., 2020. RNA-Dependent RNA Polymerase as a Target for COVID-19 Drug Discovery. SLAS Discovery 25, 1141–1151. https://doi.org/10.1177/2472555220942123

Zhu, X., He, Y., Wei, X., Kong, X., Zhang, Q., Li, J., Jin, M., Duan, Z., Key Laboratory of Medical Virology and Viral Diseases, Ministry of Health of the People's Republic of China, Beijing, China; National Institute for Viral Disease Control and Prevention, China CDC, Beijing, China, Shenzhen Center for Disease Control and Prevention, Shenzhen, Guangdong, China, 2021. Molecular Epidemiological Characteristics of Gastroenteritis Outbreaks Caused by Norovirus GII.4 Sydney [P31] Strains — China, October 2016–December 2020. China CDC Weekly 3, 1127–1132. https://doi.org/10.46234/ccdcw2021.276

Zou, L., Ruan, F., Huang, M., Liang, L., Huang, H., Hong, Z., Yu, J., Kang, M., Song, Y., Xia, J., et al., 2020. SARS-CoV-2 Viral Load in Upper Respiratory

Specimens of Infected Patients. N Engl J Med 382, 1177–1179. https://doi.org/10.1056/NEJMc2001737