

Review

Applications of Deep Mutational Scanning in Virology

Thomas D. Burton and Nicholas S. Eyre * 

College of Medicine and Public Health, Flinders University, Bedford Park, SA 5042, Australia;
burt0181@flinders.edu.au

* Correspondence: nicholas.eyre@flinders.edu.au

Abstract: Several recently developed high-throughput techniques have changed the field of molecular virology. For example, proteomics studies reveal complete interactomes of a viral protein, genome-wide CRISPR knockout and activation screens probe the importance of every single human gene in aiding or fighting a virus, and ChIP-seq experiments reveal genome-wide epigenetic changes in response to infection. Deep mutational scanning is a relatively novel form of protein science which allows the in-depth functional analysis of every nucleotide within a viral gene or genome, revealing regions of importance, flexibility, and mutational potential. In this review, we discuss the application of this technique to RNA viruses including members of the Flaviviridae family, Influenza A Virus and Severe Acute Respiratory Syndrome Coronavirus 2. We also briefly discuss the reverse genetics systems which allow for analysis of viral replication cycles, next-generation sequencing technologies and the bioinformatics tools that facilitate this research.

Keywords: deep mutational scanning; virology; virus; hepatitis; Zika; Dengue; Influenza; SARS-CoV-2



Citation: Burton, T.D.; Eyre, N.S. Applications of Deep Mutational Scanning in Virology. *Viruses* **2021**, *13*, 1020. <https://doi.org/10.3390/v13061020>

Academic Editor: Roy A. Hall

Received: 29 March 2021

Accepted: 26 May 2021

Published: 28 May 2021

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



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In its essence, deep sequencing is a tool that allows the sequencing of a genomic region multiple times. Deep mutational scanning (DMS) is a technique that utilises deep sequencing technology in combination with a library of mutant genes or genomes produced by random mutagenesis to probe the functional effects of mutations at every single nucleotide position within a gene or genome, linking genotype to phenotype in a single high-throughput experiment. This technique has been applied to various proteins to reveal residue-specific information regarding many aspects of protein biology. For example, under certain conditions, replacement of a yeast gene with a mutant library of a human orthologue can allow for determination of mutants which impact growth, and may be linked to human disease [1,2]. Unbiased selection and identification of mutants with improved properties or activities for specific requirements, such as increased solubility, have furthered the field of protein engineering [3]. Numerous other applications have also been pursued through DMS, including construction of complete functional activity landscapes of genes [4,5], establishment of quantitative evolutionary models [6,7], and contributions to structural biology [8,9]. In this review, we will predominantly focus on how DMS has been applied to the study of viral replicative fitness and immune evasion in the context of infectious virus replication cycles.

First, we should consider how residues and genomic regions of importance have been analysed by point mutagenesis in the past. Many studies have utilised alanine substitution to reveal residues or regions of importance in a viral protein. Commonly, residues for mutation are selected by analysis of protein structure, conservation with other isolates or strains, or amino acid biophysical properties [10,11]. Alternatively, large regions of interest may be interrogated by alanine scanning mutagenesis as a less targeted approach [12,13]. Alanine mutagenesis is the preferred substitution, as alanine features an inert, non-bulky methyl functional group, and does not alter main-chain conformation [14]. In these studies, mutants of interest are typically loss-of-function mutants, as this is indicative of the absence

of a functionally important residue. Experiments attempting to generate adaptive or gain-of-function mutations often rely on serial passages of a virus in cell culture or animal models [15–17], relying on an error-prone viral polymerase to generate mutants. A limiting factor is that combinations of mutations that may be required to enhance viral fitness in a given host may not be realised within the system. This method can also be time consuming as multiple rounds of infection are required. In regard to analysis of the impact of amino acid substitutions on viral protein function or viral replicative fitness, DMS can allow the substitution of a given residue with all possible amino acids, increasing the probability of identifying gain-of-function mutations that otherwise require more than one nucleotide substitution.

DMS studies in virology are generally performed using a three-step approach. First, a mutant library of a gene or genome of interest is prepared, ideally with genetic variants encoding every single possible residue change within the sequence of interest. Second, a selective pressure is applied to the library, enabling the enrichment of mutations that encode a selective advantage and the removal of deleterious mutations. Finally, the frequency of mutations within the library is quantified via next-generation sequencing (NGS) and compared before and after the application of the selective pressure. In studies of viral replicative fitness, a mutant with a cost to fitness will be selected against, while an enhancing mutant, for example, an antiviral escape mutant, will become enriched. Selective pressures, such as drug/antibody presence, stimulation of antiviral innate immunity, growth in cell types of different species, and binding potential to a host receptor have been applied to studies of many viral genes. Analysis of variants through DMS has enabled evolutionary studies, escape mutant predictions, attenuated vaccine construction and other applications. In this review, we will discuss various DMS strategies and how they have been applied in virology, with a focus on the *Flaviviridae* family of viruses, Influenza Virus A (IAV) and Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). DMS studies of RNA viruses have been made possible through the development of reverse genetics systems, or easily modified infectious clones, and advances in sequencing technologies, which we will discuss herein.

1.1. Influenza A

As members of the *Orthomyxoviridae* family, IAVs are negative-sense single-stranded RNA (-ssRNA) viruses [18]. Influenza strains vary from seasonal variants which cause a significant health care burden, to pandemics such as the Spanish Flu which killed an estimated 50 million people from 1918 to 1919 [19]. Influenza circulates not only in the human population, but in species such as birds, pigs and cats [20]. Two major obstacles towards effective Influenza protection and treatment are antigenic shift and antigenic drift, which are mechanisms which, through mutation or gene reassortment, can functionally change Influenza biology. Antigenic drift, which also applies to other viruses such as *Flaviviridae* and *Coronaviridae* family members, is the accumulation of point mutations due to the error-prone virally encoded RNA-dependent RNA polymerase used to generate new copies of the genome during replication [21]. Antigenic shift allows the formation of new Influenza subtypes through reassortment of surface antigens haemagglutinin (HA) or neuraminidase (NA) and is facilitated by the segmented genome of Influenza. Antigenic shift occurs when an animal is infected with two or more different strains of Influenza virus, enabling reassortment of HA and NA viral RNA segments and leading to a novel Influenza strain to which the population is completely naïve [21]. This resulted in the first pandemic of the 21st Century: swine-origin IAV [22]. The prediction of mutant strains of IAV that are likely to arise is therefore imperative for prophylaxis and control, and the predictive power of DMS can be utilised for the identification of animal viruses with zoonotic potential, or to predict escape mutants that are insensitive to otherwise effective adaptive immune responses or antiviral therapies.

1.2. *Flaviviridae*

The *Flaviviridae* family of RNA viruses comprises multiple positive-sense single-stranded RNA (+ssRNA) enveloped viruses, including hepatitis C virus (HCV), Dengue virus (DENV) and Zika virus (ZIKV) [23]. These viruses have enormous impact around the globe, with DENV alone infecting approximately 390 million people each year [24]. For most *Flaviviridae* (and viruses in general), no therapeutics or vaccines exist, with exceptions for vaccines being yellow fever virus [25] and Japanese encephalitis virus [26]. In addition to furthering our understanding of basic *Flaviviridae* biology, DMS has enabled the construction of a vaccine candidate for ZIKV, which will be discussed later.

1.3. SARS-CoV-2

The *Coronaviridae* family of positive-sense single-stranded RNA (+ssRNA) viruses features several clinically relevant viruses including Middle East respiratory syndrome-related coronavirus (MERS-CoV), Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV), and the current pandemic SARS-CoV-2, all of which have emerged in the past two decades [27], as well as common cold-causing viruses such as HCoV-OC43 [28]. There is currently a major global research effort to develop therapeutics and vaccines against SARS-CoV-2 which ideally target regions of the virus that are incapable of mutational escape, which would otherwise reduce the effectiveness of the treatment. DMS has been employed to determine how mutants will affect SARS-CoV-2 neutralisation.

2. Reverse Genetics Systems to Study RNA Viruses

Due to the absence of DNA in the lifecycle of RNA viruses, the construction of cDNA clones of infectious RNA viruses has become an incredibly important tool in studying their lifecycle. Generally, a reverse genetics system is comprised of genomic viral RNA that is reverse-transcribed into cDNA and then cloned into a plasmid, allowing for stable propagation of a virus genome within bacteria or yeast. The plasmid can then be easily manipulated by standard molecular methods, with the introduction of mutations and tags, or the removal of segments of the genome [29].

There are multiple key requirements for an infectious (+)RNA virus cDNA clone. A common approach involves incorporation of a DNA-dependent RNA polymerase promoter at the 5' end of a cloned viral genome to enable *in vitro* transcription of infectious viral RNA via a corresponding RNA polymerase, with or without a type I 5' cap structure, if required. This is commonly achieved using bacteriophage promoters/polymerases such as those of T7 and SP6. Produced RNA can then be purified and transfected into cells to initiate the viral replication cycle. A constitutive promoter such as a human cytomegalovirus [HCMV] promoter can also be utilised, with viral RNA produced by host Polymerase II after direct transfection of a full length cDNA clone into cells. As it is necessary to produce viral RNA with precise ends, a self-cleaving hepatitis delta or hammerhead ribozyme or a T7 terminator sequence may be added to the 3' end of the viral genome to enable generation of authentic 3' ends [30,31]. Alternatively, a unique restriction endonuclease site can be inserted at the 3' end for plasmid linearization [32]. Some of the limitations of these plasmid systems are instability in bacteria, due to the presence of cryptic bacterial promoters and other factors leading to recombination during growth, and poor plasmid yields [33]. As the preparation of mutant libraries involving plasmid clones of viral cDNA often requires the pooling of a large number of bacterial colonies, it is important in DMS studies to have a plasmid with both minimal recombination, to ensure that recombination of the plasmid in bacteria does not affect cell culture experiments/analysis, as well as high transformation efficiency to ensure that a library of sufficient mutational diversity can be prepared. For these reasons, several bacterium-free approaches such as circular polymerase extension reaction (CPER) [34] and yeast artificial chromosome (YAC) approaches, including transformation-associated recombination (TAR) cloning in *Saccharomyces cerevisiae* [35], have also been employed to enable efficient propagation and manipulation of (+)RNA virus cDNA clones. Similarly, bacterial artificial chromosome (BAC) systems and modified

plasmid DNA clones with features designed to minimise viral cDNA recombination and toxicity in bacteria have also been applied to various reverse genetics systems for (+)RNA viruses [31]. The reliability and ease of manipulation of viral cDNA using these systems are important determinants of the success of DMS experiments.

Many significant hurdles were encountered during the construction of an Influenza reverse genetics system. The Influenza genome comprises eight negative-sense ssRNA segments [18]. Due to the negative-sense genome, the viral RNA is not sufficient for initiation of viral replication, with the presence of multiple viral proteins being required to begin the viral lifecycle [18]. The eight vRNAs must colocalise with this protein machinery within the nucleus for initiation of infection [18]. Many reverse genetics systems have been developed to overcome these difficulties, with the two main categories comprising helper virus-dependent and -independent systems. Please see the review by Neumann et al. for a more in-depth analysis of the history of Influenza reverse genetics systems [36]; and for detailed documentation of these systems, please see the review by Engelhardt et al. [37].

Reverse genetics systems exist for many Flaviviridae species. Unlike Influenza A, the construction of infectious clones for *Flaviviridae* was, at the time, hindered more by the lack of available technologies such as high-fidelity PCR, instead of the fundamental biology of the virus. Construction of these systems is often considered as relatively straightforward although, as detailed above, challenges in genome construction due to repetitive elements, toxicity in *E. coli* and associated instability and recombination in *E. coli* are well-documented. For a history of Flavivirus reverse genetics systems, please see the review by Aubry et al. [31].

A reverse genetics system has recently been developed for SARS-CoV-2 [35]. However, as no DMS experiments have been performed to date using a reverse genetics system for the virus, we will not discuss this further. However, we highlight this as a potential area for future DMS studies.

Reverse genetics systems have traditionally been used in low-throughput mutational studies, to analyse the effect of single point mutations. Coupling reverse genetics with deep sequencing and random mutagenesis has allowed for high-throughput mutational studies, with analysis of hundreds of thousands of mutants being made possible in a single experiment. In the following sections, we discuss the technologies which have been essential in enabling DMS studies.

3. Next-Generation Sequencing

Advances in next-generation sequencing (NGS) have been crucial to the development of DMS as a tool in the field of molecular virology. The first generation of sequencing consisted mainly of Sanger sequencing and the Maxam and Gilbert technique. The second generation of sequencing introduced mass parallelisation of reactions. The current generation of sequencers enable real-time, single molecule sequencing [38]. We will discuss three commonly used platforms in DMS studies; Pacific Biosciences (PacBio) Single Molecule Real-Time (SMRT) sequencing, Oxford Nanopore Technologies (ONT) real-time sequencing, and the Illumina short-read sequencing-by-synthesis technology, which is part of the second generation of sequencing. [39]. For further in-depth detailing of the latest advances in NGS, please see the review by Goodwin et al. [40].

3.1. Single-Molecule Long-Read Sequencing

Produced by Pacific Biosciences, PacBio SMRT technology (henceforth referred to as SMRT) utilises a sequencing by synthesis approach. Initially, the DNA is fragmented into pieces several kilobases in length and the addition of hairpin adaptors to the DNA results in the formation of a circular SMRTbell DNA conformation. Next, the circular DNA is introduced to a flow cell lined with picolitre wells with a transparent bottom (a zero-mode waveguide) and a DNA polymerase enzyme fixed to the bottom of the well. The polymerase then incorporates a fluorescently tagged nucleotide into the elongating DNA strand, and the fluorescent signal emitted by the individual nucleotide being incorporated

is recorded by a camera. Finally, each fluorophore is cleaved by the polymerase and diffuses before the next read occurs. A major advantage of this technology is that the circular SMRTbell DNA conformation allows for many rounds of sequencing of a single DNA fragment, producing an accurate circular consensus sequence [41] (Figure 1).

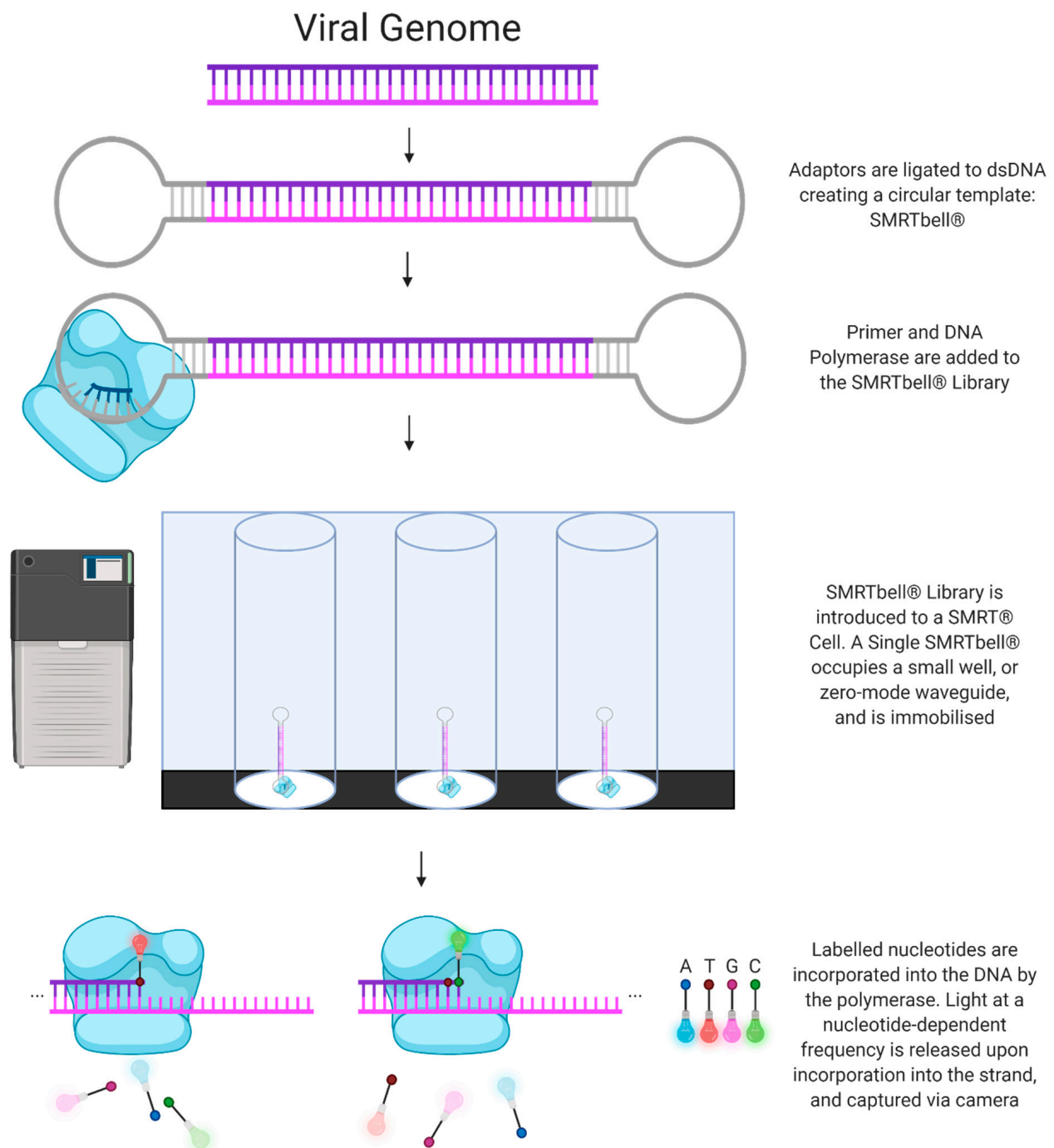


Figure 1. Next-Generation Sequencing: PacBio SMRT Technology.

The Oxford Nanopore system (henceforth referred to as ONT) sequences DNA in a unique manner. DNA is fragmented into pieces of several kilobases in length, then a motor protein and a hairpin adaptor are added to either side of the DNA. A leader sequence directs the DNA to a pore embedded in an electrically resistant membrane, and the motor protein allows ssDNA to be pulled through the aperture of the pore. Simultaneously, an electric current passes through the pore protein. As the ssDNA passes through the pore, a

characteristic disruption of the electrical current occurs, dependent on the multiple bases present in the pore. By analysis of this disruption, a DNA sequence can be identified by its unique 'k-mer', which can be translated into a sequence (for example, AAGT will have a distinct disruption compared to AGAT). The hairpin adaptor allows for bidirectional sequencing of the DNA fragment, and the read from the forward and reverse strands can be used to generate a consensus sequence [42] (Figure 2).

3.2. Illumina Sequencing by Synthesis

Used in Illumina sequencing instruments, DNA molecules are sheared to ≤ 300 base pair fragments then ligated to adapter sequences which allow for hybridisation to complementary oligonucleotides present in nanowells across a patterned flow cell. The DNA fragments are amplified via bridge amplification, resulting in clonal clusters of DNA. Subsequent addition and imaging of fluorophore-labelled terminator nucleotides allows for highly parallel sequencing [43] (Figure 3).

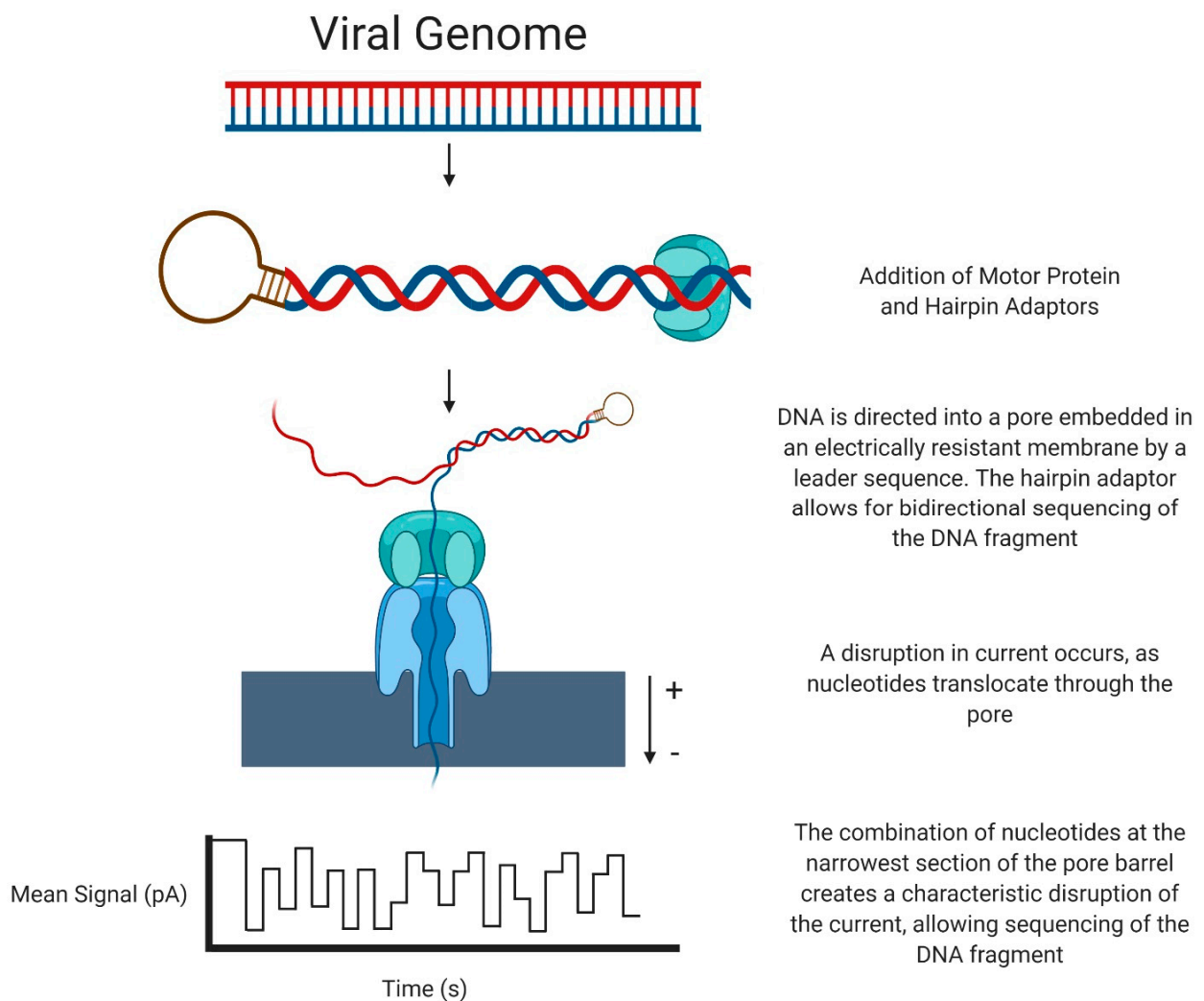


Figure 2. Next-Generation Sequencing: Oxford Nanopore Technology.

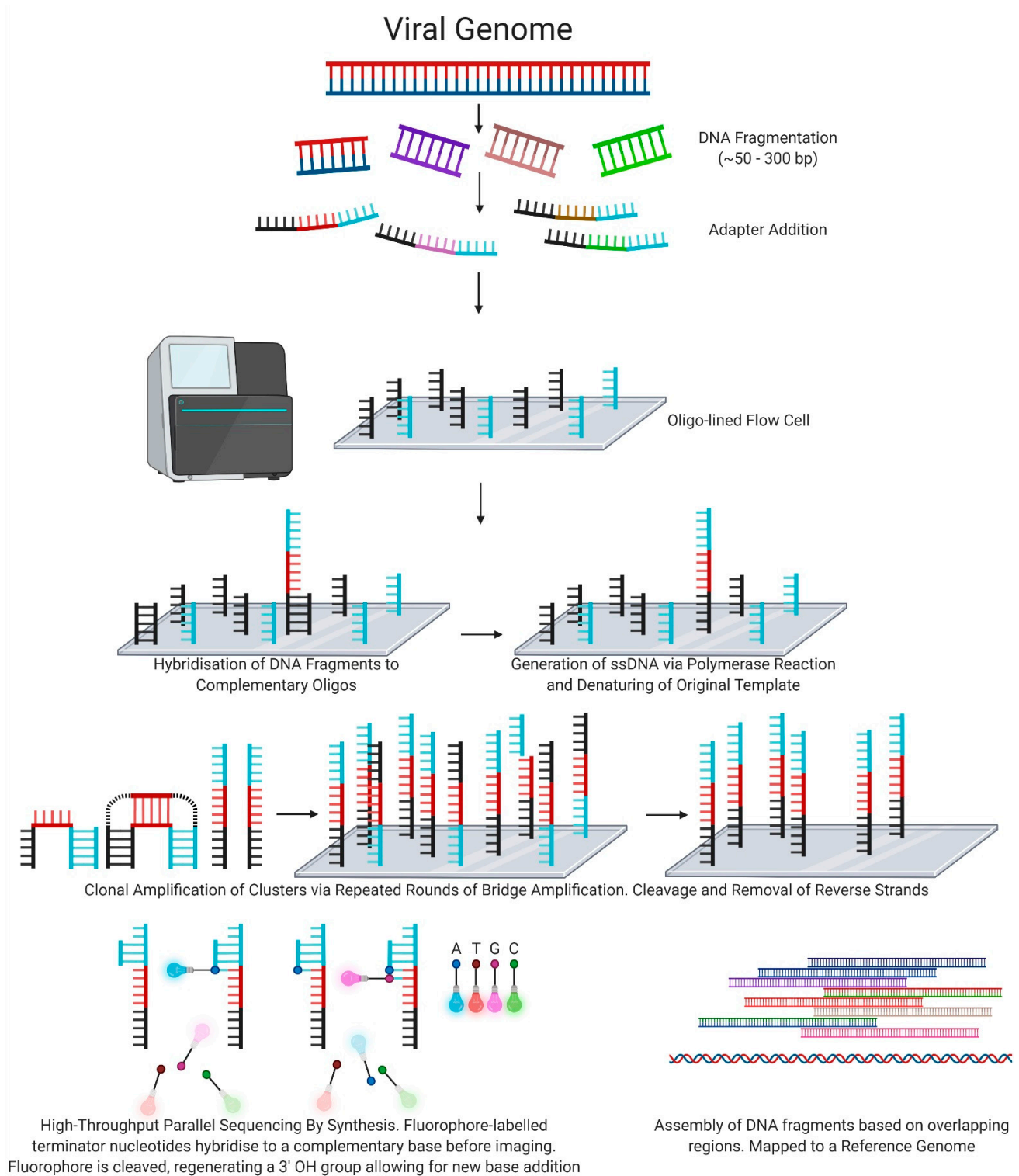


Figure 3. Next-Generation Sequencing: Illumina short-read technology.

3.3. Next-Generation Sequencing Technologies: A Comparison

The main limitation of Illumina short-read technology in DMS studies is the short-read length. Viral segments for analysis are often much longer than the read length of the instrument. If two distant mutants are present in a single viral genome, and a specific phenotype is observed, it is difficult to determine without further experimentation if they are acting in a pairwise manner or if a single mutation is wholly responsible for

the phenotype. Thus, pairwise epistatic mutations are not always resolvable using short-read technology. A partial solution to this problem is to create multiple adjacent mutant libraries of short lengths, ‘tiling’ across a gene, to restrict epistatic mutations to within a readable window [44]. Alternatively, subassembly can be utilised. In this approach, tagging each template molecule with a unique DNA barcode or ‘unique molecular identifier’ (UMI) allows for the grouping and analysis of multiple short reads on the basis of their original template molecule [45]. Resolution of distant mutants is only possible using third generation long-read sequencing strategies, offered by both SMRT and ONT.

Of the long-read sequencing platforms, SMRT is far more common than ONT in DMS studies, likely attributed to its higher accuracy. The circular SMRTbell DNA conformation allows for increased sequencing depth (how often a base is sequenced on average) to form a consensus sequence, as the SMRTbell can be repeatedly sequenced. Errors in PacBio sequencing are distributed randomly, and therefore the accuracy increases with increased reads of a single SMRTbell molecule [46]. Oxford Nanopore sequencing utilises dsDNA fragments, but does not allow continuous sequencing. A consensus sequence can be formed by sequencing one strand of dsDNA, followed immediately by the complementary strand, referred to as 1D² sequencing, though the accuracy is somewhat limited compared to PacBio and Illumina strategies [46]. Interestingly, a recent study has reported a high-throughput amplicon sequencing approach that combines UMIs with PacBio or ONT to enable generation of high-accuracy single-molecule consensus sequences for large DNA regions [47]. This and similar approaches will help to overcome compromises in accuracy that have previously been associated with the above long-read sequencing strategies. To date, however, Illumina sequencing remains the most commonly used platform in DMS studies due to its cost-effectiveness and high levels of accuracy.

4. Deep Mutational Scanning: Library Construction

As mentioned previously, the first step in a DMS experiment is the construction of a mutant library. Multiple approaches have been developed and applied to mutant library generation, which we will briefly describe below.

The simplest and most cost-effective method is error-prone PCR. In this method, one or more polymerases are employed to exponentially amplify a region of DNA, with initial template amount and cycle number varied to optimise mutation rate [48]. *Taq* DNA polymerase, a popular error-prone polymerase, has a reaction buffer-dependent mutation rate of roughly 8×10^{-6} errors/nucleotide [49]. However, the use of *Taq* polymerase alone in library construction is limited by the mutants generated being dominated by AT → GC transitions and AT → TA transversions, at approximately 2–4x the mutation rate of G and C residues. A polymerase named Mutazyme DNA polymerase has a 2–4x stronger preference for GC → AT transitions and GC → TA transversions [50]. In combination, these enzymes can be used to produce a relatively unbiased mutant library with a somewhat controllable mutation rate. A major drawback of this method is that not all amino acid residues are accessible with single-nucleotide polymorphisms [51]. Additionally, this technique can often result in multiple mutations present in a single DNA fragment that can confound results, especially when using short-read sequencing strategies as mentioned earlier.

First described in virology studies by the Bloom lab in an Influenza nucleoprotein (NP) study, synthetic oligonucleotides were designed to contain a randomised triplet for each codon present in the gene, with 16 leading and lagging nucleotides that anneal specifically to the NP gene, as well as the reverse complement of these oligonucleotides. Using a series of joining PCR steps and restriction enzyme cloning, a product pool containing all possible amino acid mutations of NP was cloned into an Influenza reverse genetics system [6]. When designing oligonucleotides with randomised triplets, codon usage should be considered. NNK degeneracy (N: Ade/Cyt/Gua/Thy, K: Gua/Thy) encodes all amino acids, while NDT (N: Ade/Cyt/Gua/Thy, D: Ade/Gua/Thy, T: Thy) and DBK (D: Ade/Gua/Thy, B: Cyt/Gua/Thy, K: Gua/Thy) each encode 12 amino acids, feature no stop codons and exhibit all major biophysical types, while potentially decreasing workload. However,

with decreased coverage, the amount of interesting variants will also be reduced [52]. This method is more costly compared to the error-prone PCR method, but can be used to generate libraries with complete mutational coverage.

A synthetic approach to mutant library generation is also available. Gene synthesis begins with oligonucleotide construction. These are designed such that adjacent oligonucleotides in the final product contain overlapping sequences. The overlap results in a DNA duplex which, through an assembly reaction with DNA polymerase, results in the construction of the gene of interest [53]. Multiple methods allow for controlled or randomised insertion of mutants into a gene. For example, the “Spiked Genes” method uses oligonucleotides spiked, or interspersed, with mutants to create a pooled mutant fragment [54]. Synthetic construction of multiple types of mutant libraries is possible through commercially available gene synthesis services. Available mutant library types include controlled randomised libraries, scanning alanine mutagenesis libraries, scanning codon mutagenesis libraries and others.

An additional type of mutant library discussed in this review is the transposon insertion library. This is facilitated by a transposase protein, such as Tn5, Tn7 or MuA that recognizes the ends of a transposon, or a mobile DNA element, and forms a protein-DNA complex termed the ‘transpososome’. This complex catalyses cleavage of target DNA at a random site (although evidence of insertion bias exists [55]), and joining reactions allow for the introduction of the transposon into a DNA template. For many applications, the transposon is engineered to feature an antibiotic resistance gene, allowing for selection of genetic elements with successful transposon integration. In these systems the transposon also features two identical restriction enzyme sites on each end of the DNA element. Upon purification, the now-unique plasmids are digested with the required restriction enzyme. Gel electrophoresis allows for the removal of the majority of the introduced DNA element by size separation, with a small insertion remaining after ligation of the plasmid backbone [56,57]. High-throughput random mutagenesis of a cloned viral genome paired with deep scanning allows for incredibly powerful studies which can probe regions of genomic flexibility and functionality.

5. Standard Deep Mutational Scanning Experiment Methodology

DMS experiments often follow a similar methodology. Using a reverse genetics system, a mutant virus library is generated via a method such as randomised mutagenesis, controlled site-saturation mutagenesis, or transposon mutagenesis. This library is often initially amplified using bacteria, with a higher number of uniquely transformed plasmids translating to increased mutational diversity. This library is transfected into cells, allowing for the propagation of RNA replication-competent viral genomes, as well as production of viral particles. Infection of naïve cells with virus-containing cell culture supernatants allows for propagation of genomes which are capable of both RNA replication and infectious particle production. Variant analysis by NGS is then key to identifying specific variants which are present at higher or lower levels at certain stages of the viral lifecycle in comparison to the initial input (Figure 4).

While not a major focus in this review, it is important to note that DMS studies of individually expressed proteins have been highly informative, particularly in the context of the interactions of viral glycoproteins with host receptors or neutralising antibodies. For example, angiotensin-converting enzyme 2 (ACE2), the host receptor for SARS-CoV-2 [58], has been analysed for mutants which enhance binding to the spike glycoprotein, to explore the use of an engineered soluble ACE2 as a potential therapeutic [59,60]. The binding of computer-generated miniprotein inhibitors [61] and ACE2 decoys [62] to the spike glycoprotein has also been effectively optimised using DMS. A similar strategy was used to optimise affinity and specificity of a computationally designed protein that targets H1N1 Influenza haemagglutinin [63,64].

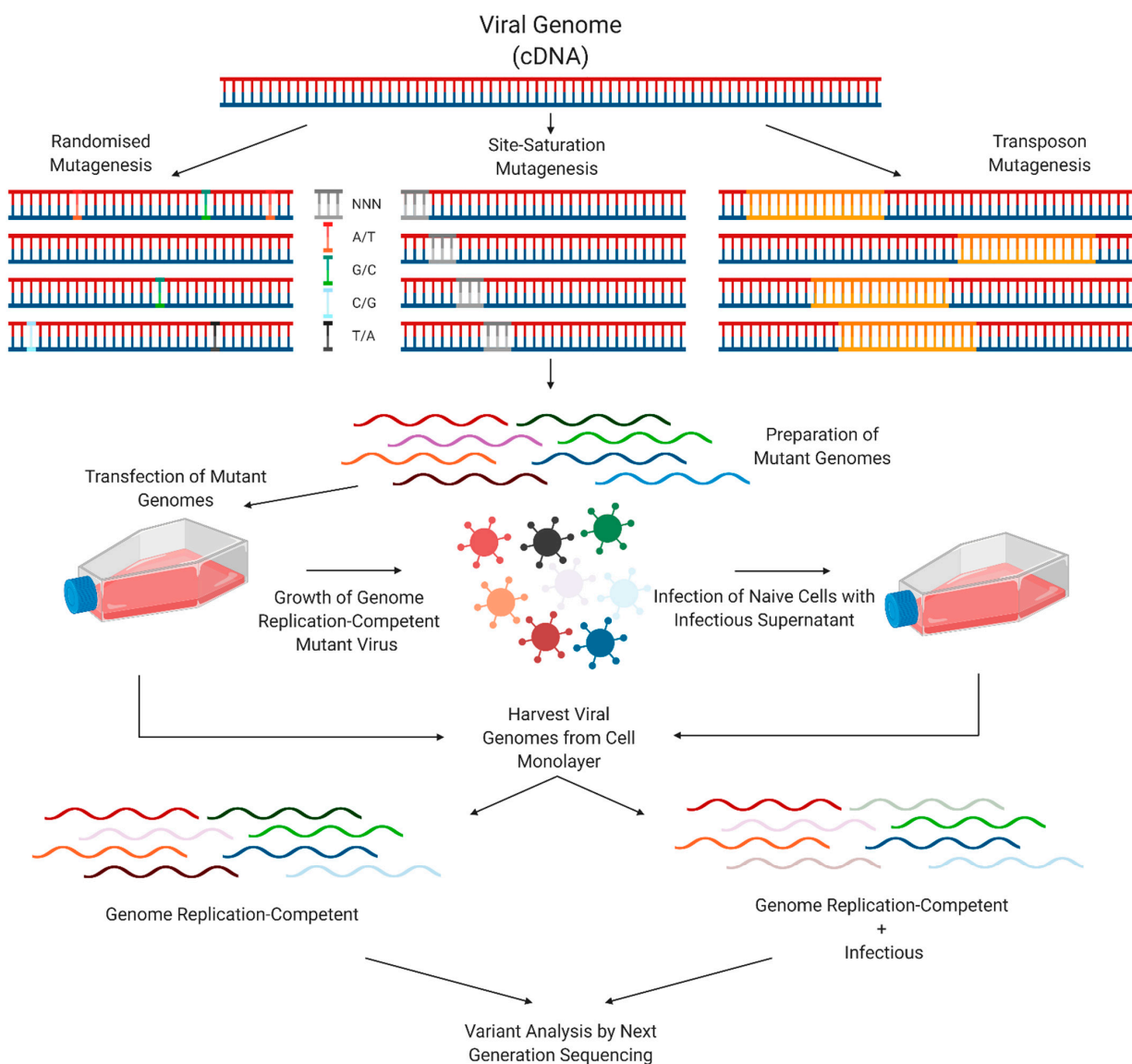


Figure 4. Standard methodology of a DMS experiment.

6. Bioinformatic Tools for Deep Mutational Scanning Data Analysis and Visualisation

Multiple software packages are available which facilitate variant analysis. ‘dms_tools’ [65] and ‘Enrich2’ [66] calculate amino acid preferences under a selective pressure. ‘dms-view’ allows straightforward visualisation of DMS data in the context of a protein structure [67]. ‘MaveDB’ (multiplex assays of variant effect database) provides a tool for sharing DMS data analyses.

7. Deep Mutational Scanning Experiments

In this section, we will briefly describe several past applications of DMS in virology, focussing on *Flaviviridae* family viruses, IAV and SARS-CoV-2. Transposon mutagenesis has been applied to several viruses to identify genomic regions tolerant to small insertions. An insertion with minimal impact on fitness indicates genomic flexibility, or high mutability, at a specific area, which may allow for adaptation of the virus to a new environment. Disruption of function from an insertion may also reveal a region of functionality within a gene or protein. The identification of a region of insertional tolerance is often exploited by incorporation of epitope tags or reporter genes for further research. DMS projects focussing on single-nucleotide polymorphisms and codon mutagenesis have focussed on

several biological aspects, including identifying determinants of viral tropism, epistatic interactions, drug and antibody escape mutants, and residues critical to several biological functions. It is important to note that the following summary is not a comprehensive list of all DMS scanning projects. DMS has also been applied to viruses including Norovirus [68], Venezuelan equine encephalitis virus [69], Measles virus [70], Paramyxoviruses [71], foot-and-mouth disease virus [72], Epstein–Barr virus [73] and Human immunodeficiency virus [74–79], and as mentioned previously, we focus mainly on studies of viral replicative fitness and immune evasion.

7.1. Hepatitis C Virus (HCV)

7.1.1. Transposon Analysis

In a comprehensive early study by Remenyi et al., the Mu-transposon system was used to insert 15 nucleotide (nt) sequences (of which 10 nt are transposon derived and 5 nt are duplicated target sites) into an HCV plasmid based on the chimeric sequence of genotype 2a J6 and JFH1 viruses [80]. This mutant library was propagated in human hepatoma Huh-7.5.1 cells, then passaged onto naïve Huh-7.5.1 cells to separately analyse genomes capable of viral RNA replication and infectious particle production. Regions tolerant to an insertion illustrate flexibility at the genomic level, and potentially highlight a region non-essential for *in vitro* growth. P7 and non-structural protein 2 (NS2), each coordinators of virus assembly [81], and envelope protein 2 (E2) were identified as potential areas for insertion of a tag or small peptide with minimal costs to replicative fitness due to the regions of high genomic flexibility uncovered. The impact of NS4B transposon insertions on infectious particle production helped to reveal a previously unknown functional region of NS4B which was suggested to play a role in the viral lifecycle post-RNA replication.

In a subsequent study by the same group, a similar methodology was used to assess insertions that confer sensitivity to the antiviral cytokine Interferon alpha (IFN- α) [82]. After two rounds of passage in Huh-7.5.1 cells in the presence or absence of IFN- α , it was observed that mutants conferring sensitivity to IFN- α were clustered in p7, the 3' untranslated region (UTR) and non-structural protein 5A (NS5A), which is important for viral replication, infectious virus particle production and modulation of host cell signalling [83]. Eight IFN- α sensitive p7 insertion mutants were constructed for validation, and revealed the role of p7 in immune evasion. An interferon-stimulated gene (ISG) cDNA expression library screen demonstrated that 13 of these ISGs were particularly antiviral against viruses with mutant p7 in comparison to wild-type p7. Amongst these ISGs, IFI6-16 was identified as a major target of p7-mediated immune evasion. Coimmunoprecipitation was then used to show a direct interaction between p7 and IFI6-16, which was previously unknown. Each of these DMS studies furthered understanding of the roles of HCV proteins in infection.

7.1.2. Variant Analysis

In an important application of DMS towards understanding emergence of viral resistance to antiviral drugs Qi et al. applied high-throughput variant analysis towards understanding resistance to Daclatasvir [84]; a potent inhibitor of NS5A [85]. An NS5A mutant library of the J6/JFH1 strain (genotype 2a) was prepared and passaged in Huh7.5.1 cells in the presence or absence of daclatasvir over 2 passages, before analysis by NGS. Drug selection resulted in selection of drug-resistant mutants at NS5A residues 28, 31, 38, 92 and 93, and increased drug resistance of some mutants at four additional residues. These escape mutants can be used to identify daclatasvir-resistant strains of HCV [86]. In support of this methodology, many of these resistance mutations have been identified in numerous *in vitro* and clinical studies of HCV resistance to Daclatasvir and related inhibitors [87].

7.2. Zika Virus (ZIKV)

7.2.1. Transposon Analysis

In the first published study involving high-throughput mutational analysis of ZIKV a transposon insertion mutant library was generated for the ZIKV MR-766 strain [88]. Human embryonic kidney 293T (HEK 293T) cells were used to propagate the virus, then naïve African green monkey kidney Vero cells were infected with viral supernatants over two rounds. The first pool of selected viral RNA, reflecting viral RNA replication requirements in HEK 293T cells, showed high levels of flexibility in the structural proteins Envelope (E), precursor membrane (prM) and capsid (C), and also non-structural protein 1 (NS1), which is essential for ER remodelling and RNA replication [89]. The third pool, consisting of infectious particle production-competent genomes, displayed most flexibility in the structural proteins. Some insertions present in NS1 rendered the virus incapable of infectious particle production while enhancing viral RNA replication capacity. In this context, an insertion after amino acid 174 resulted in >650-fold enrichment in the initial HEK 293T cell viral RNA population compared to the input RNA library yet was not present in the Vero cell populations. Together, this study revealed regions of genetic rigidity in the ZIKV genome across multiple lifecycle stages.

7.2.2. Variant Analysis

Envelope

Several mutational analyses of the ZIKV Envelope (E) protein have been performed. An envelope mutant library for the MR766 ZIKV strain was constructed by Sourisseau et al. to examine how mutants affect viral neutralisation by monoclonal antibodies [90]. The MR-766 strain was first tested for the effects of mutations on viral fitness, with a focus on infectious particle production in Vero cells. After validation, mutational antigenic profiling was performed, whereby the selective pressure of a neutralising antibody results in the enrichment of escape mutants present in the mutant library. Two antibodies were used in this screen; ZKA64, which binds domain III of recombinant E protein, and ZKA185 which neutralises ZIKV without interaction with E protein, or any confirmed epitope [91]. Domain III of E protein is responsible for binding the cellular receptor, and is a common target of potent ZIKV-neutralising antibodies [92]. The mutant ZIKV library was subjected to these antibodies individually. ZKA64 and ZKA185 treatment resulted in strong selection of several mutations that corresponded to sites within E, with mutants selected by ZKA64 present on envelope domain III, and mutants selected by ZKA185 present on envelope domain II. Antibody escape resistance mutations A333T and T335E for ZKA64, and D67A and K118R for ZKA185, were confirmed in follow-up neutralisation assays, illustrating the usefulness and accuracy of mutational antigenic profiling for escape mutant predictions, and demonstrating that antibody escape mutants can appear outside the receptor binding domain.

The impact of Envelope on tropism using ZIKV strain PRVABC59 has also been examined [93]. This study analysed mutant library growth in three cell types: C6/36 cells, derived from *Aedes albopictus* larvae, A549 adenocarcinomic human epithelial cells and hCMEC/D3 cells, a human blood–brain barrier endothelial cell line. NGS analysis revealed that a mutation at N154, an N-linked glycosylation site on the ZIKV envelope protein, resulted in a significantly higher fitness of ZIKV in mosquito cells, but not in human cells. Additionally, ablation of glycosylation through mutations surrounding the N154 position enhanced viral entry into C6/36 cells, resulting in increased levels of viral replication. DC-SIGN is a known viral entry factor for ZIKV that is not expressed in 293T cells such that overexpression of DC-SIGN in 293T cells enhances ZIKV infection [94]. A comparison between wild-type ZIKV and ZIKV glycosylation mutants showed that, with overexpression of DC-SIGN, infection of 293T cells by the glycosylation mutant ZIKV was significantly diminished compared to wild-type ZIKV, demonstrating the requirement of N-linked glycosylation of E in viral entry into mammalian cells, despite the resultant loss of fitness in mosquito cells.

In a similar study into cell tropism, a mutant library for the C-terminal of the E protein of ZIKV (PRVABC59) was prepared and grown in Vero and C6/36 cells for 8 and 13 days, respectively, to identify host-adaptive substitutions [95]. In C6/36 cells, mutants K316Q and S461G were preferentially selected. The K316Q/S461G virus replicated less efficiently in a selection of human cell lines, with lower levels of cytotoxicity compared to wild type, while no change was seen in C6/36 and Aag2 mosquito cell lines. This combinatorial mutant was determined to decrease the thermal stability of E, and as mosquito cell lines are typically grown at a lower temperature (28 °C) than human cell lines, the effect was not observed in the mosquito cell line. ZIKV infection can result in severe developmental defects in the brain [96]. The K316Q/S461G virus was used to infect induced-pluripotent stem cell-derived human brain organoids. Development of these organoids infected with the K316Q/S461G virus greatly decreased growth retardation in comparison to wild-type ZIKV. Infection of interferon alpha receptor knockout mice showed attenuated infection by the K316Q/S461G virus, and protection against a later wild-type ZIKV infection, indicating its potential as an attenuated vaccine candidate.

7.3. Dengue Virus (DENV)

Transposon Analysis

To identify regions of genetic flexibility in DENV, our group applied the Mu transposase system to generate a library of mutants containing single random 15 nt insertions in a serotype 2 Dengue virus (DENV-2) genome (strain 16681) [97]. Analysing both viral RNA replication and infectious particle production in Huh7.5 cells, non-structural protein NS1, required for RNA replication [98] and virus assembly [10], as well as structural protein capsid (C) demonstrated highest tolerance to the insertions. Lowest tolerance of insertions was seen in structural protein prM and non-structural protein 2A (NS2A), which is essential for virus assembly [99]. Across the genome, flexibility peaked at the C termini for C, E, NS1, NS2B and NS4B, and at the N termini of NS3 and NS4B. These experiments allowed for the tagging of NS1 with epitope tags and reporter proteins including FLAG, APEX2 and NLuc, with minimal loss of viral fitness in Huh7.5 cells.

A similar transposon mutagenesis study of DENV-2 (16681) was performed by Perry et al. [100]. This study identified regions of genetic flexibility and enabled identification of well-tolerated insertion sites within non-structural protein 4B (NS4B), which is involved in membrane rearrangements [101], and C, which were utilised to create a DENV-2 clone expressing HA-tagged capsid, incapable of infectious particle production, and a replication-competent infectious DENV-2 clone expressing HA-tagged NS4B. Together, these studies identified regions of genetic flexibility within DENV-2 and exploited this information to generate infectious reporter virus tools.

7.4. Influenza A Virus (IAV)

7.4.1. Transposon Analysis

Transposon mutagenesis has also been applied to great effect with the Influenza (H1N1) A/Puerto Rico/8/1934 strain [102]. Mutagenized viral genome segments were transfected into HEK 293T cells, then viruses were passaged onto naïve Madin–Darby canine kidney (MDCK) cells for two rounds. NGS analysis revealed that two genes were tolerant to insertions; haemagglutinin (HA), an attachment factor and membrane fusion protein [103], and non-structural protein 1, responsible for immune evasion [104]. The mutant library was then grown in embryonated chicken eggs, to determine how the library would perform under different environmental pressures. HA features a head and stalk domain [105]. In chicken eggs, insertions in the head of HA were preserved, while mutants present in the stalk domain were not tolerated. Analysis of these HA head insertion mutants showed no loss of fitness and were found to occur at antigenic sites. The authors hypothesised that this flexibility may reflect a strategy in evasion of host immune systems.

7.4.2. Variant Analysis Hemagglutinin and Neuraminidase

A major focus of IAV research is the functional analysis of NA, which aids virus budding, and HA, as antigenic variation of these proteins is responsible for the limited protection that vaccines provide against IAV infection [106]. Using a HA mutant library, Thyagarajan et al. illustrated that the antigenic sites of Influenza (H1N1) A/WSN/1933 HA are most tolerant to mutations, providing an explanation for Influenza's rapid antigenic evolution [107]. Similar effects were also reported in DMS studies by Wu et al. [108] and Boud et al. [109]. Influenza (H3N2) A/Perth/16/2009 HA has also been subject to analysis by DMS [110]. Interestingly, from these studies, mutational tolerance was demonstrated to be higher in the HA head for H1N1 Influenza strains, while mutational tolerance was higher in the HA stalk of H3N2 [110]. H3 virus DMS data was therefore shown to not be suitable for prediction of H1 virus evolution, while similarly H1 virus DMS data was not useful for prediction of H3 virus evolution.

Continuing research into HA and its immunogenicity, Doud et al. analysed the effects of single HA mutations on broad and narrow antibody neutralisation [111]. Broad neutralising antibodies are capable of neutralising multiple strains of a virus subtype, or neutralising multiple types of Influenza virus, due to their targeting of conserved regions of proteins, while narrow antibodies are strain specific [112]. Wild-type Influenza (H1N1) A/WSN/1933 was shown to be neutralised by broad anti-RBS and anti-stalk antibodies, and narrow anti-head antibodies. Doud et al. prepared HA mutant libraries of a lab-adapted Influenza (H1N1) A/WNS/1933 strain carrying green fluorescent protein in the PB1 segment and demonstrated that narrow anti-head antibodies as well as broad anti-RBS antibodies were highly susceptible to single mutation escape mutants, while broad anti-stalk antibodies were highly resistant to escape mutants.

Expanding on the single-strain studies of HA, multiple H3N2 IAV strains were investigated for strain-specific HA properties in a recent study by Wu et al. [113]. Five major antigenic sites (A-E) exist for H3N2 HA [114]. To investigate antigenic site B, mutant libraries of H3N2 strains A/Hong Kong/1/1968, A/Bangkok/1/1979, A/Beijing/353/1989, A/Moscow/10/1999, A/Brisbane/10/2007, and A/North Dakota/26/2016 were prepared. These libraries featured specific mutants of the HA receptor binding site at residues 156, 158, 159, 190, 193 and 196; each of which are known to affect receptor binding and viral fitness [115–118]. A local fitness landscape of each mutant for each strain was constructed, by transfection of the libraries into HEK 293T cells, and subsequent passage of the resultant virus onto MDCK-SIAT1 cells. Analysis of the fitness landscape of these strains illustrated that evolutionary constraints at these residues has changed during the natural evolution of H3N2, indicating that deep mutational scanning data is not always suitable for extrapolation across strains.

In another application of DMS to understand the evolution of IAV resistance to antiviral selective pressures, Wu et al. investigated the antiviral therapeutic Oseltamivir [119], which targets NA of various IAV strains [120]. Oseltamivir treatment has resulted in the generation of Influenza (H1N1) H274Y neuraminidase mutants which decrease sensitivity of the enzyme to oseltamivir. Despite fitness costs associated with the H274Y mutant, the mutant has quickly become widespread [121]. Studying the Influenza (H1N1) A/WSN/1933 virus with the H274Y mutant present, Wu et al. prepared a NA mutant library. After successive passages in A549 cells, four mutants were discovered which either fully (R194G, E214D) or partially (L250P, F239Y) restore H274Y A/WSN/1933 virus to wild-type fitness levels without restored susceptibility to oseltamivir. This study illustrates the power of DMS to understand and predict drug-resistant mutant strains and second site mutations that restore fitness of otherwise attenuated drug-resistant virus strains.

Polymerase

Another IAV machinery that has been the subject of DMS studies is the viral polymerase. One of the most comprehensive studies of this nature involved analysis of adap-

tation of the complex from avian to human cells by Soh et al. [122]. The core of the Influenza virus RNA polymerase consists of the polymerase acidic protein (PA) C-terminal domain, polymerase basic protein 1 (PB1) and the N-terminal of polymerase basic protein 2 (PB2) [123]. PB2 is a known virulence determinant [124] and it has been shown that a single amino acid substitution (E627K) allows avian IAV to replicate efficiently in mammalian cells [125]. Bioinformatics analyses reviewing IAV sequence databases have revealed other mutants which have facilitated host adaptation in the past [126,127]. However, these databases are likely incomplete and limited to past outbreaks. To address these ambiguities, Soh et al. constructed a PB2 mutant library to reveal mutants of Influenza (H7N9) A/Green-winged Teal/Ohio/175/1986 PB2 which improve either polymerase activity or viral replication in human A549 cells in comparison to avian embryo CCL141 cells, to predict mutants which may allow expansion of transmission towards mammalian species. Thirty-three previously undescribed adaptive mutants were discovered. These adaptive mutations were clustered in patches, potentially revealing regions of PB2 responsible for host cell interactions.

Ozawa et al. applied a novel DMS approach to investigate avian Influenza polymerase mutant residues of PA, PB1 and PB2, and their effect on host range [128]. A virus possessing vRNA of both avian Influenza (H5N1) A/Muscovy duck/Vietnam/TY93/2007 and a laboratory-adapted Influenza (H1N1) A/WSN/33, including a partial HA deletion mutant featuring GFP was constructed. Growth of the virus was therefore limited to HA-expressing cell lines. Mutant libraries were generated for PA, PB1 and PB2. The mutant virus library was grown in 293-HA cells, and the highest GFP-expressing 293-HA cells were individually sorted by flow cytometry for infection of naïve MDCK-HA cells. Analysis of expression of matrix protein by quantitative reverse transcription PCR was used to identify 90 viruses for further analysis by deep sequencing. After additional testing, 11 mutants were further characterised in lung and bronchial epithelial cells as well as chicken fibroblast cells, for viral kinetics and polymerase activity, and then for viral replication and pathogenesis in mice, identifying multiple mutants which may increase virulence in mammals.

Continuing the molecular interrogation of PA, Wu et al. coupled DMS with in silico mutant stability predictions to identify functional regions of Influenza (H1N1) A/WSN/1933 PA [129]. A PA mutant library of this strain was used to infect A549 cells, allowing identification of both functionally and structurally important residues. In addition, a protein stability predictor, Rosetta, was used to identify residues of structural importance [130]. Pairing these methods of analysis allowed exclusion of structurally important residues from residues that were deemed to be deleterious by DMS, isolating only functional effects. The study identified multiple residues of PA which abolish polymerase activity with no detectable destabilization effect, as well as residues with little impact on polymerase activity, but a substantial decrease in infectious particle production. While most mutational studies focus on conserved regions for protein functionality, which can exclude species- and strain-specific functional residues, this study by Wu et al. illustrated an unbiased means of determining residues that are strictly functionally important, as opposed to structurally important.

Assorted Viral Proteins and DMS Applications

Another application of DMS that is gaining attention involves analysis of viral determinants of sensitivity to innate immune responses. As an example of this, Ashenberg et al. explored resistance to the antiviral host factor MxA [131]. IAV nucleoprotein (NP) forms the viral ribonucleoprotein complex with vRNA and the heterotrimeric RNA-dependent RNA polymerase, and acts as an elongation factor [123,132]. Antiviral host factor MxA is known to suppress viral transcription through its interaction with polymerase PB2 subunit and NP [133]. IAV strains have been identified which either block interferon-induced production of MxA, or, through adaptive mutations, directly confer resistance to MxA [134,135]. Ashenberg et al. determined the impact of all mutant NP residues of Influenza (H3N2) A/Aichi/2/1968 on MxA resistance in MDCK-SIAT1 canine kidney cells

either wild type or constitutively expressing MxA. Twenty nine mutants impacting MxA resistance were identified via the screen, with further analysis identifying mutants which increase sensitivity, or increase resistance.

Viral resistance and sensitivity to the antiviral effects of interferons (IFNs) is another emerging application of DMS methodology, as exemplified by a recent study by Du et al. [136]. Many interactions between IAV and the host immune system, as well as host evasion strategies, have long been a focus of interest [137]. Du et al. performed a genome-wide mutant screen for Influenza (H1N1) A/WNS/1933 to discover mutants that alter interferon sensitivity. Virus was grown first in HEK 293T cells, then passaged in A549 cells in the presence or absence of type 1 interferon pre- and post-infection, directly prior to analysis of mutant abundance. Twenty six missense mutations that decreased relative fitness under interferon treatment were randomly selected for further analysis. Eight mutants from viral polymerase subunit PB2, matrix protein M1 and non-structural protein 1 were chosen for further investigation, and then combined to create an interferon-sensitive virus. This mutant IAV showed no attenuation in IFN-deficient Vero cells, but significant attenuation in IFN-competent A549 cells compared to the wild-type IAV strain. Additionally, the virus was attenuated in IFN-competent mice and ferrets, but not IFN AR-/- mice, and acted as an attenuated vaccine, protecting mice and ferrets from broad viral challenges. This study therefore allowed the identification of potential attenuated vaccine candidates.

Further exploring IAV's capacity to evade immunity, Wu et al. constructed a mutant library for the vRNA segment of Influenza (H1N1) A/WSN/1933 which encodes NS1 [138], given the critical role of NS1 as an antagonist of the innate immune response [137]. The virus was passaged in the presence or absence of type I interferon in A549 cells. Multiple SNPs displayed >2-fold higher interferon sensitivity compared to WT virus. In particular, D92Y of NS1 showed 12-fold higher sensitivity to the interferon treatment. This mutant was demonstrated to disrupt a hydrophobic pocket on NS1, which appeared to be critical for anti-interferon activity. Further research demonstrated a critical role for the domain in inhibition of RIG-I ubiquitination, a requirement for functional interferon signalling [139].

DMS is an ideal technique for discovery of single-residue changes that impact on viral fitness. However, it can be challenging to investigate paired mutants of significance, even if the mutants happen to be present on a single genome within a DMS experiment dataset. Wu et al. combined a high-throughput genetic study with analysis of naturally occurring sequences to identify a compensatory mutation in matrix protein M1 of IAV [140]. M1 acts as an endoskeleton for the virus [141] and aids virus budding [142]. Transfection of HEK 293T cells with Influenza (H1N1) A/WNS/1933 with a mutant M segment, followed by infection of A549 cells allowed identification of deleterious mutants. Thirteen deleterious mutants were validated, and it was found that 3 of these were prevalent in nature, including M1 Q214H. The software CAPS (Coevolution analysis using protein sequences) allows for identification of co-evolving amino acids with a time-dependent analysis of the evolution of pairs of amino acids from a changing sequence [143]. The analysis using CAPS revealed that M1 214 coevolves with residues 121, 207, 209 and 214. The A209T mutation was found to be compensatory for Q214H, revealing a novel pairwise epistatic interaction.

7.5. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)

Variant Analysis

The recent devastating outbreak of SARS-CoV-2 prompted the application of DMS to investigation of binding of the viral spike protein to the angiotensin-converting enzyme 2 (ACE2) receptor [144]. Specifically, this study explored the effects of all amino acid mutations of the receptor-binding domain (RBD) of a Wuhan-Hu-1 SARS-CoV-2 isolate on its interaction with angiotensin-converting enzyme 2 (ACE2) to identify potential targets of future antiviral strategies [145]. Two mutant libraries were prepared and expressed in a yeast-surface display platform. RBD mutant expression was measured using a fluorescent tag and sorted by flow cytometry based on expression levels into bins for analysis by deep

sequencing. ACE2 binding was then measured by incubation of cells with fluorescently labelled ACE2, which was separated by flow cytometry into bins for further deep sequencing. Analysis revealed considerable mutational tolerance of RBD for ACE2 binding and normal expression levels. Mutants enhancing affinity for ACE2 were found at RBD sites Q493, Q498 and N501. After publication, the SARS-CoV-2 variant VOC-202012/01, containing an N501Y mutant, spread rapidly throughout the UK and South Africa, indicating the real-world applicability of DMS [146]. In a follow up study, these RBD mutant libraries were tested for escape from 10 human monoclonal antibodies [147]. Escape mutants were identified using the yeast display platform model, with FACS determining ACE2 binding levels. Most escape mutants appeared at the receptor-binding motif, where ACE2/RBD contact residues are present. Negative-stain electron microscopy was used to determine antigenic regions of the RBD for each antibody. This information allowed prediction of effective antibody cocktails which do not allow enrichment of escape mutants

8. Concluding Remarks

DMS is an incredibly powerful tool for rapid and high-throughput mutational analysis of viral proteins and genomes, and the above studies have illustrated the plethora of applications that DMS offers to virology. DMS has allowed the construction of attenuated vaccines, evolutionary studies, identification of functional regions of proteins, prediction of mutants that increase host range and many other applications. As the feasibility of NGS increases due to decreases in the cost of such studies, we will likely have a very detailed view of the impact of point mutations for many viral genes. We hope that the recent development of bioinformatic tools and databases which facilitate visualisation and sharing of gathered DMS data will enable and encourage researchers to utilise existing data for further experimentation, akin to the Protein Data Bank. The DMS-based discovery of a SARS-CoV-2 N501 RBD mutant with enhanced ACE2-binding affinity, prior to its natural appearance as a highly transmissible variant reinforces the utility of DMS. However, it also illustrates the need for strict biosafety measures in the field, in order to prevent the accidental or intentional release of laboratory-created gain-of-function mutants. It has been suggested that publication of specific gain-of-function mutations should be regulated to ensure they are not recreated and misused, with data available on request. Alternatively, as some have argued, studies which may produce gain-of-function mutations in viruses with pandemic potential should be carefully scrutinised and regulated. For additional reading, we refer to two Editorials for further thoughts on the ethical considerations of gain-of-function experiments [148,149].

Author Contributions: Conceptualization, N.S.E.; investigation, T.D.B.; writing—original draft preparation, T.D.B.; figure preparation, T.D.B.; writing—review and editing, T.D.B. and N.S.E.; supervision, N.S.E.; funding acquisition, N.S.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Health and Medical Research Council (NHMRC) of Australia, awarded to N.S.E. (APP1163662).

Data Availability Statement: Figures are created with BioRender.com.

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

References

1. Fowler, D.M.; Stephany, J.J.; Fields, S. Measuring the activity of protein variants on a large scale using deep mutational scanning. *Nat. Protoc.* **2014**, *9*, 2267–2284. [[CrossRef](#)] [[PubMed](#)]
2. Livesey, B.J.; Marsh, J.A. Using deep mutational scanning to benchmark variant effect predictors and identify disease mutations. *Mol. Syst. Biol.* **2020**, *16*. [[CrossRef](#)] [[PubMed](#)]
3. Klesmith, J.R.; Bacik, J.-P.; Wrenbeck, E.E.; Michalczyk, R.; Whitehead, T.A. Trade-offs between enzyme fitness and solubility illuminated by deep mutational scanning. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2265–2270. [[CrossRef](#)] [[PubMed](#)]
4. Li, C.; Qian, W.; MacLean, C.J.; Zhang, J. The fitness landscape of a tRNA gene. *Science* **2016**, *352*, 837–840. [[CrossRef](#)] [[PubMed](#)]

5. Sarkisyan, K.S.; Bolotin, D.A.; Meer, M.V.; Usmanova, D.R.; Mishin, A.S.; Sharonov, G.V.; Ivankov, D.N.; Bozhanova, N.G.; Baranov, M.S.; Soylemez, O.; et al. Local fitness landscape of the green fluorescent protein. *Nature* **2016**, *533*, 397–401. [[CrossRef](#)]
6. Bloom, J.D. An Experimentally Determined Evolutionary Model Dramatically Improves Phylogenetic Fit. *Mol. Biol. Evol.* **2014**, *31*, 1956–1978. [[CrossRef](#)]
7. Canale, A.S.; Cote-Hammarlof, P.A.; Flynn, J.M.; Bolon, D.N. Evolutionary mechanisms studied through protein fitness landscapes. *Curr. Opin. Struct. Biol.* **2018**, *48*, 141–148. [[CrossRef](#)]
8. Newberry, R.W.; Leong, J.T.; Chow, E.D.; Kampmann, M.; DeGrado, W.F. Deep mutational scanning reveals the structural basis for α -synuclein activity. *Nat. Chem. Biol.* **2020**, *16*, 653–659. [[CrossRef](#)]
9. Adkar, B.V.; Tripathi, A.; Sahoo, A.; Bajaj, K.; Goswami, D.; Chakrabarti, P.; Swarnkar, M.K.; Gokhale, R.S.; Varadarajan, R. Protein Model Discrimination Using Mutational Sensitivity Derived from Deep Sequencing. *Structure* **2012**, *20*, 371–381. [[CrossRef](#)]
10. Scaturro, P.; Cortese, M.; Chatel-Chaix, L.; Fischl, W.; Bartenschlager, R. Dengue Virus Non-structural Protein 1 Modulates Infectious Particle Production via Interaction with the Structural Proteins. *PLoS Pathog.* **2015**, *11*, e1005277. [[CrossRef](#)]
11. Wang, W.; Guan, M.; Liu, Y.; Xu, Q.; Peng, H.; Liu, X.; Tang, Z.; Zhu, Y.; Wu, D.; Ren, H.; et al. Alanine scanning mutagenesis of hepatitis C virus E2 cysteine residues: Insights into E2 biogenesis and antigenicity. *Virology* **2014**, *448*, 229–237. [[CrossRef](#)] [[PubMed](#)]
12. Urabe, M.; Hasumi, Y.; Kume, A.; Surosky, R.T.; Kurtzman, G.J.; Tobita, K.; Ozawa, K. Charged-to-Alanine Scanning Mutagenesis of the N-Terminal Half of Adeno-Associated Virus Type 2 Rep78 Protein. *J. Virol.* **1999**, *73*, 2682–2693. [[CrossRef](#)] [[PubMed](#)]
13. Murray, C.L.; Jones, C.T.; Tassello, J.; Rice, C.M. Alanine Scanning of the Hepatitis C Virus Core Protein Reveals Numerous Residues Essential for Production of Infectious Virus. *J. Virol.* **2007**, *81*, 10220–10231. [[CrossRef](#)]
14. Lefevre, F. Alanine-stretch scanning mutagenesis: A simple and efficient method to probe protein structure and function. *Nucleic Acids Res.* **1997**, *25*, 447–448. [[CrossRef](#)] [[PubMed](#)]
15. Dolan, P.T.; Taguwa, S.; Rangel, M.A.; Acevedo, A.; Hagai, T.; Andino, R.; Frydman, J. Principles of dengue virus evolvability derived from genotype–fitness maps in human and mosquito cells. *eLife* **2021**, *10*. [[CrossRef](#)] [[PubMed](#)]
16. Ilyushina, N.A.; Khalenkov, A.M.; Seiler, J.P.; Forrest, H.L.; Bovin, N.V.; Marjuki, H.; Barman, S.; Webster, R.G.; Webby, R.J. Adaptation of Pandemic H1N1 Influenza Viruses in Mice. *J. Virol.* **2010**, *84*, 8607–8616. [[CrossRef](#)]
17. Choi, E.-J.; Lee, Y.J.; Lee, J.-M.; Kim, Y.-J.; Choi, J.-H.; Ahn, B.; Kim, K.; Han, M.G. The effect of mutations derived from mouse-adapted H3N2 seasonal influenza A virus to pathogenicity and host adaptation. *PLoS ONE* **2020**, *15*, e0227516. [[CrossRef](#)]
18. Bouvier, N.M.; Palese, P. The biology of influenza viruses. *Vaccine* **2008**, *26*, D49–D53. [[CrossRef](#)]
19. Phillips, H. The Recent Wave of ‘Spanish’ Flu Historiography. *Soc. Hist. Med.* **2014**, *27*, 789–808. [[CrossRef](#)]
20. Ducatez, M.; Webster, R.; Webby, R. Animal influenza epidemiology. *Vaccine* **2008**, *26*, D67–D69. [[CrossRef](#)]
21. Kim, H.; Webster, R.G.; Webby, R.J. Influenza Virus: Dealing with a Drifting and Shifting Pathogen. *Viral Immunol.* **2018**, *31*, 174–183. [[CrossRef](#)] [[PubMed](#)]
22. Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team. Emergence of a Novel Swine-Origin Influenza A (H1N1) Virus in Humans. *N. Engl. J. Med.* **2009**, *360*, 2605–2615. [[CrossRef](#)]
23. Barrows, N.J.; Campos, R.K.; Liao, K.-C.; Prasanth, K.R.; Soto-Acosta, R.; Yeh, S.-C.; Schott-Lerner, G.; Pompon, J.; Sessions, O.M.; Bradrick, S.S.; et al. Biochemistry and Molecular Biology of Flaviviruses. *Chem. Rev.* **2018**, *118*, 4448–4482. [[CrossRef](#)] [[PubMed](#)]
24. Bhatt, S.; Gething, P.W.; Brady, O.J.; Messina, J.P.; Farlow, A.W.; Moyes, C.L.; Drake, J.M.; Brownstein, J.S.; Hoen, A.G.; Sankoh, O.; et al. The global distribution and burden of dengue. *Nat. Cell Biol.* **2013**, *496*, 504–507. [[CrossRef](#)] [[PubMed](#)]
25. Chen, L.H.; Wilson, M.E. Yellow fever control: Current epidemiology and vaccination strategies. *Trop. Dis. Travel Med. Vaccines* **2020**, *6*, 1. [[CrossRef](#)] [[PubMed](#)]
26. Hegde, N.R.; Gore, M.M. Japanese encephalitis vaccines: Immunogenicity, protective efficacy, effectiveness, and impact on the burden of disease. *Hum. Vaccines Immunother.* **2017**, *13*, 1320–1337. [[CrossRef](#)] [[PubMed](#)]
27. Munster, V.; Koopmans, M.; Van Doremalen, N.; Van Riel, D.; De Wit, E. A Novel Coronavirus Emerging in China—Key Questions for Impact Assessment. *N. Engl. J. Med.* **2020**, *382*, 692–694. [[CrossRef](#)] [[PubMed](#)]
28. Lau, S.K.P.; Lee, P.; Tsang, A.K.L.; Yip, C.C.Y.; Tse, H.; Lee, R.A.; So, L.-Y.; Lau, Y.-L.; Chan, K.-H.; Woo, P.C.Y.; et al. Molecular Epidemiology of Human Coronavirus OC43 Reveals Evolution of Different Genotypes over Time and Recent Emergence of a Novel Genotype due to Natural Recombination. *J. Virol.* **2011**, *85*, 11325–11337. [[CrossRef](#)]
29. Stobart, C.C.; Moore, M.L. RNA virus Reverse Genetics and Vaccine Design. *Viruses* **2014**, *6*, 2531–2550. [[CrossRef](#)]
30. Walpita, P.; Flick, R. Reverse genetics of negative-stranded RNA viruses: A global perspective. *FEMS Microbiol. Lett.* **2005**, *244*, 9–18. [[CrossRef](#)]
31. Aubry, F.; Nougairède, A.; Gould, E.A.; de Lamballerie, X. Flavivirus reverse genetic systems, construction techniques and applications: A historical perspective. *Antivir. Res.* **2015**, *114*, 67–85. [[CrossRef](#)]
32. Yun, S.-I.; Kim, S.-Y.; Rice, C.M.; Lee, Y.-M. Development and Application of a Reverse Genetics System for Japanese Encephalitis Virus. *J. Virol.* **2003**, *77*, 6450–6465. [[CrossRef](#)]
33. Willemsen, A.; Zwart, M.P. On the stability of sequences inserted into viral genomes. *Virus Evol.* **2019**, *5*, vez045. [[CrossRef](#)] [[PubMed](#)]
34. Edmonds, J.; Van Grinsven, E.; Prow, N.; Bosco-Lauth, A.; Brault, A.C.; Bowen, R.A.; Hall, R.A.; Khromykh, A.A. A Novel Bacterium-Free Method for Generation of Flavivirus Infectious DNA by Circular Polymerase Extension Reaction Allows Accurate Recapitulation of Viral Heterogeneity. *J. Virol.* **2012**, *87*, 2367–2372. [[CrossRef](#)] [[PubMed](#)]

35. Thao, T.T.N.; Labroussaa, F.; Ebert, N.; V'kovski, P.; Stalder, H.; Portmann, J.; Kelly, J.; Steiner, S.; Holwerda, M.; Kratzel, A.; et al. Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. *bioRxiv* **2020**. [[CrossRef](#)]
36. Neumann, G. Influenza Reverse Genetics—Historical Perspective. *Cold Spring Harb. Perspect. Med.* **2020**, a038547. [[CrossRef](#)]
37. Engelhardt, O.G. Many ways to make an influenza virus—Review of influenza virus reverse genetics methods. *Influenza Other Respir. Viruses* **2012**, *7*, 249–256. [[CrossRef](#)]
38. Heather, J.M.; Chain, B. The sequence of sequencers: The history of sequencing DNA. *Genomics* **2016**, *107*, 1–8. [[CrossRef](#)] [[PubMed](#)]
39. Lee, H.; Gurtowski, J.; Yoo, S.; Nattestad, M.; Marcus, S.; Goodwin, S.; Richard McCombie, W.; Schatz, M.C. Third-generation sequencing and the future of genomics. *bioRxiv* **2016**. [[CrossRef](#)]
40. Goodwin, S.; McPherson, J.D.; McCombie, W.R. Coming of age: Ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **2016**, *17*, 333–351. [[CrossRef](#)]
41. Eid, J.; Fehr, A.; Gray, J.; Luong, K.; Lyle, J.; Otto, G.; Peluso, P.; Rank, D.; Baybayan, P.; Bettman, B.; et al. Real-Time DNA Sequencing from Single Polymerase Molecules. *Science* **2009**, *323*, 133–138. [[CrossRef](#)] [[PubMed](#)]
42. Clarke, J.; Wu, H.-C.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H. Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nanotechnol.* **2009**, *4*, 265–270. [[CrossRef](#)] [[PubMed](#)]
43. Bentley, D.R.; Balasubramanian, S.; Swerdlow, H.P.; Smith, G.P.; Milton, J.; Brown, C.G.; Hall, K.P.; Evers, D.J.; Barnes, C.L.; Bignell, H.R.; et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nat. Cell Biol.* **2008**, *456*, 53–59. [[CrossRef](#)] [[PubMed](#)]
44. Wrenbeck, E.E.; Faber, M.; Whitehead, T.A. Deep sequencing methods for protein engineering and design. *Curr. Opin. Struct. Biol.* **2017**, *45*, 36–44. [[CrossRef](#)] [[PubMed](#)]
45. Hiatt, J.B.; Patwardhan, R.P.; Turner, E.; Lee, C.; Shendure, J. Parallel, tag-directed assembly of locally derived short sequence reads. *Nat. Methods* **2010**, *7*, 119–122. [[CrossRef](#)] [[PubMed](#)]
46. Amarasinghe, S.L.; Su, S.; Dong, X.; Zappia, L.; Ritchie, M.E.; Gouil, Q. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* **2020**, *21*, 1–16. [[CrossRef](#)] [[PubMed](#)]
47. Karst, S.M.; Ziels, R.M.; Kirkegaard, R.H.; Sørensen, E.A.; McDonald, D.; Zhu, Q.; Knight, R.; Albertsen, M. High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. *Nat. Methods* **2021**, *18*, 165–169. [[CrossRef](#)]
48. Cirino, P.C.; Mayer, K.M.; Umeno, D.; Arnold, F.H.; Georgiou, G. Generating Mutant Libraries Using Error-Prone PCR. In *Directed Evolution Library Creation*; Arnold, F.H., Georgiou, G., Eds.; Humana Press: Totowa, NJ, USA, 2003; pp. 3–9.
49. Cline, J. PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res.* **1996**, *24*, 3546–3551. [[CrossRef](#)]
50. Vanhercke, T.; Ampe, C.; Tirry, L.; Denolf, P. Reducing mutational bias in random protein libraries. *Anal. Biochem.* **2005**, *339*, 9–14. [[CrossRef](#)]
51. Miyazaki, K.; Arnold, F.H. Exploring Nonnatural Evolutionary Pathways by Saturation Mutagenesis: Rapid Improvement of Protein Function. *J. Mol. Evol.* **1999**, *49*, 716–720. [[CrossRef](#)]
52. Reetz, M.T.; Carballeira, J.D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.* **2007**, *2*, 891–903. [[CrossRef](#)]
53. Hughes, R.A.; Ellington, A.D. Synthetic DNA Synthesis and Assembly: Putting the Synthetic in Synthetic Biology. *Cold Spring Harb. Perspect. Biol.* **2017**, *9*, a023812. [[CrossRef](#)] [[PubMed](#)]
54. Cárcamo, E.; Roldán-Salgado, A.; Osuna, J.; Bello-Sanmartín, I.; Yáñez, J.A.; Saab-Rincón, G.; Viadiu, H.; Gaytán, P. Spiked Genes: A Method to Introduce Random Point Nucleotide Mutations Evenly throughout an Entire Gene Using a Complete Set of Spiked Oligonucleotides for the Assembly. *ACS Omega* **2017**, *2*, 3183–3191. [[CrossRef](#)] [[PubMed](#)]
55. Green, B.; Bouchier, C.; Fairhead, C.; Craig, N.L.; Cormack, B.P. Insertion site preference of Mu, Tn5, and Tn7 transposons. *Mob. DNA* **2012**, *3*, 3. [[CrossRef](#)] [[PubMed](#)]
56. Savilahti, H.; Rice, P.A.; Mizuuchi, K. The phage Mu transpososome core: DNA requirements for assembly and function. *EMBO J.* **1995**, *14*, 4893–4903. [[CrossRef](#)] [[PubMed](#)]
57. Haapa, S.; Taira, S.; Heikkinen, E.; Savilahti, H. An efficient and accurate integration of mini-Mu transposons in vitro: A general methodology for functional genetic analysis and molecular biology applications. *Nucleic Acids Res.* **1999**, *27*, 2777–2784. [[CrossRef](#)]
58. 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. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181*, 271–280.e8. [[CrossRef](#)]
59. Chan, K.K.; Dorosky, D.; Sharma, P.; Abbasi, S.A.; Dye, J.M.; Kranz, D.M.; Herbert, A.S.; Procko, E. Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science* **2020**, *369*, 1261–1265. [[CrossRef](#)]
60. Heinzelman, P.; Romero, P.A. Discovery of human ACE2 variants with altered recognition by the SARS-CoV-2 spike protein. *bioRxiv* **2020**. [[CrossRef](#)]
61. Cao, L.; Goresnik, I.; Coventry, B.; Case, J.B.; Miller, L.; Kozodoy, L.; Chen, R.E.; Carter, L.; Walls, A.C.; Park, Y.-J.; et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **2020**, *370*, eabd9909-431. [[CrossRef](#)]
62. Linsky, T.W.; Vergara, R.; Codina, N.; Nelson, J.W.; Walker, M.J.; Su, W.; Barnes, C.O.; Hsiang, T.-Y.; Esser-Nobis, K.; Yu, K.; et al. De novo design of potent and resilient hACE2 decoys to neutralize SARS-CoV-2. *Science* **2020**, *370*, eabe0075. [[CrossRef](#)]

63. Whitehead, T.A.; Chevalier, A.; Song, Y.; Dreyfus, C.; Fleishman, S.J.; De Mattos, C.; Myers, C.A.; Kamisetty, H.; Blair, P.; Wilson, I.A.; et al. Optimization of affinity, specificity and function of designed influenza inhibitors using deep sequencing. *Nat. Biotechnol.* **2012**, *30*, 543–548. [[CrossRef](#)] [[PubMed](#)]
64. Fleishman, S.J.; Whitehead, T.A.; Ekiert, D.C.; Dreyfus, C.; Corn, J.E.; Strauch, E.-M.; Wilson, I.A.; Baker, D. Computational Design of Proteins Targeting the Conserved Stem Region of Influenza Hemagglutinin. *Science* **2011**, *332*, 816–821. [[CrossRef](#)] [[PubMed](#)]
65. Bloom, J.D. Software for the analysis and visualization of deep mutational scanning data. *BMC Bioinform.* **2015**, *16*, 168. [[CrossRef](#)] [[PubMed](#)]
66. Rubin, A.; Gelman, H.; Lucas, N.; Bajjalieh, S.M.; Papenfuss, A.T.; Speed, T.P.; Fowler, D.M. A statistical framework for analyzing deep mutational scanning data. *Genome Biol.* **2017**, *18*, 150. [[CrossRef](#)] [[PubMed](#)]
67. Hilton, S.K.; Huddleston, J.; Black, A.; North, K.; Dingens, A.S.; Bedford, T.; Bloom, J.D. *dms-view*: Interactive visualization tool for deep mutational scanning data. *bioRxiv* **2020**. [[CrossRef](#)]
68. Thorne, L.; Bailey, D.; Goodfellow, I. High-Resolution Functional Profiling of the Norovirus Genome. *J. Virol.* **2012**, *86*, 11441–11456. [[CrossRef](#)] [[PubMed](#)]
69. Beitzel, B.F.; Bakken, R.R.; Smith, J.M.; Schmaljohn, C.S. High-Resolution Functional Mapping of the Venezuelan Equine Encephalitis Virus Genome by Insertional Mutagenesis and Massively Parallel Sequencing. *PLoS Pathog.* **2010**, *6*, e1001146. [[CrossRef](#)]
70. Fulton, B.O.; Sachs, D.; Beaty, S.M.; Won, S.T.; Lee, B.; Palese, P.; Heaton, N.S. Mutational Analysis of Measles Virus Suggests Constraints on Antigenic Variation of the Glycoproteins. *Cell Rep.* **2015**, *11*, 1331–1338. [[CrossRef](#)]
71. Ikegame, S.; Beaty, S.M.; Stevens, C.; Won, S.T.; Park, A.; Sachs, D.; Hong, P.; Lee, B.; Thibault, P.A. Genome-wide transposon mutagenesis of paramyxoviruses reveals constraints on genomic plasticity. *PLoS Pathog.* **2020**, *16*, e1008877. [[CrossRef](#)]
72. Herod, M.R.; Loundras, E.-A.; Ward, J.C.; Tulloch, F.; Rowlands, D.J.; Stonehouse, N.J. Employing transposon mutagenesis to investigate foot-and-mouth disease virus replication. *J. Gen. Virol.* **2015**, *96*, 3507–3518. [[CrossRef](#)] [[PubMed](#)]
73. Procko, E.; Berguig, G.Y.; Shen, B.W.; Song, Y.; Frayo, S.; Convertine, A.J.; Margineantu, D.; Booth, G.; Correia, B.E.; Cheng, Y.; et al. A Computationally Designed Inhibitor of an Epstein-Barr Viral Bcl-2 Protein Induces Apoptosis in Infected Cells. *Cell* **2014**, *157*, 1644–1656. [[CrossRef](#)]
74. Luo, Y.; Jacobs, E.Y.; Greco, T.M.; Mohammed, K.D.; Tong, T.; Keegan, S.; Binley, J.M.; Cristea, I.M.; Fenyö, D.; Rout, M.; et al. HIV–host interactome revealed directly from infected cells. *Nat. Microbiol.* **2016**, *1*, 16068. [[CrossRef](#)] [[PubMed](#)]
75. Haddox, H.K.; Dingens, A.S.; Bloom, J.D. Experimental Estimation of the Effects of All Amino-Acid Mutations to HIV’s Envelope Protein on Viral Replication in Cell Culture. *PLoS Pathog.* **2016**, *12*, e1006114. [[CrossRef](#)] [[PubMed](#)]
76. Haddox, H.K.; Dingens, A.S.; Hilton, S.K.; Overbaugh, J.; Bloom, J.D. Mapping mutational effects along the evolutionary landscape of HIV envelope. *eLife* **2018**, *7*. [[CrossRef](#)] [[PubMed](#)]
77. Dingens, A.S.; Haddox, H.K.; Overbaugh, J.; Bloom, J.D. Comprehensive Mapping of HIV-1 Escape from a Broadly Neutralizing Antibody. *Cell Host Microbe* **2017**, *21*, 777–787.e4. [[CrossRef](#)]
78. Dingens, A.S.; Arenz, D.; Weight, H.; Overbaugh, J.; Bloom, J.D. An Antigenic Atlas of HIV-1 Escape from Broadly Neutralizing Antibodies Distinguishes Functional and Structural Epitopes. *Immunity* **2019**, *50*, 520–532.e3. [[CrossRef](#)]
79. Dingens, A.S.; Pratap, P.; Malone, K.; Hilton, S.K.; Ketas, T.; Cottrell, C.A.; Overbaugh, J.; Moore, J.P.; Klasse, P.J.; Ward, A.B.; et al. High-resolution mapping of the neutralizing and binding specificities of polyclonal rabbit serum elicited by HIV Env trimer immunization. *bioRxiv* **2020**. [[CrossRef](#)]
80. Remenyi, R.; Qi, H.; Su, S.-Y.; Chen, Z.; Wu, N.C.; Arumugaswami, V.; Truong, S.; Chu, V.; Stokelman, T.; Lo, H.-H.; et al. A Comprehensive Functional Map of the Hepatitis C Virus Genome Provides a Resource for Probing Viral Proteins. *mBio* **2014**, *5*, e01469-14. [[CrossRef](#)]
81. Alazard-Dany, N.; Denolly, S.; Boson, B.; Cosset, F.-L. Overview of HCV Life Cycle with a Special Focus on Current and Possible Future Antiviral Targets. *Viruses* **2019**, *11*, 30. [[CrossRef](#)]
82. Qi, H.; Chu, V.; Wu, N.C.; Chen, Z.; Truong, S.; Brar, G.; Su, S.-Y.; Du, Y.; Arumugaswami, V.; Olson, C.A.; et al. Systematic identification of anti-interferon function on hepatitis C virus genome reveals p7 as an immune evasion protein. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2018–2023. [[CrossRef](#)] [[PubMed](#)]
83. Ross-Thriepland, D.; Harris, M. Hepatitis C virus NS5A: Enigmatic but still promiscuous 10 years on! *J. Gen. Virol.* **2015**, *96*, 727–738. [[CrossRef](#)] [[PubMed](#)]
84. Qi, H.; Olson, C.A.; Wu, N.C.; Ke, R.; Loverdo, C.; Chu, V.; Truong, S.; Remenyi, R.; Chen, Z.; Du, Y.; et al. A Quantitative High-Resolution Genetic Profile Rapidly Identifies Sequence Determinants of Hepatitis C Viral Fitness and Drug Sensitivity. *PLoS Pathog.* **2014**, *10*, e1004064. [[CrossRef](#)] [[PubMed](#)]
85. Gao, M.; Nettles, R.E.; Belema, M.; Snyder, L.B.; Nguyen, V.N.; Fridell, R.A.; Serrano-Wu, M.H.; Langley, D.R.; Sun, J.-H.; Li, D.R.O.; et al. Chemical genetics strategy identifies an HCV NS5A inhibitor with a potent clinical effect. *Nat. Cell Biol.* **2010**, *465*, 96–100. [[CrossRef](#)]
86. Scheel, T.; Gottwein, J.; Mikkelsen, L.S.; Jensen, T.B.; Bukh, J. Recombinant HCV Variants With NS5A From Genotypes 1–7 Have Different Sensitivities to an NS5A Inhibitor but Not Interferon- α . *Gastroenterology* **2011**, *140*, 1032–1042.e6. [[CrossRef](#)]
87. Wyles, D.L.; Luetkemeyer, A.F. Understanding Hepatitis C Virus Drug Resistance: Clinical Implications for Current and Future Regimens. *Top. Antivir. Med.* **2017**, *25*, 103–109.
88. Fulton, B.O.; Sachs, D.; Schwarz, M.C.; Palese, P.; Evans, M.J. Transposon Mutagenesis of the Zika Virus Genome Highlights Regions Essential for RNA Replication and Restricted for Immune Evasion. *J. Virol.* **2017**, *91*, e00698-17. [[CrossRef](#)]

89. Ci, Y.; Liu, Z.-Y.; Zhang, N.-N.; Niu, Y.; Yang, Y.; Xu, C.; Yang, W.; Qin, C.-F.; Shi, L. Zika NS1-induced ER remodeling is essential for viral replication. *J. Cell Biol.* **2020**, *219*. [[CrossRef](#)]
90. Sourisseau, M.; Lawrence, D.J.P.; Schwarz, M.C.; Storrs, C.H.; Veit, E.C.; Bloom, J.D.; Evans, M.J. Deep Mutational Scanning Comprehensively Maps How Zika Envelope Protein Mutations Affect Viral Growth and Antibody Escape. *J. Virol.* **2019**, *93*. [[CrossRef](#)]
91. Stettler, K.; Beltramello, M.; Espinosa, D.A.; Graham, V.; Cassotta, A.; Bianchi, S.; Vanzetta, F.; Minola, A.; Jaconi, S.; Mele, F.; et al. Specificity, cross-reactivity, and function of antibodies elicited by Zika virus infection. *Science* **2016**, *353*, 823–826. [[CrossRef](#)]
92. Yang, M.; Dent, M.; Lai, H.; Sun, H.; Chen, Q. Immunization of Zika virus envelope protein domain III induces specific and neutralizing immune responses against Zika virus. *Vaccine* **2017**, *35*, 4287–4294. [[CrossRef](#)] [[PubMed](#)]
93. Gong, D.; Zhang, T.-H.; Zhao, D.; Du, Y.; Chapa, T.J.; Shi, Y.; Wang, L.; Contreras, D.; Zeng, G.; Shi, P.-Y.; et al. High-Throughput Fitness Profiling of Zika Virus E Protein Reveals Different Roles for Glycosylation during Infection of Mammalian and Mosquito Cells. *iScience* **2018**, *1*, 97–111. [[CrossRef](#)] [[PubMed](#)]
94. Hamel, R.; Dejarnac, O.; Wichit, S.; Ekcharyawat, P.; Neyret, A.; Luplertlop, N.; Perera-Lecoin, M.; Surasombatpattana, P.; Talignani, L.; Thomas, F.; et al. Biology of Zika Virus Infection in Human Skin Cells. *J. Virol.* **2015**, *89*, 8880–8896. [[CrossRef](#)]
95. Setoh, Y.X.; Amarilla, A.A.; Peng, N.Y.; Griffiths, R.E.; Carrera, J.; Freney, M.E.; Nakayama, E.; Ogawa, S.; Watterson, D.; Modhiran, N.; et al. Determinants of Zika virus host tropism uncovered by deep mutational scanning. *Nat. Microbiol.* **2019**, *4*, 876–887. [[CrossRef](#)] [[PubMed](#)]
96. Wen, Z.; Song, H.; Ming, G.-L. How does Zika virus cause microcephaly? *Genes Dev.* **2017**, *31*, 849–861. [[CrossRef](#)] [[PubMed](#)]
97. Eyre, N.S.; Johnson, S.M.; Eltahla, A.A.; Aloï, M.; Aloia, A.L.; McDevitt, C.A.; Bull, R.A.; Beard, M.R. Genome-Wide Mutagenesis of Dengue Virus Reveals Plasticity of the NS1 Protein and Enables Generation of Infectious Tagged Reporter Viruses. *J. Virol.* **2017**, *91*, e01455-17. [[CrossRef](#)]
98. Fan, J.; Liu, Y.; Yuan, Z. Critical role of Dengue Virus NS1 protein in viral replication. *Virol. Sin.* **2014**, *29*, 162–169. [[CrossRef](#)]
99. Xie, X.; Zou, J.; Zhang, X.; Zhou, Y.; Routh, A.L.; Kang, C.; Popov, V.L.; Chen, X.; Wang, Q.-Y.; Dong, H.; et al. Dengue NS2A Protein Orchestrates Virus Assembly. *Cell Host Microbe* **2019**, *26*, 606–622.e8. [[CrossRef](#)]
100. Perry, J.W.; Chen, Y.; Speliotes, E.; Tai, A.W. Functional Analysis of the Dengue Virus Genome Using an Insertional Mutagenesis Screen. *J. Virol.* **2018**, *92*, 02085–02117. [[CrossRef](#)]
101. Nemésio, H.; Palomares-Jerez, F.; Villalaín, J. NS4A and NS4B proteins from dengue virus: Membranotropic regions. *Biochim. Biophys. Acta BBA Biomembr.* **2012**, *1818*, 2818–2830. [[CrossRef](#)] [[PubMed](#)]
102. Heaton, N.S.; Sachs, D.; Chen, C.-J.; Hai, R.; Palese, P. Genome-wide mutagenesis of influenza virus reveals unique plasticity of the hemagglutinin and NS1 proteins. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 20248–20253. [[CrossRef](#)] [[PubMed](#)]
103. Russell, C.J.; Hu, M.; Okda, F.A. Influenza Hemagglutinin Protein Stability, Activation, and Pandemic Risk. *Trends Microbiol.* **2018**, *26*, 841–853. [[CrossRef](#)] [[PubMed](#)]
104. Haye, K.; Burmakina, S.; Moran, T.; García-Sastre, A.; Fernandez-Sesma, A. The NS1 Protein of a Human Influenza Virus Inhibits Type I Interferon Production and the Induction of Antiviral Responses in Primary Human Dendritic and Respiratory Epithelial Cells. *J. Virol.* **2009**, *83*, 6849–6862. [[CrossRef](#)] [[PubMed](#)]
105. Kirkpatrick, E.; Qiu, X.; Wilson, P.C.; Bahl, J.; Krammer, F. The influenza virus hemagglutinin head evolves faster than the stalk domain. *Sci. Rep.* **2018**, *8*, 10432. [[CrossRef](#)] [[PubMed](#)]
106. Krammer, F.; Palese, P. Advances in the development of influenza virus vaccines. *Nat. Rev. Drug Discov.* **2015**, *14*, 167–182. [[CrossRef](#)] [[PubMed](#)]
107. Thyagarajan, B.; Bloom, J.D. The inherent mutational tolerance and antigenic evolvability of influenza hemagglutinin. *eLife* **2014**, *3*, e03300. [[CrossRef](#)]
108. Wu, N.C.; Young, A.P.; Al-Mawsawi, L.Q.; Olson, C.A.; Feng, J.; Qi, H.; Chen, S.-H.; Lu, I.-H.; Lin, C.-Y.; Chin, R.G.; et al. High-throughput profiling of influenza A virus hemagglutinin gene at single-nucleotide resolution. *Sci. Rep.* **2015**, *4*, 4942. [[CrossRef](#)]
109. Doud, M.B.; Bloom, J.D. Accurate Measurement of the Effects of All Amino-Acid Mutations on Influenza Hemagglutinin. *Viruses* **2016**, *8*, 155. [[CrossRef](#)]
110. Lee, J.M.; Huddleston, J.; Doud, M.B.; Hooper, K.A.; Wu, N.C.; Bedford, T.; Bloom, J.D. Deep mutational scanning of hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E8276–E8285. [[CrossRef](#)]
111. Doud, M.B.; Lee, J.M.; Bloom, J.D. How single mutations affect viral escape from broad and narrow antibodies to H1 influenza hemagglutinin. *Nat. Commun.* **2018**, *9*, 1386. [[CrossRef](#)]
112. Laursen, N.S.; Wilson, I.A. Broadly neutralizing antibodies against influenza viruses. *Antivir. Res.* **2013**, *98*, 476–483. [[CrossRef](#)] [[PubMed](#)]
113. Wu, N.C.; Otwinowski, J.; Thompson, A.J.; Nycholat, C.M.; Nourmohammad, A.; Wilson, I.A. Major antigenic site B of human influenza H3N2 viruses has an evolving local fitness landscape. *Nat. Commun.* **2020**, *11*, 1233. [[CrossRef](#)] [[PubMed](#)]
114. Lee, M.-S.; Chen, J.S.-E. Predicting Antigenic Variants of Influenza A/H3N2 Viruses. *Emerg. Infect. Dis.* **2004**, *10*, 1385–1390. [[CrossRef](#)] [[PubMed](#)]
115. Wu, N.C.; Thompson, A.J.; Xie, J.; Lin, C.-W.; Nycholat, C.M.; Zhu, X.; Lerner, R.A.; Paulson, J.C.; Wilson, I.A. A complex epistatic network limits the mutational reversibility in the influenza hemagglutinin receptor-binding site. *Nat. Commun.* **2018**, *9*, 1264. [[CrossRef](#)]
116. Chen, L.-M.; Blixt, O.; Stevens, J.; Lipatov, A.S.; Davis, C.T.; Collins, B.E.; Cox, N.J.; Paulson, J.C.; Donis, R.O. In vitro evolution of H5N1 avian influenza virus toward human-type receptor specificity. *Virology* **2012**, *422*, 105–113. [[CrossRef](#)]

117. Imai, M.; Watanabe, T.; Hatta, M.; Das, S.C.; Ozawa, M.; Shinya, K.; Zhong, G.; Hanson, A.; Katsura, H.; Watanabe, S.; et al. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nat. Cell Biol.* **2012**, *486*, 420–428. [[CrossRef](#)]
118. Parker, L.; Wharton, S.A.; Martin, S.R.; Cross, K.; Lin, Y.; Liu, Y.; Feizi, T.; Daniels, R.S.; McCauley, J.W. Effects of egg-adaptation on receptor-binding and antigenic properties of recent influenza A (H3N2) vaccine viruses. *J. Gen. Virol.* **2016**, *97*, 1333–1344. [[CrossRef](#)]
119. Wu, N.C.; Young, A.P.; Dandekar, S.; Wijersuriya, H.; Al-Mawsawi, L.Q.; Wu, T.-T.; Sun, R. Systematic Identification of H274Y Compensatory Mutations in Influenza A Virus Neuraminidase by High-Throughput Screening. *J. Virol.* **2012**, *87*, 1193–1199. [[CrossRef](#)] [[PubMed](#)]
120. Lew, W.; Chen, X.; Kim, C.U. Discovery and Development of GS 4104 (oseltamivir) An Orally Active Influenza Neuraminidase Inhibitor. *Curr. Med. Chem.* **2000**, *7*, 663–672. [[CrossRef](#)]
121. Hurt, A.C.; Ernest, J.; Deng, Y.-M.; Iannello, P.; Besselaar, T.G.; Birch, C.; Buchy, P.; Chittaganpitch, M.; Chiu, S.-C.; Dwyer, D.; et al. Emergence and spread of oseltamivir-resistant A(H1N1) influenza viruses in Oceania, South East Asia and South Africa. *Antivir. Res.* **2009**, *83*, 90–93. [[CrossRef](#)] [[PubMed](#)]
122. Soh, Y.S.; Moncla, L.H.; Eguia, R.; Bedford, T.; Bloom, J.D. Comprehensive mapping of adaptation of the avian influenza polymerase protein PB2 to humans. *eLife* **2019**, *8*, 8. [[CrossRef](#)] [[PubMed](#)]
123. Velthuis, A.J.W.T.; Fodor, A.J.W.T.V.E. Influenza virus RNA polymerase: Insights into the mechanisms of viral RNA synthesis. *Nat. Rev. Genet.* **2016**, *14*, 479–493. [[CrossRef](#)]
124. Boivin, S.; Cusack, S.; Ruigrok, R.W.H.; Hart, D.J. Influenza A Virus Polymerase: Structural Insights into Replication and Host Adaptation Mechanisms. *J. Biol. Chem.* **2010**, *285*, 28411–28417. [[CrossRef](#)]
125. Subbarao, E.K.; London, W.; Murphy, B.R. A single amino acid in the PB2 gene of influenza a virus is a determinant of host range. *J. Virol.* **1993**, *67*, 1761–1764. [[CrossRef](#)]
126. Mänz, B.; De Graaf, M.; Mögling, R.; Richard, M.; Bestebroer, T.M.; Rimmelzwaan, G.F.; Fouchier, R.A.M. Multiple Natural Substitutions in Avian Influenza A Virus PB2 Facilitate Efficient Replication in Human Cells. *J. Virol.* **2016**, *90*, 5928–5938. [[CrossRef](#)]
127. Miotto, O.; Heiny, A.T.; Tan, T.W.; August, J.T.; Brusica, V. Identification of human-to-human transmissibility factors in PB2 proteins of influenza A by large-scale mutual information analysis. *BMC Bioinform.* **2008**, *9*, S18. [[CrossRef](#)] [[PubMed](#)]
128. Taft, A.S.; Ozawa, M.; Fitch, A.; DePasse, J.V.; Halfmann, P.J.; Hill-Batorski, L.; Hatta, M.; Friedrich, T.C.; Lopes, T.J.S.; Maher, E.A.; et al. Identification of mammalian-adapting mutations in the polymerase complex of an avian H5N1 influenza virus. *Nat. Commun.* **2015**, *6*, 7491. [[CrossRef](#)]
129. Wu, N.C.; Olson, C.A.; Du, Y.; Le, S.; Tran, K.; Remenyi, R.; Gong, D.; Al-Mawsawi, L.Q.; Qi, H.; Wu, T.-T.; et al. Functional Constraint Profiling of a Viral Protein Reveals Discordance of Evolutionary Conservation and Functionality. *PLoS Genet.* **2015**, *11*, e1005310. [[CrossRef](#)]
130. Das, R.; Baker, D. Macromolecular Modeling with Rosetta. *Annu. Rev. Biochem.* **2008**, *77*, 363–382. [[CrossRef](#)]
131. Ashenberg, O.; Padmakumar, J.; Doud, M.B.; Bloom, J.D. Deep mutational scanning identifies sites in influenza nucleoprotein that affect viral inhibition by MxA. *PLoS Pathog.* **2017**, *13*, e1006288. [[CrossRef](#)] [[PubMed](#)]
132. Turrell, L.; Lyall, J.W.; Tiley, L.S.; Fodor, E.; Vreede, F.T. The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. *Nat. Commun.* **2013**, *4*, 1591. [[CrossRef](#)]
133. Turan, K. Nuclear MxA proteins form a complex with influenza virus NP and inhibit the transcription of the engineered influenza virus genome. *Nucleic Acids Res.* **2004**, *32*, 643–652. [[CrossRef](#)] [[PubMed](#)]
134. Mänz, B.; Dornfeld, D.; Götz, V.; Zell, R.; Zimmermann, P.; Haller, O.; Kochs, G.; Schwemmler, M. Pandemic Influenza A Viruses Escape from Restriction by Human MxA through Adaptive Mutations in the Nucleoprotein. *PLoS Pathog.* **2013**, *9*, e1003279. [[CrossRef](#)] [[PubMed](#)]
135. Hale, B.G.; Albrecht, R.A.; García-Sastre, A. Innate immune evasion strategies of influenza viruses. *Future Microbiol.* **2010**, *5*, 23–41. [[CrossRef](#)] [[PubMed](#)]
136. Du, Y.; Xin, L.; Shi, Y.; Zhang, T.-H.; Wu, N.C.; Dai, L.; Gong, D.; Brar, G.; Shu, S.; Luo, J.; et al. Genome-wide identification of interferon-sensitive mutations enables influenza vaccine design. *Science* **2018**, *359*, 290–296. [[CrossRef](#)] [[PubMed](#)]
137. Van De Sandt, C.E.; Kreijtz, J.H.C.M.; Rimmelzwaan, G.F. Evasion of Influenza A Viruses from Innate and Adaptive Immune Responses. *Viruses* **2012**, *4*, 1438–1476. [[CrossRef](#)]
138. Wu, N.C.; Young, A.P.; Al-Mawsawi, L.Q.; Olson, C.A.; Feng, J.; Qi, H.; Luan, H.H.; Li, X.; Wu, T.-T.; Sun, R. High-Throughput Identification of Loss-of-Function Mutations for Anti-Interferon Activity in the Influenza A Virus NS Segment. *J. Virol.* **2014**, *88*, 10157–10164. [[CrossRef](#)]
139. Loo, Y.-M.; Gale, M. Immune Signaling by RIG-I-like Receptors. *Immunity* **2011**, *34*, 680–692. [[CrossRef](#)]
140. Wu, N.C.; Du, Y.; Le, S.; Young, A.P.; Zhang, T.-H.; Wang, Y.; Zhou, J.; Yoshizawa, J.M.; Dong, L.; Li, X.; et al. Coupling high-throughput genetics with phylogenetic information reveals an epistatic interaction on the influenza A virus M segment. *BMC Genom.* **2016**, *17*, 46. [[CrossRef](#)]
141. Hilsch, M.; Goldenbogen, B.; Sieben, C.; Höfer, C.T.; Rabe, J.P.; Klipp, E.; Herrmann, A.; Chiantia, S. Influenza A Matrix Protein M1 Multimerizes upon Binding to Lipid Membranes. *Biophys. J.* **2014**, *107*, 912–923. [[CrossRef](#)]
142. Liu, H.; Grantham, M.L.; Pekosz, A. Mutations in the Influenza A Virus M1 Protein Enhance Virus Budding To Complement Lethal Mutations in the M2 Cytoplasmic Tail. *J. Virol.* **2017**, *92*, e00858-17. [[CrossRef](#)]
143. Fares, M.A.; McNally, D. CAPS: Coevolution analysis using protein sequences. *Bioinformatics* **2006**, *22*, 2821–2822. [[CrossRef](#)]

144. Starr, T.N.; Greaney, A.J.; Hilton, S.K.; Ellis, D.; Crawford, K.H.; Dingens, A.S.; Navarro, M.J.; Bowen, J.E.; Tortorici, M.A.; Walls, A.C.; et al. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **2020**, *182*, 1295–1310.e20. [[CrossRef](#)] [[PubMed](#)]
145. Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367*, 1444–1448. [[CrossRef](#)] [[PubMed](#)]
146. Liu, H.; Zhang, Q.; Wei, P.; Chen, Z.; Aviszus, K.; Yang, J.; Downing, W.; Jiang, C.; Liang, B.; Reynoso, L.; et al. The basis of a more contagious 501Y.V1 variant of SARS-CoV-2. *Cell Res.* **2021**. [[CrossRef](#)] [[PubMed](#)]
147. Greaney, A.J.; Starr, T.N.; Gilchuk, P.; Zost, S.J.; Binshtein, E.; Loes, A.N.; Hilton, S.K.; Huddleston, J.; Eguia, R.; Crawford, K.H.D.; et al. Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *bioRxiv* **2020**. [[CrossRef](#)]
148. Casadevall, A.; Imperiale, M.J. Risks and Benefits of Gain-of-Function Experiments with Pathogens of Pandemic Potential, Such as Influenza Virus: A Call for a Science-Based Discussion. *mBio* **2014**, *5*, e01730-14. [[CrossRef](#)]
149. Imperiale, M.J.; Casadevall, A. Rethinking Gain-of-Function Experiments in the Context of the COVID-19 Pandemic. *mBio* **2020**, *11*. [[CrossRef](#)]