

# Deep mutational scanning and CRISPR-engineered viruses: tools for evolutionary and functional genomics studies

Mercedes Paz,<sup>1,2</sup> Gonzalo Moratorio<sup>1,2,3</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 11.

**ABSTRACT** Recent advancements in synthetic biology and sequencing technologies have revolutionized the ability to manipulate viral genomes with unparalleled precision. This review focuses on two powerful methodologies: deep mutational scanning and CRISPR-based genome editing, that enable comprehensive mutagenesis and detailed functional characterization of viral proteins. These approaches have significantly deepened our understanding of the molecular determinants driving viral evolution and adaptation. Furthermore, we discuss how these advances provide transformative insights for future vaccine development and therapeutic strategies.

**KEYWORDS** deep mutational scanning, CRISPR-engineered viruses, virus evolution

In our 2020 mSphere of Influence article (1), we unfold groundbreaking work in synthetic biology that revolutionized how we study viral genomes, introducing an approach to decode how genetic information drives function, fitness, and evolutionary trajectories. Among these, the seminal study by Coleman et al. (2) demonstrated how engineering codon pair bias in poliovirus genes created a synthetic virus with unchanged protein composition but altered nucleotide sequences. By incorporating host-unfavorable codon pairs, this virus showed reduced replication efficiency and attenuated virulence, while retaining robust immunogenicity, a milestone in designing rationally attenuated viruses.

Building on this foundation, subsequent studies expanded the scope (3–6), applying recoding strategies across diverse viral families. These efforts revealed the intricate links between genome architecture and viral fitness, paving the way for innovative approaches to studying viral evolution and developing vaccine candidates. More recently, researchers have reframed these concepts to engineer RNA viruses using genome-wide recoding strategies aimed at promoting the accumulation of the most detrimental mutations, stop codons, thereby steering viral progenies toward evolutionary “dead ends.” This approach, tested across a spectrum of RNA viruses, including human enteroviruses, influenza, chikungunya, and coronaviruses (7–9), highlights the potential of this strategy to develop live attenuated vaccines and investigate the viral capacity to buffer mutations, known as genetic robustness.

Technological advancements in computational and experimental virology have further catalyzed this progress. Cutting-edge tools now enable precise manipulation of viral genomes, encompassing genome-wide mutagenesis and targeted insertions or deletions to dissect the functional roles of proteins and sequence motifs. This Full Circle review builds upon the existing body of work on viral fitness, focusing specifically on protein-focused methods and their application in deep mutational scanning (DMS) and CRISPR-based engineered viruses. These methodologies are pivotal for dissecting viral mechanisms, developing novel therapeutics, and understanding viral evolution. Here, we highlight recent advances in these two transformative techniques, which represent the next chapter in our understanding of viral fitness. Taken together, these techniques

**Editor** Michael J. Imperiale, University of Michigan, Ann Arbor, Michigan, USA

Address correspondence to Gonzalo Moratorio, moratorio@pasteur.edu.uy.

The authors declare no conflict of interest.

See the companion article at <https://doi.org/10.1128/msphere.00179-20>.

**Published** 24 April 2025

Copyright © 2025 Paz and Moratorio. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

integrate viral genomes with protein function, showcasing the power of these tools in engineering viral genomes. The review will explore the principles behind DMS and CRISPR-based genome editing, discuss their applications in studying viral fitness, and highlight key findings from recent studies. We will also examine the challenges and future directions of these techniques in the context of viral research.

## DMS IN RNA VIRUSES

### Overview and applications

DMS is a powerful high-throughput technique that systematically investigates how genetic variation translates into phenotypic variation. It enables the creation of fitness maps that detail the effects of nearly all possible nucleotide substitutions within a viral genome or protein sequence. This approach has helped to uncover key molecular mechanisms underlying viral evolution, the roles of non-coding viral sequences, protein function, and resistance to antiviral therapies, among many other applications as summarized in Fig. 1.

Methodologically, a DMS experiment follows at least five key steps. It starts with the design and construction of a large library of sequence variants, which defines the target sequence under study. This sequence can range from an entire viral protein to a particular protein domain, such as active sites or regions involved in important molecular interactions, or even a full viral proteome. The mutational scope is also defined at this stage, including single-nucleotide or amino acid substitutions covering all 19 non-native residues per position, and codon insertions and deletions (known as indels). To generate the mutant libraries, various strategies have been employed. Error-prone PCR has been widely used due to its cost-effectiveness and scalability in high-throughput experiments for introducing random nucleotide substitutions (10–13). However, it is less suitable for experiments in which specific mutations are required and can also introduce biases in the mutation frequency depending on the polymerase used in the PCR (14). Over the past two decades, rapid advances in *de novo* DNA and RNA synthesis have made controlled mutagenesis approaches more accessible, for example, the use of a pool of oligos containing the desired mutations, ensuring precise amino acid or nucleotide substitutions (15, 16) or indels (17).

Following library construction, the mutant proteins must be expressed in a relevant experimental system which is chosen based on the biological question. Infectious virus systems, for example, are used to assess viral fitness by rescuing mutant viral genomes into live viruses and measuring replication efficiency in cells permissive to infections (18–21). Library screening tools used for high-throughput selection of desired traits based on the display of proteins or peptides on the cell surface, like yeast (22–24) or mammalian (16, 25, 26) display systems, and more recently lentiviral or vesicular stomatitis virus (VSV) based pseudoviruses, have been highly beneficial for studying envelope proteins. For example, it was used in studies of the spike (S) protein of SARS-CoV-2 and the hemagglutinin (HA) protein of H5 influenza, to characterize the effects of mutations in phenotypes critical to viral fitness, like viral cell entry, antibody neutralization, and receptor-binding assays (27, 28).

Once expressed, mutants must be filtered to identify beneficial, neutral, and deleterious mutations. For example, in experiments using live viruses, assessing the replicative fitness of mutants after a certain time post-infection or after several rounds of serial passages has been used for evaluating amino acid substitutions in the RNA-dependent RNA polymerase (RdRp) of influenza A virus (IAV) and dengue virus (DENV) (18, 19). Binding affinity assays evaluate how mutations in receptor-binding domains (RBDs) affect interactions with host receptors; for instance, SARS-CoV-2 RBD mutant S proteins are tested for binding strength to human angiotensin-converting enzyme 2 (ACE2) using labeled sorting methods (22, 29). Additionally, neutralization and immune-escape selection experiments can test how mutant envelope proteins influence viral evasion from antibodies, revealing key antigenic sites (23, 29, 30).

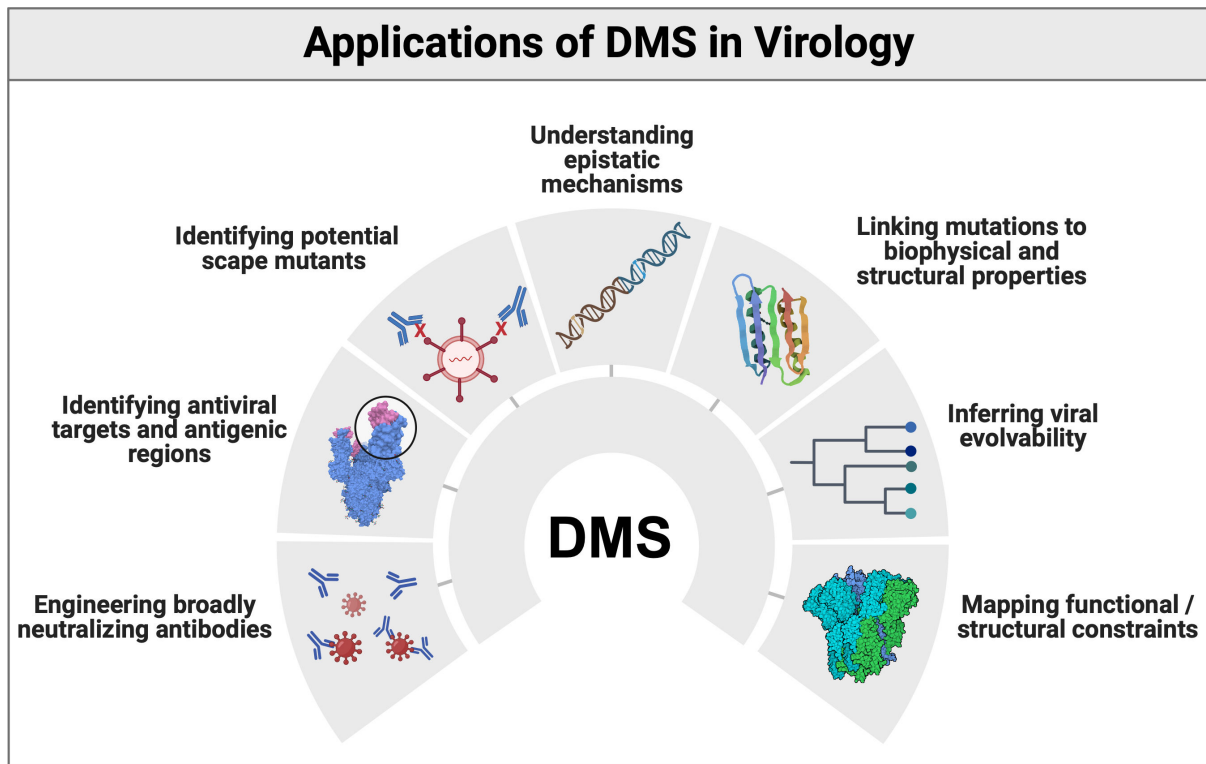


FIG 1 Summary of the main applications of deep mutational scanning in virology. Created in <https://BioRender.com>.

Following selection, next-generation sequencing is employed to quantify mutation frequencies before and after functional screening. DNA or RNA from the selected mutant populations is extracted and sequenced using high-throughput platforms. Computational analyses are then used to calculate fitness scores for each variant, categorizing them as deleterious, neutral, or beneficial based on their enrichment or depletion in the selected population (31).

The final step involves interpreting the mutational effects within a structural and evolutionary context. Mapping these changes onto known protein structures using molecular modeling, cryo-electron microscopy, or computational tools like AlphaFold helps predict how mutations alter protein stability and function (15, 16, 32).

Seminal papers published on DMS studies around 2010, originally focused on understanding protein sequence-function relationships. For example, the work by Fowler et al. (33) focused on WW domains, stable, triple-stranded beta-sheet motifs found in signaling and structural proteins that are characterized by their ability to bind proline-rich sequences and evaluated the ligand-binding performance of over 600,000 sequence variants in a single experiment. In another foundational study by Hietpas et al. (34), they developed an approach to systematically generate mutant libraries and measure fitness in a high-throughput manner. Precisely, they mapped the fitness effects of point mutations on a nine amino acid (AA) region of the yeast Hsp90 protein, a chaperone that mediates, directly or indirectly, the folding and activation of 10–15% of all proteins in yeast and humans. Another example of the applicability of this technique is the work of Starita et al. (35), where they used DMS to create a sequence-function map of the U-box domain of the murine E3 ligase Ube4b, characterizing mutations that enhance ubiquitin transfer activity. These works provided the first insights into the distribution and evolutionary weight of different types of mutations, established a tool for identifying mutations that enhance enzymatic activity for biotechnology applications, and demonstrated the potential of DMS in precision medicine, by elucidating the influence of human genome variations on disease susceptibility. As discussed in the following sections, DMS has been used extensively in virology to study viral protein function and

immune evasion mechanisms, to contribute to the development of vaccines and antiviral strategies, and to study viral evolution.

### How does DMS shape our understanding of viral proteins?

DMS has greatly been used to map fitness landscapes and constraints on RNA viral proteins. It has helped to identify which AA sites are essential for molecular interactions or enzymatic activity and which ones are more tolerant to mutation, providing a detailed picture of protein function (10, 18, 19, 36–38). Historically, this analysis has been addressed by exploring the grade of sequence conservation of available deposited sequences (39–41). Although sequencing viral genomes from naturally isolated strains has become technically straightforward, only a limited number of viruses have sufficient sequence data available for comprehensive analysis. Furthermore, this approach has been proven to obscure critical residues within highly conserved regions (10, 42). Coupled with corroborative approaches combining experimental validation and computational modeling, DMS has advanced in understanding conserved functional and structural areas, as well as epistatic mechanisms in viral proteins.

Multiple studies have applied DMS to examine viral polymerases and structural proteins. For example, different groups have focused on uncovering functional constraints on the IAV polymerase subunits. Wu et al. (10) combined experimental fitness profiling across 94% of the polymerase A (PA) AA sequence with computational protein structure stability predictions, revealing high inter-type diversity, as numerous functional residues were non-conserved among each influenza virus (A or B) type. Additionally, this integrative methodology successfully distinguished residues critical for viral replication from those essential for protein stability, outperforming conventional approaches like dN/dS analysis (nonsynonymous/synonymous substitution ratio, a method in evolutionary biology to assess selection pressure) and highlighting the limitations of relying solely on sequence conservation or phylogenetic analyses.

Li et al. (19) used DMS to study the mutational landscape of the polymerase basic 1 (PB1) protein. They generated a library covering more than 90% of all possible AA substitutions and quantified their effects on replicative fitness, measured by a viral reconstitution system in mammalian cells. The study revealed that functional and structural constraints primarily mapped to residues involved in molecular interactions with RNA or proteins rather than broader conserved subdomains. In addition to capturing mutational effects on fitness, pairwise competition assays experimentally validated more than 10 mutations, highlighting regions critical for stability and polymerase function. This comprehensive approach demonstrated how DMS can delineate essential regions of RNA virus polymerases while correlating mutational tolerance with evolutionary diversity across natural influenza populations. Similarly, Suphatrakul et al. (18) conducted a DMS study of the nonstructural protein 5 (NS5) protein of DENV, focusing on its roles as an RNA-dependent RNA polymerase and methyltransferase. They performed a saturation mutagenesis experiment targeting nearly all single AA substitutions and assessed their effects on viral replication in mammalian cells. The resulting fitness maps demonstrated high mutational constraint in the conserved RdRp active site but unexpectedly revealed tolerance for select substitutions in universally conserved residues lining the priming loop. These findings suggest potential alternative roles for these residues in regulating polymerase activity. Further integration with structural and biochemical data highlighted the enzymatic versatility of NS5, as certain constrained positions influenced both RNA synthesis and host immune antagonism. This study underscored the importance of combining DMS data with structural modeling to identify residues critical for viral replication and therapeutic targeting. Moreover, recent research on the hepatitis B virus (HBV) polymerase unveiled the mechanism behind cis-preferential reverse transcription (38). They identified conserved prolines near the polymerase's termination codon that stall ribosomes, effectively tethering the nascent polymerase to its template RNA. Conse-

quently, this interaction ensures cis-preferential RNA packaging and facilitates reverse transcription of the HBV genome.

Structural proteins also play a crucial role in many aspects of viral fitness. They dictate the assembly, infectivity, host-tropism, and survival of viruses. DMS studies on the HA and neuraminidase (NA) proteins of IAV have revealed a clear distinction between antigenic and functional/structural regions (32, 43, 44). Antigenic regions, such as the HA globular head or surface-exposed regions of NA, show high mutational tolerance, enabling immune escape by allowing the virus to evade host immune responses. In contrast, functional domains, including the HA and NA catalytic sites, exhibit low mutational tolerance due to their critical roles in maintaining essential viral functions such as receptor binding and enzymatic activity. These findings highlight the evolutionary constraints imposed by the need to balance immune evasion with functional integrity.

The M1 matrix protein is the most conserved in IAV, leading to the historical assumption that it has a low tolerance for mutations. However, a recent study demonstrated that mutational tolerance varies across the protein, despite the low sequence diversity observed in natural isolates (42). The N-terminal third of M1 exhibited low mutational tolerance, likely due to its critical role in oligomerization and interactions with other viral components. In contrast, the C-terminal two-thirds were more tolerant of AA variations, which structural analysis suggests may result from its disordered configuration, reducing constraints on AA preferences. During the COVID-19 pandemic, multiple studies focused on performing DMS to explore the mutational fitness landscape of the SARS-CoV-2 S protein, shedding light on structural and functional constraints (16, 22, 45, 46). For example, Starr et al. (22) conducted a comprehensive DMS of the RBD of the S protein, measuring the effects of all possible AA substitutions on folding stability and ACE2 binding affinity, which is the main entry point into cells for some coronaviruses including SARS-CoV-2. Their findings showed that whereas most mutations reduced binding or compromised protein folding, some substitutions improved ACE2 binding, most notably, aligning with changes observed in circulating variants. Similarly, Ouyang et al. (16) examined the spike N-terminal domain (NTD) and quantified the impact of thousands of single mutations on protein expression levels, revealing that mutational tolerance was inversely correlated with proximity to interdomain interfaces. Their findings highlighted mutations that enhanced expression without altering antigenicity, offering insights for immunogen design.

### How does DMS advance our understanding of viral evolution?

Over the past decade, DMS has also been instrumental in experimental evolution studies, identifying mutations linked to adaptation and shedding light on evolutionary mechanisms.

Several studies demonstrate the importance of epistasis, which is defined as the interactions between mutations in which one mutation's effect depends on others' presence. For example, this was done for studying SARS-CoV-2 evolution (29, 47–53), where compensatory mutations (e.g., at the interaction interface of the S protein RBD with the NTD) were demonstrated to reconfigure viral fitness. Taylor and Starr (48) showed that mutations at the RBD of the Omicron SARS-CoV-2 variants BQ.1.1 and XBB.1.5 can induce substantial epistatic effects, where earlier substitutions like N501Y reshaped the mutational fitness landscape, making previously deleterious mutations either neutral or beneficial. Using an infectious system of the whole S protein, Tamishii et al. (51) also revealed epistasis between RBD and NTD mutations, demonstrating compensatory relationships that stabilize overall S protein fitness while enabling immune evasion, such as structural compensation in the NTD for destabilizing RBD mutations observed in Omicron variants BA.1 and BA.2. Recently, epistatic interactions were observed in the Omicron BA.2.86 spike variant, where specific combinations of mutations (e.g., Q493E coupled with L455S and F456L) enhanced ACE2 binding, reversing the deleterious effect of Q493E in previous variants (49). These studies

consistently emphasize that epistasis plays an important role in modulating mutational constraints, either relaxing or tightening them based on the genetic and structural context, enabling viruses to navigate evolutionary trade-offs. Moreover, these findings also underscore the importance of investigating epistatic networks across RNA virus families to uncover conserved principles of viral evolution.

DMS has also been applied to investigate how changes in RNA viral proteins influence host tropism, the adaptation of a virus to infect different host species. The work by Thyagarajan and Bloom on the IAV HA protein (32) identified mutations that enable shifts in receptor-binding specificity, highlighting antigenic sites that enhance binding to human-like  $\alpha$ -2-6-linked sialic acids, a critical determinant for allowing IAV to infect humans. More recently, a comprehensive proteome-wide DMS analysis of coxsackievirus B3 evaluated how over 40,000 nonsynonymous mutations and almost all possible codon deletions in the nonstructural proteins affected its fitness (15). The study revealed considerable variation in mutation tolerance both within and across individual proteins, which is associated with their structural and functional roles. It was found that while many viral proteins face strict functional constraints, nonstructural proteins especially display unique mutational tolerance patterns that may relate to mechanisms of adaptation specific to the host. In another proteome-wide DMS analysis of enterovirus A71, the virus was modified with around 51,000 indels and 41,000 AA substitutions. Studies experimentally examining the effects of indels on viral fitness are scarce. Still, most have indicated that indels are lethal to the virus and can only be tolerated at certain hotspots, which coincide with regions crucial for host recognition and immune interaction. Conversely, the viral proteome shows a higher tolerance for AA substitutions compared to indels, underscoring the evolutionary pressures exerted by this type of mutation and their possible influence on phenotypic diversity (17).

Regarding arboviruses, their dual-host cycle, which involves replication in both invertebrate and vertebrate hosts, imposes a strong purifying selection on the viral genomic determinants required for replication in evolutionarily distant hosts (54). This selective pressure makes it challenging to resolve which determinants are optimal for viral fitness in each host. In a work by Setoh et al. on Zika virus (ZIKV) (55), this was addressed by coupling DMS with fitness measurements across cell types from different organisms (e.g., mosquito cells vs human cells). They characterized conserved residues in the envelope protein's RBD, two substitutions in the ZIKV E protein, K316Q and S461G, that combined, strongly favored virus replication in mosquito cells. The mutant virus 316Q/461G containing these substitutions was found to be highly attenuated in mammalian cells, organoids, and mice, demonstrating host-specific trade-offs. Similarly, constrained mutational tolerance in critical catalytic sites of nonstructural proteins often links structural stability to host-specific replication efficiency (18). These studies shed light on evolutionary selection, where host-specific evolutionary pressures, whether structural, biochemical, or immune-related, tend to exploit residues or domains that balance adaptability with the preservation of essential functions. They also uncover conserved mutational constraints and hotspots that govern cross-species adaptability in RNA viruses.

### How can DMS identify therapeutic targets?

As a powerful tool for systematically assessing the impact of mutations across viral proteins, several DMS studies have contributed insights that could be relevant to the development of different therapeutic approaches. For instance, mapping viral protein regions related to immune evasion and receptor binding may be possible to enhance the development of neutralizing antibodies (56–58). Moreover, highly conserved regions in viral proteins could be ideal targets for therapeutic and preventive strategies against RNA viruses. For example, these regions have the potential to be used for the development of broadly neutralizing antibodies that focus on conserved epitopes in envelope glycoproteins, to ensure cross-strain efficacy, such as those reported to target a discrete membrane-proximal anchor epitope of the HA stalk domain in IAV (59). In addition,

antivirals targeting critical enzymatic or functional sites in these conserved regions found by DMS experiments could have the potential to effectively inhibit viral replication, offering a promising alternative to the limited number of effective vaccines available to combat diverse RNA viruses (44).

A promising alternative approach to monoclonal antibodies and antivirals, which has capitalized on DMS studies identifying mutational rigid and conserved receptor-binding regions, involves the designing of soluble decoy receptors that mimic the virus-binding regions of host receptors. In a series of studies, a research group used deep mutagenesis of the ACE2 receptor to first engineer a decoy receptor sACE2.v2.4, with enhanced affinity for the SARS-CoV-2 S protein (60). The decoy receptor, containing three mutations (T27Y, L79T, and N330Y), showed a 35-fold increase in binding affinity and neutralized SARS-CoV-2 at levels comparable to high-affinity monoclonal antibodies. In the first study, the engineered receptor was demonstrated to be highly soluble, stable, dimeric, and catalytically active, with broad neutralization capabilities against both SARS-CoV-1 and SARS-CoV-2. In a second study (61), the researchers explored the potential for viral resistance to sACE2.v2.4 and found that no mutations within the RBD of SARS-CoV-2 redirected specificity toward the wild-type ACE2 receptor, suggesting that resistance would be rare. In another collaborative work, the *in vivo* efficacy of a derivative of the sACE2.v2.4 decoy fused with an IgG1 constant region (sACE2.v2.4-IgG1) was evaluated in K18-hACE2 mice (62). This work proved that sACE2.v2.4-IgG1 exhibited tight and persistent binding to SARS-CoV-2 and several variants of concern, and notably, that the decoy significantly reduced viral entry, lung vascular hyperpermeability, acute respiratory distress syndrome, and mortality in mice infected with both the WA-1/2020 and P.1 VOCs. These works highlight the promising therapeutic potential of an engineered decoy receptor *in vivo*. However, its clinical application is still being studied due to the possible short serum half-life, which may compromise its practical use. A recent study on the engineered decoy receptor sACE2.v2.4-IgG1 points out that sialylation is a vital factor in enhancing the *in vivo* efficacy of sACE2 decoy receptors and stresses the importance of glycosylation modifications to improve their pharmacokinetics (63).

## Virus-driven crispr screens to identify host factors modulating viral replication

### Overview

The pioneering work of Mojica et al. (64) and Barrangou et al. (65) demonstrated CRISPR-Cas9 as the adaptive immune mechanism of bacteria and archaea to fight foreign genetic material such as plasmids and phages. In prokaryotes, this system specifically recognizes and binds DNA through CRISPR RNA (crRNA), guiding Cas proteins to identify and cleave exogenous DNA with the assistance of trans-activating CRISPR RNA (tracrRNA). Subsequently, the groups of Charpentier and Doudna (66) combined crRNA and tracrRNA into a single-guide RNA (sgRNA), which more efficiently facilitated Cas9's genome-editing role *in vitro*. This groundbreaking discovery, which earned them the Nobel Prize in Chemistry in 2020, transformed our ability to investigate gene functions. In virology, the CRISPR-Cas9 system has widely been used to study virus-host interactions through whole-genome screening, using vector-based sgRNA libraries targeting the whole coding host genome, identifying host factors essential for virus replication (67–72).

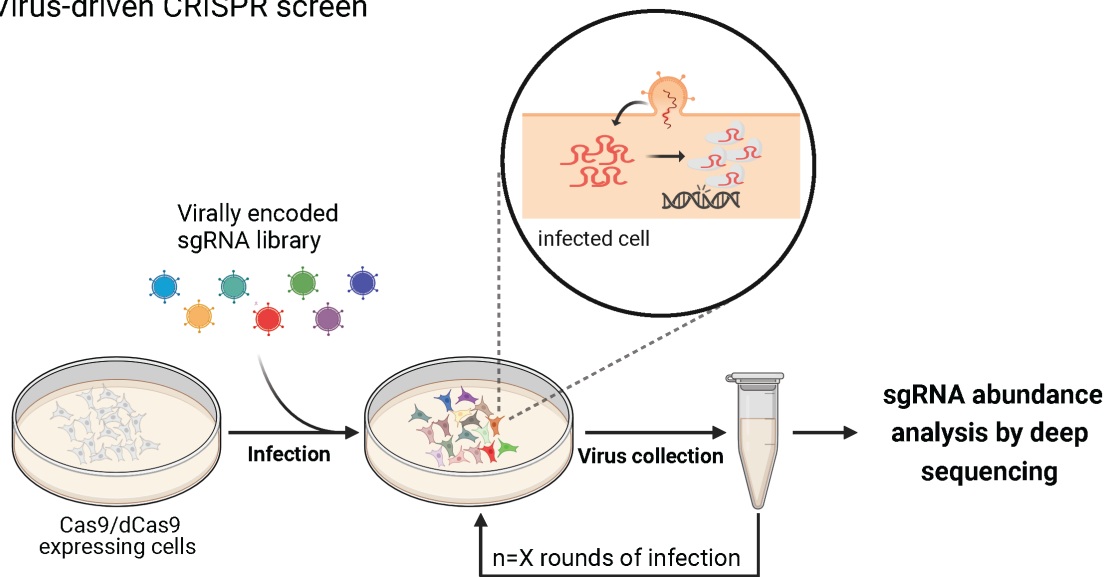
Until just a couple of years ago, most multiplexed pooled CRISPR screens in the context of infection were based on cell viability as their readout, leading to the discovery of cellular genes that allow cells to resist infection. Although extremely powerful, cell viability approaches predominantly reveal hits that block the early steps of viral replication, specifically entry, limiting the breadth of captured host factors. As reviewed in the following section, various groups have developed a novel screening system based on CRISPR-engineered replicating viral systems to expand the breadth of discoverable functional factors and increase sensitivity. In this platform, sgRNA libraries are expressed directly from the viral genome, enabling the direct follow-up of virus

propagation throughout the infection cycle. Upon infection of a cell expressing the Cas9 protein, the sgRNA encoded by the virus is released and alters the expression of cellular genes. This gene alteration affects the replication kinetics of the virus, where the sgRNA levels serve as a direct readout for viral propagation (Fig. 2A). In this way, sgRNAs that target genes that are essential or restrict virus propagation are expected to be underrepresented or enriched, respectively, in viruses budding from Cas9-expressing cells. Additionally, since thousands of viral genome copies are synthesized during each replication cycle, the signal is naturally amplified, making this screening approach highly sensitive and, potentially enabling the detection of even small positive and negative effects on viral propagation. This methodology surpasses conventional CRISPR screens by offering greater specificity, resolution, and mostly, the ability to profile later stages of the infection cycle, paving the way for novel antiviral strategies.

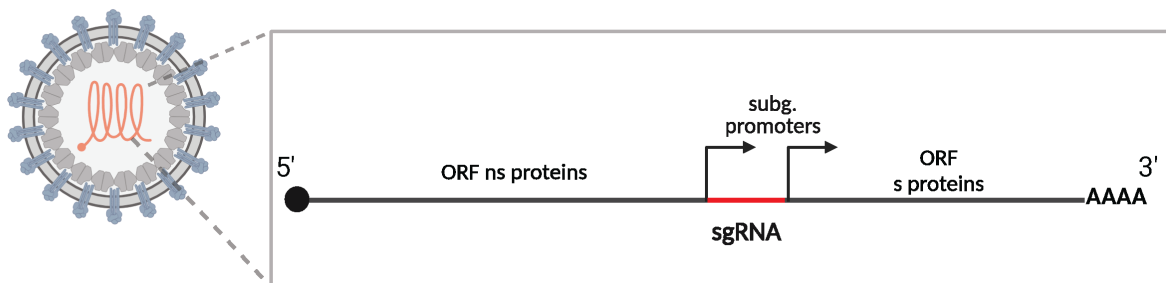
**How can viruses be engineered to conduct CRISPR screens?**

This novel approach was first applied in a study to identify interferon-stimulated genes against HIV (73). In this study, the researchers used lentiviral genomes encoding sgRNAs packaged into HIV virions, enabling the delivery and expression of Cas9 and sgRNA in target cells. Knockout cells were pooled, and deep sequencing of CRISPR RNAs in newly budded virions was used to assess HIV replication efficiency across thousands of genes in a heterogeneous population. The study identified antiviral factors like MxB, TRIM5α, IFITM1, and Tetherin, with sensitivity varying by viral strain. Additionally,

**A. Virus-driven CRISPR screen**



**B. Modified alphavirus genome encoding sgRNAs**



**FIG 2** Virus-driven CRISPR screens. (A) Overview of a virus-driven CRISPR screen. (B) Potential strategy for the development of alphaviruses encoding sgRNAs. Open reading frame (ORF), nonstructural (ns) proteins, subgenomic promoters (subg. promoters), structural (S) proteins. Created in <https://BioRender.com>.

host dependency factors such as CD169, SEC62, and TLR2 were shown to support HIV replication in THP-1 cells. In another study on HIV (74), the researchers employed a library of over 1,500 replication-competent HIV-1 constructs, each expressing a single sgRNA targeting more than 500 cellular genes, to identify antiviral restriction factors across the entire HIV-1 replication cycle. Passaging in CD4<sup>+</sup>T cells enriched sgRNAs targeting several key genes, including GRN, CIITA, EHMT2, CEACAM3, CC2D1B, and RHOA, by over 50-fold. Additionally, the technique identified IFI16 as a restriction factor specifically targeted by the HIV-1 accessory gene *nef*. Functional validation in cell lines and primary CD4<sup>+</sup>T cells confirmed that these factors impact various stages of the HIV life cycle, including entry, transcription, release, and infectivity.

Recently, Finkel et al. (75) introduced virus-encoded CRISPR-based direct readout screening (VECOS), a system leveraging human cytomegalovirus (HCMV) engineered to express sgRNAs targeting the human coding genome. This was achieved by cloning the HCMV genome into a bacterial artificial chromosome system coupled with a Gateway cloning system to streamline the process of modifying the viral genome to encode a robust sgRNA library. Primary human foreskin fibroblasts (HFFs) were employed as the experimental model due to their permissiveness to viral replication. Sequential infections were performed by harvesting viral supernatants from infected HFFs and transferring them to fresh cultures to propagate the virus over multiple passages. To quantify changes in sgRNA abundance across passages and identify host factors critical for distinct stages of the HCMV infection cycle, the researchers extracted viral DNA after each sequential passage from cell lysates and culture supernatants. The abundance of sgRNAs encoded within the viral genome was then analyzed by PCR and high-throughput sequencing. This experimental approach demonstrated that RRM2, a key enzyme in dNTP production, is a critical host factor for HCMV replication, as its depletion severely impaired viral propagation and DNA synthesis. This finding opens up a new therapeutic potential for combining RRM2 inhibitors with existing anti-HCMV treatments, such as ganciclovir, to enhance antiviral efficacy. Importantly, the study demonstrated the unique strengths of VECOS in identifying a broader range of host factors, including those critical for non-lethal stages of infection, which are often missed by traditional survival-based CRISPR screens. They also provide evidence of the poorly understood functions of ARL6IP6 and its importance in viral propagation, demonstrating that ARL6IP6 plays a critical role in the late stages of the HCMV infection cycle, likely contributing to virus assembly through its association with ER-derived structures.

Similarly to the group above, King et al. (76) developed a novel fitness-based screen named transcriptional regulation by pathogen-programmed Cas9 (TRPPC), which uses IAV to deliver single guide RNAs (sgRNAs) that program CRISPR activation or CRISPR inhibition of cellular genes exploiting the virus's ability to express non-coding RNAs. They modified the influenza virus nonstructural (NS) gene, which normally encodes two proteins, nonstructural protein 1 (NS1) and the nuclear export protein, by altering its overlapping reading frames into non-overlapping ones, separated by an artificial intron. This created a "split NS" system that allowed the expression of non-coding RNAs without disrupting viral protein production or replication. They incorporated a microRNA (miRNA) encoded within this artificial intron, which would be processed by the host's Drosha nuclease in the nucleus, along with a sgRNA downstream of the miRNA-124, releasing the sgRNA from the viral genome. After several rounds of selection of the virus-encoded sgRNA library in A549 cells, the screen identified the protein TREX1, a cytoplasmic DNA exonuclease, as a key pro-viral factor, as it degrades cytoplasmic DNA to prevent immune activation, thus aiding influenza replication. This methodology also allowed the creation of a ranking of host factors that affect IAV viral replication, providing valuable insights into host-pathogen interactions.

## Challenges and future directions

Despite their potential, the mentioned methodologies face challenges that create opportunities for refinement to address research gaps and limitations. A technique like

DMS produces a big volume of data, and the main challenge remains in analyzing and making sense of it to predict the effects of mutations. Machine learning offers a powerful approach for these methods that require the analysis of vast amounts of data and exhaustive experimental testing. For example, automated viral infectivity assays utilizing convolutional neural networks have been used to predict infection phenotypes from microscopy data, improving speed and precision in virus characterization, which could be applied for a more efficient analysis of viral variants (77–79). Moreover, machine learning has also been used to predict the functional outcomes of millions of untested mutations from DMS data sets, helping researchers to identify and filter promising variants for further study and reduce the need for exhaustive experimental testing (80–82). This approach could fill gaps in DMS by enabling real-time predictions of mutation effects in complex protein landscapes and expand the ability to explore larger mutational spaces with high accuracy and fewer laboratory resources. Another major challenge, shared by both CRISPR screening and DMS, is that they rely on cell culture as their experimental system, which can influence the interpretation of mutational effects. While these systems are used primarily for their ability to facilitate high-volume infections and vigorous viral replication, they may not fully capture the complexities of *in vivo* scenarios, where factors such as tissue-specific interactions and immune responses may play a significant role. For instance, research teams employing DMS on the same viral protein have reported inconsistent results across diverse cellular environments (42, 83). The authors discuss that discrepancies can largely be attributed to differences in viral strains, cell types, and experimental methodologies, but cannot rule out factors like population bottlenecks and epistatic interactions (42). Therefore, any conclusions drawn should be contextualized within the specifics of the viral strain and experimental setting.

Regarding DMS, although several studies have identified conserved functional regions in RNA virus proteins, such as the flavivirus NS5 or influenza PB1 polymerases (18, 19), systematic cross-family comparisons of mutational constraints are scarce. Current research mostly focuses on specific viral targets or homology-based analyses, lacking comprehensive data sets that consolidate DMS mappings across diverse RNA virus families to identify conserved functional or structural “hotspots.” This gap may largely arise from inconsistencies in experimental models and the lack of standardized fitness metrics across studies. Future research could expand DMS efforts to homologous proteins across major RNA viral families, targeting universal antiviral sites such as RdRp enzymatic cores and conserved glycoprotein domains. Such integration could inform the design of broad-spectrum inhibitors targeting residues less prone to mutational escape, paving the way for cross-family antiviral strategies, like the development of pan-viral vaccines to combat diverse RNA viruses. To support this idea, initiatives such as MaveDB (84), an open-source repository for sharing and analyzing DMS data sets, and the Atlas of Variant Effects Alliance (<https://www.varianteffect.org/>) have been established to foster collaboration and enhance the utility of mutational scanning research (85).

As part of this section discussing future directions, we hypothesize that alphaviruses, with their positive-sense, single-stranded RNA genomes that replicate in the cytoplasm, offer a distinct opportunity for exploration as a possible advancement in CRISPR/Cas9 technology. These viruses exhibit unique characteristics compared to previously employed viral systems. Specifically, the potential development of replication-competent Mayaro virus constructs encoding sgRNAs represents an innovative approach worth investigating. This idea is grounded in the capacity that alphaviruses have to tolerate the insertion of exogenous coding and non-coding sequences downstream of their nonstructural genes, expressed via a subgenomic promoter (Fig. 2B). To facilitate the release of non-coding RNAs like sgRNAs during replication, these sequences can be positioned downstream of a duplicated subgenomic promoter, followed by a self-excising ribozyme (e.g., HDV) or a miRNA sequence, as demonstrated in prior studies (76, 86). Nevertheless, RNA viruses face additional complexities, including high mutation rates and genomic instability, which increase the risk of evolutionary escape and may compromise embedded sgRNA functionality. Optimizing experimental designs and

methodologies will be essential to make CRISPR-based tools broadly applicable across multiple RNA viral families. In addition, given the recent development of this technology, establishing robust reproducibility standards is critical to ensure consistent and reliable outcomes across viral strains. Variability in experimental protocols, such as delivery methods, cell lines, and infection conditions, affects sgRNA targeting efficiency and the functional relevance of identified host factors. Addressing these issues will require standardized sgRNA libraries and multi-laboratory validation studies to benchmark findings and enhance reproducibility.

## CONCLUDING REMARKS

In this Full Circle review, we highlight how DMS and CRISPR-engineered viral systems serve as essential tools in virology, enabling researchers to map viral fitness landscapes, identify drivers of evolution, and explore host-virus dynamics. DMS has revolutionized the study of viral proteins, allowing us to create fitness maps that detail the effects of nearly all possible nucleotide substitutions within a viral genome. Combining experimental techniques with computational tools such as modeling and machine learning has provided deeper insights into mutations, protein functions, and epistatic interactions to unveil viral evolution. Global collaboration is crucial, with open-access repositories like MaveDB enabling data sharing to address broader challenges like assessing mutational landscapes across viral families or identifying pan-antiviral targets. The newly developed strategy, virus-driven CRISPR screenings, focuses on the role of host factors in viral replication, uncovering novel host-viral mechanisms of interaction and revealing new therapeutic targets. It represents a significant advancement over earlier conventional CRISPR screening, addressing limitations such as profiling later stages of the infection cycle. As this technology develops, new experimental designs and methods will broaden the application across multiple viruses, revealing new insights into viral-host dynamics and offering promising leads for antiviral development.

Looking ahead, the convergence of high-throughput technologies, artificial intelligence, and viral genome engineering will unlock an era of unprecedented precision in virology. This multidisciplinary approach will not only accelerate our understanding of viral biology but will also lead to personalized, adaptive treatment regimens, representing a new frontier in the fight against viral infections.

## ACKNOWLEDGMENTS

The authors thank Dr. Pilar Moreno for her thoughtful review of the manuscript and for providing invaluable feedback.

This work was supported by FOCEM COF 03/11, the G4 Program, the Institut Pasteur de Montevideo, and the Agencia Nacional de Investigación e Innovación.

## AUTHOR AFFILIATIONS

<sup>1</sup>Laboratory of Experimental Virus Evolution, Institut Pasteur de Montevideo, Montevideo, Uruguay

<sup>2</sup>Molecular Virology Laboratory, Faculty of Sciences, University of the Republic, Montevideo, Uruguay

<sup>3</sup>Center for Innovation in Epidemiological Surveillance, Institut Pasteur de Montevideo, Montevideo, Uruguay

## AUTHOR ORCIDs

Mercedes Paz  <http://orcid.org/0000-0002-9700-0327>

Gonzalo Moratorio  <http://orcid.org/0000-0002-4555-4926>

## AUTHOR CONTRIBUTIONS

Mercedes Paz, Writing – original draft, Writing – review and editing | Gonzalo Moratorio, Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review and editing

## REFERENCES

- Moratorio G. 2020. mSphere of influence: experimental evolution of RNA viruses. *mSphere* 5:e00179-20. <https://doi.org/10.1128/mSphere.00179-20>
- Coleman JR, Papamichail D, Skiena S, Futcher B, Wimmer E, Mueller S. 2008. Virus attenuation by genome-scale changes in codon pair bias. *Science* 320:1784–1787. <https://doi.org/10.1126/science.1155761>
- Lauring AS, Acevedo A, Cooper SB, Andino R. 2012. Codon usage determines the mutational robustness, evolutionary capacity, and virulence of an RNA virus. *Cell Host Microbe* 12:623–632. <https://doi.org/10.1016/j.chom.2012.10.008>
- Mueller S, Coleman JR, Papamichail D, Ward CB, Nimnual A, Futcher B, Skiena S, Wimmer E. 2010. Live attenuated influenza virus vaccines by computer-aided rational design. *Nat Biotechnol* 28:723–726. <https://doi.org/10.1038/nbt.1636>
- Mueller S, Stauff CB, Kalkeri R, Koidei F, Kushnir A, Tasker S, Coleman JR. 2020. A codon-pair deoptimized live-attenuated vaccine against respiratory syncytial virus is immunogenic and efficacious in non-human primates. *Vaccine (Auckl)* 38:2943–2948. <https://doi.org/10.1016/j.vaccine.2020.02.056>
- Stauff CB, Yang C, Coleman JR, Boltz D, Chin C, Kushnir A, Mueller S. 2019. Live-attenuated H1N1 influenza vaccine candidate displays potent efficacy in mice and ferrets. *PLoS One* 14:e0223784. <https://doi.org/10.1371/journal.pone.0223784>
- Moratorio G, Henningsson R, Barbezange C, Carrau L, Bordería AV, Blanc H, Beaucourt S, Poirier EZ, Vallet T, Bousnier J, Mounce BC, Fontes M, Vignuzzi M. 2017. Attenuation of RNA viruses by redirecting their evolution in sequence space. *Nat Microbiol* 2:17088. <https://doi.org/10.1038/nmicrobiol.2017.88>
- Carrau L, Rezelj VV, Noval MG, Levi LI, Megrian D, Blanc H, Weger-Lucarelli J, Moratorio G, Stapleford KA, Vignuzzi M. 2019. Chikungunya virus vaccine candidates with decreased mutational robustness are attenuated *in vivo* and have compromised transmissibility. *J Virol* 93:00775–19. <https://doi.org/10.1128/JVI.00775-19>
- Schön J, Barut GT, Trüeb BS, Halwe NJ, Berenguer Veiga I, Kratzel A, Ulrich L, Kelly JN, Brügger M, Wylezich C, et al. 2024. A safe, effective and adaptable live-attenuated SARS-CoV-2 vaccine to reduce disease and transmission using one-to-stop genome modifications. *Nat Microbiol* 9:2099–2112. <https://doi.org/10.1038/s41564-024-01755-1>
- Wu NC, Olson CA, Du Y, Le S, Tran K, Remenyi R, Gong D, Al-Mawsawi LQ, Qi H, Wu T-T, Sun R. 2015. Functional constraint profiling of a viral protein reveals discordance of evolutionary conservation and functionality. *PLoS Genet* 11:e1005310. <https://doi.org/10.1371/journal.pgen.1005310>
- Gallo G, Conceicao C, Tsigirigi C, Willett B, Graham SC, Bailey D. 2021. Application of error-prone PCR to functionally probe the morbillivirus Haemagglutinin protein. *J Gen Virol* 102:001580. <https://doi.org/10.1099/jgv.0.001580>
- Gong D, Zhang T-H, Zhao D, Du Y, Chapa TJ, Shi Y, Wang L, Contreras D, Zeng G, Shi P-Y, Wu T-T, Arumugaswami V, Sun R. 2018. High-throughput fitness profiling of zika virus E protein reveals different roles for glycosylation during infection of mammalian and mosquito cells. *iScience* 1:97–111. <https://doi.org/10.1016/j.isci.2018.02.005>
- Ikegame S, Hashiguchi T, Hung C-T, Dobrindt K, Brennand KJ, Takeda M, Lee B. 2021. Fitness selection of hyperfusogenic measles virus F proteins associated with neuropathogenic phenotypes. *Proc Natl Acad Sci U S A* 118:e2026027118. <https://doi.org/10.1073/pnas.2026027118>
- Vanhercke T, Ampe C, Tirry L, Denolf P. 2005. Reducing mutational bias in random protein libraries. *Anal Biochem* 339:9–14. <https://doi.org/10.1016/j.ab.2004.11.032>
- Álvarez-Rodríguez B, Velandia-Álvarez S, Toft C, Geller R. 2024. Mapping mutational fitness effects across the coxsackievirus B3 proteome reveals distinct profiles of mutation tolerability. *PLoS Biol* 22:e3002709. <https://doi.org/10.1371/journal.pbio.3002709>
- Ouyang WO, Tan TJC, Lei R, Song G, Kieffer C, Andrabi R, Matreyek KA, Wu NC. 2022. Probing the biophysical constraints of SARS-CoV-2 spike N-terminal domain using deep mutational scanning. *Sci Adv* 8:eadd7221. <https://doi.org/10.1126/sciadv.add7221>
- Bakhache W, Symonds-Orr W, McCormick L, Dolan PT. 2025. Deep mutation, insertion and deletion scanning across the Enterovirus A proteome reveals constraints shaping viral evolution. *Nat Microbiol* 10:158–168. <https://doi.org/10.1038/s41564-024-01871-y>
- Suphatrakul A, Posiri P, Srisuk N, Nantachokchawapan R, Onnong S, Mongkolsapaya J, Siridechadilok B. 2023. Functional analysis of flavivirus replicase by deep mutational scanning of dengue NS5. *Microbiology*. <https://doi.org/10.1101/2023.03.07.531617>
- Li Y, Arcos S, Sabsay KR, Te Velthuis AJW, Lauring AS. 2023. Deep mutational scanning reveals the functional constraints and evolutionary potential of the influenza A virus PB1 protein. *J Virol* 97:e01329-23. <https://doi.org/10.1128/jvi.01329-23>
- Haddox HK, Dings AS, Bloom JD. 2016. Experimental estimation of the effects of all amino-acid mutations to HIV's envelope protein on viral replication in cell culture. *PLoS Pathog* 12:e1006114. <https://doi.org/10.1371/journal.ppat.1006114>
- Wu NC, Young AP, Al-Mawsawi LQ, Olson CA, Feng J, Qi H, Chen S-H, Lu I-H, Lin C-Y, Chin RG, Luan HH, Nguyen N, Nelson SF, Li X, Wu T-T, Sun R. 2014. High-throughput profiling of influenza A virus hemagglutinin gene at single-nucleotide resolution. *Sci Rep* 4:4942. <https://doi.org/10.1038/srep04942>
- Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KHD, Dings AS, Navarro MJ, Bowen JE, Tortorici MA, Walls AC, King NP, Velesler D, Bloom JD. 2020. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* 182:1295–1310. <https://doi.org/10.1016/j.cell.2020.08.012>
- Gaiotto T, Hufton SE. 2016. Cross-neutralising nanobodies bind to a conserved pocket in the hemagglutinin stem region identified using yeast display and deep mutational scanning. *PLoS One* 11:e0164296. <https://doi.org/10.1371/journal.pone.0164296>
- Gai SA, Witttrup KD. 2007. Yeast surface display for protein engineering and characterization. *Curr Opin Struct Biol* 17:467–473. <https://doi.org/10.1016/j.sbi.2007.08.012>
- Javanmardi K, Chou C-W, Terrace CI, Annapareddy A, Kaoud TS, Guo Q, Lutgens J, Zorkic H, Horton AP, Gardner EC, Nguyen G, Boutz DR, Goike J, Voss WN, Kuo H-C, Dalby KN, Gollihar JD, Finkelstein IJ. 2021. Rapid characterization of spike variants via mammalian cell surface display. *Mol Cell* 81:5099–5111. <https://doi.org/10.1016/j.molcel.2021.11.024>
- Frank F, Keen MM, Rao A, Bassit L, Liu X, Bowers HB, Patel AB, Cato ML, Sullivan JA, Greenleaf M, Piantadosi A, Lam WA, Hudson WH, Ortlund EA. 2022. Deep mutational scanning identifies SARS-CoV-2 Nucleocapsid escape mutations of currently available rapid antigen tests. *Cell* 185:3603–3616. <https://doi.org/10.1016/j.cell.2022.08.010>
- Dadonaite B, Ahn JJ, Ort JT, Yu J, Furey C, Dosey A, Hannon WW, Baker ALV, Webby RJ, King NP, Liu Y, Hensley SE, Peacock TP, Moncla LH, Bloom JD. 2024. Deep mutational scanning of H5 hemagglutinin to inform influenza virus surveillance. *bioRxiv:2024.05.23.595634*. <https://doi.org/10.1101/2024.05.23.595634>
- Dadonaite B, Crawford KHD, Radford CE, Farrell AG, Yu TC, Hannon WW, Zhou P, Andrabi R, Burton DR, Liu L, Ho DD, Chu HY, Neher RA, Bloom JD. 2023. A pseudovirus system enables deep mutational scanning of the full SARS-CoV-2 spike. *Cell* 186:1263–1278. <https://doi.org/10.1016/j.cell.2023.02.001>
- Starr TN, Greaney AJ, Hannon WW, Loes AN, Hauser K, Dillen JR, Ferri E, Farrell AG, Dadonaite B, McCallum M, Matreyek KA, Corti D, Velesler D, Snell G, Bloom JD. 2022. Shifting mutational constraints in the SARS-CoV-2 receptor-binding domain during viral evolution. *Science* 377:420–424. <https://doi.org/10.1126/science.abo7896>
- Greaney AJ, Starr TN, Gilchuk P, Zost SJ, Binshtein E, Loes AN, Hilton SK, Huddleston J, Equia R, Crawford KHD, Dings AS, Nargi RS, Sutton RE, Suryadevara N, Rothlauf PW, Liu Z, Whelan SPJ, Carnahan RH, Crowe JE,

- Bloom JD. 2021. Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *Cell Host Microbe* 29:44–57. <https://doi.org/10.1016/j.chom.2020.11.007>
31. Fowler DM, Fields S. 2014. Deep mutational scanning: a new style of protein science. *Nat Methods* 11:801–807. <https://doi.org/10.1038/nmeth.3027>
  32. Thyagarajan B, Bloom JD. 2014. The inherent mutational tolerance and antigenic evolvability of influenza hemagglutinin. *Elife* 3:e03300. <https://doi.org/10.7554/eLife.03300>
  33. Fowler DM, Araya CL, Fleishman SJ, Kellogg EH, Stephany JJ, Baker D, Fields S. 2010. High-resolution mapping of protein sequence-function relationships. *Nat Methods* 7:741–746. <https://doi.org/10.1038/nmeth.1492>
  34. Hietpas RT, Jensen JD, Bolon DNA. 2011. Experimental illumination of a fitness landscape. *Proc Natl Acad Sci U S A* 108:7896–7901. <https://doi.org/10.1073/pnas.1016024108>
  35. Starita LM, Pruneda JN, Lo RS, Fowler DM, Kim HJ, Hiatt JB, Shendure J, Brzovic PS, Fields S, Klevit RE. 2013. Activity-enhancing mutations in an E3 ubiquitin ligase identified by high-throughput mutagenesis. *Proc Natl Acad Sci USA* 110. <https://doi.org/10.1073/pnas.1303309110>
  36. Carr CR, Crawford KHD, Murphy M, Galloway JG, Haddox HK, Matsen FA, Andersen KG, King NP, Bloom JD. 2024. Deep mutational scanning reveals functional constraints and antibody-escape potential of Lassa virus glycoprotein complex. *Immunity* 57:2061–2076. <https://doi.org/10.1016/j.immuni.2024.06.013>
  37. Teo QW, Wang Y, Lv H, Oade MS, Mao KJ, Tan TJC, Huan YW, Rivera-Cardona J, Shao EK, Choi D, Wang C, Tavakoli Dargani Z, Brooke CB, te Velhuis AJW, Wu NC. 2025. Probing the functional constraints of influenza A virus NEP by deep mutational scanning. *Cell Rep* 44:115196. <https://doi.org/10.1016/j.celrep.2024.115196>
  38. Yu Y, Kass MA, Zhang M, Youssef N, Freije CA, Brock KP, Aguado LC, Seifert LL, Venkittu S, Hong X, Shlomai A, de Jong YP, Marks DS, Rice CM, Schneider WM. 2024. Deep mutational scanning of hepatitis B virus reveals a mechanism for cis-preferential reverse transcription. *Cell* 187:2735–2745. <https://doi.org/10.1016/j.cell.2024.04.008>
  39. ElHefnawi M, Alaidi O, Mohamed N, Kamar M, El-Azab I, Zada S, Siam R. 2011. Identification of novel conserved functional motifs across most Influenza A viral strains. *Virology* 418:44–44. <https://doi.org/10.1016/j.viro.2011.08.044>
  40. Marsh GA, Rabadán R, Levine AJ, Palese P. 2008. Highly conserved regions of influenza A virus polymerase gene segments are critical for efficient viral RNA packaging. *J Virol* 82:2295–2304. <https://doi.org/10.1128/JVI.02267-07>
  41. Ngo V-N, Winski DP, Aho B, Kamath PL, King BL, Waters H, Zimmerberg J, Sodd A, Hess ST. 2024. Conserved sequence features in intracellular domains of viral spike proteins. *Virology (Auckl)* 599:110198. <https://doi.org/10.1016/j.viro.2024.110198>
  42. Hom N, Gentles L, Bloom JD, Lee KK. 2019. Deep mutational scan of the highly conserved influenza A virus M1 matrix protein reveals substantial intrinsic mutational tolerance. *J Virol* 93:e00161-19. <https://doi.org/10.1128/JVI.00161-19>
  43. Doud MB, Bloom JD. 2016. Accurate measurement of the effects of all amino-acid mutations on influenza hemagglutinin. *Viruses* 8:155. <https://doi.org/10.3390/v8060155>
  44. Lei R, Hernandez García A, Tan TJC, Teo QW, Wang Y, Zhang X, Luo S, Nair SK, Peng J, Wu NC. 2023. Mutational fitness landscape of human influenza H3N2 neuraminidase. *Cell Rep* 42:111951. <https://doi.org/10.1016/j.celrep.2022.111951>
  45. Dadonaite B, Brown J, McMahon TE, Farrell AG, Figgins MD, Asarnow D, Stewart C, Lee J, Logue J, Bedford T, Murrell B, Chu HY, Veleser D, Bloom JD. 2024. Spike deep mutational scanning helps predict success of SARS-CoV-2 clades. *Nature* 631:617–626. <https://doi.org/10.1038/s41586-024-07636-1>
  46. Bdeir N, Lüddecke T, Maaß H, Schmelz S, Rand U, Jacobsen H, Metzendorf K, Kulkarni U, Cossmann A, Stankov MV, Hoffmann M, Pöhlmann S, Blankenfeldt W, Dopfer-Jablonka A, Behrens GMN, Čičin-Šain I. 2025. Reverse mutational scanning of SARS-CoV-2 spike BA.2.86 identifies epitopes contributing to immune escape from polyclonal sera. *Nat Commun* 16:809. <https://doi.org/10.1038/s41467-025-55871-5>
  47. Jian F, Feng L, Yang S, Yu Y, Wang L, Song W, Yisimayi A, Chen X, Xu Y, Wang P, Yu L, Wang J, Liu L, Niu X, Wang J, Xiao T, An R, Wang Y, Gu Q, Shao F, Jin R, Shen Z, Wang Y, Wang X, Cao Y. 2023. Convergent evolution of SARS-CoV-2 XBB lineages on receptor-binding domain 455–456 synergistically enhances antibody evasion and ACE2 binding. *PLOS Pathog* 19:e1011868. <https://doi.org/10.1371/journal.ppat.1011868>
  48. Taylor AL, Starr TN. 2023. Deep mutational scans of XBB.1.5 and BQ.1.1 reveal ongoing epistatic drift during SARS-CoV-2 evolution. *bioRxiv*:2023.09.11.557279. <https://doi.org/10.1101/2023.09.11.557279>
  49. Taylor AL, Starr TN. 2024. Deep mutational scanning of SARS-CoV-2 Omicron BA.2.86 and epistatic emergence of the KP.3 variant. *Virus Evol* 10:veae067. <https://doi.org/10.1093/ve/veae067>
  50. Haddox HK, Galloway JG, Dadonaite B, Bloom JD, Matsen IV FA, DeWitt WS. 2023. Jointly modeling deep mutational scans identifies shifted mutational effects among SARS-CoV-2 spike homologs. *bioRxiv*:2023.07.31.551037. <https://doi.org/10.1101/2023.07.31.551037>
  51. Taminishi S, Li S, Higuchi Y, Kiritani Y, Motooka D, Ozaki Y, Arimori T, Ikemura N, Itoh Y, Matoba S, Okamoto T, Takagi J, Standley DM, Hoshino A. 2023. Deep mutational scanning of whole SARS-CoV-2 spike in an inverted infection system. *bioRxiv*. <https://doi.org/10.1101/2023.07.17.549430>
  52. Raisinighani N, Alshahrani M, Gupta G, Verkhivker G. 2024. AlphaFold2 modeling and molecular dynamics simulations of the conformational ensembles for the SARS-CoV-2 spike Omicron JN.1, KP2 and KP3 variants: mutational profiling of binding energetics reveals epistatic drivers of the ACE2 affinity and escape hotspots of antibody resistance. *Viruses* 16:1458. <https://doi.org/10.3390/v16091458>
  53. Raisinighani N, Alshahrani M, Gupta G, Verkhivker G. 2024. Ensemble-based mutational profiling and network analysis of the SARS-CoV-2 spike Omicron XBB lineages for interactions with the ACE2 receptor and antibodies: cooperation of binding hotspots in mediating epistatic couplings underlies binding mechanism and immune escape. *Int J Mol Sci* 25:4281. <https://doi.org/10.3390/ijms25084281>
  54. Weaver SC. 2006. Evolutionary influences in arboviral disease. *Curr Top Microbiol Immunol* 299:285–314. [https://doi.org/10.1007/3-540-26397-7\\_10](https://doi.org/10.1007/3-540-26397-7_10)
  55. Setoh YX, Amarilla AA, Peng NYG, Griffiths RE, Carrera J, Freney ME, Nakayama E, Ogawa S, Watterson D, Modhiran N, et al. 2019. Determinants of Zika virus host tropism uncovered by deep mutational scanning. *Nat Microbiol* 4:876–887. <https://doi.org/10.1038/s41564-019-0399-4>
  56. Stumpf MM, Brunetti T, Davenport BJ, McCarthy MK, Morrison TE. 2024. Deep mutationally scanned (DMS) CHIKV E3/E2 virus library maps viral amino acid preferences and predicts viral escape mutants of neutralizing CHIKV antibodies. *bioRxiv*:2024.12.04.626854. <https://doi.org/10.1101/2024.12.04.626854>
  57. Garrett ME, Itell HL, Crawford KHD, Basom R, Bloom JD, Overbaugh J. 2020. Phage-DMS: a comprehensive method for fine mapping of antibody epitopes. *iScience* 23:101622. <https://doi.org/10.1016/j.isci.2020.101622>
  58. Hanning KR, Minot M, Warrender AK, Kelton W, Reddy ST. 2022. Deep mutational scanning for therapeutic antibody engineering. *Trends Pharmacol Sci* 43:123–135. <https://doi.org/10.1016/j.tips.2021.11.010>
  59. Guthmiller JJ, Han J, Utset HA, Li L, Lan LY-L, Henry C, Stamper CT, McMahon M, O'Dell G, Fernández-Quintero ML, et al. 2022. Broadly neutralizing antibodies target a haemagglutinin anchor epitope. *Nature* 602:314–320. <https://doi.org/10.1038/s41586-021-04356-8>
  60. Chan KK, Dorosky D, Sharma P, Abbasi SA, Dye JM, Kranz DM, Herbert AS, Procko E. 2020. Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science* 369:1261–1265. <https://doi.org/10.1126/science.abc0870>
  61. Chan KK, Tan TJC, Narayanan KK, Procko E. 2021. An engineered decoy receptor for SARS-CoV-2 broadly binds protein S sequence variants. *Sci Adv* 7:eabf1738. <https://doi.org/10.1126/sciadv.abf1738>
  62. Zhang L, Dutta S, Xiong S, Chan M, Chan KK, Fan TM, Bailey KL, Lindeblad M, Cooper LM, Rong L, Gugliuzza AF, Shukla D, Procko E, Rehman J, Malik AB. 2022. Engineered ACE2 decoy mitigates lung injury and death induced by SARS-CoV-2 variants. *Nat Chem Biol* 18:342–351. <https://doi.org/10.1038/s41589-021-00965-6>
  63. Skeeters S, Bagale K, Stepanyuk G, Thieker D, Aguhob A, Chan KK, Dutzar B, Shalygin S, Shajahan A, Yang X, DaRosa PA, Frazier E, Sauer MM, Bogatzki L, Byrnes-Blake KA, Song Y, Azadi P, Tarcha E, Zhang L, Procko E. 2024. Modulation of the pharmacokinetics of soluble ACE2 decoy receptors through glycosylation. *Mol Ther Methods Clin Dev* 32:101301. <https://doi.org/10.1016/j.omtm.2024.101301>
  64. Mojica FJM, Diez-Villaseñor C, García-Martínez J, Soria E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive

- from foreign genetic elements. *J Mol Evol* 60:174–182. <https://doi.org/10.1007/s00239-004-0046-3>
65. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315:1709–1712. <https://doi.org/10.1126/science.1138140>
  66. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–821. <https://doi.org/10.1126/science.1225829>
  67. Karakus U, Thamamongood T, Ciminski K, Ran W, Günther SC, Pohl MO, Eletto D, Jeney C, Hoffmann D, Reiche S, et al. 2019. MHC class II proteins mediate cross-species entry of bat influenza viruses. *Nature* 567:109–112. <https://doi.org/10.1038/s41586-019-0955-3>
  68. Zhu Y, Feng F, Hu G, Wang Y, Yu Y, Zhu Y, Xu W, Cai X, Sun Z, Han W, Ye R, Qu D, Ding Q, Huang X, Chen H, Xu W, Xie Y, Cai Q, Yuan Z, Zhang R. 2021. A genome-wide CRISPR screen identifies host factors that regulate SARS-CoV-2 entry. *Nat Commun* 12:961. <https://doi.org/10.1038/s41467-021-21213-4>
  69. Zhang R, Kim AS, Fox JM, Nair S, Basore K, Klimstra WB, Rimkunas R, Fong RH, Lin H, Poddar S, Crowe JE, Doranz BJ, Fremont DH, Diamond MS. 2018. Mxra8 is a receptor for multiple arthritogenic alphaviruses. *Nature* 557:570–574. <https://doi.org/10.1038/s41586-018-0121-3>
  70. Richardson RB, Ohlson MB, Eitson JL, Kumar A, McDougal MB, Boys IN, Mar KB, De La Cruz-Rivera PC, Douglas C, Konopka G, Xing C, Schoggins JW. 2018. A CRISPR screen identifies IFI6 as an ER-resident interferon effector that blocks flavivirus replication. *Nat Microbiol* 3:1214–1223. <https://doi.org/10.1038/s41564-018-0244-1>
  71. Marceau CD, Puschnik AS, Majzoub K, Ooi YS, Brewer SM, Fuchs G, Swaminathan K, Mata MA, Elias JE, Sarnow P, Carette JE. 2016. Genetic dissection of *Flaviviridae* host factors through genome-scale CRISPR screens. *Nature* 535:159–163. <https://doi.org/10.1038/nature18631>
  72. Ma Y, Walsh MJ, Bernhardt K, Ashbaugh CW, Trudeau SJ, Ashbaugh IY, Jiang S, Jiang C, Zhao B, Root DE, Doench JG, Gewurz BE. 2017. CRISPR/Cas9 screens reveal Epstein-Barr virus-transformed B cell host dependency factors. *Cell Host Microbe* 21:580–591. <https://doi.org/10.1016/j.chom.2017.04.005>
  73. OhAinle M, Helms L, Vermeire J, Roesch F, Humes D, Basom R, Delrow JJ, Overbaugh J, Emerman M. 2018. A virus-packageable CRISPR screen identifies host factors mediating interferon inhibition of HIV. *Elife* 7:e39823. <https://doi.org/10.7554/eLife.39823>
  74. Prelli Bozzo C, Laliberté A, De Luna A, Pastorio C, Regensburger K, Krebs S, Graf A, Blum H, Volcic M, Sparrer KMJ, Kirchhoff F. 2024. Replication competent HIV-guided CRISPR screen identifies antiviral factors including targets of the accessory protein Nef. *Nat Commun* 15:3813. <https://doi.org/10.1038/s41467-024-48228-x>
  75. Finkel Y, Nachshon A, Aharon E, Arazi T, Simonovsky E, Dobešová M, Saud Z, Gluck A, Fisher T, Stanton RJ, Schwartz M, Stern-Ginossar N. 2024. A virally encoded high-resolution screen of cytomegalovirus dependencies. *Nature* 630:712–719. <https://doi.org/10.1038/s41586-024-07503-z>
  76. King CR, Liu Y, Amato KA, Schaack GA, Mickelson C, Sanders AE, Hu T, Gupta S, Langlois RA, Smith JA, Mehle A. 2023. Pathogen-driven CRISPR screens identify TREX1 as a regulator of DNA self-sensing during influenza virus infection. *Cell Host Microbe* 31:1552–1567. <https://doi.org/10.1016/j.chom.2023.08.001>
  77. Dodkins R, Delaney JR, Overton T, Scholle F, Frias-De-Diego A, Crisci E, Huq N, Jordan I, Kimata JT, Findley T, Goldberg IG. 2023. A rapid, high-throughput, viral infectivity assay using automated brightfield microscopy with machine learning. *SLAS Technol* 28:324–333. <https://doi.org/10.1016/j.slast.2023.07.003>
  78. Wang T-E, Chao T-L, Tsai H-T, Lin P-H, Tsai Y-L, Chang S-Y. 2020. Differentiation of cytopathic effects (CPE) induced by influenza virus infection using deep convolutional neural networks (CNN). *PLOS Comput Biol* 16:e1007883. <https://doi.org/10.1371/journal.pcbi.1007883>
  79. Werner J, Kronberg RM, Stachura P, Ostermann PN, Müller L, Schaal H, Bhatia S, Kather JN, Borkhardt A, Pandya AA, Lang KS, Lang PA. 2021. Deep transfer learning approach for automatic recognition of drug toxicity and inhibition of SARS-CoV-2. *Viruses* 13:610. <https://doi.org/10.3390/v13040610>
  80. Sarfati H, Naftaly S, Papo N, Keasar C. 2022. Predicting mutant outcome by combining deep mutational scanning and machine learning. *Proteins* 90:45–57. <https://doi.org/10.1002/prot.26184>
  81. Zhou Y, Pan Q, Pires DEV, Rodrigues CHM, Ascher DB. 2023. DDMut: predicting effects of mutations on protein stability using deep learning. *Nucleic Acids Res* 51:W122–W128. <https://doi.org/10.1093/nar/gkad472>
  82. Gelman S, Fahlberg SA, Heinzelman P, Romero PA, Gitter A. 2021. Neural networks to learn protein sequence–function relationships from deep mutational scanning data. *Proc Natl Acad Sci USA* 118:e2104878118. <https://doi.org/10.1073/pnas.2104878118>
  83. Wu NC, Du Y, Le S, Young AP, Zhang T-H, Wang Y, Zhou J, Yoshizawa JM, Dong L, Li X, Wu T-T, Sun R. 2016. Coupling high-throughput genetics with phylogenetic information reveals an epistatic interaction on the influenza A virus M segment. *BMC Genomics* 17:46. <https://doi.org/10.1186/s12864-015-2358-7>
  84. Esposito D, Weile J, Shendure J, Starita LM, Papenfuss AT, Roth FP, Fowler DM, Rubin AF. 2019. MaveDB: an open-source platform to distribute and interpret data from multiplexed assays of variant effect. *Genome Biol* 20:223. <https://doi.org/10.1186/s13059-019-1845-6>
  85. Atlas of Variant Effects Alliance. Atlas Var Eff Alliance. Available from: <https://www.varianteffect.org>. Retrieved 23 Dec 2024.
  86. Shapiro JS, Langlois RA, Pham AM, Tenover BR. 2012. Evidence for a cytoplasmic microprocessor of pri-miRNAs. *RNA* 18:1338–1346. <https://doi.org/10.1261/rna.032268.112>