

## Article

# U-CAN-seq: A Universal Competition Assay by Nanopore Sequencing

Jennifer Diaz <sup>1</sup>, John Sears <sup>1</sup>, Che-Kang Chang <sup>1</sup>, Jane Burdick <sup>2</sup>, Isabella Law <sup>2</sup>, Wes Sanders <sup>1</sup>, Colton Linnertz <sup>2</sup>, Paul Sylvester <sup>2</sup>, Nathaniel Moorman <sup>1,3</sup>, Martin T. Ferris <sup>2</sup> and Mark T. Heise <sup>1,2,3,\*</sup>

<sup>1</sup> Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA; jloome@live.unc.edu (J.D.)

<sup>2</sup> Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA

<sup>3</sup> The Rapidly Emerging Antiviral Drug Development Initiative (READDI), Chapel Hill, NC 275114, USA

\* Correspondence: mark\_heisem@med.unc.edu

**Abstract:** RNA viruses quickly evolve subtle genotypic changes that can have major impacts on viral fitness and host range, with potential consequences for human health. It is therefore important to understand the evolutionary fitness of novel viral variants relative to well-studied genotypes of epidemic viruses. Competition assays are an effective and rigorous system with which to assess the relative fitness of viral genotypes. However, it is challenging to quickly and cheaply distinguish and quantify fitness differences between very similar viral genotypes. Here, we describe a protocol for using reverse transcription PCR in combination with commercial nanopore sequencing services to perform competition assays on untagged RNA viruses. Our assay, called the Universal Competition Assay by Nanopore Sequencing (U-CAN-seq), is relatively cheap and highly sensitive. We used a well-studied N24A mutation in the chikungunya virus (CHIKV) nsp3 gene to confirm that we could detect a competitive disadvantage using U-CAN-seq. We also used this approach to show that mutations to the CHIKV 5' conserved sequence element that disrupt sequence but not structure did not affect the fitness of CHIKV. However, similar mutations to an adjacent CHIKV stem loop (SL3) did cause a fitness disadvantage compared to wild-type CHIKV, suggesting that structure-independent, primary sequence determinants in this loop play an important role in CHIKV biology. Our novel findings illustrate the utility of the U-CAN-seq competition assay.

**Keywords:** competition assay; nanopore sequencing; alphavirus; RNA virus; viral fitness



**Citation:** Diaz, J.; Sears, J.; Chang, C.-K.; Burdick, J.; Law, I.; Sanders, W.; Linnertz, C.; Sylvester, P.; Moorman, N.; Ferris, M.T.; et al. U-CAN-seq: A Universal Competition Assay by Nanopore Sequencing. *Viruses* **2024**, *16*, 636. <https://doi.org/10.3390/v16040636>

Academic Editor: Jeremy Thompson

Received: 29 February 2024

Revised: 9 April 2024

Accepted: 17 April 2024

Published: 19 April 2024



**Copyright:** © 2024 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

RNA viruses, which generally lack proofreading activity, evolve quickly by acquiring mutations in their small, labile genomes that shape multiple aspects of their biology, including viral fitness, immune escape, viral transmissibility, and host range. These changes can impact viral pathogenesis or even shape the course of epidemics [1,2]. As such epidemics unfold, researchers often try to understand how novel genotypic variants will affect viral spread, especially in the context of resistance to antiviral drugs or escape from vaccine-induced immunity. However, small changes in RNA virus genomes can have major effects on viral fitness while having no detectable growth defects in vitro. For instance, a mutation in the SARS-CoV-2 RNA-dependent RNA polymerase that conferred reduced sensitivity to the antiviral drug remdesivir did not affect virus replication in growth curves, but it did confer a competitive disadvantage compared to WT virus in a competition assay [3]. The same pattern was shown with zidovudine resistance mutations in HIV-1 [4], nevirapine resistance mutations in HIV-1 [5], and baloxivir resistance mutations in influenza A virus [6]. All of these resistance mutations caused no overt growth defects but did show disadvantages in the context of competition between viral genotypes, informing work to prevent the spread of drug resistant viruses.

Competition assays provide a method to stringently and sensitively measure the relative fitness of two virus genotypes [7–10]. During competition assays, two virus genotypes infect the same cell culture well or animal tissue, and the ratio of virus genotypes is measured in the inoculum and in the post-infection samples. As viruses compete to infect available cells and produce progeny genomes, small differences in virus replication rates become amplified over multiple rounds of replication. This allows researchers to measure subtle changes in replication that can have large effects on pathogenesis and epidemics. Competition assays require quantification of the relative abundance of two very similar genotypes of virus, often differing by only one or a few nucleotides, and this poses technical challenges. Some researchers use viruses encoding fluorescent tags, which allow detection of the genotype ratios via flow cytometry [9], or genetic tags, which allow detection of the virus genotypes by standard quantitative PCR [7]. However, tagging the viruses may confound studies by introducing small, or even large, advantages or disadvantages in replication [11–14]. RNA sequencing can detect untagged viruses [10], but it can be cost-prohibitive [15,16], require complex technical processing pipelines [17], and, in some cases, may come with multi-week delays between running experiments and getting sequencing results.

Commercial nanopore sequencing services allow the complete sequencing of plasmid or linear amplicon sequences at a relatively low cost with rapid data turnaround. We have leveraged commercial nanopore sequencing to develop a cost-effective competition assay for highly similar, untagged genotypes of RNA viruses with rapid data turn-around times. Based on this, we developed a highly accessible competition assay protocol that combines reverse transcription (RT)-PCR with commercial nanopore sequencing, which we call Universal Competition Assay by Nanopore Sequencing, or U-CAN-seq. We then used several previously described CHIKV mutants to demonstrate that the assay can rapidly and quantitatively detect fitness differences between two closely related, untagged viruses. Lastly, we used U-CAN-seq to detect a previously unrecognized fitness difference due to mutations in CHIKV RNA structures.

## 2. Materials and Methods

### 2.1. Cell Culture and Viruses

Vero81 cells (ATCC CCL-81, Manassas, VA, USA) were cultured in DMEM (Gibco, Waltham, MA, USA) + 10% fetal bovine serum (FBS, R&D Systems, Minneapolis, MN, USA) + 1% L-glutamine (Gibco, Waltham, MA, USA) at 37 °C in 5% CO<sub>2</sub>. All viruses were generated from plasmid-encoded infectious clones by *in vitro* transcription (mMessage Machine Sp6 kit, Ambion, Austin, TX, USA) to produce viral RNA, which was electroporated into BHK21 cells (ATCC CCL-10, Manassas, VA, USA). The virus was recovered from cell supernatants at 24–28 h post infection. All genotypes of virus described were on the 181/25 CHIKV background.

### 2.2. Growth Curves

We performed growth curves by infecting Vero81 cells with each virus at a multiplicity of infection (MOI) = 0.01 in 400 µL of medium in 6-well plates, rocking every 15 min for 1 h at 37 °C. The cells were then washed three times with Dulbecco's phosphate-buffered saline (DPBS, Gibco, Waltham, MA, USA) to remove unbound input virus, and 3 mL of medium was added to each well prior to incubation at 37 °C. We harvested 100 µL of supernatant from each well and replaced it with fresh medium at 1, 8, 16, and 24 h post infection (hpi), and quantified viral titers by plaque assay as described previously [18]. Briefly, supernatants were serially diluted 1:10 six times and plated in duplicate on 90% confluent Vero81 cells, the plates were rocked every 15 min for one hour, and 2 mL of 1.25% carboxymethylcellulose (Sigma-Aldrich, St. Louis, MO, USA) overlay in MEM alpha (Gibco, Waltham, MA, USA) was added. Plaques were allowed to form for 3 days before plates were fixed by adding 2 mL of 4% paraformaldehyde (Thermo Scientific, Waltham, MA, USA) washing, and staining in 0.25% crystal violet (Fisher Chemical, Pittsburgh, PA, USA).

### 2.3. Competition Assays

The full U-CAN-seq protocol is published at protocols.io (<https://www.protocols.io/private/F8AEFAE4D19711EE9D470A58A9FEAC02>, accessed on 27 February 2024). Briefly, we performed competition assays by infecting Vero81 cells at an MOI = 0.01 in 400 µL in six-well plates, rocking every 15 min for 1 h, and adding 3 mL of medium as described above. We infected the cells with each virus genotype alone, along with mixes of two genotypes (a wild-type and a mutant) in 1:1, 5:1, and 10:1 ratios, infecting five wells each. 300 µL of each inoculum was mixed with 900 µL of Trizol-LS (Invitrogen, Carlsbad, CA, USA). The infected cells were incubated for 24 h at 37 °C, at which point, we harvested 300 µL of supernatant from each well, which was then mixed with 900 µL of Trizol-LS to isolate RNA.

We isolated RNA from each sample (3 replicate inoculum samples and 5 replicate output supernatant samples per genotype ratio) according to the Trizol-LS Reagent (Invitrogen, Carlsbad, CA, USA) manufacturer's instructions. We used a gene-specific reverse primer with the Superscript III first-strand synthesis kit (Invitrogen, Carlsbad, CA, USA) to generate cDNA (Table 1). Gene-specific PCR primers were used to amplify the region around each region of interest for 35 PCR cycles (primers listed in Table 1). Amplification of the target, and lack of amplification of a mock sample, was confirmed by running the samples on an agarose gel (Figure S1). We used the DNA Clean & Concentrator-5 kit (Zymo Research Corporation, Irvine, CA, USA) to purify DNA from each sample, eluting in 12 µL. We performed one PCR replicate per sample.

**Table 1.** Primers and sequences.

	nsp3 N24A	scrSL3	scr5' CSE
Mutation location (base pairs in CHIKV genome)	4145–4147	91–145	175–217
RT primer	CGCAGTCTATGGA GATGTGCTCATC	CATGAACGGGGT TGTGTCGAACCC	CATGAACGGGGT TGTGTCGAACCC
PCR forward primer	GGTACTGGTGGGCGAC TTGACTAATCCGCC	GCAAAGCAAGA GATTAATAACCC	GCAAAGCAAGA GATTAATAACCC
PCR reverse primer	CGCAGTCTATGGA GATGTGCTCATC	CATGAACGGGGT TGTGTCGAACCC	CATGAACGGGGT TGTGTCGAACCC
WT k-mer	AGAGTGCGTGGTT AACGCCGCAAC	ACGTGGACATAGACGC TGACAGCGCCTTTT	ATCGAATGACCATGCTAATGCT AGAGCGTTCTCGCATCTAGCC
Mutant k-mer	AGAGTGCGTGGTT AACGCCGCGCC	ATGTAGACATAGACGC TGACAGCGCCTTTCT	GTCGAATGACCATGCCAACGCC AGAGCGTTTTTCGCATCTGGCT

The primers listed were used for reverse transcription and PCR. The k-mers listed were used with a Unix 'grep' search to search the raw fastq sequencing results files for genotype-specific sequences (k-mers, where k = length of the relevant sequence string).

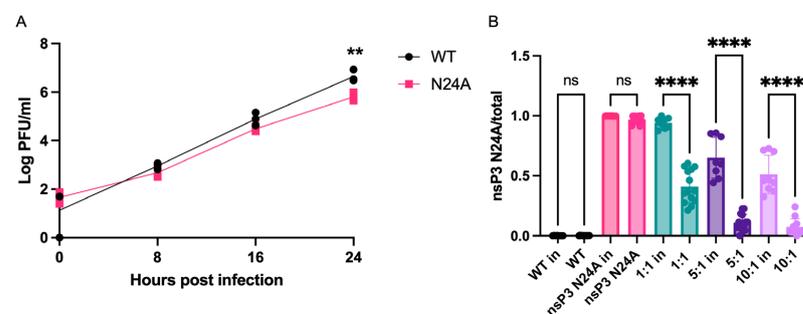
We submitted each amplicon to the Plasmidsaurus amplicon sequencing service in Louisville, KY, USA, which performs Oxford Nanopore sequencing using the v14 library prep chemistry and R10.4.1 flow cells. FASTQ files containing each raw sequencing read were returned. The FASTQ files from our amplicon sequencing were analyzed using simple command-line prompts that can be used in any Unix-based system through a command line interface (CLI). Code to perform the analysis and create an output csv file summarizing the data can be found in Figure S2. We used the Unix 'grep' command, a utility designed to search for matching strings in text files, to identify and count specific k-mers (Table 1) (10–40 nucleotides containing the site of interest, that are unique in the CHIKV genome otherwise) in our sequencing data. Note that this method works better than standard single nucleotide polymorphism (SNP) calling if at least two nucleotides differ between the two genotypes. In this way, we can avoid misalignments and misestimates of abundances of genotypes due to sequencing errors near our relevant mutant sites. We counted the

number of occurrences of each k-mer in each FASTQ file, and we calculated the proportion of each k-mer compared to the total number of wild-type (WT) and mutant k-mers counted for that file. Total k-mer occurrences in the sequencing files varied widely among samples, from 4 to 1712 (Table S1). We used one-way ANOVA followed by relevant pre-planned contrasts for pairwise comparisons to compare the mean genotype frequencies of the rarer genotypes of the supernatants to the relevant inocula for each ratio of genotypes. We used Sidak's multiple comparisons test to control for multiple comparisons. We determined results as significant if  $p < 0.05$ . (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ). We expect that genotypes that have no fitness difference compared to the wild-type should have equal genotype ratios in the inocula and in the post-infection samples, while genotypes that have a competitive advantage or disadvantage have higher or lower proportions, respectively, in the post-infection samples than in the inocula.

### 3. Results

#### 3.1. U-CAN-seq Detects a Previously Described Growth Defect in the CHIKV nsP3 N24A Genotype

To validate the U-CAN-seq protocol, we first analyzed the relative fitness of a chikungunya virus (CHIKV) mutant with a known growth defect compared to WT CHIKV. For this proof-of-concept experiment, we chose a mutation within the nsP3 macrodomain (N24A). The N24A mutation disrupts the ribosylhydrolase activity of the nsP3 macrodomain, which results in attenuated growth [19]. We first confirmed its growth defect by performing a multi-step growth curve of WT and N24A CHIKV on Vero81 cells at an MOI of 0.01 (Figure 1). By 24 h post infection (hpi), the N24A genotype had titers approximately 10-fold lower than the WT. Next, we performed a competition assay to determine if N24A had a fitness advantage or disadvantage compared to the WT virus, with the expectation that we would detect a fitness disadvantage compared to the WT. We infected Vero81 cells at a ratio of 1:1, 5:1, and 10:1 WT:N24A virus and isolated RNA from the inoculum and from cell supernatants at 24 hpi. We used RT-PCR to amplify the target region in nsP3 and submitted the amplicons for nanopore sequencing.



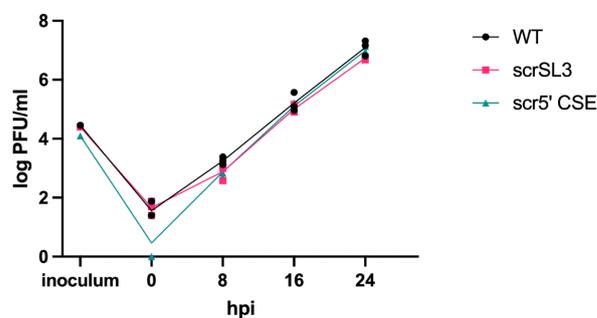
**Figure 1.** U-CAN-seq can reliably detect a fitness defect caused by the chikungunya virus (CHIKV) nsP3 N24A genotype. **(A)** Vero81 cells were infected with wild-type (WT) or nsP3 N24A CHIKV and samples were taken at 0, 8, 16, and 24 h post infection (hpi). Titers were measured by plaque assay. One representative experiment of three independent experiments shown. **(B)** Vero81 wells were infected at a multiplicity of infection of 0.01 with CHIKV WT (black), nsP3 N24A (pink), or 1:1 (turquoise), 5:1 (purple), or 10:1 (light purple) ratios of WT:N24A. RNA was extracted from the inocula and from 24 hpi cell supernatants, and reverse transcription (RT)-PCR was used to generate amplicons for nanopore sequencing. The raw sequencing files were searched for genotype-specific sequences and the proportions of the two genotypes in each sample were calculated. Combined results of three independent experiments shown. ns, not significant, \*\*  $p < 0.01$ , \*\*\*\*  $p < 0.0001$ .

We successfully amplified sequences from samples with RNA concentrations as low as 6 ng/ $\mu$ L, and although the inocula had less robust amplification on average than the supernatants of infected cells (Figure S1), our sequence of interest was consistently amplified from both supernatants and the inoculum. The N24A mutant had a lower genotype frequency in the supernatant of infected cells than it did in the inoculum, confirming that it

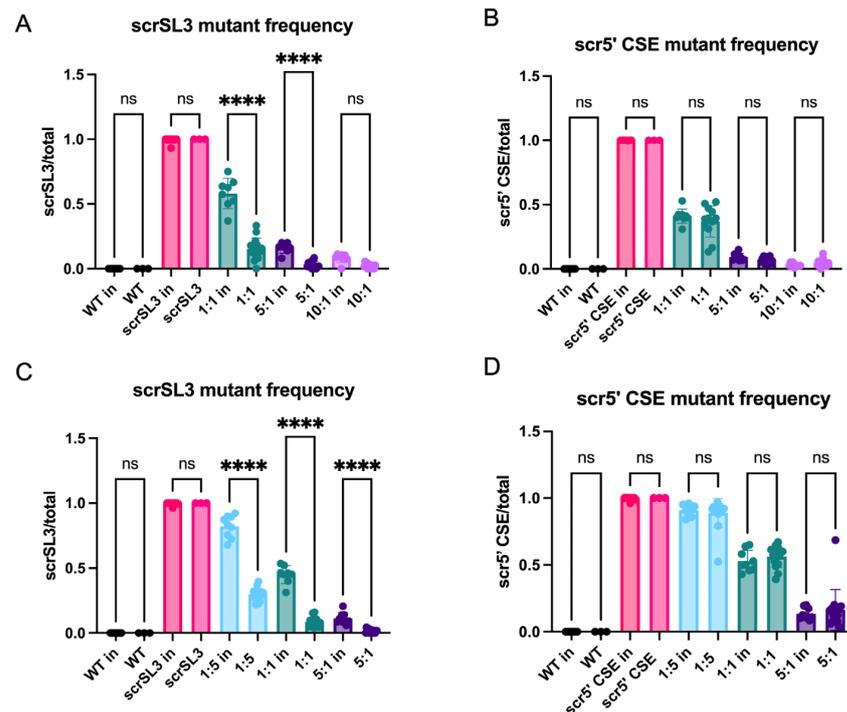
is less fit than the WT virus (Figure 1) and confirming that our assay can detect a known growth defect. Of note, the assay was also robust even when the input inoculum was inexact. For example, for the intended 1:1 WT:N24A input inoculum, we calculated an actual genotype ratio of 1:9.1. This deviation from the desired input ratios was most likely due to titering the WT and N24A viruses on separate days, leading to variation. We therefore recommend titering viruses of interest side-by-side before performing U-CAN-seq to ensure accurate titers. Regardless, N24A had a reduced allele frequency at output compared to the input. Because we statistically compare input to output rather than to an expected ratio, our assay is robust and interpretable even if input ratios vary from the intended ratios.

### 3.2. U-CAN-seq Can Detect Previously Unrecognized Fitness Differences Due to Mutations in an RNA Structure

Next, we used U-CAN-seq to analyze fitness effects of a panel of previously described mutants in the 5' end of the CHIKV genome that were used to characterize the role of RNA secondary structure in CHIKV replication [18]. We had previously shown that disruption of secondary structure in the two stem loops that comprise the alphavirus 5' conserved sequence element (5' CSE), as well as a stem loop immediately prior to the 5' CSE, SL3, resulted in attenuation of CHIKV replication [18]. Structure rescue mutations in SL3 (scrSL3) and the 5' CSE (scr5' CSE) that restored the secondary structure while changing the primary sequence had no impact on CHIKV replication compared to the wild-type virus, thereby demonstrating the importance of RNA secondary structure in these stem loops for viral replication [18]. To test whether these structure rescue mutants had more subtle fitness defects, we performed a U-CAN-seq assay. We first confirmed that scrSL3 and scr5' CSE had no detectable effect on viral growth by testing the viruses head-to-head in multistep growth curves in Vero81 cells (Figure 2). When compared to WT CHIKV in our competition assay, the scr5' CSE virus showed no differences in fitness (Figure 3B). Because we have previously shown that disruption of the predicted secondary structure of the 5' CSE reduces viral replication [18], our results here further confirm that the impact of the 5' CSE on viral replication is mediated through RNA secondary structure. However, the scrSL3 genotype made up a lower proportion of post-infection sequences than of the inoculum sequences, indicating that scrSL3 has a competitive disadvantage compared to the wild-type virus (Figure 3A). Therefore, these studies indicate that although the secondary structure of SL3 is clearly important for viral replication [18], the primary sequence in this region also plays a role in viral fitness. To confirm our initial results, we performed competition assays again on scrSL3 and scr5' CSE, this time including an input ratio in which the mutant was in the majority (WT:mutant 51:5) (Figure 3C,D). As in our initial experiment, we detected a competitive disadvantage in scrSL3 but not in scr5' CSE compared to the WT. This demonstrates U-CAN-seq is capable of detecting fitness differences between viral genotypes at a variety of input genotype ratios.



**Figure 2.** Mutations that impact the RNA sequence but not the predicted structure of the stem loop 3 (SL3) and the 5' conserved sequence element (5' CSE) have no detectable impact on the titer. Vero81 cells were infected with the WT, scrSL3, or scr5' CSE CHIKV virus, and samples were taken at 0, 8, 16, and 24 h post infection. Titers were measured by plaque assay. One representative experiment of three independent experiments shown.



**Figure 3.** U-CAN-seq detects a fitness defect in the scrSL3 mutant but not the scr5' CSE mutant. Vero81 wells were infected at a multiplicity of infection of 0.01 with CHIKV WT (black), scrSL3 (pink) (A,C), or scr5' CSE (pink) (B,D) or with the indicated ratios of WT–mutant. (A,B) WT–mutant ratios in inocula were 1:1 (turquoise), 5:1 (purple), and 10:1 (light purple). (C,D) WT–mutant ratios in inocula were 1:5 (blue), 1:1 (turquoise), and 5:1 (purple). Inocula were sampled, and supernatants were sampled at 24 h post infection. RNA was extracted, and RT-PCR was used to generate amplicons for nanopore sequencing. The raw sequencing files were searched for genotype-specific sequences and the proportions of the two genotypes in each sample were calculated. Combined results of three independent experiments shown. ns, not significant, \*\*\*\*  $p < 0.0001$ .

#### 4. Discussion

Competition assays are an important virology tool for sensitively assessing viral fitness, but many of the existing systems for performing competition assays rely either on tagged viruses or high throughput RNA sequencing approaches, which increase the required time and costs of these assays. Furthermore, while tagged viruses have been highly useful for testing virus fitness, the potential impact of these tags on viral fitness often requires the inclusion of additional controls, increasing the time and cost commitment associated with these studies. Therefore, the protocol for a Universal Competition Assay by Nanopore sequencing (U-CAN-seq) is an important improvement upon previous protocols by providing a relatively inexpensive and powerful tool to perform competition assays on untagged viruses. U-CAN-seq expands the new accessibility of high-throughput DNA sequencing to a problem that traditionally is solved using RNA sequencing, which remains expensive despite its steady decrease in price.

U-CAN-seq has a current cost of about \$950 per experiment, for 33 samples per experiment (Table 2) after RNA extraction is performed, for researchers in the continental United States. The two largest current expenses are the reverse transcriptase and sample sequencing. It may be possible to reduce expenses by using a less costly and less processive reverse transcriptase enzyme, although we validated our protocol using only the listed reagents. Importantly, U-CAN-seq also expands the accessibility of sequencing-based competition assays to researchers at small institutions without bioinformatics collaborators because it requires only a single line of code, usable on common operating systems for easy analysis. We hope that our technique will allow even undergraduate-based research laboratories to carry out competition assays.

Table 2. Costs associated with U-CAN-seq protocol.

Step	Reagent	Supplier	Product Number	Reactions	Cost of Kit	Reactions/Kit	Cost
RT	Superscript III first-strand synthesis kit	Invitrogen, Carlsbad, CA, USA	18080051	33	558	50	368.28
	Reverse primer	Invitrogen, Carlsbad, CA, USA			14	25 nmol	14
PCR (20 µL reaction)	PCR tubes	Greiner, Kremsmünster, Austria	07-000-717	33	254.5	960	8.75
	Phusion kit	NEB, Ipswich, MA, USA	M0530L	33	480	1250	12.67
	Forward primer	Invitrogen, Carlsbad, CA, USA			14	25 nmol	14
	dNTPs	NEB, Ipswich, MA, USA	N0447L	33	274	2000	4.52
	PCR tubes	Greiner, Kremsmünster, Austria	07-000-717	33	254.5	960	8.75
	Clean up	Zymo DNA Clean & Concentrator-5	Zymo, Irvine, CA, USA	D4004	33	323	200
Sequencing	Amplicon sequencing	Plasmidsaurus, Louisville, KY, USA		33	15	1	495
						SUM	979.27

As noted above, U-CAN-seq allows the researcher to investigate the relative fitness of two RNA viruses without adding tags, eliminating both the cost of creating a tagged infectious clone and any artificial fitness costs that a tag might introduce. Introducing fluorescent tags into small RNA genomes can allow for easy and cheap quantification of two viruses in a mixed population, but can also come at notable fitness costs, with many reporter systems facing issues of unstable transgenes [11–14,20]. These fitness costs may be due to disruptions in genome packaging, the diversion of resources to produce non-advantageous fluorescent proteins, or other unintended disruptions. It is also possible to insert short sequence tags into RNA viruses to allow for quantification of two genotypes by qPCR. Even if these tags do not disrupt the coding sequence or fall within untranslated regions, one must be careful about disrupting potentially important RNA secondary structures [18,21–24] or other RNA regulatory elements that might also impact viral fitness. U-CAN-seq avoids these potential issues by eliminating the need for tagged viruses.

Nanopore sequencing is notably error prone, with some single-nucleotide errors, as well as high rates of called insertions and deletions [25–27]. This makes performing local alignments somewhat challenging. Additionally, RNA viruses have relatively high mutation rates that can contribute to error in ratio assessment. For instance, in our datasets, our regions of interest differ from the exact k-mer sequences by one nucleotide in 0–2% of reads (Table S1). This suggests that despite the high error rate of nanopore sequencing, errors are introduced to our small sequence of interest at quite a low rate. Additionally, our method of searching strings of nucleotides for exact matches eliminates erroneous reads from analysis, increasing the specificity of the data. Of course, searching for perfect matches also eliminates any reads with local errors, effectively reducing the read depth of the assay by 0–2% compared to assays that tolerate single-nucleotide mismatches. However, we believe this reduction in depth is effectively countered by our increased precision in assessment of relevant genotypes.

We assessed within-treatment-group standard deviations in the mutant genotype frequency across experimental replicates, and we found that these values were relatively small (median standard deviation in input inoculum 0.0358, range 0.0011–0.1178; median in har-

vested supernatants 0.0262, range 0.0043–0.1128). The estimated variation shrinks moderately when we consider groups where all samples had at least 50 reads including either the WT or the mutant k-mer (median standard deviation in input inoculum 0.0214 range 0.0008–0.1071, or in harvested supernatants median 0.024, range = 0.0043–0.1128). We used these values to estimate the relative frequency changes we would be powered to detect using U-CAN-seq (Table 3). In general, even at low (e.g., 3) within-group sample sizes (where a sample is RNA from the supernatant of a single well of infected cells), we can detect genotype frequency changes of ~8–12%, and this number shrinks to ~4% at sample sizes of 10. The high power of this protocol makes it useful for teasing out small differences in fitness without extensive passaging of virus. We found that the largest source of variation in calculated genotype frequencies in this assay is variation in read depth during nanopore sequencing, with low read depth resulting in less precise genotype frequency measurements. We therefore recommend aiming for read depths of >100 reads in our target region.

**Table 3.** Power estimates: Expected detectable effects testing for an alpha = 0.05, beta = 0.9.

	St.Dev = 0.0358	St.Dev = 0.0214
$n = 3$	11.73%	7.97%
$n = 5$	7.67%	5.21%
$n = 10$	5.01%	3.4%

Using two different estimates of the standard deviation of genotype frequencies of our assay (the highest and lowest median standard deviation estimates), we calculated the minimum effect size our assay has the power to detect for a given sample size, assuming an alpha (likelihood of incorrectly calculating a statistically significant difference between two groups) of 0.05 and a beta (probability of incorrectly failing to identify a statistically significant difference) of 0.9.

Due to the high power of the assay, researchers can deploy U-CAN-seq to study the relative fitness of RNA viruses with small fitness differences and minimal sequence differences. In proof-of-concept studies using U-CAN-seq, we demonstrated that in CHIKV, the nsp3 N24A mutation confers a competitive disadvantage in cell culture compared to the WT. This finding is consistent with previously published data showing that in Sindbis virus, the nsp3 N24A mutation slightly reduces viral growth in cell culture [28,29]. It is also consistent with studies demonstrating that the N24A mutation reduces mono-ADP-ribosylhydrolase activity [19], inhibition of which decreases viral replication [30]. We also used U-CAN-seq to investigate more rigorously the sequence requirements of two RNA structures, showing that in Vero81 cells, disruption to the sequence but not the structure of the 5' CSE in CHIKV does not affect viral fitness, but that disruption of the sequence but not the structure in SL3 does have a fitness cost (Figure 3). This novel finding illustrates the importance of using rigorous tests such as competition assays to understand the complex evolutionary pressures on RNA virus sequence and structure, as well as the utility of our novel approach to competition assays.

We have demonstrated the effectiveness of U-CAN-seq to analyze the relative fitness of two genotypes of RNA virus in cell culture. The U-CAN-seq protocol remains to be tested in animal models and other non-cell culture systems, but there is no conceptual reason why the protocol could not be adapted beyond cell culture. The technique is specifically valuable to analyze genotypes that differ at one specific locus and can be applied to any RNA virus. However, U-CAN-seq may be especially powerful when used to study alphaviruses because cells infected with alphaviruses become resistant to further infection by alphaviruses within 15 min of initiation of the first infection [31–35]. This effect, called superinfection exclusion, does not occur in all RNA viruses and causes small differences in replication speed to confer a major competitive advantage on the faster viral genotype, increasing the effect size in a competition assay.

U-CAN-seq may be less useful for analysis of genotypes that differ at many locations across the genome. It is essential that the two genotypes of interest do not differ at the primer binding site in order to not introduce excessive bias in the PCR step. This is particularly essential because, with our method, it is not possible to determine how

many individual genomes were amplified and sequenced, so avoidance of bias is crucial. U-CAN-seq may also be less useful in the study of insertions or deletions, which could cause preferential amplification of the shorter genotype. PCR bias may still occur, causing preferential amplification of one genotype over the other. However, the inoculum and output samples are subject to the same PCR bias, and thus any differences in genotype ratio between inoculum and output should represent true biological changes. Despite these limitations, U-CAN-seq is a powerful tool that can be used by any lab, regardless of institutional sequencing facilities, to understand the relative fitness of RNA viruses.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/v16040636/s1>, Figure S1: Consistent PCR amplification from inocula and supernatants of infected cells; Figure S2: Unix code for k-mer search; Table S1: K-mer search tolerating single-nucleotide mismatch.

**Author Contributions:** Conceptualization, J.D., J.S. and M.T.H.; methodology, J.D. and J.S.; validation, J.D., J.S., C.-K.C. and P.S.; formal analysis, J.D., M.T.F., C.L. and W.S.; software, W.S., C.L. and M.T.F.; investigation, J.D., C.-K.C., J.B. and I.L.; resources, J.S.; writing—original draft preparation, J.D.; writing—review and editing, J.D., M.T.H., M.T.F., C.L., P.S., C.-K.C., J.S. and W.S.; visualization, J.D. and C.-K.C.; supervision, M.T.H. and N.M.; project administration, M.T.H.; funding acquisition, M.T.H. and N.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Institutes of Health Research, grant numbers U19 AI171292 and R21 AI138056. U19 AI171292 supported the methods development described in this study as a tool for testing the fitness of drug resistant viruses, while R21 AI128056 supported the development of the RNA structure mutants characterized in this study. J.S. received support from T32AI007419. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under grant No. DGE-2040435.

**Institutional Review Board Statement:** All work with recombinant viruses performed in this study was conducted under protocols 111755 and 111754 approved by the University of North Carolina Institutional Biosafety Committee on 2 March 2023.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be made available by the authors on request.

**Acknowledgments:** Marcia Sanders assisted with assay planning and plaque assays.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Domingo, E.; García-Crespo, C.; Lobo-Vega, R.; Perales, C. Mutation Rates, Mutation Frequencies, and Proofreading–Repair Activities in RNA Virus Genetics. *Viruses* **2021**, *13*, 1882. [[CrossRef](#)]
2. Mandary, M.B.; Masomian, M.; Poh, C.L. Impact of RNA Virus Evolution on Quasispecies Formation and Virulence. *Int. J. Mol. Sci.* **2019**, *20*, 4657. [[CrossRef](#)]
3. Szemiel, A.M.; Merits, A.; Orton, R.J.; MacLean, O.A.; Pinto, R.M.; Wickenhagen, A.; Lieber, G.; Turnbull, M.L.; Wang, S.; Furnon, W.; et al. In Vitro Selection of Remdesivir Resistance Suggests Evolutionary Predictability of SARS-CoV-2. *PLoS Pathog.* **2021**, *17*, e1009929. [[CrossRef](#)]
4. Prado, J.G.; Franco, S.; Matamoros, T.; Ruiz, L.; Clotet, B.; Menéndez-Arias, L.; Martínez, M.Á.; Martínez-Picado, J. Relative Replication Fitness of Multi-Nucleoside Analogue-Resistant HIV-1 Strains Bearing a Dipeptide Insertion in the Fingers Subdomain of the Reverse Transcriptase and Mutations at Codons 67 and 215. *Virology* **2004**, *326*, 103–112. [[CrossRef](#)]
5. Collins, J.A.; Thompson, M.G.; Paintsil, E.; Ricketts, M.; Gedzior, J.; Alexander, L. Competitive Fitness of Nevirapine-Resistant Human Immunodeficiency Virus Type 1 Mutants. *J. Virol.* **2004**, *78*, 603–611. [[CrossRef](#)]
6. Lee, L.Y.; Zhou, J.; Koszalka, P.; Frise, R.; Farrukee, R.; Baba, K.; Miah, S.; Shishido, T.; Galiano, M.; Hashimoto, T.; et al. Evaluating the Fitness of PA/I38T-Substituted Influenza A Viruses with Reduced Baloxavir Susceptibility in a Competitive Mixtures Ferret Model. *PLoS Pathog.* **2021**, *17*, e1009527. [[CrossRef](#)]
7. Carrasco, P.; Daròs, J.A.; Agudelo-Romero, P.; Elena, S.F. A Real-Time RT-PCR Assay for Quantifying the Fitness of Tobacco Etch Virus in Competition Experiments. *J. Virol. Methods* **2007**, *139*, 181–188. [[CrossRef](#)]

8. Manochewea, S.; Lanxon-Cookson, E.C.; Liu, Y.; Swain, J.V.; McClure, J.; Rao, U.; Maust, B.; Deng, W.; Sunshine, J.E.; Kim, M.; et al. Pairwise Growth Competition Assay for Determining the Replication Fitness of Human Immunodeficiency Viruses. *JoVE J. Vis. Exp.* **2015**, *99*, e52610. [[CrossRef](#)]
9. Neumann, T.; Hagmann, I.; Lohrengel, S.; Heil, M.L.; Derdeyn, C.A.; Kräusslich, H.-G.; Dittmar, M.T. T20-Insensitive HIV-1 from Naïve Patients Exhibits High Viral Fitness in a Novel Dual-Color Competition Assay on Primary Cells. *Virology* **2005**, *333*, 251–262. [[CrossRef](#)]
10. Petrie, S.M.; Butler, J.; Barr, I.G.; McVernon, J.; Hurt, A.C.; McCaw, J.M. Quantifying Relative Within-Host Replication Fitness in Influenza Virus Competition Experiments. *J. Theor. Biol.* **2015**, *382*, 259–271. [[CrossRef](#)]
11. Baker, C.; Shi, P.-Y. Construction of Stable Reporter Flaviviruses and Their Applications. *Viruses* **2020**, *12*, 1082. [[CrossRef](#)]
12. Kanai, Y.; Kawagishi, T.; Nouda, R.; Onishi, M.; Pannacha, P.; Nurdin, J.A.; Nomura, K.; Matsuura, Y.; Kobayashi, T. Development of Stable Rotavirus Reporter Expression Systems. *J. Virol.* **2019**, *93*, e01774-18. [[CrossRef](#)]
13. Ren, L.; Peng, Z.; Chen, X.; Ouyang, H. Live Cell Reporter Systems for Positive-Sense Single Strand RNA Viruses. *Appl. Biochem. Biotechnol.* **2016**, *178*, 1567–1585. [[CrossRef](#)]
14. Suphatrakul, A.; Duangchinda, T.; Jupatanakul, N.; Prasittisa, K.; Onnome, S.; Pengon, J.; Siridechadilok, B. Multi-Color Fluorescent Reporter Dengue Viruses with Improved Stability for Analysis of a Multi-Virus Infection. *PLoS ONE* **2018**, *13*, e0194399. [[CrossRef](#)]
15. Todd, E.V.; Black, M.A.; Gemmill, N.J. The Power and Promise of RNA-Seq in Ecology and Evolution. *Mol. Ecol.* **2016**, *25*, 1224–1241. [[CrossRef](#)]
16. Takele Assefa, A.; Vandesompele, J.; Thas, O. On the Utility of RNA Sample Pooling to Optimize Cost and Statistical Power in RNA Sequencing Experiments. *BMC Genom.* **2020**, *21*, 312. [[CrossRef](#)]
17. Arora, S.; Pattwell, S.S.; Holland, E.C.; Bolouri, H. Variability in Estimated Gene Expression among Commonly Used RNA-Seq Pipelines. *Sci. Rep.* **2020**, *10*, 2734. [[CrossRef](#)]
18. Madden, E.A.; Plante, K.S.; Morrison, C.R.; Kutchko, K.M.; Sanders, W.; Long, K.M.; Taft-Benz, S.; Cruz Cisneros, M.C.; White, A.M.; Sarkar, S.; et al. Using SHAPE-MaP To Model RNA Secondary Structure and Identify 3'UTR Variation in Chikungunya Virus. *J. Virol.* **2020**, *94*, e00701-20. [[CrossRef](#)]
19. Ecke, L.; Krieg, S.; Bütepage, M.; Lehmann, A.; Gross, A.; Lippok, B.; Grimm, A.R.; Kümmerer, B.M.; Rossetti, G.; Lüscher, B.; et al. The Conserved Macrod domains of the Non-Structural Proteins of Chikungunya Virus and Other Pathogenic Positive Strand RNA Viruses Function as Mono-ADP-Ribosylhydrolases. *Sci. Rep.* **2017**, *7*, 41746. [[CrossRef](#)]
20. Gadea, G.; Bos, S.; Krejbich-Trotot, P.; Clain, E.; Viranaicken, W.; El-Kalamouni, C.; Mavingui, P.; Desprès, P. A Robust Method for the Rapid Generation of Recombinant Zika Virus Expressing the GFP Reporter Gene. *Virology* **2016**, *497*, 157–162. [[CrossRef](#)]
21. Funk, A.; Truong, K.; Nagasaki, T.; Torres, S.; Floden, N.; Balmori Melian, E.; Edmonds, J.; Dong, H.; Shi, P.-Y.; Khromykh, A.A. RNA Structures Required for Production of Subgenomic Flavivirus RNA. *J. Virol.* **2010**, *84*, 11407–11417. [[CrossRef](#)]
22. Jaafar, Z.A.; Kieft, J.S. Viral RNA Structure-Based Strategies to Manipulate Translation. *Nat. Rev. Microbiol.* **2019**, *17*, 110–123. [[CrossRef](#)]
23. Robertson, M.P.; Igel, H.; Baertsch, R.; Haussler, D.; Ares Jr, M.; Scott, W.G. The Structure of a Rigorously Conserved RNA Element within the SARS Virus Genome. *PLoS Biol.* **2004**, *3*, e5. [[CrossRef](#)]
24. Simmonds, P.; Smith, D.B. Structural Constraints on RNA Virus Evolution. *J. Virol.* **1999**, *73*, 5787–5794. [[CrossRef](#)]
25. Ranasinghe, D.; Jayadas, T.T.P.; Jayathilaka, D.; Jeewandara, C.; Dissanayake, O.; Guruge, D.; Ariyaratne, D.; Gunasinghe, D.; Gomes, L.; Wijesinghe, A.; et al. Comparison of Different Sequencing Techniques for Identification of SARS-CoV-2 Variants of Concern with Multiplex Real-Time PCR. *PLoS ONE* **2022**, *17*, e0265220. [[CrossRef](#)]
26. Wang, Y.; Zhao, Y.; Bollas, A.; Wang, Y.; Au, K.F. Nanopore Sequencing Technology, Bioinformatics and Applications. *Nat. Biotechnol.* **2021**, *39*, 1348–1365. [[CrossRef](#)]
27. Zheng, P.; Zhou, C.; Ding, Y.; Liu, B.; Lu, L.; Zhu, F.; Duan, S. Nanopore Sequencing Technology and Its Applications. *MedComm* **2023**, *4*, e316. [[CrossRef](#)]
28. Abraham, R.; McPherson, R.L.; Dasovich, M.; Badiie, M.; Leung, A.K.L.; Griffin, D.E. Both ADP-Ribosyl-Binding and Hydrolase Activities of the Alphavirus nsP3 Macrod domain Affect Neurovirulence in Mice. *mBio* **2020**, *11*, e03253-19. [[CrossRef](#)]
29. Park, E.; Griffin, D.E. The nsP3 Macro Domain Is Important for Sindbis Virus Replication in Neurons and Neurovirulence in Mice. *Virology* **2009**, *388*, 305–314. [[CrossRef](#)]
30. Zhang, S.; Garzan, A.; Haese, N.; Bostwick, R.; Martinez-Gzegozewska, Y.; Rasmussen, L.; Streblow, D.N.; Heise, M.T.; Pathak, A.K.; Augelli-Szafran, C.E.; et al. Pyrimidone Inhibitors Targeting Chikungunya Virus nsP3 Macrod domain by Fragment-Based Drug Design. *PLoS ONE* **2021**, *16*, e0245013. [[CrossRef](#)]
31. Adams, R.H.; Brown, D.T. BHK Cells Expressing Sindbis Virus-Induced Homologous Interference Allow the Translation of Nonstructural Genes of Superinfecting Virus. *J. Virol.* **1985**, *54*, 351–357. [[CrossRef](#)] [[PubMed](#)]
32. Johnston, R.E.; Wan, K.; Bose, H.R. Homologous Interference Induced by Sindbis Virus. *J. Virol.* **1974**, *14*, 1076–1082. [[CrossRef](#)] [[PubMed](#)]
33. Karpf, A.R.; Lenches, E.; Strauss, E.G.; Strauss, J.H.; Brown, D.T. Superinfection Exclusion of Alphaviruses in Three Mosquito Cell Lines Persistently Infected with Sindbis Virus. *J. Virol.* **1997**, *71*, 7119–7123. [[CrossRef](#)] [[PubMed](#)]

34. Singer, Z.S.; Ambrose, P.M.; Danino, T.; Rice, C.M. Quantitative Measurements of Early Alphaviral Replication Dynamics in Single Cells Reveals the Basis for Superinfection Exclusion. *Cell Syst.* **2021**, *12*, 210–219.e3. [[CrossRef](#)]
35. Singh, I.R.; Suomalainen, M.; Varadarajan, S.; Garoff, H.; Helenius, A. Multiple Mechanisms for the Inhibition of Entry and Uncoating of Superinfecting Semliki Forest Virus. *Virology* **1997**, *231*, 59–71. [[CrossRef](#)]

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