Apr 30, 2024

# Universal Competition Assay by Nanopore Sequencing (U-CAN-Seq)

DOI
**dx.doi.org/10.17504/protocols.io.kqdg3xjdzg25/v1**

Jennifer Loome[1], John Sears[1], Wes Sanders[1], Che-Kang Chang[1], Paul Sylvester[1], Colton Linnertz[2], Nathaniel Moorman[1], Martin T. Ferris[2], Mark T. Heise[2,1]

[1]Department of Microbiology and Immunology, University of North Carolina at Chapel Hill;
[2]Department of Genetics, University of North Carolina at Chapel Hill

Jennifer Loome
University of North Carolina at Chapel Hill



**DOI: dx.doi.org/10.17504/protocols.io.kqdg3xjdzg25/v1**

**Protocol Citation:** Jennifer Loome, John Sears, Wes Sanders, Che-Kang Chang, Paul Sylvester, Colton Linnertz, Nathaniel Moorman, Martin T. Ferris, Mark T. Heise 2024. Universal Competition Assay by Nanopore Sequencing (U-CAN-Seq). **protocols.io https://dx.doi.org/10.17504/protocols.io.kqdg3xjdzg25/v1**



**Protocol status:** Working
**We use this protocol and it's working**

**Created:** February 20, 2024

**Last Modified:** April 30, 2024

**Protocol Integer ID:** 95492

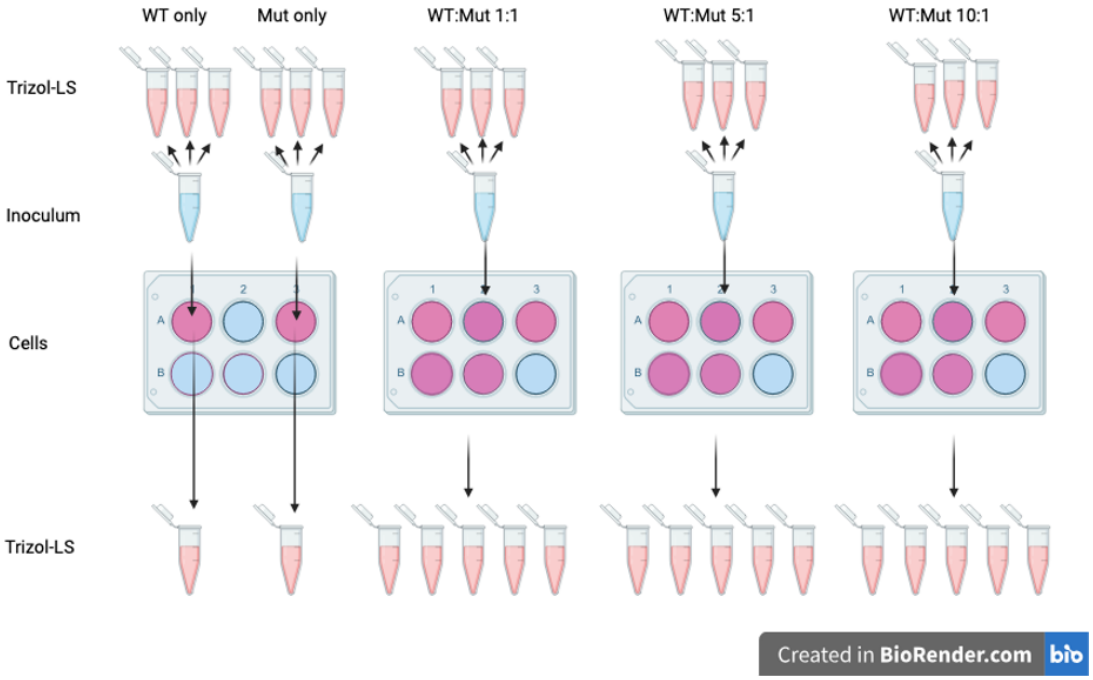
**Funders Acknowledgement:**
**National Institutes of Health**
**Grant ID: R21 AI138056**
**National Institutes of Health**
**Grant ID: U19 AI171292**

## Abstract

Competition assays are an effective and rigorous method that can be used to understand the relative fitness of different viral genotypes. In competition assays, two genotypes of virus are used to infect the same cell culture well or animal, replication occurs, and the relative abundance of each genotype is measured to determine the relative fitness of the two viral genotypes. However, it is technically challenging to distinguish between and quantify very similar genotypes of a virus. Here, we describe a protocol combining RT-PCR with newly available commercial nanopore sequencing services to perform competition assays on RNA viruses. Our assay, Universal Competition Assay by Nanopore Sequencing (U-CAN-seq), is relatively inexpensive and sensitively detects fitness differences between two similar genotypes of a virus. Competition assays using the U-CAN-seq approach can be divided into four parts: infection, RNA extraction, RT-PCR, and sequencing and analysis. First, we infect cells at a multiplicity of infection (MOI) of 0.01 using each virus genotype individually and different ratios of the two genotypes. We sample the inocula before infection and the cell supernatants at 24 hours post infection (hpi), extract RNA using Trizol-LS, and then use gene-specific primers to generate cDNA. We use PCR to amplify the region of the viral genome that differs between the two genotypes, and we submit the linear amplicons for commercial nanopore sequencing. Finally, we developed code to provide a simplified method to search for strings of sequence unique to each genotype and calculate the ratio of the genotype of interest to the total read counts at that locus.

## Guidelines

- Here, we describe a protocol for competition assays using the vaccine strain of chikungunya virus (CHIKV) 181/25 in Vero81 cells. We have also used this protocol successfully for MRC-5, C6/36, and THF cells, and in mice with different CHIKV strains (unpublished data). Different cell lines require slightly different conditions that should be optimized before the competition assay is attempted. We anticipate that the assay will work well with non-alphaviruses, but that different viruses may require different harvest times.
- This assay involves mixing two genotypes of a virus in different ratios and then measuring the genotype ratios in the inoculum and in the supernatants of infected cells. We recommend measuring titers of virus stocks side by side, soon before performing the competition assay, for the most accurate ratios of viruses in your inocula.
- We recommend measuring RNA from the output supernatants in order to capture virus that has successfully completed its replication cycle and exited the cell. However, RNA yields from supernatants are quite low. They therefore require Trizol extraction, rather than column purification, and resuspension in a very low volume. One could reasonably harvest RNA from infected cells instead, and therefore use column purification kits for ease of use, but would therefore measure RNA that has not successfully escaped the cell. In that case, the researcher might fail to detect fitness defects that affect RNA packaging and virus egress.
- We recommend using the Trizol-LS (liquid solution) reagent rather than the regular Trizol reagent. The Trizol-LS reagent is more concentrated and better for use with liquid solutions, especially to extract RNA from low concentration solutions. It is possible to mix 300 µl of sample with 900 µl Trizol LS for a total volume of 1200 µl, but it is only recommended to mix 120 µl of sample with 1080 µl of Trizol for a total volume of 1200 µl. Thus, with Trizol-LS, we can extract more RNA per sample than with Trizol.
- We use a simple line of code in the terminal application of Mac computers to search sequencing files for strings of nucleotides for the genotypes of interest. Searching for a string of nucleotides 10-40 bp (a k-mer) in length means we only count high-quality reads with no incorrect bases near the region of interest, so our total reads counted at our locus is often lower than the total read depth at the locus of interest. Our method works best with genotypes that differ by at least two nucleotides within the 10-40 nucleotide k-mer, because it eliminates noise introduced by random errors at single nucleotides in the string.
- This assay is appropriate to detect fitness differences between two stable genotypes of virus. It is not appropriate to study a genotype that rapidly acquires compensatory second-site mutations or reverts to wild-type sequence. This assay cannot distinguish between a mutant with a strong fitness disadvantage compared to wild-type and a mutant that rapidly mutates to revert to wild-type sequence.
- Sampling scheme: See diagram below. We sample each inoculum 3 times and extract RNA from each sample to account for variation in RNA extraction and sequencing. We infect 5 independent wells with each genotype ratio of virus (5 biological replicates) and harvest and extract RNA from each to account in variation from replication, RNA extraction, and sequencing.




- We use 1:1, 5:1 and 10:1 ratios of wild-type to mutant virus. Researchers can modify these ratios to fit their research questions.
- Handle RNA on ice unless otherwise noted.
- Note on primer design: Design two standard 18-29 nucleotide PCR primers with similar annealing temperatures around your region of interest (the k-mer you will search for). Make sure the primers are each at least 50 nucleotides away from the k-mer. You can use the same reverse primer for both the PCR and the reverse transcription reaction, or use nested primers for added specificity, with the PCR reverse primer more 5' than the reverse primer for reverse transcription.

## Materials

**Equipment:**

- Refrigerated microcentrifuge
- -20°C freezer
- -80°C freezer

| | A | B | C |
|---|---|---|---|
| | **Item** | **Brand** | **Reference** |
| | DMEM | Gibco | 11965092 |
| | FBS | R&D systems | S11150 |
| | L-glutamine | Gibco | 25030081 |
| | 0.05% Trypsin-EDTA | Gibco | 25300054 |
| | 1X DPBS | Gibco | 14040133 |
| | Trypan blue | Gibco | 15250061 |
| | 6 well plates | Corning | CLS3506 |
| | Trizol LS | Invitrogen | 10296010 |
| | Isopropanol | Sigma | 19516-500ml |
| | Ethanol | Sigma-Aldrich | E7023 |
| | Glyco-blue Co-precipitant | Invitrogen | AM9515 |
| | Phusion polymerase | Thermo Scientific | F-530XL |
| | Primers | Thermofisher | |
| | Moclecular biology grade water | Sigma | W4502-1L |
| | 1.5 ml tubes | Costar | 3620 |
| | PCR strip tubes | Eppendorf | 951010022 |

## Before start

- Titer your virus stocks side-by-side. If you have not titered them in more than six months, re-titer them to ensure accurate MOI calculations.
- Label sample tubes and fill with 900 µl of Trizol-LS.

## Infection

1 **Plate cells**

Plate Vero81 cells in six 6-well plates at 2*10^5 cells/well 24 hours

2 **Infect cells**

2.1 Count cells by trypsinizing 1 well in 400 µl 0.05% trypsin-EDTA, putting in the incubator until cells detach, adding 600 µl of medium and resuspending. Mix 10 µl of cell mixture with 10 µl of trypan blue and count with a hemocytometer.

# cells/#squares counted x 2 dilution factor x 10^4 = #cells/ml = #cells/well.

2.2 Calculate your inocula such that for each condition you have 3 ml of inoculum total. You want to inoculate with 400 µl/well, at an MOI = 0.01 total. (For example, for a 1:1 ratio, WT MOI= 0.005 and mutant MOI=0.005.) Make inocula forthe following conditions:
WT only
Mutant only
WT: Mutant 1:1
WT: Mutant 5:1
WT: Mutant 10:1

2.3 Remove media from cells and infect 5 wells of each condition by adding 400 µl of inoculum. Return the plate to the incubator and return every 15 minutes for 1 hour to rock the plates to ensure even
distribution of infectious particles. Once 1 hour has passed, add 3mL of medium to each well.
*NOTE: You only have to do one well each of WT only, mutant only, and mock. They are controls that will not be used statistically. These wells act as a control to ensure that cross-contamination has not occurred.

2.4 While the plates are rocking, harvest the leftover inoculum by moving 300 µl of each inoculum to 3 tubes with 900 µl of Trizol-LS reagent. Freeze overnight.

3 **Harvest supernatant**

24 hrs post infection, harvest each well once by taking 300 µl of supernatant and putting in a tube with 900 µl Trizol-LS reagent. Let incubate at least 10 min at room temp. Proceed or freeze at -20 for up to 1 month, or at -80 for up to a year, or proceed directly to RNA extraction.

## RNA extraction

4 Add 240 µl of chloroform to each tube with Trizol-LS and supernatant. Mix the tubes by shaking/vortexing and incubate 2-3 minutes at room temp and then spin 15 min 12,000g 4°C.

5 Move the aqueous phase to a new tube with 1 µl of glycogen or glycol-blue co- precipitant. Add 0.6 ml isopropanol and vortex or shake to mix. Incubate overnight at - 20.
NOTE: Do not do overnight at -80°C.

6 Spin 10 min 12000g 4°C to precipitate. You should see a pellet in the tube.

7 Remove the supernatant and wash with 1.2 ml 75% ethanol, vortexing to dislodge the pellet.
NOTE: You can pause here and freeze at -80°C indefinitely.

8 Spin again at 7500g 4°C, and remove the supernatant.
NOTE: Sometimes it can be hard to really get all the ethanol off, especially if some clings to the walls of the tube. If so, you can remove most of the ethanol, pop spin the tube for a few seconds to bring ethanol down, and then remove the rest of the ethanol with a small pipette.

9 Let the pellets air dry until slightly translucent. Then resuspend the pellet in 12 µl water by pipetting to mix, put at 55°C for ten minutes, return to ice, pipette to mix, and nanodrop.

10 Proceed, or freeze RNA at -80.

## Reverse Transcription and PCR

11 **Reverse transcription**
Use the Superscript III kit to perform reverse transcription using a gene-specific primer. Use 8 µl of RNA from the purification in a superscript III RT reaction as described in the manufacturer's protocol, except let the reaction run for 3 hrs instead of 1 hr.

11.1 Mix RNA, primer, and dNTPs. Heat mix 1 to 65°C five min, return to ice at least 1 min.

| | A | B |
|---|---|---|
| | Mix 1 | 1x |
| | Reverse primer 2 µM | 1 |
| | 10 mM dNTP | 1 |
| | RNA | 8 |

11.2 Add 10 µl mix 2/tube, heat to 50 °C 3 hrs, inactivate 85°C 5 min, infinite hold at 12°C. Proceed to PCR or freeze at -20.

| | |
|---|---|
| Mix 2 | 1x |
| 10x RT buffer | 2 |
| 25 mM MgCl2 | 4 |
| RNAse out | 1 |
| 0.1 M DTT | 2 |
| Superscript III | 1 |

12 **PCR:** Use 1 µl of each RT reaction as a template, regardless of concentration, to amplify the genome region in which your two genotypes of virus differ. Below is an example PCR protocol, but you will need to optimize your PCR conditions for your specific primers. Always include a no-template control and positive control (plasmid containing the sequence of interest.) When you are optimizing new primers, it may also be useful to isolate RNA from one set of mock-infected supernatant to use as another negative control.

| | 1x |
|---|---|
| GC buffer | 4 |
| 10 mM dNTP | .4 |
| Primer 1 10 µM | 1 |
| Primer 2 10 µM | 1 |
| DMSO | 1 |
| Template | 1 |
| Water | 11.4 |
| Phusion | 0.2 |

Lid: 105°C

Volume: 20 µl

1. 95°C, 3:00
2. 95°C, 0:30
3. 56°C, 0:30
4. 72°C, 3:00
5. GOTO step 2, 34x
6. 72°C, 5:00
7. 12°C, infinite hold

13 Run a gel: Add 4 µl of 6x loading buffer to each sample and load 4 µl of the mixture on a 1% agarose gel. Run for 1 hr at an appropriate voltage for your gel rig. Stain in EtBr to visualize bands. You should see a band at your target size.

## Analysis

14 **Analyze results: search for k-mer exact match**

Use grep search in terminal to search forstrings. Use the following example code:
for file in /Users/jenloome/Desktop/Desktop_Jennifer's_MacBook_Air/Notebooks_and_backups/0148_AV_Competition_assay_growth_curve_structure_scr_3/Loome_p9g_raw_reads/*.fastq; do { echo "$file"; echo "WT,"$(grep -o "ACGTGGACATAGACGCTGACAGCGCCTTTTT" "$file" | wc -l); echo "scrSL3,"$(grep -o "ATGTAGACATAGACGCTGACAGCGCCTTTCT" "$file" | wc -l); } >> 0148_AV_scrSL3_output.csv; done

Edit the following:
-Directory where your fastq files are at (on mac is probably /User/UserName/Downloads/PlasmidsourausRawReadDirectory/*.fastq), can be found in terminal by typing "pwd" while in that directory
-Name of string one (appears in the .csv file)
-First k-mer string that you are searching for
-Name of string two (appears in the .csv file)
-Second k-mer string that you are searching for

15 **Alternative analysis method: Search for k-mer tolerating mismatch.**

The code below functions as above to search for k-mers of interest, counting both exact matches and matches that tolerate a single nucleotide mismatch. This analysis method could be advantageous to improve read count but disadvantageous by introducing noise to the data. In our hands, both methods of analysis have yielded similar results

**#UNIX CODE**

```
#output CSV file with the following columns
echo "File,Count_Exact_String,Count_1_Mismatch_String" > Name_of_output.csv

#directory containing .fastq files to be searched
directory="/example/directory/fastqfiles"

# nucleotide string to search for
string="NucleotideString"

#Loop through each .fastq file in the directory
for file in "$directory"/*.fastq; do
  # Extract the filename from the path
  filename=$(basename "$file")

  # find occurrences of the string with up to 1 mismatch and count both
  extracted_strings=$(awk -v pattern="$string" 'BEGIN{RS="@";ORS=""}
    {
      # Check if the exact string is found
      i=index($2, pattern)
      if (i) {
        count_exact++
      } else {
        # Check for up to 1 mismatch
        for (j=1; j<=length(pattern); j++) {
          if (substr($2, j, 1) != substr(pattern, j, 1)) {
            if (++mismatches > 1) break
          }
        }
        if (mismatches <= 1) {
          count_mismatch++
        }
      }
      mismatches=0
    }
    END {
      # Print the counts as CSV
```

```
            print count_exact "," count_mismatch ","
        }' "$file")

        # output counts to the CSV (same file as above)
        echo "$filename,$extracted_strings" >> Name_of_output.csv
    done
```

## Protocol references

Extract RNA following the Trizol-LS protocol (https://tools.thermofisher.com/content/sfs/manuals/Trizol_ls_reagent.pdf).