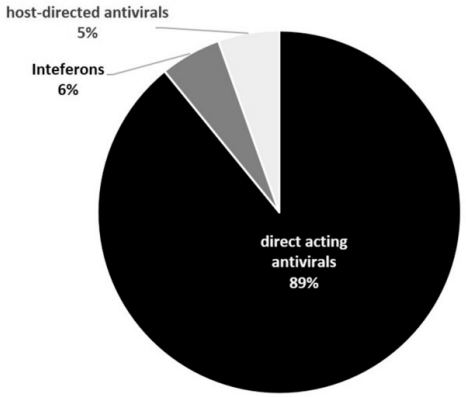
Screening viral host dependency factors of HIV via functional genomics *in silico* and *in vitro* for drug targeting

Rubendren Jamilchelvan[1,2], Riley H Tough[1,2],Michelle Perner[1,2], Xia Liu[3], Eric Enns[3], Paul J McLaren[1,2]

[1]Department of Medical Microbiology and Infectious Diseases, University of Manitoba,
[2]National HIV and Retrovirology Laboratory at the JC Wilt Infectious Diseases Research Centre, National Microbiology Labs, Public Health Agency of Canada,
[3]Bioinformatics Group, National Microbiology Labs, Public Health Agency of Canada

Background

- There is a lack of broad-acting and host-directed antivirals for HIV and other viruses
- 5% of the antivirals in the current market are host-directed and even fewer are broad-acting for more than one virus [1] (Figure 1).
- Previous studies have also shown that generally, the mutation rate in the genome is inversely proportionate to the genome size. Hence, viruses like HIV have a tendency of having a higher mutation rate than humans [2] (Figure 2) .
- When using host-directed antivirals, we could lessen the selective pressure on the virus. This may help reduce unwanted escape mutations that can render our antiviral useless



HDAs	Mechanism of action	Approved use
Imiquimod	TLR-7 agonist	HPV
Docosanol	Membrane	HSV
Ribavirin	Multiple (IMPDH, guanosine analog, immune-modulation)	HCV (other RNA and DNA viruses)
Maraviroc	CCR5 chemokine receptor	HIV (CMV)
Sinecatechins	Multiple mechanisms	HPV
Interferons	interferon alpha-2b, interferon alpha-n3, interferon alfacon-1, peginterferon Alfa-2B, and peginterferon Alfa-2A (biologics)	HCV (and HBV)

Figure 1: The proportion of antivirals found and examples of host-directed antivirals [1]

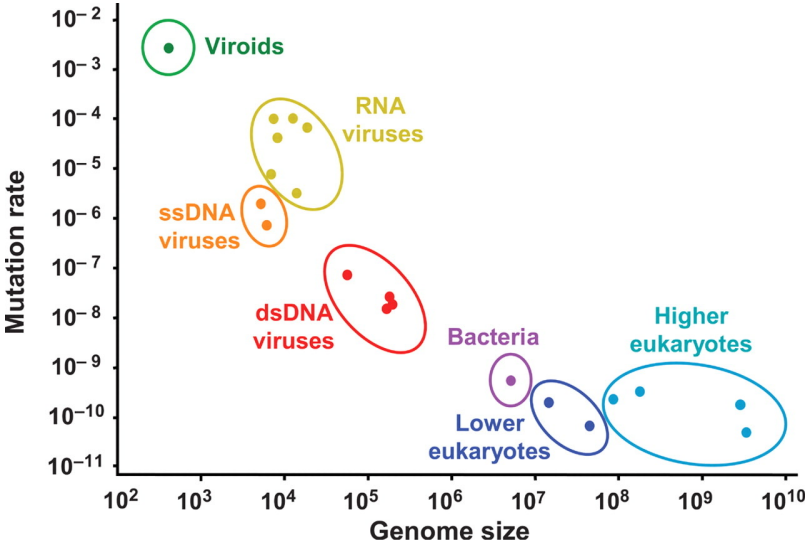


Figure 2: The relationship between mutation rate versus genome size [2]

Hypothesis

- Viruses are obligate intracellular pathogens that require many host cell components to complete their life cycles.
- Several genome-wide screens have been performed across multiple viral models that have identified hundreds of these viral host dependency factors (HDFs).
- These HDFs may be good candidates to develop novel host-directed antivirals.
- Defining which HDF may make good targets and which HDF may lead to drug toxicity is challenging. Intersecting genes found to be viral HDFs and also not essential for host function may open new avenues for the development of therapies
- **Hypothesis: Viral host dependency factors (HDFs) that can tolerate loss-of-function polymorphisms in healthy humans, may be targeted to induce antiviral effect**

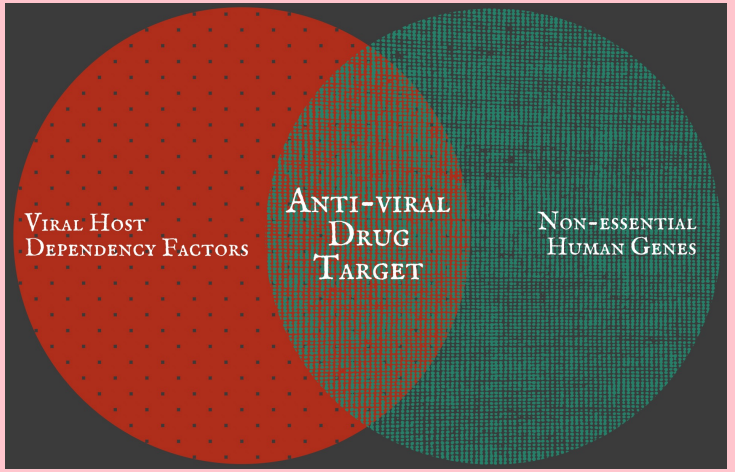


Figure 3: Visual representation of the hypothesis

Objectives

- **Pooling HDFs** - Performed a literature review of available whole-genome screens of HDFs for multiple viral models. This was followed by bioinformatic analyses of key genes and pathways that intersect between HIV and other viruses
- **HDF Essentiality** - Analyse genes and pathways identified by *in silico* screens from Objective 1 if they translate into non-essential host proteins on gnomAD (genome aggregation database)
- **HDF knockout** – Perform CRISPR gene knockouts on candidate genes and perform infection assays to gauge the change in rate of infection
- **Confirming known mutation** – The Confirmed HDF knockouts will be analysed further. Observed SNPs found on gnomAD are to be induced via CRISPR to mimic the naturally occurring SNPs to see if they lead to a successful loss-of-function mutation

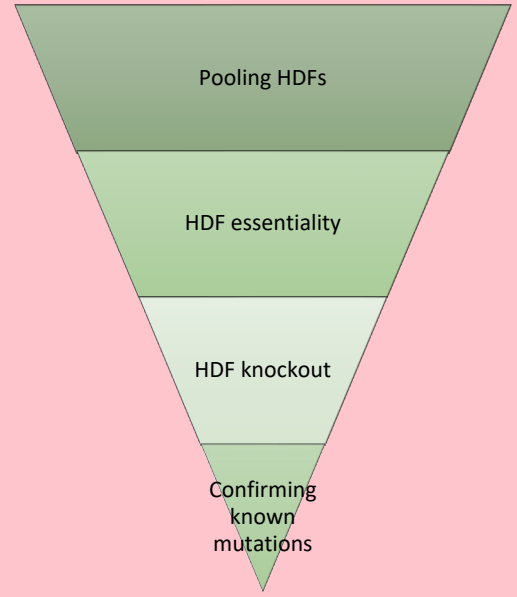


Figure 4: Visual representation of the objectives

Methods

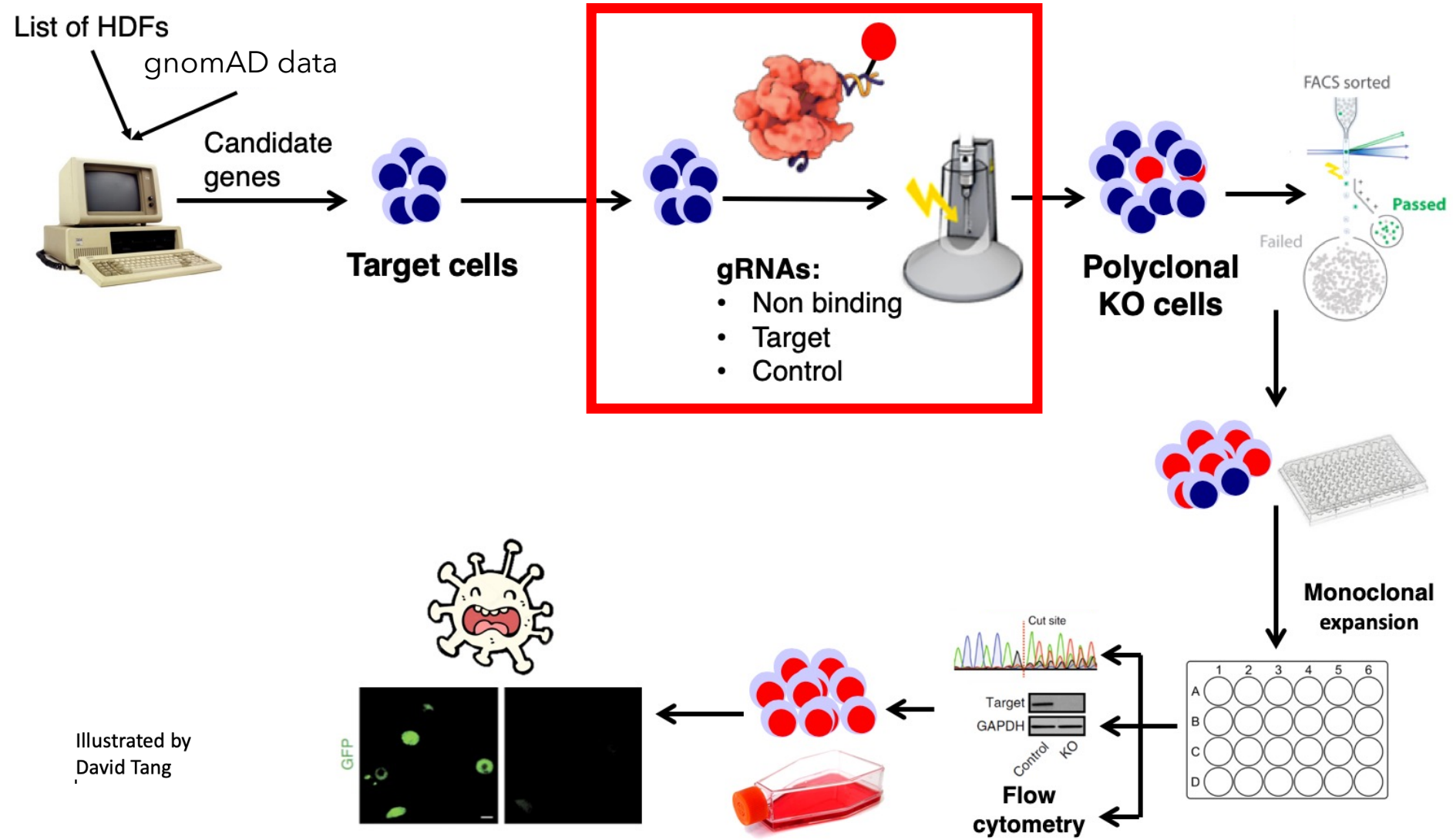


Figure 5: The main methods from objective 1 to objective 4. The red box represents the current step that is being done (Objective 3). Objectives 1 and 2 have been completed and a manuscript is being prepared for publishing.

Results

- A total of 27 papers performing whole-genome screens were found across 10 virus models – HIV, HCV, HDV, SARS-CoV-2, SARS-CoV, EBOV, IAV, ZIKV, DENV and WNV. From this, a total of 320 unique genes were seen to be overlapping across more than one virus (>10%)

Virus	Number of papers
HIV	6
SARS-CoV-2	2
SARS-CoV	1
Influenza A	8
Hepatitis C	5
Hepatitis D	1
Ebola	1
Zika	1
Dengue	1
West Nile	1
Sum	27

HDFs	Number of viruses	Percentage (%)
2928	1	90.15
268	2	8.25
43	3	1.32
9	4	0.27

Unique overlapping genes = **320**

Figure 6: The summary of all the HDF genome-wide screens found and the number of unique HDFs overlapping with more than one virus

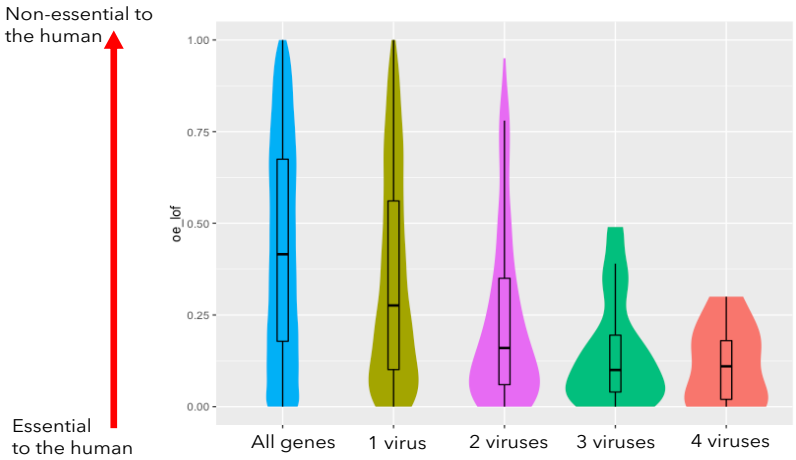


Figure 7: The distribution of gene tolerances against deleterious mutations. An O/E score of 1 is least conserved and a score of 0 is highly conserved

- Utilising 3 online databases (David Bioinformatics Resources, StringDB and Reactome) we determined that the most hijacked biological pathway by all the 10 viruses was **phagosome acidification** (Figure 8).

- From gnomAD, we also observed that as an HDF is being shared by more than one virus, it is more conserved by the body and is very intolerant to a deleterious mutation. This was determined by analysing all the O/E scores found on gnomAD (Figure 7).

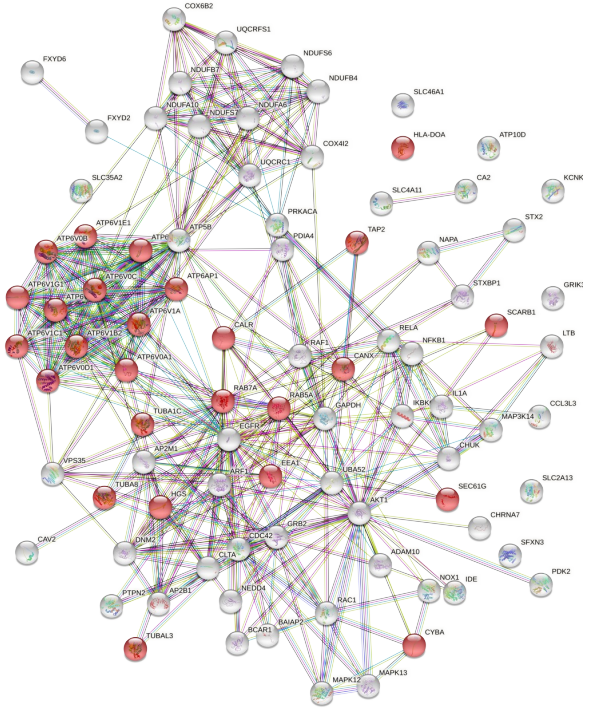


Figure 8: Enriched HDFs from HIV, HCV, HDV, EBOV and ZIKV involved in the phagosome acidification pathway

Conclusion - Candidate HDFs

- By intersecting the HDFs with the putative non-essential human genes found on gnomAD (as per the hypothesis), we narrowed down 6 HIV HDFs that are overlapping with one other virus.

Top HIV HDF candidates	Viruses sharing the HDF	Loss of function variant from gnomAD	Type of loss of function mutation	Number of Homozygous Loss of function individuals
RABEPK	HIV + Hepatitis C	p.Pro135HisfsTer44	Frameshift	3
KRBA2	HIV + Ebola	p.Arg183Ter	Stop gained	1
PI4KA	HIV + Hepatitis C	p.Leu1952CysfsTer45	Frameshift	1
MYEF2	HIV + Hepatitis D	c.1138+1G>T	Splice donor	1
USP6	HIV + West Nile	p.Arg522LysfsTer12	Frameshift	5
ERN2	HIV + Influenza A	p.Cys694TrpfsTer8	Frameshift	2

Figure 8: Six candidate genes that are HDFs for HIV and one other virus. The deleterious SNP found in gnomAD is shown for each gene and the number of putative homozygous loss of function individuals found

- As per objective 3, we are currently performing CRISPR gene knockouts on these genes (Figure 8). Once we successfully generate viable and stable monoclones with the preferred knockouts, we will perform HIV infection assays. If the infection drops, we can conclude that the gene was in an HDF for HIV with a higher confidence.

Future directions

- The confirmed HDFs will be further tested in Objective 4. We will induce the known SNP (loss-of-function variant) found on gnomAD and determine if the SNPs would lead to a loss-of-function mutation (Figure 8).
  - Infection assays will be done to test infection rates.
  - Western blots/flow cytometry will be done to confirm the presence of the protein after inducing the SNP.
- With future collaborations, we will analyse the infectivity of the overlapping virus as well. When both viruses are affected, we can be certain that the direct-acting antiviral target would also be a broad-acting target for at least one other virus.
- This study would also be tested *in vivo* in the long term.

References

1. Chitalia, V. C. & Munawar, A. H. A painful lesson from the COVID-19 pandemic: the need for broad-spectrum, host-directed antivirals. J Transl Med 18, 390 (2020).
2. Gago, S., Elena, S. F., Flores, R. & Sanjuán, R. Extremely high mutation rate of a hammerhead viroid. Science 323, 1308 (2009).

Acknowledgements

- **Thank you to the McLaren lab, the Bioinformatics Group and my committee members. This project is being funded by the GRDI. Contact information: rjamilch@myumanitoba.ca**

Rady Faculty of Health Sciences | UM

Public Health Agency of Canada

JC WILT INFECTIOUS DISEASES RESEARCH CENTRE

Genomics R&D Initiative