# Functional Genomic and Image-Based Screening Approaches for Probing Host-Pathogen Interactions

by

Rebecca J. Carlson

B.S., Chemical Engineering

Michigan State University, 2017

Submitted to the Harvard-MIT Program in Health Sciences and Technology, in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Medical Engineering and Medical Physics

at the

Massachusetts Institute of Technology

June 2023

Authored by: Rebecca J. Carlson
Harvard-MIT Program in Health Sciences and Technology
May 19, 2023

Certified by: Paul C. Blainey
Department of Biological Engineering
Thesis Supervisor

Certified by: Nir Hacohen
Harvard Medical School
Thesis Supervisor

Accepted by: Collin M. Stultz, MD, PhD
Director, Harvard-MIT Program in Health Sciences and Technology
Nina T. and Robert H. Rubin Professor in Medical Engineering and Science
Professor of Electrical Engineering and Computer Science

# Functional Genomic and Image-Based Screening Approaches for Probing Host-Pathogen Interactions

by

Rebecca J. Carlson

Submitted to the Harvard-MIT Program in Health Sciences and Technology on May 19, 2023,
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Medical Engineering and Medical Physics

## Abstract

Host-pathogen interactions represent a complex interplay between hosts and pathogens that can evolve over millions of years. Interactions between bacteria or viruses and human cells, and the resulting evolved antipathogenic signaling pathways, are processes responsible for pathologies ranging from infectious diseases to autoimmune conditions and cancer. In addition, engineered designs inspired by pathogen interactions with hosts are increasingly being used to both treat and diagnose many pathologies that need not originate from infection with a pathogen. Therefore, it is critical to build and deploy scalable tools to better understand host-pathogen dynamics in order to both better treat conditions where pathogens or antipathogenic signaling contribute directly to disease pathology as well as to engineer new treatments to address a broader range of disease states.

In this thesis, I describe approaches to leverage functional genomics and image-based screening to perturb and profile host-pathogen interactions, including responses to two RNA viruses, Sendai virus and Ebola virus. These provide case studies highlighting the utility of high-content image-based screening for revealing new genes regulating predefined phenotypes of interest as well as for generating single-cell imaging profiles that can be used to infer new genetic functions and phenotypic states directly from screening data without *a priori* specification. I also highlight an example of a genetic screen that revealed a robust negative result, leading to hypothesis and validation of a novel function of the STING protein as a proton channel.

Thesis Supervisor: Paul C. Blainey
Title: Associate Professor of Biological Engineering, Massachusetts Institute of Technology

Thesis Supervisor: Nir Hacohen
Title: Professor of Medicine, Harvard Medical School

## Acknowledgments

While the PhD journey has certainly not always been easy, the past nearly six years represent the most rewarding and transformative period of my life so far. There are many who have contributed to both preparing me for graduate school and helping me throughout this process.

First, I would like to thank my advisors, Paul Blainey and Nir Hacohen. Paul, your optimism and belief in me have given me confidence to pursue challenging projects and work through tough times. Nir, your deep knowledge about immunology and science in general always inspire me to read and think more deeply and I admire the respect for others that you display so consistently.

I am also grateful to members of the Blainey and Hacohen labs. In the Blainey lab, David Feldman and Avtar Singh first developed and introduced me to optical pooled screening (better known as "lasagna"). I am grateful to Avtar for taking the time to mentor me during my rotation and always responding to my questions and random ideas at all times of the day and night, even now that he is no longer in the lab. Thanks to Anna Le for your important feedback on project directions, figures, and lab safety but most importantly for being a friend who has always been there for me when I needed a listening ear. Thanks to Russell Walton for wide-ranging discussions on everything from science to weight training and occasionally actually going to the gym together rather than just talking about it. Thanks to Michael Leiken, the first research associate I had the opportunity to mentor, for your perseverance and patience as I learned to become a better mentor. Thanks to Brian Soong and Jake Qiu for many memorable conversations and encouragement over the years. To the rest of "team lasagna": Luke Funk, Owen Andrews, Yue Qin, Emil Hertz, and Witold Postek - thanks for working on this challenging new technology together. And to Emily Botelho, thanks for your consistent encouragement and support over the years. In the Hacohen lab, thanks to John Ray for mentoring me during my rotation and taking the time to organize an immunology journal club early on in graduate school. Thanks to Bo Li and Bing Shao Chia for help getting started on virology experiments. And thanks to Matteo Gentili and Bingxu Liu. Collaborating with you has been one of the highlights of my graduate career. You both inspire me to be a better scientist and I admire the combination of rigor and creativity you bring to your work. I sincerely hope we have the opportunity to work together again in the future. In particular, the work in Chapter 4 was co-led with Bingxu, who in addition to being a brilliant scientist was an incredibly fun and energizing collaborator on this project. The work on Ebola virus in Chapter 3 was done in collaboration with Robert Davey and Caroline Uhler's groups and Justin Patten in the Davey group led the Ebola virus infection work, I am grateful for his partnership in working on that exciting project. George Stefanakis and Adit Radhakrishnan in

Caroline Uhler's group led the deep learning model training for the Ebola project and helped give me new perspectives on analysis that I have applied to multiple projects.

Many others at the Broad have contributed to a wonderful graduate school experience but I want to especially thank David Foisy, Spike Kitowicz, Hibah Aijaz, and Liliana Jimenez for your help and great attitude in getting valuable reagents when needed. Thanks to my committee: Jonathan Weissman, Jon Kagan, and Peter Sorger for giving me useful feedback and encouragement at critical times and for making time to meet in your very busy schedules. I would also like to thank the NSF GRFP and Fannie & John Hertz Foundation for generously supporting my PhD work and providing me the flexibility to pursue exciting projects.

I have been fortunate to have met incredible scientific mentors even before my time in graduate school. Thanks to Phil Angart and S. Patrick Walton at MSU for taking a chance on a freshman who didn't know her way around a research lab. I wouldn't have pursued a PhD without your support and example of how fulfilling a career in science can be. Thanks also to Sinu John, Sharat Vayttaden, and Iain Fraser at the NIH for introducing me to immunology and imaging during a summer research program.

To my friends from undergrad and graduate school, I am grateful to more than I could name here. I would like to especially thank my friends from MSU Chemical Engineering – Ariel Rose, Sarah Thorwall, Mark Main, Henry Pan, Alafia Kanpurwala, Fatin Pathi, and Stacie Schroll. Thanks to Michelle Ramseier in HST for always being there when I have needed a friend. Thanks to Bianca Lepe, Halston Lim, and Laura Zhang for teaching me so much (and having fun along the way) through our time together in the GSC DEI committee.

And finally, thanks to my parents for supporting me in my journey to become a scientist. You made many sacrifices along the way, including missing time together as a family as I moved far away to pursue my education during high school. To my grandparents, Wendell and Rosemary Carlson, you were brave enough to host a teenager in your home for a few years so I could continue my education in the US – thank you for your constant love. To my mother- and father-in-law, Patricia Kyeremateng Berchie and Joseph Nketiah Berchie, and my sister-in-law Abena Serwaa Afum-Adjei Awuah, thank you for welcoming me into your family and always cheering me on during graduate school. To my husband Kwasi, you are the kindest person I have ever met and a brilliant scientist. Thank you for challenging me to become a better person every day.

# Table of Contents

# List of Figures

# List of Supplementary Figures

# Introduction

## 1.1 Host-Pathogen Interactions

### *1.1.1 Human Pathogens*

Human pathogens are organisms, ranging from viruses and bacteria to fungi and helminths, that can trigger signs and symptoms of disease in humans. Viruses are a particularly notable class of pathogens, as they represent approximately two-thirds of new human pathogens (Sudhan & Sharma, 2020) and are the dominant source of emerging epidemics and pandemics. High viral mutation rates also facilitate development of drug resistance, making new approaches to study interactions of viruses with human hosts critical for development of next-generation antiviral therapeutics. While all viruses rely partially on host cells for replication, viral genomes are diverse and can be composed of RNA or DNA, can be single-stranded, or double-stranded, and, among single-stranded RNA (ssRNA) viruses, positive-sense or negative-sense; in addition, retroviruses such as HIV have RNA genomes that replicate via a DNA intermediate (Durmuş & Ülgen, 2017).

### *1.1.2 Human Innate Immune Responses to DNA and RNA*

Human innate immune responses to foreign DNA and RNA rely on pattern recognition receptors (PRRs) that recognize microbe-derived products termed pathogen-associated molecular patterns (PAMPs). PRRs include toll-like receptors (TLRs) (Fitzgerald & Kagan, 2020) that recognize both non-nucleic-acid-derived pathogen products as well as dsRNA (TLR3), ssRNA (TLR7/8), and CpG DNA (TLR9) derived from pathogens. While TLRs recognizing RNA and DNA are localized to endosomes, other intracellular sensors can detect cytosolic nucleic acids. DNA sensors include absent in melanoma 2 (AIM2), which recognizes dsDNA, and cGAS, which binds to dsDNA and produces the secondary messenger 2'3'-cGAMP that in turn binds to stimulator of interferon genes (STING) (Mosallanejad & Kagan, 2022); cGAS-STING is the most used cytoplasmic DNA signaling pathway in many cell types (M.-M. Hu & Shu, 2020). RIG-I-like receptors (RLRs) RIG-I and MDA5 are the primary sensors of cytosolic RNA, recognizing 5' triphosphate dsRNA and high molecular weight RNA respectively (Martínez-Gil et al., 2013), but there are many other proteins that recognize RNA, including PKR, OASes, and ADARs (Hur, 2019).

In this thesis, we focus on nucleic acid responses to RNA and DNA that are active in many cell types: respectively, the RLRs RIG-I and MDA5 and cGAS/STING. Upon binding to RNA, RIG-I and MDA5 transition to an active conformation and associate with MAVS on mitochondrial and peroxisomal membranes (E. Dixit et al., 2010). MAVS recruits TBK1, which in turns results in activation of NF-κB and IRF3 and transcription of type I interferons (Ablasser & Hur, 2020; Chan & Gack, 2015). Activated STING also produces type I interferons via TBK1-IRF3 as well as inducing autophagy and inflammasome activation, accomplishing this by translocation upon cGAMP binding from the endoplasmic reticulum (ER) to the Golgi, where signaling occurs, and finally through the endolysosomal network, where it is degraded (Ablasser & Hur, 2020).

## 1.2 Functional Genomic Studies of Host-Pathogen Interactions

### *1.2.1 Overview of functional genomics*

Functional genomic approaches causally link genes with resulting phenotypes of interest. In the past, functional genomic screens relied on small interfering RNAs (siRNAs) or short hairpin RNAs (shRNAs) to perturb gene function; however, CRISPR-based technologies are now the dominant screening tools thanks to their ease of use and relatively high on-target editing efficiency combined with low off-target editing (Hart et al., 2015; Przybyla & Gilbert, 2022). The majority of genetic screens are loss-of-function screens, where double-stranded breaks (DSBs) induced by single-guide RNAs (sgRNAs) usually using the Cas9 endonuclease are repaired via NHEJ, resulting in out-of-frame mutations (Przybyla & Gilbert, 2022). Alternatively, CRISPR interference (CRISPRi) can be used to repress gene expression by fusing a catalytically inactive version of Cas9 (termed dCas9) to domains designed to repress transcription. CRISPRi enables screening in cells that poorly tolerate DSBs and facilitates screens for essential genes (Gilbert et al., 2014). Similarly to CRISPRi, CRISPR activation (CRISPRa) uses dCas9 fused to an effector domain that acts as a transcriptional activator to achieve gain-of-function screening. Gain-of-function screening can also be achieved by overexpression of genes of interest through introduction of exogenous gene expression vectors, while transposon-insertion sequencing is another non-CRISPR screening technique, which is able to identify both gain-of-function and loss-of-function effects (Bock et al., 2022; Puschnik et al., 2017). In addition, newer CRISPR technologies such as base editing and prime editing, as well as epigenetic editing will enable are beginning to enable screens for effects aside from a simple decrease or increase in expression of existing genes (Bock et al., 2022). Most CRISPR screens target coding genes, but CRISPR screens can also be used to interrogate non-coding regions. Screens that assay cell non-autonomous phenotypes or phenotypes that cannot be assayed in a pooled setting (e.g. secreted factors) must be done in an arrayed format but, when possible, pooling genetic perturbations results in much more tractable high-throughput screening with reduced technical variability and cost. Finally, while most screens to date have linked individual genes with phenotypes of interest, recently there has been increased effort to map genetic interactions using CRISPR approaches (Horlbeck et al., 2018).

### *1.2.2 Host-Directed Screening*

Most screens for host-pathogen interactions have focused on editing host genomes (usually human cells, although models such as mouse or green monkey cells are sometimes used as the latter have lost IFN-β expression and are easily infected (Desmyter et al., 1968; Mosca & Pitha, 1986)). Gain-of-function screens can be useful to identify cell surface receptors by, for instance, transducing a cDNA library from a cell type that is permissive to infection into cells refractory to infection; this approach identified HCV entry receptors (Evans et al., 2007; Ploss et al., 2009). Gain-of-function screens using designed ORF libraries have also identified viral regulators (S. Li et al., 2022; Saeed et al., 2015; Schoggins et al., 2011, 2014). More recently, CRISPR activation screens have also revealed regulators of host-virus interactions (Danziger et al., 2022; Heaton et al., 2017; Rebendenne et al., 2022). Gain-of-function screens are particularly attractive for revealing functions of redundant genes that would not be easily identified in loss-of-function screens, such as interferon-stimulated genes (ISGs).

Loss-of-function screens are much more common: highlights for virus screens include screens for WNV (H. Ma et al., 2015), DENV (Marceau et al., 2016), ZIKV and DENV (Savidis et al., 2016), EBV (Y. Ma et al., 2017), RSV (Hoffmann et al., 2017), EBOV (Flint et al., 2019), HIV (R. J. Park et al., 2017), IAV (B. Li et al., 2020), AAV2 (Pillay et al., 2016), and many for SARS-CoV-2, reviewed in Baggen et al., 2021. Fewer host-directed screens for bacterial pathogens have been performed: some highlights include screens in monocytes or macrophages for *S. typhimurium* (Yeung et al., 2019), *S. flexneri* (Lai et al., 2021), *L. pneumophila* (Jeng et al., 2019), and *M. bovis* BCG (Lai et al., 2020) and screens in intestinal cells for *S. typhimurium*, EHEC (Pacheco et al., 2018), *C. trachomatis* (J. S. Park et al., 2019), *V. parahaemolyticus* (Blondel et al., 2016) and HeLa cells for *S.typhimurium* (Gaudet et al., 2021).

Phenotypes screened for are usually 1) cell survival following pathogen infection, since many pathogens are cytolytic, 2) infection levels by quantifying expression of pathogen protein or RNA using flow cytometry, or 3) expression levels of a host reporter measured using flow cytometry, typically downstream interferon expression.

### *1.2.3 Pathogen-Directed Screening*

Genetic screens may also direct editing towards the pathogen genome. Transposon-insertion sequencing screens directed at pathogen genomes such as *P. falciparum* (M. Zhang et al., 2018), *M. tuberculosis* (Barczak et al., 2017), and *Y. pseudotuberculosis* (Zhou et al., 2018) have been successful but suffer from bias towards genes that have long coding regions and, since this technology relies on random insertions, these screens are always genome-scale (T. Wang et al., 2018). In bacteria, CRISPRi and dCas9-derived approaches are the most common CRISPR modality due to inefficient NHEJ repair in many species (Rousset et al., 2018; Yan et al., 2020) and because these systems are tunable, providing partial knockdown for essential genes (Todor et al., 2021). However, CRISPRi is polar, reducing expression of all downstream genes in an operon equivalently as well as decreasing upstream gene expression; therefore, it is only appropriate for investigating operon-level effects (Peters et al., 2015). In addition, some sgRNA seed sequences exhibit fitness effects in bacteria (Cui et al., 2018) and dCas9 overexpression can also be toxic in some bacteria (Cho et al., 2018; X. Li et al., 2016). Notwithstanding these considerations, a comparison in *E. coli* showed that CRISPRi outperformed transposon sequencing (T. Wang et al., 2018). Other notable bacterial CRISPRi screens include work with *M. tuberculosis* (Yan et al., 2022) and *B. subtilis* (Hawkins et al., 2020) and a more comprehensive review of bacterial CRISPRi screens can be found in Todor et al., 2021. Recent efforts have extended CRISPRi to non-model bacteria using a suite of tools termed Mobile-CRISPRi (Peters et al., 2019). Most bacterial screens are performed without the presence of mammalian host cells or where the bacteria themselves are the host cells and effects on bacteriophages are studied (Rousset et al., 2018) but, more recently, *in vivo* screens have been performed, where for instance effects of CRISPRi knockdown in bacteria are studied in the context of mouse models (X. Liu et al., 2021).

Some pathogen-directed screens have also studied other parasites such as fungi: for example, using CRISPR-Cas9-based gene drive arrays to study *C. albicans* fitness (Shapiro et al., 2018) and effects of *C. albicans* mutants on *C. elegans* hosts (Rosiana et al., 2021). Screens in protozoa are still uncommon since many protozoan parasites require an HDR repair template to repair DSBs (Bryant et al., 2019). One exception is the parasite *T. gondii*, which has an efficient NHEJ pathway, allowing for a genome-wide Cas9 fitness screen (Sidik et al., 2016) and targeted *in vivo* screens, where mice were infected with edited pools of the parasite (Young et al., 2019). In *Leishmania,* Beneke et al. screened a Cas9 knockout library of 100 mutants by providing a donor template (Beneke et al., 2017, 2019), a strategy that could be extended to other kinetoplastids.

Fewer examples of host-directed screens against viruses exist, since many viruses of interest have RNA genomes and Cas9-based technologies are appropriate for dsDNA sequences. Due to the limited number of genes in many viral genomes, often individual expression or deletion of genes without employing systematic CRISPR screening is used (Ludwig et al., 2022). dsDNA viruses can be edited by expressing Cas9 and sgRNAs in host cells infected with viruses as was done for HCMV (Hein & Weissman, 2022), KSHV (Gabaev et al., 2020; Morgens et al., 2022), and PRV (Tang et al., 2017). Some DNA viruses may not be suitable for Cas9 targeting by CRISPR, however, due to high expression levels of viral transcripts and short replication times, so experiments should be performed to validate efficient editing for each candidate virus of interest. Virus-directed screens could be useful not only for understanding viral biology but also for engineering attenuated vaccine strains or optimizing vectors for gene therapy, especially as new CRISPR tools like base editing and prime editing mature. On the other hand, screens that involve more diverse editing outcomes that could result in gain of function should be approached cautiously and precautions such as expression and editing of single proteins or fragments should be employed, especially when using viruses with pandemic potential (Lipsitch, 2018; Lipsitch & Bloom, 2012).

The vast majority of phenotypes assayed in pathogen-directed screens focus on fitness effects but some also assay fluorescent reporters aimed at measuring relative pathogen growth or metabolic state.

### *1.2.4 Image-Based Screens of Host-Pathogen Interactions*

A key limitation of traditional loss-of-function screens for host-pathogen interactions is that hits are usually biased towards factors important for only the strongest and simplest phenotypes, such as overall host or pathogen fitness or relative pathogen growth. For viruses, this often results in a strong bias towards discovery of entry factors, as these produce the strongest signals in the presence of the high infection levels typically used in screens. High infection levels increase power for factors that strongly reduce infection but often limit power for discovery of host restriction factors or antivirals that affect later stages of intracellular pathogen lifecycles. Image-based screens, on the other hand, offer the opportunity to recover more information about complex phenotypes.

High-throughput image-based arrayed genetic screens to assess host-pathogen interactions have been used to discover key regulators of these processes for over ten years. However, in many instances the naturally single-cell, multidimensional imaging datasets acquired were immediately compressed to average or low-dimensional measurements. For example, large-scale host-pathogen arrayed image-based screens measured percentage and total number of cells infected with IAV (Karlas et al., 2010) and mean infection burden as measured by fluorescence in rotavirus (Silva-Ayala et al., 2013) and coxsackievirus (H. S. Kim et al., 2018). In another case, cells infected with *L. monocytogenes* were manually categorized as having increased or decreased infection or as having a punctate infection pattern as well as manual rating of phenotype strength (weak, moderate, or strong) (Agaisse et al., 2005). In another study, genes required for colocalization of the Sindbis virus capsid protein with LC3 puncta were identified (Orvedahl et al., 2011). Additionally, a kinome-wide screen assayed IRF3 translocation in response to dsRNA transfection.

Only in a few cases were more image-based features extracted and analyzed. For example, in a screen of VACV-infected cells, over 200 features were extracted and a support vector machine (SVM) was used to classify cells into five binary phenotypes (infected/non-infected, interphase/non-interphase, mitotic/non-mitotic, apoptotic/non-apoptotic, blob/non-blob). These hundreds of features were also used to discard technical artefacts, scoring genes based on VACV infection level changes only after several normalization steps (Mercer et al., 2012). In a pathogen-directed screen in *M. smegmatis*, a CRISPRi library targeting 263 genes was introduced and multi-parameter features were extracted and used to cluster genes by their resulting ‘morphotypes’ (Wet et al., 2020).

Image-based arrayed chemical screens have also examined responses to pathogenic infection: for instance, by assaying overall RVFV infection levels after compound treatment (Mudhasani et al., 2014) or, in a more complex analysis, by testing compound effects in the context of SARS-CoV-2 infection on multiple cellular organelles using a modified Cell Painting protocol and extracting 1024 features with a CNN architecture (Heiser et al., 2020).

Arrayed image-based screens are challenging to scale and require extensive normalization due to noise introduced by plate-plate variability. In bacteria, a method termed DuMPLING (Camsund et al., 2020; Lawson et al., 2017) enabled pooled screening of 235 *E. coli* knockdowns via *in situ* genotyping following live-cell imaging. In human cells, a method termed optical pooled screening leveraged *in situ* sequencing to enable pooled screens of thousands of perturbations with image-based phenotypes (Feldman et al., 2019, 2022; Funk et al., 2022). Alternative strategies for pooled image-based screening in human cells are comprehensively reviewed in Walton et al., 2022; however, these strategies are either 1) not compatible with genome-scale screens in human cells or 2) require an enrichment step, compressing complex imaging data into a one-dimensional measurement and not allowing further analysis of the rich information acquired from images.

## 1.3 Profiling Host-Pathogen Interactions

High-content single-cell screens, with multi-dimensional outputs not necessarily restricted to imaging, are becoming routine and now scale to genome-wide perturbations even for transcriptomic readouts (Replogle et al., 2022). These high-content profiling datasets enable screens that don't require *a priori* specification of phenotypes of interest but rather result in datasets that can be used to rapidly generate and investigate hypotheses based on primary screening data alone (Bock et al., 2022). Aside from the work presented in this thesis, high-content profiling screens of host-pathogen interactions have been limited. Two recent examples of high-content transcriptome screens are a study of HCMV (Hein & Weissman, 2022) and SARS-CoV-2 (Sunshine et al., 2022).

In the HCMV study, a pooled pathogen-directed screen assessed genomic regions in HCMV that altered fibroblast fitness and a pooled host-directed screen revealed host factors altering fitness upon HCMV infection. Next, 52 host genes identified as hits in the pooled screens as well as a combination of 21 host genes and 31 viral genes were interrogated in high-content Perturb-seq screens at multiple timepoints post-infection. While host gene knockouts resulted in different infection kinetics, virus-directed perturbations altered the trajectory of infection itself. The transcriptome-level readout also enabled co-clustering of host factors that aligned with prior knowledge of relevant pathways. In the future, this approach could be used at higher throughput to impute functions of poorly characterized genes in the context of infection. In the SARS-CoV-2 study, CRISPRi perturbations targeting 195 single genes and 22 dual-gene perturbations were delivered to Calu3 cells infected with SARS-CoV-2 for 24 hours and authors found genes not routinely implicated in the interferon response (*SPNS1, KEAP1, GPR89A/B)* that co-clustered with canonical interferon genes (*IFNAR2, STAT2, IRF3, IRF9)* and found that these genes also reduced ISG expression.

Beyond genetic perturbation screens, single-cell profiling studies examining viral infection have revealed insights in a number of areas related to host-viral interactions. Cell-cell variability has been noted in multiple infection phases for diverse viral families, from the initial viral particles binding to cells, to endocytosis, cytosolic penetration, replication, and host responses including immune responses to counteract infection or prevent spread (Suomalainen & Greber, 2021). For example, a recent study leveraged live-cell imaging and scRNA-seq to show that in the context of HSV-1 infection, a rare group of abortively infected cells successfully initiated antiviral signaling, while highly infected cells, on the other hand, differentiated into an embryonic-like state (Drayman et al., 2019). Another HSV-1 study modeled cell death decision-making upon infection using live-cell imaging data (Oyler-Yaniv et al., 2021). Other profiling studies leveraged single-cell sequencing data to assess differential responses in infected vs bystander cells upon infection with DENV (Zanini, Robinson, et al., 2018), HCMV, as already discussed (Hein & Weissman, 2022), and SARS-CoV-2 (Bost et al., 2020), or to correlate viral burden with host responses in EBOV (Kotliar et al., 2020) or ZIKV infection (Zanini, Pu, et al., 2018). However, most profiling studies to date have focused on a small fraction of the viruses known to be pathogenic in humans so these methods should also be extended to less well-characterized species (Ratnasiri et al., 2022). Single-cell RNA-seq studies of intracellular bacterial infections have been similarly informative; for example, a study showed that macrophages harboring non-growing Salmonella displayed an M1 polarization state, while fast-growing Salmonella induced an M2 anti-inflammatory state (Saliba et al., 2016) and another study showed that Salmonella PhoPQ activity drove variable interferon responses (Avraham et al., 2015).

## 1.4 Optical Pooled Screens Enable High-Throughput Perturbation and Profiling of Host-Pathogen Interactions

As discussed previously, high-throughput screens for host-pathogen interactions have focused on simple, one-dimensional phenotypes compatible with fitness or flow cytometry readouts. These simple yet powerful screening approaches have recovered many key regulators of diverse processes. For instance, a flow cytometry screen drove our hypothesis and further work to demonstrate that STING itself is a proton channel (**Chapter 4**). However, more complex, multi-dimensional profiles of the effect of genetic knockouts on host-pathogen interactions can often provide further insights. Transcriptomic single-cell measurements have high information content but are often cost-prohibitive for genome-scale measurements. Optical pooled screening, a technology recently developed in the Blainey lab (Feldman et al., 2019), on the other hand, allows for genome-scale image-based profiles in tens of millions of cells. In this thesis, we present two distinct optical pooled screens that assess host responses to foreign RNA, demonstrating examples of insights that result from combined perturbation and high-content profiling at the scale of the genome. In **Chapter 2**, we assess the innate immune response to Sendai virus (SeV) infection by examining the RIG-I pathway and in **Chapter 3** we explore the host-virus interplay upon Ebola virus.

# 2 A genome-wide optical pooled screen reveals regulators of cellular antiviral responses

Authors: Rebecca J. Carlson, Michael D. Leiken, Alina Guna, Nir Hacohen*, Paul C. Blainey*

* Contributed equally

## 2.1 Abstract

The infection of mammalian cells by viruses – and innate immune responses to infection – are spatiotemporally organized processes. Cytosolic RNA sensors trigger nuclear translocation of the transcription factor IRF3 and consequent induction of host immune responses to RNA viruses. Previous genetic screens for factors involved in viral sensing did not resolve changes in the subcellular localization of host or viral proteins. Here, we increased the throughput of our optical pooled screening technology by over four-fold. This allowed us to carry out a genome-wide CRISPR knockout screen using high-resolution multiparameter imaging of cellular responses to Sendai virus infection, coupled with *in situ* cDNA sequencing by synthesis to identify 80,408 sgRNAs in 10,366,390 cells – over an order of magnitude more genomic perturbations than demonstrated previously using an *in situ* SBS readout. By ranking perturbations using human-designed and deep learning image feature scores, we identified novel regulators of IRF3 translocation, Sendai virus localization, and peroxisomal biogenesis. Among the hits, we found that ATP13A1, an ER-localized P5A-type ATPase, is essential for viral sensing and is required for targeting of MAVS to mitochondrial membranes where MAVS must be localized for effective signaling through RIG-I/DDX58. The ability to carry out genome-wide pooled screens with complex high-resolution image-based phenotyping dramatically expands the scope of functional genomics approaches.

## 2.2 Significance Statement

Our study demonstrates the first genome-wide optical pooled profiling screen. Profile-based screening is a uniquely powerful phenotype-based genetic tool that enables retrospective assignment of genetic perturbations to multiple phenotypic categories. Here, we demonstrated the power of optical pooled screening to support analysis of 80,000 perturbations from single-cell images of about 10 million cells. We leverage the results to generate new insights into innate immune responses to viral infection including the essentiality of ATP13A1 function for signaling through RIG-I/DDX58.

## 2.3 Introduction

Host cell responses to viral infection are conserved processes critical for defense against established and emerging viruses. Intracellular responses to RNA virus infection are mediated by innate immune signaling pathways that begin with binding of viral RNA by the pattern recognition receptors (PRRs) MDA5 or RIG-I. The PRRs in turn trigger phosphorylation and nuclear translocation of the transcription factor IRF3, which induces transcription of the interferon antiviral immunity program (Chow et al., 2018). Activating mutations in these PRRs and increased spontaneous IRF3 translocation are associated with autoimmune syndromes (Buers et al., 2016; Crow & Manel, 2015), and activation of PRRs such as RIG-I, on the other hand, can boost antitumor responses (Jiang et al., 2019). Given the importance of IRF3 translocation, a better understanding of genes that regulate its activation and translocation may contribute to design of therapeutics to boost antiviral immunity, dampen autoimmunity, or support antitumor immunity.

While a number of genome-wide CRISPR screens for control of virus infection have been performed, most have assayed cell fitness (Baggen, Persoons, et al., 2021; Daniloski et al., 2021; Ding et al., 2018; Flint et al., 2019; Y. Li et al., 2019; Schneider et al., 2021; R. Wang et al., 2021; Zhu et al., 2021) or viral replication (B. Li et al., 2020; Richardson et al., 2018) in response to cellular infection by a pathogenic virus. More recently, pooled genome-wide screens have examined interferon induction in response to interferon treatment (Börold et al., 2021; Harding et al., 2021; Lumb et al., 2017) or Sendai virus (SeV) infection (Baril et al., 2013); however, none have been able to examine IRF3 translocation directly, a critical step in the sensing pathway because pooled genetic screens are not typically compatible with subcellular imaging assays. While a kinome-wide image-based arrayed siRNA screen for IRF3 translocation has been performed (Willemsen et al., 2017), genome-wide arrayed screens have not become routine due to the high cost and labor investment required for cellular imaging in an arrayed format.

We recently developed a method to perform pooled image-based screens by linking single-cell image-based phenotypes to perturbation identity in mammalian cell libraries (Feldman et al., 2019, 2022). Using targeted *in situ* sequencing, this method demultiplexes entire cell libraries containing genetic CRISPR perturbations following phenotyping. In order to scale the method to the genome-wide level, we improved the throughput by over four-fold by 1) using a microscope system with reduced stage movement latency (NIS Elements Software with a Ti2 microscope rather than MicroManager software with a Ti-E microscope) and 2) completing tiled image acquisitions for each color channel before reconfiguring the microscope for the next color channel in order to eliminate time-intensive optical filter switching at each field of view. The second advance degraded raw alignment precision across channels, necessitating development of fast and accurate computational image alignment across multi-centimeter scales (**Supplementary Fig. 2.1e**). In addition, while we previously only assayed two to four phenotyping channels due to incompatibility of other channels with Illumina's four-color sequencing-by-synthesis reagents (Feldman et al., 2019; Funk et al., 2022), here we assayed as many as seven phenotyping channels at a time by taking advantage of the higher local optical intensity of the *in situ* sequencing signals and applying our higher-throughput data acquisition protocol for phenotyping with commercially available reagents. Our multiparameter screen identified novel regulators of IRF3 translocation, peroxisome biogenesis, and spatial localization of SeV, a paramyxovirus known to induce strong interferon responses (Bedsaul et al., 2016), and demonstrates the feasibility of optical pooled screening at the genome-wide scale to discover complex, multiparameter phenotypes.

## 2.4 Results

### *2.4.1 A Genome-wide Optical Pooled Screen Identifies Regulators of IRF3 Translocation*

We performed a genome-wide optical pooled screen using a lentivirus library of 80,408 sgRNAs targeting over 20,000 genes, with 454 non-targeting sgRNAs. HeLa-TetR-Cas9 cells were stably transduced with BFP-PTS1 (**Supplementary Fig. 2.1a-b**), a peroxisomal marker, and were infected at an MOI of 0.05 with the sgRNA library. Following 7 days of doxycycline Cas9 induction, cells were infected with the SeV Cantell strain at an MOI of 10 for 15 hours prior to imaging (Fig. 1a). We then performed multiplexed immunofluorescence imaging, assessing two organelles critical to antiviral signaling - peroxisomes and mitochondria (E. Dixit et al., 2010) - the RNA receptors MDA5 and RIG-I, pIRF3, SeV itself, and DNA (DAPI) in all cells (**Fig. 2.1c**) before performing reverse transcription, padlock targeting of cDNA, RCA, and in situ sequencing by synthesis (SBS) to detect sgRNAs as previously described (Feldman et al., 2022; Ke et al., 2013). We performed 12 cycles of in situ SBS to read out all sgRNAs (**Fig. 2.1d**). Due to the high intensity of the in situ SBS signal relative to immunofluorescence staining, we titrated the immunofluorescence reagents that produced signal in the phenotyping channels that overlapped with SBS spectra to mitigate interference with base calling (**Fig. 2.1b**). This strategy enabled the acquisition of as many as seven channels of phenotyping data without using custom modified antibodies (Funk et al., 2022), an advancement over the two to four phenotyping channels previously demonstrated using optical pooled screening without specialized reagents (Feldman et al., 2019). Images of single cells from our genome-wide pooled screen showed robust translocation of phosphorylated IRF3 in nontargeting control cells, while RIG-I, MAVS, and IRF3 knockouts profoundly impair translocation across multiple sgRNAs (**Fig. 2.1e,f**) and were the three highest scoring genes in the screen (**Fig. 2.1g-i**). To confirm the results of our primary screen, we performed focused secondary pooled screens (using 6 sgRNAs/gene) of 342 genes that scored for either increasing or decreasing IRF3 translocation in our genome-wide screen using antibody staining or an IRF3 reporter (**Fig. 2.1g**). Secondary screen translocation scores were well-correlated with translocation scores from the primary genome-wide screen (**Fig. 2.1h**), underscoring the considerable statistical power achieved for IRF3 translocation in the primary genome-wide screen. 51 genes were confirmed by stringent criteria to score in both of our secondary screens, including the known pathway members MAVS, DDX58, RNF135, and IRF3 (Rehwinkel & Gack, 2020), **Fig. 2.1i**.

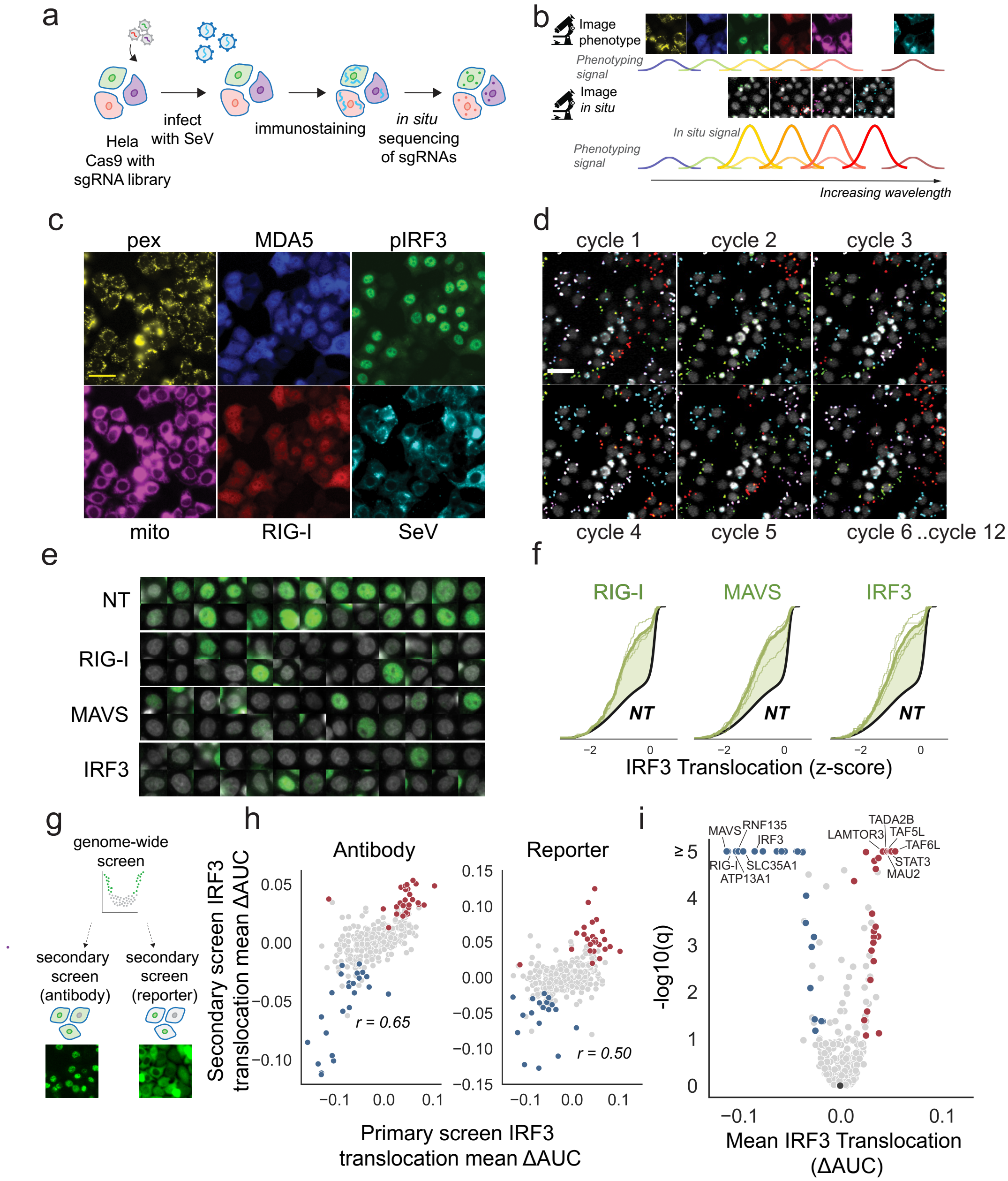


**Figure 2.1.**
A genome-wide optical pooled screen reveals regulators of pIRF3 translocation. (A) Genome-wide screening workflow. (B) Bright in situ signals enable imaging of phenotyping fluorescence signals that overlap with in situ sequencing by synthesis fluorescence signals. (C) Example images for each of the six immunofluorescence channels imaged in the genome-wide screen (DAPI was additionally imaged). Scale bar 20 μm. (D) First six in situ sequencing cycles (of 12 total cycles) for the same group of cells shown in C. (E) Single-cell images from genome-wide screen show a decrease in IRF3 translocation in positive control genes. (F) Histograms of z-scored IRF3 translocation for positive control genes. Individual traces

correspond to unique sgRNAs. (G) Hits from the genome-wide screen were assessed using a new library of 342 genes with 6 sgRNA/gene in two secondary screens with different IRF3 detection assays. (H) Secondary translocation screens are well-correlated with the primary screen for both antibody staining (r = 0.65, p = 2.6e-42) and reporter (r = 0.50, p = 5.22e-23). Blue dots indicate significant hits with decreased IRF3 translocation upon knockout while red dots indicate genes with increased IRF3 translocation. (I) Volcano plot of secondary screen results for antibody staining with hits confirmed via secondary screening marked in blue or red. Two-sided p-values were calculated as described in the Methods, corrected using the Benjamini-Hochberg procedure, and aggregated on the gene level using Fisher's method.

We next examined regulators of IRF3 translocation confirmed by secondary screening using STRING analysis (Szklarczyk et al., 2019) and found that the majority are predicted interactors (minimum confidence 0.4, 13/24 of the genes found to decrease translocation, p = 8.16e-7, **Fig. 2.2a** and 18/27 genes that increase translocation, p = 1.11e-16, **Fig. 2.2b**). Among genes found to decrease translocation, we identified a set containing genes known to be involved in RIG-I induction of IFN, as well as genes involved in ubiquitin transferase activity, which is intriguing since ubiquitin is known to be involved in regulating several pathway members but only *FBXW7* has been previously identified to have a direct role in this pathway (Song et al., 2017). In addition, we identified a cluster of three genes associated with sialic acid metabolism (*CMAS, SLC35A1*, and *GNE*) that decreased IRF3 translocation likely secondary to inhibition of SeV entry, as viral load was greatly diminished in these knockouts (**Fig. 2.2c**), and viral entry is dependent on expression of sialic acids on the cell surface (Markwell & Paulson, 1980). Novel positive regulators also included many genes not previously associated with IRF3 translocation, including *ATP13A1*, which was one of the top three genes that decreased translocation. *IKBKG* (NEMO) knockout also decreased IRF3 translocation and was previously reported to be required for MAVS phosphorylation, which in turn induces IRF3 activation (S. Liu et al., 2015). Among putative negative regulators nominated by the screen were the Ragulator complex (*LAMTOR2,3,5* and associated genes *FLCN* and *RRAGC)*, which is required for mTOR activation (Sancak et al., 2010) and whose activity was found to increase SeV load (**Fig. 2.2c**). An mTOR inhibitor, rapamycin, has been shown to increase SeV load in infected cells (Subramanian et al., 2018) but the Ragulator complex has not been previously associated with a role in SeV infection or IRF3 activation. Other significant complexes that negatively regulated IRF3 translocation include Mediator-associated genes (*MED13, MED16, MED24, CCNC,* and *BTAF1),* STAGA complex members *(TADA2B, TAF5L,* and *TAF6L),* and components of the HOPS complex (*VPS11, VPS16,* and *VPS39),* which mediates endo-lysosomal fusion, a process that likely reduces SeV replication, as it is inhibited in infected cells (Wozniak et al., 2016). Next, we also performed a focused optical pooled screen with the same set of genes upon vesicular stomatitis virus (VSV) infection, in order to determine the effect of each perturbation on IRF3 translocation upon infection with an orthogonal RNA virus, as VSV belongs to a different virus family (**Fig. 2.2e-f**). Notably, despite reduced effect sizes in the VSV screen, *ATP13A1* and *ATP2B1* decreased IRF3 translocation upon knockout in both settings.

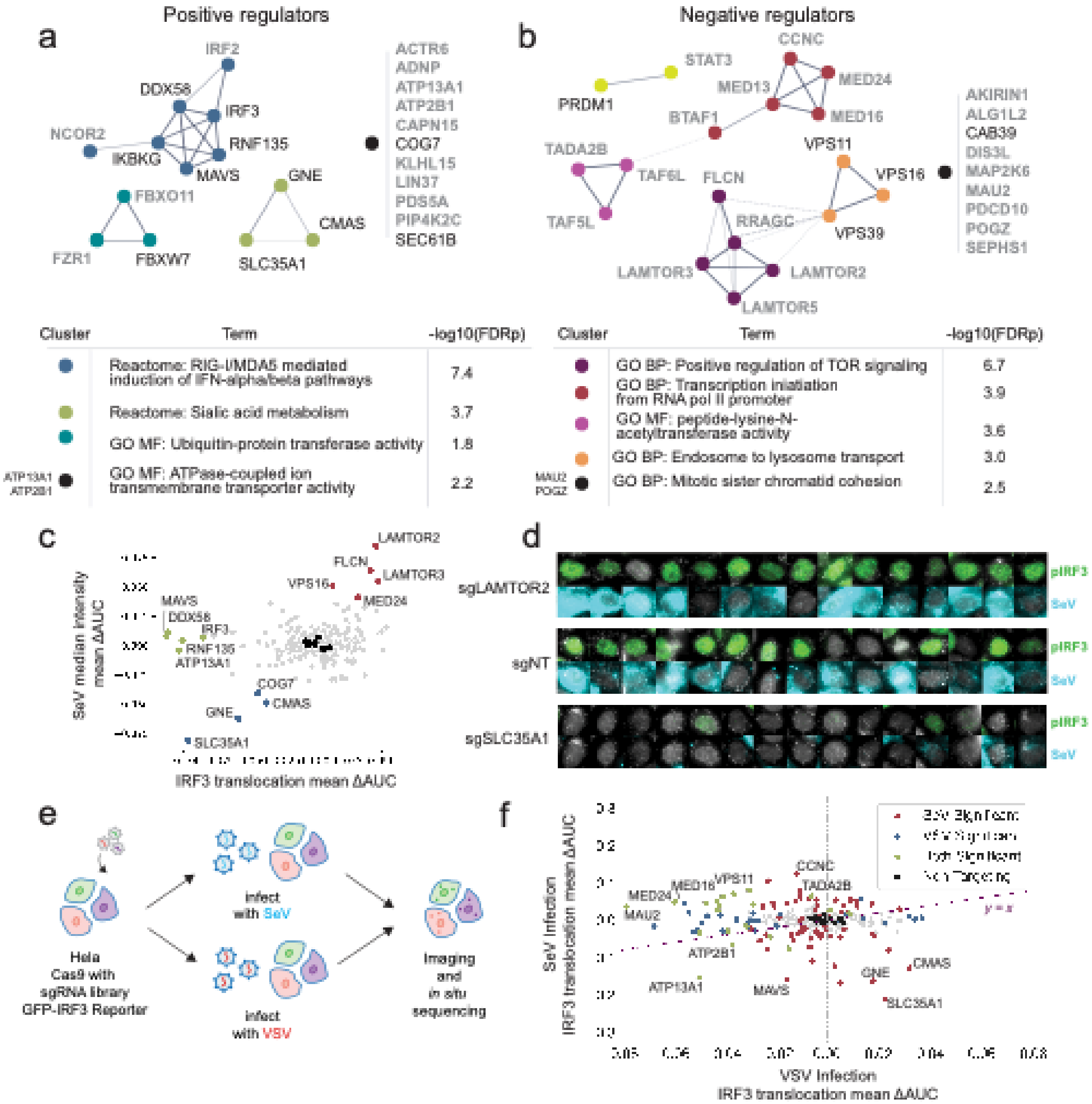


**Figure 2.2.**
Analysis of combined hits from secondary screens (A). STRING clusters for hits showing decreased translocation using MCL inflation parameter 3 for clustering. Novel genes are in gray and hits that did not cluster with other hits are listed by the black cluster marker. Table: Top significant GO/Reactome terms associated with each cluster among putative positive regulators. P-values in a, b are computed via Enrichr (Kuleshov et al., 2016) using the Fisher exact test and adjusted using the Benjamini-Hochberg procedure. (B) STRING clusters as described in (A) for combined hits that increased IRF3 translocation. Table contains top significant GO/Reactome terms associated with each cluster among putative negative regulators. (C) IRF3 Translocation score plotted against SeV intensity scores for all genes in the antibody staining secondary screen. Black dots indicate non-targeting controls. (D) Example single-cell images for non-

targeting control cells and the top gene for increased SeV or decreased SeV burden. For each gene/control, pIRF3 staining appears in green in the top row and SeV staining in cyan in the bottom row; each overlaid with DAPI nuclear staining. (E) Workflow for secondary screening to assess IRF3 reporter translocation differences upon either Sendai or VSV infection. (F) Comparison of per-gene translocation effects between VSV and Sendai virus infections, significance indicates genes with an FDR-adjusted p-value <0.01.

### *2.4.2 Deep Learning and High-Dimensional Screen Analysis Reveals Novel Gene Functions*

In addition to extraction of targeted features such as translocation and intensity from single-cell images, we also input cropped images of individual cells separately for each channel into deep convolutional neural networks (CNNs) to enable unbiased identification of additional features. We extracted features from a CNN pretrained on ImageNet (Deng et al., 2009) and separately trained an autoencoder on our dataset (**Fig. 2.3a, Supplementary Fig. 2.2a-b**). Interestingly, we found that both the features extracted from the pretrained network as well as features defined in the autoencoder's latent representation space identified confirmed hits from the secondary screens with substantially higher accuracy than translocation alone (**Fig. 2.3b**, ROC AUC improves from 0.71 to as high as 0.93; both CNNs were naïve to the secondary screening data). We therefore selected 14 genes predicted to decrease translocation upon knockout by combined autoencoder and transfer learning features that had not scored highly enough to be included in our secondary screens. We found that 8/14 of these genes did indeed decrease translocation via arrayed testing (**Fig. 2.3c**), confirming the hypothesized increased predictiveness of machine learning features even for a defined translocation phenotype.

In addition to IRF3 translocation, we also extracted peroxisome (**Fig. 2.3d**) and SeV (**Fig. 2.3g**) intensity features from our genome-wide single-cell-resolution screen dataset. Our screen also detected genes affecting peroxisomal biogenesis (*PEX13, PEX2, PEX14, PEX1, PEX12, PEX10, PEX6, PEX5, PEX26, PEX11B, PEX3)* as decreasing peroxisome intensity in our CRISPR knockout screen, while known peroxisomal fission genes (*DNM1L, MFF*) increased peroxisomal intensity as expected. Genes showing a decrease in peroxisome intensity selected using peroxisome intensity (targeted), transfer learning, and autoencoder features, all were significantly enriched for the peroxisome GO term (**Fig. 2.3e**), while genes showing increased peroxisome intensity selected using autoencoder features but not targeted or transfer learning features were significantly enriched for the GO term heme biosynthesis, which has been linked to peroxisome biogenesis through PGC-1α in some studies (Handschin et al., 2005, 2005), demonstrating the overlapping but also distinct contributions of these three methods of feature extraction. We selected three genes that showed reduced peroxisome intensity using all three feature sets, and confirmed via arrayed validation that, even in the absence of SeV, peroxisome intensity was indeed decreased in these genes (**Fig. 2.3f**).

We next examined genes that showed changes in SeV intensity in our screen (**Fig. 2.3g**). In addition, we analyzed transfer learning and autoencoder features extracted from the Sendai virus channel to identify perturbations that significantly altered Sendai virus subcellular localization. Interestingly, when we performed dimensionality reduction using PHATE (Moon et al., 2019) on the transfer learned features from the Sendai virus channel, we found that the top 30 genes with decreased Sendai virus intensity (marked in orange in **Figure 2.3h**) did not cluster closely together, indicating that the transfer learned features were capturing significant changes in the Sendai virus images other than changes in our intensity metric. We therefore analyzed the top 30 genes most different from non-targeting controls in the transfer learned features (**Figure 2.3i**). Enrichment analysis of this set showed that genes related to intracellular trafficking (Deffieu et al., 2021; Kehl et al., 2020) scored highly for increased SeV burden (*VPS11, RAB14, TMED10)*, while genes involved in glycosylation and sialic acid metabolism *(SLC35A1, GNE,* and *CMAS)* were highly ranked for a reduction in SeV intensity using targeted and autoencoder features (**Fig. 2.3j**). On the other hand, the transfer learning features preferentially identified genes involved in mRNA processing (*SART1, NUP155* - already shown to interact with Sendai's M protein (Pentecost et al., 2015)*, DHX9* - shown to facilitate HIV, HCV, and FMDV replication but not previously associated with SeV (Fullam & Schröder, 2013)*, PRPF31*), and single-cell images of cells enriched in transfer learning showed diffuse cytoplasmic Sendai antibody staining (**Fig. 2.3j-k**), likely indicating that RNA processing genes required for the formation of cytoplasmic replication foci were not functioning effectively.

We next used features extracted from our autoencoder to cluster genes from our secondary screen. When we clustered genes using features from the IRF3 and RIG-I channels (**Fig. 2.3i**), we found that strong positive regulators we had identified (blue text) and negative regulators (red text) clustered apart from non-targeting controls (black), cluster membership indicated by marker colors. When we clustered using features derived from the Sendai virus channel, on the other hand, we found that some of the positive regulators (e.g. *MAVS, DDX58*, *ATP13A1*) did not cause changes in the Sendai channel relative to non-targeting controls, while other genes known to be involved in sialic acid biosynthesis and the major negative regulators did alter Sendai virus localization. Interestingly, positive regulators *FBXO11* and *FBXW7* clustered differently from non-targeting controls in Sendai virus features as well as features from the nuclear and peroxisome channels (**Fig. 2.3n-o**), indicating that apart from their effect on IRF3 translocation, they induce significant cellular perturbations. Notably, genes known to be involved in nuclear organization (*LMNB1)* or peroxisome-related metabolic processes (*SUCLG1)* clustered apart from non-targeting controls in only the nuclear or peroxisomal features respectively.

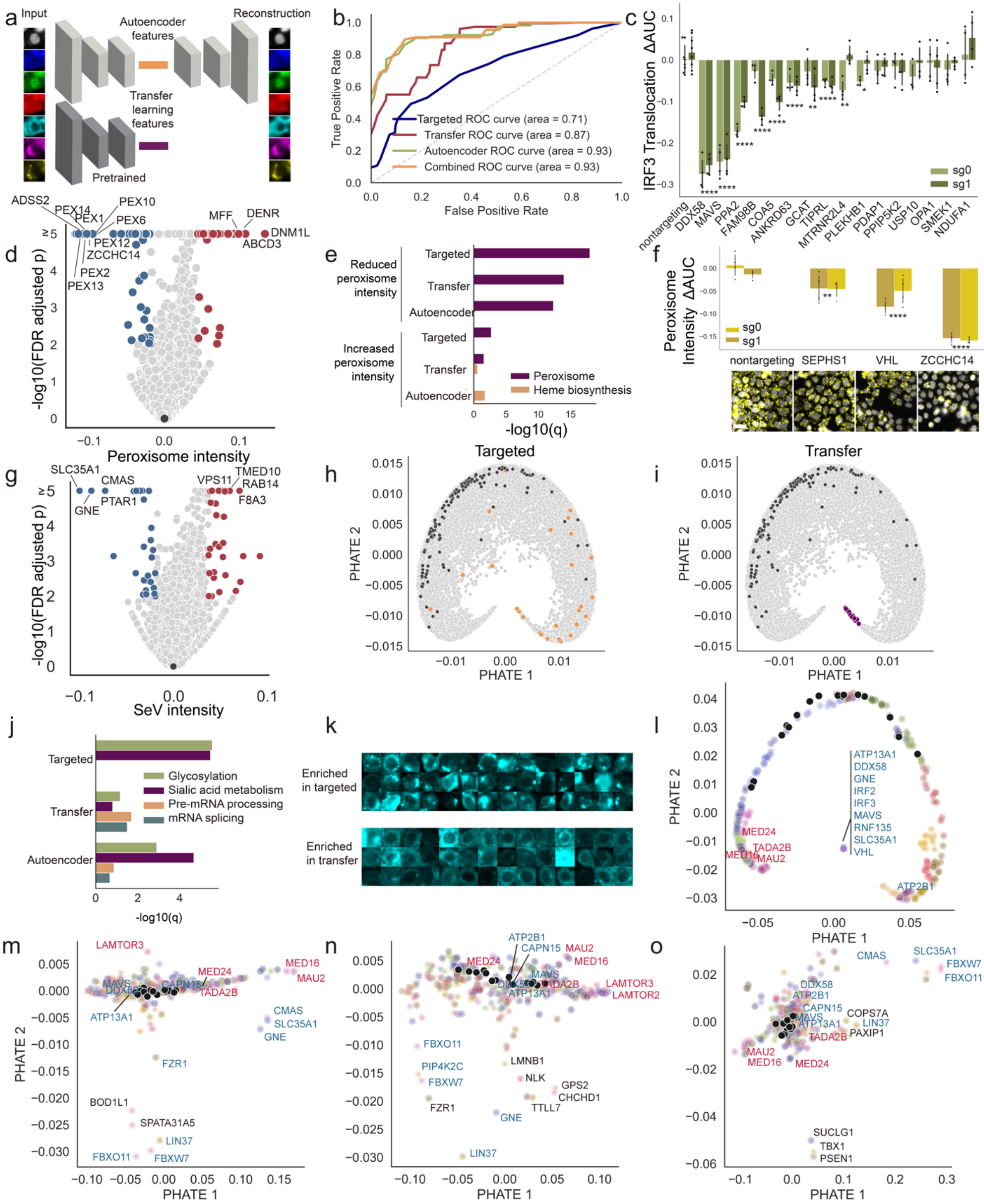


**Figure 2.3.**
Deep learning and multidimensional analysis reveal novel regulators of IRF3, peroxisomes, and SeV. (A) Schematic showing extraction of features from representative single-cell crops where each channel is processed by either a pretrained neural network (Xception) or an autoencoder that yielded the example reconstructions shown on the right. (B) Both transfer learning and the autoencoder, as well as combined features, improved prediction of true IRF3 translocation hits based on secondary screen results compared to translocation alone using random forest models. (C) Eight out of fourteen genes predicted to decrease translocation by combined deep learning features showed significantly decreased translocation in arrayed testing with two sgRNAs. * indicates adjusted $p < .05$, ** $p < .01$, *** $p < .001$, **** $p < .0001$. Two-sided p-values for c and f were calculated by computing the delta AUC for the feature of interest between cells in each well and non-targeting control cells (minimum $n = 3{,}000$); delta AUCs for all wells containing cells with an sgRNA of interest were then compared to non-targeting delta AUCs using a t test, p values were combined using Fisher's method and corrected using the Benjamini-Hochberg procedure. (D) Mean peroxisome intensity delta AUC plotted against FDR adjusted p-values for genome-wide screen. Two-sided p-values were calculated as described in the Methods and corrected using the Benjamini-Hochberg procedure. (E) Genes with reduced peroxisome intensity for targeted, transfer, autoencoder, or all three (13 genes in Venn diagram in **Supplementary Fig. 2.2c**) features showed significant enrichment of GO:0005777 (Peroxisome), while genes showing increased peroxisome intensity based on autoencoder features, unlike targeted features, were enriched for GO:0006783 (Heme biosynthesis). Enrichr p-values for e and h are computed using the Fisher exact test and adjusted using the Benjamini-Hochberg procedure. (F) Three genes with predicted reduction in peroxisome intensity using 2 sgRNAs per gene showed decreased peroxisome intensity in arrayed testing. Scale bar 20 μm. * indicates adjusted $p < .05$, ** $p < .01$, *** $p < .001$, **** $p < .0001$. (G) Mean SeV intensity delta AUC plotted against FDR adjusted p-values for genome-wide screen. Two-sided p-values were calculated as described in the Methods and corrected using the Benjamini-Hochberg procedure. (H) PHATE plot of genes derived from transfer learned features from the SeV channel with the top 30 genes that reduced SeV intensity marked in orange (non-targeting controls marked in black). (I) Same as (H) with the top 30 transfer learned perturbations most distant from non-targeting controls (black points) marked in purple. (J) Genes with reduced SeV intensity using targeted and autoencoder features were enriched for R-HSA-446203 (Glycosylation) and R-HSA-4085001 (Sialic acid metabolism), while genes showing reduced intensity in transfer learning showed stronger enrichment of features related to RNA processing (R-HSA-72203, pre-mRNA processing and R-HSA-72163, mRNA splicing). (K) Images of single cells that scored in the 99th (bottom) or 1st (top) percentile for the SeV transfer learning feature that most differentiated cells with perturbations shown in purple in panel I from non-targeting control cells. (L) PHATE plot of genes in secondary screen clustered using Leiden clustering on autoencoder features from the IRF3 and RIG-I channels, with select genes that decreased IRF3 translocation upon knockout marked with blue text and genes that increased translocation marked in red text; cluster membership is denoted by marker color. (M) Same as I but using autoencoder features from the Sendai virus channel (N) Same as I but using autoencoder features from the nuclei/DAPI channel (O) Same as I but using autoencoder features from the peroxisome channel.

### *2.4.3 Arrayed Validation and Transcriptional Analysis of Novel IRF3 Translocation Hits*

Following confirmation via secondary screening, we selected seven genes that either decreased (*ATP13A1, CAPN15, ATP2B1*) or increased (*MAU2, MED16, MED24, TADA2B*) IRF3 translocation when knocked out and confirmed the expected effects on IRF3 translocation via arrayed knockout with two sgRNAs (**Fig. 2.4a**, knockout efficiency validated in **Supplementary Fig. 2.3**, supporting translocation images in **Supplementary Fig. 2.4a**). We also assayed RIG-I activation (using an antibody against the activated form) and confirmed that 5/7 genes showed differences in RIG-I that paralleled the IRF3 translocation defect (**Fig. 2.4b**). Interestingly, *ATP2B1* showed a reduction in IRF3 translocation but an increase in RIG-I activation, indicating that it may regulate a pathway step between RIG-I and IRF3. Next, we assayed the effect of these knockouts in the U937 cell line, a cell model of monocytes, which are critical for the innate immune response. While none of the putative negative regulators significantly altered translocation in U937 cells, all three putative positive regulators reproduced a loss of translocation in these cells (**Fig. 2.4c, Supplementary Fig. 2.4b**). Interestingly, the putative negative regulators are all known to modulate transcription, which may be more co-variable with changes in the cellular context or the infection timepoint selected, as we were only able to observe infected U937 cells for a shorter period (7hr) than HeLa cells (15hr) due to high levels of cell death observed at later timepoints. We next quantified the levels of the active form of RIG-I in HeLa cells in response to a transfected synthetic hairpin RNA (hpRNA) stimulation in order to determine whether altered responses in the context of Sendai infection were secondary to modulation of the infection or, alternatively, were directly regulating antiviral sensing as would be predicted if defects were conserved in response to hpRNA treatment. We again found that responses to knockout of negative regulators was less well-conserved, perhaps due to their smaller effect on overall IRF3 translocation (**Fig. 2.4a**), and saw no alteration in active RIG-I in these knockouts. However, the two strongest novel positive regulators, *ATP13A1* and *CAPN15*, reproduced the defect in RIG-I activation in response to hpRNA stimulation (**Fig. 2.4d**).

We then performed RNA sequencing of five novel regulators to characterize transcriptional effects of these genetic knockouts in an unbiased manner. We found that all five novel regulator knockouts had significantly decreased (*ATP13A1*, *CAPN15*) or increased *(MED16, MED24*, and *TADA2B)* interferon-stimulated gene (ISG) expression in response to both SeV stimulation at two timepoints and hpRNA stimulation (**Fig. 2.4e**). In addition, none of the genes showed significant differences in SeV RNA levels relative to nontargeting controls (**Supplementary Fig. 2.4c**). Interestingly, a subset of ISGs was upregulated in *ATP13A1* and *CAPN15* knockout cells at later infection timepoints (**Fig. 2.4f**). We performed GSEA analysis to further investigate altered pathways in negative regulators of translocation and found that all three had increased serine-type endopeptidase activity (GO:004252), while *TADA2B* also had increased NOD-like receptor signaling and both Mediator members had increased lysosomal lumen component RNA expression (**Fig. 2.4g**). To further examine novel positive regulators, we performed gene ontology analysis on genes significantly decreased in each regulator relative to both *DDX58* and *MAVS* in the presence of SeV. We found that *CAPN15* had significantly decreased expression of nucleolar

RNAs and nucleolus components, while among genes with decreased expression in the *ATP13A1* knockout, transcripts encoding mitochondrial components were significantly perturbed (**Fig. 2.4h**). Finally, we also show that active RIG-I protein is significantly reduced in the absence of *ATP13A1* and *CAPN15* upon influenza (IAV) infection and for *ATP13A1* but not *CAPN15* following respiratory syncytial virus (RSV) infection (**Supplementary Fig. 2.4e-f**).

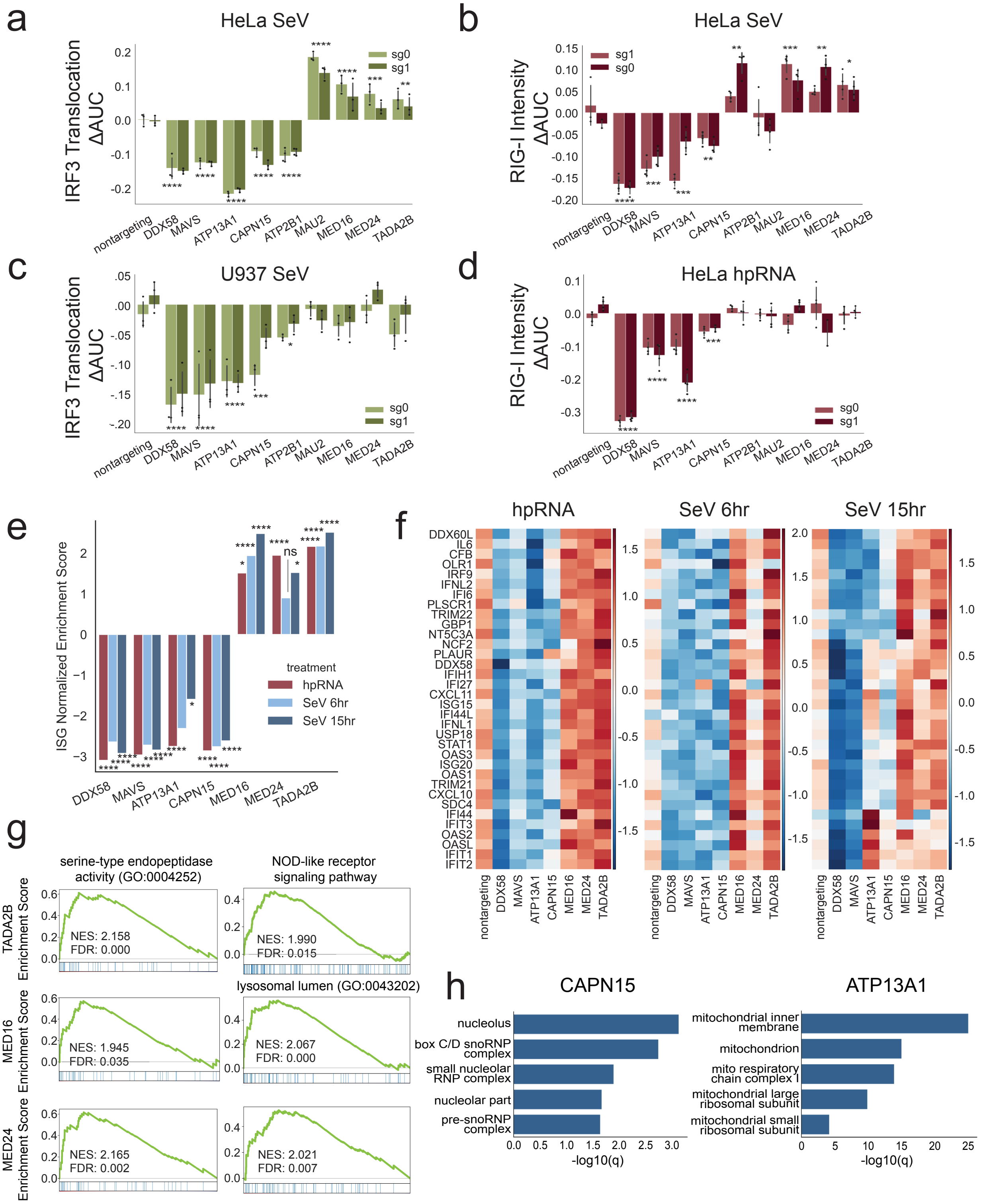

a
HeLa SeV
IRF3 Translocation ΔAUC
sg0
sg1
b
HeLa SeV
RIG-I Intensity ΔAUC
sg1
sg0
c
U937 SeV
IRF3 Translocation ΔAUC
sg0
sg1
d
HeLa hpRNA
RIG-I Intensity ΔAUC
sg0
sg1
nontargeting
DDX58
MAVS
ATP13A1
CAPN15
ATP2B1
MAU2
MED16
MED24
TADA2B
e
ISG Normalized Enrichment Score
treatment
hpRNA
SeV 6hr
SeV 15hr
ns
f
hpRNA
SeV 6hr
SeV 15hr
DDX60L
IL6
CFB
OLR1
IRF9
IFNL2
IFI6
PLSCR1
TRIM22
GBP1
NT5C3A
NCF2
PLAUR
DDX58
IFIH1
IFI27
CXCL11
ISG15
IFI44L
IFNL1
USP18
STAT1
OAS3
ISG20
OAS1
TRIM21
CXCL10
SDC4
IFI44
IFIT3
OAS2
OASL
IFIT1
IFIT2
g
serine-type endopeptidase activity (GO:0004252)
NOD-like receptor signaling pathway
lysosomal lumen (GO:0043202)
TADA2B Enrichment Score
MED16 Enrichment Score
MED24 Enrichment Score
NES: 2.158
FDR: 0.000
NES: 1.990
FDR: 0.015
NES: 1.945
FDR: 0.035
NES: 2.067
FDR: 0.000
NES: 2.165
FDR: 0.002
NES: 2.021
FDR: 0.007
h
CAPN15
nucleolus
box C/D snoRNP complex
small nucleolar RNP complex
nucleolar part
pre-snoRNP complex
-log10(q)
ATP13A1
mitochondrial inner membrane
mitochondrion
mito respiratory chain complex I
mitochondrial large ribosomal subunit
mitochondrial small ribosomal subunit
-log10(q)

**Figure 2.4.**

Arrayed validation and RNA sequencing of novel IRF3 translocation hits. (A) IRF3 translocation in HeLa cells in response to SeV stimulation for the indicated genes. * indicates adjusted $p < .05$, ** $p < .01$, *** $p < .001$, **** $p < .0001$ throughout. Two-sided p-values for a-d were calculated by computing the delta AUC for the feature of interest between cells in each well and non-targeting control cells (minimum n = 3,000); delta AUCs for all wells containing cells with an sgRNA of interest were then compared to non-targeting delta AUCs using a t test, p values were combined using Fisher's method and corrected using the Benjamini-Hochberg procedure. (B) RIG-I activation in HeLa cells in response to SeV stimulation for the indicated genes. (C) IRF3 translocation defects in U937 cells in response to SeV stimulation for the indicated genes. (D) RIG-I activation in HeLa cells in response to hpRNA stimulation for the indicated genes. (E) Normalized enrichment score (NES) computed via GSEA analysis for ISGs (genes significantly enriched in non-targeting control cells with transfected hairpin RNA relative to unstimulated non-targeting control cells). P-values were obtained from the gseapy package using 1000 permutations over all gene sets considered and corrected using the Benjamini-Hochberg procedure. (F) Heatmaps showing z-scored gene expression for ISGs in (E) in response to hpRNA or SeV treatment for 6 or 15 hours. (G) GSEA results for show increased serine-type endopeptidase activity (*TADA2B, MED16, and MED24*), NOD-like receptor signaling (*TADA2B*), and lysosomal lumen components (*MED16, MED24*). P-values were obtained from the gseapy package using 1000 permutations over all gene sets considered and corrected using the Benjamini-Hochberg procedure. (H). Enrichr results for genes that were decreased compared to positive controls *DDX58* and *MAVS* in response to SeV stimulation (15hrs) show significant enrichment of nucleolus components and small nucleolar RNAs for *CAPN15*, while *ATP13A1* had a significant decrease in mitochondrial components. Enrichr p-values are computed using the Fisher exact test and adjusted using the Benjamini-Hochberg procedure.

### *2.4.4 MAVS is mislocalized in the absence of functional ATP13A1*

We further tested the putative role of *ATP13A1* in regulating IRF3 localization by asking whether *ATP13A1* overexpression could rescue appropriate IRF3 translocation. Indeed, we observed restoration of the translocation defect we observed in the knockout when *ATP13A1* was overexpressed (**Fig. 2.5a**). *ATP13A1* encodes a P5A-type ATPase localized to the ER (Cohen et al., 2013; Sørensen et al., 2018) whose yeast homolog *spf1* has been shown to be involved in the unfolded protein response (Jonikas et al., 2009). However, in our mammalian HeLa *ATP13A1* knockout, we found no increased ER stress (**Supplementary Fig. 2.5c**). A putative transporter, it has been hypothesized to be an ion pump for divalent cations such as manganese, calcium, and magnesium (Cohen et al., 2013; Sørensen et al., 2015); however, adding divalent cations to the culture media did not rescue IRF3 translocation in ATP13A1 knockout cells (**Fig. 2.5b**). Recently, *ATP13A1* was shown to be a protein dislocase, responsible for the proper targeting of mitochondrial tail-anchored proteins (McKenna et al., 2020). Since MAVS is a tail-anchored mitochondrial protein critical for robust IRF3 translocation, we examined MAVS localization in *ATP13A1* knockout cells, finding that indeed, the fraction of mitochondrial and peroxisomal MAVS was reduced in *ATP13A1* knockout cells, while the fraction of MAVS in the ER was elevated (**Fig. 2.5c,d**). We further confirmed this effect by examining *MAVS* mitochondrial localization in hTERT-immortalized primary BJ1 cells by confocal microscopy, finding that MAVS mitochondrial localization was greatly reduced in *ATP13A1* knockouts as well as overall MAVS levels, in line with observations from a recent paper (McKenna et al., 2022) (**Fig. 2.5e,f**).

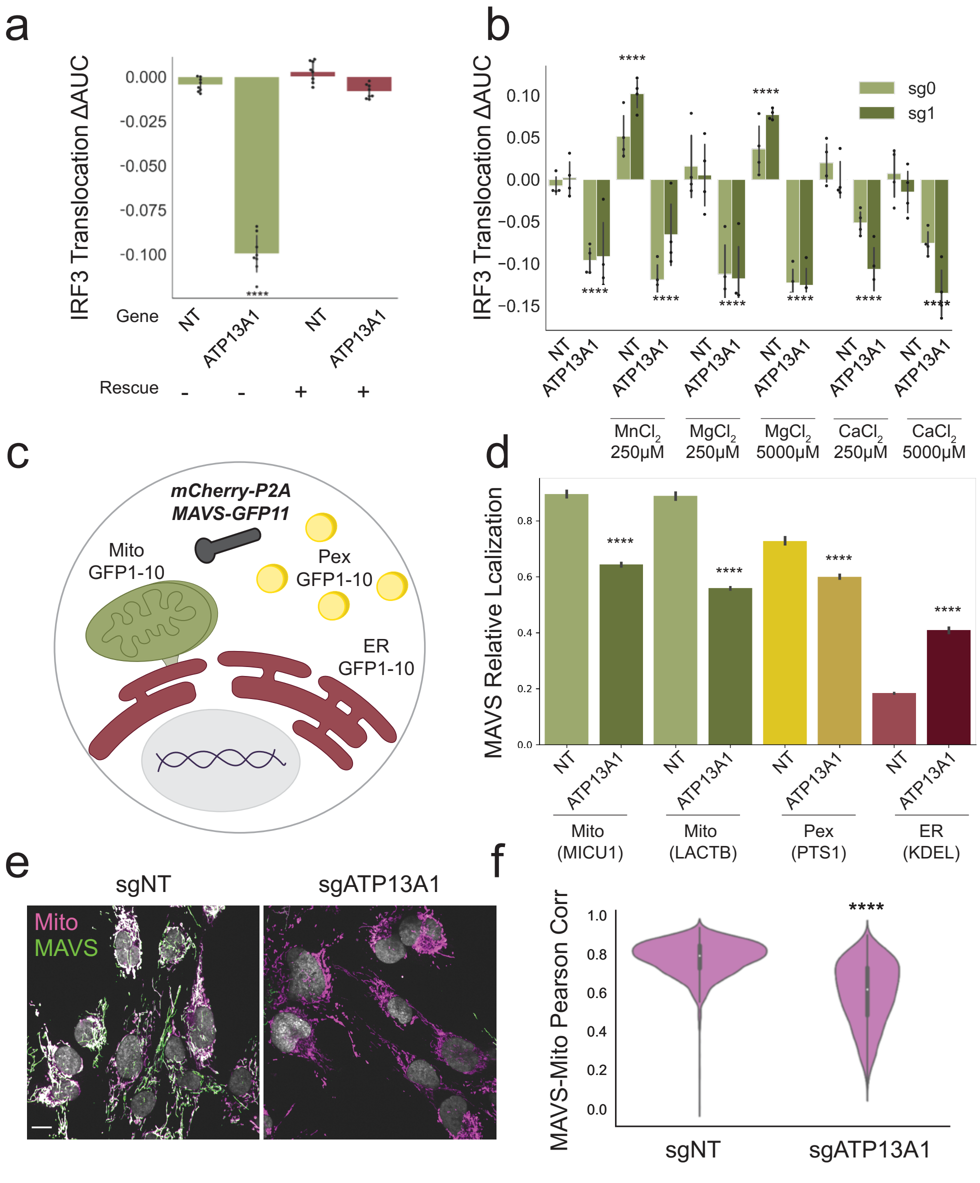


**Figure 2.5.**
MAVS is mislocalized in the absence of functional ATP13A1. (A) Overexpressed ATP13A1 rescues IRF3 translocation in ATP13A1 knockout HeLa cells. **** indicates $p < .0001$. Two-sided p-values for a, b, d, and f were calculated by computing the delta AUC for the feature of interest between cells in each well and non-targeting control cells (minimum n = 3,000); delta AUCs for all wells containing cells with an

sgRNA of interest were then compared to non-targeting delta AUCs using a t test, p values were combined using Fisher's method and corrected using the Benjamini-Hochberg procedure. (B) IRF3 translocation is not rescued by the addition of exogenous divalent cations in HeLa cells, cytokines added 24 hours prior to SeV infection. **** indicates p < .0001 (C) Diagram showing experimental setup for split-GFP experiment, where MAVS is overexpressed in HeLa cells along with either mitochondrial, peroxisomal, or ER GFP1-10. (D) In the presence of an ATP13A1 knockout, MAVS targeting to mitochondria and peroxisomes is reduced, while targeting to the ER is increased in HeLa cells. **** indicates p < .0001. (E) In hTERT-immortalized primary BJ1 cells. Scale bar 20 μm. (F) Quantification of experiment in (E), **** indicates p < .0001 using a t-test.

## 2.5 Discussion

In this study, we applied optical pooled screens at the genome-wide scale for the first time, scaling up our original optical pooled screening technology from screens of a few thousand genes to over 20,000. We used the resulting high-dimensional, single-cell resolution dataset to identify novel regulators of IRF3 subcellular localization. Secondary screens were highly correlated with primary screen data despite modest cellular sampling in the primary screen (median 484 cells per gene), highlighting the remarkable robustness and statistical power achievable with the optical pooled screening method at large scales. A sensitivity analysis (**Supplementary Fig. 2.1i**) revealed that for the IRF3 translocation phenotype, an effect size 50% that of strong positive controls could be detected with 200 cells/gene in a genome-wide screen (FDR adjusted p-value < .001), and 89% of genes in our screen had at least this number of cells. We also applied deep learning analysis to identify latent morphological features in our image-based genome-wide dataset, which helped identify three novel peroxisome regulators (*VHL, SEPHS1*, and *ZCCHC14*, validated via arrayed knockouts) and a group of RNA processing genes that alter the cytoplasmic localization of SeV.

We confirmed seven novel regulators of IRF3 translocation via arrayed experiments and investigated five of these regulators through RNA sequencing, confirming expected (signed) differences in ISG induction by SeV. Importantly, we also confirmed ISG induction in response to a synthetic hairpin RNA, showing that these genes indeed act to directly regulate IRF3 pathway induction rather than simply modulating SeV entry or replication. *ATP13A1*, a strong novel positive regulator that has recently been identified as a mitochondrial tail-anchored protein dislocase (McKenna et al., 2020), resulted in profound loss of IRF3 translocation upon knockout. We hypothesized and confirmed that MAVS, a mitochondrial tail-anchored protein critical for robust IRF3 translocation, is mislocalized in the absence of *ATP13A1*, which results in loss of MAVS from both mitochondrial and peroxisomal membranes with a concomitant increased localization to the ER. We also identified *CAPN15,* a cysteine protease and member of the *SOL* calpain family, as a novel positive regulator. Little is known about *CAPN15*, whose proteolytic targets have not been identified, but pathogenic variants of the gene result in microphthalmia and cataracts in humans (Mor-Shaked et al., 2021). Through RNA sequencing we found that snoRNP components were downregulated in the presence of the *CAPN15* knockout. This may decrease snoRNA production, thereby contributing to reduced activation of RIG-I and consequently, IRF3 (Chiang et al., 2018). Among novel negative regulators, we identified and validated two members of the Mediator complex, *MED16* and *MED24*. In addition to the known role of these Mediator components in transcriptional termination regulation (**Supplementary Fig. 2.3d**) (Takahashi et al., 2020), we found that knockouts lead to higher IRF3 translocation and ISG expression and further identified consequences of knockouts using RNA-sequencing and pathway analysis. Specifically, we found decreased SUMOylation activity, which negatively regulates IRF3/7 (Kubota et al., 2008), and increased expression of lysosomal lumen components, which can increase antiviral gene expression (Hasan et al., 2013).

In summary, we improved the imaging throughput of optical pooled screening by over four-fold and applied it at the genome-wide scale for the first time, discovering novel positive and negative regulators of viral sensing upstream of IRF3 translocation as well as modulators of SeV protein localization and peroxisomal biogenesis. Optical pooled screening was previously used to assay two to four channels but here we image as many as seven channels at a time, enabling a highly multidimensional assessment of spatially defined molecular and cell morphological features at the single-cell level. As an example, we took advantage of our multidimensional data to identify genes that altered IRF3 translocation without greatly affecting SeV load (**Fig. 2.2e**), as expected for genes that regulate IRF3 directly. We also identified genes that regulate SeV protein localization, providing a feasible approach to identify host factors affecting viral trafficking and the formation of viral replication foci. Single-cell multichannel images also enable advanced post-hoc analysis of the screening data, which we demonstrated here by applying a pre-trained neural network model and a second model trained via an autoencoder to extract unbiased feature sets that helped identify regulators of peroxisomes, and factors impacting mRNA processing and SeV localization. The implementation of widely applicable optical pooled screening protocols with increased throughput and phenotype dimensionality – as well as analysis techniques to process the complex datasets produced - will enable many discoveries to be made in future high-dimensional genome-wide optical pooled screens. The results will systematically link genes to important cellular functions and processes by leveraging high-content phenotypic readouts including morphology, subcellular molecule localization, and additional human- and machine-readable molecular and cellular features derived from high-resolution image measures.

## 2.6 Materials and Methods

*Library cloning, lentivirus production, transduction, and next generation sequencing of libraries*

Libraries were cloned as previously described (Feldman et al., 2022) into a CROP-seq-puro-v2 (Addgene #127458) backbone. Lentivirus was then produced and transduced as previously described (Feldman et al., 2022). For library transductions, multiplicity of infection was estimated by counting colonies after sparse plating and antibiotic selection. Genomic DNA was also extracted for NGS validation of library representation. Genomic DNA was extracted using an extraction mix as described above. Barcodes and sgRNAs were amplified by PCR from a minimum of 100 genomic equivalents per library using JumpStart 2X Master Mix (initial denaturation for 5 minutes at 98°C, followed by 28 cycles of annealing for 10 s at 65°C, extension for 25 s at 72°C, and denaturation for 20 s at 98°C).

*Sendai virus infection, phenotyping, and in situ sequencing for genome-wide IRF3 screen and secondary screen*

HeLa-TetR-Cas9 cells were transduced with BFP-PTS1 and a single clone expressing the reporter was selected. For screening, cells were selected with puromycin (1 µg/mL) for 3 days after transduction and library representation was validated by NGS. Cas9 expression was induced with 1 µg/mL doxycycline for 1 week and cells were then seeded in six 6-well glass-bottom dishes at 400,000 cells/well two days prior to fixation. Sendai virus (Cantell strain, ATCC VR-907) was added at an MOI of 10 in 400ul media/well for 45 minutes at 4C with rotation, the plate was quickly washed with media, and fresh media added, allowing the virus to replicate for fifteen hours prior to fixation. Thirty minutes before fixation, 50% of the media was exchanged, and MitoTracker DeepRed added to cells at 37°C at a final concentration of 7nM. Cells were fixed by removing media and adding 4% paraformaldehyde (Electron Microscopy Sciences 15714) in PBS for 30 minutes.

Peroxisomes and mitochondria were imaged prior to permeabilization with 100% ice-cold methanol for 20 minutes. The permeabilization solution was then carefully exchanged with PBS-T wash buffer (PBS + 0.05% Tween-20) by performing six 50% volume exchanges followed by three quick washes. Cells were stained with primary antibodies in 3% BSA (VWR Cat# 97061-422) in PBS overnight for pIRF3 (1:250, CST Cat# 29047, RRID:AB_2773013), RIG-I (1:650, AdipoGen Cat# AG-20B-0009, RRID:AB_2490189), and SeV (1:2500, Abcam Cat# ab33988, RRID:AB_777877) in 1:75 Ribolock, washed 3x in PBS-T, and stained with secondary antibodies for 1 hour with 1:1000 anti-rabbit AF532 (Thermo Fisher Scientific Cat# A-11009, RRID:AB_2534076), 1:1000 anti-mouse AF594 (Thermo Fisher Scientific Cat# A-21203, RRID:AB_141633), 1:300 anti-chicken Dylight 755 (Thermo Fisher Scientific Cat# SA5-10075, RRID:AB_2556655) in 1:200 Ribolock, washed 6x in PBS-T, and stained with YF488-conjugated anti-MDA5 (1:1000, custom conjugation of Proteintech Cat# 21775-1-AP, RRID:AB_10734593) in 1:200 Ribolock. Cells were then stained with DAPI (500ng/mL in 2x SSC with 1:75 Ribolock) and imaged.

After phenotyping, the sgRNA sequence was reverse transcribed in situ for 9.5 hours at 37°C using 1x RevertAid RT buffer, 250 μM dNTPs, 0.2 mg/mL BSA, 1 μM RT primer, 0.8 U/μL Ribolock Rnase inhibitor, and 4.8 U/μL RevertAid H minus reverse transcriptase in 750 μL/well. After reverse transcription, cells were washed 5x with PBS-T and post-fixed using 3% paraformaldehyde and 0.1% glutaraldehyde in PBS for 30 minutes, followed by washing with PBS-T 3 times. Samples were then incubated in a padlock probe and extension-ligation reaction mixture (1x Ampligase buffer, 0.4 U/μL Rnase H, 0.2 mg/mL BSA, 100 nM padlock probe, 0.02 U/μL TaqIT polymerase, 0.5 U/μL Ampligase and 50 nM dNTPs) for 5 minutes at 37°C and 90 minutes at 45°C, and finally washed 2 times with PBS-T. Circularized padlocks were amplified using a rolling circle amplification mix (1x Phi29 buffer, 250 μM dNTPs, 0.2 mg/mL BSA, 5% glycerol, and 1 U/μL Phi29 DNA polymerase) at 30°C overnight. *In situ* sequencing was performed as previously described using sequencing primer oSBS_CROP-seq for 12 cycles. Secondary screening was performed similarly to the genome-wide screen but phenotyping was restricted to DAPI, pIRF3, RIG-I, and SeV and 7 cycles of *in situ* sequencing were performed.

### *Fluorescence microscopy*

All in situ sequencing images were acquired using a Ti-2 Eclipse inverted epifluorescence microscope (Nikon) with automated XYZ stage control and hardware autofocus. An LED light engine (Lumencor CELESTA Light Engine) was used for fluorescence illumination and all hardware was controlled using NIS elements software. *In situ* sequencing cycles were imaged without filter switching to increase throughput (that is, one channel imaged across the entire plate prior to switching to the subsequent channel) using a 10X 0.45 NA CFI Plan Apo λ objective (Nikon) with the following filters (Semrock) and exposure times for each base: G (546 nm laser at 30% power, emission 575/30 nm, dichroic 552nm, 200 ms); T (546 nm laser at 30% power, emission 615/24 nm, dichroic 565 nm, 200 ms); A (637 nm laser at 30% power, emission 680/42 nm, dichroic 660 nm, 200 ms); C (637 nm laser at 30% power, emission 732/68 nm, dichroic 660 nm, 200 ms). Phenotyping images were also imaged without filter switching and acquired using a 20X 0.75 NA CFI Plan Apo λ objective (Nikon) with the following filters (Semrock unless otherwise noted) and exposure times: BFP-PTS1 (405 nm laser at 10% power, Chroma Multi LED set #89402, 200ms), DAPI (405 nm laser at 5% power, Chroma Multi LED set #89402, 50ms), CF488 (477 nm laser at 30% power, Chroma Multi LED set #89402, 200ms), AF532 (546 nm laser at 10% power, emission 575/30 nm, dichroic 552nm, 200ms), AF594 (546 nm laser at 10% power, emission 615/24 nm, dichroic 565 nm, 200ms), MitoTracker Deep Red (637 nm laser at 10% power, emission 680/42 nm, dichroic 660 nm, 200ms), Dylight 755 (749 nm laser at 10% power, emission 820/110 nm, dichroic 765 nm, 200ms).

*IRF3 reporter Sendai and VSV secondary screen*

For live-cell screening, HeLa-TetR-Cas9 cells were transduced with pTRIP-GFP-IRF3 (Addgene #127663). Fluorescent cells were sorted by FACS (Sony SH800) and re-sorted to select for cells with stable expression. Cells were selected with puromycin (1 µg/mL) for 3 days and Cas9 expression was induced with 1 µg/mL doxycycline for 1 week. Cells were then seeded onto 6-well cover glass-bottom plates 2 days prior to translocation experiments. Cells were stimulated with SeV for 15 hours or VSV at MOI 2.5 (ATCC Cat# VR-1238) for 18 hours prior to fixation with 4% paraformaldehyde for 30 minutes and initiation of the *in situ* sequencing protocol. After phenotyping, cells were fixed and the *in situ* sequencing protocol (reverse transcription, gap-fill, and rolling circle amplification) was performed, followed by 7 bases of sequencing-by-synthesis.

*Arrayed validation*

Novel top-ranking genes confirmed from the pooled secondary screen were validated with individual sgRNAs. For each gene, 2 sgRNAs were tested. HeLa-TetR-Cas9 cells expressing pTRIP-GFP-IRF3 (Addgene #127663) were prepared and assayed as in the pooled screen, except that assays were carried out in 96-well glass plates and cells were seeded at 8,000 cells/well. For hpRNA stimulation, cells were seeded in 96-well plates and transfected with 1µg/mL hpRNA (Invivogen tlrl-hprna-100) in 10µl/well Lyovec (Invivogen lyec-1) for 24hr. For IAV (ATCC Cat# VR-95) and RSV (ATCC Cat# VR-26PQ) infection, cells were treated with virus at an MOI of 1 for 15 hours. For arrayed validation of genes with peroxisome defects, HeLa-TetR-Cas9-BFP-PTS1 cells were used. For arrayed validation in U937 cells (ATCC Cat# CRL-1593.2, RRID:CVCL_0007), cells were first transduced with lentiCas9-Blast (Addgene #52962) and selected using 20 µg/mL (Thermo Fisher Scientific A1113903), then transduced with sgRNAs and selected using 2 µg/mL puromycin (Thermo Fisher Scientific A1113803) for 5 days. Prior to infection with SeV, U937 cells were plated in 96 well plates at 40,000 cells/well with 100ng/mL PMA (Invivogen tlrl-pma). After 24 hours, PMA was removed and fresh RPMI with 10% FBS, 1% PenStrep was added. Media was again exchanged 24 hours after PMA removal and, 41 hours after PMA removal, SeV was added at an MOI of 10 to the cells, which were fixed 48 hours after PMA removal.

### *BJ1 MAVS Localization*

For experiments with BJ1 cells (BJ-5ta - CRL-4001), cells were transduced with Cas9 and sgRNAs as described for U937 cells but with puromycin selection for 3 days. BJ1 cells were seeded at 8,000 cells/well and stained for mitochondria using MitoTracker Deep Red at 100 nM for 30minutes at 37°C prior to fixation. Cells were permeabilized with ice-cold Methanol for 20 minutes, washed with PBS-T, and stained with a YF488 conjugated MAVS (Proteintech Cat# 14341-1-AP, RRID:AB_10548408) at 1:300 for 2 hours before imaging. Cells were imaged on an Andor Dragonfly Confocal Microscope.

### *Split-GFP Localization Experiments*

Subcellular localization of MAVS was determined using a split-GFP approach as previously described (Hyun et al., 2015). Targeting sequences for *MICU1* (Le Vasseur et al., 2021)*, LACTB*, and the targeting sequences PTS1 and KDEL (Inglis et al., 2020) GFP1-10 lentiviral vectors were transduced into Cas9-expressing HeLa cells. Transduced cells were transiently transfected with *MICU1,* PTS1, or *VAMP2* GFP11 plasmids to identify cells expressing the mitochondrial, peroxisomal, or ER targeting constructs, respectively, which were isolated via flow cytometry. Cells were then transduced with a full-length mCherry *MAVS* GFP11 lentiviral vector and MAVS-positive cells were sorted based on mCherry expression. Finally, cells were transduced with a non-targeting or *ATP13A1* sgRNA in the CROP-seq-puro-v2 (Addgene #127458) backbone and selected via puromycin prior to plating and assessment of relative *MAVS* localization.

### *Indel Sequencing of Arrayed Knockouts*

Genomic DNA was extracted as previously described (Feldman et al., 2022) and targeted PCR performed to amplify target gene regions for each sgRNA. Following next-generation sequencing, editing rates were assessed using CRISPResso version 2.038 (Pinello et al., 2016) with parameters –flexiguide_homology 80–quantification_window_size 10 –quantification_window_center -3 –exclude_bp_from_left 0 –exclude_bp_from_right 0

### *Western Blotting of ATP13A1 and MAVS*

Cells were pelleted and resuspended in lysis buffer (150 mM NaCl, 1% Triton X-100, 50 mM Tris-HCl pH 7.5 with one protease inhibitor tablet added to 10 mL buffer, Sigma cat #11873580001) at 1 million cells per 100 µl. Following incubation for 30 minutes at 4°C, lysates were centrifuged at 10,000xg for 5 minutes and supernatant was stored at -80°C. Lysates were then thawed on ice and Laemmli 6x sample buffer, SDS, reducing (Boston Bioproducts Cat #BP-111R) was added prior to loading on NuPAGE 4-12% Bis-Tris Gels (Thermo Fisher Cat #NP0321). Samples were transferred to nitrocellulose membranes using iBlot2 (Thermo Fisher) and blocking was performed in TBS Tween with 5% non-fat milk. Primary antibodies (1:500 ATP13A1 Proteintech Cat #16244-1-AP and 1:1000 MAVS Proteintech Cat #14341-1-AP) were incubated with samples overnight and secondary rabbit antibody was added for 1 hour prior to visualization with chemiluminescence.

### *RNA-seq analysis of clonal knockout cell lines*

Clonal knockout cell lines and wild-type control cells were plated in 96-well format at a density of 8,000 cells/well. After 24 hours, cells were transfected with 1µg/mL hpRNA (Invivogen tlrl-hprna-100) in 10µl/well Lyovec (Invivogen lyec-1) while cells infected with Sendai virus at an MOI of 10 were infected after 33 and 42 hours for 15 hour and 6 hour Sendai virus incubations. After 48 hours, cells were washed with PBS, and lysed at room temperature for 5 minutes in 21 µl TCL buffer (QIAGEN, supplemented with 1% beta-mercaptoethanol) per well, then stored at –80°C. Smart-seq2 was performed as described (Trombetta et al., 2014). Libraries were sequenced on a NovaSeq 6000 (Illumina) using the v1.0 S4 300-cycle kit (paired end, 150 cycles forward, 150 cycles reverse).

### *Quantification and Statistical Analysis*

### *Image analysis*

Images of cell phenotype and in situ sequencing of perturbations were manually aligned during acquisition using nuclear masks to calibrate the plate position to each of the four corner wells during screening. Alignment was then refined computationally via cross-correlation of DAPI across cycles or between individual imaging channels within sequencing cycles, since these were acquired separately. For SBS and phenotyping, alignment across channels and imaging acquisitions (12 SBS cycles and 2 phenotyping acquisitions) was performed using cross-correlation and a high upsample factor (5) to ensure alignment precision. Unlike our previous work, this alignment was then further refined using ORB feature detection (skimage.feature.ORB, n_keypoints = 200, fast_threshold = 0.05) and RANSAC to filter detected ORB features (skimage.measure.ransac(EuclideanTransform, min_samples = 2, residual_threshold = 1, max_trials = 200). This allowed for refinement of image alignment and additional image rotation unlike the cross-correlation algorithm which allowed translation alone. Nuclei and cells were detected and segmented as previously described (Feldman et al., 2019). IRF3 translocation was quantified by calculating pixel-wise correlations between the nuclear DAPI channel and the IRF3 channel (reporter or antibody staining). Mitotic cells were removed based on maximum DAPI signal and cell area. *In situ* sequencing read calling was performed as previously described (Feldman et al., 2019). Data analysis functions were written in Python, using Snakemake for workflow control (Köster & Rahmann, 2012).

### *Optical pooled screen analysis*

Cells were segmented by thresholding the DAPI signal to identify nuclei and expanding the resulting regions using the watershed method as previously described (Feldman et al., 2019). Only cells with a minimum of one read matching a barcode in the library were analyzed. Mitotic or apoptotic cells were removed by filtering cells with unusually high or low nuclear/cell area and DAPI signal. IRF3 translocation was calculated by determining the pixel-wise correlation between DAPI and the IRF3 antibody signal or reporter signal within the segmented nuclear area. For the genome-wide screen, only genes with a minimum of one read matching an sgRNA in the library and 2 sgRNAs with at least 30 cells/sgRNA were considered for analysis. Features were normalized on a per-cell basis relative to non-targeting control cells in the same well by subtracting the median for non-targeting control cells and dividing by the non-targeting control MAD x 1.4826 (M.-A. Bray et al., 2016) and scores for features relative to non-targeting controls were determined by calculating differences in cumulative AUCs (shaded area in **Fig. 2.1f**). These delta AUCs were averaged over sgRNAs for a given gene and significance was determined by comparing delta AUCs for individual sgRNAs to distributions bootstrapped from non-targeting control cells (bootstrapped 100,000 times). Gene-level p-values were aggregated across sgRNAs using Fisher's method and then corrected using the Benjamini-Hochberg procedure.

### *Deep learning analysis*

For both autoencoder and transfer learning, 46x46 single-cell cropped images of individual channels were generated for cells whose center was at least 13 pixels away from the nearest cell center. For transfer learning, these single-cell crops were then resized to 299x299 images and each channel was repeated three times to generate 299x299x3 images of the size required to extract features using the Xception network model (Chollet, 2017) provided by Keras. Image intensities were renormalized to be between 0 and 255 and features were extracted from the final layer ('avg_pool') of the Xception network with a feature size of 2048. Features with less than 0.1 coefficient of variation were dropped before further analysis. Only genes with >=30 cells/sgRNA >=2 sgRNAs were considered for further analysis.

The autoencoder model was adapted from a previously published model (K. D. Yang et al., 2020) and the latent feature size was set to 2048. The model was run separately for each channel of the 7-channel image dataset for up to 250 epochs and the model with the lowest test loss was selected for each channel. The model was trained and tested on 1% of the cells randomly selected from the larger set with 0.25% of the cells as test dataset, and 0.75% were used for training. The trained model was then used to extract latent features from the entire dataset. 2048-length feature vectors for each cell were then combined as described for the transfer learning.

PHATE clustering was performed with knn = 5 and other parameters at default settings and Leiden clustering performed with resolution = 3.

### *Arrayed validation*

For each sgRNA, the IRF3 translocation or RIG-I activation change was computed relative to non-targeting control cells by computing the difference in AUCs between the cumulative distributions for non-targeting and sgRNA of interest. Delta AUCs for 3-4 replicate well per sgRNA in each experiment were then compared using a t-test and corrected using Benjamini-Hochberg procedure to assess statistical significance.

### *RNA-seq quantification*

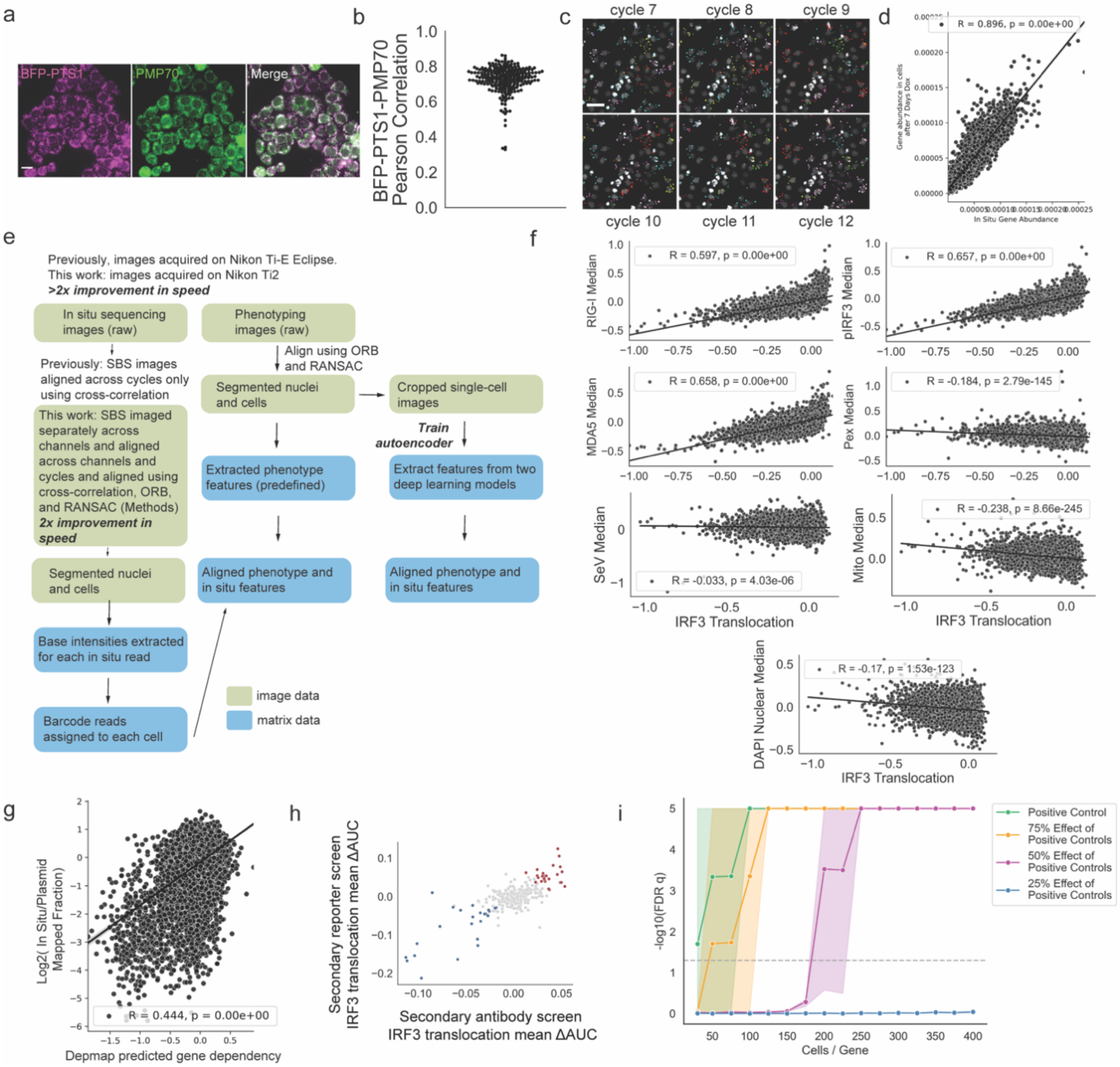
Kallisto was used to quantify transcript abundance using AB856846.1 for the Sendai genome. EdgeR was used to assess differential gene expression with default parameters for the estimateDisp and exactTest functions (N. L. Bray et al., 2016; Robinson et al., 2010). Only transcripts with cpm > 1 in at least two samples were considered. Biological replicates were defined as replicate stimulations with at least two sgRNAs for novel genes affecting IRF3 translocation. GSEA was performed on VST-transformed data using the DESeq2 package (Love et al., 2014). ISGs were defined as genes with minimum expression >3 logCPM that also demonstrated logFC > 3 and FDR < 0.01 between non-targeting unstimulated and non-targeting stimulated with synthetic hpRNA for 24hr.

## *Data and Code Availability*

Code is available at https://github.com/beccajcarlson/IRF3OpticalPooledScreen. RNA-seq data has been deposited in the Gene Expression Omnibus at GSE179288. Image data is publicly available on Google Cloud at gs://opspublic-east1/IRF3OpticalPooledScreen and one example field of view has been uploaded to the GitHub repository. The BFP-PTS1 plasmid has been deposited to Addgene (#199433).

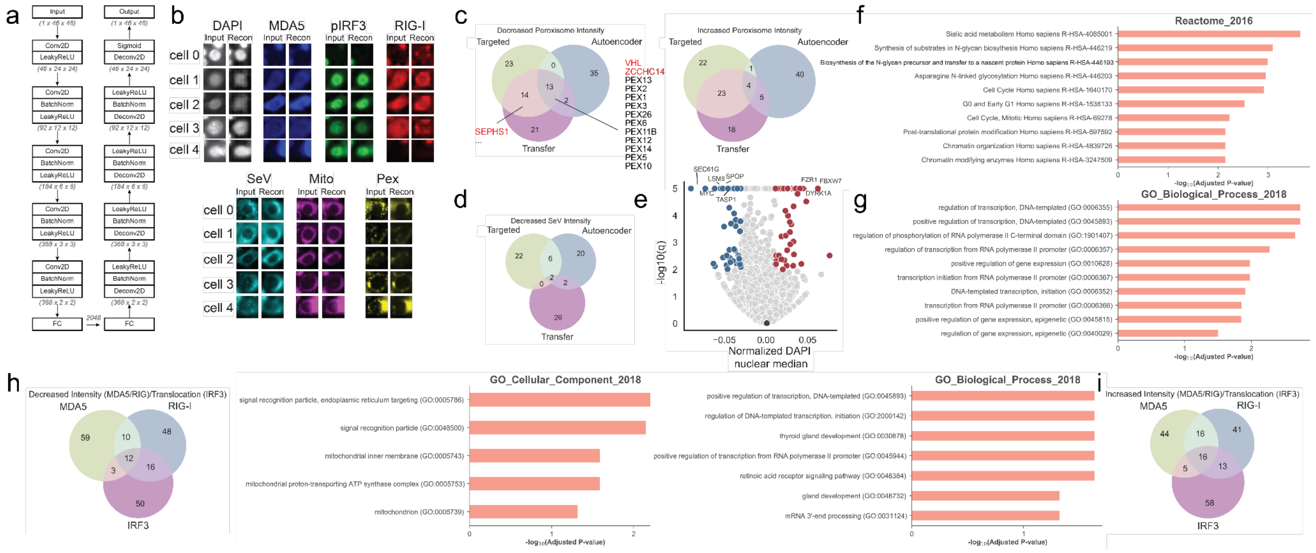
.

## 2.7 Supplementary Figures

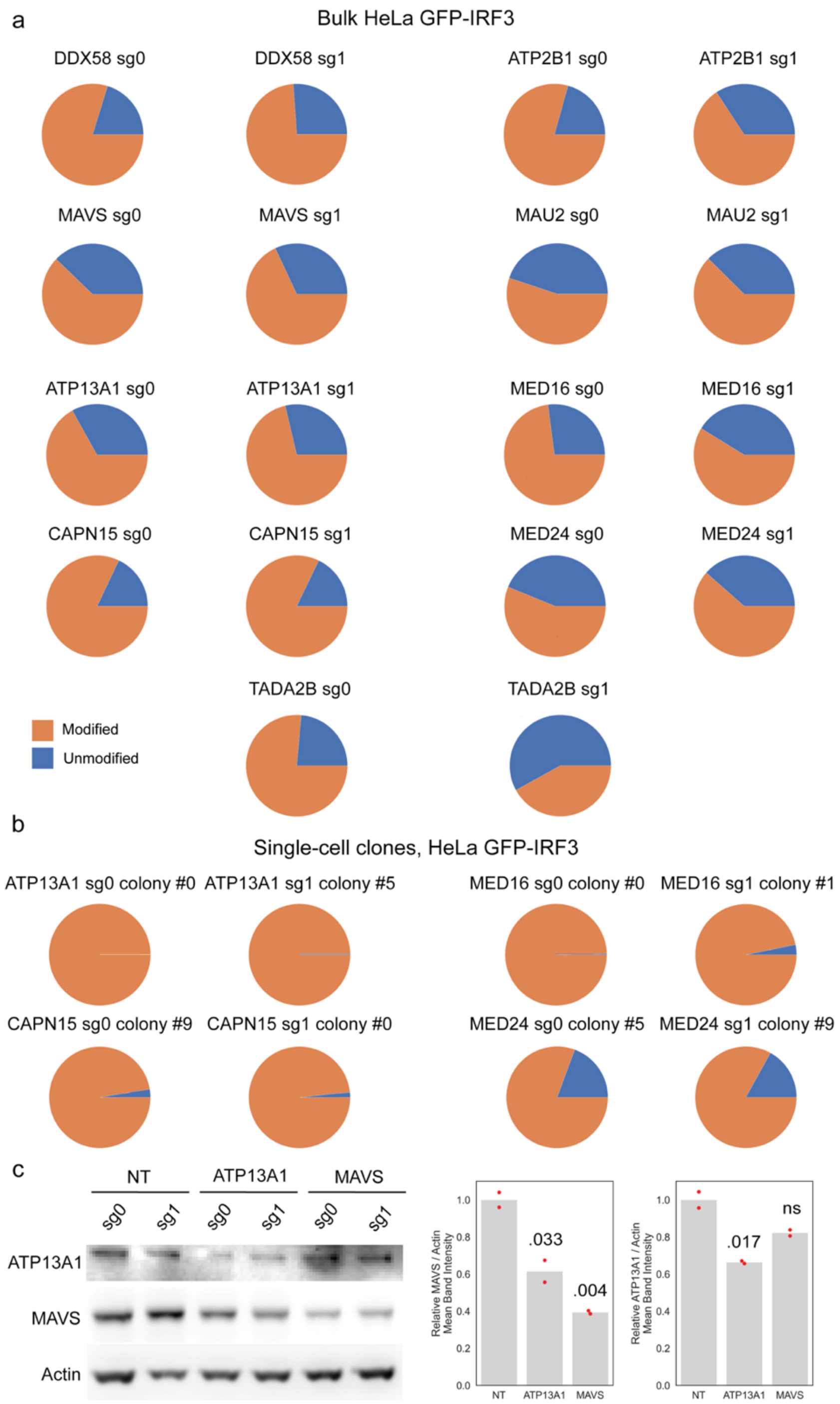


**Supplementary Figure 2.1.**

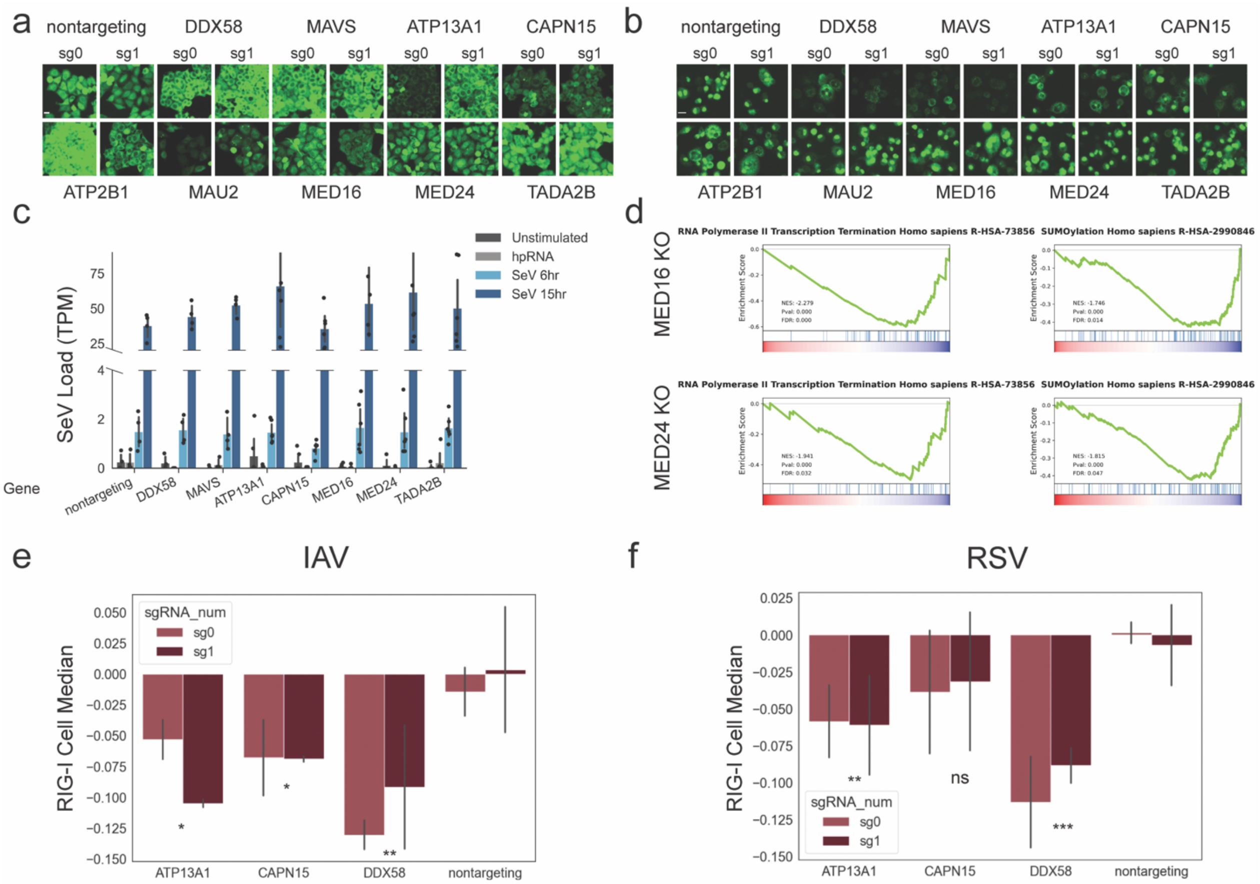
(A) BFP-PTS1 in HeLa cells significantly overlaps with PMP70 antibody staining. Scale bar 20 μm. (B) Quantification of BFP-PTS1-PMP70 correlation in HeLa cells. (C) Remaining *in situ* sequencing cycles from Fig. 2.1d. (D) *In situ* gene abundance (combined sgRNA abundance) is highly correlated with gene abundance in cells after 7 days of doxycycline treatment. Pearson r correlation and two-tailed p value are denoted on plot (n = 20,389 genes). (E) Genome-wide screen image acquisition and analysis workflow, allowing for a >4-fold improvement in throughput over our previous optical pooled screening design. See Methods for further details. (F) RIG-I, pIRF3, and MDA5 intensities are significantly correlated with pIRF3 translocation levels, while peroxisome, mitochondria, and Sendai intensities are not positively correlated, indicating coordinated pathway responses. Pearson r correlations and two-tailed p values are denoted on plots (n = 19,093 genes). (G) Predicted Depmap dependency (1, 2) is significantly correlated with the *in situ* mapped gene-level fraction relative to the plasmid library mapped fraction. Pearson r correlation and two-tailed p value are denoted on plot (n = 15,917 genes). (H) Correlation between secondary screen translocation scores is significant (Pearson r = 0.74, two-tailed p = 4.7e-60, n = 342 genes). (I) Sensitivity analysis showing number of cells per gene required to significantly identify genes with effect sizes compared to those for strong positive controls *DDX58, MAVS,* and *IRF3.*



**Supplementary Figure 2.2.**
(A) Architecture for an autoencoder used to extract 2048 features from images. (B) Example inputs and reconstructions for 5 single cells across 7 channels of interest. (C) Venn diagram showing overlap of top 50 genes using a targeted feature (decreased peroxisome intensity), autoencoder, or transfer learning. Overlapping genes are enriched for peroxisome biogenesis genes as well as *VHL,* and *ZCCHC14,* which were in the top 50 using all three feature sets, while *SEPHS1* was in the top 50 for all but the autoencoder features. (D) Venn diagram as shown in (C) but for genes that increased peroxisome intensity. (E) Volcano plot for genes affecting the DAPI nuclear median. Two-sided p-values were calculated as described in the Methods and corrected using the Benjamini-Hochberg procedure (n = 19,900 genes). (F) Significant Reactome terms for genes among the top 50 for decreased DAPI intensity in targeted, autoencoder, and transfer feature sets. Enrichr p-values are computed using the Fisher exact test and adjusted using the Benjamini-Hochberg procedure. (G) Significant GO BP terms for genes among the top 50 for increased DAPI intensity in targeted, autoencoder, and transfer feature sets. Enrichr p-values are computed using the Fisher exact test and adjusted using the Benjamini-Hochberg procedure. (H) Venn diagram of top genes from all feature sets showing decreased intensity (for MDA5/RIG-I) or translocation (IRF3) shows distinct genes influencing the receptors and IRF3. Genes uniquely affecting IRF3 are enriched for GO CC terms relating to mitochondria and ER, organelles likely affecting signaling downstream of PRRs, while genes uniquely decreasing RIG-I intensity are enriched for GO BP terms that include retinoic acid signaling, which the receptor is responsive to. Enrichr p-values are computed using the Fisher exact test and adjusted using the Benjamini-Hochberg procedure. (I) Venn diagram as in (H) showing genes that increased intensity or translocation for MDA5/RIG-I or IRF3, respectively.

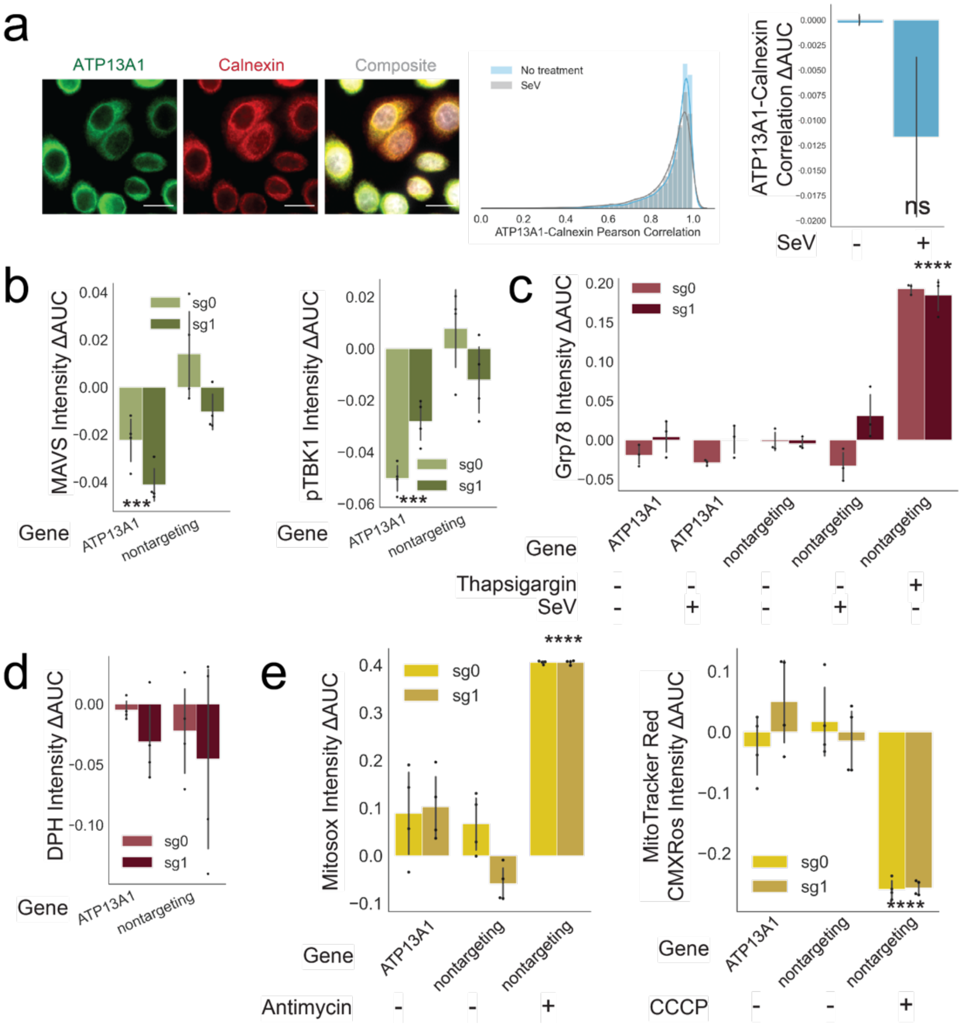


**Supplementary Figure 2.3.**
(A) Crispresso analysis of amplicon sequencing for bulk knockouts of 10 genes in HeLa GFP-IRF3 cells shows a high proportion of edited sequences. (B) Crispresso analysis of edited sequence fraction for single-cell clones used in bulk RNA sequencing experiment. (C) Western blot shows a decrease in ATP13A1 and MAVS protein levels in the respective knockouts, p values calculated using Student's t-test.



**Supplementary Figure 2.4.**

(A) Representative images of cells from Fig. 2.4a. Scale bar 20 μm. (B) Representative images of cells from Fig. 2.4c. Scale bar 20 μm. (C) Sendai virus load measured by RNA-seq is not significantly different in genetic knockouts relative to non-targeting controls at either 6 hours or 15 hours post-infection. Significance was determined by assessing Benjamini-Hochberg adjusted p values from t tests between non-targeting and each sgRNA for each condition. (D). *MED16* and *MED24* GSEA results show significant decrease in RNA polymerase II transcription termination and SUMOylation. P-values were obtained from the gseapy package using 1000 permutations over all gene sets considered and corrected using the Benjamini-Hochberg procedure. (E) RIG-I intensities upon IAV infection in perturbed cells relative to non-targeting controls. Two-sided p-values for e/f were calculated by computing the delta AUC for RIG-I per-cell median intensity between cells in each well and non-targeting control cells (minimum n = 3,000); delta AUCs for all wells containing cells with an sgRNA of interest were then compared to non-targeting delta AUCs using a 2-sided t test, p values were combined using Fisher's method and corrected using the Benjamini-Hochberg procedure; * indicates p < .05, ** indicates p < .01. (F). RIG-I intensities upon RSV infection in perturbed cells relative to non-targeting controls; ** indicates p < .01, *** indicates p < .001.



**Supplementary Figure 2.5.**

(A) An ATP13A1 overexpression vector in HeLa cells is localized to the ER, as evidenced by colocalization with calnexin and localization is not significantly different in the presence of Sendai virus infection (normalized to median of no-infection condition). Two-sided p-values for a-e were calculated by computing the delta AUC for the feature of interest between cells in each well and non-targeting control cells (minimum n = 3,000); delta AUCs for all wells containing cells with an sgRNA of interest were then compared to non-targeting delta AUCs using a t test, p values were combined using Fisher's method and corrected using the Benjamini-Hochberg procedure. (B) MAVS and pTBK1 are decreased in HeLa cells in the presence of an ATP13A1 knockout during Sendai virus infection (C) Grp78 intensity is not increased in ATP13A1 knockout cells relative to non-targeting control cells in the presence or absence of Sendai virus treatment. 1 µM thapsigargin treatment for 24 hours was used as a positive control (D) Lipid droplet content, measured by 2 µM DPH staining, is not significantly different in ATP13A1 knockout cells relative to non-targeting control cells. (E). Mitochondrial ROS (Mitosox) and membrane potential (MitoTracker Red CMXRos) are not significantly altered in ATP13A1 cells relative to non-targeting control cells. 120 µM antimycin A and 200 µM CCCP treatments for 1 hour served as positive controls.

# 3 Single-cell image-based genetic screens systematically identify regulators of Ebola virus subcellular infection dynamics

Authors: Rebecca J. Carlson,*, Justin Patten*, George Stefanakis, Brian Soong, Adit Radhakrishnan, Avtar Singh, Daisy Leung, Naveen Thakur, Christopher F. Basler, Gaya Amarasinghe, Caroline Uhler, Nir Hacohen, Robert Davey^, Paul C. Blainey^

## 3.1 Abstract

Ebola virus (EBOV) is a high-consequence filovirus that results in frequent epidemics with high case fatality rates and few therapeutic options. Here, we apply genome-wide image-based genetic screening to systematically identify regulators of Ebola virus infection in nearly 40 million single cells. We identify hundreds of regulators of viral infection and validate key genes in multiple cell lines and related filovirus species. Using a deep learning model trained on single-cell images, we show that UQCRB knockout is a post-entry regulator of Ebola virus replication, and demonstrate that inhibition of this gene with an FDA-approved small molecule reduces overall Ebola virus infection. In addition, we use a random forest model to identify STRAP, a spliceosome-associated factor, as a differential regulator of Ebola RNA and protein levels.

## 3.2 Introduction

Ebolaviruses such as Zaire Ebola virus (EBOV), Sudan virus (SUDV), and the related Marburg virus (MARV), are single-stranded, negative-sense filoviruses responsible for epidemics with high case fatality rates of up to 66%, predominantly in west or equatorial Africa (Ilunga Kalenga et al., 2019; Lo et al., 2017). EBOV infection results in Ebola virus disease (EVD), characterized by inflammatory responses, immunosuppression, and major fluid losses (Malvy et al., 2019). Monoclonal antibody therapy has demonstrated limited efficacy, reducing EVD case fatality rates to 30% (Mulangu et al., 2019), while a recently approved live-attenuated recombinant VSV vaccine expressing EBOV glycoprotein (GP) resulted in variable protection (Heppner et al., 2017; Mulangu et al., 2019) and no vaccine against MARV and SUDV currently exists.

Genetic screens are high-throughput methods that enable identification of host targets that modulate viral infection. However, due to the challenges of screening high-consequence viruses, previous genetic screens for Ebolaviruses often relied on use of pseudotyped viruses (Bruchez et al., 2020; Carette et al., 2011; Cheng et al., 2015), limiting discovery of post-entry viral modulators, or on overexpression of viral proteins via plasmids in a minigenome screen (Martin et al., 2018), which may not fully recapitulate the live Ebola virus life cycle. Two genome-wide screens have used live Ebola virus (Filone et al., 2015; Flint et al., 2019); however, these both assessed virus-induced cell death, a simple phenotypic output that is biased towards recovery of strong entry inhibitors. For instance, CAD, an EBOV genome replication regulator identified in a minigenome screen (Martin et al., 2018), was not recovered in any EBOV survival screens (Flint et al., 2019). Furthermore, previous studies with live virus had limited power, as Filone et al. only reported results for a handful of genes and Flint et al. found only 8 genes with an FDR < 0.3. Genetic screens based on host cell survival following viral infection capture a one-dimensional compressed view of the effects of genetic knockouts on the host-viral interplay, yielding an ordered list of factors that alter host cell survival upon infection that is challenging to interpret and leverage for further investigation.

Recently, we developed optical pooled screens (OPS) (Feldman et al., 2019), an approach that enables image-based pooled genetic screens in tens of millions of cells (Carlson et al., 2023; Funk et al., 2022) and which yields high-resolution images of single cells paired with each cell's genetic perturbation identity via targeted *in situ* sequencing. Unlike survival screens, these image-based genetic screens result in immediate insights into gene function and co-regulation (Carlson et al., 2023; Funk et al., 2022). Here, we present the first genome-wide multiparametric genetic screen for Ebola virus, representing an extensive landscape of the effects of hundreds of host genes on Ebola replication. We apply machine learning and deep learning approaches to our image-based dataset of nearly 40 million single cells and identify regulators of distinct stages of the Ebola lifecycle, from entry, to inclusion body formation, and viral RNA transcription and replication. We then perform targeted follow-up screens in two cell lines across three related filoviruses (EBOV, MARV, and SUDV) to robustly validate and contextualize our results and interpretation of cell- and virus-specific effects. We further show that UQCRB, a gene identified as a post-entry regulator of Ebola infection by deep learning, can be targeted by an FDA-approved small molecule with therapeutic potential, resulting in decreased viral replication in cells. Finally, we show that STRAP is a spliceosome-associated factor that increases viral RNA relative to protein levels upon knockout.

## 3.3 Results

### *3.3.1 A Genome-Wide Image-Based Genetic Screen Reveals Regulators of Distinct Indicators of Ebola Virus Infection*

Ebola virus is taken up into cells via macropinocytosis within endolysosomes. Following acidification, cathepsins from host cells process the viral glycoprotein (GP) into an active form that binds to the intracellular host receptor NPC1, resulting in membrane fusion and cytoplasmic release of viral nucleocapsids (Hoenen et al., 2019). Primary transcription and production of viral proteins leads to formation of inclusion bodies, cytoplasmic foci that serve as sites for viral replication (Hoenen et al., 2019; Nanbo et al., 2013). The Ebola nucleocapsid protein (NP) induces formation of inclusion bodies (IBs) (Miyake et al., 2020), sites critical for RNA synthesis (Hoenen et al., 2012). The viral polymerase cofactor VP35 interacts and colocalizes with NP to regulate IB formation (Miyake et al., 2020). At later stages of infection, these proteins exhibit a diffuse cytoplasmic localization pattern and, finally, localize to the cell periphery during virus budding (Nanbo et al., 2013). These distinct stages of infection are not readily distinguishable by the use of simple intensity-based measurements such as pooled flow cytometry-based screens, and genome-scale arrayed image-based screens are prohibitively costly and labor-intensive.

Here we applied OPS to gain an unbiased view of Ebola virus subcellular replication dynamics at the genome-wide scale, obtaining nearly 40 million single-cell images of Ebola-infected HeLa-TetR-Cas9 cells transduced with a custom pool of ~80,000 sgRNAs targeting ~20,000 genes, including 454 non-targeting control sgRNAs (**Fig. 3.1a**). Cells were infected with live Ebola Zaire virus at BSL4; subsequently, we assayed the Ebola VP35 protein using immunofluorescence (IF) and a previously described antibody (Mori et al., 2022) and the Ebola VP35 positive-sense RNA, which predominantly represents Ebola mRNA transcripts (Galão et al., 2022), using fluorescence *in situ* hybridization (FISH) (**Supplementary Fig. 3.1a**). Ebola infections were optimized to obtain >90% infection rates as measured by both VP35 protein and RNA levels (**Supplementary Fig. 3.1b**). In addition to measuring viral protein and RNA, we stained for the host transcription factor c-Jun, whose activity is increased in Ebola- and Marburg-infected cells (Hölzer et al., 2016; Wynne et al., 2017), as well as LAMP1, a lysosomal protein, and vimentin for cell segmentation (**Fig. 3.1b**). To enable measurement of viral RNA transcripts, we optimized detection of RNA molecules via FISH in combination with our *in situ* sequencing workflow for the first time (**Supplementary Fig. 3.1a**). To do so, we hybridized primary probes to viral RNA transcripts prior to targeted reverse transcription of sgRNAs. We then amplified viral RNA signal via hybridization chain reaction (HCR) as previously described (Choi et al., 2018) but with the omission of dextran sulfate, which inhibits the reverse transcriptase and polymerase activity critical for the sgRNA *in situ* sequencing reaction (Viennois et al., 2013). As expected, cells receiving sgRNAs targeting the Ebola virus receptor NPC1 demonstrated robust loss of VP35 RNA and protein signal, as well as loss of the nuclear host transcription factor c-Jun signal (**Fig. 3.1c, d**).

We identified 998 genes with significant changes in overall Ebola virus infection (FDR-adjusted p-value < 0.0001) as measured by VP35 protein intensity, using the difference in cumulative AUCs (delta AUC) between per-cell VP35 median intensity for each sgRNA relative to non-targeting control cells as a robust metric that flexibly captures shifts across the distribution of single-cell phenotype scores (**Fig. 3.1e**). As expected, knockout of the essential Ebola virus receptor NPC1 resulted in the strongest loss of VP35 protein. In addition, other genes previously shown to be required for Ebola virus infection were identified, including all six members of the HOPS complex which regulates endolysosomal fusion (all six scored in the top 60 genes). Only one previous Ebola screen recovered all six members (Carette et al., 2011), highlighting the robustness of our screen and its sensitivity in the genome-wide context. CTSB and CTSL, required for GP processing prior to NPC1 binding, and other previously identified Ebola entry regulators SPNS1, GNPTAB, UVRAG, PIKFYVE, FIG4, and EXT1 also scored (Carette et al., 2011; Cheng et al., 2015; Filone et al., 2015; Flint et al., 2019), as well as CAD, an enzyme critical for pyrimidine biosynthesis that was previously only identified in a minigenome screen (Martin et al., 2018).

Our analysis of cellular VP35 protein levels identified many genes and complexes not previously reported to regulate Ebola infection (**Supplementary Fig. 3.1c**), including PIK3C3, TIMM10, a member of the mitochondrial inner membrane translocase previously shown to interact with VP40 (Batra et al., 2018), the entire conserved oligomeric Golgi complex (COG1-8), the entire GET complex (GET1-4), the entire GARP complex (VPS51-54), retromer (VPS26A, 29, 35; previously shown to be required for pseudotyped Ebola entry) (Poston et al., 2022), the CIA complex (MMS19 and CIAO2B), genes involved in heparan sulfate synthesis (NDST and UGDH, likely required for virus attachment to cells (O'Hearn et al., 2015), as well as poorly characterized proteins TM9SF2 and PTAR1 (**Fig. 3.1f**). Unlike previous survival-based genetic screens, our intensity-based assay also enabled recovery of 57 negative regulators, none of which were previously shown to negatively regulate Ebola infection. Negative regulators were enriched for chaperones (HSP90B1, MESDC2, UNC45A, previously shown to interact with Ebola (J. Fang et al., 2022), integrin-related genes (such as ITGB1, ITGB5, and ITGAV), mRNA deadenylases (CNOT10, 11) and SAGA complex members (TADA2, TAF5L, TAF6L, TADA2B, SUPT20H), which modulate replication of other RNA viruses (Carlson et al., 2023; Guo et al., 2020) (**Fig. 3.1f**).

Given that our image-based pooled screen produced much additional multi-parametric single-cell information, we next sought to identify genes that differentially regulated nuclear levels of host c-Jun or VP35 RNA levels relative to VP35 protein levels, reasoning that such multi-parametric analysis may support more detailed functional inference of hits than the VP35 protein intensity analysis alone. We trained random forest regression models on features from single cells for each sgRNA in our screen, predicting either the c-Jun translocation relative to hundreds of single-cell VP35 protein and RNA features (**Fig. 3.1g, h**) or VP35 RNA levels relative to hundreds of single-cell VP35 protein features (**Fig. 3.1i**). This approach enabled us to systematically identify factors that separately regulate these three indicators of Ebola virus infection. As expected, factors where the c-Jun protein levels were poorly predicted by viral VP35 and RNA levels included Jun itself as well as known MAP kinase pathway members (**Fig. 3.1h**). The gene with the strongest reduction in VP35 RNA relative to VP35 protein levels was the mRNA binding protein PABPC1, which interacts with EBOV NP (García-Dorival et al., 2016; Morwitzer et al., 2019) and was downregulated in EBOV-infected NHP monocytes (Kotliar et al., 2020). A number of other genes with related functions based on STRING analysis also decreased Ebola RNA relative to protein levels (**Fig. 3.1j**), while only two genes, SRPK1 and STRAP, increased Ebola RNA relative to protein (**Fig. 3.1i**). Notably, SRPK1 was previously shown to regulate Ebola transcription through phosphorylation of VP30 (Takamatsu et al., 2020).

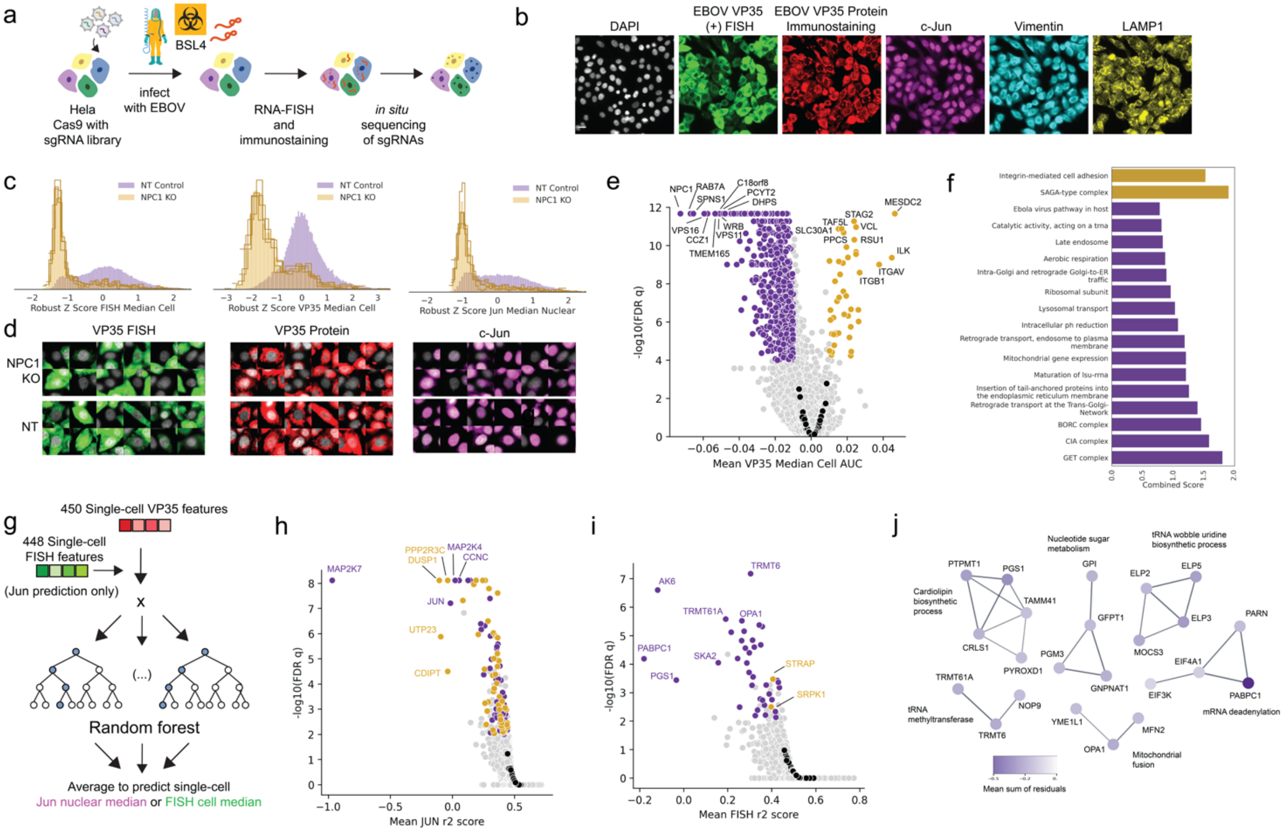


**Figure 3.1.**
**Genome-wide optical pooled screening reveals regulators of multiple responses to Ebola virus infection.** (A) Workflow for genome-wide optical pooled screen. (B) Example images of infected cells assayed across six distinct markers in the screen. Scale bar 20 µm. (C) Histograms of VP35 RNA FISH, VP35 protein, and c-Jun transcription factor intensity levels in non-targeting or NPC1 KO cells from the screen, each histogram trace represents a distinct sgRNA targeting NPC1. (D) Randomly selected cells transduced with a non-targeting sgRNA or NPC1-targeting sgRNA show reduced levels of VP35 RNA (FISH), VP35 protein, and c-Jun in NPC1 knockout cells relative to non-targeting cells. (E) Volcano plot of the per-cell median VP35 protein intensity delta AUC between each gene and non-targeting control cells, black points represent distinct non-targeting control sgRNAs. (F) Enrichr gene ontology analysis of top terms significantly enriched in genes that showed reduced VP35 protein intensity upon knockout (purple) or increased VP35 protein intensity (gold). (G) Workflow for random forest regression trees trained on single-cell features from non-targeting control cells. Models were designed to predict either Jun nuclear intensity given VP35 RNA and protein features or VP35 RNA intensity given VP35 protein features. (H) Volcano plot of random forest regression model coefficients of determination for the c-Jun prediction task, black points represent distinct non-targeting control sgRNAs; genes in purple had a negative mean sum of residuals, indicating decreased c-Jun relative to model prediction, while genes in gold had a positive mean sum of residuals. (I) Volcano plot of random forest regression model coefficients of determination for the VP35 RNA prediction task, black points represent distinct non-targeting control sgRNAs. (J) STRING analysis of genes that had a negative mean sum of residuals for the VP35 RNA FISH prediction task; purple shade denotes magnitude of mean sum of residuals, indicating the amount that Ebola VP35 positive-sense RNA was decreased relative to protein levels. Edge thickness corresponds to confidence score; only interactions with a confidence score >= 0.7 in the full STRING v11.5 were considered.

### 3.3.2 *A Supervised Deep Neural Network Reveals Regulators of Ebola Virus Subcellular Protein Localization*

Next, we sought to identify regulators of Ebola virus subcellular infection dynamics using deep learning models trained on single-cell images instead of relying on extraction of predefined features from images. Initially, we trained an autoencoder to reconstruct six-channel masked images of single cells (**Fig. 3.2a**). We then performed supervised training to fine-tune the encoder and latent space representation, using the prediction of four distinct categories of Ebola VP35 protein subcellular localization for training. We defined four classes of interest: 1) faint, indicative of uninfected cells, 2) punctate, which represents cells with viral inclusion bodies at early stages of infection, 3) diffuse, representing a later stage of infection with diffuse protein localization, and 4) peripheral, representing viral budding from infected cells and manually annotated over 3,000 cells for training. Next, we used the PHATE algorithm (Moon et al., 2019) to perform dimensionality reduction on the single-cell latent feature vectors from the unsupervised autoencoder and the autoencoder subjected to additional supervised training to visualize the encoding spaces of these two deep learning approaches. The unsupervised latent space showed a lack of separation of the four phenotypic classes, while the supervised features separated the four classes in the expected biological order (faint, punctate, diffuse, then peripheral) despite no ordinal specification during model training (**Fig. 3.2b**). Since these four categories represent ordered stages in the Ebola virus life cycle, we next performed an ordinal chi square test to identify genes that significantly altered the proportion of cells at earlier or later stages of infection relative to non-targeting controls (**Fig. 3.2c**). As expected, the genes that scored most significantly in this test also showed changes in infection levels using a simple intensity-based metric (e.g. the intracellular receptor NPC1, the HOPS complex, and negative regulators MESDC2 and ITGB1).

We next assessed the proportion of cells for each gene predicted to have a faint or punctate VP35 protein localization phenotype by our supervised neural network (**Fig. 3.2d**) in order to identify regulators not readily distinguishable from non-targeting controls by intensity alone. Genes with a high proportion of faint cells included many known entry or early infection regulators such as NPC1, as well as novel regulators such as PIK3C3 and the GET complex that were previously identified using our analysis of VP35 intensity delta AUCs (**Fig. 3.2e**). On the other hand, many of the genes associated with a significant increase in punctate VP35 localization have not been previously linked to Ebola infection or were linked to regulation of post-entry viral replication (**Fig. 3.2e**). Key enzymes in *de novo* pyrimidine biosynthesis (DHODH, CAD, and UMPS) previously shown to positively regulate Ebola virus replication (Luthra et al., 2018; Martin et al., 2018) scored significantly as enriching for a punctate phenotype, indicating that upon knockout of these genes Ebola was likely able to enter cells and begin formation of inclusion bodies but not continue replication as efficiently as non-targeting controls. Indeed, CAD was previously shown to colocalize with Ebola NP in inclusion bodies to regulate Ebola genome replication and transcription (Brandt et al., 2020). In addition to pyrimidine biosynthesis genes, we also identified genes involved in purine biosynthesis (MTHFD1, PAICS, PPAT, ATIC, and GART) that resulted in an enrichment of the punctate phenotype upon knockout and had not previously been associated with Ebola replication. MTHFD1 knockdown and treatment with the MTHFD1 inhibitor carolacton inhibited ZIKV, MuV, and SARS-CoV-2 replication without affecting entry, further validating the ability of our deep learning approach to separate effects of genetic knockouts by viral replication stage using the primary screening data without a requirement for further experiments (Anderson et al., 2021). Knockdown of Elongator complex genes (ELP2-6) also led to an increase in punctate VP35 (**Fig. 3.2e**). The Elongator complex modifies tRNA molecules critical for translational efficiency and perhaps also required optimal viral replication past the inclusion body formation stage (Hawer et al., 2018).

Finally, many mitochondrial genes significantly increased inclusion body formation, including genes from mitochondrial ribosomes, mitochondrial tRNA synthetases, and mitochondrial respiratory chain complex III and IV members. Example single-cell images and PHATE embeddings of single-cell profiles show clear enhancement of punctate VP35 protein localization in knockouts of UQCRB, a complex III subunit, and HARS2, a mitochondrial tRNA synthetase (**Fig. 3.2f, g**), the two genes with the most enhanced punctate phenotype.

Mitochondrial function has not been previously linked to Ebola infection; however, Ebola VP30 and VP25 proteins physically interact with the mitochondrial ribosome and inner membrane components (Batra et al., 2018) and oxidative phosphorylation and expression of mitochondrial translation genes was found to increase upon Ebola virus infection (Woolsey et al., 2019). Many of the mitochondrial genes identified have a role in oxidative phosphorylation (OXPHOS), either directly (such as the complex III and IV subunits) or less directly, as a previous genetic screen demonstrated the importance of mitochondrial ribosome components and tRNA synthetases for OXPHOS (Arroyo et al., 2016). Oxidative phosphorylation plays a role in propagation of other RNA viruses such as IAV (Bercovich-Kinori et al., 2016). In order to further investigate the effect of UQCRB inhibition on Ebola infection, we treated Ebola-infected HeLa cells with terpestacin, a small molecule inhibitor of UQCRB (Jung et al., 2010), and observed clear reduction in Ebola infection with an IC50 ~5 μM.

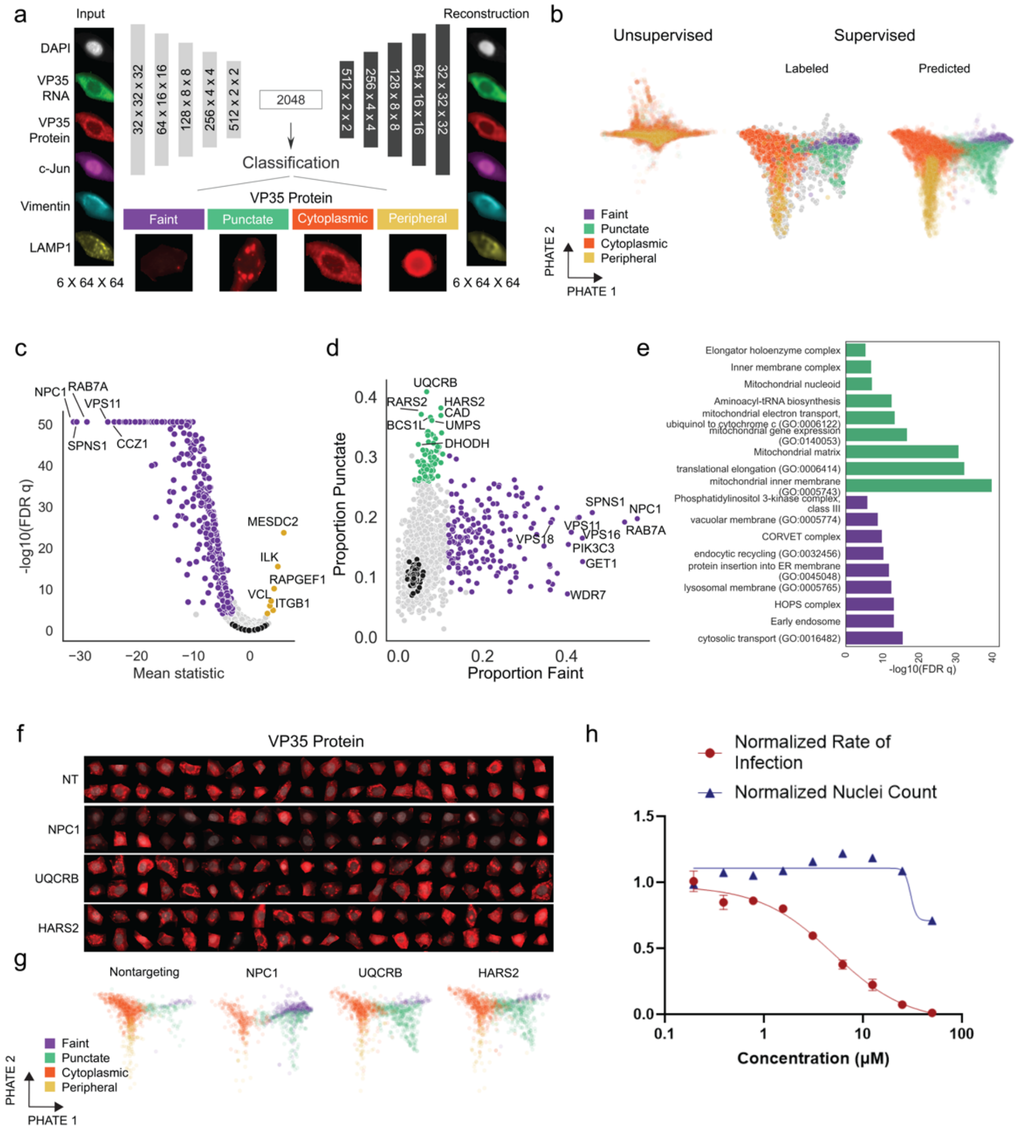


**Figure 3.2.**

**A supervised deep neural network reveals regulators of Ebola virus VP35 protein subcellular localization.** (A) Architecture for the neural network, trained first as an autoencoder (example input cell and autoencoder reconstruction shown) and subsequently tuned to classify cells based on four defined classes of VP35 protein localization, with one example shown for each class. (B) PHATE dimensionality reduction of latent single-cell features (2048/cell) from the unsupervised autoencoder and supervised model; labeled points indicate hand-classified ground truth labels while predicted points denote supervised neural network predictions. (C) Volcano plot of per-gene mean ordinal chi square statistics and FDR-corrected p-values, black points represent distinct non-targeting sgRNAs. (D) Scatterplot of proportion of faint vs punctate cells for each gene, black points denote individual non-targeting sgRNAs. (E) Gene ontology results of enriched terms for genes with a high proportion of faint cells (purple) or punctate cells (green). (F) Images of VP35 protein (red) overlaid with nuclear mask (gray) in randomly selected cells with sgRNAs targeting the indicated genes. (G) Deep learning latent single-cell features embedded using PHATE for cells with indicated gene knockouts; colors indicate labels predicted by supervised neural network. (H) Ebola infection and cell survival (nuclei count) in HeLa cells treated with terpestacin.

### *3.3.3 Phenotypic Profile Clustering and Matching Reveal Relationships Between Modulators of Ebola Virus Infection*

We next used gene-level average phenotypic profiles by taking the delta AUC between cells from each sgRNA and non-targeting control cells prior to averaging over sgRNAs for each gene. We used both unsupervised and supervised deep learning network latent dimension features to separately cluster 1500 genes that significantly altered overall virus infection as measured by VP35 protein levels and infer distinct functional groups of genes in an unbiased manner (**Fig. 3.3a**). Unsupervised learning generated a larger number of clusters significantly enriched for at least one ontology term using Enrichr, including clusters with low Jaccard similarity relative to any other cluster generated using supervised learning, demonstrating the utility of these complementary feature sets in distinct settings (**Fig. 3.3b**).

Single-cell images of genetic knockouts from distinct clusters demonstrate examples of phenotypes distinguishable using these deep learning features extracted directly from images (**Fig. 3.3c**). Some genes that scored significantly for increased VP35 protein load, such as integrins ITGAV (**Fig. 3.3c**) and ITGB1 (**Supplementary Fig. 3.3b**) showed increased intensity in channels not directly related to viral load, such as vimentin and LAMP1, as well as cell rounding. We therefore hypothesized that the increase in viral load for some genes in our screens could represent not a specific enhancement of viral replication, but rather a non-specific cell health phenotype.

Since we extracted thousands of features from single-cell images in our screen, we devised an approach to systematically define non-specific morphological scores for genes that significantly altered viral protein load, enabling better prioritization of genes for further follow-up studies. Specifically, we computed delta AUCs for each of thousands of features across all six image channels assayed and each genetic perturbation relative to a random sample non-targeting control cells (unmatched) or relative to a sample of non-targeting control cells whose distribution of VP35 cell median intensities matched the sgRNA of interest (**Fig. 3.3d**). We then plotted the correlation of delta AUCs for these features without matching compared to the matched condition (**Fig. 3.3e**). If the phenotypic effect of the genetic perturbation is primarily to alter viral infection levels, we expect that the delta AUCs for matched features will be lower than in the unmatched condition; that is, single-cell images of the genetic perturbation of interest closely match images of non-targeting control cells with a similar infection level. Indeed, we see that features for NPC1 knockout cells are much more modestly correlated between matched and unmatched conditions relative to the integrin ITGAV, which exhibits non-specific phenotypic effects (**Fig. 3.3e**). Using this approach, we calculated Pearson correlations between matched and unmatched conditions for thousands of genes (**Fig. 3.3f**).

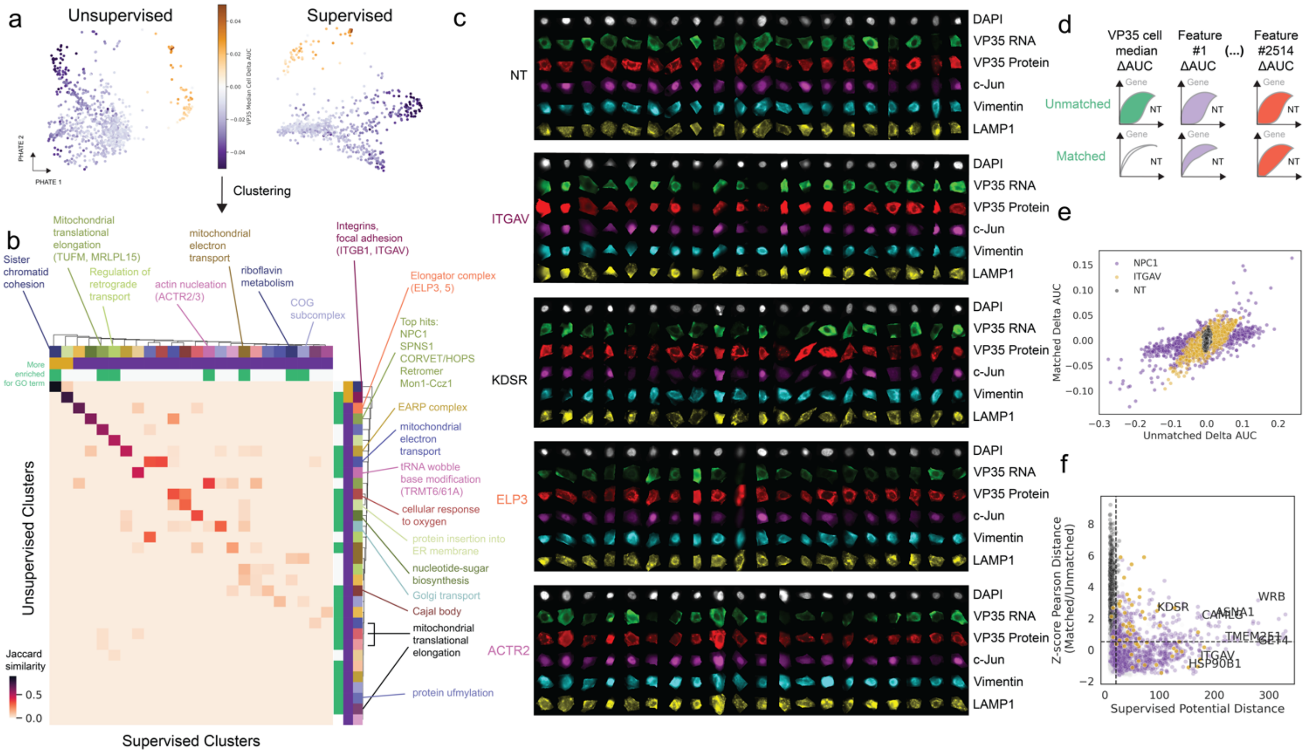


**Figure 3.3.**

**Phenotypic profile clustering and matching on infection level reveal relationships between Ebola virus infection modulators.** (A) PHATE clustering of gene-level average latent features derived from unsupervised or supervised deep neural networks. (B) Heatmap of Jaccard similarities between cluster memberships for all clusters with at least one significant GO term identified using unsupervised or supervised deep learning models. (C) Single-cell images of representative genes from distinct clusters. (D) Schematic of recalculation of cumulative delta AUCs for features between each sgRNA and non-targeting control cells after matching on VP35 protein levels in non-targeting controls. (E) Correlations between delta AUCs calculated without matching on VP35 protein levels compared to with matching on VP35 protein levels for non-targeting controls, NPC1 knockout, or ITGAV knockout cells; each point represents one individual feature. (F) Correlation between the mean PHATE potential distance from non-targeting controls in the supervised clustering (indicative of the Ebola infection phenotype strength) and the z-scored Pearson correlation distance between matched and unmatched features.

### *3.3.4 Targeted Image-Based Genetic Screens Identify Cell- and Virus-Specific Filovirus Regulators*

In order to rigorously validate modulators of infection identified in our genome-wide screen, we performed targeted image-based genetic follow-up screens, assaying 113 genes with 6 sgRNAs per gene in twelve distinct conditions: two timepoints, both earlier than the primary screen to increase power for identification of negative regulators of viral replication, two cell lines (HeLa cells used in the genome-wide screen and Huh7 cells), and three distinct filoviruses from the Ebolavirus family (Ebola virus, Marburg virus, and Sudan virus) (**Fig. 3.4a**). As in our primary screen, we assayed c-Jun, filovirus protein (VP35 for EBOV and SUDV; VP40 for MARV), and VP35 RNA via probes designed against the VP35 sequences of each virus.

PCA of non-targeting control sgRNAs from each experimental condition shows separation of conditions by cell type across PCA 1 (46% variance explained), and many conditions were most similar to the screening condition performed with the same cell and virus strain but at a different timepoint. To assess the reproducibility of our screening system, we compared cumulative delta AUCs from our genome-wide screen with the 24 hour infection (late) timepoint in the targeted secondary screen) and saw strong correlations (Pearson r >= 0.87) for both VP35 protein and RNA levels (**Fig. 3.4c**). Key known regulators such as NPC1 and SPNS1 were concordant across cell lines, viruses, and timepoints (**Fig. 3.4d, e**), as well as retromer complex members (VPS26A and VPS35), which was more recently associated with Ebola entry (Poston et al., 2022), and newly associated genes identified in this study, including PIK3C3, GET4, and the GARP complex (VPS51, 52, 53, and 54) (**Fig. 3.4d**). The ubiquitin ligase COP1 scored as a negative regulator in 10/12 screening conditions (**Fig. 3.4d, e**). The mechanism for COP1's regulation of Ebola levels is unclear; however, COP1 targets c-Jun for degradation, among other substrates, so increased c-Jun levels in COP1 knockout cells may favor filovirus replication (Migliorini et al., 2011).

We next examined virus-specific regulators (**Fig. 3.4f, g**), identifying knockout of TAF5L as increasing EBOV and SUDV infection levels while decreasing infection levels of the less closely related MARV (**Fig. 3.4f, g**). We also found that VPS35L strongly decreased Ebola infection levels but had modest effects on SUDV and MARV infection, while VPS35 knockout reduced replication of all filoviruses in multiple cell types (**Fig. 3.4d**). VPS35L and VPS35 are members of the retriever and retromer endosomal recycling complexes, respectively (McNally et al., 2017). The distinct effects of VPS35 and VPS35L knockouts may indicate differential specificity in filovirus utilization of retromer and retriever.

Finally, we investigated cell-type-specific regulators (**Fig. 3.4h, i**). PCYT2, a phospholipid synthesis enzyme which positively regulates formation of lipid droplets (Roberts et al., 2022) increased EBOV infection in Huh7 cells, while decreasing infection in HeLa cells. Interestingly, this gene is more highly expressed in Huh7 cells (Barretina et al., 2012), which have higher lipid droplet content than HeLa cells (Monson et al., 2018), indicating that the differences we observed in viral infection could be due to underlying distinct cellular lipid environments. Multiple cell-type-specific regulators were related to mitochondrial function (BCS1L, PRMT1, and HARS2, among others) (**Fig. 3.4h, Supplementary Fig. 3.4f**). Notably, HeLa cells rely on OXPHOS rather than glycolysis as the primary mode of energy production (Moreno-Sánchez et al., 2007), while Huh7 cells rely more on glycolysis (Y.-C. Shen et al., 2013); this is known as the Warburg effect. This difference could explain the strong defect in viral replication observed in HeLa cells upon loss of components related to mitochondrial respiration that was not observed in Huh7 cells.

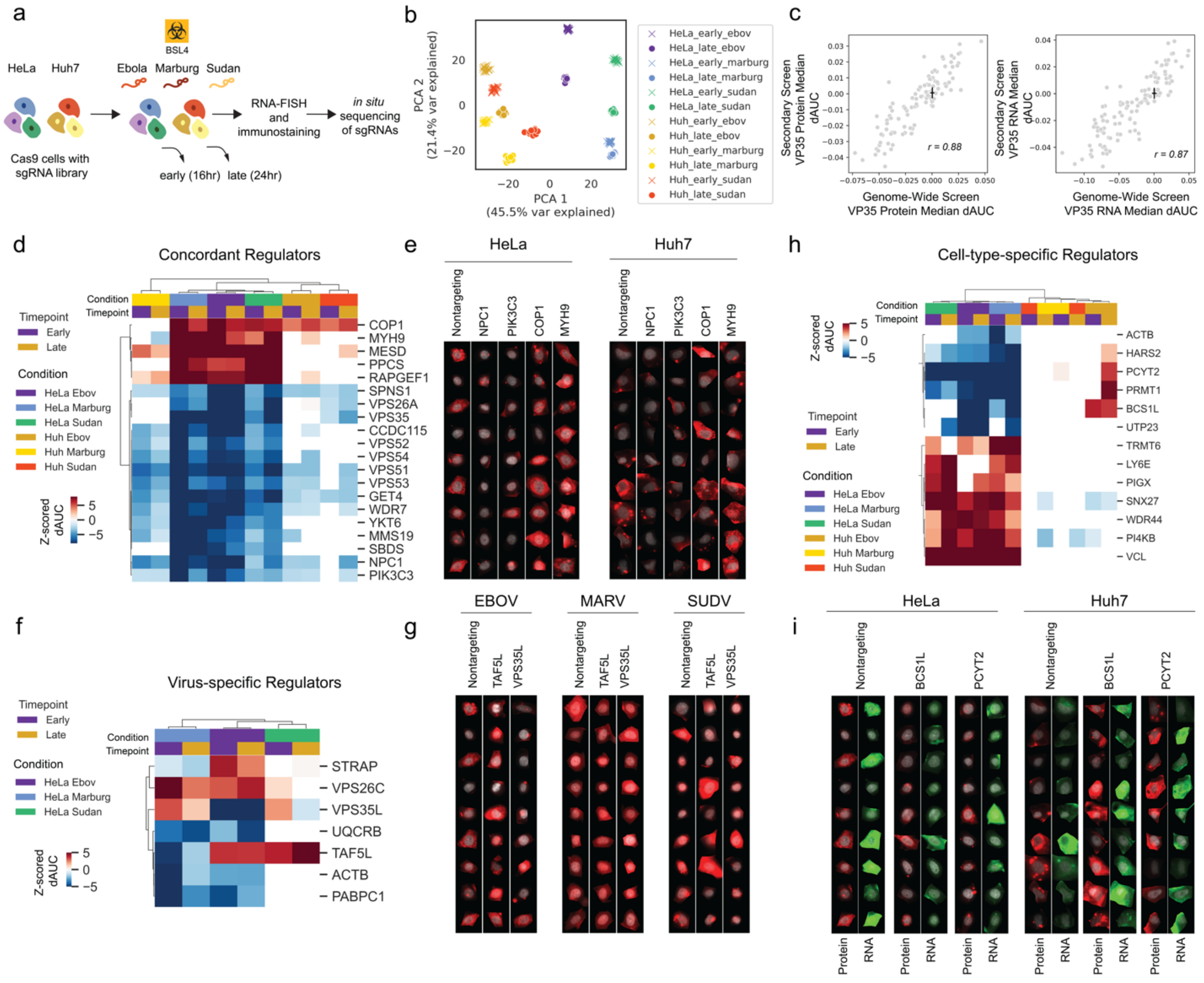


**Figure 3.4.**

**Targeted follow-up screens identify concordant and cell- and virus-specific ebolavirus regulators.** (A) Workflow for targeted secondary screens. (B) PCA of non-targeting control sgRNA phenotypic profiles from each of the twelve screening conditions. (C) Correlation between genome-wide and secondary screen VP35 protein and RNA delta AUCs (dAUCs); black lines indicate standard deviation for non-targeting control sgRNAs in each screen centered around the mean value for non-targeting sgRNAs in the screen. (D) Heatmap showing z-scored dAUC values for genes concordant across screen conditions (white cells indicate conditions where $p > 0.05$ relative to non-targeting controls in the same condition). Z-scores were calculated on delta AUC values for all genes in each screen condition relative to means and standard deviations for non-targeting sgRNAs. Hierarchical clustering performed using Pearson correlations. (E) Single-cell images from the secondary screen of select concordant genes (DAPI in gray, VP35 protein in red). (F) Heatmap as in (D) for genes with virus-specific effects (white cells indicate conditions where $p > 0.05$). (G) Single-cell images from the secondary screen of select virus-specific genes (DAPI in gray, VP35 protein for EBOV/SUDV or VP40 protein for MARV in red). (H) Heatmap as in (D) for genes with cell type-specific effects (white cells indicate conditions where $p > 0.05$). (I) Single-cell images from the secondary screen of select cell type-specific genes (DAPI in gray, VP35 protein in red, VP35 RNA in green).

## 3.4 Discussion

Ebola virus disease remains poorly understood, despite frequent outbreaks with persistent high case fatality rates. No genome-wide genetic screens to date have assessed phenotypes aside from cell fitness following infection with Ebola virus. Here, we used an image-based pooled screening technology to directly link genetic knockouts with infection levels of live Ebola virus and thousands of image parameters at the genome-wide scale. Through targeted follow-up screens, we confirmed novel roles for PIK3C3, which can be targeted by small molecule inhibitors, the GARP complex, and the GET complex in reducing replication of three filoviruses in two cell lines. Unlike previous cell survival screens, we also identified several negative regulators of filovirus replication under multiple conditions, such as COP1, MYH9, and RAPGEF1. We also present a new workflow for integration of targeted RNA FISH with our *in situ* sequencing protocol, enabling future image-based screens seeking to integrate host or pathogen RNA levels and localization into cellular phenotypic profiles.

Furthermore, we leverage the single-cell level information from our image-based dataset to better understand effects of genetic knockouts on phenotypes of interest directly from screening data. In particular, we extract features directly from images in an unbiased manner through unsupervised and supervised deep learning approaches, which we apply to identify genes that specifically modulate the subcellular localization of Ebola virus VP35 protein. We show that terpestacin, a small molecule inhibitor of UQCRB, a gene that our deep learning classifier predicted as increasing the fraction of punctate Ebola VP35, is able to decrease Ebola infection in cells, presenting an actionable new hypothesis Ebola virus therapy. We also applied random forest regression to identify genes that alter Ebola RNA levels relative to protein levels, decoupling these two steps of the Ebola virus life cycle. We further investigate STRAP, a spliceosome-associated gene that our model identified as increasing Ebola virus RNA relative to protein levels, and show that it co-localizes with the Ebola virus nucleocapsid protein (NP).

High-content single-cell screens with multi-dimensional outputs are becoming routine and now scale to genome-wide perturbations even for transcriptomic readouts (Replogle et al., 2022). However, aside from our study, high-content profiling screens of host-pathogen interactions have been limited. Notable recent examples include targeted transcriptome screens that revealed perturbation-induced changes in transcriptional trajectories upon HCMV (Hein & Weissman, 2022) and SARS-CoV-2 (Sunshine et al., 2022) infection. Importantly, high-content perturbation screens like these and that we report here did not require *a priori* specification of phenotypes of interest but rather resulted in rich data resources that can be mined to rapidly generate and investigate hypotheses based on primary screening data alone (Bock et al., 2022), provided adequate quantitative performance of the perturbation and phenotyping steps.

Our image-based genetic screens and analytical approaches revealed multiple genes that modulate distinct stages of the Ebola virus life cycle, from viral entry, to Ebola RNA replication and the formation of viral inclusion bodies, including several genes already targetable by small molecules with therapeutic potential. This work also serves as a general framework for systematically identifying regulators of distinct steps in host-virus interaction dynamics directly from data produced in a single genetic screen, enabling rapid identification of a broader set of potential therapeutic targets for diverse pathogens.

## 3.5 Methods

### *Library cloning, lentivirus production, and transduction*

Libraries were cloned into a CROP-seq-puro-v2 (Addgene #127458) backbone and lentivirus was then produced and transduced as previously described (Feldman et al., 2022). Multiplicity of infection for library transductions was estimated by counting colonies following sparse plating and antibiotic selection with puromycin.

### *Virus infection, phenotyping, and in situ sequencing for genome-wide screen*

HeLa-TetR-Cas9 clonal cells previously described (Feldman et al., 2019) were used for primary screening. Following transduction, cells were selected with puromycin (1 µg/mL) for 3 days after transduction and library representation was validated by NGS. Cas9 expression was induced with 1 µg/mL doxycycline for 1 week and cells were then seeded in ten 6-well glass-bottom dishes at 400,000 cells/well two days prior to fixation. Zaire ebolavirus (strain Mayinga) was added at an MOI of ~3 in 2mL media/well for 28 hours prior to fixation. Cells were fixed by removing media and adding 10% neutral-buffered formalin (Fisher Scientific LC146705) for >6 hours.

Cells were permeabilized with 100% methanol for 20 minutes; subsequently, the permeabilization solution was exchanged with PBS-T wash buffer (PBS + 0.05% Tween-20) by performing six 50% volume exchanges followed by three quick washes. Cells were incubated for 30 minutes at 37°C with probe hybridization buffer (30% formamide 5x SSC 0.1% Tween), then incubated with primary probes against VP35 positive-sense RNA (purchased from

Molecular Instruments and diluted 1:250 in probe hybridization buffer with 1:100 Ribolock) at 37°C for 4 hours. Samples were then washed 4x5 minutes in probe hybridization buffer at 37°C, washed 3x in PBS-T, and incubated in reverse transcription mix at 37°C overnight as previously described . After reverse transcription, cells were washed 5x with PBS-T and post-fixed using 3% paraformaldehyde and 0.1% glutaraldehyde in PBS for 30 minutes, followed by washing with PBS-T 3 times. Samples were then incubated in amplification buffer (5x SSC 0.1% Tween) at room temperature for 30 minutes. Meanwhile, HCR hairpins (Molecular Instruments B1 probes conjugated to Alexa Fluor 488) were separately prepared by heating at 95°C for 90 seconds and then cooling to room temperature in the dark for 30 minutes. Next, samples were incubated with probes diluted 1:125 in probe amplification buffer for 2 hours at room temperature. Following incubation, excess hairpins were removed by washing 5 times for 5 minutes with probe amplification buffer. Primary antibodies against VP35 (1:3200 dilution), c-Jun (1:1800 dilution, Cell Signaling Technology Cat# 9165, RRID:AB_2130165), and vimentin (1:1300 dilution, Abcam Cat# ab24525, RRID:AB_778824) were added by incubating samples for 3.5 hours at 37°C in 3% BSA (VWR Cat# 97061-422) in PBS. Samples were then washed 3x in PBS-T for 3 minutes, and incubated with secondary antibodies: 1:1800 donkey anti-mouse antibody (Jackson ImmunoResearch Labs Cat# 715-006-151, RRID:AB_2340762) disulfide-linked to Alexa Fluor 594 (Thermo Fisher A10270) via a custom conjugation, 1:1800 donkey anti-rabbit antibody (Jackson ImmunoResearch Labs Cat# 711-006-152, RRID:AB_2340586) disulfide-linked to Alexa Fluor 647 (Thermo Fisher Scientific A10277) via a custom conjugation, and goat anti-chicken DyLight 755 (Thermo Fisher Scientific SA5-10075), in 3% BSA for 3 hours at 37°C. Finally, samples were washed 6x with PBS-T for 3 minutes each, and 2x SSC with 200ng/mL DAPI was added to visualize nuclei to image samples. Following imaging, Alexa Fluor 594 and 647 antibodies were destained with 50mM TCEP in 2x SSC for 45 minutes at room temperature and FISH signal removed through treatment with 80% formamide in 2x SSC for 30 minutes at room temperature. In situ amplification of sgRNA sequences was then completed by incubating samples in a padlock probe and extension-ligation reaction mixture (1x Ampligase buffer, 0.4 U/μL RNase H, 0.2 mg/mL BSA, 100 nM padlock probe, 0.02 U/μL TaqIT polymerase, 0.5 U/μL Ampligase and 50 nM dNTPs) for 5 minutes at 37°C and 90 minutes at 45°C, and washing 2 times with PBS-T. Circularized padlocks were amplified using a rolling circle amplification mix (1x Phi29 buffer, 250 μM dNTPs, 0.2 mg/mL BSA, 5% glycerol, and 1 U/μL Phi29 DNA polymerase) at 30°C overnight. LAMP1 (Cell Signaling Technology Cat# 58996, RRID:AB_2927691) was visualized via incubation for 2 hours at 37°C in 3% BSA at 1:500 dilution. Following imaging, in situ sequencing was performed as previously described using sequencing primer oSBS_CROP-seq for 12 cycles (Feldman et al., 2019).

### *Virus infection, phenotyping, and in situ sequencing for secondary screen*

HeLa-TetR-Cas9 clonal cells used in the genome-wide screen and Huh7 311 Cas9 (Plasmid Addgene #96924) polyclonal cells were used for secondary screening. Ebola virus (strain Mayinga), Marburg virus (strain Musoke), Sudan virus (strain Gulu) were added at MOIs of ~3 in 2mL media/well for 16 hours or 24 hours prior to fixation. Viral RNA was amplified as described for the primary screen using virus strain-specific probes purchased from Molecular Instruments for positive-sense VP35. Antibody staining and in situ sequencing was performed as described for the primary screen except for Marburg virus, where an antibody against VP40 (1:1000, Integrated BioTherapeutics 0203-012) rather than VP35 was used since no antibody for Marburg VP35 was available. In addition, antibodies for vimentin and LAMP1 were omitted to allow for slower higher-magnification imaging. For Sudan virus, the same VP35 antibody was used as it recognized both Ebola and Sudan VP35 protein. *In situ* sequencing was performed for 6 cycles.

### *Fluorescence microscopy for primary genome-wide and secondary targeted follow-up screens*

All in situ sequencing images were acquired using a Ti-2 Eclipse inverted epifluorescence microscope (Nikon) with automated XYZ stage control and hardware autofocus. An LED light engine (Lumencor CELESTA Light Engine) was used for fluorescence illumination and all hardware was controlled using NIS elements software with the JOBS module. In situ sequencing cycles were imaged using a 10X 0.45 NA CFI Plan Apo λ objective (Nikon) with the following filters (Semrock) and exposure times for each base: G (546 nm laser at 40% power, emission 575/30 nm, dichroic 552nm, 200 ms); T (546 nm laser at 40% power, emission 615/24 nm, dichroic 565 nm, 200 ms); A (637 nm laser at 40% power, emission 680/42 nm, dichroic 660 nm, 200 ms); C (637 nm laser at 40% power, emission 732/68 nm, dichroic 660 nm, 200 ms). For the genome-wide primary screen, phenotyping images were acquired using a 20X 0.75 NA CFI Plan Apo λ objective (Nikon) with the following filters (Semrock unless otherwise noted) and exposure times: DAPI (405 nm laser at 5% power, Chroma Multi LED set #89402, 50ms), AF488 (477 nm laser at 30% power, Chroma Multi LED set #89402, 200ms), AF594 (546 nm laser at 10% power, emission 615/24 nm, dichroic 565 nm, 200ms), AF647 (637 nm laser at 10% power, emission 680/42 nm, dichroic 660 nm, 200ms), Dylight 755 (749 nm laser at 10% power, emission 820/110 nm, dichroic 765 nm, 200ms).

For the secondary screen, phenotyping images were acquired using a 40x 0.95 NA CFI Plan Apo λ objective (Nikon) with the following filters and exposure times: DAPI (405 nm laser at 5% power, Chroma Multi LED set #89402, 50ms), AF594 (546 nm laser at 10% power, emission 615/24 nm, dichroic 565 nm, 200ms), and AF647 (637 nm laser at 10% power, emission 680/42 nm, dichroic 660 nm, 200ms).

## *Quantification and Statistical Analysis*

### *Image analysis*

Images of cell phenotype and *in situ* sequencing of perturbations were manually aligned during acquisition using nuclear masks to calibrate the plate position to each of the four corner wells during screening. Alignment was then refined computationally via cross-correlation of DAPI signal between imaging acquisitions. Nuclei and cells were detected and segmented as previously described and *in situ* sequencing read calling was performed as previously described (Feldman et al., 2022). Data analysis functions were written in Python, using Snakemake for workflow control (Köster & Rahmann, 2012). Image analysis code is available on GitHub to enable reproducibility. Briefly, for segmentation of phenotyping images from the primary screen, nuclei were segmented using the following parameters: nuclei smooth = 4, nuclei radius = 15, nucleus area 90-1200, DAPI intensity threshold = 1350. Cells were segmented using signal in the vimentin channels, at an intensity threshold = 3000. For segmentation of *in situ* sequencing images from the primary and secondary screens, nuclei were segmented using the following parameters: nuclei smooth = 1.15, nuclei radius = 15, nucleus area 20-400, DAPI intensity threshold = 1000-2000, adjusted differently for each plate. Cells were segmented using signal in the four sequencing channels at intensity thresholds adjusted for each plate, between 2500 and 4200. For segmentation of phenotyping images from the secondary screen, nuclei were segmented using the following parameters: nuclei smooth = 9, nuclei radius = 100, nucleus area 200-18000 for HeLa cells or 200-50000 for Huh7 cells, DAPI intensity threshold = 4000 for HeLa cells and 3000 for Huh7 cells. Cells were segmented using background cell signal in the Jun channel, at an intensity threshold = 1525-1900 for Hela cells (depending on the assay plate) and 1625-1825 for Huh7 cells. All other parameters used for analysis were set to default settings.

### *Optical pooled screen analysis*

Only cells with a minimum of one read matching a barcode in the library were analyzed. For the genome-wide screen, only genes with a minimum of one read matching an sgRNA in the library and 2 sgRNAs with at least 50 cells/sgRNA were considered for analysis. Features were normalized on a per-cell basis relative to cells in the same field of view by subtracting the median and dividing by the MAD x 1.4826 (M.-A. Bray et al., 2016) and scores for features relative to non-targeting controls were determined by calculating differences in cumulative AUCs. These delta AUCs were averaged over sgRNAs for a given gene and significance was determined by comparing delta AUCs for individual sgRNAs to distributions bootstrapped from non-targeting control cells (bootstrapped 100,000 times). Gene-level p-values were calculated using Stouffer's method and then corrected using the Benjamini-Hochberg procedure. Random forest regression models were trained on 50,0000 randomly selected non-targeting control cells using sklearn.ensemble.RandomForestRegressor with scikit-learn 1.1.3, random state set to 7, n_estimators = 100, and max_features = 'sqrt' (Pedregosa et al., 2011). Statistical significance was determined as described for the cumulative AUC analysis above.

### *Dimensionality Reduction, Clustering, and Gene Enrichment Analysis*

In Figures 1 and 2, Enrichr results (Kuleshov et al., 2016) were determined using gseapy 0.14.0 (Z. Fang et al., 2023) with the 2021 KEGG and GO gene sets and 2016 Reactome gene sets. PHATE 1.0.10 (Moon et al., 2019) was used to cluster single cells with Euclidean distance, cosine mds distance, gamma = 1, knn = 5, 20 PCs in Figure 2. For gene-level clustering in Figure 3, PHATE with Euclidean distance, cosine mds distance, gamma = 1, knn = 3 and the number of PCs giving 95% of the variance (282 for unsupervised and 372 for supervised features) were used.

### *Deep learning analysis*

### *Unsupervised Training*

First, we performed unsupervised training on nearly 40 million 64x64 six-channel images from the screen, with a latent space embedding size of 2048. Watershed cell segmentation masks were used to mask cell regions of interest in each cropped image. For training, we used an 80/20 train/test split. In each epoch of training, the minimum of {256 and the number of cells in the field of view} cells were sampled from each training field of view. We then fed the image through the encoder/decoder architecture specified in Figure 3.2a. We propagated the loss as per a custom mean-squared error (MSE) loss function which takes the MSE over all reconstructed pixels within the watershed mask for each cell of interest. The motivation for this loss function was to avoid local minima wherein low-loss is achieved by perfectly reconstructing the blank background of the cell image, while not producing a high-fidelity reconstruction of the cell itself. We used the Adam optimizer with a learning rate of 0.001, trained for 50 epochs with a random seed. The training and test loss profile of this model is shown in Supplementary Figure 3.2a.

### *Supervised Training*

We then took the output model from the unsupervised training, which we term the “pretrained model” and fine-tuned its encoder and latent space representation by training a classification head, as shown in Figure 2a. We used 3,889 cell images manually labeled with one of four phenotypic categories in the set {faint, punctate, cytoplasmic, peripheral}. We used the scikit-learn (Pedregosa et al., 2011) stratified dataset splitting functionality to ensure the distributions of phenotypes between classes were balanced. Firstly, we reserved 25% of the dataset, or 973 cell images, as a test set. The remaining 2,916 images were split into 4 stratified folds (train/validation) and used for fine-tuning the autoencoder.

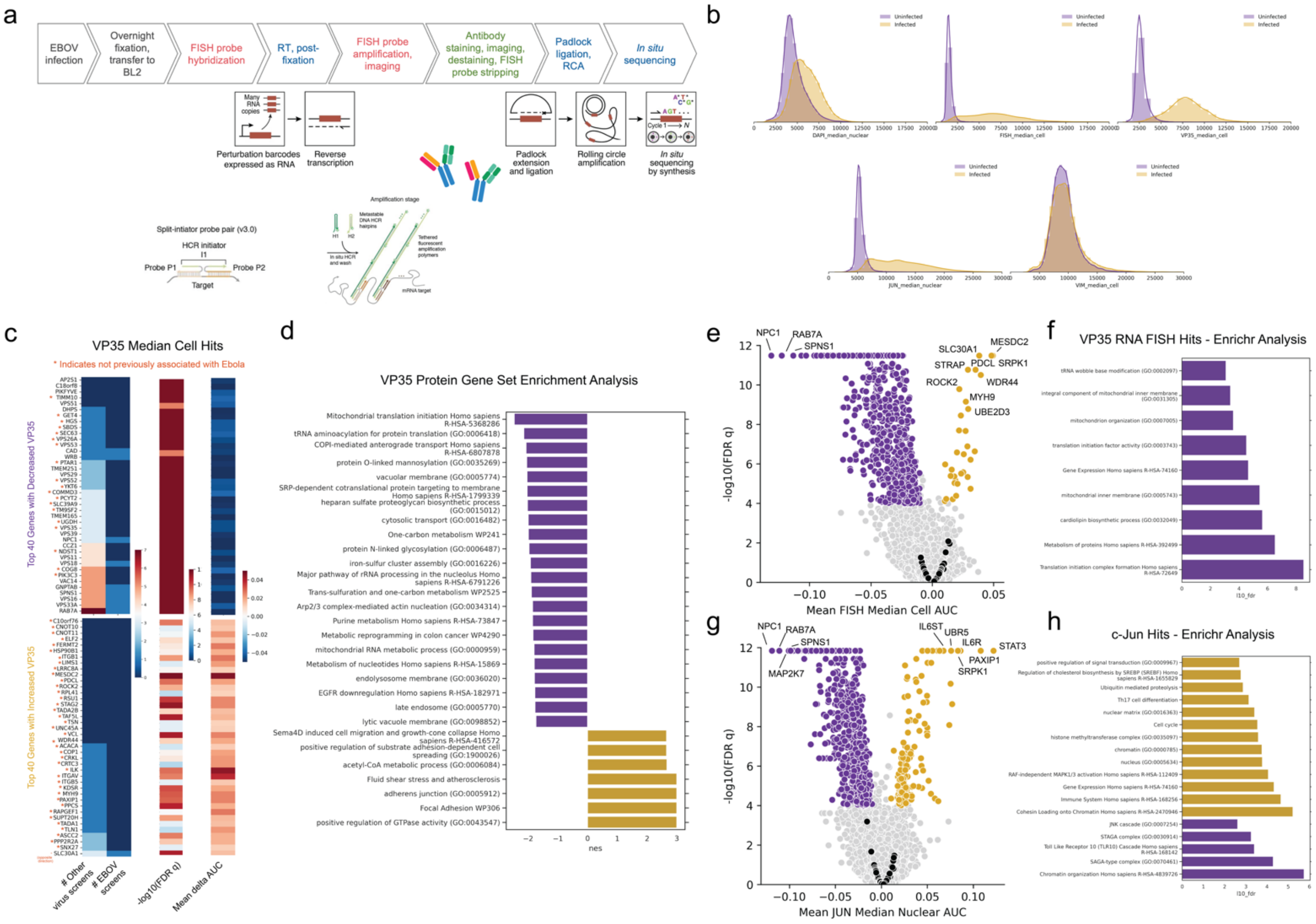
For each of the folds, we fine-tuned the autoencoder model using a specific procedure designed to amplify the utility of the small training set. We leverage the property that sufficiently minor transformations applied to a particular cell image do not change its class label. In particular, a cell which has been rotated, axis-wise mirrored, slightly blurred, or slightly distorted, or any combination of those, should retain its original label. We used PyTorch’s image transformation functions (Paszke et al., 2019) to generate a large pool of labeled samples by applying mild transformations to the original images. The model training then proceeds as follows. While the dataset of training samples does not have a uniform phenotype distribution, our objective is to provide the model with as close to a uniform distribution as possible, so as not to impose a prior on the model training. Moreover, we sought to determine whether the classifier itself would learn a shift towards a particular class. To enforce this uniform distribution, we consider the fine-tuning as a series of epochs and batches. For each batch, we select 5 cell samples at random from each phenotype. We then apply five random sequences of transformations, from the preselected set detailed above, to each image. This results in 25 cell samples per phenotype, for a total batch size of 100 cell images. We define each epoch as containing 100 batches, and thus the model is trained on 10000 images per epoch. Furthermore, we track the accuracy of the classification head on the validation set, across the run. For model selection, we store both the balanced and absolute accuracy on a held-out validation set, for each epoch of training. We use these statistics to select candidate models for further analysis. The specific model we have chosen as the best-performing version was the model having the highest balanced accuracy over 50 epochs of training. For evaluation, we split the held-out 25% of the dataset into 5 groups, and trained a Support Vector Machine (SVM) classifier with a linear kernel on 95% of the cells, while holding out 5% as a test set. The SVM classifier is trained using the latent space representation for the autoencoder and ResNet, and using the normalized features for the baselines. We then averaged the balanced accuracy on the train and test sets across all five SVMs and reported these in Supplementary Figure 3.2d. Of the five folds for the selected semi-supervised autoencoder model, we selected the fold with the highest individual balanced test accuracy and used this SVM to generate labels for the entirety of the dataset.

To determine sgRNAs that significantly altered the proportion of cells in each phenotypic category, an ordinal chi squared test was performed using R 4.2.2 with the coin package (v1.0.9) and results were combined at the gene level using Stouffer's method.

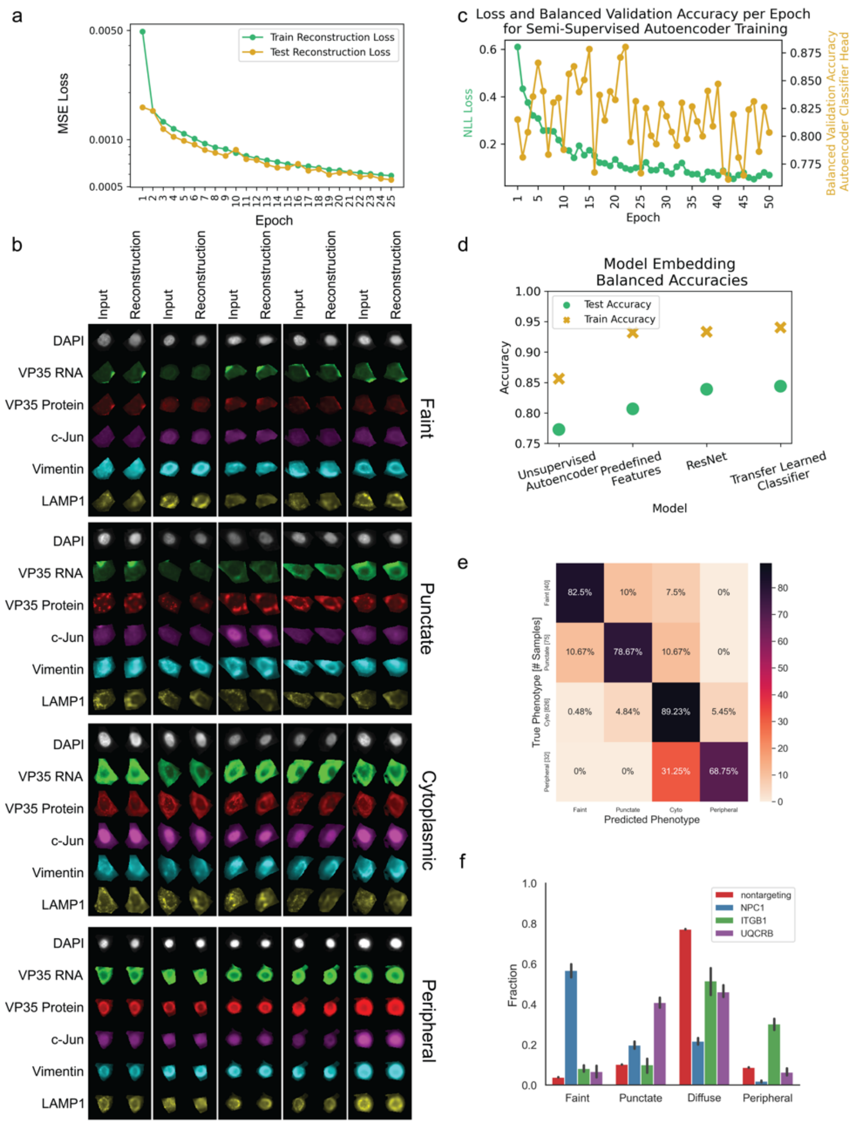
### *Data and Code Availability*

Code is available at https://github.com/beccajcarlson/EBOVOpticalPooledScreen.

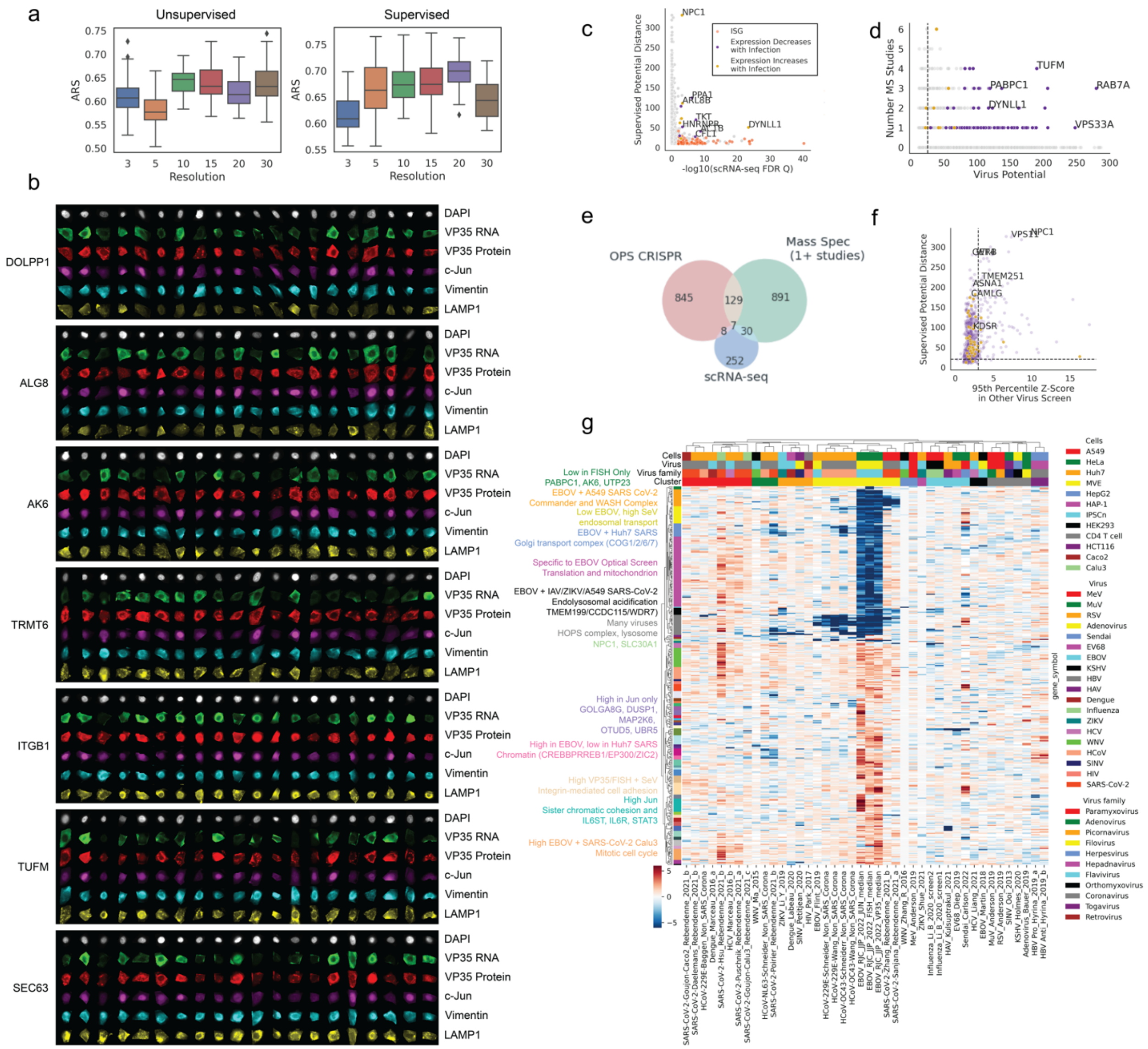
## 3.6 Supplementary Figures



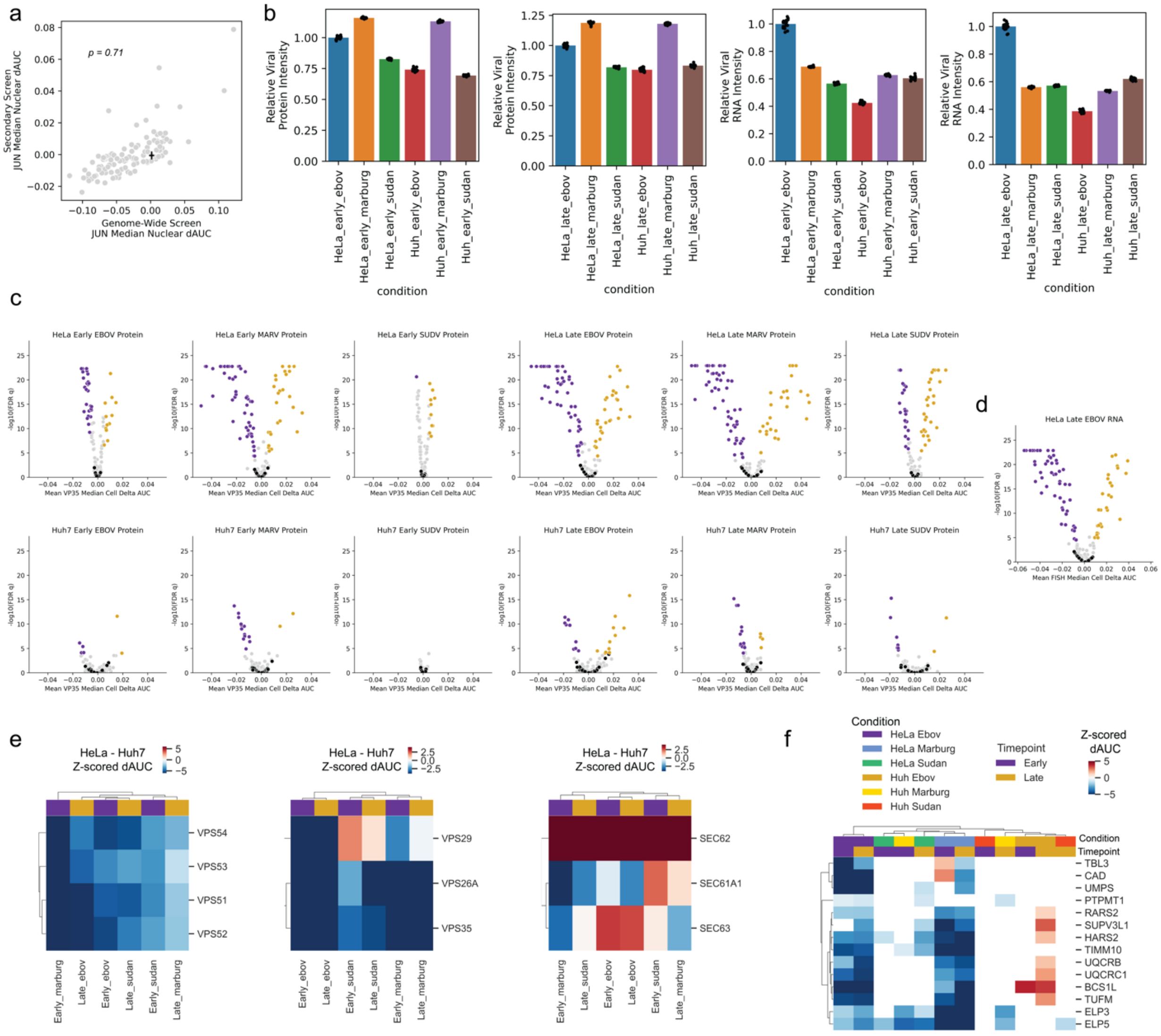
**Supplementary Figure 3.1.**
(A) Integration of optical pooled screening workflow with RNA FISH detection using HCR amplification. (B) Histograms of intensity features in five channels for non-targeting controls cells that were infected or not infected in the genome-wide optical pooled screen. (C) Top 40 hits with increased or decreased VP35 protein levels and the number of non-Ebola virus genetic screens or Ebola-specific genetic screens they scored in. Genes not previously associated with Ebola in the literature are marked with an orange asterisk. (D) Gene set enrichment analysis of genes with significantly decreased (purple) or increased (gold) Ebola virus VP35 protein levels. (E) Volcano plot showing genes that scored significantly for changes in VP35 RNA levels by FISH. (F) Enrichr analysis of gene ontology terms significantly enriched in genes that reduced VP35 RNA levels. (G) Volcano plot showing genes that scored significantly for changes in c-Jun levels. (H) Enrichr analysis of gene ontology terms significantly enriched in genes that reduced or enhanced c-Jun levels.



**Supplementary Figure 3.2.**
(A) Unsupervised autoencoder reconstruction losses for training and test sets across 25 epochs. (B) Examples of manually labeled faint, punctate, cytoplasmic, and peripheral input cell images with accompanying unsupervised autoencoder reconstructions. (C) Supervised autoencoder negative log likelihood loss and balanced validation accuracy across 50 epochs of training. (D) Balanced model train and test set accuracies for the VP35 protein localization prediction task using SVMs on latent embeddings from the unsupervised autoencoder, predefined features, a Resnet-50 architecture trained on the prediction task, or the transfer learned classifier. Predefined features include intensity, correlation, and texture morphological features similar to those previously described for Cell Painting (Bray et al., 2016). (E) Confusion matrix of model predictions vs manually labeled classifications on model test set. (F) Proportion of cells in each VP35 localization category for non-targeting controls and the genes with the largest proportion of faint (NPC1), punctate (UQCRB), and peripheral (ITGB1) cells. Error bars indicate SEM across sgRNAs targeting the same gene.



**Supplementary Figure 3.3.**
(A) Adjusted Rand score for Leiden clustering at different resolutions. (B) Additional single-cell images of select genetic knockouts from the genome-wide optical pooled screen. (C) Correlation between the PHATE potential distance from the supervised clustering and the adjusted FDR p-value from the Kotliar study, noting genes whose expression significantly increased or decreased along with infection. (D) Correlation between the PHATE potential distance from the supervised clustering and the number of mass spectrometry studies that identified the genes as an interactor with an Ebola virus protein. (E) Venn diagram showing overlap between top optical pooled screen hits, genes that were present in at least one mass spectrometry study, and differentially expressed genes from Kotliar et al's single-cell RNA sequencing study. (F) Correlation between the PHATE potential distance from the supervised clustering and the 95th percentile z-score for each gene in other virus screens. (G) Unsupervised clustering of hits from genome-wide virus screens.



**Supplementary Figure 3.4.**

(A) Correlation between genome-wide c-Jun median nuclear delta AUC scores and secondary screen delta AUC scores; black lines indicate standard deviation for non-targeting control sgRNAs in each screen centered around the mean value for non-targeting sgRNAs in the screen. ((B) Secondary screen mean viral protein (VP35 for EBOV and SUDV or VP40 for MARV) and RNA intensities in non-targeting control sgRNAs relative to HeLa cells infected with EBOV. (C) Volcano plots for VP35 (EBOV, SUDV) or VP40 (MARV) protein expression in each of the twelve screening conditions. (D) Volcano plot for viral VP35 RNA levels in HeLa cells at the late timepoint condition. (E) Heatmaps showing the difference between HeLa cell and Huh7 cell z-scored delta AUCs for members of the GARP, retromer, and the Sec61 complex. Hierarchical clustering performed using Euclidean distance. (F) Heatmap showing z-scored delta AUC values for genes identified as enriched for a punctate phenotype in the genome-wide screen and also included in secondary screens (white cells indicate conditions where $p > 0.05$). Hierarchical clustering performed using Pearson correlations.

# 4 Human STING is a proton channel

Authors: Bingxu Liu†, Rebecca J. Carlson†, Ivan Susin Pires^, Matteo Gentili^, Ellie Feng, Quentin Hellier, Marc A. Schwartz, Paul C. Blainey, Darrell J. Irvine*, Nir Hacohen*

† Contributed equally
^ Contributed equally
* Contributed equally

## 4.1 Abstract

Proton leakage is a common signal for non-canonical LC3B lipidation and NLRP3 inflammasome activation, processes induced upon STING activation. Based on analysis of STING structure, we hypothesized that human STING is a proton channel. Indeed, we found that STING activation induced a pH increase in the Golgi, and the proton flux was recapitulated by reconstituting STING in liposomes. C53, a small-molecule STING agonist that binds the putative channel interface, blocked STING-induced proton flux both in the Golgi and in liposomes. STING-induced LC3B lipidation and inflammasome activation by cGAMP or diABZI were also inhibited by C53, suggesting that STING's channel activity is critical for these two processes. Our work provides a new paradigm for decoupling STING's induction of interferon from LC3B lipidation and inflammasome activation.

## 4.2 Main Text

STimulator of INterferon Genes (STING) is a conserved mammalian cytoplasmic receptor that is essential for sensing cyclic dinucleotides derived directly from bacteria (Burdette et al., 2011) or synthesized by cyclic GMP-AMP (cGAMP) synthase (cGAS) upon recognition of cytosolic DNA (Sun et al., 2013; Wu et al., 2013). Upon binding to its native ligand cGAMP, STING undergoes a conformational change and translocates from the endoplasmic reticulum (ER) to the Golgi and endosomes, where it carries out multiple biological functions, including interferon induction (Ishikawa & Barber, 2008), non-canonical LC3B lipidation (Gui et al., 2019)5/17/23 8:40:00 PM, and NLRP3 inflammasome activation (Gaidt et al., 2017). While interferon is induced by STING-mediated TBK1-IRF3 activation (C. Zhang et al., 2019; Zhao et al., 2019)5/17/23 8:40:00 PM, the mechanisms by which STING activates non-interferon functions, in particular non-canonical LC3B lipidation and inflammasome activation, are still unclear.

STING induces FIP200-independent non-canonical LC3B lipidation, which involves Conjugation of ATG8 to Single Membranes (CASM) (Fischer et al., 2020; Xu et al., 2022). This process, sometimes termed non-canonical autophagy (Hooper et al., 2022; Xu et al., 2019), is important for bacterial control , and is known to be initiated by ion release into the cytoplasm from acidic organelles (such as Golgi and endosomes) via multiple mechanisms including organelle membrane damage (Xu et al., 2019), pathogen-derived ion channels such as the influenza M2 protein (Ulferts et al., 2021), or proton ionophores (Hooper et al., 2022). This led us to ask whether proton leakage from organelles is also involved in STING-induced LC3B lipidation and, if so, how STING activation leads to such ion transport.

To test whether STING activation leads to proton transport out of acidic compartments, we constructed genetically-encoded ratiometric pH sensors targeted to several organelles. As a sensor, we employed superecliptic pHluorin (SEP), a variant of GFP whose brightness increases with pH (Miesenböck et al., 1998; Sankaranarayanan et al., 2000), fused to pH-insensitive mRuby3. This ratiometric sensor was targeted to the cis/medial-Golgi (via fusion to MGAT), the trans-Golgi (via fusion to GALT), or endolysosomes (via fusion to LAMP1) (Linders et al., 2022). These sensors were expressed in human BJ1 fibroblasts and SEP to mRuby3 fluorescence ratios were correlated with intracellular pH values using calibration data (**Supplementary Fig. 4.1a, b**). Upon treatment with both the positive control V-ATPase inhibitor Bafilomycin A1 (BafA1) and the STING agonist diABZI, we observed that the ratio of SEP to mRuby3 fluorescence increased in both the cis/medial- and trans-Golgi compartments (**Fig. 4.1a, b, Supplementary Fig. 4.1c**). By contrast, in endolysosomal compartments a pH increase was observed upon BafA1 treatment but not upon diABZI treatment (**Supplementary Fig. 4.1d, e**). However, SEP has a pK ~ 7.1 (Sankaranarayanan et al., 2000) and our own pH calibration data showed low sensitivity to changes in pH below 6.5 (**Supplementary Fig. 4.1 a, b**), so it remains possible that STING activation could elicit an endolysosomal pH increase that is below the sensor's limit of detection.

We next sought to systematically identify genes that mediate the pH increase observed in the Golgi compartment upon STING activation. Given that non-canonical LC3B lipidation is activated by proton leakage from acidic organelles, we reasoned that screening for genes that modulate STING-induced LC3B lipidation would also identify potential channel proteins responsible for the observed proton flux. We therefore carried out a genome-wide CRISPR fluorescence-activated cell sorting (FACS) screen, using HEK293T cells transduced to express the autophagy-associated protein LC3B fused to red fluorescent protein (RFP) and HA-tagged STING. To reduce the background lipidated LC3B signal derived from basal canonical autophagy, we knocked out *FIP200* (Fischer et al., 2020). Following transduction with the Brunello genome-wide lentiviral library (Sanson et al., 2018), cells were stimulated with the STING agonist diABZI and permeabilized to remove LC3B that was not lipidated, further reducing background fluorescence (Eng et al., 2010). STING-HA$^+$ cells were sorted into LC3B$^-$ and LC3B$^+$ bins (**Fig. 4.1c**) to specifically identify STING-induced LC3B lipidation regulators that did not impair STING expression. The screen showed strong technical reproducibility (**Supplementary Fig. 4.1f, g**) and identified critical STING-induced LC3B lipidation regulators, including most of the known V-ATPase components as well as non-canonical autophagy factors such as *ATG16L1* (**Fig. 4.1d**). As a general mechanism of STING-induced LC3B lipidation, V-ATPase senses proton leakage from acidic vesicles through recruitment of V1 subunits to V0 complexes that together act as a scaffold for recruitment of ATG16L1, which initiates LC3B lipidation, through a process that is independent of V-ATPase's proton pumping function (Hooper et al., 2022; Xu et al., 2019, 2022). Despite the high recovery rate of V-ATPase components and non-canonical autophagy factors from our screen, no other known channel protein significantly inhibited STING-dependent LC3B lipidation in our screen (**Fig. 4.1d**). Therefore, we hypothesized that STING itself may mediate the observed Golgi proton leakage, and thereby trigger V-ATPase assembly and subsequent recruitment of ATG16L1 to initiate LC3B lipidation (Xu et al., 2019).

In order to identify the domains of STING involved in LC3B lipidation, we first tested whether the STING ligand-binding domain (LBD), which has been proposed to recruit LC3B through its LR motifs (Gui et al., 2019; D. Liu et al., 2019), could induce LC3B lipidation upon translocation to Golgi or endosomes. We measured LC3B lipidation in 293T cells expressing wild-type (WT) STING, a STING oligomerization-deficient variant (A277Q/Q273A STING, termed AQQA) (Shang et al., 2019), or an endolysosome-localized STING fusion protein (the endolysosomal protein TMEM192 fused to the STING LBD) (Gentili et al., 2023). Following stimulation with the STING agonist diABZI, the AQQA variant exhibited impaired translocation, phosphorylation, and LC3B lipidation (**Fig. 4.1e, f**). By contrast, TMEM192-STING-LBD did not induce LC3B lipidation despite its endolysosomal localization and strong induction of STING phosphorylation (**Fig. 4.1e, f**). Since translocation of the STING LBD domain was not sufficient to induce LC3B lipidation, we hypothesized that STING's transmembrane domain could play an important role in LC3B lipidation upon STING translocation.

Given the essentiality of STING translocation from the ER to the Golgi for STING-induced LC3B lipidation and the known role of a pH increase in acidic organelles as a common trigger for this process, we hypothesized that STING could generate Golgi ion leakage by either inducing membrane damage, resulting in a secondary ion leakage, or by directly acting as an ion channel through its transmembrane domain. Previous research has shown that STING translocation induced LC3B lipidation without formation of Galectin-3 puncta (Xu et al., 2022), suggesting that STING activation does not result in membrane damage. We therefore investigated the possibility that STING directly acts as a channel for proton release into the cytosol upon translocation to the Golgi, an acidic compartment (Linders et al., 2022). STING-dependent induction of LC3 lipidation is an ancestral function of the sensor conserved from *Homo sapiens* to *Nematostella vectensis* (Gui et al., 2019). Therefore, we hypothesized that if STING functions as a channel, this activity should be structurally conserved in non-human organisms.

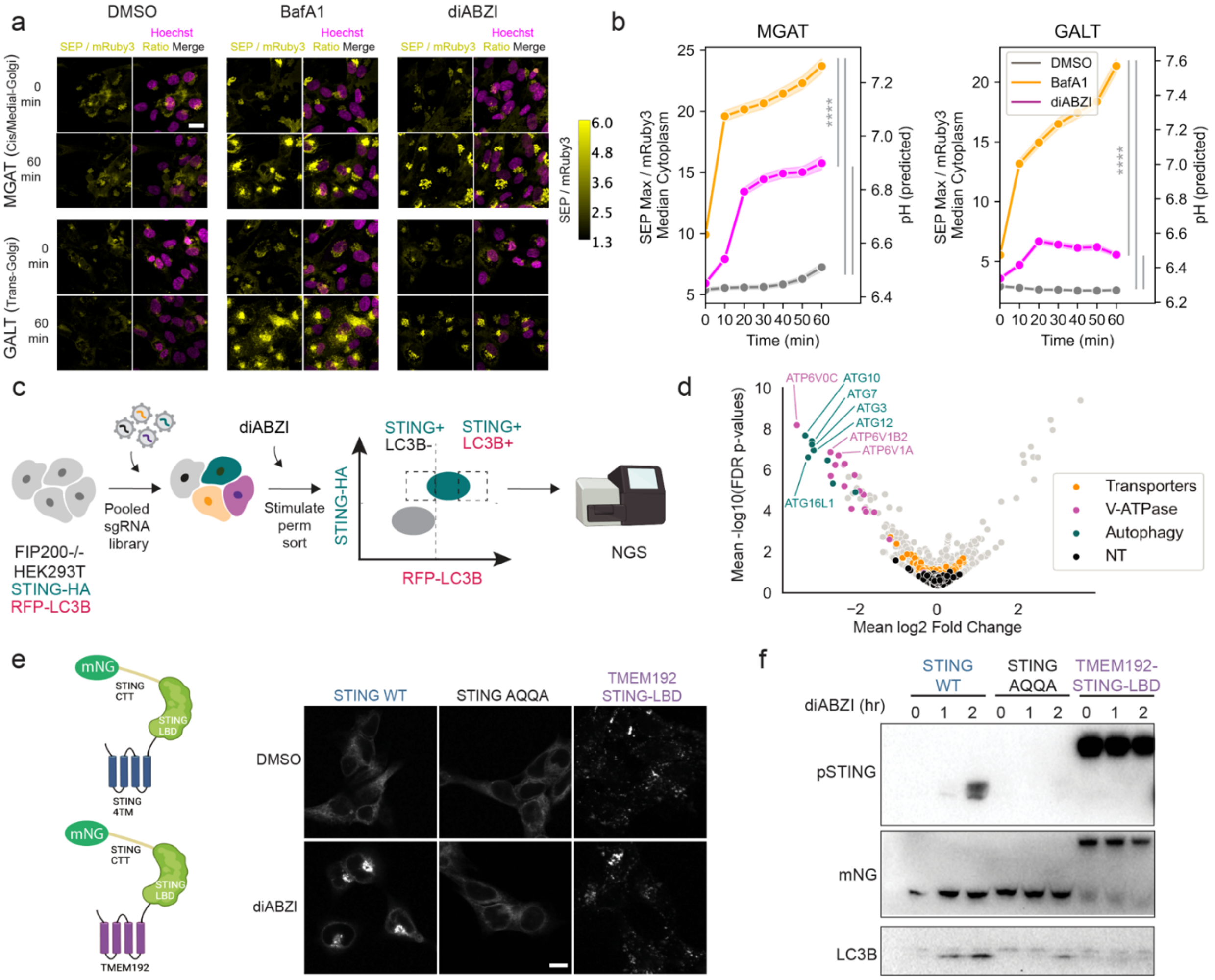


**Figure 4.1.**

**STING activation leads to a pH increase in the Golgi and a genome-wide screen for regulators of STING-induced LC3B lipidation did not identify transporters that could mediate this effect.** (A) Representative images of BJ1 cells expressing a ratiometric SEP and mRuby3 reporter localized to MGAT or GALT at 0 and 60 minutes post 1 μM diABZI or 1 μM BafA1 stimulation. Scale bar 20 μm. (B) Quantification of experiment in (A), data combined from three independent biological replicates. pH predicted using linear regression model in Supplementary Fig. 4.1b. One-way ANOVA followed by Tukey's

HSD at 60 minute timepoint. **** indicates p < .0001. (C) Workflow for the genome-wide CRISPR screen. (D) Volcano plot of genome-wide CRISPR screen results across two replicates: V-ATPase, non-canonical autophagy components, and known ion transporters (GO:0015075, ion transmembrane transporter activity) are highlighted. NT indicates non-targeting control sgRNAs. (E) STING-mNeonGreen (mNG) constructs and representative images of STING mNG localization in 293T cells expressing WT STING, STING AQQA, or TMEM192-STING-LBD and stimulated with DMSO or 1 μM diABZI for 1 hr, scale bar 10 μm. One representative experiment of n = 2 experiments. (F) Western blotting of phosphorylated STING (pSTING) and LC3B lipidation in 293T cells expressing WT STING, STING AQQA, TMEM192-STING-LBD stimulated as in (E). One representative experiment of n = 3 experiments.

To investigate whether STING could function as an ion channel, we analyzed previously published cryo-EM structures of chicken STING using MOLEonline, a tool for automated detection and characterization of channels in macromolecules (Pravda et al., 2018). When we analyzed the cryo-EM structures of ligand-free apo STING (PDB structure: 6NT6) compared to STING bound to its native ligand, cGAMP (PDB structure: 6NT7), the tool identified a pore spanning the lipid bilayer (1.3 Å bottleneck radius, 29.9 Å length) in ligand-bound STING which was absent in ligand-free apo STING; the latter showed a central cavity but not spanning the whole membrane (**Fig. 4.2a and Supplementary Fig. 4.2a**). A recently discovered STING agonist, compound 53 (C53) (Lu et al., 2022), binds to the STING transmembrane domain in the area of the putative pore. We therefore hypothesized that C53 could be used as a tool for inhibition of the proposed ion channel function of STING. Indeed, we found that the STING-mediated Golgi pH increase observed upon treatment with agonists diABZI or cGAMP alone was dramatically reduced when cells were co-treated with C53 along with diABZI or cGAMP (**Fig. 4.2b, c, Supplementary Fig. 4.2b**), consistent with the hypothesis that the proposed pore region of STING is required to produce a pH change within the Golgi. In order to exclude a potential role for downstream non-canonical autophagy factors in mediating the observed pH increase, we knocked out ATG16L1, which is essential for STING-induced LC3B lipidation (Fischer et al., 2020), in BJ1 pH reporter cells. As expected, we found no significant inhibition of STING agonist-induced Golgi pH increase in these cells (**Supplementary Fig. 4.2c, d**). To further confirm STING's mediation of the observed pH increase, we knocked out endogenous STING in BJ1 cells expressing the cis/medial Golgi pH reporter (**Supplementary Fig. 4.2e**) and overexpressed STING-miRFP680. We then stimulated these cells with diABZI and used live-cell super-resolution Airyscan imaging to image STING translocation to individual Golgi vesicles. We observed that STING preferentially translocated to cis/medial Golgi vesicles with a pH increase, indicated by higher SEP relative to mRuby3 signal (**Fig. 4.2d, e**). Quantification revealed an increase in STING signal over time in vesicles with a high SEP to mRuby3 ratio, with little change in STING intensity at vesicles with a low SEP to mRuby3 ratio (**Fig. 4.2d**) as well as increasing correlation between

STING and vesicle SEP to mRuby3 ratio over time (**Supplementary Fig. 4.2f**), as expected if STING translocation to acidic Golgi compartments induces proton leakage.

In order to explore the sufficiency of STING for mediating proton transport, we purified full-length human STING (**Supplementary Fig. 4.2g**) (Shang et al., 2019) and reconstituted STING onto liposomes (Dezi et al., 2013; H.-H. Shen et al., 2013; Tsai & Miller, 2013). We used the pH-sensitive dye 9-Amino-6-Chloro-2-Methoxyacridine (ACMA) to measure proton flux into liposomes. ACMA will be sequestered in liposomes and its fluorescence quenched upon pH changes induced by proton transport from the external buffer into the vesicles, enabling quantification of proton flux based on a reduction in total ACMA fluorescence (**Fig. 4.2f**). Proton flux was observed in STING proteoliposomes and was reduced in the presence of C53, while control liposomes formed with identical solutions devoid of protein did not show proton flux (**Fig. 4.2g**), consistent with the hypothesis that STING is sufficient to transport protons across lipid membranes. In contrast to the behavior in live cells, where we found that STING-mediated proton leakage is induced by STING agonists like diABZI or cGAMP, in liposomes STING mediated proton leakage similarly in the presence or absence of diABZI (**Fig. 4.2g**). The dispensability of diABZI for proton flux in this reductionist liposome assays suggests that an environment with a voltage difference or pH gradient (such as the Golgi) alone could be sufficient to induce an open conformation of STING and enable proton transport. In this case, agonist binding in cells is mainly essential for translocation of STING to this acidic organelle. Importantly, C53 still significantly reduced STING-driven proton transport *in vitro*.

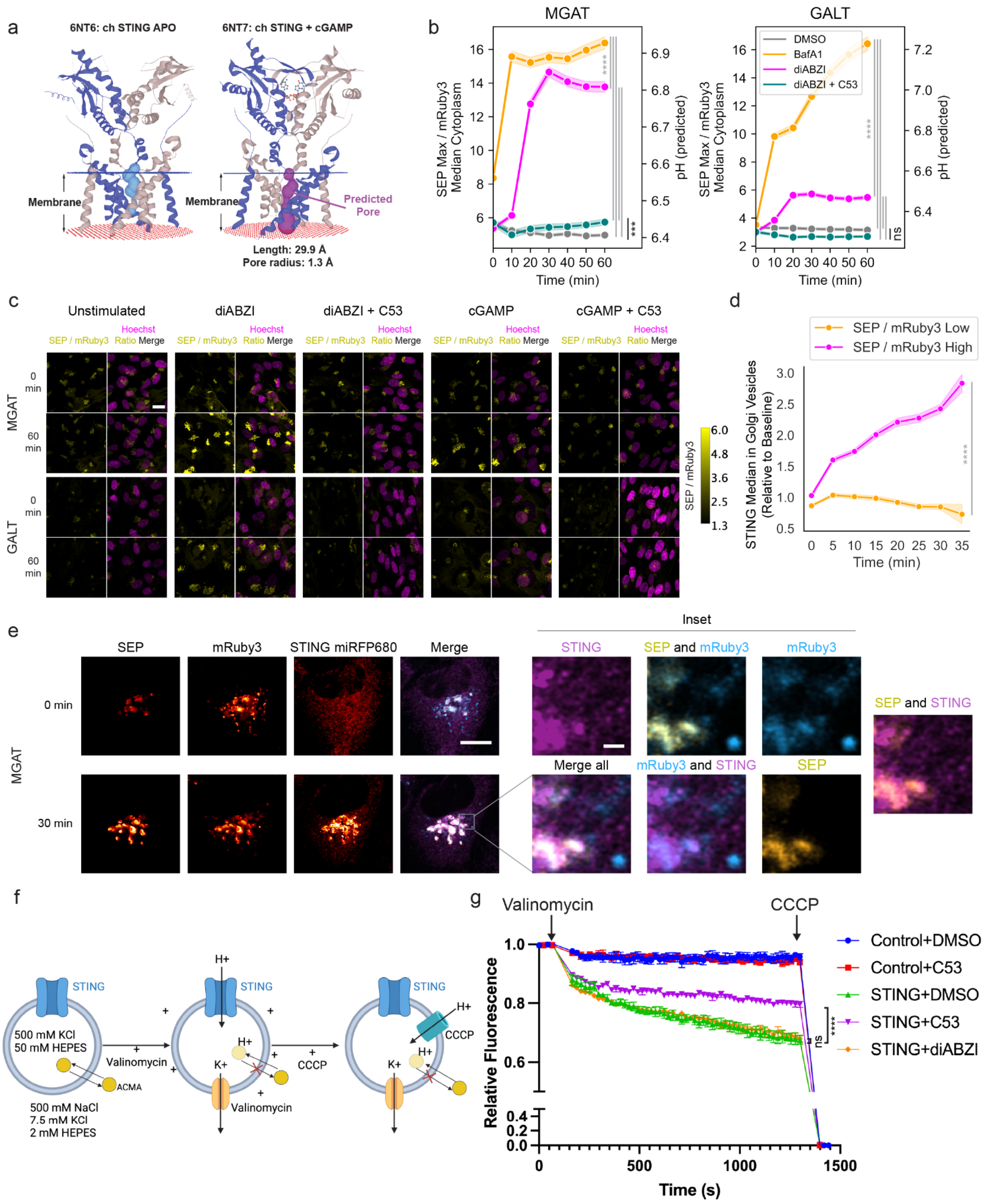

a
6NT6: ch STING APO
6NT7: ch STING + cGAMP
Membrane
Membrane
Predicted Pore
Length: 29.9 Å
Pore radius: 1.3 Å
b
MGAT
GALT
SEP Max / mRuby3 Median Cytoplasm
pH (predicted)
Time (min)
DMSO
BafA1
diABZI
diABZI + C53
****
***
ns
c
Unstimulated
diABZI
diABZI + C53
cGAMP
cGAMP + C53
SEP / mRuby3 Ratio
Hoechst Merge
MGAT
GALT
0 min
60 min
SEP / mRuby3
6.0
4.8
3.6
2.5
1.3
d
SEP / mRuby3 Low
SEP / mRuby3 High
STING Median in Golgi Vesicles (Relative to Baseline)
Time (min)
****
e
SEP
mRuby3
STING miRFP680
Merge
Inset
STING
SEP and mRuby3
mRuby3
SEP and STING
Merge all
mRuby3 and STING
SEP
MGAT
0 min
30 min
f
STING
500 mM KCl
50 mM HEPES
ACMA
500 mM NaCl
7.5 mM KCl
2 mM HEPES
Valinomycin
H+
K+
CCCP
g
Valinomycin
CCCP
Relative Fluorescence
Time (s)
Control+DMSO
Control+C53
STING+DMSO
STING+C53
STING+diABZI
****
ns

**Figure 4.2.**

**A pore-binding small molecule inhibits pH increase in cells stimulated with STING agonist and STING transports protons in an *in vitro* liposome assay.** (A) Predicted pore for chicken cGAMP-bound STING but not the apo conformation; key parameters calculated using MOLEonline. (B) Quantification of pH increase in BJ1 cells from 0 to 60 minutes post 1 μM diABZI or 1 μM BafA1 stimulation with or without 10 μM C53, data from three biological replicates combined. One-way ANOVA followed by Tukey's HSD at the endpoint measurements. **** indicates $p < .0001$, *** $p < .001$, ns $p > 0.05$. (C) Representative images of BJ1 cells in (B) and Supplementary Fig. 4.2b, scale bar 20 μm. (D) Quantification of super-resolution Airyscan images of BJ1 MGAT SEP mRuby3 STING KO cells overexpressing STING WT miRFP680 stimulated with 1 μM diABZI, representing four biological replicates and five individual cells. STING intensity was normalized to the per-cell baseline median intensity. Two-tailed student's t-test at the endpoint measurements, **** indicates $p < 0.0001$. (E) Representative super-resolution Airsycan images of BJ1 cell from (D) at 0 minutes and 30 minutes post 1 μM diABZI stimulation. Scale bar 10 μm, inset scale bar 1 μm. (F) Schematic of the ACMA-based fluorescence flux assay. A $K^+$ gradient is generated by different $K^+$ concentrations in the external and internal buffers. If STING conducts $H^+$, then the addition of the $K^+$ ionophore valinomycin will drive $H^+$ influx through STING. The fluorescent ACMA dye is sequestered and quenched upon pH changes induced by proton transport from the external buffer to the internal liposomes, resulting in reduced total ACMA fluorescence in the well. The protonophore carbonyl cyanide m- chlorophenyl hydrazone (CCCP), is added to the end of the assay to allow $H^+$ influx in all liposomes as a positive control. (G) ACMA-based fluorescence influx assay using preformed liposomes loaded with STING protein (protein:lipid at a 1:200 mass ratio) or matched detergent micelle containing buffer (Control). Loaded liposomes were treated with DMSO, 100 μM C53, or 1 μM diABZI. One representative experiment of n = 4 experiments using two distinct batches of purified STING protein. Two-way ANOVA followed by Tukey's HSD at the endpoint measurements. **** indicates $p < .0001$, ns indicates $p > 0.05$. For multiple comparisons, only "STING + diABZI" vs "STING + DMSO" and "Control + DMSO" vs "Control + C53" have ns p value, comparison between other groups all have $p < .0001$.

Given the observed impairment of STING-mediated ion leakage upon treatment with C53 both in cells and *in vitro*, we next asked if C53 could inhibit other downstream functions of STING activation. We first tested whether STING-induced LC3B lipidation could also be inhibited by C53. Indeed, we found that treatment with both cGAMP and non-cyclic dinucleotide agonists MSA2 or diABZI induced LC3B lipidation, while co-treatment with C53 strongly impaired LC3B lipidation without associated inhibition of STING phosphorylation or STING translocation (**Fig. 4.3a-c, Supplementary Fig. 4.3a, b**). By contrast, C53 co-treatment did not greatly inhibit LC3B lipidation induced by nigericin, an ionophore that induces non-canonical LC3B lipidation independently of STING (Jacquin et al., 2017) (**Fig. 4.3b, c**), suggesting that C53's activity is specific to STING-dependent LC3B lipidation. To further exclude a STING-independent effect for C53, we knocked out endogenous STING in BJ1 cells expressing the cis/medial Golgi pH reporter and overexpressed wild-type STING or STING S53L (**Supplementary Fig. 4.3c**), a STING variant previously shown to have reduced binding to C53 (Lu et al., 2022). We then measured pH changes upon stimulation with diABZI and observed that C53 co-treatment inhibited agonist-mediated pH increases in cells expressing wild-type STING but had no significant effect in cells expressing STING S53L (**Fig. 4.3d, e**). Similarly, we found that 293T cells stably transduced with STING S53L exhibited reduced sensitivity to C53-mediated impairment of LC3B lipidation induced by diABZI treatment relative to cells expressing wild-type STING (**Fig. 4.3f**).

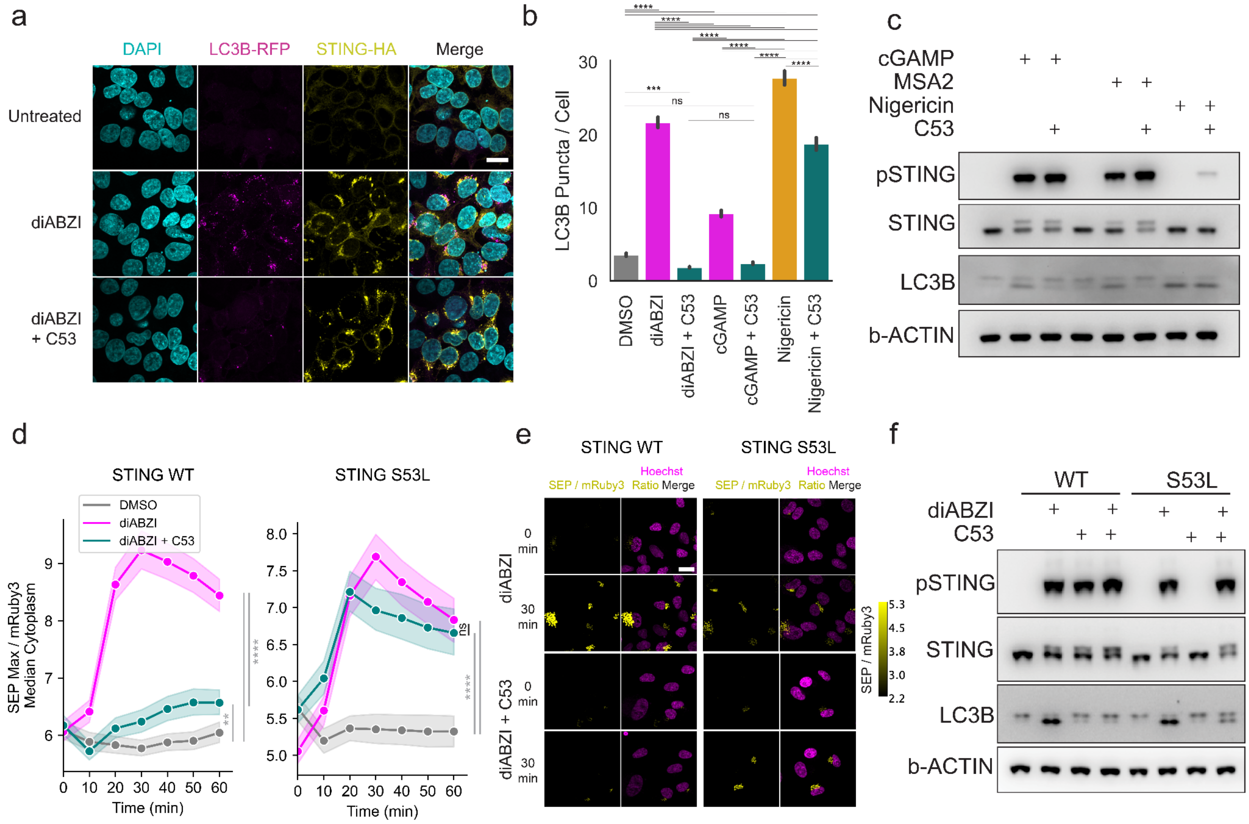


**Figure 4.3.**

**STING-induced LC3B lipidation is inhibited by C53 and STING S53L is less sensitive to C53-mediated inhibition of pH increase and LC3B lipidation.** (A) Representative images of stably expressed RFP-LC3B and STING-HA in FIP200 KO 293T cells upon 1 μM diABZI stimulation for 1 hour with or without 10 μM C53 co-treatment. Scale bar 20 μm. (B) Quantification of experiment in (A), representing three biological replicates combined. One-way ANOVA followed by Tukey's HSD, **** indicates $p < .0001$, *** indicates $p < .001$, ns $p > 0.05$. (C) Immunoblots for indicated proteins in BJ1 cells with or without co-treatment with 10 μM C53 upon 20 μg/ml cGAMP (permeabilized with 5 μg/ml digitonin), 40 μM MSA2, or 2 μM nigericin stimulation. One representative experiment of n = 3 experiments. (D) Quantification of pH change from 0 to 60 minutes post 1 μM diABZI stimulation with or without 10 μM C53, data from three biological replicates combined. STING was knocked out in BJ1 cells followed by overexpression of STING WT (left) or STING S53L (right). One-way ANOVA followed by Tukey's HSD at the endpoint measurements, **** indicates $p < .0001$, ** $p < .01$, ns $p > 0.05$. (E) Representative images of BJ1 cells assayed in (D), scale bar 20 μm. (F). Immunoblots of indicated proteins in 293T cells expressing STING WT or STING S53L treated with 1 μM diABZI with or without 10 μM C53 for 1 hour. One representative experiment of n = 3 experiments.

In addition to induction of interferon and LC3B lipidation, STING activates the NLRP3 inflammasome in human myeloid cells (Gaidt et al., 2017), but the mechanism remains unclear. The influenza virus M2 pore protein induces non-canonical LC3B lipidation by inducing proton leakage (Ulferts et al., 2021), while also activating the NLRP3 inflammasome (Ichinohe et al., 2010), so we hypothesized that STING might activate the inflammasome in a similar manner, and that C53 would block this activity. Upon activation, NLRP3 translocates from the cytosol to Golgi vesicles, where it initiates downstream inflammasome activation (J. Chen & Chen, 2018). Using an NLRP3-mNeonGreen reporter, we found that NLRP3 formed puncta upon stimulation with the STING agonist diABZI (**Fig. 4.4a**), similar to when cells were stimulated with the NLRP3 agonist nigericin (**Supplementary Fig. 4.4a**) (J. Chen & Chen, 2018). In addition, we found that NLRP3 colocalized with STING and pSTING on these puncta (**Fig. 4.4a**). Consistent with the hypothesis that STING-induced proton leakage is the driver of downstream NLRP3 activation, when cells were treated with both diABZI and C53, we observed a significant reduction in NLRP3 translocation together with an enhancement in STING phosphorylation (**Fig. 4.4b, c**). We also tested whether STING-induced LC3B lipidation could have a role in STING-induced inflammasome activation by knocking out ATG16L1 in BLaER1 cells (**Supplementary Fig. 4.4b**), a human cell line that can be transdifferentiated to monocytes and in which STING activation leads to NLRP3-dependent IL-1β release (Gaidt et al., 2017). We found that knockout of ATG16L1 did not impair diABZI-induced IL-1β release (**Supplementary Fig. 4.4c**). This suggests that STING-induced inflammasome activation is independent from STING-induced LC3B lipidation. Finally, we tested whether C53 could block STING-induced inflammasome activation as measured by IL-1β release and cell death in primary human CD14$^+$ monocytes (**Fig. 4.4d**). In agreement with our findings in HEK293T cells, C53 co-treatment significantly impaired STING-induced inflammasome activation, inhibiting IL-1β release (**Fig. 4e, f**) and cell death (**Supplementary Fig. 4.4d**) in primary human monocytes stimulated with cGAMP or diABZI to a level similar to that of monocytes treated with the NLRP3 inhibitor MCC950. Importantly, C53 did not affect IL-1β release (**Fig. 4.4e, f**) or cell death (**Supplementary Fig. 4.4c, d**) when the NLRP3 inflammasome was activated by nigericin, further indicating that C53 impairs activation of the NLRP3 inflammasome in a STING-specific manner. These results suggest that, similar to the influenza protein M2, STING activates the NLRP3 inflammasome through induction of a proton leakage.

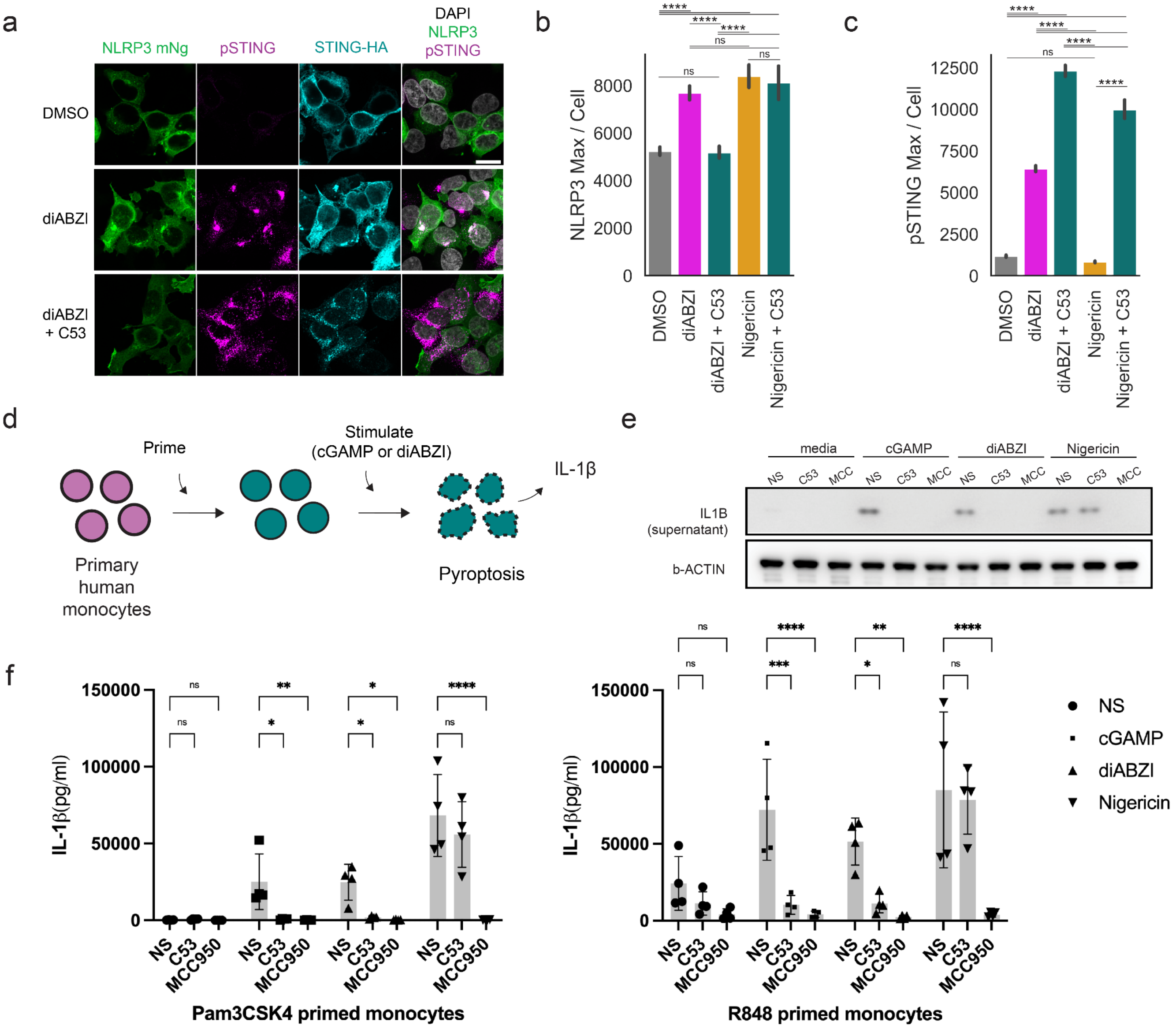


**Figure 4.4.**

**STING-induced NLRP3 inflammasome activation and IL-1β release is inhibited by C53.** (A) Representative images of pSTING, STING, and NLRP3 in HEK293T cells expressing STING-HA and NLRP3-mNeongreen (NLRP3-mNg) treated with DMSO and 1 μM diABZI with or without 10 μM C53 for 1 hour, scale bar 20 μm. (B) NLRP3 translocation quantified as the per-cell maximum NLRP3 intensity from experiment in (A) from three biological replicates combined. One-way ANOVA followed by Tukey's HSD. **** indicates $p < 0.0001$, ns $p > 0.05$. (C) Same as (B) but quantifying pSTING intensity. (D) Experimental workflow for inflammasome induction in primary human monocytes. (E) Immunoblots of processed IL-1β from human monocytes (primed with R848) upon NS (no stimulus), 10 μg/ml cGAMP, 1 μM diABZI, or 6.7 μM nigericin stimulation in the absence or presence of 10 μM C53 or the 5 μM NLRP3 inhibitor MCC950 (MCC). One representative donor of n=3 donors tested. (F) Supernatant cytokine measurement from stimulated human monocytes (left: Pam3CSK4 primed, right: R848 primed) of processed IL-1β upon no stimulation (NS), 10 μg/ml cGAMP, 1 μM diABZI, or 6.7 μM nigericin stimulation in the absence or presence of 10 μM C53 or the 5 μM NLRP3 inhibitor MCC950, each data point represent one donor with total n = 4 donors. Two-way ANOVA followed by Dunnett's multiple comparisons test was used. "ns" indicates $p > .05$, * indicates $p < .05$, *** indicates $p < .001$ **** indicates $p < .0001$.

Taken together, in this work we show that STING activation induces proton leakage at the Golgi through a channel formed at the interface of the STING homodimer's transmembrane domains. This STING-mediated pH increase is inhibited by the small molecule C53, while no inhibition of pH increase was observed in cells expressing STING S53L, a STING variant with reduced binding to C53 (Lu et al., 2022). We further show that STING transports protons in an *in vitro* liposome assay and C53 treatment inhibits STING proton transport *in vitro*.

In addition to demonstrating proton transport through purified STING *in vitro* and STING-dependent pH changes in cells, we also show that STING's channel activity is critical for downstream activation of LC3B lipidation and of the NLRP3 inflammasome, as treatment with C53 impairs these activities without reducing STING phosphorylation. Notably, C53 treatment also inhibits NLRP3 inflammasome activation downstream of STING in primary human monocytes. These findings provide an avenue for decoupling STING phosphorylation from induction of LC3B lipidation and inflammasome activation induced by STING. Comparing the effects of agonists that bind STING's natural binding pocket or the pore-associated pocket, such as C53, could help determine the relative importance of STING phosphorylation versus channel-mediated downstream functions in diverse biological contexts. From a therapeutic perspective, C53's binding to STING's proposed pore area and its dual function of activating STING translocation while blocking STING-induced proton flux provides a unique paradigm for designing additional STING agonists that enhance STING-mediated interferon production without inducing downstream effects mediated by STING's channel activity.

## 4.3 Materials and Methods

### Cell lines and constructs

293T (CRL-3216) and hTert-BJ1 (BJ-5ta - CRL-4001) were from ATCC. 293T and BJ1 were cultured in DMEM (Corning) supplemented with 10% FBS (VWR), 1X GlutaMax (Thermo Fisher) and 1X Penicillin/Streptomycin (Corning). The plasmids psPAX2 (#12260) and pCMV-VSV-G (#8454) were from Addgene. pTRIP-SFFV/PGK-Blast-2A-STING-HA was from an earlier publication(*1*). pTRIP-SFFV-Blast-2A-STING(H50R)/(S53L)-HA were cloned through PCR primers containing mutant sequences. pTRIP-SFFV-Blast-2A-NLRP3-mNG and pTRIP-SFFV-Hygro-2A-RFP-LC3B were cloned from Gibson assembly using gBlocks. Plasmids for pH imaging are described in detail in the "Organelle pH Measurements and Image Acquisition" section. *FIP200* KO 293T cells used for genetic screens were generated by transfecting 293T cells with pXPR_023 (Addgene #52961), encoding Cas9 and an sgRNA against *FIP200 (1)* and selecting cells with 2 µg/mL puromycin (Thermo Fisher Scientific #A1113803) for 2 days. *FIP200* KO 293T cells were then transduced with pTRIP-PGK-Hygro-P2A-RFP-LC3B, treated with 33 µg/mL hygromycin (Invivogen, #ant-hg-1) for 3 days and then sorted to obtain a homogenous RFP+ population.

### Genome-wide CRISPR screen

RFP-LC3B *FIP200* KO 293T cells were transduced with pTRIP-PGK-Blast-P2A-STING-HA and selected with 20 µg/mL blasticidin (Thermo Fisher Scientific #A1113903). For screening, 200M RFP-LC3+ STING-HA+ *FIP200* KO 293T cells were transduced with Cas9-sgRNA all-in-one Brunello library at MOI=0.4 and selected with 2 µg/mL puromycin for 2 days. 8 days after transduction, 200M transduced cells were treated with 1 µM diABZI (Invivogen, #tlrl-diabzi) for 2.5hr, and then permeabilized with 1X perm buffer (PBS with 0.05% saponin and 0.1% glycine) for 4 min *(2)*. Permeabilized cells were then washed with 1X PBS, and fixed, permeabilized further, and stained using BD Cytofix/Cytoperm kit (#554714) following the instruction with anti-HA Alexa647 (Biolegend, #682404) 1:200 dilution.

### Mutagenesis screen

RFP-LC3B *FIP200* KO 293T cells were infected with saturation mutagenesis ORF library lentivirus at MOI = 0.2, and selected with 20 µg/ml blasticidin for 2 days. 150M transduced cells were treated with diABZI for 1.5hr, and then permeabilized with 1X perm buffer (PBS with 0.05% saponin and 0.1% glycine) for 4 min. Permeabilized cells were then washed with 1X PBS, and fixed, permeabilized further, and stained using BD Cytofix/Cytoperm kit (#554714) following the instruction with anti-pSTING Alexa647 (CST, #43499) at 1:100 dilution.

### Human STING purification

The process described here is adopted from a previous study(*2*). Briefly, Expi293F cells were transduced with SFFV-Blast-P2A-STING-FLAG and culture in Expi293 media (Thermo) with density between 4-6 million/mL for harvest. For each batch of protein purification, 3 L of cells were pelleted and resuspended in 50 mL buffer A (20 mM Tris-HCl pH 8.0, 150 mM NaCl, with protease inhibitor (Roche)), and then sonicated (20s total, 50% on time, 5 power using a Branson digital sonicator, repeated 3 times to break down all cells efficiently). Sonicated cells were centrifuged at 5000 g for 10 minutes to remove debris, and all the supernatant was collected and ultracentrifuged (100,000 g, 1 hour using Beckman Optima XPN90). After removing the supernatant, the pellets were resuspended with 25 mL Buffer A with 1.5% 10:1 DDM/CHS (Anatrace, D310-CH210), and then centrifuged at 5000 g, 10 min to remove all insoluble debris. The supernatant was collected and added to 1 mL of Buffer A pre-washed FLAG M2 beads (Sigma, 2426-1ML or M8823) and incubated at 4ºC for 1 hr. Beads were then washed with wash buffer (25 mM Tris-HCl pH 8.0, 300 mM NaCl, 10:1 DDM/CHS 0.1%) 3 times and then eluted with 0.5 mL elution buffer (25 mM Tris-HCl pH 8.0, 200 mM NaCl, 10:1 DDM/CHS 0.1%, 0.6 mg FLAG peptide (Sigma)) at 4ºC for 3 hr. Eluted samples were loaded on an FPLC (AKTA) to perform Size Exclusion Chromatography (SEC) using a Superdex 200 increase 10/300 column (Cytiva) using a running buffer (25 mM Tris-HCl pH 8.0, 150 mM NaCl, 10:1 DDM/CHS 0.03%).

### Liposome synthesis.

Liposomes were generated via the thin film hydration method. Briefly, a mixture of POPE (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine) and POPG [1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1’-rac-glycerol)] (3:1, wt/wt) in chloroform (Avanti Polar Lipids) was dried in a rotovap (Rotovapor RII, BUCHI) to generate a lipid film(*3*). The film was hydrated with a high potassium buffer (500 mM KCl, 50 mM HEPES, pH 7.3) at 100 mg/mL and sonicated at 65°C until monodisperse liposomes were formed. Liposomes underwent 7 cycles of freeze-thawing and were sonicated in a water bath until they reached 200-300 nm in size. Size was measured via dynamic light scattering using a Zetasizer Nano ZSP (Malvern).

### Protein incorporation into lipid vesicles.

STING was loaded onto preformed liposomes as described previously for other channel proteins(*4*). Briefly, liposomes were diluted to 10 mg/mL using a low potassium buffer (500 mM NaCl and 2 mM HEPES, pH 7.3), then mixed with purified STING protein on detergent micelles. The detergent concentration was above its critical micelle concentration (CMC) to enable protein loading onto liposomes, but not sufficient to disrupt the lipid bilayer (*5*). Samples were then allowed to incubate for 30 minutes prior to assaying for proton flux. Control liposomes were generated via the same protocol, but were mixed only with the same buffer containing empty detergent micelles.

Proteoliposome proton flux assay.

The proton flux assay was performed following previously established protocols(*6*). Liposomes or proteoliposomes were diluted to 1 mg/mL or less in the low potassium buffer with 2 μM of ACMA (9-Amino-6-Chloro-2-Methoxyacridine, Thermo Fisher #A1324). The ACMA fluorescence (excitation at 419 nm and emission at 490 nm) was measured on a black 384-well plate for at least 1 minute to obtain a baseline fluorescence on an Infinite M200 pro (TECAN) plate reader. To initiate proton flux, 0.04 μM of valinomycin (Thermo Fisher, #V1644) was added to each sample and mixed. Fluorescence readings were then performed at 20-30 second intervals until no substantive changes were further observed (~10-15 minutes). To terminate the assay, 4 μM of CCCP (carbonyl cyanide 3-chlorophenylhydrazone, Sigma #C2759) was added to each sample and mixed and the fluorescence readings were measured for at least 1 minute. Relative fluorescence was calculated based on the difference between the measured fluorescence and the final fluorescence relative to the difference between the initial fluorescence and the final fluorescence. The same process was performed on the control liposomes.

Organelle pH Measurements and Image Acquisition

hTERT-immortalized BJ1 cells (ATCC CRL-2522) were transduced with lentiviral ratiometric reporters targeted to MGAT, GALT, or LAMP1 constructed based on previously reported designs(*7*) (modified from Addgene plasmid # 171718, Addgene plasmid # 171719, and Addgene plasmid # 171720) with superecliptic pHluorin (Addgene plasmid # 32001) and mRuby3 (Addgene plasmid #127808). Transduced cells were sorted based on mRuby3 expression using a Sony MA900 sorter. For imaging assays, BJ1 SEP mRuby3 cells were plated in 24-well glass-bottom plates (Greiner Bio-One) at 40,000 cells/well. After 48 hours, cells were stained for 45 minutes at 37°C with 0.5 μg/ml Hoechst 34580 (Thermo Fisher Scientific, cat. #H21486). Cells were then washed and incubated in Fluorobrite DMEM (Thermo Fisher Scientific, cat.

#A1896701) medium supplemented with 10% FBS, 1% Pen-strep, and 1x GlutaMAX (Thermo Fisher Scientific, cat. #35050061). For pH calibration experiments, cells were treated with buffers at defined pH values supplemented with nigericin and valinomycin following manufacturer's instructions (Thermo Fisher Scientific cat. # P35379). Additional buffers at pH 7, 6, and 5 were created by titration with HCl. For time-course experiments, cells were stimulated with 1 µM Bafilomycin A1 (Santa Cruz Biotechnology cat. #sc-201550), 1 µM diABZI (Invivogen, #tlrl-diabzi), or 20 µg/mL cGAMP (Invivogen, #tlrl-nacga23-1) with 5 ng/µL digitonin (Promega, #G9441) for 1 hr with or without the addition of 10 µM C53 (Cayman, #37354). For experiments assaying the effect of STING S53L, hTERT-immortalized BJ1 cells expressing SEP mRuby3 targeted to cis/medial Golgi (MGAT) were transduced with pXPR023 (lentiCRISPRv2) expressing an sgRNA targeting STING and selected with 0.1 µg/mL puromycin for 5 days. Cells were then transduced with blasticidin-STING-HA (WT or S53L) and selected using 10 µg/mL blasticidin HCl for 5 days. All images were acquired using a Ti2-E inverted epifluorescence microscope (Nikon) with automated XYZ stage control, hardware autofocus, and a Yokogawa CSU-W1 confocal spinning disk unit with Zyla 4.2 PLUS sCMOS camera. An Okolab cage incubator was set to 37°C with 5% $CO_2$. 405, 488, 561, and 640 nm laser lines were used for fluorescence illumination and all hardware was controlled using NIS Elements software. Images were acquired using a 40X 0.95 NA CFI Plan Apo λ objective (Nikon MRD70470) with the following lasers and filters: Hoechst (405 nm laser, Chroma Multi LED set #89401), superecliptic pHluorin (488 nm laser, Chroma Multi LED set #89401), and mRuby3 (561 nm laser, Chroma Multi LED set #89401), assaying three z planes per field of view with 1.25 µm spacing. Fields of view were selected using NIS Elements software coordinates without manual preselection.

### Autophagy Induction and Image Acquisition

RFP-LC3B and STING-HA-expressing FIP200 KO 293T cells were seeded on Fibronectin bovine plasma coated 24-well glass-bottom plates (Greiner Bio-One) the night before stimulation. Cells were then stimulated with 20 µg/ml cGAMP (Invivogen, #tlrl-nacga23-1) with 5 µg/ml digitonin (Promega, #G9441) or 1 µM DIABZI (Invivogen, #tlrl-diabzi) with or without the addition of 10 µM C53 (Cayman, #37354) for 1 hour. Cells were then fixed with 2% Paraformaldehyde (Electron Microscopy Sciences, #15710) in PHEM buffer (Electron Microscopy Sciences, #11162) for 30 minutes at 37°C, washed three times with PBS and quenched with freshly prepared 0.1M Glycine (Sigma, #50046) for 10 minutes. Cells were permeabilized in 100% methanol for 30 minutes and stained with anti-HA (Millipore, #11867423001) for 1 hour at room temperature in 3% BSA, washed 5 times, and then stained with Alexa 647 anti-rat IgG (H+L) (Thermo, A-21247) in 3% BSA for 1 hour. After five washes, cells were incubated in 2X SSC with 200 ng/mL DAPI (Thermo

Fisher) and imaged using the Nikon microscope used for organelle pH images. Images were acquired using a 60X 1.40 NA Plan Apo λ oil immersion objective (Nikon MRD01605) with Nikon type F immersion oil with the following lasers and filters: DAPI (405 nm laser, Chroma ET455/50), RFP-LC3B (561 nm laser, Chroma ET605/52), and STING-HA (640 nm laser, Chroma ET705/72), assaying five z planes per field of view with 0.625 μm spacing. Fields of view were selected using NIS Elements software coordinates without manual preselection.

### NLRP3 Stimulation and Image Acquisition

HEK293T cells transduced to express NLRP3-mNeonGreen and STING-HA were plated in 24-well glass-bottom plates (Greiner Bio-One) and, after 24 hours, stimulated with 2 μM nigericin or 1 μM diABZI with or without the addition of 10 μM C53 for 1 hour. Cells were then fixed with 2% Paraformaldehyde (Electron Microscopy Sciences, #15710) in PHEM buffer (Electron Microscopy Sciences, #11162) for 30 minutes at 37°C, washed three times with PBS and quenched with freshly prepared 0.1M Glycine (Sigma, #50046) for 10 minutes. Cells were permeabilized in 100% methanol for 30 minutes and stained with anti-HA (Millipore, #11867423001) and anti p-STING (Cell Signaling Technology cat. #19781s) for 1 hour at room temperature in 3% BSA, washed 5 times, and then stained with Alexa 647 anti-rat IgG (H+L) (Thermo, A-21247) and Alexa 555 plus anti-rabbit (Thermo, A32732) in 3% BSA for 1 hour. After five washes, cells were incubated in 2X SSC with 200 ng/mL DAPI (Thermo Fisher) and imaged using the Nikon microscope used for organelle pH images. Images were acquired using a 60X 1.40 NA Plan Apo λ oil immersion objective (Nikon MRD01605) with Nikon type F immersion oil with the following lasers and filters: DAPI (405 nm laser, Chroma ET455/50), NLRP3 mNeonGreen (488nm laser, Chroma ET525/36) pSTING (561 nm laser, Chroma ET605/52), and STING-HA (640 nm laser, Chroma ET705/72), assaying five z planes per field of view with 0.625 μm spacing. Fields of view were selected using NIS Elements software coordinates without manual preselection.

### Super-resolution Airyscan Imaging

hTERT-immortalized BJ1 cells (ATCC CRL-2522) expressing SEP mRuby3 targeted to cis/medial Golgi (MGAT) were transduced with pXPR023 (lentiCRISPRv2) expressing an sgRNA targeting STING and selected with 0.1 μg/mL puromycin for 5 days. Cells were then transduced with blasticidin-STING-miRFP680 (constructed using Addgene plasmid # 136557) and selected using 10 μg/mL blasticidin HCl for 5 days. Cells were plated in 96-well glass-bottom plates (Greiner Bio-One) at 6,000 cells/well. After 48 hours, cells were incubated in Fluorobrite DMEM (Thermo

Fisher Scientific, cat. #A1896701) medium supplemented with 10% FBS, 1% Pen-strep, and 1x GlutaMAX (Thermo Fisher Scientific, cat. #35050061) and stimulated with 1 μM diABZI (Invivogen, #tlrl-diabzi). All images were acquired using an LSM980 with Airyscan2 (Zeiss) with 37°C and 5% $CO_2$ incubation. 8 z-stacks were acquired with 0.15 μm z-step. Images were acquired using a 63X 1.40 NA DIC M27 objective with Immersol 518F 37°C oil. Acquired images were Airyscan processed and then analyzed as described in the image analysis section.

### Western Blotting

Cells were seeded at 0.15 million cells (BJ1) or 0.2 million cells (293T) per well in a 24 well plate the night before stimulation. Cells were then stimulated with 1 μM diABZI (Invivogen, #tlrl-diabzi), 2 μM Nigericin (Invivogen, #tlrl-nig), 20 μg/mL cGAMP (Invivogen, #tlrl-nacga23-1) with 5 μg/mL digitonin (Promega, #G9441) or 40 μM MSA2 (Invivogen, #tlrl-diabzi) with or without the addition of 10 μM C53 (Cayman, #37354) for 1 hour. Cells were then harvested and lysed in 100 μL 1X SDS sample buffer (Boston Bio, #BP111NR) and boiled for 10 mins. Samples were run on NuPAGE 4 to 12% Bis-Tris Gels (Thermo Fisher) and transferred on nitrocellulose membrane with an iBlot2 (Thermo Fisher). Membranes were blocked in 5% non-fat milk in TBS Tween. Antibodies against phospho proteins were incubated in 5% BSA TBS tween. ECL signal was recorded on a ChemiDoc Biorad Imager. Data was analyzed with ImageLab (Biorad).

### Monocyte Experiments

CD14+ monocytes were isolated from peripheral adult human blood using CD14 Microbeads (Miltenyi) as previously described *(3)*. CD14+ monocytes were cultured in RPMI (Gibco) supplemented with 10% FBS (VWR), 1X GlutaMax (Thermo Fisher), and 1X Penicillin/Streptomycin (Corning). Monocytes were plated at 400,000 cells/well and stimulated with indicated priming reagents (2μg/ml Pam3CSK4, 1μg/ml R848) for 2h, and then stimulated with indicated stimuli (10μg/ml cGAMP, 1μM DiABZI, 10uM C53, 5μM MCC950 ) for 6h. 6.7μM nigericin was added in the last 2h of stimulation. All the stimuli were from Invivogen except C53 (Cayman). Stimulated cells were then centrifuged to separate cells and supernatant for PI staining, Legendplex assay and Immunoblot. Cytokine quantification was performed following standard Legendplex protocol using anti-IL1b beads (Biolegend, #740812). PI staining was performed in 120μl/well with a 1:200 PI (Biolegend) dilution. Cells for PI staining and beads for Legendplex were acquired on a Cytoflex LX (Beckman Coulter). Data was analyzed with FlowJo v10.7 (BD). IL-1β concentration was calculated using Graphpad Prism based on mean fluorescence intensity of a standard curve.

### BlaER1 Experiments

BLaER1 were transduced with pXPR023-ATG16L1_g01 or pXPR023-ATG16L1_g02, selected with 2 µg/mL Puromycin and then trans-differentiated into monocytes as described in Gaidt et al 2018(*8*). Briefly, 70,000 cells/well were plated in a 96 well flat bottom with 10 ng/mL IL-3 (Peprotech), 10 ng/mL, M-CSF (Biolegend), 100 nM β-Estradiol (Sigma) in RPMI for 5 days. Before stimulation, the media was replaced and cells were pre-stimulated with 2 µg/ml Pam3CSK4 (Invivogen) for 1 hour. Cells were then stimulated with the combination of drugs indicated in the figure legend for 5 hours. Supernatant was then collected and IL-1β was measured via Legendplex following manufacturer's instructions.

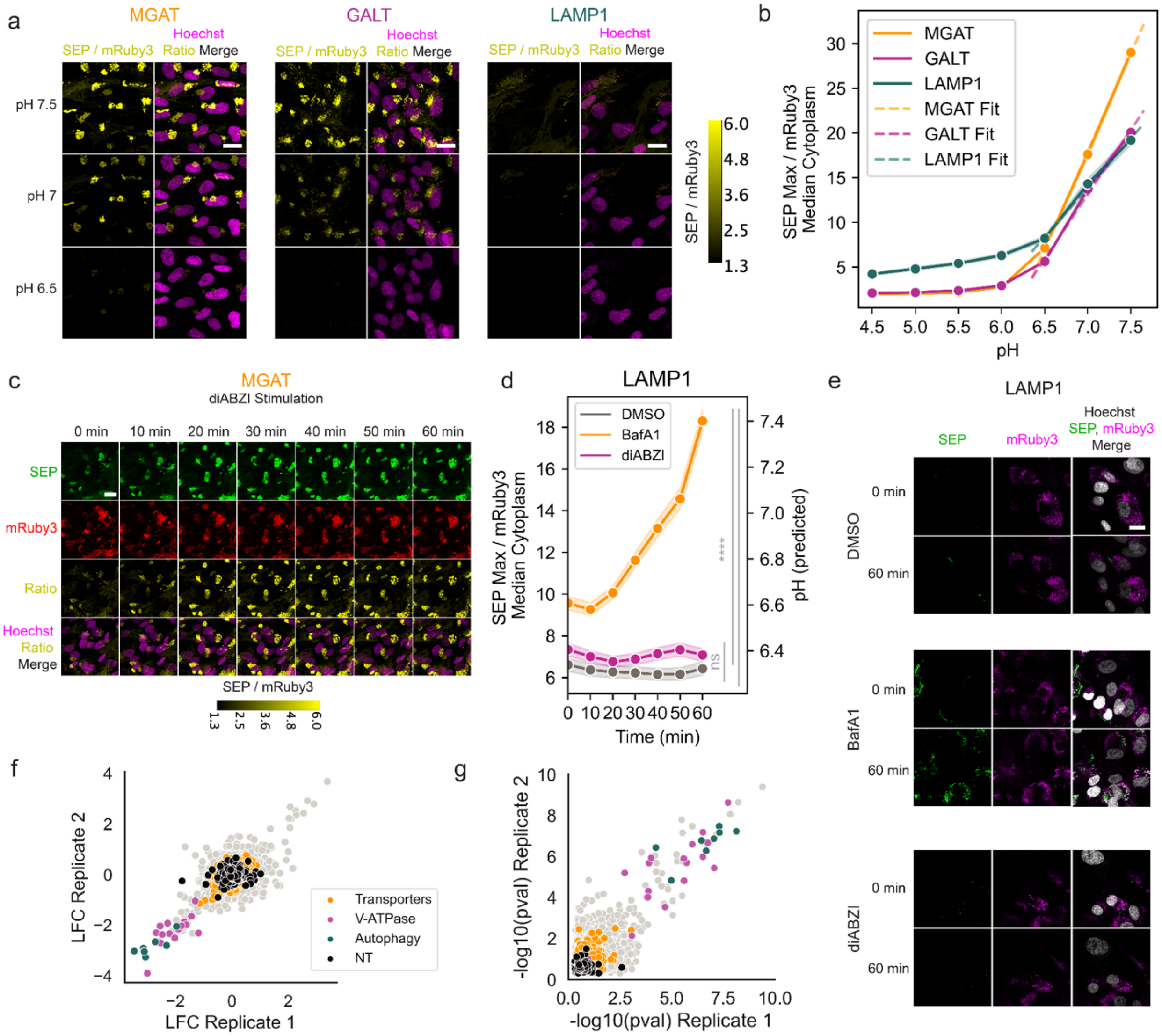
### FACS Screen Analysis

Guide RNA abundances were extracted from FASTQ files using poolq 3.3.2 with fixed row and barcode policies. Resulting log-abundances were then subtracted between sorting bins from each experimental condition and analyzed using the Broad Genetic Perturbation Platform screen analysis tool (https://portals.broadinstitute.org/gpp/public/analysis-tools/crispr-gene-scoring) using a hypergeometric analysis.

### Image Analysis

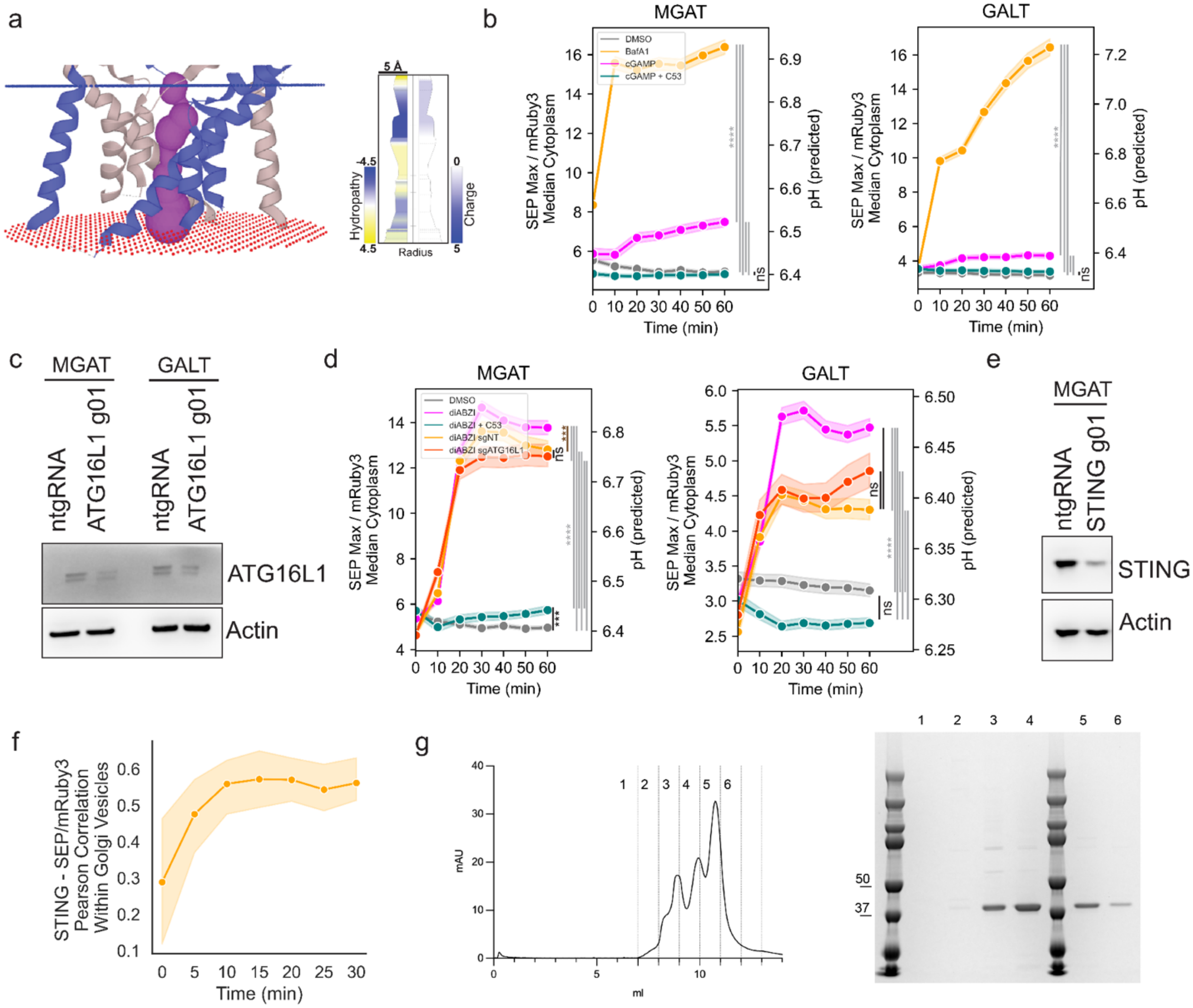
Analysis was performed using code adapted from *(4)*, and all analysis code is available on GitHub. For live-cell confocal imaging experiments with BJ1 cells expressing mRuby3 SEP, analysis was performed using z-stack maximum projection images. Cells were segmented by thresholding the Hoechst signal to identify nuclei and expanding the resulting regions to define cell areas using the seeded watershed method based on a defined threshold in the mRuby3 channel. All experiments used the same mRuby3 thresholds, while Hoechst thresholds were varied as appropriate to account for technical differences in staining levels. Segmented cells were tracked across timepoints and the ratio of maximum superecliptic pHluorin to median mRuby3 in the cytoplasm (defined as the difference between the cell and nuclear masks) was extracted in tracked cells. For pH calibration experiments, a single cell mask was defined at pH = 7.5 using the mRuby3 channel and applied to define cell regions at lower pH values that were subsequently measured. Linear regression (scikit-learn linear_model.LinearRegression) was used to separately determine correspondence between pH values and SEP max / mRuby3 median ratios for each genetically encoded reporter using measurements at pH values 6.5, 7, and 7.5 (the linear range for these reporters). For super-resolution Airyscan analysis, cells were manually masked based on background fluorescence in the STING miRFP680 channel. A single z-stack was selected for analysis and individual Golgi vesicles were segmented at each timepoint based on thresholding in the mRuby3 channel. SEP, mRuby3, and STING miRFP680 intensities were separately extracted from each segmented vesicle. For each cell, initial SEP/mRuby3 ratios and STING intensity were normalized relative to baseline and segmented vesicles were defined as SEP/mRuby3 high if the ratio was >= 1.5-fold relative to the mean baseline ratio for each cell. For fixed cell analysis of LC3B lipidation and inflammasome induction experiments, nuclei were segmented by thresholding the DAPI signal and cells were segmented by expanding the resulting regions using the seeded watershed method. Cells were segmented based on background fluorescence in the STING or NLRP3 mNeonGreen channels for LC3B lipidation and inflammasome experiments, respectively. For the LC3B lipidation experiments, LC3B puncta were segmented based on a defined intensity threshold and area range in that channel. For the inflammasome experiments, around 50% of the cells were removed based on lack of expression of NLRP3 mNeonGreen or STING, as not all cells expressed both constructs. Across all experiments, the only analysis settings varied across replicates were the Hoechst threshold for nuclear segmentation of live-cell imaging experiments.

## 4.4 Supplementary Figures



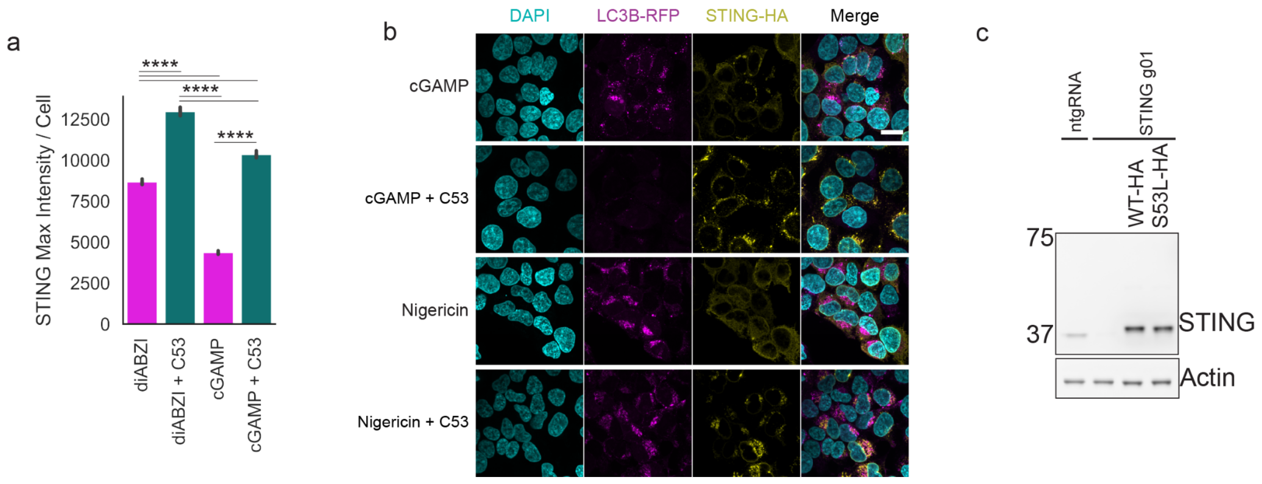
**Supplementary Figure 4.1.**

(A) Representative images of BJ1 cells expressing a ratiometric SEP and mRuby3 reporter localized to MGAT, GALT, or LAMP1 treated with 1 μM nigericin and 1 μM valinomycin and buffers with defined pH. Scale bar 20 μm. (B) Quantification of experiment in (A), data from three biological replicates combined, linear regression fits from pH 6.5 to 7.5 shown. (C) Cropped image from Fig. 1a across all timepoints assayed showing each channel separately as well as the SEP to mRuby3 ratio and Hoechst merge. (D) Quantification of changes in LAMP1 reporter pH upon 1 μM diABZI or 1 μM BafA1 stimulation, data from three biological replicates combined. One-way ANOVA followed by Tukey's HSD at 60 minute timepoint. **** indicates $p < .0001$, ns = $p > 0.05$ (E) Representative images of BJ1 cells in (D), scale bar 20 μm. (F) Replicate log2-fold-change correlations for genome-wide CRISPR screens. (G) Replicate -log10(FDR adjusted p value) correlations for genome-wide CRISPR screens. (A) Representative images of BJ1 cells expressing a ratiometric SEP and mRuby3 reporter localized to MGAT, GALT, or LAMP1 treated with 1 μM nigericin and 1 μM valinomycin and buffers with defined pH. Scale bar 20 μm. (B) Quantification of experiment in (A), data from three biological replicates combined, linear regression fits from pH 6.5 to 7.5 shown. (C) Cropped image from Fig. 4.1a across all timepoints assayed showing each channel separately as well as the SEP to mRuby3 ratio and Hoechst merge. (D) Quantification of changes in LAMP1 reporter pH upon 1 μM diABZI or 1 μM BafA1 stimulation, data from three biological replicates combined. One-way ANOVA followed by Tukey's HSD at 60 minute timepoint. **** indicates $p < .0001$, ns = $p > 0.05$ (E) Representative images of BJ1 cells in (D), scale bar 20 μm. (F) Replicate log2-fold-change correlations for genome-wide CRISPR screens. (G) Replicate -log10(FDR adjusted p value) correlations for genome-wide CRISPR screens.

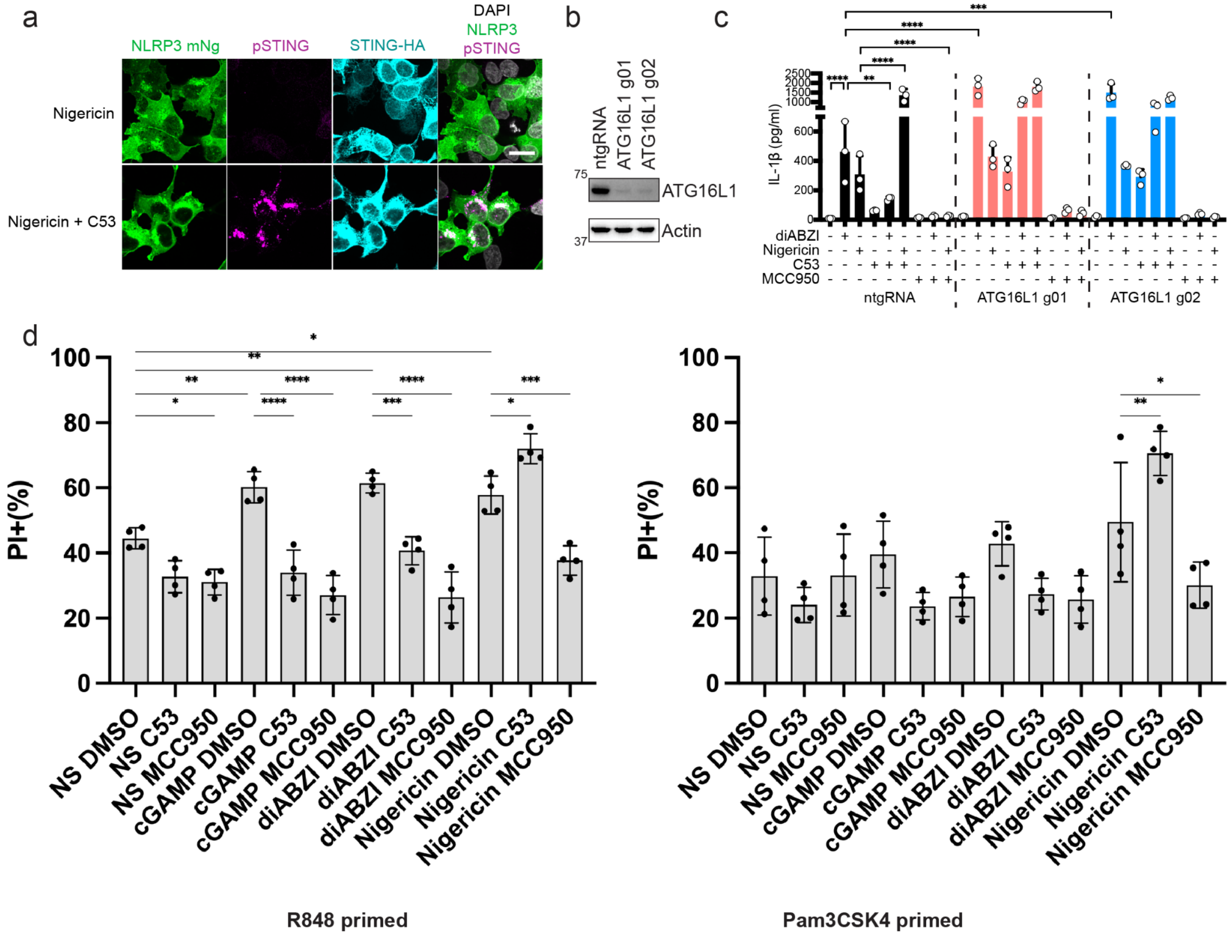


**Supplementary Figure 4.2.**
(A) Radius, charge and hydropathy of the proposed channel area of chicken STING bound with cGAMP (PBD: 6NT7). (B) Quantification of pH in BJ1 SEP mRuby3 MGAT or GALT reporter cells from 0 to 60 minutes post 20 μg/mL cGAMP with 5 μg/ml digitonin or 1 μM BafA1 stimulation, with or without 10 μM C53; data from three biological replicates combined. One-way ANOVA followed by Tukey's HSD at 60 minute timepoint. **** indicates $p < .0001$, ns $p > 0.05$. (C) Immunoblots of indicated proteins in BJ1 SEP mRuby3 MGAT or GALT reporter cells expressing Cas9 and nontargeting sgRNA (ntgRNA) or ATG16L1 targeting sgRNA (ATG16L1 g01). (D) Quantification of pH in BJ1 SEP mRuby3 MGAT or GALT reporter cells from 0 to 60 minutes post 1 μM diABZI stimulation with or without 10 μM C53 and a non-targeting sgRNA or an sgRNA targeting ATG16L1 (g01), data from three biological replicates combined. One-way ANOVA followed by Tukey's HSD at 60 minute timepoint. **** indicates $p < .0001$, ns $p > 0.05$. (E) Immunoblots of indicated proteins in BJ1 SEP mRuby3 MGAT reporter cells expressing Cas9 and nontargeting sgRNA (ntgRNA) or STING targeting sgRNA (STING g01). (F) Quantification of super-resolution Airyscan images of BJ1 MGAT SEP mRuby3 STING KO cells overexpressing STING WT miRFP680 stimulated with 1 μM diABZI, representing four biological replicates and five individual cells. STING intensity and SEP/mRuby3 ratios were normalized to per-cell baseline median intensities and

Pearson correlations for each cell between STING and SEP/mRuby3 in Golgi vesicles were calculated at each timepoint. (G) Size exclusion chromatography of eluted STING protein (left) with coomassie staining (right) of each fractions. One representative experiment of n=3 experiments. Fractions 5 and 6 were combined for proteoliposome experiments.



**Supplementary Figure 4.3.**
(A) STING translocation quantified as the per-cell maximum STING intensity for experiment in Figure 3A, representing three biological replicates combined. One-way ANOVA followed by Tukey's HSD, **** indicates p < .0001. (B) Representative images of stably expressed RFP-LC3B and STING-HA in FIP200 KO 293T cells upon 20 µg/mL cGAMP with 5 µg/ml digitonin or 2 µM nigericin stimulation with or without 10 µM C53 co-treatment. Scale bar 20 µm. (C) Immunoblots of indicated proteins from BJ1 SEP mRuby3 MGAT cells transduced with Cas9 with non-targeting sgRNA (ntgRNA) or Cas9 with STING targeting sgRNA (STING g01) with indicated STING WT or STING S53L expression.



**Supplementary Figure 4.4.**

(A) Representative images of pSTING, STING, and NLRP3 in HEK293T cells expressing STING-HA and NLRP3-mNeonGreen and treated with 2 µM nigericin with or without 10 µM C53 for 1 hour. Scale bar 20 µm. (B) Immunoblots of indicated proteins in BLaER1 cells expressing Cas9 and nontargeting sgRNA (ntgRNA) or ATG16L1 targeting sgRNAs (ATG16L1 g01 and g02). (C) IL-1β quantification in supernatant of differentiated BLaER1 KO cells in (B) pre-treated for 1 hour with 2 µg/ml Pam3CSK4 and stimulated for 5 hours with the indicated combinations of drugs (1 µM diABZI, or 6.7 µM nigericin stimulation in the absence or presence of 10 µM C53 or the 5 µM NLRP3 inhibitor MCC950). Data was quantified using three combined biological replicates. One-way ANOVA with post-hoc Tukey test on log-normalized data. ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$. (D) PI staining of R848 (left) or Pam3CSK4 (right) primed monocytes upon NS (no stimulus), 10 µg/ml cGAMP, 1 µM diABZI, or 6.7 µM nigericin stimulation in the absence or presence of 10 µM C53 or the 5 µM NLRP3 inhibitor MCC950. One-way ANOVA with post-hoc Tukey test between all DMSO-treated groups (DMSO plus NS, cGAMP, diABZI, and nigericin) and between matched DMSO-C53/MCC950 groups with the same stimulus (NS, cGAMP, diABZI, and nigericin). * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$, non-significant tests not shown. NS: no stimulation. Each data point represents one donor, total n = 4 donors.

# 5 Conclusions and Outlook

While the studies in this work clearly show the power of functional genomics and, in particular, optical pooled screening approaches to reveal regulators of host-pathogen interactions, further advances are needed to extend the accessibility and utility of this technology to answer similar questions.

First, the imaging setup recommended for optical pooled screening encompasses several affordable options (many wide-field fluorescence microscopes and illuminators are compatible, while specific filter sets may need to be purchased; suggested configurations are discussed in Feldman et al., 2022). However, the imaging time needed for screening is substantial (~1.5 imaging hours per cycle per plate, with genome-wide screens ranging from six to twelve or more 6-well plates and twelve sequencing cycles) so screening on a shared instrument is practically challenging. Furthermore, while costs per cell are much lower than transcriptome sequencing, reagent costs for *in situ* amplification and sequencing are in excess of $1,000 per 6-well plate. Personnel time for screening is also a challenge, as sequencing cycles are currently performed manually on thermocyclers. Finally, custom image analysis and experience with cloud computing is required to analyze large image-based screens. We have released code to call *in situ* sequencing reads from cells (Feldman et al., 2022) but this code will likely require custom adjustments for different imaging configurations, cell types, or perturbation modalities. For optical pooled screening to become accessible to a broad range of researchers, academic or commercial groups offering services covering some or all of the components of optical pooled screening will be required.

In terms of technology development, throughput is currently sufficient for genome-scale screens in tractable cell models such as HeLa cells or other cancer cell lines that can be cultured at high density and have a small per-cell area, but is not well-suited to genome-scale screens in models that require much more surface area per cell (e.g. larger cell lines such as fibroblasts and screening for cell-cell interactions or other phenotypes that may require sparse cell plating) or screens with $>1e^5$ perturbations. The major throughput bottleneck, apart from cost considerations, is automation of *in situ* sequencing cleavage, incorporation, and washing between cycles, which would approximately double throughput. Other optimizations include 1) reducing imaging time (examples include externally-triggered hardware, imaging at lower than 10x magnification, using a larger field-of-view camera, using 2-color rather than 4-color SBS chemistry) and 2) reducing sequencing chemistry time (examples include optimizing washing conditions or using alternative sequencing chemistries with reduced background). Furthermore, some cell types result in higher *in situ* sequencing efficiency than others, which is likely partially explained by differences in mRNA expression levels. Cell-type-specific promoters or optimization of fixation and enzymatic conditions for the cell model of interest may improve efficiency but may also require substantial time investment.

## 5.1 Outlook for Analysis of Optical Pooled Screens

### *5.1.1 Pooled Genetic Loss-of-Function Screen Analysis*

There are numerous approaches for analyzing pooled genetic loss-of-function screens, including BAGEL2 (Kim & Hart, 2021), a Z-score model (Lenoir et al., 2021), Ceres (Morgens et al., 2017), and Chronos (Dempster et al., 2021). Benchmarking studies in one instance showed that Chronos, which explicitly models cell proliferation after gene knockout and performs copy number correction, outperformed other methods in identifying essential genes (Dempster et al., 2021), while in another instance revealed that BAGEL2 and Ceres outperformed Chronos in the construction of co-essentiality networks (Gheorghe & Hart, 2022). Further work is needed to determine strengths and weaknesses of existing algorithms and further improve analysis even in the case of relatively straightforward screens with a single phenotypic output. While analysis of optical pooled screens comes with additional considerations, some advances in analysis of other pooled genetic screens such as copy-number correction could also be incorporated into analysis of optical pooled screens.

### *5.1.2 Analysis of Single-Cell Screens*

While the number of single-cell genetic perturbation-based datasets is relatively small, these are rapidly growing and now include one instance of a genome-scale single-cell transcriptome-wide dataset (Replogle et al., 2022). There are a few more examples of single-cell datasets with both genetic and non-genetic perturbations, resulting in 44 publicly available datasets recently collected in the scPerturb database (Peidli et al., 2022). Desired goals for analysis of these high-dimensional screens vary and generally focus on 1) predicting mechanisms of action, 2) predicting effects of unseen perturbations, or 3) predicting perturbation interactions (Ji et al., 2021). As single-cell perturbation screens become more routine, algorithms to leverage information in these high-resolution single-cell datasets are urgently needed as current pipelines have very different approaches and often do not present comparisons between distinct analysis approaches (A. Dixit et al., 2016; Duan et al., 2019; Papalexi et al., 2021; L. Wang, 2021; L. Yang et al., 2020).

### *5.1.3 Analysis of Image-Based Profiling Screens*

Unlike many other single-cell screens, image-based profiling screens encode spatial information about cell locations that could be used as either a phenotype of interest or to remove noise from a primary phenotype with no spatial variation of interest above the subcellular level. Examples of these spatial features include cell density on multiple scales and phenotypes of neighboring cells (example analysis approaches in Kramer et al., 2022; Snijder et al., 2012), as well as position within the well to account for technical within-well effects of uneven stimuli, enzymatic reactions, or phenotyping reagents, among other factors. In addition, technical effects related to image acquisition like the image acquisition order, position of cells within the field of view (Peng et al., 2017), and focus quality could also be taken into account in analyses.

## 5.2 Challenges and Opportunities in Screening for Host-Pathogen Interactions

### *5.2.1 Reproducibility and robustness*

A critical challenge for CRISPR screening in general and especially in the context of screening for host-pathogen interactions is the lack of reproducibility of screening results. Biological replicates are often not performed, analysis is not standardized, and screen sequencing results or even tables with scores for all perturbations are not always available, making it difficult to determine whether the cause of discordant results is technical or biological. At minimum, raw sequencing data for screens with an NGS-based output should be deposited to GEO or SRA. The research community should also seek to establish minimum information guidelines for CRISPR screens, as suggested by Bock et al., 2022.

### *5.2.2 Selection of Experimental System*

The primary genome-wide screens presented in this thesis relied on cancer cell lines or immortalized myeloid cell lines, since high-throughput screening in primary or more physiologically relevant cellular models is often intractable due to 1) difficulty obtaining a large enough cell population, 2) difficulty transducing sufficient cells, 3) difficulty maintaining cells in culture for multiple weeks while maintaining the phenotypic states of interest, and 4) challenges in assay compatibility with certain cell types (e.g. for optical pooled screens: *in situ* sequencing efficiency, surface area required, plating evenness; for flow cytometry screens: challenges generating single-cell suspensions; for all screens: challenge obtaining reliable high-penetrance phenotypes needed for large-scale screens). These challenges are not trivial and, even if some could be overcome, genes that have similar functions in multiple cell models may not be recovered in the less tractable cell model due to reduced power derived from technical effects.

Selection of the pathogen studied is also critical. In some cases, using a less-virulent pathogen such as a pseudotyped virus, bacterial component, viral replicon (self-replicating viral DNA or RNA) or minigenomes (plasmid-driven minimal replication systems) (Hannemann, 2020) may be more practical than use of a high-containment pathogen. Even in cases where use of the unmodified pathogen is available, pilot experiments should be performed to determine whether technical noise, often greater with the true pathogen, renders a large screen impractical. On the other hand, pilot experiments should also test whether known factors affecting host-pathogen interactions are replicated in simplified experimental systems and ensure use of experimental systems appropriate for the biological question being asked: for instance, pseudotyped viruses are typically best for identification of entry factors while minigenome assays are more suited for identification of post-entry replication modifiers.

As CRISPR technologies become more routine and sequencing costs decrease, it may become more common to perform screens in multiple experimental models in order to systematically compare effects across cell types and pathogens. In a meta-analysis of SARS-CoV-2 screens, the cellular model was identified as the primary source of variability (Rebendenne et al., 2022). Furthermore, gene hits in Calu3 lung epithelial cells, a more physiologically relevant cell model that does not require overexpression of ACE2 and TMPRSS2 receptors for SARS-CoV-2 infection, found that top hits in these cells were more highly expressed in COVID-19 patient epithelial cells, ACE2$^+$TMPRSS2$^+$ ciliated lung epithelial cells, and more differentially expressed in COVID-19 vs healthy epithelial cells than top hits from screens in less relevant cell types like Huh7 and A549 cells (Biering et al., 2022). This example highlights the importance of the selection of an appropriate cell model.

## 5.3 Optical Pooled Screens in Multicellular Ecosystems

### *5.3.1 In vitro Image-Based Screens in Multicellular Ecosystems*

In addition to extending optical pooled screens to more physiologically relevant cells as already discussed, screens for cell-cell interactions are also more challenging but of great interest. For instance, multiple genome-wide *in vitro* CRISPR screens in organoids have already been successful: a screen for regulators of iPSC differentiation in kidney organoids (Ungricht et al., 2022) and a screen that identified genes mediating resistance to TGF-β growth restriction in intestinal organoids (Ringel et al., 2020). Organoid models that could be adapted to two-dimensional systems compatible with high-throughput image-based screening or optically cleared to enable volumetric imaging could be used to screen for regulators of cellular interactions and spatial organization using optical pooled screening technology. Automated imaging and organoid culture will also be critical; for instance, a recent example imaged 300 organoids per hour (Beghin et al., 2022).

### *5.3.2 In vivo Spatial Screens*

Many phenotypes may not be properly recapitulated in any *in vitro* system and would therefore require *in vivo* screens. A key biological limitation here is that *in vivo* screens are generally performed in mouse models and the closest tractable model to humans for these screens would be the use of non-human primates (NHPs), which are often impractical. Therefore, if the phenomenon of interest is not recapitulated in mouse models, a profiling approach of human samples may be more appropriate than a perturbation-based approach as a first step.

*In vivo* pooled CRISPR screens are reviewed in Kuhn et al., 2021 and have been performed with fitness, flow cytometry, and scRNA-seq based phenotypic outputs. Spatial *in vivo* screens could be used to causally identify genetic determinants of *in vivo* multicellular organization. Four main approaches for *in vivo* spatial screening could be considered and the ideal approach in the near future will likely depend on the specific application. First, *in situ* sequencing approaches already used in tissue samples could be adapted to read out sgRNA sequences (X. Chen et al., 2019; Ke et al., 2013); however, these approaches are expensive and may not be compatible with tissues that have high autofluorescence. Second, *ex situ* sequencing spatial transcriptomics technologies (Slide-seq (Rodriques et al., 2019; Stickels et al., 2021) and HDST (Vickovic et al., 2019), among others) could be used to incorporate reads from sgRNAs, but costs would be substantial. Third, approaches to extract omics information from cell populations of interest could be used to perform enrichment-based spatial screens *in vivo* (Genshaft et al., 2021; K. H. Hu et al., 2020; Kishi et al., 2022; Mangiameli et al., 2022; van der Leun et al., 2021). Finally, since *in vivo* screens to assess perturbation-induced changes in cell positioning will likely involve small library sizes, these might be best encoded via protein barcodes like Pro-Codes (Wroblewska et al., 2018), which enabled detection of 364 populations, or EPICodes, enabling 46,656 combinations by using nuclear or mitochondrial protein localization to expand barcoding capacity (Kudo et al., 2022), a strategy that, however, might be challenging to apply to dense tissue sections. Pro-Codes have already been applied to tissue sections to identify 120 cancer cell populations and detect effects of knock-out of 35 genes in a mouse model of lung concer (Dhainaut et al., 2022; Wroblewska et al., 2018). These systems are attractive due to the stability and relative ease and low-cost detection of protein barcodes but, unlike the other strategies, would require construction of bespoke libraries.

## References

Ablasser, A., & Hur, S. (2020). Regulation of cGAS- and RLR-mediated immunity to nucleic acids. *Nature Immunology*, *21*(1), Article 1. https://doi.org/10.1038/s41590-019-0556-1

Agaisse, H., Burrack, L. S., Philips, J. A., Rubin, E. J., Perrimon, N., & Higgins, D. E. (2005). Genome-Wide RNAi Screen for Host Factors Required for Intracellular Bacterial Infection. *Science*, *309*(5738), 1248–1251. https://doi.org/10.1126/science.1116008

Anderson, D. E., Cui, J., Ye, Q., Huang, B., Tan, Y., Jiang, C., Zu, W., Gong, J., Liu, W., Kim, S. Y., Yan, B. G., Sigmundsson, K., Lim, X. F., Ye, F., Niu, P., Irving, A. T., Zhang, H., Tang, Y., Zhou, X., … Tan, X. (2021). Orthogonal genome-wide screens of bat cells identify MTHFD1 as a target of broad antiviral therapy. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(39), e2104759118. https://doi.org/10.1073/pnas.2104759118

Arroyo, J. D., Jourdain, A. A., Calvo, S. E., Ballarano, C. A., Doench, J. G., Root, D. E., & Mootha, V. K. (2016). A Genome-wide CRISPR Death Screen Identifies Genes Essential for Oxidative Phosphorylation. *Cell Metabolism*, *24*(6), 875–885. https://doi.org/10.1016/j.cmet.2016.08.017

Avraham, R., Haseley, N., Brown, D., Penaranda, C., Jijon, H. B., Trombetta, J. J., Satija, R., Shalek, A. K., Xavier, R. J., Regev, A., & Hung, D. T. (2015). Pathogen

Cell-to-Cell Variability Drives Heterogeneity in Host Immune Responses. *Cell*, *162*(6), 1309–1321. https://doi.org/10.1016/j.cell.2015.08.027

Baggen, J., Persoons, L., Vanstreels, E., Jansen, S., Van Looveren, D., Boeckx, B., Geudens, V., De Man, J., Jochmans, D., Wauters, J., Wauters, E., Vanaudenaerde, B. M., Lambrechts, D., Neyts, J., Dallmeier, K., Thibaut, H. J., Jacquemyn, M., Maes, P., & Daelemans, D. (2021). Genome-wide CRISPR screening identifies TMEM106B as a proviral host factor for SARS-CoV-2. *Nature Genetics*, *53*(4), 435–444. https://doi.org/10.1038/s41588-021-00805-2

Baggen, J., Vanstreels, E., Jansen, S., & Daelemans, D. (2021). Cellular host factors for SARS-CoV-2 infection. *Nature Microbiology*, *6*(10), Article 10. https://doi.org/10.1038/s41564-021-00958-0

Barczak, A. K., Avraham, R., Singh, S., Luo, S. S., Zhang, W. R., Bray, M.-A., Hinman, A. E., Thompson, M., Nietupski, R. M., Golas, A., Montgomery, P., Fitzgerald, M., Smith, R. S., White, D. W., Tischler, A. D., Carpenter, A. E., & Hung, D. T. (2017). Systematic, multiparametric analysis of Mycobacterium tuberculosis intracellular infection offers insight into coordinated virulence. *PLOS Pathogens*, *13*(5), e1006363. https://doi.org/10.1371/journal.ppat.1006363

Baril, M., Es-Saad, S., Chatel-Chaix, L., Fink, K., Pham, T., Raymond, V.-A., Audette, K., Guenier, A.-S., Duchaine, J., Servant, M., Bilodeau, M., Cohen, É., Grandvaux, N., & Lamarre, D. (2013). Genome-wide RNAi Screen Reveals a New Role of a WNT/CTNNB1 Signaling Pathway as Negative Regulator of Virus-induced Innate Immune Responses. *PLOS Pathogens*, *9*(6), e1003416. https://doi.org/10.1371/journal.ppat.1003416

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., Wilson, C. J., Lehár, J., Kryukov, G. V., Sonkin, D., Reddy, A., Liu, M., Murray, L., Berger, M. F., Monahan, J. E., Morais, P., Meltzer, J., Korejwa, A., Jané-Valbuena, J., … Garraway, L. A. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*, *483*(7391), Article 7391. https://doi.org/10.1038/nature11003

Batra, J., Hultquist, J. F., Liu, D., Shtanko, O., Von Dollen, J., Satkamp, L., Jang, G. M., Luthra, P., Schwarz, T. M., Small, G. I., Arnett, E., Anantpadma, M., Reyes, A., Leung, D. W., Kaake, R., Haas, P., Schmidt, C. B., Schlesinger, L. S., LaCount, D. J., … Krogan, N. J. (2018). Protein Interaction Mapping Identifies RBBP6 as a Negative Regulator of Ebola Virus Replication. *Cell*, *175*(7), 1917-1930.e13. https://doi.org/10.1016/j.cell.2018.08.044

Bedsaul, J. R., Zaritsky, L. A., & Zoon, K. C. (2016). Type I Interferon-Mediated Induction of Antiviral Genes and Proteins Fails to Protect Cells from the Cytopathic Effects of Sendai Virus Infection. *Journal of Interferon & Cytokine Research*, *36*(11), 652–665. https://doi.org/10.1089/jir.2016.0051

Beghin, A., Grenci, G., Sahni, G., Guo, S., Rajendiran, H., Delaire, T., Mohamad Raffi, S. B., Blanc, D., de Mets, R., Ong, H. T., Galindo, X., Monet, A., Acharya, V., Racine, V., Levet, F., Galland, R., Sibarita, J.-B., & Viasnoff, V. (2022).

Automated high-speed 3D imaging of organoid cultures with multi-scale phenotypic quantification. *Nature Methods*, *19*(7), Article 7. https://doi.org/10.1038/s41592-022-01508-0

Beneke, T., Demay, F., Hookway, E., Ashman, N., Jeffery, H., Smith, J., Valli, J., Becvar, T., Myskova, J., Lestinova, T., Shafiq, S., Sadlova, J., Volf, P., Wheeler, R. J., & Gluenz, E. (2019). Genetic dissection of a Leishmania flagellar proteome demonstrates requirement for directional motility in sand fly infections. *PLOS Pathogens*, *15*(6), e1007828. https://doi.org/10.1371/journal.ppat.1007828

Beneke, T., Madden, R., Makin, L., Valli, J., Sunter, J., & Gluenz, E. (2017). A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. *Royal Society Open Science*, *4*(5), 170095. https://doi.org/10.1098/rsos.170095

Bercovich-Kinori, A., Tai, J., Gelbart, I. A., Shitrit, A., Ben-Moshe, S., Drori, Y., Itzkovitz, S., Mandelboim, M., & Stern-Ginossar, N. (2016). A systematic view on influenza induced host shutoff. *ELife*, *5*, e18311. https://doi.org/10.7554/eLife.18311

Biering, S. B., Sarnik, S. A., Wang, E., Zengel, J. R., Leist, S. R., Schäfer, A., Sathyan, V., Hawkins, P., Okuda, K., Tau, C., Jangid, A. R., Duffy, C. V., Wei, J., Gilmore, R. C., Alfajaro, M. M., Strine, M. S., Nguyenla, X., Van Dis, E., Catamura, C., … Hsu, P. D. (2022). Genome-wide bidirectional CRISPR screens identify mucins as host factors modulating SARS-CoV-2 infection. *Nature Genetics*, *54*(8), Article 8. https://doi.org/10.1038/s41588-022-01131-x

Blondel, C. J., Park, J. S., Hubbard, T. P., Pacheco, A. R., Kuehl, C. J., Walsh, M. J., Davis, B. M., Gewurz, B. E., Doench, J. G., & Waldor, M. K. (2016). CRISPR/Cas9 Screens Reveal Requirements for Host Cell Sulfation and Fucosylation in Bacterial Type III Secretion System-Mediated Cytotoxicity. *Cell Host & Microbe*, *20*(2), 226–237. https://doi.org/10.1016/j.chom.2016.06.010

Bock, C., Datlinger, P., Chardon, F., Coelho, M. A., Dong, M. B., Lawson, K. A., Lu, T., Maroc, L., Norman, T. M., Song, B., Stanley, G., Chen, S., Garnett, M., Li, W., Moffat, J., Qi, L. S., Shapiro, R. S., Shendure, J., Weissman, J. S., & Zhuang, X. (2022). High-content CRISPR screening. *Nature Reviews Methods Primers*, *2*(1), Article 1. https://doi.org/10.1038/s43586-021-00093-4

Börold, J., Eletto, D., Busnadiego, I., Mair, N. K., Moritz, E., Schiefer, S., Schmidt, N., Petric, P. P., Wong, W. W.-L., Schwemmle, M., & Hale, B. G. (2021). BRD9 is a druggable component of interferon-stimulated gene expression and antiviral activity. *EMBO Reports*, *22*(10), e52823. https://doi.org/10.15252/embr.202152823

Bost, P., Giladi, A., Liu, Y., Bendjelal, Y., Xu, G., David, E., Blecher-Gonen, R., Cohen, M., Medaglia, C., Li, H., Deczkowska, A., Zhang, S., Schwikowski, B., Zhang, Z., & Amit, I. (2020). Host-Viral Infection Maps Reveal Signatures of Severe COVID-19 Patients. *Cell*, *181*(7), 1475-1488.e12. https://doi.org/10.1016/j.cell.2020.05.006

Brandt, J., Wendt, L., Bodmer, B. S., Mettenleiter, T. C., & Hoenen, T. (2020). The Cellular Protein CAD is Recruited into Ebola Virus Inclusion Bodies by the Nucleoprotein NP to Facilitate Genome Replication and Transcription. *Cells*, *9*(5), 1126. https://doi.org/10.3390/cells9051126

Bray, M.-A., Singh, S., Han, H., Davis, C. T., Borgeson, B., Hartland, C., Kost-Alimova, M., Gustafsdottir, S. M., Gibson, C. C., & Carpenter, A. E. (2016). Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nature Protocols*, *11*(9), 1757–1774. https://doi.org/10.1038/nprot.2016.105

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525–527. https://doi.org/10.1038/nbt.3519

Bruchez, A., Sha, K., Johnson, J., Chen, L., Stefani, C., McConnell, H., Gaucherand, L., Prins, R., Matreyek, K. A., Hume, A. J., Mühlberger, E., Schmidt, E. V., Olinger, G. G., Stuart, L. M., & Lacy-Hulbert, A. (2020). MHC class II transactivator CIITA induces cell resistance to Ebola virus and SARS-like coronaviruses. *Science*, *370*(6513), 241–247. https://doi.org/10.1126/science.abb3753

Bryant, J. M., Baumgarten, S., Glover, L., Hutchinson, S., & Rachidi, N. (2019). CRISPR in Parasitology: Not Exactly Cut and Dried! *Trends in Parasitology*, *35*(6), 409–422. https://doi.org/10.1016/j.pt.2019.03.004

Buers, I., Nitschke, Y., & Rutsch, F. (2016). Novel interferonopathies associated with mutations in RIG-I like receptors. *Cytokine & Growth Factor Reviews*, *29*, 101–107. https://doi.org/10.1016/j.cytogfr.2016.03.005

Burdette, D. L., Monroe, K. M., Sotelo-Troha, K., Iwig, J. S., Eckert, B., Hyodo, M., Hayakawa, Y., & Vance, R. E. (2011). STING is a direct innate immune sensor of cyclic di-GMP. *Nature*, *478*(7370), Article 7370. https://doi.org/10.1038/nature10429

Camsund, D., Lawson, M. J., Larsson, J., Jones, D., Zikrin, S., Fange, D., & Elf, J. (2020). Time-resolved imaging-based CRISPRi screening. *Nature Methods*, *17*(1), Article 1. https://doi.org/10.1038/s41592-019-0629-y

Carette, J. E., Raaben, M., Wong, A. C., Herbert, A. S., Obernosterer, G., Mulherkar, N., Kuehne, A. I., Kranzusch, P. J., Griffin, A. M., Ruthel, G., Dal Cin, P., Dye, J. M., Whelan, S. P., Chandran, K., & Brummelkamp, T. R. (2011). Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature*, *477*(7364), 340–343. https://doi.org/10.1038/nature10348

Carlson, R. J., Leiken, M. D., Guna, A., Hacohen, N., & Blainey, P. C. (2023). A genome-wide optical pooled screen reveals regulators of cellular antiviral responses. *Proceedings of the National Academy of Sciences*, *120*(16), e2210623120. https://doi.org/10.1073/pnas.2210623120

Chan, Y. K., & Gack, M. U. (2015). RIG-I-like receptor regulation in virus infection and immunity. *Current Opinion in Virology*, *12*, 7–14. https://doi.org/10.1016/j.coviro.2015.01.004

Chen, J., & Chen, Z. J. (2018). PtdIns4P on dispersed trans-Golgi network mediates NLRP3 inflammasome activation. *Nature*, *564*(7734), Article 7734. https://doi.org/10.1038/s41586-018-0761-3

Chen, X., Sun, Y.-C., Zhan, H., Kebschull, J. M., Fischer, S., Matho, K., Huang, Z. J., Gillis, J., & Zador, A. M. (2019). High-Throughput Mapping of Long-Range Neuronal Projection Using In Situ Sequencing. *Cell*, *179*(3), 772-786.e19. https://doi.org/10.1016/j.cell.2019.09.023

Cheng, H., Koning, K., O'Hearn, A., Wang, M., Rumschlag-Booms, E., Varhegyi, E., & Rong, L. (2015). A parallel genome-wide RNAi screening strategy to identify host proteins important for entry of Marburg virus and H5N1 influenza virus. *Virology Journal*, *12*, 194. https://doi.org/10.1186/s12985-015-0420-3

Chiang, J. J., Sparrer, K. M. J., van Gent, M., Lässig, C., Huang, T., Osterrieder, N., Hopfner, K.-P., & Gack, M. U. (2018). Viral unmasking of cellular 5S rRNA pseudogene transcripts induces RIG-I mediated immunity. *Nature Immunology*, *19*(1), 53–62. https://doi.org/10.1038/s41590-017-0005-y

Cho, S., Choe, D., Lee, E., Kim, S. C., Palsson, B., & Cho, B.-K. (2018). High-Level dCas9 Expression Induces Abnormal Cell Morphology in Escherichia coli. *ACS Synthetic Biology*, *7*(4), 1085–1094. https://doi.org/10.1021/acssynbio.7b00462

Choi, H. M. T., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A., & Pierce, N. A. (2018). Third-generation in situ hybridization chain reaction: Multiplexed, quantitative, sensitive, versatile, robust. *Development*, *145*(12), dev165753. https://doi.org/10.1242/dev.165753

Chollet, F. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. *ArXiv:1610.02357 [Cs]*. http://arxiv.org/abs/1610.02357

Chow, K. T., Gale, M., & Loo, Y.-M. (2018). RIG-I and Other RNA Sensors in Antiviral Immunity. *Annual Review of Immunology*, *36*(1), 667–694. https://doi.org/10.1146/annurev-immunol-042617-053309

Cohen, Y., Megyeri, M., Chen, O. C. W., Condomitti, G., Riezman, I., Loizides-Mangold, U., Abdul-Sada, A., Rimon, N., Riezman, H., Platt, F. M., Futerman, A. H., & Schuldiner, M. (2013). The Yeast P5 Type ATPase, Spf1, Regulates Manganese Transport into the Endoplasmic Reticulum. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0085519

Crow, Y. J., & Manel, N. (2015). Aicardi–Goutières syndrome and the type I interferonopathies. *Nature Reviews Immunology*, *15*(7), Article 7. https://doi.org/10.1038/nri3850

Cui, L., Vigouroux, A., Rousset, F., Varet, H., Khanna, V., & Bikard, D. (2018). A CRISPRi screen in E. coli reveals sequence-specific toxicity of dCas9. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-04209-5

Daniloski, Z., Jordan, T. X., Wessels, H.-H., Hoagland, D. A., Kasela, S., Legut, M., Maniatis, S., Mimitou, E. P., Lu, L., Geller, E., Danziger, O., Rosenberg, B. R., Phatnani, H., Smibert, P., Lappalainen, T., tenOever, B. R., & Sanjana, N. E. (2021). Identification of Required Host Factors for SARS-CoV-2 Infection in Human Cells. *Cell*, *184*(1), 92-105.e16. https://doi.org/10.1016/j.cell.2020.10.030

Danziger, O., Patel, R. S., DeGrace, E. J., Rosen, M. R., & Rosenberg, B. R. (2022). Inducible CRISPR activation screen for interferon-stimulated genes identifies OAS1 as a SARS-CoV-2 restriction factor. *PLOS Pathogens*, *18*(4), e1010464. https://doi.org/10.1371/journal.ppat.1010464

Deffieu, M. S., Cesonyte, I., Delalande, F., Boncompain, G., Dorobantu, C., Song, E., Lucansky, V., Hirschler, A., Cianferani, S., Perez, F., Carapito, C., & Gaudin, R. (2021). Rab7-harboring vesicles are carriers of the transferrin receptor through the biosynthetic secretory pathway. *Science Advances*, *7*(2), eaba7803. https://doi.org/10.1126/sciadv.aba7803

Dempster, J. M., Boyle, I., Vazquez, F., Root, D. E., Boehm, J. S., Hahn, W. C., Tsherniak, A., & McFarland, J. M. (2021). Chronos: A cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects. *Genome Biology*, *22*(1), 343. https://doi.org/10.1186/s13059-021-02540-7

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Desmyter, J., Melnick, J. L., & Rawls, W. E. (1968). Defectiveness of Interferon Production and of Rubella Virus Interference in a Line of African Green Monkey Kidney Cells (Vero). *Journal of Virology*, *2*(10), 955–961.

Dezi, M., Di Cicco, A., Bassereau, P., & Lévy, D. (2013). Detergent-mediated incorporation of transmembrane proteins in giant unilamellar vesicles with controlled physiological contents. *Proceedings of the National Academy of Sciences*, *110*(18), 7276–7281. https://doi.org/10.1073/pnas.1303857110

Dhainaut, M., Rose, S. A., Akturk, G., Wroblewska, A., Nielsen, S. R., Park, E. S., Buckup, M., Roudko, V., Pia, L., Sweeney, R., Le Berichel, J., Wilk, C. M., Bektesevic, A., Lee, B. H., Bhardwaj, N., Rahman, A. H., Baccarini, A., Gnjatic, S., Pe'er, D., … Brown, B. D. (2022). Spatial CRISPR genomics identifies regulators of the tumor microenvironment. *Cell*, *185*(7), 1223-1239.e20. https://doi.org/10.1016/j.cell.2022.02.015

Ding, S., Diep, J., Feng, N., Ren, L., Li, B., Ooi, Y. S., Wang, X., Brulois, K. F., Yasukawa, L. L., Li, X., Kuo, C. J., Solomon, D. A., Carette, J. E., & Greenberg, H. B. (2018). STAG2 deficiency induces interferon responses via cGAS-STING pathway and restricts virus infection. *Nature Communications*, *9*(1), 1485. https://doi.org/10.1038/s41467-018-03782-z

Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C. P., Jerby-Arnon, L., Marjanovic, N. D., Dionne, D., Burks, T., Raychowdhury, R., Adamson, B., Norman, T. M., Lander, E. S., Weissman, J. S., Friedman, N., & Regev, A. (2016). Perturb-Seq:

Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell*, *167*(7), 1853-1866.e17. https://doi.org/10.1016/j.cell.2016.11.038

Dixit, E., Boulant, S., Zhang, Y., Lee, A. S. Y., Odendall, C., Shum, B., Hacohen, N., Chen, Z. J., Whelan, S. P., Fransen, M., Nibert, M. L., Superti-Furga, G., & Kagan, J. C. (2010). Peroxisomes are signaling platforms for antiviral innate immunity. *Cell*, *141*(4), 668–681. https://doi.org/10.1016/j.cell.2010.04.018

Drayman, N., Patel, P., Vistain, L., & Tay, S. (2019). HSV-1 single-cell analysis reveals the activation of anti-viral and developmental programs in distinct sub-populations. *ELife*, *8*, e46339. https://doi.org/10.7554/eLife.46339

Duan, B., Zhou, C., Zhu, C., Yu, Y., Li, G., Zhang, S., Zhang, C., Ye, X., Ma, H., Qu, S., Zhang, Z., Wang, P., Sun, S., & Liu, Q. (2019). Model-based understanding of single-cell CRISPR screening. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-10216-x

Durmuş, S., & Ülgen, K. Ö. (2017). Comparative interactomics for virus–human protein–protein interactions: DNA viruses versus RNA viruses. *FEBS Open Bio*, *7*(1), 96–107. https://doi.org/10.1002/2211-5463.12167

Eng, K. E., Panas, M. D., Hedestam, G. B. K., & McInerney, G. M. (2010). A novel quantitative flow cytometry-based assay for autophagy. *Autophagy*, *6*(5), 634–641. https://doi.org/10.4161/auto.6.5.12112

Evans, M. J., von Hahn, T., Tscherne, D. M., Syder, A. J., Panis, M., Wölk, B., Hatziioannou, T., McKeating, J. A., Bieniasz, P. D., & Rice, C. M. (2007). Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature*, *446*(7137), 801–805. https://doi.org/10.1038/nature05654

Fang, J., Pietzsch, C., Tsaprailis, G., Crynen, G., Cho, K. F., Ting, A. Y., Bukreyev, A., de la Torre, J. C., & Saphire, E. O. (2022). Functional interactomes of the Ebola virus polymerase identified by proximity proteomics in the context of viral replication. *Cell Reports*, *38*(12), 110544. https://doi.org/10.1016/j.celrep.2022.110544

Fang, Z., Liu, X., & Peltz, G. (2023). GSEApy: A comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics*, *39*(1), btac757. https://doi.org/10.1093/bioinformatics/btac757

Feldman, D., Funk, L., Le, A., Carlson, R. J., Leiken, M. D., Tsai, F., Soong, B., Singh, A., & Blainey, P. C. (2022). Pooled genetic perturbation screens with image-based phenotypes. *Nature Protocols*, *17*(2), 476–512. https://doi.org/10.1038/s41596-021-00653-8

Feldman, D., Singh, A., Schmid-Burgk, J. L., Carlson, R. J., Mezger, A., Garrity, A. J., Zhang, F., & Blainey, P. C. (2019). Optical Pooled Screens in Human Cells. *Cell*, *179*(3), 787-799.e17. https://doi.org/10.1016/j.cell.2019.09.016

Filone, C. M., Dower, K., Cowley, G. S., Hensley, L. E., & Connor, J. H. (2015). Probing the virus host interaction in high containment: An approach using pooled short hairpin RNA. *Assay and Drug Development Technologies*, *13*(1), 34–43. https://doi.org/10.1089/adt.2014.613

Fischer, T. D., Wang, C., Padman, B. S., Lazarou, M., & Youle, R. J. (2020). STING induces LC3B lipidation onto single-membrane vesicles via the V-ATPase and ATG16L1-WD40 domain. *The Journal of Cell Biology*, *219*(12), e202009128. https://doi.org/10.1083/jcb.202009128

Fitzgerald, K. A., & Kagan, J. C. (2020). Toll-like Receptors and the Control of Immunity. *Cell*, *180*(6), 1044–1066. https://doi.org/10.1016/j.cell.2020.02.041

Flint, M., Chatterjee, P., Lin, D. L., McMullan, L. K., Shrivastava-Ranjan, P., Bergeron, É., Lo, M. K., Welch, S. R., Nichol, S. T., Tai, A. W., & Spiropoulou, C. F. (2019). A genome-wide CRISPR screen identifies N-acetylglucosamine-1-phosphate transferase as a potential antiviral target for Ebola virus. *Nature Communications*, *10*, 285. https://doi.org/10.1038/s41467-018-08135-4

Fullam, A., & Schröder, M. (2013). DExD/H-box RNA helicases as mediators of anti-viral innate immunity and essential host factors for viral replication. *Biochimica et Biophysica Acta. Gene Regulatory Mechanisms*, *1829*(8), 854–865. https://doi.org/10.1016/j.bbagrm.2013.03.012

Funk, L., Su, K.-C., Ly, J., Feldman, D., Singh, A., Moodie, B., Blainey, P. C., & Cheeseman, I. M. (2022). The phenotypic landscape of essential human genes. *Cell*, *0*(0). https://doi.org/10.1016/j.cell.2022.10.017

Gabaev, I., Williamson, J. C., Crozier, T. W. M., Schulz, T. F., & Lehner, P. J. (2020). Quantitative Proteomics Analysis of Lytic KSHV Infection in Human Endothelial Cells Reveals Targets of Viral Immune Modulation. *Cell Reports*, *33*(2), 108249. https://doi.org/10.1016/j.celrep.2020.108249

Gaidt, M. M., Ebert, T. S., Chauhan, D., Ramshorn, K., Pinci, F., Zuber, S., O'Duill, F., Schmid-Burgk, J. L., Hoss, F., Buhmann, R., Wittmann, G., Latz, E., Subklewe, M., & Hornung, V. (2017). The DNA Inflammasome in Human Myeloid Cells Is Initiated by a STING-Cell Death Program Upstream of NLRP3. *Cell*, *171*(5), 1110-1124.e18. https://doi.org/10.1016/j.cell.2017.09.039

Galão, R. P., Wilson, H., Schierhorn, K. L., Debeljak, F., Bodmer, B. S., Goldhill, D., Hoenen, T., Wilson, S. J., Swanson, C. M., & Neil, S. J. D. (2022). TRIM25 and ZAP target the Ebola virus ribonucleoprotein complex to mediate interferon-induced restriction. *PLOS Pathogens*, *18*(5), e1010530. https://doi.org/10.1371/journal.ppat.1010530

García-Dorival, I., Wu, W., Armstrong, S. D., Barr, J. N., Carroll, M. W., Hewson, R., & Hiscox, J. A. (2016). Elucidation of the Cellular Interactome of Ebola Virus Nucleoprotein and Identification of Therapeutic Targets. *Journal of Proteome Research*, *15*(12), 4290–4303. https://doi.org/10.1021/acs.jproteome.6b00337

Gaudet, R. G., Zhu, S., Halder, A., Kim, B.-H., Bradfield, C. J., Huang, S., Xu, D., Mamiñska, A., Nguyen, T. N., Lazarou, M., Karatekin, E., Gupta, K., & MacMicking, J. D. (2021). A human apolipoprotein L with detergent-like activity kills intracellular pathogens. *Science*, *373*(6552), eabf8113. https://doi.org/10.1126/science.abf8113

Genshaft, A. S., Ziegler, C. G. K., Tzouanas, C. N., Mead, B. E., Jaeger, A. M., Navia, A. W., King, R. P., Mana, M. D., Huang, S., Mitsialis, V., Snapper, S. B., Yilmaz, Ö. H., Jacks, T., Van Humbeck, J. F., & Shalek, A. K. (2021). Live cell tagging tracking and isolation for spatial transcriptomics using photoactivatable cell dyes. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-25279-y

Gentili, M., Liu, B., Papanastasiou, M., Dele-Oni, D., Schwartz, M. A., Carlson, R. J., Al'Khafaji, A. M., Krug, K., Brown, A., Doench, J. G., Carr, S. A., & Hacohen, N. (2023). ESCRT-dependent STING degradation inhibits steady-state and cGAMP-induced signalling. *Nature Communications*, *14*(1), Article 1. https://doi.org/10.1038/s41467-023-36132-9

Gheorghe, V., & Hart, T. (2022). Optimal construction of a functional interaction network from pooled library CRISPR fitness screens. *BMC Bioinformatics*, *23*(1), 510. https://doi.org/10.1186/s12859-022-05078-y

Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H., Guimaraes, C., Panning, B., Ploegh, H. L., Bassik, M. C., Qi, L. S., Kampmann, M., & Weissman, J. S. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, *159*(3), 647–661. https://doi.org/10.1016/j.cell.2014.09.029

Gui, X., Yang, H., Li, T., Tan, X., Shi, P., Li, M., Du, F., & Chen, Z. J. (2019). Autophagy induction via STING trafficking is a primordial function of the cGAS pathway. *Nature*, *567*(7747), Article 7747. https://doi.org/10.1038/s41586-019-1006-9

Guo, R., Jiang, C., Zhang, Y., Govande, A., Trudeau, S. J., Chen, F., Fry, C. J., Puri, R., Wolinsky, E., Schineller, M., Frost, T. C., Gebre, M., Zhao, B., Giulino-Roth, L., Doench, J. G., Teng, M., & Gewurz, B. E. (2020). MYC Controls the Epstein-Barr Virus Lytic Switch. *Molecular Cell*, *78*(4), 653-669.e8. https://doi.org/10.1016/j.molcel.2020.03.025

Handschin, C., Lin, J., Rhee, J., Peyer, A.-K., Chin, S., Wu, P.-H., Meyer, U. A., & Spiegelman, B. M. (2005). Nutritional regulation of hepatic heme biosynthesis and porphyria through PGC-1alpha. *Cell*, *122*(4), 505–515. https://doi.org/10.1016/j.cell.2005.06.040

Hannemann, H. (2020). Viral replicons as valuable tools for drug discovery. *Drug Discovery Today*, *25*(6), 1026–1033. https://doi.org/10.1016/j.drudis.2020.03.010

Harding, A. T., Goff, M. A., Froggatt, H. M., Lim, J. K., & Heaton, N. S. (2021). GPER1 is required to protect fetal health from maternal inflammation. *Science*, *371*(6526), 271–276. https://doi.org/10.1126/science.aba9001

Hart, T., Chandrashekhar, M., Aregger, M., Steinhart, Z., Brown, K. R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S., Mero, P., Dirks, P., Sidhu, S., Roth, F. P., Rissland, O. S., Durocher, D., Angers, S., & Moffat, J. (2015). High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-

Specific Cancer Liabilities. *Cell*, *163*(6), 1515–1526. https://doi.org/10.1016/j.cell.2015.11.015

Hasan, M., Koch, J., Rakheja, D., Pattnaik, A. K., Brugarolas, J., Dozmorov, I., Levine, B., Wakeland, E. K., Lee-Kirsch, M. A., & Yan, N. (2013). Trex1 regulates lysosomal biogenesis and interferon-independent activation of antiviral genes. *Nature Immunology*, *14*(1), 61–71. https://doi.org/10.1038/ni.2475

Hawer, H., Hammermeister, A., Ravichandran, K. E., Glatt, S., Schaffrath, R., & Klassen, R. (2018). Roles of Elongator Dependent tRNA Modification Pathways in Neurodegeneration and Cancer. *Genes*, *10*(1), 19. https://doi.org/10.3390/genes10010019

Hawkins, J. S., Silvis, M. R., Koo, B.-M., Peters, J. M., Osadnik, H., Jost, M., Hearne, C. C., Weissman, J. S., Todor, H., & Gross, C. A. (2020). Mismatch-CRISPRi Reveals the Co-varying Expression-Fitness Relationships of Essential Genes in Escherichia coli and Bacillus subtilis. *Cell Systems*, *11*(5), 523-535.e9. https://doi.org/10.1016/j.cels.2020.09.009

Heaton, B. E., Kennedy, E. M., Dumm, R. E., Harding, A. T., Sacco, M. T., Sachs, D., & Heaton, N. S. (2017). A CRISPR Activation Screen Identifies a Pan-avian Influenza Virus Inhibitory Host Factor. *Cell Reports*, *20*(7), 1503–1512. https://doi.org/10.1016/j.celrep.2017.07.060

Hein, M. Y., & Weissman, J. S. (2022). Functional single-cell genomics of human cytomegalovirus infection. *Nature Biotechnology*, *40*(3), Article 3. https://doi.org/10.1038/s41587-021-01059-3

Heiser, K., McLean, P. F., Davis, C. T., Fogelson, B., Gordon, H. B., Jacobson, P., Hurst, B., Miller, B., Alfa, R. W., Earnshaw, B. A., Victors, M. L., Chong, Y. T., Haque, I. S., Low, A. S., & Gibson, C. C. (2020). Identification of potential treatments for COVID-19 through artificial intelligence-enabled phenomic analysis of human cells infected with SARS-CoV-2. *BioRxiv*, 2020.04.21.054387. https://doi.org/10.1101/2020.04.21.054387

Heppner, D. G., Kemp, T. L., Martin, B. K., Ramsey, W. J., Nichols, R., Dasen, E. J., Link, C. J., Das, R., Xu, Z. J., Sheldon, E. A., Nowak, T. A., Monath, T. P., Heppner, D. G., Kemp, T. L., Martin, B. K., Ramsey, W. J., Nichols, R., Dasen, E. J., Fusco, J., … Adams, M. (2017). Safety and immunogenicity of the rVSVΔG-ZEBOV-GP Ebola virus vaccine candidate in healthy adults: A phase 1b randomised, multicentre, double-blind, placebo-controlled, dose-response study. *The Lancet Infectious Diseases*, *17*(8), 854–866. https://doi.org/10.1016/S1473-3099(17)30313-4

Hoenen, T., Groseth, A., & Feldmann, H. (2019). Therapeutic strategies to target the Ebola virus life cycle. *Nature Reviews Microbiology*, *17*(10), Article 10. https://doi.org/10.1038/s41579-019-0233-2

Hoffmann, H.-H., Schneider, W. M., Blomen, V. A., Scull, M. A., Hovnanian, A., Brummelkamp, T. R., & Rice, C. M. (2017). Diverse Viruses Require the Calcium Transporter SPCA1 for Maturation and Spread. *Cell Host & Microbe*, *22*(4), 460-470.e5. https://doi.org/10.1016/j.chom.2017.09.002

Hölzer, M., Krähling, V., Amman, F., Barth, E., Bernhart, S. H., Carmelo, V. A. O., Collatz, M., Doose, G., Eggenhofer, F., Ewald, J., Fallmann, J., Feldhahn, L. M., Fricke, M., Gebauer, J., Gruber, A. J., Hufsky, F., Indrischek, H., Kanton, S., Linde, J., … Marz, M. (2016). Differential transcriptional responses to Ebola and Marburg virus infection in bat and human cells. *Scientific Reports*, *6*, 34589. https://doi.org/10.1038/srep34589

Hooper, K. M., Jacquin, E., Li, T., Goodwin, J. M., Brumell, J. H., Durgan, J., & Florey, O. (2022). V-ATPase is a universal regulator of LC3-associated phagocytosis and non-canonical autophagy. *The Journal of Cell Biology*, *221*(6), e202105112. https://doi.org/10.1083/jcb.202105112

Horlbeck, M. A., Xu, A., Wang, M., Bennett, N. K., Park, C. Y., Bogdanoff, D., Adamson, B., Chow, E. D., Kampmann, M., Peterson, T. R., Nakamura, K., Fischbach, M. A., Weissman, J. S., & Gilbert, L. A. (2018). Mapping the Genetic Landscape of Human Cells. *Cell*, *174*(4), 953-967.e22. https://doi.org/10.1016/j.cell.2018.06.010

Hu, K. H., Eichorst, J. P., McGinnis, C. S., Patterson, D. M., Chow, E. D., Kersten, K., Jameson, S. C., Gartner, Z. J., Rao, A. A., & Krummel, M. F. (2020). ZipSeq:

Barcoding for real-time mapping of single cell transcriptomes. *Nature Methods*, *17*(8), Article 8. https://doi.org/10.1038/s41592-020-0880-2

Hu, M.-M., & Shu, H.-B. (2020). Innate Immune Response to Cytoplasmic DNA: Mechanisms and Diseases. *Annual Review of Immunology*, *38*, 79–98. https://doi.org/10.1146/annurev-immunol-070119-115052

Hur, S. (2019). Double-Stranded RNA Sensors and Modulators in Innate Immunity. *Annual Review of Immunology*, *37*, 349–375. https://doi.org/10.1146/annurev-immunol-042718-041356

Hyun, S.-I., Maruri-Avidal, L., & Moss, B. (2015). Topology of Endoplasmic Reticulum-Associated Cellular and Viral Proteins Determined with Split-GFP. *Traffic (Copenhagen, Denmark)*, *16*(7), 787–795. https://doi.org/10.1111/tra.12281

Ichinohe, T., Pang, I. K., & Iwasaki, A. (2010). Influenza virus activates inflammasomes via its intracellular M2 ion channel. *Nature Immunology*, *11*(5), Article 5. https://doi.org/10.1038/ni.1861

Ilunga Kalenga, O., Moeti, M., Sparrow, A., Nguyen, V.-K., Lucey, D., & Ghebreyesus, T. A. (2019). The Ongoing Ebola Epidemic in the Democratic Republic of Congo, 2018-2019. *The New England Journal of Medicine*, *381*(4), 373–383. https://doi.org/10.1056/NEJMsr1904253

Inglis, A. J., Page, K. R., Guna, A., & Voorhees, R. M. (2020). Differential Modes of Orphan Subunit Recognition for the WRB/CAML Complex. *Cell Reports*, *30*(11), 3691-3698.e5. https://doi.org/10.1016/j.celrep.2020.02.084

Ishikawa, H., & Barber, G. N. (2008). STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature*, *455*(7213), Article 7213. https://doi.org/10.1038/nature07317

Jacquin, E., Leclerc-Mercier, S., Judon, C., Blanchard, E., Fraitag, S., & Florey, O. (2017). Pharmacological modulators of autophagy activate a parallel noncanonical pathway driving unconventional LC3 lipidation. *Autophagy*, *13*(5), 854–867. https://doi.org/10.1080/15548627.2017.1287653

Jeng, E. E., Bhadkamkar, V., Ibe, N. U., Gause, H., Jiang, L., Chan, J., Jian, R., Jimenez-Morales, D., Stevenson, E., Krogan, N. J., Swaney, D. L., Snyder, M. P., Mukherjee, S., & Bassik, M. C. (2019). Systematic Identification of Host Cell Regulators of Legionella pneumophila Pathogenesis Using a Genome-wide CRISPR Screen. *Cell Host & Microbe*, *26*(4), 551-563.e6. https://doi.org/10.1016/j.chom.2019.08.017

Ji, Y., Lotfollahi, M., Wolf, F. A., & Theis, F. J. (2021). Machine learning for perturbational single-cell omics. *Cell Systems*, *12*(6), 522–537. https://doi.org/10.1016/j.cels.2021.05.016

Jiang, X., Muthusamy, V., Fedorova, O., Kong, Y., Kim, D. J., Bosenberg, M., Pyle, A. M., & Iwasaki, A. (2019). Intratumoral delivery of RIG-I agonist SLR14 induces robust antitumor responses. *Journal of Experimental Medicine*, *216*(12), 2854–2868. https://doi.org/10.1084/jem.20190801

Jonikas, M. C., Collins, S. R., Denic, V., Oh, E., Quan, E. M., Schmid, V., Weibezahn, J., Schwappach, B., Walter, P., Weissman, J. S., & Schuldiner, M. (2009). Comprehensive Characterization of Genes Required for Protein Folding in the Endoplasmic Reticulum. *Science*, *323*(5922), 1693–1697. https://doi.org/10.1126/science.1167983

Jung, H. J., Shim, J. S., Lee, J., Song, Y. M., Park, K. C., Choi, S. H., Kim, N. D., Yoon, J. H., Mungai, P. T., Schumacker, P. T., & Kwon, H. J. (2010). Terpestacin inhibits tumor angiogenesis by targeting UQCRB of mitochondrial complex III and suppressing hypoxia-induced reactive oxygen species production and cellular oxygen sensing. *The Journal of Biological Chemistry*, *285*(15), 11584–11595. https://doi.org/10.1074/jbc.M109.087809

Karlas, A., Machuy, N., Shin, Y., Pleissner, K.-P., Artarini, A., Heuer, D., Becker, D., Khalil, H., Ogilvie, L. A., Hess, S., Mäurer, A. P., Müller, E., Wolff, T., Rudel, T., & Meyer, T. F. (2010). Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature*, *463*(7282), 818–822. https://doi.org/10.1038/nature08760

Ke, R., Mignardi, M., Pacureanu, A., Svedlund, J., Botling, J., Wählby, C., & Nilsson, M. (2013). In situ sequencing for RNA analysis in preserved tissue and cells. *Nature Methods*, *10*(9), 857–860. https://doi.org/10.1038/nmeth.2563

Kehl, A., Göser, V., Reuter, T., Liss, V., Franke, M., John, C., Richter, C. P., Deiwick, J., & Hensel, M. (2020). A trafficome-wide RNAi screen reveals deployment of early and late secretory host proteins and the entire late endo-/lysosomal vesicle fusion machinery by intracellular Salmonella. *PLoS Pathogens*, *16*(7), e1008220. https://doi.org/10.1371/journal.ppat.1008220

Kim, E., & Hart, T. (2021). Improved analysis of CRISPR fitness screens and reduced off-target effects with the BAGEL2 gene essentiality classifier. *Genome Medicine*, *13*(1), 2. https://doi.org/10.1186/s13073-020-00809-3

Kim, H. S., Lee, K., Kim, S.-J., Cho, S., Shin, H. J., Kim, C., & Kim, J.-S. (2018). Arrayed CRISPR screen with image-based assay reliably uncovers host genes required for coxsackievirus infection. *Genome Research*, *28*(6), 859–868. https://doi.org/10.1101/gr.230250.117

Kishi, J. Y., Liu, N., West, E. R., Sheng, K., Jordanides, J. J., Serrata, M., Cepko, C. L., Saka, S. K., & Yin, P. (2022). Light-Seq: Light-directed in situ barcoding of biomolecules in fixed cells and tissues for spatially indexed sequencing. *Nature Methods*, *19*(11), Article 11. https://doi.org/10.1038/s41592-022-01604-1

Köster, J., & Rahmann, S. (2012). Snakemake—A scalable bioinformatics workflow engine. *Bioinformatics*, *28*(19), 2520–2522. https://doi.org/10.1093/bioinformatics/bts480

Kotliar, D., Lin, A. E., Logue, J., Hughes, T. K., Khoury, N. M., Raju, S. S., Wadsworth, M. H., Chen, H., Kurtz, J. R., Dighero-Kemp, B., Bjornson, Z. B., Mukherjee, N., Sellers, B. A., Tran, N., Bauer, M. R., Adams, G. C., Adams, R., Rinn, J. L., Melé, M., … Bennett, R. S. (2020). Single-Cell Profiling of Ebola Virus Disease In Vivo

Reveals Viral and Host Dynamics. *Cell*, *183*(5), 1383-1401.e19. https://doi.org/10.1016/j.cell.2020.10.002

Kramer, B. A., Sarabia del Castillo, J., & Pelkmans, L. (2022). Multimodal perception links cellular state to decision-making in single cells. *Science*, *377*(6606), 642–648. https://doi.org/10.1126/science.abf4062

Kubota, T., Matsuoka, M., Chang, T.-H., Tailor, P., Sasaki, T., Tashiro, M., Kato, A., & Ozato, K. (2008). Virus Infection Triggers SUMOylation of IRF3 and IRF7, Leading to the Negative Regulation of Type I Interferon Gene Expression. *The Journal of Biological Chemistry*, *283*(37), 25660–25670. https://doi.org/10.1074/jbc.M804479200

Kudo, T., Lane, K., & Covert, M. W. (2022). A multiplexed epitope barcoding strategy that enables dynamic cellular phenotypic screens. *Cell Systems*, *13*(5), 376-387.e8. https://doi.org/10.1016/j.cels.2022.02.006

Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., McDermott, M. G., Monteiro, C. D., Gundersen, G. W., & Ma'ayan, A. (2016). Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*, *44*(W1), W90-97. https://doi.org/10.1093/nar/gkw377

Lai, Y., Babunovic, G. H., Cui, L., Dedon, P. C., Doench, J. G., Fortune, S. M., & Lu, T. K. (2020). Illuminating Host-Mycobacterial Interactions with Genome-wide

CRISPR Knockout and CRISPRi Screens. *Cell Systems*, *11*(3), 239-251.e7. https://doi.org/10.1016/j.cels.2020.08.010

Lai, Y., Cui, L., Babunovic, G. H., Fortune, S. M., Doench, J. G., & Lu, T. K. (2021). High-Throughput CRISPR Screens To Dissect Macrophage-Shigella Interactions. *MBio*, *12*(6), e02158-21. https://doi.org/10.1128/mBio.02158-21

Lawson, M. J., Camsund, D., Larsson, J., Baltekin, Ö., Fange, D., & Elf, J. (2017). In situ genotyping of a pooled strain library after characterizing complex phenotypes. *Molecular Systems Biology*, *13*(10), 947. https://doi.org/10.15252/msb.20177951

Le Vasseur, M., Friedman, J., Jost, M., Xu, J., Yamada, J., Kampmann, M., Horlbeck, M. A., Salemi, M. R., Phinney, B. S., Weissman, J. S., & Nunnari, J. (2021). Genome-wide CRISPRi screening identifies OCIAD1 as a prohibitin client and regulatory determinant of mitochondrial Complex III assembly in human cells. *ELife*, *10*, e67624. https://doi.org/10.7554/eLife.67624

Lenoir, W. F., Morgado, M., DeWeirdt, P. C., McLaughlin, M., Griffith, A. L., Sangree, A. K., Feeley, M. N., Esmaeili Anvar, N., Kim, E., Bertolet, L. L., Colic, M., Dede, M., Doench, J. G., & Hart, T. (2021). Discovery of putative tumor suppressors from CRISPR screens reveals rewired lipid metabolism in acute myeloid leukemia cells. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-26867-8

Li, B., Clohisey, S. M., Chia, B. S., Wang, B., Cui, A., Eisenhaure, T., Schweitzer, L. D., Hoover, P., Parkinson, N. J., Nachshon, A., Smith, N., Regan, T., Farr, D., Gutmann, M. U., Bukhari, S. I., Law, A., Sangesland, M., Gat-Viks, I., Digard, P., … Hacohen, N. (2020). Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection. *Nature Communications*, *11*(1), 164. https://doi.org/10.1038/s41467-019-13965-x

Li, S., Qian, N., Jiang, C., Zu, W., Liang, A., Li, M., Elledge, S. J., & Tan, X. (2022). Gain-of-function genetic screening identifies the antiviral function of TMEM120A via STING activation. *Nature Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-021-27670-1

Li, X., Jun, Y., Erickstad, M. J., Brown, S. D., Parks, A., Court, D. L., & Jun, S. (2016). tCRISPRi: Tunable and reversible, one-step control of gene expression. *Scientific Reports*, *6*, 39076. https://doi.org/10.1038/srep39076

Li, Y., Muffat, J., Omer Javed, A., Keys, H. R., Lungjangwa, T., Bosch, I., Khan, M., Virgilio, M. C., Gehrke, L., Sabatini, D. M., & Jaenisch, R. (2019). Genome-wide CRISPR screen for Zika virus resistance in human neural cells. *Proceedings of the National Academy of Sciences*, *116*(19), 9527–9532.

Linders, P. T. A., Ioannidis, M., Ter Beest, M., & van den Bogaart, G. (2022). Fluorescence Lifetime Imaging of pH along the Secretory Pathway. *ACS Chemical Biology*, *17*(1), 240–251. https://doi.org/10.1021/acschembio.1c00907

Lipsitch, M. (2018). Why Do Exceptionally Dangerous Gain-of-Function Experiments in Influenza? In Y. Yamauchi (Ed.), *Influenza Virus: Methods and Protocols* (pp. 589–608). Springer. https://doi.org/10.1007/978-1-4939-8678-1_29

Lipsitch, M., & Bloom, B. R. (2012). Rethinking Biosafety in Research on Potential Pandemic Pathogens. *MBio*, *3*(5), e00360-12. https://doi.org/10.1128/mBio.00360-12

Liu, D., Wu, H., Wang, C., Li, Y., Tian, H., Siraj, S., Sehgal, S. A., Wang, X., Wang, J., Shang, Y., Jiang, Z., Liu, L., & Chen, Q. (2019). STING directly activates autophagy to tune the innate immune response. *Cell Death and Differentiation*, *26*(9), 1735–1749. https://doi.org/10.1038/s41418-018-0251-z

Liu, S., Cai, X., Wu, J., Cong, Q., Chen, X., Li, T., Du, F., Ren, J., Wu, Y.-T., Grishin, N. V., & Chen, Z. J. (2015). Phosphorylation of innate immune adaptor proteins MAVS, STING, and TRIF induces IRF3 activation. *Science*, *347*(6227), aaa2630. https://doi.org/10.1126/science.aaa2630

Liu, X., Kimmey, J. M., Matarazzo, L., Bakker, V. de, Maele, L. V., Sirard, J.-C., Nizet, V., & Veening, J.-W. (2021). Exploration of Bacterial Bottlenecks and Streptococcus pneumoniae Pathogenesis by CRISPRi-Seq. *Cell Host & Microbe*, *29*(1), 107-120.e6. https://doi.org/10.1016/j.chom.2020.10.001

Lo, T. Q., Marston, B. J., Dahl, B. A., & De Cock, K. M. (2017). Ebola: Anatomy of an Epidemic. *Annual Review of Medicine*, *68*, 359–370. https://doi.org/10.1146/annurev-med-052915-015604

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lu, D., Shang, G., Li, J., Lu, Y., Bai, X., & Zhang, X. (2022). Activation of STING by targeting a pocket in the transmembrane domain. *Nature*, *604*(7906), Article 7906. https://doi.org/10.1038/s41586-022-04559-7

Ludwig, C. H., Thurm, A. R., Morgens, D. W., Yang, K. J., Tycko, J., Bassik, M. C., Glaunsinger, B. A., & Bintu, L. (2022). *High-Throughput Discovery and Characterization of Viral Transcriptional Effectors in Human Cells* (p. 2022.12.16.520835). bioRxiv. https://doi.org/10.1101/2022.12.16.520835

Lumb, J. H., Li, Q., Popov, L. M., Ding, S., Keith, M. T., Merrill, B. D., Greenberg, H. B., Li, J. B., & Carette, J. E. (2017). DDX6 Represses Aberrant Activation of Interferon-Stimulated Genes. *Cell Reports*, *20*(4), 819–831. https://doi.org/10.1016/j.celrep.2017.06.085

Luthra, P., Naidoo, J., Pietzsch, C. A., De, S., Khadka, S., Anantpadma, M., Williams, C. G., Edwards, M. R., Davey, R. A., Bukreyev, A., Ready, J. M., & Basler, C. F. (2018). Inhibiting pyrimidine biosynthesis impairs Ebola virus replication through depletion of nucleoside pools and activation of innate immune responses. *Antiviral Research*, *158*, 288–302. https://doi.org/10.1016/j.antiviral.2018.08.012

Ma, H., Dang, Y., Wu, Y., Jia, G., Anaya, E., Zhang, J., Abraham, S., Choi, J.-G., Shi, G., Qi, L., Manjunath, N., & Wu, H. (2015). A CRISPR-Based Screen Identifies

Genes Essential for West-Nile-Virus-Induced Cell Death. *Cell Reports*, *12*(4), 673–683. https://doi.org/10.1016/j.celrep.2015.06.049

Ma, Y., Walsh, M. J., Bernhardt, K., Ashbaugh, C. W., Trudeau, S. J., Ashbaugh, I. Y., Jiang, S., Jiang, C., Zhao, B., Root, D. E., Doench, J. G., & Gewurz, B. E. (2017). CRISPR/Cas9 Screens Reveal Epstein-Barr Virus-Transformed B Cell Host Dependency Factors. *Cell Host & Microbe*, *21*(5), 580-591.e7. https://doi.org/10.1016/j.chom.2017.04.005

Malvy, D., McElroy, A. K., de Clerck, H., Günther, S., & van Griensven, J. (2019). Ebola virus disease. *Lancet (London, England)*, *393*(10174), 936–948. https://doi.org/10.1016/S0140-6736(18)33132-5

Mangiameli, S., Chen, H., Earl, A. S., Dobkin, J., Lesman, D., Buenrostro, J., & Chen, F. (2022). *Photoselective sequencing: Microscopically-guided genomic measurements with subcellular resolution* (p. 2022.09.12.507600). bioRxiv. https://doi.org/10.1101/2022.09.12.507600

Marceau, C. D., Puschnik, A. S., Majzoub, K., Ooi, Y. S., Brewer, S. M., Fuchs, G., Swaminathan, K., Mata, M. A., Elias, J. E., Sarnow, P., & Carette, J. E. (2016). Genetic dissection of Flaviviridae host factors through genome-scale CRISPR screens. *Nature*, *535*(7610), 159–163. https://doi.org/10.1038/nature18631

Markwell, M. A., & Paulson, J. C. (1980). Sendai virus utilizes specific sialyloligosaccharides as host cell receptor determinants. *Proceedings of the*

*National Academy of Sciences*, *77*(10), 5693–5697. https://doi.org/10.1073/pnas.77.10.5693

Martin, S., Chiramel, A. I., Schmidt, M. L., Chen, Y.-C., Whitt, N., Watt, A., Dunham, E. C., Shifflett, K., Traeger, S., Leske, A., Buehler, E., Martellaro, C., Brandt, J., Wendt, L., Müller, A., Peitsch, S., Best, S. M., Stech, J., Finke, S., … Hoenen, T. (2018). A genome-wide siRNA screen identifies a druggable host pathway essential for the Ebola virus life cycle. *Genome Medicine*, *10*(1), 58. https://doi.org/10.1186/s13073-018-0570-1

Martínez-Gil, L., Goff, P. H., Hai, R., García-Sastre, A., Shaw, M. L., & Palese, P. (2013). A Sendai Virus-Derived RNA Agonist of RIG-I as a Virus Vaccine Adjuvant. *Journal of Virology*, *87*(3), 1290–1300. https://doi.org/10.1128/JVI.02338-12

McKenna, M. J., Adams, B. M., Chu, V., Paulo, J. A., & Shao, S. (2022). ATP13A1 prevents ERAD of folding-competent mislocalized and misoriented proteins. *Molecular Cell*, *82*(22), 4277-4289.e10. https://doi.org/10.1016/j.molcel.2022.09.035

McKenna, M. J., Sim, S. I., Ordureau, A., Wei, L., Harper, J. W., Shao, S., & Park, E. (2020). The endoplasmic reticulum P5A-ATPase is a transmembrane helix dislocase. *Science*, *369*(6511). https://doi.org/10.1126/science.abc5809

McNally, K. E., Faulkner, R., Steinberg, F., Gallon, M., Ghai, R., Pim, D., Langton, P., Pearson, N., Danson, C. M., Nägele, H., Morris, L. L., Singla, A., Overlee, B. L., Heesom, K. J., Sessions, R., Banks, L., Collins, B. M., Berger, I., Billadeau, D.

D., … Cullen, P. J. (2017). Retriever is a multiprotein complex for retromer-independent endosomal cargo recycling. *Nature Cell Biology*, *19*(10), Article 10. https://doi.org/10.1038/ncb3610

Mercer, J., Snijder, B., Sacher, R., Burkard, C., Bleck, C. K. E., Stahlberg, H., Pelkmans, L., & Helenius, A. (2012). RNAi Screening Reveals Proteasome- and Cullin3-Dependent Stages in Vaccinia Virus Infection. *Cell Reports*, *2*(4), 1036–1047. https://doi.org/10.1016/j.celrep.2012.09.003

Miesenböck, G., De Angelis, D. A., & Rothman, J. E. (1998). Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature*, *394*(6689), 192–195. https://doi.org/10.1038/28190

Migliorini, D., Bogaerts, S., Defever, D., Vyas, R., Denecker, G., Radaelli, E., Zwolinska, A., Depaepe, V., Hochepied, T., Skarnes, W. C., & Marine, J.-C. (2011). Cop1 constitutively regulates c-Jun protein stability and functions as a tumor suppressor in mice. *The Journal of Clinical Investigation*, *121*(4), 1329–1343. https://doi.org/10.1172/JCI45784

Miyake, T., Farley, C. M., Neubauer, B. E., Beddow, T. P., Hoenen, T., & Engel, D. A. (2020). Ebola Virus Inclusion Body Formation and RNA Synthesis Are Controlled by a Novel Domain of Nucleoprotein Interacting with VP35. *Journal of Virology*, *94*(16), e02100-19. https://doi.org/10.1128/JVI.02100-19

Monson, E. A., Crosse, K. M., Das, M., & Helbig, K. J. (2018). Lipid droplet density alters the early innate immune response to viral infection. *PLoS ONE*, *13*(1), e0190597. https://doi.org/10.1371/journal.pone.0190597

Moon, K. R., van Dijk, D., Wang, Z., Gigante, S., Burkhardt, D. B., Chen, W. S., Yim, K., Elzen, A. van den, Hirn, M. J., Coifman, R. R., Ivanova, N. B., Wolf, G., & Krishnaswamy, S. (2019). Visualizing structure and transitions in high-dimensional biological data. *Nature Biotechnology*, *37*(12), Article 12. https://doi.org/10.1038/s41587-019-0336-3

Moreno-Sánchez, R., Rodríguez-Enríquez, S., Marín-Hernández, A., & Saavedra, E. (2007). Energy metabolism in tumor cells. *The FEBS Journal*, *274*(6), 1393–1418. https://doi.org/10.1111/j.1742-4658.2007.05686.x

Morgens, D. W., Nandakumar, D., Didychuk, A. L., Yang, K. J., & Glaunsinger, B. A. (2022). A Two-tiered functional screen identifies herpesviral transcriptional modifiers and their essential domains. *PLOS Pathogens*, *18*(1), e1010236. https://doi.org/10.1371/journal.ppat.1010236

Morgens, D. W., Wainberg, M., Boyle, E. A., Ursu, O., Araya, C. L., Tsui, C. K., Haney, M. S., Hess, G. T., Han, K., Jeng, E. E., Li, A., Snyder, M. P., Greenleaf, W. J., Kundaje, A., & Bassik, M. C. (2017). Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens. *Nature Communications*, *8*(1), Article 1. https://doi.org/10.1038/ncomms15178

Mori, H., Connell, J. P., Donahue, C. J., Boytz, R., Nguyen, Y. T. K., Leung, D. W., LaCount, D. J., & Davey, R. A. (2022). CAPG Is Required for Ebola Virus Infection by Controlling Virus Egress from Infected Cells. *Viruses*, *14*(9), Article 9. https://doi.org/10.3390/v14091903

Mor-Shaked, H., Salah, S., Yanovsky-Dagan, S., Meiner, V., Atawneh, O. M., Abu-Libdeh, B., Elpeleg, O., & Harel, T. (2021). Biallelic deletion in a minimal CAPN15 intron in siblings with a recognizable syndrome of congenital malformations and developmental delay. *Clinical Genetics*, *99*(4), 577–582. https://doi.org/10.1111/cge.13920

Morwitzer, M. J., Tritsch, S. R., Cazares, L. H., Ward, M. D., Nuss, J. E., Bavari, S., & Reid, S. P. (2019). Identification of RUVBL1 and RUVBL2 as Novel Cellular Interactors of the Ebola Virus Nucleoprotein. *Viruses*, *11*(4), 372. https://doi.org/10.3390/v11040372

Mosallanejad, K., & Kagan, J. C. (2022). Control of innate immunity by the cGAS-STING pathway. *Immunology & Cell Biology*, *100*(6), 409–423. https://doi.org/10.1111/imcb.12555

Mosca, J. D., & Pitha, P. M. (1986). Transcriptional and posttranscriptional regulation of exogenous human beta interferon gene in simian cells defective in interferon synthesis. *Molecular and Cellular Biology*, *6*(6), 2279–2283.

Mudhasani, R., Kota, K. P., Retterer, C., Tran, J. P., Whitehouse, C. A., & Bavari, S. (2014). High Content Image-Based Screening of a Protease Inhibitor Library

Reveals Compounds Broadly Active against Rift Valley Fever Virus and Other Highly Pathogenic RNA Viruses. *PLoS Neglected Tropical Diseases*, *8*(8). https://doi.org/10.1371/journal.pntd.0003095

Mulangu, S., Dodd, L. E., Davey, R. T., Tshiani Mbaya, O., Proschan, M., Mukadi, D., Lusakibanza Manzo, M., Nzolo, D., Tshomba Oloma, A., Ibanda, A., Ali, R., Coulibaly, S., Levine, A. C., Grais, R., Diaz, J., Lane, H. C., Muyembe-Tamfum, J.-J., PALM Writing Group, Sivahera, B., … PALM Consortium Study Team. (2019). A Randomized, Controlled Trial of Ebola Virus Disease Therapeutics. *The New England Journal of Medicine*, *381*(24), 2293–2303. https://doi.org/10.1056/NEJMoa1910993

Nanbo, A., Watanabe, S., Halfmann, P., & Kawaoka, Y. (2013). The spatio-temporal distribution dynamics of Ebola virus proteins and RNA in infected cells. *Scientific Reports*, *3*(1), Article 1. https://doi.org/10.1038/srep01206

O'Hearn, A., Wang, M., Cheng, H., Lear-Rooney, C. M., Koning, K., Rumschlag-Booms, E., Varhegyi, E., Olinger, G., & Rong, L. (2015). Role of EXT1 and Glycosaminoglycans in the Early Stage of Filovirus Entry. *Journal of Virology*, *89*(10), 5441–5449. https://doi.org/10.1128/JVI.03689-14

Orvedahl, A., Sumpter, R., Xiao, G., Ng, A., Zou, Z., Tang, Y., Narimatsu, M., Gilpin, C., Sun, Q., Roth, M., Forst, C. V., Wrana, J. L., Zhang, Y. E., Luby-Phelps, K., Xavier, R. J., Xie, Y., & Levine, B. (2011). Image-based genome-wide siRNA screen identifies selective autophagy factors. *Nature*, *480*(7375), 113–117. https://doi.org/10.1038/nature10546

Oyler-Yaniv, J., Oyler-Yaniv, A., Maltz, E., & Wollman, R. (2021). TNF controls a speed-accuracy tradeoff in the cell death decision to restrict viral spread. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-23195-9

Pacheco, A. R., Lazarus, J. E., Sit, B., Schmieder, S., Lencer, W. I., Blondel, C. J., Doench, J. G., Davis, B. M., & Waldor, M. K. (2018). CRISPR Screen Reveals that EHEC's T3SS and Shiga Toxin Rely on Shared Host Factors for Infection. *MBio*, *9*(3), e01003-18. https://doi.org/10.1128/mBio.01003-18

Papalexi, E., Mimitou, E. P., Butler, A. W., Foster, S., Bracken, B., Mauck, W. M., Wessels, H.-H., Hao, Y., Yeung, B. Z., Smibert, P., & Satija, R. (2021). Characterizing the molecular regulation of inhibitory immune checkpoints with multimodal single-cell screens. *Nature Genetics*, *53*(3), Article 3. https://doi.org/10.1038/s41588-021-00778-2

Park, J. S., Helble, J. D., Lazarus, J. E., Yang, G., Blondel, C. J., Doench, J. G., Starnbach, M. N., & Waldor, M. K. (2019). A FACS-Based Genome-wide CRISPR Screen Reveals a Requirement for COPI in Chlamydia trachomatis Invasion. *IScience*, *11*, 71–84. https://doi.org/10.1016/j.isci.2018.12.011

Park, R. J., Wang, T., Koundakjian, D., Hultquist, J. F., Lamothe-Molina, P., Monel, B., Schumann, K., Yu, H., Krupzcak, K. M., Garcia-Beltran, W., Piechocka-Trocha, A., Krogan, N. J., Marson, A., Sabatini, D. M., Lander, E. S., Hacohen, N., &

Walker, B. D. (2017). A genome-wide CRISPR screen identifies a restricted set of HIV host dependency factors. *Nature Genetics*, *49*(2), 193–203. https://doi.org/10.1038/ng.3741

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems*, *32*. https://papers.nips.cc/paper_files/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830.

Peidli, S., Green, T. D., Shen, C., Gross, T., Min, J., Garda, S., Taylor-King, J. P., Marks, D. S., Luna, A., Blüthgen, N., & Sander, C. (2022). *scPerturb: Information Resource for Harmonized Single-Cell Perturbation Data* (p. 2022.08.20.504663). bioRxiv. https://doi.org/10.1101/2022.08.20.504663

Peng, T., Thorn, K., Schroeder, T., Wang, L., Theis, F. J., Marr, C., & Navab, N. (2017). A BaSiC tool for background and shading correction of optical microscopy images. *Nature Communications*, *8*(1), Article 1. https://doi.org/10.1038/ncomms14836

Pentecost, M., Vashisht, A. A., Lester, T., Voros, T., Beaty, S. M., Park, A., Wang, Y. E., Yun, T. E., Freiberg, A. N., Wohlschlegel, J. A., & Lee, B. (2015). Evidence for Ubiquitin-Regulated Nuclear and Subnuclear Trafficking among Paramyxovirinae Matrix Proteins. *PLOS Pathogens*, *11*(3), e1004739. https://doi.org/10.1371/journal.ppat.1004739

Peters, J. M., Koo, B.-M., Patino, R., Heussler, G. E., Hearne, C. C., Qu, J., Inclan, Y. F., Hawkins, J. S., Lu, C. H. S., Silvis, M. R., Harden, M. M., Osadnik, H., Peters, J. E., Engel, J. N., Dutton, R. J., Grossman, A. D., Gross, C. A., & Rosenberg, O. S. (2019). Enabling genetic analysis of diverse bacteria with Mobile-CRISPRi. *Nature Microbiology*, *4*(2), Article 2. https://doi.org/10.1038/s41564-018-0327-z

Peters, J. M., Silvis, M. R., Zhao, D., Hawkins, J. S., Gross, C. A., & Qi, L. S. (2015). Bacterial CRISPR: Accomplishments and prospects. *Current Opinion in Microbiology*, *27*, 121–126. https://doi.org/10.1016/j.mib.2015.08.007

Pillay, S., Meyer, N. L., Puschnik, A. S., Davulcu, O., Diep, J., Ishikawa, Y., Jae, L. T., Wosen, J. E., Nagamine, C. M., Chapman, M. S., & Carette, J. E. (2016). An essential receptor for adeno-associated virus infection. *Nature*, *530*(7588), 108–112. https://doi.org/10.1038/nature16465

Pinello, L., Canver, M. C., Hoban, M. D., Orkin, S. H., Kohn, D. B., Bauer, D. E., & Yuan, G.-C. (2016). Analyzing CRISPR genome-editing experiments with CRISPResso. *Nature Biotechnology*, *34*(7), Article 7. https://doi.org/10.1038/nbt.3583

Ploss, A., Evans, M. J., Gaysinskaya, V. A., Panis, M., You, H., de Jong, Y. P., & Rice, C. M. (2009). Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. *Nature*, *457*(7231), 882–886. https://doi.org/10.1038/nature07684

Poston, D., Weisblum, Y., Hobbs, A., & Bieniasz, P. D. (2022). VPS29 Exerts Opposing Effects on Endocytic Viral Entry. *MBio*, *13*(2), e03002-21. https://doi.org/10.1128/mbio.03002-21

Pravda, L., Sehnal, D., Toušek, D., Navrátilová, V., Bazgier, V., Berka, K., Svobodová Vařeková, R., Koča, J., & Otyepka, M. (2018). MOLEonline: A web-based tool for analyzing channels, tunnels and pores (2018 update). *Nucleic Acids Research*, *46*(W1), W368–W373. https://doi.org/10.1093/nar/gky309

Przybyla, L., & Gilbert, L. A. (2022). A new era in functional genomics screens. *Nature Reviews Genetics*, *23*(2), Article 2. https://doi.org/10.1038/s41576-021-00409-w

Puschnik, A. S., Majzoub, K., Ooi, Y. S., & Carette, J. E. (2017). A CRISPR toolbox to study virus–host interactions. *Nature Reviews Microbiology*, *15*(6), Article 6. https://doi.org/10.1038/nrmicro.2017.29

Ratnasiri, K., Wilk, A. J., Lee, M. J., Khatri, P., & Blish, C. A. (2022). Single-cell RNA-seq methods to interrogate virus-host interactions. *Seminars in Immunopathology*. https://doi.org/10.1007/s00281-022-00972-2

Rebendenne, A., Roy, P., Bonaventure, B., Chaves Valadão, A. L., Desmarets, L., Arnaud-Arnould, M., Rouillé, Y., Tauziet, M., Giovannini, D., Touhami, J., Lee, Y., DeWeirdt, P., Hegde, M., Urbach, S., Koulali, K. E., de Gracia, F. G., McKellar, J., Dubuisson, J., Wencker, M., … Goujon, C. (2022). Bidirectional genome-wide CRISPR screens reveal host factors regulating SARS-CoV-2, MERS-CoV and seasonal HCoVs. *Nature Genetics*, *54*(8), Article 8. https://doi.org/10.1038/s41588-022-01110-2

Rehwinkel, J., & Gack, M. U. (2020). RIG-I-like receptors: Their regulation and roles in RNA sensing. *Nature Reviews Immunology*, *20*(9), 537–551. https://doi.org/10.1038/s41577-020-0288-3

Replogle, J. M., Saunders, R. A., Pogson, A. N., Hussmann, J. A., Lenail, A., Guna, A., Mascibroda, L., Wagner, E. J., Adelman, K., Lithwick-Yanai, G., Iremadze, N., Oberstrass, F., Lipson, D., Bonnar, J. L., Jost, M., Norman, T. M., & Weissman, J. S. (2022). Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq. *Cell*, *185*(14), 2559-2575.e28. https://doi.org/10.1016/j.cell.2022.05.013

Richardson, R. B., Ohlson, M. B., Eitson, J. L., Kumar, A., McDougal, M. B., Boys, I. N., Mar, K. B., De La Cruz-Rivera, P. C., Douglas, C., Konopka, G., Xing, C., &

Schoggins, J. W. (2018). A CRISPR screen identifies IFI6 as an ER-resident interferon effector that blocks flavivirus replication. *Nature Microbiology*, *3*(11), 1214–1223. https://doi.org/10.1038/s41564-018-0244-1

Ringel, T., Frey, N., Ringnalda, F., Janjuha, S., Cherkaoui, S., Butz, S., Srivatsa, S., Pirkl, M., Russo, G., Villiger, L., Rogler, G., Clevers, H., Beerenwinkel, N., Zamboni, N., Baubec, T., & Schwank, G. (2020). Genome-Scale CRISPR Screening in Human Intestinal Organoids Identifies Drivers of TGF-β Resistance. *Cell Stem Cell*, *26*(3), 431-440.e8. https://doi.org/10.1016/j.stem.2020.02.007

Roberts, M. A., Deol, K. K., Lange, M., Leto, D. E., Mathiowetz, A. J., Stevenson, J., Hashemi, S. H., Morgens, D. W., Easter, E., Heydari, K., Nalls, M. A., Bassik, M. C., Kampmann, M., Kopito, R. R., Faghri, F., & Olzmann, J. A. (2022). *Parallel CRISPR-Cas9 screens reveal mechanisms of PLIN2 and lipid droplet regulation* (p. 2022.08.27.505556). bioRxiv. https://doi.org/10.1101/2022.08.27.505556

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Rodriques, S. G., Stickels, R. R., Goeva, A., Martin, C. A., Murray, E., Vanderburg, C. R., Welch, J., Chen, L. M., Chen, F., & Macosko, E. Z. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science*, *363*(6434), 1463–1467. https://doi.org/10.1126/science.aaw1219

Rosiana, S., Zhang, L., Kim, G. H., Revtovich, A. V., Uthayakumar, D., Sukumaran, A., Geddes-McAlister, J., Kirienko, N. V., & Shapiro, R. S. (2021). Comprehensive genetic analysis of adhesin proteins and their role in virulence of Candida albicans. *Genetics*, *217*(2), iyab003. https://doi.org/10.1093/genetics/iyab003

Rousset, F., Cui, L., Siouve, E., Becavin, C., Depardieu, F., & Bikard, D. (2018). Genome-wide CRISPR-dCas9 screens in E. coli identify essential genes and phage host factors. *PLOS Genetics*, *14*(11), e1007749. https://doi.org/10.1371/journal.pgen.1007749

Saeed, M., Andreo, U., Chung, H.-Y., Espiritu, C., Branch, A. D., Silva, J. M., & Rice, C. M. (2015). SEC14L2 enables pan-genotype HCV replication in cell culture. *Nature*, *524*(7566), 471–475. https://doi.org/10.1038/nature14899

Saliba, A.-E., Li, L., Westermann, A. J., Appenzeller, S., Stapels, D. A. C., Schulte, L. N., Helaine, S., & Vogel, J. (2016). Single-cell RNA-seq ties macrophage polarization to growth rate of intracellular Salmonella. *Nature Microbiology*, *2*(2), Article 2. https://doi.org/10.1038/nmicrobiol.2016.206

Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A. L., Nada, S., & Sabatini, D. M. (2010). Ragulator-Rag Complex Targets mTORC1 to the Lysosomal Surface and Is Necessary for Its Activation by Amino Acids. *Cell*, *141*(2), 290–303. https://doi.org/10.1016/j.cell.2010.02.024

Sankaranarayanan, S., De Angelis, D., Rothman, J. E., & Ryan, T. A. (2000). The use of pHluorins for optical measurements of presynaptic activity. *Biophysical Journal*, *79*(4), 2199–2208. https://doi.org/10.1016/S0006-3495(00)76468-X

Sanson, K. R., Hanna, R. E., Hegde, M., Donovan, K. F., Strand, C., Sullender, M. E., Vaimberg, E. W., Goodale, A., Root, D. E., Piccioni, F., & Doench, J. G. (2018). Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-07901-8

Savidis, G., McDougall, W. M., Meraner, P., Perreira, J. M., Portmann, J. M., Trincucci, G., John, S. P., Aker, A. M., Renzette, N., Robbins, D. R., Guo, Z., Green, S., Kowalik, T. F., & Brass, A. L. (2016). Identification of Zika Virus and Dengue Virus Dependency Factors using Functional Genomics. *Cell Reports*, *16*(1), 232–246. https://doi.org/10.1016/j.celrep.2016.06.028

Schneider, W. M., Luna, J. M., Hoffmann, H.-H., Sánchez-Rivera, F. J., Leal, A. A., Ashbrook, A. W., Le Pen, J., Ricardo-Lax, I., Michailidis, E., Peace, A., Stenzel, A. F., Lowe, S. W., MacDonald, M. R., Rice, C. M., & Poirier, J. T. (2021). Genome-Scale Identification of SARS-CoV-2 and Pan-coronavirus Host Factor Networks. *Cell*, *184*(1), 120-132.e14. https://doi.org/10.1016/j.cell.2020.12.006

Schoggins, J. W., MacDuff, D. A., Imanaka, N., Gainey, M. D., Shrestha, B., Eitson, J. L., Mar, K. B., Richardson, R. B., Ratushny, A. V., Litvak, V., Dabelic, R., Manicassamy, B., Aitchison, J. D., Aderem, A., Elliott, R. M., García-Sastre, A., Racaniello, V., Snijder, E. J., Yokoyama, W. M., … Rice, C. M. (2014). Pan-viral specificity of IFN-induced genes reveals new roles for cGAS in innate immunity. *Nature*, *505*(7485), Article 7485. https://doi.org/10.1038/nature12862

Schoggins, J. W., Wilson, S. J., Panis, M., Murphy, M. Y., Jones, C. T., Bieniasz, P., & Rice, C. M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature*, *472*(7344), Article 7344. https://doi.org/10.1038/nature09907

Shang, G., Zhang, C., Chen, Z. J., Bai, X., & Zhang, X. (2019). Cryo-EM structures of STING reveal its mechanism of activation by cyclic GMP–AMP. *Nature*, *567*(7748), 389–393. https://doi.org/10.1038/s41586-019-0998-5

Shapiro, R. S., Chavez, A., Porter, C. B. M., Hamblin, M., Kaas, C. S., DiCarlo, J. E., Zeng, G., Xu, X., Revtovich, A. V., Kirienko, N. V., Wang, Y., Church, G. M., & Collins, J. J. (2018). A CRISPR–Cas9-based gene drive platform for genetic interaction analysis in Candida albicans. *Nature Microbiology*, *3*(1), Article 1. https://doi.org/10.1038/s41564-017-0043-0

Shen, H.-H., Lithgow, T., & Martin, L. L. (2013). Reconstitution of Membrane Proteins into Model Membranes: Seeking Better Ways to Retain Protein Activities. *International Journal of Molecular Sciences*, *14*(1), 1589–1607. https://doi.org/10.3390/ijms14011589

Shen, Y.-C., Ou, D.-L., Hsu, C., Lin, K.-L., Chang, C.-Y., Lin, C.-Y., Liu, S.-H., & Cheng, A.-L. (2013). Activating oxidative phosphorylation by a pyruvate dehydrogenase kinase inhibitor overcomes sorafenib resistance of hepatocellular carcinoma. *British Journal of Cancer*, *108*(1), Article 1. https://doi.org/10.1038/bjc.2012.559

Sidik, S. M., Huet, D., Ganesan, S. M., Huynh, M.-H., Wang, T., Nasamu, A. S., Thiru, P., Saeij, J. P. J., Carruthers, V. B., Niles, J. C., & Lourido, S. (2016). A Genome-wide CRISPR Screen in Toxoplasma Identifies Essential Apicomplexan Genes. *Cell*, *166*(6), 1423-1435.e12. https://doi.org/10.1016/j.cell.2016.08.019

Silva-Ayala, D., López, T., Gutiérrez, M., Perrimon, N., López, S., & Arias, C. F. (2013). Genome-wide RNAi screen reveals a role for the ESCRT complex in rotavirus cell entry. *Proceedings of the National Academy of Sciences*, *110*(25), 10270–10275. https://doi.org/10.1073/pnas.1304932110

Snijder, B., Sacher, R., Rämö, P., Liberali, P., Mench, K., Wolfrum, N., Burleigh, L., Scott, C. C., Verheije, M. H., Mercer, J., Moese, S., Heger, T., Theusner, K., Jurgeit, A., Lamparter, D., Balistreri, G., Schelhaas, M., De Haan, C. A. M., Marjomäki, V., … Pelkmans, L. (2012). Single-cell analysis of population context advances RNAi screening at multiple levels. *Molecular Systems Biology*, *8*, 579. https://doi.org/10.1038/msb.2012.9

Song, Y., Lai, L., Chong, Z., He, J., Zhang, Y., Xue, Y., Xie, Y., Chen, S., Dong, P., Chen, L., Chen, Z., Dai, F., Wan, X., Xiao, P., Cao, X., Liu, Y., & Wang, Q. (2017). E3 ligase FBXW7 is critical for RIG-I stabilization during antiviral responses. *Nature Communications*, *8*(1), Article 1. https://doi.org/10.1038/ncomms14654

Sørensen, D. M., Holemans, T., Veen, S. van, Martin, S., Arslan, T., Haagendahl, I. W., Holen, H. W., Hamouda, N. N., Eggermont, J., Palmgren, M., & Vangheluwe, P. (2018). Parkinson disease related ATP13A2 evolved early in animal evolution. *PLOS ONE*, *13*(3), e0193228. https://doi.org/10.1371/journal.pone.0193228

Sørensen, D. M., Holen, H. W., Holemans, T., Vangheluwe, P., & Palmgren, M. G. (2015). Towards defining the substrate of orphan P5A-ATPases. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1850*(3), 524–535. https://doi.org/10.1016/j.bbagen.2014.05.008

Stickels, R. R., Murray, E., Kumar, P., Li, J., Marshall, J. L., Di Bella, D. J., Arlotta, P., Macosko, E. Z., & Chen, F. (2021). Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nature Biotechnology*, *39*(3), 313–319. https://doi.org/10.1038/s41587-020-0739-1

Subramanian, G., Kuzmanovic, T., Zhang, Y., Peter, C. B., Veleeparambil, M., Chakravarti, R., Sen, G. C., & Chattopadhyay, S. (2018). A new mechanism of interferon's antiviral action: Induction of autophagy, essential for paramyxovirus replication, is inhibited by the interferon stimulated gene, TDRD7. *PLOS Pathogens*, *14*(1), e1006877. https://doi.org/10.1371/journal.ppat.1006877

Sudhan, S. S., & Sharma, P. (2020). Human Viruses: Emergence and Evolution. *Emerging and Reemerging Viral Pathogens*, 53–68. https://doi.org/10.1016/B978-0-12-819400-3.00004-1

Sun, L., Wu, J., Du, F., Chen, X., & Chen, Z. J. (2013). Cyclic GMP-AMP Synthase Is a Cytosolic DNA Sensor That Activates the Type I Interferon Pathway. *Science*, *339*(6121), 786–791. https://doi.org/10.1126/science.1232458

Sunshine, S., Puschnik, A. S., Replogle, J. M., Laurie, M. T., Liu, J., Zha, B. S., Nuñez, J. K., Byrum, J. R., McMorrow, A. H., Frieman, M. B., Winkler, J., Qiu, X., Rosenberg, O. S., Leonetti, M. D., Ye, C. J., Weissman, J. S., DeRisi, J. L., & Hein, M. Y. (2022). *Systematic functional interrogation of SARS-CoV-2 host factors using Perturb-seq* (p. 2022.07.15.500120). bioRxiv. https://doi.org/10.1101/2022.07.15.500120

Suomalainen, M., & Greber, U. F. (2021). Virus Infection Variability by Single-Cell Profiling. *Viruses*, *13*(8), 1568. https://doi.org/10.3390/v13081568

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Mering, C. von. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, *47*(D1), D607–D613. https://doi.org/10.1093/nar/gky1131

Takahashi, H., Ranjan, A., Chen, S., Suzuki, H., Shibata, M., Hirose, T., Hirose, H., Sasaki, K., Abe, R., Chen, K., He, Y., Zhang, Y., Takigawa, I., Tsukiyama, T., Watanabe, M., Fujii, S., Iida, M., Yamamoto, J., Yamaguchi, Y., … Hatakeyama, S. (2020). The role of Mediator and Little Elongation Complex in transcription termination. *Nature Communications*, *11*(1), 1063. https://doi.org/10.1038/s41467-020-14849-1

Takamatsu, Y., Krähling, V., Kolesnikova, L., Halwe, S., Lier, C., Baumeister, S., Noda, T., Biedenkopf, N., & Becker, S. (2020). Serine-Arginine Protein Kinase 1 Regulates Ebola Virus Transcription. *MBio*, *11*(1), e02565-19. https://doi.org/10.1128/mBio.02565-19

Tang, Y.-D., Liu, J.-T., Wang, T.-Y., Sun, M.-X., Tian, Z.-J., & Cai, X.-H. (2017). Comparison of Pathogenicity-Related Genes in the Current Pseudorabies Virus Outbreak in China. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-08269-3

Todor, H., Silvis, M. R., Osadnik, H., & Gross, C. A. (2021). Bacterial CRISPR screens for gene function. *Current Opinion in Microbiology*, *59*, 102–109. https://doi.org/10.1016/j.mib.2020.11.005

Trombetta, J. J., Gennert, D., Lu, D., Satija, R., Shalek, A. K., & Regev, A. (2014). Preparation of Single-Cell RNA-Seq Libraries for Next Generation Sequencing. *Current Protocols in Molecular Biology*, *107*, 4.22.1-17. https://doi.org/10.1002/0471142727.mb0422s107

Tsai, M.-F., & Miller, C. (2013). Substrate selectivity in arginine-dependent acid resistance in enteric bacteria. *Proceedings of the National Academy of Sciences*, *110*(15), 5893–5897. https://doi.org/10.1073/pnas.1301442110

Ulferts, R., Marcassa, E., Timimi, L., Lee, L. C., Daley, A., Montaner, B., Turner, S. D., Florey, O., Baillie, J. K., & Beale, R. (2021). Subtractive CRISPR screen identifies the ATG16L1/vacuolar ATPase axis as required for non-canonical LC3 lipidation. *Cell Reports*, *37*(4), 109899. https://doi.org/10.1016/j.celrep.2021.109899

Ungricht, R., Guibbal, L., Lasbennes, M.-C., Orsini, V., Beibel, M., Waldt, A., Cuttat, R., Carbone, W., Basler, A., Roma, G., Nigsch, F., Tchorz, J. S., Hoepfner, D., & Hoppe, P. S. (2022). Genome-wide screening in human kidney organoids identifies developmental and disease-related aspects of nephrogenesis. *Cell Stem Cell*, *29*(1), 160-175.e7. https://doi.org/10.1016/j.stem.2021.11.001

van der Leun, A. M., Hoekstra, M. E., Reinalda, L., Scheele, C. L. G. J., Toebes, M., van de Graaff, M. J., Chen, L. Y. Y., Li, H., Bercovich, A., Lubling, Y., David, E., Thommen, D. S., Tanay, A., van Rheenen, J., Amit, I., van Kasteren, S. I., & Schumacher, T. N. (2021). Single-cell analysis of regions of interest (SCARI) using a photosensitive tag. *Nature Chemical Biology*, *17*(11), Article 11. https://doi.org/10.1038/s41589-021-00839-x

Vickovic, S., Eraslan, G., Salmén, F., Klughammer, J., Stenbeck, L., Schapiro, D., Äijö, T., Bonneau, R., Bergenstråhle, L., Navarro, J. F., Gould, J., Griffin, G. K., Borg, Å., Ronaghi, M., Frisén, J., Lundeberg, J., Regev, A., & Ståhl, P. L. (2019). High-definition spatial transcriptomics for in situ tissue profiling. *Nature Methods*, *16*(10), 987–990. https://doi.org/10.1038/s41592-019-0548-y

Viennois, E., Chen, F., Laroui, H., Baker, M. T., & Merlin, D. (2013). Dextran sodium sulfate inhibits the activities of both polymerase and reverse transcriptase: Lithium chloride purification, a rapid and efficient technique to purify RNA. *BMC Research Notes*, *6*, 360. https://doi.org/10.1186/1756-0500-6-360

Walton, R. T., Singh, A., & Blainey, P. C. (2022). Pooled genetic screens with image-based profiling. *Molecular Systems Biology*, *18*(11), e10768. https://doi.org/10.15252/msb.202110768

Wang, L. (2021). Single-cell normalization and association testing unifying CRISPR screen and gene co-expression analyses with Normalisr. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-26682-1

Wang, R., Simoneau, C. R., Kulsuptrakul, J., Bouhaddou, M., Travisano, K. A., Hayashi, J. M., Carlson-Stevermer, J., Zengel, J. R., Richards, C. M., Fozouni, P., Oki, J., Rodriguez, L., Joehnk, B., Walcott, K., Holden, K., Sil, A., Carette, J. E., Krogan, N. J., Ott, M., & Puschnik, A. S. (2021). Genetic Screens Identify Host Factors for SARS-CoV-2 and Common Cold Coronaviruses. *Cell*, *184*(1), 106-119.e14. https://doi.org/10.1016/j.cell.2020.12.004

Wang, T., Guan, C., Guo, J., Liu, B., Wu, Y., Xie, Z., Zhang, C., & Xing, X.-H. (2018). Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-04899-x

Wet, T. J. de, Winkler, K. R., Mhlanga, M., Mizrahi, V., & Warner, D. F. (2020, November 6). *Arrayed CRISPRi and quantitative imaging describe the morphotypic landscape of essential mycobacterial genes*. ELife; eLife Sciences Publications Limited. https://doi.org/10.7554/eLife.60083

Willemsen, J., Wicht, O., Wolanski, J. C., Baur, N., Bastian, S., Haas, D. A., Matula, P., Knapp, B., Meyniel-Schicklin, L., Wang, C., Bartenschlager, R., Lohmann, V., Rohr, K., Erfle, H., Kaderali, L., Marcotrigiano, J., Pichlmair, A., & Binder, M. (2017). Phosphorylation-Dependent Feedback Inhibition of RIG-I by DAPK1 Identified by Kinome-wide siRNA Screening. *Molecular Cell*, *65*(3), 403-415.e8. https://doi.org/10.1016/j.molcel.2016.12.021

Woolsey, C., Menicucci, A. R., Cross, R. W., Luthra, P., Agans, K. N., Borisevich, V., Geisbert, J. B., Mire, C. E., Fenton, K. A., Jankeel, A., Anand, S., Ebihara, H., Geisbert, T. W., Messaoudi, I., & Basler, C. F. (2019). A VP35 Mutant Ebola Virus Lacks Virulence but Can Elicit Protective Immunity to Wild-Type Virus Challenge. *Cell Reports*, *28*(12), 3032-3046.e6. https://doi.org/10.1016/j.celrep.2019.08.047

Wozniak, A. L., Long, A., Jones-Jamtgaard, K. N., & Weinman, S. A. (2016). Hepatitis C virus promotes virion secretion through cleavage of the Rab7 adaptor protein RILP. *Proceedings of the National Academy of Sciences*, *113*(44), 12484–12489. https://doi.org/10.1073/pnas.1607277113

Wroblewska, A., Dhainaut, M., Ben-Zvi, B., Rose, S. A., Park, E. S., Amir, E.-A. D., Bektesevic, A., Baccarini, A., Merad, M., Rahman, A. H., & Brown, B. D. (2018).

Protein Barcodes Enable High-Dimensional Single-Cell CRISPR Screens. *Cell*, *175*(4), 1141-1155.e16. https://doi.org/10.1016/j.cell.2018.09.022

Wu, J., Sun, L., Chen, X., Du, F., Shi, H., Chen, C., & Chen, Z. J. (2013). Cyclic GMP-AMP Is an Endogenous Second Messenger in Innate Immune Signaling by Cytosolic DNA. *Science*, *339*(6121), 826–830. https://doi.org/10.1126/science.1229963

Wynne, J. W., Todd, S., Boyd, V., Tachedjian, M., Klein, R., Shiell, B., Dearnley, M., McAuley, A. J., Woon, A. P., Purcell, A. W., Marsh, G. A., & Baker, M. L. (2017). Comparative Transcriptomics Highlights the Role of the Activator Protein 1 Transcription Factor in the Host Response to Ebolavirus. *Journal of Virology*, *91*(23), e01174-17. https://doi.org/10.1128/JVI.01174-17

Xu, Y., Cheng, S., Zeng, H., Zhou, P., Ma, Y., Li, L., Liu, X., Shao, F., & Ding, J. (2022). ARF GTPases activate Salmonella effector SopF to ADP-ribosylate host V-ATPase and inhibit endomembrane damage-induced autophagy. *Nature Structural & Molecular Biology*, *29*(1), Article 1. https://doi.org/10.1038/s41594-021-00710-6

Xu, Y., Zhou, P., Cheng, S., Lu, Q., Nowak, K., Hopp, A.-K., Li, L., Shi, X., Zhou, Z., Gao, W., Li, D., He, H., Liu, X., Ding, J., Hottiger, M. O., & Shao, F. (2019). A Bacterial Effector Reveals the V-ATPase-ATG16L1 Axis that Initiates Xenophagy. *Cell*, *178*(3), 552-566.e20. https://doi.org/10.1016/j.cell.2019.06.007

Yan, M.-Y., Li, S.-S., Ding, X.-Y., Guo, X.-P., Jin, Q., & Sun, Y.-C. (2020). A CRISPR-Assisted Nonhomologous End-Joining Strategy for Efficient Genome Editing in Mycobacterium tuberculosis. *MBio*, *11*(1), e02364-19. https://doi.org/10.1128/mBio.02364-19

Yan, M.-Y., Zheng, D., Li, S.-S., Ding, X.-Y., Wang, C.-L., Guo, X.-P., Zhan, L., Jin, Q., Yang, J., & Sun, Y.-C. (2022). Application of combined CRISPR screening for genetic and chemical-genetic interaction profiling in Mycobacterium tuberculosis. *Science Advances*, *8*(47), eadd5907. https://doi.org/10.1126/sciadv.add5907

Yang, K. D., Damodaran, K., Venkatachalapathy, S., Soylemezoglu, A. C., Shivashankar, G. V., & Uhler, C. (2020). Predicting cell lineages using autoencoders and optimal transport. *PLOS Computational Biology*, *16*(4), e1007828. https://doi.org/10.1371/journal.pcbi.1007828

Yang, L., Zhu, Y., Yu, H., Cheng, X., Chen, S., Chu, Y., Huang, H., Zhang, J., & Li, W. (2020). ScMAGeCK links genotypes with multiple phenotypes in single-cell CRISPR screens. *Genome Biology*, *21*(1), 19. https://doi.org/10.1186/s13059-020-1928-4

Yeung, A. T. Y., Choi, Y. H., Lee, A. H. Y., Hale, C., Ponstingl, H., Pickard, D., Goulding, D., Thomas, M., Gill, E., Kim, J. K., Bradley, A., Hancock, R. E. W., & Dougan, G. (2019). A Genome-Wide Knockout Screen in Human Macrophages

Identified Host Factors Modulating Salmonella Infection. *MBio*, *10*(5), e02169-19. https://doi.org/10.1128/mBio.02169-19

Young, J., Dominicus, C., Wagener, J., Butterworth, S., Ye, X., Kelly, G., Ordan, M., Saunders, B., Instrell, R., Howell, M., Stewart, A., & Treeck, M. (2019). A CRISPR platform for targeted in vivo screens identifies Toxoplasma gondii virulence factors in mice. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-11855-w

Zanini, F., Pu, S.-Y., Bekerman, E., Einav, S., & Quake, S. R. (2018). Single-cell transcriptional dynamics of flavivirus infection. *ELife*, *7*, e32942. https://doi.org/10.7554/eLife.32942

Zanini, F., Robinson, M. L., Croote, D., Sahoo, M. K., Sanz, A. M., Ortiz-Lasso, E., Albornoz, L. L., Rosso, F., Montoya, J. G., Goo, L., Pinsky, B. A., Quake, S. R., & Einav, S. (2018). Virus-inclusive single-cell RNA sequencing reveals the molecular signature of progression to severe dengue. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(52), E12363–E12369. https://doi.org/10.1073/pnas.1813819115

Zhang, C., Shang, G., Gui, X., Zhang, X., Bai, X., & Chen, Z. J. (2019). Structural basis of STING binding with and phosphorylation by TBK1. *Nature*, *567*(7748), Article 7748. https://doi.org/10.1038/s41586-019-1000-2

Zhang, M., Wang, C., Otto, T. D., Oberstaller, J., Liao, X., Adapa, S. R., Udenze, K., Bronner, I. F., Casandra, D., Mayho, M., Brown, J., Li, S., Swanson, J., Rayner,

J. C., Jiang, R. H. Y., & Adams, J. H. (2018). Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science*, *360*(6388), eaap7847. https://doi.org/10.1126/science.aap7847

Zhao, B., Du, F., Xu, P., Shu, C., Sankaran, B., Bell, S. L., Liu, M., Lei, Y., Gao, X., Fu, X., Zhu, F., Liu, Y., Laganowsky, A., Zheng, X., Ji, J.-Y., West, A. P., Watson, R. O., & Li, P. (2019). A conserved PLPLRT/SD motif of STING mediates the recruitment and activation of TBK1. *Nature*, *569*(7758), Article 7758. https://doi.org/10.1038/s41586-019-1228-x

Zhou, P., She, Y., Dong, N., Li, P., He, H., Borio, A., Wu, Q., Lu, S., Ding, X., Cao, Y., Xu, Y., Gao, W., Dong, M., Ding, J., Wang, D.-C., Zamyatina, A., & Shao, F. (2018). Alpha-kinase 1 is a cytosolic innate immune receptor for bacterial ADP-heptose. *Nature*, *561*(7721), Article 7721. https://doi.org/10.1038/s41586-018-0433-3

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. *Nature Communications*, *12*(1), 961. https://doi.org/10.1038/s41467-021-21213-4