

## Systematic Discovery of Pathogen Effector Functions across Human Pathogens and Pathways

Tomas Pachano<sup>1\*</sup>, He Leng<sup>2\*</sup>, Guillaume Dugied<sup>2</sup>, Travis Tribble<sup>2,3</sup>, Vincent Loubiere<sup>1</sup>, Felix Rauh<sup>1</sup>, Yeojin Lee<sup>4</sup>, Alexander Schleiffer<sup>1</sup>, Veronika Young<sup>5</sup>, Benjamin Weller<sup>5</sup>, Eleanor A. Lyons<sup>6</sup>, Matthew R. Hass<sup>7</sup>, Leah C. Kottayan<sup>7,8</sup>, Matthew T. Weirauch<sup>7,8</sup>, Juan I. Fuxman Bass<sup>9</sup>, Hayley J. Newton<sup>10</sup>, Alexander W. Ensminger<sup>3,11</sup>, Pascal Falter-Braun<sup>5,12</sup>, Daniel Schramek<sup>3,4</sup>, Alexander Stark<sup>1,13#</sup>, Mikko Taipale<sup>2,3#</sup>

- 1) Research Institute of Molecular Pathology (IMP), Vienna Biocenter (VBC), Vienna, Austria.
- 2) Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Ontario, Canada.
- 3) Department of Molecular Genetics, University of Toronto, Ontario M5S 3E1, Canada
- 4) Centre for Molecular and Systems Biology, Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada.
- 5) Institute of Network Biology (INET), Molecular Targets and Therapeutics Center (MTTC), Helmholtz Munich, German Research Center for Environmental Health, Neuherberg, Germany.
- 6) Department of Microbiology and Immunology, University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia.
- 7) Center for Autoimmune Genomics and Etiology, Division of Allergy & Immunology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA.
- 8) Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA.
- 9) Biology Department and Bioinformatics Program, Boston University, Boston, MA, USA.
- 10) Infection Program, Monash Biomedicine Discovery Institute and Department of Microbiology, Monash University, Clayton, Victoria, Australia.
- 11) Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada.
- 12) Microbe-Host Interactions, Faculty of Biology, Ludwig-Maximilians-Universität (LMU) München, Planegg-Martinsried, Germany.
- 13) Medical University of Vienna, Vienna Biocenter (VBC), Vienna, Austria.

\* These authors contributed equally

# Co-corresponding authors

Correspondence to Mikko Taipale ([mikko.taipale@utoronto.ca](mailto:mikko.taipale@utoronto.ca)) or Alexander Stark ([alexander.stark@imp.ac.at](mailto:alexander.stark@imp.ac.at)).

## SUMMARY

Pathogens deploy effector proteins to exploit host cell biology, and most pathogen open reading frames (ORFs) are rapidly evolving and lack functional annotation. We developed the eORFeome, a scalable functional genomics platform encompassing 3,835 effector ORFs from diverse viruses, bacteria, and parasites. High-throughput barcoded screens across NF $\kappa$ B, apoptosis, p53, cGAS–STING and MHC-I pathways revealed functions for hundreds of uncharacterized eORFs, unexpected new activities for known effectors, and distinct pathway-specific functions encoded by single ORFs. Illustrating the power of the approach, we identify HHV6A U14 as a p53 antagonist, HHV7 U21 as a dual-function STING antagonist and MHC-I antigen display inhibitor, and adenoviral 13.6K/i-leader protein as a *de novo* evolved TAP inhibitor that suppresses MHC-I display. These results establish a general framework for systematic effector annotation, uncover new mechanisms of host–pathogen interaction across kingdoms, and highlight pathogen effectors as a versatile toolkit for rewiring and probing human cellular pathways.

## INTRODUCTION

Pathogens express specialized gene products, known as effectors, to manipulate host molecular processes and establish conditions that are favorable for their survival and transmission.<sup>1-3</sup> Many bacteria and eukaryotic parasites have evolved secretion systems that inject effector proteins into the host cell during infection, whereas viral effectors are mostly produced by the host ribosomes after viral entry (**Figure 1A**). By interfacing directly with host cellular machinery, effector proteins modulate immune responses, alter signaling pathways, and hijack metabolic processes to create a conducive environment for the pathogen.

At least 2,000 pathogenic viruses, bacteria, and parasites are known to infect humans<sup>4-6</sup> and the increasing availability of pathogen genome sequences has facilitated the identification of millions of viral open-reading frames (ORFs)<sup>7,8</sup> and the computational prediction of thousands of bacterial effectors.<sup>9-12</sup> However, with the exception of a few well-studied bacteria and viruses,<sup>13-17</sup> the vast majority of pathogen effectors remains uncharacterized. This leaves a large reservoir of pathogen proteins that are highly enriched for novel functions shaped by the unremitting arms race between pathogens and their hosts. Indeed, most viral proteins lack structural homologs in the AlphaFold database,<sup>18</sup> pointing to unique functions and mechanisms of action. Similarly, studies of secreted bacterial effectors have revealed many unexpected molecular mechanisms, including unique post-translational modifications.<sup>19</sup>

The fact that thousands of sequenced pathogen ORFs have not been characterized has far-reaching consequences for many fields, both scientific and societal. Pathogenic effector proteins target functions and pathways relevant to cell biology and human health, particularly to infection, inflammation, and cancer.<sup>20-23</sup> Systematic functional characterization of these effectors could uncover novel host vulnerabilities and reveal potential therapeutic entry points, thus pre-emptively illuminating mechanisms by which emerging pathogens might manipulate human cells. Moreover, the effectors constitute a vast untapped resource in novel tools to modulate and perturb host cell processes, potentially via entirely novel targets and mechanisms of action. As *evolutionarily optimized perturbagens*, they might moreover be more effective in modulating cellular functions than more traditional perturbations such as knocking out or overexpressing human genes.

Several previous studies have characterized host protein interactions or perturbations by multiple pathogen effectors, but they have mostly covered a single pathogen species or relatively small collections of effectors.<sup>24–34</sup> Thus, the discrepancy between the number of sequenced effector genes and those with known functions constitutes a major and critical knowledge gap for both fundamental discovery and potential therapeutics. To address this gap, we present a novel approach for investigating at an unprecedented scale how diverse effectors hijack host cell functions. Rather than focusing on a single pathogen or pathway, we developed scalable assays that enable the systematic functional profiling of thousands of effectors from a broad range of pathogens across a spectrum of host cellular pathways. This approach allowed us to uncover novel effectors that modulate key host proteins and provided general lessons for such approaches. Specifically, we uncovered novel functions for entirely uncharacterized ORFs, unexpected additional functions for known effectors, examples of close homologs with identical and with strongly divergent functions, as well as functions for recently evolved viral ORFs affecting a wide variety of pathways, including the NF- $\kappa$ B and p53 pathways, MHC-I antigen presentation, cGAS-STING signaling, and apoptosis. This work introduces a powerful framework for the annotation of host perturbations elicited by pathogen effectors and the discovery of novel protein function more generally.

## RESULTS

### The effector ORFeome collection 1.0

To systematically discover the function of thousands of eORFs, we assembled a large collection of effector ORFs. Currently, this collection, which we call the effector ORFeome (eORFeome), contains 3,835 putative effector ORFs (eORFs) from viruses, bacteria, and parasites (**Figure 1B**, **Figure S1** and **Table S1**). The collection includes virus-encoded ORFs from a broad phylogenetic range of viruses with representatives from all Baltimore classes, including dsDNA viruses (e.g. *Adenoviridae*, *Herpesviridae* and *Poxviridae*), ssDNA viruses (e.g. *Parvoviridae*, *Anelloviridae*), dsRNA viruses (e.g. *Reoviridae*, *Polyomaviridae*) and ssRNA viruses (e.g. *Coronaviridae*, *Orthomyxoviridae*, *Picornaviridae*). Our collection includes any virus-encoded ORF, including those encoding structural proteins. For polyproteins that are proteolytically processed into individual polypeptides, we included clones that correspond to the final cleaved proteins. For bacteria, the eORFeome comprises ORFs from curated lists of effectors secreted into host cells by type III, type IV or type VI secretion systems, such as those from *Legionella*, *Coxiella*, and *Chlamydia*.<sup>35–37</sup> We also included eORFs from the human microbiome effector ORFeome,<sup>38</sup> which encompasses effectors from commensal microbes from the human gut. Finally, we included 20 ORFs encoding predicted or known secreted effectors from the parasite *Toxoplasma gondii*.<sup>39</sup> In total, the eORFeome contains 3,835 effectors from 25 viral, 12 bacterial and 1 parasitic family, representing a comprehensive collection of diverse pathogens and their niches.

The eORFeome exhibits a significant sequence diversity: the 3,835 eORFs cluster into 3,208 groups at 80% sequence identity and 2,118 groups at 10% identity (**Figure 1C**). We recently showed that even highly similar effectors can have distinct interaction profiles and hence likely different functions<sup>38</sup>. Therefore, for certain pathogens (e.g. HPV, HAAdV, Enteroviruses), we included highly similar homologous ORFs from multiple strains (**Figure S1**). Notably, only 7% of the eORFs have direct experimental evidence for protein function and structure (i.e. UniProt annotation score of 5) compared to 72% of the human proteome (**Figure 1D**). The eORFs also are diverse in length, ranging from 22 to 2,225 amino acids (aa), with an average size of 345.5 aa (**Figure S2A**). Collectively, the eORFeome represents a functionally underexplored resource and a unique opportunity for large-scale characterization of pathogen effectors.

## **A barcoded lentiviral library and platform for inducible eORF expression and high-throughput functional screening**

We took advantage of the eORFeome to systematically identify pathogen effectors that hijack cellular functions using high-throughput selection-based screens. Inspired by screens based on human ORFeome overexpression,<sup>22,23</sup> we constructed a lentiviral eORFeome library in an all-in-one vector that enables both doxycycline (dox)-inducible expression and detection of barcoded eORFs by next-generation sequencing (**Figure S2B**). Briefly, each eORF is followed by a unique DNA barcode that allows identification of the eORF independent of the eORF length and, moreover, increases the statistical power of screens due to an average of 200 unique barcodes per eORF (**Figure S2C and S2D**). Introducing this library into a starting cell population at low multiplicity of infection (MOI) ensures that most cells express only a single eORF. Controlled eORF expression by dox treatment coupled to phenotypic selection of cells by either reporter gene expression and FACS-based sorting or proliferation will enrich for cells that express eORFs with the respective function of interest. Next-generation sequencing of the barcodes allows the identification of enriched or depleted eORFs in selected cells compared to the starting population.

## **A proof-of-principle screen for eORFs modulating the NF- $\kappa$ B pathway**

The NF- $\kappa$ B pathway regulates genes controlling host immunity, cell proliferation and apoptosis, making it a common target for viruses, bacteria and parasites.<sup>40-44</sup> Pathogens are known to inhibit NF- $\kappa$ B to suppress host immune response or, conversely, activate the pathway to avoid apoptosis (**Figure 2A**).<sup>40</sup> Thus, the NF- $\kappa$ B pathway provides an ideal setting for a proof-of-principle screen of the eORFeome.

We engineered an A549 reporter cell line in which an NF- $\kappa$ B response element drives GFP expression (see Methods). We tested the ability of the reporter to detect NF- $\kappa$ B inhibition by pathogen effectors by expressing two known secreted NF- $\kappa$ B inhibitors from *Salmonella*. GtgA cleaves the NF- $\kappa$ B p65 subunit, whereas SseK3 inhibits the pathway by covalently modifying the TRADD receptor with N-acetylglucosamine.<sup>43,45</sup> As expected, both eORFs strongly reduced GFP signal after stimulating the NF- $\kappa$ B pathway with TNF $\alpha$  (**Figure 2B and S3A**). Moreover, the reporter was induced in the absence of TNF $\alpha$  by two known viral NF- $\kappa$ B activators with distinct cellular targets and mechanisms: Kaposi's sarcoma herpesvirus (KSHV) vFLIP that activates the I $\kappa$ B kinase (IKK) complex, and KSHV ORF74, a ligand-

independent vGPCR that activates the TAK1 kinase (**Figure 2B** and **S3A**).<sup>46,47,47,48</sup> These results validate the suitability of the reporter for NF- $\kappa$ B modulator screens.

To identify both inhibitors and activators, we transduced the reporter cells with the lentiviral eORFeome library and conducted two parallel screens (**Figure 2C**). In the inhibitor screen, we stimulated the pathway with TNF $\alpha$  and used FACS to isolate cells with low GFP levels. In the activator screen, we omitted TNF $\alpha$  and isolated cells with high GFP expression. Following isolation, we quantified the eORF-associated barcodes in each population by deep sequencing and used MAGeCK (see Methods) to compute enrichment scores.<sup>49</sup>

In total, the screens identified 29 activators and 42 inhibitors of NF- $\kappa$ B at a false discovery rate (FDR) of 0.05 and fold change > 1.5 (**Table S2** provides enrichment and FDR values for all eORFs), including the positive controls (**Figure 2D**). The screens recovered multiple known NF- $\kappa$ B modulators (**Figure 2D**). These included viral and bacterial effectors, such as NleC proteins from enterobacteria and 3C proteases from enteroviruses, which inhibit NF- $\kappa$ B by directly cleaving the p65 subunit.<sup>50-52</sup> We identified three ligand-independent vGPCRs that mimic host chemokine receptors to activate NF- $\kappa$ B: Epstein-Barr virus EBV BILF1, cytomegalovirus (CMV) US28 and CMV UL138.<sup>53-55</sup> Additionally, we found two glycoproteins implicated in NF- $\kappa$ B activation: HHV2 gD, a homolog of the NF- $\kappa$ B activator from HHV1,<sup>56</sup> and EBV gH, which activates the pathway via interaction with TLR2.<sup>57</sup> The activation screen also enriched for E6 proteins from multiple HPV types, which are known to induce NF- $\kappa$ B activity through an unknown mechanism.<sup>58</sup>

The screens also identified novel NF- $\kappa$ B regulators. To experimentally validate these candidates, we cloned a subset of novel hits and individually transduced them into the NF- $\kappa$ B reporter cell line. All five tested novel activator candidates strongly induced the NF- $\kappa$ B reporter in the absence of TNF $\alpha$  (**Figure 2E**). Furthermore, all five novel inhibitor candidates significantly suppressed NF- $\kappa$ B reporter activation, with four showing strong inhibition and one displaying weaker but still statistically significant inhibition compared to control (**Figure 2F**).

We selected the novel inhibitor RavE to elucidate the mechanisms by which it inhibits the NF- $\kappa$ B pathway activity. RavE (Lpg0195) is a Type IV Secretion System (T4SS) effector from *Legionella pneumophila* with an unknown domain architecture and function. RavE

expression in A549 cells prevented the degradation of I $\kappa$ B $\alpha$ , the cytoplasmic inhibitor of NF- $\kappa$ B, and the subsequent phosphorylation of NF- $\kappa$ B (**Figure 2G**) after TNF $\alpha$  stimulation. Furthermore, RavE also inhibited IL-1 $\beta$ -mediated activation of NF- $\kappa$ B (**Figure 2H**). Since both the TNF $\alpha$  and IL-1 $\beta$  pathways converge at the IKK complex,<sup>59</sup> these results suggest that RavE targets the step of IKK complex activation and I $\kappa$ B $\alpha$  degradation to block signaling rather than an upstream event.

Foldseek<sup>60</sup> revealed RavE structural similarity with TIKI superfamily proteins<sup>61</sup> and, in particular, with the *Pseudomonas syringae* effector HopBA1 that can activate host cell death in plants (**Figure 2I**).<sup>62</sup> The TIKI superfamily comprises diverse esterases and proteases that share a putative active site.<sup>61</sup> Structural alignment suggested that RavE residues E19 and H21, which are conserved among RavE homologs in other *Legionella* species (**Figure S3B**), might be catalytic residues. Indeed, mutating these residues to alanines completely abolished RavE's ability to inhibit NF- $\kappa$ B activation by either TNF $\alpha$  or IL-1 $\beta$  (**Figure 2J and S3C,D**). Thus, RavE represents a novel *Legionella* effector that potently inhibits NF- $\kappa$ B pathway through highly conserved residues. More broadly, these results establish the eORFeome screens as a powerful screening and discovery platform to deliver assay-specific hits and uncover previously unknown functional host-pathogen interactions.

### **eORFeome screens uncover known and novel effectors for four additional pathways**

Building on the proof-of-principle NF- $\kappa$ B screens, we performed four additional screens representing pathways that are broadly relevant not only to infection but also to cancer, inflammation, and development: p53 signaling, STING signaling, apoptosis, and MHC-I antigen presentation. We selected these pathways because they play a central role both in normal cellular homeostasis and in host responses to infection. The p53 tumor suppressor regulates cell cycle progression, apoptosis, and DNA damage response and is a frequent target of viral oncogenes.<sup>63</sup> The cGAS-STING pathway acts as the major innate immune sensor for cytoplasmic dsDNA. It is essential for detecting DNA viruses and intracellular bacteria, but it also plays an important role in tumor immunology and inflammation.<sup>64</sup> Apoptosis is not only a conserved host defense mechanism against infection but also a key regulator of organismal development, tissue homeostasis and tumorigenesis.<sup>65</sup> Finally, antigen presentation by MHC-I enables recognition of infected or transformed cells by cytotoxic T cells and thereby is a central target of immune evasion strategies employed by both pathogens and cancer cells.<sup>66</sup>

In each case, we employed well-established reporters or markers to identify pathway inhibitors. For p53, we used an A549 cell line with a p53 response element driving GFP expression (see Methods). After activating the pathway with the MDM2/p53 interaction inhibitor nutlin, we used FACS to sort low GFP cells. For STING, after activating the pathway with the STING ligand 2',3'-cGAMP in U937 monocytes, we used an anti-IFIT1 antibody to sort cells that did not induce IFIT1 expression, a known cGAS-STING pathway target gene.<sup>67,68</sup> We induced apoptosis in A549 cells with staurosporine followed by sequencing of eORF barcodes in the surviving population. Finally, for MHC-I surface display, we used a pan-MHC-I antibody and FACS to identify pathway inhibitors.

In each screen, we identified several known effectors, such as HPV31 E6 inhibiting p53, Vaccinia (VACV) pox protein inhibiting STING signalling, EBV BHRF1 inhibiting apoptosis, or HHV7 U21 inhibiting MHC-I display (**Figure 3A, Table S2**). Moreover, the screens also uncovered hundreds of effectors with previously unrecognized activities, including both entirely uncharacterized proteins as well as known effectors with unexpected additional functions.

### **Parallel eORFeome screens uncover pathway-specific and pleiotropic effectors as well as shared and distinct functions for closely related eORF homologs**

Our screens revealed 164 distinct effector hits across the six screens. Although the screens targeted distinct pathways, 24% (40/164) of the hits scored in multiple screens, suggesting that a large fraction of pathogenic effectors target multiple processes. To discriminate pathway-specific from more pleiotropic effects, we used self-organizing map clustering to group the hits found in any of the screens into ten groups with distinct patterns of high to low scores in one or several screens (**Figure 3B, Table S3**). Most clusters were specific to a single pathway. Suggesting conserved functions for pathogen effectors, the functional clusters uncovered multiple homologous proteins among screen hits. For example, the apoptosis inhibition screen identified multiple Bcl2-like proteins, including five homologs of adenoviral E1B proteins known to suppress apoptosis (**Figure 3C**).<sup>69</sup> Similarly, among the top hits of the p53 screen were several E6 proteins from papillomaviruses that are associated with a high risk of cervical cancer or head and neck cancer. In contrast, none of the E6 proteins from low-risk HPV strains were identified as screen hits (**Figure 3C**), consistent

with the well-established oncogenic role of E6 from high-risk strains in degrading host p53.<sup>70,71</sup> Thus, the screens identified both conserved and strain-specific phenotypes for effectors, highlighting the power of unbiased, sensitive screens in functional discovery.

We identified several protein families (defined as homology groups of eORFs sharing at least 30% sequence identity) that were enriched in specific clusters (adjusted p-value < 0.05, **Figure 3D**). Some protein families were already linked to their respective cellular pathway, including the NF- $\kappa$ B-degrading zinc proteases from enterobacteria,<sup>43,72,73</sup> E6 proteins from HPVs<sup>70,71</sup> and E1B proteins from HAdV.<sup>69</sup> However, involvement of other protein families – including HHV U54 tegument proteins in apoptosis, and adenoviral i-leader/13.6K proteins in MHC-I display – has not been reported and could not have been predicted e.g. by sequence similarity. Thus, parallel eORFeome screens can functionally characterize hundreds of pathogen effector proteins, uncovering not only novel activities for individual effectors but also previously unrecognized functions for entire protein families as well as additional functions for known effectors. Next, we set out to validate and characterize novel hits corresponding to each of these general lessons, selected from three different cellular pathways.

#### **Discovery of a novel and unexpected function: HHV6A U14 is a potent p53 antagonist**

One of the most prominent hits in our p53 screen was the U14 protein from human herpesvirus 6A (HHV6A) (**Figure 4A**). A prior study reported that U14 interacts with p53 but concluded that it does not inhibit p53 activity.<sup>74</sup> We therefore investigated U14 in more detail to understand its mechanism of action and to assess if it is a true p53 antagonist.

We expressed U14 in a p53 reporter cell line and observed a marked suppression of p53 reporter activation, validating U14's role as a potent p53 antagonist (**Figure 4B**). Additionally, we validated the physical interaction between p53 and V5-tagged U14 with both affinity-purification coupled to mass spectrometry (AP-MS) and proximity-dependent biotinylation (BioID) (**Figure 4C and Table S4**). AlphaFold2 predicted that U14 forms a dimer with two principal regions of p53 contact (**Figure 4D and S4A**): Region 1 (R1) comprises residues D124 and R121 of one U14 protomer, predicted to interact with the core domain of p53 residues 177–181. Region 2 (R2) includes residues 356–376 of the second U14 protomer that are positioned to interact with the DNA-binding domain of p53 (aa 241–283). Thus, it is likely that both regions interfere with the binding of p53 to its cognate DNA

elements. Consistent with these predictions, point mutations in R1 (D124R) or R2 (D374R, D376R) abolished U14's ability to inhibit p53 activity (**Figure 4E** and **S4B,C**), thus supporting the structural model and highlighting the critical role of each of the two regions.

The U14-p53 interaction surface is also targeted by the Large T antigen (LTA<sub>g</sub>) from SV40,<sup>75</sup> a well-studied p53 antagonist (**Figure 4F**) that is unrelated to U14. LTA<sub>g</sub> stabilizes and inhibits p53 DNA binding,<sup>76</sup> in contrast to HPV E6 that degrades p53.<sup>70,71</sup> In line with the shared interaction surface, U14 did not reduce p53 protein levels, but rather stabilized p53 (**Figure 4G**). Furthermore, U14, but not the D124R mutant, impaired p21 induction following nutlin treatment in A549 cells (**Figure 4H**). Strikingly, U14 expression also led to p53 sequestration into distinct nuclear foci that overlap with U14 localization (**Figure 4I**). This relocalization depends on direct U14-p53 interaction, as the U14-D124R mutant, while still forming nuclear foci, could not recruit p53. Together, these findings suggest a functional evolutionary convergence between dissimilar effectors of two unrelated viruses, both capable of sequestering and stabilizing p53 while suppressing its trans-activation capacity.

Next, we hypothesized that U14 would interfere with p53-mediated tumor suppression and cooperate with the Kras<sup>G12D</sup> oncogene to drive tumor formation.<sup>77</sup> To test U14's biological impact *in vivo*, we engineered lentiviral vectors that contain the FLP recombinase and a double-floxed inverted orientation (DIO) cassette of U14 or eGFP (LV-Flp-DIO-U14 or LV-Flp-DIO-eGFP), which we delivered through intranasal inhalations into the lungs of Frt-Stop-Frt (FSF)-Kras<sup>G12D</sup>; Rosa26-CreERT2 mice<sup>78,79</sup>, where a premature stop codon prevents Kras<sup>G12D</sup> expression in the absence of recombination. FLP recombinase clonally activates the FSF-Kras<sup>G12D</sup> allele within the lung of inhaled mice, while administration of tamoxifen allows temporal control of the expression of U14 or eGFP via CreERT2. In alignment with our hypothesis, U14-expressing mice developed nearly double the tumor area compared to the eGFP controls (**Figure 4J**). This result provides strong evidence that U14 functions as a potent p53 antagonist *in vivo*, accelerating tumorigenesis by dismantling the cell's primary defense against oncogenic transformation.

### **Discovery of a new function for a known viral effector: HHV7 MHC-I inhibitor U21 is also a potent inhibitor of STING signaling**

Similar to the p53 screen, our screen for STING pathway inhibitors identified multiple known modulators of the pathway. One of the top hits was VACV pox, which was recently

described as a nuclease that degrades cGAMP (**Figure 5A**).<sup>80</sup> Other known STING pathway inhibitors included CMV UL42, which inhibits STING translocation,<sup>81</sup> and the Rabies virus P phosphoprotein that inhibits TBK1 and IRF3 function downstream of STING.<sup>82,83</sup> These results suggested that other hits would also be physiologically relevant modulators of the pathway. In particular, among the hits were two herpesvirus proteins, CMV US30 and HHV7 U21, with no previous connection to cGAS-STING signaling.

Activation of STING signaling with STING agonists in U937 cells leads to cell death that is dependent on an intact STING pathway.<sup>84,85</sup> We therefore assessed the ability of dox-inducible, GFP-tagged VACV poxin, CMV US30, and HHV7 U21 to suppress cGAMP-induced cell death. Consistent with the screen, all proteins potently suppressed cGAMP-induced apoptotic cell death when expressed individually (**Figure 5B** and **S5A**). However, in contrast to poxin, which could suppress cell death induced by cGAMP but not by the non-cyclic dinucleotide STING agonist diABZI,<sup>86</sup> the other two viral effectors counteracted the effect of both ligands (**Figure S5B**). Thus, the novel effectors function through a different mechanism than Vaccinia poxin.

All three effectors also inhibited STING phosphorylation as well as phosphorylation of downstream signaling components TBK1 and IRF3 after cGAMP treatment (**Figure 5C**). Interestingly, however, US30 and U21 had different effects: US30 inhibited STING phosphorylation without affecting its levels, whereas U21 expression led to lower STING protein levels. U21 is a single-pass transmembrane domain protein with an MHC-Ib-like and Ig-like luminal domains and a 50 amino acid unstructured cytoplasmic tail (**Figure 5D**).<sup>87</sup> U21 was originally identified as an immunoevasin that routes MHC-I molecules to lysosomes to evade recognition by cytotoxic T cells.<sup>88,89</sup> The luminal domain is required for MHC-I redirection to lysosomes but the cytoplasmic tail is dispensable for this activity,<sup>90</sup> which we also confirmed (**Figure 5E**). However, in contrast to this, we observed that the cytoplasmic tail was indispensable for STING pathway inhibition, indicating that these activities are genetically separable. Deleting the tail abolished the ability of U21 to suppress cGAMP-induced cell death (**Figure 5F** and **S4C**), downregulate STING levels, and inhibit cGAMP-induced STING, TBK1, and IRF3 phosphorylation (**Figure 5G**). There was no difference in STING mRNA levels, ruling out a transcriptional downregulation (**Figure S5D**). Moreover, U21 could downregulate STING levels independent of cGAMP treatment, indicating that it modulates basal STING levels (**Figure S5E**).

To investigate if U21 induces protein degradation in a more general manner, we conducted a whole-cell proteomics experiment in U937 cells expressing full-length GFP-tagged U21, HHV7 U21 lacking the cytoplasmic tail (U21 $\Delta$ tail), or Renilla luciferase as a control. After inducing the expression for 24 hours with dox, HHV7 U21 significantly upregulated only 5 proteins and downregulated 5 proteins, including STING (**Figure 5H** and **Table S5**). In contrast, U21 $\Delta$ tail did not have an effect on STING levels (**Figure 5I**). Thus, U21 has a limited but very specific effect on the global proteome and it uniquely downregulates STING via its C-terminal tail. Furthermore, it represents a dual-function effector that can inhibit both adaptive and innate immune responses by targeting MHC-I molecules and STING, respectively, with topologically distinct domains.

### **Functional annotation of a recently evolved non-canonical eORF: Adenoviral i-leader/13.6K proteins inhibit surface MHC-I expression via TAP transporter inhibition**

Our MHC-I surface display screen identified multiple known regulators of MHC-I display. These included HHV7 U21 characterized above, CMV US11 and US2 proteins that dislocate MHC-I molecules from the ER to the cytosol for proteasomal degradation,<sup>91,92</sup> and KSHV MIR2 E3 ligase that ubiquitinates and degrades MHC-I (**Figure 6A**).<sup>93</sup> However, the most prominent novel hits were 13.6K proteins from four distinct human adenovirus serotypes HAdV-A12, HAdV-B21, HAdV-D9, and HAdV-F41 (**Figure 6A**).

The 13.6K protein (also known as i-leader protein) is encoded on the opposite strand of the essential E2B polymerase by an open reading frame located between the second and third exons of the tripartite leader (TPL).<sup>94</sup> The TPL is a constant 5' leader sequence for all late adenoviral transcripts (**Figure 6B**), and some leader splicing events include the intervening leader (i-leader) with the open reading frame encoding 13.6K (**Figure 6B**).<sup>95</sup> 13.6K is predicted to have two transmembrane domains with both termini located on the cytoplasmic side (**Figure 6B**) but its function is unknown. Interestingly, the i-leader ORF is restricted to Old World monkey and hominid adenoviruses, as indicated by the absence of a start codon or the presence of in-frame stop codons in the i-leader ORF in other lineages (**Figure 6C**). Thus, 13.6K represents a recently evolved viral protein with an unknown function.

We validated that 13.6K proteins from the four human adenovirus strains identified in the screen downregulated MHC-I display in A549 and HeLa cells when expressed individually fused to C-terminal GFP (**Figure 6D** and **6E**). We also cloned the ORF from two additional serotypes (HAdV-C5, HAdV-E4) that were not included in the original screen. Both HAdV-C5 and HAdV-E4 13.6K robustly downregulated surface MHC-I, indicating that this is a conserved function of the protein (**Figure 6F**). 13.6K also inhibited surface display of  $\beta 2$  microglobulin, a constant subunit of all MHC class I molecules (**Figure S6A** and **S6B**). However, HAdV-F41 13.6K expression had no effect on the surface levels of two other plasma membrane proteins EGFR and CD47, suggesting that the effect is specific to MHC-I rather than 13.6K generally interfering with trafficking to the plasma membrane (**Figure S6C**). Moreover, 13.6K proteins were significantly more effective in MHC-I downregulation than the previously characterized adenoviral immunoevasin E3 19K from HAdV-B21 or HAdV-D9 (**Figure 6G**).<sup>96</sup>

13.6K could reduce MHC-I surface levels by interfering with membrane trafficking or by inducing MHC-I degradation. To differentiate between these options, we stained 13.6K-GFP or RLuc-GFP expressing HeLa cells with a pan-MHC-I antibody. In control cells expressing RLuc-GFP, MHC-I was localized to the plasma membrane as expected (**Figure 6H**). In contrast, in 13.6K expressing cells, there was no plasma membrane staining of MHC-I. Instead, we detected MHC-I in the endoplasmic reticulum together with 13.6K-GFP (**Figure 6H**). Moreover, 13.6K proteins did not affect total MHC-I levels as assessed by western blotting (**Figure 6I**), indicating that they are ER-associated membrane proteins that inhibit MHC-I membrane trafficking rather than stability (**Figure 6I**).

To understand how 13.6K disrupts antigen presentation, we characterized the interactomes of HAdV-D9 and HAdV-F41 13.6K in HeLa cells with AP-MS and BioID. We tagged them at the C termini with UltraID-3xFLAG and used HHV7 U21 as a positive control for an effector that downregulates MHC-I, and EGFP as a negative control. As previously reported, HHV7 U21 associated with HLA-A, HLA-B and HLA-C by both BioID and AP-MS, whereas 13.6K proteins did not significantly associate with these proteins (**Figure 6J** and **Table S5**). In contrast, HAdV-D9 and HAdV-F41 13.6K associated with the heterodimeric TAP peptide transporter complex subunits TAP1 and TAP2 in both BioID and AP-MS (**Figure 6J**). Endogenous TAP1 also co-immunoprecipitated with GFP-tagged HAdV-D9 13.6K, validating the interaction (**Figure 6K**). The TAP complex transports cytosolic peptides to the

ER for loading them onto MHC-I molecules as part of the peptide loading complex (PLC). Many viral effector proteins are known to target TAP to inhibit MHC-I loading and trafficking to the membrane,<sup>97</sup> making the TAP complex a strong candidate as the cellular target of 13.6K proteins.

To test if 13.6K proteins could directly inhibit TAP peptide transporter activity, we expressed EGFP-tagged 13.6K proteins from four serotypes (A12, B21, D9, and F41) in HeLa cells by lentiviral infection and assessed TAP activity with a fluorescent peptide transporter assay.<sup>98</sup> We permeabilized cells expressing GFP-tagged 13.6K or Renilla luciferase with a mild detergent and incubated them with a fluorescently labeled TAP substrate peptide RRYQNSTC(AF647)L in the presence of ATP.<sup>99</sup> In these conditions, TAP can transport the substrate peptide into the ER, which can be detected by flow cytometry. All 13.6K proteins significantly inhibited TAP activity (**Figure 6L**), establishing that they represent a novel family of viral TAP inhibitors.

## DISCUSSION

### The eORFeome as a discovery platform

The eORFeome constitutes a versatile discovery platform that bridges pathogen genomics and functional biology. By combining a comprehensive, sequence-diverse library of almost 4,000 viral, bacterial, and parasitic effector ORFs with barcoded, inducible expression and multiplexed phenotypic screening, this platform enables systematic mapping of effector activities across diverse host pathways. Unlike CRISPR-based perturbations (CRISPR-KO, CRISPRi, and CRISPRa), which rely on the one-dimensional modulation of endogenous gene activity within a single species, the eORFeome assays the functional potential of thousands of exogenous proteins optimized by evolution to rewire host processes in diverse, potentially multidimensional ways. This approach captures mechanisms inaccessible to genetic perturbation of host genes alone and complements CRISPR-based strategies by identifying exogenous regulators that mimic, oppose, or reveal latent host regulatory principles. Thus, eORFs exploit, as *evolved perturbagens*, a wide variety of cellular target proteins and protein-protein interfaces critical for host cell function and central to the modulation of host cell regulatory networks. The modular design of the eORFeome screening platform, encompassing pathway-specific reporters, survival-based selections, and scalable barcoded

readouts, makes it adaptable to virtually any cellular process that can be measured at population scale.

### **General lessons from diverse eORFeome screens**

Functionally, the eORFeome screens illuminate general principles of pathogen-host interaction and, more broadly, of protein evolution and function. First, they establish at a systematic scale that entirely unannotated proteins, often with no structural or sequence similarity to known proteins, can execute highly specific and potent cellular functions. Second, many effectors are multifunctional, influencing distinct pathways through separable domains or interfaces, exemplified by herpesviral U21, which employs distinct regions to inhibit both adaptive immunity via suppressing MHC-I antigen presentation and innate immunity by inducing STING degradation. Third, closely related homologs can diverge sharply in activity, such as the high-risk versus low-risk HPV E6 proteins with distinct abilities to inhibit p53. Yet despite this divergent, rapidly evolving sequence space, diverse effectors recurrently converge on a limited number of host pathways and proteins, highlighting evolutionary “hotspots” of host vulnerability and suggesting that the host proteome might impose a finite set of exploitable interfaces.

### **Specific findings from diverse eORFeome screens**

Among the specific discoveries, our work identifies (i) U14 from HHV6A as a p53 antagonist that sequesters p53 without degrading it, revealing a mode of inactivation analogous to the unrelated SV40 large T antigen; (ii) the herpesviral U21 as a dual-function effector that blocks both antigen presentation and innate DNA sensing, via separable structural modules; and (iii) adenoviral 13.6K proteins as a novel family of TAP inhibitors that suppress MHC-I surface expression. The newly uncovered role of 13.6K proteins is a particularly illuminating example of the function and evolution of pathogenic effectors. 13.6K protein is expressed at high levels during infection,<sup>100</sup> but no function had been identified. The ability of adenoviruses to evade the adaptive immune system has so far been attributed to the E3 19K immunoevasin that has been shown to interact with and inhibit the surface trafficking of MHC-I molecules. However, human adenovirus A, F, and G species lack E3 19K, leading to the suggestion that these species might evade host immunity by other means and targets.<sup>101</sup> In contrast, 13.6K is present in all HAdV clades, is more abundant than E3 19K during infection,<sup>102</sup> and is more potent in downregulating MHC-I surface presentation (**Figure 6G**), suggesting that it might be a more critical immunoevasin for adenoviral fitness. Strikingly,

the 13.6K protein emerged during primate adenovirus evolution by a combination of a novel alternative splicing event in the tripartite leader, yielding the i-leader exon, and the acquisition of an ORF in the opposite strand but in the same reading frame as the highly essential E2B polymerase. This remarkable evolutionary path highlights how pathogens can evolve novel functional proteins *de novo* even within exceptionally highly constrained sequences. More generally, our findings not only define new effector classes and mechanisms but also extend the conceptual framework of viral evolution, showing how new ORFs can emerge through splicing innovation and quickly acquire immunomodulatory activity.

### **Outlook and vision: a novel way of discovery**

Our work defines a new paradigm for protein function discovery, one that treats pathogen effectors from across kingdoms as an evolutionary library of host modulators and functional building blocks, i.e. *evolved perturbagens*. By studying individual eORFs outside their pathogenic contexts, the approach decouples effector biology from disease models, enabling systematic discovery of molecular functions in a controlled and safe setting across evolutionarily unrelated pathogens. This cross-pathogen perspective not only accelerates annotation of previously uncharacterized proteins but also generates foundational data for using eORFs as molecular tools and for training emerging sequence-to-function models. Moreover, it provides an inspiration for engineering minimal domains or synthetic proteins that modulate host pathways with precision for next-generation therapeutics, pathway engineering, and beyond.

In the longer term, coupling large-scale functional genomics with eORF collections that span highly divergent, unannotated protein sequences will enable direct mapping from sequence to function and sequence-to-function protein models. Such data-driven modelling will establish a route toward predictive and design-based biology, where the landscape of natural effector mechanisms serves as a training ground for understanding, predicting, and ultimately engineering cellular control.

### **ACKNOWLEDGEMENTS**

We thank Tanisha Teelucksingh and Mandy Lam from the Taipale lab and Kevin Sabath, Kavya Shetty, Michaela Pagani, and Katharina Bergauer from the Stark lab for experimental help and suggestions during the project. We are also indebted to Samantha Gruenheid and Lindsay Burns (McGill University, Montreal), John Brumell (SickKids, Toronto), Marc

Vidal, David Hill, and Michael Calderwood (DFCI, Boston), Fritz Roth (University of Pittsburgh), and Robert Kozak (Sunnybrook Research Institute, Toronto) for sharing effector ORF plasmids or pathogen DNA for the construction of the eORFeome. We also thank Alexei Savchenko and Deepak Patel (University of Calgary) for their advice on Legionella RavE and Victor Manon and Jue Chen (Rockefeller University) for discussions on the TAP complex.

We thank Cassandra Wong of the Network Biology Collaborative Centre Proteomics Facility at the Lunenfeld-Tanenbaum Research Institute for mass spectrometry analysis. The facility is supported by the Canada Foundation for Innovation and the Ontario Government. Tomas Pachano was supported by the EMBO (906-2022) and VIP-2 Postdoctoral Fellowships, He Leng by a CIHR Postdoctoral Fellowship (FRN 187814), Guillaume Dugied by the EPIC Convergence Postdoctoral Fellowship Award, and Yeojin Lee by the CIHR REDI fellowship (#ED6-190718). Research in the Stark group is supported by Boehringer Ingelheim GmbH, the Austrian Research Promotion Agency (FFG, FO999902549), the Austrian Science Fund (10.55776/P36971, 10.55776/PAT3564423) and the WWTF (10.47379/LS24012). For the purpose of Open Access, the author has applied a CC-BY public copyright license to any Author Accepted Manuscript version arising from this submission. The Taipale group was supported by a Pathway Grant by the Temerty Faculty of Medicine of the University of Toronto and the Canada Research Chairs Program. The Schramek group was supported by a Terry Fox Research Institute Program Projects Grant (TFRI Project #1107), a Project Grant from the Canadian Institute of Health Research (CIHR PJT-496918) and the Canada Research Chairs Program. The Weirauch and Kottayan groups were supported by National Institutes of Health (NIH) R01 HG010730, R01 NS099068, R01 AI024717, and R01 AI148276.

## STAR\* METHODS

### Key RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Alexa Fluor™ 647 anti-human EGFR Antibody	BioLegend	352917
Alexa Fluor™ 647 anti-human HLA-A,B,C Antibody	BioLegend	311414
Anti-IFIT1 antibody [EPR27276-63]	Abcam	ab305301
APC Annexin V	BioLegend	640920
APC anti-human CD47 Antibody	BioLegend	323123
APC anti-human HLA-A,B,C, Clone W6/32	BioLegend	311410; RRID: AB_314879
APC anti-human $\beta$ 2-microglobulin	BioLegend	395712
Goat anti-Mouse IgG Alexa Fluor™ 488	Invitrogen	A32723
Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 647	Invitrogen	A-21245
Goat anti-Rabbit IgG Alexa Fluor™ 594	Invitrogen	A-11012
H3	Abcam	ab1791
Hoechst 33342	Sigma-Aldrich	B2261
Human TruStain FcX	BioLegend	422302; RRID: AB_2818986
IKK	Cell Signaling Technology	8943
IKK $\alpha$	Cell Signaling Technology	11930
IRF-3 Antibody	Cell Signaling Technology	4302
I $\kappa$ B $\alpha$	Cell Signaling Technology	4814
MHC Class I Recombinant Rabbit Monoclonal Antibody	Invitrogen	MA5-53005
Monoclonal ANTI-FLAG(R) M2-Peroxidase (HRP) antibody	Sigma-Aldrich	A8592, RRID:AB_439702
Monoclonal Anti-HA	Sigma-Aldrich	H3663; RRID:AB_262051
Mouse Anti-GFP	Roche	11814460001; RRID:AB_390913
Mouse Anti-HSP90alpha/beta (F-8)	Santa Cruz Biotechnology	sc-13119, RRID:AB_675659
NF- $\kappa$ B p65	Cell Signaling Technology	8242
p2	Santa Cruz Biotechnology	sc-6246
p53	Santa Cruz Biotechnology	sc-126
Phospho-IRF-3 (Ser396)	Cell Signaling Technology	4947
Phospho-NF- $\kappa$ B p65 (Ser536)	Cell Signaling Technology	93H1
Phospho-STING (Ser366)	Cell Signaling Technology	19781
Phospho-TBK1/NAK (Ser172)	Cell Signaling Technology	5483
STING (D2P2F) Rabbit mAb	Cell Signaling Technology	13647
TAP-1 antibody	ProteinTech	11114-1-AP
TBK1/NAK Antibody	Cell Signaling Technology	3013
V5-tag (mouse)	Thermo Fisher	R960-25
V5-tag (rabbit)	abcam	ab309485
$\beta$ -Actin Antibody	Cell Signaling Technology	4967
<b>Bacterial and virus strains</b>		
ccdB-resistant MegaX <i>E. coli</i>	homemade	
DH5 $\alpha$ competent Cells	homemade	
MegaX <i>E. coli</i>	homemade	
NEB® 10-beta Electrocompetent <i>E. coli</i>	NEB	C3020K
Stbl3™ Chemically Competent <i>E. coli</i>	homemade	
<b>Chemicals, peptides, and recombinant proteins</b>		
2',3'-cGAMP sodium	MCE	HY-100564A
Bafilomycin A1	InvivoGen	inh-bfa
Chloroquine	InvivoGen	tlrl-chq-4
Clarity Western ECL Substrate	Bio-Rad	#1705061
DAPI	Sigma-Aldrich	D9542-1MG
diABZI (compound 3)	InvivoGen	tlrl-diabzi-2
Dimethyl sulfoxide (DMSO)	Fisher	D128-500
DMEM	homemade	
Doxycycline (dox)	Sigma-Aldrich	D9891
Fetal Bovine Serum	Gibco	S0615
L-glutamine	Gibco	25030081
Laemmli SDS sample buffer, reducing (4X)	Thermo Fisher	J60015.AC

Lenti-X concentrator	Takara	631232
Lipofectamine $\square$ 3000 Transfection Reagent	Invitrogen	L3000008
MEM Non-Essential Amino Acids Solution (100X)	Thermo Fisher	11140035
Opti-MEM I Reduced Serum Medium	Gibco	31985062
Penicillin-streptomycin	Gibco	15140122
Protease Inhibitor Cocktail	Roche	11836153001
Puromycin Dihydrochloride	Gibco	A1113803
Recombinant human IL-1 $\beta$ protein	Invivogen	rcyc-hil1b
Red ANTI-FLAG $\square$ M2 Affinity Gel	Sigma-Aldrich	F2426, RRID: AB_2616449
Roti-Phenol/Chloroform/Isoamyl alcohol	Thermo Fisher	FD1464
Staurosporine	Sigma-Aldrich	569397
TNF alpha	Bio-Techne	210-TA-005
Trypan Blue Stain (0.4%)	Invitrogen	T10282
<b>Critical commercial assays</b>		
Dynabeads M-280 Protein G	Thermo Fisher	10003D
GFP-Trap Magnetic Agarose beads	Chromotek	gtma
iTaq <sup>TM</sup> Universal SYBR $\square$ Green One-Step Kit	Bio-rad	1725150
KAPA HiFi HotStart ReadyMix	Roche	KK2602
Monarch Spin DNA Gel Extraction Kit	NEB	T1120
Monarch Spin PCR & DNA Cleanup	NEB	T1135
QIAamp DNA Mini Kit	Qiagen	51304
RNeasy Plus Mini Kit	Qiagen	74136
V5-Trap Magnetic Agarose beads	Chromotek	v5tma
<b>Experimental models: Cell lines</b>		
A549	ATCC	CRL-185
HEK293T (female) parental	ATCC	CRL-3216
HeLa	ATCC	CCL-2
LentiX, female	Clontech	632180
U937		
<b>Experimental models: Organisms/strains</b>		
FSF-KrasG12D (Krastm1Dsa)	Dieter Saur Lab	
Rosa-CreERT2 (Gt(ROSA)26Sortm1(cre/ERT2)Tyj/J)	Dieter Saur Lab	
<b>Deposited data</b>		
MS data of BioID and AP-MS	This manuscript	[To be requested by reviewers]
Raw eORFeome screen sequencing files	This manuscript	[To be upload to GEO]
<b>Oligonucleotides</b>		
PCR primer sequences, see Table S6	This manuscript	
Real-time PCR primer sequences, see Table S6	This manuscript	
sgRNA and shRNA sequences, see Table S6	This manuscript	
<b>Recombinant DNA</b>		
pCMV-Eco	Cell Biolabs	
pCMV-VSV-G	Addgene	#8454
pCMVR8.74	Addgene	#22036
pPGK_EcoRec-Zeo	This manuscript	
pPGK-Slc7a1-ZeoR	This manuscript	
psPAX2	Addgene	#12260
pSTV6-TetO-ccdB-eGFP	Anne-Claude Gingras Lab	
pTRE_TurboID_FLAG	This manuscript	
pVSVG	Addgene	#12259
<b>Software and algorithms</b>		
Adobe Illustrator	Adobe	2021
Clustal-Omega	Sievers et al. <sup>103</sup>	<a href="https://www.ebi.ac.uk/jdispatcher/msa/clustalo">https://www.ebi.ac.uk/jdispatcher/msa/clustalo</a>
Custom analyses	This manuscript	<a href="https://github.com/vloubiere/git_eORFscreen">https://github.com/vloubiere/git_eORFscreen</a>
DESeq2 version 1.48.2	Love et al. <sup>104</sup>	<a href="https://bioconductor.org/packages/release/bioc/html/DESeq2.html">https://bioconductor.org/packages/release/bioc/html/DESeq2.html</a>
FlowJo	BD Biosciences	Version 10.8
Foldseek release 10-941cd33	van Kempen et al. <sup>60</sup>	<a href="https://github.com/steineggerlab/foldseek">https://github.com/steineggerlab/foldseek</a>
GraphPad Prism	GraphPad	Version 10.0
ImageJ	NIH	Version 1.49v
iTOL version 7 web server	Letunic and Bork <sup>105</sup>	<a href="https://itol.embl.de/">https://itol.embl.de/</a>
MMseqs2	Steinegger and Söding <sup>106</sup>	<a href="https://github.com/soedinglab/MMseqs2">https://github.com/soedinglab/MMseqs2</a>
<b>Other</b>		
Thermal Cycler C1000	Bio-Rad	
FACSAria III cell sorter (BD Biosciences)	BD Biosciences	
iQue Screener PLUS	Intellicyt	

### **Materials availability**

All plasmids and cell lines generated in this study are available from the authors upon request.

### **Data and code availability**

All raw sequencing will be deposited in the Gene Expression Omnibus (GEO). Mass spectrometry datasets has been deposited with the ProteomeXchange Consortium via the MassIVE repository (<https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp>) under identifier PXD069484 (BioID, token: 8AIWP6z4zZNH01iC), PXD069483 (FLAP-AP, token: 4jp66PQsf7IXJQAZ), PXD069487 (total proteome, token: jRIUb68t8CZ0610b). The datasets will be made public upon acceptance of the manuscript. The code used in this study can be found at [https://github.com/vloubiere/git\\_eORFscreen](https://github.com/vloubiere/git_eORFscreen). Any additional information required to reanalyse the data reported in this paper is available upon reasonable request.

## **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

### **Cell Lines**

A549 cells (ATCC, CRL-185), HeLa (ATCC, CCL-2), HEK293T (female, ATCC CRL-3216) and Lenti-293T lentiviral packaging cells (female, Clontech, 632180) were cultured in Dulbecco's modified Eagle medium (Sigma-Aldrich) supplemented with 10% fetal bovine serum, 2 mM L-glutamine (Sigma-Aldrich) and 1× penicillin-streptomycin (Sigma-Aldrich).

U937 cells (gift from Jason Moffat lab) were cultured in RPMI supplemented with 10% FBS.

Cells were passaged every 2 days using TrypLE Express (Thermo Fisher, 12605010) for dissociation and maintained at <70% confluency. All cells were cultured at 37 °C in a humidified incubator with 5% CO<sub>2</sub>. All the cell lines were determined negative for mycoplasma. Cells were used for experiments within 10 passages from thawing.

### **Animals**

Animal husbandry, ethical handling of mice, and all animal work were carried out according to the guidelines approved by the Canadian Council on Animal Care and under protocols approved by the Centre for Phenogenomics Animal Care Committee (18-0272H). The animals used in this study were Rosa-CreERT2 (Gt(ROSA)26Sortm1(cre/ERT2)Tyj/J, in C57/Bl6 background)<sup>79</sup> and FSF-KrasG12D (Krastm1Dsa, in FVBN background),<sup>78</sup> kindly provided by Dr. Dieter Saur at Technische Universität München, Munich, Germany. Genotyping was performed by PCR using genomic DNA prepared from

mouse ear punches. For CreERT2-injected mice experiments, intragastric injections of approximately 75 mg tamoxifen/kg body weight (a standard dose of 50  $\mu$ L tamoxifen/corn oil solution) were administered on postnatal days (P) 4 and 5.

## **METHOD DETAILS**

### **Virus production**

Lenti-X cells were co-transfected with lentiviral plasmids, pCMVR8.74 helper (Addgene plasmid no. 22036) and pCMV-VSV-G (Addgene plasmid no. 8454) or pCMV-Eco (Cell Biolabs) envelope plasmids using polyethylenimine (PEI) transfection (MW 25,000, Polysciences) as previously described.<sup>107</sup> Virus containing supernatant was clarified by centrifugation and filtered through a 0.45  $\mu$ m filter. Target cells were infected at limiting dilutions in the presence of 4  $\mu$ g/mL of polybrene.

### **Generation of A549 Cells Expressing the Ecotropic Receptor**

To enable screening of the Lenti-eORFeome library in human cells under Biosafety Level-1 (BSL-1) conditions, we engineered the A549 cell line to express the mouse-specific ecotropic receptor, Slc7a1. This strategy renders the cells permissive to mouse ecotropic-pseudotyped lentivirus while ensuring that the viral particles cannot infect receptor-negative human cells, which minimizes risks of vector instability and the generation of replication-competent lentivirus.<sup>108,109</sup> To achieve this, the lentiviral plasmid pTP36\_EcoRec-Zeo was constructed to express Slc7a1 under the control of the PGK promoter, with a downstream internal ribosome entry site (IRES) driving a Zeocin resistance gene. A549 cells were then transduced with VSV-g pseudotyped lentivirus produced from this plasmid. Following selection with Zeocin, a clonal cell line, designated A549-Eco, was established by sorting single cells into 96-well plates using a FACSAria III cell sorter (BD Biosciences).

### **Generation of clonal reporter cell lines**

To generate reporter cell lines for NF- $\kappa$ B and p53 signalling, two lentiviral constructs were created. The NF- $\kappa$ B reporter construct contained four tandem copies of an NF- $\kappa$ B response element driving a minimal promoter and a destabilized enhanced green fluorescent protein (eGFP-mODC(d2)). Similarly, the p53 reporter construct contained 13 tandem copies of a p53 RE driving a minimal promoter and mCherry. Both plasmids included a PGK promoter driving a neomycin resistance gene for selection.

Ecotropic pseudotyped lentivirus was produced for each construct and used to transduce A549-Eco cells. Following selection with G418, populations were FACS-purified by sorting single cells with low basal fluorescence (eGFP or mCherry) into 96-well plates to establish clonal lines. Clones were subsequently validated for reporter functionality. The A549-Eco-NF- $\kappa$ B-eGFP clones were treated

with 20 ng/mL TNF $\alpha$ , and the clone with the highest eGFP induction was selected. Likewise, A549-Eco-p53-mCherry clones were treated with 2.5  $\mu$ M Nutlin, and the clone exhibiting the strongest mCherry induction was chosen for subsequent experiments.

### **eORFeome library**

Effector open reading frames as Gateway-compatible entry clones were sourced from previously published pathogen-specific ORF collections<sup>35–38,110,111</sup> or ordered from DNASU Plasmid Repository (Tempe, AZ). Additional clones were from cDNA collections of viral genes<sup>112,113</sup> that were amplified with PCR and cloned into pDONR221 Gateway entry vector. Salmonella effector ORFs were amplified from Salmonella genomic DNA, provided by John Brumell (SickKids, Toronto). E. coli and C. rodentium effectors were amplified from genomic DNA provided by Lindsay Burns and Samantha Gruenheid (McGill University, Montreal). Toxoplasma effectors were cloned by RT-PCR from Toxoplasma total RNA provided by Sebastian Lourido (Whitehead Institute, Cambridge, MA).

### **Generation of a *Coxiella burnetii* Gateway library**

A library of *Coxiella burnetii* effector sequences was cloned into pDONR221 as previously described<sup>35–37</sup>. Briefly, 147 sequences were selected from the C. burnetii Nine Mile Phase II RSA439 genome based on their homology to eukaryotic domains, known effectors, or direct observation of translocation.<sup>114–116</sup> Each ORF was PCR amplified using primers with flanking attB sites, with each product including an in-frame ATG-start site and an amber stop codon (F: ggggacaagttgtacaaaaagcaggcttcATG...; R: ggggaccactttgtacaagaaagctgggtcATC...). Each was then cloned into the pDONR221-ccdB vector (Invitrogen) using Gateway BP clonase II (Invitrogen) per manufacturer's instructions and sequence verified.

### **Generation of barcoded backbone**

We developed the pCW57.3-TRE3G destination vector from pCW57.1 (Addgene #41393) by replacing the TET promoter with a TRE3G promoter to minimize leakage expression of pathogen effectors in our library.

The library backbone was prepared by digesting the pCW57.3 plasmid with AgeI-HF and Sali-HF (NEB). The digestion products were separated by agarose gel electrophoresis, and the resulting vector backbone and insert template fragments were excised and purified using the Monarch DNA Gel Extraction Kit (NEB). The vector backbone was subjected to two additional rounds of cleanup using the Monarch PCR & DNA Cleanup Kit (NEB).

The insert template was barcoded using a two-step PCR strategy. First, a 50-cycle linear PCR was performed with a forward primer containing a semi-random 30-bp barcode (NNNNNWSWSWSWSWSNNNNNWSWSWSWSWS, where N is any base, S is G or C, and W is A or T; see Table S6 for primer details). This reaction used KAPA HiFi HotStart ReadyMix (Roche)

and was followed by a single fill-in cycle with a reverse primer to generate double-stranded DNA. The barcoded insert was then amplified for four cycles using primers that added the necessary overhangs for Gibson assembly. PCR products were purified at each step, first using an Exo-SAP (NEB) treatment followed by the Monarch PCR & DNA Cleanup Kit, and finally with AMPure XP beads (Beckman Coulter).

The barcoded insert was assembled into the prepared vector backbone using an in-house Gibson Assembly Master Mix at a 5-fold molar excess of insert to vector. The reaction was incubated at 50°C for 1 h, and the final assembled library was purified. The library was then transformed into electrocompetent ccdB-resistant MegaX E. coli (homemade) via electroporation. After recovery in SOC medium, the transformants were pooled and expanded in LB medium containing ampicillin and chloramphenicol. Finally, the bacterial culture was harvested, and the plasmid library was purified using a Plasmid Plus Giga/Mega Kit (Qiagen).

### **Gateway cloning of eORFs into barcoded backbone**

Entry clones from the eORFeome collection were organized into 18 standardized subpools, each containing ~384 eORFs, and subcloned into the barcoded pCW57.3 (see above). LR recombination reactions were assembled in duplicate using 150 ng of each entry ORF subpool and 150 ng of destination vector with 1 µL Gateway LR Clonase II in a total volume of 5 µL. Reactions were incubated overnight at 25°C in TE buffer. On each of the subsequent two days, an additional 150 ng of destination vector and 1 µL of LR enzyme were added to each reaction, and the volume was adjusted to 5 µL with TE buffer.

LR products were electroporated into NEB 10-beta electrocompetent cells and plated on LB agar containing carbenicillin (100 µg/mL). Plates were incubated overnight at 30°C. Colony counts confirmed a coverage of ~200 fold for each eORF. Colonies were pooled, resuspended in SOC on ice, pelleted, and plasmid DNA was prepared using an endotoxin-free midiprep kit. For downstream lentivirus production, structural and non-structural effectors were maintained as two separate libraries (eORFeome-L1 and eORFeome-L2).

### **Illumina sequencing of the eORF-barcode dictionary**

The eORFeome plasmid library (1 µg) was first linearized by digestion with I-SceI (NEB) and subsequently purified using the Monarch PCR & DNA Cleanup Kit (NEB).

To assemble the Tn5 transposome, equimolar amounts of Tn5ME-fw and Tn5MErev oligonucleotides were annealed. The resulting adapters were then incubated with an in-house Tn5 transposase at room temperature for 30 min. The tagmentation reaction was performed by incubating 12 ng of the linearized plasmid library with the assembled Tn5 transposome in a 2x tagmentation buffer at 55°C

for 8 min. The reaction was immediately stopped and purified using AMPure XP beads (Beckman Coulter).

The tagged DNA was size-selected by separating the products on a 2% agarose gel and excising fragments larger than 150 bp. DNA was extracted using the Monarch DNA Gel Extraction Kit. This size-selected library was then used as a template for PCR amplification using KAPA HiFi HotStart ReadyMix (Roche) with a primer pair where one primer was biotinylated.

The biotinylated PCR products were captured on M280 streptavidin beads (Thermo Fisher Scientific). The bead-bound DNA was then used as a template for a final PCR step to add Illumina sequencing adapters. The final library was subjected to two consecutive rounds of gel extraction, isolating fragments between 150 and 300 bp. The library was given a final purification with the Monarch PCR & DNA Cleanup Kit before being submitted for Illumina paired-end sequencing.

## Screens

### Pooled eORFeome screens

The pooled-eORFeome library was generated by packaging into ecotropic pseudotyped lentivirus through the transfection of Lenti-X cells with PEI. Viral supernatant was subsequently cleared of cellular debris by filtration through a 0.45- $\mu$ m PES filter. Target cells, including A549-Eco-NF- $\kappa$ B-eGFP, A549-Eco-p53-mCherry, A549-Eco, or U937-Eco lines, were then transduced at a multiplicity of infection (MOI) below 0.2 to ensure single-copy integration while maintaining a library representation of at least 10,000-fold. Twenty-four h post-transduction, cells were selected with 1  $\mu$ g/mL puromycin until a non-transduced control plate of cells was eliminated, a process that typically required 4–5 days, during which cells were washed twice daily with 1 $\times$  PBS and received fresh medium. Following selection, eORF expression was induced with 1  $\mu$ g/mL dox for 24 h.

Screen-specific protocols were then implemented. For the reporter-based screens, the NF- $\kappa$ B activation screen involved collecting a baseline sample of 15 million cells as the input control before the remaining cells were sorted for the top 5% of the GFP-positive population using a BD FACSAria III sorter. Conversely, for the inhibition screens, cells were first treated with either 20 ng/mL TNF $\alpha$  (NF- $\kappa$ B inhibition) or 2.5  $\mu$ M Nutlin (p53 inhibition) for 16 h. After this treatment, a 15 million cell input sample was collected, and the remaining cells were sorted to isolate the GFP-low or mCherry-low populations (bottom 5%), respectively. For the cGAS-STING inhibition screen, transduced U937 cells were treated with 10  $\mu$ g/mL cGAMP, resuspended in PBS with 2% FBS containing FVD-780 viability dye (1:1000), and incubated at 4°C for 30 min in the dark. After two washes, the pellet was fixed using the BD Cytofix/Cytoperm™ Fixation Buffer, followed by incubation at 4°C for 20 min. Cells were then washed twice with 1 $\times$  BD Permeabilization Buffer and incubated overnight at 4°C with an anti-IFIT1 primary antibody (abcam, ab305301, 1:500). The following day, cells were washed twice and incubated with an Alexa Fluor™ 647-conjugated goat anti-rabbit secondary antibody (Invitrogen, A-21245; 1:300) for 30 min at 4°C. After two final washes, cells were resuspended in

FACS sorting buffer for analysis. In the MHC-I inhibition screen, transduced A549 cells were first incubated with human TruStain FcX (BioLegend, 422302; 1:200) for 10 min at room temperature to block non-specific binding, followed by staining with an APC-conjugated anti-HLA-ABC antibody (BioLegend, 311410; 1:200) for 40 min at 4°C. Following three washes in FACS buffer, the cells were subjected to FACS. For the apoptosis inhibition screen, transduced A549-Eco cells were split post-induction into a control group maintained in dox and a treatment group cultured with both dox and 0.1 µM etoposide. These populations were maintained for 10 days and were split at a 1:4 ratio upon reaching confluency. Similarly, for the cGAMP survival screen, transduced U937 cells were initially treated with 10 µg/mL cGAMP, with the concentration increasing by 1 µg/mL every two days over a 10-day period. To select for strong suppressors, the screen concluded with the addition of 50 µg/mL cGAMP for 2 days. Finally, for every screen, genomic DNA (gDNA) was isolated from all collected cell populations (input, sorted or treated) for subsequent enrichment analysis. All screens were performed in at least two independent biological replicates.

### **DNA Isolation, Barcode PCR/eORFeome PCR, and Sequencing**

gDNA was isolated by lysing cell pellets overnight at 55°C in a buffer containing 10 mM Tris-HCl (pH 8.0), 5 mM EDTA, 100 mM NaCl, 1% SDS, and 0.5 mg/mL proteinase K. The lysate was then treated with RNase A (Qiagen) for 2 h at 37°C. Subsequently, gDNA was purified through two sequential extractions with phenol:chloroform:isoamyl alcohol followed by a final extraction with chloroform:isoamyl alcohol.

Sequencing libraries were prepared using one of two strategies, dictated by whether the eORF sequence itself or its associated barcode was amplified for quantification. For the MHC-I and cGAS-STING inhibitor screens, the eORF sequences were directly amplified from gDNA and processed as previously described.<sup>117,118</sup> For all other screens, eORF-associated barcodes were recovered from gDNA. To this end, the entire gDNA yield from sorted/selected samples and 40 µg of gDNA from input samples were used as templates. To maintain library complexity, 40 individual PCR reactions were performed for each sample. Amplification was carried out for 25 cycles (65°C annealing) using KAPA HiFi HotStart ReadyMix (Roche). The primers were designed to anneal to the Illumina adaptor sequences directly flanking the barcode region. To enable sample demultiplexing, these primers included Illumina i5 and idx index sequences as 5' overhangs.

Following amplification, all PCR reactions for a given sample were pooled. The pooled amplicons were concentrated and purified using the Monarch PCR & DNA Cleanup Kit (NEB), using six columns for each input sample and one column for each sorted/selected sample. The final purified barcode amplicons were gel-extracted and submitted for next-generation sequencing.

### **Hit validations**

To validate individual hits, candidate eORFs were subcloned from entry vectors into the barcoded pCW57.3 (for hits from the NF- $\kappa$ B or p53 screens) or pSTV6-eORF-C-GFP-Puro lentiviral plasmid (for hits from the MHC-I or cGAS-STING screens) via Gateway cloning. Lentivirus for each candidate was produced (see above) and used to transduce the appropriate cells in the presence of 4  $\mu$ g/mL Polybrene. eORF-expressing cell populations were selected with 1  $\mu$ g/mL puromycin. For functional validation, each candidate was tested in parallel with and without dox. To identify activators, cells were cultured for 24 h before analysis. To identify inhibitors, cells were cultured for 24 h (with or without dox) and then stimulated for 16 h with the relevant pathway agonist. In all validation experiments, cells expressing a control eORF (mCherry, eGFP or NanoLuc) were included as a negative control. Reporter fluorescence was quantified by flow cytometry using iQue Screener PLUS (Intellicyt).

### **RNA extraction and quantitative PCR**

48 h post-induction, U937 cells expressing U21-HHV7 or NanoLuc, alongside uninduced control cells, were harvested. Total RNA was extracted using the RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol. Subsequently, 1  $\mu$ g of total RNA was reverse transcribed using the iScript cDNA Synthesis Kit (Bio-Rad). Gene expression levels were normalized to the housekeeping gene *ACTB*. All primer sequences are listed in Table S5.

### **Western Blots**

Cells were lysed directly in Laemmli buffer containing 10%  $\beta$ -mercaptoethanol and boiled at 98°C for 5 min. Protein lysates were resolved by SDS-PAGE on 4–15% Mini-PROTEAN TGX Precast Gels (Bio-Rad) and transferred to Immobilon-P PVDF membranes (Merck Millipore) using a wet-transfer system. Membranes were probed with the following primary antibodies: anti-V5-tag (Thermo Fisher, R960-25, 1:1,000), anti-IKK $\alpha$  (Cell Signaling Technology, 11930), anti-IKK $\beta$  (Cell Signaling Technology, #8943), anti-Phospho-NF- $\kappa$ B p65 (Ser536) (Cell Signaling Technology, 93H1), anti-I $\kappa$ B $\alpha$  (Cell Signaling Technology, 4814), anti-NF- $\kappa$ B p65 (Cell Signaling Technology, 8242), anti-p53 (Santa Cruz Biotechnology, sc-126), anti-p21 (Santa Cruz Biotechnology, sc-6246), anti-human EGFR (BioLegend, #323123), anti-MHC Class I (Invitrogen, MA5-53005), anti-STING (D2P2F) (Cell Signaling Technology, 13647), anti-Phospho-STING (Ser366) (Cell Signaling Technology, 19781), anti-TBK1/NAK (Cell Signaling Technology, 3013), anti-Phospho-TBK1/NAK (Ser172) (Cell Signaling Technology, 5483), anti-IRF-3 (Cell Signaling Technology, 4302), anti-Phospho-IRF-3 (Ser396) (Cell Signaling Technology, 4947), anti-GFP (Roche, 11814460001; RRID:AB\_390913), anti-HA (Sigma-Aldrich, H3663; RRID:AB\_262051), anti-H3 (Abcam, ab1791), anti- $\beta$ -Actin (Cell Signaling Technology, 967) and anti-HSP90 $\alpha$ / $\beta$  (F-8) (Santa Cruz Biotechnology, sc-13119; RRID:AB\_675659) and anti-MHC-I (Invitrogen, MA5-53005).

Detection was performed using HRP-conjugated secondary antibodies (anti-mouse HRP, 7076; anti-rabbit HRP, 074; Cell Signaling Technology, 1:10,000) and Clarity Western ECL Substrate (Bio-Rad). Blots were imaged using a ChemiDoc Imaging System running Image Lab software (Bio-Rad).

### **Immunofluorescence**

For U14 experiments, cells expressing empty vector, U14-V5, or the U14 D124R-V5 mutant were seeded in 96-well plates. To induce the expression of the V5-tagged constructs, cells were incubated with dox for 24 h, followed by a 1 h treatment with 1.5  $\mu$ M Nutlin. Following treatment, cells were seeded in 96-well plates, washed once with PBS, and fixed for 10 min at room temperature with 3.7% formaldehyde in PBS. Following fixation, cells were washed with cold PBS and permeabilized with 0.1% Triton X-100 in PBS for 10 min at room temperature. The cells were then washed three times with PBS. Non-specific binding was blocked by incubating the cells in 5% bovine serum albumin (BSA) in PBS for 30 min at room temperature. Subsequently, cells were incubated overnight at 4°C with primary antibodies diluted in 5% BSA/PBS. The primary antibodies used were mouse anti-p53 (Santa Cruz Biotechnology, sc-126) and rabbit anti-V5-tag (abcam, ab309485) for the detection of V5-tagged U14 or V5-tagged U14 D124R constructs. After primary antibody incubation, cells were washed three times with PBS containing 0.1% Tween-20. Cells were then incubated for 1 h at room temperature in darkness with secondary antibodies diluted 1:1000 in 1% BSA/PBS. The secondary antibodies were goat anti-mouse IgG conjugated to Alexa Fluor™ 488 and goat anti-rabbit IgG conjugated to Alexa Fluor™ 594. Following secondary antibody incubation, cells were washed three times with PBS. Nuclei were counterstained with DAPI for 5 min at room temperature. Finally, cells were rinsed with PBS before imaging. For quantification, the total number of cells for each cell type was determined by counting DAPI-stained nuclei. The number of cells containing p53 foci (Alexa Fluor™ 488) and V5 foci (Alexa Fluor™ 594) was then counted. The percentage of cells positive for each type of foci was calculated relative to the total cell count. This analysis was performed for three independent experiments, and the data are presented in graphs as the mean  $\pm$  standard deviation (SD). For 13.6k experiments, HeLa cells stably expressing 13.6k-GFP were seeded at 8,000 cells per well (96-well format). The following day, expression of 13.6k was induced using dox. 24h after induction, cells were fixed with 4% paraformaldehyde for 15 min and washed with PBS three times. MCHI staining was performed using anti-MCHI fused with AF647 antibody (1:20) for 1h at RT. Cells were then washed with PBS three times and incubated with Hoescht (1mg/ml) in 2x SSC for 10 min prior to imaging. Imaging was performed using an Opera Parking Elmer Phenix automated confocal microscope. Images were taken using a 40X water objective.

### **Luminescent Cell Viability Assay**

For cell viability assays, U937 cells were seeded into 96-well plates at a density of 2,000 cells per well. The cells were induced with 1  $\mu$ g/mL dox for 24 h, followed by stimulation with the indicated

STING agonists for an additional 24 h. To measure viability, 100  $\mu$ L of CellTiter-Glo reagent (Promega) was added to the 100  $\mu$ L of cell culture in each well. The plates were agitated on an orbital shaker for 5 min and then incubated for 5 min at room temperature to stabilize the signal. Luminescence was subsequently measured using a BioTek multimode microplate reader.

### **TAP transporter assay**

HeLa cells stably expressing C-terminal eGFP-tagged 13.6k constructs were seeded (100,000 cells per well, 48-well format) in a total volume of 500  $\mu$ L. The following day, expression of the effectors was induced by adding 1  $\mu$ g/mL of dox. 48h after induction, the cells were washed with PBS, prior to semi-permeabilization with 0.25 mg/mL saponin in PBS 1x for 15 min. The semi-permeabilization buffer was removed and 100  $\mu$ L of transport buffer solution of 10 mM ATP, 10 mM MgCl<sub>2</sub> and 1 nM fluorescence peptide in PBS 1x was added for 15 min. The reaction was stopped by adding 900  $\mu$ L of PBS 1x supplemented with 20 nM EDTA. The samples were directly analyzed by flow cytometry.

### **U14-V5 AP-MS**

For U14 experiments, A549-Eco cells were transduced at a high MOI with lentivirus expressing either U14-V5 or an Empty-V5 control and were subsequently selected with 1  $\mu$ g/mL puromycin. For each biological replicate (n=3), cells from a confluent 15 cm plate were washed twice with ice-cold PBS and harvested.

Cell pellets were lysed in IP lysis buffer (20 mM HEPES pH 7.3, 150 mM NaCl, 2 mM MgCl<sub>2</sub>, 0.25% NP-40, 10% glycerol, 0.3% Triton X-100) supplemented with cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche). Lysates were incubated on ice for 30 min and sonicated. The lysates were then clarified by centrifugation at 20,000 x g for 5 min at 4°C.

The cleared supernatants were incubated overnight at 4°C with rotation alongside V5-Trap Magnetic Agarose beads (Chromotek) that had been pre-washed with IP lysis buffer. The following day, the beads were washed three times for 10 min each with IP lysis buffer, followed by four 5-minute washes with a no-detergent buffer (20 mM Tris pH 7.5, 130 mM NaCl). The final washed beads were submitted to the Vienna BioCenter Core Facilities (VBCF) Proteomics facility for on-bead digestion and subsequent mass spectrometry analysis.

For 13.6k experiments, co-immunoprecipitation experiments were performed using HeLa stably expressing 13.6k-GFP constructs. 48h after induction with dox, the cells were washed with 1x PBS and lysed in NP40 lysis buffer (10mM Tris/HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA and 0.5% NP40) with protease inhibitors on ice for 30 min. The lysate was centrifuged at 15,000g for 15 min at 4°C. Proteins were immunoprecipitated using GFP-Trap beads (ProteinTech, gtma) at 4°C for 1 h. Beads were then washed 3 times using wash buffer (10mM Tris/HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA and 0.05% NP40) to remove non-specific interactions. The immunoprecipitated proteins were eluted from beads by adding loading buffer and heating at 95°C for 5 min. Samples were then loaded

on 4%-12% Bis-Tris PAGE and analyzed by Western Blotting on nitrocellulose membranes. Membranes were blotted with anti-GFP (11814460001, Roche, 1:1000) and anti-TAP1 (11114-1-AP, ProteinTech, 1:1000).

#### **U14 BioID**

For the BioID proximity labelling experiment, A549-Eco cells were transduced at a high MOI with lentivirus expressing either U14-TurboID or an Empty-TurboID control and were subsequently selected with 1  $\mu\text{g}/\text{mL}$  puromycin. For each biological replicate ( $n=3$ ), cells from a confluent 15 cm plate were treated with 500  $\mu\text{M}$  biotin for 10 min at 37°C. The cells were then washed twice with ice-cold PBS and harvested.

Cell pellets were lysed in RIPA buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100) supplemented with cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche) and 1 mM DTT. After a 1 h incubation at 4°C, lysates were clarified by centrifugation at 18,000  $\times$  g for 10 min.

The cleared supernatants were incubated with streptavidin magnetic beads (Dynabeads™ M-280 Streptavidin, Thermo Fisher) to capture biotinylated proteins. The beads were then subjected to a stringent series of washes: once with RIPA buffer, once with 2% SDS, once with Buffer 2 (50 mM HEPES pH 7.4, 500 mM NaCl, 1 mM EDTA, 0.1% sodium deoxycholate, 1% Triton X-100), once with Buffer 3 (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 250 mM LiCl, 0.5% sodium deoxycholate, 1% NP-40), and finally, four times with a non-detergent wash buffer (20 mM Tris-HCl, 100 mM NaCl). After the final wash, the beads were flash-frozen in liquid nitrogen and submitted to the proteomics facility for on-bead digestion and mass spectrometry analysis.

#### **Whole-cell proteomics and analysis by Tandem Mass Tagging-based proteomics**

For whole proteome analysis, we collected 2 million U937 cells expressing either wild-type U21-HHV7, the truncation mutant U21-HHV7 $\Delta$ tail or RLuc. Cells were lysed with 5% SDS, 50mM TEAB. Lysates were sonicated for 3 cycles of 5sec on, 3 sec off at 30% amplitude using a Sonicator with 1/8" microtip. 25  $\mu\text{g}$  of protein material was reduced at 20mM DTT for 10 min at 95C and alkylated with 40mM iodoacetamide for 30 min in the dark. Samples were brought up to final concentration of 5% SDS and phosphoric acid was added to a final concentration of 1.2%. 165  $\mu\text{l}$  of S-Trap protein binding buffer (90% methanol, 100mM TEAB) was added to 27.5  $\mu\text{l}$  of acidified lysate. Resulting mixture was passed through the micro column at 4000xg. The micro column was washed 4 times with the S-Trap protein binding buffer. Each sample was digested with 4  $\mu\text{g}$  of trypsin (in 20  $\mu\text{l}$  of 50mM TEAB) for 1hr at 47C. Prior to elution, 40  $\mu\text{l}$  of 50mM TEAB pH 8 was added to the column. Peptides were eluted by centrifugation at 4000xg. Peptides were eluted 2 more times with 40  $\mu\text{l}$  0.2% formic acid and 40  $\mu\text{l}$  of 50% acetonitrile+0.2% formic acid. Eluted peptides were dried down and stored at -40°C. For data-dependent acquisition (DDA) LC-MS/MS, 250ng protein

equivalent of digested peptides were analyzed using a nano-HPLC (High-performance liquid chromatography) coupled to MS. The sample was loaded onto Evotip Pure per manufacturer instructions. Peptides were eluted from the column (cat#: EV-1137, 15cmx150  $\mu$ m with 1.5  $\mu$ m beads) with the 30SPD pre-formed acetonitrile gradient generated by an Evosep One system and analyzed on a timsTOF Pro 2. The Evosep was coupled to the timsTOF Pro 2 using a 10  $\mu$ m diameter emitter tip. The column toaster was set to 40C. The total DIA protocol is 44 min. The MS1 scan had a mass range of 100-1700Da in dia-PASEF. TIMS settings were accumulation and ramp time of 100ms, and within the mobility range (1/K0) of 0.6 to 1.6V $\cdot$ s/cm<sup>2</sup>. Cycle time 2s. For MS2, 1 mobility window with 2 ramps was used for 32 mass windows, 29.8 Da wide with 5Da mass overlap. The mobility range was from 0.61/K0 to 1.451/K0. This was at a duty cycle of 100% and a ramp rate of 9.52Hz. 1+ ions are excluded from fragmentation using a polygonal filter. The auto calibration was off.

### **13.6K and U21 UltraID and AP-MS**

HAdV 13.6K and HHV7 U21 clones were transferred from the pDONR221 entry vector into a destination vector carrying a C-terminal UltraID-FLAG tag via Gateway recombination. Stable HeLa cell lines expressing the effector-UltraID-FLAG fusion proteins were then generated. Cells were grown to 70% confluence in 15 cm dishes before inducing gene expression with 1  $\mu$ g/mL dox for 24 h. 50  $\mu$ M biotin was then added to each plate for 30 min. Cell pellets were collected for downstream UltraID and AP-MS applications.

For biotinylation of UltraID analysis, the cells were treated with 50  $\mu$ M biotin (Sigma, B4501) for 1 hr, rinsed, and harvested and frozen in PBS buffer. UltraID (n=2) samples were prepared for MS analysis by the NBCC Proteomics Core at the Lunenfeld-Tanenbaum Research Institute (Toronto, Canada). Frozen cell pellets were lysed in 1:4 (cell pellet weight:lysis buffer) using RIPA-lysis buffer (50mM Tris-HCl (pH 7.5), 150mM NaCl, 0.1% (w/v) SDS, 1% NP-40, 1mM MgCl<sub>2</sub>, 1mM EGTA, 0.5mM EDTA, 0.4% (w/v) sodium deoxycholate, 1mM PMSF, and 1x Sigma protease inhibitors). The lysate was sonicated (3 x 5sec, 2 sec off) at 30% amplitude using a 1/8" microtip. 250 units TurboNuclease and 10  $\mu$ g RNase was added to each sample and incubated, with rotation, at 4°C for 30 min. Additional SDS was added to bring the final concentration to 0.4% SDS. Each sample was mixed well and centrifuged at 14,000rpm for 20 min at 4°C. 200  $\mu$ l lysate was applied to 20  $\mu$ l of pre-washed 50% slurry streptavidin beads (Cytiva 17-5113-01), rendered trypsin-resistant, and incubated at 4°C, rotating, for 3 h. After incubation, supernatant was removed and beads were moved to a new tube in RIPA-wash buffer (50mM Tris-HCl pH 7.4, 150mM NaCl, 1mM EDTA, 1% NP-40, 0.1% (w/v) SDS, 0.4% (w/v) sodium deoxycholate). Beads were washed one time with 2% SDS, 50mM Tris-HCl pH 7.5, two times with RIPA-wash buffer, one time with TNNE-wash buffer (25mM Tris-HCl pH 7.4, 150mM NaCl, 0.1% NP-40, 1mM EDTA), and three times with 50mM ammonium bicarbonate. Proteins on beads were digested with 0.5  $\mu$ g of trypsin in 50  $\mu$ l of 50mM ammonium

bicarbonate, overnight at 37°C. 0.25 µg trypsin was added to the peptides and incubated at 37°C for 3 h. Peptides were moved to a new tube, beads were washed with 50 µl water and this wash was combined into the same tube. Formic acid was added to a final concentration of 5%. The solution was centrifuged at 14,000rpm for 5 min, 80% of the solution was moved to a new tube. Peptides were dried down and stored at -40°C until mass spectrometry analysis.

For AP-MS, frozen cell pellets were lysed in 1:4 (cell pellet weight:lysis buffer) using lysis buffer (50mM HEPES-NaOH pH 8.0, 100mM KCl, 2mM EDTA, 0.1% NP-40, 10% glycerol, 1mM DTT, 1mM PMSF, and 1x Sigma protease inhibitors). The lysate was sonicated (3 x 5sec, 2 sec off) at 30% amplitude using a 1/8" microtip. 250 units TurboNuclease and 10 µg RNase was added to each sample and incubated, with rotation, at 4°C for 15 min. Each sample was mixed well and centrifuged at 14,000rpm for 20 min at 4°C. 400 µl lysate was applied to 25 µl of pre-washed 50% slurry anti-FLAG M2 magnetic beads (Sigma-Aldrich M8823) and incubated at 4°C, rotating, for 3 h. After incubation, supernatant was removed and beads were moved to a new tube in lysis buffer, with no PMSF, DTT, or protease inhibitors. Beads were washed one time with FLAG rinsing buffer (20mM Tris-HCl pH 8.0, 2mM CaCl<sub>2</sub>). The last wash was removed and proteins on beads were digested with 7.5 µl of trypsin (100 ng/µl in 20 mM Tris-HCl pH 8.0), overnight at 37°C. Peptides were moved to a new tube and 2.5 µl trypsin (same concentration) was added to the peptides and incubated at 37°C for 3 h. Formic acid was added to a final concentration of 5%. Peptides were dried down and stored at -40°C until mass spectrometry analysis.

## **Histology**

For H&E staining, the mice were euthanized 9 weeks after the induction of U14 or eGFP. The lungs were inflated with intratracheal injection of 4% PFA and fixed overnight in 4% PFA. The fixed lungs were transferred to 70% ethanol, embedded in paraffin, and sectioned at 5 µm for H&E staining at the Histology core at University Health Network (UHN), Toronto, Canada. For quantification of tumor area, NDP.view.2 software was used.

## **QUANTIFICATION AND STATISTICAL ANALYSIS**

### **Screen Data analysis**

After initial quality checks using fastqc, raw sequencing reads amplified from integrated eORFs after phenotypic selection (screen) or not (input) were trimmed using Trim Galore (v0.6.10) in order to remove adapter sequences and filter out low-quality bases.<sup>119</sup> Trimmed reads were then aligned to custom indexes that either directly contained eORF sequences (MHC-I and cGAS-STING screens) or a custom dictionary of non-ambiguous barcodes (all other screens) associated to each tested eORF (see Dictionary analysis) using Bowtie2 (v1.3.1).<sup>120</sup> Reads with low mapping quality – that could not be unambiguously aligned to a unique ORF/Barcode – were discarded using SAMtools (v1.22.1).

Finally, reads associated to each ORF/Barcode were counted using the Rsamtools package (v3.21). Resulting matrices were used to identify hits that were significantly enriched after phenotypic selection (screen) compared to unselected/unsorted (input) samples.

For MHC-I and cGAS-STING screens, where integrated eORFs were directly sequenced, hits were identified using DESeq2 (v1.38.3),<sup>104</sup> and only the eORFs with at least 10 reads in at least two of the compared conditions were considered. For all the other screens, where the abundance of eORFs was inferred from their associated BCs, MAGeCK (v0.5.9)<sup>49</sup> was used, with the following parameters: --sort-criteria 'pos' --norm-method median --remove-zero none. Of note, only the BC/ORFs combinations with at least 3 reads across all screen/input sample pairs were considered, after adding a pseudocount of 1.

The complete scripts detailing the analysis pipeline described above are publicly available at the following GitHub repository: [https://github.com/vloubiere/git\\_eORFscreen](https://github.com/vloubiere/git_eORFscreen).

### Dictionary analysis

To link eORFs to their corresponding barcodes, we used the eORF-Barcode dictionary which was prepared and sequenced as described in the earlier 'Illumina sequencing of the eORF-Barcode dictionary' section. This library consists of DNA fragments where each eORF is positioned at the 5' end and its associated barcode is at the 3' end (fragment structure: Tn5\_adaptor>eORF>AttB2>22nt\_spacer>illuminaFwd>30nt\_BC>illuminaRev. Of note, the Tn5 adaptor sequence was used for sequencing, not the illumina Fwd adapter which was only used for screening). Consequently, Read 1 captured the 3' end of the eORF sequence, while Read 2 captured the barcode sequence, and were used to identify the BCs that were unambiguously assigned to each eORF sequence

First, the 30nt barcode (BC) sequence was extracted from read 2 using cutadapt to trim the illuminaFwd and IlluminaRev adaptor sequences from the 5' and the 3' ends (5' trimming parameters: -g GCTCTTCCGATCT; 3' trimming parameters: -a AGATCGGAAGAGC -m 25 -M 30). Only barcode sequences that were between 25 and 30nt long and matched the expected sequence pattern  $([GC][AT])\{4\}[GCAT]\{5\}([GC][AT])\{5\}[GCAT]\{2\}$  were considered for downstream analyses.

To extract the eORF sequence, three strategies were used: (i) read1 were trimmed using cutadapt (5' trimming parameters: -g AAAAAAGTTGGCA; 3' trimming parameters: -a GCCCAACTTTCTT -m 25) or 2 (ii) read 1 were trimmed using Trim Galore<sup>119</sup> (parameters: --hardtrim5 30) or (iii) read 2 were trimmed using trim\_galore (parameters: --hardtrim3 30). eORF sequences were then aligned to the sequences present in the eORFeome library using bowtie2 (v1.3.1, default parameters)<sup>120</sup> and, for each read ID, only the best alignment (based on mapping quality was retained), and only the reads that were unambiguously aligned (mapping quality  $\geq 30$ ) were eventually retained.

Then, confident BC/eORF pairs were retrieved using the read ID, and only the BCs that were supported by at least three read counts and were systematically assigned to the same unique eORF

were considered. These barcode sequences were finally used to generate the custom Bowtie2 index used to align the sequencing results from the phenotypic screens.

The complete scripts detailing this analysis is publicly available at the following GitHub repository:

[https://github.com/vloubiere/git\\_eORFscreen](https://github.com/vloubiere/git_eORFscreen) .

### **Phylogenetic distribution of eORFs**

Each eORF was first assigned the taxonomic identifier (TaxID) of its source pathogen from the NCBI Taxonomy database (<https://www.ncbi.nlm.nih.gov/taxonomy>). These TaxIDs were then used as input to generate a phylogenetic tree based on the NCBI taxonomic hierarchy, which was exported in Newick format. The resulting tree was visualized using the Interactive Tree Of Life (iTOL) web server.<sup>105</sup>

### **Mass Spectrometry Analysis**

For data-dependent acquisition (DDA) LC-MS/MS, one-sixteenth of digested peptides were analyzed using a nano-HPLC (High-performance liquid chromatography) coupled to MS. The sample was loaded onto Evotip Pure per manufacturer instructions. Peptides were eluted from the Performance column (cat#: EV-1109, 8cmx150  $\mu$ m with 1.5  $\mu$ m beads), heated at 40C) with the 60SPD pre-formed acetonitrile gradient generated by an Evosep One system, and analyzed on a timsTOF Pro 2. The Evosep was coupled to timsTOF Pro 2 using a 20um diameter emitter tip. The column toaster was set to 40C. The total DDA protocol is 22 min. The MS1 scan had a mass range of 100-1700Da in PASEF mode. TIMS settings were accumulation and ramp time of 100ms (with 4 PASEF ramps and active exclusion at 0.4min), and within the mobility range (1/K0) of 0.85 to 1.3V·s/cm<sup>2</sup>. This was at a cycle time of 0.53s. The target intensity was set to 17,500 and intensity threshold set to 1750. 1+ ions are excluded from fragmentation using a polygonal filter. The auto calibration was off.

For Whole-cell proteomics analysis, MS Data Analysis with Spectronaut, Spectronaut v19 directDIA+ workflow was used to search the data with the Spectronaut generated Human spectral library (Human\_PDB\_2023). Parameters for the search were default. Differential abundance testing used was unpaired t-test.

For UltraID and AP-MS, Mass spectrometry data generated were stored, searched and analyzed using ProHits laboratory information management system (LIMS) platform. Within ProHits, MSFragger 4.1 was used to search the data using a FASTA database from UP000005640 human Uniprot proteome, no isoforms, with a custom list of contaminants and decoy appended. Acetylated protein N-term and oxidated methionine were set as a variable modifications. Precursor mass tolerance was set to 20ppm on either side. Fragment mass tolerance was set to 20ppm. Enzymatic cleavage was set to trypsin with 2 missed cleavages. MSBooster and Percolator were turned on. Percolator required a minimum probability of 0.5 and did not remove redundant peptides. The target-decoy competition method was

used to assign q-values and PEPs. For ProteinProphet, the maximum peptide mass difference was set to 30ppm. When generating the final report, the protein FDR filter was set to 0.01. FDR was estimated by using both filtered PSM and protein lists. Razor peptides were used for protein FDR scoring. All other parameters were default.

### **Homology clusters**

To group eORFs by sequence similarity, protein sequences were clustered using MMseqs2 v13.45111<sup>106</sup> across a range of identity thresholds (0.8, 0.5, 0.3, 0.15 and 0.1).

### **UniProt annotation scores**

UniProt annotation scores for eORFs were retrieved from UniProt using their assigned UniProt, CRC64 or RefSeq IDs. In case it matches with more than one entry, the highest score was kept. To note, some viral entries may map to the precursor polyprotein rather than the specific eORF segment. eORFs lacking an annotation score indicate no exact sequence match to any UniProt entry. The human proteome analysis was restricted to reviewed Homo sapiens entries (organism ID: 9606; reviewed:true; n = 20,420).

### **Statistical analysis**

GraphPad Prism v10 and RStudio with R v4.2.0 were used for all statistical analysis. Statistical tests are described in figure legends.

### **Clustering of eORF hits**

For the clustering, only eORFs significantly enriched in any of the barcoded analyzed with MAGeCK<sup>49</sup> (FDR  $\leq$  0.05 and LFC  $\geq$   $\log_2(1.5)$ ) or non-barcoded screens analyzed with DESeq2<sup>104</sup> (FDR  $\leq$  0.01 and LFC  $\geq$  1) were considered. Log fold change (LFC) values were clipped at the 10th and 90th percentiles and LFC of non-significant hits was set to zero. Hits were then clustered using a two-layer self-organizing map using the supersom function from the kohonen R package (v3.0.10) with a 4x5 hexagonal, toroidal grid. The first layer contained the clipped LFC values (user.weight = 1), and the second layer used a binary matrix indicating significant hits (user.weight = 10).

### **Enrichment of homologs**

For each cluster of eORF hits defined in the previous section, significantly over-represented protein homology groups (defined as having  $\geq$ 30% sequence homology using MMseqs2,<sup>106</sup> as detailed above) were identified using a one-sided Fisher's exact test (alternative= 'greater'), with all the other clusters constituting the control group. Of note, only protein homology groups containing  $\geq$  2 ORFs were considered. Resulting p-values were adjusted for multiple hypothesis testing across all cluster-

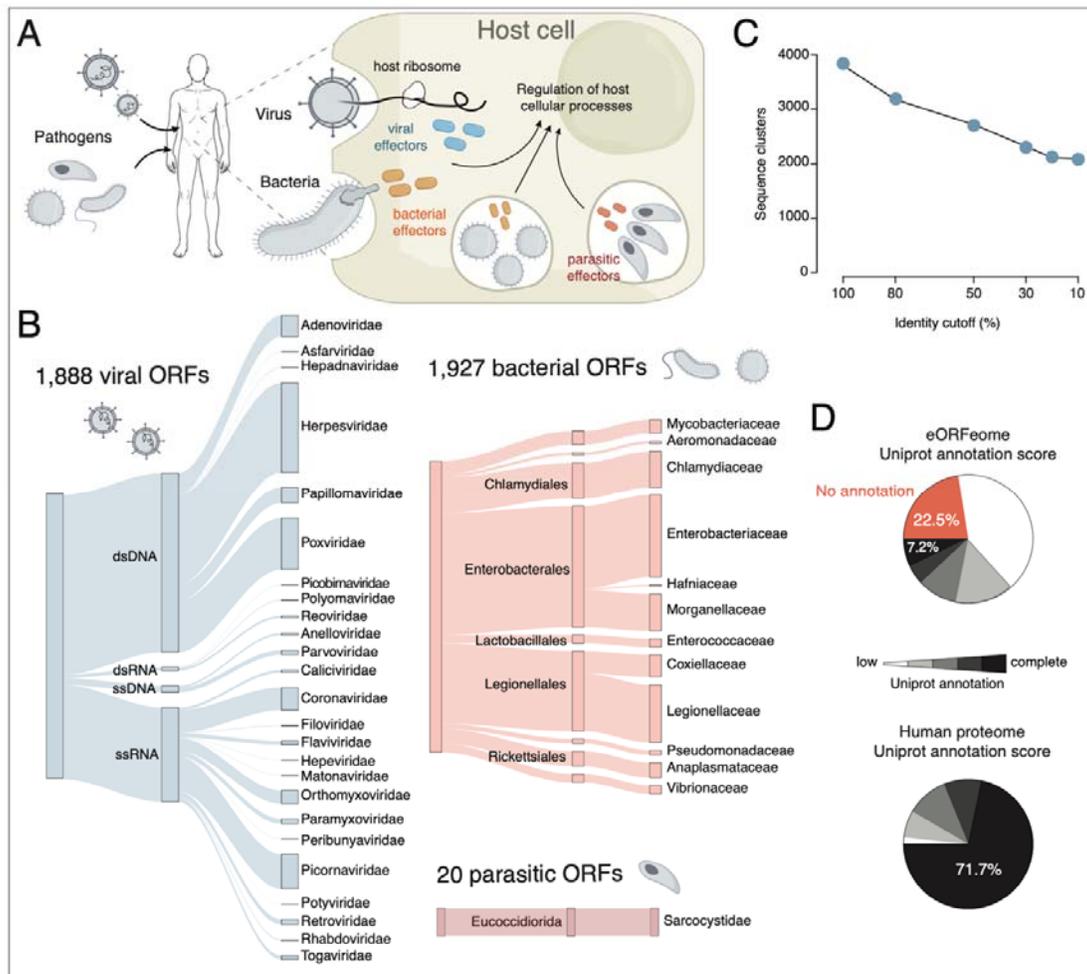
homology group comparisons using the Benjamini-Hochberg method to control the false discovery rate (FDR).

### **Structural modelling**

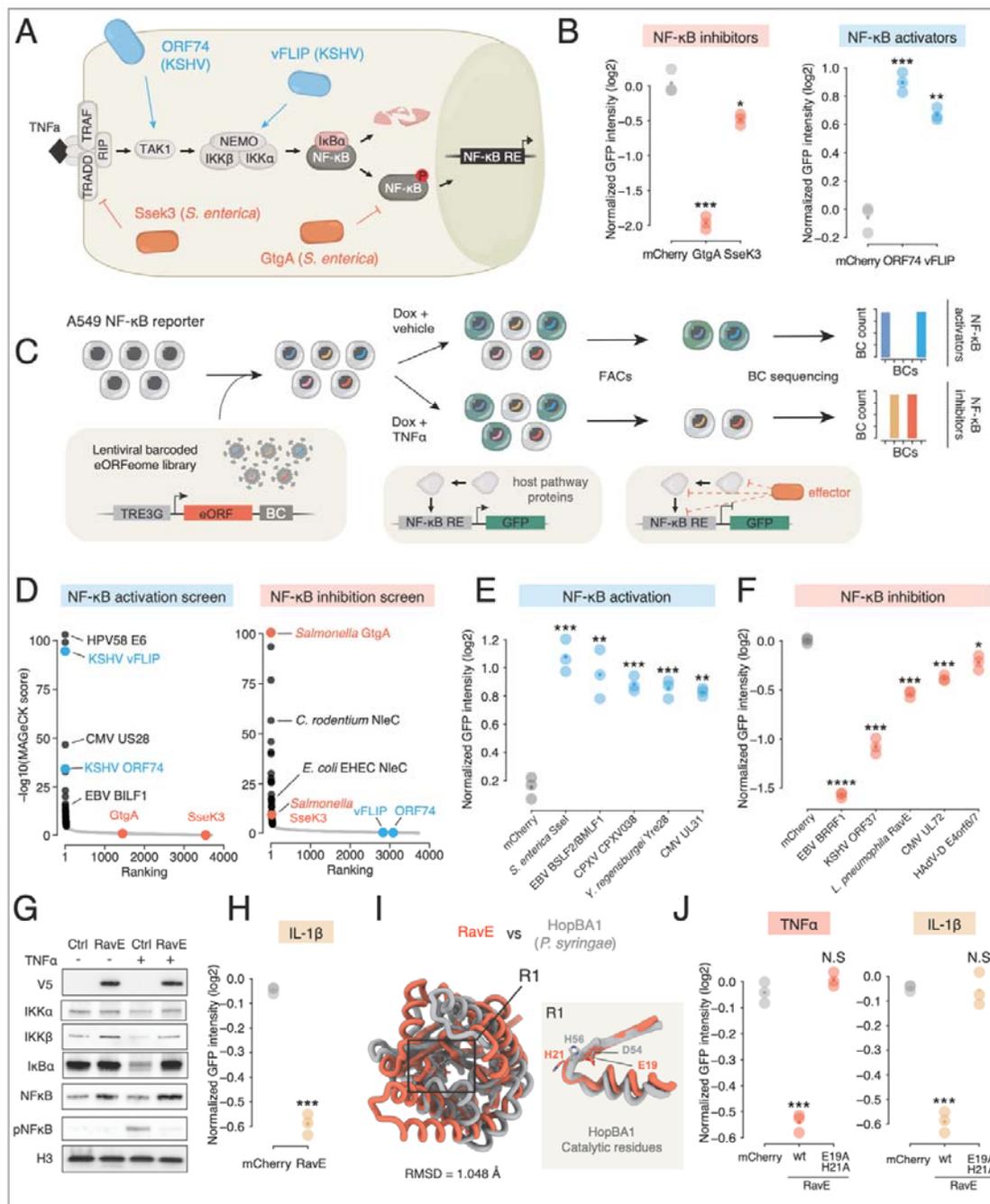
Predicted three-dimensional structures of the U14-p53 complex and of monomeric RavE were generated using the AlphaFold 3 server (<https://alphafoldserver.com/>).<sup>121</sup> The UniProt accession numbers Q69549 (U14) and P04637 (p53) were provided as input for the complex prediction, while Q5ZZ16 was used for RavE. PDB database (<https://www.rcsb.org/>) was used to retrieve the models of DNA-p53 (PDB ID: 2AC0), SV40 LargeT-p53 (PDB ID:2H1L) and HPV16 E6-p53 (PDB ID: 8GCR). All the resulting structural models were visualized using UCSF ChimeraX (v1.10).<sup>122</sup>

To identify structural homologs, the predicted structure of RavE (amino acids 1-234; UniProt: Q5ZZ16) was used as a query to search the PDB100 database (release 20240101) using Foldseek.<sup>60</sup> The search was taxonomically restricted to Bacteria.

## FIGURES

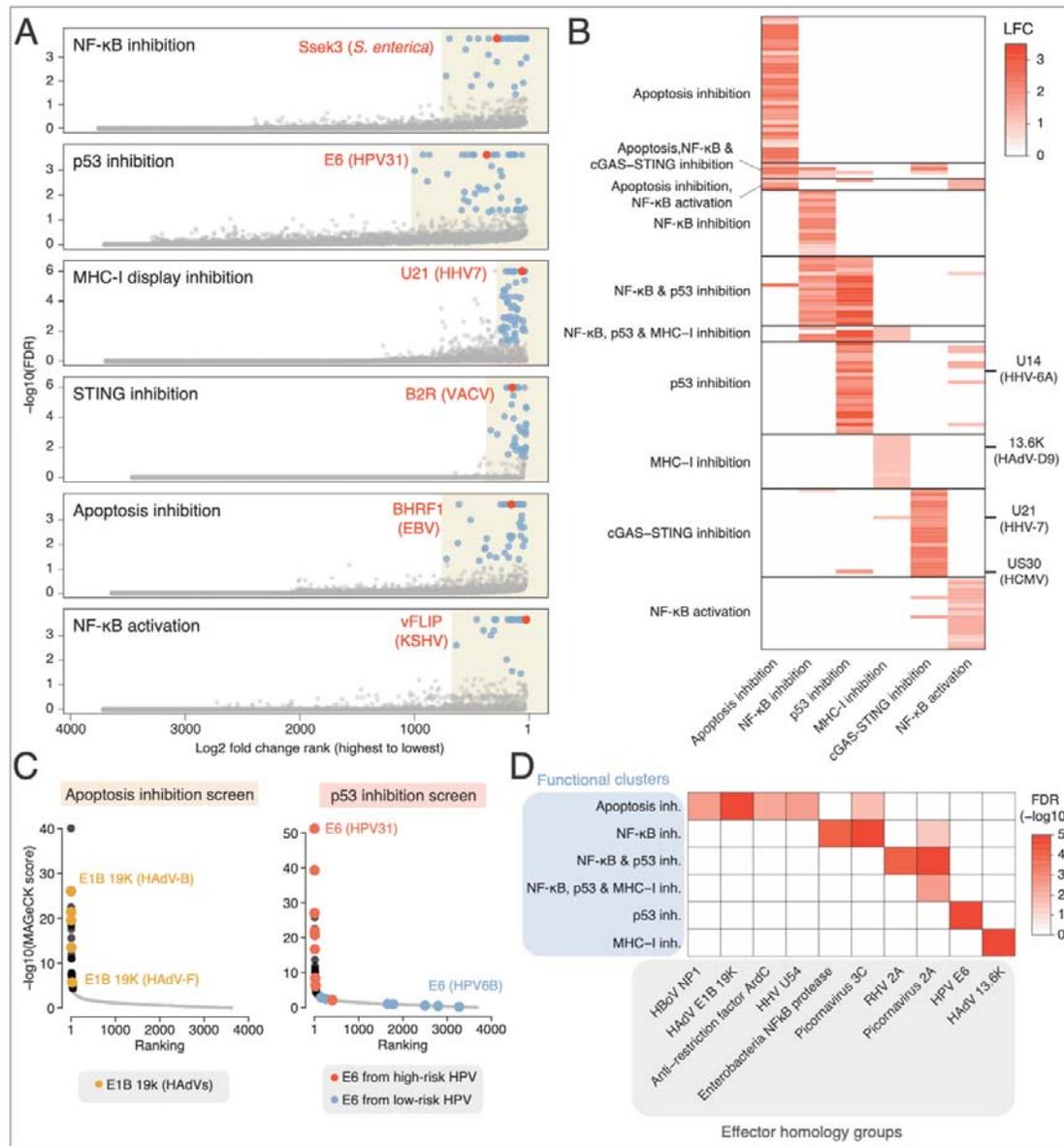


**Figure 1. The Pan-Pathogen eORFeome Library.** (A) Schematic of host cell manipulation by pathogen effectors. Diverse pathogens, including viruses, bacteria, and parasites, deploy effector proteins into host cells to hijack cellular processes, thereby promoting pathogen survival and subverting immune responses. (B) Alluvial diagrams illustrate the phylogenetic representation of the viral, bacterial and parasitic eORFs present in the eORFeome library. The flow diagrams connect taxonomic ranks - from Baltimore classification to Family for viruses, and from Order to Family for bacteria and parasites- with stream width corresponding to the abundance of ORFs from each group. (C) Sequence homology analysis of the eORFeome library. eORFs were clustered using MMseqs2<sup>106</sup> at varying sequence identity cutoffs. The plot shows the total number of resulting homology clusters for each identity threshold. (D) Characterization of the eORFeome library (top) or human proteome (bottom) by UniProt annotation score. The pie chart shows the percentage of ORFs corresponding to each score. The grayscale colour key indicates the annotation level, from score 1 (white) to score of 5 (black). Orange corresponds to no entry in the UniProt database.



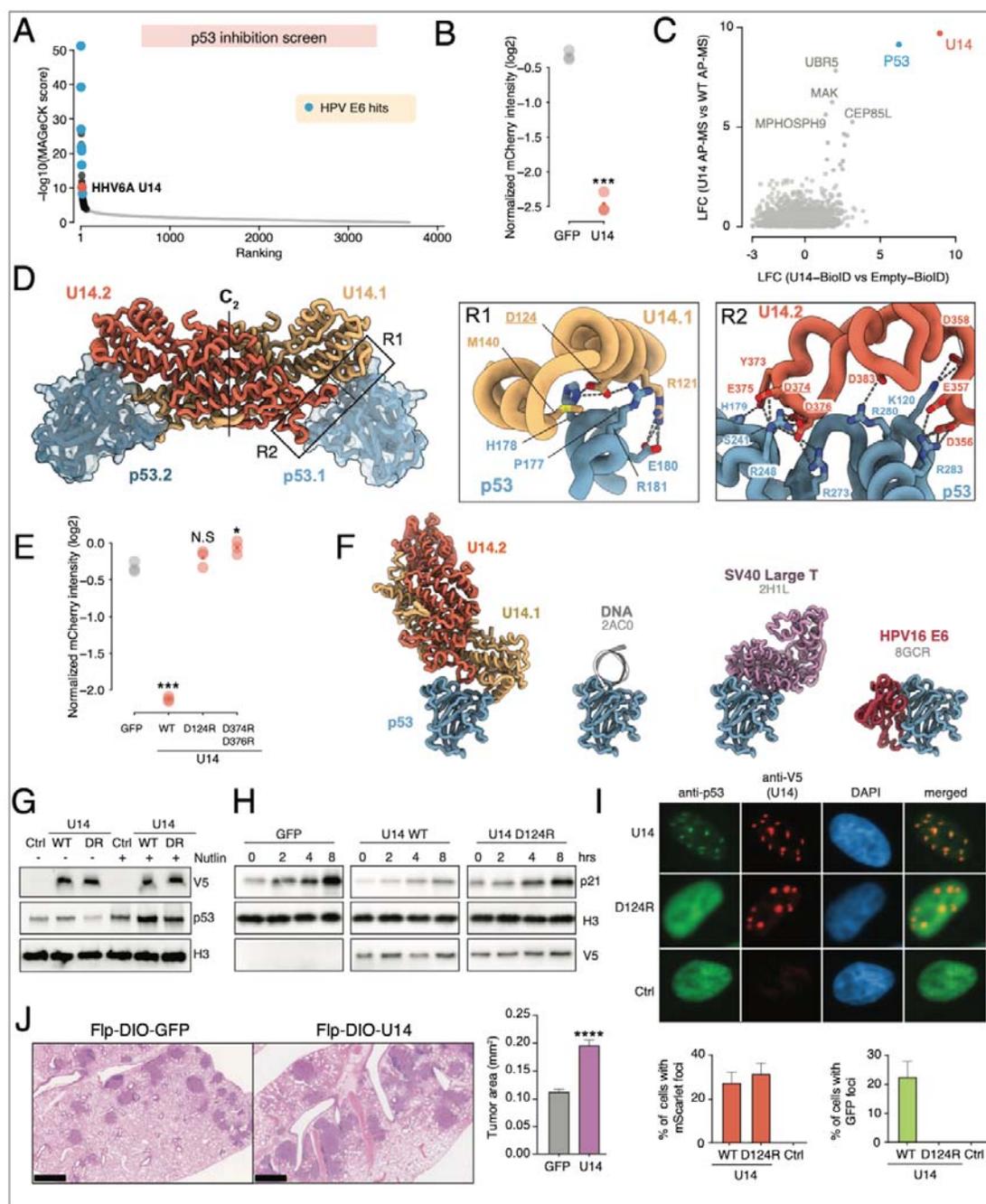
**Figure 2. A proof-of-principle screen identifies known and novel NF- $\kappa$ B modulators.** (A) Schematic of the NF- $\kappa$ B pathway. Points of intervention are shown for known inhibitory eORFs (*S. enterica* Ssek3 and GtgA) and activating eORFs (KSHV vFLIP and ORF74). (B) GFP intensity in an A549 NF- $\kappa$ B reporter cell line. Expression of the indicated ORFs was induced with dox. (Left) Cells expressing GtgA and Ssek3 from *Salmonella enterica*, following treatment with 20 ng/mL TNF $\alpha$  for 16 hours to induce reporter activity. (Right) Cells expressing ORF74 and vFLIP from KSHV without TNF $\alpha$  treatment. The plots show the median GFP intensity, measured by flow cytometry and normalized to parallel cultures that were not treated with dox. The mCherry ORF serves as the negative control. Statistical

significance was determined by comparing each ORF to the mCherry control using multiple paired t-tests. Dots indicate the mean of replicates. (C) Schematic of the screening platform for identifying eORFs that modulate NF- $\kappa$ B. The A549-NF- $\kappa$ B reporter cell line was used to monitor pathway activity. The reporter population was transduced at a low multiplicity of infection with a dox-inducible, barcoded lentiviral eORFeome library, ensuring single eORF expression per cell. For the activation screen (top), eORF expression was induced with dox, and GFP cells (indicative of pathway activation) were enriched by FACS. For the inhibition screen (bottom), eORF expression was induced with dox, followed by stimulation with TNF $\alpha$  (20 ng/mL) for 16 h, and cells with reduced GFP fluorescence were enriched by FACS. For both screens, eORF-associated barcodes from sorted and input populations were recovered from gDNA and quantified by deep sequencing. MAGeCK analysis was used to aggregate enrichment scores to systematically identify candidate eORFs with significant activating or inhibitory effects on NF- $\kappa$ B activity. (D) Ranking plot of enrichment scores for eORFs from the NF- $\kappa$ B activation (left) and inhibition (right) screens. Candidate hits (black dots) are distinguished from non-hits (grey) by a threshold of FDR < 0.05 and a fold change > 1.5. Positive controls for activation (vFLIP, ORF74; blue) and inhibition (Ssek3, GtgA; orange) are indicated, and other known regulators examples are labelled. (E) Validation of candidate eORFs from the NF- $\kappa$ B activation screen: SseI (*S. enterica*), BSLF2+BMLF1 (EBV), CPXV038 (Cowpox virus), Yre28/WP\_006817680.1 (*Y. regensburgi*) and UL31 (HCMV). The assay was performed under basal conditions as in (B, right). (F) Validation of candidate ORFs from the NF- $\kappa$ B inhibition screen: BRRF1 (EBV), ORF37 (KSHV), RavE (*L. pneumophila*), UL72 (HCMV), and E4orf6/7 (HAdV-D). The assay was performed with TNF $\alpha$  stimulation as in (B, left). (G) Western blot analysis of NF- $\kappa$ B pathway proteins in cells expressing the novel inhibitor RavE from *L. pneumophila*. Cells expressing V5-tagged RavE or an mCherry control were treated with or without TNF $\alpha$  (20 ng/mL) for 30 min. Lysates were immunoblotted with the indicated antibodies. An anti-V5 antibody was used to confirm RavE expression, and H3 served as a loading control. (H) Inhibitory activity of RavE assessed following pathway activation with IL-1 $\beta$  (1 ng/mL) for 16 h. (I) The structured domain of RavE (aa 1-234; orange) is shown aligned with the structure of the *P. syringae* effector HopBA1 (grey). A highlighted region, R1, indicates the alignment of catalytic residues from HopBA1 and RavE. The Root Mean Square Deviation (RMSD) indicates the degree of structural similarity between the two proteins. (J) Comparison of the inhibitory activity of wild-type RavE and a RavE (H21A,E19A) mutant following pathway activation with either TNF $\alpha$  (left) or IL-1 $\beta$  (right). \*\*\*\*p < 0.0001, \*\*\*p < 0.001, \*\*p < 0.01, \*p < 0.05; N.S, not significant. See also Figure S3 for representative examples of the single-cell GFP fluorescence distribution.



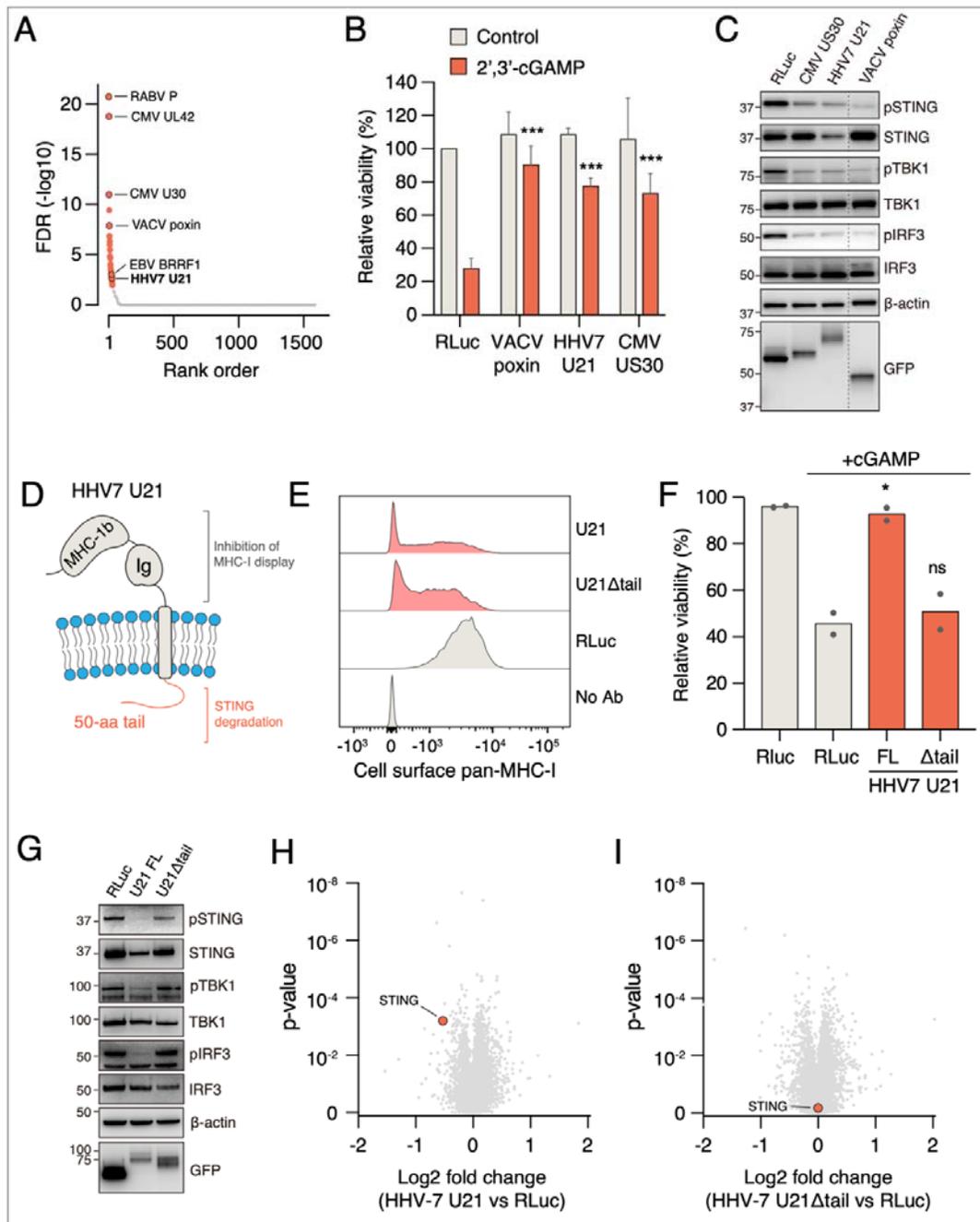
**Figure 3. Parallel eORFeome screens uncover pathway-specific and pleiotropic effector functions.** (A) Volcano plots with the results from six independent eORFeome screens. Each plot displays the enrichment (log<sub>2</sub> fold-change; LFC) ranking on the x-axis (from highest/less enriched to lowest/more enriched) against the statistical significance ( $-\log_{10}$  FDR) on the y-axis. Individual eORFs are represented as dots. Significant hits (FDR < 0.05 and FC > 1.5) are highlighted in blue, while a known pathway regulator is marked in orange for each screen as a positive control. (B) Self-organizing map (SOM) clustering of all significant hits (FDR < 0.05, LFC > 1) from the screens. The analysis groups eORFs into 10 distinct clusters based on their enrichment profiles (log-fold change values) across the different screens. (C) Enrichment ranking plot of eORFs from the apoptosis-inhibition screen (left) or p53-inhibition screen (right). Significant hits are marked in black. E1B 19K (HADVs), E6 (high-risk HPVs) and E6 (low-risk HPVs) protein homologs are highlighted. (D) Enrichment of protein families (homology groups) within the functional clusters from (B). Homology groups were defined by clustering ORFs with a 30% sequence identity cutoff (MMseqs2<sup>106</sup>).

The heatmap displays the FDR values for the enrichment of each homology group within each functional cluster. Only homology groups containing more than one ORF and with a significant enrichment ( $FDR < 0.05$ ) in at least one functional cluster are shown.



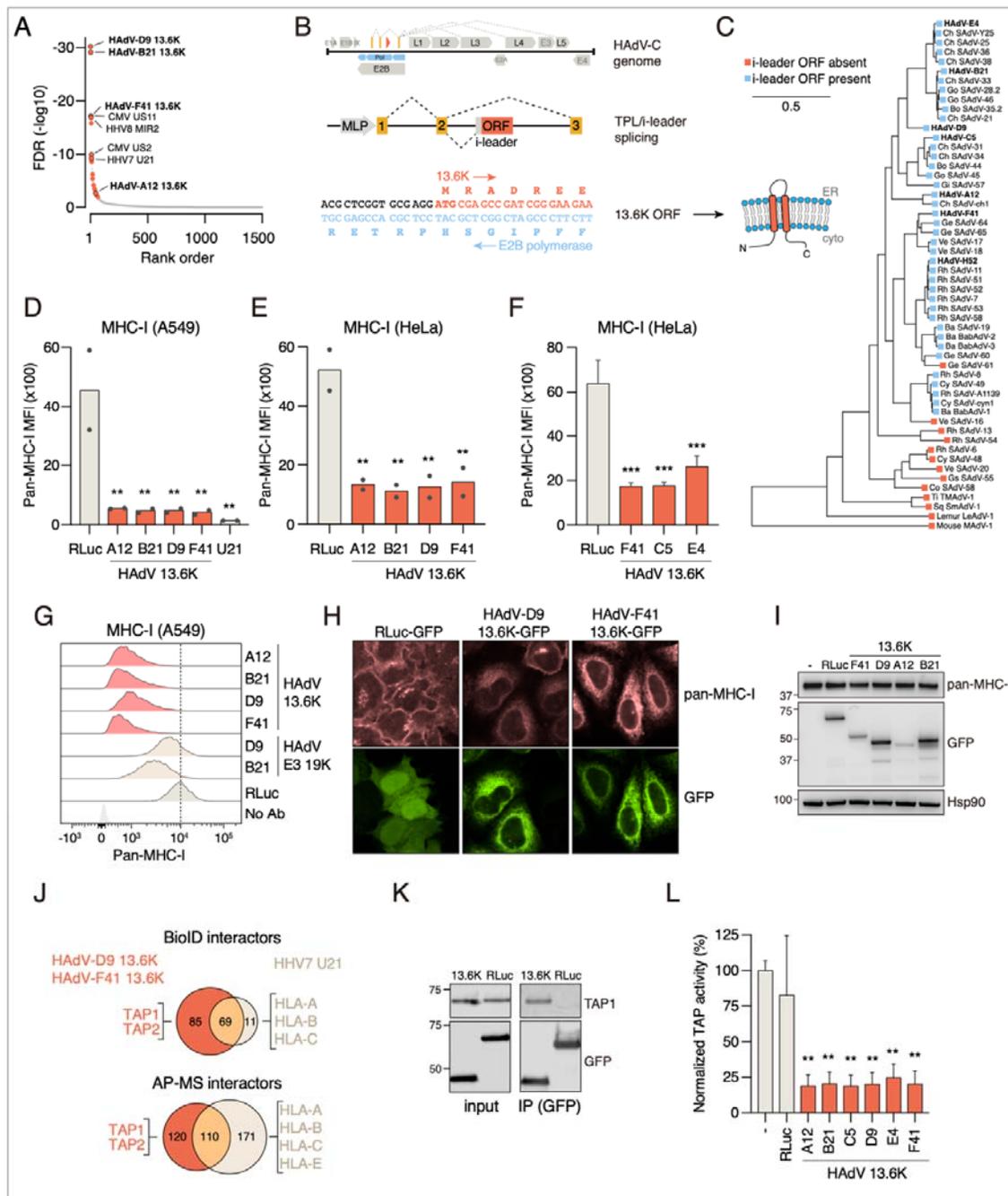
**Figure 4. U14 from HHV6A is a potent p53 antagonist.** (A) Enrichment ranking plot of eORFs from the p53-inhibition screen using a p53-mCherry reporter in A549 cells. Significant hits (FDR < 0.05 and FC > 1.5) are highlighted in black. HPV E6 proteins that scored as hits are highlighted in blue; U14 from HHV6A is highlighted in orange. All other eORFs are shown in grey. (B) mCherry intensity in an A549 p53 reporter cell line. Expression of the novel candidate U14 (HHV-6A) was induced with dox, following treatment with nutlin (2.5  $\mu\text{M}$ ) for 16 hours to induce reporter activity. The plots show the median mCherry intensity, measured by flow cytometry and normalized to parallel cultures that were not treated with dox. The GFP ORF serves as the negative control. Statistical significance was determined by comparing each ORF to the GFP control using multiple paired t-tests.

Dots indicate the mean of replicates. (C) Scatter plot of protein enrichment scores from BioID (x-axis) and IP-MS (y-axis) for U14 relative to an empty control. (D) AlphaFold 3 model of the U14-p53 complex. The U14 dimer is shown in orange and yellow, bound to two p53 proteins in blue and light blue. Black rectangles (R1 and R2) mark the interaction regions. The right panels show enlarged views of the interaction regions R1 and R2 from the structural model. Key amino acid residues are shown and labelled. (E) Comparison of the inhibitory activity of wild-type U14, U14 D124R mutant or U14 D374R,D376R mutant following pathway activation with nutlin (2.5  $\mu$ M) for 16 hours. (F) Structural models of p53 in complex with various binding partners. Shown are PDB structures of p53 bound to DNA,<sup>123</sup> SV40 LTag<sup>75</sup> and HPV E6,<sup>124</sup> alongside an AlphaFold 3 model of the predicted interaction between U14 dimers and p53. (G) Western blot analysis of cells expressing GFP, V5-tagged U14, or V5-tagged U14 D124R mutant, treated with or without 2.5  $\mu$ M nutlin for 1 hour. Blots were probed with antibodies against V5, p53, and Histone H3 as a loading control. (H) Western blot analysis of cells expressing GFP, V5-U14, or V5-U14 D124R, treated with nutlin for the indicated times (0-8 hours). Blots were probed for p21, V5, and H3 loading control. (I) Immunofluorescence of cells expressing an empty control, V5-tagged U14, or V5-tagged U14 D124R. Cells were stained with antibodies against p53 and V5, with DAPI for nuclear visualization. Bottom: Quantification of the percentage of cells with p53 foci (left) and U14 foci (right). Data are shown as mean  $\pm$  SD from three independent experiments. (J) In vivo evaluation using a Rosa-CreERT2; FSF-KrasG12D lung adenocarcinoma mouse model. Lentivirus co-expressing U14 or GFP control was delivered via inhalation, followed by tamoxifen induction. Lungs were harvested 9 weeks later, H&E stained, and tumour area quantified. FACS data in all panels is shown as mean  $\pm$  SD from three independent experiments. Statistical significance is calculated relative to the mCherry (negative control) + dox condition for all panels, unless otherwise specified by lines indicating alternative comparisons, and is denoted as follows: \**padj* < 0.05, \*\**padj* < 0.01, \*\*\**padj* < 0.001, \*\*\*\**padj* < 0.0001. See also Figure S4 for representative examples of the single-cell mCherry fluorescence distribution.



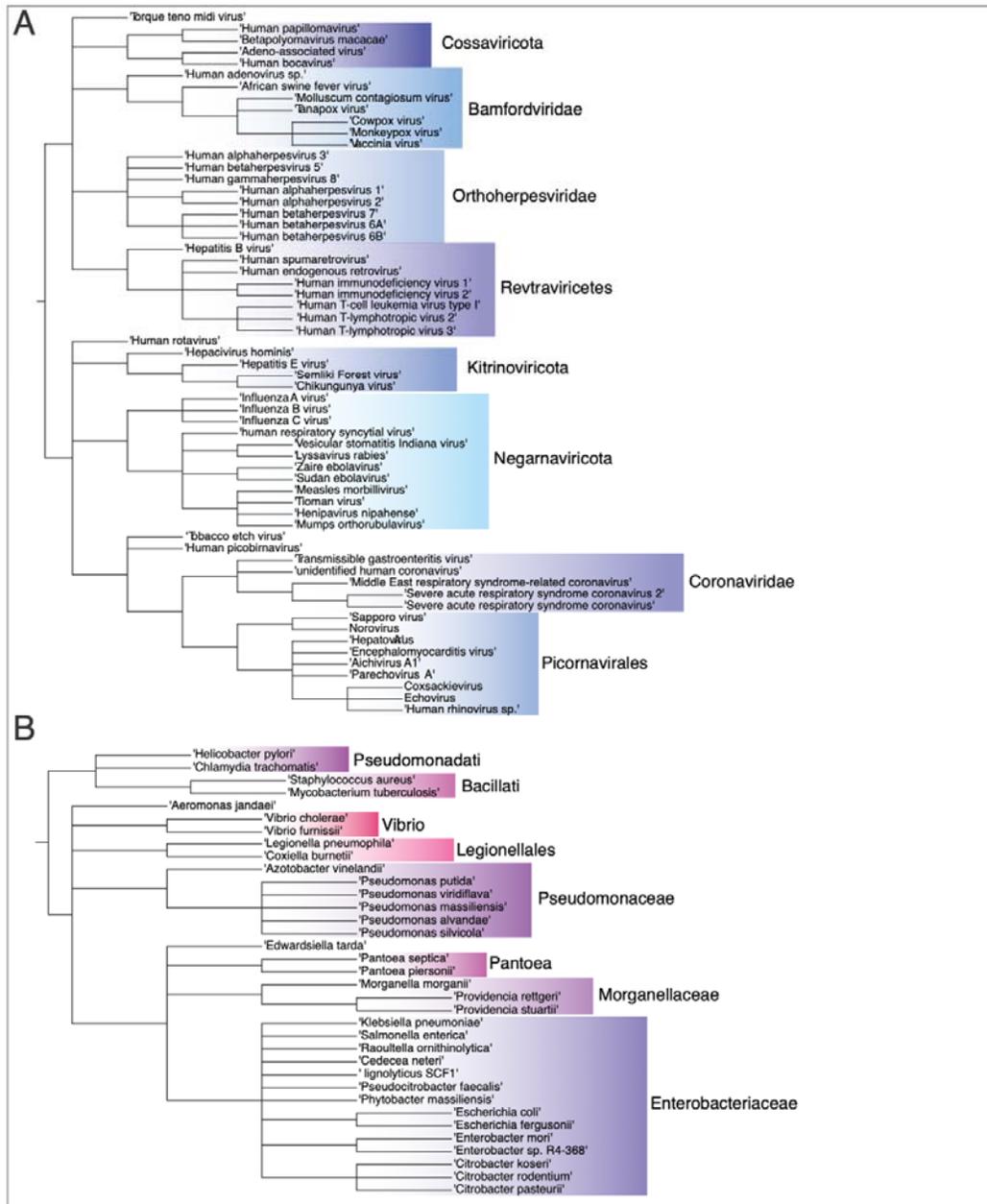
**Figure 5. HHV7 U21 antagonizes cGAS-STING signaling by inducing STING degradation.** (A) Enrichment ranking plot of eORFs from the STING inhibition screen in U937 cells. Significant hits are highlighted in red. Notable known inhibitors of STING signaling are labeled, with U21 highlighted as a novel STING inhibitor. All other eORFs are shown in grey. (B) U937 cells infected with indicated lentiviral dox-inducible GFP-tagged constructs, treated with dox and 40  $\mu$ g/ml 2',3'-cGAMP or vehicle and assessed for viability with CellTiter-Glo. (C) Indicated GFP-tagged proteins were expressed in cGAMP-treated U937 cells followed by western blotting for STING pathway components. (D) Schematic of HHV7 U21 protein. (E) U937 cells expressing full-length U21 (FL), U21 lacking the C-terminal tail ( $\Delta$ tail, aa 1-380), or Renilla luciferase were stained for cell surface MHC-I and

analyzed by flow cytometry. (F) Full-length U21 and U21 $\Delta$ tail were assessed for suppression of cGAMP-induced cell death as in (B). (G) Full-length U21 and U21 $\Delta$ tail were assessed for inhibition of STING pathway by western blotting as in (C). (H-I) Full-length HHV7 U21 (H) or U21 $\Delta$ tail (I) were expressed in U937 cells, which were subjected to whole cell proteomics.

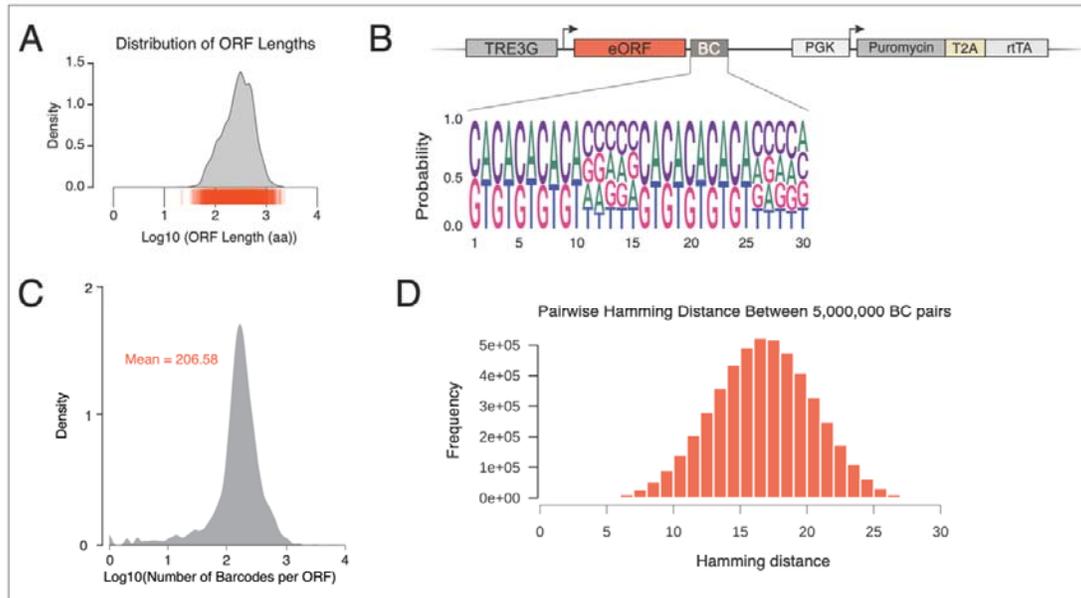


**Figure 6. Adenoviral 13.6K/i-leader protein is a recently evolved inhibitor of the TAP transporter.** (A) Enrichment ranking plot of eORFs from the MHC-I display screen in A549 cells. Significant hits are highlighted in red. Notable known inhibitors of STING signaling are labeled, with 13.6K clones present in the library highlighted as novel MHC-I regulators. All other eORFs are shown in grey. (B) 13.6K is encoded by the i-leader ORF located between the second and third adenoviral tripartite leader exons, on the opposite strand of the E2B polymerase. (C) Phylogenetic analysis of i-leader ORF presence in primate adenoviruses. Human adenoviruses are represented by a single strain from each clade whereas primate adenoviruses include representatives of 98% sequence identity clusters. Ch, chimpanzee; Bo, bonobo; Go, gorilla; Gi, gibbon; Ge, gelada baboon; Ve, vervet monkey;

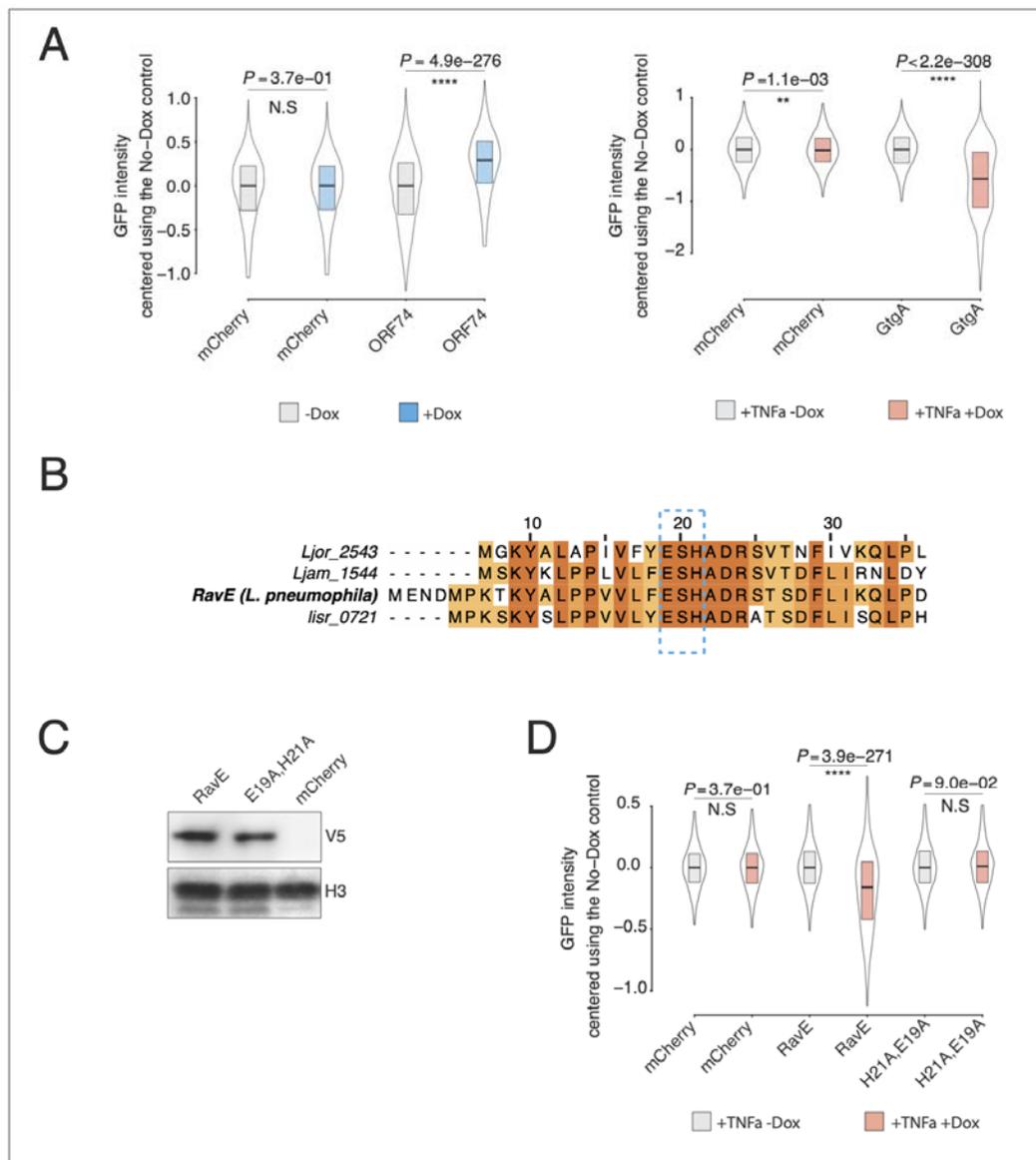
Rh, rhesus macaque; Ba, baboon; Cy, cynomolgus macaque; Gs, Golden snub-nosed monkey; Co, Colobus monkey; Ti, titi monkey; Sq, squirrel monkey. (D) A549 cells stably expressing dox-inducible GFP-tagged 13.6K proteins or HHV7 U21 were treated with dox for 24 hours followed by flow cytometry analysis of cell surface MHC-I levels. (E) The same 13.6K-GFP constructs were analyzed in HeLa cells as in (D). (F) 13.6K-GFP proteins from HAdV-F41 and HAdV-C5 were similarly analyzed in HeLa cells. (G) Comparison of MHC-I surface levels in A549 cells expressing 13.6K-GFP or the previously known immunoevasin E3 19K-GFP from HAdV-D9 or HAdV-B21. (H) HeLa cells expressing HAdV-D9 or HAdV-F41 13.6K-GFP or Renilla luciferase fused to GFP were stained for MHC-I. The GFP channel shows localization of the GFP-tagged constructs. (I) Western blot analysis of total MHC-I levels in HeLa cells expressing indicated 13.6K-GFP constructs. (J) HAdV-D9 or HAdV-F41 were tagged with miniTurbo-3xFLAG in HeLa cells and analyzed for physical and proximity interactors with BioID and AP-MS, respectively. HHV7 U21 served as a control for an immunoevasin that inhibits surface trafficking of MHC-I molecules. Venn diagrams show the overlap of 13.6K and U21 interactors. (K) HAdV-D9 13.6K-GFP or Renilla luciferase-GFP expressed in HeLa cells was immunoprecipitated with an anti-GFP antibody followed by western blotting for endogenous TAP1 or GFP. (L) GFP-tagged Renilla luciferase or 13.6K from indicated adenoviral serotypes were expressed in HeLa cells and semi-permeabilized followed by a translocation assay of fluorescently labeled TAP substrate peptide into the ER in the presence of ATP and  $Mg^{2+}$ . Transported peptide was detected by flow cytometry.



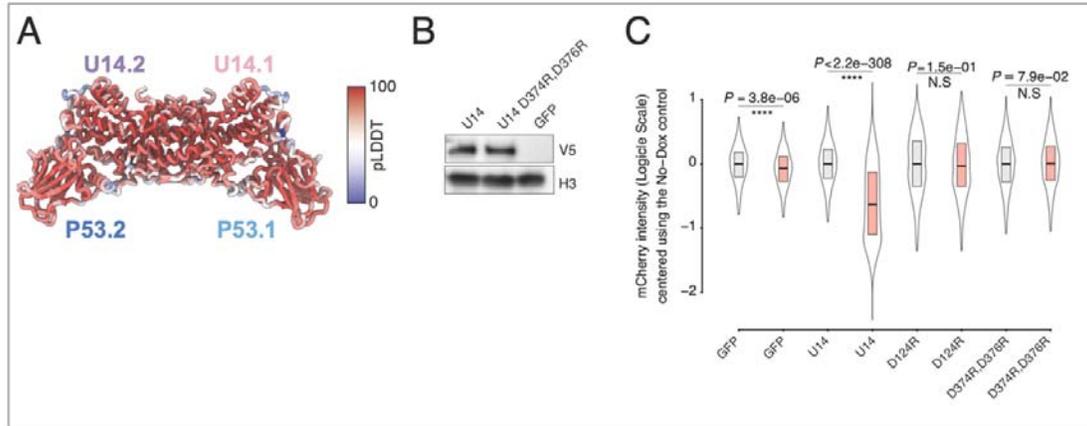
**Supplementary Figure 1. Evolutionary relationships of eORFs in the eORFeome library.** Phylogenetic tree of the viral (A) or bacterial (B) eORFs present in the eORFeome library.



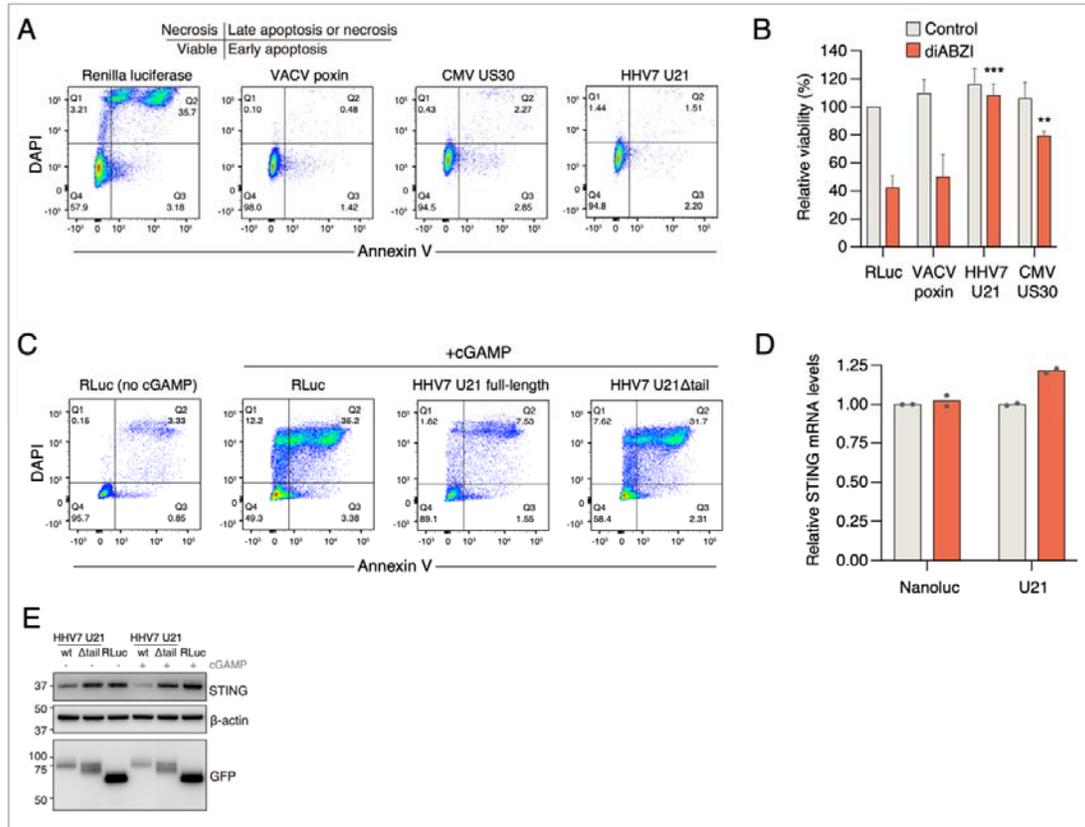
**Supplementary Figure 2. Characterization of the eORFeome library.** (A) Distribution of eORF lengths. The density plot is based on the Log10-transformed length (in amino acids) of each ORF in the library. The underlying orange lines marks the position of each individual eORF. (B) Top: Scheme of the lentiviral construct, detailing the arrangement of its key components: a dox-inducible TRE3G promoter drives the expression of an eORF, which is followed by a unique barcode (BC). Downstream of this, a constitutive promoter ensures the expression of both puromycin for cell selection and rtTA, which is essential for the dox-inducible activation of the TRE3G promoter. Bottom: position weight matrix that describes the semi-random nature of the 30-nucleotide barcode (BC), which consists of the pattern [(SW)×5 + N×5] repeated twice, where S represents G or C, W represents A or T, and N represents any nucleotide. (C) Density plot of the Log10 distribution of the number of barcodes per ORF, with a mean of 206.58 BC per ORF. (D) Histogram of the Hamming distance distribution, where the frequency (y-axis) of a given Hamming distance (x-axis) was calculated from 5 million randomly sampled barcode pairs from the library.



**Supplementary Figure 3. Validation of eORF hits from NF- $\kappa$ B Screens.** (A) Boxplots showing representative examples of the single-cell GFP fluorescence distribution, complementing the median-based analysis presented in Figure 2B. ORF expression was induced with dox (blue or orange) or left untreated (grey) and where specified, cells were treated with 20 ng/mL TNF $\alpha$  for 16 hours to stimulate NF- $\kappa$ B activity. The y-axis represents the Logicle-transformed GFP fluorescence, centred to the median of the corresponding no-dox control population. Boxplots display the median and interquartile range. mCherry serves as the negative control. Statistical significance was determined by comparing dox-treated versus untreated samples for each ORF. (B) Multiple sequence alignment of four RavE homologs. Conserved residues are coloured in orange. (C) Western blot analysis of cells expressing V5-tagged RavE, V5-tagged RavE E19A,H21A mutant, or an mCherry control. Lysates were immunoblotted with an anti-V5 antibody to confirm that the mutant protein is stably expressed. H3 served as a loading control. (D) Boxplots showing single-cell GFP fluorescence distributions, generated using the same analysis method described in (A). These data complement the median-based analysis presented in Figure 2J.

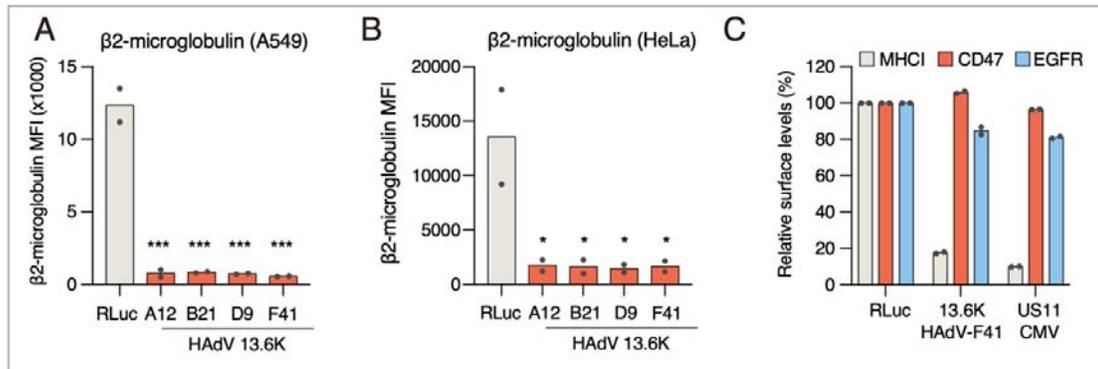


**Supplementary Figure 4. Validation of U14 from HHV6A as a p53 antagonist.** (A) Multiple sequence alignment of four RavE homologs. Conserved residues are coloured in orange. (B) Western blot analysis of cells expressing V5-tagged U14, V5-tagged U14 D124R mutant, or an GFP control. Lysates were immunoblotted with an anti-V5 antibody to confirm that the mutant protein is stably expressed. H3 served as a loading control. (C) Predicted local distance difference test (pLDDT) scores for the U14–p53 complex structural model generated by AlphaFold 3. (B) Western blot analysis of cells expressing V5-tagged U14, V5-tagged U14 D374R, D376R mutant, or an mCherry control. Lysates were immunoblotted with an anti-V5 antibody to confirm that the mutant protein is stably expressed. H3 served as a loading control. (C) Boxplots showing representative examples of the single-cell mCherry fluorescence distribution, complementing the median-based analysis presented in Figure 4. ORF expression was induced with dox (orange) or left untreated (grey) and where specified, cells were treated with 2.5uM nutlin for 16 hours to stimulate p53 activity. The panels show results for wild-type U14 with its inactive mutants (D124R or D374R, D376R). The y-axis represents the Logicle-transformed mCherry fluorescence, centred to the median of the corresponding no-dox control population. Boxplots display the median and interquartile range. GFP serves as the negative control. Statistical significance was determined by comparing dox-treated versus untreated samples for each ORF. \* $padj < 0.05$ , \*\* $padj < 0.01$ , \*\*\* $padj < 0.001$ , \*\*\*\*  $padj < 0.0001$ .



**Supplementary Figure 5. Validation of U21 as a potent inhibitor of STING signaling.**

(A) U937 cells stably expressing dox-inducible GFP-tagged Renilla luciferase, VACV poxin, CMV US30, or HHV7 U21 were induced with dox followed by 2',3'-cGAMP treatment and analyzed for apoptosis with Annexin V and DAPI staining. (B) The same constructs were analyzed for viability CellTiter-Glo after treating the cells with the non-nucleotide STING agonist diABZI. (C) Cells expressing full-length HHV7 U21 or U21Δtail were analyzed for apoptosis after cGAMP treatment as in (A). (D) STING mRNA levels were assessed by qRT-PCR in cells expressing Nanoluc-GFP or U21-GFP. (E) STING protein levels were analyzed by western blotting in cells expressing the indicated GFP-tagged constructs after treating the cells with cGAMP or vehicle.



**Supplementary Figure 6. Validation of 13.6K proteins as novel TAP inhibitors.** (A-B) A549 cells (A) and HeLa cells (B) expressing indicated GFP-tagged constructs were assessed for cell surface  $\beta$ 2-microglobulin levels by flow cytometry. (C) HeLa cells expressing GFP-tagged HAdV-F41 13.6K or CMV US11, a known inhibitor of MHC-I surface display, were analyzed for cell surface MHC-I, EGFR, or CD47 levels by flow cytometry.

## REFERENCES

1. Tomita, T., Guevara, R.B., Shah, L.M., Afrifa, A.Y., and Weiss, L.M. (2021). Secreted Effectors Modulating Immune Responses to *Toxoplasma gondii*. *Life (Basel)* *11*, 988. <https://doi.org/10.3390/life11090988>.
2. Shah, P.S., Beesabathuni, N.S., Fishburn, A.T., Kenaston, M.W., Minami, S.A., Pham, O.H., and Tucker, I. (2022). Systems Biology of Virus-Host Protein Interactions: From Hypothesis Generation to Mechanisms of Replication and Pathogenesis. *Annu Rev Virol* *9*, 397–415. <https://doi.org/10.1146/annurev-virology-100520-011851>.
3. Mattoo, S., Lee, Y.M., and Dixon, J.E. (2007). Interactions of bacterial effector proteins with host proteins. *Current Opinion in Immunology* *19*, 392–401. <https://doi.org/10.1016/j.coi.2007.06.005>.
4. Bartlett, A., Padfield, D., Lear, L., Bendall, R., and Vos, M. (2022). A comprehensive list of bacterial pathogens infecting humans. *Microbiology (Reading)* *168*. <https://doi.org/10.1099/mic.0.001269>.
5. Woolhouse, M.E.J., and Brierley, L. (2018). Epidemiological characteristics of human-infective RNA viruses. *Sci Data* *5*, 180017. <https://doi.org/10.1038/sdata.2018.17>.
6. Woolhouse, M.E.J., and Gowtage-Sequeria, S. (2005). Host Range and Emerging and Reemerging Pathogens. *Emerg. Infect. Dis.* *11*, 1842–1847. <https://doi.org/10.3201/eid1112.050997>.
7. Siew, N., and Fischer, D. (2003). Analysis of singleton ORFans in fully sequenced microbial genomes. *Proteins* *53*, 241–251. <https://doi.org/10.1002/prot.10423>.
8. Edgar, R.C., Taylor, B., Lin, V., Altman, T., Barbera, P., Meleshko, D., Lohr, D., Novakovsky, G., Buchfink, B., Al-Shayeb, B., et al. (2022). Petabase-scale sequence alignment catalyses viral discovery. *Nature* *602*, 142–147. <https://doi.org/10.1038/s41586-021-04332-2>.
9. Jehl, M.-A., Arnold, R., and Rattei, T. (2011). Effective--a database of predicted secreted bacterial proteins. *Nucleic Acids Res* *39*, D591-595. <https://doi.org/10.1093/nar/gkq1154>.
10. Arnold, R., Brandmaier, S., Kleine, F., Tischler, P., Heinz, E., Behrens, S., Niinikoski, A., Mewes, H.-W., Horn, M., and Rattei, T. (2009). Sequence-Based Prediction of Type III Secreted Proteins. *PLoS Pathog* *5*, e1000376. <https://doi.org/10.1371/journal.ppat.1000376>.
11. McDermott, J.E., Corrigan, A., Peterson, E., Oehmen, C., Niemann, G., Cambronne, E.D., Sharp, D., Adkins, J.N., Samudrala, R., and Heffron, F. (2011). Computational prediction of type III and IV secreted effectors in gram-negative bacteria. *Infect Immun* *79*, 23–32. <https://doi.org/10.1128/IAI.00537-10>.
12. Sonah, H., Deshmukh, R.K., and Bélanger, R.R. (2016). Computational Prediction of Effector Proteins in Fungi: Opportunities and Challenges. *Front Plant Sci* *7*, 126. <https://doi.org/10.3389/fpls.2016.00126>.

13. Callaway, E. (2025). These are the 20 most-studied bacteria - the majority have been ignored. *Nature* *637*, 770–771. <https://doi.org/10.1038/d41586-025-00038-x>.
14. Schneewind, O., and Missiakas, D.M. (2012). Protein secretion and surface display in Gram-positive bacteria. *Philos Trans R Soc Lond B Biol Sci* *367*, 1123–1139. <https://doi.org/10.1098/rstb.2011.0210>.
15. Gordon, D.E., Hiatt, J., Bouhaddou, M., Rezelj, V.V., Ulferts, S., Braberg, H., Jureka, A.S., Obernier, K., Guo, J.Z., Batra, J., et al. (2020). Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* *370*, eabe9403. <https://doi.org/10.1126/science.abe9403>.
16. Kulanayake, S., and Tikoo, S.K. (2021). Adenovirus Core Proteins: Structure and Function. *Viruses* *13*, 388. <https://doi.org/10.3390/v13030388>.
17. Damania, B., Kenney, S.C., and Raab-Traub, N. (2022). Epstein-Barr virus: Biology and clinical disease. *Cell* *185*, 3652–3670. <https://doi.org/10.1016/j.cell.2022.08.026>.
18. Nomburg, J., Doherty, E.E., Price, N., Bellieny-Rabelo, D., Zhu, Y.K., and Doudna, J.A. (2024). Birth of protein folds and functions in the virome. *Nature* *633*, 710–717. <https://doi.org/10.1038/s41586-024-07809-y>.
19. Mak, H., and Thurston, T.L.M. (2021). Interesting Biochemistries in the Structure and Function of Bacterial Effectors. *Front Cell Infect Microbiol* *11*, 608860. <https://doi.org/10.3389/fcimb.2021.608860>.
20. Bhavsar, A.P., Guttman, J.A., and Finlay, B.B. (2007). Manipulation of host-cell pathways by bacterial pathogens. *Nature* *449*, 827–834. <https://doi.org/10.1038/nature06247>.
21. Krump, N.A., and You, J. (2018). Molecular mechanisms of viral oncogenesis in humans. *Nat Rev Microbiol* *16*, 684–698. <https://doi.org/10.1038/s41579-018-0064-6>.
22. Casanova, J.-L., and Abel, L. (2021). Mechanisms of viral inflammation and disease in humans. *Science* *374*, 1080–1086. <https://doi.org/10.1126/science.abj7965>.
23. Medzhitov, R. (2021). The spectrum of inflammatory responses. *Science* *374*, 1070–1075. <https://doi.org/10.1126/science.abi5200>.
24. Weigele, B.A., Orchard, R.C., Jimenez, A., Cox, G.W., and Alto, N.M. (2017). A systematic exploration of the interactions between bacterial effector proteins and host cell membranes. *Nat Commun* *8*, 532. <https://doi.org/10.1038/s41467-017-00700-7>.
25. Stukalov, A., Girault, V., Grass, V., Karayel, O., Bergant, V., Urban, C., Haas, D.A., Huang, Y., Oubraham, L., Wang, A., et al. (2021). Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV. *Nature* *594*, 246–252. <https://doi.org/10.1038/s41586-021-03493-4>.
26. Scaturro, P., Stukalov, A., Haas, D.A., Cortese, M., Draganova, K., Płaszczycza, A., Bartenschlager, R., Götz, M., and Pichlmair, A. (2018). An orthogonal proteomic survey uncovers novel Zika virus host factors. *Nature* *561*, 253–257. <https://doi.org/10.1038/s41586-018-0484-5>.

27. Pichlmair, A., Kandasamy, K., Alvisi, G., Mulhern, O., Sacco, R., Habjan, M., Binder, M., Stefanovic, A., Eberle, C.-A., Goncalves, A., et al. (2012). Viral immune modulators perturb the human molecular network by common and unique strategies. *Nature* *487*, 486–490. <https://doi.org/10.1038/nature11289>.
28. Nobre, L.V., Nightingale, K., Ravenhill, B.J., Antrobus, R., Soday, L., Nichols, J., Davies, J.A., Seirafian, S., Wang, E.C., Davison, A.J., et al. (2019). Human cytomegalovirus interactome analysis identifies degradation hubs, domain associations and viral protein functions. *Elife* *8*, e49894. <https://doi.org/10.7554/eLife.49894>.
29. Luchetti, G., Roncaioli, J.L., Chavez, R.A., Schubert, A.F., Kofoed, E.M., Reja, R., Cheung, T.K., Liang, Y., Webster, J.D., Lehoux, I., et al. (2021). Shigella ubiquitin ligase IpaH7.8 targets gasdermin D for degradation to prevent pyroptosis and enable infection. *Cell Host Microbe* *29*, 1521–1530.e10. <https://doi.org/10.1016/j.chom.2021.08.010>.
30. Jäger, S., Cimermancic, P., Gulbahce, N., Johnson, J.R., McGovern, K.E., Clarke, S.C., Shales, M., Mercenne, G., Pache, L., Li, K., et al. (2012). Global landscape of HIV–human protein complexes. *Nature* *481*, 365–370. <https://doi.org/10.1038/nature10719>.
31. Coyaud, E., Ranadheera, C., Cheng, D., Gonçalves, J., Dyakov, B.J.A., Laurent, E.M.N., St-Germain, J., Pelletier, L., Gingras, A.-C., Brumell, J.H., et al. (2018). Global Interactomics Uncovers Extensive Organellar Targeting by Zika Virus. *Mol Cell Proteomics* *17*, 2242–2255. <https://doi.org/10.1074/mcp.TIR118.000800>.
32. Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O’Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* *583*, 459–468. <https://doi.org/10.1038/s41586-020-2286-9>.
33. Embry, A., Baggett, N.S., Heisler, D.B., White, A., de Jong, M.F., Kocsis, B.L., Tomchick, D.R., Alto, N.M., and Gammon, D.B. (2024). Exploiting bacterial effector proteins to uncover evolutionarily conserved antiviral host machinery. *PLoS Pathog* *20*, e1012010. <https://doi.org/10.1371/journal.ppat.1012010>.
34. Davis, Z.H., Verschueren, E., Jang, G.M., Kleffman, K., Johnson, J.R., Park, J., Von Dollen, J., Maher, M.C., Johnson, T., Newton, W., et al. (2015). Global mapping of herpesvirus-host protein complexes reveals a transcription strategy for late genes. *Mol Cell* *57*, 349–360. <https://doi.org/10.1016/j.molcel.2014.11.026>.
35. Losick, V.P., Haenssler, E., Moy, M.-Y., and Isberg, R.R. (2010). LnaB: a Legionella pneumophila activator of NF-kappaB. *Cell Microbiol* *12*, 1083–1097. <https://doi.org/10.1111/j.1462-5822.2010.01452.x>.
36. Urbanus, M.L., Quaile, A.T., Stogios, P.J., Morar, M., Rao, C., Di Leo, R., Evdokimova, E., Lam, M., Oatway, C., Cuff, M.E., et al. (2016). Diverse mechanisms of metaeffector activity in an intracellular bacterial pathogen, Legionella pneumophila. *Mol Syst Biol* *12*, 893. <https://doi.org/10.15252/msb.20167381>.
37. Roan, N.R., Gierahn, T.M., Higgins, D.E., and Starnbach, M.N. (2006). Monitoring the T cell response to genital tract infection. *Proc Natl Acad Sci U S A* *103*, 12069–12074. <https://doi.org/10.1073/pnas.0603866103>.

38. Young, V., Dohai, B., Hitch, T.C.A., Hyden, P., Weller, B., Van Heusden, N.S., Saha, D., Macgregor, J.F., Maseko, S.B., Lin, C.-W., et al. (2023). A gut meta-interactome map reveals modulation of human immunity by microbiome effectors. Preprint, <https://doi.org/10.1101/2023.09.25.559292> <https://doi.org/10.1101/2023.09.25.559292>.
39. Barylyuk, K., Koreny, L., Ke, H., Butterworth, S., Crook, O.M., Lassadi, I., Gupta, V., Tromer, E., Mourier, T., Stevens, T.J., et al. (2020). A Comprehensive Subcellular Atlas of the Toxoplasma Proteome via hyperLOPIT Provides Spatial Context for Protein Functions. *Cell Host & Microbe* 28, 752-766.e9. <https://doi.org/10.1016/j.chom.2020.09.011>.
40. Rahman, M.M. (2011). Modulation of NF- $\kappa$ B signalling by microbial pathogens. *Nat Rev Microbiol* 9, 291–306. <https://doi.org/10.1038/nrmicro2539>.
41. De Oliveira, D.E., Ballon, G., and Cesarman, E. (2010). NF- $\kappa$ B signaling modulation by EBV and KSHV. *Trends in Microbiology* 18, 248–257. <https://doi.org/10.1016/j.tim.2010.04.001>.
42. Finlay, B.B., and McFadden, G. (2006). Anti-Immunology: Evasion of the Host Immune System by Bacterial and Viral Pathogens. *Cell* 124, 767–782. <https://doi.org/10.1016/j.cell.2006.01.034>.
43. Sun, H., Kamanova, J., Lara-Tejero, M., and Galán, J.E. (2016). A Family of Salmonella Type III Secretion Effector Proteins Selectively Targets the NF- $\kappa$ B Signaling Pathway to Preserve Host Homeostasis. *PLoS Pathog* 12, e1005484. <https://doi.org/10.1371/journal.ppat.1005484>.
44. Chadha, A., and Chadee, K. (2021). The NF- $\kappa$ B Pathway: Modulation by Entamoeba histolytica and Other Protozoan Parasites. *Front. Cell. Infect. Microbiol.* 11, 748404. <https://doi.org/10.3389/fcimb.2021.748404>.
45. Günster, R.A., Matthews, S.A., Holden, D.W., and Thurston, T.L.M. (2017). SseK1 and SseK3 Type III Secretion System Effectors Inhibit NF- $\kappa$ B Signaling and Necroptotic Cell Death in Salmonella-Infected Macrophages. *Infect Immun* 85, e00010-17. <https://doi.org/10.1128/IAI.00010-17>.
46. Shimizu, A., Baratchian, M., Takeuchi, Y., Escors, D., Macdonald, D., Barrett, T., Bagneris, C., Collins, M., and Noursadeghi, M. (2011). Kaposi's Sarcoma-Associated Herpesvirus vFLIP and Human T Cell Lymphotropic Virus Type 1 Tax Oncogenic Proteins Activate I $\kappa$ B Kinase Subunit  $\gamma$  by Different Mechanisms Independent of the Physiological Cytokine-Induced Pathways. *J Virol* 85, 7444–7448. <https://doi.org/10.1128/jvi.02337-10>.
47. Vischer, H.F., Siderius, M., Leurs, R., and Smit, M.J. (2014). Herpesvirus-encoded GPCRs: neglected players in inflammatory and proliferative diseases? *Nat Rev Drug Discov* 13, 123–139. <https://doi.org/10.1038/nrd4189>.
48. Montaner, S., Sodhi, A., Servitja, J.-M., Ramsdell, A.K., Barac, A., Sawai, E.T., and Gutkind, J.S. (2004). The small GTPase Rac1 links the Kaposi sarcoma-associated herpesvirus vGPCR to cytokine secretion and paracrine neoplasia. *Blood* 104, 2903–2911. <https://doi.org/10.1182/blood-2003-12-4436>.

49. Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., and Liu, X.S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* 15. <https://doi.org/10.1186/s13059-014-0554-4>.
50. Lei, X., Sun, Z., Liu, X., Jin, Q., He, B., and Wang, J. (2011). Cleavage of the adaptor protein TRIF by enterovirus 71 3C inhibits antiviral responses mediated by Toll-like receptor 3. *J Virol* 85, 8811–8818. <https://doi.org/10.1128/JVI.00447-11>.
51. Pearson, J.S., Riedmaier, P., Marchès, O., Frankel, G., and Hartland, E.L. (2011). A type III effector protease NleC from enteropathogenic *Escherichia coli* targets NF- $\kappa$ B for degradation. *Molecular Microbiology* 80, 219–230. <https://doi.org/10.1111/j.1365-2958.2011.07568.x>.
52. Yen, H., Ooka, T., Iguchi, A., Hayashi, T., Sugimoto, N., and Tobe, T. (2010). NleC, a Type III Secretion Protease, Compromises NF- $\kappa$ B Activation by Targeting p65/RelA. *PLoS Pathog* 6, e1001231. <https://doi.org/10.1371/journal.ppat.1001231>.
53. Davis-Poynter, N., and Farrell, H.E. (2022). Constitutive Signaling by the Human Cytomegalovirus G Protein Coupled Receptor Homologs US28 and UL33 Enables Trophoblast Migration In Vitro. *Viruses* 14, 391. <https://doi.org/10.3390/v14020391>.
54. Le, V.T.K., Trilling, M., and Hengel, H. (2011). The Cytomegaloviral Protein pUL138 Acts as Potentiator of Tumor Necrosis Factor (TNF) Receptor 1 Surface Density To Enhance UL b'-Encoded Modulation of TNF- $\alpha$  Signaling. *J Virol* 85, 13260–13270. <https://doi.org/10.1128/JVI.06005-11>.
55. Beisser, P.S., Verzijl, D., Gruijthuisen, Y.K., Beuken, E., Smit, M.J., Leurs, R., Bruggeman, C.A., and Vink, C. (2005). The Epstein-Barr Virus BILF1 Gene Encodes a G Protein-Coupled Receptor That Inhibits Phosphorylation of RNA-Dependent Protein Kinase. *J Virol* 79, 441–449. <https://doi.org/10.1128/JVI.79.1.441-449.2005>.
56. Medici, M.A., Sciortino, M.T., Perri, D., Amici, C., Avitabile, E., Ciotti, M., Balestrieri, E., De Smaele, E., Franzoso, G., and Mastino, A. (2003). Protection by Herpes Simplex Virus Glycoprotein D against Fas-mediated Apoptosis. *Journal of Biological Chemistry* 278, 36059–36067. <https://doi.org/10.1074/jbc.M306198200>.
57. Boehme, K.W., Guerrero, M., and Compton, T. (2006). Human cytomegalovirus envelope glycoproteins B and H are necessary for TLR2 activation in permissive cells. *J Immunol* 177, 7094–7102. <https://doi.org/10.4049/jimmunol.177.10.7094>.
58. Vandermark, E.R., Deluca, K.A., Gardner, C.R., Marker, D.F., Schreiner, C.N., Strickland, D.A., Wilton, K.M., Mondal, S., and Woodworth, C.D. (2012). Human papillomavirus type 16 E6 and E7 proteins alter NF- $\kappa$ B in cultured cervical epithelial cells and inhibition of NF- $\kappa$ B promotes cell growth and immortalization. *Virology* 425, 53–60. <https://doi.org/10.1016/j.virol.2011.12.023>.
59. Vallabhapurapu, S., and Karin, M. (2009). Regulation and Function of NF- $\kappa$ B Transcription Factors in the Immune System. *Annu. Rev. Immunol.* 27, 693–733. <https://doi.org/10.1146/annurev.immunol.021908.132641>.

60. Van Kempen, M., Kim, S.S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C.L.M., Söding, J., and Steinegger, M. (2024). Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* 42, 243–246. <https://doi.org/10.1038/s41587-023-01773-0>.
61. Sanchez-Pulido, L., and Ponting, C.P. (2013). Tiki, at the head of a new superfamily of enzymes. *Bioinformatics* 29, 2371–2374. <https://doi.org/10.1093/bioinformatics/btt412>.
62. Nishimura, M.T., Anderson, R.G., Cherkis, K.A., Law, T.F., Liu, Q.L., Machius, M., Nimchuk, Z.L., Yang, L., Chung, E.-H., El Kasmi, F., et al. (2017). TIR-only protein RBA1 recognizes a pathogen effector to regulate cell death in Arabidopsis. *Proc Natl Acad Sci U S A* 114, E2053–E2062. <https://doi.org/10.1073/pnas.1620973114>.
63. Oren, M., and Prives, C. (2024). p53: A tale of complexity and context. *Cell* 187, 1569–1573. <https://doi.org/10.1016/j.cell.2024.02.043>.
64. Zhang, B., Xu, P., and Ablasser, A. (2025). Regulation of the cGAS-STING Pathway. *Annu Rev Immunol* 43, 667–692. <https://doi.org/10.1146/annurev-immunol-101721-032910>.
65. Carneiro, B.A., and El-Deiry, W.S. (2020). Targeting apoptosis in cancer therapy. *Nat Rev Clin Oncol* 17, 395–417. <https://doi.org/10.1038/s41571-020-0341-y>.
66. Dhatchinamoorthy, K., Colbert, J.D., and Rock, K.L. (2021). Cancer Immune Evasion Through Loss of MHC Class I Antigen Presentation. *Front Immunol* 12, 636568. <https://doi.org/10.3389/fimmu.2021.636568>.
67. Elsner, C., Ponnurangam, A., Kazmierski, J., Zillinger, T., Jansen, J., Todt, D., Döhner, K., Xu, S., Ducroux, A., Kriedemann, N., et al. (2020). Absence of cGAS-mediated type I IFN responses in HIV-1-infected T cells. *Proc Natl Acad Sci U S A* 117, 19475–19486. <https://doi.org/10.1073/pnas.2002481117>.
68. Khan, H., Sumner, R.P., Rasaiyaah, J., Tan, C.P., Rodriguez-Plata, M.T., Van Tulleken, C., Fink, D., Zuliani-Alvarez, L., Thorne, L., Stirling, D., et al. (2020). HIV-1 Vpr antagonizes innate immune activation by targeting karyopherin-mediated NF- $\kappa$ B/IRF3 nuclear transport. *Elife* 9, e60821. <https://doi.org/10.7554/eLife.60821>.
69. Rao, L., Debbas, M., Sabbatini, P., and White, E. (1992). The adenovirus E1A proteins induce apoptosis, which is inhibited by the E1B 19-kDa and Bc1-2 proteins. *Proc Natl Acad Sci U S A* 89, 7742–7746.
70. Thomas, M., Pim, D., and Banks, L. (1999). The role of the E6-p53 interaction in the molecular pathogenesis of HPV. *Oncogene* 18, 7690–7700. <https://doi.org/10.1038/sj.onc.1202953>.
71. Hoppe-Seyler, K., Bossler, F., Braun, J.A., Herrmann, A.L., and Hoppe-Seyler, F. (2018). The HPV E6/E7 Oncogenes: Key Factors for Viral Carcinogenesis and Therapeutic Targets. *Trends in Microbiology* 26, 158–168. <https://doi.org/10.1016/j.tim.2017.07.007>.
72. Nahar, K., Matsumoto, I., Taguchi, F., Inagaki, Y., Yamamoto, M., Toyoda, K., Shiraishi, T., Ichinose, Y., and Mukaiharu, T. (2014). *Ralstonia solanacearum* type III secretion system effector Rip36 induces a hypersensitive response in the nonhost wild eggplant *Solanum torvum*. *Mol Plant Pathol* 15, 297–303. <https://doi.org/10.1111/mpp.12079>.

73. Baruch, K., Gur-Arie, L., Nadler, C., Koby, S., Yerushalmi, G., Ben-Neriah, Y., Yogev, O., Shaulian, E., Guttman, C., Zarivach, R., et al. (2011). Metalloprotease type III effectors that specifically cleave JNK and NF- $\kappa$ B. *EMBO J* 30, 221–231. <https://doi.org/10.1038/emboj.2010.297>.
74. Takemoto, M., Koike, M., Mori, Y., Yonemoto, S., Sasamoto, Y., Kondo, K., Uchiyama, Y., and Yamanishi, K. (2005). Human Herpesvirus 6 Open Reading Frame U14 Protein and Cellular p53 Interact with Each Other and Are Contained in the Virion. *J Virol* 79, 13037–13046. <https://doi.org/10.1128/JVI.79.20.13037-13046.2005>.
75. Lilyestrom, W., Klein, M.G., Zhang, R., Joachimiak, A., and Chen, X.S. (2006). Crystal structure of SV40 large T-antigen bound to p53: interplay between a viral oncoprotein and a cellular tumor suppressor. *Genes Dev.* 20, 2373–2382. <https://doi.org/10.1101/gad.1456306>.
76. Deppert, W., Steinmayer, T., and Richter, W. (1989). Cooperation of SV40 large T antigen and the cellular protein p53 in maintenance of cell transformation. *Oncogene* 4, 1103–1110.
77. Wahl, S.G.F., Dai, H.Y., Emdal, E.F., Berg, T., Halvorsen, T.O., Ottestad, A.L., Lund-Iversen, M., Brustugun, O.T., Førde, D., Paulsen, E.-E., et al. (2021). The Prognostic Effect of KRAS Mutations in Non-Small Cell Lung Carcinoma Revisited: A Norwegian Multicentre Study. *Cancers (Basel)* 13, 4294. <https://doi.org/10.3390/cancers13174294>.
78. Schönhuber, N., Seidler, B., Schuck, K., and Saur, D. (2014). A next-generation dual-recombinase system for time- and host-specific targeting of pancreatic cancer. *Nat Med* 20, 1340–1347. <https://doi.org/10.1038/nm.3646>.
79. Ventura, A., Kirsch, D.G., McLaughlin, M.E., Tuveson, D.A., Grimm, J., Lintault, L., Newman, J., Reczek, E.E., Weissleder, R., and Jacks, T. (2007). Restoration of p53 function leads to tumour regression in vivo. *Nature* 445, 661–665. <https://doi.org/10.1038/nature05541>.
80. Eaglesham, J.B., Pan, Y., Kupper, T.S., and Kranzusch, P.J. (2019). Viral and metazoan poxins are cGAMP-specific nucleases that restrict cGAS-STING signalling. *Nature* 566, 259–263. <https://doi.org/10.1038/s41586-019-0928-6>.
81. Fu, Y.-Z., Guo, Y., Zou, H.-M., Su, S., Wang, S.-Y., Yang, Q., Luo, M.-H., and Wang, Y.-Y. (2019). Human cytomegalovirus protein UL42 antagonizes cGAS/MITA-mediated innate antiviral response. *PLoS Pathog* 15, e1007691. <https://doi.org/10.1371/journal.ppat.1007691>.
82. Scrima, N., Le Bars, R., Nevers, Q., Glon, D., Chevreux, G., Civas, A., Blondel, D., Lagaudrière-Gesbert, C., and Gaudin, Y. (2023). Rabies virus P protein binds to TBK1 and interferes with the formation of innate immunity-related liquid condensates. *Cell Rep* 42, 111949. <https://doi.org/10.1016/j.celrep.2022.111949>.
83. Brzózka, K., Finke, S., and Conzelmann, K.-K. (2005). Identification of the rabies virus alpha/beta interferon antagonist: phosphoprotein P interferes with phosphorylation of interferon regulatory factor 3. *J Virol* 79, 7673–7681. <https://doi.org/10.1128/JVI.79.12.7673-7681.2005>.

84. Ritchie, C., Cordova, A.F., Hess, G.T., Bassik, M.C., and Li, L. (2019). SLC19A1 Is an Importer of the Immunotransmitter cGAMP. *Mol Cell* 75, 372-381.e5. <https://doi.org/10.1016/j.molcel.2019.05.006>.
85. Sivick, K.E., Desbien, A.L., Glickman, L.H., Reiner, G.L., Corrales, L., Surh, N.H., Hudson, T.E., Vu, U.T., Francica, B.J., Banda, T., et al. (2018). Magnitude of Therapeutic STING Activation Determines CD8+ T Cell-Mediated Anti-tumor Immunity. *Cell Rep* 25, 3074-3085.e5. <https://doi.org/10.1016/j.celrep.2018.11.047>.
86. Ramanjulu, J.M., Pesiridis, G.S., Yang, J., Concha, N., Singhaus, R., Zhang, S.-Y., Tran, J.-L., Moore, P., Lehmann, S., Eberl, H.C., et al. (2018). Design of amidobenzimidazole STING receptor agonists with systemic activity. *Nature* 564, 439-443. <https://doi.org/10.1038/s41586-018-0705-y>.
87. Weaver, G.C., Arya, R., Schneider, C.L., Hudson, A.W., and Stern, L.J. (2022). Structural Models for Roseolovirus U20 And U21: Non-Classical MHC-I Like Proteins From HHV-6A, HHV-6B, and HHV-7. *Front Immunol* 13, 864898. <https://doi.org/10.3389/fimmu.2022.864898>.
88. Dirck, A.T., Whyte, M.L., and Hudson, A.W. (2020). HHV-7 U21 exploits Golgi quality control carriers to reroute class I MHC molecules to lysosomes. *Mol Biol Cell* 31, 196-208. <https://doi.org/10.1091/mbc.E19-07-0363>.
89. Hudson, A.W., Howley, P.M., and Ploegh, H.L. (2001). A human herpesvirus 7 glycoprotein, U21, diverts major histocompatibility complex class I molecules to lysosomes. *J Virol* 75, 12347-12358. <https://doi.org/10.1128/JVI.75.24.12347-12358.2001>.
90. Hudson, A.W., Blom, D., Howley, P.M., and Ploegh, H.L. (2003). The ER-luminal domain of the HHV-7 immunoevasin U21 directs class I MHC molecules to lysosomes. *Traffic* 4, 824-837. <https://doi.org/10.1046/j.1398-9219.2003.0137.x>.
91. Wiertz, E.J., Jones, T.R., Sun, L., Bogyo, M., Geuze, H.J., and Ploegh, H.L. (1996). The human cytomegalovirus US11 gene product dislocates MHC class I heavy chains from the endoplasmic reticulum to the cytosol. *Cell* 84, 769-779. [https://doi.org/10.1016/s0092-8674\(00\)81054-5](https://doi.org/10.1016/s0092-8674(00)81054-5).
92. Wiertz, E.J., Tortorella, D., Bogyo, M., Yu, J., Mothes, W., Jones, T.R., Rapoport, T.A., and Ploegh, H.L. (1996). Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction. *Nature* 384, 432-438. <https://doi.org/10.1038/384432a0>.
93. Coscoy, L., Sanchez, D.J., and Ganem, D. (2001). A novel class of herpesvirus-encoded membrane-bound E3 ubiquitin ligases regulates endocytosis of proteins involved in immune recognition. *J Cell Biol* 155, 1265-1273. <https://doi.org/10.1083/jcb.200111010>.
94. Falvey, E., and Ziff, E. (1983). Sequence arrangement and protein coding capacity of the adenovirus type 2 "i" leader. *J Virol* 45, 185-191. <https://doi.org/10.1128/JVI.45.1.185-191.1983>.

95. Ramke, M., Lee, J.Y., Dyer, D.W., Seto, D., Rajaiya, J., and Chodosh, J. (2017). The 5'UTR in human adenoviruses: leader diversity in late gene expression. *Sci Rep* 7, 618. <https://doi.org/10.1038/s41598-017-00747-y>.
96. Burgert, H.G., and Kvist, S. (1985). An adenovirus type 2 glycoprotein blocks cell surface expression of human histocompatibility class I antigens. *Cell* 41, 987–997. [https://doi.org/10.1016/s0092-8674\(85\)80079-9](https://doi.org/10.1016/s0092-8674(85)80079-9).
97. Lee, J., Manon, V., and Chen, J. (2025). Structurally diverse viral inhibitors converge on a shared mechanism to stall the antigen transporter TAP. *Proc Natl Acad Sci U S A* 122, e2516676122. <https://doi.org/10.1073/pnas.2516676122>.
98. Fischbach, H., Döring, M., Nikles, D., Lehnert, E., Baldauf, C., Kalinke, U., and Tampé, R. (2015). Ultrasensitive quantification of TAP-dependent antigen compartmentalization in scarce primary immune cell subsets. *Nat Commun* 6, 6199. <https://doi.org/10.1038/ncomms7199>.
99. Brunnberg, J., Herbring, V., Günther Castillo, E., Krüger, H., Wieneke, R., and Tampé, R. (2021). Light control of the peptide-loading complex synchronizes antigen translocation and MHC I trafficking. *Commun Biol* 4, 430. <https://doi.org/10.1038/s42003-021-01890-z>.
100. Symington, J.S., Lucher, L.A., Brackmann, K.H., Virtanen, A., Pettersson, U., and Green, M. (1986). Biosynthesis of adenovirus type 2 i-leader protein. *J Virol* 57, 848–856. <https://doi.org/10.1128/JVI.57.3.848-856.1986>.
101. Oliveira, E.R.A., and Bouvier, M. (2019). Immune evasion by adenoviruses: a window into host-virus adaptation. *FEBS Lett* 593, 3496–3503. <https://doi.org/10.1002/1873-3468.13682>.
102. Källsten, M., Gromova, A., Zhao, H., Valdés, A., Konzer, A., Pettersson, U., and Lind, S.B. (2017). Temporal characterization of the non-structural Adenovirus type 2 proteome and phosphoproteome using high-resolving mass spectrometry. *Virology* 511, 240–248. <https://doi.org/10.1016/j.virol.2017.08.032>.
103. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7, 539. <https://doi.org/10.1038/msb.2011.75>.
104. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
105. Letunic, I., and Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research* 52, W78–W82. <https://doi.org/10.1093/nar/gkae268>.
106. Steinegger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* 35, 1026–1028. <https://doi.org/10.1038/nbt.3988>.

107. Michlits, G., Jude, J., Hinterndorfer, M., de Almeida, M., Vainorius, G., Hubmann, M., Neumann, T., Schleiffer, A., Burkard, T.R., Fellner, M., et al. (2020). Multilayered VBC score predicts sgRNAs that efficiently generate loss-of-function alleles. *Nat Methods* *17*, 708–716. <https://doi.org/10.1038/s41592-020-0850-8>.
108. Schambach, A., Galla, M., Modlich, U., Will, E., Chandra, S., Reeves, L., Colbert, M., Williams, D.A., Von Kalle, C., and Baum, C. (2006). Lentiviral vectors pseudotyped with murine ecotropic envelope: Increased biosafety and convenience in preclinical research. *Experimental Hematology* *34*, 588–592. <https://doi.org/10.1016/j.exphem.2006.02.005>.
109. Vogt, B., Roscher, S., Abel, B., Hildinger, M., Lamarre, A., Baum, C., and Von Laer, D. (2001). Lack of Superinfection Interference in Retroviral Vector Producer Cells. *Human Gene Therapy* *12*, 359–365. <https://doi.org/10.1089/10430340150503984>.
110. Rozenblatt-Rosen, O., Deo, R.C., Padi, M., Adelmant, G., Calderwood, M.A., Rolland, T., Grace, M., Dricot, A., Askenazi, M., Tavares, M., et al. (2012). Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. *Nature* *487*, 491–495. <https://doi.org/10.1038/nature11288>.
111. Bian, X., Wiktor, P., Kahn, P., Brunner, A., Khela, A., Karthikeyan, K., Barker, K., Yu, X., Magee, M., Wasserfall, C.H., et al. (2015). Antiviral antibody profiling by high-density protein arrays. *Proteomics* *15*, 2136–2145. <https://doi.org/10.1002/pmic.201400612>.
112. Alzhanova, D., Corcoran, K., Bailey, A.G., Long, K., Taft-Benz, S., Graham, R.L., Broussard, G.S., Heise, M., Neumann, G., Halfmann, P., et al. (2021). Novel modulators of p53-signaling encoded by unknown genes of emerging viruses. *PLoS Pathog* *17*, e1009033. <https://doi.org/10.1371/journal.ppat.1009033>.
113. Tschärke, D.C., Karupiah, G., Zhou, J., Palmore, T., Irvine, K.R., Haeryfar, S.M.M., Williams, S., Sidney, J., Sette, A., Bennink, J.R., et al. (2005). Identification of poxvirus CD8+ T cell determinants to enable rational design and characterization of smallpox vaccines. *J Exp Med* *201*, 95–104. <https://doi.org/10.1084/jem.20041912>.
114. Carey, K.L., Newton, H.J., Lührmann, A., and Roy, C.R. (2011). The *Coxiella burnetii* Dot/Icm system delivers a unique repertoire of type IV effectors into host cells and is required for intracellular replication. *PLoS Pathog* *7*, e1002056. <https://doi.org/10.1371/journal.ppat.1002056>.
115. Lifshitz, Z., Burstein, D., Peeri, M., Zusman, T., Schwartz, K., Shuman, H.A., Pupko, T., and Segal, G. (2013). Computational modeling and experimental validation of the *Legionella* and *Coxiella* virulence-related type-IVB secretion signal. *Proc Natl Acad Sci U S A* *110*, E707–715. <https://doi.org/10.1073/pnas.1215278110>.
116. Lifshitz, Z., Burstein, D., Schwartz, K., Shuman, H.A., Pupko, T., and Segal, G. (2014). Identification of novel *Coxiella burnetii* Icm/Dot effectors and genetic analysis of their involvement in modulating a mitogen-activated protein kinase pathway. *Infect Immun* *82*, 3740–3752. <https://doi.org/10.1128/IAI.01729-14>.

117. Poirson, J., Cho, H., Dhillon, A., Haider, S., Imrit, A.Z., Lam, M.H.Y., Alerasool, N., Lacoste, J., Mizan, L., Wong, C., et al. (2024). Proteome-scale discovery of protein degradation and stabilization effectors. *Nature* 628, 878–886. <https://doi.org/10.1038/s41586-024-07224-3>.
118. Alerasool, N., Leng, H., Lin, Z.-Y., Gingras, A.-C., and Taipale, M. (2022). Identification and functional characterization of transcriptional activators in human cells. *Molecular Cell* 82, 677-695.e7. <https://doi.org/10.1016/j.molcel.2021.12.008>.
119. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j.* 17, 10. <https://doi.org/10.14806/ej.17.1.200>.
120. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357–359. <https://doi.org/10.1038/nmeth.1923>.
121. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630, 493–500. <https://doi.org/10.1038/s41586-024-07487-w>.
122. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H., and Ferrin, T.E. (2023). UCSF CHIMERA-X: Tools for structure building and analysis. *Protein Science* 32, e4792. <https://doi.org/10.1002/pro.4792>.
123. Kitayner, M., Rozenberg, H., Kessler, N., Rabinovich, D., Shaulov, L., Haran, T.E., and Shaked, Z. (2006). Structural basis of DNA recognition by p53 tetramers. *Mol Cell* 22, 741–753. <https://doi.org/10.1016/j.molcel.2006.05.015>.
124. Wang, J.C.K., Baddock, H.T., Mafi, A., Foe, I.T., Bratkowski, M., Lin, T.-Y., Jensvold, Z.D., Preciado López, M., Stokoe, D., Eaton, D., et al. (2024). Structure of the p53 degradation complex from HPV16. *Nat Commun* 15, 1842. <https://doi.org/10.1038/s41467-024-45920-w>.