A novel antiviral function of AMOTL2 enhances the human type I interferon response against Zika virus

Alexandra Willcox

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Michael Emerman, Chair

Julie M. Overbaugh

Adam P. Geballe

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

A novel antiviral function of AMOTL2 enhances the human type I interferon response against Zika virus

Alexandra Willcox

Chair of the Supervisory Committee:

Michael Emerman

Department of Microbiology

Zika virus (ZIKV) re-emerged over the last several decades to cause large outbreaks in the Asian Pacific and Americas. ZIKV disease, though usually mild, can cause severe neurological complications, including developmental defects in infants born to a ZIKV-infected person. Much remains to be learned about ZIKV host-pathogen interactions, including how the innate immune system fights the early stages of infection in human cells. The type I interferon (IFN) response plays a particularly critical role in the innate immune response to viruses, including ZIKV. Type I IFN is released from virus-infected cells as a warning signal to surrounding cells, setting off a chain of signaling events in the target cells, including the phosphorylation of STAT1 and STAT2, their formation of a complex with IRF9, and the translocation of this complex to the nucleus, where it acts as a transcription factor driving the upregulation of hundreds of interferon-stimulated genes (ISGs). ISGs work collectively to potently restrict viral replication.

While the type I IFN response is clearly important in restricting ZIKV, there has not been a systematic assessment of which host genes are responsible for this potent effect. This thesis describes a CRISPR knockout screen to identify genes that contribute to type I IFN restriction of ZIKV in A549 cells. Eleven gene hits were identified, including IFI6, a previously described ISG with described antiviral activity against the *Orthoflavivirus* genus of viruses, of which ZIKV is a member. I showed that inactivation of this gene led to increased ZIKV replication in the presence of IFN-β, confirming its role as a restriction factor against ZIKV. Because previous studies on IFI6 used different cell types than A549 cells, my results with IFI6 knockout in A549s support the idea that IFI6 is an important anti-ZIKV restriction factor in diverse cell types. The top hit in the screen was AMOTL2, a gene with no previously described role in innate immunity. Surprisingly, AMOTL2 is not upregulated by type I IFN in A549 cells, but nonetheless experiments in single-gene knockout cells confirmed its IFN-specific antiviral phenotype, meaning that knockout of AMOTL2 increased ZIKV replication in the presence, but not absence, of type I IFN. I found that inactivation of AMOTL2 caused inhibition of key steps in the IFN signaling pathway including STAT1 phosphorylation and nuclear translocation, which corresponded with blunting of ISG upregulation and increased ZIKV replication. Interestingly, inactivation of AMOTL2 constitutively increased the levels of unphosphorylated STAT1 (U-STAT1), which has been reported to prevent activation of the type I IFN signaling pathway. I propose a model in which the presence of AMOTL2 in the cell suppresses high constitutive levels of U-STAT1, enabling a potent antiviral response to type I IFN.

Interestingly, a known binding partner of AMOTL2, TAZ, has been reported to antagonize multiple innate immune signaling pathways. I found that AMOTL2 and TAZ interact in A549s, and that inactivation of TAZ increases type I IFN restriction of ZIKV in A549 cells, adding to

published work that reported a proviral function of TAZ towards other viruses in other cell types. However, an expression construct of AMOTL2 lacking the TAZ binding domain was still able to rescue AMOTL2's antiviral phenotype, suggesting that AMOTL2's antiviral mechanism does not depend on TAZ binding.

In summary, I have defined a new antiviral factor, AMOTL2, that acts through a novel mechanism to enhance the type I IFN response. These studies define a new pathway of type I IFN restriction of ZIKV.

Acknowledgments

I feel so lucky to have completed my PhD training under the mentorship of Julie Overbaugh. Julie, thank you for your endless support, mentorship, and encouragement—whether on my next experiment or my weekend plans in the mountains. Your leadership has honed an amazing lab environment, where we all feel highly motivated to learn and grow as scientists without sacrificing balance. Thank you for your detailed feedback on every aspect of science—from experimental design, writing, oral presentations, and data interpretation, to collaboration, maintaining optimism and resilience, and so much more. I have grown more than I could have imagined and have emerged from my PhD feeling much more confident and independent as a scientist. Thank you!

To Dara Lehman, thank you for your endless kindness, encouragement, and reassurance. I always feel better after talking to you. To all of the wonderful people I overlapped with in the Overbaugh/Lehman lab, including Meghan Garrett, Joshua Marceau, Carolyn Fish, Nell Baumgarten, Kevin Sung, Haidyn Weight, Mackenzie Shipley, Feli Ruiz, Ryan Yucha, Hannah Itell, Zak Yaffe, Caitlin Stoddard, Caroline Phan, Michelle Lilly, Morgan Litchford, Jamie Guenthoer, Delphine Depierreux, Vrasha Chohan, Sahil Bhatti, Elise Kang, and Rachel Fish—thank you all for sitting through so many practice talks, and for all of the collaborative crosswords, happy hours, lab dinners, and dog meet-ups. I've had so much fun being in the lab with all of you! To Caitlin, thank you for your friendship and advice, for sitting with me through hard times, for all of the dog walks and writing weekends—I'm so lucky to have shared these years in the lab with you. To Ted Gobillot and Caroline Kikawa, thank you for your work on this project and all of your help and advice as I learned the ropes.

I'm grateful to my thesis committee (both the final and and all previous iterations!): Julie Overbaugh, Michael Emerman, Adam Geballe, Daniel Blanco-Melo, Bruce Torbett, Michael Gale Jr., Leslie Goo, and Erick Matsen. Thank you all for being accessible and for your encouragement and advice on my project. Special thanks to Julie, Michael, and Adam for their valuable feedback on this written thesis. To everyone at Thursday Morning Virus Meeting and in the Hutch virology community, thank you for your advice and encouragement over the years. I have learned so much from all of you.

Thank you to the directors of the UW MSTP for all you do: Marshall Horwitz, Mary-Claire King, Stephen Tapscott, and Heather Cheng. Thank you for always having our backs, listening and being receptive to student-led initiatives, and for being so communicative and helpful—I feel so supported in this program and I can't thank you enough. Huge thanks to the UW MSTP administration, including Sara Carlson, Marcie Buckner, and Kacee French—none of this would be possible without your hard work behind the scenes. Huge thank you also to the directors and administrators of UW MCB, including Nina Salama, Celeste Berg, Julian Simon, Maia Low, Alice Ven, Denise Barnes, Carolina Chambers, Andrea Brocato, Maura Do, and Mel Leavens. To Jasmine Gonzalez and Hannah Snider: you are both just so good at your job, and it made my PhD a lot easier! We are so lucky to get to work with you both.

To all of my former mentors—Steve Meshnick, Matthew Collins, Lionel Almeras, Kazutoyo Miura, and Carole Long—thank you for taking a chance on me and investing so much time in making me a better scientist. I would like to especially thank Steve for letting me chase my scientific whims as an undergrad, letting me experience all the thrill and challenges of the scientific process. It worked—I was immediately hooked. I would not have chased this dream

without your encouragement, and I wish you were still here to celebrate this milestone with me today.

To all my friends near and far—I am so lucky to have you all in my life. To the E-19 MSTP cohort, I feel so lucky to get to call you my classmates and best friends, and to be sharing this journey with you. To Rechel Geiger, for always checking in on me, and for your endless love and encouragement. To Miguel Paredes, for giving me a shoulder to cry on, for laughing with me, for countless scary movie nights, for all the good and bad times we have faced together. To Linnea Lieth, for always being there for me, even from far away.

Thank you to my parents, Bill and Smaranda Willcox, for believing I was capable long before I did, and for always encouraging me on this path. To my brother, Mihnea Mangalea, for your kindness and support over the years—I'm so glad we became friends even though I annoyed you endlessly as a kid. To my partner, Nick Franko, for your unwavering love, commitment, encouragement, and validation. I couldn't have done this without you. To Nick's family, especially his parents, Stephanie Bower and Rich Franko, for your generosity and kindness over the last few years. And last but not least, to my dog and best friend, Coyote: thank you for your unconditional love and for all the adventures we've shared together. I love you!

Dedication

To my father, Bill Willcox. You have been with me every step of the way.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Basics of Zika virus: a summary overview of the topic

Zika virus (ZIKV) is a positive-sense, single-stranded RNA virus in the *Orthoflavivirus* genus within the *Flaviviridae* family, related to Spondweni virus, dengue virus (DENV), yellow fever virus (YFV), West Nile virus, and Japanese encephalitis virus (Figure 1-1). ZIKV shares enough similarity to other flaviviruses that serological diagnosis is complicated by cross-reactive antibodies in individuals with prior flavivirus exposure, which is a particular issue in DENV-endemic areas where ZIKV now spreads [32]. ZIKV was originally isolated in Africa, and strains are phylogenetically classified in the African lineage or the more recently evolved Asian/American lineage (Figure 1-1). ZIKV is an arbovirus that is transmitted between vertebrate hosts by *Aedes spp.* mosquitoes in sylvatic and urban transmission cycles. In the sylvatic cycle, ZIKV circulates between arboreal *Aedes* mosquitoes and non-human primates in forested areas, with occasional human spillover events. In the urban cycle, domestic species of *Aedes* transmit ZIKV between humans in densely populated areas, causing large human outbreaks [1].

ZIKV is a pathogen of great global health concern. The clinical outcome of infection with ZIKV is variable, with most infected adults exhibiting asymptomatic or mild disease, but a subset of individuals experiencing severe neurological outcomes. ZIKV is of particular concern during pregnancy as it can undergo maternal-fetal transmission, causing microcephaly and other developmental defects in infants born to someone infected with the virus during their pregnancy [2]. In addition to mosquito-borne and maternal-fetal transmission, ZIKV can also be transmitted through blood transfusion and sexual contact [3].

Though it was first isolated in the 1940s, ZIKV rarely caused human disease until the 21st century, with a series of increasingly frequent human outbreaks culminating in a ZIKV pandemic in South and Central America in 2016, as described in more detail below. ZIKV is currently endemic to many areas and is predicted to continue causing large outbreaks in humans. Despite this ongoing threat, much remains unknown about ZIKV host-pathogen interactions and the human immune response to ZIKV infection, and we have no approved vaccines or other new tools to combat its spread. Thus, there is a need to understand host factors that regulate ZIKV infection as an approach to understand how we can curb ZIKV spread and pathogenesis.

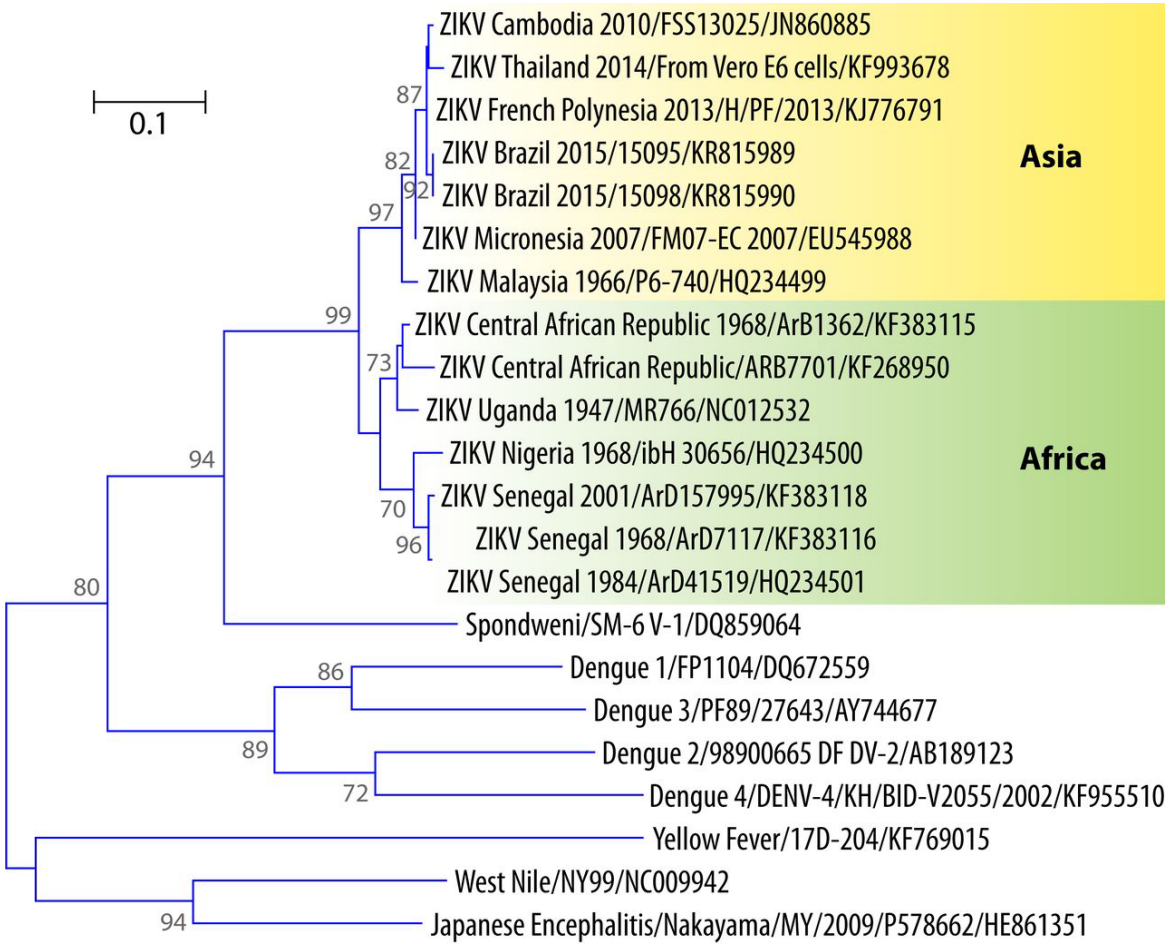

Figure 1-1. ZIKV phylogeny showing its African and Asian/American lineages and relation to other orthoflaviviruses.
Figure from: Musso, D., & Gubler, D. J. (2016). Zika Virus. *Clinical Microbiology Reviews*, *29*(3), 487–524. https://doi.org/10.1128/CMR.00072-15. CC license: 1588539-1

1.2 Zika virus history and epidemiology

Discovery in Africa

In 1947, as part of a YFV surveillance program funded by the Rockefeller Foundation, rhesus macaques were kept in cages in the canopy of the Zika forest of Uganda and monitored for signs of infection. On April 20th, Rhesus 766 developed a fever, and a sample of the monkey's blood contained the very first evidence of a new arbovirus, dubbed "Zika" for the location of its discovery (Figure 1-2) [4]. ZIKV was the seventh novel virus to be discovered as part of the surveillance program over a ten-year period in Uganda, beginning with West Nile Virus in 1937 [2].

Over the next several decades, active ZIKV infection was reported very sporadically in humans, with the first three confirmed human cases on the African continent reported in Nigeria in 1954 [5] and the first human cases in Asia reported in Indonesia in 1977 [6]. Between its discovery in 1947 and the first large human outbreak in 2007, there were only 14 confirmed human cases of ZIKV [2]. Serosurveys from the 1940s onward paint a slightly different picture, suggesting varying levels of ZIKV endemicity in humans across much of the African continent and some countries in Asia [2]. Though the exact numbers should be interpreted with caution due to the potential for cross-reactive antibodies within the *Orthoflavivirus* genus, a large body of evidence suggests that ZIKV has been infecting humans on the African continent for some time. Coupled with the exceedingly low detection of active infections, the finding of high seropositivity suggests that ZIKV infections were either mild and therefore undetected or else frequently misdiagnosed as malaria or other common causes of febrile illness, which is one

outcome of ZIKV infection. With little cause for concern regarding ZIKV's pathogenicity, no papers were published on the virus between 1983 and 2006 [7].

Outbreaks in the Pacific Islands

In April 2007, physicians on Yap Island in the Federated States of Micronesia reported an uptick in patients presenting with symptoms of fever, rash, conjunctivitis, arthritis, and arthralgia [8]. Despite some serum samples testing positive for DENV IgM antibodies, suggesting recent DENV infection, the clinical presentation of the outbreak was distinct enough from DENV to warrant further investigation. The patients' serum was found to contain ZIKV RNA, while testing negative for nucleic acids of other arboviruses, including DENV. By the end of the outbreak, an estimated 73% of the population of Yap Island aged 3 or older (~5,000 individuals) had been infected with ZIKV based on household serosurveys that tested for ZIKV IgM antibodies, which are an indication of recent infection [8]. Of the ~5,000 estimated infections, 18% reported a clinical illness that was likely attributable to ZIKV. Though no hospitalizations or deaths were reported, the scale of the outbreak alarmed the biomedical community and renewed scientific interest in ZIKV.

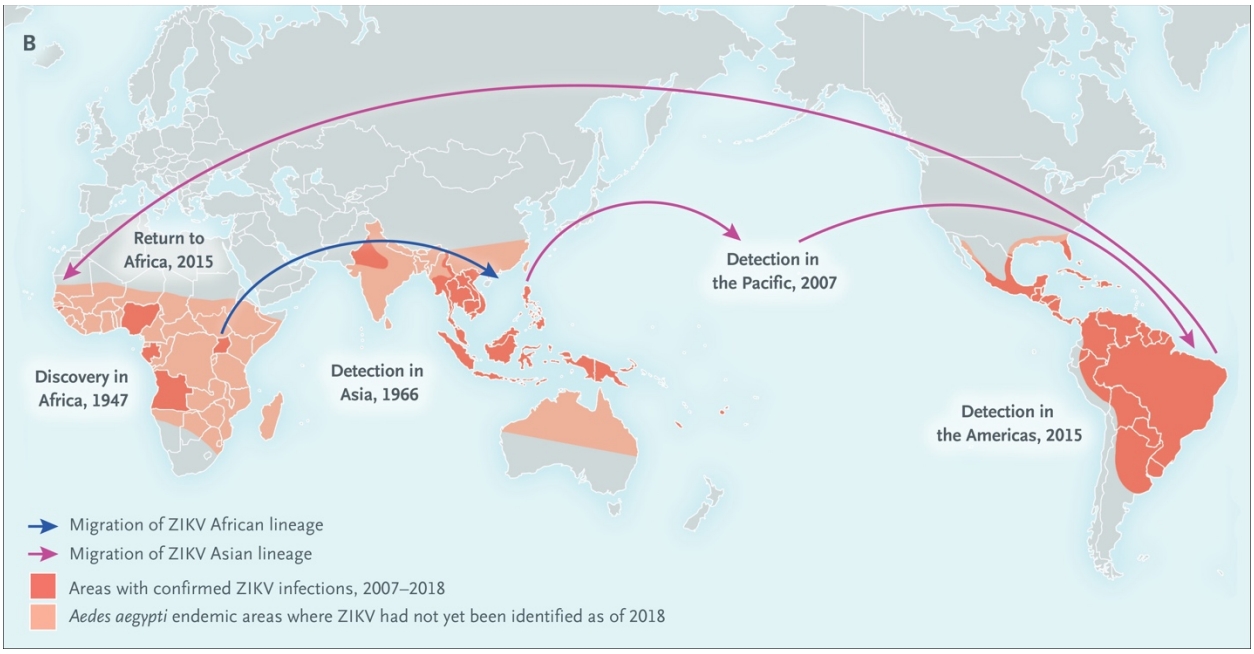

Figure 1-2. Map of known ZIKV spread since its discovery in 1947.
Reproduced with permission from Musso, D., Ko, A. I., & Baud, D. (2019). Zika Virus Infection - After the Pandemic. *The New England Journal of Medicine*, 381(15), 1444–1457. https://doi.org/10.1056/NEJMra1808246. Copyright Massachusetts Medical Society.

The next large outbreak of ZIKV occurred in 2013 in French Polynesia and, similar to the outbreak on Yap Island, resulted in much of the population seroconverting (49%) [9]. The French Polynesian outbreak was important for our understanding of the virus' pathogenicity and non-vector-borne transmission routes in several regards [10]:

- It marked the first known cases of Guillain-Barré syndrome attributed to ZIKV infection, and the incidence of Guillain-Barré in French Polynesia was 20-fold higher than baseline by the end of the outbreak. (I describe Guillain-Barré in section 1.2.)
- The outbreak saw the first detection of ZIKV in semen (suggesting possible sexual transmission).

- The first known maternal-fetal transmission events occurred during the outbreak, though the link between ZIKV infection and developmental defects in infants was not realized until several years later.

In 2014 and 2015, ZIKV continued to cause outbreaks in the Pacific Islands of New Caledonia, Easter Island, Fiji, and others. In contrast to Yap Island and French Polynesia, these nearby outbreaks were generally small. These differences underscore how we are still striving to understand which factors dictate ZIKV spread, such as prior immunity to other flaviviruses, differences in *Aedes* mosquito vector populations, and ZIKV evolution.

Emergence in the Americas

ZIKV is estimated to have landed in the Americas in 2013, more than one year before the recognition of an outbreak in Brazil [11] (Figure 1-2, Figure 1-3). In late 2014 and early 2015, northeastern Brazil reported outbreaks of "exanthematic disease" affecting thousands of patients, and the first locally acquired ZIKV infection of the Americas was confirmed in May 2015. By the end of that year, suspected ZIKV cases were estimated between 440,000 – 1,300,000, in stark contrast to the scale of prior Asian outbreaks [12]. The American outbreak was caused by a new subclade of ZIKV descended from an Asian strain (Figure 1-1) [13].

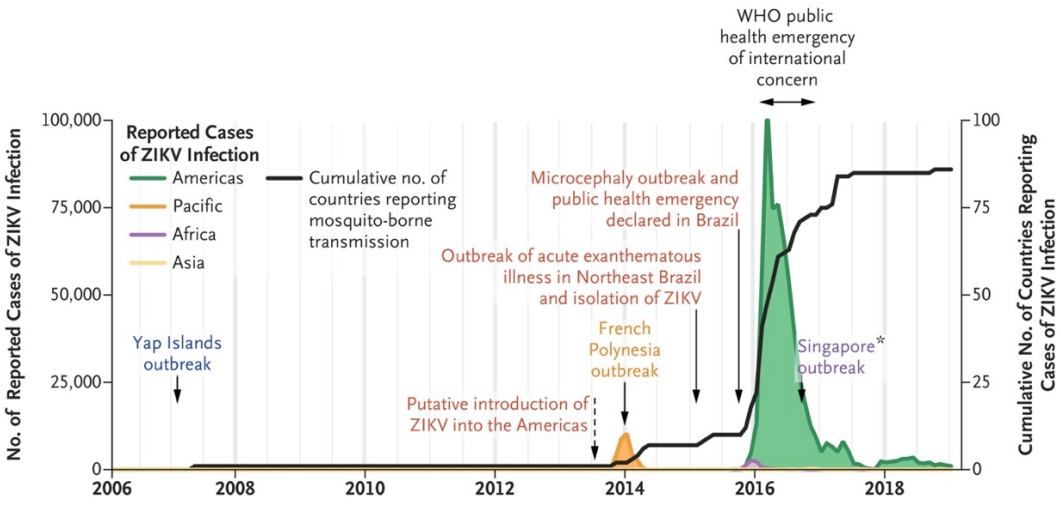

Figure 1-3. Timeline of ZIKV spread from January 2007 through December 2018.
Reproduced with permission from Musso, D., Ko, A. I., & Baud, D. (2019). Zika Virus Infection - After the Pandemic. *The New England Journal of Medicine*, 381(15), 1444–1457. https://doi.org/10.1056/NEJMra1808246. Copyright Massachusetts Medical Society.

By the time the World Health Organization declared a Public Health Emergency of International Concern in February 2016, ZIKV had been detected in 22 countries in Latin America and the Caribbean [14], and there was strong evidence for a link between ZIKV infection during pregnancy and severe birth defects in infants [15]. Modeling has estimated that by the time the pandemic waned at the end of 2018, 132 million people across the Americas had been infected with ZIKV [16], and thousands of families were caring for a child with severe developmental defects caused by the virus. The rapid spread and severe phenotypes of ZIKV in the Americas were likely the result of a combination of competent mosquito vectors in highly populated urban areas, a uniformly susceptible population, and possibly evolution of the virus, though the latter is still not well understood.

Current situation and the threat of future outbreaks

It is largely accepted that the development of herd immunity in human populations of South and Central America likely drove the end of the pandemic in 2018, though pockets of susceptible populations unaffected by the pandemic maintain low-level endemic transmission in the region [3,17]. In addition to its continued presence in the Americas, ZIKV also causes sporadic outbreaks in India [18] and continues to circulate in Africa [2,19,20]. Like other arboviruses such as DENV, YFV, and chikungunya virus, large ZIKV outbreaks can be expected to occur cyclically with the waxing and waning of herd immunity (i.e., as people born after the pandemic make up an increasing proportion of the population with time) [17]. A modeling study using data from Nicaragua predicted that the risk of a new ZIKV outbreak will surpass 50% by

2047, assuming lifelong immunity in individuals infected during the pandemic, which is a conservative assumption [21]. Indeed, a recent study provides evidence for ZIKV re-infection events, potentially challenging the assumption of lifelong sterilizing immunity [22]. Another study using data from the 2013-2014 French Polynesia outbreak predicted re-emergence in that population between 2026 and 2034 [23]. Furthermore, climate change will increase suitable habitats for ZIKV's *Aedes* mosquito vectors, potentially exposing up to 1.3 billion new individuals to suitable temperatures for transmission by 2050 [24] (Figure 1-4). Despite this growing threat, we have no approved ZIKV vaccines or therapeutics to prepare us for the next outbreak.

Figure 1-4. Populations newly at risk for ZIKV transmission by 2050.
Red circles indicate millions of new individuals predicted to be exposed to suitable temperatures for ZIKV transmission by 2050 assuming worst-case scenario emissions. Figure from: Ryan, S. J., Carlson, C. J., Tesla, B., Bonds, M. H., Ngonghala, C. N., Mordecai, E. A., Johnson, L. R., & Murdock, C. C. (2021). Warming temperatures could expose more than 1.3 billion new people to Zika virus risk by 2050. *Global change biology*, *27*(1), 84–93. https://doi.org/10.1111/gcb.15384. CC license: 5987850858493

1.3 Clinical manifestations and pathophysiology of Zika virus infection

As discussed above, most ZIKV infections are mild or asymptomatic. Those with mild disease usually exhibit fever, rash, arthralgia, myalgia, and conjunctivitis and recover without complications [3]. However, when complications do occur, they can be severe. For example, ZIKV is a trigger for Guillain-Barré syndrome, which occurs in 2 to 3 cases per 10,000 ZIKV

infections [3]. Guillain-Barré is an acute immune-mediated polyneuropathy that disrupts motor function. Guillain-Barré is the leading global cause of non-traumatic paralysis and tends to occur following certain viral infections [25]. The pathophysiology is incompletely understood. It has been suggested that since ZIKV directly infects and causes apoptosis of neurons, antigen presenting cells that engulf and present proteins from these infected neurons will trigger high levels of autoantibodies against neuronal proteins (specifically gangliosides) [25].

ZIKV infection poses a particularly high risk during pregnancy (Figure 1-5). Congenital Zika syndrome is a spectrum of developmental defects that may appear in infants born to someone infected with the virus during pregnancy, and includes microcephaly in addition to other brain and skull abnormalities, ocular manifestations, seizures, vision loss, hearing loss, movement abnormalities, and various other symptoms related to disrupted central nervous system development [3]. Vertical transmission occurs in 20-30% of pregnant people infected with ZIKV, and of those infected fetuses, congenital Zika syndrome occurs in 5-14% [3]. The risk is highest when infection occurs during the first trimester. We continue to learn about the long-term neurological outcomes by following the cohort of infants born during the 2015-2016 pandemic, as several studies have suggested that neurodevelopmental abnormalities can become apparent in the first few years of life even if defects were not present at birth [26-28].

Fetal loss occurs in a subset of ZIKV infections during pregnancy. Interestingly, several lines of evidence in mice and rhesus macaque infection models suggest that African-lineage ZIKV strains are more pathogenic to the developing fetus, tending to cause a high proportion of silent miscarriages rather than live births with visible developmental abnormalities associated with Asian and American strains [29-31]. This suggests that the apparent reduction in ZIKV's

pathogenicity in its evolution from African-lineage to Asian- and American-lineage strains may have paradoxically opened our eyes to its propensity to cause severe birth outcomes.

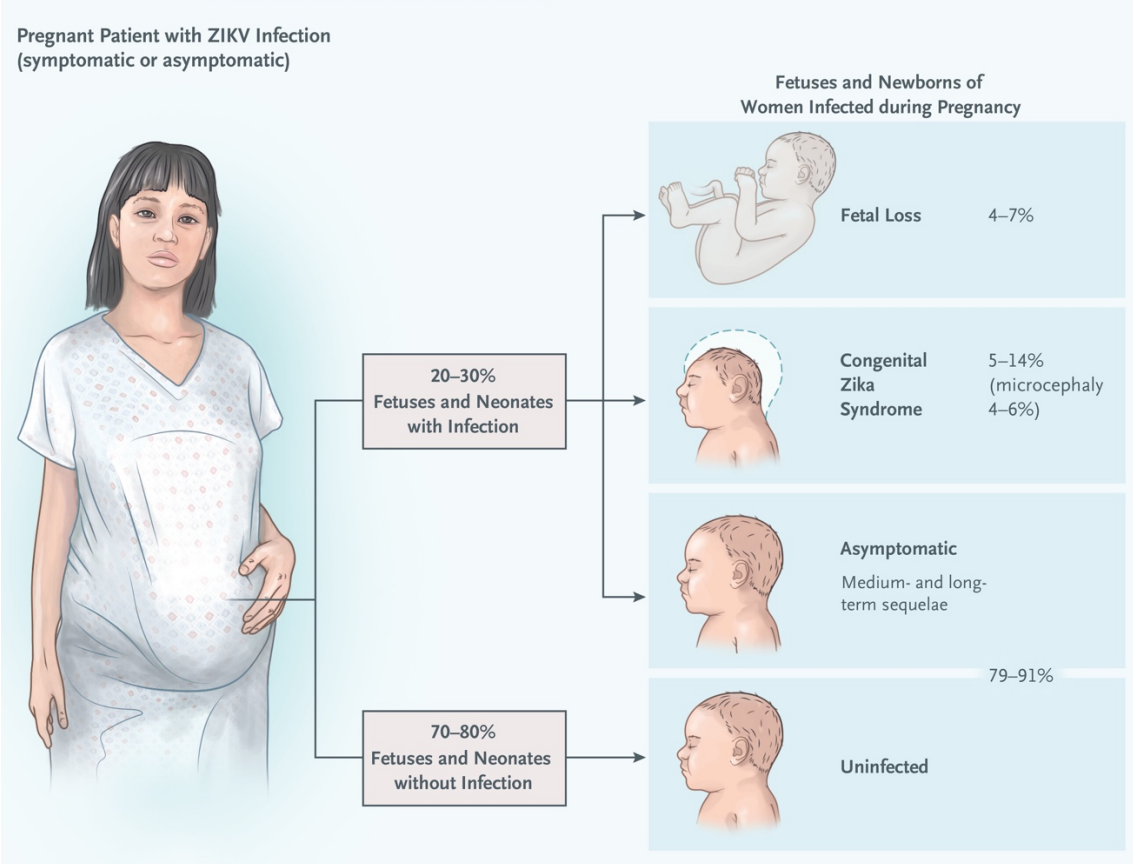

Figure 1-5. Estimated proportions of different clinical outcomes following ZIKV infection in pregnancy. Reproduced with permission from Musso, D., Ko, A. I., & Baud, D. (2019). Zika Virus Infection - After the Pandemic. *The New England Journal of Medicine*, 381(15), 1444–1457. https://doi.org/10.1056/NEJMra1808246. Copyright Massachusetts Medical Society.

1.4 Zika virus life cycle

The ZIKV genome is ~11kb and contains a single open reading frame (ORF) encoding a polyprotein that is subsequently cleaved into 10 proteins by host and viral proteases (Figure 1-6). The ORF is flanked by two untranslated regions (UTRs). The main functions of ZIKV's three structural and seven non-structural proteins are described in Table 1-1 [33-36].

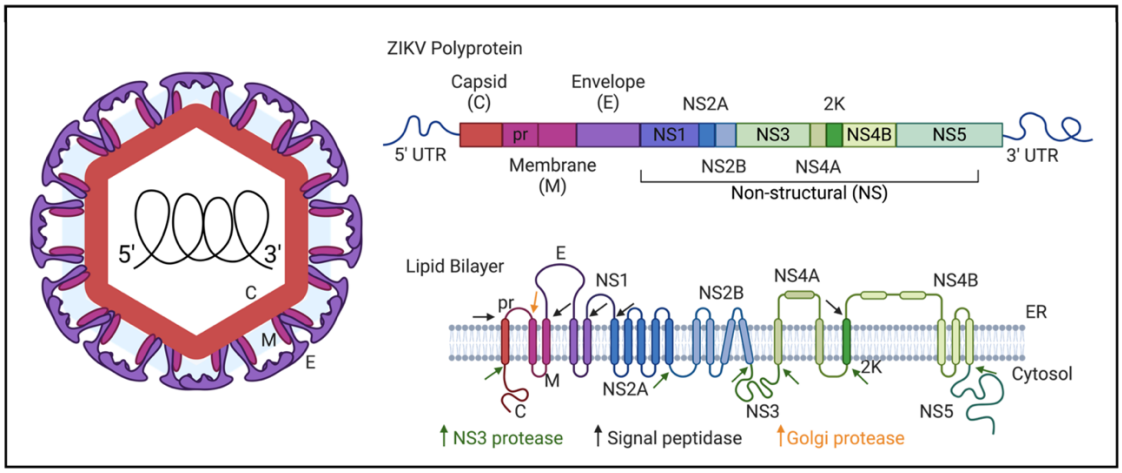

Figure 1-6. Organization of ZIKV's genome.
Figure shows the organization of ZIKV's genome and sites of processing by host and viral proteases. Figure from: Metzler, A.D. & Tang, H. (2024). Zika Virus Neuropathogenesis—Research and Understanding. *Pathogens*, 13(7), 555. https://doi.org/10.3390/pathogens13070555.

Structural proteins	
Premembrane (prM)	With E protein, makes up the outer membrane of the virion. Host furin-like proteases cleave prM into mature membrane (M) in the acidic environment of the trans-Golgi network, yielding the mature and fully infectious virion.
Envelope (E)	Class II fusion protein. With prM/M, makes up the outer membrane of the virion. Facilitates fusion with the host membrane. Major target of neutralizing antibodies.
Capsid (C)	Binds and encapsulates viral genomic RNA.
Non-structural proteins	
NS1	Essential for RNA replication (remodels the ER membrane). Secreted NS1 in the bloodstream is essential for infectivity of mosquitoes during a blood meal. Capable of binding endothelial cells and causing vascular dysfunction in the umbilical vein and brain endothelium, contributing to severe disease.
NS2A	Involved in assembly of viral replication complexes.
NS2B	Involved in assembly of viral replication complexes. Cofactor for NS3.
NS3	N-terminus is a protease that helps to process the viral polyprotein after translation. C-terminus is a helicase that separates the daughter strand from the template strand during viral genome replication
NS4A	Involved in RNA replication. Cofactor for NS3.
NS4B	Involved in RNA replication.
NS5	RNA-dependent RNA polymerase. Has methyltransferase activity to add 5' cap on viral mRNA to mimic host mRNA.

Table 1-1. ZIKV proteins and their functions.

ZIKV binds a target cell by interacting with host attachment factors such as DC-SIGN, TIM-1, and members of the TAM family. The presence of these host factors increases the efficiency of ZIKV binding and infection, but they are not receptors in the traditional sense; i.e., they are not strictly required for infection as they do not induce necessary conformational changes in the virus, but rather increase the likelihood of viral entry by clathrin-mediated endocytosis [36,37]. pH-dependent fusion between the viral and endosomal membrane results in release of the genomic RNA into the host cell cytoplasm. On the ER membrane, the positive-sense RNA is translated into a polyprotein and cleaved by the NS3 viral protease and host proteases to yield the 10 ZIKV proteins (Table 1-1). Accumulation of viral proteins is thought to stimulate a shift from translation to replication of the viral RNA [36,37]. The non-structural proteins contribute to remodeling of the ER membrane, forming replication organelles where double-stranded RNA genomes are synthesized. These invaginations in the ER membrane help to shield the replicating virus from detection by the host innate immune system and increase the local concentration of viral components. Immature virions bud from the ER and move through the trans-Golgi network, where the acidic environment triggers prM cleavage and maturation of the virion, which is released from the cell by exocytosis. The steps of the ZIKV replication cycle are summarized in Figure 1-7.

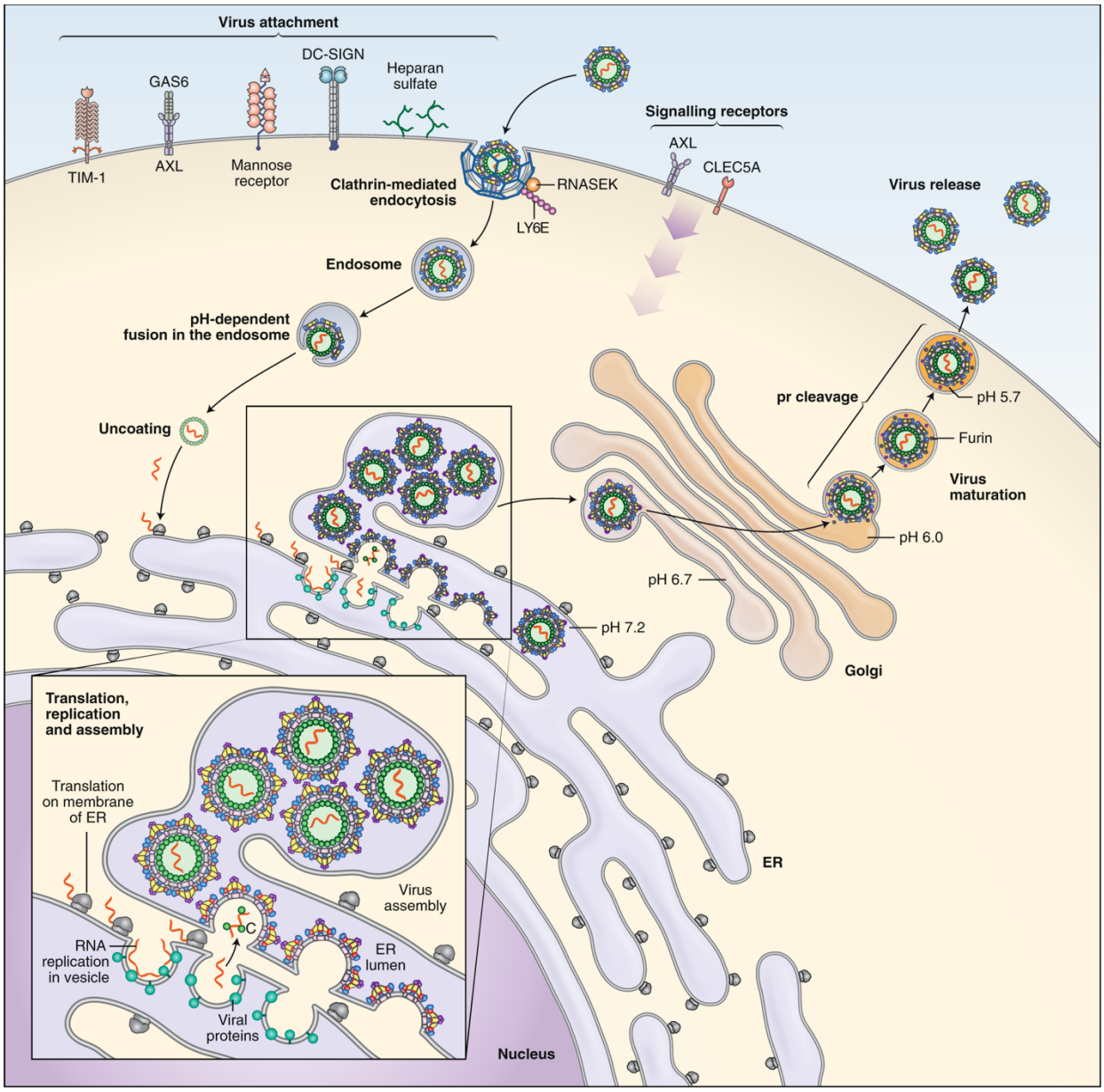

Figure 1-7. The flavivirus replication cycle.
Figure from: Pierson, T.C., Diamond, M.S. (2020). The continued threat of emerging flaviviruses. *Nature Microbiology*, 5, 796–812. https://doi.org/10.1038/s41564-020-0714-0. CC license: 5987851452067

1.5 The type I interferon response

The type I interferon (IFN) response is an arm of the vertebrate innate immune system that provides a defense system against all viruses, including ZIKV. It constitutes an extremely rapid response to infection that helps to curb the initial spread of the virus. Human cells detect viral infection through specific pattern recognition receptors (PRRs), including the endosomal

toll-like receptors 3 (TLR3) and 7 (TLR7), the cytoplasmic RIG-like receptors RIG-I and MDA5, and the cGAS-STING pathway [38]. Virus detection by these PRRs sets off signaling pathways that converge on the phosphorylation of transcription factors IRF3 and IRF7, which then bind to the promoters of IFN genes, causing their transcription and translation. IFN is released from the cell and binds its cell-surface receptor in an autocrine and paracrine fashion, setting off an intracellular signaling pathway that sets up antiviral defense systems in the target cell.

The major type I IFNs in humans are IFN-β and the 13 IFN-α subtypes; there are, in addition, the lesser-studied IFN-ε, IFN-κ, and various other IFN genes. IFN-α and IFN-β are the best-studied and most broadly expressed subtypes [39], and of these, IFN-β is often considered the "primary" type I IFN for its exceptionally strong affinity for the type I IFN receptor [40]. The other major classes of IFN are type II and III IFNs, and they signal through distinct receptors from type I IFN and from each other. Type II IFN (IFN-γ) is proinflammatory and has immunomodulatory properties that are quite distinct from type I IFNs. Type III IFN (IFN-λ) is expressed exclusively at epithelial barriers and exerts similar antiviral effects to the type I IFNs, though with less potency [41].

The heterodimeric type I IFN receptor, composed of IFNAR1 and IFNAR2 subunits, is a transmembrane receptor that is expressed by nearly all cells in the human body. The intracellular tails of the receptor subunits are associated with the JAK tyrosine kinases JAK1 and TYK2. In the canonical type I IFN signaling pathway (Figure 1-8), binding of IFN to the receptor triggers the auto-phosphorylation of JAK1 and TYK2, leading to the phosphorylation of the intracellular tails of the IFNAR receptor subunits, which promotes the recruitment of STAT1 and STAT2. STAT1/2 are phosphorylated by JAK1 and TYK2, and the phospho-STAT1/phospho-STAT2

complex comes together with IRF9, forming a complex known as IFN-stimulated gene factor 3 (ISGF3). The ISGF3 complex translocates to the nucleus and binds conserved sequence motifs called IFN-stimulated response elements (ISREs) in gene promoters. The genes that are upregulated in response to ISRE binding are collectively referred to as IFN-stimulated genes (ISGs) [39,42].

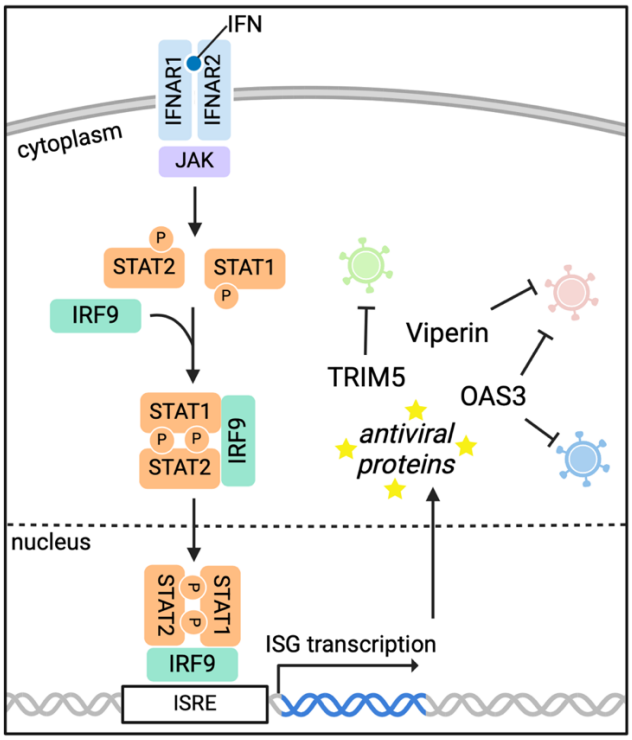

Figure 1-8. The canonical type I interferon signaling pathway.
Created using BioRender.

Hundreds of ISGs with diverse functions are upregulated in response to IFN, and their collective action creates an antiviral state that protects the cell from infection [43]. The upregulation of ISGs is non-specific (i.e., all ISGs are upregulated regardless of which virus in infecting the cell), but there is specificity in which ISGs contribute to inhibition of which pathogens (Figure 1-8) as some ISGs have evolved to directly inhibit the replication cycle of a certain pathogen. For example, TRIM5 is a well-characterized ISG that is specific for retroviruses, due to its direct recognition and destabilization of the retroviral core through

interactions with capsid protein assemblies (CA) [44]. Other antiviral ISGs have mechanisms that make them less specific; for example, the oligoadenylate (OAS) enzymes recognize double-stranded RNA (generated during genome replication by diverse viruses) and activate RNase L to non-specifically degrade RNA. Due to its widely applicable mechanism, OAS enzymes restrict viruses across diverse families [45,46]. In addition to direct viral inhibitors, ISGs also include positive regulators of IFN signaling such as STAT1/2, which both regulate and are regulated by IFN, and the pattern recognition receptors MDA5 and RIG-I, as well as the negative regulators SOCS and USP18 [43].

1.6 Interactions between the human type I interferon response and Zika virus

The IFN response is active against ZIKV infection, as evidenced by extensive *in vitro* data on IFN's potent effect against ZIKV [47-49] and the fact that mouse models of ZIKV pathogenesis require ablation of IFN signaling for productive infection [50-52]. Several studies have revealed some of the ISGs that contribute to IFN-mediated restriction of ZIKV. The ISGs that show evidence of inhibiting ZIKV replication include IFI6 [53,54], Viperin [55,56], PARP11 and PARP12 [57,58], the IFITMs [59-61], ISG15 [62,63], ISG20 [64], SAT1 [65], cholesterol-25-hydroxylase [66], XAF1 [67], Shiftless (C19orf66) [68,69], SERTAD3 [70], Schlafen 11 [71], TRIM22 [72], and CMPK2 [73]. While these studies have greatly improved our understanding of how type I IFN restricts ZIKV, knowledge gaps remain. Most notably, almost all of the genes listed above were chosen for study based on existing knowledge of their role as viral restriction factors rather than unbiased systematic screening. Additionally, some of these ZIKV-restricting host genes were identified through overexpression studies [54,74], which may limit the biological relevance of the results, and has resulted in some findings not being reproducible. This was demonstrated for IFITM3, which was shown in overexpression studies to

inhibit ZIKV but did not have an effect when knocked out or expressed at endogenous levels [47].

One prior study performed a knockdown screen in microglial cells by repurposing a small siRNA library of 386 genes that was enriched for ISGs [75]. In addition to identifying some ISGs that modulate ZIKV infection, this study also unexpectedly identified a non-ISG, MTA2, that appears to contribute to type I IFN restriction of ZIKV. This finding suggests that the focus on ISGs alone may be limiting, as non-ISGs can also contribute to type I IFN restriction of ZIKV. A better understanding of the larger milieu of genes relevant to IFN-mediated restriction of ZIKV, including ISGs and non-ISGs alike, would improve our understanding of ZIKV host-pathogen interactions and inform drug development.

While vertebrate evolution has clearly resulted in a robust type I IFN response against viruses including ZIKV, so too has ZIKV developed a plethora of ways to antagonize host immunity through both passive and active mechanisms. For example, ZIKV NS5 encodes a methyltransferase that adds a methyl group to the 5' viral mRNA cap. This prevents IFIT protein family members, antiviral ISGs that bind and sequester uncapped viral RNA, from recognizing ZIKV mRNA as foreign [76]. Another passive method used by ZIKV and other flaviviruses is the sequestration of their dsRNA replication intermediates in virus-induced replication organelles in the endoplasmic reticulum membrane, where they are shielded from detection by the innate immune system. Each of ZIKV's non-structural proteins have also evolved secondary roles that actively target various steps upstream of type I IFN transcription in order to reduce IFN production [77-79]. In addition, downstream of type I IFN production, ZIKV NS5 binds and targets human STAT2 for degradation [80].

Despite this multifaceted approach to type I IFN antagonism, the type I IFN response still effectively restricts ZIKV *in vivo,* as supported by the fact that mutations in TLR3 and TLR7 are associated with more severe clinical outcomes in ZIKV-infected newborn humans, suggesting that successful triggering of IFN production has an antiviral role within the human immune response to ZIKV infection [81,82]. Additionally, ablation of IFN signaling in mouse models drastically increases ZIKV pathogenesis, and intact IFN signaling specifically at the placenta has been shown to restrict ZIKV infection and improve outcomes for both mother and fetus [83]. Thus, understanding the mechanisms used by both sides of the host-virus arms race that is ongoing between the human type I interferon response and ZIKV will be informative, with the goal of leveraging those findings to tip the balance further in favor of a potent type I IFN response and target ZIKV therapeutically.

1.7 The role of unphosphorylated STAT1 in the type I IFN response

One aspect of the type I IFN signaling pathway of particular relevance to this thesis is the role of unphosphorylated STAT proteins, and particularly unphosphorylated STAT1 (U-STAT1) in the type I IFN response. Unphosphorylated STAT proteins were historically considered to be latent transcription factors that awaited IFN stimulation in the cytoplasm to undergo phosphorylation and formation of the activating nuclear complex (ISGF3). This view held that they had no functional activity of their own, absent phosphorylation. Now, it is understood that U-STAT1 and unphosphorylated STAT2 (U-STAT2) can interact with IRF9, translocate to the nucleus, and mediate expression of a subset of the genes stimulated by ISGF3 (Figure 1-9). Because STAT1 and STAT2 are themselves target genes of ISGF3, the levels of the unphosphorylated proteins increase in cells in the days following type I IFN treatment. U-STAT1 and U-STAT2 therefore prolong the expression of certain interferon-induced genes long after the

phosphorylated forms that promoted their expression have become undetectable, contributing to a continued antiviral state that lasts for days after resolution of the acute response to type I IFN mediated by phosphorylated STAT1 and STAT2 [84,85].

Although the U-STAT proteins contribute to the antiviral state conferred by type I IFN, their expression can also, somewhat paradoxically, cause reduced responsiveness of cells to type I IFN (Figure 1-9). For example, hepatitis C virus (HCV)-infected livers exhibit constitutive expression of U-STAT1 and the ISGs induced by U-STAT1, and high levels of these basal ISGs correlate with poor responses to IFN-α, which until recently was a first-line treatment for chronic HCV [86,87]. This phenomenon may occur because some of the ISGs upregulated by U-STAT1 are negative regulators of IFN signaling. One study extended this observation mechanistically to show that exogenous overexpression of U-STAT1 in hepatic cells caused reduced activation of STAT1 and STAT2 upon IFN-α treatment, and the extent of the reduction correlated with the length of time that U-STAT1 was overexpressed prior to IFN treatment [87]. Specifically, STAT1 and STAT2 phosphorylation and ISRE reporter activity were all decreased [87]. In a similar vein, long-term priming of cells with low doses of IFNs causes reduced activation of the STAT proteins in response to type I IFN treatment (several examples are reviewed in [88]). Importantly, various studies suggest that the role of U-STAT1 in IFN signaling and pathogen restriction may be cell type- and/or pathogen-specific.

1.8 Goals for this thesis

The overall goal of this thesis was to identify and characterize host genes contributing to the type I IFN response to ZIKV using large-scale screening and a loss-of-function approach. Chapter 2 will describe the results of a large CRISPR screen to identify human genes that contribute to type I IFN-mediated restriction of ZIKV. Specifically, I report the identification of

a new antiviral gene, AMOTL2, that improves type I IFN restriction of ZIKV through an effect on U-STAT1 levels and STAT1 activation in the IFN signaling pathway. Chapter 3 will discuss the implications and future directions related to this work.

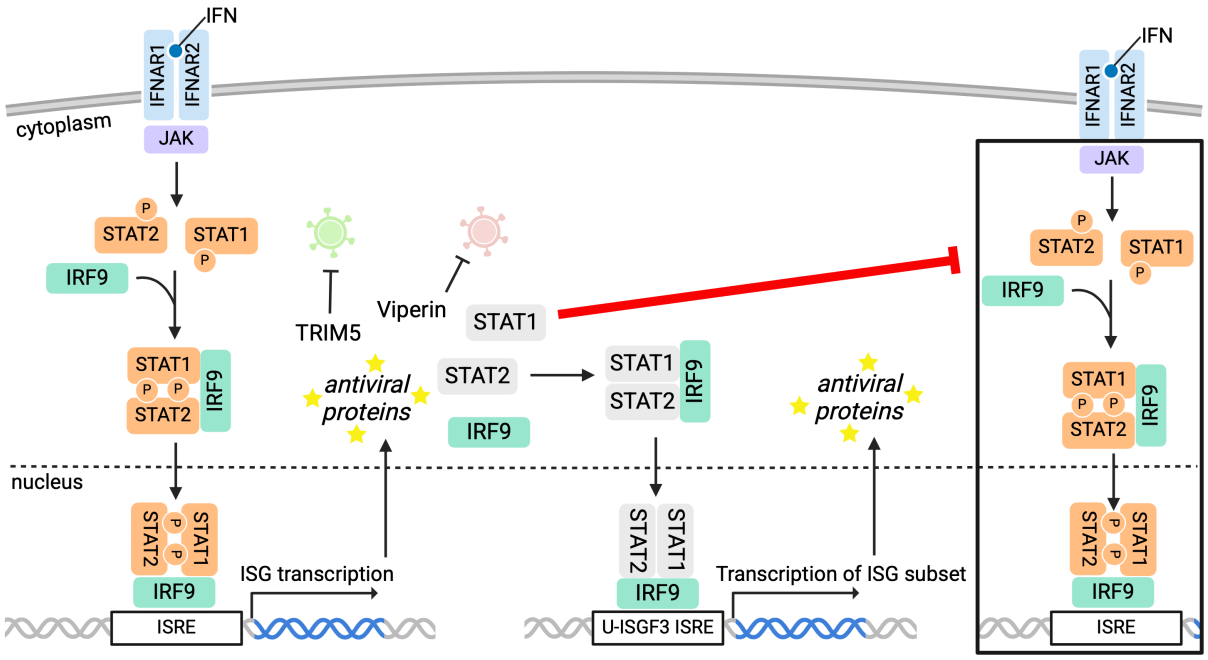

Figure 1-9. Roles for unphosphorylated STAT1 (U-STAT1) in the type I IFN response.
Created using BioRender.

Chapter 2 A novel antiviral function of AMOTL2 enhances the human type I interferon response against Zika virus

This chapter has been modified slightly from the following manuscript:

Willcox, A. C., Gobillot, T.A., Kikawa, C., Baumgarten, N., Stoddard, C.I., Sung, K., Marceau, J., Humes, D., and Overbaugh, J. (2025). A novel antiviral function of AMOTL2 enhances the human type I interferon response against Zika virus. *In preparation.*

2.1 Abstract

Zika virus (ZIKV) has caused multiple human outbreaks, with more recent epidemics associated with severe outcomes in infants. Today, ZIKV is endemic to many countries and presents a persistent threat for future epidemics. The host innate immune proteins that regulate ZIKV replication are incompletely defined. We developed a CRISPR knockout screen to identify host factors that impact ZIKV replication, resulting in the discovery of angiomotin-like protein 2 (AMOTL2), a protein that inhibits ZIKV by regulating the host type I interferon (IFN) response. AMOTL2 affects type I IFN signaling by modulating STAT1 levels and activation in response to type I IFN. Thus, AMOTL2, which has largely been studied for its role in cancer, represents a novel antiviral protein that interacts with the IFN signaling pathway to promote downstream expression of IFN stimulated genes, resulting in restriction of ZIKV.

2.2 Introduction

Zika virus (ZIKV) is a mosquito-borne flavivirus that was first discovered in a rhesus monkey in Uganda in 1947 [4]. While ZIKV caused several human outbreaks in the 20th century, the burden of disease appeared to be limited until ZIKV's emergence in the 21st century in the Pacific Islands (2007-2015) and South America (2015-2016) [2]. In 2016 alone, there were over 500,000 suspected symptomatic cases of ZIKV across the Americas [89]. Though ZIKV infection is usually mild or asymptomatic, it was discovered during the 2016 outbreak that ZIKV

can cause microcephaly and other serious developmental anomalies in prenatal infections [90]. Furthermore, ZIKV can be a trigger for Guillain-Barre syndrome, an autoimmune-mediated paralysis that can cause chronic neurologic impairment [91,92]. Today, ZIKV remains endemic to the Americas, where population-level immunity from the 2015-2016 outbreak is thought to be preventing larger outbreaks [3,21]. ZIKV outbreaks occur sporadically in India [18], and there is evidence that ZIKV circulates at low levels in regions of Africa where the host mosquito vector is present [2,19,20], potentially providing a reservoir for future outbreaks. Thus, ZIKV remains a serious threat to global public health.

Type I interferons (IFNs) are crucial mediators of the innate immune response against viruses, including ZIKV. Type I IFNs, including IFN-α and IFN-β, are cytokines that are released from infected cells upon detection of virus. They bind to the ubiquitously expressed IFNAR1 and IFNAR2 receptor subunits in an autocrine and paracrine response, resulting in receptor dimerization, autophosphorylation of associated JAK tyrosine kinases, and phosphorylation and activation of STAT1 and STAT2 proteins. Activated STAT1 and STAT2 form a complex with IRF9 to make IFN-stimulated gene factor 3 (ISGF3), which translocates to the nucleus to bind IFN-stimulated response elements (ISREs) in gene promoter regions, resulting in the transcription of hundreds of IFN-stimulated genes (ISGs) [93]. A subset of these ISGs encodes proteins that act in diverse ways to limit viral replication, with many ISGs having evolved specificity to certain viruses [39]. Many ISGs that directly inhibit viral proteins have been described [39]. In addition, given the central role of type I IFN in regulating an antiviral state, any host genes that modulate IFN signaling also have the potential to exert a potent effect on viral replication by broadly affecting the expression of some or all antiviral ISGs.

The type I IFN response effectively inhibits ZIKV infection, as evidenced by extensive *in vitro* data on IFN's potent effect against ZIKV [47-49] and the fact that mouse models of ZIKV pathogenesis require ablation of IFN signaling for productive infection [50-52]. Additionally, polymorphisms in toll-like receptors 3 and 7, which normally trigger IFN production, are associated with severe clinical outcomes in ZIKV-infected newborn humans [81,82]. Several studies have characterized ISGs that block ZIKV [53-63,65,66,68], but only one loss-of-function screen has been performed to identify the larger milieu of host factors that contribute to type I IFN inhibition of ZIKV [75].

In this study, we performed a CRISPR knockout screen to identify genes important to type I IFN-mediated restriction of ZIKV. Our screen results validated a known flavivirus restriction factor, IFI6, as an antiviral ISG with activity against ZIKV. We also identified a gene, AMOTL2, that is constitutively expressed yet possesses antiviral activity against ZIKV only when cells are treated with IFN-β. AMOTL2 is known to be involved in angiogenesis, tight junction formation, cell polarity, the actin cytoskeleton, and regulation of Hippo signaling [94], but has no described role in innate immunity. We found that AMOTL2 modulates total STAT1 levels, resulting in enhanced nuclear translocation of phosphorylated STAT1 and heightened ISG transcription. AMOTL2 thus represents a novel host factor contributing to restriction of ZIKV.

2.3 Development of a CRISPR screen based on ZIKV-induced cell death

We optimized a CRISPR knockout screen to identify genes contributing to IFN-β restriction of ZIKV. A549 epithelial cells were selected because they are highly susceptible to IFN-induced inhibition of ZIKV [47]. We repurposed the previously described PIKA CRISPR library of single guide RNAs (sgRNAs) targeting 1,905 putative ISGs identified in diverse cell types [95], including genes identified using microarray data from IFN-stimulated A549 cells

[96]. To further verify that ISGs in A549 cells are well represented in the library, we performed bulk RNA-Seq on A549 cells stimulated with IFN-β for 24 hours. Comparison with genes upregulated at least 4-fold upon IFN-β stimulation of A549 cells indicated that the library contains sgRNAs targeting 84% of these genes (n=174). Thus, the library is enriched for ISGs of interest, but also includes many other genes that are not significantly upregulated by IFN-β in A549 cells (Fig 2-1A). This afforded the potential to identify both ISGs and non-ISGs that contribute to IFN-β restriction of ZIKV.

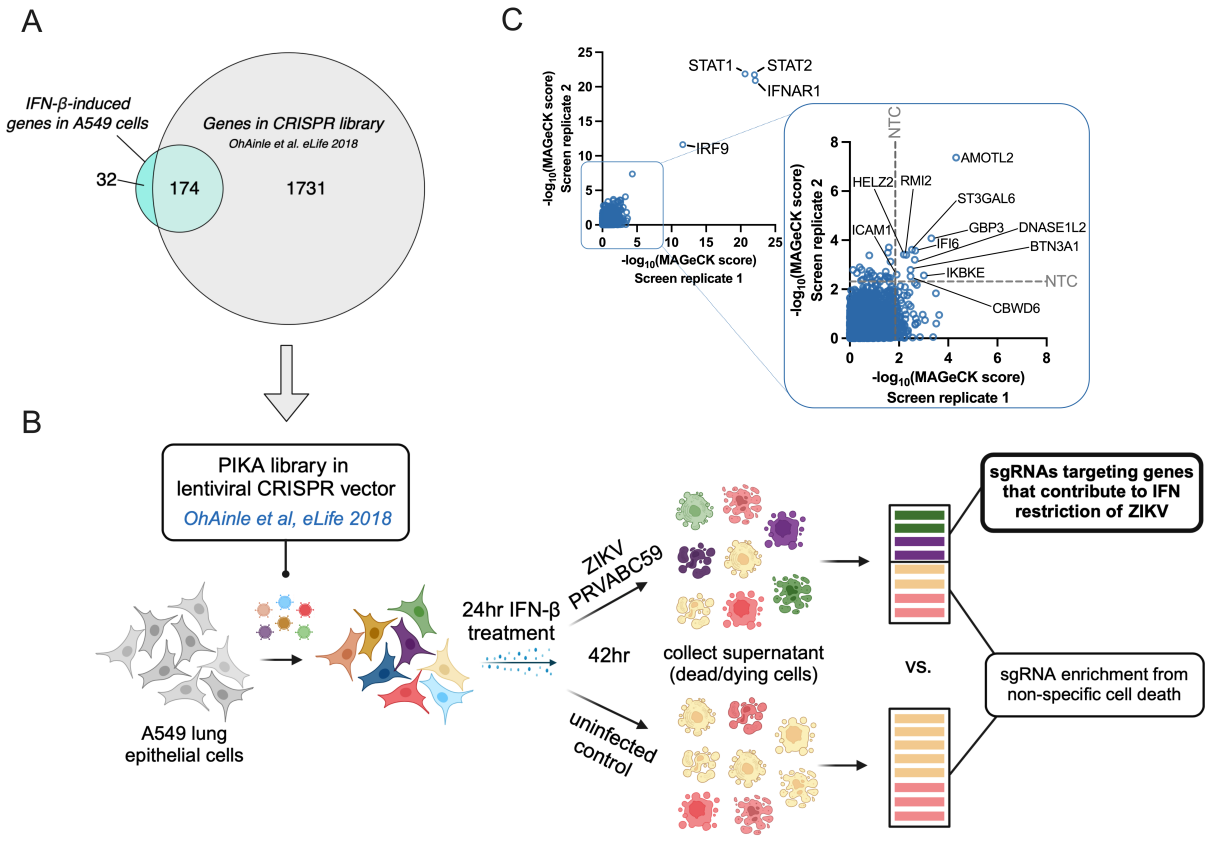

Figure 2-1. Development of a CRISPR knockout screen to identify antiviral genes against ZIKV based on ZIKV-induced cell death.
(A) Overlap between genes in the PIKA library and genes upregulated at least 4-fold by 24 hours of IFN-β treatment in A549 cells. (B) Schematic of CRISPR screen method. (C) CRISPR screen results from two independent replicates. Gene-level MAGeCK enrichment scores were correlated between screen replicates. The dotted lines indicate the most-enriched NTC sgRNA in each screen.

Because ZIKV infection causes apoptosis in A549 cells [48], we leveraged ZIKV-induced apoptotic cell death to detect cells that were permissive to ZIKV infection in the

presence of IFN-β. We hypothesized that if a ZIKV antiviral gene was inactivated, it would be enriched in a population of IFN-treated, ZIKV-infected dead cells, which would be no longer adherent. To define the optimal window to capture dying cells in the supernatant, untreated and IFN-β-treated A549 KO library cells (A549 cells transduced with the PIKA library) were infected with ZIKV and stained with Annexin-V every 6 hours from 24 to 48 hours post infection (hpi) (Fig S1). This approach was based on prior studies using Annexin-V, which labels apoptotic cells by binding to phosphatidylserine present on the outer leaflet of dead and dying cell membranes, to detect A549 and neural progenitor cells succumbing to DENV- and ZIKV-induced cell death, respectively [97,98]. We observed that ZIKV infection caused an appreciable increase in Annexin-V+ cells in the supernatant after ~30 hours of infection in the absence of IFN treatment. By contrast, cells pre-treated with IFN-β showed similar levels of dead cells as in non-infected cells, consistent with the protective effect of IFN-β on ZIKV replication (Fig S1). From these data, we chose 42 hpi to conduct the CRISPR screen, as the level of ZIKV-induced cell death was clearly above (by ~3-fold) the IFN-treated and uninfected conditions. This would allow us to maximize our chances of detecting cells that preferentially succumbed to ZIKV-induced cell death due to inactivation of antiviral genes important in restricting ZIKV.

2.4 CRISPR knockout screen reveals candidate genes that contribute to IFN-β restriction of ZIKV

With conditions optimized for a ZIKV-induced cell death-based screen readout, we performed CRISPR screens with the A549 KO library cells. Cells were pre-treated with IFN-β for 24 hours prior to ZIKV infection, followed by replenishment with IFN-β-containing media. In parallel, we included a control condition of A549 KO library cells that were IFN-treated but uninfected to account for cell death from factors other than ZIKV infection. The supernatant

from each condition was collected at 42 hpi and pelleted to collect dead/dying cells. Enriched CRISPR-targeted genes in ZIKV-infected, IFN-treated dead/dying cells were detected by sequencing the sgRNAs and comparing to the sgRNAs detected in the uninfected, IFN-treated dead/dying cells (Fig 2-1B). sgRNAs targeting members of the type I IFN signaling pathway (IFNAR1, STAT1/2, IRF9), included in the PIKA library as positive controls, were the most enriched in two independent replicates of the screen (Fig 2-1C), confirming that this method can successfully identify genes that are critical for IFN-β restriction of ZIKV. Apart from these controls, 11 other genes fell above the most-enriched non-targeting control sgRNA (NTC) (Fig 2-1C, Table S1). Of these 11 hits, only one is a previously identified ZIKV restriction factor (IFI6) [53,54].

2.5 IFI6 and AMOTL2 validate as antiviral genes in out-of-screen experiments

A subset of six of the 11 hits identified in the screen were selected for validation experiments in single-gene KO cells based on both their enrichment score and published data that might support their antiviral potential. IFI6 was tested to determine if its inactivation caused reduced IFN inhibition of ZIKV in our system, as suggested by previous studies that were largely based on overexpression experiments [53,54]. The top hit, AMOTL2, was chosen for follow-up based on its large effect size in the screen (Fig 2-1C). In addition to AMOTL2 and IFI6, the genes HELZ2, GBP3, and BTN3A1 were also selected for follow-up studies based on their known roles in restricting other viruses [99-102]. Finally, the gene RMI2 was included because a prior study showed that IFN-α treatment causes an increase in RMI2 mRNA expression, suggesting RMI2 may play a role in the type I IFN response [103]. Inactivation of IRF9 was performed as a positive control, and a non-targeting control (NTC) sgRNA was included as a negative control. For each of these six genes and two controls, single-gene KO cell pools were

generated by lentiviral transduction. Knockout was assessed by Synthego's Inference of CRISPR Edits (ICE) tool, which predicted at least 60% functional KO of each gene (see Figure 2 legend). The single-gene KO cells were pre-treated with or without IFN-β prior to infection with ZIKV, which was quantified as ZIKV RNA in the cell supernatant [104]. Viral RNA measurements quantified using this assay correlate closely with levels of infectious virus as measured by $TCID_{50}$ (Fig S2).

IFN-treated IRF9 KO cells showed levels of ZIKV replication identical to those in cells that were not treated with IFN, demonstrating complete ablation of IFN restriction of ZIKV (Fig 2-2A). Both AMOTL2 KO and IFI6 KO increased ZIKV infection in the presence of IFN-β, and these differences were evident throughout a 72-hour time course of infection (Fig 2-2B,C). At 72 hours post infection in IFN-β treated cells, ZIKV infection was increased by 49-fold in IRF9 KO cells, 6.7-fold in IFI6 KO cells, and 3.9-fold in AMOTL2 KO cells relative to IFN-treated NTC cells (Fig 2-2H). None of these KOs showed any considerable difference in ZIKV replication from NTC cells in the absence of IFN (Fig 2-2A-C, H), demonstrating an IFN-dependent effect. We observed a small effect with HELZ2 KO (Fig 2-2D, H), which increased ZIKV infection by 2.2-fold in IFN-treated cells while having no effect in the absence of IFN-β (Fig 2-2H). GBP3, RMI2, and BTN3A1 KO had less than two-fold average effects on ZIKV RNA levels with or without IFN-β treatment (Fig 2-2E-H).

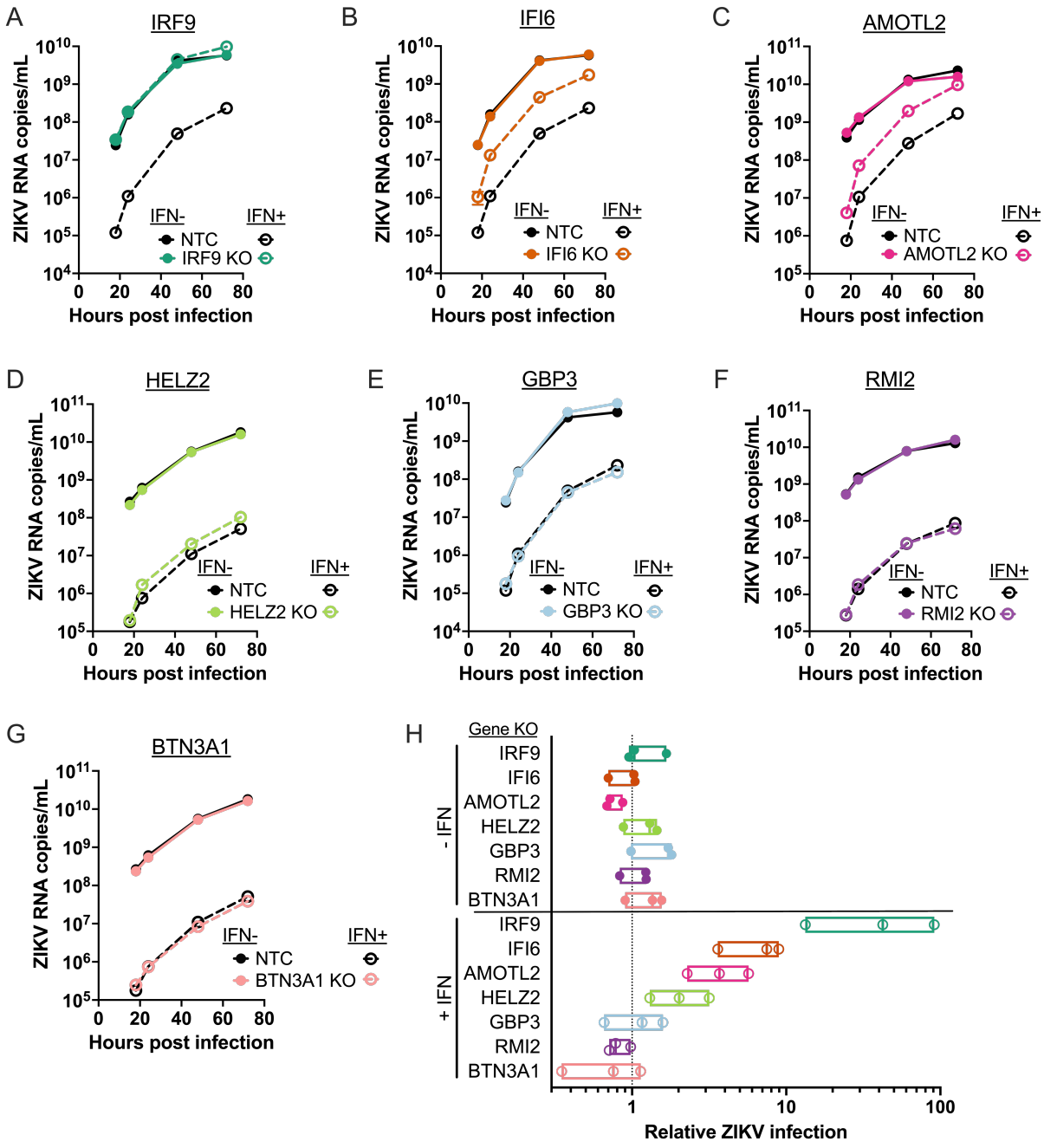

Figure 2-2. Validation experiments on selected CRISPR screen hits confirm role of AMOTL2 and IFI6 in type I IFN restriction of ZIKV.

(A-G) Each indicated gene KO was tested in ZIKV infection experiments in at least three independent replicates; a representative experiment is shown for each KO condition. Cells were pre-treated with or without 1000U/mL IFN-β for 24 hours, inoculated with ZIKV PRVABC59 MOI=0.5 for four hours, and replenished with appropriate media (IFN- closed symbols; or IFN+ open symbols). Cell supernatant was collected at multiple time points throughout a 72-hour infection and tested by RT-qPCR for ZIKV RNA levels. (H) Aggregate data for three independent experiments using a single cell pool with each indicated gene KO. ZIKV infection is portrayed relative to the values in IFN-treated or IFN-untreated NTC cells. For IFI6, two different cell pools were used. The KO scores (Synthego ICE analysis) are as follows: IRF9, 77; IFI6, 61 and 95; AMOTL2, 100; HELZ2, 96; GBP3, 88; RMI2, 100; BTN3A1, 100.

2.6 AMOTL2 contributes to IFN-β restriction of ZIKV

Because AMOTL2 had a statistically significant effect on ZIKV replication and it has not previously been described to have an antiviral function, we focused on further characterization of this gene. To increase our confidence that the phenotype observed was due to AMOTL2 KO, we generated two additional AMOTL2 KO cell pools by lentiviral transduction using distinct sgRNA sequences and validated the KOs by ICE and/or Western blot, which revealed >90% KO by genotyping and/or visible protein depletion by Western (Fig 2-3A). We also generated AMOTL2 KO cells by nucleofection of Cas9 and the original AMOTL2 sgRNAs in four independent nucleofections and likewise observed visible protein depletion in these cells by Western.

Across 14 independent biological replicates that included these various KO cell pools, IFN-treated AMOTL2 KO cells had an average of 5.7-fold higher ZIKV RNA levels at 72 hours post infection compared to IFN-treated NTC cells ($p=0.0027$, Fig 2-3B). Of note, AMOTL2 KO did not significantly change ZIKV RNA levels in the absence of IFN-β treatment (Fig 2-3B). In the presence of IFN-β, the effect of AMOTL2 KO did not reach the effect of IRF9 KO (127-fold increase in ZIKV RNA, $p<0.0001$, Fig 2-3B), suggesting that AMOTL2 contributes a portion of the IFN restrictive effect against ZIKV.

To further validate this phenotype, we complemented the AMOTL2 KO cells with a stably expressed AMOTL2 expression vector and confirmed reconstitution of AMOTL2 protein by Western blot (Fig 2-3C). Relative to complementation with a control open reading frame (ORF stuffer), AMOTL2 complementation resulted in an 8.1-fold decrease in ZIKV infection in the presence of IFN-β at 72 hours post infection ($p=0.0286$, Fig 2-3D-E). We also observed a very small (1.3-fold average) increase in infection in the AMOTL2 KO cells in the absence of IFN-β

(p=0.0286, Fig 2-3D-E). Overall, the direction of the phenotype observed with AMOTL2 complementation in IFN-treated cells is opposite what we observed with AMOTL2 KO, further supporting a role for AMOTL2 as a protein that contributes to the type I IFN response against ZIKV.

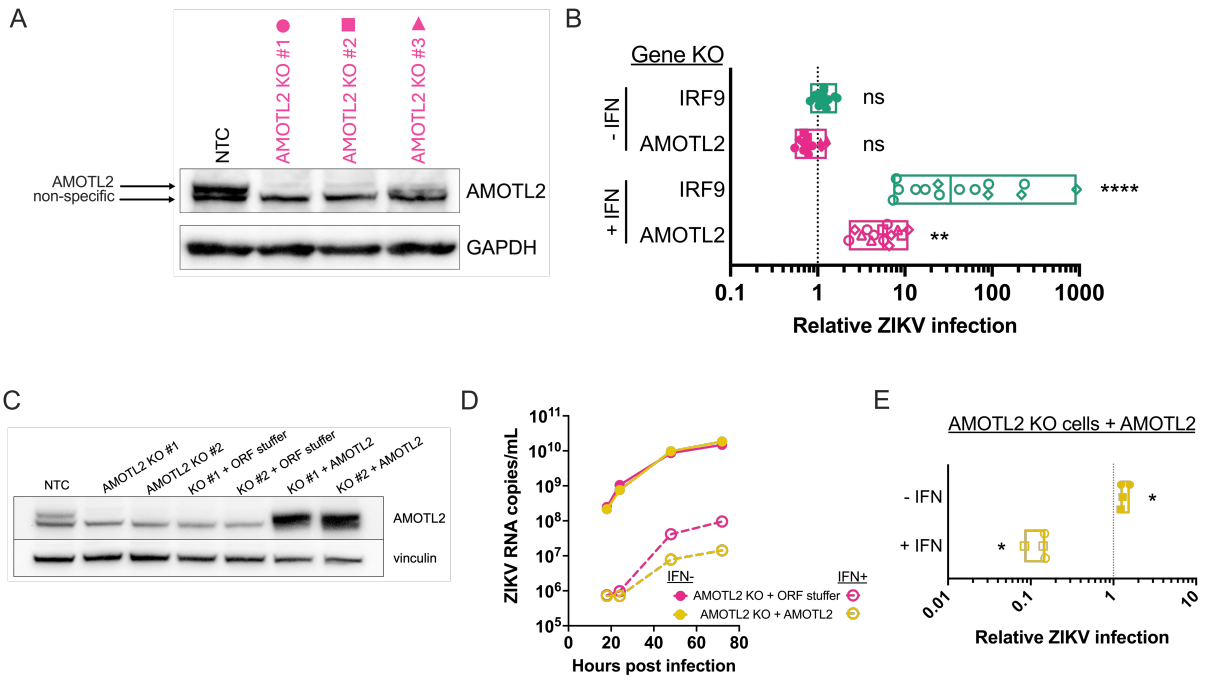

Figure 2-3. AMOTL2 contributes to IFN-β restriction of ZIKV.
(A) Western blot showing AMOTL2 protein reduction in three AMOTL2 KO cell pools generated using different sgRNAs. The antibody cross-reacts with an unknown protein that is slightly smaller than AMOTL2. (B) Aggregate data from ZIKV infection experiments with and without IFN-β treatment in IRF9 KO and AMOTL2 KO cells. Data are normalized to NTC cells. Data include the three AMOTL2 KO cell pools generated by lentiviral transduction shown in (A), as well as four replicates of AMOTL2 KO cells generated by Cas9 RNP nucleofection (indicated by diamonds), for a total of 14 biological replicates. p values are Kruskal-Wallis test with Dunn's multiple comparisons comparing NTC cells to each KO, using separate tests for the IFN- and IFN+ conditions. (C) Western blot showing AMOTL2 protein levels in two AMOTL2 KO cell pools generated using different sgRNAs, as well as in the cells complemented with ORF stuffer or AMOTL2 expression vectors. (D) Representative ZIKV infection experiment with and without IFN-β treatment in AMOTL2 KO cells complemented with AMOTL2. Cells were pre-treated with or without 1000U/mL IFN-β for 24 hours, inoculated with ZIKV PRVABC59 MOI=0.5 for four hours, and replenished with appropriate media (IFN- or IFN+). Cell supernatant was collected at multiple time points throughout a 72-hour infection and tested by RT-qPCR for ZIKV RNA levels. (E) Aggregate data from four replicates of the experiment shown in (D). Data are normalized to AMOTL2 KO cells complemented with ORF stuffer. p values are Mann-Whitney tests comparing the AMOTL2-complemented cells to the ORF-complemented cells for each condition (IFN- or IFN+). P values in this figure are denoted by asterisks as follows: ns = >0.05, * = <0.05, ** = <0.01, and **** = <0.0001.

2.7 AMOTL2 is a constitutively expressed gene that is associated with enhanced ISG expression following IFN-β treatment

The single-gene KO studies with AMOTL2 demonstrate that it has an IFN-dependent effect on ZIKV replication, which would be consistent with AMOTL2 being an ISG. To determine if IFN-β regulates AMOTL2 expression in wildtype A549 cells, we examined our RNA-Seq data from cells treated with IFN-β for 24 hours, which is the length of time of IFN-β pre-treatment in our infection assays. To our surprise, AMOTL2 transcript levels were unaffected by IFN-β treatment in A549 cells, suggesting that AMOTL2 was not upregulated at the time of infection in this cell type (Fig 2-4A). By contrast, upregulation was readily detected for the known ISGs IFIT1, OAS1, and IFITM3, with 25-400 fold upregulation observed for these genes in the same experiment (Fig 2-4A).

Since transcription of many ISGs can occur within hours, we also examined AMOTL2 transcript levels at an earlier timepoint. For this purpose, we performed RT-qPCR on A549 cells treated with IFN-β for nine hours and again saw minimal change in AMOTL2 mRNA levels (average 1.3x increase), despite high upregulation of the ISGs IFIT1, GBP1, and Mx1 (700-fold, 76-fold, and 1150-fold, respectively), which are known to peak quickly following IFN-β stimulation based on ISG kinetics experiments (data not shown) (Fig 2-4B). Additionally, AMOTL2 protein is detected in non-IFN-treated A549 cells, and the protein levels do not change with IFN-β treatment (Fig 2-4C). These findings presented a paradox: that AMOTL2's antiviral phenotype is IFN-dependent, but its expression is not regulated by IFN. We therefore hypothesized that AMOTL2 may be interacting with the IFN signaling pathway itself rather than serving as a downstream effector.

To examine whether AMOTL2 impacts the type I IFN signaling pathway, we measured transcript levels of two ISGs highly upregulated by IFN (IFIT1 and Mx1) in AMOTL2 KO cells and NTC cells with and without IFN-β treatment. In AMOTL2 KO cells, both genes were 2-fold less upregulated by IFN-β treatment compared their level of upregulation in NTC cells ($p=0.0286$ for IFIT1, $p=0.0286$ for Mx1, Fig 2-4D). To determine whether other ISGs are also affected, we used NanoString nCounter technology to assess the expression of a panel of innate immune genes. Of the ten genes on the panel with >10-fold upregulation by IFN-β in unedited NTC cells, all were less upregulated by IFN-β in AMOTL2 KO cells compared to NTC cells (ranging from 1.3- to 2-fold reduction) (Fig S3). This result supports the hypothesis that AMOTL2 KO broadly suppresses upregulation of many ISGs, suggesting that AMOTL2 may interact with the type I IFN signaling pathway upstream of ISG transcription to increase ISG expression.

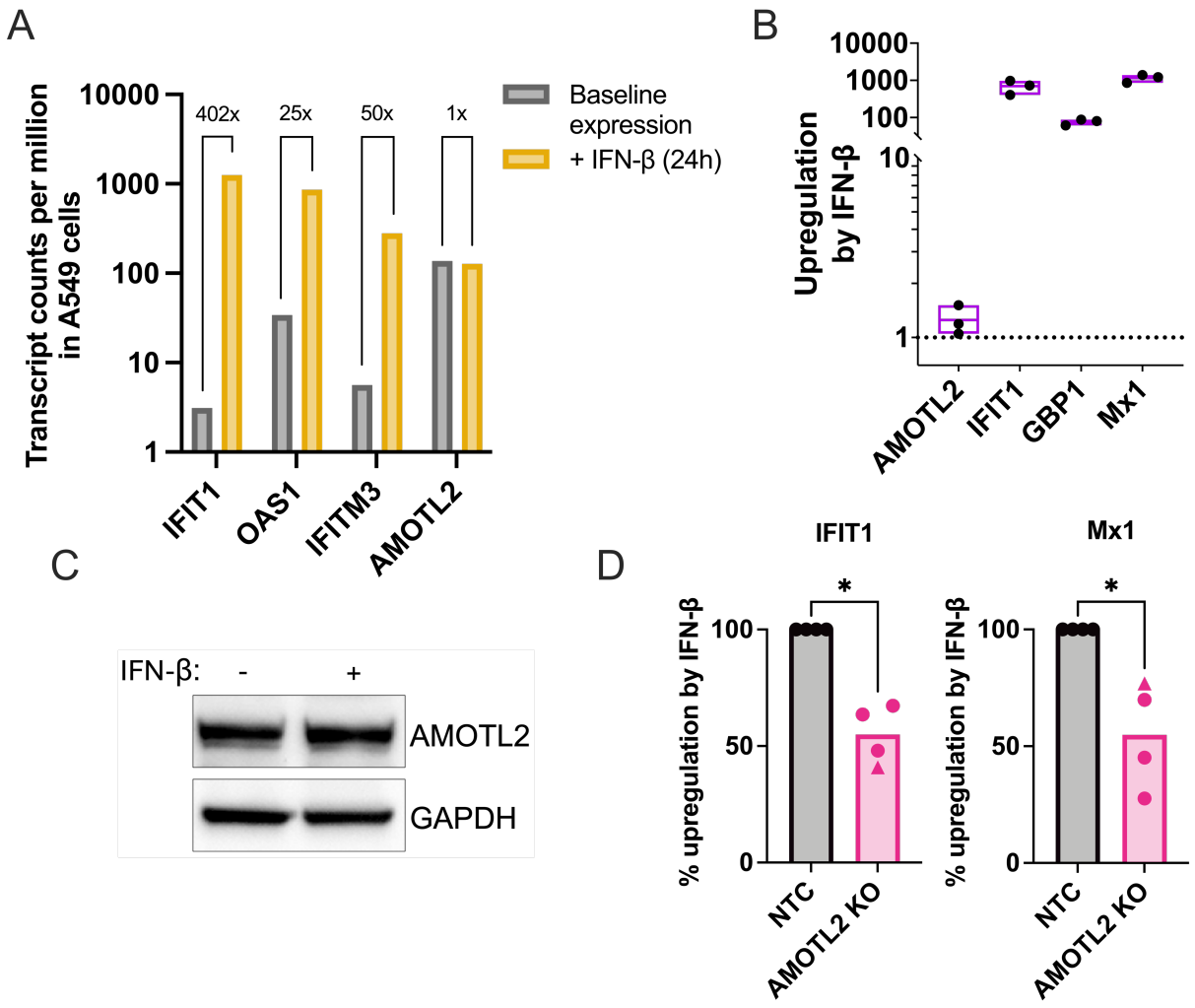

Figure 2-4. AMOTL2 is a constitutively expressed gene that enhances ISG expression following IFN-β treatment.
(A) Bulk RNA-Seq data showing transcripts per million of each indicated gene in wildtype A549 cells with no IFN treatment or with 24 hours of 1000U/mL IFN-β treatment. (B) Upregulation of AMOTL2, IFIT1, GBP1, and Mx1 in NTC cells treated with 1000U/mL IFN-β for 9 hours. Data points represent three independent experiments. (C) Western blot showing AMOTL2 protein levels in NTC cells with and without 24 hours of 1000U/mL IFN-β treatment. (D) RT-qPCR data assessing IFIT1 and Mx1 upregulation in response to 3 hours of 1000U/mL IFN-β treatment of AMOTL2 KO cells, relative to NTC cells. Data points represent independent experiments, and shapes represent different AMOTL2 KO cell pools. p values are Mann-Whitney test. P values are Mann-Whitney test (* = <0.05).

2.8 AMOTL2 reduces phosphorylation and nuclear translocation of STAT1 and increases total STAT1 levels

We next tested whether AMOTL2 affects specific steps of the IFN signaling pathway, beginning with the phosphorylation of STAT1 and STAT2. We detected a lower (1.7-fold)

fraction of phosphorylated STAT1 relative to total STAT1 in IFN-stimulated AMOTL2 KO cells compared to NTC cells ($p=0.0238$, Fig 2-5A, representative blot in Fig S4A). In contrast, there was no difference in phosphorylated STAT2 levels (Fig S4A,B). We also measured nuclear translocation of phosphorylated STAT1 and STAT2, which is a key step leading to ISG transcription (Fig 2-5B). Across three independent AMOTL2 KO cell pools, we observed higher protein levels of phosphorylated STAT1 in the cytoplasmic fraction of AMOTL2 KO cells compared to NTC cells, and a corresponding decrease in the nuclear fraction (Fig 2-5B,C). Specifically, there was a statistically significant reduction in the nuclear to cytoplasmic ratio (1.8-fold average) of phosphorylated STAT1 in AMOTL2 KO cells following IFN treatment compared to NTC cells ($p=0.0091$, Fig 2-5C). Interestingly, there was no difference in the nuclear translocation of phosphorylated STAT2 (Fig S4C,D). This suggests that AMOTL2 is associated with increased phosphorylation and nuclear translocation of STAT1.

Using a STAT1 antibody that preferentially detects unphosphorylated STAT1 (Cell Signaling 9176), we detected 2.4-fold higher levels of total, unphosphorylated STAT1 in the cytoplasm of AMOTL2 KO cells compared to NTC cells, and this difference existed even before IFN treatment ($p=0.0091$ both with and without IFN treatment, Fig 2-5B,D). There was no difference in total STAT2 levels (Fig S4C,E). To determine whether STAT1 transcripts were also elevated in the AMOTL2 KO cells, we measured STAT1 mRNA in AMOTL2 KO cells at baseline (without IFN treatment) and found 1.6-fold higher levels compared to NTC cells ($p<0.0001$, Fig 2-5E). Finally, we performed co-immunoprecipitations of endogenously expressed STAT1 and STAT2 and found that AMOTL2 KO had no effect on STAT1/2 binding compared to NTC cells (Fig 2-5F,G). Taken together, these results suggest that AMOTL2 is associated with reduced

basal STAT1 expression and increased activation of STAT1 upon IFN-β treatment, and thus may affect IFN signaling by modulating STAT1 levels and activation.

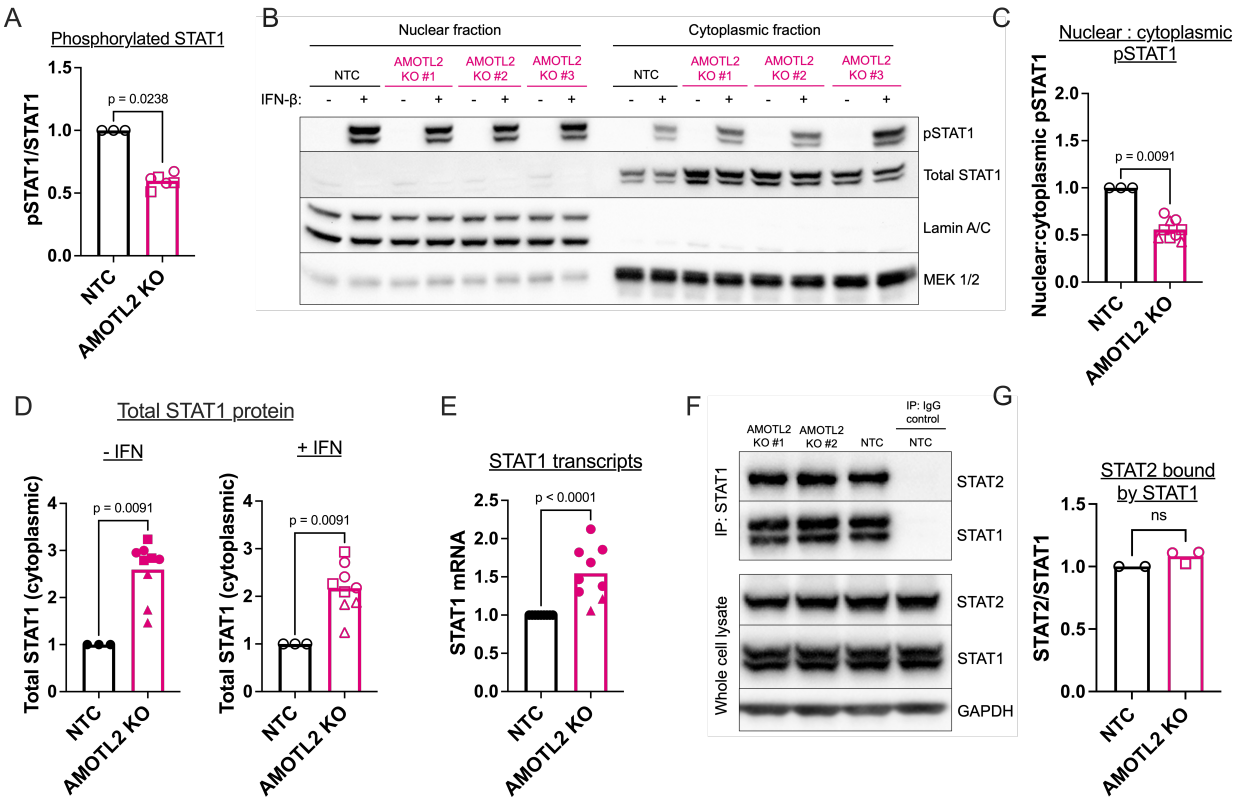

Figure 2-5. AMOTL2 reduces phosphorylation and nuclear translocation of STAT1 and increases total STAT1 levels.
(A) Phosphorylated STAT1 levels relative to total STAT1 levels in whole cell lysates from AMOTL2 KO cell pools stimulated with 1000U/mL IFN-β for 1 hour, normalized to NTC cells. Data show three independent replicates. (B) Representative western blot on nuclear and cytoplasmic fractions of non-IFN treated and IFN-treated NTC and AMOTL2 KO A549 cells. IFN-β treatment was 1000U/mL for 2 hours. The total STAT1 antibody is known to favor detection of unphosphorylated STAT1. Lamin A/C is a nuclear marker, and MEK 1/2 is a cytoplasmic marker. (C) Nuclear to cytoplasmic ratio of phosphorylated STAT1 for three replicates of the experiment shown in (B), normalized to NTC cells. (D) Total cytoplasmic STAT1 levels for three replicates of the experiment shown in (B), normalized to NTC cells. (E) STAT1 transcripts in non-IFN-stimulated AMOTL2 KO cells, normalized to NTC cells. (F) Co-immunoprecipitation assay using A549 cells treated with IFN-β for 24hr and analyzed by Western blot. An anti-STAT1 antibody was used as the capture antibody. (G) Quantification of STAT2 pulldown relative to STAT1 pulldown in the STAT1 IP represented in (F) from two independent experiments. In all panels, shapes represent different AMOTL2 KO cell pools, open circles represent IFN-treated conditions, closed circles represent non-IFN-treated conditions, and p values are Mann-Whitney test.

2.9 Identification of AMOTL2 interacting proteins and critical domain for AMOTL2 antiviral function

AMOTL2 is known to interact with YAP and TAZ [105,106], transcriptional regulators that have been reported to antagonize the innate immune response through several mechanisms [107]. We hypothesized that AMOTL2 may exert its antiviral function via this interaction with YAP and/or TAZ. To test whether YAP and TAZ affect the type I IFN response against ZIKV in A549 cells, we generated YAP and TAZ KO A549 cells and validated the KOs by Synthego ICE and/or Western blot, which revealed 100% KO by genotyping and/or visible protein depletion by Western. Neither gene KO affected ZIKV levels in the absence of IFN treatment (Fig 2-6A-C). In the presence of IFN, YAP KO had little if any appreciable effect (1.5-fold reduction in ZIKV infection at 72 hours post infection, not significant, Fig 2-6A,C), while TAZ KO caused an 8.7-fold average reduction in infection at 72 hours ($p=0.0011$, Fig 2-6B,C), in addition to causing a consistent reduction at all earlier time points tested (Fig 2-6B). We therefore focused our attention on TAZ. Because the interaction between AMOTL2 and TAZ has not been examined in A549 cells, we tested the interaction between the endogenous proteins by co-IP in cells either untreated or treated with IFN-β for various lengths of time. Similar levels of AMOTL2 were detected in immunoprecipitants using TAZ as bait under all conditions in wildtype cells, suggesting that binding between TAZ and AMOTL2 is independent of IFN-β treatment (Fig 2-6D).

To test whether TAZ affects AMOTL2's antiviral activity, we next generated AMOTL2/TAZ double KO cells (Fig 2-6E). In ZIKV infection experiments at 72 hours post infection, we observed that ZIKV replication in the double KO cells was significantly lower than in the AMOTL2 single KO cells ($p=0.0448$, Fig 2-6F), but not compared to the TAZ single KO cells

(Fig 2-6F). This suggests that TAZ's effect is dominant, which could mean that AMOTL2's antiviral activity is dependent on an interaction with TAZ, or that TAZ's proviral function occurs upstream of AMOTL2's antiviral function.

To further explore these possibilities, we expressed a version of AMOTL2 with a mutation in the TAZ-binding domain (the PPQY motif at amino acids 210-213 [105]) in AMOTL2 KO cells. In parallel, we complemented AMOTL2 KO cells with wildtype AMOTL2, AMOTL2 Y213A (ablating the PPQY motif), a control ORF stuffer coding sequence, and several other AMOTL2 mutation or deletion constructs that disrupt various functional domains of the protein. These included AMOTL2 P105A (disrupting AMOTL2's LPTY motif, responsible for AMOTL2's interaction with MAGI-1 and actin [108,109]), AMOTL2 CCdel (deleting AMOTL2's coiled-coil domain, responsible for AMOTL2 homo-oligomerization and localization to tight junctions [108]), AMOTL2 PDZdel (deleting AMOTL2's C-terminal PDZ-binding motif, involved in cell polarity and migration [110]), and S159A (removing a serine that is phosphorylated by LATS1/2 [94]). AMOTL2 protein levels were markedly increased relative to ORF stuffer-complemented AMOTL2 KO cells in all the expression constructs save CCdel (a deletion of the amino acids 314-572) (Fig 2-6G). This is likely because the AMOTL2 antibody was raised against an AMOTL2 recombinant protein spanning a region encompassed within the coiled-coil domain (amino acids 401-480). Additionally, STAT1 levels were visibly reduced relative to ORF stuffer-complemented AMOTL2 KO cells in all the expression constructs save CCdel (Fig 2-6G). To confirm that the AMOTL2 CCdel cells were expressing AMOTL2, we measured AMOTL2 mRNA levels by qPCR using primers that bind all of our AMOTL2 constructs, including CCdel. All of the AMOTL2-complemented cells including CCdel exhibited ~10-fold higher AMOTL2

mRNA levels relative to the ORF stuffer-complemented cells, suggesting that the CCdel-complemented cells were indeed expressing AMOTL2 (Fig 2-6H).

Compared to ORF stuffer, complementation with either wildtype AMOTL2 or AMOTL2 Y213A reduced ZIKV replication measured at 48h and 72h in the presence, but not absence, of IFN-β (Fig 2-6I,K). This difference was statistically significant in the IFN-treated cells at 72h in four independent replicates (Fig 2-6K, 8.1-fold reduction in ZIKV infection for wildtype, 9.1-fold reduction in ZIKV infection for Y213A, $p<0.05$ for both), with no significant difference observed in the absence of IFN. By contrast, complementation with AMOTL2 CCdel had no effect on ZIKV infection compared to ORF stuffer (Fig 2-6J,K; no significant difference across four replicates in the presence or absence of IFN). Complementation with the other AMOTL2 expression variants (AMOTL2 P105A, AMOTL2 PDZdel, and AMOTL2 S159A) decreased ZIKV infection relative to ORF stuffer by 6.8- to 9.4-fold in IFN-treated cells, similar to the effect of complementation with wildtype AMOTL2 (Fig 2-6K, $p<0.05$ for all). Complementation with AMOTL2 Y213A, AMOTL2 P105A, and AMOTL2 PDZdel caused small but statistically significant increases in ZIKV infection in the absence of IFN (1.4-fold, $p<0.05$ for all, Fig 2-6K). These results suggest that TAZ binding is not required for AMOTL2's antiviral activity, and that AMOTL2's antiviral mechanism does not likely act via TAZ. Furthermore, AMOTL2's coiled-coil domain is essential for its antiviral phenotype.

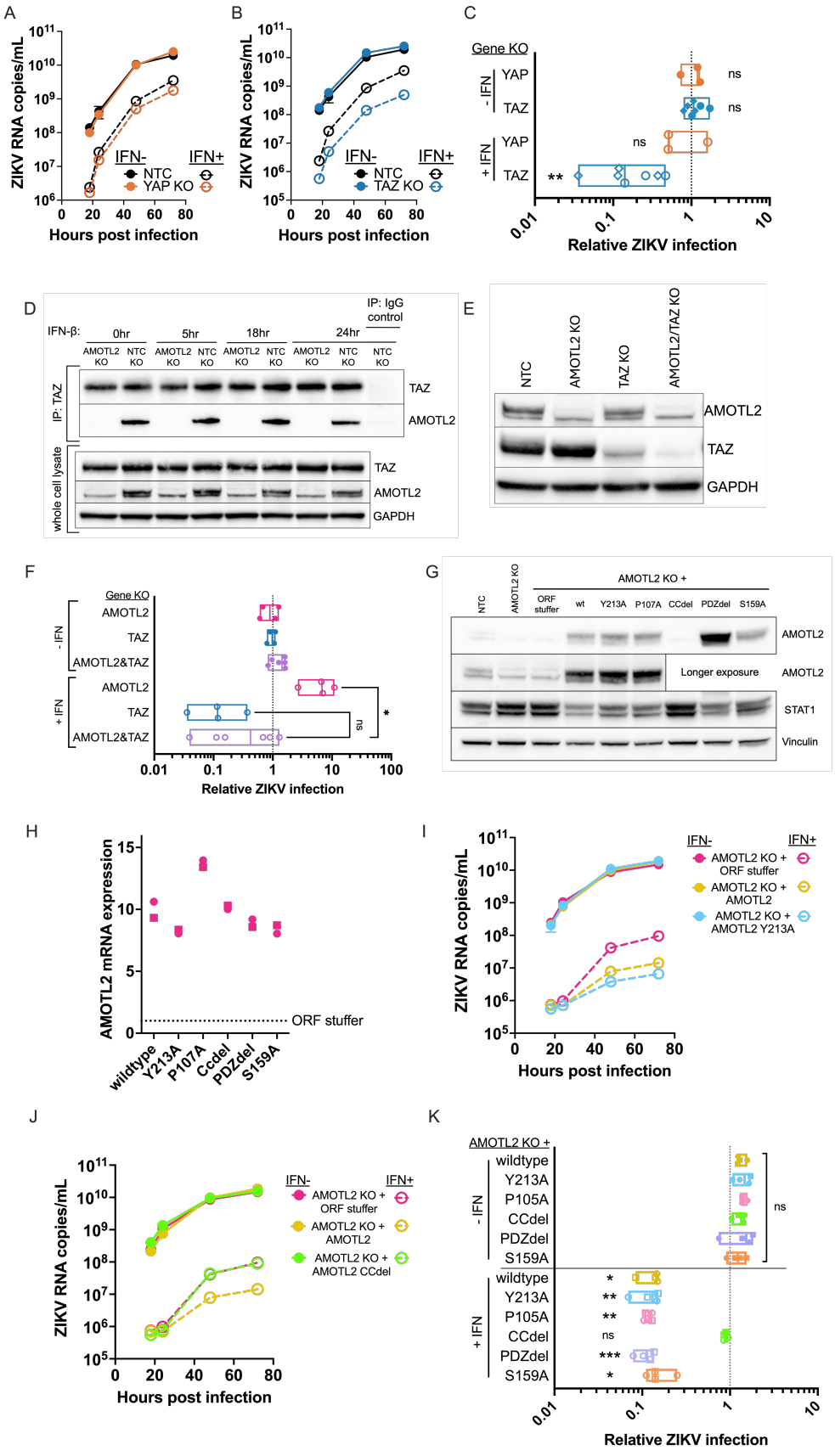
A
ZIKV RNA copies/mL
Hours post infection
IFN-
IFN+
NTC
YAP KO
B
TAZ KO
C
Gene KO
- IFN
+ IFN
YAP
TAZ
ns
**
Relative ZIKV infection
D
IP: IgG control
IFN-β:
0hr
5hr
18hr
24hr
AMOTL2 KO
NTC KO
IP: TAZ
whole cell lysate
TAZ
AMOTL2
GAPDH
E
NTC
AMOTL2 KO
TAZ KO
AMOTL2/TAZ KO
F
AMOTL2
AMOTL2&TAZ
su
*
G
AMOTL2 KO +
ORF stuffer
wt
Y213A
P107A
CCdel
PDZdel
S159A
Longer exposure
STAT1
Vinculin
H
AMOTL2 mRNA expression
ORF stuffer
wildtype
Y213A
P107A
CCdel
PDZdel
S159A
I
AMOTL2 KO + ORF stuffer
AMOTL2 KO + AMOTL2
AMOTL2 KO + AMOTL2 Y213A
J
AMOTL2 KO + AMOTL2 CCdel
K
P105A
***

Figure 2-6. Identification of AMOTL2 interacting proteins and critical domain for AMOTL2 antiviral function.
(A,B) Representative ZIKV infection experiments with and without IFN-β treatment in YAP KO (A) and TAZ KO (B) cells. Cells were pre-treated with or without 1000U/mL IFN-β for 24 hours, inoculated with ZIKV PRVABC59 MOI=0.5 for four hours, and replenished with appropriate media (IFN- or IFN+). Cell supernatant was collected at multiple time points throughout a 72-hour infection and tested by RT-qPCR for ZIKV RNA levels. (C) Aggregate data from at least three independent replicates of the experiments shown in (A) and (B). YAP KO cells were generated by lentiviral transduction, and TAZ KO cells were generated by lentiviral transduction or Cas9 RNP nucleofection (the latter indicated by diamonds). Data are normalized to NTC KO cells. p values are Kruskal-Wallis with Dunn's multiple comparisons comparing NTC KO cells to each KO, with separate tests performed for IFN- and IFN+ conditions. (D) Co-immunoprecipitation assay using A549 cells treated with IFN-β for 0hr, 5hr, 18hr, or 24hr and analyzed by Western blot. An anti-TAZ antibody was used as the capture antibody. (E) Western blot showing AMOTL2 and TAZ protein levels in single-gene and double-gene KO cells used in infection assays in (F). GAPDH is a loading control. (F) Aggregate data from at least four independent replicates of IFN treatment and ZIKV infection assays using the AMOTL2 and TAZ double KO cells. Data are normalized to NTC cells. p values are Kruskal-Wallis with Dunn's multiple comparisons comparing IFN-treated AMOTL2/TAZ double KO cells to IFN-treated AMOTL2 single KO and IFN-treated TAZ single KO cells. (G) Western blot showing AMOTL2 and STAT1 levels in AMOTL2 KO cells complemented with AMOTL2 expression vectors or ORF stuffer. (H) AMOTL2 transcripts in the panel of AMOTL2-complemented AMOTL2 KO cells, relative to AMOTL2 transcript expression in ORF stuffer-complemented AMOTL2 KO cells. (I-J) Representative ZIKV infection experiments with and without IFN-β treatment in AMOTL2 KO cells complemented with AMOTL2 Y213A (I) and AMOTL2 CCdel (J), compared to cells complemented with ORF stuffer and cells complemented with wildtype AMOTL2. Cells were pre-treated with or without 1000U/mL IFN-β for 24 hours, inoculated with ZIKV PRVABC59 MOI=0.5 for four hours, and replenished with appropriate media (IFN- or IFN+). Cell supernatant was collected at multiple time points throughout a 72-hour infection and tested by RT-qPCR for ZIKV RNA levels. (K) Aggregate data from four independent replicates of the experiments shown in (I-J) in addition to data from other AMOTL2 expression constructs not shown in (I-J). Shapes represent different AMOTL2 KO cells that were transduced with the expression vectors. Data are normalized to ORF stuffer-complemented AMOTL2 KO cells. p values are Kruskal-Wallis with uncorrected Dunn's test comparing ORF stuffer cells to each AMOTL2-complemented cell pool, with separate tests performed for IFN- and IFN+ conditions. P values in this figure are denoted by asterisks as follows: ns = >0.05, * = <0.05, ** = <0.01, and *** = <0.001.

2.10 Discussion

In this study, we present a CRISPR screening platform that uses virus-induced cell death to identify genes with innate antiviral activity against ZIKV. To our knowledge, this is the first CRISPR knockout screen to be performed to identify genes contributing to IFN restriction of ZIKV. Using the gene inactivation approach allowed us to demonstrate that IFI6 KO reduces type I IFN restriction of ZIKV in A549 cells, reinforcing results of prior overexpression and knockdown studies in other cell types and strengthening the evidence for IFI6 as a major anti-flavivirus ISG. Importantly, we uncovered a new antiviral host gene, AMOTL2, that is not an

ISG in A549 cells, but nonetheless inhibits ZIKV in an IFN-dependent manner. Our studies suggest that AMOTL2 acts by regulating STAT1 levels and downstream ISG expression.

Prior screens for ZIKV restriction factors include a knockdown screen using a limited siRNA library of 386 genes [75] and two other studies that used genome-wide gain-of-function (overexpression) libraries [54,74]. While these latter studies are useful starting points, relying on gene overexpression may limit the biological relevance of the results, as has been demonstrated for IFITM3 [47]. Our screen results validated IFI6, a known ISG and restriction factor that blocks flavivirus replication by preventing the formation of the viral replication organelles in the endoplasmic reticulum membrane [53,54]. In our validation experiments, we also observed a small effect on IFN restriction with HELZ2 KO, though it was modest in comparison to other screen hits. HELZ2 is an ISG that has been reported to restrict two other flaviviruses, dengue virus and hepatitis C virus [99,100]. Future work is required to determine whether HELZ2 is a true ZIKV restriction factor.

In addition to containing sgRNAs targeting most ISGs in A549 cells, our CRISPR library also targeted more than 1,700 other human genes, affording us the opportunity to discover non-ISGs that are important in the type I IFN response. Indeed, the top experimental hit in our screen was *AMOTL2*, a gene that we show is not regulated by IFN-β in A549 cells yet contributes to IFN-mediated restriction of ZIKV. AMOTL2 is a motin family protein of 780 amino acids with known functions in cancer pathogenesis, angiogenesis, tight junction formation, cell polarity, the actin cytoskeleton, and regulation of Hippo signaling [94,111]. We confirmed the role of AMOTL2 in the IFN response against ZIKV in validation experiments using several distinct AMOTL2 KO cell pools targeting different CRISPR cut sites. We also demonstrated that

reconstituting AMOTL2 expression can rescue IFN restriction of ZIKV in the AMOTL2 KO cells. Together, these findings confirm that AMOTL2 has an antiviral effect against ZIKV.

We found that AMOTL2 knockout results in broad blunting of ISG transcription following IFN-β stimulation. In follow-up mechanistic studies, we identified specific defects in the IFN signaling pathway in AMOTL2 knockout cells that could explain this effect, including reduced STAT1 phosphorylation and reduced nuclear translocation of phosphorylated STAT1, with no effect on STAT2. Interestingly, levels of total, unphosphorylated STAT1 (U-STAT1) were increased in AMOTL2 KO cells, even in the absence of IFN-β stimulation. U-STAT1 and unphosphorylated STAT2 plays an important role in maintaining the antiviral state triggered by type I IFN after levels of phosphorylated STAT1 and STAT2 decline. However, paradoxically, resting high levels of U-STAT1 have also been reported to reduce the responsiveness of cells to activation by type I IFN treatment. We therefore propose a model in which AMOTL2 suppresses high levels of basal STAT1, relieving the block on IFN-induced STAT1 phosphorylation and nuclear translocation (Fig 2-7). In an AMOTL2-low environment, levels of basal U-STAT1 are high, and STAT1 activation by type I IFN treatment is low. ISG transcription is reduced, and the cell is less able to control ZIKV replication. In an AMOTL2-high environment, U-STAT1 levels are kept in check, and STAT1 can be potently activated by type I IFN treatment, resulting in high ISG transcription and an antiviral state that better controls ZIKV replication. We also found that STAT1 mRNA levels are higher in AMOTL2 KO cells. While this could suggest that AMOTL2 KO increases STAT1 transcription, it is also possible that AMOTL2 KO increases STAT1 protein levels and this feeds forward to drive more STAT1 transcription, as overexpression of exogenous U-STAT1 has been shown to stimulate transcription of STAT1 in other cell types [85].

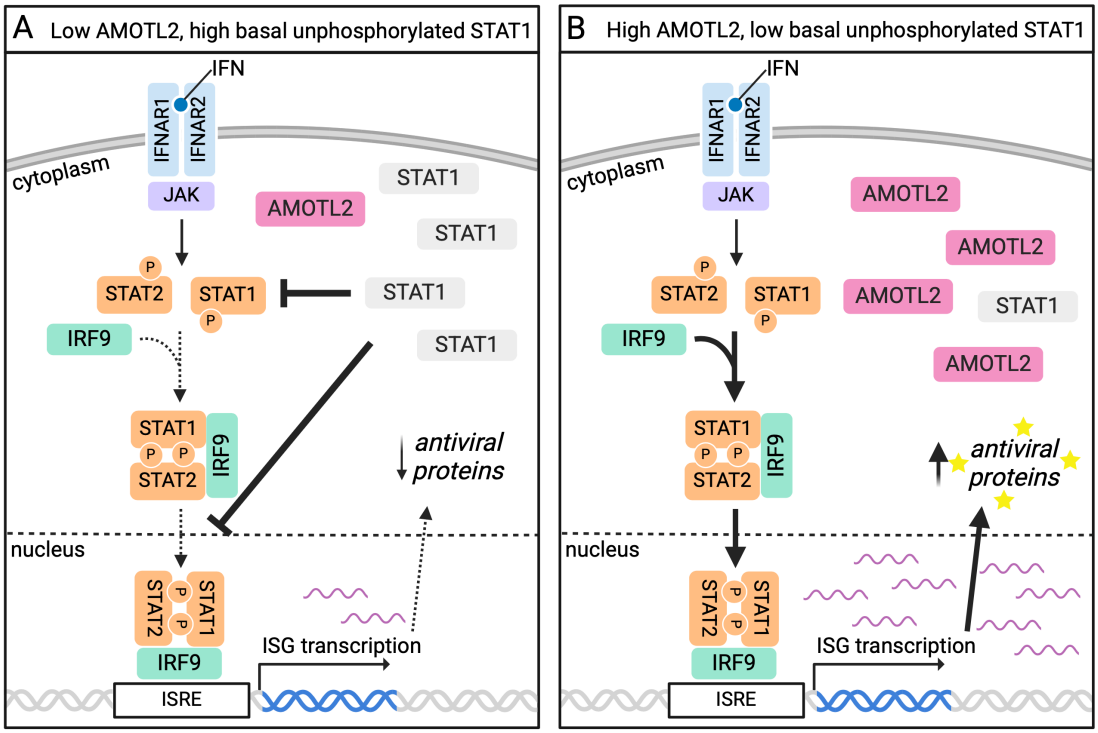

Figure 2-7. Working model of the impact of AMOTL2 on the type I IFN signaling pathway.
(A) AMOTL2 knockout results in high basal levels of unphosphorylated STAT1. Upon IFN treatment, the cell exhibits reduced STAT1 phosphorylation and nuclear translocation and low ISG transcription, resulting in increased ZIKV replication. (B) With high levels of AMOTL2, basal unphosphorylated STAT1 levels are low, and STAT1 activation occurs normally. ISG transcription is high and ZIKV replication is better controlled. Created using BioRender.

AMOTL2 is known to interact with TAZ and YAP, transcriptional co-activators that stimulate the upregulation of genes involved in diverse biological processes, including cell proliferation, stemness, epithelial-to-mesenchymal transition, angiogenesis, and other processes resulting in cancerous transformation of cells [112]. When bound to AMOTL2, TAZ and YAP are sequestered in the cytoplasm and blocked from engaging with transcription in the nucleus [105,106]. Interestingly, TAZ and YAP also antagonize the innate immune response through a variety of mechanisms, including by antagonizing the IFN signaling pathway [107,113]. We reasoned that TAZ and/or YAP may be involved in AMOTL2's phenotype, given their tendency

to interrupt innate immune signaling pathways [107], and AMOTL2's tendency to bind and block TAZ and YAP.

We found that inactivation of TAZ reduced ZIKV infection in IFN-treated cells, suggesting that TAZ is a proviral gene that antagonizes IFN restriction of ZIKV. We also demonstrated that TAZ and AMOTL2 interact in A549 cells in an IFN-independent manner. In AMOTL2/TAZ double KO cells, ZIKV replication resembled TAZ single KO cells more so than AMOTL2 single KO cells, suggesting that TAZ's mechanism occurs upstream of AMOTL2's or that AMOTL2 is dependent on TAZ for its function. To definitively test whether TAZ is involved in AMOTL2's antiviral mechanism, we complemented the AMOTL2 KO cells with a panel of AMOTL2 variants with mutations in various functional domains, including in the TAZ-binding PPQY motif. We showed that AMOTL2's antiviral phenotype was rescued with wildtype AMOTL2 to a similar extent as it was rescued with the PPQY-mutated version, suggesting that TAZ binding is not involved in AMOTL2's mechanism. On the other hand, an AMOTL2 mutant with a deletion of the coiled-coil domain failed to rescue the phenotype, suggesting that this domain mediates AMOTL2's antiviral function. This finding provides a direction for future detailed studies of the mechanism and cellular partners contributing to AMOTL2's ability to modulate STAT1 levels and enhance IFN signaling.

Despite being largely studied in the context of cancer, AMOTL2 mRNA is broadly detectable in primary human tissue, including in the placenta, a site highly relevant to ZIKV infection *in vivo* (The Human Protein Atlas). Interestingly, several transcriptomic datasets indicate that AMOTL2 is upregulated by type I IFN treatment in primary monocyte-derived macrophages (MDMs) (interferome.org v2.01), but it is not reported as an ISG in any other cell type, suggesting that AMOTL2 might play a role in IFN restriction in MDMs. Thus, future

studies might study the effect of AMOTL2 inactivation in primary cells such as MDMs, which are also a major target of ZIKV infection *in vivo* [114,115].

In this study, we uncovered a novel role for AMOTL2, a gene with no previously described function in innate immunity, in IFN-mediated restriction of ZIKV. This work adds to our understanding of the cellular players that act in concert to enact the antiviral state mediated by type I IFN. Additionally, based on AMOTL2's impact on IFN-stimulated STAT1 activation, it is likely that AMOTL2 is broadly antiviral and not specific to ZIKV. Future studies are needed to better define the mechanism and breadth of AMOTL2's antiviral function. This work improves our basic understanding of type I IFN-mediated protection against ZIKV and could inform the development of new therapeutics.

2.11 Supplementary Material

Gene	-log10(MAGeCK score), screen replicate 1	-log10(MAGeCK score), screen replicate 2
STAT1	22.11	20.91
STAT2	21.98	21.74
IFNAR1	20.64	21.86
IRF9	11.63	11.63
AMOTL2	4.32	7.37
GBP3	3.32	4.08
IKBKE	3.01	2.56

IFI6	2.66	3.58
DNASE1L2	2.64	3.20
ST3GAL6	2.52	3.61
CBWD6	2.47	2.52
BTN3A1	2.46	2.79
RMI2	2.29	3.40
HELZ2	2.19	3.41
ICAM1	1.91	2.60

Table S1. CRISPR screen hits

Gene Target	Knock out ID	Knockout	sgRNA #1	sgRNA #2	Forward primer to check gene editing	Reverse primer to check gene editing	ICE KO score	KO checked by Western blot
N/A	NTC KO #1	LentiCRISPR v2	ATCTCGGGTC GACTGCGGAT	N/A	N/A	N/A	N/A	N/A
N/A	NTC KO #2	LentiCRISPR v2	ATCTCGGGTC GACTGCGGAT	GCAAACCCGA GTGACACGTC	N/A	N/A	N/A	N/A
IRF9	IRF9 KO #1	LentiCRISPR v2	ACAATTCCACA GGCCAGCCA	N/A	CCTGCATAATC CCTTCTGAGC	AGGAAGCAGAA ACTCCAGGG	77	
IRF9	IRF9 KO #2	LentiCRISPR v2	ACAATTCCACA GGCCAGCCA	GAGGGAGTCC TGGAGCACAG	CCTGCATAATC CCTTCTGAGC	AGGAAGCAGAA ACTCCAGGG	100	
YAP	YAP KO	LentiCRISPR v2	ACATCGATCAG ACAACAACA	GACACTGTAAT TACAGCACA	GACTTTTGGGG TTTTGTGGTGT	AACCACAAAGG GAAGAGGGT	100	
WWTR1	TAZ KO	LentiCRISPR v2	ACATAGAAAAA ATCACCACA	AGGCTTACCGA GATTTGGCT	GGAATACAAGC TCCACGGGC	GGCAGAAGGG TAACACGGT	100	✓
GBP3	GBP3 KO	LentiCRISPR v2	GTCTGCCATTA CACAGCCTG	N/A	CTTTTGGCCAT CCTGGGGAA	ACCTTAAAGAT CAGTGTAAACT TGC	88	
HELZ2	HELZ2 KO	LentiCRISPR v2	CATCTCGCGG CACTTCTACG	CCGCTGGCCT ATGCCTCGCA	TTTAGCGGACA GCACTGAGG	ACACAGATGCA CCTCTGCTC	96	
RMI2	RMI2 KO	LentiCRISPR v2	CGGACCGAGA AGTCCCCGCT	TCCGAGGTCG CCGCCACTCA	GTAGGTGATCT GCATCCACGC	AAACGAAGGTG CCGGAGTC	100	
BTN3A1	BTN3A1 KO	LentiCRISPR v2	CCTTCTTCAGG AGCGCCCAG	AAGCAAGAACA AAGCACAAG	ACTGGGTGGAT GGGAGACTT	ACTTCTCCATC CTGCAACGG	100	
AMOTL2	AMOTL2 KO #1	LentiCRISPR v2	GCGGCGCCAT CGAGGACCAG	GGGCCTTGAG TTCCTGCTCC	CCCAGTTGTTA CTGCAAGGGT	CTAGAGGTCAA GGACAACTAGC	100	✓
AMOTL2	AMOTL2 KO #2	LentiCRISPR v2	CCAGGGCGGT GAGAACCACC	GACCCACGGT ACCGTGCCCG	CCCATCCACCC CTAGACTGA	CCAGGCATTAG CTCTCCCTT	not analyzed	✓
AMOTL2	AMOTL2 KO #3	LentiCRISPR v2	ACTGTCCATCT TGTTCCGCA	N/A	CCTGTTTGGTG GCCTCTCAT	CACACAGGCCA GGCTTCTTA	93	✓
IFI6	IFI6 KO	LentiCRISPR v2	CTGCTGCTCTT CACTTGCAG	N/A	GATGCGGAGA GGAAACGGAT	CCTGCAGGAG AGCAGACAAA	61	
IFI6	IFI6 KO #2	LentiCRISPR v2	CTGACCTTCAT GGCCGTCGG	GAGCAGACAAA GCGTAGTGG	TGTGGGTAAG GATGCAGGT	CCATTCAGGAT CGCAGACCA	95	
N/A	NTC KO #2 (nf)	Cas9 RNP nucleofection	ATCTCGGGTC GACTGCGGAT	GCAAACCCGA GTGACACGTC	N/A	N/A	N/A	N/A
AMOTL2	AMOTL2 KO #1 (nf)	Cas9 RNP nucleofection	GCGGCGCCAT CGAGGACCAG	GGGCCTTGAG TTCCTGCTCC	CCCAGTTGTTA CTGCAAGGGT	CTAGAGGTCAA GGACAACTAGC	100	✓
WWTR1	TAZ KO (nf)	Cas9 RNP nucleofection	ACATAGAAAAA ATCACCACA	AGGCTTACCGA GATTTGGCT	GGAATACAAGC TCCACGGGC	GGCAGAAGGG TAACACGGT	100	✓
AMOTL2/WWTR1	AMOTL2/TAZ KO (nf)	Cas9 RNP nucleofection	AMOTL2 KO #1 and TAZ KO sgRNAs		AMOTL2 KO #1 and TAZ KO primers		100, 100	✓

Table S2. sgRNA sequences and ICE primers

PCR to amplify edited loci from genomic DNA for ICE analysis

Temp	Time	Cycles
95C	4 min	1 cycle
95C	20 s	14 cycles
55C (-0.5C ea cycle)	30 s	
72C	30 s	
95C	20 s	20 cycles
55C	30 s	
72C	30 s	
72C	10 min	1 cycle
Final store		

RT-qPCR protocol to amplify ZIKV RNA from viral supernatants

Temp	Time	Cycle
50C	10 min	1 cycle
95C	3 min	1 cycle
95C	15 s	40 cycles
60C	1 min	
Final store		

Reverse transcription protocol

Temp	Time	
25C	5 min	
50C	60 min	
70C	15 min	
Final store		

qPCR protocol for human gene expression quantification

Temp	Time	Cycle
95C	30 s	1 cycle
95C	15 s	40 cycles
58C	1 min	
Melt curve analysis		

Table S3. Thermocycler programs

Primers and probe to amplify ZIKV RNA from viral supernatants

Name	Sequence
forward Zika-PRV/BRA-F	5′-TGAGGCATCAATATCAGACATG-3′
reverse Zika-Dual Rev	5′-GTTCTTTTGCAGACATATTGAGTG-3′
probe	5′ FAM −TGCCCAACA/ZEN/C-AAGGTGAAGCCTACCT-IAbkBkFQ

Primers for human gene expression by RT-qPCR

Gene	Forward primer	Reverse primer
AMOTL2	CACAGGCATCAGGAGATGGAAAG	GCGCTGCTGAAGGACCTTG
IFIT1	TTGATGACGATGAAATGCCTGA	CAGGTCACCAGACTCCTCAC
Mx1	AAGAGCTCCGTGTTGGAGG	TGGTAACTGACCTTGCCTCTC
GBP1	CGCTCTTAAACTTCAGGAACAGGA	GTACATGCCTTTCGTCGTCTC
STAT1	GTGATCTCCAACGTCAGCCA	TGGCGTTAGGACCAAGAAGC
GAPDH	GAAGGTCGGAGTCAACGGATTT	GAATTTGCCATGGGTGGAAT

Table S4. Other primers

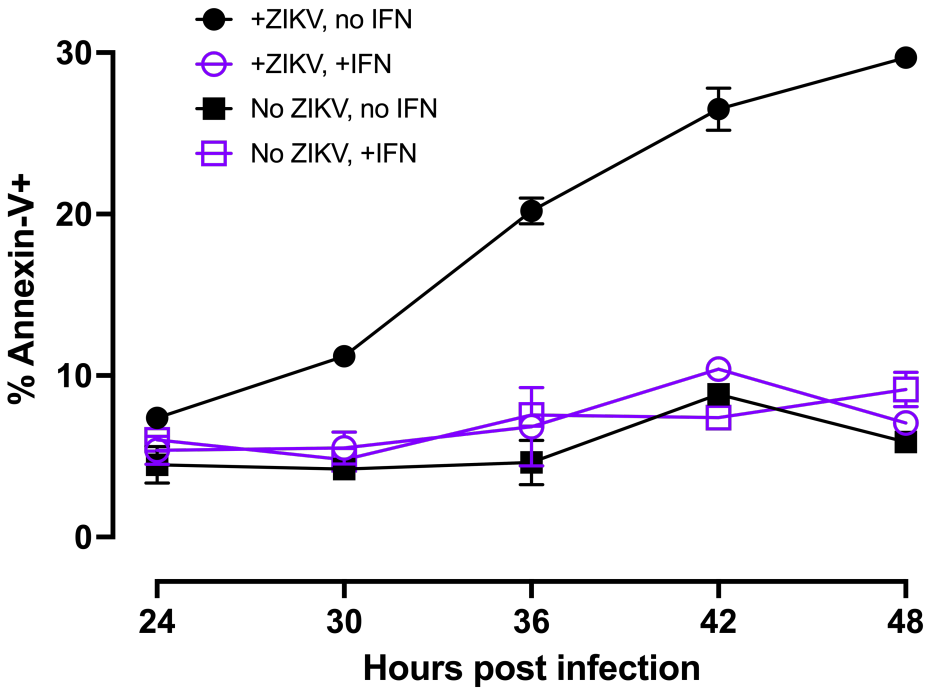

Figure S1. IFN-β treatment suppresses ZIKV-induced cell death.
Annexin-V staining over a time course of ZIKV infection in untreated and IFN-β-treated A549 KO library cells. The percent Annexin-V+ cells in the supernatant of untreated or IFN-β-treated, ZIKV-infected or uninfected A549 KO library cells is shown every 6 hours from 24 to 48 hours post infection. Data points represent mean of technical replicates, and error bars indicate range.

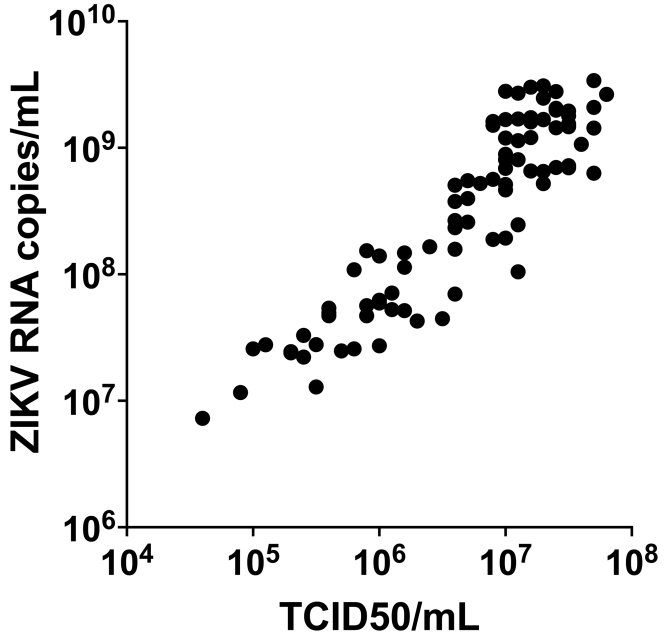

Figure S2. ZIKV infection quantification by TCID50 vs. RTq-PCR.
Each data point represents an individual ZIKV-infected cell supernatant sample quantified by both methods. Spearman r=0.87, p<0.0001

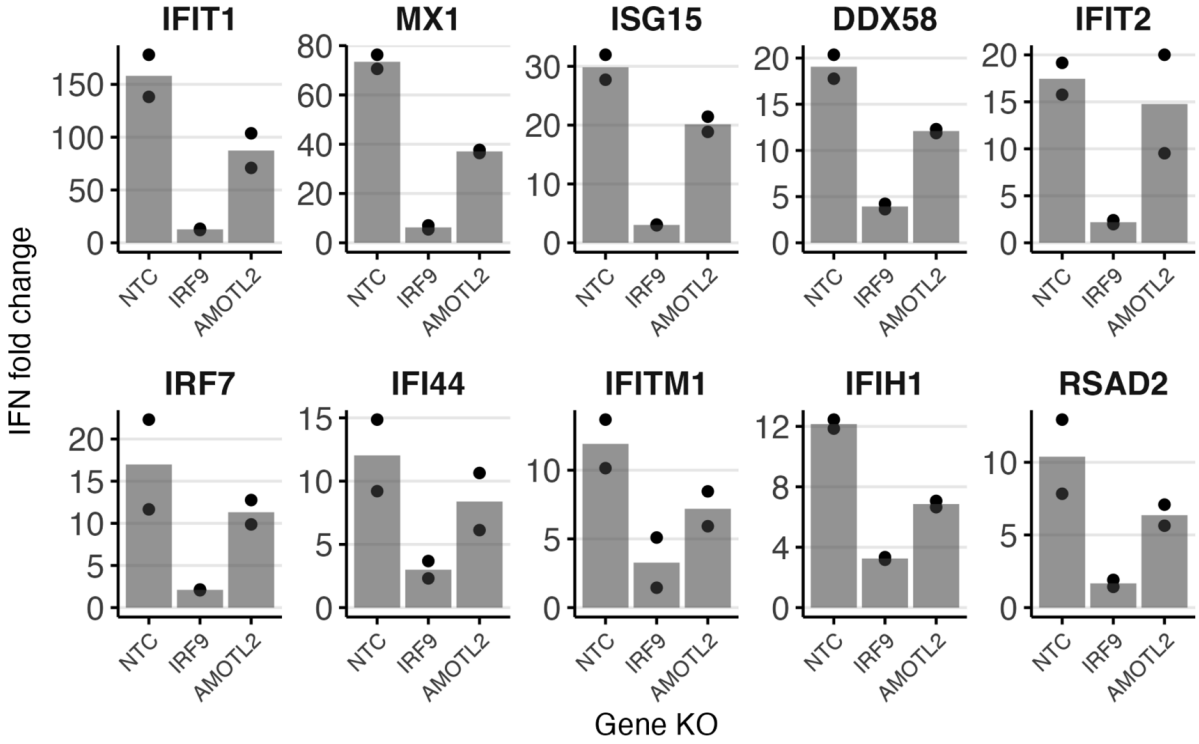

Figure S3. Nanostring nCounter quantification of ISG transcripts in AMOTL2 KO cells.
The 10 genes on the panel with at least 10-fold upregulation in response to 3 hours of 1000U/mL IFN-β treatment in NTC cells are shown. Upregulation of each gene in response to IFN-β treatment is shown for NTC, IRF9 KO, and AMOTL2 KO cells. Data points represent biological replicates.

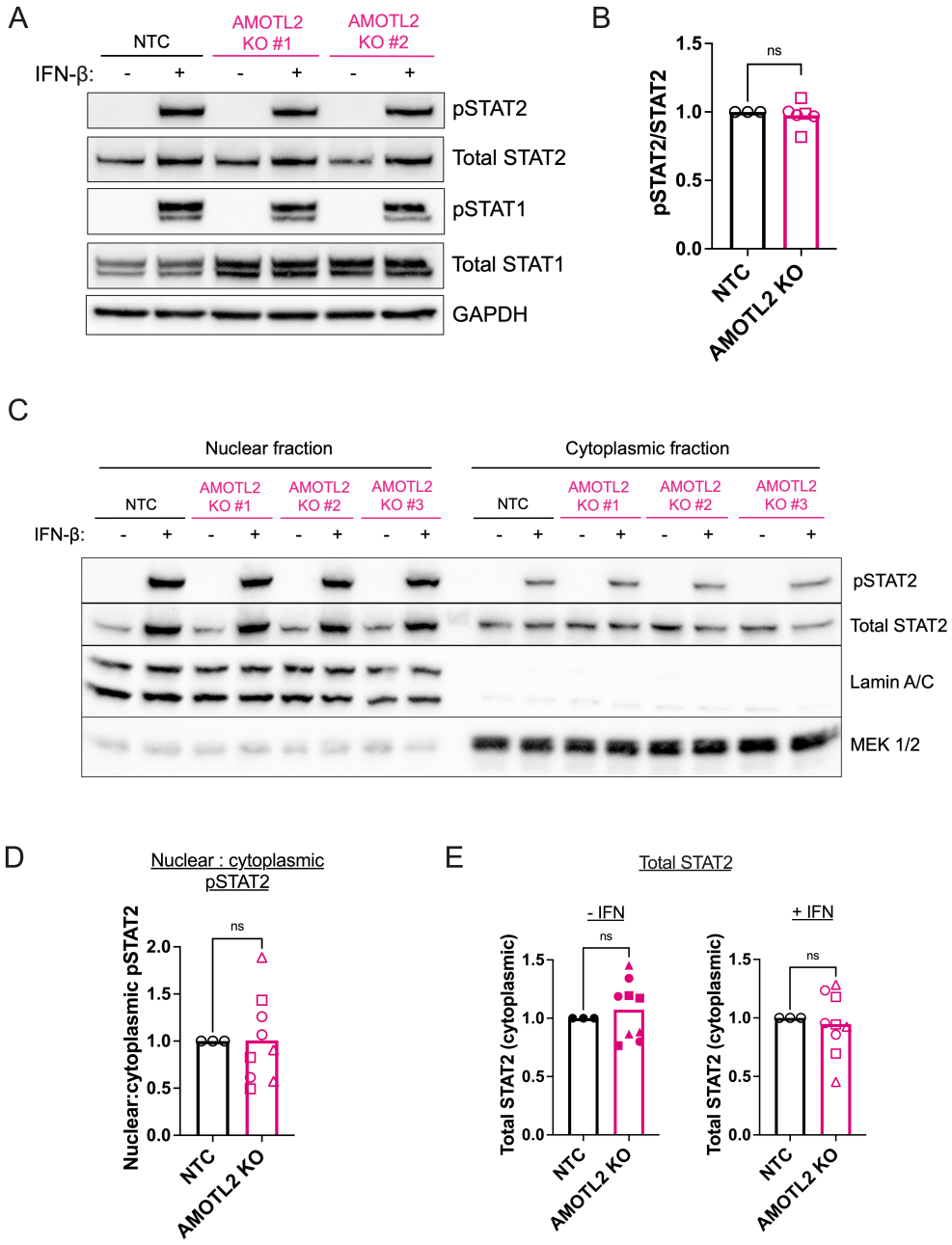

Figure S4. Levels of STAT2 phosphorylation in AMOTL2 KO cells, and total and phosphorylated STAT2 in nuclear and cytoplasmic fractions of AMOTL2 KO cells.
(A) Representative western blot corresponding to the data shown in Fig 2-5A. Whole cell lysates from NTC and AMOTL2 KO cell pools were non-IFN-treated or stimulated with 1000U/mL IFN-β for 1 hour and probed with antibodies to phosphorylated STAT1/2 and total STAT1/2. (B) Phosphorylated STAT2 levels relative to total STAT2 levels for the experiment shown in (A), normalized to NTC cells. Data show three independent replicates. (C) Representative western blot on nuclear and cytoplasmic fractions of non-IFN treated and IFN-treated NTC and AMOTL2 KO A549 cells, showing phosphorylated STAT2 and total STAT2 levels. IFN-β treatment was 1000U/mL for 2 hours. Lamin A/C is a nuclear marker, and MEK 1/2 is a cytoplasmic marker. (D) Nuclear to cytoplasmic ratio of phosphorylated STAT2 for three replicates of the experiment shown in (C), normalized to NTC cells. (E) Total cytoplasmic STAT2 levels for three replicates of the experiment shown in (C), normalized to NTC cells. In all panels, shapes represent different AMOTL2 KO cell pools, open circles represent IFN-treated conditions, closed circles represent non-IFN-treated conditions, and p values are Mann-Whitney test.

2.12 Materials and Methods

Cell culture and reagents

Human lung epithelial A549 cells (provided by A. Berger; ATCC CCL-185) were cultured in Gibco RPMI 1640 (ThermoFisher 22400089) supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, and 1x Gibco antibiotic/antimycotic ("RPMI complete media"), ThermoFisher 15240062). ZIKV infections were performed in RPMI without added FBS, L-glutamine, or antibiotic/antimycotic ("serum-free RPMI"). Vero cells (provided by A. Geballe) and HEK293T cells were maintained in Gibco DMEM (ThermoFisher 11965092) supplemented with 10% FBS, 2 mM L-glutamine, and 1x antibiotic/antimycotic ("DMEM complete media"). The identity of the A549 cells and HEK293T cells was confirmed using STR CODIS fingerprinting prior to beginning experiments.

The A549 cells were regularly tested for mycoplasma and found to be mycoplasma-free by the Research Cell Bank shared resource at the Fred Hutchinson Cancer Center.

ZIKV strain PRVABC59 was obtained from BEI Resources. ZIKV PRVABC59 was propagated in Vero cells at an MOI of 0.01, as previously described [116]. Viral titers were determined by the TCID50 assay described below.

Plasmids

pMD2.G (Addgene #12259, vesicular stomatitis virus glycoprotein [VSV-G] envelope plasmid) and psPAX2 (Addgene #12260, an HIV-based packaging plasmid) were gifts from Didier Trono. lentiCRISPRv2 (encoding puromycin resistance) was a gift from Feng Zhang (Addgene #52961) [117]. The expression plasmids encoding AMOTL2 (NCBI reference sequence NM_016201.4) and ORF stuffer (amino acids 2-83 of *E. coli* beta-galactosidase) were synthesized by VectorBuilder (Vector IDs VB240108-1277gwz and VB230727-1905uxr).

Whole plasmid sequencing was performed by Plasmidsaurus using Oxford Nanopore Technology for all plasmid preps used in this study.

CRISPR sgRNA library

The PIKA CRISPR library was designed as previously described [95]. In brief, ISGs were identified from gene expression profiles of IFN-stimulated cells, including microarray data from A549 cells [96,118]. A total of eight sgRNAs were selected for each of 1,905 gene targets. In total, 15,348 unique sgRNA sequences were synthesized, in addition to 200 unique non-targeting control (NTC) sgRNAs. The library was synthesized (Twist Biosciences) and cloned into pLentiCRISPRv2.

Generation of A549 KO library cells

Pseudoviruses for lentiviral transduction of the PIKA CRISPR library were made by transfection of HEK293T cells, as previously described [95], with the following specifics: $5x10^5$ HEK293T cells were plated in 2 mL in 6-well plates and the following day were transfected with 667 ng lentiCRISPRv2 plasmid, 500 ng psPAX2, and 162.5 ng pMD2.G. The comprehensive lentiviral CRISPR library was titered by colony forming unit assay on A549 cells and used to transduce A549 cells at an MOI of 0.5. Cells were selected in 1ug/mL puromycin for 14 days. Following selection, aliquots of $1x10^7$ A549 KO library cells/mL (> 500-fold representation) viably frozen until CRISPR screening. When cultured, cells were maintained in RPMI complete media with the addition of 1 μg/mL puromycin. Cells were maintained at ≥ 500-fold representation to avoid bottlenecking the library.

Annexin-V staining

A549 KO library cells per well were plated at a density of $3x10^5$ in 6-well dishes, either with or without 1000U/mL IFN-β (PBL Assay sciences 11410). Twenty-four hours later, cells were infected with ZIKV PRVABC59 at MOI=3 for 4-6 hours; uninfected cells were maintained in parallel. The viral inoculum was removed, and cells were replenished with RPMI complete media; in the IFN-treated condition, the refreshed media included 1000U/mL IFN- β. The cell supernatant was collected every six hours from 24 hours post infection to 48 hours post infection. A 2mL PBS wash was performed on the adherent cells and added to the supernatant fraction to capture any additional non adherent cells. Annexin-V staining was performed on the supernatant cells using the FITC Annexin V Apoptosis Detection Kit I (BD Biosciences) as per the manufacturer's protocol. Cells were analyzed by flow cytometry on a BD FACSCanto II flow cytometer. All data were analyzed using FlowJo v10 software.

CRISPR screen

To perform the CRISPR screen, $2.5x10^7$ A549 KO library cells were plated in a 5-layer flask (Corning Falcon) in 125 mL RPMI complete media containing 1000 U/mL IFN-β (PBL Assay Sciences) and placed at 37˚C. Twenty-four hours after plating and IFN-β pre-treatment, cells were infected with ZIKV PRVABC59 at MOI=3 for four hours in a total of 50 mL of serum-free RPMI inoculum. For the uninfected control condition, the method was identical except the cells received 50mL of serum-free RPMI without virus added. The inoculum was removed and replaced with 125 mL of fresh RPMI complete media containing 1000 U/mL IFNβ and placed at 37˚C. At 42 hours post infection, the supernatant was collected. A PBS wash was performed to capture any remaining dead/dying cells not adhered to the flask, and the PBS was collected and added to the cell supernatant. The supernatant including the PBS wash was spun at

1500 x g for 10 mins at room temperature. Following the spin, the supernatant was poured off and the pellets containing dead/dying cells were resuspended in PBS and spun again at 1500 x g for 10 mins. Genomic DNA was isolated using a Qiagen QIAamp DNA blood mini kit. Illumina library preparation was performed by two rounds of PCR using Herculase II Fusion DNA Polymerase (Agilent 600679). The round 1 primer sequences were 5'-GAGGGCCTATTTCCCATGATTCCTTCA-3' ("PLC-seq-R1-F") and 5'-AACTTCTCGGGGACTGTGG-3' ("PLC-Seq-R1-R"). Thermocycling conditions for round 1 were 95°C for 2 minutes; 12 rounds of 95*C for 15 seconds, 60°C for 20 seconds, and 72°C for 30 seconds; and 72°C for 3 minutes. A maximum of 2 µg of gDNA was used per first-round PCR reaction. Following the first-round PCR, all PCRs corresponding to the same sample were pooled and up to 12 first-round PCR reactions were purified on a single Qiaquick PCR purification column (QIAGEN 28104). Four second-round PCRs were performed per sample (10 µL PCR-purified first-round product as template) to add identifier sequences to each sample using indexing primers (see [95]) and the PLC-Seq-R2-RS primer (CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCACTTTTTCAAGTTGATAACGGACT). Thermocycling conditions for round 2 were 95°C for 2 minutes; 20 rounds of 95*C for 15 seconds, 60°C for 20 seconds, and 72°C for 30 seconds; and 72°C for 3 minutes. The second-round PCR products were pooled and run on a 2% TBE gel to ensure successful amplification of a ~230 bp product. A 0.7x right-sided AMPure XP (Beckman-Coulter) bead purification was performed as per the manufacturer's protocol using a 96-well plate magnet (Thermo Fisher). The bead-purified products were run on a 2% TBE gel to verify clean-up and their concentration was quantified using the QuBit dsDNA HS assay kit as per the manufacturer's protocol. An equimolar pool was made from all the samples, gel-purified

using a 2% TBE gel, and sequenced on an Illumina HiSeq (Fred Hutch Shared Resources Genomics Shared Resource).

Bulk RNA-Seq in A549 cells

Wildtype A549 cells were plated in triplicate in RPMI complete media with or without 1000U/mL IFN-β. 24 hours later, cells were harvested and total RNA was extracted with the RNeasy Plus Mini Kit (Qiagen 74134). RNA quality was assessed with RNA TapeStation (Agilent) to ensure acceptable RNA integrity values, and RNA-seq was performed with 50bp paired-end reads on an Illumina machine (Fred Hutch Shared Resources Genomics Shared Resource).

Generation of single- and double-gene KOs in A549 cells

Single-gene KOs were made by lentiviral transduction for the genes IRF9, IFI6, AMOTL2, GBP3, RMI2, HELZ2, BTN3A1, WWTR1 (TAZ), and YAP. Individual sgRNAs (Table S2) were cloned into the pLentiCRISPRv2 plasmid. Non-targeting sgRNA sequences were used to generate NTC cells in parallel. HEK293T cells were plated at a density of $2x10^6$ in 10mL in T-75 flasks and the following day were transfected with 5ug lentiCRISPRv2 plasmid with the appropriate sgRNA sequence, 5ug psPAX2, and 2.5ug pMD2.G. Viral supernatants were collected two days later, clarified by centrifugation to remove any cells in solution, and concentrated using 100kDa MWCO Amicon filter units (MilliPore Sigma UFC910096) until ~250uL of concentrated virus remained. A549 cells were plated in 6-well dishes at $2x10^5$ cells per well in 2mL of RPMI. The following day, cells were transduced with the concentrated lentivirus by spinoculation. In cases where single-gene KOs were generated with more than one sgRNA (see Table S2), the cells were transduced with the lentiviruses for both sgRNAs. Successfully transduced cells were selected by culturing in RPMI complete media supplemented

with 2ug/mL puromycin (Sigma Aldrich P9620) for 1-2 weeks or until flasks with non-transduced control cells were killed by puromycin. All KOs were verified by Synthego's ICE genomic editing analysis and/or Western blotting.

For AMOTL2/TAZ double KOs, we generated KOs by successive nucleofection of Cas9/sgRNA ribonucleoproteins (RNPs). We opted for this approach rather than the constitutive editing provided by the lentiviral system described above, since AMOTL2 and WWTR1 are both located on the same chromosome, and we wanted to avoid the possibility of a large intergenic deletion. The same two sgRNAs per gene as used in dual-guide lentiviral editing experiments were ordered as synthetic sgRNAs (Synthego CRISPRevolution sgRNA EZ Kit, modified with 2'-O-Methyl at 3 first and last bases and 3' phosphorothioate bonds between first 3 and last 2 bases). RNPs were assembled on the day of nucleofection by combining 4 µL of 50 pmol/µL pooled sgRNAs (Synthego), 20 µL of SF Cell Line Complete Nucleofector solution (Lonza V4XC-2032), and 1 µL of 20 µM Cas9-NLS protein (UC Berkeley QB3 MacroLab). Complexes were gently mixed and incubated for at least 10 min at room temperature. $2x10^5$ A549 cells were pelleted and washed once with PBS prior to resuspension in 25uL RNPs and immediate transfer to a 16-well Nucleocuvette Strip in an Amaxa Nucleofector (Lonza). Nucleofection was done using pulse code CM-130. Cells were immediately supplemented with 80 µL of RPMI complete media and allowed to recover in the cuvette for at least 10 min at 37°C. Cells were then transferred to 6-well dishes and supplemented with an additional 3mLs of RPMI complete media. The second RNP nucleofection was done in the same way between 3 to 4 days later. For cells not receiving a second edit (e.g., single KO control cells and NTC cells), the second nucleofection was performed with NTC RNPs. Since cells generated by RNP nucleofection do not have a selectable marker, we performed fresh nucleofections for each biological replicate to

minimize outgrowth of wildtype cells. All KOs were verified by Synthego's ICE genomic editing analysis and/or Western blotting.

Genomic editing analysis

A549 KO cell pellets were collected for genomic DNA extraction (QIAGEN 51104). Edited loci were amplified from the genomic DNA of edited and control cells using primers specific to the region around each sgRNA cut site, Herculase II Fusion DNA Polymerase (Agilent 600679), and the thermocycler program detailed in Table S3. PCR products were purified (QIAGEN 28104) and Sanger sequenced (Fred Hutch Shared Resources Genomics Core). Results were analyzed using the ICE analysis tool (Synthego) to estimate the proportion of DNA with a functional KO of the gene (KO score). These values are reported in Table S2.

IFN treatment and ZIKV infection of A549 cells

A549 KO cells were diluted to $8x10^4$ cells/mL in RPMI complete media, and 1mL of cells was plated for each IFN-negative well. IFN-β (PBL Assay sciences 11410) was added to the remaining cells at 1000U/mL, and 1mL of these cells was plated as each IFN-treated well. All conditions were performed in technical duplicate. Twenty-four hours later, cells were infected with 250uL PRVABC59 ZIKV at an MOI of 0.5 in serum-free RPMI. Four hours later, the viral inoculum was removed and cells were washed once with PBS, then replenished with fresh IFN-negative or IFN-treated RPMI complete media. For quantification of infectious virus by $TCID_{50}$, all viral supernatant was collected at the infection endpoint and frozen at -80°C. For quantification of ZIKV RNA by RT-qPCR, 12uL of viral supernatant was removed from each well at various time points throughout the infection time course. The viral supernatant was heat-inactivated at 60°C for 30 minutes and stored at -80°C.

ZIKV quantification by $TCID_{50}$

ZIKV titers were determined by $TCID_{50}$ assay on Vero cells as previously described [47]. Vero cells were seeded in 100 uL of DMEM complete media in a flat-bottomed 96-well plate at $8x10^3$ cells per well. The next day, for each condition tested, seven serial 10-fold dilutions of viral supernatants were prepared, starting at a concentration of 1 uL viral supernatant/well, with each dilution including 10 replicate wells. Cells were infected with 50 uL of each viral dilution in serum-free DMEM for 4–6 h, before being replenished with 100 uL of DMEM with 3% FBS. On day 5 post-infection the wells were examined by light microscope for the presence or absence of cytotoxicity, and the $TCID_{50}$/mL was calculated using the Spearman–Karber method.

ZIKV quantification by RT-qPCR

Viral supernatants were diluted 1:3 in water and used as the template in a one-step RT-qPCR reaction. In a 384-well MicroAmp Optical plate (Applied Biosystems 4309849), 3.95uL of diluted supernatant was mixed with 5uL iTaq Universal Probes PCR reaction mix (BioRad), 0.25uL iScript reverse transcriptase (BioRad), 0.3uL of each 10uM primer, and 0.2uL of 10uM probe targeting the NS1 region of the ZIKV PRVABC59 genome (Table S4) (adapted from [104]). Thermocycling conditions are listed in Table S3. Plates were read on an Applied Biosystems QuantStudio 7 Real-Time PCR machine. ZIKV RNA copies per mL were quantified using a standard curve of a positive control DNA fragment.

Human gene expression quantification by RT-qPCR

For assessment of AMOTL2 and ISG upregulation in unedited A549 cells, $4x10^5$ NTC cells were plated in technical duplicate in 6-well dishes in 2mL RPMI complete media, with or without 1000U/mL IFN-β (PBL Assay Sciences). Cells were lysed in buffer RLT+ (Qiagen) with

beta-mercaptoethanol and stored at -80°C until RNA extraction and cDNA synthesis as described below.

For assessment of IFIT1 and Mx1 upregulation in AMOTL2 KO cells, $4x10^5$-$6x10^5$ AMOTL2 KO and NTC A549 cells were plated in technical duplicate in 6-well dishes in 2mL RPMI complete media, with or without 1000U/mL IFN-β (PBL Assay Sciences). Cells were lysed in buffer RLT+ (Qiagen) with beta-mercaptoethanol and stored at -80°C until RNA extraction and cDNA synthesis as described below.

For assessment of basal STAT1 transcript levels in AMOTL2 KO cells, cells nearing confluency were lysed in buffer RLT+ (Qiagen) with beta-mercaptoethanol and stored at -80°C until RNA extraction and cDNA synthesis as described below.

RNA was extracted using the RNeasy Plus Mini Kit (Qiagen 74134). In addition to the genomic DNA eliminator spin columns included in the kit, an additional DNase treatment was performed either during the extraction protocol with an on-column DNase digest (Qiagen 79254, following the manufacturer's protocol) or following the extraction (Roche 4716728001). For the latter, 20uL RNA was mixed with 2.5uL 10X incubation buffer, 0.5uL RNAase OUT, 0.5uL DNase I, and 1.5uL water and incubated at 25°C for 1 hour, then 75°C for 5 minutes (Roche 4716728001). cDNA synthesis was performed by first incubating 12uL of extracted RNA with 1uL of 10mM dNTPs, 0.75uL of random hexamer primers (Thermo Fisher N8080127), and 1.25uL of nuclease-free water at 65°C for 5 minutes to denature RNA and allow for primer annealing. The reaction was placed on ice for 2 minutes, then mixed with 4.4uL of 5x First Strand Buffer (Thermo Fisher 18080093) 1uL 0.1M DTT (Thermo Fisher 18080093), 1uL RNAaseOut (Thermo Fisher 10777019), and 1.5uL Superscript III reverse transcriptase (Thermo Fisher 18080093). Reverse transcription was carried out using the thermocycling conditions

described in Table S3. A "no-RT" control, in which reverse transcriptase was replaced with water, was included for a subset of samples from each RNA extraction and run in the qPCR reaction to ensure adequate genomic DNA removal. For qPCR reactions, 4.4uL of the cDNA template was mixed with 5uL iTaq universal SYBR Green mix (Bio-Rad 1725120) and 0.3uL of each 10uM primer for the gene of interest (Table S4). GAPDH was used as a housekeeping gene. Plates were read on an Applied Biosystems QuantStudio 7 Real-Time PCR machine, and relative quantification was calculated using the $\Delta\Delta Ct$ method [119].

Assessment of ISG upregulation by Nanostring nCounter

$4x10^5$ AMOTL2 KO, IRF9 KO and NTC A549 cells were plated in technical duplicate in 6-well dishes in 2mL RPMI complete media, with or without 1000U/mL IFN-β (PBL Assay Sciences). Cells were lysed in buffer RLT+ (Qiagen) with beta-mercaptoethanol, and RNA was extracted using the RNeasy Plus Mini Kit (Qiagen 74134) following the manufacturer's protocol. RNA samples were prepared for NanoString nCounter analysis as described previously [120]. Briefly, 6uL of 20ng/uL RNA was hybridized to reporters and capture probes from a custom innate immune probe set (NanoString) [120]. Hybridization was performed for 16 hours at 65°C, and sample preparation was completed on an nCounter MAX prep station according to the manufacturer's protocol (NanoString). RNA transcript counts were quantified on a NanoString nCounter (NanoString).

AMOTL2 complementation

Lentiviral expression plasmids encoding AMOTL2 and ORF stuffer were synthesized by VectorBuilder. To allow for successful complementation of AMOTL2 in the lentiCRISPR AMOTL2 KO cells, the lentiviral expression plasmid encoding wildtype AMOTL2 was edited by site-directed mutagenesis (SDM) to mutate the sgRNA sequence and/or PAM sites for each

sgRNA expressed by the AMOTL2 KO cells. An additional mutation was made to mutate the reference sequence's glutamic acid at position 732 to an aspartic acid, as we noticed the A549 cells express an aspartic acid at this position according to RNA-Seq reads. Then, individual AMOTL2 point (P105A, Y213A, S159A) and deletion (CCdel, PDZdel) mutants were synthesized by SDM. The primers to generate all mutations were designed using TakaraBio's In-Fusion Cloning Primer Design Tool and used in PCR reactions with Phusion High-Fidelity DNA polymerase according to the manufacturer's instructions. PCR products underwent DpnI treatment to remove the template plasmid, followed by PCR purification (Qiagen 28104). The purified PCR product was transformed into One Shot Stbl3 Chemically Competent E. coli (Thermo Fisher C737303), and single colonies were swabbed and grown overnight. Purified plasmid preparations (Qiagen 27104) were sequenced by Plasmidsaurus to check for the intended mutation(s).

$2x10^6$ 293T cells were plated in 10mL in T-75 flasks and the following day were transfected with 2.5ug lentiviral expression plasmid, 5ug psPAX2, and 2.5ug pMD2.G. Viral supernatants were collected two days later, clarified by centrifugation to remove any cells in solution, and concentrated using 100kDa MWCO Amicon filter units (MilliPore Sigma UFC910096) until ~250uL of concentrated virus remained. AMOTL2 KO cells were plated in 6-well dishes at $2x10^5$ cells per well in 2mL of RPMI. The following day, cells were transduced with the concentrated pseudotyped lentivirus by spinoculation. Successfully transduced cells were selected by culturing in RPMI supplemented with 100ug/mL hygromycin for 1-2 weeks or until flasks with non-transduced control cells were killed by hygromycin. AMOTL2 protein expression was confirmed by RT-qPCR and Western blot.

Co-immunoprecipitation

A549 cells with the indicated gene KOs and treatments were harvested at the indicated time points and lysed in Pierce IP buffer (87787) containing protease inhibitor (Thermo Fisher A32955) or protease/phosphatase inhibitor (Cell Signaling 5872). Samples were flash-frozen and stored at -80°C until further processing. Upon thawing, lysates were sonicated, then spun at maximum speed for 15min at +4°C. Protein concentrations were quantified using the DC Protein Assay (Bio-Rad 5000116) against a standard curve of bovine serum albumin (BSA). Samples were diluted to equivalent concentrations and volumes. Lysates were pre-cleared by incubating with protein G beads at +4°C for at least 30 minutes. Meanwhile, protein G beads were incubated with the indicated IP antibody at +4°C for at least 30 minutes. The antibodies used for IP were: anti-TAZ (mouse IgG2a monoclonal antibody, Cell Signaling 71192), anti-STAT1 (mouse IgG1 monoclonal antibody, Cell Signaling 9176), anti-phospho-STAT1 Tyr701 (mouse IgG2a monoclonal antibody, Thermo Fisher 33-3400), mouse IgG2a isotype control antibody (Cell Signaling 61656), and mouse IgG1 isotype control antibody (Cell Signaling 5415)**.** Pre-cleared whole cell lysates were then incubated with the antibody-beads overnight at +4°C. The beads were washed thrice in lysis buffer before eluting off the bound proteins by resuspending beads in LDS sample buffer (Invitrogen NP007) containing 50mM DTT and heating to 70°C for 10 minutes. Immunoblotting was performed as described below.

Nuclear fractionation

Confluent T-75 flasks of A549 NTC and AMOTL2 KO cells were treated with 1000U/mL IFN-β for 2 hours or left untreated. Nuclear fractionation was performed using the NuCLEAR™ Extraction Kit (Sigma-Aldrich) according to the manufacturer's protocol, with the following changes: cells were trypsinized instead of scraped, and the nuclear pellet was washed

twice in PBS to remove residual cytoplasmic proteins prior to nuclear extraction. Cytoplasmic and nuclear extracts were quantified and processed for immunoblotting as described in the section below.

Immunoblot analysis

For assessment of protein levels in whole cell lysates without subcellular fractionation or immunoprecipitation, cells were trypsinized, pelleted, washed once in PBS, and stored at -80 until lysis in RIPA buffer (Cell Signaling 9806) with protease/phosphatase inhibitor (Cell Signaling 5872). Lysates were vortexed on a thermomixer at +4°C for 15 minutes, then spun at maximum speed at +4°C for 15 minutes. Protein concentrations were quantified using the DC Protein Assay (Bio-Rad 5000116) against a standard curve of bovine serum albumin (BSA). Samples were diluted to equivalent concentrations and heated to 70°C for 10 minutes in LDS sample buffer (Invitrogen NP007) containing 50mM DTT. The protein extracts were separated on NuPAGE Bis-Tris Mini Protein gels (Invitrogen NP0321), transferred to nitrocellulose membranes (Invitrogen LC2001), blocked with 3% BSA, and probed with the indicated antibodies overnight at +4°C. The following primary antibodies were used: anti-AMOTL2 (Thermo Fisher PA5-78770, 1:1000), anti-TAZ (Cell Signaling 4883, 1:1000 or 83669, 1:1000), anti-STAT1 (Cell Signaling 9176, 1:1000 or 9172, 1:1000), anti-pSTAT1 (Cell Signaling 9167, 1:1000), anti-STAT2 (Cell Signaling 72604, 1:1000), anti-pSTAT2 (Cell Signaling 88410, 1:1000), anti-lamin A/C (Cell Signaling 4777, 1:2000), anti-MEK1/2 (Cell Signaling 8727, 1:1000). Membranes were probed with secondary antibody for 1 hour at room temperature using HRP-linked anti-rabbit IgG (Cell Signaling 7074, 1:10,000) or HRP-linked anti-mouse IgG (Cytiva NA931, 1:10,000) and visualized with chemiluminescence (Cytiva RPN2232). For GAPDH staining, membranes were probed with HRP-linked anti-GAPDH antibody (Bio-Rad

MCA4739P, 1:1000). When necessary, membranes were stripped (Thermo Scientific 46430) and re-probed with a different antibody. Membranes were washed in TBS with 0.1% Tween-20 between each step.

Illustrations

Figures were generated using BioRender and eulerr.co.

CRISPR screen analysis

Deep sequencing reads were demultiplexed, trimmed, and aligned to the sgRNA library using Bowtie v1.0.0 [121]. The aligner's native –trim5 and –trim3 options were used to trim down the raw reads to 20nts, which is the expected guide size. In an uninfected, IFN-treated control condition used to assess sgRNA coverage of the A549 KO library cells at baseline and performed in parallel to the experimental conditions of the screen, 97.6% of genes had at least six of their eight sgRNAs sequenced at high representation. This ensured to us that the setup of the screen allowed for high representation of the sgRNAs in the PIKA library. For screen analysis, relative enrichment of sgRNAs (based on all eight sgRNAs targeting each gene) and genes were analyzed and compared between groups of interest using the Model-based Analysis of Genome wide CRISPR/Cas9 Knockout (MAGeCK) screens algorithm, v0.5.9.2 [122,123]. Statistical scores were assigned to identify the most enriched genes, indicating factors important for IFN-mediated restriction of ZIKV infection.

Bulk RNA-seq analysis

The mapping of RNA-seq reads to genome was performed with the Spliced Transcripts Alignment to a Reference (STAR) 2.7.9a software [124]. We aligned against the GRCh38 human genome reference using annotations from GENCODE [125]. To generate the list of ISGs in

A549 cells, the gene counts were then analyzed with edgeR [126]: we applied the filterByExpr function requiring a gene have at least 10 counts in some samples and 15 counts in total across all samples; we used the glmQLFTest function to test if the variances for the gene between IFN-treated and untreated samples were significantly different. ISGs were required to pass the edgeR filter and have $p < 0.05$. Furthermore, we selected genes that had counts per million (CPM) greater than 1 in at least half the samples across both IFN-treated and untreated conditions. Then, taking the average CPMs under each condition, we computed the log (base 2) fold change of the IFN-treated condition over the untreated. A gene was considered differentially expressed if the log fold change was greater than 2. The code to select differentially expressed genes was implemented in Python 3.8.10 [127] using the Numpy 1.23.5 [128] and Pandas 1.5.2 [129] packages.

NanoString data analysis

NanoString data were processed as previously described [120,130]. Briefly, raw reads were assessed for technical quality control flags using nSolver software (Nanostring) and transferred into RStudio (Boston, MA, USA). We used count data from housekeeping genes on the panel as an internal reference, against which we compared our genes of interest. We measured the coefficient of variation for all genes in the probe set, including genes of interest, housekeeping genes, and positive and negative control genes. All housekeeping genes had small coefficients of variation (CVs) between all samples in the probe set, thus we opted to calculate the geometric mean of all housekeeping genes to use as the within-sample reference value. All raw count data from each sample was “anchored” to the within-sample housekeeping gene average in ratio form to generate gene expression values. IFN fold change values were calculated as the ratio between IFN-treated and non-IFN treated gene expression values for each gene

within each sample. The data were filtered to include only ISGs (defined as genes with an average IFN fold change of at least 10 in NTC cells).

Other data analysis

Prism 10 was used for statistical tests and display of quantitative data. Digital images were processed with ImageJ.

Chapter 3 Perspectives and future directions

3.1 A novel CRISPR screening system leads to the discovery of an antiviral gene that contributes to IFN-mediated restriction of ZIKV despite not being IFN-induced

ZIKV remains a persistent threat to global public health due to its potential to cause devastating health outcomes in neonates and its predicted increasing burden of disease in coming decades. While our understanding of ZIKV has certainly improved since the 2016 American outbreak, much remains to be discovered about the virus, including which human innate immune genes respond to ZIKV infection. For my thesis work, I set out to identify and characterize human genes that contribute to type I IFN-mediated restriction of ZIKV by following up on the results of large CRISPR screens performed by a previous graduate student in our lab, Ted Gobillot, MD, PhD. To our knowledge, this study represents the largest-scale loss-of-function screen to identify human antiviral factors against ZIKV. Prior large-scale screens have used overexpression or gene activation, which may lead to the identification of genes that are antiviral when overexpressed but are not relevant to the antiviral response at their endogenous expression levels [54,74]. Other studies have chosen antiviral factors to test against ZIKV based on their known roles in the restriction of other viruses, which likely leaves other unknown antiviral genes undiscovered. In the largest loss-of-function screen to identify genes relevant to IFN restriction of ZIKV to date, a 2022 study used siRNAs to screen 386 human genes for their contribution to IFN restriction of ZIKV in microglial cells. Surprisingly, their work identified a non-ISG that contributes to IFN restriction of ZIKV and other flaviviruses, highlighting the potential for the discovery of unexpected, novel antiviral genes through loss-of-function screening [75]. Our method used the gold standard of CRISPR knockout to screen a much larger gene set enriched for innate immune genes (1905 genes) in A549 lung epithelial cells, a commonly used cell type

for studying ZIKV infection. Importantly, there is a robust type I IFN antiviral response to ZIKV in these cells.

Our large CRISPR library encompassed more than 80% of the genes upregulated by type I IFN in A549 cells—the subset that we thought was most likely to contain antiviral genes—but also included over 1,700 other human genes. Unexpectedly, the top hit in our screen was AMOTL2, which we showed is not an ISG in A549 cells but nonetheless has an antiviral phenotype that is dependent on the presence of type I IFN. My experiments uncovered that AMOTL2 enhances the strength of the type I IFN signaling pathway itself, rather than acting as a downstream effector like an ISG. AMOTL2 knockout constitutively increases levels of U-STAT1, reduces STAT1 phosphorylation and nuclear translocation after type I IFN stimulation, reduces the upregulation of ISGs, and increases ZIKV replication. This intriguing result suggests that AMOTL2's antiviral effect stems from enhancement of the type I IFN response itself and raises many further questions, which are detailed later in this chapter.

Our CRISPR screening method, developed by Dr. Gobillot, takes advantage of the cytopathic effect of ZIKV in A549 cells to screen for cells that succumb to ZIKV-induced cell death. We show that the screen successfully enriched for key regulators of type I IFN signaling, and I further demonstrated that one of the top hits, IFI6, which is an ISG and known flavivirus restriction factor in other cell types [53,54], was antiviral against ZIKV in single-gene KO experiments in A549 cells. As CRISPR methods continue to improve and enable editing in primary cells and organoid models, this simple and innovative screening platform lays the groundwork for future, even larger-scale knockout screens to identify host factors that modulate ZIKV under a variety of conditions and in diverse cell types relevant to ZIKV infection. Moreover, it can be adapted to other viruses that cause cytopathic effects.

3.2 Limitations of and possible improvements to the ZIKV cell death CRISPR screen

While our CRISPR screening method resulted in an exciting novel finding, limitations of the method prevented us from identifying all the genes that contribute to type I IFN restriction of ZIKV. Additionally, some previously identified ZIKV-restricting ISGs that were included in our CRISPR library were not hits in the screen. These discrepancies can sometimes be explained by differences in the experimental system (i.e., previous studies used overexpression instead of knockout, and/or used different cell types than A549s). However, at least one described ZIKV-restricting ISG, PARP12, was previously shown to increase ZIKV infection when knocked out in A549 cells, yet it was not a hit in our screen [57]. This suggests that there is room for improvement in our CRISPR method.

One limitation may be the use of ZIKV-induced cell death as the readout in the screen. While we successfully identified two genes that play a role in the type I IFN response against ZIKV – AMOTL2 and IFI6 – we may have missed other genes with a smaller effect. This could be due to the fact that some knockouts increased ZIKV replication, but not enough to cause cell death at the 42-hour timepoint that was used in the screen. In this vein, it is possible that changing the timing of the screen would reveal other hits that act faster or slower than AMOTL2 and IFI6. Future work could develop a screening protocol based on ZIKV-positivity of cells (e.g., by flow cytometry and sorting for cells expressing ZIKV proteins). This method might be more sensitive than the cell death readout, and earlier time points could be collected (prior to the development of cytopathic effect).

Another possibility is that because each ISG is thought to contribute a small proportion of the overall IFN restrictive effect, single gene knockouts have small effect sizes that the screen is not sensitive enough to pick up. Additionally, maybe some ISGs work in tandem, and we would

detect a much greater effect if both genes were knocked out at once (e.g., a synergistic effect). To overcome this, future screens could perform double- or triple- gene knockouts by using a lentiviral CRISPR plasmid with two to three (or more) sgRNA sequences. The results would tell us which combinations of genes work together to have the greatest effect on IFN restriction of ZIKV.

3.3 AMOTL2 affects unphosphorylated STAT1 levels

One of the most surprising results of my thesis work is the finding that AMOTL2 knockout constitutively increases levels of U-STAT1, both at baseline and in the presence of type I IFN. This indicates that AMOTL2 decreases U-STAT1 expression. U-STAT1, the form of STAT1 that is not phosphorylated, is an ISG that is normally upregulated following IFN treatment and along with unphosphorylated STAT2 and IRF9, can serve as a transcription factor to upregulate certain genes in the days following IFN treatment. Because levels of phosphorylated STAT1 and STAT2 peak and decline quickly in the hours following IFN stimulation, the unphosphorylated complex is thought to drive late-phase ISG expression to maintain a longer-acting antiviral state. Despite this important role in the antiviral response, constitutively high levels of U-STAT1 can paradoxically make cells *less* responsive to type I IFN treatment by blunting the activation of STAT1 and STAT2 in the canonical IFN signaling pathway. I propose a model in which AMOTL2's presence in the cell suppresses high levels of U-STAT1, allowing for a more robust response to type I IFN treatment, higher ISG expression, and better control of viral replication (Figure 2-7). My data do not suggest any effect of AMOTL2 on basal STAT2 levels or STAT2 activation following IFN treatment, revealing an effect specific to STAT1. This suggests that perhaps the AMOTL2 knockout cells are forming a known non-canonical alternative to ISGF3 – a STAT2 homodimer bound to IRF9 – at higher

levels than wildtype cells [131]. Even with normal STAT2 activation, the AMOTL2 knockout cells cannot make up for the effect of blunted STAT1 activation on ISG transcription and viral restriction. This would suggest that the non-canonical complex is less potent than the canonical ISGF3. Clearly, future work is needed to investigate these possibilities.

Much remains unknown about this newly identified association between AMOTL2 and STAT1. Firstly, it is unclear how exactly AMOTL2 modulated STAT1 levels and whether other cellular players are involved. While I found that STAT1 mRNA transcripts are elevated in AMOTL2 KO cells, this does not necessarily mean that AMOTL2 affects STAT1 transcription, due to the fact that U-STAT1 itself may be increasing transcription of STAT1, as has been reported in other cell types [85]; therefore AMOTL2 could be affecting STAT1 protein levels, which feed forward to drive transcription of more STAT1. To clarify this point, it would be helpful to perform assays to determine whether AMOTL2 affects STAT1 transcript stability or protein stability. The former could be tested by measuring STAT1 mRNA half-life in the presence and absence of AMOTL2 in actinomycin D treated cells (which prevents RNA polymerase activity). Similarly, protein half-life could be measured in cycloheximide treated cells (which prevents protein synthesis). It would also be informative to test whether AMOTL2 can directly bind to STAT1. I was unable to detect binding between the endogenous proteins by co-immunoprecipitation in preliminary experiments (data not shown), but it would be prudent to also test binding of exogenously overexpressed AMOTL2 and STAT1, as endogenous co-immunoprecipitations can lack sensitivity for weaker interactions. Proximity-dependent labeling techniques like BioID could also be used to detect proteins that associate with AMOTL2 and STAT1 in an unbiased manner, and any overlapping hits could reveal proteins that link AMOTL2 to STAT1.

Because U-STAT1 is thought to maintain low-level expression of a subset of ISGs, it is interesting that I observed no difference in viral replication dynamics in the absence of type I IFN treatment, given that U-STAT1 levels are high at baseline in AMOTL2 KO cells. In other words, one might expect the presence of U-STAT1 itself to result in *reduced* viral replication in the non-IFN-treated condition. This perhaps suggests that the level of U-STAT1 is high enough to prevent activation of the canonical IFN signaling pathway (as in the IFN-treated condition), but not high enough to exert any appreciable antiviral effects of its own, at least in the time points I used in this study. To investigate this idea, future work might look at later time points of viral replication in the AMOTL2 KO cells, when phosphorylated STAT1 levels are undetectable and U-STAT1 levels are at their highest (U-STAT1 levels have been reported to remain high in response to IFN treatment for more than 72 hours in other cell types [85]). It is at this point that we may see AMOTL2 KO cells' relatively high U-STAT1 levels, relative to NTC cells, begin to cause antiviral effects. My data at 72-hour ZIKV infection time points may begin to hint at this trend, as we see AMOTL2 KO cause a subtle antiviral effect in the non-IFN-treated condition at 72 hours, while AMOTL2 reconstitution causes a subtle proviral effect in the non-IFN-treated condition at 72 hours (Figs 2-2C, 2-3). This proposed increase in U-STAT1 levels in the non-IFN-treated condition at 72+ hours of infection would be the result of virus-induced endogenous IFN secretion by the A549 cells. This idea is portrayed in Figure 3-1. If this hypothesis is true, an interesting question is then whether the ultimate outcome of AMOTL2 KO and elevated U-STAT1 levels *in vivo* is antiviral or proviral, which could be studied by moving this experimental system to animal models of ZIKV infection.

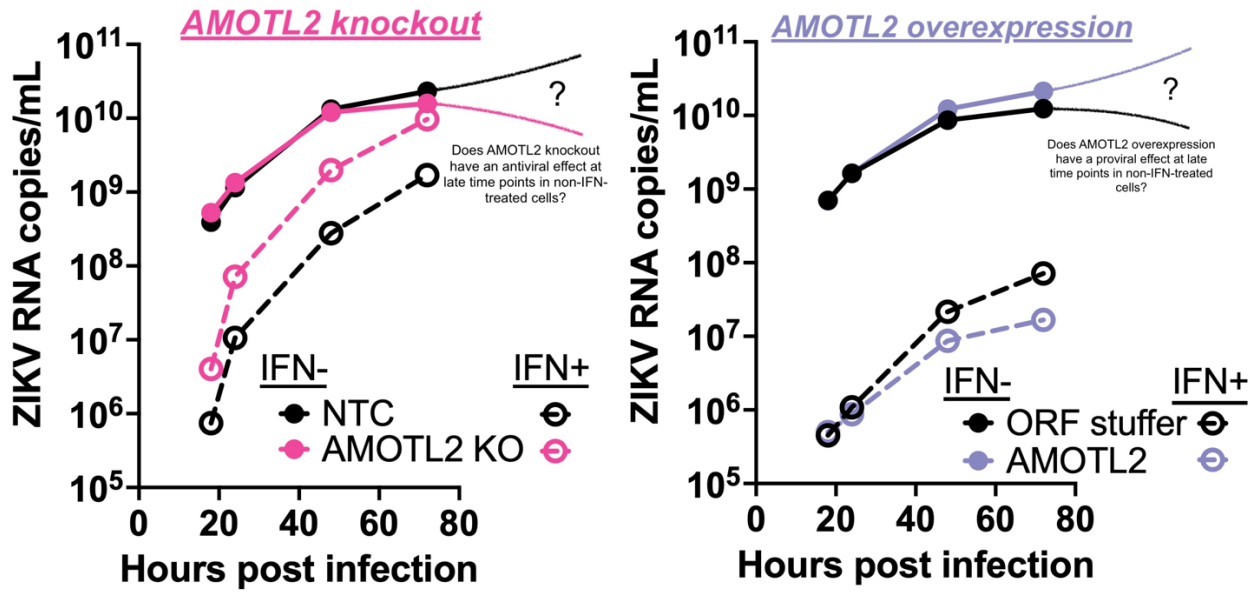

Figure 3-1. Hypothesis for the effect of AMOTL2 KO and increased U-STAT1 levels on viral replication at late time points.
Data portrayed in thick lines are experimental results, taken from Figures 2-2C and 2-3C. Thin lines after the 72 hour infection time point are hypothetical data if the trend observed at 72 hours were to continue. In this hypothesis, increased U-STAT1 levels in the AMOTL2 KO cells exert a late-stage antiviral effect in the non-IFN-treated condition, in contrast to the observed proviral effect caused by AMOTL2 KO and high U-STAT1 levels in earlier time points in the IFN-treated condition.

A useful experiment to confirm the association between AMOTL2 KO, U-STAT1, and reduced responsiveness to type I IFN would be to perform RNA-seq on AMOTL2 KO cells vs. NTC cells, at baseline and at multiple time points following type I IFN treatment. The subset of ISGs upregulated by U-STAT1 has been defined in other cell types and includes IFI27, BST2, OAS2, and IFI44 [85], which are not genes I explicitly examined in my thesis work. Compared to NTC cells, I would expect RNA-seq to reveal higher levels of U-STAT1 ISGs in AMOTL2 KO cells at baseline, lower levels of fast-acting ISGs (as I have already observed by qPCR and Nanostring) in the hours following IFN treatment due to the inhibitory effect of U-STAT1 on IFN responsiveness, and potentially higher levels of U-STAT1 ISGs at later time points of IFN treatment, as the constitutive effect of U-STAT1 becomes dominant again following resolution

of the acute IFN response. RNA-seq would allow an unbiased look at how AMOTL2 KO affects expression of all human genes, which would be helpful given the complexity of the proposed model. However, this question could also be answered with a carefully designed gene panel by qPCR.

3.4 Impact of AMOTL2 on type I IFN restriction of ZIKV in other cell types

I chose to use A549 lung epithelial cells for this work because they are a commonly used model for ZIKV infection, are highly responsive to type I IFN treatment, exhibit cytopathic effect in response to ZIKV infection (as needed for the CRISPR screen read-out), and are easily editable by CRISPR [47,48]. With my discovery that AMOTL2 has antiviral function against ZIKV in A549 cells, an important next step is to test AMOTL2's role in ZIKV restriction in more biologically relevant cell types. For example, primary monocyte-derived macrophages (MDMs) would be a useful system in which to study this question. Macrophages are a major target of ZIKV infection *in vivo* [114,115] and are also highly relevant to IFN signaling, as macrophages are major IFN producers and responders *in vivo* [39]. Interestingly, AMOTL2 is an ISG in primary MDMs according to online gene expression datasets from three independent studies (interferome.org v2.01), with AMOTL2 increasing 6-24 fold in response to 3-4 hours of type I IFN stimulation, depending on the dataset. AMOTL2 is not reported to be an ISG in any other cell type. This suggests that AMOTL2 may be important in the type I IFN response mounted by macrophages in response to infection. Primary monocytes isolated from human blood can be efficiently edited by CRISPR [132] before differentiation into macrophages with the cytokine M-CSF, and these cells are highly sensitive to *in vitro* IFN treatment [133]. In preliminary work towards moving my experimental system into primary MDMs, I found that these cells can be productively infected with ZIKV (strain PRVABC59; Figure 3-2). The next

steps towards this goal would be to inactivate AMOTL2 in the MDMs by CRISPR and observe the effect on ZIKV replication in the presence and absence of exogenous type I IFN treatment.

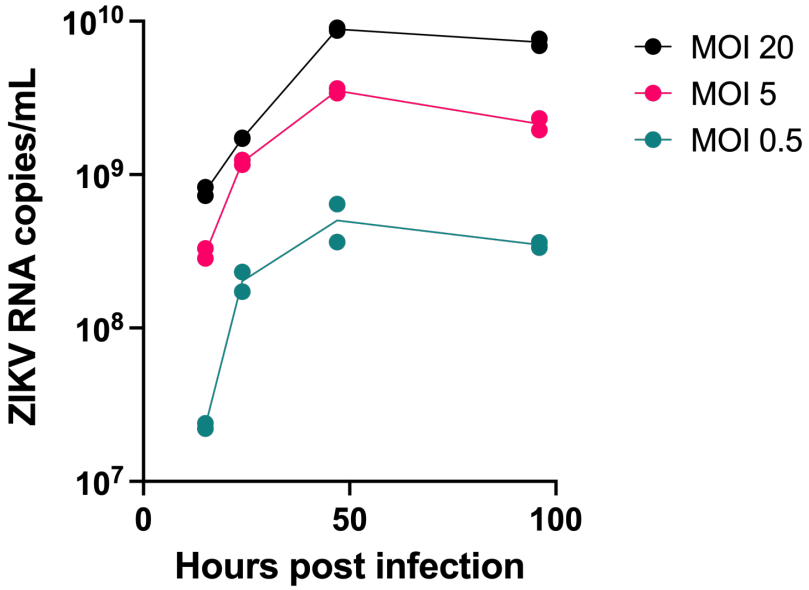

Figure 3-2. ZIKV successfully infects primary monocyte-derived macrophages.
Data are preliminary from one experiment. Data points represent technical replicates.

3.5 Impact of AMOTL2 on type I IFN restriction of other viruses

Based on my findings that AMOTL2 enhances the potency of the type I IFN response itself, rather than directly binding and inhibiting ZIKV, it seems likely that AMOTL2 contributes to type I IFN restriction of diverse viruses. In parallel to the ZIKV infection experiments in primary MDMs that I describe above, it would be useful to infect my existing AMOTL2 KO A549 cells with a panel of diverse viruses to decipher the limits, if any, of AMOTL2's antiviral activity. I would propose to begin with a small panel of viruses of different families that includes enveloped and non-enveloped viruses with RNA and DNA genomes, and that encompasses viruses with a diversity of known interactions with the human type I interferon response. For example, to supplement ZIKV (an enveloped, positive-sense RNA flavivirus), this panel could include influenza A virus (an enveloped, negative-sense RNA orthomyxovirus), herpes simplex

virus (an enveloped DNA alphaherpesvirus), and adenovirus 5 (a non-enveloped DNA adenovirus). Between the primary MDM experiments I propose above (testing AMOTL2's contribution to IFN restriction of ZIKV in a different cell type) and the experiments proposed here (testing AMOTL2's contribution to IFN restriction of other viruses in the same cell type), the results should cleanly reveal whether AMOTL2's antiviral phenotype is cell-type specific, virus-specific, both, or neither.

3.6 AMOTL2's antiviral mechanism is independent of TAZ binding

In my thesis work, I pursued experiments to address the possibility that AMOTL2 is acting through YAP and/or TAZ to exert its antiviral effect. YAP and TAZ are well-described binding partners of AMOTL2 with numerous described roles in antagonism of innate immune signaling pathways. In other contexts, AMOTL2 sequesters YAP/TAZ away from performing their described canonical functions as transcriptional activators. Thus, it seemed plausible that AMOTL2 may exert its antiviral role by inhibiting YAP/TAZ antagonism of the type I IFN signaling pathway. I found that TAZ, but not YAP, has a proviral role towards ZIKV in IFN-treated cells. Previous work showed a role for a short isoform of TAZ in antagonizing IFN restriction of vesicular stomatitis virus (VSV) and herpes simplex virus (HSV) in different cell types, so my finding that this phenotype extends to ZIKV infection in A549 cells is novel and expands our understanding of the breadth of TAZ's proviral function. I also showed that AMOTL2 and TAZ interact in A549 cells. To definitively say whether TAZ was involved in AMOTL2's phenotype, I complemented the AMOTL2 KO cells with various AMOTL2 constructs lacking different functional domains. My results clearly show that the TAZ binding mutant rescued the antiviral phenotype to the same extent as wildtype AMOTL2, suggesting that TAZ is not involved in AMOTL2's mechanism. This suggests that AMOTL2's mechanism is

wholly novel and not based on inhibition of TAZ. It is interesting that these two proteins that are well-described binding partners are involved in the type I interferon response via distinct mechanisms.

In parallel with the TAZ binding (PPQY) mutant, I tested other AMOTL2 mutation/deletion constructs, including most of the domains shown in Figure 3-3. Interestingly, mutating AMOTL2's coiled-coil domain (blue box in Figure 3-3) failed to rescue the phenotype, implicating this domain in the antiviral function. Coiled-coil domains are structural motifs that involve supercoiled alpha-helices, are found in roughly 5% of proteins, and have diverse functions, but often facilitate protein-protein interactions. Homo- or hetero-oligomerization can be mediated by coiled-coil domain interactions, and AMOTL2 has been shown to homo-oligomerize using its coiled-coil domain [108]. Interestingly, STAT1 also has a coiled-coil domain, thus it is possible that AMOTL2 and STAT1 interact directly via these domains; however, as coiled-coil domains are not uncommon, this could also be coincidence. It is also possible that a specific site within the coiled-coil domain (such as a phosphorylation or ubiquitination site [94]), and not the structure itself, is responsible for AMOTL2's antiviral phenotype. This question could be pursued by making AMOTL2 expression constructs with point mutations to putative phosphorylation and ubiquitination sites located within the coiled-coil domain and examining STAT1 expression and ZIKV infection in the cells.

Figure 3-3. Major domains of AMOTL2.
Black boxes represent LPTY/PPxY domains. The green box represents a phosphorylation site at serine 159. The light blue region is a coiled-coil domain. The dark green region at the C-terminus is a PDZ-binding motif. From Wigerius, M., Quinn, D., & Fawcett, J. P. (2020). Emerging roles for angiomotin in the nervous system. *Science signaling*, *13*(655), eabc0635. https://doi.org/10.1126/scisignal.abc0635. Reprinted with permission from AAAS.

3.7 Conclusion

From the results of a large CRISPR screen, I discovered that AMOTL2 plays an important role in the potency of IFN-mediated restriction of ZIKV. This work improves our understanding of the human innate immune response against ZIKV and raises new, interesting questions about the possible role of AMOTL2 in restriction of other viruses. Given the ongoing threat posed by ZIKV's predicted re-emergence and ability to cause severe disease, there remain many unknowns regarding the immune response to ZIKV and ZIKV pathogenesis, and we have no approved vaccine nor specific treatment for the virus. Host genes involved in the type I IFN response have been shaped over millions of years of evolution in the host-virus arms race. These host factors therefore represent an attractive target for better understanding the basic biology of host-pathogen interactions and may help us to leverage pathways used by the host to limit viral replication. Ultimately, this knowledge could inform the development of new therapeutics.

Chapter 4 (Appendix A): Derivation of an HIV Risk Score for African Women Who Engage in Sex Work

This appendix describes work that is independent from the main topic of this thesis. A version of this work is published in *AIDS and Behavior*:

Willcox, A. C., Richardson, B. A., Shafi, J., Kabare, E., Kinuthia, J., Jaoko, W., Mandaliya, K., Overbaugh, J., & McClelland, R. S. (2021). Derivation of an HIV Risk Score for African Women Who Engage in Sex Work. *AIDS and Behavior*, 25(10), 3292–3302. https://doi.org/10.1007/s10461-021-03235-7

4.1 Abstract

No tool exists to stratify HIV risk in contemporary African female sex worker (FSW) populations. Data from a cohort of HIV-negative FSWs in Mombasa, Kenya from 2010 to 2017 were used to conduct a survival analysis assessing predictors of HIV infection. Stepwise regression was used to construct a multivariable model that formed the basis for the score. Seventeen HIV infections occurred over 1247 person-years of follow-up contributed by 670 women. Using depot medroxyprogesterone acetate (DMPA), having a curable sexually transmitted infection (STI), and being married contributed points to the score. HIV incidence was 0.85/100 person-years in a lower-risk group and 3.10/100 person-years in a higher-risk group. In a cohort with overall HIV incidence < 1.50/100 person-years, this risk score identified a subgroup of FSWs with HIV incidence > 3.00/100 person-years, which is the threshold used by the World Health Organization for initiating pre-exposure prophylaxis (PrEP). If validated in an external population, this tool could be useful for targeted PrEP promotion among higher-risk FSWs.

4.2 Introduction

African cisgender women who engage in sex work (female sex workers; FSWs) have been disproportionately burdened by HIV since the epidemic began [134]. Antiretroviral therapy (ART) has reduced mortality among HIV-infected individuals, but slower progress has been made in limiting new infections: AIDS-related mortality and HIV incidence declined by 39% and 23% from 2010 to 2019, respectively [135]. Major gaps remain in delivery of HIV prevention services to FSWs [136], who are 13 times more likely to acquire HIV than other adult women [137].

As pre-exposure prophylaxis (PrEP) availability expands in Africa, several studies have reported low PrEP uptake, adherence, and retention among FSWs, though results vary in different contexts [138-140]. Qualitative data have revealed several barriers that hinder PrEP use among African FSWs, including lack of knowledge about PrEP, lack of confidence about its efficacy, stigma associated with taking an HIV drug, concerns about side effects, and difficulty adhering to daily medication [141-144]. Importantly, risk perception may improve PrEP uptake and adherence. Studies in several populations have shown that perceived risk of acquiring HIV is associated with self-reported PrEP adherence [139], and PrEP users align periods of greater adherence with engagement in risky behaviors [145,146]. These data suggest that improved risk assessment combined with risk-appropriate counseling could improve PrEP use in at-risk individuals.

Among African countries, Kenya, Côte d'Ivoire, Eritrea, Nigeria, South Africa and South Sudan recommend offering PrEP for HIV prevention to all FSWs [147]. By contrast, the World Health Organization (WHO) encourages use of an incidence-based system, recommending that PrEP be offered in populations with an HIV incidence at or above three per 100 person-years

[148]. With incidence rates on the decline, FSW populations no longer uniformly qualify for PrEP based on the WHO threshold [149-154]. Given this background, it may no longer be useful to group all FSWs together as uniformly high-risk.

Risk scores are useful for identifying individuals at increased risk of acquiring HIV. These scores have been developed for adults in the United States [155], men who have sex with men in the United States [156-158] and Kenya [159], heterosexual serodiscordant couples in Africa [160], adults in Uganda [161], pregnant and postpartum women in Kenya [162], and African women [163]. Given the increasing number of FSW populations with an HIV incidence below the WHO threshold for PrEP initiation [149-154], the need to focus PrEP resources on those at greatest risk, and the possibility that a more accurate perception of increased risk could improve PrEP use, development of a risk score for African FSWs holds potential for informing PrEP delivery for this key population. The analyses presented in this study utilize data from the Mombasa Cohort, an open cohort study of FSWs in Mombasa, Kenya, to derive an HIV risk score algorithm for this population. Once externally validated, this tool may be helpful for identifying women with elevated risk in other FSW populations with overall incidences below the WHO threshold of three per 100 person-years.

4.3 Results

A total of 670 women contributed 1,757 risk intervals in this study. A median of two (interquartile range [IQR] 1-4) risk intervals were contributed per woman. Seventeen HIV infections occurred over 1,247 person-years of follow-up, for an incidence of 1.36 per 100 person-years (95% CI 0.79-2.18). For the survival analysis, 33 visits with missing data were excluded. A complete case analysis was performed using the 1,724 risk intervals, contributed by 660 women, which remained after exclusion of the visits with incomplete data.

Baseline characteristics of the 660 women are presented in Table 4-1. Their median age was 31 (IQR 26-37) years. Only nine (1.4%) reported that they were married. Of the 547 (82.9%) who reported sexual activity in the past week, 390 (71.3%) reported 100% condom use. Among those sexually active in the past week, the median frequency of sex was three (IQR 2-5), and the median number of partners was two (IQR 1-4). The most common form of modern non-barrier contraception was DMPA with 110 (16.7%) women reporting use. Sexually transmitted infections (STIs) were common, with 68 (10.3%) women having one or more of the four curable STIs (*N. gonorrhoeae, C. trachomatis, T. vaginalis,* and syphilis), and 417 (63.2%) women with HSV-2 seropositivity.

Variable	**Median (IQR) or number (%)***
Age	31 (26, 37)
Education (years)	8 (7, 12)
Married	9 (1.4%)
Alcohol use	523 (79.2%)
Sexually active in past week	547 (82.9%)
100% condom use in past week†	390 (71.3%)
Sex acts in past week†	3 (2, 5)
Sex partners in past week†	2 (1, 4)
Contraceptive method	
None/condoms only	425 (64.4%)
DMPA	110 (16.7%)
Implant	61 (9.2%)
Intrauterine contraceptive device	16 (2.4%)
Tubal ligation	11 (1.7%)
Other	37 (5.6%)
Vulvitis	107 (16.2%)
Genital ulcer disease	4 (0.6%)
Cervical mucopus	16 (2.4%)
Any curable STI‡	68 (10.3%)
Chlamydia trachomatis	36 (5.5%)
Neisseria gonorrhoeae	20 (3.0%)
Trichomonas vaginalis	14 (2.1%)
Syphilis	5 (0.8%)
HSV-2 seropositive	
Yes	417 (63.2%)
Unknown	134 (20.3%)

Table 4-1: Baseline characteristics of 660 HIV-seronegative women in Mombasa Cohort, 2010-2017

STI, sexually transmitted infection; DMPA, depot medroxyprogesterone acetate; HSV, herpes simplex virus.
* For women who contributed more than one risk interval, only data from the first visit are included here.
† Calculated among women who were sexually active in the past week.
‡ Any one or more of *Chlamydia trachomatis, Neisseria gonorrhoeae, Trichomonas vaginalis,* or syphilis.

Of the variables assessed for an association with HIV infection risk, seven were included in the forward stepwise model building based on having a univariable HR p-value less than 0.15 (Table 4-2). These included DMPA use (HR 3.51, 95% CI 1.33-9.28, Wald chi-square 6.43, p=0.01), any curable STI (HR 3.81, 95% CI 1.26-11.54, Wald chi-square 5.58, p=0.02), marital status (HR 6.32, 95% CI 0.78-50.99, Wald chi-square 2.99, p=0.08), genital ulcer disease (HR 5.81, 95% CI 0.78-43.29, Wald chi-square 2.95, p=0.09), alcohol use (HR 0.43, 95% CI 0.16-1.18, Wald chi-square 2.68, p=0.10), vulvitis (HR 2.37, 95% CI 0.83-6.74, Wald chi-square 2.60, p=0.11), and cervical mucopus (HR 4.82, 95% CI 0.66-35.17, Wald chi-square 2.41, p=0.12) (Table 4-2). The final multivariable model used to derive point values for the risk score retained DMPA (adjusted HR [aHR] 3.81, 95% CI 1.42-10.19, Wald chi-square 7.08, p=0.01), any curable STI (aHR 3.99, 95% CI 1.32-12.05, Wald chi-square 6.04, p=0.01), and marital status (aHR 11.44, 95% CI 1.31-99.90, Wald chi-square 4.86, p=0.03). All HRs were calculated using a GEE model that accounted for multiple measurements of risk factors in cases where women contributed more than one risk interval to the analysis.

A risk score was calculated for each woman at the start of each risk interval by summing the risk score point values shown in Table 4-2. Of the 1,724 risk intervals evaluated, 1,312 (76.1%) had a risk score of zero, 374 (21.7%) had a risk score of one, and 38 (2.2%) had a risk score of two. Higher scores were possible, but were not observed, as no woman who reported being married also had a curable STI or was using DMPA. The HIV incidence per 100 person-years was 0.85 (95% CI 0.37-1.68) in women with a risk score of zero, 1.91 (95% CI 0.62-4.46)

in women with a risk score of one, and 13.76 (95% CI 3.70-35.21) in women with a risk score off two (Table 4-3).

The ROC curve of the risk score in this dataset had an AUC of 0.67 (95% CI 0.52-0.82; Figure 4-1). This is higher than the AUCs of any of the individual predictors: 0.63 (95% CI 0.48-0.77) for DMPA, 0.58 (95% CI 0.43-0.73) for STIs, and 0.53 (95% CI 0.38-0.67) for marriage. To simplify use of the score for identifying a group of FSWs who would qualify for PrEP under the WHO recommendations, a conservative binary cut-point was chosen to classify women into either a "lower-risk" (score = 0) or "high-risk" group (score ≥ 1). The binary cut-point corresponds to the blue circle on the ROC curve in Figure 4-1 and has a sensitivity and specificity for HIV acquisition of 53% and 76%, respectively, and a PPV and NPV of 2.2% and 99.4%, respectively. The lower-risk group had an HIV incidence of 0.85 per 100 person-years, while the high-risk group had an incidence of 3.10 per 100 person-years.

Variable	Group	Person-years of follow-up*	HIV infections	Incidence† (95% CI)	HR (95% CI), Wald test (1 df)	Univariable p-value	Multivariable regression coefficient	Risk score points
Demographic and behavioral								
Age‡	<25	104	2	1.92 (0.22, 6.92)	1.00			
	≥25	1123	15	1.34 (0.75, 2.20)	0.71 (0.16, 3.15), 0.20	0.66		
Years of education§‖	≤8	703	9	1.28 (0.58, 2.43)	1.00			
	>8	524	8	1.53 (0.66, 3.01)	1.16 (0.45, 3.01), 0.10	0.76		
Years in sex work‖	≤8	608	10	1.65 (0.79, 3.03)	1.00			
	>8	619	7	1.13 (0.45, 2.33)	0.72 (0.28, 1.9), 0.43	0.51		
Workplace¶	Bar#	580	9	1.55 (0.71, 2.95)	1.00			
	Other**	647	8	1.24 (0.53, 2.44)	0.79 (0.3, 2.04), 0.24	0.62		
Married††	No	1215	16	1.32 (0.75, 2.14)	1.00			0
	Yes	12	1	8.01 (0.10, 44.57)	6.32 (0.78, 50.99), 2.99	0.08	2.44	2
Number of live births‖	≤2	926	12	1.30 (0.67, 2.26)	1.00			
	>2	301	5	1.66 (0.54, 3.88)	1.32 (0.47, 3.76), 0.28	0.60		
Alcohol use††	No	231	6	2.60 (0.95, 5.66)	1.00			
	Yes	996	11	1.10 (0.55, 1.98)	0.43 (0.16, 1.18), 2.68	0.10		
Vaginal washing§§	None	527	6	1.14 (0.42, 2.48)	1.00			
	Any‡‡	700	11	1.57 (0.78, 2.81)	1.34 (0.50, 3.62), 0.34	0.56		
Sexual activity and contraception								
Condomless sex§§	No	957	11	1.15 (0.57, 2.06)	1.00			
	Yes	270	6	2.22 (0.81, 4.84)	1.92 (0.72, 5.17), 1.68	0.20		
Sexual partners‖§§	≤1	664	9	1.36 (0.62, 2.57)	1.00			
	>1	563	8	1.42 (0.61, 2.80)	1.01 (0.39, 2.63), 0.00	0.98		
Sexual acts‖§§	≤2	754	12	1.59 (0.82, 2.78)	1.00			
	>2	473	5	1.06 (0.34, 2.47)	0.64 (0.23, 1.82), 0.70	0.40		
Contraception ††	None/ other	1022	10	0.98 (0.47, 1.80)	1.00			0
	DMPA	205	7	3.42 (1.37, 7.04)	3.51 (1.33, 9.28), 6.43	0.01	1.34	1
Clinical diagnoses								
Vulvitis††	No	1039	12	1.15 (0.60, 2.02)	1.00			
	Yes	188	5	2.66 (0.86, 6.20)	2.37 (0.83, 6.74), 2.60	0.11		
Genital ulcer disease††	No	1214	16	1.32 (0.75, 2.14)	1.00			
	Yes	13	1	7.44 (0.10, 41.39)	5.81 (0.78, 43.29), 2.95	0.09		
Cervical mucopus††	No	1211	16	1.32 (0.75, 2.15)	1.00			
	Yes	16	1	6.22 (0.08, 34.60)	4.82 (0.66, 35.17), 2.41	0.12		
Laboratory diagnoses								
Bacterial vaginosis	Nugent <7	876	11	1.26 (0.63, 2.25)	1.00			
	Nugent ≥7	351	6	1.71 (0.62, 3.72)	1.36 (0.50, 3.68), 0.37	0.54		
Vaginal candidiasis	No	1087	17	1.56 (0.91, 2.50)	1.00			
	Yes	140	0	0.00 (0.00, 2.62)	—‖‖			
Endocervical PMNs	<30	1196	17	1.42 (0.83, 2.28)	1.00			
	≥30	31	0	0.00 (0.00, 11.70)	—‖‖			
Any curable STI††	No	1137	13	1.14 (0.61, 1.96)	1.00			0
	Yes	90	4	4.44 (1.20, 11.37)	3.81 (1.26, 11.54), 5.58	0.02	1.38	1
HSV-2	No	127	1	0.79 (0.01, 4.38)	1.00			
	Yes/unknown	1100	16	1.45 (0.83, 2.36)	1.86 (0.25, 13.92), 0.36	0.55		

Table 4-2: Univariable and multivariable analysis of potential risk factors for HIV acquisition and their contributions to the HIV risk score.

HIV, human immunodeficiency virus; HR, hazard ratio; CI, confidence interval; DMPA, depot medroxyprogesterone acetate; PMNs, polymorphonuclear cells; STI, sexually transmitted infection; HSV-2, herpes simplex virus type 2
* Person-years are rounded to nearest integer

† Incidence per 100 person-years
‡ Dichotomized at age 25, as in previous studies [163-166]
§ Dichotomized at break-point between primary and secondary school
‖ Dichotomized at median value
¶ Dichotomized based on previous findings [167]
# Includes bar, restaurant, and guesthouse
** Includes nightclub, home-based sex work, and other
†† Variables assessed for inclusion in multivariable model based on univariable p-value < 0.15
‡‡ Includes washing with water only, based on previous findings [168]
§§ In last week
‖‖ Model did not converge

Score	Risk intervals*	Person-years of follow-up	HIV infections	Incidence per 100 person-years (95% CI)	
0	1312	936	8	0.85	(0.37-1.68)
1	374	262	5	1.91	(0.62-4.46)
2**	38	29	4	13.76	(3.70-35.21)
Total	**1724**	**1227**	**17**	**1.39**	**(0.81-2.22)**

Table 4-3: HIV incidence by risk score
HIV, human immunodeficiency virus; CI, confidence interval
*Each participant could contribute more than one follow-up period (risk interval).
**The group of women with a score of 2 due to concomitant DMPA use and STI positivity totaled 22 risk intervals, 17 person-years, and 3 infections. The group of women with a score of 2 due to being married totaled 16 risk intervals, 12 person-years, and 1 infection.

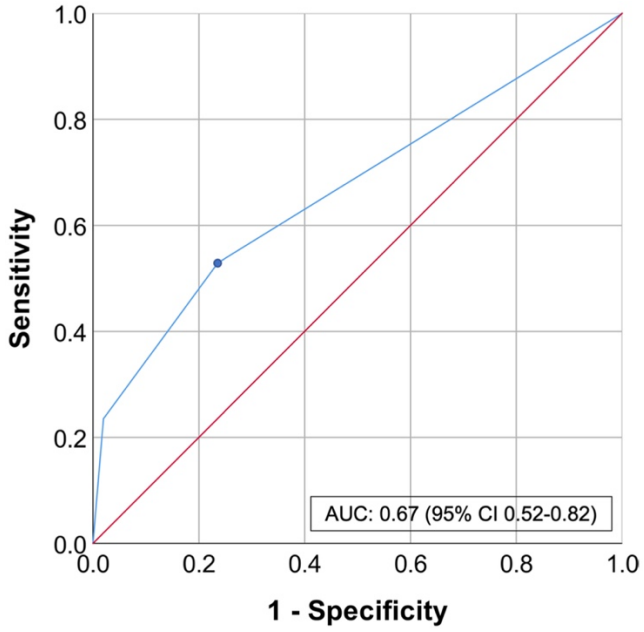

Figure 4-1. Receiver operating characteristic curve of the risk score.
ROC, receiver operating characteristic; AUC, area under the curve; CI, confidence interval. ROC curve showing sensitivity and specificity for the two possible cut-points of the risk score. The blue circle represents the cut-point chosen for a binary risk score

4.4 Discussion

In this study, a simple risk score stratified FSWs in Mombasa into a lower-risk or high-risk group based on the presence of one or more risk factors including using DMPA, being married, and having a curable STI (*N. gonorrhoeae, C. trachomatis, T. vaginalis,* or syphilis). If this tool is externally validated, it could provide a useful means of stratifying risk in populations of FSWs where the overall risk of HIV infection is less than 3% per year, allowing targeted PrEP counseling and delivery for those women most likely to benefit.

In the binary risk categories created by this risk score, the high-risk group included women with a score of one or two, with the latter category comprising a small group of women with a very high incidence of 13.76 per 100 person-years. Women in this group were either married or had both DMPA use and STI positivity as risk factors. Though a conservative approach was applied here to consider all women with a score greater than zero as high-risk, the highest-risk subcategory of women could also be considered separately, especially as overall incidence rates continue to decline.

This risk score is unique in targeting African FSWs, a particularly vulnerable population. Notably, the risk factors described here differ from those used in risk scores for other populations of African women. For example, in scores developed for general population African women [161,163], risk points are added for younger age, more risky sexual practices (e.g., sexual concurrency, having a primary partner with other partners), and being unmarried. In contrast, in this cohort of FSWs, reporting being married, though rare, was an important contributor to a higher risk score. Age and sexual practices did not contribute to the score. These important differences illustrate the necessity of developing an HIV risk score adapted to the unique risk context of women who report engaging in transactional sex. A 2014 study estimated that 4% of

adult women in Kenya may engage in transactional sex [169], indicating that this risk score could be applicable to a substantial proportion of the population.

While the link between marriage and higher HIV risk may initially seem surprising, multiple studies have established that FSWs use condoms less frequently with partners who are long-term, intimate, cohabitating, or non-paying [170-174]. In a qualitative study of FSWs in the Mombasa Cohort, an important emerging theme was that women perceived more power in negotiating condom use with short-term clients [175]. Some women even reported avoiding long-term sexual relationships because of concern that they would not be able to insist on condom use. Several studies have also shown that HIV prevalence is higher among FSWs' boyfriends than their clients [176,177]. These data underscore that while emphasis is often placed on the importance of consistent condom use with clients, non-paying partners may contribute more to incident HIV infections in FSWs. It is also possible that women who report both transactional sex and being married are a subgroup experiencing more extreme financial pressure or are simply willing to take more risks.

Current use of DMPA was the most statistically significant predictor of infection contributing to the risk score in this FSW population. There has been a consistent association between DMPA use and HIV risk in the Mombasa Cohort for over 20 years [167,178,179]. In addition, numerous studies have examined the association between DMPA use and HIV acquisition in other populations. Many, though not all, have observed an association similar to that seen in the Mombasa Cohort. A 2016 systematic review and meta-analysis of studies examining this association reported an overall increased risk of HIV among women using DMPA, but noted that results of the individual studies varied widely [180]. Recently, the only study that randomly assigned women at risk for HIV to different contraceptive methods found no

statistically significant increased risk associated with DMPA use compared to the copper IUD or the levonorgestrel implant [181]. Taken together, these data indicate that regardless of whether DMPA has a biological effect on HIV susceptibility, it remains a fairly consistent marker of women's risk of HIV acquisition.

Having any of the most common curable STIs (*N. gonorrhoeae, C. trachomatis, T. vaginalis,* and syphilis) predicted HIV infection and contributed to the risk score developed in this analysis. It is well established that STIs are associated with higher risk of both acquiring and transmitting HIV [182-185]. Global STI incidence remains extremely high, with an estimated 376 million new infections of these four common STIs every year [186]. This evidence makes a strong case for addressing STIs and HIV concurrently and lends support to the benefits of performing regular and sensitive STI testing in populations at risk for HIV.

This study had several strengths. The Mombasa Cohort is a large, well-characterized cohort with data available to examine many variables as possible contributors to a risk score. The overall incidence of HIV in the cohort is well below the WHO threshold of three per 100 person-years for PrEP initiation, so this group of women is representative of the type of FSW population that might benefit from a targeted approach for PrEP delivery. Despite the low overall incidence, the risk score identified a subgroup of women with an incidence above three per 100 person-years. The score also has acceptable predictive value, given that its AUC of 0.67 falls in the range of AUCs of other published HIV risk scores (0.66 to 0.85) [155-161,163]. Another strength was the use of NAATs to diagnose *N. gonorrhoeae* and *C. trachomatis*, as these methods are considerably more sensitive than classical microbiological approaches [187].

Several limitations of this study are important to consider. Stepwise regression has several drawbacks, including the production of biased regression coefficients and the potential

for the model to perform poorly when tested out-of-sample [188]. This method of model selection was used for consistency with the methods of many other risk scores, including most published HIV risk scores [156,158-163]. In addition, variables in the final model were carefully reviewed by the investigators and were noted to be behaviorally and biologically plausible risk factors for HIV acquisition that align closely with known risk factors for FSWs. Given only seventeen HIV infections, the power of the analysis to identify correlates of HIV infection was limited. In addition, several demographic variables including marriage were collected only at enrollment, so there may have been misclassification of these variables at entry into the risk intervals. In general, this non-differential misclassification would be most likely to attenuate the observed associations with HIV acquisition. An additional limitation related to the marriage data was the very low proportion of women who reported that they were married, resulting in the high incidence among married women being driven by a single HIV infection and a small number of person-years. This resulted in large confidence intervals and casts some uncertainty over the utility of marriage as a risk score variable for other populations of FSWs. Nonetheless, given the data supporting an association between regular partnerships and riskier sexual behaviors in FSWs [170-174], it seems plausible that marriage may be a useful variable in the risk score. Finally, the small number of infections precluded internal validation of the score by bootstrap resampling analysis or splitting the dataset into a derivation and validation subset. As such, there remains a need to validate this risk score in an independent cohort.

The impetus leading to development of this risk score was the fact that most FSWs in the Mombasa Cohort decline PrEP. The goal was to identify women who would most benefit from PrEP so that effort can be focused on those women in whom it will have the greatest impact on transmission. In this analysis, the low- and high-risk groups had 8 and 9 infections over 936 and

291 years of follow-up, respectively. This suggests that a similar number of infections could be averted in a smaller number of person-years by focusing on women with particular risk factors. The fact that many groups of FSWs now fall below the 3% annual HIV incidence threshold suggested by WHO does not necessarily mean that PrEP should not be offered to "lower-risk" FSWs. Indeed, there may be an added prevention benefit at the population level, given their likelihood of transmitting HIV to others if they do become infected. Rather, the risk stratification system developed here may allow programs to decide how best to allocate time, energy, and resources for promoting and delivering PrEP in FSW populations with an overall HIV incidence below 3%.

In conclusion, this study presents the derivation of a risk score that could be applied to identify those FSWs who would receive the greatest benefit from PrEP. The next step is to validate this tool either with prospective data from the Mombasa Cohort or in other populations. In FSW populations where overall HIV incidence is substantially below 3% annually, identifying and targeting individuals with a predicted risk ≥ 3% may offer important HIV prevention benefits at a population level and would be a more cost-effective way of delivering PrEP [189-191]. In addition, this risk score may provide a framework to guide individualized counseling with FSWs to support PrEP initiation, adherence, and retention.

4.5 Materials and Methods

Population and procedures

Since 1993, the Mombasa Cohort has enrolled women who report exchanging sex for cash or in-kind payment in a long-term open cohort study of risk factors for HIV acquisition. Women are continuously recruited through outreach activities in local bars, nightclubs, and brothels. At enrollment, women provide extensive demographic and behavioral data in a

standardized interview with a study nurse. Women are invited to return for monthly follow-up visits during which they are interviewed on recent sexual and medical history, have blood drawn for laboratory testing, and undergo a pelvic examination that includes collection of specimens for laboratory diagnosis of genital tract infections. Women presenting with symptoms of STIs are offered free syndromic management during the visit. Asymptomatic laboratory-confirmed infections are treated using pathogen-directed therapy when women return for test results 7 days later. HIV incidence in the cohort has declined dramatically over time, from 17.1 per 100 person-years in 1993 to below 2 per 100 person-years since 2010 [149,192]. Further details of the cohort have been described previously [149,193,194].

For this study, data from January 2010 through June 2017 were analyzed. Eligible participants included women who had at least one clinic visit on or after January 1, 2010, at which they tested negative for HIV and had STI testing performed for gonorrhea, chlamydia, trichomoniasis, and syphilis (described in "Laboratory methods" below). In addition, participants eligible for this cohort analysis had to return for at least one follow-up visit with HIV testing within one year of their baseline visit. All women provided written informed consent to participate in the research. This study was approved by the Institutional Review Boards at Kenyatta National Hospital and the University of Washington.

Laboratory methods

HIV serostatus was determined by ELISA at every visit. The Detect HIV 1-2 test (Biochem Immunosystems, Montreal, Canada) was used through January 2010, and the Pishtaz HIV 1,2 ELISA (Pishtaz Teb Diagnostics, Tehran, Iran) was used from February 2010 through April 16, 2014. Vironostika HIV-1 Uni-Form II Ag/Ab (bioMerieux, Marcy I'Etoile, France) was used as a confirmatory test from January 2010 through April 16, 2014. On April 17, 2014,

all tests were switched to fourth generation HIV 1/2 rapid testing platforms, with the Determine Assay (Alere International Ltd, Galway, Ireland) as the initial test and the Unigold assay (Trinity Biotech, Bray, Ireland) as the confirmatory test. Confirmatory tests were only performed on individuals with a positive initial test.

For women who seroconverted, HIV viral load was measured retrospectively in stored samples (collected at every previous visit) to determine the date of the first HIV RNA-positive sample using the Hologic/Gen-Probe quantitative HIV-1 assay (Aptima, Hologic/Gen Probe, San Diego, CA) to more closely define the time of infection, as described previously [195]. Plasma samples from each HIV-seronegative visit were stored at -80°C prior to testing.

Neisseria gonorrhoeae and *Chlamydia trachomatis* were detected by nucleic acid amplification testing (NAAT; Aptima, Hologic/Gen Probe, San Diego, CA). *Trichomonas vaginalis* and vaginal candidiasis were detected by microscopic examination of vaginal wet mounts. Syphilis was defined as a positive rapid plasma reagin (RPR) blood test with confirmation based on a positive *Treponema pallidum* hemagglutination (TPHA) test. Herpes simplex virus type 2 (HSV-2) infections were detected by ELISA (HerpeSelect 2; Focus Diagnostics, Cypress, California, USA), with an OD value of 2.1 or above considered positive [194]. Bacterial vaginosis (BV) was detected by microscopic examination of a Gram stain and defined based on the criteria of Nugent and Hillier [196]. Cervicitis was defined as an average of 30 or more polymorphonuclear leukocytes per high-power field of Gram-stained cervical secretions.

Statistical analyses

The approach to analysis was guided with the explicit goal of creating a risk score. Many variables were tested that may serve as good markers of risk, regardless of whether these

variables are the direct cause of HIV acquisition (e.g., workplace, marital status, years in sex work). Such variables are often easier to measure and less subject to recall or social desirability bias than variables more directly affecting HIV acquisition (e.g., condom use). All continuous variables were dichotomized to simplify use of the score. Additionally, since the goal of a risk score is to use a set of variables assessed at baseline to predict risk prospectively, we did not use time-varying exposures (in contrast to many previous studies analyzing data from the Mombasa Cohort).

Variables assessed for an association with HIV acquisition included known risk factors from the literature, published risk scores, previous studies of and the investigators' in-depth knowledge of the Mombasa Cohort, and discussions with other investigators leading FSW cohorts. The following self-reported variables were considered: marital status, alcohol use, condomless sex in the last week, sexual partners in the last week, sexual frequency in the last week, number of live births, age, years of education, workplace (bar/restaurant, nightclub, or other), depot medroxyprogesterone acetate (DMPA) use, vaginal washing in the last week, and years in sex work. The following variables assessed by physical examination or laboratory testing were also considered: vulvitis, genital ulcer disease, cervical mucopus visualized on speculum-assisted examination, laboratory-defined cervicitis, any curable STI (*Neisseria gonorrhoeae, Chlamydia trachomatis, Trichomonas vaginalis,* or syphilis), HSV-2 infection, BV, and vaginal candidiasis. Data for marital status, alcohol use, number of live births, years of education, and workplace were collected only at enrollment in the cohort. Continuous variables were dichotomized at the median value from all follow-up periods (risk intervals) contributed, unless there was a breakpoint that was better supported by published data. Prior to the decision to dichotomize all continuous variables, multiple categorizations of continuous variables were

examined. Changing the number of categories and the break points for categories did not change the association of any variable with HIV acquisition.

A survival analysis was performed to evaluate the association between baseline variables and HIV acquisition during a maximum follow-up of one year. After censoring at one year, women who remained uninfected were able to re-enter this analysis on or after their date of censoring and contribute additional risk intervals.

For women who became infected with HIV and had a positive viral load first measured at or after seroconversion, the date of infection was estimated as the midpoint between the last seronegative and first seropositive visit. Though women were invited to return for monthly follow-up visits, some returned less often. Misestimating the date of HIV infection by several months would be expected to have a minimal effect on the calculated hazard ratios, given the large number of total person-years at risk and the relatively small number of HIV infections for which this misestimation occurs. For women with a detectable viral load prior to seroconversion (i.e., HIV RNA was detected but antibodies were not), indicating a very recent infection, the date of infection was estimated to be 17 days prior to the positive viral load, as described previously [195,197,198].

Follow-up time was censored at the estimated date of HIV infection (for women who became infected) or date of the last HIV test within one year of the baseline visit (for women who remained uninfected), for each risk interval. Incidence rates are reported with Taylor series confidence intervals (CIs).

A generalized estimating equation (GEE) model with a Poisson (log) link, independent correlation structure, and time as an offset variable was used to calculate hazard ratios (HRs) with Wald CIs for exposure variables. This approach accounts for within-participant correlation

in women who contributed more than one risk interval. Variables with univariable HR p-values of less than 0.15 were assessed for inclusion in a multivariable model by manual forward stepwise model-building. Variables were added to the model in order of ascending univariable p-values. Newly added variables were retained in the model if the p-values of existing variables in the model remained below 0.15 [199]. The final model retained only multivariable p-values less than 0.05. A backward stepwise model-building approach resulted in the same model.

The final multivariable model created using GEE formed the basis for the HIV risk score. Points were assigned to individual predictors by dividing each regression coefficient by the smallest regression coefficient in the model and rounding to the nearest integer.

The risk score's performance in the Mombasa Cohort was assessed by calculating HIV incidences for each risk score group and generating a receiver operating characteristic (ROC) curve to assess the score's performance and select an optimal binary cut-point. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the binary score were calculated using units of risk intervals.

Analyses were performed using IBM SPSS 26.0 (IBM, Kirkland, Washington, USA) and OpenEpi [200].

Chapter 5 (Appendix B): Detailed analysis of antibody responses to SARS-CoV-2 vaccination and infection in macaques

This appendix describes work that is independent from the main topic of this thesis. A version of this work is published in *PLOS Pathogens*:

Willcox, A. C., Sung, K., Garrett, M. E., Galloway, J. G., Erasmus, J. H., Logue, J. K., Hawman, D. W., Chu, H. Y., Hasenkrug, K. J., Fuller, D. H., Matsen Iv, F. A., & Overbaugh, J. (2022). Detailed analysis of antibody responses to SARS-CoV-2 vaccination and infection in macaques. *PLOS Pathogens*, *18*(4), e1010155. https://doi.org/10.1371/journal.ppat.1010155

5.1 Abstract

Macaques are a commonly used model for studying immunity to human viruses, including for studies of SARS-CoV-2 infection and vaccination. However, it is unknown whether macaque antibody responses resemble the response in humans. To answer this question, we employed a phage-based deep mutational scanning approach (Phage-DMS) to compare which linear epitopes are targeted on the SARS-CoV-2 Spike protein in convalescent humans, convalescent (re-infected) rhesus macaques, mRNA-vaccinated humans, and repRNA-vaccinated pigtail macaques. We also used Phage-DMS to determine antibody escape pathways within each epitope, enabling a granular comparison of antibody binding specificities at the locus level. Overall, we identified some common epitope targets in both macaques and humans, including in the fusion peptide (FP) and stem helix-heptad repeat 2 (SH-H) regions. Differences between groups included a response to epitopes in the N-terminal domain (NTD) and C-terminal domain (CTD) in vaccinated humans but not vaccinated macaques, as well as recognition of a CTD epitope and epitopes flanking the FP in convalescent macaques but not convalescent humans. There was also considerable variability in the escape pathways among individuals within each group. Sera from convalescent macaques showed the least variability in escape overall and

converged on a common response with vaccinated humans in the SH-H epitope region, suggesting highly similar antibodies were elicited. Collectively, these findings suggest that the antibody response to SARS-CoV-2 in macaques shares many features with humans, but with substantial differences in the recognition of certain epitopes and considerable individual variability in antibody escape profiles, suggesting a diverse repertoire of antibodies that can respond to major epitopes in both humans and macaques. Differences in macaque species and exposure type may also contribute to these findings.

5.2 Introduction

The COVID-19 pandemic has created a pressing need to understand immunity to SARS-CoV-2, both in the setting of vaccination and infection. This has prompted numerous studies in non-human primates (NHPs), which are considered the most relevant animal model for studying many infectious diseases of humans. Various NHP models have been employed to study the immunogenicity and protective efficacy of SARS-CoV-2 vaccine candidates, with most studies using macaque species including rhesus macaques (*Macaca mulatta*) [201-223], cynomolgus macaques (*Macaca fascicularis*) [208,224-232], and pigtail macaques (*Macaca nemestrina*) [222,233-235]. Some of these models have also been used to study infection and re-infection [235-239]. In the NHP model, studies typically measure virus neutralizing antibody responses to vaccination or infection. However, no study has investigated the fine binding specificities of both neutralizing and non-neutralizing SARS-CoV-2 antibodies in macaques and how they compare to the human responses they are meant to model.

Coronaviruses such as SARS-CoV-2 enter host cells using their Spike glycoprotein, which is composed of trimeric S1 and S2 subunits. Receptor-binding S1 homotrimers protrude out from the surface of the virion like a crown, giving this family of viruses its name, while the

fusion-mediating S2 trimers anchor the protein to the viral membrane. On S1, the receptor-binding domain (RBD) of SARS-CoV-2 Spike protein binds to angiotensin-converting enzyme 2 (ACE2) on host cells [240,241]. For subsequent membrane fusion to occur, the Spike protein must be cleaved by host cell proteases at the S1/S2 boundary and at an S2' site located just upstream of the fusion peptide (FP) of S2 [242], leading to substantial conformational changes that likely unmask new epitopes of S2 to immune cells [243].

Antibodies to SARS-CoV-2 Spike protein are especially interesting as a potential correlate of protection, as they have the capacity to block infection and kill infected cells [244-247]. There has understandably been great interest in studying neutralizing antibodies against the RBD, given that such antibodies can directly block interaction with host cells. While RBD-directed antibodies indeed contribute disproportionately to neutralization [248], the majority of the anti-Spike plasma IgG response in convalescent individuals is directed to epitopes outside of the RBD [249,250]. RBD-directed antibodies are also less likely to maintain activity against future viral strains, given the increasing number of variants of concern that harbor mutations in the RBD and have reduced sensitivity to neutralization by immune plasma [251]. Additionally, growing evidence from studies in humans and animal models indicates that non-neutralizing antibodies play a role in protection [252-254].

Previous studies have used Phage-DMS [255], a tool that combines phage display of linear epitopes with deep mutational scanning, to interrogate the fine binding specificities and escape profiles of binding antibodies against all domains of Spike in infected and vaccinated humans [256,257]. These studies have shown that infection-induced human polyclonal antibodies consistently bind linear epitopes in the FP and stem helix-heptad repeat 2 (SH-H) epitope regions, with patient-to-patient variability in escape profiles [256]. Comparatively,

mRNA vaccination induces a broader antibody response across Spike protein with more consistent escape profiles [257].

In this study, we built on this foundation by using Phage-DMS to study the binding and escape profiles of antibodies in repRNA-vaccinated pigtail macaques and convalescent (re-infected) rhesus macaques in comparison to mRNA-vaccinated humans and convalescent humans. Our data reveal broad overlap in some major epitopes targeted by both macaques and humans, though neither vaccinated nor convalescent macaques perfectly model the human response. We also find considerable variability in individuals' antibody escape pathways in most epitope regions in both macaques and humans. The broadest responses were seen in vaccinated humans and re-infected rhesus macaques, groups that also share more concordant escape profiles. These results have implications for the interpretation of COVID-19 macaque research studies.

5.3 Sample details

Four groups were included in this study: vaccinated pigtail macaques, vaccinated humans, convalescent (re-infected) rhesus macaques, and convalescent humans (Table 5-1). The vaccinated macaques received a replicating mRNA (repRNA) vaccine encoding the full-length wildtype (not pre-fusion stabilized) SARS-CoV-2 A.1 lineage Spike protein formulated with a cationic nanocarrier [235,258]. The vaccine was delivered as a prime-only 25μg (n=3) or 250μg (n=6) dose or prime-boost 50μg dose (n=2), with plasma collected 42 days after the first dose (n=9) or 14 days after the second dose (n=2). The vaccinated humans received two doses of the 100μg Moderna mRNA-1273 vaccine encoding the pre-fusion stabilized full-length SARS-CoV-2 A.1 lineage Spike protein and formulated with a lipid nanoparticle. Serum was collected from human vaccinees 36 days after the first dose (7 days after the second dose). The convalescent

macaques were depleted of CD4+ T cells, CD8+ T cells, CD4+ and CD8+ T cells, or neither as part of a previous study. They were infected twice with SARS-CoV-2, with infections spaced six weeks apart and serum collected 56 days after the first infection (14 days after the second infection). The T cell depleted animals were not excluded based on a detailed analysis of the humoral response in these macaques, which suggested that neither CD4+ nor CD8+ T cells were critical for the development of anamnestic antibody responses, neutralizing antibodies, or protection from re-infection [239]. The convalescent humans were naturally infected once with SARS-CoV-2 and exhibited mild disease, with a median of 67 days between symptom onset and sample collection. Details of individual participants are available in Table S1.

Group	**Number of samples**	**Age range (years)**	**Treatment**	**Time of sample collection**
Vaccinated pigtail macaques	11	3 ½ - 6	repRNA vaccine encoding full-length SARS-CoV-2 Spike[a]	42 days post 1st dose (prime-only, n=9) or 14 days post 2nd dose (prime-boost, n=2)
Vaccinated humans	15	18 - 55	100µg mRNA vaccine encoding full-length pre-fusion stabilized SARS-CoV-2 Spike (Moderna)	36 days post 1st dose
Convalescent rhesus macaques	12	2 ½ - 5	Infected twice with SARS-CoV-2 six weeks apart[a]	56 days post 1st infection (14 days post 2nd infection)
Convalescent humans	12	28 - 52	Naturally infected once with SARS-CoV-2 (mild disease)	Median 67 (IQR 62, 70) days post symptom onset

Table 5-1. Details of samples used in the current study.
[a]Within each group of macaques, subgroups received slightly different treatments (described in Table S1).

5.4 Enrichment of wildtype peptides

To compare which regions of Spike protein are recognized by human and macaque antibodies, we examined the enrichment of wildtype peptides by antibodies from each individual

(Fig 5-1A). Broadly speaking, binding was observed in the NTD, CTD, FP, and stem helix-HR2 epitope regions as reported previously in human studies [256,257]. Epitope regions (shown as different colors on Fig 5-1) were defined as previously [257]: NTD, amino acid 285-305; FP, 805-835; stem helix-HR2 (SH-H), 1135-1170. For the CTD, the bounds of epitope regions were expanded and altered from previous studies based on macaque antibodies recognizing a wider area than previously seen in humans: CTD-N', 526-593; CTD-C', 594-685 (S1A Fig). Several additional epitopes that flank previously-defined regions were also identified in this analysis: pre-FP, 777-804; post-FP, 836-855 (S1B Fig); and HR2, 1171-1204 (S1C Fig). Specific epitope regions can be visualized on the structure of a Spike protein monomer in Fig 5-1B. In addition to these defined regions, we noted that one convalescent rhesus macaque appeared to weakly recognize an epitope at the beginning of the S2 subunit (amino acid 686-710, Fig 5-1A).

In general, we did not detect responses in the RBD because many epitopes in this region are known to be conformational, and Phage-DMS only has the power to detect epitopes that include linear sequences. Epitopes in the RBD have been extensively detailed elsewhere [259,260]. However, we did detect strong binding to an RBD epitope in some vaccinated pigtail macaques (Fig 5-1A). This same region was enriched in samples from before vaccination in four of the five pigtail macaques with baseline samples available (S2 Fig). Pre-infection serum from the twelve rhesus macaques did not show such a consistent response, though a few individuals did show strong enrichment of certain peptides (S2 Fig). Because these responses were present prior to vaccination with SARS-CoV-2 Spike or infection with SARS-CoV-2, we did not investigate them further as part of this study. These findings are likely the result of prior exposure to a different coronavirus, as has been documented for human infections.

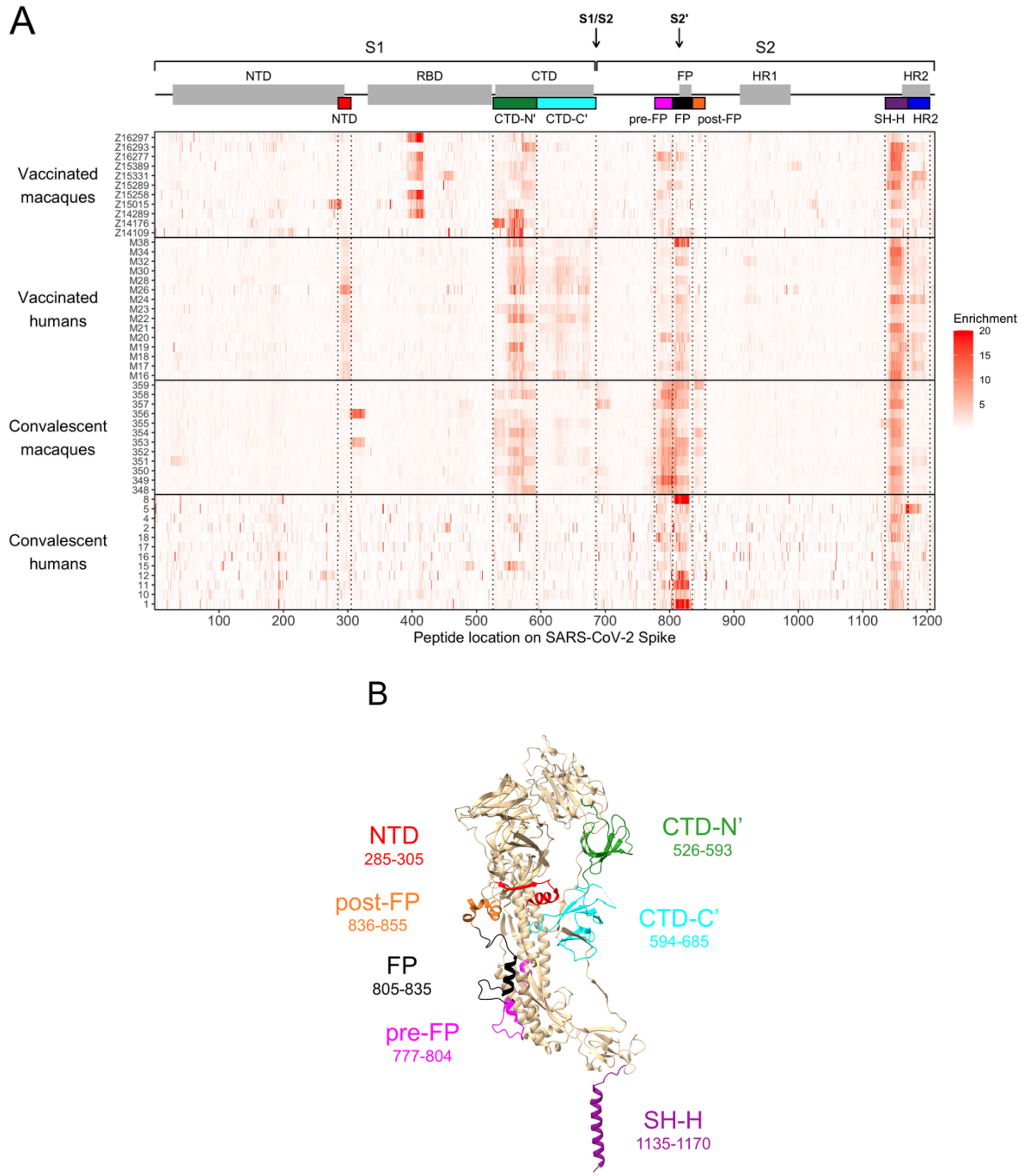

Figure 5-1. Enrichment of wildtype peptides.
(A) The x axis indicates each peptide's location along SARS-CoV-2 Spike protein, and each entry on the y axis is an individual sample. All enrichment values over 20 are plotted as 20 to better show the lower range of the data. Above the heatmap, domains of Spike are shown with grey boxes, with the S1/S2 and S2' cleavage sites indicated with arrows. The epitope regions defined in the current study are shown as colored boxes (from left to right: NTD in red, CTD-N' in green, CTD-C' in cyan, pre-FP in pink, FP in black, post-FP in orange, SH-H in purple, and HR2 in blue). (B) Defined epitope regions shown on a structure of one monomer of SARS-CoV-2 Spike in the pre-fusion conformation (PDB 6XR8). The amino acid loci spanned by each epitope are listed. The HR2 epitope (AA 1171–1204) could not be resolved on the structure and is not shown.

To quantify differences in the epitopes targeted by different groups, the enrichment of wildtype peptides was summed across each epitope region for every individual. Because the main research question is whether responses in macaques model those in humans, two comparisons were performed: vaccinated macaques vs. vaccinated humans and convalescent macaques vs. convalescent humans (Fig 5-2).

In concordance with a qualitative assessment of the enrichment heatmap in Fig 5-1A, vaccinated humans preferentially recognized the following epitope regions compared to vaccinated macaques: NTD (Mann-Whitney $p \leq 0.01$), CTD-C' ($p \leq 0.0001$), and FP ($p \leq 0.05$) (Fig 2A). Meanwhile, convalescent macaques recognized the following epitope regions more than convalescent humans: CTD-N' ($p \leq 0.01$), pre-FP ($p \leq 0.001$), and post-FP ($p \leq 0.01$) (Fig 5-2B). All groups consistently recognized the SH-H epitope region (Fig 5-2). While vaccination appeared to induce a stronger response against HR2 than infection (Fig 5-1A), there were no significant differences in response driven by species (Fig 5-2). Within each group of macaques (vaccinated and convalescent), subgroups received slightly different treatments (Table S1), so similar analyses were performed comparing these subgroups; no comparisons were significant at a threshold of $p=0.05$ (Kruskal-Wallis test, S3 Fig).

Taken together, these findings indicate: 1) vaccinated humans were the only group to consistently recognize peptides from both the NTD and CTD-C' epitope regions, which are in close physical proximity to one another (Fig 5-1B); 2) convalescent humans had a limited response to the CTD-N'; 3) compared to other groups, convalescent macaques had a notably more robust response to regions upstream and downstream of the main FP epitope region; 4) vaccinated macaques did not recognize the FP as strongly as other groups; and 5) vaccination seemed to induce a stronger response against HR2 than infection in both macaques and humans.

These findings may also be explained by other differences between the vaccinated and convalescent groups, including the number, dose, and type of exposures.

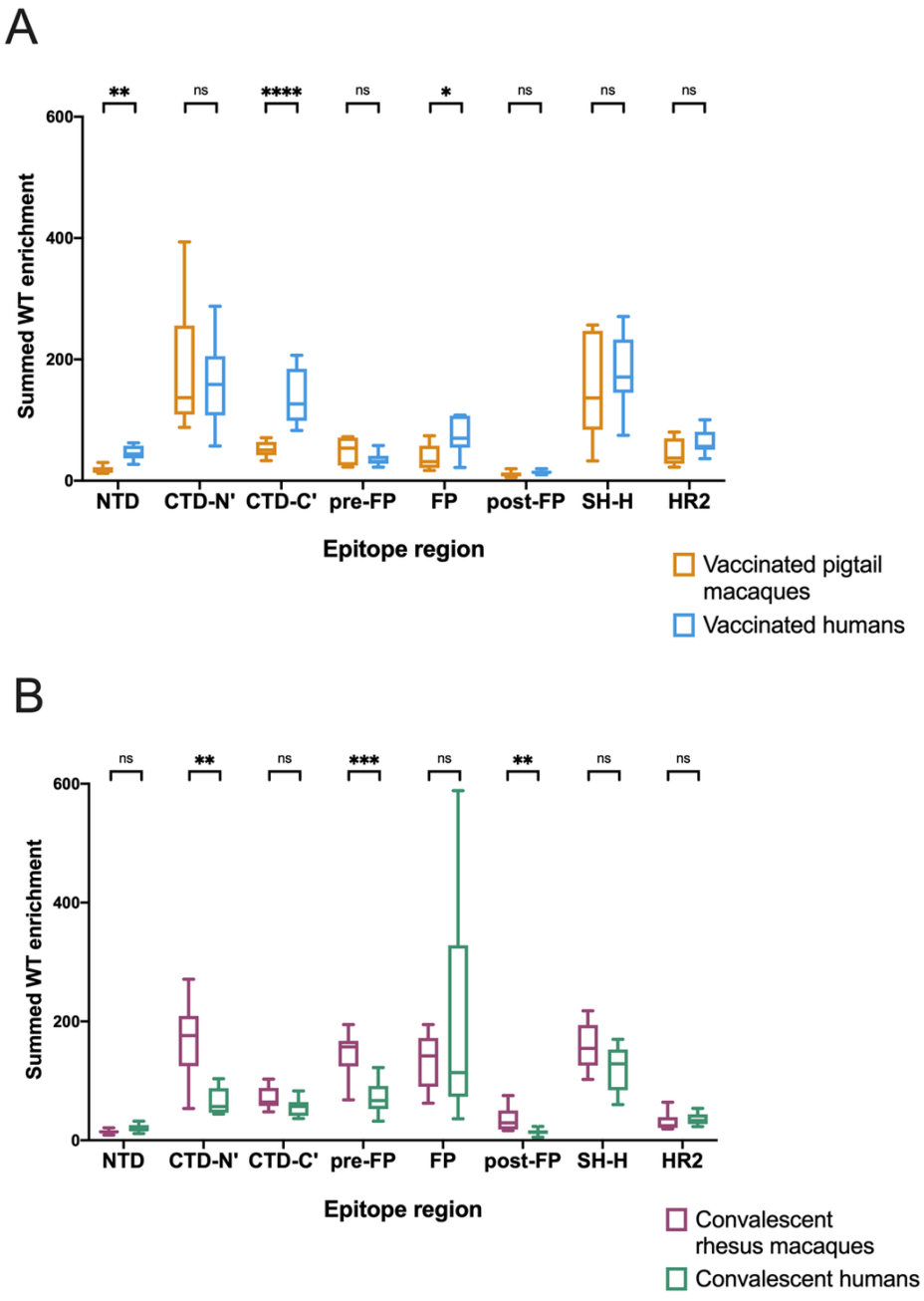

Figure 5-2. Differences in enrichment of wildtype peptides by group.
For each individual, wildtype enrichment values were summed for all peptides within each epitope region. Boxplots summarize the data for all individuals in each sample group. The box represents median and interquartile range (IQR), the lower whisker represents the lowest data point above Q1-1.5IQR, and the upper whisker represents the highest data point below Q3+1.5IQR. (A) compares vaccinated pigtail macaques to vaccinated humans, while (B) compares convalescent rhesus macaques to convalescent humans. Multiple Mann-Whitney U tests were performed, with p values corrected for the number of comparisons in each plot (8) using the Bonferroni-Dunn method. ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$.

5.5 Defining and comparing escape pathways

To assess differences in the binding characteristics of human and macaque antibodies on a more granular level, we next examined the mutations in Spike that reduced antibody binding in each epitope region of interest. Because the antibody escape pathways for vaccinated humans have been described previously [257], we did not examine the NTD and CTD-C', which are exclusively recognized by this group. Instead, we focused on comparing escape profiles between groups in the following epitope regions: CTD-N', FP, and SH-H. We first represent the data as scaled differential selection values in logo plot form, as commonly shown in previous studies. Importantly, scaled differential selection is highly correlated with peptide binding as measured by competition ELISA [255]. To summarize the data represented by the logo plots by group, summed differential selection values across each epitope region were also calculated. This metric represents the overall magnitude of escape at each locus regardless of the specific amino acid substitution, with negative values indicating a decrease in binding compared to the wildtype amino acid, and positive values indicating enhanced binding (see "Materials and Methods"). Finally, escape similarity scores were calculated between pairs of individuals to quantify similarity in escape profiles (see "Materials and Methods" and S4 Fig).

CTD-N'

Vaccinated macaques, vaccinated humans, and convalescent macaques recognized peptides in the CTD-N' (AA 526-593), whereas convalescent humans generally did not (Fig 5-2B). Within this epitope region, the individual escape profiles showed notable variability both within and between groups (S5 Fig). For example, across all groups, some individuals showed relatively high sensitivity to mutations between sites 558-567, while others had a response

focused more downstream around AA 577-586. There was also substantial variability in which loci in the CTD-N' had the highest relative magnitude of escape, and sometimes even in the directionality of scaled differential selection at a given locus. For example, some individuals had antibodies that bound mutated peptides better than wildtype at AA 555 (e.g., convalescent macaque 353) while others exhibited reduced binding to mutated peptides (e.g., convalescent macaque 358). The same was true for site 560 (e.g., vaccinated humans M24 and M26 exhibited improved and disrupted binding to mutated peptides, respectively).

To summarize the trends observed in the individual findings, we calculated summed differential selection values for each individual at each site and generated boxplots by group (Fig 5-3A). In addition to the aforementioned regions of escape common to all groups, convalescent macaques also showed considerable escape between AA 529-535, with vaccinated macaques also showing a less consistent response in this area (Fig 5-3A and S5). The complexity and variability of the escape pathways also prompted us to quantify the similarity in escape between and within groups. Escape similarity scores largely corresponded to areas of high magnitude of escape. Sites with low-magnitude summed differential selection values indicate loci where mutations have no notable impact, and therefore those escape profiles reflect fluctuations in peptide enrichments due to noise, which drives a lower escape similarity score at those sites (Fig 5-3A, lower panel). At some sites (e.g., 560, as described above), low scores were also the result of some samples showing negative differential selection and others showing positive differential selection, a comparison that was assigned the highest cost in our escape similarity score algorithm.

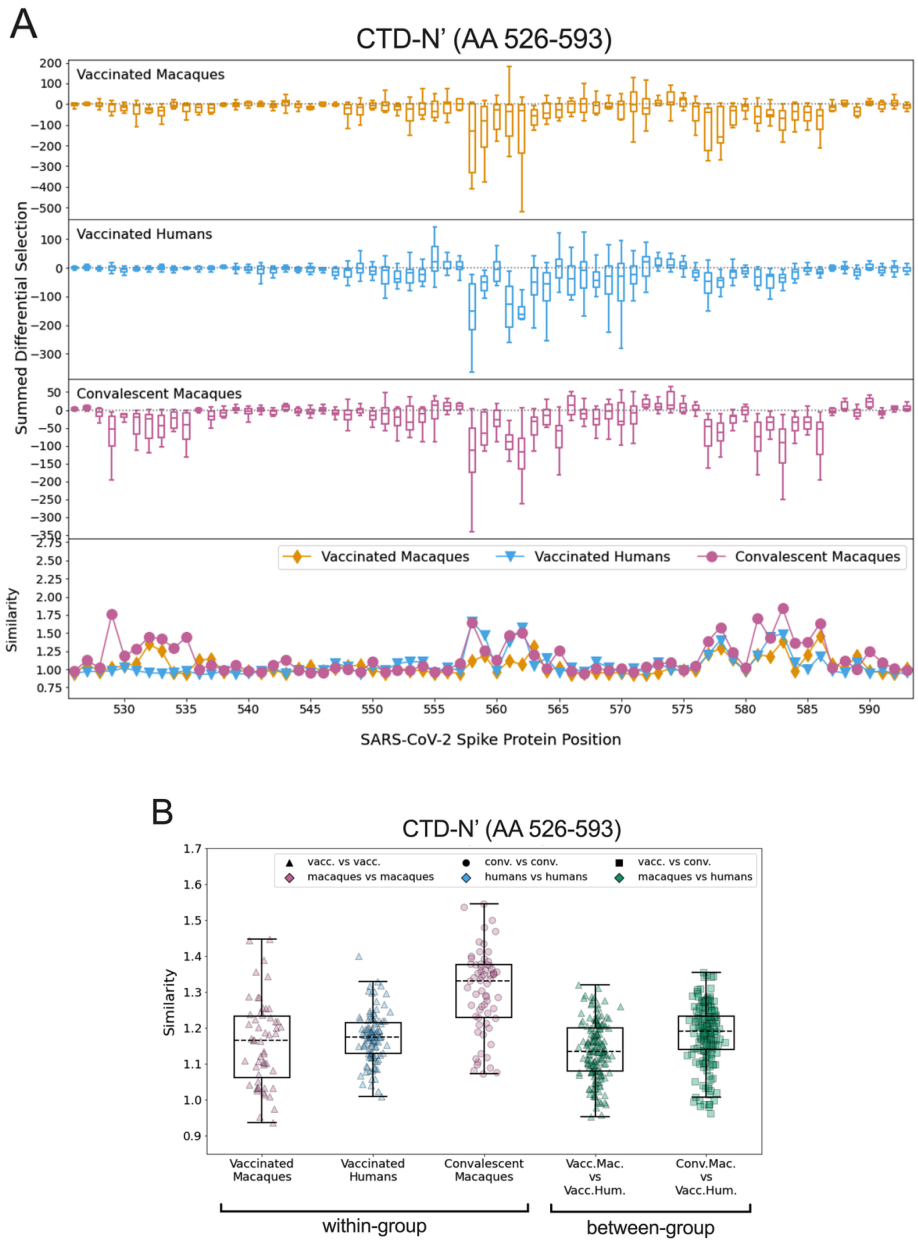

Figure 5-3. Comparison of escape profiles in the CTD-N'.
(A) The top three panels show boxplots depicting the summed differential selection values of all individuals in a group at each locus, with each data point representing a different individual. The box represents median and interquartile range (IQR), the lower whisker represents the lowest data point above Q1-1.5IQR, and the upper whisker represents the highest data point below Q3+1.5IQR. Negative values represent sites where the binding interaction between antibody and peptide was weakened when peptides were mutated, whereas positive values represent enhanced binding. The bottom panel shows the mean escape similarity score for all pairwise comparisons between samples in each group, calculated at every locus. See S4 Fig for a description of the escape similarity score algorithm. (B) Within- and between-group region-wide escape similarity scores, with each point representing a pairwise comparison between two samples. The box represents median and interquartile range (IQR), the lower whisker represents the lowest data point above Q1-1.5IQR, and the upper whisker represents the highest data point below Q3+1.5IQR. The contribution of a site's score to the total escape similarity score is weighted based on its relative contribution to the summed differential selection values across the region. P values are not computed due to lack of independence between data points.

To test the similarity of escape profiles across the CTD-N' epitope region, escape similarity scores were aggregated across the region and computed both within and between groups. These are shown as boxplots, with each point representing a pairwise comparison between individual samples (Fig 5-3B). For example, every vaccinated macaque was compared to every other vaccinated macaque (a within-group comparison) and to every vaccinated human (a between-group comparison). We included a comparison of convalescent macaques and vaccinated humans, given visual similarities between their patterns of escape (Fig 5-3A). Convalescent macaques showed the highest within-group similarity in escape profiles, meaning their escape profiles were more consistent than those of the vaccinated macaques or vaccinated humans (Fig 5-3B). Between-group escape similarity scores were on par with the within-group scores for the vaccinated macaques and humans, indicating that although there was substantial variability in individual profiles, this was not driven by sample groups.

FP

Escape profiles were examined in the FP epitope region (AA 805-835) for the three groups that showed significant wildtype enrichment in this area: vaccinated humans, convalescent macaques, and convalescent humans. As in the CTD-N', overall there was variability in individual escape profiles, though the convalescent macaques showed a more consistent pattern of escape than other groups (S6 Fig). Within the FP, most sites of escape fell between AA 811-825 for all groups (Fig 5-4A). The convalescent macaques again exhibited the highest escape similarity scores (Fig 5-4B). The median within-group escape similarity scores in the FP were on par with those in the CTD-N' (Fig 5-3B), indicating approximately equal variability in antibody escape in these epitope regions. The between-group escape similarity scores were generally similar to each other and to the human within-group scores (Fig 5-4B).

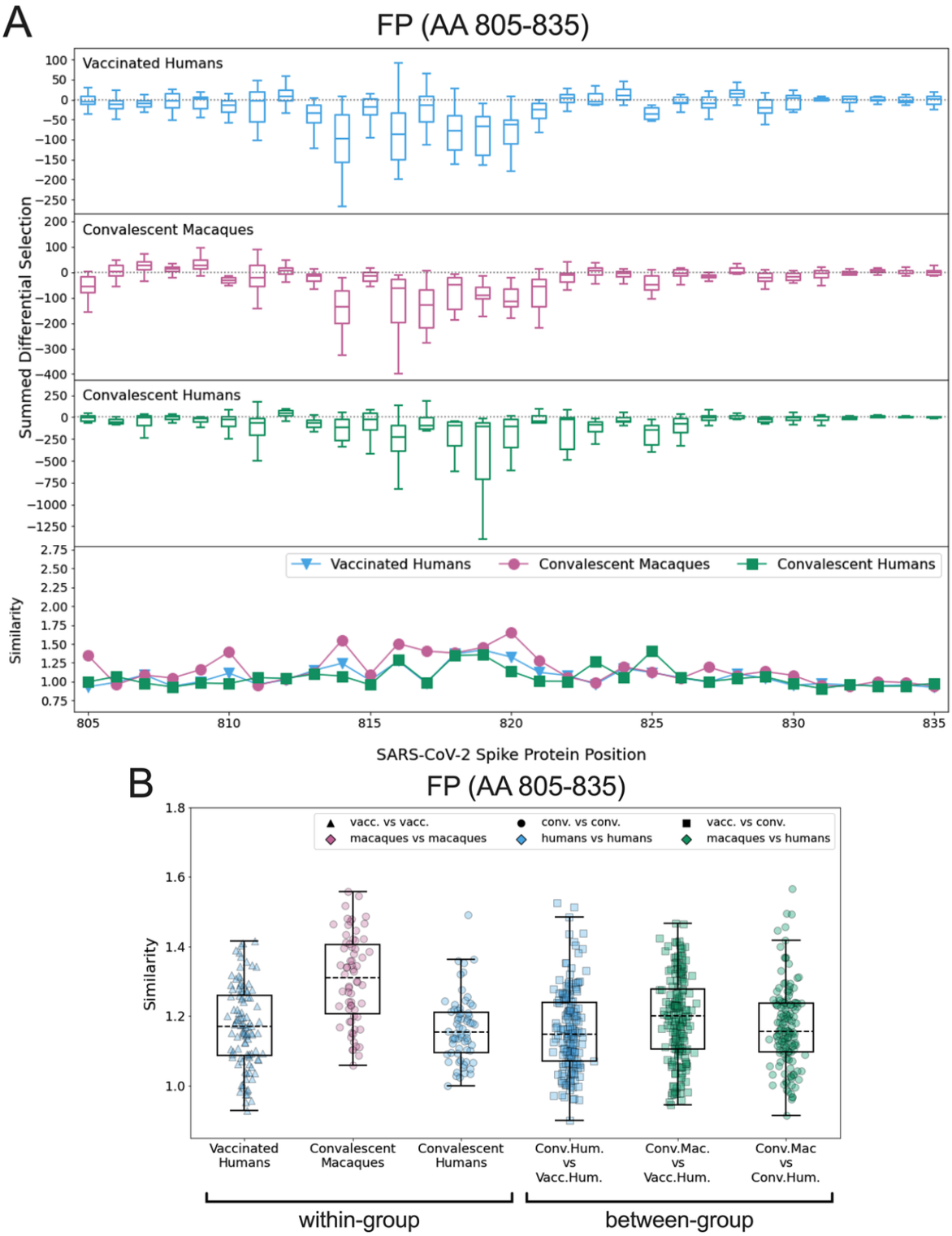

Figure 5-4. Comparison of escape profiles in the fusion peptide (FP).
(A) and (B) Data are shown as described in Figure 5-3.

SH-H

All four groups consistently recognized peptides spanning the SH-H epitope region (AA 1135-1170). Major sites of escape were located between AA 1145-1158 for all groups (Fig 5-5A). The individual logo plots in the SH-H suggested a consistent response among vaccinated humans and convalescent macaques, with more variability in the remaining groups (S7 Fig). This finding is supported by the within-group escape similarity scores for those groups trending higher across the epitope region (Fig 5-5A lower panel and 5-5B). The median epitope region-wide escape similarity scores for vaccinated humans and convalescent macaques were also higher in the SH-H than in the CTD-N' or FP, confirming a more concordant response. The median between-group escape similarity score for vaccinated humans and convalescent macaques was on par with their median within-group scores, indicating that the escape profile of a vaccinated human looks as similar to that of a convalescent macaque as it does to another vaccinated human (Fig 5-5B). The similarity between these two groups was higher than the similarity between convalescent macaques and humans, as well as between vaccinated macaques and humans (Fig 5-5B). Despite this overall trend, two vaccinated humans had more unique escape profiles (S7 Fig, M26 and M19) and are responsible for a cluster of lower-similarity outlier points (Fig 5-5B, "Vaccinated Humans" and "Conv. Mac. vs. Vacc. Hum.").

The pairwise comparison between participant 352 (a convalescent macaque) and M21 (a vaccinated human) generated an escape similarity score closest to the median for all comparisons between these groups. Logo plots for these individuals are shown in Fig 5-5C as a representative example of the striking between-group similarity. The most consistent sites of escape for both groups were AAs 1148, 1152, 1155, and 1156 (Fig 5-5A and S7). While some differences exist, there was not nearly as much variability as in the CTD-N' (S5 Fig) and FP (S6 Fig).

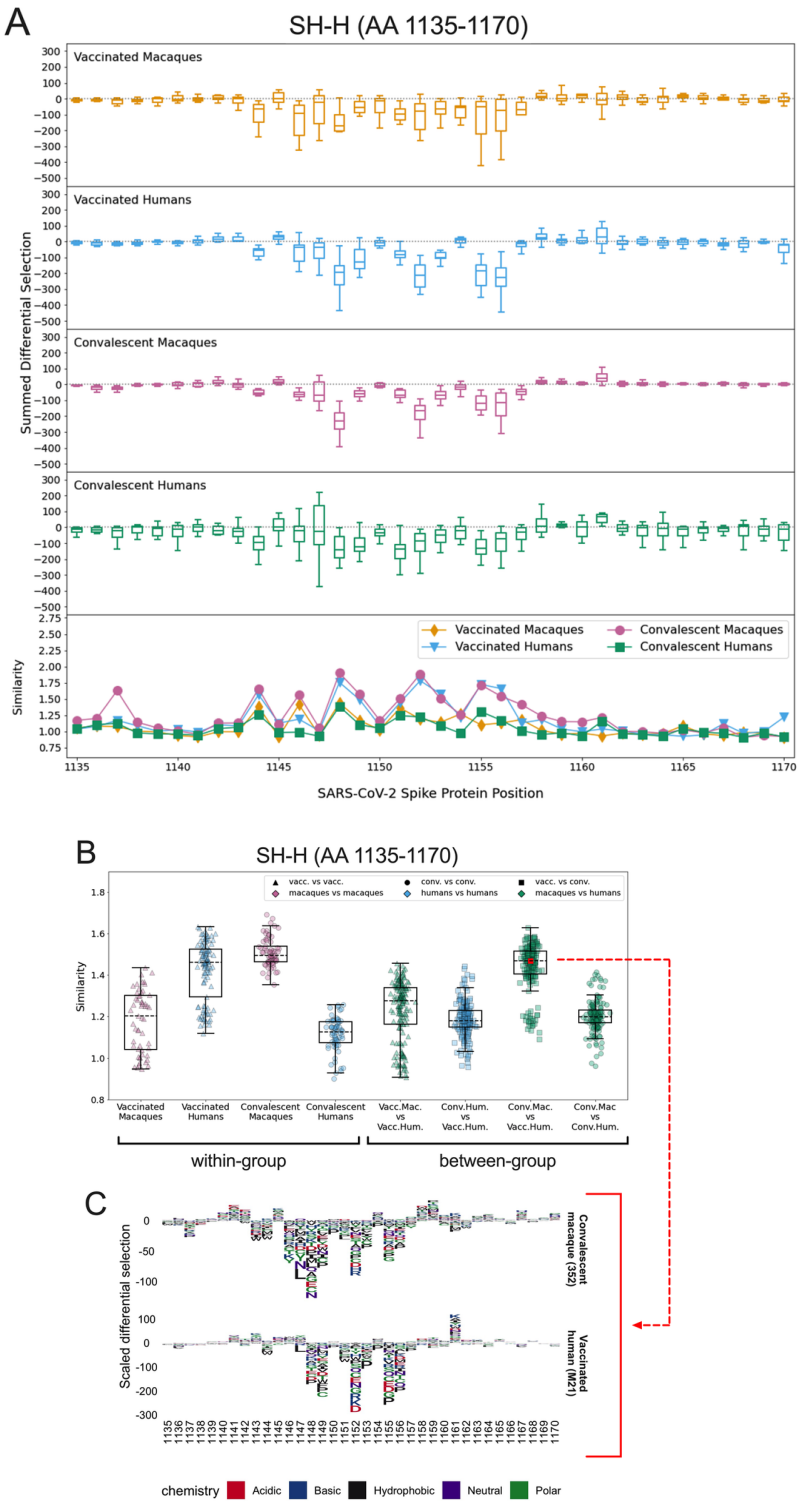

Figure 5-5. Comparison of escape profiles in the stem helix-HR2 region (SH-H).
(A) and (B) Data are shown as described in Figure 5-3. (C) Logo plots for participant 352 (a convalescent macaque) and M21 (a vaccinated human) showing the effect of specific mutations on antibody binding at each site. The comparison between these samples had an escape similarity score closest to the median value for all pairwise convalescent macaque vs. vaccinated human comparisons and thus can be considered representative of the similarity between these groups. The 352 –M21 comparison is shown in red on (B).

Other epitope regions

In addition to the epitope regions described above, the convalescent macaques strongly recognized the pre-FP and post-FP, which were not targeted by human antibody responses (S8 Fig). Escape profiles in the pre-FP appeared highly consistent among individual macaques, with major sites of escape at AAs 795, 798, 800, and 802. Profiles were more variable in the post-FP, likely due in part to low enrichment of wildtype peptides in this epitope region for some individuals (S8 Fig).

5.6 Comparison of vaccinated humans and convalescent macaques

It was notable that the vaccinated humans and convalescent macaques showed the most similarity in escape profiles across all epitope regions, most strikingly in the SH-H. Thus, we also asked whether they showed similarity in the epitopes they targeted by comparing the enrichment of wildtype peptides in these groups in each epitope region (S9 Fig). Vaccinated humans recognized the following epitope regions more strongly than convalescent macaques: NTD (Mann-Whitney $p \leq 0.0001$), CTD-C' ($p \leq 0.0001$), and HR2 ($p \leq 0.001$). Convalescent macaques preferentially recognized the pre-FP ($p \leq 0.0001$) and post-FP ($p \leq 0.001$) epitope regions. This suggests some diversity in the epitopes targeted, but similarity of antibody escape patterns within epitopes targeted by both groups.

5.7 Discussion

In this study, we aimed to assess whether the antibody binding specificities to SARS-CoV-2 Spike in macaques are a useful model for the human response. Our results indicate important similarities between macaques and humans; for example, both have antibodies that recognize major epitopes in the CTD, FP, and SH-H. However, many differences are also

apparent, with some groups showing responses to unique epitopes, such as two physically proximal epitopes in the NTD and CTD that are recognized by antibodies from vaccinated humans but not macaques. Additionally, epitope regions flanking the FP were recognized by antibodies from convalescent macaques, while antibodies from convalescent humans did not recognize the flanking regions but showed a strong response within the FP itself. We found considerable diversity in the pathways of escape between individuals, and this was not specific to either macaques or humans, suggesting a diverse repertoire of antibodies that can respond to the major epitopes in both groups. Overall, these results suggest that macaques and humans share recognition of certain major epitopes. The differences that exist could be due to species (macaque vs. human), but could also be influenced by differences in the specific type and number of exposures to antigen in each group.

Other studies have characterized human monoclonal antibodies against some of the epitopes we report here, many of them with neutralizing or other activities. As previously reported by our group [257], we found that antibodies from vaccinated humans bound peptides spanning a 30 amino acid segment at the C-terminus of the NTD. Interestingly, most if not all neutralizing human mAbs targeting the SARS-CoV-2 NTD to date have been shown to target a single supersite on the "tip" of Spike, distinct from the epitope we detected at the C-terminus [249,261-267]. An NTD mAb with Fc effector function [254], as well as several NTD mAbs that enhance infection in vitro [262,268], also bind sites upstream of the C-terminal epitope. Therefore, future studies are warranted to investigate the function of antibodies binding the new NTD epitope detected by Phage-DMS. In the CTD, we detected broad antibody binding, with vaccinated macaques, vaccinated humans, and convalescent macaques enriching peptides in the CTD-N' epitope region, and vaccinated humans also recognizing peptides spanning the

remainder of this domain (CTD-C'). Polyclonal antibodies targeting sites within the CTD-N' and CTD-C' have been isolated from human sera and shown to have neutralizing activity [269]. Interestingly, the neutralizing epitope on the CTD-C' (AA 625-636) [269] is physically adjacent to the NTD epitope we describe (AA 285-305), raising the possibility that a conformational epitope extending to the NTD is recognized by neutralizing antibodies from vaccinated humans. Depleting human serum of FP-binding antibodies reduced its neutralization capacity [270]; these antibodies are of high interest, both due to their potential to block membrane fusion, and given the high sequence conservation among the FPs of diverse coronaviruses [271,272]. We found that convalescent rhesus macaque sera strongly recognized the pre- and post-FP epitope regions, but to our knowledge, functional antibodies against these regions have not been previously described. Finally, the SH-H epitope region we describe is in the stem helix, a region known to be highly conserved across coronaviruses. Broadly neutralizing [273-275] stem helix antibodies have been isolated and suggest an avenue for rational design of a pan-coronavirus vaccine. Interestingly, a mAb raised against the MERS-CoV stem region protected mice against SARS-CoV-2 challenge, despite having no neutralizing activity against SARS-CoV-2 in vitro [276]. The detection of broad antibody binding across Spike supports the continued investigation of non-RBD epitopes, which remain understudied. Some of the epitopes we describe may also be the target of non-neutralizing Fc-effector antibodies [277], and/or antibodies that enhance infection via Fc-independent [268] or Fc-dependent [278] mechanisms. This latter concept may be important in the pathogenesis of COVID-19, though this remains speculative.

Previous work elucidated that pathways of antibody escape to SARS-CoV-2 Spike protein can be quite variable in convalescent humans, with vaccination inducing a more consistent response [257]. In the current study, we found considerable variability in escape

profiles in the FP and CTD-N' in both macaques and humans, though the convalescent rhesus macaques had more concordant escape profiles than other groups. Variability in escape patterns suggests that a diversity of antibodies are targeting these epitopes. Intra-species germline diversity in immunoglobulin genes may help explain why individuals with similar exposures often mount distinct responses [279,280]. On the other hand, escape profiles were more consistent in the SH-H, where the responses of convalescent macaques and vaccinated humans appeared to converge. This conservation of response suggests that highly similar antibodies are dominating the antibody repertoire against this epitope. Convergent antibody responses to SARS-CoV-2 have been reported within human populations [281-283], and our findings here suggest that antibodies from different species may also be able to converge on the same "public" antibody repertoires in a functional sense, despite genetic differences. While a shared escape profile among individuals could suggest that viral escape mutations are more likely to emerge on a population level, another factor to consider is the effect of the mutations on viral fitness. Key domains of the S2 subunit (such as the SH-H epitope) have essential functions and high sequence conservation, suggesting a low tolerance for mutation and thus for escape. Indeed, previous work determined that sites of escape identified by Phage-DMS are not typically mutated at a high frequency in circulating strains of SARS-CoV-2 [256].

While our focus was on understanding how macaques and humans respond to a similar exposure (i.e., vaccination or infection), we also noted similarities in response between re-infected macaques and vaccinated humans. These groups both exhibited the broadest recognition across Spike, although the epitope regions they targeted were somewhat different. As described above, these groups also had highly similar antibody escape profiles in the SH-H. The vaccinated humans and re-infected macaques both received two exposures to high doses of antigen. It is

plausible that re-exposure directed initially diverse antibodies to converge on a more focused response in both scenarios. While it is known that vaccination and infection induce distinct humoral responses against Spike [257,284,285], our data suggest that a second exposure may generate antibodies that better match the vaccine-induced response.

This study had several limitations. Because the Phage-DMS library displays peptides 31AA in length, discontinuous or conformational epitopes are not readily detected using this method. Additionally, epitopes that may normally be glycosylated are exposed for antibody binding in Phage-DMS. There also are known germline-encoded differences in the properties of immunoglobulin subclasses and Fc receptors between macaques and humans, leading to differences in antibody function that cannot be assayed using Phage-DMS [286]. Additionally, our sample set includes variables that limit our ability to draw conclusions about species-specific (macaque vs. human) differences in antibody response. The vaccinated macaques and humans both received RNA vaccines encoding full-length Spike protein, but there were differences in vaccine technology, including: 1) the use of mRNA in the human vaccine vs. repRNA in the macaque vaccine, 2) the stabilization of Spike in its pre-fusion state in the human vaccine, 3) the dosage and number of doses delivered, and 4) the formulation used to deliver the RNA. Despite these differences, we found commonalities in some of the epitopes targeted by antibodies from both groups. The convalescent rhesus macaques also underwent T cell depletion as part of another study, which may have altered the epitope specificity of their antibodies, although we did not find significant differences between depleted and control animals in our analysis (S3 Fig). Additionally, the convalescent rhesus macaques were experimentally infected twice with high titers of virus, compared to the convalescent humans who were naturally infected once. This important discrepancy could be the reason why the response in re-infected macaques aligned

more closely with vaccinated humans than convalescent humans. Studies of re-infected humans would help address this possibility. Finally, we found pre-existing antibody responses in both groups of macaques. Many of our human samples also likely contain cross-reactive antibodies from prior endemic coronavirus infections. These pre-existing responses are difficult to control for but may have influenced our results.

Our findings suggest that while vaccinated and convalescent macaques and humans share recognition of some major epitopes, each group has a unique antibody binding profile. Antibody escape profiles suggest a diversity of individual responses to most epitopes. Important avenues for future study include comparing macaque and human responses to the RBD and evaluating species differences in antibody function. Continued investigation of immunogenic epitopes in conserved regions of Spike is also warranted to inform the development of immunity that is more robust in the face of viral escape.

5.8 Supplementary Material

All supplemental tables and figures for this study are available for download at https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1010155#sec025.

5.9 Materials and Methods

Samples

Vaccinated pigtail macaques

Plasma was collected from 11 pigtail macaques immunized with a replicating RNA (repRNA) vaccine expressing full-length SARS-CoV-2 Spike protein. A subset of these animals was previously described [235]. All animals were housed at the Washington National Primate Research Center (WaNPRC), an accredited facility of the American Association for the

Accreditation of Laboratory Animal Care International (AAALAC). All procedures were approved by the University of Washington's Institutional Animal Care and Use Committee (IACUC) (IACUC #4266-14). Individual macaques received the vaccine by intramuscular injection in either a Lipid InOrganic Nanoparticle (LION) [235] or a Nanostructured Lipid Carrier (NLC) [258] formulation, delivered in a single priming dose of 25μg (n=3) or 250μg (n=6) or in a prime-boost regimen with 50μg doses spaced 4 weeks apart (n=2). All samples were collected 6 weeks post-prime immunization. A subset of these animals also previously received an experimental hepatitis B vaccine as part of another study (n=5).

Convalescent rhesus macaques

Serum was collected from 12 rhesus macaques housed at the Rocky Mountain Laboratories (National Institutes of Health [NIH]), 14 days after the second of two SARS-CoV-2 infections spaced 42 days apart. The SARS-CoV-2 isolate used for infection was nCoV-WA1-2020 (MN985325.1), which was provided by the Centers for Disease Control and Prevention and propagated as described previously [239]. This isolate came from a COVID-19 patient in Washington state in January 2020 and therefore represents an ancestral strain, prior to the emergence of variants. Prior to infection, macaques were variably depleted of CD4+ T cells, CD8+ T cells, CD4+ and CD8+ T cells, or neither, as part of another study. Details of macaque treatment and regulatory approvals are as published previously [239].

Vaccinated humans

We obtained serum from 15 individuals who received two 100μg doses of the Moderna mRNA-1273 vaccine as part of a phase I clinical trial (NCT04283461) [287]. All participants provided written informed consent. Phage-DMS results from these samples were reported previously [257]. Because samples were de-identified, this study was approved by the Fred

Hutchinson Cancer Research Center Institutional Review Board as non-human subjects research. Only samples from individuals aged 18-55 years were included in the current study to better match the young age range of the macaques.

Convalescent humans

Plasma was collected from 12 individuals post-mild COVID-19 illness as part of the Hospitalized or Ambulatory Adults with Respiratory Viral Infections (HAARVI) study in Seattle, WA. The date of symptom onset for these individuals ranged from February - March 2020, representing infections with early circulating strains of SARS-CoV-2. Phage-DMS results from these samples were reported previously [256,257]. This research was approved by the University of Washington Institutional Review Board (IRB number STUDY00000959). Electronic informed consent was obtained for all participants. Again, the sample set was restricted to only include individuals aged 18-55 years to better match other sample groups. All plasma and sera were heat inactivated at 56°C for 1 hour prior to use. Full details of all samples are available in Table 5-1 and S1.

Phage-DMS, Illumina library preparation and deep sequencing

The experimental protocol was performed exactly as described previously [256]. Briefly, an oligonucleotide pool was synthesized that contains sequences coding for peptides of 31 amino acids that tile along the length of the Wuhan-Hu-1 Spike protein sequence [288] in 1 amino acid increments. For each peptide with the wildtype sequence, 19 variations were included that have a single mutation at the middle amino acid, resulting in a total library size of 24,820 unique sequences. The oligonucleotide pool was cloned into T7 phage, followed by amplification of the phage library; this step was performed twice independently to generate biological duplicate phage libraries. The phage library was incubated with a serum or plasma sample, then bound

antibody-phage complexes were immunoprecipitated using Protein A and Protein G Dynabeads (Invitrogen). Bound phage were lysed, and DNA was amplified by PCR and cleaned prior to sequencing on an Illumina MiSeq or HiSeq 2500 with single end reads. Demultiplexing and read alignment were also performed as described previously [257].

Replicate curation

Biological replicates were analyzed in parallel to assess reproducibility of results. For simplicity, results from only one biological replicate are shown and described, with the same figures generated with the second biological replicate available to view online at https://github.com/matsengrp/phage-dms-nhp-analysis. Within each biological replicate, "in-line" technical replicates were run for some samples. In these cases, the technical replicate with the highest mapped read count was selected for analysis.

Wildtype enrichment and defining epitope regions

The enrichment of wildtype peptides was calculated as described previously to quantify the proportion of each peptide in an antibody-selected sample relative to the proportion of that peptide in the input phage library [255]. On enrichment plots, the locus of each peptide is defined by its middle amino acid. Enrichment values of wildtype peptides were summed across epitope regions of interest for statistical comparisons between groups ("Summed WT enrichment" on figures).

Escape profile comparison

The effect of a mutation on antibody-peptide binding was quantified as "differential selection," which is the log fold change in the enrichment of a mutation-containing peptide compared to the wildtype peptide. This number is multiplied by the average of the wildtype

peptide enrichments at that site and its two adjacent sites to get a "scaled differential selection" value, as described previously [257]. The enrichment values of the adjacent wildtype peptides are included in this calculation to make the analysis less susceptible to noise. Negative differential selection values represent reduced binding compared to wildtype, while positive differential selection values indicate that the mutation enhanced binding. "Summed differential selection" is the sum of the 19 scaled differential selection values for all mutations at a site, and gives a sense of the overall magnitude of escape at that site.

The comparison of two escape profiles is quantified by an escape similarity score computed in the framework of an optimal transport problem [289]; this algorithm was described in detail at https://matsengrp.github.io/phippery/esc-prof.html. An overview of the method is shown in S4 Fig. Escape profiles are commonly portrayed as logo plots using scaled differential selection values (S4A Fig). At each site, escape data in logo plot form can instead be represented as binned distributions, with each mutation making some contribution to the total amount of escape at that site based on its scaled differential selection value (S4B Fig). For each site, an optimal transport problem computes the most efficient way to transform one individual's escape distribution into that of a different individual (S4C Fig). The cost to "exchange" amino acid contributions between profiles is based on the similarity between the amino acids being exchanged, as defined by the BLOSUM62 matrix [290]. More "movement" between dissimilar amino acids drives up the total cost of the transport; therefore, a higher cost indicates less similar profiles. Escape similarity scores are the inverse of the total cost of transforming one profile into another. Scores were calculated between pairwise combinations of individuals to compare escape profile variability within and between sample groups.

Protein structure

The structure of a SARS-CoV-2 Spike glycoprotein monomer in the closed state (PDB 6XR8) was examined to visualize epitope regions [291]. Coloring was added using UCSF ChimeraX-1.2.5, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases [292].

Statistical analysis and plotting

For comparison of summed wildtype enrichment values across an epitope region between sample groups (Fig 5-2 and S9), multiple Mann-Whitney U tests were performed, with p values corrected for the number of comparisons in each plot using the Bonferroni-Dunn method. Asterisks represent the following corrected p values: ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$. For comparison of summed wildtype enrichment values between macaque sub-groups (S3 Fig), a Kruskal-Wallis test was used with a significance threshold of $p=0.05$. Boxplots are used to summarize summed wildtype enrichment values, summed differential selection values, and similarity scores within a sample group. For all boxplots, the box represents the median and interquartile range (IQR), the lower whisker represents the lowest data point above Q1-1.5IQR, and the upper whisker represents the highest data point below Q3+1.5IQR.

Code, software, and data availability

All analyses were performed in RStudio version 1.3.1093, Python version 3.6.12, GraphPad Prism version 9.0.1, and the phip-flow and phippery software suite (https://matsengrp.github.io/phippery/). The phip-flow tools perform read alignment using

Bowtie2 [293] in a Nextflow [294] pipeline script. The escape profile comparisons are done with phippery in Python 3.6.12 and depend on the NumPy [128], pandas [295,296], xarray [297], POT [298], and biopython [299] packages. All code and instructions for running this analysis are available at https://github.com/matsengrp/phage-dms-nhp-analysis.

References

1. Vasilakis N, Weaver SC. Flavivirus transmission focusing on Zika. Curr Opin Virol. 2017; 22:30-35.
2. Musso D, Gubler DJ. Zika Virus. Clin Microbiol Rev. 2016; 29(3):487-524.
3. Musso D, Ko AI, Baud D. Zika Virus Infection - After the Pandemic. N Engl J Med. 2019; 381(15):1444-1457.
4. Dick GW, Kitchen SF, Haddow AJ. Zika virus. I. Isolations and serological specificity. Trans R Soc Trop Med Hyg. 1952; 46(5):509-20.
5. Macnamara FN. Zika virus: a report on three cases of human infection during an epidemic of jaundice in Nigeria. Trans R Soc Trop Med Hyg. 1954; 48(2):139-45.
6. Olson JG, Ksiazek TG, Suhandiman, Triwibowo. Zika virus, a cause of fever in Central Java, Indonesia. Trans R Soc Trop Med Hyg. 1981; 75(3):389-93.
7. Paixao ES, Barreto F, Teixeira Mda G, Costa Mda C, Rodrigues LC. History, Epidemiology, and Clinical Manifestations of Zika: A Systematic Review. Am J Public Health. 2016; 106(4):606-12.
8. Duffy MR, Chen TH, Hancock WT, Powers AM, Kool JL, Lanciotti RS, et al. Zika virus outbreak on Yap Island, Federated States of Micronesia. N Engl J Med. 2009; 360(24):2536-43.
9. Aubry M, Teissier A, Huart M, Merceron S, Vanhomwegen J, Roche C, et al. Zika Virus Seroprevalence, French Polynesia, 2014-2015. Emerg Infect Dis. 2017; 23(4):669-672.
10. Musso D, Bossin H, Mallet HP, Besnard M, Broult J, Baudouin L, et al. Zika virus in French Polynesia 2013-14: anatomy of a completed outbreak. Lancet Infect Dis. 2018; 18(5):e172-e182.
11. Faria NR, Azevedo R, Kraemer MUG, Souza R, Cunha MS, Hill SC, et al. Zika virus in the Americas: Early epidemiological and genetic findings. Science. 2016; 352(6283):345-349.
12. Hennessey M, Fischer M, Staples JE. Zika Virus Spreads to New Areas - Region of the Americas, May 2015-January 2016. MMWR Morb Mortal Wkly Rep. 2016; 65(3):55-8.
13. Zanluca C, Melo VC, Mosimann AL, Santos GI, Santos CN, Luz K. First report of autochthonous transmission of Zika virus in Brazil. Mem Inst Oswaldo Cruz. 2015; 110(4):569-72.
14. Yakob L. Zika Virus after the Public Health Emergency of International Concern Period, Brazil. Emerg Infect Dis. 2022; 28(4):837-840.
15. Schuler-Faccini L, Ribeiro EM, Feitosa IM, Horovitz DD, Cavalcanti DP, Pessoa A, et al. Possible Association Between Zika Virus Infection and Microcephaly - Brazil, 2015. MMWR Morb Mortal Wkly Rep. 2016; 65(3):59-62.
16. Moore SM, Oidtman RJ, Soda KJ, Siraj AS, Reiner RC, Jr., Barker CM, et al. Leveraging multiple data types to estimate the size of the Zika epidemic in the Americas. PLoS Negl Trop Dis. 2020; 14(9):e0008640.
17. Ribeiro GS, Hamer GL, Diallo M, Kitron U, Ko AI, Weaver SC. Influence of herd immunity in the cyclical nature of arboviruses. Curr Opin Virol. 2020; 40:1-10.

18. WHO. Zika epidemiology update - February 2022. 2022.
19. Gobillot TA, Kikawa C, Lehman DA, Kinuthia J, Drake AL, Jaoko W, et al. Zika Virus Circulates at Low Levels in Western and Coastal Kenya. J Infect Dis. 2020; 222(5):847-852.
20. Willcox AC, Collins MH, Jadi R, Keeler C, Parr JB, Mumba D, et al. Seroepidemiology of Dengue, Zika, and Yellow Fever Viruses among Children in the Democratic Republic of the Congo. Am J Trop Med Hyg. 2018; 99(3):756-763.
21. Counotte MJ, Althaus CL, Low N, Riou J. Impact of age-specific immunity on the timing and burden of the next Zika virus outbreak. PLoS Negl Trop Dis. 2019; 13(12):e0007978.
22. da Costa Castilho M, de Filippis AMB, Machado LC, de Lima Calvanti TYV, Lima MC, Fonseca V, et al. Evidence of Zika Virus Reinfection by Genome Diversity and Antibody Response Analysis, Brazil. Emerg Infect Dis. 2024; 30(2):310-320.
23. Kucharski AJ, Funk S, Eggo RM, Mallet HP, Edmunds WJ, Nilles EJ. Transmission Dynamics of Zika Virus in Island Populations: A Modelling Analysis of the 2013-14 French Polynesia Outbreak. PLoS Negl Trop Dis. 2016; 10(5):e0004726.
24. Ryan SJ, Carlson CJ, Tesla B, Bonds MH, Ngonghala CN, Mordecai EA, et al. Warming temperatures could expose more than 1.3 billion new people to Zika virus risk by 2050. Glob Chang Biol. 2021; 27(1):84-93.
25. Rivera-Correa J, de Siqueira IC, Mota S, do Rosario MS, Pereira de Jesus PA, Alcantara LCJ, et al. Anti-ganglioside antibodies in patients with Zika virus infection-associated Guillain-Barre Syndrome in Brazil. PLoS Negl Trop Dis. 2019; 13(9):e0007695.
26. Rice ME, Galang RR, Roth NM, Ellington SR, Moore CA, Valencia-Prado M, et al. Vital Signs: Zika-Associated Birth Defects and Neurodevelopmental Abnormalities Possibly Associated with Congenital Zika Virus Infection - U.S. Territories and Freely Associated States, 2018. MMWR Morb Mortal Wkly Rep. 2018; 67(31):858-867.
27. Einspieler C, Utsch F, Brasil P, Panvequio Aizawa CY, Peyton C, Hydee Hasue R, et al. Association of Infants Exposed to Prenatal Zika Virus Infection With Their Clinical, Neurologic, and Developmental Status Evaluated via the General Movement Assessment Tool. JAMA Netw Open. 2019; 2(1):e187235.
28. Nielsen-Saines K, Brasil P, Kerin T, Vasconcelos Z, Gabaglia CR, Damasceno L, et al. Delayed childhood neurodevelopment and neurosensory alterations in the second year of life in a prospective cohort of ZIKV-exposed children. Nat Med. 2019; 25(8):1213-1217.
29. Rosinski JR, Raasch LE, Barros Tiburcio P, Breitbach ME, Shepherd PM, Yamamoto K, et al. Frequent first-trimester pregnancy loss in rhesus macaques infected with African-lineage Zika virus. PLoS Pathog. 2023; 19(3):e1011282.
30. Aubry F, Jacobs S, Darmuzey M, Lequime S, Delang L, Fontaine A, et al. Recent African strains of Zika virus display higher transmissibility and fetal pathogenicity than Asian strains. Nat Commun. 2021; 12(1):916.
31. Raasch LE, Yamamoto K, Newman CM, Rosinski JR, Shepherd PM, Razo E, et al. Fetal loss in pregnant rhesus macaques infected with high-dose African-lineage Zika virus. PLoS Negl Trop Dis. 2022; 16(8):e0010623.
32. World Health Organization. Laboratory testing for Zika virus and dengue virus infections: interim guidance. July 14, 2022 (https://www.who.int/publications/i/item/WHO-ZIKV_DENV-LAB-2022.1).

33. Pierson TC, Diamond MS. The emergence of Zika virus and its new clinical syndromes. Nature. 2018; 560(7720):573-581.
34. Liu ZY, Shi WF, Qin CF. The evolution of Zika virus from Asia to the Americas. Nat Rev Microbiol. 2019; 17(3):131-139.
35. Puerta-Guardo H, Glasner DR, Espinosa DA, Biering SB, Patana M, Ratnasiri K, et al. Flavivirus NS1 Triggers Tissue-Specific Vascular Endothelial Dysfunction Reflecting Disease Tropism. Cell Rep. 2019; 26(6):1598-1613 e8.
36. Brett D. Lindenbach GR, Ralf Bartenschlager, Charles M. Rice. Fields Virology Seventh Edition Volume 1: Emerging Viruses. Chapter 7: Flaviviridae: The Viruses and Their Replication2020.
37. Pierson TC, Diamond MS. The continued threat of emerging flaviviruses. Nat Microbiol. 2020; 5(6):796-812.
38. Serman TM, Gack MU. Evasion of Innate and Intrinsic Antiviral Pathways by the Zika Virus. Viruses. 2019; 11(10).
39. McNab F, Mayer-Barber K, Sher A, Wack A, O'Garra A. Type I interferons in infectious disease. Nat Rev Immunol. 2015; 15(2):87-103.
40. Wittling MC, Cahalan SR, Levenson EA, Rabin RL. Shared and Unique Features of Human Interferon-Beta and Interferon-Alpha Subtypes. Front Immunol. 2020; 11:605673.
41. Lazear HM, Schoggins JW, Diamond MS. Shared and Distinct Functions of Type I and Type III Interferons. Immunity. 2019; 50(4):907-923.
42. Platanias LC. Mechanisms of type-I- and type-II-interferon-mediated signalling. Nat Rev Immunol. 2005; 5(5):375-86.
43. Schneider WM, Chevillotte MD, Rice CM. Interferon-stimulated genes: a complex web of host defenses. Annu Rev Immunol. 2014; 32:513-45.
44. Ganser-Pornillos BK, Pornillos O. Restriction of HIV-1 and other retroviruses by TRIM5. Nat Rev Microbiol. 2019; 17(9):546-556.
45. Schoggins JW, Rice CM. Interferon-stimulated genes and their antiviral effector functions. Curr Opin Virol. 2011; 1(6):519-25.
46. Kristiansen H, Gad HH, Eskildsen-Larsen S, Despres P, Hartmann R. The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities. J Interferon Cytokine Res. 2011; 31(1):41-7.
47. Gobillot TA, Humes D, Sharma A, Kikawa C, Overbaugh J. The Robust Restriction of Zika Virus by Type-I Interferon in A549 Cells Varies by Viral Lineage and Is Not Determined by IFITM3. Viruses. 2020; 12(5):503.
48. Frumence E, Roche M, Krejbich-Trotot P, El-Kalamouni C, Nativel B, Rondeau P, et al. The South Pacific epidemic strain of Zika virus replicates efficiently in human epithelial A549 cells leading to IFN-beta production and apoptosis induction. Virology. 2016; 493:217-26.
49. Hamel R, Dejarnac O, Wichit S, Ekchariyawat P, Neyret A, Luplertlop N, et al. Biology of Zika Virus Infection in Human Skin Cells. J Virol. 2015; 89(17):8880-96.
50. Lazear HM, Govero J, Smith AM, Platt DJ, Fernandez E, Miner JJ, et al. A Mouse Model of Zika Virus Pathogenesis. Cell Host Microbe. 2016; 19(5):720-30.

51. Dowall SD, Graham VA, Rayner E, Atkinson B, Hall G, Watson RJ, et al. A Susceptible Mouse Model for Zika Virus Infection. PLoS Negl Trop Dis. 2016; 10(5):e0004658.
52. Rossi SL, Tesh RB, Azar SR, Muruato AE, Hanley KA, Auguste AJ, et al. Characterization of a Novel Murine Model to Study Zika Virus. Am J Trop Med Hyg. 2016; 94(6):1362-1369.
53. Richardson RB, Ohlson MB, Eitson JL, Kumar A, McDougal MB, Boys IN, et al. A CRISPR screen identifies IFI6 as an ER-resident interferon effector that blocks flavivirus replication. Nat Microbiol. 2018; 3(11):1214-1223.
54. Dukhovny A, Lamkiewicz K, Chen Q, Fricke M, Jabrane-Ferrat N, Marz M, et al. A CRISPR Activation Screen Identifies Genes That Protect against Zika Virus Infection. J Virol. 2019; 93(16):e00211-19.
55. Panayiotou C, Lindqvist R, Kurhade C, Vonderstein K, Pasto J, Edlund K, et al. Viperin Restricts Zika Virus and Tick-Borne Encephalitis Virus Replication by Targeting NS3 for Proteasomal Degradation. J Virol. 2018; 92(7):e02054-17.
56. Van der Hoek KH, Eyre NS, Shue B, Khantisitthiporn O, Glab-Ampi K, Carr JM, et al. Viperin is an important host restriction factor in control of Zika virus infection. Sci Rep. 2017; 7(1):4475.
57. Li L, Zhao H, Liu P, Li C, Quanquin N, Ji X, et al. PARP12 suppresses Zika virus infection through PARP-dependent degradation of NS1 and NS3 viral proteins. Sci Signal. 2018; 11(535):eaas9332.
58. Li L, Shi Y, Li S, Liu J, Zu S, Xu X, et al. ADP-ribosyltransferase PARP11 suppresses Zika virus in synergy with PARP12. Cell Biosci. 2021; 11(1):116.
59. Savidis G, Perreira JM, Portmann JM, Meraner P, Guo Z, Green S, et al. The IFITMs Inhibit Zika Virus Replication. Cell Rep. 2016; 15(11):2323-30.
60. Monel B, Compton AA, Bruel T, Amraoui S, Burlaud-Gaillard J, Roy N, et al. Zika virus induces massive cytoplasmic vacuolization and paraptosis-like death in infected cells. EMBO J. 2017; 36(12):1653-1668.
61. Spence JS, He R, Hoffmann HH, Das T, Thinon E, Rice CM, et al. IFITM3 directly engages and shuttles incoming virus particles to lysosomes. Nat Chem Biol. 2019; 15(3):259-268.
62. Singh PK, Guest JM, Kanwar M, Boss J, Gao N, Juzych MS, et al. Zika virus infects cells lining the blood-retinal barrier and causes chorioretinal atrophy in mouse eyes. JCI Insight. 2017; 2(4):e92340.
63. Singh PK, Singh S, Farr D, Kumar A. Interferon-stimulated gene 15 (ISG15) restricts Zika virus replication in primary human corneal epithelial cells. Ocul Surf. 2019; 17(3):551-559.
64. Ding J, Aldo P, Roberts CM, Stabach P, Liu H, You Y, et al. Placenta-derived interferon-stimulated gene 20 controls ZIKA virus infection. EMBO Rep. 2021; 22(10):e52450.
65. Mounce BC, Poirier EZ, Passoni G, Simon-Loriere E, Cesaro T, Prot M, et al. Interferon-Induced Spermidine-Spermine Acetyltransferase and Polyamine Depletion Restrict Zika and Chikungunya Viruses. Cell Host Microbe. 2016; 20(2):167-77.

66. Li C, Deng YQ, Wang S, Ma F, Aliyari R, Huang XY, et al. 25-Hydroxycholesterol Protects Host against Zika Virus Infection and Its Associated Microcephaly in a Mouse Model. Immunity. 2017; 46(3):446-456.
67. Han Y, Bai X, Liu S, Zhu J, Zhang F, Xie L, et al. XAF1 Protects Host against Emerging RNA Viruses by Stabilizing IRF1-Dependent Antiviral Immunity. J Virol. 2022; 96(17):e0077422.
68. Wu Y, Yang X, Yao Z, Dong X, Zhang D, Hu Y, et al. C19orf66 interrupts Zika virus replication by inducing lysosomal degradation of viral NS3. PLoS Negl Trop Dis. 2020; 14(3):e0008083.
69. Hanners NW, Mar KB, Boys IN, Eitson JL, De La Cruz-Rivera PC, Richardson RB, et al. Shiftless inhibits flavivirus replication in vitro and is neuroprotective in a mouse model of Zika virus pathogenesis. Proc Natl Acad Sci U S A. 2021; 118(49).
70. Sun N, Zhang RR, Song GY, Cai Q, Aliyari SR, Nielsen-Saines K, et al. SERTAD3 induces proteasomal degradation of ZIKV capsid protein and represents a therapeutic target. J Med Virol. 2023; 95(2):e28451.
71. Valdez F, Salvador J, Palermo PM, Mohl JE, Hanley KA, Watts D, et al. Schlafen 11 Restricts Flavivirus Replication. J Virol. 2019; 93(15).
72. Zu S, Li C, Li L, Deng YQ, Chen X, Luo D, et al. TRIM22 suppresses Zika virus replication by targeting NS1 and NS3 for proteasomal degradation. Cell Biosci. 2022; 12(1):139.
73. Pawlak JB, Hsu JC, Xia H, Han P, Suh HW, Grove TL, et al. CMPK2 restricts Zika virus replication by inhibiting viral translation. PLoS Pathog. 2023; 19(4):e1011286.
74. Li S, Qian N, Jiang C, Zu W, Liang A, Li M, et al. Gain-of-function genetic screening identifies the antiviral function of TMEM120A via STING activation. Nat Commun. 2022; 13(1):105.
75. Lesage S, Chazal M, Beauclair G, Batalie D, Cerboni S, Couderc E, et al. Discovery of Genes that Modulate Flavivirus Replication in an Interferon-Dependent Manner. J Mol Biol. 2022; 434(6):167277.
76. Coutard B, Barral K, Lichiere J, Selisko B, Martin B, Aouadi W, et al. Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. J Virol. 2017; 91(5).
77. Xia H, Luo H, Shan C, Muruato AE, Nunes BTD, Medeiros DBA, et al. An evolutionary NS1 mutation enhances Zika virus evasion of host interferon induction. Nat Commun. 2018; 9(1):414.
78. Riedl W, Acharya D, Lee JH, Liu G, Serman T, Chiang C, et al. Zika Virus NS3 Mimics a Cellular 14-3-3-Binding Motif to Antagonize RIG-I- and MDA5-Mediated Innate Immunity. Cell Host Microbe. 2019; 26(4):493-503 e6.
79. Wu Y, Liu Q, Zhou J, Xie W, Chen C, Wang Z, et al. Zika virus evades interferon-mediated antiviral response through the co-operation of multiple nonstructural proteins in vitro. Cell Discov. 2017; 3:17006.
80. Grant A, Ponia SS, Tripathi S, Balasubramaniam V, Miorin L, Sourisseau M, et al. Zika Virus Targets Human STAT2 to Inhibit Type I Interferon Signaling. Cell Host Microbe. 2016; 19(6):882-90.

81. Santos CNO RD, Cardoso Alves J, Cazzaniga RA, Magalhães LS, de Souza MSF, Fonseca ABL, Bispo AJB, Porto RLS, Santos CAD, da Silva ÂM, Teixeira MM, de Almeida RP, de Jesus AR. Association Between Zika Virus Microcephaly in Newborns With the rs3775291 Variant in Toll-Like Receptor 3 and rs1799964 Variant at Tumor Necrosis Factor-α Gene. J Infect Dis. 2019; 220(11):1797-1801.
82. Santos CNO, Magalhaes LS, Fonseca ABL, Bispo AJB, Porto RLS, Alves JC, et al. Association between genetic variants in TREM1, CXCL10, IL4, CXCL8 and TLR7 genes with the occurrence of congenital Zika syndrome and severe microcephaly. Sci Rep. 2023; 13(1):3466.
83. Alippe Y, Wang L, Coskun R, Muraro SP, Zhao FR, Elam-Noll M, et al. Fetal MAVS and type I IFN signaling pathways control ZIKV infection in the placenta and maternal decidua. J Exp Med. 2024; 221(9).
84. Cheon H, Holvey-Bates EG, Schoggins JW, Forster S, Hertzog P, Imanaka N, et al. IFNbeta-dependent increases in STAT1, STAT2, and IRF9 mediate resistance to viruses and DNA damage. EMBO J. 2013; 32(20):2751-63.
85. Cheon H, Stark GR. Unphosphorylated STAT1 prolongs the expression of interferon-induced immune regulatory genes. Proc Natl Acad Sci U S A. 2009; 106(23):9373-8.
86. Sung PS, Cheon H, Cho CH, Hong SH, Park DY, Seo HI, et al. Roles of unphosphorylated ISGF3 in HCV infection and interferon responsiveness. Proc Natl Acad Sci U S A. 2015; 112(33):10443-8.
87. Radaeva S, Jaruga B, Kim WH, Heller T, Liang TJ, Gao B. Interferon-gamma inhibits interferon-alpha signalling in hepatic cells: evidence for the involvement of STAT1 induction and hyperexpression of STAT1 in chronic hepatitis C. Biochem J. 2004; 379(Pt 1):199-208.
88. van Boxel-Dezaire AH, Rani MR, Stark GR. Complex modulation of cell type-specific signaling in response to type I interferons. Immunity. 2006; 25(3):361-72.
89. PAHO/WHO. 2016: the year Zika evolved from an emergency into a long-term public health challenge. 2016.
90. Rasmussen SA, Jamieson DJ, Honein MA, Petersen LR. Zika Virus and Birth Defects--Reviewing the Evidence for Causality. N Engl J Med. 2016; 374(20):1981-7.
91. Cao-Lormeau VM, Blake A, Mons S, Lastere S, Roche C, Vanhomwegen J, et al. Guillain-Barre Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. Lancet. 2016; 387(10027):1531-1539.
92. van den Berg B, Walgaard C, Drenthen J, Fokke C, Jacobs BC, van Doorn PA. Guillain-Barre syndrome: pathogenesis, diagnosis, treatment and prognosis. Nat Rev Neurol. 2014; 10(8):469-82.
93. Mazewski C, Perez RE, Fish EN, Platanias LC. Type I Interferon (IFN)-Regulated Activation of Canonical and Non-Canonical Signaling Pathways. Front Immunol. 2020; 11:606456.
94. Moleirinho S, Guerrant W, Kissil JL. The Angiomotins--from discovery to function. FEBS Lett. 2014; 588(16):2693-703.
95. OhAinle M, Helms L, Vermeire J, Roesch F, Humes D, Basom R, et al. A virus-packageable CRISPR screen identifies host factors mediating interferon inhibition of HIV. Elife. 2018; 7:e39823.

96. Boonyaratanakornkit JB, Bartlett EJ, Amaro-Carambot E, Collins PL, Murphy BR, Schmidt AC. The C proteins of human parainfluenza virus type 1 (HPIV1) control the transcription of a broad array of cellular genes that would otherwise respond to HPIV1 infection. J Virol. 2009; 83(4):1892-910.
97. Hsu YL, Shi SF, Wu WL, Ho LJ, Lai JH. Protective roles of interferon-induced protein with tetratricopeptide repeats 3 (IFIT3) in dengue virus infection of human lung epithelial cells. PLoS One. 2013; 8(11):e79518.
98. Souza BS, Sampaio GL, Pereira CS, Campos GS, Sardi SI, Freitas LA, et al. Zika virus infection induces mitosis abnormalities and apoptotic cell death of human neural progenitor cells. Sci Rep. 2016; 6:39775.
99. Fusco DN, Pratt H, Kandilas S, Cheon SS, Lin W, Cronkite DA, et al. HELZ2 Is an IFN Effector Mediating Suppression of Dengue Virus. Front Microbiol. 2017; 8:240.
100. Fusco DN, Brisac C, John SP, Huang YW, Chin CR, Xie T, et al. A genetic screen identifies interferon-alpha effector genes required to suppress hepatitis C virus replication. Gastroenterology. 2013; 144(7):1438-49, 1449 e1-9.
101. Nordmann A, Wixler L, Boergeling Y, Wixler V, Ludwig S. A new splice variant of the human guanylate-binding protein 3 mediates anti-influenza activity through inhibition of viral transcription and replication. FASEB J. 2012; 26(3):1290-300.
102. Bamford CGG, Aranday-Cortes, E., Sanchez-Velazquez, R., Catrina Mullan 1,, Alain Kohl 1 AHP, Sam J.Wilson 1 and John McLauchlan. A Human and Rhesus Macaque Interferon-Stimulated Gene Screen Shows That Over-Expression of ARHGEF3/XPLN Inhibits Replication of Hepatitis C Virus and Other Flavivirids. Viruses. 2022; 14(1655).
103. Colli ML, Ramos-Rodriguez M, Nakayasu ES, Alvelos MI, Lopes M, Hill JLE, et al. An integrated multi-omics approach identifies the landscape of interferon-alpha-mediated responses of human pancreatic beta cells. Nat Commun. 2020; 11(1):2584.
104. Goebel S, Snyder B, Sellati T, Saeed M, Ptak R, Murray M, et al. A sensitive virus yield assay for evaluation of Antivirals against Zika Virus. J Virol Methods. 2016; 238:13-20.
105. Lucci V, Di Palma T, D'Ambrosio C, Scaloni A, Zannini M. AMOTL2 interaction with TAZ causes the inhibition of surfactant proteins expression in lung cells. Gene. 2013; 529(2):300-6.
106. Wang W, Huang J, Chen J. Angiomotin-like proteins associate with and negatively regulate YAP1. J Biol Chem. 2011; 286(6):4364-70.
107. Wang S, Zhou L, Ling L, Meng X, Chu F, Zhang S, et al. The Crosstalk Between Hippo-YAP Pathway and Innate Immunity. Front Immunol. 2020; 11:323.
108. Patrie KM. Identification and characterization of a novel tight junction-associated family of proteins that interacts with a WW domain of MAGI-1. Biochim Biophys Acta. 2005; 1745(1):131-44.
109. Hildebrand S, Hultin S, Subramani A, Petropoulos S, Zhang Y, Cao X, et al. The E-cadherin/AmotL2 complex organizes actin filaments required for epithelial hexagonal packing and blastocyst hatching. Sci Rep. 2017; 7(1):9540.
110. Wang Y, Li Z, Xu P, Huang L, Tong J, Huang H, et al. Angiomotin-like2 gene (amotl2) is required for migration and proliferation of endothelial cells during angiogenesis. J Biol Chem. 2011; 286(47):41095-104.

111. Bratt A, Wilson WJ, Troyanovsky B, Aase K, Kessler R, Van Meir EG, et al. Angiomotin belongs to a novel protein family with conserved coiled-coil and PDZ binding domains. Gene. 2002; 298(1):69-77.
112. Pocaterra A, Romani P, Dupont S. YAP/TAZ functions and their regulation at a glance. J Cell Sci. 2020; 133(2).
113. Fang C, Li J, Qi S, Lei Y, Zeng Y, Yu P, et al. An alternatively transcribed TAZ variant negatively regulates JAK-STAT signaling. EMBO Rep. 2019; 20(6).
114. Reynoso GV, Gordon DN, Kalia A, Aguilar CC, Malo CS, Aleshnick M, et al. Zika virus spreads through infection of lymph node-resident macrophages. Cell Rep. 2023; 42(2):112126.
115. Quicke KM, Bowen JR, Johnson EL, McDonald CE, Ma H, O'Neal JT, et al. Zika Virus Infects Human Placental Macrophages. Cell Host Microbe. 2016; 20(1):83-90.
116. Contreras D, Arumugaswami V. Zika Virus Infectious Cell Culture System and the In Vitro Prophylactic Effect of Interferons. J Vis Exp. 2016; (114).
117. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods. 2014; 11(8):783-784.
118. Li J, Ding SC, Cho H, Chung BC, Gale M, Jr., Chanda SK, et al. A short hairpin RNA screen of interferon-stimulated genes identifies a novel negative regulator of the cellular antiviral response. mBio. 2013; 4(3):e00385-13.
119. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001; 25(4):402-8.
120. Lu AY, Gustin A, Newhouse D, Gale M, Jr. Viral Protein Accumulation of Zika Virus Variants Links with Regulation of Innate Immunity for Differential Control of Viral Replication, Spread, and Response to Interferon. J Virol. 2023; 97(5):e0198222.
121. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10(3):R25.
122. Wang B, Wang M, Zhang W, Xiao T, Chen CH, Wu A, et al. Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute. Nat Protoc. 2019; 14(3):756-780.
123. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. 2014; 15(12):554.
124. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29(1):15-21.
125. Frankish A, Diekhans M, Jungreis I, Lagarde J, Loveland JE, Mudge JM, et al. Gencode 2021. Nucleic Acids Res. 2021; 49(D1):D916-D923.
126. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26(1):139-40.
127. Van Rossum G, Drake FL. Python 3 Reference Manual. Scotts Valley, CA: CreateSpace; 2009.
128. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature. 2020; 585(7825):357-362.
129. The pandas development team. pandas-dev/pandas: Pandas. Zenodo. 2020.

130. Davis MA, Voss K, Turnbull JB, Gustin AT, Knoll M, Muruato A, et al. A C57BL/6 Mouse Model of SARS-CoV-2 Infection Recapitulates Age- and Sex-Based Differences in Human COVID-19 Disease and Recovery. Vaccines (Basel). 2022; 11(1).
131. Blaszczyk K, Olejnik A, Nowicka H, Ozgyin L, Chen YL, Chmielewski S, et al. STAT2/IRF9 directs a prolonged ISGF3-like transcriptional response and antiviral activity in the absence of STAT1. Biochem J. 2015; 466(3):511-24.
132. Freund EC, Lock JY, Oh J, Maculins T, Delamarre L, Bohlen CJ, et al. Efficient gene knockout in primary human and murine myeloid cells by non-viral delivery of CRISPR-Cas9. J Exp Med. 2020; 217(7).
133. Rempel H, Sun B, Calosing C, Pillai SK, Pulliam L. Interferon-alpha drives monocyte gene expression in chronic unsuppressed HIV-1 infection. AIDS. 2010; 24(10):1415-23.
134. Kreiss JK, Koech D, Plummer FA, Holmes KK, Lightfoote M, Piot P, et al. AIDS virus infection in Nairobi prostitutes. Spread of the epidemic to East Africa. N Engl J Med. 1986; 314(7):414-8.
135. UNAIDS. Fact Sheet—World AIDS Day 2020. https://www.unaids.org/en/resources/documents/2020/UNAIDS_FactSheet.
136. Shannon K, Crago AL, Baral SD, Bekker LG, Kerrigan D, Decker MR, et al. The global response and unmet actions for HIV and sex workers. Lancet. 2018; 392(10148):698-710.
137. UNAIDS. Miles to go—closing gaps, breaking barriers, righting injustices. 2018. https://www.unaids.org/en/resources/documents/2018/global-aids-update.
138. Eakle R, Gomez GB, Naicker N, Bothma R, Mbogua J, Cabrera Escobar MA, et al. HIV pre-exposure prophylaxis and early antiretroviral treatment among female sex workers in South Africa: Results from a prospective observational demonstration project. PLoS Med. 2017; 14(11):e1002444.
139. Koss CA, Charlebois ED, Ayieko J, Kwarisiima D, Kabami J, Balzer LB, et al. Uptake, engagement, and adherence to pre-exposure prophylaxis offered after population HIV testing in rural Kenya and Uganda: 72-week interim analysis of observational data from the SEARCH study. Lancet HIV. 2020; 7(4):e249-e261.
140. Glick JL, Russo R, Jivapong B, Rosman L, Pelaez D, Footer KHA, et al. The PrEP Care Continuum Among Cisgender Women Who Sell Sex and/or Use Drugs Globally: A Systematic Review. AIDS Behav. 2020; 24(5):1312-1333.
141. Restar AJ, Tocco JU, Mantell JE, Lafort Y, Gichangi P, Masvawure TB, et al. Perspectives on HIV Pre- and Post-Exposure Prophylaxes (PrEP and PEP) Among Female and Male Sex Workers in Mombasa, Kenya: Implications for Integrating Biomedical Prevention into Sexual Health Services. AIDS Educ Prev. 2017; 29(2):141-153.
142. Eakle R, Bothma R, Bourne A, Gumede S, Motsosi K, Rees H. "I am still negative": Female sex workers' perspectives on uptake and use of daily pre-exposure prophylaxis for HIV prevention in South Africa. PLoS One. 2019; 14(4):e0212271.
143. Shea J, Bula A, Dunda W, Hosseinipour MC, Golin CE, Hoffman IF, et al. "The Drug Will Help Protect My Tomorrow": Perceptions of Integrating PrEP into HIV Prevention Behaviors Among Female Sex Workers in Lilongwe, Malawi. AIDS Educ Prev. 2019; 31(5):421-432.

144. Pillay D, Stankevitz K, Lanham M, Ridgeway K, Murire M, Briedenhann E, et al. Factors influencing uptake, continuation, and discontinuation of oral PrEP among clients at sex worker and MSM facilities in South Africa. PLoS One. 2020; 15(4):e0228620.
145. Haberer JE, Kidoguchi L, Heffron R, Mugo N, Bukusi E, Katabira E, et al. Alignment of adherence and risk for HIV acquisition in a demonstration project of pre-exposure prophylaxis among HIV serodiscordant couples in Kenya and Uganda: a prospective analysis of prevention-effective adherence. J Int AIDS Soc. 2017; 20(1):21842.
146. Pyra M, Brown ER, Haberer JE, Heffron R, Celum C, Bukusi EA, et al. Patterns of Oral PrEP Adherence and HIV Risk Among Eastern African Women in HIV Serodiscordant Partnerships. AIDS Behav. 2018; 22(11):3718-3725.
147. World Health Organization. Report on key populations in African HIV/AIDS: National Strategic Plans. 2018. https://www.who.int/hiv/pub/toolkits/key-population-report-afro/en/.
148. World Health Organization. Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. 2015. https://www.who.int/hiv/pub/guidelines/earlyrelease-arv/en/.
149. McClelland RS, Richardson BA, Cherutich P, Mandaliya K, John-Stewart G, Miregwa B, et al. A 15-year study of the impact of community antiretroviral therapy coverage on HIV incidence in Kenyan female sex workers. AIDS. 2015; 29(17):2279-86.
150. Diabate S, Chamberland A, Geraldo N, Tremblay C, Alary M. Gonorrhea, Chlamydia and HIV incidence among female sex workers in Cotonou, Benin: A longitudinal study. PLoS One. 2018; 13(5):e0197251.
151. Bazzi AR, Rangel G, Martinez G, Ulibarri MD, Syvertsen JL, Bazzi SA, et al. Incidence and Predictors of HIV and Sexually Transmitted Infections Among Female Sex Workers and Their Intimate Male Partners in Northern Mexico: A Longitudinal, Multilevel Study. Am J Epidemiol. 2015; 181(9):723-31.
152. Li J, Zhang H, Shen Z, Zhou Y, Fang N, Wang L, et al. Screening for acute HIV infections and estimating HIV incidence among female sex workers from low-grade venues in Guangxi, China. PLoS One. 2014; 9(6):e99522.
153. Sharifi H, Mirzazadeh A, Shokoohi M, Karamouzian M, Khajehkazemi R, Navadeh S, et al. Estimation of HIV incidence and its trend in three key populations in Iran. PLoS One. 2018; 13(11):e0207681.
154. Wang H, Reilly KH, Brown K, Jin X, Xu J, Ding G, et al. HIV incidence and associated risk factors among female sex workers in a high HIV-prevalence area of China. Sex Transm Dis. 2012; 39(11):835-41.
155. Haukoos JS, Lyons MS, Lindsell CJ, Hopkins E, Bender B, Rothman RE, et al. Derivation and validation of the Denver Human Immunodeficiency Virus (HIV) risk score for targeted HIV screening. Am J Epidemiol. 2012; 175(8):838-46.
156. Smith DK, Pals SL, Herbst JH, Shinde S, Carey JW. Development of a clinical screening index predictive of incident HIV infection among men who have sex with men in the United States. J Acquir Immune Defic Syndr. 2012; 60(4):421-7.
157. Menza TW, Hughes JP, Celum CL, Golden MR. Prediction of HIV acquisition among men who have sex with men. Sex Transm Dis. 2009; 36(9):547-55.

158. Hoenigl M, Weibel N, Mehta SR, Anderson CM, Jenks J, Green N, et al. Development and validation of the San Diego Early Test Score to predict acute and early HIV infection risk in men who have sex with men. Clin Infect Dis. 2015; 61(3):468-75.
159. Wahome E, Thiong'o AN, Mwashigadi G, Chirro O, Mohamed K, Gichuru E, et al. An Empiric Risk Score to Guide PrEP Targeting Among MSM in Coastal Kenya. AIDS Behav. 2018; 22(Suppl 1):35-44.
160. Kahle EM, Hughes JP, Lingappa JR, John-Stewart G, Celum C, Nakku-Joloba E, et al. An empiric risk scoring tool for identifying high-risk heterosexual HIV-1-serodiscordant couples for targeted HIV-1 prevention. J Acquir Immune Defic Syndr. 2013; 62(3):339-47.
161. Kagaayi J, Gray RH, Whalen C, Fu P, Neuhauser D, McGrath JW, et al. Indices to measure risk of HIV acquisition in Rakai, Uganda. PLoS One. 2014; 9(4):e92015.
162. Pintye J, Drake AL, Kinuthia J, Unger JA, Matemo D, Heffron RA, et al. A Risk Assessment Tool for Identifying Pregnant and Postpartum Women Who May Benefit From Preexposure Prophylaxis. Clin Infect Dis. 2017; 64(6):751-758.
163. Balkus JE, Brown E, Palanee T, Nair G, Gafoor Z, Zhang J, et al. An Empiric HIV Risk Scoring Tool to Predict HIV-1 Acquisition in African Women. J Acquir Immune Defic Syndr. 2016; 72(3):333-43.
164. Bowring AL, Ketende S, Rao A, Mfochive Njindam I, Decker MR, Lyons C, et al. Characterising unmet HIV prevention and treatment needs among young female sex workers and young men who have sex with men in Cameroon: a cross-sectional analysis. Lancet Child Adolesc Health. 2019; 3(7):482-491.
165. Napierala S, Chabata ST, Fearon E, Davey C, Hargreaves J, Busza J, et al. Engagement in HIV Care Among Young Female Sex Workers in Zimbabwe. J Acquir Immune Defic Syndr. 2018; 79(3):358-366.
166. UN Women. Facts and figures: HIV and AIDS. 2018. https://www.unwomen.org/en/what-we-do/hiv-and-aids/facts-and-figures.
167. Martin HL, Jr., Nyange PM, Richardson BA, Lavreys L, Mandaliya K, Jackson DJ, et al. Hormonal contraception, sexually transmitted diseases, and risk of heterosexual transmission of human immunodeficiency virus type 1. J Infect Dis. 1998; 178(4):1053-9.
168. McClelland RS, Lavreys L, Hassan WM, Mandaliya K, Ndinya-Achola JO, Baeten JM. Vaginal washing and increased risk of HIV-1 acquisition among African women: a 10-year prospective study. AIDS. 2006; 20(2):269-73.
169. Odek WO, Githuka GN, Avery L, Njoroge PK, Kasonde L, Gorgens M, et al. Estimating the Size of the Female Sex Worker Population in Kenya to Inform HIV Prevention Programming. PLoS ONE. 2014; 9(3).
170. Murray L, Moreno L, Rosario S, Ellen J, Sweat M, Kerrigan D. The role of relationship intimacy in consistent condom use among female sex workers and their regular paying partners in the Dominican Republic. AIDS Behav. 2007; 11(3):463-70.
171. Deering KN, Bhattacharjee P, Bradley J, Moses SS, Shannon K, Shaw SY, et al. Condom use within non-commercial partnerships of female sex workers in southern India. BMC Public Health. 2011; 11 Suppl 6(Suppl 6):S11.
172. Voeten HA, Egesah OB, Varkevisser CM, Habbema JD. Female sex workers and unsafe sex in urban and rural Nyanza, Kenya: regular partners may contribute more to HIV transmission than clients. Trop Med Int Health. 2007; 12(2):174-82.

173. Luchters S, Richter ML, Bosire W, Nelson G, Kingola N, Zhang XD, et al. The contribution of emotional partners to sexual risk taking and violence among female sex workers in Mombasa, Kenya: a cohort study. PLoS One. 2013; 8(8):e68855.
174. Goldenberg SM, Gallardo Cruz M, Strathdee SA, Nguyen L, Semple SJ, Patterson TL. Correlates of unprotected sex with female sex workers among male clients in Tijuana, Mexico. Sex Transm Dis. 2010; 37(5):319-24.
175. McClelland L, Wanje G, Kashonga F, Kibe L, McClelland RS, Kiarie J, et al. Understanding the Context of HIV Risk Behavior Among HIV-positive and HIV-negative Female Sex Workers and Male Bar Clients Following Antiretroviral Therapy Rollout in Mombasa, Kenya. AIDS Educ Prev. 2011; 23(4):299-312.
176. Lowndes CM, Alary M, Gnintoungbe CA, Bedard E, Mukenge L, Geraldo N, et al. Management of sexually transmitted diseases and HIV prevention in men at high risk: targeting clients and non-paying sexual partners of female sex workers in Benin. AIDS. 2000; 14(16):2523-34.
177. Cote AM, Sobela F, Dzokoto A, Nzambi K, Asamoah-Adu C, Labbe AC, et al. Transactional sex is the driving force in the dynamics of HIV in Accra, Ghana. AIDS. 2004; 18(6):917-25.
178. Baeten JM, Benki S, Chohan V, Lavreys L, McClelland RS, Mandaliya K, et al. Hormonal contraceptive use, herpes simplex virus infection, and risk of HIV-1 acquisition among Kenyan women. AIDS. 2007; 21(13):1771-7.
179. Lavreys L, Baeten JM, Martin HL, Jr., Overbaugh J, Mandaliya K, Ndinya-Achola J, et al. Hormonal contraception and risk of HIV-1 acquisition: results of a 10-year prospective study. AIDS. 2004; 18(4):695-7.
180. Polis CB, Curtis KM, Hannaford PC, Phillips SJ, Chipato T, Kiarie JN, et al. An updated systematic review of epidemiological evidence on hormonal contraceptive methods and HIV acquisition in women. AIDS. 2016; 30(17):2665-2683.
181. Ahmed K, Baeten JM, Beksinska M, Bekker LG, Bukusi EA, Donnell D, et al. HIV Incidence Among Women Using Intramuscular Depot Medroxyprogesterone Acetate, a Copper Intrauterine Device, or a Levonorgestrel Implant for Contraception: A Randomised, Multicentre, Open-Label Trial. Lancet. 2019; 394(10195):303-13.
182. Masha SC, Cools P, Sanders EJ, Vaneechoutte M, Crucitti T. Trichomonas vaginalis and HIV infection acquisition: a systematic review and meta-analysis. Sex Transm Infect. 2019; 95(1):36-42.
183. Ward H, Ronn M. Contribution of sexually transmitted infections to the sexual transmission of HIV. Curr Opin HIV AIDS. 2010; 5(4):305-10.
184. Fleming DT, Wasserheit JN. From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection. Sex Transm Infect. 1999; 75(1):3-17.
185. Galvin SR, Cohen MS. The role of sexually transmitted diseases in HIV transmission. Nat Rev Microbiol. 2004; 2(1):33-42.
186. World Health Organization. Report on global sexually transmitted infection surveillance. 2018. https://www.who.int/reproductivehealth/publications/stis-surveillance-2018/en/.

187. Boyadzhyan B, Yashina T, Yatabe JH, Patnaik M, Hill CS. Comparison of the APTIMA CT and GC Assays With the APTIMA Combo 2 Assay, the Abbott LCx Assay, and Direct Fluorescent-Antibody and Culture Assays for Detection of Chlamydia Trachomatis and Neisseria Gonorrhoeae. J Clin Microbiol. 2004; 42(7):3089-93.
188. Smith G. Step away from stepwise. Journal of Big Data. 2018; 5.
189. Verguet S, Stalcup M, Walsh JA. Where to Deploy Pre-Exposure Prophylaxis (PrEP) in sub-Saharan Africa? Sex Transm Infect. 2013; 89(8):628-34.
190. McGillen JB, Anderson SJ, Hallett TB. PrEP as a Feature in the Optimal Landscape of Combination HIV Prevention in sub-Saharan Africa. J Int AIDS Soc. 2016; 19(7(Suppl 6)):21104-21104.
191. Gomez GB, Borquez A, Case KK, Wheelock A, Vassall A, Hankins C. The Cost and Impact of Scaling Up Pre-Exposure Prophylaxis for HIV Prevention: A Systematic Review of Cost-Effectiveness Modelling Studies. PLoS Med. 2013; 10(3):e1001401-e1001401.
192. Baeten JM, Richardson BA, Martin JHL, Nyange PM, Lavreys L, Ngugi EN, et al. Trends in HIV-1 Incidence in a Cohort of Prostitutes in Kenya: Implications for HIV-1 Vaccine Efficacy Trials. J Acquir Immune Defic Syndr. 2000; 24(5):458-64.
193. Sabo MC, Richardson BA, Lavreys L, Martin JHL, Jaoko W, Mandaliya K, et al. Does Bacterial Vaginosis Modify the Effect of Hormonal Contraception on HIV Seroconversion. AIDS. 2019; 33(7):1225-30.
194. Masese L, Baeten JM, Richardson BA, Bukusi E, John-Stewart G, Jaoko W, et al. Incident Herpes Simplex Virus Type 2 Infection Increases the Risk of Subsequent Episodes of Bacterial Vaginosis. J Infect Dis. 2014; 209(7):1023-7.
195. Lavreys L, Baeten JM, Kreiss JK, Richardson BA, Chohan BH, Hassan W, et al. Injectable Contraceptive Use and Genital Ulcer Disease During the Early Phase of HIV-1 Infection Increase Plasma Virus Load in Women. J Infect Dis. 2004; 189(2):303-11.
196. Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J Clin Microbiol. 1991; 29(2):297-301.
197. Richardson BA, Mbori-Ngacha D, Lavreys L, John-Stewart GC, Nduati R, Panteleeff DD, et al. Comparison of human immunodeficiency virus type 1 viral loads in Kenyan women, men, and infants during primary and early infection. J Virol. 2003; 77(12):7120-3.
198. Busch MP, Lee LL, Satten GA, Henrard DR, Farzadegan H, Nelson KE, et al. Time course of detection of viral and serologic markers preceding human immunodeficiency virus type 1 seroconversion: implications for screening of blood and tissue donors. Transfusion. 1995; 35(2):91-7.
199. Penn State Eberly College of Science. STAT 501—regression methods—stepwise regression. https://online.stat.psu.edu/stat501/lesson/10/10.2.
200. Dean AG, Sullivan KM, Soe MM. Open Epi: Open Source Epidemiologic Statistics for Public Health, Version 3.01 www.openepi.com.
201. Corbett KS, Flynn B, Foulds KE, Francica JR, Boyoglu-Barnum S, Werner AP, et al. Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates. N Engl J Med. 2020; 383(16):1544-1555.
202. Vogel AB, Kanevsky I, Che Y, Swanson KA, Muik A, Vormehr M, et al. BNT162b vaccines protect rhesus macaques from SARS-CoV-2. Nature. 2021; 592(7853):283-289.

203. Mercado NB, Zahn R, Wegmann F, Loos C, Chandrashekar A, Yu J, et al. Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. Nature. 2020; 586(7830):583-588.
204. van Doremalen N, Lambe T, Spencer A, Belij-Rammerstorfer S, Purushotham JN, Port JR, et al. ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques. Nature. 2020; 586(7830):578-582.
205. Gao Q, Bao L, Mao H, Wang L, Xu K, Yang M, et al. Development of an inactivated vaccine candidate for SARS-CoV-2. Science. 2020; 369(6499):77-81.
206. Yu J, Tostanoski LH, Peter L, Mercado NB, McMahan K, Mahrokhian SH, et al. DNA vaccine protection against SARS-CoV-2 in rhesus macaques. Science. 2020; 369(6505):806-811.
207. Yang J, Wang W, Chen Z, Lu S, Yang F, Bi Z, et al. A vaccine targeting the RBD of the S protein of SARS-CoV-2 induces protective immunity. Nature. 2020; 586(7830):572-577.
208. Wang H, Zhang Y, Huang B, Deng W, Quan Y, Wang W, et al. Development of an Inactivated Vaccine Candidate, BBIBP-CorV, with Potent Protection against SARS-CoV-2. Cell. 2020; 182(3):713-721.e9.
209. Feng L, Wang Q, Shan C, Yang C, Feng Y, Wu J, et al. An adenovirus-vectored COVID-19 vaccine confers protection from SARS-COV-2 challenge in rhesus macaques. Nat Commun. 2020; 11:1-11.
210. Ma X, Zou F, Yu F, Li R, Yuan Y, Zhang Y, et al. Nanoparticle Vaccines Based on the Receptor Binding Domain (RBD) and Heptad Repeat (HR) of SARS-CoV-2 Elicit Robust Protective Immune Responses. Immunity. 2020; 53(6):1315-1330.e9.
211. Sui Y, Li J, Zhang R, Prabhu SK, Andersen H, Venzon D, et al. Protection against SARS-CoV-2 infection by a mucosal vaccine in rhesus macaques. JCI Insight. 2021; 6(10):e148494.
212. Harris PE, Brasel T, Massey C, Herst CV, Burkholz S, Lloyd P, et al. A Synthetic Peptide CTL Vaccine Targeting Nucleocapsid Confers Protection from SARS-CoV-2 Challenge in Rhesus Macaques. Vaccines (Basel). 2021; 9(5):520.
213. Yadav PD, Ella R, Kumar S, Patil DR, Mohandas S, Shete AM, et al. Immunogenicity and protective efficacy of inactivated SARS-CoV-2 vaccine candidate, BBV152 in rhesus macaques. Nat Commun. 2021; 12(1):1386.
214. Garrido C, Curtis AD, 2nd, Dennis M, Pathak SH, Gao H, Montefiori D, et al. SARS-CoV-2 vaccines elicit durable immune responses in infant rhesus macaques. Sci Immunol. 2021; 6(60):eabj3684.
215. Routhu NK, Cheedarla N, Gangadhara S, Bollimpelli VS, Boddapati AK, Shiferaw A, et al. A modified vaccinia Ankara vector-based vaccine protects macaques from SARS-CoV-2 infection, immune pathology, and dysfunction in the lungs. Immunity. 2021; 54(3):542-556.e9.
216. Li H, Guo L, Zheng H, Li J, Zhao X, Li J, et al. Self-Assembling Nanoparticle Vaccines Displaying the Receptor Binding Domain of SARS-CoV-2 Elicit Robust Protective Immune Responses in Rhesus Monkeys. Bioconjug Chem. 2021; 32(5):1034-1046.
217. Li Y, Bi Y, Xiao H, Yao Y, Liu X, Hu Z, et al. A novel DNA and protein combination COVID-19 vaccine formulation provides full protection against SARS-CoV-2 in rhesus macaques. Emerg Microbes Infect. 2021; 10(1):342-355.

218. Arunachalam PS, Walls AC, Golden N, Atyeo C, Fischinger S, Li C, et al. Adjuvanting a subunit COVID-19 vaccine to induce protective immunity. Nature. 2021; 594(7862):253-258.
219. Liang JG, Su D, Song TZ, Zeng Y, Huang W, Wu J, et al. S-Trimer, a COVID-19 subunit vaccine candidate, induces protective immunity in nonhuman primates. Nat Commun. 2021; 12(1):1346.
220. Luo S, Zhang P, Liu B, Yang C, Liang C, Wang Q, et al. Prime-boost vaccination of mice and rhesus macaques with two novel adenovirus vectored COVID-19 vaccine candidates. Emerg Microbes Infect. 2021; 10(1):1002-1015.
221. Solforosi L, Kuipers H, Jongeneelen M, Rosendahl Huber SK, van der Lubbe JEM, Dekking L, et al. Immunogenicity and efficacy of one and two doses of Ad26.COV2.S COVID vaccine in adult and aged NHP. J Exp Med. 2021; 218(7):e20202756.
222. Walls AC, Miranda MC, Schafer A, Pham MN, Greaney A, Arunachalam PS, et al. Elicitation of broadly protective sarbecovirus immunity by receptor-binding domain nanoparticle vaccines. Cell. 2021; 184(21):5432-5447.e16.
223. King HAD, Joyce MG, Lakhal-Naouar I, Ahmed A, Cincotta CM, Subra C, et al. Efficacy and breadth of adjuvanted SARS-CoV-2 receptor-binding domain nanoparticle vaccine in macaques. Proc Natl Acad Sci U S A. 2021; 118(38):e2106433118.
224. Guebre-Xabier M, Patel N, Tian JH, Zhou B, Maciejewski S, Lam K, et al. NVX-CoV2373 vaccine protects cynomolgus macaque upper and lower airways against SARS-CoV-2 challenge. Vaccine. 2020; 38(50):7892-7896.
225. Sanchez-Felipe L, Vercruysse T, Sharma S, Ma J, Lemmens V, Van Looveren D, et al. A single-dose live-attenuated YF17D-vectored SARS-CoV-2 vaccine candidate. Nature. 2021; 590(7845):320-325.
226. Zhang NN, Li XF, Deng YQ, Zhao H, Huang YJ, Yang G, et al. A Thermostable mRNA Vaccine against COVID-19. Cell. 2020; 182(5):1271-1283.e16.
227. Li T, Zheng Q, Yu H, Wu D, Xue W, Xiong H, et al. SARS-CoV-2 spike produced in insect cells elicits high neutralization titres in non-human primates. Emerg Microbes Infect. 2020; 9(1):2076-2090.
228. Brouwer PJM, Brinkkemper M, Maisonnasse P, Dereuddre-Bosquet N, Grobben M, Claireaux M, et al. Two-component spike nanoparticle vaccine protects macaques from SARS-CoV-2 infection. Cell. 2021; 184(5):1188-1200.e19.
229. Hong SH, Oh H, Park YW, Kwak HW, Oh EY, Park HJ, et al. Immunization with RBD-P2 and N protects against SARS-CoV-2 in nonhuman primates. Sci Adv. 2021; 7(22):eabg7156.
230. Sun S, He L, Zhao Z, Gu H, Fang X, Wang T, et al. Recombinant vaccine containing an RBD-Fc fusion induced protection against SARS-CoV-2 in nonhuman primates and mice. Cell Mol Immunol. 2021; 18(4):1070-1073.
231. Capone S, Raggioli A, Gentile M, Battella S, Lahm A, Sommella A, et al. Immunogenicity of a new gorilla adenovirus vaccine candidate for COVID-19. Mol Ther. 2021; 29(8):2412-2423.
232. Kalnin KV, Plitnik T, Kishko M, Zhang J, Zhang D, Beauvais A, et al. Immunogenicity and efficacy of mRNA COVID-19 vaccine MRT5500 in preclinical animal models. NPJ Vaccines. 2021; 6(1):61.

233. Walls AC, Fiala B, Schafer A, Wrenn S, Pham MN, Murphy M, et al. Elicitation of Potent Neutralizing Antibody Responses by Designed Protein Nanoparticle Vaccines for SARS-CoV-2. Cell. 2020; 183(5):1367-1382.e17.
234. Tan HX, Juno JA, Lee WS, Barber-Axthelm I, Kelly HG, Wragg KM, et al. Immunogenicity of prime-boost protein subunit vaccine strategies against SARS-CoV-2 in mice and macaques. Nat Commun. 2021; 12(1):1403.
235. Erasmus JH, Khandhar AP, O'Connor MA, Walls AC, Hemann EA, Murapa P, et al. An Alphavirus-derived replicon RNA vaccine induces SARS-CoV-2 neutralizing antibody and T cell responses in mice and nonhuman primates. Sci Transl Med. 2020; 12(555):eabc9396.
236. Hewitt JA, Lutz C, Florence WC, Pitt MLM, Rao S, Rappaport J, et al. ACTIVating Resources for the COVID-19 Pandemic: In Vivo Models for Vaccines and Therapeutics. Cell Host Microbe. 2020; 28(5):646-659.
237. Deng W, Bao L, Liu J, Xiao C, Liu J, Xue J, et al. Primary exposure to SARS-CoV-2 protects against reinfection in rhesus macaques. Science. 2020; 369(6505):818-823.
238. Chandrashekar A, Liu J, Martinot AJ, McMahan K, Mercado NB, Peter L, et al. SARS-CoV-2 infection protects against rechallenge in rhesus macaques. Science. 2020; 369(6505):812-817.
239. Hasenkrug KJ, Feldmann F, Myers L, Santiago ML, Guo K, Barrett BS, et al. Recovery from Acute SARS-CoV-2 Infection and Development of Anamnestic Immune Responses in T Cell-Depleted Rhesus Macaques. mBio. 2021; 12(4):e0150321.
240. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell. 2020; 181(2):281-292.e6.
241. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020; 367(6483):1260-1263.
242. Li F. Structure, Function, and Evolution of Coronavirus Spike Proteins. Annu Rev Virol. 2016; 3(1):237-261.
243. Fan X, Cao D, Kong L, Zhang X. Cryo-EM analysis of the post-fusion structure of the SARS-CoV spike glycoprotein. Nat Commun. 2020; 11(1):3618.
244. Khoury DS, Cromer D, Reynaldi A, Schlub TE, Wheatley AK, Juno JA, et al. Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. Nat Med. 2021; 27(7):1205-1211.
245. Lumley SF, O'Donnell D, Stoesser NE, Matthews PC, Howarth A, Hatch SB, et al. Antibody Status and Incidence of SARS-CoV-2 Infection in Health Care Workers. N Engl J Med. 2021; 384(6):533-540.
246. Corbett KS, Nason MC, Flach B, Gagne M, O' Connell S, Johnston TS, et al. Immune correlates of protection by mRNA-1273 vaccine against SARS-CoV-2 in nonhuman primates. Science. 2021; 373(6561):eabj0299.
247. Gilbert PB, Montefiori DC, McDermott A, Fong Y, Benkeser DC, Deng W, et al. Immune correlates analysis of the mRNA-1273 COVID-19 vaccine efficacy clinical trial. Science. 2022; 375(6576):43-50.
248. Rogers TF, Zhao F, Huang D, Beutler N, Burns A, He WT, et al. Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. Science. 2020; 369(6506):956-963.

249. Voss WN, Hou YJ, Johnson NV, Delidakis G, Kim JE, Javanmardi K, et al. Prevalent, protective, and convergent IgG recognition of SARS-CoV-2 non-RBD spike epitopes. Science. 2021; 372(6546):1108-1112.
250. Greaney AJ, Loes AN, Crawford KHD, Starr TN, Malone KD, Chu HY, et al. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. Cell Host Microbe. 2021; 29(3):463-476.e6.
251. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol. 2021; 19(7):409-424.
252. Tauzin A, Nayrac M, Benlarbi M, Gong SY, Gasser R, Beaudoin-Bussieres G, et al. A single dose of the SARS-CoV-2 vaccine BNT162b2 elicits Fc-mediated antibody effector functions and T cell responses. Cell Host Microbe. 2021; 29(7):1137-1150.e6.
253. Brunet-Ratnasingham E, Anand SP, Gantner P, Dyachenko A, Moquin-Beaudry G, Brassard N, et al. Integrated immunovirological profiling validates plasma SARS-CoV-2 RNA as an early predictor of COVID-19 mortality. Science Advances. 2021; 7(48):eabj5629.
254. Beaudoin-Bussières G, Chen Y, Ullah I, Prévost J, Tolbert WD, Symmes K, et al. A Fc-enhanced NTD-binding non-neutralizing antibody delays virus spread and synergizes with a nAb to protect mice from lethal SARS-CoV-2 infection. Cell Reports. 2022; 38(7):110368.
255. Garrett ME, Itell HL, Crawford KHD, Basom R, Bloom JD, Overbaugh J. Phage-DMS: A Comprehensive Method for Fine Mapping of Antibody Epitopes. iScience. 2020; 23(10):101622.
256. Garrett ME, Galloway J, Chu HY, Itell HL, Stoddard CI, Wolf CR, et al. High-resolution profiling of pathways of escape for SARS-CoV-2 spike-binding antibodies. Cell. 2021; 184(11):2927-2938.e11.
257. Garrett ME, Galloway JG, Wolf C, Logue JK, Franko N, Chu HY, et al. Comprehensive characterization of the antibody responses to SARS-CoV-2 Spike protein after infection and/or vaccination. Elife. 2022:e73490.
258. Erasmus JH, Khandhar AP, Guderian J, Granger B, Archer J, Archer M, et al. A Nanostructured Lipid Carrier for Delivery of a Replicating Viral RNA Provides Single, Low-Dose Protection against Zika. Mol Ther. 2018; 26(10):2507-2522.
259. Yuan M, Liu H, Wu NC, Wilson IA. Recognition of the SARS-CoV-2 receptor binding domain by neutralizing antibodies. Biochem Biophys Res Commun. 2021; 538:192-203.
260. Niu L, Wittrock KN, Clabaugh GC, Srivastava V, Cho MW. A Structural Landscape of Neutralizing Antibodies Against SARS-CoV-2 Receptor Binding Domain. Front Immunol. 2021; 12:647934.
261. Chi X, Yan R, Zhang J, Zhang G, Zhang Y, Hao M, et al. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. Science. 2020; 369(6504):650-655.
262. Li D, Edwards RJ, Manne K, Martinez DR, Schafer A, Alam SM, et al. In vitro and in vivo functions of SARS-CoV-2 infection-enhancing and neutralizing antibodies. Cell. 2021; 184(16):4203-4219.e32.
263. Liu L, Wang P, Nair MS, Yu J, Rapp M, Wang Q, et al. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. Nature. 2020; 584(7821):450-456.

264. Wang N, Sun Y, Feng R, Wang Y, Guo Y, Zhang L, et al. Structure-based development of human antibody cocktails against SARS-CoV-2. Cell Res. 2021; 31(1):101-103.
265. Cerutti G, Guo Y, Zhou T, Gorman J, Lee M, Rapp M, et al. Potent SARS-CoV-2 neutralizing antibodies directed against spike N-terminal domain target a single supersite. Cell Host Microbe. 2021; 29(5):819-833.e7.
266. McCallum M, De Marco A, Lempp FA, Tortorici MA, Pinto D, Walls AC, et al. N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. Cell. 2021; 184(9):2332-2347.e16.
267. Suryadevara N, Shrihari S, Gilchuk P, VanBlargan LA, Binshtein E, Zost SJ, et al. Neutralizing and protective human monoclonal antibodies recognizing the N-terminal domain of the SARS-CoV-2 spike protein. Cell. 2021; 184(9):2316-2331.e15.
268. Liu Y, Soh WT, Kishikawa JI, Hirose M, Nakayama EE, Li S, et al. An infectivity-enhancing site on the SARS-CoV-2 spike protein targeted by antibodies. Cell. 2021; 184(13):3452-3466.e18.
269. Li Y, Lai DY, Zhang HN, Jiang HW, Tian X, Ma ML, et al. Linear epitopes of SARS-CoV-2 spike protein elicit neutralizing antibodies in COVID-19 patients. Cell Mol Immunol. 2020; 17(10):1095-1097.
270. Poh CM, Carissimo G, Wang B, Amrun SN, Lee CY, Chee RS, et al. Two linear epitopes on the SARS-CoV-2 spike protein that elicit neutralising antibodies in COVID-19 patients. Nat Commun. 2020; 11(1):2806.
271. Tang T, Bidon M, Jaimes JA, Whittaker GR, Daniel S. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. Antiviral Res. 2020; 178:104792.
272. Madu IG, Roth SL, Belouzard S, Whittaker GR. Characterization of a highly conserved domain within the severe acute respiratory syndrome coronavirus spike protein S2 domain with characteristics of a viral fusion peptide. J Virol. 2009; 83(15):7411-7421.
273. Zhou P, Yuan M, Song G, Beutler N, Shaabani N, Huang D, et al. A human antibody reveals a conserved site on beta-coronavirus spike proteins and confers protection against SARS-CoV-2 infection. Science Translational Medicine. 2022.
274. Pinto D, Sauer MM, Czudnochowski N, Low JS, Tortorici MA, Housley MP, et al. Broad betacoronavirus neutralization by a stem helix-specific human antibody. Science. 2021; 373(6559):1109-1116.
275. Li W, Chen Y, Prévost J, Ullah I, Lu M, Gong SY, et al. Structural Basis and Mode of Action for Two Broadly Neutralizing Antibodies Against SARS-CoV-2 Emerging Variants of Concern. Cell Reports. 2022; 38(2):110210.
276. Hsieh CL, Werner AP, Leist SR, Stevens LJ, Falconer E, Goldsmith JA, et al. Stabilized coronavirus spike stem elicits a broadly protective antibody. Cell Rep. 2021; 37(5):109929.
277. Zohar T, Alter G. Dissecting antibody-mediated protection against SARS-CoV-2. Nat Rev Immunol. 2020; 20(7):392-394.
278. Lee WS, Wheatley AK, Kent SJ, DeKosky BJ. Antibody-dependent enhancement and SARS-CoV-2 vaccines and therapies. Nat Microbiol. 2020; 5(10):1185-1191.

279. Mikocziova I, Greiff V, Sollid LM. Immunoglobulin germline gene variation and its impact on human disease. Genes Immun. 2021; 22(4):205-217.
280. Ramesh A, Darko S, Hua A, Overman G, Ransier A, Francica JR, et al. Structure and Diversity of the Rhesus Macaque Immunoglobulin Loci through Multiple De Novo Genome Assemblies. Front Immunol. 2017; 8:1407.
281. Robbiani DF, Gaebler C, Muecksch F, Lorenzi JCC, Wang Z, Cho A, et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. Nature. 2020; 584(7821):437-442.
282. Chen EC, Gilchuk P, Zost SJ, Suryadevara N, Winkler ES, Cabel CR, et al. Convergent antibody responses to the SARS-CoV-2 spike protein in convalescent and vaccinated individuals. Cell Rep. 2021; 36(8):109604.
283. Nielsen SCA, Yang F, Jackson KJL, Hoh RA, Roltgen K, Jean GH, et al. Human B Cell Clonal Expansion and Convergent Antibody Responses to SARS-CoV-2. Cell Host Microbe. 2020; 28(4):516-525.e5.
284. Greaney AJ, Loes AN, Gentles LE, Crawford KHD, Starr TN, Malone KD, et al. Antibodies elicited by mRNA-1273 vaccination bind more broadly to the receptor binding domain than do those from SARS-CoV-2 infection. Sci Transl Med. 2021; 13(600):eabi9915.
285. Amanat F, Thapa M, Lei T, Ahmed SMS, Adelsberg DC, Carreno JM, et al. SARS-CoV-2 mRNA vaccination induces functionally diverse antibodies to NTD, RBD, and S2. Cell. 2021; 184(15):3936-3948.e10.
286. Crowley AR, Ackerman ME. Mind the Gap: How Interspecies Variability in IgG and Its Receptors May Complicate Comparisons of Human and Non-human Primate Effector Function. Front Immunol. 2019; 10:697.
287. Jackson LA, Anderson EJ, Rouphael NG, Roberts PC, Makhene M, Coler RN, et al. An mRNA Vaccine against SARS-CoV-2 - Preliminary Report. N Engl J Med. 2020; 383(20):1920-1931.
288. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020; 579(7798):265-269.
289. Monge G. Mémoire sur la théorie des déblais et des remblais. Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année. 1781:666-704.
290. Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci U S A. 1992; 89(22):10915-10919.
291. Cai Y, Zhang J, Xiao T, Peng H, Sterling SM, Walsh RM, Jr., et al. Distinct conformational states of SARS-CoV-2 spike protein. Science. 2020; 369(6511):1586-1592.
292. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. 2021; 30(1):70-82.
293. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; 9(4):357-359.
294. Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. Nextflow enables reproducible computational workflows. Nat Biotechnol. 2017; 35(4):316-319.

295. The pandas development team. pandas-dev/pandas: Pandas. 2020. doi:10.5281/zenodo.4067057.
296. McKinney W. Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference. 2010; 445:56-61.
297. Hoyer S, Hamman J. xarray: N-D labeled Arrays and Datasets in Python. Journal of Open Research Software. 2017; 5(1):10.
298. Flamary R, Courty N, Gramfort A, Alaya MZ, Boisbunon A, Chambon S, et al. POT: Python Optimal Transport. JMLR. 2021; 22(78):1-8.
299. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009; 25(11):1422-1423.