THE UNIVERSITY *of* EDINBURGH

Does Positive Selection of Sialic Acid Regulatory Genes Affect Bats' Influenza A Virus Carriage?

Eilidh Campbell

S2089210

Infectious Diseases Master's by Research (MScR)

University of Edinburgh

2025

Declaration

I declare that this dissertation is my own work and has not been submitted for any other degree or professional qualification. Unless otherwise stated, all results presented are from my own work.

Signed: Eilidh Campbell, 30th of August 2025.

Acknowledgements

Firstly, I would like to thank Dr Richard Sloan for his incredible guidance and support, not only throughout the past year, but since undertaking my undergraduate research project in his lab. I would also like to thank my lab colleagues, Sole, Nanne, Anjali, Jordan and Erin for providing invaluable experimental advice that made the work in this project possible. Finally, I need to thank my friends and family, specifically my mum, dad and Oliver, for all their encouragement and for spending the past year listening to all my scientific dilemmas without growing tired of me.

Abstract

In the past two decades, four novel influenza A virus (IAV) strains have been identified in bats. Given bat species' established role in viral emergence and cross-species transmission, a deeper understanding of host factors affecting their IAV carriage is required to prevent future spillover events. B4GALNT2, a host cell factor known to inhibit avian IAV entry by preventing α2,3 sialic acid receptor-mediated binding has been identified as under positive selection in bats. Evidence that bat B4GALNT2 is subject to positive selection suggests bat species may be evolving their sialic acid availability to regulate IAV infection. Here, we ask if bat B4GALNT2 alleles can regulate IAV H5 infection and if alleles subject to positive selection demonstrate enhanced functionality. Our work shows the first instance of non-human B4GALNT2-mediated antiviral activity. *R. aegyptiacus* B4GALNT2 alleles demonstrate the highest level of inhibition, yet considerable variations in B4GALNT2-mediated antiviral exist between splice variants. Conversely, despite displaying signatures of positive selection, *R. trifoliatus* B4GALNT2 expression does not appear to influence IAV infection. Moreover, our analysis identifies several residues that may be important for bat B4GALNT2 antiviral functionality. The results generated throughout this study will enable further research into the antiviral role of B4GALNT2 and the differential patterns of inhibition observed between bat species' B4GALNT2 genes. By increasing comparative analysis between the genetic factors underlying zoonosis, we will improve our understanding of cross-species transmission and leverage these findings to prevent future viral emergence.

Lay summary

Influenza A virus (IAV), commonly known as the flu, is a respiratory virus that is known to infect a wide range of different animals, including humans, birds, cattle, pigs, horses and bats. Bat species are known to carry various viruses, many of which have spread into opposing animal and human populations, such as Ebola, Rabies and COVID-19. By understanding how bat species immune systems have evolved to withstand viral infection, and applying these findings to humans, we may be able to prevent the spread of further viruses, such as IAV, into human populations. B4GALNT2 is a human protein that can prevent infection of host cells with IAVs by altering the expression of the host cell receptors required for viral entry. B4GALNT2 has been reported to be evolving rapidly in Trefoil horseshoe bats, suggesting this gene might have enhanced antiviral functionality. During this study, we tested a range of bat B4GALNT2 genes, including B4GALNT2 from Trefoil horseshoe bats, against an artificial avian flu virus. Our results demonstrate that some bat B4GALNT2 proteins prevent IAV infection more effectively than human B4GALNT2. However, despite signs of evolution in B4GALNT2 from Trefoil horseshoe bats, its expression appeared to have no effect on virus entry. Overall, the results gathered in this study will help further our understanding of how bat immune systems can effectively manage viral infection and contribute towards preventing the future spread of bat-origin viruses.

Contents

Figures and Tables

Introduction

Zoonotic Disease

60% of all infectious diseases and 75% of emerging diseases are zoonotic in origin, with many of the world's most well-known diseases, including HIV, Influenza, Bovine tuberculosis, and SARS-CoV-2, emerging from cross-species spillover events (Di Bari *et al.*, 2023; Rahman *et al.*, 2020; WHO, 2020b). Viruses are the most well represented of all zoonotic pathogens, causing almost all global disease pandemics since the early 19th century (Singh *et al.*, 2022; Sampath *et al.*, 2021).

Viruses are embedded in the global environment, found even in the most extreme habitats, such as polar regions, hot springs and hypersaline water (Hou *et al.*, 2024; Abbasi & Alam, 2025; Yau *et al.*, 2019). Given the continual contact between all living organisms and viral particles, the emergence and spread of novel viruses should be anticipated. In 2022, an AI-based metagenomic study identified more than 160,000 novel putative RNA virus species from across a range of diverse ecosystems, the majority with unknown host species (Hou *et al.*, 2024). RNA viruses are more likely to spill over into opposing species than DNA viruses due to their high mutation and genetic reassortment rates, and ability to replicate in the cellular cytoplasm (Tomori & Oluwayelu, 2023; Pauciullo *et al.*, 2024). The continual discovery of novel viruses, many with zoonotic potential necessitates further understanding of the molecular mechanisms underpinning cross-species transmission in order to prevent future viral outbreaks and possible pandemics.

Bats as a unique zoonotic species

Bats were first established as important viral reservoirs in the early 20th century after the isolation of rabies virus from common vampire bats in Trinidad (Pawan, 1936). Since this first association, bats have been implicated in the emergence and spread of numerous viruses from across several families, including filoviruses, coronaviruses, and paramyxoviruses (Letko *et al.*, 2020). Nevertheless, the most severe outcome of bat viral carriage was observed less than a decade ago, when SARS-CoV-2 (COVID-19) emerged into human populations.To date, COVID-19 has claimed more than 7.1 million lives and is estimated to have cost the global economy upwards of $16 trillion (WHO, 2025; Cutler & Summer, 2020).

Bats are found across every continent except Antarctica and are among the most diverse mammalian species, with upwards of 1400 species known to exist (Pennisi, 2020; Teeling *et al.*, 2018). Bats are the only species capable of true flight, permitting some species to travel over 800 miles during seasonal migration (Teeling *et al.*, 2018; Gonzalez & Banerjee, 2022). Their wide geographic distribution, coupled with the use of powered flight, enables bats to encounter a diverse range of animals, including a large number of other bat species (Gonzalez & Banerjee, 2022). For instance, various bat species across the *Rhinolophus* genus often share the same roosting cave. Coronavirus sequences isolated from opposing *Rhinolophus* species demonstrate evidence of interspecies transmission, suggesting that during roosting, viruses may be exchanged between species (Cui *et al.*, 2007). However, contact between bats and other mammalian species is increasing with global warming and anthropogenic deforestation (Becker *et al.*, 2022; Wilson *et al.*, 2002; Chua *et al.*, 2002). The emergence of Nipah virus from Malaysian fruit bats is thought to be driven by mass deforestation which occurred less than a year prior to the outbreak. Bats were likely forced into closer proximity with human and animal populations in search of alternative food sources, creating opportunities conducive to viral spillover events, and ultimately resulting in the transmission of Nipah virus to novel host species (Chua *et al.*, 2002). Moreover, periods of nutritional stress have been associated with increased viral virulence and emergence of novel viral infections (Beck *et al.*, 2004; Honce & Schultz-Cherry, 2020). The nutritional stress associated with global warming and deforestation is demonstrated to increase Hendra virus

seroprevalence in *Pteropus* bat species, promoting viral shedding and consequently viral transmission (Plowright *et al*., 2008).

Despite an established role in viral emergence, bat species can effectively tolerate viral infection, rarely displaying symptoms of disease. This unique suppression of viral pathogenicity is owed to a range of metabolic, physiological and immunological adaptations. Over 50% of bats' daily energy expenditure is dedicated to maintaining flight. As a consequence, many genes involved in energy metabolism are subject to positive selection, including reactive oxygen species (ROS) (Shen *et al*., 2010). ROS expression is known to induce cellular damage resulting in the cytosolic release of self-DNA. To avoid consistent activation of innate inflammatory pathways, bats have employed several counter strategies that appear to indirectly influence their viral carriage (Xie *et al*., 2018). STING is a signalling protein that initiates various intracellular innate DNA sensing pathways. In bats, a key serine residue is substituted, preventing STING-mediated interferon activation (Xie *et al*., 2018). Similarly, the inflammasome sensor, NLRP3 is attenuated in bats, inhibiting inflammasome formation (Ahn *et al*., 2019). Such modifications suppress the harmful inflammatory responses associated with viral infection, enabling bat species to effectively manage infection (Xie *et al*., 2018; Ahn *et al*., 2019). In the absence of IFN signalling, RNAi pathways appear to be enhanced in bats (Dai *et al*., 2024). Previously, RNAi was only thought to attenuate viral RNA in human embryonic or stem cells prior to the development of the type I interferon (IFN) response (Li *et al*., 2013; Poirier *et al*., 2021). However, cleavage of dsRNA by bat RNAi can activate an alternative interferon-independent antiviral pathway that appears to reduce RNA virus infection. Evidence of an improved RNAi pathway in bats indicates their reservoir status may have driven the development of IFN-independent antiviral mechanisms (Dai *et al*., 2024).

Although bat viral carriage is not the exclusive driver of immune selection, systematically identifying instances of adaptive selection at both the species and molecular levels may better provide insight into the mechanisms governing bats' viral tolerance and potential spillover events. The Red Queen hypothesis developed by Van Valen in the early 1970s suggests that species must be subject to constant evolution in order to keep up with the selection pressures imposed by other organisms (Van Valen, 1973). In host-pathogen interactions, this is reflected in the continual adaptation of both host defence mechanisms and pathogen virulence. Known coronavirus reservoirs, Rhinolophid and Hipposiderid bats, have been shown to exhibit significant evolutionary changes in ISG15. ISG15 is an interferon-stimulated gene (ISG) that is upregulated in response to several viral infections (Morales *et al*., 2025; Perng & Lenschow, 2018). In severe human SARS-CoV-2 infection, ISG15 can promote hyperinflammation, resulting in poorer infection outcomes (Morales *et al*., 2025; Sarkar *et al*., 2023). However, in Rhinolophid and Hipposiderid bats, substitution of several residues have been demonstrated to increase ISG15-mediated viral-host binding, promoting an antiviral phenotype (Morales *et al*., 2025). Similarly, positive selection of four distinct residues in bat ASC2 has been found to drive its enhanced antiviral functionality. Inducing these four amino acid substitutions into human ASC2 reduces inflammasome formation to levels associated with the expression of bat ASC2. Notably, expression of bat ASC2 in mouse models improves influenza A infection outcomes by 50% (Ahn *et al*., 2023).

Bat-origin Influenza Viruses

Influenza A

Influenza A viruses are a group of negative-sense RNA viruses found across a diverse range of species, including humans, cattle, swine, poultry and bats (Russell *et al.,* 2021; Sarker *et al*., 2022; Medina *et al*., 2011). Influenza A virus (IAV) strains are categorised by their surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA), which are known are known to bind host sialic acid receptors to mediate cellular entry and exit. Consequently, HA-receptor specificity determines IAV host range (Russell *et al*., 2021; Medina *et al*., 2011; Shao *et al*., 2017). For instance, avian-adapted IAV strains preferentially bind to α2,3 sialic acid-linkages present on host cells, whereas human-adapted IAV strains bind to α2,6 sialic acid-linkages (Russell, 2021). However, IAV possesses a segmented RNA genome that allows internal viral segments to be exchanged between avian- and mammalian-origin IAV strains (Shao *et al*., 2017). Such reassortment events typically occur in intermediatory host species, such as swine, that simultaneously express both α2,3- and α2,6- sialic acid receptors (Medina *et al*., 2011; Skehel & Wiley, 2000). IAV genetic reassortment often results in the rapid emergence of novel IAV strains with the potential to infect new host species, such as humans and poultry (Shao *et al*., 2017).

Emergence of High-Pathogenicity Avian Influenza (HPAI)

In 2024, highly pathogenic avian influenza (HPAI) H5N1 clade 2.3.4.4b was first detected in American dairy cattle. Since then, clade 2.3.4.4b has been found in various other mammalian species, such as dogs, pigs, and humans (CDC, 2025; USDA, 2025; Caserta *et al*., 2024). Presently, H5N1 is unable to preferentially bind α2,6-linked sialic acids, preventing onwards transmission of the virus in human populations. Nevertheless, evidence of H5N1 in intermediary host species, such as pigs, indicates genetic reassortment may occur at any time to facilitate emergence of a α2,6-binding strain with the potential to cause wide-scale outbreaks or even a global pandemic (Charostad *et al*., 2023).

Ferrets share the same distribution of sialic acid receptors throughout their upper and lower respiratory tracts as humans, making them an attractive model for IAV research (Basu Thakur *et al*., 2024). Consequently, the ability of IAV strains to spread amongst ferret populations is often used as a determinant for influenza pandemic risk (El-Shesheny *et al*., 2024; WHO, 2020a). Comparative analysis between clade 2.3.4.4b viruses isolated prior to and following the 2024 cattle outbreak have demonstrated that a newly isolated human strain can be transmitted between ferret populations despite maintaining a preference for α2,3-mediated binding (Pulit-Penaloza *et al*., 2024). Moreover, bovine clade 2.3.4.4b can be transmitted to cats through consumption of infected raw milk samples (Burrough *et al*., 2024). Under laboratory conditions, opposing H5N1 strains have been demonstrated to survive for up to seven days in raw milk samples indicating clade 2.3.4.4b H5N1 could be transmitted to human populations through indirect transmission routes, such as consumption of untreated dairy products (Schafers *et al*., 2025). Evidence that H5N1 clade 2.3.4.4b has quickly established various infection routes in several mammalian species highlights the requirement for continued and improved surveillance of emerging IAV strains to prevent further cross-species spillover events.

H17N10, H18N11 and H18N12

Several South American bat species (*Sturnira lilium*, *Artibeus planirostris, Artibeus obscurus* and *Artibeus lituratus)* have been identified as carriers of two novel IAV strains, namely H17N10 and H18N11. Initial next-generation sequencing (NGS) of the isolated viruses confirmed that both possess the eight RNA segments consistent with IAVs. However, the HA and NA genes are completely distinct from all known avian- and mammalian-origin IAV strains (Tong *et al*., 2012; Tong *et al*., 2013). Subsequent structural analysis demonstrated that the HA proteins of H17 and H18 subtypes have an altered sialic acid binding site, thereby preventing both strains from binding to both α2,6- and α2,3-linked sialic receptors (Sun *et al*., 2013). Furthermore, many residues required for sialic acid cleavage are substituted in the NA proteins (Tong *et al*., 2012; Tong *et al*., 2013; Zhu *et al*., 2012). Following transcriptomic profiling, major histocompatibility complex (MHC) class II leukocyte antigen DR isotype (HLA-DR) was identified as the key entry determinant for H17N10 and H18N11 (Karakus *et al*., 2019). Human leukocyte antigen (HLA), alternatively known as major histocompatibility complex (MHC), plays a crucial role in antigen presentation to $CD4^+$ T-cells (Roche & Furuta, 2015). Previous work has implicated HLA-DR as a HIV-1 viral attachment factor and as a co-receptor for Epstein-Barr virus entry (Mann *et al*., 1988; Spriggs *et al*., 1996).

In early 2025, a further novel bat-origin IAV strain was reported. H18N12, isolated from South American fishing bats (*Noctilio albiventris*), shares 90% sequence identity with H18N11, differing only in the NA gene. Given the recent isolation of H18N12, its pathogenicity has yet to be assessed (Echeverri-De la Hoz *et al.*, 2025). The discovery of H17N10, H18N11 and H18N12, all utilising an alternative entry mechanism to previously isolated IAV strains, indicates a novel subset of IAVs may be circulating in South American bat species (Echeverri-De la Hoz *et al.*, 2025; Tong *et al*., 2012; Tong *et al*., 2013; Yang *et al*., 2021).

Bat-origin H9N2

Conversely, a novel H9N2 strain (A/bat/Egypt/381-OP/2017) has been discovered in oral and rectal swabs from Old World Egyptian fruit bats (*Rousettus aegyptiacus)* (Kandeil *et al.,* 2019). *R. aegyptiacus* are the primordial reservoirs of numerous human-infective viruses, including Marburg, Ebola and Kasokero virus (Pourrut *et al*., 2009; Kalunda *et al*., 1986). Unlike H17 and H18 subtypes, A/bat/Egypt/381-OP/2017 appears to use α2,3-linked sialic acid receptors for host cell entry. Moreover, serological analysis demonstrates cross-reactivity against other H9N2 subtypes, suggesting A/bat/Egypt/381-OP/2017 was likely introduced by an avian H9 IAV strain. The evidence of A/bat/Egypt/381-OP/2017 in a bat species with a known history in viral emergence, coupled with the use of faecal-oral transmission routes, suggests that if A/bat/Egypt/381-OP/2017 undergoes reassortment to bind α2,6-linked sialic acid receptors, there is a significant risk of transmission into other mammalian and human populations (Kandeil *et al*., 2019).

Potential of spillover of bat-origin IAVs

Despite a preference for α2,3 sialic acid receptors, A/bat/Egypt/381-OP/2017 can infect human respiratory cell lines, suggesting human infection may be possible (El-Shesheny *et al*., 2024). During ex-vivo infection, A/bat/Egypt/381-OP/2017 replicates to low titres in human bronchus cells, similar to many avian IAV strains. However, higher titres have been observed in human lung cells, with levels of viral replication equivalent to that of cells infected with 2009 pandemic H1N1. Furthermore, within the same study, A/bat/Egypt/381-OP/2017 was demonstrated to be transmissible between male ferrets (El-Shesheny *et al*., 2024). Similarly, H18 and H17 pseudotyped viruses are able infect human glioblastoma (U-87MG) and cancer cell lines (Calu-3). Both pseudotyped viruses can also enter 293T cells expressing human, mice and pig HLA-DR receptors (Karakus *et al.,* 2019). However, interspecies transmission of H17N10 and H18N11 has yet to be reported in non-bat species (Yang *et al*., 2021; Ciminski *et al*., 2019).

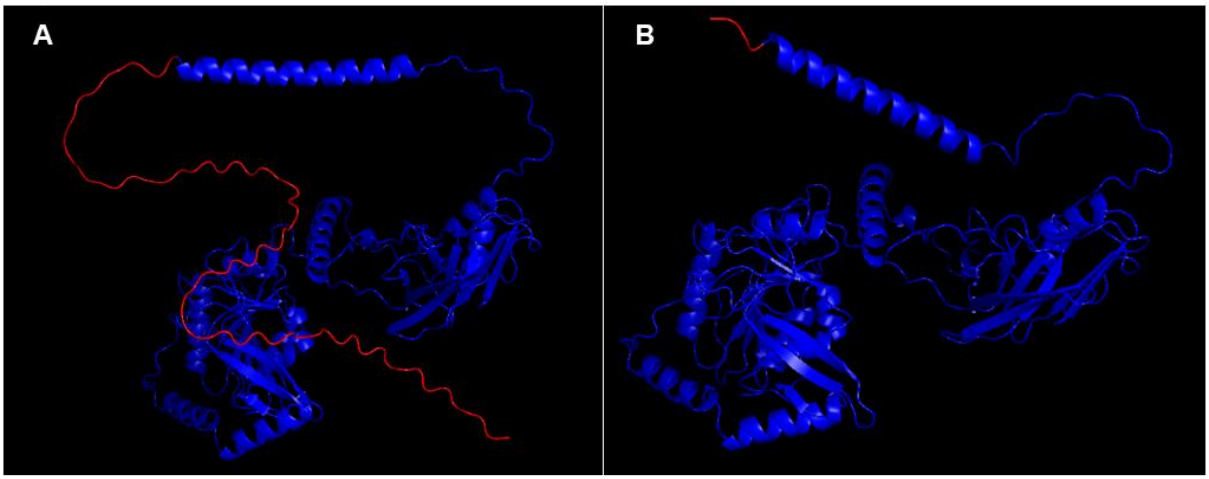
B4GALNT2

The continual discovery of IAV strains in novel host species necessitates the systematic elucidation of host factors involved in processes underlying viral attachment and entry as these are likely significant barriers to viral spillover. In 2017, a genome-wide CRISPR/Cas9 overexpression screen identified Beta-1,4-N-acetyl-galactosaminyltransferase 2 (B4GALNT2) as an IAV inhibitory factor (Heaton *et al.*, 2017).

B4GALNT2 was initially identified in 1980 as the glycosyltransferase responsible for catalysing the transfer of an N-acetyl galactosamine (GalNAc) residue onto sialic acid-containing glycans, thereby completing the synthesis of the Sd^a antigen (Groux-Degroote *et al.*, 2018). The Sd^a antigen is a blood group antigen located on >90% of Caucasian individuals' erythrocytes that plays several important roles in cell-cell recognition and binding, particularly within the colon (Duca *et al.*, 2023; Morton *et al.*, 1970). As such, B4GALNT2 expression is often associated with improved outcomes in colon cancer patients (Pucci *et al.*, 2020). During colon cancer, B4GALNT2 expression is downregulated, preventing Sd^a antigen formation. In the absence of Sd^a, the Sialyl-Lewis X (sLe^x) antigen is expressed, which is known to increase the metastatic potential of leukocyte cells by promoting transmigration and tumour formation (Malagolini *et al.*, 2007; Groux-Degroote *et al.*, 2021; Duca *et al.*, 2023). Moreover, B4GALNT2 can alter the glycan composition of cells located within the intestinal epithelia, indirectly preventing or promoting the binding of invading bacterial pathogens. For example, expression of mouse B4GALNT2 can promote *Morganella morganii* infection, while preventing *Citrobacter rodentium* entry and infection (Duca *et al.,* 2023; Vallier *et al.*, 2023; Suwandi *et al.*, 2022).

Two non-soluble human B4GALNT2 isoforms exist, one comprising 566 amino acids (B4GALNT2 X1) and the other 506 amino acids (B4GALNT2 X2). Both isoforms are localised in the Golgi, featuring a cytosolic N-terminus, transmembrane helix, stem region and a C-terminal catalytic domain. However, the 66-amino acid cytoplasmic tail of the long form contains targeting signals that facilitate both Golgi localisation and sorting to the cell surface (Figure 1) (Groux-Degroote *et al.*, 2018). Previous analysis of opposing glycosyltransferases demonstrates signals located in the cytoplasmic and transmembrane regions are essential for sub-cellular localisation, and subsequent enzymatic activity (Grabenhorst & Conradt, 1999; Tang *et al.*, 1992).

Figure 1: AlphaFold predicted structure of human B4GALNT2 alleles.



N-terminal cytoplasmic tails highlighted in red. Made with PyMol. (A) Alphafold2 prediction of B4GALNT2 X1 (amino acids 1-566) (GenBank accession - NM_153446.3). (B) Alphafold2 prediction of B4GALNT2 X2 (amino acids 1-506) (GenBank accession - NM_001159387.2).

B4GALNT2 expression in viral infection

Upon initial analysis, Heaton *et al*. found that B4GALNT2 expression reduced H7N2, H9N2, H5N9 and H10N4 infection and replication, likely by altering sialic acid expression and hence viral attachment. B4GALNT2-mediated enzymatic activity does not remove or alter α2,3-linked sialic acid receptor abundance but merely prevents glycan binding (Heaton *et al*., 2017). The transfer of GalNAc residues by B4GALNT2 to regions in close proximity to the receptor binding site creates steric hindrance, disrupting interactions between viral HA and host α2,3-linked sialic acids (Heaton *et al*., 2017; Wu *et al*., 2025). As B4GALNT2 only acts to convert α2,3-sialic acid-containing glycans to Sd^a epitopes, the protective effects of B4GALNT2 are restricted to α2,3-binding IAVs (Heaton *et al*., 2017; Wong *et al*., 2019; Wu *et al*., 2025). However, the molecular mechanisms underpinning B4GALNT2 inhibition of avian IAVs have not been completely resolved.

Intriguingly, overexpression of B4GALNT2 has been used to gain a deeper understanding of IAV binding dynamics. Previously thought to solely bind α2,6-linked sialic acid receptors, the entry and replication of H1N1 pandemic strains, A/California/04/2009 (Cal04) and A/Hong Kong/415742/2009 (HK09) is inhibited in MDCK cells overexpressing B4GALNT2. Consequently, indicating that α2,3 binding is required by both strains to establish viral infection. Conversely, the abolition of α2,3 binding in B4GALNT2-expressing cells has highlighted A/WSN/33 (WSN) as an α2,6 binding only virus. (Wong *et al*., 2019).

Despite the ability to inhibit avian-origin IAVs, the majority of avian species do not encode a B4GALNT2 gene within their genome. Previously, the absence of known human IAV restriction factors, such as BTN3A3, in avian species has been considered an important parameter for determining avian IAV spill over into human populations (Pinto *et al*., 2023). Transfection of human B4GALNT2 into chicken fibroblast cells significantly reduces the binding and replication of various viruses that utilise α2,3-sialic acid receptors, such as H5N8, H9N2, and Newcastle Disease virus (NDV) (Park *et al*., 2023). Although artificially expressing B4GALNT2 in poultry and other avian species may provide an approach to limiting IAV outbreaks, widespread avian B4GALNT2 expression may inadvertently promote viral spillover events (Park *et al*., 2023). B4GALNT2 may act as a selective pressure driving avian IAV strains

towards α2,6 -sialic acid receptor usage, thereby accelerating the emergence of mammalian-adapted avian IAV strains.

Aims and hypotheses

The continual emergence of novel IAV strains highlights the need for enhanced surveillance of zoonotic species and circulating strains of interest to mitigate future pandemic risks. Within the past two decades, four novel bat-origin IAV strains have been isolated, with the most recent earlier this year (Tong *et al*., 2012; Tong *et al*., 2013; Kandeil *et al.,* 2019; Echeverri-De la Hoz *et al*., 2025). Given bat species established role in viral emergence, defining the host-specific factors that regulate IAV infections in them is needed.

Recent research has demonstrated bat immune genes and proviral factors display signatures of positive selection. For example, in certain bats species, ISG15 and IRF3, display improved antiviral functionality (Morales *et al*., 2025; Banerjee *et al*., 2020). While, human B4GALNT2 has been identified as an avian IAV inhibitory factor, its antiviral functionality has never been explored in bats (Heaton *et al*., 2017). B4GALNT2 has been found to be under positive selection in the Chiropteran family, suggesting an adaptive advantage of these genes. Therefore, we hypothesise bat B4GALNT2 genes will demonstrate enhanced antiviral activity against avian-origin IAV pseudotype infection. To test this hypothesis, we developed several aims:

1. Clone bat B4GALNT2 genes into suitable expression vectors.
2. Overexpress our candidate bat B4GALNT2 alleles in mammalian cell lines.
3. Characterise the antiviral activity of bat B4GALNT2 against H5 pseudotype infection.

Methods

Phylogenetic analysis

The coding sequences of *Homo sapiens* B4GALNT2 (*hs*B4GALNT2) transcript variants X1 and X2 (GenBank: NM_153446.3 and NM_001159387.2), *Rousettus aegyptiacus* B4GALNT2 (*ra*B4GALNT2) transcript variants X1–X6 (GenBank: XM_016158492.2, XM_016158493.2, XM_016158494.2, XM_016158495.2, XM_016158496.2 and XM_016158497.2), and *Rhinolophus sinicus* B4GALNT2 (*rs*B4GALNT2) transcript variants X1-4 (Genbank: XM_019740559.1, XM_019740560.1, XM_019740561.1 and XM_019740562.1) were retrieved from the National Centre for Biotechnology Information (NCBI).

The coding sequence of *Rhinolophus trifoliatus* B4GALNT2 (*rt*B4GALNT2) was kindly provided by Dr Aaron Trent-Irving (University of Zhejiang, Bat1K project).

The ClustalW multiple sequence alignment programme on MegaX was used to construct nucleotide and amino acid alignments (Thompson *et al.*, 1994; Kumar *et al.*, 2018). All alignment visualised and annotated on Jalview software (Waterhouse *et al.*, 2009). Phylogenetic trees were constructed in MEGA X using the maximum likelihood (ML) method, and a bootstrap test of 1000 replicates (Kumar *et al.*, 2018; Feldstein, 1985). All phylogenetic tress formatted using the Interactive Tree of Life (iTOL) (Letunic & Bork, 2024).

Positive selection analyses

Mammalian B4GALNT2 genes and phylogenetic tree branches under positive selection were identified by examining the supplementary information (Supplementary Table 5 and Supplementary Table 6) from a previously published study from the Bat1K consortium (Morales *et al.*, 2023; Teeling *et al.*, 2018).

Individual B4GALNT2 codons displaying signatures of positive selection were identified by using MEME (Mixed Effects Model of Evolution) (cutoff p value <0.05), FEL (Fixed Effects Likelihood) (cutoff p value <0.05) and FUBAR (Fast, Unconstrained Bayesian AppRoximation for Inferring Selection) (posterior probability >0.9) models on the DataMonkey web server (Murrell *et al*, 2013; Kosakovsky Pond *et al.*, 2021; Murrell *et al*, 2012; Weaver *et al.,* 2018).

Gene and plasmid synthesis

The empty *pLVX-IRES-mCherry* lentiviral expression vector was kindly provided by Dr Aaron Trent-Irving (Zhejiang University). *hs*B4GALNT2 transcript variants X1 and X2, *ra*B4GALNT2 transcript variants X1–X4, *rs*B4GALNT2 transcript variants X1 and X3 and *rt*B4GALNT2 (Bat1K project) were synthesised as gBlocks (Gene fragments) by Integrated DNA Technologies (IDT). The coding sequence of the haemagglutinin (HA) gene of A/bat/Egypt/381OP/2017 was retrieved from NCBI (GenBank: MH376902.1) and synthesised as a gBlock by IDT.

pLVX-IRES-mCherry-rax2B4GALNT2 and *pLVX-IRES-mCherry-rax3B4GALNT2* plasmids were synthesised by Thermo Fisher Scientific using the GeneArt™ Custom Gene Synthesis service (Regensburg, Germany).

PCR

The complementary nucleotide sequences required for HiFi assembly were attached to synthesised gBlocks by PCR. Primer sequences are listed in Supplementary Table 1. Each reaction mixture contained 4µL of Phusion HF or GC buffer, 0.4µL of 10mM dNTPs, 1µL of 10µM forward primer, 1µL of 10µM reverse primer, 0.6µL of DMSO, 0.6µL of Mg^{2+}, 0.2µL of Phusion polymerase, 100-150ng of template DNA and nuclease-free water to a final volume of 20µL. The cycling conditions for all reactions included an initial denaturation step of 98°C for 30 seconds, followed by 35 cycles at 98°C for 10 seconds, 59-65°C for 30 seconds (specific annealing temperatures listed in Supplementary Table 1), 72°C for 30 seconds, and a final incubation at 72°C for 10 minutes. All PCRs were performed in the MJ Mini™ Gradient Thermal Cycler (BioRad).

DNA gel extraction

1-5µg of *pVAX1-BMP2-H5HA* was digested with NotI-HF and EcoRI in rCutsmart buffer (NEB) at 37°C for 15 minutes. The complete digest was loaded onto an agarose gel. Following electrophoresis, the band corresponding to the empty vector backbone was excised from the gel. The DNA was purified from the gel using either the QIAquick Gel Extraction Kit (QIAgen) or the Monarch DNA Gel Extraction Kit (NEB) according to the supplied protocols.

Gel electrophoresis

1% agarose gels were prepared using 1% TAE and Invitrogen SYBR Safe DNA Gel Stain (Invitrogen). All gels included a 1Kb Plus DNA ladder (Invitrogen), were resolved at 100 V and visualised on the ChemiDoc Imaging System (BioRad).

pVLX-IRES-mCherry-B4GALNT2 and *pCMV-Myc-H9HA* cloning

Vector linearisation

To linearise the vector backbones prior to HiFi assembly, 1µg of *pLVX-IRES-mCherry* was digested with BamHI in rCutsmart buffer at 37°C for 30 minutes. BsmBI was added, and the digest was incubated for a further hour at 55°C. 1µg of *pCMV-Myc* was digested with EcoRI and NotI in rCutsmart buffer at 37°C for 30 minutes. All restriction enzymes used were obtained from NEB. Linearised vectors were dephosphorylated with Antarctic phosphatase (NEB) and purified using the Monarch DNA Clean up kit (NEB) according to the supplied protocols.

HiFi assembly

The PCR amplified B4GALNT2 inserts were cloned into the linear *pVLX-IRES-mCherry* vectors, and the PCR amplified A/bat/Egypt/381OP/2017 H9 inserts were cloned into linear *pCMV-Myc*. All assembly reactions performed using the NEBuilder® HiFi DNA Assembly Cloning Kit (NEB) according to the manufacturer supplied protocol. A vector-to-insert ratio of 0.067:0.133 was used throughout all assemblies.

Bacterial transformation

2µL of each HiFi assembly product was added to 10µL of NEB 5-alpha chemically competent cells (NEB) and incubated on ice for 30 minutes. The bacterial cells were lysed at 42°C for 30 seconds before being returned to ice for 2 minutes. 200µL of SOC medium was added, and the mixtures were incubated in an orbital shaker set to 37°C with shaking at 200 rpm for one hour. The transformed cells were then plated on ampicillin selective Luria-Bertani (LB) agar and incubated for overnight at 37°C.

Colony PCR

Following HiFi assembly and subsequent bacterial transformation, colony PCR was used to verify plasmid insert expression. Half a bacterial colony was collected using a sterile pipette tip and added to the corresponding PCR master mix. Prior to the typical cycling steps, an extended initial denaturation step of 98°C for 5 minutes was included in order to lyse the bacterial cells. All colony PCRs were performed using the same primers, reaction and cycling conditions as listed above.

Plasmid purification

The remainder of the colonies showing insert expression during colony PCR were added to 5mL of LB-agar broth with 5µL of ampicillin and incubated at 37°C overnight in an orbital shaker set to 200 rpm. Plasmid DNA isolated and purified using the Monarch® Spin Plasmid Miniprep Kit (NEB).

Plasmid DNA stock expanded by transforming plasmid DNA isolated by Miniprep into subcloning efficiency *Escherichia coli* DH5α (NEB). Colonies were added to 250ml of LB-agar broth with 250µl of ampicillin and incubated overnight using a 37°C orbital shaker set at 200rpm. Plasmid DNA isolated using the PureYield™ Plasmid Maxiprep System (Promega).

Plasmid restriction digests

To confirm successful cloning, 250-500ng of isolated plasmid DNA was digested with restriction enzymes (NEB). Where possible, one cut site was located in the vector backbone and the other only present in the insert.

Whole Genome Plasmid Sequencing (WGS)

1.25µg of plasmid DNA was sent to GeneWiz (Azenta Life Sciences) for nanopore WGS. Sequencing results were aligned with B4GALNT2 sequences using ClustalW on MEGA X to confirm cloning success. Plasmid maps were visualised on Benchling.

Cell culture

293T (immortalised human embryonic kidney cells), MDCK (Madin-Darby canine kidney cells) and A549 (human alveolar basal epithelial) cells were cultured and maintained in Dulbecco's Modified Eagle Medium (DMEM) (Gibco), supplemented with 10% fetal bovine serum (FBS) (Gibco), 1% L-glutamine (Gibco), 1% penicillin-streptomycin (P/S) (Gibco) and 100µg/mL Normocin (Invivogen) at 37°C and 5% CO_2. Cells used in downstream assays were maintained in antibiotic- and Noromocin-free DMEM.

Influenza A lentiviral pseudotype production

H5

8×10^5 293T cells were seeded in a 6-well plate and incubated overnight at 37°C until reaching 70-90% confluency. Cells were transfected with 6µg of FuGene (Promega), 500ng of *p8.91* packaging plasmid (packaging plasmid; expresses HIV-1 Gag and Pol), 750ng of reporter plasmid (*pCSFLW* for luciferase or *pLenti-GFP-Puro*), 250ng of *pHAT* (expresses Human Airway Trypsin-like protease (HAT)) and 500ng of *pVAX1-BMP2-H5HA* (Influenza A virus haemagglutinin H5). Media was replaced with fresh DMEM following overnight incubation. At 24 hours post-transfection, 0.5 units of Neuraminidase (NA) from *Clostridium perfringens* (Sigma-Aldrich) was added. Supernatants were collected at 48-hours post-transfection, passed through a 0.45µm filter and stored at -80°C.

VSV-G bearing lentiviral pseudotype particles with GFP reporters were produced as a control using 6µg of FuGene (Promega), 500ng of *p8.91*, 450ng *VSV-g* (envelope) and 750ng of reporter plasmid (*pLVX-IRES-mCherry-B4GALNT2*, *pCSFLW*, or *pLenti-GFP-Puro).*

Bat-derived hemagglutinin H9

8×10^5 293T cells were seeded in a 6-well plate and incubated overnight at 37°C until reaching 70-90% confluency. Cells were transfected with 6µg of FuGene (Promega), 500ng of *p8.91*, 650-950ng of reporter plasmid (*pCSFLW* for luciferase or *pLenti-GFP-Puro*), 150-400ng of *pHAT,* and 300-900ng of *pCMV-myc-H9HA* (Influenza A virus haemagglutinin H9). The cells were incubated overnight, and media replaced the following day with fresh DMEM. 0.5 units of NA from *Clostridium perfringens* (Sigma-Aldrich) was added to chosen wells at 24 hours post-transfection. Supernatants were collected at 48-hours post-transfection, passed through a 0.45µm filter and stored at -80°C.

VSV-mediated delivery of *pVLX-IRES-mCherry*

8 x 10^5 293T cells were seeded in a 6-well plate and incubated overnight at 37°C in DMEM until reaching 70-90% confluency. The cells were transfected with 5.1-8.85µg of polyethylenimine (PEI), 500ng of *p8.91* (packaging plasmid), 450ng *VSV-G* (expresses VSV glycoprotein) and 750ng-2µg of reporter plasmid (*pLVX-IRES-mCherry* or *pCSFLW* for luciferase). The cells were incubated for approximately 8 hours, thereafter the media was replaced with fresh DMEM. Supernatants were harvested 48- and 72- hours post-transfection, passed through a 0.45µM filter and stored at -80°C.

Lentiviral pseudotype transduction

4 × 10^4 293T cells were seeded in a 96-well plate and incubated overnight at 37°C until reaching 70-90% confluency. The cells were treated with 2-fold serial dilutions of viral pseudotyped particles and incubated for 48 hours. All viral transductions performed in triplicates.

Luciferase assays were performed 48-hours post-transduction. Supernatants were removed from transduced cells and replaced with a 1:1 ratio of Bright-Glo™ reagent (Promega) and DMEM. Cells were incubated for 3 minutes before the solution was transferred to an empty opaque 96-well plate. The resulting luminescence was measured using the BioTek Cytation 3 Plate Reader (Aligent).

pLVX-IRES-mCherry-B4GALNT2 transfections

293T, MDCK and A549 cells were seeded in 6-well or 48-well plates and incubated overnight at 37°C until reaching 70-90% confluency. Prior to transfection, DMEM was replaced in all wells with OPTIMEM. In 6-well plates, cells were transfected with 1.35-4.5µg of PEI, Lipofectamine 2000 (Invitrogen) or FuGene (Promega) and 450-1500ng of *pLVX-IRES-mCherry-B4GALNT2* plasmids. In select transfections, a 1:1 ratio of *pLVX-IRES-mCherry-B4GALNT2* construct to *p8.91* was included. A control well transfected with 450ng of *pEGFP-C3* and 1.35µg of PEI, Lipofectamine 2000 (Invivogen) or FuGene (Promega) was included during all transfections. Transfection in 48-wells performed using 250ng of each *pLVX-IRES-mCherry-B4GALNT2* plasmid or *pEGFP-C3* and 750ng of PEI. All transfections for downstream infection assays performed in triplicates. Media was replaced in all wells 7-8 hours post-transfection with fresh DMEM.

Fluorescence microscopy

Fluorescence was visualised using the Axio Vert A1 inverted microscope using 540-580nm LED and Filter set 00 for mCherry expression and 470nm LED and Filter set 38 for Green fluorescent protein (GFP) expression (Ziess). Images captured using the AxioCam Erc 5s colour digital camera and edited using AxioVision LE software (Zeiss).

H5-GFP pseudotype transduction

At 48-hours post-transfection, 293T cells overexpressing B4GALNT2 were transduced with either a 1:2 or 1:4 dilution of H5 pseudotype particles expressing GFP to a final volume of 200µL. Plates were centrifuged at 1500 x g for 90 minutes at room temperature. Following centrifugation, cells were incubated at 37°C for 48 hours. Cells were harvested using 0.25% trypsin-EDTA solution and fixed with 2% PFA for 15 minutes prior to flow cytometry analysis. All H5-GFP pseudotype transductions were performed in triplicates.

Lectin binding assay to measure α2,3- or α2,6-linked sialic acid binding

293T cells overexpressing *pVLX-IRES-mCherry* or non-transfected MDCK cells were harvested using 0.25% trypsin-EDTA solution. The cells were transferred to ice, blocked using 1% BSA, and then incubated with at 4°C with 5-30µg/ml of Maackia amurensis Lectin (MAA/MAL II) conjugated to FITC (US Biological) or 0.5-10µg/ml Sambucus Nigra (Elderberry Bark) Lectin (SNA, EBL) conjugated to FITC (Invitrogen) for 30 minutes to 1 hour. Cells were fixed using 2% PFA for 15 minutes and then analysed by flow cytometry.

293T cells overexpressing *pVLX-IRES-mCherry* were harvested using 0.25% trypsin-EDTA solution and incubated with 0.1-0.25U/ml neuraminidase (NA) from *Arthrobacter ureafaciens* (Sigma Aldrich) or *Clostridium perfringens* (Sigma Aldrich). Cells incubated at 37°C for either 20 hours or 1 hour prior to biotinylated lectin probe analysis.

Flow cytometry

All flow cytometry analyses was conducted in and supported by the IRR Flow Cytometry and Cell Sorting Facility (University of Edinburgh). The Attune™ Cytpix Flow Cytometer (ThermoFisher Scientific) was used for all analyses. Data was visualised and processed using the Attune™ NxT software.

MOI calculations

Multiplicity of infection (MOI) was calculated using the inverse Poisson distribution; m=−ln(1−p), where p is equal to the number of GFP-positive cells divided by the total number of cells (Figliozzi *et al.*, 2016).

Statistical analysis

All statistical analyses were performed using one-way analysis of variance (ANOVA) followed by Tukey's multiple comparisons tests on Prism 8 software (GraphPad, 2019). For all figures, *p <0.05, ** p <0.01, *** p <0.001. and **** p <0.001.

Results

Selection of B4GALNT2 as candidate gene

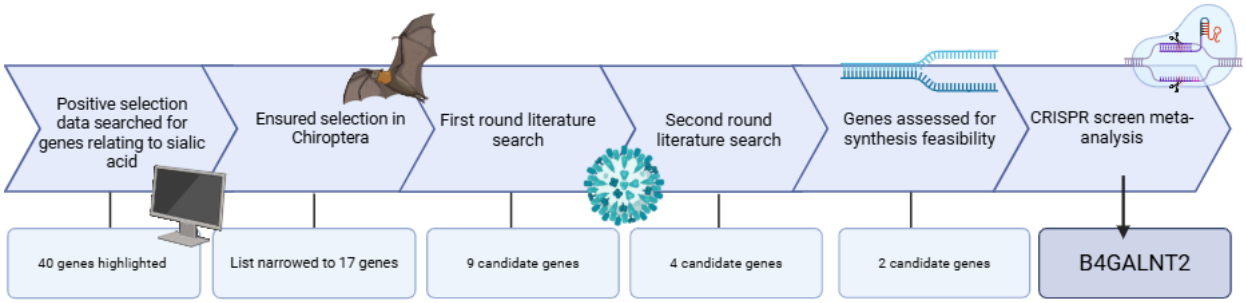
Understanding how bat species are able to tolerate viral infection without disease is essential. Given the importance of sialic acid receptors in mediating viral entry, we postulated that genes regulating sialic acid receptor expression/functionality may be under selection in bat species.

To identify such genes, we searched a previously published list of positively selected mammalian immune genes. 'Sialic acid' was used as a search term to screen the disease and gene association and molecular function columns of Supplementary Table 6 from the Morales *et al*. study (Morales *et al*., 2025). The initial search returned 40 genes, of which 17 are reported to be under positive selection across the Chiropteran order. The 17 positively selected bat genes were searched on PubMed alongside the term 'sialic acid'. Genes not directly related to sialic acid regulation were discarded from our analysis. The following genes underwent further analysis: B4GALNT2, CD22, CD27, HOGA1, LDLR, SIGLECL1, ST3GAL3, ST3GAL6 and NEU3.

The candidate gene list was further narrowed by searching each gene on PubMed and Google Scholar with keywords, such as 'virus', 'influenza A' and 'IAV' to understand if they had been studied in a viral context. This limited the candidate gene list to four genes, namely, CD22, NEU3, B4GALNT2 and SIGLECL1. Synthesis feasibility for downstream functional assays was next assessed by constructing nucleotide alignments of the candidate genes. CD22 and SIGELC1 appeared too large to affordably synthesise (average amino acid length ~865 and ~1710), restricting the prospective genes to NEU3 and B4GALNT2. At the time of analysis, 14 genome-wide CRISPR Influenza A screens had been performed (Munir *et al*., 2024; Heaton *et al*., 2017; Sharon *et al*., 2020; Gao *et al*., 2022; Chia *et al*., 2020; Han *et al*., 2018; Li *et al*., 2020; Tran *et al*., 2020; Anderson *et al*., 2021; Song *et al*., 2021; Zhou *et al*., 2021; Yi *et al*., 2022; King *et al*., 2023; Ma *et al*., 2023). Each screen was examined for mention of B4GALNT2 or NEU3. B4GALNT2 appeared in two screens, exhibiting the highest enrichment score across both, whereas NEU3 did not appear in any of the screens (Heaton *et al*., 2017; Munir *et al.,* 2024). Based on

this systematic analysis, B4GALNT2 was selected as our primary gene for downstream analysis (Figure 2).

Figure 2: Schematic of gene selection workflow.

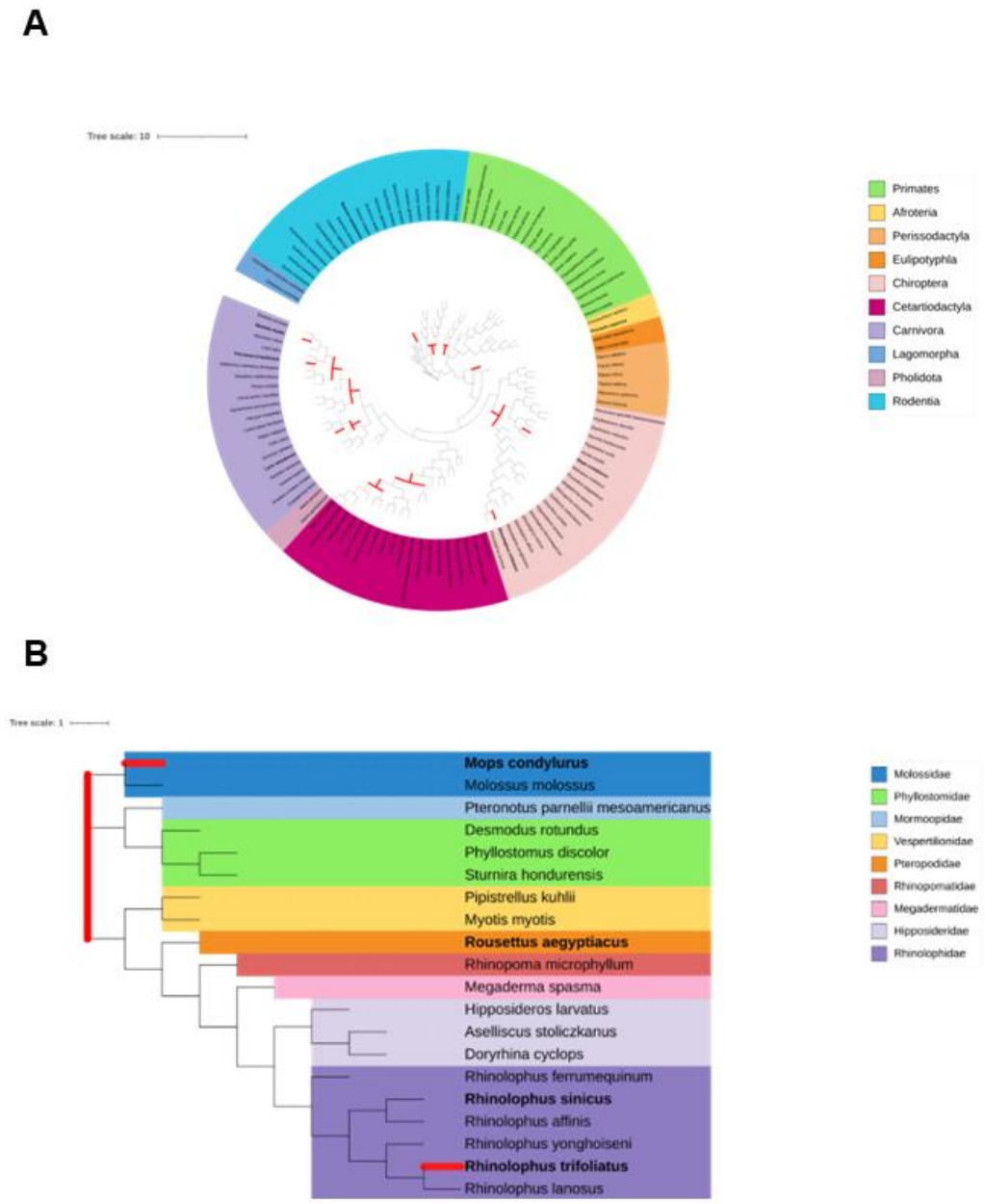


Bat immune genes relating to sialic acid receptor use identified using signatures of genetic positive selection as per Morales *et al*. Subsequent literature searches, alignments and CRISPR-cas9 meta-analysis used to narrow down candidate gene list. Schematic created using BioRender.com.

Bat B4GALNT2 genetic analysis

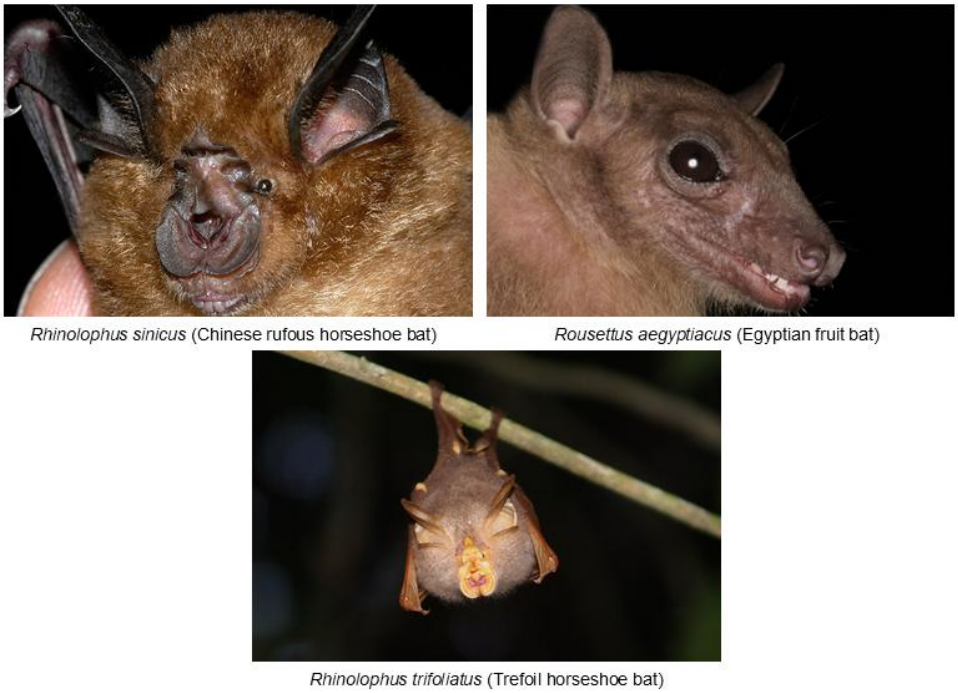
As B4GALNT2 is widely expressed across mammalian species, we wanted to visualise the genetic relationship between positively selected mammalian B4GALNT2 genes by constructing a phylogenetic tree (Figure 3A). The 115 mammalian B4GALNT2 sequences used to construct the tree, including those from 20 bat species, were provided by the Bat1K consortium (Morales *et al*., 2025). Despite appearing to be under selection across the entire Chiropteran order, branch-specific analysis reveals B4GALNT2 selection is most evident in two species, *Rhinolophus trifoliatus* (Trefoil horseshoe bat) and *Mops condylurus* (Angolan free-tailed bat) (Morales *et al*., 2025) (Figure 3B). Due to a high frequency of ambiguous bases, *M. condylurus* B4GALNT2 was excluded from our analysis. This ambiguity likely arose from poor quality sequencing. In addition to *R. trifoliatus* B4GALNT2, two bat species' B4GALNT2 genes not subject to diversifying selection were included in our analysis (Figure 4). *Rousettus aegyptiacus* (Egyptian fruit bat) B4GALNT2 was selected due to their carriage of a novel bat-origin H9N2 IAV strain (Figure 4). *Rhinolophus sinicus* (Chinese rufous horseshoe bat) B4GALNT2 was chosen due to their close genetic relationship to *R. trifoliatus* as a Rhinophid family member and status as SARS-like coronavirus reservoirs (Figure 3B and 4) (Yuan *et al*., 2010).

Figure 3: Phylogenetic analysis of mammalian and bat B4GALNT2 orthologs.



Both trees generated using MEGA X and formatted using iTOL. (A) Phylogenetic tree constructed using codon alignment of 115 mammalian B4GALNT2 sequences. Branches under positive selection are highlighted in red. (B) Phylogenetic tree of bat B4GALNT2 genes. Branches and species under positive selection highlighted in red. Genes studied throughout this analysis highlighted in bold.

Figure 4: Bat species studied throughout this analysis.



Rhinolophus sinicus (Chinese rufous horseshoe bat)

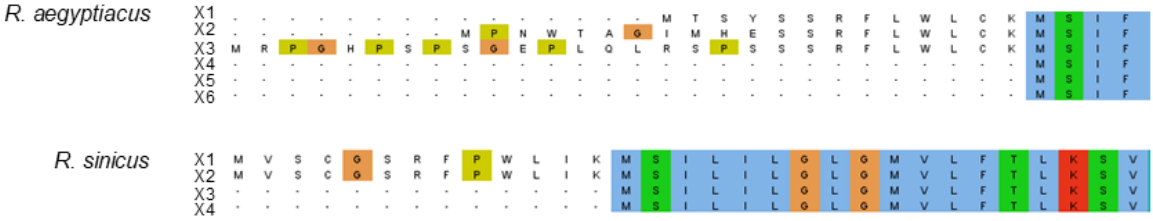
Rousettus aegyptiacus (Egyptian fruit bat)

Rhinolophus trifoliatus (Trefoil horseshoe bat)

Rhinolophus sinicus (University of Bristol). *Rousettus aegyptiacus* (iNaturalist). *Rhinolphus trifoliatus* (iNaturalist).

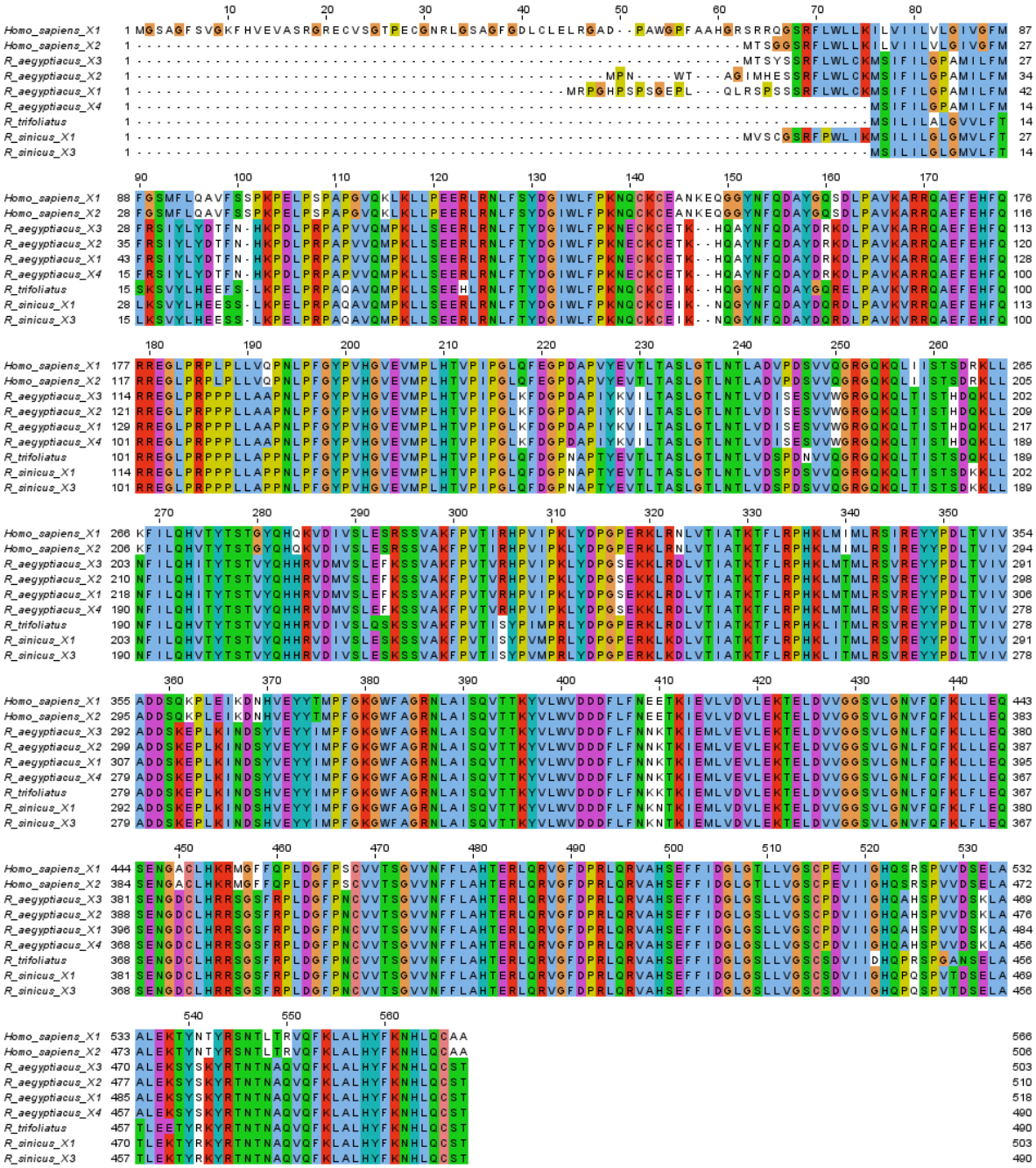
Next, we aimed to assess whether multiple coding variants are spliced from the mRNA of our candidate B4GALNT2 genes, as such variants may demonstrate differential antiviral activity. As previously reported, two non-soluble human B4GALNT2 (*hs*B4GALNT2) isoforms exist; *hsx1*B4GALNT2 and *hsx2*B4GALNT2, (Groux-Degroote *et al.*, 2018). We identified six *R. aegyptiacus* B4GALNT2 splice variants. *rax1-3*B4GALNT2 are distinct from each other and all other variants; however, *rax4-6*B4GALNT2 are conserved (Figure 5). Consequently, *rax1-4*B4GALNT2 were synthesised for downstream analysis. Similarly, *R. sinicus* potentially generates four splice variants; *rsx1*B4GALNT2 and *rsx2*B4GALNT2 which are identical, and *rsx3*B4GALNT2 and *rsx4*B4GALNT2 which also appear identical (Figure 5). Therefore, *rsx1*B4GALNT2 and *rsx3*B4GALNT2 were synthesised. At the time of analysis, no alternatively spliced *R. trifoliatus* B4GALNT2 coding mRNA variants were publicly available. Intriguingly, all splice variants of our candidate B4GALNT2 genes differentiate solely in the early N-terminal region (Figure 6). Altogether, two *rs*B4GALNT2 alleles, four *ra*B4GALNT2 alleles and one *rt*B4GALNT2 variant were included in downstream analyses.

Figure 5: Amino acid alignment of the N-terminal region of *R. aegyptiacus* and *R. sinicus* B4GALNT2 alleles



Codon sequences were aligned using ClustalW in MegaX. Sequences visualised on Jalview with the Clustal colouring scheme. Blue – small, hydrophobic residues. Cyan – aromatic residues. Green - polar residues. Red - basic residues. Magenta - acidic residues. Pink – cysteines. Orange – glycines. Yellow - prolines. White – unconserved.

Figure 6: Amino acid alignment of all candidate B4GALNT2 genes



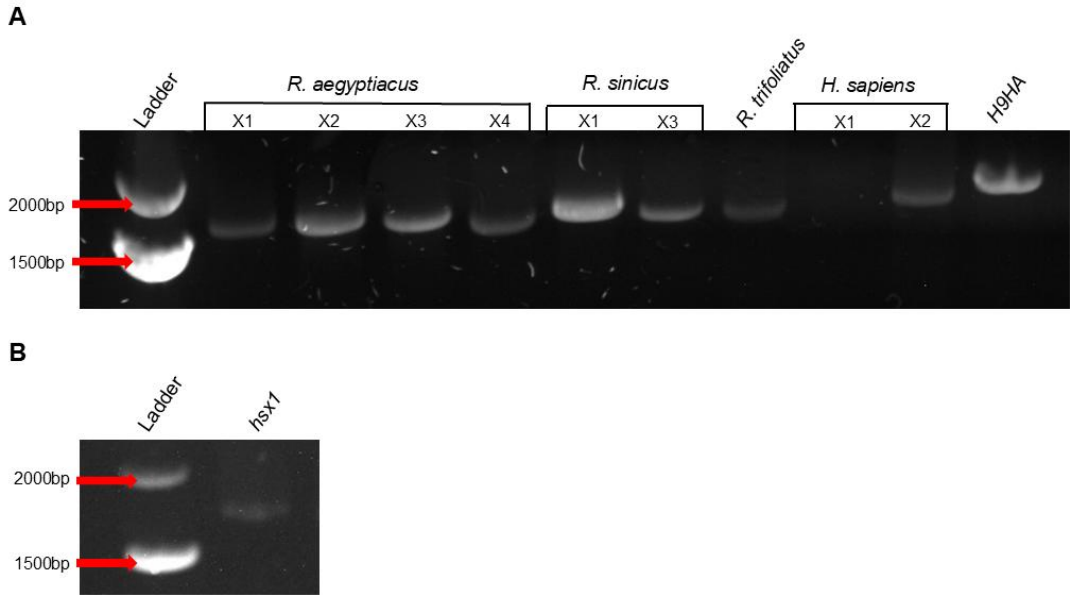
Codon sequences aligned using ClustalW in MegaX. Sequences visualised on Jalview with the Clustal colouring scheme. Blue – small, hydrophobic residues. Cyan – aromatic residues. Green - polar residues. Red - basic residues. Magenta - acidic residues. Pink – cysteines. Orange – glycines. Yellow - prolines. White – unconserved.

Construction of *pVLX-IRES-mCherry-B4GALNT2* plasmids

We first synthesised human, *R. trifoliatus, R. aegyptiacus* and *R. sinicus* B4GALNT2 alleles, and then sought to express them in mammalian cell lines via a lentiviral bicistronic expression vector. The plasmid *pLVX-IRES-mCherry* allows the simultaneous expression of a protein-of-interest, in this study, B4GALNT2, and mCherry fluorescent protein, which is expressed via an internal ribosomal entry site (Clontech Laboratories Inc., 2009).

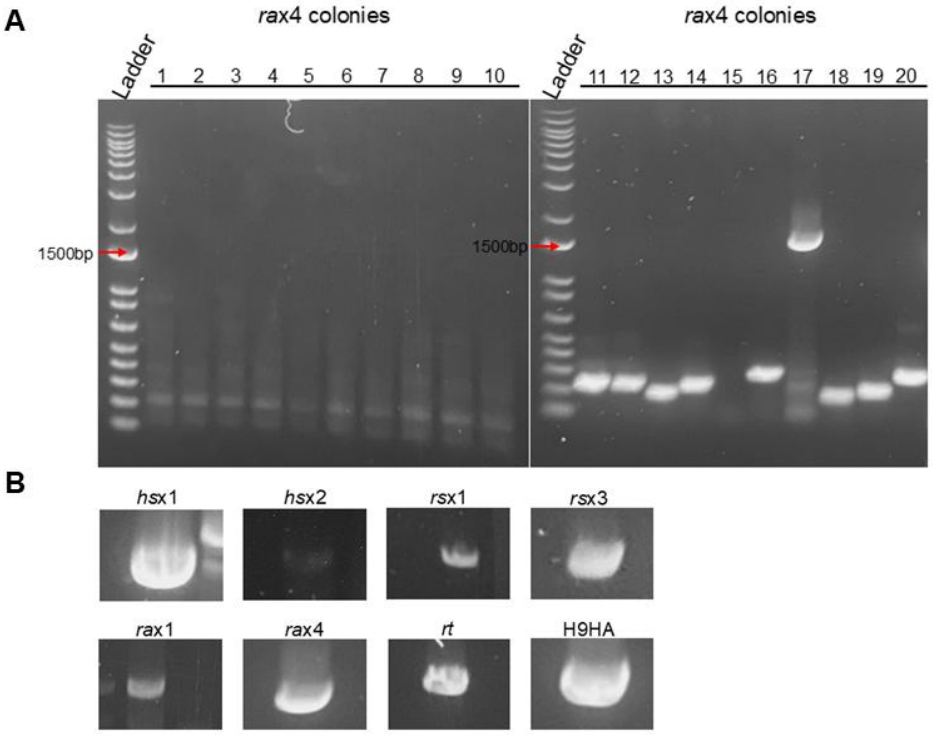
HiFi DNA assembly was used to engineer our B4GALNT2 constructs. Firstly, we performed a PCR to attach the 15 overlapping base pairs required for HiFi assembly to the synthesised B4GALNT2 genes (Figure 7). All alleles except for *hs*x1B4GALNT2 were amplified during the first PCR performed (Figure 7A). To successfully amplify *hs*x1B4GALNT2, an increased amount of template DNA was required (Figure 7B). Following assembly, constructs were transformed into bacterial cells. Approximately one in twenty colonies (~5%) contained plasmid DNA with the correct insert (Figure 8A). This low rate of assembly was most apparent for *rax2*B4GALNT2 and *rax3*B4GALNT2 plasmids as after screening >50 colonies we were unable to confirm insert expression. Subsequently, the decision was made to synthesise these plasmids entirely, as opposed to the synthesis of gene inserts as previously pursued. The remaining B4GALNT2 inserts were identified following multiple rounds of colony PCR (Figure 8B).

Figure 7: Amplified PCR products with attached overhangs for HiFi assembly.



B4GALNT2 genes expressing the attached overhangs for HiFi assembly separated by DNA gel electrophoresis. (A) First round PCR using 100ng template DNA. (B) Second round PCR using 150ng template DNA. Band sizes expected to be ~1400-1700bp.

Figure 8: Colony PCR amplification



PCR amplification of B4GALNT2 inserts from *pVLX-IRES-mCherry-B4GALNT2* transformed cells separated by gel electrophoresis. (A) Gel images of different *rax*4B4GALNT2 transformants. (B) Confirmation of B4GALNT2 insert expression in *pVLX-IRES-mCherry-B4GALNT2* transformants.

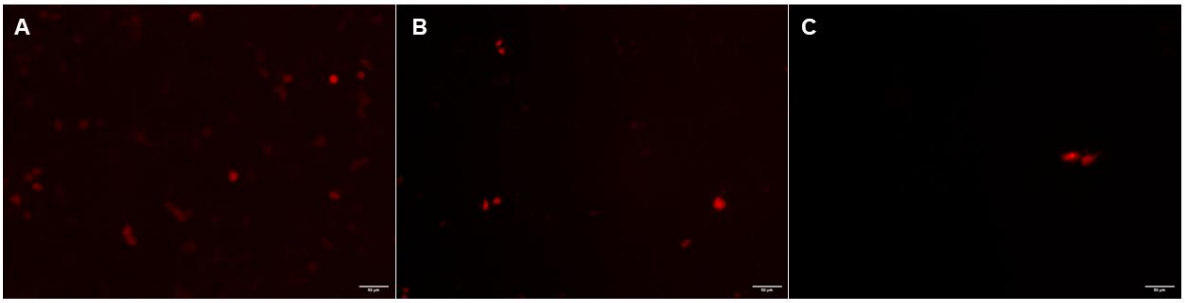
To ensure the B4GALNT2 inserts had been incorporated into the *pLVX-IRES-mCherry* backbones correctly, isolated plasmid DNA was validated by restriction digest. Where possible, digests were performed with one cut site present only in the insert and one in *pLVX-IRES-mCherry* vector backbone. Despite appearing positive for B4GALNT2 insert expression during colony PCR, plasmid DNA isolated from many of these colonies failed to be confirmed by restriction digest. Consequently, multiple rounds of screening and validation were required to isolate all *pLVX-IRES-mCherry-B4GALNT2* plasmids. Prior to initial transfections, all *pLVX-IRES-mCherry-B4GALNT2* plasmids were confirmed by whole genome sequencing (WGS) (Supplementary Figure 1).

Expression of *pLVX-IRES-mCherry-B4GALNT2* in mammalian cell lines

Initial transfections in 293T, MDCK and A549 cells

After the *pLVX-IRES-mCherry-B4GALNT2* constructs were verified by whole plasmid sequencing, we wanted to ensure they produced a fluorescent signal in 293T cells. Addition of the B4GALNT2 inserts into the *pVLX-IRES-mCherry* vector resulted in a considerable reduction in fluorescence intensity. mCherry expression appeared to decrease by >50% in cells transfected with *pVLX-IRES-mCherry-B4GALNT2* constructs compared to that of the empty vector (Figure 9). Following confirmation of fluorescence in 293T cells, the constructs were transfected into MDCK and A549 cells as these cell lines were intended for downstream infection assays with influenza A virus. However, no visible fluorescence was detected in either cell line even upon titration of the plasmid DNA. As previously published studies have successfully delivered B4GALNT2 into MDCK cells via VSV-G-mediated lentiviral transduction, we attempted to use this approach (Wong *et al*., 2019).

Figure 9: Initial transfection of *pLVX-IRES-mCherry-B4GALNT2* constructs into 293T cells

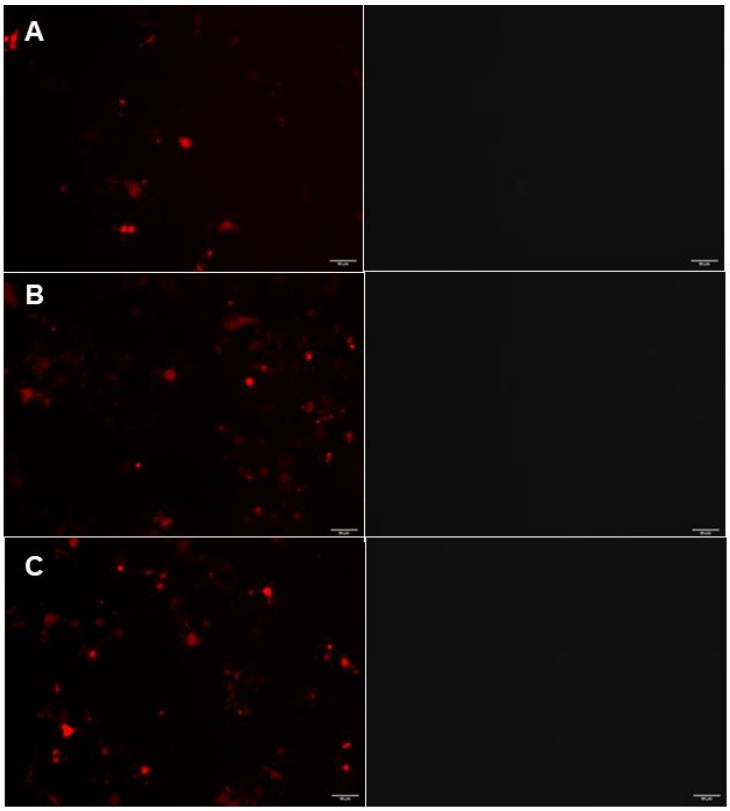


Transfection performed using 450ng of plasmid DNA with PEI. Cells visualised using 20X magnification. Scale bar 50µm. (A) Empty *pLVX-IRES-mCherry.* (B) *pLVX-IRES-mCherry-hsx1B4GALNT2*. (C) *pLVX-IRES-mCherry-rtB4GALNT2.*

Attempts at VSV-G-mediated transduction

To generate *pLVX-IRES-mCherry*-expressing VSV-G pseudotype particles, the *pLVX-IRES-mCherry* vector was co-transfected into 293T cells with *p8.91* and VSV-G. Co-transfection appeared to notably improve the mCherry fluorescence signal; however, transduction of both MDCK and 293T cells with the mCherry-expressing VSV-G pseudotyped particles did not produce detectable fluorescence. To troubleshoot, we increased the amount of *pLVX-IRES-mCherry* construct co-transfected. mCherry fluorescence appeared to increase in dose-dependent manner during transfection, but no fluorescence was observed upon transduction with any pseudotypes produced (Figure 10). The challenges encountered during the cloning workflow and optimisation of VSV-G-mediated lentiviral transduction imposed significant time constraints on the project. Consequently, we decided instead to optimise the transfection in 293T cells to proceed with downstream infection assays.

Figure 10: Production and transduction of VSV pseudotype particles expressing *pLVX-IRES-mCherry*

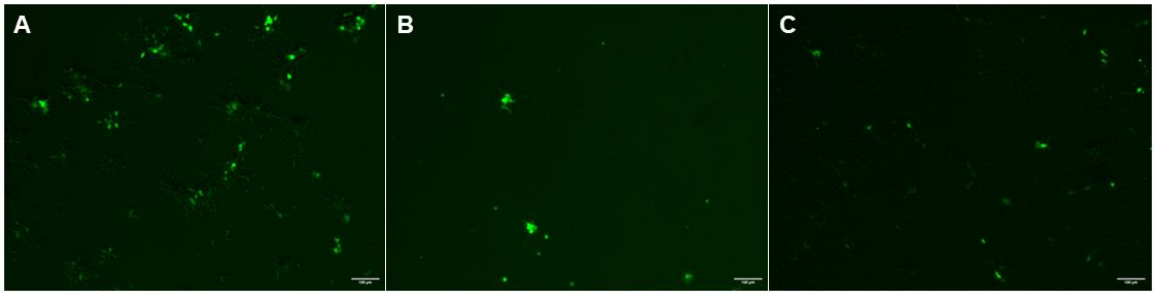


Cells visualised by 20X magnification. Scale bar 50µm. Panels on left - transfection of *pLVX-IRES-mCherry, p8.91* and VSV-G in 293T cells. Panels on right - 293T cells transduced with the supernatant collected from the transfected cells. (A) 1ug empty *pLVX-IRES-mCherry* (B) 1.5ug empty *pLVX-IRES-mCherry* (C) 2ug empty *pLVX-IRES-mCherry.*

Optimised transfections in 293T cells

As previous transfections in 293T cells with the *pLVX-IRES-mCherry-B4GALNT2* constructs gave a low fluorescent signal, we began optimising the transfection to improve efficiency. Titrating *pLVX-IRES-mCherry-B4GALNT2* DNA appeared to have no effect on fluorescent signal. However, the control transfections performed with *pEGFP-C3* also did not produce any detectable fluorescence, indicating an issue with experimental protocol or reagents. Consequently, lower passage number 293T cells and fresh PEI stock were thawed, yet neither appeared improved efficiency. Next, we tested several transfection reagents, including FuGene and Lipofectamine 2000. Although no mCherry fluorescence was observed upon transfection of *pLVX-IRES-mCherry* for any transfection reagent, PEI proved to be the most effective reagent for transfection of the *pEGFP-C3* control in 293T cells (Figure 11).

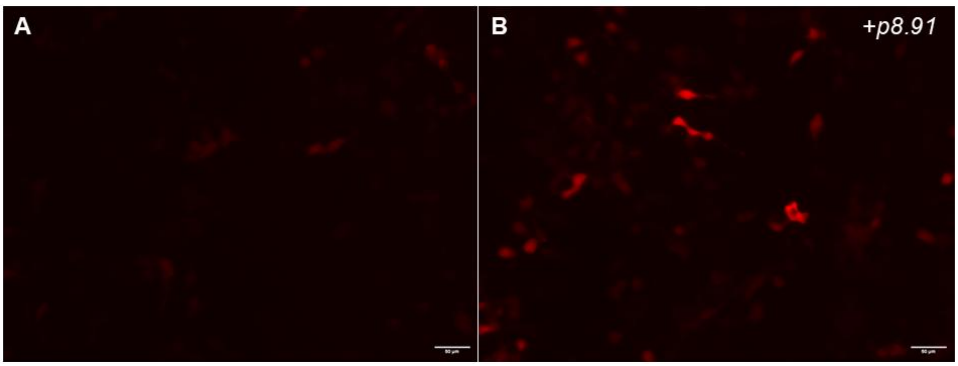
Figure 11: Comparison between PEI, Lipofectamine2000 and FuGene transfection reagents.



293T cells transfected with 450ng of *pEGFP-C3* using various transfection reagents. Cells visualised at 40X magnification. Scale bar 100µm. (A) PEI (B) Lipofectamine2000 (C) FuGene.

As successful transfection of the *pEGFP-C3* control had now been achieved, we assumed that the issues in efficiency were associated with the *pLVX-IRES-mCherry* vector backbone. During attempts to deliver the B4GALNT2 genes via VSV-G-mediated transduction, we noticed co-transfection with *p8.91* (providing HIV-1 Gag and Pol) and VSV-G rescued mCherry expression in a dose-dependent manner. Thus, a 1:1 ratio of *pLVX-IRES-mCherry-B4GALNT2* constructs to *p8.91* was transfected into 293T cells. mCherry fluorescence was visualised at 24 hours, with expression peaking at 48 hours post-transfection (Figure 12). As we were only able to achieve reasonable mCherry fluorescence from the *pLVX-IRES-mCherry-B4GALNT2* constructs during co-transfection with *p8.91,* we decided to proceed with transfections using this method. Nevertheless, we continued to optimise transfection conditions for the *pLVX-IRES-mCherry-B4GALNT2* constructs alone. Sole transfection of the *pVLX-IRES-mCherry-B4GALNT2* constructs was achieved while optimising the parameters for the H5-GFP pseudotype infectivity assays. Prior to the successful transfections, a new batch of earlier passage stock 293T cells was thawed. Changing the cell batch allowed us to achieve a transfection efficiency of between 40-50% for both the empty vector and those containing B4GALNT2 inserts. Thus, all viral pseudotype infections were conducted with cells exclusively expressing our *pVLX-IRES-mCherry-B4GALNT2* constructs and without Gag/Pol expression.

Figure 12: 293T cells co-transfected with empty *pVLX-IRES-mCherry* and



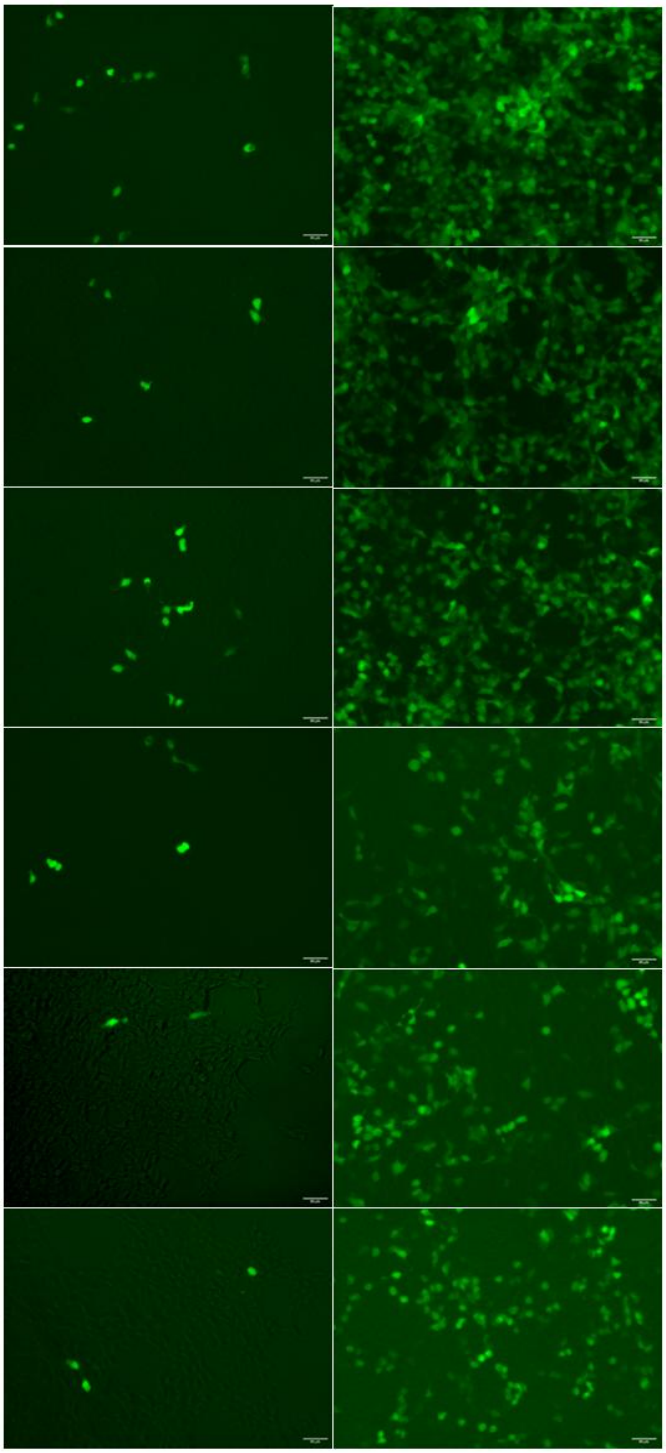
Cells visualised by 20X magnification. Scale bar 50µm. Images taken at 48 hours post-transfection. (A) Transfection with empty *pVLX-IRES-mCherry*. (B) Transfection with empty *pVLX-IRES-mCherry* and *p8.91.*

IAV pseudotype production

GFP-expressing H5 pseudotype particles

Now that B4GALNT2 genes had been cloned and could be reasonably expression in target cells, in order to address the primary aim of this study, H5 pseudotypes expressing reporter green fluorescent protein (GFP) needed to be generated. 293T cells were transfected with *p8.91* (packaging plasmid), *pVAX1-BMP2-H5HA* (Influenza A virus H5 haemagglutinin), *pHAT* (a protease plasmid used to cleave the HA precursor into the functional subunits required for infection (Wang *et al*., 2008)) and *pLenti-GFP-Puro* (GFP reporter transgene expression plasmid). VSV-G pseudotype particles delivering GFP reporters were also produced as a control throughout both transfection and transduction. Following production, two-fold serial dilutions of H5-GFP or VSV-G-GFP pseudotype particles were applied to 293T cells, 50% being the highest dilution, and 1.5625% the lowest. 293Ts transduced with VSV-G-pseudotyped particles showed significant GFP expression, whereas those transduced with H5 pseudotyped particles demonstrated a marked decrease in GFP fluorescent signal (Figure 13). During infection assays, 293T cells transduced with H5-GFP pseudotype particles were subjected to spinoculation, which appeared to considerably improve infectivity (data not shown).

Figure 13: H5-GFP pseudotype production



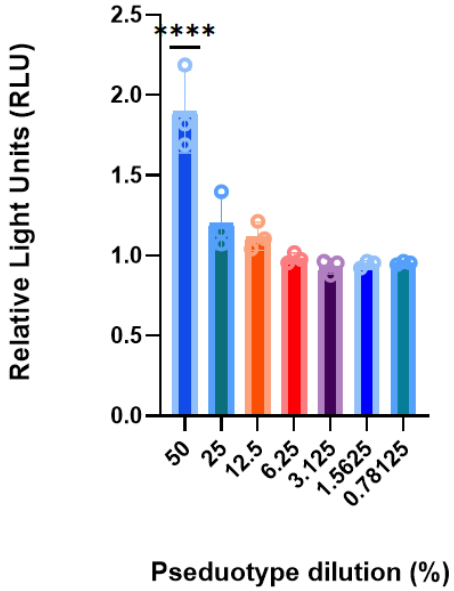
Cells visualised with 20X magnification. Scale bar 50µm. 2-fold serial dilution of H5-GFP pseudotype (left panels) and VSV-GFP pseudotype (right panels) Pseudotype particles diluted from 1:2 ratio (top) to 1:64 ratio (bottom).

Bat-origin H9 vector construction and transduction attempts

As outlined in the introduction, *R. aegytiacus* are the known carriers of a novel bat H9N2 IAV strain. Therefore, we wanted to test our candidate B4GALNT2 proteins against lentiviral pseudotypes expressing the haemagglutinin gene (H9) of A/bat/Egypt/381OP/2017(H9N2). To construct the plasmid, we first attempted to insert the HA gene into the *pVAX1-BMP2* vector as this backbone successfully supports the production of H5 pseudotypes. We were unable to obtain the empty *pVAX1-BMP2* vector, thus DNA gel extraction was used to isolate the linearised vector backbone from the *pVAX1-BMP2-H5HA* plasmid. Despite testing several DNA gel extraction kits, DNA yields from the excised bands were negligible (<10µg/µL). Consequently, an alternative vector backbone (*pCMV-Myc*) was used to express the H9 gene of A/bat/Egypt/381OP/2017(H9N2). Following assembly, we successfully isolated the *pCMV-Myc-H9HA* plasmid to be used for pseudotype production.

To generate the pseudotyped particles, *pCMV-Myc-H9HA* was co-transfected with *p8.91*, *pHAT* and *pLenti-GFP-Puro* or *pCSFLW* according to the optimised protocol used to produce H5 pseudotype virus. Transduction of the luciferase-expressing pseudotype at a 50% dilution resulted in a 1.8-fold increase in luminescence signal over background (Figure 14). Despite also transducing 293T cells with neat GFP-expressing H9 pseudotype particles, no fluorescent signal was observed at any dilution (data not shown). In attempts to improve infectivity, titrations of *pCMV-Myc-H9HA* and *pHAT* plasmids were tested, and exogenous NA was reduced or excluded. However, no modification tested resulted in pseudotyped particles expressing GFP signal upon transduction into 293T cells (data not shown). Although bat H9 was successfully cloned into a vector that appeared appropriate for pseudotype production, we were unable to produce pseudotype particles that could infect 293T cells. Thus, further analysis of bat H9 was not pursued in this study.

Figure 14: H9 pseudotype transduction

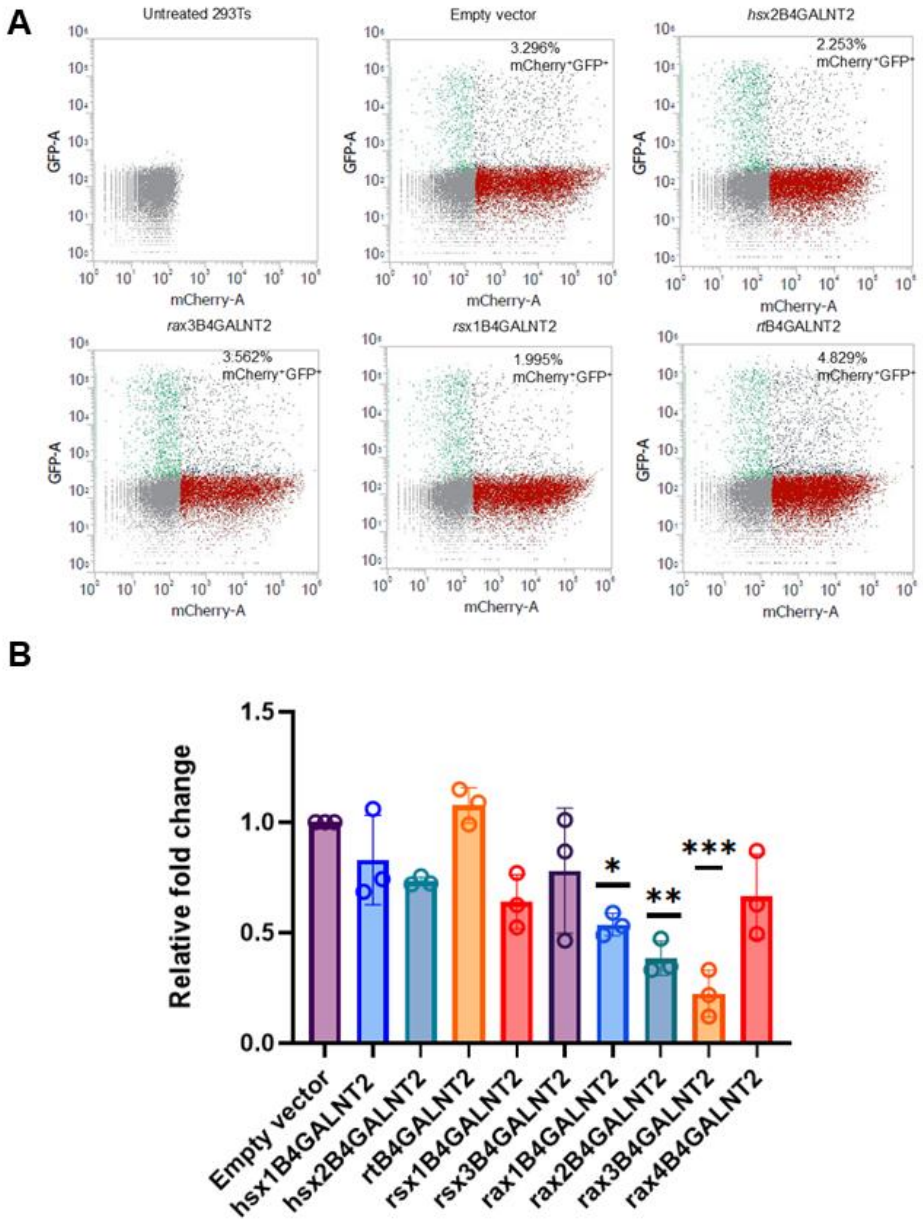


Luciferase assay of different H9 pseudotype dilutions in 293T cells performed 48-post transduction. Results normalised to non-transduced control. Three technical replicates conducted for each test group. Statistical significance indicated using asterisks, where **** p <0.0001.

R. aegyptiacus B4GALNT2 alleles inhibit H5 pseudotype infection

Our primary aim when undertaking this analysis was to investigate whether our candidate bat B4GALNT2 genes display antiviral activity against avian influenza pseudotypes. To test this aim, cells overexpressing B4GALNT2 were transduced with H5 pseudotype particles that would deliver reporter GFP. 293T cells transfected with empty *pVLX-IRES-mCherry* vectors and subsequently transduced with H5-GFP pseudotyped virus served as a control (Figure 15A). The majority of B4GALNT2 genes were tested at an MOI of ~0.08. However, *rax2*B4GALNT2 and *rax3*B4GALNT2 were initially tested at an MOI 0.01. Consequently, we repeated the assay for these genes at a higher MOI (0.11) and reported these results in our analysis.

Figure 15: H5-GFP pseudotype infection results



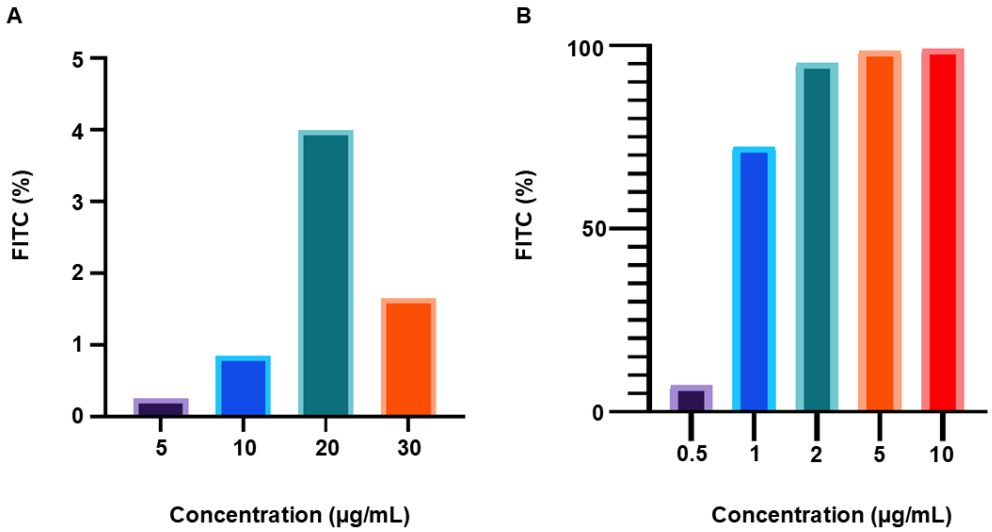
293T cells overexpressing *pVLX-IRES-mCherry-B4GALNT2* constructs transduced with H5-GFP pseudotype particles. At 48 hours post-transduction, mCherry and GFP expression was measured by flow cytometry. (A) Representative flow cytometry plots. Analysis performed using Attune™ NxT software. B) Effect of human and bat B4GALNT2 expression on H5 pseudotype infection. All data is presented relative to their respective empty vector control. The control values are normalised to 1. Three technical replicates conducted for each test group. Statistical significance indicated using asterisks, where *p <0.05, ** p <0.01, and *** p <0.001.

Human B4GALNT2 alleles demonstrated very weak antiviral activity, restricting H5 pseudotype entry by less than 2-fold. *R. aegyptiacus* B4GALNT2 alleles showed the highest levels of inhibition of all genes tested, with *rax3*B4GALNT2 the most potent, reducing infection by almost 5-fold. However, significant variation in antiviral activity between *R. aegyptiacus* alleles was observed. Conversely, both *R. sinicus* alleles and *R. trifoliatus* do not exhibit antiviral activity, appearing to have no effect on H5 pseudotype infection (Figure 15B). Overall, no inhibitory phenotype was observed for human, *R. trifoliatus* and *R. sinicus* B4GALNT2 alleles, with a significantly stronger inhibitory effect observed in cells over expressing *ra*B4GALNT2 alleles.

Optimisation of lectin probe analysis to measure α2,3 and α2,6 sialic acid binding

Following the evidence of *R. aegyptiacus* B4GALNT2-mediated viral inhibition, we wanted to assess whether changes in α2,3-sialic acid receptor binding were responsible for this enhanced phenotype. 293T cells overexpressing empty *pLVX-IRES-mCherry* were treated with lectin probes that bind to either α2,3- or α2,6-sialic acid linkages. At 10µg/mL, α2,6 probe binding was >99%, suggesting a high rate of non-specificity (>99% cells FITC-positive), whereas α2,3 probe binding was negligible (<1% cells FITC-positive) (data not shown). As MDCK cells are known to express higher levels of α2,3-linked sialic acid receptors, we attempted to optimise the experiment in this cell line without prior transfection of *pLVX-IRES-mCherry* (Barnard *et al.,* 2021; Nelson *et al*., 2019). Despite the change in cell line, α2,3 probe binding remained <1% and α2,6 probe binding >99%. Consequently, probe concentration was titrated in 293T cells (Figure 16). 20µg/mL of the α2,3 probe resulted in the highest binding observed (3.99% cells FITC-positive), whereas near saturation of α2,6-linked sialic acid binding sites was achieved using 2µg/mL of the α2,6 probe (96.7% cells FITC-positive) (Figure 16). Therefore, these concentrations were identified as the most appropriate to identify B4GALNT2-mediated changes in sialic acid binding.

Figure 16: Lectin probe optimisation



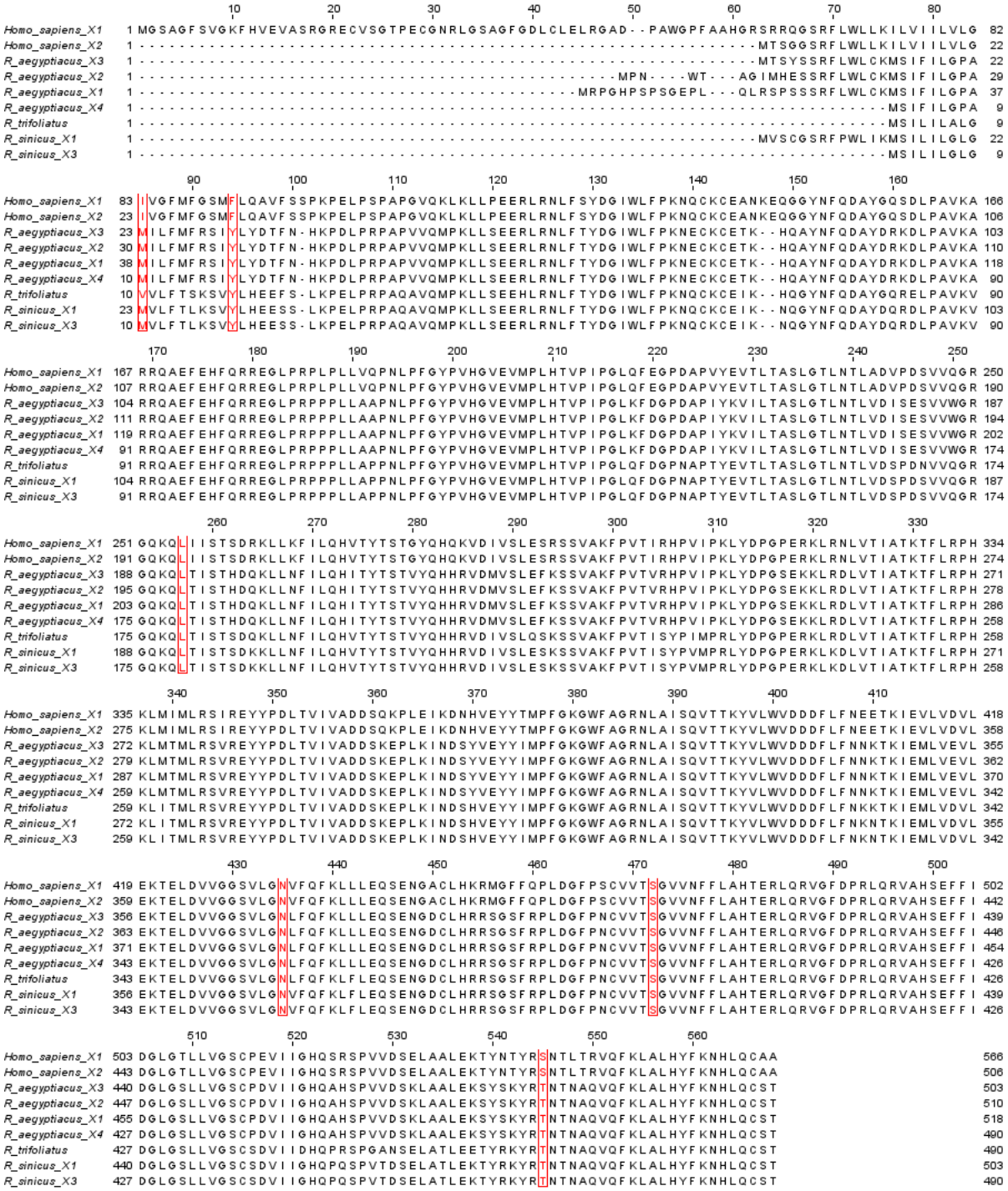
pVLX-IRES-mCherry expressing 293T cells treated with varying concentrations of FITC-conjugated Sambucus nigra agglutinin (SNA) and FITC-conjugated Maackia amurensis lectin (MAL II). (A) Cells treated with Maackia amurensis lectin (MAL II). (B) Cells treated with FITC-conjugated Sambucus nigra agglutinin (SNA).

To ensure the lectin probes were binding specifically to α2,3- and α2,6-linked sialic acids, a negative control was included. 0.1-0.25U/mL of each sialidase was incubated with cells for 30 minutes to 1 hour. No concentration tested appeared to remove sialic acid receptors as binding with respective probes remained unaffected. Following a previously published protocol, incubation time with each sialidase was increased to 20 hours, but this also did not appear alter probe binding (Ishigaki & Itoh, 2022). Thus, despite some of the B4GALNT2 alleles disrupting H5 IAV transduction, we were unable to relate this data to pattens of sialic acid expression.

Identification of positively selected codons in B4GALNT2 alleles

Given the evidence of antiviral activity associated with the expression of some bat B4GALNT2 variants, we carried out several analyses to identify specific sites subject to selection that may be important for B4GALNT2 functionality. Using the alignment of 115 mammalian species B4GALNT2 genes, we employed three methods from the Datamonkey webserver (Weaver *et al*., 2018). MEME (Mixed Effects Model of Evolution) detected a total of 28 sites ($p < 0.05$), whereas the FEL model (Fixed Effect Likelihood) detected 12 sites ($p < 0.05$) and the FUBAR model only 6 sites (posterior probability <0.9). MEME identified the highest number of sites of all methods used as the model detects both episodic and pervasive diversifying selection and can identify selection limited to only a few branches of the phylogenetic tree (Murrell *et al*., 2012). Conversely, FEL and FUBAR identify instances of pervasive selection across the entire phylogeny (Murrell *et al*., 2013; Kosakovsky Pond *et al*., 2021). In our analysis, only sites detected by FUBAR are reported (Figure 17).

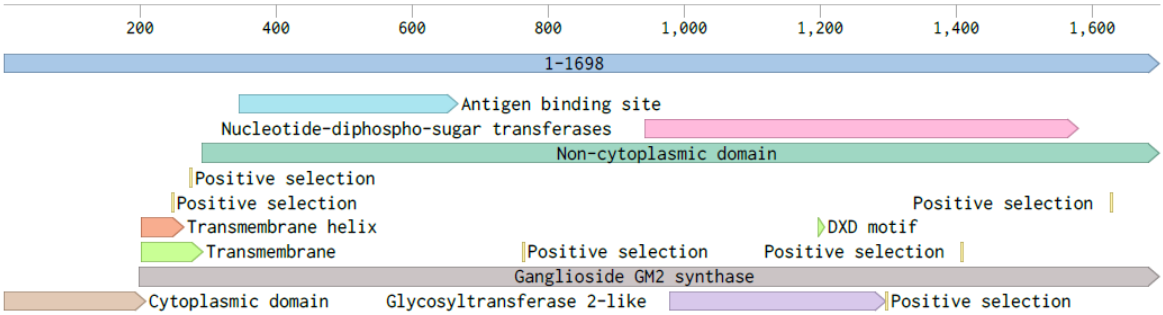
Figure 17: Amino acid alignment of candidate B4GALNT2 alleles, highlighting residues under positive selection



Alignment created using ClustalW on MegaX and visualised on Jalview. Residues identified by FUBAR analysis are highlighted in red.

The majority of sites identified to be under positive selection appear to be concentrated in the transmembrane domain or the C-terminal region (Figure 18). Sites 432 and 472 are conserved amongst all candidate B4GALNT2 alleles; however, every other site identified varies between bats and humans (Figure 17). Two sites are recognised to be within the N-terminal domain, with selection at site 85 and 95, in the reported transmembrane region (Figure 18). Selection at site 85 is the most varied of all sites identified. At this site, human isoforms express an isoleucine (I) residue whereas all *R. sinicus* and *R. aegyptiacus* alleles express a methionine (M). Conversely, *R. trifoliatus* B4GALNT2 encodes a valine (V) at this position (Figure 15). Selection at sites 95 and 545 differs between human and bat alleles yet is conserved amongst all bat alleles tested (Figure 17). Collectively, this data identifies several residues that may be important for B4GALNT2 antiviral functionality and could be of value in future analysis.

Figure 18: Reported domain structure of human B4GALNT2 X1



Domains identified using UniProt.org and GeneCards.com. Residues reported to be under selection from FUBAR analysis are denoted by yellow tabs. Schematic created with Benchling.

Discussion

The overall aim of this study was to determine if B4GALNT2 may be evolving in certain bat species to limit viral infection. The frequent use of sialic acid binding by viruses suggests that many genes involved in these processes may be subject to host-pathogen genetic conflict and are being selected to tolerate viral infection (Valero-Rello *et al.*, 2024). A large pan-mammalian and bat enriched dataset published by the Bat1K consortium has provided a novel method for studying adaptations that may be important for bat antiviral immunity (Morales *et al.*, 2025). Given the recent discovery of a novel subset of IAVs circulating in bat species, we used this dataset to identify sialic-acid regulatory genes under positive selection in bats.

Positive selection of B4GALNT2 has been reported across the Chiropteran order, with selection being driven by two bat species (Morales *et al.*, 2025). Throughout this study we aimed to assess whether positive selection may improve bat B4GALNT2 antiviral activity. To address this hypothesis, we cloned several bat species B4GALNT2 genes and expressed them in mammalian cells to assess their antiviral functionality against influenza A virus H5 pseudotype infection. In addition to our primary aims, we further identified various amino acids that may be important for B4GALNT2 antiviral functionality, and we successfully cloned the haemagglutinin gene from bat-origin H9N2 into an expression vector for future generation of bat-derived H9 pseudotyped particles.

While our analysis was unable to confirm the potent inhibition of influenza A virus by human B4GALNT2, we demonstrate the first instance of viral suppression by non-human B4GALNT2 proteins in Egyptian fruit bats (*R. aegyptiacus*). Bat B4GALNT2, although weak or modest in some instances, more potently inhibited avian IAV pseudotype infection than human B4GALNT2. Moreover, we found that positive selection of *R. trifoliatus* B4GALNT2 does not exhibit enhanced restriction of H5 pseudotype entry as we originally postulated.

Technical challenges in *pVLX-IRES-mCherry-B4GALNT2* plasmid construction

Vector linearisation

To assess the antiviral functionality of bat B4GALNT2 against GFP-expressing avian IAV pseudotypes, we first cloned the genes into a bicistronic lentiviral expression vector using HiFi assembly. As previously stated, several issues were encountered while trying to isolate *pLVX-IRES-mCherry* clones expressing B4GALNT2 inserts. Such difficulties likely arose from the empty *pLVX-IRES-mCherry* vector harbouring a mutated multiple cloning site (MCS). As only three cut sites were available, we were required to use BamHI and BsmBI recognition sequences to linearise the empty vector. BsmBI-v2 and BamHI-HF restriction enzymes exhibit variations in optimal activity in all NEB-supplied buffers. Given our buffer availability, we were required to proceed with rCutsmart buffer, in which BamHI-HF demonstrates 100% efficiency but BsmBI-v2 only demonstrates 25% efficiency (NEB., no date b). Consequently, the incubation time with BsmBI-v2 was extended in attempts to improve digestion. Nonetheless, the lack of bacterial clones containing B4GALNT2 is likely the result of poor vector linearisation prior to the assembly reaction.

These issues have highlighted the importance of assessing restriction enzyme compatibility during the experimental design process. We designed and received HiFi primers prior to identifying the disparities between the cutting efficiency of BamHI and BsmBI. When using this vector in future plasmid assemblies, only the BamHI recognition sequence should be used for linearisation. Alternatively, CRISPR/Cas9 gene editing could be employed to precisely cut the vector (Madsen *et al.*, 2025). Utilising CRISPR/Cas9 in this manner would avoid the restrictions imposed by the mutated MCS, ensuring complete digestion prior to assembly.

Improvements to HiFi assembly efficiency

Further issues arose when verifying B4GALNT2 expression and orientation using restriction digest. Many genes appeared to be partially or incorrectly incorporated in the vector backbone, suggesting a low rate of successful HiFi assembly. HiFi and Gibson assembly master mixes are typically composed of three enzymes, including T5 exonuclease, which is used to create the 3' overhangs necessary for assembly. T5 exonuclease overactivity can result in the cleaved single-stranded 3' overhangs being damaged or removed (Rabe & Cepko, 2020; Sayers & Eckstein, 1991). Addition of a single-stranded DNA binding protein (e.g. ET SSB) to Gibson assembly reactions appears to protect single-stranded 3' DNA ends from additional cleavage by T5 exonucleases, improving assembly efficiency by approximately 60% (Rabe & Cepko, 2020). Alternatively, extending the length of complementary overhangs may protect against any additional cleavage by T5 exonucleases. For instance, an increase in overhang length from 15 base pairs to 20+ base pairs has been demonstrated to significantly improve assembly reaction efficiency (NEB, no date a).

pLVX-IRES-mCherry-B4GALNT2 expression in mammalian cell lines

Attempting to transfect the *pVLX-IRES-mCherry-B4GALNT2* constructs proved to be the rate limiting step in this project. mCherry is 43% dimmer than that of enhanced GFP (eGFP) (Chen *et al.,* 2023). Consequently, mCherry fluorescence proved to be more difficult and time consuming to visualise even during optimised transfections. The extended time required to detect mCherry expression during initial transfection and VSV-G-mediated transduction attempts may have inadvertently resulted in photobleaching, causing us to overlook weak fluorescence or believe no mCherry signal was produced (Remmel, 2024). Nevertheless, issues surrounding mCherry visualisation likely also resulted from poor transfection efficiency.

Vector size plays an important role in plasmid DNA transfection. Both chemical transfection and electroporation of larger plasmids negatively impacts cell survivability and DNA uptake (Lesueur *et al*., 2016; McLenachan *et al*., 2007). The decreased transfection efficiency associated with larger DNA constructs may be due to reduced trafficking to the nucleus for gene expression (McLenachan *et al*., 2007). Given the *pLVX-IRES-mCherry-B4GALNT2* constructs ranged from 9.5kb to 9.7kb, it is possible that construct size was limiting transfection efficiency.

Following numerous unsuccessful attempts to improve gene uptake by altering various experimental parameters, mCherry fluorescence was eventually visualised during co-transfection with *p8.91*. The improved gene delivery observed with co-transfection of *p8.91* implied the lentiviral backbone of *pLVX-IRES-mCherry* may benefit from HIV-1 Gag and/or Pol expression. However, such a phenomenon has never been reported, suggesting the beneficial effect was likely related to an increase in extra chromosomal DNA. Previous research has demonstrated that co-transfection of an additional smaller plasmid can increase efficiency by up to 40% (Søndergaard *et al*., 2020). Despite the observed increase in transfection, co-expression of Gag and Pol during downstream assays may have resulted in off-target biological effects, such as activation of host immune pathways (Wang *et al.,* 2023). Given the aim of this study, such effects may have misrepresented the antiviral effect of B4GALNT2 expression if we persisted with Gag/Pol expression as a transfection rescue.

Fortunately, improved transfection was ultimately achieved without *p8.91* co-expression following a change in cell batch. Various cellular factors can affect transfection efficiency, including cell health, confluency and passage number. Cells of higher passage number demonstrate slower growth rates and exhibit significant morphological and epigenetic changes that can affect DNA uptake (Briske-Anderson *et al*., 1997). During transfection, the use of cells of an early passage number can improve efficiency by up to 30% compared to that of later passages (de Los Milagros Bassani Molinas *et al*., 2013).

As outlined, the transfection troubleshooting took a significant amount of time during the project, limiting time available for downstream infection experiments. Various techniques or other transfection approaches could have been trialled to improve DNA uptake. For instance, plate centrifugation following transfection can increase efficiency by approximately 5.46-fold and the use of electroporation for large plasmids can yield cellular uptake of 90-99% (Yang *et al.*, 2024; Lesueur *et al.*, 2016). Alternatively, in future, a different bicistronic expression vector could be used. Other students within the lab reported the successful cloning, transfection, and visualisation of *pLVX-IRES-ZsGreen*, which required minimal troubleshooting (C. Hird, personal communication). Though if this vector was used to express the B4GALNT2 genes, a different fluorescent protein would need to be incorporated into IAV pseudotypes which may cause similar visualisation issues.

Bat-origin H9 lentiviral pseudotype production

In addition to our primary aims, we planned to test the bat B4GALNT2 genes against a bat-origin IAV pseudotype. At the time of writing, no published studies have reported the construction or use of a A/bat/Egypt/381OP/2017 pseudotype. We successfully cloned the HA gene of A/bat/Egypt/381OP/2017 into a suitable expression vector but were unable to produce pseudotype particles for infection assays.

Our unsuccessful attempts at modifying envelope and protease plasmid ratios suggested pseudotyped particles may require an alternative extracellular protease or cleavage post-production to produce infective particles. The HA precursor (HA_0) of replication-competent and pseudotyped IAV particles requires cleavage to form the surface (HA_1) and transmembrane (HA_2) HA subunits that facilitate receptor binding (Wang *et al.*, 2008). Human airway trypsin-like protease (HAT) has been previously used within our lab to create infective H5 pseudotypes and has been reported to cleave various other IAV pseudotypes yet, no concentration of HAT tested appeared to cleave the A/bat/Egypt/381OP/2017 pseudotype (Wang *et al.*, 2008; Sawoo *et al.*, 2014; Delgadillo *et al.*, 2022). Various other human proteases can cleave replication competent viruses and pseudotyped IAV particles. For instance, replication competent avian-origin H9N2 strains can be cleaved with a range of proteases,

including, HAT, transmembrane protease, serine S1 member 2 (TMPRSS2) and matriptase. However, simultaneous use of HAT and TMPRSS2 in avian H9N2 production promotes multicycle replication and results in an increase in viral protein expression over use of only one protease plasmid (Baron *et al.*, 2013). Thus, the use of an alternative protease or combination of such may facilitate A/bat/Egypt/381OP/2017 pseudotype cleavage. Addition of exogenous trypsin can also effectively cleave HA_0 into HA_1 and HA_2 subunits (Böttcher-Friebertshäuser, *et al.*, 2011). Attempts could also be made to cleave the H9 pseudo particles post-production with trypsin-TPCK.

Alternatively, it is possible that infective H9 pseudotyped particles were being produced and limitations associated with cellular transduction may have limited infection. As mentioned, bat-origin H9N2 enters through α2,3 sialic acid receptors, consequently, attempting to transduce cells with higher basal expression of α2,3-linked receptors, such as MDCK or A549 cell lines, may promote A/bat/Egypt/381OP/2017 pseudotype entry. Nevertheless, transduction efficiency of MDCK and A549 with lentiviral-based influenza pseudotypes has been reported to be lower than that of 293T cells (Temperton *et al.*, 2007; Ferrara *et al.*, 2021; Ao *et al.*, 2008).

Testing our genes against a bat-origin IAV is imperative, if the alternative methods outlined above does not successfully produce infective A/bat/Egypt/381OP/2017 pseudotypes, cloning in an alternative vector backbone with a known ability to produce infective IAV pseudotyped particles should be considered, such as *pCAGGS* or *pVAX1-BMP2* (Ferrara *et al.*, 2021). Construction of a replication-deficient influenza strain expressing the bat-origin haemagglutinin gene could also be attempted. The use of six internal genes from A/Puerto Rico/8/1934 (PR8) combined with the haemagglutinin genes from avian-origin H5 and H7 strains have generated low virulence viruses that are unable to replicate when intranasally administered to mice (Tian *et al.*, 2021). Such recombinant viruses can be engineered to stably express a reporter gene, such as GFP (Victor *et al.*, 2012; Ozawa *et al.*, 2011). Although PR8 is considered a biosafety level 2 (BSL-2) virus, recombinant viruses containing genes or segments from highly pathogenic and emerging IAV strains, such as HPAI H5N1, are considered BSL-3 (GOV UK, 2025; NIH, 2024). Given bats established role in viral emergence and the cross-species

transmission potential of A/bat/Egypt/381OP/2017, creating a recombinant strain with features of avian IAVs may inadvertently create a novel IAV strain. Consequently, the construction of such a pseudotype would likely require a higher level of safety and containment, and would require a strong justification for study, making its generation challenging (GOV UK, 2025).

Bat B4GALNT2 demonstrates variations in inhibitory activity against H5 pseudotype infection

This work demonstrates the first instance of non-human B4GALNT2-mediated inhibition of IAV. *R. aegyptiacus* B4GALNT2 alleles demonstrate the highest levels of inhibition of all B4GALNT2 alleles tested. However, cells expressing *R. sinicus* and human alleles appear to be almost unaffected by H5 pseudotype infection, and contrary to our hypothesis, *R. trifoliatus* B4GALNT2 appears to have no influence on H5 pseudotype transduction.

The differences in B4GALNT2-mediated restriction between bat species could potentially be explained by their distinct viral carriage. The association between *R. aegyptiacus* and A/bat/Egypt/381OP/2017 H9N2 could have exerted a selection pressure, driving the enhanced antiviral activity of genes involved in sialic acid regulation, such as B4GALNT2. However, this hypothesis would require a consistent virus-driven pressure in *R. aegyptiacus*, and currently the full extent and timescale of H9N2 infection in this bat population remains unclear (Karamendin *et al.*, 2024; Yang *et al.*, 2021). Moreover, *R. aegyptiacus* B4GALNT2 did not emerge in the positive selection analysis, indicating various immunoregulatory genes or mechanisms may be governing their viral tolerance. Bat species immune systems have likely evolved to be universally antiviral, rather than in response to specific pathogens. For instance, many bats constitutively express immune genes such as IFNs and ISGs which are broadly antiviral (Li *et al.*, 2024; Schountz *et al.*, 2017). *R. sinicus* carries a range of viruses, including SARS-related coronaviruses, chikungunya virus, Japanese encephalitis virus and even a novel poxvirus (Peng *et al.*, 2022; Wang *et al.*, 2024). Both Japanese encephalitis virus and SARS-CoV-2 infection utilise sialic acids as viral attachment factors (He *et al.*, 2024; Baker *et al.*, 2020). Therefore, the inhibition mediated by bat B4GALNT2 genes may not be limited to avian IAVs but any viruses that utilise α2,3-linked receptors for attachment or entry. Previously,

human B4GALNT2 expression has been demonstrated to not only reduce cellular infection with avian IAV, but also paramyxoviruses that preferentially bind α2,3-linked sialic acid receptors (Park *et al.*, 2023; Wu *et al.*, 2025). Evidence implicating *R. trifoliatus* as a viral reservoir is limited. The absence of a viral selection pressure in *R. trifoliatus* may be driving the lack of B4GALNT2-mediated inhibition observed, yet this is more likely due to limited sequencing efforts. Between 2018 and 2022, 72 novel viruses were isolated from bat various species; however only 8 bat families featured in the analyses (Mo *et al.*, 2024). As bat metagenomic analyses improve to cover more species, we will likely observe an increase in viruses found in bat species previously unknown to be reservoirs, such as *R. trifoliatus*. Future studies should endeavour to test bat B4GALNT2 genes against a broader range of α2,3-binding viruses, such as further avian-origin influenza viruses and paramyxoviruses, but also bat-origin H9N2. This will determine whether bat B4GALNT2 genes exhibit enhanced potency against bat-origin α2,3-binding IAVs, or if their antiviral activity is conserved across diverse α2,3-binding viruses.

Effect of alternative splicing on *R. aegyptiacus* B4GALNT2

The variation in antiviral activity demonstrated across *R. aegypticaus* B4GALNT2 alleles suggests alternative splicing may have been employed to create variants with differential functionality. Positively selected genes, proviral factors and restriction factors are more likely to be spliced (unpublished data from within our lab group). *R. sinicus* IFITM3 is known to undergo alternative splicing to create two distinct alleles with the same structure but varied antiviral activity against a range of viruses including IAV, Nipah virus, SARS- and MERS-related coronaviruses (Mak *et al.*, 2024). Similarly, OAS2 can encode two variants with differential activity against human seasonal coronavirus OC43 (HCoV-OC43) and picornavirus Cardiovirus A (EMCV), with this antiviral variation dependent on changes observed in the C-terminal region of each isoform (Davies *et al.*, 2025). *R. aegyptiacus* may be utilising a similar mechanism to create a repertoire of B4GALNT2 alleles to prevent α2,3-mediated viral entry, while maintaining core physiological enzymatic function. Testing each of the alleles against opposing α2,3-binding viruses, such as A/bat/Egypt/381OP/2017 H9N2, NDV and Human respirovirus 3 (hPIV3) may undercover whether variants have arisen with distinct antiviral preferences. Future

studies of antiviral genes should endeavour to cover all splice variants given the functional diversity observed in this study and others. Sequence databases describing a group of splice variants as a single, canonical variant likely misrepresents the complexity of an mRNA population. Such populations may have alternative functionalities depending on tissue- and cell-dependent contexts.

While it is currently unknown whether *R. trifoliatus* B4GALT2 undergoes alternative splicing to generate opposing transcript variants, all other bat B4GALNT2 genes studied in this analysis have multiple variants. Given the aforementioned role of alternative splicing in innate immunity regulation and diversification, uncovering novel transcript variants may provide further understanding of the enhanced viral tolerance in bats (Mak *et al*., 2024; Davies *et al*., 2025). Ongoing efforts by the Bat1K consortium to sequence every bat species genome may result in the coding sequences of alternative B4GALNT2 transcripts becoming available in the near future ((Teeling *et al*., 2018). At which time, the antiviral activity of all alleles should be assessed to understand if the phenotypes observed in this study are conserved across all variants.

All splice variants of the same B4GALNT2 gene vary only in the N-terminal region and exhibit differential antiviral activity against H5 pseudotype infection, indicating this region may be important for antiviral function. As mentioned, the extended N-terminal domain of *hs*x1B4GALNT2 is required for its Golgi localisation but expression at the plasma membrane (Groux-Degroote *et al*., 2018). Moreover, the N-terminal region of IFITM3 has previously been shown to be essential for its endosomal localisation and subsequent antiviral activity against IAV (Jia *et al*., 2012). To investigate whether the N-terminal region of bat B4GALNT2 influences subcellular localisation, confocal microscopy should be used. Changes to sub-cellular localisation may affect B4GALNT2's interaction with α2,3 sialic acid-containing glycans, explaining the patterns of antiviral restriction observed.

H5 pseudotype infection assay limitations

Our results demonstrate a significantly weaker antiviral phenotype than that of previous published data. Initial analysis by Heaton *et al*. demonstrated a 100-fold reduction in viral replication when using replication competent H9N2, H10N4, H7N2 and H5N9 IAV strains in plaque assays, whereas our results demonstrate an extremely weak and nearly entirely absent 1.2-fold and 1.32-fold reduction (Heaton *et al*., 2017). It is possible that pseudotype infection may not encapsulate all the molecular mechanisms required for viral binding and B4GALNT2-mediated antiviral activity. Additionally, it could be that B4GALNT2 antiviral activity is not entirely dependent on its enzymatic activity, and further B4GALNT2-mediated mechanisms may be acting to inhibit viral replication at later infection stages. Nevertheless, to address the disparity in results, we should perform plaque assays using replication competent H9N2, H10N4, H7N2 and H5N9 avian IAV strains (Heaton *et al*., 2017). Additional testing using replication competent IAVs will also allow us to confirm the antiviral phenotype of *R. aegyptiacus* B4GALNT2 is not restricted to pseudotype infection. However, testing live bat-origin IAV strains, such as A/bat/Egypt/381OP/2017 would likely require a higher BSL given bat species' extensive zoonotic history.

The increased MOI used when testing *rax2*B4GALNT2 and *rax3*B4GALNT2 may have saturated the antiviral activity of B4GALNT2 gene expression. However, the pattern of inhibition observed remained consistent across both MOIs tested, indicating MOI did not significantly impact B4GALNT2-mediated antiviral activity. In future, testing all B4GALNT2 alleles at a range of MOIs would ensure the consistency of results.

Lectin probe analysis to measure α2,3- and α2,6- sialic acid receptor binding

Given the evidence of antiviral restriction by bat B4GALNT2, we wanted to assess whether the enhanced restriction or lack thereof is mediated by changes to α2,3 sialic receptor binding. Previously, biotinylated lectin probe analysis has demonstrated human B4GALNT2 prevents α2,3-sialic acid receptor binding by promoting Sd^a antigen formation (Wong *et al.*, 2019) Although we managed to optimise the concentration of lectin probes needed to bind to α2,3- or α2,6 sialic acid linkages, time constraints on the project did not allow us to assess the impact of B4GALNT2 on sialic acid receptor expression. Future work should begin by using the concentrations outlined in this study to complete the assay. Using an alternative cell line with increased α2,3 sialic acid receptor expression, such as A549 or MDCK may improve the specificity of α2,3 probe binding and provide a more accurate model for studying the antiviral effect of B4GALNT2. Future assays should endeavour to include an appropriate negative control. As we were seemingly unable to remove sialic acids using the relevant sialidases, performing a western blot would indicate whether efforts to remove sialic acids were successful and confirm if the probes were binding specifically to α2,3- and α2,6-linked receptors.

Identification of codons under positive selection in mammalian B4GALNT2

Following initial confirmation of bat B4GALNT2 antiviral activity, we used a combination of MEME, FEL and FUBAR models to detect specific residues displaying signatures of positive selection (Murrell *et al.*, 2012; Murrell *et al.*, 2013; Kosakovsky Pond *et al.*, 2021). We reported the residues identified by FUBAR as the Bayesian framework ensures a high rate of specificity and this model appears to be the most stringent of the three models used in this analysis (Le Corf *et al.*, 2025; Murrell *et al.* 2013). Moreover, all residues identified by FUBAR were detected by MEME and FEL models, suggesting a low false positive rate.

Residues reported under positive selection by our analysis are primarily located in the transmembrane and C-terminal regions of B4GALNT2. However, the highest level of amino acid diversity between non-functional and functional B4GALNT2 variants is located in the N-terminal region. Codons 85 and 95 are located in the N-terminal transmembrane region, indicating these residues could be important for B4GALNT2 subcellular localisation as previously mentioned (Groux-Degroote *et al.*, 2018). Codon 85 demonstrates the highest level of diversity between all species assessed in this study (Isoleucine residue expressed in humans, a methionine residue in *R. sinicus* and *R. aegyptiacus* alleles and a valine residue in *R. trifoliatus*). Substituting residues between each B4GALNT2 allele via site-directed mutagenesis and testing against IAV pseudotype infection will allow us to discern the impact of this region on the observed differences in B4GALNT2-mediated inhibition. Moreover, introducing targeted mutations in the N-terminus of B4GALNT2 may indicate whether this region is important in influencing subcellular localisation as previously mentioned. The remaining residues located in the C-terminal region were either conserved between all bat species or across all species included in the analysis, suggesting their expression does not influence the observed species-specific differences in B4GALNT2-mediated antiviral activity. Nonetheless, one study has reported the importance of two aspartic acid residues in C-terminal domain for B4GALNT2 enzymatic activity; however, this activity is not exclusive to antiviral responses (Heaton *et al.*, 2017). Similarly, various mutations located in the C-terminal region of B4GALNT2 have highlighted its requirement for enzyme function and subsequent Sd^a antigen formation (Stenfelt *et al.*, 2019). By introducing targeted changes via site-directed mutagenesis or CRISPR/Cas9 into the sites under positive selection located in the C-terminal region, we could use α2,3-binding probes to assess impact of this region on B4GALNT2 enzyme function.

Implications for determining novel antiviral genetic targets

Positive selection analyses as a method of identify genes of antiviral interest

Identifying bat immune genes with enhanced functionality is beginning to give insight into their sustained viral tolerance; however, further efforts are required to fully elucidate this unique phenotype. The dataset provided by Morales *et al*. offers a useful way to use positively selected genes to identify those that might relate to antiviral immunity. However, to specifically identify genes previously reported to be involved in coronavirus infection, the dataset was corrected using genome-wide studies of coronavirus infection outcomes (Morales *et al*., 2025). With the same thinking, we conducted an independent IAV CRISPR meta-analysis to find genes previously reported to be involved in IAV infection. In future, MAIC (Meta-Analysis by Information Content) analysis of GWAS and CRISPR screens (and other omic-level datasets) could be employed to find suitable genetic targets with better resolution or to refine hits identified by positive selection analysis (Li *et al*., 2020).

Although we chose to study B4GALNT2, several other sialic-regulatory genes were identified as subject to positive selection early in this analysis (e.g. NEU3, CD22, SIGLECL1, etc). Similarly, many host proteins related to the attachment and entry of alternative viral species, such as ANPEP and CTSB, have been identified to be under positive selection in bat species (Morales *et al*., 2025). Future studies should endeavour to assess these genes antiviral functionality in order to build a larger picture of the mechanisms governing bats' IAV carriage. Nevertheless, it remains to be understood if empirical screening or hypothesis-led approaches are more suitable for identifying novel antiviral genes. In this analysis, the highest levels of B4GALNT2-mediated inhibition was observed in *R. aegyptiacus*, a species which exhibited no signatures of positive selection. Thus, presently, we may only have the resolution to identify genes with improved functionality across an entire species, with specific tools, such as alignments and codon selection analyses, allowing species- and allele-specific inferences to be made. Such resolution will likely improve with ongoing sequencing efforts and the further omic-level analyses.

Determining the role of B4GALNT2 expression in bat immunity

During our study we had to limit the scope of bat B4GALNT2 genes that we would be able to successfully study during the timeframe of the project. At present, there is a limited number of bat B4GALNT2 coding sequences publicly available. As such, alongside the positively selected *R. trifoliatus* gene, we made the decision to study *R. aegyptiacus* and *R. sinicus* B4GALNT2. As the Bat1K consortia continues their focus on sequencing all living bat species genomes, more sequences will likely become available to study (Teeling *et al*., 2018). Investigating the B4GALNT2 genes of additional bat IAV reservoirs, such as *Sturnira lilium* and *Artibeus* species, may help determine whether B4GALNT2 acts as selection pressure on bat-origin IAVs. Pressure on IAV sialic acid binding may explain the shift from sialic acid-mediated entry towards the HLA-DR-mediated entry observed in H17N10 and H18N11 strains (Karakus *et al*., 2019). Moreover, the antiviral functionality of *Mops condylurus* B4GALNT2 should be assessed given the evidence of diversifying selection generated in this work. This will allow us the confirm if the absence of viral entry inhibition related to *R. trifolatus* B4GALNT2 expression is species-specific or if the positive selection associated with bat B4GALNT2 genes may benefit an alternative biological process.

Alternative roles for *R. trifoliatus* positive selection

The absence of *R. trifoliatus* B4GALNT2-mediated antiviral activity suggests an opposing biological process may be driving the observed genetic selection. Sialic acids play several roles in immunity, primarily influencing cell-cell, cell-matrix and cell-pathogen interactions (Varki & Gagneux, 2012). Consequently, sialic acids can activate or suppress various pro-inflammatory and metabolic pathways. For instance, α2,3-sialic binding has been linked to the expression of metabolic genes associated with ATPase and ROS expression in monocyte-derived dendritic cells (Lübbers *et al*., 2021). One of the best understood examples is the expression of the sLe^x antigen, which is important in mediating leukocyte trafficking (Zhu *et al*., 2024). Leukocyte trafficking and infiltration is one of the hallmark events in initiating an inflammatory response. Thus, the decreased α2,3-sialic acid expression facilitated by bat B4GALNT2 may be employed to alleviate the harmful metabolic inflammatory responses associated with bat-winged flight (Shen *et al*., 2010). As previously mentioned, downregulation of B4GALNT2 is associated with the loss of Sd^a antigen formation, thereby promoting the expression of glycan patterns related to tumorigenesis and metastasis. Consequently, improved bat B4GALNT2 functionality may be indirectly contributing to the enhanced cancer resistance observed in many bat species (Hua *et al*., 2024). Overall, the lack of bat B4GALNT2 antiviral functionality coupled with the evidence of selection in *R. trifoliatus* suggests bat B4GALNT2 may protect against various forms of inflammatory damage commonly observed in bats. This may be parameter evident across Chiroptera, given prior descriptions of positive selection on known immune-regulatory genes (Morales *et al*., 2025; Ahn *et al*., 2023; Xie *et al*., 2023).

Concluding remarks

Bat species' status as a recurrent source of zoonotic infections necessitates further understanding of the immune mechanisms governing their enhanced viral tolerance. Studying bat immune genes subject to positive selection has uncovered a range of genes with improved antiviral or anti-inflammatory functionality. While further exploration of such genes will provide a comprehensive view of how bats are able to exist as asymptomatic viral reservoirs, it will also aid the identification of antiviral genes that may be conserved across human and animal populations. Discovery of antiviral gene products and pathways with improved activity may provide the foundation for novel therapeutic modalities, such as in the case of bat ASC2 (Paratus, 2025). As viral binding to host receptor presents the first barrier to cross-species transmission, studying genes involved in bat receptor function is imperative to prevent further spillover events and improve pandemic preparedness.

Throughout this analysis we aimed to assess whether expression of positively selected bat B4GALNT2 genes would restrict avian IAV infection. Although our results do not support this hypothesis, this study provides the first demonstration of non-human B4GALNT2 antiviral activity. We show that bats known to harbour various α2,3-binding viruses exhibit enhanced B4GALNT2-mediated antiviral activity and that alternative splicing of B4GALNT2 genes may provide a mechanism to generate variants with enhanced antiviral functionality. Altogether, the results presented in this study will assist future research into the role of B4GALNT2 in viral tolerance and emergence.

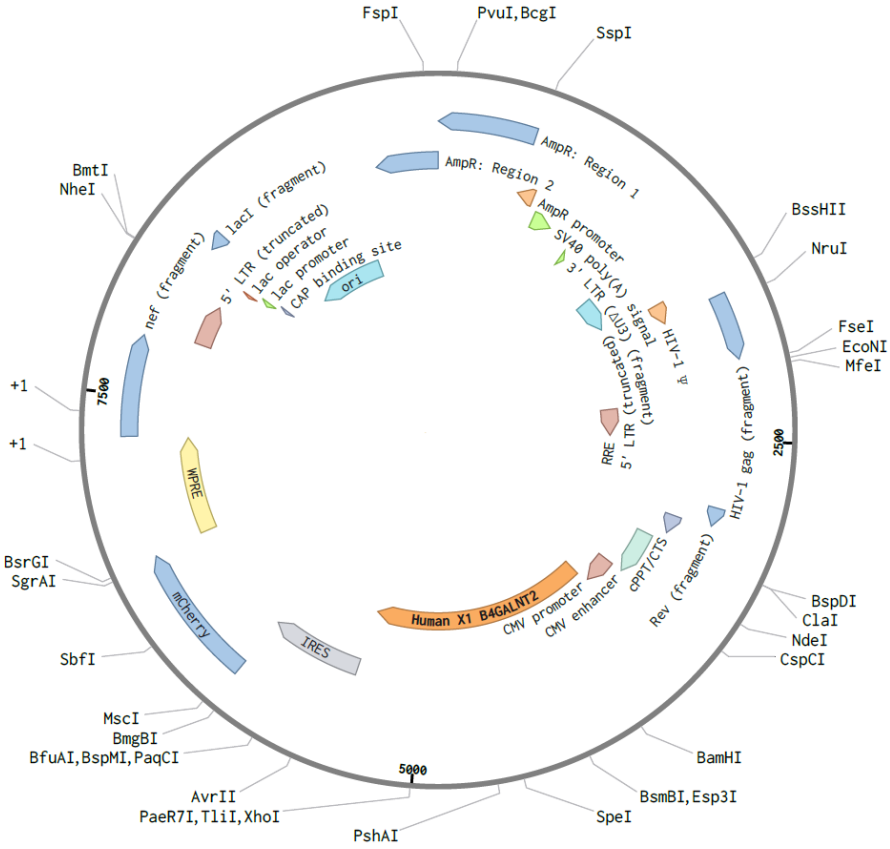
Supplementary Figures

Supplementary Table 1: Primers and annealing temperatures used in all PCRs

gBlock	Forward primer	Reverse primer	Annealing temperature
*hsx*1	gtttagtgaaccgtcagatcatggg atctgccggattcag	gggggagggagaggggcggtgc ggcgcactgcaagtg	65°C
*hsx*2	gtttagtgaaccgtcagatcatgac gagcggcggatcac	gggggagggagaggggcggtgc tgcacactgcaagtg	64°C
rt	gtttagtgaaccgtcagatcatgag tatccttatcttggcg	gggggagggagaggggcggttat gtagagcactgaagatgattttta aa	61°C
*rsx*1	gtttagtgaaccgtcagatcatggt atcatgtggttctcg	gggggagggagaggggcggtta ggtggaacactggag	59°C
*rsx*3	gtttagtgaaccgtcagatcatgag cattctgatactcgg	gggggagggagaggggcggtta ggtactgcactgcaag	61°C
*rax*1	gtttagtgaaccgtcagatcatgcg ccccgggcacccc	gggggagggagaggggcggtta cgttgaacactggaggtggtttttaaa gta	69°C
*rax*2	gtttagtgaaccgtcagatcatgcc gaactggacggcc	gggggagggagaggggcggtta cgtgctgcattggagatg	64°C
*rax*3	gtttagtgaaccgtcagatcatgac atcctattcctcaagatttc	gggggagggagaggggcggtta agttgaacattggaggtg	60°C
*rax*4	gtttagtgaaccgtcagatcatgtc aatatttatacttggtcctg	gggggagggagaggggcggtta cgttgagcattggagatg	60°C
H9	cttatggccatggaggcccgatgg aagtcaaaatttttataatacttc	tatcatgtctggatccccgcttatatgc agatagtgcagcg	59°C

Sequences highlighted in red indicate the nucleotide bases with homology to the linearised vector backbone (*pVLX-IRES-mCherry* for B4GALNT2 inserts and *pCMV-Myc* for H9 insert). Annealing temperatures were calculated by NEB's Tm Calculator.

Supplementary Figure 1: *pVLX-IRES-mCherry-hsx1B4GALNT2* plasmid map.



Whole genome sequencing (WGS) results confirming the insertion of human B4GALNT2 X1 into the *pLVX-IRES-mCherry* backbone. *pVLX-IRES-mCherry-hsx1B4GALNT2* plasmid map representative of all B4GALNT2 constructs cloned/synthesised in this study. Plasmid map visualised using Benchling.

References

Abbasi, M. and Alam, M. (2025) “Understanding the roles of viruses as key players in environmental dynamics and ecosystem functioning,” *Discover Viruses*, 2(1), p. 10. Available at: https://doi.org/10.1007/s44370-025-00015-y.

Ahn, M. *et al.* (2019) “Dampened NLRP3-mediated inflammation in bats and implications for a special viral reservoir host,” *Nature Microbiology*, 4(5), pp. 789–799. Available at: https://doi.org/10.1038/s41564-019-0371-3.

Ahn, M. *et al.* (2023) “Bat ASC2 suppresses inflammasomes and ameliorates inflammatory diseases,” *Cell*, 186(10), pp. 2144-2159.e22. Available at: https://doi.org/10.1016/j.cell.2023.03.036.

Anderson, D.E. *et al.* (2021) “Orthogonal genome-wide screens of bat cells identify MTHFD1 as a target of broad antiviral therapy,” *Proceedings of the National Academy of Sciences*, 118(39). Available at: https://doi.org/10.1073/pnas.2104759118.

Ao, Z. *et al.* (2008) “Characterization of a trypsin-dependent avian influenza H5N1-pseudotyped HIV vector system for high throughput screening of inhibitory molecules,” *Antiviral Research*, 79(1), pp. 12–18. Available at: https://doi.org/10.1016/j.antiviral.2008.02.001.

Baker, A.N. *et al.* (2020) “The SARS-COV-2 Spike Protein Binds Sialic Acids and Enables Rapid Detection in a Lateral Flow Point of Care Diagnostic Device,” *ACS Central Science*, 6(11), pp. 2046–2052. Available at: https://doi.org/10.1021/acscentsci.0c00855.

Banerjee, A. *et al.* (2020) “Positive Selection of a Serine Residue in Bat IRF3 Confers Enhanced Antiviral Protection,” *iScience*, 23(3), p. 100958. Available at: https://doi.org/10.1016/j.isci.2020.100958.

Barnard, K.N. *et al.* (2021) “Sequence dynamics of three influenza A virus strains grown in different MDCK cell lines, including those expressing different sialic acid receptors,” *Journal of Evolutionary Biology*, 34(12), pp. 1878–1900. Available at: https://doi.org/10.1111/jeb.13890.

Baron, J. *et al.* (2013) “Matriptase, HAT, and TMPRSS2 Activate the Hemagglutinin of H9N2 Influenza A Viruses,” *Journal of Virology*, 87(3), pp. 1811–1820. Available at: https://doi.org/10.1128/JVI.02320-12.

Basu Thakur, P. *et al.* (2024) “Ferrets as a Mammalian Model to Study Influenza Virus-Bacteria Interactions,” *The Journal of Infectious Diseases*, 229(2), pp. 608–615. Available at: https://doi.org/10.1093/infdis/jiad408.

Beck, M., Handy, J. and Levander, O. (2004) “Host nutritional status: the neglected virulence factor,” *Trends in Microbiology*, 12(9), pp. 417–423. Available at: https://doi.org/10.1016/j.tim.2004.07.007.

Becker, D.J. *et al.* (2023) “Ecological conditions predict the intensity of Hendra virus excretion over space and time from bat reservoir hosts,” *Ecology Letters*, 26(1), pp. 23–36. Available at: https://doi.org/10.1111/ele.14007.

Böttcher-Friebertshäuser, E. *et al.* (2011) “Inhibition of Influenza Virus Infection in Human Airway Cell Cultures by an Antisense Peptide-Conjugated Morpholino Oligomer Targeting the Hemagglutinin-Activating Protease TMPRSS2,” *Journal of Virology*, 85(4), pp. 1554–1562. Available at: https://doi.org/10.1128/JVI.01294-10.

Briske-Anderson, M.J., Finley, J.W. and Newman, S.M. (1997) “The Influence of Culture Time and Passage Number on the Morphological and Physiological Development of Caco-2 Cells,” *Experimental Biology and Medicine*, 214(3), pp. 248–257. Available at: https://doi.org/10.3181/00379727-214-44093.

Burrough, E.R. *et al.* (2024) “Highly Pathogenic Avian Influenza A(H5N1) Clade 2.3.4.4b Virus Infection in Domestic Dairy Cattle and Cats, United States, 2024,” *Emerging Infectious Diseases*, 30(7). Available at: https://doi.org/10.3201/eid3007.240508.

Caserta, L.C. *et al.* (2024) “Spillover of highly pathogenic avian influenza H5N1 virus to dairy cattle,” *Nature*, 634(8034), pp. 669–676. Available at: https://doi.org/10.1038/s41586-024-07849-4.

CDC (2025) *H5 Bird Flu: Current Situation*, *U.S. Centres for Disease Control and Prevention*. Available at: https://www.cdc.gov/bird-flu/situation-summary/index.html?cove-tab=0 (Accessed: August 6, 2025).

Charostad, J. *et al.* (2023) “A comprehensive review of highly pathogenic avian influenza (HPAI) H5N1: An imminent threat at doorstep,” *Travel Medicine and Infectious Disease*, 55, p. 102638. Available at: https://doi.org/10.1016/j.tmaid.2023.102638.

Chen, Y.-L. *et al.* (2023) “Research Progresses and Applications of Fluorescent Protein Antibodies: A Review Focusing on Nanobodies,” *International Journal of Molecular Sciences*, 24(5), p. 4307. Available at: https://doi.org/10.3390/ijms24054307.

Chia, B.S. *et al.* (2020) “Loss of the Nuclear Protein RTF2 Enhances Influenza Virus Replication,” *Journal of Virology*, 94(22). Available at: https://doi.org/10.1128/JVI.00319-20.

Chua, K.B., Chua, B.H. and Wang, C.W. (2002) “Anthropogenic deforestation, El Niño and the emergence of Nipah virus in Malaysia.,” *The Malaysian journal of pathology*, 24(1), pp. 15–21.

Ciminski, K. *et al.* (2019) “Bat influenza viruses transmit among bats but are poorly adapted to non-bat species,” *Nature Microbiology*, 4(12), pp. 2298–2309. Available at: https://doi.org/10.1038/s41564-019-0556-9.

Clontech Laboratories, Inc. (2009) *pLVX-IRES-mCherry Vector Information* , *Clontech*.

Cui, J. *et al.* (2007) “Evolutionary Relationships between Bat Coronaviruses and Their Hosts,” *Emerging Infectious Diseases*, 13(10), pp. 1526–1532. Available at: https://doi.org/10.3201/eid1310.070448.

Cutler, D.M. and Summers, L.H. (2020) “The COVID-19 Pandemic and the $16 Trillion Virus,” *JAMA*, 324(15), p. 1495. Available at: https://doi.org/10.1001/jama.2020.19759.

Dai, Y. *et al.* (2024) “Increased viral tolerance mediates by antiviral RNA interference in bat cells,” *Cell Reports*, 44(7), p. 114581. Available at: https://doi.org/10.1016/j.celrep.2024.114581.

Davies, E.L. *et al.* (2025) “Alternative splicing broadens antiviral diversity at the human *OAS2* locus.” Available at: https://doi.org/10.1101/2025.02.24.639105.

de los Milagros Bassani Molinas, M. *et al.* (2014) “Optimizing the transient transfection process of HEK-293 suspension cells for protein production by nucleotide ratio monitoring,” *Cytotechnology*, 66(3), pp. 493–514. Available at: https://doi.org/10.1007/s10616-013-9601-3.

Delgadillo-Gutiérrez, K. *et al.* (2022) “Characterization and use in neutralization assays of avian influenza codon-optimized H5 and H7 retroviral pseudotypes,” *Journal of Virological Methods*, 300, p. 114391. Available at: https://doi.org/10.1016/j.jviromet.2021.114391.

di Bari, C. *et al.* (2023) “The global burden of neglected zoonotic diseases: Current state of evidence,” *One Health*, 17, p. 100595. Available at: https://doi.org/10.1016/j.onehlt.2023.100595.

Duca, M., Malagolini, N. and Dall’Olio, F. (2023) “The story of the Sda antigen and of its cognate enzyme B4GALNT2: What is new?,” *Glycoconjugate Journal*, 40(1), pp. 123–133. Available at: https://doi.org/10.1007/s10719-022-10089-1.

Echeverri-De la Hoz, D. *et al.* (2025) “Genomics of novel influenza A virus (H18N12) in bats, Caribe Colombia,” *Scientific Reports*, 15(1), p. 6507. Available at: https://doi.org/10.1038/s41598-025-91026-8.

El-Shesheny, R. *et al.* (2024) “Cross-species spill-over potential of the H9N2 bat influenza A virus,” *Nature Communications*, 15(1), p. 3449. Available at: https://doi.org/10.1038/s41467-024-47635-4.

Felsenstein, J. (1985) “CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP,” *Evolution*, 39(4), pp. 783–791. Available at: https://doi.org/10.1111/j.1558-5646.1985.tb00420.x.

Ferrara, F. *et al.* (2021) “Development of Lentiviral Vectors Pseudotyped With Influenza B Hemagglutinins: Application in Vaccine Immunogenicity, mAb Potency, and Sero-Surveillance Studies,” *Frontiers in Immunology*, 12. Available at: https://doi.org/10.3389/fimmu.2021.661379.

Figliozzi, R.W. *et al.* (2016) “Using the inverse Poisson distribution to calculate multiplicity of infection and viral replication by a high-throughput fluorescent imaging system,” *Virologica Sinica*, 31(2), pp. 180–183. Available at: https://doi.org/10.1007/s12250-015-3662-8.

Gao, P. *et al.* (2022) “Apolipoprotein E mediates cell resistance to influenza virus infection,” *Science Advances*, 8(38). Available at: https://doi.org/10.1126/sciadv.abm6668.

Gonzalez, V. and Banerjee, A. (2022) “Molecular, ecological, and behavioral drivers of the bat-virus relationship,” *iScience*, 25(8), p. 104779. Available at: https://doi.org/10.1016/j.isci.2022.104779.

GOV UK (2025) *Avian influenza and influenza of avian origin: diagnostic testing, controls and reporting obligations*, *Animal and Plant Health Agency*. Available at: https://www.gov.uk/government/publications/listed-diseases-in-animals-case-definitions-testing-and-reporting/avian-influenza-and-influenza-of-avian-origin-diagnostic-testing-controls-and-reporting-obligations?utm_source=chatgpt.com#influenza-a-viruses (Accessed: August 12, 2025).

Grabenhorst, E. and Conradt, H.S. (1999) “The Cytoplasmic, Transmembrane, and Stem Regions of Glycosyltransferases Specify Their in Vivo Functional Sublocalization and Stability in the Golgi,” *Journal of Biological Chemistry*, 274(51), pp. 36107–36116. Available at: https://doi.org/10.1074/jbc.274.51.36107.

GraphPad (2019) *GraphPad Prism 8.3.0 Release Notes*. Available at: https://www.graphpad.com/updates/prism-830-release-notes (Accessed: August 20, 2025).

Groux-Degroote, S. *et al.* (2018) “The extended cytoplasmic tail of the human B4GALNT2 is critical for its Golgi targeting and post-Golgi sorting,” *The FEBS Journal*, 285(18), pp. 3442–3463. Available at: https://doi.org/10.1111/febs.14621.

Groux-Degroote, S. *et al.* (2021) “B4GALNT2 Controls Sd ^{a} and SLe ^{x} Antigen Biosynthesis in Healthy and Cancer Human Colon,” *ChemBioChem*, 22(24), pp. 3381–3390. Available at: https://doi.org/10.1002/cbic.202100363.

Han, J. *et al.* (2018) “Genome-wide CRISPR/Cas9 Screen Identifies Host Factors Essential for Influenza Virus Replication,” *Cell Reports*, 23(2), pp. 596–607. Available at: https://doi.org/10.1016/j.celrep.2018.03.045.

He, Y. *et al.* (2024) “Sialic acids as attachment factors in mosquitoes mediating Japanese encephalitis virus infection,” *Journal of Virology*, 98(5). Available at: https://doi.org/10.1128/jvi.01959-23.

Heaton, B.E. *et al.* (2017) “A CRISPR Activation Screen Identifies a Pan-avian Influenza Virus Inhibitory Host Factor,” *Cell Reports*, 20(7), pp. 1503–1512. Available at: https://doi.org/10.1016/j.celrep.2017.07.060.

Honce, R. and Schultz-Cherry, S. (2020) “They are what you eat: Shaping of viral populations through nutrition and consequences for virulence,” *PLOS Pathogens*, 16(8), p. e1008711. Available at: https://doi.org/10.1371/journal.ppat.1008711.

Hou, X. *et al.* (2024) “Using artificial intelligence to document the hidden RNA virosphere,” *Cell*, 187(24), pp. 6929-6942.e16. Available at: https://doi.org/10.1016/j.cell.2024.09.027.

Hua, R. *et al.* (2024) “Experimental evidence for cancer resistance in a bat species,” *Nature Communications*, 15(1), p. 1401. Available at: https://doi.org/10.1038/s41467-024-45767-1.

iNaturalist (no date a) *Egyptian Rousette*, *iNaturalist* . Available at: https://www.inaturalist.org/taxa/75042-Rousettus-aegyptiacus (Accessed: August 22, 2025).

iNaturalist (no date b) *Trefoil Horseshoe Bat* , *iNaturalist*. Available at: https://www.inaturalist.org/taxa/40692-Rhinolophus-trifoliatus (Accessed: August 25, 2025).

Ishigaki, H. and Itoh, Y. (2022) “Detection of Sialic Acids on the Cell Surface Using Flow Cytometry,” in, pp. 31–35. Available at: https://doi.org/10.1007/978-1-0716-2635-1_3.

Jia, R. *et al.* (2012) “The N-Terminal Region of IFITM3 Modulates Its Antiviral Activity by Regulating IFITM3 Cellular Localization,” *Journal of Virology*, 86(24), pp. 13697–13707. Available at: https://doi.org/10.1128/JVI.01828-12.

Kalunda, M. *et al.* (1986) “Kasokero Virus: a New Human Pathogen from Bats (Rousettus Aegyptiacus) in Uganda,” *The American Journal of Tropical Medicine and Hygiene*, 35(2), pp. 387–392. Available at: https://doi.org/10.4269/ajtmh.1986.35.387.

Kandeil, A. *et al.* (2019) “Isolation and Characterization of a Distinct Influenza A Virus from Egyptian Bats,” *Journal of Virology*, 93(2). Available at: https://doi.org/10.1128/JVI.01059-18.

Karakus, U. *et al.* (2019) “MHC class II proteins mediate cross-species entry of bat influenza viruses,” *Nature*, 567(7746), pp. 109–112. Available at: https://doi.org/10.1038/s41586-019-0955-3.

Karamendin, K., Kydyrmanov, A. and Fereidouni, S. (2024) “Has avian influenza virus H9 originated from a bat source?,” *Frontiers in Veterinary Science*, 10. Available at: https://doi.org/10.3389/fvets.2023.1332886.

King, C.R. *et al.* (2023) “Pathogen-driven CRISPR screens identify TREX1 as a regulator of DNA self-sensing during influenza virus infection,” *Cell Host & Microbe*, 31(9), pp. 1552-1567.e8. Available at: https://doi.org/10.1016/j.chom.2023.08.001.

Kosakovsky Pond, S.L. *et al.* (2021) “Contrast-FEL—A Test for Differences in Selective Pressures at Individual Sites among Clades and Sets of Branches,” *Molecular Biology and Evolution*, 38(3), pp. 1184–1198. Available at: https://doi.org/10.1093/molbev/msaa263.

Kumar, S. *et al.* (2018) “MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms,” *Molecular Biology and Evolution*, 35(6), pp. 1547–1549. Available at: https://doi.org/10.1093/molbev/msy096.

le Corf, A. *et al.* (2025) “Genomic and functional adaptations in guanylate-binding protein 5 (GBP5) highlight specificities of bat antiviral innate immunity.” Available at: https://doi.org/10.1101/2025.02.11.637683.

Lesueur, L.L., Mir, L.M. and André, F.M. (2016) “Overcoming the Specific Toxicity of Large Plasmids Electrotransfer in Primary Cells In Vitro,” *Molecular Therapy - Nucleic Acids*, 5, p. e291. Available at: https://doi.org/10.1038/mtna.2016.4.

Letko, M. *et al.* (2020) “Bat-borne virus diversity, spillover and emergence,” *Nature Reviews Microbiology*, 18(8), pp. 461–471. Available at: https://doi.org/10.1038/s41579-020-0394-z.

Letunic, I. and Bork, P. (2024) “Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool,” *Nucleic Acids Research*, 52(W1), pp. W78–W82. Available at: https://doi.org/10.1093/nar/gkae268.

Li, B. *et al.* (2020) “Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection,” *Nature Communications*, 11(1), p. 164. Available at: https://doi.org/10.1038/s41467-019-13965-x.

Li, R. *et al.* (2024) “The Unique Immune System of Bats: An Evolutionary Analysis and Bibliometric Study,” *Ecology and Evolution*, 14(11). Available at: https://doi.org/10.1002/ece3.70614.

Li, Y. *et al.* (2013) “RNA Interference Functions as an Antiviral Immunity Mechanism in Mammals,” *Science*, 342(6155), pp. 231–234. Available at: https://doi.org/10.1126/science.1241911.

Lübbers, J. *et al.* (2021) “α2-3 Sialic acid binding and uptake by human monocyte-derived dendritic cells alters metabolism and cytokine release and initiates tolerizing T cell programming,” *Immunotherapy Advances*, 1(1). Available at: https://doi.org/10.1093/immadv/ltab012.

Ma, T. *et al.* (2023) “Genome-wide CRISPR screen identifies GNE as a key host factor that promotes influenza A virus adsorption and endocytosis,” *Microbiology Spectrum*, 11(6). Available at: https://doi.org/10.1128/spectrum.01643-23.

Madsen, C.K. *et al.* (2025) “Rapid one-step CRISPR-cas vector assembly by isothermal spacer removal linearization and sequence-ligation independent cloning (ISRL-SLIC),” *MethodsX*, 15, p. 103567. Available at: https://doi.org/10.1016/j.mex.2025.103567.

Mak, N.S.C. *et al.* (2024) “Alternative splicing expands the antiviral IFITM repertoire in Chinese rufous horseshoe bats,” *PLOS Pathogens*, 20(12), p. e1012763. Available at: https://doi.org/10.1371/journal.ppat.1012763.

Malagolini, N. *et al.* (2007) “Biosynthesis and expression of the Sda and sialyl Lewis x antigens in normal and cancer colon,” *Glycobiology*, 17(7), pp. 688–697. Available at: https://doi.org/10.1093/glycob/cwm040.

Mann, D.L. *et al.* (1988) “HLA-DR is involved in the HIV-1 binding site on cells expressing MHC class II antigens.,” *The Journal of Immunology*, 141(4), pp. 1131–1136. Available at: https://doi.org/10.4049/jimmunol.141.4.1131.

McLenachan, S., Sarsero, J.P. and Ioannou, P.A. (2007) “Flow-cytometric analysis of mouse embryonic stem cell lipofection using small and large DNA constructs,” *Genomics*, 89(6), pp. 708–720. Available at: https://doi.org/10.1016/j.ygeno.2007.02.006.

Medina, R.A. and García-Sastre, A. (2011) “Influenza A viruses: new research developments,” *Nature Reviews Microbiology*, 9(8), pp. 590–603. Available at: https://doi.org/10.1038/nrmicro2613.

Mo, Y., Lim, L.-S. and Ng, S.K. (2024) “A systematic review on current approaches in bat virus discovered between 2018 and 2022,” *Journal of Virological Methods*, 329, p. 115005. Available at: https://doi.org/10.1016/j.jviromet.2024.115005.

Morales, A.E. *et al.* (2025) “Bat genomes illuminate adaptations to viral tolerance and disease resistance,” *Nature*, 638(8050), pp. 449–458. Available at: https://doi.org/10.1038/s41586-024-08471-0.

Morton, J.A., Pickles, M.M. and Terry, A.M. (1970) “The Sda Blood Group Antigen in Tissues and Body Fluids,” *Vox Sanguinis*, 19(5–6), pp. 472–482. Available at: https://doi.org/10.1111/j.1423-0410.1970.tb01779.x.

Munir, M. *et al.* (2024) “Genome-wide CRISPR activation screen identifies JADE3 as an antiviral activator of NF-kB–dependent IFITM3 expression,” *Journal of Biological Chemistry*, 300(4), p. 107153. Available at: https://doi.org/10.1016/j.jbc.2024.107153.

Murrell, B. *et al.* (2012) “Detecting Individual Sites Subject to Episodic Diversifying Selection,” *PLoS Genetics*, 8(7), p. e1002764. Available at: https://doi.org/10.1371/journal.pgen.1002764.

Murrell, B. *et al.* (2013) “FUBAR: A Fast, Unconstrained Bayesian AppRoximation for Inferring Selection,” *Molecular Biology and Evolution*, 30(5), pp. 1196–1205. Available at: https://doi.org/10.1093/molbev/mst030.

NEB (no date a) *How does overlap length affect NEBuilder HiFi DNA Assembly efficiency?*, *New England Biolabs*. Available at: https://www.neb.com/en-gb/faqs/2021/10/13/how-does-overlap-length-affect-nebuilder-hifi-dna-assembly-efficiency?srsltid=AfmBOopwJ3Gs634zHjL-t1KA1fENfkZRxDkRJfGqXkqygHu3cgvwIxVp (Accessed: August 10, 2025).

NEB (no date b) *NEBuffer Activity/Performance Chart with Restriction Enzymes*, *New England Biolabs*. Available at: https://www.neb.com/en-gb/tools-and-resources/usage-guidelines/nebuffer-performance-chart-with-restriction-enzymes?srsltid=AfmBOor1TXSIViSXNeNPkodUb-gqQr4lKUTgEyeJWEm1VQ-FV2DVSdtI (Accessed: August 10, 2025).

Nelson, S.W. *et al.* (2019) “Madin-Darby canine kidney cell sialic acid receptor modulation induced by culture medium conditions: Implications for the isolation of influenza A virus,” *Influenza and Other Respiratory Viruses*, 13(6), pp. 593–602. Available at: https://doi.org/10.1111/irv.12671.

NIH (2024) *NIH Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules (NIH GUIDELINES)*.

Ozawa, M. *et al.* (2011) “Replication-incompetent influenza A viruses that stably express a foreign gene,” *Journal of General Virology*, 92(12), pp. 2879–2888. Available at: https://doi.org/10.1099/vir.0.037648-0.

Paratus (2025) *Pipeline*, *Paratus Sciences* . Available at: https://paratussciences.com/pipeline/ (Accessed: August 24, 2025).

Park, J.S. *et al.* (2023) “Modification of surface glycan by expression of beta-1,4-N-acetyl-galactosaminyltransferase (B4GALNT2) confers resistance to multiple viruses infection in chicken fibroblast cell,” *Frontiers in Veterinary Science*, 10. Available at: https://doi.org/10.3389/fvets.2023.1160600.

Pauciullo, S. *et al.* (2024) “Spillover: Mechanisms, Genetic Barriers, and the Role of Reservoirs in Emerging Pathogens,” *Microorganisms*, 12(11), p. 2191. Available at: https://doi.org/10.3390/microorganisms12112191.

Pawan, J.L. (1936) “The Transmission of Paralytic Rabies in Trinidad by the Vampire Bat (*Desmodus Rotundus Murinus* Wagner, 1840),” *Annals of Tropical Medicine & Parasitology*, 30(1), pp. 101–130. Available at: https://doi.org/10.1080/00034983.1936.11684921.

Peng, C. *et al.* (2022) “Rhinolophus sinicus virome revealed multiple novel mosquito-borne zoonotic viruses,” *Frontiers in Cellular and Infection Microbiology*, 12. Available at: https://doi.org/10.3389/fcimb.2022.960507.

Pennisi, E. (2020) “How bats have outsmarted viruses—including coronaviruses—for 65 million years,” *Science* [Preprint]. Available at: https://doi.org/10.1126/science.abd9595.

Perng, Y.-C. and Lenschow, D.J. (2018) “ISG15 in antiviral immunity and beyond,” *Nature Reviews Microbiology*, 16(7), pp. 423–439. Available at: https://doi.org/10.1038/s41579-018-0020-5.

Pinto, R.M. *et al.* (2023) “BTN3A3 evasion promotes the zoonotic potential of influenza A viruses,” *Nature*, 619(7969), pp. 338–347. Available at: https://doi.org/10.1038/s41586-023-06261-8.

Plowright, R.K. *et al.* (2008) “Reproduction and nutritional stress are risk factors for Hendra virus infection in little red flying foxes (Pteropus scapulatus).,” *Proceedings. Biological sciences*, 275(1636), pp. 861–9. Available at: https://doi.org/10.1098/rspb.2007.1260.

Poirier, E.Z. *et al.* (2021) “An isoform of Dicer protects mammalian stem cells against multiple RNA viruses,” *Science*, 373(6551), pp. 231–236. Available at: https://doi.org/10.1126/science.abg2264.

Pourrut, X. *et al.* (2009) “Large serological survey showing cocirculation of Ebola and Marburg viruses in Gabonese bat populations, and a high seroprevalence of both viruses in Rousettus aegyptiacus,” *BMC Infectious Diseases*, 9(1), p. 159. Available at: https://doi.org/10.1186/1471-2334-9-159.

Pucci, M. *et al.* (2020) “High Expression of the Sda Synthase B4GALNT2 Associates with Good Prognosis and Attenuates Stemness in Colon Cancer,” *Cells*, 9(4), p. 948. Available at: https://doi.org/10.3390/cells9040948.

Pulit-Penaloza, J.A. *et al.* (2024) “Transmission of a human isolate of clade 2.3.4.4b A(H5N1) virus in ferrets,” *Nature*, 636(8043), pp. 705–710. Available at: https://doi.org/10.1038/s41586-024-08246-7.

Rabe, B.A. and Cepko, C. (2020) “A Simple Enhancement for Gibson Isothermal Assembly,” *bioRxiv* [Preprint]. Available at: https://doi.org/10.1101/2020.06.14.150979.

Rahman, Md.T. *et al.* (2020) “Zoonotic Diseases: Etiology, Impact, and Control,” *Microorganisms*, 8(9), p. 1405. Available at: https://doi.org/10.3390/microorganisms8091405.

Remmel, A. (2024) “How to keep the lights on: the mission to make more photostable fluorophores,” *Nature*, 630(8015), pp. 258–260. Available at: https://doi.org/10.1038/d41586-024-01591-7.

Roche, P.A. and Furuta, K. (2015) “The ins and outs of MHC class II-mediated antigen processing and presentation,” *Nature Reviews Immunology*, 15(4), pp. 203–216. Available at: https://doi.org/10.1038/nri3818.

Russell, C.J. (2021) “Hemagglutinin Stability and Its Impact on Influenza A Virus Infectivity, Pathogenicity, and Transmissibility in Avians, Mice, Swine, Seals, Ferrets, and Humans,” *Viruses*, 13(5), p. 746. Available at: https://doi.org/10.3390/v13050746.

Sampath, S. *et al.* (2021) “Pandemics Throughout the History,” *Cureus* [Preprint]. Available at: https://doi.org/10.7759/cureus.18136.

Sarkar, L., Liu, G. and Gack, M.U. (2023) “ISG15: its roles in SARS-CoV-2 and other viral infections,” *Trends in Microbiology*, 31(12), pp. 1262–1275. Available at: https://doi.org/10.1016/j.tim.2023.07.006.

Sarker, A. *et al.* (2022) “Influenza-existing drugs and treatment prospects,” *European Journal of Medicinal Chemistry*, 232, p. 114189. Available at: https://doi.org/10.1016/j.ejmech.2022.114189.

Sawoo, O. *et al.* (2014) “Cleavage of Hemagglutinin-Bearing Lentiviral Pseudotypes and Their Use in the Study of Influenza Virus Persistence,” *PLoS ONE*, 9(8), p. e106192. Available at: https://doi.org/10.1371/journal.pone.0106192.

Sayers, J.R. and Eckstein, F. (1991) “A single-strand specific endonuclease activity copurifies with overexpressed T5 D15 exonuclease,” *Nucleic Acids Research*, 19(15), pp. 4127–4132. Available at: https://doi.org/10.1093/nar/19.15.4127.

Schafers, J. *et al.* (2025) “Stability of influenza viruses in the milk of cows and sheep.” Available at: https://doi.org/10.1101/2025.05.28.25328508.

Schountz, T. *et al.* (2017) “Immunological Control of Viral Infections in Bats and the Emergence of Viruses Highly Pathogenic to Humans,” *Frontiers in Immunology*, 8. Available at: https://doi.org/10.3389/fimmu.2017.01098.

Shao, W. *et al.* (2017) “Evolution of Influenza A Virus by Mutation and Re-Assortment.,” *International journal of molecular sciences*, 18(8). Available at: https://doi.org/10.3390/ijms18081650.

Sharon, D.M. *et al.* (2020) “A pooled genome-wide screening strategy to identify and rank influenza host restriction factors in cell-based vaccine production platforms,” *Scientific Reports*, 10(1), p. 12166. Available at: https://doi.org/10.1038/s41598-020-68934-y.

Shen, Y.-Y. *et al.* (2010) “Adaptive evolution of energy metabolism genes and the origin of flight in bats,” *Proceedings of the National Academy of Sciences*, 107(19), pp. 8666–8671. Available at: https://doi.org/10.1073/pnas.0912613107.

Singh, B.B., Ward, M.P. and Dhand, N.K. (2022) “Inherent virus characteristics and host range drive the zoonotic and emerging potential of viruses,” *Transboundary and Emerging Diseases*, 69(4). Available at: https://doi.org/10.1111/tbed.14361.

Skehel, J.J. and Wiley, D.C. (2000) “Receptor Binding and Membrane Fusion in Virus Entry: The Influenza Hemagglutinin,” *Annual Review of Biochemistry*, 69(1), pp. 531–569. Available at: https://doi.org/10.1146/annurev.biochem.69.1.531.

Søndergaard, J.N. *et al.* (2020) “Successful delivery of large-size CRISPR/Cas9 vectors in hard-to-transfect human cells using small plasmids,” *Communications Biology*, 3(1), p. 319. Available at: https://doi.org/10.1038/s42003-020-1045-7.

Song, Y. *et al.* (2021) “A genome-wide CRISPR/Cas9 gene knockout screen identifies immunoglobulin superfamily DCC subclass member 4 as a key host factor that promotes influenza virus endocytosis,” *PLOS Pathogens*, 17(12), p. e1010141. Available at: https://doi.org/10.1371/journal.ppat.1010141.

Spriggs, M.K. *et al.* (1996) “The extracellular domain of the Epstein-Barr virus BZLF2 protein binds the HLA-DR beta chain and inhibits antigen presentation,” *Journal of Virology*, 70(8), pp. 5557–5563. Available at: https://doi.org/10.1128/jvi.70.8.5557-5563.1996.

Stenfelt, L. *et al.* (2019) “Missense mutations in the C-terminal portion of the B4GALNT2-encoded glycosyltransferase underlying the Sd(a−) phenotype,” *Biochemistry and Biophysics Reports*, 19, p. 100659. Available at: https://doi.org/10.1016/j.bbrep.2019.100659.

Sun, X. *et al.* (2013) “Bat-Derived Influenza Hemagglutinin H17 Does Not Bind Canonical Avian or Human Receptors and Most Likely Uses a Unique Entry Mechanism,” *Cell Reports*, 3(3), pp. 769–778. Available at: https://doi.org/10.1016/j.celrep.2013.01.025.

Suwandi, A. *et al.* (2022) “B4galnt2-mediated host glycosylation influences the susceptibility to Citrobacter rodentium infection,” *Frontiers in Microbiology*, 13. Available at: https://doi.org/10.3389/fmicb.2022.980495.

Tang, B.L. *et al.* (1992) “The transmembrane domain of N-glucosaminyltransferase I contains a Golgi retention signal.,” *Journal of Biological Chemistry*, 267(14), pp. 10122–10126. Available at: https://doi.org/10.1016/S0021-9258(19)50208-3.

Teeling, E.C. *et al.* (2018) “Bat Biology, Genomes, and the Bat1K Project: To Generate Chromosome-Level Genomes for All Living Bat Species,” *Annual Review of Animal Biosciences*, 6(1), pp. 23–46. Available at: https://doi.org/10.1146/annurev-animal-022516-022811.

Temperton, N.J. *et al.* (2007) “A sensitive retroviral pseudotype assay for influenza H5N1-neutralizing antibodies,” *Influenza and Other Respiratory Viruses*, 1(3), pp. 105–112. Available at: https://doi.org/10.1111/j.1750-2659.2007.00016.x.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) “CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice,” *Nucleic Acids Research*, 22(22), pp. 4673–4680. Available at: https://doi.org/10.1093/nar/22.22.4673.

Tian, X. *et al.* (2021) “A Replication-Defective Influenza Virus Harboring H5 and H7 Hemagglutinins Provides Protection against H5N1 and H7N9 Infection in Mice,” *Journal of Virology*, 95(3). Available at: https://doi.org/10.1128/JVI.02154-20.

Tomori, O. and Oluwayelu, D.O. (2023) “Domestic Animals as Potential Reservoirs of Zoonotic Viral Diseases,” *Annual Review of Animal Biosciences*, 11(1), pp. 33–55. Available at: https://doi.org/10.1146/annurev-animal-062922-060125.

Tong, S. *et al.* (2012) “A distinct lineage of influenza A virus from bats,” *Proceedings of the National Academy of Sciences*, 109(11), pp. 4269–4274. Available at: https://doi.org/10.1073/pnas.1116200109.

Tong, S. *et al.* (2013) “New World Bats Harbor Diverse Influenza A Viruses,” *PLoS Pathogens*, 9(10), p. e1003657. Available at: https://doi.org/10.1371/journal.ppat.1003657.

Tran, V. *et al.* (2020) “Influenza virus repurposes the antiviral protein IFIT2 to promote translation of viral mRNAs,” *Nature Microbiology*, 5(12), pp. 1490–1503. Available at: https://doi.org/10.1038/s41564-020-0778-x.

University of Bristol (no date) *Chinese rufous Horseshoe Bat*, *University of Bristol* .

USDA (2025) H5N1 Influenza, Animal and Plant Health Inspection Service*, U.S. DEPARTMENT OF AGRICULTURE*. Available at: https://www.aphis.usda.gov/h5n1-hpai (Accessed: August 6, 2025).

Valero-Rello, A. *et al.* (2024) “Cellular receptors for mammalian viruses,” *PLOS Pathogens*, 20(2), p. e1012021. Available at: https://doi.org/10.1371/journal.ppat.1012021.

Vallier, M. *et al.* (2023) “Pathometagenomics reveals susceptibility to intestinal infection by Morganella to be mediated by the blood group-related B4galnt2 gene in wild mice,” *Gut Microbes*, 15(1). Available at: https://doi.org/10.1080/19490976.2022.2164448.

van Valen, L. (1973) “A New Evolutionary Law,” *Evolutionary Theory* , 1, pp. 1–30.

Varki, A. and Gagneux, P. (2012) “Multifarious roles of sialic acids in immunity,” *Annals of the New York Academy of Sciences*, 1253(1), pp. 16–36. Available at: https://doi.org/10.1111/j.1749-6632.2012.06517.x.

Victor, S.T. *et al.* (2012) “A Replication-Incompetent PB2-Knockout Influenza A Virus Vaccine Vector,” *Journal of Virology*, 86(8), pp. 4123–4128. Available at: https://doi.org/10.1128/JVI.06232-11.

Wang, W. *et al.* (2008) “Establishment of retroviral pseudotypes with influenza hemagglutinins from H1, H3, and H5 subtypes for sensitive and specific detection of neutralizing antibodies,” *Journal of Virological Methods*, 153(2), pp. 111–119. Available at: https://doi.org/10.1016/j.jviromet.2008.07.015.

Wang, Y. and Shao, W. (2023) “Innate Immune Response to Viral Vectors in Gene Therapy,” *Viruses*, 15(9), p. 1801. Available at: https://doi.org/10.3390/v15091801.

Wang, Y. *et al.* (2024) “Unveiling bat-borne viruses: a comprehensive classification and analysis of virome evolution,” *Microbiome*, 12(1), p. 235. Available at: https://doi.org/10.1186/s40168-024-01955-1.

Waterhouse, A.M. *et al.* (2009) “Jalview Version 2—a multiple sequence alignment editor and analysis workbench,” *Bioinformatics*, 25(9), pp. 1189–1191. Available at: https://doi.org/10.1093/bioinformatics/btp033.

Weaver, S. *et al.* (2018) “Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes,” *Molecular Biology and Evolution*, 35(3), pp. 773–777. Available at: https://doi.org/10.1093/molbev/msx335.

WHO (2020a) Tool for Influenza Pandemic Risk Assessment *(TIPRA) 2nd Edition*. Available at: https://www.who.int/publications/i/item/tool-for-influenza-pandemic-risk-assessment-(tipra)-2nd-edition (Accessed: August 20, 2025).

WHO (2020b) Zoonoses, World Health Organisation. Available at: https://www.who.int/news-room/fact-sheets/detail/zoonoses (Accessed: August 6, 2025).

WHO (2025) WHO COVID-19 dashboard, *World Health Organisation*. Available at: https://data.who.int/dashboards/covid19/deaths (Accessed: August 6, 2025).

Wilson, B.A., Neldner, V.J. and Accad, A. (2002) “The extent and status of remnant vegetation in Queensland and its implications for statewide vegetation management and legislation,” *The Rangeland Journal*, 24(1), p. 6. Available at: https://doi.org/10.1071/RJ02001.

Wong, H.H., Fung, K. and Nicholls, J.M. (2019) “MDCK-B4GalNT2 cells disclose a α2,3-sialic acid requirement for the 2009 pandemic H1N1 A/California/04/2009 and NA aid entry of A/WSN/33,” *Emerging Microbes & Infections*, 8(1), pp. 1428–1437. Available at: https://doi.org/10.1080/22221751.2019.1665971.

Wu, X. *et al.* (2025) “Identification of β4GALNT2 as an anti-hPIV3 factor through genome-wide CRISPR/Cas9 library screening,” *Emerging Microbes & Infections*, 14(1). Available at: https://doi.org/10.1080/22221751.2025.2529895.

Xie, J. *et al.* (2018) “Dampened STING-Dependent Interferon Activation in Bats,” *Cell Host & Microbe*, 23(3), pp. 297-301.e4. Available at: https://doi.org/10.1016/j.chom.2018.01.006.

Xie, J., Tan, B. and Zhang, Y. (2023) “Positive Selection and Duplication of Bat TRIM Family Proteins,” *Viruses*, 15(4), p. 875. Available at: https://doi.org/10.3390/v15040875.

Yang, S. *et al.* (2024) “Plate centrifugation enhances the efficiency of polyethylenimine-based transfection and lentiviral infection,” *Journal of Virological Methods*, 330, p. 115039. Available at: https://doi.org/10.1016/j.jviromet.2024.115039.

Yang, W., Schountz, T. and Ma, W. (2021) “Bat Influenza Viruses: Current Status and Perspective,” *Viruses*, 13(4), p. 547. Available at: https://doi.org/10.3390/v13040547.

Yau, S. and Seth-Pasricha, M. (2019) “Viruses of Polar Aquatic Environments,” *Viruses*, 11(2), p. 189. Available at: https://doi.org/10.3390/v11020189.

Yi, C. *et al.* (2022) “Genome-wide CRISPR-Cas9 screening identifies the CYTH2 host gene as a potential therapeutic target of influenza viral infection,” *Cell Reports*, 38(13), p. 110559. Available at: https://doi.org/10.1016/j.celrep.2022.110559.

Yuan, J. *et al.* (2010) “Intraspecies diversity of SARS-like coronaviruses in Rhinolophus sinicus and its implications for the origin of SARS coronaviruses in humans,” *Journal of General Virology*, 91(4), pp. 1058–1062. Available at: https://doi.org/10.1099/vir.0.016378-0.

Zhou, A. *et al.* (2021) “Porcine Genome-Wide CRISPR Screen Identifies the Golgi Apparatus Complex Protein COG8 as a Pivotal Regulator of Influenza Virus Infection.,” *The CRISPR journal*, 4(6), pp. 872–883. Available at: https://doi.org/10.1089/crispr.2021.0054.

Zhu, W. *et al.* (2024) “Biological function of sialic acid and sialylation in human health and disease,” *Cell Death Discovery*, 10(1), p. 415. Available at: https://doi.org/10.1038/s41420-024-02180-3.

Zhu, X. *et al.* (2012) “Crystal structures of two subtype N10 neuraminidase-like proteins from bat influenza A viruses reveal a diverged putative active site,” *Proceedings of the National Academy of Sciences*, 109(46), pp. 18903–18908. Available at: https://doi.org/10.1073/pnas.1212579109.