

Glycosylation and the global virome

Authors

Cassandra L. Pegg^{1,*}, Benjamin L. Schulz¹, Benjamin A. Neely², Gregory F Albery³, and Colin J. Carlson^{3,4,5,*}

Affiliations

1. School of Chemistry and Molecular Biosciences, The University of Queensland, St Lucia QLD, 4072, Australia
2. National Institute of Standards and Technology, Charleston, SC 29412, U.S.A.
3. Department of Biology, Georgetown University, Washington, D.C. 20007, U.S.A.
4. Department of Microbiology and Immunology, Georgetown University Medical Center, Washington, D.C. 20007, U.S.A.
5. Center for Global Health Science and Security, Georgetown University Medical Center, Washington, D.C. 20007, U.S.A.

*Address correspondence to c.peggl@uq.edu.au or Colin.Carlson@georgetown.edu

ABSTRACT

The sugars that coat the outsides of viruses and host cells are key to successful disease transmission, but they remain understudied compared to other molecular features. Understanding the comparative zoology of glycosylation - and harnessing it for predictive science - could help close the molecular gap in zoonotic risk assessment.

INTRODUCTION

Due to recent encounters with zoonotic viruses like Ebola virus and SARS-CoV-2, efforts to forecast the zoonotic risk of wildlife viruses - and, more broadly, to understand the biological constraints on cross-species transmission - are increasingly appealing (1). To date, most of these efforts rely on easily-observed traits of hosts, like morphology, diet, or phylogeny (2, 3). Despite their distance from the molecular determinants of transmission, these traits can be used to build models that have surprising predictive accuracy (4).

Many microbiologists have expressed a healthy skepticism of these approaches, which often entirely lack predictors that consider the molecular biology of hosts or *any* viral traits, and therefore, only coarsely infer the cellular processes of infection through other proxies that are correlated across evolutionary space. As a rare exception, genomic approaches are increasingly being used to close this gap (5, 6), and can help identify salient mechanisms of host-virus interactions (e.g., CpG dinucleotide depletion in vertebrate viruses appears to help them evade innate immune responses like the zinc finger antiviral protein (7)). However, as predictive features, genomic traits are often confounded by evolutionary signals (8, 9), and genomic data only offer limited insights into the actual three-dimensional structural compatibility of viral and host cell surfaces, a “lock-and-key” type process. This lock-and-key interaction not only allows efficient infection of susceptible hosts, but also limit cross-species viral transmission. Structural modeling approaches have been used to examine the binding of viral proteins and host cell receptors, often in the context of research on SARS-CoV-2, but many of these simulations neglect key information: the glycosylation of these structures. Indeed, the handful that do address this aspect have revealed unexpected and important roles for glycosylation in these processes (10, 11).

The sugars, or glycans, that decorate host cell surface macromolecules are often critical ligands that viruses associate with to enter cells (12). Viral proteins from different virus families exhibit substantial variation in binding affinity towards host glycan receptors, and the compatibility between different viral proteins and host glycosylation varies between host and virus species, across tissues and organ systems, and even over time or between individuals (12-14). Some viral surface proteins become glycosylated by host cell machinery during infection, and in the process can mimic host cell surfaces or shield proteins from antibody recognition, helping viruses evade the host’s immune system (15, 16). The glycans on viral surfaces are also

recognised by host glycan-binding proteins on immune cells that capture viruses, either preventing or promoting infection (17, 18). These aspects of host-virus compatibility can create or unlock barriers to transmission, but are poorly characterized as an underlying structural determinant of host-virus networks because glycan structures are subject to rapid regulation and are sometimes perceived as being analytically challenging. We suggest an undertaking to describe the comparative zoology of glycoproteins, and their role in structuring the global virome.

WHAT SUGARS DO, AND HOW

Glycans are a key feature of the cell surface and extracellular matrix of eukarya, archaea, and bacteria (19). Glycans can be found on proteins or lipids, and technological advances (20) in the 21st century have greatly increased our ability to identify, characterise, and manipulate glycosylation, which has in turn supported deeper insights into the multitude of diverse roles it plays (21). Glycans can constitute a substantial proportion of the molecular mass of a protein (22), contributing to their biophysical properties and influencing protein targeting, folding, structure and secretion (22, 23). Moreover, glycans are mostly located on secreted proteins and at the cell surface, and are therefore key determinants in molecular recognition events (24).

Glycans exhibit tremendous structural diversity (25). Unlike proteins, whose sequences can be predicted from gene sequences, the biosynthesis of glycans is not directly template-driven; glycans are built, modified and trimmed by an extensive network of co-expressed enzymes that are differentially expressed in cells and tissues, and can be affected by factors intrinsic and extrinsic to the cell. Glycan structures are therefore specific to various organisms, tissues, and cells, and the resulting structures can be highly heterogeneous, imparting additional complexity to the structural and functional properties of proteins (24).

Glycans act as receptors, coreceptors or attachment factors for numerous viruses (13) including HIV, dengue (26), MERS-CoV (27), influenza (28), and SARS-CoV-2 (29). For example, both avian and human-adapted influenza viruses bind to glycans that terminate with sialic acid, but avian influenza preferentially binds to sialic acid with α 2-3-linkages, while human-adapted influenza prefers α 2-6-linkages, which are expressed in the human upper respiratory tract (28, 30). When avian-origin influenza lineages jump directly into humans, infections mostly become established in the lower lungs where cells express α 2-3-linkages, leading to rare infections that are generally more severe but less transmissible (31). Animal bridge hosts that can be infected by different influenza subtypes (including avian species) or express both types of sialic acid - for example, swine - provide an environment where viral lineages may directly undergo mutation to adapt to α 2-6-binding (32), or where human and avian lineages co-circulate and undergo reassortment (Figure 1a). These hosts therefore provide an evolutionary stepping stone for avian lineages to switch to α 2-6-binding, opening a transmission route for harmful zoonoses and producing more transmissible strains that pose epidemic or pandemic threats in humans (33).

In cases like these, glycosylation is a key driver of host range and zoonotic risk, but one that is often neglected or folded into the “black box” of host-virus interactions and evolutionary dynamics in systems that are less well-characterized than influenza. However, the general importance of glycosylation in host-pathogen interactions is well established (15, 34). Given that potential differences in glycosylation presence and structure can have profound effects on molecular interactions, we suggest there is a clear need to measure the glycomes of potential hosts as part of broader efforts to describe viral ecology and emergence.

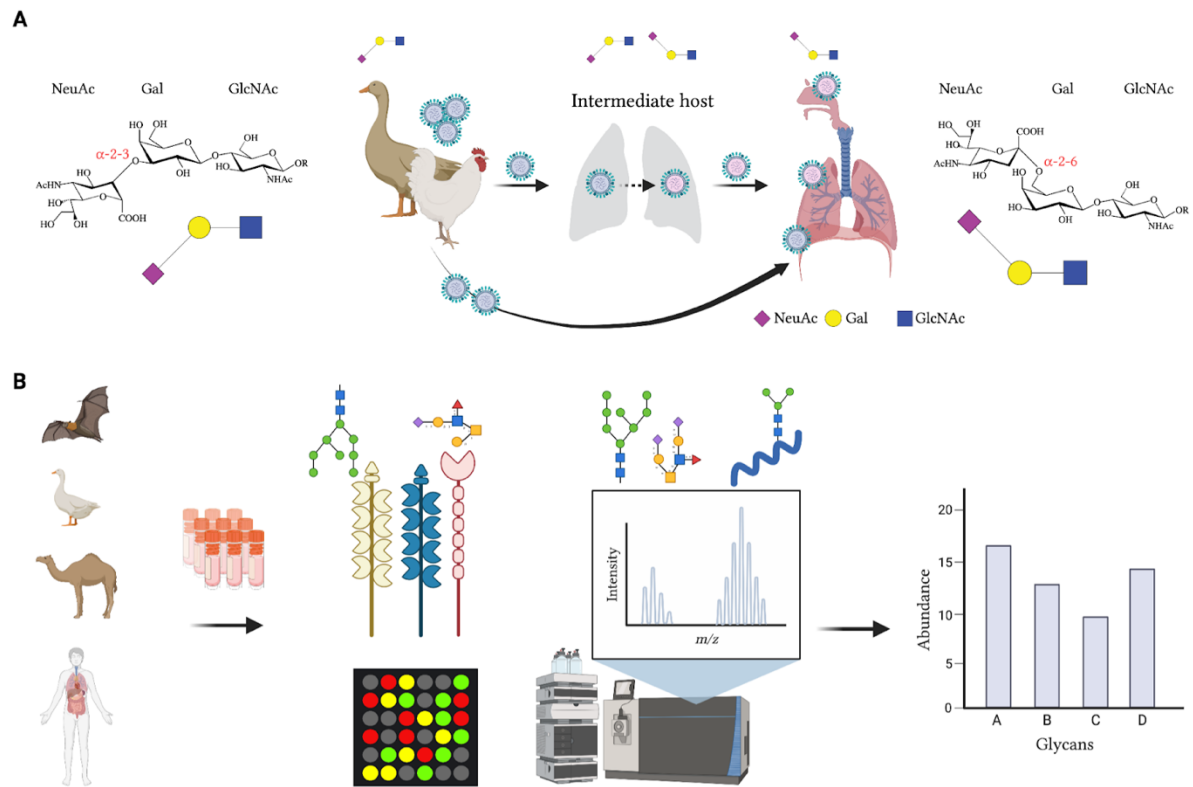


Figure 1. Glycosylation underlies the evolutionary shift towards zoonotic emergence of influenza. A. When avian influenza makes the jump directly to humans, preferential binding to α 2-3-linked sialic acid expressed in the lower lung leads to rarer infections that can be more severe but less transmissible. In intermediate hosts that express both α 2-3-linked and α 2-6-linked sialic acid, viruses can undergo an evolutionary shift that facilitates emergence in humans, who express α 2-6-linked sialic acid in the upper respiratory tract. B. An array of different tools can be used to study the glycosylation profiles of tissues, cells or proteins from a global level down to the glycan level. Figure created with BioRender (biorender.com).

CHARACTERIZING GLYCANS AT DIFFERENT SCALES

The inherent structural complexity and heterogeneity of glycans across species, individuals, organ systems, tissues, and even time and space make them an analytically challenging subject. Additionally, and in contrast to other biomolecules, they can have extraordinarily high structural complexity due the variety of monosaccharide building blocks and the multiple ways

they can attach to each other, both in bond configuration (α and β) and in the positions of the inter-saccharide linkages within the molecules (35). Nevertheless, a number of techniques can be used individually or in combination to characterise protein glycosylation, with mass spectrometry being a powerful and widely used tool that can be incorporated at various levels (Figure 1b). Generally, the overarching aim of glycosylation analyses is to deduce one or more of the following: the monosaccharide composition of the glycans, the order and branching of monosaccharides in a glycan, the types of glycosidic linkages and monosaccharide anomericity, or the location of the glycosylation sites on a protein. The functional roles of glycans can also be assessed by measuring non-covalent interactions between specific proteins and glycans. A global view can be obtained from antibody or lectin binding to selectively identify glycan epitopes or motifs in tissues, cells and proteins. Lectins can be used in array-based platforms (36) enabling high-throughput analyses with the caveat that they do not provide comprehensive structural information. The most effective way to obtain structural details of glycans is to use a glycomic workflow whereby the glycans are chemically or enzymatically released from glycoproteins. The released glycans are typically chemically labelled with fluorescent tags and analysed by liquid chromatography or capillary electrophoresis with a mass spectrometer used for enhanced detection (20). These methods for studying the glycome can provide monosaccharide composition and sequence information, and at times, linkage position. Coupling these analytical techniques with enzymes that cleave specific monosaccharide linkages provides additional precise structural information.

A limitation of glycomic approaches is that protein- and site-specificity is lost with glycan release. Nevertheless, these techniques are the most powerful for providing the basic information about glycan structure that will likely form the basis of future predictive models (see below). This is because the same glycan structures or epitopes can often be found on many

different sites and proteins from the same cell, because they share the same glycan biosynthetic pathways. In addition, a key benefit of glycomics, compared to proteomics or glycoproteomics (the study of proteins and glycosylated proteins, respectively), is that analysis of released glycans does not require prior knowledge of the genome. Although glycan biosynthesis is non-template driven and millions of possible glycan structures can be predicted (35), there are comparatively few glycan structures actually observed on glycoproteins (37), and these can be predicted or *de novo* structurally determined without knowledge of the genome (38). If the annotated genome is available, mass spectrometry glycoproteomics can also be used to identify and measure peptides with attached glycans. In this case, fine structural detail of the glycan structure is lost but the site of attachment and the level of site occupancy is retained.

BUILDING GLYCOSYLATION INTO PREDICTIVE SCIENCE

Understanding the landscape of host glycosylation might help scientists build better predictive tools to understand the broader rules of viral cross-species transmission or even the special case of zoonotic risk. This could be accomplished in a number of ways, most of which are untested. Glycosylation could be represented as data in several ways, ranging from simple (e.g., the presence or absence of a specific set of glycan structures) to complex (generating quantitative features using graph representations of the glycan structure (39)). These data can then be used several ways. For example, recent attention on SARS-CoV-2's use of the ACE2 receptor has sparked the development of models that predict host susceptibility based on receptor sequences (40), but glycosylation is a missing element; incorporating glycan structures as receptor or co-receptor "metadata" might help researchers better understand viral attachment (Figure 2a). For instance, sialic acid and heparan sulfate are key cell surface glycans that are co-receptors for SARS-CoV-2 (41, 42). Conversely, from the host perspective, similar glycosylation might help explain pathogen sharing between different animals (Figure 2b), and (according to some

preliminary evidence (43)) might even help unpack some of the microbiology inside the black box “phylogenetic distance effect” that broadly structures the viral sharing network (3). If viral glycosylation helps evade host immune system detection, these dimensions of similarity may help explain how particular bridge host “stepping stones” are possible, including in cases of zoonotic emergence. (During early characterization, glycan motifs or structures can even be proactively searched to identify homology in human and animal hosts, and linked to viruses where glycan-binding preferences have been established). Similarly, understanding the role of viral mutations that alter the viral surface glycoproteins (Figure 2c) may lead to more targeted insights about how glycosylation relates to zoonotic risk (as in the example of influenza and sialic acid; Figure 1a).

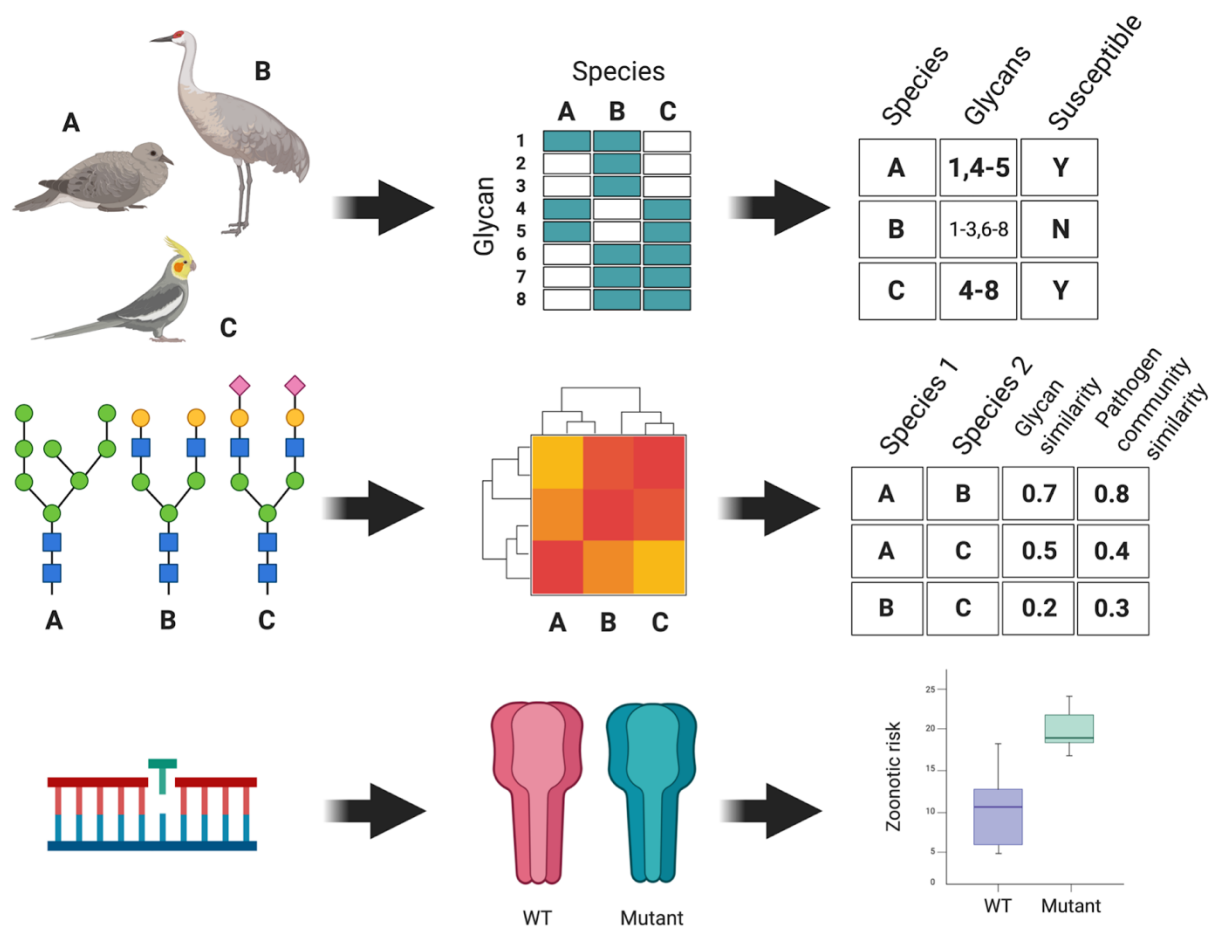


Figure 2. Adding glycosylation to microbiology-smart modeling. (*top*) Different animals have different combinations of glycans, which may help unpack how specific glycans contribute to susceptibility to a given pathogen. In this example, species B's lack of glycans 4 and 5 could explain its lack of susceptibility to the pathogen. (*middle*) When the mechanism is understood in better detail, the glycosylation of a single structure (e.g., the ACE2 receptor) might help predict cross-species transmission potential for a specific virus (e.g., SARS-CoV-2) - if the structural similarity of a given glycan can be converted into machine-readable features. (*bottom*) Understanding mechanisms in greater detail may improve other kinds of predictions about cross-species transmission: for example, mutations in the hemagglutinin structure of influenza viruses limit their binding efficiency to the glycosylation of human sialic acid receptors, allowing prediction of the zoonotic potential of specific influenza A strains based on a few point mutations. Figure created with BioRender (biorender.com).

Most of these examples are still hypothetical, but in limited cases, these types of model-based exploration have shown tremendous promise. For example, a recent study used graph representations of glycans and multiple kinds of advanced machine learning (graph convolutional neural networks and natural language models) to predict host identity and glycan immunogenicity, and was able to predict influenza and rotavirus binding affinity for host receptors from different species (43). Studies like these are exciting proofs-of-concept, and point to the idea that feature representations of glycans may eventually be useful as part of a broader palette of cell- and virus-level trait predictors used to make even more advanced (and crucially, microbiology-driven) machine learning or network models.

In order to power these kinds of approaches, more data is needed about the “global glycome.” While the human glycome is well studied (44), the glycomes of animal reservoirs are severely

understudied - a major problem when, for example, host range is extensive (e.g., influenza A virus can infect captive and wild animals including birds, dogs, cats, pigs, horses, bats, seals and even some reptiles (45)). Wild reservoir populations should be a priority for glycosylation analyses, particularly those that are endemically infected, at the human-animal interface and those likely to act as an intermediate “mixing vessel” for cross-species transmission. Tissues, cells and fluids that are the routes of entry for infectious agents or are predicted to be involved in tropism and systemic spread should be a focal point of analysis. These may include the mucosal epithelial tissue of the gastrointestinal, urogenital and respiratory tracts and cells of the skin, lymphatic system and blood vessels. The types of data to be incorporated from these sample types could range from global glycome analyses to precise glycan structure and site occupancy information. Ultimately, the types of experiments conducted will be determined by the capabilities of the laboratory. In depth protocols are available for global structural (46, 47) and protein- and site-specific (48-52) characterisation of glycans. Given the versatility of liquid chromatography mass spectrometry to study glycosylation, this approach seems most accessible with costs predicted to be in the low to medium range depending on the level of structural characterisation achieved. With unlimited sample availability and technical resources, multiple tissues from multiple species could be characterized and this data used for modeling. Tissue types may be limited due to the requirements for lethal or non-lethal sampling, the ability to dissect out specific tissues, or even the capacity to sample all species, though many field researchers have suitable samples already collected and in storage. As with most field work, sample availability will be dictated by each situation, but a balance should be struck between accessible biofluids (such as blood) and relevant tissues (such as lung tissue for respiratory viruses). On the other hand, given the ever-growing frequency of glycomic and glycoproteomic techniques (lectin-arrays or mass spectrometry-based), it is possible that the analytical work could be accomplished collaboratively at low to medium cost as part of broader

comparative zoology and viral ecology. Critically, depositing the datasets generated by this work into existing public repositories like GlycoPOST and GlyTouCan, which is becoming standard practice in the glycobiology field (53), will allow their value to grow exponentially for comparative and predictive research.

Recent viral epidemics and pandemics have highlighted a need for increased surveillance at the animal-human interface and forward planning of biochemical countermeasures (54-56). Characterizing the global glycome will help microbiologists and ecologists understand the broader dynamics of viral ecology, and these data could also easily be applied to understanding other pathogens. Moreover, as work during the COVID-19 pandemic has highlighted, understanding glycosylation as a viral phenotype is a key part of understanding pathogenesis and developing effective countermeasures (57, 58), and we suggest that building more glycomics into viral surveillance is a feasible, cost-effective, and impactful way to expand the body of basic science that forms the basis for epidemic preparedness.

ACKNOWLEDGEMENTS

This work was supported by NSF BII 2021909 and NHMRC Ideas grant APP1186699. Identification of certain commercial equipment, instruments, software, or materials does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products identified are necessarily the best available for the purpose.

CONTRIBUTIONS

C.J.C and B.A.N conceptualized the perspective. All authors contributed to the writing and review of the present paper.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

1. Albery GF, Becker DJ, Brierley L, Brook CE, Christofferson RC, Cohen LE, Dallas TA, Eskew EA, Fagre A, Farrell MJ, Glennon E, Guth S, Joseph MB, Mollentze N, Neely BA, Poisot T, Rasmussen AL, Ryan SJ, Seifert S, Sjodin AR, Sorrell EM, Carlson CJ. 2021. The science of the host-virus network. *Nat Microbiol* 6:1483-1492.
2. Han BA, Schmidt JP, Alexander LW, Bowden SE, Hayman DT, Drake JM. 2016. Undiscovered bat hosts of filoviruses. *PLoS Negl Trop Dis* 10:e0004815.
3. Albery GF, Eskew EA, Ross N, Olival KJ. 2020. Predicting the global mammalian viral sharing network using phylogeography. *Nat Commun* 11:2260.
4. Becker DJ, Albery GF, Sjodin AR, Poisot T, Bergner LM, Chen B, Cohen LE, Dallas TA, Eskew EA, Fagre AC, Farrell MJ, Guth S, Han BA, Simmons NB, Stock M, Teeling EC, Carlson CJ. 2022. Optimising predictive models to prioritise viral

- discovery in zoonotic reservoirs. *Lancet Microbe* doi:10.1016/S2666-5247(21)00245-7.
5. Babayan SA, Orton RJ, Streicker DG. 2018. Predicting reservoir hosts and arthropod vectors from evolutionary signatures in RNA virus genomes. *Science* 362:577-580.
 6. Mollentze N, Babayan SA, Streicker DG. 2021. Identifying and prioritizing potential human-infecting viruses from their genome sequences. *PLoS Biol* 19:e3001390.
 7. Takata MA, Goncalves-Carneiro D, Zang TM, Soll SJ, York A, Blanco-Melo D, Bieniasz PD. 2017. CG dinucleotide suppression enables antiviral defence targeting non-self RNA. *Nature* 550:124-127.
 8. Di Giallonardo F, Schlub TE, Shi M, Holmes EC. 2017. Dinucleotide composition in animal RNA viruses is shaped more by virus family than by host species. *J Virol* 91:e02381-16.
 9. Shackelton LA, Parrish CR, Holmes EC. 2006. Evolutionary basis of codon usage and nucleotide composition bias in vertebrate DNA viruses. *J Mol Evol* 62:551-63.
 10. Casalino L, Gaieb Z, Goldsmith JA, Hjorth CK, Dommer AC, Harbison AM, Fogarty CA, Barros EP, Taylor BC, McLellan JS, Fadda E, Amaro RE. 2020. Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent Sci* 6:1722-1734.
 11. Sztain T, Ahn SH, Bogetti AT, Casalino L, Goldsmith JA, Seitz E, McCool RS, Kearns FL, Acosta-Reyes F, Maji S, Mashayekhi G, McCammon JA, Ourmazd A, Frank J, McLellan JS, Chong LT, Amaro RE. 2021. A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nat Chem* 13:963-968.
 12. Jones JE, Le Sage V, Lakdawala SS. 2021. Viral and host heterogeneity and their effects on the viral life cycle. *Nat Rev Microbiol* 19:272-282.

13. Thompson AJ, de Vries RP, Paulson JC. 2019. Virus recognition of glycan receptors. *Curr Opin Virol* 34:117-129.
14. Maginnis MS. 2018. Virus-Receptor Interactions: The Key to Cellular Invasion. *J Mol Biol* 430:2590-2611.
15. Watanabe Y, Bowden TA, Wilson IA, Crispin M. 2019. Exploitation of glycosylation in enveloped virus pathobiology. *Biochim Biophys Acta Gen Subj* 1863:1480-1497.
16. Zhao J, Song E, Huang Y, Yu A, Mechref Y. 2021. Variability in the glycosylation patterns of gp120 proteins from different human immunodeficiency virus type 1 isolates expressed in different host cells. *J Proteome Res* 20:4862-4874.
17. Crocker PR, Paulson JC, Varki A. 2007. Siglecs and their roles in the immune system. *Nat Rev Immunol* 7:255-66.
18. Erikson E, Wratil PR, Frank M, Ambiel I, Pahnke K, Pino M, Azadi P, Izquierdo-Useros N, Martinez-Picado J, Meier C, Schnaar RL, Crocker PR, Reutter W, Keppler OT. 2015. Mouse Siglec-1 Mediates trans-Infection of Surface-bound Murine Leukemia Virus in a Sialic Acid N-Acyl Side Chain-dependent Manner. *J Biol Chem* 290:27345-27359.
19. West CM, Malzl D, Hykollari A, Wilson IBH. 2021. Glycomics, glycoproteomics, and glycogenomics: An inter-taxa evolutionary perspective. *Mol Cell Proteomics* 20:100024.
20. Everest-Dass AV, Moh ESX, Ashwood C, Shathili AMM, Packer NH. 2018. Human disease glycomics: technology advances enabling protein glycosylation analysis - part 1. *Expert Rev Proteomics* 15:165-182.
21. Moremen KW, Tiemeyer M, Nairn AV. 2012. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat Rev Mol Cell Biol* 13:448-62.
22. Varki A. 2017. Biological roles of glycans. *Glycobiology* 27:3-49.

23. Ohtsubo K, Marth JD. 2006. Glycosylation in cellular mechanisms of health and disease. *Cell* 126:855-67.
24. Schjoldager KT, Narimatsu Y, Joshi HJ, Clausen H. 2020. Global view of human protein glycosylation pathways and functions. *Nat Rev Mol Cell Biol* 21:729-749.
25. Spiro RG. 2002. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* 12:43R-56R.
26. Raman R, Tharakaraman K, Sasisekharan V, Sasisekharan R. 2016. Glycan-protein interactions in viral pathogenesis. *Curr Opin Struct Biol* 40:153-162.
27. Park YJ, Walls AC, Wang Z, Sauer MM, Li W, Tortorici MA, Bosch BJ, DiMaio F, Veerle D. 2019. Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors. *Nat Struct Mol Biol* 26:1151-1157.
28. Shinya K, Ebina M, Yamada S, Ono M, Kasai N, Kawaoka Y. 2006. Avian flu: influenza virus receptors in the human airway. *Nature* 440:435-6.
29. Yang Q, Hughes TA, Kelkar A, Yu X, Cheng K, Park S, Huang WC, Lovell JF, Neelamegham S. 2020. Inhibition of SARS-CoV-2 viral entry upon blocking N- and O-glycan elaboration. *Elife* 9.
30. Kuchipudi SV, Nelli R, White GA, Bain M, Chang KC, Dunham S. 2009. Differences in influenza virus receptors in chickens and ducks: Implications for interspecies transmission. *J Mol Genet Med* 3:143-51.
31. Stevens J, Blixt O, Glaser L, Taubenberger JK, Palese P, Paulson JC, Wilson IA. 2006. Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. *J Mol Biol* 355:1143-55.
32. Thompson AJ, Paulson JC. 2021. Adaptation of influenza viruses to human airway receptors. *J Biol Chem* 296:100017.

33. Chothe SK, Bhushan G, Nissly RH, Yeh YT, Brown J, Turner G, Fisher J, Sewall BJ, Reeder DM, Terrones M, Jayarao BM, Kuchipudi SV. 2017. Avian and human influenza virus compatible sialic acid receptors in little brown bats. *Sci Rep* 7:660.
34. Suenaga T, Arase H. 2014. Viral Interactions with glycans. *Glycoscience: Biology and Medicine* doi:10.1007/978-4-431-54841-6_152:785-794.
35. Cummings RD. 2009. The repertoire of glycan determinants in the human glycome. *Mol Biosyst* 5:1087-104.
36. Gao C, Wei M, McKittrick TR, McQuillan AM, Heimbürg-Molinaro J, Cummings RD. 2019. Glycan microarrays as chemical tools for identifying glycan recognition by immune proteins. *Front Chem* 7:833.
37. Werz DB, Ranzinger R, Herget S, Adibekian A, von der Lieth CW, Seeberger PH. 2007. Exploring the structural diversity of mammalian carbohydrates ("glycospace") by statistical databank analysis. *ACS Chem Biol* 2:685-91.
38. Kellman BP, Lewis NE. 2021. Big-data glycomics: Tools to connect glycan biosynthesis to extracellular communication. *Trends Biochem Sci* 46:284-300.
39. Alonso JM, Arroyuelo A, Garay PG, Martin OA, Vila JA. 2018. Finite dimension: A mathematical tool to analyse glycans. *Sci Rep* 8:4426.
40. Fischhoff IR, Castellanos AA, Rodrigues J, Varsani A, Han BA. 2021. Predicting the zoonotic capacity of mammals to transmit SARS-CoV-2. *Proc Biol Sci* 288:20211651.
41. Nguyen L, McCord KA, Bui DT, Bouwman KM, Kitova EN, Elaish M, Kumawat D, Daskhan GC, Tomris I, Han L, Chopra P, Yang TJ, Willows SD, Mason AL, Mahal LK, Lowary TL, West LJ, Hsu SD, Hobman T, Tompkins SM, Boons GJ, de Vries RP, Macauley MS, Klassen JS. 2021. Sialic acid-containing glycolipids mediate

- binding and viral entry of SARS-CoV-2. *Nat Chem Biol* doi:10.1038/s41589-021-00924-1.
42. Clausen TM, Sandoval DR, Spliid CB, Pihl J, Perrett HR, Painter CD, Narayanan A, Majowicz SA, Kwong EM, McVicar RN, Thacker BE, Glass CA, Yang Z, Torres JL, Golden GJ, Bartels PL, Porell RN, Garretson AF, Laubach L, Feldman J, Yin X, Pu Y, Hauser BM, Caradonna TM, Kellman BP, Martino C, Gordts P, Chanda SK, Schmidt AG, Godula K, Leibel SL, Jose J, Corbett KD, Ward AB, Carlin AF, Esko JD. 2020. SARS-CoV-2 infection depends on cellular heparan sulfate and ACE2. *Cell* 183:1043-1057 e15.
 43. Burkholz R, Quackenbush J, Bojar D. 2021. Using graph convolutional neural networks to learn a representation for glycans. *Cell Rep* 35:109251.
 44. Jia N, Byrd-Leotis L, Matsumoto Y, Gao C, Wein AN, Lobby JL, Kohlmeier JE, Steinhauer DA, Cummings RD. 2020. The human lung glycome reveals novel glycan ligands for influenza A virus. *Sci Rep* 10:5320.
 45. Short KR, Richard M, Verhagen JH, van Riel D, Schrauwen EJ, van den Brand JM, Manz B, Bodewes R, Herfst S. 2015. One health, multiple challenges: The inter-species transmission of influenza A virus. *One Health* 1:1-13.
 46. Jensen PH, Karlsson NG, Kolarich D, Packer NH. 2012. Structural analysis of N- and O-glycans released from glycoproteins. *Nat Protoc* 7:1299-310.
 47. Li Q, Xie Y, Wong M, Barboza M, Lebrilla CB. 2020. Comprehensive structural glycomic characterization of the glycocalyxes of cells and tissues. *Nat Protoc* 15:2668-2704.
 48. Kolarich D, Jensen PH, Altmann F, Packer NH. 2012. Determination of site-specific glycan heterogeneity on glycoproteins. *Nat Protoc* 7:1285-98.

49. Hart GW, Wells L. 2021. Glycoproteomics: Making the study of the most structurally diverse and most abundant post-translational modifications more accessible to the scientific community. *Mol Cell Proteomics* 20:100086.
50. Riley NM, Bertozzi CR, Pitteri SJ. 2020. A pragmatic guide to enrichment strategies for mass spectrometry-based glycoproteomics. *Mol Cell Proteomics* 20:100029.
51. Oliveira T, Thaysen-Andersen M, Packer NH, Kolarich D. 2021. The Hitchhiker's guide to glycoproteomics. *Biochem Soc Trans* 49:1643-1662.
52. Remoroza CA, Burke MC, Liu Y, Mirokhin YA, Tchekhovskoi DV, Yang X, Stein SE. 2021. Representing and Comparing Site-Specific Glycan Abundance Distributions of Glycoproteins. *J Proteome Res* 20:4475-4486.
53. Rojas-Macias MA, Mariethoz J, Andersson P, Jin C, Venkatakrishnan V, Aoki NP, Shinmachi D, Ashwood C, Madunic K, Zhang T, Miller RL, Horlacher O, Struwe WB, Watanabe Y, Okuda S, Levander F, Kolarich D, Rudd PM, Wuhrer M, Kettner C, Packer NH, Aoki-Kinoshita KF, Lisacek F, Karlsson NG. 2019. Towards a standardized bioinformatics infrastructure for N- and O-glycomics. *Nat Commun* 10:3275.
54. Munir K, Ashraf S, Munir I, Khalid H, Muneer MA, Mukhtar N, Amin S, Ashraf S, Imran MA, Chaudhry U, Zaheer MU, Arshad M, Munir R, Ahmad A, Zhao X. 2020. Zoonotic and reverse zoonotic events of SARS-CoV-2 and their impact on global health. *Emerg Microbes Infect* 9:2222-2235.
55. Anonymous. 2020. Emerging zoonoses: A one health challenge. *EClinicalMedicine* 19:100300.
56. Lurie N, Saviile M, Hatchett R, Halton J. 2020. Developing Covid-19 vaccines at pandemic speed. *N Engl J Med* 382:1969-1973.

57. Uraki R, Kawaoka Y. 2021. Host glycolipids in SARS-CoV-2 entry. *Nat Chem Biol*
doi:10.1038/s41589-021-00923-2.
58. Shiliaev N, Lukash T, Palchevska O, Crossman DK, Green TJ, Crowley MR, Frolova EI, Frolov I. 2021. Natural and recombinant SARS-CoV-2 isolates rapidly evolve in vitro to higher infectivity through more efficient binding to heparan sulfate and reduced S1/S2 cleavage. *J Virol* 95:e0135721.