



Bridging Glycomics and Genomics: New Uses of Functional Genetics in the Study of Cellular Glycosylation

Natalie Stewart^{1,2} and Simon Wisnovsky^{2*}

¹Biochemistry and Microbiology Dept, University of Victoria, Victoria, BC, Canada, ²Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, Canada

All living cells are coated with a diverse collection of carbohydrate molecules called glycans. Glycans are key regulators of cell behavior and important therapeutic targets for human disease. Unlike proteins, glycans are not directly templated by discrete genes. Instead, they are produced through multi-gene pathways that generate a heterogenous array of glycoprotein and glycolipid antigens on the cell surface. This genetic complexity has sometimes made it challenging to understand how glycosylation is regulated and how it becomes altered in disease. Recent years, however, have seen the emergence of powerful new functional genomics technologies that allow high-throughput characterization of genetically complex cellular phenotypes. In this review, we discuss how these techniques are now being applied to achieve a deeper understanding of glyco-genomic regulation. We highlight specifically how methods like ChIP-seq, RNA-seq, CRISPR genomic screening and scRNA-seq are being used to map the genomic basis for various cell-surface glycosylation states in normal and diseased cell types. We also offer a perspective on how emerging functional genomics technologies are likely to create further opportunities for studying cellular glycobiology in the future. Taken together, we hope this review serves as a primer to recent developments at the glycomics-genomics interface.

Keywords: glycomics, genomics, CRISPR, RNA-seq, ChIP-seq, CRISPR screen, sc-RNA seq

INTRODUCTION

Glycosylation is a common property of all cellular life. In human cells, a number of different monosaccharides can be covalently connected to produce a range of extended, branched oligosaccharides called glycans. These molecules can then be covalently linked to discrete asparagine (*N*-linked glycosylation) or serine/threonine (*O*-linked glycosylation) residues located on secreted or cell-surface proteins. Glycans can also be directly attached to membrane-embedded lipid molecules (Varki A, 2015–2017). Specific glycans serve as ligands for a range of carbohydrate-binding signaling receptors, which play important roles in regulating the activation and differentiation of immune cells (Thiemann and Baum, 2016; Brown et al., 2018; Duan and Paulson, 2020; Smith and Bertozzi, 2021). Like other post-translational modifications, glycosylation can also act as a molecular switch that modulates the binding of signaling receptors to their ligands (Gee et al., 2003; Li et al., 2016). Finally, glycans can mediate complex, multimolecular interactions that broadly regulate cell-surface trafficking and residency of glycosylated proteins. These molecules thus profoundly influence all aspects of cell signaling and metabolism (Demetriou et al., 2001; Lau et al., 2007; Lau and Dennis, 2008).

OPEN ACCESS

Edited by:

Kayvon Pedram,
Janelia Research Campus,
United States

Reviewed by:

Ryan Weiss,
University of Georgia, United States
Vered Padler-Karavani,
Tel Aviv University, Israel

*Correspondence:

Simon Wisnovsky
simon.wisnovsky@ubc.ca

Specialty section:

This article was submitted to
Glycoscience,
a section of the journal
Frontiers in Molecular Biosciences

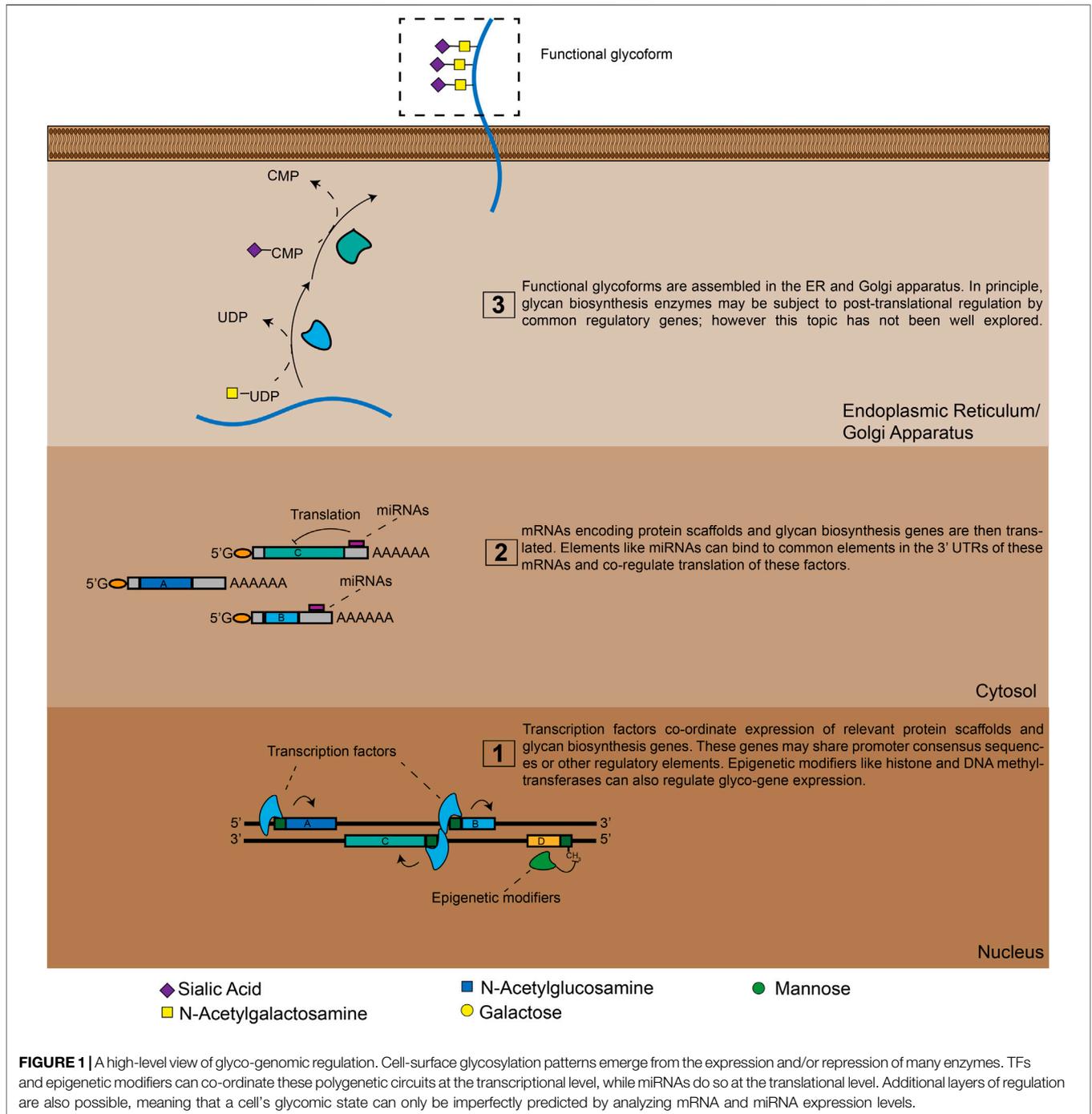
Received: 02 May 2022

Accepted: 26 May 2022

Published: 16 June 2022

Citation:

Stewart N and Wisnovsky S (2022)
Bridging Glycomics and Genomics:
New Uses of Functional Genetics in the
Study of Cellular Glycosylation.
Front. Mol. Biosci. 9:934584.
doi: 10.3389/fmolb.2022.934584



Like proteins, glycans are functional biopolymers made up of discrete, modular building blocks. However, unlike proteins, the information required to make a specific glycan is not directly encoded in the DNA sequence of a corresponding gene. Glycans are instead produced through complex circuits that involve coordinated expression of many different genes (Figure 1). At the most basic level, glycans are constructed by biosynthetic enzymes like glycosyltransferase (GT). These factors build glycan chains by selectively attaching monosaccharides to specific underlying

oligosaccharide substrates (Figure 1) (Varki A, 2015; Moremen and Haltiwanger, 2019). In the human genome there are hundreds of genes that encode GTs with different underlying substrate specificities (Narimatsu et al., 2018). The set of glycan structures produced by a cell will thus be greatly influenced by which GT genes are actively expressed in that cell type (Varki A, 2015–2017). Additionally, the ability of glycans to perform specific functions can depend on the density and/or orientation in which glycans are presented on the cell-surface

(Sako et al., 1995; Hashimoto et al., 2019). As a result, many glycans only acquire distinct functional properties when they are attached to a specific protein scaffold (Sako et al., 1995; Sackstein, 2011). Genes encoding these scaffolding molecules can therefore also be key components of glyco-regulatory genetic circuits (Figure 1). Finally, glycan biosynthesis is also broadly regulated by secondary mechanisms that tune glyco-ligand production through modulating GT localization and metabolic intermediate availability (Lau et al., 2007; Smith and Lupashin, 2008). GT is thus an emergent cellular property that arises from a staggeringly complex set of genetic interactions.

From this picture, a key question naturally arises: how are these polygenetic circuits regulated? As with other complex cellular processes, there are likely upstream “master regulators” that co-ordinate the biosynthesis of functional glycans by simultaneously modulating the activities of many different genes. In cases where aberrant glycosylation is involved in disease pathogenesis, these factors may serve as important disease biomarkers or novel therapeutic targets. Identifying new glyco-genetic circuits and the mechanisms that regulate their activity has thus always been a significant focus of glycoscience research. In particular, a defining trend of the past two decades has been a rapid increase in the reliability and throughput of various functional genomics methodologies and protocols. Instead of assessing gene function in a targeted, sequential way, researchers now can leverage next-generation sequencing (NGS) technology to routinely perform multiplexed analysis and/or perturbation of thousands of genes in a single biological experiment. Specific significant advances have included: 1) the progressive lowering of costs for whole-exome transcriptomic profiling by RNA-seq (Stark et al., 2019), 2) the methodical identification of transcription factor binding sites on a genome-wide scale by ChIP-seq (Wu et al., 2015), 3) functional genomic screening by shRNA and CRISPR/Cas9-based gene manipulation (Shalem et al., 2014), and 4) single cell gene expression profiling by single cell RNA-seq (scRNA-seq) (Yan et al., 2020; Nayak and Hasija, 2021). These advances have offered unprecedented opportunities for systematic dissection of genetically complex cellular phenotypes.

In the past few years, we have observed many new and exciting applications of these functional genetics methods towards the study of glycans in cell biology and disease. This trend of growing integration between glycomics and genomics appears to be accelerating. We believe research at this interface is poised to generate sweeping new insights into the regulation of glycosylation in human cells. In this review, we highlight some recent, significant developments in this area. While we have attempted to cover as much ground as possible, this discussion is not intended to be completely comprehensive. As these functional genomics techniques have been most broadly applied to mammalian cell models, we focus here on human rather than on microbial glycosylation. We also primarily discuss cell-surface carbohydrates rather than cytosolic (O-GlcNAc) glycosylation. While O-GlcNAc plays a fascinating set of roles in many biological processes, the biosynthesis of these glycans is subject to somewhat less complex genetic regulation when compared to cell-surface glycosylation (Yang and Qian, 2017).

Our primary goal is to give glycobiochemistry researchers a sense of how functional genomics techniques and resources can be usefully applied to general questions that may be of interest in their specific research areas. At the end of the article, we also discuss how new, leading-edge techniques may create further opportunities for glyco-genomics research in the future.

Chromatin Immunoprecipitation Sequencing and RNA-Seq in Glycoscience: Identifying Master Regulators of Glycan Biosynthesis

Transcription factors (TFs) are key regulators of gene expression in mammalian cells. TFs usually bind to specific DNA sequence motifs found in gene promoter or enhancer regions (Pan et al., 2010; de Boer et al., 2020). Subsequent TF-mediated recruitment or blockade of factors like RNA polymerases can then either activate or repress gene expression (Pan et al., 2010; de Boer et al., 2020). Different TFs also interact with one another in complex gene regulatory networks, making the function of a given TF somewhat cell- and context-dependent (Jolma et al., 2015; Stampfel et al., 2015). By binding to regulatory regions in multiple genes with shared consensus sequences, specific TFs can simultaneously modulate expression of many factors within a specific biosynthetic pathway (Figure 1). This feature has always made them attractive targets for investigation as regulators of glycan biosynthesis (Guo and Pierce, 2015; Neelamegham and Mahal, 2016). TF(s) also participate in chromatin remodeling and DNA methylation, which has been shown to play a key role in long-term (epigenetic) transcriptional repression of various glycan biosynthesis genes (Bui et al., 2010; Greville et al., 2016; Indelicato and Trinchera, 2021).

Significant effort has thus been expended by a number of groups into identifying conserved TF-binding sites within the promoter regions of various GT enzymes. Historically, these studies had to be conducted on a targeted basis. Painstaking work was often required to sequence gene promoter regions and identify consensus sites for transcription factors that had known binding motifs. Investigators could also take the complementary approach. TFs of known importance could be overexpressed in cells and changes in expression of various glycozymes and/or cell-surface glycosylation patterns could be determined using various bioanalytic methods. These types of studies have identified a range of transcriptional regulators of enzymes from the MGAT (Zhang et al., 2000; Xu et al., 2011), ST3GAL (Taniguchi et al., 2003), ST6GAL (Taniguchi et al., 2000) and FUCT (Higai et al., 2008) families, among many others. This critical foundational work is too extensive to adequately describe in detail here, but has been well-reviewed elsewhere (Guo and Pierce, 2015; Neelamegham and Mahal, 2016).

In recent years there has been a growing appreciation of the ways that aberrant glycosylation can drive cancer progression (Taylor-Papadimitriou et al., 1999; Saeland et al., 2007; Lin et al., 2014; Beatson et al., 2016; Agrawal et al., 2017). This has led to a specific focus on dissecting how various oncogenes regulate GT expression in cancer cells. One foundational study, for example, demonstrated that overexpression of mutant k-Ras increases

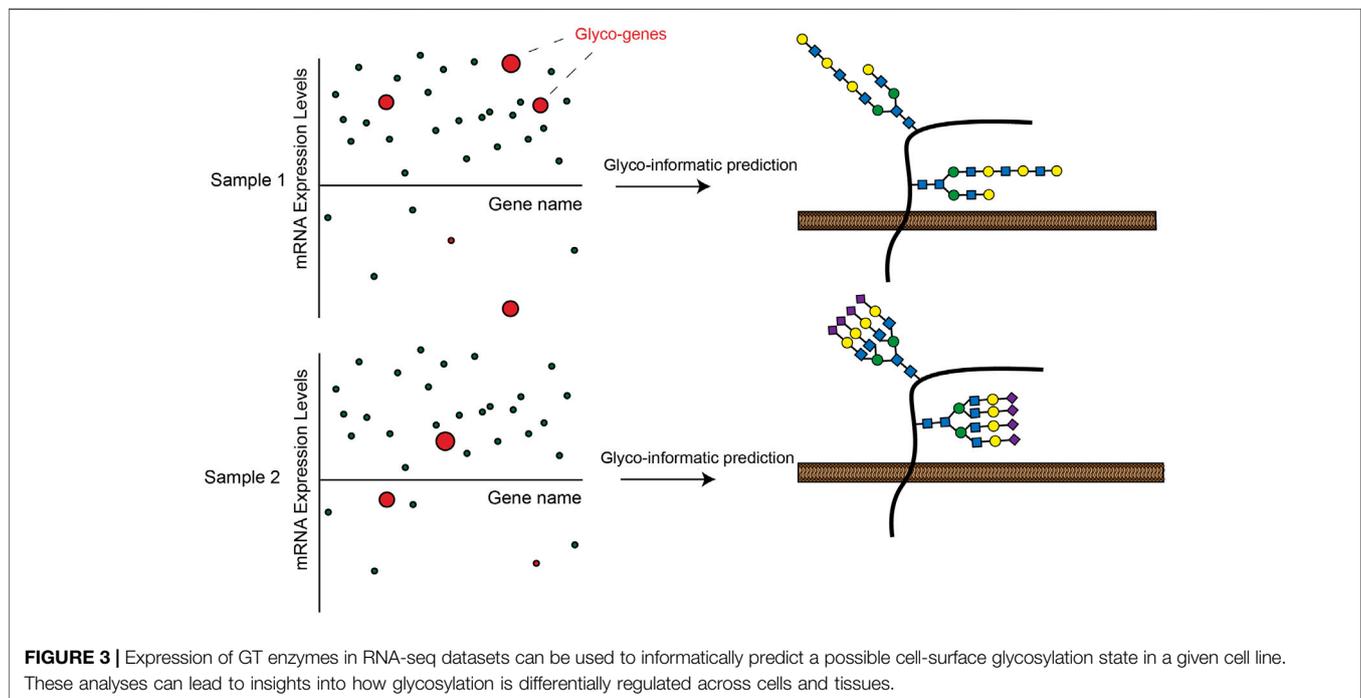
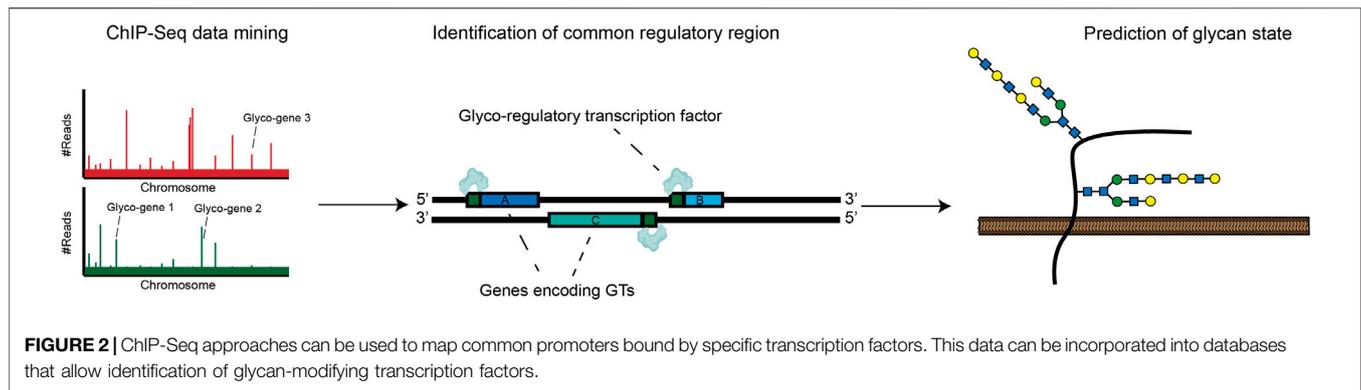
2,6-sialylation of β -integrins in several cancer cell lines. These effects were found to be mediated by downstream transcriptional activation of the ST6GAL1 GT gene (Seales et al., 2003). ST6GAL1 expression is thus likely induced by TFs downstream of k-Ras signaling. More recent studies have since identified such transcriptional regulators, most notably the TF SOX2, which has been found to drive cancer progression through driving expression of both ST6GAL1 (Dorsett et al., 2019) (in ovarian cancer) and ST3GAL1 (Pietrobono et al., 2020) (in metastatic melanoma). These advances in understanding cancer-specific regulation of 2,6-linked sialoglycan expression were recently well-reviewed (Dorsett et al., 2021). Another key study showed that the TFs c-Myc and CDX2 co-operate to induce expression of selectin ligands in colon cancer cells during epithelial-mesenchymal transition (Sakuma et al., 2012). These changes in glycosylation are mediated by co-ordinated transcriptional regulation of multiple GTs, in particular sialyl and fucosyl-transferases involved in selectin ligand biosynthesis (Sakuma et al., 2012). Separate glycoproteomics-based studies identified another member of this TF family, CDX1, as a key regulator of N-glycan biosynthesis in colorectal carcinoma (CRC) (Holst et al., 2016).

Integration of glycoproteomics and transcriptomic data was also recently used to identify several transcriptional regulators of glycosphingolipid biosynthesis in acute myeloid leukemia (AML) and O-glycan biosynthesis in CRC (Madunić et al., 2021; Wang et al., 2022). Leung et al. also recently used high-resolution mass spectrometry methods to globally map changes in glycosylation induced by overexpression of six common oncogenes (e.g., b-Raf, MEK, AKT) in human cell lines (Leung et al., 2020). These studies revealed significant changes in the biosynthesis of complex N-glycans upon overexpression of several oncogenes. Finally, a study by Möckl and Pedram et al. used super-resolution microscopy to show that k-Ras drives increased “thickness” of the cell-surface glycocalyx, the outer layer of glycosylation that coats all human cells. This effect was found to be mediated by increased expression of the GALNT7 GT enzyme (Möckl et al., 2019). Taken together, these studies illustrate how novel mechanisms of glyco-genomic regulation can be revealed through hypothesis-driven, targeted studies on known oncogenic factors.

A significant complementary trend has been the development of resources that allow for unbiased identification of glyco-regulatory transcription factors. In Chromatin Immunoprecipitation Sequencing (ChIP-seq), TF-DNA complexes are cross-linked and isolated by affinity capture with targeted antibodies. High-throughput sequencing of associated DNA allows for identification of TF-binding sites throughout the genome. Using ChIP-seq, TF-binding motifs can be identified experimentally rather than computationally, reducing the rate of false positive bioinformatic predictions. ChIP-seq is now a mature technology that has been broadly used to produce detailed maps of promoter and enhancer-binding activities for hundreds of cellular transcription factors (Park, 2009; Wu et al., 2015). These results are now readily accessible in online databases, enhancing the throughput with

which regulators of glycosylation can be identified. In a recent elegant work, Weiss et al. analyzed ChIP-seq datasets to identify conserved TF-binding sites in the promoter regions of twenty heparan sulfate (HS) biosynthesis genes. This study revealed a specific transcription factor, ZNF263, that binds to ~50% of gene promoters within this complex biosynthetic pathway. CRISPR/Cas9 knockout of ZNF263 in human cells dramatically increased expression of several enzymes involved in HS biosynthesis and enhanced sulfation of HS chains. This study implies that modulation of ZNF263 expression may be a novel strategy for optimizing cell-based production of heparin, one of the most widely used biopharmaceuticals in the clinic (Weiss et al., 2020). In another recent paper, Groth et al. applied this same type of strategy on a larger scale, mining the Cistrome Cancer Database (which integrates ChIP-Seq data from many cancers) to identify TF-glycogene relationships at a systems level. This novel pipeline identified hundreds of potential glyco-regulatory TFs, creating a rich resource to guide hypothesis generation for the cancer glycoscience community (Groth et al., 2021). These approaches represent generally applicable strategies for discovering regulators of glycan biosynthesis in a range of biological contexts (Figure 2).

Another key development of recent years has been the increasing throughput and declining cost of RNA-seq transcriptomic profiling. This technique allows multiplexed, quantitative measurement of gene expression levels across the protein coding genome. Researchers have used RNA-seq to generate huge quantities of transcriptomic data from a range of cell lines and human tissue samples (Stark et al., 2019). Many of these datasets have since been incorporated into easily searchable online databases such as the Human Protein Atlas (HPA) (Uhlén et al., 2015), The Cancer Genome Atlas Program (TCGA) (Cancer Genome Atlas Research et al., 2013) and the Broad Cancer Cell Line Encyclopedia (Barretina et al., 2012). Several recent studies have used these databases to generate new insights into tissue-specific regulation of cellular glycosylation pathways. In one groundbreaking work, Huang et al. extracted RNA-seq data from the HPA and systematically mapped mRNA expression levels for 950 glycan biosynthesis genes across a selection of human cell lines. The investigators went on to show that integrated transcriptomic analysis of glycogene expression could effectively predict the complement of glycan structures synthesized by a given cell type (Figure 3) (Huang et al. 2021b). In a follow-up study, the investigators applied this tool (GlycoMaple) to study differential expression of glycosaminoglycan (GAG)-related enzymes in various human tissues. This body of work uncovered alterations to GAG biosynthesis in a number of breast cancer subtypes (Huang et al., 2021c). Similar approaches have been used to investigate regulation of glycogene expression in CRC (Wu et al., 2021), cervical cancer (Martinez-Morales et al., 2021), hepatic cancer (Angata et al., 2020) and pluripotent stem cells (Pecori et al., 2021). To facilitate such studies, Thambu et al. also recently developed aneXVis, a visualization tool that similarly facilitates transcriptomic analysis of glycan biosynthesis genes (Thambu and Balagurunathan 2022). Systematically dissecting transcriptomic data across a broad range of tissues and disease states creates opportunities for identifying cell and tissue-specific regulators of glycosylation. In one work, for example, Zeng et al. analyzed TCGA



data to identify glycosphingolipid biosynthesis genes that are differentially expressed in triple negative breast cancer (TNBC) patients (Zeng et al., 2021). Further dissection of ChIP-seq databases identified the transcription factors AR and GATA3 as putative regulators of glyco-gene expression in TNBC (Zeng et al., 2021).

Taken together, we expect that these resources will catalyze rapid discovery of new glyco-regulatory mechanisms in the future. In particular, we still see large gains to be made from the integrated study of cancer genomics and glycomics. While the role of oncogenes in driving biosynthesis of specific cancer-promoting glycans is becoming better understood, there has been almost no work studying how loss of tumor suppressor genes affects cell-surface glycosylation. For example, out of the ~11,000 publications on PubMed that mention the tumor suppressor gene p53, we were not able to identify a single article that examined the role of this protein in regulating

glyco-gene expression or glycosylation. This is a gap in the field we anticipate will be filled in coming years.

Multi-Layered Regulation: miRNAs and the Cell-Surface Glycome

As discussed above, analyzing the expression of GT-encoding mRNAs can yield important insights into how the glycome is regulated in different cells and tissues. There is some debate, however, as to how consistently mRNA expression data can be used to predict cell-surface glycosylation patterns. Transcriptional activation and repression of glycan biosynthesis genes is clearly one key mechanism that determines which glycan structures will be synthesized by a cell (Neelamegham and Mahal, 2016). However, glycan biosynthesis is also subject to other complex forms of

regulation. The expression of an mRNA encoding a specific GT does not guarantee that the mRNA will be translated at high levels or that the corresponding GT will be enzymatically active. Predictions of cell-surface glycosylation states made using transcriptomic data must therefore be carefully calibrated. Hypotheses generated by such approaches should always be validated using state-of-the-art methods for glycan analysis (ideally mass-spectrometry-based glycoproteomics) (Huang et al., 2021b).

The role of micro RNAs (miRNAs) in regulating glycan biosynthesis provides a clear example of this additional complexity. Since the discovery of miRNAs in the early 1990s, evidence has steadily emerged implicating these factors as powerful regulators of gene expression in human cells (Saliminejad et al., 2019). miRNAs are small, non-coding RNAs, typically 22 base pairs in length, that bind to target mRNAs (usually at the 3'-UTR) through complementary base pairing. The typical effect of miRNA binding is to inhibit protein translation or destabilize the target mRNA (Macfarlane and Murphy, 2010). Because miRNA sequences must be complementary to their target mRNA, putative miRNA target sequences can be readily predicted using bioinformatic analysis. In a classic paper, Kurcon et al. informatically identified possible miRNA targets of the miRNA-200 family. The investigators found that miRNAs transcribed from this cluster co-regulate translation of multiple glycoenzymes, including the GTs B3GLCT, ST3GAL5 and ST6GALNAC5. This “miRNA proxy approach” revealed a linked role for these enzymes in regulating epithelial-mesenchymal transition (EMT) (Kurcon et al., 2015). Multiple other studies have since explored the role of various miRNAs in regulating cancer-associated activity of specific GTs (Gaziel-Sovran et al., 2011; Li et al., 2017; Wu et al., 2017; Ma et al., 2021). This extensive body of work was recently comprehensively reviewed (Thu and Mahal, 2020). In general, the capacity of miRNAs to simultaneously co-regulate translation of many glycan biosynthesis implicates them as “master regulators” of glycan biosynthesis (Figure 1).

It is thus crucial to integrate miRNA expression profiling into systems models that attempt to predict cellular glycosylation states from transcriptomic data. One important aspect of this work is developing better experimental methods for determining exactly which miRNAs regulate a given GT mRNA. Computational methods for miRNA target prediction are known to suffer from high false positive rates (Thu et al., 2021). In one recent paper, Thu et al. developed an elegant high-throughput screening platform (miRFluR) to rapidly identify miRNA clusters that regulate translation of the B3GLCT GT gene (Thu et al., 2021). This study revealed a much narrower set of miRNAs that modulate B3GLCT than would be predicted from computational studies. In the future, we expect that expansion of this approach to a wider set of GT targets will refine our understanding of miRNA-mediated glycome regulation. Even further complexity is also imaginable. Long non-coding RNAs (lncRNAs) for instance, have been shown to play a crucial role in antagonizing miRNA function through competitive inhibition (Xue et al., 2017). One interesting recent

study found that a lncRNA with homology to the ST8SIA6 gene (ST8SIA6-AS1) was significantly elevated in breast cancer tissues and promotes invasiveness of breast cancer cell lines (Fang et al., 2020). Findings like this may imply an interesting role for lncRNAs in further tuning glycan biosynthesis, although this hypothesis has yet to be comprehensively explored. These considerations further highlight the need to integrate regulatory RNAs into the next-generation of glyco-informatics models.

Applications of High-Throughput Genomic Screening in Glycoscience

Functional genetic screening is a powerful strategy for defining genetic regulators of cellular processes and phenotypes. Broadly speaking, functional genetic screening refers to experimental approaches in which the cellular expression of many genes is disrupted in a multiplexed, parallelized manner (Heynen-Genel et al., 2012; Agrotis and Ketteler, 2015). Changes in a given cellular phenotype are then assessed and linked back to specific gene disruption events. Modulation of gene expression can be performed in a variety of different cell models using a diverse array of molecular approaches. Methodologies for assessing and/or selecting for specific phenotypes similarly vary widely. Screens can also be *targeted*, in which a subset of genes of interest are assessed in parallel, or *unbiased*, in which gene function is studied on a genome-wide scale. Until the early-2000s, the model system of choice for large-scale genomic screening studies was budding yeast. Yeast cells are extremely genetically tractable and faithfully recapitulate many evolutionarily conserved aspects of human biology (Enserink, 2012; Adames et al., 2019). However, yeast cells exhibit dramatic differences in glycan structure compared with mammalian cells (Varki A, 2015–2017), which historically limited the applicability of yeast screening techniques to the study of glyco-genomic regulation. Glycoscience researchers have instead often made use of Chinese hamster ovary (CHO) cell collections with defined genetic knockouts in genes encoding GT enzymes. These resources can be effectively used to assess the genetic basis for some glycan-dependent phenotypes, such as cell-surface binding specificities of recombinant lectins (Yang et al., 2015; Nielsen et al., 2018; Nielsen and Wandall, 2022). However, CHO cells are also not a perfect model for human glycosylation, and the number of genes that can be assessed using these methods is relatively small.

The advent and commercialization of RNA interference (RNAi) technology in the early 2000s opened up new possibilities for high-throughput manipulation of gene expression in mammalian cells (Elbashir et al., 2001). These possibilities further expanded with the subsequent development of CRISPR-Cas9 gene editing in the 2010s (Garneau et al., 2010; Jinek et al., 2012). CRISPR-Cas9 allows for genes to be directly disrupted in human cells with unprecedented ease and specificity. Researchers have since adapted CRISPR-Cas9 methods to dissect and engineer cellular glycosylation pathways in living cells. In one foundational work, Stolfa et al. used CRISPR-Cas9 to engineer a small number of cell

lines with knockouts in specific glycan biosynthesis pathways (Stolfa et al., 2016). These cell line tools were then used to dissect the contribution of various glycan subtypes to leukocyte adhesion. In a more comprehensive work, CRISPR-Cas9 gene editing was used to generate a panel of isogenic HEK293 cell lines with knockouts in a wide range of key GT enzymes (Narimatsu et al., 2018; Narimatsu et al., 2019). Using this “cell-based glycan array”, recombinant lectins can be rapidly screened to determine specific glycan structures required for binding in living human cells (Büll et al., 2020). One study, for example, applied this approach to identify GTs whose knockout ablates cell-surface binding of Siglecs, immune receptors emerging as targets of interest for cancer immunotherapy (Büll et al., 2021). These methods can also be used in more sophisticated model systems. One recent work used CRISPR-Cas9 to engineer a library of 3D organoid tissues to exhibit various cell-surface glycosylation states (Dabelsteen et al., 2020). This study revealed a key role for specific glycan structures in regulating efficient skin formation (Dabelsteen et al., 2020). Such resources create rich opportunities for array-based functional screening focused on the portion of the genome that encodes glycan biosynthesis genes like GTs.

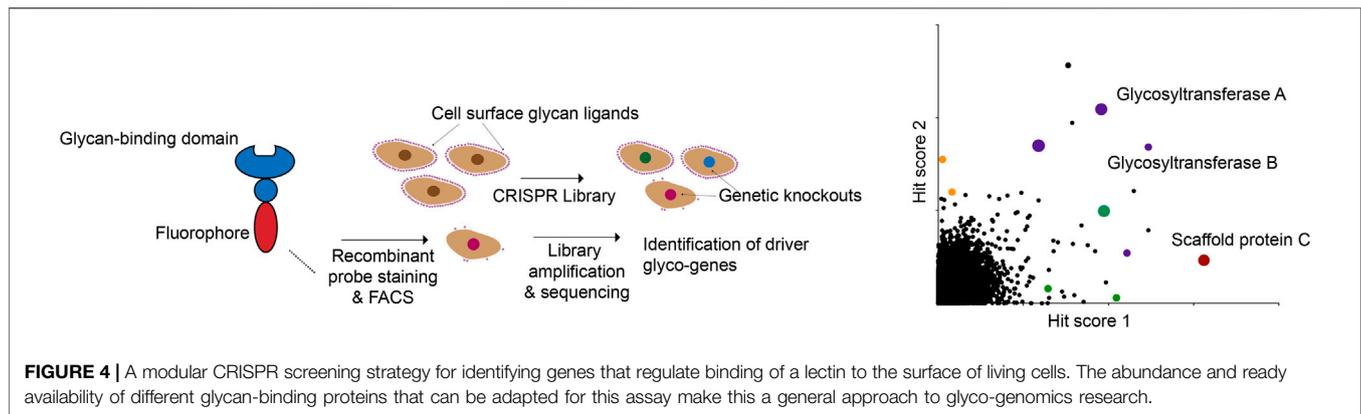
Soon after its discovery, CRISPR-Cas9 technology was quickly integrated into pooled screening protocols, permitting functional genomics experiments to be conducted on a much larger scale (Shalem et al., 2014; Wang et al., 2014). In CRISPR-based genome-wide screening, a cell line expressing Cas9 is transduced in bulk with a mixed plasmid library containing thousands of sgRNAs. This process creates a pooled, genetically heterogeneous cell population in which every cell exhibits a different knockout in a single functional gene. This “cell library” can then be subjected to various types of phenotypic selection, most commonly through exposure to cytotoxic agents or sorting of specific cell populations by fluorescence-activated cell sorting (FACS). NGS is then used to assess how the relative abundance of each sgRNA in the library changes in response to selection. sgRNAs whose abundance shifts significantly in response to selection likely target genes that are functionally linked to the phenotype of interest (Sanjana, 2017; Doench, 2018).

Pooled screening does have some disadvantages relative to arrayed screening approaches. Pooled screens can suffer from high false discovery rates and must be subjected to careful statistical analysis in order to reveal true hits (Sanjana, 2017; Doench, 2018). They also cannot be used to investigate non-cell intrinsic and non-selectable cell phenotypes. However, pooled screening also has some crucial advantages. Most notably, these experiments are: 1) unbiased, in that they assess the function of every gene in the genome, 2) can be completed in a reasonable time frame and 3) use resources commonly found in a typical molecular biology laboratory. Pooled CRISPR screening has thus become an engine of discovery in the life sciences, helping to decode the genetic basis for many cellular phenotypes. Reagents and protocols for genome-wide screening are now broadly available. Recent years have seen the first applications of these powerful methods to the study of glycobiology.

The first of these CRISPR screening studies made use of bacterial toxins (e.g., ricin and Shiga-like toxins) that enter cells through binding to glycan receptors on the cell-surface (Morgens et al., 2017; Tian et al., 2018; Yamaji et al., 2019). Such toxins are highly lethal to human cells, meaning that pooled CRISPR libraries can be easily selected to identify sgRNAs that produce toxin resistance. In several studies, this approach revealed a number of novel genes that are essential for biosynthesis and trafficking of cell-surface glycoconjugates. In addition to recovering known GT enzymes, these screens also implicated some poorly characterized lysosomal (e.g., LAPTM4A) and Golgi (e.g., TMEM165 and TM9SF2) proteins as key regulators of glycan expression (Tian et al., 2018; Yamaji et al., 2019). These screens thus suggest novel mechanisms of glycan synthesis regulation within the secretory pathway, providing interesting directions for future research.

Subsequent work has used CRISPR-Cas9 screening to probe more specific glycan biosynthesis pathways. In one study, cells were infected with a genome-wide library of sgRNAs and incubated with a fluorescently-labelled ligand that binds selectively to HS. Cell sorting techniques (FACS) were then used to isolate cells exhibiting reduced binding. Subsequent sgRNA sequencing revealed the KDM2B gene as a top hit. In follow-up work, KDM2B inactivation was shown to inhibit expression of multiple sulfotransferase enzymes and upregulate expression of sulfatases, implicating this TF as a novel “master regulator” of HS biosynthesis (Weiss et al., 2021). Other studies have applied a similar FACS-based screening approach to identify regulators of glycosylation in cancer. In one recent work, the investigators used a genome-wide CRISPR interference (CRISPRi) screening strategy to identify genes that reduce binding of recombinant Siglec receptors to the surface of leukemia cells (Wisnovsky et al., 2021). This approach revealed a cluster of genes including 1) the GTs C1GALT1 and ST6GALNAC1 and 2) the cell-surface protein CD43 to be drivers of Siglec-7 ligand expression. Further investigation revealed that Siglec-7 binds to a distinct glycopeptide motif containing clusters of disialylated O-linked glycans (Wisnovsky et al., 2021). The GTs identified in the screen biosynthesize these glycans, while the CD43 protein acts as a scaffold that presents these carbohydrates at an unusually high spatial density. The combination of these two factors mediates selective Siglec-7 binding. These types of studies exemplify how genomic screening can be used to dissect the complex polygenic circuits that regulate glycan ligand expression.

These proof-of-principle studies open up numerous possibilities for the future. Decades of glycobiology research have uncovered dozens of lectins with various specificities for different classes of glycans (Bojar et al., 2022). In principle, any of these lectins (many of which are commercially available in large quantities) can be easily used to identify genetic regulators of specific glycans using FACS-based screening (**Figure 4**). In anticipation of these possibilities, CRISPR screening resources targeted specifically at glycobiology users are now being developed. A recent study, for example, described the generation of a small, pooled CRISPR library targeting 347 known glycan biosynthesis genes (Zhu et al., 2021). The



investigators used this library to identify genes essential for the cell-surface binding of several known lectins. Using a compact sgRNA library allows for screens to be performed quickly and at low cost, while still permitting multiplexed functional annotation of glycosylation-related genes. While genome-wide libraries will still be the tool of choice when an unbiased approach is needed, these types of targeted resources will also be of great use to the glycobiology community.

Dissecting Glycomic Heterogeneity: Glycan Analysis at Single-Cell Resolution

Finally, the most recent interesting trend to emerge in glyco-genomics has been the application of single cell profiling methods to the study of glycobiology. Cell-surface glycosylation notoriously displays significant biochemical heterogeneity. For example, glycoproteomics studies performed on bulk cell mixtures frequently discover that many different glycan structures can be appended to a single glycosylation site (Oliveira et al., 2021). In part, these findings likely reflect differential regulation of glycan biosynthesis by different cells within complex populations. The recent development of techniques like scRNA-seq now offer new opportunities for determining the genetic basis for these distinct cellular “glyco-states”. In scRNA-seq, single cell suspensions are disassociated from cultured cells or tissue and physically separated using a variety of methods (Nayak and Hasija, 2021). Microfluidics-based systems, in which separate cells are captured within individual microfluidic droplets, have gained particular popularity in recent years (Nayak and Hasija, 2021). Individual cells are frequently sorted *via* FACS to further isolate cells exhibiting expression of specific cell-surface markers. mRNA from individual cells is then amplified by RT-PCR and sequenced using NGS. These techniques can be used to build detailed gene expression maps for thousands of individual cells, providing a high-resolution look into the heterogeneity of gene expression within a cell population.

One important recent study pioneered a new technique (SUGAR-Seq) that integrates scRNA-seq and glycomics analysis. In this approach, tumor-infiltrating T-cells (TILs) were isolated from primary tumors and labelled with a lectin (L-Pha) that binds to complex branched N-glycans (Kearney

et al., 2021). Single TILs were then sorted based on L-Pha staining and analyzed by scRNA-seq to reveal differences in glycosylation amongst different TIL subsets. This study revealed several insights. Most notably, it was found that cells exhibiting regulatory and exhausted gene expression signatures were more densely N-glycosylated than active memory or effector T-cells. Follow-up glycoproteomic and lectin proximity labelling studies also indicated changes in glycosylation of immune checkpoint receptors and ligands commonly expressed on T-cells. These findings may thus have important implications for understanding T-cell regulation in cancer and for the design of new immune checkpoint inhibitors.

Another set of recently-published protocols also applied single cell profiling to study glycosylation in human induced pluripotent stem cells (Minoshima et al., 2021; Odaka et al., 2022). Single-cell suspensions were stained with several lectins, sorted by FACS and analyzed by scRNA-seq. These studies revealed that differences in glycosylation were associated with differential expression of genes associated with neural differentiation (Minoshima et al., 2021). The protocol used for this work was also separately described in a detailed methods article (Odaka et al., 2022). As with CRISPR screening, the availability of a wide range of lectins for conducting these types of studies creates a range of exciting possibilities. We thus anticipate that “single cell glycomics” will continue to develop in the future.

Future Perspectives: What’s Next for Glyco-Genomics?

In this review, we hope we have adequately summarized some of the most significant recent advances at the interface of genomics and glycomics. In the future, we believe continued progress in this area will allow researchers to build sophisticated, systems-oriented models of how glycosylation is regulated in living cells. In addition to the perspectives we have already offered, we see several aspects of this research that have significant potential for growth.

As has been discussed, transcriptomics datasets can now be harnessed to predict cellular glycosylation patterns with amazing speed and considerable accuracy. However, it is clear that quantitating GT expression levels is not sufficient to analyze

the glycome in its full complexity. Systems analyses of cellular glycosylation must account for other mechanisms by which glycosylation may be regulated. miRNA-mediated regulation is clearly relevant, but other areas remain chronically underexplored. Several CRISPR screening studies, for example, have identified specific ER and Golgi-relevant chaperones as key players in glycan biosynthesis (Tian et al., 2018; Yamaji et al., 2019; Wisnovsky et al., 2021). The tissue-specific expression and/or function of these types of molecules has yet to be analyzed in a systematic way. Another gap in the field is understanding the post-translational regulation of glycan biosynthesis. To take one example, many GTs and other Golgi-localized proteins have been found to be phosphorylated in living cells (Yu and Bieberich, 2001; Bergeron et al., 2017). However, there has not been significant recent exploration of how these phospho-sites are dynamically altered (e.g., by cancer-associated kinases) or how they functionally modulate GT activity. To develop truly comprehensive models of how glycosylation is regulated, we will also need to study these types of mechanisms and closely integrate them with insights from genomics studies.

As discussed above, we view the application of single-cell profiling techniques as a leading edge of this research area. In addition to scRNA-seq, other optimized technologies like single cell ATAC-seq (for analyzing epigenomic states) and Perturb-seq (integrating functional genetic screening with scRNA-seq) are ready to be applied to questions in glycobiology (Dixit et al., 2016; Fang et al., 2021). Current single cell profiling methods do face some technical and financial limits on sequencing depth (Haque et al., 2017; Nayak and Hasija, 2021). These factors can make it difficult to accurately profile expression of transcripts with low copy numbers. In the context of glyco-genomics, this means that integrated expression analysis of glycan biosynthesis genes (which is now routine for bulk RNA-seq) is likely not yet feasible at single cell resolution. However, single cell sequencing technologies are advancing at a rapid clip, and we may well see new technical breakthroughs that solve some of these challenges.

Finally, the incredible proliferation of different CRISPR-Cas9-based gene manipulation technologies is poised to create new breakthroughs in glycoscience. We conclude by suggesting several novel applications. Firstly, in CRISPR activation (CRISPRa) screening, sgRNA-mediated overexpression of

target genes is induced by recruitment of a dCas9-VPR fusion protein (VPR is a transcriptional activator). This technique allows for gain-of-function gene overexpression screens to be easily performed in human cell lines. CRISPRa can induce expression of genes that are not normally active in most model cell lines, allowing for broad functional annotation of tissue and context-specific genetic factors to be performed in a single cell type (Maeder et al., 2013; Gilbert et al., 2014). In future, we envision that CRISPRa screening could be usefully applied to dissect functions of genetically redundant, complementary glycogenes that would be missed by traditional loss-of-function approaches. Secondly, in multiplexed CRISPR screening, multiple sgRNAs are co-expressed from the same promoter, permitting simultaneous knockout of multiple genes in the same cell. Construction of unique barcoded combinatorial libraries allows for rapid screening of pairwise genetic interactions (Wong et al., 2016; Han et al., 2017). These types of techniques are ideal for studying the complex, polygenetic networks that regulate glycosylation. Finally, CRISPR screening in organoid cultures (Han et al., 2020), primary cells (Shifrut et al., 2018) and *in vivo* (Huang H. et al., 2021) also offer new options for dissecting glycan regulation in sophisticated model systems. The possibilities here are near limitless, and we are excited to see what new frontiers in glyco-genomics emerge from this type of research.

AUTHOR CONTRIBUTIONS

NS conducted an extensive literature review of recent manuscripts in the areas of glycomics and genomics that served as the basis for this article. NS and SW collaboratively wrote and edited the manuscript.

FUNDING

SW is supported by funding from the Canadian Institutes of Health Research (PJ4-179805) and the Natural Sciences and Engineering Research Council of Canada (RGPIN-2022-04448).

REFERENCES

- Adames, N. R., Gallegos, J. E., and Peccoud, J. (2019). Yeast Genetic Interaction Screens in the Age of CRISPR/Cas. *Curr. Genet.* 65, 307–327. doi:10.1007/s00294-018-0887-8
- Agrawal, P., Fontanals-Cirera, B., Sokolova, E., Jacob, S., Vaiana, C. A., Argibay, D., et al. (2017). A Systems Biology Approach Identifies FUT8 as a Driver of Melanoma Metastasis. *Cancer Cell.* 31, 804–819. e807. doi:10.1016/j.ccell.2017.05.007
- Agrotis, A., and Ketteler, R. (2015). A New Age in Functional Genomics Using CRISPR/Cas9 in Arrayed Library Screening. *Front. Genet.* 6, 300. doi:10.3389/fgene.2015.00300
- Angata, K., Sawaki, H., Tsujikawa, S., Ocho, M., Togayachi, A., and Narimatsu, H. (2020). Glycogene Expression Profiling of Hepatic Cells by RNA-Seq Analysis for Glyco-Biomarker Identification. *Front. Oncol.* 10, 1224. doi:10.3389/fonc.2020.11224
- Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., et al. (2012). The Cancer Cell Line Encyclopedia Enables Predictive Modelling of Anticancer Drug Sensitivity. *Nature* 483, 603–607. doi:10.1038/nature11003
- Beatson, R., Tajadura-Ortega, V., Achkova, D., Picco, G., Tsourouktsoglou, T.-D., Klausung, S., et al. (2016). The Mucin MUC1 Modulates the Tumor Immunological Microenvironment through Engagement of the Lectin Siglec-9. *Nat. Immunol.* 17, 1273–1281. doi:10.1038/ni.3552
- Bergeron, J. J. M., Au, C. E., Thomas, D. Y., and Hermo, L. (2017). Proteomics Identifies Golgi Phosphoprotein 3 (GOLPH3) with A Link between Golgi Structure, Cancer, DNA Damage and Protection from Cell Death. *Mol. Cell. Proteomics* 16, 2048–2054. doi:10.1074/mcp.mr117.000068
- Bojar, D., Meche, L., Meng, G., Eng, W., Smith, D. F., Cummings, R. D., et al. (2022). A Useful Guide to Lectin Binding: Machine-Learning Directed Annotation of 57 Unique Lectin Specificities. *ACS Chem. Biol.* doi:10.1021/acscchembio.1c00689

- Brown, G. D., Willment, J. A., and Whitehead, L. (2018). C-type Lectins in Immunity and Homeostasis. *Nat. Rev. Immunol.* 18, 374–389. doi:10.1038/s41577-018-0004-8
- Bui, C., Ouzzine, M., Talhaoui, I., Sharp, S., Prydz, K., Coughtrie, M. W. H., et al. (2010). Epigenetics: Methylation-associated Repression of Heparan Sulfate 3-O-sulfotransferase Gene Expression Contributes to the Invasive Phenotype of H-EMC-SS Chondrosarcoma Cells. *FASEB J.* 24, 436–450. doi:10.1096/fj.09-136291
- Büll, C., Joshi, H. J., Clausen, H., and Narimatsu, Y. (2020). Cell-Based Glycan Arrays-A Practical Guide to Dissect the Human Glycome. *Star. Protoc.* 1, 100017. doi:10.1016/j.xpro.2020.100017
- Büll, C., Nason, R., Sun, L., Van Coillie, J., Madriz Sørensen, D., Moons, S. J., et al. (2021). Probing the Binding Specificities of Human Siglecs by Cell-Based Glycan Arrays. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2026102118. doi:10.1073/pnas.2026102118
- Cancer Genome Atlas Research, N., Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R., Ozenberger, B. A., et al. (2013). The Cancer Genome Atlas Pan-Cancer Analysis Project. *Nat. Genet.* 45, 1113–1120. doi:10.1038/ng.2764
- Dabelsteen, S., Pallesen, E. M. H., Marinova, I. N., Nielsen, M. I., Adamopoulou, M., Rømer, T. B., et al. (2020). Essential Functions of Glycans in Human Epithelia Dissected by a CRISPR-Cas9-Engineered Human Organotypic Skin Model. *Dev. Cell.* 54, 669–684. doi:10.1016/j.devcel.2020.06.039
- De Boer, C. G., Vaishnav, E. D., Sadeh, R., Abeyta, E. L., Friedman, N., and Regev, A. (2020). Deciphering Eukaryotic Gene-Regulatory Logic with 100 Million Random Promoters. *Nat. Biotechnol.* 38, 56–65. doi:10.1038/s41587-019-0315-8
- Demetriou, M., Granovsky, M., Quaggin, S., and Dennis, J. W. (2001). Negative Regulation of T-Cell Activation and Autoimmunity by Mgat5 N-Glycosylation. *Nature* 409, 733–739. doi:10.1038/35055582
- Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C. P., Jerby-Arnon, L., et al. (2016). Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell.* 167, 1853–1866. doi:10.1016/j.cell.2016.11.038
- Doench, J. G. (2018). Am I Ready for CRISPR? A User's Guide to Genetic Screens. *Nat. Rev. Genet.* 19, 67–80. doi:10.1038/nrg.2017.97
- Dorsett, K. A., Jones, R. B., Ankenbauer, K. E., Hjelmeland, A. B., and Bellis, S. L. (2019). Sox2 Promotes Expression of the ST6Gal-I Glycosyltransferase in Ovarian Cancer Cells. *J. Ovarian Res.* 12, 93. doi:10.1186/s13048-019-0574-5
- Dorsett, K. A., Marciel, M. P., Hwang, J., Ankenbauer, K. E., Bhalerao, N., and Bellis, S. L. (2021). Regulation of ST6GAL1 Sialyltransferase Expression in Cancer Cells. *Glycobiology* 31, 530–539. doi:10.1093/glycob/cwaa110
- Duan, S., and Paulson, J. C. (2020). Siglecs as Immune Cell Checkpoints in Disease. *Annu. Rev. Immunol.* 38, 365–395. doi:10.1146/annurev-immunol-102419-035900
- Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs Mediate RNA Interference in Cultured Mammalian Cells. *Nature* 411, 494–498. doi:10.1038/35078107
- Enserink, J. M. (2012). Chemical Genetics: Budding Yeast as a Platform for Drug Discovery and Mapping of Genetic Pathways. *Molecules* 17, 9258–9273. doi:10.3390/molecules17089258
- Fang, K., Hu, C., Zhang, X., Hou, Y., Gao, D., Guo, Z., et al. (2020). LncRNA ST8SIA6-AS1 Promotes Proliferation, Migration and Invasion in Breast Cancer through the P38 MAPK Signalling Pathway. *Carcinogenesis* 41, 1273–1281. doi:10.1093/carcin/bgz197
- Fang, R., Preissl, S., Li, Y., Hou, X., Lucero, J., Wang, X., et al. (2021). Comprehensive Analysis of Single Cell ATAC-Seq Data with SnapATAC. *Nat. Commun.* 12, 1337. doi:10.1038/s41467-021-21583-9
- Garneau, J. E., Dupuis, M.-È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., et al. (2010). The CRISPR/Cas Bacterial Immune System Cleaves Bacteriophage and Plasmid DNA. *Nature* 468, 67–71. doi:10.1038/nature09523
- Gaziel-Sovran, A., Segura, M. F., Di Micco, R., Collins, M. K., Hanniford, D., Vega-Saenz de Miera, E., et al. (2011). miR-30b/30d Regulation of GalNAc Transferases Enhances Invasion and Immunosuppression during Metastasis. *Cancer Cell.* 20, 104–118. doi:10.1016/j.ccr.2011.05.027
- Gee, K., Kozlowski, M., and Kumar, A. (2003). Tumor Necrosis Factor- α Induces Functionally Active Hyaluronan-Adhesive CD44 by Activating Sialidase through P38 Mitogen-Activated Protein Kinase in Lipopolysaccharide-Stimulated Human Monocytic Cells. *J. Biol. Chem.* 278, 37275–37287. doi:10.1074/jbc.m302309200
- Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell.* 159, 647–661. doi:10.1016/j.cell.2014.09.029
- Greville, G., Mccann, A., Rudd, P. M., and Saldova, R. (2016). Epigenetic Regulation of Glycosylation and the Impact on Chemo-Resistance in Breast and Ovarian Cancer. *Epigenetics* 11, 845–857. doi:10.1080/15592294.2016.1241932
- Groth, T., Gunawan, R., and Neelamegham, S. (2021). A Systems-Based Framework to Computationally Describe Putative Transcription Factors and Signaling Pathways Regulating Glycan Biosynthesis. *Beilstein J. Org. Chem.* 17, 1712–1724. doi:10.3762/bjoc.17.119
- Guo, H., and Pierce, J. M. (2015). “Transcriptional Regulation of Glycan Expression,” in *Glycoscience: Biology and Medicine*. Editors N. Taniguchi, T. Endo, G. W. Hart, P. H. Seeberger, and C.-H. Wong (Tokyo: Springer Japan), 1173–1180. doi:10.1007/978-4-431-54841-6_79
- Han, K., Jeng, E. E., Hess, G. T., Morgens, D. W., Li, A., and Bassik, M. C. (2017). Synergistic Drug Combinations for Cancer Identified in a CRISPR Screen for Pairwise Genetic Interactions. *Nat. Biotechnol.* 35, 463–474. doi:10.1038/nbt.3834
- Han, K., Pierce, S. E., Li, A., Spees, K., Anderson, G. R., Seoane, J. A., et al. (2020). CRISPR Screens in Cancer Spheroids Identify 3D Growth-specific Vulnerabilities. *Nature* 580, 136–141. doi:10.1038/s41586-020-2099-x
- Haque, A., Engel, J., Teichmann, S. A., and Lönnberg, T. (2017). A Practical Guide to Single-Cell RNA-Sequencing for Biomedical Research and Clinical Applications. *Genome Med.* 9, 75. doi:10.1186/s13073-017-0467-4
- Hashimoto, N., Ito, S., Tsuchida, A., Bhuiyan, R. H., Okajima, T., Yamamoto, A., et al. (2019). The Ceramide Moiety of Disialoganglioside (GD3) Is Essential for GD3 Recognition by the Sialic Acid-Binding Lectin SIGLEC7 on the Cell Surface. *J. Biol. Chem.* 294, 10833–10845. doi:10.1074/jbc.ra118.007083
- Heynen-Genel, S., Pache, L., Chanda, S. K., and Rosen, J. (2012). Functional Genomic and High-Content Screening for Target Discovery and Deconvolution. *Expert Opin. Drug Discov.* 7, 955–968. doi:10.1517/17460441.2012.711311
- Higai, K., Miyazaki, N., Azuma, Y., and Matsumoto, K. (2008). Transcriptional Regulation of the Fucosyltransferase VI Gene in Hepatocellular Carcinoma Cells. *Glycoconj J.* 25, 225–235. doi:10.1007/s10719-008-9114-z
- Holst, S., Deuss, A. J. M., Van Pelt, G. W., Van Vliet, S. J., Garcia-Vallejo, J. J., Koelman, C. A. M., et al. (2016). N-glycosylation Profiling of Colorectal Cancer Cell Lines Reveals Association of Fucosylation with Differentiation and Caudal Type Homeobox 1 (CDX1)/Villin mRNA Expression. *Mol. Cell. Proteomics* 15, 124–140. doi:10.1074/mcp.m115.051235
- Huang, H., Zhou, P., Wei, J., Long, L., Shi, H., Dhungana, Y., et al. (2021a). In Vivo CRISPR Screening Reveals Nutrient Signaling Processes Underpinning CD8+ T Cell Fate Decisions. *Cell.* 184, 1245–1261. doi:10.1016/j.cell.2021.02.021
- Huang, Y.-F., Aoki, K., Akase, S., Ishihara, M., Liu, Y.-S., Yang, G., et al. (2021b). Global Mapping of Glycosylation Pathways in Human-Derived Cells. *Dev. Cell.* 56, 1195–1209. doi:10.1016/j.devcel.2021.02.023
- Huang, Y.-F., Mizumoto, S., and Fujita, M. (2021c). Novel Insight Into Glycosaminoglycan Biosynthesis Based on Gene Expression Profiles. *Front. Cell. Dev. Biol.* 9, 709018. doi:10.3389/fcell.2021.709018
- Indellicato, R., and Trincherà, M. (2021). Epigenetic Regulation of Glycosylation in Cancer and Other Diseases. *Int. J. Mol. Sci.* 22, 2980. doi:10.3390/ijms22062980
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337, 816–821. doi:10.1126/science.1225829
- Jolma, A., Yin, Y., Nitta, K. R., Dave, K., Popov, A., Taipale, M., et al. (2015). DNA-dependent Formation of Transcription Factor Pairs Alters Their Binding Specificity. *Nature* 527, 384–388. doi:10.1038/nature15518
- Kearney, C. J., Vervoort, S. J., Ramsbottom, K. M., Todorovski, I., Lelliott, E. J., Zethoven, M., et al. (2021). SUGAR-seq Enables Simultaneous Detection of Glycans, Epitopes, and the Transcriptome in Single Cells. *Sci. Adv.* 7, eabe3610. doi:10.1126/sciadv.abe3610
- Kurcon, T., Liu, Z., Paradkar, A. V., Vaiana, C. A., Koppolu, S., Agrawal, P., et al. (2015). miRNA Proxy Approach Reveals Hidden Functions of Glycosylation. *Proc. Natl. Acad. Sci. U.S.A.* 112, 7327–7332. doi:10.1073/pnas.1502076112
- Lau, K. S., and Dennis, J. W. (2008). N-Glycans in Cancer Progression. *Glycobiology* 18, 750–760. doi:10.1093/glycob/cwn071

- Lau, K. S., Partridge, E. A., Grigorian, A., Silvescu, C. I., Reinhold, V. N., Demetriou, M., et al. (2007). Complex N-Glycan Number and Degree of Branching Cooperate to Regulate Cell Proliferation and Differentiation. *Cell* 129, 123–134. doi:10.1016/j.cell.2007.01.049
- Leung, K. K., Wilson, G. M., Kirkemo, L. L., Riley, N. M., Coon, J. J., and Wells, J. A. (2020). Broad and Thematic Remodeling of the Surfaceome and Glycoproteome on Isogenic Cells Transformed with Driving Proliferative Oncogenes. *Proc. Natl. Acad. Sci. U.S.A.* 117, 7764–7775. doi:10.1073/pnas.1917947117
- Li, C.-W., Lim, S.-O., Xia, W., Lee, H.-H., Chan, L.-C., Kuo, C.-W., et al. (2016). Glycosylation and Stabilization of Programmed Death Ligand-1 Suppresses T-Cell Activity. *Nat. Commun.* 7, 12632. doi:10.1038/ncomms12632
- Li, Y., Sun, Z., Liu, B., Shan, Y., Zhao, L., and Jia, L. (2017). Tumor-suppressive miR-26a and miR-26b Inhibit Cell Aggressiveness by Regulating FUT4 in Colorectal Cancer. *Cell. Death Dis.* 8, e2892. doi:10.1038/cddis.2017.281
- Lin, C.-Y., Ho, J.-Y., Hsieh, M.-T., Chiang, H.-L., Chuang, J.-M., Whang-Peng, J., et al. (2014). Reciprocal Relationship of Tn/NF- κ B and sTn as an Indicator of the Prognosis of Oral Squamous Cell Carcinoma. *Histopathology* 64, 713–721. doi:10.1111/his.12309
- Ma, M., Guo, D., Tan, Z., Du, J., Guan, F., and Li, X. (2021). Fucosyltransferase 8 Regulation and Breast Cancer Suppression by Transcription Factor Activator Protein 2 γ . *Cancer Sci.* 112, 3190–3204. doi:10.1111/cas.14987
- Macfarlane, L.-A., and R. Murphy, P. (2010). MicroRNA: Biogenesis, Function and Role in Cancer. *Cg* 11, 537–561. doi:10.2174/138920210793175895
- Madunic, K., Zhang, T., Mayboroda, O. A., Holst, S., Stavenhagen, K., Jin, C., et al. (2021). Colorectal Cancer Cell Lines Show Striking Diversity of Their O-Glycome Reflecting the Cellular Differentiation Phenotype. *Cell. Mol. Life Sci.* 78, 337–350. doi:10.1007/s00018-020-03504-z
- Maeder, M. L., Linder, S. J., Cascio, V. M., Fu, Y., Ho, Q. H., and Joung, J. K. (2013). CRISPR RNA-Guided Activation of Endogenous Human Genes. *Nat. Methods* 10, 977–979. doi:10.1038/nmeth.2598
- Martinez-Morales, P., Morán Cruz, I., Roa-De La Cruz, L., Maycotte, P., Reyes Salinas, J. S., Vazquez Zamora, V. J., et al. (2021). Hallmarks of Glycogene Expression and Glycosylation Pathways in Squamous and Adenocarcinoma Cervical Cancer. *PeerJ* 9, e12081. doi:10.7717/peerj.12081
- Minoshima, F., Ozaki, H., Odaka, H., and Tateno, H. (2021). Integrated Analysis of Glycan and RNA in Single Cells. *iScience* 24, 102882. doi:10.1016/j.isci.2021.102882
- Möckl, L., Pedram, K., Roy, A. R., Krishnan, V., Gustavsson, A. K., Dorigo, O., et al. (2019). Quantitative Super-Resolution Microscopy of the Mammalian Glycocalyx. *Dev. Cell.* 50, 57–72 e6. doi:10.1016/j.devcel.2019.04.035
- Moremen, K. W., and Haltiwanger, R. S. (2019). Emerging Structural Insights into Glycosyltransferase-Mediated Synthesis of Glycans. *Nat. Chem. Biol.* 15, 853–864. doi:10.1038/s41589-019-0350-2
- Morgens, D. W., Wainberg, M., Boyle, E. A., Ursu, O., Araya, C. L., Tsui, C. K., et al. (2017). Genome-scale Measurement of Off-Target Activity Using Cas9 Toxicity in High-Throughput Screens. *Nat. Commun.* 8, 15178. doi:10.1038/ncomms15178
- Narimatsu, Y., Joshi, H. J., Nason, R., Van Coillie, J., Karlsson, R., Sun, L., et al. (2019). An Atlas of Human Glycosylation Pathways Enables Display of the Human Glycome by Gene Engineered Cells. *Mol. Cell.* 75, 394–407. doi:10.1016/j.molcel.2019.05.017
- Narimatsu, Y., Joshi, H. J., Yang, Z., Gomes, C., Chen, Y.-H., Lorenzetti, F. C., et al. (2018). A Validated gRNA Library for CRISPR/Cas9 Targeting of the Human Glycosyltransferase Genome. *Glycobiology* 28, 295–305. doi:10.1093/glycob/cwx101
- Nayak, R., and Hasija, Y. (2021). A Hitchhiker's Guide to Single-Cell Transcriptomics and Data Analysis Pipelines. *Genomics* 113, 606–619. doi:10.1016/j.ygeno.2021.01.007
- Neelamegham, S., and Mahal, L. K. (2016). Multi-level Regulation of Cellular Glycosylation: from Genes to Transcript to Enzyme to Structure. *Curr. Opin. Struct. Biol.* 40, 145–152. doi:10.1016/j.sbi.2016.09.013
- Nielsen, M. I., Stegmayr, J., Grant, O. C., Yang, Z., Nilsson, U. J., Boos, I., et al. (2018). Galectin Binding to Cells and Glycoproteins with Genetically Modified Glycosylation Reveals Galectin-Glycan Specificities in a Natural Context. *J. Biol. Chem.* 293, 20249–20262. doi:10.1074/jbc.ra118.004636
- Nielsen, M. I., and Wandall, H. H. (2022). “Dissecting Context-Specific Galectin Binding Using Glycoengineered Cell Libraries,” in *Dissecting Context-specific Galectin/Galectins Binding Using Glycoengineered Cell Libraries/Galectins: Methods and Protocols*. Editors S. R. Stowell, C. M. Arthur, and R. D. Cummings (New York, NY: Springer US), 205–214. doi:10.1007/978-1-0716-2055-7_12
- Odaka, H., Ozaki, H., and Tateno, H. (2022). scGR-Seq: Integrated Analysis of Glycan and RNA in Single Cells. *Star. Protoc.* 3, 101179. doi:10.1016/j.xpro.2022.101179
- Oliveira, T., Thaysen-Andersen, M., Packer, N. H., and Kolarich, D. (2021). The Hitchhiker's Guide to Glycoproteomics. *Biochem. Soc. Trans.* 49, 1643–1662. doi:10.1042/bst20200879
- Pan, Y., Tsai, C.-J., Ma, B., and Nussinov, R. (2010). Mechanisms of Transcription Factor Selectivity. *Trends Genet.* 26, 75–83. doi:10.1016/j.tig.2009.12.003
- Park, P. J. (2009). ChIP-seq: Advantages and Challenges of a Maturing Technology. *Nat. Rev. Genet.* 10, 669–680. doi:10.1038/nrg2641
- Pecori, F., Yokota, I., Hanamatsu, H., Miura, T., Ogura, C., Ota, H., et al. (2021). A Defined Glycosylation Regulatory Network Modulates Total Glycome Dynamics during Pluripotency State Transition. *Sci. Rep.* 11, 1276. doi:10.1038/s41598-020-79666-4
- Pietrobono, S., Anichini, G., Sala, C., Manetti, F., Almada, L. L., Pepe, S., et al. (2020). ST3GAL1 Is a Target of the SOX2-GLI1 Transcriptional Complex and Promotes Melanoma Metastasis through AXL. *Nat. Commun.* 11, 5865. doi:10.1038/s41467-020-19575-2
- Sackstein, R. (2011). The Biology of CD44 and HCELL in Hematopoiesis: the 'step 2-bypass Pathway' and Other Emerging Perspectives. *Curr. Opin. Hematol.* 18, 239–248. doi:10.1097/moh.0b013e3283476140
- Saeland, E., Van Vliet, S. J., Bäckström, M., Van Den Berg, V. C. M., Geijtenbeek, T. B. H., Meijer, G. A., et al. (2007). The C-type Lectin MGL Expressed by Dendritic Cells Detects Glycan Changes on MUC1 in Colon Carcinoma. *Cancer Immunol. Immunother.* 56, 1225–1236. doi:10.1007/s00262-006-0274-z
- Sako, D., Comess, K. M., Barone, K. M., Camphausen, R. T., Cumming, D. A., and Shaw, G. D. (1995). A Sulfated Peptide Segment at the Amino Terminus of PSGL-1 Is Critical for P-Selectin Binding. *Cell* 83, 323–331. doi:10.1016/0092-8674(95)90173-6
- Sakuma, K., Aoki, M., and Kannagi, R. (2012). Transcription Factors C-Myc and CDX2 Mediate E-Selectin Ligand Expression in Colon Cancer Cells Undergoing EGF/bFGF-induced Epithelial-Mesenchymal Transition. *Proc. Natl. Acad. Sci. U.S.A.* 109, 7776–7781. doi:10.1073/pnas.1111135109
- Saliminejad, K., Khorram Khorshid, H. R., Soleymani Fard, S., and Ghaffari, S. H. (2019). An Overview of microRNAs: Biology, Functions, Therapeutics, and Analysis Methods. *J. Cell. Physiology* 234, 5451–5465. doi:10.1002/jcp.27486
- Sanjana, N. E. (2017). Genome-scale CRISPR Pooled Screens. *Anal. Biochem.* 532, 95–99. doi:10.1016/j.ab.2016.05.014
- Seales, E. C., Jurado, G. A., Singhal, A., and Bellis, S. L. (2003). Ras Oncogene Directs Expression of a Differentially Sialylated, Functionally Altered β 1 Integrin. *Oncogene* 22, 7137–7145. doi:10.1038/sj.onc.1206834
- Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., et al. (2014). Genome-scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* 343, 84–87. doi:10.1126/science.1247005
- Shifrut, E., Carnevale, J., Tobin, V., Roth, T. L., Woo, J. M., Bui, C. T., et al. (2018). Genome-wide CRISPR Screens in Primary Human T Cells Reveal Key Regulators of Immune Function. *Cell* 175, 1958–1971. doi:10.1016/j.cell.2018.10.024
- Smith, B. a. H., and Bertozzi, C. R. (2021). The Clinical Impact of Glycobiology: Targeting Selectins, Siglecs and Mammalian Glycans. *Nat. Rev. Drug Discov.* 20(3), 217–243. doi:10.1038/s41573-020-00093-1
- Smith, R. D., and Lupashin, V. V. (2008). Role of the Conserved Oligomeric Golgi (COG) Complex in Protein Glycosylation. *Carbohydr. Res.* 343, 2024–2031. doi:10.1016/j.carres.2008.01.034
- Stampfel, G., Kazmar, T., Frank, O., Wienerroither, S., Reiter, F., and Stark, A. (2015). Transcriptional Regulators Form Diverse Groups with Context-dependent Regulatory Functions. *Nature* 528, 147–151. doi:10.1038/nature15545
- Stark, R., Grzelak, M., and Hadfield, J. (2019). RNA Sequencing: the Teenage Years. *Nat. Rev. Genet.* 20, 631–656. doi:10.1038/s41576-019-0150-2
- Stolfá, G., Mondal, N., Zhu, Y., Yu, X., Buffone, A., and Neelamegham, S. (2016). Using CRISPR-Cas9 to Quantify the Contributions of O-Glycans, N-Glycans and Glycosphingolipids to Human Leukocyte-Endothelium Adhesion. *Sci. Rep.* 6, 30392. doi:10.1038/srep30392

- Taniguchi, A., Hasegawa, Y., Higai, K., and Matsumoto, K. (2000). Transcriptional Regulation of Human α -galactoside 2,6-sialyltransferase (hST6Gal I) Gene during Differentiation of the HL-60 Cell Line. *Glycobiology* 10, 623–628. doi:10.1093/glycob/10.6.623
- Taniguchi, A., Saito, K., Kubota, T., and Matsumoto, K. (2003). Characterization of the Promoter Region of the Human Gal β 1,3(4)GlcNAc α 2,3-sialyltransferase III (hST3Gal III) Gene. *Biochimica Biophysica Acta (BBA) - Gene Struct. Expr.* 1626, 92–96. doi:10.1016/s0167-4781(03)00021-6
- Taylor-Papadimitriou, J., Burchell, J., Miles, D. W., and Dalziel, M. (1999). MUC1 and Cancer. *Biochimica Biophysica Acta (BBA) - Mol. Basis Dis.* 1455, 301–313. doi:10.1016/s0925-4439(99)00055-1
- Thambu, K., and Balagurunathan, K. (2022). anexVis: A Transcriptome Tool to Visualize Organ/Tissue-Specific Glycosaminoglycan Biosynthetic and Catabolic Pathways in Human Health and Diseases. *Methods Mol. Biol.* 2303, 389–401. doi:10.1007/978-1-0716-1398-6_31
- Thiemann, S., and Baum, L. G. (2016). Galectins and Immune Responses—Just How Do They Do Those Things They Do? *Annu. Rev. Immunol.* 34, 243–264. doi:10.1146/annurev-immunol-041015-055402
- Thu, C. T., Chung, J. Y., Dhawan, D., Vaiana, C. A., and Mahal, L. K. (2021). High-Throughput miRFluR Platform Identifies miRNA Regulating B3GLCT That Predict Peters' Plus Syndrome Phenotype, Supporting the miRNA Proxy Hypothesis. *ACS Chem. Biol.* 16, 1900–1907. doi:10.1021/acscchembio.1c00247
- Thu, C. T., and Mahal, L. K. (2020). Sweet Control: MicroRNA Regulation of the Glycome. *Biochemistry* 59, 3098–3110. doi:10.1021/acs.biochem.9b00784
- Tian, S., Muneeeruddin, K., Choi, M. Y., Tao, L., Bhuiyan, R. H., Ohmi, Y., et al. (2018). Genome-wide CRISPR Screens for Shiga Toxins and Ricin Reveal Golgi Proteins Critical for Glycosylation. *PLoS Biol.* 16, e2006951. doi:10.1371/journal.pbio.2006951
- Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Proteomics. Tissue-Based Map of the Human Proteome. *Science* 347, 1260419.
- A. Varki, D. R. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, G. W. Hart, et al. (2015). *Essentials of Glycobiology*. 4th edition ((NY): Cold Spring Harbor Laboratory Press).
- Wang, D., Zhang, T., Madunić, K., De Waard, A. A., Blöchl, C., Mayboroda, O. A., et al. (2022). Glycosphingolipid-Glycan Signatures of Acute Myeloid Leukemia Cell Lines Reflect Hematopoietic Differentiation. *J. Proteome Res.* 21, 1029–1040. doi:10.1021/acs.jpoteome.1c00911
- Wang, T., Wei, J. J., Sabatini, D. M., and Lander, E. S. (2014). Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science* 343, 80–84. doi:10.1126/science.1246981
- Weiss, R. J., Spahn, P. N., Chiang, A. W. T., Liu, Q., Li, J., Hamill, K. M., et al. (2021). Genome-wide Screens Uncover KDM2B as a Modifier of Protein Binding to Heparan Sulfate. *Nat. Chem. Biol.* 17, 684–692. doi:10.1038/s41589-021-00776-9
- Weiss, R. J., Spahn, P. N., Toledo, A. G., Chiang, A. W. T., Kellman, B. P., Li, J., et al. (2020). ZNF263 Is a Transcriptional Regulator of Heparin and Heparan Sulfate Biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 117, 9311–9317. doi:10.1073/pnas.1920880117
- Wisnovsky, S., Möckl, L., Malaker, S. A., Pedram, K., Hess, G. T., Riley, N. M., et al. (2021). *Genome-wide CRISPR Screens Reveal a Specific Ligand for the Glycan-Binding Immune Checkpoint Receptor Siglec-7*. USA: Proceedings of the National Academy of Sciences, 118. doi:10.1073/pnas.2015024118
- Wong, A. S. L., Choi, G. C. G., Cui, C. H., Pregernig, G., Milani, P., Adam, M., et al. (2016). Multiplexed Barcoded CRISPR-Cas9 Screening Enabled by CombiGEM. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2544–2549. doi:10.1073/pnas.1517883113
- Wu, D. Y., Bittencourt, D., Stallcup, M. R., and Siegmund, K. D. (2015). Identifying Differential Transcription Factor Binding in CHIP-Seq. *Front. Genet.* 6, 169. doi:10.3389/fgene.2015.00169
- Wu, H., Chen, J., Li, D., Liu, X., Li, L., and Wang, K. (2017). MicroRNA-30e Functions as a Tumor Suppressor in Cervical Carcinoma Cells through Targeting GALNT7. *Transl. Oncol.* 10, 876–885. doi:10.1016/j.tranon.2017.08.006
- Wu, Z. Y., He, Y. Q., Wang, T. M., Yang, D. W., Li, D. H., Deng, C. M., et al. (2021). Glycogenes in Oncofetal Chondroitin Sulfate Biosynthesis Are Differently Expressed and Correlated With Immune Response in Placenta and Colorectal Cancer. *Front. Cell. Dev. Biol.* 9, 763875. doi:10.3389/fcell.2021.763875
- Xu, Q., Akama, R., Isaji, T., Lu, Y., Hashimoto, H., Kariya, Y., et al. (2011). Wnt/ β -Catenin Signaling Down-Regulates N-Acetylglucosaminyltransferase III Expression. *J. Biol. Chem.* 286, 4310–4318. doi:10.1074/jbc.m110.182576
- Xue, M., Zhuo, Y., and Shan, B. (2017). MicroRNAs, Long Noncoding RNAs, and Their Functions in Human Disease. *Methods Mol. Biol.* 1617, 1–25. doi:10.1007/978-1-4939-7046-9_1
- Yamaji, T., Sekizuka, T., Tachida, Y., Sakuma, C., Morimoto, K., Kuroda, M., et al. (2019). A CRISPR Screen Identifies LAPTMA and TM9SF Proteins as Glycolipid-Regulating Factors. *iScience* 11, 409–424. doi:10.1016/j.isci.2018.12.039
- Yan, F., Powell, D. R., Curtis, D. J., and Wong, N. C. (2020). From Reads to Insight: a Hitchhiker's Guide to ATAC-Seq Data Analysis. *Genome Biol.* 21, 22. doi:10.1186/s13059-020-1929-3
- Yang, X., and Qian, K. (2017). Protein O-GlcNAcylation: Emerging Mechanisms and Functions. *Nat. Rev. Mol. Cell. Biol.* 18, 452–465. doi:10.1038/nrm.2017.22
- Yang, Z., Wang, S., Halim, A., Schulz, M. A., Frodin, M., Rahman, S. H., et al. (2015). Engineered CHO Cells for Production of Diverse, Homogeneous Glycoproteins. *Nat. Biotechnol.* 33, 842–844. doi:10.1038/nbt.3280
- Yu, R. K., and Bieberich, E. (2001). Regulation of Glycosyltransferases in Ganglioside Biosynthesis by Phosphorylation and Dephosphorylation. *Mol. Cell. Endocrinol.* 177, 19–24. doi:10.1016/s0303-7207(01)00457-9
- Zeng, R., Mohamed, A., Khanna, K. K., and Hill, M. M. (2021). Differential Regulation of Lacto-/Neolacto- Glycosphingolipid Biosynthesis Pathway Reveals Transcription Factors as Potential Candidates in Triple-Negative Breast Cancer. *Cancers (Basel)* 13, 3330. doi:10.3390/cancers13133330
- Zhang, W., Revers, L., Pierce, M., and Schachter, H. (2000). Regulation of Expression of the Human β -1,2-N-acetylglucosaminyltransferase II Gene (MGAT2) by Ets Transcription Factors. *Biochem. J.* 347, 511–518. doi:10.1042/bj3470511
- Zhu, Y., Groth, T., Kelkar, A., Zhou, Y., and Neelamegham, S. (2021). A GlycoGene CRISPR-Cas9 Lentiviral Library to Study Lectin Binding and Human Glycan Biosynthesis Pathways. *Glycobiology* 31, 173–180. doi:10.1093/glycob/cwaa074

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a past collaboration with the author SW.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Stewart and Wisnovsky. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.