A systematic framework to derive \( N \)-glycan biosynthesis process and the automated construction of glycosylation networks

Wenpin Hou1*, Yushan Qiu3, Nobuyuki Hashimoto2, Wai-Ki Ching1 and Kiyoko F. Aoki-Kinoshita2

From 12th Annual Biotechnology and Bioinformatics Symposium (BIOT-2015)
Provo, UT, USA.10-11 December 2015

Abstract

Background: Abnormalities in glycan biosynthesis have been conclusively related to various diseases, whereas the complexity of the glycosylation process has impeded the quantitative analysis of biochemical experimental data for the identification of glycoforms contributing to disease. To overcome this limitation, the automatic construction of glycosylation reaction networks in silico is a critical step.

Results: In this paper, a framework K2014 is developed to automatically construct \( N \)-glycosylation networks in MATLAB with the involvement of the 27 most-known enzyme reaction rules of 22 enzymes, as an extension of previous model KB2005. A toolbox named Glycosylation Network Analysis Toolbox (GNAT) is applied to define network properties systematically, including linkages, stereochemical specificity and reaction conditions of enzymes. Our network shows a strong ability to predict a wider range of glycans produced by the enzymes encountered in the Golgi Apparatus in human cell expression systems.

Conclusions: Our results demonstrate a better understanding of the underlying glycosylation process and the potential of systems glycobiology tools for analyzing conventional biochemical or mass spectrometry-based experimental data quantitatively in a more realistic and practical way.

Keywords: \( N \)-glycan, Glycosylation reaction networks construction, Glycosylation enzyme activity, Mass spectrum, Glycobiology

Background

Glycosylation is an important and highly complex post-translational modification that generates an extensive functional capability from a limited set of genes and encompasses the biosynthesis of sugar moieties in the endoplasmic reticulum (ER) and Golgi apparatus [1–3]. Glycans are highly variable and structurally diverse compounds consisting of a large number of monosaccharides, including mannose, fucose, and galactose, linked through an enzymatic process called glycosylation [4]. Unlike protein structures, glycan structures are neither directly encoded in the genome nor arranged in a simple linear chain [5]. Instead, the structure of secreted and membrane-bound glycans is determined during their assembly in the endoplasmic reticulum and the Golgi apparatus by a controlled sequence of glycosyltransferase and glycosidase processing reactions [6].

One of the major types of glycans attached to asparagine residues of proteins, \( N \)-linked glycans, is determined by a manageable number of enzymes that catalyze monosaccharide attachment. \( N \)-linked glycosylation occurs cotranslationally in endoplasmic reticulum compartments. Glycoproteins migrate into the Golgi apparatus once the
protein finishes folding and some residues in the glycan trim successfully [7]. Many of these enzymes can generally accept several N-linked glycans as substrates, therefore generating a large number of glycan products and their glycosylation pathways [7]. Processing involves the removal of mannnose groups, which is facilitated by mannosidases, and the addition of diverse monosaccharides driven by specific glycosyltransferases to the substrate glycan. Therefore, the glycosylation pathways of N-linked glycans comprise consecutive enzymatic steps, which are determined by the glycan structures produced by the previous enzyme, to produce a new glycan structure as the substrate of the next glycosylation reaction [7].

From research conducted in the past decades, it is clear that the glycosylation of diseased cells and healthy cells often results in different glycan changes that contribute to pathological progression, leading to the possibility that disease-specific glycan structures exist [8–12]. This has potential medical applications; for example, the potential to distinguish benign forms of prostate cancer from highly malignant cancer based on the changes in enzymes’ activities and intracellular processing events [13, 14]. Effective engineering of glycosylation pathways can potentially lead to an improved therapeutic performance of glycoprotein products. Considerable progress has been made in Prostate-Specific Antigen (PSA) research; and analytic approaches in analyzing large data sets have also been utilized to expand analyses from PSA to numerous cancers and diseases known to have abnormalities in glycosylation. However, more research is still needed in acquiring and interpreting datasets to completely characterize glycosylation, including enzymatic profiles involved in glycosylation and the large quantities of glycans produced by enzymes.

Fortunately, a wealth of data is available from glycosylation-specific databases such as the Consortium for Functional Glycomics (CFG) website [15] and GlycomeDB [16]. Also, Liquid Chromatography (LC) and Mass Spectrometric (MS) techniques have been emerging as enabling and important techniques in glycomics. A number of LC/MS methods have been incorporated into glycomics workflows for permethylated and amidated glycans reduction [17, 18]. Statistical methods have been proposed to predict glycan structures from gene expression data [19–21]. However, these traditional and qualitative methods in biochemistry or cell biology research do not provide a detailed understanding of the complex glycosylation mechanism quantitatively.

Systems biology-based mathematical models have been developed to overcome this limitation [6, 22–26]. In this regard, the construction of glycosylation reaction networks in silico is an important step that can enable the quantitative analysis of biochemical experimental data. Whereas several studies have been done to construct glycosylation reaction networks automatically on computers, they are limited by the lack of a systematic definition of the linkage, stereocchemical specificity and reaction conditions of enzymes that are involved in the reactions.

Liu and Neelamegham [27] made a significant contribution by designing an open-source MATLAB-based toolbox, Glycosylation Network Analysis Toolbox (GNAT), for studies of systems glycobiology. This toolbox enables a streamlined machine-readable definition for the glycosylation enzyme class and the construction as well as adjustment of glycosylation reaction networks [27].

This paper extends Liu and Neelamegham’s work [27] to predict a wider range of glycans produced by enzymes encountered in human cell expression systems. In addition, our model can be applied to a larger range of experimental conditions that might be encountered in a cell culture environment. We expand the scope by inclusion of additional enzyme classes involved in gene expression data. We extend the framework of KB2005 [6] through involving 22 enzymes (27 enzyme reaction rules) in our network.

To the best of our knowledge, these 22 enzymes (27 enzyme reaction rules) are all enzymes associated with N-glycan that exist in Golgi compartments. Networks constructed can be used to relate the observed mass spectrometric measurements to the underlying gene expression weights. This relationship can be used to quantitatively understand how changes in enzymes’ activities affect the profile of glycan structures produced in the biosynthesis process.

**Methods**

We apply the Glycosylation Network Analysis Toolbox (GNAT) as the platform to facilitate the modeling. Generated by Liu and Neelamegham [27], GNAT is a MATLAB-based toolbox for the automated construction of glycosylation reaction networks.

**Enzyme definitions using biological experimental data**

To analyze biological experimental data in the glycomics field quantitatively, the construction of glycosylation reaction networks, which can describe the glycan biosynthesis process in silico, is rather critical. Some efforts have been made to build systems-level cellular glycosylation reaction networks on O-linked glycan [26] or N-linked glycan [23] formation. However, this work has always been hindered by the lack of a complete system for the specificity of detailed glycosylation rules. In this paper, we address this problem by defining enzymes in a machine-friendly way using a MATLAB-based toolbox called Glycosylation Network Analysis Toolbox (GNAT).

In GNAT, we can define generic enzymes using the Enz class. Since the properties of transferases and hydrolases...
are very different, \textit{TfEnz} and \textit{HlEnz} are used as subclasses of the \textit{Enz} class to define them, respectively, as follows:

Transferase reaction:

\[ \text{donor} + \text{acceptor} \rightleftharpoons \text{donorProd} + \text{acceptorProd} \]

Hydrolase reaction:

\[ \text{substrate} + \text{H}_2\text{O} \rightleftharpoons \text{prod}_{\text{oh}} + \text{prod}_{\text{h}} \]

where \textit{donor}, \textit{donorProd}, \textit{acceptor} and \textit{acceptorProd} are variables used to specify properties in transferase (\textit{TfEnz}) reactions and \textit{prod}_{\text{oh}} and \textit{prod}_{\text{h}} are for hydrolase (\textit{HlEnz}) reactions as named.

In addition, variables such as \textit{resfuncgroup}, \textit{linkFG}, \textit{resAtt}_{2\text{FG}} and \textit{linkAtt}_{2\text{FG}} are used to specify functional groups and linkage specificity for enzymes. More specifically, \textit{resfuncgroup} is the residue of functional group transferred, \textit{resAtt}_{2\text{FG}} is the residue attaching to functional group, \textit{linkFG} refers to the linkage between attaching residue and functional group and \textit{linkresAtt}_{2\text{FG}} stands for the linkage between the attaching residue and its next neighboring residue.

Most importantly, eight substrate specificity functions are available in the glycosyltransferase and glycosidase classes: \textit{substMinStruct}, \textit{substMaxStruct}, \textit{substNABranch}, \textit{substNAClass}, \textit{targetBranch}, \textit{targetNABranch}, and \textit{iSTerminalTarget}. They define the properties of substrates or products in reactions upon which a specific enzyme acts. Only if all conditions defined by these functions are satisfied can the corresponding enzyme act on the substrate; that is, the reaction will happen. In this work, firstly, we collect 27 enzyme reaction rules of 22 enzymes from databases, such as GlycomeDB [16], KEGG Glycan, IUBMB (International Union of Biochemistry and Molecular Biology), GlycoGene Database (ggdb), the Consortium for Functional Glycomics (CFG) website [25], and previous work by Sandra V. Bennun et al. [24] (See Table 1). Secondly, we define enzymes by specifying their properties in reactions.

| Index | Enzyme | EC No. | Substrate | Product | Constraint |
|-------|--------|--------|-----------|---------|------------|
| 1     | ManI   | [3.2.1.113] | (Ma2Ma) | (Ma) | \(\sim^*2\text{Ma}...(\ldots\text{Ma6})\) |
| 2     | ManI   | [3.2.1.113] | (Ma3(Ma2Ma3(Ma6)Ma6)Ma6) | (Ma3(Ma3(Ma6)Ma6)) | \(\sim^*2\text{Ma}...(\ldots\text{Ma6})\) |
| 3     | ManII  | [3.2.1.114] | (Ma3(Ma6)Ma6) | (Ma6Ma6) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 4     | ManII  | [3.2.1.114] | (Ma6Ma6) | (Ma6) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 5     | FUT8   | [2.4.1.68] | Gnb4GN | Gnb4(Fa6)GN | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 6     | MGAT1  | [2.4.1.101] | (Ma3(Ma3(Ma6)Ma6)Mb4) | (Gnb2Ma3(Ma3(Ma6)Ma6)Mb4) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 7     | MGAT2  | [2.4.1.103] | (Gnb2(Ma3(Ma6)Mb4)Mb4) | (Gnb2(Ma3(Ma6)Mb4)Mb4) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 8     | MGAT3  | [2.4.1.144] | Gnb2(Ma3) | Gnb2(Ma3(Ma6)Mb4) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 9     | MGAT4  | [2.4.1.145] | Gnb2Ma3 | Gnb2(Ma3(Ma6)Mb4) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 10    | MGAT5  | [2.4.1.145] | Gnb2Ma6 | Gnb2(Ma3(Ma6)Mb4) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 11    | iGnT   | [2.4.1.149] | (Ab4GN) | Gnb3Ab4GN | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 12    | b4GalT | [2.4.1.38] | (GN) | (Ab4GN) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 13    | a3SiaT | [2.4.99.6] | (Ab4GN) | (NNa3Ab4GN) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 14    | IGNT   | [2.4.1.150] | (Ab4GNb3Ab) | (Ab4GNb3(Gnb6)Ab) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 15    | a6SiaT | [2.4.99.1] | (Ab4GN) | (NNa6Ab4GN) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 16    | b3GalT | [2.4.1.62] | (GN) | (Ab4GN) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 17    | FUT3   | [2.4.1.65] | Ab3Gnb | Ab3(Fa4)Gnb | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 18    | FUT3   | [2.4.1.65] | (Fa3(\ldots\text{Ab4})Gnb | (Fa3(\ldots\text{Ab4})Gnb | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 19    | FUT1   | [2.4.1.69] | Ab3Gnb | (Fa2Ab3Gnb | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 20    | FUT1   | [2.4.1.69] | Ab4Gnb | (Fa2Ab4Gnb | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 21    | a3FucT | [2.4.1.152] | (\ldots\text{Ab4)Gnb} | (\ldots\text{Ab4)Gnb} | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 22    | Ga1NAcT-A | [2.4.1.40] | (Fa2Ab) | (Fa2(\ldots\text{NnA3Ab3})Ab | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 23    | Ga1T-B | [2.4.1.37] | (Fa2Ab) | (Fa2(\ldots\text{NnA3Ab3})Ab | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 24    | b3GALT6 | [2.4.1.34] | Ab4GN | Ab3Ab4GN | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 25    | b3GALT6 | [2.4.1.34] | Ab4A | Ab3Ab4A | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 26    | c1GALT1 | [2.4.1.122] | AN | Ab3AN | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
| 27    | st3galI | [2.4.1.99.4] | (Ab3GN) | (NNa3Ab4GN) | \(\sim^*\text{Gnb6}(\ldots\text{Ma3}\sim\text{Gnb6})\) |
biosynthesis reactions utilizing the functions above (See Figs. 1 and 2).

Automated construction of glycosylation reaction networks

We integrate glycosyltransferase and glycosidase enzymatic data from databases into across-the-board enzyme classes: GHEnz and GTEnz. Automated construction of glycosylation reaction networks is enabled by the definition of the glycosidase (GHEnz) and glycosyltransferase (GTEnz) classes. If the substrate and the enzyme are determined, the product of glycosidase and glycosyltransferase reactions can be automatically generated by function product inference. Here we choose the function forward network inference, which can consider the products generated as the substrates of the next reaction so that reactions will happen in a sequential manner if the reaction conditions conform to the enzyme reaction rules. This is repeated until no additional new products are generated by the set of enzymes specified.

In this work, we define all enzymes in file createEnzDb.m (See “Availability of data and materials” section for the MATLAB codes) and then load it in MATLAB. Next, we place all enzymes in a CellArrayList object called enzArray. Starting with the glycan CellArrayList objects 9-mannose and 8-mannose as substrates, all possible reaction pathways and products are determined by function inferGlyForwPath. This process will continue indefinitely if there are always reactions and products. In this process, duplicate reactions and glycan product species are removed while being consolidated into the Pathway object. In the network generated by our model (called K2014), nodes represent the glycan species and the biochemical reactions between two glycans are denoted as edges. Setting the value of the variable count, we can easily limit the size of the network generated as needed.

Results

We applied the 27 current reaction rules of biosynthesis reactions catalyzed by 22 enzymes to define all these enzymes. These definitions were then used to infer an N-linked glycosylation reaction network automatically in MATLAB. The network generated shows all enzymes catalyzing each reaction in different colors. All reactions are denoted as arrows, which are determined by the enzyme that catalyzed them. Glycan structures involving linkages and monosaccharides that constitute the glycan are also visual. We call this network K2014, as shown in Fig. 3. The comparison among previous models UB1997 [22], KB2005 [23], and our model K2014 is shown in Table 2 (only in the first seven rounds because from the eighth round, our network becomes too big to be presented here). Since pathways can be generated endlessly if reactions are still possible to happen according to the reaction
rules, we can set round in the for loop and get a network of the size we need. At the same time, we can also operate the network generated by removing glycans that do not exist in the Golgi apparatus using the pathway class.

**Discussion**

As can be seen from Table 2, our model generates a much more complex glycosylation reaction network. Here we consider one horizontal step as one round. Take round 7 as an example. In the model UB1997 [22] where only eight enzymes are considered, there are only four glycan structures in the network. Three monosaccharides are operated in this network: mannose, N-acetylglucosamine, and galactose. While in 2005, 11 enzymes are involved in the model KB2005 and 14 structures can be shown in round 7. In this model, the network shows the glycosylation process containing five monosaccharides being added to or cut from the visual glycan structures. However, in reality, the structures of glycan can be much more complicated, and more monosaccharides can be added to a glycan during the glycosylation process regarding enzyme specificity.

Concerning the fact that more enzyme reaction rules have been updated by biologists, we take the most up-to-date reaction rules into concern and involve them all in the construction of a glycosylation reaction network. Thus, the model K2014 is built. There are many more glycan structures than we have space to show here. As can be seen from Fig. 3, 31 glycan structures have been generated by the catalysis of these 22 enzymes (totally 27 enzyme reaction rules). Note that it provides a more complete understanding of the glycosylation process. The enzymatic rules collected from databases and literature show that these enzymes exist in cells and are involved in the glycosylation process. Concerning that our network includes more enzymes than the original GNAT and the model KB2005 [23], it is closer to reality. Thus, our model provides a result that better represents the underlying glycosylation process.

**Conclusions**

In this work, we extend the model KB2005 [23] with application of the toolbox GNAT to construct a relatively larger N-linked glycosylation network. While GNAT only involves the presence of 9 enzymes, we include 22 enzymes (totally 27 enzyme reaction rules) whose information is collected from the literature, glycosylation-specific databases such as GlycomeDB [16], the Consortium for Functional Glycomics (CFG) website [25], KEGG Glycan and ggd. Our framework K2014 can automatically construct a relatively larger glycosylation reaction network using 27 streamlined enzyme reaction rules. Accordingly, this framework advances the study of the in silico cell processes and potentially has significant benefits.
Fig. 3 Reaction network generated by model K2014. This is the glycosylation reaction network automatically generated by our model K2014 which takes M9 as the substrate and involves all 27 enzyme reaction rules (only the first seven rounds are shown here).
With this glycosylation network and related algorithms, a variety of network analysis strategies can be implemented to analyze the components of the overall glycosylation network. With our framework K2014, it is possible to analyze conventional biochemical or mass spectrometry-based experimental data quantitatively in a more realistic and practical way. Examples include, but are not limited to, a dynamic mathematical model for monoclonal antibodies (mAbs) glycosylation to estimate unknown enzymes and transport proteins (TPs) concentration profile parameters [47], a framework to quantitatively understand how changes in enzymes activities affect the profile of glycan structures produced in the biosynthesis process [24], and the simulations on N-glycan processing in accessing whether a homogeneous glycan profile can be created through metabolic engineering [48]. Future research can be conducted on the above-related issues using both conventional biochemical resources and high-throughput MS experimental data.

Acknowledgements
The authors would like to thank the referees and the editors for their helpful comments and suggestions. This research work was supported by Research Grants Council of Hong Kong under Grant Number 17301214, HKU CERG Grants, Hung Hing Ying Physical Research Grant, the Natural Science Foundation of China under Grant No. 11271144. Publication of this article was funded by Soka University (S2021200).

Declarations
This article has been published as part of BMC Bioinformatics Volume 17 Supplement 7, 2016: Selected articles from the 12th Annual Biotechnology and Bioinformatics Symposium: Bioinformatics. The full contents of the supplement are available online at https://bmcbioinformatics.biomedcentral.com/articles/supplements/volume-17-supplement-7.

Availability of data and materials
Here follows the MATLAB codes for the construction of glycosylation networks. File 1 – createEnzDb.m. This file defines all enzymes. Please refer to this link to download it: http://hkumath.hku.hk/~wkc/Hou/createEnzDb.m.

File 2 – glyenzDB.mat. This file should be loaded before running codes. Please refer to this link to download it: http://hkumath.hku.hk/~wkc/Hou/glyenzDB.mat.zip.

File 3 – runme. This file contains the codes we should run on MATLAB. Please refer to this link to download it: http://hkumath.hku.hk/~wkc/Hou/runme.m.

File 4 – readme.txt. Please refer to the link to download it: http://hkumath.hku.hk/~wkc/Hou/readme.txt.

Authors’ contributions
KK designed the research and gave basic ideas for the method. WH developed and implemented the method, performed theoretical analyses and computational experiments, and drafted and revised the manuscript. NH assisted WH in implementing the method. WKC participated in discussions and gave valuable comments to improve the draft. YQ assisted in data presentation. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Published: 25 July 2016

References
1. Pattison RJ, Amtmann A. N-glycan production in the endoplasmic reticulum of plants. Trends Plant Sci. 2009;14(2):92–9.
2. Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. Cell. 2006;126(5):855–67.
3. Brooks SA. Protein glycosylation in diverse cell systems: implications for modification and analysis of recombinant proteins. Expert Rev Proteomics. 2006;3(3):345–59.
4. IUPAC Gold Book- Glycans. http://goldbook.iupac.org/G02645.html.
5. von der Lieth CW, Bohne-Lang A, Lohmann KK, Frank M. Bioinformatics for glycomics: status, methods, requirements and perspectives. Brief Bioinform. 2004;5(2):164–78.
6. Krambeck FJ, Bunnun SV, Nanag S, Choi S, Yaremka J, Benbaug MAJ. A mathematical model to derive N-glycan structures and cellular enzymes activities from mass spectrometric data. Glycobiology. 2009;19(11):1163–75.
7. Kim PJ, Lee DY, Jeong H. Centralized modularity of N-linked glycosylation pathways in mammalian cells. PLoS One. 2009;4(10):7317.
8. Sel S. Cancer-associated carbohydrates identified by monoclonal antibodies. Hum Pathol. 1990;21(10):1003–19.
9. Hakomori S. Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines. Adv Exp Med Biol. 2001;491:369–407.
10. Fuster MM, Esko JD. The sweet and sour of cancer: glycans as novel therapeutic targets. Nat Rev Cancer. 2005;5(7):526–42.
11. Tong L, Baskaran G, Jones MB, Rhee J, Yaremka JK. Glycosylation changes as markers for the diagnosis and treatment of human disease. Biotechnol Genet Eng Rev. 2003;20(1):199–244.
12. Dennis JW, Granovsky M, Warren CE. Glycoprotein glycosylation and cancer progression. Biochrom Biophys Acta. 1999;1473(1):21–34.
13. Tajiri M, Ohyama C, Wada Y. Oligosaccharide profiles of the prostate specific antigen in free and complexed forms from the prostate cancer patient serum and in seminal plasma: A glycopeptide approach. Glycobiology. 2008;18(1):2–8.
14. Meany DL, Zhang Z, Sokoll LJ, Zhang H, Chan DW. Glycopeptides for prostate cancer detection: Changes in serum PSA glycosylation patterns. J Proteome Res. 2009;8(2):613–9.
15. Raman R, Venkataraman M, Ramakrishnan S, Lang W, Raguram S, Sasisekharan R. Advancing glycomics: Implementation strategies at the consortium for functional glycomics. Glycobiology. 2006;16(5):82–90.
16. Ranzinger R, Herget S, Lieth CWVD. Frank M. GlycomeDBA unified database for carbohydrate structures. Nucleic Acids Res. 2011;39 (Database issue):373–6.
17. Zaia J. Mass spectrometry and the emerging field of glycomics. Chem Biol. 2008;15(9):881–92.
18. Novotny MV, Alley WR. Recent trends in analytical and structural glycomics. Curr Opin Chem Biol. 2013;17(5):832–40.
19. Sug A, Yamashii N, Hashimoto K, Goto S, Kaneliths M. An improved scoring scheme for predicting gycan structures from gene expression data. Genome Inform. 2007;18:237–46.
20. Hashimoto K, Goto S, Kawano S, Aoki-Kinoshita KF, Ueda N, Hamajima M, Kawasaki T, Kanehisa M. KEGG as a glycome informatics resource. Glycobiology. 2006;16(5):63–70.
21. Kawano S, Hashimoto K, Miyama T, Goto S, Kanehisa M. Prediction of glycan structures from gene expression data based on glycosyltransferase reactions. Bioinformatics. 2005;21(21):3976–82.

22. Umama P, Bailey JE. A mathematical model of N-linked glycoform biosynthesis. Biotechnol Bioeng. 1997;55(6):890–908.

23. Krambeck FJ, Betenbaugh MJ. A mathematical model of N-linked glycosylation. Biotechnol Bioeng. 2005;92(6):711–28.

24. Bennun SV, Yarema KJ, Betenbaugh MJ, Krambeck FJ. Integration of the transcriptome and glycome for identification of glycan cell signatures. PLoS Comput Biol. 2013;9(1):1002813.

25. Puri A, Neelamegham S. Understanding glycomechanics using mathematical modeling: A review of current approaches to simulate cellular glycosylation reaction networks. Ann Biomed Eng. 2012;40(4):816–27.

26. Liu G, Marathe DD, Matta KL, Neelamegham S. Systems-level modeling of cellular glycosylation reaction networks: O-linked glycan formation on natural selectin ligands. Bioinformatics. 2008;24(23):2740–7.

27. Liu G, Neelamegham S. A computational framework for the automated construction of glycosylation reaction networks. PLoS One. 2014;9(6):100939.

28. Yip B, Chen S, Mulder H, Hoppener J, Schachter H. Organization of the human β-2,6-acetylglucosaminyltransferase I gene (MGAT1), which controls complex and hybrid N-glycan synthesis. Biochem J. 1997;321:465–74.

29. Tan J, D’Agostaro G, Bendik B, Reck F, Sarkar M, Squire J, Leong P, Schachter H. The human UDP-N-acetylglucosamine:α-6-D-mannoside-β-1,2-N-acetylglucosaminyltransferase II gene (MGAT2). Eur J Biochem. 1995;231:317–28.

30. Yoshida A, Minowa M, Takamatsu S, Hara T, Oguri S, Ikenaga H, Takeuchi M. Tissue specific expression and chromosomal mapping of a human UDP-N-acetylglucosamine:α-1,3-D-mannoside-β-1,4-N-acetylglucosaminyltransferase GlcBiology. 1999;93(9):303–10.

31. Yoshida A, Minowa M, Takamatsu S, Hara T, Ikenaga H, Takeuchi M. A novel second isozyme of the human UDP-N-acetylglucosaminyltransferase β1,3-D-mannoside-β1,4-N-acetylglucosaminyltransferase family: cDNA cloning, expression, and chromosomal assignment. Glycoconj J. 1998;15:1115–23.

32. Larsen R, Ernst L, Nair R, Lowe J. Molecular cloning, sequence, and expression of a human GDP-L-fucose:β-D-galactoside 2-α-L-fucosyltransferase cDNA that can form the H blood group antigen. Proc Nat Acad Sci USA. 1990;87:6674–8.

33. Amado M, Almeida R, Carneiro F, Levery S, Holmes E, Nomoto M, Hollingsworth M, Hassan H, Schwentek T, Nielsen P, Bennett E, Clausen H. A family of human β3-galactosyltransferases. J Biol Chem. 1993;268(12):12770–8.

34. Ju T, Brewer K, D’Souza A, Cummings R, Canfield W. Closing and expression of human core 1 β1,3-galactosyltransferase. J Biol Chem. 2002;277(11):178–86.

35. Ihara Y, Nishikawa A, Tohma T, Soejima H, Nikawa N, Taniguchi N. cDNA cloning, expression, and chromosomal localization of human β3-n-acetylglucosaminyltransferase III (β3-NTIII). J Biochem. 1993;113:692–8.

36. Oulmouden A, Wieringa A, Pettij R, Costache M, Palcic M, Mollicone R, Oriol R, Julien R. Molecular cloning and expression of a bovine α(1,3)-fucosyltransferase gene homologous to a putative ancestor gene of the human FUT3-FUT5-FUT6 cluster. J Biol Chem. 1997;272(12):8764–73.

37. Shintani N, Natsume A, Togayachi A, Endo T, Akashima T, Yamada Y, Imai N, Nakagawa S, Kozumi S, Sekine S, Narimatsu H, Sasaki K. Identification and characterization of three novel β1,3-N-acetylglucosaminyltransferases structurally related to the β1,3-galactosyltransferase family. J Biol Chem. 2001;276(5):3498–507.

38. Inaba N, Hiruma T, Togayachi A, Iwasaki H, Wang X, Fukunaka Y, Sumi R, Kuro T, Fujimura K, Iwai T, Gotoh M, Nakamura M, Narimatsu H. A novel l-branched β1,6-N-acetylglucosaminyltransferase involved in human blood group I antigen expression. Blood. 2003;101(7):2870–6.

39. Inamori K, Endo T, Ide Y, Fuji S, Gu J, Honke K, Taniguchi N. Molecular cloning and characterization of human GNT-IX, a novel β1,6-N-acetylglucosaminyltransferase that is specifically expressed in the brain. J Biol Chem. 2003;278(44):43102–9.

40. Almeida R, Amado M, David L, Levery S, Holmes E, Merx G, van Kessel AG, Bygaard E, Hassan H, Bennett E, Clausen H. A family of human β4-galactosyltransferases. J Biol Chem. 1997;272(51):31979–91.

41. Voynow J, Kaiser R, Scanlin T, Glick M. Purification and characterization of GDP-L-fucose-N-acetyl β-D-glycosaminide α1 → 6 fucosyltransferase from cultured human skin fibroblasts. J Biol Chem. 1991;266(32):21575–7.

42. Nakayama F, Nishihara S, Iwaki S, Kudo T, Okubo R, Kaneko M, Nakamura M, Karube M, Sasaki K, Narimatsu H. CD15 expression in mature granulocytes is determined by α1,3-fucosyltransferase IX, but in promyelocytes and monocytes by α1,3-fucosyltransferase IV. J Biol Chem. 2001;276(19):16100–6.

43. Bai X, Zhou D, Brown J, Crawford B, Hennet T, Esko J. Biosynthesis of the linkage region of glycosaminoglycans. J Biol Chem. 2001;276(51):48189–95.

44. Takashima S, Tsuji S, Tsujimoto M. Characterization of the second type of human β-galactoside α2,6-sialyltransferase (ST6Gal II), which sialylates Galβ1,4GlcNAc structures on oligosaccharides preferentially. J Biol Chem. 2002;277(48):45719–28.

45. Kitagawa H, Paulson J. Cloning of a novel α2,3-sialyltransferase that sialylates glycoprotein and glycolipid carbohydrate groups. J Biol Chem. 1994;269(2):1394–401.

46. Kitagawa H, Paulson J. Differential expression of five sialyltransferase genes in human tissues. J Biol Chem. 1994;269(27):17872–8.

47. del Val U, Nagy JM, Kontoravdi C. A dynamic mathematical model for monoclonal antibody N-linked glycosylation and nucleotide sugar donor transport within a maturing Golgi apparatus. Biotechnol Prog. 2011;27(6):1730–43.

48. Hossler P, Mulukutla B, Hu W. Systems analysis of N-glycan processing in mammalian cells. PLoS One. 2007;2(8):713.