Glycoconjugate Journal 21, 17–24, 2004

Construction of a human glycogene library and comprehensive functional analysis

Hisashi Narimatsu

Glycogene Function Team, Research Center for Glycoscience, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, 305-8568, Japan

Eighteen years have passed after the first mammalian glycosyltransferase was cloned. At the beginning of April, 2001, 110 genes for human glycosyltransferases, including modifying enzymes for carbohydrate chains such as sulfotransferases, had been cloned and analyzed. We started the Glycogene Project (GG project) in April 2001, a comprehensive study on human glycogenes with the aid of bioinformatic technology. The term glycogene includes the genes for glycosyltransferases, sulfotransferases adding sulfate to carbohydrates and sugar-nucleotide transporters, etc. Firstly, as many novel genes, which are the candidates for glycogenes, as possible were searched using bioinformatic technology in databases. They were then cloned and expressed in various expression systems to detect the activity for carbohydrate synthesis. Their substrate specificity was determined using various acceptors.
Published in 2004.

Keywords: **glycogene, glycogene project, glycosyltransferase, sulfotransferase, sugar-nucleotide transporter**

Introduction

Human genome analysis has almost been completed, and proteome analysis, that is, comprehensive functional analysis of proteins, is underway in research laboratories throughout the world as a post-genomic strategy. However, in many cases, proteins undergo posttranslational modifications such as phosphorylation and glycosylation. Since these modifications confer various functions on proteins as biosubstances, the consideration of protein functions should include posttranslational modifications. A vast amount of information on the sites and functions of phosphorylation has accumulated. Despite the estimate that 50% or more of proteins undergo glycosylation, research on glycosylation has been delayed by the extreme difficulty in the analysis of sugar chains due to their diversity and complexity.

The genes involved in the addition of the sugar portions of glycoconjugates are generically named glycogenes, which include the genes for (1) glycosyltransferases, (2) glycolytic enzymes, (3) sugar nucleotide synthetases, (4) sugar nucleotide transporters, and, in a broader sense, sugar chain-recognizing molecules such as (5) lectins and (6) sugar-chain receptors for the synthesized sugar chains to fulfill functions. One gene is involved in the synthesis of the protein moiety of a glycoprotein, whereas dozens of genes are involved in the synthesis of its sugar chain moiety. Thus, a single glycoprotein is the joint product of dozens of genes. Comprehensive identification and functional analyses of glycogenes are expected to dramatically advance: (1) the elucidation of the functions of sugar chains as bioactive substances including carrier molecules, (2) the elucidation of control mechanisms of carrier molecules by sugar chains, (3) the development of analytical techniques for sugar chain structures, and (4) the development of automated synthesis techniques for sugar chains. Under the project "construction of a glycogene library", we have been performing a comprehensive search for glycogenes in the human genome and cDNA databases utilizing bioinformatics, particularly glycosyltransferase genes playing central roles in sugar chain biosynthesis, and have been analyzing their functions. This paper reviews the current status of glycogene cloning and functional analysis, and assesses the future prospects for sugar chain research.

To whom correspondence should be addressed: Hisashi Narimatsu, Glycogene Function Team, Research Center for Glycoscience, National Institute of Advanced Industrial Science and Technology (AIST), OSL C-2, 1-1-1 Umezono, Tsukuba, Ibaraki, 305-8568, Japan. E-mail: h.narimatsu@aist.go.jp

Search of novel glycosyltransferases using bioinformatics

More than 160 human glycosyltransferases including sulfotransferases have been cloned to date, and the number of cloned enzymes is increasing, and may reach 200 within one year. This is mainly because of the expansion of the genome database. Since glycosyltransferases are not necessarily expressed in sufficient amounts, they were frequently not registered in the EST

or cDNA database. However, since the primary sequence analysis of the human genome has been completed, gene search is now possible from the database. In addition, the improved accuracy of gene prediction software has also contributed to the acceleration of the discovery of glycosyltransferases.

In the project "Construction of a human glycogene library and comprehensive functional analysis" (the GG project) supported by the New Energy and Industrial Technology Development Organization (NEDO) of the Ministry of Economy, Trade and Industry (METI) of Japan, we constructed a system for predicting novel glycosyltransferase genes from the primary sequence characteristics shared by glycosyltransferases. This system has functions such as an automated BLAST search, automatic removal of known sequences, assembly of EST search results (EST sequence hits are assembled with Phrap to eliminate duplicate hits for better search efficiency), and prediction of gene regions in genome sequence search results (in the genome sequence search, gene regions are predicted with GENSCAN, and the amino acid sequence for the entire ORF is obtained), in combination with the below described program to identify glycosyltransferases. The general characteristics of glycosyltransferases are as follows: (1) The *N*-terminus is short and within the cytoplasm; (2) The trans-Golgi membrane region (a sequence of approximately 18–20 residues rich in hydrophobic amino acids) is in the vicinity of the amino terminus; (3) This region is followed by the stem domain that is rich in proline (an amino acid forming a no high-order structure) and in the O-glycan-linked amino acid residues serine and threonine; (4) The *C*-terminus is followed by an approximately 300–400-residue catalytic site; (5) The catalytic site contains DxD or DxH, an amino acid sequence necessary for binding divalent cations [1–4]. In addition to these characteristics, several motif regions were set up for every existing gene family using MEME [5], then a program was designed to assess six items including whether the motif regions are present, and has been used to determine possible glycosyltransferases. However, since this system is only capable of detecting genes with some degree of homology to the existing ones, we are in the process of expanding its functions, such as the incorporation of software capable of comparing high-order structures. These systems are sufficiently applicable to the analysis of not only glycosyltransferase genes but also various other genes including lectin genes [4].

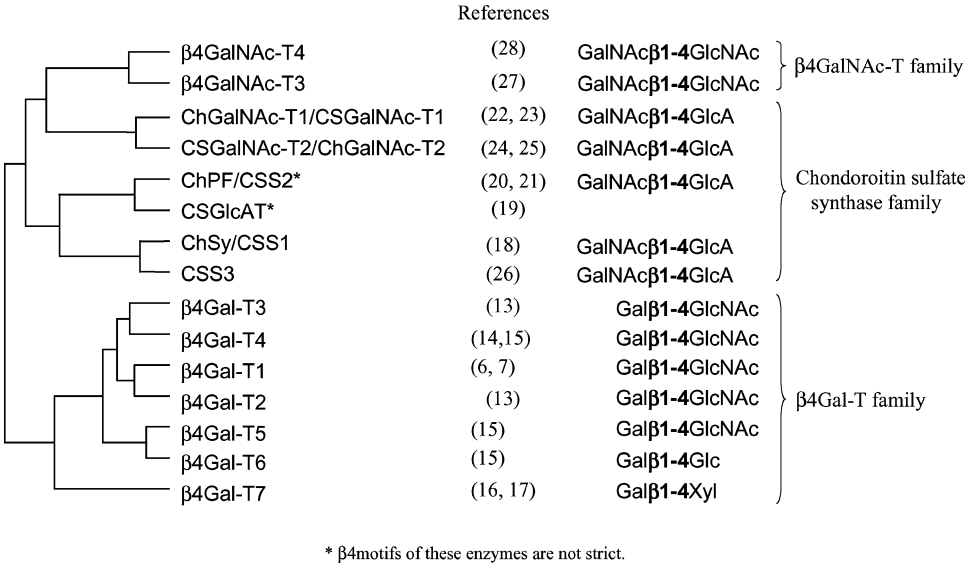
Current status of glycogene research

Since the first mammalian glycosyltransferase gene was cloned by two groups independently [6,7], more than 160 human glycosyltransferase genes have been cloned to date, mostly by Japanese researchers [8]. In animals other than humans and mice, glycogenes have been actively cloned in lower animals such as fruit flies and nematodes [9,10], in which gene deletion mutants can be easily produced. The sequence of studies is as follows: (1) Glycogenes are cloned, and the encoded enzymes are expressed by various methods; (2) The enzymes are analyzed for the synthesis of sugar chain structures *in vitro* or *in vivo*. Subsequently, the synthesized sugar chain structures are examined for their biological functions as follows: (3) Glycogenes are transfected into cultured cells, and the functions of the cells whose sugar chain structures have undergone changes are analyzed; (4) Transgenic mice are produced for functional analysis; (5) The genetic basis of human diseases and blood-types [11] is explored with SNPs analysis of glycogenes identified by chromosomal mapping of the human genetic diseases and blood-types; (6) Knockout mice are produced to analyze the functions of glycogenes in the individual. Mice knocked out in terms of genes for glycosyltransferases synthesizing the essential portion of sugar chains are lethal. Deletion of enzymes synthesizing the intermediate portion results in the development of serious symptoms, and deletion of those for the terminal portion leads to an apparently normal phenotype. However, a detailed analysis reveals abnormal phenotypes reflecting individual characteristics. Recent advances in the siRNA technique have enabled knock-down at the cellular level. Although biological function analysis is in progress in parallel with glycogene cloning as described above, further analyses of sugar chain structures and functions are required to look at the changes of entire glycoconjugate molecules.

The characteristics of a few glycosyltransferase gene families and their substrate specificities are presented below along with the results of the GG project. For more details, the reader is referred to our cited publications.

Glycosyltransferase gene family with β4-glycosyltransferase (β4GT) motif

The β4-galactosyltransferase (β4Gal-T) gene was the first glycosyltransferase cloned [6,7]. β4Gal-T forms a family including seven enzymes [12–17], β4Gal-T1 through β4Gal-T7, which share the same characteristic of transferring galactose from the sugar donor UDP-Gal to the sugar acceptor via a β1,4 linkage, but differ in using sugar acceptors such as sugar chains in glycoproteins, glycolipids, or glycosaminoglycans (Figure 1). Using the amino acid sequences of this family as query sequences, we performed database searches and found eight new members sharing a β4GT motif, GWGXED. The results of phylogenetic analysis of these gene families are shown in Figure 1. The results reported by Kitagawa *et al.* and us showed that six novel glycosyltransferases were involved in the synthesis of chondroitin sulfate [18–26]. Chondroitin sulfate occurs as a sugar chain of proteoglycans in which dozens of the disaccharide repeating unit (-GlcAβ1,3GalNAcβ1,4-) are bound to a serine residue of the core protein via a linkage tetrasaccharide (GlcAβ1,3Galβ1,3Galβ1,4Xyl). Of these six enzymes, the catalytic domain containing the β4GT motif, GWGXED, probably functions in transferring GalNAc from UDP-GalNAc to the sugar acceptor GlcA via a β1,4 linkage. Three of them, CSS1, CSS2 and CSS3 in Figure 1, have



* β4motifs of these enzymes are not strict.

Figure 1. A phylogenetic tree of the human β4-GT family. The references reported are in the middle. The sugar linkage catalyzed by each enzyme is on the right.

another catalytic domain containing a β3glycosyltransferase motif (β3GT motif as described later) which catalyzes the transfer of GlcA to GalNAc with a β3-linkage.

In addition, two novel members were found recently and their substrate specificity was analyzed [27,28]. β4GalNAc-T3 and -T4 effectively synthesized N,N'-diacetyllactosediamine (LacdiNAc), GalNAcβ1-4GlcNAc, at non-reducing termini of various acceptors derived not only from N-glycans, but also from O-glycans. They were identified to be the enzymes responsible for the synthesis of the LacdiNAc structure. All members in Figure 1, the β4-galactosyltransferase gene, the chondroitin sulfate-synthesizing enzyme gene and the LacdiNAc synthase families have the amino acid sequence GWGXED in common, but differ in their sugar donors and acceptors, suggesting that this motif is the sequence specifying the β1,4 linkage. Many other β1,4-linked sugar-chain structures have been reported to occur in the body, but the corresponding glycosyltransferases remain unknown. We predict that the genes for the glycosyltransferases synthesizing them will probably be included in this gene family.

Glycosyltransferase gene family with β3-glycosyltransferase (β3GT) motif

The glycosyltransferase gene family with the β3 motif includes five β3-galactosyltransferases (β3Gal-T), six β3-*N*-acetylglucosaminyltransferases (β3Gn-T) and two β3-*N*-acetylgalactosaminyltransferases (β3GalNAc-T). The results of phylogenetic analysis of this gene family are shown in Figure 2. The glycosyltransferases in this family transfer sugars via a β1,3 linkage, and conserve three motifs in their amino acid sequences. Motif 2 contains DxD, and motif 3 is thought to specify the β1,3 linkage. The first β3Gal-T1 gene was isolated in 1994 by Sasaki *et al*. by expression cloning [29]. The enzyme for synthesizing GD1b/GM1/GA1 was first cloned by Miyazaki *et al*. by an expression cloning method [30]. This enzyme had the β3GT motifs, and was named β3Gal-T4. Later, the expansion of the genome database led to the report of the β3Gal-T2 through β3Gal-T4 genes [31,32]. Later on, the specificity of β3Gal-T3 was precisely analyzed, and corrected to be β3GalNAc-T1 which synthesizes globoside (Gb4) [33]. By a cloning technique using degenerate primers corresponding to the β3GT motifs, we isolated and reported the β3Gal-T5 gene that synthesizes type I sugar-chain antigens including the CA-19-9 antigen (sialyl Lewis a; sLe^a) [34]. β3Gal-T5 also synthesizes a stage-specific embryonic antigen-3 (SSEA-3) [35]. β3Gal-T5 was found to be expressed specifically in the gastrointestinal tract, and to be very important in the expression of the above-described antigens in gastrointestinal cancer cells [34]. Interestingly, transcriptional regulation of the β3Gal-T5 gene is controlled by the intestinal homeoproteins, *i.e.* members of caudal-related homeobox protein (Cdx) and hepatocyte nuclear factor (HNF) families [36]. In addition, β3Gal-T6 was reported to synthesize the linkage region in proteoglycan sugar chains [37].

Zhou *et al*. reported β3Gn-T1 which exerts a GlcNAc-transfer to Gal with a β3-linkage [38]. Thereafter, we reported three members of β3Gn-Ts, *i.e.* β3Gn-T2, -T3 and -T4 [39]. After our publication, Zhou *et al.* corrected the sequence of β3Gn-T1 to be identical to that of β3Gn-T2. Therefore, β3Gn-T1 and -T2 are identical. Of β3Gn-T2 through β3Gn-T4 reported by us, β3Gn-T2 clearly shows stronger activity for the sugar acceptor polylactosamine structures, suggesting that this is a polylactosamine-synthesizing enzyme [40]. β3Gn-T3 was further identified to mainly transfer GlcNAc to the core 1 structure [41]. We performed an *in silico* analysis using the database,

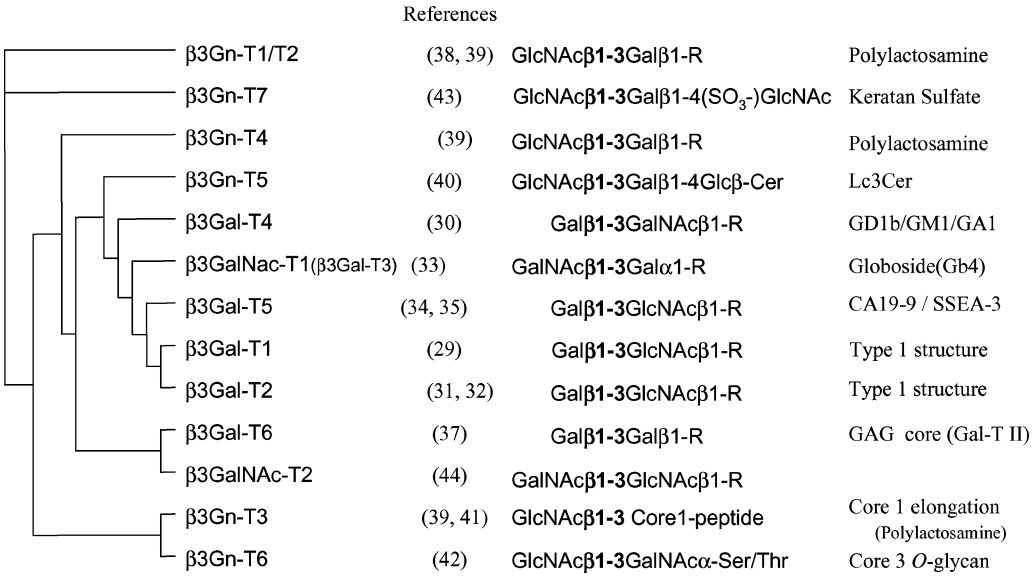


Figure 2. A phylogenetic tree of the human β3-GT family. The references reported are in the middle. The sugar linkage catalyzed by each enzyme is on the right.

successfully cloned the Lc_3Cer-synthesizing enzyme gene, β3Gn-T5 [40], and the O-liked sugar-chain core 3-synthesizing enzyme gene, β3Gn-T6 [42], and reported their substrate specificities. β3Gn-T6 is restrictively localized in the epithelia of the stomach and colon [42], in which the major form of O-glycans is the core 3 structure. Very recently, β3Gn-T7 was identified as being responsible for the synthesis of keratan sulfate [43]. A novel enzyme, named β3GalNAc-T2, was found by *in silico* cloning and characterized [44]. β3GalNAc-T2 transfers GalNAc to GlcNAc with a β3-linkage on the termini of N- and O-glycans. Although the GalNAcβ1-3GlcNAcβ1-R structure has not been reported in humans and other mammals, it must exist where the enzyme is expressed. Interestingly, the β3GalNAc-T2 transcripts were most highly expressed in the germ cells of testis in a stage specific manner [44]. This may be involved in the maturation of sperm.

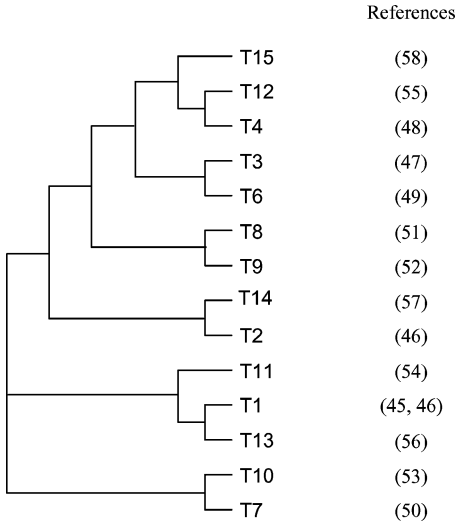


Figure 3. A phylogenetic tree of the human pp-GalNAc-T family. The references reported are on the right.

Polypeptide *N*-acetylgalactosaminyltransferase gene family

Polypeptide *N*-acetylgalactosaminyltransferase (pp-GalNAc-T) is a group of glycosyltransferases transferring GalNAc to serine or threonine residues in O-glycans through an α-linkage. At the present time, 15 members of this family have been published (Figure 3) [45–58]. pp-GalNAc-T5 is not listed in Figure 3, because its activity has not been demonstrated. We cloned five of these genes, pp-GalNAc-T10, -T12, -T13, -T14 and -T15, and analyzed their functions [53,55–58]. The sequences of this enzyme family are characterized by the presence of conserved motifs called the GT1 and Gal/GalNAc-T motifs, and have homology throughout the entire lengths, unlike the β1,3 and β1,4 glycosyltransferase families. Thus, it is easy to find homologous members in the databases. Besides 15 members published, an additional four or five members can be found in the databases. Most pp-GalNAc-T enzymes have a lectin-like domain in the *C*-terminal portion, suggesting that it influences the enzyme substrate specificity. Fourteen kinds of pp-GalNAc-T differ in relative activities determined with different polypeptide-chain sequences; however, few studies have analyzed which polypeptide chains are used as acceptor substrates by the respective enzymes. Thus, the identification of sugar acceptor proteins of pp-GalNAc-T awaits further study. An example of such analyses is the case of multiple O-linked sugar chains in the hinge region of IgA-1. Based on the tissue distribution of expression of the pp-GalNAc-T family and the differences in their relative activity determined with the hinge

region peptide, we reported that pp-GalNAc-T2 transfers *N*-acetylgalactosamine to the IgA-1 hinge region [59].

Other glycogenes characterized in the GG project

A β1,6-*N*-acetylglucosaminyltransferase, GnT-IX, that is homologous to GnT-V, was cloned and characterized by Inamori *et al.* [60,61]. This was cloned almost at the same time by the others [62]. GnT-IX is expressed specifically in the brain and catalyzes the transfer of GlcNAc to the 6-OH position of the mannose residue of GlcNAcβ1,2-Manalpha on both the α1,3- and α1,6-linked mannose arms in the core structure of N-glycan. This enzyme also transfers GlcNAc to GlcNAcβ1,2Manα in O-mannosyl glycan as a β1,6-linkage.

Two novel β-1,6-*N*-acetylglucosaminyltransferases (IGnT), which are homologous to IGnT1, have been cloned, analyzed and named IGnT2 and IGnT3 [63]. They showed differential expression in tissues, and only IGnT-3 was expressed in reticulocytes and determined the human Ii blood group.

An enzyme homologous to the core 1 synthase [64] has been cloned and characterized [65]. We named this gene C1Gal-T2 because of the following results. By the transfection experiments with the cells lacking the core 1 synthase activity, the cell homogenates expressing C1Gal-T2 apparently contained the core 1 synthesizing activity. However, Ju *et al.* reported that C1Gal-T2 is not a enzyme, but a molecular chaperone that helps the folding of C1Gal-T1 [66].

Heparan sulfate D-glucosaminyl 3-O-sulfotransferases (3-OSTs) catalyze the transfer of sulfate from 3′-phosphoadenosine 5′-phosphosulfate (PAPS) to position 3 of the glucosamine residue of heparan sulfate and heparin. A sixth member of the human 3-OST family, named 3-OST-5, was characterized in the GG project [67]. Recombinant 3-OST-5 only exhibited sulfotransferase activity toward heparan sulfate and heparin. 3-OST-5 can synthesize a tetrasulfated disaccharide unit. To our knowledge, this is the first report to describe the enzymatic formation of this novel structure.

Two glycogenes encoding sugar-nucleotide transporters have been cloned and analyzed [68,69]. In the reference 68, we demonstrated the activity of PAPS transporter for the first time.

Current status of research on glycosyltransferases and their functional analysis

Glycosyltransferases are membrane proteins localized in the rough endoplasmic reticulum and Golgi apparatus. About half of the proteins synthesized in ribosomes successively undergo several steps of glycosylation by glycosyltransferases in the rough endoplasmic reticulum and Golgi apparatus cisternae, and become biologically active mature proteins. Glycosylation by glycosyltransferases requires sugar-donor substrates (sugar nucleotides), sugar-acceptor substrates and, in many cases, divalent cations. Sugar chains vary in structure and length according to tissue (cell) type, developmental stage, differentiation, and malignant transformation, resulting in sugar chains of heterogeneous structures on the same carrier molecule. This complexity of sugar chain synthesis is produced by a close control of the expression of a large number of enzymes and the substrate specificities of enzymes themselves. However, *in vitro* activity and substrate specificity may not reflect the actual synthesis of sugar chains *in vivo*. Frequently, glycosyltransferases specifically recognize the protein and lipid portions, or sugar chains in sugar acceptor substrates. If such a complex of substrate specificity is involved, it is almost impossible to characterize it with *in vitro* experiments. The more complex the acceptor sugar-chains or glycoconjugates, the less possible their organic chemical synthesis, necessitating the purification of substrates from biological materials. However, this is difficult to perform and, moreover, yields are very low. On the other hand, rapid advances in the cloning of glycogenes have enabled the modification or conversion of the sugar-chain portions of glycoproteins; however, at this time, it is far from a total synthesis. Sugar-chain structures that can be enzymatically synthesized are limited. Even if they can be synthesized, the problem of low yields remains. In addition, it is not easy to analyze what sugar-chain structures are added to what amino acids in proteins by current sugar-chain analysis techniques. Under these circumstances, an extremely laborious procedure is needed to clarify the details of sugar-chain biosynthesis by a group of glycogenes. In addition, recent studies have reported that more than one glycosyltransferase form complexes and transfer sugars, and have suggested the presence of a molecular chaperone [66]. In such a case, attention to only one glycosyltransferase may not allow the detection of enzyme activity.

The substrate specificities of glycosyltransferases are usually first analyzed *in vitro*, but there is no guarantee that substrate specificities *in vitro* reflect those *in vivo*. Therefore, glycosyltransferase genes are transfected into cultured cell lines to express these enzymes at the same levels as those *in vivo* and structural changes in sugar chains should be analyzed. Sugar-chain structural analysis in knockout mice or human natural variants is another approach. However, the methods of identifying sugar-chain structures are highly laborious and time-consuming. Moreover, the number of sugar-chain structures resulting from a complex combination of many groups of glycosyltransferases is enormous, making it impossible to analyze all of them at present. Despite the piling up of many such problems, sugar chain science is a fascinating area of research to be advanced over a long period.

Conclusion

At present, sugar-chain engineering has been given the status of one of the key techniques of biotechnology in Japan. There is an urgent need for the development of technology in the three areas of sugar-chain, *i.e.* synthesis, structure analysis, and function analysis to advance sugar-chain research. All of these must be developed hand in hand, and the lack of any one of them

becomes an obstacle. Therefore, delineation of glycogenes with an emphasis on glycosyltransferases is a beginning. In an era when gene cloning and functional analysis have advanced, we have no doubt about that new research on sugar chains will be started with the advances in next-generation technologies.

Acknowledgment

The GG project was performed as part of the R&D Project of Industrial Science and Technology Frontier Program (R&D for Establishment and Utilization of a Technical Infrastructure for Japanese Industry) supported by the New Energy and Industrial Technology Development Organization (NEDO).

References

1 Paulson JC, Colley KJ, Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation, *J Biol Chem* **264**, 17615–8 (1989).
2 Narimatsu H, Recent progress in molecular cloning of glycosyltransferase genes of eukaryotes, *Microbiol Immunol* **38**, 489–504 (1994).
3 Martinez-Duncker I, Mollicone R, Candelier JJ, Breton C, Oriol R, A new superfamily of protein-*O*-fucosyltransferases, α2-fucosyltransferases, and α6-fucosyltransferases: Phylogeny and identification of conserved peptide motifs, *Glycobiology* **13**, 1–5 (2003).
4 Kikuchi N, Kwon YD, Gotoh M, Narimatsu H, Comparison of glycosyltransferase families using the profile hidden Markov model, *Biochem Biophys Res Commun* **310**, 574–9 (2003).
5 Bailey TL, Elkan C, Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, AAAI Press, pp. 28–36, 1994.
6 Shaper NL, Shaper JH, Meuth JL, Fox JL, Chang H, Kirsch IR, Hollis GF, Bovine galactosyltransferase: Identification of a clone by direct immunological screening of a cDNA expression library, *Proc Natl Acad Sci USA* **83**, 1573–7 (1986).
7 Narimatsu H, Sinha S, Brew K, Okayama H, Qasba PK, Cloning and sequencing of cDNA of bovine *N*-acetylglucosamine (β1-4)galactosyltransferase, *Proc Natl Acad Sci USA* **83**, 4720–4 (1986).
8 Taniguchi N, Honke K, Fukuda M (Eds.), *Handbook of Glycosyltransferases and Related Genes* (Springer-Verlag Tokyo Berlin Heidelberg New York, 2001).
9 Kawar ZS, Van Die I, Cummings RD, Molecular cloning and enzymatic characterization of a UDP-GalNAc:GlcNAcβ-R β1,4-*N*-acetylgalactosaminyltransferase from *Caenorhabditis elegans*, *J Biol Chem* **277**, 34924–32 (2002).
10 Zheng Q, Van Die I, Cummings RD, Molecular cloning and characterization of a novel α 1,2-fucosyltransferase (CE2FT-1) from *Caenorhabditis elegans*, *J Biol Chem* **277**, 39823–32 (2002).
11 Blumenfeld O, *Blood Group Antigen Gene Mutation Databases*. http://www.bioc.aecom.yu.edu/bgmut/index.htm
12 Lo NW, Shaper JH, Pevsner J, Shaper NL, The expanding β4-galactosyltransferase gene family: Messages from the databanks, *Glycobiology* **8**, 517–26 (1998).
13 Almeida R, Amado M, David L, Levery SB, Holmes EH, Merkx G, van Kessel AG, Rygaard E, Hassan H, Bennett E, Clausen H, A family of human β4-galactosyltransferases. Cloning and expression of two novel UDP-galactose: β-*N*-acetylglucosamine β1, 4-galactosyltransferases, β4Gal-T2 and β4Gal-T3, *J Biol Chem* **272**, 31979–91 (1997).
14 Schwientek T, Almeida R, Levery SB, Holmes EH, Bennett E, Clausen H, Cloning of a novel member of the UDP-galactose:β-*N*-acetylglucosamine β1,4-galactosyltransferase family, β4Gal-T4, involved in glycosphingolipid biosynthesis, *J Biol Chem* **273**, 29331–40 (1998)
15 Amado M, Almeida R, Schwientek T, Clausen H, Identification and characterization of large galactosyltransferase gene families: Galactosyltransferases for all functions, *Biochim Biophys Acta* **1473**, 35–53 (1999).
16 Okajima T, Yoshida K, Kondo T, Furukawa K, Human homolog of *Caenorhabditis elegans sqv-3* gene is galactosyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans, *J Biol Chem* **274**, 22915–8 (1999).
17 Almeida R, Levery SB, Mandel U, Kresse H, Schwientek T, Bennett EP, Clausen H, Cloning and expression of a proteoglycan UDP-galactose:β-xylose β1,4-galactosyltransferase I. A seventh member of the human β4-galactosyltransferase gene family, *J Biol Chem* **274**, 26165–71 (1999).
18 Kitagawa H, Uyama T, Sugahara K, Molecular cloning and expression of a human chondroitin synthase, *J Biol Chem* **276**, 38721–6 (2001).
19 Gotoh M, Yada T, Sato T, Akashima T, Iwasaki H, Mochizuki H, Inaba N, Togayachi A, Kudo T, Watanabe H, Kimata K, Narimatsu H, Molecular cloning and characterization of a novel chondroitin sulfate glucuronyltransferase that transfers glucuronic acid to *N*-acetylgalactosamine, *J Biol Chem* **277**, 38179–88 (2002).
20 Kitagawa H, Izumikawa T, Uyama T, Sugahara K, Molecular cloning of a chondroitin polymerizing factor that cooperates with chondroitin synthase for chondroitin polymerization, *J Biol Chem* **278**, 23666–71 (2003).
21 Yada T, Gotoh M, Sato T, Shionyu M, Go M, Kaseyama H, Iwasaki H, Kikuchi N, Kwon YD, Togayachi A, Kudo T, Watanabe H, Narimatsu H, Kimata K, Chondroitin sulfate synthase-2. Molecular cloning and characterization of a novel human glycosyltransferase homologous to chondroitin sulfate glucuronyltransferase, which has dual enzymatic activities, *J Biol Chem* **278**, 30235–47 (2003).
22 Uyama T, Kitagawa H, Tamura Ji J, Sugahara K, Molecular cloning and expression of human chondroitin *N*-acetylgalactosaminyltransferase: The key enzyme for chain initiation and elongation of chondroitin/dermatan sulfate on the protein linkage region tetrasaccharide shared by heparin/heparan sulfate, *J Biol Chem* **277**, 8841–6 (2002).
23 Gotoh M, Sato T, Akashima T, Iwasaki H, Kameyama A, Mochizuki H, Yada T, Inaba N, Zhang Y, Kikuchi N, Kwon YD, Togayachi A, Kudo T, Nishihara S, Watanabe H, Kimata K, Narimatsu H, Enzymatic synthesis of chondroitin with a novel chondroitin sulfate *N*-acetylgalactosaminyltransferase that transfers *N*-acetylgalactosamine to glucuronic acid in initiation and elongation of chondroitin sulfate synthesis, *J Biol Chem* **277**, 38189–96 (2002).
24 Sato T, Gotoh M, Kiyohara K, Akashima T, Iwasaki H, Kameyama A, Mochizuki H, Yada T, Inaba N, Togayachi A, Kudo T, Asada

M, Watanabe H, Imamura T, Kimata K, Narimatsu H, Differential roles of two *N*-acetylgalactosaminyltransferases, CSGalNAcT-1, and a novel enzyme, CSGalNAcT-2. Initiation and elongation in synthesis of chondroitin sulfate, *J Biol Chem* **278**, 3063–71 (2003).

25 Uyama T, Kitagawa H, Tanaka J, Tamura J, Ogawa T, Sugahara K, Molecular cloning and expression of a second chondroitin *N*-acetylgalactosaminyltransferase involved in the initiation and elongation of chondroitin/dermatan sulfate, *J Biol Chem* **278**, 3072–78 (2003).

26 Yada T, Sato T, Kaseyama H, Gotoh M, Iwasaki H, Kikuchi N, Kwon YD, Togayachi A, Kudo T, Watanabe H, Narimatsu H, Kimata K, Chondroitin sulfate synthase-3. Molecular cloning and characterization, *J Biol Chem* **278**, 39711–25 (2003).

27 Sato T, Gotoh M, Kiyohara K, Kameyama A, Kubota T, Kikuchi N, Ishizuka Y, Iwasaki H, Togayachi A, Kudo T, Ohkura T, Nakanishi H, Narimatsu H, Molecular cloning and characterization of a novel human β1,4-*N*-acetylgalactosaminyltransferase, β4GalNAc-T3, responsible for the synthesis of *N*, *N*′-diacetyllactosediamine, GalNAc β1-4GlcNAc, *J Biol Chem* **278**, 47534–44 (2003).

28 Gotoh M, Sato T, Kiyohara K, Kameyama A, Kikuchi N, Kwon YD, Ishizuka Y, Iwai T, Nakanishi H, Narimatsu H, Molecular cloning and characterization of β1,4-*N*-acetylgalactosaminyltransferases IV synthesizing *N*,*N*′-diacetyllactosediamine, *FEBS Lett* **562**, 134–40 (2004).

29 Sasaki K, Sasaki E, Kawashima K, Hanai N, Nishi T, Hasegawa M, Japan patent 0618759 A 940705 (1994).

30 Miyazaki H, Fukumoto S, Okada M, Hasegawa T, Furukawa K, Expression cloning of rat cDNA encoding UDP-galactose: GD2 β1,3-galactosyltransferase that determines the expression of GD1b/GM1/GA1, *J Biol Chem* **272**, 24794–9 (1997).

31 Hennet T, Dinter A, Kuhnert P, Mattu TS, Rudd PM, Berger EG, Genomic cloning and expression of three murine UDP-galactose: β-*N*-acetylglucosamine β1,3-galactosyltransferase genes, *J Biol Chem* **273**, 58–65 (1998).

32 Amado M, Almeida R, Carneiro F, Levery SB, Holmes EH, Nomoto M, Hollingsworth MA, Hassan H, Schwientek T, Nielsen PA, Bennett EP, Clausen H, A family of human β3-galactosyltransferases. Characterization of four members of a UDP-galactose:β-*N*-acetylglucosamine/ β-*N*-acetylgalactosamine β-1,3-galactosyltransferase family, *J Biol Chem* **273**, 12770–8 (1998).

33 Okajima T, Nakamura Y, Uchikawa M, Haslam DB, Numata SI, Furukawa K, Urano T, Furukawa K, Expression cloning of human globoside synthase cDNAs. Identification of β3Gal-T3 as UDP-*N*-acetylgalactosamine: Globotriaosylceramide β1,3-*N*-acetylgalactosaminyltransferase, *J Biol Chem* **275**, 40498–503 (2000).

34 Isshiki S, Togayachi A, Kudo T, Nishihara S, Watanabe M, Kubota T, Kitajima M, Shiraishi N, Sasaki K, Andoh T, Narimatsu H, Cloning, expression, and characterization of a novel UDP-galactose: β-*N*-acetylglucosamine β1,3-galactosyltransferase (β3Gal-T5) responsible for synthesis of type 1 chain in colorectal and pancreatic epithelia and tumor cells derived therefrom, *J Biol Chem* **274**, 12499–507 (1999).

35 Zhou D, Henion TR, Jungalwala FB, Berger EG, Hennet T, The β1,3-galactosyltransferase β3GalT-V is a stage-specific embryonic antigen-3 (SSEA-3) synthase, *J Biol Chem* **275**, 22631–4 (2000).

36 Isshiki S, Kudo T, Nishihara S, Ikehara Y, Togayachi A, Furuya A, Shitara K, Kubota T, Watanabe M, Kitajima M, Narimatsu H, Lewis type 1 antigen synthase (β3Gal-T5) is transcriptionally regulated by homeoproteins, *J Biol Chem* **278**, 36611–20 (2003).

37 Bai X, Zhou D, Brown JR, Crawford BE, Hennet T, Esko JD, Biosynthesis of the linkage region of glycosaminoglycans: Cloning and activity of galactosyltransferase II, the sixth member of the β1,3-galactosyltransferase family (β3GalT6), *J Biol Chem* **276**, 48189–95 (2001).

38 Zhou D, Dinter A, Gutierrez Gallego R, Kamerling JP, Vliegenthart JF, Berger EG, Hennet T, β-1,3-*N*-acetylglucosaminyltransferase with poly-*N*-acetyllactosamine synthase activity is structurally related to β-1,3-galactosyltransferases, *Proc Natl Acad Sci USA* **96**, 406–11 (1999).

39 Shiraishi N, Natsume A, Togayachi A, Endo T, Akashima T, Yamada Y, Imai N, Nakagawa S, Koizumi 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* **276**, 498–507 (2001).

40 Togayachi A, Akashima T, Ookubo R, Kudo T, Nishihara S, Iwasaki H, Natsume A, Mio H, Inokuchi J, Irimura T, Sasaki K, Narimatsu H, Molecular cloning and characterization of UDP-GlcNAc:lactosylceramide β1,3-*N*-acetylglucosaminyltransferase (β3Gn-T5), an essential enzyme for the expression of HNK-1 and Lewis X epitopes on glycolipids, *J Biol Chem* **276**, 22032–40 (2001).

41 Yeh JC, Hiraoka N, Petryniak B, Nakayama J, Ellies LG, Rabuka D, Hindsgaul O, Marth JD, Lowe JB, Fukuda M, Novel sulfated lymphocyte homing receptors and their control by a Core1 extension β1,3-*N*-acetylglucosaminyltransferase, *Cell* **105**, 957–69 (2001).

42 Iwai T, Inaba N, Naundorf A, Zhang Y, Gotoh M, Iwasaki H, Kudo T, Togayachi A, Ishizuka Y, Nakanishi H, Narimatsu H, Molecular cloning and characterization of a novel UDP-GlcNAc:GalNAc-peptide β1,3-*N*-acetylglucosaminyltransferase (β3Gn-T6), an enzyme synthesizing the core 3 structure of O-glycans, *J Biol Chem* **277**,12802–9 (2002).

43 Seko A, Yamashita K, β1,3-*N*-acetylglucosaminyltransferase-7 (β3Gn-T7) acts efficiently on keratan sulfate-related glycans, *FEBS Lett* **556**, 216–20 (2004).

44 Hiruma T, Togayachi A, Okamura K, Sato T, Kikuchi N, Kwon YD, Nakamura A, Fujimura K, Gotoh M, Tachibana K, Ishizuka Y, Noce T, Nakanishi H, Narimatsu H, A novel human β1,3-*N*-acetylgalactosaminyltransferase that synthesizes a unique carbohydrate structure, GalNAcβ1-3GlcNAc, *J Biol Chem* **279**, 14087–95 (2004).

45 Homa FL, Hollander T, Lehman DJ, Thomsen DR, Elhammer AP, Isolation and expression of a cDNA clone encoding a bovine UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase, *J Biol Chem* **268**, 12609–16 (1993).

46 White T, Bennett EP, Takio K, Sorensen T, Bonding N, Clausen H, Purification and cDNA cloning of a human UDP-*N*-acetyl-alpha-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase, *J Biol Chem* **270**, 24156–65 (1995).

47 Bennett EP, Hassan H, Clausen H cDNA cloning and expression of a novel human UDP-*N*-acetyl-alpha-D-galactosamine. Polypeptide *N*-acetylgalactosaminyltransferase, GalNAc-T3, *J Biol Chem* **271**, 17006–12 (1996).

48 Bennett EP, Hassan H, Mandel U, Mirgorodskaya E, Roepstorff P, Burchell J, Taylor-Papadimitriou J, Hollingsworth MA, Merkx G, van Kessel AG, Eiberg H, Steffensen R, Clausen H, Cloning of a human UDP-*N*-acetyl-alpha-D-Galactosamine:polypeptide *N*-acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat, *J Biol Chem* **273**, 30472–81 (1998).

49 Bennett EP, Hassan H, Mandel U, Hollingsworth MA, Akisawa N, Ikematsu Y, Merkx G, van Kessel AG, Olofsson S, Clausen H, Cloning and characterization of a close homologue of human UDP-*N*-acetyl-alpha-*D*-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase-T3, designated GalNAc-T6. Evidence for genetic but not functional redundancy, *J Biol Chem* **274**, 25362–70 (1999).

50 Bennett EP, Hassan H, Hollingsworth MA, Clausen H, A novel human UDP-*N*-acetyl-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase, GalNAc-T7, with specificity for partial GalNAc-glycosylated acceptor substrates, *FEBS Lett* **460**, 226–30 (1999).

51 White KE, Lorenz B, Evans WE, Meitinger T, Strom TM, Econs MJ, Molecular cloning of a novel human UDP-Gal NAc:polypeptide *N*-acetylgalactosaminyltransferase, GalNAc-T8, and analysis as a candidate autosomal dominant hypophosphatemic rickets (ADHR) gene, *Gene* **246**, 347–56 (2000).

52 Toba S, Tenno M, Konishi M, Mikami T, Itoh N, Kurosaka A, Brain-specific expression of a novel human UDP-Gal NAc:polypeptide *N*-acetylgalactosaminyltransferase (GalNAc-T9), *Biochim Biophys Acta* **1493**, 264–8 (2000).

53 Cheng L-M, Tachibana K, Zhang Y, Guo J-M, Tachibana K, Kameyama A, Wang H, Hiruma T, Iwasaki H, Narimatsu H, Molecular Cloning and Characterization of a Human UDP-GalNAc: Polypeptide *N*-acetylgalactosaminyltransferase, pp-GalNAc-T10, *FEBS Lett* **531**, 115–21 (2002).

54 Schwientek T, Bennett EP, Flores C, Thacker J, Hollmann M, Reis CA, Behrens J, Mandel U, Keck B, Schafer MA, Haselmann K, Zubarev R, Roepstorff P, Burchell JM, Taylor-Papadimitriou J, Hollingsworth MA, Clausen H, Functional conservation of subfamilies of putative UDP-*N*-acetylgalactosamine:polypeptide *N*-acetylgalactosaminyltransferases in *Drosophila*, *Caenorhabditis elegans*, and mammals. One subfamily composed of l(2)35Aa is essential in Drosophila, *J Biol Chem* **277**, 22623–38 (2002).

55 Guo J-M, Zhang Y, Cheng L, Iwasaki H, Wang H, Kubota T, Tachibana K, Narimatsu H, Molecular cloning and characterization of a novel member of the UDP-GalNAc: Polypeptide *N*-acetylgalactosaminyltransferase family, pp-GalNAc-T12, *FEBS Lett* **524**, 211–8 (2002).

56 Zhang Y, Iwasaki H, Wang H, Kudo T, Kalka TB, Hennet T, Kubota T, Cheng L, Inaba N, Gotoh M, Togayachi A, Guo J, Hisatomi H, Nakajima K, Nishihara S, Nakamura M, Marth JD, Narimatsu H, Cloning and characterization of a new human UDP-*N*-Acetyl-?-D-galactosamine: Polypeptide *N*-acetylgalactosaminyltransferase, designated pp-GalNAc-T13, that is specifically expressed in neurons and synthesizes Tn antigen, *J Biol Chem* **278**, 573–84 (2003).

57 Wang H, Tachibana K, Zhang Y, Iwasaki H, Kameyama A, Cheng L, Guo J, Hiruma T, Togayachi A, Kudo T, Kikuchi N, Narimatsu H, Cloning and characterization of a novel UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase, pp-GalNAc-T14, *Biochem Biophys Res Commun* **300,** 738–44 (2003).

58 Cheng L, Tachibana K, Iwasaki H, Kameyama A, Zhang Y, Kubota T, Hiruma T, Tachibana K, Kudo T, Guo J, Narimatsu H, Characterization of a novel human UDP-GalNAc transferase, pp-GalNAc-T15, *FEBS Lett* **566**, 17–24 (2004).

59 Iwasaki H, Zhang Y, Tachibana K, Gotoh M, Kikuchi N, Kwon YD, Togayachi A, Kudo T, Kubota T, Narimatsu H, Initiation of O-glycan synthesis in IgA1 hinge region is determined by a single enzyme, UDP-*N*-acetyl-α-D-galactosamine: Polypeptide *N*-acetylgalactosaminyltransferase 2; pp-GalNAc-T2, *J Biol Chem* **278**, 5613–21 (2003).

60 Inamori K, Endo T, Ide Y, Fujii 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* **278**, 43102–9 (2003).

61 Inamori K, Endo T, Gu J, Matsuo I, Ito Y, Fujii S, Iwasaki H, Narimatsu H, Miyoshi E, Honke K, Taniguchi N, *N*-acetylglucosaminyltransferase IX acts on the GlcNAc β1,2-Man α1-Ser/Thr moiety, forming a 2,6-branched structure in brain O-mannosyl glycan, *J Biol Chem* **279**, 2337–40 (2004).

62 Kaneko M, Alvarez-Manilla G, Kamar M, Lee I, Lee JK, Troupe K, Zhang W, Osawa M, Pierce M, A novel β(1,6)-*N*-acetylglucosaminyltransferase V (GnT-VB)(1), *FEBS Lett* **554**, 515–9 (2003).

63 Inaba N, Hiruma T, Togayachi A, Iwasaki H, Wang XH, Furukawa Y, Sumi R, Kudo T, Fujimura K, Iwai T, Gotoh M, Nakamura M, Narimatsu H, A novel I-branching β-1,6-*N*-acetylglucosaminyltransferase involved in the human blood group I antigen expression, *Blood* **101**, 2870–6 (2003).

64 Ju T, Brewer K, D'Souza A, Cummings RD, Canfield WM, Cloning and expression of human core 1 β1,3-galactosyltransferase, *J Biol Chem* **277**, 178–86 (2002).

65 Kudo T, Iwai T, Iwasaki H, Hiruma T, Inaba N, Zhang Y, Gotoh M, Togayachi A, Narimatsu H, Molecular cloning and characterization of a novel UDP-Gal: GalNAc?-peptide β1,3-galactosyltransferase (C1Gal-T2), an enzyme synthesizing a core 1 structure of O-glycan, *J Biol Chem* **277**, 47724–31 (2002).

66 Ju T, Cummings RD, A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase, *Proc Natl Acad Sci USA* **99**, 16613–8 (2002).

67 Mochizuki H, Yoshida K, Gotoh M, Sugioka S, Kikuchi N, Kwon YD, Tawada A, Maeyama K, Inaba N, Hiruma T, Kimata K, Narimatsu H, Characterization of a heparan sulfate 3-*O*-sulfotransferase-5, an enzyme synthesizing a tetrasulfated disaccharide, *J Biol Chem* **278**, 26780–7 (2003).

68 Kamiyama S, Suda T, Ueda R, Suzuki M, Okubo R, Kikuchi N, Chiba Y, Goto S, Toyoda H, Saigo K, Watanabe M, Narimatsu H, Jigami Y, Nishihara S, Molecular cloning and identification of 3′-phosphoadenosine 5′-phosphosulfate transporter, *J Biol Chem* **278**, 25958–63 (2003).

69 Suda T, Kamiyama S, Suzuki M, Kikuchi N, Nakayama KI, Narimatsu H, Jigami Y, Aoki T, Nishihara S, Molecular cloning and characterization of a human multi-substrate specific nucleotide-sugar transporter homologous to Drosophila fringe connection, *J Biol Chem* (2004) (in press).

Received 3 June 2004; accepted 11 June 2004