Glycobiology vol. 6 no. 7 pp. 683–689, 1996

MINI REVIEW

Expression cloning of glycosyltransferases

Minoru Fukuda[1], Marti F.A.Bierhuizen[2] and Jun Nakayama[3]

Glycobiology Program, La Jolla Cancer Research Center, The Burnham Institute, La Jolla, CA 92037, USA

[1]To whom correspondence should be addressed
[2]Present address: Department of Hematology, Erasmus University, Rotterdam, The Netherlands
[3]Present address: Central Clinical Laboratories, Shinshu University Hospital, Matsumoto 390, Japan
This article is dedicated to the 65th birthday of Dr Robert L.Hill.

Cloning of cDNAs encoding glycosyltransferases enters into a new era because of the advent of expression cloning for glycosyltransferases. This mini-review summarizes a short historical view on the development of this method and surveys various improvements over the original method. This review also emphasizes the advantages of those recently developed improvements with a brief description on their background. Finally, the review provides some prospects based on these developments in the methodology. It is expected that increasing numbers of newly cloned glycosyltransferases will become critical tools for understanding the roles of carbohydrates.

Key words: glycosyltransferase/carbohydrate/cDNA

Introduction

Carbohydrates are not primary products of gene expression but are sequentially synthesized by glycosyltransferases. In order to understand the roles of carbohydrates and the regulatory mechanisms for their synthesis, it is thus essential to isolate genes (cDNAs) encoding those glycosyltransferases. This has not been an easy task, since the majority of glycosyltransferases are present in minute amounts. Because the Golgi apparatus is such an efficient subcellular functional unit, only a relatively small number of molecules are necessary to perform each reaction for the vast number of molecules that go through the Golgi complex. Such a low abundance of glycosyltransferases makes it harder to isolate a glycosyltransferase with enough quantity to derive the amino acid sequence. Because of this low abundance, only relatively abundant glycosyltransferases such as β-galactosyltransferase were cloned in the early stages of studies (Narimatsu *et al.*, 1986; Shaper *et al.*, 1986; Masri *et al.*, 1988). For a while, only α-2,6-sialyltransferase was cloned among sialyltransferases (Weinstein *et al.*, 1987), because the other sialyltransferases are present in a much lower abundance (Wen *et al.*, 1992).

This situation was dramatically changed after expression cloning was introduced to the field of glycobiology by John Lowe's group (Larsen *et al.*, 1989). It is now possible to clone a cDNA encoding a glycosyltransferase without the isolation of an enzyme, so long as a suitable reagent(s) recognizing a specific carbohydrate is provided. It is not an exaggeration to state that the introduction of expression cloning revolutionized the cloning strategy of glycosyltransferases. In this article, I describe a short historical aspect in the development of the technology for expression cloning of glycosyltransferases, and summarize the most improved methods, describing prospects in this field. For comprehensive surveys of glycosyltransferases cloned, readers are encouraged to read other reviews (Schachter, 1994; Field and Wainwright, 1995; Harduin-Lepers *et al.*, 1995).

Expression cloning—early development and basic techniques

Expression cloning of a desired gene was originally developed by Seed and Aruffo in 1987. Before this development, Berg and his colleagues constructed a vector that enabled the expression of a gene in mammalian cells (Mulligan and Berg, 1980). Moreover, it was established that the SV40 and polyoma virus large T antigens bind to the replication origin of the SV40 and polyoma virus; therefore, plasmids containing those replication origins are replicated more than several hundredfold in those cells (Francke and Eckhart, 1973; Gluzman, 1981). Such an enormous multiplication of plasmids can be achieved in mammalian cells into which plasmids are introduced. Such amplified expression enables isolation of a cDNA, even when it is present in minute amounts. Seed and Aruffo thus constructed a pCDM8 vector (Figure 1) that contained the replication origins for both the SV40 and polyoma virus (Seed and Aruffo, 1987; Aruffo and Seed, 1987). As a resistant marker, the vector contained *sup*F suppressor tRNA that suppresses amber mutations in ampicillin and tetracycline resistant genes in the presence of P3 episome (Little *et al.*, 1983). By using this vector, Seed's group has cloned many cDNAs that encode cell surface proteins. The basic procedure of expression cloning is as follows.

By using antibody specific to an antigen to be cloned, those cells that were transfected with a cDNA expression library are enriched by their reaction to the specific antibody. From those cells expressing a desired antigen, plasmids can be isolated using the Hirt procedure, which was originally developed for the isolation of viral DNA from infected mammalian cells (Hirt, 1967). Recovered plasmids are then divided into pools of plasmids, and each pool is tested for its capability to express the antigen. Once a particular pool of plasmids is identified to direct the expression of the desired gene, plasmids in that pool will be divided again into subpools containing a smaller number of plasmids. Each subpool is then tested for its ability to direct the expression of the desired gene. Further narrowing down of the plasmid pool eventually results in the isolation of a cDNA encoding the desired antigen. For the enrichment of transfected cells that express a desired antigen, the transfected cells are incubated first with the primary antibody recognizing

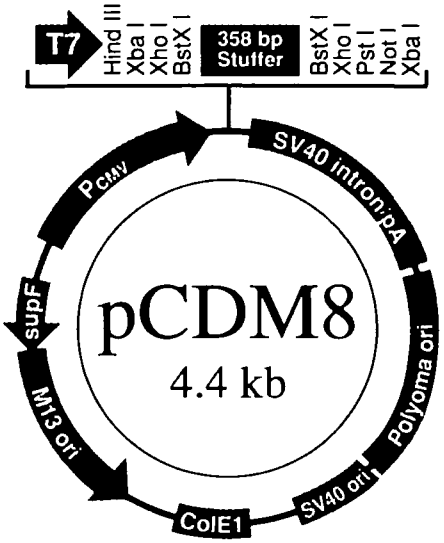

Fig. 1. Schematic structure of pCDM8. pCDM8 is based on enhancer-promoter sequences of the human cytomeglovirus (CMV) and contains both SV40 and polyoma virus origins. *Sup*F suppresser tRNA allows its replication in *Escherichia coli* strains that carry the P3 episome (for example, MC1061/P3), which allows the bacteria to grow in the presence of ampicillin and tetracycline. (Based on Seed and Aruffo, 1987.) This figure is taken from the Invitrogen catalog with permission

the antigen, and then overlaid on bacterial dishes coated with a secondary antibody ('panning') (Wysocki and Sato, 1978). From those cells attached to dishes, plasmids are recovered by the Hirt procedure. This method is very powerful in the majority of expression cloning. However, cell sorting by fluorescence-activated cell sorters (FACS) is now replacing the panning procedure, because of powerful cell sorters now available.

Expression cloning of glycosyltransferases—basic protocol

In order to isolate cDNAs encoding glycosyltransferases, the protocol for expression cloning needs to be slightly modified. Here I describe an example that utilized a specific antibody to Le^x, Galβ1→4(Fucα1→3)GlcNAcβ→R, to clone an α-1,3-fucosyltransferase (Kukowska-Latallo *et al.*, 1990).

For successful cloning of this enzyme, the following conditions need to be met. First, the recipient cells to which a cDNA library is introduced must lack carbohydrate products that are synthesized by a desired gene. Second, the recipient cells, however, need to express a precursor carbohydrate, that is, Galβ1→4GlcNAcβ1→R in the case for the cloning of α-1,3-fucosyltransferase. Third, the recipient cells need to express polyoma or SV40 large T antigen that enables introduced plasmids to replicate. For expression of α-1,3-fucosyltransferase, COS-1 cells met all these requirements. After transfection of a cDNA library constructed from A431 human epidermoid carcinoma cell line, the transfected cells were selected by anti-SSEA-1 antibody (Solter and Knowles, 1978), which reacts with Le^x structure. A431 cells were chosen for the source of mRNAs because they express different kinds of blood group antigens (Childs *et al.*, 1984). Plasmids were isolated from the transfected COS-1 cells which were attached to plates coated with the secondary antibody (panning). These enriched plasmids were again introduced to COS-1 cells and enrichment of COS-1 cells with anti-SSEA-1 antibody was repeated. After the third enrichment, plasmids were isolated from the transfected COS-1 cells that were enriched for SSEA-1 antigen, the plasmids were divided into pools of smaller numbers of plasmids ('sibling selection'), and eventually a single plasmid directing SSEA-1 expression was cloned. The cloned enzyme, named fucosyltransferase III, transfers a fucose residue to both Galβ1→4GlcNAc and Galβ1→3GlcNAc forming Le^x, Galβ1→4(Fuc α1→3) GlcNAc and Le^a, Galβ1→3 (Fucα1→4) GlcNAc. This dual specificity confirmed the previous discovery that the Le^a enzyme also transfers a fucose to Galβ1→4GlcNAc acceptor (Prieels *et al.*, 1981). This cloning had great significance since it resulted in the isolation of various α-1,3-fucosyltransferases, including fucosyltransferase VII, that form the E-selectin ligand sialyl Le^x in leukocytes (Natsuka *et al.*, 1994). This cloning was possible by cross-hybridization because all of these fucosyltransferases are highly homologous to each other.

Recipient cells expressing a carrier glycoprotein and polyoma large T antigen

Our laboratory has been studying the biosynthesis of *O*-glycans in leukocytes and discovered that the structures of *O*-glycans are characteristic to different cell lineages and different maturation stages of a given cell lineage. In particular, *O*-glycans attached to leukosialin, a mucin-like glycoprotein in T-lymphocytes, change dramatically during T-cell activation; resting T cells express tetrasaccharides NeuNAcα2→3Galβ1→3 (NeuNAcα2→6) GalNAc while activated T cells express hexasaccharides NeuNAc α2→3Galβ1→3 (NeuNAcα2→3Galβ1→4GlcNAcβ1→6) GalNAc (Piller *et al.*, 1988). This conversion is due to the new expression of Galβ1→3GalNAc (GlcNAc→GalNAc) β-1,6-N-acetylglucosaminyltransferase, core 2 GnT. Increased expression of core 2 GnT can be also observed in immunodeficiencies such as the Wiskott–Aldrich syndrome (Higgins *et al.*, 1991; Piller *et al.*, 1991), AIDS, and leukemia (Brockhausen, 1991; Saitoh *et al.*, 1991). Since *O*-glycans attached to leukosialin play roles such as adhesion of thymocytes to thymus epithelium (Baum *et al.*, 1995), we decided to clone core 2 GnT by expression cloning.

Initially, we used COS-1 cells as recipient cells since almost no core 2 GnT activity was detected in this cell line. By using T305 antibody as a selection marker, we isolated a cDNA that directs the expression of T305 antigen on COS-1 cells. However, the nucleotide sequence of the cloned cDNA indicated that we actually cloned cDNA encoding leukosialin (Bierhuizen *et al.*, 1992, 1994). From these early stages of the study, we learned three things. First, expression cloning is working in our hands. Second, COS-1 cells contain a small amount of core 2 GnT, although it may be difficult to detect. Third and most importantly, T305 preferentially binds to the hexasaccharides attached to leukosialin. This property of T305 antibody is similar to that of anti-M and -N antibodies which recognize both the peptide sequence of glycophorins and their attached oligosaccharides (Sadler *et al.*, 1987). Because of these findings, we chose Chinese hamster ovary (CHO) cells as recipient cells since CHO cells were shown to contain Galβ1→3GalNAc structure, but lack the hexasaccharide NeuNAcα2→3Galβ1→3 (NeuNAcα2→3Galβ1→4GlcNAcβ1→6) GalNAc (Sasaki *et al.*, 1987) (see Figure 2). CHO cells, however, do not express large T antigens that multiply plasmids. Since amplification of

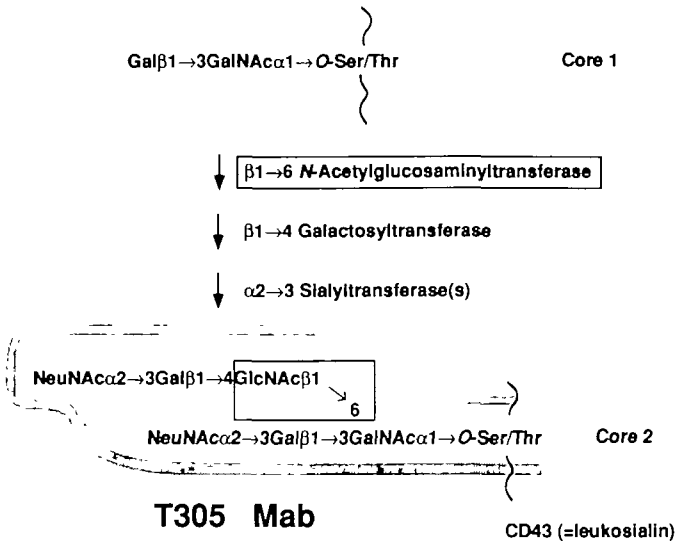

Fig. 2. Schematic representation of core 2 biosynthesis. Core 2 GnT adds β-1,6-linked N-acetylglucosamine to GalNAc in Galβ1→3GalNAc, core 1 structure. Following the addition of β-1,4-linked galactose and α-2,3-linked sialic acids, core 2–based hexasaccharide is formed. The hexasaccharides attached to leukosialin can be recognized by T305 antibody while the tetrasaccharide, NeuNAc α2→3Galβ1→3 (NeuNAcα2→6) GalNAc attached to leukosialin is not recognized by the same antibody (based on Bierhuizen *et al.*, 1992, 1994; Maemura and Fukuda, 1992).

plasmids are essential for successful transient expression cloning, we transfected CHO cells with a vector harboring polyoma large T cDNA (Figure 3). In parallel, we have made two more vectors, but the vector shown in Figure 3 was found to be the best. A very similar vector was developed in Dennis's laboratory (Heffernan and Dennis, 1991).

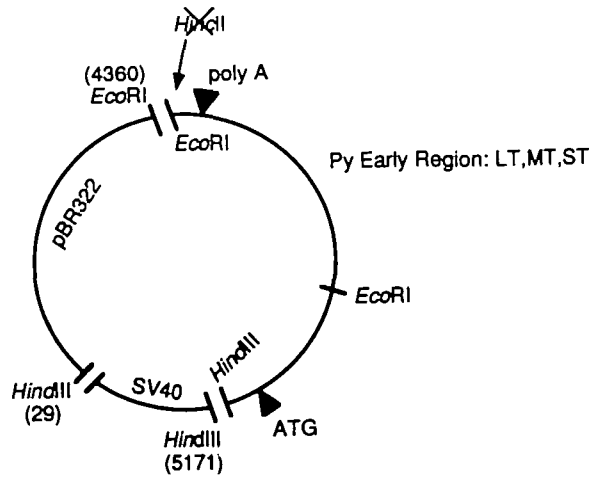

Fig. 3. Schematic representation of pPSVE1-PyE. Briefly, pPSVE1-PyE was prepared as follows. pPSVE1 was prepared by using pPSG4 DNA (ATCC no. 37337) and simian virus 40 viral DNA (Bethesda Research Laboratories). pPyCT-1DNA (ATCC no. 41043) was then digested with *Eco*RI/*Hin*cII to generate the COOH-terminal coding sequences for the polyoma large T antigen. *Eco*RI site at the 3′ end was introduced by blunt-end ligation of this *Eco*RI/*Hin*cII polyoma large T fragment with phosphorylated *Eco*RI linkers (Stratagene). Plasmid pPSVE1-PyE was then generated by ligating the COOH-terminal coding sequence for the T antigen into the unique *Eco*RI site of plasmid pPSVE1 (from Bierhuizen and Fukuda, 1992).

The reason behind the construction of the vector shown in Figure 3 is as follows. The promoter of polyoma large T itself has a genomic element that prevents the multiplication of polyoma large T antigen by feedback mechanisms (Zhu *et al.*, 1984). Because of this, the polyoma virus promoter was replaced by a SV40 early promoter.

Because T305 binds strongly with the hexasaccharides attached to leukosialin, we transfected CHO cells with both vectors that harbor leukosialin and polyoma large T antigen, resulting in CHO-Py•leu cells. By using these CHO cells which stably express leukosialin and polyoma large T antigen, the cDNA encoding core 2 GnT was cloned (Bierhuizen *et al.*, 1992).

It was also shown that CHO cells lack large I antigen, Galβ1→4GlcNAcβ1→3 (Galβ1→4GlcNAcβ1→6)Galβ1→R (Sasaki *et al.*, 1987; Smith *et al.*, 1990). The conversion of i to I antigen is due to the appearance of another β-1,6-N-acetylglucosaminyltransferase, IGnT. By using this CHO-Py•leu cell line and human anti-I antiserum, a cDNA encoding IGnT was cloned (Bierhuizen *et al.*, 1993). Although the expression of leukosialin was not necessary, the CHO-Py•leu cell line prepared had high efficient expression of polyoma large T, which made the cells suitable for expression cloning.

Transient expression of carrier molecules and large T antigens

An alternative method to the above described one is to transiently express a carrier molecule and/or large T antigen together with a cDNA library. When plasmids are recovered from transfected and sorted (or panned) cells, bacteria containing plasmids that are derived from the cDNA library will be selectively grown in the presence of ampicillin and tetracycline. This is possible when a cDNA library is constructed in a vector such as pCDM8 and pcDNAI since these vectors contain *sup*F suppressor tRNA that allows the bacteria to grow in the presence of ampicillin and tetracycline. Since other vectors usually have only an ampicillin resistant gene, cDNA harbored in those vectors will not be rescued.

This method has been used in several studies (Inoue *et al.*, 1993; Eckhardt *et al.*, 1995; Nakayama *et al.*, 1995). I will summarize here our strategy for the cloning of polysialyltransferase (PST). Polysialic acid is a developmentally regulated carbohydrate composed of a linear homopolymer of α-2,8–linked sialic acid residues. Polysialic acid (PSA) is mainly attached to N-CAM and is suggested to regulate the function of N-CAM. The polysialylated form of N-CAM is abundant in embryonic brain, while in adult brain, the majority of N-CAM lacks this unique glycan. The presence of this large negatively charged carbohydrate modulates the adhesive property of N-CAM, and the removal of PSA increases N-CAM–mediated binding between cells. Due to its presence in tissues undergoing synaptic rearrangement and cell migration, PSA is implicated in reducing N-CAM–mediated adhesion and thus perhaps allowing neurite outgrowth and cellular mobility (for reviews, see Edelman *et al.*, 1985; Rutishauser *et al.*, 1988).

Because of our experience in the cloning of core 2 GnT, it was obvious for us to use recipient cells that express N-CAM. First, COS-1 cells transfected with N-CAM cDNA were found to be negative for PSA despite the fact that they expressed N-CAM. The results enabled us to use COS-1 cells as recipient cells for cloning of PST. COS-1 cells were thus transiently cotransfected with pHβAPr-1-neo-N-CAM and a human fetal

brain cDNA library constructed in pcDNAI. After transfection, PSA-positive COS-1 cells were isolated by fluorescence-activated cell sorting (FACStar, Becton Dickinson) using antibody 735 (Frosch *et al.*, 1985). Plasmid DNA was isolated from the sorted cells using the Hirt procedure (Figure 4). Sibling selection of isolated plasmids resulted in the isolation of a single plasmid that directs the expression of PSA in COS-1 and HeLa cells (Nakayama *et al.*, 1995). During the early stages of this cloning, it was not possible to detect the expression of PSA despite the fact that the tested pool of plasmids contained PST unless N-CAM was expressed as well. When the cloned cDNA was introduced into COS-1 or HeLa cells, however, those transfected cells expressed PSA in the absence of N-CAM (Nakayama and Fukuda, 1996). It is thus evident that N-CAM is necessary for obtaining a detectable amount of PSA expression, when only a small amount of PST is expressed in transfected cells. It was thus critical to coexpress N-CAM to clone the enzyme.

Another critical development took place in expressing a large T antigen. After the studies on the expression of polyoma large T antigen was published, a vector for the expression of SV40 large T antigen was reported for transient expression cloning of cDNAs (de Chasseval and de Villartay, 1991). We have used this SV40 large T antigen vector in our laboratory and found that this particular vector, similar to that for polyoma large T antigen (see Figure 2), actually amplifies vectors such as pcDNAI, pcDNA3, and pCDM8 even in CHO cells.

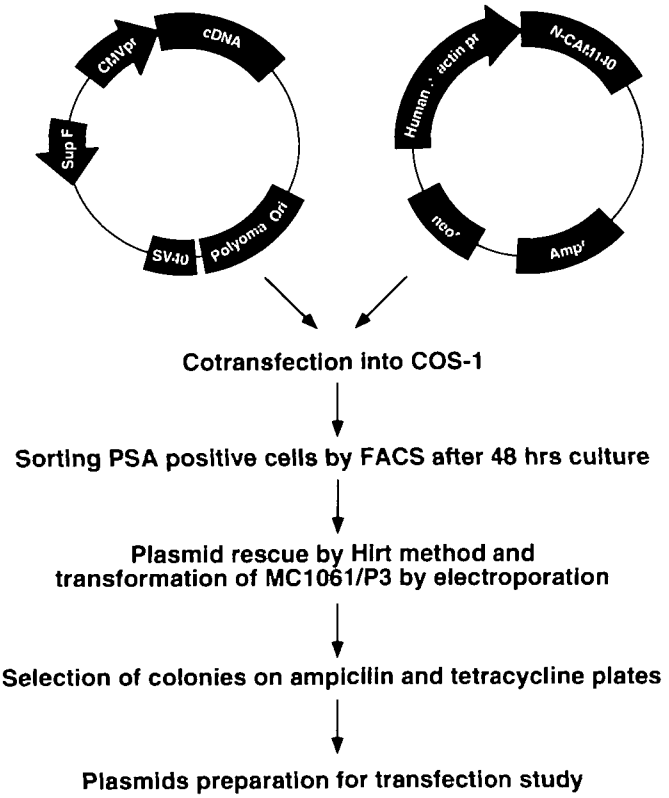

Fig. 4. Strategy for cloning of a cDNA encoding polysialyltransferase Two different vectors are used for cloning cDNA encoding polysialyltransferase, PST. Plasmid harboring N-CAM cDNA contains only ampicillin resistant gene. In contrast, plasmids harboring a human fetal brain cDNA library contain *Sup*F that allows those bacteria containing the plasmids to be rescued after ampicillin and tetracycline treatment (based on Nakayama *et al.*, 1995).

Although we have not tested many other cell lines, it is possible that this SV40 large T antigen may be capable of replicating vectors in various mammalian cells.

If that is the case, it may not be necessary to restrict the usage of SV40 and polyoma large T antigens to monkey or human cells and rodent cells, respectively. Further studies are necessary to determine whether or not the efficiency of polyoma and SV40 large T antigens in replicating plasmids depends on cell types where these antigens are expressed.

Cloning of a cDNA that encodes a desired glycosyltransferase using recipient cells expressing the same gene

During the expression cloning of GT3 synthase, we faced the problem that the recipient cells used for transient expression cloning expressed GT3 when the GD3 synthase was introduced to express the precursor glycolipid GD3. COS-1 cells lack GD3 synthase but a greater proportion of the transfected COS-1 cells, COS-1•GD3 express GT3 as well (Nakayama *et al.*, 1996). After cell sorting using anti-GT3 antibody (M6703), those cells expressing a minimum amount of GT3 were used as recipient cells for expression cloning. Plasmids were rescued from those COS-1 cells that were positive for GT3 staining using M6703 antibody.

When COS-1•GD3 cells were transiently transfected with a mixture of the above isolated plasmids, it was not possible with immunofluorescent staining to distinguish the cells that newly became GT3-positive from the cells that were endogenously GT3-positive. This failure was thought to be due to the high background expression of GT3 in COS-1•GD3 cells. In order to overcome this problem, the plasmid DNAs were transfected into HeLa cells that express a small amount of GD3 but no detectable amount of GT3. By using HeLa cells as indicator cells, it was possible to isolate a cDNA that encodes GT3 synthase (Nakayama *et al.*, 1996). Surprisingly, GT3 synthase was found to be identical to GD3 synthase (Haraguchi *et al.*, 1994; Nara *et al.*, 1994; Sasaki *et al.*, 1994b), therefore named GD3/GT3 synthase. It was no wonder that COS-1•GD3 cells also expressed GT3. It was also demonstrated that the cloned enzyme can synthesize GT3 much less efficiently than GD3, and only those cells expressing abundant GD3/GT3 synthase synthesize GT3 (Nakayama *et al.*, 1996).

This cloning taught us that a cDNA can be cloned even though the majority of recipient cells express the gene that encodes the desired protein. This success was probably due to two factors. First, only those transfected cells expressing substantial amounts of the antigen were collected for the source of the isolated plasmids. By setting up the sorting region where almost none of the COS-1•GD3 cells were present, only those cells that newly became positive were presumably isolated. Second, as long as indicator cells are available, a plasmid directing the expression of a desired gene can be isolated. This application should be expanded further by incorporating the SV40 large T antigen system shown above, when enriched plasmids are introduced into testing (indicator) cells.

Stable expression cloning using the Epstein–Barr virus system.

As an alternative to large T antigen–mediated replication of plasmids, the replication of the Epstein–Barr virus can be obtained by the nuclear antigen (EBNA-1) of the Epstein–Barr

virus. EBNA-1 binds to the replication origin of this virus and thus replicates a plasmid containing the replication origin (Margolskee *et al.*, 1988). In general, recipient cells for expression cloning are those that were stably transformed by the Epstein–Barr virus.

A great advantage in this system is that plasmids containing the replication origin of the Epstein–Barr virus can be replicated stably in an episomal state. Thus, the transfected cells can be continuously cultured for a long time and yet plasmids can be isolated from those transfected cells by the Hirt procedure. Because of this property, stably transfected cells can be subjected to cell sorting by FACS and those sorted cells can be cultured and again subjected to cell sorting. Since plasmids introduced into the recipient cells will not be lost after culturing for at least a few months, the enrichment of a plasmid that encodes a desired gene is possible after a continuous culture. Isolation of plasmids after each cell sorting can be thus avoided.

Sasaki *et al.* (1993) utilized this advantage in cloning ST3GalIV transferase using the lectin-resistant property of transfected cells. Transfected Namalwa KJM-1 were incubated with RCA_{120}, which binds to Galβ1→4GlcNAc, for 4–5 weeks in 96 wells. Plasmids isolated from those cells were tested to determine if they direct the resistance to RCA_{120}. By sibling selection, a single plasmid harboring α-2,3-sialyltransferase that adds α-2,3-linked sialic acid to Galβ1→4GlcNAc and Galβ1→3GlcNAc was cloned. In separate studies, the transfected Namalwa KJM-1 cells were sorted in FACS using anti-sialyl Le^x. By enriching the transfected cells expressing sialyl Le^x, fucosyltransferase VII that is responsible for sialyl Le^x formation in leukocytes was cloned (Sasaki *et al.*, 1994a). Similarly, repeated cell sorting of transfected Namalwa KJM-1 cells using anti-GD3 antibody led into the isolation of NeuNAcα2→3Galβ1→4Glc-Cer (NeuNAcα2→8NeuNAc) α-2,8-sialyltransferase (GD3 synthase) (Sasaki *et al.*, 1994b).

In these studies, B lymphocytes that have been transformed by the Epstein–Barr virus were used as recipient cells. Alternatively, cotransfection of a cDNA library and a vector encoding EBNA-1 should be feasible. If the expression of EBNA-1 in those transfected cells replicates the vector well, such cells can be used as recipient cells.

As an alternative and more traditional method, genomic fragments are introduced to recipient cells, and genomic fragments that direct the expression of a desired gene can be rescued. For this cloning, recipient cells need to have a mutation(s) that makes the gene inactive. Because CHO cells have a vast array of mutants that contain inactive glycosyltransferases or other carbohydrate processing enzymes (Stanley, 1984), these CHO cells are suited to clone these enzymes. For example, CHO cells were transfected with genomic DNA of HL-60 cells that contain α-1,3-fucosyltransferase. Stable transfectants were then cloned by limiting dilution, and those clones expressing Le^x antigen were selected. Those cell lines were also found to contain an *Alu* repeat sequence, which yielded 7.5 kb *Eco*RI fragment that hybridized to an *Alu* probe. *Alu* repeat sequences are genomic sequences of ~300 bp and inserted in human genome at high frequency (Jelinek *et al.*, 1980). Because this sequence is unique to human genomes, the presence of human genomic sequences in CHO genomes can be detected by hybridization of an *Alu* specific sequence to the total cellular genome sequences. Genomic DNA in the cloned CHO cells was digested by *Eco*RI, and DNAs of ~6–8 kb in size were isolated. A genomic library from those genomic sequences was prepared and screened for the presence of *Alu*-positive clones. Among two groups of *Alu*-positive clones, one group with an insert of ~6.3 kb was found to confer α-1,3-fucosyltransferase. After using various restriction enzymes, DNA fragments were tested for the expression of α-1,3-fucosyltransferase, yielding a single genomic fragment that directs the expression of α-1,3-fucosylated structure on the cell surface (Kumar *et al.*, 1991).

Theoretically, this method should be useful for many studies to isolate a genomic DNA that encodes a desired gene. However, the transient expression cloning or stable episomal expression cloning is much less laborious and also does not require additional factors such as the fortuitous presence of an *Alu* repeat. It is thus likely that these two new methods will entirely replace the classical genomic gene transfer method.

Prospects

As shown above, we now have many methods to clone a cDNA encoding a glycosyltransferase. However, three recently developed methods will be expected to play the most significant roles. First, it is essential to have suitable recipient cells for expression cloning. Occasionally, this has been a limiting factor since such cells may not be available in those cell lines so far established. In order to overcome this problem, mutant cell lines can be generated by using mutagens followed by selection using the antibodies detecting a desired product. To this cell line a cDNA library can be introduced together with a vector encoding polyoma large T or SV40 large T antigen. Since these large T antigens apparently replicate plasmids containing the replication origin in a wide variety of cells, a large number of mammalian cells can be used for recipient cells. In fact, hamster PST was cloned using a CHO mutant cell line that was generated exactly as described above. After complementation of those mutant cells, one particular cell line was found to be defective in PST, but not for other activities necessary for polysialic acid formation. By transfecting a CHO cDNA library and the vector encoding polyoma large T antigen (Bierhuizen and Fukuda, 1992) into this CHO mutant cell line, this defect was corrected and CHO cells expressing polysialic acid were sorted. From those sorted cells, a plasmid, which encodes the hamster PST, was eventually isolated (Eckhardt *et al.*, 1995). By a similar approach, hyaluronan synthase and ceramide glucosyltransferase were cloned using respective mutant cells which were generated for this purpose (Ichikawa *et al.*, 1996; Itano and Kimata, 1996). I believe that this type of expression cloning will become important.

The second critical method is the use of COS-1 or CHO-Py cells as recipient cells, but the plasmids isolated will be tested by other cells that express the precursor but not the product. This was successfully employed during the cloning of GT3 synthase as shown above (Nakayama *et al.*, 1996). This method also has an advantage in that the recipient cells do not need to lack the product structure for which the enzyme is sought.

The third critical method is to utilize stable episomal replication such as the Epstein–Barr virus system. In this system, continuous expression of introduced genes, which are episomal, provides an advantage that repeated selection of high expressors can be selected after only one transfection. So long as this episomal replication does not selectively amplify those vectors that contain only the replication origin, this method is extremely useful, as already shown (Sasaki *et al.*, 1993,

1994a,b). Although the recipient cells for this type of cloning appeared to restrict to those cells transformed by the Epstein–Barr virus, it is worth attempting to use nontransformed cells but stably expressing EBNA-1 as recipient cells.

Lastly, the most important new development took place in the usage of a functional assay in expression cloning. The first example of this was carried out by Goelz *et al.* (1990). In this study, a monoclonal antibody C_2E_5 was developed that inhibits the adhesion of HL-60 cells to E-selectin. This antibody was thus thought to bind to a ligand for E-selectin. By using the antibody C_2E_5, a cDNA that directs the expression of C_2E_5 antigen was cloned. The results indicate that this cDNA encodes fucosyltransferase IV. Since fucosyltransferase IV produces almost exclusively Le^x structure and a small amount of VIM-2 structure (Sueyoshi *et al.*, 1994), these results suggest that Le^x may also support E-selectin–mediated adhesion. Although it has not been supported by others, it was also reported that Le^x structure is a ligand for P-selectin (Larsen *et al.*, 1990).

More recently, the adhesion of transfected COS-1 cells to endometrial adenocarcinoma SNG-M cells was utilized to enrich plasmids encoding an adhesion molecule that apparently allows the attachment of trophoblasts to the endometrial epithelium. After the introduction of a cDNA library derived from trophoblastic teratocarcinoma, HT-H cells into COS-1 cells, single suspension cells of the transfected COS-1 cells were overlaid to a fixed monolayer of SNG-M cells. Plasmid DNAs were recovered from the twice-adhered COS-1 cells and subjected to the second round of screening. Plasmids obtained thereafter were then subjected to sibling selection, resulting in isolation of two cDNAs encoding trophinin and tastin. These two cDNAs need to be expressed together for efficient adhesion (Fukuda *et al.*, 1995). This study, carried out by Michiko Fukuda's laboratory, is monumental in that a functional assay was used to enrich plasmids encoding a desired gene. The work also provided an extraordinary finding that two independent, but mutually supportive genes can be cloned by expression cloning at the same time.

Just several years ago, only a handful of genes for glycosyltransferases had been isolated. The advent of expression cloning and its continuous improvement in methods will likely provide us with a possibility that all glycosyltransferases will be cloned in foreseeable years. We are thus entering an exciting era where we can obtain all necessary cDNAs that encode glycosyltransferases. Such success will undoubtedly enhance our understanding of the biological functions of carbohydrates.

Acknowledgements

We thank our colleagues and collaborators who played keyed roles in obtaining the data for the work described. We also thank Dr. Edgar Ong for critical reading of the manuscript and Susan Greaney for organizing the manuscript. The work in our laboratory has been supported by Grants R37CA33000, RO1CA33895, RO1CA48737, PO1CA71932, and a grant from the Mizutani Foundation for Glycoscience.

Abbreviations

FACS, fluorescence-activated cell sorters; CHO, Chinese hamster ovary; PST, polysialyltransferase; PSA, polysialic acid.

References

Aruffo,A. and Seed,B. (1987) Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system. *Proc. Natl Acad Sci. US A*, **84**, 8573–8577.

Baum,L.G., Pang,M., Perillo,N.L.,Wu, T., Delegeane,A., Uittenbogaart,C.H., Fukuda,M. and Seilhamer,J.J. (1995) Human thymic epithelial cells express an endogenous lectin, galectin-1, which binds to core 2 *O*-glycans on thymocytes and T lymphoblastoid cells. *J. Exp. Med.*, **181**, 877–887.

Bierhuizen,M.F.A. and Fukuda,M. (1992) Expression cloning of cDNA encoding UDP-GlcNAc: Galβ1–3GalNAc-R (GlcNAc to GalNAc) β-1,6-*N*-acetylglucosaminyltransferase by gene transfer into CHO cells expressing polyoma large T antigen. *Proc. Natl. Acad. Sci. USA*, **89**, 9326–9330.

Bierhuizen,M F.A., Mattei,M.-G. and Fukuda,M. (1993) Expression of the developmental I antigen by a cloned human cDNA encoding a member of a β-1,6-*N*-acetylglucosaminyltransferase gene family. *Genes Dev.* **7**, 468–478.

Bierhuizen,M.F., Maemura,K. and Fukuda,M. (1994) Expression of a differentiation antigen and poly-*N*-acetyllactosaminyl *O*-glycans directed by a cloned core 2 β-1,6-*N*-acetylglucosaminyltransferase. *J. Biol. Chem.*, **269**, 4473–4479

Brockhausen,I., Kuhns,W., Schachter,H., Matta,K.L., Sutherland,D.R. and Baker,M.A. (1991) Biosynthesis of *O*-glycans in leukocytes from normal donors and from patients with leukemia: increase in *O*-glycan core 2 UDP-GlcNAc. Galβ3GalNacα-R(GlcNAc to GalNAc)β(1–6)-*N*-acetylglucosaminyltransferase in leukemic cells *Cancer Res.*, **51**, 1257–1263.

Childs,R.A., Gregoriou,M., Scudder,P., Thorpe,S.J., Rees,A.R and Feizi,T. (1984) Blood group-active carbohydrate chains on the receptor for epidermal growth factor of A431 cells *EMBO J.*, **3**, 2227–2233.

de Chasseval,R. and de Villartay,J.-P. (1991) High level transient gene expression in human lymphoid cells by SV40 large T antigen boost. *Nucleic Acids Res.*, **20**, 245–250.

Eckhardt,M., Muhlenhoff,M., Bethe,A., Koopman,J., Frosch,M. and Gerardy-Schahn,R. (1995) Molecular characterization of eukaryotic polysialyltransferase-1. *Nature*, **373**, 715–718.

Edelman,G.M. (1985) Cell adhesion and the molecular processes of morphogenesis. *Annu. Rev. Biochem.*, **54**, 135–169.

Field,M.C. and Wainwright,L.J. (1995) Molecular cloning of eukaryotic glycoprotein and glycolipid glycosyltransferases: a survey. *Glycobiology*, **5**, 463–472 [published erratum appears in *Glycobiology*, **6**, 5 (1996)].

Francke,B and Eckhart,W. (1973) Polyoma gene function required for viral DNA synthesis. *Virology*, **55**, 127–135.

Frosch,M., Gorgen,I., Boulnois,G.J., Timmis,K.N. and Bitter-Suermann,D (1985) NZB mouse system for production of monoclonal antibodies to weak bacterial antigens: isolation of an IgG antibody to the polysaccharide capsules of *Escherichia coli* K1 and group B meningococci. *Proc. Natl. Acad. Sci. USA*, **82**, 1194–1198.

Fukuda,M.N., Sato,T., Nakayama,J., Klier,G., Mikami,M., Aoki,D. and Nozawa,S. (1995) Trophinin and tastin, a novel cell adhesion molecule complex with potential involvement in embryo implantation *Genes Dev.*, **9**, 1199–1210.

Gluzman,Y. (1981) SV40-transformed simian cells support the replication of early SV40 mutants. *Cell*, **23**, 175–182.

Goelz,S.E , Hession,C., Goff,D., Griffiths,B., Tizard,R., Newman,B., Chi-Rosso,G. and Lobb,R. (1990) ELFT: a gene that directs the expression of an ELAM-1 ligand. *Cell*, **63**, 1349–1356.

Haraguchi,M., Yamashiro,S., Yamamoto,A., Furukawa,K , Takamiya,K., Lloyd,K.O. and Shiku,H. (1994) Isolation of GD3 synthase gene by expression cloning of GM3 α-2,8-sialyltransferase cDNA using anti-GD2 monoclonal antibody. *Proc. Natl. Acad. Sci. USA*, **91**, 10455–10459.

Harduin-Lepers,A., Recchi,M.-A. and Delannoy,P (1995) 1994, the year of sialyltransferases. *Glycobiology*, **5**, 741–758.

Heffernan,M. and Dennis,J.W. (1991) Polyoma and hamster papovirus large T antigen-mediated replication of expression shuttle vectors in Chinese hamster ovary cells. *Nucleic Acids Res*, **19**, 85–92.

Higgins,E.A., Siminovitch,K A., Zhuang,D., Brockhausen,I and Dennis,J.W. (1991) Aberrant *O*-linked oligosaccharide biosynthesis in lymphocytes and platelets from patients with the Wiskott–Aldrich syndrome. *J. Biol. Chem.*, **266**, 6280–6290

Hirt,B. (1967) Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol Biol.*, **26**, 365–369

Ichikawa,S , Sakiyama,H., Suzuki,G., Hidari,K.I.-P.J. and Hirabayashi,Y. (1996) Expression cloning of a cDNA for human ceramide glucosyltransferase that catalyzes the first glycosylation step of glycosphingolipid synthesis. *Proc. Natl. Acad. Sci. USA*, **93**, 4638–4643.

Inoue,N., Kinoshita,T., Orii,T. and Takeda,J. (1993) Cloning of a human gene, PIG-F, a component of glycosylphosphatidylinositol anchor biosynthesis, by a novel expression cloning strategy. *J. Biol. Chem.*, **268**, 6882–6885.

Itano,N. and Kimata,K. (1996) Expression cloning and molecular characterization of HAS protein, a eukaryotic hyaluronan synthase. *J. Biol. Chem.*, **271**, 9875–9878.

Jelinek,W R , Toomey,T.P., Leinwand,L., Duncan,C.H., Biro,P.A., Choudary,P V., Weissman,S M., Rubin,C M., Houck,C M., Deininger,P.L. and Schmid,C.W. (1980) Ubiquitous, interspersed repeated sequences in mammalian genomes. *Proc. Natl. Acad. Sci. USA*, **77**, 1398–1402.

Kukowska-Latallo,J.F., Larsen,R.D., Nair,R.P. and Lowe,J.B. (1990) A cloned human cDNA determines expression of a mouse stage-specific embryonic antigen and the Lewis blood group α(1,3/1,4) fucosyltransferase. *Genes Dev.*, **4**, 1288–1303.

Kumar,R., Potvin,B., Muller,W.A. and Stanley,P. (1991) Cloning of a human α1,3-fucosyltransferase gene that encodes ELFT but does not confer ELAM-1 recognition on Chinese hamster ovary cell transfectants. *J. Biol. Chem.*, **266**, 21777–21783.

Larsen,E., Palabrica,T., Sajer,S., Gilbert,G.E., Wagner,D.D., Furie,B.C. and Furie,B. (1990) PADGEM-dependent adhesion of platelets to monocytes and neutrophils is mediated by a lineage-specific carbohydrate, LNF III (CD15). *Cell*, **63**, 467–474.

Larsen,R.D., Rajan,V.P., Ruff,M.M., Kukowska-Latallo,J., Cummings,R.D. and Lowe,J.B. (1989) Isolation of a cDNA encoding a murine UDP-galactose:β-D-galactosyl-1,4-*N*-acetyl-D-glucosaminide α-1,3-galactosyltransferase: expression cloning by gene transfer. *Proc Natl. Acad. Sci. USA*, **86**, 8227–8231.

Little,P.F., Treisman,R , Bierut,L., Seed,B. and Maniatis,T (1983) Plasmid vectors for the rapid isolation and transcriptional analysis of human β-globin gene alleles. *Mol. Biol. Med*, **1**, 473–488.

Maemura,K. and Fukuda,M. (1992) Poly-*N*-acetyllactosaminyl *O*-glycans attached to leukosialin. The presence of sialyl Le^x structures in *O*-glycans. *J. Biol. Chem.*, **267**, 24379–24386.

Margolskee,R F., Kavathas,P. and Berg,P. (1988) Epstein–Barr virus shuttle vector for stable episomal replication of cDNA expression libraries in human cells. *Mol. Cell Biol.*, **8**, 2837–2847.

Masri,K.A., Appert,H.E and Fukuda,M N. (1988) Identification of the full-length coding sequence for human galactosyltransferase (β-*N*-acetylglucosaminide: β 1,4-galactosyltransferase). *Biochem. Biophys Res. Commun.*, **157**, 657–663.

Muller,W.J., Naujokas,M.A and Hassell,J.A. (1984) Isolation of large T antigen–producing mouse cell lines capable of supporting replication of polyomavirus-plasmid recombinants *Mol. Cell. Biol.*, **4**, 2406–2412.

Mulligan,R.C. and Berg,P. (1980) Expression of a bacterial gene in mammalian cells. *Science*, **209**, 1422–1427.

Nakayama,J. and Fukuda,M. (1996) A human polysialyltransferase directs *in vitro* synthesis of polysialic acid. *J. Biol. Chem.*, **271**, 1829–1832.

Nakayama,J., Fukuda,M.N., Fredette,B., Ranscht,B. and Fukuda,M. (1995) Expression cloning of a human polysialyltransferase that forms the polysialylated neural cell adhesion molecule present in embryonic brain. *Proc. Natl. Acad. Sci. USA*, **92**, 7031–7035

Nakayama,J., Fukuda,M.N., Hirabayashi,Y., Kanamori,A., Sasaki,K., Nishi,T. and Fukuda,M. (1996) Expression cloning of a human GT3 synthase: GD3 and GT3 are synthesized by a single enzyme. *J. Biol Chem.*, **271**, 3684–3691.

Nara,K., Watanabe,Y., Maruyama,K , Kasahara,K., Nagai,Y. and Sanai,Y. (1994) Expression cloning of a CMP-NeuAc:NeuAc α 2–3Gal β 1–4Glc β 1–1′Cer α 2,8-sialyltransferase (GD3 synthase) from human melanoma cells. *Proc Natl. Acad. Sci. USA*, **91**, 7952–7956.

Narimatsu,H., Sinha,S., Brew,K , Okayama,H. and Qasba,P K (1986) Cloning and sequencing of cDNA of bovine N-acetylglucosamine (β-1–4)galactosyltransferase. *Proc. Natl. Acad. Sci. USA*, **83**, 4720–4724

Natsuka,S , Gersten,K.M., Zenita,K., Kannagi,R. and Lowe,J.B. (1994) Molecular cloning of a cDNA encoding a novel human leukocyte α-1,3-fucosyltransferase capable of synthesizing the sialyl Lewis x determinant *J Biol. Chem.* **269**, 16789–16794 [published erratum appears in *J. Biol. Chem.*, **269**, 20806 (1994)].

Piller,F., Piller,V , Fox,R.I. and Fukuda,M. (1988) Human T-lymphocyte activation is associated with changes in *O*-glycan biosynthesis *J. Biol. Chem.*, **263**, 15146–15150.

Piller,F., Le Deist,F., Weinberg,K., Parkman,R. and Fukuda,M. (1991) Altered *O*-glycan synthesis in lymphocytes from patients with Wiskott–Aldrich syndrome. *J. Exp. Med.*, **173**, 1501–1510.

Prieels,J.P., Monnom,D., Dolmans,M., Beyer,T A. and Hill,R.L. (1981) Copurification of the Lewis blood group N-acetylglucosaminide α 1 goes to 4 fucosyltransferase and an N-acetylglucosaminide α 1 goes to 3 fucosyltransferase from human milk. *J. Biol. Chem.*, **256**, 10456–10463.

Rutishauser,U., Acheson,A., Hall,A.K., Mann,D.M. and Sunshine,J. (1988) The neural cell adhesion molecule (NCAM) as a regulator of cell–cell interactions. *Science*, **240**, 53–57.

Sadler,J.E., Paulson,J.C. and Hill,R.L (1979) The role of sialic acid in the expression of human MN blood group antigens. *J. Biol. Chem.*, **254**, 2112–2119.

Saitoh,O., Piller,F., Fox,R.I. and Fukuda,M. (1991) T-Lymphocytic leukemia express complex, branched *O*-linked oligosaccharides on a major sialoglycoprotein, leukosialin. *Blood*, **77**, 1491–1499.

Sasaki,H., Bothner,B., Dell,A. and Fukuda,M. (1987) Carbohydrate structure of erythropoietin expressed in Chinese hamster ovary cells by a human erythropoietin cDNA. *J. Biol. Chem.*, **262**, 12059–12076.

Sasaki,K., Watanabe,E., Kawashima,K., Sekine,S., Dohi,T., Oshima,M., Hanai,N., Nishi,T. and Hasegara,M. (1993) Expression cloning of a novel Galβ (1–3/1–4) GlcNAc α2,3-sialyltransferase using lectin resistance selection. *J. Biol. Chem.*, **268**, 22782–22787.

Sasaki,K., Kurata,K., Funayama,K , Nagata,M., Watanabe,E., Ohta,S., Hanai,N. and Nishi,T. (1994a) Expression cloning of a novel α 1,3-fucosyltransferase that is involved in biosynthesis of the sialyl $Lewis^x$ carbohydrate determinants in leukocytes. *J. Biol. Chem.*, **269**, 14730–14737.

Sasaki,K., Kurata,K., Kojima,N., Kurosawa,N., Ohta,S., Hanai,N., Tsuji,S. and Nishi,T. (1994b) Expression cloning of a G_{M3}-specific α-2,8-sialyltransferase (G_{D3} synthase) *J. Biol. Chem.*, **269**, 15950–15956.

Schachter,H. (1994) Molecular cloning of glycosyltransferase genes In Fukuda,M. and Hindsgaul,O. (eds), *Molecular Glycobiology*. IRL Press, Oxford, pp 88–162, 258–259.

Seed,B and Aruffo,A. (1987) Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. *Proc. Natl. Acad. Sci. USA*, **84**, 3365–3369.

Shaper,N.L., Shaper,J.H., Meuth,J.L., Fox,J.L., Chang,H., Kirsch,I.R. and Hollis,G.F. (1986) Bovine galactosyltransferase: identification of a clone by direct immunological screening of a cDNA expression library. *Proc. Natl Acad. Sci USA*, **83**, 1573–1577.

Smith,D.F., Larsen,R D., Mattox,S., Lowe,J.B and Cummings,R.D. (1990) Transfer and expression of a murine UDP-Gal:β-D-Gal-α 1,3-galactosyltransferase gene in transfected Chinese hamster ovary cells. Competition reactions between the α 1,3-galactosyltransferase and the endogenous α 2,3-sialyltransferase. *J. Biol. Chem.*, **265**, 6225–6234.

Solter,D. and Knowles,B.B. (1978) Monoclonal antibody defining a stage-specific mouse embryonic antigen (SSEA-1). *Proc. Natl Acad. Sci USA*, **75**, 5565–5569

Stanley,P. (1984) Glycosylation mutants of animal cells. *Annu Rev. Genet.*, **18**, 525–552

Sueyoshi,S., Tsuboi,S., Sawada-Hirai,R., Dang,U.N., Lowe,J.B. and Fukuda,M. (1994) Expression of distinct fucosylated oligosaccharides and carbohydrate-mediated adhesion efficiency directed by two different α-1,3-fucosyltransferases. Comparison of E- and L-selectin-mediated adhesion. *J Biol. Chem.*, **269**, 32342–32350.

Weinstein,J., Lee,E.U., McEntee,K., Lai,P H and Paulson,J.C. (1987) Primary structure of β-galactoside α-2,6-sialyltransferase. Conversion of membrane-bound enzymes to soluble forms by cleavage of the NH2 terminal signal anchor. *J. Biol. Chem.*, **262**, 17735–17743.

Wen,D.X., Livingston,B.D., Medzihradszky,K.F., Kelm,S., Burlingame,A.L and Paulson,J.C. (1992) Primary structure of Galβ-1,3(4)GlcNAc α-2,3-sialyltransferase determined by mass spectrometry sequence analysis and molecular cloning. Evidence for a protein motif in the sialyltransferase gene family. *J. Biol. Chem.*, **267**, 21011–21019.

Wysocki,L.J. and Sato,V.L. (1978) Panning for lymphocytes a method for cell selection. *Proc. Natl. Acad Sci. USA*, **75**, 2844–2848.

Zhu,Z.Y., Veldman,G.M., Cowie,A., Carr,A., Schaffhausen,B. and Kamen,R. (1984) Construction and functional characterization of polyomavirus genomes that separately encode the three early proteins. *J. Virol.*, **51**, 170–180.

Received on April 10, 1996, accepted on July 28, 1996