SHORT COMMUNICATION

Toru Furukawa · Emile M. Youssef · Toshimasa Yatsuoka
Tadaaki Yokoyama · Naohiko Makino · Hiroko Inoue
Shinichi Fukushige · Masato Hoshi · Yutaka Hayashi
Makoto Sunamura · Akira Horii

Cloning and characterization of the human UDP-*N*-acetylglucosamine: α-1,3-D-mannoside β-1,4-*N*-acetylglucosaminyltransferase IV-homologue (*hGnT-IV-H*) gene

Abstract A novel human gene determining a polypeptide product of 478 residues with an estimated molecular mass of 55 kDa, which has significant homology and structural similarity to Bos UDP-*N*-acetylglucosamine: α-1,3-D-mannoside β-1,4-*N*-acetylglucosaminyltransferase (GnT-IV), was cloned from the commonly deleted region in pancreatic cancer at 12q21. The gene is composed of at least six exons, and the last three exons cover the open reading frame. Different sized transcripts, 3.8-kb in the heart, brain, and fetal brain and 2.8-kb and 1.7-kb in the testis were observed by Northern blot analysis. By reverse transcription-polymerase chain reaction, expression was also observed in the adult brain, liver, and adrenal gland, but not in pancreas. Because of its significant homology and structural similarity to Bos GnT-IV, it is potentially the gene encoding human GnT-IV or its homologue, which had been one of two genes remaining to be cloned in the human GnT family.

Key words N-acetylglucosaminyltransferase · Fetal brain · Adult brain · Heart · Testis · Chromosomal band 12q21

Introduction

Chromosome band 12q21 is a locus implicated in cancers of the pancreas and stomach (Fey et al. 1980; Sano et al. 1991; Kimura et al. 1996; Fukushige et al. 1997; Kimura et al. 1998). We previously identified a commonly deleted region in pancreatic cancer at 12q21 between D12S81 and D12S1719 (Kimura et al. 1998).

Methods and results

In the search for a candidate tumor suppressor gene, the deleted region was covered with bacterial artificial chromosome (BAC) clones and cosmid clones subcloned from yeast artificial chromosome (YAC) clones. Then shotgun sequencing was carried out to determine the nucleotide (nt) sequences of the deleted region. Several expressed sequence tags (ESTs) were found to be located in this area. One of the ESTs, AA385488 (GenBank; http://www.ncbi.nlm.nih.gov/Web/Genbank), was located in the deleted region. A paired primer set, primers 5′-GGAAACATCCACACATCAACTG-3′ and 5′-TGCTAGGTGAACCACCACTG-3′, was used to amplify a 257-bp fragment corresponding to a partial sequence of the EST by reverse transcription-polymerase chain reaction (RT-PCR) according to methods described previously (Mori et al. 1997). Using the PCR product as a probe, we isolated 15 independent cDNA clones from a fetal brain cDNA library (Clontech, Palo Alto, CA, USA). After analysis of nucleotide sequences determined with an ABI prism BigDye Terminator Cycle Sequencing Ready Reaction FS Kit and an ABI prism 377 DNA sequencer (Perkin-Elmer Applied Biosystems, Foster City, CA, USA), two types of sequences were obtained. These were generated by alternative splicing mechanisms (see below).

The first sequence (type 1 transcript, accession no. AB024729) was composed of 2655-nt with open reading frame encoding for a polypeptide of 478 amino acid residues with an estimated molecular mass of 55 kDa. The second sequence (type 2 transcript, accession no. AB024730) lacked a sequence consisting of 121-nt within the 5′ untranslated region of the first sequence. In-frame termination codons were found upstream of the predicted initiation codon. The homology search for the predicted protein utilized the National Center of Biotechnology Information

T. Furukawa · E.M. Youssef · T. Yatsuoka · T. Yokoyama · N. Makino · H. Inoue · S. Fukushige · M. Hoshi · A. Horii (✉)
Department of Molecular Pathology, Tohoku University School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai 980-8575, Japan
Tel. +81-22-717-8042; Fax +81-22-717-8047
e-mail: horii@mail.cc.tohoku.ac.jp

T. Yatsuoka · T. Yokoyama · H. Inoue · M. Sunamura
First Department of Surgery, Tohoku University School of Medicine, Sendai, Japan

M. Hoshi · Y. Hayashi
Department of Pediatric Surgery, Tohoku University School of Medicine, Sendai, Japan

(NCBI) position specific iterated basic local alignment search tool (PSI-BLAST) (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-psi_blast). We found that the predicted product had high homology with Bos UDP-*N*-acetylglucosamine: α-1,3-D-mannoside β-1,4-*N*-acetylglucosaminyltransferase (GnT-IV) (accession no. AB000628). The identity of the predicted amino acid sequence of the gene to bovine GnT-IV protein (Bos GnT-IV) was 27%, and the similarity was 45% (see Fig. 1). It contained a highly hydrophobic region close to the N-terminus between Val-26 and Tyr-42, followed by a hydrophilic region between Ser-44 and Ser-77; this could be a transmembrane domain and a stem region. A similar structure including a large hydrophobic region with a subsequent hydrophilic region is observed in the N-terminal region between Val-8 and Tyr-25 of Bos GnT-IV. This structure corresponds to type II transmembrane protein, which is commonly observed in other human GnTs (see below). Because of its significant homology to Bos GnT-IV, the cDNA we isolated is thought to have been derived from the gene encoding human GnT-IV or its homologue. We named this gene human GnT-IV homologue (*hGnT-IV-H*). It is notable that the predicted protein product of the *hGnT-IV-H* gene lacked a proline-rich sequence in the predicted stem region (similar to Bos GnT-IV), which exists in other *N*-acetylglucosaminyltransferases (GnTs).

GnT-IV is an essential enzyme for the synthesis of the multiantennary structure of asparagine (Asn)-linked oligosaccharides, which are mainly contained in erythropoietin and α1 acid glycoprotein (Gleeson and Schachter 1983; Minowa et al. 1998). Together with GnT-V, this protein contributes to phenotypic changes in neoplastic transformed cells by inducing a structural change in cell surface carbohydrates, known as increased sialylation, and by β1-6-linked branching of complex-type Asn-linked oligosaccharides (Schachter, 1986). This structural change is directly associated with the metastatic potential of neoplastic transformed cells (Dennis et al. 1987). There are six different human GnTs known so far; GnT-I (UDP-*N*-acetylglucosamine: α-1,3-D-mannoside β-1,2-*N*-acetylglucosa-

Fig. 1. Protein alignment between the human GnT-IV-homologue (*hGnT-IV-H*) and the bovine GnT-IV (*Bos GnT-IV*). Identical residues are indicated by *bold letters with gray boxes* and conservative residues by *plain letters with open boxes*

Fig. 2. Primary structure of *hGnT-IV-H*. Exons are indicated by *boxes with their sizes in nucleotide length. Open boxes*, Nontranslated region; *closed boxes*, coding region. Sizes of the introns were not determined. Codon numbers are indicated *below the boxes*. Splicing variation is also shown. *ATG*, Translation initiation codon; *TAG*, Translation termination codon
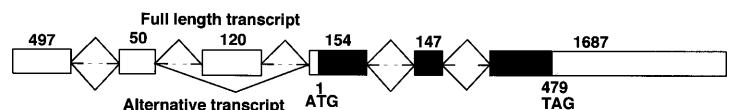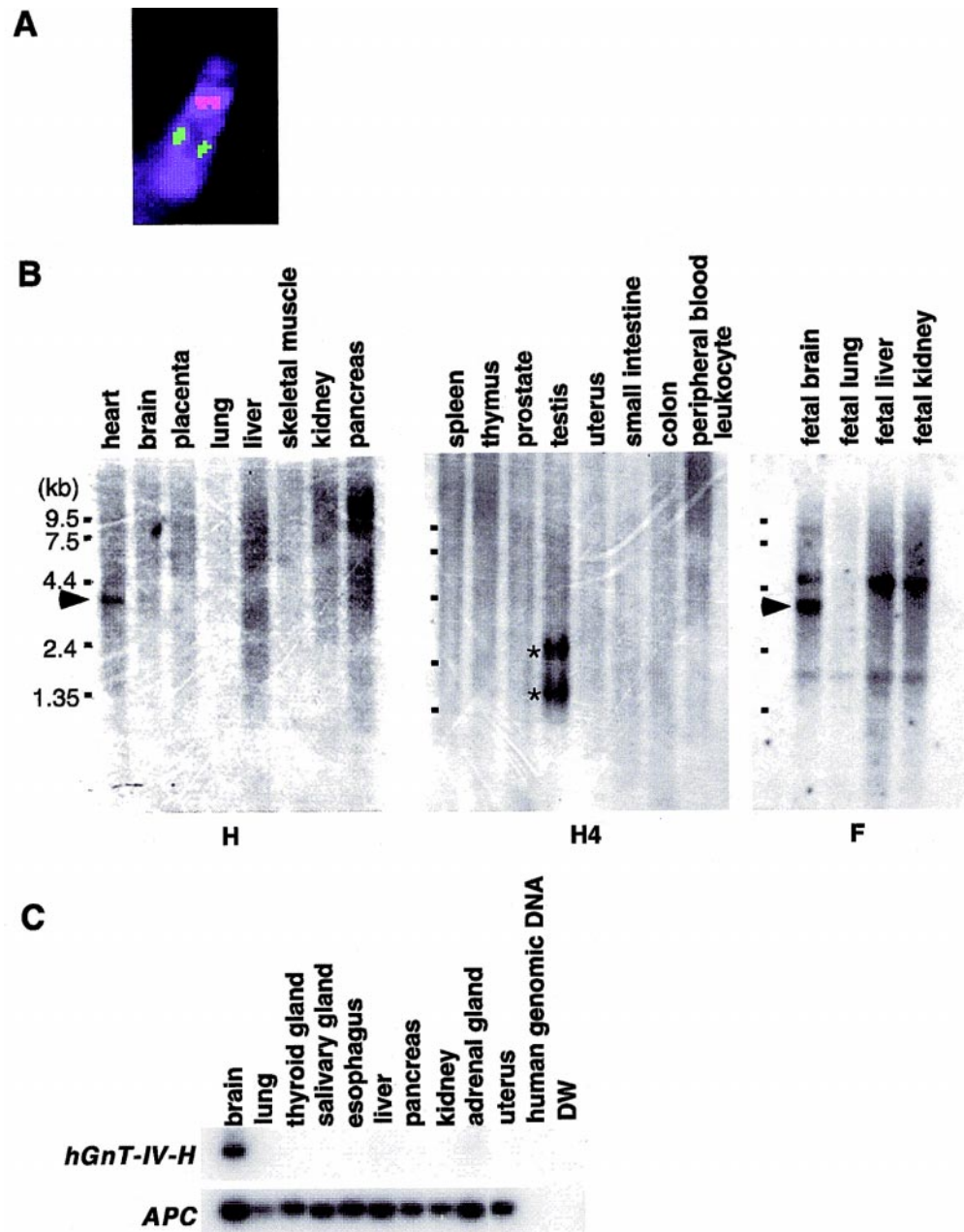
Fig. 3. A Dual-color fluorescence in situ hybridization (FISH) analysis of *hGnT-IV-H*. The *red signal* indicates the centromere of chromosome 12 (pα12H8) (Looijenga et al. 1990). The *green signals* indicate the human *GnT-IV-H* (BAC759H8), which was located at 12q21. **B** Northern blot analysis of *hGnT-IV-H*. A radiolabelled cDNA fragment containing the C-terminal coding region and 3′ untranslated region (UTR) was used for hybridization of Human MTN Blot (*H*) and MTN Blot IV (*H4*), and a cDNA fragment harboring the entire coding region was used for hybridization of the Fetal MTN Blot II (*F*) (Clontech for all Blots). Different sized transcripts, 3.8-kb in the heart, brain, and the fetal brain (indicated by *arrowheads*), and 2.8-kb and 1.7-kb in the testis (indicated by *asterisks*) were observed. Nonspecific hybridizations of 28S and 18S ribosomal RNAs were observed in the blot of fetal tissues. **C** Reverse

transcription-polymerase chain reaction (RT-PCR) for *hGnT-IV-H* and the adenomatous polyposis coli (*APC*) gene in various normal adult tissues, including brain, lung, thyroid gland, salivary gland, esophagus, liver, pancreas, kidney, adrenal gland, and uterus. Primers for *hGnT-IV-H* were 5′-TGATTCTGCGTGAGAAAGGC-3′ in exon 1 and 5′-GAACGTTTTCTCAGGCATCTC-3′ in exon 4. Results of Southern hybridization using a radiolabelled oligonucleotide, 5′-AAGACGCTGTCATAATCTGTGC-3′ in exon 2 are shown. RT-PCR for the *APC* gene as a control was performed using primers 5′-TGGGTTCTTCTGGACAGATT-3′ and 5′-TCAGCAAGAAGCAATGACCT-3′. Southern hybridization using a radiolabelled oligonucleotide 5′-GGACTGCACTCTCCAGAAC-3′ is shown

minyltransferase), GnT-II (UDP-*N*-acetylglucosamine: α-1,6-D-mannoside β-1,2-*N*-acetylglucosaminyltransferase), GnT-III (UDP-*N*-acetylglucosamine: β-D-mannoside β-1,4-*N*-acetylglucosaminyltransferase) and GnT-V (UDP-*N*-acetylglucosamine: α-1,6-D-mannoside β-1,6-*N*-acetylglucosaminyltransferase V) have been cloned and characterized. GnT-I is an essential molecule for mouse development because its knockout is lethal (Ioffe and

Stanley, 1994). GnT-II is the putative gene for carbohydrate-deficient glycoprotein syndrome type II (Tan et al. 1996). GnT-III is assumed to be related to the promotion and metastasis of certain tumors, including hepatic tumor and mouse melanoma (Yoshimura et al. 1995; Bhaumik et al. 1998). Thus, GnTs are important enzymes in regard to both physiology and pathology. In the human GnTs family, GnT-IV and GnT-VI remain to be cloned and characterized.

To determine the genomic structure of *hGnT-IV-H*, the BAC clone 759H8 (Research Genetics, Huntsville, AL, USA) harboring the EST fragment was sequenced. The gene was composed of at least six exons, and exon 4 through exon 6 harbored an open reading frame (see Fig. 2). Sequences of exon-intron junctions were deposited in the database (accession nos. AB024906 through AB024911). After comparison of the genomic and cDNA sequences, we found that the type 2 transcript lacked the entire exon 3, by alternative splicing. The BAC clone also contained a microsatellite marker, D12S88, indicating that this gene is very close to this marker. Localization of the gene at 12q21 was confirmed by fluorescence in situ hybridization (FISH) (see Fig. 3A) according to a method described previously (Fukushige et al. 1997).

Expression of *hGnT-IV-H* was analyzed by Northern blot and RT-PCR. Human Multiple Tissue Northern (MTN) Blot, Human MTN Blot IV, and Human Fetal MTN Blot II (Clontech) were hybridized according to the manufacturer's instructions with a ^{32}P-labelled cDNA fragment containing the C-terminal coding region and the 3′ untranslated region (UTR) for the former two blots, and a cDNA fragment harboring the entire coding region for the fetal blot. A 3.8-kb transcript was detected in the heart, brain, and fetal brain tissues. Two different sized transcripts, 2.8-kb and 1.7-kb in size, were detected in the testis (see Fig. 3B). To study the expression of the two types of isolated transcripts, we carried out RT-PCR for cDNAs generated from the mRNAs of various normal adult tissues, including brain, lung, thyroid gland, salivary gland, esophagus, liver, pancreas, kidney, adrenal gland, and uterus, according to a method described previously (Mori et al. 1997). A primer pair localizing in exons 1 and 4 was used. By Southern hybridization of the RT-PCR products, using an oligonucleotide in exon 2, we observed relatively abundant transcripts in adult brain, whereas only trace amounts of transcripts were detected in liver and adrenal gland (see Fig. 3C). These transcripts corresponded to the type 2 transcript, which produced a 167-bp product.

Discussion

Because the *hGnT-IV-H* gene was located at a frequently deleted locus in pancreatic cancer, there was a possibility that this gene had a role in pancreatic carcinogenesis. However, expression of *hGnT-IV-H* was not detected in pancreatic tissue by Northern blot or RT-PCR, as shown in Fig. 3B,C. RT-PCR with a different primer pair located in exons 5 and 6 within the coding region also failed to detect expression of the gene in the pancreas (data not shown). Hence, it is very likely that the *hGnT-IV-H* gene does not play any role in pancreatic carcinogenesis. We also analyzed 19 pancreatic cancer cell lines, but no expression was detected in these specimens by RT-PCR.

We have cloned a novel human gene located in the region at 12q21 deleted in pancreatic cancer. The gene encoded a polypeptide of 478 residues with an estimated molecular mass of 55 kDa. The encoded polypeptide product showed a significant homology with and structural similarities to Bos GnT-IV and is therefore the gene potentially encoding human GnT-IV or its homologue. Its biological and pathological roles should be clarified in future investigations.

The nucleotide sequence data reported in this manuscript will appear in the DDBJ, EMBLE, and GenBank nucleotide sequence databases with the following accession numbers: AB024729, AB024730, and AB024906 through AB024911.

References

Bhaumik M, Harris T, Sundaram S, Johnson L, Guttenplan J, Rogler C, Stanley P (1998) Progression of hepatic neoplasms is severely retarded in mice lacking the bisecting N-acetylglucosamine on N-glycans: evidence for a glycoprotein factor that facilitates hepatic tumor progression. Cancer Res 58:2881–2887

Dennis JW, Laferte S, Waghorne C, Breitman ML, Kerbel RS (1987) Beta 1–6 branching of Asn-linked oligosaccharides is directly associated with metastasis. Science 236:582–585

Fey MF, Hesketh C, Wainscoat JS, Gendler S, Thein SL (1989) Clonal allele loss in gastrointestinal cancers. Br J Cancer 59:750–754

Fukushige S, Waldman F, Kimura M, Abe T, Furukawa T, Sunamura M, Kobari M, Horii A (1997) Frequent gain of copy number on the long arm of chromosome 20 in human pancreatic adenocarcinoma. Genes Chromosom Cancer 19:161–169

Gleeson PA, Schachter H (1983) Control of glycoprotein synthesis. J Biol Chem 258:6162–6173

Ioffe E, Stanley P (1994) Mice lacking N-acetylglucosaminyl-transferase I activity die at mid-gestation, revealing an essential role for complex or hybrid N-linked carbohydrates. Proc Natl Acad Sci USA 91:728–732

Kimura M, Abe T, Sunamura M, Matsuno S, Horii A (1996) Detailed deletion mapping on chromosome arm 12q in human pancreatic adenocarcinoma: identification of a 1-cM region of common allelic loss. Genes Chromosom Cancer 17:88–93

Kimura M, Furukawa T, Abe T, Yatsuoka T, Youssef EM, Yokoyama T, Ouyang H, Ohnishi Y, Sunamura M, Kobari M, Matsuno S, Horii A (1998) Identification of two common regions of allelic loss in chromosome arm 12q in human pancreatic cancer. Cancer Res 58:2456–2460

Looijenga LH, Smit VT, Wessels JW, Mollevanger P, Oosterhuis JW, Cornelisse CJ, Devilee P (1990) Localization and polymorphism of a chromosome 12-specific alpha satellite DNA sequence. Cytogenet Cell Genet 53:216–218

Minowa MT, Oguri S, Yoshida A, Hara T, Iwamatsu A, Ikenaga H,

Takeuchi M (1998) cDNA cloning and expression of bovine UDP-N-acetylglucosamine: alpha1, 3-D-mannoside beta1,4-N-acetylglucosaminyltransferase IV. J Biol Chem 273:11556–11562

Mori Y, Shiwaku H, Fukushige S, Wakatsuki S, Sato M, Nukiwa T, Horii A (1997) Alternative splicing of hMSH2 in normal human tissues. Hum Genet 99: 590–595

Sano T, Tsujino T, Yoshida K, Nakayama H, Haruma K, Ito H, Nakamura Y, Kajiyama G, Tahara E (1991) Frequent loss of heterozygosity on chromosomes 1q, 5q, and 17p in human gastric carcinomas. Cancer Res 51:2926–2931

Schachter H (1986) Biosynthetic controls that determine the branching and microheterogeneity of protein-bound oligosaccharides. Biochem Cell Biol 64:163–181

Tan J, Dunn J, Jaeken J, Schachter H (1996) Mutations in the MGAT2 gene controlling complex N-glycan synthesis cause carbohydrate-deficient glycoprotein syndrome type II, an autosomal recessive disease with defective brain development. Am J Hum Genet 59:810–817

Yoshimura M, Nishikawa A, Ihara Y, Taniguchi S, Taniguchi N (1995) Suppression of lung metastasis of B16 mouse melanoma by N-acetylglucosaminyltransferase III gene transfection. Proc Natl Acad Sci USA 92:8754–8758