Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 303 (2003) 868–876
BBRC
www.elsevier.com/locate/ybbrc

The molecular genetics of the mouse *I* β-1,6-*N*-acetylglucosaminyltransferase locus☆

Yuh-Ching Twu, Ming-Lun Chou, and Lung-Chih Yu*

Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan

Received 25 February 2003

Abstract

The I antigen and its precursor, the i antigen, are carbohydrate structures and are found on the surface of most mammalian cells. Conversion of the i to the I structure requires I β-1,6-*N*-acetylglucosaminyltransferase activity. The present investigation demonstrates a novel transcript form expressed from the mouse *I* locus and elucidates the molecular genetics and the genomic organization of the mouse *I* locus. The mouse *I* locus was demonstrated to express three transcript forms, one newly identified and two previously reported, which have a different exon 1 but identical exons 2 and 3. The three transcripts were shown to express differentially in various mouse tissues, and all their protein products demonstrated GlcNAc-transferring activity in enzyme function assay. The molecular genetics proposed for the mouse *I* locus shows that it is homologous to the human *I* locus. It has been established recently that a defect in the human *I* locus may lead to the development of congenital cataracts. It was demonstrated that the mouse and the human *I* transcripts expressed in the epithelium cells of the mouse and human lens, respectively, are homologous forms.

Keywords: I locus; β-1,6-*N*-acetylglucosaminyltransferase; IGnT; Lens

The i and I antigens are carbohydrate structures characterized as linear and branched repeats of *N*-acetyllactosamine, Galβ1-4GlcNAcβ1-3Gaβ1-4GlcNAc-R and Galβ1-4GlcNAcβ1-3(Galβ1-4GlcNAcβ1-6)Galβ1-4GlcNAc-R, respectively, carried on glycolipids and glycoproteins [1–5]. The *N*-acetyllactosamine repeats are synthesized by the sequential action of β-1,3-*N*-acetylglucosaminyltransferase and β-1,4-galactosyltransferase. Conversion of the i antigen into the I-active structure requires the action of a third enzyme, the I-branching β-1,6-*N*-acetylglucosaminyltransferase (I β6GlcNAcT) [2,6,7]. The i and I antigens were first identified on human red blood cells (RBCs) [8–12], but are known to be present on the surface of most human cells and also those of many other mammals [13]. They were also found to be present on soluble glycoproteins in various body fluids, including milk [14], saliva [15], plasma [15], amniotic fluid, urine, and ovarian cyst fluid [5,16]. The expression of the I antigen was found to be developmentally regulated in human RBCs [12,17], and changes in the expression patterns of the I and i antigens during the development of mouse embryos have also been observed [18]. Furthermore, altered expression patterns of I and i antigens have often been noted during the oncogenic processes [19], and thus, I and i are considered to be onco-developmental antigens [20].

In 1993, Bierhuizen et al. [21] successfully cloned the cDNA encoding the human I-branching β6GlcNAcT using an expression cloning technique. The gene, designated *IGnT*, is located at chromosome 6p24 [22,23]. Recently further investigations have provided evidences proving that the *IGnT* is the gene locus responsible for I-antigen expression on human RBCs [24,25], and have demonstrated the unusual molecular genetics of the human *I* locus: three different transcript forms, designated *IGnTA*, *IGnTB*, and *IGnTC*, each possessing a different exon 1 but identical exon 2 and 3 coding regions, are expressed from the human *I* locus. Furthermore it has

☆ The nucleotide sequences reported in this paper have been deposited in the GenBank/EBI Data Bank with Accession Nos. AY236873, AY236874, and AY236875.

* Corresponding author. Fax: +886-2-2363-5038.
E-mail address: yulc@ntu.edu.tw (L.-C. Yu).

0006-291X/03/$ - see front matter © 2003 Published by Elsevier Science (USA).
doi:10.1016/S0006-291X(03)00443-1

been established that a defect in the human *I* locus may directly lead to the development of congenital cataracts [25].

After the expression cloning of the human *IGnT* gene, identification of the *IGnT* homologue in mouse followed [26]. In 2000, Chen et al. [27] further reported a variant form of the *IGnT* transcript expressed from the mouse *IGnT* locus. The original and the variant forms of the mouse *IGnT* transcripts were designated as *IGnT A* and *IGnT B*, respectively, and were shown to have a different exon 1 but identical exon 2 and 3 coding regions.

The present report demonstrates a third form of the *IGnT* transcript expressed from the mouse *I* locus and shows the homologous molecular genetics and genomic organization of the mouse and human *I* loci. To distinguish the *IGnT* genes of the two species, the *IGnT* transcripts from the mouse and human *I* loci are designated as *mIGnT* and *hIGnT*, respectively. The newly identified mouse *IGnT* transcript, designated *mIGnTC*, has a different exon 1 but identical exons 2 and 3 to the previously reported mouse transcripts *IGnT A* and *IGnT B*, designated *mIGnTA* and *mIGnTB* in this paper, respectively. The homology between the mouse and the human *I*-gene loci and the possible physiological significance of the *I*-gene product in the lens of the eye are discussed.

Materials and methods

5′- and 3′-rapid amplification of cDNA end to establish the mIGnT transcript structures. The 5′- and 3′-rapid amplification of cDNA end (RACE) were performed on Marathon-Ready cDNA derived from mouse heart (BALB/c, males, aged 9–11 weeks, Clontech Laboratories, Palo Alto, CA) to establish the cDNA structures of the *mIGnTA*, *mIGnTB*, and *mIGnTC* transcripts. The primers used were: mIR32 (5′-GCCATGGACGTAGTGACCATGGCAGCCTCC-3′, complementary to nucleotides 1003–1032, spanning the exon 2–3 junction of the three *mIGnT* cDNAs, codon for initiation methionine as nucleotides 1–3), mIAR22 (5′-GTGATGGATGACCATGACATACGCCAAAG G-3′, for *mIGnTA*, complementary to nucleotides 280–309), mIBR24 (5′-GGATAGTAAGTGTGAAAGCCAGAGGGAAGC-3′, for *mIGnTB*, complementary to nucleotides 275–304), and mICR26 (5′-CA CCATGATGTAAGCCAGGGGAAAGGCAGC-3′, for *mIGnTC*, complementary to nucleotides 271–300) for 5′-RACE, and mIF23 (5′-TGGAGGCTGCCATGGTCACTACGTCCATGG-3′, nucleotides 1002–1031, spanning the exon 2–3 junction) and mIF27 (5′-ACACA TACCCTCTTACCGTGGAATGCCTGG-3′, nucleotides 1109–1138) for 3′-RACE. The polymerase chain reaction (PCR) was performed following the manufacturer's protocol. Amplified DNA fragments from RACE were cloned into the pCRII-TOPO vectors by a TOPO TA Cloning Kit (Invitrogen, Groningen, The Netherlands). Their DNA sequences were determined using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) and an Applied Biosystems sequencer. The full cDNA structures of the three *mIGnT* forms were assembled from the sequences of the longest 5′- and 3′-RACE products.

Reverse transcription-PCR analysis of the mIGnT transcript expression. cDNAs prepared from poly(A)$^+$ RNAs purified from mouse heart, brain, spleen, lung, liver, skeletal muscle, kidney, and testis (Clontech) were used. PCR amplification was performed using the forward primers: mIAFa (5′-CAAGCTTGACGGACTCTGACAT CGGAGACC-3′, for *mIGnTA*, nucleotides −41 through −12, codon for initiation methionine as nucleotides 1–3), mIBFb (5′-AACGA GCCAAGGAGCTTTGCTCATCAGAGC-3′, for *mIGnTB*, nucleotides −155 through −126), and mICFc (5′-CGCCTCCTAGGAGATT CAAGCTTCAGATGC-3′, for *mIGnTC*, nucleotides −257 through −228). The reverse primer, mIRc (5′-GTCTTCAATTTCCATTTA GGCCGGAGCTGC-3′, antisense sequence, seven nucleotides downstream from the stop codon in exon 3), was common for the three transcripts. The PCR program included 5 min at 94 °C followed by 40 cycles of 30 s at 94 °C, 30 s at 68 °C (for *mIGnTA* and *mIGnTC*) or 64 °C (for *mIGnTB*), and 1.5 min at 72 °C. The PCR products were analyzed by 1.5% agarose gel electrophoresis.

Twenty-eight lenses were obtained from adult mice (ICR, female). Total RNA from the lens-epithelium cells was purified using the RNeasy Mini Kit (Qiagen GmbH, Hilden, Germany). The first-strand cDNA was primed by oligo(dT) priming and synthesized by Sensiscript reverse transcriptase (Qiagen). PCR amplification, as described above, was then performed.

Functional analyses of the enzymes encoded by the mIGnT cDNAs. The cDNA fragment encompassing the region of nucleotides 88–1203, which encodes amino acid residues 30–401, of the *mIGnTA* gene was prepared by reverse transcription (RT)-PCR using the primers mIA-Fsfi (5′-aattggcccagccggccGATCAAAGCTACCAGAAGCTGAACA TCTCAG-3′) and mIRapa (5′-ttaagggcccGAAATACCAGCTCGGC TGTATGGCGATCTC-3′, antisense sequence). Primer of mIBFsfi (5′-aattggcccagccggccAGAAAATTATGGAAGAACTATCATTTCC CGAGGG-3′) and primer mIRapa were used to prepare the *mIGnTB* cDNA fragment encompassing nucleotides 73–1203, which encodes amino acid residues 25–401. The *mIGnTC* cDNA fragment encompassing nucleotides 97–1203, which encodes amino acid residues 33–401, was amplified by the primer mICFsfi (5′-aattggcccagccggccCAG CAGCTCAACAGCTCCAGCGAAAGG-3′) and primer mIRapa. The primers contained *Sfi*I and *Apa*I recognition sequences (underlined) at the 5′ ends, respectively. The template used was cDNA prepared from poly(A)$^+$ RNA purified from mouse heart (Clontech). The amplified cDNA fragments were cloned into *Sfi*I and *Apa*I sites of the mammalian expression vector pSecTaq2A (Invitrogen), which is designed for secretion of expressed protein using the N-terminal secretion signal from the V-J2-C region of the mouse Igκ chain. Vectors containing *mIGnTA*, *mIGnTB*, and *mIGnTC* cDNA inserts were selected and their sequences were confirmed. Control pSecTaq2A and the three constructed plasmids were prepared using the EndoFree Plasmid Kit (Qiagen) for transfection.

The control and the constructed plasmids were transfected into COS-7 cells, and after expression, the culture medium was harvested, concentrated using a Centriplus YM-10 (Millipore Intertech, Bedford, MA), and then directly used for GlcNAc-transferase assays. The GlcNAc-transferase assay was performed using ^{3}H-labelled UDP-GlcNAc as the donor substrate and LS-tetrasaccharide c (NeuNAcβ2-6Galβ1-4GlcNAcβ1-3Galβ 1-4Glc, Oxford GlycoSystems, Abingdon, UK) as the acceptor substrate. After incubation, the free and bound [^{3}H]GlcNAc was separated by being passed through a mixed bed of AG1-X8 (AcO^-) and AG50W-X8 (H^+) resins (Bio-Rad Laboratories, Hercules, CA) as described in detail in an earlier publication [24].

Results

The mouse I locus expresses three mIGnT transcripts, each with a different exon 1, but with identical exon 2 and exon 3 regions

It has been previously reported that mouse expresses the *mIGnTA* and *mIGnTB* transcripts, which

have different exon 1 regions, designated exon 1A and exon 1B, respectively, but have identical exon 2 and 3 coding regions [27]. Both of the two transcripts are made up of 1206 coding nucleotides, which are divided into three exons with respective coding nucleotides of 919, 96, and 191 bp. The mouse *IGnT* gene has been found to be located on chromosome 13. BLAST analysis of the gene databases at the National Center for Biotechnology Information (National Institute of Health, Bethesda, MD) with the exon 1A sequence revealed, in addition to the exon 1B region, another novel open reading frame with significant homology within the mouse *IGnT* locus. This novel region is designated 1C (Fig. 1A). The relative positions of the exon 1A, 1B, and 1C regions to the exons 2 and 3 of the *mIGnT* gene were known from the supercontig Mm13_WIFeb01_256 from mouse chromosome 13 (GenBank Accession No. NW_000075). From the 5′-RACE and RT-PCR analyses, it was demonstrated that the homologous 1C region is transcribed and also processed to link with the exon 2 and exon 3 regions, as for exons 1A and 1B. The 5′ splice site of the intron sequence at the exon–intron junction of exon 1C conforms to the GT consensus.

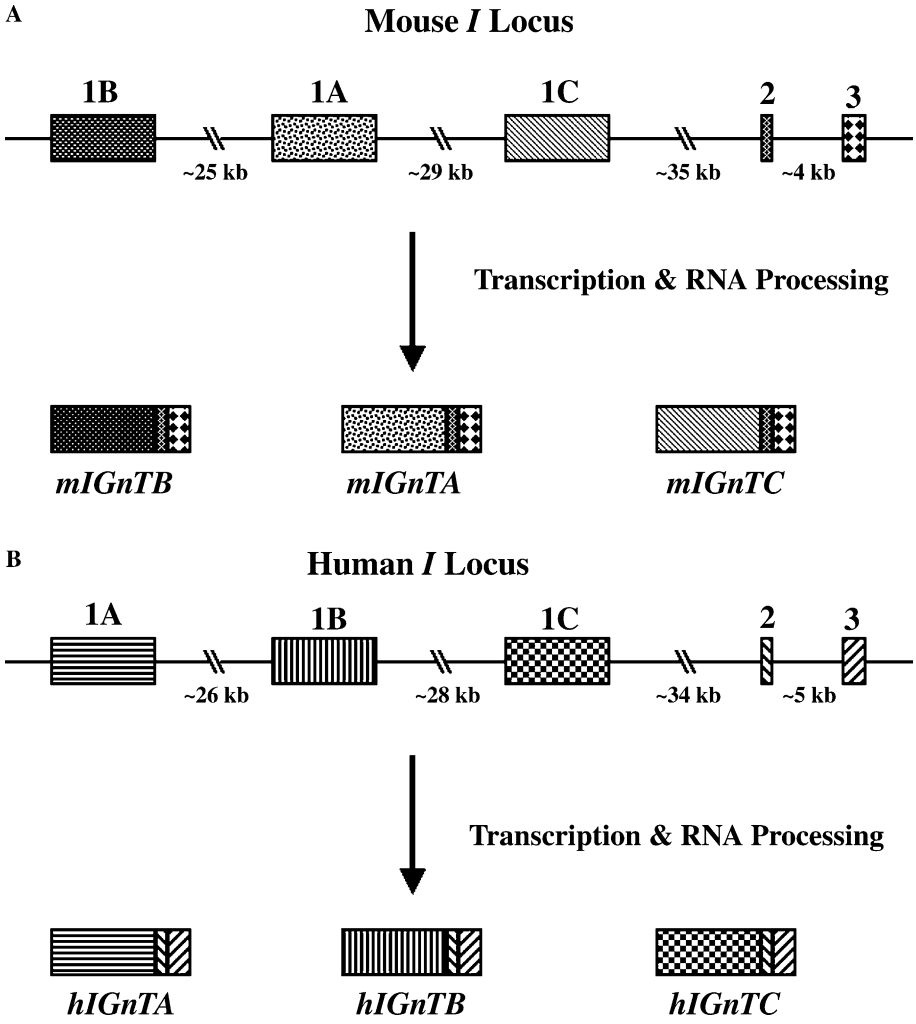

Fig. 1. (A) Schematic representation of the organization of the mouse *I* locus and the structures of the expressed *mIGnT* gene. Three *mIGnT* transcripts, *mIGnTA*, *mIGnTB*, and *mIGnTC*, which have a different exon 1, but identical exon 2 and exon 3, are expressed from the mouse *I* locus. The coding nucleotides of exons 1A, 1B, and 1C are 919 bp, and the common exons 2 and 3 have coding nucleotides of 96 and 191 bp, respectively. All the three transcripts predict protein products of 401 amino acid residues. The *mIGnTA* and *mIGnTB* transcripts have been reported previously and were formerly designated as *IGnT* [26] or *IGnTA*, and *IGnTB* [27], respectively. The previously designated exon 1B region is found to be located 5′ to the exon 1A region. (B) The organization of the human *I* locus and the structures of the expressed *hIGnT* gene. The coding nucleotides of the exons 1A, 1B, and 1C are 925, 919, and 925 bp, respectively, and the common exons 2 and 3 have coding nucleotides of 93 and 191 bp, respectively. Thus the *hIGnTA*, *hIGnTB*, and *hIGnTC* transcripts encode protein products with 402, 400, and 402 amino acid residues, respectively (adapted from Yu et al. [25]).

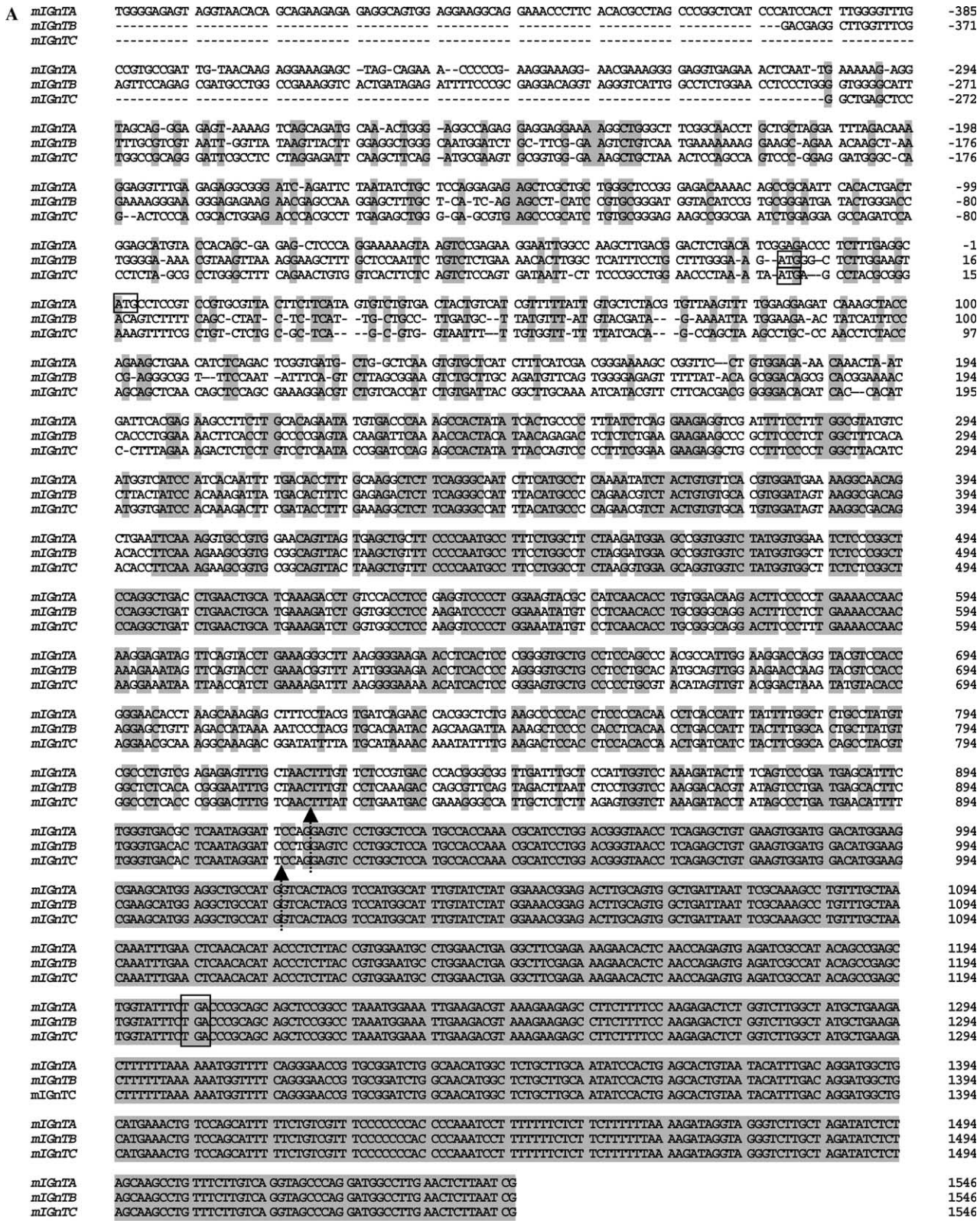

Fig. 2. (A) Nucleotide sequences of the *mIGnTA*, *mIGnTB*, and *mIGnTC* cDNAs. The full-length cDNA structures of the three *mIGnT* forms were established using 5′- and 3′-RACE performed on cDNA derived from mouse heart. The translation start codons are numbered 1–3. The translation start and stop codons are boxed. The exon 1–exon 2 and exon 2–exon 3 junctions are indicated by arrows. The sequences for *mIGnTA*, *mIGnTB*, and *mIGnTC* cDNAs have been deposited in the GenBank/EBI Data Bank with Accession Nos. AY236873, AY236874, and AY236875, respectively. (B) Amino acid sequences are deduced from the *mIGnTA*, *mIGnTB*, and *mIGnTC* cDNAs. The nine conserved cysteine residues are indicated by asterisk. Solid lines underline the hydrophobic segments at the N-terminals.

Fig. 2. (*continued*)

Thus, three *mIGnT* transcripts with different exon 1, but identical exon 2 and 3 regions, are expressed from the mouse *I* locus. The previously designated exon 1B region was found to be located 5′ to the exon 1A; however, we retain using the nomenclature. A very similar molecular genetic organization, as illustrated in Fig. 1B, has recently been demonstrated for the human *I* locus.

To establish the structures of the three *mIGnT* cDNAs, 5′- and 3′-RACE were performed on cDNA derived from mouse heart. The full cDNAs (Fig. 2A) were assembled from the sequences of the longest 5′- and 3′-RACE products. From 5′-RACE, the longest 5′ untranslated regions identified were 484, 387, and 282 bp for the *mIGnTA*, *mIGnTB*, and *mIGnTC* cDNAs, respectively, and the longest 3′ untranslated region obtained was 340 bp. The full cDNA structures assembled were further demonstrated by RT-PCR using the synthetic primers which anneal to the extreme 5′ and 3′ ends of the cDNA sequences (data not shown).

Exon 1C also consists of 919 coding nucleotides like exons 1A and 1B, and thus all three transcripts have open reading frames of 1206 nucleotides, predicting protein products of 401 amino acid residues (Fig. 2B). The protein product encoded by the novel *mIGnTC* cDNA has potential hydrophobic transmembrane segments at N-terminals, as do the *mIGnTA*- and *mIGnTB*-encoded β6GlcNAcTs, and shares 70% and 73% sequence identity with the mIGnTA and mIGnTB β6GlcNAcTs, respectively. Sixty-three percentage overall sequence identity is demonstrated for the three mIGnT proteins. Nine cysteine residues are conserved across all three of them.

Expression profiles of the mIGnT transcripts in various mouse tissues

Using RT-PCR analysis, differential expression profiles for the three *mIGnT* transcripts in the mouse tissues were revealed (Fig. 3A). For instance, a relatively strong expression of the *mIGnTA* transcript, with weak expression of the *mIGnTC* transcript and only trace amounts of the *mIGnTB* transcript, was demonstrated in heart. In lung tissue, significant expression of all the three transcripts was observed, whereas only the *mIGnTB* transcript, but not the other two, was detected in liver.

The expression of the three *mIGnT* transcripts in epithelium cells of the mouse lens was examined. As shown in Fig. 3B, only the *mIGnTA* transcript, but not the other two forms, was detected in the RNA sample purified from epithelium cells of the mouse lens.

Activity of the enzymes encoded from the mIGnT cDNAs

The protein product encoded from the novel *mIGnTC* cDNA was expressed in mammalian cells (COS-7), and its potential GlcNAcT activity was examined and compared with those of the I β6GlcNAcTs encoded from the *mIGnTA* and *mIGnTB* cDNAs. As shown in Table 1,

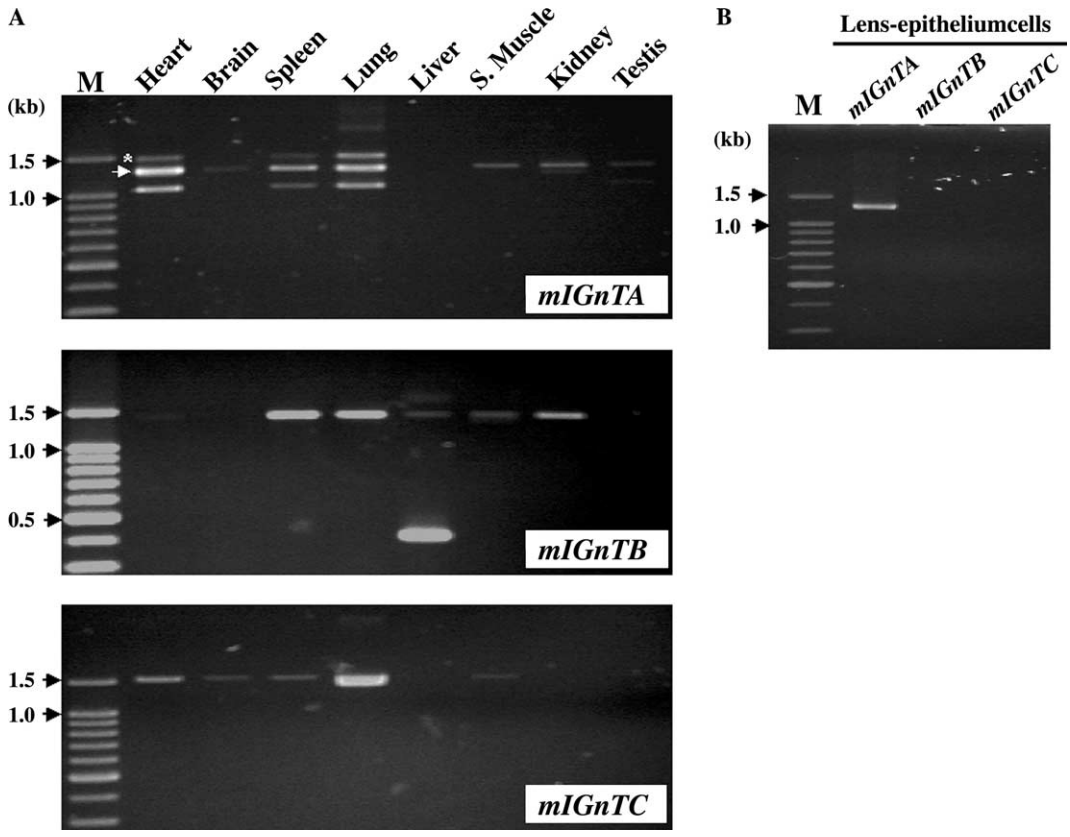

Fig. 3. The expression profiles for the *mIGnT* transcripts in various mouse tissues (A) and the epithelium cells of the mouse lens (B). cDNAs prepared from poly(A)$^+$ RNAs purified from mouse heart, brain, spleen, lung, liver, skeletal muscle, kidney, and testis were used for the former. PCR using gene-specific forward primer and common reverse primer was performed, as described in Materials and methods. The RT-PCR products were analyzed using 1.5% agarose gel electrophoresis. The expected sizes of the products from the *mIGnTA*, *mIGnTB*, and *mIGnTC* transcripts were 1283, 1397, and 1499 bp, respectively. In the RT-PCR for the *IGnTA*, together with the 1283-bp products (indicated by an arrow), additional products of approximately 1.1 kb in size were observed. This smaller product, believed to result from alternative splicing, has been demonstrated by sequencing to consist of a shorter exon 1A region conjoining with exon 2–3 regions, and does not have a correct reading frame relative to the exon 2–3 coding sequence. The bands (indicated by an asterisk), approximately 1.5 kb in size, were the hybrid complex of the *IGnTA* 1283-bp and the alternatively spliced 1.1 kb products. An alternatively spliced *IGnTB* transcript, approximately 400 bp in size, consisting of a very short exon 1B region, was observed in the lung tissue. The expression of the three *mIGnT* transcripts in the epithelium cells of the mouse lens was examined. The first-strand cDNA, primed by oligo(dT) primer, was synthesized by Sensiscript reverse transcriptase using total RNA purified from the epithelium cells of a number of mouse lenses as the target for PCR amplification.

Table 1
The GlcNAcT activity of the enzymes encoded from the *mIGnTA*, *mIGnTB*, and *mIGnTC* cDNAs

	Vector	*mIGnTA*	*mIGnTB*	*mIGnTC*
GlcNAc transferred (pmol)	76.0 ± 6.1	2488.3 ± 115.4	4225.9 ± 399.0	4158.2 ± 186.1

The results of the average and standard deviation of four tests are shown. Endogenous transfer of GlcNAc in the absence of acceptor substrate was corrected for each test. The amounts of transferred GlcNAc in the vector control, pSecTaq2A, indicate the background levels of the assay and are believed to result from the addition of the acceptor substrate.

GlcNAcT activity was detected for the expressed mIGnTC protein product as well as for the *mIGnTA*- and *mIGnTB*-encoded products. It has been clearly demonstrated that the enzymes encoded by the *mIGnTA* and *mIGnTB* transcripts show I β6GlcNAcT activity [26,27]. Compared with the mIGnTA and mIGnTB β6GlcNAcTs, comparable or even higher GlcNAcT activity was demonstrated for the mIGnTC enzyme with the same acceptor substrate, LS-tetrasaccharide c, which mimics the i structure, the precursor of I epitope, and has been shown to be a good acceptor substrate for β6GlcNAcT activity assay [27].

Discussion

The present study has demonstrated that, in addition to the previously reported *mIGnTA* and *mIGnTB*

Table 2
The homology for the amino acid sequences of the mouse and human IGnTs

	mIGnTB	mIGnTA	mIGnTC	hIGnTA	hIGnTB	hIGnTC
mIGnTB	100					
mIGnTA	69	100				
mIGnTC	73	70	100			
hIGnTA	78	68	69	100		
hIGnTB	69	86	70	73	100	
hIGnTC	68	69	75	71	73	100

The identity percentages for the amino acid sequences are represented. The amino acid sequences were aligned, using GeneWorks software, and different alignment parameters were applied to obtain the highest identity percentage. The highest three of the identity percentages among the comparisons of the different IGnTs are boxed. Note that the mIGnTB is ordered in front of the mIGnTA, according to their relative positions in the genomic structure.

transcripts, a third *mIGnTC* transcript is expressed from the mouse *I* locus. The three *mIGnT* transcript forms have different exon 1 but identical exons 2 and 3. The enzymes encoded from the *mIGnTA* and *mIGnTB* cDNAs have been shown to exhibit I β6GlcNAcT activity. The results shown in enzyme function assay suggest that the *mIGnTC*-encoded enzyme also has I β6GlcNAcT activity. Nevertheless, further carbohydrate structure analysis of the products generated from mIGnTC enzyme activity is required to further characterize the GlcNAcT activity and to demonstrate the specific nature of the I β6GlcNAcT activity of the newly identified mIGnTC enzyme. More detailed comparison of the enzyme kinetics of the three mIGnT enzymes and whether this newly identified enzyme possesses any other glycosyltransferase specificity other than the I β6GlcNAcT activity are also worthy of further investigation.

The three *mIGnT* transcripts were shown to exhibit differential expression across various mouse tissues. The 5′-RACE analyses showed that the three *mIGnT* cDNAs did not have a common 5′ region, indicating that transcription of the *mIGnT* forms may be determined by different DNA regulatory regions or by different regulatory mechanisms. It will be of interest to elucidate the regulatory mechanisms for each of the three *mIGnT* forms and to thus gain an understanding of how the differential regulation of the *mIGnT*s may be manipulated.

It should be noted that the molecular genetics proposed for the mouse *I* locus is homologous with that for the human *I* locus, which has been recently revealed to express three transcripts, *hIGnTA*, *hIGnTB*, and *hIGnTC*, each also with different exon 1 but identical exon 2 and 3 coding regions [25]. Furthermore, the genomic organizations of the exon regions for the mouse and human *I* loci show a high degree of homology (Figs. 1A and B). The existence of such an unusual molecular genetic mechanism for *I* locus, which expresses three different transcripts through utilizing different promoter regions, both in human and mouse, and the highly homologous *I*-locus genomic organization in the two species suggest conservation of the *I*-gene locus during evolution. Indeed, the similarities between the homologous IGnTs for the two species are higher than the similarities between the IGnT forms in the same species. The similarities between the six IGnT proteins were compared by sequence alignment analysis (GeneWorks version 2.3, IntelliGenetics, Campbell, CA) and are presented in Table 2. The mIGnTB protein with hIGnTA, mIGnTA with hIGnTB, and mIGnTC with hIGnTC have identities in amino acid sequence of 78%, 86%, and 75%, respectively, which are the highest three among all comparisons of the six IGnTs. The sequence identities among the three mouse IGnTs range between 69% and 73%, while those among the human IGnTs range between 71% and 73%. The UPGMA (unweighted pair-group method using arithmetic means) distance tree [28], based on these amino acid sequence identity comparisons, showed clustering of mIGnTB with hIGnTA, mIGnTA with hIGnTB, and mIGnTC with hIGnTC (data not shown), suggesting a homologous origin for each pair. The UPGMA trees established using the nucleotide sequences, full length or only the exon 1 coding region, also gave similar clustering patterns.

Although the I antigen has been referred to as oncodevelopmental antigen due to its altered expression at different cell stages during development and in oncogenesis, the I carbohydrate structure or the I β6GlcNAcT enzyme has never been clearly given any physiological significance. However, the human *I* locus has been correlated with the development of congenital cataracts, as the partial association of congenital cataracts with the adult i phenotype, a rare blood type where RBCs do not express common I epitope due to a defect in the I β6GlcNAcT activity, has been observed [29–34]. Recently, through molecular genetic studies and pedigree analyses of the individuals with the adult i phenotype, the molecular genetic background explaining the partial association of the adult i phenotype with congenital cataracts has been elucidated [24,25], and this has indicated that a defect in the human *I* locus may lead directly to the development of congenital cataracts. The *hIGnTB* was found to be the only one of the three *hIGnT* transcripts expressed in the epithelium cells of the human lens

[25], suggesting that the hIGnTB form of the enzyme may play a functional role in lens transparency. Interestingly the *mIGnTA* transcript was the only one of the three *mIGnT* forms detected in RNA purified from the epithelium cells of the mouse lens (Fig. 3B). The mouse *IGnTA* form is the homologue of the human *IGnTB* form, judging from their similar genomic organization position and the high sequence identity (86%) of the gene products.

The molecular genetics proposed for the *I* locus, in mouse and in human, offers a new perspective on the formation and expression of the I antigen in different tissues and cells. In attempting to understand the expression profiles of the I antigen in different tissues and cells and the mechanisms controlling the appearance and disappearance of the I antigen, the functional roles of the three individual *IGnT* forms should be considered. The homology of the molecular genetics of the *I* locus in mouse and human suggests that mouse might be an appropriate animal model for the study of the biological significance of the I carbohydrate structure and the I β6GlcNAcT activity. It will be especially interesting to understand the mechanism underlying the *I*-gene defect associated with the formation of cataracts and the functional role that the *I*-gene product may play in maintaining lens transparency.

Acknowledgments

This work was partly supported by National Science Council Grants NSC 91-2314-B-002-405 and National Research Program for Genomic Medicine of National Science Council Grant NSC 91-3112-B-002-028.

References

[1] H. Neimann, K. Watanabe, S. Hakomori, R.A. Childs, T. Feizi, Blood group i and I activities of 'lacto-*N*-nor-hexaosylceramide' and its analogues: the structural requirements for i-specificity, Biochem. Biophys. Res. Commun. 81 (1978) 1286–1293.

[2] K. Watanabe, S. Hakomori, R.A. Childs, T. Feizi, Characterization of a blood group I-active ganglioside, J. Biol. Chem. 254 (1979) 3221–3228.

[3] J. Koscielak, E. Zdebska, Z. Wilczynska, H. Miller-Podraza, W. Dzierzkowa-Borodej, Immunochemistry of Ii-activity glycosphingolipids, Eur. J. Biochem. 96 (1979) 331–337.

[4] P.D. Issitt, D.J. Anstee, Applied Blood Group Serology, Montgomery Scientific Publications, Durham, NC, 1998.

[5] G. Daniels, Human Blood Groups, second ed., Blackwell Science, Oxford, England, 2002.

[6] M. Fukuda, M. Fukuda, S. Hakomori, Developmental change and genetic defect in the carbohydrate structure of band 3 glycoprotein of human erythrocyte membrane, J. Biol. Chem. 254 (1979) 3700–3703.

[7] F. Piller, J.-P. Cartron, Biosynthesis of blood group I antigens, J. Biol. Chem. 259 (1984) 13385–13390.

[8] A.S. Wiener, L.J. Unger, L. Cohen, J. Feldman, Type-specific cold auto-antibodies as a cause of acquired hemolytic anemia and hemolytic transfusion reactions: biological test with bovine red cells, Ann. Intern. Med. 44 (1956) 221–240.

[9] W.J. Jenkins, W.L. Marsh, J. Noades, P. Tippett, R. Sanger, R.R. Race, The I antigen and antibody, Vox Sang. 5 (1960) 97–106.

[10] P. Tippett, J. Noades, R. Sanger, R.R. Race, L. Sausais, C.A. Holman, J. Buttimer, Further studies of the I antigen and antibody, Vox Sang. 5 (1960) 107–121.

[11] W.L. Marsh, W.J. Jenkins, Anti-i: a new cold antibody, Nature 188 (1960) 753.

[12] W.L. Marsh, Anti-i: a cold antibody defining the Ii relationship in human red cells, Br. J. Haematol. 7 (1961) 200–209.

[13] A.S. Wiener, J. Moor-Jankowski, E.B. Gordon, J. Davis, The blood factors I and i in primates including man, and in lower species, Am. J. Phys. Anthropol. 23 (1965) 389–396.

[14] W.L. Marsh, M.E. Nichols, F.H. Allen, Inhibition of anti-I sera by human milk, Vox Sang. 18 (1970) 149–154.

[15] P. Rouger, G. Juszczak, C. Doinel, C. Salmon, Relationship between I and H antigens, I. A study of the plasma and saliva of a normal population, Transfusion 20 (1980) 536–539.

[16] M.E. Reid, C. Lomas-Francis, The Blood Group Antigen Facts Book, Academic Press, San Diego, CA, 1997.

[17] W.L. Marsh, M.E. Nichols, M.E. Reid, The definition of two I antigen components, Vox Sang. 20 (1971) 209–217.

[18] A. Kapadia, T. Feizi, M.J. Evans, Changes in the expression and polarization of blood group I and i antigens in post-implantation embryos and teratocarcinomas of mouse associated with cell differentiation, Exp. Cell Res. 131 (1981) 185–195.

[19] S. Hakomori, R. Kannagi, Glycosphingolipids as tumor-associated and differentiation markers, J. Natl. Cancer Inst. 71 (1983) 231–251.

[20] T. Feizi, Demonstration by monoclonal antibodies that carbohydrate structures of glycoproteins and glycolipids are onco-developmental antigens, Nature 314 (1985) 53–56.

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

[22] J.-C. Yeh, E. Ong, M. Fukuda, Molecular cloning and expression of a novel β-1,6-*N*-acetylglucosaminyltransferase that forms core 2, core 4, and I branches, J. Biol. Chem. 274 (1999) 3215–3221.

[23] T. Schwientek, M. Nomoto, S.B. Levery, G. Merkx, A.G. van Kessel, E.P. Bennett, M.A. Hollingsworth, H. Clausen, Control of *O*-glycan branch formation, J. Biol. Chem. 274 (1999) 4504–4512.

[24] L.-C. Yu, Y.-C. Twu, C.-Y. Chang, M. Lin, Molecular basis of the adult i phenotype and the gene responsible for the expression of the human blood group I antigen, Blood 98 (2001) 3840–3845.

[25] L.-C. Yu, Y.-C. Twu, M.-L. Chou, M.E. Reid, A.R. Gray, J.M. Moulds, C.-Y. Chang, M. Lin, The molecular genetics of the human *I* locus and molecular background explain the partial association of the adult i phenotype with congenital cataracts, Blood 101 (2003) 2081–2087.

[26] A.D. Magnet, M. Fukuda, Expression of the large I antigen forming β-1,6-*N*-acetylglucosaminyltransferase in various tissues of adult mice, Glycobiology 7 (1997) 285–295.

[27] G.-Y. Chen, N. Kurosawa, T. Muramatsu, A novel variant form of murine β-1,6-*N*-acetylglucosaminyltransferase forming branches in poly-*N*-acetyllactosamines, Glycobiology 10 (2000) 1001–1011.

[28] P.H. Sneath, R.R. Sokal, Numerical Taxonomy, W.H. Freeman, San Francisco, CA, 1973.

[29] H. Yamaguchi, Y. Okubo, M. Tanaka, A note on possible close linkage between the Ii blood locus and a congenital cataract locus, Proc. Jpn. Acad. 48 (1972) 625–628.

[30] H. Ogata, Y. Okubo, T. Akabane, Phenotype i associated with congenital cataract in Japanese, Transfusion 19 (1979) 166–168.

[31] M. Lin-Chu, R.E. Broadberry, Y. Okubo, M. Tanaka, The i phenotype and congenital cataracts among Chinese in Taiwan, Transfusion 31 (1991) 676–677.
[32] W.L. Marsh, H. DePalma, Association between the Ii blood group and congenital cataract, Transfusion 22 (1982) 337–338.
[33] E.B. Macdonald, R. Douglas, P.A. Harden, A Caucasian family with the i phenotype and congenital cataracts, Vox Sang. 44 (1983) 322–325.
[34] P.L. Page, S. Langevin, R.A. Petersen, M.S. Kruskall, Reduced association between the Ii blood group and congenital cataracts in white patients, Am. J. Clin. Pathol. 87 (1987) 101–102.