J. Biochem. 132, 387–393 (2002)

Phylogenetic Development of a Regulatory Gene for the Core 2 GlcNAc Transferase in *Mus musculus*

Michiko Sekine,* Yoshiaki Kikkawa,* Sumiyo Takahama,* Kaoru Tsuda,* Hiromichi Yonekawa,* and Akemi Suzuki†,1

*Department of Laboratory Animal Science, The Tokyo Metropolitan Institute of Medical Science, 3-18-22, Honkomagome, Bunkyo-ku, Tokyo 113-8613; and †Sphingolipid Expression Laboratory, RIKEN Frontier Research System, 2-1, Hirosawa, Wako, Saitama 351-0198

Received May 15, 2002; accepted June 25, 2002

In our previous study, we identified a mouse gene, *Gsl5*, that controls the expression of a glycolipid, GL-Y [Galβ1-4(Fucα1-3)GlcNAcβ1-6(Galβ1-3)GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer], and the core 2 structure of *O*-linked glycans of glycoproteins, GlcNAcβ1-6(Galβ1-3)GalNAcα-Ser/Thr, in a kidney tubular cell-specific manner through the regulation of UDP-GlcNAc β-1,6-GlcNAc transferase (GNT). Regulation by the *Gsl5* gene occurs at the level of GNT mRNA and the recessive allele of *Gsl5* is rare and carried by DBA/2 and its related strains. Here, we report a sequence comparison of the 5′ flanking region of the GNT gene among 5 laboratory strains and 10 wild-derived strains, demonstrating that the DBA/2 allele sequence is similar to the sequence carried by Asian *Mus m. musculus* and differs substantially from the East European *M. m. musculus*. These results suggest that the DBA/2 allele of *Gsl5* was introduced into laboratory mouse strains by Asian wild-derived mice. Phylogenetic comparison of the 5′ flanking region sequences between the recessive and dominant *Gsl5* alleles indicates that mutations to create a functional *Gsl5* gene occurred approximately one million years ago during the subspeciation of *M. musculus*, and provides a case for studies on the creation of functional genes involved in tissue-specific transcriptional regulation.

Key words: glycolipid expression, kidney, mouse, polymorphism, transcriptional regulation.

The diverse carbohydrate structures of glycoconjugates on the cell surface are essential in cell–cell and cell–matrix interactions, cellular differentiation, and malignant transformation (*1*). The expression of these diverse structures is precisely regulated by the synergy of multiple gene products responsible for the biosynthesis or degradation of glycoconjugates, including glycosyltransferases (*2, 3*), glycosidases (*4, 5*), nucleotide-sugar transporters (*6, 7*), and related molecules. Identification of the mechanisms regulating these processes will facilitate our understanding of the physiological role of these diverse structures.

In the course of our study to clarify these mechanisms, we discovered polymorphic differences in the glycolipids of various tissues among inbred strains of mice (*8*). Genetic analysis of kidney neutral glycolipids identified a single autosomal gene regulating the expression of an 8-sugar glycolipid, GL-Y [Galβ1-4(Fucα1-3)GlcNAcβ1-6(Galβ1-3)GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer] (*9–12*). This gene, designated *Gsl5* (glycosphingolipid regulatory gene-5), controls the expression of GL-Y by regulating β-1,6-GlcNAc transferase (GNT) activity (*13*). We purified the GNT protein, which transfers β-GlcNAc to the 6-position of the penultimate GalNAc of Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer (GL-X), from mouse kidney (*14*). We then cloned two mouse cDNAs encoding GNT that differ only in their 5′-untranslated sequences, suggesting the usage of an alternative promoter and the involvement of alternative splicing (*15*). An mRNA corresponding to one of the two cDNAs is expressed ubiquitously in various tissues and the other is only detected in the kidney. The latter transcript (kidney-type mRNA) is present in the kidneys at levels 20-times greater than the ubiquitous-type mRNA (*16*). Analysis of the GNT gene revealed the presence of 6 exons, production of three transcripts, and one exon encoding the entire open reading frame (ORF). The ORF, therefore, is shared by all three transcripts.

Among inbred strains of mice, the DBA/2 strain lacks detectable levels of GL-Y and GNT activity in the kidney, instead expressing large amounts of the precursor, GL-X. We confirmed that the DBA/2 strain carries a defective *Gsl5* allele by mating experiments. The lack of the kidney-type GNT transcript in DBA/2 mice indicates that *Gsl5* regulates the transcription of kidney-type mRNA (Fig. 1). Using AKXD/Ty recombinant inbred strains, we demonstrated that the *Gsl5* gene (accession ID, MGI: 95847) is localized on chromosome 19 and is closely linked to the gene for β-1,6-GlcNAc transferase, GNT (*Gcnt1*, accession

[1] To whom correspondence should be addressed. Tel: +81-48-467-9615, Fax: +81-48-462-4692, E-mail: aksuzuki@postman.riken.go.jp

Abbreviations: GL-Y, Galβ1-4(Fucα1-3)GlcNAcβ1-6(Galβ1-3)GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer; GL-X, Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer; GNT, UDP-GlcNAcβ-1,6-*N*-acetylglucosaminyltransferase; C2GnT1, core 2 β-1,6-*N*-acetylglucosaminyltransferase 1; *Gsl5*, glycosphingolipid regulatory gene-5; *Gcnt*, the locus for GNT; SSLP, simple sequence length polymorphism; RFLP, restriction fragment length polymorphism.

ID, MGI: 95676) (*15*).

Most common laboratory mice, including the DBA/2 strain, were established by interbreeding between European and Asian fancy mouse stocks in the last 100 years (*17*), and carry the main genetic background of a European subspecies of *Mus musculus*, *M. m. domesticus*, with some genes derived from the Asian subspecies, *M. m. musculus* (*18, 19*). These lines of evidence suggested that the DBA/2 allele of the GNT gene regulatory element was derived from *M. m. musculus* and a comparison between recessive and dominant alleles of the GNT gene regulatory element would provide an interesting example for the evolutional process involved in the production of functional glycosyltransferases or the tissue-specific regulation of glycosyltransferases.

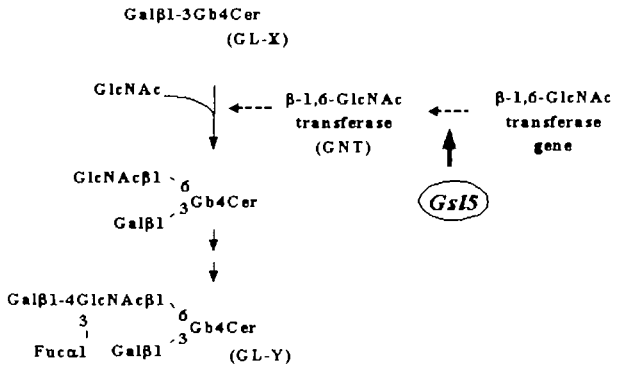


Fig. 1. **Biosynthetic pathway of GL-Y in mouse kidney.** *Gsl5* regulates the transcription of the GNT gene in a kidney-specific manner.

In this study, we analyzed kidney glycolipids of mice derived from Asian wild mice, sequenced about 1.2 kb of the 5′ upstream region of the GNT gene, and compared these sequences with those of the *Gsl5* gene-wild type allele. The comparison suggested an interesting relation between the wild and recessive types of the *Gsl5* alleles. Here, we report the lineage of *Gsl5* genes among wild and laboratory mouse strains and discuss their phylogenetic relationships.

MATERIALS AND METHODS

Mice—Wild mice were kindly provided by Drs. Moriwaki, Shiroishi, and Tsuchiya. Some of these mice were derived from the *M. musculus* subspecies, including *M. m. musculus* or *M. m. molossinus*, maintained at the National Institute of Genetics (Mishima, Shizuoka) or Miyazaki Medical College (Kiyotake-cho, Miyazaki). Their origin and subspecies nomenclature are detailed in Table I.

Analysis of Kidney Neutral Glycolipids—Kidney neutral glycolipids were isolated and identified as described (*9*). Briefly, total glycolipids were extracted from kidneys in a chloroform–methanol mixture. The neutral glycolipid frac-

TABLE I. **The expression of GL-X and GL-Y in mouse kidney.**

Strain	GL-X	GL-Y	Subspecies	Strain	Collection locality	GL-X	GL-Y
DBA/2	+	–	*M. m. molossinus*	MSM/Ms	Mishima, Japan	+	–
DBA/1J	+	–		Kgs	Kagoshima, Japan	–	+
CBA/J	+	–		Hkz	Hakozaki, Japan	+	–
CBA/CaHN	+	–		MOM/Nga	Nagoya, Japan	+	–
TT6/Le	+	–		JF1/Ms[a]	(Japanese fancy mouse)	+	–
BALB/c	–	+	*M. m. musculus*	Bjn	Beijing, China	+	–
C57BL/10	–	+		Shh	Shanghai, China	+	–
C57BL/6	–	+		CHD/Ms	Chengdu, China	+	–
WHT/Ht	–	+		KJR/Ms	Kojuri, Korea	+	–
AKR/J	–	+		SWN/Ms	Suwon, Korea	+	–
SWR/J	–	+		NJL/Ms	Northern Jutland, Denmark	+	–
C57L/J	–	+		BLG2/Ms	Toshevo, Bulgaria	+	–
C3H/He	–	+		M.MAN	Heilongjiang, China	+	–
P/J	–	+		M.KYA	Kyachta, Russia	+	–
DM/Shi	–	+		M.V-K	Vyasovka × Komosomolsk, Russia	+	–
I/LnJ	–	+		M.w.AST	Astrakhan, Russia	+	–
SJL/J	–	+		M.NIK	Nikolaevsk, Russia	+	–
SM/J	–	+		M.S-K	Sovgavan × Komosomolsk, Russia	+	–
IQI/Jic	–	+		M.w-KAZ	Aktyubinsk, Kazakhstan	+	–
TF/GnLe	–	+		M.W.AKT	Akutubinsk, Kazakhstan	+	–
WB/ReJ-W	–	+		M.KAZ	Balkash Lake, Kazakhstan	+	–
RIIIs/J	–	+		M.UZB	Tashkent, Uzbekistan	+	–
HRS/J	–	+	*M. m. bactrianus*	Iran	Mashad, Iran	+	–
MS/MyJ	–	+	*M. m. domesticas*	BFM/2[c]	Montpellier, France	–	+
129/J	–	+		SK/Cam[c]	Skokholm Island, England	–	+
A/WgSnJ	–	+		PGN2/Ms	Pegion, Canada	–	+
NZB/BINJ	–	+		Sey	Seychells	–	+
C57BL/6ByJ	–	+	*M. m. castaneus*	Bgr	Bogor, Indonesia	–	+
C58/J	–	+		Mal[b]	Malaysia	–	+
C57BR/cdJ	–	+		CASA/Rk[c]	Thailand	–	+
RFM/MnNrS	–	+		CAST/Ei[c]	Thailand	–	+
PT	–	+		MYS/Ms	Misore, India	–	+
			M. m. homourus	M.HOM-9	Tukuche, India	–	+
				M.HOM-11	Tukuche, India	–	+
			M. spretus	SEG/Pas[b]	Spain	–	–
			M. spicilegus	ZBN/Ms[b]	Bulgaria	–	+

[a]See Ref. 27. [b]Exact collection locations are not known. [c]See Ref. 30.

tion was eluted from a DEAE-Sephadex A-25 column and subjected to mild alkaline treatment. After neutralization, the neutral glycolipid fractions were desalted on a Sephadex LH-20 column and analyzed by TLC. Glycolipids were visualized by spraying the plates with an orcinol-sulfuric acid reagent and heating at 110°C.

RT-PCR Analysis of β-1,6-GlcNActransferase (GNT) Transcripts—Total RNA was obtained from the kidneys of DBA/2, MSM, JF1, CHD/Ms, KJR/Ms, SWN/Ms, BLG2/Ms, AKR/J, C57BL/6, and BALB/c mice as described previously (*15*). First strand cDNA synthesis was performed using Omniscript Reverse Transcriptase (QIAGEN Inc., Valencia, CA) with 100 ng total RNA as a template. cDNAs were amplified by PCR using sense primers (primer A: 5′CTGGAACCCAATGAAACCTT3′, primer B: 5′AGCCGATCCAGCAGTGTTTC3′, and primer C: 5′CTGGTGCTTGATAGGAACTT3′) based on either the 5′ sequences of kidney- and ubiquitous-type cDNAs or a region common to both cDNAs, and an antisense primer (primer D: 5′GTCACCCTGTAAAATCTTGG3′) based on the coding sequence. PCR products were analyzed by agarose gel electrophoresis.

Comparison of the 5′ Upstream Sequences of GNT Genes—The 5′ upstream region of the kidney-specific GNT cDNA was amplified by PCR using the primers UP1 (5′GTGGTAGCCTGAGCATTTCT3′) and UP2 (5′GGCTAGGGATTGCTGGTTAT3′). Genomic DNA obtained from DBA/2, MSM/Ms, JF1/Ms, CHD/Ms, KJR/Ms, SWN/Ms, BLG2/Ms, CBA/J, AKR/J, C57BL/6, BALB/c, PGN2/Ms, M.HOM-9, MYS/Ms, and SEG/Pas (*M. spretus*) were used as templates. PCR products were sequenced directly using a BigDye Terminator Cycle-Sequencing kit (Applied Biosystems, Foster City, CA) and an ABI PRISM 377 DNA sequencer.

PCR-SSLP Analysis for Microsatellite Markers on Chromosome 19—Simple sequence length polymorphism (SSLP) testing was performed using primer sets purchased from Research Genetics (Huntsville, AL) identifying microsatellite markers. PCR-amplified fragments of the genomic DNA obtained from C57BL/6, BALB/c, DBA/2, MSM/Ms, and JF1/Ms mice were analyzed by agarose gel electrophoresis.

RESULTS

Kidney Neutral Glycolipids—TLC analysis of kidney neutral glycolipids identified two categories of mice, one expressing GL-Y, such as BALB/c, C57BL/10, C57BL/6, and WHT/Ht, and another expressing GL-X rather than GL-Y, including DBA/2 (Fig. 2A). Five of the 32 inbred strains tested, DBA/2, DBA/1, TT6/Le, CBA/J, and CBA/CaHN, do not express GL-Y (Table I).

As shown in Fig. 2B, the wild-derived strains, MSM/Ms derived from *M. m. molossinus* and BLG2/Ms from *M. m. musculus*, express GL-X instead of GL-Y, whereas PGN2/Ms derived from *M. m. domesticus* and CASA/Rk from *M. m. castaneus* express GL-Y. SEG/Pas, an inbred strain derived from *M. spretus*, expresses neither GL-Y nor GL-X, representing an interesting phenotype that corresponds well to the position of *M. spretus* in the *Mus* genus, *i.e.*, *M. spretus* males and laboratory strain females are bred to produce viable offspring of both sexes, and although all male hybrids are sterile, the female hybrids are fully fertile and can be backcrossed to either *M. musculus* or *M. spretus* males to obtain fully viable second-generation offspring (*19*). This phenotype is not dealt with further in this paper, but the sequence of the 5′ upstream region of the *M. spretus* GNT gene was used as the outgroup to construct a phylogenetic tree.

The whole results on inbred and wild-derived strains are summarized in Table I. All strains of *M. m. musculus* and an Iranian mouse, *M. m. bactrianus*, exhibit the same phenotype as DBA/2, whereas *M. m. domesticus*, *M. m. castaneus*, and *M. m. homourus* express GL-Y. In contrast, Japanese wild mice, *M. m. molossinus*, include both classes: MSM from Mishima, Shizuoka prefecture, Hkz from Hakozaki, Fukuoka prefecture, and MOM/Nga from Nagoya, Aichi prefecture do not express GL-Y, while Kgs from Kago-

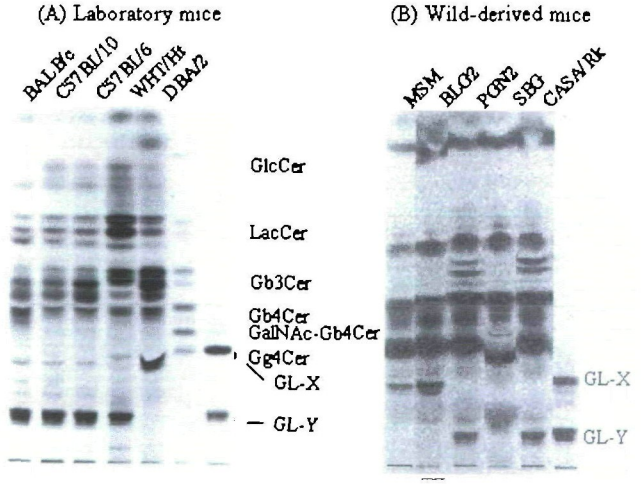


Fig. 2. **Neutral glycolipids isolated from the kidneys of (A) laboratory mice and (B) wild-derived mice.** The right two lanes of the TLC (A) display the standard glycolipids: GlcCer, glucosylceramide; LacCer, lactosylceramide; Gb3Cer, globotriaosylceramide; Gb4Cer, globotetraosylceramide (globoside); GalNAc-Gb4Cer, *N*-acetylgalactosaminylglobotetraosylceramide (Forssman glycolipid); and Gg4Cer, gangliotetraosylceramide (asialo GM1).

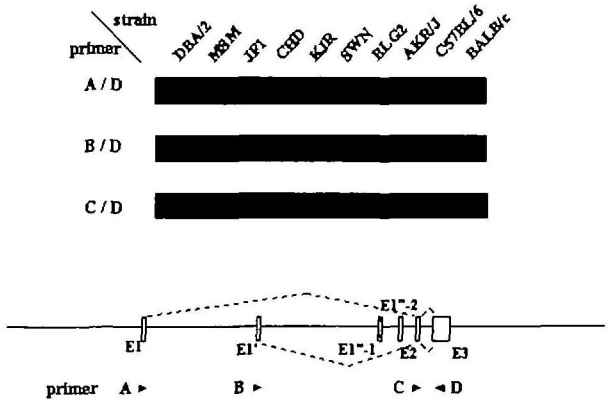


Fig. 3. **RT-PCR analysis of GNT transcripts in the mouse kidney.** In the genomic organization of the GNT gene (shown below), open boxes indicate exons. E1, exon 1; E1′, exon 1′; E1″-1, exon 1″-1; E1″-2, exon 1″-2; E2, exon 2; and E3, exon 3. Kidney-type mRNA utilizes E1 and the ubiquitous-type mRNA utilizes E1′. E1″-1 and -2 have been reported by Warren (*31*). E2 and E3 are common to these mRNAs.

shima, Kagoshima prefecture expresses GL-Y.

These results clearly show that the recessive phenotype of the *Gsl5* gene in DBA/2 mice is derived from Asian wild mice, *M. m. musculus*, and that both phenotypes, dominant and recessive phenotypes, in terms of the GL-Y expression are found in Japan.

RT-PCR Analysis of GNT Transcripts—RT-PCR performed with two sets of primers differentiates the kidney-type from the ubiquitous-type transcript of GNT. The kidney-type transcript, amplified with the A and D primers, is clearly detected in the kidneys of AKR/J, C57BL/6, and BALB/c. This transcript is not detected in the kidneys of CHD/Ms, KJR/Ms, SWN/Ms, and BLG2/Ms mice, derived from *M. m. musculus*, or the kidneys of JF1/Ms and MSM/Ms mice, derived from *M. m. molossinus*. This defect is the same as in the case of the DBA/2 strain (Fig. 3). The ubiquitous-type transcript, amplified by the B and D primers, is detected in all strains. These results suggest that both *M. m. musculus* and *M. m. molossinus* carry recessive *Gsl5* alleles identical to that of DBA/2.

The 5′ Upstream Sequences of the GNT Gene—We sequenced 1.2-kb-long DNA fragments amplified with primers UP1 and UP2. MSM/Ms, JF1/Ms, CHD/Ms, KJR/Ms, and SWN/Ms mice, derived from Asian *M. m. musculus* and *M. m. molossinus*, and the inbred strain CBA/J, carry a sequence identical to that in DBA/2 mice (Fig. 4). In addi-

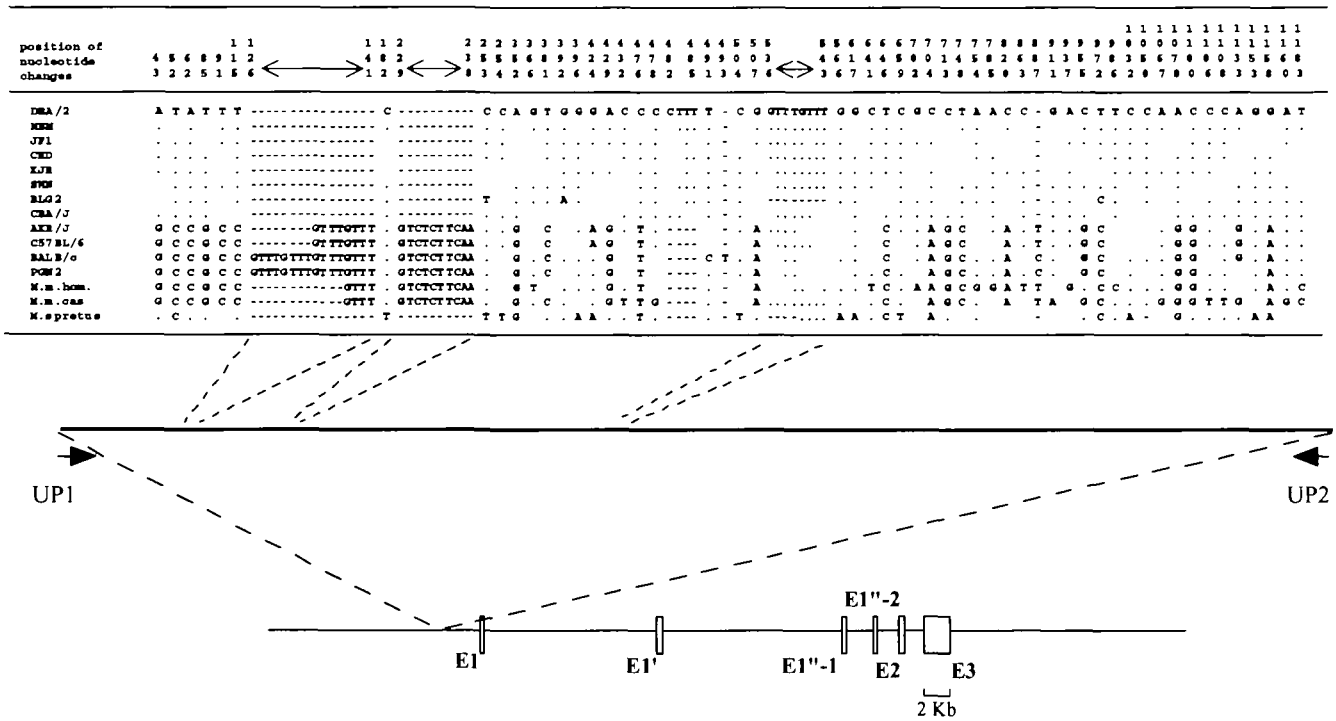


Fig. 4. **Substitutions in the 5′ upstream sequence of the GNT gene.** The upper table shows nucleotide changes compared to the DBA/2 strain. Dots represent nucleotides identical to those of DBA/2 mice, and dashes indicate deletions.

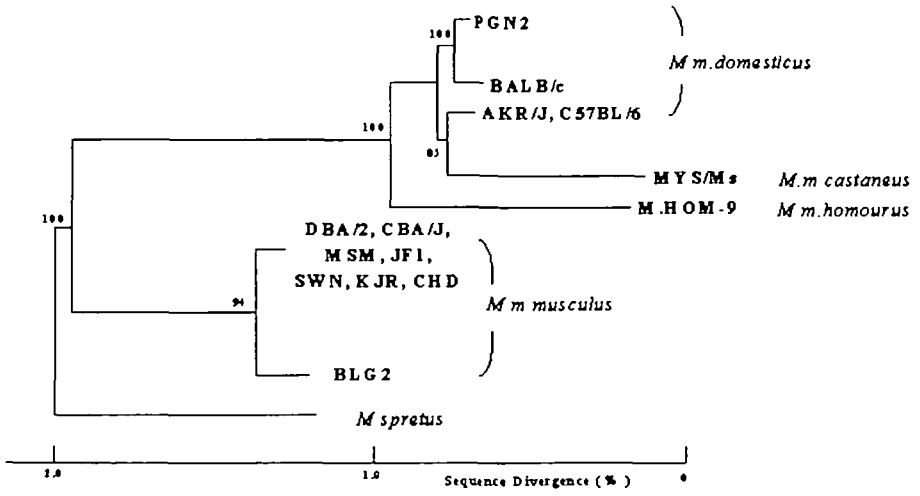
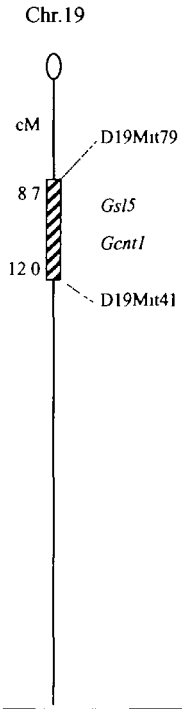


Fig. 5. **Phylogenetic tree based on polymorphisms in the 5′ upstream sequences of the GNT genes.**

TABLE II. **Polymorphism of microsatellite markers observed in SSLP analysis.**



Marker	cM	Strain polymorphism
D19Mit32	0.0	B6=BALB=D2>MSM=JF1
D19Mit93	0.0	B6=D2>BALB>MSM=JF1
D19Mit78	2.2	B6=BALB=D2=MSM=JF1
D19Mit68	3.3	B6>D2>BALB<MSM=JF1
D19Mit22	7.7	B6=BALB=D2>MSM>JF1
D19Mit79	8.7	B6=**D2=MSM=JF1**<BALB
D19Mit28	9.8	B6=BALB=D2=MSM=JF1
D19Mit128	10.9	**D2=MSM=JF1**>B6>BALB
D19Mit41	12.0	B6>BALB=**D2=MSM=JF1**
D19Mit110	13.1	MSM=JF1>B6=BALB>D2
D19Mit15	16.4	B6=BALB=D2=MSM=JF1
D19Mit40	17.5	B6=D2>BALB>MSM=JF1
D19Mit85	18.6	MSM=JF1>B6>BALB=D2
D19Mit30	18.6	BALB=MSM=JF1<B6=D2
D19Mit86	18.6	B6=BALB=D2<MSM=JF1
D19Mit39	19.7	B6=BALB=D2<MSM<JF1
D19Mit12	24.0	B6=BALB=D2<MSM=JF1
D19Mit13	24.0	B6=BALB=D2<MSM=JF1
D19Mit11	28.4	B6=BALB=D2=MSM=JF1
D19Mit9	35.0	B6=BALB=D2>MSM=JF1
D19Mit4	35.0	B6=BALB=D2>MSM=JF1
D19Mit72	57.9	B6=BALB=D2>MSM=JF1

B6, C57BL/6; BALB, BALB/c; D2, DBA/2; MSM, MSM/Ms; and JF1, JF1/MS. =, < and > indicate that PCR product is the same size, shorter and longer, respectively. cM, centimorgan. The shaded box on chromosome 19 indicates the region in which the SSLP allele is shared by DBA/2, MSM and JF1. D19Mit78, D19Mit28, D19Mit15, and D19Mit11 are not polymorphic among the five strains, and can not be used as markers differentiating lineage.

tion, BLG2/Ms mice, which exhibit the recessive phenotype and are derived from the Bulgarian *M. m. musculus*, carry three substitutions and one deletion of 8 nucleotides comprising of two GTTT repeats. The other inbred strains, including AKR/J, C57BL/6, BALB/c, and PGN2/Ms, and the wild-derived strains *M. m. homourus* and *M. m. castaneus* contain considerable numbers of substitutions in comparison with DBA/2. These sequences have been submitted to DDBJ as accession numbers, AB074465 to AB074479. These results confirm that the DBA/2 allele, in terms of the 5′ upstream sequences of GNT, is identical to that of *M. m. musculus* and *M. m. molossinus*.

Based on these sequence data, we calculated the sequence divergence by the two-parameter method of Kimura (*20*) using DNADIST on PHYLIP software (Version 3.5c) (*21*). The sequence of *M. spretus* was used as an outgroup and we constructed a phylogenetic tree using the Neighbor-Joining method (*22*). The bootstrap value and the length of each branch were calculated and included in the phylogenetic tree (Fig. 5). DBA/2 and CBA/J are included within the same group as MSM/Ms, JF1/Ms, SWN/Ms, KJR/Ms, and CHD/Ms mice. AKR/J, C57BL/6, BALB/c, and PGN2/Ms mice are located on an additional branch, closer to *M. m. castaneus* than to either *M. m. musculus* or DBA/2. The phylogenetic tree demonstrates that mutations to differentiate the dominant and recessive types of the *Gsl5* gene occurred approximately one million years ago during the subspeciation of *M. musculus*, if *M. spretus* is taken as the outgroup for the calculation.

TABLE III. **A matrix diagram of SSLP between two strains.**

P/T %	JF1	MSM	DBA/2	C57BL/6	BALB/c
JF1		9.0	68.2	77.3	72.7
MSM	2/22		68.2	77.3	72.7
DBA/2	15/22	15/22		22.7	31.8
C57BL/6	17/22	17/22	5/22		29.6
BALB/c	16/22	16/22	7/22	8/22	

P: number of polymorphic markers between two strains, T: total number of microsatellite markers examined. The right upper half displays the calculated frequency of polymorphism (%).

Simple Sequence Length Polymorphism (SSLP) Analysis—Gsl5 was mapped at 12 cM from the centromere on chromosome 19 in our previous study (*15*). PCR-SSLP analysis of 22 microsatellite markers on chromosome 19, as shown in Table II, reveals no polymorphic differences among DBA/2, MSM/Ms, and JF1/Ms in the segment covered by D19Mit79 and D19Mit41, which are located 8.7 cM and 12.0 cM from the centromere, respectively. However, as to markers other than those of this segment, the difference between DBA/2 and MSM/Ms or JF1/Ms is greater than that between DBA/2 and either C57BL/6 or BALB/c. A summarization of the results is shown in Table III as the polymorphic differences of chromosome 19 markers in total. These results together indicate that the major genetic background of DBA/2 is *M. m. domesticus*, like C57BL/6 and BALB/c, except for a small segment spanning from D19Mit79 to D19Mit41. This segment includes the 5′ upstream region of GNT and the GNT gene, and the segment in DBA/2 mice was introduced from the Asian subspecies, *M. m. musculus*.

DISCUSSION

In the mouse, *Gsl5* controls the transcription of GNT, resulting in the regulation of the kidney-specific expression of a unique glycolipid, GL-Y (Fig. 1). A few inbred strains, including DBA/2, possess a recessive allele preventing the expression of GL-Y in the kidney. DBA/1 and DBA/2 substrains were established from crosses of an ancestral DBA strain in 1929–1930, and the DBA-type allele was passed to in both substrains. CBA/J, CBA/CaHN, and TT6/Le were established by mating with DBA. Therefore, it is concluded that the recessive phenotype of GL-Y in these strains was introduced from DBA (Table I).

The DBA-type allele is common in Asian wild-derived mice. The present study of wild-derived strains and stocks reveals that the mice of a subspecies, *M. m. musculus*, distributed from Eastern Europe to Far East Asia and Central Asia, exclusively possess the DBA-type allele. Japanese wild mice, *M. m. molossinus*, are known to be a hybrid of *M. m. castaneus* and *M. m. musculus*, according to RFLP analysis of mitochondrial DNA (*23*). *M. m. castaneus* exclu-

sively possesses the dominant allele of *Gsl5*. Mice of the Iranian *M. m. bactrianus* subspecies also carry the DBA-type allele. mtDNA RFLP analysis demonstrated that the Iranian mouse possesses an mtDNA haplotype derived from *M. m. musculus* (unpublished data), suggesting that the crossing of this mouse with *M. m. musculus* occurred at the border of the territory between *M. m. musculus* and *M. m. bacterianus*. It is reasonable to conclude that the origin of the recessive allele in DBA/2 mice is *M. m. musculus* because historical documents prove that Japanese fancy mice were exported to and bred in Europe.

In the establishment of laboratory inbred strains, Japanese fancy mice, possibly derived from *M. m. molossinus*, were crossed with European or North American mice. In that process, some of their alleles were introduced into current laboratory inbred mice. Most classical common inbred strains carry an *M. m. musculus* Y chromosome (*24, 25*). The original *p* allele of pink-eyed dilution (*26*) and the piebald mutation are also derived from Asian mice, most likely the *M. m. molossinus* strain (*27*). The *Gsl5* allele is an additional instance where an allele, the *Gsl5* allele of the DBA/2 strain, originated from *M. m. musculus*.

The comparison of 5′ upstream GNT gene sequences confirms the above conclusion. The 5′ upstream 1.2 kb sequence of DBA/2 is identical to that of CBA/J and the other wild-derived strains carrying the DBA-type allele, except for BLG2. BLG2 exhibits three single base substitutions and these would be a good control to estimate the mutation rate in the same group of the allele. On the contrary, extensive polymorphisms were discovered between the sequences of the DBA- and the dominant-type alleles. We constructed a phylogenetic tree by the Neighbor-Joining method using single base substitutions (Fig. 5). At least two major clades exist in the tree; one clade includes *M. m. musculus* and the other includes *M. m. domesticus*, *M. m. castaneus*, and *M. m. homourus*. The DBA-type allele is restricted to the former clade, supporting the conclusion that the DBA allele is derived from *M. m. musculus*.

Based on DNA polymorphisms, the simultaneous subspeciation of the four major subspecies, *M. m. musculus*, *M. m. domesticus*, *M. m. castaneus*, and *M. m. bactrianus*, is estimated to have occurred approximately one million years ago (*18, 19, 28*). During subspeciation, multiple mutations accumulated in the respective genomes. Mutation of the DBA-type *Gsl5* allele may have occurred in ancestors of the *M. m. musculus* subspecies following subspeciation, and the mutation would have then spread to the *M. m. musculus* population.

Gsl5 controls the expression of both GL-Y and glycoproteins carrying a specific carbohydrate structure, Galβ1-4(Fucα1-3)GlcNAcβ1-6(Galβ1-3)GalNAcα1-Ser/Thr (core 2-Le^x), in the proximal tubular cells of mouse kidney (*16*). The key enzyme, GNT, is a mouse homologue of the human core 2 β-1,6-GlcNAc transferase 1 (C2GnT1) (*29*). C2GnT1 reacts with the core 1 structure (Galβ1-3GalNAcα-Ser/Thr) of *O*-linked glycoproteins to create the core 2 structure [GlcNAcβ1-6(Galβ1-3)GalNAcα-Ser/Thr]. The branch extending from the GlcNAc moiety can carry terminal determinant groups such as Lewis^x, sialyl-Lewis^x and 6′sulfo-sialyl Lewis^x structures, the ligands of the selectin family of carbohydrate recognition molecules. DBA/2 mice cannot express the core 2 structure in a kidney-specific manner because of the defective *Gsl5* gene. The physiological role of *Gsl5* in mouse kidney is as yet undetermined, because so far we have not found any apparent kidney dysfunctions in DBA/2 mice. However, we consider that close examination focusing on the functions of particular molecules, especially under some stressful conditions, may reveal functional differences in the molecules modified by the *Gsl5* gene.

Both loci of the GNT gene (*Gcnt1*) and *Gsl5* have been mapped to mouse chromosome 19 with no recombination, suggesting that the mutation(s) responsible for the *Gsl5* phenotype must occur in the kidney-specific promoter region of the GNT gene (*15*). Recent BAC-transgenesis supports this hypothesis; the DBA/2 phenotype was rescued *in vivo* by the transgenesis of a 150 kb BAC clone, which includes the GNT gene and the *Gsl5* dominant gene (unpublished data). The present paper suggests that mutations to create the functional *Gsl5* gene occurred approximately one million years ago during the subspeciation of *M. musculus*, and provides a case for the elucidation of mutations responsible for the creation of functional genes involved in tissue-specific transcriptional regulation. The elucidation of the mutation(s) responsible for the *Gsl5* phenotype is the goal of our current investigation.

We are grateful to Drs. Moriwaki (RIKEN BioResource Center), Tsuchiya (Miyazaki Medical College), and Shiroishi (National Institute of Genetics) for providing various wild-derived mice.

REFERENCES

1. Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**, 97–130
2. Tsuji, S. (1996) Molecular cloning and functional analysis of sialyltransferases. *J. Biochem.* **120**, 1–13
3. Shaper, N.L., Harduin-Lepers, A., and Shaper, J.H. (1994) Male germ cell expression of murine β4-galactosyltransferase. A 796-base pair genomic region containing two cAMP-responsive element (Cre)-like elements, mediates male germ cell-specific expression in transgenic mice. *J. Biol. Chem.* **269**, 25165–25171
4. Kolter, T. and Sandhoff, K. (1998) Recent advances in the biochemistry of sphingolipidosis. *Braun Pathol.* **8**, 79–100
5. Fukuda, M.N., Masri, K.A., Dell, A., Luzzatto, L., and Moremen, K.W. (1990) Incomplete synthesis of *N*-glycans in congenital dyserythropoietic anemia type II caused by a defect in the gene encoding α-mannosidase II. *Proc. Natl. Acad. Sci. USA* **87**, 7443–7447
6. Abeijon, C., Robbins, P.W., and Hirschberg, C.B. (1996) Molecular cloning of the Golgi apparatus uridine diphosphate-*N*-acetylglucosamine transporter from *Kluyveromyces lactis*. *Proc. Natl. Acad. Sci. USA* **93**, 5963–5968
7. Miura, N., Ishida, N., Hoshino, M., Yamauchi, M., Hara, T., Ayusawa, D., and Kawakita, M. (1996) Human UDP-galactose translocator: Molecular cloning of a complementary DNA that complements the genetic defect of a mutant cell line deficient in UDP-galactose translocator. *J. Biochem.* **120**, 236–241
8. Suzuki, A., Sekine, M., Nakamura, K., Kawano, T., Sakaizumi, M., and Kozutsumi, Y. (1994) Genes controlling the expression of glycolipids characterized by wild mice. *Genetics in Wild Mice* (K. Moriwaki *et al.*, eds.) pp. 57–65, Japan Scientific Societies Press/Karger, Tokyo
9. Sekine, M., Suzuki, M., Inagaki, F., Suzuki, A., and Yamakawa, T. (1987) A new extended globoglycolipid carrying the stage specific embryonic antigen-1 (SSEA-1) determinant in mouse kidney. *J. Biochem.* **101**, 553–562
10. Sekine, M., Yamakawa, T., and Suzuki, A. (1987) Genetic control of the expression of two extended globoglycolipids carrying either the stage specific embryonic antigen-1 or -3 determinant in mouse kidney. *J. Biochem.* **101**, 563–568
11. Sekine, M., Nakamura, K., Suzuki, M., Inagaki, F., Yamakawa,

T., and Suzuki, A. (1988) A single autosomal gene controlling the expression of the extended globoglycolipid carrying SSEA-1 determinant is responsible for the expression of two extended globogangliosides. *J. Biochem.* **103**, 722–729
12. Sekine, M., Sakaizumi, M., Moriwaki, K., Yamakawa, T., and Suzuki, A. (1989) Two genes controlling the expression of extended globoglycolipids in mouse kidney are closely linked to each other on chromosome 19. *J. Biochem.* **105**, 680–683
13. Sekine, M., Hashimoto, Y., Inagaki, F., Yamakawa, T., and Suzuki, A. (1990) The expression of IV^{6}β[Galβ1-4(Fucα1-3)-GlcNAc]-Gb5Cer in mouse kidney is controlled by the *Gsl5* gene through regulation of UDP-GlcNAc:Gb5Cer β1-6 *N*-acetylglucosaminyltransferase activity. *J. Biochem.* **108**, 103–108
14. Sekine, M., Hashimoto, Y., Suzuki, M., Inagaki, F., Takio, K., and Suzuki, A. (1994) Purification and characterization of UDP-GlcNAc:IV^{3}βGal-Gb4Cer β-1,6-GlcNAc transferase from mouse kidney. *J. Biol. Chem.* **269**, 31143–31148
15. Sekine, M., Nara, K., and Suzuki, A. (1997) Tissue-specific regulation of mouse core2 β-1,6-*N*-acetylglucosaminyltransferase. *J. Biol. Chem.* **272**, 27246–27252
16. Sekine, M., Taya, C., Kikkawa, Y., Yonekawa, H., Takenaka, M., Matsuoka, Y., Imai, E., Izawa, M., Kannagi, R., and Suzuki, A. (2001) Regulation of mouse kidney tubular epithelial cell-specific expression of core2 GlcNAc transferase. *Eur. J. Biochem.* **268**, 1129–1135
17. Morse, H.C. (1981) The laboratory mouse—A historical perspective in *The Mouse in Biomedical Research* (Foster, H.L., Small, J.D., and Fox, J.G., eds.) Vol. I, pp. 1–16, Academic Press, New York
18. Yonekawa, H., Moriwaki, K., Gotoh, O., Watanabe, J., Hayashi, J., Miyashita, N., Petras, M.L., and Tagashira, Y. (1980) Relationship between laboratory mice and the subspecies *Mus musculus domesticus* based on restriction endonuclease cleavage patterns of mitochondrial DNA. *Jpn. J. Genet.* **55**, 289–296
19. Silver, L.M. (1995) *Mouse Genetics, Concepts and Applications*., Oxford Univ. Press, Oxford
20. Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120
21. Felsenstein, J. (1993) "PHYLIP (phylogenetic influence package)," version 3.5c, distributed by the author, Department of Genetics, University of Washington, Seattle
22. Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425
23. Yonekawa, H., Moriwaki, K., Gotoh, O., Miyashita, N., Matsushima, Y., Shi, L., Cho, W.S., Zhen, X-L., and Tagashira, Y. (1988) Hybrid origin of Japanese mice "*Mus musculus molossinus*": Evidence from restriction analysis of mitochondrial DNA. *Mol. Biol. Evol.* **5**, 63–78
24. Bishop, C.E., Boursot, P., Baron, B., Bonhomme, F., and Hatat, D. (1985) Most classical *Mus musculus domesticus* laboratory mouse strains carry a *Mus musculus musculus* Y chromosome. *Nature* **315**, 70–72
25. Nagamine, C.M., Nishioka, Y., Moriwaki, K., Boursot, P., Bonhomme, F., and Lau, Y.F.C. (1992) The *musculus*-type Y chromosome of the laboratory mouse is of Asian origin. *Mamm. Genome* **3**, 84–91
26. Brilliant, M.H., Ching, A., Nakatsu, Y., and Eicher, E.M. (1994) The original pink-eyed dilution mutation (*p*) arose in Asiatic mice: Implications for the *H4* minor histocompatibility antigen, *Myod1* regulation and the origin of inbred strains. *Genetics* **138**, 203–211
27. Koide, T., Moriwaki, K., Uchida, K., Mita, A., Sagai, T., Yonekawa, H., Katoh, H., Miyashita, N., Tsuchiya, K., Nielsen, T.J., and Shiroishi, T. (1998) A new inbred strain JF1 established from Japanese fancy mouse carrying the classic piebald allele. *Mamm. Genome* **9**, 15–19
28. Kikkawa, Y., Miura, I., Takahama, S., Wakana, S., Yamazaki, Y., Moriwaki, K., Shiroishi, T., and Yonekawa, H. (2001) Microsatellite database for MSM/Ms and JF1/Ms, *molossinus*-derived inbred strains. *Mamm. Genome* **12**, 750–752
29. Bierhuizen, M.F. and Fukuda, M. (1992) Expression cloning of a cDNA encoding UDP-GlcNAc:Galβ1-3GalNAc-R(GlcNAc to GalNAc)β1-6GlcNAc transferase by gene transfer into CHO cells expressing polyoma large tumor antigen. *Proc. Natl. Acad. Sci. USA* **89**, 9326–9330
30. Lyon, M.F. and Searle, A.G. (1989) *Genetic Variants and Strains of the Laboratory Mouse*, 2nd ed., Oxford Univ. Press, Oxford
31. Warren, C.E. (1995) Direct submission to GeneBank, accession number: MMU19265