Murine Equivalent of the Human Histo-blood Group ABO Gene Is a cis-AB Gene and Encodes a Glycosyltransferase with Both A and B Transferase Activity*

Received for publication, November 30, 2000, and in revised form, January 12, 2001
Published, JBC Papers in Press, January 23, 2001, DOI 10.1074/jbc.M010805200

Miyako Yamamoto‡, Xiao-Hong Lin‡, Yoshihiko Kominato§, Yukiko Hata§, Reiko Noda¶, Naruya Saitou‡, and Fumiichiro Yamamoto‡‡

From ‡The Burnham Institute, La Jolla Cancer Research Center, La Jolla, California 92037, the §Department of Legal Medicine, Faculty of Medicine, Toyama Medical & Pharmaceutical University, Toyama 930-01, Japan, and the ¶Laboratory of Evolutionary Genetics, National Institute of Genetics, Mishima 411, Japan

We have cloned murine genomic and complementary DNA that is equivalent to the human ABO gene. The murine gene consists of at least six coding exons and spans at least 11 kilobase pairs. Exon-intron boundaries are similar to those of the human gene. Unlike human A and B genes that encode two distinct glycosyltransferases with different donor nucleotide-sugar specificities, the murine gene is a cis-AB gene that encodes an enzyme with both A and B transferase activities, and this cis-AB gene prevails in the mouse population. Cloning of the murine AB gene may be helpful in establishing a mouse model system to assess the functionality of the ABO genes in the future.

Histo-blood group A/B antigens are clinically important antigens in blood transfusion and organ transplantation. These antigens are oloigosaccharide antigens whose immunodominant structures are defined as GalNAc α1→3 (Fuc α1→2) Gal- and Gal α1→3 (Fuc α1→2) Gal- for A and B antigen, respectively. Functional alleles at the ABO locus encode enzymes that catalyze the final step of synthesis. A alleles encode for A transferase, which transfers the GalNAc residues from the UDP-GalNAc nucleotide-sugar to the galactose residue of the acceptor H substrates defined by Fuc α1→2 Gal-. B alleles encode for B transferase that transfers the galactose residue from UDP-galactose to the same H substrates. O alleles are nonfunctional, null alleles. During the past decade, we have been studying the molecular genetic basis of the histo-blood group ABO system (1). From a human gastric carcinoma cell line cDNA library, we were able to clone human A transferase cDNA (2) based on the partial amino acid sequence of the soluble form of A transferase purified from human lung (3). Using cross-hybridization with A transferase cDNA probes, we then cloned B transferase cDNA and nonfunctional O allelic cDNA from cDNA libraries made with RNA from colon adenocarcinoma cell lines that exhibited different ABO phenotypes (4). Possible allele-specific mutations were identified. Four amino acid substitutions were discovered between A and B transferases. O alleles were more homologous to A alleles than to B alleles. A single base deletion was found near the N-terminus of the coding sequence in most of the O alleles, which caused the codon frame to shift. This resulted in a truncated protein without glycosyltransferase activity. In addition to the three major alleles (A1, B, and O), we also identified mutations that modified the enzymatic activity by determination of the partial nucleotide sequences of subgroup alleles (A2, A3, A3, and B3) (5–7). We also elucidated the molecular mechanisms of two phenomena named cis-AB and B(A) (7, 8). Although the incidence was low, another type of O allele was discovered that lacked the single base deletion but contained an amino acid substitution at the residue crucial for nucleotide-sugar recognition/binding (9). Although no functional analyses have been performed to disprove polymorphism, others have reported additional alterations (10–15). The nucleotide and deduced amino acid sequences of a variety of ABO alleles are posted on the Blood Group Antigen Gene Mutation Database developed by Blumenfeld and colleagues (available on the World Wide Web).

A/B antigens are not restricted to humans but are widely present in nature (16). We therefore investigated the presence/absence of homologous sequence(s) in the genomes of other species of organisms (17). Hybridization of zoo blot, using the radiolabeled human A transferase cDNA probe, showed weak signals in chicken genomic DNA but strong signals, comparable with the signal detected in human DNA, in genomic DNA from mice and other mammals. No signals were detected in genomic DNA from lower species of organisms in the evolutionary tree. We next determined the partial nucleotide sequences of the primate ABO genes (17). The glycosyltransferases responsible for A or B phenotypes in primates were shown to conserve amino acid substitutions corresponding to codons 266 and 268 in humans. A similar study was also reported by others (18). Through comparative sequence analyses of the ABO genes from humans and apes, we and others proposed a convergent hypothesis of evolution that ABO genes arose from independent mutations after the speciation of humans and apes (19, 20).

No apparent disadvantages are recognized among any of the phenotypes involving the ABO polymorphism. Hemolytic disease of newborns may be a natural selection against specific combinations of blood groups between the mother and fetus. However, serious incompatibility cases are rare with ABO, since the natural antibodies against A and B antigens are mostly IgM and do not cross the placenta. Although some anti-A, B antibodies are IgG and capable of crossing the pla-

* This study was supported in part by funds from The Burnham Institute, the Department of Defense Breast Cancer Research Program (DAMD 17-98-1-8168), and from Kyowa Hakko Kogyo Co. Ltd. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

The nucleotide sequence reported in this paper has been submitted to the DDBJ/GenBank™/EMBL Data Bank with accession numbers AB041038 and AB041039.
centa, A/B antigens are not well developed in fetuses. Therefore, little damage is done. There should be some reason for the existence of ABO polymorphism in the population. It has been speculated that the possible role of the ABO system is to provide resistance against infection (21). Actually, Le<sup>b</sup> (Fucα1–2 Gal β1–3 (Fucα1–4) GlcNAc<sup>−</sup>), an ABO-related structure, was demonstrated to be the receptor for a Gram-negative bacillus, *Helicobacter pylori*, a causative agent for gastritis, peptic ulcer, and possibly gastric cancer (22). A and B transferases modify the Le<sup>b</sup> structure into ALe<sup>b</sup> and BLE<sup>b</sup> structures, which *H. pylori* does not bind to in vitro. This may explain the earlier observation that group O individuals have a higher incidence of stomach ulcer than individuals in any other group (21). Antibodies against the α1–3 Gal epitope (Gal α1–3 Gal–) were demonstrated experimentally to block the interspecies infection of certain retroviruses (23). From this result, anti-A and anti-B antibodies have been suspected to play a role in inhibiting the epidemics of certain infections. Some type of selection based on the advantage/disadvantage of playing antigens/antibodies may have been operating at the ABO locus to secure the survival of species from extinction during evolution. The report that an anti-A monoclonal antibody neutralized human immunodeficiency virus particles produced by lymphocytes from group A individuals but not from group B or O individuals (24) may support this hypothesis. To experimentally assess the functionality of the ABO genes, establishing an animal model is critical. As an initial step, we cloned and characterized the murine ABO gene equivalent.

**EXPERIMENTAL PROCEDURES**

**Materials**

Mouse genomic DNA library (ML1044), which was constructed by replacing the internal BamHI–BamHI stuffer fragment of the λ EMBL3 SP6/T7 vector with Mol1-partially cleaved genomic DNA fragments of the BALB/c strain of mouse, was purchased from CLONTECH (Palo Alto, CA). A Marathon cDNA Amplification Kit and an AdvanTage PCR<sup>Ⅰ</sup> cloning kit were also from CLONTECH. A GeneClean kit was purchased from Bio 101 (La Jolla, CA), and pT7T3α18 plasmid vector, α-S-dGTP, α-S-dCTP, and S-300 MicroSpin columns were from Amersham Pharmacia Biotech. LipofectAMINE was purchased from Life Technologies, Inc. [<sup>14</sup>C]UDP-GalNAc, [<sup>3</sup>H]UDP-galactose, and [α-<sup>32</sup>P]dCTP were from PerkinElmer Life Sciences, and 2'-deoxy- and 2'-fluorosylactosamine was from Oxford Glycosystems (Rosedale, NY) and from Calbiochem. Murine anti-A and anti-B monoclonal antibody mixtures were purchased from Ortho Diagnostic Systems (Raritan, NJ). Biotinylated *Ulex europaeus* agglutinin I, Vectastain ABC kit, and 4-chloro-1-naphthol chromogen were from Vector Laboratories (Beverly, MA), or Roche Molecular Biochemicals. dRhodamine dye terminator cycle sequencing ready reagents and Southern hybridization transfer media were then arranged using a human A transferase cDNA probe by the plaque hybridization method (25). Radiolabeled probe was prepared by the random hexamer primer method, using a Prime-It II kit and [α-<sup>32</sup>P]dCTP (26). After four rounds of screening, individual clones were isolated. Phage DNA was prepared, cleaved with restriction endonucleases, gel-electrophoresed, and Southern hybridized. This hybridization was then performed to construct restriction enzyme cleavage maps.

**Subcloning, Nested Deletion, and DNA Sequencing**—DNA from MAB0616 phage clone was cleaved with *Hin*dIII and SalI and subcloned into the BlueScript SKM13<sup>Ⅱ</sup> vector. Nested deletions were constructed by the Exo I-mung bean nuclease method (27). Wherever no unique 3'–overhang restriction sites were available, the thioderivatives were prepared with Klenow enzyme using α-S-dGTP and α-S-dCTP before *Exo*I treatment. After transformation of *E. coli* XL1-blue strain, plasmid DNA was prepared from individual clones and analyzed for insert size. The nucleotide sequences were determined by Sanger's dideoxy chain termination method using the dRhodamine dye terminator cycle sequencing ready reaction kit (28). Sequences were aligned using Lasergene SeqMan II sequencing project management software.

5'-RACE cDNA Cloning—RNA from the CMT-93 rectal carcinoma cell line (ATCC 223-CCL), established from a C57BL strain of mouse, was prepared and used for the 5'-RACE experiments (29). We followed the Marathon cDNA amplification protocol provided by the manufacturer. Briefly, the first strand of cDNA was synthesized from RNA using Moloney murine leukemia virus reverse transcriptase and Marathon cDNA synthesis primer. After the second strand was synthesized with RNase H, *E. coli* DNA polymerase I, and *E. coli* DNA ligase, the Marathon cDNA adaptor was ligated. Nested PCR was performed, first with M13 primer and then with AP1 and AP2 primers and then with MY-1 and MY-2 primers. AP1 and AP2 primers were provided in the kit. The nucleotide sequences of MY-1 and MY-2 primers were complementary to the sequences in the coding region of the murine ABO genes. Their sequences were as follows: 5'TTATTTTCTCTGGCTTGGCTCTTGGACAC and 5'TCAGTGGACACAGGCTCAATGCCGT for MY-1 and 2, respectively. PCR products were electrophoresed through a 3% agarose gel, and the DNA was gel-purified using the GeneClean kit. DNA fragments were then ligated with pT-Adv vector from the AdvanTage PCR cloning kit by the T-A cloning method. Nucleotide sequences of the inserts were determined.

**Preparation of Eukaryotic Expression Constructs**—The BamHI–XhoI fragment containing the last coding exon of the murine ABO gene was first subcloned from a murine ABO genomic clone, MAB0616, into the pSP6.1 plasmid vector. The BamHI site was located in the intron preceding the last coding exon of the murine ABO gene. The XhoI site was in the λ EMBL3 SP6/T7 vector next to the BamHI site used to accommodate genomic DNA. This construct was then digested with *Sst*I and *Sna*BI. The *Sst*I site was within the pT7T3α18 plasmid vector, and the *Sna*BI site was located downstream of the stop codon of the mouse ABO gene. The *Sst*I–*Sna*BI fragment containing the coding sequence in the last coding exon of the mouse ABO gene was then isolated. The human B transferrase expression construct with intron, pBBI (30), was cleaved with BamHI, blunt-ended by the Klenow filling-in reaction, and then digested with *Sst*I. The *Sst*I site was in the intron preceding coding exon 7 of the human ABO gene; *Bam*HI was in the eukaryotic expression vector (originally pBS-5). The *Sst*I–*Bam*HI vector fragment containing the human B transferrase cDNA sequence of exons 1–6 was then ligated to the mouse *Sst*I–*Sna*BI fragment to produce the human-mouse chimeric gene (*p*Human-mouse chimera). A murine cDNA eukaryotic expression construct was then constructed by replacing the EcoRI–A/HII fragment from the chimeric construct with the EcoRI–A/IHI fragment from the 5'-RACE clone in the *Bam*HI vector. The EcoRI site of the clone was in the plasmid and located 5'–upstream of the *cDNA* end. The *A/HII* was in the last coding exon. The EcoRI site in the pBS-5 vector was located downstream of the SV40 early promoter and upstream of the human *cDNA* sequence. The resultant construct (*p*Mouse) contained the entire coding sequence of the mouse ABO gene cDNA.

**DNA Transfection and Enzymatic Assays**—Plasmid DNA was prepared by the SDS-alkaline method (25). The HeLa cell line derived from a human adenocarcinoma of uterus was used as a recipient of transient DNA transfection analyses. The HeLa cells express H antigens on their cell surfaces and have been successfully used in similar transfection experiments of A and B transferrase expression constructs (5, 30, 31). Following the manufacturer’s protocol, we used LipofectAMINE for transfection. Seventy-two hours after transfection, the cells were incubated in buffer (0.1 M NaCl, 25 mM sodium cacodylate, 10 mM MnCl<sub>2</sub>, and 0.1% Triton X-100). A/B transfection activity was determined by measuring the transfer of carbon-14 from [<sup>14</sup>C]UDP-GalNAc or [<sup>14</sup>C]UDP-galactose to the acceptor substrate 2'-fucosylactose, as described previously (31). After incubation, the reaction products were separated from unincorporated nucleoside-sugars by AG1-X8 anion exchange column chromatography.
The incorporation of radioactivity was determined using a scintillation counter.

PCR Amplification and Nucleotide Sequence Determination of the ABO Gene Fragments from Several Strains of Mouse Species and Subspecies—Murine submaxillary glands were used for the expression analyses of A and B transferases. A and B transferase activities were measured by the incorporation of carbon-14 from [14C]UDP-GalNAc or [14C]UDP-galactose to 2-fucosyllactose. Although the same reaction conditions were used, the reaction product was separated from the precursor substrate by paper chromatography rather than column chromatography. Genomic DNA was prepared by the proteinase K-SDS method and used to amplify a DNA fragment derived from the murine ABO gene. The names and nucleotide sequences of the primers used were as follows: SN-16, 5′-GAGACTGCAAGAACACACTT; SN-17, 5′-CAATTGCCGTTGGCCCTGTC. The PCR-amplified DNA fragments were purified through chromatography using S-300 MicroSpin columns and subjected to direct DNA sequencing reactions with the BigDye cycle sequencing kit. After the sequencing reaction, DNA was purified and then analyzed using an ABI Prism 377 automatic DNA sequencer.

Detection of ABH Antigen Expression—Expression of the ABH antigens in murine submaxillary glands was examined immunologically using extracts spotted on a nitrocellulose membrane. Murine anti-A and anti-B monoclonal antibody mixtures were biotinylated using EZ-Link Sulfo-NHS-LC-Biotin, following the protocol provided by the manufacturer. After biotinylation, the unincorporated biotin was removed using Microcon 30 centrifugal filter devices. The submaxillary glands from C57BL and ICR strains of mice were homogenized in buffer containing 20 mM Tris-HCl (pH 7.5), 0.15 M NaCl, and 1% Triton X-100. After centrifugation, the supernatant was diluted with buffer containing 25 mM Tris-HCl (pH 7.5) and 0.1% SDS. The extract was then spotted onto a Duralose-UV membrane. As controls, the extracts similarly obtained from human colon adenocarcinoma SW48 cells (AB phenotype) and from group A and O porcine submaxillary glands were also spotted on the membrane. After drying for 15 min, the membrane was treated with 0.3% hydrogen peroxide and 0.3% fetal calf serum in phosphate-buffered saline for 5 min to block endogenous peroxidase activity. After washing, the membrane was incubated overnight in phosphate-buffered saline containing 4% bovine serum albumin at 4 °C. The membrane was then cut into four pieces, which were individually incubated with either biotinylated murine anti-A monoclonal antibody mixture, biotinylated murine anti-B monoclonal antibody mixture, biotinylated Ultrex europaeus agglutinin I, or bovine serum albumin (negative control) for 1 hr at room temperature. The filters were washed separately and then incubated collectively with the Elite ABC reagents for 15 min. After washing with phosphate-buffered saline, the membranes were treated with 4-chloro-1-naphthol substrate for color development.

RESULTS

Mouse Genome Contains the Human ABO Gene Equivalent—Certain mammalian cells exhibit a-3 Gal epitopes. The cDNA encoding a1,3-galactosyltransferase that synthesizes this epitope was cloned from cow (32), mouse (33), and pig (34, 35). Humans do not exhibit this epitope but possess the antibody against the epitope in sera (36). Human sequence corresponding to this gene was shown to be a pseudogene due to frameshifts and nonsense mutations (32, 37). A/B transferases utilize the galactose substrate with fucose, whereas a1,3-galactosyltransferase utilizes the substrate without fucose. ABO genes and a1,3-galactosyltransferase genes share significant homology at both the nucleotide and deduced amino acid sequence levels (30). Cloned canine cDNA encoding Forsmann glycolipid synthetase (UDP-GalNAc:globoside a1,3-N-acetyl-D-galactosaminyltransferase) also exhibited sequence homology (38). Therefore, these genes are believed to have derived from the same ancestral gene and constitute the ABO gene family. Southern hybridization experiments of murine genomic DNA showed different banding patterns when murine a1,3-galactosyltransferase cDNA probe and human A transferase cDNA probe were used (17, 33). Accordingly, the ABO gene sequence was assumed to exist in the mouse genome (17). Results from our cloning experiments of the murine ABO gene described here concluded that mice actually do possess an ABO gene equivalent.

Organization of the Murine Gene Is Similar to the Human Counterpart—We cloned the genomic DNA sequence encompassing most of the murine ABO structural gene. By screening 1 million phage plaques from a murine genomic DNA library, we obtained a total of nine independent clones that hybridized with the human A transferase CDNA probe. A preliminary mapping showed that two phage clones named MABO11 and MABO16 contained the entire coding sequence in the last cloning exon. Since the MABO16 clone contained the sequence farther upstream, this clone was used for the nucleotide sequence determination. The MABO11 clone containing the farther downstream sequence was used to construct a human-mouse ABO gene chimeric expression construct as well as a murine gene expression construct. We sequenced the entire insert in the MABO16 clone (~11.2 kilobase pairs) with more than 99.9% accuracy. Almost all of the coding sequence was contained in the sequenced region. The exon-intron boundaries were determined and are shown in Fig. 1, and A and B, represents two probable splicing patterns, although other possibilities still exist because the sequence encoding the first few amino acid residues has not yet been identified. There are six coding exons in Fig. 1A and seven in Fig. 1B. Approximately 4.0 kilobase pairs and 70 base pairs upstream of the splicing acceptor site of coding exon 2 (cEXON 2) in Fig. 1A, there was a CTTCAG sequence and a TGAATTCAG sequence, respectively. These sequences may be portions of the coding sequence, since they are found upstream in the cDNA preceding the sequence in cEXON 2. Fig. 1B depicts the case where GAATCTCAG of the latter TGAATTCAG sequence represents the sequence in the preceding exon. In that case, the acceptor site of cEXON 3 needs to shift 2 nucleotides upstream, which would break up the GT-AG rule of splice junctions. We determined the entire nucleotide sequence contained in the MABO16 clone, which included ~5.0 kilobase pairs of sequence upstream of the splicing acceptor site of cEXON2 in Fig. 1A. No sequence corresponding to the 5′-untranslated region was found. Therefore, the promoter region of the murine ABO genes must reside farther upstream. The sequence corresponding to human coding exons 3 and 4 was found in one exon in the mouse gene. However, the number of amino acid residues (19 amino acids) was much smaller than that of human exons 3 and 4 combined (53 amino acids). Further studies are needed, since there may be an alternative splicing that would divide this small exon into two smaller exons with an intron in between. The entire insert sequence in the MABO16 clone and the entire cDNA sequence have been deposited in the DNA Data Bank of Japan (DDBJ) (accession numbers AB401388 and AB401399).

The Nucleotide and Deduced Amino Acid Sequences of the Mouse ABO Gene Equivalent—The nucleotide and deduced amino acid sequences in the coding region of the murine cDNA were aligned with those of human A1-1 (A101) allele (accession number AF134412 in GenBank™) by combining the Clustal method (40) and the J. Hein method (41) using the MegAlign software. Results are shown in Fig. 1. Especially high homology was observed in the coding sequence in the last two coding exons. The percentages of identical nucleotide and amino acid residues in the last two coding exons were 78% (642/822) and 81% (222/273) between the two species, respectively. The amino acid sequence of the murine gene was also aligned with the amino acid sequences of human A and B transferases, mouse a1,3-galactosyltransferase, and canine Forsmann glycolipid synthetase. Results are shown in Fig. 2A. The percentages of the identical amino acid residues of the coding sequences in the last two coding exons are 47% (127 of 272) between the mouse ABO and a1,3-galactosyltransferase genes and 49% (132 of 272) between the mouse ABO gene and the dog Forsmann synthetase gene. Fig. 2B highlights the amino acid se-
sequences of the region important for the recognition/binding of nucleotide-sugars. The phylogenetic tree is shown in Fig. 2C. The cloned mouse gene was evolutionarily mapped closest to the human ABO gene. It was also mapped closer to the canine Forssman gene than the murine α1,3-galactosyltransferase gene.

Murine ABO Gene Encodes an Enzyme with both A and B Transferase Activities—We examined whether the isolated mouse ABO gene sequence could encode a functional glycosyltransferase. We first constructed a human-mouse chimeric construct in an eukaryotic expression vector pSG-5. A DNA fragment containing coding sequence in the last coding exon of the mouse genomic sequence was linked downstream of the human cDNA sequence of exons 1–6 in the human B transferase expression construct, pBBBB (30). When DNA from the chimera construct was transiently transfected into HeLa cells, the appearance of both A and B transferase activity was observed (Table I). Because there may be inactivating mutations in the upstream...
sequence, as observed in human O alleles, we constructed a mouse cDNA expression construct. The upstream sequence of the mouse AB transferase was compared with those of the human A and B transferases, mouse α1,3-galactosyltransferase (mouse GaIT), and canine Forsmann glycolipid synthetase (dog Forsman). The mouse AB transferase was 47 and 49% identical to the mouse α1,3-galactosyltransferase and the dog Forsman synthetase gene in the last two coding exons, respectively. B, the coding sequences around the amino acid residues important for the nucleotide-sugar recognition/binding are compared. The sequences from the human O03 allele that contains a missense mutation around this area (9) and the cis-AB01 allele with dual transferase activity (8) are also included. C, a phylogenetic tree of the ABO gene family is shown. The following sequences in GenBank were used for the computation (46) and neighbor-joining tree (39): marmoset GaIT (S71333), bovine GaIT (J04989), pig GaIT (L36152), mouse GaIT (M85153), and dog Forsman synthetase (U66140). Numbers on interior branches denote bootstrap probabilities in percent. Branch lengths are proportional to the numbers of amino acid substitutions, and a scale is given at the bottom of the tree. The dotted line denotes the probable root node of this tree. This node also corresponds to a gene duplication event producing both the α1,3-galactosyltransferase gene and ABO-related gene. The latter lineage experienced another gene duplication, resulting in creation of both Forsman glycolipid synthetase gene and ABO gene. The murine AB gene was mapped closest to the human ABO genes. It also mapped closer to the canine Forsman gene than the murine α1,3-galactosyltransferase gene.

**Murine ABO Gene Equivalent**

![Diagram](http://www.jbc.org)

**Fig. 2. Comparisons of the amino acid sequences of the coding regions among murine AB gene, human A1-1 (A101) and B (B101) genes, murine α1,3-galactosyltransferase gene, and canine Forsmann synthetase gene and an evolutionary tree of the ABO and related genes.** A, the amino acid sequence of the mouse AB transferase was compared with those of the human A and B transferases, mouse α1,3-galactosyltransferase (mouse GaIT), and canine Forsmann glycolipid synthetase (dog Forsman). The mouse AB transferase was 47 and 49% identical to the mouse α1,3-galactosyltransferase and the dog Forsman synthetase gene in the last two coding exons, respectively. B, the coding sequences around the amino acid residues important for the nucleotide-sugar recognition/binding are compared. The sequences from the human O03 allele that contains a missense mutation around this area (9) and the cis-AB01 allele with dual transferase activity (8) are also included. C, a phylogenetic tree of the ABO gene family is shown. The following sequences in GenBank were used for the computation (46) and neighbor-joining tree (39): marmoset GaIT (S71333), bovine GaIT (J04989), pig GaIT (L36152), mouse GaIT (M85153), and dog Forsman synthetase (U66140). Numbers on interior branches denote bootstrap probabilities in percent. Branch lengths are proportional to the numbers of amino acid substitutions, and a scale is given at the bottom of the tree. The dotted line denotes the probable root node of this tree. This node also corresponds to a gene duplication event producing both the α1,3-galactosyltransferase gene and ABO-related gene. The latter lineage experienced another gene duplication, resulting in creation of both Forsman glycolipid synthetase gene and ABO gene. The murine AB gene was mapped closest to the human ABO genes. It also mapped closer to the canine Forsman gene than the murine α1,3-galactosyltransferase gene.

**Murine ABO Gene Equivalent**

![Diagram](http://www.jbc.org)

**Fig. 2. Comparisons of the amino acid sequences of the coding regions among murine AB gene, human A1-1 (A101) and B (B101) genes, murine α1,3-galactosyltransferase gene, and canine Forsmann synthetase gene and an evolutionary tree of the ABO and related genes.** A, the amino acid sequence of the mouse AB transferase was compared with those of the human A and B transferases, mouse α1,3-galactosyltransferase (mouse GaIT), and canine Forsmann glycolipid synthetase (dog Forsman). The mouse AB transferase was 47 and 49% identical to the mouse α1,3-galactosyltransferase and the dog Forsman synthetase gene in the last two coding exons, respectively. B, the coding sequences around the amino acid residues important for the nucleotide-sugar recognition/binding are compared. The sequences from the human O03 allele that contains a missense mutation around this area (9) and the cis-AB01 allele with dual transferase activity (8) are also included. C, a phylogenetic tree of the ABO gene family is shown. The following sequences in GenBank were used for the computation (46) and neighbor-joining tree (39): marmoset GaIT (S71333), bovine GaIT (J04989), pig GaIT (L36152), mouse GaIT (M85153), and dog Forsman synthetase (U66140). Numbers on interior branches denote bootstrap probabilities in percent. Branch lengths are proportional to the numbers of amino acid substitutions, and a scale is given at the bottom of the tree. The dotted line denotes the probable root node of this tree. This node also corresponds to a gene duplication event producing both the α1,3-galactosyltransferase gene and ABO-related gene. The latter lineage experienced another gene duplication, resulting in creation of both Forsman glycolipid synthetase gene and ABO gene. The murine AB gene was mapped closest to the human ABO genes. It also mapped closer to the canine Forsman gene than the murine α1,3-galactosyltransferase gene.

**Murine ABO Gene Equivalent**

![Diagram](http://www.jbc.org)

**Fig. 2. Comparisons of the amino acid sequences of the coding regions among murine AB gene, human A1-1 (A101) and B (B101) genes, murine α1,3-galactosyltransferase gene, and canine Forsmann synthetase gene and an evolutionary tree of the ABO and related genes.** A, the amino acid sequence of the mouse AB transferase was compared with those of the human A and B transferases, mouse α1,3-galactosyltransferase (mouse GaIT), and canine Forsmann glycolipid synthetase (dog Forsman). The mouse AB transferase was 47 and 49% identical to the mouse α1,3-galactosyltransferase and the dog Forsman synthetase gene in the last two coding exons, respectively. B, the coding sequences around the amino acid residues important for the nucleotide-sugar recognition/binding are compared. The sequences from the human O03 allele that contains a missense mutation around this area (9) and the cis-AB01 allele with dual transferase activity (8) are also included. C, a phylogenetic tree of the ABO gene family is shown. The following sequences in GenBank were used for the computation (46) and neighbor-joining tree (39): marmoset GaIT (S71333), bovine GaIT (J04989), pig GaIT (L36152), mouse GaIT (M85153), and dog Forsman synthetase (U66140). Numbers on interior branches denote bootstrap probabilities in percent. Branch lengths are proportional to the numbers of amino acid substitutions, and a scale is given at the bottom of the tree. The dotted line denotes the probable root node of this tree. This node also corresponds to a gene duplication event producing both the α1,3-galactosyltransferase gene and ABO-related gene. The latter lineage experienced another gene duplication, resulting in creation of both Forsman glycolipid synthetase gene and ABO gene. The murine AB gene was mapped closest to the human ABO genes. It also mapped closer to the canine Forsman gene than the murine α1,3-galactosyltransferase gene.

**Murine ABO Gene Equivalent**

![Diagram](http://www.jbc.org)

**Fig. 2. Comparisons of the amino acid sequences of the coding regions among murine AB gene, human A1-1 (A101) and B (B101) genes, murine α1,3-galactosyltransferase gene, and canine Forsmann synthetase gene and an evolutionary tree of the ABO and related genes.** A, the amino acid sequence of the mouse AB transferase was compared with those of the human A and B transferases, mouse α1,3-galactosyltransferase (mouse GaIT), and canine Forsmann glycolipid synthetase (dog Forsman). The mouse AB transferase was 47 and 49% identical to the mouse α1,3-galactosyltransferase and the dog Forsman synthetase gene in the last two coding exons, respectively. B, the coding sequences around the amino acid residues important for the nucleotide-sugar recognition/binding are compared. The sequences from the human O03 allele that contains a missense mutation around this area (9) and the cis-AB01 allele with dual transferase activity (8) are also included. C, a phylogenetic tree of the ABO gene family is shown. The following sequences in GenBank were used for the computation (46) and neighbor-joining tree (39): marmoset GaIT (S71333), bovine GaIT (J04989), pig GaIT (L36152), mouse GaIT (M85153), and dog Forsman synthetase (U66140). Numbers on interior branches denote bootstrap probabilities in percent. Branch lengths are proportional to the numbers of amino acid substitutions, and a scale is given at the bottom of the tree. The dotted line denotes the probable root node of this tree. This node also corresponds to a gene duplication event producing both the α1,3-galactosyltransferase gene and ABO-related gene. The latter lineage experienced another gene duplication, resulting in creation of both Forsman glycolipid synthetase gene and ABO gene. The murine AB gene was mapped closest to the human ABO genes. It also mapped closer to the canine Forsman gene than the murine α1,3-galactosyltransferase gene.
the sequenced and unsequenced regions of the gene in at least those five strains of mice examined.

**A and B Antigens Are Expressed in Mice**—We have shown that the mouse equivalent of the human ABO gene encodes a protein capable of transferring both GalNAc and galactose through in vitro enzymatic assays using extracts from HeLa cells transfected with the eukaryotic expression constructs of the murine gene. We have also shown the presence of A and B transferase activity in the murine submaxillary glands by use of in vitro assays of tissue extracts. However, the expression of the protein does not necessarily prove the expression of A and B antigens, since the enzymatic reactions require the appropriate substrates and reaction conditions. Therefore, we next analyzed the expression of A and B antigens. Two laboratory strains of *M. musculus domesticus* mice were analyzed. We initially examined the agglutination of murine red blood cells using murine monoclonal antibody mixtures against A and B antigens under the regular agglutination conditions used for the ABO blood group typing of human red blood cells. No agglutination was observed (data not shown). Because we observed A and B transferase activity in the submaxillary gland extract, we next performed the inhibition study, using boiled extracts of the murine submaxillary glands. No inhibition of the reference human red blood cell agglutination was observed, although the treatment of murine monoclonal antibody mixtures with the control group A porcine submaxillary gland extract resulted in some inhibition (data not shown). These results suggested that A and B antigens are not expressed in abundance if at all. We therefore examined the expression of these antigens using the more sensitive immunological method of nitrocellulose spotting. As shown in Fig. 3, both A and B antigens were detected in the murine submaxillary glands. Apparently, higher expression was observed with A antigens than with B antigens. This may be attributed to the decreased availability of UDP-galactose substrate, resulting from the competition between B transferase and $\alpha_{1,3}$-galactosyltransferase for the same donor nucleotide-sugar. The low expression of A and B antigens may be similarly explained by the competition for the same acceptor substrate between $\alpha_{1,3}$-galactosyltransferase and $\alpha_{1,2}$-fucosyltransferase. Since $\alpha_{1,3}$-Gal epitope produced by the $\alpha_{1,3}$-galactosyltransferase is abundantly expressed in murine tissues, it is likely that the $\alpha_{1,3}$-galactosyltransferase has higher affinity for the acceptor substrates than the $\alpha_{1,2}$-fucosyltransferase that synthesizes H antigens.

**DISCUSSION**

We have cloned murine genomic DNA containing most of the coding sequence that was equivalent to the human ABO gene.

### Table I

**In vitro enzymatic assays of extracts from HeLa cells transfected with eukaryotic expression constructs**

DNA transfection was performed using LipofectAMINE. Human A and B transferase expression constructs, pAAAA and pBBBB (30), were used as positive controls; nonfunctional pA(arginine) construct (31) was used as a negative control. A and B transferase activity was determined by measuring the transfer of $^{14}\text{C}$ from radiolabeled nucleotide-sugars to the acceptor substrate 2'-fucosyllactose. The values represent the counts in cpm after subtracting the values obtained without 2'-fucosyllactose from the values obtained with 2'-fucosyllactose in the reactions. The activity of the respective constructs was summarized and is shown in the rightmost column.

<table>
<thead>
<tr>
<th>Construct</th>
<th>$^{14}\text{C}$GalNAc incorporation</th>
<th>$^{14}\text{C}$Galactose incorporation</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>No DNA</td>
<td>9</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>pA(arginine)</td>
<td>30</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>pAAAA</td>
<td>2700</td>
<td>75</td>
<td>A</td>
</tr>
<tr>
<td>pBBBB</td>
<td>10</td>
<td>450</td>
<td>B</td>
</tr>
<tr>
<td>pHuman-mouse chimera</td>
<td>660</td>
<td>610</td>
<td>AB</td>
</tr>
<tr>
<td>pMouse</td>
<td>630</td>
<td>600</td>
<td>AB</td>
</tr>
</tbody>
</table>

### Table II

**Nucleotide and amino acid substitutions in mouse species and subspecies**

The nucleotide and the deduced amino acid sequences of the murine equivalents of the human ABO gene were compared among several strains of mouse species and subspecies. The nucleotide and the deduced amino acid sequences in the region between nucleotide 360 and 807 in Fig. 1 were examined. The nucleotide sequence substitutions are shown in boldface type, and the resulting amino acid substitutions are shown in parentheses. A and B transferase activity was examined using extracts from submaxillary glands. + and ND denote the presence of activity and not determined, respectively. In these enzymatic assays, A transferase activity of 0.004–0.093 nmol/h/mg of protein (1.287 for the positive control human gastric carcinoma cell line KATOIII cell extract) and B transferase activity of 0.008–0.036 nmol/h/mg (0.035 for the positive control human gastric carcinoma cell line KATOIII cell extract) were detected among different strains of mice.

<table>
<thead>
<tr>
<th>Name</th>
<th>Nucleotide position</th>
<th>Transferase activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>418</td>
<td>542</td>
</tr>
<tr>
<td><em>M. musculus domesticus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BALB/c strain</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>B10 strain</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>p52 strain</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td><em>M. musculus molossinus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSF/Msf strain</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td><em>M. musculus musculus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLG/Msf strain</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td><em>M. musculus subspecies</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWN/Msf strain</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td><em>M. musculus breviostris</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJL/Msf strain</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>BFM/2Msf strain</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>(R181H)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. musculus castaneus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAST/El strain</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>(T206A)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMI/Msf strain</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>(T206A)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. spicilegus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZBN strain</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>(Q140K)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Murine ABO Gene Equivalent

The sequence encoding the first few amino acid residues and the promoter region sequence farther upstream remain to be cloned. In humans, ABH antigens are widely expressed on a variety of cell surface molecules in many tissues, depending on the ABO genotypes of the individuals. These include glycoproteins and glycolipids on mucous cells, nerve cells, red cells, epidermis, and vascular endothelium. ABH antigen expression seems to be more highly restricted in lower mammals (42, 43). To understand the differential expression mechanism of A/B transferase activity, we undertook of Genetics, Mishima, Japan) for supplying various strains of mice.

FIG. 3. Immunological detection of A and B antigens in murine submaxillary glands. The expression of A and B antigens was analyzed by the immunostaining of the extracts spotted onto a nitrocellulose membrane. The extracts were derived from the following sources: 1) human colon adenocarcinoma SW48 cells (AB phenotype); 2) group A porcine submaxillary gland; 3) group O porcine submaxillary gland; 4) submaxillary gland from a C57BL strain of mouse; and 5) submaxillary gland from an ICR strain of mouse. **BSA**, bovine serum albumin.

The expression of A and B antigens was analyzed by the immunostaining of the extracts spotted onto a nitrocellulose membrane. The extracts were derived from the following sources: 1) human colon adenocarcinoma SW48 cells (AB phenotype); 2) group A porcine submaxillary gland; 3) group O porcine submaxillary gland; 4) submaxillary gland from a C57BL strain of mouse; and 5) submaxillary gland from an ICR strain of mouse. **BSA**, bovine serum albumin.

**REFERENCES**


Acknowledgments—We thank Emily N. Wang for editing the manuscript and Dr. Sandra S. Matsumoto for critically reviewing the manuscript. We are also indebted to Dr. Yoshikichi Shiraiishi (National Institute of Genetics, Mishima, Japan) for supplying various strains of mice.
Murine ABO Gene Equivalent