Munere 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 Saitout‡, and Fumihiro 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 oligosaccharide 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- and 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- and B transferase for B substrate. O 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 adeno-
carcinoma 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, A4, and B2) (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 blots, 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 priate 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 1744 solely to indicate this fact.

The nucleotide sequence reported in this paper has been submitted to the DDBJ/GenBankTM/EMBL Data Bank with accession numbers AB041038 and AB041039.

† To whom correspondence should be addressed: The Burnham Institute, La Jolla Cancer Research Center, 10901 N. Torrey Pines Rd., La Jolla, CA 92037. Tel.: 858-646-3116; Fax: 858-646-3173; E-mail: fyamamoto@burnham.org.

This paper is available on line at http://www.jbc.org

13701
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, Leb (Fucα1–2 Gal β1–3 [Fuc α1–4] GlcNac), 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 Leb structure into ALε and BLε 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 having 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 MsoI-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 cloning kit were also from CLONTECH. A GeneClean kit was purchased from Bio101 (La Jolla, CA), and pRTT3x18 plasmid vector, α-S-dGTP, α-S-dCTP, and S-300 Microspin columns were from Amer sham Pharmacia Biotech. LifeprobeAMINE was purchased from Life Technologies, Inc. [14C]UDP-GalNAc, [14C]UDP-galactose, and [α-32P]dCTP were from PerkinElmer Life Sciences, and 2′-fucosyllactose 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, Vectorstain Elite ABC kit, and 4-chloro-1-napthol chromogen kit were from Vector Laboratories, Burlington, CA, and EZ-Link Sulfo-NHS-LC-Biotin was purchased from Pierce. Bluescript SKM13 vector, Prime-It II kit, Duralose-UV membranes, and frozen competent XL1-blue strain of *Escherichia coli* were procured from Stratagene (La Jolla, CA). The ULTRAhyb hybridization buffer was obtained from Ambion (Austin, TX). Restriction endonucleases and nucleic acid-modifying enzymes were from Life Technologies, New England Biolabs (Beverly, MA), or Roche Molecular Biochemicals. dRhodamine dye terminator cycle sequencing ready reaction kit (28). 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′-TTAGTTTCTGATGCTAGGCTCCTTGGCAC and 5′-TCATGCACACAGCCCTCACTGGCGT 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.

**Methods**

**Isolation of Genomic DNA Clones and Construction of Restriction Endonuclease Cleavage Maps**—Under low stringency conditions, 1 million independent plasmid clones from the pBS/T7 phage DNA library were screened 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-32P]dCTP (26). After four rounds of screening, individual clones were isolated. Phage DNA was prepared, cleaved with restriction endonucleases, gel-electrophoresed, and Southern transferred. Hybridization was then performed to construct restriction enzyme cleavage maps.

**Subcloning, Nested Deletion, and DNA Sequencing**—DNA from MAB0416 phage clone was cleaved with HindIII and SalI and subcloned into the Bluescript SKM13 vector. Nested deletion constructs were obtained by the ExoFill PCR and exonuclease method (27). When no unique 3′ overhang restriction sites were available, the thioderivative fill-in reactions were performed with Klenow enzyme using α-S-dGTP and α-S-dCTP before EcoIII treatment. After transformation of *E. coli* XLI-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 MY-1 adapter primer and then with AP2 and MY-2 primers. DNA fragments were then ligated with pT-Adv vector from the Advanced 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, MAB0416, into the pBS/T7 phage 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 SstI and SnaBI. The SstI site was within the pRTT3x18 plasmid vector, and the SnaBI site was located downstream of the stop codon of the mouse ABO transferase. The resulting sequence from the purification of the phage DNA fragments was then ligated into the SstI–SnaBI fragment containing the coding sequence in the last coding exon of the mouse ABO gene was then isolated. The human B transferase expression construct with intron, pBBBB (30), was cleaved with BamHI, blunt-ended by the Klenow filling-in reaction, and then digested with SstI. The SstI site was in the intron preceding coding exon 7 of the human ABO gene; BamHI was in the eukaryotic expression vector (originally pBG-5). The SstI–BamHI vector fragment containing the human B transferase cDNA sequence of exons 1–6 was then ligated to the mouse SstI–SnaBI fragment to produce the human-mouse chimeric gene (pHuman-mouse chimera). A murine cDNA eukaryotic expression construct was then constructed by replacing the EcoRI–A/III fragment from the chimeric construct with the EcoRI–A/II fragment from a 5′-RACE clone in the XhoI–Ad vector. The EcoRI site of the clone was in the plasmid and located 5′ upstream of the cDNA end. The A/III was in the last coding exon. The EcoRI site in the pSg-5 vector was located downstream of the SV40 early promoter and upstream of the human cDNA sequence. The resultant construct (pMouse) 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 human antigens on their cell surfaces and have been successfully used in similar transfection experiments of A and B transferase expression constructs (5, 30, 31). Following the manufacturer’s protocol, we used LipofectAMINE for transfection. Seventy-two hours after transfection, the cells were washed and harvested. Cells were then lysed in buffer (0.1 M NaCl, 25 mM sodium cacodylate, 10 mM Mops, and 0.1% Triton X-100). A/B transferase activity was determined by measuring the transfer of carbon-14 from [14C]UDP-GalNAc or [14C]UDP-galactose to the acceptor substrate 2′-fucosyllactose, as described previously (31). After incubation, the reaction products were separated from unincorporated nucleotide-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/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'-GAGACTTCGACAGAAACACTT; SN-17, 5'-CAATGGCTTGGCCGTTGTC. 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-Cl (pH 7.5) and 0.1% SDS. The extract was then spotted onto a Durable-UVE 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 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 A1→3Gal epitopes. The cDNA encoding α1,3-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 anti-A epitope. This epitope was cloned from cow (32), mouse (33), and pig (34, 35). Therefore, these genes are believed to have derived from the same ancestral gene and constitute the ABO gene family. Fig. 2 highlights the amino acid 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 α1,3-galactosyltransferase, and canine Forsmann glycolipid synthetase. Results are shown in Fig. 2A. The percentages of identical amino acid residues of the coding sequences in the last two coding exons are 47% (127 of 272) between the mouse ABO and α1,3-galactosyltransferase genes and 49% (132 of 272) between the mouse ABO gene and the dog Forsmann synthetase gene. Fig. 2B highlights the amino acid sequence...
quences 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 \(\alpha_1,3\)-galactosyltransferase gene.

**Murine Genes Encode 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

**Figure 1.** Comparison of the exon-intron boundaries and the nucleotide and deduced amino acid sequences of the coding region of the murine \(\alpha\)-A1-1 (A101) allele. The nucleotide sequence of the mouse gene in the MAB0\(\alpha\)16 phage clone and the sequence of the 5'-RACE-amplified cDNA were used to determine the exon-intron boundaries. Almost all of the coding sequence was contained in the cloned genomic DNA fragment. Since the sequence encoding the first few amino acid residues is still missing, the earlier exon-intron boundaries have not yet been completely determined. In Fig. 1, A and B, two possible splicing patterns are presented. There are six coding exons in A and seven in B. In A, ~70 base pairs upstream of the splicing acceptor site in coding exon 2 (cEXON 2), there was a GAATTCAG sequence. If the acceptor site is shifted upstream by 2 nucleotides, that sequence may represent the preceding exon sequence as shown in B. This shift, however, breaks up the GT-AG rule of splice junctions. In addition, no 5'-untranslated region sequence is found upstream. Therefore, A may represent what most likely occurs. XXXXXX, the nucleotide sequence is unknown. Unlike the human sequence, the sequence in human exons 3 and 4 was contained in a single exon in the mouse gene. The numbers shown in parentheses next to the nucleotide numbers of the coding sequence indicate the nucleotide residue numbers of the MAB0\(\alpha\)16 phage clone insert deposited to DDBJ (accession number AB041039). The nucleotide sequence of the murine ABO gene cDNA has also been deposited (accession number AB041038). Also shown is the comparison of the nucleotide and deduced amino acid sequences in the coding region of the murine gene with the human ABO (A101) gene sequences (accession number AF134412 in GenBank). A combined method incorporating the Clustal method (40) and the J. Hein method (41) was used with the MegAlign program. High homology was observed in the coding sequence in the last two coding exons. The percentages of the identical nucleotide and amino acid residues were 78 and 81% between the two sequences, respectively.
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 GalT), and canine Forsmann glycolipid synthetase (dog Forsman). The mouse AB transferase was 47 and 49% identical to the mouse α1,3-galactosyltransferase gene and the dog Forsman synthetase gene in the 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 of evolutionary distance (46) and neighbor-joining tree (39): marmoset GalT (S71333), bovine GalT (J04989), pig GalT (L36152), mouse GalT (M85153), and dog Forsman synthetase (U68140). 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 further 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 mouse α1,3-galactosyltransferase gene.

**Murine ABO Gene Equivalents**

### A

| mouse AB | human A | human B | human AB | dog Forsman |
|----------|---------|---------|----------|-------------|
| N---------E----- | GPKPCNPLML | -------- | GPKPCNPLML | MTLPLMLVFLFGYCLV | 33 |
| E---------L----- | MAELVRLTA | -------- | GPKCHALRP | ------------ | 38 |
| L---------G----- | MIIMLMLVFLFGYCLV | -------- | GPKCHALRP | ------------ | 38 |
| F---------E----- | GPKCHALRP | -------- | GPKCHALRP | ------------ | 38 |
| Y---------C----- | GPKCHALRP | -------- | GPKCHALRP | ------------ | 38 |

### B

| mouse AB | human A | human B | human AB | dog Forsman |
|----------|---------|---------|----------|-------------|
| FRSQNL--- | PTKLPRC | -------- | PTKLPRC | ------------ | 60 |
| Q---------L----- | PTKLPRC | -------- | PTKLPRC | ------------ | 60 |
| S---------E----- | PTKLPRC | -------- | PTKLPRC | ------------ | 60 |
| K---------Q----- | PTKLPRC | -------- | PTKLPRC | ------------ | 60 |
| E---------G----- | PTKLPRC | -------- | PTKLPRC | ------------ | 60 |

### C

| marmoset GalT | bovine GalT | pig GalT | dog Forsman | mouse AB | human A |
|---------------|-------------|---------|-------------|---------|---------|
| 100            | 66          | 99      | 70          | 100     | 0.100   |

**AB Gene Is Prevailing in Mice**—We have determined the partial nucleotide sequences of the coding region in the last coding exon of the murine ABO gene using genomic DNA from several species and subspecies of mice. The results are summarized in Table II. Mouse strains analyzed were *Mus musculus domesticus* (B10 and png2 strains), *Mus musculus molossinus* (MSF/Msf strain), *Mus musculus musculus* (BLG/Msf strain), *Mus musculus breviostris* (BFM/2Msf and NJL/Msf strains), *Mus musculus castaneus* (CAST/Ei and HMI/Msf strains), and *Mus spicilegus* (ZBN strain). Except for *Mus spicilegus*, all others were subspecies of *Mus musculus*. The results identified several nucleotide substitutions, some of which resulted in amino acid substitutions. Therefore, there are no mutations that change the donor nucleotide-sugar specificity in

---

FIG. 2. Comparisons of the amino acid sequences of the coding regions among murine AB gene, human A-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 GalT), and canine Forsmann glycolipid synthetase (dog Forsman). The mouse AB transferase was 47 and 49% identical to the mouse α1,3-galactosyltransferase gene 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 of evolutionary distance (46) and neighbor-joining tree (39): marmoset GalT (S71333), bovine GalT (J04989), pig GalT (L36152), mouse GalT (M85153), and dog Forsman synthetase (U68140). 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 further 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 mouse α1,3-galactosyltransferase gene.

**Murine ABO Gene Equivalents**

### A

| mouse AB | human A | human B | human AB | dog Forsman |
|----------|---------|---------|----------|-------------|
| MN--------ER----- | GPKPCNPLML | -------- | GPKPCNPLML | MTLPLMLVFLFGYCLV | 33 |
| human A | MAELVRLTA | -------- | GPKCHALRP | ------------ | 38 |
| human B | MAELVRLTA | -------- | GPKCHALRP | ------------ | 38 |

### B

| mouse AB | human A | human B | human AB | dog Forsman |
|----------|---------|---------|----------|-------------|
| FRSQNL--- | PTKLPRC | -------- | PTKLPRC | ------------ | 60 |
| human A | PTKLPRC | -------- | PTKLPRC | ------------ | 60 |
| human B | PTKLPRC | -------- | PTKLPRC | ------------ | 60 |

### C

| marmoset GalT | bovine GalT | pig GalT | dog Forsman | mouse AB | human A |
|---------------|-------------|---------|-------------|---------|---------|
| 100            | 66          | 99      | 70          | 100     | 0.100   |
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 mouse 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 α1,3-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 α1,3-galactosyltransferase and α1,2-fucosyltransferase. Since α1→3-Gal epitope produced by the α1,3-galactosyltransferase is abundantly expressed in murine tissues, it is likely that the α1,3-galactosyltransferase has higher affinity for the acceptor substrates than the α1,2-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, pAA and pBBB (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 [14C] from radiolabeled nucleotide-sugars to the acceptor substrate 2′-fucosyllactose. The values represent the counts in cpm after subtracting the values obtained with 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.

| Construct                        | [14C]GalNAc incorporation | [14C]galactose incorporation | Activity |
|----------------------------------|---------------------------|------------------------------|----------|
| No DNA                           | −9                        | +40                          |          |
| pA(arginine)                     | −30                       | −110                         |          |
| pAA                              | +2700                     | +75                          | A        |
| pBBB                             | +10                       | +450                         | B        |
| pHuman-mouse chimera             | +660                      | +610                         | AB       |
| pMouse                           | +630                      | +600                         | AB       |

**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 MKN 45 cell extract) and B transferase activity of 0.008–0.030 nmol/h/mg (0.035 for the positive control human gastric carcinoma cell line KATOIII cell extract) were detected among different strains of mice.

| Name                     | Nucleotide position | Transferase activity |
|--------------------------|---------------------|----------------------|
|                          | 418     | 542    | 543    | 615    | 616    | 735    | 802    | A | B |
| *M. musculus domesticus* |          |        |        |        |        |        |        |   |   |
| BALB/c strain            | C       | G      | T      | A      | A      | A      | A      | + | + |
| B10 strain               | C       | G      | T      | A      | A      | A      | A      | ND| ND|
| png2 strain              | C       | G      | T      | A      | A      | A      | A      | ND| ND|
| *M. musculus molossinus* |          |        |        |        |        |        |        |   |   |
| MSF/Msf strain           | C       | G      | T      | A      | A      | A      | A      | ND| ND|
| *M. musculus musculus*   |          |        |        |        |        |        |        |   |   |
| BLG/Msf strain           | C       | G      | T      | A      | A      | A      | A      | ND| ND|
| *M. musculus subspecies* |          |        |        |        |        |        |        |   |   |
| SWN/Msf strain           | C       | G      | T      | A      | A      | A      | A      | ND| ND|
| *M. musculus breviostris*|          |        |        |        |        |        |        |   |   |
| NJL/Msf strain           | C       | G      | T      | A      | A      | A      | A      | ND| ND|
| BFM/2Msf strain          | C       | A      | T      | A      | A      | A      | A      | + | + |
| (R181H)                  |          |        |        |        |        |        |        |   |   |
| *M. musculus castaneus*  |          |        |        |        |        |        |        |   |   |
| CAST/EI strain           | C       | G      | C      | G      | G      | G      | G      | + | + |
| (T206A)                  |          |        |        |        |        |        |        |   |   |
| HMI/Msf strain           | C       | G      | C      | G      | G      | G      | G      | + | + |
| (T206A)                  |          |        |        |        |        |        |        |   |   |
| *M. spicilegus*          |          |        |        |        |        |        |        |   |   |
| ZBN strain               | A       | G      | T      | A      | G      | G      | G      | + | + |
| (Q140K)                  |          |        |        |        |        |        |        |   |   |
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 antigens between humans and mice, cloning of the murine gene promoter region will be necessary.

Functional analysis of the murine gene has shown that the human ABO gene is crucially essential for the production of ABO antigens. In our prior studies of the murine ABO gene, we determined the molecular causes of two interesting phenomena named cis-AB (8) and Bi(A) (7). Human cis-AB alleles encode a protein with strong A and weak B transferase activity, whereas B(A) alleles encode a protein with strong B and weak A transferase activity. A mutation was found in these alleles at each of the four amino acid substitutions that distinguish human A transferase (AAAA) from B transferase (BBBB). The cis-AB alleles were represented as AAAB, whereas the B(A) alleles were represented as BABB in those four substitution positions. By employing DNA transfection assays of A/B transferase chimeras and in vitro mutagenized expression constructs, we showed that these amino acid substitutions might cause changes in specificity and activity of the enzyme (30, 31). As shown in Fig. 28, the amino acid residues at codons 245 and 247 of the murine gene, which corresponded to human codons 266 and 268 (the third and fourth positions of amino acid substitutions between A and B transfersases), were glycosylated and alanylated. The alanyl residue is identical to that of human B transferase, but the glycine residue is much smaller than the methionine residue of human B transferase. Therefore, the mouse enzyme is expected to possess a larger space for donor nucleotide-sugar substrate recognition/binding. Since N-acetyl-β-galactosamine is more bulky than galactose, the increase in space may allow for the accommodation of not only the galactose portion of UDP-galactose but also the GaINAc portion of UDP-GaINAc, which may account for the dual transferase activity of the murine enzyme. Dog Forssman glycolipid synthetase contains glycosyl and alanylated, the same amino acid residues as those of the murine ABO gene equivalent, at codons 261 and 263. However, the codon 264 in the Forssman glycolipid synthetase is valine rather than phenylalanine in the murine AB gene. This difference in size (valine is smaller than phenylalanine) may possibly render dog Forssman glycolipid synthetase able to catalyze the transfer of only GaINAc residues and not galactose residues.

To understand the meaning and role of ABO polymorphism during evolution, the cloned murine AB gene fragments will be useful in the production of knockout (group O) mice at the ABO locus. Production of group A and B mice may also be possible by knocking in the genes after modifying the specificity of the enzymes so that only A or B transferase activity is retained. Additional manipulation(s) may be a prerequisite to enhance the expression of A and B antigens. These may include the knockout of the α1,3-galactosyltransferase gene (44) or the introduction of the α1,2-fucosyltransferase gene under strong promoter (45) to either abolish or lessen the expression of the α1,3-galactosyltransferase. These mice with different ABO phenotypes in the same genetic background may clarify the functionality of the ABO system in the future.

**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. Yoshitaka Shirushi (National Institute of Genetics, Mishima, Japan) for supplying various strains of mice.

**REFERENCES**

1. Yamamoto, F. (2000) *Vox Sang.* 78, 91–103
2. Yamamoto, F., Marken, J., Tsuchi, T., White, T., Clausen, H., and Hakomori, S. (1990) *J. Biol. Chem.* 265, 1146–1151
3. Clausen, H., White, T., Takio, K., Titani, K., Stroud, M., Holmes, E., Karkov, J., Thim, L., and Hakomori, S. (1990) *J. Biol. Chem.* 265, 1139–1145
4. Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) *Nature* 354, 229–233
5. Yamamoto, F., McNeill, P. D., and Hakomori, S. (1992) *Biochem. Biophys. Res. Commun.* 187, 360–374
6. Yamamoto, F., McNeill, P. D., Yamamoto, M., Hakomori, S., Harris, T., Judd, D. W., and Davenport, R. D. (1993) *Vox Sang.* 64, 116–119
7. Yamamoto, F., McNeill, P. D., Yamamoto, M., Hakomori, S., and Harris, T. (1993) *Vox Sang.* 64, 171–174
8. Yamamoto, F., McNeill, P. D., Kominato, Y., Yamamoto, M., Hakomori, S., Ishimoto, S., Nishida, S., Shima, M., and Fujimura, Y. (1993) *Vox Sang.* 64, 130–133
9. Yamamoto, F., McNeill, P. D., Yamamoto, M., Hakomori, S., Bromilow, I. M., and Duguid, J. K. (1993) *Vox Sang.* 64, 175–178
10. Ogawasawa, K., Yabe, R., Uchikawa, M., Saitou, N., Bannai, N., Nakata, K., Takenaka, M., Fujisawa, K., Ishikawa, Y., Juji, T., and Tokunaga, K. (1996) *Blood* 88, 2372–2377
11. Ogawasawa, K., Bannai, M., Saitou, N., Yabe, R., Nakata, K., Takenaka, M., Fujisawa, K., Uchikawa, M., Ishikawa, Y., Juji, T., and Tokunaga, K. (1996) *Hum. Genet.* 97, 777–783
12. Olsson, M. L., and Chester, M. A. (1998) *Transfus. Med.* 8, 231–238
13. Olsson, M. L., Thuresson, B., and Chester, M. A. (1995) *Biochem. Biophys. Res. Commun.* 216, 642–647
14. Olsson, M. L., and Chester, M. A. (1996) *Vox Sang.* 71, 113–117
15. Olsson, M. L., and Chester, M. A. (1996) *Biochem. Biophys. Res. Commun.* 226, 599–613
16. Rabah, E. (1969) *Blood* and its Chemistry and Immunology, Academic Press, Inc., New York
17. Kominato, Y., McNeill, P. D., Yamamoto, M., Russell, M., Hakomori, S., and Yamamoto, F. (1992) *Biochem. Biophys. Res. Commun.* 193, 154–164
18. Martinko, J. M., Vincent, V., Klein, D., and Klein, J. (1993) *Immunogenetics* 37, 274–278
19. Saitou, N., and Yamamoto, F. (1997) *Med. Biol. Educ.* 19, 399–411
20. O’Huigin, C., Sato, A., and Klein, J. (1997) *Hum. Genet.* 101, 141–148
21. Pittiglio, D. H. (1986) in *Genetics and Biochemistry of A, B, H, and Lewis Antigens* (Wallace, M. F., and Gibbs, F. L., eds), pp. 1–56, American Association of Blood Banks, Arlington, VA
22. Boren, T., Falk, P., Roth, K. A., Larson, G., and Normark, S. (1993) *Science* 262, 1892–1895
23. Takeuchi, Y., Porter, C. D., Strahan, K. M., Freeno, A. F., Gustafsson, K., Cosset, F. L., Weiss, R. A., and Collins, M. K. (1996) *Nature* 379, 85–88
24. Arendrup, M., Hansen, J. E., Clausen, H., Nielsen, C., Mathiesen, L. R., and Nielsen, J. O. (1991) *AIDS* 5, 441–444
25. Manipol, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
26. Feinberg, A. P., and Vogelstein, B. (1983) *Anal. Biochem.* 132, 6–13
27. Chang, D. W., Tartof, K. D., and Yeung, A. T. (1993) *Genes Immun.* 127, 95–98
28. Sanger, F., Nicklen, S., and Coulson, A. R. (1992) *Bio/Technology* 10, 104–108
29. Barnes, W. M. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91, 2216–2220
30. Yamamoto, F., and Hakomori, S. (1990) *J. Biol. Chem.* 265, 19257–19262
31. Yamamoto, F., and McNeill, P. D. (1996) *J. Biol. Chem.* 271, 10515–10520
32. Joziasse, D. H., Shaper, J. H., Van den Eijnden, D. H., Van Tunen, A. J., and Shaper, N. L. (1989) *J. Biol. Chem.* 264, 14290–14297
33. Larsen, R. D., Rajan, V. P., Ruff, M. M., Kukowska-Latallo, J., Cummings, R. D., and Lowe, J. B. (1989) *Proc. Natl. Acad. Sci. U. S. A.* 86, 8227–8231
34. Strahan, K. M., Wu, P., Anderson, L., and Gustafsson, K. (1995) *Transplant.*
Murine ABO Gene Equivalent

35. Strahan, K. M., Gu, F., Preece, A. F., Gustavsson, I., Andersson, L., and Gustafsson, K. (1995) Immunogenetics 41, 101–105
36. Galili, U., Macher, B. A., Buehler, J., and Shohet, S. B. (1985) J. Exp. Med. 162, 573–582
37. Larsen, R. D., Rivera-Marrero, C. A., Ernst, L. K., Cummings, R. D., and Lowe, J. B. (1990) J. Biol. Chem. 265, 7055–7061
38. Haslam, D. B., and Baenziger, J. U. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 10697–10702
39. Saitou, N., and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425
40. Higgins, D. G., and Sharp, P. M. (1988) Gene (Amst.) 73, 237–244
41. Hein, J. (1990) Methods Enzymol. 183, 626–645
42. Oriol, R., Le Pendu, J., and Mollicone, R. (1986) Vox Sang. 51, 161–171
43. Oriol, R., Mollicone, R., Coullin, P., Dalix, A. M., and Candelier, J. J. (1992) APMIS Suppl. 27, 28–38
44. Tearle, R. G., Tange, M. J., Zanettino, Z. L., Katerelos, M., Shinkel, T. A., Van Denderen, B. J., Lonie, A. J., Lyons, I., Nettle, M. B., Cox, T., Becker, C., Pours, A. M., Wigley, P. L., Crawford, R. J., Robins, A. J., Pearse, M. J., and d’Apice, A. J. (1996) Transplantation 61, 13–19
45. Costa, C., Zhao, L., Burton, W. V., Bondioli, K. R., Williams, B. L., Hoagland, T. A., Ditullio, P. A., Ebert, K. M., and Fodor, W. L. (1999) FASEB J. 13, 1762–1773
46. Kimura, M. (1983) The Neutral Theory of Molecular Evolution, Cambridge Press, Cambridge, United Kingdom