Expression of the histo-blood group *B* gene predominates in *AB*-genotype cells

Yuh-Ching Twu, Chuang-Yi Hsieh, and Lung-Chih Yu

BACKGROUND: It has been demonstrated that the 43-bp minisatellite sequence in the 5′ region of the *ABO* gene plays an important role in its transcriptional regulation. It was determined in previous investigations that the structure of the minisatellite enhancer was specific to *A*, *B*, and *O* alleles.

STUDY DESIGN AND METHODS: Real-time polymerase chain reaction (PCR) detection and a PCR-restriction fragment length polymorphism (RFLP) strategy were used to compare the quantities of the *A* and *B* transcripts in *AB*-genotype cells, including peripheral blood cells and cancer cell line with the group AB phenotype. The 5′ 3.7-kb regions of the *A* and *B* genes were cloned and the sequences compared. The transcriptional activities of the 5′ segments of the *A* and *B* genes were compared with luciferase reporter assay.

RESULTS: Both real-time PCR and PCR-RFLP analyses show that there is evidently more of the *B* transcript in the *AB*-genotype cells. It was demonstrated that the 5′ segment of the *B* gene had a markedly higher transcription-activation activity relative to the *A* gene. This difference in transcription capability appears to result from the variation in minisatellite-enhancer structures in the *A* and *B* genes, which contain one and four repeats of the 43-bp enhancer unit, respectively.

CONCLUSION: Our study indicates that the majority of steady-state mRNA within *AB*-genotype cells is composed of the *B* transcript and that this phenomenon is due to the predominant expression of the *B* gene relative to the *A* gene.

From the Institute of Biochemical Sciences, National Taiwan University, Taipei; and the Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan.

Address reprint requests to: Lung-Chih Yu, PhD, the Institute of Biochemical Sciences, National Taiwan University, PO Box 23-106, Taipei 106, Taiwan; e-mail: yulc@ntu.edu.tw.

Supported in part by Grant NSC 94-2320-B-002-088 from the National Science Council, Taiwan.

Received for publication January 8, 2006; revision received April 9, 2006, and accepted April 12, 2006.

doi: 10.1111/j.1537-2995.2006.01007.x

TRANSFUSION 2006;46:1988-1996.

The human blood group A and B antigens, which have been characterized as the carbohydrate determinants, GalNAcα1-3(Fucα1-2)Galβ1-R and Galα1-3(Fucα1-2)Galβ1-R, respectively, are the most significant histo-blood group antigens in transfusion medicine.[1-3] In 1990, the molecular genetic basis of the ABO system was first elucidated by Yamamoto and coworkers.[4,5] They characterized the respective nucleotide sequences of the three major alleles, A^1, *B*, and *O*, of the *ABO* locus (Blood Group Antigen Gene Mutation Database,[6] http://www.ncbi.nlm.nih.gov/projects/mhc/xslcgi.fcgi?cmd=bgmut/home), and found that their cDNA structures are highly homologous. The coding regions of the A^1 and *B* allelic cDNAs have seven nucleotide dissimilarities (nucleotides 297, 526, 657, 703, 796, 803, and 930), and this results in differences at four amino acids (residues 176, 235, 266, and 268).[5,7] It is believed that these differences are responsible for the variation in donor-substrate specificities between the A and B transferases.[8-12]

The A and B antigens are present not only on the red blood cells (RBCs), but they are also widely distributed in the human body and its various mucous secretions.[2,3] It has been demonstrated that the expression of the A and B antigens is characterized by tissue-specific and temporally variant patterns. Marked change in these antigens has been observed for different embryo-development stages[13] and in cell differentiation.[14,15] More interestingly, association between changes in A and B antigen expression and the processes underlying oncogenesis have also been noted. The absence of A and B antigen expression has been observed in many types of carcinoma, including lung, oral, ovarian, pancreas, stomach, and uterine cervix cancers[16-21] and hematologic malignancy.[22] Further, it has been demonstrated that loss of these antigens precedes tumor metastasis and is associated with tumor grade[16,18,19] and survival rates of patients.[17,18,20] Thus, understanding the regulatory mechanics of *ABO* gene expression is of particular interest as it is a prerequisite for further elucidation of the determinants that lead to altered expression of this gene during oncogenesis.

Elaboration of the genomic organization of the *ABO* gene rapidly followed elucidation of the molecular genetic basis of the ABO system.[23,24] It was demonstrated that the

ABO gene consisted of seven exons that span approximately 20 kb of genomic DNA. Recently, through utilization of an alternative promoter another variant form of *ABO* gene expression, which contains a different starting exon 1 region, was identified.[25,26] The establishment of these basics of *ABO* gene organization has allowed analysis of its genomic sequence[27,28] and exploration of the regulatory mechanisms of its expression. It has been demonstrated that DNA methylation of the *ABO*-gene promoter region suppresses expression of the *ABO* gene and plays an important role in the regulation of its expression.[29] Nonetheless, Kominato and associates[30] proved that the activation of the *ABO* gene transcription is dependent on binding of the transcription factor CBF/NF-Y to the minisatellite sequence of four tandem repeats of a 43-bp unit, located –3.7 kb from the transcription initiation site, in KATO III gastric cancer cell line.[30] Previous investigations, however, have demonstrated that the structure of this minisatellite-enhancer region is specific to *A*, *B*, and *O* alleles.[31-33] The structure with the four tandem repeats of the 43-bp unit is present in the *B* gene, but not in the A^1 gene. These works demonstrated that the corresponding region of the A^1 gene possesses only a single 43-bp unit. Transcriptional activities contributed by the two different enhancer structures of the A^1 and *B* genes were compared with reporter assays, with the results demonstrating much greater transcription-induction activity in the *B*-gene minisatellite-enhancer structure relative to the *A*-gene analog.[33]

Based on these previous observations, the expression of the *A* and *B* genes was compared in the present study. It was demonstrated that there was evidently more of the *B* transcript relative to the *A* transcript in the RNA samples prepared from *AB*-genotype cells. It is suggested that the main reason for the variant expression of the *A* and *B* genes is their different minisatellite-enhancer structures.

MATERIALS AND METHODS

Semiquantitative comparison of the *A* and *B* transcripts by real-time polymerase chain reaction

Total RNA samples of six unrelated individuals with the group A_1B phenotype were prepared from their peripheral blood cells with a RNA blood mini kit (QIAamp, Qiagen GmbH, Hilden, Germany). Total RNA of the SW48 colorectal adenocarcinoma cell line (American Type Culture Collection, Manassas, VA), which originated from a group AB patient, was prepared with the a mini kit (RNeasy, Qiagen). The RNA samples were treated with RNase-free DNase I before synthesis of cDNA. The first-strand cDNAs were primed with oligo(dT) primer and synthesized by reverse transcriptase (SuperScrip III, Invitrogen, Carlsbad, CA).

Based on the polymorphic nucleotides at positions 796 and 803 of *A* and *B* cDNAs, primer sets with a common forward primer, F68 (GCCATCAAGAAATACGTGGCT, nucleotides 364 through 384 of *ABO* cDNA, translational start ATG as nucleotides +1 to +3), and the reverse primers, AR5 (GAAGAACCCCCCCAGGTA, complementary to nucleotides 793 through 810 of *A* cDNA) and BR7 (GAAGAACGCCCCCATGTA, complementary to nucleotides 793 through 810 of *B* cDNA), were used to specifically amplify the *A* and *B* cDNAs, respectively. The AR5 and BR7 primers were designed to anneal to the *A* gene with the 796C and 803G nucleotides and to the *B* gene with the 796A and 803C nucleotides, respectively. Serial dilution of the pCRII-TOPO plasmid vectors (Invitrogen) containing *A* or *B* cDNA segments (nucleotides 226 through 1062) was used to generate standard curves for quantification of *A* and *B* cDNAs, respectively. The cDNA sample and 5 pmol of each forward (F68) and reverse (AR5 or BR6) primer were added to 20 µL of polymerase chain reaction (PCR) mixture combined from a fast start DNA master SYBR Green I kit (LightCycler, Roche Diagnostics GmbH, Mannheim, Germany). Equal amounts of each cDNA sample were used to quantify the *A* and *B* transcripts with real-time PCR detection, which was accomplished with the LightCycler system (Roche). The PCR program included 10 minutes at 95°C followed by 45 cycles of 0 seconds at 95°C, 5 seconds at 63°C, and 15 seconds at 72°C.

Semiquantitative comparison of the *A* and *B* transcripts by PCR-restriction fragment length polymorphism analysis

Based on the presence of the polymorphic nucleotide at position 526 of *A* and *B* cDNAs, a PCR-based restriction fragment length polymorphism (RFLP) analysis was developed to compare the *A* and *B* transcripts in cDNAs prepared from group A_1B individuals. The 526C nucleotide in the *A* gene creates a *Bss*HII recognition sequence, whereas the 526G nucleotide in the *B* gene creates a *Bsa*HI recognition sequence. The cDNA sample and 5 pmol of each forward (GTGTTTGCCATCAAGAAATACGTGGCT TTCC, nucleotides 358 through 388 of *ABO* cDNA) and reverse (TCGATGCCGTTGGCCTGGTCGACCATCAT, complementary to nucleotides 862 through 890 of *ABO* cDNA) primer were combined in 12.5 µL of PCR buffer containing 0.2 mmol per L of dNTP and 0.625 U of DNA polymerase (Expand High Fidelity^{PLUS} DNA polymerase, Roche). The standard pCRII-TOPO plasmid vectors containing *A* or *B* cDNA segments, as used in real-time PCR detection, served as control templates. The PCR program included 2 minutes at 95°C followed by 40 cycles of 10 seconds at 95°C and 1 minute at 72°C. After the PCR procedure, each reacted mixture was supplemented with an equal amount of fresh PCR mixture containing all the PCR reagents, except template cDNA, and then subjected to 2 minutes at 95°C and 10 minutes at 72°C. This supplemental step was used to eliminate the formation of heteroduplex product. The 533-bp PCR products were subjected to

digestion by *Bss*HII and *Bsa*HI restriction endonucleases separately and then analyzed by 2.0 percent agarose gel electrophoresis. After staining with ethidium bromide, the gel image was captured with a gel documentation system (UVIdoc, UVItec Limited, Cambridge, UK), and the fluorescent illumination of the digested PCR products was quantified with computer software (Science Laboratory 2001 Image Gauge 4.0, Fuji Photo Film Co. Ltd, Tokyo, Japan).

Sequence analysis of 5′ regions of the *A* and *B* genes

The 5′ segments of nucleotides –3753 through +23 (numbered according to the *A* gene and translational start ATG as +1 to +3) of the *A* and *B* genes were amplified by PCR and cloned, and the sequences determined. Genomic DNA samples from one individual homozygous for the A^{1v} (*ABO*A102*) gene and another homozygous for the *B* (*ABO*B101*) gene were used as templates to obtain the 5′ regions of the *A* and *B* genes, respectively. Fifty nanograms of genomic DNA and 5 pmol of each forward (F10, aattacgcGTTTAAGCATTAGCCTAAATCCTACCCCTACAACC) and reverse (R11, ttaagctAGCGTCCGCAACACCTCGGCCATGGCT, antisense sequence) primer were combined in 12.5 μL of PCR buffer containing 0.2 mmol per L dNTP and 0.625 U of Expand High Fidelity^{PLUS} DNA polymerase. The F10 and R11 primers contained recognition sequences for *Mlu*I and *Nhe*I restriction endonuclease (underlined) at their 5′ ends, respectively. The PCR program included 2 minutes at 95°C followed by 40 cycles of 10 seconds at 94°C, 30 seconds at 68°C, and 4 minutes at 72°C. The PCR products were cloned into the pCRII-TOPO vectors with a TOPO TA cloning kit (Invitrogen). DNA sequences were determined with a cycle sequencing kit (BigDye Terminator, Applied Biosystems, Foster City, CA). Multiple clones from two batches of PCR products were sequenced to identify PCR errors from actual sequence polymorphisms.

Reporter assay

The DNA fragments encompassing the 5′ –3753 to +23-bp regions of the *A* and *B* genes, as described above, were inserted into the *Mlu*I and *Nhe*I sites of the pGL3-basic vector (Promega Corp., Madison, WI), which contains the reporter of the firefly *luciferase* (*luc*) gene. Recombinant plasmids with the 5′ 3.7-kb segments of the *A* and *B* genes introduced upstream of the *luc* reporter gene were selected and designated as plasmids A3.7 and B3.7, respectively. The minisatellite-enhancer region of the *ABO* gene locates at position –3708 to –3666. The DNA fragments encompassing the 5′ 3.6-kb regions (nucleotides –3639 through +23) of the *A* and *B* genes were amplified by PCR with the forward primer (aattacgcGTTCAGTGTGCTCTTGGGATTGTAAC), which contained *Mlu*I recognition sequence (underlined) at the 5′ end, and the R11 reverse primer. The amplified 3.6-kb fragments of the *A* and *B* genes were cloned into the *Mlu*I and *Nhe*I sites of the pGL3-basic vector, yielding the reporter plasmids A3.6 and B3.6, which were devoid of the minisatellite-enhancer region of the *A* and *B* genes, respectively. To construct reporter plasmids containing the 5′ 3.7-kb regions of the *A* and *B* genes, but with the *A*- and *B*-genes' minisatellite-enhancer regions exchanged, the A3.7 and B3.7 reporter plasmids were digested with *Mlu*I (recognition position at –3753) and *Nsi*I (recognition position at –3004) restriction endonucleases. The excised 750-bp fragments from the A3.7 and B3.7 plasmids were exchanged and inserted into the cleaved A3.7 and B3.7 plasmids, respectively, yielding the reporter plasmids A3.7-4X43 and B3.7-1X43, which had undergone an exchange of *A*- and *B*-genes' minisatellite regions.

The regions between –117 bp and the transcription start site are believed to function as the *ABO*-gene promoter.[30] The *ABO*-gene promoter region (–118 to +32 bp) and the 152- and 281-bp fragments encompassing the minisatellite enhancers of the *A* and *B* genes, respectively, were PCR-amplified as described previously.[33] The amplified *ABO*-gene promoter fragment was then inserted into the *Bgl*II site of the pGL3-basic vector, with the orientation of the insert of the selected recombinant plasmid verified by DNA sequence analysis. This constructed the reporter plasmid ABOP with the *ABO*-gene promoter region introduced upstream of the *luc* reporter. The 152- and 281-bp PCR products of the *A*- and *B*-gene minisatellite-enhancer regions, respectively, were introduced separately into the *Kpn*I and *Sac*I sites of the ABOP plasmid, constructing the reporter plasmids AEP and BEP with the introduction of *A*- and *B*-gene minisatellite-enhancer regions upstream to the *ABO*-gene promoter, respectively.

The constructed and the mock pGL3-basic plasmids were prepared with a plasmid kit (EndoFree, Qiagen) for transfection. The human gastric carcinoma cell line KATO III (American Type Culture Collection) was used as a host in the reporter assay. Transfection of the cells was performed with transfection reagent (Lipofectamine, Invitrogen; supplemented with Plus reagent). Cells were split into six-well culture plates at a density of 2×10^5 cells per mL and transfected with 1.0 μg of reporter plasmid. After incubation for 48 hours after transfection, cells were harvested to analyze the activity of the expressed luciferase with a luciferase assay system (Bright-Glo, Promega).

RESULTS

Evident predominance of *B* transcript relative to *A* transcript in *AB*-genotype RNA samples

As described under Materials and Methods, the primer pairs of F68+AR5 and F68+BR7, designed based on the

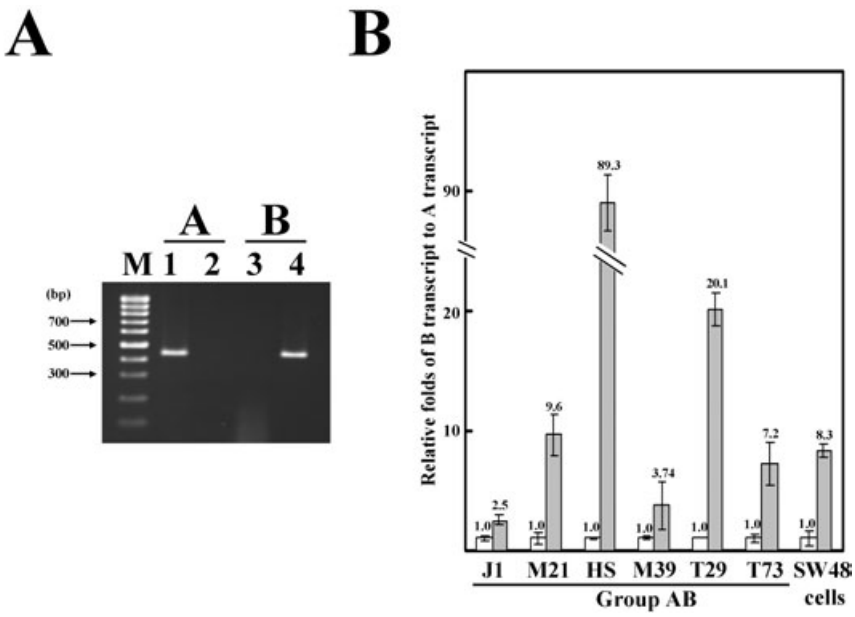

Fig. 1. Quantitative comparison of the *A* and *B* transcripts in *AB*-genotype samples by real-time PCR. (A) Specificity of the real-time PCR detection for the *A* and *B* transcripts. The primer pairs F68+AR5 (Lanes 1 and 3) and F68+BR7 (Lanes 2 and 4) were used in the real-time PCR, and plasmid vectors bearing *A* (Lanes 1 and 2) or *B* (Lanes 3 and 4) cDNA segments were used as templates. As shown in the figure, the F68+AR5 and F68+BR7 primers specifically amplified the *A* and *B* cDNA segments, respectively. (B) The mean multiples of *B* transcript to *A* transcript in *AB*-genotype cDNAs. Standard curves for quantification of *A* and *B* cDNAs were generated from serial dilution of the plasmid vectors containing *A* or *B* cDNA segments, respectively. Equal amounts of each cDNA sample, including six samples from unrelated group A_1B individuals (J1, M21, HS, M39, T29, T73) and the sample from the SW48 colon cancer cell line, which originated from a group AB patient, were used to quantify the *A* and *B* transcripts. The mean multiple of the *B* transcript (gray bar) to *A* transcript (open bar, arbitrarily given the value of 1.0) for each sample is shown. Data were obtained from three detections and standard deviations (SDs) are shown.

polymorphic nucleotides 796 and 803, specifically amplified the *A* and *B* cDNAs, respectively, in real-time PCR (Fig. 1A). The cDNA samples, which were prepared from six group AB individuals and from the SW48 colon cancer cell line, were used for quantification and comparison of the *A* and *B* transcripts. The results are summarized in Fig. 1B, with the ratio of the *B* and *A* transcripts in each sample presented. Although the *B*/*A* transcript ratio varied between samples, it was shown that the amount of *B* transcript was severalfold higher than that of the *A* transcript for all six of the cDNA samples from the group AB individuals, with the lowest and highest *B*/*A* ratios being 2.5 (J1 sample) and 89.3 (HS), respectively. A *B*/*A* transcript ratio of 8.3 was demonstrated in the SW48 cells.

A PCR-RFLP analysis, which was based on the polymorphic nucleotide 526, was developed to further demonstrate the different amounts of *A* and *B* transcripts in these samples. The design and procedures of this PCR-RFLP strategy are illustrated in Fig. 2A. Plasmid vectors bearing *A* or *B* cDNA segments (Fig. 2B, Lanes A and B, respectively) served as control templates. Gel image analysis of the cleaved products amplified from a mixture of equal amounts of the two control plasmids (Lanes A + B) showed that the *B*/*A*-cDNA ratios were close to 1 (0.95 and 1.01 in Lanes Bss and Bsa, respectively), demonstrating that the PCR amplification did not prefer either template. Five cDNA samples from the group AB individuals were subjected to analysis, and significant *B* transcript but scarcely detectable *A* transcript was demonstrated in four of these (M21, HS, M39, and T29). Although the *A* transcript was apparent in the J1 sample, image analysis showed that there was about 25 percent more of the *B* cDNA than the *A* cDNA. This observation of the J1 sample is consistent with the result of the real-time PCR analysis, in which the lowest *B*/*A* ratio was exhibited for the J1 cDNA. Both results of these two analyses demonstrate that the majority of the steady-state mRNA in the *AB*-genotype RNA samples is comprised of the *B* allele transcript.

Sequence comparison of the 5′ regions of the *A* and *B* genes

The 5′ regions encompassing nucleotides –3753 through +23 of the *A* and *B* genes were cloned and the sequence determined (GenBank/EBI DataBank Accession Numbers DQ231473 and DQ231474, respectively). Ten different positions were identified between the 5′ 3.7-kb regions of the *A* and *B* genes, as illustrated in Fig. 3 and described in its legend. The previously demonstrated difference between *A* and *B* alleles in the minisatellite-enhancer region, locating nucleotides –3708 through –3666, was confirmed. The minisatellite-enhancer region of the *B* gene is composed of four tandem repeats of the 43-bp unit. By contrast, only one 43-bp unit is present in the *A* gene, and an A>G change was present at the 41st nucleotide of the 43-bp consensus.

Different minisatellite sequences in the 5′ regions of the *A* and *B* genes contribute to specific transcription-activation activities in vitro

Reporter assay was employed to compare the transcription-activation activities of the 5′ regions of the *A* and *B*

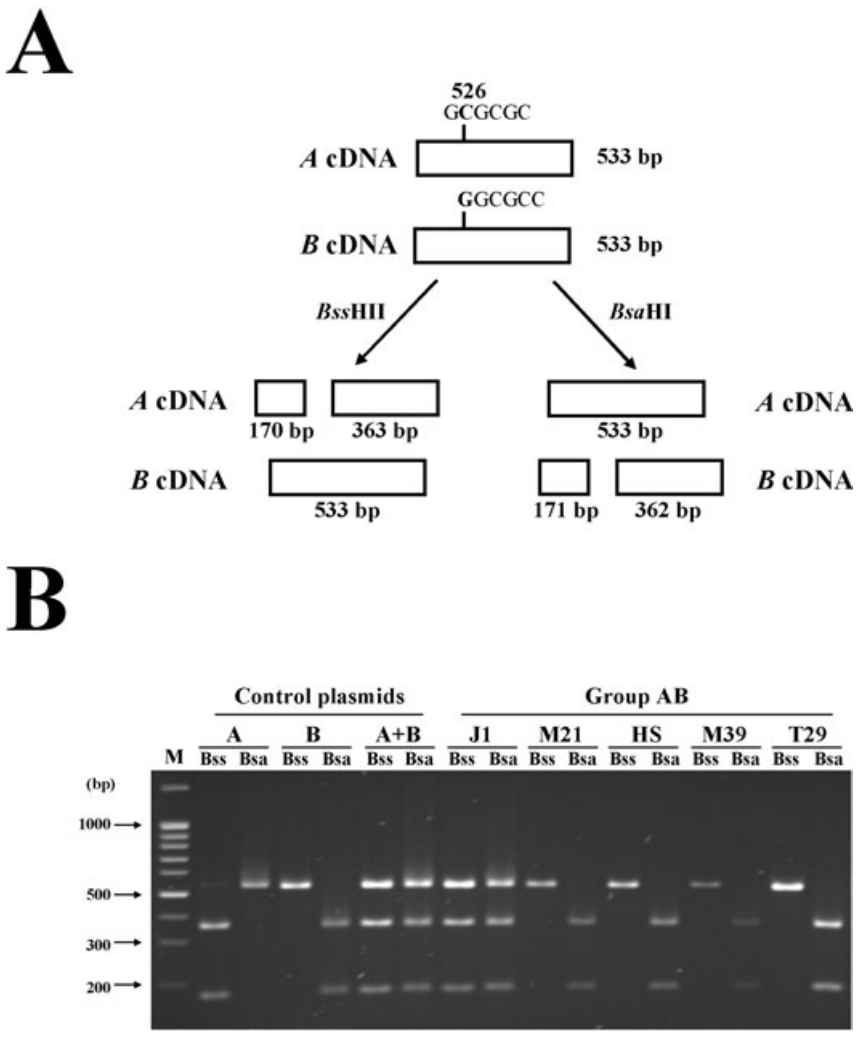

Fig. 2. Quantitative comparison of the *A* and *B* transcripts in *AB*-genotype samples by PCR-RFLP. (A) Schematic representation of the PCR-RFLP strategy. The 526C and 526G nucleotides in the respective *A* and *B* cDNAs create *Bss*HII (GCGCGC) and *Bsa*HI (GGCGCC) restriction recognition sequences, respectively. *Bss*HII digestion cleaves the 533-bp PCR product amplified from *A* cDNA into 363- and 170-bp fragments and leaves the PCR product amplified from *B* cDNA undigested, while *Bsa*HI digestion cleaves the PCR product amplified from *B* cDNA into 362- and 171-bp fragments and leaves the *A* cDNA PCR product intact. (B) Quantitative comparison of the *A* and *B* transcripts in *AB*-genotype samples by PCR-RFLP. Plasmid vectors bearing *A* or *B* cDNA segments served as control templates. The 533-bp PCR products, amplified from plasmid with *A* cDNA (lanes A), plasmid with *B* cDNA (lanes B), mixture of equal amounts of the two control plasmids (Lanes A + B), and five cDNA samples from group A_1B individuals, were subjected to digestion by *Bss*HII and *Bsa*HI restriction enzymes (Lanes Bss and Bsa, respectively) and analyzed by 2.0 percent agarose gel electrophoresis.

genes. As shown in Fig. 4A, the reporter plasmid B3.7, which had the introduction of the 5′ 3.7-kb region of the *B* gene upstream to the *luc* reporter of the pGL3-basic plasmid, exhibited markedly higher transcription-activation activity than the construct A3.7, which had the introduction of the *A*-gene 5′ 3.7-kb region. When the introduced fragments were shortened to about 3.6-kb, as in the constructs of A3.6 and B3.6, transcriptional activity was reduced, especially for the latter. The luciferase activity generated from the B3.6 construct was reduced to a level similar to those from the A3.7 and A3.6 plasmids. Of the 10 different positions identified in the 5′ 3.7-kb regions of the *A* and *B* genes, 9 were retained in the A3.6 and B3.6 constructs, and only the minisatellite-enhancer regions were deleted when compared with the A3.7 and B3.7 analogs. The results suggest that the minisatellite structure of the *B* genes is responsible for the high transcription-activation activity. Analysis of the A3.7-4X43 and B3.7-1X43 constructs, which were built with the minisatellite-enhancer regions of the *A* and *B* genes exchanged, further supports the above suggestion.

Another set of reporter plasmids, with *ABO*-gene promoter and the *A*- or *B*-gene minisatellite-enhancer regions inserted, was constructed. Introduction of the *ABO*-gene promoter to the upstream of the *luc* reporter (ABOP construct) resulted in a 12-fold increase in luciferase activity in the transfected KATO III cells when compared with the pGL3-basic vector (Fig. 4B). Further addition of the *A*-gene enhancer region of one 43-bp unit to the upstream of the *ABO* promoter increased the luciferase activity about 2.3-fold. When the *B*-gene enhancer region of four 43-bp units was added, however, the luciferase activity was increased approximately 200-fold relative to the ABOP plasmid. Reporter assay of these constructs demonstrated the high transcriptional activity of the four tandem repeats of the 43-bp unit of the *B* gene. This finding is consistent with the results of our previous report.[33]

DISCUSSION

In this study, we have demonstrated apparently greater quantities of *B* transcript relative to *A* transcript in RNA samples prepared from group A_1B individuals and from an *AB*-genotype cell line. The predominance of the *B* allele transcript was demonstrated in each of the samples

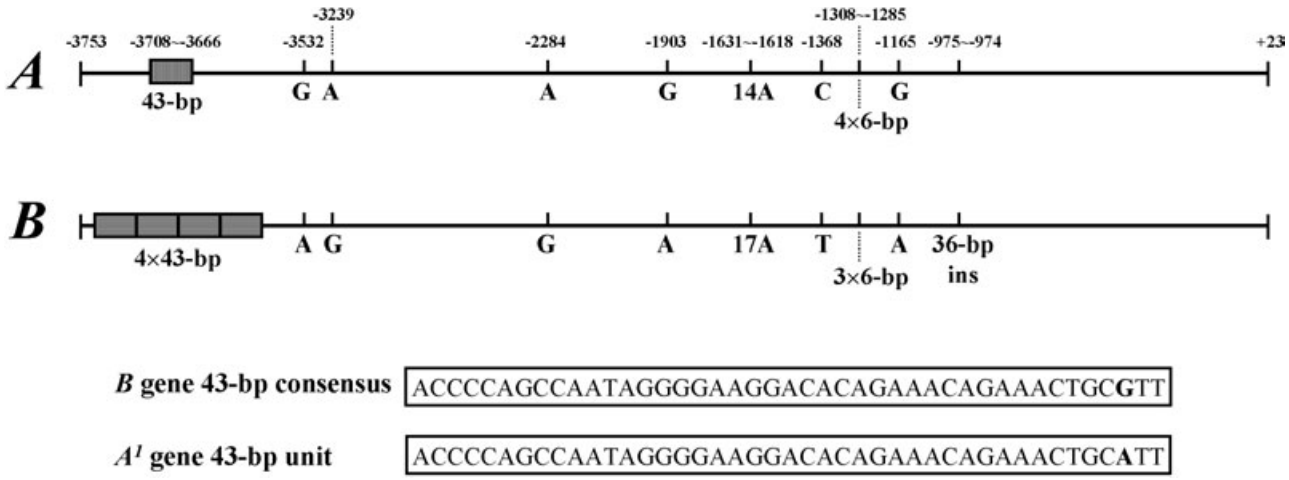

Fig. 3. Sequence comparison of the 5′ regions of the *A* and *B* genes. Ten different positions were identified in the 5′ 3.7-kb regions of the *A* and *B* genes and are schematically represented. The nucleotides are numbered according to the *A* gene and the translational start ATG are as nucleotides +1 to +3. Among the 10 different positions, six positions (–3532, –3239, –2284, –1903, –1368, and –1165) are single-nucleotide change. At the polymorphic poly(A) region of nucleotides –1631 through –1618, 14A and 17A nucleotides were found in the clones from *A* and *B* genes, respectively. At the region of nucleotides –1308 to –1285, four and three repeats of a 6-bp (ATTTTT) unit were identified in the *A* and *B* genes, respectively. In contrast to the *A* gene sequence, there is a 36-bp (TGAGGAATTGCCACAATTTTTTCCTGCGCCTGCACC) insertion between nucleotides –975 and –974 in the *B* gene. The minisatellite-enhancer regions, which are located at nucleotides –3708 through –3666, of the *A* and *B* genes are different. The region in the *B* gene comprises four tandem repeats of the 43-bp unit. Only one 43-bp unit is present in the region of the *A* gene, and an A>G change was present at the 41st nucleotide of the 43-bp consensus. The sequence of the 43-bp consensus in the *B* gene and the sequence of the 43-bp unit in the *A* gene are shown. The 5′ region of the A^{1v} (*ABO*A102*) gene was investigated and compared with that of the *B* (*ABO*B101*) gene in this study. A genomic sequence derived from the A^1 (*ABO*A101*) gene has been deposited in GenBank (Accession No. AC000397), and the 5′ sequence of the A^{1v} gene was found to be identical to that of the A^1 gene, except for the number of A nucleotides at the polymorphic poly(A) region.

assessed in this study, and both results of the two strategies, real-time PCR detection and PCR-RFLP analysis, demonstrate that the majority of the steady-state mRNA from the *AB*-genotype cells is composed of the *B* allele transcript. It was found, however, that the *B*/*A* transcript ratio varies drastically between group AB individuals. The underlying reason for this phenomenon awaits elucidation.

A previous brief report has indicated that the amount of *O* transcript is reduced relative to that of the *A* or *B* transcripts in *AO*- and *BO*-genotype RNA samples, and it has been suggested that the decreased stability of the *O* allele transcript, as resulted from the presence of a premature stop codon, accounts for this phenomenon.[34] There is no tenable basis, however, for inference of differential stability between the *A* and *B* transcripts, although the possibility cannot be totally excluded. Results obtained from sequence comparison of the 5′ regions of the *A* and *B* genes and reporter assays suggest that the increased amount of *B* allele transcript in the steady-state *AB*-genotype mRNA is due to the markedly higher transcriptional activity of the *B* gene relative to the *A* gene and that this is a consequence of the different minisatellite-enhancer structures in the two genes, which contain one and four repeats of 43-bp unit, respectively.

Minisatellite-enhancer specificity has been previously revealed for *A*, *B*, and *O* alleles.[31-33] The major difference is the number of repeats of the 43-bp unit to which the transcription factor CBF/NF-Y binds. In our previous investigation, excluding the A^1- and *B*-gene structures described above, a third minisatellite structure, four repeats of the 43-bp unit with a G>C substitution at the 41st nucleotide of the first of the 4 units, has been demonstrated in the O^{1v} gene and part of the O^1 gene.[33] It has been demonstrated that the A^2 allele contains four repeats of the 43-bp unit[31,32] and the infrequent O^2 allele contain one repeat of the 43-bp unit.[32] Except for the minisatellite enhancer region, other allele-specific polymorphisms in the 5′ regions of the *ABO* genes have also been documented.[35] Transient reporter assay has proved noticeably higher transcriptional activity for the enhancer with four repeats of the 43-bp unit, as derived from the *B* gene, compared to that of the analog with a single 43-bp unit, as derived from A^1 gene. Further, enhancement of transcriptional activity has been demonstrated with an increase in the repeat number of the 43-bp unit.[33]

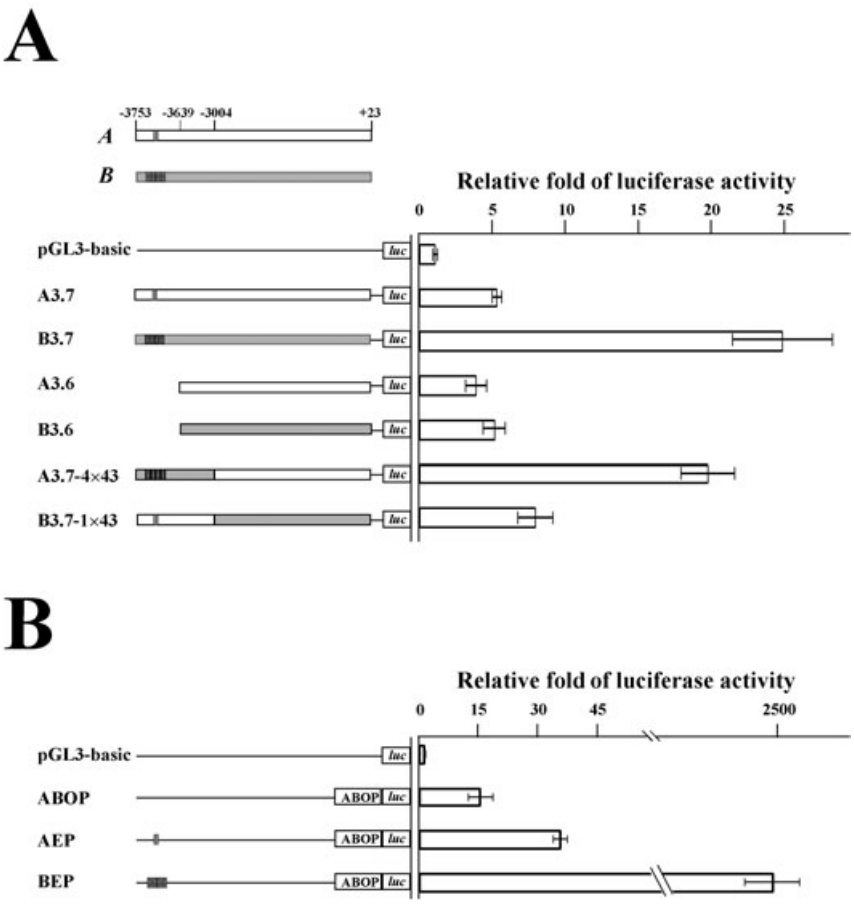

Fig. 4. Comparison of transcriptional activities of the 5′- and minisatellite-enhancer regions of the *A* and *B* genes. (A) Comparison of transcriptional activities of the 5′ regions of the *A* and *B* gene. The –3753 to +23-bp segments of the *A* and *B* genes and the positions of the minisatellite-enhancer regions are schematically represented, with the –3639- and –3004-nucleotide positions indicated. The A3.7 and B3.7 constructs have the –3753- to +23-bp regions of the *A* and *B* genes introduced upstream of the *luc* reporter gene of the pGL3-basic vector, respectively. The A3.6 and B3.6 constructs have the insertion shortened to the -3639-bp position. The A3.7-4X43 and B3.7-1X43 constructs were built with exchange of the segments from nucleotides –3753 through –3004 of the A3.7 and B3.7 constructs. The results are presented as mean multiples of luciferase activities from three repetitions compared with those from the pGL3-basic vector, which is arbitrarily given the value of 1.0 SDs are shown. (B) Comparison of transcriptional activities of the *A*- and *B*-gene minisatellite enhancers. The ABOP construct has the *ABO*-gene promoter region introduced upstream of the *luc* reporter of the pGL3-basic vector. The AEP and BEP constructs have the *A*- and *B*-gene enhancer regions, with respective one and four 43-bp units, inserted upstream of the *ABO*-gene promoter, respectively. The results are shown as mean multiples of luciferase activity compared with analogous activities for the pGL3-basic vector, which is arbitrarily given the value of 1.0. Data were obtained from three repetitions, and SDs are shown.

Doubts have been raised about the capability of the CBF/NF-Y factor to individually determine the complex expression patterns of the *ABO* gene, as CBF/NF-Y is known as a constitutive and ubiquitous transcription factor. In further investigations, it has been suggested that several other DNA segments in the 5′ region of the *ABO* gene play a regulatory role in its expression. These include the potential negative elements present within the –3.6- to –3.4-kb and –3.2- to –2.3-kb regions,[30] the negative regulatory N box in the –202- to –118-bp region,[36] and the Sp1-binding site in the promoter region.[37] As the minisatellite-enhancer structures for the *A* and *B* alleles differ, the expression of the allelic *A* and *B* genes should be influenced differentially by the CBF/NF-Y transcription factor. As shown in our reporter assay, deletion of the minisatellite enhancer in the *A* and *B* genes resulted in 25 and 80 percent reductions in luciferase activity, respectively (a comparison of the results for the A3.7, A3.6 and B3.7, B3.6 reporter plasmids provided in Fig. 4A). This demonstrates that the *A* gene possesses a very weak 43-bp minisatellite enhancer compared to the *B* gene in the cell line used in this study. Consequently other regulatory elements in the 5′ region might exert their function on the regulation of *A* gene expression more influentially. Thus, the *A* and *B* genes should be considered separately when examining the roles of the other regulatory elements, as those regulatory elements may have different weights in terms of the influence on the expression of the two genes. Furthermore, it should be noted that our reporter assays were performed in a gastric cancer cell line. Analysis for the transcriptional activity performed in erythroid precursor cells may be required to elucidate the influence of those regulatory elements in the transcriptions of the *A* and *B* genes in hematopoiesis.

It may be of interest to further determine whether the quantities of A- and B-transferase products differ in *AB*-genotype cells, given the differences in the quantities of the *A* and *B* transcripts. It should be noted, however, that, contradictory to the observation of greater expression of the *B* allele than the *A* allele in this study, previous investigation has shown that more A antigens than B antigens are present on group A_1B RBCs and that more A antigens are present on group A_1 RBCs than B

antigens on group B RBCs.[38] It is known that a series of complicated cellular events, including translation, post-translational modification, and glycosyltransferase reaction, is required to lead to the final expression of the A and B antigenic structures on cell membranes from the transcribed *A* and *B* mRNAs and that many factors may affect the processes of these cellular events. Further, the A and B antigenic structures are synthesized through the different reactions of the A and B transferases, which utilize specific donor substrates. Thus, it is interpretable that the predominance of the *B* allele transcript does not imply more final products of the B antigenic structure.

REFERENCES

1. Watkins WM. Molecular basis of antigenic specificity in the ABO, H and Lewis blood-group systems. In: Montreuil J, Schachter H, Vliegenthart JF, editors. Glycoproteins. Amsterdam: Elsevier Science; 1995. p. 313-90.
2. Daniels G. Human blood groups. Oxford: Blackwell Science; 2002.
3. Reid ME, Lomas-Francis C. The blood group antigen factsbook. San Diego: Elsevier Ltd.; 2004.
4. Yamamoto F, Marken J, Tsuji T, et al. Cloning and characterization of DNA complementary to human UDP-GalNAc: Fucα1→2Galα1→3GalNAc transferase (histo-blood group A transferase) mRNA. J Biol Chem 1990;265:1146-51.
5. Yamamoto F, Clausen H, White T, Marken J, Hakomori S. Molecular genetic basis of the histo-blood group ABO system. Nature 1990;345:229-33.
6. Blumenfeld OO, Patnaik SK. Allelic genes of blood group antigens: a source of human mutations and cSNP documented in the blood group antigen gene mutation database. Hum Mutat 2004;23:8-16.
7. Yamamoto F. Molecular genetics of the ABO histo-blood group system. Vox Sang 1995;69:1-7.
8. Yamamoto F, Hakomori S. Sugar-nucleotide donor specificity of histo-blood group A and B transferases is based on amino acid substitutions. J Biol Chem 1990;265:19257-62.
9. Yamamoto F, McNeil PD. Amino acid residue at codon 268 determines both activity and nucleotide-sugar donor substrate specificity of human histo-blood group A and B transferases. J Biol Chem 1996;271:10515-20.
10. Seto NO, Palcic MM, Compston CA, et al. Sequential interchange of four amino acids from blood group B to blood group A glycosyltransferase boosts catalytic activity and progressively modifies substrate recognition in human recombinant enzymes. J Biol Chem 1997;272:14133-8.
11. Seto NO, Compston CA, Evans SV, et al. Donor substrate specificity of recombinant human blood group A, B and hybrid A/B glycosyltransferases expressed in *Escherichia coli*. Eur J Biochem 1999;259:770-5.
12. Patenaude SI, Seto NO, Borisova SN, et al. The structural basis for specificity in human ABO(H) blood group biosynthesis. Nat Struct Biol 2002;9:685-90.
13. Szulman AE. The histological distribution of the blood group substances in man as disclosed by immunofluorescence. J Exp Med 1964;119:503-16.
14. Dabelsteen E, Vedtofte P, Hakomori S, Young WW. Carbohydrate chains specific for blood group antigens in differentiation of human oral epithelium. J Invest Dermatol 1982; 79:3-7.
15. Bony V, Gane P, Bailly P, Cartron JP. Timecourse expression of polypeptides carrying blood group antigens during human erythroid differentiation. Br J Haematol 1999;107: 263-74.
16. Davidsohn I. Early immunologic diagnosis and prognosis of carcinoma. Am J Clin Pathol 1972;57:715-30.
17. Lee JS, Ro JY, Sahin AA, et al. Expression of blood-group antigen A—a favorable prognostic factor in non-small-cell lung cancer. N Engl J Med 1991;324:1084-90.
18. Matsumoto H, Muramatsu H, Shimotakahara T, et al. Correlation of expression of ABH blood group carbohydrate antigens with metastatic potential in human lung carcinomas. Cancer 1993;72:75-81.
19. Welshinger M, Finstad CL, Venkatraman E, et al. Expression of A, B, and H blood group antigens in epithelial ovarian cancer: relationship to tumor grade and patient survival. Gynecol Oncol 1996;62:106-12.
20. Graziano SL, Tatum AH, Gonchoroff NJ, Newman NB, Kohman LJ. Blood group antigen A and flow cytometric analysis in resected early-stage non-small cell lung cancer. Clin Cancer Res 1997;3:87-93.
21. Dabelsteen E, Gao S. ABO blood-group antigens in oral cancer. J Dent Res 2004;84:21-8.
22. Bianco T, Farmer BJ, Sage RE, Dobrovic A. Loss of red cell A, B, and H antigens is frequent in myeloid malignancies. Blood 2001;97:3633-9.
23. Yamamoto F, McNeil PD, Hakomori S. Genomic organization of human histo-blood group ABO genes. Glycobiology 1995;5:51-8.
24. Bennett EP, Steffensen R, Clausen H, Weghuis DO, van Kessel AG. Genomic cloning of the human histo-blood group ABO locus. Biochem Biophys Res Commun 1995; 206:318-25;erratum in Biochem Biophys Res Commun 1995;211:347-8.
25. Kominato Y, Hata Y, Takizawa H, et al. Alternative promoter identified between a hypermethylated upstream region of repetitive elements and a CpG island in human ABO histo-blood group genes. J Biol Chem 2002;277:37936-48.
26. Hata Y, Kominato Y, Takizawa H, et al. Transcription starting from an alternative promoter leads to the expression of the human ABO histo-blood group antigen. Transfusion 2003;43:656-62.
27. Seltsam A, Hallensleben M, Kollmann A, Blasczyk R. The nature of diversity and diversification at the *ABO* locus. Blood 2003;102:3035-42.
28. Hosseini-Maaf B, Hellberg Å, Rodrigues MJ, Chester MA, Olsson ML. *ABO* exon and intron analysis in individuals with the $A_{weak}B$ phenotype reveals a novel O^{1v}-A^2 hybrid allele that

causes four missense mutation in the A transferase. BMC Genet 2003;4:17.
29. Kominato Y, Hata Y, Takizawa H, et al. Expression of human histo-blood group ABO genes is dependent upon DNA methylation of the promoter region. J Biol Chem 1999; 274:37240-50.
30. Kominato Y, Tsuchiya T, Hata N, Takizawa H, Yamamoto F. Transcription of human ABO histo-blood group genes is dependent upon binding of transcription factor CBF/NF-Y to minisatellite sequence. J Biol Chem 1997;272:25890-8.
31. Shimada I, Komonato Y, Hata N, Takizawa H. DNA polymorphisms in the 5′-flanking sequence of the human ABO blood group and their association with the alleles for the common ABO phenotypes. Legal Med 1999;1:217-25.
32. Irshaid NM, Chester MA, Olsson ML. Allele-related variation in minisatellite repeats involved in the transcription of the blood group *ABO* gene. Transfus Med 1999;9:219-26.
33. Yu LC, Chang CY, Twu YC, Lin M. Human histo-blood group *ABO* glycosyltransferase genes: Different enhancer structures with different transcriptional activities. Biochem Biophys Res Commun 2000;273:459-6.
34. O'Keefe DS, Dobrovic A. Decreased stability of the O allele mRNA transcript of the ABO gene. Blood 1996;87:3061-2.
35. Thuresson B, Hosseini-Maaf B, Chester MA, Olsson ML. Allele-related motifs in the regulating region of the *ABO* gene. Vox Sang 2004;87(Suppl 3):15.
36. Kominato Y, Hata Y, Matsui K, et al. Transcription regulation of the human *ABO* histo-blood group gene is dependent on the N box upstream of the proximal promoter. Transfusion 2004;44:1741-9.
37. Hata Y, Kominato Y, Yamamoto F, Takizawa H. Characterization of the human ABO gene promoter in erythroid cell lineage. Vox Sang 2002;82:39-46.
38. Economidou J, Hughes-Jones NC, Gardner B. Quantitative measurements concerning A and B antigen sites. Vox Sang 1967;12:321-8.