BMC Genetics

BioMed Central

Research article

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 mutations in the A transferase

Bahram Hosseini-Maaf[1], Åsa Hellberg[1], Maria J Rodrigues[2], M Alan Chester*[1] and Martin L Olsson[1]

Address: [1]Dept. of Transfusion Medicine, Institute of Laboratory Medicine, Lund University & Blood Centre, University Hospital, SE-221 85 Lund, Sweden and [2]Centro Regional de Sangue de Lisboa/IPS, Lisboa, Portugal

Email: Bahram Hosseini-Maaf - bahram.hosseini_maaf@transfumed.lu.se; Åsa Hellberg - asa.hellberg@transfumed.lu.se; Maria J Rodrigues - mjrodrigues@crsl.ips.min-saude.pt; M Alan Chester* - alan.chester@transfumed.lu.se; Martin L Olsson - Martin_L.Olsson@transfumed.lu.se

* Corresponding author

Published: 17 November 2003

BMC Genetics 2003, **4**:17

Received: 05 September 2003
Accepted: 17 November 2003

This article is available from: http://www.biomedcentral.com/1471-2156/4/17

Abstract

Background: Since the cloning in 1990 of cDNA corresponding to mRNA transcribed at the blood-group ABO locus, polymorphisms due to ethnic and/or phenotypic variations have been reported. Some subgroups have been explained at the molecular level, but unresolved samples are frequently encountered in the reference laboratory.

Results: ABO blood grouping discrepancies were investigated serologically and by *ABO* genotyping [duplex polymerase-chain-reaction (PCR) – restriction-fragment-length-polymorphism (RFLP) and PCR – allele-specific-primer (ASP) across intron 6] and DNA sequencing of the *ABO* gene and its proposed regulatory elements. Blood samples from five individuals living in Portugal, Switzerland, Sweden and the USA were analysed. These individuals were confirmed to be of Black ethnic origin and had the unusual $A_{weak}B$ phenotype but appeared to have the A^2B genotype without previously reported mutations associated with weak A or B expression. Sequencing of this *A* allele (having 467C>T and 1061delC associated with the common A^2 [A201] allele) revealed three mutations regularly encountered in the O^{1v} [O02] allele: 106C>T (Val36Phe), 188G>A (Arg63His), 220C>T (Pro74Ser) in exons 3, 4 and 5, respectively. The additional presence of 46G>A (Ala16Thr) was noted, whilst 189C>T that normally accompanies 188G>A in O^{1v} was missing, as were all O^{1v}-related mutations in exons 6 and 7 (261delG, 297A>G, 646T>A, 681G>A, 771C>T and 829G>A). On screening other samples, 46G>A was absent, but two new *O* alleles were found, a Jordanian O^1 and an African O^{1v} allele having 188G>A but lacking 189C>T. Sequencing of introns 2, 3, 4 and 5 in common alleles (A^1 [A101], A^2, *B* [B101], O^1, $O^{1v}$and O^2 [O03]) revealed 7, 12, 17 and 8 polymorphic positions, respectively, suggesting that alleles could be defined by intronic sequences. These polymorphic sites allowed definition of a breakpoint in intron 5 where the O^{1v}-related sequence was fused with A^2 to form the new hybrid. Intron 6 has previously been sequenced. Four new mutations were detected in the hybrid allele and these were subsequently also found in intron 6 of A^2 alleles in other Black African samples.

Conclusions: A novel O^{1v}-A^2 hybrid was defined by ABO exon/intron analysis in five unrelated individuals of African descent with the $A_{weak}B$ blood group phenotype.

Background

The ABO blood group system is the most clinically significant system in transfusion and transplantation medicine. The *A* and *B* genes are co-dominant, so individuals can be phenotyped as A, B, AB or O. A common dimorphism in some populations leads to the division of blood group A into A_1 and A_2, the latter showing weaker antigenicity. On occasion, A_2 activity is weakened further when competition due to a co-dominant *B* gene occurs (A_2B phenotype).

ABO allele nomenclature poses significant problems that are still under consideration by the International Society of Blood Transfusion (ISBT). In the absence of an officially agreed terminology, alleles are referred to here by their serological activity and an alternative allele name[1] is given in square brackets.

The *ABO* gene contains seven exons and a 1062 base pair (bp) sequence codes for a glycosyltransferase, of which the A^1 [A101] allele (usually described as the consensus allele) product adds a monosaccharide (*N*-acetyl-α-D-galactosamine) to a specific acceptor glycoconjugate. The *B* [B101] allele produces a similar protein, differing only in four amino acids, but these changes result in a modified enzyme specificity [2]. The same acceptor glycoconjugate is utilised, but the B-specific monosaccharide, α-D-galactose, is added (for a review, see [3]). The coding region of the A^2 [A201] allele differs from A^1 by a seemingly innocuous 467C>T (Pro156Leu) substitution and a 1061delC that causes a frame-shift and extends the reading frame by 64 nucleotides [4]. Many other mutations have been described [5] that weaken these activities and result in weak A or B subgroups. Totally inactivating mutations result in the blood group O phenotype, the most common of which is a deletion at nucleotide (nt.) 261 (261delG) in exon 6 that results in a frame-shift and premature termination of translation [2]. Two major alleles of this type exist. The first allele (O^1 [O01]) differs from the consensus *A* [A101] allele in all seven exons by only this mutation. The second allele (O^{1v} [O02]) [6,7] has nine further mutations spread across exons 3–7 in addition to 261delG and an additional 13 mutations have been found amongst the 1052 nucleotides in intron 6 [8,9].

Other alleles have been described that are predicted to lead to amino acid substitutions or frame-shifts [5], or alternative splicing [10,11] and a dimorphic enhancer region has been reported [12-14]. Several alleles at the *ABO* locus appear to arise by crossing over between two dissimilar alleles. First described in 1996 [15], a *B* allele was continued as an O^{1v} allele from downstream of the middle of the last and largest exon (exon 7). The B-enzyme-determining mutations occur after the cross-over point and the O^{1v} sequence after this point only differs from the consensus (A^1) sequence by two point mutations, of which only one would lead to an amino acid change (Val277Met). This change presumably weakened the A enzyme activity, as the resulting phenotype was A_2, albeit from a *B*-O^{1v} hybrid allele [A204]. Other examples on the same theme include several A^x alleles based on the 5'-ends of *A* or *B* alleles in combination with the 3'-end of the O^{1v} allele [9,16]. However, many of the hybrid alleles described so far are hybrids of O^1 and O^{1v}, have 261delG, and hence are *O* (i.e. inactive) alleles [17].

We describe here a new hybrid *A* allele having O^{1v}-A^2 allele characteristics in five individuals of $A_{weak}B$ phenotype from different parts of the world, but having a common ethnic background. This is the first allele described in which the effect on enzyme activity of the mutations in the O^{1v} allele prior to exon 6 can be shown.

Results

Blood group serology

The five individuals from Portugal (origin: Guinea Bissau), Sweden (n = 2, origins: Senegal and Zimbabwe), Switzerland (origin: the Dominican Republic), and the USA were all Black and of African or Afro-American descent. On routine typing they were found to express the A antigen weakly on their red blood cells (RBC) whilst the B antigen was expressed normally. Fresh blood samples from all five subjects were subjected to an extended serological analysis at the referring transfusion centres. Three of the samples were analysed with a panel of commercially available polyclonal and monoclonal anti-A reagents (described previously in [18]) for further characterisation of the weak A phenotype. One monoclonal reagent (Seraclone) agglutinated the RBC almost completely (3+ reaction with few unagglutinated cells) but the other reagents gave mixed field reactions (macroscopic reading ranging from negative to 1+/2+). All reactions were also read microscopically and weak agglutination with mixed field was verified also in the reagents yielding macroscopically negative readings. Adsorption of red cells with anti-A and subsequent elution yielded an eluate containing anti-A (only performed in one case). Anti-A_1 was present in serum from all subjects but one also had a weak anti-A reactive only after

incubation at 4°C. Saliva testing was not performed due to lack of saliva samples and/or non-secretor status as judged by Lewis blood group phenotyping.

Blood samples from the parents of the Portuguese individual were also investigated. Whilst the father had the common B phenotype, the phenotype of the mother was A_{weak} but with a strength of reactions suggesting an intermediate form of A antigen expression, weaker than A_2 but stronger than the one observed in the $A_{weak}B$ cases studied here.

Blood group genotyping

All samples were found on initial genotype screening [19] to have the *A*^{2}*B* genotype and lacked the majority of previously reported mutations associated with weak A or B expression [16,20]. Further analysis by *A*^{2}-specific PCR-ASP [18] established the presence of the 1061delC mutation [4] and *B*-specific PCR-ASP [16] indicated the presence of a normal *B* allele [21]. In addition, enhancer minisatellite PCR analysis indicated that four of these samples were homozygous for four 43 bp-repeats in the CBF-NF/Y-binding domain approximately 4 kbp upstream from the translation start codon, as expected [13]. Interestingly, the individual originating from the Dominican Republic was heterozygous for one and four repeats, thus deviating from the rule [13].

A *BstUI*-based PCR-RFLP test [7] results in cleavage between 188C and 189G in exon 4 of all known alleles except *O*^{1v} (and some minor variants of *O*^{1} and *O*^{1v} [5]), which has two mutations at these positions (188G>A and 189C>T). The indication that these five individuals already genotyped as *A*^{2}*B* were also heterozygous for an *O*^{1v} allele-specific mutation led us to sequence exons 1–7 to investigate the reason for this anomaly.

Based on information obtained about the geographic and ethnic origin of the individuals with the anomalous *A* allele, we also performed genomic typing of the *FY* and *RHD* blood group loci according to published methods [22,23]. The silent *Fy* allele (based on the *Fy*^{b} sequence with a disrupted GATA-1-binding motif in the promoter region [24]) commonly found in individuals of Black ethnic origin but not in Caucasians was detected in the homozygous or heterozygous state in all five cases whilst the *RHD* pseudogene [23] exclusively found in some Blacks was detected in the Portuguese sample only.

DNA was also isolated from blood from the parents of the Portuguese $A_{weak}B$ individual. The mother was genotyped as *A*^{2}*O*^{1v} and the father as *BO*^{1}. Further investigation of the mother's *A*^{2}-like allele showed the same hybrid allele as her son and the other four index cases, as expected.

ABO *exon sequences*

The complete coding region (all seven exons, comprising 1128 bp) of the *A*^{2} allele was sequenced in all five samples.

The nucleotide sequence in exons 6 and 7 was identical to the non-Asian *A*^{1} consensus allele except for the characteristic mutations (467C>T and 1061delC in exon 7) found in the common *A*^{2} allele.

The sequences in exons 1–5 were similar to the consensus allele, except for 106G>T (exon 3), 188G>A (exon 4) and 220C>T (exon 5), all consistent with the presence of an *O*^{1v} allele. Surprisingly, the *O*^{1v}-specific mutation, 189C>T, was absent, whereas a hitherto undescribed mutation in exon 2 (46G>A) was found in all five samples. As opposed to the situation in *O*^{1v}, the effect of these missense mutations can be seen in the translated transferase, as the *O*^{1}/*O*^{1v}-specific 261delG mutation is not present in this allele. The amino acids predicted to occur as well as their relative locations in the translated protein are shown in Figure 1.

A possible hybrid allele breakpoint was suspected somewhere between the end of exon 5 and the beginning of exon 6. We therefore sequenced the intervening intron 5 (see below) in an attempt to localise the cross-over region.

None of the more than 100 samples screened by PCR-ASP (i.e. at least 200 alleles, comprising over 50 *O*^{1} and *O*^{1v} alleles and at least 10 *A*^{1}, *A*^{2}, *B*, *O*^{2} and *O* hybrid alleles, as well as several weak *A* alleles) had the new mutation, 46G>A (Ala16Thr). Amongst all these samples one *O*^{1} allele from a Jordanian individual was found to have the *O*^{1v}-characteristic 188G>A (Arg63His) mutation. One *O*^{1v} allele from an African individual lacked the *O*^{1v}-characteristic 189C>T (silent) mutation.

ABO *intron sequences*

Intron 5

Intron 5 from a number of individuals with the common alleles, *A*^{1} (two homozygotes), *A*^{2} (two homozygotes), *B* (one homozygote, one heterozygote), *O*^{1} (two homozygotes), *O*^{1v} (two homozygotes),*O*^{2} [O03] (three heterozygotes) and the new suspected hybrid alleles was sequenced. Three *A*^{2} and three *O*^{1v} alleles from Black African donors were also sequenced for comparison. Eight polymorphic nucleotide positions were found in the 554 bp sequence and these are shown in Table 1. Intron 5 sequences in the Black African samples were identical to the respective alleles of the other (Caucasian) control material. The sequence in the new allele was consistent with a crossover between an *O*^{1v} and an *A*^{2} allele after nt. 103insCCC but before 306C in intron 5. This is shown schematically in Figure 2.

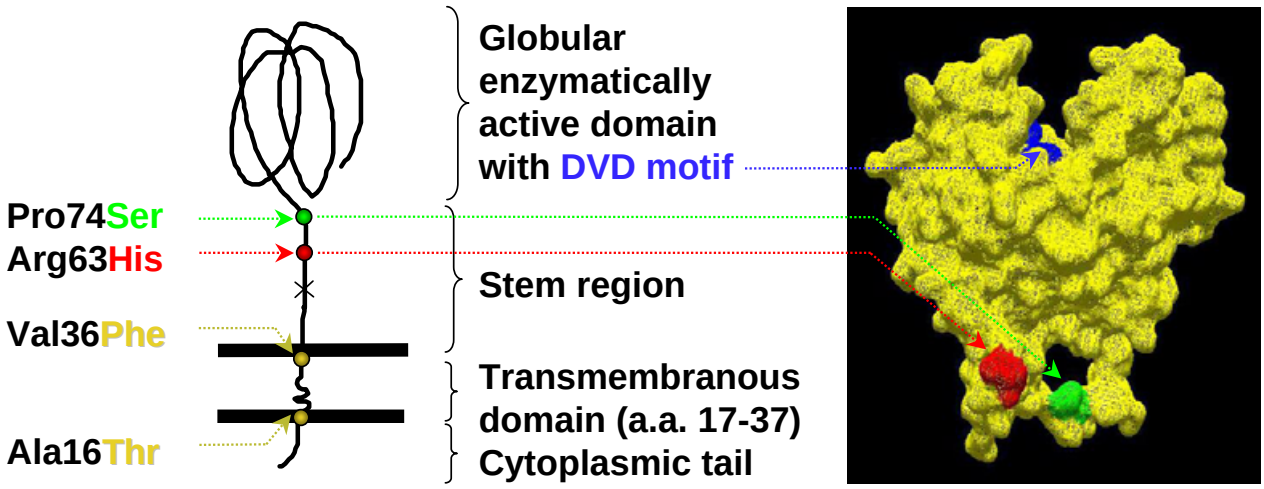

Figure 1
Depiction of the relative locations of amino acid substitutions caused by the novel allele. A schematic representation of the A transferase resulting from transcription and translation of the novel allele is shown on the left side whilst a computed 3D-model on black background is displayed on the right side. In the former, all amino acid substitutions resulting from the four missense mutations in the O^{1v}-A^2 hybrid are shown as coloured circles highlighted by dotted arrows. In the latter, only two of the mutated positions could be displayed in the model since the two N-terminal substitutions were not included in the expressed soluble portion of the crystallized A transferase [32] on which the template structure (molecular coordinate file 1LZ0 in the NCBI structure database) was based. The 3D surface model was created with the DeepView Swiss Pdb Viewer version 3.7, an interactive molecular graphics programme for viewing and analyzing protein structure (http://www.expasy.org/spdbv and [33]) using the molecular surface mode and the 3D rendering display option. No effort was made to calculate the possible structural alterations caused by the amino acid changes. The original sequence expressed for the crystallographic study [32] was used for the analysis. Residue positions 63 and 74 are highlighted in red and green, respectively, to show their relative locations in the model. In addition, the DXD motif (present in almost all glycosyltransferases, here as DVD capable of binding the UDP part of the nucleotide-sugar substrate) is shown in blue in order to highlight the catalytic cleft. The different regions of the glycosyltransferase are shown according to Paulson and Colley [34] and the transmembrane domain (amino acids 17–37) is based on the hydrophobicity plot and amino acid composition originally reported by Yamamoto et al. [35]. The black X represents the approximate location of a proteolytic cleavage site for generation of the soluble glycosyltransferase found in body fluids.

Introns 2, 3 and 4

Since the new allele has a novel mutation in exon 2 (46G>A) and an unexpected lack of the O^{1v}-specific mutation in exon 4 (189C>T), other introns in the A^2-like allele of the five $A_{weak}B$ samples were also sequenced and all gave identical results. Samples from the same 13 Caucasian and three Black individuals examined for intron 5 above were also analysed (Table 1). Intron 1 was not examined due to its size (approx. 13,000 bp). We found 7, 12 and 17 polymorphic sites (point mutations, deletions and insertions) in introns 2, 3 and 4, respectively, in the common alleles (Table 1).

The intron 2–4 sequences of the novel allele were identical to the O^{1v} alleles sequenced except for three adjacent mutations in intron 3. The first, 205T>C, caused a reversion to the consensus from O^{1v}-specific, whereas 354G>A and 399G>A were mutations not previously encountered in any allele. Intron 3 in the O^{1v} alleles from the control

Table 1: Polymorphic nucleotide positions in introns 2–5 of the *ABO* gene.

	Intron 2 (724 bp) nt. position						
Allele	126	209	362	369	396	437	539
dbSNP rs			2073828	2073827	6872809	687621	2073826
*A*1	A	C	C	C	T	C	C
*A*2	A	C	C	C	T	C	C
B	A	C	C	C	T	C	C
*O*1	A	C	***T***	***G***	***C***	***T***	***A***
*O*2	A***A***	C/***T******	***T***	***G***	***C***	***T***	***A***
*O*1v	A	C	C	C	***C***	***T***	C
new hybrid	A	C	C	C	***C***	***T***	C

	Intron 3 (1451 bp) nt. position													
Allele	205	354	399	479	525	575	969	1063	1095	1096	1170	1244	1278	1357
dbSNP rs	493211		8176701	579622	579483	8176702	575259	574347			8176704	613423	551322	624601
*A*1	C	G	G	C	A	C	T	A	C	A	C	C	T	C
*A*2	C	G	G	C	A	C	T	A	C	A	***T***	C	T	C
B	C	G	G	C	A	C	T	A	C	A	C	C	T	C
*O*1	C	G	G	C	A	***T***	T	A	C	A	C	C	T	C
*O*2	C	G	G	C	A	C	T	A	***del***	***del***	C	C	T	C
*O*1v	***T***	G	G	***T***	***T***	C	***C***	***G***	C	A	C	***T***	***C***	***T***
new hybrid	C	***A***	***A***	***T***	***T***	C	***C***	***G***	C	A	C	***T***	***C***	***T***

	Intron 4 (1686 bp) nt. position																
Allele	28	73	102	114	163	215	216	346	375	738	1092	1176	1467	1496	1511	1623	1678
dbSNP rs	7848145	6151203	8176707	625593	8176708	547495	626035	626792		638756		517414	514708	641943	641959	8176712	4962040
*A*1	G	C	C	C	G	A	A	T	C	T	C	C	G	T	T	C	T
*A*2	G	C	C	C	G	A	A	T	C	T	C	C	G	T	T	C	T
B	G	C	C	C	G	A	A	T	C	T	C	C	G	T	T	C	T
*O*1	G	C	C/***A***	C	G	A	A	T	C	T	C	C	G	T	T	C	***C***
*O*2	G	C	C	C	G	A	A	T	***T***	T	***T***	C	***A***	***G***	***G***	C	T
*O*1v	***C***	**	C	***T***	***A***	***G***	***C***	***G***	C	***G***	C	***T***	***A***	***G***	***G***	C***C***	T
new hybrid	***C***	**	C	***T***	***A***	***G***	***C***	***G***	C	***G***	C	***T***	***A***	***G***	***G***	C***C***	T

	Intron 5 (554 bp) nt. position							
Allele	91	103	266	306	336	450	527	530
dbSNP rs		8176713		8176714	8176715	8176717	8176718	
*A*1	T	C	G	C	G	C	G	A
*A*2	T	C	G	C	G	C	G	A
B	T	C	G	C	G	C	G	A/***G***
*O*1	T	C	G	C	***A***	C	G	A
*O*2	***C***	C***CCC***	***A***	C	G	C	G	A
*O*1v	T	C***CCC***	G	***T***	G	***A***	***A***	A
new hybrid	T	C***CCC***	G	C	G	C	G	A

* C/T etc. indicates dimorphism at this position in the same allele from different individuals. ** Insertion of GTGTGGACAGAAG between nt. 72C and 73C of the consensus (*A*1 [A101]) sequence in intron 4. Deviations from the consensus sequence are highlighted (bold italics).

donors of African origin was sequenced and in each case was identical to Caucasian *O*1v alleles.

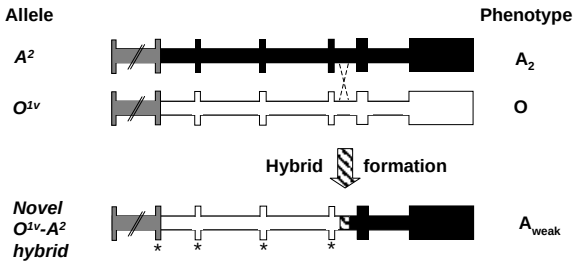

Figure 2
Representation of the exons and introns in A^2 and O^{1v} alleles and the new hybrid allele. Exons are represented as thick boxes and introns as thin boxes. They are drawn approximately to scale except for the approx 13 kbp long intron 1. Filled regions indicate an A^2 sequence and unfilled regions an O^{1v} sequence. The hatched region in the hybrid allele indicates the cross-over region. The grey area is presumed to have an O^{1v} allele sequence, but lack of known mutations in exons 1 and 2 and the unsequenced intron 1 does not allow proof of this. The asterisks indicate the localisation of the four missense mutations compared to the common A^2 allele.

Intron 6

Allele-specific variations in intron 6 have already been described [8,9]. Four new mutations were found in the new hybrid allele when compared to the *A* consensus sequence (277A>G, 286C>T, 911G>T and 952A>G). However, intron 6 in three normal A^2 alleles from Black Africans with the common A^2 phenotype also had these mutations.

Discussion

Detailed analysis of the alleles at the blood group ABO locus is shedding light on the effect of polymorphism in different regions of the translated products, the blood group A and B glycosyltransferases, and ultimately the clinically important ABO phenotype of red cells. Factors with the potential to influence the glycosyltransferase activities include base insertions, deletions and substitutions mainly in exons 6 and 7 (for review see [5,25]), hybrid alleles [17], splice-site mutations [10,11], variations in enhancer activity [12-14], promoter methylation [26], promoter mutations [27] and alternative promoter regions [28].

The new allele described here is unusual in several respects. The O^{1v}-characteristic mutations, 106C>T, 188G>A and 220C>T in exons 3, 4 and 5, respectively, in combination with the common A^2-specific sequence in exon 7 suggested that the allele is an O^{1v}-A^2 [O02-A201] hybrid having a crossing-over point after nt. 220 in exon 5 and before nt. 261 in exon 6. A deletion at nt. 261 is the most common inactivating event creating *O* alleles and is present in both O^1 and O^{1v} alleles and hybrid variants of these as shown in Figure 3. This mutation is absent from the five hybrid alleles described here and hence the cross-over should occur upstream of this position. In an attempt to determine the cross-over point more precisely and confirm the identity of the two contributing alleles we sequenced the intervening intron (intron 5) in the hope that allele-specific mutations were present in this intron, by analogy with our previous findings in intron 6 [9]. As Table 1 shows, eight of the 554 nucleotides in the intron were polymorphic. In this intron the A^1, A^2, *B* and O^1 alleles are very similar, differing only at nt. 336 in the O^1 allele and at nt. 529 in only one of the three *B* alleles tested. More pronounced differences were observed in the O^2 alleles that differed at each of the first three polymorphic positions (the complex variations in the O^2 allele will be presented elsewhere), and in the O^{1v} alleles, that differed at four of these sites. This latter information allowed us to determine the crossing-over point to occur in intron 5 between nt. 103 and 306, and that the allele is indeed an O^{1v}-A^2 hybrid (Figure 2).

This new allele showed some additional interesting characteristics. The mutation, nt. 46G>A (Ala16Thr) in exon 2, has not been described in any other context, nor could we detect this mutation when we analysed 260 alleles from individuals of diverse ethnic background (of whom about 40 were African). Its occurrence in all five index samples, albeit from individuals of Black African descent, collected from diverse parts of the world is surprising.

Few mutations have been found in exon 4. Both 188G>A and 189C>T have hitherto been exclusive characteristics of the O^{1v} allele; 190 G>A has only been found in some Brazilian Black O^1 alleles [29]; and 203G>C occurred in some Scandinavian A^{weak} alleles [16]. The finding of 188G>A without 189C>T in the new allele was also unexpected. When we screened more than 200 alleles at these positions we found one otherwise normal O^{1v} allele (African) having 188G>A but lacking 189C>T and one otherwise normal O^1 allele (Jordanian) having 188G>A. Although infrequent, additional genetic diversity obviously exists in individuals with common ABO phenotypes.

Several hybrid alleles at the *ABO* locus have been reported previously (reviewed in [17] and summarised in Figure 3). About half of these alleles are *O* alleles containing the common inactivating deletion at nt. 261 contributed by

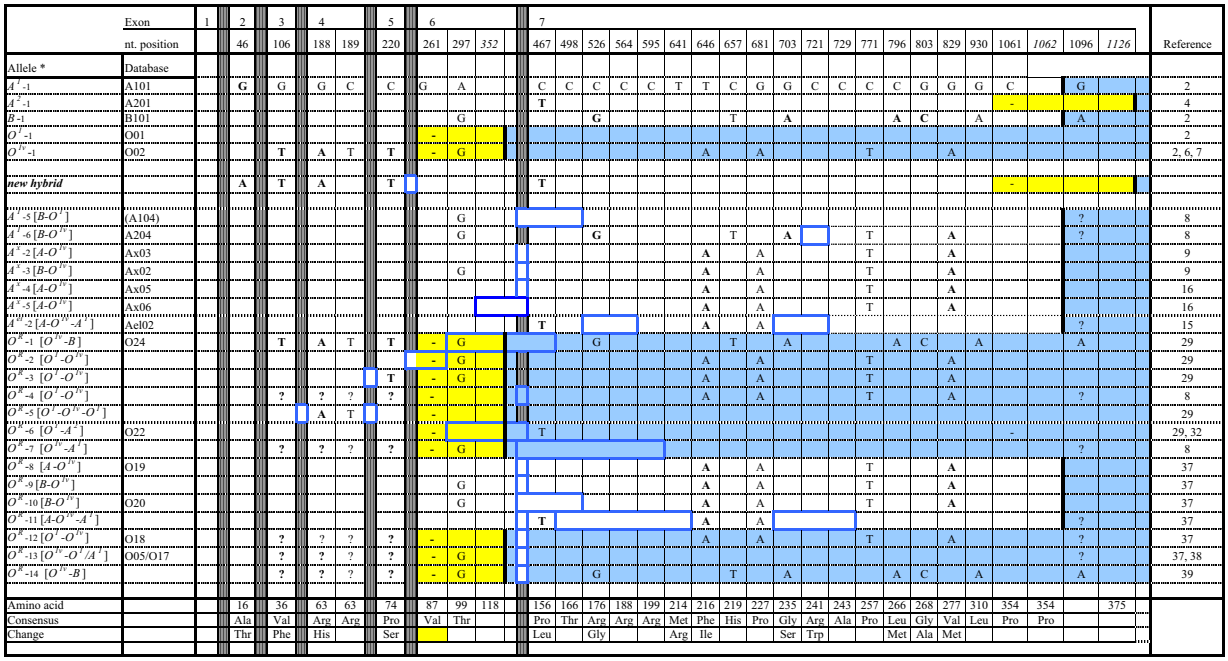

Figure 3
Known hybrid alleles at the *ABO* locus. The common alleles are also shown for comparison. Only changes from the consensus (*A*^{1}-*1*) sequence are shown. Mutations causing amino acid changes are shown in bold face. The yellow areas indicate a reading frame shift. The blue areas indicate untranslated regions. The introns are represented by the thick, dark vertical bars. The blue rectangles indicate the region of the allele where a crossing over event occurred. ?, indicates that the nucleotide at this position was not described.

the O^1 or O^{1v} part of the hybrid. The others described so far are *A* alleles of varying activity, containing initial *A* or *B* sequences followed by O^1 or O^{1v} sequences, where the critical nt. 261 is supplied by the *A* or *B* allele and is thus not deleted. The decreased A antigen expression is attributed to mutations in exon 7 (which encodes the enzyme's active site) contributed by the O^{1v} allele. This novel A^{weak} allele described here is the first example of a "reverse" hybrid, where the weakening of enzyme activity from A_2 to A_{weak} may be due to the concerted effect of the mutations in exon 2 at 46G>A (Ala16Thr), exon 3 at 106G>T (Val36Phe), exon 4 at 188G>A (Arg63His) and exon 5 at 220C>T (Pro74Ser). However, theoretically the A_{weak} phenotype may depend on only one or a combination of some of these four mutations.

Based on the linear structure, the former two substitutions presumably occur in or near the predicted transmembrane domain whilst the latter two localise to the postulated stem region of the Golgi-membrane-anchored enzyme. None of them is thought to occur in the proximity of the active site, which would be expected to have the normal A_2 enzyme structure based on the sequence of exon 7. In order to visualize this prediction, we created a three-dimensional computer model of the A transferase for mapping of two of the mutated amino acid positions onto the crystal-structure-based molecular surface. The result appears to confirm that residues 63 and 74 are localised far from the enzymatically active site in a stalk-like part of the protein (Figure 1). It should be emphasized that we did not attempt to determine the effect that these mutations would have on enzymatic function or fine structure of the gene product. Expression studies beyond the scope of this study or the future finding of other samples carrying other combinations of the above mutations may be able to address this issue.

The mutations found early on in intron 3 of this new hybrid allele (the consensus 205C rather than the O^{1v}-specific 205C>T, as well as the unique 354G>A and 399G>A)

appear to be a specific characteristic of this new allele since intron 3 of the three control O^{1v} alleles from Black African individuals did not differ from Caucasian samples. On the other hand, the four new mutations found in intron 6 of the new hybrid allele were also found in the three Black African A^2 alleles, which led us to conclude that the alterations observed in intron 6 may simply reflect a common ancestral A^2 allele of African evolutionary lineage.

The A and B glycosyltransferases compete for the same acceptor glycoconjugates. However, the weaker A activity due to this new allele was also observed serologically, although to a lesser degree, in the mother of one of these $A_{weak}B$ individuals who had inherited the hybrid allele in combination with an *O* allele and hence lacked a *B* gene. This may indeed be the reason why all five index samples studied here were $A_{weak}B$. Obviously, the weakening effect of this hybrid is relatively mild so that the A_{weak} phenotype, produced when the A hybrid glycosyltransferase is allowed to convert the available acceptor glycoconjugates to A without any competition from a B transferase, will sometimes escape detection in routine laboratories, especially if automated blood grouping equipment is used.

Conclusions

A new hybrid, A^{weak} allele with O^{1v} and A^2 characteristics with a crossing over point in intron 5 has been found at the blood group ABO locus in five individuals of diverse Black African backgrounds. To our knowledge this is the only defined ABO subgroup allele so far associated with an African ethnic origin.

Sequencing of all seven exons of these A^{weak} alleles showed two major exons (exons 6 and 7) were identical to the most frequent A^2 allele, although the intervening intron 6 had mutations only found so far in Black Africans. Exons 1–5 had sequences consistent with the O^{1v} allele, except for two novel changes, the most important one of which (46G>A) results in an amino acid substitution in the putative trans-membrane region of the translated protein (glycosyltransferase). The other novel change (lack of 189C>T) led further to the identification of two new *O* alleles.

This new hybrid allele shows how mutations in early exons, far from the enzyme product's active site, can affect expression of the blood group A antigen on the erythrocyte surface.

Methods

Blood samples and blood group serology

Blood samples from individuals living in Portugal, Sweden, Switzerland and the USA were referred to our laboratory for genomic analysis due to unclear phenotyping. Blood samples from first-degree relatives were only obtained in one case.

One hundred and thirty blood samples available at the Blood Centre, University Hospital, Lund from blood donors and other apparently healthy individuals with mixed phenotypes and mixed ethnicity (Africans, Europeans, Jordanians) were used for screening purposes. Their ABO group was determined according to current practice [30].

Routine ABO *genotyping*

All oligonucleotide primers used were synthesized by DNA Technology ApS (Aarhus, Denmark). DNA was prepared in Lund using a simple salting-out method [31].

The initial *ABO* genotyping comprised duplex PCR-RFLP and PCR-ASP analysis of exons 6 and 7 across intron 6 and subsequently DNA sequencing of the *ABO* gene and its regulatory elements was performed [7,16,19,20].

Selected homozygous and heterozygous DNA (samples homozygous for O^2 and *B* allele were not available) from blood donors and other apparently healthy individuals at the reference laboratory in Lund was used for identification of introns 2 to 5 from the common alleles A^1, A^2, *B*, O^1, O^{1v} and O^2.

PCR amplification of the ABO *gene for DNA sequencing*

Primers used to amplify DNA fragments and for allele-specific direct sequencing of the seven exons and intron 6 are described elsewhere [16].

Alternatively, polymerase chain reaction (PCR) was carried out using Expand High Fidelity PCR system (Roche Molecular Systems, Pleasanton, CA, USA) to amplify different intron fragments and for screening for mutations 188G>A and 46A>G. Amplifications were performed with primer pairs as shown in Table 2. An internal positive control primer pair was used in each PCR reaction. Amplification was performed in a reaction volume of 20 µl with 0.5 µmol/L of each primer, 2 nmol of each dNTP, and 100 ng of genomic DNA. As thermostable enzyme we used 0,5U from the Expand High Fidelity PCR System in the supplied buffer 2 with a final Mg^{2+} concentration of 1.5 mM according to the manufacturer (Roche Molecular Systems). After an initial denaturation step at 95°C for 2 min followed 10 cycles of denaturation (94°C for 20 s), annealing (65°C for 45 s) and extension (72°C for 1.5 min), then 25 cycles at 94°C for 20 s, 61°C for 30 s and 72°C for 1 min and a final extension for 5 min.

PCR products were excised from 3% agarose gels (Seakem, FMC Bioproducts, Rockland, ME, USA) stained with ethidium bromide (0.56 mg/l gel, Sigma Chemicals, St.

Table 2: Oligonucleotide primers used for PCR amplification, sequencing and screening in this study.

Primer name	F/R*	Nucleotide sequence (5' → 3')	Position	Function
mo-21s	F[1–3]	GGTGAGAGAAGGAGGGTGAG	intron 1	amplification-sequencing of intron 2 for all alleles
mo-31r	R[1,4]	CCAGCACCCCGGCCAGCA	intron 3	amplification-sequencing of intron 2 /screening of 46G>A
ABO-46A-F	F[4]	CCAGGAAAACCAAAATGCCACA	exon 2	screening of 46G>A
ABO-106G-R	R[2]	TAGACTTCTGGGGCTTAGGAC	exon 3	amplification of *O*[2] allele in intron 2
ABO-106T-R	R[3]	AGACTTCTGGGGCTTAGGAAC	exon 3	amplification of new hybrid allele in intron 2
ABO-inII-123F	F	GTTATCAGGGTCCTAAGGACAG	intron 2	sequencing of intron 2
ABO-inII-660R	R	CTGCCTGTTGGTCCCTTCCTC	intron 2	sequencing of intron 2
ABO-133s	F[5–8]	GCCCCAGAAGTCTAATGCCAG	exon 3	amplification-sequencing of introns 3 and 4
ABO-202 cons-R	R[5]	GGGAGGCACTGACATTATACC	intron 4/exon 4	amplification-sequencing of intron 3 (except *O*[2], *B* and hybrid alleles**)
ABO-188A-R	R[6,9]	ATACCTTGGCAACGAGACGT	intron 4/exon 4	amplification-sequencing of hybrid allele in intron 3 and screening
ABO-220 T-R	R[7,10]	CCACGGTGTCAGCACCTTTGA	exon 5	amplification-sequencing of *O*[2] allele in introns 3 and 4
ABO-220 C-R	R[8,11]	CACGGTGTCAGCACCTTTGG	exon 5	amplification-sequencing of *B* allele in introns 3 and 4
ABO-inIII-F	F[9]	GCTGGCCGTTACAGGGTCTG	intron 3	screening of 188G>A mutation in exon 4
ABO-inIII-425F	F	GCTGCCCTCATCTCTGTGACA	intron 3	sequencing of intron 3
ABO-inIII-672F	F	GTGCTATGGCCTCTGTTGGG	intron 3	sequencing of intron 3
ABO-inIII 916F	F	CTCTGGCAGTTGATGCTGGC	intron 3	sequencing of intron 3
mo-41s	F[10,11]	TAAATCCTGCTCCTAGACTAAAC	intron 3	amplification-sequencing of intron 4
ABO-inIV 170F	F	GACTTGGCCCTCGTCCTGCA	intron 4	sequencing of intron 4
ABO-inIV-429R	R	GACTAGCCTGGCCAACATGG	intron 4	sequencing of intron 4
ABO-inIV-852F	F	TAGCAACTCCATTTTCCCTCCC	intron 4	sequencing of intron 4
ABO-inIV 877R	R	GTTGGAGTAGCAGACTCATAACA	intron 4	sequencing of intron 4
ABO-inIV 1122F	F	CTCCTGAGCCTCTACAATCCT	intron 4	sequencing of intron 4
ABO-inIV 1411F	F	CTCTACGTCCCTCCCAGCCT	intron 4	sequencing of intron 4
ABO-229F	F[12,13]	CTACCCCCAGCCAAAGGTGC	exon 5	amplification-sequencing of all alleles in intron 5
ABO-297A-R	R[12]	GTTGAGGATGTCGATGTTGAAT	exon 6	amplification-sequencing of *A*[1], *A*[2], *O*[1] and hybrid allele in intron 5
ABO-297G-R	R[13]	GTTGAGGATGTCGATGTTGAAC	exon 6	amplification-sequencing of *B*, *O*[1v] and *O*[2] alleles in intron 5
mo-101s	F[4,9]	CCGTCCGCCTGCCTTGCAG	intron 6	internal control primer pair in screening
EPB-79R	R[4,9]	TACTTGTTCAGGTGGCTCTCGTC	exon 7	internal control primer pair in screening

* Forward/reverse primer. ** *O*[2], *B* and hybrid alleles were examined in heterozygous samples. The numeral superscripts denote primer combinations. Primers with the same number were used in the same primer mixes for amplification of *ABO* gene fragments.

Louis, MO, USA) following high-voltage electrophoresis and purified using the Qiaquick gel extraction kit (Qiagen GmbH, Hilden, Germany).

The Big Dye Terminator Cycle Sequencing kit (Applied Biosystems, Foster City, CA, USA) and an ABI PRISM 310 Genetic Analyser (Applied Biosystems) were used for direct DNA sequencing with capillary electrophoresis and automated fluorescence-based detection according to the manufacturer's instructions. Sequence analysis was performed with SeqEd software 1.03 (Applied Biosystems).

Authors' Contributions

Authors BHM and ÅH performed the molecular biology experimentation including PCR and DNA sequencing. Author MJR collected blood samples and performed serological blood group studies. Authors BHM, AC and MLO conceived and coordinated the study and drafted the manuscript. All authors read and approved the submitted manuscript.

Note

[1]Nomenclature used in the Blood Group Antigen Gene Mutation Database http://www.bioc.aecom.yu.edu/bgmut/abo.htm.

Acknowledgements

Dr. P. Marinho, Clinica Médica Dr J. Chaves, and Dr. T. Chabert, Centro Regional de Sangue de Lisboa, Portugal, Dr. Viveka Stiller, Division of Clinical Immunology and Transfusion Medicine, Karolinska Hospital (present affiliation: Huddinge University Hospital), Stockholm, Sweden, Ms. Nicole Eicher, Blood Transfusion Service SRC, Bern, Switzerland and Ms. Marilyn Moulds, Immucor/Gamma Reference Laboratories, Atlanta, GA/Houston, TX, USA are thanked for referral of blood samples for genetic analysis after performing the initial serological investigation. We are also grateful to Dr. Geoff Daniels and Dr. Nidal Irshaid for providing samples of African and Jordanian blood donors, respectively. This work was supported financially by the Swedish Research Council (project no. K2002-71X-14251-01A), the Medical Faculty at Lund University/ALF, the Lund University Hospital Donation Funds, the Claes Högman SAGMAN-stipendium.

Part of this work was presnted at the 8th Regional European Congress of the International Society of Blood Transfusion held in Istanbul, Turkey on July, 5-9 2003 and published in absract form [1].

References

1. Hosseini-Maaf B, Hellberg Å, Eicher N, Rodrigues MJ, Chester MA, Olsson ML: **A novel O ^{1v}-*A*^{2} hybrid allele at the ABO blood group locus is associated with the A_wB Phenotype [abstract].** *Proceedings (Abstract Book) of the 8th Congress of the International Society of Blood Transfusion* 2003, **120:**P341.
2. Yamamoto F, Clausen H, White T, Marken J, Hakomori S: **Molecular genetic basis of the histo-blood group ABO system.** *Nature* 1990, **345:**229-233.
3. Watkins WM: **Biochemistry and genetics of the ABO, Lewis and P blood group systems.** In *Advances in Human Genetics* Edited by: *Harris H, Hirschhorn K. New York: Plenum Press*; 1980:1-136.
4. Yamamoto F, McNeill PD, Hakomori S: **Human histo-blood group A^{2} transferase coded by A^{2} allele, one of the A subtypes, is characterized by a single base deletion in the coding sequence, which results in an additional domain at the carboxyl terminal.** *Biochem Biophys Res Commun* 1992, **187:**366-374.
5. Chester MA, Olsson ML: **The ABO blood group gene – A locus of considerable genetic diversity.** *Transfus Med Rev* 2001, **11:**295-313.
6. Yamamoto F, McNeill PD, Yamamoto M, Hakomori S, Harris T: **Molecular genetic analysis of the ABO blood group system: 3. A^{x} and B$^{(A)}$ alleles.** *Vox Sang* 1993, **64:**171-174.
7. Olsson ML, Chester MA: **Frequent occurrence of a variant O^{1} gene at the blood group ABO locus.** *Vox Sang* 1996, **70:**26-30.
8. Suzuki K, Iwata M, Tsuji H, Takagi T, Tamura A, Ishimoto G *et al.*: **A de novo recombination in the ABO blood group gene and evidence for the occurrence of recombination products.** *Hum Genet* 1997, **99:**454-461.
9. Olsson ML, Chester MA: **Heterogeneity of the blood group *A*^{x} allele: genetic recombination of common alleles can result in the A_x phenotype.** *Transfus Med* 1998, **8:**231-238.
10. Olsson ML, Irshaid NM, Kuosmanen M, Pirkola A, Chester MA: **A splice-site mutation defines the *Afinn* allele at the blood group ABO locus [abstract].** *Transfusion* 2000, **40(10S):**90S.
11. Yu LC, Twu YC, Chou ML, Chang CY, Wu CY, Lin M: **Molecular genetic analysis for the B(3) allele.** *Blood* 2002, **100:**1490-1492.
12. 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:**25890-25898.
13. 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:**219-226.
14. 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:**459-466.
15. Ogasawara K, Yabe R, Uchikawa M, Saitou N, Bannai M, Nakata K *et al.*: **Molecular genetic analysis of variant phenotypes of the ABO blood group system.** *Blood* 1996, **88:**2732-2737.
16. Olsson ML, Irshaid NM, Hosseini-Maaf B, Hellberg A, Moulds MK, Sareneva H *et al.*: **Genomic analysis of clinical samples with serologic ABO blood grouping discrepancies: identification of 15 novel A and B subgroup alleles.** *Blood* 2001, **98:**1585-1593.
17. Olsson ML, Chester MA: **Polymorphism and recombination events at the ABO locus: a major challenge for genomic ABO blood grouping strategies.** *Transfus Med* 2001, **11:**295-313.
18. Olsson ML, Chester MA: **Polymorphisms at the ABO locus in subgroup A individuals.** *Transfusion* 1996, **36:**309-313.
19. Olsson ML, Chester MA: **A rapid and simple ABO genotype screening method using a novel B/O^{2} versus A/O^{1} discriminating nucleotide substitution at the ABO locus.** *Vox Sang* 1995, **69:**242-247.
20. Olsson ML, Hosseini-Maaf B, Hellberg Å, Chester MA: **Allele-specific primer PCR across exon 6 resolves potential genotyping errors caused by recombinant hybrid alleles at the *ABO* locus [abstract].** *Transfusion* 1998, **38(10S):**3S.
21. 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:**19257-19262.
22. Olsson ML, Hansson C, Avent ND, Akesson IE, Green CA, Daniels GL: **A clinically applicable method for determining the three major alleles at the Duffy (FY) blood group locus using polymerase chain reaction with allele-specific primers.** *Transfusion* 1998, **38:**168-173.
23. Singleton BK, Green CA, Avent ND, Martin PG, Smart E, Daka A *et al.*: **The presence of an RHD pseudogene containing a 37 base pair duplication and a nonsense mutation in africans with the Rh D-negative blood group phenotype.** *Blood* 2000, **95:**12-18.
24. Tournamille C, Colin Y, Cartron JP, Le Van KC: **Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals.** *Nat Genet* 1995, **10:**224-228.
25. Yip SP: **Sequence variation at the human ABO locus.** *Ann Hum Genet* 2002, **66:**1-27.
26. Kominato Y, Hata Y, Takizawa H, Tsuchiya T, Tsukada J, Yamamoto F: **Expression of human histo-blood group ABO genes is dependent upon DNA methylation of the promoter region.** *J Biol Chem* 1999, **274:**37240-37250.
27. 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.
28. Kominato Y, Hata Y, Takizawa H, Matsumoto K, Yasui K, Tsukada J *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:**37936-37948.
29. Olsson ML, Guerreiro JF, Zago MA, Chester MA: **Molecular analysis of the O alleles at the blood group ABO locus in populations of different ethnic origin reveals novel crossing-over events and point mutations.** *Biochem Biophys Res Commun* 1997, **234:**779-782.
30. Vengelen-Tyler V: *Technical Manual* 13th edition. *Bethesda, MD, USA: American Association of Blood Banks*; 1999.
31. Miller SA, Dykes DD, Polesky HF: **A simple salting out procedure for extracting DNA from human nucleated cells.** *Nucleic Acids Res* 1988, **16:**1215.
32. Patenaude SI, Seto NO, Borisova SN, Szpacenko A, Marcus SL, Palcic MM *et al.*: **The structural basis for specificity in human ABO(H) blood group biosynthesis.** *Nat Struct Biol* 2002, **9:**685-690.
33. Guex N, Peitsch MC: **SWISS-MODEL and the Swiss-Pdb-Viewer: an environment for comparative protein modeling.** *Electrophoresis* 1997, **18:**2714-2723.
34. Paulson JC, Colley KJ: **Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation.** *J Biol Chem* 1989, **264:**17615-17618.
35. Yamamoto F, Marken J, Tsuji T, White T, Clausen H, Hakomori S: **Cloning and characterization of DNA complementary to human UDP-GalNAc:Fuc-alpha1-2Gal alpha1-3GalNAc transferase (histo-blood group A transferase) mRNA.** *J Biol Chem* 1990, **265:**1146-1151.
36. Gassner C, Schmarda A, Nussbaumer W, Schonitzer D: **ABO glycosyltransferase genotyping by polymerase chain reaction using sequence-specific primers.** *Blood* 1996, **88:**1852-1856.
37. Ogasawara K, Yabe R, Uchikawa M, Nakata K, Watanabe J, Takahashi Y *et al.*: **Recombination and gene conversion-like events may**

contribute to ABO gene diversity causing various phenotypes. *Immunogenetics* 2001, **53:**190-199.
38. Ogasawara K, Bannai M, Saitou N, Yabe R, Nakata K, Takenaka M *et al.*: **Extensive polymorphism of ABO blood group gene: three major lineages of the alleles for the common ABO phenotypes.** *Hum Genet* 1996, **97:**777-783.
39. Bugert P, Rutten L, Goerg S, Kluter H: **Characterization of a novel O^1 variant allele at the ABO blood group locus.** *Tissue Antigens* 2001, **58:**422-424.