Vox Sanguinis (2006) **90**, 198–203

DOI: 10.1111/j.1423-0410.2006.00746.x

ORIGINAL PAPER

Sequence variation in the 5′ untranslated region of the human *A4GALT* gene is associated with, but does not define, the P1 blood-group polymorphism

L. Tilley, C. Green & G. Daniels

Bristol Institute for Transfusion Sciences, National Blood Service, Bristol, UK

Vox Sanguinis

Background and Objective The gene responsible for the P1 polymorphism of the P blood-group system remains unidentified, although the *A4GALT* gene, whose product is responsible for the production of P^k, has been implicated. No coding differences in *A4GALT* account for the P1 polymorphism, but homozygosity for two polymorphisms (–551_–550insC and –160A>G) in the 5′ untranslated region of the gene has been reported to be unique to Japanese P1– individuals. This study aimed to confirm this correlation in a larger number of British individuals.

Materials and Methods Serologically confirmed P1+ (n = 35) and P1– (n = 15) individuals were genotyped for polymorphisms in the 5′ untranslated region of *A4GALT*.

Results In addition to those previously reported, a further polymorphism, –164C>T, was identified. All P1– individuals were homozygous for –551_–550insC and –160G as compared with 10 of 35 (29%) P1+ individuals (P = 0·000003, two-tailed Fisher's exact test). Allele frequencies for all polymorphisms and estimated haplotype frequencies across the region differed significantly between P1+ and P1– groups.

Conclusions Homozygosity for the *A4GALT* –551_–550insC and –160G allele is significantly associated with, but not restricted to, the P1– phenotype. No single *A4GALT* genotype or haplotype was unique to P1– individuals. Thus, *A4GALT* cannot be unequivocally confirmed as the gene responsible for the P1 phenotype.

Received: 4 October 2005, revised 7 December 2005, accepted 9 December 2005, published online 16 February 2006

Key words: *A4GALT* gene, P1 blood group, polymorphism.

Introduction

The P blood groups – P1, P^k and P – on red cells are glycosphingolipids of the paragloboside (P1) and globoside (P^K and P) series [1]. The genetic relationship of P1 to P^k and P is still not clear, and the gene responsible for the P1 polymorphism is the sole remaining gene of the 29 blood group systems still to be identified [2].

In most people, lactosylceramide (Gb2) is converted to P^k antigen (Gb3 or CD77) through the addition of a terminal galactosyl residue by an α1,4-galactosyltransferase, the product of the *A4GALT* gene (Fig. 1). However, P^k is barely detectable on red cells of common phenotypes because P^k is further converted to P (globoside or Gb4) by the *N*-acetylgalactosaminyltransferase product of the inappropriately named *B3GALT3* gene. Rare 'null' phenotypes can result from homozygosity for inactivating mutations in either gene. Inactivity of *A4GALT* or its product prevents production of P^k and, consequently, P, giving rise to the p phenotype [3–7]. Inactivity of *B3GALT3* or its product prevents conversion of P^k to P, resulting in the P^k phenotype, in which there is strong expression of P^k, but no P antigen [7,8].

P1 is polymorphic: about 80% of Caucasians and 30% of eastern Asians have the P1+ red cell phenotype [1]. The product of the gene responsible for P1 production is predicted to be an α1,4-galactosyltransferase, which converts paragloboside to galactosylparagloboside, P1 (Fig. 1). Red cells of the p phenotype (inactive *A4GALT*) are always P1 negative.

Correspondence: Louise Tilley, Bristol Institute for Transfusion Sciences, National Blood Service, Southmead Road, Bristol BS10 5ND, UK
E-mail: louise.tilley@nbs.nhs.uk
The authors were supported by a grant from the UK NHS R&D Directorate.

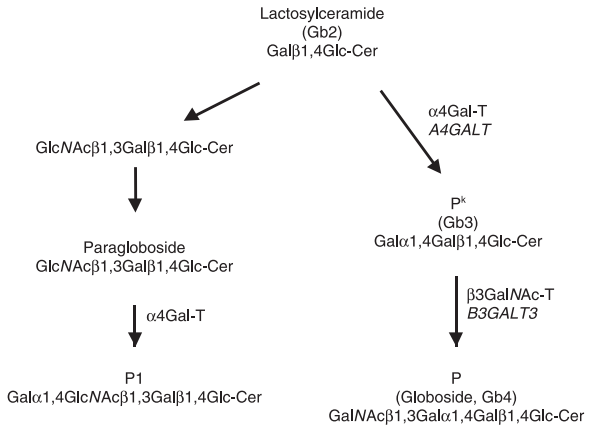

Fig. 1 Pathways involved in the biosynthesis of P^k, P and P1 antigens from lactosylceramide, showing some of the transferases involved and the genes that encode them. Cer, ceramide; Glc, glucose; Gal, galactose; Glc*N*Ac, *N*-acetylglucosamine; Gal*N*Ac, *N*-acetylgalactosamine; T, transferase.

Somatic cell hybridization studies and linkage analyses have located the gene responsible for the P1 polymorphism to chromosome 22q11.2-qter [9–11], and *A4GALT* is at 22q13.2 [3]. In the absence of any closely linked *A4GALT* homologue, it is feasible that *A4GALT* could encode an α1,4-galactosyltransferase that catalyses the biosynthesis of both P^k from lactosylceramide and P1 from paragloboside. However, no sequence differences in the coding region of the *A4GALT* gene account for the P1 polymorphism [3].

Iwamura *et al.* [12] showed that a mouse fibroblast cell line transfected with *A4GALT* cDNA expressed P1 in the cytoplasm, and that red cells serotyped as P1– had P1 antigen inside the cells, although the levels were lower than in P1+ cells. No P1 was detected in p cells. As P1– red cells had lower levels of *A4GALT* mRNA than P1+ cells, Iwamura *et al.* [12] analysed 5′-upstream promoter regions in *A4GALT* and found two mutations in the same allele: these mutations were either in the homozygous state in 10 P1– individuals, or in the heterozygous state (with the wild-type sequence) in seven P1+ individuals. The present study attempted to confirm this correlation in a larger number of P1+ and P1– individuals, to establish whether the polymorphisms reported by Iwamura *et al.* [12] in the 5′-upstream promoter regions could be used for P1 genotyping. We found an association between several polymorphisms in the *A4GALT* 5′ untranslated region and P1 phenotype, but concordance was incomplete.

Materials and methods

Whole-blood samples from 39 random blood donors and 11 known P1– donors were obtained from the processing, testing and issues department, Blood Services South-west, Bristol, UK.

Serological methods

Red cells were washed three times and a 4% suspension was prepared in phosphate-buffered saline (PBS).

Monoclonal anti-P1 (OSK17) (Gamma Biologicals, Houston, TX) and goat anti-P1 were used in agglutination tests. Standard tube techniques were used. Agglutination test results were read microscopically, on cold slides, after incubation for 1 h at 4 °C.

Molecular genetic analyses

Genomic DNA was isolated from whole blood using a commercial genomic DNA isolation kit (Promega, Madison, WI).

The 5′ untranslated region of the *A4GALT* gene was amplified by polymerase chain reaction (PCR), using two sets of primers and conditions as follows.

(1) Forward primer: 5′-GAGGTTCCATTTTCTCAGTGG-3′ (nucleotides –811 to –791) and reverse primer 5′-GCGGAATTCCAGTTATTTGCT-3′ (nucleotides –461 to –481). PCR consisted of initial denaturation at 94 °C for 5 min followed by 35 cycles of incubation at 94 °C for 30 s, annealing at 55 °C for 30 s and extension at 72 °C for 30 s, with a final extension at 72 °C for 7 min (method adapted from Iwamura *et al.*) [12].

(2) Forward primer: 5′-GGTGACAGCGTCCCTTCTCCA-3′ (nucleotides –381 to –361) and reverse primer 5′-GCCCGTCCCGACCTACCCT-3′ (nucleotides –118 to –136). PCR consisted of initial denaturation at 94 °C for 5 min followed by 35 cycles of a two-step protocol of incubation at 94 °C for 45 s followed by annealing/extension at 68 °C for 1 min, with a final extension at 72 °C for 7 min.

All PCR amplifications were carried out using 200 ng of genomic DNA in a 20-μl final volume of buffer (supplied by the polymerase manufacturer) containing 0·2 mM dNTPs, 2·5 mM $MgCl_2$, 0·5 μM each forward and reverse primer, and 0·25 U BioTaq polymerase (Bioline Ltd, London, UK).

PCR products were prepared for sequencing using QuickClean (Bioline Ltd) and subjected to direct sequencing using both forward and reverse PCR primers. An additional internal reverse primer [5′-TCCTGGTGCTTTTACTTCTT-3′ (nucleotides –687 to –668)] was used for sequencing products of PCR reaction 1 to avoid the insertion/deletion mutation in this fragment [12]. The numbering of nucleotides was adopted from Iwamura *et al.* [12], where +1 refers to the transcription initiation site.

Statistical methods

Allele and genotype frequencies were calculated and compared between P1+ and P1– groups using χ^2-tests or Fisher's exact tests (where cell numbers were insufficient to perform

χ^2-tests), performed using SISA statistical analysis [13]. Haplotype analysis was performed using the HAP program for haplotype reconstruction from genotype data [14].

Results

The 11 donors known to be P1– were confirmed as such by serological typing. In addition, a further four of the random donors were found to have the P1– phenotype following serological testing.

Amplification and sequencing of the 5′ untranslated region of *A4GALT* in 35 P1+ and 15 P1– individuals identified six polymorphic sites (–773C>G, –770C>T, –638A>C, –551_–550insC, –212C>G and –160A>G) as previously described [12]. In addition, a further polymorphism, –164C>T, was identified. The genotypes of P1+ and P1– individuals for each of these polymorphisms are shown in Table 1.

The –160G allele and –551_–550C insertion are in complete linkage disequilibrium. All P1– individuals studied were homozygous for the –551_–550C insertion and –160G alleles compared with 10 of 35 (29%) P1+ individuals (P = 0·000003, two-tailed Fisher's exact test). Although there is a highly significant difference in frequency of homozygosity for the –551_–550C insertion and –160G alleles between P1+ and P1– groups, in contrast to previous findings [12] homozygosity was not found to be unique to P1– individuals.

Allele and genotype frequencies for each polymorphism were calculated in P1+ and P1– sample groups (Table 2). Allele frequencies for the –773 polymorphism [χ^2 = 14·3, 1 degree of freedom (d.f.), P = 0·0001] and the –551_550/–160 polymorphisms (P = 0·000006, two-tailed Fisher's exact test) differ significantly between the P1+ and P1– sample groups. Allele frequencies for the –770 (χ^2 = 4·3, 1 d.f., P = 0·04), –638 (χ^2 = 4·0, 1 d.f., P = 0·04), –212 (χ^2 = 4·9, 1 d.f., P = 0·03) and –164 (χ^2 = 5·7, 1 d.f., P = 0·02) polymorphisms also differ between the P1+ and P1– groups, but with marginal significance levels. No correction for multiple testing has been applied as the polymorphisms are all located on the same region of DNA and thus linkage between the loci means that the observations at each locus are not independent. Applying standard Bonferroni correction for multiple testing would therefore be too stringent. However, even with this correction, the frequencies of the –773 and –551_–550/–160 polymorphisms would clearly still be highly significantly different between the P1+ and P1– groups.

Haplotype analysis, generating estimated haplotypes from genotypic data, predicted 11 haplotypes, five of which are only seen in single individuals and have been grouped together. Table 3 shows the estimated frequencies of the common haplotypes in P1+ and P1– sample groups. The CCA–CCA (possessing –773C, –770C, –638A, –212C, –164C and –160A alleles and lacking the –551_–550C insertion) haplotype is the most common in P1+ samples, with a frequency of 0·39, but is not observed in P1– samples (P = 0·00001, two-tailed Fisher's exact test). In contrast, the common haplotype in P1– samples, GTA+GCG, with a frequency of 0·27, is rare in P1+ samples, with a frequency of 0·06 (P = 0·006, two-tailed Fisher's exact test). However, although statistically significant differences are observed in estimated haplotype frequencies between P1+ and P1– sample groups, no single haplotype or diplotype is unique to the P1– individuals. For example, samples 34 and 45 have identical genotypes across all seven polymorphisms, yet sample 34 is P1+ and sample 45 is P1–. Therefore, although the polymorphisms in the 5′ untranslated region of the *A4GALT* gene are significantly associated with the P1 phenotype, they are not the sole determining factor.

Discussion

The P blood group system contains one antigen, P1 (003001), and two phenotypes, P1+ and P1– (often called P_2). The frequency of P1 is about 80% in Caucasians, but may be as low as 30% in Japanese [1]. P1– individuals occasionally produce anti-P1, but this is usually a weak immunoglobulin M (IgM) antibody that is only active at low temperatures. Only very rarely has anti-P1 been produced in response to stimulation by transfusion or pregnancy [1].

The gene governing the P1 polymorphism remains the only gene of the 29 blood group systems still to be identified. The chemical structure of P1 suggests that the P1 gene will encode an α1,4-galactosyltransferase, and the invariable absence of P1 from red cells of the p phenotype has led to speculation that the α1,4-galactosyltransferase product of *A4GALT*, which catalyses the biosynthesis of P^k antigen from lactosylceramide, might also catalyse synthesis of P1 from paraglobside (Fig. 1) [15,16]. This proposal was very strongly supported by the discovery by Iwamura *et al.* [12], of two polymorphisms upstream of the *A4GALT* translation initiation site, in which all 10 P1– individuals were homozygous for one allele and all seven P1+ individuals were heterozygous. Confirmation of this correlation in a larger sample would have permitted the P^k antigen (209002) to become the second antigen of the P system.

Our tests on 35 P1+ and 15 P1– English blood donors have failed to confirm an absolute correlation between the P1 phenotype and any of seven *A4GALT* polymorphic sites. Although all P1– individuals were homozygous for the –551_–550insC and –160G alleles, as found by Iwamura *et al.* [12], over 25% of the 35 P1+ individuals had the same genotype. Furthermore, only three of the 35 P1+ individuals were homozygous for the absence of the –551_–550C insertion and –160A, far fewer than would be expected if that represented the $P1^+$ allele. Consequently, these polymorphic sites in *A4GALT* do not define the P1 polymorphism and cannot

Table 1 Genotypes at seven polymorphic loci in the 5′ untranslated region of *A4GALT* in P1+ and P1− individuals

Sample	Phenotype	−773C>G	−770C>T	−638A>C	−551_−550insC[a]	−212C>G	−164C>T	−160A>G
1	P1+	CC	CT	AA	+ −	CG	CC	AG
2	P1+	CG	CT	AC	+ −	CC	CC	AG
3	P1+	CC	CC	AA	+ −	CC	CT	AG
4	P1+	CG	CT	AA	+ −	CG	CC	AG
5	P1+	CC	CC	AA	++	CC	CT	GG
6	P1+	CC	CC	AA	− −	CC	CC	AA
7	P1+	CC	TT	AA	++	GG	CC	GG
8	P1+	CC	CT	AA	+ −	CG	CC	AG
9	P1+	CC	CC	AA	+ −	CC	CT	AG
10	P1+	CC	CT	AA	+ −	CG	CC	AG
11	P1+	CC	CC	AA	+ −	CC	CT	AG
12	P1+	CG	CT	AA	++	CG	CC	GG
13	P1+	CG	TT	AC	++	CG	CT	GG
14	P1+	CG	CT	AC	+ −	CG	CC	AG
15	P1+	CC	CC	AA	+ −	CC	CT	AG
16	P1+	CC	CC	AA	+ −	CC	CT	AG
17	P1+	CC	CC	AA	− −	CC	CC	AA
18	P1+	CG	CT	AC	+ −	CG	CC	AG
19	P1+	CC	CC	AA	++	CC	CT	GG
20	P1+	CG	TT	AA	++	GG	CC	GG
21	P1+	CC	CC	AA	+ −	CC	CT	AG
22	P1+	CC	CC	AA	+ −	CC	CT	AG
24	P1+	CC	CC	AA	+ −	CC	CT	AG
26	P1+	CC	CC	AA	++	CC	CT	GG
27	P1+	CG	TT	AC	++	GG	CC	GG
28	P1+	CC	CT	AA	+ −	CG	CC	AG
29	P1+	CC	CT	AA	+ −	CG	CC	AG
30	P1+	CC	CC	AA	+ −	CC	CT	AG
31	P1+	CC	CC	AA	+ −	CC	CT	AG
32	P1+	CG	CT	AA	+ −	CG	CC	AG
33	P1+	CC	CC	AA	− −	CC	CC	AA
34	P1+	CC	CT	AA	++	CG	CT	GG
35	P1+	CC	CT	AA	+ −	CC	CT	AG
38	P1+	CC	CC	AA	+ −	CC	CC	AG
39	P1+	CG	CT	AC	++	CG	CC	GG
23	P1−	CC	CC	AA	++	CC	TT	GG
25	P1−	CC	CC	AA	++	CC	TT	GG
36	P1−	GG	TT	AC	++	GG	CC	GG
37	P1−	CG	CT	AC	++	CG	CT	GG
40	P1−	GG	TT	AA	++	GG	CC	GG
41	P1−	CG	CT	AC	++	CG	CT	GG
42	P1−	CG	CT	AC	++	GG	CC	GG
43	P1−	CG	CT	AA	++	CG	CT	GG
44	P1−	CG	CT	AC	++	CG	CT	GG
45	P1−	CC	CT	AA	++	CG	CT	GG
46	P1−	GG	TT	AA	++	GG	CC	GG
47	P1−	CG	CT	AA	++	CG	CT	GG
48	P1−	CG	CT	AC	++	CC	TT	GG
49	P1−	CG	CT	AA	++	CC	CT	GG
50	P1−	CG	CT	AC	++	CG	CT	GG

[a]Genotypes for the −551_−550insC polymorphism, shown as ++ for homozygosity for C insertion, − − for homozygosity for lack of insertion and + − for heterozygotes.

Table 2 Genotype and allele frequencies for seven polymorphic loci in the 5′ untranslated region of *A4GALT* in P1+ and P1– sample groups

	Genotypes	P1+ (*n* = 35) *n* (frequency)	P1– (*n* = 15) *n* (frequency)	Alleles	P1+ (*n* = 35) frequency	P1– (*n* = 15) frequency
–773C>G	CC	25 (0·71)	3 (0·20)	*C*	0·86	0·50
	CG	10 (0·29)	9 (0·60)	*G*	0·14	0·50
	GG	0	3 (0·20)			
–770C>T	CC	17 (0·49)	2 (0·13)	*C*	0·69	0·47
	CT	14 (0·40)	10 (0·67)	*T*	0·31	0·53
	TT	4 (0·11)	3 (0·20)			
–638A>C	AA	29 (0·83)	8 (0·53)	*A*	0·91	0·77
	AC	6 (0·17)	7 (0·47)	*C*	0·09	0·23
	CC	0	0			
–551_–550insC[a]	++	10 (0·29)	15 (1·0)	+	0·60	1·0
	+ –	22 (0·63)	0	–	0·40	0
	– –	3 (0·09)	0			
–212C>G	CC	19 (0·54)	4 (0·27)	*C*	0·73	0·50
	CG	13 (0·37)	7 (0·47)	*G*	0·27	0·50
	GG	3 (0·09)	4 (0·27)			
–164C>T	CC	19 (0·54)	4 (0·27)	*C*	0·77	0·53
	CT	16 (0·46)	8 (0·53)	*T*	0·23	0·47
	TT	0	3 (0·20)			
–160A>G	AA	3 (0·09)	0	*A*	0·40	0
	AG	22 (0·63)	0	*G*	0·60	1·0
	GG	10 (0·29)	15 (1·0)			

[a]Genotypes for the –551_–550insC polymorphism, shown as ++ for homozygosity for C insertion, – – for homozygosity for lack of insertion and + – for heterozygotes.

Table 3 Estimated haplotype frequencies for the –773C>G, –770C>T, –638A>C, –551_–550insC (+/–), –212C>G, –164C>T and –160A>G polymorphisms in the 5′ untranslated region of A4GALT in P1+ and P1– sample groups

Haplotype	P1+ *n* (frequency)	P1– *n* (frequency)
CCA–CCA	27 (0·39)	0
CCA+CTG	14 (0·20)	12 (0·40)
GTA+GCG	4 (0·06)	8 (0·27)
CTA+GCG	11 (0·16)	1 (0·03)
GTC+GCG	4 (0·06)	5 (0·17)
CCA+CCG	7 (0·10)	1 (0·03)
Others	3 (0·04)	3 (0·10)

Each haplotype covers seven polymorphic loci and is thus represented by seven characters representing the bases predicted to be present at each position. For example, the first letter of the haplotype represents the base present at the –773C>G locus, whilst the last letter represents the base present at the –160A>G locus. The –551_–550insC polymorphism at the fourth position of the haplotype is represented by + or –, according to the presence or absence, respectively, of the insertion in the haplotype.

be used for P1 genotyping. These data concur with those reported briefly by Hellberg [17] (see the Addendum).

The question remains as to whether *A4GALT* encodes the enzyme responsible for the biosynthesis of P1. The high level of correlation reported here and elsewhere [12,17], between *A4GALT* sequence polymorphisms and the P1 phenotype, strongly suggests that either *A4GALT* or a very closely linked gene governs the P1 polymorphism. Consequently, *A4GALT* remains the prime candidate in the apparent absence of any other glycosyltransferase gene in the region of chromosome 22 occupied by *A4GALT*. The genetics of oligosaccharide biosynthesis is often complex, with the picture being clouded by competition between different glycosyltransferases for limited quantities of the same donor or acceptor substrates. It is feasible that polymorphisms in the promoter region of *A4GALT* might result in reduced levels of gene product that are sufficient for P^k synthesis, but not for the synthesis of detectable levels of P1. If the –551_–550 insertion and –160 polymorphisms were directly responsible for variation in α1,4-galactosyltransferase activity, then other factors affecting substrate levels might account for those individuals who are homozygous for –551_–550insC and –160G, yet express P1 on their red cells. However, Iwamura *et al.* [12] failed to demonstrate any significant effect of the polymorphisms on promoter activity. Alternatively, the –551_–550C insertion and –160 polymorphisms could be in partial linkage disequilibrium with an undetected *A4GALT* polymorphism that correlates completely with the P1 phenotype. Clearly, more research is required before *A4GALT* can be confirmed as the gene responsible for the P1 phenotype.

Addendum

Whilst this paper was in preparation, this abstract data was expanded to a paper reporting similar data to that reported in this study.

Hellberg A, Chester, MA, Olsson, ML. Two previously proposed P1/P2-differentiating and nine novel polymorphisms at the A4GALT (Pk) locus do not correlate with the presence of the P1 blood group antigen. BMC Genet 2005; 6:49.

References

1 Daniels G: *Human Blood Groups*, 2nd edn. Oxford, Blackwell Science, 2002

2 Daniels GL, Fletcher A, Garratty G, Henry S, Jørgensen J, Judd WJ, Levene C, Lomas-Francis C, Moulds JJ, Moulds JM, Moulds M, Overbeeke M, Reid ME, Rouger P, Scott M, Sistonen P, Smart E, Tani Y, Wendel S, Zelinski T: Blood group terminology 2004. From the ISBT Committee on Terminology for Red Cell Surface Antigens. *Vox Sang* 2004; **87**:304–316

3 Steffensen R, Carlier K, Wiels J, Lavery SB, Stroud M, Cedergren B, Nilsson Sojka B, Bennet EP, Jersild C, Clausen H: Cloning and expression of the histo-blood group P^k UDP-galactose: Galβ1-4Glcβ1-Cer α1,4-Galactosyltransferase. Molecular basis of the p phenotype. *J Biol Chem* 2000; **275**:16723–16729

4 Furukawa K, Iwamura K, Uchikawa M, Sojka BN, Wiels J, Okajima T, Urano T, Furukawa K: Molecular basis for the p phenotype. Identification of distinct and multiple mutations in the α1,4-Galactosyltransferase gene in Swedish and Japanese individuals. *J Biol Chem* 2000; **275**:37752–37756

5 Koda Y, Soejima M, Sato H, Maeda Y, Kimura H: Three-base deletion and one-base insertion of the α(1,4) galactosyltransferase gene responsible for the p phenotype. *Transfusion* 2002; **42**:48–51

6 Hellberg Å, Steffensen R, Yahalom V, Sojka BN, Heier HE, Levene C, Poole J, Olsson ML: Additional molecular bases of the clinically important, p blood group phenotype. *Transfusion* 2003; **43**:899–907

7 Hellberg Å, Ringressi A, Yahalom V, Säfwenberg J, Reid ME, Olsson ML: Genetic heterogeneity at the glycosyltransferase loci underlying the GLOB blood group system and collection. *Br J Haematol* 2004; **125**:528–536

8 Hellberg A, Poole J, Olsson ML: Molecular basis of the globoside-deficient P^k blood group phenotype. *J Biol Chem* 2002; **277**: 29455–29459

9 Eiberg H, Møller N, Mohr J, Nielsen LS: Linkage of transcobalamin II (TC2) to the P blood group system and assignment to chromosome 22. *Clin Genet* 1986; **29**:354–359

10 Julier C, Lathrop GM, Reghis A, Szajnert M-F, Lalouel J-M, Kaplan J-C: A linkage and physical map of chromosome 22, and some applications to gene mapping. *Am J Hum Genet* 1988; **42**:297–308

11 Kaplan J-C, Emanuel BS: Report of the committee on the genetic constitution of chromosome 22. Human Gene Mapping 10. *Cytogenet Cell Genet* 1989; **51**:372–383

12 Iwamura K, Furukawa K, Uchikawa M, Sojka BN, Kojima Y, Wiels J, Shiku H, Urano T, Furukawa K: The blood group P1 synthase gene is identical to the Gb3/CD77 synthase gene: a clue to the solution of the P1/P2/p puzzle. *J Biol Chem* 2003; **45**:44429–44438

13 Uitenbroek D: SISA-Binomial 1997; Available at http://home.clar.net/sisa/binomial.htm (cited 2005 June 6).

14 Halperin E, Eskin E: Haplotype reconstruction from genotype data using imperfect phylogeny. *Bioinformatics* 2004; **20**:1842–1849

15 Naiki M, Marcus DM: An immunochemical study of the human blood group P_1, P, and P^k glycosphingolipid antigens. *Biochemistry* 1975; **14**:4837–4841

16 Graham HA, Williams AN: A genetic model for the inheritance of the P, P_1 and P^k antigens. *Immunol Commun* 1980; **9**:191–201

17 Hellberg A: Sixteen polymorphic sites in the 5′/3′-UTR of the P^k gene do not correlate with P1/P2 phenotypes (Abstract). *Vox Sang* 2004; **87**:81