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

Åsa Hellberg, M Alan Chester and Martin L Olsson*

Address: Division of Hematology and Transfusion Medicine, Department of Laboratory Medicine, Lund University and the Blood Centre, Lund University Hospital, Lund, Sweden

Email: Åsa Hellberg - asa.hellberg@med.lu.se; M Alan Chester - Alan.Chester@med.lu.se; Martin L Olsson* - Martin_L.Olsson@med.lu.se

* Corresponding author

Abstract

Background: The molecular genetics of the P blood group system and the absence of P₁ antigen in the p phenotype are still enigmatic. One theory proposes that the same gene encodes for both the P₁ and Pₖ glycosyltransferases, but no polymorphisms in the coding region of the Pₖ gene explain the P₁/P₂ phenotypes. We investigated the potential regulatory regions up- and downstream of the A4GALT (Pₖ) gene exons.

Results: P₁ (n = 18) and P₂ (n = 9) samples from donors of mainly Swedish descent were analysed by direct sequencing of PCR-amplified 5'- and 3'-fragments surrounding the Pₖ coding region. Seventy-eight P₁ and P₂ samples were investigated with PCR using allele-specific primers (ASP) for two polymorphisms previously proposed as P₂-related genetic markers (-551_-550insC, -160A>G). Haplotype analysis of single nucleotide polymorphisms was also performed with PCR-ASP. In ~1.5 kbp of the 3'-untranslated region one new insertion and four new substitutions compared to a GenBank sequence (AL049757) were found. In addition to the polymorphisms at positions -550 and -160, one insertion, two deletions and one substitution were found in ~1.0 kbp of the 5'-upstream region. All 20 P₂ samples investigated with PCR-ASP were homozygous for -550insC. However, so were 18 of the 58 P₁ samples investigated. Both the 20 P₂ and the 18 P₁ samples were also homozygous for -160G.

Conclusion: The proposed P₂-specific polymorphisms, -551_-550insC and -160G, found in P₂ samples in a Japanese study were found here in homozygous form in both P₁ and P₂ donors. Since P² is the null allele in the P blood group system it is difficult to envision how these mutations would cause the P₂ phenotype. None of the novel polymorphisms reported in this study correlated with P₁/P₂ status and the P₁/p mystery remains unsolved.

Background

The P-related blood groups include four antigens that are predominantly of glycolipid nature and occur in related biosynthetic pathways [1].

The GLOB blood group system [International Society of Blood Transfusion (ISBT) number 028] comprises the P antigen and the GLOB collection (ISBT number 209) includes the Pₖ antigen and also LKE that is not discussed.
The P1 antigen is assigned to ISBT system number 003. Five phenotypes depending on the presence or absence of the three antigens, P1, P, and Pk, are known (Table 1). The presence of all three antigens results in the P1 phenotype but absence of the P1 antigen causes the P2 phenotype. If both P1 and P are absent the phenotype P2k arises. Absence of P but presence of P1 and Pk results in the P1k phenotype. Absence of all three antigens results in the p phenotype. In each of the phenotypes naturally occurring-antibodies can arise against the missing antigen, invariably so in the case of P and Pk but less frequently for P1. These phenotypes can be explained biochemically by the presence or absence of some of the enzymes shown to catalyse the pathways shown in Figure 1.

The P1 antigen is present on hematopoietic cells [3,4] and other cells [5]. The strength of the antigen expression can differ from one person to another and it seems to be dependent on gene dosage [1]. Frequencies of the P1 phenotype vary in different ethnic groups, for example, ~80%
of Caucasians are P1, compared to only ~20% of Asians [1]. This may be due to selective pressure since P1 and related antigens can act as cellular receptors for microorganisms and biotoxins [5].

Two genes that code for the Pk- and P-synthesizing enzymes, 4-α-galactosyltransferase (α4GalT, Gb3 synthase) [6] and 3-β-N-acetylgalactosaminyltransferase (β3GalNAcT, globoside/Gb4 synthase) [7], respectively, have been cloned. Subsequently, numerous mutations have been found to explain the rare P1k, P2k [8,9] and p [6,9-12] phenotypes.

The molecular genetic background of the P1 antigen remains unknown. Several theories exist, including one model suggesting that the same α4GalT is able to transfer galactosyl residues to both lactosylceramide and paragloboside but in order to use the latter as the acceptor, a regulatory protein is required [13]. Another model postulates the existence of two different enzymes, and thus two genes, requiring both of them to be inactivated to cause the p phenotype [13]. A third model proposes a single gene with three alleles, one allele coding for an α4GalT that can utilise lactosylceramide and paragloboside as acceptors, one allele using lactosylceramide only and the third allele coding for an inactive form of the transferase [14]. However, none of the known polymorphisms (109A>G, 903G>C, 987G>A) in the coding region of the Pk gene explains the P1/P2 phenotypes [6].

Recently, Iwamura et al. [15] suggested that transcriptional regulation caused by two polymorphisms (-551_-550insC, -160A>G) in the 5'-upstream region of the Pk gene might be the reason for the P1/P2 phenotypes.

The Pk gene was originally thought to comprise two exons, but recent GenBank depositions indicate the presence of three exons with the whole coding region in exon 3, as shown in Figure 2. Various publications have considered different transcription starting points resulting in different numbering of the same nucleotide positions. The numbers used here are described in the legend to Figure 2.

In this study we have investigated an extended sequence surrounding the coding region of the Pk gene including untranslated exons and intronic portions as well as potentially regulatory regions 5' or 3' of the transcribed region. Contrary to a previous report [15], we found no clear-cut correlation with the P1/P2 phenotype, neither between previously described polymorphisms in the 5'-regulatory
Results

Screening for the -551_-550insC and -160A>G polymorphisms by PCR-ASP

Seventy-eight samples were screened for the two genetic markers previously [15] suggested to cause the P2 phenotype. The results are summarized in Table 2. Two haplotypes, -550T;160A and -551_-550insC;160G, were found whilst the other theoretically possible haplotypes, 550T;160G and -551_-550insC;160A, were not detected in this study. Each of the 20 samples, that were phenotyped as P2 were homozygous for both -551_-550insC and -160A>G, which could indicate that these polymorphic positions were indeed P2-specific as proposed. However, 18 of the 58 P1 samples investigated were also homozygous for the same polymorphic markers. Of the remaining P1 samples, 32 were heterozygous at both nucleotide positions and only eight samples were homozygous for the proposed P1-specific combination, -550T;160A. Twenty-nine of the 32 heterozygous samples were available for analysis by two haplotype-specific PCR reactions (-550T;160A and -551_-550insC;160G). All samples were positive in both PCR reactions, indicating that the samples were heterozygous for the combinations -550T;160A and -551_-550insC;160G and not for -550T;160G and -551_-550insC;160A. Additionally, 26 samples with the rare phenotypes P1k (n = 3), P2k (n = 3) and p (n = 20) were screened. One of the P1k samples was homozygous for the genetic markers reported to be associated with the P2 phenotype, and the other two were heterozygous. Of the P2k samples two had the expected polymorphisms (-551_-550insC;160G), but the third P2k was heterozygous and thus the first reported sample that lacks the P1 antigen in spite of a genotype not homozygous for the -551_-550insC and -160G markers. Sixteen of the 20 p samples were homozygous for -551_-550insC;160G, thus consistent with their lack of P1, but one was heterozygous and three were homozygous for -550T;160A. Due to lack of available DNA from these rare individuals the haplotype-specific PCR was only run on one of the four heterozygous samples. As above, this sample was also positive in both PCR reactions.

Amplification of the 5’- , 3’- and coding regions of the Pk gene for DNA sequencing

Ten of the 18 P1 samples with P2-associated polymorphisms (-551_-550insC;160G) were chosen for further investigation by sequencing the 5’-region and the 3’-region, as were all eight P1(-550T;160A) and nine of the 20 P2 samples. In three samples from each category the whole Pk gene reading frame, located in exon 3, was also sequenced.

In the 5’-region upstream of exon 1, four novel polymorphisms were detected compared to a sequence deposited in GenBank (accession number AL049757). These findings comprised a substitution (-770C>T), an insertion (-107_-106insG) and two deletions (-907_-903del and -17_8del). Interestingly, the latter deletion is located across the border of the 5’-region and exon 1.

In the 3’-UTR, five new polymorphisms were found, four of which were substitutions (1409G>A, 1495C>A, 1523G>A, 1697G>A) and one was an insertion (1592dupG). Figure 2 shows the relative positions of the polymorphisms investigated.

The distribution of the polymorphisms in the three different categories, P1(-550T;160A), P1(-551_-550insC;160G) and P2 are shown in Figure 3. As can be seen, none of the polymorphisms is a genetic marker specific for the P2 phenotype. On the other hand, individuals from the P1(-550T;160A) category are homozygous for 12 of the 16 investigated polymorphisms. However, when individuals from the P1(-551_-550insC;160G) samples are included such pattern is no longer evident. It can also be noted that the frequency of some of the variants appears

---

### Table 2: The occurrence of the -550T;160A and -551_-550insC;160G haplotypes in individuals with various P phenotypes.

| Phenotype | No. | -550T;160A (homozygous) | -550T;160A -551_-550insC;160G (heterozygous) | -551_-550insC;160G (homozygous) |
|-----------|-----|--------------------------|------------------------------------------|----------------------------------|
| P1        | 58  | 8                        | 32                                       | 18                               |
| P1        | 20  | 0                        | 0                                        | 20                               |
| P1k       | 3   | 0                        | 2                                        | 1                                |
| P2k       | 3   | 0                        | 1                                        | 2                                |
| p         | 20  | 3                        | 1                                        | 16                               |
Figure 3

The distribution of the polymorphic variants in the three different sample categories: P1(-550T;-160A), P1(-551_-550insC;-160G) and P2. The diagram shows data of 16 polymorphisms investigated. The data for the 5'- and the 3'-regions were obtained by sequencing while most of the data for exon 3 originated from PCR-ASP analysis. Eight, P1(-550T;-160A), ten P1(-551_-550insC;-160G) and nine P2 samples were investigated. However, for the 903C>G polymorphism the corresponding numbers were five, five and seven samples. In position -107_-106 all investigated samples had three guanosines instead of two as reported in the reference sequence (AL049757). We therefore do not consider this position as polymorphic in our analysis and it is consequently not included in this figure. cons. = consensus.
to be relatively low, less than ~10%, for six of the polymorphic sites analyzed.

While sequencing exon 3, an unexpected mutation, 441G>A, was encountered in a P1 sample of African origin. No other samples, including two of African descent, examined in this study had this particular mutation, which would not alter the amino acid sequence, or any other new polymorphisms.

Haplotype analysis of SNPs in the reading frame and a polymorphism in the 3'-UTR using PCR-ASP

We also performed PCR-ASP utilising polymorphisms in exon 3 and in the 3'-UTR of the \textit{Pk} gene to determine the\ cis/trans\ linkage of SNPs and to establish if combinations of polymorphisms correlated with specific alleles. The SNPs chosen were two of those previously described, i.e. 109A>G and 987G>A. Analysis of the polymorphism 903C>G was considered unnecessary due to its proximity to nt. 987. In addition, of the new polymorphisms the one furthest downstream (1697G>A) was analyzed. The results are summarized in Table 3. None of the haplotypes correlated well with the P\textsubscript{1}/P\textsubscript{2} phenotypes although 83% of the alleles among the P2 samples and 75% of the P1(-550 insC;-160T) category samples had the A-1 haplotype, whilst in the P1(-550T;-160A) category samples only 12.5% had this partial haplotype (B-1). The remaining alleles in this latter category consisted of 12.5% of the rare B-3 and 75% of the B-2 haplotype.

**Discussion**

The molecular background of the P blood group system has been a subject of speculation since its discovery. Antigens formed as a result of closely related biosynthetic pathways are now known to arise from independent genetic loci. However, the current paradox revolves around the 4-\textalpha galactosyltransferase (\textalpha 4GalT) that synthesises the P\textsuperscript{k} antigen and possibly also the P1 antigen. If it is indeed exactly the same enzyme synthesizing both, then the P\textsubscript{2} and P\textsubscript{2k} phenotypes should not exist, as the P1-synthesizing \textalpha 4GalT that is lacking in these individuals should also result in the loss of the P\textsuperscript{k} and, subsequently, P antigens. The absence of the P1 antigen is not due to a defect in the biosynthesis of paragloboside since this glycolipid is a major precursor of the erythrocyte ABH antigens that are unaffected by the P\textsubscript{1}/P\textsubscript{2} status. The absence of P1 antigen in the P\textsubscript{2} phenotype is likely to be caused by inefficient or absent glycosyltransferase activity that can be due to a structural defect in the enzyme itself or more indirectly due to various other factors necessary to ensure the enzyme's optimal expression, localisation and efficiency[1]. Studies are currently in progress to analyse \textit{A4GALT} mRNA by real-time PCR and other methods.

The P1 and P\textsuperscript{k} antigens both share the same terminal carbohydrate structure and therefore the question arises whether the same enzyme is, or is not, synthesizing both antigens. Various theories exist but so far none has been proven correct. Recently Iwamura et al [15] proposed that the P1 synthase is identical to the P\textsuperscript{k} synthase and that the phenotypic difference may be caused by two polymorphisms, -160A>G and -551_-550insC, found in the 5' upstream regulatory region. Unfortunately, these authors were not able to show any functional effects of these two polymorphisms, at least not when tested in a fibroblast cell line. As they commented themselves, a haematopoietic cell line, preferentially even an erythropoietic one, would have been the optimal choice for the challenge but was technically difficult. Iwamura et al. also demonstrated the presence of cryptic and intracellular P1 antigen in cells from individuals known to have the P\textsubscript{1} red blood cell phenotype. Whether this surprising finding is method-dependent or may differ between populations is currently unclear.

| Phenotype | No. of samples | Polymorphism | Haplotype | No. of haplotypes |
|-----------|----------------|--------------|-----------|------------------|
| P\textsubscript{1} | 10 | insC G A G G | A-1 | 15 |
| P\textsubscript{1} | 8 | T A A G G | B-1 | 2 |
| P\textsubscript{2} | 9 | insC G A G G | A-1 | 15 |

Table 3: Correlation of phenotype and haplotype
Our data indicate that an individual's P1/P2 status is not due to the -551_-550insC;160G haplotype, contrary to the previously published indication [15]. However, it is striking that all (except one P2k) samples with the P2 phenotype tested here have the same -551_-550insC;160G haplotype as all ten P2 samples tested by Iwamura et al. This clearly argues that the genetic linkage between the A4GALT locus and P2 status must be relatively strong. The results shown in Figure 3 indicate that none of the 16 polymorphic markers investigated can predict the serologically determined P1/P2 status. It is also interesting to note the high frequency with which SNPs occur in the A4GALT gene, as opposed to the gene coding for the next glycosyltransferase in this biosynthetic pathway, the P synthase, in which we have found no variations at all. The only exceptions are rare mutations causing the P1k and P2k phenotypes [8,9]. Despite the apparent variability in the A4GALT gene, the current study suggests that a limited number of haplotypes, with the A-1 and B-2 haplotypes being the predominant ones in the Swedish population, may constitute genetic clusters within which further variation has arisen (as judged by the lack of homogeneity within each group, see Table 3), still without obvious correlation to the blood group phenotype.

However, all this does not differentiate between the one-gene theory and the possibility of a tightly coupled independent locus responsible for P1 antigen expression. Since there are no apparent α4GalT homologues in this genetic region this may imply that a hypothetical second closely linked gene would give rise to either a regulatory molecule modifying the acceptor specificity of the Pk-synthesizing α4GalT (analogous to lactose synthetase [16]) or a chaperone type of molecule to make a fraction of the α4GalT molecules more suitably located/positioned for P1 synthesis. The finding of intracellular P1 antigen in P2 individuals [15], would tend to support the latter possibility. Chaperones have been shown to be involved in processes related to glycosyltransferase action [17] but it is somewhat difficult (although not impossible) to imagine an α4GalT-specific chaperone to be the solution for this long-standing enigma.

Conclusion
The study of potential regulatory regions surrounding the Pk gene coding sequence revealed nine previously unreported polymorphisms but none of them correlated with the P1/P2 red blood cell phenotypes. Two polymorphisms, -551_-550insC and -160A>G, suggested to cause the P2 phenotype in Japanese individuals [15] were found in homozygous form also in P1 samples in this study and since the P2 is the null phenotype of this blood group system, it is therefore very unlikely that these mutations cause the P2 phenotype.

Methods

Blood samples and DNA preparation
Samples with the P1 (n = 58) and P2 (n = 20) phenotypes were chosen from our in-house panel of test erythrocytes. The majority of the donors are of Swedish origin but a few are of Asian or African descent. Three P1k, three P2k and 20 p samples, genetically characterized in our laboratory, were also included for screening purposes [8,9,12,18]. The erythrocyte phenotype was determined by standard serological techniques. DNA was prepared from EDTA blood using a simple salting-out method for small volumes modified from Miller et al. [19], or Qiagen QIAmp Blood Extraction kit (Qiagen GmbH, Hilden, Germany). The DNA was dissolved in H2O at a concentration of 100 ng/µl.

Screening for the -551_-550insC and -160A>G polymorphisms by PCR-ASP
All oligonucleotide primers used in the study were synthesized by DNA Technology ApS (Aarhus, Denmark) and the sequences are shown in Table 4. PCR with allele-specific primers designed to detect the polymorphisms at -551_-550insC and -160A>G, described to cause the P2 phenotype [15] was performed. For all heterozygous samples double allele-specific amplification, -551_-550insC;160G and -550T;160A, were performed. The primer combinations used are listed in Table 5 and the locations in Figure 4. Primers were mixed with 100 ng of genomic DNA, 2 nmol of each dNTP, 2% glycerol, 1% cresol red and 0.5 U of AmpliTaq Gold (Perkin Elmer/Roche Molecular Systems, Branchburg, NJ, USA) in 10 × PCR buffer with 15 mM MgCl2. The final reaction volume was 11 µl. Thermocycling was undertaken in GeneAmp PCR system 2400/2700 (Perkin-Elmer/Cetus, Norwalk, CT, USA) under PCR conditions described in Table 5.

Amplification of the 5'-, 3'- and coding regions of the Pk gene for DNA sequencing
The coding region, exon 3, in the Pk gene was amplified with the primer pair Pk-(-140)-F and Pk-1120-R in the Expand High Fidelity PCR System (Roche Molecular Systems, Pleasanton, CA, USA) and sequenced as previously described [12].

The 5'-regulatory region of the Pk gene was amplified with primers Pk-5'(-1056)-F and Pk-int1-160-R and Pk-5'(-131)-F and Pk-int1-160-R. Amplification was performed in a reaction volume of 22 µl with four pmol of each primer, 2 nmol of each dNTP, 100 ng of genomic DNA, GC-rich enzyme mix (0.5 U per reaction), GC-rich resolution solution and buffer with a final MgCl2 concentration of 1.5 mM (GC-rich PCR System, Roche Diagnostics GmbH, Mannheim, Germany). Thermocycling was undertaken in GeneAmp PCR system 2400/2700 (Perkin-
Elmer/Cetus): Initial denaturation at 96°C for 7 min followed by 10 cycles at 94°C for 30 s, 62°C for 30 s and 72°C for 1 min and then 25 cycles at 94°C for 30 s, 60°C for 30 s and 72°C for 1 min.

For the 3’-region of the Pk gene 5 pmol of primers Pk-(1006)-F and Pk-(1881)-R were mixed with 100 ng of genomic DNA, 2 nmol of each dNTP, 2% glycerol, 1% cresol red and 0.5 U of AmpliTaq Gold (Perkin Elmer/ Roche Molecular Systems) in 10 × PCR buffer with 15 mM MgCl₂. The final reaction volume was 11 µl. PCR was run at 96°C for 7 min followed by 35 cycles at 94°C for 30 s, 64°C for 30 s and 72°C for 1 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. Louis, MO, USA) following high-voltage electrophoresis and purified using Qiaquick gel extraction kit (Qiagen). 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

### Table 4: Oligonucleotide primers used in this study.

| Primer name         | Nucleotide sequence (5’→3’)                                                                 | Function |
|---------------------|---------------------------------------------------------------------------------------------|----------|
| Pk-5’-(-1056)-F†    | ACAGCCTGTGATGGGAATGAC                                                                     | a, b, e  |
| Pk-5’-(-740)-R      | TTGAGTGCTAGCGGCGAC                                                                         | b        |
| Pk-5’-(-834)-F      | GGAGAATTTACCAATACGCAATATGT                                                               | e        |
| Pk-5’-(-550)-R      | CCTGCCCCCATCTTCAGGCG                                                                     | e        |
| Pk-5’-(-550insC)-R  | ACAAATTTACCAATAGCAATATGCT                                                                | e        |
| Pk-5’-(-477)-F      | GCGGCTTTAAGGATACGCAA                                                                     | b        |
| Pk-5’-(-477)-R      | CTGATCCACCGGCGCTG                                                                         | b        |
| Pk-5’-(-235)-F      | GCCGCTCTACGCTGCC                                                                           | b        |
| Pk-5’-(-131)-F      | GGACCGGAGCGGCGGA                                                                         | a, b     |
| Pk-5’-(-160Gmis)-R  | CCGGCTCCCGAGACCT                                                                         | e        |
| Pk-5’-(-160Amis)-R  | CCGGCTCCCGAGACCT                                                                         | e        |
| Pk-5’-(-160G)-R     | CCGGCTCCCGAGACCT                                                                         | e        |
| Pk-5’-(-160A)-R     | CCGGCTCCCGAGACCT                                                                         | e        |
| Pk-int1-35-R        | CGTCCCCCGCAGGCC                                                                            | b        |
| Pk-int1-160-R       | GCACAAATCTGCGTCCGTC                                                                      | a, b     |
| Pk-ex2-74_75insC-F  | AGGTCGCCGTGCTGAC                                                                         | e        |
| Pk-int2-R           | GGGTGCAACCGTGATTGTAAG                                                                    | e        |
| Pk-109G-F           | TCACTGTGTTGCTGCTGCC                                                                       | e        |
| Pk-109A-F           | TCACTGTGTTGCTGCTGCC                                                                       | e        |
| Pk-987G-F           | CACCCGGTCACCGGCG                                                                         | e        |
| Pk-987A-F           | CACCCGGTCACCGGCG                                                                         | e        |
| Pk-1120-R           | GGAAGGGCGGGCCAGTC                                                                        | e        |
| Pk-1006-F           | CACGGACGACTGCTGCC                                                                        | c, d     |
| Pk-1253-F           | GACAGTCTGGCTGCTGCTG                                                                     | d        |
| Pk-1697G-R          | CCTGGCTGGAGGAGGAGGAG                                                                     | e        |
| Pk-1697A-R          | CCTGGCTGGAGGAGGAGGAG                                                                     | e        |
| Pk-1791-35-R        | TTATTTATTTATATTTATTTATTTGCG                                                              | d        |
| Pk-1881-R           | CCCGCTACGAGAAATGAGGAG                                                                    | c, d     |
| HGH-F               | TGCCGTTACACCAAGCT                                                                         | f        |
| HGH-R               | CCACTACGAGGATTTGCTGTTGTTGTTG                                                           | f        |
| JK-781L-F           | GCATGCGTGCGAGGATGCAA                                                                     | f        |
| JK-943L-R           | GACGCAAGGAGGGGTGGTC                                                                     | f        |
| MO-21               | GGGTGGAGGAGGAGGAGGAG                                                                    | f        |
| MO-31               | CCAAGACCCCCGCCAGCA                                                                      | f        |

† F = Forward. R = Reverse. The underlined nucleotide is mismatched for better specificity.

- a. used for amplification of the 5’ regulatory region.
- b. used for sequencing of the 5’ regulatory region.
- c. used for amplification of the 3’-UTR.
- d. used for sequencing the 3’-UTR.
- e. used for PCR-ASP.
- f. used as internal control.
direct DNA sequencing with capillary electrophoresis and automated fluorescence-based detection according to the manufacturer’s instructions. Besides the PCR primers, internal primers were used as sequencing primers, see Table 4. To avoid detection of artefacts, sequencing was performed on both strands and using independently obtained fragments.

**Table 5: Oligonucleotide primer combinations and PCR conditions (previously not described) used for PCR-ASP in this study.**

| Primer combination | Specific primer (pmol) | Control primer | DMSO (%) | Annealing temperature (°C) | Extension time (sec) |
|-------------------|------------------------|----------------|----------|----------------------------|---------------------|
|                   | HGH-F/R (pmol)         | JK-L-F/R (pmol) | MO-21/31 (pmol) |                          |                     |
| Pk-5’(-1056)-F    | 5                      | 1              |           | 64                         | 40                  |
| Pk-5’(-550)-R     | 5                      |                |           |                            |                     |
| Pk-5’(-1056)F     | 5                      | 1              |           | 64                         | 40                  |
| Pk-5’(-550insC)-R | 5                      |                |           |                            |                     |
| Pk-5’(-834)-F     | 7.5                    | 0.5            | 4        | 62                         | 60                  |
| Pk-5’(-160A-mis)-R| 7.5                    |                |          |                            |                     |
| Pk-5’(-834)-F     | 5                      | 0.5            | 3        | 62                         | 60                  |
| Pk-5’(-160G-mis)-R| 5                      |                |          |                            |                     |
| Pk-5’(-550T)-F    | 7.5                    | 0.5            |          | 63                         | 40                  |
| Pk-5’(-160A)-R    | 7.5                    |                |          |                            |                     |
| Pk-109A-F or      | 6                      | 0.4            | 2        | 65                         | 90                  |
| Pk-109G-F         |                        |                |          |                            |                     |
| Pk-1697A-R or     | 6                      | 0.4            |          |                            |                     |
| Pk-1697G-R*       |                        |                |          |                            |                     |
| Pk-987A-F         | 5                      | 0.4            |          | 66                         | 60                  |
| Pk-1697A-R        | 5                      |                |          |                            |                     |
| Pk-987G-F         | 5                      | 0.4            |          | 66                         | 60                  |
| Pk-1697G-R        | 5                      |                |          |                            |                     |
| Pk-109A-F or      | 7.5                    | 0.5            |          | 64                         | 60                  |
| Pk-109G-F         |                        |                |          |                            |                     |
| Pk-1120-R**       |                        |                |          |                            |                     |

Denaturation was carried out at 96°C for 7 min followed by 35 cycles at 94°C for 30 s. The annealing time was 30 s.

* Four double allele-specific amplifications were performed.

** Detection of an insertion in exon 2 and linkage of SNPs in the reading frame with a polymorphism in the 3’-UTR using PCR-ASP **

PCR-ASP for a previously found insertion, 75dupC in exon 2 (then believed to be exon 1), was also performed as described [12]. PCR-ASP was performed to investigate if the polymorphisms at nt. 109, 987 and 1697 were present in an allele-specific pattern. The reaction mixtures comprised 100 ng of genomic DNA, 2 nmol of each dNTP, 2% glycerol, 1% cresol red and 0.5 U of AmpliTaq
Gold (Perkin Elmer/Roche Molecular Systems) in 10 × PCR buffer with 15 mM MgCl₂. The final reaction volume was 11 µl. Primer combinations and PCR conditions are described in Figure 4 and Table 5.

List of abbreviations
ASP, allele specific primer
bp, base pairs
ins, insertion
del, deletion
kb, kilo bases
PCR, polymerase chain reaction
SNP, single nucleotide polymorphism
UTR, untranslated region
ISBT, International Society of Blood Transfusion.

Authors’ contributions
ÅH carried out the experimental studies and participated in the discussion and preparation of the manuscript. AC participated in the discussion and preparation of the manuscript. MLO contributed to the design and coordination of the study and participated in preparation of the manuscript. All authors read and approved the final manuscript.

Acknowledgements
Sources of support: This work was supported in part by the Swedish Research Council (project no. K2005-71X-14251), the Medical Faculty at Lund University, governmental ALF research grants to Lund University Hospital, the Inga & John Main Foundation for Medical Research and the Lund University Hospital Donation Funds.

Part of this study has previously been reported in abstract form at the 28th Congress of the International Society of Blood Transfusion held in Edinburgh, UK, on 11–15 July 2004 [20].

References
1. Daniels G: Human Blood Groups 2nd edition. Oxford, UK, Blackwell Scientific; 2002.
2. Daniels GL, Fletcher A, Garratty G, Henry S, Jorgensen J, Judd WJ, Levene C, Lomas-Francis C, Moulds JJ, Moulds JM, Moulds M, Overbeeke M, Reid ME, Roquer P, Scott M, Sistonen P, Smart E, Tani Y, Wendel S, Zelinski T: Blood group terminology 2004: from the International Society of Blood Transfusion committee on terminology for red cell surface antigens. Vox Sang 2004, 87:304-316.
3. Dunstan RA: Status of major red cell blood group antigens on neutrophils, lymphocytes and monocytes. Br J Haematol 1986, 62:301-309.
4. Dunstan RA, Simpson MB, Rosse WF: Presence of P blood group antigens on human platelets. Am J Clin Pathol 1985, 83:71-73.
5. Spitalnik PF, Spitalnik SL: The P blood group system: biochemical, serological, and clinical aspects. Transfus Med Rev 1995, 9:110-122.
6. Steffensen R, Carlier K, Wiels J, Levery SB, Stroud M, Cedergren B, Nilsson SB, Bennett EP, Jersild C, Clausen H: Cloning and expression of the histo-blood group Pk UDP-galactose: Gal-beta 1,4Gal-beta 1-Cer alpha 1,4-galactosyltransferase. Molecular genetic basis of the p phenotype. J Biol Chem 2000, 275:16723-16729.
7. Okajima T, Nakamura Y, Uchikawa M, Haslam DB, Numata SI, Furukawa K, Urano T, Furukawa K: Expression cloning of human globoside synthase cDNAs. Identification of beta3Gal-T3 as UDP-N-acetylgalactosamimeliglobotriaosylceramide beta1,3-N-acetylgalactosaminytransferase. Molecular genetic basis of the p phenotype. J Biol Chem 2000, 275:40498-40503.
8. Hellberg Å, Poole J, Olsson ML: Molecular basis of the globoside-deficient Pk blood group phenotype. Identification of four inactivating mutations in the UDP-N-acetylgalactosamimeglobotriaosylceramide 3-beta-N-acetylgalactosaminyltransferase gene. J Biol Chem 2002, 277:29455-29459.
9. Hellberg Å, Ringressi A, Yahalom V, Salwenberg 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.
10. Furukawa K, Iwamura K, Uchikawa M, Sojka BN, Wiels J, Okajima T, Urano T, Furukawa K: Molecular basis for the p phenotype.

Figure 4
Schematic representation of the primers used to identify the polymorphisms at -551_-550insC and -160A>T, described to cause the P₂ phenotype. The figure also shows combinations and positions for the primers used to detect a linkage phase of SNPs in the reading frame with a polymorphism in the 3’-UTR. For size reference, see Figure 2.
Identification of distinct and multiple mutations in the alpha1,4-galactosyltransferase gene in Swedish and Japanese individuals. J Biol Chem 2000, 275:37752-37756.

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

12. Hellberg A, Steffensen R, Yahalom Y, 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.

13. Naiki M, Marcus DM: An immunochemical study of the human blood group P1, P, and PK glycosphingolipid antigens. Biochemistry 1975, 14:4837-4841.

14. Graham HA, Williams AN: A genetic model for the inheritance of the P, P1 and PK antigens. Immunol Commun 1980, 9:191-201.

15. Iwamura K, Furukawa K, Uchikawa M, Sojka BN, Kojima Y, Wielis J, Shiku H, Uran 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, 278:44429-44438.

16. Hill RL, Brew K: Lactose synthetase. Adv Enzymol Relat Areas Mol Biol 1975, 43:411-490.

17. Ju T, Cummings RD: A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase. Proc Natl Acad Sci USA 2002, 99:16613-16618.

18. Hellberg A, McGrath C, Reid ME, Olsson ML: A novel missense mutation in the A4GALT gene of Amish individuals with the p phenotype. Transfusion 2003, 43:99A.

19. Miller SA, Dykes DD, Polesky HF: A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988, 16:1215.

20. Hellberg A, Chester MA, Olsson ML: Sixteen polymorphic sites in the 5'/3'-UTR of the Pk gene do not correlate with P1/P2 phenotypes. Vox Sang 2004, 87:p18.11.

21. Cedergren B: Population studies in northern Sweden. IV. Frequency of the blood type p. Hereditas 1973, 73:27-30.