Genetics and Molecular Biology, 25, 1, 13-19 (2002)

www.sbg.org.br

Identification and characterization of polymorphisms at the HSA α_1-*acid glycoprotein* (*ORM**) gene locus in Caucasians

Catherine M. Owczarek[1], Aleksander L. Owczarek[2] and Philip G. Board[3]

[1]*Centre for Functional Genomics and Human Disease, Monash Institute of Reproduction and Development, Monash University, Melbourne, Australia.*
[2]*Department of Mathematics and Statistics, University of Melbourne, Parkville, Australia.*
[3]*Molecular Genetics Group, John Curtin School of Medical Research, Australia.*

Abstract

Human α_1-acid glycoprotein (AGP) or orosomucoid (ORM) is a major acute phase protein that is thought to play a crucial role in maintaining homeostasis. Human AGP is the product of a cluster of at least two adjacent genes located on HSA chromosome 9. Using a range of restriction endonucleases we have investigated DNA variation at the locus encoding the *AGP* genes in a panel of healthy Caucasians. Polymorphisms were identified using *Bam*HI, *Eco*RI, *Bgl*II, *Pvu*II, *Hin*dIII, *Taq*I and *Msp*I. Non-random associations were found between the *Bam*HI, *Eco*RI, *Bgl*II RFLPs. The RFLPs detected with *Pvu*II, *Taq*I and *Msp*I were all located in exon 6 of both AGP genes. The duplication of an *AGP* gene was observed in 11% of the indiviuals studied and was in linkage disequilibrium with the *Taq*I RFLP. The identification and characterization of these polymorphisms will prove useful for other population and forensic studies.

Key words: Human α_1-acid glycoprotein, RFLP, linkage disequilibrium.

Received: March 5, 2002; accepted: March 25, 2002.

Introduction

HSA α_1-acid glycoprotein (AGP, orosomucoid, ORM) is an abundantly expressed plasma protein whose levels rise dramatically during the acute phase response. A member of the lipocalin protein family, it is thought to function mainly as a transport protein for basic drugs although several other functions have been ascribed (Flower 1996). The expression of the ORM protein product in most individuals is controlled by two genes, *AGP1* and *AGP2* (Dente *et al.* 1987; Merritt and Board 1988), that are closely linked on HSA chromosome 9q31-q34.1 (Webb *et al.* 1987). A third gene, structurally identical to *AGP2* has been reported to exist in some individuals (Dente *et al.* 1987), and duplication of the *AGP1* gene has been demonstrated to occur in the Japanese population at an appreciable frequency (Nakamura *et al.* 2000). Considerable variation in the ORM polypeptide chain has been described. In addition to the two common alleles *ORM1*F* and *ORM1*S* (Johnson *et al.* 1969) a large number of variants have been identified in different populations (Yuasa *et al.* 1993).

Send correspondence to Catherine M. Owczarek. Monash Institute of Reproduction and Development, Monash University, Melbourne, Victoria 3168, Australia. E-mail: catherine.owczarek@med.monash.edu.au.

In this study we have used RFLP analysis to investigate DNA variation at the *AGP* gene locus. We have demonstrated the existence of RFLPs in the region upstream of the *AGP* gene locus as well as polymorphisms within the *AGP* gene cluster and have examined linkage disequilibrium between these sites. A duplication of one of the *AGP* genes was observed in the population studied and was strongly linked with the presence of a *Taq*I polymorphism.

Materials and Methods

Genomic DNA samples

The samples of DNA used for screening for RFLPs were from a set of 97 unrelated Caucasian blood donors recruited at the Canberra Red Cross Blood Transfusion Centre. An additional 20 random controls were obtained from healthy staff members at the John Curtin School of Medical Research. Family material was obtained from healthy Caucasian volunteers.

Genomic RFLP analysis

High molecular weight genomic DNA was extracted from the buffy coat from 10 mL peripheral blood (Grunebaum *et al.* 1984). Approximately 10 µg of genomic

DNA was digested with the following enzymes according to the manufacturer's specifications: *Bam*HI, *Eco*RI, *Bgl*II, *Pvu*II, *Hin*dIII, *Msp*I, and *Taq*I and electrophoresed through a 0.8% agarose gel. After Southern blotting of the DNA (Reed and Mann 1985) on to Gene-Screen *Plus* (Dupont) nylon membranes, the filters were hybridized overnight at 65 °C with an α ^{32}P-dCTP labeled α_1-AGP cDNA probe (Board *et al.* 1986).

Hardy-Weinberg and linkage disequilibrium analysis

Standard χ^2 tests were used to compare observed genotype frequencies with those expected under the Hardy-Weinberg equilibrium (Weir 1996). In order to test for linkage disequilibrium between the alleles of the different polymorphisms, contingency tables were used, with standard χ^2 tests and Fisher's exact tests, producing identical results. When using Fisher's exact test we used a two-sided p-value that was the minimum of 1 and twice the one sided p-value. Because the data cannot distinguish the two possible double heterozygotes gametic frequencies could not be inferred (Weir 1996), and hence the standard test for linkage equilibrium using hapolotype frequencies was not possible. In such cases various alternative approaches are possible. However, because in all cases when the data was analyzed using 2x2 contingency tables between the two less common alleles, the results were extreme in one direction or the other the inferences of linkage disequilibrium/equilibrium were unequivocal.

Results

Polymorphisms detected by *Bam*HI, *Eco*RI and *Bgl*II

When human genomic DNA digested with *Bam*HI was hybridized to the α_1-AGP cDNA probe, two constant bands: 4.5 kb and 2.5 kb and two variable bands: 14.8 kb (B1) and 13.6 kb (B2) were observed (Figure 1a). Co-dominant segregation was observed in Families D and M for the B1 and B2 alleles of the *Bam*HI RFLP (Figure 3). Mapping data (Merritt and Board 1988) indicated that these hybridizing bands corresponded to exons 1-5 of *AGP2* and exon 6 of both *AGP* genes respectively placing the polymorphic *Bam*HI site approximately 11 kb upstream of the *AGP1* gene.

Digestion of human DNA with *Eco*RI detected a two-allele polymorphism with bands at 12.6 kb (E1) and 11.6 kb (E2) and invariant bands at 16.8 kb and 6.9 kb (Figure 1b). Co-dominant segregation of the *Eco*RI RFLP was demonstrated in Family D (Figure 4). The 6.9 kb invariant band contains exon 1 of *AGP1* and exons 1-5 of *AGP2* (Merritt and Board 1988). Hybridization with a probe specific to exon 1 of *AGP1/2* (data not shown) indicated that the polymorphic *Eco*RI site was located upstream of *AGP1*.

*Bgl*II digestion also detected a two-allele polymorphism. Fragment lengths of 9.8 kb (Bg1) and 8.5 kb (Bg2) with invariant bands at 12.0 kb, 5.2 kb 0.8 kb and 0.7 kb were observed. Genetic transmission of the *Bgl*II RFLP was observed in Family D (Figure 3). The position of the polymorphic *Bgl*II site was determined from the nucleotide sequence (Merritt and Board 1988) and additional mapping experiments (data not shown). The 0.8 kb and 0.7 kb fragments corresponded to exons 2-3, and exons 4-5 respectively of both *AGP* genes. The 5.2 kb fragment contained exon 6 of *AGP1* and exon 1 of *AGP2* plus intergenic sequence. The polymorphic *Bgl*II band was detected with an exon 1-specific probe indicating that it was located upstream of the *AGP1* gene.

Polymorphisms detected by *Taq*I, *Hin*dIII, *Pvu*II and *Msp*I

*Taq*I digestion generated a two allele polymorphism consisting of either a 3.02 kb band (T2) or a 2.88 kb band (T1) with invariant bands at 4.5 kb, 1.4 kb 1.2 kb 0.84 kb

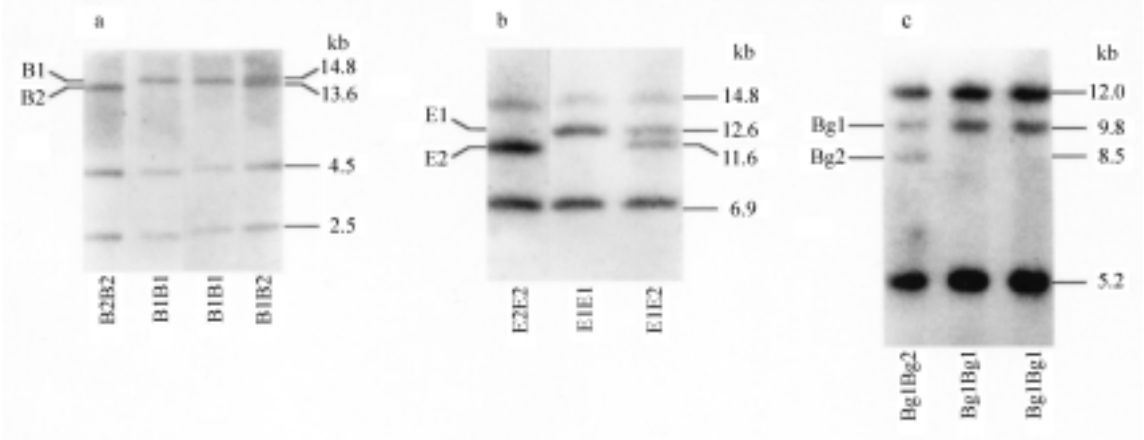

Figure 1 - Southern blot analysis of the DNA polymorphisms generated by digestion with: a. *Bam*HI (B); b. *Eco*RI (E); c. *Bgl*II (Bg) and probed with the cDNA probe pα_1AGP. The phenotype is given beneath each sample. The sizes of the polymorphic fragments are indicated and described in the text.

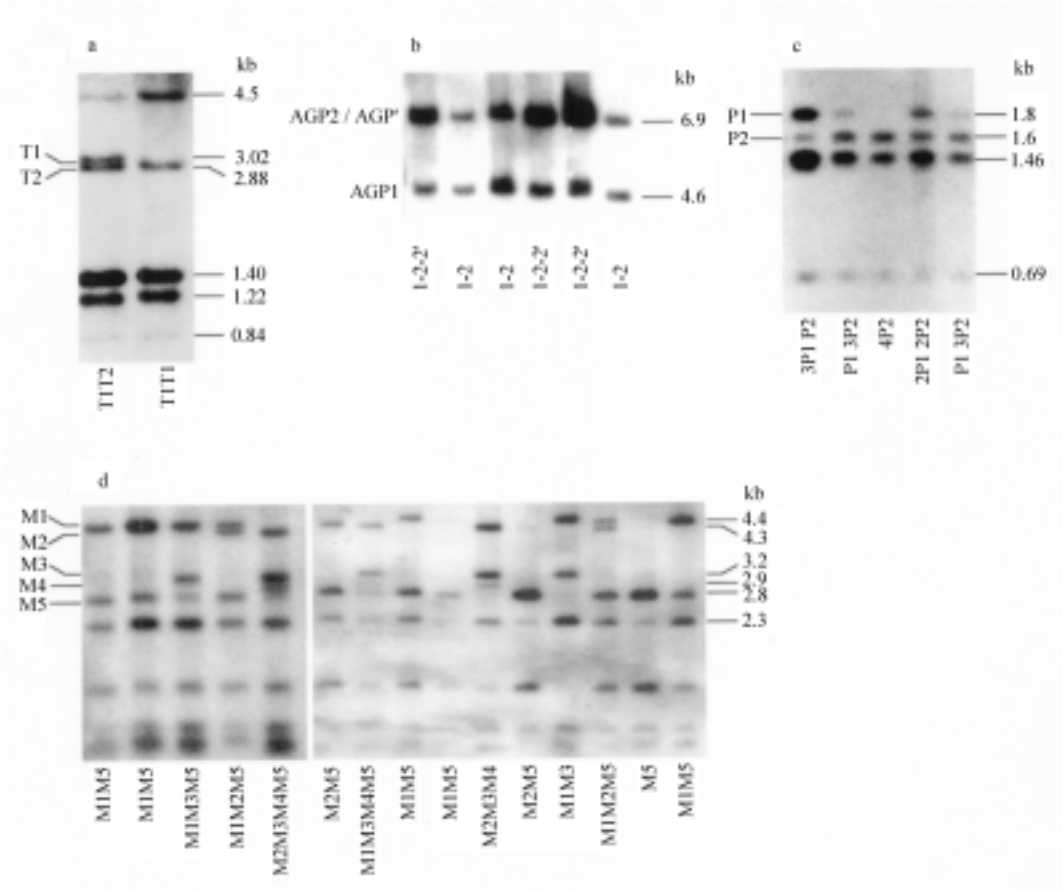

Figure 2 - Southern blot analysis of the DNA polymorphisms generated by digestion with: a. *Taq*I (T); b. *Hin*dIII. The more intense 6.9 kb band is assigned *AGP2-AGP'* because the exact identities of the duplicated *AGP* genes are not known in these samples. (1-2, normal two *AGP* gene array; 1-2-2', three *AGP* gene array); c. *Pvu*II (P); d. *Msp*I (M) and probed with the cDNA probe pα_1AGP. The phenotype is given beneath each sample. The sizes of the polymorphic fragments are indicated and described in the text.

and 0.285 kb (Figure 2a). Co-dominant segregation for the *Taq*I RFLP was observed in two informative families (Figure 3). Hybridization experiments and analysis of sequence data (Merritt and Board 1988) indicated that the *Taq*I polymorphic site located on an exon-6 containing fragment.

When human genomic DNA was digested with *Hin*dIII and hybridized to the α_1-AGP cDNA probe two bands at 4.6 kb (*AGP1*) and 6.9 kb (*AGP2*) were detected (Figure 2b) but the 6.9 kb band relative to the 4.6 kb band was more intense in 7 out of 65 individuals examined (11%) (Table I). The greater intensity of the 6.9 kb *Hin*dIII band relative to the 4.6 kb band in some individuals has been previously noted (Dente *et al.* 1985; Merritt and Board 1988) and correlates with an extra *AGP* gene (Dente *et al.* 1987; Nakamura *et al.* 2000). Individuals were scored either as 1-2 (*AGP1-AGP2* on each chromosome) or 1-2-2' indicating the presence of an extra *AGP* gene on one or both chromosomes (since homozygotes and heterozygotes would be indistinguishable under the conditions used in this study). Co-dominant segregation was observed in Family M (Figure 3) where the father had the intense 6.9 kb *Hin*dIII band and the mother had 6.9 kb and 4.5 kb *Hin*dIII bands of equal intensity. Two siblings have inherited the intense 6.9 kb *Hin*dIII band and the other two have *Hin*dIII bands of equal intensity indicating that the father was heterozygous for the presence of a third *AGP* gene and the mother was homozygous for the more common two *AGP* gene arrangement. Family M had the same pattern of inheritance for the *Taq*I RFLP.

A complex polymorphism was detected in human genomic DNA digested with *Pvu*II (Figure 2c). A total of four bands of different intensities were detected. Alleles P1 and P2 were defined by 1.8 kb and 1.6 kb bands respectively and invariant bands were observed at 1.46 kb, 1.38 kb and 0.69 kb. Hybridization experiments (data not shown) indicated that the polymorphic *Pvu*II site was present in fragments containing exon 6 of either *AGP* gene. Because of the duplicated and sometimes triplicated genes in the *AGP* gene locus it was not possible to determine exactly which *AGP* gene contained the polymorphic allele/s or what the allelic distribution was in a given individual. In the case of equal intensities of bands P1 and P2 could be the consequence of several possible arrangements of the P1 and P2 alleles *e.g.* P1 could arise from *AGP1* and P2 from

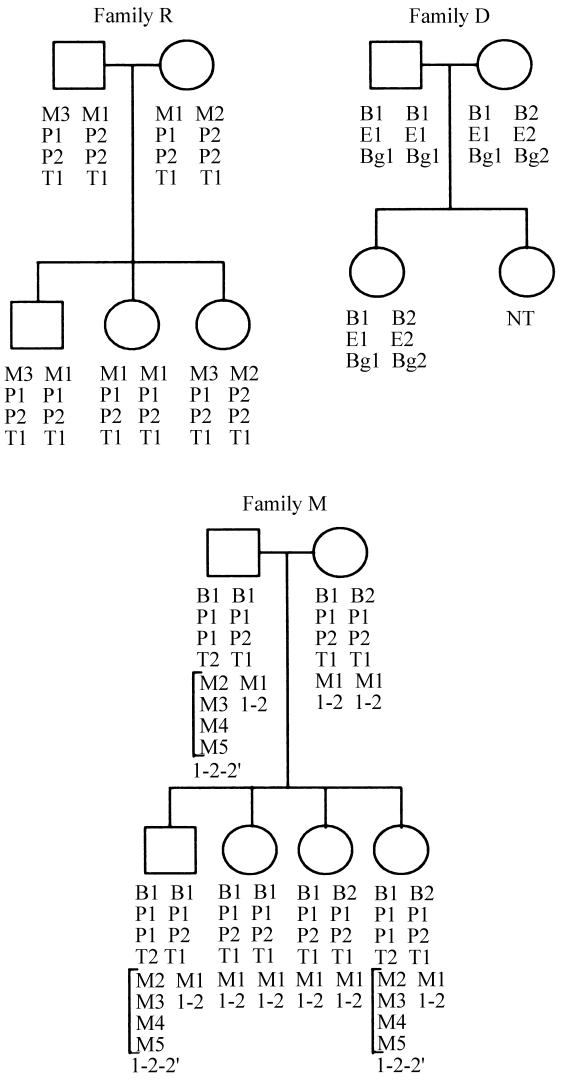

Figure 3 - Segregation of RFLP haplotypes in three nuclear families. Polymorphic fragment abbreviations: B, *Bam*HI; E, *Eco*RI; Bg, *Bgl*II, T, *Taq*I; P, *Pvu*II; M, *Msp*I; 1-2, two *AGP* genes; 1-2-2', three *AGP* gene array. The order of the haplotypes is not intended to represent their relationship within the *AGP* gene cluster.

AGP2 on one chromosome (or vice versa). The same number of P1 and P2 alleles will be present on the other chromosome but the arrangement could either be reversed or the same. Five phenotypic classes were therefore assigned based on the intensity of each allele (P1, P2) relative to each other on an autoradiogram (Figure 2c).

1. 4P1	P1 present, P2 absent
2. 3P1 1P2	P1 band more intense than P2
3. 2P1 2P2	P1 and P2 bands equally intense
4. 1P1 3P2	P2 band more intense than P1
5. 4P2	P2 present, P1 absent

Co-dominant segregation was observed in Families R and M (Figure 3). In Family R, the father was 2P1 2P2 and the mother P1 3P2. One child was 2P1 2P2 indicating that a P1 and a P2 allele are co-segregating in the father and two P2 alleles are co-segregating in the mother. The two other children in this family were 2P1 2P2. Family M demonstrated Mendelian inheritance of a different phenotypic class where the father was 3P1 P2 and the mother was 2P1 2P2. The phenotypic class of the offspring indicated that a P1 and a P2 allele must be co-segregating in the mother whilst the father is P1/P1 on one chromosome and P1/P2 on the other.

Digestion of human genomic DNA with *Msp*I and hybridization to the α_1-AGP cDNA probe also resulted in a complex band pattern. The variant fragments were designated M1 to M5 in order of decreasing size: 4.4 kb (M1), 4.3 kb (M2), 3.2 kb (M3), 2.9 kb (M4), and 2.8 kb (M5). The polymorphic fragments all hybridized to an exon-6 specific probe. A total of 47 random individuals were screened and several different arrangements of the variant *Msp*I fragments were observed. Some representative combinations are presented in Figure 2d. The various alleles that these fragments represent were not examined for deviation from Hardy-Weinberg equilibrium since extensive family studies would be necessary in order to determine the correct number of alleles in a particular individual.

Table I - Genotype and Allele Frequencies of the α_1-acid Glycoprotein RFLPs.

Restriction site	Genotype			N	Allele frequency		χ^2	p	d.f.
*Bam*HI	B1B1	B1B2	B2B2		B1	B2			
	72	21	1	94	0.88	0.12	0.15	0.7	1
*Eco*RI	E1E1	E1E2	E2E2		E1	E2			
	68	15	2	85	0.89	0.11	1.05	0.3	1
*Bgl*II	Bg1Bg1	Bg1Bg2	Bg2Bg2		Bg1	Bg2			
	61	13	1	75	0.90	0.10	0.1	0.7	1
*Hin*dIII	1-2	1-2-2'							
	58	7		65	0.94*	0.06*	-	-	-
*Taq*I	T1T1	T1T2	T2T2		T1	T2			
	78	9	0	87	0.95	0.05	0.26	0.6	1
*Pvu*II	4P1	3P1P2	2P12P2	P13P2	4P2		P1	P2	
	0	16	39	17	4	76	0.47#	0.53#	9.3

χ^2 tests the goodness of fit to Hardy-Weinberg expectations. (d.f. = degrees of freedom). *Inferred allele frequencies under the assumption of Hardy-Weinberg equilibrium. #Allele frequencies under the assumption that the frequency at each loci is the same.

However, a segregation pattern consistent with Mendelian inheritance was observed in Families R and M (Figure 3). In Family R each allele were interpreted as being alleles at a separate locus. The father had bands M1, M3 and the mother had bands M1 and M2. These segregated independently in each parent to give siblings with allele distributions of M1 M3, M1 M1, and M3 M2. The pattern of inheritance in Family M was more complex. Fragment M1 in the father segregated independently from bands M2, M3, M4, and M5. The mother was homozygous for M1. Two of the siblings have inherited an M1 band from either parent. The remaining two siblings derived an M1 band from their mother and the M2, M3, M4 and M5 bands from their father. The bands appeared to be transmitted as a single allele although they most likely represented multiple closely linked sites on a single chromosome. Interestingly, the M2, M3, M4, M5 band arrangement had the same pattern of inheritance as the *Hin*dIII and *Taq*I polymorphisms in this family.

The allele frequencies for the *Bam*HI, *Eco*RI, *Bgl*II, *Taq*I, *Pvu*II RFLPs and occurrence of an extra *AGP* gene in a sample of Caucasians are given in Table I. The observed frequency distribution of genotypes for the *Bam*HI, *Eco*RI, *Bgl*II and *Taq*I RFLPs did not differ significantly from those expected on the basis of a Hardy-Weinberg equilibrium (Table I). As mentioned above, tests were carried out using the standard χ^2 statistic. As each RFLP was concordant with Hardy-Weinberg equilibrium using that statistic the more conservative Fisher's exact test was unnecessary (despite some small expected counts).

The distribution of genotypes for the *Pvu*II RFLP was examined for deviation from "Hardy-Weinberg equilibrium" under the assumption that the probability of the occurrence of similar alleles at each of the polymorphic sites is identical. In this case the null hypothesis of Hardy-Weinberg equilibrium (Table I) amongst the alleles at the two loci is rejected on the basis of a χ^2 statistic using a standard significance level of 0.05. It must be remembered that this disequilibrium could be due to one or more causes. Firstly, one or more of the loci separately may be out of Hardy-Weinberg equilibrium, or secondly, that the probability of the occurrence of similar alleles at each of the polymorphic sites is not identical, or thirdly, that the two loci are in linkage disequilibrium. The experimental data does not allow reasonable discrimination between these alternatives since the two loci cannot be distinguished.

Analysis of linkage disequilibrium between polymorphic sites

The distributions of the alleles of the polymorphic loci were analyzed for linkage disequilibrium. The null hypothesis that each of the polymorphic sites were in linkage equilibrium was tested by both the χ^2 statistic and using Fisher's exact test on 2x2 contingency tables. The 2x2 tables were constructed from the 3x3 tables given in Table II by simply summing the second and third columns and rows respectively. For example, the *Bam*HI and *Bgl*II 2x2 contingency table had entries 40 and 1 (being the sum of 1 and 0) in the first row, and 1 (being the sum of 1 and 0) and 10 (being the sum of 9, 0, 1 and 0) in its second row. This 2x2 table then gives the counts of the combinations of the absence and presence of the less-common allele of the two RFLPs in its entries. This approach was necessary because of the very low expected counts in the third columns and rows, due to the small frequencies of the homozygotes of the less common allele. P-values from the χ^2 statistic with 1 degree of freedom and Fisher's exact test are both given. In all cases the two tests lead to the same conclusions. This is because when linkage equilibrium is observed using the χ^2 statistic Fisher's exact test, being a more conservative test, naturally leads to the same conclusion, and where linkage disequilibrium is inferred the results are so extreme, with p-values of less than 10^{-6} that no ambiguity occurs. Linkage disequilibrium was observed between the following RFLPs: *Bam*HI and *Eco*RI*; Bam*HI and *Bgl*II*; Bgl*I and *Eco*RI (Table II).

In order to test for linkage disequilibrium between the RFLP genotypes and the presence of multiple *AGP* genes (as detected by a 6.9 kb *Hin*dIII band that was more intense relative to the 4.5 kb *Hin*dIII band) 2x2 contingency tables were constructed from the 3x2 tables given in Table III in an analogous fashion to that described above. These 2x2 tables were then tested for association using, as above, the χ^2 statistic and Fisher's exact test. Linkage disequilibrium was observed between the presence of the relatively intense 6.9 kb *Hin*dIII band (1-2-2') and hence extra *AGP* gene(s), and the *Taq*I RFLP (Table III).

Discussion

Since the initial studies (of Johnson *et al.* 1969) there have been many reports of genetic variation at the human α_1-acid glycoprotein or ORM locus. In this study we have investigated DNA variation in and around the human *AGP* genes in a Caucasian population using RFLP analysis. RFLPs were detected with the use of restriction enzymes *Bam*HI, *Eco*RI, *Bgl*II, *Pvu*II, *Hin*dIII, *Msp*I, and *Taq*I.

RFLPs detected by enzymes *Bam*HI, *Eco*RI, *Bgl*II, were located at least 11 kb upstream of the *AGP* gene cluster and were in linkage disequilibrium with each other. This group of polymorphic loci did not deviate from a random association with the *Taq*I RFLP that was located within the *AGP* gene cluster. Two complex polymorphisms were detected within non-coding regions of the *AGP* gene cluster using *Pvu*II and *Msp*I. Interestingly, the *Pvu*II, *Msp*I, and *Taq*I polymorphisms could be detected with an exon 6-specific probe indicating a higher degree of recombination in this region of the *AGP* genes.

Previous studies have suggested that the two *AGP* gene array seen in the majority of individuals arose as a

Table II - Tables of the distribution of α_1-*AGP* RFLP genotypes with values of the χ^2 statistic and its p-value p_χ and the p-value from Fisher's exact test p_E based on analysis of combined 2x2 contingency tables.

*Bam*HI and *Bgl*II

*Bam*HI	*Bgl*II Bg1Bg1	Bg1Bg2	Bg2Bg2
B1B1	40	1	0
B1B2	1	9	0
B2B2	0	0	1

$\chi^2 = 41$, p_χ, $p_E < 10^{-6}$.

*Bam*HI and *Eco*RI

*Bam*HI	*Eco*RI E1E1	E1E2	E2E2
B1B1	66	0	0
B1B2	0	15	1
B2B2	0	0	1

$\chi^2 = 83$, p_χ, $p_E < 10^{-6}$.

*Bam*HI and *Taq*I

*Bam*HI	*Taq*I T1T1	T1T2	T2T2
B1B1	60	6	0
B1B2	15	1	0
B2B2	1	0	0

$\chi^2 = 0.18$, $p_\chi = 0.7$, $p_E = 1$.

*Bgl*II and *Eco*RI

*Bgl*II	*Eco*RI E1E1	E1E2	E2E2
Bg1Bg1	36	0	0
Bg1Bg2	0	9	0
Bg2Bg2	0	0	1

$\chi^2 = 46$, p_χ, $p_E < 10^{-6}$.

*Bgl*II and *Taq*I

*Bgl*II	*Taq*I T1T1	T1T2	T2T2
Bg1Bg1	35	6	0
Bg1Bg2	9	1	0
Bg2Bg2	1	0	0

$\chi^2 = 0.23$, $p_\chi = 0.63$, $p_E = 1$.

*Eco*RI and *Taq*I

*Eco*RI	*Taq*I T1T1	T1T2	T2T2
E1E1	56	8	0
E1E2	11	1	0
E2E2	2	0	0

$\chi^2 = 0.32$, $p_\chi = 0.6$, $p_E = 0.98$.

The results are expressed as number of observed subjects.

Table III - Contingency tables of the distribution of α_1-*AGP* RFLP genotypes and presence of multiple *AGP* genes as detected by digestion with *Hin*dIII.

*Bam*HI

*Bam*HI	*Hin*dIII 1-2	1-2-2'
B1B1	43	7
B1B2	12	0
B2B2	1	0

$\chi^2 = 2$, $p_\chi = 0.15$, $p_E = 0.36$.

*Bgl*II

*Bgl*II	*Hin*dIII 1-2	1-2-2'
Bg1Bg1	24	6
Bg1Bg2	6	0
Bg2Bg2	1	0

$\chi^2 = 1.7$, $p_\chi = 0.2$, $p_E = 0.5$.

*Eco*RI

*Eco*RI	*Hin*dIII 1-2	1-2-2'
E1E1	42	7
E1E2	11	0
E2E2	1	0

$\chi^2 = 1.9$, $p_\chi = 0.16$, $p_E = 0.39$.

*Taq*I

*Taq*I	*Hin*dIII 1-2	1-2-2'
T1T1	50	0
T1T2	1	7
T2T2	0	0

$\chi^2 = 50$, p_χ, $p_E < 10^{-6}$.

The results are expressed as number of observed subjects.
1-2 indicates 6.9 kb (*AGP2)* and 4.6 kb (*AGP1) Hin*dIII bands of equal intensity.
1-2-2' indicates 6.9 kb **(***AGP2)* band more intense than the 4.6 kb (*AGP1) Hin*dIII band.

consequence of gene duplication subsequent to the divergence of humans from rodents (Merritt *et al.* 1990). Individuals containing three *AGP* genes have been reported (Dente *et al*. 1987; Merritt and Board 1988; Nakamura *et al.* 2000). These three gene arrays (AGP1-AGP2-AGP2 or AGP1-AGP1-AGP2) represent polymorphisms in the populations studied and are the result of further crossover events that must have occurred relatively recently since there were no changes in the duplicated genes studied. In the Caucasian population studied here linkage disequilibrium was observed between the presence of an intense 6.9 kb *Hin*dIII fragment (1-2-2' and hence multiple *AGP1* or *AGP2* genes) and the *Taq*I RFLP. The simplest explanation for the origin of the *Taq*I polymorphism would be a point mutation that caused the loss of the *Taq*I site in the region between the *AGP1* and *AGP2* genes. However, if one considers that unequal crossing over events generated three member *AGP* gene arrays (Dente *et al.* 1987; Merritt and Board 1988; Merritt *et al.* 1990; Nakamura *et al.* 2000) it is possible that a crossover event could cause the loss of a *Taq*I site. In this study individuals who may be homozygous for the presence of a third *AGP* gene would be indistinguishable from heterozygotes since increased intensity of the 6.9 kb band relative to the 4.5 kb was used as the basis for scoring. Interestingly, however, Family M, (Figure 3.) was informative for both the *Taq*I polymorphism and the presence of an extra *AGP* gene and those individuals that were heterozygous for the *Taq*I polymorphism were also heterozygous for an extra *AGP* gene. Furthermore, in the population studied there were no T2T2 individuals suggesting that there were no individuals homozygous for an extra *AGP* gene. Further sequence analysis of the *AGP* locus from the individuals studied would be required to confirm the genetic basis for observed linkage between the *Taq*I polymorphism and the presence of multiple *AGP* genes and to determine if the particular duplicated gene was *AGP1* or *AGP2*.

The HSA orosomucoid polymorphisms (Yuasa *et al.* 1993; Yuasa *et al.* 1997) have been widely studied in a range of populations. The results presented in this survey provide evidence for further variation at the *AGP* gene locus and the polymorphisms described may be potentially useful as genetic markers in a variety of forensic, linkage and population studies.

References

Board PG, Jones IM and Bentley AK (1986) Molecular cloning and nucleotide sequence of human α_1-acid glycoprotein cDNA. Gene 44:127-131.

Dente L, Ciliberto G and Cortese R (1985) Structure of the human α_1-acid glycoprotein gene: Sequence homology with other acute phase proteins. Nucl. Acid Res. 13:3941-3952.

Dente L, Pizza MG, Metspalu A and Cortese R (1987) Structure and expression of the genes coding for human α_1-acid glycoprotein. EMBO J 6:2289-2296.

Flower DR (1996) The lipocalin protein family: structure and function. Biochem J 318:1-14.

Grunebaum L, Casenave J-P, Camerino G, Kloepfer C, Mandel J-L, Tolstoshev P, Jaye M, De la Salle H and Lecocq J-P (1984) Carrier detection of hemophilia B by using a restriction site polymorphism associated with the coagulation factor IX gene. J Clin Invest 73:1491-1495.

Johnson AM, Schmid K and Alper CA (1969) Inheritance of human α_1-acid glycoprotein (orosomucoid) variants. J Clin Invest 48:2293-2299.

Merritt CM and Board PG (1988) Structure and characterization of a duplicated human α_1-acid glycoprotein gene. Gene 66:97-106.

Merritt CM, Easteal S and Board PG (1990) Evolution of human α_1-acid glycoprotein genes and surrounding Alu repeats. Genomics 6:659-665.

Nakamura H, Yuasa I, Umetsu K, Nakagawa M, Nanba E and Kimura K (2000) The rearrangement of the human α_1-acid glycoprotein/Orosomucoid gene: Evidence for tandemly triplicated genes consisting of two AGP1 and one AGP2. Biochem Biophys.Res Commun 276:779-784.

Reed KC and Mann DA (1985) Rapid transfer of DNA from agarose gels to nylon membranes. Nucl Acids Res 13:7202-7221.

Webb GC, Earle EA, Merritt CM and Board PG (1987) Localization of human α_1-acid glycoprotein genes to 9q31-34.1. Cytogenet Cell Genet 47:18-21.

Weir BS (1996) Genetic Data Analysis II, 2nd ed. Sinauer Associates, Sunderland, MA.

Yuasa I, Weidinger S, Umetsu K, Suenaga K, Ishimoto G, Eap BC, Duche J-C and Baumann P (1993) Orosomucoid System: 17 additional orosomucoid variants and proposal for a new nomenclature. Vox Sang 64:47-55.

Yuasa I, Umetsu K, Vogt U, Nakamura H, Nanba E, Tamaki N and Irizawa Y (1997). Human orosomucoid polymophis: molecular basis of the three common ORM1 alleles ORM1*F1, ORM1*F2, and ORM1*S. Hum Genet 99:393-398.