Human immunoglobulin heavy chain \( \alpha_2 \) gene allotype determination by restriction fragment length polymorphism

Marie-Paule Lefranc* and T.H.Rabbitts

MRC Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH, UK

Received 23 December 1983; Accepted 11 January 1984

ABSTRACT

The human immunoglobulin heavy chain \( \alpha_2 \) genes have two allelic forms or allotypes called \( \text{A2m}(1) \) and \( \text{A2m}(2) \). Previously, these allotypic markers have only been distinguishable by serology. Studies of the \( \alpha_2 \) genes, however, show that it is possible to differentiate between the allotypes by restriction enzyme site polymorphisms, both in the protein coding regions and in flanking regions. These polymorphic sites have been used to determine the \( \alpha_2 \) allotypes of several human DNAs.

INTRODUCTION

The IgA class of human antibodies is defined by the presence of the \( \alpha \) heavy (H) chain and is the second most abundant of the five classes of human immunoglobulin. It is the major immunoglobulin in secretory fluids where it occurs as a dimer containing the J-chain and the secretory component (1). There are two subclasses (or isotypes) of human IgA (IgA1, which carries the \( \alpha_1 \), and IgA2, the \( \alpha_2 \) heavy chain) which differ both chemically and antigenically (2,3). Human serum IgA consists of about 90% IgA1 and 10% IgA2 and in secretory fluids IgA2 is predominant (4). The complete amino acid sequence of the constant region of the \( \alpha_1 \) and \( \alpha_2 \) heavy chains (5-8) as well as the corresponding nucleotide sequence have been reported (9).

Genetic polymorphism has been described for the IgA2 subclass in man. There are two allotypes or genetic markers, designated \( \text{A2m}(1) \) and \( \text{A2m}(2) \) (10-12). The frequency of the \( \text{A2m}(2) \) allotype is high in Africa and Asia (0.5 to 0.8) and very low in Europe (0.01 to 0.02) (13-15). An isoallotype, designated \( \text{nA2m}(2) \) has also been identified serologically (16); this determinant is present on all \( \alpha_1 \) chains and on the \( \alpha_2 \) chains of the \( \text{A2m}(1) \) allotype. It is due to the identity in sequence of the last domain (C\( _H^3 \)) of the \( \alpha_1 \) chain and the \( \text{A2m}(1) \) \( \alpha_2 \) chain (8).

The allotypes of the \( \alpha_2 \) H-chain are important both in population studies of gene frequency (14,17) and also in family studies where inherited patterns...
of a\textsubscript{2} genes are relevant (18). Furthermore, the a gene allotypic markers will be associated with, as yet unknown, genes downstream of the human H-chain locus on chromosome 14. Until now it has only been possible to determine the a\textsubscript{2} allotype by serology. We now describe two types of restriction enzyme fragment length polymorphism (RFLP) in the a\textsubscript{2} genes which allow the allotype to be determined by the Southern filter hybridisation method using a gene probes. Utilising this observation we confirmed the a allotypes derived from serology of several people as well as determining that of four randomly selected examples where no serological data was available.

**METHODS**

Construction and screening of a \(\lambda\) phage library from TOUII-5 DNA

TOUII-5 is a Tunisian shown by testing of IgA allotypic markers, using the haemagglutination inhibition method, to be homozygous for the A2m(2) allotype (19,20). TOUII-5 DNA was prepared from total peripheral white blood cells. The DNA was partially digested with MboI and size-fractionated on a sucrose gradient. Restriction fragments of 15-20 kb were ligated into BamHI-digested \(\lambda\)2004 (21) and packaged in vitro. Recombinant phages were screened by the method of Benton and Davis (22) according to the procedure previously described (23). The probes, radiactively labelled with \(a^{-32}\)P-dATP and \(a^{-32}\)P-dCTP by nick-translation (24) were for the a gene, the clone \(a\)XP8, which contains a 600 bp fragment of the a\textsubscript{2} gene extending from an XhoI site near the start of the second exon to a PstI site in the middle of the third exon (25) and for the \(\epsilon\) gene, the clone \(\epsilon\)1.2BP25, a 2.1 kb BamHI-PstI fragment, containing the entire human \(\epsilon\) coding region (26). To localise the switch region (\(S_a\)) of the a gene an \(S_\mu\) probe designated C76R51 (27) was used.

**DNA analysis**

DNA from various sources was prepared either from total peripheral blood cells (BAR IX-3, BAR X-1, BAR15, B.S.) (28,29) or from cell lines JI, LY67 (30) and Daudi (31) and digested to completion either with EcoRI plus HindIII enzymes or with EcoRI plus PstI enzymes. Fragments were separated on 0.8% agarose gels for the EcoRI-HindIII digests or on 1.4% agarose gels for the EcoRI-PstI digests and transferred to cellulose nitrate filters (32). The hybridisation was carried out at 65°C using 2 x SSC, 0.1% SDS, 0.2% Ficoll400, 0.2% bovine serum albumin, 0.2% polyvinylpyrrolidone and 50 µg ml\(^{-1}\) sonicated, denatured, salmon sperm DNA, followed by washes in 0.1 x SSC, 0.1% SDS at 65°C before autoradiography at -70°C with prefogged X-ray film (33). The probes were nick-translated to a specific activity of approximately 10\(^8\) cpm per µg.
RESULTS AND DISCUSSION

Cloning the $\alpha_2$A2m(2) gene

An example of an $\alpha_2$A2m(2) gene was prepared from a $\lambda$ phage library containing genomic DNA from an individual (designated TOUII-5) who is homozygous for this gene (19,20). One clone, containing a 19.5 kb piece of genomic DNA (XTOUca), was isolated and selected for characterisation. A restriction map of this clone is shown in Figure 1. The genomic segment in $\lambda$TOUca was found to contain both $C_e$ and $C_\alpha$ genes, as shown by hybridisation to probes for the respective $C_H$-regions, separated by about 11 kb of DNA. The $S_\mu$ probe hybridised to the region near the $C_\alpha$, as indicated in the figure, which represents the $S_\alpha$ region.

Since previous studies have shown the absence of $\alpha_1$ and $\psi_\alpha$ genes from the genome of TOUII-5 (19,20) the $C_H$ genes in $\lambda$TOUca must be the $C_{a2}$ and the active $C_e$ [which are known to be closely linked in human DNA (25,34)]. Furthermore, another clone $\lambda$TOUta has been sequenced and shown to contain the $\alpha_2$A2m(2) $C_H$ gene (9) as expected since the serology had shown TOU to be homozygous for this gene. Thus the $\alpha_2$ gene in $\lambda$TOUca must also be the $\alpha_2$A2m(2) allotypic gene; this is confirmed by restriction enzyme polymorphism described below (polymorphic sites for EcoRI and HindIII are marked with asterisks in Figure 1).

Restriction fragment length polymorphism of the human $\alpha$ genes

We previously showed that the restriction enzyme PstI yielded different fragments containing $\alpha_1$ (1.2 kb) and $\alpha_2$ (2 kb) genes (25) when probed with the clone $\alpha_2$XP8 (which encompassed most of the $C_H2$ exon, the beginning of the $C_H3$

![Fig. 1. Restriction map of the $\lambda$TOU insert. The location of the four exons (protein coding regions) of the $C_e$ gene and of the three exons of the $C_\alpha2$ gene are shown by open boxes. The rectangles below the insert indicate the fragments used as $\epsilon$ and $\alpha$ probes. The rectangle in dotted lines shows the region of the insert hybridising with the $S$ probe (see text). The two restriction sites surmounted by an asterisk are characteristic of the nucleotide sequence of the $\alpha_2$A2m(2) $C_H$ gene and of its 5' flanking region. B = BamHI, H = HindIII, S = SacI, R = EcoRI, X = XhoI.](image-url)
Restriction site polymorphisms associated with amino acid substitutions in \( \alpha \) genes.

**TABLE 1A**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Amino acid and residue number</th>
<th>Codon</th>
<th>Sequence</th>
<th>Enzyme restriction site</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_2 )</td>
<td>Cys(^{133})</td>
<td>TGC</td>
<td>CfTGC(\sim)G</td>
<td>PstI</td>
</tr>
<tr>
<td>( \alpha_2 )(^{A2m(1)} )</td>
<td>Asp(^{133})</td>
<td>GAC</td>
<td>C</td>
<td>GAC</td>
</tr>
</tbody>
</table>

**TABLE 1B**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Amino acid and residue number</th>
<th>Codon</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_1 )</td>
<td>Pro(^{212})</td>
<td>CCC</td>
<td>GAATCC</td>
</tr>
<tr>
<td>( \alpha_2 )(^{A2m(1)} )</td>
<td>Ser(^{212})</td>
<td>TCC</td>
<td>GAATCC</td>
</tr>
<tr>
<td>( \alpha_2 )(^{A2m(2)} )</td>
<td>Ser(^{212})</td>
<td>TCC</td>
<td>GAATCC</td>
</tr>
</tbody>
</table>

The difference in PstI fragments is due to two base differences at codon 133 of the \( \alpha_H \) domain (numbering according to the protein sequence \( \beta_1 \) chain (5)) resulting in a PstI site in the \( \alpha_1 \) sequence (Table 1A). A further PstI polymorphic restriction site is found 131 basepairs upstream of the \( \alpha_H \) exon of \( \alpha_1 \) (Fig. 2A). Figure 2 and Table 1B also illustrate restriction site polymorphisms which exist between the two allotypic genes for \( \alpha_2 \). One of these, an EcoRI site, occurs in the \( \alpha_H \) exon of \( \alpha_2 \)\(^{A2m(2)} \) resulting from a C-T difference. This base difference gives rise to one of the two substitutions which occur only in the \( A2m(2) \) \( \alpha_2 \) polypeptide chain prior to the hinge segment. A further significant polymorphism is the occurrence of a HindIII site in the \( A2m(2) \) \( \alpha_2 \) about 4 kb upstream of the \( \alpha_2 \) sequences (indicated by an asterisk in Fig. 1).

Figure 2 compares the relevant restriction enzyme sites of the three \( \alpha \) genes plus the relative location of \( \epsilon \) and \( \alpha \) genes [the \( \epsilon_1 \) gene is upstream of \( \alpha_2 \) (25)]. The differences in restriction fragments which exist between the \( \alpha_1 \) and the two \( \alpha_2 \) allotypic genes can be deduced from this figure. All three genes possess PstI sites in the \( \alpha_3 \) exon (codons 484 and 485) whilst the \( \alpha_1 \) gene has a PstI site in \( \alpha_1 \) giving a 1.2 kb fragment distinct from the 2 kb
Fig. 2. Comparison of the restriction maps of the Ca_1, Ca_2 A2m(1) and Ca_2 A2m(2) genes and of their flanking region (A, B, C, respectively, in the figure). The only enzymes and fragment sizes indicated are those used for the study of the polymorphism. The first line of A, B and C shows the restriction fragments resulting from the EcoRI-HindIII digests, whereas the second line illustrates those observed in the EcoRI-PstI digests. The four exons of the C_e gene and the three exons of the different C_a genes are indicated by open boxes, the pseudogene C_{αc}1 is shown by two rectangles in dotted lines.

H = HindIII, P = PstI, R = EcoRI.

Fragment of both α_2 allotypic genes detectable with the α probe. The two α_2 genes can, however, be distinguished by a combination of EcoRI plus PstI digestion, since the EcoRI site only exists in the α_2 A2m(2) gene. EcoRI-PstI digestion therefore yields a unique 0.9 kb hybridising α_2 A2m(2) whereas the usual 1.2 kb and 2 kb fragments are obtained for the α_1 and α_2 A2m(1) genes, respectively. The presence of the EcoRI site allows a further distinction between the α_2 allotypes since combined EcoRI plus HindIII digestion yields a unique 3.8 kb band representing the α_2 A2m(2) gene whilst the α_1 and α_2 A2m(1) genes migrate at about 18 kb.
Fig. 3. Southern filter hybridisation of human α probe to genomic DNA from various sources. The DNA digests were: A, EcoRI-PstI; B, PstI; C, EcoRI-HindIII. B.S. is DNA from peripheral blood of a hypogammaglobulinaemic patient. Size estimations were made using λ cut with HindIII.

**Distinguishing α genes by filter hybridisation**

The use of these RFLPs to determine the α2 allotypes of human individuals is illustrated in Figure 3. Here we have examined the hybridisation of DNA from three individuals whose α allotype is known [BAR IX-3, BAR X-1 and BAR15 (28,29)] and from other randomly selected DNAs of unknown haplotype. Three patterns of hybridisation are apparent characterising DNA of individuals who
are either homozygous for A2m(1) or A2m(2) allotypes or heterozygous A2m(1,2). Figure 3 shows that EcoRI-PstI-digested DNA from BAR IX-3 (serologically defined as homozygous A2m(1)) as well as JI, Daudi and B.S. yield 1.2 kb and 2 kb bands corresponding to the PstI fragments of the $a_1$ and $a_2$ A2m(1) genes, respectively. The absence of the 0.9 kb EcoRI-PstI fragment indicates that, like BAR IX-3, DNAs from JI, Daudi and B.S. are homozygous for A2m(1). The DNAs from BAR15 (serologically defined as homozygous A2m(2)) as well as from LY67 display the $a_1$ gene fragment (1.2 kb) plus the 0.9 kb EcoRI-PstI fragment representative of the A2m(2) gene; the absence of the 2 kb band shows that these DNAs are homozygous A2m(2). The simultaneous presence of the 2 kb $a_2$ A2m(1) fragment and of 0.9 kb $a_2$ A2m(2) in BAR X-1 shows that this individual is heterozygous for Am allotypes. The hybridisation pattern obtained with PstI digestion alone for Daudi DNA is illustrated for comparison in Figure 3B.

The A2m allotypes of the individuals deduced from the hybridisation patterns are written underneath Figure 3A and are consistent with the known allotypes derived by serological studies on the BAR family (29). The result is further confirmed in Figure 3C which shows the hybridisation to EcoRI-HindIII-digested DNA; here the three individuals with the $a_2$ A2m(2) gene possess a band of 3.8 kb which represents this gene. In the other cases the hybridisation is restricted to the largest restriction fragment containing $a_1$ and $a_2$ A2m(1).

Implications

The practical advantages of conducting genetic studies by directly studying the genomes of individuals has been dramatically illustrated recently by the determination of restriction enzyme polymorphism linkage to disease loci such as that responsible for Huntingdon's Chorea (35). The importance, therefore, of a RFLP map of the human genome is clear. The experiments in this paper demonstrate a simple method for the detection of polymorphic sites associated with the human $a_2$ genes. These should be useful for population studies on the distribution of these alleles and may ultimately be useful as genetic markers for as yet undetermined loci on chromosome 14 of man. Further, the study of RFLP in the $a_2$ gene is generally applicable in laboratories, whereas the serological determination requires very specific reagents which cannot be generally available. It is also possible by this procedure to determine the $a$ allotypes of individuals even when no serum is available for study (e.g. Daudi).

Acknowledgements

We thank G. Lefranc for providing us with the blood samples of BAR IX-3,
BAR X-1 and BAR15, Erna van Loghem and Gerda de Lange for the determination of the A2m allotypes of these samples. This work was supported in part by an EMBO Fellowship (M.P.L.).

*Permanent address: Laboratoire d’Immunogénétique, Université des Sciences et Techniques du Languedoc, 34060 Montpellier Cedex, France

REFERENCES
