The Human Apolipoprotein(a)/Plasminogen Gene Cluster Contains a Novel Homologue Transcribed in Liver

Christopher D. Byrne, Karen Schwartz, Kelli Meer, Jan-Fang Cheng, Richard M. Lawn

Abstract Lipoprotein(a) is an atherogenic lipoprotein whose function and plasma concentration reflect the structure and regulation of the apolipoprotein(a) gene. Apolipoprotein(a) is a close homologue of plasminogen, and their genes are tightly linked on chromosome 6. To further characterize these genes, we analyzed overlapping human genomic yeast artificial chromosome clones, which revealed a cluster of four highly homologous genes encoding apolipoprotein(a), plasminogen, and two apolipoprotein(a)-related genes (rg) or pseudogenes. Hybridization analysis and reverse transcriptase polymerase chain reaction showed that one of these novel genes, designated apolipoprotein(a)rg-C, has a domain structure similar to apolipoprotein(a) and is transcribed in human liver. Three additional homologues designated as plasminogen-related genes are shown to be unlinked to this gene cluster and reside on chromosomes 2 and 4. (Arterioscler Thromb. 1994;14:534-541.)

Key Words • lipoprotein(a) • apolipoprotein(a) • apolipoprotein genes • plasminogen-related genes

Elevated plasma levels of lipoprotein(a) (Lp[a]) represent a major inherited risk for atherosclerosis (for reviews, see References 1 through 3). Lp(a) resembles low-density lipoprotein in lipid content and the presence of apolipoprotein (apo) B-100, but it is distinguished by the presence of apo(a). Apo(a) is a close homologue of plasminogen, with a high number of repeated kringle domains. Plasminogen contains five kringle pairs, whereas alleles of apo(a) contain between 10 and 40 copies of a kringle that resembles kringle 4 of plasminogen. Apo(a) also contains a single homologue of plasminogen kringle 5 and an inactive protease-like domain. The high degree of homology between human apo(a) and plasminogen (between 70% and 100% for various domains) suggests that the genes arose from a common precursor during the evolution of primates. The nearly identical 5' regions of both genes suggest that gene conversion events followed initial gene duplication. Recently, hybridization and cloning studies have shown that six or more genes and/or pseudogenes are present in the human genome that contain over 90% DNA sequence homology in the region surrounding the translation initiation site. All six of these gene fragments are sufficiently similar to retain the ability to promote transcription in vitro. Although previously only apo(a) and plasminogen were known to correspond to a common precursor, the recently described apo(a)rg-C gene that contains several kringle and a protease-like domain.

Methods

Preparation and Analysis of YAC Clones

Polymerase chain reaction (PCR) protocols followed those described by Kawasaki. Electrophoresis, blot hybridization, and other standard molecular biological procedures essentially followed the protocols of Maniatis et al. YAC clones were identified from the library of the Centre d'Etude du Polymorphisme Humain, Paris, France. DNA suitable for analysis by Southern blotting and PCR was prepared as described, and YAC DNA plugs for pulsed-field gel electrophoresis (PFGE) were prepared according to the method of Game et al. The library was screened with a PCR-generated 483-bp band designated pcrA (primers 6 and 7), which hybridizes to the 5' region of apo(a), plasminogen, and their close homologues. The pcrA fragment was also used to characterize the clones; other radiolabeled probes used for characterization included a cDNA fragment containing only kringle 4 sequences, the 3' end of the protease domain, and a cDNA fragment containing the first three kringle of plasminogen. The gene cluster was mapped by complete and partial digestion using PFGE (CHEF; Bio-Rad) and hybridization to the above specific probes as well as left- and right-arm YAC vector probes generated by Pvu II–BamHI digestion of pBR322.

Specific Oligonucleotide Hybridization

Characterization of the plasminogen gene– and plasminogen-related gene (PRG)–containing clones was undertaken by

Received November 17, 1993; accepted January 4, 1994.

From the Falk Cardiovascular Research Center (C.D.B., K.S., K.M., R.M.L.), Stanford University School of Medicine, Stanford, Calif, and the Lawrence Berkeley Laboratory (J.-F.C), University of California, Berkeley.

Correspondence to Richard M. Lawn, PhD, Falk Cardiovascular Research Center, Stanford University School of Medicine, 300 Pasteur Dr, Stanford, CA 94305.
amplification, blot hybridization, gel isolation, and thermal cycle sequencing (NE Biolabs) of a 483-bp PCR product flanking the ATG (primers 6 and 7). Specific oligonucleotides were generated for use as end-labeled probes: plasminogen, PRG-A, PRG-B, and PRG-D (our data).

Identification and Characterization of Apo(a)rg-C-Containing PI Clone

A bacteriophage P1 clone containing part of the apo(a)rg-C gene was identified after screening a P1 library by PCR (primers 37 and 38). One positive clone was obtained, and the specific 395-bp PCR product was sequenced to confirm its identity to the corresponding region of the apo(a)rg-C gene. The P1 clone was further characterized by standard electrophoresis and PFGE and blot hybridization with the same radiolabeled probes used for mapping the YAC clones. Selected restriction fragments were isolated and sequenced as above.

Cloning and Partial Sequencing of Apo(a)rg-C cDNA and Gene

RNA was prepared from human liver, and a 332-bp band was identified after reverse transcriptase PCR (RT-PCR) (primers 22 and 64) by hybridization with a sequence-specific apo(a)rg-C oligonucleotide (ASO 77) corresponding to the 5' untranslated region of a predicted transcript. The band was excised, cloned into a TA cloning vector (Invitrogen), and sequenced by the Sanger dideoxy method using the Sequenase kit (United States Biochemical). A 2.6-kb EcoRI fragment of the P1-C genomic clone hybridizing to the KC probe was isolated and cloned into pBluescript for sequence analysis.

Chromosomal Location of PRGs

Chromosomal assignment of the PRGs was made with a panel of 14 Chinese hamster/human cell lines that was a gift from Uta Francke, Stanford University. A 483-bp PCR product from the plasminogen and PRGs was generated with primers 6 and 7 and as described above for YAC clones. To distinguish between the homologous genes, the products were electrophoresed, blotted, and hybridized with the plasminogen, PRG-A, PRG-B, and PRG-D oligonucleotides as described above. Hybridization and wash conditions were as described in Fig 1. No corresponding PCR product was obtained from total Chinese hamster DNA. YAC DNA known to contain these genes was used for positive and negative controls.

The Table lists the relevant PCR primers, hybridization oligonucleotides, and the YAC clones code.

Results

Identification of YAC Clones

A YAC library containing human genomic DNA was screened with a PCR amplification product of 483 bp (PCR) corresponding to the homologous 5' region that surrounds the transcription start site of the plasminogen and apo(a) genes. Twenty-four YAC clones were identified, and 23 were characterized by EcoRI digestion and Southern blot hybridization. We have previously shown that the apo(a) gene contains a 4.8-kb fragment that hybridizes to this probe. Apo(a)rg-B and apo(a)rg-C contained 3.6-kb and 3.1-kb hybridizing EcoRI fragments, respectively, and plasminogen and two or more PRGs (PRG-A and PRG-B) contained 1.3-kb hybridizing EcoRI fragments. Figure 1B shows that individual YAC clones contain up to three of these fragments. These results demonstrate a close clustering of at least four of the related genes and suggest their order on the chromosome.

Thirteen of the YAC clones had only a 1.3-kb hybridizing band (Fig 1A) and hybridized with the cDNA probe corresponding to plasminogen kringle 1 through 3 but not with kringle 4 or with the protease probes (Fig 1B). This suggested that these clones do not contain the bona fide plasminogen gene and probably contain PRGs or pseudogenes (PRG-A and PRG-B), which were not linked to the cluster containing the plasminogen, apo(a), apo(a)rg-B, and apo(a)rg-C genes.

To test whether these genes contain kringle and protease domains, the filter used in Fig 1A was sequentially rehybridized with probes that detected the kringle 4-like domains of plasminogen and apo(a), the protease domain of both genes, and the kringle 1 through 3 region of plasminogen. The results are summarized in Fig 1B. With the kringle 4 probe, clones 2, 14, 19, and 20 contained the pattern of multiple hybridizing bands, including extremely intense bands at 6.0 kb that are the hallmark of the apo(a) gene. Clones 7 and 22 contained weakly hybridizing bands of 2.6, 6.5, and 7.0 kb only, while clone 14 contained these bands plus the apo(a) pattern of hybridization. apo(a)rg-B hybridized to the apo(a)rg-C gene, and YAC 14 contained these bands plus the apo(a)rg-C gene, which has several kringle 4-like domains. Clone 14 contained both apo(a) and apo(a)rg-C genes. Hybridizations indicated that the apo(a)rg-C gene also contained a protease-like domain located in a 12-kb EcoRI fragment but did not contain a kringle 1 through 3 domain. In contrast, the apo(a)rg-B gene appeared to contain no close homologues of any of these coding regions.

Map of the Apo(a)/Plasminogen Gene Cluster

A map of the cluster of the four homologous genes shown in Fig 2 was determined by standard and pulse-field blot hybridization of complete and partial digests of YAC clones that had been hybridized with probes from the left and right arms of the YAC vectors in addition to the probes used in Fig 1. The relative location and orientation of these genes are shown, as well as all Sal I restriction sites in the cluster plus the Cla I sites surrounding the apo(a)rg-C gene that were used to determine its orientation.

The apo(a) and plasminogen genes are aligned in opposite orientation with respect to transcription. This was determined by analysis of clones terminating within either the apo(a) gene (YACs 5, 6, and 8) or the plasminogen gene (YAC 14). These two sets of clones overlap in the middle and extend in opposite directions. Clones 5, 6, and 8 contain the plasminogen protease domain (9-kb EcoRI fragment) but not the apo(a) protease (11-kb EcoRI fragment), whereas clone 14 contains the protease of apo(a) but not of plasminogen. This is consistent with analysis of clone 2, which terminates within both genes and contains the 5' flanking regions of the apo(a) and plasminogen genes but lacks the protease regions of each.

The apo(a) and apo(a)rg-C genes are oriented in the same direction. This was determined by partial and complete Cla I and Sal I digests and mapping with vector arm and gene domain probes of YACs 7 and 22, which extend from within the apo(a) gene through apo(a)rg-C, and YAC 14, which contains all of both genes. Further characterization of the apo(a)rg-C gene was made possible by isolating a clone from a human
FG 1. A, Biot. Three micrograms yeast artificial chromosome (YAC) DNA was digested with EcoRI and analyzed by electrophoresis and blot hybridization to a 5' region probe (porA) in 30% formamide, 5× saline–sodium phosphate–EDTA (SSPE), and 0.1% sodium dodecyl sulfate (SDS) at 42°C and washed in 0.1% saline–sodium citrate and 0.1% SDS. B, Characterization of YAC clones using probes including kringle 4 (K4), an apolipoprotein (apo) (a) cDNA partial clone containing only kringle 4-like sequences; Pro, a 600-bp Kpn I restriction fragment containing the 3' region of the apo(a) protease cDNA; and K1-K3, a 666-bp EcoRI/Rsa I restriction fragment of plasminogen (pmg) cDNA. Oligonucleotides (oligo) of plasminogen and the plasminogen-related genes (PRG) show the assignment of the 1.3-kb EcoRI fragments hybridizing to porA. + indicates hybridization detected with indicated probe; •, K1-K3 hybridization pattern different from that of plasminogen gene; and rg, related gene. C, A 483-bp fragment amplified from YAC DNA (0.6 μg) with 0.5 μmol/L primers 6 and 7 with 1.5 mmol/L MgCl₂ and 200 μmol/L dNTPs in a buffer containing 50 mmol/L KCl and 10 mmol/L Tris(hydroxymethyl)aminomethane-HCl, pH 8.3, cycling for 30 cycles (95°C, 1.5 minutes; 37°C, 1 minute; and 65°C, 3 minutes). Polymerase chain reaction products were blot hybridized with specific 17-base oligonucleotides to part of the 5' flanking region of plasminogen, PRG-A, PRG-B, and PRG-D genes. Filters were probed with [γ-32P]ATP-labeled plasminogen, PRG-A, and PRG-B oligonucleotides (GIBCO BRL) in 5× SSPE, 5× Denhardt's solution, and 0.5% SDS at 34°C and washed in 5× SSPE and 0.1% SDS at 47°C. PRG-D hybridization and washing was at 48°C.

genomic library in bacteriophage P1 with a PCR-generated probe that only amplifies the 5' flanking region of this gene (see below).

The apo(a)rg-B is located at one end of the cluster. Failure of this region to hybridize with either kringle 1 through 3, kringle 4, or protease domain probes suggested that this gene or pseudogene represents a fragmentary homologue of only the 5' exon of the plasminogen and apo(a) genes that is unlinked to homologous coding domains.

The four genes and/or pseudogenes are tightly clustered in a region of about 250 kb. For individual genomes, the size of this cluster will vary by 100 to 200 kb, depending on allelic variation in the number of kringle domains and the resulting size of the apo(a) gene. The entire cluster is flanked by Not I restriction sites. The
PCR Primers, Hybridization Oligonucleotides, and YAC Clones Code

PCR primers

<table>
<thead>
<tr>
<th>PCR</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>GCCTGTTGAAACCTTG (coding orientation)</td>
</tr>
<tr>
<td>7</td>
<td>AGTAGAAGACACCTTC (reverse complement orientation)</td>
</tr>
<tr>
<td>22</td>
<td>CGACCTCGCTGCATAGA (reverse complement orientation)</td>
</tr>
<tr>
<td>37</td>
<td>CGACCTCAGCTGACCTCGA (reverse complement orientation)</td>
</tr>
<tr>
<td>38</td>
<td>CCCTTGATTGCTGACAT (coding orientation)</td>
</tr>
<tr>
<td>64</td>
<td>CTGACCCAGTGCAATGCTTTC (coding orientation)</td>
</tr>
</tbody>
</table>

Hybridization oligonucleotides

<table>
<thead>
<tr>
<th>PRG</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>TCAAGCTTATGTTTGAA</td>
</tr>
<tr>
<td>B</td>
<td>TCAAGGTCTGTTTGAA</td>
</tr>
<tr>
<td>D</td>
<td>TCAAGGTCATGTTTGAA</td>
</tr>
<tr>
<td>PMG</td>
<td>TCAAGGTACTGTTTGAA</td>
</tr>
<tr>
<td>ASO77</td>
<td>AGAAATAACAGAAGTAG</td>
</tr>
<tr>
<td>KC</td>
<td>AGAAGGCCCTGTCTCAGTCGGTGCTGACTTCAG</td>
</tr>
</tbody>
</table>

YAC clones code

<table>
<thead>
<tr>
<th>YAC clone</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1B1</td>
<td>1. 54G4</td>
</tr>
<tr>
<td>204B2</td>
<td>8. 219E11</td>
</tr>
<tr>
<td>347E7</td>
<td>14. 366H2</td>
</tr>
<tr>
<td>431B1</td>
<td>20. 431C2</td>
</tr>
<tr>
<td>517D10</td>
<td>23. 517D10</td>
</tr>
<tr>
<td>514E4</td>
<td>22. 514E4</td>
</tr>
<tr>
<td>179F1</td>
<td>6. 179F1</td>
</tr>
<tr>
<td>219E11</td>
<td>9. 220G4*</td>
</tr>
<tr>
<td>366H2</td>
<td>15. 376G6</td>
</tr>
<tr>
<td>380D10</td>
<td>17. 393C8</td>
</tr>
<tr>
<td>400D1</td>
<td>24. 400D1</td>
</tr>
<tr>
<td>431C2</td>
<td>22. 431C2</td>
</tr>
<tr>
<td>517D10</td>
<td>23. 517D10</td>
</tr>
</tbody>
</table>

PCR indicates polymerase chain reaction; YAC, yeast artificial chromosome; PRG, plasminogen-related gene. Laboratory clone numbers 1 through 24 correspond to the Centre d'Etude du Polymorphisme Humain, Paris, France, YAC library designations.

*For technical reasons clone 9 was not extensively characterized.

Allelic variation in the size of this Not I fragment is readily shown in pulse-field blotting with any of the probes we have used (data not shown and Reference 20).

Characterization of the PRGs

The plasminogen gene and PRGs cannot be distinguished by the size of the EcoRI fragment that hybridizes with the 5' pcrA probe, since it is 1.3 kb in each case. This entire region contains 96% DNA sequence identity with the plasminogen, PRG-A, and PRG-B genes.5,7 Thus, we synthesized 17-base oligonucleotides that allowed discrimination between them. YAC DNA was amplified by PCR with oligonucleotides that generate a 483-bp fragment from each of these genes containing the 17-base region. Seven YAC clones hybridized with only the plasminogen-specific oligonucleotide (Fig 1C). These were the same clones that generated the plasminogen gene pattern of EcoRI fragments hybridizing with the plasminogen kringle 1 through 3 probe (Fig 3). Unique patterns of hybridization were associated with clones containing plasminogen, PRG-A, and PRG-B. The pattern of hybridization of clones containing PRG-D was the same as that obtained with PRG-A. On the basis of this result alone, it was not possible to determine whether PRG-D was a distinct gene or another allelic variant of PRG-A. However, subsequent chromosome mapping (see below) found PRG-A and PRG-D to be distinct.

In summary, the analysis of YAC clones revealed a cluster of plasminogen, apo(a), apo(a)rg-B, and apo(a)rg-C genes, while two or more PRGs are not closely linked. Hybridization with coding region probes showed that PRG-A, PRG-B, and PRG-D contain sequences related to the initial domains of plasminogen but not to its kringle 4 or protease coding regions.

Chromosome Assignment of the PRGs

Studies that demonstrate the assignment of the apo(a) and plasminogen genes to chromosome region 6q26-ter have used hybridization probes that only detect kringle and protease domains of the two genes.9-11 Since our results implied that some of these coding regions are absent from the PRGs and that they are not included in any of the YAC clones that contained the apo(a) or plasminogen genes, their localization remained uncertain. To determine chromosome assignment, a panel of human/Chinese hamster somatic cell hybrids was subjected to PCR amplification and oligonucleotide hybridization as described above. Human and hamster DNA and representative distinguishing YAC clones were included in the experiments to control for the specificity of the hybridization. As expected, the plasminogen-specific oligonucleotide hybridized only
with cell lines containing human chromosome 6. However, this was not the case for the PRGs. Analysis of the hybrid cell panel showed that PRG-A was located on human chromosome 4, while PRG-B and PRG-D were located on chromosome 2 (not shown). This result, in combination with the kringle 1 through 3 hybridization of EcoRI digests of PRG-containing clones, indicated that PRG-B and PRG-D were separate genes, since clones containing both genes gave distinct patterns of hybridization with the kringle 1 through 3 probe (Fig 3). Furthermore, although PRG-A- and PRG-D-containing clones gave the same pattern of hybridization with the kringle 1 through 3 probe, they were located on different chromosomes.

Characterization of Apo(a)rg-C and Evidence for Its Transcription in Human Liver

To further characterize the apo(a)rg-C gene, a human genomic library in bacteriophage PI was screened by PCR with a pair of oligonucleotides that generate a 395-bp fragment of the 5' flanking region of this gene but none of the other homologous genes. The isolated clone, designated P1-C, contained 5' flanking and all kringle 4 domains identified in YAC clones containing the apo(a)rg-C gene. Complete and partial restriction digest and blot hybridization produced the map of this gene shown in Fig 2A. As noted from the analysis of YAC clones, the apo(a)rg-C gene contains several kringle 4-hybridizing fragments. EcoRI fragments of 7.0 kb, 6.5 kb, and 2.6 kb hybridized with the kringle 4 probe, as did five EcoRI plus Sst I double-digest fragments of the clone. These results, plus the derived restriction-site map, indicated that approximately five kringle 4-like coding domains are present. Since the P1 clone failed to hybridize with the kringle 1 through 3-like probe, apo(a)rg-C thus resembles a smaller version of an apo(a)-like gene.

In a previous sequence analysis, we found that the 5' end of apo(a)rg-C contains a region that is highly homologous (≈96%) to the first exon of both apo(a) and plasminogen genes.7 In those two genes, the first exon encodes 5' untranslated sequences and 17 amino acids of secretion signal peptide. Since the 5' flanking region of apo(a)rg-C can drive transcription of a reporter gene when transfected into cultured HepG2 cells,7 we investigated whether this gene is transcribed in human liver. cDNA was synthesized from human liver mRNA and amplified by RT-PCR with primers 64 and 22. Primer 64 matches the sequence from the 5' untranslated region of apo(a) and all its related genes, and primer 22 matches the reverse complement of sequence to the identically repeated apo(a) kringle. The 335-bp fragment that was produced was the predicted size for a product of an mRNA whose mature protein coding region begins in an apo(a)-like fashion. Specific oligonucleotide hybridization probing of the PCR fragments (as described above) showed that both the apo(a) and apo(a)rg-C sequences were represented in this product. Subcloning and sequencing identified a product of an apo(a)rg-C transcript by the following criteria: the initial 169 bp following the PCR primer sequence exactly matched the corresponding region of the apo(a)rg-C sequence but not the sequence of plasminogen, apo(a), or apo(a)rg-B. Following the signal peptide, the encoding sequence was 126 bp that was related to, but distinctly different from, the apo(a) kringle sequence (19% nucleotide differences, including a 3-base deletion compared with the first kringle of apo(a)). However, the sequence retained a kringle 4-like character, including the position of two cysteine
Fig. 3. Blot showing kringle 1 through 3 probe hybridization with plasminogen gene- and plasminogen-like gene-containing clones. Three micrograms yeast artificial chromosome (YAC) DNA was digested with EcoRI and analyzed by electrophoresis and blot hybridization to the kringle 1 through 3 cDNA probe in 5 x saline-sodium phosphate-EDTA, 0.1% sodium dodecyl sulfate (SDS), and 30% formamide at 42°C and washed in 0.1% saline-sodium citrate and 0.1% SDS at 42°C. Representative YAC clones containing plasminogen (YAC 6), plasminogen-related gene (PRG) A (YACs 3 and 13), PRG-B (YACs 10 and 15), and PRG-D (YACs 4 and 11) are shown.

codons (Fig 4). Subsequently, the same procedure yielded an RT-PCR product from RNA isolated from a different human liver sample. Two nucleotides differed from the previous sequence, which could have resulted from individual variation or PCR artifact. Thus, a product of the transcription of the apo(a)-rg-C gene that was distinct from the apo(a) or plasminogen sequences was recovered from the two human liver samples that were tested. We took advantage of the distinct sequence of these RNA-derived fragments to design a 33-base oligonucleotide that contained 15 and 17 differences with homologous kringle sequences in apo(a) and plasminogen, respectively. Hybridization with fragments of the P1-C clone highlighted one fragment of this genomic clone, which appeared to have a precise match with the probe sequence. It was subcloned and sequenced to further confirm the identity of the liver RNA-derived sequences and the P1-C genomic clone. It exactly
matches the sequence of the first RT-PCR product discussed above (Fig 4).

Discussion

The apo(a) gene belongs to an extended superfamily of genes related to an ancestral serine protease. This superfamily has evolved through duplication, deletion, exon shuffling, and base substitutions, producing genes responsible for widely different functions. A branch of this superfamily comprises genes encoding a serine protease–like domain and multiple kringle domains (kringles contain approximately 100 amino acids and have internal disulfide-linked residues). Plasminogen, a fibrinolytic zymogen containing five kringle domains, was the first member of this branch to be recognized. Recently several other related genes and their protein products have been discovered. In 1989, Nakamura and colleagues reported the cloning and sequencing of the DNA for hepatocyte growth factor, which has four kringles, an inactive protease-like domain, and 38% amino acid identity with plasminogen. Hepatocyte growth factor is a mitogen for epithelial and other cell types and a stimulator of epithelial cell motility. Its gene is located on chromosome 7. Subsequently, Han et al described a distinct gene located on chromosome 3 that is expressed in liver and has identical domain structure and 50% amino acid identity with hepatocyte growth factor. It was subsequently shown that this gene encodes macrophage-stimulating protein, which enhances the response of macrophages to the chemoattractant C5a. The apo(a) gene is a closer relative of plasminogen than these two genes. The human apo(a) and plasminogen genes have 99% DNA sequence identity of their 5' flanking regions, 80% in kringle 4 domains, 91% in kringle 5 domains, and 94% in their protease domains. If the apo(a) gene arose from a recent duplication of the plasminogen gene, it is not alone, since we have now found that the human genome contains at least seven distinct genes and/or pseudogene products of this duplication.

The close linkage of some but not all of these genes/pseudogenes is reminiscent of the globin family. In humans, the α- and β-globin families form clusters of adjacent homologous genes on two different chromosomes. The globin clusters contain both active genes and pseudogenes, and within the blocks of similar DNA sequence there are nearly identical regions thought to be the product of subsequent gene conversions. It is likely that the members of the apo(a)/plasminogen gene family also arose from initial duplications and subsequent base substitutions followed by more recent gene conversion events involving portions of the duplicated sequence. It is worth noting that the 5' regions of apo(a), plasminogen, apo(a)rg-B, apo(a)rg-C, and the PRGs retain sufficient identity that they can all drive expression of a reporter transcript in transfected liver cells. Such duplication of coding and control regions not only allows the new gene family members to be transcribed but could also lead to a situation in which expression of a tandemly duplicated gene is affected by regulatory sequences that are a considerable distance apart. Here, transcription occurs near homologous sequences in tandem, duplicated genes and may act as long-distance enhancers. An example of this phenomenon may be found in the β-globin locus, where a control region flanking the e-globin gene profoundly affects the expression of the β-globin gene, even though it is separated by over 50 kb and five intervening genes and pseudogenes.

Protein and restriction fragment length polymorphisms and metaphase chromosome in situ hybridization have previously been used to show that the human plasminogen and apo(a) genes are closely linked in the region 6q26-ter. Our hybridization data suggested that the probes used in those studies might have limited sequence in common with PRG-A, PRG-B, and PRG-D. Since these genes were not found in any YAC clones that also contained plasminogen or apo(a), they were not necessarily linked. In fact, we localized PRG-B and PRG-D to chromosome 2 and PRG-A to chromosome 4. Frank and colleagues report that in addition to the plasminogen locus on chromosome 6, a second signal was detected on chromosome 2 when hybridization stringency was reduced. This locus, which they designate PLGL, might correspond to PRG-B and/or PRG-D.

The detailed structure of apo(a)rg-B and apo(a)rg-C and their possible functions are not known. Our previous analysis revealed that the 5' region of apo(a)rg-B contains over 90% identity with the apo(a) gene and plasminogen genes, with a sequence nearly identical to the first exon of these genes that codes for the secretion prepeptide. However, analysis of the YAC clones presented here suggested that it is a potential pseudogene that does not contain homologues of the other domains of plasminogen and apo(a). Although the 5' exon could be linked to coding sequences not recognized by our probes or that are located beyond the extent of the YAC clones, apo(a)rg-B may represent a partial gene duplication with no function. There was a stronger indication that apo(a)rg-C is a functioning gene. Hybridization showed that apo(a)rg-C has an apo(a)-like rather than a plasminogen-like structure. The apo(a) gene has deleted the preactivation and kringle 1 through 3 domains of plasminogen, directly linking the signal sequence to the first of many tandem domains with 80% identity with the kringle 4 of plasminogen. The domain content of apo(a)rg-C appears to be similar. We designed a PCR strategy to identify potential transcripts of this gene based on the sequence of its first exon and these hybridization results. The primer pair PCR 64 and PCR 22 would generate a fragment of =340 bp if the 5' untranslated and signal peptide region were directly linked to a kringle with homology to the repeated kringles of apo(a). Such a fragment was produced from RNA prepared from the liver of two different individuals. Its sequence proved to match that of the first exon of apo(a)rg-C, continuing into a sequence that resembled plasminogen kringle 4 but with numerous differences from those in apo(a) or plasminogen. On the basis of this sequence, an oligonucleotide probe was designed that hybridized with only one of the kringle-containing restriction fragments of the PI-C genomic clone. Sequence analysis of this fragment of the apo(a)rg-C genomic clone confirmed its identity with the transcript derived from liver RNA. The apo(a)rg-C RNA may be present in relatively low levels; attempts to detect it by Northern blot hybridization were ambiguous. The size of its transcript was not determined, since various probes detected a number of
bands, which may be derived from transcripts of other members of this gene family as well. Since it contains an estimated five kringle sequences, it is the same size as plasminogen and thus be obscured by the 2.9-kb plasminogen transcript detected by many of the probes used (data not shown). The definitive analysis of the apo(a)/rg-C sequence therefore awaits the cloning of its full-length cDNA.

Detection of transcripts from novel members of the apo(a)/plasminogen gene family is not without precedent. The plasminogen-like genes PRG-A and PRG-B were first described by Ichinose, who isolated partial genomic clones in the course of characterizing plasminogen and apo(a) genes. At the time, they were considered to be pseudogenes, since they contain an early stop codon. However, a transcript corresponding to the PRG-B sequence has been detected in malignant cancer cells. It is predicted to encode an 8800-Da protein homologous to only the preactivation peptide domain of plasminogen. Thus, the apo(a)/plasminogen gene family is now known to contain at least seven genes and/or pseudogenes with regions of 80% to 100% sequence identity, and at least two of these genes may encode novel proteins.

Acknowledgments

This research was supported by National Institutes of Health program project grant HL 48638-01 and University of California/Tobacco-Related Disease Research program grant 2RT0339 to Dr Lawn. Dr Byrne is supported by fellowships from the Peel Medical Trust & Parke-Davis/Cambridge University, UK, and Dr Cheng is supported by Department of Energy grant DE-AC03-76SF00098. We thank Uta Francke and Xu Li for DNA samples of Chinese hamster/human hybrid cell lines and David Wade for aid in characterizing YAC clones.

References

17. Sternberg N. Bacteriophage Pl cloning system for the isolation, amplification and recovery of DNA fragments as large as 100 kilobase pairs. Proc Natl Acad Sci U S A. 1990;87:103-107.
28. Slightom JL, Bleich A, Smithies O. Human fetal a- and g-globin genes: complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell. 1980;21:627-634.

Byrne et al Apolipoprotein(a)/Plasminogen Gene Cluster 541
The human apolipoprotein(a)/plasminogen gene cluster contains a novel homologue transcribed in liver.

C D Byrne, K Schwartz, K Meer, J F Cheng and R M Lawn

Arterioscler Thromb Vasc Biol. 1994;14:534-541
doi: 10.1161/01.ATV.14.4.534

Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1994 American Heart Association, Inc. All rights reserved.
Print ISSN: 1079-5642. Online ISSN: 1524-4636

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://atvb.ahajournals.org/content/14/4/534