Apolipoprotein A5, a Newly Identified Gene That Affects Plasma Triglyceride Levels in Humans and Mice

Len A. Pennacchio, Edward M. Rubin

Abstract—Apolipoprotein A5 (APOA5) is a newly described member of the apolipoprotein gene family whose initial discovery arose from comparative sequence analysis of the mammalian APOA1/C3/A4 gene cluster. Functional studies in mice indicated that alteration in the level of APOA5 significantly affected plasma triglyceride concentrations. Mice that overexpressed human APOA5 displayed significantly reduced triglycerides, whereas mice that lacked apoa5 had a large increase in this lipid parameter. Studies in humans have also suggested an important role for APOA5 in determining plasma triglyceride concentrations. In these experiments, polymorphisms in the human gene were found to define several common haplotypes that were associated with significant changes in triglyceride concentrations in multiple populations. Several separate clinical studies have provided consistent and strong support for the effect with 24% of whites, 35% of blacks, and 53% of Hispanics who carry APOA5 haplotypes associated with increased plasma triglyceride levels. In summary, APOA5 represents a newly discovered gene involved in triglyceride metabolism in both humans and mice whose mechanism of action remains to be deciphered.

Genomics and the Identification of APOA5

Recent accomplishments by the Human Genome Project have facilitated genome-wide strategies to uncover functional regions of the mammalian genome. With the increasing availability of genomic sequences from multiple species, comparative genomic approaches have proven to be a powerful means for annotating human sequence.\(^1\)\(^2\)\(^3\) A basic underlying hypothesis of comparative genomics is that evolutionarily conserved sequences are functionally important. Based on this hypothesis, biologic experimentation has focused on sequences that are highly conserved between vertebrate species. Recently, this strategy was applied to the well-characterized APOA1/C3/A4 gene cluster to identify additional functional elements and resulted in the identification of a new apolipoprotein gene (APOA5).\(^2\)

Whereas the human APOA1/C3/A4 gene cluster sequence has long been available, the recent generation of the orthologous mouse sequence provided a comparison-based method of analyzing the human interval (Figure 1). In addition to the extensive evolutionary conservation of human/mouse exons in the region, several noncoding conserved elements were also uncovered. The largest of these sequences was located \(\approx 30\) kb downstream of APOA4 and spanned several thousand base pairs, with a high percentage identity between humans and mice. These striking features prompted further analyses of the interval.\(^2\)\(^3\) Experimental studies with sequences from this region indicated that the interval was transcribed in liver tissue from both humans and mice and suggested that this was a previously unappreciated gene. Closer examination of the genomic sequence alignment predicted 4 exons in both species with a highly conserved open reading frame. Translation of this sequence revealed high levels of protein identity with APOAIV, the neighboring genes’ product. This paralogous relation resulted in the naming of this newly identified gene as APOA5.\(^2\)\(^3\) These studies further highlight the utility of comparative genomic approaches and led to the further characterization of this newly identified gene.
APOA5 is located within the previously well-defined APOA1/C3/A4 gene cluster on human chromosome 11q23 and in an orthologous block of mouse chromosome 9.4,5 Numerous studies support the concept that members of the apolipoprotein gene family arose by gene duplication events,6,7 although the precise evolutionary events leading to the present-day cluster are not understood. The finding that this 4-member apolipoprotein gene cluster is present in both humans and mice indicates that the evolutionary duplication events predate the last common ancestor of these 2 mammalian species.8 It is worth noting that recent studies in our laboratory have identified the same organization of this gene cluster in the chicken genome, further supporting the idea that the original gene duplication events occurred before the mammalian/avian evolutionary split (Pennacchio and Rubin, unpublished data, 2002).

Amino acid analyses of APOAV show a high level of sequence identity between humans, mice, and chickens (Figure 2A). For instance, human and mouse APOAV display 71% amino acid identity and 78% similarity (Figure 2B). The lengths of the proteins are also similar and are composed of 366, 368, and 355 amino acids in humans, mice, and chickens, respectively. These lengths are slightly shorter than those for APOAIV in these species, which are 396, 394, and 366 amino acids, respectively. Sequence alignment and phylogenetic analysis indicate a closer sequence relation among APOAV in humans, mice, and chickens compared with APOAIV from any of these species (Figure 2C). Again, this supports an ancient duplication event(s) that predate the last common ancestor of mammals and birds.

A unique feature of APOA5 compared with its evolutionarily related apolipoprotein paralogues is its multiple transcripts of ~1.3 and 1.9 kb, as determined by Northern blot analysis.8 Alternative transcripts have not been described for other members within this gene cluster. Examination of expressed sequence tags in GenBank indicates that the 2 transcripts are the result of alternative polyadenylation. The functional relevance, if any, of these 2 transcripts remains to be uncovered.

It is worth noting that APOA5 was also identified in rats, where it was found to be upregulated after liver regeneration.9 In that work, differentially expressed genes were identified 6 hours after a 70% rat hepatectomy. Based on the feature of its being the most highly overexpressed gene, it was independently named “regeneration-associated protein 3” (RAP3) and was proposed to be important in the early phase of liver regeneration. Studies in apoa5 knockout mice after hepatectomy should reveal whether apoav has an important role in liver regeneration.

Genetically Engineered Mice Reveal a Role for APOA5 in Triglyceride Homeostasis

The sequence similarity between APOAV and other apolipoproteins suggested that APOAV functions in the plasma with a role in lipid transport.8 Indeed, structural predictions indicate that APOAV contained a signal peptide export sequence for transport from the liver to plasma and that the
mature APOAV protein contains several amphipathic helical domains. These motifs are characteristics of lipid-binding molecules and are generally present in apolipoproteins.

To define the true in vivo function of APOA5, 2 different engineered mouse lines were generated. First, a human 26-kb XhoI restriction fragment was isolated that was predicted to contain only APOA5 and its flanking sequence. This genomic piece of DNA was subsequently used to generate human APOA5 transgenic mice (APOA5 overexpressers). Second, the endogenous mouse apoA5 gene was deleted from the mouse genome (apoA5 knockouts) through standard embryonic stem-cell techniques. Examination of plasma lipid levels in both APOA5 overexpressers and knockouts revealed profound effects on triglyceride concentrations. Specifically, APOA5 transgenics displayed a 66% decrease in triglycerides. In contrast, apoA5 knockouts had 4-fold higher triglyceride levels than controls. In both models, consistent changes in ApoB and VLDL particle quantities were noted, although no differences were seen in plasma total or HDL cholesterol levels.

Recent studies in mice with adenoviral vectors containing APOA5 confirmed the strong effect of overexpressing APOA5 on plasma triglyceride levels. In that work, the authors estimate ~20-fold higher plasma APOAV levels in vector-treated animals and observed an ~70% reduction in plasma triglyceride concentrations, similar to that in the original APOA5 transgenic report. Interestingly, although these 2 independent overexpressing lines both showed an ~70% reduction in triglycerides, the adenovirus-treated mice also had a 40% reduction in plasma cholesterol levels, a lipid parameter not previously noted to change in response to alterations in APOA5 levels in mice. It is possible that the extremely high level of APOAV in the adenovirus-treated mice creates an artificial condition that accounts for this novel finding.

These studies in mice provide convincing evidence that APOAV plays a role in plasma triglyceride homeostasis. The triglyceride-lowering effect of APOAV is quite distinct from the influence of several other apolipoprotein transgenes (APOCI, CII, CIII, APOB), in which increased protein levels led to higher triglyceride levels. The mechanism by which alterations in APOAV affect triglyceride concentrations remains to be deciphered.

Some clues to the function of APOAV have been suggested on the basis of computational analysis of the proteins’ amino acid sequences. A direct comparison with human APOAIV indicates that APOAV has a higher global hydrophobicity, contains a greater amount of α-helical structures, and is predicted to have a higher interfacial exclusion pressure (Weinberg, unpublished data, 2002). This suggests that APOAV should display a very high affinity for lipid interfaces. It has been proposed that apolipoproteins with moderate lipid affinity and high elasticity, such as APOAIV, facilitate triglyceride-rich particle assembly by stabilizing nascent particles as they acquire triglyceride and expand in the second stage of assembly. Conversely, apolipoproteins with very high lipid affinity, such as APOAV, could impede
this process, thereby functioning as an intracellular “brake” on hepatic lipid export. Functional studies are needed to test this hypothesis.

**Association Studies Extend APOA5’s Role in Triglyceride Homeostasis to Humans**

The experimental studies described thus far provide convincing evidence for APOAV’s role in plasma triglyceride homeostasis in mice. These findings suggested that sequence alterations in the human APOA5 gene might also contribute to differences in plasma lipid levels in humans. To date, from extensive resequencing of the APOA5 gene, severe mutations in humans have not been reported. However, several studies have consistently shown that common genetic variations that map to the APOA5 locus in humans are strongly associated with quantitative differences in plasma triglyceride levels of normolipidemic individuals in the general population. A detailed description of APOA5 polymorphisms, the haplotypes that they define, and the various genetics-associated studies are provided below.

**APOA5 Polymorphism Identification and Haplotype Analysis**

Extensive sequencing of the APOA5 interval in humans has been performed in several studies to identify common polymorphisms for subsequent genetic association studies. Initially, a set of 4 common polymorphisms (SNPs1 through 4; also named c.1259T>C, IVS3+476G>A, −1131T>C, and −12.238T>C, respectively) were identified within the human APOA5 interval. Statistical analysis indicated that the minor alleles SNPs1 through 3 formed a relatively common haplotype that is found in ∼15% of whites.

Subsequently, a more exhaustive screen for APOA5 polymorphisms was undertaken. Through direct DNA sequencing of the gene in 116 hyperlipidemic individuals, 9 additional SNPs were identified. One of the polymorphisms (c.−3A>G) was found to be in strong linkage disequilibrium with the minor alleles for SNPs1 through 3, and this haplotype was named APOA5*2. In addition, a second common polymorphism was also identified, which results in a C-to-G nonsynonymous substitution (c.56C>G) that changes codon 19 from serine to tryptophan. Further haplotype analysis in whites indicated that the minor allele of this polymorphism defines a third common APOA5 haplotype (APOA5*3). Similar to APOA5*2, this haplotype was also found in ∼15% of whites. The remaining 7 polymorphisms from this study were either uncommon or not obviously associated with triglycerides.

Thus, polymorphism discovery and haplotype analysis in whites defined 3 common haplotypes in the APOA5 interval and provided detailed information for genetic association studies in humans. In these analyses, the −1131C allele (SNP3) was used as a marker to define APOA5*2, whereas the c.56G allele (W19) was used to define APOA5*3. Studies of these 2 minor alleles indicated that they are also present at a high frequency in blacks and Hispanics and support the notion that these polymorphisms arose early in the evolutionary history of humankind.

**Berkeley Lipid Study Population**

The initial genetic association study with markers that defined APOA5*2 and APOA5*3 was performed in 500 random, unrelated normolipidemic white individuals. Strong associations were found between these 2 minor haplotypes and increased triglyceride concentrations (Figure 3). Specifically, each of these haplotypes was associated with a ∼30% increase in triglyceride concentrations compared with individuals who lacked the minor alleles that define APOA5*2 and APOA5*3. Further studies in these 500 individuals found no significant association of triglyceride levels with an SstI polymorphism in APOC3 (located ∼40 kbp upstream of APOA5), which has been previously associated with severe hypertriglyceridemia. This finding supports the contention that these APOA5 haplotypes are associated with altered triglyceride levels independent of the APOC3 SstI polymorphism.

**Stratified Population Study**

As a follow-up to the initial association, an independent white population was examined with a different experimental design. In this study, the allele frequencies of −1131T>C (SNP3; APOA5*2) and c.56C>G (S19W; APOA5*3) were compared in an unrelated group of whites stratified according to plasma triglyceride levels. The 2 groups represented (1) several hundred individuals with triglyceride levels in the top tenth percentile and (2) several hundred individuals with triglyceride levels from the bottom tenth percentile of a larger population. For both polymorphisms, an ∼3-fold overrepresentation of the minor alleles was found in individuals from the high- versus the low-plasma-triglyceride group.

**The Dallas Heart Disease Prevention Program**

Recently, a third study tested for a genetic association between APOA5 and triglyceride levels in ethnic groups other than whites. The Dallas population study comprised ∼2600 randomly selected individuals representing blacks and Hispanics in addition to whites. Once again, strong genetic associations were found between both the −1131T>C and c.56C>G polymorphisms and triglyceride concentrations. For c.56C>G (APOA5*3), the effect was seen in both men and women from each ethnic group. For c.56C>G (APOA5*2), increased plasma triglyceride concentrations were found in Hispanic men and women and in white men but not in black men and women or white women. Whether the lack of an association in this subset of samples is due to small sample size or significant sex- and ethnicity-specific effects remains to be determined.

**Northwick Park Heart Study (NPHSII)**

In addition, a detailed genetic analysis was performed with 9 polymorphisms that span the APOC3/A4 cluster as well as APOA5. Confirming the previous findings, in 2800 white males, individuals homozygous for APOA5 19W (APOA5*2) or SNP3 (−1131C; APOA5*2) had 52% and 40% higher triglyceride levels, respectively (P<0.003), compared with individuals homozygous for the common allele. To understand the relative contribution of genes in the cluster toward the effects on triglycerides, a 9-SNP-haplotype analysis was
undertaken. This confirmed that rare alleles of APOA5 (S19W) and APOC3 (−482T>C) independently define the major triglyceride-raising haplotypes. The triglyceride-lowering effect of the rare allele of APOA4, T347S, reflects the strong linkage disequilibrium with the triglyceride-lowering, common alleles of APOA5 (S19W) and APOC3 (−482T>C). Thus, common variants of APOA5, independent of APOC3, contributed to triglyceride determination in these healthy, middle-aged men.

Japanese School Children Study
In a final normolipidemic population study, Japanese school children were examined for an association between APOA5 SNP3 (−1131C; APOA5*2) and triglycerides.23 Once again, higher triglyceride levels (∼15%) were found in individuals who carried minor alleles for SNP3 compared with those homozygous for the major allele. This study supports an age-independent effect of APOA5 polymorphisms on plasma triglyceride concentrations and further extends its effect to an additional ethnic group. Of additional significance is the extremely high population frequency of the SNP3 minor allele in the Japanese. In this work, 61% of individuals carried at least 1 copy of the SNP3 minor allele, in comparison with the 15%, 30%, and 22% found in whites, Hispanics, and blacks, respectively. The c.56C>G (APOA5*3) polymorphism was not examined in this population.

APOA5 and Familial Combined Hyperlipidemia
In addition to genetic studies in the general normolipidemic population, recent work explored APOA5 variation and its relation to familial combined hyperlipidemia (FCHL).24 This population represented 16 families with 42 FCHL individuals and 61 first-degree relative controls. In this sample, the SNP3 (−1131T>C) APOA5 polymorphism was examined. Analysis of both affected and unaffected groups combined indicated 45% higher triglycerides in individuals who carried the −1131C allele (1.82 mmol/L) compared with those lacking it (1.26 mmol/L). In addition, the minor allele was nearly 3-fold more common in the FCHL group compared with controls. These data support the concept that polymorphism in the APOA5 interval contributes to the complex condition of FCHL.

Summary of Human Genetic Studies
These findings establish that the APOA5 locus contributes significantly to interindividual variation in plasma triglyceride levels in the general human population. Together, the APOA5*2 and APOA5*3 haplotypes are found in 25% to 50% of blacks, Hispanics, and whites, highlighting the large fraction of individuals with increased triglyceride levels that are due solely to the effect of APOA5 polymorphisms. In addition to the strong effect of APOA5 haplotypes on plasma triglycerides from various age, sex, and ethnic groups, these chromosomal regions also appear to be important contributors to the common condition of FCHL.

Functional Cause of the APOA5 Genetic Association
Genetic association studies between APOA5 polymorphisms and triglyceride levels have provided convincing evidence for a relation between these 2 parameters. However, it remains unclear which variant(s) in the APOA5 chromosomal region is responsible for this association. Although functional studies are needed to definitely prove the culprit variant(s), there is a single, strong candidate within each of the minor haplotypes (APOA5*2 and APOA5*3) associated with triglyceride levels.16

APOA5*2 comprises several minor alleles that extend throughout the APOA5 gene. One of these polymorphisms (c.−3A>G) is found 3 bp upstream from the predicted start codon for APOA5 in a functionally important base of the putative Kozak consensus sequence. The compilation of numerous vertebrate gene sequences upstream from the start codon indicates a strong bias in the consensus base at the −3-bp position, with the nucleotide A being found in 61% of cases.25 Thus, the c.−3A>G polymorphism, which changes the common allele A to G at −3 bp, could potentially result in a decreased rate of APOA5 mRNA translation and thereby lead to lower APOAV plasma levels. Mechanistically, this result would be consistent with the finding of increased triglycerides in mice lacking apoA5 (apoA5 knockouts).8

In contrast to APOA5*2, APOA5*3 is defined by a single minor allele within the APOA5 coding sequence. The c.56C>G sequence variant results in a nonconservative change of serine to tryptophan at codon 19. Apolipoproteins and other polypeptides that function in plasma are known to contain N-terminal export signal sequences. Indeed, computational analyses for APOAV predicted a strong export consensus sequence with a likely export cleavage site between amino acids 23 and 24.26 The change of a serine to a bulky tryptophan residue at position 19 could thus reduce the rate of APOAV export from the liver and result in higher triglycerides in humans.

Future Perspectives
The large generation of DNA sequences by the Human Genome Project has accelerated our discovery of new genes. This is readily apparent in the discovery of APOA5 through the use of publicly available genomic sequence surrounding the APOAI/C3/A4 gene cluster in humans and mice.8 On the basis of this success, should we expect that other members of the mammalian apolipoprotein gene family remain to be discovered? Although possible, the answer is unlikely. Currently, the sequence for the vast majority of the human genome is available, and electronic searches for additional evolutionarily related apolipoprotein sequences have failed to uncover additional family members. Be that as it may, why was APOA5 previously missed in the well-studied area of plasma apolipoproteins? Part of the explanation may be due to the neighboring gene and paralogous relation between APOA4 and APOA5. These features may have shielded their view and made these 2 separate genes appear like 1. The low levels of APOAV in mammalian plasma may also have contributed to its lack of earlier discovery.

A striking feature of APOA5 is its strong effect on triglyceride levels in both humans and mice. It is clear that alterations in APOA5 levels in mice are inversely correlated with triglyceride levels. It remains to be determined why APOAV strongly influences plasma triglyceride and not
cholesterol levels yet is found primarily on HDL particles. In addition to these mouse studies, common human sequence variation in APOA5 has also been significantly associated with triglyceride levels in the general population. Surprisingly, the minor APOA5 haplotypes associated with increased triglycerides are found in ~25% to 50% of whites, blacks, and Hispanics, indicating their wide-reaching effects. These findings support the concept that common genetic variation contributes to common quantitative phenotypes in the general population. This holds promise for future genome-wide association strategies aimed at uncovering common genetic contributors to quantitative and disease phenotypes in humans. Future experimentation will determine whether the significant increase in plasma triglyceride levels translates into an increased risk for cardiovascular disease and the specific mechanisms by which APOAV affects this important plasma lipid parameter.

Acknowledgment
This work was supported in part by the NIH-NHLBI Programs for Genomic Application grant HL66681 (to E.M.R.) through the US Department of Energy under contract No. DE-AC03 to 76SF00098.

References
Apolipoprotein A5, a Newly Identified Gene That Affects Plasma Triglyceride Levels in Humans and Mice
Len A. Pennacchio and Edward M. Rubin

Arterioscler Thromb Vasc Biol. 2003;23:529-534; originally published online December 26, 2002;
doi: 10.1161/01.ATV.0000054194.78240.45

Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2002 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/23/4/529