Apolipoprotein A5: A Newly Identified Gene Impacting Plasma Triglyceride Levels in Humans and Mice

Len A. Pennacchio and Edward M. Rubin

Genome Sciences Department, Lawrence Berkeley National Laboratory,

Berkeley, CA., 94720, USA.

Address for correspondence and reprints: Dr. Len A. Pennacchio, Department of Genome Sciences, MS 84-171, One Cyclotron Road, Lawrence Berkeley National Laboratory, Berkeley, CA. 94720. Email: LAPennacchio@lbl.gov. Phone: (510) 486-7498. Fax: (510) 486-4229.

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 impacted plasma triglyceride concentrations. Mice over-expressing human *APOA5* displayed significantly reduced triglycerides, while mice lacking *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 Caucasians, 35% of African-Americans and 53% of Hispanics carrying *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[1-5]. A basic underlying hypothesis of comparative genomics is that *evolutionary conserved sequences are functionally important*. Based on this hypothesis, biological 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*)[3].

While the human *APOA1/C3/A4* gene cluster sequence has long been available, the recent generation of the orthologous mouse sequence provided a comparative-based method of analyzing the human interval (Figure 1). In addition to the extensive evolutionary conservation of human/mouse exons in the region, several non-coding conserved elements were also uncovered. The largest of these sequences was located ~30 kb downstream of *APOA4* and spanned several thousand basepairs with a high percent identity between humans and mice. These striking features prompted further analyses of the interval[3]. Experimental studies with sequences from this region indicated the interval was transcribed in liver tissue from both humans and mice and suggested this was a previously unappreciated gene. Closer examination of the genomic sequence alignment predicted four 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 relationship resulted in the naming of this newly identified gene as "*APOA5*"[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[6-8]. Numerous studies support the concept that members of the apolipoprotein gene family arose by gene duplication events[9,10]; though the precise evolutionary events leading to the present day cluster are not understood. The finding that this four-member apolipoprotein gene cluster is present in both humans and mice indicate that the evolutionary duplication events predate the last common ancestor of these two mammalian species[11]. 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 original gene duplication events were prior to the mammalian/avian evolutionary split (Pennacchio and Rubin, unpublished data, 2002).

Amino acid analyses of APOAV show a high level of sequence identity between human, mouse and chicken (Figure 2A). For instance, human and mouse APOAV display 71% amino acid identity and 78% similarity (Figure 2B). The length of the proteins are also similar and are composed of 366, 368 and 355 amino acids in human, mouse and chicken, 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 indicates a closer sequence relationship among APOAV in humans, mice and chicken compared to APOAIV from any of these species (Figure 2C). Again, this supports an ancient duplication event(s) that predates the last common ancestor of mammals and birds.

A unique feature of *APOA5* compared to its evolutionary-related apolipoprotein paralogs is its multiple transcripts of approximately 1.3- and 1.9-kilobases as determined by Northern blot analysis[11]. Alternative transcripts have not been described for other members within this gene cluster. Examination of expressed sequence tags in Genbank indicates the two transcripts are the result of alternative polyadenylation. The functional relevance, if any, of these two transcripts remains to be uncovered.

It is worth noting that *APOA5* was also identified in rats where it was found to be up-regulated following liver regeneration[12]. In this work, differentially expressed genes were identified six hours after a 70% rat hepatectomy. Based on the feature of being the most highly over-expressed gene, it was independently named "Regeneration Associated Protein 3 (RAP3)" and was proposed to be important in the early phase of 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[11]. 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[11,13-15]. These motifs are characteristics of lipid binding molecules and are generally present in apolipoproteins.

To define the true *in vivo* function of *APOA5*, two different engineered mouse lines were generated[11]. First, a human 26kb *Xho*I 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* over-expressors"). Second, the endogenous mouse *apoA5* gene was deleted from the mouse genome ("*apoA5* knockouts") through standard embryonic stem cell technology. Examination of plasma lipid levels in both *APOA5* "over-expressors" and "knockouts" revealed profound effects on triglyceride concentrations. Specifically, *APOA5* transgenics displayed a 66% decrease in triglycerides. In contrast, *apoA5* knockouts had four-fold higher triglycerides than controls. In both models consistent changes in APOB and very low-density lipoprotein (VLDL) particle quantities were noted, though no differences were seen in plasma- or HDL-cholesterol levels.

Recent studies in mice using adenoviral vectors containing *APOA5* confirmed the strong effect of over-expressing *APOA5* on plasma triglyceride levels[16]. In this 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[11]. Interestingly, while these two independent over-expressing lines both showed an~70% reduction in triglycerides, the adenoviral 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 adenoviral 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 impact of several other apolipoprotein transgenes (APOCI , CII, CIII, APOB) where increased protein levels led to higher triglyceride levels[17-19]. The mechanism by which alterations in APOAV impact on triglyceride concentrations remains to be deciphered.

Some clues to the function of APOAV have been suggested based on the computational analysis of the proteins' amino acid sequence. A direct comparison with human APOAIV indicates that APOAV has a higher global hydrophobicity, contains a greater amount of α-helical structure, 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[20,21]. 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 showed that common genetic variations mapping to the human *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 they define, and the various genetic 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[11,22]. Initially, a set of four common polymorphisms (SNPs1-4, 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 of SNPs1-3 formed a relatively common haplotype which is found in ~15% of Caucasians[11,22].

Subsequently, a more exhaustive screen for *APOA5* polymorphisms was undertaken[22]. Through the direct DNA sequencing of the gene in 116 hyperlipidemic individuals, nine additional SNPs were identified. One of the polymorphisms (c.-3A>G) was found to be in strong linkage disequilibrium with the minor alleles for SNPs1-3 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) changing codon 19 from serine to tryptophan. Further haplotype analysis in Caucasians 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 Caucasians. The remaining seven polymorphisms from this study were either uncommon or not obviously associated with triglycerides[22].

Thus, polymorphism discovery and haplotype analysis in Caucasians defined three 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, while the c.56G allele (W19) was used to define *APOA5*3*[22]. Studies of these two minor alleles indicated that they are also present at a high frequency in African-Americans and

Hispanics and support 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 normo-lipidemic Caucasian individuals[11,22]. Strong associations were found between these two minor haplotypes and increased triglyceride concentrations (Figure 3). Specifically, each of these haplotypes was associated with a ~30% increase in triglyceride concentrations compared to individuals lacking the minor alleles that define *APOA5**2 and *APOA5*3*. Similar association was also found for the related plasma parameter VLDL mass. Further studies in these 500 individuals found no significant association of triglyceride levels with a *Sst*1 polymorphism in *APOC3* (located ~40 kbp upstream of *APOA5*) which has been previously associated with severe hyper-triglyceridemia[23-34]. This finding supports that these *APOA5* haplotypes are associated with altered triglyceride levels independent of the *APOC3 Sst*1 polymorphism.

Stratified Population Study

As a follow-up to the initial association, an independent Caucasian population was examined using a different experimental design[11,22]. 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 Caucasians stratified according to plasma triglyceride levels. The two 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 approximately three-fold over-representation of the minor alleles was found in individuals from the high- versus low-plasma triglyceride group.

The Dallas Heart Disease Prevention Program

Recently, a third study tested for genetic association between *APOA5* and triglycerides in ethnic groups beyond Caucasians[22]. The Dallas population studied comprised ~2600 randomly selected individuals representing African-Americans and Hispanics, in addition to Caucasians. Once again, strong genetic associations were found between both the -1131T>C and c.56C>G polymorphisms and triglycerides. For c.56C>G (*APOA5*3*), the effect was seen in both men and women from each ethnic group. While for c56.C>G (*APOA5**2), increased plasma triglyceride concentrations were found in Hispanic men and women, and in Caucasian men, but not in African-American men and women or in Caucasian women. Whether the lack of an association in this subset of samples is due to a small sample size or significant gender- and ethnic-specific effects remains to be determined.

Northwick Park Heart Study (NPHSII)

In addition, a detailed genetic analysis was performed with several polymorphisms within the *APOA1/C3/A4* cluster as well as *APOA5*[35]. In 2800 Caucasian males, individuals homozygous for *APOA5* 19W (*APOA5*3*) or SNP3 (-1131C; *APOA5**2) had 52% and 40% higher triglycerides ($p < 0.003$) compared to individuals homozygous for the common allele, respectively. Further examination of a SNP (T347S) in the neighboring *APOA4* gene indicated that 347SS men had 23% lower triglycerides than 347TT men. Statistical analyses of these SNPs and four *APOC3* SNPs (-2845T>G, -482C>T, 1100C>T, and 3238C>G) revealed that the effects of *APOA5* S19W ($p<0.0001$), *APOC3* –482C>T ($p<0.002$) and *APOA4* T347S ($p<0.007$) most significantly accounted for the triglyceride differences in this male population. Additional haplotype analysis defined the triglyceride-raising alleles as being *APOA5* W19 and *APOC3* –482T. The conclusion from this study was that variation in both *APOA5* and *APOA4*, in addition to effects previously reported for *APOC3*, are associated with differences in triglycerides in healthy men.

Japanese School Children Study

In a final normo-lipidemic population study, Japanese school children were examined for association between *APOA5* SNP3 (-1131C; *APOA5**2) and triglycerides[36]. Once again, higher triglyceride levels (~15%) were found in individuals containing minor alleles for SNP3 compared to those homozygous for the major allele. This study supports an age-independent effect of *APOA5* polymorphisms on plasma triglyceride concentrations and further extends its' affect 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 contained at least one copy of the SNP3 minor allele, in comparison to the 15%, 30% and 22% found in Caucasians, Hispanics, and African-Americans, 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 normo-lipidemic population, studies have also been performed in individuals with abnormal plasma lipid levels. A common disorder of lipid metabolism in humans is combined hyperlipidemia (CHL) which is characterized by increased triglycerides and/or cholesterol and leads to an increased risk of cardiovascular disease. The severe triglyceride phenotype in this complex condition suggested that *APOA5* might contribute to the disease etiology.

To test whether variation at the *APOA5* locus contributes to the transmission of CHL, linkage and linkage disequilibrium (LD) tests were performed on a large cohort of families (n=128) with familial CHL (FCHL). Indeed, the linkage data produced evidence for linkage of the *APOA1/C3/A4/A5* genomic interval to FCHL (NPL = 1.7, P = 0.042). Further studies revealed that two independent rare alleles, *APOA5* 19W (c.56G) and the nearby *APOC3* c.386G were over-transmitted in FCHL (P = 0.004 and 0.007, respectively). The *APOA5 19W* allele was associated with increased plasma triglycerides in FCHL probands, while in contrast the second and independent *APOC3* c.386G allele was associated with increased plasma triglyceride levels in FCHL pedigree founders. This study supported that common polymorphism in the *APOA1/C3/A4/A5* gene cluster contributes to the transmission of FCHL in a substantial proportion of affected families.

In a second independent study, a single *APOA5* polymorphism (SNP3; -1131T>C) was examined in a set of 16 FCHL families[37]. These families represented 42 FCHL individuals and 61 first-degree relative controls. Analysis of both groups combined indicated 45% higher triglycerides in individuals containing the -1131C allele (1.82 mmol/L) compared to those lacking it (1.26

mmol/L). In addition, the minor allele was nearly three-fold more common in the FCHL group compared to controls. Again, these data support that polymorphism in the *APOA5* interval contributes to FCHL.

Summary of Human Genetic Studies

These findings establish that the *APOA5* locus contributes significantly to inter-individual variation in plasma triglyceride levels in the general human population. Together, the *APOA5**2 and *APOA5**3 haplotypes are found in 25 to 50 percent of African-Americans, Hispanics, and Caucasians, highlighting the large fraction of individuals with increased triglycerides solely due to the effect of *APOA5* polymorphisms. In addition to the strong effect of *APOA5* haplotypes on plasma triglycerides from various age-, gender- 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 relationship between these two parameters. However, it remains unclear which variant(s) in the *APOA5* chromosomal region are responsible for this association. While 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[22].

*APOA5**2* is composed of several minor alleles that extend throughout the *APOA5* gene. One of these polymorphisms (c.-3A>G) is found three basepairs upstream of the predicted start codon for *APOA5* in a functionally important base of the putative Kozak consensus sequence[38-40]. The compilation of numerous vertebrate gene sequences upstream of the start codon indicates a strong bias in the consensus base at the -3bp position with the nucleotide A being found in 61% of cases[39]. Thus, the c.-3A>G polymorphism which changes the common allele A to G at -3bp 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")[11].

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 non-conservative 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 predict a strong export consensus sequence with a likely export cleavage site between amino acids 23 and 24[41]. 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 sequence 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 *APOA1/C3/A4* gene cluster in humans and mice[11]. Based on this success should we expect that other members of the mammalian apolipoprotein gene family remain to be discovered? While possible, the answer is unlikely. Currently, 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. With 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 relationship between *APOA4*and *APOA5*. These features may have shielded their view and made these two separate gene appear like one. 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. 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 ~25-50 % of Caucasians, African-Americans, 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 if the significant increase in plasma triglyceride levels translates into an increased risk for cardiovascular disease and the specific mechanisms by which APOAV impacts on this important plasma lipid parameter.

ACKNOWLEDGEMENTS

We thank XXXX for thoughtful discussions. This work was supported in part by the NIH-NHLBI Programs for Genomic Application Grant HL66681 (EMR) through the U.S. Department of Energy under contract no. DE-AC03-76SF00098 (EMR), and an appointment to the Alexander Hollaender Distinguished Postdoctoral Fellowship Program sponsored by the U.S. Department of Energy, Office of Biological and Environmental Research, and administered by the Oak Ridge Institute for Science and Education (LAP).

REFERENCES

1. Hardison RC. Conserved noncoding sequences are reliable guides to regulatory elements [In Process Citation]. *Trends Genet*. 2000;16:369-72.
2. Hardison RC, Oeltjen J, Miller W. Long human-mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. *Genome Res*. 1997;7:959-66.
3. Pennacchio LA, Rubin EM. Genomic strategies to identify mammalian regulatory sequences. *Nat Rev Genet*. 2001;2:100-9.
4. Rubin EM, Tall A. Perspectives for vascular genomics. *Nature*. 2000;407:265-9.
5. Wasserman WW, Palumbo M, Thompson W, Fickett JW, Lawrence CE. Human-mouse genome comparisons to locate regulatory sites. *Nat Genet*. 2000;26:225-228.
6. Bruns GA, Karathanasis SK, Breslow JL. Human apolipoprotein A-I--C-III gene complex is located on chromosome 11. *Arteriosclerosis*. 1984;4:97-102.
7. Karathanasis SK, McPherson J, Zannis VI, Breslow JL. Linkage of human apolipoproteins A-I and C-III genes. *Nature*. 1983;304:371-3.
8. Karathanasis SK. Apolipoprotein multigene family: tandem organization of human apolipoprotein AI, CIII, and AIV genes. *Proc Natl Acad Sci U S A*. 1985;82:6374-8.
9. Boguski MS, Birkenmeier EH, Elshourbagy NA, Taylor JM, Gordon JI. Evolution of the apolipoproteins. Structure of the rat apo-A-IV gene and its relationship to the human genes for apo-A-I, C-III, and E. *J Biol Chem*. 1986;261:6398-407.
10. Scott J. The human apolipoprotein genes. *Oxf Surv Eukaryot Genes*. 1987;4:168-97.
11. Pennacchio LA, Olivier M, Hubacek JA, Cohen JC, Cox DR, Fruchart JC, Krauss RM, Rubin EM. An apolipoprotein influencing triglycerides in

humans and mice revealed by comparative sequencing. *Science*. 2001;294:169-73.

12. van der Vliet HN, Sammels MG, Leegwater AC, Levels JH, Reitsma PH, Boers W, Chamuleau RA. Apolipoprotein A-V: a novel apolipoprotein associated with an early phase of liver regeneration. *J Biol Chem*. 2001;276:44512-20.
13. Lupas A, Van Dyke M, Stock J. Predicting coiled coils from protein sequences. *Science*. 1991;252:1162-4.
14. Lupas A. Prediction and analysis of coiled-coil structures. *Methods Enzymol*. 1996;266:513-25.
15. Prestridge DS. SIGNAL SCAN 4.0: additional databases and sequence formats. *Comput Appl Biosci*. 1996;12:157-60.
16. van der Vliet HN, Schaap FG, Levels JH, Ottenhoff R, Looije N, Wesseling JG, Groen AK, Chamuleau RA. Adenoviral overexpression of apolipoprotein A-V reduces serum levels of triglycerides and cholesterol in mice. *Biochem Biophys Res Commun*. 2002;295:1156-9.
17. Shachter NS, Hayek T, Leff T, Smith JD, Rosenberg DW, Walsh A, Ramakrishnan R, Goldberg IJ, Ginsberg HN, Breslow JL. Overexpression of apolipoprotein CII causes hypertriglyceridemia in transgenic mice. *J Clin Invest*. 1994;93:1683-90.
18. Shachter NS, Ebara T, Ramakrishnan R, Steiner G, Breslow JL, Ginsberg HN, Smith JD. Combined hyperlipidemia in transgenic mice overexpressing human apolipoprotein Cl. *J Clin Invest*. 1996;98:846-55.
19. Ito Y, Azrolan N, O'Connell A, Walsh A, Breslow JL. Hypertriglyceridemia as a result of human apo CIII gene expression in transgenic mice. *Science*. 1990;249:790-3.
20. Weinberg RB, Anderson RA, Cook VR, Emmanuel F, Denefle P, Hermann M, Steinmetz A. Structure and interfacial properties of chicken apolipoprotein A-IV. *J Lipid Res*. 2000;41:1410-8.

21. Weinberg RB, Cook VR, DeLozier JA, Shelness GS. Dynamic interfacial properties of human apolipoproteins A-IV and B-17 at the air/water and oil/water interface. *J Lipid Res*. 2000;41:1419-27.

22. Pennacchio LA, Olivier M, Hubacek JA, Krauss RM, Rubin EM, Cohen JC. Two independent apolipoprotein A5 haplotypes influence human plasma triglyceride levels. *Hum Mol Genet*. 2002;*in press*.

23. Groenendijk M, Cantor RM, Blom NH, Rotter JI, de Bruin TW, Dallinga-Thie GM. Association of plasma lipids and apolipoproteins with the insulin response element in the apoC-III promoter region in familial combined hyperlipidemia. *J Lipid Res*. 1999;40:1036-44.

24. Dammerman M, Sandkuijl LA, Halaas JL, Chung W, Breslow JL. An apolipoprotein CIII haplotype protective against hypertriglyceridemia is specified by promoter and 3' untranslated region polymorphisms. *Proc Natl Acad Sci U S A*. 1993;90:4562-6.

25. Ordovas JM, Civeira F, Genest J, Jr., Craig S, Robbins AH, Meade T, Pocovi M, Frossard PM, Masharani U, Wilson PW, et al. Restriction fragment length polymorphisms of the apolipoprotein A-I, C-III, A-IV gene locus. Relationships with lipids, apolipoproteins, and premature coronary artery disease. *Atherosclerosis*. 1991;87:75-86.

26. Shoulders CC, Harry PJ, Lagrost L, White SE, Shah NF, North JD, Gilligan M, Gambert P, Ball MJ. Variation at the apo AI/CIII/AIV gene complex is associated with elevated plasma levels of apo CIII. *Atherosclerosis*. 1991;87:239-47.

27. Shoulders CC, Grantham TT, North JD, Gaspardone A, Tomai F, de Fazio A, Versaci F, Gioffre PA, Cox NJ. Hypertriglyceridemia and the apolipoprotein CIII gene locus: lack of association with the variant insulin response element in Italian school children. *Hum Genet*. 1996;98:557-66.

28. Surguchov AP, Page GP, Smith L, Patsch W, Boerwinkle E. Polymorphic markers in apolipoprotein C-III gene flanking regions and hypertriglyceridemia. *Arterioscler Thromb Vasc Biol*. 1996;16:941-7.

29. Zeng Q, Dammerman M, Takada Y, Matsunaga A, Breslow JL, Sasaki J. An apolipoprotein CIII marker associated with hypertriglyceridemia in Caucasians also confers increased risk in a west Japanese population. *Hum Genet*. 1995;95:371-5.

30. Stocks J, Paul H, Galton D. Haplotypes identified by DNA restriction-fragment-length polymorphisms in the A-1 C-III A-IV gene region and hypertriglyceridemia. *Am J Hum Genet*. 1987;41:106-18.

31. Tybjaerg-Hansen A, Nordestgaard BG, Gerdes LU, Faergeman O, Humphries SE. Genetic markers in the apo AI-CIII-AIV gene cluster for combined hyperlipidemia, hypertriglyceridemia, and predisposition to atherosclerosis. *Atherosclerosis*. 1993;100:157-69.

32. Paul H, Galton D, Stocks J. DNA polymorphic patterns and haplotype arrangements of the apo A-1, apo C-III, apo A-IV gene cluster in different ethnic groups. *Hum Genet*. 1987;75:264-8.

33. Ahn YI, Valdez R, Reddy AP, Cole SA, Weiss KM, Ferrell RE. DNA polymorphisms of the apolipoprotein AI/CIII/AIV gene cluster influence plasma cholesterol and triglyceride levels in the Mayans of the Yucatan Peninsula, Mexico. *Hum Hered*. 1991;41:281-9.

34. Tas S. Strong association of a single nucleotide substitution in the 3'-untranslated region of the apolipoprotein-CIII gene with common hypertriglyceridemia in Arabs. *Clin Chem*. 1989;35:256-9.

35. Talmud P, Hawe E, Martin S, Olivier M, Miller GJ, Rubin EM, Pennacchio LA, Humphries SE. Relative contribution of variation within the *APOC3-A4-A5* gene cluster in determining plasma triglycerides. *Hum Mol Genet*. 2002;*in press*.

36. Endo K, Yanagi H, Araki J, Hirano C, Yamakawa-Kobayashi K, Tomura S. Association found between the promoter region polymorphism in the apolipoprotein A-V gene and the serum triglyceride level in Japanese schoolchildren. *Hum Genet*. 2002.
37. Ribalta J, Figuera L, Fernandez-Ballart J, Vilella E, Castro Cabezas M, Masana L, Joven J. Newly identified apolipoprotein AV gene predisposes to high plasma triglycerides in familial combined hyperlipidemia. *Clin Chem*. 2002;48:1597-600.
38. Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*. 1986;44:283-92.
39. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res*. 1987;15:8125-48.
40. Kozak M. An analysis of vertebrate mRNA sequences: intimations of translational control. *J Cell Biol*. 1991;115:887-903.
41. Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng*. 1997;10:1-6.
42. Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer KA, Pachter LS, Dubchak I. VISTA : visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics*. 2000;16:1046-7.

FIGURE LEGENDS

FIGURE 1. Human, mouse and rabbit comparative sequence analysis of the *APOA1/C3/A4/A5* gene cluster. In each panel 30 kb of contiguous human sequence is illustrated horizontally. Above each panel arrows correspond to known genes and their orientation with each exon depicted by a box (gene names are indicated above each arrow). Dark blue boxes represent coding regions of exons and light blue boxes untranslated regions. The VISTA graphical plot displays the level of homology between human and the orthologous mouse sequence in the top portion and between human and rabbit in the bottom portion of each panel (http://pga.lbl.gov)[42]. Human sequence is represented on the x-axis and the percent similarity with the mouse or rabbit sequence is plotted on the y-axis (ranging from 50-100% identity). Blue shading indicate conserved exon sequences and red shading highlights conserved noncoding sequences. The *APOA5* gene was identified ~30 kb downstream of *APOA4* based on its high level of conservation between humans and mice.

FIGURE 2. Computational analyses of the human, mouse and chicken APOAV proteins. **(A)** Amino acid sequence alignment of the human, mouse and chicken APOAV proteins. **(B)** Amino acid percent similarity and identity between the human, mouse and chicken APOAV proteins. **(C)** Phylogenetic relationships of the human, mouse and chicken APOAV and APOAIV proteins.

FIGURE 3. Plasma triglyceride concentrations in a random sample of ~500 normo-lipidemic Caucasians separated by *APOA5* genotype (Berkeley Lipid Study Population)[22]. The *APOA5**2 and *APOA5*3* haplotypes were defined by the -1131C and c. 56G minor alleles, respectively.