Linkage and Association of Haplotypes at the *APOA1/C3/A4/A5* Gene Cluster to Familial Combined Hyperlipidemia.

Sophie Eichenbaum-Voline[1*], Michael Olivier[4*], Emma L Jones[1], Rossitza P. Naoumova[1], Bethan Jones[1], Brian Gau[4], Mary Seed[2], D. John Betteridge[5], David J Galton[6], Edward M Rubin[7], James Scott[1,3], Carol C Shoulders[1**] and Len A Pennacchio[7].

[1]Genomic and Molecular Medicine Group, Medical Research Council (MRC) Clinical Sciences Centre, Hammersmith Hospital and [2]Department of Cardiovascular Medicine, Charing Cross Hospital and [3]Genetics and Genomics Research Institute, Imperial College Faculty of Medicine, London, UK. [4]Human and Molecular Genetics Centre, Medical College of Wisconsin, Milwaukee, Wisconsin, USA. [5]Department of Medicine, Royal Free and University College Medical School, University College London, UK. [6]Department of Metabolism and Genetics, St Bartholomew's Hospital, London, UK. [7]Genome Sciences Department, Lawerence Berkeley National Laboratory, Berkeley, California, USA.

Running title: Independent *APOA1/C3/A4/A5* haplotypes confer susceptibility to FCHL

*these two authors contributed equally.

**Corresponding author: Dr Carol C Shoulders. Genomic and Molecular Medicine Group, MRC Clinical Sciences Centre, Imperial College Faculty of Medicine, Hammersmith Hospital, DuCane Rd, London W12 ONN UK. Phone: +(44)-208-383-4318; Fax: +(44)-208-383 2028.Email: carol.shoulders@csc.mrc.ac.uk.

Abstract

Combined hyperlipidemia (CHL) is a common disorder of lipid metabolism that leads to an increased risk of cardiovascular disease. The lipid profile of CHL is characterised by high levels of atherogenic lipoproteins and low levels of high-density-lipoprotein-cholesterol. Apolipoprotein (*APO*) *A5* is a newly discovered gene involved in lipid metabolism located within 30kbp of the *APOA1/C3/A4* gene cluster. Previous studies have indicated that sequence variants in this cluster are associated with increased plasma lipid levels. To establish whether variation at the *APOA5* gene contributes to the transmission of CHL, we performed linkage and linkage disequilibrium (LD) tests on a large cohort of families (n=128) with familial CHL (FCHL). The linkage data produced evidence for linkage of the *APOA1/C3/A4/A5* genomic interval to FCHL (NPL = 1.7, P = 0.042). The LD studies substantiated these data. Two independent rare alleles, $APOA5^{c.56G}$ and $APOC3^{c.386G}$ of this gene cluster were over-transmitted in FCHL (P = 0.004 and 0.007, respectively), and this was associated with a reduced transmission of the most common *APOA1/C3/A4/A5* haplotype (frequency 0.4425) to affected subjects (P = 0.013). The $APOA5^{c.56G}$ allele was associated with increased plasma triglyceride levels in FCHL probands, whereas the second, and independent, $APOC3^{c.386G}$ allele was associated with increased plasma triglyceride levels in FCHL pedigree founders. Thus, this allele (or an allele in LD) may mark a quantitative trait associated with FCHL, as well as representing a disease susceptibility locus for the condition. This study establishes that sequence variation in the *APOA1/C3/A4/A5* gene cluster contributes to the transmission of FCHL in a substantial proportion of affected families, and that these sequence variants may also contribute to the lipid abnormalities of the metabolic syndrome, which is present in up to 40% of persons with cardiovascular disease.

Introduction

Combined hyperlipidemia (CHL) affects 1-2% of individuals in Western society, and is present in up to 20% of patients with premature coronary heart disease ([1-4]). The term familial CHL (FCHL) was coined by ([1]) to describe a pattern of lipid abnormalities in 47 Seattle pedigrees, which was simultaneously observed in two other datasets ([2,5]). FCHL was originally described as a dominant disorder with incomplete penetrance until the third decade, and to primarily affect blood triglyceride levels, and secondarily cholesterol levels ([1]). This was proposed because plasma triglyceride levels were bimodally distributed in the first-degree relatives of affected probands that were above the age of 20 years, and less than one half of the offspring of affected family members were hyperlipidemic ([1]). However, more recent segregation analyses ([6-11]) and genome-wide linkage studies ([12,13]); Naoumova et al. submitted) have suggested a more complex inheritance pattern involving contributions from major genes and the environment.

The lipid profile in FCHL is characterised by increased plasma triglyceride and/or cholesterol levels, decreased HDL-cholesterol levels, and the presence of small-dense-low-density lipoprotein particles ([3,7,14]). Importantly, the lipid abnormalities of FCHL may also be present in persons with the metabolic syndrome (World Health Organisation Expert Committee 1985), which is a major cause of morbidity and mortality worldwide, and present in up to 40% of patients with premature coronary heart disease ([15]).

Apolipoprotein 5 (*APOA5*) is a newly identified gene involved in lipid metabolism ([16]), and represents a candidate for conferring susceptibility to FCHL. *APOA5* was discovered by a comparative sequence analysis of human and mouse genomic DNA sequences spanning the

APOA1/C3/A4 gene cluster ([16]), and subjected to functional analyses. In mice, disruption of the gene in "knock out" animals resulted in a four-fold increase in plasma triglyceride levels. Conversely, transgenic mice that independently over-expressed the human *APOA5* gene had a 66% reduction in plasma triglyceride levels. Furthermore, in humans single nucleotide polymorphisms (SNPs) across the *APOA5* locus were associated with increased plasma triglyceride levels in two independent datasets ([16]). For example in one dataset, healthy individuals with the minor allele at three neighbouring loci ($APOA5^{-1,131T>C}$, $APOA5^{IVS3+476G>T}$, $APOA5^{c.1,259T>C}$) had on average 20-30% higher plasma triglyceride levels than individuals homozygous for the major allele at these loci.

APOA5 resides ~30kbp downstream of the *APOA1/C3/A4/A5* cluster on chromosome 11q23, which has been extensively studied in mice ([17-21]) and humans ([22-26]). The *APOA1* gene contains a rare allele at the $APOA1^{-3,031C>T}$ locus, and this has been associated with combined hyperlipidemia (CHL) and/or FCHL in some studies ([27-34]). Likewise, a series of associations between rare alleles of the *APOC3* gene (e.g. the rare allele at the $APOC3^{c.386C>G}$ locus ([35]) and hypertriglyceridemia have been reported ([36-48]). However, several investigators have failed to detect linkage of the *APOA1/C3/A4* locus to any form of hyperlipidemia ([31,49-51]), which may reflect population specific differences in LD between markers and causal variants.

In the present study, we show that there is extensive LD between alleles of the *APOA1/C3/A4/A5* gene cluster in white British FCHL probands and pedigree founders. Furthermore, we establish that two distinct sequence variants (or alleles in LD) within this genomic interval contribute to the transmission of FCHL in a substantial proportion of affected white British families. These findings suggest that combinational therapies that specifically target the *APOA1/C3/A4/A5* gene cluster may provide a clinically significant

strategy for the treatment of the lipid abnormalities of FCHL, and the associated metabolic syndrome.

Materials and Methods

Ascertainment of Datasets

A description of the families used in this study, and phenotype cutoff values have been deposited at http://www.csc.mrc.ac.uk. In brief, white British probands were recruited through London-based tertiary referral specialised lipid clinics at Northwick Park Hospital, Hammersmith Hospital, Charing Cross Hospital, University College London Hospital and St Bartholomew's. FCHL probands were required to have cholesterol and triglyceride levels > age-sex-specific 95th and 90th percentiles, respectively, and a blood relative with either raised plasma cholesterol, triglyceride or both > age-sex-specific 90th percentile. In the absence of published British age and sex related percentile data for total cholesterol and triglyceride levels, the "first visit" percentile points of the Lipid Research Clinics ([52]) were used. Exclusion criteria for probands and family members were: < 16 years of age, other forms of genetic hyperlipidaemia (e.g. familial hypercholesterolemia), secondary hyperlipidaemia caused by either obesity (BMI > 30 kg/m^2), diabetes mellitus, untreated hypothyroidism, liver and kidney disease, alcohol intake > 21units/week, or drugs known to interfere with lipid metabolism. The Ethical Committees of the participating centres approved the study design. All participants gave written informed consent. Fasting levels of total cholesterol, triglyceride and HDL-cholesterol were determined by standard automated methods (Beckman Instruments, Inc, Galway, Ireland) using commercial kits, and interassay controls.

Genotyping

The forward and reverse oligonucleotide primers for PCR amplification of microsatellite marker D11S1998 and D11SAPOC3 were 5'AGCCATCAACTAGCTTTCCCT3' and 5'GAGGCACCAACAGATGGATG3', and 5'GAGTTGAGACTGCATTCCTCC3' and 5'GATGGCACCACTGCACTCCA3', respectively. PCR-products were pooled for analysis on an Applied Biosystems 3700 DNA sequencer, and sized with Genescan and Genotyper, version 2.0 software (ABI). SNP genotyping was performed with the PCR Invader assay (Third Wave Technologies, Madison, WI) as described previously ([16]);Olivier et al. submitted for publication. Accession numbers for entries in the SNPdb are: *APOA5*$^{58,892C>T}$, ss4472666; *APOA5*$^{c.56C>G}$, ss4383597; *APOA5*$^{c.-3A>G}$, ss4383596; *APOA5*$^{-1,131T>C}$, ss3199915 and *APOA5*$^{-12,238T>C}$, ss319916. Accession numbers for *APOA1*$^{-3,031C>T}$ and *APOC3*$^{c.386C>G}$ sequences are X67732 and NM_000040, respectively.

Statistical Analysis

Linkage analysis was performed on 86 FCHL families that contained an affected relative pair for two correlated, standard diagnostic criteria: 1) CHL phenotype, defined as cholesterol and triglyceride levels >90th age-sex-specific values; and 2) triglyceride trait, triglyceride >90th percentile age-sex-specific values. Data were analysed with GENEHUNTER-PLUS ([53,54]). This calculates a non parametric linkage score (NPL^{+}) to correct for the conservativeness of datasets which contain incomplete information on descent. Estimates of allele sharing were based on marker allele frequencies of pedigree founders (i.e. the "married-ins"), and equal weights assigned to each family. Nominal P-values were derived from the Gaussian distribution approximating the NPL distribution. We assessed the evidence for linkage and LD of individual SNPs and of haplotypes to FCHL with the pedigree disequilibrium test

(PDT), version 3.11, a transmission disequilibrium test for general pedigrees that tests for both linkage and LD ([55]). The sumPDT statistic gives more weight to larger families within the dataset, while the avePDT gives equal weight to all families ([56]). Haplotypes within families were reconstructed with Merlin, under the assumption of the most likely pattern of gene flow([57]). Families (n=4) that produced more than one haplotype solution were excluded from the PDT. Reconstructed parental haplotypes were only used in the PDT if complete (i.e. both haplotypes could be reconstructed). Founder haplotype frequencies were determined with the associated Fugue software ([58]), and used to compute the normalised disequilibrium statistic (D') values between two diallelic loci ([59]). D' was chosen as a measure of LD because it is relatively insensitive to allele frequencies ([60]). Values of D' are negative if the rare allele at one locus is associated with the common allele at the second, and positive when the rare allele at each locus are associated. Associations between individual SNP loci and triglyceride levels were performed on log-transformed values, adjusted for the effect of body mass index, age and sex. The Student's t-test was used to test the significance of differences between genotypes.

Haplotype Notation

Seven-locus haplotypes were constructed. Allele '1' refers to the common allele in the sequence (*APOA5*$^{58,892C>T}$, *APOA5*$^{c.56C>G}$, *APOA5*$^{c.-3A>G}$, *APOA5*$^{-1,131T>C}$, *APOA5*$^{-12,238T>C}$, *APOC3*$^{c.386C>G}$, *APOA1*$^{-3,031C>T}$), and '2' corresponds to the rare allele. The rare alleles at the *APOA5*$^{-1,131T>C}$ and *APOA5*$^{-12,238T>C}$ loci have previously been referred to as SNP 3 and 4, respectively ([16]). Alternative nomenclature for the *APOC3*$^{c.386G}$ and *APOA1*$^{-3,031T}$ alleles include the S2 allele of *APOC3* ([40]) and the X2 allele of *APOA1* ([27]), respectively.

Electronic Databases

Merlin: www.sph.umich.edu/csg/abecasis

http://www.ncbi.nlm.nih.gov/SNP/

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi

Results

Increased Transmission of Rare Alleles at the APOA1/C3/A4/A5 Genomic Interval in FCHL

To establish the contribution of sequence variation at the *APOA1/C3/A4/A5* genomic interval to FCHL susceptibility, we performed linkage and LD tests on a cohort of white British families (Table 1). For the linkage test, 86 extended families were genotyped with two markers: D11SAPOC3 which resides within the third intron of *APOC3*, and D11S1998, which is located approximately 1.7 Mbp downstream of *APOA5* (Figure 1). The families contained 177 and 270 affected relative pairs for the CHL phenotype and the triglyceride trait of FCHL, respectively (Table 1). The linkage analysis produced a NPL^{+} value of 1.72 (P = 0.042) at the D11SAPOC3 genetic marker for the triglyceride trait of FCHL, and this was attributable to a positive NPL value in a subset (n=36) of the 86 families.

To substantiate the evidence for linkage of the *APOA1/C3/A4/A5* genomic interval to FCHL, we performed a PDT on 115 white British families (Table 1) using seven SNPs that span an interval of 108kbp (Figure 1). The analysis used information from nuclear families comprising an affected offspring and two parents, one of whom had to be heterozygous for the marker under investigation, and discordant sibpairs who had different marker genotypes.

The PDT produced evidence for increased transmission of the rare alleles at the *APOA5*$^{c.56C>G}$ and *APOC3*$^{c.386C>G}$ loci to affected subjects (Table 2). Thus, the *APOA5*$^{c.56G}$ and *APOC3*$^{c.386G}$ alleles were respectively transmitted 1.95- and 1.45-fold more frequently to affected family members with the triglyceride trait of FCHL than to unaffected family members. The corresponding values for the CHL trait were 1.95 and 1.33, respectively. The rare alleles at the *APOA5*$^{c.-3A>G}$ and *APOA5*$^{-1,131T>C}$ loci were also transmitted 1.28 and 1.40 fold-more frequently to affected individuals with the CHL phenotypes of FCHL than to unaffected individuals (P = 0.039 and 0.033) (Table 2).

The rare alleles at the *APOA5*$^{c.56C>G}$, *APOC3*$^{c.386C>G}$, *APOA5*$^{-1,131T>C}$ and *APOA5*$^{c.-3A>G}$ loci were also present at increased frequencies in FCHL probands versus pedigree founders (i.e. "married ins") (Table 3). For example, the rare allele at the *APOA5*$^{c.56C>G}$ locus was present in 21% of the probands compared to 13% of the normolipidemic pedigree founders, whereas the rare allele at the *APOC3*$^{c.386C>G}$ locus was present in 29% of the probands and 15% of the normolipidemic pedigree founders. Importantly, the results from this case-control study and the PDT complemented each other. For the example, the frequencies of the rare alleles at the *APOA5*$^{c.56C>G}$, *APOA5*$^{c.-3A>G}$, *APOA5*$^{-1,131T>C}$ and *APOC3*$^{c.386C>G}$ loci in FCHL probands and affected FCHL sibs were remarkably similar (i.e. 0.1200, 0.1144, 0.1111 and 0.1486, respectively versus 0.1114, 0.1156, 0.1296 and 0.1566) (Table 2). Likewise, the frequencies of the rare alleles at the *APOA5*$^{c.56C>G}$ and *APOC3*$^{c.386C>G}$ loci were similar in the pedigree founders and unaffected sibs (0.0694 and 0.1024, respectively versus 0.0511 and 0.1048) (Table 2). Thus, the case-control data support the evidence that the *APOA5*$^{c.56G}$ and *APOC3*$^{c.386G}$ alleles (or alleles in LD) are preferentially transmitted in FCHL.

Probands with the rare allele *APOA5*$^{c.56G}$ had higher mean triglyceride levels than probands homozygous for the major allele at this locus (Table 3), and this was particularly

evident in those individuals that were homozygous for this rare allele (n=5). Thus, mean plasma triglyceride levels in probands with the *APOA5*$^{c.56G}$ allele were on average 2.2 fold higher than in probands homozygous for the *APOA5*$^{c.56C}$ allele, and ~1.8 fold higher relative to the heterozygote probands. By contrast, the *APOA5*$^{c.56G}$ allele had no major impact on triglyceride levels in heterozygote pedigree founders (Table 3), and this was also the case when all individuals with the rare allele at the *APOC3*$^{c.386C>G}$ locus were excluded from the analyses (data not shown). Only one pedigree founder was homozygous for the *APOA5*$^{c.56G}$ allele, precluding an assessment of the impact of the homozygous state of this allele (or an allele in LD) on plasma triglyceride in the pedigree founders of white British families with FCHL.

The *APOC3*$^{c.386G}$ allele (or an allele in LD) had a modest impact on triglyceride levels in probands and pedigree founders (Table 3). On average pedigree founders with the *APOC3*$^{c.386G}$ allele had plasma triglyceride levels that were 31% higher than pedigree founders without this allele (P = 0.001), and this increased to a value of 38% (P = 0.001) when we considered only those individuals with the common allele at the *APOA5*$^{c.56C>G}$ locus (data not shown). Similar increases in plasma triglyceride levels were also observed in pedigree founders with the rare alleles at the *APOA5*$^{-1,131T>C}$ and *APOA5*$^{c.-3A>G}$ loci (data not shown). In a complementary analysis, increased frequencies of these rare alleles were observed in pedigree founders that had plasma cholesterol and triglyceride levels >75th percentile age-sex-specific values relative to the rest (Table 3). This trend was not observed for the rare allele at the *APOA5*$^{c.56C>G}$ locus, indicating that this allele resides on a different *APOA1/C3/A4/A5* haplotype than the rare alleles at the *APOC3*$^{c.386C>G}$ and *APOA5*$^{-1,131T>C}$ and *APOA5*$^{c.-3A>G}$ loci.

The Rare Alleles at the APOA5$^{c.56C>G}$ and APOC3$^{c.386C>G}$ Loci Define Different Haplotypes

To examine the extent of LD between alleles in the *APOA5* and *APOA1/C3/A4* genomic interval, seven-locus haplotypes were constructed in the pedigree founders. Alleles at all loci were in LD (Table 4), and 67% of the chromosomes in the dataset were accounted for by two major haplotypes (Table 5, Figure 2). The most common haplotype, designated *APOA1/C3/A4/A5*1* (111111), was estimated to be present at a frequency of 0.4425 in the pedigree founders, while the second most common, designated *APOA1/C3/A4/A5*2* (1111211), was present at a frequency of 0.2246. Pertinent to this study, the rare allele at the *APOA5*$^{c.56C>G}$ locus was rarely observed on the same haplotype that contained the rare alleles at the *APOC3*$^{c.386C>G}$, *APOA5*$^{-1,131T>C}$ and *APOA5*$^{c.-3A>G}$ loci, while the rare alleles at the *APOA5*$^{-1,131T>C}$ and *APOA5*$^{c.-3A>G}$ loci were essentially restricted to three haplotypes that contained the rare allele at the *APOC3*$^{c.386C>G}$ locus (Table 5, Figure 2). Thus, these data indicate that the *APOA5*$^{c.56G}$ and *APOC3*$^{c.386G}$ alleles define independent haplotypes (Figure 2). Furthermore, that the *APOC3*$^{c.386G}$, *APOA5*$^{-1,131C}$ and *APOA5*$^{c.-3G}$ alleles are in strong LD in the white British population.

Distorted Transmission of Haplotypes at the APOA1/C3/A4/A5 Genomic Interval in FCHL

To further test/delineate/examine the for preferential transmission of the *APOA5*$^{c.56G}$ and *APOC3*$^{c.386G}$ alleles in FCHL (Table 2), we repeated the PDT in our families with haplotype data for the *APOA1/C3/A4/A5* genomic interval. The transmission of haplotypes to affected members was distorted at a global level (P = 0.013), and this was largely attributable to the reduced transmission of the common *APOA1/C3/A4/A5*1* (1111111) haplotype to affected FCHL subjects, and increased transmission of two distinct haplotypes, namely *APOA1/C3/A4/A5*4* (1211111) and *APOA1/C3/A4/A5*7* (2122121) (Table 5). Haplotype

*APOA1/C3/A4/A5*4* (1211111) contains the rare allele at the *APOA5*$^{c.56C>G}$ locus, and was transmitted 2.6 more frequently to an affected sib than to a non-affected sib, while *APOA1/C3/A4/A5*7* (12122121), which contains the rare alleles at the *APOC3*$^{c.386C>G}$, *APOA5*$^{-1,131T>C}$ and *APOA5*$^{c.-3A>G}$ loci, was transmitted 1.3 fold more frequently to affected sibs (Table 5).

The finding that haplotypes that contained the rare alleles at the *APOA5*$^{c.56C>G}$, *APOC3*$^{c.386C>G}$, *APOA5*$^{-1,131T>C}$ and *APOA5*$^{c.-3A>G}$ loci were over-transmitted in families with FCHL raised the issue ofwhether they might represent causal sequence variants. To address this issue, we examined the additional haplotypes that contained the minor allele for either *APOA5*$^{c.56C>G}$ or *APOC3*$^{c.386C>G}$ to determine if they were also over-transmitted in FCHL. The most common haplotypes containing the *APOA5*$^{c.56G}$ allele, namely *APOA1/C3/A4/A5*3* (1211112), and *APOA1/C3/A4/A5*5* (1211121) [we need to mention that these are the most common c.56G haplotypes when we exclude hap*4. The sentence above may make what we are trying to say clearer. also should we remove the following two snps from the first sentence of the paragraph , *APOA5*$^{-1,131T>C}$ and *APOA5*$^{c.-3A>G}$ since we don't address if they are causative? It seems we are only testing two of the snps, one for av and one for c3] (Figure 2) were transmitted to affected subjects at increased frequencies (Table 5), but these increases did not reach statistical significance (P = 0.540 and 0.086). These data may indicate that there was insufficient power in our dataset to obtain a significant result or alternatively that the *APOA5*$^{c.56G}$ allele is not a causal sequence variant for FCHL. Simulation studies based on 5,000 datasets, each comprising 250 families, have shown that the power of the PDT is poor for alleles/haplotypes, such as the *APOA1/C3/A4/A5*3* and *APOA1/C3/A4/A5*5* haplotypes, with frequencies below 0.15 (Martin et al. 2000).

Finally, the PDT produced no evidence for increased transmission of the second most common *APOC3*$^{c.386G}$ haplotype, designated *APOA1/C3/A4/A5*8* (1111121) (Figure 2) in FCHL (Table 5). A defining difference between this haplotype and the more common *APOA1/C3/A4/A5*7* haplotype relates to the alleles at the *APOA5*$^{c.-3A>G}$ and *APOA5*$^{-1,131T>C}$ loci (Figure 2). Haplotype *APOA1/C3/A4/A5*7* contains the rare alleles at the *APOA5*$^{c.-3A>G}$ and *APOA5*$^{-1,131T>C}$ loci, and was over-transmitted in our FCHL families as described above. In contrast, *APOA1/C3/A4/A5*8* contains the common alleles at these *APOA5* loci, and was not over-transmitted in FCHL, suggesting that the *APOC3*$^{c.386G}$ allele itself may not represent a causal sequence variant for this condition. Alternatively, the *APOC3*$^{c.386G}$ allele may only confer susceptibility to FCHL in the context of an *APOA1/C3/A4/A5* haplotype that contains the rare alleles at the *APOA5*$^{c.-3A>G}$ and *APOA5*$^{-1,131T>C}$ loci.

Discussion

In the present study we have performed genetic analyses in a substantial cohort of white British families with FCHL to derive information on the contribution of sequence variation at *APOA5* in the transmission of this condition, and to examine these data in the context of variation at the linked *APOA1/C3/A4* gene complex. The results establish that there is extensive allelic association in the *APOA1/C3/A4/A5* cluster, and that two independent rare alleles (i.e. *APOA5*$^{c.56G}$ and *APOC3*$^{c.386G}$) contained within this genomic interval are linked to an increased genetic risk (should we say this explicitly? or rather say that they are simply over-transmitted in FCHL. The data support/suggest they are risk factors but we haven't provded it) of FCHL. Specifically, the *APOA5*$^{c.56G}$ marks a disease susceptibility locus for FCHL, while the second independent allele (*APOC3*$^{c.386G}$) marks a locus that influences a

quantitative trait associated with the disorder.[again you might consider using milder words than marks. supports/suggests/ provides evidence ect.]

The evidence that allelic variation at the *APOA1/C3/A4/A5* genomic interval contributes to the genetic risk of FCHL derives from the results of a traditional-linkage based analysis that tests for excess allele sharing in affected relative pairs, and the PDT, that tests for both linkage and association in general pedigrees. The PDT was performed on 115 families with seven individual SNPs and seven-locus haplotypes, and produced comparable results. The SNP data set comprised genetic information from 88 trios and 307 discordant sibpairs, and produced evidence that the $APOA5^{c.56G}$ (P = 0.004) and $APOC3^{c.386G}$ (P = 0.007) alleles were over-transmitted in FCHL {yes, i like this tone}. The haplotype analysis extracted genotype data from 153 trios and 280 discordant sibpairs, and included only those individual where the phase between markers could be reliably assigned. Thus, a proportion of families used in the first PDT produced no data for the second PDT, and visa versa. This observation adds credence for the evidence of increased transmission in FCHL of haplotype *APOA1/C3/A4/A5*4*, which contains the $APOA5^{c.56G}$ allele and of haplotype *APOA1/C3/A4/A5*7* that bears the $APOC3^{c.386G}$ allele, and of reduced transmission of the most common *APOA1/C3/A4/A5*1* haplotype, in this condition.[this sentence is a bit clunky. What exactly are we trying to say? I don't think we want to add credence to evidence. Should we remove or simplify the sentence?]

The present study provides compelling evidence that the rare $APOA5^{c.56G}$ allele (or an allele in LD) confers susceptibility to FCHL, rather than influencing a quantitative trait associated with the disorder. This result is consistent with the genetic analyses of Goldstein et al (1973) and the complex segregation analyses of Cullen et al. (1994), which allowed for a major gene(s) and locus heterogeneity to produce the best genetic model for FCHL. In the

present study, we found that the *APOA5*$^{c.56G}$ allele was present in a subset of FCHL families (45 out of 128), and that this allele was transmitted more frequently to affected individuals than to unaffected individuals. The rare *APOA5*$^{c.56G}$ allele was also associated with markedly higher triglyceride levels in FCHL probands, and significantly under-represented in the pedigree founders of white British families with FCHL. This finding may partly explain the lack of association of the *APOA5*$^{c.56G}$ allele with plasma triglyceride levels in FCHL pedigree founders, which is at variance with the results found in community-based samples in other populations (Pennacchio et al. submitted). In the current study, FCHL probands homozygous for the *APOA5*$^{c.56G}$ allele had some of the highest plasma triglyceride levels in our dataset. In previous studies ([28,29,61]), similar results were obtained for the *APOA1*$^{-3,031T}$ allele, and this we suggest may be attributable to the significant LD between this allele and the *APOA5*$^{c.56G}$allele. In one study, the *APOA1*$^{-3,031T}$ allele displayed linkage to FCHL, in a series of families ascertained through a proband with this same allele ([29]). In another study, FCHL probands with the *APOA1*$^{-3,031T}$ allele had mean plasma cholesterol and triglyceride levels that were respectively, two- and eight-fold higher than probands without this allele ([32]). In a third study involving patients with peripheral vascular disease, all five patients homozygous for the *APOA1*$^{-3,031T}$ allele had CHL ([28]).

The proposition that the rare alleles at the *APOC3*$^{c.386C>G}$, *APOA5*$^{c.-3A>G}$ and *APOA5*$^{-1,131T>C}$ loci mark an independent locus affecting the triglyceride component of FCHL as a quantitative trait is consistent with the results of the present and previous studies ([22,23,32,42,45,62,63]). The association of the *APOC3*$^{c.386G}$ allele with hypertriglyceridemia (i.e. plasma triglyceride > 220mg/dl, cholesterol < 240mg/dl), CHL and/or severe hypertriglyceridemia (i.e. plasma triglyceride > 1000mg/dl) has also been amply demonstrated ([27,32,36-38,40,41,43,44,48,64,64]). The present study extend these data to provide the first evidence

that the *APOC3*$^{c.386G}$ allele marks a sequence variant for a quantitative trait associated with FCHL in a substantial subset of white British families. Furthermore, that this sequence variant, due to extensive LD across the APOA1/C3/A4/A5 genomic interval, may reside at some considerable distance from *APOC3*, and perturb the expression of either *APOA5, APOC3,* or both. This would accord with the results of transgenic mice experiments. Manipulations that perturb the expression of either these genes influence plasma triglyceride levels in a quantitative manner ([16,17,21]).

In summary, the current study confirms that FCHL is indeed a heterogeneous disorder ([8]) involving multiple genetic determinants, and that two of these determinants reside within the *APOA1/C3/A4/A5* genomic interval. What remains to be established is whether the determinants affect lipid levels through their effect on *APOA5*, *APOC3*, or both. Additional genetic studies aimed at resolving this issue could provide specific therapeutic targets for future combinational drug therapies for the treatment of FCHL, and the associated metabolic syndrome.

Acknowledgements

The work was supported by the British Heart Foundation (Grants PG/98159, PG/2001015), the Medical Research Council and the Hammersmith Hospital NHS Trust, London UK, the U.S. Department of Energy (contract number DE-AC03-76SF00098(EMR) through the NIH-NHLBI program for Genomic Applications (Grant HL66681(EMR)), and a Hollaender Distinguished Postdoctoral award to LAP. We are also indebted to all of the patients and families who have participated in this study, Mrs Saro Nithyananathan, Susan Earl-Mitchell, Ms Stephanie A Bonney, Ms Regina Cole, Sister Clare K.Y. Neuwirth, Ms Hetal N Patel, Mr Daniel Savic and Ms Linda Smith for key technical support, and Ms Tracey Greiner for excellent secretarial assistance. We also thank Dr Martin Farrall (Statistical Genetics Group, Wellcome Trust Centre for Human Genetics, University of Oxford, UK) and Dr Goncalo Abecasis (J. Mendez Hospital, Buenos Aires, Argentina) for help and advice with the software programs, Fugue and Merlin and Drs Meredyth Bass and Eden R. Martin (Center for Human Genetics, Duke University Medical Center, Durham, North Carolina, USA) for help with the PDT algorithm.

References

1. Goldstein,J.L., Schrott,H.G., Hazzard,W.R., Bierman,E.L., and Motulsky,A.G. (1973) Hyperlipidemia in coronary heart disease. II. Genetic analysis of lipid levels in 176 families and delineation of a new inherited disorder, combined hyperlipidemia. *J.Clin Invest*, **52**,1544-1568.

2. Nikkila,E.A., Aro,A. (1973) Family study of serum lipids and lipoproteins in coronary heart-disease. *Lancet*, **1**,954-959.

3. Brunzell,J.D., Albers,J.J., Chait,A., Grundy,S.M., Groszek,E., and McDonald,G.B. (1983) Plasma lipoproteins in familial combined hyperlipidemia and monogenic familial hypertriglyceridemia. *J.Lipid Res.*, **24**,147-155.

4. Grundy,S.M., Greenland,P., Herd,A., Huebsch,J.A., Jones,R.J., Mitchell,J.H., and Schlant,R.C. (1987) Cardiovascular and risk factor evaluation of healthy American adults. A statement for physicians by an Ad Hoc Committee appointed by the Steering Committee, American Heart Association. *Circulation*, **75**,1340A-1362A.

5. Rose,H.G., Kranz,P., Weinstock,M., Juliano,J., and Haft,J.I. (1973) Inheritance of combined hyperlipoproteinemia: evidence for a new lipoprotein phenotype. *Am J.Med.*, **54**,148-160.

6. Austin,M.A., Brunzell,J.D., Fitch,W.L., and Krauss,R.M. (1990) Inheritance of low density lipoprotein subclass patterns in familial combined hyperlipidemia. *Arteriosclerosis*, **10**,520-530.

7. Austin,M.A., Horowitz,H., Wijsman,E., Krauss,R.M., and Brunzell,J. (1992) Bimodality of plasma apolipoprotein B levels in familial combined hyperlipidemia. *Atherosclerosis*, **92**,67-77.

8. Cullen,P., Farren,B., Scott,J., and Farrall,M. (1994) Complex segregation analysis provides evidence for a major gene acting on serum triglyceride levels in 55 British families with familial combined hyperlipidemia. *Arterioscler.Thromb.*, **14**,1233-1249.

9. Jarvik,G.P., Brunzell,J.D., Austin,M.A., Krauss,R.M., Motulsky,A.G., and Wijsman,E. (1994) Genetic predictors of FCHL in four large pedigrees. Influence of ApoB level major locus predicted genotype and LDL subclass phenotype. *Arterioscler.Thromb.*, **14**,1687-1694.

10. Bredie,S.J., Vogelaar,J.M., Demacker,P.N., and Stalenhoef,A.F. (1996) Apolipoprotein E polymorphism influences lipid phenotypic expression, but not the low density lipoprotein subfraction distribution in familial combined hyperlipidemia. *Atherosclerosis*, **126**,313-324.

11. Juo,S.H., Han,Z., Smith,J.D., Colangelo,L., and Liu,K. (2000) Common polymorphism in promoter of microsomal triglyceride transfer protein gene influences cholesterol, ApoB, and triglyceride levels in young african american men:

results from the coronary artery risk development in young adults (CARDIA) study. *Arterioscler.Thromb.Vasc.Biol.*, **20**,1316-1322.

12. Pajukanta,P., Terwilliger,J.D., Perola,M., Hiekkalinna,T., Nuotio,I., Ellonen,P., Parkkonen,M., Hartiala,J., Ylitalo,K., Pihlajamaki,J. et al. (1999) Genomewide scan for familial combined hyperlipidemia genes in finnish families, suggesting multiple susceptibility loci influencing triglyceride, cholesterol, and apolipoprotein B levels. *Am J.Hum.Genet.*, **64**,1453-1463.

13. Aouizerat,B.E., Allayee,H., Cantor,R.M., Davis,R.C., Lanning,C.D., Wen,P.Z., Dallinga-Thie,G.M., de Bruin,T.W., Rotter,J.I., and Lusis,A.J. (1999) A genome scan for familial combined hyperlipidemia reveals evidence of linkage with a locus on chromosome 11. *Am J.Hum.Genet.*, **65**,397-412.

14. Porkka,K.V., Nuotio,I., Pajukanta,P., Ehnholm,C., Suurinkeroinen,L., Syvanne,M., Lehtimaki,T., Lahdenkari,A.T., Lahdenpera,S., Ylitalo,K. et al. (1997) Phenotype expression in familial combined hyperlipidemia. *Atherosclerosis*, **133**,245-253.

15. Carlson,L.A., Bottiger,L.E. (1980) Triglycerides and coronary heart disease. *N.Engl.J.Med.*, **303**,1061-1062.

16. Pennacchio,L.A., Olivier,M., Hubacek,J.A., Cohen,J.C., Cox,D.R., Fruchart,J.C., Krauss,R.M., and Rubin,E.M. (2001) An apolipoprotein influencing triglycerides in humans and mice revealed by comparative sequencing. *Science*, **294**,169-173.

17. Ito,Y., Azrolan,N., O'Connell,A., Walsh,A., and Breslow,J.L. (1990) Hypertriglyceridemia as a result of human apo CIII gene expression in transgenic mice. *Science*, **249**,790-793.

18. Aalto-Setala,K., Fisher,E.A., Chen,X., Chajek-Shaul,T., Hayek,T., Zechner,R., Walsh,A., Ramakrishnan,R., Ginsberg,H.N., and Breslow,J.L. (1992) Mechanism of hypertriglyceridemia in human apolipoprotein (apo) CIII transgenic mice. Diminished very low density lipoprotein fractional catabolic rate associated with increased apo CIII and reduced apo E on the particles. *J.Clin Invest*, **90**,1889-1900.

19. Chen,M., Breslow,J.L., Li,W., and Leff,T. (1994) Transcriptional regulation of the apoC-III gene by insulin in diabetic mice: correlation with changes in plasma triglyceride levels. *J.Lipid Res.*, **35**,1918-1924.

20. Maeda,N., Li,H., Lee,D., Oliver,P., Quarfordt,S.H., and Osada,J. (1994) Targeted disruption of the apolipoprotein C-III gene in mice results in hypotriglyceridemia and protection from postprandial hypertriglyceridemia. *J.Biol.Chem.*, **269**,23610-23616.

21. Masucci-Magoulas,L., Goldberg,I.J., Bisgaier,C.L., Serajuddin,H., Francone,O.L., Breslow,J.L., and Tall,A.R. (1997) A mouse model with features of familial combined hyperlipidemia. *Science*, **275**,391-394.

22. Ordovas,J.M., Civeira,F., Genest,J., Jr., Craig,S., Robbins,A.H., Meade,T., Pocovi,M., Frossard,P.M., Masharani,U., Wilson,P.W. et al. (1991) 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*, **87**,75-86.

23. Hegele,R.A., Brunt,J.H., and Connelly,P.W. (1995) Multiple genetic determinants of variation of plasma lipoproteins in Alberta Hutterites. *Arterioscler.Thromb.Vasc.Biol.*, **15**,861-871.

24. Zerba,K.E., Ferrell,R.E., and Sing,C.F. (1998) Genetic structure of five susceptibility gene regions for coronary artery disease: disequilibria within and among regions. *Hum.Genet.*, **103**,346-354.

25. Vu-Dac,N., Gervois,P., Torra,I.P., Fruchart,J.C., Kosykh,V., Kooistra,T., Princen,H.M., Dallongeville,J., and Staels,B. (1998) Retinoids increase human apo C-III expression at the transcriptional level via the retinoid X receptor. Contribution to the hypertriglyceridemic action of retinoids. *J.Clin Invest*, **102**,625-632.

26. Waterworth,D.M., Ribalta,J., Nicaud,V., Dallongeville,J., Humphries,S.E., and Talmud,P. (1999) ApoCIII gene variants modulate postprandial response to both glucose and fat tolerance tests. *Circulation*, **99**,1872-1877.

27. Hayden,M.R., Kirk,H., Clark,C., Frohlich,J., Rabkin,S., McLeod,R., and Hewitt,J. (1987) DNA polymorphisms in and around the Apo-A1-CIII genes and genetic hyperlipidemias. *Am J.Hum.Genet.*, **40**,421-430.

28. Monsalve,M.V., Young,R., Wiseman,S.A., Dhamu,S., Powell,J.T., Greenhalgh,R.M., and Humphries,S.E. (1989) Study of DNA polymorphisms of the apolipoprotein AI-CIII-AIV gene cluster in patients with peripheral arterial disease. *Clin Sci.(Lond)*, **76** ,221-228.

29. Wojciechowski,A.P., Farrall,M., Cullen,P., Wilson,T.M., Bayliss,J.D., Farren,B., Griffin,B.A., Caslake,M.J., Packard,C.J., Shepherd,J. et al. (1991) Familial combined hyperlipidaemia linked to the apolipoprotein AI-CII- AIV gene cluster on chromosome 11q23-q24. *Nature*, **349**,161-164.

30. Tybjaerg-Hansen,A., Nordestgaard,B.G., Gerdes,L.U., Faergeman,O., and Humphries,S.E. (1993) Genetic markers in the apo AI-CIII-AIV gene cluster for combined hyperlipidemia, hypertriglyceridemia, and predisposition to atherosclerosis. *Atherosclerosis*, **100**,157-169.

31. Xu,C.F., Talmud,P., Schuster,H., Houlston,R., Miller,G., and Humphries,S. (1994) Association between genetic variation at the APO AI-CIII-AIV gene cluster and familial combined hyperlipidaemia. *Clin Genet.*, **46**,385-397.

32. Dallinga-Thie,G.M., Bu,X.D., Linde-Sibenius,T.M., Rotter,J.I., Lusis,A.J., and de Bruin,T.W. (1996) Apolipoprotein A-I/C-III/A-IV gene cluster in familial combined hyperlipidemia: effects on LDL-cholesterol and apolipoproteins B and C- III. *J.Lipid Res.*, **37**,136-147.

33. Ribalta,J., La Ville,A.E., Vallve,J.C., Humphries,S., Turner,P.R., and Masana,L. (1997) A variation in the apolipoprotein C-III gene is associated with an increased

number of circulating VLDL and IDL particles in familial combined hyperlipidemia. *J.Lipid Res.*, **38**,1061-1069.

34. Groenendijk,M., Cantor,R.M., Funke,H., and Dallinga-Thie,G.M. (2001) Two newly identified SNPs in the APO AI-CIII intergenic region are strongly associated with familial combined hyperlipidaemia. *Eur.J.Clin Invest*, **31**,852-859.

35. Karathanasis,S.K., Zannis,V.I., and Breslow,J.L. (1985) Isolation and characterization of cDNA clones corresponding to two different human apoC-III alleles. *J.Lipid Res.*, **26**,451-456.

36. Rees,A., Shoulders,C.C., Stocks,J., Galton,D.J., and Baralle,F.E. (1983) DNA polymorphism adjacent to human apoprotein A-1 gene: relation to hypertriglyceridaemia. *Lancet*, **1**,444-446.

37. Aalto-Setala,K., Kontula,K., Sane,T., Nieminen,M., and Nikkila,E. (1987) DNA polymorphisms of apolipoprotein A-I/C-III and insulin genes in familial hypertriglyceridemia and coronary heart disease. *Atherosclerosis*, **66**,145-152.

38. Henderson,H.E., Landon,S.V., Michie,J., and Berger,G.M. (1987) Association of a DNA polymorphism in the apolipoprotein C-III gene with diverse hyperlipidaemic phenotypes. *Hum.Genet.*, **75**,62-65.

39. Paulweber,B., Friedl,W., Krempler,F., Humphries,S.E., and Sandhofer,F. (1988) Genetic variation in the apolipoprotein AI-CIII-AIV gene cluster and coronary heart disease. *Atherosclerosis*, **73**,125-133.

40. Shoulders,C.C., Ball,M.J., and Baralle,F.E. (1989) Variation in the apo AI/CIII/AIV gene complex: its association with hyperlipidemia. *Atherosclerosis*, **80**,111-118.

41. Tas,S. (1989) Strong association of a single nucleotide substitution in the 3'-untranslated region of the apolipoprotein-CIII gene with common hypertriglyceridemia in Arabs. *Clin Chem.*, **35**,256-259.

42. Shoulders,C.C., Harry,P.J., Lagrost,L., White,S.E., Shah,N.F., North,J.D., Gilligan,M., Gambert,P., and Ball,M.J. (1991) Variation at the apo AI/CIII/AIV gene complex is associated with elevated plasma levels of apo CIII. *Atherosclerosis*, **87**,239-247.

43. Dammerman,M., Sandkuijl,L.A., Halaas,J.L., Chung,W., and Breslow,J.L. (1993) An apolipoprotein CIII haplotype protective against hypertriglyceridemia is specified by promoter and 3' untranslated region polymorphisms. *Proc.Natl.Acad.Sci.U.S.A*, **90**,4562-4566.

44. Zeng,Q., Dammerman,M., Takada,Y., Matsunaga,A., Breslow,J.L., and Sasaki,J. (1995) An apolipoprotein CIII marker associated with hypertriglyceridemia in Caucasians also confers increased risk in a west Japanese population. *Hum.Genet.*, **95**,371-375.

45. Shoulders,C.C., Grantham,T.T., North,J.D., Gaspardone,A., Tomai,F., de Fazio,A., Versaci,F., Gioffre,P.A., and Cox,N.J. (1996) Hypertriglyceridemia and the apolipoprotein CIII gene locus: lack of association with the variant insulin response element in Italian school children. *Hum.Genet.*, **98**,557-566.

46. Surguchov,A.P., Page,G.P., Smith,L., Patsch,W., and Boerwinkle,E. (1996) Polymorphic markers in apolipoprotein C-III gene flanking regions and hypertriglyceridemia. *Arterioscler.Thromb.Vasc.Biol.*, **16**,941-947.

47. Hoffer,M.J., Sijbrands,E.J., De Man,F.H., Havekes,L.M., Smelt,A.H., and Frants,R.R. (1998) Increased risk for endogenous hypertriglyceridaemia is associated with an apolipoprotein C3 haplotype specified by the SstI polymorphism. *Eur.J.Clin Invest*, **28**,807-812.

48. Sijbrands,E.J., Hoffer,M.J., Meinders,A.E., Havekes,L.M., Frants,R.R., Smelt,A.H., and De Knijff,P. (1999) Severe hyperlipidemia in apolipoprotein E2 homozygotes due to a combined effect of hyperinsulinemia and an SstI polymorphism. *Arterioscler.Thromb.Vasc.Biol.*, **19**,2722-2729.

49. Marcil,M., Boucher,B., Gagne,E., Davignon,J., Hayden,M., and Genest,J., Jr. (1996) Lack of association of the apolipoprotein A-I-C-III-A-IV gene XmnI and SstI polymorphisms and of the lipoprotein lipase gene mutations in familial combined hyperlipoproteinemia in French Canadian subjects. *J.Lipid Res.*, **37**,309-319.

50. Wijsman,E.M., Brunzell,J.D., Jarvik,G.P., Austin,M.A., Motulsky,A.G., and Deeb,S.S. (1998) Evidence against linkage of familial combined hyperlipidemia to the apolipoprotein AI-CIII-AIV gene complex. *Arterioscler.Thromb.Vasc.Biol.*, **18**,215-226.

51. Tahvanainen,E., Pajukanta,P., Porkka,K., Nieminen,S., Ikavalko,L., Nuotio,I., Taskinen,M.R., Peltonen,L., and Ehnholm,C. (1998) Haplotypes of the ApoA-I/C-III/A-IV gene cluster and familial combined hyperlipidemia. *Arterioscler.Thromb.Vasc.Biol.*, **18**,1810-1817.

52. Heiss,G., Tamir,I., Davis,C.E., Tyroler,H.A., Rifkand,B.M., Schonfeld,G., Jacobs,D., and Frantz,I.D., Jr. (1980) Lipoprotein-cholesterol distributions in selected North American populations: the lipid research clinics program prevalence study. *Circulation*, **61**,302-315.

53. Kruglyak,L., Daly,M.J., Reeve-Daly,M.P., and Lander,E.S. (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J.Hum.Genet.*, **58**,1347-1363.

54. Kong,A., Cox,N.J. (1997) Allele-sharing models: LOD scores and accurate linkage tests. *Am J.Hum.Genet.*, **61**,1179-1188.

55. Martin,E.R., Monks,S.A., Warren,L.L., and Kaplan,N.L. (2000) A test for linkage and association in general pedigrees: the pedigree disequilibrium test. *Am J.Hum.Genet.*, **67**,146-154.

56. Martin,E.R., Bass,M.P., and Kaplan,N.L. (2001) Correcting for a potential bias in the pedigree disequilibrium test. *Am J.Hum.Genet.*, **68**,1065-1067.

57. Abecasis,G.R., Martin,R., and Lewitzky,S. (2001) Estimation of haplotype frequencies from diploid data. *American Journal of Human Genetics*, **69**,114.

58. Abecasis,G.R., Cherny,S.S., Cookson,W.O., and Cardon,L.R. (2002) Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat.Genet.*, **30**,97-101.

59. Lewontin,R.C. (1988) On measures of gametic disequilibrium. *Genetics*, **120**,849-852.

60. Devlin,B., Risch,N. (1995) A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics*, **29**,311-322.

61. Dallinga-Thie,G.M., Linde-Sibenius,T.M., Rotter,J.I., Cantor,R.M., Bu,X., Lusis,A.J., and de Bruin,T.W. (1997) Complex genetic contribution of the Apo AI-CIII-AIV gene cluster to familial combined hyperlipidemia. Identification of different susceptibility haplotypes. *J.Clin Invest*, **99**,953-961.

62. Anderson,R.A., Burns,T.L., Lee,J., Swenson,D., and Bristow,J.L. (1989) Restriction fragment length polymorphisms associated with abnormal lipid levels in an adolescent population. *Atherosclerosis*, **77**,227-237.

63. Mendis,S., Shepherd,J., Packard,C.J., and Gaffney,D. (1990) Genetic variation in the cholesteryl ester transfer protein and apolipoprotein A-I genes and its relation to coronary heart disease in a Sri Lankan population. *Atherosclerosis*, **83**,21-27.

64. Marcais,C., Bernard,S., Merlin,M., Ulhmann,M., Mestre,B., Rochet-Mingret,L., Revol,A., Berthezene,F., and Moulin,P. (2000) Severe hypertriglyceridaemia in Type II diabetes: involvement of apoC- III Sst-I polymorphism, LPL mutations and apo E3 deficiency. *Diabetologia*, **43**,1346-1352.