New technologies for delineating and characterizing the lipid exome: prospects for understanding familial combined hyperlipidemia

Stuart D. Horswell, Helen E. Ringham, and Carol C. Shoulders

Medical Research Council Clinical Sciences Centre, Imperial College London, Hammersmith Hospital, Du Cane Rd., London, W12 0NN United Kingdom

Abstract This review summarizes the progress made in cutting through the biological and genetic complexity of the Gordian knot that is familial combined hyperlipidemia. We particularly focus on how the application of new genomic technologies, especially massively parallel sequencing and high-throughput genotyping platforms, promise to accelerate the gene discovery process in this common, highly atherogenic disorder, with important diagnostic and therapeutic implications. — Horswell, S. D., H. E. Ringham, and C. C. Shoulders. New technologies for delineating and characterizing the lipid exome: prospects for understanding familial combined hyperlipidemia. J. Lipid Res. 2009. 50: S370–S375.

Supplementary key words linkage • genetic heterogeneity

This review considers the potential power of new and emerging genomic technologies to solve a clinically important yet seemingly intractable problem, the genetic dissection and consequent biological understanding of familial combined hyperlipidemia (FCHL). The reason for this focus is three-fold. First, the genetic complexity of this condition is undoubtedly rooted in the diversity of the intra- and inter-cellular processes of the various organs that govern serum cholesterol and triglyceride levels (1–4). As such, only the application of the new generation of genomic tools, in particular massively parallel sequencing (Mps) (5–8) and microarray technologies (9–11), has the capability to comprehensively identify and characterize the genomic variants that contribute to the transmission and manifestation of this common, highly atherogenic disorder (12, 13). Second, the wealth of data emerging from the application of these tools will provide important genetic information to supplement the traditional family-based, biochemical diagnosis of FCHL, thereby promising to resolve the current controversy over diagnostic criteria (14–17), as well as offering the prospect of early therapeutic intervention. Third, comparative analysis of the data from multiple human genomes, obtained with the new tools, will enormously expand our knowledge (and perhaps even our understanding) of the contribution that lipids make to a diverse range of cellular activities from conception to grave, including sperm capacitation (18), fetal and embryonic development (19, 20), osteoblast differentiation (21), and immunity (22). We may therefore find many examples, as in sickle cell anemia (23), where Mother Nature gives with one hand and takes away with her other.

Current data suggest that both common and rare variants underlie FCHL, with the specific contributions of each form varying from cohort to cohort, family to family, and even from family member to family member (15, 24–26). In short, the mixed patterns of hyperlipidemia in FCHL families may result from either specific combinations of common variants (Fig. 1), with each allele having only a small influence on serum lipid values (common-disease-common-variant hypothesis), or multiple, rare, functionally deleterious variants each individually affecting a handful of pedigrees (common-disease-multiple-rare-variant hypothesis), or a combination of both.

CHARACTERIZING THE LIPID EXOME

The per-unit cost and high throughput of the new genomic technologies now offer the possibility of sequencing the entire lipid exome (operationally defined as exons of all known “lipid metabolism” genes and liberally extended to include those regulating their pre- and posttranscript-

Supported by the Medical Research Council and the British Heart Foundation.

Manuscript received 10 November 2008 and in revised form 10 November 2008.

Published, JLR Papers in Press, November 20, 2008.

DOI 10.1194/jlr.R800081-JLR200

Abbreviations: asp, age-sex percentile; BMI, body mass index; FCHL, familial combined hyperlipidemia; HDL-C, HDL-cholesterol; LDL-C, LDL-cholesterol; Mps, massively parallel sequencing; SNP, single nucleotide polymorphism.

1 To whom correspondence should be addressed.

e-mail: carol.shoulders@csc.mrc.ac.uk

Copyright © 2009 by the American Society for Biochemistry and Molecular Biology, Inc.

This article is available online at http://www.jlr.org
individuals. The genes comprised two groups: A) 21 whose disruption caused obesity in either mice or humans; and B) 37 involved in regulating food intake, adipogenesis, energy expenditure, or lipid metabolism but with no genetic evidence for involvement in obesity. Ahituv et al. found that the cumulative frequency of unique nonsynonymous variants across all group A genes was significantly higher in the obese (46 variants in 41 individuals) subjects compared with the lean (26 in 27 individuals) participants. Moreover, it was discerned that a higher proportion of those variants in the obese group were likely to be functionally deleterious than in the lean group (19 vs. 4; \( P < 0.002 \)). By contrast, the number of functionally deleterious variants in the group B control genes did not differ significantly between obese and lean individuals (21 vs. 27). In summary, the data from these three studies suggest that a highly heterogeneous spectrum of nonsynonymous coding variants is likely to lurk in the genomes of FCHL patients, with a high proportion of the functionally deleterious changes residing in their lipid exomes.

To assess the common-disease-multiple-rare-variant hypothesis for the dyslipidemias, several investigators have already applied yesterday’s technologies, yielding some tantalizing results. For example, Wang et al. (30) showed that an \( LPL \) or \( APOC2 \) coding mutation was present on one allele in 10\% of hypertriglyceridemic patients, while Cohen et al. (31) demonstrated that nonsynonymous changes in \( ABCA1, APOAI, \) and \( LCAT \) were more common (16\% vs. 2\%) in people with low HDL cholesterol (HDL-C) levels [<5th age-sex percentile (asp)] than in those with high HDL-C (>95th asp). Additionally, Civeira et al. (32) found that 19.6\% of their 143 Spanish FCHL patients had a \( LDLR \) mutation and understandably argued that these individuals might be more accurately diagnosed as having the FCHL phenotype rather than the disease per se. In detail, using microarray chips containing 203 detrimental \( LDLR \) variants, they established that 23 of their patients had a mutation on one allele, while five bore two such variants (presumably on the same allele). Moreover, as would be anticipated, they showed that the patient group bearing a \( LDLR \) mutation (22 different variants) had higher mean total cholesterol, LDL cholesterol (LDL-C), and apoB levels, plus smaller waist circumferences, than those who did not. Disappointingly, Civeira et al. (32) did not examine whether the mutations segregated with hypercholesterolemia in their FCHL families, which would have provided more information on the potential overlap between monogenic familial hypercholesterolemia and the more complex disorder, FCHL. In a similar study, Gianfalone et al. (26) screened 41 unrelated French-Canadian hyper-apoB/FCHL patients for mutations in the acylation stimulating protein receptor gene, \( C5L2 \). A rare variant (i.e., \( 0/2176 \) control chromosomes) was detected on one allele (Ser323Ile) in one patient, and subsequently in two of his siblings who also had marked combined hyperlipidemia. This allele was transmitted to four offspring: two who were normolipidemic (i.e., <90th asp), one who was hypertriglyceridemic (i.e., >90th asp), and one who had hypercholesterolemia (i.e., >90th asp). Thus, although the data for involvement of \( C5L2 \) variants in the etiology of FCHL are not yet persua-

---

**Fig. 1.** Different combinations of rare and common variants may underlie transmission of FCHL-lipid traits. Hypothetical mode of transmission of the FCHL-cholesterol (TC), triglyceride (TG), and combined hyperlipidemia (CHL) traits. Family was ascertained through a white British proband (indicated by arrow) with serum TC>95th and TG>90th age-sex-specific percentiles. Exclusion criteria were as described (15). Color coding of variants: TC, red shades; TG, blue; CHL, purple. Phenotypic effect of variants proportional to tile size, frequency is inversely proportional.
sive, this study stresses the importance of analyzing large family cohorts with sufficient power to delineate the contribution of rare variants in the transmission of FCHL-lipid abnormalities.

Turning to the common-disease-common-variant hypothesis, there is compelling evidence that relatively common variants in the ApoA1/C3/A4/A5 complex contribute to FCHL (15, 33), although, as discussed in two excellent reviews (34, 35), the exact mechanisms by which these single nucleotide polymorphisms (SNPs) alter serum lipid levels are not fully resolved. By sequencing the retinoid X receptor γ (RXRγ) gene in 180 hyperlipidemic patients, Nohara et al. (36) found the common variant Gly14Ser. They showed it was more prevalent in FCHL probands (9/60; 15%) compared with peers with other forms of hyperlipidemia (5/120; 4%) and the general Japanese population (15/298; 5%). Rather unexpectedly, they did not find the RXRγ Ser14 variant to be more prevalent in patients (age 58 ± 7 years) suspected of having coronary heart disease (9/175; 5%). Nonetheless, in this patient group, the Ser14 allele was associated with higher plasma triglyceride (1.91 ± 0.42 vs. 1.62 ± 0.82 mmol/L; \( P < 0.05 \)), lower HDL-C levels (1.04 ± 0.26 vs. 1.32 ± 0.41 mmol/L; \( P < 0.05 \)), and decreased postheparin LPL activities (0.08 ± 0.03 vs. 0.11 ± 0.06 U/L; \( P < 0.05 \)) and mass (149 ± 38 vs. 222 ± 85 ng/ml; \( P < 0.01 \)). They also showed that in vitro, the RXRγ Ser14 allele repressed LPL promoter activity more than the wild type. Consequently, we keenly await further genetic and functional evidence that the Ser14 allele, or indeed other RXRγ variants, confers susceptibility to FCHL.

After SNPs, insertion/deletions (indels) are the most common type of genetic variation in the human exome (27), and indeed we have known for some time that common indels within the lipid exome (e.g., apoB, LPL) quantitatively influence serum lipid levels (37, 38). Less clear is whether there are rare indels making more substantial contributions to dyslipidemia, and it is this question that Mps, combined with high-throughput genotyping methodologies, can readily address. In this respect, the rare PCSK9 indel, designated p.L15_L16ins2L (39), is of interest. It was identified in 2 out of 25 (8%) French-Canadian FCHL families, absent from all 100 controls, and present in just 7 (1.8%) Americans participating in the Lipoprotein Coronary Atherosclerosis study [372 individuals; LDL-C: 2.96–4.90 mmol/L (40)]. In the affected families, heterozygotes (\( n = 10 \)) with the p.L15_L16ins2L allele had higher mean total cholesterol and LDL-C levels than the 9 relatives homozygous for the common allele (6.59 ± 0.97 vs. 5.18 ± 0.81 mmol/L, \( P = 0.0067 \); 4.06 ± 0.82 vs. 3.13 ± 0.70 mmol/L, \( P = 0.0167 \)). However, not all heterozygote family members were hyperlipidemic and, conversely, not all family members who were hyperlipidemic had the p.L15_L16ins2L allele. Thus, further studies are required to establish whether this indel only contributes to the FCHL-cholesterol trait on a specific genetic background(s) and if so, by what mechanism(s).

Recent technical and statistical advances (41–44) indicate that it is now timely to examine the copy number variant/polymorphism content of FCHL-patients’ lipid exome, the rationale being that copy number variants contribute to autism (45) and neurodevelopmental pathways in schizophrenia (46). Additionally, psoriasis is associated with increased β-defensin genomic copy number (47).

**REVISITING FCHL LINKAGE INTERVALS**

As spelled out by Kahvejian et al. (5), the new sequencing technologies have yet to reach the nucleotide unit-cost that would permit submission of a grant application to sequence the entire genome of, say, 200 FCHL patients and controls without automatically being filed under fantasy. Thus, for now, much high-density, sequence-driven FCHL research will necessarily be confined to analyzing those genomic intervals (lipid exomic or otherwise) linked to its component traits (15, 48). Despite this restriction (which until recently would itself have seemed a luxury), this approach, combined with targeted transcriptome/expression analyses, promises to clarify, to an unprecedented degree, the causes of FCHL and the cellular processes that run amok in this condition. Recent results from Lee et al. (49) support this stance. Proceeding from a well-documented linkage interval, they painstakingly accrued genotype data from family cohorts and two case-control samples, finding a common variant [or allele(s) in linkage disequilibrium] within the a priori biologically unconnected tumor suppressor gene, WWOX (WW-Domain-Containing Oxidoreductase) that modifies HDL-C levels.

In outline, Lee et al. (49) tracked the transmission of 1,318 SNPs to individuals with low HDL-C levels (<10th asp) in 50 Finnish dyslipidemic families (33 FCHL, 17 low-HDL-C pedigrees) displaying evidence for linkage of this trait to the chromosome 16q23-24 interval. They found a series of SNPs showing association (\( P < 0.01 \)) with the low HDL-C trait, 25 of which were analyzed in a further 52 dyslipidemic families. In the combined dataset of 102 dyslipidemic families, the rs2548861 T allele (frequencies 0.49 and 0.50) had lower HDL-C levels than individuals without the allele (\( P = 0.003 \) and 0.004), with one copy accounting for ~1.5% of the variance in HDL-C values. Hence, this study nicely demonstrates the efficacy of a SNP-based approach to detect modifier variants that have risen to high frequencies in the human population, presumably because they were not evolutionarily disadvantageous. It also highlights that a more in-depth analysis of the WWOX locus is warranted, specifically, to establish whether the rs2548861 T allele is merely acting as a marker for a rarer functional variant(s) that has a more substantial impact on an individual’s HDL-C than current estimates and, additionally, to identify whether rare variants at the WWOX locus are enriched in patient groups with low-HDL-C levels.

Many investigators (15) have used genome-wide linkage analyses with mixed success to identify the chromo-
performed on the datasets. Multipoint LOD scores are mean maximum chromosomes (UC) were allocated independently of affectations status. Marker characteristics were taken from observed data and 300 replicates (Chr). Each family was randomly assigned one such pair or phenocopy status. Alleles at unassigned locus pairs and loci on completely unlinked chromosomes (UC) were allocated independently of affectations status. Marker characteristics were taken from observed data and 300 replicates performed on the datasets. LOD scores were also generally higher in the chromosomal regions allocated two susceptibility variants (e.g., chromosome 1), and in the single region carrying 3 variants (chromosome 11), LOD scores were higher still. This heterogeneous region was the only one to consistently return values exceeding the classical thresholds (50) for declaring linkage.

We present the results from the simulations performed on real family structures to make a fundamental point: genuine FCHL susceptibility variants will exist in some of the chromosomal regions that have returned LOD scores just missing threshold levels for declaration of linkage (50), simply because obtaining significant/confirmed LOD scores is difficult in the face of genetic heterogeneity on linkage signals.

Simulations were carried out on extended white British FCHL pedigrees with two sample sizes: 60 families to emulate the size of previous genome-wide FCHL linkage scans (15, 48) and 181 families corresponding to our most recent study. To this end, we declared [based on published (15, 48) and our unpublished data] that the component traits of this disorder were caused by variants at any one of 9 distinct pairs of susceptibility loci and applied a phenocopy rate of 10%. Chromosomes assigned no susceptibility loci returned low LOD scores (<0.25) (representative results summarized in Table 1). At susceptibility loci, the 181 family datasets (300 simulations) generated mean LOD scores ~2-fold higher than those observed in the 60 family datasets. LOD scores were also generally higher in the chromosomal regions allocated two susceptibility variants (e.g., chromosome 1), and in the single region carrying 3 variants (chromosome 11), LOD scores were higher still. This heterogeneous region was the only one to consistently return values exceeding the classical thresholds (50) for declaring linkage.

We thank the Medical Research Council and the British Heart Foundation for funding our research, and Drs. Andrew Dean, Margaret Town, Max Salm, and Bethan Jones for critically reading the review.

TABLE 1. Summary of representative results of simulation study

| Chr | Families | LOD | Proportion of replicates | Proportion of replicates | Proportion of replicates | Proportion of replicates |
|-----|----------|-----|--------------------------|--------------------------|--------------------------|--------------------------|
|     |          |     | >2.2                     | >3.6                     | >5.4                     | >2.2                     | >3.6                     | >5.4                     |
| UC  | 60       | 0.14 ± 0.21 | 0.00 | 0.00 | 0.00 | 181 | 0.17 ± 0.23 | 0.00 | 0.00 | 0.00 | 0.19 ± 0.26 | 0.00 | 0.00 | 0.00 |
| 1   | 181      | 0.36 ± 0.89 | 0.18 | 0.02 | 0.00 | 181 | 0.21 ± 0.89 | 0.12 | 0.02 | 0.00 | 0.76 ± 0.70 | 0.05 | 0.00 | 0.00 |
| 2   | 60       | 1.30 ± 1.24 | 0.56 | 0.02 | 0.03 | 181 | 1.21 ± 1.34 | 0.63 | 0.20 | 0.04 | 1.21 ± 0.92 | 0.15 | 0.02 | 0.00 |
| 4   | 60       | 0.71 ± 0.55 | 0.03 | 0.00 | 0.00 | 181 | 3.20 ± 1.31 | 0.75 | 0.35 | 0.06 | 1.98 ± 1.11 | 0.38 | 0.09 | 0.00 |
| 8   | 181      | 0.98 ± 0.90 | 0.21 | 0.03 | 0.00 | 181 | 0.80 ± 1.41 | 0.80 | 0.40 | 0.09 | 0.70 ± 0.78 | 0.47 | 0.08 | 0.01 |
| 9   | 60       | 0.82 ± 0.61 | 0.02 | 0.00 | 0.00 | 181 | 1.04 ± 0.71 | 0.06 | 0.01 | 0.00 | 0.72 ± 0.87 | 0.12 | 0.02 | 0.00 |
| 10  | 60       | 1.55 ± 0.86 | 0.21 | 0.02 | 0.00 | 181 | 1.34 ± 0.85 | 0.16 | 0.01 | 0.00 | 0.73 ± 0.63 | 0.03 | 0.00 | 0.00 |
| 11  | 60       | 3.67 ± 1.32 | 0.88 | 0.48 | 0.10 | 181 | 4.20 ± 1.51 | 0.91 | 0.63 | 0.20 | 2.77 ± 1.12 | 0.68 | 0.20 | 0.02 |
| 16  | 60       | 10.52 ± 2.27 | 1.00 | 1.00 | 0.99 | 181 | 1.00 ± 0.84 | 0.10 | 0.02 | 0.00 | 0.72 ± 0.87 | 0.07 | 0.02 | 0.00 |
| 17  | 60       | 0.57 ± 0.58 | 0.02 | 0.00 | 0.00 | 181 | 1.45 ± 0.85 | 0.17 | 0.02 | 0.00 | 0.51 ± 0.55 | 0.01 | 0.00 | 0.00 |
| 20  | 60       | 0.95 ± 0.69 | 0.04 | 0.01 | 0.00 | 181 | 1.52 ± 0.95 | 0.19 | 0.03 | 0.01 | 0.65 ± 0.83 | 0.13 | 0.02 | 0.00 |
| 21  | 60       | 0.84 ± 0.71 | 0.05 | 0.00 | 0.00 | 181 | 2.02 ± 1.09 | 0.40 | 0.09 | 0.00 | 1.25 ± 0.94 | 0.15 | 0.02 | 0.00 |
case-control association testing with copy number variation. Nat. Genet. 40: 1245–1252.
42. Korn, J. M., F. G. Kuruvilla, S. A. McCarroll, A. Wysoker, J. Nemesh, S. Cavley, E. Hubbell, J. Veitch, P. J. Collins, K. Darvish, et al. 2008. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat. Genet. 40: 1253–1260.
43. McCarroll, S. A. 2008. Extending genome-wide association studies to copy-number variation. Hum. Mol. Genet. 17: R135–R142.
44. Wang, K., Z. Chen, M. G. Tadesse, J. Glessner, S. F. Grant, H. Hakonarson, M. Bucan, and M. Li. 2008. Modeling genetic inheritance of copy number variations. Nucleic Acids Res. 36: e138.
45. Sebat, J., B. Lakshmi, D. Malhotra, J. Troge, C. Lese-Martin, T. Walsh, B. Yamrom, S. Yoon, A. Krasnitz, J. Kendall, et al. 2007. Strong association of de novo copy number mutations with autism. Science. 316: 445–449.
46. Walsh, T., J. M. McClellan, S. E. McCarthy, A. M. Addington, S. B. Pierce, G. M. Cooper, A. S. Nord, M. Kusenda, D. Malhotra, A. Bhandari, et al. 2008. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science. 320: 539–543.
47. Hollox, E. J., U. Huffmeier, P. I. Zeeuwen, R. Palla, J. Lascorz, D. Rodijk-Olthuis, P. C. van de Kerkhof, H. Traupe, G. de Jongh, M. den Heijer, et al. 2008. Psoriasis is associated with increased beta-defensin genomic copy number. Nat. Genet. 40: 23–25.
48. Pollex, R. L., and R. A. Hegele. 2005. Complex trait locus linkage mapping in atherosclerotic time to take a step back before moving forward? Arterioscler. Thromb. Vasc. Biol. 25: 1541–1544.
49. Lee, J. C., D. Weissglas-Volkov, M. Kytala, Z. Dastani, R. M. Cantor, E. M. Sobel, C. L. Plaisier, J. C. Engert, M. M. van Grevenbroek, J. P. Kane, et al. 2008. WW-domain-containing oxidoreductase is associated with low plasma HDL-C levels. Am. J. Hum. Genet. 83: 180–192.
50. Lander, E., and L. Kruglyak. 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat. Genet. 11: 241–247.