Deciphering Unexplained Familial Dyslipidemias
Do We Have the Tools?

Heribert Schunkert, MD; Felix Bourier, MD

Familial dyslipidemias may increase the risk of atherosclerosis. As a role model, familial hypercholesterolemia (FH) is characterized by marked elevations of plasma low-density lipoprotein (LDL) cholesterol levels and myocardial infarction risk. Variants in LDL-receptor, apolipoprotein B, proprotein convertase subtilisin/kexin type 9, and STAP1 explain, in aggregate, ≈60% of FH cases and account for ≈5% of all cases with premature myocardial infarction. The large number of patients with apparently monogenic dyslipidemias but without molecular explanation suggests that mutations in other genes, which have not been identified to date, may also cause such disorders.

Why Aiming at a Molecular Genetic Diagnosis?
Standard risk algorithms (eg, SCORE, Framingham) may underestimate the coronary heart disease risk in FH patients. Thus, knowledge of the mutation improves prognostic evaluation and justifies early, intensive treatment at younger age in hypercholesterolemia patients. Moreover, once a causal variant has been identified cascade screening in the patient’s relatives is simplified and treatment can be started more promptly. Based on these and other considerations, the European and American guidelines recommend a cascade screening of the entire family for obtaining a molecular diagnosis, identification of all affected family members, and early initiation of medical therapy.

In the 8 expert laboratories collaborating with Stitziel et al, such molecular genetic testing for causal mutations had been carried before whole-exome sequencing. Unfortunately, we do not know in how many families candidate gene sequencing had been successful (which precluded them from entering the current study). Thus, the article does not inform us about the overall percentage of families with dyslipidemia, in which a molecular diagnosis can be obtained successfully. Rather, 1 conclusion of the work is that next-generation sequencing may sometimes identify mutations in candidate genes, which had undergone futile conventional sequencing (3 of the 41 families).

Why Did Exome Sequencing Fail to Identify More Genes?
In ≈80% of the selected families with apparently Mendelian inheritance of dyslipidemia, whole-exome sequencing failed to identify the causal mutation. Stitziel et al offer 3 reasons for the lack of novel gene discovery: (1) an inability to identify potentially causal variants because of imperfect sequencing coverage; (2) an inability to identify the causal variant among hundreds of shared variants within families; and (3) an inability to identify the effect of complex genetics using exome sequencing.

First, using an Illumina GA-II sequencer, which produced reads with a length of ≈75 base pairs, ≈4% of the exomes of known lipid genes were covered with <10 reads, that is, not adequately to exclude a potentially causal mutation. Albeit the algorithm to further investigate a variant required that a putative mutation was found in all affected family members (which may increase the chance to miss a relevant mutation because of poor coverage), it is unlikely that the small number of mutation discoveries in these families was because of poor sequencing coverage.

Inherently, exome sequencing can only identify mutations within the coding sequence and some adjacent segments. Thus, whole genome sequencing offers a much broader view. However, genetic variants underlying a Mendelian inheritance pattern are expected to result in massive alterations of protein function such that it is questionable whether whole genome sequencing, that is, adding more noncoding sequence to the already known coding sequence, of these families would change the overall picture.

Second, Stitziel et al identified an average of 12,544 non-synonymous single-nucleotide variants and 802 insertion/deletions—per individual! Thus, rather too many variants than too few were a problem. Preferentially, a large family size, which allows adequate linkage analysis, may help to discriminate
causal variants in such situation. However, because of the small size of the pedigrees studied by Stitziel et al, traditional linkage for obtaining logarithm of the odds scores was not feasible. Thus, given the thousands of private variants present in families, it seems that exome sequencing may compare to loosing sight of the forest for the trees. For future studies in the era of next-generation sequencing, this finding implies that large families with multiple affected individuals are still a fundamental prerequisite for identification of novel genes involved in a disease.

A method that may circumvent linkage for gene identification is the so-called burden test. This test aims to document clustering of mutations in a given gene in individuals that share a specific phenotype. However, by far larger studies like the one of Stitziel performed in 41 families are required to apply the burden test and overcome the limitation of small family size.

Third, common LDL modifying SNPs can modulate the extent of LDL elevation in mutation carriers or even copy the phenotype by clustering of multiple variants, each affecting LDL-cholesterol by a small extent, in a given individual. Such phenocopy might have been of relevance in as many as 20% of the families studied by Stitziel et al, as well as FH patients, of a previous study. These findings exemplify the seemingly contradictory hereditary properties of discrete traits (eg, Mendel’s peas or FH) and continuous traits (eg, Galton’s height or population-wide variability in LDL cholesterol). About a century ago, Fisher showed that the 2 were consistent if quantitative trait variation is caused by a combination of many genetic loci, each with a variable effect and inherited in a Mendelian manner, together with environmental effects. The substantial variability of LDL and high-density lipoprotein cholesterol levels brought about by common variants may even obscure the view on monogenic factors (Figure). Adjustment for the effects of common variants may improve the cosegregation signal and help to discriminate a gene affecting the phenotype more profoundly.

Many Linkage Studies Fail to Uncover Apparently Mendelian Traits

Stitziel et al are certainly not alone in that they failed to explain in a number of families with an apparently Mendelian inheritance the molecular nature of disease. Indeed, many cardiovascular geneticists, like us, share this obstacle. Most of the failed attempts will not be published such that the overall dilemma is hard to quantify. Indeed, even the
Thus, a challenge for modern genetics will be (1) to determine the extent of apparently Mendelian but unexplained inheritance and (2) to tackle the puzzle of so-called missing heritability.

Dynamic gene–gene or gene–environment interactions may explain a sizable proportion of the failure with a mono- genic view on disease pathogenesis. As an example, a recent study on a large family with 32 coronary artery disease patients produced maximal logarithm of the odds scores of only 1.0 to 1.5 by single locus analysis. However, 2 functionally interacting genes both carrying mutations produced a 2-point logarithm of the odds score of 5.68. In this case, the interaction of 2 specific genes explained the disease best, whereas each individual mutation had insufficient penetrance to raise suspicion.

One may hypothesize that such gene–gene interaction may be the rule, rather than the exception in families with apparently Mendelian inheritance but no conclusive signal from variants at known candidate genes. Such epistasis may largely reside even outside the protein-coding exome, as common regulatory variants affect either transcription of coding and noncoding RNAs (regulatory SNPs or rSNPs) or RNA functions and processing (structural RNA SNPs or srSNPs).

The Figure illustrates how the effect of a single dominant variant increasing LDL cholesterol by 100 mg/dL may be diluted in subsequent generations. Assuming that the grandparents carry no other allele that affects the phenotype, the distribution in the grandchildren will be identical (Figure, left panel). However, we know already of multiple alleles that elevate or decrease LDL by 2 mg/dL. If we add 100 of such alleles to the imputation model, LDL distribution looks different (Figure, middle panel). If we add epistasis to the model, assuming that 3% of all possible interactions between the 101 alleles will amplify the effect of the interaction partner by 20%, for example, increasing the effect of a given allele from 2 to 2.2 mg/dL, the dominant allele is almost hidden in the distribution curve (right panel).

Genome-Wide Association Studies May Help
Interestingly, the genes in which the mutations identified by Stitziel reside had all produced signals in previous genome-wide association studies. Indeed, this applies to almost all genes involved in familial dyslipidemias, perhaps with the exception of the newly discovered STAP1. In other words, these genes harbor both common variants with effects sufficiently large to be detectable in genome-wide association studies as well as mutations with profound effects on plasma lipids to display a Mendelian pattern of inheritance. Thus, one may hypothesize that discovery of mutations showing Mendelian inheritance might as well be focused on genome-wide association studies hits in the first place.

Disclosures
None.

References

Key Words: Editorials ▪ coronary artery disease ▪ familial hypercholesterolemia ▪ molecular epidemiology
Deciphering Unexplained Familial Dyslipidemias: Do We Have the Tools?
Heribert Schunkert and Felix Bourier

_Circ Cardiovasc Genet._ 2015;8:250-252
doi: 10.1161/CIRCGENETICS.115.001066
_Circulation: Cardiovascular Genetics_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2015 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/8/2/250

**Permissions:** Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation: Cardiovascular Genetics_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

**Reprints:** Information about reprints can be found online at:
http://www.lww.com/reprints

**Subscriptions:** Information about subscribing to _Circulation: Cardiovascular Genetics_ is online at:
http://circgenetics.ahajournals.org//subscriptions/