Published in final edited form as:

Coron Artery Dis. 2016 June ; 27(4): 253–254. doi:10.1097/MCA.0000000000000367.

Whole-Exome Sequencing for the Discovery of Rare Genetic Variants That Protect from Coronary Artery Disease

Jane A. Leopold, MD

Keywords

genome wide association studies; whole exome sequencing; coronary artery disease

It has long been recognized that heritable factors contribute to the pathogenesis of coronary artery disease (CAD) and explain its high prevalence in certain families [1]. Population-based epidemiological studies have also confirmed the importance of genetic factors for the development of CAD, albeit indirectly, by establishing family history as an important and independent risk factor for myocardial infarction [2]. For example, in the Framingham Study a family history of CAD was associated with a 2.6-fold increased risk of CAD for men and a 2.3-fold increased risk for women [3]. It is, therefore, not surprising that when considered together, findings from epidemiological and family studies have suggested that “genetic risk” may account for up to ~50% of an individual’s susceptibility for developing CAD [2]. To define this “genetic risk,” investigators have embarked on large-scale analyses of the human genome in the search for genes that predispose some individuals to CAD.

Genome wide association studies (GWAS) sought to determine whether or not there were associations between common genetic variants and CAD using a case-control study design. This approach ultimately identified a variant on chromosome 9p21 that was shown to impart a >2-fold increased risk of myocardial infarction independent of traditional risk factors on carriers of the allele. As approximately 4×10^9 individuals are estimated to carry one or two alleles, the 9p21.3 variant itself has been considered as an independent risk factor for CAD [2,4,5]. In the GWAS studies that followed, more than 50 genetic loci associated with CAD have been identified; however, they account for only 10% of disease heritability and only 33% of these loci were associated with traditional CAD risk factors [2,6]. This suggested that while GWAS provide valuable information, they are limited by the technology (microarrays containing prevalent single nucleotide polymorphisms as DNA markers at base pair intervals spanning the genome), the fact that the results do not always provide information about a relevant disease gene, and the low likelihood of detecting rare variants.

Whole-exome sequencing (WES), a next generation sequencing technology, uses massively parallel sequencing of DNA for high throughput but focuses only the exons that encode proteins. Exons comprise only ~1% of the genome but are believed to harbor ~85% of

disease-related mutations [7,8]. The advantage of WES is the ability to detect rare genetic variants that have a large effect on the CAD phenotype. The power of WES to identify rare CAD-related variants as compared to GWAS has been demonstrated. For example, GWAS found frequent variants that mapped near the *GUCY1A3* gene in individuals with CAD. This gene encodes the $\alpha 1$ -subunit of soluble guanylyl cyclase, which is a key component of nitric oxide signaling and plays a role in regulating vasomotor tone and thrombus formation [9]. In contrast to the GWAS results that identified a loci near *GUCY1A3*, WES of only 3 members of an extended family with a high prevalence of CAD found loss-of-function mutations in *GUCY1A3* (p.Leu163Phefs*24) and a second gene *CCT7* (p.Ser525Leu), which encodes a protein that stabilizes soluble guanylyl cyclase. The presence of both mutations was shown to decrease soluble guanylyl cyclase protein levels and activity *in vitro* and correlated *in vivo* with a shorter time to thrombus formation after vascular injury in a mouse model. Individuals that harbored both mutations, presumably with significant defects in nitric oxide-soluble guanylyl cyclase signaling, were found to have 100% risk of developing of CAD [10].

Other WES studies in the field similarly focused on finding genetic variants associated with CAD leading them to ignore the flip side of the coin, namely to investigate why some individuals are protected from CAD despite the presence of several risk factors. In this issue of *Coronary Artery Disease*, Abramowitz et al. hone in on issue and use WES to identify putative atheroprotective variants [11]. By studying 17 individuals with angiographically normal coronary arteries despite a high burden of CAD risk factors and 17 matched controls with multivessel disease, they found 51 genes with a significant single nucleotide variant burden in patients with normal coronary arteries. Gene ontology mapping showed that these genes are related to the regulation of the JAK-STAT signaling cascade and ventricular development both of which are biologically relevant to the cellular processes that contribute to CAD. To identify rare variants, the investigators used a stringent selection process that resulted in discovery of 19 variants present in 16 genes in patients without CAD that were not present in patients with multivessel disease. Gene set enrichment analysis revealed that there was a higher than expected representation in categories related to the cell cytoskeleton and sarcomere assembly. This led to the identification of the candidate genes *SPTBN5*, *NID2*, and *ADAMTSL4*, which all encode proteins that are important for the extracellular matrix, a key factor in vascular health and disease [11]. Although further biological characterization of these genes and variants was beyond the scope of this study, it will be of interest to know whether or not they are confirmed in other WES studies.

The utility of WES as a mechanism to identify atheroprotective genes is an area of active interest and two other large studies have investigated this concept via different approaches. The first sequenced the exome of 3,734 individuals of European or African ancestry found 4 rare mutations (3 loss-of-function and 1 missense) in *APOC3*, the gene encoding apolipoprotein C3. This was associated with triglyceride levels that were 39% lower and a risk of CAD that was 40% lower in individuals who are carriers of the variants (~1 in 150 people) compared to noncarriers [12]. The second study used the rationale that rare variants that led to loss-of-function in known drug targets (i.e., Neimann-Pick C1-like 1 (NPC1L1) protein) and mimicked the effects of cardioprotective drugs (i.e., ezetimibe) would have similar benefit. The study performed WES in 7,364 patients with documented CAD and

14,728 control subjects and identified 15 inactivating mutations in NPC1L1 that occurred in 1 out of every 650 individuals. Carriers had significantly lower mean total cholesterol (– 13 mg/dL) and low-density lipoprotein cholesterol levels (– 12 mg/dL) compared to noncarriers and this was associated with a 53% relative reduction of CAD [13]. Although neither of these targets was identified in the Abramowitz et al. study [11], this is likely due to differences in the sample size as well as the patient and control populations.

As next generation sequencing technologies, such as WES, advance and it becomes easier to sequence larger patient populations, there will be a rapid pace of discovery of rare variants linked to CAD. This will likely lead to the identification of genetic profile associated with increased risk of CAD that is comprised of a panel of genes with rare variants that could also be used to provide a personalized risk assessment. Finally, data from these types of studies could be used to explore the druggable genome and identify drug-gene interactions or find new pharmacotherapeutic targets. It is important, however, to recognize that WES is limited to protein coding regions and has the potential to miss some genetic contributors to CAD, such as intronic long noncoding RNAs, some microRNAs, and risk-associated loci such as 9p21. Although these are the early days of large-scale WES studies, it is important to recognize that smaller studies, such as the one by Abramowitz et al. [11], will make a large contribution to our understanding of what protects us from and what increases our risk for CAD.

Acknowledgments

Source of Funding: NIH/NHLBI U01 125215 and the Thomas W. Smith, MD Foundation

References

1. Stitzel NO, MacRae CA. A clinical approach to inherited premature coronary artery disease. *Circ Cardiovasc Genet*. 2014; 7(4):558–564. [PubMed: 25140063]
2. Roberts R, Stewart AF. 9p21 and the genetic revolution for coronary artery disease. *Clin Chem*. 2012; 58(1):104–112. [PubMed: 22015375]
3. Lloyd-Jones DM, Nam BH, D'Agostino RB Sr, Levy D, Murabito JM, Wang TJ, et al. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA*. 2004; 291(18):2204–2211. [PubMed: 15138242]
4. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*. 2007; 316(5830):1491–1493. [PubMed: 17478679]
5. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science*. 2007; 316(5830):1488–1491. [PubMed: 17478681]
6. Lee SE, Kim HS. Unraveling new therapeutic targets of coronary artery disease by genetic approaches. *Circ J*. 2015; 79(1):8–14. [PubMed: 25501952]
7. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*. 2009; 461(7261):272–276. [PubMed: 19684571]
8. Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A*. 2009; 106(45):19096–19101. [PubMed: 19861545]

9. Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet.* 2013; 45(1):25–33. [PubMed: 23202125]
10. Erdmann J, Stark K, Esslinger UB, Rumpf PM, Koesling D, de Wit C, et al. Dysfunctional nitric oxide signalling increases risk of myocardial infarction. *Nature.* 2013; 504(7480):432–436. [PubMed: 24213632]
11. Abramowitz Y, Roth A, Keren G, Isakov O, Shomron N, Laitman Y, et al. Whole-exome sequencing in individuals with multiple cardiovascular risk factors and normal coronary arteries. *Coronary Artery Disease.* 2016
12. Crosby J, Peloso GM, Auer PL, Crosslin DR, Stitzel NO, Lange LA, et al. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *N Engl J Med.* 2014; 371(1):22–31. [PubMed: 24941081]
13. Stitzel NO, Won HH, Morrison AC, Peloso GM, Do R, Lange LA, et al. Inactivating mutations in NPC1L1 and protection from coronary heart disease. *N Engl J Med.* 2014; 371(22):2072–2082. [PubMed: 25390462]