# Genetic testing for atherosclerosis risk: Inevitability or pipe dream?

Matthew Lanktree BSc, Jisun Oh MD, Robert A Hegele MD FRCPC FACP

**M Lanktree, J Oh, RA Hegele. Genetic testing for atherosclerosis risk: Inevitability or pipe dream? Can J Cardiol 2008;24(11): 851-854.**

Family history is a risk factor for coronary artery disease (CAD). However, defining this risk at the DNA level has been elusive. In 2007, four genome-wide association studies reported a strong association between CAD and a region on chromosome 9p21. The high-risk genotype was identified in up to 30% of individuals, creating the potential for a clinical genetic test to assist in the calculation of a patient's CAD risk. However, the reported effect size of the association is modest (OR of approximately 1.3). The present paper examines the feasibility of including DNA tests in CAD risk prediction algorithms. The greatest contribution from the 9p21 association is likely yet to come, as further studies identify the mechanistic basis for the association, possibly leading to additional insights into the progression, prevention and treatment of CAD.

**Key Words:** *Atherosclerosis; Cardiovascular risk factors; Coronary artery disease; Genetic testing*

## Test génétique pour le risque d'athérosclérose : Inévitabilité ou solution à la « *Pipe Dream* »?

Si les antécédents familiaux constituent bel et bien un facteur de risque à l'égard de la coronaropathie, sa détermination à l'échelle de l'ADN nous élude toujours. En 2007, quatre études d'associations pangénomiques ont fait état d'un lien solide entre la coronaropathie et une région du chromosome 9p21. On a observé le génotype associé au risque élevé chez jusqu'à 30 % des sujets, ce qui ouvre la porte à l'éventuelle mise au point d'un test génétique unique pour faciliter le calcul du risque coronarien d'un patient. Or, l'effet de taille rapporté de ce lien est modeste (RR environ 1,3). Le présent article analyse l'applicabilité des tests d'ADN dans les algorithmes de prédiction du risque coronarien. La contribution la plus importante de ce lien avec le chromosome 9p21 n'a pas encore été identifiée, puisque d'autres études portent sur les fondements mécanistes du lien, ce qui pourrait mener à d'autres pistes pour l'étude de la progression, de la prévention et du traitement de la coronaropathie.

Coronary artery disease (CAD) risk factors have been incorporated into algorithms for risk assessment, such as the Prospective Cardiovascular Münster (PROCAM) study (1) and Framingham risk equations (2,3). However, many CAD events occur in the absence of known risk factors (4,5), while most individuals who never develop CAD have one or more risk factors (6). For this reason, identifying new risk factors is considered to be important for enhancing sensitivity and specificity of risk algorithms (4-6). In 2007, several reports (7-9) brought the potential use of genetic markers in CAD risk assessment and stratification closer to reality.

## GENETICS OF CAD

Family history of CAD in first-degree relatives is associated with an increased risk of CAD (10). However, defining this risk at the DNA level has been elusive. Genetic predisposition to CAD may comprise multiple, relatively common genetic variants, each with small to modest effects that, alone or in combination with modifier genes or environmental factors, modulate the risk of disease (11-14). This has been called the 'common disease-common variant' model (15). An alternative model, called the 'heterogeneity' model, maintains that genetic predisposition to common diseases in some patients is caused by rare genetic variants (14). Such rare variants explain the extreme versions of quantitative traits related to the disease. Because carriers of such mutations are found at the extremes of the distribution of the trait, this model would explain risk for a relatively small proportion of patients (10). In contrast, the 'common disease-common variant' model invokes common genetic variants that are found at a higher frequency and may affect a larger proportion of patients. Each individual variant has a small effect, but cumulatively, variants may exert a large effect and, thus, may explain susceptibility among individuals clustered around the centre of Gaussian distribution of a quantitative trait (16). In reality, a blend of both models likely explains the genetic basis of CAD.

## EVALUATING GENETIC DETERMINANTS OF DISEASE

Genetic linkage and association studies have endeavoured to implicate genes that may predispose patients to CAD end points, such as myocardial infarction. Historically, the most common approach was the study of a single gene using case-control designs. Hundreds of DNA polymorphisms in scores of genes encoding proteins involved in lipid metabolism, thrombosis and vascular biology have been associated with increased CAD risk (17,18). Limitations of candidate gene association studies are well known (18). Recent genome-wide association (GWA) studies suggest the existence of genetic markers consistently associated with CAD that may be translated into useful predictive tests (7,19,20). In addition, unbiased GWA studies have the ability to detect previously unsuspected associations (7,19). Finally, GWA studies examining the genetic basis of traditional risk factors, such as lipid concentrations, may also contribute to a predictive test (20,21). Technological advances include availability of dense genotyping microarrays examining $10^5$ to $10^6$ single nucleotide polymorphisms (SNPs); large, well-characterized clinical samples often pooled between different groups (13); and evolved bioinformatic capacity.

*Robarts Research Institute and Schulich School of Medicine and Dentistry, University of Western Ontario, London, Ontario*
*Correspondence: Dr Robert A Hegele, Blackburn Cardiovascular Genetics Laboratory, Vascular Biology Research Group, Robarts Research Institute, #406-100 Perth Drive, London, Ontario N6A 5K8. Telephone 519-663-3461, fax 519-663-3037, e-mail hegele@robarts.ca*
*Received for publication February 29, 2008. Accepted May 14, 2008*

**TABLE 1**
**Summary of 9p21 locus identified in genome-wide association studies of coronary artery disease**

| SNP identification number | Position (base pairs) | Helgadottir et al (8) 4587 cases, 12,767 controls | McPherson et al (9) 3989 cases, 18,808 controls | Samani et al (19) 875 cases, 1644 controls | WTCCC (7) 1926 cases ~3000 controls |
|---|---|---|---|---|---|
| rs10116277 | 22071397 | 1.24 (1.17–1.30) $1.8\times10^{-15}$ | – | – | – |
| rs6475606 | 22071850 | – | – | 1.28 (1.13–1.44) $5.2\times10^{-5}$ | 1.37 (1.26–1.48) $4.4\times10^{-14}$ |
| rs1333040 | 22073404 | 1.24 (1.17–1.30) $4.1\times10^{-15}$ | – | – | – |
| rs4977574 | 22088574 | – | – | 1.36 (1.20–1.53) $5.2\times10^{-7}$ | 1.35 (1.24–1.46) $4.3x10^{-13}$ |
| rs10757274 | 22086055 | – | 1.27 (1.21–1.33) $1.1\times10^{-22}$ | – | – |
| rs2891168 | 22088619 | – | – | 1.36 (1.20–1.53) $4.9\times10^{-7}$ | 1.35 (1.24–1.46) $5.9\times10^{-13}$ |
| rs1333042 | 22093813 | – | – | – | 1.34 (1.23–1.45) $2.5\times10^{-12}$ |
| rs2383206 | 22105026 | – | 1.24 (1.18–1.30) $1.1\times10^{-18}$ | – | – |
| rs2383207 | 22105959 | 1.25 (1.18–1.31) $2.0\times10^{-16}$ | – | – | – |
| rs10757278 | 22114477 | 1.28 (1.22–1.35) $1.2\times10^{-20}$ | – | – | – |
| rs1333048 | 22115347 | – | – | – | 1.36 (1.25–1.48) $1.3\times10^{-13}$ |
| rs1333049 | 22115503 | – | – | 1.33 (1.18–1.51) $3.4\times10^{-6}$ | 1.37 (1.26–1.48) $1.8\times10^{-14}$ |

*Position was taken from National Centre for Biotechnology Information reference sequence build 36.2. OR (with 95% CI range) and P is shown for at-risk allele (8,9) and at-risk genotype (7,19). SNP Single nucleotide polymorphism; WTCCC Wellcome Trust Case Control Consortium*

**Chromosome 9p21: A CAD locus discovered by GWA studies**

The Wellcome Trust Case Control Consortium explored genetic associations of several complex diseases, including CAD (7,19). A powerful association of CAD was observed with a region on chromosome 9p21.3. The strongest signal was seen with SNP rs1333049 ($P=1.8\times10^{-14}$), but association was seen for other SNPs across more than 100 kilobases. The RR for the susceptibility genotypes was approximately 1.4. This region harbours genes encoding two cyclin-dependent kinase inhibitors, *CDKN2A* (encoding $p16^{INK4a}$) and *CDKN2B* ($p15^{INK4b}$), although SNP rs1333049 was outside both genes. Eight other loci were found to be strongly associated with CAD (7). A further analysis of the Wellcome Trust Case Control Consortium and a sample of German subjects (19) identified several CAD-associated loci, with chromosome 9p21.3 SNP rs1333049 having the strongest association in both samples ($P=1.80\times10^{-14}$ and $P=3.40\times10^{-6}$, respectively) (19).

Two other GWA studies from 2007 supported the association of CAD with markers on chromosome 9p21: a 58-kilobase interval was associated with CAD in six samples (9) and an SNP in this region was associated with myocardial infarction (8). Each study showed a relative CAD risk of approximately 1.4, with strongest associations observed with SNPs outside the known genes. A recent report (22) confirmed the CAD association with 9p21 in a Korean population with a risk ratio of approximately 1.3. There is no biological explanation yet for the chromosome 9p21 SNP associations with CAD; indeed, new aspects of biology may be involved (9). Despite such modest ORs, the findings' consistency suggests a special relationship between CAD and chromosome 9p21 (GWA data are summarized in Table 1).

## HOW MANY GENOTYPES MAY BE REQUIRED TO PREDICT CAD RISK?

The population-attributable fraction (PAF) is defined as the proportion of disease cases in a population resulting from a particular risk factor (16). A risk factor cluster with PAF of 30% to 50% would be considered to have excellent potential for clinical use. But how many genetic variants will be required to produce such a PAF for CAD? Very large numbers – up to 100 or more – of rare genotypes (eg, those with a frequency of less than 1%) would be required to explain a PAF of 50%, even if the risk ratios for each variant are large (ie, 10 to 20). On the other hand, approximately 20 genotypes are needed to explain 50% of the burden of a disease in the population if the predisposing genotypes are common (more than 25%), even if the individual risk ratios are relatively small (ie, 1.2 to 1.5) (16).

As a further illustration, imagine a genotype with a frequency of 10% in the population that has a risk ratio of 1.5 for development of CAD. The number of markers with such a risk ratio and genotype frequency needed to explain a PAF of 50% ranges between 15 and 20. In contrast, if the at-risk genotype frequency is 30% and the risk ratio is 1.5, the number of different genotype markers to explain a PAF of 50% decreases to approximately seven (16). This illustrates the potential use of relatively small numbers of common genotype combinations in predicting CAD, assuming that such markers exist.

## WILL GENETIC TESTING BE CLINICALLY USEFUL?

For a diagnostic test to enter the clinic, it must provide additional predictive power over and above traditional risk factors that are easily and inexpensively measured (23). One method to evaluate added

diagnostic or predictive value of a test is the receiver-operating characteristic curve (24). The predictive power of a diagnostic test can be evaluated by the area under the receiver-operating characteristic curve (AUC) (24). A perfect test has an AUC of 1, while a test with no discriminatory power has an AUC of 0.5 (24). The discriminatory predictive power of the Framingham and PROCAM CAD risk algorithms (1,3) have AUC of 0.62 and 0.63, respectively (25). Hence, established risk indexes have moderate predictive power for CAD. Humphries et al (23) studied the apolipoprotein E gene and showed limited additional predictive power over classic CAD risk factors. The reported CAD OR is approximately 1.4 for carriers of the E4 allele, with a prevalence of 15% (23,25). Compared with PROCAM, adding the apolipoprotein E gene genotype nonsignificantly increased the AUC from 0.63 to 0.67 (P=0.11) (23). Thus, the inclusion of a single genetic risk factor with OR of 1.4 and prevalence of 15% may have limited predictive power over established risk factors.

Three recent studies (26-28) have attempted to incorporate genotype information from multiple markers into a single genetic test (26-28). Drenos et al (26) found individuals with six, and seven or more risk genotypes had a higher risk of CAD than those with only three or four risk genotypes (OR approximately 1.7 and 4.5, respectively), but did not compare results with traditional risk factors alone. Humphries et al (27) created a weighted model for CAD susceptibility based on four genotypes and three environmental interaction components. The resulting AUC was significantly better than that using traditional risk factors alone (AUC of 0.72 versus 0.62, P=0.01). However, as the authors stated, evaluating the effectiveness of a model based on the data that produced the model is perilous, and replication in different samples is required (27,29). Kathiresan et al (28) did not observe a significant change in the AUC between a 14 clinical covariate risk-prediction model with or without the addition of genotype information in nine cholesterol-associated genes. In a second retrospective analysis, Kathiresan et al (28) examined the number of individuals who would have been reclassified from the middle- to high-risk group (34 of 340), and five of these 34 subsequently had an event using a new statistic called the 'net reclassification improvement' (30). Many statisticians are supporting the validity of the statistic but caution that it does not fully address all factors required to determine whether a test should be in clinical use (31-33).

## SUMMARY

Advances in molecular genetics have made large-scale GWA studies of CAD a reality. The CAD-associated locus on chromosome 9p21 has generated considerable excitement. However, this locus may represent the 'low hanging fruit' for genetic susceptibility to CAD. Subsequent studies to find additional significantly associated markers with even more marginal RR ratios (eg, approximately 1.10 or smaller) require hundreds of thousands of patients; the clinical relevance of such markers would be even less obvious. Clinical practice guidelines currently dictate that a patient's management is based on the evaluation of traditional risk factors. It is possible that the inclusion of a genetic test could shift an individual between risk strata, affecting treatment decisions. However, before any such test is applicable to clinical practice, the finding must be applicable across populations, not to mention the implications on cost for testing and medication, possible negative consequences of over-treatment, and testing of the success of potential therapies. Further research to understand the architecture of genetic susceptibility, characterize the pathophysiological mechanisms underlying genetic associations, define interactions between genetic variants and the environment, discover new forms of variation and validate genetic associations across additional populations are all required before routine genotyping for CAD risk can be considered ready for 'prime time'.

**FUNDING:** This work was supported by operating grants from the Canadian Institutes of Health Research (FRN-13430 and MOP-79533), the Heart and Stroke Foundation of Ontario (T6018, NA6059 and PRG5967) and Genome Canada through the Ontario Genomics Institute. Dr Hegele is a Career Investigator of the Heart and Stroke Foundation of Ontario and holds the Edith Schulich Vinet Canada Research Chair (Tier I) in Human Genetics and the Jacob J Wolfe Distinguished Medical Research Chair. Matthew Lanktree is supported by the Canadian Institute of Health Research MD/PhD Studentship Award.

## REFERENCES

1. Assmann G, Cullen P, Schulte H. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Münster (PROCAM) study. Circulation 2002;105:310-5. (Erratum in 2002;105:900).
2. D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P; CHD Risk Prediction Group. Validation of the Framingham coronary heart disease prediction scores: Results of a multiple ethnic groups investigation. JAMA 2001;286:180-7.
3. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837-47.
4. Khot UN, Khot MB, Bajzer CT, et al. Prevalence of conventional risk factors in patients with coronary heart disease. JAMA 2003;290:898-904.
5. Ridker PM, Brown NJ, Vaughan DE, Harrison DG, Mehta JL. Established and emerging plasma biomarkers in the prediction of first atherothrombotic events. Circulation 2004;109(Suppl 1):IV6-19.
6. Greenland P, Knoll MD, Stamler J, et al. Major risk factors as antecedents of fatal and nonfatal coronary heart disease events. JAMA 2003;290:891-7.
7. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661-78.
8. Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 2007;316:1491-3.
9. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. Science 2007;316:1488-91.
10. Arnett DK, Baird AE, Barkley RA, et al; American Heart Association Council on Epidemiology and Prevention; American Heart Association Stroke Council; Functional Genomics and Translational Biology Interdisciplinary Working Group. Relevance of genetics and genomics for prevention and treatment of cardiovascular disease: A scientific statement from the American Heart Association Council on Epidemiology and Prevention, the Stroke Council, and the Functional Genomics and Translational Biology Interdisciplinary Working Group. Circulation 2007;115:2878-901.
11. Glazier AM, Nadeau JH, Aitman TJ. Finding genes that underlie complex traits. Science 2002;298:2345-9.
12. Guttmacher AE, Collins FS. Genomic medicine – a primer. N Engl J Med 2002;347:1512-20.
13. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 2005;6:95-108.
14. Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: Theoretical and practical concerns. Nat Rev Genet 2005;6:109-18.
15. Reich DE, Lander ES. On the allelic spectrum of human disease. Trends Genet 2001;17:502-10.
16. Yang Q, Khoury MJ, Friedman J, Little J, Flanders WD. How many genes underlie the occurrence of common complex diseases in the population? Int J Epidemiol 2005;34:1129-37.
17. Broeckel U, Hengstenberg C, Mayer B, et al. A comprehensive linkage analysis for myocardial infarction and its related risk factors. Nat Genet 2002;30:210-4.
18. Hegele RA. SNP judgments and freedom of association. Arterioscler Thromb Vasc Biol 2002;22:1058-61.

19. Samani NJ, Erdmann J, Hall AS, et al; WTCCC and the Cardiogenics Consortium. Genomewide association analysis of coronary artery disease. N Engl J Med 2007;357:443-53.
20. Willer CJ, Sanna S, Jackson AU, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat Genet 2008;40:161-9.
21. Kathiresan S, Melander O, Guiducci C, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. Nat Genet 2008;40:189-97.
22. Shen GQ, Li L, Rao S, et al. Four SNPs on chromosome 9p21 in a South Korean population implicate a genetic locus that confers high cross-race risk for development of coronary artery disease. Arterioscler Thromb Vasc Biol 2008;28:360-5.
23. Humphries SE, Ridker PM, Talmud PJ. Genetic testing for cardiovascular disease susceptibility: A useful clinical management tool or possible misinformation? Arterioscler Thromb Vasc Biol 2004;24:628-36.
24. Janssens AC, Pardo MC, Steyerberg EW, van Duijn CM. Revisiting the clinical validity of multiplex genetic testing in complex diseases. Am J Hum Genet 2004;74:585-8,588-9.
25. Cooper JA, Miller GJ, Humphries SE. A comparison of the PROCAM and Framingham point-scoring systems for estimation of individual risk of coronary heart disease in the Second Northwick Park Heart Study. Atherosclerosis 2005;181:93-100.
26. Drenos F, Whittaker JC, Humphries SE. The use of meta-analysis risk estimates for candidate genes in combination to predict coronary heart disease risk. Ann Hum Genet 2007;71:611-9.
27. Humphries SE, Cooper JA, Talmud PJ, Miller GJ. Candidate gene genotypes, along with conventional risk factor assessment, improve estimation of coronary heart disease risk in healthy UK men. Clin Chem 2007;53:8-16.
28. Kathiresan S, Melander O, Anevski D, et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. N Engl J Med 2008;358:1240-9.
29. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361-87.
30. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157-72,207-12.
31. Kraemer HC. Comments on 'Evaluating the added predictive ability of a new marker' by M. Pencina, R. D'Agostino, R. D'Agostino Jr, R. Vasan, Statistics in Medicine (DOI: 10.1002/sim.2929). Stat Med 2008;27:196-8.
32. Ware JH, Cai T. Comments on 'Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond' by M. J. Pencina et al., Statistics in Medicine (DOI: 10.1002/sim.2929). Stat Med 2008;27:185-7.
33. Chi YY, Zhou XH. The need for reorientation toward cost-effective prediction: Comments on 'Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond' by Pencina et al., Statistics in Medicine (DOI: 10.1002/sim.2929). Stat Med 2008;27:182-4.