Amanda Doran, MD, PhD
PGY-1
CRC Project
September 1, 2011

Research Question:

Is polymorphism in the novel murine atherosclerotic factor Id3 causally associated with cardiovascular disease in humans?

A. Study Purpose and Rationale

Atherosclerosis is a chronic, inflammatory disease in which lipids, cells and fibrous elements accumulate in the intimal layer of large arteries[1, 2]. This process begins in adolescence and follows a variable clinical course that may ultimately result in myocardial infarction or stroke[3, 4]. Because of this, it is estimated that atherosclerosis is the underlying cause of 50% of all deaths in Westernized society[1]. Despite the magnitude of this problem, understanding mechanisms whereby specific genes or gene pathways modulate atherosclerosis in humans remains challenging. This is largely attributed to the fact that atherosclerosis, in its common form, is a multifactorial disorder. While recent studies have identified potential candidate genes in humans, few studies have been able to determine the impact of specific human gene variants on protein function and disease modulation.

Recent studies have identified susceptibility loci for atherosclerosis in both humans and mice[5]. Using genomewide linkage analysis in a Caucasian American population, Wang et al. reported a novel significant susceptibility locus for premature myocardial infarction at 1p34-36 (12-40 Mbp)[6]. Welch et al., used the $LDLR^{-/-}$ mouse in an interspecific genetic cross to identify a murine atherosclerosis susceptibility locus on mouse chromosome 4 (27-154 Mbp), which they named Athsq1[7]. Intriguingly, comparison of human 1p34-36 with the Athsq1 locus in mice reveals that one major gene common to both loci is Inhibitor of Differentiation-3 (Id3).

Id3 is a member of the basic helix-loop-helix (bHLH) family of proteins. While Id3 contains the HLH domain required for protein:protein dimerization, it lacks the basic DNA-binding domain possessed by other members of the family, making it incapable of binding to DNA. Instead, Id3 binds to another subset of bHLH factors known as the E-proteins, including E12 and E47, thereby preventing their dimerization with tissue-specific bHLH factors and inhibiting subsequent DNA binding. These Id:E-protein dimers have been demonstrated to regulate many genes in a variety of cell types including B cells, T cells, adipocytes and smooth muscle cells[8-10].

Recently, our group has demonstrated an association between polymorphism at rs11574 in the *ID3* gene and carotid intima-medial thickness (IMT) in participants from the Diabetes Heart Study. Biochemical analyses performed in parallel showed that mutation of the major allele of the human *ID3* gene at rs11574 to the risk allele resulted in attenuated Id3 function, rendering it unable to interact with its bHLH partner proteins. Moreover, deletion of the *ID3* gene resulted in a significant increase in atherosclerosis formation in Western-fed $ApoE^{-/-}$ mice (an atherosclerosis-prone mouse model). Together, these results suggest that Id3 is an important atheroprotective factor in mice and in humans. It is unknown whether *ID3* polymorphism is associated with any cardiovascular outcomes such as MI and whether Id3 may play a causal role in the development of cardiovascular disease (CVD).

B. Study Design and Statistical Analysis

To determine the relationship between rs11574 and cardiovascular outcomes, a retrospective case-control mendelian randomization study will be performed using data compiled by the Coronary ARtery DIsease Genome-wide Replication And Meta-analysis (CARDIoGRAM) consortium. This dataset was compiled from 14 different studies, each of which was either a prospective cohort or case-control study. Of these, nine studies utilized technology that captures the *ID3* SNP of interest, rs11574. This yields roughly 16729 cases and 58006 controls within the aggregate data set (see table 1 below, section G). Cases were defined as patients who had angina or MI, and in some cases also included chest pain/symptoms necessitating PCI or CABG.

Our hypothesis is that within the population of CVD cases, the rs11574 will be more prevalent, suggesting a causal association between Id3 and CVD. Based on prior data, the carrier frequency of the rs11574 risk allele is 25-30% with approximately 10% of the population being homozygous for the risk allele. In order to detect an increase of 1% in the frequency of the SNP within cases, at 80% power and testing at p=0.05, the following would be necessary:

χ^2 estimation:

$n = 8 \times (p_1q_1 + p_2q_2/effect^2) + (2/effect) + 2$

$n = 8 \times ((0.1)(0.9) + (0.11)(0.89)/(0.01)^2) + (2/0.01) + 2$

$n = 15234$

Given the available data from the CARDIoGRAM consortium, in which there are far greater numbers of controls than cases, this estimate can be revised to approximately 9542 cases and 33112 controls (http://www.biomath.info/crc/) necessary for 80% power and p=0.05. Therefore, the CARDIoGRAM consortium data should be adequately powered to detect differences. If, however, the allele frequency and/or proportion of patients homozygous for the risk allele are significantly different than those in our original study population (Diabetes Heart Study), then this study may be underpowered.

Once genotypic data is obtained for cases and controls, an odds ratio can be calculated as follows:

	Homozygous WT allele	Homozygous risk allele	
Cases	A	B	~16729
Controls	C	D	~58006
	~67261	~7474	

$$OR = (A/C)/(B/D)$$

If rs11574 is indeed association with atherosclerosis, we would anticipate an OR of approximately 1.1, which is consistent with other genes found to be risk factors for atherosclerotic disease.

Because our preliminary study using the Diabetes Heart Study was limited to caucasians, it is possible that no effect will be seen in other ethnic groups and the population heterogeneity present in the CARDIoGRAM sample will obscure any effect of the SNP. Therefore, we propose to also examine the effect of rs11574 in a Caucasian subpopulation of the consortium data. Based on published data from these studies,

approximately 85% of the cases and controls were of European ancestry, which, given the large number of patients available in this data set, still provides an adequate number of subjects.

There are several limitations to this study. First, it is possible that Id3, by its nature as a transcriptional regulator upstream of a variety of genes, may influence other factors that can affect CVD. It is also possible that Id3 is in linkage dysequilibrium with another gene that is associated with CVD, particularly given its proximity to other genes that have been proposed to be associated with atherosclerosis in GWAS. It is possible that since CVD is multifactorial and most factors are responsible for only a small amount of variation, rs11574 will have too small an effect to be detected in the outcomes. False negative results can also arise from pleiotropy, canalization or gene-environment interaction, although these are difficult to determine or control.

C. Study Procedure
There are no additional procedures being performed on these patients.

D. Study Drugs
Not applicable.

E. Medical Device
Not applicable.

F. Study Questionnaires
Not applicable.

G. Study Subjects
This study will involve the use of a subset of data collected by 14 different prospective cohort or case-control studies, which has been compiled by the CARDIoGRAM consortium. A brief summary of the study subjects from each of the studies is included below (adapted from Preuss et al[11]). As outlined above, not all of the studies have complete data for the SNP of interest, therefore the total number of subjects for the analysis will more likely approximate 16729 cases and 58006 controls from 9 different studies (estimated based on the type of chip used for analysis in each study; actual numbers not available prior to request for data from the consortium).

Study	#Cases/Controls	%Female	Age	BMI	CVD Definition	Control Definition	Ref
CADomics	2078/2952	21.9/50.5	60.8(10.1)/55.3(10.8)	29.4(5.1)/27.0(4.7)	MI (based on EKG and enzymes)	Population sample with no history of MI	11
CHARGE	2287/22024	33.4/59.6	60.0(7.9)/63.1(8.0)	28.1(7.4)/27.5(8.0)	Definite/probable MI, or PTCA/CABG	No history of CVD defining condition	12
deCODE	6640/27611	36.3/61.9	74.8(11.8)/53.7(21.5)	27.7(4.7)/27.0(5.4)	MI (MONICA criteria), angina, CABG or PCI	Population sample with no history of CVD	13
GERMIFS II	1222/1287	33.1/48.3	51.4(7.5)/51.2(11.9)	29.0(3.8)/27.4(4.6)	MI (MONICA criteria)	Population sample with no history of CVD	14
GERMIFS III	1157/1748	20.1/48.9	58.6(8.7)/55.9(10.7)	27.0(3.6)/27.1(4.5)	MI (MONICA criteria)	Population sample with no history of	14

						CVD	
Luric 1	652/213	20.3/46.0	61.0(11.8)/58.3(12.1)	27.7(4.4)/27.4(4.2)	Angina, NSTEMI or STEMI	No or minor coronary lesions	15
Luric 2	486/296	23.4/48.6	63.7(9.4)/56.4(12.7)	27.1(3.8)/26.8(4.0)	Angina, NSTEMI or STEMI	No or minor coronary lesions	15
MIGen	1274/1407	37.2/39.9	42.4(6.6)/43.0(7.8)	27.6(5.2)/25.8(4.4)	MI	Hospital and community based controls	16
PennCATH	933/468	23.7/51.9	52.7(7.6)/61.7(9.6)	29.8(5.6)/28.9(6.4)	MI, CABG or PCI	Normal angiography	17

Table 1. Summary of studies from which aggregate data was compiled by the CARDIoGRAM consortium to be used in the proposed studies.

H. Recrultment of Subjects
Subjects have already been recruited to the individual aforementioned studies. No new subjects will be recruited.

I. Confidentiallty of Study Data
Data made available to outside investigators by the individual studies or the CARDIoGRAM consortium is already de-identified and coded, so issues regarding patient confidentiality should not arise.

J. Potential Conflict of Interest
None.

K. Location of the Study
Not applicable (no ongoing recruitment)

L. Potential Risks
As data has previously been collected and de-identified, there should be no risks specifically associated with this study.

M. Potential Benefits
There will be no direct benefits to patients.

N. Alternative Therapies
Not applicable.

O. Compensation to SubJects
Not applicable.

P. Costs to SubJects
Not applicable.

Q. Minors as Research Subjects
Not applicable.

R. Radiation or Radioactive Substances
Not applicable.

REFERENCES

1. Lusis AJ. Atherosclerosis. *Nature*. 2000;407:233-241
2. Geng YJ, Libby P. Progression of atheroma: A struggle between death and procreation. *Arterioscler Thromb Vasc Biol*. 2002;22:1370-1380
3. Stary HC. Natural history and histological classification of atherosclerotic lesions: An update. *Arterioscler Thromb Vasc Biol*. 2000;20:1177-1178
4. Berenson GS, Srinivasan SR, Bao W, Newman WP, 3rd, Tracy RE, Wattigney WA. Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults. The bogalusa heart study. *N Engl J Med*. 1998;338:1650-1656
5. Lusis AJ, Mar R, Pajukanta P. Genetics of atherosclerosis. *Annu Rev Genomics Hum Genet*. 2004;5:189-218
6. Wang Q, Rao S, Shen GQ, Li L, Moliterno DJ, Newby LK, Rogers WJ, Cannata R, Zirzow E, Elston RC, Topol EJ. Premature myocardial infarction novel susceptibility locus on chromosome 1p34-36 identified by genomewide linkage analysis. *Am J Hum Genet*. 2004;74:262-271
7. Welch CL, Bretschger S, Latib N, Bezouevski M, Guo Y, Pleskac N, Liang CP, Barlow C, Dansky H, Breslow JL, Tall AR. Localization of atherosclerosis susceptibility loci to chromosomes 4 and 6 using the ldlr knockout mouse model. *Proc Natl Acad Sci U S A*. 2001;98:7946-7951
8. Doran AC, Meller N, Cutchins A, Deliri H, Slayton RP, Oldham SN, Kim JB, Keller SR, McNamara CA. The helix-loop-helix factors id3 and e47 are novel regulators of adiponectin. *Circ Res*. 2008;103:624-634
9. Kumar MS, Hendrix JA, Johnson AD, Owens GK. Smooth muscle alpha-actin gene requires two e-boxes for proper expression in vivo and is a target of class i basic helix-loop-helix proteins. *Circ Res*. 2003;92:840-847
10. Murre C. Helix-loop-helix proteins and lymphocyte development. *Nat Immunol*. 2005;6:1079-1086
11. Preuss M, Konig IR, Thompson JR, Erdmann J, Absher D, Assimes TL, Blankenberg S, Boerwinkle E, Chen L, Cupples LA, Hall AS, Halperin E, Hengstenberg C, Holm H, Laaksonen R, Li M, Marz W, McPherson R, Musunuru K, Nelson CP, Burnett MS, Epstein SE, O'Donnell CJ, Quertermous T, Rader DJ, Roberts R, Schillert A, Stefansson K, Stewart AF, Thorleifsson G, Voight BF, Wells GA, Ziegler A, Kathiresan S, Reilly MP, Samani NJ, Schunkert H. Design of the coronary artery disease genome-wide replication and meta-analysis (cardiogram) study: A genome-wide association meta-analysis involving more than 22 000 cases and 60 000 controls. *Circ Cardiovasc Genet*. 2010;3:475-483
12. Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, Uitterlinden AG, Harris TB, Witteman JC, Boerwinkle E. Cohorts for heart and aging research in genomic epidemiology (charge) consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet*. 2009;2:73-80
13. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, Sigurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levey AI, Backman VM, Matthiasdottir S, Jonsdottir T, Palsson S, Einarsdottir H, Gunnarsdottir S, Gylfason A, Vaccarino V, Hooper WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgeirsson G, Thorsteinsdottir U, Kong A, Stefansson K. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*. 2007;316:1491-1493

14. Erdmann J, Grosshennig A, Braund PS, Konig IR, Hengstenberg C, Hall AS, Linsel-Nitschke P, Kathiresan S, Wright B, Tregouet DA, Cambien F, Bruse P, Aherrahrou Z, Wagner AK, Stark K, Schwartz SM, Salomaa V, Elosua R, Melander O, Voight BF, O'Donnell CJ, Peltonen L, Siscovick DS, Altshuler D, Merlini PA, Peyvandi F, Bernardinelli L, Ardissino D, Schillert A, Blankenberg S, Zeller T, Wild P, Schwarz DF, Tiret L, Perret C, Schreiber S, El Mokhtari NE, Schafer A, Marz W, Renner W, Bugert P, Kluter H, Schrezenmeir J, Rubin D, Ball SG, Balmforth AJ, Wichmann HE, Meitinger T, Fischer M, Meisinger C, Baumert J, Peters A, Ouwehand WH, Deloukas P, Thompson JR, Ziegler A, Samani NJ, Schunkert H. New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat Genet*. 2009;41:280-282
15. Winkelmann BR, Marz W, Boehm BO, Zotz R, Hager J, Hellstern P, Senges J. Rationale and design of the luric study--a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. *Pharmacogenomics*. 2001;2:S1-73
16. Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, Anand S, Engert JC, Samani NJ, Schunkert H, Erdmann J, Reilly MP, Rader DJ, Morgan T, Spertus JA, Stoll M, Girelli D, McKeown PP, Patterson CC, Siscovick DS, O'Donnell CJ, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Melander O, Altshuler D, Merlini PA, Berzuini C, Bernardinelli L, Peyvandi F, Tubaro M, Celli P, Ferrario M, Fetiveau R, Marziliano N, Casari G, Galli M, Ribichini F, Rossi M, Bernardi F, Zonzin P, Piazza A, Yee J, Friedlander Y, Marrugat J, Lucas G, Subirana I, Sala J, Ramos R, Meigs JB, Williams G, Nathan DM, MacRae CA, Havulinna AS, Berglund G, Hirschhorn JN, Asselta R, Duga S, Spreafico M, Daly MJ, Nemesh J, Korn JM, McCarroll SA, Surti A, Guiducci C, Gianniny L, Mirel D, Parkin M, Burtt N, Gabriel SB, Thompson JR, Braund PS, Wright BJ, Balmforth AJ, Ball SG, Hall AS, Linsel-Nitschke P, Lieb W, Ziegler A, Konig I, Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann HE, Schreiber S, Ouwehand W, Deloukas P, Scholz M, Cambien F, Li M, Chen Z, Wilensky R, Matthai W, Qasim A, Hakonarson HH, Devaney J, Burnett MS, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Knouff CW, Waterworth DM, Walker MC, Mooser V, Epstein SE, Scheffold T, Berger K, Huge A, Martinelli N, Olivieri O, Corrocher R, McKeown P, Erdmann E, Konig IR, Holm H, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Do R, Xie C, Siscovick D. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet*. 2009;41:334-341
17. Lehrke M, Millington SC, Lefterova M, Cumaranatunge RG, Szapary P, Wilensky R, Rader DJ, Lazar MA, Reilly MP. Cxcl16 is a marker of inflammation, atherosclerosis, and acute coronary syndromes in humans. *J Am Coll Cardiol*. 2007;49:442-449