Published in final edited form as:

Hum Genet. 2010 January ; 127(1): 101–105. doi:10.1007/s00439-009-0761-3.

The transcription factor GATA-2 does not associate with angiographic coronary artery disease in the Ottawa Heart Genomics and Cleveland Clinic GeneBank Studies

Sonny Dandona,

The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

Li Chen,

The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

Meng Fan,

The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

Md Afaque Alam,

The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

Olivia Assogba,

The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

Melanie Belanger,

The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

Kathryn Williams,

The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

George A. Wells,

The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

W. H. Wilson Tang,

Cleveland Clinic Foundation, Cleveland, OH, USA

Stephen G. Ellis,

Cleveland Clinic Foundation, Cleveland, OH, USA

Stanley L. Hazen,

Cleveland Clinic Foundation, Cleveland, OH, USA

Ruth McPherson,

The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

Robert Roberts, and

The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

Alexandre F. R. Stewart

The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada astewart@ottawaheart.ca

Abstract

The transcription factor GATA2 was reported to associate with coronary artery disease (CAD) in the family-based Genecard sample (Connelly et al. in *PLoS Genet* 2:e139, 2006). We asked whether GATA2 associates with sporadic cases of CAD in the Ottawa Heart Genomics Study (OHGS) and Cleveland Clinic (CC) populations. We genotyped the lead single nucleotide polymorphism (SNP) from Genecard, rs2713604 which is located in intron 5–6 of GATA2 in 600 CAD cases and 625 controls, as well as a tag SNP rs1573949 ($r^2 = 0.87$ in Caucasians of European ancestry in Utah from HapMap) in 1,136 cases and 1,162 controls in the OHGS1 population. A further 1,838 CAD cases and 913 controls derived from an independent sample combining genotypes from CC and OHGS2 populations were genotyped for rs1573949. Neither of the genotyped SNPs associates with CAD in the OHGS1 or CC/OHGS2 populations. Our data suggest that GATA2 does not contribute to the development of angiographic CAD among sporadic cases.

Introduction

Coronary artery disease (CAD) is the leading killer in industrialized nations (Murray and Lopez 1997). It represents a major economic burden to health care delivery systems; the estimated annual expenditure for the management of its consequences is in excess of \$100 billion in the United States (Lloyd-Jones et al. 2009). Coronary artery disease demonstrates a strong familial component even after adjustment for traditional risk factors (Murabito et al. 2005; Rissanen 1979; Shea et al. 1984; ten Kate et al. 1982; Thomas and Cohen 1955). More nuanced screening of the population will aid in early identification of patients at risk for development of CAD and initiation of treatment paradigms that could prevent its late sequelae. This is contingent upon defining variation in the genetic blueprint that confers risk for development of disease. Unlike, simple Mendelian traits where the genetic variant is not only sufficient but necessary for the development of the phenotype of interest, coronary artery disease is a complex trait characterized by genetic modifiers that interact with environmental cues to only modulate the risk of disease. Thus far, many of the variants that have been identified to confer risk for coronary artery disease have not been successfully replicated in subsequent populations. Examples include the initial descriptions of disease association and subsequent non-association with polymorphisms in lymphotoxin- α (Clarke et al. 2006; Ozaki et al. 2002) and TNFSF4 (tumor necrosis factor superfamily, member 4) (Koch et al. 2008; Wang et al. 2005). While this may reflect a weak biological effect, it may also be secondary to the significant phenotypic heterogeneity employed between studies. These sequence variants must be distinguished from those that show consistent association across populations, such as polymorphisms in PCSK9 (proprotein convertase subtilisin/kexin type 9) (Cohen et al. 2006; Kathiresan 2008) and sequence variation at chromosome 9p21.3 (Helgadottir et al. 2007; Samani et al. 2007; McPherson et al. 2007). 9p21.3 has been successfully replicated across CAD subphenotypes (Helgadottir et al. 2007; Samani et al. 2007; McPherson et al. 2007). It is of considerable interest to attempt to replicate loci that have been associated with coronary artery disease in independent populations across phenotypes. This aids in validating the locus in question in addition to pinpointing the biology that underlies the association. That is, those loci that associate with angiographic coronary artery disease and not myocardial infarction, likely affect the mechanisms that govern deposition of atheroma, whereas those that are associated with

myocardial infarction but not atheromatous coronary artery disease likely affect plaque rupture and/or thrombosis. Recently, sequence variants in GATA2 were shown to associate with familial coronary artery disease (Connelly et al. 2006). We asked whether the same variants associate with sporadic angiographically defined CAD.

Methods

This study was approved by the Research Ethics Board of the University of Ottawa Heart Institute and written informed consent was obtained from all study participants. Informed consent was approved from the CC institutional review board for the CC cohort. Males below the age of 55 years and females below the age of 65 years with at least one epicardial stenosis >50% were eligible to be cases. Controls were those who did not carry a diagnosis of CAD by the age of 65 years for male and 70 years for females. Specimens from Cleveland Clinic were obtained from subjects who were prospectively enrolled in GeneBank, a large and well-characterized clinical repository with clinical and longitudinal data comprised from consenting subjects undergoing elective diagnostic cardiac catheterization. All GeneBank participants gave written informed consent and the Institutional Review Board of the Cleveland Clinic approved the study protocol. Genomic DNA was purified from peripheral blood in all study participants. For the lead GATA2 SNP rs2713604, 600 cases and 625 controls were genotyped using the following primers: forward, 5'-CCC AGC CAC AAT ATC CAG C-3' and reverse, 5'-CAT AGT CCC ATC ACC AGA TGG C-3' by the polymerase chain reaction on a Roche Lightcycler (Roche Applied Science, Laval, QC, Canada). The tagging SNP rs1573949 was genotyped by the Affymetrix 6.0 array in 1,136 cases and 1,162 controls. Samples analyzed by the 6.0 arrays yielded an average call rate by the Birdseed algorithm of 99.3%. Population stratification was analyzed by the principal components method using Eigenstrat (Price et al. 2006). Prior to association analysis, SNPs were excluded if deviated from Hardy-Weinberg equilibrium ($p < 0.001$) in the control samples, or yielded an individual SNP call rate of <95%, or had MAF < 0.05 in either cases or controls. Associations between SNPs and CAD were analyzed with additive tests using a threshold of $p < 0.001$. Analyses were adjusted for gender, BMI, and smoking.

Results

Table 1 lists baseline characteristics for individuals genotyped for the GATA2 SNP rs2713604 by Lightcycler® and for the linked SNP rs1573949 using Affymetrix® 6.0 arrays, respectively, in the OHGS 1 population. Table 1 also lists the baseline characteristics for the CC/OHGS 2 sample. The lead SNP rs2713604 and the tagging SNP rs1573949 demonstrate higher frequencies among controls than cases in the OHGS1 population and are not associated with risk for CAD (Table 2). Similarly, in an independent sample, the CC/OHGS2, the tagging SNP rs1573949 again demonstrates a higher frequency in controls than in cases and is not associated with risk for CAD (Table 2). To show that these samples accurately detect variants associated with CAD, we confirmed rs1333049 at 9p21 to be highly associated with CAD in both the OHGS1 and CC/OHGS2 samples.

Discussion

Neither the lead SNP, rs2713604, nor the tagging SNP, rs1573949, ($r^2 = 0.87$) are associated with angiographic coronary artery disease in the OHGS or the OHGS2/CC cohorts. Moreover, the frequency of each of the SNPs is higher among controls than in cases (i.e. in our dataset, the relative frequencies between cases and controls would suggest a protective effect of this variant, if any). In the report by Connelly et al. the association was seen in a family-based sample, Genecard, where it is conceivable that genetic factors play a relatively greater role in determination of phenotype than in cohorts comprised by sporadic cases. The authors validate

rs2713604 in a sporadic case-based sample. Furthermore, they have attempted to replicate this finding in the Intermountain Heart Collaborative Study. While the association was reported as suggestive, it did not meet the $p < 0.05$ threshold among those with the primary case phenotype versus controls; its significance is further undermined by the fact that Bonferroni correction would demand a p -value of <0.0045 given that 11 SNPs were genotyped simultaneously (Horne et al. 2009). While the association is nominally significant in the Cathgen sample, it, too, does not achieve statistical significance once corrected for multiple hypothesis testing by the Bonferroni method. The authors compared 656 Cathgen spontaneous cases to 410 controls and found a nominally significant association (Connelly et al. 2006). We have compared 3,574 cases and 2,700 controls and therefore our inability to detect association is not due to insufficient statistical power. Moreover, it is noteworthy that in each of our subpopulations, the frequency of the polymorphisms in GATA2 is higher among controls than cases, suggesting a protective rather than a risk-conferring action. The adequacy of our populations to detect polymorphisms associated with CAD is confirmed by the robust association seen with sequence variation at 9p21 with angiographic CAD in both OHGS1 and CC/OHGS2 (Table 3).

Few loci have been replicated across populations. The great majority have yielded association that is far less consistent. This may be due to phenotypic heterogeneity between studies; while some studies interrogated CAD populations defined by angiography, others used myocardial infarction-based populations. In this instance, the case definitions are fairly similar in that in both the OHGS and the Cathgen /Genecard samples the cases are those with significant atheromatous burden detected by coronary angiography. However, in the Cathgen sample the age of 55 years is used as a threshold for both males and females, while in the OHGS, a threshold of 65 years is used for females. It is highly unlikely that the lack of even nominal association could be explained on this basis in the OHGS. It should be noted that the Cathgen and Genecard samples are ethnically heterogeneous, including a significant number of American Indians, African Americans and Asians; the OHGS and CC populations are exclusively Caucasian-derived. Moreover, at inception diabetics were excluded from the OHGS and CC populations in order to identify genetic variants that predisposed uniquely to CAD; the Cathgen and Genecard populations include a significant number of diabetics. Though unlikely, it is possible that the observation noted by the Genecard/Cathgen investigators represents a race-specific or diabetes dependent association.

A weak biological effect of the variant of interest can also result in the inability to consistently replicate association; if each cycle of replication was undertaken in large sample sets, then statistically significant associations could be replicated. Alternatively, meta-analyses can be performed that incorporate data from multiple sources. This assumes, of course, that the variant does elicit an effect, albeit mild, and that the seminal association is not spurious. Once genetic association is established, the mechanisms that underlie the association must be established as this will improve our understanding of the pathogenesis of CAD and will possibly lead to the identification of novel therapeutic targets. Secondly, the use of genetic markers in the clinical arena may be useful in clinical risk assessment. This would require demonstrating the incremental benefit of testing for the variant of interest allied with standard tools of risk estimation such as the Framingham risk score in prospective cohorts. Before such studies are undertaken consistent association in a case-control format is necessary.

In genetic association studies, it is possible that the variant of interest is merely a surrogate for the bona fide biological agent and its eventual identification may facilitate subsequent replication as the statistical association may well be more robust. This may be the case with GATA2 although it remains a good candidate, given its expression in the aortic endothelial and smooth muscle cells.

In order to develop a comprehensive inventory of the genetic variants that affect CAD-related phenotypes, alliances will need to be formed in order to develop large datasets limited to specific CAD-related phenotypes. This will facilitate the reliable identification disease-causing variants, including those that elicit a mild biologic action.

Acknowledgments

This research was supported by grants from the Canadian Institutes of Health Research #MOP82810, #NA6650, #MOP77682, the Canada Foundation for Innovation CFI #11966, the Heart and Stroke Foundation of Ontario #NA6001, the National Institutes of Health grants P01HL087018, P01 HL076491 and P01 HL077107 (S.L.H.) and the Cleveland Clinic Clinical Research Unit of the Cleveland Clinic/Case Western Reserve University CTSA (1UL1RR024989).

References

- Clarke R, Xu P, Bennett D, Lewington S, Zondervan K, Parish S, Palmer A, Clark S, Cardon L, Peto R, Lathrop M, Collins R. Lymphotoxin-alpha gene and risk of myocardial infarction n 6, 928 cases and 2, 712 controls in the ISIS case-control study. *PLoS Genet* 2006;2:e107. [PubMed: 16839190]
- Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med* 2006;354:1264–1272. [PubMed: 16554528]
- Connelly JJ, Wang T, Cox JE, Haynes C, Wang L, Shah SH, Crosslin DR, Hale AB, Nelson S, Crossman DC, Granger CB, Haines JL, Jones CJ, Vance JM, Goldschmidt-Clermont PJ, Kraus WE, Hauser ER, Gregory SG. GATA2 is associated with familial early-onset coronary artery disease. *PLoS Genet* 2006;2:e139. [PubMed: 16934006]
- Helgadóttir A, Thorleifsson G, Manolescu A, Gretarsdóttir S, Blondal T, Jonasdóttir A, Sigurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levey AI, Backman VM, Matthiassdóttir S, Jonsdóttir T, Palsson S, Einarssdóttir H, Gunnarsdóttir S, Gylfason A, Vaccarino V, Hooper WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgerisson G, Thorsteinsdóttir U, Kong A, Stefansson K. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 2007;316:1491–1493. [PubMed: 17478679]
- Horne BD, Hauser ER, Wang L, Muhlestein JB, Anderson JL, Carlquist JF, Shah SH, Kraus WE. Validation study of genetic associations with coronary artery disease on chromosome 3q13-21 and potential effect modification by smoking. *Ann Hum Genet* 2009;73(Pt 6):551–558. [PubMed: 19706030]
- Kathiresan S. A PCSK9 missense variant associated with a reduced risk of early-onset myocardial infarction. *N Engl J Med* 2008;358:2299–2300. [PubMed: 18499582]
- Koch W, Hoppmann P, Mueller JC, Schomig A, Kastrati A. Lack of support for association between common variation in TNFSF4 and myocardial infarction in a German population. *Nat Genet* 2008;40:1386–1387. author reply 1387–1388. [PubMed: 19029970]
- Lloyd-Jones D, Adams R, Carnethon M, De Simone G, Ferguson TB, Flegal K, Ford E, Furie K, Go A, Greenlund K, Haase N, Hailpern S, Ho M, Howard V, Kissela B, Kittner S, Lackland D, Lisabeth L, Marelli A, McDermott M, Meigs J, Mozaffarian D, Nichol G, O'Donnell C, Roger V, Rosamond W, Sacco R, Sorlie P, Stafford R, Steinberger J, Thom T, Wasserthiel-Smoller S, Wong N, Wylie-Rosett J, Hong Y. Heart disease and stroke statistics—2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* 2009;119:480–486. [PubMed: 19171871]
- McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC. A common allele on chromosome 9 associated with coronary heart disease. *Science* 2007;316:1488–1491. [PubMed: 17478681]
- Murabito JM, Pencina MJ, Nam BH, D'Agostino RB Sr, Wang TJ, Lloyd-Jones D, Wilson PW, O'Donnell CJ. Sibling cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults. *JAMA* 2005;294:3117–3123. [PubMed: 16380592]
- Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet* 1997;349:1498–1504. [PubMed: 9167458]

- Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, Sato H, Hori M, Nakamura Y, Tanaka T. Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nat Genet* 2002;32:650–654. [PubMed: 12426569]
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38:904–909. [PubMed: 16862161]
- Rissanen AM. Familial occurrence of coronary heart disease: effect of age at diagnosis. *Am J Cardiol* 1979;44:60–66. [PubMed: 453047]
- Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, Barrett JH, König IR, Stevens SE, Szymczak S, Tregouet DA, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth AJ, Baessler A, Ball SG, Strom TM, Braenne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H. Genome-wide association analysis of coronary artery disease. *N Engl J Med* 2007;357:443–453. [PubMed: 17634449]
- Shea S, Ottman R, Gabrieli C, Stein Z, Nichols A. Family history as an independent risk factor for coronary artery disease. *J Am Coll Cardiol* 1984;4:793–801. [PubMed: 6481018]
- ten Kate LP, Boman H, Daiger SP, Motulsky AG. Familial aggregation of coronary heart disease and its relation to known genetic risk factors. *Am J Cardiol* 1982;50:945–953. [PubMed: 7137046]
- Thomas CB, Cohen BH. The familial occurrence of hypertension and coronary artery disease, with observations concerning obesity and diabetes. *Ann Intern Med* 1955;42:90–127. [PubMed: 13229192]
- Wang X, Ria M, Kelmenson PM, Eriksson P, Higgins DC, Samnegard A, Petros C, Rollins J, Bennet AM, Wiman B, de Faire U, Wennberg C, Olsson PG, Ishii N, Sugamura K, Hamsten A, Forsman-Semb K, Lagercrantz J, Paigen B. Positional identification of TNFSF4, encoding OX40 ligand, as a gene that influences atherosclerosis susceptibility. *Nat Genet* 2005;37:365–372. [PubMed: 15750594]

Table 1

Baseline characteristics

	OHGS1 samples genotyped by PCR		OHGS1 samples genotyped on Affymetrix 6.0		CC/OHGS2 samples genotyped on Affymetrix 6.0	
	CAD cases n = 600	Controls n = 625	CAD cases n = 1,136	Controls n = 1,162	CAD cases n = 1,912	Controls n = 933
Male (%)	73.5	47.3	73.4	49.7	75.9	49.7
Age (years)	49.5 ± 7.1	73.3 ± 6.3	49.6 ± 7.4	74.6 ± 5.5	48.9 ± 7.3	74.3 ± 5.8
MI (%)	53.1	NA	55.2	NA	60.5	NA
PCI (%)	32.2	NA	32.5	NA	44.0	NA
CABG (%)	27.0	NA	29.8	NA	41.5	NA
Cholesterol (mmol/L)	5.9 ± 1.2	5.8 ± 1.1	6.0 ± 1.3	5.7 ± 1.1	6.0 ± 1.3	5.7 ± 1.0
Triglycerides (mmol/L)	2.5 ± 1.7	1.5 ± 1.2	2.4 ± 1.8	1.4 ± 0.9	2.4 ± 1.9	1.4 ± 1.0
LDL-C (mmol/L)	3.7 ± 1.1	3.7 ± 1.5	3.8 ± 1.1	3.6 ± 1.2	3.9 ± 1.2	3.6 ± 1.4
HDL-C (mmol/L)	1.2 ± 0.6	1.5 ± 0.4	1.1 ± 0.4	1.5 ± 0.5	1.1 ± 0.36	1.47 ± 0.45
Hypertension (%)	52.6	43.9	54.2	42.7	68.9	43.7
Cigarette smoking (ever) (%)	75.8	51.8	76.0	51.9	76.2	48.9
BMI (kg/m ²)	28.4 ± 4.8	26.6 ± 4.1	28.4 ± 5.2	26.7 ± 6.6	29.1 ± 5.4	26.8 ± 7.3

For cases age at diagnosis; for controls age at recruitment

NA not applicable

Table 2

Associations between CAD and GATA2

	MAF	AA	AG	GG	Odds ratio (95% CI)	Trend <i>p</i>
OHGS1						
rs2713604						
Cases <i>n</i> = 600	0.338	73	260	267	0.85 (0.71, 1.01)	0.068
Controls <i>n</i> = 625	0.359	89	271	265		
OHGS2						
rs1573949						
Cases <i>n</i> = 1,136	0.346	145	496	495	0.99 (0.88, 1.13)	0.925
Controls <i>n</i> = 1,162	0.350	159	495	508		
rs1573949						
Cases <i>n</i> = 1,838	0.335	232	769	837	0.92 (0.82, 1.02)	0.193
Controls <i>n</i> = 913	0.353	133	378	402		

Adjusted for age, gender, BMI and smoking history

Table 3

Associations between CAD and 9p21

	MAF	CC	CG	GG	Odds ratio (95% CI)	Trend <i>p</i>
OHGS1						
rs1333049						
Cases <i>n</i> = 1,523	0.568	483	765	275	1.44 (1.29,1.62)	2.13 × 10 ⁻¹⁰
Controls <i>n</i> = 1,454	0.474	327	725	402		
CC/OHGS2						
rs1333049						
Cases <i>n</i> = 1,900	0.525	597	954	349	1.44 (1.29,1.62)	1.70 × 10 ⁻⁹
Controls <i>n</i> = 932	0.435	212	461	259		

Adjusted for age, gender, BMI, and smoking history