

RESEARCH ARTICLE

Rare variants in previously identified linkage regions associated with carotid plaque in Dominican Republic families

Nicole D. Dueker¹, Ashley Beecham¹, Liyong Wang^{1,2}, Chuanhui Dong³, Ralph L. Sacco^{2,3,4,5}, Susan H. Blanton^{1,2}, Tatjana Rundek^{3,4,5*}

1 John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL, United States of America, **2** Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, FL, United States of America, **3** Department of Neurology, Miller School of Medicine, University of Miami, Miami, FL, United States of America, **4** Department of Public Health Sciences, Miller School of Medicine, University of Miami, Miami, FL, United States of America, **5** Evelyn F. McKnight Brain Institute, Department of Neurology, University of Miami, Miami, FL, United States of America

* trundek@med.miami.edu



OPEN ACCESS

Citation: Dueker ND, Beecham A, Wang L, Dong C, Sacco RL, Blanton SH, et al. (2022) Rare variants in previously identified linkage regions associated with carotid plaque in Dominican Republic families. PLoS ONE 17(1): e0250799. <https://doi.org/10.1371/journal.pone.0250799>

Editor: Regie Lyn Pastor Santos-Cortez, University of Colorado Denver - Anschutz Medical Campus, UNITED STATES

Received: March 4, 2021

Accepted: April 13, 2021

Published: January 12, 2022

Copyright: © 2022 Dueker et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data is now publicly available: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs002406.v1.p1.

Funding: This work was supported by grants from the National Institutes of Health/National Institute for Neurological Disorders and Stroke (R01NS040807 and R01NS29993) and the Evelyn F. McKnight Brain Institute.

Abstract

Carotid plaque is a subclinical measure of atherosclerosis. We have previously shown measures of carotid plaque to be heritable in a sample of 100 Dominican families and found evidence for linkage and association of common variants (CVs) on 7q36, 11p15, 14q32 and 15q23 with plaque presence. Our current study aimed to refine these regions further and identify rare variants (RVs) influencing plaque presence. Therefore, we performed targeted sequencing of the one LOD unit down region on 7q36, 11p15, 14q32 and 15q23 in 12 Dominican families with evidence for linkage to plaque presence. Gene-based RV analyses were performed using the Sequence Association Test for familial data (F-SKAT) under two filtering algorithms; 1. all exonic RVs and 2. non-synonymous RVs. Replication analyses were performed using a sample of 22 Dominican families and 556 unrelated Dominicans with Exome Array data. To identify additional non-synonymous RVs influencing plaque, we looked for co-segregation of RVs with plaque in each of the sequenced families. Our most strongly associated gene with evidence for replication was *AMPD3* which showed suggestive association with plaque presence in the sequenced families (exonic RV $p = 0.003$, non-synonymous RV $p = 0.005$) and replication families (exonic RV $p = 0.04$, nonsynonymous RV $p = 0.02$). Examination of the sequenced family pedigrees revealed two missense variants on chromosome 11 which co-segregated with plaque presence in one of our families; rs61751342 (located in *DENND2B*), and rs61760882 (located in *RNF141*). The rs61751342 missense variant is an eQTL for *SCUBE2* in the atrial appendage. Notably, *SCUBE2* encodes a protein which interacts with vascular endothelial growth factor (VEGF) receptor 2 to regulate VEGF-induced angiogenesis, thus providing biologic plausibility for this gene in atherosclerosis. In conclusion, using targeted sequencing of previously-identified linkage regions, we have identified suggestive evidence for the role of RVs in carotid plaque pathogenesis.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Carotid plaque, a well-established subclinical marker of atherosclerosis, is a known predictor of stroke and myocardial infarction [1]. Heritability studies show evidence for a genetic contribution to plaque [2–5], with heritability estimates ranging from 13%–52% in various populations. A recent study also showed evidence for shared genes between various subclinical measures of atherosclerosis, including plaque, and clinical outcomes [6]. Further evidence for a genetic contribution to plaque comes from genome-wide linkage and association studies. The largest genome-wide association study (GWAS) identified two loci significantly associated with plaque presence (*PIK3CG* and *EDNRA*) and two loci that were suggestively associated (*LDLR*, and *LRIG1*) [7]. Smaller studies of plaque phenotypes have implicated additional regions including 9p21, 10q24 [8] and chromosome 17 [5].

In addition to these studies, we have taken a multi-stage approach to identify genetic variants influencing plaque presence in a large sample of extended families from the Dominican Republic. Within this sample, we have shown plaque presence to be heritable ($h^2 = 50\%$) and genome-wide linkage analysis identified four chromosomal regions linked to plaque presence; 7q36, 11p15, 14q32 and 15q23. Further, we have shown evidence for the association of common variants (CVs) within these regions [2, 9]. Our current study aimed to further characterize these regions by sequencing the regions under the linkage peaks to identify rare variants contributing to carotid plaque in these families.

Materials and methods

Study samples

Study participants were drawn from two studies; the Family Study of Stroke Risk and Carotid Atherosclerosis and the Northern Manhattan Stroke Study (NOMAS), a population-based cohort study. Details of both study designs have been published previously [10, 11]. Briefly, NOMAS enrolled 3,298 stroke-free individuals between 1993 and 2001. A subset of Caribbean Hispanic probands from NOMAS who were at high risk for cardiovascular were subsequently enrolled in the family study. To be eligible for inclusion in the family study, probands needed to have a high risk of cardiovascular disease (defined as 1. Having a sibling with a history of myocardial infarction or stroke, or 2. Having two of the following three risk phenotypes; maximal carotid plaque thickness, left ventricular mass or homocysteine level above the 75th percentile in NOMAS). They also needed to have the ability to provide a family history, obtain family members' permission to be contacted by research staff, and have at least three first-degree relatives able to participate [10]. Probands for the family study were enrolled in northern Manhattan and family members were enrolled at two locations: New York at Columbia University and in the Dominican Republic (DR) at the Clinicas Corazones Unidos in Santo Domingo. All participants provided written informed consent and the study was approved by the Institutional Review Boards of Columbia University, the University of Miami, the National Bioethics Committee and the Independent Ethics Committee of Instituto Oncologico Regional del Cibao in the DR (20070478 and 20072012).

A total of 34 families with a family-specific LOD score >0.1 at one of the four QTLs for plaque located at chromosomes 7q36, 11p15, 14q32 and 15q23, were selected for our current study; 12 families were sequenced as part of our discovery analyses and 22 families were genotyped on the Exome Array for our replication analyses. Details of the family selection have been published previously [12]. Additional replication analyses were performed using an independent sample of 556 unrelated Dominicans from NOMAS.

Carotid plaque phenotypes and risk factor measurements

High-resolution B-mode 2-dimensional carotid ultrasound imaging was performed to measure atherosclerotic plaque at the internal and common carotid arteries and bifurcations, according to standardized scanning and reading protocols [13]. Details of our carotid plaque phenotyping have been previously published [2, 13]. Briefly, in-depth imaging of plaques was performed in long axes and multiple angles. The optimized and normalized images were analyzed offline by automated computerized edge detection system M'ATH (Intelligence in Medical Technologies, Inc, Paris, France) and area of each plaque was measured. For each participant, the sum of all plaque areas (mm^2) within each subject was calculated and expressed as a total carotid plaque area. Plaque presence was then defined as an area of focal wall thickening over 50% greater than surrounding wall thickness in millimeters.

Traditional risk factor data on body mass index (BMI), waist-hip ratio (WHR), diabetes, hypertension and pack-years smoking, was collected during a standardized interview [14]. BMI was defined as weight (kg) divided by height (m)² and WHR was defined as waist circumference divided by hip circumference. Presence of diabetes was defined as fasting blood glucose level ≥ 126 mg/dl or self-reported history of diabetes. Presence of hypertension was defined as self-reported history of high blood pressure, systolic blood pressure ≥ 140 mmHg, diastolic blood pressure ≥ 90 mmHg, or self-reported use of antihypertensive medication. Pack-years of smoking were calculated as the number of cigarettes smoked per day divided by 20 and then multiplied by the number of years smoking. All data collection procedures were standardized across NOMAS and the Family Study.

Discovery sample sequencing and quality control

Genomic DNA was isolated from whole blood in family study participants. Targeted sequencing of the exons in genes within the one LOD unit down region on 7q36 (chr7:150-160Mb), 11p15 (chr11:7-25 Mb), 14q32 (chr14:78-102 Mb), and 15q23 (chr15: 70-95 Mb), as well as targeted sequencing beyond the exons for candidate genes identified previously [9], was performed using a customized Agilent SureSelect Enrichment Kit. Details of our sequencing methods have been previously described [15]. Briefly, DNA libraries were sequenced on an Illumina HiSeq2000 and the Burrows-Wheeler Aligner was used to align the raw sequencing reads to the human reference sequence hg19 [16]. The Genome Analysis ToolKit was used to call variants and potential functional annotations of variants were obtained using ANNOVAR v. 2016Feb01 and SeattleSequtation 138.

Quality control was conducted at both variant and sample levels, as described previously [12, 15]. Within each sample, variants having depth < 4 or Phred-Like score < 100 were set to missing. Variants with VQSLOD < -4 or call rate $< 75\%$ were removed from analysis. Individuals missing plaque measures ($n = 4$) were removed prior to segregation analyses and an additional 13 individuals missing covariate values were removed prior to association analyses. All individuals had high concordance ($\geq 95\%$) between the sequencing data and available genotype data. For the remaining family study samples, pedigree structure was confirmed using the Graphical Relationship Representation software v. 1.2.1.41. Mendelian error checking was performed, and Mendelian errors were set to missing for all the variants called using PLATO v. 0.84 [17].

Replication sample genotyping and quality control

A total of 698 NOMAS participants and 378 individuals within our 22 replication families were genotyped using the Illumina HumanExome-24v1_B Beadchip, at the Hussman Institute for Human Genomics in the Center for Genome Technology (Miami, FL). Our Exome Array

included custom exonic variants selected on the basis of sequencing data obtained in the discovery family data set. Details of the variant selection have been described previously [15].

We removed 39 NOMAS individuals and 18 replication family members due to unexpected duplication or relatedness, gender discrepancy, or low call rate (<98%). Relatedness was assessed in NOMAS using PLINK and all pairs of individuals had π -hat < 0.18. For our current study, individuals missing plaque measures and/or covariate values were also removed prior to association analyses (103 NOMAS participants; 18 replication family members), leaving us with a final sample of 556 Dominican NOMAS participants and 342 replication family members. At the variant level, we removed SNVs with a call rate <95%, monomorphic SNVs and non-exonic SNVs, leaving us with 3,418 exonic rare single nucleotide variants (RVs) in our regions for Exome Array analysis in our replication samples. Mendelian error checking was performed in the replication families and Mendelian errors were set to missing for all the variants called using PLATO v. 0.84 [17].

Statistical methods

Gene-based discovery analyses. Rare SNVs were defined based on frequencies from our NOMAS Dominican participants as previously described [15], with SNVs having an $MAF < 5\%$ being considered rare. Gene-based analyses were performed using F-SKAT, an extension of the sequence kernel association test for familial samples with binary traits [18]. Two analyses were performed based on ANNOVAR and SeattleSeq annotations; one analyzing all exonic rare variants (UTR 3', UTR 5', synonymous, missense, nonsense and splice site variants) and a second restricting analyses to nonsynonymous variants (missense, nonsense and splice site variants). All analyses were adjusted for BMI, waist-hip ratio, diabetes, hypertension, pack-years of smoking and age. Covariates were selected for being previously associated with plaque in a polygenic screen implemented in SOLAR [2]. Our analyses were additionally adjusted for the first principal component obtained via principal components analyses implemented in PC-AiR to account for population substructure [19]. PC-AiR analyses were performed using Exome Array data and were restricted to autosomal variants with call rate > 95% and $MAF \geq 5\%$. Additionally, variants in pairwise linkage disequilibrium were removed (PLINK option `indep-pairwise 100 2 0.8`). Gene-based analyses were restricted to RVs with a call rate = 100% as required by F-SKAT and to genes with at least two polymorphic RVs. A $p < 9.3 \times 10^{-5}$ was considered significant based on a Bonferroni correction of 534 tests (399 exonic RV genes and 135 nonsynonymous RV genes).

Replication analyses. Gene-based RV association analyses using Exome Array data were performed using SKAT-O v. 1.1.2 in the NOMAS sample and F-SKAT v. 1.0 in the replication family sample. These analyses were performed using the same two filtering algorithms employed in the discovery analyses; all exonic RVs and then a subset of nonsynonymous RVs only. Analyses were restricted to genes with at least two polymorphic variants. Family sample replication analyses were further restricted to RVs with call rate = 100% due to requirements of our analysis program, F-SKAT. Replication was defined as genes with $p < 0.05$. All association analyses were adjusted for the same covariates included in our discovery analyses. To account for population substructure in NOMAS, these analyses were adjusted for the first principal component obtained via principal components analyses implemented in Eigenstrat. Details of our Eigenstrat analyses and results have been previously reported [20].

Segregation analyses. To investigate rare nonsynonymous variants further, we examined discovery family pedigrees with a family-specific LOD score > 0.4 for co-segregation of variants with plaque. Within each family, we filtered to identify variants that were present in all individuals with plaque and were not present in all individuals without plaque.

Table 1. Characteristics of the 12 sequenced Dominican families.

Family	Individuals per Family	Family-Specific LOD Score				% Plaque presence	% Female	% Diabetes	% Hypertension	Age $\mu \pm$ SD	BMI $\mu \pm$ SD	Pack-years $\mu \pm$ SD	WHR $\mu \pm$ SD
		Chr7	Chr11	Chr14	Chr15								
253	13	0.38	0.05	-0.05	0.21	46.2	61.5	30.8	69.2	49.54 \pm 20.72	27.50 \pm 4.91	0.88 \pm 3.03	0.87 \pm 0.07
1917	10	-0.16	0.54	0.14	0.06	50.0	80.0	20.0	60.0	45.80 \pm 16.82	33.79 \pm 7.17	12.41 \pm 16.88	0.92 \pm 0.07
2235	15	0.06	0.04	0.45	0.08	33.3	60.0	13.3	60.0	46.80 \pm 16.90	28.62 \pm 3.85	6.04 \pm 11.04	0.92 \pm 0.12
2783	13	0.22	0.46	-0.03	-0.21	23.1	61.6	15.4	61.5	43.31 \pm 14.63	28.92 \pm 4.91	0.92 \pm 1.99	0.93 \pm 0.09
3603	12	-0.03	-0.06	0.07	0.37	25.0	66.7	8.3	25.0	45.08 \pm 20.15	26.56 \pm 5.60	4.87 \pm 11.08	0.89 \pm 0.11
3630	13	0.01	0.21	0.15	0.29	15.4	69.2	0.0	30.8	60.54 \pm 15.15	24.68 \pm 4.92	5.08 \pm 10.83	0.89 \pm 0.08
3631	17	0.14	-0.03	0.13	-0.10	29.4	70.6	11.8	35.3	40.53 \pm 15.86	31.53 \pm 6.98	4.70 \pm 11.07	0.91 \pm 0.06
4641	12	0.05	-0.006	0.28	0.09	25.0	83.3	58.3	25.0	44.58 \pm 17.71	31.75 \pm 8.70	5.73 \pm 7.00	0.96 \pm 0.08
5103	36	0.06	0.36	-0.10	-0.25	19.4	36.1	8.3	44.4	43.11 \pm 18.41	26.72 \pm 5.39	3.43 \pm 8.19	0.92 \pm 0.07
5279	19	0.30	0.06	-0.12	-0.08	52.6	42.1	15.8	42.1	54.42 \pm 17.76	25.18 \pm 4.17	4.71 \pm 14.06	0.95 \pm 0.05
5569	17	0.88	0.04	0.27	0.06	52.9	70.6	5.9	52.9	46.24 \pm 18.05	26.64 \pm 2.91	7.95 \pm 10.21	0.88 \pm 0.05
6081	28	0.19	0.78	-0.30	1.01	21.4	60.7	7.1	42.9	47.82 \pm 18.53	27.72 \pm 5.41	0.38 \pm 1.41	0.89 \pm 0.09

*Bold family-specific LOD score indicates family was included in discovery analyses for that QTL.

<https://doi.org/10.1371/journal.pone.0250799.t001>

Results

Participant and SNV characteristics

Twelve families containing a total of 205 individuals were included in our discovery analyses. Characteristics of these families are in Table 1.

Sequencing within these families identified 3,232 exonic RVs across the four QTLs, as shown in Table 2. Within the chromosome 7 QTL, our smallest sequenced region, 143 polymorphic exonic RVs were identified, of which 11 were novel (7.7%) and 16 were nonsynonymous (11.3%). In the chromosome 11 QTL, 987 exonic variants were identified with 129 being novel (13.0%) and 209 being nonsynonymous (21.2%). Within the chromosome 14 QTL, 856 variants were identified with 109 being novel (12.7%) and 159 being nonsynonymous (18.6%). In the chromosome 15 QTL, our largest sequenced region, 1,246 exonic variants were identified with 150 being novel (12.0%) and 268 being nonsynonymous (21.5%).

RV association results

Gene-based analyses were performed under two filtering algorithms; all exonic RVs and then restricted to nonsynonymous RVs. Exonic RV analyses revealed *NRXN3* to be significantly

Table 2. Polymorphic exonic rare variants identified through sequencing, stratified by function*.

Function*	QTL			
	Chr 7	Chr 11	Chr 14	Chr 15
Missense	16	209	157	263
Nonsense	0	0	1	2
Splice site	0	0	1	3
Synonymous	18	168	145	252
UTR 3' or UTR 5'	90	507	439	619
ncRNA	19	103	113	107
Total	143	987	856	1,246

*Function based on Annovar and SeattleSeq annotation.

<https://doi.org/10.1371/journal.pone.0250799.t002>

Table 3. Gene-based association results for rare variant analysis in the four linkage regions, for genes with $p < 0.05$ in the discovery families and $p < 0.05$ in the replication families or NOMAS.

Gene	Chr	MB Start	Exonic RVs						Nonsynonymous RVs					
			Discovery Families		Replication Families		NOMAS		Discovery Families		Replication Families		NOMAS	
			# SNVs	Pval	# SNVs	Pval	# SNVs	Pval	# SNVs	Pval	# SNVs	Pval	# SNVs	Pval
<i>AMPD3</i>	11	10.47	11	0.003	10	0.04	16	0.30	4	0.005	3	0.02	8	0.46
<i>NOM1</i>	7	156.74	23	0.007	7	0.003	30	0.55	6	0.03	3	0.002	16	0.85
<i>MICAL2</i>	11	12.11	15	0.009	8	0.01	18	0.74	7	0.007	6	0.008	15	0.83
<i>KCNC1</i>	11	17.75	18	0.01	8	4.0x10⁻⁴	17	0.35	0	N/A	0	N/A	1	N/A
<i>DPP6</i>	7	153.74	14	0.02	10	0.05	12	0.74	1	N/A	3	0.10	3	0.50
<i>C11orf58</i>	11	16.76	8	0.02	4	8.3x10⁻⁴	4	0.55	0	N/A	0	N/A	0	N/A
<i>ADM</i>	11	10.32	7	0.03	4	0.006	4	0.28	2	0.21	2	0.009	2	0.44
<i>LMO1</i>	11	8.24	4	0.05	2	0.02	2	0.21	0	N/A	1	N/A	1	N/A
<i>NRDE2</i>	14	90.74	10	0.07	9	0.004	18	0.26	7	0.03	6	6.5x10⁻⁴	14	0.08
<i>FAM181A</i>	14	94.39	5	0.27	6	0.01	5	0.10	2	0.04	4	0.03	4	0.25

<https://doi.org/10.1371/journal.pone.0250799.t003>

associated with plaque ($p = 6.8 \times 10^{-5}$), although this finding did not replicate in the replication families ($p = 0.82$) or NOMAS ($p = 0.56$). In addition to *NRXN3*, 34 genes showed suggestive association with plaque in the discovery families (S1 Table), of which eight genes showed evidence for replication in the replication families (*AMPD3*, *NOM1*, *MICAL2*, *KCNC1*, *DPP6*, *C11orf58*, *ADM*, *LMO1*) (Table 3). No genes showed evidence for replication in NOMAS (Table 3 and S1 Table). When analyses were restricted to nonsynonymous RVs, five genes showed suggestive association with plaque in the discovery families ($p < 0.05$) and had evidence for replication in replication families (*AMPD3*, *NOM1*, *MICAL2*, *NRDE2* and *FAM181A*) (Table 3). No genes showed evidence for replication in NOMAS (Table 3 and S1 Table).

Segregation analysis results

When examining rare nonsynonymous variants for co-segregation with plaque, we identified two missense variants within the chr11 QTL in F1917: rs61751342 and rs61760882 (Fig 1). The rs61751342 missense variant is located in DENN domain containing 2B gene, *DENND2B*, and is an eQTL for *SCUBE2* in the atrial appendage ($p = 3.5 \times 10^{-5}$, GTEx Consortium) and *AKIP1* in subcutaneous adipose tissue ($p = 2.0 \times 10^{-5}$, GTEx Consortium). The frequency of the minor allele of rs61751342 in our study was 3.2% in NOMAS and 6.4% in our sequenced families. The following frequencies were observed in other populations according to gnomAD [21]: non-Finnish European = 5.8%, African = 0.95%, Latino = 2.0% and South Asian = 2.5%. The rs61760882 missense variant is located in the ring finger protein 141 gene, *RNF141* and had a minor allele frequency of 0.2% in NOMAS and 6.4% in our sequenced families. The following frequencies of rs61760882 were observed in other populations according to gnomAD: non-Finnish European = 1.8%, African = 0.04%, Latino = 0.07% and South Asian = 0.1%.

Discussion

Using targeted re-sequencing in extended families from the Dominican Republic, we have identified a suggestive role for RVs within previously-identified linkage regions on chromosomes 7q36, 11p15, 14q32 and 15q23 in carotid plaque. Our study found *AMPD3* to be our most strongly associated gene with evidence for replication in both exonic and nonsynonymous RV analyses, followed by *MICAL2* which was our second most strongly associated gene with evidence for replication in nonsynonymous RV analyses. Notably, these genes showed association with carotid plaque in the replication families but not NOMAS which could reflect

F1917

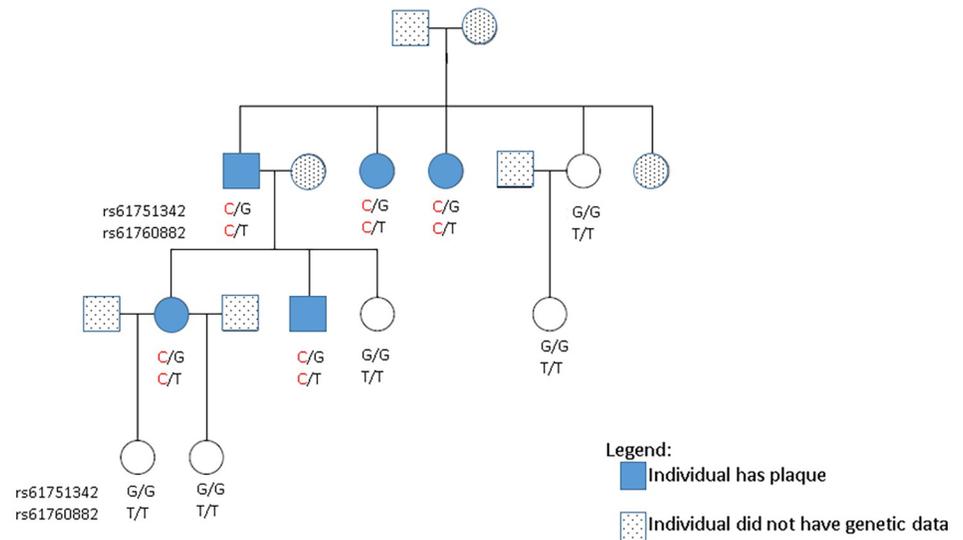


Fig 1. Genotypes for rs61751342 (*DENND2B*) and rs61760882 (*RNF141*), two missense variants found to cosegregate with plaque presence in family 1917.

<https://doi.org/10.1371/journal.pone.0250799.g001>

several things. First, the associations may be specific to the environmental and genetic context of the families. Families were selected to be at high risk for CVD whereas NOMAS participants were not selected for study inclusion based on CVD risk. Therefore, these RV associations may only be observed in this setting. Second, the association may have been specific to the RVs included in the discovery family analyses. However, in exploratory analyses restricting gene-based analyses to only those variants found in the discovery families, the NOMAS gene associations remained non-significant (S2 Table). Third, the association observed in the discovery families may be a false positive, although association was observed in the replication families lending support to the involvement of *AMPD3* and *MICAL2* in carotid plaque.

AMPD3 encodes the erythrocyte isoforms of AMP-deaminase. Two additional isoforms of AMP-deaminase are encoded by *AMPD1* and *AMPD2*, of which, the C34T variant in *AMPD1* has been implicated in CVD [22, 23]. Although the role of *AMPD3* in carotid plaque is unknown, previous studies have found common variants in *AMPD3* to be associated with HDL cholesterol [24, 25]. HDL cholesterol is inversely related to CVD risk [26] and may help prevent atherosclerosis [27, 28]. The *MICAL2* gene encodes [F-actin]-monooxygenase *MICAL2*, a regulator of the SRF signaling pathway. Published GWAS have shown common variants in *MICAL2* to be significantly associated with resting heart rate [29] and white blood cell count [30], traits which can influence atherosclerosis [31] and vascular events [32]. Further, a study investigating gene expression patterns found *MICAL2* to be downregulated in atherosclerotic CAD patients compared to controls [33]. Together, these findings in conjunction with our current study, offer biologic plausibility for *AMPD3* and *MICAL2* in atherosclerosis.

Exonic RV analyses also identified *DPP6* as a candidate gene for carotid plaque presence, a gene we previously identified in common variant analyses [9]. Therefore, the current study validates and extends this finding to suggest that RVs in *DPP6* also play an important role in carotid plaque pathogenesis. The *DPP6* gene encodes dipeptidyl aminopeptidase-like protein 6 which is a β subunit that conducts the transient-outward K^+ current $I_{(to)}$ in the human heart

[34]. Variants within *DPP6* have been implicated in familial idiopathic ventricular fibrillation [35] and Bazett's corrected QT intervals (msec) in African Americans [36], a measure which may be a marker for subclinical atherosclerosis [37]. Using these data we can hypothesize that *DPP6* is involved in severe heart disease that is likely caused by gene-related atherosclerotic mechanisms.

Further candidates were identified when examining rare nonsynonymous variants for co-segregation with plaque in individual families. Through this analysis, two missense variants on chromosome 11 were found to co-segregate with plaque. Interestingly, one of these variants, rs61751342 (located in *DENND2B*), was found to be an eQTL for *SCUBE2* in the atrial appendage. The protein encoded by *SCUBE2* is known to be expressed in vascular endothelial cells and interacts with vascular endothelial growth factor (VEGF) receptor 2 to regulate VEGF-induced angiogenesis [38]. In addition, thrombi commonly form in the atrial appendage in individuals with atrial fibrillation, mitral valve disease and other conditions [39], making this protein a potential therapeutic target for heart disease.

Our study has several limitations. First, only a subset of our extended families were sequenced; however, the selected families showed the highest evidence for linkage to each of our four QTLs. Second, our replication analyses were performed using Exome Array data, therefore, these analyses were restricted to those variants found on the Array. However, we added custom content to our Exome Array which included RVs found in our sequenced DR families but not on the commercially designed array. Fourth, our study population was from the Dominican Republic and therefore findings may not extend to other populations, particularly if the associated RVs have different allele frequencies in other populations [40]. Finally, due to our sample sizes, we had reduced power to detect RV association in the population-based NOMAS cohort, which might attribute to the lack of replication of the family findings in the NOMAS. Study strengths included our carefully defined phenotyping that was performed using standardized protocols, the use of two validation samples and our family study design which minimizes population stratification and increases our ability to investigate RVs.

In conclusion, using a unique sample of extended families from the Dominican Republic, we have extended our previous findings to implicate RVs in previously identified linkage regions in carotid plaque. Our study suggests that rare variation in genes including *AMPD3*, *MICAL2*, *DPP6*, and *DENND2B* may play a role in subclinical atherosclerosis, thereby making them promising novel candidates for the development of anti-atherosclerotic therapies. However, functional studies are warranted to elucidate the specific effects of identified variants on carotid plaque pathogenesis.

Supporting information

S1 Table. Association results for genes with $p < 0.05$ in exonic or nonsynonymous RV analyses in the discovery families.

(XLSX)

S2 Table. Association results for genes with $p < 0.05$ in exonic or nonsynonymous RV analyses in the discovery families. Replication analyses restricted to only RVs included in discovery family analyses.

(XLSX)

Acknowledgments

The authors are grateful to the families, participants and research staff who participated in the study.

Author Contributions

Conceptualization: Ralph L. Sacco, Susan H. Blanton, Tatjana Rundek.

Data curation: Liyong Wang, Chuanhui Dong.

Formal analysis: Nicole D. Dueker, Ashley Beecham.

Funding acquisition: Ralph L. Sacco, Tatjana Rundek.

Investigation: Ralph L. Sacco, Tatjana Rundek.

Resources: Ralph L. Sacco, Tatjana Rundek.

Supervision: Susan H. Blanton, Tatjana Rundek.

Visualization: Liyong Wang, Susan H. Blanton, Tatjana Rundek.

Writing – original draft: Nicole D. Dueker.

Writing – review & editing: Nicole D. Dueker, Ashley Beecham, Liyong Wang, Chuanhui Dong, Ralph L. Sacco, Susan H. Blanton, Tatjana Rundek.

References

1. Johnsen SH, Mathiesen EB. Carotid plaque compared with intima-media thickness as a predictor of coronary and cerebrovascular disease. *Curr Cardiol Rep* 2009 Jan; 11(1):21–27. <https://doi.org/10.1007/s11886-009-0004-1> PMID: 19091171
2. Dong C, Beecham A, Slifer S, Wang L, Blanton SH, Wright CB, et al. Genomewide linkage and peak-wide association analyses of carotid plaque in Caribbean Hispanics. *Stroke* 2010 Dec; 41(12):2750–2756. <https://doi.org/10.1161/STROKEAHA.110.596981> PMID: 20966410
3. Moskau S, Golla A, Grothe C, Boes M, Pohl C, Klockgether T. Heritability of carotid artery atherosclerotic lesions: an ultrasound study in 154 families. *Stroke* 2005 Jan; 36(1):5–8. <https://doi.org/10.1161/01.STR.0000149936.33498.83> PMID: 15569868
4. Hunt KJ, Duggirala R, Goring HH, Williams JT, Almasy L, Blangero J, et al. Genetic basis of variation in carotid artery plaque in the San Antonio Family Heart Study. *Stroke* 2002 Dec; 33(12):2775–2780. <https://doi.org/10.1161/01.str.0000043827.03966.ef> PMID: 12468769
5. Kuipers AL, Wojczynski MK, Barinas-Mitchell E, Minster RL, Wang L, Feitosa MF, et al. Genome-wide linkage analysis of carotid artery traits in exceptionally long-lived families. *Atherosclerosis* 2019 Dec; 291:19–26. <https://doi.org/10.1016/j.atherosclerosis.2019.10.008> PMID: 31634740
6. Franceschini N, Giambartolomei C, de Vries PS, Finan C, Bis JC, Huntley RP, et al. GWAS and colocalization analyses implicate carotid intima-media thickness and carotid plaque loci in cardiovascular outcomes. *Nat Commun* 2018 Dec 3; 9(1):5141–018-07340-5. <https://doi.org/10.1038/s41467-018-07340-5> PMID: 30510157
7. Bis JC, Kavousi M, Franceschini N, Isaacs A, Abecasis GR, Schminke U, et al. Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque. *Nat Genet* 2011 Sep 11; 43(10):940–947. <https://doi.org/10.1038/ng.920> PMID: 21909108
8. Pott J, Burkhardt R, Beutner F, Horn K, Teren A, Kirsten H, et al. Genome-wide meta-analysis identifies novel loci of plaque burden in carotid artery. *Atherosclerosis* 2017 Apr; 259:32–40. <https://doi.org/10.1016/j.atherosclerosis.2017.02.018> PMID: 28282560
9. Dong C, Beecham A, Wang L, Blanton SH, Rundek T, Sacco RL. Follow-up association study of linkage regions reveals multiple candidate genes for carotid plaque in Dominicans. *Atherosclerosis* 2012 Jul; 223(1):177–183. <https://doi.org/10.1016/j.atherosclerosis.2012.03.025> PMID: 22503546
10. Sacco RL, Sabala EA, Rundek T, Juo SH, Huang JS, DiTullio M, et al. Design of a family study among high-risk Caribbean Hispanics: the Northern Manhattan Family Study. *Ethn Dis* 2007 Spring; 17(2):351–357. PMID: 17682370
11. Sacco RL, Anand K, Lee HS, Boden-Albala B, Stabler S, Allen R, et al. Homocysteine and the risk of ischemic stroke in a triethnic cohort: the Northern Manhattan Study. *Stroke* 2004 Oct; 35(10):2263–2269. <https://doi.org/10.1161/01.STR.0000142374.33919.92> PMID: 15345803

12. Dueker ND, Beecham A, Wang L, Blanton SH, Guo S, Rundek T, et al. Rare Variants in NOD1 Associated with Carotid Bifurcation Intima-Media Thickness in Dominican Republic Families. *PLoS One* 2016 Dec 9; 11(12):e0167202. <https://doi.org/10.1371/journal.pone.0167202> PMID: 27936005
13. Rundek T, Arif H, Boden-Albala B, Elkind MS, Paik MC, Sacco RL. Carotid plaque, a subclinical precursor of vascular events: the Northern Manhattan Study. *Neurology* 2008 Apr 1; 70(14):1200–1207. <https://doi.org/10.1212/01.wnl.0000303969.63165.34> PMID: 18354078
14. Elkind MS, Sciacca R, Boden-Albala B, Rundek T, Paik MC, Sacco RL. Moderate alcohol consumption reduces risk of ischemic stroke: the Northern Manhattan Study. *Stroke* 2006 Jan; 37(1):13–19. <https://doi.org/10.1161/01.STR.0000195048.86810.5b> PMID: 16306464
15. Wang L, Beecham A, Dueker N, Blanton SH, Rundek T, Sacco RL. Sequencing of candidate genes in Dominican families implicates both rare exonic and common non-exonic variants for carotid intima-media thickness at bifurcation. *Hum Genet* 2015 Oct; 134(10):1127–1138. <https://doi.org/10.1007/s00439-015-1592-z> PMID: 26319989
16. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010 Mar 1; 26(5):589–595. <https://doi.org/10.1093/bioinformatics/btp698> PMID: 20080505
17. Grady BJ, Torstenson E, Dudek SM, Giles J, Sexton D, Ritchie MD. Finding unique filter sets in PLATO: a precursor to efficient interaction analysis in GWAS data. *Pac Symp Biocomput* 2010:315–326. PMID: 19908384
18. Yan Q, Tiwari HK, Yi N, Gao G, Zhang K, Lin WY, et al. A Sequence Kernel Association Test for Dichotomous Traits in Family Samples under a Generalized Linear Mixed Model. *Hum Hered* 2015; 79(2):60–68. <https://doi.org/10.1159/000375409> PMID: 25791389
19. Conomos MP, Miller MB, Thornton TA. Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. *Genet Epidemiol* 2015 May; 39(4):276–293. <https://doi.org/10.1002/gepi.21896> PMID: 25810074
20. Dueker ND, Guo S, Beecham A, Wang L, Blanton SH, Di Tullio MR, et al. Sequencing of Linkage Region on Chromosome 12p11 Identifies PKP2 as a Candidate Gene for Left Ventricular Mass in Dominican Families. *G3 (Bethesda)* 2018 Feb 2; 8(2):659–668. <https://doi.org/10.1534/g3.117.300358> PMID: 29288195
21. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfdi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 2020 May; 581(7809):434–443. <https://doi.org/10.1038/s41586-020-2308-7> PMID: 32461654
22. Feng AF, Liu ZH, Zhou SL, Zhao SY, Zhu YX, Wang HX. Effects of AMPD1 gene C34T polymorphism on cardiac index, blood pressure and prognosis in patients with cardiovascular diseases: a meta-analysis. *BMC Cardiovasc Disord* 2017 Jul 3; 17(1):174-017-0608-0. <https://doi.org/10.1186/s12872-017-0608-0> PMID: 28673246
23. Smolenski RT, Rybakowska I, Turyn J, Romaszko P, Zabielska M, Taegtmeyer A, et al. AMP deaminase 1 gene polymorphism and heart disease—a genetic association that highlights new treatment. *Cardiovasc Drugs Ther* 2014 Apr; 28(2):183–189. <https://doi.org/10.1007/s10557-013-6506-5> PMID: 24431031
24. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010 Aug 5; 466(7307):707–713. <https://doi.org/10.1038/nature09270> PMID: 20686565
25. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013 Nov; 45(11):1274–1283. <https://doi.org/10.1038/ng.2797> PMID: 24097068
26. Toth PP, Barter PJ, Rosenson RS, Boden WE, Chapman MJ, Cuchel M, et al. High-density lipoproteins: a consensus statement from the National Lipid Association. *J Clin Lipidol* 2013 Sep-Oct; 7(5):484–525. <https://doi.org/10.1016/j.jacl.2013.08.001> PMID: 24079290
27. Calabresi L, Gomasarshi M, Franceschini G. Endothelial protection by high-density lipoproteins: from bench to bedside. *Arterioscler Thromb Vasc Biol* 2003 Oct 1; 23(10):1724–1731. <https://doi.org/10.1161/01.ATV.0000094961.74697.54> PMID: 12969988
28. Vergeer M, Holleboom AG, Kastelein JJ, Kuivenhoven JA. The HDL hypothesis: does high-density lipoprotein protect from atherosclerosis? *J Lipid Res* 2010 Aug; 51(8):2058–2073. <https://doi.org/10.1194/jlr.R001610> PMID: 20371550
29. Eppinga RN, Hagemeyer Y, Burgess S, Hinds DA, Stefansson K, Gudbjartsson DF, et al. Identification of genomic loci associated with resting heart rate and shared genetic predictors with all-cause mortality. *Nat Genet* 2016 Dec; 48(12):1557–1563. <https://doi.org/10.1038/ng.3708> PMID: 27798624
30. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, et al. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am J Hum Genet* 2019 Jan 3; 104(1):65–75. <https://doi.org/10.1016/j.ajhg.2018.11.008> PMID: 30595370

31. Tardif Jean-Claude, Coronary artery disease in 2010, *European Heart Journal Supplements*, Volume 12, Issue suppl_C, August 2010, Pages C2–10, <https://doi.org/https%3A//doi.org/10.1093/eurheartj/suq014>
32. Lee CD, Folsom AR, Nieto FJ, Chambless LE, Shahar E, Wolfe DA. White blood cell count and incidence of coronary heart disease and ischemic stroke and mortality from cardiovascular disease in African-American and White men and women: atherosclerosis risk in communities study. *Am J Epidemiol* 2001 Oct 15; 154(8):758–764. <https://doi.org/10.1093/aje/154.8.758> PMID: 11590089
33. Abdullah MH, Othman Z, Noor HM, Arshad SS, Yusof AK, Rahman AR. Peripheral blood gene expression profile of atherosclerotic coronary artery disease in patients of different ethnicity in Malaysia. *J Cardiol* 2012 Sep; 60(3):192–203. <https://doi.org/10.1016/j.jjcc.2012.05.009> PMID: 22738689
34. Radicke S, Cotella D, Graf EM, Ravens U, Wettwer E. Expression and function of dipeptidyl-aminopeptidase-like protein 6 as a putative beta-subunit of human cardiac transient outward current encoded by Kv4.3. *J Physiol* 2005 Jun 15; 565(Pt 3):751–756. <https://doi.org/10.1113/jphysiol.2005.087312> PMID: 15890703
35. Ten Sande JN, Postema PG, Boekholdt SM, Tan HL, van der Heijden JF, de Groot NM, et al. Detailed characterization of familial idiopathic ventricular fibrillation linked to the DPP6 locus. *Heart Rhythm* 2016 Apr; 13(4):905–912. <https://doi.org/10.1016/j.hrthm.2015.12.006> PMID: 26681609
36. Jeff JM, Ritchie MD, Denny JC, Kho AN, Ramirez AH, Crosslin D, et al. Generalization of variants identified by genome-wide association studies for electrocardiographic traits in African Americans. *Ann Hum Genet* 2013 Jul; 77(4):321–332. <https://doi.org/10.1111/ahg.12023> PMID: 23534349
37. Festa A, D'Agostino R Jr, Rautaharju P, O'Leary DH, Rewers M, Mykkanen L, et al. Is QT interval a marker of subclinical atherosclerosis in nondiabetic subjects? The Insulin Resistance Atherosclerosis Study (IRAS). *Stroke* 1999 Aug; 30(8):1566–1571. <https://doi.org/10.1161/01.str.30.8.1566> PMID: 10436102
38. Lin YC, Chao TY, Yeh CT, Roffler SR, Kannagi R, Yang RB. Endothelial SCUBE2 Interacts With VEGFR2 and Regulates VEGF-Induced Angiogenesis. *Arterioscler Thromb Vasc Biol* 2017 Jan; 37(1):144–155. <https://doi.org/10.1161/ATVBAHA.116.308546> PMID: 27834687
39. Al-Saady NM, Obel OA, Camm AJ. Left atrial appendage: structure, function, and role in thromboembolism. *Heart* 1999 Nov; 82(5):547–554. <https://doi.org/10.1136/hrt.82.5.547> PMID: 10525506
40. Gravel S, Henn BM, Gutenkunst RN, Indap AR, Marth GT, Clark AG, et al. Demographic history and rare allele sharing among human populations. *Proc Natl Acad Sci U S A* 2011 Jul 19; 108(29):11983–11988. <https://doi.org/10.1073/pnas.1019276108> PMID: 21730125