

ARIC Manuscript Proposal #1947

PC Reviewed: 5/29/12
SC Reviewed: _____

Status: A
Status: _____

Priority: 2
Priority: _____

Population Architecture using Genomics and Epidemiology (PAGE)

Ver. 06/14/10

PAGE Manuscript Proposal Template

Submit proposals by email to the PAGE Coordinating Center at Purn@biology.rutgers.edu

*All sections must be completed; incomplete applications will be returned.
Do not exceed 3 pages in length (not including references).*

PAGE Ms. Number: _____ Submission Date : May 2012 [Approval Date: _____]

Title of Proposed Ms.: Genetic association of Ankle Brachial Index (ABI) and Lower Extremity Peripheral Artery Disease (PAD) in African American, European, and Hispanic individuals using the MetaboChip array: The PAGE Study

I. INVESTIGATOR INFORMATION:

Name of Lead Author: Christina L. Wassel (CALiCo/SOL) | **Junior Investigator?** Yes
Email Address: cwassel@ucsd.edu
Telephone Number: 858-822-7668

Names, affiliations and email address of PAGE Investigators proposed as co-authors:

N, N	Affiliation in PAGE	Email
Nora Franceschini	CALiCo/ARIC and SHS	noraf@unc.edu
Charles Kooperberg	WHI	clk@fhcrc.org
Cara Carty	WHI	ccarty@whi.org
Alicia Young	WHI	ayoung@whi.org
Petra Buzkova	CALiCo/CHS	buzkova@u.washington.edu
Gerardo Heiss	CALiCo/ARIC	gerardo_heiss@unc.edu
Linda Kao	CALiCo/ARIC	wkao@jhsp.edu
Nancy Jenny	CALiCo/CHS	Nancy.Jenny@uvm.edu
Steve Buyske	PAGE CC	buyske@stat.rutgers.edu
Dana Crawford	EAGLE	dana.crawford@chgr.mc.vanderbilt.edu
Ebony Bookman	NHGRI	bookmane@mail.nih.gov
Lucia Hindorff	NHGRI	hindorffl@mail.nih.gov
Matthew A. Allison	SOL	mallison@ucsd.edu

Partner studies in PAGE not collaborating in this ms. proposal:

Study	Contacted? Y/N	Declined? / Other?
MESA	N	
NOMAS	N	
eMERGE network	N	

Names, affiliations, email address of non-PAGE investigators proposed as co-authors: **None at this time, but will likely be added as we contact the studies**

II. SCIENTIFIC RATIONALE (Please be specific and concise) Peripheral artery disease (PAD) is associated with an increased risk for incident cardiovascular disease events and mortality^{1, 2}. In the Reduction of Atherothrombosis for Continued Health (REACH) registry, almost two-thirds of the individuals with PAD had concomitant clinically-evident atherosclerotic disease in the cerebrovascular or coronary artery disease (CAD) territories whereas only one-quarter of the individuals with coronary disease had clinically-evident atherosclerotic involvement of other arterial beds³. While PAD and CAD share many common risk factors, cigarette smoking and type 2 diabetes are stronger risk factors for PAD than for coronary artery disease⁴. The variable distribution of the burden of atherosclerosis across vascular beds among subjects at risk suggests that other factors exist, including possibly genetic factors, that may contribute to the predilection of atherosclerosis to develop in a given anatomic location. Currently, little is known about the genetic susceptibility to PAD but familial aggregation of PAD and heritability estimates suggest a significant genetic contribution⁵⁻¹⁰.

The ankle brachial index (ABI) is an easy and reliable diagnostic test used to detect symptomatic as well as asymptomatic PAD¹¹. A genome-wide linkage scan for ABI identified several potential candidate genes under six linkage signals in pathways of inflammation, coagulation, blood pressure regulation, and lipid metabolism⁶. A recent large genome-wide association study (GWAS) meta-analysis of European descent participants found variants in the 9p21 locus significantly associated with ABI¹². Additionally, a recent large-scale candidate gene analysis found a variant in *TCF7L2* significantly associated with ABI in the discovery phase, but it was not replicated¹³. However, overall there have been few large scale candidate gene studies conducted concentrating on the ABI and PAD phenotypes, and most of which have been limited by small sample size, lack of ethnic diversity, lack of modest number of genes examined, and lack of robust independent replication of initial findings¹⁴. In particular, very little is known regarding possible genetic variants for ABI and PAD in Hispanics.

III. OBJECTIVES AND PLAN (Please be specific and concise)

a. Study Questions/Hypotheses.

1. For loci represented in the MetaboChip array, are there *possible* novel associations with ankle brachial index and peripheral artery disease in African American and Hispanic individuals (potential new loci)? Currently very little is known about genetic variants for ABI and/or PAD in Hispanics. **We will also examine this association in European Americans so that they can be used as a comparison group.**
2. For known loci associated with atherosclerotic, subclinical cardiovascular disease (CVD) and clinical CVD traits and represented in the MetaboChip, are there one of more independent signals in African Americans and/or Hispanics (generalization and fine mapping)? **We will**

also use European results as necessary here to aid in generalization and fine-mapping efforts.

b. Study populations, study design for each

All PAGE study populations with MetaboChip data and measured ankle brachial index and/or peripheral artery disease, which includes ARIC, CHS, WHI, HCHS/SOL, and EAGLE.

c. Variant/SNPs (Specify)

We will evaluate all the variants in the MetaboChip that passed quality control (QC) from a range of allele frequencies (0.05% or more allele frequency) for our discovery effort and fine mapping of known loci. QC will be performed at the PAGE CC.

d. Phenotype(s) (Specify)

The phenotypes to be used for analyses are:

1. The ABI as a quantitative trait, i.e. the systolic blood pressure in the ankle divided by the systolic blood pressure in the arm. Those with $ABI > 1.4$ will be removed. This phenotype will be harmonized as closely as possible among all of the studies. The lead author has previously performed ABI harmonization for several of the PAGE studies (as well as possible replication studies) already as part of the Candidate Gene Association Resource (CARE) consortium.
2. PAD as a binary trait, including participants with $ABI < 0.90$, hospitalization for PAD (WHI and SOL), or previous leg revascularization. Those with $ABI > 1.4$ will be removed. If information on intermittent claudication is available in enough studies (as assessed by the Rose claudication questionnaire), and the sample size is sufficient, this may also be either added to the PAD definition as a sensitivity analysis, or examined as separate outcome.

e. Covariates (Specify)

Continuous age, sex, study center or region (if needed), and ancestry principal components. We will also adjust for local ancestry estimates if available. Additional levels of adjustment for traditional cardiovascular disease risk factors (i.e. smoking, hypertension, lipids, body mass index, diabetes) may be performed to determine whether any novel loci identified for ABI or PAD are actually working through another closely associated trait or disease (i.e. mediation).

f. Main statistical analysis methods

Analyses will be stratified by self-reported race and within each PAGE study. We will use linear regression models for ABI and logistic regression models for the PAD. We will use additive genetic models adjusted for age, sex (except WHI), study center or region (if needed), and ancestry principal components (PCs). PCs will be centrally generated at the PAGE CC to ensure consistency across analyses and manuscripts. Adjustment for local ancestry estimates will take place as needed, and if available in time. The number of PCs adjusted for will be guided via QQ plots, and also strength of the associations of the PCs with the kidney traits of interest. It is difficult to determine beforehand, without the actual data, how many PCs we will need to adjust for, exactly what type of analyses may need to be conducted, and how this may or may not differ among the

ethnic Hispanic groups (i.e. Dominican, Mexican, Cuban). Additional levels of adjustment may be carried out as described above in the “Covariates” section. Those with ABI > 1.4 will be removed from both continuous ABI and binary PAD analyses, as this generally indicates arterial stiffening or medial calcification, and it cannot be effectively determined with ABI only whether these participants have atherosclerosis in the lower extremities. Additionally, WHI will not contribute to ABI analyses, as they do not have ABI but rather PAD determined from hospital records. Because individuals were recruited using a probability sampling in the SOL, all analyses of this study will use weights to account for identification and selection of community areas and random selection of households within those areas¹⁵. Study-specific association results will be combined across race-specific samples using a fixed effects inverse variance meta-analysis approach implemented in METAL¹⁶ or PLINK¹⁷.

For aim 1. Possible Discovery. We will perform single variant test for all variants with minor allele frequency (MAF) >1% and use Bonferroni correction to declare a significant p-value threshold based on multiple testing. Although we acknowledge that power may be suboptimal in certain groups for this aim (i.e. not all studies have ABI (WHI)), there is still potential to discover new loci in groups such as Hispanics, which are very understudied with regard to genetic variants for ABI and PAD. For example of result reporting see Example Table 1.

For aim 2. Generalization and fine mapping. We will perform comprehensive analyses of all known loci for atherosclerotic, subclinical CVD and clinical CVD traits represented in the MetaboChip. We will use single variant test approach for variants with MAF>1% and adjust for the number of independent tests using a Monte Carlo approach that accounts for linkage disequilibrium (LD) between variants across the tested loci¹⁸. We will perform conditional analyses using the most significant variant as covariate if there is suggestive evidence for secondary independent signals in these regions. For low frequency variants (0.1%<MAF<5%), we will explore novel methods for collapsing variants across regions including a recent method developed by Lin et al¹⁹, as well as rare variants burden tests of Madsen and Browning²⁰, Li and Leal²¹, Liu and Leal²², and Pan and Shen²³. In order to adequately evaluate SNPs within the 0.1-5% MAF range for PAD as a binary trait, we will use the methods of Li et al²⁴, which are appropriate for analysis of rare variants from GWAS and imputed GWAS data. Briefly, these methods include a weighted haplotype-based test, and an imputation-based test. These methods do not rely on the availability of external sequence data, but it can be incorporated, and also improve the power over traditional rare variant methods that are more appropriate for sequence data. For novel variants identified in known loci, we will further use bioinformatics tools for functional characterization of the variants. For example of result reporting see Example Table 2 and Example Figures.

We are currently working on possible replication groups for this study, including the Multi-Ethnic Study of Atherosclerosis (MESA) for African American and Hispanic individuals, possibly the Northern Manhattan Study (NOMAS) for Hispanics, as well as the eMERGE network, and will be contacting these studies in the near future.

g. Ancestry information used? No ___ Yes__X_ How is it used in the analyses?

We will use estimates of global ancestry (principal components) to adjust for population stratification. Because global ancestry may not fully account for population stratification in African Americans and Hispanics, we will also adjust for local ancestry, as available.

h. Anticipated date of draft manuscript to P&P: 6-8 months after the data are available. Analyses should be done in 1-2 months and remaining time is for replication of main findings in collaborating studies and for manuscript writing.

- i. What manuscript proposals listed on www.pagestudy.org/index.php/manuscripts/ are most related to the work proposed here? Approved PAGE ms. numbers: _
 – If any: Have the lead authors of these proposals been contacted for comments and/or collaboration? Yes X No ___

IV. SOURCE OF DATA TO BE USED (Provide rationale for any data whose relevance to this manuscript is not obvious): **Check all that apply:**

Aggregate/summary data to be generated by investigators of the study(ies) mentioned:

☒ EAGLE; ☒ CALiCO; ☐ MEC; ☒ WHI; ☐ CC; ☐ Other: _____
 If CALiCo, specify ☒ ARIC; ☐ CARDIA; ☒ CHS; ☐ SHS-Fam; ☐ SHS-Cohort; ☒ SOL

Table of ABI/PAD availability in PAGE

	ABI	PAD
ARIC	Yes	Yes
CHS	Yes	Yes
CARDIA	No	No
SHS	No metabochip data	No metabochip data
EAGLE	Yes	Yes
MEC	No	No
WHI	No	Yes
SOL	Yes	Yes

NOTE: As indicated above, our MAIN definition of PAD will be ABI<0.90, hospitalization for PAD (WHI and SOL), or previous leg revascularization. We will perform sensitivity analysis to determine whether we obtain similar results when using hospital records, i.e. excluding WHI.

I, **Christina Wassel**, affirm that this proposal has been reviewed and approved by all listed investigators.

V. REFERENCES

1. Ankle Brachial Index Collaboration, Fowkes FG, Murray GD, Butcher I, Heald CL, Lee RJ, Chambless LE, Folsom AR, Hirsch AT, Dramaix M, deBacker G, Wautrecht JC, Kornitzer M, Newman AB, Cushman M, Sutton-Tyrrell K, Fowkes FG, Lee AJ, Price JF, d'Agostino RB, Murabito JM, Norman PE, Jamrozik K, Curb JD, Masaki KH, Rodriguez BL, Dekker JM, Bouter LM, Heine RJ, Nijpels G, Stehouwer CD, Ferrucci L, McDermott MM, Stoffers HE, Hooi JD, Knottnerus JA, Ogren M, Hedblad B, Witteman JC, Breteler MM, Hunink MG, Hofman A, Criqui MH, Langer RD, Fronek A, Hiatt WR, Hamman R, Resnick HE, Guralnik J, McDermott MM. Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis. *JAMA*. 2008; 300: 197-208.
2. Steg PG, Bhatt DL, Wilson PW, D'Agostino RS, Ohman EM, Rother J, Liao CS, Hirsch AT, Mas JL, Ikeda Y, Pencina MJ, Goto S, REACH Registry Investigators. One-year cardiovascular event rates in outpatients with atherothrombosis. *JAMA*. 2007; 297: 1197-1206.
3. Bhatt DL, Steg PG, Ohman EM, Hirsch AT, Ikeda Y, Mas JL, Goto S, Liao CS, Richard AJ, Rother J, Wilson PW, REACH Registry Investigators. International prevalence, recognition, and

treatment of cardiovascular risk factors in outpatients with atherothrombosis. *JAMA*. 2006; 295: 180-189.

4. Hirsch AT, Haskal ZJ, Hertzner NR, Bakal CW, Creager MA, Halperin JL, Hiratzka LF, Murphy WR, Olin JW, Puschett JB, Rosenfield KA, Sacks D, Stanley JC, Taylor LM, Jr, White CJ, White J, White RA, Antman EM, Smith SC, Jr, Adams CD, Anderson JL, Faxon DP, Fuster V, Gibbons RJ, Hunt SA, Jacobs AK, Nishimura R, Ornato JP, Page RL, Riegel B, American Association for Vascular Surgery, Society for Vascular Surgery, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, Society of Interventional Radiology, ACC/AHA Task Force on Practice Guidelines Writing Committee to Develop Guidelines for the Management of Patients With Peripheral Arterial Disease, American Association of Cardiovascular and Pulmonary Rehabilitation, National Heart, Lung, and Blood Institute, Society for Vascular Nursing, TransAtlantic Inter-Society Consensus, Vascular Disease Foundation. ACC/AHA 2005 Practice Guidelines for the management of patients with peripheral arterial disease (lower extremity, renal, mesenteric, and abdominal aortic): a collaborative report from the American Association for Vascular Surgery/Society for Vascular Surgery, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, Society of Interventional Radiology, and the ACC/AHA Task Force on Practice Guidelines (Writing Committee to Develop Guidelines for the Management of Patients With Peripheral Arterial Disease): endorsed by the American Association of Cardiovascular and Pulmonary Rehabilitation; National Heart, Lung, and Blood Institute; Society for Vascular Nursing; TransAtlantic Inter-Society Consensus; and Vascular Disease Foundation. *Circulation*. 2006; 113: e463-654.

5. Carmelli D, Fabsitz RR, Swan GE, Reed T, Miller B, Wolf PA. Contribution of genetic and environmental influences to ankle-brachial blood pressure index in the NHLBI Twin Study. National Heart, Lung, and Blood Institute. *Am J Epidemiol*. 2000; 151: 452-458.

6. Kullo IJ, Turner ST, Kardina SL, Mosley TH, Jr, Boerwinkle E, de Andrade M. A genome-wide linkage scan for ankle-brachial index in African American and non-Hispanic white subjects participating in the GENOA study. *Atherosclerosis*. 2006; 187: 433-438.

7. Murabito JM, Guo CY, Fox CS, D'Agostino RB. Heritability of the ankle-brachial index: the Framingham Offspring study. *Am J Epidemiol*. 2006; 164: 963-968.

8. Valentine RJ, Guerra R, Stephan P, Scoggins E, Clagett GP, Cohen J. Family history is a major determinant of subclinical peripheral arterial disease in young adults. *J Vasc Surg*. 2004; 39: 351-356.

9. Valentine RJ, Verstraete R, Clagett GP, Cohen JC. Premature cardiovascular disease is common in relatives of patients with premature peripheral atherosclerosis. *Arch Intern Med*. 2000; 160: 1343-1348.

10. Wassel CL, Loomba R, Ix JH, Allison MA, Denenberg JO, Criqui MH. Family history of peripheral artery disease is associated with prevalence and severity of peripheral artery disease: the San Diego population study. *J Am Coll Cardiol*. 2011; 58: 1386-1392.

11. Newman AB, Shemanski L, Manolio TA, Cushman M, Mittelmark M, Polak JF, Powe NR, Siscovick D. Ankle-arm index as a predictor of cardiovascular disease and mortality in the Cardiovascular Health Study. The Cardiovascular Health Study Group. *Arterioscler Thromb Vasc Biol*. 1999; 19: 538-545.

12. Murabito JM, White CC, Kavousi M, Sun YY, Feitosa MF, Nambi V, Lamina C, Schillert A, Coassin S, Bis JC, Broer L, Crawford DC, Franceschini N, Frikke-Schmidt R, Haun M, Hollewijn S, Huffman JE, Hwang SJ, Kiechl S, Kollerits B, Montasser ME, Nolte IM, Rudock ME, Senft A, Teumer A, van der Harst P, Vitart V, Waite LL, Wood AR, Wassel CL, Absher DM, Allison MA, Amin N, Arnold A, Asselbergs FW, Aulchenko Y, Bandinelli S, Barbalic M, Boban M, Brown-Gentry K, Couper DJ, Criqui MH, Dehghan A, den Heijer M, Dieplinger B, Ding J, Dorr M, Espinola-Klein C, Felix SB, Ferrucci L, Folsom AR, Fraedrich G, Gibson Q, Goodloe R, Gunjaca G, Haltmayer M, Heiss G, Hofman A, Kieback A, Kiemeny LA, Kolcic I, Kullo IJ, Kritchevsky SB, Lackner KJ, Li X, Lieb W, Lohman K, Meisinger C, Melzer D, Mohler ER,3rd, Mudnic I, Mueller T, Navis G, Oberhollenzer F, Olin JW, O'Connell J, O'Donnell CJ, Palmas W, Penninx BW, Petersmann A, Polasek O, Psaty BM, Rantner B, Rice K, Rivadeneira F, Rotter JI, Seldenrijk A, Stadler M, Summerer M, Tanaka T, Tybjaerg-Hansen A, Uitterlinden AG, van Gilst WH, Vermeulen SH, Wild SH, Wild PS, Willeit J, Zeller T, Zemunik T, Zgaga L, Assimes TL, Blankenberg S, Boerwinkle E, Campbell H, Cooke JP, de Graaf J, Herrington D, Kardina SL, Mitchell BD, Murray A, Munzel T, Newman A, Oostra BA, Rudan I, Shuldiner AR, Snieder H, van Duijn CM, Volker U, Wright AF, Wichmann HE, Wilson JF, Witteman JC, Liu Y, Hayward C, Borecki IB, Ziegler A, North KE, Cupples LA, Kronenberg F. Association Between Chromosome 9p21 Variants and the Ankle-Brachial Index Identified by a Meta-Analysis of 21 Genome-Wide Association Studies. *Circ Cardiovasc Genet*. 2011; .
13. Wassel CL, Lamina C, Nambi V, Coassin S, Mukamal KJ, Ganesh SK, Jacobs DR,Jr, Franceschini N, Papanicolaou GJ, Gibson Q, Yanek LR, van der Harst P, Ferguson JF, Crawford DC, Waite LL, Allison MA, Criqui MH, McDermott MM, Mehra R, Cupples LA, Hwang SJ, Redline S, Kaplan RC, Heiss G, Rotter JI, Boerwinkle E, Taylor HA, Eraso LH, Haun M, Li M, Meisinger C, O'Connell JR, Shuldiner AR, Tybjaerg-Hansen A, Frikke-Schmidt R, Kollerits B, Rantner B, Dieplinger B, Stadler M, Mueller T, Haltmayer M, Klein-Weigel P, Summerer M, Wichmann HE, Asselbergs FW, Navis G, Leach IM, Brown-Gentry K, Goodloe R, Assimes TL, Becker DM, Cooke JP, Absher DM, Olin JW, Mitchell BD, Reilly MP, Mohler ER,3rd, North KE, Reiner AP, Kronenberg F, Murabito JM. Genetic determinants of the ankle-brachial index: A meta-analysis of a cardiovascular candidate gene 50K SNP panel in the candidate gene association resource (CARE) consortium. *Atherosclerosis*. 2012; .
14. Knowles JW, Assimes TL, Li J, Quertermous T, Cooke JP. Genetic susceptibility to peripheral arterial disease: a dark corner in vascular biology. *Arterioscler Thromb Vasc Biol*. 2007; 27: 2068-2078.
15. Lavange LM, Kalsbeek WD, Sorlie PD, Aviles-Santa LM, Kaplan RC, Barnhart J, Liu K, Giachello A, Lee DJ, Ryan J, Criqui MH, Elder JP. Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol*. 2010; 20: 642-649.
16. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010; 26: 2190-2191.
17. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81: 559-575.
18. Lin DY. An efficient Monte Carlo approach to assessing statistical significance in genomic studies. *Bioinformatics*. 2005; 21: 781-787.

19. Lin DY, Tang ZZ. A general framework for detecting disease associations with rare variants in sequencing studies. *Am J Hum Genet.* 2011; 89: 354-367.
20. Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. *PLoS Genet.* 2009; 5: e1000384.
21. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet.* 2008; 83: 311-321.
22. Liu DJ, Leal SM. A novel adaptive method for the analysis of next-generation sequencing data to detect complex trait associations with rare variants due to gene main effects and interactions. *PLoS Genet.* 2010; 6: e1001156.
23. Pan W, Shen X. Adaptive tests for association analysis of rare variants. *Genet Epidemiol.* 2011; 35: 381-388.
24. Li Y, Byrnes AE, Li M. To identify associations with rare variants, just WHaIT: Weighted haplotype and imputation-based tests. *Am J Hum Genet.* 2010; 87: 728-735.

Appendix. Example of PAGE Metabochip fine-mapping figures and tables using European and African American data (for SOL P&P)*

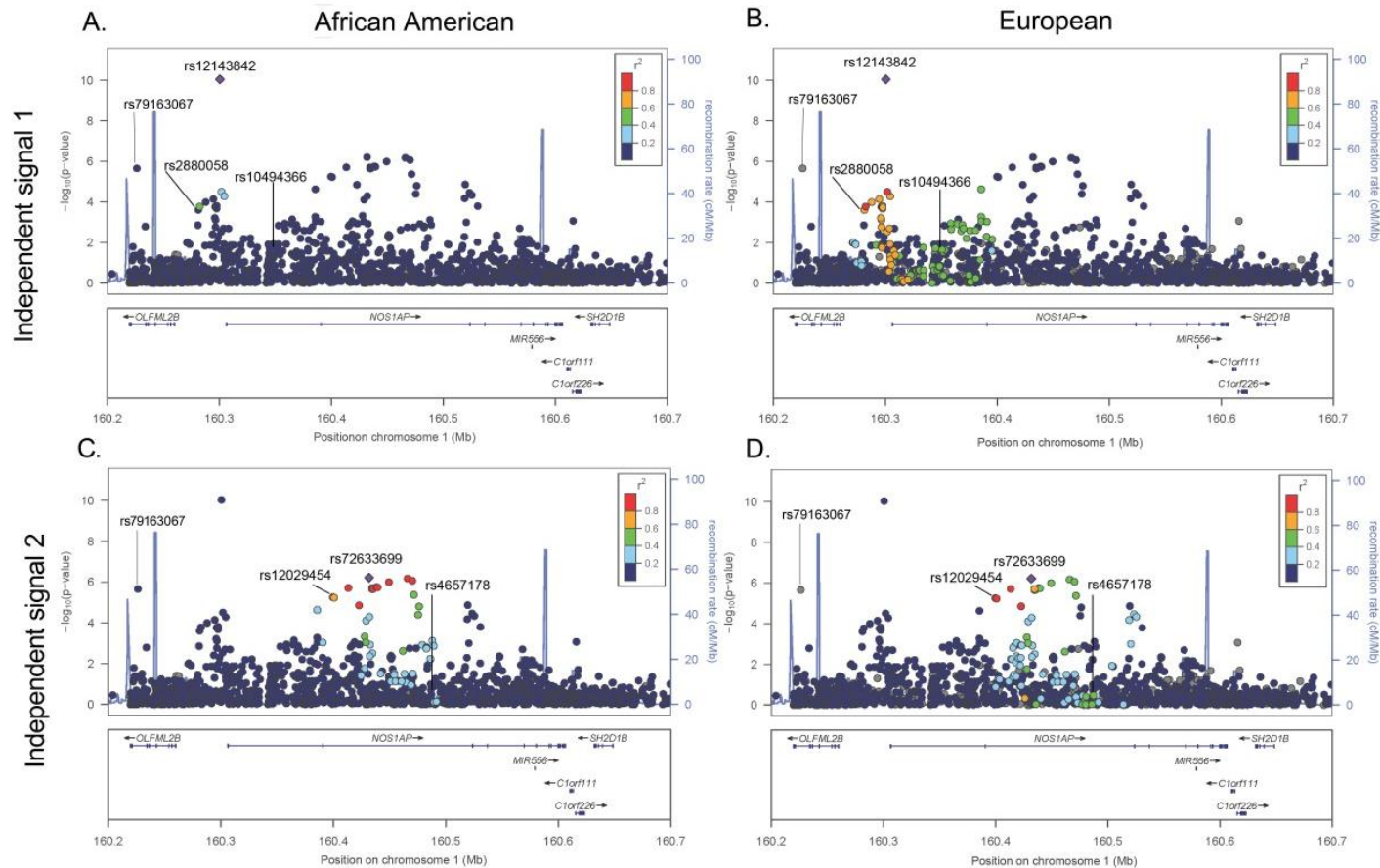


FIGURE 1. $-\log P$ plot for common SNPs at the gene independent signal 1 and 2 loci. P -values are estimated in African Americans are plotted using linkage disequilibrium estimates from African Americans (panels A and C) and Europeans (panels B and D). SNPs are represented by *circles*, lines indicate index SNPS previously identified in GWA studies of European and Indian Asian populations, and the *large blue diamond* is the best marker in African Americans. Circle color represents correlation with the best marker in African Americans: *blue* indicates weak correlation and *red* indicates strong correlation. Recombination rate is plotted in the background and annotated genes are shown at the bottom of the plot. ***Similar figures will be generated for each ethnic group included in the proposal**

Example of TABLE 1. Novel associations with ABI and PAD in SOL participants.*

SNP	Locus	Chr	Position (Build 36)	Coded/other allele	Coded allele frequency	Beta (se) or OR (se)	<i>P</i> -value
Ankle Brachial Index							
rsXXX							
Peripheral Artery Disease							

(we will considered showing this results by country of origin)

***Similar results tables will be generated for each ethnic group included in the proposal**

Example of Descriptive TABLE. Demographic characteristics of SOL and WHI Hispanic participants *

Characteristic	SOL	<u>WHI PAGE</u>	
		Wave 1	Wave 2
N			
Age, years, mean (SD)			
Sex, female, N (%)			
QT duration, ms, mean (SD)			
Heart rate, bpm, mean (SD)			

WHI, Women's Health Initiative; Study of Latinos.

***Similar results tables will be generated for each ethnic group included in the proposal**