ARIC Manuscript Proposal # 1308

PC Reviewed: _11/13_/07	**Status: _A___**	**Priority: __2_**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: *NOS1AP* Association with Type 2 Diabetes

b. Abbreviated Title (Length 26 characters): *NOS1AP* and diabetes

2. Writing Group:
Writing group members:

Audrey Chu
Linda Kao
Josef Coresh
Jim Pankow
Eric Boerwinkle
Wendy Post
Dan Arking
Aravinda Chakravarti
Peter Spooner
Gordon Tomaselli
Others

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. __AC___ **[please confirm with your initials electronically or in writing]**

First author: Audrey Chu
Address: 2024 E. Monument Street
Suite 1-500K
Baltimore, MD 21205

Phone: 310-387-5558 Fax:
E-mail: achu@jhsph.edu

Corresponding/senior author (if different from first author correspondence will be sent to both the first author & the corresponding author):
Address: Linda Kao
Department of Epidemiology
Johns Hopkins School of Public Health
615 N. Wolfe Street
Rm. W6513
Baltimore, MD 21205

Phone: 410-614-0945 Fax: 410-955-0863
E-mail: wkao@jhsph.edu

3. Timeline:

Will begin analysis when proposal is approved.

4. Rationale:

Type 2 diabetes (T2D) is currently estimated to affect more than 170 million individuals worldwide, with estimates increasing to 320 million in 2030 (1). Discovering the genes underlying the development of T2D would provide better understanding of the molecular etiology of T2D and guide development of treatment and prevention.

Initial identification of the *NOS1AP* gene as a candidate for T2D was facilitated by replication of a strong linkage signal from the chromosome 1q21-q25 region from several studies (2, 3). Subsequent fine mapping studies in 7 populations confirmed that variants in *NOS1AP* are associated with T2D in the 1q Consortium populations (U.K. whites, French whites, Shanghai Chinese, Hong Kong Chinese, Utah whites, Old Order Amish, Pima Indians and Arkansas whites) (4-10). *NOS1AP* lies directly under the peak at 1q23 and is a strong biologic candidate, with function related to nitric oxide (NO) synthesis and release (11); studies of mice with functional mutations in proteins essential for NO synthesis present with insulin resistance (12, 13). Experimental evidence of the effect of NO on glucose metabolism in humans is not available yet, but results from observational studies show lowered NO availability in those with T2D (14).

We propose to perform a cross-sectional analysis of prevalent T2D with *NOS1AP* variants in the ARIC study, as well as a cohort analysis of incident T2DM. 21 SNPs have already been genotyped in the 5' region of the *NOS1AP* gene.

5. Main Hypothesis/Study Questions:

This manuscript will address the following questions:

Are any *NOS1AP* variants associated with prevalent or incident T2D in both black and white ARIC participants?

Are any *NOS1AP* variants associated with glucose, insulin levels (both from visit 1), and insulin resistance in both black and white non-diabetic ARIC participants?

6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

Study design: Prospective follow-up from visit 1 (1987-89) to visit 4 (1996-98) of ARIC participants meeting inclusion criteria. 21 *NOS1AP* variants have been genotyped in all black and white ARIC participants who agreed to future genetic research.

Exclusion criteria:

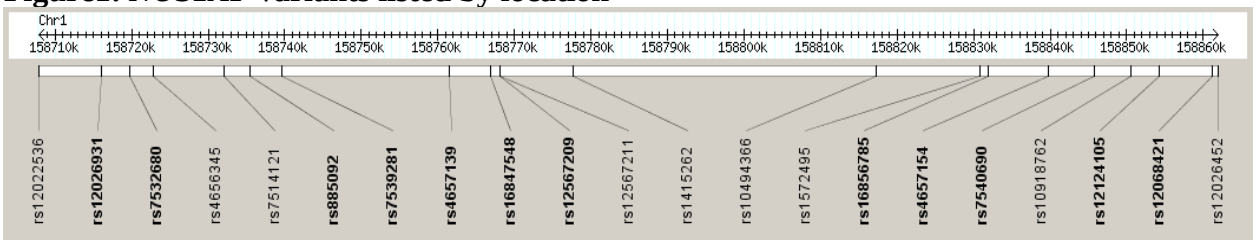
1) No consent to genetic research
2) Non-white or non-black
3) Missing genotype for the SNP being analyzed
4) Missing diabetes status at baseline
5) Missing glucose or insulin
6) For secondary outcome analysis (fasting glucose, fasting insulin, insulin resistance):
   a) Did not fast for at least 8 hours
   b) Self-report of T2D medication

Outcome: Analysis of the primary outcome, T2D, will be defined by any of the following conditions: 1) fasting plasma glucose (FPG) >126 mg/dL, 2) casual plasma glucose >200 mg/dL, 3) self-report of physician diagnosed diabetes, 4) self-report of diabetes medication. Prevalent T2D will include all participants with diabetes diagnosis up to visit 1; incident T2D will include all participants who are free of diabetes at visit 1, but subsequently developed diabetes in visits 2-4.

For secondary analysis, we will investigate FPG from visit 1 and fasting insulin from visit 1, in separate models; insulin will be log-transformed due to non-normality. In addition, we will calculate HOMA-IR (a measure of insulin resistance) from FPG and fasting insulin from visit 1 (HOMA-IR = [FPG*Fasting Insulin]/22.5).

Exposure: The *NOS1AP* variants included in this aim are rs12026931, rs12567209, rs10494366, rs10918762, rs12022536, rs12026452, rs12068421, rs12124105, rs12567211, rs1415262, rs1572495, rs16847548, rs16856785, rs4656345, rs4657139, rs4657154, rs7514121, rs7532680, rs7539281, rs7540690 and rs885092, listed in location order in figure 1.

Figure1: *NOS1AP* variants listed by location



These SNPs were selected based on coverage for *NOS1AP* in the block containing rs10494366 and rs4657139 (significant SNPs from a study of cardiac repolarization in whites where *NOS1AP* was first identified (15)) and neighboring LD blocks in a population of U.S. whites with ancestry from northern and western Europe (CEU) in

HapMap. Twenty-one SNPs were selected using Tagger with criteria of R-square>0.65 and minor allele frequency (MAF) >0.05 in the CEU population.

Data analysis:
Note: All analyses will be stratified by race.
Data checking and cleaning: Hardy-Weinburg equilibrium (HWE) will be checked within each race; we will exclude SNPs with $p<0.001$. We will examine genotype distribution for each SNP by inclusion/exclusion criteria, as well as T2D distribution by those with missing genotype.

Analysis of prevalent cases: Logistic regression will be used to examine the association between each SNP and T2D diagnosis in a cross-sectional design.

Analysis of incident cases: Proportional hazards Cox regression will be used to investigate incident T2D. Follow up time will be determined by 1) the interpolated date for those who develop diabetes, or 2) last study visit for those who simply did not return to the study or administratively censored. We are able to estimate diabetes incidence date by using linear interpolation of glucose values at the visit which T2D was diagnosed and the preceding visit, as described in Duncan et al, (16). Again, each SNP will be modeled additively for genotypic association analysis and coded 0, 1 or 2 dependent on the number of risk alleles present in the genotype.

SNP modeling: For both prevalent and incident T2D analyses, each SNP will be modeled additively for genotypic association analysis and coded 0, 1 or 2 dependent on the number of risk alleles present in the genotype.

Haplotype modeling: Haplotype association analysis will be carried out by a 2-4 SNP windowing approach for all SNPs for prevalent T2D. We will implement the Haplostats package created by Schaid et al. (17, 18) for this analysis. Haplostats estimates the probabilities of possible haplotypes for an individual and uses the probabilities as weights to iteratively update coefficients from the regression and probabilities for haplotypes while simultaneously adjusting for covariates and interactions of interest.

However, Haplostats does not currently feature an option for use in survival analysis, so we will perform two sets of analyses: 1) treat the prospective portion as a case-control and use logistic regression as implemented in Haplo.glm and 2) apply a two-step method where haplotype are estimated from PHASE, a Baysian-based haplotype method by Stephens et al (19, 20) which will then be used as exposures in a Cox regression. We will select best estimated haplotypes.

In Haplostats and PHASE, the same method to code genotype will be used to code haplotypes (0: no at-risk haplotype, 1: one copy of the at-risk haplotype, 2: two copies of the at-risk haplotype).

Multivariable modeling: The crude model for both prevalent and incident T2D analyses will include age and sex. In addition, tiers of multivariate models will be constructed

consisting of the crude model in addition to family history of diabetes, factors related to obesity (BMI, waist circumference, physical activity level, energy intake) and education. We considered Bonferroni correction to adjust for multiple comparisons and to determine significance. However, Bonferroni correction will be overly conservative because many of the SNPs are in LD; therefore, we will still consider SNPs with uncorrected p-values <0.05 as significant.

Analysis of pooled prevalent and incident cases: If the association between genotypes and diabetes are similar in both the cross-sectional and the prospective analyses, then a pooled analysis will be performed using logistic regression with prevalent T2D cases defined at visit 4. Although the T2D cases in visit 1 are prevalent, this combined analysis is appropriate since the genetic component should precede the disease; logistic regression is also the most appropriate way to combine prevalent and incident cases. Genotype and haplotype analyses will be performed as described in previous sections; we will use Haplo.glm from the Haplostats package to estimate haplotypes for this analysis. We will construct models as previously described.

Analysis of secondary outcomes: We will perform linear regression with FPG, log-transformed fasting insulin and HOMA-IR as separate outcomes; we will again perform genotypic and haplotypic associations. We will be using Haplostats for haplotype analyses, as linear regression is available in the Haplostats package. In multivariable analysis, we will be adjusting for age, sex, family history of diabetes, obesity and factors relating to obesity (BMI, waist circumference, physical activity and energy intake) as well as education.

Population stratification: Results from our analyses in blacks will be compared to those adjusting for within-population stratification using 1,536 SNPs informative for ancestry. If substructure exists, we will examine the association between SNPs and outcomes either stratified or adjusted for membership into each subpopulation.

The ARIC study is the ideal cohort in which to perform this analysis; DNA has already been isolated, prevalent and incident T2D has been assessed. However, a limitation of this analysis is that we do not currently know the causal *NOS1AP* SNP which influences risk of T2DM. However, the SNPs which are associated in this analysis will either be the causal SNP or in high linkage disequilibrium with the causal SNP. In addition, these SNPs were selected based on their association with another phenotype, QT interval. The SNPs do not "cover" the entire gene, but are primarily located in the 5' and first intron region. The region potentially associated with T2D may not be in this region. We are considering genotyping the 1q Diabetes Consortium SNP, rs7548169, in the ARIC participants.

Strengths of this aim include the large number of T2DM cases in whites and blacks in ARIC and well-phenotyped participants in ARIC.

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ Yes __x__ No

b. If Yes, is the author aware that the file ICTDER02 must be used to exclude persons with a value RES_OTH = "CVD Research" for non-DNA analysis, and for DNA analysis RES_DNA = "CVD Research" would be used? ____ Yes ____ No
(This file ICTDER02 has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? __x__ Yes ____ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER02 must be used to exclude those with value RES_DNA = "No use/storage DNA"?
__x__ Yes ____ No

9.The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status. ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/ARIC/search.php

___x___ Yes _____ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?

Manuscript Proposal #1141: (Yan) TCF7L2 and diabetes mellitus
Manuscript Proposal #1161: (Morrison) KIF6 and diabetes mellitus
Manuscript Proposal #1152: (Post) Genomic predictors of sudden death

11. a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? __x__ Yes ____ No

11.b. If yes, is the proposal
_x__ A. primarily the result of an ancillary study (list number* __1152__)
___ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* _____ _____ _____)

*ancillary studies are listed by number at http://www.cscc.unc.edu/aric/forms/

12. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

References:

1. **Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 2004;27:1047-53.**
2. **Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ. A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. Diabetes 1999;48:1175-82.**
3. **Hanson RL, Ehm MG, Pettitt DJ, et al. An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. Am J Hum Genet 1998;63:1130-8.**
4. **Langefeld CD, Wagenknecht LE, Rotter JI, et al. Linkage of the metabolic syndrome to 1q23-q31 in Hispanic families: the Insulin Resistance Atherosclerosis Study Family Study. Diabetes 2004;53:1170-4.**
5. **Ng MC, So WY, Lam VK, et al. Genome-wide scan for metabolic syndrome and related quantitative traits in Hong Kong Chinese and confirmation of a susceptibility locus on chromosome 1q21-q25. Diabetes 2004;53:2676-83.**
6. **Vionnet N, Hani EH, Dupont S, et al. Genomewide search for type 2 diabetes-susceptibility genes in French whites: evidence for a novel susceptibility locus for early-onset diabetes on chromosome 3q27-qter and independent replication of a type 2-diabetes locus on chromosome 1q21-q24. Am J Hum Genet 2000;67:1470-80.**
7. **Wiltshire S, Hattersley AT, Hitman GA, et al. A genomewide scan for loci predisposing to type 2 diabetes in a U.K. population (the Diabetes UK Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q. Am J Hum Genet 2001;69:553-69.**
8. **Das SK, Hasstedt SJ, Zhang Z, Elbein SC. Linkage and association mapping of a chromosome 1q21-q24 type 2 diabetes susceptibility locus in northern European Caucasians. Diabetes 2004;53:492-9.**
9. **Hsueh WC, St Jean PL, Mitchell BD, et al. Genome-wide and fine-mapping linkage studies of type 2 diabetes and glucose traits in the Old Order Amish: evidence for a new diabetes locus on chromosome 14q11 and confirmation of a locus on chromosome 1q21-q24. Diabetes 2003;52:550-7.**
10. **Zeggini E. Replicated Association between Type 2 Diabetes and Variants in CAPON Revealed by High Density Linkage Disequilibrium Mapping on Chromosome 1q. American Diabetes Association, 2006.**
11. **Jaffrey SR, Snowman AM, Eliasson MJ, Cohen NA, Snyder SH. CAPON: a protein associated with neuronal nitric oxide synthase that regulates its interactions with PSD95. Neuron 1998;20:115-24.**
12. **Duplain H, Burcelin R, Sartori C, et al. Insulin resistance, hyperlipidemia, and hypertension in mice lacking endothelial nitric oxide synthase. Circulation 2001;104:342-5.**
13. **Shankar RR, Wu Y, Shen HQ, Zhu JS, Baron AD. Mice with gene disruption of both endothelial and neuronal nitric oxide synthase exhibit insulin resistance. Diabetes 2000;49:684-7.**
14. **Kashyap SR, Roman LJ, Lamont J, et al. Insulin resistance is associated with impaired nitric oxide synthase activity in skeletal muscle of type 2 diabetic subjects. J Clin Endocrinol Metab 2005;90:1100-5.**
15. **Arking DE, Pfeufer A, Post W, et al. A common genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization. Nat Genet 2006;38:644-51.**
16. **Duncan BB, Schmidt MI, Pankow JS, et al. Low-grade systemic inflammation and the development of type 2 diabetes: the atherosclerosis risk in communities study. Diabetes 2003;52:1799-805.**
17. **Lake SL, Lyon H, Tantisira K, et al. Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. Hum Hered 2003;55:56-65.**
18. **Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002;70:425-34.**
19. **Stephens M, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 2003;73:1162-9.**

20. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 2001;68:978-89.