

ARIC Manuscript Proposal #2989

PC Reviewed: 06/06/17
SC Reviewed: _____

Status: _____
Status: _____

Priority: 2
Priority: _____

1.a. Full Title: SNP-Heritability analysis of nontraditional markers of hyperglycemia in the Atherosclerosis Risk in Communities Study

b. Abbreviated Title (Length 26 characters): Heritability of glucose markers

2. Writing Group:

Writing group members: Stephanie Loomis, Priya Duggal, Liz Selvin, Adrienne Tin, Anna Kottgen, Joe Coresh, Eric Boerwinkle, James Pankow, Xiaoming Liu; others welcome

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

First author: **Stephanie Loomis**
Address: Johns Hopkins University
615 N. Wolfe St, W6021
Baltimore, MD 21287

Phone: 860 912-7169 Fax:
E-mail: sloomis5@jhu.edu

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).

Name: **Elizabeth Selvin**
Address: Johns Hopkins University
2024 E. Monument Street, Suite 2-600
Baltimore, MD 21287

Phone: 410-955-0495 Fax: 410-955-0476
E-mail: eselvin@jhu.edu

3. Timeline:

Data are available. We anticipate a rapid timeline for this project and aim to have a first draft of the manuscript to co-authors in <6 months.

4. Rationale:

Type 2 diabetes is defined by elevated blood glucose levels, or hyperglycemia. There are multiple ways to evaluate glucose levels: fasting glucose and hemoglobin A1c (HbA1c) are traditional glycemia biomarkers that are commonly used clinically¹, and fructosamine, glycated albumin and 1,5-AG are more recently proposed nontraditional glycemia biomarkers.^{2,3} While each of these biomarkers aim to capture blood glucose levels, they vary in their molecular structure, timespan and limitations. Fasting glucose is a direct measure of serum glucose after an 8 hour fast, representing instantaneous blood glucose levels, but has high intra-individual variability and is affected by factors such as acute illness, recent physical activity and time of day.⁴ HbA1c is formed as glucose binds to hemoglobin molecules within erythrocytes, and represents average blood glucose over the erythrocyte lifespan, 2-3 months.² Factors that impact erythrocyte turnover, such as hemolytic anemia or severe kidney disease, as well as rare hemoglobin variants, alter HbA1c levels in a manner not related to blood glucose levels.^{5,6} Fructosamine is glucose bound to total serum protein; glycated albumin is glucose bound to serum albumin, and is similar to fructosamine, as the majority of serum protein is comprised of albumin. Both represent average blood glucose over the previous 2-3 weeks. Fructosamine and glycated albumin levels can be affected by changes in serum protein and serum albumin metabolism, respectively.⁷ 1,5-AG is a molecule structurally similar to glucose that competes with glucose for reabsorption in the kidney at high concentrations of glucose and competes with glucose for enternal uptake among persons without diagnosed diabetes (Loomis et al, manuscript in preparation). It represents glycemic excursions over the previous 1-2 weeks.⁸

These glycemic biomarkers are under both environmental and genetic control. Previous studies have estimated the heritability (a measure of the proportion of total variance in a phenotype explained by genetics) of fasting glucose to range from 30-70%, and HbA1c to range from 20-75%, although most studies that evaluated both traits found lower heritability for fasting glucose than for HbA1c.⁹⁻¹⁷ This provides strong evidence that fasting glucose and HbA1c are under moderate to substantial genetic influence. Recent studies have identified genetic variants associated with fructosamine, glycated albumin and 1,5-AG^{11,18}, indicating some genetic impact on these biomarkers, and several studies have estimated heritability of 1,5-AG as part of a large metabolome panel,^{11,18} but no study has quantified the genetic control of the clinical assays of fructosamine, glycated albumin and 1,5-AG through heritability estimation. Consequently, no studies have had the ability to compare heritability across the traditional and nontraditional glycemic biomarkers. Heritability is population specific, affected by the relative genetic and environmental impact on the variance of an outcome, as well as by the method used to calculate it. Thus, calculating heritability for different traits in the same population will allow for a direct evaluation of their relative genetic impact. This will put previous genetic variant associations with these biomarkers into context, determining how much variability they explain and hence how much is left to uncover.

Heritability is traditionally measured in related individuals, but recently developed methods have allowed for heritability estimation among unrelated individuals^{19,20}. This permits large cohort studies such as ARIC to contribute to quantification of trait genetic control. In this analysis, we will calculate SNP heritability for fasting glucose, HbA1c, fructosamine, glycated albumin and 1,5-AG in ARIC participants and compare heritabilities across the different glycaemic biomarkers.

5. Main Hypothesis/Study Questions:

In this study, we will quantify the amount of SNP heritability for fructosamine, glycated albumin and 1,5-AG using participants from the ARIC study. We will also estimate heritability for fasting glucose and HbA1c and compare heritabilities across all biomarkers of hyperglycemia.

Hypothesis: Fructosamine, glycated albumin and 1,5-AG are under genetic control, the portion of which due to common genetic variants can be quantified using SNP heritability.

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 population: GWAS data is available for approximately 10,000 (8,000 white, 2,000 black) ARIC participants.

Study design: Analysis using glycaemic biomarker (fructosamine, glycated albumin and 1,5-AG) data collected at ARIC visit 2 (1990-1992). Blood for genetic data was collected at visit 1 (1987-1989), but as the DNA sequence does not change over time, it is acceptable to collect exposure (DNA) and outcome (markers of glycaemia) variables at different study visits.

Inclusion/exclusion ARIC individuals with consent for genetics studies, and with GWAS data that has passed standard (ARIC approved) quality thresholds will be included. We will exclude individuals without valid fructosamine, glycated albumin and 1,5-AG data available and individuals with prevalent diabetes at visit 2, (defined by self-reported physician diagnosed diabetes or taking diabetes medication) when fructosamine, glycated albumin and 1,5-AG were measured. We will also exclude one from each pair of related individuals.

Exposure variables: Imputed genotypes from GWAS data

GWAS data

DNA was extracted from blood collected at visit 1 from ARIC participants. Genotyping was done using the Affymetrix 6.0 array and imputed to 1000 Genomes Phase I reference panel for 37 million SNPs. Standard quality control measures were applied.

Outcomes: Fructosamine, glycated albumin, 1,5-AG

Fructosamine (Roche Diagnostics, Indianapolis IN, USA), glycated albumin (Asashi Kasei GA-L, Tokyo, Japan) and 1,5-AG (GlycoMark assay implemented on the Roche ModP, Wiston-

Salem, NC) were measured in 2012-2013 using a Roche Modular P800 system from serum collected at visit 2 (1990-92) and stored at -70°C.²¹

Covariates: Age (years) at visit 2, sex, study center and significant ($p < 0.05$) principal components.

Data analysis:

SNP heritability

We will utilize genomic-relatedness-based restriction maximum likelihood (GREML) to calculate SNP heritability for fructosamine, glycated albumin, fasting glucose and HbA1c.^{19,20} This method uses unrelated individuals, excluding related individuals to remove the potential for residual similarities in environmental exposures, and calculates a relatedness matrix (GRM) based on genotype similarities across all genotyped SNPs and all pairs of individuals in the study. It then regresses the outcome on this weighted matrix as a random effect in a mixed linear model (fixed effects include covariates and principal components), and the variance due to genotyped SNPs is calculated using restriction maximum likelihood estimation. This variance is the variance in the phenotype due to additive genetic effects of SNPs in the dataset, so dividing it by the overall variance results in a heritability estimate.

This method was originally developed by Yang and Visscher in 2010 using the GCTA software package and widely implemented for various phenotypes.^{20,22,23} One limitation with the original method is that linkage disequilibrium (LD) structure may affect calculations of genetic relatedness and must also be taken into account. SNPs in the same LD block are likely tagging the same causal variant but in the original Yang-Visscher method, the SNPs are given equal weight, hence upweighting the influence of that particular causal SNP. Speed et al attempted to remedy this with weighting by LD structure (eg, if 5 SNPs are in strong LD with each other, each SNP will contribute 1/5, totaling a contribution of 1 SNP) in a software package called LDAK.²⁴ Subsequent versions of the Yang-Visscher method, GCTA-LDMS stratify by allele frequency and LD structure in an attempt to remove the influence of LD structure.²⁵ There is debate in the literature as to which method is best, thus we will use both LDAK and GCTA-LDMS to estimate heritability.

We will compare heritabilities across different measures of hyperglycemia, fasting glucose, HbA1c, fructosamine, glycated albumin, and 1,5-AG and we will determine how much of heritability is explained by the significant SNPs. We will use genome-wide data to calculate SNP heritability separately in blacks and whites to minimize confounding by ancestry, and will control for principal components to further reduce confounding by ancestry.

Limitations:

SNP heritability is calculated using only the genetic variants for which we have high quality data, thus it reflects the percent of phenotypic variation due to the causal SNPs tagged by the SNPs in our data. Genome-wide SNP data and subsequent imputation aim to capture a large

portion of the genome, but do not well cover rare variants and thus do not account for all genetic variation across the entire genome. Thus, SNP heritability may underestimate true heritability.

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

b. If Yes, is the author aware that the file ICTDER03 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 ICTDER 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? 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 ICTDER03 must be used to exclude those with value RES_DNA = "No use/storage DNA"? 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.csc.c.unc.edu/ARIC/search.php>

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)?

#2387 Comparative genetics of fructosamine, glycated albumin, and 1,5- anhydroglucitol in the Atherosclerosis Risk in Communities Study

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

11.b. If yes, is the proposal

A. primarily the result of an ancillary study (list number*2006.02)

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.csc.c.unc.edu/atic/forms/>

12a. 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.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PUBMED Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from <http://publicaccess.nih.gov/> are posted in <http://www.csc.unc.edu/aric/index.php>, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to Pubmed central.

13. Per Data Use Agreement Addendum for the Use of Linked ARIC CMS Data, approved manuscripts using linked ARIC CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes No.

References:

1. Marathe PH, Gao HX, Close KL. American Diabetes Association Standards of Medical Care in Diabetes 2017. *J Diabetes*. 2017;40 suppl:1-142. <http://www.ncbi.nlm.nih.gov/pubmed/24357209>.
2. Goldstein DE, Little RR, Lorenz RA, et al. Tests of glycemia in diabetes. *Diabetes Care*. 2004;27(7):1761-1773. <http://www.ncbi.nlm.nih.gov/pubmed/15220264>.
3. True MW. Circulating biomarkers of glycemia in diabetes management and implications for personalized medicine. *J diabetes Sci Technol*. 2009;3(4):743-747.
4. Parrinello CM, Sharrett AR, Maruthur NM, et al. Racial differences in and prognostic value of biomarkers of hyperglycemia. *Diabetes Care*. 2016;39(4):589-595. doi:10.2337/dc15-1360.
5. Sacks DB. A1C versus glucose testing: A comparison. *Diabetes Care*. 2011;34(2):518-523. doi:10.2337/dc10-1546.
6. Sacks DB, John WG. Interpretation of hemoglobin A1c values. *JAMA*. 2014;311(22):2271-2272. doi:10.1001/jama.2014.6342.
7. Armbruster DA. Fructosamine: structure, analysis, and clinical usefulness. *Clin Chem*. 1987;33(12):2153-2163. <http://www.ncbi.nlm.nih.gov/pubmed/3319287>.
8. Buse JB, Freeman JLR, Edelman S V, Jovanovic L, McGill JB. Serum 1,5-anhydroglucitol (GlycoMark): a short-term glycemic marker. *Diabetes Technol Ther*. 2003;5(3):355-363. doi:10.1089/152091503765691839.
9. Meigs JB, Panhuysen CIM, Myers RH, Wilson PWF, Cupples LA. A genome-wide scan for loci linked to plasma levels of glucose and HbA(1c) in a community-based sample of Caucasian pedigrees: The Framingham Offspring Study. *Diabetes*. 2002;51(3):833-840.
10. Pilia G, Chen W-M, Scuteri A, et al. Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet*. 2006;2(8):e132. doi:10.1371/journal.pgen.0020132.
11. Shin S-Y, Fauman EB, Petersen A-K, et al. An atlas of genetic influences on human blood

- metabolites. *Nat Genet.* 2014;46(6):543-550. doi:10.1038/ng.2982.
12. Watanabe RM, Valle T, Hauser ER, et al. Familiality of quantitative metabolic traits in Finnish families with non-insulin-dependent diabetes mellitus. *Hum Hered.* 1999;49(3):159-168. doi:10.1159/000022865.
 13. Mills GW, Avery PJ, McCarthy MI, et al. Heritability estimates for beta cell function and features of the insulin resistance syndrome in UK families with an increased susceptibility to Type 2 diabetes. *Diabetologia.* 2004;47(4):732-738. doi:10.1007/s00125-004-1338-2.
 14. Hsueh WC, Mitchell BD, Aburomia R, et al. Diabetes in the Old Order Amish: characterization and heritability analysis of the Amish Family Diabetes Study. *Diabetes Care.* 2000;23(5):595-601. <http://www.ncbi.nlm.nih.gov/pubmed/10834415>.
 15. Mitchell BD, Kammerer CM, Blangero J, et al. Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. The San Antonio Family Heart Study. *Circulation.* 1996;94(9):2159-2170. <http://www.ncbi.nlm.nih.gov/pubmed/8901667>.
 16. Simonis-Bik AMC, Eekhoff EMW, Diamant M, et al. The heritability of HbA1c and fasting blood glucose in different measurement settings. *Twin Res Hum Genet.* 2008;11(6):597-602. doi:10.1375/twin.11.6.597.
 17. Snieder H, Sawtell PA, Ross L, Walker J, Spector TD, Leslie RDG. HbA1c levels are genetically determined even in type 1 diabetes. *Diabetes.* 2001;50(December):2858-2863. doi:10.2337/diabetes.50.12.2858.
 18. Long T, Hicks M, Yu H-C, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nat Genet.* 2017;49(4):568-578. doi:10.1038/ng.3809.
 19. Yang J, Bakshi A, Zhu Z, et al. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet.* 2015;47(10):1114-1120. doi:10.1038/ng.3390.
 20. Yang J, Benyamin B, McEvoy BP, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet.* 2010;42(7):565-569. doi:10.1038/ng.608.
 21. Selvin E, Rawlings AM, Grams M, et al. Fructosamine and glycated albumin for risk stratification and prediction of incident diabetes and microvascular complications: A prospective cohort analysis of the Atherosclerosis Risk in Communities (ARIC) study. *Lancet Diabetes Endocrinol.* 2014;2(4):279-288. doi:10.1016/S2213-8587(13)70199-2.
 22. Visscher PM, Yang J, Goddard ME. A Commentary on “ Common SNPs Explain a Large Proportion of the Heritability for Human Height” by Yang et al. (2010). *Twin Res Hum Genet.* 2010;13(6):517-524.
 23. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet.* 2011;88(1):76-82. doi:10.1016/j.ajhg.2010.11.011.
 24. Speed D, Hemani G, Johnson MR, Balding DJ. Improved heritability estimation from genome-wide SNPs. *Am J Hum Genet.* 2012;91(6):1011-1021. doi:10.1016/j.ajhg.2012.10.010.
 25. Yang J, Bakshi A, Zhu Z, et al. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet.* 2015;47(10):1114-1120. doi:10.1038/ng.3390.