

ARIC Manuscript Proposal # 1380

PC Reviewed: 06/10/08
SC Reviewed: _____

Status: A
Status: _____

Priority: 2
Priority: _____

1.a. Full Title: Association of the single nucleotide polymorphism rs780094 in the glucokinase regulator gene (*GCKR*) and metabolic phenotypes in the ARIC Study

b. Abbreviated Title (Length 26 characters): GCKR and metabolic traits

2. Writing Group:

Writing group members: Anna Kottgen, Mark Bi, Josef Coresh, Eric Boerwinkle, Laura Rasmussen-Torvik, Caroline Fox, Linda Kao. Others welcome, invited: Ron Hoogeveen, Kari North. Investigators from other cohorts may be included upon inclusion of data from these additional cohorts.

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

First author: Anna Kottgen

Address: 2024 E. Monument St.
Welch Center for Prevention, Epidemiology, and Clinical Research
Baltimore, MD 21287

Phone: (410) 245-6897 Fax: (410) 955-4076
E-mail: akottgen@jhsphe.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: Linda Kao
Address: 615 N Wolfe St, Room W6513
Baltimore, MD 21205

Phone: (410) 614-0945 Fax: (410) 955-0863
E-mail: wkao@jhsphe.edu

3. Timeline: Data analysis to start immediately, first draft of the manuscript by August 2008.

4. Rationale: The clustering of cardiovascular risk factors such as glycemic traits (impaired fasting glucose, diabetes mellitus and related measures) altered lipid levels,

obesity, increased blood pressure, and several biomarkers within individuals is established in the literature.¹ While the individual risk factors have heritable components,² it is less clear whether the observed clustering has a common biologic basis.^{3,4} In the recent wave of genome-wide association studies (GWAS), susceptibility loci for some of the individual metabolic components have been discovered, such as for diabetes mellitus,^{5,6} serum triglyceride and other blood lipid levels,⁷ as well as biomarkers related to the metabolic syndrome.^{8,9} Many of these studies identified genetic loci that only seem to confer risk for the respective discovery trait, based on the data available so far. In light of the clustering of metabolic risk factors, it is therefore of particular interest to investigate genetic loci that have demonstrated pleiotropy, i.e. association to more than one phenotype. Common variants in the glucokinase regulator (*GCKR*) gene, and specifically the single nucleotide polymorphism (SNP) rs780094, have recently been reported by independent GWAS as associated with serum triglyceride levels^{10,11} and serum C-reactive protein levels.⁸ Moreover, candidate gene studies have linked this polymorphism to fasting insulin levels, impaired glucose tolerance, and type 2 diabetes mellitus.¹² We therefore propose to study the SNP rs780094 in *GCKR* in the ARIC Study. We are specifically interested in its association with metabolic traits, including all glycemic traits, obesity, lipid traits, blood pressure, and also with biomarkers related to the metabolic syndrome such as CRP and urate levels, as well as markers of kidney function in order to evaluate a common genetic/biologic basis to the clustering observed for these traits. Depending on the availability of the data from the ARIC GWAS, our study may be extended to other variants in the region of the *GCKR* gene. Furthermore, data from other cohorts may be included as independent study samples during the course of the project.

5. Main Hypothesis/Study Questions:

1. Do the recently described associations of rs780094 and serum triglycerides, glycemic traits, and serum CRP replicate in white and black participants of the ARIC Study?
2. Is there an independent association with more than one or all components of the metabolic syndrome? Is the risk allele for each such observed association the same?
3. Do the observed independent associations extend to biomarkers that have been linked to the metabolic syndrome?
4. Is the rs780094 variant associated with cross-sectional and prospective endpoints, such as cardiovascular disease, diabetes mellitus, kidney disease and gout?
5. In case genome-wide data are available, are there other variants in the *GCKR* gene region more significantly associated with the traits of interest in ARIC participants, and specifically in black 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 and inclusion/exclusion: subjects and sample size

Analyses will be conducted using data from 10,929 white and 3,960 black ARIC participants who consented to genetic research and had genotypes at rs780094 available. The cross-sectional association with metabolic traits will be evaluated at the baseline examination (visit 1) to maximize sample size, except for the traits measured at visit 4 only (CRP, OGTT, markers of kidney function and damage). Prospective analyses will be conducted using visit 1 as the baseline and incorporating follow-up data through January 1, 2005.

Exposure Measurements and Definitions

The rs780094 variant was genotyped by the central ARIC DNA laboratory in Houston. Race-specific measures of genotyping quality control will include genotyping call rate, evaluation of conformation to Hardy-Weinberg proportions, comparison of the minor allele frequency to the ones in the respective HapMap referent population, and the evaluation of genotyping of blind duplicate samples by the ARIC coordinating center.

Outcome Measurements and Definitions

Primary cross-sectional outcomes are the individual outcomes previously described as associated with rs780094. These outcomes will be used as measured at the baseline visit (triglycerides, fasting glucose) or visit 4 (CRP). Diabetes mellitus will be defined as a self-reported diagnosis of diabetes, the current intake of diabetes medications, or fasting plasma glucose ≥ 126 mg/dl or non-fasting glucose ≥ 200 mg/dl. Further primary cross-sectional outcomes include the components of the metabolic syndrome, defined following the US National Cholesterol Education Program Adult Treatment Panel III¹³ recommendations as 1) waist circumference ≥ 102 cm (male) or ≥ 88 cm (female); 2) plasma triglycerides ≥ 150 mg/dl or lipid lowering treatment; 3) plasma HDL-cholesterol < 40 mg/dl (male) or < 50 mg/dl (female); 4) systolic/diastolic blood pressure $\geq 130/85$ mmHg or the current intake of antihypertensive medications; and 5) fasting plasma glucose ≥ 100 mg/dl.

Secondary cross-sectional outcomes include a) measures of kidney function and damage, estimated glomerular filtration rate and the urinary albumin-to-creatinine ratio, calculated and defined as described previously;¹⁴ and b) serum urate measured at visit 1 and gout defined as a self-report of the disease at study visit 4.

Primary prospective outcomes are incident diabetes mellitus, defined as described previously.¹⁵ Secondary incident outcomes include incident cardiovascular disease as well as incident kidney disease.^{14, 16, 17}

Statistical Analysis and Anticipated Challenges

Analyses will be conducted stratified by race. We will use an additive genetic model, and analyses will be divided into primary and secondary analyses (see outcomes section). After the conduct of standard quality control analyses as outlined in the exposure section above, linear or logistic regression modeling will be conducted as appropriate to evaluate the cross-sectional outcomes. Cox-proportional hazards analyses will be conducted to evaluate associations with the prospective outcomes, and appropriate model checks will be conducted. First, minimally adjusted analyses (adjustment for age, sex, and study center) will be conducted to replicate previously observed associations and to examine

the association with each of the outcomes. Then, fully adjusted models will be evaluated to assess whether these different metabolic traits are independently associated with rs780094, beyond the known associations. For example, we will assess whether the association with CRP remains significant after adjusting for serum triglycerides and fasting glucose. Similar analyses will be conducted in a step-wise fashion to evaluate whether the SNP is independently associated with more than one component of the metabolic syndrome as defined above, and with other traits associated with the metabolic syndrome such as impaired kidney function. Additional secondary data analyses may be added depending on the results of the analyses.

Since several associations of rs780094 and metabolic traits have been reported already, replication of these associations in further independent study samples will not be necessary, and the significance threshold for these associations will be set at $\alpha=0.05$. In case associations with additional metabolic traits are detected, other cohort data may be included. Moreover, in case of association to additional metabolic traits, we will consider more complex analyses such as a network analysis or factor analysis.

In case of a positive association among the black ARIC participants, we will repeat these analyses adjusting for % European ancestry derived from a panel of ancestry informative markers genotyped in the ARIC Study.

The issue of testing multiple phenotypes will be addressed by dividing the analyses into primary and secondary outcomes as detailed above, and assigning a significance threshold of $\alpha=0.01$ to the secondary outcomes.

All analyses will be conducted using Stata software. Should the analysis be extended to include additional GWAS data in the *GCKR* region, we will use the software plink to analyze the data.

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 ICTDER03 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

8.c. If yes, is the author aware that the participants with RES_DNA = 'not for profit' restriction must be excluded if the data are used by a for profit group?
 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.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)?

1. MS #1237: Association between genetic variants conferring risk for type 2 diabetes mellitus and incident chronic kidney disease
2. MS #1273: Genetic risk score for type 2 diabetes
3. MS #1307: Gene-by-Environment Interaction for Type 2 Diabetes
4. MS #1343: Stage II of a Genome-Wide Association Study for Genetic Variants Associated with Uric Acid Levels and Gout

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* _____)
 B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* __ AS #2006.16 _____)

*ancillary studies are listed by number at <http://www.csc.unc.edu/anic/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. Reynolds K, He J. Epidemiology of the metabolic syndrome. *Am J Med Sci.* 2005; 330: 273-279.
2. Austin MA, Edwards KL, McNeely MJ, Chandler WL, Leonetti DL, Talmud PJ, Humphries SE, Fujimoto WY. Heritability of multivariate factors of the metabolic syndrome in nondiabetic Japanese americans. *Diabetes.* 2004; 53: 1166-1169.
3. Pollex RL, Hegele RA. Genetic determinants of the metabolic syndrome. *Nat Clin Pract Cardiovasc Med.* 2006; 3: 482-489.

4. Benyamin B, Sorensen TI, Schousboe K, Fenger M, Visscher PM, Kyvik KO. Are there common genetic and environmental factors behind the endophenotypes associated with the metabolic syndrome? *Diabetologia*. 2007; 50: 1880-1888.
5. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*. 2007; 445: 881-885.
6. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Bostrom KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney AS, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jorgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marvelle AF, Meisinger C, Midthjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JR, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjogren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ, Wellcome Trust Case Control Consortium, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet*. 2008; 40: 638-645.
7. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, Cooper GM, Roos C, Voight BF, Havulinna AS, Wahlstrand B, Hedner T, Corella D, Tai ES, Ordovas JM, Berglund G, Vartiainen E, Jousilahti P, Hedblad B, Taskinen MR, Newton-Cheh C, Salomaa V, Peltonen L, Groop L, Altshuler DM, Orho-Melander M. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet*. 2008; 40: 189-197.
8. Ridker PM, Pare G, Parker A, Zee RY, Danik JS, Buring JE, Kwiatkowski D, Cook NR, Miletich JP, Chasman DI. Loci related to metabolic-syndrome pathways including LEPR, HNF1A, IL6R, and GCKR associate with plasma C-reactive protein: the Women's Genome Health Study. *Am J Hum Genet*. 2008; 82: 1185-1192.
9. Li S, Sanna S, Maschio A, Busonero F, Usala G, Mulas A, Lai S, Dei M, Orru M, Albai G, Bandinelli S, Schlessinger D, Lakatta E, Scuteri A, Najjar SS, Guralnik J, Naitza S, Crisponi L, Cao A, Abecasis G, Ferrucci L, Uda M, Chen WM, Nagaraja R. The GLUT9 Gene Is Associated with Serum Uric Acid Levels in Sardinia and Chianti Cohorts. *PLoS Genet*. 2007; 3: e194.

10. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, Burt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*. 2007; 316: 1331-1336.
11. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, Heath SC, Timpson NJ, Najjar SS, Stringham HM, Strait J, Duren WL, Maschio A, Busonero F, Mulas A, Albai G, Swift AJ, Morken MA, Narisu N, Bennett D, Parish S, Shen H, Galan P, Meneton P, Herberg S, Zelenika D, Chen WM, Li Y, Scott LJ, Scheet PA, Sundvall J, Watanabe RM, Nagaraja R, Ebrahim S, Lawlor DA, Ben-Shlomo Y, Davey-Smith G, Shuldiner AR, Collins R, Bergman RN, Uda M, Tuomilehto J, Cao A, Collins FS, Lakatta E, Lathrop GM, Boehnke M, Schlessinger D, Mohlke KL, Abecasis GR. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet*. 2008; 40: 161-169.
12. Sparso T, Andersen G, Nielsen T, Burgdorf KS, Gjesing AP, Nielsen AL, Albrechtsen A, Rasmussen SS, Jorgensen T, Borch-Johnsen K, Sandbaek A, Lauritzen T, Madsbad S, Hansen T, Pedersen O. The GCKR rs780094 polymorphism is associated with elevated fasting serum triacylglycerol, reduced fasting and OGTT-related insulinaemia, and reduced risk of type 2 diabetes. *Diabetologia*. 2008; 51: 70-75.
13. Grundy SM, Brewer HB, Jr, Cleeman JI, Smith SC, Jr, Lenfant C, American Heart Association, National Heart, Lung, and Blood Institute. Definition of metabolic syndrome: Report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. *Circulation*. 2004; 109: 433-438.
14. Hsu CC, Kao WH, Coresh J, Pankow JS, Marsh-Manzi J, Boerwinkle E, Bray MS. Apolipoprotein E and progression of chronic kidney disease. *JAMA*. 2005; 293: 2892-2899.
15. Wang L, Folsom AR, Zheng ZJ, Pankow JS, Eckfeldt JH, ARIC Study Investigators. Plasma fatty acid composition and incidence of diabetes in middle-aged adults: the Atherosclerosis Risk in Communities (ARIC) Study. *Am J Clin Nutr*. 2003; 78: 91-98.

16. White AD, Folsom AR, Chambless LE, Sharret AR, Yang K, Conwill D, Higgins M, Williams OD, Tyroler HA. Community surveillance of coronary heart disease in the Atherosclerosis Risk in Communities (ARIC) Study: methods and initial two years' experience. *J Clin Epidemiol.* 1996; 49: 223-233.

17. Rosamond WD, Folsom AR, Chambless LE, Wang CH, McGovern PG, Howard G, Copper LS, Shahar E. Stroke incidence and survival among middle-aged adults: 9-year follow-up of the Atherosclerosis Risk in Communities (ARIC) cohort. *Stroke.* 1999; 30: 736-743.