ARIC Manuscript Proposal # 1237

PC Reviewed: __3/_13/07	**Status: __A__**	**Priority: _2__**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Association between genetic variants conferring risk for type 2 diabetes mellitus and incident chronic kidney disease

b. Abbreviated Title (Length 26 characters):
Diabetes genes & incident CKD

2. Writing Group:
Writing group members: Anna Kottgen, Wen Hong Linda Kao, Kari E. North, Josef Coresh, others welcome. Invited: Eric Boerwinkle, James S. Pankow

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]**

F**irst 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@jhsph.edu

Corresponding/senior author (if different from first author correspondence will be sent to both the first author & the corresponding author): **Josef Coresh**
Address: 2024 E. Monument St.
Welch Center for Prevention, Epidemiology, and Clinical Research
Baltimore, MD 21287

Phone: (410) 955-0495 Fax: (410) 955-4076
E-mail: coresh@jhu.edu

3. Timeline: Analyses to start immediately; additional genotyping and analyses and manuscript preparation are projected to take place over the next year.

4. Rationale:

Chronic kidney disease (CKD) has been recognized as a public health problem which affects an estimated 19 million adults in the US [1, 2]. Progression of CKD may lead to end-stage renal disease (ESRD). The yearly mortality rates for individuals treated with dialysis for ESRD exceed 20% [3]. Therefore, early identification of individuals at increased risk for CKD and effective intervention is essential. Previous studies have shown that sub-groups of individuals susceptible to kidney disease exist [4]. Additionally, multiple studies have confirmed that kidney disease is heritable [5]. Apart from major cardiovascular risk factors such as hypertension, genetic causes contribute directly to the complex disease CKD.

One of the major contributors to CKD is type 2 diabetes mellitus (T2DM). Diabetes mellitus accounts for about 45% of new ESRD cases in the US, and about 30% of individuals with type 2 diabetes develop overt kidney disease. While it has largely been assumed that the excess risk of renal complications seen among individuals with T2DM is likely the result of the interaction between inadequate glycemic control over time and susceptibility genes for renal disease, we hypothesize that genetic variants conferring T2DM risk may also contribute to the risk of CKD, independent of their effects on T2DM. Several lines of evidence have led us to this hypothesis:

First, rare monogenic diseases such as familial hypoplastic glomerulocystic kidney disease caused by mutations in the HNF1b gene [6, 7] are characterized by the development of both renal cysts and diabetes syndrome in affected individuals. In this situation renal disease is not a consequence of diabetes, but rather, both the renal and diabetes phenotypes are caused by the same mutations. This suggests that genetic variations can affect common physiologic or cellular pathways that are present in both the kidney and the pancreas. This has lead us to considering specific shared pathways (detail beyond this proposal).

Second, the central nervous system has been hypothesized to contribute to the etiology of T2DM. It is therefore also thinkable that genetic variations can affect common physiologic or cellular pathways that are present in both the central nervous system as well as the kidney, but not in the pancreas (for T2DM, there is a series of review articles from Nature Genetics Review, I believe, maybe just Nature, that have gone over this; will look it up tomorrow).

Third, previous linkage studies have identified overlapping genomic regions that have been suggestive for linkage with T2DM and GFR on chromosome 20 ([8-10] and unpublished results of GFR in Mexican Americans from FIND) and on chromosome 1q ([11-13] and unpublished GFR results in Mexican Americans from FIND).

Lastly, preliminary results from the Atherosclerosis Risk in Communities (ARIC) Study indicate that one gene known to confer risk for T2DM (TCF7L2, manuscript proposal #1141) and another gene that is apparently associated with T2DM in the ARIC Study (KIF6, manuscript proposal #1161) seem to confer risk for incident CKD beyond their effect on T2DM.

Therefore, we propose to study the association between incident CKD and well-known T2DM susceptibility alleles. The field of T2DM genetics has been progressing rapidly. Variants in genes that confer risk for T2DM, such as, *KCNJ11*, *PPARG*, *CAPN10*, and *HNF4a*, have been established from candidate gene studies / linkage studies over the past decade [14-16]. Moreover, recently result from new genome-wide association analyses have

added genes, such as *TCF7L2* [17], *SLC30A8*, and 2 loci in the *IDE–KIF11–HHEX* and *EXT2–ALX4* regions [18] to the list of T2DM susceptibility genes. We propose here to study variations in *TCF7L2, KIF6, PPARG, CAPN10, HNF4a, KCNJ11, SLC30A8* and variations in the *IDE–KIF11–HHEX* and *EXT2–ALX4* regions. In addition, we propose to study variants in the less well established candidate *KIRREL3*. Evidence for association of variants within the *KIRREL3* gene with T2DM exists from one genome wide association study [18], and its transcription product (Neph2) localizes to the renal glomerulus where it is an integral part of the glomerular filtration barrier [19]. The protein product of the SLC30A8 gene has been described to be exclusively located within the beta cells of the pancreas. Typing a variant within this gene and including it into our study should serve as a control of any observed effect: if consistent with our hypothesis, one would expect to observe an effect of variants in diabetes genes and kidney function beyond their effect through diabetes, however, this should not be the case for variants within the SLC30A8 gene.

Association studies in large study populations provide greater power for identifying variants responsible for such traits [20]. The hypothesis being proposed is novel and biologically driven. The large representative sample, bi-racial population, information about and measurement of cardiovascular and renal risk factors and function, and the extended follow-up of the ARIC Study make it possible to prospectively study genetic risk factors that may confer susceptibility to both T2DM and CKD. In addition, studying large prospective cohorts like the ARIC cohort allows for quantification both of the relative risk as well as the attributable risk associated with CKD susceptibility genes.

5. Main Hypothesis/Study Questions:

Main hypothesis: Some of the genetic variants that have been associated with T2DM in previous studies will be associated with incident CKD in the ARIC Study.

Study questions:

1. Will there be an effect of T2DM risk variants on incident CKD? Will this effect extend beyond the variants' effect on T2DM and hyperglycemia?
2. Will the association of these genetic variants and incident CKD be consistently present for different definitions of incident CKD, as well as cross-sectionally with measures of kidney function (estimated glomerular filtration rate) and kidney damage (albuminuria)?
3. How will such associations be influenced by adjustment for or stratification on T2DN, obesity, and hyperglycemia?
4. Will the association of such genetic variants and CKD be present in both African American and Caucasian individuals? Furthermore, can differences in frequencies of such genetic variants between these two populations account for part of the differences in disease risks?

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 of all ARIC participants meeting the inclusion criteria from baseline (visit 1, 1987-1989) through January 1, 2003.

Some of the single nucleotide polymorphisms (SNPs) to be examined (KIF6, 2 variants) have been genotyped as part of a panel proposed by Celera diagnostics in collaboration with Dr. Boerwinkle as described in ARIC Ancillary Study 2004.11 and manuscript proposal #1161. 5 variants in the TCF7L2 gene have already been typed in Dr. Boerwinkle's lab as contract work, and are the variants proposed to study in MP #1141. The *PPARG* variant Pro12Ala was typed as part of the GxE in all ARIC participants. Variants in the *KCNJ11, CAPN10s* and *KIF6* genes are on the ARIC MRI list. SNP43 of *CAPN10* was typed in 1800 African Americans only. The ARIC DNA laboratory has typically genotyped the most promising diabetes variants in the entire cohort. We will coordinate with Dr. Boerwinkle in proposing the relevant genotypes which will be useful for this proposal as well as others (relevant phenotypes justifying genotyping include diabetes, obesity, hyperglycemia and cancer).

Inclusions/exclusions: Participants who did not consent to genotyping will be excluded from analysis (use of DNA data distributed by the Coordinating Center, confirmation by using the variable "res_dna" in datafile "ictder02" (n = 45 individuals did not consent to DNA use for the type of study outlined here)). Moreover, individuals reporting race other than African American or white will be excluded (n = 48), as will be individuals missing variables needed to calculate estimated glomerular filtration rate (eGFR) as a measure of kidney function at visit 1 (n = 150, all due to missing serum creatinine (variable chma09)). Depending on the definition of incident CKD used (see below), individuals with severe hypercreatinemia (n = 40, serum creatinine ≥2.0 mg/dl for men, ≥1.8 mg/dl for women) or those with eGFR < 60 at visit 1 (n = 462) will also be excluded from analyses. Individuals lacking information on diabetes mellitus at visit 1 (n = 116) will be excluded as well. The number of individuals excluded due to missing or unknown genotype will depend on the SNP under investigation.

Outcome: The primary outcome will be CKD defined as defined by a rise in serum creatinine of at least 0.4 mg/dl above baseline (n = 1,201) or a hospitalization discharge or death coded for chronic renal disease (*International Classification of Diseases, Ninth Revision* [*ICD-9*] codes 581-583 or 585-588), hypertensive renal disease (*ICD-9* code 403), hypertensive heart and renal disease (*ICD-9* code 404), unspecified disorder of kidney and ureter (*ICD-9* code 593.9), diabetes with renal manifestations (*ICD-9* code 250.4), kidney transplantation, renal dialysis, or adjustment/fitting of catheter (*ICD-9* codes V42.0, V45.1, or V56), hemodialysis (*ICD-9* code 39.95) or peritoneal dialysis (*ICD-9* code 54.98), without acute renal failure (*ICD-9* codes 584, 586, 788.9, and 958.5) as the primary or secondary hospitalization code, all from ARIC surveillance datasets (c02occ1, c02celb1, c02dtha1).

In additional analyses, incident CKD as defined by a decrease in eGFR from ≥60 ml/min/1.73m^2 at baseline to <60 ml/min/1.73m^2 at the second or fourth follow-up visit, or CKD hospitalization or death (defined as above, n = 1,616) will be explored.
Additionally, albuminuria at visit 4 as well as eGFR at visit 1 and 4 will be investigated cross-sectionally.

eGFR as a measure of kidney function will be calculated using the abbreviated Modification of Diet in Renal Disease (MDRD) Study formula: eGFR (ml/min/1.73m^2 = 186.3 * (serum creatinine)$^{-1.154}$ * age$^{-0.203}$ * (0.742 if female) * (1.21 if black) [21].

Other variables of interest: Variables needed to calculate eGFR: serum creatinine from visit 1, 2, and 4 (chma09, chmb08, lipd6a), age (v1age01), race (racegrp), and gender (gender). The variable "ACRv2" will be used to assess the albumin-to-creatinine ratio (ACR) at visit 4. The SNPs to be explored are contained in the datafiles "Celera AS" (*KIF6*, 1 SNP, variable cv14), "snp_p.sas7bdat" (*KIF6*, 1 SNP and *PPARG*, 2 SNPs) and "uc4598_p.sas7bdat" (*TCF7L2*, 5 SNPs).
Diabetes mellitus at each study visit (1-4) will be defined as present if fasting glucose ≥126 mg/dl after an 8-hour fast (variable fast0802), non-fasting glucose ≥200 mg/dl (variable glucos01), reported physician diagnosis of DM (variable HOM10E), or the reported intake of DM medication in the previous 2 weeks (variable msra08f). In addition, DM at visit 4 will be updated incorporating information of the oral glucose tolerance test at visit 4 (variable gl2siu41), classifying individuals with a 2-hour OGTT value >11.1 mmol/l as diabetic. Information on HbA1c and duration of diabetes will be incorporated if known (all individuals with diabetes and a random subset of others). Incident DM cases and their follow-up time will be assessed as reported previously [22].
Further covariates include risk factors for CHD: blood pressure (variables sbp21a and dbp21a), use of anti-hypertensive medication (hyptmd04), HDL- and LDL-cholesterol (hdl02 and ldl01), diabetes mellitus (diabts03), smoking (cursmk01), and body mass index (bmi01) at visit 1. Body mass index will be incorporated as a time-varying variable using BMI at study visits 2, 3, and 4.
Other covariates may be selected, depending on the specific SNP of interest and the hypothesized function of the corresponding gene.

Data analysis:
Data checks: Hardy-Weinberg equilibrium (HWE) in controls will be checked by race for each SNP by using Fisher exact test [23].
Differences in the distribution of genotypes between individuals included in the study and those excluded from our analyses will be conducted. Moreover, differences in the distribution of CKD and DM cases between those missing genotype data and those with available genotype data will be assessed using chi-square tests.
Exploratory and cross-sectional data analyses: The distribution of baseline characteristics in the study population by genotype as well as by outcome will be computed using t-tests, chi-square tests and ANOVA as applicable. All association analyses will first be examined within each self-reported race groups as well as stratified by DM status. For cross-sectional analyses of the association of genotypes with eGFR and albuminuria, mean eGFR and albumin-to-creatinine ratio (ACR) will be estimated and compared for the three genotypes at each SNP using ANOVA. Multiple linear

regression will be used to examine the association of genotypes and continous eGFR and ACR. Models will be constructed to account for effects of potential confounders similar to the construction of Cox proportional hazards models (see below).

Survival analyses: For CKD cases, follow-up time will be counted from visit 1 until the visit date at which the creatinine rise / eGFR decline occurred, or the date of CKD hospitalization discharge or death, or the earlier of the two dates for participants meeting both definitions. Non-cases will be censored at the earlier of the date of last contact (or date of non-CKD death) or December 31, 2002. Incidence rates of CKD will be calculated using person-time methods. Kaplan-Meier estimates of mortality will be computed, and log-rank tests will be used to compare survival curves among the genotypes. These analyses will be carried out adjusted for baseline DM status as well as stratified by baseline DM status.

In regression analyses, an additive genetic model will be assumed unless indicated otherwise by results of the analysis, or unless the allele frequency of a given candidate variant is low, in which case a dominant model combining the risk of heterozygotes and homozygotes for the rare allele will be used. Genotype will be coded as 0 (zero copies of the risk increasing allele), 1 (one copy of the risk increasing allele), or 2 (two copies of the risk increasing allele). Stratified analyses will be conducted to examine the potential for interaction and effect modification between the covariate (e.g., DM, BMI, hypertension) and the association of the SNP and incident CKD.

Each SNP will be tested for association with incident CKD in crude Cox proportional hazard analyses, by race if applicable. Those with a p-value of < 0.1 in these analyses will be considered for further analyses. Cox proportional hazard regression will then be used to estimate the effect size (RH of incident CKD) and 95% confidence interval for each SNP. Subsequent multivariate models will include basic variables (age, sex, study center, and race if no interaction by race is observed), co-morbidities thought to act as potential confounders, and relevant potential intermediate variables depending on the putative function of the gene in which the SNP is located.

Control for diabetes: First, analyses will be examined stratified by diabetes status. Individuals will contribute person time to the diabetes group they are in after the visit at which diabetes status is known, e.g. if someone develops diabetes at visit 3 they will be considered non-diabetic until that point and diabetic thereafter. In additional analyses, individuals with prevalent DM will be excluded from analyses, and those with incident DM will be censored at the time of their DM event should it occur prior to any CKD event. Additionally, individuals with DM by visit 4 will be followed for incident CKD after visit 4 stratified by categories of albuminuria (defined as normo, micro-, and macroalbuminuria based on the ACR) at visit 4. If the increased risk of CKD by carriers of the high risk genotype is present before but not after adjustment for diabetes we will conclude that the risk of CKD is mediated by diabetes. If the association is substantially attenuated but still significant we will need to consider adjustment for measurement error in level of glycemic control and the ability to determine whether an independent effect on CKD risk if present will likely be limited. If the association is similar before and after stratification or adjustment for diabetes we will conclude that the genotype is likely to confer risk of CKD through an independent pathway of hyperglycemia and diabetes.

Determination of statistical significance:

We recognize the limitations of screening a number of genetic variants as risk factors for CKD. Determining statistical significance in the face of multiple testing is controversial [23]. In this case, we will use the False Positive Report Probability (FPRP) [24] to adjust for the number of SNPs tested as we have proposed previously (MP #1203).
The power in ARIC is high for moderate associations. A simplified analysis of CKD cases vs. non-cases (ignoring follow-up time) shows that assuming alpha = 0.05, an OR of 1.3 (or 1.6 if larger genetic effects are hypothesized) and allele frequencies of 0.2, 0.1, 0.05 and 0.01 results in calculated power of 98.5% (100%), 88.3% (100%), 64.4% (98.9%), and 19.0% (53.7%).

Limitations:
An important limitation will be our inability to exclude the possibility that any positive findings are due to residual confounding. Repeated measurements of blood glucose are only available through study visit 4, and we cannot exclude the possibility that any observed association between variations in a DM gene and incident CKD beyond the effect conferred by diabetes could be due to T2DM developed after the last study visit. In addition, the duration of fasting, the physician diagnosis of T2DM, and the intake of T2DM medication are based on self report. To address this concern, sensitivity analyses will be conducted limiting events to those occurring before study visit 4.
Another limitation to the proposed research project is the problem of multiple comparisons that we will address as outlined above. Another problem is the possible impact of undetected population substructures on the study findings. However, the impact of population stratification on false-positive findings has been reported as low [23]. In addition, we will be getting data from ancestry informative makers from CIDR shortly. This information can then be incorporated into our analyses.
If there are positive findings, they will need to be scrutinized in terms of biological mechanism and tested in further studies.

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)?

MP 1141: TCF7L2 and diabetes mellitus
MP 1161: KIF6 and diabetes mellitus
MP 1203: Association between candidate genetic variants and incident chronic kidney disease: The Atherosclerosis Risk in Communities (ARIC) Study

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
___ A. primarily the result of an ancillary study (list number*)__
Some genetic data from ARIC ancillary study 2004.11,
albuminuria data from ARIC ancillary study 2002.02
___ B. primiarly 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. Coresh J, Astor BC, Greene T, Eknoyan G, Levey AS. Prevalence of chronic kidney disease and decreased kidney function in the adult US population: Third National Health and Nutrition Examination Survey. *Am J Kidney Dis.* 2003; 41: 1-12.

2. National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis.* 2002; 39: S1-266.

3. National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis.* 2002; 39: S1-266.

4. Klag MJ, Whelton PK, Randall BL, Neaton JD, Brancati FL, Ford CE, Shulman NB, Stamler J. Blood pressure and end-stage renal disease in men. *N Engl J Med.* 1996; 334: 13-18.

5. Freedman BI, Satko SG. Genes and renal disease. *Curr Opin Nephrol Hypertens.* 2000; 9: 273-277.

6. Bingham C, Bulman MP, Ellard S, Allen LI, Lipkin GW, Hoff WG, Woolf AS, Rizzoni G, Novelli G, Nicholls AJ, Hattersley AT. Mutations in the hepatocyte nuclear factor-1beta gene are associated with familial hypoplastic glomerulocystic kidney disease. *Am J Hum Genet.* 2001; 68: 219-224.

7. Decramer S, Parant O, Beaufils S, Clauin S, Guillou C, Kessler S, Aziza J, Bandin F, Schanstra JP, Bellanne-Chantelot C. Anomalies of the TCF2 Gene Are the Main Cause of Fetal Bilateral Hyperechogenic Kidneys. *J Am Soc Nephrol.* 2007; .

8. Ghosh S, Watanabe RM, Valle TT, Hauser ER, Magnuson VL, Langefeld CD, Ally DS, Mohlke KL, Silander K, Kohtamaki K, Chines P, Balow Jr J, Birznieks G, Chang J, Eldridge W, Erdos MR, Karanjawala ZE, Knapp JI, Kudelko K, Martin C, Morales-Mena A, Musick A, Musick T, Pfahl C, Porter R, Rayman JB. The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. I. An autosomal genome scan for genes that predispose to type 2 diabetes. *Am J Hum Genet.* 2000; 67: 1174-1185.

9. Rotimi CN, Chen G, Adeyemo AA, Furbert-Harris P, Parish-Gause D, Zhou J, Berg K, Adegoke O, Amoah A, Owusu S, Acheampong J, Agyenim-Boateng K, Eghan BA,Jr, Oli J, Okafor G, Ofoegbu E, Osotimehin B, Abbiyesuku F, Johnson T, Rufus T, Fasanmade O, Kittles R, Daniel H, Chen Y, Dunston G, Collins FS, the Africa America Diabetes Mellitus (AADM) Study. A genome-wide search for type 2 diabetes susceptibility genes in West Africans: the Africa America Diabetes Mellitus (AADM) Study. *Diabetes.* 2004; 53: 838-841.

10. Hani e, Zouali H, Philippi A, Beaudoin JC, Vionnet N, Passa P, Demenais F, Froguel P. Indication for genetic linkage of the phosphoenolpyruvate carboxykinase (PCK1) gene region on chromosome 20q to non-insulin-dependent diabetes mellitus. *Diabetes Metab.* 1996; 22: 451-454.

11. Vionnet N, Hani E, Dupont S, Gallina S, Francke S, Dotte S, De Matos F, Durand E, Lepretre F, Lecoeur C, Gallina P, Zekiri L, Dina C, Froguel P. 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: 1470-1480.

12. 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: 1175-1182.

13. Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake D, Foroud T, Kobes S, Baier L, Burns DK, Almasy L, Blangero J, Garvey WT, Bennett PH, Knowler WC. 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: 1130-1138.

14. Yen CJ, Beamer BA, Negri C, Silver K, Brown KA, Yarnall DP, Burns DK, Roth J, Shuldiner AR. Molecular scanning of the human peroxisome proliferator activated receptor gamma (hPPAR gamma) gene in diabetic Caucasians: identification of a Pro12Ala PPAR gamma 2 missense mutation. *Biochem Biophys Res Commun.* 1997; 241: 270-274.

15. Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M, Hinokio Y, Lindner TH, Mashima H, Schwarz PE, del Bosque-Plata L, Horikawa Y, Oda Y, Yoshiuchi I, Colilla S, Polonsky KS, Wei S, Concannon P, Iwasaki N, Schulze J, Baier LJ, Bogardus C, Groop L, Boerwinkle E, Hanis CL, Bell GI. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat Genet.* 2000; 26: 163-175.

16. Silander K, Mohlke KL, Scott LJ, Peck EC, Hollstein P, Skol AD, Jackson AU, Deloukas P, Hunt S, Stavrides G, Chines PS, Erdos MR, Narisu N, Conneely KN, Li C, Fingerlin TE, Dhanjal SK, Valle TT, Bergman RN, Tuomilehto J, Watanabe RM, Boehnke M, Collins FS. Genetic variation near the hepatocyte nuclear factor-4 alpha gene predicts susceptibility to type 2 diabetes. *Diabetes.* 2004; 53: 1141-1149.

17. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, Helgason A, Stefansson H, Emilsson V, Helgadottir A, Styrkarsdottir U, Magnusson KP, Walters GB, Palsdottir E, Jonsdottir T, Gudmundsdottir T, Gylfason A, Saemundsdottir J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Gudnason V, Sigurdsson G, Thorsteinsdottir U, Gulcher JR, Kong A, Stefansson K. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet.* 2006; 38: 320-323.

18. 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; .

19. Gerke P, Sellin L, Kretz O, Petraschka D, Zentgraf H, Benzing T, Walz G. NEPH2 is located at the glomerular slit diaphragm, interacts with nephrin and is cleaved from podocytes by metalloproteinases. *J Am Soc Nephrol.* 2005; 16: 1693-1702.

20. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science.* 1996; 273: 1516-1517.

21. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Ann Intern Med.* 1999; 130: 461-470.

22. 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.

23. Balding DJ. A tutorial on statistical methods for population association studies. *Nat Rev Genet.* 2006; 7: 781-791.

24. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst.* 2004; 96: 434-442.

Addendum to ARIC Manuscript Proposal # 1237

PC Reviewed: ___/___/07	**Status: _____**	**Priority: ____**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: **Addendum to MP #1237:** Association between genetic variants conferring risk for type 2 diabetes mellitus and incident chronic kidney disease

b. Abbreviated Title (Length 26 characters): Diabetes genes & incident CKD

2. Writing Group:

Writing group members: Anna Kottgen, W.H. Linda Kao, Kari E. North, James S. Pankow, Eric Boerwinkle, Josef Coresh

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]**

F**irst 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@jhsph.edu

Corresponding/senior author (if different from first author correspondence will be sent to both the first author & the corresponding author): **Josef Coresh**
Address: 2024 E. Monument St.
Welch Center for Prevention, Epidemiology, and Clinical Research
Baltimore, MD 21287

Phone: (410) 955-0495 Fax: (410) 955-4076
E-mail: coresh@jhu.edu

Additionally, we propose to investigate the single nucleotide polymorphism rs9939609 in the *FTO* gene (1, 2). **This SNP is being typed / will be typed in the ARIC Study as part of the regular contract work. The rs9939609 variant in the *FTO* gene has been detected in a genome-wide association study for susceptibility loci for type 2 diabetes** (1)**. Mice in which the *FTO* region is deleted exhibit a phenotype closely resembling the one observed in several inherited kidney disease syndromes**

(3)**. We therefore plan to study the association of rs9939609 with the kidney traits as outlined in manuscript proposal #1237.**

The rest of the proposal will be as outlined in MP #1237.

3. Timeline:

4. Rationale:

5. Main Hypothesis/Study Questions:

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).

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 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? ____ 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"?
____ Yes ____ No

10. 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

_____ 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)?

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)* _____ _____ _____)

*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. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science. 2007;316(5826):889-94.

2. Dina C, Meyre D, Gallina S, Durand E, Korner A, Jacobson P, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. Nat Genet. 2007;39(6):724-6.

3. Groop L. From fused toes in mice to human obesity. Nat Genet. 2007;39(6):706-7.