ARIC Manuscript Proposal #1952

PC Reviewed: 5/29/12	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: "Next Generation Sequencing to Identify Susceptibility Variants for Chronic Kidney Disease (CKD) and its quantitative traits"

b. Abbreviated Title (Length 26 characters): **Sequence analysis of CKD**

2. Writing Group:

Writing group members: Adrienne Tin, Lawrence Shimmin, Anna Kottgen, Jim Hixson, Eric Boerwinkle, Linda Kao, Josef Coresh (we invite other interested ARIC investigators)

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

F**irst author**: **Adrienne Tin**

Address:
Johns Hopkins Bloomberg School of Public Health
615 N Wolfe St, Room W6521
Baltimore, MD 21205

E-mail: atin@jhsph.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:
Johns Hopkins Bloomberg School of Public Health
615 N Wolfe St, Room W6513
Baltimore, MD 21205

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

3. Timeline: Data analysis to start immediately, completion of data analysis and drafting of the manuscript over the next year.

4. Rationale: Recent genome-wide association studies (GWAS) have identified multiple loci for serum uric acid levels and gout[1,2]. Many of the index GWAS SNPs identified are either in intronic or intergenic regions with unknown function. Moreover, together, the SNPs at these loci explain only a small proportion of the variance in serum urate, suggesting that additional variants, including rare variants, at these known loci, and new loci remain to be identified. Thus, sequencing approaches are necessary in order to further characterize variants at known loci and to identify novel, rare variants that may account for the "missing heritability."

Therefore, we propose to use next generation sequencing to identify additional variants and analyze their association with serum uric acid and gout. This will be a multi-layer project, beginning with the sequencing of exons, promoters, and flanking regions of candidate genes identified from previous GWAS of urate (Jim Hixon's approved ancillary proposal). Subsequently, we will analyze the exome chip and exome and whole genome data generated from CHARGE-S for association with serum urate and gout. Results from the aforementioned analyses will be combined with other cohorts from either the CHARGE Consortium or other collaborators from the Global Urate Genetics Consortium for meta-analysis.

5. Main Hypothesis/Study Questions:
Common and rare SNP are associated with serum urate levels and gout. In addition to the known GWAS loci, novel loci will be identified through sequencing.

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).

The design of the sequence study are case-control (Hixson's ancillary study) and case-cohort (CHARGE-S). In the case-control, cases were defined as those with self-reported gout and controls were those with serum urate $<5^{th}$ race-specific pentile and reported no gout.

We will exclude individuals without genetic consent, has no genotype data, or has no serum urate data.

A brief description of the statistical analysis plan is provided below:

- Single SNP (MAF>1%) analyses: we will perform linear (urate) or logistic (gout) regression for each outcome on additive coding of SNP genotype. Additional exact p-value or permutation p-value will be provided for rare variants (MAF<5%). Number of replicates in the permutation should >1/pval from regression.
- Secondary analyses of aggregated effects of multiple nonsynonymous SNPs in each gene.

- Weighted sum approach by Madsen and Browning (A Groupwise Association Test for Rare Mutations Using a Weighted Sum Statistic[3].
- Fixed threshold test (T1, T5) and variable threshold test by Price et al.[4].
- Software for all methods can be downloaded from http://genetics.bwh.harvard.edu/rare_variants
- Kernel-based association test which is more powerful than the above tests when a region contain protective, deleterious and null variants and can incorporate covariates by Wu et al.[5].

- General approach following up a SNP of interest. We will apply the following criteria to narrow down our list of SNPs for follow up: 1) non-syn coding SNPs; 2) SNPs observed in cases but not controls; 3) SNPs identified in controls but not cases; 4) SNPs with high statistical significance even in analyses of the sequenced individuals; 5) novel SNPs (not in any public databases); 6) SNPs associated with a splice site; 7) SNPs predicted to be functional; 8) SNPs that are in LD with the initial GWAS index SNP but have predicted functional significance; 9) SNPs in regulatory regions of the gene.
- When using CHARGE-S data, we will apply the appropriate sampling weight in our analysis of the continuous serum urate trait.

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 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? _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 ICTDER03 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)?

#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? _X Yes ____ No

11.b. If yes, is the proposal

X A. primarily the result of an ancillary study (list number* 2006.03, 2007.02**)**

___ 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.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.

1. Dehghan A, Kottgen A, Yang Q, et al. Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. *Lancet.* 2008;372(9654):1953-1961.
2. Kottgen A, Glazer NL, Dehghan A, et al. Multiple loci associated with indices of renal function and chronic kidney disease. *Nat.Genet.* 2009;41(6):712-717.
3. Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. *PLoS.Genet.* 2009;5(2):e1000384.
4. Price AL, Kryukov GV, de Bakker PI, et al. Pooled association tests for rare variants in exon-resequencing studies. *Am.J.Hum.Genet.* 2010;86(6):832-838.
5. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *American journal of human genetics.* Jul 15 2011;89(1):82-93.