ARIC Manuscript Proposal #1887

PC Reviewed: 01/10/12  Status: A  Priority: 2
SC Reviewed: _________  Status: _____  Priority: ____

PAGE Ms. Number: _____  Submission Date: _November 2011_____  [Approval Date: ___]

Title of Proposed Ms.: Genetic association of kidney traits in African American, European, Hispanic, and Asian individuals using the MetaboChip array: Discovery and Fine Mapping in the PAGE Consortium

I. INVESTIGATOR INFORMATION:

Name of Lead Author: Christina L. Wassel  |  Junior Investigator? Yes
Email Address: cwassel@ucsd.edu
Telephone Number: 858-822-7668

Names, affiliations and email address of PAGE Investigators proposed as co-authors:

<table>
<thead>
<tr>
<th>N, N</th>
<th>Affiliation in PAGE</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nora Franceschini</td>
<td>CALiCo-SHS</td>
<td><a href="mailto:noraf@unc.edu">noraf@unc.edu</a></td>
</tr>
<tr>
<td>Kari E North</td>
<td>CALiCo-ARIC</td>
<td><a href="mailto:Kari_north@unc.edu">Kari_north@unc.edu</a></td>
</tr>
<tr>
<td>Joachim H. Ix</td>
<td>CALiCo-CHS</td>
<td><a href="mailto:arnolda@u.washington.edu">arnolda@u.washington.edu</a></td>
</tr>
<tr>
<td>Myriam Fornage</td>
<td>CALiCo-CARDIA</td>
<td><a href="mailto:Myriam.Fornage@uth.tmc.edu">Myriam.Fornage@uth.tmc.edu</a></td>
</tr>
<tr>
<td>Sarah Pendergrass</td>
<td>EAGLE</td>
<td><a href="mailto:sarah.pendergrass@chgr.mc.vanderbilt.edu">sarah.pendergrass@chgr.mc.vanderbilt.edu</a></td>
</tr>
<tr>
<td>TBN</td>
<td>WHI</td>
<td></td>
</tr>
<tr>
<td>Steve Buyske</td>
<td>PAGE CC</td>
<td><a href="mailto:buyske@stat.rutgers.edu">buyske@stat.rutgers.edu</a></td>
</tr>
<tr>
<td>Anastasia Wise</td>
<td>NHGRI</td>
<td><a href="mailto:anastasia.wise@nih.gov">anastasia.wise@nih.gov</a></td>
</tr>
<tr>
<td>Linda Kao</td>
<td>CALiCo-ARIC</td>
<td><a href="mailto:wkao@jhsph.edu">wkao@jhsph.edu</a></td>
</tr>
</tbody>
</table>

PAGE ms. proposal template; 082409  1 of 7
Partner studies in PAGE not collaborating in this ms. proposal:

<table>
<thead>
<tr>
<th>Study</th>
<th>Contacted? Y/N</th>
<th>Declined? / Other?</th>
</tr>
</thead>
<tbody>
<tr>
<td>GenNet</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>MESA</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>HyperGEN</td>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>

Names, affiliations, email address of non-PAGE investigators proposed as co-authors: **None at this time, but will likely be added as we contact the studies**

II. SCIENTIFIC RATIONALE  (Please be specific and concise)

Chronic kidney disease (CKD), defined by a reduced kidney function (GFR) or presence of increased albuminuria, is an emerging public health problem with an estimated prevalence of 9.6% (or 19.2 million) of the adult US population\(^1\),\(^2\). CKD is a strong risk factor for cardiovascular disease\(^3\), is associated with increased morbidity and mortality\(^4\),\(^5\), increased health care resource use\(^6\) and a decreased quality of life\(^7\). Hypertension and type 2 diabetes are the most common causes of CKD and other known risk factors for kidney disease are metabolic syndrome\(^8\), hyperlipidemia\(^9\),\(^10\) and smoking\(^11\),\(^12\). CKD disproportionally affects minorities, who also more often progress to end-stage renal disease (ESRD) than those of European descent\(^1\),\(^13\). African-Americans have a disproportionate burden of CKD and albuminuria compared to non-Hispanic whites\(^1\), and Mexican Americans have similar prevalence of CKD and albuminuria compared to non-Hispanic whites\(^13\). However, little data exists on other Hispanic subgroups such as Puerto Ricans, Dominicans, and Cubans.

To date, a number of significant loci have been discovered and replicated for a variety of kidney related traits in European populations\(^14\)-\(^16\). In African-Americans, fewer loci have been identified and replicated, through both GWAS\(^17\) and admixture mapping\(^18\)-\(^20\) studies. However, in Hispanics, few genetic variants or loci have been discovered for kidney traits, and existing studies are primarily from candidate gene and linkage studies\(^21\)-\(^25\). Genetic ancestry, or admixture, has been significantly associated with albuminuria among Hispanics\(^26\), indicating that the ancestral background could be an important indicator of kidney function, and that there may be genetic variants not yet discovered that are important in renal function among Hispanic groups.

The Metabochip is a high-density custom array of 200K SNPs that captures DNA variations at regions identified in well-powered genome wide association studies (GWAS) for a variety of traits relevant to type 2 diabetes and cardiovascular diseases. The array includes new DNA variation emerging from the 1000 Genomes Projects, thus providing a fine-mapping in few target loci. In addition, it contains ancestry informative markers (AIMs) that may allow for better characterization of population stratification.

Using data from minorities with diverse genomic background, we expect to identify new variants in known loci and novel loci associated with kidney traits, especially in Hispanic groups, where little work has been done to date.

III. OBJECTIVES AND PLAN  (Please be specific and concise)

a. Study Questions/Hypotheses.

1. For loci represented in the MetaboChip array, are there novel associations with kidney traits (i.e. eGFR, albuminuria) in African American, Asian and Hispanic individuals (new loci)
2. For known loci associated with kidney traits and represented in the MetaboChip, are there one of more independent signals in African Americans, Asians and/or Hispanics (generalization and fine mapping).

b. Study populations, study design for each

All PAGE study populations with MetaboChip data and measured serum creatinine, urine creatinine, urine albumin/urine creatinine ratio, estimated GFR.

c. Variant/SNPs (Specify)

We will evaluate all the variants in the MetaboChip that passed quality control (QC) from a range of allele frequencies (0.05% or more allele frequency) for our discovery effort and fine mapping of known loci. QC will be performed at the PAGE CC.

d. Phenotype(s) (Specify)

The phenotypes to be used for analyses are:

1. Quantitative traits are estimated GFR (calculated from serum creatinine, age, race and sex by CKD-Epi equation\(^{27}\), urine albumin to urine creatinine ratio. The CKD-Epi equation has been previously validated in Hispanics by Stevens et al\(^{28}\).

2. Secondary phenotypes (binary traits) are chronic kidney disease (eGFR≤60 ml/min/1.73m\(^2\) vs eGFR>60 ml/min/1.73m\(^2\)), chronic kidney disease severity/categories (eGFR<45 ml/min/1.73m\(^2\), 45≤eGFR≤60 ml/min/1.73m\(^2\), 60≤eGFR≤90 ml/min/1.73m\(^2\), eGFR>90 ml/min/1.73m\(^2\)), albuminuria (≥30mg/g vs <30 mg/g), and albuminuria categories (<30 mg/g, 30-300 mg/g, >300 mg/g). Secondary analyses will also explore gender-specific cut points for albuminuria, i.e. <17 mg/g, 17-250 mg/g and > 250 mg/g for men, and <25 mg/g, 25-355 mg/g and > 355 mg/g for women.

e. Covariates (Specify)

Continuous age, sex, study center or region (if needed), and ancestry principal components. We will also adjust for local ancestry estimates if available. Additional levels of adjustment for traditional cardiovascular and kidney disease risk factors (i.e. smoking, blood pressure, lipids, body mass index, diabetes) may be performed to determine whether any novel loci identified for kidney traits are actually working through another closely associated trait or disease (i.e. mediation).

f. Main statistical analysis methods

Analyses will be stratified by self-reported race and within each PAGE study. We will use linear regression models for quantitative traits and logistic regression models for the binary traits of CKD and albuminuria. We will use additive genetic models adjusted for age, sex (except WHI), study center or region (if needed), and ancestry principal components. Additional levels of adjustment may be carried out as described above in the "Covariates" section. Urine albumin/creatinine ratio will be log transformed. Additionally, creatinine will be calibrated as necessary for eGFR calculations to ensure comparability across studies. Because individuals were recruited using a probability sampling in the SOL, all analyses of this study will use weights to account for identification and selection of community areas and random selection of households within those areas\(^{29}\). Study-specific association results will be combined across race-specific samples using a fixed effects inverse variance meta-analysis approach implemented in PLINK\(^{30}\).
For aim 1. Discovery. We will perform single variant test for all variants with minor allele frequency (MAF) >1% and use Bonferroni correction to declare a significant p-value threshold based on multiple testing. Although we acknowledge that power may be suboptimal in certain groups for this aim (i.e. not all studies will have albuminuria, Asians may comprise a smaller group), there is still potential to discover new loci in groups such as Hispanics, which are very understudied with regard to genetic variants for kidney disease phenotypes.

For aim 2. Generalization and fine mapping. We will perform comprehensive analyses of all known loci for kidney traits represented in the MetaboChip. We will use single variant test approach for variants with MAF>1% and adjust for the number of independent tests using a Monte Carlo approach that accounts for linkage disequilibrium (LD) between variants across the tested loci. We will perform conditional analyses using the most significant variant as covariate if there is suggestive evidence for secondary independent signals in these regions. For low frequency variants (0.1%<MAF<5%), we will explore novel methods for collapsing variants across regions including a recent method developed by Lin et al, as well as rare variants burden tests of Madsen and Browning, Li and Leal, Liu and Leal, and Pan and Shen. In order to adequately evaluate SNPs within the 0.1-5% MAF range for binary traits, CKD and albuminuria, we will use the methods of Li et al, which are appropriate for analysis of rare variants from GWAS and imputed GWAS data. Briefly, these methods include a weighted haplotype-based test, and an imputation-based test. These methods do not rely on the availability of external sequence data, but it can be incorporated, and also improve the power over traditional rare variant methods that are more appropriate for sequence data. For novel variants identified in known loci, we will further use bioinformatics tools for functional characterization of the variants.

We are currently working on possible replication groups for this study, including the Multi-Ethnic Study of Atherosclerosis (MESA), HyperGEN and GenNet for African American individuals, and will be contacting those studies in the near future. For Hispanic individuals, we plan to collaborate with MESA.

g. Ancestry information used? No __ Yes ___ X__ How is it used in the analyses?

We will use estimates of global ancestry (principal components) to adjust for population stratification. Because global ancestry may not fully account for population stratification in African Americans and Hispanics, we will also adjust for local ancestry, as available.

h. Anticipated date of draft manuscript to P&P: ___ 6-8 months after the data are available __. Analyses should be done in 1-2 months and remaining time is for replication of main findings in collaborating studies and for manuscript writing.

i. What manuscript proposals listed on www.pagesudy.org/index.php/manuscripts/ are most related to the work proposed here? Approved PAGE ms. numbers: ___

   – If any: Have the lead authors of these proposals been contacted for comments and/or collaboration? Yes __ X__ No ___

IV. SOURCE OF DATA TO BE USED (Provide rationale for any data whose relevance to this manuscript is not obvious): Check all that apply:

Aggregate/summary data to be generated by investigators of the study(ies) mentioned:

[ X ] EAGLE; [ X ] CALiCO; [ ] MEC; [ X ] WHI; [ ] CC; [ ] Other: __________________________
If CALiCo, specify [ X ] ARIC; [ X ] CARDIA; [ X ] CHS; [ ] SHS-Fam; [ X ] SHS-Cohort; [X] SOL

I, Christina Wassel, affirm that this proposal has been reviewed and approved by all listed investigators.

V. REFERENCES


