ARIC Manuscript Proposal #2428

PC Reviewed: 9/9/14  Status: A  Priority: 2
SC Reviewed: _________  Status: _____  Priority: _____

1.a. Full Title: Meta-analysis of 1000 Genomes imputed variants and COPD in the International COPD Genetics Consortium (ICGC)

b. Abbreviated Title (Length 26 characters): 1000G and COPD: the ICGC

2. Writing Group:

Analysis of 1000 Genomes imputed variants and COPD in ARIC will be led by Annah Wyss and Stephanie London. The meta-analysis of 1000 Genomes imputed variants and COPD in the International COPD Genetics Consortium (ICGC) will be led by Michael Cho (michael.cho@channing.harvard.edu) and Marike Boezen (h.m.boezen@umcg.nl). Since this is a consortium paper, a set number of authors per cohort will also be included. The current understanding is up to four authors per cohort. Laura Loehr was the lead author for ARIC on the earlier airflow obstruction manuscript but has declined to be involved in this one for lack of time. Thus we expect to be able to include two additional ARIC authors.

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

First author:

Annah Wyss
Address: NIEHS, PO Box 12233, MD A3-05. RTP NC 27709

Phone: 919-541-0420  Fax:
E-mail: annah.wyss@nih.gov

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):

Stephanie London
Address: NIEHS, PO Box 12233, MD A3-05. RTP NC 27709

Phone: 919-541-5772  Fax:
E-mail: london2@niehs.nih.gov

4. **Rationale:**

Chronic obstructive pulmonary disease (COPD) is a leading cause of morbidity and mortality (WHO The global burden of disease 2004). It is influenced by both environmental, particularly cigarette smoking, and genetic factors. A previous GWAS in subjects of European ethnicity, which included individuals from ARIC and other CHARGE and SpiroMeta Consortium cohorts, implicated 2 loci (CHRNA5/3 and HTR4) associated with airflow obstruction (Wilk et al. AJRCCM 2012).

A large international consortium (International COPD Genetics Consortium) has formed to perform a larger meta-analysis of COPD (or airflow obstruction in studies which did not perform bronchodilator testing needed for the definition of COPD). We wish to participate in this larger meta-analysis. This meta-analysis is based on 1000G imputation. Thus we will redo our previous GWAS in the ARIC whites using the 1000G imputed data as well as the slightly different case definition required by this project. The meta-analysis includes population based cohorts such as ARIC and case-control studies of COPD. The meta-analysis will be performed by Michael Cho at Harvard Medical School.

5. **Main Hypothesis/Study Questions:**

To identify novel loci for COPD in a large international meta-analysis of COPD using 1000G imputed data.

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).**

ICGC analysis plan is attached. Per agreement with Dr. Cho, we will not perform analysis C of COPD severity because of potential overlap with our ongoing 1000g analysis of quantitative spirometric traits.

7.a. Will the data be used for non-CVD analysis in this manuscript?

   _X_ Yes  ____ No

7.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?

   _X_ 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?

   _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.csc.c.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)?

We have an ongoing approved project for a CHARGE meta-analysis of spirometric traits which is number 2397. That plan does not include COPD/airflow obstruction.

This project follows up our earlier GWAS meta-analysis of airflow obstruction that resulted in Wilk et al. AJRCCM 2012 (PMC3480517). This paper was split off from an original approved manuscript proposal so both numbers 1357 and 1360 may be attached to this proposal.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data?  
_____ Yes  _X_ 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.csc.c.unc.edu/aric/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.c.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.
International COPD Genetics Consortium (ICGC)
Genetic Analysis Plan

Each participating cohort is requested to carry out the following genetic analyses and to send the results to Michael Cho (michael.cho@channing.harvard.edu) and Marike Boezen (h.m.boezen@umcg.nl) by September 1, 2014.

A. Genetic Collaborations and Validation

Rationale: Obtain GWAS information using standard phenotypes, a standard panel of imputed genotypes, and a standard approach for population stratification adjustment, which will facilitate collaborative projects. Phenotypes were selected to include measurements likely to be available in all study populations.

Please perform GWAS within your study cohort(s) using the following parameters:

1) Phenotypes:
   a. Case-Control Analysis of Moderate-to-Severe COPD:
      Cases: Pre-bronchodilator FEV₁ < 80% predicted and FEV₁/FVC < 0.7
      Controls: Pre-bronchodilator FEV₁ ≥ 80% predicted and FEV₁/FVC ≥ 0.7
      Note: analyses performed using different definitions of affection status (e.g. Wilk et al AJRCCM 2012) are acceptable.
   b. Case-Control Analysis of Severe COPD:
      Cases: Pre-bronchodilator FEV₁ < 50% predicted and FEV₁/FVC < 0.7
      Controls: Pre-bronchodilator FEV₁ ≥ 80% predicted and FEV₁/FVC ≥ 0.7
   c. COPD Severity: Pre-bronchodilator FEV₁ (in liters) adjusted for age, sex, height, ever smoking status, current smoking status, and pack-years of smoking within COPD cases (defined as pre-bronchodilator FEV₁ < 80% predicted and FEV₁/FVC <0.7).

2) Genotypes: We strongly suggest the inclusion of imputed genotypes (preferably using the 1000 Genomes project; for methods see under 4) to allow for comparison across different genotyping platforms. If imputation has not been performed, genotyping results alone will be acceptable; we will be prepared to offer technical assistance in completing imputation.

3) Association Analysis: Perform GWAS as follows: Use an additive genetic model and adjust for population stratification as optimal for your cohort. Adjust for covariates of age, ever smoking status, current smoking status, pack-years of smoking, and sex. For X chromosome SNPs, indicate method used for analysis.

4) Results:
   a. For each analysis, report:
      i. Cohort / phenotype
         1. Number of cases and controls
         2. Means and standard deviations for covariates (or % for sex and smoking status)
ii. Association analysis
   1. Genotyping platform
   2. Quality control parameters (cutoffs for missingness, HWE, duplicate or related samples, etc)
   3. Population stratification adjustment (outlier subject removal, number of principal components, lambdaGC)
   4. Imputation program (e.g. MaCH / minimac, IMPUTE) and reference population used (e.g. 1000 Genomes European, Cosmopolitan, other)
   5. Software used (e.g. GenABEL, plink)

b. For each SNP association, report
   i. rs number or 1000 Genomes identifier (chromosome: position, referencing hg19/b37)
   ii. strand orientation (preferably +)
   iii. effect allele
   iv. other allele
   v. frequency of the effect allele
   vi. beta coefficient (log odds ratio)
   vii. standard error
   viii. Imputation accuracy metric (e.g. r-square, info)