ARIC Manuscript Proposal #2803

PC Reviewed: 7/12/16  Status: A  Priority: 2
SC Reviewed: ________  Status: _____  Priority: ____

1.a. Full Title: RHOGEN2

b. Abbreviated Title (Length 26 characters): Chromosome Y and multiple traits

2. Writing Group:
ARIC co-authors: Nora Franceschini, Alanna C. Morrison, Eric Boerwinkle, Kari E North, Stephanie London, Dan Arking, Aravinda Chakravarti, Misa Graf, Jim Pankow

Writing group members from RhoGen2: Jim Wilson, Tõnu Esko, Peter Joshi, Ilaria Gandin

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

First author:
Jim Wilson, DPhil FRCPE
Reader in Population and Disease Genetics
Usher Institute for Population Health Sciences and Informatics
University of Edinburgh
Teviot Place
Edinburgh
EH8 9AG
Scotland

Email: Jim Wilson jim.wilson@ed.ac.uk

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: Nora Franceschini
Address: 137 E Franklin St # 306, Chapel Hill, NC 27514

Phone: 919-966-1305  Fax: 919-9669800
E-mail: noraf@unc.edu
3. **Timeline:** June/July 2016 for analysis, June/July of 2017 for submission of paper to a journal.

4. **Rationale:**
This proposal is an extension of the already approved RHOGEN proposal (Ms 2324) and aims to study investigate the effects of runs of homozygosity on human quantitative traits.

5. **Main Hypothesis/Study Questions:** Runs of homozgosity, as a genomic measure of inbreeding, will negatively affect health-related indicators and quantitative traits.

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).**

   This consortium will be run in a very similar manner to the ROHgen consortium, for which ARIC participated (Joshi et al, Nature 523: 459-62). Like in ROHgen, we will consider a broad spread of quantitative traits. A pipeline of scripts will be distributed which will perform all the analyses. Extensive beta testing in Dr. Wilson’s lab and others has already been conducted.

   Meta-analysis of all available resources which can provide both relevant genotypic and phenotypic data.

   For the 2016 round of analysis, we will focus on the following quantitative traits of public health importance, in models adjusted for .

   We have written a pipeline of scripts which can run the entire analysis given the genotypes, traits, covariates, plink, R and unix. A cookbook accompanies the scripts describing the methods used phenotype modelling and how to run the pipeline. Participating groups will share descriptions of the initial work they have carried out including genotyping and data quality control; will report summary statistics (e.g. means, standard deviations, maxima and minima of traits, covariates and ROH distributions); and will share the cohort-specific analysis results (e.g. effect sizes, standard errors and P values). Meta-analysis will then be performed by central ROHgen2 analysts. There may be follow up analyses in addition to the primary work.

   ROHgen2 will analyse 41 quantitative traits, including the 16 previously analyzed in ROHgen1 (see below for full list). The new traits include several components of fitness where theory and model organism data suggest an effect of inbreeding depression.

   **Effect of different ROH lengths**
   We will use multivariate linear models to investigate the effect of different lengths of ROH on each quantitative trait. In addition, we will compare the effect of ROH to other measures of inbreeding based on individual SNP homozygosity. This may provide insight into the nature of the casual variants responsible for the signals observed in ROHgen1.
Regional ROH effects
ROHgen2 tests the effect of ROH burden in each of approximately 1000 3 Mb-wide windows along the genome. This aims to identify regions of the genome where homozygosity has a large effect on any particular trait.

Sex-specific effects
In ROHgen2 we will quantify the effect of ROH on each trait for males and females separately.

Trait variance effects
For a purely additive trait, theory predicts a directly proportional increase in genetic variance with inbreeding coefficient. If however, some of the genetic variance is caused by rare recessive alleles then the increase in genetic variance may be significantly greater. ROHgen2 estimates the increase in trait variance with f_roh and may permit the quantification of dominance even in the absence of directionality.

Coding will be provided if requested.

7.a. Will the data be used for non-CVD analysis in this manuscript? __x__ 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? ___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.cscu.unc.edu/ARIC/search.php

   ___x__ Yes  _______ No Manuscript for RhoGen proposal already published (MS 2324)
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)? None

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.cscc.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.cscc.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.

13. Per Data Use Agreement Addendum for the Use of Linked ARIC CMS Data, approved manuscripts using linked ARIC CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes __x__ No.