ARIC Manuscript Proposal # 3340

PC Reviewed: 2/12/19   Status: _____   Priority: 2
SC Reviewed: _________   Status: _____   Priority: ____

1.a. Full Title: Using ARIC controls for Whole Genome Sequence Analysis of Intracerebral Hemorrhage

b. Abbreviated Title (Length 26 characters): WGS of ICH

2. Writing Group:

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

First author: Christopher D. Anderson, MD, MMSc
Address: MGH Center for Genomic Medicine
185 Cambridge Street, CPZN 6818
Boston, MA 02114

Phone: 312-498-9443       Fax: cdanderson@partners.org
E-mail: cdanderson@partners.org

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: Myriam Fornage, PhD
Address: University of Texas Health Science Center at Houston

Phone: 713-500-2463       Fax: myriam.fornage@uth.tmc.edu
E-mail: myriam.fornage@uth.tmc.edu

3. Timeline:
2/30/19: Identify ARIC black and white referents for ERICH black and white ICH cases
2/2019: Identify SoL Hispanic referents for ERICH Hispanic ICH cases
3/2019: Generate joint call-set of ERICH ICH cases with ARIC&SoL referents
5/2019: Perform case/control analyses of ICH risk and prepare manuscript
4. **Rationale:** Stroke is a leading cause of death and disability worldwide. Intracerebral hemorrhage (ICH) is a common stroke subtype and is associated with high morbidity and mortality and a substantial risk of recurrence. To date, few genetic loci have been identified for ICH. Whole genome sequencing (WGS) analysis provides an opportunity to examine the full spectrum of contribution of genetic variants to ICH. To this end, a collaborative effort between the ISGC, NHGRI, and the Baylor Human Genome Sequencing Center has provided resources for WGS analysis of ICH. In order to maximize power of identifying novel ICH genes, these resources were devoted almost exclusively to cases (N~3000), while control data would be obtained from previously sequenced cohort data. The ARIC study represents a suitable source of controls due to its large, well-characterized population with WGS generated using the same protocols, including common variant calling, the ISGC ICH cases. Additional controls of Hispanic ethnicity will be sought concurrently from SoL, which has similar advantages to ARIC as a suitable referent population.

5. **Main Hypothesis/Study Questions:**
We propose to use the ARIC WGS data as control data to perform WGS analysis of ICH risk. An important consideration in designing a case-control study is that cases and controls come from comparable underlying populations and data be generated with the same methods so that any differences observed between the groups is not due to confounders such as technical artefacts or population stratification. As previously stated, ICH cases and ARIC data were generated under the same variant calling protocol, using the same QC procedures at the same sequencing center. Issues of population stratification will be addressed as follow:
A small panel of 15 genetic variants have already been genotyped in all ICH cases, which will allow classification by ancestry using principal components analysis among the three represented recruitment groups of white European, African American, and American Hispanic. Based on this data, ICH-free referents will be chosen from ARIC and SoL at 3:1 within each of these population groups. Given the heterogeneity of African American and Hispanic ancestral backgrounds, the use of this genetic ancestry data will allow accurate representation of case ancestries among the selected controls.

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?** __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?** _____ 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.cscc.unc.edu/aric/mantrack/maintain/search/dtSearch.html

___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)?
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 https://www2.cscc.unc.edu/aric/approved-ancillary-studies

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.