ARIC Manuscript Proposal #2124

PC Reviewed: 5/14/13  Status: A  Priority: 2
SC Reviewed: _________  Status: _____  Priority: ____

Candidate Gene Association Resource (CARe) Project
Database Application

NHLBI CARe Project Number: TBD

Submission Date:

All proposals must be submitted by email to the CARe DAC COORDINATOR at care-proposals@broad.mit.edu

All sections of this application must be completed. Incomplete applications will be returned.

Sections I-IV should be no more than 3 pages in length, not including references.

I. INVESTIGATOR INFORMATION:

Title of Proposed Project: Meta Analysis of Genome Wide Association Data for Type 2 Diabetes Quantitative Traits in African Americans including the Candidate Gene Association Resource (CARe) Project: Evaluation of Gene x Gene and Gene x Environment interactions

Name of proposed *Lead Author: TBD by the consortium writing group
Email Address: Telephone Number: Fax Number: Mailing Address:

Name of proposed *Senior Author: TBD by the consortium writing group. Initially Senior and ultimately responsible for the project: James B Meigs MD MPH
Email Address: jmeigs@partners.org Telephone Number: 617-724-3203 Fax Number: 617-724-3544 Mailing Address: James B Meigs MD MPH MGH GMD 50 Staniford St 9th Floor Boston, MA 02114

Institution/Company of Principal Investigator: Massachusetts General Hospital

Name of Lead Analysts / Statisticians: Josée Dupuis, PhD With: Eric Kolaczyk PhD, Ching-Ti Liu PhD, Denis Rybin, MS, Han Chen, MS, Chen Lu, MS

Josée Dupuis PhD:
Email Address: dupuis@bu.edu Telephone Number: (617) 638-5880 Fax Number: 617-638-5299
II. SCIENTIFIC RATIONALE (~250 WORDS)

A. Please provide an abstract describing the rationale and design of the proposed research project. The abstract must include major hypotheses, an outline of the research methods and analytical approach, and phenotypes to be studied. It should state clearly the objectives of the proposed project and provide the background rationale that would justify them. It should also address why the CARe database is appropriate for answering the research question. (Abstracts for approved projects will be posted, with the name of the Principal Investigator, on the NHLBI/NCBI CARe website.)

Over 2.8 million African Americans have type 2 diabetes mellitus (T2D). This represents approximately 13% of the African American population and a significant proportion of the 20 million Americans believed to be living with diabetes. On average, an African American (AA) individual is twice as likely to have T2D as a European Ancestry (EA) peer. AA individuals are also more likely to have sub-diabetic hyperglycemia compared with EA individuals, and have higher HbA1c levels at diagnosis, even
accounting for clinical characteristics and access to screening. Thus, sub-diabetic levels of diabetes-related quantitative traits like fasting glucose, fasting insulin and HbA1c are of interest for analysis to understand disparities in T2D physiology, pathogenesis and disease detection. Important related quantitative traits needed for statistical control in analyses of these variables or for alternate diabetes-related traits include age, sex, BMI, waist circumference, fasting triglycerides, total and HDL cholesterol, blood pressure and when available, family history of diabetes.

Meta-analyses of genome wide association studies (GWAS) of T2D and diabetes-related quantitative traits have had a revolutionary impact on our understanding of genetic contributions to diabetes susceptibility. To date published studies have largely examined EA populations, with a few GWAS conducted in AA populations. However, associated variants discovered to date only accounts for a fraction of the total heritability, in both EA and AA populations. A possible source of the unexplained heritability may reside in gene x gene and gene x environment interaction effects.

In a prior approved CARE manuscript proposal, we proposed to conduct a meta-analysis of GWAS for diabetes-related quantitative traits in the CARe and other samples. In addition to CARe, other GWAS studies in AA are underway for meta-analysis now allowing consortium formation for the purposes of the research. Many of these samples have been genotyped with the Affy 6.0 chip, thus further facilitating imputation development, in silico comparisons and meta analyses. Diabetes diagnosis is widely available in the AA subjects in CARe and relatively easy to harmonize between studies, so that a non-diabetic sample can easily be defined for quantitative traits analysis.

In this new manuscript proposal, we propose to investigate gene x gene and gene x environment interactions as the next step because T2D is no doubt underlain by complex genetic pathways and interactions. We propose to apply a novel network guided approach for this purpose. The novel method not only evaluates SNP effects, but also allow for gene x gene (GxG) and gene x environment (GxE) interaction effects to expand the public health understanding and impact of modern T2D related genetics to race/ethnic groups so far under-represented in T2D genetics research studies yet disproportionately affected by T2D and its morbid consequences.

We have developed and implemented network-based, penalized regression methods for genome-wide association analysis, aimed at detecting GxE and GxG interactions. Our proposed methods are extensions of the Lasso paradigm for penalized regression with high-dimensional variable selection, where the usual penalty will be replaced by one that accounts for GxE and GxG pairwise interactions and gene grouping (e.g., in pathways). This novel approach will be applied to glycemic traits that we have been analyzing under this CARe proposal (e.g., fasting glucose and fasting insulin) to detect genes whose effects are modified by the adipose environment (e.g., as measured by BMI). We are proposing to first look for such interactions using data from the 5 African American CARe cohorts already studied in, for example, Liu et al. Transferability and fine-mapping of glucose and insulin quantitative trait loci across populations: CARe, the Candidate Gene Association Resource. Diabetologia. 2012 Nov;55(11):2970-84. PMID: 22893027.

We have developed a two-stage meta-analysis strategy for integrating analyses across CARe cohorts, ultimately allowing for substantially increased power to detect interaction. We will perform a multi-SNP analysis rather than taking a one SNP at a time or pairs of SNPs when investigating interactions. We will seek replication of interesting findings in other African American cohorts if warranted, or publish results along with the methods development if there is not sufficient evidence that including interaction enhance our ability to detect important genes for these glycemic traits.

III. PRIOR EXPERIENCE OF THE PI AND ASSOCIATES (~250 WORDS)
Please describe the experience and expertise of your team to complete the research project.

The Diabetes phenotype writing group is composed of individuals, listed in this application, with extensive expertise and experience in the genetics, epidemiology and physiology of T2D, related phenotypes and related quantitative traits. In addition, these investigators have worked with the individual level cohort data that will be used for these analyses and have actively participated in GWAS...
studies, including the MAGIC analyses of diabetes-related quantitative traits in EA cohorts. The proposed senior author (Dr. Meigs) and colleagues are either the PI or participants in all the cohorts listed herein.

IV. DATA REQUESTED FOR THE PROPOSED ANALYSES (Provide rationale for any requested data whose relevance to these analyses is not obvious):

Genotype data (check all that apply):

[ ] CARe IBC Candidate Gene array

[ ] Affymetrix 6.0

[X] Both IBC array and Affymetrix 6.0

*The principal and currently only planned analysis will be with genotyped and imputed SNPs from the Affymetrix 6.0 array. However, if we find loci of interest after staged GWAS with meta-analysis that have been mapped on the IBC array, then we would like to have this data at hand to immediately pursue fine-mapping and integration with other data. We have no specific plan at this time for IBC array analysis.*

[ ] Other (please specify; e.g. SNPs within candidate genes, specific regions, etc.)

Specification of “other” genotype data: __________ Ancestry informative marker data

Phenotype data:

From each CARe cohort we request the following traits (if available, as already defined by the Diabetes Phenotype Group), with values from the exam where quantitative traits in this analysis were collected (when possible):

1. Sex
2. Age
3. BMI (derived from weight and height)
4. Waist circumference
5. Diabetes diagnosis (yes/no)
6. Hypertension diagnosis (yes/no)
7. Medication: diabetes oral medication (yes/no)
8. Medication: diabetes insulin therapy (yes/no)
9. Medication: hypertension medication (yes/no)
10. Medication: cholesterol medication (yes/no)
11. Family history of diabetes (yes/no)
12. Fasting glucose (mmol/l)
13. Fasting insulin (pmol/l)
14. HOMA-IR (derived from fasting glucose and insulin)
15. HOMA-B (derived from fasting glucose and insulin)
16. McAuley Index: glucose disposal corrected for FFM - [http://care.diabetesjournals.org/content/24/3/460.full](http://care.diabetesjournals.org/content/24/3/460.full)
17. HbA1c (%)
18. Fasting triglycerides
19. Total cholesterol
20. HDL cholesterol
21. LDL cholesterol
V. ORGANIZATION APPLYING FOR DATA ACCESS:

Name of Applicant Organization: Boston University School of Public Health

Address of Applicant Organization:
715 Albany Street, Talbot Building
Boston, MA 02118

Name and Title of Applicant Organization’s Authorized Institutional Business Official:
Jane F. Kinsel, Ph.D. MBA ; Director, Office of Sponsored Programs.

Name of Each Institution/Company whose investigators will receive access to the requested data: TBN

I, James B. Meigs, PI for the proposed project, request the following data:

a. [x] Yes  [ ] No:  Genotype Data
b. [x] Yes  [ ] No:  Phenotype Data

Please answer the following questions:

a. [ ] Yes [x] No:  This research using the CARe database may be used for development of a commercial product or for commercial purposes.

b. [ ] Yes [x] No:  Data accessed through this application will be used by or shared with individuals from a for-profit company.

NOTE: Some participants in the CARe Study have not provided consent for use of their CARe data for development of a commercial product or to be accessed by a for-profit entity. For each approved project, a dataset will be generated by a computer algorithm that incorporates the consent options of each of the CARe Cohorts, such that data access will only be provided as specified by the informed consent document of each study participant.

VI. ANNUAL REPORTING OF RESULTS

Continued access to the CARe database after one year will only be permitted if an annual report is submitted to the DAC Coordinator (care-dac@nhlbi.nih.gov) that describes the product of your research using the CARe database, and includes a listing of presentations and publications resulting from that research. Please complete the following:

[x] Yes  [ ] No:  I understand that use of the CARe dataset includes the timely completion of an annual report.

Complete results of CARe analyses are being web-posted by the National Center for Biotechnology Information (NCBI). Please complete the following:

[x] Yes  [ ] No:  I am willing to provide my completed analysis results for such web-posting.

VII. ADDITIONAL DOCUMENTS:
Return by email to the DAC Coordinator: care-proposals@broad.mit.edu

(Signed or other non-electronic documents can be faxed.)

Please include:

(1) Application Form

(2) Supporting documentation, including current human studies training certification for all key personnel*

(3) IRB Approval (Note: Full board or expedited review is required; “exempt” status is not acceptable.)*

(4) Data Distribution Agreement*

*Can be submitted after DAC Committee approval

Applications can be submitted at any time. Applicants will be notified of the DAC Committee decision within four weeks.

Upon approval, a Data Recipient at each institution that will receive data will be required to submit a signed Data Distribution Agreement to the CARe DAC Coordinator. The CARe Data Distribution Agreement can be obtained from Website to be added. The Data Recipient(s) will be responsible for obtaining signatures on behalf of the Recipient Entities and returning the signed Data Distribution Agreement to the CARe DAC Coordinator before access can be provided to the CARe database.

VII. ADDITIONAL INFORMATION (Required only for “CARe Investigators”)

Name and CARe Cohort Affiliation(s) of proposed Lead Author(s): TBD

Names and CARe Cohort Affiliation(s) of proposed Senior Author(s): James B. Meigs, FHS

Table of CARe Cohort Representation: Please see the Table in: “Name(s) of all other Professional Participants.”

Reason(s) for CARe cohorts not represented: CARe cohorts are not represented in this proposal if they have not been phenotyped for diabetes-related quantitative traits.