ARIC Manuscript Proposal # 1366

PC Reviewed: 05/13/08  Status: A  Priority: 2  
SC Reviewed: _________  Status: _____  Priority: ____

1.a. Full Title: Transcription factor 7-like 2 (TCF7L2) variants, Diabetes, Hemoglobin A1c, and fasting glucose: The Atherosclerosis Risk in Communities (ARIC) MRI Study

b. Abbreviated Title (Length 26 characters): TCF7L2 and diabetes traits

2. Writing Group:
   Writing group members: Suzette J. Bielinski, James S. Pankow, Weihong Tang, Aaron R. Folsom, Kari E. North, Eric Boerwinkle

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

First author: Suzette J. Bielinski  
Address: University of Minnesota  
1300 South Second Street, Suite 300  
Minneapolis, Minnesota 55116  
Phone: 612-624-1899  Fax: 612-624-0315  
E-mail: suzetteb@umn.edu

Corresponding/senior author (if different from first author correspondence will be sent to both the first author & the corresponding author):

Aaron Folsom  
Address: University of Minnesota  
1300 South Second Street, Suite 300  
Minneapolis, Minnesota 55116  
Phone: 612-626-8862  Fax: 612-624-0315  
E-mail: folsom@epi.umn.edu

3. Timeline:  
   Starting Analyses: July 1, 2008  
   First Draft: September 15, 2008  
   Submission for Publication: November 2008
4. **Rationale:**
Diabetes is a major risk factor for cardiovascular disease (CVD) increasing the risk of stroke, angina, MI, and coronary heart disease\(^1\). Transcription factor 7-like 2 (TNF7L2) polymorphisms have been associated with diabetes in numerous studies\(^2\-^4\). The T-allele of rs7903146 is a common variant in the population and is hypothesized to be either the risk variant or the closest correlate discovered to date\(^5\). In the whole ARIC population, only five TCF7L2 SNPs were genotyped including rs7903146. These SNPs were studied in relation to diabetes traits and incident cardiovascular disease. In the ARIC MRI subset the number of TCF7L2 SNPs was expanded to over 50, allowing for a more thorough investigation of the variation in this gene. In addition, hemoglobin A1c (HbA1c) was measured in this group. Therefore, we propose to study the relationship of these SNPs and diabetes, HbA1c, and fasting glucose in the ARIC MRI cohort.

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

Are any of the newly genotype variants in the TCF7L2 gene associated with diabetes, HbA1c, and fasting glucose?

6. **Data (variables, time window, source, inclusions/exclusions):**

**Design:** Cross Sectional study  
**Outcome:** Prevalent diabetes, HbA1c, fasting glucose  
**Exposure:** 57-TCF7L2 SNPs genotyped in ARIC MRI  
**Covariates** include, but are not limited to, traditional risk factors including age, sex, race, adiposity measures, lipid levels, blood pressure medication use, smoking status and amount, and physical activity.

**Analysis Plan**
1. Hardy Weinberg equilibrium among genotypes will be calculated using the chi-square test on race-specific datasets.
2. SNP genotypes will be modeled as categorical variables, defined as having 0, 1, or 2 copies of the minor allele. If appropriate given the results, an additive or dominant genetic model may be used.
3. Logistic regression methods will be used to assess associations with prevalent diabetes. Associations between individual SNPs and levels of HbA1c and fasting glucose will be assessed with linear regression methods following exclusion of those study subjects taking diabetes medications.
4. Because systematic differences in ancestry can produce spurious associations, all analyses will be stratified by race to account for systematic allele frequency differences between racial groups. However, we will also need to account for population substructure within racial groups. While there are a number of methods available to the analyst, we can take advantage of GWAS data most effectively using the principal components analysis method developed by Price and colleagues (Price, Patterson et al. 2006), implemented in the software EIGENSOFT. This method explicitly models ancestry and has higher power to detect true associations than other methods. Principal component factor scores will be incorporated into each model as covariates (number of actual covariates to be determined) to account for population
stratification in each of the samples. Appropriate weights will be used to account for the sample selection used for the ARIC MRI Study.

5. Permutation methods will be used to estimate an empirical p-value to account for multiple testing.

7.a. Will the data be used for non-CVD analysis in this manuscript?
   ____ Yes  ____X__ No

   b. If Yes, is the author aware that the file ICTDER02 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 ICTDER02 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 ICTDER02 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/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)?

   Manuscript #1141 Transcription factor 7-like 2 (TCF7L2) gene and type 2 diabetes

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/

12. Manuscript preparation is expected to be completed in one to three years. If a
The 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.
References


