ARIC Manuscript Proposal # 1227

PC Reviewed: _2_/_13/07	**Status: __A__**	**Priority: __2_**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: TCF7L2 variants and cancer risk

b. Abbreviated Title (Length 26 characters): TCF7L2 variants and cancer

2. Writing Group:

Writing group members: Aaron Folsom, James Peacock, Sue Bielinski, James Pankow, Gerardo Heiss. We also propose Eric Boerwinkle but he has not yet replied to our request.

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

F**irst author**: Aaron Folsom
Address:

Phone: 612-626-8862 Fax: 612-624-0315
E-mail: folsom@epi.umn.edu

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

Phone: Fax:
E-mail:

3. Timeline:

Analysis will begin once we have the TCF7L2 snp data

4. Rationale:

Diabetes is a risk factor for several neoplasms, including endometrial cancer, colon cancer, pancreatic cancer. The reasons for increased risk are not fully known. Obesity is one possible link, but the association of diabetes with neoplasms often is independent of BMI. Increased levels of IGF-1 or other growth factors also seem plausible. Prostate

cancer risk is reduced in men with diabetes, probably due to low testosterone in diabetic men. Of course, the diabetes-cancer association may not be causal and studies employing a "Mendelian randomization" design could be helpful.

Transcription factor 7-like 2 (TCF7L2) microsatellite DG10S478, and correlated T allele of SNP rs12255372 and rs7903146, are common and numerous studies have linked these variations with increased risk of type 2 diabetes. A case control study of diabetes in three diverse samples, Danish, Icelandic, and West African, concluded that the T-allele of rs7903146 is either the risk variant or the closest known correlate (1) The mechanisms for the diabetes association are still being elucidated. However, TCF7L2 is involved the Wnt/B-catenin pathway, which has oncogenic effects (1). It therefore seems possible that TCF7L2 variation could affect risk of diabetes-related cancers. On the other hand, the absence of an association between TCF7L2 could reaffirm that diabetes has no causal link to cancer. Only one study on this topic was found in Pub Med. A case control study of familial breast cancer (without BRCA 1 or 2) found TCF7L2 rs12255372 statistically significantly associated OR=1.19, 95% CI=1.01-1.42 and a dose response with number of alleles (2).

The ARIC DNA lab measured 5 SNPs for TCF7L2, including rs7903146. We propose to study their association with common cancers in ARIC. As part of an ancillary study, ARIC cohort members' cancers through 2000 were identified. Numbers of first primaries were: Forsyth (n=500), Mpls (n=502), Wash Co (n=468), and Jackson (n=425). There were roughly equal numbers of breast, prostate, colon, and lung cancer.

(1) Helgason et al. Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. Nature Genetics 2007: Jan 7 [Epub ahead of print]

(2) Burwinkel B, et al. Transcription factor 7-like 2 (TCF7L2) variant is associated with familial breast cancer risk. BMC Cancer 2006; 6:268.

5. Main Hypothesis/Study Questions:
TCF7L2 variation is related to risk of total cancer and several common sites: colon, breast, and prostate. (Too few cases of pancreatic and endometrial cancer.)

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).

Dependent variable: total cancer incidence; look at subtypes: colon, prostate and breast

Independent variable: TCF7L2 SNPs (rs12255372, rs7901695, rs11196205, rs7903146, and rs7895340)

Covariates: should be little confounding by other factors, but will consider main risk factors: age, sex, smoking, BMI, HRT, drinking status, etc.

Analysis:

Hardy Weinberg equilibrium among genotypes will be calculated using the chi-square test on race-specific datasets.

An additive genetic model will be assumed unless indicated otherwise by the results. Therefore, genotypes will be coded as 0 (0 copies of candidate allele), 1 (1 copy), or 2 (2 copies). If appropriate given the results, a dominant model combining homozygotes and heterozygotes will be used.

Will verify no association between cancer risk factors and TCF7L2 variants, in which case, we mainly will do simple proportional hazards models with TCF72 as the predictor of incident cancer. Analysis will also stratify on diabetes status, if numbers permit, so that there would be four risk groups: variant positive (yes/no) crossed by diabetes (yes/no).

If TCF7L2 is associated with cancer, then there may be some later analyses that adjust for diabetes in an attempt to tease out indirect effects (through diabetes) versus direct effects of the gene on cancer risk. If so, such analyses will need to include all potential confounders of the association between diabetes (the intermediate variables) and CVD, such as lifestyle factors, even if those potential confounders are themselves not associated with TCF7L2.

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 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? __x__ 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)?

1141 is on TCF7L2 and diabetes. We have asked Dr. Heiss to identify appropriate coauthors for this paper.

11. a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? __x__ Yes ____ No

11.b. If yes, is the proposal

_xx__ A. primarily the result of an ancillary study (list number* _1995.04__)

___ B. primiarly 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 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.