ARIC Manuscript Proposal #1536

PC Reviewed: 7/14/09	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: GWAS-by-Physical Activity Interaction for Type 2 Diabetes Related Traits for Meta-Analysis of Glucose and Insulin related traits Consortium (MAGIC)

b. Abbreviated Title (Length 26 characters): MAGIC: GWAS x PhysActivity

2. Writing Group:
Writing group members:

Audrey Chu
Linda Kao
James Pankow
Eric Boerwinkle
Liz Selvin
Man Li
Other MAGIC contributors and ARIC GWAS diabetes working group members

At the time of submission of this manuscript proposal we have committed to full participation in a manuscript as part of an international consortium of over a dozen epidemiologic cohorts investigating diabetes-related traits, called the Meta-Analysis of Glucose and Insulin Consortium (MAGIC). We will be providing meta-analysis results for gene-physical activity interactions and association with fasting glucose and insulin, 2-hour glucose, and hemoglobin A1C from the ARIC study.

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

First author: Audrey Chu
Address: 615 N. Wolfe Street
Rm W6507
Baltimore, MD 21205

Phone: 310-387-5558 Fax:
E-mail: achu@jhsph.edu

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

Johns Hopkins School of Public Health
615 N. Wolfe Street
Rm. W6513
Baltimore, MD 21205

Phone: 410-614-0945 Fax: 410-955-0863
E-mail: wkao@jhsph.edu

3. Timeline:

Will begin analysis immediately (May 2009).

4. Rationale: A total of 18 genetic variants have been identified in association with type 2 diabetes (1-6) and diabetes-related traits (7-9) using candidate gene studies and genome-wide association studies (GWAS). ARIC has begun participation in the Meta-Analysis of Glucose and Insulin related traits Consortium (MAGIC), an international consortium formed to analyze diabetes-related traits in the context of large-scale GWAS. Physical activity has been long implicated as a modifiable risk factor for diabetes as well as many diabetes-related traits. Examination of gene-physical activity interactions with diabetes-associated quantitative traits may discover new loci which induce risk of type 2 diabetes, provide insight into the pathophysiology of diabetes and identify a subgroup of individuals at higher risk of developing diabetes which could be targeted for intensive interventions. Currently, there are no MAGIC investigations of gene-physical activity interactions and diabetes-related traits. We would like to systematically investigate the following:

I. The interaction between SNP and physical activity in association with fasting glucose and insulin, HOMA-IR and HOMA-%β, hemoglobin A1c and post challenge 2hr glucose levels on a genome-wide level.

II. How multiple loci of proven small effect in the overall population interact with physical activity by investigating the effect modification of a genetic score scale and fasting glucose and insulin, HOMA-IR and HOMA-%β, hemoglobin A1c (HbA1c) and post challenge 2hr glucose levels by physical activity.

5. Main Hypothesis/Study Questions:

I. We propose to study physical activity X SNP interactions with diabetes-related traits using a two-tiered approach:

A. To study the interaction between established diabetes and diabetes-related trait loci and physical activity with diabetes related-traits.

B. To study the interaction of ~ 2.5 million genotyped and imputed SNPs from the Affymetrix 6.0 array and physical activity with diabetes-related traits.

II. We propose to investigate physical activity X genetic score interactions with diabetes-related phenotypes.

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).

All ARIC GWAS analyses will be performed by ARIC authors and results will be sent to MAGIC collaborators for the full meta-analysis. The following sections apply to the ARIC analyses only.

Participants who did not consent to genetic research and those who self-report race other than "white" will be excluded. If the last round of genotyping data becomes available before the analyses begin, we will include all white individuals who meet the inclusion criteria. However, if this data is not available, we will perform analyses with the currently available set of genotypes. A cleaned dataset (with relatives, genetic outliers, sex inconsistencies and samples with poor DNA quality removed) will be used for this analysis. Our final dataset will include ~.7 million genotyped and 1.8 million imputed SNPs.

Fasting glucose and insulin measures will be derived from visit 1. HOMA-IR and HOMA-%β will be estimated using fasting glucose and insulin from visit 1. Those with prevalent diabetes at visit 1 will be excluded from the analyses of these traits. Additionally, subjects who did not report fasting for 8 hours or more will be excluded from these analyses. Measures of HbA1C will be taken from ARIC visit 2 (visit 1 and visit 2 diabetic participants will be excluded). 2-hour post-challenge glucose will be taken from the 75g oral-glucose tolerance test administered at ARIC visit 4. Those with documented diabetes at any of the ARIC visits will be excluded from this analysis, as will subjects who did not fast for more than 10 hours and subjects with technical problems during the OGTT (eg. problem with venipuncture, less than 50% of glucose solution ingested, vomiting, length of OGTT <110 min or >130min).

Physical activity measures will be taken from baseline (visit 1) for analysis of fasting glucose and insulin, HOMA-IR, HOMA-%β and HbA1c. Physical activity measures from visit 3 will be used for analysis of 2-hour post-challenge glucose and insulin.

GWAS interaction analyses will be completed in ProbABEL. All traits will be modeled continuously and all analyses will adjust for age, sex and ARIC center. We will use the genotypes from the dosage file in which the SNPs are modeled continuously between 0 and 2. Initially, physical activity will be dichotomized into sedentary/not sedentary. We will further categorize physical activity into three categories: inactive, moderately active and active. In addition to the SNP and the indicator of sedentary/not sedentary, an interaction term of SNP*physical activity variable will be included in each test of association. A likelihood ratio test (LRT) will be performed to assess the statistical

significance of the interaction; p-values $<5x10^{-8}$ will be considered significant. The interaction p-value will be the primary result from this analysis and will be reported to MAGIC. Interaction p-values which are small, but do not pass the $<5x10^{-8}$ threshold may also be reported to MAGIC.

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 Proposal #1198 (Monda, et al) investigated genes, environment, and their interactions to explore determinants of metabolic risk factors, but examined genes associated with cardiovascular disease and the metabolic syndrome.

Manuscript Proposal #1307 (Chu et al) Gene-by-environment interaction for type 2 diabetes. Examines gene-environment interaction using previously reported diabetes susceptibility genes and many environmental variables (including physical activity) using logic regression. We will be examining interaction with SNPs previously reported to be associated with diabetes and diabetes-related traits, but will also be investigating interaction with all genotyped and imputed SNPs available from the Affymetrix 6.0 array.

Manuscript Proposal #1409 (Pankow et al) Genome-wide Association Study of Diabetes-Related Quantitative Traits in ARIC Whites. The lead author of this manuscript proposal (Dr. James Pankow) is a coauthor for this proposal.

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

_X__ 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 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:

Saxena R, Voight BF, Lyssenko V, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007;316:1331-6.

Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007;316:1341-5.

Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 2007;316:1336-41.

Sladek R, Rocheleau G, Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007;445:881-5.

Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661-78.

Zeggini E, Scott LJ, Saxena R, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008;40:638-45.

Chen WM, Erdos MR, Jackson AU, et al. Variations in the G6PC2/ABCB11 genomic region are associated with fasting glucose levels. J Clin Invest 2008;118:2620-8.

Bouatia-Naji N, Bonnefond A, Cavalcanti-Proenca C, et al. A variant near MTNR1B is associated with increased fasting plasma glucose levels and type 2 diabetes risk. Nat Genet 2009;41:89-94.

Bouatia-Naji N, Rocheleau G, Van Lommel L, et al. A polymorphism within the G6PC2 gene is associated with fasting plasma glucose levels. Science 2008;320:1085-8.

Ruczinski I, Kooperberg C, LeBlanc M. Logic regression. J Comput Graph Stat 2003;12:475-511.