ARIC Manuscript Proposal # 1358

PC Reviewed: 04/08/08	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Interaction between FTO genotype and physical activity level on adiposity: The Atherosclerosis Risk in Communities (ARIC) Study

b. Abbreviated Title (Length 26 characters): FTO, Physical Activity, and Adiposity

2. Writing Group:

Writing group members:

Ellen Demerath
Kari North
Keri Monda
James Pankow
Aaron Folsom

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

First author: Ellen Demerath
Address: Epidemiology and Community Health
University of Minnesota School of Public Health
1300 S. Second St., Suite 300
Minneapolis, MN 55454
Phone: (612) 624-8231 Fax: (612) 624-0315
E-mail: ewd@umn.edu

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).

Aaron Folsom, MD
Epidemiology and Community Health
University of Minnesota School of Public Health
1300 S. Second St., Suite 300
Minneapolis, MN 55454
Phone: (612) 626-8862 Fax: (612) 624-0315
E-mail: folso001@umn.edu

3. Timeline:

Statistical analyses: May – July, 2008

Manuscript preparation: August – October 2008
Manuscript revision: November, 2008
Manuscript submission: December, 2009

4. Rationale:

Epidemiologic Evidence

FTO was first identified as a human obesity-related gene via a 500K SNP genome-wide association study of diabetes cases and controls in the Welcome Trust Case-Control Consortium, in which a 1.67 fold risk of obesity was found in individuals with the AA genotype as compared to the TT genotype of SNP rs9939609 (Frayling et al., 2007). That same publication demonstrated replication of the association in 7 adult and 2 juvenile cohorts comprising in total ~38,000 individuals (Frayling et al., 2007), and suggested additive effects of the A allele on BMI. While the association is not found at birth (Frayling et al., 2007; Lopez-Bermejo et al., 2007), the elevated adiposity of A allele carriers develops by two weeks of age (Lopez-Bermejo et al., 2007). Similar associations between adiposity traits and variations in the FTO gene have been subsequently reported in over 12 additional independent cohorts to date (Al-Attar et al., 208; Grant et al., 2008; Do et al., 2008; Lopez-Bermejo et al., 2008; Jacobsson et al., 2008; Hunt et al., 2008; Price et al., 2008; Hinney et al., 2007; Peeters et al., 2008; Andreasen et al., 2007; Scuteri et al., 2007; Dina et al., 2007), making it among the most widely-replicated obesity genes yet identified, and a prime example of the use of the genome-wide association approach to reveal previously unknown genes influencing common complex diseases.

Ethnic differences in association between FTO and BMI

Most of these studies have been in subjects of European or Hispanic ancestry; neither rs9939609, nor other common SNPs in this region of FTO are associated with obesity in East Asian or Oceanic populations (Li et al., 2007; Ohashi et al., 2007; Al-Attar et al., 2008). FTO is a large (>400 mb) gene on chromosome 16q12.2, and the first intronic region, where many of these obesity-associated SNPs reside, tends to demonstrate high linkage disequilibrium (LD), at least in Caucasians (Dina et al., 2007; Grant et al., 2008). It is possible that ethnic differences in LD structure may explain the differences in findings across racial and ethnic groups. Of interest to this proposal and to ARIC are three reports in African American cohorts. Whereas a handful of highly correlated SNPs in strong LD with rs9939609 were all associated with BMI and obesity in a sample of unrelated Sardinians, as well as the European-American and Hispanic-American families in GenNet, they were not associated with BMI or obesity in African Americans (Scuteri et al., 2007). In ARIC, Bressler and colleagues (ARIC ms 1269, in preparation), replicated the findings of Frayling et al. (2007) and others by showing a significant association of rs9939609 with diabetes, explained by its association with BMI in whites ($p<0.0001$). However, they found only a weak association of rs9939609 with obesity in African American members of the cohort (BMI>30: OR= 1.10 (1.00-1.21) p=0.04) and no significant association with continuously distributed BMI (p=0.37). A recent study by Grant et al. (2008) found a SNP (rs3751812) in FTO that was significantly associated with obesity in both white (OR=1.27) and African-American (OR=1.31) children. Interestingly, this SNP was in LD with rs9939609 in whites ($r^2=1.0$) but not in African Americans ($r^2=0.058$). The opportunity to examine SNPs other than rs9939609 in the African American subjects in ARIC may aid the search for the causal variant/s involved in obesity risk.

Putative Function of FTO

Experiments of Gerken et al (2007) and Sanchez-Pulido et al. (2007) suggest that FTO encodes an Fe(II)- and 2oxoglutarate-dependent dioxygenase. This family of enzymes is involved in DNA repair, fatty acid metabolism, and post translational (epigenetic) modification of DNA, including histone lysine demethylation. The putative Fto enzyme localizes to the cell nucleus in mice, and can demethylate single-stranded DNA (Gerken et al., 2007). In both mice (Gerken et al., 2007; Stratigopoulos et al., 2008) and humans (Dina et al., 2007; Wahlen et al., 2008), FTO is ubiquitous but is most highly expressed in adipocytes, the adrenal glands, and in the brain, particularly in the hypothalamus. From a functional standpoint, both fasting (Gerken et al., 2007; Stratigopoulos et al., 2008) and cold exposure (Stratigopoulos et al., 2008) reduce the expression of FTO in the arcuate nucleus, suggesting that FTO may be a metabolic sensor involved in regulation of food intake and energy expenditure. The obesity protective allele in the

FTO gene in its homozygous form (TT) is also associated with increased adipocyte lipolytic activity, suggesting that it may also regulate body fat mass through lipolysis (Wahlen et al., 2008).

Although FTO expression is increased by 20% in the adipocytes of obese compared to non-obese women (Wahlen et al., 2008), FTO rs9939609 genotype apparently does not relate to FTO gene expression levels in the adipocyte (Wahlen et al., 2008). As it is not clear that this SNP is the causal variant, nor that FTO's gene action resides foremost within the adipocyte, the relationship between FTO genotype and gene expression requires further study. While both epidemiologic and physiologic evidence mounts for FTO's role in human obesity, many steps in the pathway/s between FTO genotype and the accrual of excess body fat are still unclear. Elucidating its role in appetite, food intake, and energy expenditure in humans will be particularly important.

Two papers are of specific relevance to the proposed manuscript. One is the paper now in preparation for submission (Bressler et al., ARIC ms 1269), discussed above, showing a significant main effect for rs9939609 on obesity in both African-American and white subjects in ARIC ($p<0.0001$ for whites and $p=0.04$ for African Americans). BMI among individuals with the AA genotype was ~1 kg/m^2 higher than those with the TT genotype in whites ($p<0.0001$). The second is Andreasen et al. (2007) which reported a significant interaction between FTO genotype and self-reported physical activity level in the Danish population-based Inter99 cohort (N=5,722; mean age 46 y.). Among subjects with a low physical activity level, BMI was 1.95 kg/m^2 higher in AA vs TT genotypes, among subject with moderate physical activity BMI was 0.7 kg/m^2 higher in AA vs TT, and among subjects with high physical activity, BMI was not significantly different between AA vs TT genotypes (p for interaction = 0.007). This suggests that being physically active may significantly reduce the elevated risk of obesity associated with FTO genotype, which has public health significance.

By addressing this hypothesis in the ARIC cohort, we have the opportunity to examine the genotypic differences in the protective effects of relatively vigorous exercise on obesity and adiposity for the first time in the US in a biracial cohort of older adults, using more direct measures of adiposity (circumferences and skinfolds) than did Andreasen et al. (2007). Given the 4 study exams at which anthropometric variables were collected, we also have the opportunity to examine the impact of genotype and genotype*physical activity level interactions on the rate of change in adiposity measures over a period of up to 9 years in an aging cohort.

5. Main Hypothesis/Study Questions:

1) To test the interaction effect between individual FTO SNPs genotyped in ARIC (rs9939609, rs1421085, rs8050136, and rs17817449) and baseline self-reported physical activity level on baseline adiposity traits in ARIC.

2) To test the interaction effect between FTO SNP haplotypes and baseline physical activity level on baseline adiposity traits.

3) To test the interaction effect between FTO SNPs genotypes (or haplotypes, if significant) and self-reported physical activity level on longitudinal changes in adiposity traits.

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).

Subjects and Sample size: The usual DNA consent restriction and missing data exclusion criteria will be used. There are ~10,900 Caucasian subjects and 3755 African Americans with FTO genotype (rs9939609), self-reported physical activity, and BMI at baseline. Given the previously reported allele frequencies reported for rs9939609 in ARIC (q=0.41 in whites and q=0.48 in African Americans), approximately 17% of whites (N=1681) and 23% of African Americans (N= 864) will be homozygous for the risk allele, of which one third will be defined as having high, moderate, and low physical activity, respectively.

Power Calculation: We estimated our power to detect various effect sizes for the interaction of interest (Gauderman and Morrison, 2006; Quanto ver 1.1). As a starting place, we used the magnitude of the difference in BMI reported by Andreasen et al. (2007) of TT vs AA genotypes across three levels of PA. Among subjects with a low physical activity level, BMI was 1.95 kg/m^2 higher in AA vs TT genotypes; among subjects with high physical activity, BMI was 0.5 kg/m^2 higher in AA vs TT genotypes. Our power to detect this relatively large magnitude of effect is >99% in both Caucasians and African Americans. Furthermore, we have 80% power to detect an interaction accounting for as little as 0.41 kg/m^2 in Caucasians and as little as 0.72 kg/m^2 in African Americans.

Definitions and Treatment of Variables:

1) Genotype: Based on previous findings of additive effects of the FTO risk alleles on adiposity traits, we plan to test models comparing homozygous risk allele carriers and heterozygotes to homozygous common allele carriers. Haplotypes capturing the most common co-occurances of multiple SNPs in the population will be inferred and reconstructed using the program PHASE (Stephens et al., 2001; Stephens and Donnelly, 2003), assuming a codominant model.

2) Physical Activity: There are four indices of physical activity calculated using responses to the Baecke Questionnaire of Habitual Physical Activity: Sport Activity, Work Activity, Leisure Activity, and the sum (Total Activity). Based upon our observation that the Sport Index provides the most robust effects of the four on cardiovascular disease risk factors and outcomes in ARIC, we propose to primarily examine the gene-by-PA interaction using this measure. Work, Leisure, and Total Activity Indices will also be examined. In the remainder of this proposal, we will refer generally to tests of "PA" (physical activity), which refers to all of the indexes mentioned above. We propose to categorize PA as "high", "moderate", and "low" based on tertile values. Given that women and African Americans have lower PA levels than men and whites, we may also define PA tertiles within sex-race groups. We will also examine variables derived from the Baecke questionnaire (including metabolic equivalents: METmins/week for total PA, moderately vigorous PA, and vigorous PA) as continuously distributed forms of the PA variable, which will increase power to detect interaction.

3) Adiposity Measures: Baseline adiposity measures in ARIC (BMI, waist circumference, hip circumference, waist-hip ratio, waist/stature ratio, and sum of triceps and subscapular subcutaneous skinfolds) will be used primarily as continuous variables. If these variables exhibit significant skewness or kurtosis, a variety of transformations will be tested (e.g., log-transform, inverse normal transform) to achieve normality prior to analysis. ~95% of the white subjects and 75% of African American subjects in ARIC have at least 2 longitudinal measures of BMI, waist circumference, and hip circumference.

Statistical analysis: All analyses will be conducted separately in whites and African Americans due to the likelihood of different genetic structure in the two races, particularly with regard to the haplotype analysis. To address our primary aims of detecting gene-by-PA

interaction, a general linear ANOVA model will be used to test both the main effects of FTO genotype or haplotype and PA tertile, as well as the genotype*PA interaction effect on continuously distributed values of BMI, Waist, Waist/Hip, Waist/Stature, and the sum of subcutaneous skinfolds. We will assume an additive model, as the studies published to date indicate that mean BMI and waist circumference values for FTO heterozygotes are midway between the values for the homozygotes. Physical activity will be represented primarily by tertiles of the self-reported Sport activity index derived from responses to the Baecke Physical Activity Questionnaire. Minimally adjusted models (with age, sex, and field center as covariates) and more extensively adjusted models (with age, sex, field center, stature, current cigarette smoking (baseline yes/no), and alcohol exposure as covariates) will be tested. If the results for the individual SNP models differ in the two races, we will formally test that by pooling the races and adding a genotype x race interaction term. We will also conduct exploratory analysis of sex- and age- specific effects on the genotype x PA interaction by conducting stratified analyses. The Work activity index and the Leisure activity index will also be examined as alternatives to the Sport activity index. The Benjamini and Hochberg methods will be used to correct for multiple testing, considering both the number of traits and the number of genetic models tested.

As a secondary aim, we will examine the association of FTO genotype on changes in adiposity measures, and the interaction of baseline PA with FTO genotype on changes in adiposity measures. We will employ ANOVA models first using annualized changes in BMI, waist circumference, hip circumference, waist/stature ratio, waist-hip ratio, and the sum of skinfolds (=difference between first and last measure / # years between first and last measure) for subjects with >1 observation will be used, followed by mixed effects models incorporating all repeated adiposity trait observations. In the mixed effects models, sex, age, race, field center, genotype, physical activity, time, genotype*time, PA*time, genotype*PA and genotype*PA*time interaction terms will be considered fixed effects and the subject-specific intercepts and slopes will be considered random effects. A variety of repeated measures covariance structures will be compared using likelihood ratio testing to determine the covariance structure with the best fit to the data. The likelihood of models incorporating a genotype*time and genotype*PA*time interactions (ie., testing for differences by genotype in the change over time in adiposity, and differences by physical activity level within genotype in the changes in adiposity, respectively) will be compared to constrained models in which these terms are not estimated and will address the hypothesis that changes in adiposity in the ARIC participants is influenced by FTO genotype/s. The AIC and BIC will be used along with the likelihood ratio test to determine the best-fitting and most parsimonious models.

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 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 ICTDER03 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/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)?

#1269 "Risk of Type 2 Diabetes is Associated with Variation in the *FTO* Gene in White but Not in African-American Participants in the Atherosclerosis Risk in Communities Study" (Lead author: Jan Bressler, U.T. Houston Health Science Center) (*ms in preparation)*

Lead author of ms #1269 has been contacted for collaboration.

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* AS#1995.07_)

___ 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.