1.a. Full Title: Candidate-gene association study of variants influencing age at menarche and their association with childhood growth and BMI trajectories

b. Abbreviated Title (Length 26 characters): Genetic variants influencing menarche and childhood growth

2. Writing Group: Ellen Demerath, Javier Nieto, Nora Franceschini, Aaron Folsom, Eric Boerwinkle, others welcome

Collaboration within CHARGE Consortium will be explored.

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: Nora Franceschini, MD, MPH
Address: Department of Epidemiology
University of North Carolina Chapel Hill
Bank Of America Center
137 E. Franklin St., Suite 306
CB #8050
Chapel Hill, NC 27514
Phone: 919-966-1305
Fax: 919-966-9800
E-mail: noraf@unc.edu

Senior author: Ellen Demerath, Ph.D.
Associate Professor,
Division of Epidemiology and Community Health
School of Public Health, University of Minnesota
1300 South Second St., Suite 300
Minneapolis, MN 55454-1015
(612) 624-8231 (phone)
(612) 624-0315 (fax)
e-mail: ewd@umn.edu
3. **Timeline:** Analyses will begin immediately upon approval of this proposal.

4. **Rationale:**

   **Background.** Earlier age at menarche is associated with reduced stature and increased risk of obesity \(^1,^2,^3\), among other outcomes, although the temporal relationships are not clear, as studies conflict on whether early onset obesity may lower the age at menarche while others have shown that timing of menarche is associated only with subsequent weight status. Although age at menarche is under relatively strong genetic control, with heritability of approximately 50% (reviewed in \(^4\)), environmental factors are also important. A secular decline in the mean age at menarche has occurred over the past 100 years \(^5,^6,^7\), likely as a result of improved childhood nutrition. In the ARIC cohort, age at menarche has decreased by 0.02 years per year of birth from 1923-1942 (p< 0.0001).

Several recent genome-wide association studies (GWAS) have identified single nucleotide polymorphisms (SNPs) in genes for age of menarche \(^8,^9\). Liu et al. identified associations of SNPs in the *SPOCK* gene with age of menarche \(^8\). In our own GWAS within the CHARGE Consortium involving a sample size of 17,510 women of European descent, we identified two genes associated with age of menarche \(^9\) (in press). The strongest signal was at 9q31.2 (P = 1.7x10^-9) where the nearest genes include *TMEM38B*, *FKTN*, *FSD1L*, *TAL2* and *ZNF462*. The next best signal was near the *LIN28B* gene (rs7759938; P = 7.0x10^-9), which also influences adult height. We were unable to replicate the association findings from the study of Liu et al.

**Aim.** Given the epidemiologic association of timing of menarche and childhood growth traits and the pleiotropic effects of age at menarche-associated SNPs on adulthood height reported in the CHARGE meta-analysis, we aim to examine the pleiotropic effects of genes associated with menarche with childhood weight and height growth trajectories and to examine whether the main SNP effects on age of menarche are modified by childhood BMI and weight trajectories.

**Approach.** We will take advantage of data on childhood growth in 724 ARIC participants that were collected as part of the Hagerstown Growth Study (HGS; Javier Nieto, PI) to perform a candidate-gene (locus) association study. Dr. Nieto will provide the link between the ARIC id and the HGS ids to link childhood growth data and genotype data, and Drs. Franceschini and Demerath will perform the phenotypic and genetic analyses on the merged dataset. The SNPs we plan to examine include the top hits for age at menarche identified in CHARGE (N~2), those identified by Liu and colleagues (N~3), as well as SNPs identified in the Nurses’ Health Study+ WGHS and DeCODE (in press). The number of SNPs to be examined from the Nurses Health +WGHS report is not yet known but the manuscript is expected to be published shortly in Nature Genetics alongside that from CHARGE. Age at menarche SNPs in other loci that may be reported in the near future will also be included. A total of ~10 SNPs representing...
independent loci influencing normal variation in age at menarche are expected to be tested in this analysis.

5. Main Hypothesis/Study Questions:
Hypothesis 1: Genetic variants associated with timing of menarche also influence weight, height, Tanner staging, and other available developmental data in childhood. Because these genes may have effects on development of both boys and girls, we will perform combined and sex-stratified analyses for this aim.
Hypothesis 2: Childhood BMI and weight status (or trajectories) modify the association of these genetic variants with age at menarche.

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: Because >99% of the HGS subjects were of European ancestry, we will include only white participants in this analysis, and restrict the analysis to those with relevant genotype and phenotypic data. Both boys and girls will be included, as the SNPs in question were related not only to age at menarche, but also to general growth (height) and to a lesser degree, with BMI, so it is possible that may also influence growth and development traits in males.

Variables (phenotypes): Pre-pubertal and post-pubertal height, weight, and BMI; serial weight and height changes; birth weight; age at peak height velocity (if available), age at initiation of puberty (Tanner stage=2), if available; other developmental traits in the HGS, if available.

Inclusion: All white ARIC/HGS subjects with available genotype and phenotype data
Exclusions: Non-white participants, due to low N (<1%) in HGS
Exposure: Candidate SNPs (measured and imputed) passing QC from the Affymetrix 6.0 SNP chip that have been previously identified in GWAS for age of menarche (NSNPs~10)
Models:

Primary analysis:
1) We will use linear regression models with additive genetic effects to model the association of candidate SNPs and childhood growth traits.
   a. Sex-stratified and combined models will be tested

2) We will also use linear regression models to test (in girls only) the interaction of candidate SNPs and childhood BMI on age at menarche.

Population stratification will be assessed using principal components (PC) analysis where PCs that are significantly associated with the trait will be included.

We have contacted the PIs of the Fels Longitudinal Study to assess their interest in performing in silico replication of any significant findings from this analysis, and will consider extending the analysis to other cohorts in a meta-analysis, including the ALSPAC study, if appropriate.
**Covariates:** Sex, year of birth

**Statistical significance:** We will correct for the number of independent tests performed using Bonferroni correction.

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)?

ARIC Manuscript Proposal # 1472. Genome-wide association analysis of age of menarche: the CHARGE Consortium

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*2006.03 (Stampede and Geneva genotype funding in Caucasians).  
____ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* __________ __________ __________)
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:


