1.a. Full Title: Replication Request For: Interactions between Sugar-Sweetened Beverage Intake with Genetic Variants in the CHREBP-FGF21 Pathway on Fasting Glucose and Insulin

b. Abbreviated Title (Length 26 characters): SSB-genetic variants interactions on glycemic traits

2. Writing Group:
   Mariaelisa Graff, Kristin Young, Kari E North, others welcome

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal.

MG

First Authors: Mariaelisa Graff

Address: 137 E. Franklin Str, Chapel Hill, NC 27514
Phone: 919-843-6249
E-mail: migraff@email.unc.edu

Corresponding author: same as first author

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).

Name: Kari E North
Address: Department of Epidemiology
University of North Carolina at Chapel Hill
137 E. Franklin St, Suite 306
CB #8050
Chapel Hill, NC 27514
Phone: 919-966-2148 Fax: 919-966-9800
E-mail: kari_north@unc.edu

3. Timeline: ~6 months (June 2016-October 2016)
   Replication cohort statistical analyses: June 2016
   Interpretation and meta-analyses: July 2016
   Manuscript preparation: July 2016
   Manuscript submission: July 2016
4. **Rationale:**

Epidemiological evidence suggests that excess sugar-sweetened beverage (SSB) intake is associated with increased risk of metabolic syndrome [1, 2] and type 2 diabetes [3]. SSB intake may also contribute to weight gain [2, 4, 5], in particular the accumulation of visceral and ectopic fat depots [1, 3], associated with greater metabolic dysregulation. Sucrose (table sugar) and high-fructose corn syrup are the most common forms of sugar in SSBs, and they are composed of nearly equal amounts of glucose and fructose [6]. Evidence from both animal [7] and human studies [8, 9] suggests that the fructose moiety is particularly harmful to cardiometabolic health. Given that SSBs are the major dietary source of fructose intake in the US, this is of significant public health importance [10].

Over the past three decades, the prevalence of type 2 diabetes has been steadily increasing worldwide with an estimated 8.5% (422 million) adults having this chronic disease in 2014 [11]. Currently, an estimated 37% of US adults have prediabetes with underlying insulin resistance, indicative of a trajectory path to diabetes that could be altered by lifestyle changes [12]. Excess sugar intake, particularly in the form of sugar-sweetened beverage, may impair glucose homeostasis. A recent meta-analysis of 17 cohorts found that daily intake of SSBs was associated with a 13% greater risk of T2D, independent of overall adiposity [3]. A few studies have linked SSB intake to measures of glucose and insulin metabolism [13-15], yet the influence of genetic variation to the underlying susceptibility of SSB induced metabolic disease in humans remains unknown.

Carbohydrate Responsive Element-Binding Protein (ChREBP, also known as Mlxipl) is a transcription factor that responds to intracellular carbohydrate metabolites and is a principal mediator of carbohydrate-induced gene expression in key metabolic tissues including the liver [2, 16-18]. Recent data indicate that hepatic ChREBP is particularly responsive to fructose intake [19] and contributes to fructose-induced metabolic diseases [20]. We have recently demonstrated that fructose ingestion acutely increases circulating levels of the novel metabolic hormone fibroblast growth factor 21 (FGF21) in a hormone-like manner [21]. FGF21 has pleiotropic effects on carbohydrate and lipid metabolism [22]. We and others have recently reported that single nucleotide polymorphisms (SNPs) in the FGF21 locus which associate with higher circulating FGF21 concentrations also associate with higher carbohydrate intake relative to fat in humans [23, 24]. Moreover, it has recently been confirmed in rodents and non-human primates that sucrose-induced FGF21 regulates sweet taste preference [25, 26]. Together, these data indicate that a ChREBP-FGF21 hormonal axis mediates an adaptive response to sugar consumption and may contribute to regulation of metabolic traits in the context of sugar consumption.

To date, no gene-diet study has examined interactions between SSBs and genes on indices of carbohydrate metabolism. We hypothesized that common SNPs associated with genes involved in fructose metabolism and the ChREBP-FGF21 pathway may interact with SSB intake to influence glycemic traits. The study will include up to 34,748 participants of European descent from the following 11 US and European cohort studies (6 discovery cohorts and 5 replication cohorts) of the CHARGE Consortium Nutrition Working Group. Aric will be part of the replication aspect.

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

The aim of the current investigation is to (1) evaluate the relationship between SSB intake and fasting plasma glucose (FG) and fasting insulin (FI) in 11 cohorts from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium totaling 34,748 subjects and (2) to examine whether these associations were modified by SNPs related to ChREBP function.
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).

Study design:

A. Subjects and Sample Size

Discovery Cohorts
The discovery cohorts were Cardiovascular Heart Study (CHS), Framingham Heart Study (FHS), Multi-Ethnic Study of Atherosclerosis (MESA), Rotterdam Study I (RS1), Rotterdam Study II (RS2), and the Cardiovascular Risk in Young Finns Study (YFS).

Replication cohorts
The replication cohorts will include Atherosclerosis Risk In Communities (ARIC) Study, Malmö Study, Netherlands Epidemiology in Obesity (NEO) Study, Nurses’ Health Study (NHS), and Western Australia Pregnancy Cohort (Raine) Study.

For the ARIC replication data set, the study participants will include European Americans (N=8,758) who have provided informed consent and have measured fasting insulin, glucose profiles, genetic data, and sugar-sweetened beverage (SSB) data at visit 1.

Inclusion:
- Adults of ≥ 18 years of age
- European ancestry

Sugar-sweetened beverages:
One serving of SSB is defined as 360 ml or 12 fl. oz., and included the following: (1) Coke, Pepsi, or other cola with sugar; (2) caffeine free Coke, Pepsi, or other cola with sugar; (3) other carbonated beverage with sugar (e.g., 7-Up, ginger ale); (4) Hawaiian Punch, lemonade, or other non-carbonated fruit drinks. Fruit intake, vegetable intake, whole grain intake, and fish intake in servings/day, alcohol intake in grams/day, and saturated fatty acid as percentage of total energy intake (where 1 g saturated fat has 9 kcal) were further quantified and used as covariates in the present analysis. SSB intake will be used continuously and further dichotomized into low (<1 serving/day) and high (≥1 serving/day) intakes, whereas all remaining dietary variables were considered continuously only.

Outcome:
Glycemic biomarkers, fasting glucose and insulin, will be used if measured after ≥8 hours of fasting.

Genotype data:
- For the discovery results, SNPs were selected that previously showed genome-wide (i.e. P <5E-08) or sub-genome-wide (i.e. P <5E-06) significant associations with hypertriglyceridemia or low-HDL cholesterol in humans and were found within the ChREBP locus or within the loci of genes predicted to regulate either ChREBP or the biological response to ChREBP activation.

For replication in ARIC, we will use SNPs that met nominal significance in the interaction results based on the discovery findings. We will pull these SNPs from imputed data based on HapMap II.

Summary data analysis:
Analytical Methods in Parent Study

In the discovery cohorts, inverse-variance weighted, random-effect meta-analyses were conducted using the ‘metafor’ R package for 1) main associations of SSB-intake on FG and FI, and fixed-effect meta-analyses using METAL (version released 2011-03-25) for 2) main associations of the selected SNPs on outcomes, and 3) interactions between SSB intake and nominally significant SNPs on glycemic outcomes. Statistical significance for association/interaction tests was defined at a level of 0.001, based on Bonferroni correction for 36 (=18 independent SNPs x 2 glycemic outcomes) total tests.

Analytical Methods for Replication in ARIC

For nominally significant interaction results (i.e. $P < 0.05$) from the discovery cohorts analyses, we will further investigate via meta-analysis in replication cohorts the 1) main effect associations of SSB intake with FG and FI concentrations, 2) main associations between nominally significant SNPs and glycemic outcomes, and 3) interactions between SSB intake and nominally significant SNPs on glycemic outcomes.

Analytical Methods for Meta-Analysis with other HCHS/SOL Replication Cohorts

Finally, we will further conduct joint meta-analyses (combined discovery cohorts and replication cohorts) for the described analyses. Heterogeneity across studies will be tested by using Cochran’s Q statistic and quantified using the $I^2$ statistic. Analyses with moderate heterogeneity ($I^2 > 30\%$) will be further assessed for potential sources of heterogeneity by conducting meta-regression and sensitivity analyses. Meta-regression analyses were conducted using the R metafor package (R version 3.1.0) to assess the effect of the following moderator variables on heterogeneity of association/interaction: geographical location (U.S. vs. northern Europe vs. Australian), BMI (<27 vs. ≥27 kg/m²), and sample size ($n < 1000$ vs. ≥1000).

A. Example Figures and Tables for Manuscript

Meta-analyzed interactions between SSB and SNPs on glycemic traits in discovery cohorts

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr</th>
<th>Gene</th>
<th>Alleles†</th>
<th>Fasting Glucose (mmol/L) β (SE)</th>
<th>$P$</th>
<th>Fasting Insulin (ln-pmol/L) β (SE)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10819937</td>
<td>9</td>
<td>ALDOB</td>
<td>C/G</td>
<td>-0.0067 (0.0147)</td>
<td>0.65</td>
<td>-0.0025 (0.0152)</td>
<td>0.87</td>
</tr>
<tr>
<td>rs10819931</td>
<td>9</td>
<td>ALDOB</td>
<td>T/C</td>
<td>-0.0229 (0.0215)</td>
<td>0.29</td>
<td>0.0023 (0.0212)</td>
<td>0.92</td>
</tr>
<tr>
<td>rs174546</td>
<td>11</td>
<td>FADS1</td>
<td>T/C</td>
<td>0.0018 (0.0116)</td>
<td>0.87</td>
<td>0.0070 (0.0115)</td>
<td>0.54</td>
</tr>
<tr>
<td>rs838133</td>
<td>19</td>
<td>FGF21</td>
<td>A/G</td>
<td>-0.0149 (0.0128)</td>
<td>0.24</td>
<td>-0.0176 (0.013)</td>
<td>0.17</td>
</tr>
<tr>
<td>rs4607517</td>
<td>7</td>
<td>GCK</td>
<td>A/G</td>
<td>-0.0020 (0.0145)</td>
<td>0.89</td>
<td>0.0101 (0.0144)</td>
<td>0.48</td>
</tr>
<tr>
<td>rs1260326</td>
<td>2</td>
<td>GCKR</td>
<td>C/T</td>
<td>-0.0015 (0.0105)</td>
<td>0.89</td>
<td>-0.0088 (0.0104)</td>
<td>0.40</td>
</tr>
<tr>
<td>rs2119026</td>
<td>2</td>
<td>HKH</td>
<td>C/T</td>
<td>-0.0111 (0.0119)</td>
<td>0.35</td>
<td>0.0059 (0.0119)</td>
<td>0.62</td>
</tr>
<tr>
<td>rs1542423</td>
<td>4</td>
<td>KLB</td>
<td>T/C</td>
<td>-0.0105 (0.0111)</td>
<td>0.34</td>
<td>0.0302 (0.0110)</td>
<td>0.006</td>
</tr>
<tr>
<td>rs799166</td>
<td>7</td>
<td>MLXIPL</td>
<td>C/G</td>
<td>-0.0084 (0.0175)</td>
<td>0.63</td>
<td>0.0183 (0.0187)</td>
<td>0.33</td>
</tr>
<tr>
<td>rs799168</td>
<td>7</td>
<td>MLXIPL</td>
<td>G/A</td>
<td>-0.0229 (0.0164)</td>
<td>0.16</td>
<td>0.0173 (0.0176)</td>
<td>0.33</td>
</tr>
<tr>
<td>rs799160</td>
<td>7</td>
<td>MLXIPL</td>
<td>T/C</td>
<td>-0.0149 (0.0159)</td>
<td>0.35</td>
<td>0.0091 (0.0165)</td>
<td>0.58</td>
</tr>
<tr>
<td>rs11974409</td>
<td>7</td>
<td>TBL2</td>
<td>A/G</td>
<td>-0.0095 (0.0139)</td>
<td>0.50</td>
<td>0.0086 (0.0144)</td>
<td>0.55</td>
</tr>
<tr>
<td>rs11920090</td>
<td>3</td>
<td>SLC2A2</td>
<td>A/T</td>
<td>0.0214 (0.0165)</td>
<td>0.19</td>
<td>0.0123 (0.0165)</td>
<td>0.46</td>
</tr>
<tr>
<td>rs11924032</td>
<td>3</td>
<td>SLC2A2</td>
<td>A/G</td>
<td>0.0142 (0.0117)</td>
<td>0.23</td>
<td>-0.0023 (0.0117)</td>
<td>0.84</td>
</tr>
<tr>
<td>rs5438</td>
<td>1</td>
<td>SLC2A5</td>
<td>A/G</td>
<td>0.0220 (0.0283)</td>
<td>0.78</td>
<td>-0.011 (0.0283)</td>
<td>0.70</td>
</tr>
<tr>
<td>rs3820034</td>
<td>1</td>
<td>SLC2A5</td>
<td>C/T</td>
<td>0.0140 (0.0140)</td>
<td>0.32</td>
<td>0.0009 (0.0142)</td>
<td>0.95</td>
</tr>
<tr>
<td>rs5840</td>
<td>1</td>
<td>SLC2A5</td>
<td>T/C</td>
<td>0.0109 (0.0112)</td>
<td>0.33</td>
<td>0.0002 (0.0112)</td>
<td>0.98</td>
</tr>
<tr>
<td>rs2954029</td>
<td>8</td>
<td>TRIB1</td>
<td>A/T</td>
<td>0.0186 (0.0111)</td>
<td>0.09</td>
<td>0.0069 (0.0112)</td>
<td>0.54</td>
</tr>
</tbody>
</table>
Example Table 2. List of all SNPxSSB interactions on the indicated glycemic trait.

Example Table 3. Significant loci that reach $P<5\times10^{-8}$ from the meta-analysis of the discovery plus replication data sets.

<table>
<thead>
<tr>
<th>Trait</th>
<th>SNP</th>
<th>Chr</th>
<th>Gene</th>
<th>Alleles†</th>
<th>n</th>
<th>$\beta$ (SE)</th>
<th>$P$</th>
<th>$I^2$</th>
<th>Discovery Cohorts</th>
<th>Replication Cohorts</th>
<th>All Cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fasting Glucose</td>
<td>rs10819937</td>
<td>9</td>
<td>ALDOB</td>
<td>C/G</td>
<td>15590</td>
<td>0.0302</td>
<td>0.006</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fasting Glucose</td>
<td>rs10819931</td>
<td>9</td>
<td>ALDOB</td>
<td>T/C</td>
<td>15590</td>
<td>0.0230</td>
<td>0.1</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fasting Insulin</td>
<td>rs838133</td>
<td>19</td>
<td>FGF21</td>
<td>A/G</td>
<td>15590</td>
<td>0.0308</td>
<td>0.09</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fasting Insulin</td>
<td>rs1542423</td>
<td>4</td>
<td>KLB</td>
<td>T/C</td>
<td>15590</td>
<td>0.0340</td>
<td>0.26</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example Figure 1. Example forest plot of main association between SSB intake and glycemic trait.
References (50 max):


beverages is positively related to insulin resistance and higher plasma leptin concentrations in men and nonoverweight women. J Nutr 144:1099–1105. doi: 10.3945/jn.114.195230


34. Austin GL, Krueger PM (2013) Increasing the percentage of energy from dietary sugar, fats, and alcohol in adults is associated with increased energy intake but has minimal association with biomarkers of cardiovascular risk. J Nutr 143:1651–1658. doi: 10.3945/jn.113.180067


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?

___X__ Yes
____ No

(This file ICTDER 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.cscexp.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)?

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/

12a. 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.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is your responsibility to upload manuscripts to PubMed Central whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.