Do Genotype Patterns Predict Weight Loss Success for Low Carb vs. Low Fat Diets?

Analysis Plan (effective March 2013)

Scientific background and explanation of rationale;

BACKGROUND: Genomics research is advancing rapidly, and links between genes and obesity continue to be discovered and better defined. A growing number of single nucleotide polymorphisms (SNPs) in multiple genes have been shown to alter an individual's response to dietary macronutrient composition. Based on prior genetic studies evaluating the body's physiological responses to dietary carbohydrates or fats, we identified multi-locus genotype patterns with SNPs from three genes (FABP2, PPARG, and ADRB2): a low carbohydrate-responsive genotype (LCG) and a low fat-responsive genotype (LFG). In a preliminary, retrospective study (using the A TO Z weight loss study data), we observed a 3-fold difference in 12-month weight loss for initially overweight women who were determined to have been appropriately matched vs. mismatched to a low carbohydrate (Low Carb) or low fat (Low Fat) diet based on their multi-locus genotype pattern.

OBJECTIVE: The primary objective of this study is to confirm and expand on the preliminary results and determine if weight loss success can be increased if the dietary approach (Low Carb vs. Low Fat) is appropriately matched to an individual's genetic predisposition (LCG vs. LFG) toward those diets. This study will target both women and men (the A TO Z study involved only women), and address a set of specific aims intended to further elaborate on potential mechanisms and the clinical utility of these results. A new secondary aim has been added to this resubmitted application that will involve a rigorous exploratory investigation of additional SNPs that have shown genome-wide significant associations with obesity and metabolic phenotypes that might improve on the 3-SNP signature.

DESIGN: The main study is a randomized trial employing a 2X2 parallel design to test the central hypothesis that there will be greater weight loss when 320 overweight/obese nondiabetic adults are matched vs. mismatched by genetic predisposition (LCG vs. LFG) to a 12-month Low Carb vs. Low Fat weight loss diet (n=80/cell). Participants will be genotyped prior to randomization, and blinding will be maintained for the genotyping results for both participants and data collectors during the study. Other than the primary outcome of weight change, which will be assessed monthly, primary data collection will occur at 0, 3, 6, and 12 months and include energy intake (3-day unannounced 24-hour recalls, NDS-R), appetite/satiety/hunger, energy expenditure (resting energy expenditure), body composition (DEXA), and blood variables (lipids, insulin, glucose, OGTT).

IMPACT: If the intriguing preliminary retrospective results are confirmed in this full scale study, the results will demonstrate that inexpensive DNA testing could help dieters predict whether they will have greater weight loss success on a Low Carb or a Low Fat diet. Commensurate with increasing scientific interest in personalized medicine approaches to intervention development, this would provide an example of the potentially substantial health impacts that could be obtained through understanding specific gene-environment interactions that have been anticipated from the unraveling of the human genome.

Specific objectives or hypotheses;

Specific Aim #1 – Determine, using an experimental design, if matching vs. mismatching overweight women and men to a Low Carb or a Low Fat weight loss diet based on their

predetermined genotype pattern (LCG vs. LFG) has a significant and differential impact on 12-month weight loss.

Specific Aim #2 – Explore whether additional gene variants that are robustly associated with obesity or related metabolic phenotypes may improve upon the LCG/LFG prediction of the optimal response to Low Carb vs. Low Fat weight loss diets.

Specific Aim #3 – Examine putative mechanisms that could explain differential weight loss success (e.g., energy intake, appetite/satiety/hunger, energy expenditure, insulin/glucose homeostasis).

Description of trial design including allocation ratio;

Parallel design
12-month protocol
Randomly assigned 50:50 to Low-Fat vs. Low-Carb weight loss diets

Eligibility and exclusion criteria for participants;

Participants will be recruited from the local community primarily through media advertisements. Pre-menopausal women and men ages 18-50 years will be invited to enroll if BMI is 27-40 kg/m^2, body weight is stable over the previous 2 months, and medications are stable for ≥ 3 months. Potential participants will be excluded if they self-reported: hypertension, type I or II diabetes mellitus, heart, renal, or liver disease, cancer or active neoplasms, hyperthyroidism unless treated and under control, taking any medications known to affect weight/energy expenditure, blood pressure, or blood lipids, smoking, alcohol intake ≥ 3 drinks/day, pregnancy, lactation, no menstruation for the previous 12 months, or plans to become pregnant within the next year.

Interventions for each group with sufficient details to allow replication, including how and when they are going to be administered;

The intervention will be a class-based education program led by two health educators (HEs). Participants will be assigned to groups of 15-20 per class to follow either a Low-Fat or Low-Carb diet. There will be 22 one-hour classes over 12 months; once/week for 8 weeks, then once every other week for 8 weeks, then once every third week for 8 weeks, and then finally once/month for the last 6 months of the protocol. The focus of the first 8 classes will be on strategies for dramatically lowering total grams of fat or carbohydrate, depending on diet assignment. From the 9^{th} week through 6 months the focus will be on similar topics for both diet groups – holiday eating, family member situations – both supportive and unsupportive, deviations from adherence and strategies for getting back on track, and much more. For the last 6 months of the study most of the topics will involve identifying and addressing any barriers to making these long-term and lasting dietary changes.

Dietary strategy. There will be four central components to the dietary strategy. The first is "How Low Can You Go" (Limbo). Low-Fat participants will be instructed to cut back to 20 grams/day of total fat, and for Low-Carb to 20 grams/day of digestible carbohydrate. The goal is to achieve the lowest level of fat or carbohydrate intake within the first eight weeks. The second stage (Titrate) is to slowly add fat or carbohydrate back to the diet in increments of five grams/day (e.g., from 20 to 25 grams/day) and then hold it at that amount for 1-4 weeks before adding another 5 grams/day. The third component is to identify the lowest level of fat or carbohydrate

intake participants feel could be maintained long term, potentially for the rest of their lives. The fourth strategy is to promote high nutrient density (Quality). Other *Quality* concepts included "real food," "minimally processed," "seasonal," "organic," "grass-fed," "whole grain," and "pasture-raised," depending on diet assignment. Both diet groups will receive similar instructions to drink water and to minimize added sugars, refined white flour products, and sources of *trans* fats. All participants will be encouraged to take an active role in making food choices; by preparing their own foods at home, reading labels, and asking for appropriate modifications for restaurant menu items.

In summary, the diet strategy for both Low-Fat and Low-Carb is a "*Limbo-Titrate-Quality*" approach designed to motivate participants to achieve the lowest possible level of fat or carbohydrate intake with maximal overall nutritional quality and a dietary pattern that could be continued for a lifetime.

Beyond fat and carbohydrate lowering. Notably, there are no calorie restriction targets in the intervention. Participants will be encouraged to track their intake using daily food journals and computer tracking programs. While the first 8 weeks of classes focused specifically on separate strategies to lower fat or carbohydrate intake, the subsequent four months of classes addressed more global topics for both diet groups, similarly, such as mindful eating, adequate sleep, body acceptance, and sugar addition.

Completely defined pre-specified primary and secondary outcome measures,

Primary outcome is weight change (likely to be looked at as ***12m change in BMI***, but will also calculate the trajectory of the combined 3m, 6m and 12m difference from baseline).

Secondary outcomes include:

Insulin resistance (assessed by repeated OGTTs)
Lipid profile
Percent body fat (as assessed by DXA)
Blood Pressure
Cytokines

Description of the similarity of interventions, if relevant;

- Both diet groups will meet for the same number of classes – 22 over the course of a year.
- Both diet groups will be encouraged to be physically active, and to meet or exceed national standards for physical activity.
- Both groups will be told that there is no specific calorie restriction involved, but they are allowed to choose this and follow it themselves if they want to.
- Both groups will be given initial diet goals of decreasing current fat or carbohydrate intake to 20 grams/day, which we consider to be comparably ambitious.

Settings and locations where the data are to be collected;

All settings and locations are on the Stanford University campus (i.e., single site trial).

Data collection:
1070 Arastradero Road:
DEXA, self-reported weight, and some others collected at 1070 Arastradero Rroad, the human

intervention trial site of our division in the Department of Medicine – the Stanford Prevention Research Center.

CTRU – Clinical Translational Research Unit (NIH, funded through CTSA)
Metabolic cart – Resting Energy Expenditure and Respiratory Quotient
Blood sampling:
    Oral Glucose Tolerance Test (T0, T30, T60 and T120 minutes) – insulin and glucose
    Lipids (all fasting)

Diet data:
Three unannounced 24-hour recalls at each major time point (Baseline, 3m, 6m, 12m), administered by trained dietitians.

Questionnaire data: On-line

Dates defining the periods of recruitment and follow-up;

Recruitment will start in January of 2013, and end in May of 2015
    Cohort 1: Spring 2013 – Spring 2014
    Cohort 2: Fall 2013 – Fall 2014
    Cohort 3: Spring 2014 – Spring 2015
    Cohort 4: Fall 2014 – Fall 2015
    Cohort 5: Spring 2015 – Spring 2016
Follow-up will end in Spring of 2016

Description of how and when the pre-specified primary and secondary outcome measures will be assessed;

Weight – measured at CTRU by nurses
Percent body fat – measured by DXA
Metabolic profile of insulin, glucose, and lipid parameters (from blood samples taken at CTRU)
    Glucose – analyzed by CTRU
    Insulin – analyzed by NORC at Washington University in St Louis
    Lipids – analyzed by Northwest Lipid Lab

Time points: baseline, 3, 6 and 12 months (except for OGTT – only 0, 6 and 12 months)

Description of how sample size will be determined, including statistical power analysis and a predetermined stopping point;

Sample Size/Power Calculations. Based on our previous experience conducting weight loss trials of this length, for the primary analysis (Aim 1) we will assume an attrition rate of ~20% among the n=320 randomized in the study; we expect ~260 participants will complete the study and be included in the complete case analysis. With 260 participants we will have more than 95% power to detect differences of 1 unit change in BMI (~6.5 lbs, or ~3 kg) between the appropriate and inappropriate groups in our primary aim (i.e., a moderate effect size of 0.5). This is based on a two-sided t-test, assumes a type I error rate of 0.05, and a standard deviation of 12 month weight change of 2 BMI units (~13 lbs). The sample size for the secondary analysis (aim 2) will be slightly higher since all participants are included without consideration of the 3-locus genotype (n=400 of which about 320 should have complete follow-up), but given the exploratory nature, we will not make any strong inferences at alpha=0.05; instead, these

analyses will only be seen as a first selection screen for choosing/prioritizing additional markers for future study.

No stopping point.

Description of statistical methods to be used to compare groups for primary and secondary outcomes (t-tests of group means, permutation tests, intent-to-treat analysis, using randomization as an instrumental variable to recover the local average treatment effect, etc.);

Primary analysis will test the hypothesis that there is a diet X genotype interaction for weight loss.

STATISTICAL ANALYSES

Preliminary Checking of Data and Underlying Assumptions. Descriptive statistics such as means, medians, standard deviations, and interquartile ranges will be generated for all continuous variables, while frequency distributions will be provided for discrete and categorical variables. Graphical tools such as histograms and QQ-plots will be used to assess distributional properties of continuous variables. In cases where nonnormality is suspected, a transformation of the outcome or a non-parametric approach may be considered. We will also evaluate whether, despite the randomization procedure, baseline variables differ between the 2 intervention arms and the genotype strata and adjust for those variables where significant differences are observed. Finally, where possible, we will describe differences between those with missing and observed data on key variables.

Plan of Analyses

Primary Analysis. The primary analysis will address Specific Aim 1. Using 12-month weight change as the primary outcome, we will assess whether there are differences in weight change between those appropriately assigned to intervention group or not using the information from the LCG and LFG strata. Regression techniques such as ANOVA will be used to estimate differences between groups (both with and without adjustment for baseline characteristics).

Secondary Analyses. The secondary analyses will address Specific Aim 2. Using 12-month weight change as the outcome, we will use group lasso methods to assess whether additional SNPs are predictive of weight change and/or can improve upon the ability of the multi-locus genotype pattern to predict weight change. The additional SNPs we will consider are those that have been previously robustly documented to have genome-wide significant association with weight, waist circumference or metabolic phenotypes (lipid levels, type 2 diabetes, insulin resistance) in previous studies. The list of genes will be compiled based on the latest update of the Catalog of Genome-Wide Association studies, which is continuously updated by NHGRI18, 93 and will include all independent genetic variants with corresponding p-values smaller than 5x10-8 for any of these phenotypes. These are only exploratory analyses and will only serve to identify additional potential gene loci that may regulate response to specific diets and should be studied in further, larger studies in the future. There are currently almost 200 such independent genetic variants that have been discovered and we anticipate that approximately 250 or more may be available in the next year or so. The group lasso approach is more sophisticated than one that assesses each SNP alone. First, it avoids the multiple comparison issue involved with testing 200-250 hypotheses. Second, we expect some degree of correlation among the SNPs and are particularly interested in whether we can enhance the predictive ability of the previously

identified genotype pattern with additional SNP information. We therefore would like to jointly consider or incorporate additional SNP information. Group lasso improves upon the SNP-by-SNP approach by jointly penalizing regressions so that overfitting is minimized. The variables considered will be diet assignment, each SNP, and its interaction with diet assignment. Those SNPs (and possibly their interactions with diet assignment) that are exceptionally useful for predicting 12-month weight change will be identified. The use of cross-validation to select the penalization parameter assures us that the model will not be overfit to the data. The sparsity of the coefficients makes the results easier to interpret. The SNPs and their interactions with diet that have the largest non-zero coefficients would be identified as the most important factors in a parsimonious description of the association. This analysis will thus reveal whether some group of SNPs & their interaction with diet assignment accounts for a significant portion of the variability in weight change. A further critical step (assuming the 3-locus genotype is found to be an important predictor in Aim 1) will then be to evaluate the contribution of the newly identified SNPs (expressed as a weighted linear combination or score as informed by group lasso) in predicting weight change in the presence of the multi-locus genotype pattern. Using cross-validation methods, we will assess whether the prediction error corresponding to a model with the multi-locus genotype, diet and their interaction that additionally includes the newly derived score identified by the group lasso method is significantly lower than the error resulting from a model that only includes the multi-locus genotype, diet assignment and their interaction.

The objective of Aim 3 is to provide further insight into possible mechanisms of action by identifying potential mediators of the relation between weight loss and appropriate assignment. For this purpose we will use contemporary mediation analysis techniques.136-138 Selected mediators for these analyses will focus on energy balance: energy intake (Kcal/day, appetite, satiety, hunger), and energy expenditure (resting energy expenditure, physical activity), and on insulin sensitivity/resistance.

Missing Data, Drop-Outs, and Intent-to-Treat. Analyses that do not account for missing data can lead to biased and inefficient estimates. To address such issues, we will perform both a complete-case analysis that excludes individuals missing at least one variable in the model as well as a multiple imputation-based model. Multiple imputation provides statistically valid results when the data are missing at random (i.e., the reason for missingness is related to observed variables only).139 Although the data may not be missing at random (e.g., if people who lose less weight are more likely to drop out), conditioning on baseline weight and/or subsequent weight measurements available make the missing at random assumption more reasonable. We will compare results from both strategies as a sensitivity analysis. Regardless of the level of dietary adherence to the intervention diets, we will analyze data collected from all randomized subjects and consider randomized intervention group as the intervention assignment, in accordance with intent-to-treat principles.

Explanation of any interim analyses that might be planned, when applicable;

None.

Description of the method used to generate the random allocation sequence;

Random assignments will be generated using the blockrand function from the blockrand package in R.

Specification of the type of randomization, such as blocked or paired, details of any restriction (such as blocking and block size) and disclosure and explanation if

optimization is used rather than randomization;

Randomization will be blocked, and the block size is 4 subjects.

Description of the mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned;

Randomized assignments will be determined and stored on the server of our Quantitative Sciences Unit (QSU) until we request participant randomization. Assignments are "blue" or "purple" to signify one of two diets, low-carb or low-fat. Participants will not be randomized until the majority of their baseline data collection is completed. Following randomization, participants will only be told of their date and time of class. They will not be told of their assigned diet until the first night of class.

Description of who will generate the random allocation sequence, who will enroll participants, and who will assign participants to interventions;

QSU will generate random allocation sequence to determine which subjects will be assigned which intervention. The Study Coordinator will tell participants of their class date and time. The Health Educator will reveal their diet assignment in class.

Description of who will be blinded after assignment to interventions (e.g., participants, care providers, those assessing outcomes) and how, if applicable;

Staff not involved in delivering the intervention will be blinded.
All staff from laboratories doing analyses of blood samples will be blinded.
Health educators, participants and study coordinator are the only individuals unblinded to diet.
All other team members, (including data collection, statistical teams, and PI) will remain blinded to the diet and just know assigned color.