PRECISION NUTRITION:

SEX, STRAIN, AND DIET DEPENDENT RESPONSE TO CARBOHYDRATE RESTRICTION IN MICE

A Dissertation

by

ANNA CAMILLE SALVADOR

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	David W. Threadgill
Committee Members,	Karen Kubena
	Stephen B. Smith
	Chaodong Wu
Head of Department,	David W. Threadgill

December 2021

Major Subject: Nutrition

ABSTRACT

The nutrition care process encompasses the provision of nutrition care among four domains: Assessment, Diagnosis, Intervention, and Monitoring & Evaluation. As a component of precision medicine, precision nutrition will build upon the nutrition care process by taking genetic background into consideration along with the typical information collected during Assessment. We searched for genetic variants underlying differences in the responses to American and ketogenic diets between C57BL/6J (B6) and FVB/NJ (FVB) mice. We used an F2 population derived from these strains to investigate the genetic origin of differential response to carbohydrate restriction. We identified three loci regulating fat mass gained during the feeding trial (*Fmgq1, Fmgq2,* and *Fmgq3*). The confidence interval for *Fmgq1* overlaps with a locus regulating serum HDL cholesterol concentration (*Hdlq1*) that harbors *Apoa2* and has previously been associated with serum HDL cholesterol concentration. *Fmgq1* may influence fat gain through an intermediate change in serum cholesterol. We also identified candidate genes at *Fmgq1* and *Fmgq2* associated with male hormone secretion. Linkage analysis for microbial traits identified genotype specific loci regulating microbial abundances as well as genotype-by-diet interactions and genotype-by-sex interactions. Lastly, we used our model to implement precision nutrition by exposing B6 and FVB mice to an American diet prior to the introduction of a ketogenic dietary intervention and observed low plasticity of the B6 male obese phenotype. These results demonstrate how precision nutrition will be advanced through integration of genetic variation and sex in physiological responses to diets varied in carbohydrate composition.

DEDICATION

T. COYNE, Popi & Dee, and Charlie

ACKNOWLEDGEMENTS

My parents have provided endless encouragement and support throughout my academic career. My dad has lived through every single moment of this experience with me, big and small. My mom instilled me with skills for strategic planning and self-discipline from an early age. My brother, the first scientist in our family, has always graciously allowed me to follow in his footsteps. My friends and family have tolerated a great deal of my absence over the last five years and have never waivered in their support of my goals. This work would not have been possible without that support.

I am thankful for the mentorship that I received during my undergraduate career from Ms. Eljeana Quebedeaux, Dr. Mark Merchant, and Dr. Jeffrey Stevens as well as Ms. Karen Geismar before and during my dietetic internship. Their guidance helped to make this dream a reality. My dissertation committee has donated countless hours of their time to making me a better researcher and I value the time I was able to spend with them.

Gratitude is not a strong enough word for the way I feel about having had the opportunity to work with Dr. David Threadgill. His laboratory quickly became my home and he and my lab-mates have become a part of my extended family. I have been in the company of excellent scientists and even better people. During this time, I have had the opportunity to work with several bright students including Anthony Gacasan, Ryan McGovern, Laura Deus Ramirez, Edeline D'Souza, Xavier Balboa, Gabrielle Manno, Audrey Watts, Thomas Wong, Hannah Field, Jace McGovern,

and Carly De La Garza. I would like to especially acknowledge Anthony Gacasan and Ryan McGovern for their patience and dedication to these projects during my early graduate training.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Dr. David Threadgill (advisor), Dr. Karen Kubena, Dr. Stephen B. Smith, and Dr. Chaodong Wu of the Department of Nutrition. Dr. Aniko Sabo donated her time and expertise along with the committee from the Department of Molecular and Human Genetics at Baylor College of Medicine.

Funding Sources

This work was supported by National Institutes of Health (NIH) grants RM1HG008529 and P30ES029067.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I: INTRODUCTION

Introduction

The nutrition care process is a global standard for the provision of nutrition care by Registered Dietitians. Provision of nutrition care is divided among four domains: Nutrition Assessment, Nutrition Diagnosis, Nutrition Intervention, and Nutrition Monitoring and Evaluation. During nutrition assessment an individual's eating behaviors, their access to foods, physical activity levels, physiological responses to diets, and body composition are compared to population level, comparative standards[1,2].

This information is used to develop a nutrition diagnosis that is distinct from a medical diagnosis. A medical diagnosis of chronic kidney disease may come along with a nutrition diagnosis like excessive consumption of potassium. A nutrition intervention is developed to directly address the problem identified by the nutrition diagnosis. The intervention might involve modification of food and nutrient delivery, nutrition education, nutrition counseling, or coordination of nutrition care from other members of the healthcare team. The nutrition care process is a cyclic process with continuous monitoring and evaluation of nutrition care indicators to assess the appropriateness and effectiveness of a chosen nutrition intervention as well as to determine when and if nutrition intervention can be discontinued.

The nutrition care process is often described as an individualized or personalized provision of nutrition care. While unique data is collected from each individual, the guidelines and comparative standards used to evaluate the data are developed at the

population level. Empirically, we see that the population appears to segregate into responders and non-responders to nutrition interventions. Occasionally non-response to a dietary intervention is as simple as a lack of adherence. Sometimes an individual is unable or unwilling to comply with the chosen nutrition intervention because of a lack of resources or lack of support. More recently, we are interested in elucidating genetic and environmental factors to explain the lack of response in two thirds of the population and ideally, work towards generating individualized guidelines to better meet their needs. Precision medicine has successfully taken individual characteristics such as genetics, diet, and lifestyle into consideration for the treatment of disease. As a component of precision medicine, precision nutrition will build upon the nutrition care process by taking genetic background into consideration along with the traditional information collected during Nutrition Assessment.

Carbohydrate restriction is a widely used dietary intervention for treatment of Diabetes Mellitus and other chronic disease states associated with metabolic syndrome. The ketogenic diet represents the most severe case of carbohydrate restriction and is being adopted rapidly by metabolically healthy and unhealthy individuals. A historical study demonstrated that only one third of individuals with Diabetes Mellitus responded well enough to carbohydrate restriction that no additional pharmaceutical intervention would be necessary[3]. These results are not surprising. We often summarize findings such as these as an indication that the dietary intervention is likely to be effective for the average individual especially if a higher proportion of individuals respond to carbohydrate restricted diets rather than another comparative diet. However, if one considers that the

same report is indicating that *two thirds of individuals failed to respond to carbohydrate restriction,* we find profound evidence for the risk of ineffective, population level dietary recommendations for two thirds of the population.

Meta-analyses have consistently shown that both carbohydrate restricted and fat restricted diets are effective means of reducing body weight with minimal differences, if any, in efficacy of the two interventions[4]. The Preventing Overweight Using Novel Dietary Strategies (POUNDS LOST) Trial was a 2-year randomized clinical trial of four energy-restricted diets that varied in macronutrient composition including, a fat restricted and carbohydrate restricted diet. Numerous subgroup analyses have been performed on the POUNDS LOST trial data where the population studied has been stratified by genotype at loci that have been associated with metabolic perturbations and obese phenotypes.

Stratification of the study population by *IRS1* (rs2943641) genotype has shown that individuals carrying the CC genotype respond best to a low-fat dietary intervention, while stratification of the study population by *PPM1K* (rs1440581) genotype has shown that individuals harboring the T allele will respond best to a carbohydrate restricted diet[5–7]. The observations made in the POUNDS LOST trial are landmarks along the path towards identifying subgroups of the population for which to develop individualized guidelines. However, as promising as these results were, the results are confounded by energy-restriction and not all prospective gene-diet interactions resulted in sustained weight loss throughout the 2-year period. To date, no single-gene diet interactions like the ones tested by POUNDS LOST have had the utility to combat the obesity epidemic. As with other

responses to other environmental factors, responses to diet are likely not due to the action of single genes with high impact but instead the culmination of multiple alleles, interactions among alleles, epigenetic modifications, and sex. Weight re-gain is common after trials of dietary intervention and may reflect high heritability and low plasticity of obese phenotypes and metabolic perturbations[8].

Precision nutrition will build upon the nutrition care process by taking individual characteristics such as genetics and sex in addition to the typical information collected about disease status, dietary intake, and lifestyle into consideration for the treatment of nutrition-related disease. Precision nutrition is poised to revolutionize the practice of dietetics. Nutrition scientists and forward-thinking dietetics practitioners have been in search of molecular mechanisms that characterize responders from non-responders for decades.

After efforts over the last 40 years, several gene-diet interactions have been identified, but no robust molecular signature of responders and non-responders to carbohydrate restriction have been identified[9]. To distinguish between responders and non-responders to dietary interventions, we must steer away from population level comparative standards like the ones that are currently used during Nutrition Assessment and conduct data analyses at the individual level, incorporating as much information about genetics, sex, disease status, dietary intake, and lifestyle as is available to us. Current direct-to-consumer tests claiming to provide dietary guidelines based on genetic markers lack

scientific rigor but demonstrate that there is high interest among consumers for these services.

The transition towards precision nutrition will require more than refining reference ranges and comparative standards. The search for a robust molecular signature of response to diets varied in macronutrient and even micronutrient continues to be a priority. An easily accessible biomarker must be incorporated into the Nutrition Assessment to parse the population into subgroups for which reference ranges and comparative standards can be developed. This information would be used to identify a molecular signature during Nutrition Diagnosis and integrate genetic counseling into the Nutrition Intervention with truly individualized recommendations.

The mouse represents a unique opportunity to model the implementation of the precision nutrition paradigm. However, translating preclinical research to clinically actionable recommendations has been difficult since until recently, studies rarely consider the impact of genetic variation on diet response. Most studies use a single inbred mouse strain, resulting in over-estimation of individual effects of genes and lack of translation to humans.

More recently, our laboratory has used multiple, genetically distinct, inbred mouse strains to characterize inter-strain differences in response to variable amounts of dietary carbohydrate as seen in response to American (high fat, high carbohydrate) and ketogenic (high fat, very low carbohydrate) diets. We have used this information to

characterize the role of sex, diet, genetics, and interactions among these factors in response to carbohydrate restriction in C57BL6/J (B6) and FVB/NJ (FVB) mice, especially B6 males exposed to the ketogenic diet.

We used an F2 population derived from B6 and FVB mice to investigate the genetic origin of differential response to carbohydrate restriction and sensitivity to high fat diets observed in our preliminary studies. Genome-wide linkage analysis in this population revealed multiple quantitative trait loci (QTL) regulating metabolic and microbial traits in mice exposed to American and ketogenic diets. Metabolic traits in our analyses included changes to body composition during the feeding trial and serum cholesterol concentration after exposure to these high fat diets that are varied in carbohydrate content. Microbial abundances were measured by 16S ribosomal RNA sequencing in feces collected at the end of the feeding trial.

We identified three loci regulating fat mass gained during the feeding trial (*Fmgq1, Fmgq2,* and *Fmgq3*). While *Fmgq2* and *Fmgq3* affect fat gain directly, it appears that *Fmgq1* may influence fat gain through an intermediate change in serum cholesterol. The confidence interval for *Fmgq1* overlaps with a locus regulating serum HDL cholesterol concentration (*Hdlq1*), which harbors *Apoa2* and has previously been associated with serum HDL cholesterol concentration in these strains on high fat diets[10,11]. The most significant locus associated with the amount of fat mass gained during the feeding trial, *Fmgq2*, explains a greater proportion of variation in male F2s exposed to the ketogenic diet. This supports sex differences observed in parental and F1 populations of B6 and

FVB mice from our previous work. We identified candidate genes at *Fmgq1* and *Fmgq2* associated with male hormone secretion that may explain the differential response to carbohydrate restriction or the consequences associated with mixing of B6 and FVB genetic backgrounds during high fat diet exposures.

Linkage analysis for microbial traits identified genotype specific loci regulating microbial abundances as well as genotype-by-diet interactions and genotype-by-sex interactions. We distinguished operational taxon units (OTU) that were regulated at each of these levels. Microbial abundances that are sensitive to genotype-by-diet interactions would be most clinically relevant for identifying subgroups of the population to develop individualized dietary guidelines to promote or reduce abundances of these organisms.

Lastly, we used our model to implement precision nutrition by exposing B6 and FVB mice to an American diet to induce obese phenotypes prior to the introduction of a ketogenic dietary intervention. In contrast to the lifetime dietary exposures that are typically used in preclinical research, this approach is more relevant to human clinical trials where participants are enrolled with obesity and/or another metabolic perturbation and interventions are trialed to attenuate the consequences of exposure the unhealthy diets.

Considering all preliminary evidence, we expected that B6 males would respond best to the dietary intervention but found that the obese phenotype appears to be more plastic in B6 females. Fat mass at the end of the feeding trial was highly correlated to the amount of fat mass after three months on the American diet, just prior to the dietary intervention

rather that caloric intake across all strain, sex, and diet combinations. The severe amount of adipose accumulated in B6 males prior to the dietary intervention may molecularly or physiologically reduce the plasticity of the dietary intervention. Future work will investigate molecular and physiological reasons behind the reduced plasticity of the obese phenotype in B6 males and investigate timing and synergy among interventions to reduce fat mass gained after the three-month exposure to the American diet.

References

1. Lacey K, Pritchett E. Nutrition Care Process and Model: ADA adopts road map to quality care and outcomes management. *Journal of the American Dietetic Association*. 2003;103(8). doi:10.1016/S0002-8223(03)00971-4
2. Hammond MI, Myers EF, Trostler N. Nutrition care process and model: An academic and practice odyssey. *Journal of the Academy of Nutrition and Dietetics*. 2014;114(12). doi:10.1016/j.jand.2014.07.032
3. Perkins JR, West TET, Sönksen PH, Lowy C, Iles C. The effects of energy and carbohydrate restriction in patients with chronic Diabetes Mellitus. *Diabetologia*. 1977;13:607-614.
4. Johnston BC, Kanters S, Bandayrel K, et al. Comparison of weight loss among named diet programs in overweight and obese adults: A meta-analysis. *JAMA - Journal of the American Medical Association*. 2014;312(9). doi:10.1001/jama.2014.10397
5. Heianza Y, Qi L. Gene-diet interaction and precision nutrition in obesity. *International Journal of Molecular Sciences*. 2017;18(4). doi:10.3390/ijms18040787
6. Qi Q, Bray GA, Smith SR, Hu FB, Sacks FM, Qi L. Insulin receptor substrate 1 gene variation modifies insulin resistance response to weight-loss diets in a 2-year randomized trial the preventing overweight using novel dietary strategies (POUNDS LOST) Trial. *Circulation*. 2011;124(5). doi:10.1161/CIRCULATIONAHA.111.025767

7. Xu M, Qi Q, Liang J, et al. Genetic Determinant for Amino Acid Metabolites and Changes in Body Weight and Insulin Resistance in Response to Weight-Loss Diets. *Circulation*. 2013;127(12). doi:10.1161/circulationaha.112.000586

8. Bradley P. Refined carbohydrates, phenotypic plasticity and the obesity epidemic. *Medical Hypotheses*. 2019;131. doi:10.1016/j.mehy.2019.109317

9. Wylie-Rosett J, Aebersold K, Conlon B, Isasi CR, Ostrovsky NW. Health effects of low-carbohydrate diets: Where should new research go? *Current Diabetes Reports*. Published online 2013. doi:10.1007/s11892-012-0357-5

10. Wang X, Paigen B. Genetics of variation in HDL cholesterol in humans and mice. *Circulation Research*. Published online 2005. doi:10.1161/01.RES.0000151332.39871.13

11. Wang X, Korstanje R, Higgins D, Paigen B. Haplotype analysis in multiple crosses to identify a QTL gene. *Genome Research*. Published online 2004. doi:10.1101/gr.2668204

CHAPTER II: SEX-SPECIFIC GENETIC ARCHITECTURE IN RESPONSE TO AMERICAN AND KETOGENIC DIETS*

Abstract

There is a growing appreciation for individual responses to diet. In a previous study, mouse strain-specific responses to American and ketogenic diets were observed. In the current study, we searched for genetic variants underlying differences in the responses to American and ketogenic diets between C57BL/6J (B6) and FVB/NJ (FVB) mouse strains. Genetic mapping of fat and lean mass gain revealed QTLs on Chromosome (Chr) 1 at 191.6 Mb (*Fmgq1*) ($p<0.001$, CI=180.2-194.4 Mb), Chr5 at 73.7 Mb (*Fmgq2*, *Lmgq1)* ($p<0.001$, CI=66.1-76.6 Mb), and Chr7 at 40.5 Mb (*Fmgq3*) ($p<0.01$, CI=36.6-44.5Mb). Analysis of serum HDL cholesterol concentration identified a significant ($p<0.001$, CI=160.6-176.1 Mb) QTL on Chr1 at 168.6 Mb (*Hdlq1*). Causal network inference suggests that HDL cholesterol and fat mass gain are both linked to *Fmgq1*. Strong sex effects were identified at both *Fmgq2 and Lmgq1*, which are are also diet-dependent. Interestingly, *Fmgq2 and Fmgq3* affect fat gain directly, while *Fmgq1* influences fat gain directly and possibly via an intermediate change in serum cholesterol. These results demonstrate how precision nutrition will be advanced through integration of genetic variation and sex in physiological responses to diets varied in carbohydrate composition.

*Reprinted with permission from "Sex-specific genetic architecture in response to American and ketogenic diets" by A.C. Salvador, et. al., 2021. *International Journal of Obesity*, Volume 45, 1284-1297.

with small effect sizes that are not clinically actionable (1–3). As with studies investigating responses to other environmental factors, responses to diet are likely not due to the action of single alleles with a large effect, but rather the sum of multiple small effect alleles, interactions among alleles, epigenetic modifications, and sex.

Recently, our group observed strong mouse strain-specific differences in response to feeding different human-relevant diets (4–6). This initial study provided evidence for striking differences between C57BL/6J (B6) and FVB/NJ (FVB) mice in response to high fat diets varying in carbohydrate content. Although studies increasingly consider the role of sex as a biological variable, the role of sex in the response to diets with varied macronutrient contents has been understudied (7).

To further investigate the genetic origin of differential response to carbohydrate restriction and sensitivity to high fat diets, we generated an intercross population (F2) between B6 and FVB. All F2s were fed either an American (35% of energy from fat, 50% from carbohydrate) or a ketogenic (84% of energy from fat, 0% from carbohydrate) diet and changes to body composition and serum cholesterol concentrations were measured. In this population, we performed a genome-wide linkage analysis to elucidate the genetic architecture contributing to differential responses to the specific diets. The data obtained provide evidence for genetic loci that not only directly affect body composition response, but also loci that indirectly affect response through differences in serum cholesterol concentration. Additionally, significant sex differences in the effects of the identified quantitative trait loci (QTL) were detected.

Methods

Animals and diets

Initially, we screened 6 week-old B6 and FVB mice for their response to American (35% of energy from fat, 50% from carbohydrates) and ketogenic (84% of energy from fat, 0% from carbohydrates) diets after a 6-month feeding trial. These strains have previously exhibited significantly different responses to these two diets (4). Detailed diet compositions are provided in Appendix Table A1. B6 females were crossed with FVB males to generate F1 mice and subsequently intercrossed to generate the F2 population. Four-week-old F1s and 3-5 week-old F2s were screened for their response to American and ketogenic diets during a 3-month feeding trial.

For the feeding trials, mice were randomly assigned to one of the two diet groups. Half of the B6, FVB, F1, and F2 mice were placed on American diet (B6: 11 males, 9 females, FVB: 10 males, 10 females, F1: 6 males, 6 females, F2: 102 males, 122 females) and half on ketogenic diet (B6: 9 males, 10 females, FVB: 10 males, 10 females, F1: 6 females, 9 males, F2: 126 males, 119 females). Researchers were not blinded to diet assignments. All animals were maintained in accordance with Texas A&M University Institution Animal Care and Use Committee guidelines at 22 °C under a 12-hour light cycle. At the end of the feeding trial, mice were euthanized by carbon dioxide asphyxiation, blood was collected, and tissues were harvested and immediately flash frozen in liquid nitrogen.

Phenotyping

Echo Magnetic Resonance Spectroscopy (MRI) (EchoMRI, Houston, TX, USA) was used to measure fat and lean mass of all individuals. During the initial screen of the B6 and FVB strains, body weight and body composition were measured at a 3-month time-point of the 6-month feeding trial. In the F1 population, body weight and body composition were measured at the beginning and end of the 3-month feeding trial. Fat percentage of total body weight was calculated at the 3-month time-point for B6, FVB, and F1 populations. Fat percentage is the percentage of total body fat mass measured by MRI relative to body weight at the time of the MRI measurement. In the F2 population, body weight and body composition from before and after the 3-month feeding trial allowed for changes in fat and lean mass to be calculated. Fat and lean mass gains were calculated as the difference in fat and lean mass prior to the feeding trial and after 3 months on the assigned diet.

In the F2 population, total cholesterol, HDL and LDL fractions, as well as APOA2 were measured in serum obtained from blood at sacrifice at the end of the feeding trial. Total cholesterol, HDL, and LDL measurements were performed in duplicate using the EnzyChrom AF HDL and LDL/VLDL Assay kit (BioAssay Systems, Hayward, CA, USA). APOA2 measurements were performed in duplicate in a subset of the F2 population with the highest and lowest serum HDL cholesterol concentration (11 males, 27 females), using the Mouse Apolipoprotein A2 ELISA kit (ABclonal cat # RK02605, Woburn, MA, USA).

Genotyping

The F2 population was genotyped on the Mouse Universal Genotyping Array (MUGA) that includes 7854 SNP markers (8). Markers that were not polymorphic between B6 and FVB were removed from the data set and uncertain genotype calls for individuals (GenCall score quality metric <0.7) were set to missing. The remaining markers were used to generate a genetic map to check for problematic markers and/or sample DNAs. After all corrections, 1667 markers were used for the association analyses. Updated MUGA marker annotation was obtained from Dr. Karl Broman (https://kbroman.org/MUGAarrays/new_annotations.html).

Heritability calculations

Broad-sense heritability for body fat percentage was calculated as the ratio of total genetic variance (variance in the F2 population) to environmental variance (variance in the F1 population).

Quantitative Trait Loci Analysis

Outliers can have a strong influence on the results of QTL analyses. We observed biologically implausible errors in data that were greater than 3 standard deviations from the mean and suspect that these reflect technical error in the measurements. As such, outliers were defined as individuals with phenotypes that were more than three standard deviations away from the mean for each sex and phenotype in the F2 population. Outliers were set to missing. This procedure was repeated iteratively to discover all outliers in the

data. Pearson's correlation between phenotypes was determined after correcting for sex and diet effects (Appendix Figure A1). Code available by request.

The combined model includes all F2s, both sexes, and both diets (y~ sex * diet + marker). QTL peaks with a logarithm of the odds (LOD) greater than thresholds determined by 10,000 permutations were considered genome-wide significant ($p<0.05$, LOD > 3.90) or highly significant ($p<0.01$, LOD > 4.70). A LOD drop of 1.5 LOD from the top marker was used to determine the 95% confidence intervals, or support intervals, for each QTL. Linear models using ANOVA was used to check for any interactions between sex and/or diet with the top markers of each QTL. The variance explained by the top markers at each QTL was calculated by dividing the sum of squares of the model including the top marker by the total sum of squares of the model without QTL.

Candidate gene analysis using KEGG

KEGG pathways are a collection of pathway maps that reflect known genetic and metabolic relationships. All genes within each significant QTL confidence interval were annotated with KEGG pathway identifiers. We further characterized our candidate genes by KEGG pathways related to glucose, insulin, fatty acids, adipocytes, cholesterol, obesity, diabetes mellitus, metabolic syndrome, and digestion and absorption of carbohydrates, fats, and proteins. A comprehensive list of KEGG pathway queries is provided in Appendix Table A2.

Conditioned and Unconditioned Linkage Analysis

We performed a set of conditional QTL scans using traits as covariates in the analysis of other traits based on the individual QTL that were identified for each trait and suspected biological relationships. If conditioning the genome scan with a covariate resulted in a significant increase or decrease in the absolute value of the LOD score, it was interpreted that the traits are causally related to one another. When comparing conditioned and unconditioned QTL scans, an increase or decrease in the LOD of at least 2.0 corresponds to a 5% type I error rate (9,10). Only conditioned genome scans that resulted in a change of 2.0 LOD were considered pleiotropic, or shared, QTL between the two traits.

Causal Network Inferences

For traits with overlapping QTL we made causal inferences in networks based on methods described elsewhere (11). Briefly, the first trait (T1) was regressed on the second trait (T2) and T2 was regressed on T1 in order to obtain the residual of each trait after adjusting for the other (R1 and R2). A bivariate t-test between R1, R2 and the shared locus was used to infer the causal network among them. P-values were Holm-Bonferroni corrected for the number of tests (*i.e.* number of residuals tested = 2) where $p = \alpha_{0.05}$ / (number of tests – rank of i^{th} hypothesis + 1). A p-value < 0.05 is considered significant. An initial pathway describing the relationship between T1 and T2 was defined based on the inferred causal networks, and the predicted causal models were compared with the Akaike's information criterion (AIC) and Bayesian information criterion (BIC). A lower AIC and BIC indicates a better model.

Structural Equation Modeling

Structural equation modeling was used to illustrate observed and unobserved relationships between the QTL models. The models were refined until all path coefficients were significantly different from zero. Linear models using ANOVA was used to check for the amount of variation explained by each predictor in the structural model. The proportion of variation explained with the predictors modeled was calculated by dividing the sum of squares of the model including the predictor by the total sum of squares without the predictor.

Quantitative PCR (qPCR)

RNA from flash-frozen gonadal fat and liver were extracted using the simplyRNA Tissue kit on a Maxwell AS3000 (Promega). cDNA was generated using the Transcriptor First Strand cDNA Synthesis kit (Roche). qPCRs were performed on a LightCycler 480 (Roche) and CFX384 (BioRad) with Light Cycler 480 SYBR Green I Master reagent (Roche). *B2m* was used as a housekeeping gene to correct for starting amounts of cDNA in both gonadal fat and liver. Primers were purchased from Integrated DNA Technologies as custom DNA oligos and sequences are provided in Appendix Table A3.

Results

Genetic background and sex modulate parental strain fat differences

The fat percentage of B6 males consuming an American diet was 1.77-fold higher compared to B6 males consuming a ketogenic diet (B6 male American: 27.4% +/- 5.2%, B6 male ketogenic: 15.5% +/- 7.1%, p=0.001, CI=5.8 - 18.0%; Figure 1, Appendix Table

A4). Conversely, B6 females, and FVB males and females fed the same two diets did not show diet-dependent differences of body fat percentage (B6 female American: 20.7 +/- 8.8%, B6 female ketogenic: 17.2% +/- 5.7%; FVB male American: 18.1% +/- 6.9%, FVB male ketogenic: 20.4% +/- 7.5%; FVB female American: 13.4% +/- 6.7%, FVB female ketogenic: 16.2% +/- 6.4%; Figure 1, Appendix Table A4). The mean fat percentage in B6 males on the ketogenic diet is lower than the fat percentage in FVB males on the ketogenic diet, although the difference is not statistically significant (Figure 1, Appendix Table A4).

Sex modulates hybrid population fat differences

The F1 males, but not females, also responded to ketogenic diet. F1 males on the American diet had 1.25-fold higher body fat percentage than on the ketogenic diet (F1 male American: 32.5% +/- 3.2%, F1 male ketogenic: 25.9% +/- 4.7%, p=0.020, CI=1.3 – 11.9%; Figure 1, Appendix Table A4). We also observed that the F1 males on the ketogenic diet had a 1.67-fold higher fat percentage than that observed in B6 males on the ketogenic diet (F1 male ketogenic: 25.9% +/- 4.7%, B6 male ketogenic: 15.5% +/- 7.1% p=0.005, CI=3.7 – 15.6%; Figure 1, Appendix Table A4).

Sex and diet modulate phenotypes in the F2 population

As expected, sex had a profound effect on the phenotypes used for analysis (Appendix Table A4). This effect was greater than the effect of diet for all phenotypes (Appendix Table A4). Regardless of the carbohydrate composition of the diet, females had lower fat mass gain than males (F2 females, 4.5 g +/- 3.0 g, F2 males, 8.2 g +/- 4.5 g, p<0.001,

CI=2.8-4.2 g), lower lean mass gain than males (F2 females, 5.5 g +/- 1.6 g, F2 males, 9.0 g +/- 3.2 g, $p<0.001$, CI=3.0-3.9g), and lower serum HDL cholesterol concentration than males (F2 females, 157.6 ng/mL +/- 58.3 ng/mL, F2 males, 201.6 ng/mL +/- 32.5 ng/mL, $p < 0.001$, CI=34.9 – 53.0 ng/mL).

Diet had a significant effect on lean mass gain and serum HDL cholesterol (Appendix Table A4). Irrespective to sex, F2s on the ketogenic diet gained less lean mass than F2s on the American diet (F2 ketogenic, 6.9 ng/mL +/- 2.8 ng/mL, F2 American, 7.4 ng/mL +/- 3.2 ng/mL, $p=0.002$, CI = 0.3 – 1.2 ng/mL) and had lower serum HDL cholesterol concentration on the ketogenic diet compared to F2s on the American diet (F2 ketogenic, 175.1 ng/mL +/- 52.2 ng/mL, F2 American, 183.3 ng/mL +/- 52.4 ng/mL, $p=0.027$, CI=1.2-19.2 ng/mL).

A significant interaction of sex and diet (Appendix Table A4) was observed for lean mass gain where F2 males on the ketogenic diet gain significantly less lean mass than F2 males on the American diet (F2 male ketogenic, 8.2 g +/- 3.1 g, F2 male American, 10.0 g +/- 3.0 g, $p<0.001$, CI=0.9-2.7 g).

The amount of fat mass gain was highly heritable in the F2 population. Broad sense heritability for the trait was 81% in males and 71% in females.

Linkage analysis reveals QTLs for fat mass gain during the feeding trial

In the combined analysis that incorporates all F2s on both diets (y~ sex * diet + marker), we detected QTL for fat mass gain on Chr1 at 191.6 Mb (*Fmgq1*; p<0.001, CI=180.2-194.4 Mb), Chr5 at 73.7 Mb (*Fmgq2*; p<0.001, CI=66.1-76.6 Mb), and Chr7 at 40.5 Mb (*Fmgq3*; p<0.01, CI=36.6-44.5 Mb) (Figure 2A, Table 1). At *Fmgq1*, the FVB allele contributes to higher fat mass gain, and the top marker (UNC010475128) accounts for 3.44% of the variance. *Fmgq2* is a highly significant QTL where the FVB allele contributes to lower fat mass gain. The top marker (backupUNC050383757) accounts for 6.3% of the total variance in the F2 population (Table 1). The top marker explains 22.8% of the variation in males on the ketogenic diet, while it only explaining 5.9% of the variation in females on the ketogenic diet and 4.1% and 1.6% of the variation in males and females on the American diet, respectively. At *Fmgq3* the FVB allele contributes to lower fat mass gain, and the top marker (JAX00150446) accounts for 3.4% of the variance in fat mass gain.

We expected that FVB alleles would drive higher fat mass gain across the genome based on the initial screen of B6 and FVB showing that B6 males have a differential response to American and ketogenic diets, while FVB mice do not have a differential response to these two high fat diets. Surprisingly, the FVB allele contributes to lower fat mass gain in response to the ketogenic diet at the most highly significant QTL at *Fmgq2* as well as *Fmgq3*, but higher fat mass gain at *Fmgq1*.

In the initial analysis of fat percentage in B6, FVB, and F1 populations, we also observed that the male response to the ketogenic diet was greater in the F1 population compared to the B6 parent strain (Figure 1). This prompted us to test if the combination of parent alleles at *Fmgq1* with either *Fmgq2* or *Fmgq3* resulted in higher fat mass gain in the F2 population. We found an interaction between *Fmgq2* and *Fmgq1* (p=0.007, CI=1.58 – 9.07 g) affecting fat mass gain. The interaction effect was specific to males on the ketogenic diet. Males that are homozygous for the B6 allele at *Fmgq2* and that carry at least one FVB allele at *Fmgq1* gained about 5g more fat on ketogenic diet than males homozygous for the B6 allele at both loci (Figure 2B, C)

Linkage analysis reveals QTLs for lean mass gain during the feeding trial

In the combined model that includes male and female F2s on both diets (y~ sex * diet + marker), we detected a significant QTL for lean body mass gain during the 3-month feeding trial (*Lmgq1;* p<0.001, CI = 66.1-76.6Mb) (Table 1, Figure 2D). *Lmgq1* overlapped with *Fmgq2*. We observe again that males on the ketogenic diet drive this QTL. In the combined model, the top marker (backupUNC050383757) explains 3.3% of the variance in lean mass gain. In males on the ketogenic diet, the top marker (JAX00131700) explains 19.0% of the variation in lean mass gain while it only explaining 1.2% of the variation in females on the ketogenic diet and 8.2% and 4.1% of the variation in males and females on the American diet, respectively

Linkage analysis reveals QTLs for serum HDL concentration at the end of the feeding trial

In the combined model that includes male and female F2 animals on both diets (y~ sex * diet + marker), we detected a highly significant QTL on Chr1 at 168.6 Mb (*Hdlq1*) for serum HDL concentration after the 3-month feeding trial (p<0.001, CI=160.6-176.1 Mb; Figure 2E). All genotype classes (B6/B6 155.4 ng/mL +/- 48.8 ng/mL; B6/FVB 179.2 ng/mL +/- 51.5 ng/mL; FVB/FVB 200.6 ng/mL +/- 48.5 ng/mL) are significantly different from each other at this locus (B6/B6 : B6/FVB p<0.001, CI=9.7-38.1 ng/mL; B6J/B6J : FVB/FVB p<0.001, CI=28.7-61.8 ng/mL; FVB/FVB : B6/FVB p<0.001, CI=7.4-35.4 ng/mL; Appendix Figure A5). This locus harbors the *Apoa2* gene which has previously been associated with serum HDL cholesterol concentration (12–14). Therefore, we measured serum APOA2 concentrations in F2s with the highest and lowest serum HDL cholesterol concentrations. These measurements confirm significant differences between homozygous genotype classes at *Hdlq1* (B6/B6 2.2 mg/dL +/- 0.8 mg/dL; FVB/FVB 3.0 mg/dL +/- 0.9 mg/dL, p = 0.004, CI = 0.3-2.2 mg/dL; Figure 3A).

KEGG pathway annotation of candidate genes at Fmgq1 and Fmgq2 highlights steroid hormone biosynthesis

We searched for candidate genes that might elucidate the functional interaction between *Fmgq1* and *Fmgq2*. Out of 98 positional candidate genes, 4 genes at *Fmgq1* are found on metabolically relevant KEGG pathways: *Hsd11b1* (steroid hormone biosynthesis; mmu00140), *Ephx1* (bile secretion; mmu004976), *Camk1g* (aldosterone synthesis and

secretion; mmu004925), and *Ppp2r5a* (AMPK signaling pathway; mmu004152) (Table 2). Out of 482 positional candidate genes, 2 genes at *Fmgq2* are found on metabolically relevant KEGG pathways: *Srd5a3* (steroid hormone biosynthesis; mmu00140) and *Cox7b2* (non-alcoholic fatty liver disease; mmu004932) (Table 2).

On the steroid hormone biosynthesis pathway, *Srd5a3* converts inactive testosterone to active dihydrotestosterone. This makes *Srd5a3* a particularly interesting candidate at *Fmgq2* given the sex specificity we see at this QTL and the functional interaction we observed between *Fmgq2* and *Fmgq1*. Consistent with *Srd5a3* underlying *Fmgq2*, we observed lower expression of *Srd5a3* in gonadal fat of F2s that are homozygous for the FVB allele at *Fmgq2* relative to F2s that are homozygous for the B6 allele at *Fmgq2* (Figure 3B). Interestingly, we also observe a 1.59-fold increase in *Srd5a3* expression in F2s that are B6/B6 at *Fmgq2* and carry FVB/FVB alleles at *Fmgq1* relative to F2s that are B6/B6 at both loci (Figure3C). This makes *Srd5a3* a strong candidate gene at *Fmgq2* and for the functional interaction between *Fmgq1* and *Fmgq2*.

At *Fmgq1,* we did not observe any genotype dependent differences in expression of *Hsd11b1* (liver)*, Ephx1* (liver), or *Ppp2r5a* (gonadal fat) (Appendix Figure A6ABC). We did however, observe a 1.61-fold increase in *Ppp2r5a* expression in F2s that are B6/B6 at *Fmgq2* and carry FVB/FVB alleles at *Fmgq1* relative to F2s that are B6/B6 at both loci (Figure 3D). This makes *Ppp2r5a* a strong candidate gene at *Fmgq1* for the functional interaction between *Fmgq1* and *Fmgq2*.

The remaining candidate at *Fmgq1, Camk1g*, is exclusively expressed in brain tissue, which was unavailable to measure expression. Given involvement of *Camk1g* in aldosterone synthesis and secretion, we instead measured serum aldosterone concentration (after the feeding trial) and observed no genotype dependent differences (Appendix Figure A6D). The remaining candidate at *Fmgq2*, *Cox7b2* is expressed exclusively in the testis, which was unavailable to measure its expression.

Conditioned linkage analysis of fat and lean mass gain

Fmgq2 and *Lmgq1* overlap in the combined model of F2s that includes both sexes and diets. For both traits, this QTL appears to be driven largely by males on the ketogenic diet. This shared QTL prompted us to model the relationship between fat and lean mass gain. There is no significant change to the LOD score on Chr5 at 73.7 Mb when fat mass gain is conditioned on lean mass gain, nor when lean mass gain is conditioned on fat mass gain. Thus, the relationship between fat and lean mass gain at this locus could not be elucidated in this model. This would suggest that the overlapping QTL region harbors one or two tightly linked genes that affect both lean and fat mass in the same direction.

Conditioned linkage analysis of fat mass gain and serum HDL cholesterol concentration

The close proximity of *Hdlq1* and *Fmgq1* in the combined models prompted us to model the relationship between fat mass gain and serum HDL cholesterol concentration at these loci. We observed that when fat mass gain is conditioned on serum HDL cholesterol concentration, there is no significant change to the LOD score on Chr1 at 196.1 Mb (Table 3). Thus, the relationship between fat mass gain and serum HDL cholesterol

concentration could not be elucidated in this model. Similarly, when serum HDL cholesterol concentration is conditioned on fat mass gain, the LOD score on Chr1 at 168.6 Mb drops by only 1.08 LOD suggesting that *Hdlq1* is not shared between fat mass gain and serum HDL cholesterol concentration (Table3).

Causal network between Fmgq1, fat mass gain, and serum HDL cholesterol concentration

We also explored causal networks between fat mass gain and serum HDL cholesterol concentration at *Fmgq1*. After adjusting for multiple testing, the inferred network shows that serum HDL cholesterol concentration and fat mass gain are independently related to *Fmgq1* (Table 3). The AIC and BIC model selection scores suggest that the model in which serum HDL cholesterol concentration occurs upstream of fat mass gain at *Fmgq1* is most consistent with the data (Table 3).

Structural equation modeling of fat mass gain and serum HDL cholesterol concentration

We built a structural model to illustrate the magnitude of the effects of each predictor in the models of fat mass gain and serum HDL cholesterol concentration (Figure 3E). The path coefficients are all significantly different from zero (Table 3). The proportion of variation explained with the predictors modeled for fat mass gain and serum HDL cholesterol concentration is 29.95% and 27.00%, respectively.

Discussion

This study provides evidence that individual responses to high fat diet significantly depends upon three factors: (1) the presence or absence of carbohydrates in the high-fat diet; (2) the combination of alleles occurring in the study population; and (3) sex. In our experiment, B6 males have a lower percentage of body fat in response to the high fat, no carbohydrate, ketogenic diet, while in contrast, B6 females and FVB males and females do not have a differential response to American and ketogenic diets. This sex-specific response to carbohydrate restriction on ketogenic diet observed in B6 males persists in F1s males, and we observed that the combination of B6 and FVB alleles in the hybrids results in a higher percentage of body fat in response to the feeding trial. The difference in ages at the beginning of the parent and F1 feeding trials might also contribute to the higher percentage of body fat we observed in the F1s. In the F2 population, we observed a very high heritability for fat mass gain in both sexes and were able to identify significant QTLs for fat mass gain at *Fmgq1*, *Fmgq2*, and *Fmgq3*. This raises the question of whether these loci could be used to make predictions about individual response to carbohydrate restriction. *Fmgq2* explains the most variation in fat mass gain in males on the ketogenic diet. We observed that the FVB allele drives lower fat mass gain at *Fmgq2* and *Fmgq3*, while at *Fmgq1* the FVB allele drives higher fat mass gain. Additionally, we provided evidence that male F2s exposed to the ketogenic diet that are homozygous for B6 alleles at *Fmgq2* and carry at least one FVB allele at *Fmgq1* gain the most fat mass during the feeding trial. We observed that *Fmgq2* and *Lmgq1* represent the same QTL on Chr5 for fat and lean mass gain. Likely, a single gene or two tightly linked genes are responsible for the change in lean and fat mass.

The QTL for serum HDL cholesterol concentration at *Hdlq1* is significant in both males and females and contains *Apoa2. Apoa2* has been described repeatedly for its relationship to serum HDL cholesterol concentration, especially in these two strains (12–14). The initial observation of the *Apoa2* locus affecting serum concentrations of HDL cholesterol was made in an association study of females using an advanced intercross line between B6 and NZB/BINJ (12). Later, in a panel of inbred strains that included B6 and FVB mice, it was determined that the amino acid substitution Ala61Val in APOA2 led to increased serum HDL concentrations (13). B6 carry Ala61 while FVB mice carry amino acid substitution 61Val. Both males and females of all strains carrying the 61Val had increased serum HDL cholesterol concentrations. The results of our linkage analysis for serum cholesterol concentrations are consistent with the amino acid differences. They show that *Hdlq1* affects HDL in males and females. We confirmed that FVB alleles at *Hdlq1* drive higher concentrations of serum HDL cholesterol and serum APOA2.

In our causal network of the relationship between fat mass gain and serum HDL cholesterol concentration, we found that serum HDL cholesterol concentration and fat mass gain are independently linked to *Fmgq1*. The close proximity of the *Fmgq1* and *Hdlq1* QTLs put them in linkage disequilibrium and this makes their direct and indirect effects difficult to detangle. In general, we would have expected that serum HDL cholesterol concentration would have a negative relationship with fat mass gain given the well-established relationship between serum HDL cholesterol and obesity in humans. It is likely that the two traits are independently linked to *Fmgq1* because of the linkage

disequilibrium. However, the AIC and BIC scores for the models suggest that differences in serum HDL cholesterol concentration occur upstream of the differences we observe in fat mass gain at *Fmgq1*.

Overlaying genes within *Fmgq1* and *Fmgq2* onto metabolically relevant KEGG pathways revealed candidate genes involved in biogenesis of cortisol (*Hsd11b1*), aldosterone (*Camk1g*), and testosterone (*Srd5a3*). Cholesterol falls upstream of the production of each of these steroid hormones. *Ephx1* resides inside of *Fmgq1 and* plays a role in the transfer of bile acids to the liver, a process that is critical for the endogenous synthesis of cholesterol. Each of these candidate genes has in common, a relationship with cholesterol. This offers insight into the role these particular genes could play in the causal network we inferred that showed serum HDL cholesterol concentration occurring upstream of fat mass gain. The inferred network might be reflective of the relationship between fat mass gain and one of these candidate genes more so than a direct relationship between serum HDL cholesterol concentration and fat mass gain.

Further, our candidate genes offer insight into the sex specific interaction in males on the ketogenic diet that we observed between *Fmgq1* and *Fmgq2*. Our most prominent QTL at *Fmgq2* harbors *Srd5a3* that converts inactive testosterone to active dihydrotestosterone. We confirmed that F2s that are homozygous for FVB alleles at *Fmgq2* express *Srd5a3* less than F2s that are homozygous for B6 alleles. At this QTL, FVB alleles also drive lower fat mass gain. This suggests that more active dihydrotestosterone promotes more fat mass gain at *Fmgq2.* The functional interaction

we observed between *Fmgq1* and *Fmgq2* showed us that male F2s that are homozygous B6 at *Fmgq2* and carry at least one FVB allele at *Fmgq1* gained the most fat mass. Interestingly, we observed that F2s that carry this interaction have a 1.59-fold increase in expression of *Srd5a3* relative to animals that are homozygous B6 at both *Fmgq1* and *Fmgq2.* This provides further evidence that more active dihydrotestosterone promotes more fat mass gain in this population. *Ppp2r5a* at *Fmgq1* has been associated with polycystic ovary syndrome (PCOS) through pathway analysis of genes enriched in genome wide association studies (GWAS). *Ppp2r5a* is a member of the oocyte meiosis KEGG pathway that was significantly associated with PCOS, a pathway affected by the hyperandrogenemia that frequently occurs with PCOS (15). The association *of Ppp2r5a* with PCOS makes *Srd5a3,* and *Ppp2r5a* primary candidate genes of interest for further investigation.

Each of the remaining candidate genes has a potential relationship with *Srd5a3* and active dihydrotestosterone levels. *Hsd11b1* at *Fmgq1* reversibly converts active cortisol to inactive cortisone. Epidemiological data in humans shows that the ratio of androgens to glucocorticoids is critical to metabolic homeostasis and alterations in this ratio can lead to increased incidence of obesity and metabolic syndrome (16). Additionally, excessive levels of active dihydrotestosterone have been show to stimulate aldosterone secretion by a mechanism that is dependent upon the calmodulin/calmodulin-dependent protein kinase and protein kinase C intracellular signaling pathways (17). *Camk1g* at *Fmgq1* is a member of this pathway and people with obesity often experience hyperaldosteronism (18). *Ephx1* at *Fmgq1* has previously been associated with other *Srd5a* isoforms in

GWAS of PCOS (19), and similarly, *Cox7b2* at *Fmgq2,* has been identified in another GWAS for PCOS (20). The association with PCOS establishes a probable link between both *Ephx1* and *Cox7b2* and abnormal androgen secretion.

Sequence variation between B6 and FVB result in amino acid changes in EPHX1*,* CAMK1G*,* and PPP2R5A*,* sometimes at multiple locations within the proteins*.* In HSD11B1 and SRD5A3 no amino acid changes have been documented, but sequence variation in *Hsd11b1* and *Srd5a3* does result in non-coding transcript variants that might be of interest in further investigation. These non-coding transcript variants would suggest differences in expression of relevant candidate genes like we observed for *Srd5a3*. However, we did not observe genotype dependent differences in expression of *Hsd11b1,* but did find in the F2 population that mice that are B6/B6 at *Fmgq2* and carry FVB/FVB alleles at *Fmgq1* express *Ppp2r5a* 1.61-fold more than F2s that are B6/B6 at both loci.

Overall, our data parallels a clinical trial showing a greater fat mass loss among men than women on a very low energy, ketogenic diet (21–23). The differences could not be attributed to attenuation of appetite stimulating hormones, and concentrations of serum steroid hormones were not considered in the study. Unfortunately, studies in humans that utilize very low carbohydrate, ketogenic diets typically lack power to detect sex differences or fail to incorporate realistic control diets (24–28). Nonetheless, meta-analyses of human, genome-wide association studies of body mass index and waist-to-hip circumference have identified several sex specific loci (29,30). Future sex specific

genome-wide association studies that are powered to detect differential responses to macronutrients between men and women are needed.

Conclusions

It is possible that the genetic architecture for fat gain in response to high fat diets is more complex in females than in males, and as such, the genetic component we are able to detect in our model contributes more to the overall response in males than in females. This observation highlights the importance of sex specific analyses during the development of individualized dietary guidelines. We identified a known QTL at *Hdlq1* for serum HDL cholesterol concentration that harbors *Apoa2.* Interestingly, the QTLs at *Fmgq2* and *Fmgq3* affect fat gain directly while *Fmgq1* seems to influence fat gain directly and possibly via an intermediate physiological change in serum cholesterol related to the strain specific *Apoa2* phenotypes. We have shown that genotypes at *Fmgq2* alter expression of *Srd5a3* in a sex and genotype specific manner. It remains unclear if these loci explain the differential response to carbohydrate restriction we observed in the initial parent strain screen or if instead, they characterize the increased fat mass gain that we observe after the parental genomes are combined in hybrid and F2 populations. Further investigation of the candidate genes presented here will elucidate their roles in response to carbohydrate restriction in the parent, hybrid, and F2 populations.

Acknowledgements

We thank Dr. Andrew Hillhouse and the Texas A&M Institute for Genome Sciences and Society (TIGSS) for assistance in the TIGSS Molecular Genomics Workspace and TIGSS

Rodent Preclinical Phenotyping Core; Dr. Karl Broman for his discussion and guidance during the QTL analysis as well as for organizing and managing annotation files for the MUGA arrays; Daniel Genung and Aaron Van Wettering for assistance with mouse husbandry and phenotyping; Anthony Gacasan and Ryan McGovern for assistance with preparing samples for genotyping; and all members of the laboratory for their helpful insights. This work was supported by National Institutes of Health (NIH) grants RM1HG008529 and P30ES029067.

Competing Interests

The authors have no competing interests to declare.

References

1. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-analysis of genome-wide association studies for height and body mass index in ~700 000 individuals of European ancestry. *Hum Mol Genet.* 2018;27(20):3641–9.
2. Hoffmann TJ, Choquet H, Yin J, Banda Y, Kvale MN, Glymour M, et al. A large multiethnic genome-wide association study of adult body mass index identifies novel loci. *Genetics*. 2018;
3. Heianza Y, Qi L. Gene-diet interaction and precision nutrition in obesity. Vol. 18, *International Journal of Molecular Sciences.* 2017.
4. Barrington WT, Wulfridge P, Wells AE, Rojas CM, Howe SYF, Perry A, et al. Improving metabolic health through precision dietetics in mice. *Genetics*. 2018;208(1):399–417.
5. Wells A, Barrington WT, Dearth S, May A, Threadgill DW, Campagna SR, et al. Tissue Level Diet and Sex-by-Diet Interactions Reveal Unique Metabolite and Clustering Profiles Using Untargeted Liquid Chromatography-Mass Spectrometry on Adipose, Skeletal Muscle, and Liver Tissue in C57BL6/J Mice. *J Proteome Res*. 2018;
6. Cuomo D, Porreca I, Ceccarelli M, Threadgill DW, Barrington WT, Petriella A, et al. Transcriptional landscape of mouse-aged ovaries reveals a unique set of non-coding RNAs associated with physiological and environmental ovarian dysfunctions. *Cell Death Discov*. 2018;
7. Comitato R, Saba A, Turrini A, Arganini C, Virgili F. Sex Hormones and Macronutrient Metabolism. *Critical Reviews in Food Science and Nutrition*. 2015.

8. Morgan AP, Fu CP, Kao CY, Welsh CE, Didion JP, Yadgary L, et al. The mouse universal genotyping array: From substrains to subspecies. *G3 Genes, Genomes, Genet.* 2016;
9. Brockmann GA, Tsaih SW, Neuschl C, Churchill GA, Li R. Genetic factors contributing to obesity and body weight can act through mechanisms affecting muscle weight, fat weight, or both. *Physiol Genomics*. 2009;
10. Li R, Tsaih S-W, Shockley K, Stylianou IM, Wergedal J, Paigen B, et al. Structural Model Analysis of Multiple Quantitative Traits. *PLoS Genet.* 2006;
11. Li Y, Tesson BM, Churchill GA, Jansen RC. Critical reasoning on causal inference in genome-wide linkage and association studies. *Trends in Genetics.* 2010.
12. Wang X, Le Roy I, Nicodeme E, Li R, Wagner R, Petros C, et al. Using advanced intercross lines for high-resolution mapping HDL cholesterol quantitative trait loci. *Genome Res*. 2003;
13. Wang X, Korstanje R, Higgins D, Paigen B. Haplotype analysis in multiple crosses to identify a QTL gene. *Genome Res.* 2004;
14. Wang X, Paigen B. Genetics of variation in HDL cholesterol in humans and mice. *Circulation Research.* 2005.
15. Shim U, Kim HN, Lee H, Oh JY, Sung YA, Kim HL. Pathway analysis based on a genome-wide association study of polycystic ovary syndrome. PLoS One. 2015;
16. Alemany M. Do the interactions between glucocorticoids and sex hormones regulate the development of the metabolic syndrome? *Front Endocrinol (Lausanne).* 2012;
17. Yanes LL, Romero DG. Dihydrotestosterone stimulates aldosterone secretion by

H295R human adrenocortical cells. Mol Cell Endocrinol. 2009;

18. Kawarazaki W, Fujita T. The role of aldosterone in obesity-related hypertension. *American Journal of Hypertension.* 2016.

19. Sang Q, Li X, Wang H, Wang H, Zhang S, Feng R, et al. Quantitative methylation level of the EPHX1 promoter in peripheral blood DNA is associated with polycystic ovary syndrome. *PLoS One.* 2014;

20. Lee H, Oh JY, Sung YA, Chung H, Kim HL, Kim GS, et al. Genome-wide association study identified new susceptibility loci for polycystic ovary syndrome. *Hum Reprod*. 2015;

21. Nymo S, Coutinho SR, Jørgensen J, Rehfeld JF, Truby H, Kulseng B, et al. Timeline of changes in appetite during weight loss with a ketogenic diet. *Int J Obes.* 2017;

22. Nymo S, Coutinho SR, Rehfeld JF, Truby H, Kulseng B, Martins C. Physiological Predictors of Weight Regain at 1-Year Follow-Up in Weight-Reduced Adults with Obesity. *Obesity*. 2019;

23. Lyngstad A, Nymo S, Coutinho SR, Rehfeld JF, Truby H, Kulseng B, et al. Investigating the effect of sex and ketosis on weight-loss-induced changes in appetite. *Am J Clin Nutr.* 2019;

24. Foster GD, Wyatt HR, Hill JO, McGuckin BG, Brill C, Mohammed BS, et al. A randomized trial of a low-carbohydrate diet for obesity. *N Engl J Med.* 2003;

25. Samaha FF, Iqbal N, Seshadri P, Chicano KL, Daily DA, McGrory J, et al. A low-carbohydrate as compared with a low-fat diet in severe obesity. *N Engl J Med.* 2003;

26. Gardner CD, Kiazand A, Alhassan S, Kim S, Stafford RS, Balise RR, et al. Comparison of the Atkins, Zone, Ornish, and LEARN diets for change in weight and

related risk factors among overweight premenopausal women: The A to Z weight loss study: A randomized trial. *J Am Med Assoc.* 2007;

27. Hall KD, Chen KY, Guo J, Lam YY, Leibel RL, Mayer LES, et al. Energy expenditure and body composition changes after an isocaloric ketogenic diet in overweight and obese men. *Am J Clin Nutr.* 2016;
28. Rosenbaum M, Hall KD, Guo J, Ravussin E, Mayer LS, Reitman ML, et al. Glucose and Lipid Homeostasis and Inflammation in Humans Following an Isocaloric Ketogenic Diet. *Obesity*. 2019;
29. Winkler TW, Justice AE, Graff M, Barata L, Feitosa MF, Chu S, et al. The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PLoS Genet.* 2015;
30. Randall JC, Winkler TW, Kutalik Z, Berndt SI, Jackson AU, Monda KL, et al. Sex-stratified Genome-wide Association Studies Including 270,000 Individuals Show Sexual Dimorphism in Genetic Loci for Anthropometric Traits. *PLoS Genet*. 2013;
31. McGrice M, Porter J. The effect of low carbohydrate diets on fertility hormones and outcomes in overweight and obese women: A systematic review. *Nutrients*. 2017.
32. Gupta L, Khandelwal D, Kalra S, Gupta P, Dutta D, Aggarwal S. Ketogenic diet in endocrine disorders: Current perspectives. *Journal of Postgraduate Medicine.* 2017.
33. Mavropoulos JC, Yancy WS, Hepburn J, Westman EC. The effects of a low-carbohydrate, ketogenic diet on the polycystic ovary syndrome: A pilot study. *Nutr Metab.* 2005;

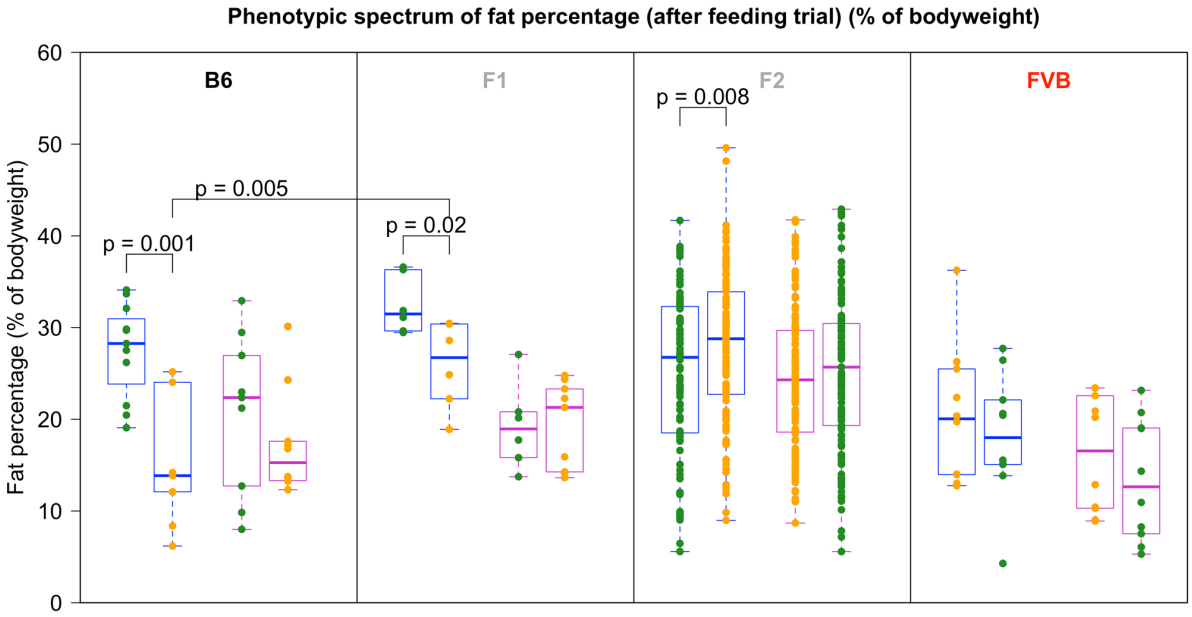

Figure 1. Fat percentage (after feeding trial) (% of body weight). Orange dots indicate ketogenic diet. Green dots indicate American diet. Blue (male) and pink (female) boxes denote sex. B6 males on the ketogenic diet have a lower percentage of body fat than B6 males on the American diet (Welch's two sample t-test when variances are unequal). This trend persists in the F1 population where F1 males on the ketogenic diet have a lower percentage of body fat than F1 males on the American diet. The F1 males on the ketogenic diet also have a significantly higher percentage of body fat than B6 males on the ketogenic diet.

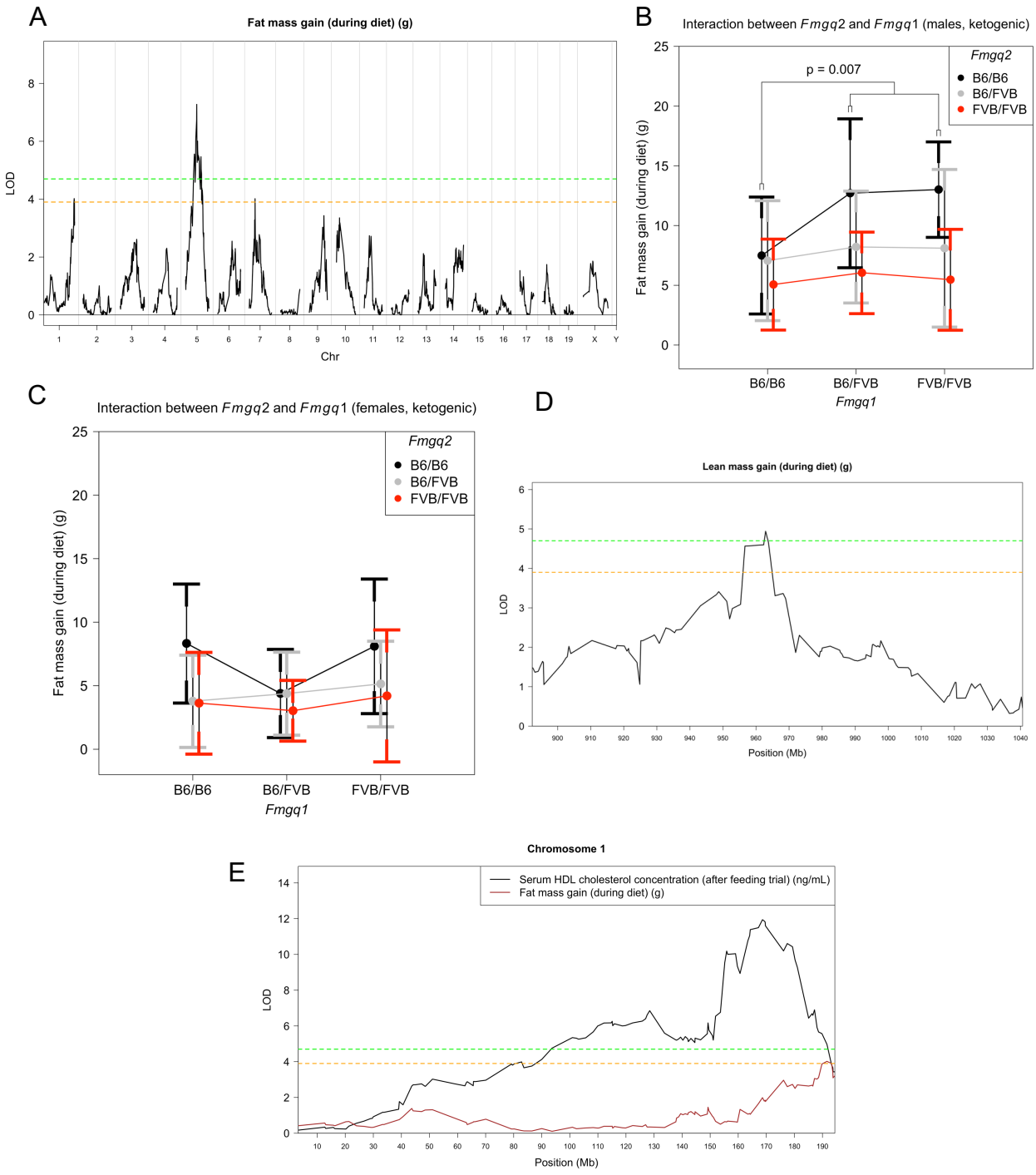

Figure 2. QTL profiles and interaction plots. The QTL profiles show the logarithmic p-values across the whole genome. Positive values indicate that the FVB allele increases the trait while negative values indicate that the FVB allele decreases the trait. The horizontal lines present the genome-wide thresholds of high significance ($p < 0.01$, green) and significance ($p<0.05$, orange) based on 10,000 permutations of the data. **A)** QTL profile for fat mass gain during the feeding trial. (**B)** Interaction plots for *Fmgq2* and *Fmgq1* in ketogenic males and **(C)** ketogenic females represented as mean +/- standard deviation. **D)** QTL profile on Chr5 for lean mass gain during the 3-month feeding trial. **E)** QTL profile on Chr1 for serum HDL cholesterol concentration after the 3-month feeding trial

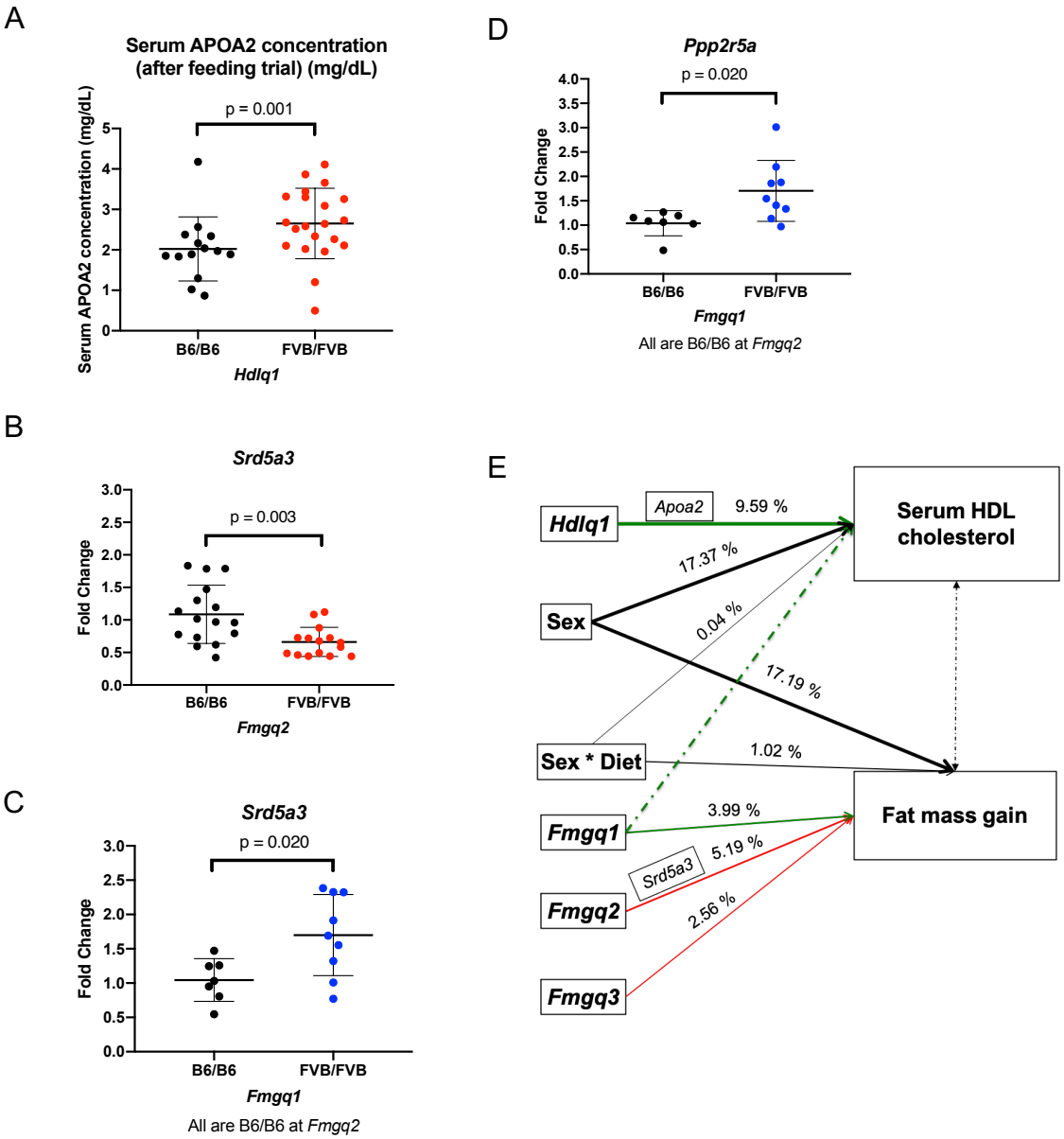

Figure 3. Candidate genes and models. A) Serum *APOA2* concentration for F2s which are homozygous for the alternative alleles at *Hdlq1*. **B)** Fold change in *Srd5a3* expression among F2s of genotype FVB/FVB relative to B6/B6 at *Fmgq2*. **C)** Fold change in *Srd5a3* expression among F2s of that are homozygous for the FVB allele at *Fmgq1* and B6/B6 at relative to those that are homozygous for the B6 allele at both loci. **D)** Fold change in *Ppp2r5a* expression among F2s of that are homozygous for the FVB allele at *Fmgq1* and B6/B6 at relative to those that are homozygous for the B6 allele at both loci. **E)** Graphical representation of SEM for fat mass gain during the feeding trial and serum HDL cholesterol concentrations after the feeding trial in the combined model. Solid arrows indicate the direction of paths and the weight of each arrow is proportional to the path coefficient from the predictor to the variable and the percentage of variation in the variable that is explained by each predictor. Positive effects (green arrows) indicate that the FVB allele increases the trait; negative effects (red arrows) indicate that the FVB allele decreases the trait. The single-headed dashed arrow represents the inferred causal pathway. The double-headed dashed arrow represents a covariate pathway detected by the structural model between fat mass gain and serum HDL cholesterol concentration.

Table 1. Significant QTLs for fat mass and lean mass gain during feeding trial and for serum HDL cholesterol concentration after feeding trial.

Phenotype	Model	Chr	Left marker (Mb)	Top marker (Mb)	Right marker (Mb)	LODξ	Effect of B6/FVB Alleles ξ	Effect of FVB/FVB Alleles ξ	Variance explained (%) ξ
Fat mass gain (during feeding trial) (g)	Combined **(y ~ sex * diet +marker)**	1	180.2	191.6	194.4	4.02 [b]	1.44	2.32	3.44
	Combined **(y ~ sex * diet +marker)**	5	66.1	73.7	76.6	7.28 [a]	-1.95	-2.78	6.29
	Combined **(y ~ sex * diet +marker)**	7	36.6	40.5	44.5	4.02 [b]	0.23	-1.70	3.43
Lean mass gain (during feeding trial) (g)	Combined (both sexes, both diets) **(y ~ sex * diet +marker)**	5	66.1	73.7	76.6	4.94 [a]	-0.78	-1.62	3.34
Serum HDL concentration (after feeding trial) (ng/mL)	Combined (both sexes, both diets) **(y ~ sex * diet +marker)**	1	160.6	168.6	176.1	11.94 [a]	26.88	48.01	9.93

ξ LOD scores, allele effects, and variance explained are provided for the top marker in each confidence interval. Effects of the B6/FVB and FVB/FVB alleles are given relative to B6/B6 alleles at this location.

[a] Genome-wide significance threshold of high significance ($p < 0.01$) based on 10,000 permutations.

[b] Genome-wide significance threshold of significance ($p < 0.05$) based on 10,000 permutations.

Table 2. Candidate genes in QTL confidence intervals for *Fmgq1* and *Fmgq2*. Candidate genes in QTL confidence intervals for *Fmg1* and *Fmg2* were annotated with KEGG pathways. Those genes that contribute in metabolically relevant KEGG pathways are shown here.

Phenotype	Model	QTL	KEGG pathway	Gene Symbol (MGI)
Fat mass gain (during feeding trial) (g)	Combined (both sexes, both diets) **(y ~ sex * diet +marker)**	*Fmgq1*	Steroid hormone biosynthesis (mmu00140)	*Hsd11b1*
			Bile secretion (mmu04976)	*Ephx1*
			Aldosterone synthesis and secretion (mmu04925)	*Camk1g*
			AMPK signaling pathway (mmu04152)	*Ppp2r5a*
	Combined (both sexes, both diets) **(y ~ sex * diet +marker)**	*Fmgq2*	Steroid hormone biosynthesis (mmu00140)	*Srd5a3*
			Non-alcoholic fatty liver disease (mmu04932)	*Cox7b2*

Table 3. Structural equation model for fat mass gain during feeding trial and serum HDL cholesterol concentration after the feeding trial.

Conditioned and Unconditioned QTL Scans				
Phenotype	Model	Condition	QTL	LOD[ξ]
Fat mass gain (during feeding trial) (g)	Combined **(y ~ sex * diet +marker)**	Unconditioned	*Fmgq1*	4.02
	Combined **(y ~ sex * diet +m + HDL)**	Serum HDL cholesterol concentration (after feeding trial) (ng/mL)	*Fmgq1*	4.58
Serum HDL concentration (after feeding trial) (ng/mL)	Combined **(y ~ sex * diet +marker)**	Unconditioned	*Hdlq1*	11.94
	Combined **(y ~ sex * diet +marker + fat mass gain)**	Fat mass gain (during feeding trial) (g)	*Hdlq1*	10.86
Causal Network Analysis				
Relationship	Model	p-value	AIC	BIC
Fat mass gain ← serum HDL cholesterol	(Fat mass gain ~ HDL) ~ sex * diet + Top marker at *Fmgq1*	0.0046 [a]	2181	2209
Fat mass gain → serum HDL cholesterol	(HDL ~ Fat mass gain) ~ sex * diet + Top marker at *Fmgq1*	0.0252 [a]	4204	4232
Structural Model				
	Variable	Predictor	Path coefficient (% variation explained)	t-statistic of Path Coefficient
	Fat mass gain (during feeding trial) (g)	Sex	0.43 (17.19%)	9.89
		Sex * Diet	-0.09 (1.02%)	-2.03
		Fmgq1	0.19 (3.99%)	4.40
		Fmgq2	-0.21 (5.19%)	-4.87
		Fmgq3	-0.11 (2.56%)	-2.55
	Serum HDL cholesterol concentration (after feeding trial) (ng/mL)	Sex	0.44 (17.37%)	10.09
		Sex * Diet	-0.01 (0.04%)	-2.03
		Hdlq1	0.31 (9.59%)	7.20

[ξ] LOD (-log10(p-values)) is provided for the top marker in each confidence interval.
[a] Causal inference is significant at Holm-Bonferroni corrected $p < 0.05$

CHAPTER III: SEX, STRAIN, AND DIET DEPENDENT MODULATION OF GUT MICROBIOTA

Abstract

The microbiome is modulated by a combination of host genetics, diet, and sex effects. The magnitude of these effects and interactions among them is important to understanding inter-individual variability in gut microbiota. In a previous study, mouse strain-specific responses to American and ketogenic diets were observed along with several QTL for metabolic traits. In the current study, we searched for genetic variants underlying differences in the gut microbiome in response to American and ketogenic diets between C57BL/6J (B6) and FVB/NJ (FVB) mouse strains. Genetic mapping of microbial traits revealed 11 loci that were genotype specific, 6 loci that were genotype and diet specific, and 3 loci that were genotype and sex specific. For many microbial traits, irrespective to which quantitative trait loci model was used, diet or the interaction between diet and a genotype were the strongest predictors of the abundance of each microbial trait. Causal network inference suggests that serum HDL cholesterol concentration and abundances of microbiota are independently linked to a multiple genotype specific QTLs. Irrespective to genetic background, diet has a profound ability to modulate gut microbiota. Sex, while important to the analyses, was not as strong of a predictor for microbial abundances. These results demonstrate the importance of characterizing the magnitude of the effects that sex, diet, and genetic background have on inter-individual differences in gut microbiota. Precision nutrition will be advanced through integration of genetic

variation, microbiota variation, and sex in response to diets varied in carbohydrate composition.

Introduction

The gut microbiome has emerged as a key component of precision nutrition and individualized dietary response. Gut microbiota utilizes nutrients passing through the gastrointestinal tract to perform biological functions that impact host digestion, absorption, and metabolism of nutrients[1]. There is a consensus that a relationship exists between the microbes and their host, but the impact of inter-individual variability in gut microbiota composition remains unclear[2,3].

The composition of gut microbiota is influenced by both host genetics and environmental factors such as diet composition[4,5], which is considered one of the most potent regulators of gut microbiota composition. However, changes to bacterial abundance do not occur uniformly in response to diets varied in macronutrient composition because of differences in substrate utilization between bacterial taxa[6]. These complexities make it difficult to detangle the effects of host genetics and diet on the composition of the microbiome to determine what the “ideal” microbiome would be[6,7]. Few studies have considered the extent to which the combination of host genetics and diet modulate the abundance of specific bacterial taxa and even fewer have considered how sex might add an additional layer of complexity to describing inter-individual variation in microbiota composition.

A previous study demonstrated strong mouse strain-specific differences in response to American and ketogenic diets[8–10]. These initial studies provided evidence that C57BL/6J (B6) and FVB/NJ (FVB) mice have strikingly different responses to American and ketogenic diets. To investigate the strain, sex, and diet dependent modulation of the gut microbiota, an intercross population (F2) was generated between B6 and FVB. F2s were fed either an American (35% of energy from fat, 50% from carbohydrate) or a ketogenic (84% of energy from fat, 0% from carbohydrate) diet and fecal microbiota was quantified. The results provide evidence for quantitative trait loci (QTL) that affect microbiota composition, but also significant diet and sex differences in the effects of QTL. In many cases these were genotype-specific QTL and in other cases these were genotype and diet dependent or genotype and sex dependent QTL, which allowed for characterization of the extent to which the hosts genetics, sex, and diet impact specific gut microbiota operational taxon units (OTU).

Methods

Animals and diets

B6 females were crossed with FVB males to generate F1 mice and subsequently intercrossed to generate an F2 population. F2s, 3-5 weeks-old, were screened for their response to American (35% of energy from fat, 50% from carbohydrates) and ketogenic (84% of energy from fat, 0% from carbohydrates) diets during a 3-month feeding trial. Detailed diet compositions are provided in Appendix Table A1.

For the feeding trials, mice were randomly assigned to one of the two diet groups. Half of the F2 mice were placed on American diet (102 males, 122 females) and half on ketogenic diet (126 males, 119 females). All animals were maintained in accordance with Texas A&M University Institution Animal Care and Use Committee guidelines at 22 °C under a 12-hour light cycle. At the end of the feeding trial, mice were euthanized by carbon dioxide asphyxiation, blood was collected, and tissues were harvested and immediately flash frozen in liquid nitrogen.

Microbiota Phenotypes

Stool microbiota was analyzed by 16S rRNA V4 sequencing methodology as reported previously[11]. In brief, Total stool DNA was extracted using ZymoBIOMICS™ 96 MagBead DNA kit (Zymo Research, Irvine, CA) with an automated epMotion (Eppendorf, Hamburg, Germany) robotic system. About 100 mg of stool samples were placed in the ZR BashingBead™ Lysis Tube and homogenized using FastPerp24 bead beater (Millipore, Hayward, CA) at 6.5 HZ for 2 min. The lysate was centrifuged at ≥10,000xg for 1min and 200µl supernatant from lysis tube was transferred to 96 deep-well plate (Eppendorf, Hamburg, Germany) and loaded in an epMotion 5075t robotic system. Using epMotion, 600µl ZymoBIOMICS™ MagBinding Buffer and 25µl of ZymoBIOMICS™ MagBinding Beads were added to each well and was mixed well. After mixing, the plate was placed on a magnetic stand and the supernatant was discarded. MagBinding Beads were washed with MagWash 1 and MagWash 2 and the DNA was eluted using 50 µl ZymoBIOMICS™ DNase/RNase free water. The DNA concentration was measured using NanoDrop One (Thermo Scientific, Petaluma, CA).

Mixed template amplicon library for the 16S variable region 4 (V4) was prepared according to the protocol from Earth Microbiome Project (http://www.earthmicro biome.org/emp-standard-

protocols/) using the extracted stool total DNA and the primer sets (515F and barcoded 806R)[12]. The PCR master mix, primer, and samples were plated using an automated epMotion robotic system (Eppendorf, Hamburg, Germany). Appropriate NTC, extraction control and pooled fecal sample were added to each plate. The PCR master mix was prepared that consisting of 37.5µl of GoTaq Green Master Mix (Promega, Madison, WI), 3µl of 25mM $MgCl_2$, 1.5µl of 10µM forward primer 515F, and 25.5 µL of nuclease-free water. Then, 1.5µl of 10µM barcode specific reverse primer 806R and 6µl of extracted stool DNA were added. PCR was performed in duplicate of 25µL under the following conditions: denaturation (1 cycle) at 94°C for 3 min; amplification of 25 cycles at 94°C for 45 s, 50°C for 60s, and 72°C for 90s; and a final extension step cycle at 72°C for 10min. Amplicon DNA was multiplexed and purified using Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI). The amplicon library was sequenced using the Illumina MiSEQ platform with 2x250bp paired-end sequencing. Sequences were de-multiplexed and amplicon sequence variance (ASV) was determined using the open-source software QIIME2-DADA2 pipeline[13]. A total of 11,316,115 sequences with an average of 26,074 ± 13,697 (mean ± SD) sequences per sample were recovered after demultiplexing. Taxonomy was assigned using the SILVA 132 reference database [14] customized for 16s V4 (515F/806R) region of sequences at the threshold of 99% pairwise identity. ASV belonging to mitochondria and chloroplast were filtered out from the ASV table. We performed a single rarefaction at a sequence depth of 4,500 sequences per sample. α-diversity (Shannon diversity index, observed species, and Faith's PD) and β-diversity (unweighted UniFrac, weighted UniFrac, and Bray Curtis) were calculated from the unfiltered ASV table. Any ASV not seen more than 5 times in at least 5% of the samples were removed for calculating differential bacteria abundance.

Microbiota Statistical Analysis

Statistical analyses were performed using R version v3.6.1 for Windows[15]. Differences in the homogeneity of microbial composition dispersions between FMT groups were determined by using PERMDISP2 function of R Package Vegan[16] with 999 permutations. Differences in microbial community β-diversity were tested by ADONIS (perMANOVA) in the R Package Vegan. Principal coordinate (PCoA) analysis was carried out by PhyloSeq[17]. Differential microbiota abundance was analyzed by ANCOM (Mandal et al., 2015) using R package ancom.R with default settings and FDR correction. Graphs were prepared by GGplot2[18] and GraphPad Prism (GraphPad Software, Inc., CA, U.S.A.).

Metabolic Phenotypes

The data analysis and collection methods for fat mass gain and serum HDL cholesterol concentration have been described previously[19].

Genotyping

The genotyping analysis and collection methods have been described previously[19].

Quantitative Trait Loci Analysis

For microbiota phenotypes, a core measurable microbiota (CMM) was defined as those traits present in at least 20% of the individuals at the genus and species level of taxonomy. The CMM consists of 137 microbial taxonomies referred to as operational taxonomic units (OTU). The lowest available classification in the taxonomic hierarchy was defined as the OTU for each organism used for linkage analysis. Pearson's correlation between phyla for each diet group was determined after correcting for sex effects.

The combined model (y~ marker + sex * diet), the diet specific model (y~ sex + diet * marker), and the sex specific model (y~ diet + sex * marker) include all F2s, from both sexes and both diets. QTL peaks with a logarithm of the odds (LOD) greater than thresholds determined by 10,000 permutations were considered genome-wide significant ($p<0.05$, LOD > 3.95) or highly significant ($p<0.01$, LOD > 4.79). A LOD drop of 1.5 LOD from the top marker was used to determine the 95% confidence intervals, or support intervals, for each QTL. Linear models using ANOVA was used to check for any interactions between sex and/or diet with the top markers of each QTL. The variance explained by the top markers at each QTL in the combined model was calculated by dividing the sum of squares of the model including the top marker by the total sum of squares of the model without QTL. The variance explained by the top markers at each QTL in the interactive models was calculated by dividing the sum of squares of the model including the interaction between diet and the top marker or sex and the top marker by the total sum of squares of the model without QTL.

Candidate gene analysis using KEGG

All genes within each significant QTL confidence interval were annotated with KEGG pathway identifiers. We further characterized candidate genes by KEGG pathways related to glucose, insulin, fatty acids, adipocytes, cholesterol, obesity, diabetes mellitus, metabolic syndrome, and digestion and absorption of carbohydrates, fats, and proteins. A comprehensive list of KEGG pathway queries is provided in Appendix Table A2.

Structural Equation Modeling

A set of conditional QTL scans were performed using traits as covariates in the analysis of other traits based on the individual QTL that were identified for each trait and suspected biological relationships. If conditioning the genome scan with a covariate resulted in a significant increase or decrease in the absolute value of the LOD score, it was interpreted that the traits are causally related to one another. When comparing conditioned and unconditioned QTL scans, an increase or decrease in the LOD of at least 2.0 corresponds to a 5% type I error rate[20,21]. Only conditioned genome scans that resulted in a change of 2.0 LOD were considered pleiotropic, or shared, QTL between the two traits.

Causal Network Inferences

For all overlapping QTL, causal inferences were made in networks based on methods described elsewhere[22]. Briefly, the first trait (T1) was regressed on the second trait (T2) and T2 was regressed on T1 in order to obtain the residual of each trait after adjusting for the other (R1 and R2). A bivariate t-test between R1, R2, and the shared locus was used to infer the causal network among them. P-values were Holm-Bonferroni corrected for the number of tests (*i.e.* number of residuals tested = 2) where p= $\alpha_{0.05}$ / (number of tests – rank of i^{th} hypothesis + 1). A p-value < 0.05 was considered significant. An initial pathway describing the relationship between T1 and T2 was defined based on the inferred causal networks, and the predicted causal models were compared with the Akaike's information criterion (AIC) and Bayesian information criterion (BIC). A lower AIC and BIC indicates a better model.

Results

Diet influences Phyla in F2

Diet explains a large proportion of variation in the abundance of microbiota at the Phyla level irrespective to genetic background. Diet explains 64.79% of variation in abundance of Actinobacteria, 25.49% of variation in the abundance of Bacteroidetes, and 61.22% of variation in the abundance of Firmicutes (Appendix Table A5). The relative abundance of Firmicutes in F2s on the ketogenic diet is nearly twice as high as in F2s on the American diet (Figure 4A). This increase in Firmicutes in F2s on the ketogenic diet appears to occur at the expense of the relative abundance of Actinobacteria and Bacteroidetes (Figure 4A).

Diet differentially modulates correlations among Phyla in F2s on American and ketogenic diets (Figure 4B). The abundance of Firmicutes has a slight positive correlation with Actinobacteria in F2s on the ketogenic diet while the abundance of Firmicutes is negatively correlated to Actinobacteria in F2s on the American diet (ketogenic: r^2= 0.12, p= 0.063; American: r^2= -0.66, p< 0.001). Unique to the ketogenic diet, the abundance of Firmicutes is also negatively associated with the abundance of Verrucomicrobia (r^2= -0.71, p < 0.001) and Proteobacteria (r^2= -0.27, p< 0.001). While Firmicutes are negatively associated with Bacteroidetes in F2s on both diets, the correlation is much stronger in F2s on the ketogenic diet (American: r^2= -0.44, p< 0.001, ketogenic: r^2= -0.79, p< 0.001).

Sex and an interaction between sex and diet influence Phyla in F2

At the Phyla level, sex describes a very low proportion of variation in microbiota (Appendix Table A5). Although the effect of sex is significant for the abundance of Verrucomicrobia,

sex only explains 0.99% of the variation in this microbial trait (p = 0.031, 95% CI = -106.79 – (-5.15) abundance). Furthermore, interaction between sex and diet has no significant influence of F2 microbiota at the phyla level.

Linkage analysis reveals QTLs for Microbial Traits

In the combined analysis that incorporates all F2s on both diets (y~ marker + sex * diet), 14 distinct QTL were detected for microbial traits at the OTU level (Figure 5, Table 4). The first of these QTL being for the genus and an unidentified species of *Coriobacteriaceae* (*Coriobacteriaceae.UCG.002*) on Chr2 at 65.8 Mb (*Mtq1*; $p < 0.05$, CI= 60.3-77.4 Mb) and Chr8 at 34.7Mb (*Mtq2*; $p < 0.05$, CI = 22.0-55.1), followed by an unidentified *Bacteroidales* bacterium in the family *Muribaculaceae* on Chr8 at 51.0Mb (*Mtq3;* $p < 0.05$, CI= 43.4-62.8 Mb), the genus *Lactobacillus* on Chr1 at 65.6 Mb (*Mtq4*; $p < 0.05$, CI= 51.1-79.2), the genus *Streptococcus* on Chr8 at 68.3 Mb (*Mtq5*; $p < 0.01$, CI= 58.6-84.1 Mb), the genus *Lachnospiraceae* FCS020 Group on Chr2 at 40.1 Mb (Mtq6; $p < 0.05$, CI= 24.1-60.3 Mb), the genus *Roseburia* on Chr9 at 65.6 Mb (Mtq7; $p < 0.05$, CI = 60.1-85.4 Mb), the unidentified genus of the family *Lachnospiraceae* on Chr1 at 181.1 Mb (Mtq8; $p < 0.01$, CI = 151.9-186.6 Mb), the genus and an unidentified species of *Romboutsia* on Chr13 at 18.4 Mb (*Mtq9*; $p < 0.05$, CI = 5.4-37.5 Mb), the genus and an unidentified species of *Butyricicoccus* on Chr1 at 137.8 Mb (*Mtq10;* $p < 0.05$, CI= 133.7-144.5 Mb), the genus and an unidentified species of *Dubosiella* on Chr3 at 102.3 Mb (*Mtq11;* $p < 0.05$, CI=89.1-130.4 Mb), the unidentified genus of the family *Erysipelotrichaceae* on Chr19 at 14.8 Mb (*Mtq12*; $p < 0.05$, CI= 10.2-24.1 Mb), and the genus and unidentified species of *Bilophila* on Chr1 at 188.6 Mb (*Mtq13*; $p < 0.05$, CI=

144.5-193.3) and Chr9 at 76.4 Mb (*Mtq14*; $p < 0.05$, CI= 42.3-80.8). *Mtq8* for the unidentified genus of the family *Lachnospiraceae* overlaps with *Mtq13* for the genus and unidentified species *Bilophila*.

With the exception of *Coriobacteriaceae* (*Mtq1* and *Mtq2*) and *Lachnospiraceae* FCS020 (*Mtq6*), diet appears to explain a significant proportion of the variation in the abundance of these OTUs despite these QTL not being diet specific. The unidentified *Bacteroidales* bacterium from the *Muribaculaceae* family (*Mtq3*) is the only OTU for which the top marker explains a greater proportion of the variation in the abundance of the organism despite there being a significant effect of diet as well (Table 4). For *Lactobacillus* (*Mtq4*), *Streptococcus* (*Mtq5*), *Roseburia* (*Mtq7*), *Lachnospiraceae* (*Mtq8*) *Romboutsia* (*Mtq9*), *Butyricoccus* (Mtq10), *Dubosiella* (*Mtq11*), *Erysipelotricaceae* (*Mtq12*), and *Bilophila* (*Mtq13* and *Mtq14*), diet explains a greater proportion of the variation than the top marker does at each QTL. This is particularly clear for *Streptococcus* where diet explains a striking 24.67% of the variation in abundance *of Streptococcus* while the top marker at *Mtq5* explains 4.42% of the variation. For *Lachnospiraceae* diet explains 17.56% of the variation while the top marker at *Mtq8* explains 3.97% of the variation, for *Dubosiella* diet explains 10.64% of the variation in abundance while the top marker at *Mtq11* explains 3.86% of the variation, for *Erysipelotrichaceae* diet explains 9.55% of variation in abundance while the top marker at *Mtq12* explains 3.84% of the variation, and for *Bilophila* diet explains over 15% of variation in abundance while the top marker at *Mtq13* explains 3.74% of the variation and the top marker at *Mtq14* explains 3.66% of the variation (Table 4).

Linkage analysis reveals diet specific QTLs for Microbial Traits

In the diet specific model that includes male and female F2s on both diets (y~ sex + diet * marker), six QTL were detected for microbial traits at the OTU level (Figure 5, Table 4). Five of these QTL were distinct from the ones identified in the combined model. The first of these distinct QTL being for the genus and an unidentified species of *Alistipes* on Chr16 at 79.4 Mb (*Mtq15*; $p < 0.05$, CI=72.8-95.8 Mb), followed by the genus and unidentified species of the *Rikenellaceae* RC9.gutgroup on Chr16 at the same location (*Mtq16;* $p < 0.05$, CI= 72.8-96.5 Mb), the genus *Streptococcus* on Chr8 at 117.1 Mb (*Mtq15;* $p < 0.05$, CI=102.4-122.5 Mb), the genus *Erysipelatoclostridium* on Chr6 at 77.1 Mb (*Mtq18;* $p < 0.05$, CI=54.6-81.5 Mb), and the unidentified genus of the family *Erysipelotrichaceae* on Chr9 at 110.5 Mb (*Mtq19*; $p < 0.05$, CI= 98.5-115.6 Mb).

The remaining diet specific QTL is another locus for the unidentified genus of the family *Erysipelotrichaceae* and nearly identical to *Mtq12* that was identified in the combined model for the same family. The only difference being that the top marker at *Mtq12* in the combined model was JAX00471367 at 14.8Mb while the top marker at *Mtq9* in the diet specific model was backupJAX00471466 at 15.1Mb. The 95% confidence interval is unchanged between the combined and diet specific models so this QTL continues to be referred to as *Mtq12*.

Interestingly, diet alone explains a greater proportion of the variation than the interaction between diet and the top marker at each QTL for *Alistipes* (*Mtq12*)*, Rikenellaceae*

(*Mtq13*), *Streptococcus* (*Mtq15*), and *Erysipelotrichaceae* (*Mtq16* and *Mtq9*) (Table 4). At *Mtq14,* the interaction between diet and the top marker explains 4.61% of the variation in the abundance of *Eryispelatoclostridium* while diet explains 3.44% of the variation (Table 5).

Linkage analysis reveals sex specific QTLs for microbial traits

In the sex specific model that includes male and female F2 animals on both diets (y~ diet + sex * marker), three QTL were detected for a microbial traits at the OTU level (Figure 5, Table 4). The first sex specific QTL is for the genus and unidentified species of *Alistipes* on Chr13 at 18.4 Mb (*Mtq20;* $p < 0.05$, CI = 12.5-43.7 Mb), followed by two loci for the genus *Lactobacillus* on Chr7 at 34.6 Mb (*Mtq21*; $p < 0.05$, CI = 18.9-41.9 Mb) and 54.0 Mb (*Mtq22* ; $p < 0.05$, CI = 44.5-72.3 Mb).

Sex explains over 2% of variation in the abundance of *Alistipes* while the interaction between sex and the top marker at *Mtq20* explains over 4% of variation in the abundance of this OTU (Table 4). *Alistipes* is the only organism that was associated with both a diet specific QTL (*Mtq15*) and a sex specific QTL (*Mtq20*). Even in the sex specific model of *Mtq20*, diet explains a greater proportion of the variation than either sex or the interaction between sex and the genotype. *Lactobacillus* is the only organism that was associated with both a genotype specific QTL (*Mtq4*) and multiple sex specific QTL (*Mtq21* and *Mtq22*). Sex describes a significant proportion of the variation at the genotype and sex specific QTLs (Table 4).

Conditioned Linkage Analysis of Lachnospiraceae and Bilophila

Mtq8 and *Mtq13* overlap in the combined model of F2s that includes both sexes and diets. This overlapping QTL prompted modeling of the relationship between *Lachnospiraceae* (*Mtq8*) and *Bilophila* (*Mtq13*). However, there is no significant change to the LOD score at *Mtq8* or *Mtq13* when the abundance of *Bilophila* is conditioned on the abundance of *Lachnospiraceae*, nor when the abundance of *Lachnospiraceae* is conditioned on the abundance of *Bilophila*. Thus, the relationship between *Bilophila* and *Lachnospiraceae* at these loci could not be elucidated in this model. Interestingly, *Mtq8* and *Mtq13* not only overlap with each other but also two QTL previously identified for fat mass gain during the feeding trial (*Fmgq1;* Chr1 180.2 Mb- 194.4 Mb) and serum HDL cholesterol concentration after the feeding trial *(Hdlq1;* Chr1 160.6-176.1 Mb) (Figure 5).

However, there is no significant change to the LOD score after conditioning on serum HDL cholesterol concentration, nor when conditioning serum HDL cholesterol concentration on either OTU. The same is true for these OTU and fat mass gain. Therefore, the relationship between these OTU and metabolic traits could not be elucidated. This would suggest that the overlapping QTL regions harbor one or two tightly linked genes that affect the OTU and metabolic traits independently.

Conditioned Linkage Analysis of Roseburia and Bilophila

The proximity of *Mtq7* and *Mtq14* in the combined models prompted modeling of the relationship between abundance of *Roseburia* (*Mtq7*) and *Bilophila* (*Mtq14*). However, there is no significant change to the LOD score after conditioning the abundance of these

OTU on one another. This again would suggest that the overlapping QTL regions harbor one or two tightly linked genes that affect both OTU.

Causal network analysis of overlapping QTL

Causal networks were explored for all overlapping QTL. After adjusting for multiple testing, the inferred networks revealed that *Bilophila* and *Roseburia* are independently related to both *Mtq7* and *Mtq14*, as well as that *Lachnospiraceae* and serum HDL cholesterol concentration are independently related to *Mtq8* and *Hdlq1*, and that *Bilophila* and serum HDL cholesterol concentration are independently linked to *Mtq13* (Table 5). The AIC and BIC scores suggest that the model in which abundance of *Lachnospiraceae* occurs upstream of the concentration of serum HDL cholesterol concentration at *Hdlq1* is most consistent with the data (Table 5). Otherwise, the AIC and BIC model selection scores for the remaining models support the independent relationship between each of the traits and the QTL.

Finally, a structural model was built to illustrate the magnitude of the effects of each predictor in the models of *Roseburia*, *Lachnospiraceae*, *Bilophila, and* serum HDL cholesterol concentration (Figure 6). The path coefficients are all significantly different from zero (Table 5). The proportion of variation explained with the predictors in the structural model for *Roseburia*, *Lachnospiraceae*, *Bilophila*, and serum HDL cholesterol concentration is 10.06%, 26.91% 22.85%, and 29.98% respectively.

KEGG pathway annotation of candidate genes at Mtq7 and Mtq14

Candidate genes were investigated that might elucidate the relationship between *Bilophila*, *Roseburia*, and *Mtq14*. Positional candidates at *Mtq7* and *Mtq14 that* overlap with one or more metabolic KEGG pathways are summarized in Table 6. Out of 180 positional candidates at *Mtq7*, seven genes overlap with one or more metabolic KEGG pathways. Out of 398 positional candidates at *Mtq14*, 25 genes overlap with one or more metabolic KEGG pathways. All seven genes that were annotated with KEGG pathways at *Mtq7* are present in the confidence interval for *Mtq14*. These genes include *Aqp9*, *Col12a1*, *Cox7a2*, *Hmgcll1*, *Map2k1*, *Onecut1*, and *Slc51b*. The remaining 18 genes that are unique to the *Mtq14* confidence interval can be found in Table 6. Seventeen out of these 25 positional candidates at *Mtq14* harbor a non-synonymous transcript variant. The presence of these non-synonymous transcript variants makes *Col12a1*, *Map2k1*, *Slc51b*, *Adpgk*, *Apoa1*, *Apoa4*, *Apoa5*, *Apoc3*, *Cyp11a1*, *Cyp19a1*, *Cyp1a1*, *Cyp1a2*, *Dlat*, *Fxyd2*, *Ppp2r1b*, *Sc5d*, and *Slc37a4* primary candidate genes of interest in this region.

KEGG pathway annotation of candidate genes at Mtq8, Mtq13, and Hdlq1

Candidate genes that might elucidate the relationship between serum HDL cholesterol concentration, *Bilophila*, *Lachnospiraceae*, *Hdlq1*, *Mtq13*, and *Mtq8* were also investigated. Positional candidates at Hdlq1, *Mtq8*, and *Mtq13* that overlap with one or more metabolic KEGG pathways are summarized in Table 6. Out of 205 positional candidates at Hdlq1, 11 genes overlap with one or more KEGG pathways. Out of 667 positional candidates at *Mtq8*, 15 genes overlap with one or more KEGG pathways. Out of 405 positional candidates at *Mtq13*, 18 genes overlap with one or more KEGG

pathways. All ten genes that were annotated with KEGG pathways at *Hdlq1* and all 15 genes that were annotated with KEGG pathways at *Mtq8* are present in the confidence interval for *Mtq13*. The remaining three genes that are unique to *Mtq13* are *Pla2g4a*, *Ppp2r5a*, and *Ptgs2*. Fifteen out of the 18 positional candidates harbor a non-synonymous transcript variant. The presence of these non-synonymous transcript variants makes *Akt3*, *Apoa2*, *Atp1a2*, *Atp1a4*, *Atp1b1*, *Ephx1*, *Hsd17b7*, *Ndufs2*, *Pex19*, *Pla2g4a*, *Ppp2r5a*, *Ptgs2*, *Rxrg*, *Sdhc*, and *Soat1* the primary candidate genes of interest in this region.

Discussion

This study provides evidence that abundances of gut microbiota are driven by unique combinations of effects from the host's genetics, response to high fat diets varied in carbohydrate content, and sex. In this study, 14 genotype specific QTL were identified at *Mtq1* and *Mtq2* for the genus and unidentified species of *Coriobacteriaceae, Mtq3* for the unidentified *Bacteroidales* bacterium of the family *Muribaculaceae, Mtq4* for the family *Lactobacillus, Mtq5* for the family *Streptococcus, Mtq6* for the genus *Lachnospiraceae* FCS020 Group, *Mtq7* for the genus *Roseburia*, *Mtq8* for the unidentified genus of the family of *Lachnospiraceae*, Mtq9 for the genus *Romboutsia*, *Mtq10* for the genus *Butyricoccus, Mtq11* for the genus *Dubosiella, Mtq12* for the unidentified genus of the family *Erysipelotrichaceae*, and both *Mtq13* and *Mtq14* for the genus and unidentified species of *Bilophila*.

Out of these genotype specific QTL, *Coriobacteriaceae* and *Lachnospiraceae* FCS020 Group seem to be the least influenced by diet. *Coriobacteriaceae* has previously been associated with host genetics and QTL regulating immune function and susceptibility to carcinoma and tumor development in mice[23,24]. Coriobacteriaceae has been described as a dominant species in the mammalian gut and is positively correlated with hepatic triglyceride concentration and non-HDL cholesterol concentration in mice[25]. In humans, the family *Lachnospiraceae* is more similar among monozygotic twins than dizygotic twins and less influenced by environmental factors[2,24]. *Lachnospiraceae* FCS020 Group has been associated with VLDL particles, small HDL particles, and triglyceride concentration within HDL [26]. Larger HDL particles are associated with lower cardiovascular disease risk.

A significant proportion of the variation in all other OTU with genotype specific QTLs is driven by diet, especially for *Streptococcus* at *Mtq5*, the unidentified genus of the family *Lachnospiraceae* at *Mtq8*, *Dubosiella* at *Mtq11*, *Erysipelotrichaceae* at *Mtq12*, and *Bilophila* at both *Mtq13* and *Mtq14*. Each of these OTU belong to the Firmicutes phyla. Fiber is a particularly important dietary component for modulating abundance of Firmicutes. When animals switch from a low fat/fiber rich plant diet to a high fat/high sugar diet, they experience a significant increase in the Firmicutes phylum along with a decrease in *Bacteroidetes*[27]. Dramatic shifts were observed in these phyla between American and ketogenic diet F2s irrespective to their genetic backgrounds. The relative abundance of Firmicutes in F2s exposed to the ketogenic diet is nearly twice as high as F2s exposed to the ketogenic diet. It appears that this increase in Firmicutes in F2s exposed to the ketogenic diet coincides with a decrease in the relative abundance of

Bacteroidetes. The ketogenic diet is composed of twice as much soluble and insoluble fiber as the American diet, and this likely drives many of the differences in the abundance of OTUs from these phyla. Limited evidence suggests that a higher Firmicutes to Bacteroidetes ratio is positively correlated with obesity while a decrease in this ratio has been associated with inflammatory bowel disease however, much controversy surrounds the association of the Firmicutes to Bacteroidetes ratio and health status[28].

Aside from the influence dietary fiber has on the gut microbiota, there is an abundance of literature supporting the potent effects of diet on the abundance of gut microbiota [5,27,29–33]. Other studies have demonstrated that the effect of abnormal diets on gut microbiota might stifle the underlying effect of single gene mutations because diets are such a potent regulator of microbial abundances[33,34]. These authors have called for further study of diets varied in macronutrient content and study of more complex genetic models. The current study has demonstrated that high fat diets varied in carbohydrate content continue to be commanding predictors of abundances of gut microbial abundances even in a more complex genetic model.

The only OTU for which the top marker explains a greater proportion of the variation in the abundance of the organism than diet when diet is a significant predictor in the model is *Muribaculaceae*. There is evidence that abundance of *Muribaculaceae* is particularly sensitive to genetic abnormalities like the single gene mutation in the leptin receptor harbored by Ob/Ob mice. In this model, abundance of *Muribaculaceae* was reduced in homozygous Ob/Ob mice relative to B6 when both strains were exposed to HFD[33]. This

suggests that *Muribaculaecae* is one organism for which a genotype would be a commanding predictor of microbial abundance even in the presence of abnormal diets. *Muribaculaceae* are capable of producing enzymes that degrade complex carbohydrates and modulate energy metabolism in mice[35].

Six diet specific QTL were also identified, all of which are for microbial traits from either the Bacteroidetes or Firmicutes phyla. These QTL include, *Mtq15* for the genus and unidentified species of *Alistipes* (Bacteroidetes)*, Mtq16* for the genus and unidentified species of *Rikenellaceae* (Bacteroidetes), *Mtq17* for the family *Streptococcus* (Firmicutes)*, Mtq18* for the genus *Erysipelatoclostridium* (Firmicutes), and *Mtq19* and *Mtq12* for the family *Erysipelotrichaceae* (Firmicutes). Diet is the strongest predictor in these models for these OTU except for the abundance of *Erysipelatoclostridium* (Firmicutes)*.*

Many previous studies have compared the effects of control mouse diets to high fat diets where one or two representative ingredients contribute to the total fat, carbohydrate, and protein content of the diet[8]. Our American and ketogenic diets recapitulate the diversity of ingredients found in human diets[8]. The fat component of the ketogenic diet is composed of equal parts butter and lard with a small portion of corn and menhaden oils, while the fat component of the American diet is a more diverse mixture of primarily butter as well as corn, menhaden, flaxseed, and olive oil. Lard-derived fat has been shown to reduce the abundance of *Streptococcus* [5]. The American diet contains multiple sources of animal proteins, some of which contribute to the total fat content of the diet. In contrast, casein

is the main protein source in the ketogenic diet and the total fat content is primarily derived from butter and lard. *Alistipes* is particularly influenced by animal-based protein diets [5,31,36]. This observation could reflect differences in dietary fat intake as the saturated fat content varies between animal-based proteins. *Rikenellaceae* RC9 gut group is a member of the *Rikenellaceae* family like *Alistipes* and may be similarly impacted by animal proteins and lipid metabolism[37,38]. *Alistipes* and *Rikenellaceae* are both thought to be related to type 2 diabetes mellitus[35]*.* Many of the effects of diet for these diet specific QTL are likely influenced by the differences in soluble and insoluble fiber in the American and ketogenic diets as well.

Alistipes is an excellent example of the context dependent nature of the diet specific QTLs. In the context of non-human-comparable high fat and control mouse chows, others have identified QTL for *Alistipes* that were not diet dependent[29]. With the knowledge that the abundance of *Alistipes* is particularly influenced by animal-based proteins, it becomes more clear why our American and ketogenic diets and their diverse set of ingredients contributing to the protein-content of the diet, lead to the detection of a diet specific QTL regulating the abundance of *Alistipes*. This makes it challenging to apply knowledge from QTL models to different contexts, especially when non-human comparable diets are used.

For one of the loci picked up by the sex specific QTL model, *Mtq20* for *Alistipes,* the strongest predictor in the model was again diet. Sex specificity for abundance of *Alistipes* has been established in studies of pre- and-post-menopausal women and men. Men were more likely than pre- or post-menopausal women to have higher abundances of

Alistipes in their fecal samples[39]. Diet has a significant effect on the other sex specific QTLs for *Lactobacillus* at *Mtq21* and *Mtq22*. However, sex and the interaction between sex and the genotype at *Mtq21* describe a greater proportion of the variation in the abundance of *Lactobacillus* than diet does at this locus. *Lactobacillus* belongs to the family *Lactobacillaceae* which has been observed in higher abundance among post-menopausal women when compared to men and is negatively associated with circulating plasma testosterone levels[39]. Lactobacillus is one of many probiotics associated with positive health outcomes. The realized importance of sex as a biological variable has increased attention paid to the role of steroid hormones in development of obesity and Metabolic Syndrome[19]. Plasma testosterone has also been linked to microbial traits in men, and the post-menopausal microbiome becomes more similar to the male microbiome over time[39].

QTL for an unidentified genera of *Lachnospiraceae* (*Mtq8*), the genera and an unidentified species of *Bilophila* (*Mtq13*), fat mass gained during the feeding trial (Fmgq1), and serum HDL cholesterol concentration (*Hdlq1*) overlap on the distal part of Chr1 and QTL for the genera *Roseburia* (*Mtq7*) and the genera and an unidentified species of *Bilophila* (*Mtq14*) overlap on Chr9. American or westernized diets are associated with increased abundances of *Bilophila wadsworthia* which coincides with increased LDL cholesterol concentration and links this species of *Bilophila* to dyslipidemia and increased inflammation[40]. Lachnospiraceae is increased in many metabolic diseases including diabetes and non-alcoholic fatty liver disease[41]. We expected to find causal relationships between these metabolic and microbial traits and even between multiple microbial traits,

however, conditioned linkage analysis and causal networks suggest that there is an independent relationship between each of these overlapping loci and the traits that associate with them.

Our inability to detect causal relationships between metabolic and microbial traits mapping to similar regions of the genome at *Fmgq1*, *Hdlq1*, *Mtq8*, and *Mtq13* as well as between microbial traits at *Mtq7* and *Mtq14* may be the result of underlying latent variables. Latent variables are those that are not directly observable in a model but can be inferred from other variables. As mentioned previously, gut microbiota utilize nutrients passing through the gastrointestinal tract. Microbial metabolism of these nutrients produces metabolites and microbial-derived metabolites are known to impact metabolic health[1]. These metabolites may represent latent variables linking the genomic region underlying *Fmgq1*, *Hdlq1*, *Mtq8*, and *Mtq13* and each of their associated traits as well as *Mtq7* and *Mtq14* and their microbial traits.

Conclusions

The current experiment identified organisms for which genetic background was the strongest predictor for bacterial abundances and other organisms for which irrespective to genetic background, diet was the strongest modulator of gut microbiota, as well as organisms for which combinations of sex, diet, and genotypes modulated the gut microbiota. These results demonstrate the importance of characterizing the magnitude of effects that sex, diet, and genetic background have on inter-individual differences in gut microbiota. Precision nutrition will be advanced through integration of genetic variation, microbiota variation, and sex in response to diets varied in carbohydrate composition.

Acknowledgements

We thank Dr. Andrew Hillhouse and the Texas A&M Institute for Genome Sciences and Society (TIGSS) for assistance in the TIGSS Molecular Genomics Workspace and TIGSS Rodent Preclinical Phenotyping Core; Dr. Karl Broman for his discussion and guidance during the QTL analysis as well as for organizing and managing annotation files for the MUGA arrays; Daniel Genung and Aaron Van Wettering for assistance with mouse husbandry and phenotyping; and Ryan McGovern for assistance with preparing samples for genotyping; and all members of the laboratory for their helpful insights. This work was supported by National Institutes of Health (NIH) grants RM1HG008529 and P30ES029067.

Competing Interests

The authors have no competing interests to declare.

References

1. Lim MY, You HJ, Yoon HS, et al. The effect of heritability and host genetics on the gut microbiota and metabolic syndrome. *Gut*. 2017;66(6):1031-1038. doi:10.1136/gutjnl-2015-311326
2. Goodrich JK, Waters JL, Poole AC, et al. Human genetics shape the gut microbiome. *Cell*. 2014;159(4):789-799. doi:10.1016/j.cell.2014.09.053
3. Bonder MJ, Kurilshikov A, Tigchelaar EF, et al. The effect of host genetics on the gut microbiome. *Nature Genetics*. 2016;48(11):1407-1412. doi:10.1038/ng.3663
4. Snijders AM, Langley SA, Kim YM, et al. Influence of early life exposure, host genetics and diet on the mouse gut microbiome and metabolome. *Nature Microbiology*. 2016;2. doi:10.1038/nmicrobiol.2016.221
5. Singh RK, Chang HW, Yan D, et al. Influence of diet on the gut microbiome and implications for human health. *Journal of Translational Medicine*. 2017;15(1). doi:10.1186/s12967-017-1175-y
6. Gentile CL, Weir TL. The gut microbiota at the intersection of diet and human health. *Science*. 2018;362(6416). doi:10.1126/science.aau5812
7. Cabrera-Mulero A, Tinahones A, Bandera B, Moreno-Indias I, Macías-González M, Tinahones FJ. Keto microbiota: A powerful contributor to host disease recovery. *Reviews in Endocrine and Metabolic Disorders*. 2019;20(4). doi:10.1007/s11154-019-09518-8
8. Barrington WT, Wulfridge P, Wells AE, et al. Improving metabolic health through precision dietetics in mice. *Genetics*. 2018;208(1):399-417. doi:10.1534/genetics.117.300536

9. Wells A, Barrington WT, Dearth S, et al. Tissue Level Diet and Sex-by-Diet Interactions Reveal Unique Metabolite and Clustering Profiles Using Untargeted Liquid Chromatography-Mass Spectrometry on Adipose, Skeletal Muscle, and Liver Tissue in C57BL6/J Mice. *Journal of Proteome Research*. Published online 2018. doi:10.1021/acs.jproteome.7b00750

10. Cuomo D, Porreca I, Ceccarelli M, et al. Transcriptional landscape of mouse-aged ovaries reveals a unique set of non-coding RNAs associated with physiological and environmental ovarian dysfunctions. *Cell Death Discovery*. Published online 2018. doi:10.1038/s41420-018-0121-y

11. Nazmul Huda M, Winnike JH, Crowell JM, O'Connor A, Bennett BJ. Microbial modulation of host body composition and plasma metabolic profile. *Scientific Reports*. 2020;10(1). doi:10.1038/s41598-020-63214-1

12. Caporaso JG, Lauber CL, Walters WA, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME Journal*. 2012;6(8). doi:10.1038/ismej.2012.8

13. Bolyen E, Rideout JR, Dillon M, et al. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ*. Published online 2018. doi:10.7287/peerj.preprints.27295

14. Quast C, Pruesse E, Yilmaz P, et al. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*. 2013;41(D1). doi:10.1093/nar/gks1219

15. R Core Team (2020). R: A language and environment for statistical computing. *R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria*. Published online 2020.
16. Oksanen J, Kindt R, Legendre P, et al. The vegan package. Community ecology package 10. Published online 2007.
17. McMurdie PJ, Holmes S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*. 2013;8(4). doi:10.1371/journal.pone.0061217
18. Wickham H. ggplot2: elegant graphics for data analysis. Published online 2009.
19. Salvador AC, Arends D, Barrington WT, Elsaadi AM, Brockmann GA, Threadgill DW. Sex-specific genetic architecture in response to American and ketogenic diets. *International Journal of Obesity*. 2021;45(6). doi:10.1038/s41366-021-00785-7
20. Brockmann GA, Tsaih SW, Neuschl C, Churchill GA, Li R. Genetic factors contributing to obesity and body weight can act through mechanisms affecting muscle weight, fat weight, or both. *Physiological Genomics*. Published online 2009. doi:10.1152/physiolgenomics.90277.2008
21. Li R, Tsaih S-W, Shockley K, et al. Structural Model Analysis of Multiple Quantitative Traits. *PLoS Genetics*. Published online 2006. doi:10.1371/journal.pgen.0020114
22. Li Y, Tesson BM, Churchill GA, Jansen RC. Critical reasoning on causal inference in genome-wide linkage and association studies. *Trends in Genetics*. Published online 2010. doi:10.1016/j.tig.2010.09.002

23. Benson AK, Kelly SA, Legge R, et al. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proceedings of the National Academy of Sciences of the United States of America*. 2010;107(44):18933-18938. doi:10.1073/pnas.1007028107

24. Dabrowska K, Witkiewicz W. Correlations of host genetics and gut microbiome composition. *Frontiers in Microbiology*. 2016;7(AUG). doi:10.3389/fmicb.2016.01357

25. Clavel T, Desmarchelier C, Haller D, et al. Intestinal microbiota in metabolic diseases: From bacterial community structure and functions to species of pathophysiological relevance. *Gut Microbes*. 2014;5(4). doi:10.4161/gmic.29331

26. Vojinovic D, Radjabzadeh D, Kurilshikov A, et al. Relationship between gut microbiota and circulating metabolites in population-based cohorts. *Nature Communications*. 2019;10(1). doi:10.1038/s41467-019-13721-1

27. Senghor B, Sokhna C, Ruimy R, Lagier JC. Gut microbiota diversity according to dietary habits and geographical provenance. *Human Microbiome Journal*. 2018;7-8:1-9. doi:10.1016/j.humic.2018.01.001

28. Magne F, Gotteland M, Gauthier L, et al. The firmicutes/bacteroidetes ratio: A relevant marker of gut dysbiosis in obese patients? *Nutrients*. 2020;12(5). doi:10.3390/nu12051474

29. Leamy LJ, Kelly SA, Nietfeldt J, et al. *Host Genetics and Diet, but Not Immunoglobulin A Expression, Converge to Shape Compositional Features of the Gut Microbiome in an Advanced Intercross Population of Mice*.; 2014. http://genomebiology.com/2014/15/12/552

30. den Hartigh LJ, Gao Z, Goodspeed L, et al. Obese mice losing weight due to trans-10,cis-12 conjugated linoleic acid supplementation or food restriction harbor distinct gut microbiota. *Journal of Nutrition*. 2018;148(4):562-572. doi:10.1093/jn/nxy011

31. David LA, Maurice CF, Carmody RN, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature*. 2014;505(7484):559-563. doi:10.1038/nature12820

32. Hansen NW, Sams A. The microbiotic highway to health—New perspective on food structure, gut microbiota, and host inflammation. *Nutrients*. 2018;10(11). doi:10.3390/nu10111590

33. Wang JH, Shin NR, Lim SK, et al. Diet control more intensively disturbs gut microbiota than genetic background in wild type and ob/ob mice. *Frontiers in Microbiology*. 2019;10(JUN). doi:10.3389/fmicb.2019.01292

34. Zhang C, Zhang M, Wang S, et al. Interactions between gut microbiota, host genetics and diet relevant to development of metabolic syndromes in mice. *ISME Journal*. 2010;4(2):232-241. doi:10.1038/ismej.2009.112

35. Wang B, Kong Q, Li X, et al. A high-fat diet increases gut microbiota biodiversity and energy expenditure due to nutrient difference. *Nutrients*. 2020;12(10). doi:10.3390/nu12103197

36. Zheng X, Huang F, Zhao A, et al. Bile acid is a significant host factor shaping the gut microbiome of diet-induced obese mice. *BMC Biology*. 2017;15(1). doi:10.1186/s12915-017-0462-7

37. Sun L, Jia H, Li J, et al. Cecal Gut Microbiota and Metabolites Might Contribute to the Severity of Acute Myocardial Ischemia by Impacting the Intestinal Permeability,

Oxidative Stress, and Energy Metabolism. *Frontiers in Microbiology*. 2019;10. doi:10.3389/fmicb.2019.01745

38. Zhou L, Xiao X, Zhang Q, et al. Improved glucose and lipid metabolism in the early life of female offspring by maternal dietary genistein is associated with alterations in the gut microbiota. *Frontiers in Endocrinology*. 2018;9(SEP). doi:10.3389/fendo.2018.00516

39. Mayneris-Perxachs J, Arnoriaga-Rodríguez M, Luque-Córdoba D, et al. Gut microbiota steroid sexual dimorphism and its impact on gonadal steroids: Influences of obesity and menopausal status. *Microbiome*. 2020;8(1). doi:10.1186/s40168-020-00913-x

40. Kazemian N, Mahmoudi M, Halperin F, Wu JC, Pakpour S. Gut microbiota and cardiovascular disease: Opportunities and challenges. *Microbiome*. 2020;8(1). doi:10.1186/s40168-020-00821-0

41. Vacca M, Celano G, Calabrese FM, Portincasa P, Gobbetti M, de Angelis M. The controversial role of human gut lachnospiraceae. *Microorganisms*. 2020;8(4). doi:10.3390/microorganisms8040573

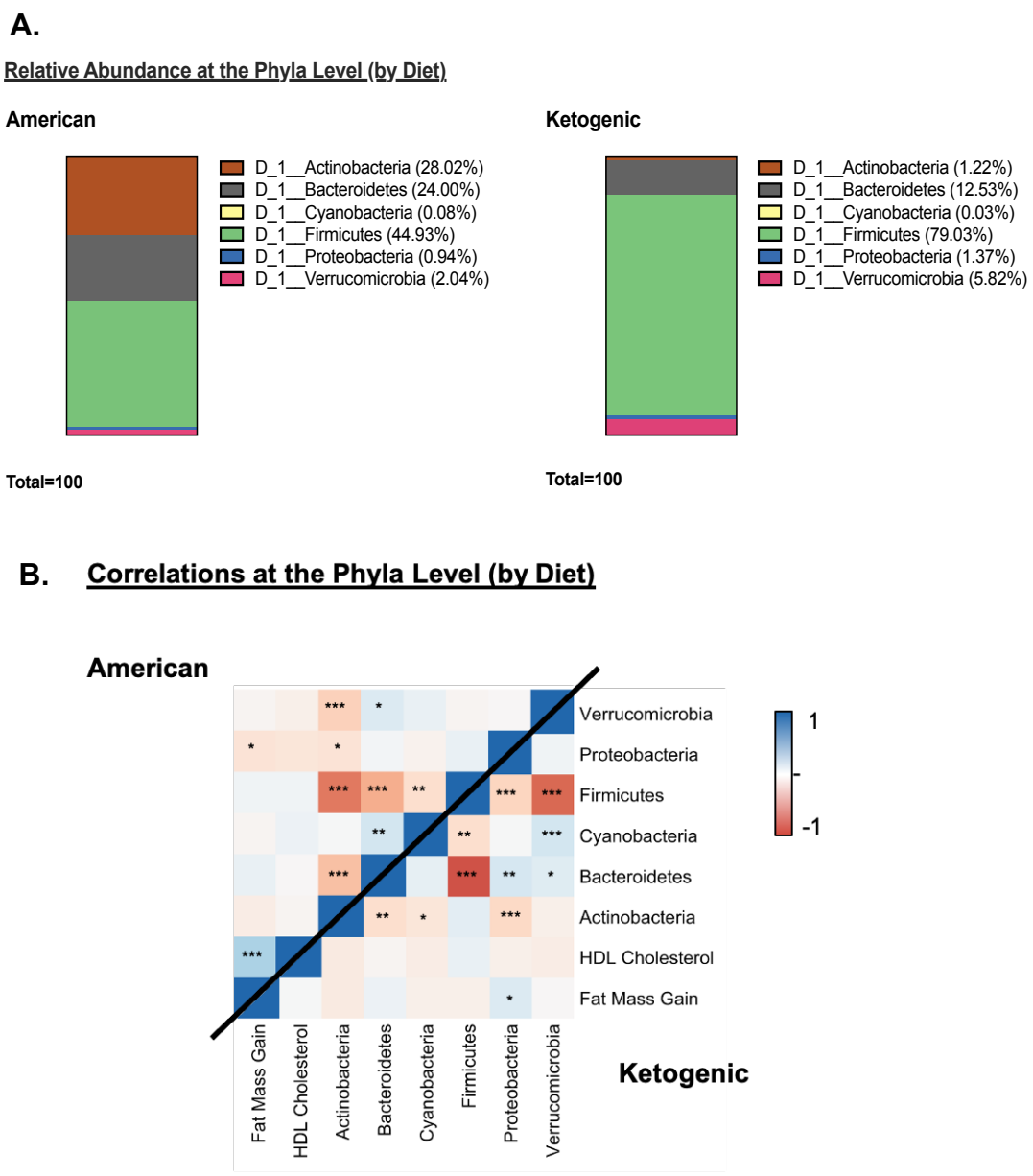

Figure 4. Relative abundance of microbiota and correlations at the phyla level (by diet). A. Relative abundance of Firmicutes in F2s on the ketogenic diet is nearly twice as high as the relative abundance of Firmicutes in F2s on the American diet at the expense of Actinobacteria, and Bacteroidetes for which the relative abundances are lower in F2s on the ketogenic diet. **B.** Abundance of Firmicutes is positively associated with Actinobacteria in F2s on the ketogenic diet (r^2= 0.12, p= 0.063) while abundance of Firmicutes is negatively associated with abundance of Actinobacteria in F2s on the American diet (r^2= -0.66, p< 0.001). Firmicutes in F2s on the ketogenic diet are also negatively associated with Verrucomicrobia (r^2= -0.71, p < 0.001). Firmicutes are negatively associated with Proteobacteria in F2s on the ketogenic diet (r^2= -0.27, p< 0.001). Firmicutes are negatively associated with Bacteroidetes in F2s on both diets however; this correlation is much stronger in F2s on the ketogenic diet (American: r^2= -0.44, p< 0.001, ketogenic: r^2= -0.79, p< 0.001). Fat mass gain is negatively correlated with the abundance of Proteobacteria in F2s on the American diet (r^2= -0.17, p= 0.025) while fat mass gain is positively correlated with Proteobacteria in F2s on the ketogenic diet (r^2= 0.18, p= 0.011).

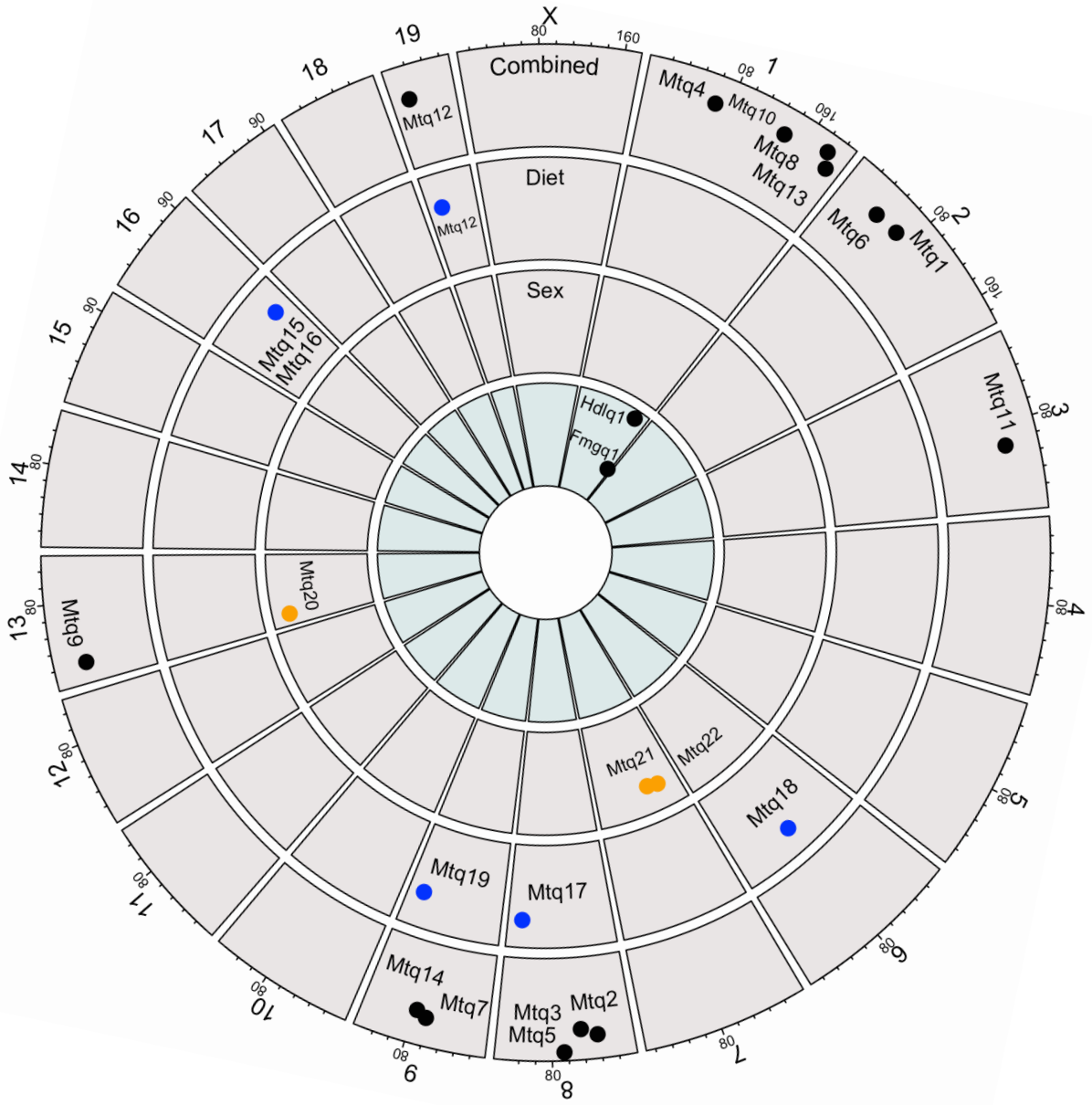

Figure 5. Microbial QTL. Outer ring to inner ring: Significant QTL in the combined model (black) for *Coriobacteriaceae* on Chr2 at *Mtq1* and *Mtq2*, an unidentified *Bacteroidales* bacterium from the *Muribaculaceae* family on Chr8 at *Mtq3*, *Lactobacillus* on Chr1 at *Mtq4*, *Streptococcus* on Chr8 at *Mtq5*, *Lachnospiraceae* FCS020 Group on Chr2 at *Mtq6, Roseburia* on Chr9 at *Mtq7,* unidentified *Lachnospiraceae* on Chr1 at *Mtq8*, *Romboutsia* on Chr13 at *Mtq9, Butyricicoccus* on Chr1 at *Mtq10*, *Dubosiella* on Chr3 at *Mtq11*, *Erysipelotrichaceae* on Chr19 at *Mtq12*, and *Bilophila* on Chr1 at *Mtq13* and Chr9 at *Mtq14*; Significant QTL in the diet specific model (blue) for *Alistipes* on Chr16 at *Mtq15*, *Rikenellaceae* on Chr16 at *Mtq16*, *Streptococcus* on Chr8 at *Mtq17, Erysipelatoclostridium* on Chr6 at *Mtq18*, and unidentified *Erysipelotrichaceae* on Chr9 at *Mtq19* as well as on Chr19 again at *Mtq12.* Significant QTL in the sex specific model (orange) for *Alistipes* on Chr13 at *Mtq20* and for *Lactobacillus* on Chr7 at both *Mtq21* and *Mtq22*; Previously identified QTL for metabolic traits in the combined model (black) for fat mass gain during the feeding trial on Chr1 at *Fmgq1* and serum HDL cholesterol concentration after the feeding trial on Chr1 at *Hdlq1*. *Fmgq1* and *Hdlq1* overlap the same region of the genome as *Mtq8* for unidentified *Lachnospiraceae* and *Mtq13* for *Bilophila*.

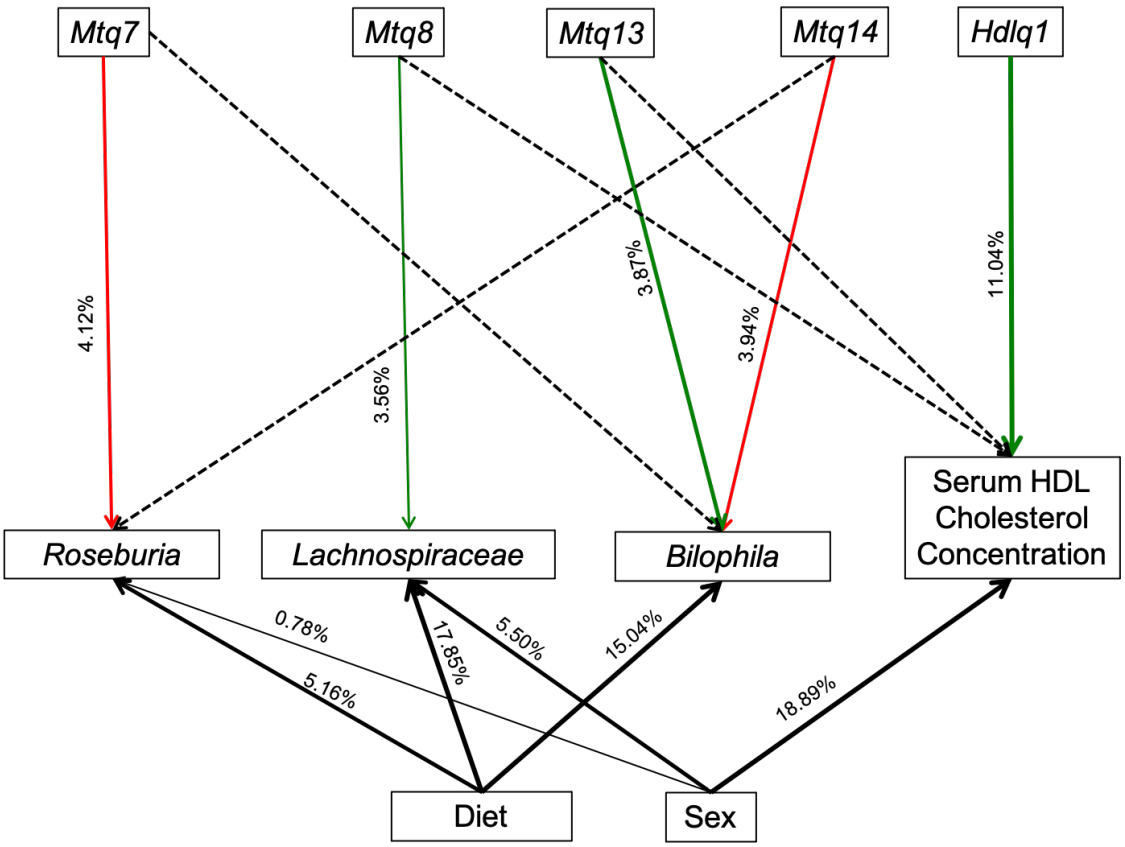

Figure 6. Graphical Representation of SEM for abundance of *Roseburia*, *Lachnospiraceae*, *Bilophila* and serum HDL cholesterol concentration after the feeding trial. Solid, single headed arrows indicate the direction of paths and the weight of each arrow is proportional to the path coefficient from the predictor to the variable and the percentage of variation in the variable that is explained by each predictor. Positive effects (green arrows) indicate that the FVB allele increases the trait; negative effects (red arrows) indicate that the FVB allele decreases the trait. The dashed black arrow represents the inferred causal network at loci which are associated with multiple OTU.

Table 4. Linkage analysis for microbial traits in the combined, diet specific, and sex specific models

Phylum	O T U	Model	QTL Identifier	Chr	Start (Mb)	Top (Mb)	End (Mb)	LOD §	Effect of B6/FVB Alleles §	Effect of FVB/FVB Alleles §	Top Marker (% Variance Explained)	Sex (% Variance Explained)	Diet (% Variance Explained)	Sex * Diet (% Variance Explained)
Actinobacteria	*Coriobacteriaceae UCG 002*	Combined y~m+sex*diet	*Mtq1*	2	60.3	65.8	77.4	4.11[b]	21.22	-1.71	4.12***	5.03***	0.8	1.59**
Actinobacteria	Unidentified *Coriobacteriaceae UCG 002*	Combined y~m+sex*diet	*Mtq1*	2	60.3	65.8	77.4	4.11[b]	21.22	-1.71	4.12***	5.03***	0.8	1.59**
Actinobacteria	*Coriobacteriaceae UCG 002*	Combined y~m+sex*diet	*Mtq2*	8	22.0	34.7	55.1	4.70[b]	-14.30	-33.02	4.66***	5.02***	0.67	1.14*
Actinobacteria	Unidentified *Coriobacteriaceae UCG 002*	Combined y~m+sex*diet	*Mtq2*	8	22.0	34.7	55.1	4.70[b]	-14.30	-33.02	4.66***	5.02***	0.67	1.14*
Bacteroidetes	Unidentified *Bacteroidales bacterium*	Combined y~m+sex*diet	*Mtq3*	8	43.4	51.0	62.8	4.25[b]	0.73	5.97	4.24***	2.98***	3.25***	0.26
Firmicutes	*Lactobacillus*	Combined y~m+sex*diet	*Mtq4*	1	51.1	65.6	79.2	4.39[b]	42.40	131.81	4.23***	4.43***	4.29***	0.4
Firmicutes	*Streptococcus*	Combined y~m+sex*diet	*Mtq5*	8	58.6	68.3	84.1	5.74[b]	3.23	7.99	4.42***	1.14**	24.67****	1.64**
Firmicutes	*Lachnospiraceae FCS020 Group*	Combined y~m+sex*diet	*Mtq6*	2	24.1	40.1	60.3	3.99[b]	0.51	3.52	4.21***	< 0.01	0.85	0.34
Firmicutes	*Roseburia*	Combined y~m+sex*diet	*Mtq7*	9	60.1	65.6	85.4	4.41[b]	-18.92	-35.40	4.40***	0.88*	4.96***	0.17
Firmicutes	Unidentified *Lachnospiraceae*	Combined y~m+sex*diet	*Mtq8*	1	151.9	181.1	186.6	5.08[a]	57.91	75.05	3.97***	5.24***	17.56***	3.70***
Firmicutes	*Romboutsia*	Combined y~m+sex*diet	*Mtq9*	13	5.4	18.4	37.5	4.08[b]	-31.03	-75.89	3.63***	6.33***	6.09***	4.12***
Firmicutes	Unidentified *Romboutsia*	Combined y~m+sex*diet	*Mtq9*	13	5.4	18.4	37.5	4.05[b]	-24.97	-60.52	3.60***	6.60***	5.97***	4.26***
Firmicutes	*Butyricicoccus*	Combined y~m+sex*diet	*Mtq10*	1	133.7	137.8	144.5	4.51[b]	6.11	12.21	4.49***	0.74	7.18***	0.38
Firmicutes	Unidentified *Butyricoccus*	Combined y~m+sex*diet	*Mtq10*	1	133.7	137.8	144.5	4.51[b]	6.11	12.21	4.49***	0.74	7.18***	0.38
Firmicutes	Unidentified *Dubosiella*	Combined y~m+sex*diet	*Mtq11*	3	89.1	102.3	130.4	4.07[b]	-96.29	-108.69	3.86***	0.24	10.60***	0.02
Firmicutes	Unidentified *Erysipelotrichaceae*	Combined y~m+sex*diet	*Mtq12*	19	10.2	14.8	24.1	4.00[b]	3.51	-1.08	3.84***	0.01	9.55***	0.2
Proteobacteria	*Bilophila*	Combined y~m+sex*diet	*Mtq13*	1	144.5	188.6	193.3	4.10[b]	5.95	16.96	3.74***	0.39	15.91***	0.17
Proteobacteria	Unidentified *Bilophila*	Combined y~m+sex*diet	*Mtq13*	1	144.5	188.6	193.3	4.10[b]	5.95	16.96	3.74***	0.39	15.91***	0.17
Proteobacteria	*Bilophila*	Combined y~m+sex*diet	*Mtq14*	9	42.3	76.4	80.8	4.05[b]	-9.88	-15.28	3.66***	0.23	15.94***	0.11
Proteobacteria	Unidentified *Bilophila*	Combined y~m+sex*diet	*Mtq14*	9	42.3	76.4	80.8	4.05[b]	-9.88	-15.28	3.66***	0.23	15.94***	0.11

Table 5. Conditioned linkage analysis, causal network analysis, and structural modeling of overlapping QTL

Conditioned and Unconditioned QTL Scans				
Phenotype	**Model**	**Condition**	**QTL**	**LOD**
Roseburia	Combined (y~ sex * diet + marker)	Unconditioned	*Mtq7*	4.41
	Combined (y~ sex * diet + marker + *Bilophila*)	*Bilophila*	Mtq7	4.67
Lachnospiraceae	Combined (y~ sex * diet + marker)	Unconditioned	*Mtq8*	5.08
	Combined (y~ sex * diet + marker + *Bilophila*)	*Bilophila*	*Mtq8*	6.91
Lachnospiraceae	Combined (y~ sex * diet + marker)	Unconditioned	*Mtq8*	5.08
	Combined (y~ sex * diet + marker + fat mass gained)	Fat mass gained	*Mtq8*	5.04
Lachnospiraceae	Combined (y~ sex * diet + marker)	Unconditioned	*Mtq8*	5.08
	Combined (y~ sex * diet + marker + serum HDL cholesterol)	Serum HDL cholesterol	*Mtq8*	4.2
Bilophila	Combined (y~ sex * diet + marker)	Unconditioned	*Mtq13*	4.1
	Combined (y~ sex * diet + marker + fat mass gained)	Fat mass gained	*Mtq13*	3.54
Bilophila	Combined (y~ sex * diet + marker)	Unconditioned	*Mtq13*	4.1
	Combined (y~ sex * diet + marker + serum HDL cholesterol)	Serum HDL cholesterol	*Mtq13*	3.7
Bilophila	Combined (y~ sex * diet + marker)	Unconditioned	*Mtq13*	4.1
	Combined (y~ sex * diet + marker + *Lachnospiraceae*)	*Lachnospiraceae*	Mtq13	5.65
Bilophila	Combined (y~ sex * diet + marker)	Unconditioned	*Mtq14*	4.05
	Combined (y~ sex * diet + marker + *Roseburia*)	*Roseburia*	*Mtq14*	4.28
Fat mass gained (during the feeding trial) (g)	Combined (y~ sex * diet + marker)	Unconditioned	*Fmgq1*	4.02
	Combined (y~ sex * diet + marker + *Bilophila*)	*Bilophila*	*Fmgq1*	4.05
Fat mass gained (during the feeding trial) (g)	Combined (y~ sex * diet + marker)	Unconditioned	*Fmgq1*	4.02
	Combined (y~ sex * diet + marker + *Lachnospiraceae*)	*Lachnospiraceae*	*Fmgq1*	4.02
Serum HDL cholesterol concentration (after the feeding trial) (ng/mL)	Combined (y~ sex * diet + marker)	Unconditioned	*Hdlq1*	11.94
	Combined (y~ sex * diet + marker + *Bilophila*)	*Bilophila*	*Hdlq1*	12.01
Serum HDL cholesterol concentration (after the feeding trial) (ng/mL)	Combined (y~ sex * diet + marker)	Unconditioned	*Hdlq1*	11.94
	Combined (y~ sex * diet + marker + *Lachnospiraceae*)	*Lachnospiraceae*	*Hdlq1*	12.04
Causal Network Analysis				
Relationship	**Model**	**p-value**	**AIC**	**BIC**
Bilophila <-- *Roseburia*	(*Bilophila* ~ *Roseburia*) ~ sex * diet + *Mtq7*	0.015	3697	3725
Bilophila --> *Roseburia*	(*Roseburia* ~ *Bilophila*) ~ sex * diet + *Mtq7*	0.015	4261	4289
Lachnospiraceae <-- serum HDL cholesterol	(*Lachnospiraceae* ~ HDL) ~ sex * diet + *Mtq8*	p < 0.001	4816	4843
Lachnospiraceae --> serum HDL cholesterol	(HDL ~ *Lachnospiraceae*) ~ sex * diet + *Mtq8*	p < 0.001	4097	4124
Serum HDL cholesterol <-- *Lachnospiraceae*	(HDL ~ *Lachnospiraceae*) ~ sex * diet + *Hdlq1*	p < 0.001	4093	4121
Serum HDL cholesterol --> *Lachnospiraceae*	(*Lachnospiraceae* ~ HDL) ~ sex * diet + *Hdlq1*	p < 0.001	4829	4856
Bilophila <-- serum HDL cholesterol	(*Bilophila* ~ HDL) ~ sex * diet + *Mtq13*	p < 0.001	3611	3638
Bilophila --> serum HDL cholesterol	(HDL ~ *Bilophila*) ~ sex * diet + *Mtq13*	p < 0.001	4021	4049
Bilophila <-- *Roseburia*	(*Bilophila* ~ *Roseburia*) ~ sex * diet + *Mtq14*	0.016	3627	3654
Bilophila --> *Roseburia*	(*Roseburia* ~ *Bilophila*) ~ sex * diet + *Mtq14*	0.01	4194	4221
Structural Model				
	Variable	**Predictor**	**Path coefficient (% variation explained)**	**t-statistic of Path Coefficient**
	Roseburia	Sex	0.11 (0.78%)	2.27
		Diet	-0.22 (5.16%)	-4.38
		Mtq7	-0.17 (4.12%)	-3.78
	Lachnospiraceae	Sex	0.27 (5.50%)	7.04
		Diet	-0.42 (17.85%)	-9.52
		Mtq8	0.15 (3.56%)	3.75
	Bilophila	Diet	-0.38 (15.04%)	-8.26

Table 6. KEGG Pathway annotation of positional candidate genes at *Mtq7*, *Mtq8*, *Mtq13*, *Mtq14*, and *Hdlq1*.

QTL	Model	Phylum	O T U	Gene Symbol (MGI)	KEGG Pathway
Mtq7	Combined y~m+sex*diet	Firmicutes	*Roseburia*	*Aqp9*	Bile secretion (mmu04976)
				Col12a1	Protein digestion and absorption (mmu04974)
				Cox7a2	Non-alcoholic fatty liver disease (NAFLD) (mmu04932)
				Hmgcll1	Peroxisome (mmu04146)
				Map2k1	Insulin signaling pathway (mmu04910)
				Onecut1	Maturity onset diabetes of the young (MODY) (mmu04950)
				Slc51b	Bile secretion (mmu04976)
Mtq8	Combined y~m+sex*diet	Firmicutes	*Lachnospiraceae*	*Akt3*	AMPK signaling pathway (mmu04152)
					Insulin signaling pathway (mmu04910)
					Adipocytokine signaling pathway (mmu04920)
					Regulation of lipolysis in adipocytes (mmu04923)
					Insulin resistance (mmu04931)
					Non-alcoholic fatty liver disease (NAFLD) (mmu04932)
					Carbohydrate digestion and absorption (mmu04973)
				Aldh9a1	Glycolysis/Gluconeogenesis (mmu00010)
				Apoa2	PPAR signaling pathway (mmu03320)
				Atp1a2	Carbohydrate digestion and absorption (mmu04973)
					Protein digestion and absorption (mmu04974)
					Bile secretion (mmu04976)
				Atp1a4	Carbohydrate digestion and absorption (mmu04973)
					Protein digestion and absorption (mmu04974)
					Bile secretion (mmu04976)
				Atp1b1	Carbohydrate digestion and absorption (mmu04973)
					Protein digestion and absorption (mmu04974)
					Bile secretion (mmu04976)
				Cacna1e	Type II diabetes mellitus (mmu04930)
				Ephx1	Bile secretion (mmu04976)
				Fasl	Non-alcoholic fatty liver disease (NAFLD) (mmu04932)
					Type I diabetes mellitus (mmu04940)
				Hsd17b7	Steroid biosynthesis (mmu00100)
					Steroid hormone biosynthesis (mmu00140)
					Ovarian steroidogenesis (mmu04913)
				Ndufs2	Non-alcoholic fatty liver disease (NAFLD) (mmu04932)
				Pex19	Peroxisome (mmu04146)
				Rxrg	PPAR signaling pathway (mmu03320)
					Adipocytokine signaling pathway (mmu04920)
				Sdhc	Non-alcoholic fatty liver disease (NAFLD) (mmu04932)
				Soat1	Steroid biosynthesis (mmu00100)

Table 6. KEGG Pathway annotation of positional candidate genes at *Mtq7*, *Mtq8*, *Mtq13*, *Mtq14*, and *Hdlq1* (continued).

QTL	Model	Phylum	O T U	Gene Symbol (MGI)	KEGG Pathway
Mtq14	Combined y~m+sex*diet	Proteobacteria	*Bilophila*	*Acsbg1*	Adipocytokine signaling pathway (mmu04920)
					Fatty acid biosynthesis (mmu00061)
					PPAR signaling pathway (mmu03320)
				Adpgk	Glycolysis/Gluconeogenesis (mmu00010)
				Apoa1	Fat digestion and absorption (mmu04975)
					PPAR signaling pathway (mmu03320)
				Apoa4	Fat digestion and absorption (mmu04975)
				Apoa5	PPAR signaling pathway (mmu03320)
				Apoc3	PPAR signaling pathway (mmu03320)
				Aqp9	Bile secretion (mmu04976)
				Cbl	Insulin signaling pathway (mmu04910)
				Col12a1	Protein digestion and absorption (mmu04974)
				Cox5a	Non-alcoholic fatty liver disease (NAFLD) (mmu04932)
				Cox7a2	Non-alcoholic fatty liver disease (NAFLD) (mmu04932)
				Cyp11a1	Aldosterone synthesis and secretion (mmu04925)
					Ovarian steroidogenesis (mmu04913)
					Steroid hormone biosynthesis (mmu00140)
				Cyp19a1	Ovarian steroidogenesis (mmu04913)
					Steroid hormone biosynthesis (mmu00140)
				Cyp1a1	Ovarian steroidogenesis (mmu04913)
					Steroid hormone biosynthesis (mmu00140)
				Cyp1a2	Steroid hormone biosynthesis (mmu00140)
				Dlat	Glycolysis/Gluconeogenesis (mmu00010)
				Fxyd2	Bile secretion (mmu04976)
					Carbohydrate digestion and absorption (mmu04973)
					Protein digestion and absorption (mmu04974)
				Hmgcll1	Peroxisome (mmu04146)
				Map2k1	Insulin signaling pathway (mmu04910)
				Onecut1	Maturity onset diabetes of the young (MODY) (mmu04950)
				Pkm	Glycolysis/Gluconeogenesis (mmu00010)
					Type II diabetes mellitus (mmu04930)
				Ppp2r1b	AMPK signaling pathway (mmu04152)
				Sc5d	Steroid biosynthesis (mmu00100)
					Non-alcoholic fatty liver disease (NAFLD) (mmu04932)
				Slc37a4	Carbohydrate digestion and absorption (mmu04973)
				Slc51b	Bile secretion (mmu04976)

Table 6. KEGG Pathway annotation of positional candidate genes at *Mtq7*, *Mtq8*, *Mtq13*, *Mtq14*, and *Hdlq1* (continued).

QTL	Model	Phylum	O T U	Gene Symbol (MGI)	KEGG Pathway
Mtq14	Combined y~m+sex*diet	Proteobacteria	*Bilophila*	*Acsbg1*	Adipocytokine signaling pathway (mmu04920)
					Fatty acid biosynthesis (mmu00061)
					PPAR signaling pathway (mmu03320)
				Adpgk	Glycolysis/Gluconeogenesis (mmu00010)
				Apoa1	Fat digestion and absorption (mmu04975)
					PPAR signaling pathway (mmu03320)
				Apoa4	Fat digestion and absorption (mmu04975)
				Apoa5	PPAR signaling pathway (mmu03320)
				Apoc3	PPAR signaling pathway (mmu03320)
				Aqp9	Bile secretion (mmu04976)
				Cbl	Insulin signaling pathway (mmu04910)
				Col12a1	Protein digestion and absorption (mmu04974)
				Cox5a	Non-alcoholic fatty liver disease (NAFLD) (mmu04932)
				Cox7a2	Non-alcoholic fatty liver disease (NAFLD) (mmu04932)
				Cyp11a1	Aldosterone synthesis and secretion (mmu04925)
					Ovarian steroidogenesis (mmu04913)
					Steroid hormone biosynthesis (mmu00140)
				Cyp19a1	Ovarian steroidogenesis (mmu04913)
					Steroid hormone biosynthesis (mmu00140)
				Cyp1a1	Ovarian steroidogenesis (mmu04913)
					Steroid hormone biosynthesis (mmu00140)
				Cyp1a2	Steroid hormone biosynthesis (mmu00140)
				Dlat	Glycolysis/Gluconeogenesis (mmu00010)
				Fxyd2	Bile secretion (mmu04976)
					Carbohydrate digestion and absorption (mmu04973)
					Protein digestion and absorption (mmu04974)
				Hmgcll1	Peroxisome (mmu04146)
				Map2k1	Insulin signaling pathway (mmu04910)
				Onecut1	Maturity onset diabetes of the young (MODY) (mmu04950)
				Pkm	Glycolysis/Gluconeogenesis (mmu00010)
					Type II diabetes mellitus (mmu04930)
				Ppp2r1b	AMPK signaling pathway (mmu04152)
				Sc5d	Steroid biosynthesis (mmu00100)
					Non-alcoholic fatty liver disease (NAFLD) (mmu04932)
				Slc37a4	Carbohydrate digestion and absorption (mmu04973)
				Slc51b	Bile secretion (mmu04976)

CHAPTER IV: SEX AND GENETIC BACKGROUND REGULATE RESPONSE TO A KETOGENIC DIETARY INTERVENTION IN MICE

Abstract

Efforts to use genetic information to predict dietary responses continue to proliferate. In previous studies, we have demonstrated differential responses to American and ketogenic diets in C57BL/6J (B6) and FVB/NJ (FVB) mice during lifetime feeding trials. In this study we characterize the utility of the ketogenic diet as a dietary intervention in B6 and FVB mice. B6 females that received the ketogenic diet as a dietary intervention after 3 months of exposure to the American diet gained less fat mass than B6 females that remained on the American diet for the duration of the feeding trial (B6 female American: 15.94g, +/- 5.00g ; B6 female Reversal: 9.67g, +/- 3.98g; $p = 0.007$, 95% CI: 1.91-10.63g). B6 males exposed to the dietary intervention gained more fat mass than B6 males that remained on the American diet for the duration of the feeding trial (B6 male American: 19.69g, +/- 1.42g; B6 male Reversal: 23.86g, +/- 1.28g; $p < 0.001$, 95% CI: -5.43 – (-2.89) g). The amount of fat mass gained on the American diet prior to the introduction of the dietary intervention was highly correlated to the amount of fat mass gained at the end of the feeding trial for all strain and diet combinations ($r = 0.878$, $p < 0.001$, 95% CI: 0.81-0.92). By performing a sex, strain, and diet specific analysis were we able to determine that the ketogenic dietary intervention was most effective for B6 females. The strong relationship observed between the amount of fat mass gained prior to the introduction of the dietary intervention and the amount of fat mass gained at the end of the feeding trial suggests that there are mechanisms contributing to obesity that are not attenuated by

simple dietary intervention even within genetic backgrounds predicted to respond to carbohydrate restriction.

Introduction

Studies of diets varied in macronutrient composition have limitations in both clinical and pre-clinical populations. Human clinical trials examining the effects of carbohydrate restriction are generally underpowered to perform both sex and genotype specific analyses and thus choose to stratify data by either sex[1] or genotype[2–8]. These experiments are typically weakened by limited adherence from participants and short durations of experimental conditions[2–4,9,10]. Furthermore, most studies purported to determine the effects of specific variations in macronutrient content purposefully or inadvertently limit and/or otherwise alter energy intake in the context of the dietary intervention[2,3,7,9–13]. Thus, the effects of macronutrient distributions cannot be detangled from the effects of caloric restriction when conclusions are drawn from these studies.

Pre-clinical studies in mice overcome many limitations encountered in human clinical trials. Mice are an excellent resource for conducting high-powered studies for sex specific, and genotype specific analyses. Adherence to experimental diets in pre-clinical work is 100% and experiments can be conducted over long periods of time. Energy consumption can be accurately measured in the context of dietary interventions that are varied in macronutrient content, which allows for the detangling of effects of caloric intake and macronutrient distribution in analyses. However, animals are typically exposed to one diet for the majority of their lifetimes and predictions/conclusions are based on unrealistic

dietary conditions. In human populations dietary interventions are generally not trialed until an individual desires weight loss. Therefore, it is critical to test the utility of pre-clinical data in the context of a dietary intervention and not just in the context of a lifetime exposure to experimental diets.

Previous work has demonstrated striking differences between C57BL/6J (B6) and FVB/NJ (FVB) mice in response to high-fat diets varying in carbohydrate content[14–17]. Recently, the genetic origin was investigated of differential response to carbohydrate restriction and sensitivity to high fat diets utilizing an intercross population (F2) between B6 and FVB. Genetic loci were identified that contribute to differential responses to American (35% of energy from fat, 50% from carbohydrate) and ketogenic (84% of energy from fat, 0% from carbohydrate) diets and directly or indirectly affect body composition, serum cholesterol concentration, and abundances of gut microbiota within various strain, sex, and diet groups.

To further investigate the utility of genetic information in predicting diet response, B6 and FVB mice were exposed to the American diet for 3 months before providing the ketogenic diet to half of the animals as a dietary intervention. In this population, the ketogenic dietary intervention was determined to be most effective for B6 females. A strong relationship was observed between the amount of fat mass gained prior to the introduction of the dietary intervention and the amount of fat mass gained at the end of the feeding trial, which appears to limit weight loss in B6 males after the introduction of the dietary intervention. The importance of sex as a biological variable was demonstrated, especially

in conjunction with genetic data, to determine response to dietary interventions and highlight the implications for conclusions when data analysis is conducted at the population and individual level.

Methods

Animals and Diets

Four-to six-week-old B6 and FVB mice were exposed to an American (35% of energy from fat, 50% from carbohydrates) diet for 3 months. After 3 months on the American diet, half of the animals (10 male B6, 10 female B6, 10 male FVB, and 10 female FVB) were reversed to a ketogenic (84% of energy from fat, 0% from carbohydrates) dietary intervention for the remainder of the 6-month feeding trial while the other half were left on the American diet. These strains have previously exhibited significantly different responses to these two diets[14]. This initial study provided evidence for differential responses to American and ketogenic diets in B6 while FVB were observed to have similar responses to both high fat diets. Detailed diet compositions are provided in Appendix Table A1. Animals that received the dietary intervention were weaned onto the ketogenic diet over 2 weeks by mixing the American and ketogenic diets and gradually reducing the amount of American diet at each feeding. This weaning procedure eliminates spurious weight loss associated with stress induced by abrupt introduction of the dietary intervention.

All animals were maintained in accordance with Texas A&M University Institution Animal Care and Use Committee guidelines at 22 °C under a 12-hour light cycle. At the end of

the feeding trial, mice were euthanized by carbon dioxide asphyxiation, blood was collected, and tissues were harvested and immediately flash frozen in liquid nitrogen.

Phenotyping

Echo Magnetic Resonance Spectroscopy (MRI) (EchoMRI, Houston, TX, USA) was used to measure fat and lean mass of all individuals prior to the feeding trial, after 3 months on the American diet, 2-weeks post intervention, 1 month post intervention, 2 months post-intervention, and 3 months post-intervention. Fat and lean mass gains were calculated as the difference in fat and lean mass prior to the feeding trial and after each timepoint on the assigned diet.

Mice were singly housed in Phenomaster Metabolic Chambers (TSE Systems) at 3 months (prior to dietary intervention) and again at the end of the 6-month feeding trial. After a 24-hour acclimation period, data collection included two 12-hour day and two 12-hour night cycles. Heat expenditure was calculated per hour during the data collection period. Calorie consumption was measured during the last 24 hours of the data collection period.

Statistical Analysis

Linear models using ANOVA was used to detect interactions among strain, sex, and diet assignment. Tukey HSD was applied to determine which differences among means were significant. A non-parametric Welches Two Sample t-test was used to compare means for non-normally distributed traits. Pearson's correlation was determined to assess associations with fat mass gained during the feeding trial.

Early deaths occurred in four animals, 3 from unknown cause (1 FVB Male assigned to the dietary intervention, 1 FVB Female assigned to the dietary intervention, 1 FVB Female assigned to the American diet) and 1 that occurred during an unrelated data collection procedure (1 B6 Female assigned to the dietary intervention). These animals were excluded from all analyses.

Results

Fat and lean mass gained during the feeding trial

At baseline, no significant differences were observed between fat mass gained during the feeding trial after 3 months on the American diet and prior to the dietary intervention (Figure 7A; B6 male American: 13.29g, +/- 1.71g, B6 male Reversal: 13.17g, +/- 2.29g; p = 0.897, 95% CI: -1.79 - 2.03g; FVB male American: 4.22g, +/- 1.00g, FVB male Reversal: 6.07g, +/- 2.39g; p = 0.054, 95% CI: -3.75-0.04g; B6 female American: 4.81g, +/- 2.47g, B6 female Reversal: 4.38g, +/- 1.69g; p = 0.665, 95% CI: -1.61 - 2.46g; FVB female American: 0.91g, +/- 1.42g, FVB female Reversal: 1.38g, +/- 0.98g; p = 0.427, 95% CI: -1.70-0.76g). By the end of the feeding trial, carbohydrate restriction did not reduce fat mass in B6 males but did in B6 females. B6 males that received the dietary intervention gained more weight by the end of the feeding trial than B6 males that remained on the American diet for the duration of the trial (Figure 7A; B6 male American: 19.69g, +/- 1.42g; B6 male Reversal: 23.86g, +/- 1.28g; $p < 0.001$, 95% CI: -5.43 – (-2.89) g). In contrast, B6 females that received the dietary intervention gained significantly less weight by the end of the feeding trial than B6 females that remained on the American diet for the

duration of the feeding trial (Figure 7A; B6 female American: 15.94g, +/- 5.00g; B6 female Reversal: 9.67g, +/- 3.98g; p = 0.007, 95% CI: 1.91-10.63g). As expected based on previous studies, FVB males and females did not demonstrate a differential response to the American diet or dietary intervention by the end of the feeding trial (Figure 7A; FVB male American: 11.74g, +/- 0.872g, FVB male Reversal 10.32g, +/- 3.60g; p = 0.280, 95% CI: -1.38-4.21g; FVB female American: 5.97g, +/- 2.59g, FVB female Reversal: 6.91g, +/- 2.27g; p = 0.424, 95% CI: -3.38-1.50g).

After the feeding trial, B6 males that received the dietary intervention had a higher percentage of body fat than B6 males that remained on the American diet for the duration of the feeding trial (Figure 7B; B6 male American: 42%, +/- 2%; B6 male Reversal: 48%, +/- 1%; p < 0.001, 95% CI: -7%-3%). B6 males that received the dietary intervention gained 1.17g less lean mass during the feeding trial than B6 males that remained on the American diet for the duration of the feeding trial (Figure 7C; B6 male American: 9.85g, +/- 0.93g; B6 male Reversal: 8.68g, +/- 1.15g; p = 0.023, 95% CI: 0.18-2.16g). B6 females that received the dietary intervention had a lower percentage of body fat than B6 females that remained on the American diet for the duration of the feeding trial (Figure 7B; B6 female American: 44%, +/- 8%; B6 female Reversal: 35%, +/- 8%; p = 0.028, 95% CI: 1%-17%). However, there was no difference in the amount of lean mass gained by B6 females that received the dietary intervention and B6 females that remained on the American diet for the duration of the feeding trial (Figure 7C; B6 female American: 7.33g, +/- 1.01g; B6 female Reversal: 6.59g, +/- 0.69; p = 0.078, 95% CI: -0.09-1.57g). Lastly, there was no differential response in terms of body fat percentage or lean mass gained

during exposure to the American diet or dietary intervention for FVB males or females (Figure 7B; FVB male American: 34%, +/- 1%, FVB male Reversal 32%, +/- 6%; p = 0.412, 95% CI: -3-6%; FVB female American: 28%, +/- 6%, FVB female Reversal: 30%, +/- 5%; p = 0.329, 95% CI: -8 – 3%); Figure 7C; FVB male American: 7.71g, +/- 1.33g, FVB male Reversal 6.74g, +/- 0.94g; p = 0.083, 95% CI: -0.14 – 2.08g; FVB female American: 5.43g, +/- 1.00g, FVB female Reversal: 4.88g, +/- 1.00g; p = 0.253, 95% CI: -0.44 – 1.56g.

Associations with fat mass gained during the feeding trial

The observed differences in fat mass gained during the feeding trial at 6 months do not have a strong association with caloric intake. A weak, inverse relationship between the amount of calories consumed and the amount of fat mass gained in each diet, strain, and sex specific group was observed (Figure 8 (A); r = -0.454, p < 0.001, 95% CI: -0.62- (-0.25)). Heat expenditure and fat mass gained during the feeding trial at 6 months (3 months post-intervention) were inversely correlated at the end of the feeding trial (Figure 8 (B); r= -0.832, p < 0.001; 95% CI: -0.89- (-0.75)). This demonstrates that there is a stronger association between heat expenditure and fat mass gained during the feeding trial than observed between caloric intake and fat mass gained during the feeding trial. Interestingly, fat mass gained during the feeding trial at 3 months, prior to the dietary intervention was also highly correlated to the amount of fat mass gained during the feeding trial at 6 months (3 months post-intervention) (Figure 8 (C); r = 0.878, p < 0.001, 95% CI: 0.81- 0.92). This was the strongest of the three associations. B6 males gained more weight than any other strain and sex combination after 3 months on the American

diet, prior to the dietary intervention (Figure 7A; B6 males: 13.23g, +/- 1.97g; B6 females: 4.61g, +/- 2.09g; FVB males: 5.10g, +/- 1.99g; FVB females: 1.15g, +/- 1.21g; $p < 0.001$, 95% CI = 8.45-10.68g).

No significant differences in caloric intake was observed among any of the strain and sex specific diet groups. (Figure 8 (D); B6 male American: 10.95kcal, +/- 3.94kcal, B6 male Reversal 8.30kcalg, +/- 3.30kcal; $p = 0.121$, 95% CI: -0.78 -6.07kcal; FVB male American: 11.49kcal, +/- 3.16kcal, FVB male Reversal 14.20kcal, +/- 3.43kcal; $p = 0.093$, 95% CI: -5.91-0.51kcal; B6 female American: 9.60kcal, +/- 2.94kcal, B6 female Reversal: 13.53kcal, +/- 6.19kcal; $p = 0.110$, 95% CI: -8.90 – 1.04kcal; FVB female American: 14.08kcal, +/- 3.01kcal, FVB female Reversal: 13.53kcal, +/- 5.93kcal; $p = 0.809$, 95% CI: -4.29- 5.38kcal).

B6 females that received the dietary intervention demonstrate significantly higher heat expenditure than B6 females that remained on the American diet for the duration of the feeding trial (Figure 8 (E); B6 female American: 13.16kcal/kg/hr, +/- 2.41 kcal/kg/hr, B6 female Reversal: 17.58 kcal/kg/hr, +/- 3.09 kcal/kg/hr; $p = 0.004$, 95% CI: -7.14 – (-1.68)kcal). There were no differences in heat expenditure observed among the other strain and sex specific diet groups (Figure 8 (E); B6 male American: 11.69kcal/kg/hr, +/- 0.95kcal/kg/hr, B6 male Reversal: 11.28kcal/kg/hr, +/- 0.60kcal/kg/hr, $p = 0.257$, 95% CI = -0.34-1.17kcal/kg/hr; FVB male American 13.61kcal/kg/hr, +/- 2.41kcal/kg/hr, FVB male Reversal: 13.08kcal/kg/hr, +/- 1.97kcal/kg/hr, $p = 0.490$, 95% CI = -1.08-2.13kcal/kg/hr;

FVB female American 18.14kcal/kg/hr, +/- 1.63kcal/kg/hr, FVB female Reversal 17.93kcal/kg/hr, +/- 1.80kcal/kg/hr, p = 0.799, 95% CI = -1.51-1.93kcal/kg/hr).

Population vs Individual Level Analyses

Classical approaches to analyzing the response to a dietary intervention examine the data at the population level and may or may not include a sex specific analysis. Using these classical approaches to analyze data at the population level, no significant difference was observed between animals that received the dietary intervention when compared to animals that continued the American diet for the duration of the feeding trial (Figure 9; American: 13.53g +/- 5.82g, Reversal: 12.99g +/- 7.38g, p = 0.729, 95% CI: -2.52-3.59). Similarly, no significant difference was observed in males exposed to either diet (Figure 9; Male American: 15.72g +/- 4.24g, Male Reversal 17.44g +/- 7.40g, p = 0.382, 95% CI: -5.71-2.25g) or females exposed to either diet (Figure 9; Female American: 11.22g +/- 6.45g, Female Reversal: 8.29g +/- 3.45g, p = 0.737, 95% CI: -0.53-6.39g). In fact, even when strain is incorporated into a population level analyses, there are no differences observed between the two treatments with respect to genetic background (Figure 9; B6 American: 17.82g, +/- 4.06g, B6 Reversal: 17.14g, +/- 7.80g, p = 0.737, 95% CI: -3.44-4.80g; FVB American: 9.01g, +/- 3.48g, FVB Reversal: 8.62g, +/- 3.41g, p = 0.733, 95% CI: -1.91-2.69). It is not until the data is stratified by diet, strain, and sex that individual effects of the dietary intervention are observed that would be the most effective for B6 females, rather than B6 males.

Discussion

B6 females that received the dietary intervention gained less weight than B6 females that remained on the American diet for the duration of the feeding trial. In contrast, B6 males that received the dietary intervention gained more weight than B6 males that remained on the American diet for the duration of the feeding trial. This was a surprising result given prior data suggesting that B6 males would respond best to a ketogenic dietary intervention[17]. We observed that B6 males that received the intervention have a higher percentage of body fat at the end of the feeding trial and gain less lean mass during the feeding trial than B6 males that remained on the American diet for the duration of the feeding trial. B6 females that received the intervention had a lower percentage of body fat and gained a similar amount of lean mass as B6 females that remained on the American diet for the duration of the feeding trial. As expected, no differences were observed between FVB males or females that received the dietary intervention when compared to their American diet counterparts.

Obesity is associated with reductions in lean body mass and decreased capacity for fatty acid oxidation as skeletal muscle plays a critical role in beta oxidation, especially in response to high fat diets[18–20]. It has been shown that exercise training results in increased capacity for fatty acid oxidation, even in individuals with obesity[18]. This is likely due to maintenance or even increases in lean body mass. The reduced amount of lean mass gained by B6 males that received the dietary intervention may reduce their capacity for fatty acid oxidation and decrease the efficacy of the dietary intervention. However, B6 females that received the dietary intervention gain similar amounts of lean mass as B6

females that remained on the American diet for the duration of the feeding trial. This suggests there the efficacy of the dietary intervention is not dependent upon increased capacity of fatty acid oxidation in B6 females.

Most studies have largely used lifetime feeding trials to demonstrate the effectiveness of ketogenic diets in B6 mice for reducing fat mass gained during feeding trials[21]. In contrast, Kennedy et. al. demonstrated the effects of ketogenic diet on high fat diet induced obesity in B6. Similar to the current study, Kennedy et. al. exposed B6 to high fat diet for 3 months prior to introducing the ketogenic diet as a dietary intervention. They reported that B6 males and females reversed to the ketogenic diet lost all excess body weight within 1 month after the introduction of the dietary intervention. One striking difference between the current study and Kennedy et. al. is that the dietary intervention was introduced by weaning the animals onto the ketogenic diet over 2 weeks to eliminate spurious changes to body composition associated with stress induced by the abrupt introduction of the dietary intervention. In our pilot studies, it was observed that >1 month after the dietary intervention, the spurious changes to the percentage of body fat observed early on eventually plateaued when the intervention was introduced without a weaning period (Appendix Figure A7). The weight loss observed in the study conducted by Kennedy et. al. might have been the result of this stress from the diet transition.

B6 males gained dramatically more fat mass on the American diet during the 3 months prior to the introduction of the dietary intervention, consistent with observations that B6 males are particularly sensitive to high fat diets. Casimiro et. al. has recently shown that

after 3 months on a western diet comparable to the current American diet, B6 males gained significantly more fat mass on the western diet than they did on a low fat diet[22]. This group observed no difference for B6 females on western and low fat diets. From their study, it is visually apparent that males gained more fat mass than females while on the western diet, although this was not a direct comparison that they reported in their findings. Fat mass gain at 3 months was strongly correlated with fat mass gain at 6 months (3 months after the dietary intervention was introduced). This observation suggests that there may be a mechanism contributing to obesity in B6 males on the American diet that cannot be reversed by a simple dietary intervention. It appears as if once B6 males gain sufficient fat mass, it may be harder to lose the fat mass and/or stop the cycle of fat gain even with a personalized dietary intervention. Although the genetic signal is stronger in B6 males according to previous work, the phenotype may be more plastic in B6 females.

This observation adds a layer of complexity to the development of techniques to implement precision nutrition and provide personalized dietary recommendations. Lifetime dietary exposures in pre-clinical models may not be effective for making predictions of dietary response, even in pre-clinical models of dietary intervention. Pre-clinical research should be conducted or at least validated in the context of dietary intervention to demonstrate the utility of the results in achieving desired outcomes[20].

A noteworthy observation was made regarding the weak and inverse relationship between caloric intake and fat mass gained during the feeding trial. This observation highlights a major weakness of clinical trial data for response to carbohydrate restriction

where many studies have purposefully or inadvertently restricted caloric intake as well as carbohydrate intake and confounded the conclusions regarding which intervention is achieving the outcomes observed. Here, carbohydrate restriction, without caloric restriction, was able to prevent further weight gain in B6 females that received the dietary intervention. Kennedy et. al. also showed that B6 mice on a ketogenic diet consume the same amount of calories as B6 mice on a high fat diet and that caloric intake is not correlated to mouse weight[23].

In human trials, several groups have suggested that weight loss is similar on low fat and low carbohydrate dietary interventions[10,12,24]. A striking number of these trials investigating carbohydrate restriction are confounded by hypocaloric experimental diets. It remains unclear if weight loss achieved in studies of this nature is related to the macronutrient distribution of experimental diets or to purposeful caloric restriction. Notably, fat mass loss was not observed in B6 females that received the dietary intervention, but instead, there was less fat mass gained in B6 females that received the dietary intervention than B6 females that remained on the American diet for the duration of the feeding trial. This suggests that the dietary intervention will not result in weight or fat mass reduction until another component of the environment is altered.

A tight, inverse relationship was observed between heat expenditure and fat mass gained during the feeding trial. Of note, B6 females that received the dietary intervention generated significantly more heat by the end of the feeding trial than B6 females that remained on the American diet for the duration of the feeding trial. To understand the

relationship between heat expenditure and the response to this dietary intervention, future work will include a systematic evaluation of candidate genes identified in an intercross population (B6 x FVB F2s) exposed to American and ketogenic diets.

In order to highlight implications for conclusions drawn when data analysis is conducted at the population level, a diet specific, diet and sex specific, and diet and strain specific analysis of the data was presented. These represent efforts of many trials that analyze the response to diet without considering males and females as separate entities, those that analyze males and females separately without the use of any genetic information, and those that utilize genetic information but lack power within genotype classes or fail to incorporate sex as a biological variable[25,26]. Occasionally, genetic markers are identified under baseline conditions and applied to individuals consuming specific diets where they appear to have no effect on diet response[27]. In the current experiment, it was observed that there were no significant differences between diet groups until the data was stratified by strain, sex, and diet. Once the data is stratified, groups of animals where the dietary intervention would be most effective was discerned. This represents important information that is lost if conclusions are drawn from the population level.

Conclusions

By parsing the data set by strain, sex, and diet, the ketogenic dietary intervention was determined most effective for B6 females. Interestingly, there appears to be a strong relationship between the amount of fat mass gained prior to the introduction of the dietary intervention and the amount of fat mass gained at the end of the feeding trial. This

observation is particularly important for B6 males as the amount of fat mass gained prior to the introduction of the dietary intervention appears to limit weight loss in B6 males once the intervention is introduced. This suggests that there are mechanisms contributing to obesity that are not attenuated by simple dietary intervention, even within genetic backgrounds that should respond to carbohydrate restriction. It is imperative to include sex as a biological variable, especially in conjunction with genetic data, to determine response to dietary interventions and generate datasets that will develop Precision Nutrition and individualized dietary recommendations.

Acknowledgements

We thank Dr. Andrew Hillhouse and the Texas A&M Institute for Genome Sciences and Society (TIGSS) for assistance in the TIGSS Molecular Genomics Workspace and TIGSS Rodent Preclinical Phenotyping Core; and all members of the laboratory for their helpful insights. This work was supported by National Institutes of Health (NIH) grants RM1HG008529 and P30ES029067.

Competing Interests

The authors have no competing interests to declare.

References

1. Gerdts E, Regitz-Zagrosek V. Sex differences in cardiometabolic disorders. *Nature Medicine*. 2019;25(11). doi:10.1038/s41591-019-0643-8
2. Deluis DA, Sagrado MG, Aller R, Izaola O, Conde R. Effects of C358A missense polymorphism of the degrading enzyme fatty acid amide hydrolase on weight loss, adipocytokines, and insulin resistance after 2 hypocaloric diets. *Metabolism: Clinical and Experimental*. 2010;59(9). doi:10.1016/j.metabol.2009.12.029
3. Grau K, Hansen T, Holst C, et al. Macronutrient-specific effect of FTO rs9939609 in response to a 10-week randomized hypo-energetic diet among obese Europeans. *International Journal of Obesity*. 2009;33(11). doi:10.1038/ijo.2009.159
4. de Luis DA, Izaola O, Primo D, Aller R. Polymorphism rs16147 of the neuropeptide y gene modifies the response of cardiovascular risk biomarkers and adipokines to two hypocaloric diets. *Journal of Nutrigenetics and Nutrigenomics*. 2017;10(1-2). doi:10.1159/000478528
5. Qi Q, Bray GA, Smith SR, Hu FB, Sacks FM, Qi L. Insulin receptor substrate 1 gene variation modifies insulin resistance response to weight-loss diets in a 2-year randomized trial the preventing overweight using novel dietary strategies (POUNDS LOST) Trial. *Circulation*. 2011;124(5). doi:10.1161/CIRCULATIONAHA.111.025767
6. Zhang X, Qi Q, Zhang C, Hu FB, Sacks FM, Qi L. FTO genotype and 2-year change in body composition and fat distribution in response to weight-loss diets: The POUNDS LOST trial. *Diabetes*. 2012;61(11). doi:10.2337/db11-1799

7. Heianza Y, Ma W, Huang T, et al. Macronutrient intake-associated FGF21 genotype modifies effects of weight-loss diets on 2-year changes of central adiposity and body composition: The POUNDS lost trial. *Diabetes Care*. 2016;39(11). doi:10.2337/dc16-1111

8. Gardner CD, Hauser M, Gobbo LD, et al. Neither insulin secretion nor genotype pattern modify 12-month weight loss effects of healthy low-fat vs. healthy low-carbohydrate diets among adults with obesity. *Circulation Conference: American Heart Association's Epidemiology and Prevention/Lifestyle and Cardiometabolic Health*. 2017;135(Supplement 1).

9. Bradley U, Spence M, Courtney CH, et al. Low-fat versus low-carbohydrate weight reduction diets - Effects on weight loss, insulin resistance, and cardiovascular risk: A randomized control trial. *Diabetes*. 2009;58(12). doi:10.2337/db09-0098

10. Lasker DAW, Evans EM, Layman DK. Moderate carbohydrate, moderate protein weight loss diet reduces cardiovascular disease risk compared to high carbohydrate, low protein diet in obese adults: A randomized clinical trial. *Nutrition and Metabolism*. 2008;5(1). doi:10.1186/1743-7075-5-30

11. Wycherley TP, Brinkworth GD, Keogh JB, Noakes M, Buckley JD, Clifton PM. Long-term effects of weight loss with a very low carbohydrate and low fat diet on vascular function in overweight and obese patients: Original Article. *Journal of Internal Medicine*. 2010;267(5). doi:10.1111/j.1365-2796.2009.02174.x

12. Tay J, Brinkworth GD, Noakes M, Keogh J, Clifton PM. Metabolic Effects of Weight Loss on a Very-Low-Carbohydrate Diet Compared With an Isocaloric High-

Carbohydrate Diet in Abdominally Obese Subjects. *Journal of the American College of Cardiology*. 2008;51(1). doi:10.1016/j.jacc.2007.08.050

13. Henke C, Haufe S, Ziehl D, et al. Low-fat hypocaloric diet reduces neprilysin in overweight and obese human subjects. *ESC Heart Failure*. 2021;8(2). doi:10.1002/ehf2.13220
14. Barrington WT, Wulfridge P, Wells AE, et al. Improving metabolic health through precision dietetics in mice. *Genetics*. 2018;208(1):399-417. doi:10.1534/genetics.117.300536
15. Wells A, Barrington WT, Dearth S, et al. Tissue Level Diet and Sex-by-Diet Interactions Reveal Unique Metabolite and Clustering Profiles Using Untargeted Liquid Chromatography-Mass Spectrometry on Adipose, Skeletal Muscle, and Liver Tissue in C57BL6/J Mice. *Journal of Proteome Research*. Published online 2018. doi:10.1021/acs.jproteome.7b00750
16. Cuomo D, Porreca I, Ceccarelli M, et al. Transcriptional landscape of mouse-aged ovaries reveals a unique set of non-coding RNAs associated with physiological and environmental ovarian dysfunctions. *Cell Death Discovery*. Published online 2018. doi:10.1038/s41420-018-0121-y
17. Salvador AC, Arends D, Barrington WT, Elsaadi AM, Brockmann GA, Threadgill DW. Sex-specific genetic architecture in response to American and ketogenic diets. *International Journal of Obesity*. 2021;45(6). doi:10.1038/s41366-021-00785-7
18. Fritzen AM, Lundsgaard AM, Kiens B. Tuning fatty acid oxidation in skeletal muscle with dietary fat and exercise. *Nature Reviews Endocrinology*. 2020;16(12). doi:10.1038/s41574-020-0405-1

19. Turner N, Bruce CR, Beale SM, et al. Excess Lipid Availability Increases Mitochondrial Fatty Acid Oxidative Capacity in Muscle. *Diabetes*. 2007;56(8). doi:10.2337/db07-0093

20. Rasmussen BB, Wolfe RR. Regulation of fatty acid oxidation in skeletal muscle. *Annual Review of Nutrition*. 1999;19. doi:10.1146/annurev.nutr.19.1.463

21. Roberts MN, Wallace MA, Tomilov AA, et al. A Ketogenic Diet Extends Longevity and Healthspan in Adult Mice. *Cell Metabolism*. 2017;26(3). doi:10.1016/j.cmet.2017.08.005

22. Casimiro I, Stull ND, Tersey SA, Mirmira RG. Phenotypic sexual dimorphism in response to dietary fat manipulation in C57BL/6J mice. *Journal of Diabetes and its Complications*. 2021;35(2). doi:10.1016/j.jdiacomp.2020.107795

23. Kennedy AJ, Ellacott KLJ, King VL, Hasty AH. Mouse models of the metabolic syndrome. *Disease Models & Mechanisms*. 2010;3:156-166. doi:10.1242/dmm.003467

24. Hernandez TL, Sutherland JP, Wolfe P, et al. Lack of suppression of circulating free fatty acids and hypercholesterolemia during weight loss on a high-fat, low-carbohydrate diet. *American Journal of Clinical Nutrition*. 2010;91(3). doi:10.3945/ajcn.2009.27909

25. Gardner CD, Trepanowski JF, Gobbo LCD, et al. Effect of low-fat VS low-carbohydrate diet on 12-month weight loss in overweight adults and the association with genotype pattern or insulin secretion the DIETFITS randomized clinical trial. *JAMA - Journal of the American Medical Association*. 2018;319(7). doi:10.1001/jama.2018.0245

26. Coletta AM, Sanchez B, O'Connor A, et al. Alignment of diet prescription to genotype does not promote greater weight loss success in women with obesity participating in an exercise and weight loss program. *Obesity Science and Practice*. 2018;4(6). doi:10.1002/osp4.305

27. Baratali L, Mean M, Marques-Vidal P. Impact of dietary and obesity genetic risk scores on weight gain. *The American Journal of Clinical Nutrition*. Published online 2021. doi:10.1093/ajcn/nqab069

Response to Dietary Intervention

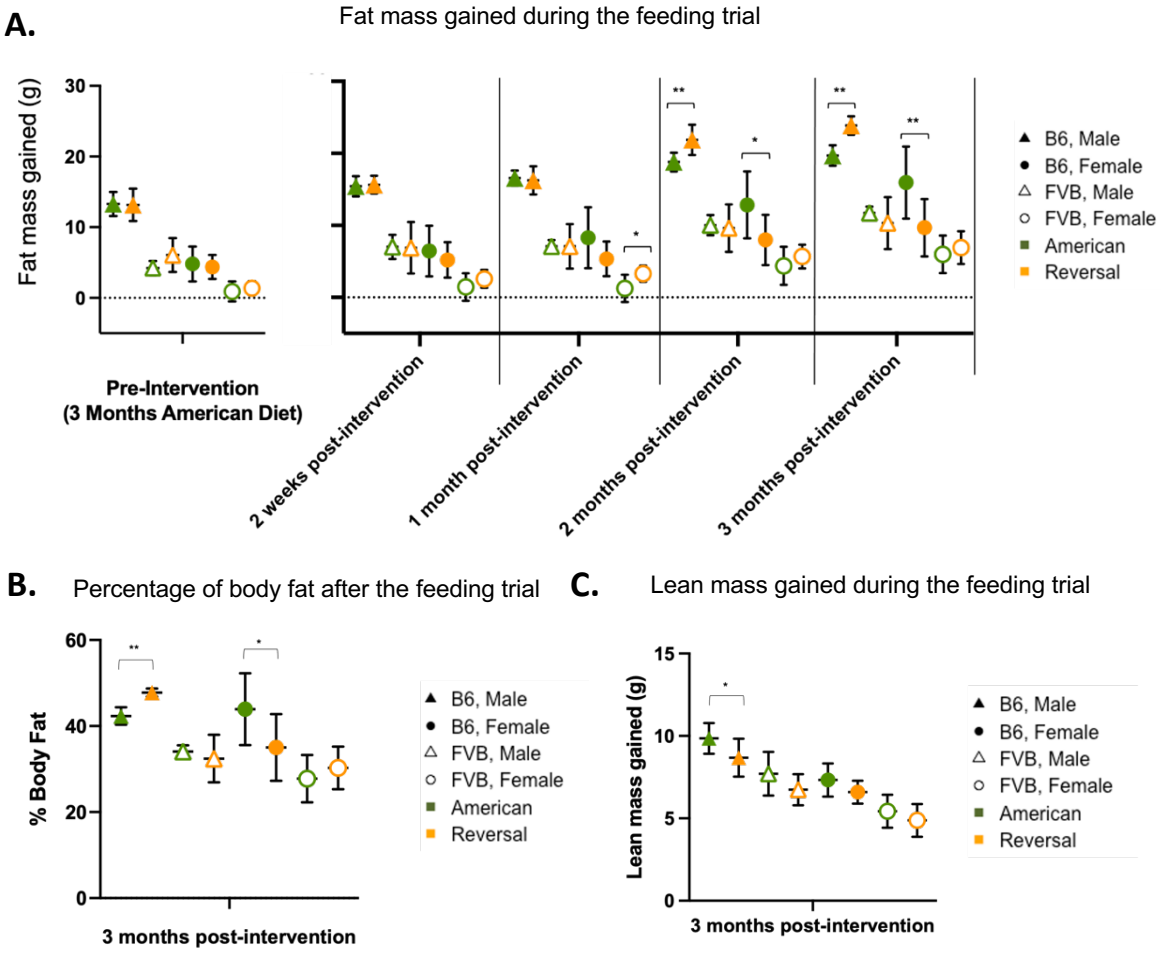

Figure 7. Response to Dietary Intervention. A. Fat mass gained during the feeding trial. No significant differences in baseline fat mass gained during the feeding trial after 3 months on the American diet, prior to intervention. On average, B6 males that received the dietary intervention gained 6.89g more fat by the end of the feeding trial than B6 males that remained on the American diet for the duration of the trial. On average, B6 females that received the dietary intervention gained 3.04g less fat by the end of the feeding trial than B6 females that remained on the American diet for the duration of the feeding trial. No differential response to the American diet or dietary intervention by the end of the feeding trial (3 months post-intervention) for FVB males or females. **B.** Percentage of body fat after the feeding trial. On average, B6 males that received the dietary intervention had a higher percentage of body fat by the end of the feeding trial than B6 males that remained on the American diet for the duration of the feeding trial. On average, B6 females that received the dietary intervention had a lower percentage of body fat by the end of the feeding trial than B6 females that remained on the American diet for the duration of the feeding trial. No differential response to the American diet or dietary intervention by the end of the feeding trial (3 months post-intervention) for FVB males or females. **C.** Lean mass gained during the feeding trial. On average, B6 males that received the dietary intervention gained 1.17 g less lean mass by the end of the feeding trial than B6 males that remained on the American diet for the duration of the feeding trial. No differential response to the American diet or dietary intervention by the en do the feeding trial for B6 females or FVB males and females.

Associations with fat mass gained during the feeding trial

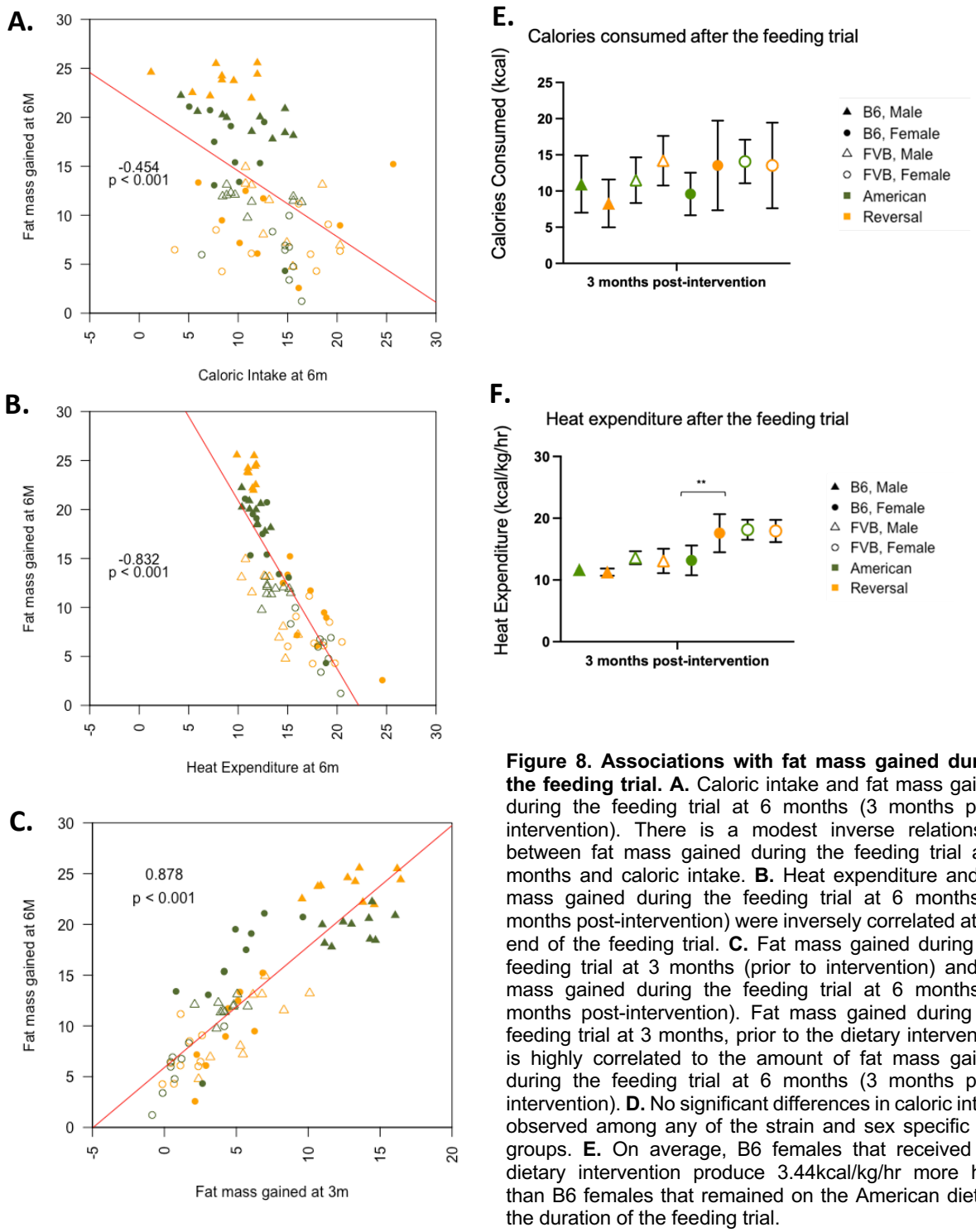

Figure 8. Associations with fat mass gained during the feeding trial. A. Caloric intake and fat mass gained during the feeding trial at 6 months (3 months post-intervention). There is a modest inverse relationship between fat mass gained during the feeding trial at 6 months and caloric intake. **B.** Heat expenditure and fat mass gained during the feeding trial at 6 months (3 months post-intervention) were inversely correlated at the end of the feeding trial. **C.** Fat mass gained during the feeding trial at 3 months (prior to intervention) and fat mass gained during the feeding trial at 6 months (3 months post-intervention). Fat mass gained during the feeding trial at 3 months, prior to the dietary intervention is highly correlated to the amount of fat mass gained during the feeding trial at 6 months (3 months post-intervention). **D.** No significant differences in caloric intake observed among any of the strain and sex specific diet groups. **E.** On average, B6 females that received the dietary intervention produce 3.44kcal/kg/hr more heat than B6 females that remained on the American diet for the duration of the feeding trial.

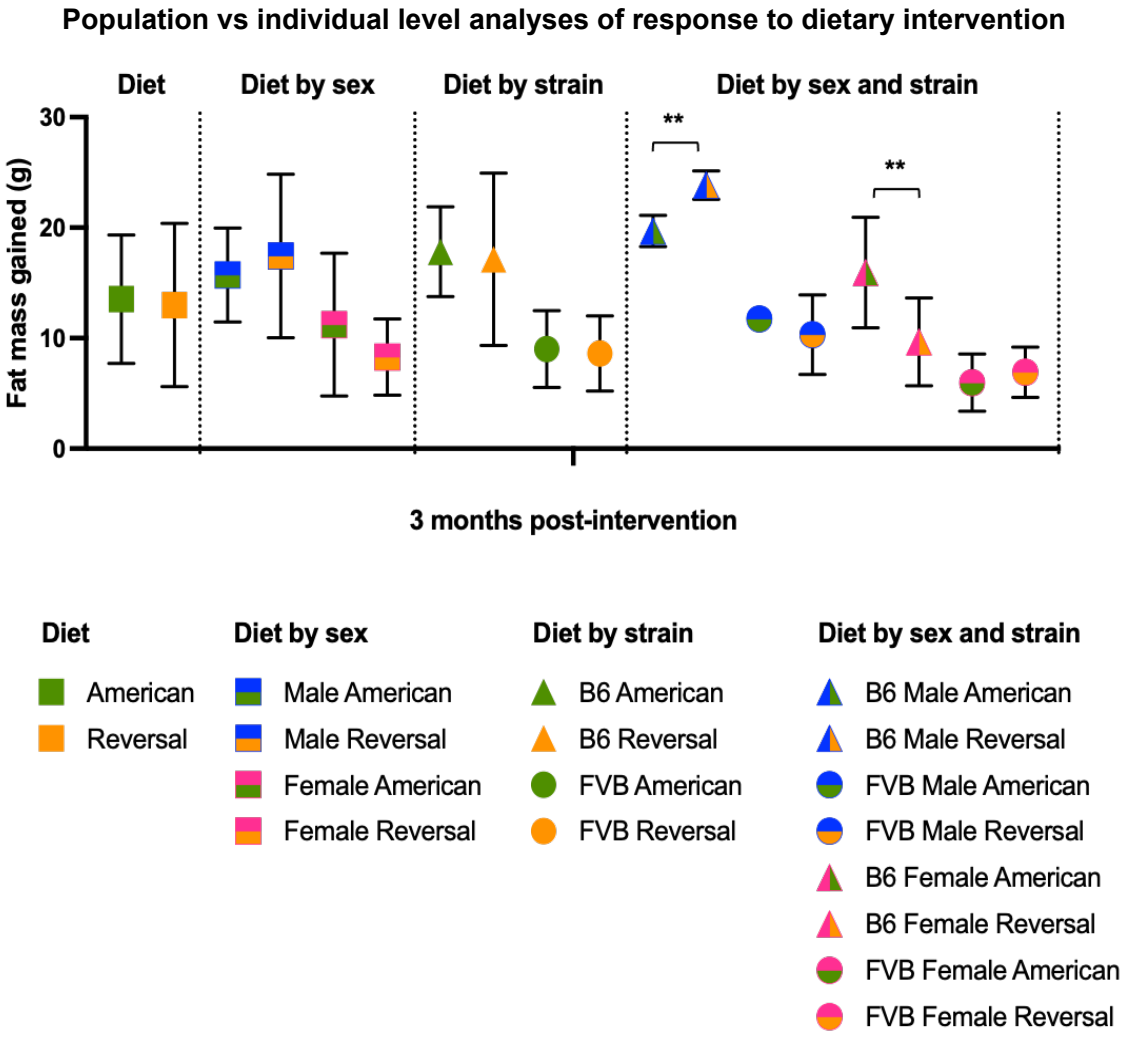

Figure 9. Population vs individual level analyses of response to dietary intervention. If the analysis were conducted at the population level, we would observe no significant difference between the diet treatments, no significant difference in males exposed to either diet or females exposed to either diet, and no differences observed between the two treatments with respect to genetic background. Not until the data is stratified by diet, strain, and sex do we observe which individuals the dietary intervention would be the most effective for.

CHAPTER V: CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions and Future Directions

Precision nutrition is poised to revolutionize the practice of dietetics by taking individual characteristics such as genetics, sex, disease status, dietary intake, and lifestyle into consideration for the treatment of nutrition-related disease. Despite the decades of efforts to identify gene-diet interactions, no robust molecular signatures of responders and non-responders to particular diets have been established for use in routine clinical settings. Studies must incorporate as much information about genetics, sex, disease status, dietary intake, and lifestyle as is available to advance precision nutrition and lay the groundwork for developing individualized dietary guidelines.

The mouse provides a unique opportunity to model the development and application of precision nutrition. Mice are commonly used to investigate metabolic disease, especially metabolic syndrome. Spontaneous mutations occurring in obese mouse strains have been used for decades to model the effects of single genes[1]. The availability of inbred mouse strains makes the mouse an excellent resource for multi-strain genetic studies. We have observed mouse strain-specific differences in response to feeding different human-relevant diets, and documented a differential response to American (35% of energy from fat, 50% from carbohydrate) and ketogenic (84% of energy from fat, 0% from carbohydrate) diets in C57BL/6J (B6) mice. Further sex and strain specific analyses revealed that the differential response to carbohydrate restriction was specific to B6

males, but not B6 females, FVB/NJ (FVB) males, or FVB females during lifetime dietary exposures.

This sex specific pattern persisted in F1 and F2 populations derived from B6 and FVB mice where again, we observed a differential response to the two diets in males rather than females. Of the three genetic loci we identified for fat mass gained during the feeding trial, two of them affect fat gain directly (*Fmgq2* and *Fmgq3*), while *Fmgq1* appears to influence fat gain directly and possibly via an intermediate change in serum cholesterol. The genetic component regulating the amount of fat mass gained during the feeding trial that we were able to detect in our model contributed more to the overall response in males than in females. We also noted that the response to the ketogenic diet becomes more severe in the F1 and F2 populations. This prompted us to consider interactions between loci identified for fat mass gain that might explain why the mixture of B6 and FVB genomes would result in more severe responses to the ketogenic diet.

We identified a functional interaction between *Fmgq1* and *Fmgq2*; male F2s that are homozygous B6 at *Fmgq2* that carry at least one FVB allele at *Fmgq1* gained the most fat mass. *Ppp2r5a* and *Srd5a3* are strong candidate genes at *Fmgq1* and *Fmgq2*, respectively, and represent two candidates that have been previously associated male hormone secretion, making them primary candidate genes of interest for future work. It remains unclear if the loci we identified explain the differential response to carbohydrate restriction we observed in the parent screen or, if instead, they characterize the increased fat mass gain that we observe after the parental genomes are combined in hybrid and F2

populations. Further investigation of the candidate genes will help elucidate their roles in response to carbohydrate restriction in the parent, hybrid, and F2 populations.

The gut microbiome is an emerging component of precision nutrition. Abundances of gut microbiota are influenced by sex, diet, genetics, and interactions among these factors. At the phyla level, we observed strong effects of diet that were irrespective to genetic background especially with regard to the ratio of Firmicutes to Bacteroidetes. Some sources suggest that a higher Firmicutes to Bacteroidetes ratio is positively correlated with obesity and a decrease in this ratio is associated with inflammatory bowel disease[2]. This would imply that population level dietary guidelines would be an effective tool for modulating the gut microbiota of individuals irrespective of the host's sex or genetics. Metabolic effects stemming from modulation of the gut microbiota do not always occur at the phyla level, though. More often we are interested in identifying changes at the genus and species level and how these microbiota are related to metabolic traits. At the genus and species level, we observed that there are particular operational taxonomic units (OTU) for which sex, strain, diet, and interactions among these factors differentially modulate gut microbiota.

We have distinguished between OTU that were primarily regulated by genetics, those that were primarily regulated by interactions between diet and genetics, and those that were primarily regulated by interactions between sex and genetics. This distinction has important implications for developing individualized dietary guidelines for modulating the gut microbiota. For example, we observed that the abundance of the genus and an

unidentified species of *Coriobacteriaceae* and the genus *Lachnospiraceae* (FCS020 Group) were least influenced by diet but have been positively associated with unfavorable metabolic consequences for lipid profiles and cardiovascular disease risk. We would expect that developing dietary guidelines to reduce the abundance of these particular OTU would be ineffective. It may still be clinically relevant to understand which individuals are genetically predisposed to increased or decreased abundances of the organisms. However, the immediate utility of this information is unclear since we would not expect a dietary intervention to have a profound effect on their abundances.

The most clinically relevant distinction we can make is for those OTU that are driven by an interaction between diet and genetics. We identified an interaction between genotype and diet for the genus *Alistipes*, an unidentified species of *Alistipes*, the genus *Rikenellaceae* (RC9 gut group), the genus *Streptococcus*, the genus *Erysipelatoclostridium*, and an unidentified species of *Erysipelotrichaceae*. These are the organisms that would be sensitive to dietary intervention that we could begin to develop individualized dietary recommendations around for increasing or decreasing their abundance in the subset of the population harboring the sensitive genotypes.

However, the observations of microbial abundances are incredibly context dependent and are easily illustrated by comparing our observation of diet and sex specific loci regulating the abundance of *Alistipes* to literature which detected sex specific loci but no diet dependent loci for this organism. Our human comparable American diet contains a diverse set of ingredients contributing to the total protein content of the diet and *Alistipes*

is particularly sensitive to animal sources of dietary protein. Thus, we are able to detect a diet and genotype interaction in our model. In literature where others have been unable to identify diet specificity for loci regulating the abundance of *Alistipes*, the classical laboratory chow diets that were used were composed with a single source of protein. Thus, they did not detect an interaction between the diets and individual genotypes[3]. This should be considered when selecting experimental diets for analyses of microbiota abundance.

QTL for an unidentified genera of *Lachnospiraceae* (*Mtq8*), the genera and an unidentified species of *Bilophila* (*Mtq13*), fat mass gained during the feeding trial (Fmgq1), and serum HDL cholesterol concentration (*Hdlq1*) overlap on the distal part of Chr1 and QTL for the genera *Roseburia* (*Mtq7*) and the genera and an unidentified species of *Bilophila* (*Mtq14*) overlap on Chr9. These organisms have been associated with lipid profiles, non-alcoholic fatty liver disease, and type 2 diabetes however, we were unable to identify causal relationships between these metabolic and microbial traits. This may be the result of latent variables that are not directly observable in the model but might be inferred from other variables. Microbiota utilize nutrients passing through the gastrointestinal tract and metabolism of the nutrients produces metabolites. Microbial-derived metabolites are known to impact metabolic health and may represent latent variables linking the genomic regions underlying *Fmgq1*, *Hdlq1*, *Mtq8*, and *Mtq13* and each of there associated traits as well as *Mtq7* and *Mtq14* and their metabolic traits. We could consider microbial-derived metabolites in our linkage analysis to elucidate these relationships.

We performed sex, strain, and diet specific analysis to determine whether the ketogenic diet is an effective dietary intervention for B6 and FVB males and females that have been exposed to the American diet. As has been discussed, all preliminary evidence suggested that B6 males would respond best to carbohydrate restriction. Interestingly, we found that the dietary intervention was most effective at reducing the amount of fat mass gained during the feeding trial in B6 females that received the dietary intervention rather than B6 males. Conversely, we observed that B6 males that received the dietary intervention began to gain fat mass at a greater rate than those that remained on an American diet for the duration of the feeding trial. B6 males that received the dietary intervention gained less lean mass than B6 males that remained on the American diet for the duration of the feeding trial. Skeletal muscle plays a critical role in fatty acid oxidation and a reduction in lean mass gained by B6 males that received the intervention may reduce their capacity for fatty acid oxidation. However, B6 females that received the dietary intervention gained similar amounts of lean mass as B6 females that remained on the American diet for the duration of the feeding trial which suggests that the efficacy of the dietary intervention is not solely dependent upon increased capacity of fatty acid oxidation in B6 females.

Historically, mouse models have used lifetime dietary exposures to assess response to experimental diets. Most human trials of diet response enroll participants that have been diagnosed with one or more metabolic perturbations and expose them to experimental diets in order to improve the metabolic perturbation. We noted a strong relationship observed between the amount of fat mass gained prior to the introduction of the dietary intervention and the amount of fat mass gained at the end of the feeding trial. B6 males

gain more fat mass than any other strain and sex combination while on the American diet for the three months leading up to the dietary intervention. This suggests that there are mechanisms contributing to obesity that are not attenuated by simple dietary intervention in the B6 males, even within a genetic background predicted to respond to carbohydrate restriction. This appears to limit the ability of the dietary intervention to reduce the amount of fat mass gained in B6 males that received the dietary intervention. Although the genetic signal we detected in earlier studies in our intercross population was stronger in males, it is possible that the phenotypes are more plastic in females.

A 2019 medical hypothesis touches on mechanisms contributing to obesity that are resistant to change and highly heritable[4]. The author describes non-adaptive manifestations of obesity and suggests that insulin resistance occurs to protect against reactive hypoglycemia induced by poor diet. This protection is critical for the brain and central nervous system that rely on glucose for fuel. He cites several examples of conditions that result in acute or chronic insulin resistance to mitigate conditions associated with high risk of hypoglycemia including diabetes, pregnancy, infection, burns, and shock. If the plasticity of obese phenotypes is grounded in adaptive and non-adaptive responses to diets, it becomes more intuitive that there are mechanisms contributing to obesity that are not reversible by simple dietary intervention. A hypoglycemic episode is much more dangerous in the short-term than a hyperglycemic episode. The pathophysiology we associate with many chronic diseases may in fact reflect purposeful adaptive and non-adaptive responses to environmental stimuli to mitigate higher acuity risks like reactive hypoglycemia. According to this medical hypothesis, being at a more

severe obese state prior to the introduction of the dietary intervention would make the non-adaptive manifestations of obesity less plastic in B6 males, just like we have observed.

Recently, interest has piqued in transcriptome profiling of extracellular vesicles (EV). EV are membranous pouches released from cells that contain lipids, proteins, and nucleic acids, including miRNAs among a diverse set of RNAs. Transfer of these components is a type of intercellular communication and influences cell and tissue physiology[5]. Collection and transcriptome profiling of EV has led to discoveries suggesting that miRNA secretions from liver-derived EV enhance adipocyte lipid accumulation and correlate with body mass index in humans[6]. In our model, we suspect that accumulated lipid prior to the introduction of the dietary intervention may continue to secrete EV that prevent the reversal of the metabolic perturbation. To investigate this, we could investigate EV secretions resulting from severe lipid accumulation prior to the introduction of the dietary intervention and their role in the reduced plasticity of the B6 male obese phenotype by transcriptome profiling serum collected before and after the introduction of the dietary intervention. With sufficient preliminary evidence for the role of EV in our model, we could then test the efficacy of surgical removal of adipose tissue at the time of dietary intervention to reverse the obese phenotype with the introduction of the intervention to evaluate if this underlies the lack of response to the dietary intervention in B6 males.

In addition to EV collection and profiling, we could also consider the temporal component to reaching this perceived “point of no return” during lipid accumulation as well as synergy

among interventions like diet and exercise and their efficacy to reduce the amount of fat mass and in turn, modify the EV secretions regulating cell and tissue physiology. We could use our model to test a 2-week exposure to the American diet prior to the dietary intervention and a 1-month exposure to the American diet prior to the dietary intervention to determine if by intervening at an earlier timepoint can increase the plasticity of the obese phenotype of B6 males after exposure to an American diet. Similarly, we could repeat our study of the 3-month exposure to the American diet prior to the introduction of the dietary intervention and add an exercise component to the intervention phase to determine if we observe increased plasticity of the obese phenotype in B6 males. The expectation is that response to the intervention requires a genetic predisposition and either limited lipid accumulation prior to the introduction of the intervention or an effective synergy among interventions to effectively reduce the accumulated lipid during the intervention phase.

To highlight the consequences of data analysis at the population level, we conducted our analyses of the response to the ketogenic dietary intervention in a diet specific, diet and sex specific, and diet and strain specific way. We observed that there were no significant differences between diet groups until the data was stratified by strain, sex, and diet and this has important implications for trials that report findings on response to diet without considering males and females as separate entities, those that may consider sex differences but fail to utilize genetic information, and those that utilize genetic information but lack power within genotype classes or fail to incorporate both sex and genetics as biological variables.

This work reflects an early attempt at modeling precision nutrition in mice. We have refined our assessment criteria by incorporating as much information about genetic background and sex as was available to us (Figure 10A-E). Systematic evaluation of positional candidate genes regulating fat mass gained during the feeding trial, as well as serum cholesterol concentration and microbial abundances after the feeding trial could reveal mechanisms of obesity unique to subgroups of the population. Once these mechanisms are revealed, molecular signatures of them could serve as biomarkers and easily accessible biomarkers could be used to further refine assessment criteria and identify individuals within each subgroup during Nutrition Assessment. After a nutrition diagnosis is made, these are the types of individuals that we could develop unique dietary guidelines for and select individualized nutrition interventions for. As we establish a better understanding of the timing of dietary intervention and synergy among interventions, this approach to precision nutrition could dramatically reduce the rate of obesity and improve health outcomes.

References

1. Kennedy AJ, Ellacott KLJ, King VL, Hasty AH. Mouse models of the metabolic syndrome. *Disease Models & Mechanisms*. 2010;3:156-166. doi:10.1242/dmm.003467
2. Magne F, Gotteland M, Gauthier L, et al. The firmicutes/bacteroidetes ratio: A relevant marker of gut dysbiosis in obese patients? *Nutrients*. 2020;12(5). doi:10.3390/nu12051474
3. Leamy LJ, Kelly SA, Nietfeldt J, et al. *Host Genetics and Diet, but Not Immunoglobulin A Expression, Converge to Shape Compositional Features of the Gut Microbiome in an Advanced Intercross Population of Mice*.; 2014. http://genomebiology.com/2014/15/12/552
4. Bradley P. Refined carbohydrates, phenotypic plasticity and the obesity epidemic. *Medical Hypotheses*. 2019;131. doi:10.1016/j.mehy.2019.109317
5. Amorim MG, Valieris R, Drummond RD, et al. A total transcriptome profiling method for plasma-derived extracellular vesicles: Applications for liquid biopsies. *Scientific Reports*. 2017;7(1). doi:10.1038/s41598-017-14264-5
6. Zhao Y, Zhao MF, Jiang S, et al. Liver governs adipose remodelling via extracellular vesicles in response to lipid overload. *Nature Communications*. 2020;11(1). doi:10.1038/s41467-020-14450-6

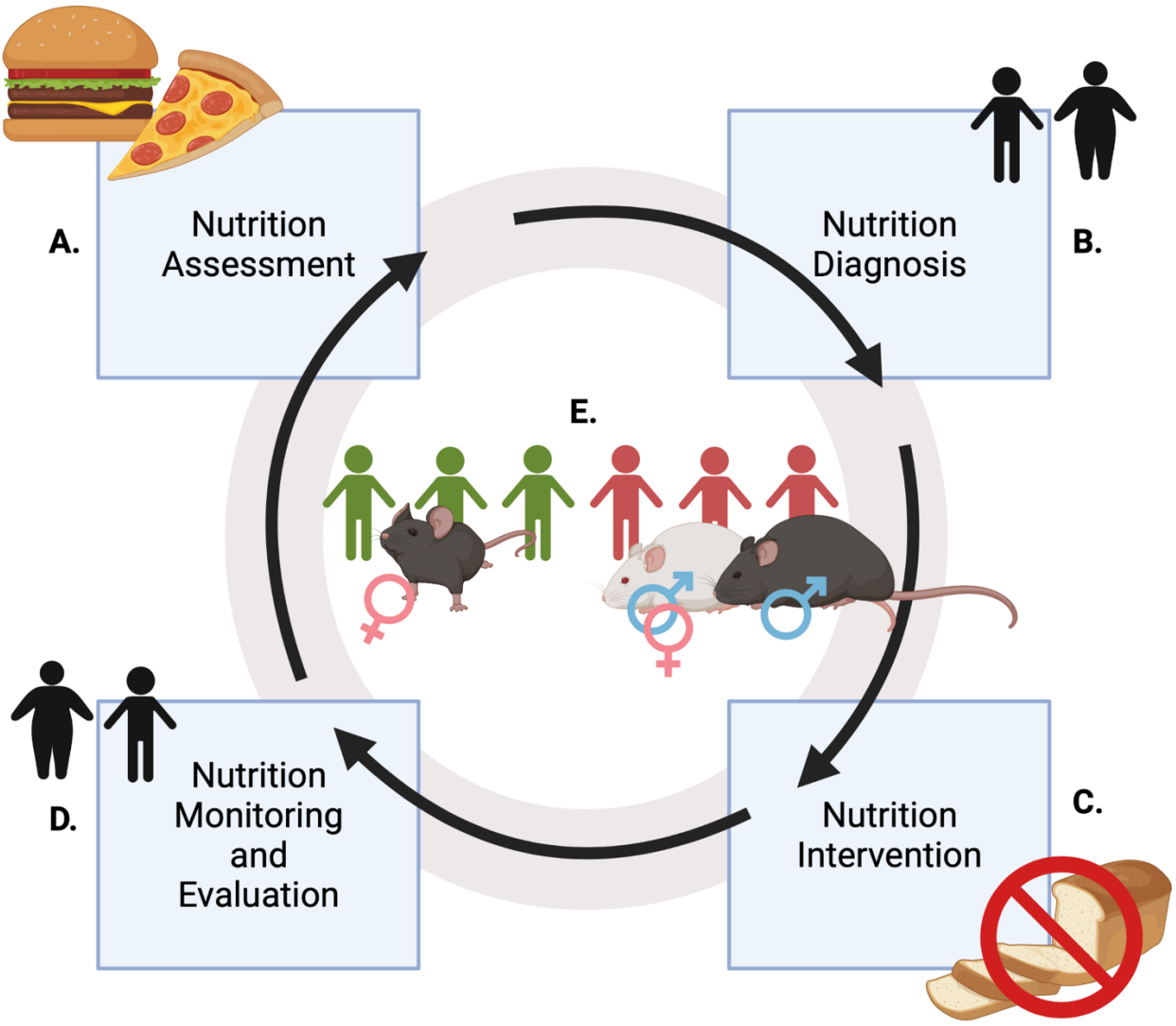

Figure 10. Implementing precision nutrition. A. Sex, strain, and diet specific analyses during Nutrition Assessment revealed B6 and FVB females gain fat during exposure to an American diet. B6 males gain more fat than any other sex and strain combination during the initial exposure to the American diet. **B.** Based on the data collected in the Nutrition Assessment we documented overweight and obesity nutrition diagnoses for each strain and sex combination after exposure to the American diet. **C.** To determine the efficacy a carbohydrate restricted nutrition intervention in multiple genetic backgrounds we reversed half of the animals to a ketogenic diet. **D.** Nutrition monitoring and evaluation of changes to body composition after the introduction of the dietary intervention revealed which sex and strain combinations responded to the nutrition intervention. **E.** The ketogenic dietary intervention was most effective at reducing the amount of fat mass gained during the feeding trial in B6 females that received the intervention rather than B6 males. As expected, no differential response to the two diets was detected in FVB males or females.

APPENDIX

Table A1. Diet compositions

	D12052705C (American)	D12052706 (Ketogenic)
Ingredient	(g)	(g)
Casein	38.5	160
Soy Protein Isolate, Supro 661	0	0
Fish Protein Isolate	8.5	0
Egg White	55	0
Beef, Cooked	76.9	0
L-Cystine	3	3
Corn Starch	**30**	0
Wheat Starch	195	0
Potato Starch	**30**	0
Sucrose	205	0
Fructose	22	0
Cellulose, BW200	18.2	37.5
Inulin	**6**	12.5
Corn Oil	34.4	8
Menhaden Oil (299 ppm tBHQ)	1	8
Butter, Anhydrous	54.1	161
Lard	0	161
Flaxseed Oil	1	0
Olive Oil	27.5	0

*Reprinted with permission from "Sex-specific genetic architecture in response to American and ketogenic diets" by A.C. Salvador, et. al., 2021. *International Journal of Obesity*, Volume 45, 1284-1297. Copyright 2021 The Author(s), under exclusive license to Springer Nature Limited part of Springer Nature 2021

t-BHQ	0.0047	0.0033
Mineral Mix S10026	10	10
Dicalcium Phosphate	13	13
Calcium Carbonate	5.5	5.5
Potassium Citrate, 1 H2O	16.5	16.5
Vitamin Mix V10001	10	10
Biotin (1%)	0.1	0
Choline Bitartrate	2	2
Cholesterol	1.5	1.5
FD&C Red Dye #5	0	0.025
FD&C Blue Dye #1	0.05	0.025
Total	864.7547	609.5533
(g)		
Protein	140.7	142.2
Carbohydrate	455.3	3.1
Fat	139.7	339.9
Cholesterol	1.81	2.09
Fiber	22.7	46.9
% of Total Weight (g)		
Protein	16.3	23.3
Carbohydrate	52.6	0.5
Fat	16.2	55.8
Cholesterol	0.21	0.342
Fiber	2.6	7.7
Energy (kcal)		
Protein	563	569
Carbohydrate	1821	13
Fat	1257	3059
Total	3641	3641

% of Total Energy (kcal)		
Protein	15	16
Carbohydrate	50	0
Fat	35	84
% of total weight (g) Contribution to Protein (g)		
Casein	23.8	97.9
Fish Protein Isolate	5.7	0
Egg White	31.7	0
Beef, Cooked	36.6	0
L-Cystine	2.1	2.1
% of total weight (g) Contribution to Carbohydrate		
Corn Starch	5.9	0
Wheat Starch	38.4	0
Potato Starch	5.3	0
Sucrose	45.6	0
Fructose	4.9	0
Lipid Composition		
C2, Acetic	0	0
C4, Butyric	1.7	5.2
C6, Caproic	1	3.1
C8, Caprylic	0.6	1.8
C10, Capric	1.4	4.1
C12, Lauric	1.5	4.7
C14, Myristic	6.1	18.5
C14:1, Myristoleic	1	2.4
C15:0	0.1	0.2
C16, Palmitic	25.9	75.3
C16:1, Palmitoleic	2.4	6.7
C16:2	0	0.1
C16:3	0	0.1

C16:4	0	0.1
C17:0	0.2	0.6
C17:1, n-9	0.2	0
C18, Stearic	10.2	36.9
C18:1, Oleic, n-9	50.5	96.1
C18:1, Vaccenic, n-7	0.9	0
C18:2, Linoleic, n-6	26.5	47.9
C18:3, gamma-Linolenic, n-6	0	0
C18:3, alpha-Linolenic, n-3	2	4.6
C18:4, Stearidonic, n-3	0	0.2
C20, Arachidic	0.7	1.8
C20:1, n-9	0.2	1.1
C20:2, Eicosadienoic, n-6	0	1.3
C20:3, n-6	0	0.2
C20:3, n-3	0	0
C20:4, Arachidonic, n-6	0.1	0.6
C20:4, n-3	0	0
C20:5, Eicosapentaenoic, n-3	0.1	1.1
C21:0	0	0
C21:5, n-3	0	0.1
C22, Behenic	0	0
C22:1, Erucic	0	0
C22:4, Clupanodonic, n-6	0	0
C22:5, n-3	0	0.4
C22:5, n-6	0	0
C22:6, Docosahexaenoic, n-3	0.1	0.8
C24, Lignoceric	0	0
C24:1	0	0
Lipid Profile		
Saturated (g)	49.1	151.3
Monounsaturated (g)	54.1	106.4
Polyunsaturated (g)	29	57.3

Saturated (%)	36.7	47.9
Monounsaturated (%)	40.5	33.6
Polyunsaturated (%)	21.7	18.1
Total Omega-6 (g)	26.7	50.1
Total Omega-6 (%)	20	15.9
Total Omega-3 (g)	2.4	7.2
Total Omega-3 (%)	1.8	2.3
n6 : n3 ratio	11.3	6.9

Table A2. Comprehensive list of all KEGG queries used for candidate gene associations.

Inclusion Criteria	KEGG id	KEGG pathway title
Glucose and insulin related pathways	mmu00010	Glycolysis/Gluconeogenesis
	mmu04910	Insulin signaling pathway
	mmu04931	Insulin resistance
Diabetes Mellitus related pathways	mmu04940	Type I diabetes mellitus
	mmu04930	Type II diabetes mellitus
	mmu04950	Maturity onset diabetes of the young (MODY)
Fatty acid and adipocyte related pathways	mmu00061	Fatty acid biosynthesis
	mmu04920	Adipocytokine signaling pathway
	mmu04923	Regulation of lipolysis in adipocytes
Digestion and absorption related pathways	mmu04973	Carbohydrate digestion and absorption
	mmu04974	Protein digestion and absorption
	mmu04975	Fat digestion and absorption
Cholesterol related pathways	mmu04979	Cholesterol metabolism
	mmu00140	Steroid hormone biosynthesis
	mmu00120	Primary bile acid biosynthesis
	mmu04927	Cortisol synthesis and secretion
	mmu04976	Bile secretion
	mmu04913	Ovarian steroidogenesis
	mmu00100	Steroid biosynthesis
	mmu04925	Aldosterone synthesis and secretion
	mmu04934	Cushing syndrome
	mmu03320	PPAR signaling pathway
	mmu04152	AMPK signaling pathway
Additional Obesity and Metabolic Syndrome related pathways	mmu04932	Non-alcoholic fatty liver disease (NAFLD)
	mmu04714	Thermogenesis
	mmu04371	Apelin signaling pathway
	mmu04146	Peroxisome

Table A3. qPCR primer sequences.

Gene	Forward Primer	Reverse Primer
B2m	GGTCTTTCTATATCCTGGCTCAC	ACATGTCTCGATCCCAGTAGA
Srd5a3	GGATGATAAGAATGTGTATGTTCTGG	TTTATACTGATGGGCGGATGAC
Hsd11b1	TTGGCCTCATAGACACAGAAAC	TGTGCCTTTGATGATCTCCAG
Ephx1	CTATGGCTTCAACTCCAGCTAC	GATGTCCAGCCCTTCAATCTT
Ppp2r5a	AAGTTTGTCCAACAGCTCCT	TGCTTTCTGATGAACGCTCT

Table A4. Phenotypes in each study population.

Population	Phenotype	Sex (p-value)	Diet (p-value)	Sex * Diet (p-value)	Diet	Sex	Mean	+/-
B6	Fat percentage (after feeding trial) (% of body weight)	0.143	0.001	0.059	American	Male	27.4	5.2
						Female	20.7	8.8
					Ketogenic	Male	15.5	7.1
						Female	17.2	5.7
FVB	Fat percentage (after feeding trial) (% of body weight)	0.048	0.251	0.916	American	Male	18.1	6.9
						Female	13.4	6.7
					Ketogenic	Male	20.4	7.5
						Female	16.2	6.4
F1	Fat percentage (after feeding trial) (% of body weight)	< 0.001	0.007	0.088	American	Male	32.5	3.2
						Female	19.2	4.7
					Ketogenic	Male	25.9	4.7
						Female	19.3	4.7
F2	Fat mass gain (during feeding trial) (g)	< 0.001	0.373	0.120	American	Male	7.4	4.3
						Female	4.6	3.0
					Ketogenic	Male	8.5	4.6
						Female	4.4	2.9
	Lean mass gain (during feeding trial) (g)	< 0.001	0.003	< 0.001	American	Male	10.0	3.0
						Female	5.4	1.6
					Ketogenic	Male	8.2	3.1
						Female	5.6	1.7
	Serum HDL cholesterol (after feeding trial) (ng/mL)	< 0.001	0.027	0.956	American	Male	207.2	31.1
						Female	163.1	58.1
					Ketogenic	Male	197.3	33.0
						Female	152.6	58.3

Table A5. Effects of sex and diet on phyla abundance

Phylum	Sex (p-value)	Diet (p-value)	Sex*Diet (p-value)	% Variance Explained by Sex	% Variance Explained by Diet	% Variance Explained by Sex*Diet
Actinobacteria	0.246	< 0.001	0.209	0.11	64.79	0.13
Bacteroidetes	0.315	< 0.001	0.794	0.18	25.49	0.01
Cyanobacteria	0.055	< 0.001	0.137	0.82	3.56	0.5
Firmicutes	0.411	< 0.001	0.330	0.06	61.22	0.09
Proteobacteria	0.814	< 0.001	0.638	0.01	3.02	0.05
Verrucomicrobia	0.031	< 0.001	0.708	0.99	9.21	0.03

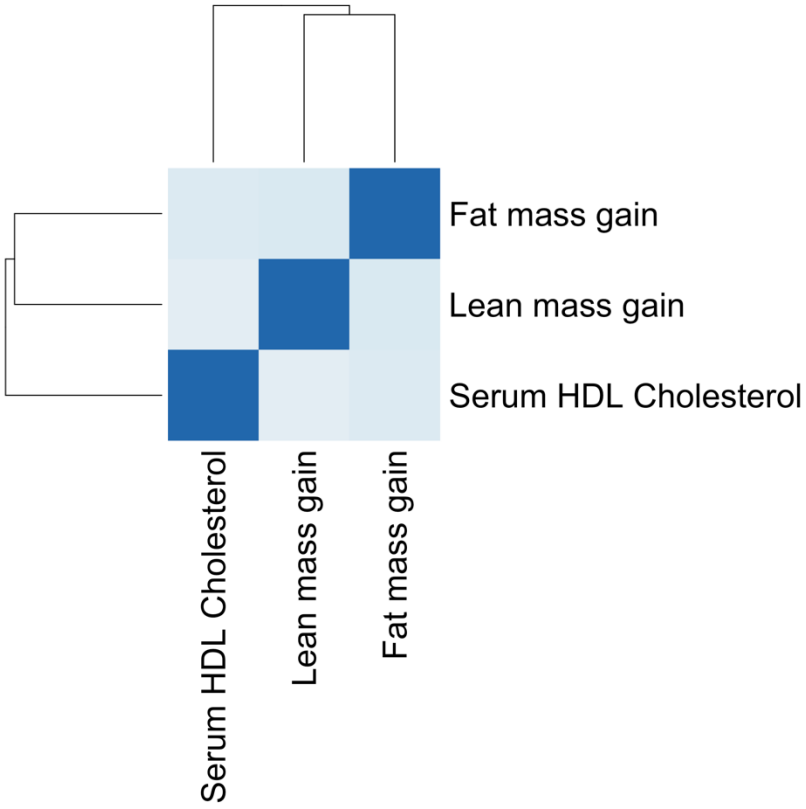

Figure A1. Pearson's correlation between phenotypes. Plot depicting the correlation between examined phenotypes. Range of correlation coefficients: blue: r = 1, red: r = -1.

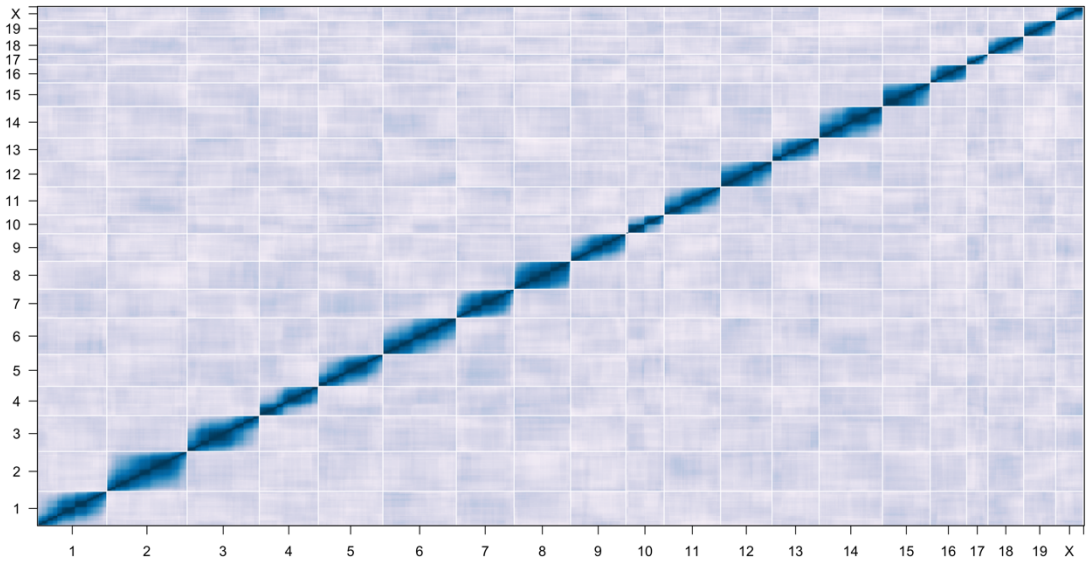

Figure A2. Genetic map after phasing to the founder lines. Genetic map of 16667 informative SNPs after phasing to the founder lines.

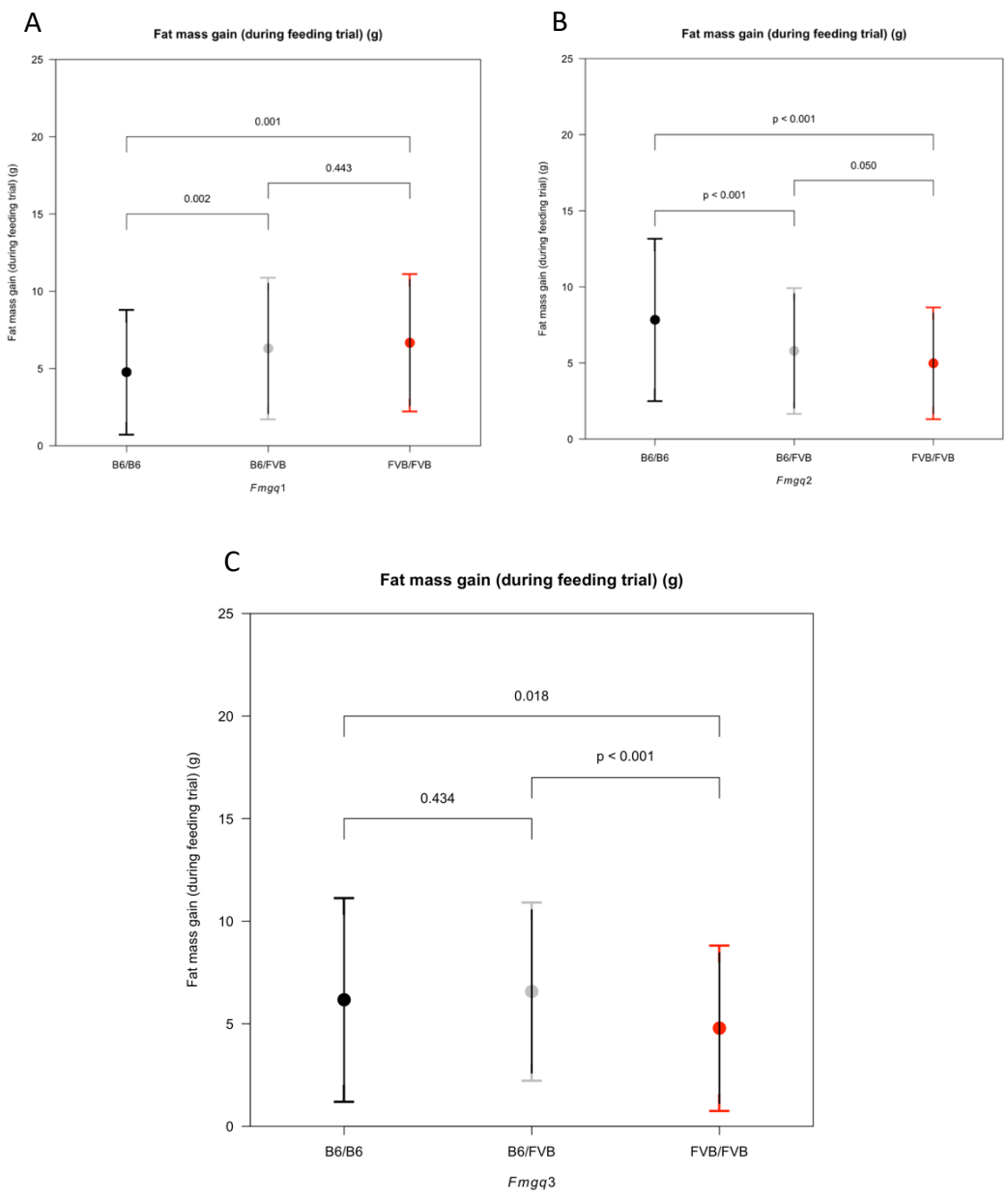

Figure A3. Fat mass gain effect plots. Fat mass gain effect plots. **A)** *Fmgq1,* combined model. **B)** *Fmgq2,* combined model. **C)** *Fmgq3,* combined model.

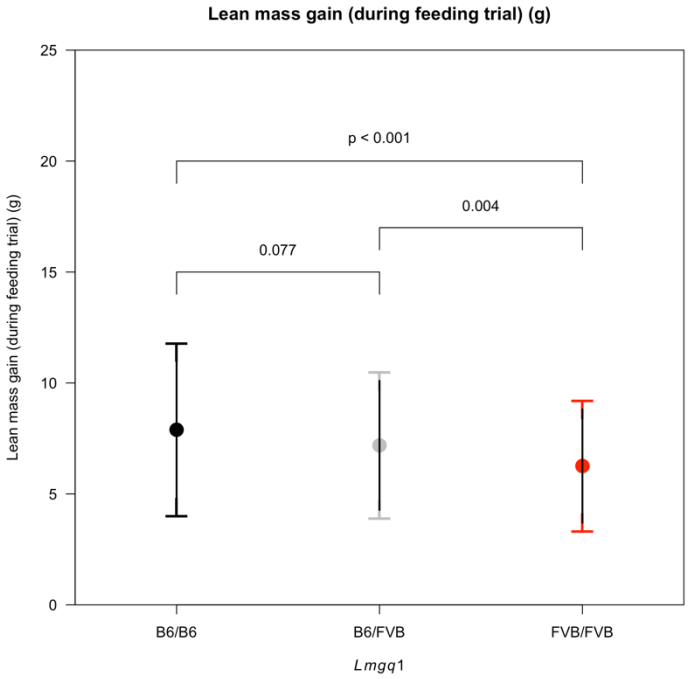

Figure A4. Lean mass gain effect plot. Lean mass gain effect plot for *Lmgq1* in the combined model.

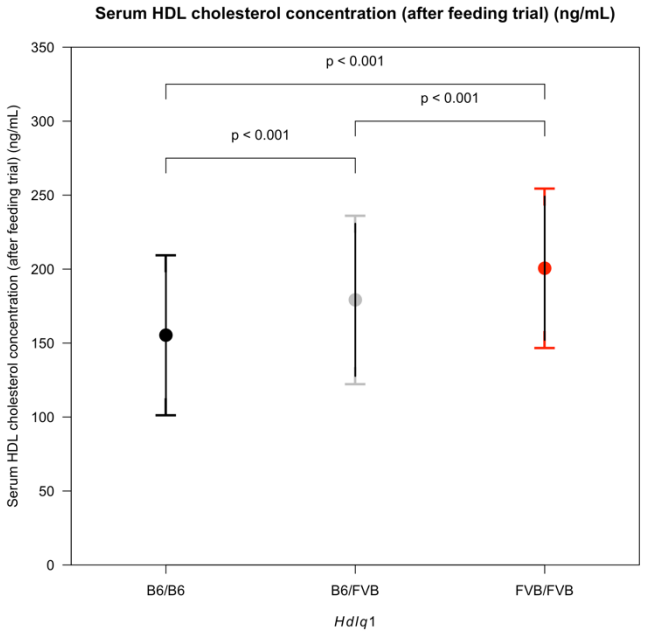

Figure A5. Serum HDL cholesterol concentration effect plot. Serum HDL cholesterol concentration effect plots for *Hdlq1* in the combined model.

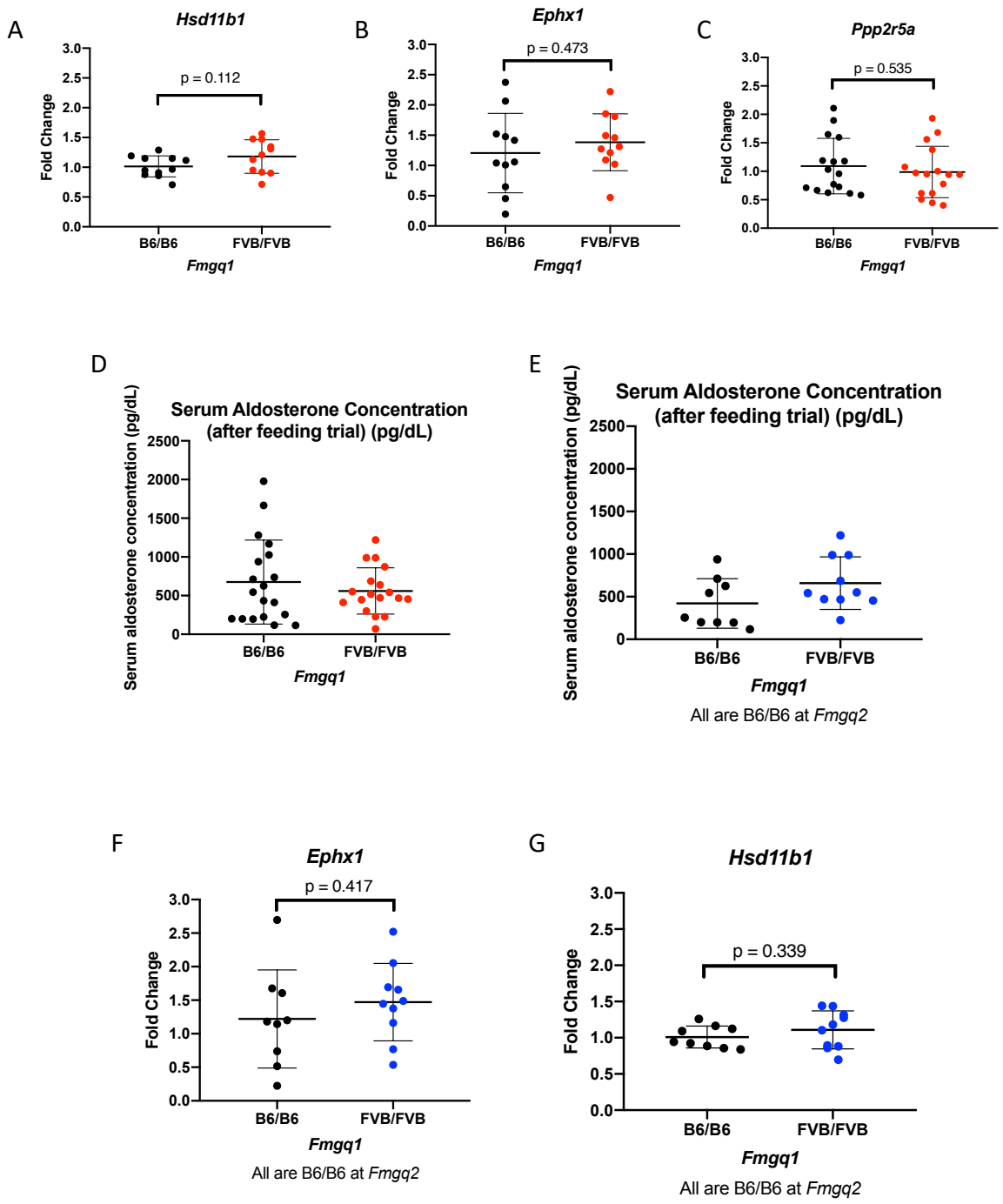

Figure A6. Candidate genes. Candidate gene expression and association with top QTL markers.

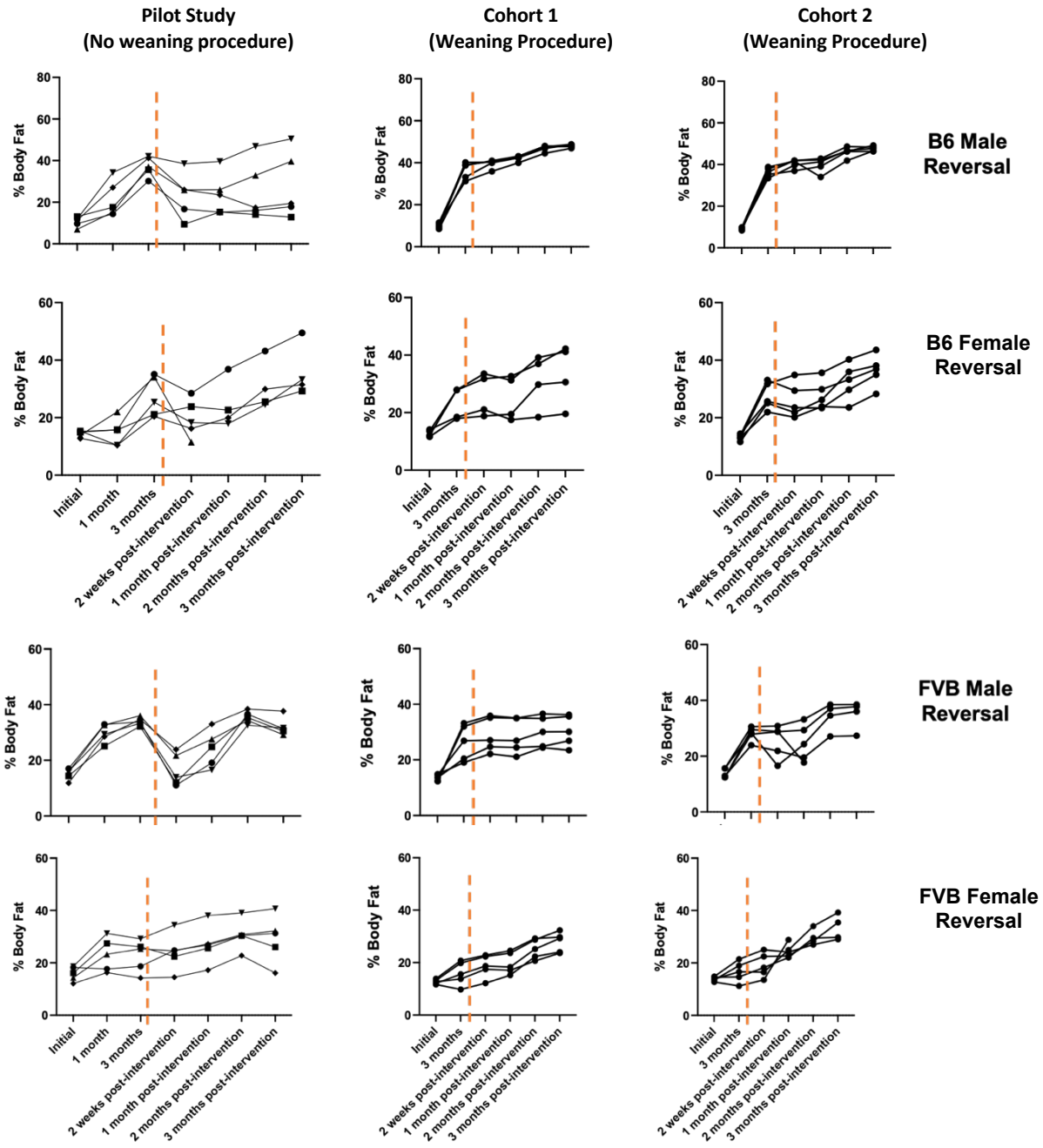

Figure A7. Percentage of body fat during the feeding trial with and without a weaning procedure during the dietary intervention. During the pilot study, no weaning procedure was used during the dietary intervention. Abrupt introduction of the dietary intervention resulted in stress-induced fat loss that eventually plateaued. With the addition of the 2-week weaning procedure, the percentage of body fat remained stable throughout the 2-week weaning period. The dashed orange line indicates where the dietary intervention was introduced.