SYSTEMS GENETICS APPROACH TO DEFINING GENETIC AND MICRORNA ASSOCIATION WITH DIET-ASSOCIATED ATHEROSCLEROSIS

Alisha Renee Coffey

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum in Genetics in Molecular Biology in the School of Medicine.

Chapel Hill
2018

Approved by:

Brian Bennett

Terry Furey

Samir Kelada

Nobuyo Maeda

Karen Mohlke

Praveen Sethupathy

ABSTRACT

Alisha Renee Coffey: Systems Genetics Approach to Defining Genetic and MicroRNA Association with Diet-Associated Atherosclerosis
(Under the direction of Brian J. Bennett and Praveen Sethupathy)

Atherosclerosis is a progressive condition that can develop over the course of decades, and eventually lead to cardiovascular disease, which is the leading cause of death in the world. Metabolic imbalance in the liver can cause dyslipidemia (abnormal levels of lipids) and aberrant levels of metabolites such as trimethylamine N-oxide (TMAO), both of which are associated with increased risk of atherosclerosis. Understanding the genetic influences and molecular networks that govern these processes is necessary to identify candidates for more effective therapeutics.

MicroRNAs (miRNAs) have emerged as important post-transcriptional regulators of gene expression in various biological and pathological processes, and have been shown to be dysregulated in atherosclerosis. Most previous mouse studies of miRNAs in atherosclerosis have a common limitation in that they have utilized single inbred strains to study individual miRNAs of interest. These have left important knowledge gaps in the field that motivate new studies to understand the response of miRNAs to atherogenic diets in diverse genetic backgrounds. Also, because miRNAs are purported to function cooperatively in groups, it is important to not simply focus investigative efforts on the actions of single miRNAs. My research directly addresses these gaps in the field. Specifically, I utilized an unbiased systems genetics approach in

a genetically diverse mouse cohort from the Diversity Outbred (DO) resource to identify co-regulated modules of hepatic miRNAs, or rather groups of miRNAs in the liver that exhibit a shared response to a pro-atherogenic diet. I also quantified the extent of association between the miRNA groups, gene modules, and several cardiometabolic traits. Notably, we found that one particular group of miRNAs, comprising members such as miR-146, miR-27, miR-24, miR-199, and miR-181, is strongly correlated with circulating levels of low-density lipoprotein cholesterol.

The majority of previous studies on the cardiovascular disease risk factor and liver metabolite TMAO focus on its regulation by the gut microbiota as well as some host liver enzymes. However, there have been very few studies to investigate the effects of host genetics on circulating TMAO levels. To understand the genetic influences on TMAO, I performed quantitative trait loci mapping in the same DO cohort mentioned above. I identified a novel association between a locus on chromosome 12 and circulating TMAO levels. Also, I identified miR-146 as a strong candidate regulator of TMAO, which was corroborated in additional animal models (a mouse and non-human primate model) of cardiometabolic dysfunction.

The findings from this body of work contribute to the field by addressing two major knowledge gaps, and warrant further study and validation in the context of atherosclerosis.

To all of those who came before me – who fought, suffered, hoped, prayed,
or sacrificed their lives to make this possible for me.
I am my ancestors' wildest dreams.

ACKNOWLEDGEMENTS

I'd first like to thank God for blessing me throughout this entire journey. This was the best and hardest time in my life, but I made it through.

I also would like to thank my mentors and advisors, Brian Bennett and Praveen Sethupathy. Brian has always been so encouraging and helpful. Even with the miles and miles of distance between us, he always was dedicated to being a present mentor, and made absolutely sure that I had everything that I needed, from lab equipment to IT help; ending every conversation with, "Is there anything else you need from me?" Praveen has also encouraged me to strive to do better, and was always willing to teach, and make sure I understood concepts. He always reminded me to "get out of the weeds" and always question the biological meaning from a broader perspective. I thank him for taking the chance on accepting me into his lab, and fully stepping into the role of mentor without hesitation. I know I may not have been the perfect graduate student for either Brian or Praveen, but I know that they always made sure to let me know that they were there if ever I needed anything. I often compare my advisors with others, and could not have been happier with anyone else. Thank you for making the graduate experience and training that much easier.

To my lab members, past and present, thank you all for helping me, giving me suggestions and advice, finding random tiny mistakes in my presentations, teaching

me lab techniques, sharing your code, and never making me feel excluded. You guys were a great lab family!

I want to acknowledge my dissertation committee members, Terry Furey, Nobuyo Maeda, Karen Mohlke, and Samir Kelada. Everyone was always so understanding of who I am as a student, and so often willing to give suggestions and advice. Thank you for making me feel comfortable, yet challenging me in all of my committee meetings and individual meetings with you.

I'd like to thank my friends for seeing me through this journey. Kim Stratford and Nicole Fleming especially got me through these last couple of years of graduate school. Thanks for literally making me keep the faith, and for understanding the struggle. To Lennora Folkes, I owe thanks for lots of laughs. Even though we are on opposite sides of the country, she always made sure to check on me, listen to my rants, and put a smile on my face. To two of my oldest friends, Ashley Mays Agbemaple and Lauren Royal McDonald Moss, I thank them for randomly calling or texting me, and still managing to include me in their lives, even when I was unplugged and seemingly missing in action. To various other friends that may not have gotten a personal mention here, you are not forgotten. I thank you for making these past few years great. I honestly would not have survived grad school without *all* of you!

To my parents, I love you and thank you for always being a support system. To my mom who laughed, cried, and shared every experience with me even though she was hundreds of miles away. Mommy, you truly are a great mom, and I thank

you for always being a phone call away. Daddy, thank you for always praying for me, and supporting me through it all. You were always willing to talk things out and help me plan my next steps. To my sister, Jonaee, I thank you for being the type of big sister that I could share things with, and have deep or silly conversations with. I also would like to thank my grandmothers for their love, and for calling to make sure I was doing well. To my aunts, I thank them for their support and love. To the rest of my family, thank you for all of the encouragement and well wishes.

I am grateful for everyone in the Office of Graduate Education. I am most grateful for Initiative for Maximizing Student Diversity, especially Ashalla Freeman, who was always willing to hear how I was doing, and offer advice for moving forward. In the department administration team, I would like to extend my gratitude to John Cornett, Kathy Cornett, and Cara Marlow for always being so nice to me, even when I asked them long questions, or had to reschedule a room three times. I appreciate all of you!

PREFACE

The work in Chapter 2 previously appeared as an article in the journal Physiological Genomics, in 2017, under the direction of Dr. Brian J. Bennett and Dr. Praveen Sethupathy. The citation for this manuscript is as following: Coffey AR, Smallwood TL, Albright J, Hua K, Kanke M, Pomp D, Bennett BJ, Sethupathy P. Systems genetics identifies a co-regulated module of liver microRNAs associated with plasma LDL cholesterol in murine diet-induced dyslipidemia. Physiol Genomics, DOI: 10.1152/physiolgenomics.00050.02017.

Experiments performed on LIRKO mice and African Green Monkeys in Chapter 3 were executed by individuals in the labs of Sudha Biddinger and Ryan Temel, respectively, as denoted by the appropriate references.

The work highlighted in this dissertation was supported through several funding sources including National Science Foundation Graduate Research Fellowship, National Institutes of Health, and the University of North Carolina Nutrition Research Institute.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

3’ UTR	3’ untranslated region
ABCA1	ATP-binding cassette A1
ABCG1	ATP-binding cassette G1
ACAT1	Acetyl-CoA acetyltransferase 1
Apoa1	Apolipoprotein A1
Apoe	Apolipoprotein E
CC	Collaborative Cross
CVD	Cardiovascular disease
DO	Diversity Outbred
eQTL	Expression quantitative trait loci
GWAS	Genome-wide association studies
HDL	High-density lipoprotein
HFCA	High fat diet with added cholic acid
HFHC	High fat, high cholesterol diet
HMDP	Hybrid mouse diversity panel
HP	High protein diet
IL-1β	Interleukin 1 beta
LDL-C	Low-density lipoprotein cholesterol
LDLR	Low-density lipoprotein receptor
LDLRAP1	Low-density lipoprotein receptor adaptor protein 1
LIRKO	Liver insulin receptor knockout

LPGAT1	Lysophosphatidylglycerol acyltransferase 1
LPL	Lipoprotein lipase
miRNA	micoRNA. Individual miRNAs are identified as “miR”, “mmu-miR”, or “hsa-miR” followed by a number
MMP-13	Matrix metalloproteinase 13
mRNA	messenger RNA
MTP	Microsomal triglyceride transport protein
MYC	Myelocytomatosis gene
NAFLD	Non-alcoholic fatty liver disease
oxLDL	Oxidized LDL
PCSK9	Proprotein convertase subtilisin/kexin type 9
PPARα	Peroxisome proliferator-activated receptor alpha
Pre-miRNA	miRNA precursor
Pri-miRNA	Primary miRNA
QTL	Quantitative trait loci
SNP	Single nucleotide polymorphism
TMA	Trimethylamine
TMAO	Trimethylamine-N-oxide
TOM	Topological overlap measure
VLDL-C	Very low-density lipoprotein cholesterol
WGCNA	Weighted Gene Co-expression Network Analysis

CHAPTER 1: INTRODUCTION

1.1 ATHEROSCLEROSIS

According to the World Health Organization, cardiovascular disease (CVD) is the number one cause of death worldwide, with ~30% of all deaths resulting from complications of CVD in 2015 [WHO]. CVD is not one disease, but can be one (or more) of multiple systemic conditions including, but not limited to angina, coronary heart disease, and peripheral artery disease. Atherosclerosis, a condition involving the gradual hardening and narrowing of the major or medium arteries caused by fatty deposit build-up, is a common cause and hallmark of CVD [AHA].

Atherosclerosis is a progressive condition that can initiate as early as within the first ten years of life [Lusis 2000]. Atherosclerotic plaques begin as fatty streaks that are mainly composed of monocyte-derived macrophages, and T lymphocytes [Ross et al. 1995]. These tend to form in the regions of arteries that have an increase in blood pressure-induced tensile stresses such as in branching sites, or in sites where there are changes in the direction of blood flow such as in the aortic arch [Langille et al. 1997]. Over time and in favorable conditions, the streak can grow to become an advanced fibrous lesion consisting of smooth muscle cells, extracellular matrix, debris such as calcium and connective tissue, and a lipid-rich necrotic core

[Lusis 2000]. The advanced lesion, if left untreated, can lead to ischemia, heart attack, stroke and death.

A number of environmental factors are known to promote the development of atherosclerosis. These include having a lifestyle that lacks exercise, overuse of alcohol, chronic smoking, and (of particular importance) an unhealthy diet that is high in saturated fats and cholesterol [AHA, WHO]. Of these, diet is one of the largest contributors to atherosclerosis and CVD. A health comparison of Japanese people living in Japan versus those living in the United States showed that individuals that adopted an American diet and lifestyle had higher incidence of CVD, and that the frequency of incidence was comparable to that of other Americans; thus, highlighting the relevance and importance of diet in CVD [Kagan et al.1974]. An unhealthy, unbalanced diet consisting of large quantities of high fat, high cholesterol foods can cause a type of dyslipidemia known as hyperlipidemia, or abnormally high lipid levels. In addition, supplementation of the high fat diet with cholic acid, a type of hydrophobic bile acid, causes an increase in lipid absorption in the intestines, an inflammatory response, and increased cholesterol excretion [Fickert et al. 2001]. To that end, cholic acid is often added to a high fat diet to induce hypercholesterolemia in mice. Elevated concentrations of certain cholesterol fractions, such as low-density lipoprotein cholesterol (LDL-C) or triglycerides, often leads to atherosclerosis and many forms of CVD [Lusis 2004].

1.2 LIPID HOMEOSTASIS, THE LIVER, AND ATHEROSCLEROSIS

Lipid homeostasis is a balancing act between lipid intake, distribution, storage, and elimination. Although there are several organs and tissues that have a role in lipid homeostasis, the liver is one of the central organs in lipid metabolism and distribution. It is responsible for the synthesis, export, and clearance of lipids from circulation. Liver metabolism is meticulously regulated by the quality and quantity of diet, and energy availability.

Dietary lipids, as well as lipids from other sources in the body (e.g. adipose tissue, macrophages), are sent to the liver as components of lipoproteins [Ponziani et al. 2015]. However, atherogenic dyslipidemia can occur when there is a deviation from normal uptake, export of lipids, or as a result of chronically consuming foods high in fat and cholesterol [Austin et al. 1990]. After large amounts of dietary lipids are transported to the liver, an overproduction of very low-density lipoprotein (VLDL; composed of apolipoprotein B, phospholipids, triglycerides and cholesteryl esters) often occurs, and this is a common hallmark of atherogenic dyslipidemia [Adiels et al. 2008, Hebbachi et al. 2001]. VLDL is converted to LDL by a series of processes that involves an initial transfer of triglycerides by CETP, and subsequent lipase activity [Ginsberg et al. 1994, Manjunath et al. 2013]. In fact, increased amounts of circulating LDL-C, which is especially rich in cholesterol, tend to accumulate in the arterial wall, and undergo oxidation, which initiates an inflammatory response that causes recruitment of monocytes to the area [Helkin et al. 2016]. Once there, the

monocytes differentiate into macrophages, which proceed to engulf and digest the oxidized LDL-C (oxLDL). However, oxLDL overload often occurs in macrophages, thereby causing them to become foam cells: cells typically found in abundance in atherosclerotic plaques [Helk et al. 2016]. Thus, an increase in circulating lipids initiated in large part by the liver, is extremely important in the onset and progression of atherosclerosis.

A metabolic imbalance in the liver often results in a shift in lipid homeostasis. For instance, non-alcoholic fatty liver disease (NAFLD) has been implicated as a predictor for several cardiometabolic conditions, including CVD [Lonardo et al. 2017]. Not only is NAFLD a strong risk factor for CVD, but also, evidence suggests that it may instigate the premature progression of fatty streaks to advanced lesions, which ultimately increases the incidence of cardiac events (e.g., heart attack, stroke) and mortality irrespectively of other CVD risk factors [Lonardo et al. 2006]. Furthermore, it has been shown to cause chronic inflammation, which is another hallmark of atherosclerosis and CVD [Sinn et al. 2017, Zeb et al. 2016]. NAFLD also can be used as a predictor for hypertension (another important risk factor for CVD), and the severity of NAFLD-induced fibrosis in the liver can additionally serve as an indicator of CVD mortality [Ekstedt et al. 2015, Li et al. 2017, Lopez-Suarez et al. 2011].

The liver is also responsible for processing and producing a number of metabolites, including pro-atherogenic trimethylamine N-oxide (TMAO), which is associated with a 2.5-fold increase in the risk of developing CVD [Tang et al. 2013].

The microbiota first convert dietary choline (from sources like red meat, eggs, and fish) to trimethylamine (TMA) gas, which is then absorbed and transported to the liver where it is oxidized into TMAO, predominately by flavin monooxygenase 3 (FMO3) [Wang et al. 2011]. Studies in atherosclerosis-prone mice show that a regular chow diet supplemented with TMAO or choline is enough to cause formation of atherosclerotic plaques in the aorta, whereas treatment of the same mice with antibiotics prevented this increase in plaque formation [Wang et al. 2011]. A similar study in humans showed a marked reduction in TMAO levels after treatment with antibiotics, and a rebound after antibiotic treatment was complete [Tang et al. 2013]. Further studies in humans comparing vegetarians and omnivores showed that omnivores had higher baseline levels of TMAO and a greater capacity to produce TMAO after L-carnitine feeding than the vegetarians; thereby, highlighting the differences in diet-mediated gut microbiota composition between meat-eaters and vegetarians [Koeth et al. 2013].

While these studies suggest a causal relationship between the microbiota, TMAO, and atherosclerosis, they do not address the relationship between TMAO levels and host genetics. A genome-wide association study (GWAS) done in the Hybrid Mouse Diversity Panel identified a locus for TMAO levels that overlaps with an expression quantitative trait locus (eQTL) on chromosome 3 for *Slc30a7* expression [Hartiala et al. 2014]. In their comparative GWAS performed in humans, the association was found to be merely suggestive. Although the association did not replicate when the lead SNP was genotyped in the individuals, it is still possible that

this gene is involved with modulating TMAO levels [Hartiala et al. 2014]. Additionally, in a recent variant analysis done in individuals with trimethylaminuria disorder – which prevents individuals from converting TMA to TMAO – rare SNPs and insertion/deletions were identified in genes other than FMO3, thus providing further evidence that TMA metabolism and TMAO levels may be affected by other genes and genetic variation [Guo et al. 2017]. This provides the motivation for additional studies to further elucidate genetic influences, whether at the genomic level or molecular level, on TMAO.

1.3 MICRORNAS

A prominent species of RNA, known as microRNAs (miRNAs), has emerged as important regulators of gene expression, has been shown to function within molecular regulatory networks, and has been implicated in many biological processes that control lipid and metabolite metabolism. Here, it is important to highlight their actions and significance in CVD as they are part of the foundation of the work presented in this dissertation.

1.3.1 Overview of MicroRNA Biogenesis

MiRNAs are small non-coding RNAs that are post-transcriptional repressors of messenger RNA (mRNA) expression. They are encoded in the genome in intergenic regions, or within exonic or intronic regions of host genes [Altuvia et al. 2005]. Primary miRNA transcripts (pri-miRNAs) are transcribed predominantly by

RNA Polymerase II, although there are some transcribed by RNA Polymerase III [Borchert et al. 2006, Lee et al. 2004]. The pri-miRNAs, which can be over 1 kilobase in length, are trimmed by nuclear RNase III Drosha to reveal one or more miRNA precursors (pre-miRNAs), which are stem loop intermediates that are around 70 nucleotides in length [Lee et al. 2002, Lee et al. 2003]. The pre-miRNAs are then transported to the cytoplasm by the Ran-GTP-dependent RNA binding protein, Exportin-5, where they are further processed by RNase III Dicer into a ~22 nucleotide duplex [Yi et al. 2003]. One of the strands of the duplex is incorporated into the RISC (RNA-induced silencing complex), more specifically with an Argonaute protein, to locate an accessible complementary target sequence commonly found in the 3' untranslated region (UTR) of specific mRNAs. Once the mature miRNA's seed is sufficiently hybridized with a target sequence, the RISC can prevent translation by deadenylation, recruiting P body components to the mRNA, ribosomal drop off during translational elongation, and possibly other proposed methods [Eulalio et al. 2008, Wakiyama et al. 2007, Parker and Sheth 2007, Petersen et al. 2006].

1.3.2 Regulatory Function of MicroRNAs

MiRNAs were first studied in *Caenorhabditis elegans* (*C. elegans*), in the context of post-embryonic development, and were regarded as molecular switches. Lin-4, which was originally presumed to be a protein-coding gene, was found to be a miRNA that heterochronically decreases expression of *lin-14* during the L1 larval stage [Ambros and Horvitz 1987, Lee et al. 1993, Wightman et al. 1993]. Without lin-

4, the lingering high levels of *lin-14* cause the larva to undergo a repeat of previous developmental stages, thus causing an abnormal adult lacking critical tissues. These findings demonstrated that miRNAs can turn essential genes that progress developmental programs on or off. Beyond development, miRNAs have been shown to perform this on/off switch-like function in other pathways, such as those associated with cancer cell proliferation. For instance, overexpression of let-7a was shown to cause a decrease in expression of *MYC* mRNA and protein, and decrease proliferation of lymphoma cells [Sampson et al. 2007]. More recently, researchers have harnessed and utilized the molecular on/off regulatory power of miRNAs in the detection and purification of cell types for which specific antigens have not been identified or extensively validated [Miki et al. 2015]. Overall, miRNA function as a molecular switch demonstrates that the effect of a single miRNA can have a robust regulatory effect on the expression of highly impactful genes.

Although there are some contexts in which miRNAs operate as molecular switches, they more often operate as fine tuners of gene expression. On their own, miRNAs tend to have moderate effects on target gene expression. Most miRNAs merely stymie the expression of their target mRNAs to maintain levels below a certain threshold [Mukherji et al. 2011]. In other words, the miRNAs function to decrease noisy expression of their targets in order to enhance the robustness of some other switch-like molecule [Cohen et al. 2006, Mukherji et al. 2011, Siciliano et al. 2013]. This buffering function is likely the reason why the dysregulation of a miRNA's expression can lead to drastic physiological outcomes, including

atherosclerosis and cardiometabolic dysfunction. The dysregulated miRNA's function can obscure its target genes' expression thresholds, thereby rendering the other molecular switch ineffective.

Many studies across the field of miRNA biology present results in such a way that may lead one to assume that a single miRNA may regulate a single mRNA transcript at a time in a particular tissue or disease context. However, generally, miRNAs can have hundreds of target genes, many of which may be simultaneously expressed in the cell at any given time [Bartel 2009]. In addition, the vast majority of mRNAs, on average, contain four or more target sites in their 3' UTRs, thus allowing a single mRNA to be targeted by multiple miRNAs [Friedman et al. 2009, John et al. 2004]. Furthermore, these target sites are often conserved together, and many genes operating within the same networks have the same miRNA target sites [Chan 2005]. All of this supports the notion that miRNAs can regulate gene expression in a combinatorial manner. In fact, there is evidence to indicate that multiple miRNAs can cooperatively target the same mRNAs to regulate their expression [Coronello et al. 2013, Doench et al. 2004, Hashimoto et al. 2013, Hon et al. 2007, Krek et al. 2005, Wu et al. 2010]. Even beyond this concept, it has been shown that networks of miRNAs can coordinately target entire networks of mRNAs, many of which are part of the same or related biological processes [Georges et al. 2008, Gusev et al. 2007, Joung et al. 2007]. Transcription factors can promote the expression of several miRNAs and their targets thereby creating complicated expression networks such as feedforward loops and feedback loops [Vera et al. 2012].

1.4 MICRORNAS IN ATHEROSCLEROSIS

While processes of atherosclerosis and CVD progression encompass several tissues including immune cells, endothelial cells, the intestine, and adipose, multiple miRNAs in the liver have been implicated as being drivers or responders to atherosclerosis or cardiometabolic dysfunction. Yet, there are a precious few that have been extensively studied. In the following paragraphs, I will briefly describe them and their importance in atherosclerosis.

1.4.1 miR-33

miR-33 is arguably the most well-studied miRNA when it comes to lipid metabolism and atherosclerosis. Its genomic location in mice is within an intron in the SREBF2 gene, which encodes a transcription factor that is essential for turning on the expression of cholesterol biosynthesis genes [Horton et al. 2002, Rayner et al. 2010]. MiR-33 is co-expressed with SREBP-2 in macrophages, the liver, and other tissues in the absence of cholesterol or after treatment with a statin, and is decreased in cholesterol feeding [Rayner et al. 2010]. It has been shown to target *ABCG1*, and have 3 functional target sites in the 3'UTR of *ABCA1*, both of which encode membrane proteins that control cellular cholesterol efflux [Rayner et al. 2010]. MiR-33 also targets *SREBP1*, which is an activator of fatty acid synthesis genes [Horie et al. 2013]. MiR-33 knockout (KO) mice were shown to have elevated SREBP-1, as well as higher incidences of obesity and hepatic steatosis (fatty liver)

[Horie et al. 2013]. This action is conserved in humans, however, humans have two members of the miR-33 family, miR-33a and miR-33b, that differ by only 2 nucleotides; whereas mice only have one, which is homologous to human miR-33a. Antagonism of miR-33 results in an increase in fatty acid oxidation genes and high-density lipoprotein (HDL), and a decrease in expression of fatty acid synthesis genes and circulating VLDL triglyceride levels in non-human primates [Rayner et al. 2011]. In mice, treatment of antagomiR-33 or miR-33 KO ultimately caused the regression of atherosclerotic plaques in two traditional models of atherosclerosis, *Ldlr*$^{-/-}$ and *Apoe*$^{-/-}$, respectively, which led to the consideration of miR-33 as a therapeutic target [Horie et al. 2012, Rotllan et al. 2013]. However, long-term inhibition of miR-33 showed no change in the state of disease in *Ldlr*$^{--/-}$ mice, and even had detrimental effects in wild-type C57BL/6 mice fed a high fat diet (HFD) [Marquart et al. 2013, Goedeke et al. 2014].

1.4.2 miR-30c

miR-30c is a more recently discovered miRNA in the context of atherosclerosis. In 2013, it was discovered to have atherosclerosis and hyperlipidemia-reducing effects when overexpressed by lentivirus in the livers of C57BL/6 mice fed a western diet [Soh et al. 2013]. Injecting miR-30c mimic was shown to have similar effects in *Ldlr*$^{-/-}$ mice, as well as *ob/ob*, and *db/db* mice, which are other models of cardiometabolic dysfunction [Irani et al. 2016, Irani et al. 2017]. miR-30c targets *MTP*, or microsomal triglyceride transport protein, which assists in

lipoprotein assembly and lipidation to form LDL. In an MTP activity assay, miR-30c decreased MTP activity by 50%, whereas other miR-30 family members had no effect [Soh et al. 2013]. It exerts its lipid synthesis reducing activity by also targeting *Lpgat1*. This activity, along with its lack of steatosis or other harmful side effects makes miR-30c an attractive candidate for use as a therapeutic agent against atherosclerosis [Irani et al. 2015, Irani et al. 2016, Soh et al. 2013].

1.4.3 miR-148a

miR-148a is an intergenic miRNA that regulates lipid metabolism in the liver, adipose tissue, and macrophages by regulating *LDLR* and *ABCA1* [Baldan and Fernandez-Hernando 2016]. A high-throughput screen revealed that miR-148a mimic led to a decrease in LDLR [Goedeke et al. 2016]. Further investigation showed that miR-148a expression is modulated by SREBP1, which is in turn modulated by lipid levels [Goedeke et al. 2016]. When miR-148a is inhibited, LDLR and ABCA1 are increased, which causes a decrease in circulating LDL-C and an increase in circulating HDL-C [Goedeke et al. 2016, Wagschal et al. 2015]. In the livers of mice and non-human primates fed a HFD, miR-148a is found to be upregulated, and a human GWAS solidified a connection between SNPs near miR-148a and abnormal circulating lipids [Goedeke et al. 2016, Wagschal et al. 2015]. All of these findings lend the suggestion of miR-148a as another potential therapeutic target for dyslipidemia and atherosclerosis.

1.4.4 miR-27

Yet another miRNA that plays a role in lipid metabolism and atherosclerosis is miR-27. It is another miRNA that targets *LDLR* and *ABCA1*, but also targets other genes involved in lipid transport [Alvarez et al. 2015]. Overexpression of miR-27a both directly and indirectly causes a decrease in LDLR activity, and an increase in PCSK9, a negative regulator of LDLR [Alvarez et al. 2015]. MiR-27a and miR-27b both regulate cholesterol transport in and out of THP-1 macrophages by targeting ABCA1, LPL, and ACAT1, and they also indirectly regulate the expression of apoA1 *in vitro* [Zhang et al. 2014]. *In vitro* experiments also showed that overexpression of miR-27b decreased MMP-13 expression both directly and via targeting of IL-1β, which may promote plaque formation during atherosclerosis progression [Akhtar et al. 2010, Deguchi et al. 2005, Prescott et al. 1999,]. Yet another mechanism of regulation facilitated by miR-27b is that of PPARα, a transcription factor that is highly expressed in the liver that activates expression of a slew of lipid metabolism genes [Kida et al. 2011]. MiR-27 also regulates cholesterol synthesis by targeting *HMGCR*, which is the rate-limiting enzyme in the pathway [Selitsky et al. 2015]. Furthermore, miR-27 has been shown to regulate a number of genes in other pathways involved with atherosclerosis including inflammation, angiogenesis, and apoptosis [Chen et al. 2012]. With regulatory control over such a wide array of genes, it is no wonder that miR-27 has been identified as a regulatory hub in lipid metabolism, and that its

expression in the liver is altered in dyslipidemia and atherosclerosis [Vickers et al. 2013].

In addition to the four that have been highlighted here, there are a number of other hepatic miRNAs that have been implicated in atherosclerosis. Many of these such as miR-29, miR-146, miR-24, and miR-122 have been identified as regulators of lipid metabolism and have been shown to be abnormally expressed in the disease state [Baldan and Fernandez-Hernando 2016, Kurtz et al. 2015, Rotllan et al. 2016, Willeit et al. 2016]. The growing list of miRNAs relevant to lipid metabolism and homeostasis provides the motivation to identify groups of miRNAs that modulate the underlying processes of CVD.

1.5 ANIMAL MODELS USED FOR STUDYING ATHEROSCLEROSIS

The laboratory mouse has been an invaluable tool for studying the onset, progression, and molecular mechanisms involved in atherosclerosis and its risk factors. Since mice are naturally resistant to atherosclerosis, it was necessary to develop mouse models that are prone to developing atherosclerosis and hyperlipidemia. These are generally on the C57BL/6J background, and commonly harbor a null mutation in a single gene -- the two most widely-used KO mouse models being $ApoE^{-/-}$ and $Ldlr^{-/-}$ mice [Ishibashi et al.1994, Zhang et al. 1992].

ApoE KO mice have a disruption in cholesterol uptake (especially that of cholesterol-rich chylomicrons and VLDL-C) and, therefore, have elevated levels of plasma cholesterol. This phenotype is further exacerbated by HFD or Western diet

feeding [Plump et al. 1992]. Ldlr deficient mice also have disrupted cholesterol uptake (especially that of apoB and apoE-containing lipoproteins), and have a phenotype similar to that seen in humans with familial hypercholesterolemia [Defesche 2004]. Similar to ApoE deficient mice, when Ldlr deficient mice are fed a HFD, the severity of hyperlipidemia and atherosclerosis is worsened. Much of what is known about genes and mechanisms influencing atherosclerosis have been identified using these mouse models. It is worth noting that there are other models, including transgenic mouse models such as ApoE3-Leiden that are also popular for studying cardiometabolic dysfunction, but may have slightly different phenotypic characteristics [van den Maagdenberg et al. 1993].

Although these mouse models have contributed a great deal to what we know about atherosclerosis and hyperlipidemia, utilizing them in experiments comes along with some limitations. For instance, $ApoE^{-/-}$ mice typically have elevated VLDL-C and rarely have plaque rupture without the aid of a perivascular cuff. This is different from the elevated LDL-C and myocardial infarction/stroke-causing plaques that are seen in human atherosclerosis [Emini Veseli et al. 2017]. Also, in both ApoE and Ldlr null mice, the metabolic dysfunction is caused by a disruption in only one gene. The role that genetic variation plays in CVD is, in a way, omitted since the studies utilize mice that are all on the same genetic background and harboring the same mutation. That single mutation, while detrimental in one context, may produce a completely opposite effect in a different mouse strain [Shi et al. 2002]. However

successful these models are, their limitations and shortcomings suggest that a model that more directly mimics the genetic diversity seen in humans is needed.

1.6 SYSTEMS GENETICS APPROACHES FOR STUDYING ATHEROSCLEROSIS

Many of the risk factors have significant genetic components associated with strong heritability and increased susceptibility for developing atherosclerosis. These risk factors, including total cholesterol, HDL-C, triglycerides, body mass index, blood pressure, and type 2 diabetes, together have heritability measures anywhere from approximately 25-80% [Lusis 2004]. These factors, when measured in the population, often convey normal distributions, which are indicative of polygenic traits, or rather traits controlled by multiple genes [Lusis 2004].

Since many cardiometabolic traits are polygenic and influenced by pleiotropic molecules, studying them should involve a method that includes a more comprehensive view of the systems and networks underpinning them. Instead of using a knockout or transgenic animal model that involves a single mutation like in conventional studies, systems genetics studies -- in which there is integration of genetic variation, molecular profiles, and phenotype information -- are becoming increasingly popular as a method for relating genetics with molecular and physiological phenotypes [Sieberts and Schadt 2007]. In these methods, a genetically diverse population is necessary for studying molecular interactions and drawing conclusions that are applicable to multiple genetic backgrounds. In this

case, the genetic diversity acts as multiple genetic perturbations, which modifies the proper function of molecular networks and initiates disease [Civelek and Lusis 2014, Sieberts and Schadt 2007]. Overall, by integrating genetic and genomic data, we can explore how natural variation in a population influences disease and disease-related phenotypes, and eventually identify candidates that exert causal control over them [Kulp and Jagalur 2006].

There are few requirements when choosing a model to use in a systems genetics study. As mentioned, a systems genetics study utilizes natural DNA variation as perturbations, so the model must have high levels of genetic diversity. However, other sources of variation such as environment or population structure must be controlled for. Diversity outbred mice, or DO mice, are a mouse resource that were created for the exact purpose of performing systems genetics studies [Churchill et al. 2012]. The population was created by strategically breeding eight parental mouse strains, including both inbred and wild-derived strains to first form the Collaborative Cross (CC) mice, which are recombinant inbred strains of mice [Churchill et al. 2012]. Then, mice from different CC mouse strains were randomly outbred in a way that ensured that mating pairs were unrelated; thus, resisting genetic drift or loss of variation [Churchill et al. 2012, Schmidt 2015]. The breeding scheme allows for a high level of allelic diversity, an even distribution of variants in gene coding and regulatory regions, as well as small haplotype blocks that decrease with each successive generation, thereby allowing for increasingly high-resolution mapping [Churchill et al. 2012, Gatti et al. 2014, Svenson et al. 2012]. With such a

high degree of naturally-occurring allelic variation, the genetic makeup, phenotypic variability, and susceptibility to various chemicals or diseases mimics that which is seen in humans [Schmidt 2015]. To that end, the DO mouse resource is an idyllic animal model for systems genetics and genetic mapping studies. Specifically, it is a resource that is ideal for performing unbiased systems genetics studies, in which transcriptional signatures are identified in a particular context without the specific modification of the organism's genome. By this method, we can begin to answer explicit questions regarding genetics and regulation, as well as gain a more global view of the system.

1.7 SUMMARY AND SIGNIFICANCE OF RESEARCH PRESENTED

The research within this dissertation features work that is an expansion of a previous study in which DO mice were used to perform quantitative trait loci (QTL) mapping studies to identify novel QTL for clinical phenotypes measured either before or after specific diet treatments [Smallwood 2015]. The work presented here delves a bit deeper by investigating the molecular profiles of gene and miRNA expression in the livers of the same ~300 DO mice. Having the molecular information as well as genotype and phenotypic measurements from the genetically diverse mice, we endeavored to find miRNAs and miRNA co-regulated modules that are associated with cardiometabolic risk factors. In addition, we also attempted to identify genetic variation that regulates these parameters via genes or miRNAs.

Although there is a large amount of evidence supporting groups of miRNAs regulating networks or modules of genes, miRNA studies encompassing cardiometabolic dysfunction commonly involve the investigation of single miRNAs in a single inbred mouse strain, or mice on an inbred strain background. Those studies are not conducive to getting a complete picture of how the fine-tuning action of miRNA groups, or modules, can facilitate the development of particular physiological outcomes. With the work included in Chapter 2, we addressed this deficiency by performing an unbiased systems genetics approach to identifying miRNA modules associated with VLDL/LDL-C in a large and genetically diverse cohort of outbred mice on an atherogenic diet. By utilizing the gene and miRNA expression data and phenotype data, I performed a co-expression network analysis to identify co-regulated modules (or groups) of miRNAs that are most responsive to the atherogenic diet and genetic variation. To do this, I used the weighted gene co-expression network analysis (WGCNA) software. Briefly, WGCNA identifies modules of possible co-regulated miRNAs by calculating the extent of correlation, adjacency, and topological overlap measures for each item to another. This not only groups miRNAs that have similar expression patterns, but also calculates how connected each miRNA is to another, thereby allowing for the identification of potential hubs, or master regulators, of each module. Utilizing this method led to the identification of a particular miRNA module that is both associated with post-diet VLDL/LDL-C, and with gene modules that are also associated with VLDL/LDL-C. With these data, I also identified possible hub miRNAs, miR-199a, miR-181b, miR-24, miR-27a, and

miR-21 isomiR. This work not only addresses the aforementioned shortcomings in the field, but also highlights the importance of using genetically diverse animal models in such studies, as some results did not coincide with findings from studies using only inbred mice; thus, suggesting that the results from the inbred mouse studies may be ungeneralizable to genetically diverse organisms such as humans.

To take advantage of the genetic diversity of the mouse cohort, and to address our question on genetic regulatory action, I integrated the data sets to perform QTL and eQTL mapping on the cardiometabolic parameters, genes and miRNAs. Since not much is known about the host genetic regulation of circulating TMAO, in chapter 3, I utilized the mapping results to shed some light on the molecular underpinnings governing the cardiometabolic trait, TMAO. My work peels back one of the layers of this regulation by identifying miR-146 as a potential regulator of TMAO levels as it is one of the most highly aberrantly expressed miRNAs in the livers of the DO mice, as well as in the livers of two other animal models of cardiometabolic dysfunction with elevated TMAO levels, including non-human primates. Its expression is also significantly correlated with post-diet TMAO thereby supporting miR-146 as a likely driver of, or responder to, elevated circulating TMAO levels (a marker for cardiometabolic dysfunction). Also, QTL mapping done on pre-diet and post-diet levels of TMAO introduces another layer of genetic regulation of TMAO. We identified candidate genes with eQTL overlapping the QTL for TMAO -- a step in the direction of further elucidating the underlying host genetics that modulate circulating TMAO levels.

The work presented herein is unique in that it incorporates multiple datasets with phenotypic, genotypic, and molecular profile information from a cohort of genetically diverse mice. In Chapter 4, I end the exposition of my work with a discussion of how with these data, we have been able to answer some key scientific questions on genetics and regulation in cardiometabolic dysfunction, and develop a basis for asking and answering even more.

CHAPTER 2: SYSTEMS GENETICS IDENTIFIES A CO-REGULATED MODULE OF LIVER MICRORNAS ASSOCIATED WITH PLASMA LDL CHOLESTEROL IN MURINE DIET-INDUCED DYSLIPIDEMIA

2.1 INTRODUCTION

Dyslipidemia, or the state of having chronically altered lipid levels in the blood, is a major risk factor for developing atherosclerosis and cardiovascular disease [Nelson 2013, Zhang et al. 2015]. The liver is the primary organ regulating plasma lipid levels, and dysfunction in certain hepatic processes has been shown to be a main contributor to dyslipidemia [Hojland et al. 2016, Zhang et al. 2015]. Thus, understanding the underlying molecular mechanisms in the liver that cause or respond to dyslipidemia is important for ultimately identifying novel therapeutic targets.

MiRNAs, which are small non-coding RNAs that fine-tune gene expression primarily at the post-transcriptional level, have emerged as key players in many processes, including those involved with lipid homeostasis. Several hepatic miRNAs have been associated with atherosclerosis and hyperlipidemia, including miR-27 [Shirasaki et al. 2013, Vickers et al. 2013], miR-122 [Elman et al. 2008], miR-148a [Goedeke et al 2015], miR-33 [Horie et al. 2012, Rayner et al. 2010, Rayner et al. 2011], and miR-30c [Soh et al. 2013].

A number of these miRNAs, including miR-33 and miR-30c, have been identified as potential therapeutic targets for atherosclerosis and hyperlipidemia [Christopher et al. 2016, Irani and Hussain 2015]. However, there are at least two limitations shared by most of these studies. First, the vast majority of the studies of miRNAs in lipid-related disorders have been performed in C57BL/6 mice, an inbred mouse strain. Although these have produced promising results, the relevance of the findings to genetically diverse outbred populations (like humans) is unclear. Second, all of these studies involve a focused effort to understand an individual miRNA, and its contextually relevant target genes. While this is a reasonable approach, it is not conducive to understanding the roles of miRNAs within a network of other miRNAs and genes. This is an important limitation since there is evidence to support the idea that miRNAs often work in cooperative groups to regulate gene expression [Krek et al. 2005, Lai et al. 2012, Xu et al. 2011].

One approach to addressing both of these limitations is to utilize a systems genetics strategy wherein transcript levels are quantified in tissues of interest, integrated with underlying genetic information, and related to clinical traits of interest. This has been successfully performed in a number of studies focused on gene networks in glucose and lipid metabolism in humans [Plaisier et al. 2009], mice [Yang et al. 2009], and flies [Cermelli et al. 2006]. Subsequent studies have consistently demonstrated that the candidates identified with these approaches are critical mediators of these processes [Fan et al. 2014, Musunuru et al. 2010, Zhang et al. 2016].

A resource that is ideal for such a study is the Diversity Outbred (DO) mouse population. The DO mice were created by strategic outbreeding of eight parental strains of mice, the same ones used to generate the Collaborative Cross (CC) which are distinct lines of mice that are maintained as recombinant inbred strains [Churchill et al. 2004]. While DO mice are similar to CC mice in that they represent mosaics of the eight founder lines, they are different from the CC in that each DO mouse has a unique, non-reproducible genome with dramatically increased levels of accumulated recombination [Churchill et al. 2012]. With each mouse harboring around 45 million variants in its genome, there is a high degree of allelic and phenotypic variation within the population. The extent of genotypic and phenotypic diversity, as well as the high frequency of recombination events, is very useful for identifying genetic contributions to traits of interest with high resolution [Svenson et al. 2012].

In the present study, we utilized a cohort of almost 300 DO mice to interrogate the hepatic network of miRNAs associated with circulating lipid levels in diet-induced dyslipidemia. We identify a key co-regulated module of miRNAs that is strongly associated with LDL cholesterol (LDL-C), which is a significant risk factor for many downstream morbidities, including atherosclerosis and metabolic disease.

2.2 MATERIALS AND METHODS

2.2.1 Animals, Diets, and Phenotyping

Details on the origin, housing, husbandry and treatment of the Diversity Outbred (DO) animals, diet compositions, and measurement of total cholesterol,

triglycerides, glucose, and insulin have been provided previously [Smallwood et al. 2014]. For HDL precipitation, 100µL aliquots of blood plasma samples were diluted 1:4 with PBS, combined with 9 µL of Heparin-$MnCl_2$, and centrifuged. Supernatant was removed and combined with working cholesterol reagent (mixture of reagent, HDCBS, Cholesterol Oxidase, Cholesterol Esterase, and Horseradish Peroxidase) [Puppione and Charugundla 1994]. Samples were run in triplicate on 96-well flat bottom plates, and absorbance was read at 515nm using BioTek plate reader and Gen5 software. Absorbance values were averaged across triplicates, and concentrations were calculated. In order to determine VLDL-C/LDL-C levels, HDL-C was subtracted from total cholesterol. Markers of liver inflammation, alanine aminotransferase (ALT) and aspartate aminotransferase (AST), were measured using a Biolis 24i Analyzer (Carolina Liquid Chemistries, Winston Salem, NC).

2.2.2 RNA Extraction

Livers were flash frozen in liquid nitrogen and subsequently stored at -80 degrees Celsius until their use. Total RNA was isolated by automated instrumentation from approximately 25 milligrams of liver tissue per sample using Norgen Total RNA Purification Kit (Norgen, Ontario, Canada, Catalog No. 24300). Quant-iT^{TM} RiboGreen from ThermoFisher Scientific (Waltham, MA, Catalog No. R11490) was used to measure RNA concentration by fluorometry. RNA integrity was determined by Bioanalyzer from Bio-Rad Technologies. Only samples with RNA

Quality Indicator (RQI) of at least 7.5 or greater were used for microarray and sequencing.

2.2.3 Microarray

High quality RNA was available from livers of 268 of the 292 DO mice and was used for microarray gene expression analysis. The RNA was hybridized to Affymetrix Mouse Gene 2.1 ST 96-Array Plate using the GeneTitan Affymetrix instrument according to standard manufacturer's protocol. Robust Multiarray Average (RMA) method was used to estimate normalized expression levels of transcripts (median polish and sketch-quantile normalization). Affymetrix Expression Console software was used for quality control assessment, and as a result, six of the mice were removed for not passing tests leaving 262 samples to be analyzed. All probes containing known SNPs from the eight founder inbred mouse strains of the DO mouse population were masked (165,204 probes) during normalization by downloading the SNPs from the Sanger sequencing website [http://www.sanger.ac.uk/science/data/mouse-genomes-project], and overlapping them with probe sequences. All control probes (190 probes), reporter probes (82 probes), and normalization probes (6,683 probes) were removed from the probe sets before running WGCNA. Probes were filtered using an expression threshold of a minimum RMA of 4 in at least one-quarter of the samples, which left 15,105 probes. Differential expression analysis was performed by Student's t-test, and p-values were corrected using the Bonferroni method. Microarray data is available on the GEO repository, accession number GSE99561.

2.2.4 Small RNA-sequencing

High quality RNA was available from livers of 269 of 292 of the DO mice was used for small RNA sequencing (smRNA-seq). Libraries were created using New England Biosciences NEBNext Multiplex Small RNA Library Prep Set for Illumina, and 50bp single-read sequencing was carried out on the Illumina HiSeq platform resulting in an average of over 16 million reads per sample. miRquant 2.0 [Kanke et al. 2016] was used to trim off adapter sequences, align reads to the mouse genome, and quantify miRNAs and their isoforms (termed isomiRs). A previous study in Collaborative Cross mice has shown that miRNAs do not contain variants across founder strains within their seed regions, so reads were aligned to the mm9 mouse genome [Rutledge et al. 2015]. Reads were normalized to reads per millions mapped to miRNAs (RPMMMs). An expression threshold of at least 50 RPMMMs in at least one-quarter of all samples was set to filter out the lowly expressed miRNAs, which resulted in a set of 246 robustly expressed miRNAs. The results were consistent when repeated with an expression threshold of at least 100 RPMMMs in at least a quarter of all samples. Hierarchical clustering of the samples' expression profiles was performed based on several dataset characteristics (library prep date, plate, etc.) to ensure there were no batch affects. Differential expression of miRNAs was performed by Student's t-test, and p-values were corrected using the Benjamini-

Hochberg method. Small RNA-seq data is available on the GEO repository, accession number GSE99561.

2.2.5 WGCNA

The WGCNA R package was used to identify the co-regulated modules (CRMs) for miRNAs and genes [Langfelder and Horvath 2008]. Only expression data from the HFCA-fed mouse samples were used to identify the CRMs.

For the miRNA network analysis, we matched the smRNA-seq samples with the mice for which phenotypic data was measured, and were left with 256 DO mice, of which 135 were HFCA-fed mice. The RPMMMs were transformed to log2(x+1) scale. The soft threshold was chosen by running the *pickSoftThreshold* function to find the best fit to a scale-free topology, and beta was set to 14 because it fit with an R^2 value ≥ 0.8, and connectivity measures suggested the possibility of identifying hubs. An adjacency matrix was created using Pearson correlations. From the adjacency matrix, the topological overlap measure (TOM) was calculated using the signed method. The dissimilarity measure was calculated by 1-TOM, and this was used to create a dendrogram according to the Ward's hierarchical clustering method. We use Ward's method instead of the default average method because it considers the variance in expression between miRNAs before choosing to put them in a clade together. Thus, miRNAs within the same clade have the lowest variance in expression possible, which is meaningful for identifying clusters of co-regulated miRNAs. The hybrid tree-cutting algorithm was used to form the modules, which

were left unmerged. Module eigenmiRs (MEms) were calculated using the *moduleEigengenes* function, and were correlated with each phenotype measured in the mice using the biweight midcorrelation. Module significance was also calculated using biweight midcorrelation method. Modules with the highest correlation or inverse correlation (coefficient >|0.4|) were taken as those of interest. miRNAs that were found to have the highest aggregated TOMs within the respective CRMs, the highest kWithin (intramodular connectivity measure) and the highest Pearson correlations to their MEms were identified as hub miRNAs. Network files were exported to Cytoscape [Shannon et al. 2003] for visualization.

For WGCNA on mRNAs, we matched microarray samples with the mice we have measured phenotypes for, and we were left with 249 DO mice, of which 135 were HFCA-fed mice. We parsed out the top 5,000 most variably expressed mRNAs within the HFCA samples from the 15,105. We found that a large number of the mRNAs were not separated into CRMs, but were allocated to an undefined group, which is where genes are assigned when they cannot be placed into any module. We gradually reduced the number of mRNAs from 5,000 to 3,000. Using the top 3,000 most variably expressed mRNAs within the HFCA samples allowed for each gene to be assigned to a specific CRM. Soft threshold beta was set to 9 because it fit a scale-free topology with an R^2 value > 0.8, and connectivity measures suggested the possibility of identifying hub genes. An adjacency matrix was calculated with Pearson correlations, and the TOM was calculated using the signed method. The 1-TOM distance measure was used to create a dendrogram according

to the Ward's method of hierarchical clustering. The hybrid tree-cutting algorithm was used to form the modules, which were left unmerged. Module eigengenes were calculated and correlated with each phenotype measured in the mice using the biweight midcorrelation. The same method was used to calculate correlations between gene CRMs and the brown miRNA CRM. Modules with the highest correlation or inverse correlation (coefficient >|0.4|) were recognized as those of potential interest.

2.2.6 Pathway Enrichment Analysis

Enrichr was used to perform pathway enrichment analysis and gene ontology analysis on the genes within each gene CRM [Chen et al. 2013, Kuleshov et al. 2016].

2.2.7 MicroRNA Target Site Enrichment Analysis

miRhub [Selitsky et al. 2015] was used to perform target site enrichment analysis for miRNAs. Briefly, miRhub employs a Monte Carlo simulation strategy to determine which miRNAs, if any, have an over-representation of predicted target sites at a specified level of conservation in a set of input genes. We ran miRhub on genes up- and down-regulated in the liver from HFCA-fed mice and required positional conservation of predicted target sites in at least two mammalian species in addition to mouse.

2.3 RESULTS

2.3.1 Effects of a high-fat cholic acid, dyslipidemia-inducing diet on plasma lipoprotein cholesterol levels in the diversity outbred mouse population

In a previous study [Smallwood et al. 2014], we demonstrated the value of a specific multi-parental mouse population, the diversity outbred (DO) resource, for mapping quantitative trait loci (QTL) and identifying candidate genes and potential therapeutic targets for dyslipidemia and atherosclerosis. An initial cohort of 292 DO mice comprised of 146 sibling pairs was fed either a high-fat cholic acid containing diet (HFCA), which induces dyslipidemia, or a calorie-matched high protein diet (HP) for eighteen weeks. We identified QTLs for atherosclerotic lesion size, pre-diet circulating triglycerides, and post-diet circulating total cholesterol. In the present study (Figure 2.1), we have analyzed plasma samples from the same mouse cohort for several additional cardio-metabolic endpoints, with a primary focus on Very Low Density Lipoprotein and Low Density Lipoprotein Cholesterol (VLDL/LDL-C), and High Density Lipoprotein Cholesterol (HDL-C) prior to and after the eighteen-week diet exposure. We found that the average levels of alanine aminotransferase (ALT), aspartate aminotransferase (AST), VLDL/LDL-C, but not HDL-C, are significantly elevated in HFCA-fed DO mice relative to the HP-fed DO mice (Figure 2.2A-D).

Notably, all but three of the HFCA-fed mice had higher circulating VLDL/LDL-C levels than the average level among HP-fed mice. Moreover, VLDL/LDL-C levels were highly variable among the HFCA-fed mice, ranging from <10 mg/dL to > 300mg/dL (Figure 2C). This finding indicates that effects of the HFCA diet on plasma VLDL/LDL-C is highly dependent on genetic background.

2.3.2 Effects of the HFCA diet on liver microRNA expression in the DO cohort

Given the importance of the liver in maintaining cholesterol and lipoprotein homeostasis, and the growing appreciation for miRNAs in the control of cholesterol metabolism, we reasoned that hepatic miRNAs may associate with the observed variation in lipid phenotypes, particularly VLDL/LDL-C, across the mice in the DO cohort. To test this hypothesis, we performed small RNA sequencing (smRNA-seq) on liver tissue from 269 of the same DO mice at an average depth of 16 million reads per sample (range 7,662,595 – 30,8861,518 reads). The reads were mapped to the mouse genome (mm9), and miRNAs and their isoforms (referred to as isomiRs) were annotated and quantified using miRquant 2.0 [Kanke et al. 2016]. Detailed information on the mapping statistics are provided in online supplemental tables at http://www.physiology.org/doi/full/10.1152/physiolgenomics.00050.2017.

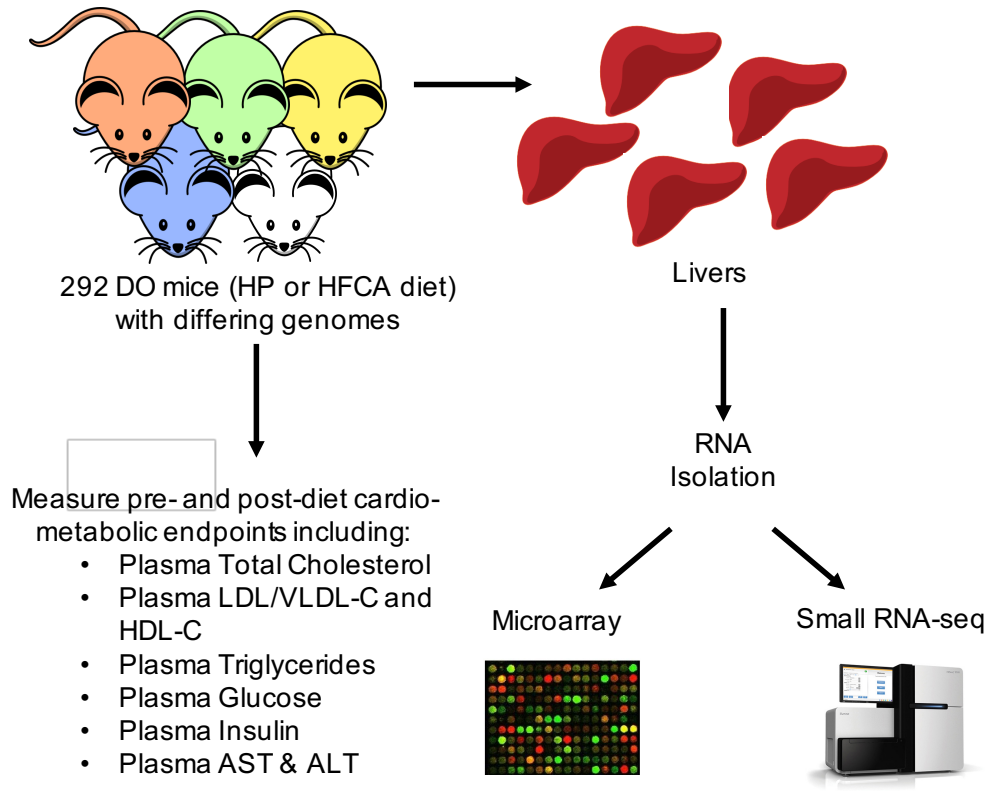

Figure 2.1 Summary Diagram of Study Design. 292 DO mice, each having a different composite of the 8 founder mouse strain genomes, were fed either HP or HFCA diet. Cardio-metabolic endpoints were measured before and after diet intervention. RNA was isolated from the livers of each of the mice, and used for microarray analysis to measure gene expression, and small RNA sequencing.

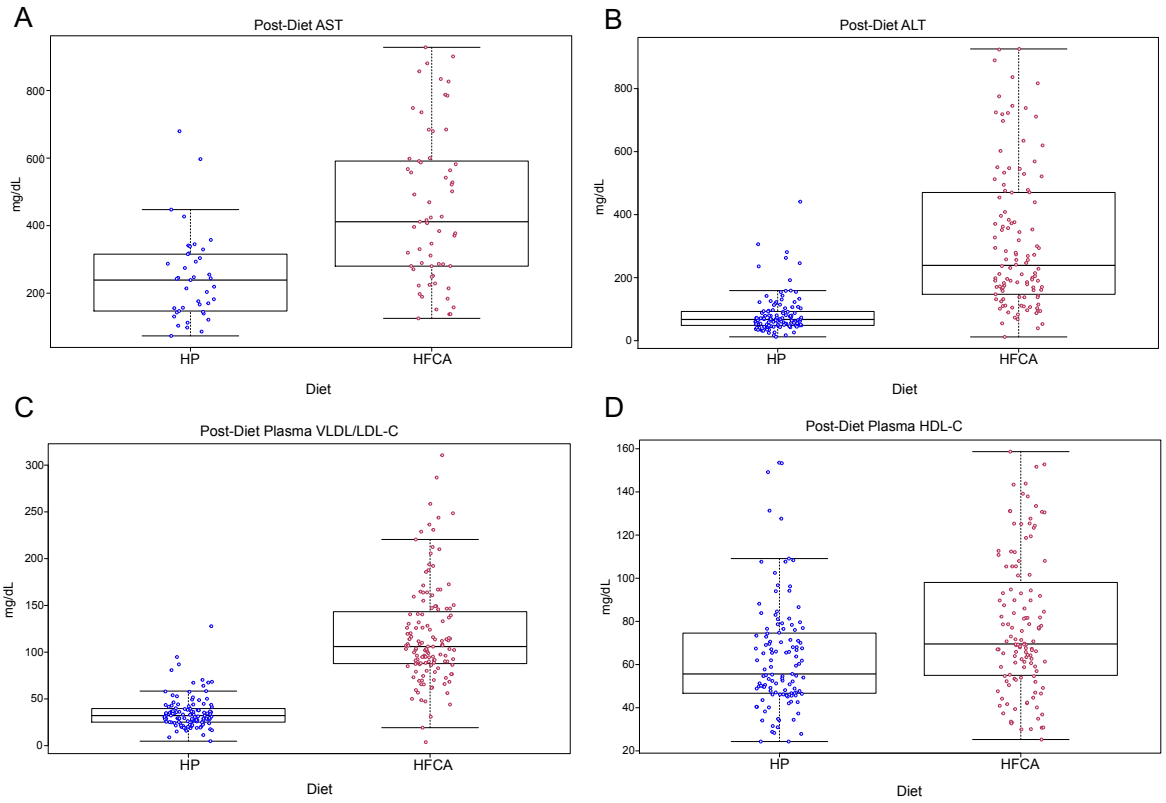

Figure 2.2: Plasma vLDL/LDL, AST and ALT are dramatically affected by HFCA diet. Box plot of **(A)** post-diet AST (mg/dL), **(B)** post-diet ALT (mg/dL), **(C)** post-diet plasma VLDL/LDL-C (mg/dL), and **(D)** post-diet plasma HDL-C (mg/dL) concentration in HP-fed DO mice and HFCA-fed DO mice. Each dot represents one mouse in the respective diet. Hinges of boxplots represent the first and third quartile of expression.

To normalize miRNA expression, we used the reads per millions mapped to miRNAs (RPMMM) method. After filtering out those miRNAs with low levels of expression across the majority of samples, 246 miRNAs remained (Methods). More than one-third of these were significantly elevated (fold-change ≥ 1.5, FDR < 0.05) in the livers of the mice fed the dyslipidemic HFCA diet relative to those fed the HP diet (Figure 2.3). Specifically, 45 miRNAs were significantly up-regulated, and 39 miRNAs were significantly down-regulated in the HFCA-fed mice. These differentially expressed miRNAs included several that have previously been implicated in the development and/or progression of atherosclerosis or hyperlipidemia. One such miRNA is miR-34a, which is up-regulated in the plasma of ApoE knockout (KO) C57BL/6 mice [Han et al. 2015], a well-established animal model of atherosclerosis; in the liver tissue of high-fat diet-fed C57BL/6 mice [Ding et al. 2015, Fu et al. 2012], a model of hepatic steatosis and obesity; as well as in the plaques of humans with coronary artery disease [Raitoharju et al. 2011]. We found that liver miR-34a expression levels in most HFCA diet-fed mice were significantly greater than in the HP diet-fed mice (Figure 2.4A). Notably, miR-34a expression in all but two (98.7%) of the HFCA samples was above the average expression in the HP-fed mice. In addition, analysis of miR-34a expression within the sibling pairs shows the majority of the pairs follow the same trend of an increase in miR-34a expression as a result of HFCA feeding (Figure 2.4B). We also identified other miRNAs significantly altered by HFCA diet, including several that have not been previously associated with regulation of lipid levels, such as miR-874, which was

down-regulated by almost 2-fold, and miR-1247-5p, which was up-regulated by >4-fold (Figure 2.4C, D).

The HFCA diet did not universally alter expression for every miRNA, including miR-30c, which has been shown to be significantly down-regulated in the liver of the ApoE C57BL/6 KO model. Over-expression of miR-30c mimic can mitigate hyperlipidemia and regress atherosclerosis in both wild-type and ApoE KO C57BL/6 mice [Soh et al. 2013]. However, in our DO cohort, while there is an expected trend toward lower liver miR-30c expression in the HFCA-fed mice, the difference is not substantial and the expression distributions among HFCA- and HP-fed mice are largely overlapping (Figure 4E). The response of the miR-30c levels to the HFCA diet is much more mixed than what is seen for miR-34a (Figure 4F). These data suggest that while the dyslipidemia-inducing diet has a robust effect on liver miR-34a, the effect on miR-30c may be more dependent on genetic composition than diet. The extensive variation in hepatic expression that we observed for miRNAs across the DO mice that were fed the same HFCA diet is likely due to interactions between the underlying genetics and diet, and it provided a unique opportunity to identify groups or modules of miRNAs that exhibit highly similar genotype-dependent responses to the HFCA diet.

2.3.3 Identification of co-regulated microRNA modules and correlation with lipid phenotypes

We next sought to identify groups, or co-regulated modules (CRMs), of hepatic miRNAs across the HFCA-fed DO mice. We applied the weighted gene co-expression network analysis (WGCNA) [Langfelder and Horvath 2008] to the 246 miRNAs robustly expressed in the liver (Methods). This analysis identified five main miRNA CRMs (mCRMs) with anywhere from 32 to 60 miRNAs in each CRM (Figure 2.5A). We then used the biweight midcorrelation (bicor) analysis to assess the extent of correlation between each of these modules and the various endpoints that were measured in these mice (Methods). We found that one particular module, which we will refer to as the brown module, is comprised of 34 miRNAs/isomiRs, and is highly correlated with important endpoints, most notably post-diet circulating VLDL/LDL-C (bicor coefficient 0.49) (Figure 2.5B). Furthermore, all of the top 20 correlations between any of the 246 miRNAs and any of the endpoints involve miRNAs from the brown module, and almost all are associated with the diet-induced change in plasma VLDL/LDL-C (Table 2.1).

As a comparison, we performed the WGCNA analysis with the HP-fed mice using the liver expression data for the same 246 miRNAs. Of the five mCRMs that were identified, none were strongly correlated with any of the cardio-metabolic endpoints that were measured in the mice (Figure 2.5C). Also, none of the mCRMs in the HP-only analysis exhibited substantial overlap with any of the mCRMs identified in the HFCA analysis, with an average of only ~8 shared miRNAs (which

represents an average of just ~20% shared between any one HP mCRM and any one HFCA mCRM). This indicates that the HFCA diet leads to a very specific, robust re-wiring of the regulatory networks governing miRNA expression.

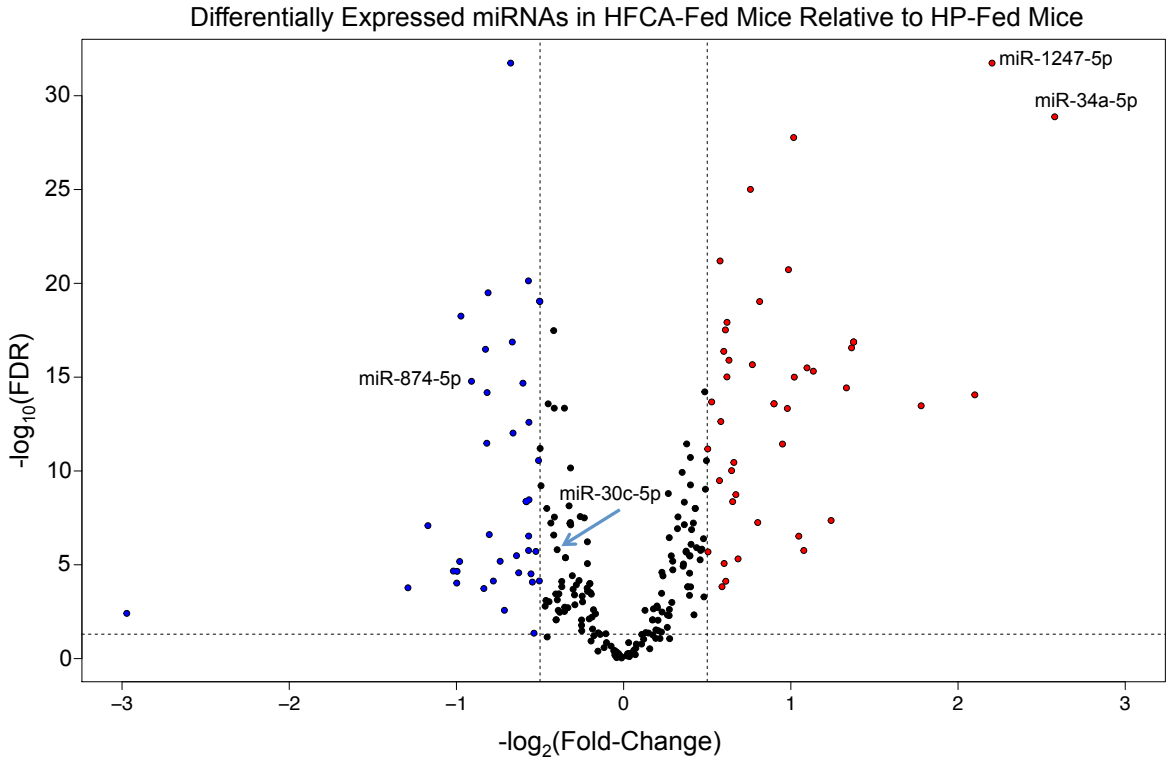

Figure 2.3: Diet alters miRNA expression. Volcano plot of differentially expressed liver miRNAs between HFCA-fed and HP-fed DO mice after filtering out those that were lowly expressed. Each dot represents one miRNA. Red dots are miRNAs that are up-regulated in HFCA-fed mice relative to HP-fed mice with a fold-change of 1.5 or more and an FDR ≤ 0.05. Blue dots are miRNAs that are down-regulated in HFCA-fed mice relative to HP-fed mice with a fold-change of 1.5 or more and an FDR ≤ 0.05. Horizontal dashed line denotes FDR = 0.05. Vertical dashed lines denote fold change of -1.5 (left) and 1.5 (right).

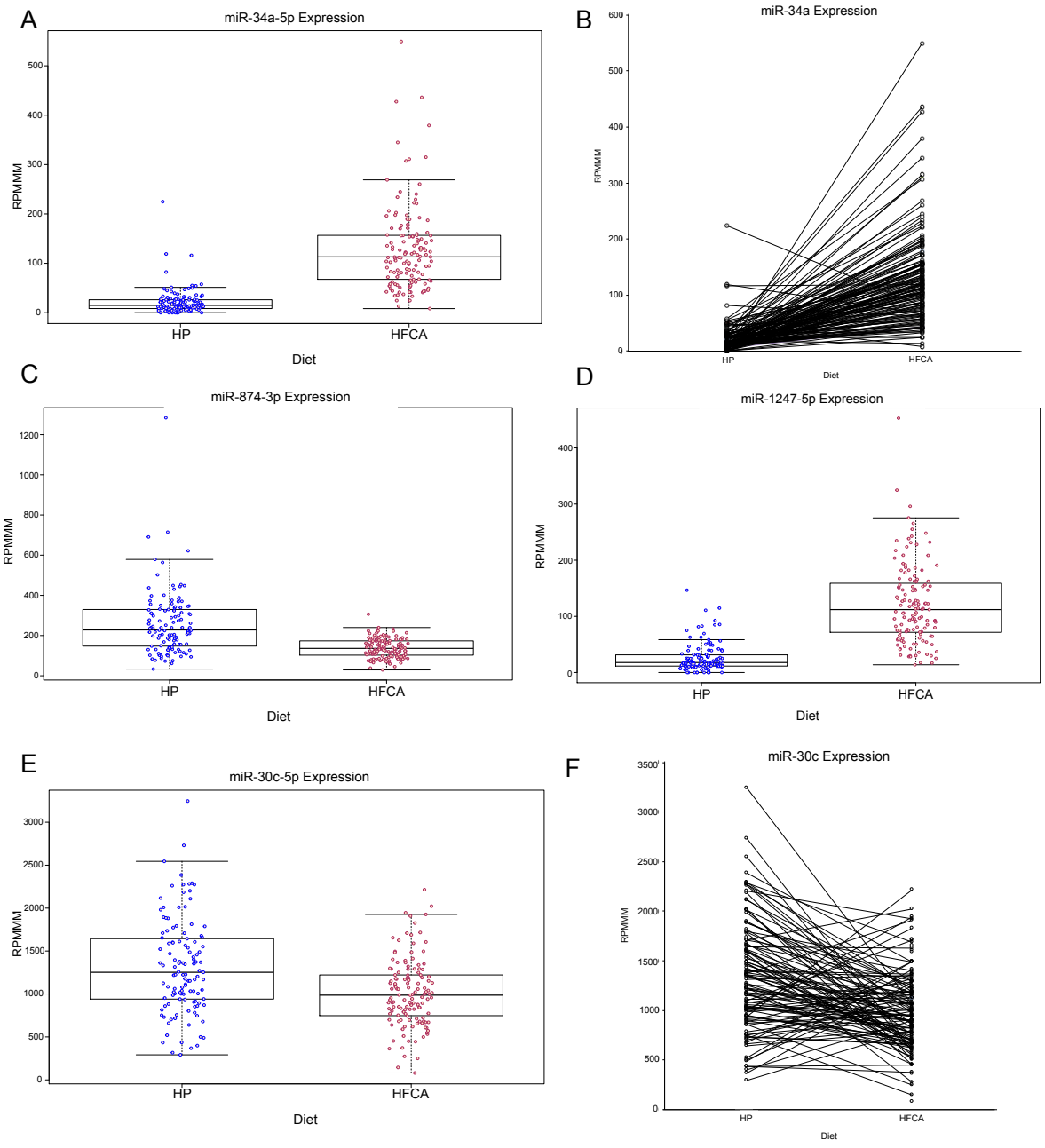

Figure 2.4: Diet alters expression of specific miRNAs. (A) Box plot of miR-34a-5p. **(B)** miR-34a-5p expression with lines connecting DO mouse sibling pairs in either diet group. **(C-E)** Box plot of miR-874-3p, miR-1247-5p, and miR-30c-2-5p expression in HP-fed and HFCA-fed DO mice. For all box plots, each dot represents one mouse in the respective diet. Hinges of boxplots represent the first and third quartile of expression. **(F)** miR-30c-2-5p expression with lines connecting DO mouse sibling pairs in either diet group.

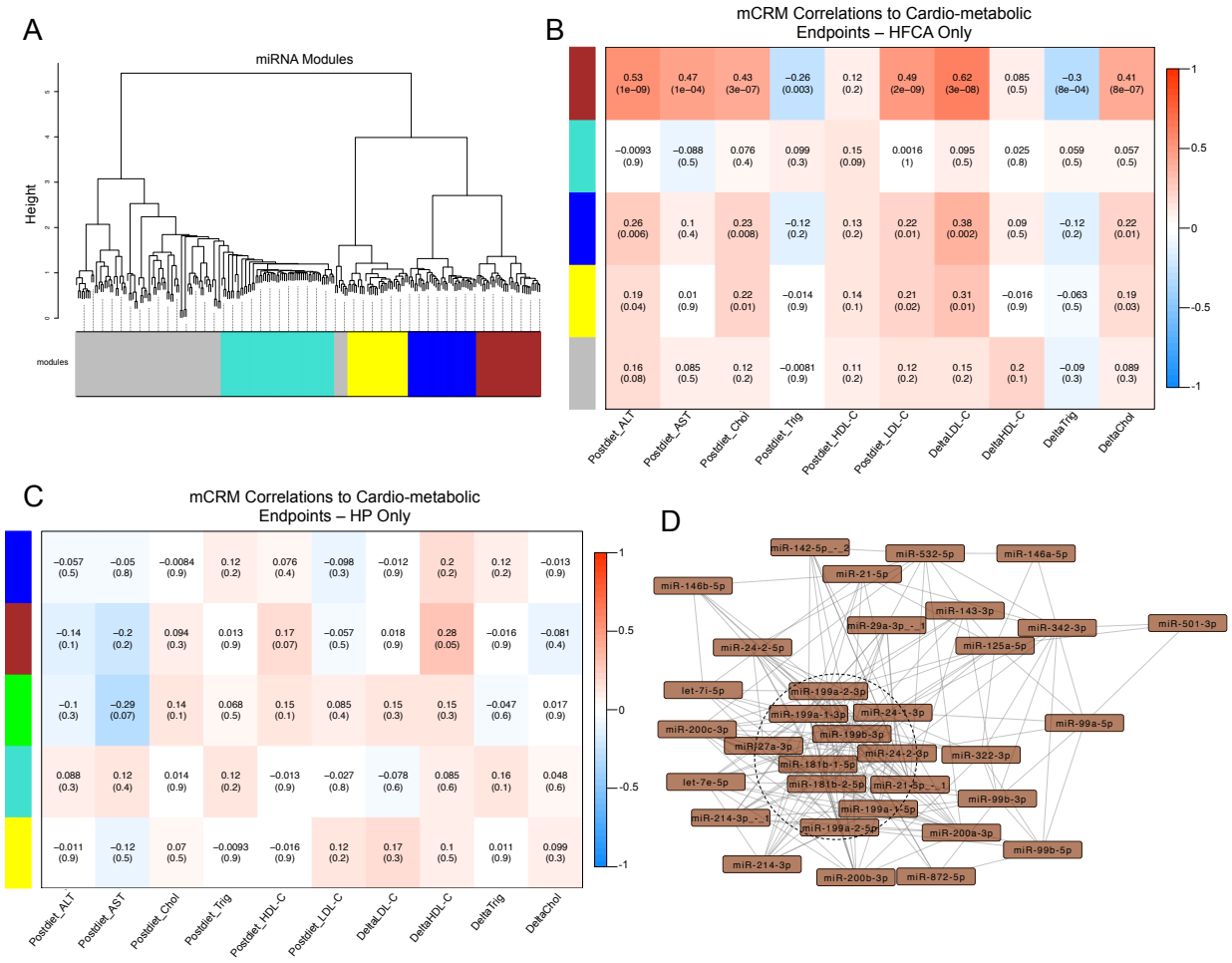

Figure 2.5: MiRNA co-expression analysis identifies module associated with metabolic traits. (A) miRNA modules formed using WGCNA. Only HFCA mice were used during analysis. RPMMMs were converted using log2(x+1) and used to calculate Pearson correlations. Dendrogram was created using the 1-TOM, and Ward's method of hierarchical clustering. Modules were formed using the hybrid tree-cutting function in the WGCNA software package. **(B,C)** Heatmap of miRNA module eigenmiRs correlated to cardio-metabolic endpoints measured in the (B) HFCA-fed DO mice and (C) HP-fed DO mice. EigenmiRs were calculated using the WGCNA function *moduleEigengenes*, and correlated using the biweight midcorrelation to normalized endpoint values. The intensity of orange or blue denotes how close the correlation coefficient is to 1 or -1, respectively. Top numbers are biweight midcorrelation coefficients, bottom numbers are p-values. **(D)** Cytoscape visualization of brown mCRM. Each node represents one miRNA. Each edge represents high co-correlation. The dashed circle highlights the hub miRNAs in this module as determined by number of connections and weight of connections.

2.3.4 Identification of microRNA "hubs" in the brown module correlated with post-diet plasma VLDL/LDL-C

The strength of connection between miRNAs in an mCRM is indicative of the extent of co-correlation. The miRNAs with the highest connectivity scores are defined as "hubs." To define the miRNA "hubs" in the brown mCRM, we ranked the miRNAs according to their topological overlap measure (TOM) scores and intramodular connectivity measures, which are the WGCNA metrics for interconnectedness. The miRNAs in the 75^{th} percentile were identified as "hubs" in the brown mCRM: miR-199a, miR-181b, miR-27a, miR-21_-_1, and miR-24. Each of these miRNAs was very highly correlated with the module eigenmiR, or first principal component, of the brown module (Figure 2.5D, Table 2.2), with miR-199a being the most highly correlated (Pearson coefficient = 0.93).

2.3.5 Analysis of gene expression data and identification of several gene co-regulated modules associated with post-diet LDL-C and inversely correlated with the brown microRNA module

To determine the effects of the HFCA diet on gene expression, we performed microarray analysis on 262 liver samples from the DO mice (34,390 mRNAs detected, 15,105 mRNAs with RMA ≥ 4 in at least one-quarter samples). Differential expression analysis revealed that 4236 genes were significantly (corrected p-value < 1.20×10^{-6}) up-regulated, of which 401 exhibited a fold-change greater than 2, and 3603 genes significantly (corrected p-value < 1.20×10^{-6}) down-regulated, of which

140 exhibited a fold-change less than -2, in the liver tissue from HFCA-fed mice compared to HP-fed mice (Figure 2.6A). We found that lipid processing and metabolism genes such as lipoprotein lipase (*Lpl*) are up-regulated (+14.8-fold), whereas cholesterol biosynthesis genes such as squalene epoxidase (*Sqle*) are down-regulated (-19.5-fold). These results are expected in response to HFCA, which leads to a dramatic increase in dietary lipid and cholesterol, eliciting a suppression of endogenous lipid/cholesterol synthesis and increase in lipid metabolic activity.

Using the tool miRhub we demonstrated that the 3603 genes significantly down-regulated in the liver from HFCA-fed mice are significantly enriched for predicted conserved target sites for the brown module hub miRNAs miR-27a (empirical p-value = 0.008), miR-199a (p=0.015), miR-24 (p=0.020), and miR-181b (p=0.045), each of which is significantly up-regulated in the liver in response to HFCA. This finding held for miR-27a (p=0.017) upon miRhub analysis of the subset of significantly down-regulated genes that are altered by more than 2-fold (n=140). We and others have shown previously that the miR-27 family is involved in the control of lipid balance in part through regulation of genes in the lipid synthesis and uptake pathways. We confirmed that the previously validated targets of miR-27, *Hmgcr* [Selitsky et al. 2015] and *Ldlr* [Alvarez et al. 2015], which encode proteins critical for cholesterol biosynthesis and LDL-C uptake, respectively, are indeed among the 140 genes significantly down-regulated in the liver from HFCA-fed mice, and are significantly inversely correlated with miR-27a levels (Figure 2.6B). We additionally identified a predicted conserved target site for miR-27 in the 3' UTRs of

Acly and *Lpin1*, each of which is reduced in the liver by >2-fold in HFCA-fed mice and is inversely correlated with miR-27a levels (Figure 6C). Acly catalyzes an early step in fatty acid synthesis and Lpin1 mediates one of the final steps in triglyceride biosynthesis in the liver. Analysis of published Argonaute (Ago) CLIP-seq data from human Huh7 hepatoma cells provided strong experimental support for a regulatory interaction between miR-27 and *Acly*, and to a slightly lesser extent, miR-27 and *Lpin1* [Luna et al. 2015], thereby adding to the evidence that miR-27 is a critical diet-responsive regulator of lipid homeostasis.

As we did with miRNA expression data, we next analyzed the gene expression data via WGCNA. Using the top 3,000 most variable genes within the HFCA-fed mice, we identified twenty groups, or gene CRMs (gCRMs), each of which contains highly co-correlated mRNAs that are likely co-regulated (Figure 2.7A). Among these there are two gCRMs, pink and midnight blue, that are relatively highly correlated with post-diet plasma LDL-C (pink bicor coefficient = 0.5, midnight blue bicor coefficient = 0.58), and three gCRMs, magenta, tan and turquoise, which are strongly inversely correlated with post-diet LDL-C (bicor coefficients -0.56, -0.56, and -0.51, respectively) (Figure 2.7B). As expected, individual miRNAs from the brown mCRM (e.g., miR-199a) exhibited strong inverse correlation with the genes in the magenta, light cyan, tan, and turquoise gCRMs, whereas miRNAs in other mCRMs did not (e.g., miR-151), suggestive of a unique association in diet-induced dyslipidemia between miRNAs in the brown mCRM and genes in these four gCRMs

(Figure 2.7C). Genes in these gCRMs that are candidate targets of one or more of the miRNAs in the brown mCRM are listed in Table A.1.[1]

Table 2.1: Top correlations between miRNAs and cardio-metabolic endpoints. Biweight midcorrelation coefficients of miRNAs as they are related to the cardio-metabolic endpoints. Only correlation values ≥ 0.45 and ≤ -0.45 are shown.

miRNA	Associated Phenotype	Cor_coef	p-value
mmu-mir-24-2-5p	Delta VLDL/LDL-C	0.66	2.88E-09
mmu-mir-24-1-3p	Delta VLDL/LDL-C	0.63	2.48E-08
mmu-mir-24-2-3p	Delta VLDL/LDL-C	0.63	2.48E-08
mmu-let-7i-5p	Delta VLDL/LDL-C	0.61	5.35E-08
mmu-let-7i-5p	Post-diet ALT	0.61	5.61E-13
mmu-mir-29a-3p_-_1	Delta VLDL/LDL-C	0.60	9.75E-08
mmu-let-7i-5p	Post-diet AST	0.59	7.19E-07
mmu-mir-181b-1-5p	Delta VLDL/LDL-C	0.58	3.95E-07
mmu-mir-214-3p_-_1	Delta VLDL/LDL-C	0.58	4.18E-07
mmu-mir-21-5p	Delta VLDL/LDL-C	0.58	4.19E-07
mmu-mir-501-3p	Post-diet ALT	0.58	1.42E-11
mmu-mir-143-3p	Delta VLDL/LDL-C	0.58	5.18E-07
mmu-mir-99b-3p	Delta VLDL/LDL-C	0.58	5.32E-07
mmu-mir-199a-2-3p	Delta VLDL/LDL-C	0.57	5.94E-07
mmu-mir-199a-1-3p	Delta VLDL/LDL-C	0.57	6.22E-07
mmu-mir-199b-3p	Delta VLDL/LDL-C	0.57	6.22E-07
mmu-mir-181b-2-5p	Delta VLDL/LDL-C	0.56	9.75E-07
mmu-mir-99b-5p	Delta VLDL/LDL-C	0.56	1.02E-06
mmu-let-7e-5p	Delta VLDL/LDL-C	0.56	1.04E-06
mmu-mir-21-5p_-_1	Delta VLDL/LDL-C	0.56	1.33E-06
mmu-mir-24-2-5p	Post-diet LDL-C	0.55	6.37E-12
mmu-mir-214-3p	Delta VLDL/LDL-C	0.54	2.77E-06
mmu-mir-21-5p	Post-diet ALT	0.54	5.40E-10
mmu-mir-146b-5p	Post-diet AST	0.54	7.10E-06
mmu-mir-146b-5p	Post-diet ALT	0.54	5.85E-10
mmu-mir-342-3p	Post-diet ALT	0.54	6.13E-10
mmu-mir-143-3p	Post-diet ALT	0.54	7.01E-10
mmu-mir-146a-5p	Post-diet ALT	0.53	1.49E-09
mmu-mir-21-5p	Post-diet LDL-C	0.52	1.07E-10
mmu-mir-501-3p	Post-diet AST	0.52	1.67E-05
mmu-mir-27a-3p	Delta VLDL/LDL-C	0.52	9.24E-06
mmu-mir-24-2-5p	Post-diet ALT	0.52	4.08E-09
mmu-mir-146a-5p	Delta VLDL/LDL-C	0.52	1.12E-05
mmu-mir-99a-5p	Delta VLDL/LDL-C	0.51	1.14E-05
mmu-mir-191-5p_+_1	Post-diet ALT	0.51	4.83E-09
mmu-let-7e-5p	Post-diet AST	0.51	2.56E-05
mmu-mir-1839-5p	Post-diet ALT	0.51	7.11E-09
mmu-mir-1839-5p_+_1	Post-diet ALT	0.51	8.68E-09
mmu-mir-199a-1-5p	Delta VLDL/LDL-C	0.50	1.93E-05
mmu-mir-199a-2-5p	Delta VLDL/LDL-C	0.50	1.94E-05

[1] Oversized tables and figures are presented in Appendix A.

mmu-mir-532-5p	Post-diet ALT	0.50	1.27E-08
mmu-mir-191-5p	Delta VLDL/LDL-C	0.50	2.39E-05
mmu-mir-322-3p	Post-diet ALT	0.50	2.12E-08
mmu-mir-21-5p	Post-diet AST	0.49	5.07E-05
mmu-let-7i-5p	Post-diet LDL-C	0.49	1.65E-09
mmu-mir-140-5p	Delta VLDL/LDL-C	0.49	3.78E-05
mmu-let-7e-5p	Post-diet ALT	0.49	3.95E-08
mmu-mir-191-5p_+_1	Delta VLDL/LDL-C	0.49	4.09E-05
mmu-mir-322-3p	Delta VLDL/LDL-C	0.48	4.28E-05
mmu-mir-425-5p	Delta VLDL/LDL-C	0.48	4.77E-05
mmu-mir-200b-3p	Post-diet AST	0.48	8.57E-05
mmu-mir-10a-5p_+_1	Post-diet ALT	0.48	7.24E-08
mmu-mir-143-3p	Post-diet LDL-C	0.48	7.54E-09
mmu-mir-674-3p	Delta VLDL/LDL-C	0.48	6.14E-05
mmu-mir-21-5p_-_1	Post-diet AST	0.47	0.0001225
mmu-mir-99b-5p	Post-diet ALT	0.47	1.43E-07
mmu-mir-21-5p_-_1	Post-diet ALT	0.47	1.43E-07
mmu-mir-143-3p	Post-diet AST	0.47	0.00015041
mmu-mir-146b-5p	Delta VLDL/LDL-C	0.47	8.95E-05
mmu-mir-21-5p_-_1	Post-diet LDL-C	0.46	2.32E-08
mmu-mir-10a-5p_+_1	Delta VLDL/LDL-C	0.46	0.00010479
mmu-mir-186-5p	Delta VLDL/LDL-C	0.46	0.00010562
mmu-mir-10a-5p	Post-diet AST	0.46	0.00017707
mmu-mir-24-2-5p	Delta Cholesterol	0.46	2.84E-08
mmu-mir-24-2-5p	Post-diet Cholesterol	0.46	3.15E-08
mmu-mir-532-5p	Post-diet AST	0.45	0.00023692
mmu-mir-140-3p_+_1	Delta VLDL/LDL-C	0.45	0.00014825
mmu-mir-122-5p	Post-diet ALT	-0.45	4.22E-07
mmu-mir-122-5p	Delta VLDL/LDL-C	-0.56	1.24E-06

2.4 DISCUSSION

To our knowledge, this is the first large-scale study aimed at using a genetically diverse mouse population to identify hepatic miRNAs associated with a variety of different cardio-metabolic endpoints pertinent to diet-induced dyslipidemia and metabolic dysfunction. Our study utilizes a population of outbred mice to delineate the miRNA regulatory networks associated with diet-induced metabolic dysfunction. This systems-level approach yielded at least four novel findings. Firstly, we determined that certain metabolic phenotypes in DO mice, notably plasma

VLDL/LDL-C, are far more variable in response to a dyslipidemia-inducing diet than other phenotypes. Secondly, we identified specific miRNAs (e.g., miR-34a) that exhibit a dramatic response to the dyslipidemic diet irrespective of host genotype; these are miRNAs for which the diet had the most dominant effect. Thirdly, we showed that some miRNAs (e.g., miR-30c), though previously identified as dramatically altered in diet-induced dyslipidemia, were not altered by HFCA in most DO mice, indicating the prominent contribution of genetics to the responses of certain liver miRNAs to diet. Fourthly, we identified a co-regulated module of miRNAs that is strongly associated with circulating levels of VLDL/LDL-C – a significant risk factor for atherosclerosis and other related cardiovascular conditions. Moreover, we predicted that miRNAs in this module together target about one-third of the genes that are inversely correlated with VLDL/LDL-C.

We found that some miRNAs, such as miR-34a, are altered by high-fat/high-cholesterol diet across most strains in a manner that is consistent with expectation based on previous studies in C57BL/6 mice. However, others such as miR-30c, exhibited much more genotype-dependent behaviors, indicating that previous implications of miR-30c as a potential therapeutic in dyslipidemia and atherosclerosis based on studies in C57BL/6 mice are not necessarily generalizable to different genetic backgrounds. This finding underscores the importance and value of performing studies across diverse genetic backgrounds in order to more broadly define the variability in miRNA responses to diet or other perturbations. Multi-

parental resources such as the DO and CC appear to be especially promising in this regard.

It is worth noting that another prominent atherosclerosis-related miRNA, miR-33 [Horie et al. 2012, Marquart et al. 2010, Najafi-Shoushtari et al. 2010, Rayner et al. 2010, Rayner et al. 2011a, Rayner et al. 2011b], was not included in the set of miRNAs considered in this analysis because the levels at which it was detected did not reach our threshold for robust expression. This may be because the library preparation protocol we used in this study is biased against the detection of some miRNAs such as miR-33 (most likely due to adapter ligation bias, as has been reported previously [Baran-Gale et al. 2015]. Due to the importance of miR-33 in lipid biology and atherosclerosis, it may be worth re-analyzing these samples in the future with alternate detection methods such as RT-qPCR as well as the Bioo Scientific NextFlex V3 library preparation protocol, which is intended to mitigate adapter ligation bias [Baran-Gale et al. 2015].

To identify those mCRMs that are most affected by HFCA diet interactions with genetic composition, we used only the HFCA-fed mouse samples in the WGCNA process. Since all of these mice were fed the same impactful HFCA diet, and each harbored a distinct genome, we reasoned that the use of just these samples would best accentuate the CRMs affected most by genetics and the dyslipidemic diet. Indeed, running the analysis with just the HP-fed mouse samples resulted in mCRMs that are poorly correlated with the cardio-metabolic endpoints. Furthermore, the compositions of the HP mCRMs were very different from the HFCA

mCRMs, with the mean overlap between modules being approximately 8. Taken together, these data demonstrate that HFCA leads to a rearrangement of the regulatory networks controlling miRNA expression.

In order to identify hub members of gene and miRNA CRMs, we calculated the total TOM for each member within the respective module, ranked them and found those with the top aggregated TOMs and intramodular connectivity measures. A hub's expression is believed to influence the expression of the other members within their module, so their TOMs and scaled intramodular connectivities are relatively high because of their strong interconnectedness to the other members in the CRM. Hubs are also most similar to the eigenvalues of the modules itself, so correlations to the eigenmiRs and eigengenes should be very high. In regards to the brown mCRM, we identified miR-199a, miR-181b, miR-27a, miR-21_-_1, and miR-24 as hubs due to their high TOMs and scaled kWithin, and their high correlations with the brown mCRM eigenmiR. In this context, these are miRNAs that are purported to have the most regulatory influence in the module. miRNA hubs may modulate the expression of regulatory genes, such as those encoding transcription and/or biogenesis factors, which subsequently effect the expression and/or stability of other members of the module.

Several of the hub miRNAs we identified have been studied in the context of hyperlipidemia and related metabolic diseases. miR-27 is involved in the regulation of lipid synthesis and metabolic pathways [Vickers et al. 2013], and has been implicated in the etiology of viral hepatitis-induced steatosis [Singaravelu et al. 2014]

and atherosclerosis [Chen et al. 2012]. Although miR-21_-_1 (a 5'-shifted isomiR of miR-21) is not well-studied, the canonical miR-21 is more highly expressed in the liver tissue of the DO mice, and is also a member of the brown mCRM. Notably, hepatic miR-21 is a key partner of miR-27 in the negative regulation of factors mediating cholesterol synthesis [Selitsky et al. 2015] and lipid metabolism [Kida et al. 2011], and was very recently proposed as a driver of metabolic disorders associated with diet-induced obesity as well as fatty liver disease [Calo et al. 2016]. Our findings underscore the importance of these three miRNAs in contributing to diet-induced dyslipidemia. miR-24 also has been connected previously to the development of hyperlipidemia. It is known to be up-regulated in the livers of C57BL/6 mice fed a high-fat diet, and to directly target *Insig1* which leads to hepatic lipid accumulation and hyperlipidemia [Ng et al. 2014]. In the DO mice, miR-24 is up-regulated by ~1.4-fold in the HFCA-fed mice relative to the HP-fed mice. In our gene expression microarray data, *Insig1* is down-regulated by about 2.4-fold. Hepatic miR-199a is not as well-studied in the context of dyslipidemia; however, it has been linked to a lipid-related condition as it was shown to be elevated in livers of humans with non-alcoholic fatty liver disease (NAFLD) [Li et al. 2014]. Hepatic expression of miR-181b in dyslipidemia is even less studied as it is more known for its role in liver fibrosis [Yu et al. 2016], and hepatocarcinogenesis in mice [Wang et al. 2010] and rats [Furtado et al. 2017].

We found that the brown mCRM is inversely correlated with four of the gCRMs. Moreover, we observed that for each of the gCRMs approximately 30-40%

of their genes are predicted to harbor conserved target sites for one or more brown mCRM miRNAs. The brown mCRM members may function within a larger regulatory network to cooperatively regulate gene expression pertinent to the control of circulating VLDL/LDL-C levels.

The brown mCRM and four inversely-correlated gCRMs are associated with post-diet circulating VLDL/LDL-C levels, AST and ALT levels in the HFCA-fed DO mice, but not strongly associated with any of the other endpoints. It is noteworthy that the same mCRM was not associated with endpoints such as atherosclerotic lesion size. One reason for this seemingly discordant result is the fact that long-term feeding of HFCA diet consistently results in early atheroma formation where lesions consist primarily of macrophages and immune cells [Paigen et al. 1987, Qiao et al. 1994]. Furthermore, we found DO mice to be generally resistant to atherosclerosis on the HFCA diet, but all were hyperlipidemic [Smallwood et al. 2014]. This indicated that factors other than hepatic gene expression and hyperlipidemia, perhaps those operating at the vessel wall itself, may be responsible for inhibition of lesion formation.

In utilizing the DO mice in a systems genetics study, we have identified not only miRNAs that may potentially function cooperatively within a network of other miRNAs and genes, but also hub miRNAs that may contribute substantially to the observed increase in VLDL/LDL-C levels or may contribute to the liver's adaptive response to HFCA. These miRNAs are prime candidates for future loss- and gain-of-function studies in diverse genetic strains in the context of diet-induced dyslipidemia.

Table 2.2: miRNAs with the highest interconnectedness are identified as hubs. Members of the brown mCRM ranked according to intramodular connectivity measure (kWithin). Column 2 lists kWithin scaled by the module. Column 3 lists Pearson correlations to the module eigenmiR, and the last column lists the total summed TOM for each miRNA.

Brown mCRM miRNAs	kWithin	kWithin (Scaled)	Cor to MEbrown	Summed TOM
mmu-mir-199a-2-3p	7.0038	1.0000	0.93	8.0038
mmu-mir-199a-1-3p	6.9947	0.9987	0.93	7.9947
mmu-mir-199b-3p	6.9947	0.9987	0.93	7.9947
mmu-mir-181b-1-5p	6.2731	0.8957	0.92	7.2731
mmu-mir-181b-2-5p	6.1845	0.8830	0.92	7.1845
mmu-mir-199a-2-5p	6.0431	0.8628	0.92	7.0431
mmu-mir-199a-1-5p	6.0431	0.8628	0.92	7.0431
mmu-mir-24-2-3p	6.0018	0.8569	0.89	7.0018
mmu-mir-24-1-3p	6.0018	0.8569	0.89	7.0018
mmu-mir-27a-3p	5.0559	0.7219	0.88	6.0559
mmu-mir-21-5p_-_1	5.0528	0.7214	0.89	6.0528
mmu-let-7e-5p	4.9871	0.7121	0.88	5.9871
mmu-mir-200a-3p	4.8912	0.6984	0.86	5.8912
mmu-mir-214-3p_-_1	4.8437	0.6916	0.86	5.8437
mmu-mir-214-3p	4.6433	0.6630	0.84	5.6433
mmu-let-7i-5p	4.4421	0.6342	0.86	5.4421
mmu-mir-200c-3p	4.2868	0.6121	0.81	5.2868
mmu-mir-200b-3p	4.2822	0.6114	0.81	5.2822
mmu-mir-142-5p_-_2	4.2114	0.6013	0.82	5.2114
mmu-mir-24-2-5p	4.2033	0.6001	0.86	5.2033
mmu-mir-146b-5p	3.9551	0.5647	0.84	4.9551
mmu-mir-21-5p	3.8489	0.5495	0.81	4.8489
mmu-mir-29a-3p_-_1	3.7823	0.5400	0.80	4.7823
mmu-mir-872-5p	3.6169	0.5164	0.78	4.6169
mmu-mir-99b-3p	3.4688	0.4953	0.77	4.4688
mmu-mir-99b-5p	3.4553	0.4933	0.76	4.4553
mmu-mir-146a-5p	3.3812	0.4828	0.78	4.3812
mmu-mir-143-3p	3.3189	0.4739	0.78	4.3189
mmu-mir-125a-5p	3.1855	0.4548	0.69	4.1855
mmu-mir-99a-5p	3.1538	0.4503	0.74	4.1538
mmu-mir-342-3p	3.0082	0.4295	0.72	4.0082
mmu-mir-322-3p	3.0044	0.4290	0.75	4.0044
mmu-mir-501-3p	2.9523	0.4215	0.73	3.9523
mmu-mir-532-5p	1.6634	0.2375	0.53	2.6634

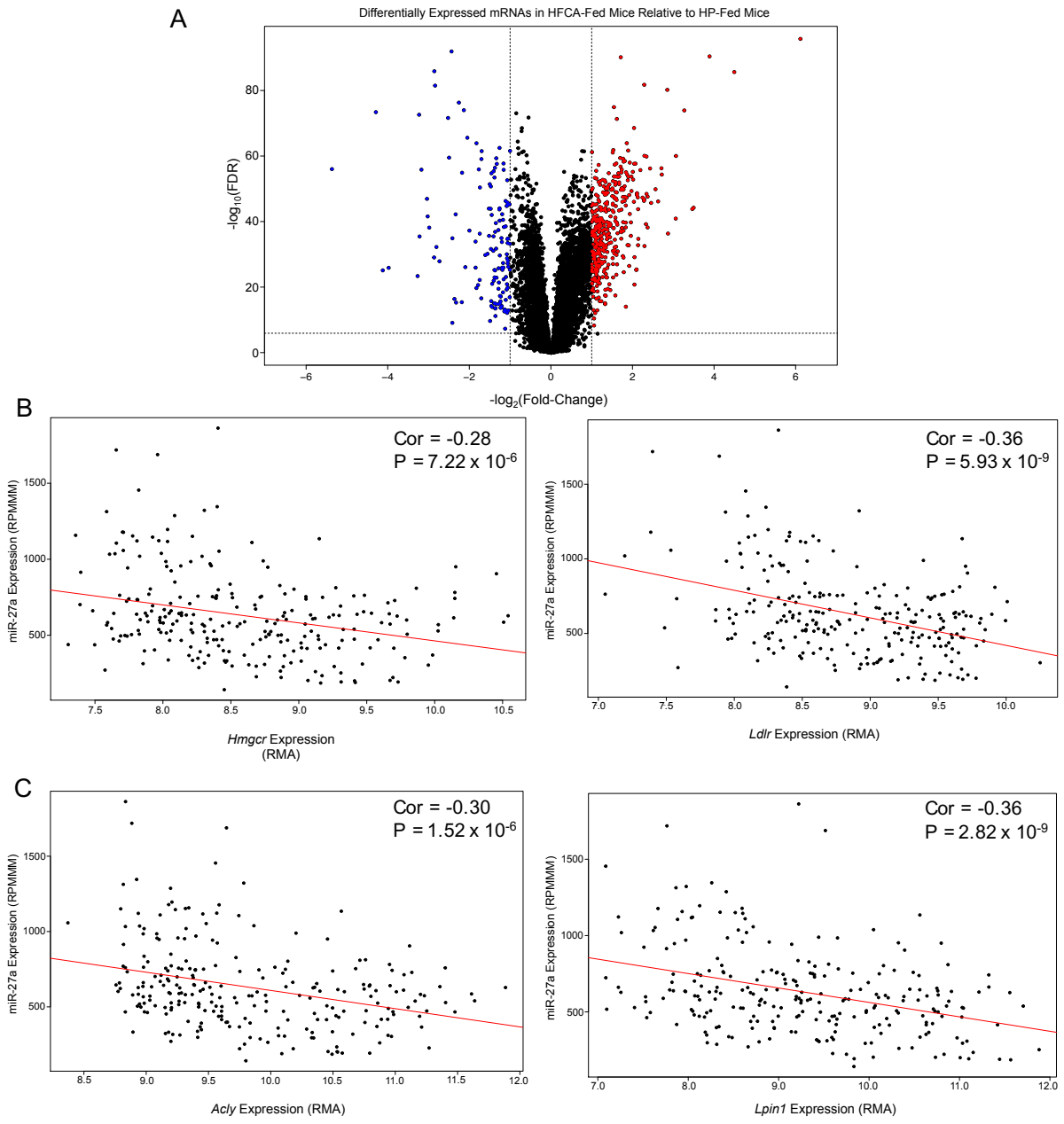

Figure 2.6: Differential expression analysis for gene expression. (A) Volcano plot of differentially expressed liver genes between HFCA-fed and HP-fed DO mice. Each dot represents one probe. Red dots are probes that are up-regulated (n=401) in HFCA-fed mice relative to HP-fed mice with a fold-change of 2 or more and a p-value ≤ 1.20e-06 (Bonferroni correction). Blue dots are probes that are down-regulated (n=140) in HFCA-fed mice relative to HP-fed mice with a fold-change of 2 or more and an p-value ≤ 1.20e-06. Horizontal dashed line denotes –log10(1.20e-06). Vertical dashed lines denote fold change of -2 (left) and 2 (right). **(B,C)** Correlation plots illustrating the inverse relationship between miR-27a and *Hmgcr*, *Ldlr*, *Acly*, and *Lpin1* expression (correlations calculated using bicor).

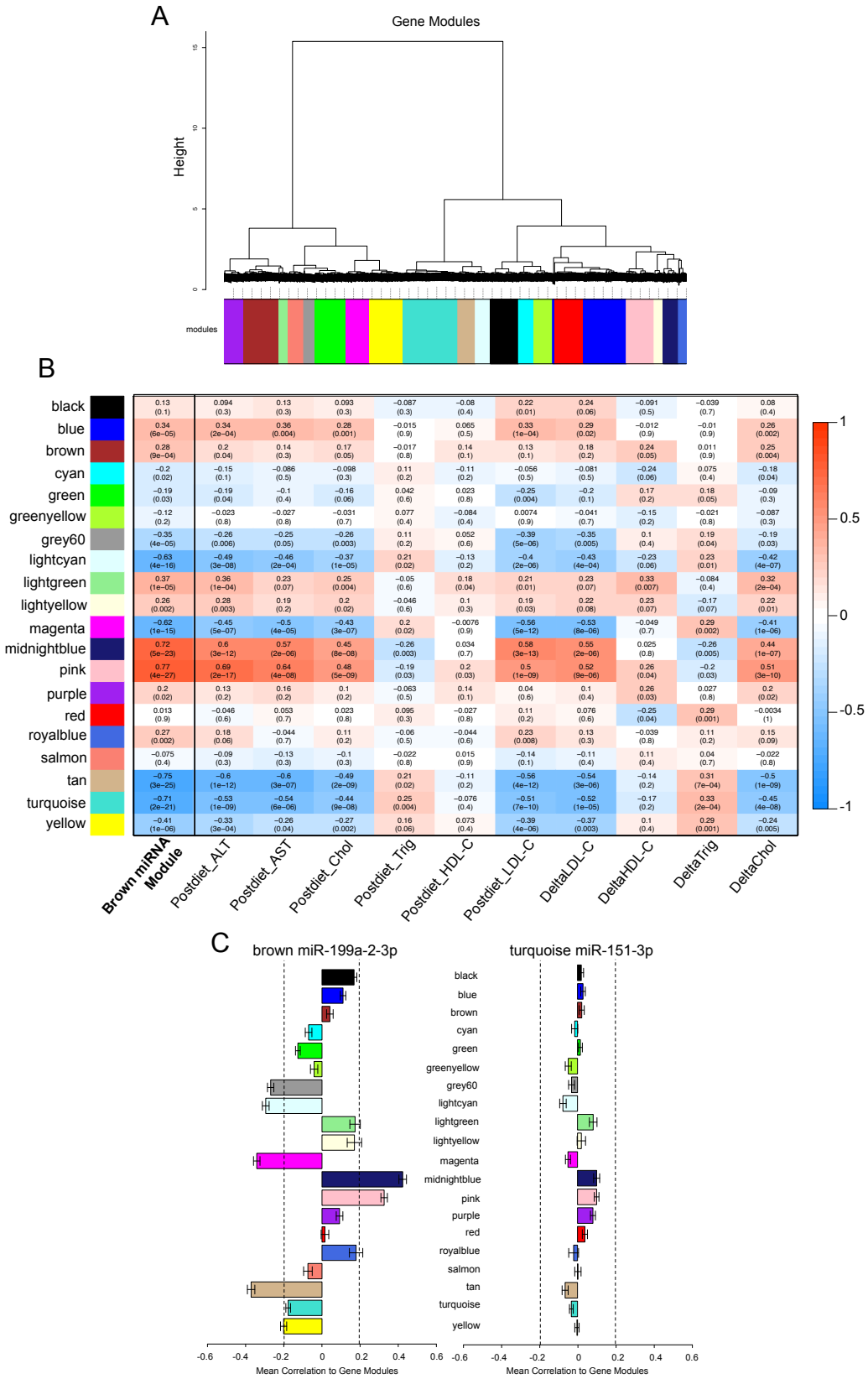
A
Gene Modules
Height
modules
B
black
blue
brown
cyan
green
greenyellow
grey60
lightcyan
lightgreen
lightyellow
magenta
midnightblue
pink
purple
red
royalblue
salmon
tan
turquoise
yellow
Brown miRNA Module
Postdiet_ALT
Postdiet_AST
Postdiet_Chol
Postdiet_Trig
Postdiet_HDL_C
Postdiet_LDL_C
DeltaLDL_C
DeltaHDL_C
DeltaTrig
DeltaChol
1
0.5
0
-0.5
-1
C
brown miR-199a-2-3p
turquoise miR-151-3p
Mean Correlation to Gene Modules
Mean Correlation to Gene Modules

Figure 2.7: Gene co-expression analysis identifies gCRMs that are correlated with the brown mCRM. (A) Gene co-regulated modules formed using WGCNA. Only HFCA mice were used during analysis. The top 3,000 most variable genes and the hybrid tree-cutting function in the WGCNA software package were used to form modules. **(B)** Heatmap of correlations between mRNA (gene) modules and brown miRNA module, and list of cardio-metabolic endpoints measured in the DO mice. Eigengenes and eigenmiRs were calculated using the WGCNA function, and correlated using the biweight midcorrelation to normalized endpoint values. The intensity of orange or blue denotes how close the correlation coefficient is to 1 or -1, respectively. Numbers in parentheses are Student p-values. **(C)** Aggregate correlation values of miRNAs to gene module members. Biweight midcorrelations were calculated between each individual miRNA and each gene. Values were averaged (mean) across gene module members. Dashed lines denote significant correlation values (-0.198, 0.198) as determined by 97.5% quantile of 1000 permutations.

CHAPTER 3: GENETIC AND MICRORNA ASSOCIATION WITH THE CARDIOMETABOLIC DISEASE RISK FACTOR TMAO

3.1 INTRODUCTION

The previous chapter covered the work on discovering miRNA modules associated with elevated LDL-C, a common hallmark and risk factor of cardiometabolic dysfunction. While the work in this chapter exploits the same gene and miRNA expression to find associations that convey potential regulatory relationships, it also incorporates the genotype information in order to elucidate the genetic architectural component regulating another cardiometabolic risk factor, TMAO.

Intestinal microbiota metabolism is now appreciated as having a profound impact on cardiovascular disease (CVD) risk [Bennett et al. 2013a, Koeth et al. 2013, Loscalzo 2011, Rak and Rader 2012, Tang et al. 2013, Wang et al. 2011]. In particular, the meta-organismal pathway involving dietary choline, gut microbiota, and TMAO, has been shown to be predictive of cardiovascular disease events in humans [Tang et al. 2013]. Just as in mice that are genetically engineered to be susceptible to atherosclerosis, dietary supplementation with choline (the dietary precursor of TMAO) or treatment with TMAO itself has been shown to promote atherosclerotic lesion development [Wang et al. 2011]. The pathway for TMAO is thought to be dependent on the microbiota, as administration of antibiotics

suppresses TMAO levels and atherosclerosis [Koeth et al. 2013, Koeth et al. 2014]. The effects of the microbiota on TMAO and atherosclerosis are transferrable as demonstrated by adoptive transfer studies in mice [Gregory et al. 2014].

The role of host genetics on the TMAO/CVD pathway has been difficult to dissect. The gut microbiota metabolize dietary choline to form an intermediate compound, TMA, which is then absorbed into the host circulation and detoxified to TMAO primarily by FMO3 in the liver [Bennett et al. 2013b]. Aside from the regulation by FMO3, little is known about how host genetics interact with and regulate circulating TMAO levels. Human cohort studies have demonstrated that common genetic variation within the *FMO3* gene is not associated with circulating TMAO levels [Hartiala et al. 2014]. Genetic studies of inbred mouse strains have demonstrated that TMAO levels are heritable [O'Conner et al. 2014] and have identified candidate loci [Hartiala et al. 2014, Bennett et al. 2015] associated with circulating TMAO. These studies indicate that the genetic architecture and regulation of TMAO levels is complex.

One possible mode of regulation of plasma TMAO levels is via miRNAs. Changes in gene and miRNA expression profiles have each been linked to the response and onset of many physiological conditions, including CVD. The studies presented here utilize the Diversity Outbred (DO) mouse population, an ideal resource for systems genetics [Churchill et al. 2013]. In this study, we focus on the discovery of novel genetic and miRNA associations with TMAO, which extends our previous work on cardiometabolic traits using DO mice [Smallwood et al. 2014], as

well as independent mouse and monkey models of atherosclerosis. We identify a novel link between a locus on Chromosome 12 and TMAO levels, as well as a robust association between miR-146a/b and TMAO potentially mediated through interactions with genes at the Chromosome 12 locus.

3.2 MATERIALS AND METHODS

3.2.1 Animals, Diets and Genotyping

Animal housing, husbandry, and handling, diet compositions, and methods of genotyping were previously reported in Chapter 2 and in Smallwood et al. 2014.

3.2.2 Measurement of TMAO Metabolite

Measurement of choline metabolites was performed by the University of North Carolina Nutrition Research Institute Metabolomics Core Facility (Kannapolis, NC). Plasma was extracted with three volumes of acetonitrile spiked with internal standards of TMAO-d9 (DLM-4779-1, Cambridge Isotope Laboratories), incubated on ice for 10 minutes, and centrifuged at 15,000 g for 2 minutes. Quantification of TMAO, was performed using liquid chromatography-stable isotope dilution-multiple reaction monitoring mass spectrometry (LC-SID-MRM/MS). Chromatographic separations were performed on an Atlantis Silica HILIC 3μm 4.6×150mm column (Waters Corp, Milford, USA) using a Waters ACQUITY UPLC system. The column was heated to 40°C, and the flow rate was maintained at 1 mL/min. The gradient was 5% A for 0.05 min, to 15% A in 0.35 min, to 20% A in 0.6 min, to 30% A in 1

min, to 45% A in 0.55 min, to 55% A in 0.05 min, at 55% A for 0.9 min, to 5% A in 0.05 min, at 5% A for 1.45 min, where A is 10% acetonitrile/90%water with 10 mM ammonium formate. The metabolites and their corresponding isotopes were monitored on a Waters TQ detector using characteristic precursor-product ion transitions: 76→58 for TMAO, 85→66 for TMAO-d9. Concentrations of TMAO in the samples were determined from its calibration curve using peak area ratio of the metabolite to its isotope.

3.2.3 Measurement of microRNA and mRNA Expression

Methods of RNA extraction from the livers of the DO mice, evaluation of RNA integrity, microarray analysis for mRNAs (genes), library prep and small RNA sequencing were as previously explained in Chapter 2. miRNA expression was transformed using the boxcox method. Bi-weight mid-correlations between genes, miRNAs, and TMAO levels were performed using WGCNA package in R [Langfelder and Horvath 2008]. Microarray and small RNA-seq data is available on the GEO repository, accession number GSE99561.

3.2.4 QTL Mapping

QTL and expression QTL (eQTL) mapping were performed using DOQTL [Gatti et al. 2014] package in R as previously described [Smallwood et al. 2014]. Haplotype reconstructions were performed utilizing genotyping data from a much larger cohort of DO mice available at the Jackson Laboratory (kindly performed by

Daniel Gatti; Jackson Laboratory). Diet was included as an additive covariate in the mapping model for measurements including those obtained from 24-week old mice after dietary treatment. Significant QTL were determined at a genome-wide p-value of <0.05 and suggestive QTL were determined at a p-value of <0.63, which corresponds to one false positive per genome scan [Lander and Kruglyak 1995]. QTL support intervals were defined by the 95% Bayesian credible interval, calculated by normalizing the area under the QTL curve on a given Chromosome [Sen and Churchill 2001]. The mapping statistic reported is log of the odds (LOD) ratio. The significance thresholds were determined via permutations of genome-wide scans by shuffling phenotypic or expression data in relation to individual genotypes. Association mapping was performed for TMAO levels or gene/miRNA expression with known SNPs within the identified QTL by imputing the founder SNPs onto the DO genomes. Candidate genes were identified by position based on the UCSC Genome Browser. To verify that QTL are robust, we repeated haplotype reconstruction and pre- and post-diet TMAO QTL mapping using the R/qtl2 package (http://kbroman.org/qtl2/), which differs from DOQTL in that it utilizes genotype calls instead of allele intensity plots for haplotype reconstruction.

TMAO measurements were transformed using the boxcox method to satisfy the model assumption of a normal distribution. Microarray RMA values were corrected for known SNPs in the DO founder strains. Each phenotype and miRNA were used to run 1000 permutations. For mRNA-eQTL LOD thresholds, 500 genes were chosen at random to perform 1000 permutations on each.

3.2.5 QTL-eQTL Overlap Analysis

Overlaps between mRNA eQTL, miRNA eQTL, and TMAO QTL suggestive and/or significant hits were performed by taking the peak SNPs of suggestive/significant QTL, and identifying SNPs across gene and miRNA eQTL datasets within a flanking 1Mb region that were above their respective suggestive or significant threshold. This 2Mb flanking region was then removed, and the next peak SNP was identified. The process was repeated until no other SNPs were found to be above the suggestive/significant threshold.

3.2.6 Bioinformatics

MicroRNA target prediction in mouse reference (NCBI build 37) 3' UTRs was performed using the TargetScan algorithm [Argwal et al. 2015, Grimson et al. 2007]. Positional conservation of a predicted target site in at least one other species was required. Correlation between miRNA levels and gene expression as well as TMAO was calculated using biweight mid-correlation analysis. All p-values take into account multiple testing unless specifically noted otherwise. Significant differences between diet groups were determined using a Student's t-test.

3.2.7 Small RNA-Sequencing in Livers of LIRKO Mice

Treatment and phenotyping of liver-specific insulin receptor knockout (LIRKO) mice has been previously reported by Miao et al. 2015. Liver tissue was isolated

from LIRKO mice (n=3) and floxed controls (n=2). RNA was extracted using Norgen Total RNA Purification Kit (Norgen Biotek, Thorold, ON, Canada). RNA yield and quality was assessed by Thermo Scientific NanoDrop 2000 (Waltham, MA) and integrity was measured by Agilent 2100 Bioanalyzer (Santa Clara, CA). Small RNA library preparation using the TriLink CleanTag kit was performed at the Genome Sequencing Facility of Greehey Children's Cancer Research Institute at the University of Texas Health Science Center at San Antonio and sequencing was carried out on the HiSeq platform at an average depth of ~40 million reads/sample (with a range across samples of 18 million to 65 million). Data processing and miRNA quantification was performed using miRquant 2.0 [Kanke et al. 2016]. On average, >75% of the trimmed reads mapped to the mouse genome, of which almost 50% corresponded to miRNA loci. The length distribution of mapped reads revealed a clear peak at the size range 21-24 nucleotides, which matches the expectation for miRNAs.

3.2.8 Small RNA-Sequencing in Livers of Monkeys

Dietary treatment, liver tissue extraction, and phenotyping of African Green Monkeys has been previously reported by Chung et al., 2014. The samples used for reporting here were from monkeys fed either chow or a high fat high cholesterol (0.4 mg/kcal cholesterol) diet. RNA was extracted using Norgen Total RNA Purification Kit (Norgen Biotek, Thorold, ON, Canada). RNA yield and quality was assessed by Thermo Scientific NanoDrop 2000 (Waltham, MA) and integrity was measured by

Agilent 2100 Bioanalyzer (Santa Clara, CA). Small RNA library preparation using the TriLink CleanTag kit was performed at the Genome Sequencing Facility of Greehey Children's Cancer Research Institute at the University of Texas Health Science Center at San Antonio and sequencing was carried out on the HiSeq platform at an average depth of ~17 million reads/sample. Data processing and microRNA quantification was performed using miRquant 2.0 [Kanke et al. 2016]. On average, ~85% of the trimmed reads mapped to the human genome, of which greater than ~65% corresponded to miRNA loci. The length distribution of mapped reads revealed a clear peak at the size range 21-24 nucleotides, which matches the expectation for miRNAs.

3.3 RESULTS

Several studies have identified a role for the microbiome in the regulation of circulating TMAO; however, relatively few studies to date have reported host genetic associations with TMAO. In this study, we leveraged a DO cohort of 288 female mice comprised of 144 sibling pairs that has been previously described in Chapter 2 to identify novel genetic regulators of circulating TMAO. In those earlier studies we reported baseline measures of metabolic parameters and molecular traits in the DO mice during feeding of a synthetic AIN-76 diet and then after feeding with either a high-fat, high cholesterol diet with added cholic acid (HFCA) or a high protein diet (HP) for 18 weeks. In the current study, we integrate hepatic messenger RNA (mRNA) expression and miRNA expression with targeted measurement of the

metabolite TMAO to identify novel candidate genetic regulators of TMAO concentrations in circulation.

3.3.1 Plasma TMAO levels in the Diversity Outbred mice and its relationship to cardiovascular risk factors

At baseline the mean circulating TMAO value in the DO mice was 1.25 mmol and ranged between 0.16 and 5.41 mmol. There was no difference in TMAO levels between the mice randomized to each of the diet groups at baseline (data not shown). However, after 18 weeks on an atherogenic diet we observed a highly significant increase in TMAO levels, 4.19±0.22 mmol ($p = 1.36 \times 10^{-22}$), whereas TMAO levels in the mice fed the HP diet did not differ significantly from the baseline levels (1.25±0.06 mmol) (Figure 3.1A).

We next assessed the relationship between TMAO and cardiometabolic risk factors. We found a significant correlation between TMAO and LDL/VLDL cholesterol ($R = 0.61$, $p = 3.56 \times 10^{-26}$) (Figure 3.1B) as well as total cholesterol ($R = 0.56$, $p = 2.18 \times 10^{-22}$) (Figure 3.1C). A simple Pearson correlation analysis across all samples indicated a moderate relationship between TMAO and atherosclerosis ($R=0.26$, $P=0.002$) similar to previous reports [Bennett et al. 2013b, Bennett et al. 2015, Wang et al. 2011], but the biweight mid-correlation between atherosclerotic lesion size and TMAO was not significant (Figure 3.1D), indicating that a relationship between TMAO and atherosclerosis is likely driven by a few of the more susceptible mice in the study.

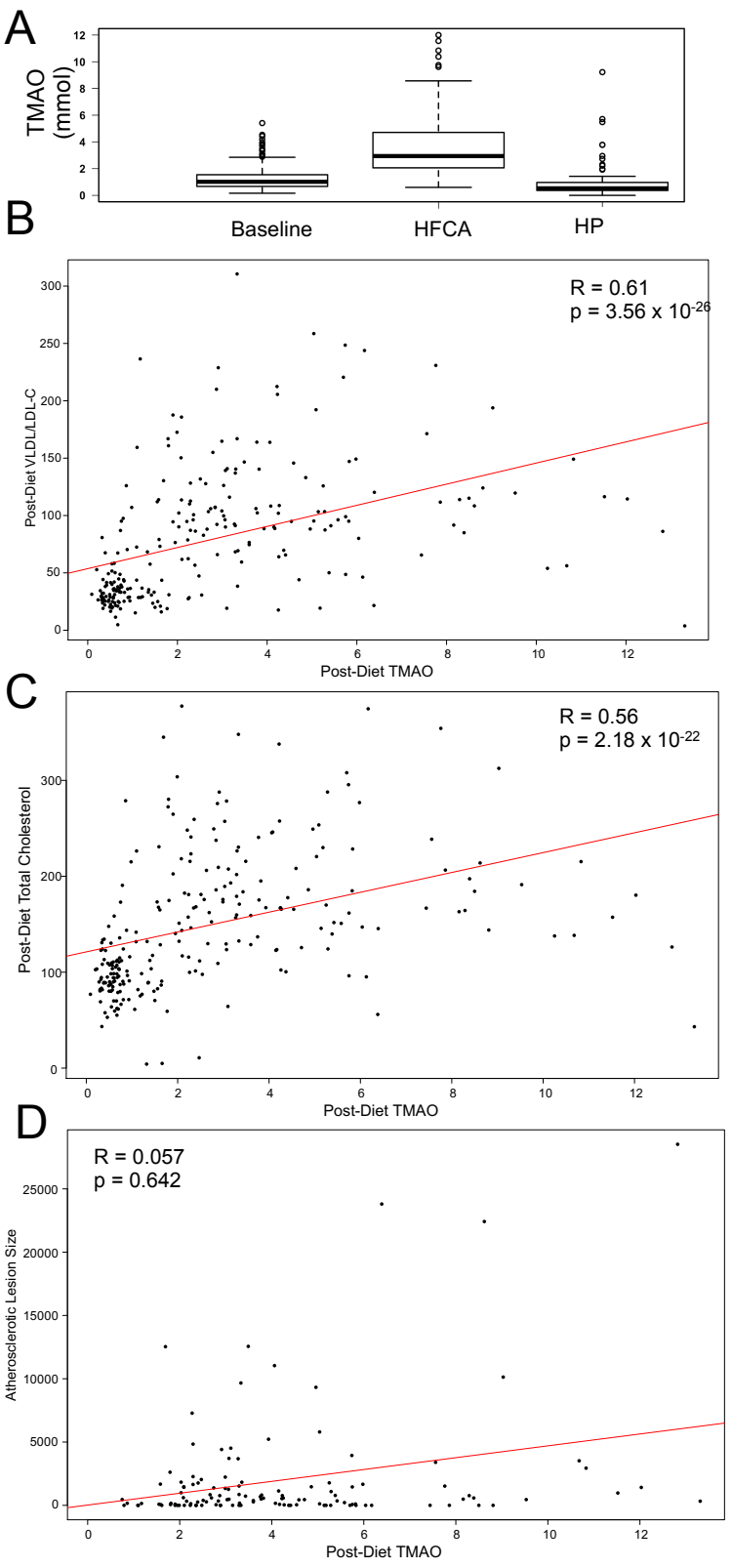

Figure 3.1: Plasma TMAO levels are altered by diet and are correlated with cardiometabolic phenotypes in the DO mice. (A) Mice were maintained on a synthetic diet for two weeks, fasted for four hours, and then phenotyped for plasma clinical chemistries at 6 weeks of age (Baseline). Following two weeks of synthetic diet, mice were transferred to either a high protein diet (HP) or an atherogenic diet (HFCA). Plasma was taken from 24-week-old mice after 18 weeks on their respective diets, and with four hours fasting, and then phenotyped for plasma clinical chemistries after diet treatment (HFCA or HP). TMAO is measured in mmol. **(B-D)** Correlation plots demonstrating the associations between post-diet TMAO and post-diet VLDL/LDL-C (B) or Total Cholesterol (C) or atherosclerotic lesion size (D). Lesion sizes were measured in Oil Red O-stained slides of cross sections of aortic sinuses from each mouse, and averaged based on number of slides per mouse. Data is only representative of HFCA-fed mice, as HP-fed mice developed no lesions.

3.3.2 Identification of Chromosome 12 QTL associated with TMAO at baseline and after dietary treatment

We identified a novel QTL on Chromosome 12 that is associated with TMAO levels at baseline and after the dietary regimen. In 6-week old mice, the Chromosome 12 QTL is significantly associated with TMAO levels (LOD=9.98, $p<0.05$) (Figure 3.2A). This QTL has a ~3.5 Mb (83.55-87.02 Mb) support interval with a peak SNP at 86.25 Mb. Based on the founder allele effects, we see that allelic contribution from NOD/ShiLtJ and 129S1/SvImJ within the interval are associated with the highest TMAO levels, while allelic contribution from CAST/EiJ and PWK/PhJ alleles are associated with the lowest TMAO levels (Figure 3.2B). According to the UCSC Genome Browser, there are 78 positional candidates within this QTL interval (Table A.2).[2]

We identified a coincident QTL on Chromosome 12 (LOD= 7.04, $p<0.01$) associated with TMAO after dietary treatment, which overlaps with the QTL on Chromosome 12 in these mice at 6 weeks of age prior to dietary treatment (Figure 3.2C). The allele effects for the QTL also indicate that allelic contribution from CAST/EiJ and PWK/EiJ are associated with low circulating TMAO levels, suggesting that a variant shared by these strains may be causal for the Chromosome 12 QTL both before and after dietary treatment (Figure 3.2D). The QTL for TMAO after dietary treatment has a ~3.6 Mb (83.46 Mb- 87.08 Mb) support interval with a peak SNP at 85.7 Mb, which is about 500 kb from the peak SNP associated with baseline

[2] Oversized tables can be found in Appendix A.

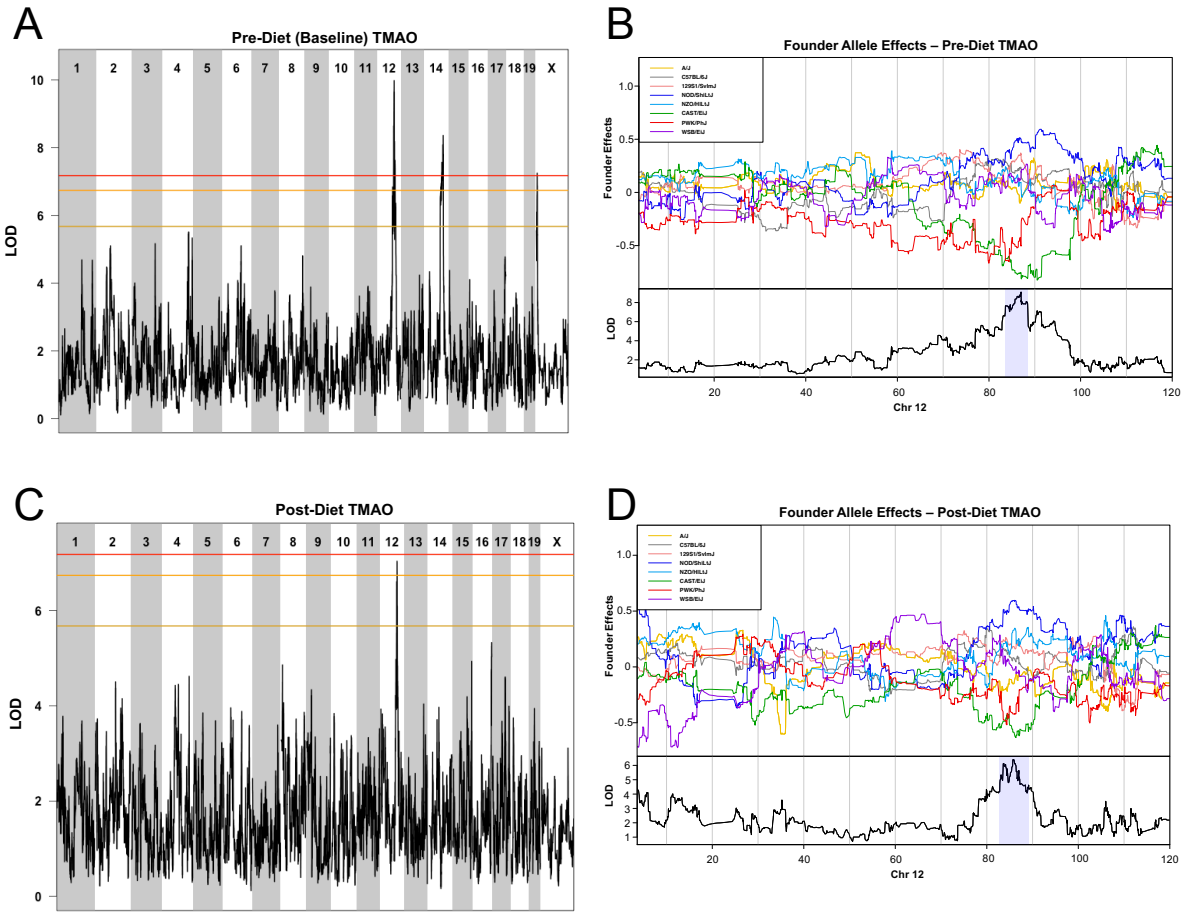

Figure 3.2: Genetic mapping of TMAO in DO mice before and after dietary treatment. Genome-wide quantitative trait locus (QTL) scan for loci affecting TMAO concentrations. Chromosomes 1 through X are represented numerically on the *x*-axis and the *y*-axis represents the logarithm of odds (LOD) score. The relative width of the space allotted for each Chromosome reflects the relative length of each Chromosome. Colored lines show permutation-derived significance thresholds (N=1000) at $P = 0.05$ (LOD=7.5, shown in red), $P = 0.10$ (LOD=7.1, shown in orange), and $P = 0.63$ (LOD=5.8, shown in yellow). **(A)** Genome-wide QTL scan for loci affecting TMAO levels at baseline. **(B)** The eight coefficients of the QTL model show the effect of each founder allele on the phenotype along Chromosome 12. **(C)** Genome-wide QTL scan for loci affecting TMAO levels after 18 weeks of dietary treatment. **(D)** Shading identifies the 95% Bayesian credible interval around the peak.

TMAO levels in mice at 6 weeks of age. The interval contains 80 positional candidates (Table A.2).

We also identified a significant QTL for baseline TMAO levels on Chromosome 14 (LOD= 8.2, $p < 0.05$) that spans a 16.3 Mb region (Figure 3.2A) and contains 21 positional candidate genes. Based on the founder allele effects we observed that C57BL/6J, 129S1/SvImJ, CAST/EiJ, and PWK/PhJ alleles in this interval are associated with higher TMAO levels, while A/J, NOD/ShiLtJ, NZO/HiLtJ, and WSB/EiJ alleles are associated with lower TMAO levels. This QTL, however, was not detected in the post-diet analysis and thus we focused only on the Chromosome 12 locus.

3.3.3 A subset of genes within the Chromosome 12 TMAO QTL have overlapping eQTL

In order to identify any candidate genes within the 95% confidence interval of the Chromosome 12 TMAO QTL whose expression is related to TMAO levels, we sought to determine whether any of the Chromosome 12 interval genes have local eQTL, or *cis*-eQTL, that overlap the TMAO QTL. First, since the pre-diet and post-diet TMAO QTL are overlapping, we extended the interval to encompass the QTL SNPs above the 90% threshold (which includes the post-diet QTL) within both intervals (83.46Mb – 87.08Mb). Then, for each of the 63 out of 80 genes in the extended interval for which a probe was present in the microarray, we associated differences in gene expression to genetic differences using the R package DOQTL

as described in the methods [Gatti et al. 2014]. We and others have previously shown that DOQTL effectively identifies QTL for quantitative traits and can estimate the allele effects for haplotypes of the 8 founders of the DO [Churchill et al. 2012, Gatti et al. 2014, Smallwood et al. 2014, Svenson et al. 2012]. Next, for each of the 63 genes with significant eQTL SNPs, we assessed whether any of the eQTLs are within 1 Mb on either side of the TMAO QTL (Methods), which we refer to as overlapping. We found significant overlapping eQTL for 14 of the 63 genes (Figure 3.3, Table 3.1).

For each of these 14 genes we next investigated the allele effects at their respective eQTL SNPs. We found that for 7 of the 14 genes, the allele effect is consistent with that of the QTL for TMAO, with 3 having CAST/EiJ and PWK/PhJ alleles associating with lower levels of expression, and 4 with the same alleles associating with higher levels of expression (Table 3.1, Figure 3.3, Figure 3.4). For example, based on the founder coefficients for each marker along Chromosome 12, we see that both CAST/EiJ and PWK/PhJ alleles are associated with higher expression of *Numb* and lower expression of *Acot4* (Figure 3.3). Several genes, such as *Arel1*, have overlapping eQTL with the TMAO QTL but do not have similar allele effect patterns. Among the 7 genes with overlapping eQTL and shared allele effect patterns with the TMAO QTL, we next sought to identify genes that are positively or inversely correlated with pre- and post-diet TMAO by performing a biweight mid-correlation analysis. We found that 4 of these 7 are significantly associated with post-diet TMAO levels (Table 3.2), including *Numb* and *Acot4*.

Taken together, these results indicate that mice contributing CAST/EiJ and PWK/PhJ alleles at the Chromosome 12 QTL have higher expression of *Numb*, lower expression of *Acot4*, and lower circulating TMAO levels.

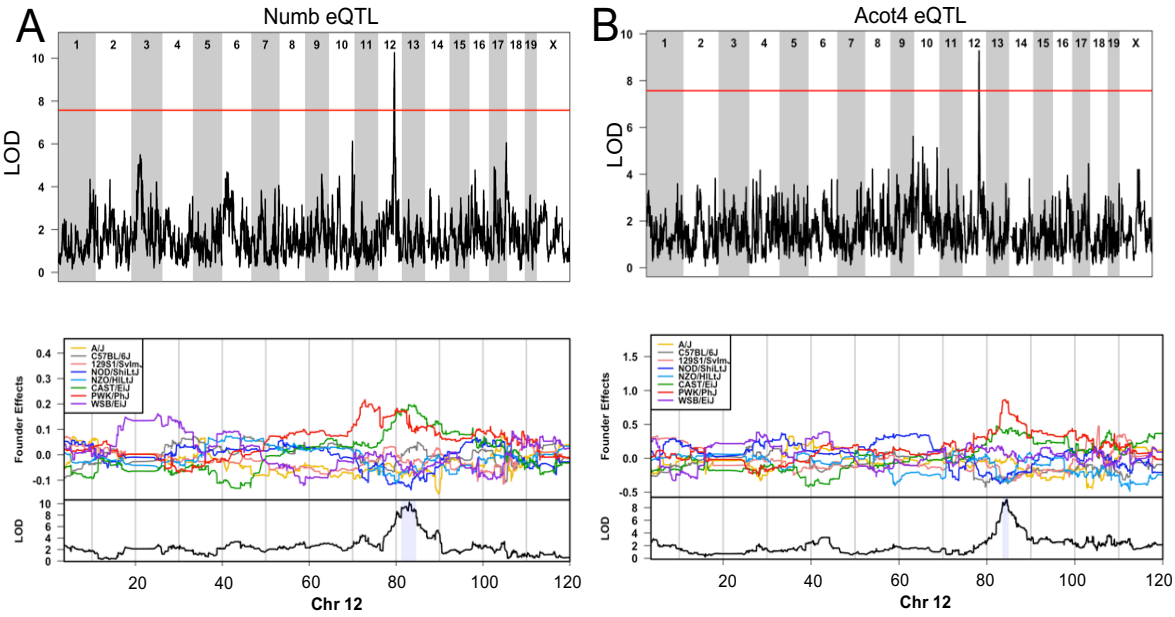

Figure 3.3: Genes within the TMAO QTL interval have overlapping eQTL. **(A,B)** Genome-wide QTL scan for loci affecting A) *Numb*, and B) *Acot4* in the livers of the DO mice. Chromosomes 1 through X are represented numerically on the *x*-axis and the *y*-axis represents the LOD score. The relative width of the space allotted for each chromosome reflects the relative length of each chromosome. Red lines in top plots show permutation-derived significance thresholds (N=1000) at $P = 0.05$ (LOD=7.2). The bottom plots show eight coefficients of the QTL model depicting the effect of each founder allele on the expression of each respective gene along Chromosome 12. Shading identifies the 95% Bayesian credible interval around the peak.

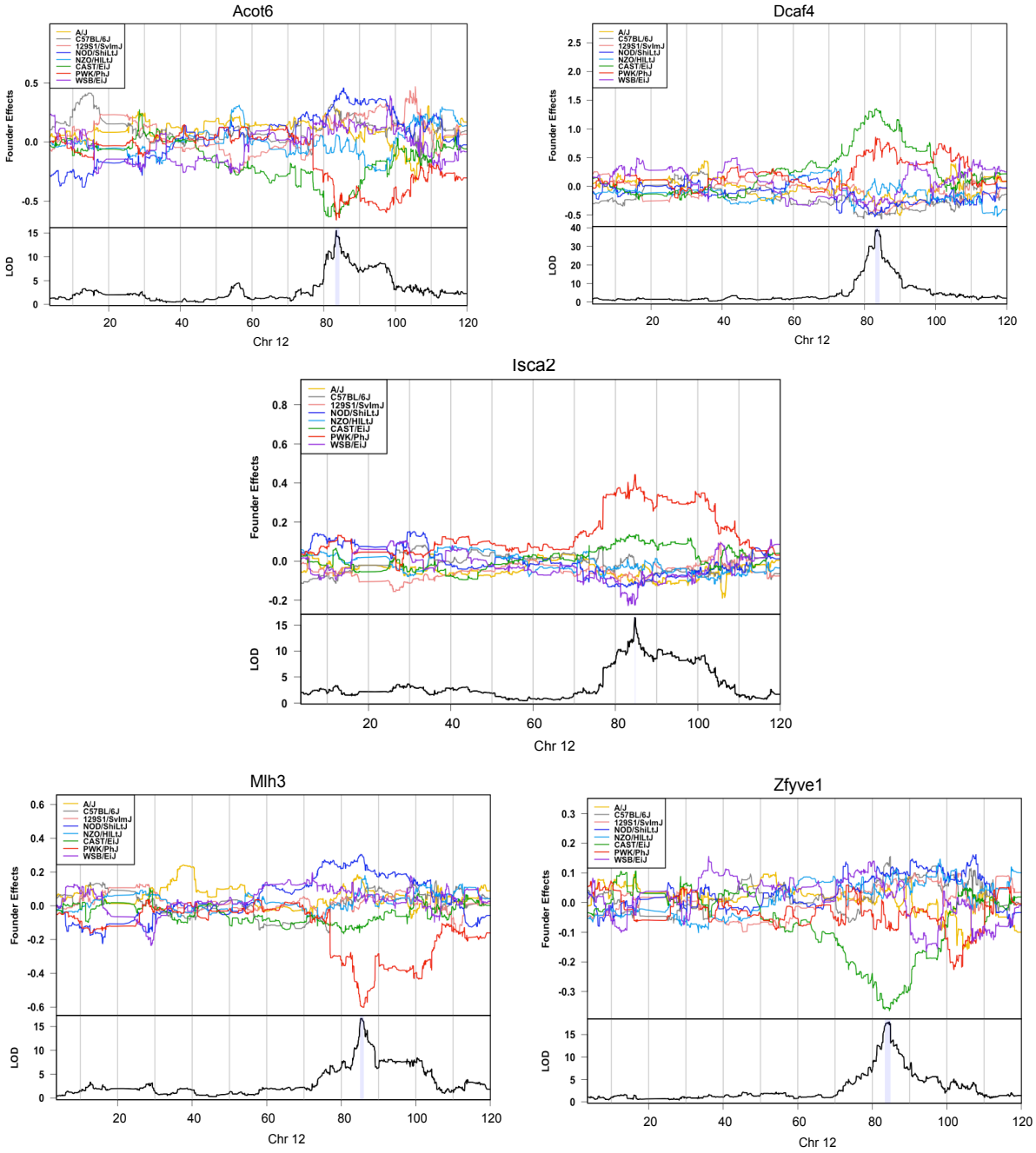

Figure 3.4: The remainder of the seven genes that have consistent founder allele effects with TMAO. Coefficients of the QTL model showing the effect of each of the eight founders' alleles on expression along Chromosome 12.

3.3.4 Evidence for robust miR-146 association with circulating TMAO levels and atherosclerosis

MiRNAs represent an important class of diet-responsive biomolecules that have emerged as prominent regulators of atherogenesis and related cardiometabolic conditions [Feinberg and Moor 2016]. While there is only one known miRNA encoded in the TMAO QTL interval, miR-6938, we found that it is not expressed in the DO mice livers. Thus, we sought to determine whether any robustly expressed liver miRNAs are associated with the 7 genes of interest in the post-diet TMAO QTL that have an overlapping eQTL and a shared allele effect pattern with the TMAO QTL, and are strongly negatively or positively correlated with circulating TMAO levels. In a previous study, we had shown that 246 miRNAs are robustly expressed in the livers of the same cohort of DO mice [Chapter 2]. Among these, we found that 4 miRNAs had predicted conserved target sites in 2 of the 7 genes of interest (Table 3.3). Using biweight mid-correlation analysis, we found that 3 of the 4 miRNAs exhibit significant negative or positive correlation with TMAO, and 2 miRNAs among these with the highest and most significant correlations with post-diet TMAO are miR-146a and miR-146b ($R = 0.42$ and 0.40, respectively, p-value = 9.26×10^{-11} and 6.42×10^{-10}, respectively, Figure 3.5). Of note, we observed that a few miRNAs that do not have predicted target sites in the 7 genes of interest, most notably miR-34a and miR-1247, are nonetheless very strongly positively correlated with TMAO ($R = 0.62$ and 0.53, respectively).

Table 3.1: Significant eQTL at the chromosome 12 TMAO locus.

Gene	Gene Chromosome	Strand	TSS	Peak SNP	Peak SNP LOD	P-value	Allele Effects Consistent with TMAO?
Acot4	12	+	84.04	UNC21525003	9.44	2.68×10^{-07}	Yes
Acot6	12	+	84.10	UNC21513484	15.45	8.66×10^{-13}	Yes
Arel1	12	-	84.92	UNC21535324	21.09	4.20×10^{-18}	No
Angel1	12	-	86.70	UNC21559085	16.80	4.72×10^{-14}	No
Dcaf4	12	+	83.52	UNC21513484	39.74	4.54×10^{-36}	Yes
Ift43	12	+	86.08	UNC21535324	8.69	1.26×10^{-06}	No
Isca2	12	+	84.77	UNC21526856	16.46	9.75×10^{-14}	Yes
Lin52	12	+	84.45	UNC21506680	10.90	1.31×10^{-08}	No
Mlh3	12	-	85.23	UNC21535914	16.83	4.42×10^{-14}	Yes
Numb	12	-	83.79	UNC21509060	10.42	3.58×10^{-08}	Yes
Ptgr2	12	+	84.29	UNC21523779	7.28	2.13×10^{-05}	No
Rnf113a2	12	+	84.42	UNC21522264	7.71	9.03×10^{-06}	No
Syndig1l	12	-	84.68	UNC21526856	35.04	1.65×10^{-31}	No
Zfyve1	12	-	83.55	UNC21525523	17.79	5.63×10^{-15}	Yes

Table 3.2: Correlations between genes and TMAO levels. Correlations were calculated using the biweight mid-correlation method with Student's p-values.

Gene	Pre-Diet TMAO	Pre-Diet TMAO P-value	Post-Diet TMAO	Post-Diet TMAO P-value
Acot4	-0.16	0.621	**-0.3**	**1.72×10^{-05}**
Acot6	0.08	0.862	**0.37**	**1.86×10^{-08}**
Angel1	0.14	0.659	0.04	0.638
Arel1	-0.05	0.919	0.05	0.593
Dcaf4	-0.18	0.543	-0.12	0.142
Ift43	-0.03	0.949	**0.52**	**2.36×10^{-16}**
Isca2	-0.11	0.757	**-0.37**	**1.42×10^{-08}**
Lin52	-0.004	0.998	**-0.19**	**0.01**
Mlh3	0.22	0.409	0.06	0.471
Numb	-0.13	0.715	**-0.27**	**0.0001**
Ptgr2	0.04	0.931	**-0.32**	**2.76×10^{-06}**
Rnf113a2	-0.09	0.821	**-0.29**	**2.58×10^{-05}**
Syndig1l	-0.03	0.959	0.05	0.547
Zfyve1	0.18	0.509	-0.04	0.626

Table 3.3: miRNAs that are predicted to target genes with eQTL overlapping the TMAO QTL and exhibit consistent allele effects. Information for predicted regulatory relationships were taken from TargetScanMouse. All miRNA target sites are conserved between mouse and at least one other species.

Genes	miRNAs
Isca2	mmu-miR-23a-3p
Numb	mmu-miR-146a-5p, mmu-miR-146b-5p, mmu-miR-31-5p

Table 3.4: Mapping statistics for small RNA-seq in LIRKO mice.

Sample_name	Flox 1	Flox 2	LIRKO 1	LIRKO 2	LIRKO 3
Total reads	18838092	48825252	65008050	45497483	18043841
Trimmed reads	4665795	16598923	25657079	14924053	6489845
% Trimmed reads	24.77	34	39.47	32.8	35.97
Short reads	13275125	856164	1456951	1427793	10606325
% Short	70.47	1.75	2.24	3.14	58.78
Exact match to genome	2376178	10930102	17218926	9795010	3520129
% EM	50.93	65.85	67.11	65.63	54.24
No exact match to genome	2289617	5668821	8438153	5129043	2969716
% NEM	49.07	34.15	32.89	34.37	45.76
Toatl mapped reads	3346853	12989102	19475840	12273850	5073373
% Mapped	71.73	78.25	75.91	82.24	78.17
Total mapped to miRs	1621051	5549923	7233770	7163434	2630898
% of total mapped to miRs	48.44	42.73	37.14	58.36	51.86
Total mapped to tRNAs	119835	940582	1798469	376758	303200
% of total mapped to tRNAs	3.58	7.24	9.23	3.07	5.98

Indeed, miR-146a, miR-146b, miR-34a, and miR-1247 are among the most significantly up-regulated miRNAs in the liver of HFCA-fed DO mice compared to the HP-fed DO mice (2.2-fold, 4.3-fold, 6-fold, and 4.6-fold, respectively; FDR < 0.05, Figure 3.6).

To test whether the miRNAs most strongly correlated with TMAO in the DO mice are also significantly elevated in an independent model documented to have elevated circulating TMAO, we examined the liver-specific insulin receptor knockout (LIRKO) mouse model [Feinberg and Moore 2016] performed small RNA-seq analysis with liver tissue from both LIRKO and floxed control mice (Table 3.4) and found that miR-1247, miR-34a, miR-146a, and miR-146b were among the top 10 most highly up-regulated hepatic miRNAs in LIRKO mice, at levels of ~35-fold, ~7-fold, ~2.2-fold and ~1.8-fold, respectively (Figure 3.6).

We next sought to provide even further validation for this result by studying a monkey model of hypercholesterolemia and atherosclerosis. Specifically, we profiled miRNAs by small RNA-seq in liver tissue from African Green Monkeys at baseline during standard chow diet feeding (n=4) and then after a 10-week high-fat/high-cholesterol (HFHC) atherogenic diet (n=5) (Table 3.5). As expected, the HFHC diet significantly elevates circulating TMAO (Figure 3.7A). We found that miR-146a is among the most robustly up-regulated miRNAs (>2-fold) in the liver in response to the HFHC diet (Figure 3.7B**,** Figure 3.8), and that miR-146b is also significantly elevated albeit to a slightly lower degree (Figure 3.7B, Table A.3). Importantly, we also found that the levels of miR-146a in the HFHC-fed monkeys are very strongly correlated with plasma TMAO levels ($R = 0.67$, $p < 0.05$), as well as total plasma cholesterol ($R = 0.58$, $p < 0.05$), LDL cholesterol ($R = 0.45$, $p < 0.05$), and HDL cholesterol ($R = 0.70$, $p < 0.05$).

According to the literature, the predicted interaction between miR-146 and *Numb* (Table 3.3) has been validated in several different tissues [Forloni et al. 2014, Hwang et al. 2014], though it has not been studied previously in the liver. We found that in the DO mice both miR-146a and miR-146b are strongly inversely correlated with *Numb* ($R = -0.40$ and -0.29, respectively; p-value = 6.44×10^{-10} and 2.05×10^{-05}, respectively, Figure 3.7C-D) across the 288 post-diet murine liver samples, supporting the previously suggested regulatory connection. Overall, these results point to a strong link between miR-146a/b and TMAO, particularly in the context of an atherogenic diet.

3.4 DISCUSSION

High circulating TMAO levels are strongly associated with an increased risk for cardiovascular disease. However, the factors that are involved in regulating TMAO levels remain incompletely characterized. In this work, we've carried out a genetic mapping study of TMAO in a cohort of 288 female Diversity Outbred mice, of which half were fed an atherogenic diet. The three main novel findings are the identification of: (i) a QTL on Chromosome 12 that is associated with circulating TMAO levels; (ii) at least 7 genes at this locus, including *Numb*, which have overlapping eQTLs for which allele effects are consistent with the TMAO QTL, and which have post-diet expression levels that are significantly negatively or positively correlated with TMAO; and (iii) a link between miR-146 and TMAO levels, possibly through regulation of one of these 7 genes, *Numb*. Numb has been implicated previously in the control of circulating cholesterol levels through its role in regulating intestinal cholesterol absorption [Li et al. 2014, Wei et al. 2014]. Moreover, a recent study reported that genetic variants in *Numb* may be associated with coronary artery disease in humans [Abudoukelimu et al. 2015]. What, if any, functions Numb has in the liver, or in immune cells recruited to the liver in atherosclerosis, remains to be determined.

The levels of miR-146a are significantly elevated in atherosclerotic plaques and rise with disease progression [Cheng et al. 2017, Raitoharju et al. 2011]. Moreover, levels of miR-146 in serum have been associated with hyperlipidemia [Simionescu et al. 2014] and were shown recently to be reduced upon statin treatment [Yang et al.

2016]. A very recent study using a knockout mouse model demonstrated that miR-146a is a critical regulator of cholesterol metabolism and inflammatory signaling in the context of an atherogenic diet [Cheng et al. 2017]. The authors showed that genetic ablation of miR-146a in bone marrow-derived cells reduces circulating LDL-C and suppresses atherogenesis, whereas deletion of miR-146a in the vascular endothelium promotes the development of atherosclerosis in part due to unrestrained inflammatory signaling. This complicated result points to the pleiotropism of miRNAs and their cell-type specific functionalities. The inverse correlation and regulatory interaction between miR-146 and *Numb* has been validated experimentally in other tissues [Forloni et al. 2014, Hwang et al. 2014], but to our knowledge this study represents the first report of its potential relevance in the context of atherosclerosis. Interestingly, miR-146a and miR-146b are members of the co-regulated module of miRNAs identified in our previous study that is associated with circulating LDL-C levels in the DO mice fed the HFCA atherogenic diet. In addition, *Numb* is a member of a co-regulated module that is inversely correlated with the LDL-C associated miRNA module identified in Chapter 2. Our current and previous findings strengthen the burgeoning connection between miR-146 and diet-associated cardiometabolic disease by demonstrating a novel correlation with circulating TMAO via a possible mediator of this relationship, Numb. It is also important to point out here that there are 80 genes in the Chr12 TMAO QTL, and genes other than *Numb* at this locus may be more relevant to TMAO levels and atherogenesis, and therefore, merit attention in future studies as well.

It has been shown previously that TMAO levels are significantly elevated in the liver-specific insulin receptor knockout (LIRKO) mouse model [Miao et al. 2015]. We demonstrate here that miR-146a and miR-146b, as well as a suite of other miRNAs including miR-34a and miR-1247, which are positively correlated with TMAO in the DO cohort, are also significantly elevated in the livers of LIRKO mice. Lastly, we demonstrate that miR-146a is among the most aberrantly elevated miRNAs in the livers of monkeys fed a high-fat/high-cholesterol diet, and is also significantly correlated with plasma TMAO levels, which provides even greater physiological relevance to diet-associated atherosclerosis in humans. The >30-fold increase in miR-1247 in the liver of LIRKO mice compared to floxed controls also warrants a deeper study, especially since the targets and functions of miR-1247 are very poorly characterized.

Overall, these findings strengthen the connection between miRNAs and dyslipidemia/atherosclerosis. Whether the rise in the levels of miR-146a/b and other miRNAs contributes to atherogenesis or is part of an adaptive response to atherogenic stimuli is not yet known and merits further investigation. As demonstrated by the Cheng et al. study (2017) of the miR-146 knockout mice, these miRNAs could serve both pro- and anti-atherogenic roles, depending on cellular context and stage of disease progression.

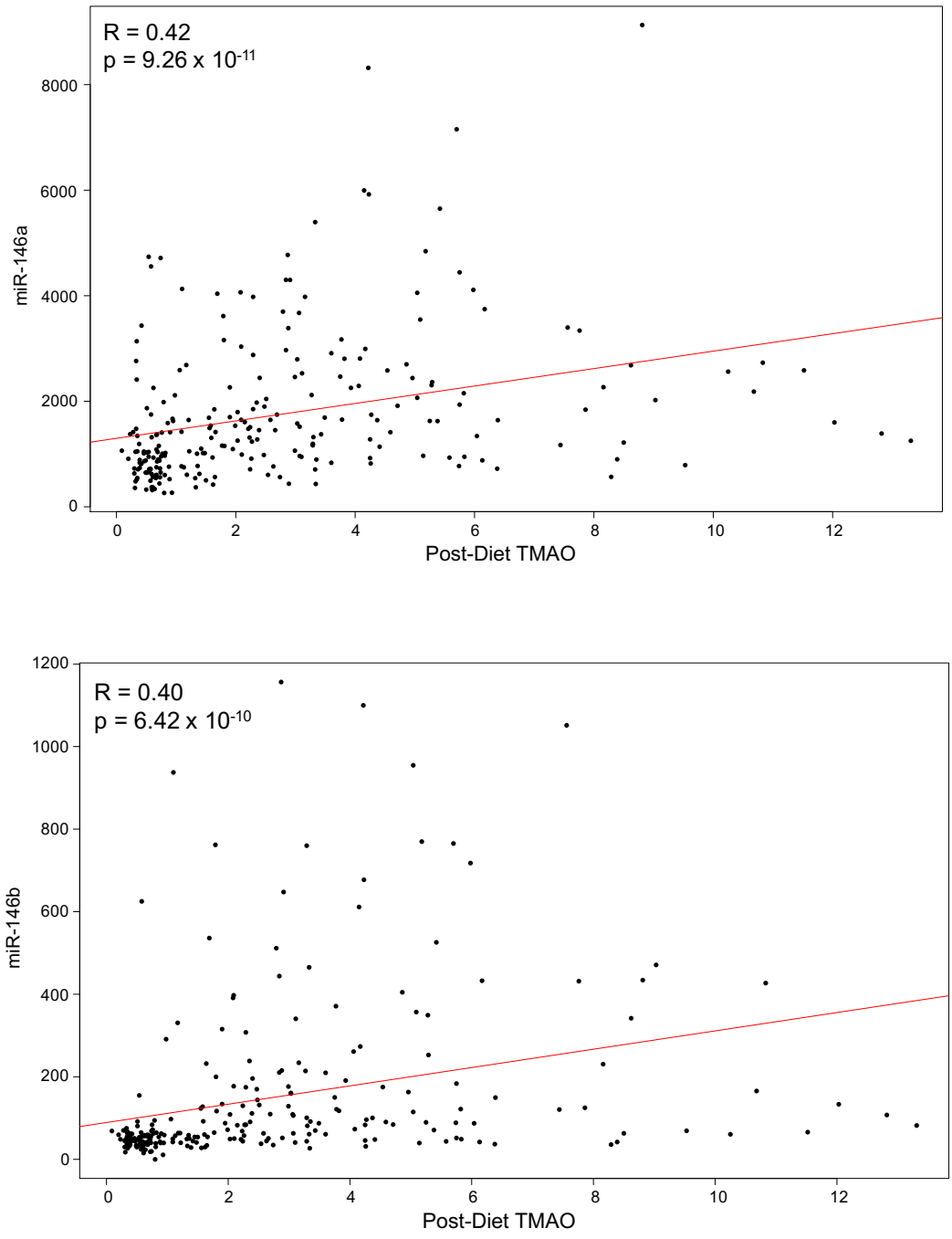

Figure 3.5: miR-146a/b has significant associations with post-diet TMAO in DO mice. Plots indicating correlations between miR-146a-5p and miR-146b-5p, respectively, and post-diet TMAO. R is the bi-weight mid-correlation coefficient. P is Student p-value. miR-146b-5p expression is in reads per million mapped to miRNAs (RPMMM).

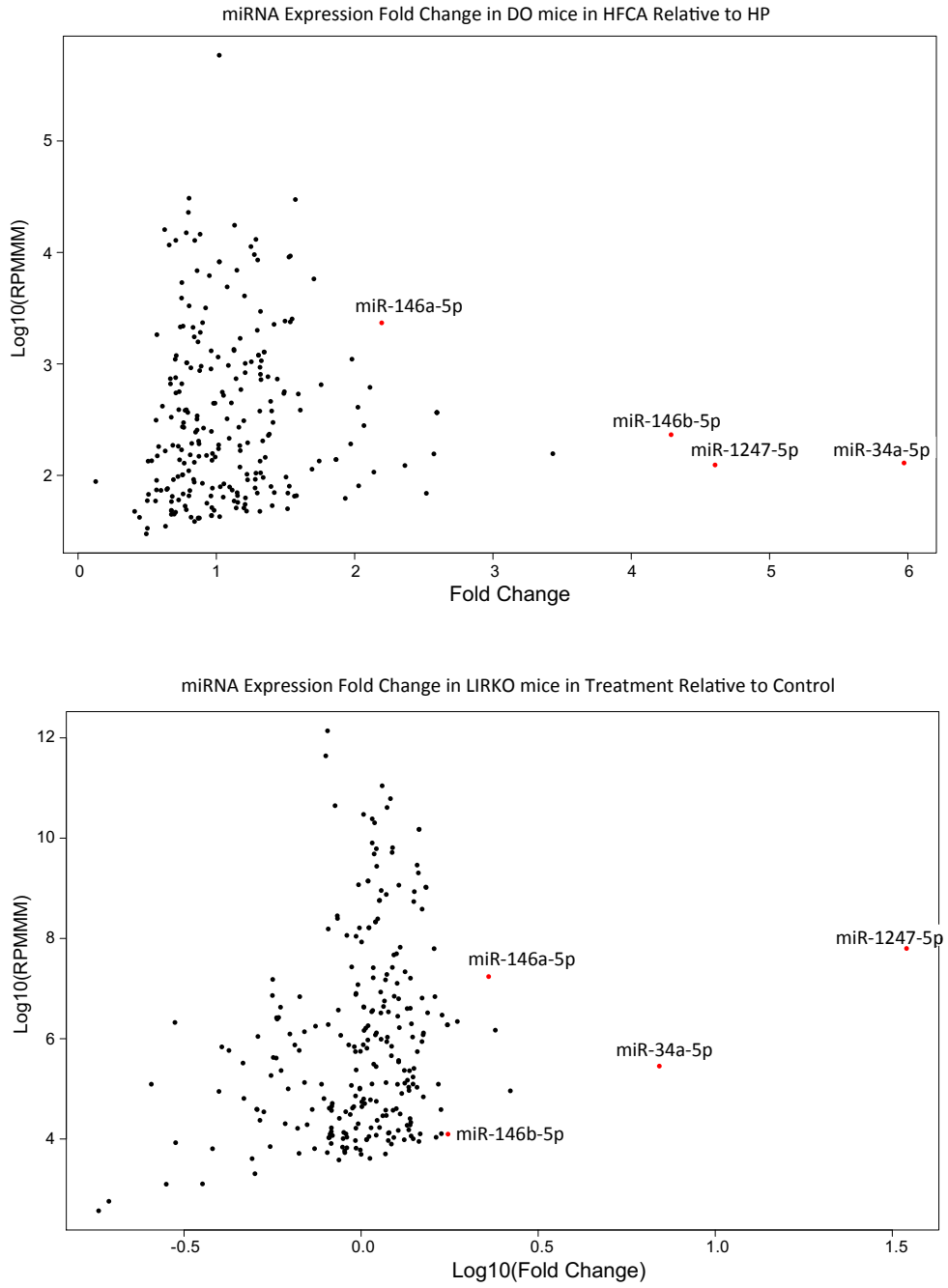

Figure 3.6: miR-146a/b, miR-34a, and miR-1247 are upregulated in DO and LIRKO mouse livers. Top: miRNA expression fold-change in the livers of the DO mice. RPMMMs are $\log_{10}$ transformed. **Bottom:** miRNA expression fold-change in the livers of the LIRKO mice (n = 2 control, n = 3 LIRKO). RPMMMs and fold-change are $\log_{10}$ transformed.

Table 3.5: Mapping statistics for small RNA-seq in African Green Monkey livers. Chow denotes monkeys fed chow diet, HFHC denotes monkeys fed a high fat high cholesterol diet.

Sample name	Chow 1	Chow 2	Chow 3	Chow 4
Total reads	6921450	14800272	22431223	13731169
Trimmed reads	6578170	13894207	20261384	12891396
% Trimmed reads	95.04%	93.88%	90.33%	93.88%
Short reads	142744	572506	1590743	557777
% Short	2.06%	3.87%	7.09%	4.06%
Exact match to genome	3509573	9778874	13617582	8470357
% EM	53.35%	70.38%	67.21%	65.71%
No exact match to genome	3068597	4115333	6643802	4421039
% NEM	46.65%	29.62%	32.79%	34.29%
Total mapped reads	5117952	12539485	18139799	10992897
% Mapped	77.80%	90.25%	89.53%	85.27%
Total mapped to miRs	1997849	8934078	12645667	6966324
% of total mapped to miRs	39.04%	71.25%	69.71%	63.37%
Total mapped to tRNAs	532787	346637	717081	307479
% of total mapped to tRNAs	10.41%	2.76%	3.95%	2.80%

Sample name	HFHC 1	HFHC 2	HFHC 3	HFHC 4	HFHC 5
Total reads	15695910	17844833	19020563	24498194	16772594
Trimmed reads	14732475	15751785	18132684	21045693	15992340
% Trimmed reads	93.86%	88.27%	95.33%	85.91%	95.35%
Short reads	613518	1659082	277830	2842095	371020
% Short	3.91%	9.30%	1.46%	11.60%	2.21%
Exact match to genome	10421206	11245175	11187318	13609384	11003906
% EM	70.74%	71.39%	61.70%	64.67%	68.81%
No exact match to genome	4311269	4506610	6945366	7436309	4988434
% NEM	29.26%	28.61%	38.30%	35.33%	31.19%
Total mapped reads	13102013	14518080	14570075	17626760	14003600
% Mapped	88.93%	92.17%	80.35%	83.75%	87.56%
Total mapped to miRs	9266037	10660269	8962880	11253808	9400558
% of total mapped to miRs	70.72%	73.43%	61.52%	63.85%	67.13%
Total mapped to tRNAs	670051	152697	1265460	954150	1593411
% of total mapped to tRNAs	5.11%	1.05%	8.69%	5.41%	11.38%

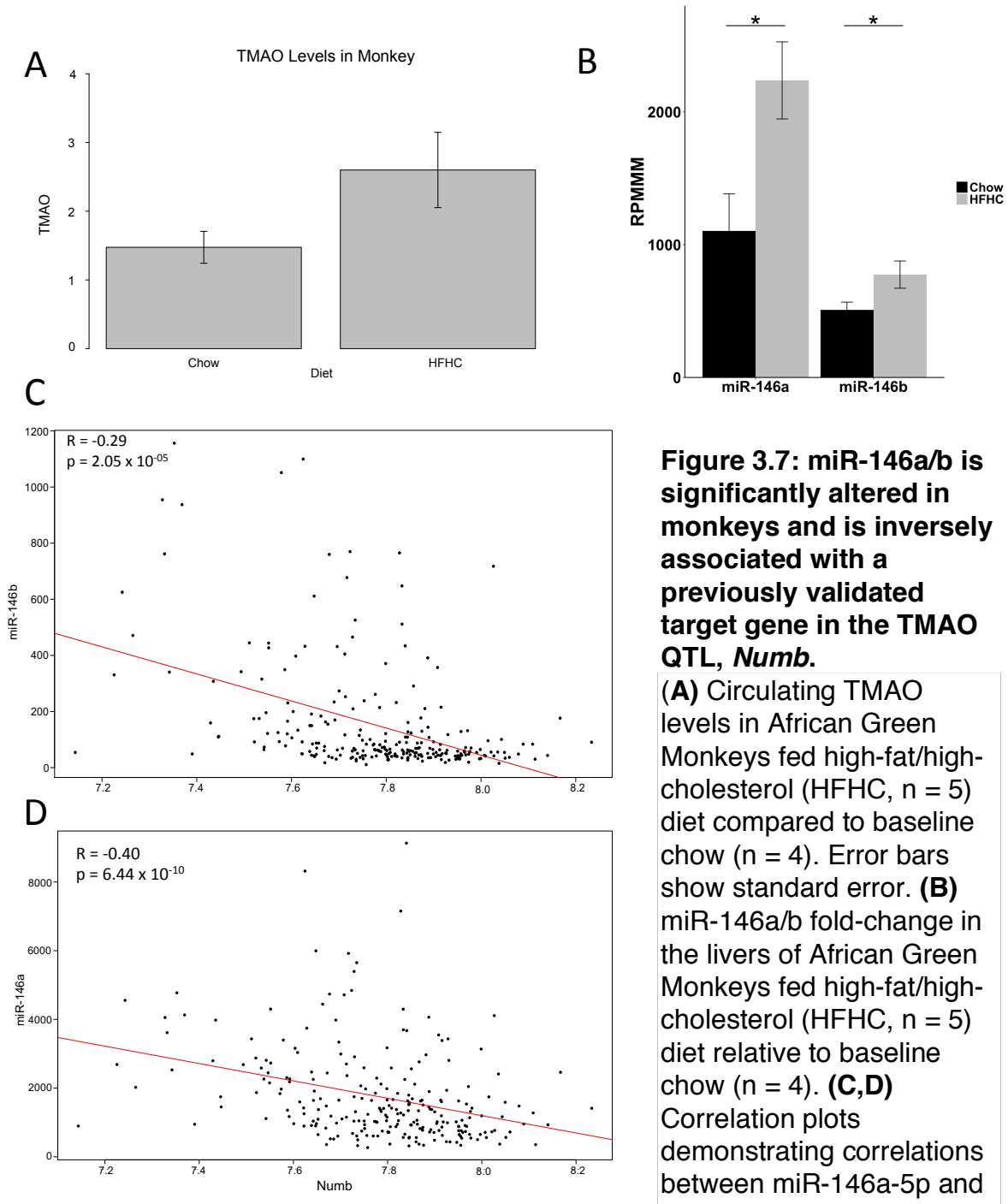

Figure 3.7: miR-146a/b is significantly altered in monkeys and is inversely associated with a previously validated target gene in the TMAO QTL, *Numb*.

(**A)** Circulating TMAO levels in African Green Monkeys fed high-fat/high-cholesterol (HFHC, n = 5) diet compared to baseline chow (n = 4). Error bars show standard error. **(B)** miR-146a/b fold-change in the livers of African Green Monkeys fed high-fat/high-cholesterol (HFHC, n = 5) diet relative to baseline chow (n = 4). **(C,D)** Correlation plots demonstrating correlations between miR-146a-5p and miR-146b-5p, respectively, and *Numb* expression in DO mice. Asterisk (*) denotes $p < 0.05$.

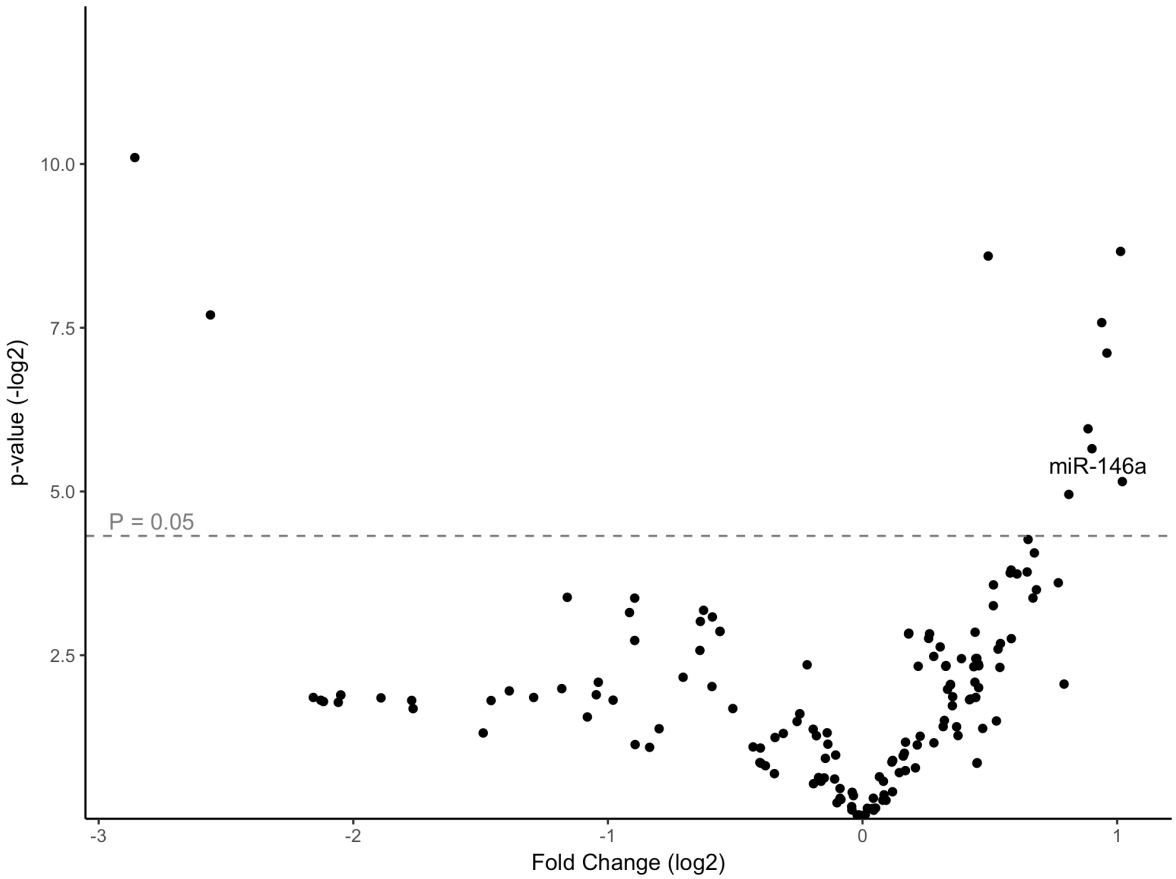

Figure 3.8: Fold change in expression of miRNAs in monkeys fed HFHC diet. miRNA expression log2 of fold change in the livers of African Green Monkeys fed HFHC diet (n=5_ relative to chow-fed monkeys (n=4).

CHAPTER 4: DISCUSSION AND FUTURE DIRECTIONS

Atherosclerosis is a progressive condition that is governed by genetics, environmental factors, or an interaction between the two. Many miRNAs have been implicated in the consequences of cardiometabolic dysfunction such as atherosclerosis and dyslipidemia. Yet, there are significantly fewer studies that convey an unbiased investigation of the global profile of hepatic miRNAs as they fit into co-regulated modules that are most responsive to genetics and diet perturbation; and to my knowledge, none that are in the context of dyslipidemia and cardiometabolic dysfunction. Furthermore, the knowledge concerning genetic components that modulate TMAO levels is still fairly sparse. The work presented here takes a step in the direction of elucidating miRNA regulation of atherosclerosis risk factors, dyslipidemia and TMAO, and provides compelling data as the foundations for future studies.

4.1 MICRORNA MODULES

With expression data and phenotype data in hand, we first asked if there are co-regulated modules associated with cardiometabolic parameters. We identified the hepatic miRNAs that not only are key players in the regulation or response to cardiometabolic dysfunction in a diverse population of mice, but also identified which

miRNAs are co-regulated and potentially function as a module. To clarify, co-regulated modules are assigned when its members have similar expression levels. When forming modules using WGCNA, it is common to use the 'average' method of hierarchical clustering. This method ensures that the distance of each clade to another is the average of each of its members' distances to members in neighboring clades. This method, while helpful for some, may not be the most intuitive when trying to identify modules of co-regulated miRNAs (or genes) since averages can be skewed by outliers. Instead, I used the Ward's method of hierarchical clustering, which is an ANOVA-based approach that puts members with the lowest possible combined standard error together into clades. Since the fundamental idea behind co-regulatory modules is that module members have shared regulatory elements, and may be co-expressed, it makes sense that using Ward's method would be the most advantageous at identifying co-regulated members. The standard error of expression patterns between each miRNA to another across the mouse samples is the lowest possible within a clade, giving a greater likelihood that the expression patterns are extremely similar.

As an answer to our primary question, we found that one miRNA module, the brown module, is associated with post-diet LDL-C and with several gene modules that are also associated with post-diet LDL-C. There are 34 members of the module, 5 of which (miR-199a, miR-181b, miR-27a, miR-24, and miR-21 isomiR) were identified as hub miRNAs. In this context, we refer to the hub miRNAs that are the most highly connected miRNAs in the module. These connections are based on the

summed TOMs (topological overlap measures), interconnectedness of each miRNA within its module, and the correlation to the module eigenmiR. A pair of miRNAs (or module members in general) has a high TOM when they are similarly connected to the same groups of miRNAs in the entire network [Li and Horvath 2009]. The intramodular connectivity is a fuzzy measure of module membership, and those members with high intramodular connectivity are the most central to the module [Horvath and Dong 2008]. The module eigenmiR (normally named eigengene, but was changed to fit a miRNA module) is a measure that is representative of the characteristic expression profile exhibited by module members [Horvath and Dong 2008]. Thus, the use of these three metrics to define the miRNA hubs is extensive and quite thorough.

After identifying them, we are forced to consider what their function is as miRNA hubs. In other words, how are they biologically relevant to the system? The following are possible roles for hub miRNAs with respect to regulating their fellow miRNA module members. One potential role could be as a regulator of a regulator, such as a transcription factor, inhibitor, chromatin modifier, kinase or phosphatase. Even slight fluctuations in the levels of any of these can result in large downstreamt differences in signaling cascades and efforts in transcription [Giorgetti et al. 2010, Ivanovska et al. 2015, McKinsey 2012, Vaquerizas et al. 2009]. Another possible method is by regulating the expression of RNA-binding proteins that may aid in the processing, function, or regulation of the other module members. The upregulation of a miRNA that targets a gene encoding an RNA-binding protein, which in turn

inhibits a direct or indirect inhibitor of another miRNA would surely effect that miRNA's expression. In reality, since miRNAs can be participants in complex regulatory motifs, there may be a combination of the two possibilities at work, or even others that have not been considered here.

With the module of miRNAs that is significantly correlated to a cardiometabolic endpoint/phenotype comes the question as to whether or not the miRNA members are cooperatively regulating the network of genes that promote the phenotype. We found that approximately 30-40% of the genes in the gene modules inversely correlated with post-diet LDL-C are predicted to have conserved target sites for at least one of the miRNA module members. Some of these genes are predicted to be targeted by several of the miRNAs, suggesting that they could also be regulated coordinately. Furthermore, several of the miRNA hubs have been studied and implicated in atherosclerosis or dyslipidemia, two of which, miR-21 and miR-27, have been shown to be regulatory partners for lipid metabolism [Kida et al. 2011]. In addition, a previous study identified miR-21, miR-27 and miR-199 as candidate master regulators of downregulated genes in viral hepatitis, of which the majority were lipid metabolism genes [Selitsky et al. 2015]. To that end, it is very likely that these miRNA module members work together to effect lipid metabolism and homeostasis, but have yet to be specifically studied together in this context. Future studies to validate this miRNA module and the role of its hubs in modulating LDL-C levels may include co-overexpressing the hubs in mice and measuring the effects on expression of genes and other miRNAs, especially those that are miRNA

module members identified in the current work. If the hub miRNAs are truly hubs of the module, the remainder of the 34 miRNAs in the module should have altered expression levels in the same direction as is seen in the work presented here. In addition, some of the genes in the correlated gene modules would also have markedly altered expression levels. In order to deeper understand whether the increased hub miRNA expression is simply responding to the atherogenic diet, or is causal or contributory to the increased circulating LDL-C levels, we could feed mice a HFCA diet, and also treat mice with a cocktail of inhibitors against the five miRNA hubs. If the miRNA hubs are causal for increased circulating LDL-C levels, the mice fed the HFCA diet and treated with the inhibitors should have circulating LDL-C levels that are reduced in comparison to the mice fed the HFCA diet alone.

Although the primary focus of this portion of the dissertation has been on miRNAs with a secondary view of gene modules, gene expression is also a major piece of the puzzle that is important to address. In the present work, we utilized microarray technology to measure gene expression differences in the DO mice, and used this information to perform a basic differential expression analysis and co-expression network analysis. While the microarray is informative, performing sequencing would be even more informative. The microarray plate incorporated SNPs in the probes so that the majority of mRNAs from the genetically diverse DO mice would be able to hybridize to them. However, there is a chance that we are inadvertently missing transcripts, such as those that are alternatively spliced due to some variant. Performing RNA-seq in the DO mice would give us more information

about the transcript differences between mice with different genotypes, and furthermore, a more sensitive account of transcript abundance. With sequencing data for both miRNAs and mRNAs, we would also be able to identify specific variants within transcripts and relate them to differences in miRNA binding between alleles.

4.2 TMAO QTL

In an effort to identify host genetic regulation of TMAO, we identified a novel QTL for plasma TMAO levels located on Chromosome 12. We focused on the Chromosome 12 locus because it was present in mapping for pre- and post-diet TMAO levels. The Chromosome 12 locus is almost perfectly overlapping between pre- and post-diet TMAO, and the two share the same founder allele effects with CAST/EiJ and PWK/PhJ alleles associated with high plasma TMAO levels. Notably, we found neither a significant nor suggestive peak on Chromosome 3, for which a GWAS locus and overlapping eQTL were identified in a previous study [Hartiala et al. 2014]. While this result was not recapitulated in our hands, this does not necessarily mean that either result is false. Major differences in the type of mice used could be the cause for the vast difference in results. Haritala et al. used an all-male cohort of HMDP mice, whereas we used an all-female cohort of DO mice. The differences caused by gender along with the differences between the two mouse resources are two possible sources of discrepancy.

The QTL peak on Chromosomes 14 and X did not replicate in post-diet TMAO QTL mapping. The pre-diet TMAO levels were measured in the same mice as were the post-diet TMAO levels, so the only difference between the two is dietary intervention with either HP or HFCA diet. This suggests that the regulatory elements on Chromosomes 14 and X seem to become inactive or less effective in modulating plasma TMAO levels after diet treatment. This may indicate that interaction between diet and genetics is heavily influential to these loci, and possibly causes a rewiring of the genetic architecture for some loci, and not at all for others.

Within the extended TMAO QTL region are 80 genes, 63 of which have probes on the microarray so their expression levels were measured. We also performed eQTL mapping on these 63 genes and found that only 14 of the 63 had overlapping eQTL, or rather, eQTL peaks that are within 1Mb of the TMAO QTL. Only 7 of the 14 genes had consistent founder allele frequencies that effected pre- and post-diet TMAO levels. However, 4 of the 7, and a few other of the 14 genes, were significantly correlated with TMAO levels. It is worth noting that even though the eQTL overlap with QTL for both pre-diet and post-diet TMAO levels, merely a subset of the genes with overlapping eQTL were significantly correlated with post-diet TMAO, and not pre-diet TMAO. This may indicate that either 1) TMAO levels before diet are modulated by different genes from post-diet TMAO that happen to be located within the same region, 2) TMAO before and after diet treatment are modulated by the same one or more genes within the region, but diet interacting with the underlying genetics exacerbates their effects, or 3) this is further evidence of a

rewiring of the genetic architecture controlling pre-diet TMAO levels versus post-diet levels. Option 1 is highly unlikely, but cannot be completely ruled out because, with the current data, we cannot analyze all 80 genes in the locus. A likely reason is option 3 given that the non-replicated Chromosome 14 and X QTL are seen in pre-diet TMAO and not post-diet TMAO; this is perhaps evidence of epistasis of genes in different loci. This can be coupled with option 2 where the diet interactions initiate the alteration in architecture and exacerbate the effects of the Chromosome 12 locus gene(s).

Although there are only 4 genes that are within the TMAO QTL region that have consistent allele frequencies, have overlapping significant eQTL, and are correlated with TMAO, it is worth considering that there are possibly other genes that may contribute to modulating TMAO levels. We identified one of the 4, *Numb*, as a likely candidate gene as it has been shown to be regulated by NPC1L1, and necessary for cholesterol absorption in the intestine [Lei et al. 2014, Wei et al. 2014]. Even with its connection to cholesterol homeostasis, it is possible that some other gene indirectly effects TMAO levels through influential actions on genes like *Numb*, and not *Numb* alone mediating TMAO. One could argue that the other gene's indirect regulatory effects would show up as a peak in the eQTL mapping. After all, there are a total of 47 transcripts (including the 14) with significant eQTL that overlap the Chromosome 12 locus. Also, *Numb* has a single highly significant peak at Chromosome 12, and three suggestive peaks on Chromsomes 3, 10, and 17. The 'other' gene could contribute such minor effects that we are too underpowered for

the results to be clearly detected or reach significance. There is also the likelihood that the 'other's' effects may be at the post-transcriptional or post-translational level, which would not be recognized by the microarray. Further investigation, perhaps using a *Numb-/-* mouse, is necessary to tease apart the potentially complex mechanism.

Since the DO mice have so much genetic variation between them, performing other types of sequencing in them would be advantageous for further studying how genetic variation effects expression, and phenotypes. Sequencing techniques such as CHIP-seq would provide further information about chromatin, and more interestingly, information about transcription factor binding differences between alleles. Relating transcription factor affinity for their respective binding sites with genotype would add another layer of information that would help elucidate the regulatory effects that variants can have on genes and miRNAs, and could help clarify any overlaps in the QTL and eQTL data.

4.3 MIR-146 AND TMAO

In addition to host genetics at the DNA level, we wanted to identify any candidate miRNA regulators of circulating TMAO levels. MiR-146a and miR-146b were identified as two of the most highly upregulated miRNAs in the livers of the DO mice fed a HFCA diet, and their expression is positively correlated with post-diet TMAO levels. Additionally, they are also two of the most highly upregulated miRNAs in the livers of LIRKO mice, which are known to have elevated plasma TMAO levels.

Lastly, it is the most highly upregulated miRNA in the livers of African Green Monkeys fed a HFHC diet who also have, on average, elevated circulating TMAO levels in comparison to their counterparts fed a chow diet. To my knowledge, miR-146 has not previously been connected with circulating TMAO levels. However, the evidence presented here in three independent animal models is highly associative and suggests that miR-146 is responsive to an atherogenic/dyslipidemic diet, and may be a regulator of TMAO.

Although the evidence is compelling, it is only associative, and more experimentation is required to solidify the notion of the regulatory relationship. To begin to definitively test for a regulatory relationship, miR-146 can be overexpressed and TMAO measured in a mouse model in an effort to recapitulate the results observed in the current study. In a second experiment, a miR-146 KO mouse can be used to determine if there is an opposite effect on TMAO levels to what is seen in miR-146 overexpression, and miR-146 mimic can later be injected or expressed in the KO mouse in an attempt to create a rescue effect. The results from such experiments would determine if TMAO is indeed under regulation by miR-146, however further experimentation would be necessary to elucidate the mechanism by which this occurs. Sequencing experiments to measure gene expression in the overexpression and knockout animals would allow us to identify any genes that are altered in both experiments, especially those that are altered in opposite directions between the two. If *Numb* is one such gene that is altered in both experiments, it would lend evidence to support our findings, and would catalyze another series of

experiments to validate it as a modulator of TMAO that is under the regulatory control of miR-146 in the liver.

We identified that miR-146 is predicted to target *Numb*. *Numb* is correlated with miR-146 and inversely correlated with post-diet TMAO levels. Of note, miR-146 is part of the brown miRNA module that was identified in Chapter 2. Numb is also a member of one of the gene modules that is inversely correlated with the brown miRNA module. Not only is this a connection between the two bodies of work, but also suggests that there may be other connections between non-hub miRNA module members and cardiometabolic traits worth investigating.

Although we identified Numb as a possible candidate that is mediating the relationship between miR-146 and TMAO, we also cannot deny that there are at least 5 other genes within the TMAO QTL region that miR-146 is predicted to target (Zfyve1, Psen1, Elmsan1, Dlst, Gpatch2l). Any one of these, or any combination of these, can be targeted by miR-146 and effect TMAO levels. Indeed, as the theme of this body of work is oriented toward a collective effort in regulation, and since TMAO is a quantitative trait, I am inclined to suspect that the latter explanation would be valid.

While I have focused here on liver molecular traits and their regulatory relationship with TMAO, it is important not to disregard the contributions of the gut and gut microbiota. Although we have endeavored to understand the relationships between genetic variation, hepatic gene and miRNA expression, and circulating TMAO levels, understanding the relationship between genotype, expression profiles,

and microbiota composition is also extremely relevant. To that end, fecal matter and intestine samples were taken from the mice in the DO cohort. Since TMAO begins as TMA, which is produced in the intestine, performing sequencing in these samples from the DO mice to have gene, miRNA, and microbiota composition profiles, and perform genetic mapping and correlation analyses would provide greater insight into the regulatory systems behind circulating TMAO, and help to fill in any gaps or confounding results we have.

The work presented here was generated by using computational analysis on experimental data from mice, and provides largely associative and compelling data that is quite sufficient to be expanded upon. The results and data presented here have contributed another piece of the cardiometabolic dysfunction regulatory puzzle, and can act as strong rationale and basis for future work.

APPENDIX A: OVERSIZED TABLES

Table A.1: Predicted target genes of the brown miRNA module members. Target prediction analysis performed on the inversely correlated gene CRMs using only the brown miRNA module members. Target predictions listed are conserved conserved in at least 2 species.

Module Color	Gene Symbol	miRNAs
lightcyan	Ppp1r15b	mmu-miR-200a-3p,mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-29a-3p
lightcyan	Aamp	mmu-miR-342-3p
lightcyan	Pan2	mmu-miR-29a-3p
lightcyan	Agpat3	mmu-miR-24-3p,mmu-miR-27a-3p
lightcyan	Trim41	mmu-let-7e-5p,mmu-let-7i-5p
lightcyan	Nfe2l1	mmu-miR-199a-5p
lightcyan	Fam134c	mmu-miR-29a-3p,mmu-miR-125a-5p,mmu-miR-27a-3p
lightcyan	Numb	mmu-miR-146a-5p,mmu-miR-146b-5p,mmu-miR-143-3p
lightcyan	Arel1	mmu-miR-199b-3p,mmu-miR-199a-3p,mmu-miR-24-3p,mmu-miR-125a-5p
lightcyan	Sel1l	mmu-miR-125a-5p,mmu-miR-181b-5p
lightcyan	Foxn3	mmu-miR-125a-5p,mmu-miR-200a-3p,mmu-miR-27a-3p,mmu-miR-200a-3p
lightcyan	Desi1	mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-199a-5p
lightcyan	Ccnt1	mmu-miR-199a-3p,mmu-miR-199b-3p,mmu-miR-27a-3p,mmu-miR-181b-5p
lightcyan	Rhot2	mmu-miR-146a-5p,mmu-miR-146b-5p,mmu-miR-125a-5p
lightcyan	Cyb5a	mmu-miR-200a-3p
lightcyan	Ak3	mmu-miR-29a-3p
lightcyan	Surf2	mmu-miR-532-5p,mmu-miR-146a-5p,mmu-miR-146b-5p
lightcyan	Rprd1b	mmu-miR-181b-5p
lightcyan	Atg13	mmu-miR-199a-5p
lightcyan	Ube2r2	mmu-miR-214-3p,mmu-miR-125a-5p
lightcyan	Nfia	mmu-miR-29a-3p,mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-200a-3p,mmu-miR-27a-3p,mmu-miR-199b-3p,mmu-miR-199a-3p
lightcyan	Bsdc1	mmu-miR-29a-3p
lightcyan	Abhd11	mmu-miR-125a-5p,mmu-let-7i-5p,mmu-let-7e-5p
lightcyan	Edem1	mmu-let-7e-5p,mmu-let-7i-5p,mmu-miR-125a-5p,mmu-miR-125a-5p
lightcyan	Necap1	mmu-miR-143-3p,mmu-miR-27a-3p

lightcyan	Phb2	mmu-miR-24-3p,mmu-miR-200a-3p
lightcyan	Pex5	mmu-miR-29a-3p,mmu-miR-200a-3p
lightcyan	Akt2	mmu-miR-200c-3p,mmu-miR-200b-3p
lightcyan	Egln2	mmu-let-7e-5p,mmu-let-7i-5p
lightcyan	Amfr	mmu-miR-24-3p,mmu-miR-29a-3p,mmu-miR-200b-3p,mmu-miR-200c-3p
lightcyan	Ormdl2	mmu-miR-532-5p
magenta	Pcmtd1	mmu-miR-199a-3p,mmu-miR-199b-3p,mmu-miR-200c-3p,mmu-miR-200b-3p
magenta	Hspd1	mmu-miR-27a-3p
magenta	Tmem183a	mmu-miR-29a-3p
magenta	Shprh	mmu-miR-27a-3p,mmu-miR-29a-3p,mmu-miR-501-3p
magenta	Ttc19	mmu-miR-501-3p,mmu-miR-199a-5p
magenta	Rhot1	mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-27a-3p
magenta	Ube2b	mmu-miR-181b-5p,mmu-miR-200c-3p,mmu-miR-200b-3p
magenta	Med1	mmu-miR-146a-5p,mmu-miR-146b-5p,mmu-miR-200b-3p,mmu-miR-200c-3p
magenta	1110057K04Rik	mmu-miR-342-3p,mmu-miR-200a-3p
magenta	Ppm1a	mmu-miR-199a-3p,mmu-miR-199b-3p
magenta	Synj2bp	mmu-miR-200b-3p,mmu-miR-200c-3p
magenta	Smek1	mmu-miR-125a-5p,mmu-miR-125a-5p
magenta	B3galnt2	mmu-miR-125a-5p
magenta	C78339	mmu-miR-200a-3p,mmu-miR-27a-3p,mmu-miR-200c-3p,mmu-miR-200b-3p
magenta	Srek1ip1	mmu-let-7e-5p,mmu-let-7i-5p
magenta	Ngly1	mmu-let-7e-5p,mmu-let-7i-5p
magenta	Samd8	mmu-miR-200a-3p
magenta	Naa30	mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-143-3p,mmu-miR-199a-5p
magenta	Ghitm	mmu-miR-181b-5p
magenta	Rnf139	mmu-miR-27a-3p
magenta	Ndufb9	mmu-miR-143-3p,mmu-miR-199b-3p,mmu-miR-199a-3p,mmu-miR-200a-3p,mmu-miR-214-3p,mmu-miR-532-5p
magenta	Crkl	mmu-miR-200c-3p,mmu-miR-200b-3p
magenta	Tomm70a	mmu-miR-24-3p,mmu-miR-200c-3p,mmu-miR-200b-3p
magenta	Gabpa	mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-146a-5p,mmu-miR-146b-5p,mmu-miR-501-3p
magenta	1600012H06Rik	mmu-miR-27a-3p
magenta	Ypel5	mmu-miR-200a-3p
magenta	Ppm1b	mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-181b-5p
magenta	Pfdn6	mmu-miR-181b-5p

magenta	Ttr	mmu-miR-200a-3p
magenta	Ammecr1l	mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-29a-3p,mmu-miR-27a-3p
magenta	Ndfip1	mmu-miR-143-3p
magenta	Pcgf5	mmu-miR-199a-3p,mmu-miR-199b-3p
magenta	Arl5b	mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-200a-3p,mmu-miR-146b-5p,mmu-miR-146a-5p,mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-29a-3p
magenta	Mtx2	mmu-miR-214-3p
magenta	Ythdf3	mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-181b-5p,mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-29a-3p
magenta	Mynn	mmu-miR-200a-3p
magenta	Nmd3	mmu-miR-146a-5p,mmu-miR-146b-5p
magenta	Mrpl24	mmu-miR-200b-3p,mmu-miR-200c-3p
magenta	Usp33	mmu-miR-181b-5p
magenta	Ankrd13c	mmu-miR-181b-5p,mmu-miR-29a-3p,mmu-miR-501-3p,mmu-miR-199a-3p,mmu-miR-199b-3p
magenta	Golim4	mmu-miR-199b-3p,mmu-miR-199a-3p,mmu-miR-200b-3p,mmu-miR-200c-3p
magenta	Ttpa	mmu-miR-143-3p
magenta	Stx17	mmu-let-7e-5p,mmu-let-7i-5p,mmu-miR-29a-3p,mmu-miR-24-3p
magenta	Gpbp1l1	mmu-miR-199b-3p,mmu-miR-199a-3p
magenta	Zyg11b	mmu-miR-200a-3p,mmu-miR-532-5p
magenta	Osbpl9	mmu-miR-125a-5p
magenta	Vkorc1l1	mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-181b-5p,mmu-miR-342-3p
magenta	Hmgb1	mmu-miR-181b-5p,mmu-miR-200a-3p
magenta	Met	mmu-miR-181b-5p,mmu-miR-199a-3p,mmu-miR-199b-3p,mmu-miR-27a-3p
magenta	Crebl2	mmu-miR-24-3p,mmu-miR-24-3p
magenta	Zranb1	mmu-miR-27a-3p
magenta	Tmem66	mmu-miR-199a-5p
magenta	4933411K20Rik	mmu-miR-200a-3p,mmu-miR-27a-3p,mmu-miR-200a-3p
magenta	Zfp617	mmu-miR-199a-5p
magenta	Hook3	mmu-miR-29a-3p,mmu-miR-200c-3p,mmu-miR-200b-3p
magenta	Pgrmc1	mmu-let-7i-5p,mmu-let-7e-5p
tan	Lypla1	mmu-let-7e-5p,mmu-let-7i-5p,mmu-miR-200a-3p,mmu-miR-29a-3p
tan	Stradb	mmu-miR-199a-5p,mmu-miR-24-3p
tan	Timm13	mmu-miR-24-3p
tan	Ube2g1	mmu-miR-146b-5p,mmu-miR-146a-5p,mmu-miR-125a-5p
tan	Gphn	mmu-miR-200a-3p

tan	Dlst	mmu-let-7e-5p,mmu-let-7i-5p,mmu-miR-146b-5p,mmu-miR-146a-5p
tan	Ndufs4	mmu-miR-27a-3p
tan	Ndrg2	mmu-miR-181b-5p
tan	Deptor	mmu-miR-146a-5p,mmu-miR-146b-5p
tan	Slc38a4	mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-27a-3p
tan	Clpp	mmu-miR-27a-3p
tan	Slc25a46	mmu-miR-532-5p
tan	Nr3c1	mmu-miR-181b-5p,mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-200a-3p
tan	Afg3l2	mmu-miR-27a-3p,mmu-miR-181b-5p
tan	Cdc37l1	mmu-miR-125a-5p
tan	Acvr2a	mmu-miR-181b-5p,mmu-let-7e-5p,mmu-let-7i-5p,mmu-miR-200a-3p,mmu-miR-125a-5p,mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-27a-3p,mmu-miR-29a-3p,mmu-miR-200a-3p,mmu-miR-199a-3p,mmu-miR-199b-3p
tan	Atrn	mmu-miR-29a-3p
tan	Gid8	mmu-miR-342-3p,mmu-miR-29a-3p,mmu-miR-143-3p
tan	Aco1	mmu-miR-200a-3p
tan	Mycbp	mmu-miR-27a-3p,mmu-let-7e-5p,mmu-let-7i-5p,mmu-miR-181b-5p
tan	Wipi2	mmu-let-7i-5p,mmu-let-7e-5p
tan	Hadha	mmu-miR-27a-3p
tan	Grsf1	mmu-miR-532-5p,mmu-miR-125a-5p
tan	Ndufb2	mmu-miR-200b-3p,mmu-miR-200c-3p
tan	Crebzf	mmu-miR-143-3p
tan	Tmem135	mmu-miR-200a-3p,mmu-miR-199a-5p
tan	Slc7a6os	mmu-miR-181b-5p
tan	Wdr59	mmu-miR-24-3p
tan	Glyctk	mmu-miR-214-3p
tan	Taz	mmu-miR-125a-5p,mmu-miR-125a-5p
turquoise	Atg16l1	mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Ccnt2	mmu-miR-200a-3p,mmu-miR-27a-3p,mmu-miR-29a-3p,mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Kdm5b	mmu-miR-29a-3p
turquoise	Tor1aip2	mmu-let-7e-5p,mmu-let-7i-5p
turquoise	Scyl3	mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Tada1	mmu-miR-27a-3p,mmu-miR-200a-3p
turquoise	Cox20	mmu-miR-125a-5p,mmu-miR-199a-5p
turquoise	Nr5a2	mmu-miR-27a-3p,mmu-miR-200b-3p,mmu-miR-200c-3p
turquoise	Heca	mmu-miR-181b-5p

turquoise	Man1a	mmu-miR-27a-3p
turquoise	Slc25a3	mmu-miR-200a-3p
turquoise	Nudt4	mmu-miR-200b-3p,mmu-miR-200c-3p
turquoise	Rab21	mmu-miR-200c-3p,mmu-miR-200b-3p
turquoise	Frs2	mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-532-5p
turquoise	Sertad2	mmu-miR-27a-3p,mmu-miR-532-5p,mmu-miR-181b-5p
turquoise	Fnip1	mmu-let-7e-5p,mmu-let-7i-5p
turquoise	Alkbh5	mmu-miR-214-3p
turquoise	Timm22	mmu-miR-199a-3p,mmu-miR-199b-3p
turquoise	Sdf2	mmu-miR-200a-3p
turquoise	Mbtd1	mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-200a-3p,mmu-miR-29a-3p,mmu-miR-27a-3p
turquoise	Msl1	mmu-miR-125a-5p
turquoise	Helz	mmu-miR-29a-3p,mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Nf2	mmu-miR-146b-5p,mmu-miR-146a-5p
turquoise	Ctns	mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Gosr1	mmu-miR-125a-5p,mmu-miR-146a-5p,mmu-miR-146b-5p
turquoise	Rps6kb1	mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-181b-5p,mmu-miR-27a-3p
turquoise	Dynll2	mmu-miR-27a-3p,mmu-miR-181b-5p,mmu-miR-143-3p
turquoise	Msi2	mmu-miR-143-3p,mmu-let-7e-5p,mmu-let-7i-5p,mmu-miR-27a-3p,mmu-miR-181b-5p
turquoise	Fbxl20	mmu-miR-200a-3p,mmu-miR-27a-3p,mmu-miR-125a-5p,mmu-miR-29a-3p
turquoise	Ccdc43	mmu-miR-199a-5p
turquoise	Smurf2	mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-29a-3p
turquoise	Exoc7	mmu-miR-29a-3p
turquoise	Slc39a9	mmu-miR-125a-5p,mmu-miR-200a-3p,mmu-miR-29a-3p
turquoise	Ccnk	mmu-miR-181b-5p,mmu-miR-27a-3p
turquoise	Yy1	mmu-miR-200a-3p,mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Eif5	mmu-miR-27a-3p,mmu-miR-200a-3p
turquoise	Atg2b	mmu-miR-143-3p
turquoise	2010107E04Rik	mmu-miR-24-3p
turquoise	Dusp22	mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Secisbp2	mmu-miR-181b-5p
turquoise	Mier3	mmu-miR-199a-5p,mmu-miR-27a-3p,mmu-miR-29a-3p,mmu-miR-181b-5p
turquoise	Slc38a9	mmu-miR-125a-5p,mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-143-3p
turquoise	Iqgap2	mmu-miR-181b-5p,mmu-miR-199a-3p,mmu-miR-199b-3p
turquoise	Nek4	mmu-miR-24-3p

turquoise	Ankrd28	mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-501-3p,mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Mapk8	mmu-miR-199a-3p,mmu-miR-199b-3p,mmu-miR-29a-3p,mmu-let-7e-5p,mmu-let-7i-5p
turquoise	Rictor	mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-27a-3p,mmu-miR-342-3p
turquoise	Tnrc6b	mmu-miR-322-3p,mmu-miR-29a-3p,mmu-miR-200a-3p,mmu-miR-181b-5p
turquoise	Letmd1	mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Yaf2	mmu-miR-200a-3p,mmu-miR-214-3p,mmu-miR-199a-5p,mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Klhl24	mmu-miR-125a-5p
turquoise	Ubxn7	mmu-miR-532-5p,mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-146b-5p,mmu-miR-146a-5p,mmu-miR-29a-3p
turquoise	Pcyt1a	mmu-miR-24-3p
turquoise	Osbpl11	mmu-miR-200a-3p,mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-27a-3p
turquoise	Zfp318	mmu-let-7e-5p,mmu-let-7i-5p
turquoise	Kat2b	mmu-miR-181b-5p,mmu-miR-27a-3p
turquoise	Zeb1	mmu-miR-199a-3p,mmu-miR-199b-3p,mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-200a-3p
turquoise	Cep120	mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-200a-3p,mmu-miR-24-3p
turquoise	March5	mmu-miR-200a-3p
turquoise	Btrc	mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-199a-5p
turquoise	Sorbs1	mmu-miR-342-3p
turquoise	Fbxw4	mmu-miR-125a-5p
turquoise	Nmt2	mmu-miR-181b-5p
turquoise	Mllt10	mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Arfgap2	mmu-miR-342-3p,mmu-let-7e-5p,mmu-let-7i-5p
turquoise	Stx16	mmu-miR-27a-3p,mmu-miR-29a-3p,mmu-miR-200a-3p
turquoise	Gtf3c4	mmu-miR-29a-3p
turquoise	Fbxw2	mmu-miR-146a-5p,mmu-miR-146b-5p,mmu-miR-29a-3p
turquoise	Rab14	mmu-miR-27a-3p,mmu-miR-29a-3p
turquoise	Tlk1	mmu-miR-532-5p,mmu-miR-181b-5p
turquoise	Srp14	mmu-miR-146b-5p,mmu-miR-146a-5p
turquoise	Zscan29	mmu-miR-146a-5p,mmu-miR-146b-5p,mmu-miR-125a-5p
turquoise	Pdrg1	mmu-miR-342-3p
turquoise	Chd6	mmu-miR-27a-3p,mmu-miR-143-3p,mmu-miR-532-5p
turquoise	Taf4a	mmu-miR-200b-3p,mmu-miR-200c-3p
turquoise	Ythdf1	mmu-miR-29a-3p
turquoise	Supt20	mmu-miR-200b-3p,mmu-miR-200c-3p
turquoise	Setdb1	mmu-miR-29a-3p

turquoise	Prkacb	mmu-miR-200a-3p,mmu-miR-200b-3p,mmu-miR-200c-3p
turquoise	Pnisr	mmu-miR-181b-5p,mmu-miR-27a-3p
turquoise	Dcaf10	mmu-miR-29a-3p,mmu-miR-125a-5p
turquoise	Pum1	mmu-miR-181b-5p
turquoise	Casp9	mmu-miR-199b-3p,mmu-miR-199a-3p
turquoise	AI314180	mmu-miR-29a-3p
turquoise	Ptprd	mmu-miR-24-3p,mmu-let-7e-5p,mmu-let-7i-5p,mmu-miR-29a-3p,mmu-miR-200a-3p,mmu-miR-143-3p
turquoise	Nfib	mmu-miR-24-3p,mmu-miR-532-5p,mmu-miR-29a-3p,mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-27a-3p,mmu-miR-125a-5p,mmu-miR-181b-5p
turquoise	S100pbp	mmu-miR-200a-3p,mmu-miR-200c-3p,mmu-miR-200b-3p
turquoise	Szrd1	mmu-miR-27a-3p,mmu-miR-125a-5p
turquoise	Rufy3	mmu-let-7e-5p,mmu-let-7i-5p
turquoise	Atp6v0a2	mmu-miR-24-3p,mmu-miR-200c-3p,mmu-miR-200b-3p
turquoise	Zfp664	mmu-miR-181b-5p
turquoise	Sfswap	mmu-miR-27a-3p
turquoise	Rabgef1	mmu-miR-181b-5p
turquoise	Cpsf4	mmu-miR-214-3p,mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Cdk8	mmu-miR-181b-5p,mmu-miR-27a-3p,mmu-miR-200a-3p
turquoise	Mphosph9	mmu-miR-200a-3p
turquoise	Stx2	mmu-miR-200a-3p
turquoise	Gtf2i	mmu-miR-27a-3p,mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Mepce	mmu-miR-27a-3p,mmu-miR-342-3p
turquoise	C1galt1	mmu-miR-27a-3p,mmu-miR-200a-3p,mmu-let-7e-5p,mmu-let-7i-5p,mmu-miR-181b-5p
turquoise	Cped1	mmu-let-7i-5p,mmu-let-7e-5p
turquoise	Luc7l2	mmu-miR-146a-5p,mmu-miR-146b-5p
turquoise	C87436	mmu-miR-200a-3p
turquoise	Mtmr14	mmu-miR-24-3p
turquoise	March8	mmu-miR-199a-5p,mmu-miR-200c-3p,mmu-miR-200b-3p
turquoise	Etnk1	mmu-miR-200a-3p,mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-199a-3p,mmu-miR-199b-3p,mmu-miR-181b-5p
turquoise	Wasl	mmu-let-7e-5p,mmu-let-7i-5p,mmu-miR-181b-5p
turquoise	Abtb1	mmu-miR-125a-5p
turquoise	Zfyve20	mmu-miR-200c-3p,mmu-miR-200b-3p
turquoise	Tmcc1	mmu-miR-27a-3p,mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-181b-5p
turquoise	Erc1	mmu-miR-29a-3p
turquoise	Tnrc6a	mmu-miR-27a-3p,mmu-miR-146a-5p,mmu-miR-146b-5p
turquoise	Inpp5a	mmu-let-7e-5p,mmu-let-7i-5p,mmu-miR-181b-5p

turquoise	Ruvbl2	mmu-miR-27a-3p
turquoise	Chd2	mmu-miR-27a-3p,mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-200a-3p
turquoise	2210018M11Rik	mmu-miR-181b-5p,mmu-miR-29a-3p
turquoise	Uvrag	mmu-miR-125a-5p
turquoise	Mrpl48	mmu-miR-146a-5p,mmu-miR-146b-5p
turquoise	Knop1	mmu-miR-125a-5p
turquoise	Otud4	mmu-miR-143-3p,mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-29a-3p,mmu-miR-501-3p,mmu-miR-181b-5p
turquoise	Papd5	mmu-miR-181b-5p,mmu-miR-200b-3p,mmu-miR-200c-3p
turquoise	Polr2c	mmu-miR-24-3p
turquoise	Nfatc3	mmu-miR-29a-3p
turquoise	Ap1g1	mmu-miR-199a-5p,mmu-miR-29a-3p,mmu-miR-24-3p,mmu-miR-181b-5p
turquoise	Aars	mmu-miR-24-3p
turquoise	Gabarapl2	mmu-miR-200a-3p,mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-532-5p
turquoise	Spg7	mmu-miR-24-3p
turquoise	Golga7	mmu-let-7i-5p,mmu-let-7e-5p,mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-29a-3p
turquoise	Nr2f6	mmu-miR-27a-3p
turquoise	Gab1	mmu-miR-29a-3p,mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-200a-3p,mmu-miR-27a-3p
turquoise	Cmtm4	mmu-miR-24-3p,mmu-miR-214-3p
turquoise	Fam96b	mmu-miR-200a-3p
turquoise	Cog8	mmu-miR-199a-5p
turquoise	Maf	mmu-miR-143-3p,mmu-miR-200c-3p,mmu-miR-200b-3p,mmu-miR-125a-5p,mmu-miR-200a-3p
turquoise	Dennd4a	mmu-miR-29a-3p,mmu-miR-181b-5p
turquoise	Zfp280d	mmu-miR-181b-5p
turquoise	Dcps	mmu-miR-214-3p
turquoise	Zbtb38	mmu-miR-200b-3p,mmu-miR-200c-3p,mmu-miR-125a-5p
turquoise	Arih2	mmu-miR-199a-5p,mmu-miR-200c-3p,mmu-miR-200b-3p
turquoise	Tab3	mmu-miR-181b-5p,mmu-miR-27a-3p,mmu-miR-532-5p,mmu-miR-199a-3p,mmu-miR-199b-3p
turquoise	Rbmx	mmu-miR-29a-3p,mmu-miR-199a-5p
turquoise	Amer1	mmu-miR-29a-3p

Table A.2: Genes in the TMAO chromosome 12 QTL interval. List of genes in the entire chromosome 12 QTL interval. Transcript Cluster ID column provides the transcript cluster identification numbers from the Affymetrix MoGene 2.1 microarray. Genes that are not present in the microarray are marked by NA. Start and Stop columns denote transcription start and stop sites, respectively. QTL region columndescribes if the gene is present in either both the pre- and post-diet TMAO QTL regions, orpost-diet TMAO region only.

Genes	Transcript Cluster ID	Chromosome	Strand	Start	Stop	QTL Region
0610007P14Rik	17282732	chr12	-	85815448	85824550	Both
2410016O06Rik	17277130	chr12	+	83950608	83952953	Both
AB347864	NA	chr12	-	85441070	85441091	Both
AB349811	NA	chr12	+	86528204	86528225	Both
AB351889	NA	chr12	-	83961999	83962047	Both
Abcd4	17282534	chr12	-	84602531	84617466	Both
Acot1	17277140	chr12	+	84009498	84017671	Both
Acot2	17277134	chr12	+	83987861	83993877	Both
Acot3	17277152	chr12	+	84052144	84059565	Both
Acot4	17277146	chr12	+	84038379	84044723	Both
Acot5	17277161	chr12	+	84069325	84076020	Both
Acot6	17277170	chr12	+	84100654	84111349	Both
Acyp1	17282689	chr12	-	85272398	85288438	Both
AK033693	NA	chr12	+	84285373	84287899	Both
AK081163	NA	chr12	-	84313853	84317076	Both
AK086375	NA	chr12	+	85086373	85088228	Both
AK132490	NA	chr12	-	84039468	84053807	Both
AK140771	NA	chr12	-	83591335	83595687	Both
AK156692	NA	chr12	+	86916531	86918956	Both
AK164313	NA	chr12	-	83822558	83827661	Both
AK165270	NA	chr12	+	83648301	83652107	Both
AK186759	NA	chr12	+	85133452	85134083	Both
Aldh6a1	17282518	chr12	-	84430717	84450950	Both
Angel1	17282760	chr12	-	86700502	86726460	Both
Arel1	17282611	chr12	-	84918148	84970900	Both
Batf	17277404	chr12	+	85686669	85709087	Both
Bbof1	NA	chr12	+	84409067	84433780	Both
Cipc	17277552	chr12	+	86947043	86965364	Both
Coq6	17277232	chr12	+	84361657	84373796	Both
D030025P21Rik	17277300	chr12	+	84875769	84879755	Both
Dcaf4	17277043	chr12	+	83520466	83541994	Post-Diet TMAO

Dlst	17277352	chr12	+	85110833	85134091	Both
Dnal1	17277177	chr12	+	84114328	84143517	Both
Dpf3	17282403	chr12	-	83213745	83487708	Post-Diet TMAO
DQ687186	NA	chr12	-	84915344	84919247	Both
Eif2b2	17277370	chr12	+	85219481	85226628	Both
Elmsan1	17282465	chr12	-	84149168	84218881	Both
Entpd5	17282498	chr12	-	84373857	84409029	Both
Esrrb	17277504	chr12	+	86361117	86521628	Both
Fam161b	17282486	chr12	-	84345309	84361833	Both
Fcf1	17277307	chr12	+	84970897	84983303	Both
Fos	17277387	chr12	+	85473890	85477273	Both
Gm805	17277484	chr12	+	86169341	86195102	Both
Gpatch2l	17277490	chr12	+	86241869	86291414	Both
Ift43	17277472	chr12	+	86082561	86162459	Both
Irf2bpl	17282776	chr12	-	86880703	86884814	Both
Isca2	17277294	chr12	+	84773270	84775089	Both
JA187517	NA	chr12	-	85218436	85218457	Both
Jdp2	17277396	chr12	+	85599105	85639878	Both
Lin52	17277270	chr12	+	84451508	84531533	Both
Lrrc74a	NA	chr12	+	86734368	86763795	Both
Ltbp2	17282570	chr12	-	84783212	84876532	Both
Mfsd7c	17277411	chr12	+	85746539	85813585	Both
Mir6938	NA	chr12	-	85245922	85245985	Both
Mlh3	17282673	chr12	-	85234520	85270591	Both
Nek9	17282698	chr12	-	85299514	85339362	Both
Npc2	17282563	chr12	-	84754559	84773112	Both
Numb	17282435	chr12	-	83794034	83921934	Both
Papln	17277101	chr12	+	83763634	83792384	Both
Pgf	17282664	chr12	-	85166639	85177296	Both
Pnma1	17282461	chr12	-	84146131	84148489	Both
Prox2	17282641	chr12	-	85086814	85106431	Both
Psen1	17277084	chr12	+	83688202	83735199	Both
Ptgr2	17277195	chr12	+	84285232	84315832	Both
Rbm25	17277063	chr12	+	83632234	83683123	Both
Rnf113a2	17277265	chr12	+	84417200	84418578	Both
Rps6kl1	17282649	chr12	-	85135596	85151264	Both
Syndig1l	17282557	chr12	-	84677278	84698807	Both
Tgfb3	17282743	chr12	-	86056743	86079041	Both
Tmed10	17282723	chr12	-	85340614	85374717	Both
Tmem63c	17277561	chr12	+	87026292	87090041	Both
Ttll5	17277424	chr12	+	85824673	86053760	Both

Vash1	17277521	chr12	+	86678700	86695681	Both
Vrtn	17277287	chr12	+	84641019	84651455	Both
Vsx2	17277278	chr12	+	84569840	84595457	Both
Ylpm1	17277319	chr12	+	84996321	85070515	Both
Zc2hc1c	17277382	chr12	+	85288591	85299358	Both
Zdhhc22	17282783	chr12	-	86980764	86988676	Both
Zfp410	17277212	chr12	+	84316859	84344439	Both
Zfyve1	17282420	chr12	-	83546941	83597147	Both

Table A.3: Fold change of robustly expressed miRNAs in African Green Monkey livers. P-values are not adjusted.

miRNAs	FOLD_CHANGE (HFHC/Chow)	P_VALUE
hsa-mir-590-3p	14.05956079	0.00018404
hsa-mir-590-3p_-_1	14.00448562	0.000403138
hsa-mir-29b-2-3p_-_1	3.047384615	0.148574234
hsa-mir-378i	3.039855048	0.15201004
hsa-mir-29a-5p	2.982051282	0.161315438
hsa-mir-29b-1-3p_-_1	2.887463557	0.168071801
hsa-mir-27b-5p	2.383482195	0.228574395
hsa-mir-29c-5p_+_1	2.236625971	0.093716366
hsa-mir-133a-1	2.211568123	0.290275654
hsa-mir-133a-2	2.211568123	0.290275654
hsa-mir-133a-1_+_1	2.174717833	0.359116771
hsa-mir-133a-2_+_1	2.174717833	0.359116771
hsa-mir-146a-5p	2.028501687	0.028141747
hsa-mir-21-5p	2.018896639	0.002463205
hsa-mir-126-5p	1.945253287	0.007223175
hsa-mir-22-3p_+_1	1.920912006	0.033197831
hsa-mir-126-3p_+_1	1.917841357	0.005233945
hsa-mir-126-5p_+_1	1.903835666	0.005819777
hsa-mir-199b-5p	1.867673365	0.103236177
hsa-mir-22-3p	1.86760805	0.019875459
hsa-mir-140-5p	1.859099479	0.043917403
hsa-mir-660-5p	1.847675999	0.016094865
hsa-mir-34a-5p_+_1	1.827690551	0.222898902
hsa-mir-223-3p	1.825247063	0.05244403
hsa-mir-1-2	1.814156873	0.267009634
hsa-mir-19a-3p	1.794781826	0.353786793
hsa-mir-126-3p_-_1	1.786809269	0.065028898
hsa-mir-126-3p_+_2	1.78615467	0.028213046
hsa-mir-429	1.777394389	0.089402862
hsa-mir-130b-3p	1.772544302	0.031027022
hsa-mir-126-3p	1.753587329	0.032240955
hsa-mir-1-1	1.752517162	0.283147029
hsa-mir-454-3p	1.748555466	0.029594536

hsa-mir-34a-5p_-_1	1.732943	0.291097907
hsa-mir-34a-5p	1.730994117	0.239921739
hsa-mir-142-5p	1.722240154	0.145609165
hsa-mir-142-5p_-_2	1.703974108	0.082119751
hsa-mir-29b-1-5p	1.677121845	0.112518902
hsa-mir-29b-1-5p	1.677121845	0.112518902
hsa-mir-32-3p	1.625300048	0.040184318
hsa-mir-100-5p	1.605511727	0.088401584
hsa-mir-186-5p	1.596572138	0.059871264
hsa-mir-203	1.594255686	0.047602284
hsa-mir-676-3p	1.590418132	0.096525075
hsa-mir-152	1.569334238	0.051930708
hsa-mir-145-3p_+_2	1.569119964	0.14999042
hsa-mir-222-3p	1.565449135	0.073267246
hsa-mir-186-5p_+_1	1.55775079	0.101428839
hsa-mir-425-5p_+_1	1.553701586	0.086151273
hsa-mir-421	1.553563142	0.068678853
hsa-mir-29b-2-5p_+_1	1.53604336	0.488797529
hsa-mir-335-5p	1.529178748	0.103644734
hsa-mir-27a-3p	1.526562368	0.029291464
hsa-mir-146b-5p	1.523028906	0.074824737
hsa-mir-30e-5p_+_1	1.505520458	0.11981439
hsa-mir-99a-5p	1.499427785	0.148326308
hsa-mir-24-1-3p	1.498813216	0.071773971
hsa-mir-411-3p	1.49627907	0.081606134
hsa-mir-24-2-3p	1.494832629	0.074036454
hsa-mir-532-5p_+_1	1.494581046	0.011727452
hsa-mir-200b-3p	1.489435031	0.109126346
hsa-mir-29c-3p	1.480466892	0.228831749
hsa-mir-29c-3p	1.480466892	0.228831749
hsa-mir-215_+_1	1.458478265	0.239470777
hsa-mir-200a-3p	1.455665967	0.156206277
hsa-mir-194-2-5p	1.453517195	0.201124112
hsa-mir-30e-5p	1.446254958	0.165771862
hsa-mir-99b-5p_+_1	1.446012724	0.359529949
hsa-mir-29a-3p	1.439797501	0.354418408
hsa-mir-29a-3p	1.439797501	0.354418408
hsa-mir-199a-1-5p	1.437815048	0.140147689
hsa-mir-199a-2-5p	1.437815048	0.140147689

hsa-mir-148b-3p	1.428242508	0.083972601
hsa-mir-221-3p	1.427797305	0.104694297
hsa-mir-369-3p	1.421589428	0.139937905
hsa-mir-15a-5p	1.420723652	0.167513236
hsa-mir-29b-2-5p_-_1	1.415086207	0.328758904
hsa-mir-27b-3p_+_1	1.414125528	0.234978036
hsa-mir-497-5p	1.413543506	0.228968752
hsa-mir-532-5p	1.407959076	0.002587625
hsa-mir-29b-2-5p	1.406064395	0.149816687
hsa-mir-29b-2-5p	1.406064395	0.149816687
hsa-mir-29a-3p_-_1	1.405670163	0.330731685
hsa-mir-29a-3p_-_1	1.405670163	0.330731685
hsa-mir-29c-5p	1.403094233	0.397490167
hsa-mir-455-5p	1.40000938	0.283511171
hsa-mir-339-3p	1.394973939	0.245223914
hsa-mir-99b-5p	1.38700713	0.383651839
hsa-mir-22-5p	1.382975436	0.061496549
hsa-mir-27b-3p_-_1	1.372458796	0.197590143
hsa-mir-128-1	1.371886339	0.249206022
hsa-mir-136-3p	1.371369933	0.127274334
hsa-mir-24-1-3p_+_2	1.370989252	0.279034865
hsa-mir-194-1-5p	1.370974037	0.194343912
hsa-mir-500a-3p	1.370888927	0.14743196
hsa-mir-19b-2-3p	1.366467118	0.551396673
hsa-mir-19b-1-3p	1.365236089	0.554644042
hsa-mir-16-2-5p	1.365130459	0.182839078
hsa-mir-122-3p	1.362714181	0.280773376
hsa-mir-128-2	1.361253431	0.276414513
hsa-mir-16-1-5p	1.361225175	0.183180056
hsa-mir-411-5p_-_1	1.358692387	0.13863353
hsa-mir-17-5p	1.357965819	0.235269448
hsa-mir-192-5p_+_1	1.353857012	0.199620487
hsa-let-7f-2-5p_+_1	1.352963671	0.165563702
hsa-mir-122-5p_-_2	1.351069133	0.14584982
hsa-mir-361-3p	1.341444435	0.28098751
hsa-mir-493-3p	1.339424201	0.355286217
hsa-mir-122-3p_-_1	1.338470623	0.421743249
hsa-mir-20a-5p	1.338102836	0.283148447
hsa-mir-143-5p	1.336878246	0.35202295

hsa-mir-106b-5p	1.331447112	0.168165316
hsa-mir-374a-5p	1.328118609	0.305859463
hsa-mir-7-1-5p	1.325289486	0.161407284
hsa-mir-485-5p	1.321185787	0.395287928
hsa-mir-342-5p	1.320744157	0.285105739
hsa-mir-7-3-5p	1.315704374	0.15821248
hsa-mir-7-2-5p	1.314315944	0.160825392
hsa-mir-664-5p_+_1	1.31035837	0.285180669
hsa-mir-411-5p	1.309104171	0.183311979
hsa-mir-361-3p_+_1	1.304090181	0.312126828
hsa-mir-192-5p_-_1	1.302133613	0.172927337
hsa-mir-1307-5p	1.299591331	0.462569893
hsa-mir-23a-3p	1.298442083	0.065885981
hsa-mir-200b-3p_+_1	1.29694803	0.389837497
hsa-mir-125a-5p	1.296918815	0.413769017
hsa-mir-424-3p	1.291901078	0.376940138
hsa-mir-132-3p	1.290301115	0.186145386
hsa-mir-203_+_1	1.286150171	0.152960898
hsa-let-7g-5p_-_1	1.283408942	0.238883648
hsa-mir-532-3p	1.283352229	0.426507071
hsa-mir-154-5p	1.282669856	0.185369823
hsa-mir-503	1.280229868	0.22630693
hsa-let-7g-5p	1.278204277	0.27429742
hsa-mir-191-5p	1.27706538	0.301748666
hsa-mir-342-5p_+_1	1.272741771	0.426201777
hsa-mir-199a-2-3p	1.270671092	0.240732865
hsa-mir-199a-1-3p	1.270423088	0.242042419
hsa-mir-199b-3p	1.270423088	0.242042419
hsa-mir-654-3p_-_2	1.268376977	0.105430057
hsa-mir-331-3p	1.26779616	0.353481405
hsa-mir-125b-2-5p	1.267676784	0.414980862
hsa-mir-125b-1-5p	1.263393727	0.419637568
hsa-mir-654-3p	1.261369557	0.17053399
hsa-mir-10a-3p	1.260497024	0.14662385
hsa-mir-27b-3p	1.260458869	0.253593872
hsa-mir-185-5p	1.260350505	0.377131321
hsa-mir-214-5p	1.255268836	0.490070753
hsa-mir-199a-2-3p_-_1	1.25525307	0.197312432
hsa-mir-199a-1-3p_-_1	1.254451483	0.198489997

hsa-mir-199b-3p_-_1	1.254451483	0.198489997
hsa-mir-425-5p	1.249701271	0.352421628
hsa-mir-505-5p	1.246734317	0.383465765
hsa-mir-191-5p_+_1	1.245541626	0.376394882
hsa-mir-125a-3p	1.244206062	0.523366474
hsa-mir-499a-5p	1.238343641	0.226861693
hsa-mir-361-5p	1.235445012	0.161895246
hsa-mir-130a-3p	1.229013254	0.194360805
hsa-mir-664-5p	1.228365058	0.436713843
hsa-mir-484	1.225257485	0.364595802
hsa-mir-155-5p	1.214515964	0.446942833
hsa-mir-26b-5p	1.213952027	0.178978622
hsa-mir-431-5p	1.203334254	0.540466492
hsa-let-7f-1-5p	1.200225671	0.140891495
hsa-mir-324-5p	1.198663107	0.491083139
hsa-mir-425-3p_-_1	1.197252234	0.392042495
hsa-let-7f-2-5p	1.197191193	0.147982457
hsa-mir-100-5p_+_3	1.196230131	0.614018964
hsa-mir-150-5p	1.19285993	0.325572809
hsa-mir-99a-5p_+_3	1.191857853	0.584605024
hsa-mir-92b-3p	1.191690203	0.490762805
hsa-mir-455-5p_+_2	1.177973623	0.363148736
hsa-mir-151a-3p_-_2	1.174630357	0.042007616
hsa-mir-29c-3p_+_5	1.174271088	0.652210495
hsa-mir-122-5p_+_1	1.169930623	0.416686995
hsa-mir-487b_+_2	1.167554102	0.53129519
hsa-mir-10a-5p_+_1	1.163655254	0.198689264
hsa-mir-30d-5p	1.160590136	0.457072646
hsa-let-7e-5p	1.154950284	0.582293349
hsa-mir-378c	1.149732779	0.293022133
hsa-mir-326	1.146568993	0.667362855
hsa-mir-193a-5p	1.141937858	0.495525682
hsa-mir-424-5p	1.141734906	0.657982709
hsa-mir-122-5p_+_2	1.140517908	0.31048089
hsa-mir-125b-1-3p	1.134912062	0.678954348
hsa-mir-26a-1-5p	1.134342208	0.140048424
hsa-mir-26a-2-5p	1.133543295	0.14116087
hsa-mir-378d-2	1.129414689	0.361072964
hsa-mir-192-5p	1.124157944	0.443833084

hsa-let-7d-5p	1.123755529	0.59980024
hsa-mir-652-3p	1.12360228	0.649846726
hsa-mir-199a-2-3p_+_1	1.120670046	0.498777898
hsa-mir-199a-1-3p_+_1	1.117205419	0.513561083
hsa-mir-199b-3p_+_1	1.117205419	0.513561083
hsa-mir-376a-1-3p	1.112828519	0.626307358
hsa-mir-376a-2-3p	1.112828519	0.626307358
hsa-mir-455-3p_-_2	1.112611887	0.590321372
hsa-mir-335-3p	1.108862809	0.404940746
hsa-mir-151b	1.107793773	0.664852425
hsa-mir-151a-5p	1.10530889	0.612364905
hsa-mir-323a-3p_-_1	1.095136418	0.779594927
hsa-mir-885-3p	1.094502247	0.735419496
hsa-mir-299-5p	1.085872284	0.693612994
hsa-let-7a-2-5p	1.085570593	0.538069832
hsa-mir-28-5p	1.085041026	0.7497052
hsa-let-7a-3-5p	1.08493054	0.541439262
hsa-mir-433	1.083420182	0.783684662
hsa-let-7a-1-5p	1.083114906	0.547740981
hsa-mir-410	1.081382597	0.598598412
hsa-mir-148a-3p_+_2	1.073945946	0.617706996
hsa-mir-378d-1	1.072552843	0.528979853
hsa-mir-339-5p	1.068362258	0.830272185
hsa-mir-95	1.066679466	0.776632459
hsa-mir-122-5p_+_3	1.066594926	0.729001104
hsa-mir-323b-3p	1.066175355	0.822866261
hsa-mir-193b-3p	1.066035783	0.82077522
hsa-mir-100-5p_+_4	1.06269237	0.851214063
hsa-mir-17-3p	1.061207012	0.844073387
hsa-mir-181b-2-5p_+_1	1.060406252	0.611985736
hsa-mir-455-3p	1.060314869	0.774614485
hsa-mir-10a-5p	1.058646217	0.671411797
hsa-mir-423-3p_-_1	1.056818898	0.820219301
hsa-mir-93-5p_+_1	1.056382493	0.634169749
hsa-mir-887	1.054452095	0.874416114
hsa-mir-98	1.047036272	0.639954629
hsa-let-7a-3-5p_-_1	1.045124535	0.8032748
hsa-mir-1180	1.043778849	0.865000183
hsa-mir-29a-3p_+_5	1.040219629	0.941835609

hsa-mir-29a-3p_+_5	1.040219629	0.941835609
hsa-mir-423-3p	1.036233536	0.890977889
hsa-mir-141-3p	1.035932024	0.960069392
hsa-mir-122-5p_+_4	1.030982175	0.911763112
hsa-mir-122-5p_-_1	1.029811166	0.802624907
hsa-mir-107	1.028328652	0.894122184
hsa-mir-99a-5p_+_4	1.025223389	0.944334005
hsa-mir-543	1.022036595	0.937105152
hsa-mir-487b	1.021102213	0.927238807
hsa-mir-140-3p	1.013308319	0.892731119
hsa-let-7g-5p_+_3	1.008268924	0.981466096
hsa-mir-134	1.007760523	0.953613374
hsa-mir-423-3p_+_1	1.007747955	0.981063467
hsa-mir-409-3p_+_1	1.005095179	0.981935934
hsa-mir-122-5p_+_5	1.00439322	0.987092246
hsa-mir-30c-1-3p	1.003907482	0.982507211
hsa-let-7c	1.00074132	0.996987368
hsa-mir-574-5p	0.999425854	0.998130433
hsa-mir-127-3p	0.996120346	0.988728895
hsa-mir-423-5p	0.995775195	0.986442174
hsa-mir-148a-3p	0.990555329	0.957953883
hsa-mir-505-3p_+_1	0.989745037	0.948067609
hsa-mir-29b-2-3p	0.988767032	0.977389577
hsa-mir-744-5p	0.985594527	0.959427523
hsa-mir-504_+_1	0.982216723	0.894306605
hsa-mir-598	0.980995862	0.945064835
hsa-mir-139-3p_-_1	0.979495915	0.929760389
hsa-mir-100-5p_+_5	0.976515473	0.947496823
hsa-mir-103a-2-3p	0.975575055	0.781333589
hsa-mir-193b-5p_-_1	0.975526604	0.942933489
hsa-mir-103a-1-3p	0.972349897	0.753963775
hsa-mir-483-5p	0.971429796	0.908836343
hsa-let-7i-5p	0.970984245	0.877629574
hsa-mir-190a	0.9700519	0.934858946
hsa-mir-200c-3p	0.96882877	0.956151664
hsa-mir-505-3p	0.962742371	0.794485712
hsa-mir-574-3p_+_1	0.96126119	0.855149997
hsa-mir-671-3p	0.959092441	0.867660375
hsa-mir-99a-3p	0.954812477	0.778521687

hsa-mir-455-3p_+_1	0.953385354	0.845041661
hsa-mir-29b-1-3p	0.952258983	0.900236961
hsa-let-7c_-_1	0.947620962	0.579046898
hsa-mir-320a_-_1	0.946295323	0.810245936
hsa-mir-654-5p	0.944619752	0.650955655
hsa-mir-340-5p	0.942878358	0.813662914
hsa-mir-485-3p	0.94171875	0.801960609
hsa-mir-23b-3p	0.94059224	0.724541162
hsa-mir-125b-2-3p_+_2	0.940583806	0.803393493
hsa-mir-214-3p	0.936117647	0.779765172
hsa-mir-370	0.932893431	0.842800093
hsa-mir-151a-3p	0.929510749	0.508529077
hsa-mir-99a-5p_+_5	0.927664124	0.851301619
hsa-mir-193b-3p_+_1	0.926981885	0.799207317
hsa-mir-1307-3p	0.92686145	0.655246388
hsa-let-7d-3p	0.918266521	0.71748926
hsa-mir-29c-3p_+_6	0.916788321	0.811378872
hsa-mir-29a-3p_+_6	0.916367651	0.865597204
hsa-mir-29a-3p_+_6	0.916367651	0.865597204
hsa-mir-92a-1-5p	0.915845274	0.852510223
hsa-mir-483-3p_+_1	0.912242599	0.625887882
hsa-mir-181b-2-5p	0.909963036	0.453208313
hsa-mir-93-5p	0.90824732	0.401810878
hsa-mir-1296	0.907727933	0.750891383
hsa-mir-129-2-3p	0.90744868	0.834412898
hsa-mir-204-5p	0.904799788	0.74847063
hsa-mir-342-3p	0.904262814	0.776181188
hsa-mir-574-3p	0.903774117	0.526457448
hsa-mir-29a-3p_+_7	0.902262057	0.858460716
hsa-mir-493-5p	0.901851439	0.553398558
hsa-let-7b-5p	0.901117811	0.647736532
hsa-mir-378a-5p	0.893246617	0.671295934
hsa-mir-505-3p_+_2	0.892887243	0.604451674
hsa-mir-30b-3p	0.8912711	0.566050946
hsa-mir-195-5p	0.888714238	0.735331684
hsa-mir-320a	0.887224922	0.645009652
hsa-mir-320a_+_1	0.887089697	0.653949948
hsa-mir-122-5p_+_6	0.884534655	0.693355773
hsa-mir-148a-5p	0.881991637	0.414483775

hsa-mir-320a_+_2	0.877640075	0.668732529
hsa-mir-382-3p	0.877099735	0.591172277
hsa-mir-328	0.875053078	0.688777162
hsa-mir-181b-1-5p	0.874338648	0.386893009
hsa-mir-483-3p	0.87171283	0.506307107
hsa-mir-378a-5p_+_1	0.869024292	0.588786926
hsa-mir-504	0.868767243	0.579359591
hsa-mir-193b-5p	0.866407263	0.640424194
hsa-mir-320b-2	0.863778428	0.561291475
hsa-mir-30d-3p	0.862030814	0.405816633
hsa-mir-181c-5p	0.86142005	0.182374539
hsa-mir-378a-3p	0.86005768	0.195632385
hsa-mir-31-5p	0.859622136	0.757829846
hsa-mir-214-3p_-_1	0.856013654	0.588723969
hsa-mir-21-3p	0.851562003	0.703861763
hsa-mir-30c-2-5p	0.843168623	0.328831441
hsa-mir-30c-1-5p	0.842910171	0.328562476
hsa-mir-197-3p	0.838247127	0.352937777
hsa-mir-127-3p_-_1	0.837539936	0.457668588
hsa-mir-140-3p_+_1	0.836774351	0.356468065
hsa-mir-139-5p	0.806055311	0.404609399
hsa-mir-409-3p	0.788058528	0.422208928
hsa-mir-145-5p	0.786718656	0.619144782
hsa-mir-374b-5p	0.774180797	0.585883539
hsa-mir-30b-5p	0.76767335	0.569322149
hsa-mir-320b-1	0.764709865	0.35888887
hsa-mir-195-3p	0.761141232	0.327542556
hsa-mir-99b-3p	0.759632344	0.42865002
hsa-mir-101-1-3p_-_1	0.758326602	0.554553994
hsa-mir-30a-5p	0.757303928	0.472061092
hsa-mir-101-2-3p_-_1	0.756219526	0.550472083
hsa-mir-664-3p	0.754801331	0.603114246
hsa-mir-23b-3p_+_1	0.751961569	0.312468156
hsa-mir-30a-5p_+_1	0.7423799	0.466584168
hsa-mir-29c-3p_+_7	0.731893266	0.510049761
hsa-mir-379-5p	0.702497858	0.31086116
hsa-mir-181a-1-5p	0.678383957	0.137425879
hsa-mir-181a-2-5p	0.678353802	0.137380816
hsa-mir-181a-1-3p	0.675378705	0.462218

hsa-mir-374a-3p	0.671097834	0.49904261
hsa-mir-432-5p	0.664511144	0.117905032
hsa-mir-409-3p_-_1	0.663752232	0.246267826
hsa-mir-1307-3p_+_1	0.658686453	0.129592738
hsa-mir-16-2-3p_-_1	0.653202039	0.297950672
hsa-mir-92a-1-3p	0.648750941	0.109917477
hsa-mir-769-5p	0.644634654	0.348807878
hsa-mir-92a-2-3p	0.643084573	0.123652716
hsa-mir-381	0.642321975	0.167999992
hsa-mir-15b-5p	0.629818557	0.190671022
hsa-mir-210	0.613601522	0.223266828
hsa-mir-382-5p_+_2	0.61331998	0.201117865
hsa-mir-181d	0.607818629	0.237165418
hsa-mir-148a-3p_+_1	0.577839134	0.358243733
hsa-mir-215	0.576546311	0.48334011
hsa-mir-143-3p	0.574768699	0.384819596
hsa-mir-106b-3p_-_2	0.560419564	0.073745256
hsa-mir-101-2-3p	0.560111016	0.468537104
hsa-mir-495	0.555594656	0.400841846
hsa-mir-140-3p_+_2	0.551444127	0.335779872
hsa-mir-200a-5p	0.550044635	0.290500228
hsa-mir-143-3p_+_3	0.54969594	0.149385376
hsa-mir-148a-5p_-_1	0.544290943	0.294244184
hsa-mir-101-1-3p	0.538456137	0.455081031
hsa-mir-194-2-3p	0.537801322	0.151264589
hsa-mir-92a-1-3p_+_1	0.537791369	0.096606346
hsa-mir-92a-2-3p_+_1	0.530293943	0.112532734
hsa-mir-378a-3p_+_1	0.507108491	0.284397705
hsa-mir-143-3p_+_1	0.497748344	0.374983327
hsa-mir-30c-2-3p	0.487134582	0.235315659
hsa-mir-375	0.48437756	0.268688375
hsa-mir-143-3p_-_1	0.472781717	0.339839842
hsa-mir-10b-5p_+_1	0.456199913	0.467460454
hsa-mir-106b-3p	0.44762806	0.095797425
hsa-mir-382-5p	0.440970859	0.251771284
hsa-mir-28-3p	0.408499963	0.276424152
hsa-mir-187-3p	0.391519728	0.30028873
hsa-mir-28-3p_+_1	0.382293179	0.257891899
hsa-mir-25-3p	0.363779923	0.285592833

hsa-mir-10b-5p	0.356073753	0.402528657
hsa-mir-25-3p_+_1	0.335104346	0.264630944
hsa-mir-889	0.294200503	0.311144676
hsa-mir-382-5p_+_1	0.293081183	0.285210492
hsa-mir-28-3p_+_2	0.291146639	0.276761116
hsa-mir-30a-3p_+_1	0.269571391	0.277655038
hsa-mir-144-5p	0.247697718	0.020885832
hsa-mir-584-5p	0.246246476	0.036640254
hsa-mir-129-1-5p	0.241643457	0.269367837
hsa-mir-129-2-5p	0.241590011	0.26951116
hsa-mir-375_-_1	0.240072439	0.296805999
hsa-mir-30a-3p	0.23998844	0.291156926
hsa-mir-30e-3p	0.230516514	0.288679839
hsa-mir-144-3p	0.230390144	0.069373664
hsa-mir-30e-3p_+_1	0.228846866	0.284684681
hsa-mir-182-5p	0.224207096	0.276432838
hsa-mir-486-3p	0.209603497	0.066229684
hsa-mir-451a	0.169453882	0.00482445
hsa-mir-490-5p	0.155332315	0.299675602
hsa-mir-486-5p	0.137932284	0.000912209
hsa-mir-486-5p_+_2	0.135503255	0.001495682
hsa-mir-3607-3p_-_5	0.082709235	0.281022079
hsa-mir-891a	0	0.292351992

REFERENCES

Abudoukelimu M, Fu ZY, Maimaiti A, Ma YT, Abudu M, Zhu Q, Adi D, Yang YN, Li XM, Xie X, Liu F, Chen BD (2015) The association of cholesterol absorption gene Numb polymorphism with Coronary Artery Disease among Han Chinese and Uighur Chinese in Xinjiang, China. *Lipids in health and disease*. 14, 120.

Adiels M, Olofsson SO, Taskinen MR, Boren J. Overproduction of very low-density lipoproteins is the hallmark of the dyslipidemia in the metabolic syndrome. *Arteriosclerosis, thrombosis, and vascular biology*. 2008;28(7):1225-36.

Agarwal V, Bell GW, Nam JW, Bartel DP (2015) Predicting effective microRNA target sites in mammalian mRNAs. *eLife*. 4.

Akhtar N, Rasheed Z, Ramamurthy S, Anbazhagan AN, Voss FR, Haqqi TM. MicroRNA-27b regulates the expression of matrix metalloproteinase 13 in human osteoarthritis chondrocytes. *Arthritis and rheumatism*. 2010;62(5):1361-71.

Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, et al. Clustering and conservation patterns of human microRNAs. *Nucleic acids research*. 2005;33(8):2697-706.

Alvarez ML, Khosroheidari M, Eddy E, Done SC. MicroRNA-27a decreases the level and efficiency of the LDL receptor and contributes to the dysregulation of cholesterol homeostasis. *Atherosclerosis*. 2015;242(2):595-604.

Ambros V, Horvitz HR. The lin-14 locus of Caenorhabditis elegans controls the time of expression of specific postembryonic developmental events. *Genes & development*. 1987;1(4):398-414.

Austin MA, King MC, Vranizan KM, Krauss RM. Atherogenic lipoprotein phenotype. A proposed genetic marker for coronary heart disease risk. *Circulation*. 1990;82(2):495-506.

Baldan A, Fernandez-Hernando C. Truths and controversies concerning the role of miRNAs in atherosclerosis and lipid metabolism. *Current opinion in lipidology*. 2016;27(6):623-9.

Baran-Gale J, Kurtz CL, Erdos MR, Sison C, Young A, Fannin EE, et al. Addressing Bias in Small RNA Library Preparation for Sequencing: A New Protocol Recovers MicroRNAs that Evade Capture by Current Methods. *Frontiers in genetics*. 2015;6:352.

Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell*. 2009;136(2):215-33.

Bennett BJ, Davis RC, Civelek M, Orozco L, Wu J, Qi H, Pan C, Packard RR, Eskin E, Yan M, Kirchgessner T, Wang Z, Li X, Gregory JC, Hazen SL, Gargalovic PS, Lusis AJ (2015) Genetic Architecture of Atherosclerosis in Mice: A Systems Genetics Analysis of Common Inbred Strains. *PLoS Genet*. 11, e1005711.

Bennett BJ, de Aduiar Vallim TQ, Wang Z, Shih DM, Meng Y, Gregory J, Allayee H, Lee R, Graham M, Crooke R, Edwards PA, Hazen SL (2013a) Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation. *Cell Metabolism*. 17, 49-60.

Bennett Brian J, Vallim Thomas Qde A, Wang Z, Shih Diana M, Meng Y, Gregory J, Allayee H, Lee R, Graham M, Crooke R, Edwards Peter A, Hazen Stanley L, Lusis Aldons J (2013b) Trimethylamine-N-Oxide, a Metabolite Associated with Atherosclerosis, Exhibits Complex Genetic and Dietary Regulation. *Cell Metabolism*. 17, 49-60.

Borchert GM, Lanier W, Davidson BL. RNA polymerase III transcribes human microRNAs. *Nature structural & molecular biology*. 2006;13(12):1097-101.

Calo N, Ramadori P, Sobolewski C, Romero Y, Maeder C, Fournier M, et al. Stress-activated miR-21/miR-21* in hepatocytes promotes lipid and glucose metabolic disorders associated with high-fat diet consumption. *Gut*. 2016;65(11):1871-81.

Cermelli S, Guo Y, Gross SP, Welte MA. The lipid-droplet proteome reveals that droplets are a protein-storage depot. *Current biology: CB*. 2006;16(18):1783-95.

Chan CS, Elemento O, Tavazoie S. Revealing posttranscriptional regulatory elements through network-level conservation. *PLoS computational biology*. 2005;1(7):e69.

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC bioinformatics*. 2013;14:128.

Chen WJ, Yin K, Zhao GJ, Fu YC, Tang CK. The magic and mystery of microRNA-27 in atherosclerosis. *Atherosclerosis*. 2012;222(2):314-23.

Cheng HS, Besla R, Li A, Chen Z, Shikatani EA, Nazari-Jahantigh M, Hammoutene A, Nguyen MA, Geoffrion M, Cai L, Khyzha N, Li T, MacParland SA, Husain M, Cybulsky MI, Boulanger CM, Temel RE, Schober A, Rayner KJ, Robbins CS,

Fish JE (2017) Paradoxical Suppression of Atherosclerosis in the Absence of microRNA-146a. *Circulation research*. 121, 354-367.

Christopher AF, Kaur RP, Kaur G, Kaur A, Gupta V, Bansal P. MicroRNA therapeutics: Discovering novel targets and developing specific therapy. *Perspectives in clinical research*. 2016;7(2):68-74.

Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD, Beatty J, et al. The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nature genetics*. 2004;36(11):1133-7.

Churchill GA, Gatti DM, Munger SC, Svenson KL. The Diversity Outbred mouse population. *Mammalian genome : official journal of the International Mammalian Genome Society*. 2012;23(9-10):713-8.

Civelek M, Lusis AJ. Systems genetics approaches to understand complex traits. *Nature reviews Genetics*. 2014;15(1):34-48.

Cohen SM, Brennecke J, Stark A. Denoising feedback loops by thresholding--a new role for microRNAs. *Genes & development*. 2006;20(20):2769-72.

Coronnello C, Benos PV. ComiR: Combinatorial microRNA target prediction tool. *Nucleic acids research*. 2013;41(Web Server issue):W159-64.

Defesche JC. Low-density lipoprotein receptor--its structure, function, and mutations. *Seminars in vascular medicine*. 2004;4(1):5-11.

Deguchi JO, Aikawa E, Libby P, Vachon JR, Inada M, Krane SM, et al. Matrix metalloproteinase-13/collagenase-3 deletion promotes collagen accumulation and organization in mouse atherosclerotic plaques. *Circulation*. 2005;112(17):2708-15.

Ding J, Li M, Wan X, Jin X, Chen S, Yu C, et al. Effect of miR-34a in regulating steatosis by targeting PPARalpha expression in nonalcoholic fatty liver disease. *Scientific reports*. 2015;5:13729.

Doench JG, Sharp PA. Specificity of microRNA target selection in translational repression. *Genes & development*. 2004;18(5):504-11.

Ekstedt M, Hagstrom H, Nasr P, Fredrikson M, Stal P, Kechagias S, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology* (Baltimore, Md). 2015;61(5):1547-54.

Elmen J, Lindow M, Schutz S, Lawrence M, Petri A, Obad S, et al. LNA-mediated microRNA silencing in non-human primates. *Nature*. 2008;452(7189):896-9.

Emini Veseli B, Perrotta P, De Meyer GRA, Roth L, Van der Donckt C, Martinet W, et al. Animal models of atherosclerosis. *European journal of pharmacology*. 2017;816:3-13.

Eulalio A, Huntzinger E, Izaurralde E. Getting to the root of miRNA-mediated gene silencing. *Cell*. 2008;132(1):9-14.

Fan YM, Hernesniemi J, Oksala N, Levula M, Raitoharju E, Collings A, et al. Upstream Transcription Factor 1 (USF1) allelic variants regulate lipoprotein metabolism in women and USF1 expression in atherosclerotic plaque. *Scientific reports*. 2014;4:4650.

Feinberg, MW, Moore, KJ (2016) MicroRNA Regulation of Atherosclerosis. *Circ Res*. 118(4): 703-720.

Fickert P, Zollner G, Fuchsbichler A, Stumptner C, Pojer C, Zenz R, et al. Effects of ursodeoxycholic and cholic acid feeding on hepatocellular transporter expression in mouse liver. *Gastroenterology*. 2001;121(1):170-83.

Forloni M, Dogra SK, Dong Y, Conte D, Jr., Ou J, Zhu LJ, Deng A, Mahalingam M, Green MR, Wajapeyee N (2014) miR-146a promotes the initiation and progression of melanoma by activating Notch signaling. *eLife*. 3, e01460.

Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome research*. 2009;19(1):92-105.

Fu T, Choi SE, Kim DH, Seok S, Suino-Powell KM, Xu HE, et al. Aberrantly elevated microRNA-34a in obesity attenuates hepatic responses to FGF19 by targeting a membrane coreceptor beta-Klotho. *Proceedings of the National Academy of Sciences of the United States of America*. 2012;109(40):16137-42.

Furtado KS, de Oliveira Andrade F, Campos A, Rosim MP, Vargas-Mendez E, Henriques A, et al. beta-ionone modulates the expression of miRNAs and genes involved in the metastatic phenotype of microdissected persistent preneoplastic lesions in rats submitted to hepatocarcinogenesis. *Molecular carcinogenesis*. 2017;56(1):184-96.

Gatti DM, Svenson KL, Shabalin A, Wu LY, Valdar W, Simecek P, et al. Quantitative trait locus mapping methods for diversity outbred mice. *G3* (Bethesda, Md). 2014;4(9):1623-33.

Georges SA, Biery MC, Kim SY, Schelter JM, Guo J, Chang AN, et al. Coordinated regulation of cell cycle transcripts by p53-Inducible microRNAs, miR-192 and miR-215. *Cancer research*. 2008;68(24):10105-12.

Ginsberg HN. Lipoprotein metabolism and its relationship to atherosclerosis. *The Medical clinics of North America*. 1994;78(1):1-20.

Giorgetti L, Siggers T, Tiana G, Caprara G, Notarbartolo S, Corona T, et al. Noncooperative interactions between transcription factors and clustered DNA binding sites enable graded transcriptional responses to environmental inputs. *Molecular cell*. 2010;37(3):418-28.

Goedeke L, Rotllan N, Canfran-Duque A, Aranda JF, Ramirez CM, Araldi E, et al. MicroRNA-148a regulates LDL receptor and ABCA1 expression to control circulating lipoprotein levels. *Nature medicine*. 2015;21(11):1280-9.

Goedeke L, Salerno A, Ramirez CM, Guo L, Allen RM, Yin X, et al. Long-term therapeutic silencing of miR-33 increases circulating triglyceride levels and hepatic lipid accumulation in mice. *EMBO molecular medicine*. 2014;6(9):1133-41.

Gregory JC, Buffa JA, Org E, Wang Z, Levison BS, Zhu W, Wagner MA, Bennett BJ, Li L, DiDonato JA, Lusis AJ, Hazen SL (2014) Transmission of Atherosclerosis Susceptibility with Gut Microbial Transplantation. *The Journal of biological chemistry*.

Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Molecular cell*. 27, 91-105.

Guo Y, Hwang LD, Li J, Eades J, Yu CW, Mansfield C, et al. Genetic analysis of impaired trimethylamine metabolism using whole exome sequencing. BMC medical genetics. 2017;18(1):11.

Gusev Y, Schmittgen TD, Lerner M, Postier R, Brackett D. Computational analysis of biological functions and pathways collectively targeted by co-expressed microRNAs in cancer. *BMC bioinformatics*. 2007;8 Suppl 7:S16.

Han H, Qu G, Han C, Wang Y, Sun T, Li F, et al. MiR-34a, miR-21 and miR-23a as potential biomarkers for coronary artery disease: a pilot microarray study and confirmation in a 32 patient cohort. *Experimental & molecular medicine*. 2015;47:e138.

Hartiala J, Bennett BJ, Tang WH, Wang Z, Stewart AF, Roberts R, et al. Comparative genome-wide association studies in mice and humans for trimethylamine N-oxide, a proatherogenic metabolite of choline and L-carnitine. *Arteriosclerosis, thrombosis, and vascular biology*. 2014;34(6):1307-13.

Hashimoto Y, Akiyama Y, Yuasa Y. Multiple-to-multiple relationships between microRNAs and target genes in gastric cancer. *PloS one*. 2013;8(5):e62589.

Hebbachi AM, Gibbons GF. Microsomal membrane-associated apoB is the direct precursor of secreted VLDL in primary cultures of rat hepatocytes. *Journal of lipid research*. 2001;42(10):1609-17.

Helkin A, Stein JJ, Lin S, Siddiqui S, Maier KG, Gahtan V. Dyslipidemia Part 1--Review of Lipid Metabolism and Vascular Cell Physiology. *Vascular and endovascular surgery*. 2016;50(2):107-18.

Hojland Ipsen D, Tveden-Nyborg P, Lykkesfeldt J. Normal weight dyslipidemia: Is it all about the liver? *Obesity* (Silver Spring, Md). 2016;24(3):556-67.

Hon LS, Zhang Z. The roles of binding site arrangement and combinatorial targeting in microRNA repression of gene expression. *Genome biology*. 2007;8(8):R166.

Horie T, Baba O, Kuwabara Y, Chujo Y, Watanabe S, Kinoshita M, et al. MicroRNA-33 deficiency reduces the progression of atherosclerotic plaque in ApoE-/- mice. *Journal of the American Heart Association*. 2012;1(6):e003376.

Horie T, Nishino T, Baba O, Kuwabara Y, Nakao T, Nishiga M, et al. MicroRNA-33 regulates sterol regulatory element-binding protein 1 expression in mice. *Nature communications*. 2013;4:2883.

Horton JD, Goldstein JL, Brown MS. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *The Journal of clinical investigation*. 2002;109(9):1125-31.

Horvath S, Dong J. Geometric interpretation of gene coexpression network analysis. *PLoS computational biology*. 2008;4(8):e1000117.

Hwang WL, Jiang JK, Yang SH, Huang TS, Lan HY, Teng HW, Yang CY, Tsai YP, Lin CH, Wang HW, Yang MH (2014) MicroRNA-146a directs the symmetric division of Snail-dominant colorectal cancer stem cells. *Nature cell biology*. 16, 268-280.

Irani S, Hussain MM. Role of microRNA-30c in lipid metabolism, adipogenesis, cardiac remodeling and cancer. *Current opinion in lipidology*. 2015;26(2):139-46.

Irani S, Iqbal J, Antoni WJ, Ijaz L, Hussain MM. microRNA-30c reduces plasma cholesterol in homozygous familial hypercholesterolemic and type 2 diabetic mouse models. *Journal of lipid research*. 2018;59(1):144-54.

Irani S, Pan X, Peck BC, Iqbal J, Sethupathy P, Hussain MM. MicroRNA-30c Mimic Mitigates Hypercholesterolemia and Atherosclerosis in Mice. *The Journal of biological chemistry*. 2016;291(35):18397-409.

Ivanovska N, Saso L, Dimitrov P. Kinase inhibitors with redox and anti-inflammatory activities. *Current topics in medicinal chemistry*. 2015;15(9):872-85.

John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human MicroRNA targets. *PLoS biology*. 2004;2(11):e363.

Joung JG, Hwang KB, Nam JW, Kim SJ, Zhang BT. Discovery of microRNA-mRNA modules via population-based probabilistic learning. *Bioinformatics* (Oxford, England). 2007;23(9):1141-7.

Kagan A, Harris BR, Winkelstein W, Jr., Johnson KG, Kato H, Syme SL, et al. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: demographic, physical, dietary and biochemical characteristics. *Journal of chronic diseases*. 1974;27(7-8):345-64.

Kanke M, Baran-Gale J, Villanueva J, Sethupathy P. miRquant 2.0: an Expanded Tool for Accurate Annotation and Quantification of MicroRNAs and their isomiRs from Small RNA-Sequencing Data. *Journal of integrative bioinformatics*. 2016;13(5):307.

Kida K, Nakajima M, Mohri T, Oda Y, Takagi S, Fukami T, et al. PPARalpha is regulated by miR-21 and miR-27b in human liver. *Pharmaceutical research*. 2011;28(10):2467-76.

Koeth RA, Levison BS, Culley MK, Buffa JA, Wang Z, Gregory JC, Org E, Wu Y, Li L, Smith JD, Tang WH, DiDonato JA, Lusis AJ, Hazen SL (2014) gamma-Butyrobetaine is a proatherogenic intermediate in gut microbial metabolism of L-carnitine to TMAO. *Cell Metab*. 20, 799-812.

Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, et al. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nature medicine*. 2013;19(5):576-85.

Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, et al. Combinatorial microRNA target predictions. *Nature genetics*. 2005;37(5):495-500.

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic acids research*. 2016;44(W1):W90-7.

Kulp DC, Jagalur M. Causal inference of regulator-target pairs by gene mapping of expression phenotypes. *BMC genomics*. 2006;7:125.

Kurtz CL, Fannin EE, Toth CL, Pearson DS, Vickers KC, Sethupathy P. Inhibition of miR-29 has a significant lipid-lowering benefit through suppression of lipogenic programs in liver. *Scientific reports*. 2015;5:12911.

Lai X, Schmitz U, Gupta SK, Bhattacharya A, Kunz M, Wolkenhauer O, et al. Computational analysis of target hub gene repression regulated by multiple and cooperative miRNAs. *Nucleic acids research*. 2012;40(18):8818-34.

Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet*. 11, 241-247.

Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics*. 2008;9:559.

Langille BL, Ojha M. Blood flow dynamics, atherosclerosis and bypass graft failure. *Trends in cardiovascular medicine*. 1997;7(4):111-8.

Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*. 1993;75(5):843-54.

Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature*. 2003;425(6956):415-9.

Lee Y, Jeon K, Lee JT, Kim S, Kim VN. MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO journal*. 2002;21(17):4663-70.

Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, et al. MicroRNA genes are transcribed by RNA polymerase II. *The EMBO journal*. 2004;23(20):4051-60.

Li A, Horvath S. Network module detection: Affinity search technique with the multi-node topological overlap measure. *BMC research notes*. 2009;2:142.

Li B, Zhang Z, Zhang H, Quan K, Lu Y, Cai D, et al. Aberrant miR199a-5p/caveolin1/PPARalpha axis in hepatic steatosis. *Journal of molecular endocrinology*. 2014;53(3):393-403.

Li PS, Fu ZY, Zhang YY, Zhang JH, Xu CQ, Ma YT, et al. The clathrin adaptor Numb regulates intestinal cholesterol absorption through dynamic interaction with NPC1L1. *Nature medicine*. 2014;20(1):80-6.

Li Y, Wang J, Tang Y, Han X, Liu B, Hu H, et al. Bidirectional association between nonalcoholic fatty liver disease and type 2 diabetes in Chinese population: Evidence from the Dongfeng-Tongji cohort study. *PloS one*. 2017;12(3):e0174291.

Lonardo A, Lombardini S, Scaglioni F, Ballestri S, Verrone AM, Bertolotti M, et al. Fatty liver, carotid disease and gallstones: a study of age-related associations. *World journal of gastroenterology*. 2006;12(36):5826-33.

Lonardo A, Nascimbeni F, Mantovani A, Targher G. Hypertension, diabetes, atherosclerosis and NASH: Cause or consequence? *Journal of hepatology*. 2017.

Lopez-Suarez A, Guerrero JM, Elvira-Gonzalez J, Beltran-Robles M, Canas-Hormigo F, Bascunana-Quirell A. Nonalcoholic fatty liver disease is associated with blood pressure in hypertensive and nonhypertensive individuals from the general population with normal levels of alanine aminotransferase. *European journal of gastroenterology & hepatology*. 2011;23(11):1011-7.

Loscalzo J (2011) Lipid Metabolism by Gut Microbes and Atherosclerosis. *Nature*. 472, 57-63.

Luna JM, Scheel TK, Danino T, Shaw KS, Mele A, Fak JJ, et al. Hepatitis C virus RNA functionally sequesters miR-122. *Cell.* 2015;160(6):1099-110.

Lusis AJ, Mar R, Pajukanta P. Genetics of atherosclerosis. *Annual review of genomics and human genetics*. 2004;5:189-218.

Lusis AJ. Atherosclerosis. *Nature*. 2000;407(6801):233-41.

Manjunath CN, Rawal JR, Irani PM, Madhu K. Atherogenic dyslipidemia. *Indian journal of endocrinology and metabolism*. 2013;17(6):969-76.

Marquart TJ, Allen RM, Ory DS, Baldan A. miR-33 links SREBP-2 induction to repression of sterol transporters. *Proceedings of the National Academy of Sciences of the United States of America*. 2010;107(27):12228-32.

Marquart TJ, Wu J, Lusis AJ, Baldan A. Anti-miR-33 therapy does not alter the progression of atherosclerosis in low-density lipoprotein receptor-deficient mice. *Arteriosclerosis, thrombosis, and vascular biology*. 2013;33(3):455-8.

McKinsey TA. Therapeutic potential for HDAC inhibitors in the heart. *Annual review of pharmacology and toxicology*. 2012;52:303-19.

Miao J, Ling AV, Manthena PV, Gearing ME, Graham MJ, Crooke RM, et al. Flavin-containing monooxygenase 3 as a potential player in diabetes-associated atherosclerosis. *Nature communications*. 2015;6:6498.

Miki K, Endo K, Takahashi S, Funakoshi S, Takei I, Katayama S, et al. Efficient Detection and Purification of Cell Populations Using Synthetic MicroRNA Switches. *Cell stem cell*. 2015;16(6):699-711.

Mukherji S, Ebert MS, Zheng GX, Tsang JS, Sharp PA, van Oudenaarden A. MicroRNAs can generate thresholds in target gene expression. *Nature genetics*. 2011;43(9):854-9.

Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, Sachs KV, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature*. 2010;466(7307):714-9.

Najafi-Shoushtari SH, Kristo F, Li Y, Shioda T, Cohen DE, Gerszten RE, et al. MicroRNA-33 and the SREBP host genes cooperate to control cholesterol homeostasis. *Science* (New York, NY). 2010;328(5985):1566-9.

Nelson RH. Hyperlipidemia as a risk factor for cardiovascular disease. *Primary care*. 2013;40(1):195-211.

Ng R, Wu H, Xiao H, Chen X, Willenbring H, Steer CJ, et al. Inhibition of microRNA-24 expression in liver prevents hepatic lipid accumulation and hyperlipidemia. *Hepatology* (Baltimore, Md). 2014;60(2):554-64.

O'Connor A, Quizon PM, Albright JE, Lin FT, Bennett BJ (2014) Responsiveness of cardiometabolic-related microbiota to diet is influenced by host genetics. *Mamm Genome*. 25, 583-599.

Paigen B, Morrow A, Holmes PA, Mitchell D, Williams RA. Quantitative assessment of atherosclerotic lesions in mice. *Atherosclerosis*. 1987;68(3):231-40.

Parker R, Sheth U. P bodies and the control of mRNA translation and degradation. *Molecular cell*. 2007;25(5):635-46.

Petersen CP, Bordeleau ME, Pelletier J, Sharp PA. Short RNAs repress translation after initiation in mammalian cells. *Molecular cell*. 2006;21(4):533-42.

Plaisier CL, Horvath S, Huertas-Vazquez A, Cruz-Bautista I, Herrera MF, Tusie-Luna T, et al. A systems genetics approach implicates USF1, FADS3, and other causal candidate genes for familial combined hyperlipidemia. *PLoS genetics*. 2009;5(9):e1000642.

Plump AS, Smith JD, Hayek T, Aalto-Setala K, Walsh A, Verstuyft JG, et al. Severe hypercholesterolemia and atherosclerosis in apolipoprotein E-deficient mice created by homologous recombination in ES cells. *Cell*. 1992;71(2):343-53.

Ponziani FR, Pecere S, Gasbarrini A, Ojetti V. Physiology and pathophysiology of liver lipid metabolism. *Expert review of gastroenterology & hepatology*. 2015;9(8):1055-67.

Prescott MF, Sawyer WK, Von Linden-Reed J, Jeune M, Chou M, Caplan SL, et al. Effect of matrix metalloproteinase inhibition on progression of atherosclerosis and aneurysm in LDL receptor-deficient mice overexpressing MMP-3, MMP-12, and MMP-13 and on restenosis in rats after balloon injury. *Annals of the New York Academy of Sciences*. 1999;878:179-90.

Puppione DL, Charugundla S. A microprecipitation technique suitable for measuring alpha-lipoprotein cholesterol. *Lipids*. 1994;29(8):595-7.

Qiao JH, Xie PZ, Fishbein MC, Kreuzer J, Drake TA, Demer LL, et al. Pathology of atheromatous lesions in inbred and genetically engineered mice. Genetic determination of arterial calcification. *Arteriosclerosis and thrombosis: a journal of vascular biology*. 1994;14(9):1480-97.

Raitoharju E, Lyytikainen LP, Levula M, Oksala N, Mennander A, Tarkka M, et al. miR-21, miR-210, miR-34a, and miR-146a/b are up-regulated in human

atherosclerotic plaques in the Tampere Vascular Study. *Atherosclerosis*. 2011;219(1):211-7.

Rak K, Rader DJ (2012) The diet-microbe morbid union. *Nature*. 472, 40-41.

Rayner KJ, Esau CC, Hussain FN, McDaniel AL, Marshall SM, van Gils JM, et al. Inhibition of miR-33a/b in non-human primates raises plasma HDL and lowers VLDL triglycerides. *Nature*. 2011;478(7369):404-7.

Rayner KJ, Sheedy FJ, Esau CC, Hussain FN, Temel RE, Parathath S, et al. Antagonism of miR-33 in mice promotes reverse cholesterol transport and regression of atherosclerosis. *The Journal of clinical investigation*. 2011;121(7):2921-31.

Rayner KJ, Suarez Y, Davalos A, Parathath S, Fitzgerald ML, Tamehiro N, et al. MiR-33 contributes to the regulation of cholesterol homeostasis. *Science* (New York, NY). 2010;328(5985):1570-3.

Rotllan N, Price N, Pati P, Goedeke L, Fernandez-Hernando C. microRNAs in lipoprotein metabolism and cardiometabolic disorders. *Atherosclerosis*. 2016;246:352-60.

Rotllan N, Ramirez CM, Aryal B, Esau CC, Fernandez-Hernando C. Therapeutic silencing of microRNA-33 inhibits the progression of atherosclerosis in Ldlr-/- mice--brief report. *Arteriosclerosis, thrombosis, and vascular biology*. 2013;33(8):1973-7.

Rutledge H, Baran-Gale J, de Villena FP, Chesler EJ, Churchill GA, Sethupathy P, et al. Identification of microRNAs associated with allergic airway disease using a genetically diverse mouse population. *BMC genomics*. 2015;16:633.

Sampson VB, Rong NH, Han J, Yang Q, Aris V, Soteropoulos P, et al. MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. *Cancer research*. 2007;67(20):9762-70.

Schmidt CW. Diversity outbred: a new generation of mouse model. *Environmental health perspectives*. 2015;123(3):A64-7.

Selitsky SR, Dinh TA, Toth CL, Kurtz CL, Honda M, Struck BR, et al. Transcriptomic Analysis of Chronic Hepatitis B and C and Liver Cancer Reveals MicroRNA-Mediated Control of Cholesterol Synthesis Programs. *mBio*. 2015;6(6):e01500-15.

Sen S, Churchill GA (2001) A statistical framework for quantitative trait mapping. *Genetics*. 159, 371-387.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*. 2003;13(11):2498-504.

Shi W, Wang X, Tangchitpiyanond K, Wong J, Shi Y, Lusis AJ. Atherosclerosis in C3H/HeJ mice reconstituted with apolipoprotein E-null bone marrow. *Arteriosclerosis, thrombosis, and vascular biology*. 2002;22(4):650-5.

Shirasaki T, Honda M, Shimakami T, Horii R, Yamashita T, Sakai Y, et al. MicroRNA-27a regulates lipid metabolism and inhibits hepatitis C virus replication in human hepatoma cells. *Journal of virology*. 2013;87(9):5270-86.

Siciliano V, Garzilli I, Fracassi C, Criscuolo S, Ventre S, di Bernardo D. MiRNAs confer phenotypic robustness to gene networks by suppressing biological noise. *Nature communications*. 2013;4:2364.

Sieberts SK, Schadt EE. Moving toward a system genetics view of disease. *Mammalian genome : official journal of the International Mammalian Genome Society*. 2007;18(6-7):389-401.

Simionescu N, Niculescu LS, Sanda GM, Margina D, Sima AV (2014) Analysis of circulating microRNAs that are specifically increased in hyperlipidemic and/or hyperglycemic sera. *Molecular biology reports*. 41, 5765-5773.

Singaravelu R, Chen R, Lyn RK, Jones DM, O'Hara S, Rouleau Y, et al. Hepatitis C virus induced up-regulation of microRNA-27: a novel mechanism for hepatic steatosis. *Hepatology* (Baltimore, Md). 2014;59(1):98-108.

Sinn DH, Kang D, Chang Y, Ryu S, Gu S, Kim H, et al. Non-alcoholic fatty liver disease and progression of coronary artery calcium score: a retrospective cohort study. *Gut*. 2017;66(2):323-9.

Smallwood TL, Gatti DM, Quizon P, Weinstock GM, Jung KC, Zhao L, et al. High-resolution genetic mapping in the diversity outbred mouse population identifies Apobec1 as a candidate gene for atherosclerosis. *G3* (Bethesda, Md). 2014;4(12):2353-63.

Soh J, Iqbal J, Queiroz J, Fernandez-Hernando C, Hussain MM. MicroRNA-30c reduces hyperlipidemia and atherosclerosis in mice by decreasing lipid synthesis and lipoprotein secretion. *Nature medicine*. 2013;19(7):892-900.

Svenson KL, Gatti DM, Valdar W, Welsh CE, Cheng R, Chesler EJ, et al. High-resolution genetic mapping using the Mouse Diversity outbred population. *Genetics*. 2012;190(2):437-47.

Tang WH, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, et al. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *The New England journal of medicine*. 2013;368(17):1575-84.

van den Maagdenberg AM, Hofker MH, Krimpenfort PJ, de Bruijn I, van Vlijmen B, van der Boom H, et al. Transgenic mice carrying the apolipoprotein E3-Leiden gene exhibit hyperlipoproteinemia. *The Journal of biological chemistry*. 1993;268(14):10540-5.

Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. *Nature reviews Genetics*. 2009;10(4):252-63.

Vera J, Lai X, Schmitz U, Wolkenhauer O. MicroRNA-regulated networks: the perfect storm for classical molecular biology, the ideal scenario for systems biology. *Advances in experimental medicine and biology*. 2013;774:55-76.

Vickers KC, Shoucri BM, Levin MG, Wu H, Pearson DS, Osei-Hwedieh D, et al. MicroRNA-27b is a regulatory hub in lipid metabolism and is altered in dyslipidemia. *Hepatology* (Baltimore, Md). 2013;57(2):533-42.

Wagschal A, Najafi-Shoushtari SH, Wang L, Goedeke L, Sinha S, deLemos AS, et al. Genome-wide identification of microRNAs regulating cholesterol and triglyceride homeostasis. *Nature medicine*. 2015;21(11):1290-7.

Wakiyama M, Takimoto K, Ohara O, Yokoyama S. Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. *Genes & development*. 2007;21(15):1857-62.

Wang B, Hsu SH, Majumder S, Kutay H, Huang W, Jacob ST, et al. TGFbeta-mediated upregulation of hepatic miR-181b promotes hepatocarcinogenesis by targeting TIMP3. *Oncogene*. 2010;29(12):1787-97.

Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature*. 2011;472(7341):57-63.

Wei J, Fu ZY, Li PS, Miao HH, Li BL, Ma YT, et al. The clathrin adaptor proteins ARH, Dab2, and numb play distinct roles in Niemann-Pick C1-Like 1 versus low density lipoprotein receptor-mediated cholesterol uptake. *The Journal of biological chemistry*. 2014;289(48):33689-700.

Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell*. 1993;75(5):855-62.

Willeit P, Skroblin P, Kiechl S, Fernandez-Hernando C, Mayr M. Liver microRNAs: potential mediators and biomarkers for metabolic and cardiovascular disease? *European heart journal*. 2016;37(43):3260-6.

Wu S, Huang S, Ding J, Zhao Y, Liang L, Liu T, et al. Multiple microRNAs modulate p21Cip1/Waf1 expression by directly targeting its 3' untranslated region. *Oncogene*. 2010;29(15):2302-8.

Xu J, Li CX, Li YS, Lv JY, Ma Y, Shao TT, et al. MiRNA-miRNA synergistic network: construction via co-regulating functional modules and disease miRNA topological features. *Nucleic acids research*. 2011;39(3):825-36.

Yang X, Deignan JL, Qi H, Zhu J, Qian S, Zhong J, et al. Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks. *Nature genetics*. 2009;41(4):415-23.

Yang ZX, Wang YZ, Jia BB, Mao GX, Lv YD, Wang GF, Yu H (2016) Downregulation of miR-146a, cyclooxygenase-2 and advanced glycation end-products in simvastatin-treated older patients with hyperlipidemia. *Geriatrics & gerontology international*. 16, 322-328.

Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development*. 2003;17(24):3011-6.

Yu F, Lu Z, Chen B, Dong P, Zheng J. Identification of a Novel lincRNA-p21-miR-181b-PTEN Signaling Cascade in Liver Fibrosis. *Mediators of inflammation*. 2016;2016:9856538.

Zeb I, Li D, Budoff MJ, Katz R, Lloyd-Jones D, Agatston A, et al. Nonalcoholic Fatty Liver Disease and Incident Cardiac Events: The Multi-Ethnic Study of Atherosclerosis. *Journal of the American College of Cardiology*. 2016;67(16):1965-6.

Zhang JY, Qin X, Park HG, Kim E, Liu G, Kothapalli KS, et al. Alternative splicing generates novel Fads3 transcript in mice. *Molecular biology reports*. 2016;43(8):761-6.

Zhang M, Wu JF, Chen WJ, Tang SL, Mo ZC, Tang YY, et al. MicroRNA-27a/b regulates cellular cholesterol efflux, influx and esterification/hydrolysis in THP-1 macrophages. *Atherosclerosis*. 2014;234(1):54-64.

Zhang QQ, Lu LG. Nonalcoholic Fatty Liver Disease: Dyslipidemia, Risk for Cardiovascular Complications, and Treatment Strategy. *J Clin Transl Hepatol*. 2015;3(1):78-84.

Zhang SH, Reddick RL, Piedrahita JA, Maeda N. Spontaneous hypercholesterolemia and arterial lesions in mice lacking apolipoprotein E. *Science* (New York, NY). 1992;258(5081):468-71.