

INTEGRATED ANALYSIS OF GENETIC MARKER, MIRNA, AND MRNA DATA TO
UNRAVEL MECHANISMS CONTROLLING GROWTH AND MEAT QUALITY TRAITS IN
PIGS

By

Kaitlyn R. Daza

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Animal Science – Doctor of Philosophy

2021

ABSTRACT

INTEGRATED ANALYSIS OF GENETIC MARKER, MIRNA, AND MRNA DATA TO UNRAVEL MECHANISMS CONTROLLING GROWTH AND MEAT QUALITY TRAITS IN PIGS

By

Kaitlyn R. Daza

Determining mechanisms regulating complex traits in pigs is essential to improve the production efficiency of this globally important protein source. MicroRNAs (miRNAs) are a class of non-coding RNAs known to post-transcriptionally regulate gene expression affecting numerous tissues and phenotypes, including those important to the pig industry. However, further research is needed to characterize the miRNAs expressed in pig skeletal muscle and assess their impact on the regulation of growth, carcass composition, and meat quality traits. Additionally, little is known about the genetic architecture controlling miRNA expression in pigs, which can be elucidated by combining high-density genotypic data with miRNA expression profiles from the same animals in the mapping of miRNA expression quantitative trait loci (miR-eQTL). This analysis reveals associations between genomic regions harboring single nucleotide polymorphisms (SNPs) and variation in the expression of miRNAs. By integrating mRNA expression profiles and phenotypic data from the same individuals, putative regulatory relationships can be revealed underlying variation in phenotypes relevant to pig production. In this study, our objectives were to profile and characterize the small RNA population present in *longissimus dorsi* (LD) skeletal muscle samples from a F₂ Duroc x Pietrain resource population, and to conduct an integrated miR-eQTL analysis to identify regulators of miRNA expression and candidate genes regulating phenotypic traits in adult pig skeletal muscle.

MicroRNA expression profiling was performed on total RNA extracted from (LD) muscle samples from 174 F₂ pigs. The composition of small RNA classes present in this dataset was characterized through a series of homology searches against human, mouse, and pig databases. MicroRNA quantification and novel miRNA prediction was conducted, profiling the abundance of 295 known mature pig miRNAs and producing 27 unique candidate novel miRNA precursors. The 295 miRNA expression profiles were subsequently used as response variables in a GBLUP-based GWA analysis. Results for associating these miRNAs with 36,292 SNPs identified 315 significant miRNA-SNP associations ($FDR < 0.05$), comprising 23 significant eQTL peaks associated with 17 unique miRNAs. Five of the 23 miR-eQTL peaks were defined as *local*-acting, meaning the genomic positions of the significantly associated SNPs comprising the miR-eQTL peak overlapped that of the miRNA precursor transcript. We then investigated the potential effects of these miRNAs through miRNA target prediction, correlation, and colocalization analyses. Notably, one miR-eQTL miRNA exhibiting a strong *local*-acting miR-eQTL, miR-874, had predicted target genes colocalizing with previously identified phenotypic QTL for 12 production traits including backfat thickness, dressing percentage, muscle pH at 24 h post-mortem, and cook yield.

The results of this study revealed putative pig-novel miRNAs for further study and validation, contributing to our understanding of the miRNA landscape present in adult pig skeletal muscle. Additionally, we identified genomic regions underlying variation in miRNA expression, and candidate miRNAs and genes for future investigation of their regulatory effects on growth, carcass composition, and meat quality traits of importance to the global pig industry.

This work is dedicated to my family, especially my nephews and niece, who brought me such joy and were my motivation to complete my PhD.

Always know that if you put your mind to it, you can accomplish anything.

ACKNOWLEDGEMENTS

There are so many people I need to thank for their support and encouragement, more so than I have room for in this document. I would like to thank Dr. Cathy Ernst for seeing my potential (before I could) and guiding me through this journey. I would also like to acknowledge the other members of my guidance committee: Dr. Juan Steibel, for his many hours of assistance in learning statistical modeling and R programming; Dr. Ron Bates, for working patiently with me when I was first starting my research journey; and Dr. Chen Chen, for his assistance in my transition to the wet lab and for always being a kind, smiling face to see in the hallway.

Special thanks to Nancy Raney for being my guardian angel of the lab countless times and for always encouraging me as I learned new techniques. To Dr. Steve Bursian, whose kindness and empathy helped me through the most difficult days of graduate school: Thank you for always providing a safe space and for being a friend to lean on. I would also like to thank Dr. Tasia Kendrick for her encouragement in developing my passion for teaching and communications.

I would be remiss not to thank my friends, whose support and comradery helped make this journey the formative experience that it was. To my lab mates Ryan Corbett, Scott Funkhouser, Andrea Luttmann, and Deborah Velez Irizarry: Thank you for supporting me and pushing me to be a better scientist (and person). Thanks also to Sebastian Casiró, Kaitlin Wurtz, and Carly O'Malley for bringing laughter and light to my darkest days.

Finally, I need to thank my family, especially my husband, Greg. You've seen me at my worst and my best and loved and encouraged me throughout. I could not have completed my PhD without you.

TABLE OF CONTENTS

LIST OF TABLES.....	ix
LIST OF FIGURES	x
CHAPTER ONE.....	1
INTRODUCTION	1
1.1 Advancements in genetic selection of pigs: Quantitative Trait Loci.....	1
1.2 Expression QTL: uncovering the genetic architecture of gene expression	2
1.3 MicroRNAs: Discovery, biology, and implications in pig skeletal muscle	4
1.4 Integrated analysis of genetic variation, mRNA, and miRNA	7
REFERENCES	12
CHAPTER TWO.....	19
PROFILING AND CHARACTERIZATION OF A <i>LONGISSIMUS DORSI</i>	
MUSCLE MICRORNA DATASET FROM AN F ₂ DUROC X PIETRAIN PIG	
RESOURCE POPULATION	19
2.1 Abstract.....	19
2.2 Direct link to deposited data	20
2.3 Experimental design, materials and methods	20
2.3.1 Sample collection, RNA isolation and sequencing	20
2.3.2 Analysis of sequencing data	21
2.3.3 Characterization of small RNAs.....	21
2.4 Data description.....	23
2.5 Acknowledgements	26
APPENDIX	29
REFERENCES	31
CHAPTER THREE	34
INTEGRATED GENOME-WIDE ANALYSIS OF MICRORNA EXPRESSION	
QUANTITATIVE TRAIT LOCI IN PIG <i>LONGISSIMUS DORSI</i> MUSCLE	34
3.1 Abstract.....	34
3.2 Introduction	35
3.3 Materials and methods.....	37
3.3.1 Data collection and sequencing	37
3.3.2 Bioinformatics analysis	38
3.3.3 MicroRNA eQTL analysis.....	39
3.3.4 Colocalization of miR-eQTL and pQTL	42
3.3.5 Genomic colocalization of miR-eQTL-correlated	
target mRNAs and pQTL	43
3.4 Results	44
3.4.1 MicroRNA eQTL analysis.....	44
3.4.1.1 Local and distant microRNA eQTL	44

3.4.1.2 Heritability of miRNA expression in pig longissimus dorsi	49
3.4.2 Peak miR-eQTL SNP conditional analysis	49
3.4.3 Colocalization of miR-eQTL and pQTL	53
3.4.4 Genomic colocalization of miR-eQTL target mRNAs and pQTL	54
3.5 Discussion.....	58
3.6 Conclusion	66
3.7 Author contributions.....	66
3.8 Acknowledgements	67
APPENDIX	68
REFERENCES	70
 CHAPTER FOUR	 77
CONCLUSIONS AND FUTURE DIRECTIONS	77
4.1 Conclusions	77
4.2 Future Directions	79
REFERENCES	83

LIST OF TABLES

Table 2.1	Candidate novel miRNA precursors predicted from miRDeep2 and characterized with BLAST+	27
Table 3.1	Summary of miRNA expression Quantitative Trait Loci.....	48
Table 3.2	Summary of miR-eQTL peak SNP conditional analysis results	52
Table 3.3	Correlations of miR-874, target genes, and co-localized phenotypes	57

LIST OF FIGURES

Figure 2.1	Characterization of small RNA sequencing data.....	24
Figure 3.1	Global plot of genomic position of miRNA transcript (Mb) versus genomic position of SNP (Mb) for miR-eQTL	45
Figure 3.2	Manhattan plots of the five local-acting miR-eQTL	47
Figure 3.3	Conditional Analyses of miR-eQTL peak SNPs	51
Figure 3.4	Manhattan plot of genomic co-localization events of miR-874 target genes with pQTL	56
Figure 4.1	Localization of CITK in early and late telophase.....	80

CHAPTER ONE

INTRODUCTION

1.1 Advancements in genetic selection of pigs: Quantitative Trait Loci

The United States pork industry produced over \$19.1 billion in gross income in 2020, making it a key contributor to the prosperity of the American agricultural economy (USDA – NASS 2021). The incorporation of genomic selection techniques has stimulated significant improvement in livestock production, including in traits of importance to pork producers and processors such as meat quality and carcass characteristics (Meuwissen et al. 2016). However, historical selection for increased carcass leanness has adversely affected meat quality phenotypes, increasing the incidence of inferior eating-quality pork and decreasing consumer satisfaction (Barbut et al. 2008; Schwab and Baas 2008; Ciobanu et al. 2011). Increased demand for high-quality products has placed growing emphasis on improving meat quality and consistency while reducing the presence of inferior quality pork (Barbut et al. 2008). Thus, understanding the complex genetic regulation of economically important production traits continues to be a priority for food animal producers.

Genetic improvement of livestock has historically been achieved using traditional animal breeding techniques, combining phenotypic data with pedigree information to estimate animals' individual breeding values. While these methods have been successful in achieving genetic gain, advancements in molecular genetic technologies facilitated the identification of quantitative trait loci (QTL); polymorphic loci or “markers” along the genome significantly associated with variation in quantitative traits. These methods integrate molecular genetics approaches into traditional animal breeding, accelerating improvement through the targeted selection of

genetically superior breeding stock. Many QTL studies have been conducted in the decades since the first identified pig QTL in 1994 (Andersson et al.), with the goal of understanding genetic mechanisms controlling variation in complex, economically important pig phenotypes. There are currently 33,143 QTL reported in the PigQTLdb database from nearly 733 publications, with over 17,000 QTL representing meat and carcass traits (<http://www.animalgenome.org/cgi-bin/QTLdb/SS/index>; release 44, April 26, 2021). Despite the success of QTL identification in pigs, due to most QTL having small effect on their associated traits, causative variants associated with known genes have been validated for relatively few of these QTL.

1.2 Expression QTL: uncovering the genetic architecture of gene expression

The availability of high-density single nucleotide polymorphism (SNP) panels and advancements in RNA sequencing technologies enable the mapping of expression quantitative trait loci (eQTL) in segregating populations (Jansen and Nap 2001; Schadt et al. 2003). In an eQTL analysis, gene expression profiles are used as response variables in genome-wide association models. By investigating variations in gene expression between individuals, significantly associated genomic loci – here, SNP markers – can be identified. These SNPs can then be mapped to the genome and identified as affecting gene expression *locally*, near the gene-of-origin, or *distantly*, far from the gene-of-origin (Jansen and Nap, 2001; Kadarmideen et al. 2006). This method may also provide insight into the genetic architecture of phenotypic variation through the integration of phenotypic QTL (pQTL) from the same population.

Many eQTL studies have been conducted in pigs, profiling tissues including skeletal muscle (Ponsuksili et al. 2008, 2010, 2012, 2014; Liaubet et al. 2011; Steibel et al. 2011; Cánovas et al. 2012, Velez-Irizarry et al. 2019). Ponsuksili and collaborators (2008) utilized

microarrays to profile gene expression in *longissimus dorsi* (LD) samples from 74 F2 pigs, identifying 104 eQTL overlapping pQTL associated with water holding capacity (Ponsuksili et al. 2008). Additionally, Steibel et al. (2011) utilized a whole-genome expression microarray to conduct an eQTL analysis in an F2 pig resource population, identifying 59 eQTL (FDR < 10%) aligned to known genomic positions. Upon co-localizing these eQTL with previously identified pQTL, the eQTL were found to regulate genes associated with lipid metabolism, DNA replication, and cell cycle regulation (Steibel et al. 2011). Ponsuksili and colleagues (2014) integrated eQTL data from LD tissue of two pig populations to further define results from a QTL scan of the same animals, identifying candidate genes influencing multiple meat quality phenotypes.

Our group has utilized the Michigan State University Pig Resource Population (MSUPRP), a F₂ Duroc x Pietrain pig population, for over a decade to identify pQTL (Edwards et al. 2008a, 2008b; Choi et al. 2010, 2011; Gualdrón-Duarte et al. 2016; Casiró et al. 2017) and eQTL (Steibel et al. 2011; Velez-Irizarry et al. 2019) associated with important production traits. These two pig breeds are utilized in commercial systems worldwide, differing in growth rate, carcass leanness and meat quality traits (Edwards et al. 2006). Most recently in 2019, Velez-Irizarry et al. (2019) identified 339 eQTL in LD tissue of 168 F2 pigs, revealing 16 genes co-localizing with 21 pQTL for meat quality, carcass composition, and growth traits. These studies demonstrate the successful identification of genomic regions harboring variants associated with gene expression variation and downstream phenotypes of economic importance to the pig production industry.

1.3 MicroRNAs: Discovery, biology, and implications in pig skeletal muscle

One mechanism of genetic regulation potentially governing complex pig phenotypes is the post-transcriptional silencing of genes via microRNAs (miRNAs). MicroRNAs are a class of single-stranded, small non-coding RNAs known to regulate various biological processes through complementarity of the miRNA “seed” sequence with target mRNA sequences. The first miRNAs were identified in the *lin-4* gene in *C. elegans* in 1993, wherein two small transcripts were found containing sequences complementary to the 3' UTR of *lin-14* mRNA. These transcripts were discovered to suppress the expression of the *lin-14* gene, and were later termed *microRNAs* (Lee et al. 1993). MicroRNAs have since been shown to regulate a multitude of biological processes through recognition of target genes via sequence complementarity and are expressed in a spatiotemporal manner, indicating that they have specific regulatory functions.

The biogenesis of miRNAs is well understood (for review see Liu et al. 2010; Ha and Kim 2014). In animals, a long (> 1 kb) primary miRNA transcript (pri-miRNA) is initially transcribed in the nucleus by RNA polymerase II. The pri-miRNA is processed by the Microprocessor complex of Drosha, an RNase III family endonuclease, and DGCR8 proteins, producing a ~70-nt hairpin pre-miRNA (Lee et al. 2003). The pre-miRNA is then exported from the nucleus by the protein exportin 5 (EXP5; Yi et al. 2003), and is cleaved by Dicer proteins in the cytoplasm to form a 19 – 24 nt duplex containing the mature miRNA “guide” and its complementary “star” strand (Ketting et al. 2001; Soifer et al. 2008). The mature miRNA is preferentially loaded onto argonaute proteins, and the RNA-Induced Silencing Complex (RISC) is assembled (Hammond et al. 2001). The RISC complex, coupled with guide miRNA, then targets mRNA(s) exhibiting complementary sequence to the 6 – 8 nt miRNA “seed” sequence,

typically located in the 3' UTR of the mRNA (for review, see Carthew and Sontheimer 2009). Post-transcriptional suppression of target mRNA expression then occurs through multiple mechanisms, including transcript degradation through mRNA cleavage or destabilization, and repression of ribosomal translation (Behm-Ansmant et al. 2006; Mohr and Mott 2015).

Research efforts began with miRNA discovery and profiling in multiple tissues. Primary miRNA discovery methods involved homology searches of known miRNA in other species, typically in combination with microarray expression data. In vertebrates, miRNAs have been most extensively studied in human and mouse. The official miRNA database, miRBase, currently contains 2,654 mature human miRNAs (Release 22.1, October 2018; <http://www.mirbase.org/index.shtml>; Griffiths-Jones 2006; Kozomara et al. 2019). More recently, the focus of human miRNA research has turned to their utilization as potential biomarkers and targeted treatments for cancers and other diseases (Mollaei et al. 2018). Discovery of miRNAs in livestock and their roles in animal agriculture are also currently being studied. The miRBase database currently contains 457 mature pig miRNAs (Release 22.1, October 2018; <http://www.mirbase.org/index.shtml>; Griffiths-Jones 2006; Kozomara et al. 2019). The first pig miRNA (mir17-92) was identified in 2005 by Sawera and colleagues using homology searching of human miRNAs and northern blot analysis of five porcine tissues. Wernersson et al. (2005) utilized sequence similarity searching of approximately 3.8 million shotgun sequences of the pig genome to identify 51 mature pig miRNAs, observing that miRNA sequences are strongly evolutionarily conserved between pig, human, and mouse species. Kim et al. (2006) also utilized homology of human and mouse sequences to identify 58 additional pig miRNAs in five pig tissues.

As expression profiling technologies advanced, larger numbers of pig miRNAs were discovered. Huang et al. (2008) found over 700 candidate pig miRNAs, 296 of which were validated in fetal porcine skeletal muscle tissue via mammalian miRNA microarrays. Xie et al. (2011) performed next-generation sequencing of a total RNA library pooled from 16 pig tissues, identifying 437 conserved and 86 candidate novel miRNAs (Xie et al. 2011). Eight candidate novel miRNAs were subsequently validated using stem-loop RT-PCR (Xie et al. 2011). Additionally, Li et al. (2012) utilized small RNA sequencing and miRNA microarrays to isolate 184 known and 521 candidate miRNAs from pooled libraries of pig skeletal muscle and adipose tissues. More recently, machine learning approaches have been developed to predict novel miRNAs in the pig genome. By adopting a semi-supervised transductive learning approach, the *eMIRNA* pipeline utilizes positive, negative, and unlabeled training datasets to overcome limitations of miRNA precursor prediction caused by the lack of miRNA annotation in the pig genome (<https://github.com/emarmolsanchez/eMIRNA/>; Mármol-Sánchez et al. 2020). Twenty novel expressed miRNAs were identified using this pipeline from small RNA sequencing data obtained from *gluteus medius* muscle samples of 48 Duroc gilts (Mármol-Sánchez et al. 2020). Advancements in miRNA prediction techniques and improved genome annotation offer possibility for the continued identification of pig-novel miRNAs.

Research efforts continue to focus on understanding miRNA biology in various tissues and developmental or physiological stages. Multiple studies have investigated miRNA expression in pig skeletal muscle. Huang et al. (2008) identified 140 differentially expressed (DE) miRNAs between two fetal time points and adult tissue in LD muscle tissue, validating five through real-time PCR. McDanel et al. (2009) profiled miRNAs in skeletal muscle satellite cells derived from 8-wk-old pigs, *biceps femoris* (BF) and LD from fetuses at 60-, 90- and 105-d

of gestation, and BF of neonate and adult pigs, identifying several muscle-specific and ubiquitous miRNAs and 12 potentially novel miRNAs. Yang and colleagues (2014) characterized the dynamic expression of miR-127 during pre- and postnatal skeletal muscle development, profiling miRNA expression in LD tissue at 26 developmental stages and eight adult tissues with stem-loop quantitative real-time PCR. Cai and colleagues (2014) discovered seven DE miRNAs in skeletal muscle between castrated and intact male pigs (fold change ≥ 2). Subsequent target prediction and integrated network analysis revealed miRNA-mRNA pairs regulating skeletal muscle contractile function and lipid synthesis, further elucidating the genetic effect of castration on skeletal muscle growth (Cai et al. 2014).

While evidence demonstrates that miRNA regulation effects multiple developmental and biological processes, further research is needed to fully understand the impacts of miRNA regulation on traits important to the pig industry. In Chapter 2 of this dissertation, we characterized the small RNA populations present in LD tissue of 176 F₂ MSUPRP pigs using RNA-sequencing. We successfully identified different classes of small RNA, profiled known pig miRNAs, and predicted putative novel miRNAs in these samples, improving our knowledge of these important regulatory RNAs in skeletal muscle of pig.

1.4 Integrated analysis of genetic variation, mRNA, and miRNA

A logical progression of research following miRNA discovery and profiling is to consider the expression of miRNA and mRNA together in disease, metabolic, and developmental states. Many studies compare miRNA-mRNA expression profiles between normal and disease states in human and mouse tissues (Boren et al. 2008; Dong et al. 2010; Lawless et al. 2013; Quitadamo et al. 2015). Kara and colleagues (2015) utilized high-throughput real-time PCR to compare

expression profiles of mRNA and select miRNAs targeting colorectal cancer-associated genes between case and control samples, identifying five genes and 18 miRNAs significantly DE between the two states. Co-expression approaches have also been used to identify regulatory pathways in immune response (Lukowski et al. 2015), evaluate drug candidates (for review, see Schmidt 2014; Gmeiner et al. 2010), and to study the regulation of mesenchymal stem cells in humans and pigs (Giraud-Triboult et al. 2011; Eirin et al. 2014).

Co-expression profiling has also been conducted in skeletal muscle of various species. Johnston et al. (2009) used microarrays to determine global changes in miRNA and mRNA expression associated with the transition from hyperplasic to hypertrophic muscle growth in zebrafish. Additionally, Gallagher et al. (2010) used microarrays in clinical trials comparing mRNA and miRNA profiles in *vastus lateralis* muscle of insulin-resistant type 2 diabetic and control patients, identifying 62 miRNAs with altered expression but no differences in the mRNA transcriptome. Szeto et al. (2014) utilized next-generation sequencing to profile mRNA and miRNA expression in two cell lines and one xenograft associated with human nasopharyngeal carcinoma. Subsequent analysis identified 533 mRNA-miRNA pairs inversely regulated in all three model systems compared to a negative control, aiding in the development of a more complete characterization of the genetic regulation of this cancer (Szeto et al. 2014).

Integrating mRNA and miRNA profiles with genetic marker data provides further insight into the mechanisms underlying complex trait phenotypes, enhancing results obtained from eQTL studies. Allelic variation in miRNA target sites or at miRNA loci have been demonstrated to contribute to phenotypic differences in both human and livestock traits (for review see Liu et al. 2010). Clop and colleagues (2006) investigated a SNP in the myostatin gene of sheep

creating illegitimate miRNA target sites for miR-1 and miR-206 and contributing to the muscular hypertrophy of the Texel breed. More recently, Lei et al. (2010) identified a SNP at the miR-27a locus significantly associated with litter size in a Meishan pig population. In humans, variations in longevity phenotypes were associated with SNPs located in the 3'UTRs of two genes; SIRT2, whose functions involve cellular metabolism and differentiation, and DRD2, whose signaling regulates multiple physiological processes including locomotion and behavior. Individuals expressing the minor allele for either of these 3' UTR SNPs observed a decreased probability of becoming long-lived, indicating that miRNA targeting may affect the human aging process (Crocco et al. 2015).

Due to the analytical complexity of studies involving integration of genetic marker, mRNA and miRNA data, fewer studies of this nature have been reported. Dong et al. (2010) used this approach to construct co-expression networks from glioblastoma patient data, identifying a large number of *cis*- and *trans*-acting eQTL. Saba et al. (2010) built upon their previous eQTL studies examining quantitative variations in animal alcohol consumption by comparing eQTL, mRNA expression, and protein expression of candidate genes in breeds of high or low alcohol-consuming inbred mice strains, subsequently investigating miRNA regulation of the candidate gene, *Gnb1*. Quitadamo and colleagues (2015) combined multiple types of data including miRNA eQTL, miRNA-target interactions, protein-protein interactions, and correlations between miRNA and gene expression to develop an integrated network elucidating the effect of miRNA expression on ovarian cancer phenotypes. More recently, Ponsuksili and colleagues (2017) identified 221 miRNA-eQTL associated with 108 miRNAs in liver samples from 209 German Landrace pigs. Putative miRNA-mediated regulatory mechanisms affecting biomedical and hematological traits of interest to human medicine (e. g.

blood urea nitrogen, total cholesterol, white blood cell count, etc.) were revealed by integration of mRNA and phenotypic data from the same individuals (Ponsuksili et al. 2017). The same group later conducted integrated miRNA-eQTL analyses in ileal epithelium samples of 482 Japanese Quail, identifying 34 miRNA-eQTL miRNAs and 62 mRNA-eQTL mRNAs. Integration of microbiome data revealed networks of genes and miRNAs affecting phosphorus utilization and microbiome composition, with the potential to minimize environmental impacts of poultry production (Ponsuksili et al. 2021). These integrated studies aid in elucidating regulatory interactions underlying complex phenotypes. However, prior to this dissertation research no studies had been published regarding the global integration of genetic marker, miRNA, mRNA, and phenotype data in traits relevant to food production.

The goal of this dissertation research was to improve our understanding of molecular mechanisms controlling phenotypic variation in growth, carcass composition, and meat quality traits in pigs. To accomplish this, we integrated high-density SNP genotype data, miRNA and mRNA RNA-seq data, and phenotype data from 176 F₂ MSUPRP pigs. Through this integration we identified genomic regions underlying variation in miRNA expression, and candidate miRNAs and genes for further assessment of their regulatory effects on traits important to the pig industry.

Specifically, our aims included:

1. Profile and characterize the small RNA population present in LD skeletal muscle samples of 176 F₂ Duroc x Pietrain pigs.

2. Conduct an integrated miRNA-eQTL analysis utilizing small RNA sequencing data to identify regulators of miRNA expression and candidate genes regulating phenotypic traits in LD skeletal muscle samples of 176 F₂ Duroc x Pietrain pigs.

In Chapter 2 of this dissertation, we characterized the composition of small RNA classes present in adult pig skeletal muscle through a series of homology searches against human, mouse, and pig databases. MicroRNA quantification and novel miRNA prediction were also conducted using the miRDeep2 software package (Friedländer et al. 2012), profiling the abundance of 295 known mature pig miRNAs and producing 27 unique candidate novel miRNA precursors.

In Chapter 3 of this work, we identified genomic regions underlying variation in miRNA expression by using 295 miRNA expression profiles as response variables in a genomic best linear unbiased prediction (GBLUP) -based genome-wide association (GWA) analysis. We identified 315 significant miRNA-SNP associations ($FDR < 0.05$), comprising 23 significant eQTL peaks associated with 17 unique miRNAs. and revealed candidate miRNAs and genes for future validation and assessment of their regulatory effects on traits of importance to the global pig industry.

REFERENCES

REFERENCES

1. Andersson, L., Haley, C. S., Ellegren, H., Knott, S. a, Johansson, M., Andersson, K., et al. (1994). Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science* 263, 1771–4.
2. Barbut, S., Sosnicki, A. A., Lonergan, S. M., Knapp, T., Ciobanu, D. C., Gatcliffe, L. J., et al. (2008). Progress in reducing the pale, soft and exudative (PSE) problem in pork and poultry meat. *Meat Sci.* 79, 46–63.
3. Behm-Ansmant, I., Rehwinkel, J., and Izaurralde, E. (2006). MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay. *Cold Spring Harb. Symp. Quant. Biol.* 71, 523–530.
4. Boren, T., Xiong, Y., Hakam, A., Wenham, R., Apte, S., Wei, Z., et al. (2008). MicroRNAs and their target messenger RNAs associated with endometrial carcinogenesis. *Gynecol. Oncol.* 110, 206–215.
5. Cai, Z., Zhang, L., Jiang, X., Sheng, Y., and Xu, N. (2014). Differential miRNA expression profiles in the longissimus dorsi muscle between intact and castrated male pigs. *Res. Vet. Sci.* 99, 99–104.
6. Cánovas, A., Pena, R. N., Gallardo, D., Ramí, O., and Amills, M. (2012). Segregation of Regulatory Polymorphisms with Effects on the Gluteus Medius Transcriptome in a Purebred Pig Population. *PLoS One* 7, 1–12.
7. Casiró, S., Velez-Irizarry, D., Ernst, C. W., Raney, N. E., Bates, R. O., Charles, M. G., et al. (2017). Genome-wide association study in an F2 Duroc x Pietrain resource population for economically important meat quality and carcass traits1. *J. Anim. Sci.* 95, 545–558.
8. Choi, I., Steibel, J. P., Bates, R. O., Raney, N. E., Rumph, J. M., and Ernst, C. W. (2011). Identification of Carcass and Meat Quality QTL in an F(2) Duroc × Pietrain Pig Resource Population Using Different Least-Squares Analysis Models. *Front. Genet.* 2, 18.
9. Choi, I., Steibel, J. P., Bates, R. O., Raney, N. E., Rumph, J. M., and Ernst, C. W. (2010). Application of alternative models to identify QTL for growth traits in an F2 Duroc x Pietrain pig resource population. *BMC Genet.* 11, 97.
10. Ciobanu, D. C., Lonergan, S. M., and Huff-Lonergan, E. J. (2011). “Genetics of meat quality and carcass traits,” in *The Genetics of the Pig*, eds M. F. Rothschild and A. Ruvinsky (New York, NY: CAB International), 355.

11. Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibé, B., et al. (2006). A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat. Genet.* 38, 813–818.
12. Crocco, P., Montesanto, a., Passarino, G., and Rose, G. (2015). Polymorphisms Falling Within Putative miRNA Target Sites in the 3'UTR Region of SIRT2 and DRD2 Genes Are Correlated With Human Longevity. *Journals Gerontol. Ser. A Biol. Sci. Med. Sci.*, 1–7.
13. Dong, H., Luo, L., Hong, S., Siu, H., Xiao, Y., Jin, L., et al. (2010). Integrated analysis of mutations, miRNA and mRNA expression in glioblastoma. *BMC Syst. Biol.* 4, 163.
14. Edwards, D. B., Ernst, C. W., Raney, N. E., Doumit, M. E., Hoge, M. D., and Bates, R. O. (2008). Quantitative trait locus mapping in an F2 Duroc × Pietrain resource population: II. Carcass and meat quality traits1. *J. Anim. Sci.* 86, 254–266.
15. Edwards, D. B., Ernst, C. W., Tempelman, R. J., Rosa, G. J. M., Raney, N. E., Hoge, M. D., et al. (2008). Quantitative trait loci mapping in an F2 Duroc x Pietrain resource population: I. Growth traits. *J. Anim. Sci.* 86, 241–253.
16. Edwards, D. B., Tempelman, R. J., and Bates, R. O. (2006). Evaluation of Duroc- vs. Pietrain-sired pigs for growth and composition1. *J. Anim. Sci.* 84, 266–275.
17. Eirin, A., Riester, S. M., Zhu, X.-Y., Tang, H., Evans, J. M., O'Brien, D., et al. (2014). MicroRNA and mRNA cargo of extracellular vesicles from porcine adipose tissue-derived mesenchymal stem cells. *Gene* 551, 55–64.
18. Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 40, 37–52.
19. Gallagher, I. J., Scheele, C., Keller, P., Nielsen, A. R., Remenyi, J., Fischer, C. P., et al. (2010). Integration of microRNA changes in vivo identifies novel molecular features of muscle insulin resistance in type 2 diabetes. *Genome Med.* 2, 9.
20. Giraud-Triboult, K., Rochon-Beaucourt, C., Nissan, X., Champon, B., Aubert, S., and Piétu, G. (2011). Combined mRNA and microRNA profiling reveals that miR-148a and miR-20b control human mesenchymal stem cell phenotype via EPAS1. *Physiol. Genomics* 43, 77–86.
21. Gmeiner, W. H., Reinhold, W. C., and Pommier, Y. (2010). Genome-Wide mRNA and miRNA Profiling of the NCI 60 Cell Line Screen and Comparison of FdUMP[10] with fluorouracil, floxuridine, and Top1 Poisons. 9, 3105–3114.
22. Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A., and Enright, A. J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 34, D140–4.

23. Gualdrón Duarte, J. L., Cantet, R. J. C., Bernal Rubio, Y. L., Bates, R. O., Ernst, C. W., Raney, N. E., et al. (2016). Refining genomewide association for growth and fat deposition traits in an F2 pig population. *J. Anim. Sci.* 94, 1387–1397.
24. Ha, M., and Kim, V. N. (2014). Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* 15, 509–524.
25. Hammond, S. M., Boettcher, S., Caudy, a a, Kobayashi, R., and Hannon, G. J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* 293, 1146–1150.
26. Huang, T. H., Zhu, M. J., Li, X. Y., and Zhao, S. H. (2008). Discovery of porcine microRNAs and profiling from skeletal muscle tissues during development. *PLoS One* 3, e3225.
27. Jansen, R. C., and Nap, J. P. (2001). Genetical genomics: The added value from segregation. *Trends Genet.* 17, 388–391.
28. Johnston, I. a, Lee, H.-T., Macqueen, D. J., Paranthaman, K., Kawashima, C., Anwar, A., et al. (2009). Embryonic temperature affects muscle fibre recruitment in adult zebrafish: genome-wide changes in gene and microRNA expression associated with the transition from hyperplastic to hypertrophic growth phenotypes. *J. Exp. Biol.* 212, 1781–1793.
29. Kadarmideen, H. N., Von Rohr, P., and Janss, L. L. G. (2006). From genetical genomics to systems genetics: Potential applications in quantitative genomics and animal breeding. *Mamm. Genome* 17, 548–564.
30. Kara, M., Yumrutas, O., Ozcan, O., Celik, O. I., Bozgeyik, E., Bozgeyik, I., et al. (2015). Differential expressions of cancer-associated genes and their regulatory miRNAs in colorectal carcinoma. *Gene* 567, 81–86.
31. Ketting, R. F., Fischer, S. E. J., Bernstein, E., Sijen, T., Sijen, T., Hannon, G. J., et al. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing of *C. elegans*. *Genes Dev.*, 2654–2659.
32. Kim, H.-J., Cui, X.-S., Kim, E.-J., Kim, W.-J., and Kim, N.-H. (2006). New porcine microRNA genes found by homology search. *Genome* 49, 1283–1286.
33. Knol, E. F., Nielsen, B., and Knap, P. W. (2016). Genomic selection in commercial pig breeding. *Anim. Front.* 6, 15–22.
34. Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S. (2019). MiRBase: From microRNA sequences to function. *Nucleic Acids Res.* 47, D155–D162.
35. Lawless, N., Foroushani, A. B. K., McCabe, M. S., O’Farrelly, C., and Lynn, D. J. (2013). Next Generation Sequencing Reveals the Expression of a Unique miRNA Profile in

Response to a Gram-Positive Bacterial Infection. *PLoS One* 8, e57543.

36. Lee, R. C. (1993). The *C. elegans* Heterochronic Gene *lin-4* Encodes Small RNAs with Antisense Complementarity to *lin-14*. *Cell* 75, 843–854.
37. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415–419.
38. Lei, B., Gao, S., Luo, L. F., Xia, X. Y., Jiang, S. W., Deng, C. Y., et al. (2010). A SNP in the *miR-27a* gene is associated with litter size in pigs. *Mol. Biol. Rep.* 38, 3725–3729.
39. Li, H.-Y., Xi, Q.-Y., Xiong, Y.-Y., Liu, X.-L., Cheng, X., Shu, G., et al. (2012). Identification and comparison of microRNAs from skeletal muscle and adipose tissues from two porcine breeds. *Anim. Genet.* 43, 704–13.
40. Liaubet, L., Lobjois, V., Faraut, T., Tircazes, A., Benne, F., Iannuccelli, N., et al. (2011). Genetic variability of transcript abundance in pig peri-mortem skeletal muscle: eQTL localized genes involved in stress response, cell death, muscle disorders and metabolism. *BMC Genomics* 12, 548.
41. Liu, H. C., Hicks, J. a., Trakooljul, N., and Zhao, S. H. (2010). Current knowledge of microRNA characterization in agricultural animals. *Anim. Genet.* 41, 225–231.
42. Lukowski, S. W., Fish, R. J., Martin-Levilain, J., Gonelle-Gispert, C., Bühler, L. H., Maechler, P., et al. (2015). Integrated analysis of mRNA and miRNA expression in response to interleukin-6 in hepatocytes. *Genomics*, 1–9.
43. Mármol-Sánchez, E., Cirera, S., Quintanilla, R., Pla, A., and Amills, M. (2020). Discovery and annotation of novel microRNAs in the porcine genome by using a semi-supervised transductive learning approach. *Genomics* 112, 2107–2118.
44. McDaneld, T. G., Smith, T. P., Doumit, M. E., Miles, J. R., Coutinho, L. L., Sonstegard, T. S., et al. (2009). MicroRNA transcriptome profiles during swine skeletal muscle development. *BMC Genomics* 10, 77.
45. Meuwissen, T., Hayes, B., and Goddard, M. (2016). Genomic selection: A paradigm shift in animal breeding. *Anim. Front.* 6, 6–14.
46. Mohr, A. M., and Mott, J. L. (2015). Overview of microRNA biology. *Semin. Liver Dis.* 35, 3–11.
47. Mollaei, H., Safaralizadeh, R., and Rostami, Z. (2019). MicroRNA replacement therapy in cancer. *J. Cell. Physiol.* 234, 12369–12384.
48. Ponsuksili, S., Du, Y., Murani, E., Schwerin, M., and Wimmers, K. (2012). Elucidating molecular networks that either affect or respond to plasma cortisol concentration in target

tissues of liver and muscle. *Genetics* 192, 1109–1122.

49. Ponsuksili, S., Jonas, E., Murani, E., Phatsara, C., Srikanchai, T., Walz, C., et al. (2008). Trait correlated expression combined with expression QTL analysis reveals biological pathways and candidate genes affecting water holding capacity of muscle. *BMC Genomics* 9, 367.
50. Ponsuksili, S., Murani, E., Schwerin, M., Schellander, K., and Wimmers, K. (2010). Identification of expression QTL (eQTL) of genes expressed in porcine M. longissimus dorsi and associated with meat quality traits. *BMC Genomics* 11, 572.
51. Ponsuksili, S., Murani, E., Trakooljul, N., Schwerin, M., and Wimmers, K. (2014). Discovery of candidate genes for muscle traits based on GWAS supported by eQTL-analysis. *Int. J. Biol. Sci.* 10, 327–37.
52. Ponsuksili, S., Oster, M., Reyer, H., Hadlich, F., Trakooljul, N., Rodehutsord, M., et al. (2021). Genetic regulation and heritability of miRNA and mRNA expression link to phosphorus utilization and gut microbiome. *Open Biol.* 11.
53. Ponsuksili, S., Trakooljul, N., Hadlich, F., Haack, F., Murani, E., and Wimmers, K. (2017). Genetic architecture and regulatory impact on hepatic microRNA expression linked to immune and metabolic traits. *Open Biol.* 7.
54. Quitadamo, A., Tian, L., Hall, B., and Shi, X. (2015). An integrated network of microRNA and gene expression in ovarian cancer. *BMC Bioinformatics* 16, S5.
55. Saba, L. M., Bennett, B., Hoffman, P. L., Barcomb, K., Ishii, T., Kechris, K., et al. (2010). A systems genetic analysis of alcohol drinking by mice, rats and men: Influence of brain GABAergic transmission. *Neuropharmacology* 60, 1269–1280.
56. Schadt, E. E., Monks, S. a, Drake, T. a, Lusis, A. J., Che, N., Colinayo, V., et al. (2003). Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422, 297–302.
57. Schmidt, M. F. (2014). Drug target miRNAs: chances and challenges. *Trends Biotechnol.* 32, 578–585.
58. Schwab, C. R., and Baas, T. J. (2008). Direct and Correlated Responses to Selection for Intramuscular Fat in Duroc Swine Fat in Duroc Swine. 654.
59. Soifer, H. S., Sano, M., Sakurai, K., Chomchan, P., Sætrom, P., Sherman, M. a., et al. (2008). A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. *Nucleic Acids Res.* 36, 6511–6522.
60. Sontheimer, E. J., and Carthew, R. W. (2005). Silence from within: Endogenous siRNAs and miRNAs. *Cell* 122, 9–12.

61. Steibel, J. P., Bates, R. O., Rosa, G. J. M., Tempelman, R. J., Rilmington, V. D., Ragavendran, A., et al. (2011). Genome-Wide Linkage Analysis of Global Gene Expression in Loin Muscle Tissue Identifies Candidate Genes in Pigs. *PLoS One* 6, e16766.
62. Szeto, C. Y. Y., Lin, C. H., Choi, S. C., Yip, T. T. C., Ngan, R. K. C., Tsao, G. S. W., et al. (2014). Integrated mRNA and microRNA transcriptome sequencing characterizes sequence variants and mRNA-microRNA regulatory network in nasopharyngeal carcinoma model systems. *FEBS Open Bio* 4, 128–140.
63. United States Department of Agriculture (2021). United States Department of Agriculture National Agricultural Statistics Service Meat Animals Production, Disposition, and Income.
64. Velez-Irizarry, D., Casiro, S., Daza, K. R., Bates, R. O., Raney, N. E., Steibel, J. P., et al. (2019). Genetic control of longissimus dorsi muscle gene expression variation and joint analysis with phenotypic quantitative trait loci in pigs. *BMC Genomics* 20, 1–19.
65. Wernersson, R., Schierup, M. H., Jørgensen, F. G., Gorodkin, J., Panitz, F., Staerfeldt, H.-H., et al. (2005). Pigs in sequence space: a 0.66X coverage pig genome survey based on shotgun sequencing. *BMC Genomics* 6, 70.
66. Xie, S.-S., Li, X.-Y., Liu, T., Cao, J.-H., Zhong, Q., and Zhao, S.-H. (2011). Discovery of Porcine microRNAs in Multiple Tissues by a Solexa Deep Sequencing Approach. *PLoS One* 6, e16235.
67. Yang, Y., Li, Y., Liang, R., Zhou, R., Ao, H., Mu, Y., et al. (2014). Dynamic Expression of MicroRNA-127 During Porcine Prenatal and Postnatal Skeletal Muscle Development. *J. Integr. Agric.* 13, 1331–1339.
68. Yi, R., Qin, Y., Macara, I. G., and Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. 3011–3016.

CHAPTER TWO

PROFILING AND CHARACTERIZATION OF A *LONGISSIMUS DORSI* MUSCLE MICRORNA DATASET FROM AN F₂ DUROC X PIETRAIN PIG RESOURCE POPULATION

This chapter has been published previously (Daza et al. 2017). The manuscript was prepared alongside co-authors Juan P. Steibel, Deborah Velez-Irizarry, Nancy E. Raney, Ronald O. Bates, and Catherine W. Ernst

2.1 Abstract

To elucidate the effects of microRNA (miRNA) regulation in skeletal muscle of adult pigs, miRNA expression profiling was performed with RNA extracted from longissimus dorsi (LD) muscle samples from 174 F₂ pigs (~5.5 months of age) from a Duroc × Pietrain resource population. Total RNA was extracted from LD samples, and libraries were sequenced on an Illumina HiSeq 2500 platform in 1 × 50 bp format. After processing, 232,826,977 total reads were aligned to the *Sus scrofa* reference genome (v10.2.79), with 74.8% of total reads mapping successfully. The miRDeep2 software package (v0.0.5) was utilized to quantify annotated *Sus scrofa* mature miRNAs from miRBase (Release 21) and to predict candidate novel miRNA precursors. Among the retained 295 normalized mature miRNA expression profiles sscmiR1, sscmiR133a3p, sscmiR378, sscmiR206, and sscmiR10b were the most abundant, all of which have previously been shown to be expressed in pig skeletal muscle. Additionally, 27 unique candidate novel miRNA precursors were identified exhibiting homologous sequence to annotated human miRNAs. The composition of classes of small RNA present in this dataset was also characterized; while the majority of unique expressed sequence tags were not annotated in any of the queried databases, the most abundantly expressed class of small RNA in this dataset was miRNAs. This data provides a resource to evaluate miRNA regulation of gene expression and

effects on complex trait phenotypes in adult pig skeletal muscle. The raw sequencing data were deposited in the Sequence Read Archive, BioProject PRJNA363073.

2.2 Direct link to deposited data

The deposited sequencing data can be found here:

<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA363073>

2.3 Experimental design, materials and methods

2.3.1 Sample collection, RNA isolation and sequencing

A subset of 176 F₂ pigs from 44 litters (two males and two females per litter) of the Michigan State University Duroc × Pietrain Pig Resource Population (MSUPRP) were selected for transcriptional profiling (Edwards et al. 2008a, 2008b; Steibel et al. 2011). Samples of *longissimus dorsi* (LD) tissue were collected from each animal at slaughter and frozen at – 80 °C. Animal care and experimental protocols were approved by the Michigan State University All University Committee on Animal Use and Care (AUF# 09/03-114-00). Total RNA was isolated from LD samples using the miRNeasy Mini Kit (QIAGEN, California, USA), and small RNA library construction and sequencing was performed at the MSU Research Technology Support Facility. Samples were prepared for sequencing utilizing the Bioo Scientific NEXTFlex™ Small RNA Sequencing Kit (v2; Bioo Scientific, Austin, TX, USA). Libraries were barcoded and multiplexed for sequencing on the Illumina HiSeq 2500 platform (Illumina, Inc.; California, USA) in 50 bp, single-end format. Two libraries failed to produce acceptable sequencing output and were removed from further analysis, yielding 174 files of 50 nt short read sequences in fastq format (86 males, 88 females).

2.3.2 Analysis of sequencing data

The 3' adaptor sequences were trimmed from raw short reads using cutadapt (cutadapt/1.4.1; Martin 2011), and trimmed reads of length 26–38 nt were filtered for quality using FASTX toolkit (FASTX/0.0.14; http://hannonlab.cshl.edu/fastx_toolkit), removing any reads in which > 50% of the nucleotides had Phred score < 30. The selected sequence length retained an additional 4 nt on both the 5' and 3' ends of each read to accommodate the randomized adaptors ligated during library preparation, which allowed the identification of PCR duplicates. Distribution of sequence read lengths was assessed, and trimmed, filtered reads were collapsed using FASTX toolkit (FASTX/0.0.14; http://hannonlab.cshl.edu/fastx_toolkit) into unique expressed sequence tags containing a sequence identifier and the numerical count of each read (read count). PCR duplicates and random adaptor sequences were removed using the ShortRead package of R prior to genome alignment and quantification (Morgan et al. 2009).

2.3.3 Characterization of small RNAs

To characterize classes of small RNA present in the dataset, unique expressed sequence tags underwent multiple queries utilizing BLAST+ (v2.2.30) prior to read alignment and quantification of known miRNAs. Unique sequences expressed at least two times were included in this analysis. Sequences were sequentially queried against: the *Sus scrofa* mature miRNA miRBase database (Release 21; Griffiths-Jones et al. 2008), the Ensembl *Sus scrofa* ncRNA database (Release 84), and the *Sus scrofa*, *Homo sapiens*, and *Mus musculus* Rfam databases (version 11) using blastn. The blastn-short parameter was implemented as it is optimal for queries of sequences < 50 nt, and an e-value threshold of 1×10^{-5} was used to declare a significant hit. Results of each query were filtered as follows to retain unique sequence read hits:

hits were retained that obtained 100% sequence identity, followed by at least 96% sequence identity, and finally hits with the maximum bitscore were retained. If multiple BLAST + hits remained for a given sequence, the hits most identical to the sequence were retained (based on percent identity). If multiple significant hits persisted with equal sequence identity, a single hit was retained for each sequence. This yielded the composition of classes of small RNA present in the small RNA sequencing dataset, based on both the unique expressed sequence tags and total read counts.

The miRDeep2 software package (v0.0.5) was used to align high-quality reads to the *Sus scrofa* reference genome (v10.2.79; Ensembl), and to quantify *Sus scrofa* mature microRNAs (miRNAs) obtained from miRBase (Release 21; Griffiths-Jones et al. 2008; Friedländer et al. 2012). The average abundance of each mature pig miRNA was adjusted for differences in sequencing depth between libraries by converting the read counts to counts per million (cpm) with the edgeR package of R, incorporating TMM normalization factors (Robinson et al. 2010). miRNAs expressed at < 1 cpm in ≥ 44 libraries were removed from the dataset prior to calculation of the normalization factors.

miRDeep2 (Friedländer et al. 2012) was also utilized to predict candidate novel miRNAs. The software was provided the known *Sus scrofa* mature and precursor miRNA sequences, and the known *Homo sapiens* mature miRNA sequences from miRBase (Release 21; Griffiths-Jones et al. 2008) to search for sequence homology for novel miRNA prediction. The resulting candidate novel precursors were filtered based on: the miRDeep2 score output by the algorithm (≥ 10), the estimated probability of the novel miRNA being a true positive ($\geq 91 \pm 2\%$), a significant Randfold *p*-value, and a minimum read count of 10 reads for both mature and star

(complementary) sequences. Retained sequences were converted to DNA alphabet and FASTA format (ShortRead (R); Morgan et al. 2009). BLAST + (v2.2.30) was utilized to further characterize the candidate novel miRNA precursors by searching for sequence homology to the database of known human precursor and mature miRNAs in miRBase (Release 21; Griffiths-Jones et al. 2008), utilizing a stringent e-value threshold of 1×10^{-5} resulting in high-confidence homologous sequences.

All code used to analyze this dataset is publicly available and can be found at <https://github.com/perrykai>.

2.4 Data description

Determining underlying mechanisms controlling complex trait phenotypes such as growth, carcass composition and meat quality in pigs is important for achieving continued genetic improvement. One genetic mechanism involved in regulating these traits is the silencing of gene expression via miRNAs. Previous studies report that miRNAs exhibit dynamic expression in pig skeletal muscle across developmental stages, physiological states, and breeds (Huang et al. 2008; McDanel et al. 2009; Nielsen et al. 2009; Xie et al. 2011; Qin et al. 2013; Siengdee et al. 2013; Cai et al. 2014). Moreover, the signaling pathways enriched for genes that these miRNAs target are shown to be involved in myogenesis and muscle regeneration (Nielsen et al. 2009). Due to the influential role of miRNAs on economically-important complex traits, it is necessary to continue characterizing miRNAs in pig skeletal muscle. Thus, the objective of this work was to profile and characterize the expression of miRNAs in LD muscle of market-age (approximately 5.5 months of age) F₂ pigs from the MSUPRP (Edwards et al. 2008a, 2008b).

Assessment of sequence read length distribution revealed 63.8% of the total reads to be 22 nt long, which is the common length of mature miRNAs. This indicates the success of small RNA sequencing at isolating miRNAs (Fig. 2.1a), and concurs with results in Xie et al. (2011) where 43.7% of total small RNA sequencing reads were 22 nt in length. In total, 232,826,977 high-quality reads (mean 1,338,086 reads per library) were aligned to the *Sus scrofa* reference genome (v10.2.79; Ensembl), with 74.8% of reads successfully mapping. Moreover, 158,672,792 reads were quantified as miRNAs from the 174 samples, corresponding to 91.2% of the mapped reads.

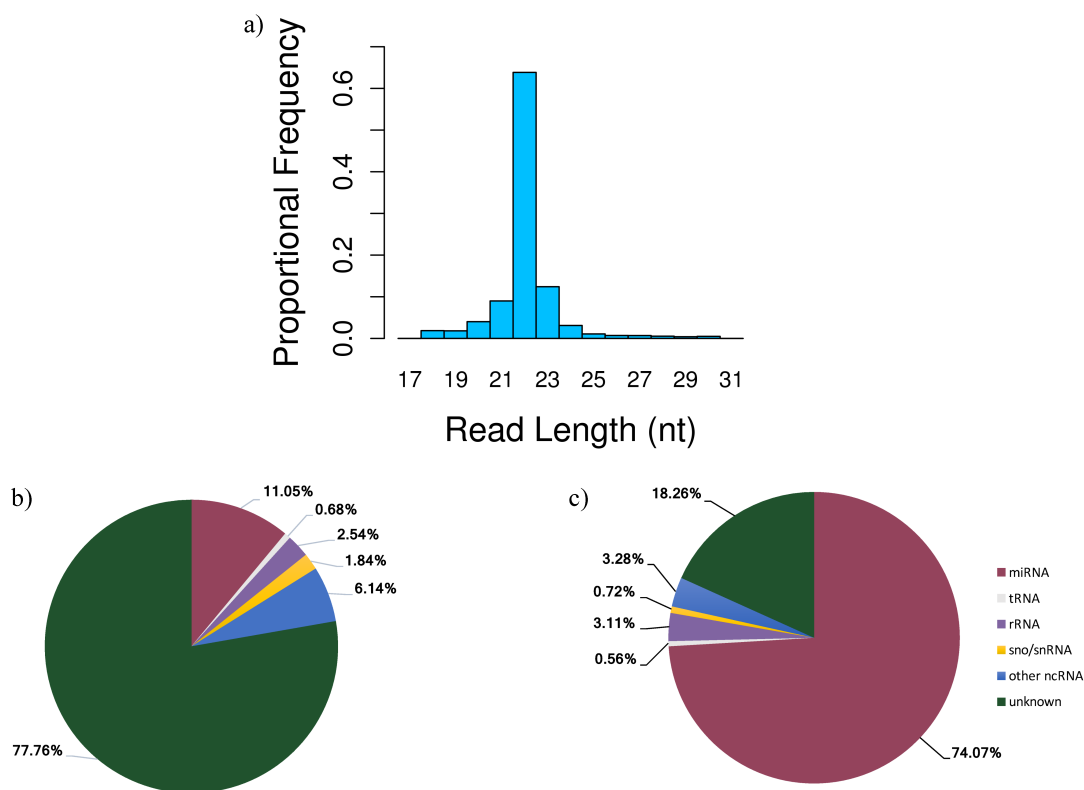


Figure 2.1 Characterization of small RNA sequencing data. a) Sequence read length distribution; composition of small RNA classes present in the small RNA sequencing dataset based on b) unique sequences, and c) total read counts.

After filtering for abundance across libraries, 295 miRNA expression profiles were obtained (Table 2.S1). The five most abundant miRNAs represented 47.9% of the total cpm in the dataset including ssc-miR-1, ssc-miR-133a3p, sscmiR-378, ssc-miR-206, and ssc-miR-10b. These miRNAs have previously been identified in pig skeletal muscle, including Nielsen et al. (2009), where ssc-miR-1 and ssc-miR-206 were the two most highly abundant miRNAs identified from sequencing LD samples from seven 1.5–2-year-old Danish Landrace/Yorkshire crossbred pigs. MiR-1, miR-206, and miR-133a are also considered the “myomiRs”, and have been well-characterized for their roles in mammalian skeletal muscle myogenesis and regeneration (for review, see Horak et al. 2016).

Results of the small RNA characterization analysis are shown in Fig. 2.1b, considering unique expressed sequence tags, and Fig. 2.1c, considering the total read count of each sequence. Most of the unique expressed sequence tags were not annotated in any of the five databases queried (“unknown”, 77.8%; Fig. 2.1b). Additionally, 11.1% of the unique expressed sequence tags (where each unique sequence was counted one time) were classified as miRNAs, and 6.1% of the unique expressed sequence tags consisted of other non-coding RNAs (Fig. 2.1b). When considering the sequences based on total read count, 74.1% of the sequences were identified as miRNAs, while 18.3% of the total reads were not annotated (Fig. 2.1c). These results show that miRNAs are the most abundant small RNA expressed in this dataset, and are consistent with results obtained in Xie et al. (2011), where miRNAs were the most abundant class of small RNAs found in a pooled RNA sample from 16 pig tissues.

There were 132 candidate novel miRNAs predicted from miRDeep2 that passed filtering steps. After BLAST + query, there were 27 unique candidate novel miRNA precursors

homologous to a human miRNA precursor (Table 2.1; additional annotation information in Table 2.S2). Eight candidate novel precursors had more than one homologous human miRNA precursor, ranging from 2 to 6 hits per candidate novel sequence. The consensus secondary structures for the 27 unique candidate novel precursors predicted by miRDeep2 can be found in Fig. 2.S1.

In summary, 295 known *Sus scrofa* mature miRNA expression profiles were obtained from next-generation sequencing of LD RNA from 174 F₂ Duroc × Pietrain pigs of the MSUPRP. Twenty-seven unique candidate novel pig miRNA precursors were predicted, and the composition of classes of small RNA present in the dataset were characterized to reveal that miRNAs were the most abundant small RNA class expressed in the skeletal muscle of these pigs. This work contributes to the further characterization of microRNAs in pig skeletal muscle, which will lead to a more complete picture of the genetic regulation of complex economically-important pig production traits.

2.5 Acknowledgements

The authors wish to thank Scott Funkhouser for his R programming support. Computer services were provided by the Michigan State University High Performance Computing Center.

This research was supported by Agriculture and Food Research Initiative Competitive Grant no. 2014-67015-21619 from the USDA National Institute for Food and Agriculture and by National Needs Fellowships from the National Institute of Food and Agriculture, Award No. 2012-38420-30199.

Table 2.1. Candidate novel miRNA precursors predicted from miRDeep2 and characterized with BLAST+.

Sequence ID ^a	miRDeep2 Score	Read Count	Predicted Mature Sequence ^a	Seq Length	miRDeep2 predicted miRbase miRNA ^a	BLAST+ Matched precursor ^b	% Identical ^b	E value ^b	Precursor Coordinate ^a
GL892871.2_43622	572746.3	1123410	uucaaguaauucaggauagguu	22	hsa-miR-26a-5p	hsa-mir-26b	100.0	2.00E-23	GL892871.2:56019..56076:+
X_39130	136780.7	268282	uaccgauugcauauccggaguug	22	hsa-miR-660-5p	hsa-mir-660	97.7	2.00E-17	X:48640826..48640884:+
3_7507	57547.6	112871	uaauuuuauguauaagcuagu	21	hsa-miR-590-3p	hsa-mir-590	96.0	1.00E-06	3:11024999..11025060:+
13_28851	40621.2	79667	uucaaguaaccaggauaggcu	22	hsa-miR-26a-5p	hsa-mir-26a-1	96.1	7.00E-20	13:24885255..24885316:-
13_28851	40621.2	79667	uucaaguaaccaggauaggcu	22	hsa-miR-26a-5p	hsa-mir-26a-2	95.8	4.00E-06	13:24885255..24885316:-
GL894231.1_44092	12387.7	24290	auaaauacaugguuaccucuuu	22	hsa-miR-655-3p	hsa-mir-655	100.0	7.00E-17	GL894231.1:23071..23131:-
GL894231.1_44082	6189.5	12132	gucauacacggcucuccucucu	22	hsa-miR-485-3p	hsa-mir-485	100.0	3.00E-25	GL894231.1:16616..16675:-
GL894231.1_44074	6021.2	11802	aucacacaaaggcaacuuuugu	22	hsa-miR-377-3p	hsa-mir-377	98.2	5.00E-24	GL894231.1:10061..10121:-
12_25369	4096.6	8028	aucaugaugauacugcaaaaca	22	hsa-miR-6516-3p	hsa-mir-6516	98.0	5.00E-21	12:4273970..4274033:+
GL894231.1_44108	3936.2	7714	aucauagaggaaaauccacau	21	hsa-miR-376a-3p	hsa-mir-376a-2	95.8	4.00E-18	GL894231.1:31685..31743:-
GL894231.1_44108	3936.2	7714	aucauagaggaaaauccacau	21	hsa-miR-376a-3p	hsa-mir-376a-1	85.2	2.00E-07	GL894231.1:31685..31743:-
1_1168	3418.9	6698	uuuagugugauaauaggcguuug	22	hsa-miR-3591-5p	hsa-mir-3591	98.0	3.00E-22	1:179916350..179916408:+
1_1168	3418.9	6698	uuuagugugauaauaggcguuug	22	hsa-miR-3591-5p	hsa-mir-122	97.9	2.00E-20	1:179916350..179916408:+
X_39122	3415.6	6691	caucccuugcaugggagggu	22	hsa-miR-188-5p	hsa-mir-188	100.0	1.00E-21	X:48632394..48632453:+
X_39256	1851.7	3631	uuacaauacaaccugauaagugc	23	-	hsa-mir-374a	96.1	6.00E-20	X:67530381..67530433:+
X_39256	1851.7	3631	uuacaauacaaccugauaagugc	23	-	hsa-mir-374a	90.5	8.00E-10	X:67530381..67530433:+
X_39256	1851.7	3631	uuacaauacaaccugauaagugc	23	-	hsa-mir-374c	88.9	3.00E-09	X:67530381..67530433:+
X_39256	1851.7	3631	uuacaauacaaccugauaagugc	23	-	hsa-mir-374c	86.4	3.00E-06	X:67530381..67530433:+
X_39256	1851.7	3631	uuacaauacaaccugauaagugc	23	-	hsa-mir-374b	88.9	3.00E-09	X:67530381..67530433:+
X_39256	1851.7	3631	uuacaauacaaccugauaagugc	23	-	hsa-mir-374b	86.4	3.00E-06	X:67530381..67530433:+
7_18912	1362.7	2664	auaagacgagcaaaaagcuugu	22	hsa-miR-208a-3p	hsa-mir-208a	100.0	3.00E-25	7:81043856..81043913:-
13_28053	986.0	1925	gcgaccuauacuugguuucaga	22	hsa-miR-551a	hsa-mir-551b	94.9	1.00E-12	13:115915962..115916023:+

Table 2.1 (cont'd)

6_14565	642.0	1250	uaauacugccugguaaugauga	22	hsa-miR-200b-3p	hsa-mir-200b	95.0	2.00E-13	6:58064309.. 58064367:+ JH118928.1:123819 ..123880:- GL894231.1:24222.. 24280:- 17:12191239.. 12191302:+ 17:12191239.. 12191302:+ 17:12191239.. 12191302:+ 1:313734674.. 313734733:- 14:133619338.. 133619395:+ 14:133619338.. 133619395:+ X:51475450.. 51475523:- 1:140919000.. 140919059:+ X:69515840.. 69515903:+ 12:1507090.. 1507153:- 12:1507090.. 1507153:- GL894231.1:20450.. 20508:- 1:302806862.. 302806927:+ 1:302806862.. 302806927:+ 1:302806862.. 302806927:+ 2:100272647.. 100272707:+
JH118928.1_41756	527.0	1036	ucauucuccuucuuuugaccaga	22	-	hsa-mir-6758	94.4	6.00E-11	
GL894231.1_44094	365.0	715	ucuuguuaaaaggcagauucu	21	-	hsa-mir-544a	97.6	6.00E-17	
17_36762	298.5	583	ccggguacugagcuggcccag	22	-	hsa-mir-486-1	91.2	7.00E-14	
17_36762	298.5	583	ccggguacugagcuggcccag	22	-	hsa-mir-486-2	91.1	3.00E-13	
17_36762	298.5	583	ccggguacugagcuggcccag	22	-	hsa-mir-486-2	100.0	3.00E-07	
1_3654	244.3	470	cauuauuacucacggucagagu	22	hsa-miR-126-5p	hsa-mir-126	100.0	7.00E-23	
14_31131	225.4	433	ccaaaccaguugugccuguag	21	hsa-miR-6715a-3p	hsa-mir-6715b	95.4	4.00E-15	
14_31131	225.4	433	ccaaaccaguugugccuguag	21	hsa-miR-6715a-3p	hsa-mir-6715a	93.5	2.00E-14	
X_40048	137.0	266	uucccugcccucuccuccagg	22	-	hsa-mir-6894	89.6	8.00E-11	
1_851	130.1	248	uggaaacacuucugcacaacu	22	hsa-miR-4261	hsa-mir-147b	95.9	1.00E-18	
X_39264	117.7	224	uguaaacaauuccuagguaaugu	23	hsa-miR-30a-5p	hsa-mir-384	88.4	6.00E-08	
12_26433	111.2	209	ucaacaaaucacugaucugga	23	hsa-miR-3065-5p	hsa-mir-338	94.7	5.00E-21	
12_26433	111.2	209	ucaacaaaucacugaucugga	23	hsa-miR-3065-5p	hsa-mir-3065	94.1	2.00E-17	
GL894231.1_44088	80.7	150	aaucacacaggacauccaguu	22	hsa-miR-154-3p	hsa-mir-487a	90.0	1.00E-08	
1_1860	61.0	111	agauguccagccacaauucucg	22	hsa-miR-219b-5p	hsa-mir-219b	100.0	3.00E-22	
1_1860	61.0	111	agauguccagccacaauucucg	22	hsa-miR-219b-5p	hsa-mir-219a-2	100.0	2.00E-20	
1_1860	61.0	111	agauguccagccacaauucucg	22	hsa-miR-219b-5p	hsa-mir-219a-1	100.0	3.00E-07	
2_5120	24.3	39	acugacaggagagcauuuuau	22	hsa-miR-3660	hsa-mir-3660	94.8	1.00E-12	

^a Results obtained from miRDeep2 algorithm

^b Results obtained from BLAST+ query against *Homo sapiens* miRBase database.

APPENDIX

APPENDIX

Supplementary materials are available at: <http://dx.doi.org/10.1016/j.gdata.2017.07.006>

Table 2.S1. *Sus scrofa* mature miRNA average abundance (cpm) in the 174 F₂ Duroc x Pietrain Michigan State University resource population pigs.

Table 2.S2. Candidate novel miRNA annotation information.

Figure 2.S1. Secondary structures of candidate novel miRNAs. Sequences are color-coded as follows: red, mature sequence; yellow, loop sequence; purple, star strand sequence observed in small RNA sequencing data; light blue, expected star sequence based on Dicer/Drosha processing (Friedländer et al. 2012).

REFERENCES

REFERENCES

1. Cai, Z., Zhang, L., Jiang, X., Sheng, Y., and Xu, N. (2014). Differential miRNA expression profiles in the longissimus dorsi muscle between intact and castrated male pigs. *Res. Vet. Sci.* 99, 99–104.
2. Edwards, D. B., Ernst, C. W., Tempelman, R. J., Rosa, G. J. M., Raney, N. E., Hoge, M. D., et al. (2008a). Quantitative trait loci mapping in an F2 Duroc x Pietrain resource population: I. Growth traits. *J. Anim. Sci.* 86, 241–253.
3. Edwards, D. B., Ernst, C. W., Raney, N. E., Doumit, M. E., Hoge, M. D., and Bates, R. O. (2008b). Quantitative trait locus mapping in an F2 Duroc × Pietrain resource population: II. Carcass and meat quality traits1. *J. Anim. Sci.* 86, 254–266.
4. Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 40, 37–52.
5. Griffiths-Jones, S., Saini, H. K., van Dongen, S., and Enright, A. J. (2008). miRBase: tools for microRNA genomics. *Nucleic Acids Res.* 36, D154–D158.
6. Horak, M., Novak, J., and Bienertova-Vasku, J. (2016). Muscle-specific microRNAs in skeletal muscle development. *Dev. Biol.* 410, 1–13.
7. Huang, T.-H., Zhu, M.-J., Li, X.-Y., and Zhao, S.-H. (2008). Discovery of Porcine microRNAs and Profiling from Skeletal Muscle Tissues during Development. *PLoS One* 3, e3225.
8. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10.
9. McDaneld, T. G., Smith, T. P., Doumit, M. E., Miles, J. R., Coutinho, L. L., Sonstegard, T. S., et al. (2009). MicroRNA transcriptome profiles during swine skeletal muscle development. *BMC Genomics* 10, 77.
10. Morgan, M., Anders, S., Lawrence, M., Aboyoun, P., Pages, H., and Gentleman, R. (2009). ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* 25, 2607–2608.
11. Nielsen, M., Hansen, J. H., Hedegaard, J., Nielsen, R. O., Panitz, F., Bendixen, C., et al. (2009). MicroRNA identity and abundance in porcine skeletal muscles determined by deep sequencing. *Anim. Genet.* 41, 159–168.

12. Qin, L., Chen, Y., Liu, X., Ye, S., Yu, K., Huang, Z., et al. (2013). Integrative Analysis of Porcine microRNAome during Skeletal Muscle Development. *PLoS One* 8, e72418.
13. Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, R25.
14. Siengdee, P., Trakooljul, N., Murani, E., Schwerin, M., Wimmers, K., and Ponsuksili, S. (2013). Transcriptional profiling and miRNA-dependent regulatory network analysis of longissimus dorsi muscle during prenatal and adult stages in two distinct pig breeds. *Anim. Genet.* 44, 398–407.
15. Steibel, J. P., Bates, R. O., Rosa, G. J. M., Tempelman, R. J., Rilmington, V. D., Ragavendran, A., et al. (2011). Genome-Wide Linkage Analysis of Global Gene Expression in Loin Muscle Tissue Identifies Candidate Genes in Pigs. *PLoS One* 6, e16766.
16. Xie, S.-S., Li, X.-Y., Liu, T., Cao, J.-H., Zhong, Q., and Zhao, S.-H. (2011). Discovery of Porcine microRNAs in Multiple Tissues by a Solexa Deep Sequencing Approach. *PLoS One* 6, e16235.

CHAPTER THREE

INTEGRATED GENOME-WIDE ANALYSIS OF MICRORNA EXPRESSION QUANTITATIVE TRAIT LOCI IN PIG LONGISSIMUS DORSI MUSCLE

This chapter has been published previously (Daza et al. 2021). The manuscript was prepared alongside co-authors Deborah Velez-Irizarry, Sebastian Casiró, Juan P. Steibel, Nancy E. Raney,

Ronald O. Bates, and Catherine W. Ernst

3.1 Abstract

Determining mechanisms regulating complex traits in pigs is essential to improve the production efficiency of this globally important protein source. MicroRNAs (miRNAs) are a class of non-coding RNAs known to post-transcriptionally regulate gene expression affecting numerous phenotypes, including those important to the pig industry. To facilitate a more comprehensive understanding of the regulatory mechanisms controlling growth, carcass composition, and meat quality phenotypes in pigs, we integrated miRNA and gene expression data from *longissimus dorsi* muscle samples with genotypic and phenotypic data from the same animals. We identified 23 miRNA expression Quantitative Trait Loci (miR-eQTL) at the genome-wide level and examined their potential effects on these important production phenotypes through miRNA target prediction, correlation, and colocalization analyses. One miR-eQTL miRNA, miR-874, has target genes that colocalize with phenotypic QTL for 12 production traits across the genome including backfat thickness, dressing percentage, muscle pH at 24 h post-mortem, and cook yield. The results of our study reveal genomic regions underlying variation in miRNA expression and identify miRNAs and genes for future validation of their regulatory effects on traits of economic importance to the global pig industry.

3.2 Introduction

Pork is the meat of choice worldwide, and global meat consumption is projected to continue to increase over the next decade. To fulfill consumer expectations, pork producers will need to continue to improve the quality and consistency of pork products in order to enhance the prosperity of this global industry (USDA, 2020). However, industry focus on increasing carcass leanness has historically neglected meat quality characteristics, increasing the incidence of inferior eating-quality pork and decreasing consumer satisfaction (Barbut et al. 2008; Schwab and Baas 2008; Ciobanu et al. 2011). Improvements in animal production have been gained through the incorporation of genomic selection techniques (Meuwissen et al. 2016). However, many economically relevant production traits are controlled by complex interactions of gene and protein expression and regulation. While genomic selection has facilitated increased progress on those traits, a better understanding of the genetic architecture controlling phenotypic expression cannot only improve selection in the long term but also enhance management interventions. Improving understanding of the complex regulation of important pork production traits including meat quality, carcass composition, and growth will continue to be a priority for scientists and livestock producers alike.

One genetic mechanism that has been investigated for its role in regulating these economically important traits is the silencing of gene expression *via* microRNAs [miRNA(s) or miR], a class of single-stranded, non-coding small RNA that post-transcriptionally regulate gene expression through sequence complementarity of an approximately 7 nt “seed” sequence with target mRNA sequences. They are known to target genes throughout the genome, influencing a multitude of biological processes. Furthermore, a single miRNA can potentially target hundreds of genes, and multiple miRNAs have the ability to target the same mRNA, acting as “fine tuners”

of gene expression and adding rich complexity to the regulation of polygenic traits. Several miRNAs expressed specifically in muscle tissue have been termed the “myomiRs” (miR-1, miR-133a, miR-133b, miR-206, and others), which play critical roles in the proliferation, differentiation, and regeneration of skeletal muscle (for review, see Horak et al. 2016). Previous studies demonstrate miRNA involvement in skeletal muscle development and function in pigs, spanning various developmental time points, physiological states, and breeds (Qin et al. 2013; Siengdee et al. 2013; Cai et al. 2015; Jing et al. 2015; Mai et al. 2016; Xi et al. 2018; Mármol-Sánchez et al. 2020). While many studies have examined effects of miRNAs on diverse phenotypes in pigs, less has been done to elucidate mechanisms regulating the expression of miRNAs themselves, and their effects on downstream phenotypes.

By combining high-density genotypic data with miRNA expression profiles, associations between regions of the genome harboring single nucleotide polymorphisms (SNPs) and the expression of a miRNA can be revealed. These associated genomic regions are termed “miRNA expression Quantitative Trait Loci” (miR-eQTL) and add another layer to our understanding of the regulation of complex phenotypes on the molecular level. Studies identifying miR-eQTL have been reported, but their objectives primarily focus on human biomedical applications, particularly using miRNAs as biomarkers or therapeutics in disease phenotypes (Dong et al. 2010; Gamazon et al. 2013; Huan et al. 2015; Wohlers et al. 2018). These studies reinforce the importance of elucidating the genetic architecture of miRNA expression and its effects on traits of interest to the pig industry.

Understanding both the influence of miRNA regulation on pig production phenotypes and factors affecting the expression of miRNAs themselves will increase our understanding of skeletal muscle tissue at the molecular level, ultimately leading to improvements in pork quality

and consistency. Therefore, the purpose of this study was to investigate the regulatory role of miRNAs in growth, carcass composition, and meat quality traits through genome-wide miR-eQTL analyses utilizing small RNA sequencing data from *longissimus dorsi* (LD) muscle samples from an F₂ Duroc × Pietrain resource population and to integrate gene expression and phenotypic data from the same animals to elucidate regulatory mechanisms controlling these complex traits.

3.3 Materials and methods

3.3.1 Data collection and sequencing

Animal housing and care protocols were evaluated and approved by the Michigan State University All University Committee on Animal Use and Care (AUF # 09/03-114-00). The 176 animals used in this study were a subset of F₂ pigs selected from the Duroc × Pietrain Michigan State University Pig Resource Population expressing extremes for loin muscle area or back fat thickness phenotypes (MSUPRP; Edwards et al. 2008a, 2008b; Steibel et al. 2011). These animals were phenotyped for over 60 traits encompassing meat quality, carcass composition, and growth traits. Genotyping was performed by Neogen Corporation—GeneSeek Operations (Lincoln NE) using Illumina PorcineSNP60 BeadChips (Ramos et al. 2009) and was previously reported by our group (Casiró et al. 2017; Velez-Irizarry et al. 2019). Resulting genotypes were filtered for markers with extreme allele frequencies calculated from the entire F₂ population of 940 animals ($MAF < 0.10$), removal of non-informative markers, and markers located on sex chromosomes. A total of 36,292 SNPs were included in subsequent analyses. Genotype data for the animals used in this analysis are publicly available and can be found at <https://www.github.com/steibelj/gwaR>.

Samples of LD tissue were collected from each animal at slaughter, frozen in liquid nitrogen, and stored at -80°C for later use. Total RNA was isolated from LD samples using the miRNeasy Mini Kit (QIAGEN, California, United States). Samples were prepared for sequencing utilizing the Bioo Scientific NEXTFlexTM Small RNA Sequencing Kit (v2; Bioo Scientific, Austin, TX, United States) and were sequenced on the Illumina HiSeq 2500 platform (Illumina, Inc.; California, United States) in 50 bp, single-end format at the Michigan State University Research Technology Support Facility (East Lansing, Michigan, United States). Sample libraries were sequenced on two flow cells (38 libraries/lane on four lanes of one flow cell; 27 libraries/lane (including 24 new libraries and three repeated libraries) on one lane of a second flow cell), and sequencing of the same library pools was repeated on two additional flow cells to increase sequencing depth. Raw counts of sequence output was merged and compiled into a single file for each sample library prior to analysis. Two libraries failed to produce acceptable sequencing output and were removed from further analysis, yielding 174 files of 50 nt short-read sequences in fastq format (86 male, 88 female). The small RNA sequencing dataset generated for this study can be found in the NCBI Sequence Read Archive under BioProject PRJNA363073.

3.3.2 Bioinformatics analysis

After adaptor trimming, read quality filtering, and removal of PCR duplicate reads (Daza et al. 2017), the miRDeep2 software package (v0.0.5; Friedländer et al. 2012) was used with default parameters to align high-quality reads to the *Sus scrofa* reference genome (v11.1; Ensembl), and to quantify *Sus scrofa* mature miRNAs obtained from miRBase¹ (Release 21; Griffiths-Jones et al. 2008; Friedländer et al. 2012). The average abundance of each mature pig miRNA was adjusted for differences in sequencing depth between libraries by converting the

read counts to counts per million (cpm) with the edgeR package (v3.12.1) of R (R Core Team 2018), incorporating trimmed mean of M (TMM) normalization factors (Robinson and Oshlack, 2010). TMM normalization has been shown to reduce the technical bias of sequencing and control the rate of false-positive associations (Dillies et al. 2013). miRNAs expressed at < 1 cpm in ≥ 44 libraries were removed from the dataset prior to calculation of the normalization factors.

3.3.3 MicroRNA eQTL analysis

The 295 mature miRNA expression profiles retained from quality control and expression filtering were normalized using the *voom* function of the *limma* R package (v3.26.9), which accounts for the mean-variance relationship of the miRNA expression profiles across samples (Law et al. 2014). The resulting log-counts per million were treated as response variables in a genomic best linear unbiased prediction (GBLUP)-based genome-wide association (GWA) analysis utilizing the *gwaR* package developed by our group² in the R statistical environment³ (R Core Team, 2018). The GBLUP model is described in model (1):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{a} + \mathbf{e}, (1)$$

$$\mathbf{a} \sim N(0, \mathbf{G}\sigma_a^2)$$

$$\mathbf{G} = \mathbf{Z}\mathbf{Z}'$$

$$\mathbf{e} \sim N(0, \sigma_e^2 \text{diag}(1/\hat{\mathbf{w}}))$$

where $\mathbf{y}$ is the vector of normalized read log-count per million (log-cpm) associated with a mature miRNA; $\mathbf{X}$ is the incidence matrix relating the observed phenotypes to the coefficients of fixed effects $\boldsymbol{\beta}$ including the population mean, sex, and “growth group,” which is a factor broken into four classifications corresponding to phenotypic extremes for the selected traits of loin muscle area or 10th rib backfat thickness (Cardoso et al. 2008). Effect $\mathbf{a}$ is the random effect of breeding value, whose variance incorporated the genomic relationship matrix $\mathbf{G}=\mathbf{Z}\mathbf{Z}'$, where $\mathbf{Z}$ is

the matrix of standardized SNP marker effects for each animal (VanRaden, 2008; Gualdrón Duarte et al. 2014). The error term $\mathbf{e}$ included a variance inversely proportional to the precision weights produced by the *voom* function, accounting for the heterogeneous variance between miRNAs with differing expression.

The narrow-sense heritability (h^2) of the miRNA expression profiles was estimated by obtaining the ratio of the additive genetic variance and total phenotypic variance parameters resulting from the GBLUP model as shown in equation (2):

$$h^2 = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_e^2}, (2)$$

A likelihood ratio test (LRT) was utilized to assess the significance of the heritability, comparing the likelihood of model (1) with the null model assuming $h^2 = 0$. False-discovery rate (FDR) was implemented for multiple test correction ($\text{FDR} < 0.05$; Storey and Tibshirani, 2003). All miRNAs were included in the subsequent GWA analysis, regardless of heritability. The SNP effects and their variances from model (1) were estimated from a linear transformation of the $\hat{\mathbf{a}}$ matrix of estimated breeding values, as described by Gualdrón Duarte et al. (2014):

$$\hat{\mathbf{g}} = \mathbf{Z}'\mathbf{G}^{-1}\hat{\mathbf{a}}, (3)$$

$$\text{var}(\hat{\mathbf{g}}) = \mathbf{Z}'\mathbf{G}^{-1}\text{var}(\hat{\mathbf{a}})\mathbf{G}^{-1}\mathbf{Z}'$$

Individual SNP effects were standardized by dividing a SNP effect by the square root of its variance, producing the GWA test statistic of the j th SNP with a miRNA:

$$\mathbf{t}_j = \frac{\hat{g}_j}{\sqrt{\text{var}(\hat{g}_j)}}, (4)$$

The p -values associated with each significance test were assessed as described in Gualdrón Duarte et al. (2014), and FDR was utilized for multiple test correction ($\text{FDR} < 0.05$).

The miRNA eQTL (miR-eQTL) were then mapped by plotting the genomic position of each miRNA precursor sequence against the associated SNP marker position(s). miR-eQTL were defined as “local” regulators of miRNA expression if the miR-eQTL peak overlapped the genomic position of the miRNA precursor transcript. Peaks not meeting this criterion were defined as “distant” regulators. This definition of local versus distant miR-eQTL is more conservative than that used in traditional mRNA GWA analyses, which generally defines a genomic region surrounding the transcription start site of the associated mRNA (in the range of kilobases) as a “*cis*-acting” eQTL. This conservative definition of local regulators was chosen with the understanding that genomic variation in the region of the miRNA precursor transcript could potentially affect miRNA biogenesis or targeting.

After completing the GBLUP-based GWA analysis using model (1), the significant miR-eQTL peaks were characterized by estimating the proportion of variance accounted for by the peak miR-eQTL SNP (selected based on minimum *p*-value). This was accomplished by conducting a conditional GBLUP-based GWA analysis (Casiró et al. 2017). The association analysis was repeated for each miR-eQTL utilizing Equation (1) and including in the model the genotypes of the most significantly associated peak SNP (determined by *p*-value) as fixed effects. Estimating the proportion of variance explained by the peak miR-eQTL SNP was conducted using equations (5) and (6), as described in Casiró et al. 2017.

$$\widehat{Var}(q) = b^2 * var(Z_{peak}), (5)$$

where $\widehat{Var}(q)$ is the estimated variance explained by the SNP effect, b is the estimated effect of the SNP, and Z_{peak} is the genotype of the most significant SNP. The proportion of variance explained by the peak SNP is then estimated by dividing the estimated variance of the SNP effect by the total phenotypic variance components as follows:

$$\frac{\widehat{var}(q)}{\hat{\sigma}_a^2 + \hat{\sigma}_e^2 + \widehat{var}(q)}, \quad (6)$$

If all other SNP in the region failed to reach significance after fixing the peak SNP for a given miR-eQTL, this indicated the miR-eQTL effect could be explained by a SNP(s) acting in that region. Additionally, if fixing a peak SNP for a peak in one region of the genome caused a SNP association to be eliminated in another genomic region, this indicated that that single SNP was not in linkage disequilibrium with neighboring SNP but was highly correlated with SNP elsewhere in the genome (Casiró et al. 2017). Finally, if significant SNP associations remain after fixing a peak SNP, the associated miRNA expression could be controlled by multiple genomic loci acting independently of one another, with a proportion of variance in miRNA expression being explained by the given genomic region. In summary, this conditional analysis enabled the estimation of the proportion of variance in miRNA abundance explained by the peak miR-eQTL SNP for each miR-eQTL and facilitated the potential discovery of multiple genomic regions significantly associated with variation in miRNA abundance that, due to the effects of linkage disequilibrium, may have originally presented as a single miR-eQTL peak.

3.3.4 Colocalization of miR-eQTL and pQTL

Previously estimated phenotypic QTL (pQTL; genomic regions significantly associated with variation in a phenotype of interest) from the MSUPRP (Velez-Irizarry et al. 2019) were colocalized with miR-eQTL peaks to identify regions of the genome affecting variation in both miRNA expression and economically important phenotypes. Using the genomic position of miR-eQTL and pQTL peaks, defined as the range between the minimum and maximum significantly associated SNP genomic positions for a given miRNA expression profile or phenotype, multiple types of colocalization events were identified: miR-eQTL peaks completely enveloped by pQTL peaks, pQTL peaks completely enveloped by miR-eQTL peaks, and pQTL peaks overlapping

miR-eQTL peaks either upstream or downstream. Once colocalization was identified based on genomic position, the genotypes of the peak SNP of the pQTL were added as a fixed effect in a replication GWA for the miR-eQTL similar to Equation (1), utilizing FDR for multiple test correction ($FDR < 0.05$). If the previously significant SNP associations of a miR-eQTL were eliminated by adjusting for the peak SNP of the pQTL, it indicated a region of the genome affecting both variation in miRNA expression and phenotypic expression, further supporting a regulatory relationship between the miRNA and phenotype.

3.3.5 Genomic colocalization of miR-eQTL-correlated target mRNAs and pQTL

To investigate the regulatory mechanisms affecting carcass composition, meat quality, and growth phenotypes in this population, the expression profiles of 15 miRNAs with miR-eQTL were correlated with each miRNA's target genes' expression. Prediction of miRNA target genes was conducted with TargetScan using human homologs of pig miRNAs and genes, and the resulting gene lists were filtered to retain genes containing the two most preferentially conserved and effective target site classes (8mer and 7mer-m8; Agarwal et al. 2015). Gene expression profiles were obtained from next-generation sequencing of mRNA from RNA extracted from a subset of the 174 LD samples, yielding 166 individuals with both miRNA and mRNA expression; for detailed methods, see Velez-Irizarry et al. (2019). Expression profiles for miRNA and mRNA datasets (6,628 unique transcripts) were separately adjusted for the same fixed and random effects used in the linear model described in Equation (1). The corrected miRNA and gene expression values output from the GBLUP model were utilized in a two-sided Kendall's rank correlation analysis in which each miR-eQTL miRNA was correlated with its target mRNAs expressed in this population, and FDR was utilized to correct for multiple tests ($FDR < 0.05$). Thus, the genomic position of the target mRNAs exhibiting significant negative

correlation (the assumed relationship between a miRNA and its targets) were compared with pQTL regions associated with meat quality, carcass composition, and growth phenotypes previously identified in the same population (Gualdrón Duarte et al. 2016; Casiró et al. 2017; Velez-Irizarry et al. 2019). Corrected colocalized target gene expression was also correlated through Pearson's correlation to the corrected expression of colocalized phenotypes. Covariates included in linear models obtaining corrected phenotypic expression varied by phenotype, following the methods of Velez-Irizarry et al. (2019). An overlap in the genomic positions of a miR-eQTL miRNA's target mRNAs and a pQTL, in addition to a correlation between the datasets may reveal novel regulatory relationships of miRNAs and mRNAs underlying economically important pig production traits. All scripts utilized in these analyses are publicly available and can be found at <https://github.com/perrykai>.

3.4 Results

3.4.1 MicroRNA eQTL analysis

3.4.1.1 Local and distant microRNA eQTL

Results of the miR-eQTL analysis are shown in Table 3.1. In total, 295 miRNA expression profiles were included as response variables in the GBLUP-based GWA analysis, and results for associating these miRNAs with 36,292 SNPs identified 315 significant miRNA-SNP associations ($FDR < 0.05$), comprising 23 significant miR-eQTL peaks. These miR-eQTL were associated with 17 unique miRNAs and mapped to 10 different chromosomes. The genomic positions of the miR-eQTL SNPs and their associated miRNA can be found in Supplementary Table 3.1. The complete map of genomic positions of miRNA transcripts compared with miR-eQTL peak SNPs is shown in Figure 3.1, with local-acting miR-eQTL in white and located along the blue diagonal line. Five of the miR-eQTL peaks were identified as local regulators,

associated with miR-184 (DIAS0000025), miR-190b (ALGA0026452), miR-429 (ALGA0106326), miR-7135-3p (ALGA0124095), and miR-874 (ALGA0016550) as seen in Figures 3.2A–E, while 16 miR-eQTL were determined to be distant regulators of miRNA expression. None of the significantly associated miR-eQTL SNPs overlapped with mature miRNA sequences. For miR-7135-3p, the significantly associated SNP in closest proximity to the mature miRNA sequence was the peak miR-eQTL SNP (ALGA0124095), lying 16.2 kb away from the mature miRNA sequence. Additionally, while miR-140-5p abundance was associated with two SNPs on chromosome 6 (the same chromosome as the miR-140-5p precursor), the significantly associated SNPs lie approximately 104.5 kb from the miRNA precursor transcript, deeming this association a distant-acting miR-eQTL. This is demonstrated in Figure 3.1 as the green point lying close to the blue diagonal line.

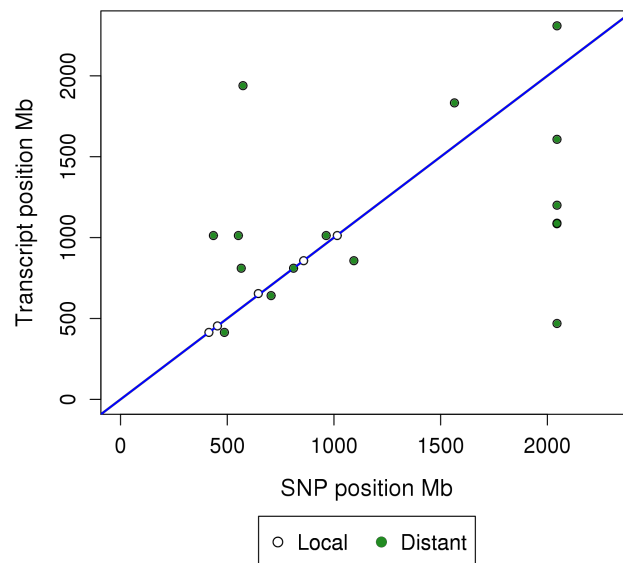


Figure 3.1. Global plot of genomic position of miRNA transcript (Mb) versus genomic position of SNP (Mb) for miR-eQTL. Significant microRNA eQTL were identified using GBLUP-based GWA models (FDR < 0.05). miR-eQTL were defined as “local” regulators of miRNA expression if the miR-eQTL peak overlapped the genomic position of the miRNA precursor transcript. Peaks not meeting this criterion were defined as “distant” regulators. The x-axis denotes the absolute position of the peak miR-eQTL SNP (Mb), while the y-axis denotes the absolute position of the miRNA precursor transcript. White points along the blue line represent *local*-acting miR-eQTL, while green points represent *distant*-acting miR-eQTL.

miR-eQTL peaks associated with two miRNAs (miR-6782-3p and miR-9785-5p) could not be classified as local or distant acting; these miRNAs are mapped to unplaced scaffolds in the current genome assembly (v11.1; Ensembl), making this distinction and their inclusion on the global miR-eQTL map in Figure 3.1 impossible. One SNP, MARC0093624 (SSC15: 122.22 Mb), was associated with five different miRNAs (let-7d-5p, let-7g, miR-1468, miR-95, and miR-9843-3p). This genomic region was found in mRNA eQTL analyses to contain a putative regulatory hotspot (Velez-Irizarry et al. 2019), indicating its potential regulatory role over multiple transcripts. In situations where a miRNA was regulated by both local- and distant-acting miR-eQTL, genomic segments surrounding local-acting miR-eQTL peaks accounted for a larger proportion of variance in its respective miRNA's expression than distant-acting peak segments.

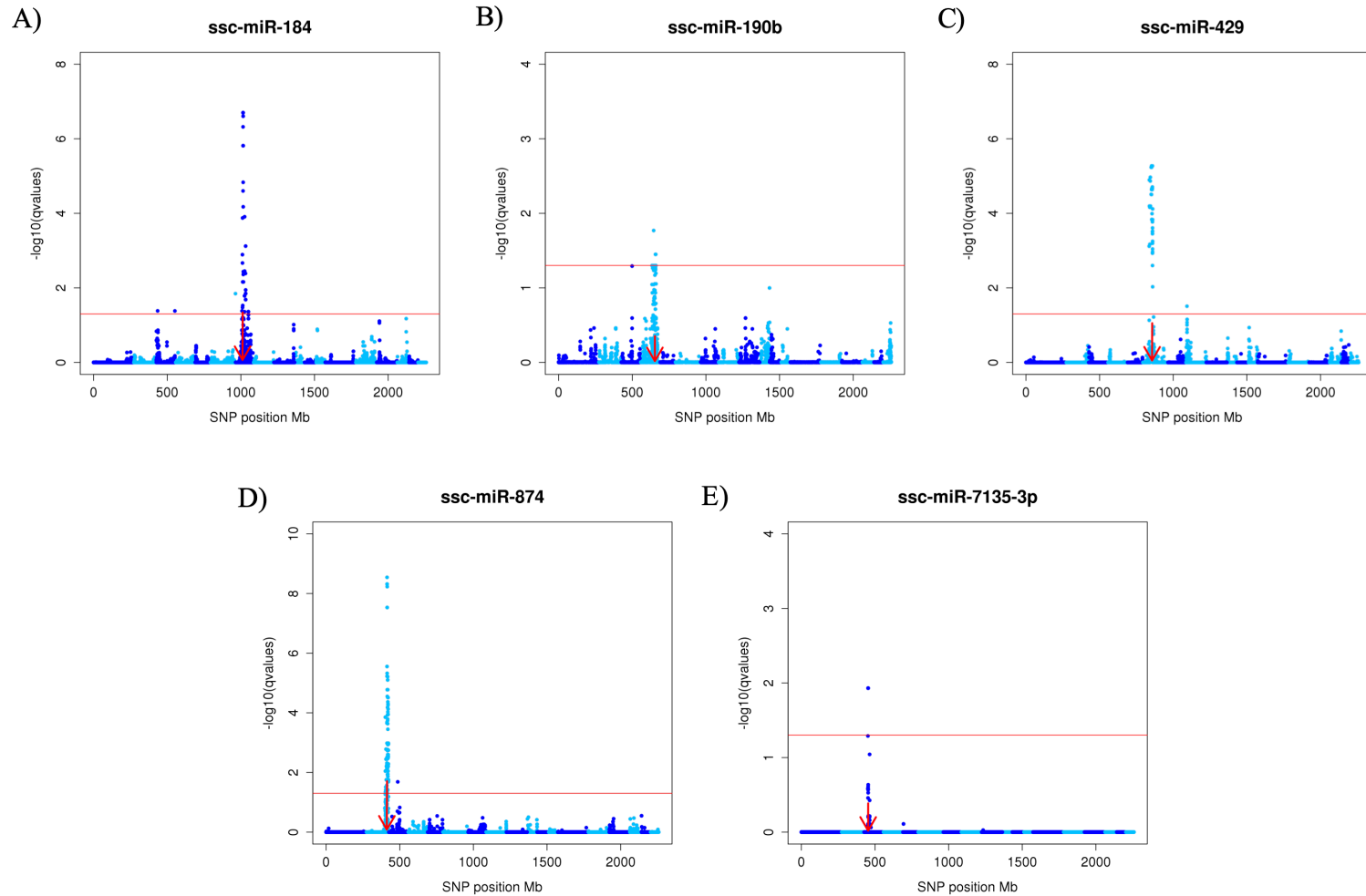


Figure 3.2 *Manhattan plots of the five local-acting miR-eQTL.* Manhattan plots depicting the five *local-acting* miR-eQTL identified through GBLUP-based GWA models. For these miR-eQTL [ssc-miR-184 (A), ssc-miR-190b (B), ssc-miR-429 (C), ssc-miR-874 (D), ssc-miR-7135-3p (E)], the position of the miR-eQTL peak overlapped the genomic position of the miRNA precursor transcript (denoted by the red arrow). Each blue point represents one miRNA-SNP association, with chromosomes differentiated by shades of blue. The x -axis denotes the absolute SNP position (Mb), while the y -axis represents the significance of the miRNA-SNP association [$-\log_{10}(q\text{ value})$]. The red line designates the threshold of significance, declared at $FDR < 0.05$.

Table 3.1 Summary of miRNA expression Quantitative Trait Loci

miRNA	miRNA Chr.	miRNA h^2 (SE)	Peak SNP ID	SNP Chr.	SNP Pos. (Mb)	Peak Range (Mb)	SNPs in Peak	Peak SNP q-value	Prop. Var. Explained Peak SNP ^a
miR-874	2	0.30 (0.15) [†]	ALGA0016550 [‡]	2	139.74	21.71	115	2.87E-09	0.36
miR-184	7	0.63 (0.17) [†]	DBWU0000430	3	9.46	0.00	1	4.14E-02	0.11
miR-184	7	0.63 (0.17) [†]	ASGA0016793	3	126.51	0.00	1	4.16E-02	0.11
miR-7135-3p	3	6.28E-8 (0.08)	ALGA0124095 [‡]	3	28.39	0.70	14	1.17E-02	0.13
miR-874	2	0.30 (0.15) [†]	ALGA0122273	3	61.02	0.00	1	2.06E-02	0.10
miR-9785-5p	NA	0.24 (0.14)	ALGA0121561	3	7.32	27.07	17	3.69E-02	0.14
miR-128	13, 15	0.24 (0.14)	ALGA0023517	4	15.33	0.00	1	4.37E-02	0.17
miR-140-5p	6	0.35 (0.16) [†]	ASGA0017748	4	7.19	0.00	1	1.39E-02	0.15
miR-190b	4	0.50 (0.17) [†]	ALGA0026452 [‡]	4	87.03	14.43	4	1.70E-02	0.19
miR-9810-3p	4	0.16 (0.13)	MARC0021620	5	16.31	0.08	2	3.27E-02	0.15
miR-140-5p	6	0.35 (0.16) [†]	ALGA0117081	6	16.97	0.39	2	4.30E-04	0.26
miR-184	7	0.63 (0.17) [†]	M1GA0026172	6	169.93	0.00	1	1.43E-02	0.10
miR-429	6	0.34 (0.16) [†]	ALGA0106326 [‡]	6	64.41	23.59	90	5.35E-06	0.28
miR-184	7	0.63 (0.17) [†]	DIAS0000025 [‡]	7	51.03	43.23	46	2.00E-07	0.34
miR-429	6	0.34 (0.16) [†]	ALGA0046283	8	7.17	0.00	1	3.10E-02	0.11
miR-6782-3p	NA	0.52 (0.17) [†]	ASGA0094215	10	24.87	9.75	4	1.55E-04	0.26
miR-1306-3p	14	0.03 (0.09)	H3GA0034702	12	52.40	0.00	1	6.13E-03	0.17
let-7d-5p	3	0.11 (0.12)	MARC0093624	15	122.22	0.35	2	1.12E-02	0.18
let-7g	13	0.10 (0.11)	MARC0093624	15	122.22	0.35	2	5.96E-03	0.18
miR-1468	X	0.19 (0.13) [†]	MARC0093624	15	122.22	0.00	1	2.02E-02	0.19
miR-345-3p	7	0.41 (0.16) [†]	H3GA0052416	15	121.81	0.07	2	3.59E-02	0.18
miR-95	8	0.42 (0.16) [†]	MARC0093624	15	122.22	0.41	3	4.65E-02	0.17
miR-9843-3p	8	0.31 (0.15) [†]	MARC0093624	15	122.22	0.41	3	3.23E-03	0.22

^a Proportion of variance explained by the peak miR-eQTL SNP (minimum p-value) when included as a fixed effect in linear models.

[†] Designates miRNAs exhibiting significant narrow-sense heritability (FDR < 0.05).

[‡] Designates local-acting miR-eQTL

3.4.1.2 Heritability of miRNA expression in pig longissimus dorsi

Results of the GBLUP analysis indicated the average h^2 of miRNA expression profiles to be 0.12, while the average h^2 of the 46 miRNAs exhibiting significantly heritable expression after the LRT was 0.34 (FDR < 0.05). The relationship between heritability of a miRNA expression profile and its significance in the LRT [$-\log_{10}(p \text{ value})$] is shown in Supplementary Figure 3.S1. The overall trend of this figure (increasing heritability associated with increased significance) indicates that the additive genetic effect is playing a significant role in the expression of these miRNAs in these samples. Heritability of a miRNA expression profile was not used as a filter for inclusion in the GWA analysis; as shown in Table 3.1, some miRNAs exhibiting non-significant heritability did have miR-eQTL, indicating that a significant association between genomic loci and variation in miRNA expression can exist regardless of the heritability of that miRNA.

3.4.2 Peak miR-eQTL SNP conditional analysis

Results of the conditional analysis (repeating the GWA analysis including the peak SNP for each miR-eQTL as a fixed effect) are shown in Table 3.2. Strong linkage disequilibrium between the peak SNP and other SNPs comprising the miR-eQTL peak caused the analysis to fail for three miR-eQTL associated with miR-184, miR-7135-3p, and miR-9810-3p (Supplementary Figure 3.S2). For 12 miR-eQTL peaks (representing 12 unique miRNAs), accounting for the peak SNP resulted in complete loss of significant associations. For six of these cases (let-7d-5p, let-7g, miR-128, miR-1306-3p, miR-1468, and miR-345-3p), there was initially only one to two statistically significant SNP association(s) and all were distant-acting. Most intriguing was the local-acting peak associated with miR-874, whose miR-eQTL peaks initially contained 115 significant SNPs (Figure 3.3A); fixing the peak significant SNP (ALGA0016550)

resulted in a complete loss of association for the local- and distant-acting signals. Similarly, for miR-429, accounting for the local-acting peak SNP (ALGA0106326) eliminated the miR-eQTL peak previously consisting of 90 SNPs ($FDR < 0.05$; Figure 3.3B). Variation in miRNA expression between segregating peak SNP genotypes for miR-874 and miR-429 are shown in Figures 3.3C, D, respectively. For the eight remaining miR-eQTL peaks, fixing the peak SNP did not eliminate all SNP associations: these consisted of distant-miR-eQTL associated with miR-140-5p (two peaks), miR-184 (three peaks), miR-429, and miR-874. Additionally, accounting for the miR-6782-3p peak SNP (ASGA0094215; SSC10) revealed significant SNP associations on SSC5 and SSC10 (peaks contain 11 and 6 SNPs, respectively), however, this miRNA lacks annotation data, making the distinction of the resulting eQTL peaks as local- or distant-acting impossible. Interestingly for miR-874, whose expression associated with both a local-acting miR-eQTL on chromosome 2 and a distant-acting SNP on chromosome 3, accounting for the distant-acting miR-eQTL peak SNP in the GWA analysis did not eliminate the local-acting signal but yielded a stronger local signal containing 126 significantly associated SNPs with ALGA0016550 remaining the peak SNP (Figure 3.3A). A similar result was seen with miR-429, where the local-acting signal remained after conditioning on the distant-acting miR-eQTL peak SNP (Figure 3.3B).

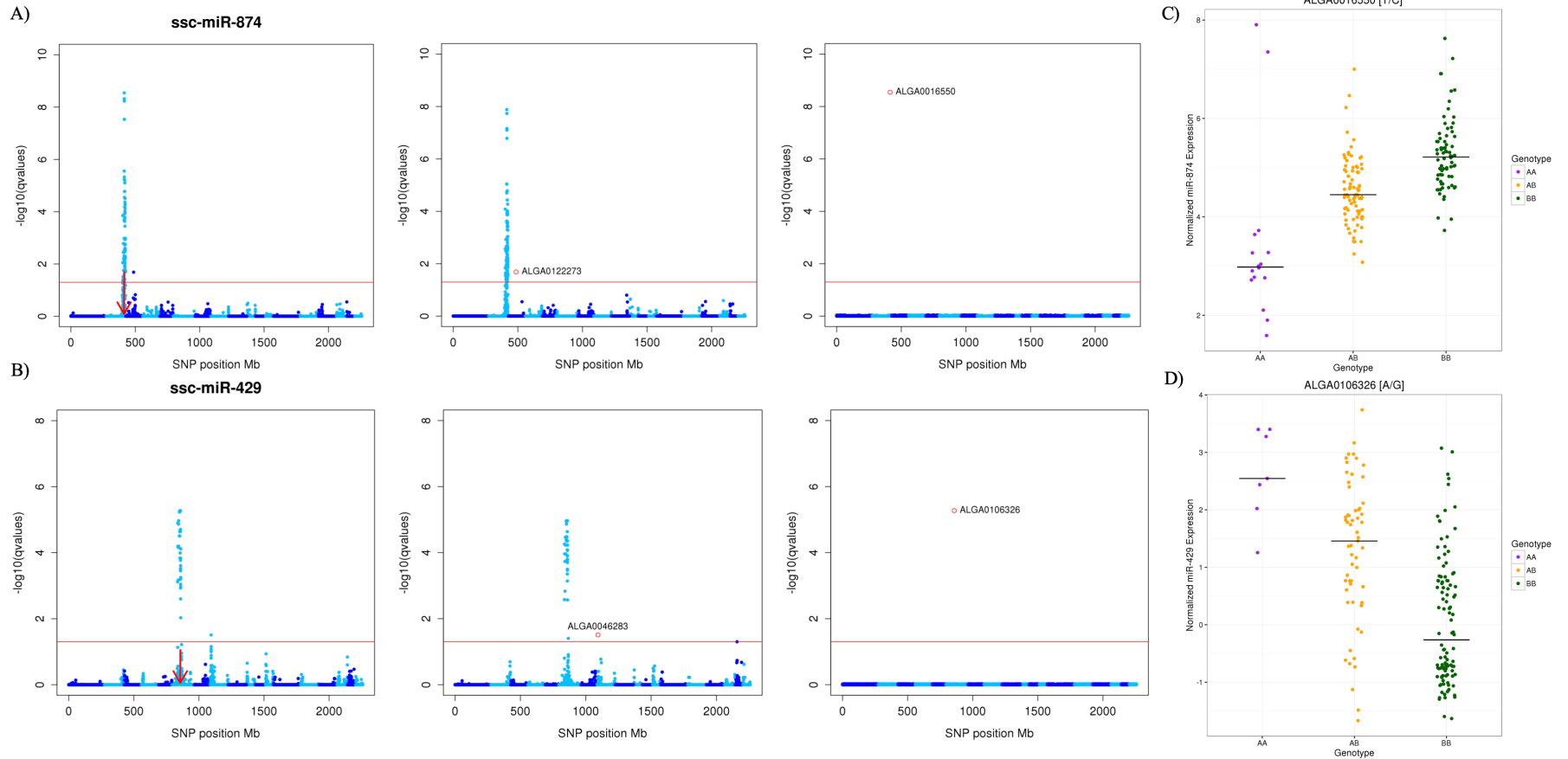


Figure 3.3 *Conditional Analyses of miR-eQTL peak SNPs.* Manhattan plots (A,B) depicting the results of the conditional analyses of miR-874 and miR-429, which included the genotype of the most significantly associated peak SNP (determined by p -value) in the GBLUP-based GWA model as a fixed effect for each miR-eQTL. (C,D) Variations of each miRNA's expression between segregating genotypes. For each Manhattan plot, the y -axis designates the $-\log_{10}(q)$ value and the x -axis indicates the absolute SNP position (Mb). Chromosomes are differentiated by shades of blue. Significance was declared at $FDR < 0.05$. (A) Conditional analysis of miR-874 miR-eQTL peak SNPs ALGA0122273 (SSC3) and ALGA0016550 (SSC2). (B) Conditional analysis of miR-429 miR-eQTL peak SNPs ALGA0046283 (SSC8) and ALGA0106326 (SSC6). Left, Manhattan plot of miR-eQTL peaks. Middle, Manhattan plot of miR-eQTL peaks after conditional analysis of *distant*-acting peak miR-eQTL SNP. Note that the distant-acting peak in each case is eliminated, but the local-acting peak remains. Right, Manhattan plot of miR-eQTL peaks, after conditional analysis of local-acting peak miR-eQTL SNP. (C) Variation in miR-874 expression between segregating SNP ALGA0016550 genotypes. The black line in each column represents the median expression of miR-874 for the respective genotype. (D) Variation in miR-429 expression between segregating SNP ALGA0106326 genotypes. The black line in each column represents the median expression of miR-429 for the respective genotype.

Table 3.2 Summary of miR-eQTL peak SNP conditional analysis results

miRNA	miRNA Chr.	Fixed SNP ID	SNP Chr.	SNPs in Peak ^a	Peak(s) Remaining ^b	SNPs in Peak(s) Remaining	Remaining Peak(s) SNP	Remaining Peak(s) SNP Chr.
miR-874	2	ALGA0016550 [‡]	2	115	0	-	-	-
miR-184	7	DBWU0000430	3	1	2	1, 63	M1GA0026172, ALGA0041952	6, 7
miR-184	7	ASGA0016793	3	1	2	1, 62	M1GA0026172, ALGA0041993	6, 7
miR-7135-3p	3	ALGA0124095 [‡]	3	14	NA	-	-	-
miR-874	2	ALGA0122273	3	1	1	126	ASGA0012467	2
miR-9785-5p	NA	ALGA0121561	3	17	0	-	-	-
miR-128	13, 15	ALGA0023517	4	1	0	-	-	-
miR-140-5p	6	ASGA0017748	4	1	1	17	ALGA0117081	6
miR-190b	4	ALGA0026452 [‡]	4	4	0	-	-	-
miR-9810-3p	4	MARC0021620	5	2	NA	-	-	-
miR-140-5p	6	ALGA0117081	6	2	1	1	ALGA0022682	4
miR-184	7	M1GA0026172	6	1	2	1, 25	H3GA0011011, ALGA0041952	3, 7
miR-429	6	ALGA0106326 [‡]	6	90	0	-	-	-
miR-184	7	DIAS0000025 [‡]	7	46	NA	-	-	-
miR-429	6	ALGA0046283	8	1	1	91	H3GA0053081	6
miR-6782-3p	NA	ASGA0094215	10	4	2	11, 6	H3GA0016379, DIAS0000707	5, 10
miR-1306-3p	14	H3GA0034702	12	1	0	-	-	-
let-7d-5p	3	MARC0093624	15	2	0	-	-	-
let-7g	13	MARC0093624	15	2	0	-	-	-
miR-1468	X	MARC0093624	15	1	0	-	-	-
miR-345-3p	7	H3GA0052416	15	2	0	-	-	-
miR-95	8	MARC0093624	15	3	0	-	-	-
miR-9843-3p	8	MARC0093624	15	3	0	-	-	-

^a Number of significantly-associated SNPs in miR-eQTL peak prior to conditional analysis^b Number of miR-eQTL peaks remaining after conditional analysis of peak miR-eQTL SNP[‡] Designates local-acting miR-eQTL

3.4.3 Colocalization of miR-eQTL and pQTL

Colocalization of miR-eQTL and pQTL based on genomic position of the miR-eQTL and pQTL peak SNPs yielded eight miR-eQTL overlapping 10 pQTL across four chromosomes (Supplementary Table 3.S2). Six distant-acting miR-eQTL for miRNAs let-7d-5p, let-7g, miR-95, miR-345-3p, miR-1468, and miR-9843-3p overlapped in genomic position with a large pQTL peak on SSC15 associated with seven meat quality phenotypes including sensory panel juiciness, sensory panel tenderness, and overall tenderness, Warner-Bratzler shear force (WBS), protein content, pH at 24 h post-mortem, driploss, and cook yield. All but one of these six miR-eQTL peaks were associated with SNP MARC0093624 (122.22 Mb); the miR-eQTL for miR-345-3p was associated with H3GA0052416, located nearby on the same chromosome (121.81 Mb; $r = -0.89$). Two local-acting miR-eQTL peaks for miR-190b (ALGA0026452: SSC4: 87.03 Mb) and miR-184 (DIAS0000025: SSC7: 51.03 Mb) colocalized with pQTL for last lumbar vertebrae backfat thickness (SSC4) and number of ribs (SSC7), respectively. The remaining 15 miR-eQTL peaks did not overlap in genomic position with any pQTL peaks.

For each miR-eQTL colocalizing with a pQTL, the pQTL SNP with the minimum p -value was identified and included as a fixed effect in a GBLUP-based GWA conditional analysis to assess the impact of the pQTL peak SNP on the associated miRNA's expression. Results of this analysis can be found in Supplementary Table 3.S3. For the local-miR-eQTL for miR-184 on SSC7, accounting for the peak pQTL SNP for number of ribs reduced the number of significantly associated SNPs. The number of ribs pQTL SNP ALGA0043983 explained 8.80% of the variance in miR-184 expression, leaving 34 SNP significantly associated with miR-184 ($FDR < 0.05$; Supplementary Table 3.S3). This means a variant in the locus containing the peak

SNP for the number of ribs pQTL is likely also having an effect on the expression of miR-184, further suggesting a regulatory relationship between the miRNA and the phenotype. For the remaining seven colocized miR-eQTL, including the peak pQTL SNP as a fixed effect in the GBLUP-based GWA eliminated all associations for that miRNA. Interestingly, for miR-190b, which colocized with last lumbar backfat thickness (SSC4), accounting for SNP ASGA0092651 in the linear model explained 15.67% of the variance in miRNA expression. For the five miR-eQTLs colocizing with pQTL on SSC15, the peak pQTL SNPs for all seven traits were either H3GA0052416 or MARC0093624 (Supplementary Table 3.S3). As noted above in Table 3.1, these are the same peak SNPs for the colocized miR-eQTL, and the two SNPs are highly correlated with one another due to their close proximity in the genome ($r = -0.89$). Conducting conditional analyses on these peak SNPs left no significantly associated SNPs and explained between 13.11 and 21.96% of the variance for a given miRNA (Supplementary Table 3.S3). For the majority of miR-eQTL colocizing with pQTL peaks, there were only one to four significant SNPs in the original miR-eQTL analysis, so these results are not surprising.

3.4.4 Genomic colocalization of miR-eQTL target mRNAs and pQTL

Significant negative correlated target genes were colocized based on genomic position with pQTL identified in the same population to identify potential novel mechanisms regulating growth, carcass composition, and meat quality traits. Significant miRNA-target correlations ranged from -0.09 to -0.18 (FDR < 0.05). MiR-874 negatively correlated with 29 target genes overlapping with pQTL for 12 phenotypes across six chromosomes (Supplementary Tables 3.S4, 3.S5 and Figure 3.4). These phenotypes include carcass 10th rib backfat thickness, loin muscle area at 16 weeks old, number of ribs, sensory panel tenderness and overall tenderness,

WBS, dressing percentage, protein content, pH at 24 h post-mortem, cook yield, driploss, and juiciness.

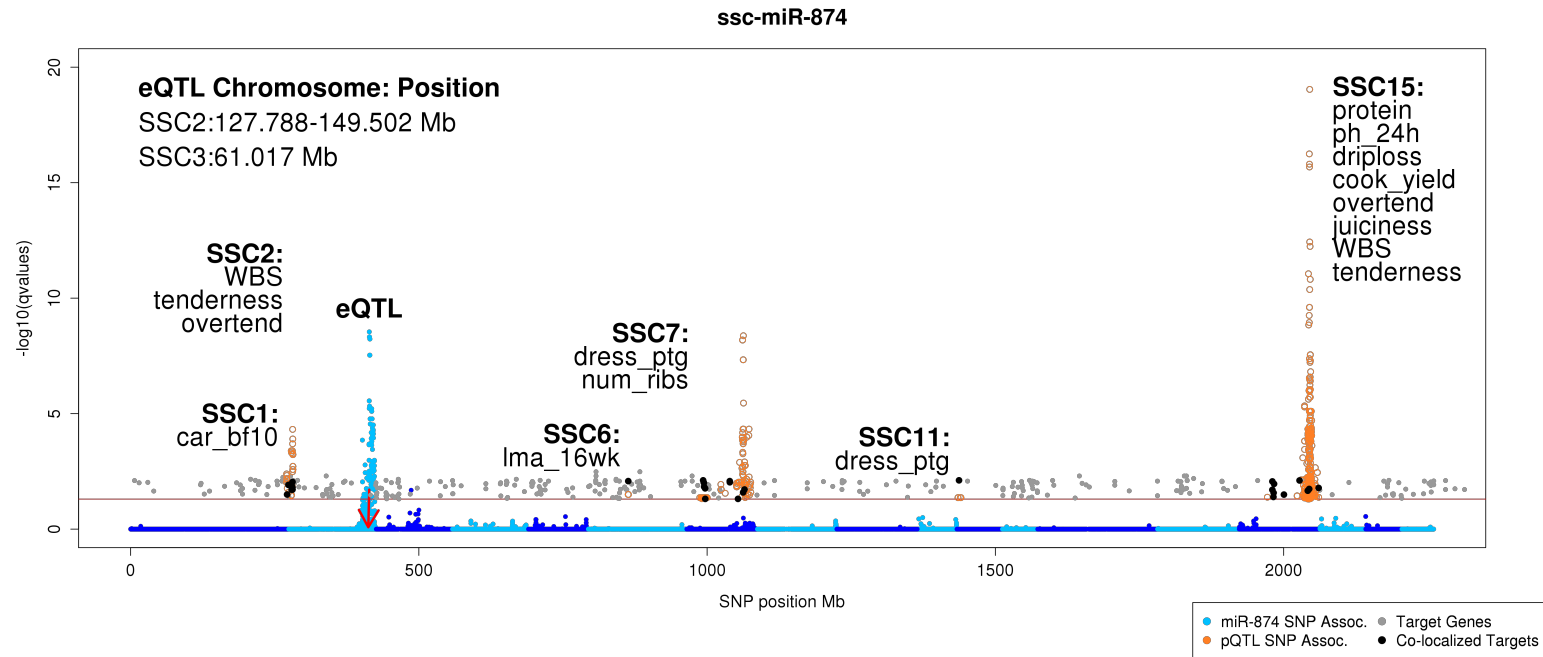


Figure 3.4. *Manhattan plot of genomic co-localization events of miR-874 target genes with pQTL.* This Manhattan plot compiles the results of the miR-eQTL, previously-identified mRNA eQTL (Velez-Irizarry et al. 2019), previously-identified pQTL (Gualdrón Duarte et al. 2016, Casiró et al. 2017, Velez-Irizarry et al. 2019), miR-874 target prediction, and correlation analyses to identify potential novel mechanisms regulating growth, carcass composition, and meat quality traits in pigs. In total, 29 unique target genes co-localized with pQTL across 6 chromosomes. Nine target genes (*PLEKHB2*, *WDR33*, *NIF3L1*, *IGFBP5*, *UGGT1*, *PRPF40A*, *METTL8*, *PTPRN*, and *KLHL30*) overlap the pQTL peak on SSC15 associated with meat quality traits. The y-axis denotes the significance of the association ($-\log_{10}(\text{q-value})$) and the x-axis indicates the absolute SNP position (Mb). Chromosomes are differentiated by shades of blue. Blue points represent miRNA-SNP associations for miR-874. The red arrow indicates the genomic position of the miR-874 precursor transcript. Orange circles represent the genomic position of pQTL. Gray points represent the genomic position of significantly negatively correlated target genes for miR-874. Black points represent the genomic position of significantly negatively correlated target genes for miR-874 whose genomic position overlaps that of a pQTL peak. Text in the upper left-hand side of plot describes the genomic position of the miR-eQTL peaks (Mb). Text in close proximity to pQTL peaks co-localizing with target genes describe the chromosome on which the pQTL resides, and which phenotypes the associated pQTL represent (car_bf10 = carcass 10th rib backfat thickness; cook_yield = cook yield; dress_ptg = dressing percentage; driploss = driploss; juiciness = sensory panel juiciness; lma_16wk = loin muscle area at 16 weeks of age; num_ribs = number of ribs; overtend = sensory panel overall tenderness; ph_24h = pH measured at 24 h post-mortem; protein = protein percentage; tenderness = sensory panel tenderness; WBS = Warner Bratzler Shear Force).

Correlation analysis between retained target genes and the colocalized phenotype yielded 11 potential targets of interest (Table 3.3). These genes were negatively correlated with miR-874 corrected expression (FDR < 0.05) in addition to either the corrected gene or corrected miR-874 expression being correlated with the colocalized phenotype ($p < 0.05$). These genes serve as potential subjects to further investigate the role of miRNAs in the regulation of traits important to the swine industry.

Table 3.3 Correlations of miR-874, target genes, and co-localized phenotypes

Gene (Chr)	τ miR-874: Gene ^a	Co- localized Phenotype ^b	R Gene: Phenotype ^c	p-value ^d	R miR-874: Phenotype ^e	p-value ^f
GPR157 (6)	-0.160	lma_16wk	-0.217	0.005*	0.242	0.002*
CDKN1A (7)	-0.088	dress_ptg	-0.157	0.044*	0.165	0.034*
ETV7 (7)	-0.126	dress_ptg	-0.116	0.135	0.165	0.034*
LEMD2 (7)	-0.148	dress_ptg	-0.126	0.106	0.165	0.034*
NUDT3 (7)	-0.132	dress_ptg	-0.147	0.059	0.165	0.034*
PFDN6 (7)	-0.151	dress_ptg	-0.156	0.045*	0.165	0.034*
ZNF451 (7)	-0.176	dress_ptg	-0.129	0.098	0.165	0.034*
ATP8A2 (11)	-0.174	dress_ptg	-0.115	0.141	0.165	0.034*
IGFBP5 (15)	-0.116	ph_24h	-0.113	0.171	0.169	0.039*
		cook_yield	-0.029	0.716	0.187	0.016*
PLEKHB2 (15)	-0.119	protein	-0.250	0.001*	0.100	0.202
PTPRN (15)	-0.123	ph_24h	-0.231	0.005*	0.169	0.039*
		cook_yield	0.020	0.801	0.187	0.016*
		WBS	0.223	0.004*	-0.009	0.905

^a Kendall's rank correlation of miR-874 to target gene (FDR < 0.05)

^b Abbreviated phenotype name: lma_16wk = loin muscle area at 16 weeks; dress_ptg = dressing percentage; ph_24h = pH measured at 24 h post-mortem; cook_yield = cook yield; protein = protein percentage; WBS = Warner Bratzler Shear Force.

^c Pearson's correlation of target gene corrected expression to co-localized corrected phenotype

^d Significance of Pearson's correlation of target gene corrected expression to co-localized corrected phenotype

^e Pearson's correlation of miR-874 corrected expression to co-localized corrected phenotype

^f Significance of Pearson's correlation of miR-874 corrected expression to co-localized corrected phenotype

3.5 Discussion

This study is the first to report miRNA-eQTL in a livestock species. The regulatory effects of miRNAs in pig skeletal muscle are not yet fully understood. Even less explored is how variation in miRNA expression affects complex traits important to the quality and consistency of pork products and the efficiency of pig production. To elucidate these relationships, we conducted a GBLUP-based GWA analysis of miRNA expression profiles in LD tissue from a F₂ Duroc × Pietrain pig population. The GBLUP model was used to estimate breeding values for each miRNA expression profile, accounting for population stratification *via* the genomic relationship matrix **G** composed of standardized marker effects for 36,292 SNPs. The GWA analysis was then conducted by estimating SNP effects through a linear transformation of the estimated breeding values and testing their association with a given miRNA expression profile. This resulted in the identification of genomic regions associated with variation in miRNA expression, or conversely, the identification of miRNAs experiencing variation in expression due to additive genetic effects in this population. Results of the miR-eQTL analysis were subsequently integrated with gene expression and phenotypic data from the same animals to elucidate the effects of miRNA regulation on target genes and downstream phenotypes. We identified 23 miR-eQTL peaks (FDR < 0.05) corresponding to 17 unique miRNAs acting both locally and distantly to regulate miRNA expression. Nine of the 17 miR-eQTL miRNAs have previously been linked to biological processes in skeletal muscle and adipose tissues (Shao et al. 2011; Jiang et al. 2013; Yan et al. 2013; Boudoukha et al. 2014; Jeanson-Leh et al. 2014; Shi et al. 2015; Siengdee et al. 2015; Zhang et al. 2015; Dai et al. 2016; Peng et al. 2016; Li et al. 2017; Xie et al. 2018; Zhang et al. 2018; Yu et al. 2019); whereas, others are novel findings for these tissues.

MicroRNAs let-7d and let-7g each exhibited distant-acting miR-eQTL overlapping a large pQTL peak for meat quality traits on SSC15. Members of the let-7 family of miRNAs have been shown to be involved in diverse biological processes including glucose metabolism, glycogen synthesis, adipogenesis, and myoblast motility in multiple species. Let-7d was identified as a direct translational repressor of the anti-inflammatory cytokine *IL13* gene, indicating its role in glucose metabolism and glycogen synthesis in human skeletal muscle (Jiang et al. 2013), while another let-7 family member, let-7g, was identified as a candidate regulator of adipogenesis during fetal skeletal muscle development in sheep (Yan et al. 2013). MicroRNA let-7g has also been shown to regulate mouse skeletal muscle myoblast motility (Boudoukha et al. 2014). These reports indicate biological mechanisms through which variation in the expression of these miRNAs could be affecting myogenesis and adipogenesis in pig skeletal muscle, ultimately affecting traits important to pork quality.

Shi et al. (2015) demonstrated that miR-128 inhibited proliferation and promoted differentiation in C2C12 myoblasts through targeted repression of *MSTN*, subsequently promoting expression of the Smad2/3 signaling pathway genes *MYF5*, *MYOG*, and *PAX 3/7*, indicating the critical role of this miRNA in regulating myogenesis. Similarly, Xie et al. (2018) investigated myogenesis-associated miRNAs in C2C12 myoblasts and identified miR-128 as a promoter of differentiation, demonstrating its role in a multi-miRNA network repressing the *MYOD*-inhibiting JNK/MAPK signaling pathway. MicroRNA-128 has also been shown to regulate bovine skeletal muscle satellite cells by targeting *SP1*, as reported by Dai et al. (2016). Upon transfection of cultured fetal hind limb muscle samples with miR-128 mimics, they observed down-regulated SP1 protein levels at late differentiation stages. The inverse relationship also held true when cells were transfected with miR-128 inhibitors, suggesting miR-

128 inhibition could promote differentiation in bovine skeletal muscle satellite cells by increasing SP1 protein levels, which promotes *MYOD* expression and the exit of cells from the cell cycle. Additionally, in a study of C2C12 myoblasts, miR-345-3p was predicted to target two genes involved in mitochondrial energy metabolism, suggesting its role in regulating ATP levels during myoblast differentiation (Siengdee et al. 2015). Yu et al. (2019) demonstrated that miR-190b expression decreased in skeletal muscle of rats after burn injury. This suppression of miR-190b resulted in increased abundance of its identified target gene, PH domain, and leucine-rich repeat protein phosphatase 1 (*PHLPP1*). Increased abundance of PHLPP1 protein, the phosphatase of Akt, indirectly enhanced the expression of cell autophagy-related proteins and promoted burn-induced skeletal muscle wasting in this study (Yu et al. 2019). MiR-95 was found to be significantly upregulated in a study profiling miRNAs in serum of golden retriever dogs with muscular dystrophy, which was validated in human patients with Duchenne muscular dystrophy (Jeanson-Leh et al. 2014). This miRNA is located in an intron of the *ABLIM2* gene in dogs and pigs, which encodes a muscle actin-binding protein. Furthermore, after observing increased miR-95 abundance in LD muscle of loss-of-function *MSTN*-mutant Meishan pigs, Li et al. (2017) utilized C2C12 myoblasts to identify its regulatory role in myogenesis. Aminoacyl-tRNA synthase complex-interacting multifunctional protein 2 (*AIMP2*) was found to be a miR-95 target gene, with overexpression of miR-95 inhibiting translation of the protein and promoting myogenic differentiation (Li et al. 2017). While further examination into the function of miRNAs -128, -345-3p, and -95 in pigs would be required, these studies all support possible roles these miRNAs exhibiting eQTL in our study could play in regulating traits of importance to the pig industry.

Skeletal muscle tissue is intercalated with intramuscular adipocytes, so it is logical to also investigate miRNAs affecting adipose tissue when studying skeletal muscle. Multiple miRNAs have been shown to regulate adipogenesis and fat deposition across species including mouse, chicken, pig, and human. MicroRNA 140-5p exhibited increased expression in adipose tissue samples from obese compared with lean mice, and miR-140-5p overexpression in cultured stromal cells promoted adipocyte formation and increased protein levels of adipogenic transcription factors and marker genes. The authors proposed a regulatory model with increased miR-140-5p expression promoting adipocyte differentiation through suppression of its target gene, transforming growth factor β receptor 1 (*TGFBRI*; Zhang et al. 2015). Investigation into the genetic regulation of meat quality traits in chicken breast muscle tissue revealed miR-140-5p as a possible regulator of intramuscular adipogenesis. Overexpression of miR-145-5p in cultured intramuscular preadipocytes resulted in increased lipid accumulation, and luciferase reporter assays confirmed the targeting relationship between miR-145-5p and retinoid X receptor gamma (*RXRG*; Zhang et al. 2018). Comparing sequences of bone morphogenic protein 5 (*BMP5*) between Large White and Meishan pigs revealed a mutation in its 3' UTR, located in miRNA binding sites for let-7c and miR-184 and associated with fatness traits in a crossbred population of the two breeds. Down-regulation of *BMP5* after transfection of recombinant vectors of let-7c and miR-184 primary sequence into pig fibroblast cells further supported the role of these miRNAs in regulating fat deposition in pigs (Shao et al. 2011). Another study of pig adipogenesis identified miR-429 as a promoter of pre-adipocyte proliferation. Peng et al. (2016) cultured both porcine subcutaneous and intramuscular pre-adipocytes and observed the down-regulation of miR-429 throughout adipogenesis. Overexpression of the miRNA prior to differentiation induction resulted in decreased lipid accumulation in both cell types, and

decreased gene and protein expression of adipogenic markers, indicating its role in regulating this process (Peng et al. 2016). MicroRNAs-184 and -429 each exhibited both local- and distant-acting miR-eQTL in this study, while miR-140-5p showed distant-acting regulation. These results indicate the complex regulation of these miRNAs that may, based on the evidence shown here, eventually influence traits related to adipogenesis and fat deposition in pigs.

These studies all indicate ways in which variation in miRNA expression could affect biological processes in skeletal muscle and potentially affect phenotypes of economic importance to the swine industry. With this in mind, in addition to studying the genomic region surrounding the miR-eQTL peaks, we identified the miRNA whose expression was being affected by genomic variants and considered what the functional effects would be on their target gene(s)' expression and related phenotypes. Target genes were colocalized based on genomic position with pQTL previously identified in the same population of pigs (Gualdrón Duarte et al. 2016; Casiró et al. 2017; Velez-Irizarry et al. 2019), revealing miR-874 as a focus of interest. The miR-874 primary transcript lies within an intron of the ubiquitously expressed *KLHL3* gene on the reverse strand, which is involved in *innate immune system* and *antigen processing and presentation* pathways and is enriched for *actin binding* and *structural molecule activity* GO terms (www.genecards.org; Stelzer et al. 2011). The peak SNP for the local-acting signal explained 36% of the variance in miR-874 expression in conditional analyses, indicating the strong likelihood of an influential variant residing in the region comprising this peak. While the miR-eQTL peak encompasses the miRNA primary transcript, the peak SNP lies 81 kb 5' of the miRNA primary transcript on the reverse strand. There are no SNPs on the SNP60 BeadChip residing between the positions of the peak SNP and the miRNA primary transcript. Thus, it is likely that the peak miR-eQTL SNP is in high linkage disequilibrium with the true variant

controlling variation in miR-874 expression. Further experiments are required to identify this candidate marker and examine the impact of variation on miR-874 expression and downstream phenotypes.

In this study, 29 unique negatively correlated miR-874 target genes colocalized with 12 pQTL across six chromosomes, including carcass backfat thickness measurements and meat quality phenotypes such as tenderness and juiciness. Intriguingly, miR-874 was previously shown to be involved in fat deposition in beef cattle. Guo et al. (2017) profiled miRNAs differentially expressed between Wagyu versus Holstein cattle, two breeds differing in intramuscular lipid content (marbling). Expression of miR-874 was lower in subcutaneous adipose tissue in Wagyu compared with Holstein cattle. Target prediction revealed *RXRA* as a likely miR-874 target, an important transcription factor in the peroxisome proliferator-activated receptor (*PPAR*) signaling pathway involved in lipid and carbohydrate metabolism (Guo et al. 2017). Additionally, Gondret et al. (2016) identified decreased expression of *RXRA* and *PPARG* (its heterodimeric partner) in perirenal and subcutaneous adipose tissues of pigs fed a high-fiber, high-fat diet compared with those fed a low-fiber, low-fat diet, indicating its role in regulating adipogenesis. *RXRA* was corroborated in our study as a miR-874 target and colocalized with a carcass 10th rib backfat pQTL on chromosome 1, supporting the hypothesis that miR-874 is involved in regulation of adipose tissue.

The cyclin-dependent kinase inhibitor 1A (*CDKN1A*) gene colocalized with dressing percentage, a trait that is affected by the overall muscularity and adiposity of an animal. *CDKN1A* functions as a regulator of cell cycle progression and is induced by myostatin in proliferating myoblasts. The upregulation of *CDKN1A* inhibits the cyclin-dependent kinase 2 (CDK2) complex, resulting in hypo-phosphorylation of retinoblastoma protein and the ultimate

arrest of myoblasts in the G1 phase of the cell cycle, promoting myoblast differentiation (Thomas et al. 2000). Conversely, in animals lacking functional myostatin, the dysregulation of *CDKN1A* and its associated genes promotes myoblast proliferation, maintaining hyperplasia and resulting in increased muscularity. Thus, variation in *CDKN1A* could contribute to the overall muscularity of an individual (a trait for which the two parental breeds of this population differ), affecting dressing percentage at harvest.

The LEM domain nuclear envelope protein 2 (*LEMD2*) gene is expressed in the inner nuclear membrane of the nuclear envelope, and in our study also colocalized with pQTL for dressing percentage. After depletion of *LEMD2* and its paralog (*LEMD3*) through RNA interference in C2C12 myoblasts, Huber et al. (2009) observed a significant reduction in differentiation as measured by myogenic index (the percentage of nuclei found in MyHC-positive myotubes). After recognizing that *LEMD2* depletion did not affect cell cycle distributions, they utilized western blotting to discover that it instead increased activation of Mitogen-Activated Protein Kinase 3 (MAPK3) a member of the MAPK signaling pathway, thereby suppressing myoblast differentiation (Huber et al. 2009). In our study, the indication that variation in *LEMD2* expression might affect muscle mass would relate to many economically relevant traits in pigs, including dressing percentage.

Insulin-like growth factor-binding protein 5 (*IGFBP5*) colocalized with pH at 24 h post-mortem in our analysis. This gene was identified in a GWAS study of Finnish Yorkshire boars as a candidate gene for 24-h meat pH (Verardo et al. 2017). The authors also discovered an eQTL peak on SSC15 encompassing *IGFBP5*, and subsequent gene-transcription factor network analysis indicated its potential for regulating this trait (Verardo et al. 2017). Offspring of transgenic mice overexpressing *IGFBP5* were shown to exhibit a reduction in skeletal muscle

tissue mass of 31% and decreased muscle fiber hypertrophy, predicted to be due to inhibition of IGF-II activity (Salih et al. 2004). Additionally, Ren et al. (2008) knocked down *IGFBP5* expression in C2C12 mouse myoblasts and observed reduced *MYOG* expression and inhibited myotube formation which was not rescuable with exogenous *MYOG* transfection in the absence of *IGFBP5*. Finally, *IGFBP5* has been shown to be a target of miR-143 in two studies, one assessing its effects on proliferation and differentiation of primary bovine muscle satellite cells (Zhang et al. 2017) and the other investigating muscle regeneration in satellite cells and primary myoblasts from mice and humans (Soriano-Arroquia et al. 2016). The above studies demonstrate the role of *IGFBP5* expression on skeletal muscle development and the potential impacts of miRNA regulation on downstream phenotypes.

Many of the predicted miR-874 target genes identified in our study have not been characterized for their roles in skeletal muscle. Further investigation into the effects of miR-874 and these target genes are needed to fully elucidate its role in pig production traits.

Due to the limited number of recombination events occurring in our F₂ population, long-range persistence of linkage disequilibrium limited the resolution of the GWA analysis resulting in large genomic segments being significantly associated with each miRNA expression profile. In some cases, this limited our ability to conduct the conditional analysis on a peak SNP due to high correlations between genotypic frequencies of SNPs in close proximity. However, we have shown here the successful utilization of this dataset for identifying local- and distant-acting regulators of miRNA expression despite the extent of linkage disequilibrium in this population. The SNP MARC0093624 (SSC15: 122.22 Mb) was associated with five different miRNAs in our GWA analysis, which may lead to the identification of this marker as a miR-eQTL hotspot. While this genomic region did produce a putative regulatory hotspot in mRNA eQTL analyses

(Velez-Irizarry et al. 2019), previous literature has reported eQTL-hotspot analyses being subject to high false-positive rates (Breitling et al. 2008). Thus, further analysis is required to confidently define MARC0093624 as a miR-eQTL hotspot. Additionally, while the recent update to the *Sus scrofa* genome assembly (v11.1; Ensembl) is a vast improvement over the previous reference genome, there still exists a lack of accurate genome assembly data for some genomic regions, making classifying a miR-eQTL as local- versus distant-acting impossible in some cases. For the majority of cases, however, we were successful in classifying the type of regulatory relationship occurring between genomic regions harboring variants and miRNAs exhibiting variation in expression.

3.6 Conclusion

No previous studies evaluating miRNA-eQTL in pigs have been reported. Our genome-wide analysis of miRNA expression profiles successfully identified genomic regions affecting the expression of 17 unique miRNAs, indicating that miRNA expression in this tissue does have a genetic component. We then examined the potential effects miRNAs experiencing variation could have on the phenotypes measured in this population, demonstrating some key miRNAs colocalizing with growth, carcass composition, and meat quality traits. Some of these miRNAs had previously been identified for their involvement in skeletal muscle and adipose tissues, whereas others represented novel findings. While more research is needed to confirm the roles of miRNA regulation in these traits, our work has contributed to the understanding of regulatory mechanisms underlying complex trait phenotypes in pigs.

3.7 Author contributions

CE, JS, and RB conceived and supervised this research. KD contributed to RNA extraction, performed the microRNA bioinformatic and statistical analysis, composed the article,

and created the figures and tables. SC developed code for and conducted the phenotypic QTL analysis. DV-I contributed to RNA extraction, developed the code used for statistical analysis, and provided results of the mRNA eQTL analysis and updated phenotypic QTL analysis. NR performed tissue and phenotypic collection and RNA extraction and oversaw RNA sequencing. All authors contributed to the article and approved the submitted version.

3.8 Acknowledgements

We wish to acknowledge the MSU Research Technology Support Facility Genomics Core for their support in sequencing of small RNA libraries, and the MSU Institute of Cyber-Enabled Research for providing the computing environment for data storage and analysis. We also wish to express gratitude to Scott Funkhouser and Ryan Corbett for their feedback on this work.

APPENDIX

APPENDIX

The Supplementary Material for this article can be found online at:

<https://www.frontiersin.org/articles/10.3389/fgene.2021.644091/full#supplementary-material>

Figure 3.S1. Heritability of miRNA expression. Narrow-sense heritability (h^2) of the 295 miRNA expression profiles was estimated by obtaining the ratio of the additive genetic variance and total phenotypic variance parameters resulting from the GBLUP model. Significance of heritability was assessed using LRTs, and FDR was implemented for multiple test correction. The x-axis denotes narrow-sense heritability of the miRNA expression profiles; the y-axis denotes the log-adjusted p -values of the LRTs. Highlighted in red are the 46 miRNAs exhibiting significantly heritable expression in this dataset ($FDR < 0.05$).

Figure 3.S2. Visualization of correlation between SNPs in miR-eQTL peaks failing conditional analysis. This figure shows the results of correlation analyses between SNPs in the miR-eQTL peak for those miR-eQTL that failed the conditional analyses, repeating the GBLUP-based GWA analysis incorporating the peak SNP for each miR-eQTL as a fixed effect. The significantly associated SNPs comprising the miR-eQTL peak for (A) miR-184, (B) miR-7135-3p, and (C) miR-9810-3p were included in each respective correlation analysis. Each square represents the correlation between a pair of SNPs, identified above and on the diagonal in each plot. The strength of each pair's correlation is depicted by increasingly saturated color; blue shades represent positive correlations and red shades represent negative correlations between two SNPs.

REFERENCES

REFERENCES

1. Agarwal, V., Bell, G. W., Nam, J. W., and Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *eLife* 4:e05005.
2. Barbut, S., Sosnicki, A. A., Lonergan, S. M., Knapp, T., Ciobanu, D. C., Gatcliffe, L. J., et al. (2008). Progress in reducing the pale, soft and exudative (PSE) problem in pork and poultry meat. *Meat Sci.* 79, 46–63.
3. Boudoukha, S., Rivera Vargas, T., Dang, I., Kropp, J., Cuvellier, S., Gautreau, A., et al. (2014). MiRNA let-7g regulates skeletal myoblast motility via Pinch-2. *FEBS Lett.* 588, 1623–1629.
4. Breitling, R., Li, Y., Tesson, B. M., Fu, J., Wu, C., Wiltshire, T., et al. (2008). Genetical genomics: spotlight on QTL hotspots. *PLoS Genet.* 4, 2–5.
5. Cai, Z., Zhang, L., Jiang, X., Sheng, Y., and Xu, N. (2015). Differential miRNA expression profiles in the longissimus dorsi muscle between intact and castrated male pigs. *Res. Vet. Sci.* 99, 99–104.
6. Cardoso, F. F., Rosa, G. J. M., Steibel, J. P., Ernst, C. W., Bates, R. O., and Tempelman, R. J. (2008). Selective transcriptional profiling and data analysis strategies for expression quantitative trait loci mapping in outbred F 2 populations. *Genetics* 180, 1679–1690.
7. Casiró, S., Velez-Irizarry, D., Ernst, C. W., Raney, N. E., Bates, R. O., Charles, M. G., et al. (2017). Genome-wide association study in an F2 Duroc x Pietrain resource population for economically important meat quality and carcass traits1. *J. Anim. Sci.* 95, 545–558.
8. Ciobanu, D. C., Lonergan, S. M., and Huff-Lonergan, E. J. (2011). “Genetics of meat quality and carcass traits,” in *The Genetics of the Pig*, eds M. F. Rothschild and A. Ruvinsky (New York, NY: CAB International), 355.
9. Dai, Y., Zhang, W. R., Wang, Y. M., Liu, X. F., Li, X., Ding, X., et al. (2016). microRNA-128 regulates the proliferation and differentiation of bovine skeletal muscle satellite cells by repressing Sp1. *Mol. Cell. Biochem.* 414, 37–46.
10. Daza, K. R., Steibel, J. P., Velez-Irizarry, D., Raney, N. E., Bates, R. O., and Ernst, C. W. (2017). Profiling and characterization of a longissimus dorsi muscle microRNA dataset from an F 2 Duroc × Pietrain pig resource population. *Genom. Data* 13, 50–53.
11. Dillies, M.-A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., et al. (2013). A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *BriefBioinform.* 14, 671–683.

12. Dong, H., Luo, L., Hong, S., Siu, H., Xiao, Y., Jin, L., et al. (2010). Integrated analysis of mutations, miRNA and mRNA expression in glioblastoma. *BMC Syst. Biol.* 4:163.
13. Edwards, D. B., Ernst, C. W., Tempelman, R. J., Rosa, G. J. M., Raney, N. E., Hoge, M. D., et al. (2008a). Quantitative trait loci mapping in an F2 Duroc x Pietrain resource population: I. Growth traits. *J. Anim. Sci.* 86, 241–253.
14. Edwards, D. B., Ernst, C. W., Raney, N. E., Doumit, M. E., Hoge, M. D., and Bates, R. O. (2008b). Quantitative trait locus mapping in an F2 Duroc × Pietrain resource population: II. Carcass and meat quality traits1. *J. Anim. Sci.* 86, 254–266.
15. Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 40, 37–52.
16. Gamazon, E. R., Innocenti, F., Wei, R., Wang, L., Zhang, M., Mirkov, S., et al. (2013). A genome-wide integrative study of microRNAs in human liver. *BMC Genomics* 14:395.
17. Gondret, F., Vincent, A., Houée-Bigot, M., Siegel, A., Lagarrigue, S., Louveau, I., et al. (2016). Molecular alterations induced by a high-fat high-fiber diet in porcine adipose tissues: variations according to the anatomical fat location. *BMC Genomics* 17:120.
18. Griffiths-Jones, S., Saini, H. K., van Dongen, S., and Enright, A. J. (2008). miRBase: tools for microRNA genomics. *Nucleic Acids Res.* 36, D154–D158.
19. Gualdrón Duarte, J. L., Cantet, R. J., Bates, R. O., Ernst, C. W., Raney, N. E., and Steibel, J. P. (2014). Rapid screening for phenotype-genotype associations by linear transformations of genomic evaluations. *BMC Bioinform.* 15:246.
20. Gualdrón Duarte, J. L., Cantet, R. J. C., Bernal Rubio, Y. L., Bates, R. O., Ernst, C. W., Raney, N. E., et al. (2016). Refining genomewide association for growth and fat deposition traits in an F2 pig population. *J. Anim. Sci.* 94, 1387–1397.
21. Guo, Y., Zhang, X., Huang, W., and Miao, X. (2017). Identification and characterization of differentially expressed miRNAs in subcutaneous adipose between Wagyu and Holstein cattle. *Sci. Rep.* 7:44026.
22. Horak, M., Novak, J., and Bienertova-Vasku, J. (2016). Muscle-specific microRNAs in skeletal muscle development. *Dev. Biol.* 410, 1–13.
23. Huan, T., Rong, J., Liu, C., Zhang, X., Tanriverdi, K., Joehanes, R., et al. (2015). Genome-wide identification of microRNA expression quantitative trait loci. *Nat. Commun.* 6:6601.
24. Huber, M. D., Guan, T., and Gerace, L. (2009). Overlapping functions of nuclear envelope proteins NET25 (Lem2) and emerin in regulation of extracellular signal-regulated kinase

- signaling in myoblast differentiation. *Mol. Cell. Biol.* 29, 5718–5728.
25. Jeanson-Leh, L., Lameth, J., Krimi, S., Buisset, J., Amor, F., Le Guiner, C., et al. (2014). Serum profiling identifies novel muscle miRNA and cardiomyopathy- related miRNA biomarkers in golden retriever muscular dystrophy dogs and duchenne muscular dystrophy patients. *Am. J. Pathol.* 184, 2885–2898.
 26. Jiang, L. Q., Franck, N., Egan, B., Sjögren, R. J. O., Katayama, M., Duque- Guimaraes, D., et al. (2013). Autocrine role of interleukin-13 on skeletal muscle glucose metabolism in type 2 diabetic patients involves microRNA let-7. *Am. J. Physiol. Metab.* 305, E1359–E1366.
 27. Jing, L., Hou, Y., Wu, H., Miao, Y., Li, X., Cao, J., et al. (2015). Transcriptome analysis of mRNA and miRNA in skeletal muscle indicates an important network for differential residual feed intake in pigs. *Sci. Rep.* 5:11953.
 28. Law, C. W., Chen, Y., Shi, W., and Smyth, G. K. (2014). Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 15:R29.
 29. Li, B., Xie, S., Cai, C., Qian, L., Jiang, S., Ma, D., et al. (2017). MicroRNA- 95 promotes myogenic differentiation by downregulation of aminoacyl-tRNA synthase complex-interacting multifunctional protein 2. *Oncotarget* 8, 111356– 111368.
 30. Mai, M., Jin, L., Tian, S., Liu, R., Huang, W., Tang, Q., et al. (2016). Deciphering the microRNA transcriptome of skeletal muscle during porcine development. *PeerJ* 4:e1504.
 31. Mármol-Sánchez, E., Ramayo-Caldas, Y., Quintanilla, R., Cardoso, T. F., González-Prendes, R., Tibau, J., et al. (2020). Co-expression network analysis predicts a key role of microRNAs in the adaptation of the porcine skeletal muscle to nutrient supply. *J. Anim. Sci. Biotechnol.* 11:10.
 32. Meuwissen, T., Hayes, B., and Goddard, M. (2016). Genomic selection: a paradigm shift in animal breeding. *Anim. Front.* 6, 6–14.
 33. Peng, Y., Chen, F.-F., Ge, J., Zhu, J.-Y., Shi, X.-E., Li, X., et al. (2016). miR-429 inhibits differentiation and promotes proliferation in porcine preadipocytes. *Int. J. Mol. Sci.* 17:2047.
 34. Qin, L., Chen, Y., Liu, X., Ye, S., Yu, K., Huang, Z., et al. (2013). Integrative analysis of porcine microRNAome during skeletal muscle development. *PLoS One* 8:e72418.
 35. R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
 36. Ramos, A. M., Crooijmans, R. P. M. A., Affara, N. A., Amaral, A. J., Archibald, A. L., Beever, J. E., et al. (2009). Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS One*

4:e6524.

37. Ren, H., Yin, P., and Duan, C. (2008). IGFBP-5 regulates muscle cell differentiation by binding to IGF-II and switching on the IGF-II auto-regulation loop. *J. Cell Biol.* 182, 979–991.
38. Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11:R25.
39. Salih, D. A. M., Tripathi, G., Holding, C., Szeszak, T. A. M., Gonzalez, M. I., Carter, E. J., et al. (2004). Insulin-like growth factor-binding protein 5 (Igfbp5) compromises survival, growth, muscle development, and fertility in mice. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4314–4319.
40. Schwab, C. R., and Baas, T. J. (2008). Direct and Correlated Responses to Selection for Intramuscular Fat in Duroc Swine Fat in Duroc Swine. *Animal Industry Report: AS 654, ASL R2351*. Ames, IA: Iowa State University.
41. Shao, G. C., Luo, L. F., Jiang, S. W., Deng, C. Y., Xiong, Y. Z., and Li, F. E. (2011). A C/T mutation in microRNA target sites in BMP5 gene is potentially associated with fatness in pigs. *Meat Sci.* 87, 299–303.
42. Shi, L., Zhou, B., Li, P., Schinckel, A. P., Liang, T., Wang, H., et al. (2015). MicroRNA-128 targets myostatin at coding domain sequence to regulate myoblasts in skeletal muscle development. *Cell. Signal.* 27, 1895–1904.
43. Siengdee, P., Trakooljul, N., Murani, E., Schwerin, M., Wimmers, K., and Ponsuksili, S. (2013). Transcriptional profiling and miRNA-dependent regulatory network analysis of longissimus dorsi muscle during prenatal and adult stages in two distinct pig breeds. *Anim. Genet.* 44, 398–407.
44. Siengdee, P., Trakooljul, N., Murani, E., Schwerin, M., Wimmers, K., and Ponsuksili, S. (2015). MicroRNAs regulate cellular ATP levels by targeting mitochondrial energy metabolism genes during C2C12 myoblast differentiation. *PLoS One* 10:e0127850.
45. Soriano-Arroquia, A., McCormick, R., Molloy, A. P., McArdle, A., and Goljanek-Whysall, K. (2016). Age-related changes in miR-143-3p:Igfbp5 interactions affect muscle regeneration. *Aging Cell* 15, 361–369.
46. Steibel, J. P., Bates, R. O., Rosa, G. J. M., Tempelman, R. J., Rilington, V. D., Ragavendran, A., et al. (2011). Genome-wide linkage analysis of global gene expression in loin muscle tissue identifies candidate genes in pigs. *PLoS One* 6:e16766.
47. Stelzer, G., Dalah, I., Stein, T., Satanower, Y., Rosen, N., Nativ, N., et al. (2011). In-silico human genomics with GeneCards. *Hum. Genom.* 5:709.

48. Storey, J. D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9440–9445.
49. Thomas, M., Langley, B., Berry, C., Sharma, M., Kirk, S., Bass, J., et al. (2000). Myostatin, a negative regulator of muscle growth, functions by inhibiting myoblast proliferation. *J. Biol. Chem.* 275, 40235–40243.
50. USDA (2020). USDA Agricultural Projections to 2029. Washington, DC: United States Department of Agriculture, 1–124.
51. VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423.
52. Velez-Irizarry, D., Casiro, S., Daza, K. R., Bates, R. O., Raney, N. E., Steibel, J. P., et al. (2019). Genetic control of longissimus dorsi muscle gene expression variation and joint analysis with phenotypic quantitative trait loci in pigs. *BMC Genomics* 20:3.
53. Verardo, L. L., Sevón-Aimonen, M.-L., Serenius, T., Hietakangas, V., and Uimari, P. (2017). Whole-genome association analysis of pork meat pH revealed three significant regions and several potential genes in finnish Yorkshire pigs. *BMC Genet.* 18:13.
54. Wohlers, I., Bertram, L., and Lill, C. M. (2018). Evidence for a potential role of miR-1908-5p and miR-3614-5p in autoimmune disease risk using integrative bioinformatics. *J. Autoimmun.* 94, 83–89.
55. Xi, Y., Liu, H., Zhao, Y., Li, J., Li, W., Liu, G., et al. (2018). Comparative analyses of longissimus muscle miRNAsomes reveal microRNAs associated with differential regulation of muscle fiber development between Tongcheng and Yorkshire pigs. *PLoS One* 13:e0200445.
56. Xie, S. J., Li, J. H., Chen, H. F., Tan, Y. Y., Liu, S. R., Zhang, Y., et al. (2018). Inhibition of the JNK/MAPK signaling pathway by myogenesis-associated miRNAs is required for skeletal muscle development. *Cell Death Differ.* 25, 1581–1597.
57. Yan, X., Huang, Y., Zhao, J.-X., Rogers, C. J., Zhu, M.-J., Ford, S. P., et al. (2013). Maternal obesity downregulates microRNA let-7g expression, a possible mechanism for enhanced adipogenesis during ovine fetal skeletal muscle development. *Int. J. Obes.* 37, 568–575.
58. Yu, Y., Yang, L., Han, S., Wu, Y., Liu, L., Chang, Y., et al. (2019). MIR-190B alleviates cell autophagy and burn-induced skeletal muscle wasting via modulating PHLPP1/Akt/FoxO3A signaling pathway. *Shock* 52, 513–521.
59. Zhang, M., Li, D. H., Li, F., Sun, J. W., Jiang, R. R., Li, Z. J., et al. (2018). Integrated analysis of miRNA and genes associated with meat quality reveals that gga-miR-140-5p affects intramuscular fat deposition in chickens. *Cell. Physiol. Biochem.* 46, 2421–2433.

60. Zhang, W. R., Zhang, H. N., Wang, Y. M., Dai, Y., Liu, X. F., Li, X., et al. (2017). miR-143 regulates proliferation and differentiation of bovine skeletal muscle satellite cells by targeting IGFBP5. *Vitr. Cell. Dev. Biol. Anim.* 53, 265–271.
61. Zhang, X., Chang, A., Li, Y., Gao, Y., Wang, H., Ma, Z., et al. (2015). miR-140-5p regulates adipocyte differentiation by targeting transforming growth factor- β signaling. *Sci. Rep.* 5:18118.

CHAPTER FOUR

CONCLUSIONS AND FUTURE DIRECTIONS

4.1 Conclusions

MicroRNAs, colloquially termed the “fine tuners” of gene expression, post-transcriptionally regulate gene expression through sequence complementarity with target genes. While it is understood that miRNAs regulate numerous biological processes, further investigation is required to understand the impacts of miRNA regulation in pig skeletal muscle and their potential effects on traits important to pig production. The purpose of this dissertation research was to improve our understanding of regulatory interactions underlying complex meat quality, carcass composition, and growth traits in pigs. Specifically, we strove to 1) Profile and characterize the small RNA population expressed in *longissimus dorsi* (LD) skeletal muscle of 174 F₂ Duroc x Pietrain pigs from the MSU Pig Resource Population (MSUPRP), and 2) Conduct an integrated miR-eQTL analysis, to identify genomic regions underlying variation in miRNA abundance in the same individuals and reveal candidate genes regulating complex traits relevant to the pig production industry.

The composition of small RNA classes present in the skeletal muscle of selected F₂ MSUPRP pigs were characterized through sequence homology searches with pig, human, and mouse databases. We also quantified the expression of 295 known mature miRNAs and predicted 27 unique candidate pig-novel miRNA precursors from small RNA sequencing data of 174 adult pig LD samples. Small RNA sequencing data were deposited in the Sequence Read Archive, BioProject PRJNA363073 (Daza et al. 2017), and the results of these analyses were presented in Chapter 2 of this dissertation.

The regulatory effects of miRNAs in pig skeletal muscle are not yet fully understood. Even less explored is how variation in miRNA expression affects complex traits important to the quality and consistency of pork products and the efficiency of pig production. In Chapter 3 of this dissertation, we conducted a GBLUP-based GWA analysis of 295 miRNA expression profiles in pig skeletal muscle to elucidate these relationships. We identified 23 miR-eQTL peaks ($\text{FDR} < 0.05$) associated with variation in the expression of 17 unique miRNAs, acting both *locally* and *distantly* (Daza et al. 2021). This indicates that additive genetic effects significantly impact the expression of miRNAs in this population. The miR-eQTL results were then integrated with gene expression and phenotypic data from the same individuals to reveal potential regulatory relationships affecting traits encompassing animal growth, carcass composition, and meat quality phenotypes. Notably, miR-184, miR-429, and miR-874 exhibited strong *local*-acting miR-eQTL, meaning the positions of the SNPs comprising the miR-eQTL peak overlapped the genomic position of the miRNA precursor transcript. Nine of the miR-eQTL miRNAs had previously been linked to biological processes in skeletal muscle and adipose tissues, while the roles of the remaining eight miRNAs were novel.

The miR-eQTL peaks were further characterized through colocalization analyses with previously identified pQTL (Velez-Irizarry et al. 2019), and predicted miR-eQTL miRNA target genes negatively correlated with miRNA abundance were co-localized based on genomic position with pQTL. Notably, miR-874 exhibited negative correlation with 29 target genes overlapping pQTL for 12 phenotypes across six chromosomes; 11 of these became genes of interest by either miR-874 or gene expression being significantly correlated with the colocalized phenotype ($p < 0.05$).

This work produced putative novel pig miRNAs for further study and validation, contributing to our understanding of the miRNA landscape in pig skeletal muscle. Additionally, we generated a resource of known miRNA expression profiles to use in evaluating the genetic architecture of miRNA expression and impacts of miRNA regulation of gene expression on complex traits in adult pigs. The identification of the candidate miRNAs and predicted target genes demonstrate the ability of this work to generate hypothesis-driven research elucidating the roles of miRNA regulation on gene expression and phenotypes of interest to pig production.

4.2 Future directions

There are many avenues this research could pursue to further improve our understanding of the genetic regulation of complex traits in pigs. For instance, further investigation into genomic regions harboring SNPs associated with miRNA variation could yield valuable information regarding *local* and *distant* genetic regulation of miRNA expression. Combining proteomic data with miRNA and mRNA transcriptomic data from the same individuals could also strengthen predicted regulatory relationships, as miRNA regulation is post-transcriptional in nature and commonly affects protein abundance. Also, as mentioned previously, our GWAS study revealed candidate miRNAs for further investigation into their regulatory functions in skeletal muscle tissue.

Our group has begun to investigate the predicted miR-874 target gene, citron Rho-interacting serine/threonine kinase (*CIT*; protein name CITK) and its effects on proliferation and differentiation of C2C12 mouse myoblasts in culture. We were drawn to this gene for multiple reasons. First, the *CIT* predicted target site was one of the most highly conserved types (7mer-m8; nucleotides two through seven of the seed sequence are complementary, plus position 8). Next, *CIT* exhibited a *distant*-acting eQTL in GWAS of the same animals conducted by Velez-

Irizarry et al. (2019). This eQTL included a putative hotspot SNP on SSC15 (H3GA0052416), which was also the peak pQTL SNP for meat quality and protein percentage traits in the same dataset (Velez-Irizarry et al. 2019). Finally, expression profiling of genes in skeletal muscle of Yorkshire x Landrace pig fetuses across gestational stages revealed *CIT* to be differentially expressed, decreasing in abundance from 41 to 70 days gestation (unpublished data from R. Corbett). These data indicate that *CIT* abundance is variable in fetal skeletal muscle development and adult tissue and may be associated with traits of importance to pig production.

The CITK protein is known to act in the late stages of cell division, performing multiple roles including binding proteins comprising the actomyosin contractile ring and central spindle (e. g. anillin, myosin, and rhoA (Ras Homolog Family Member A), etc.) and maintaining the organization of proteins at the midbody, the cellular structure contained in the intercellular bridge that organizes the proteins regulating abscission (Fig 4.1). Additionally, CITK has been shown to recruit casein kinase 2 alpha (CK2 α), phosphorylating tubulin beta 3 (TUBB3) to promote the stabilization of microtubules at the midbody. Regardless of mechanism, dysregulation of CITK expression results in abscission failure (D’Avino 2017 (commentary); Dema et al. 2018).

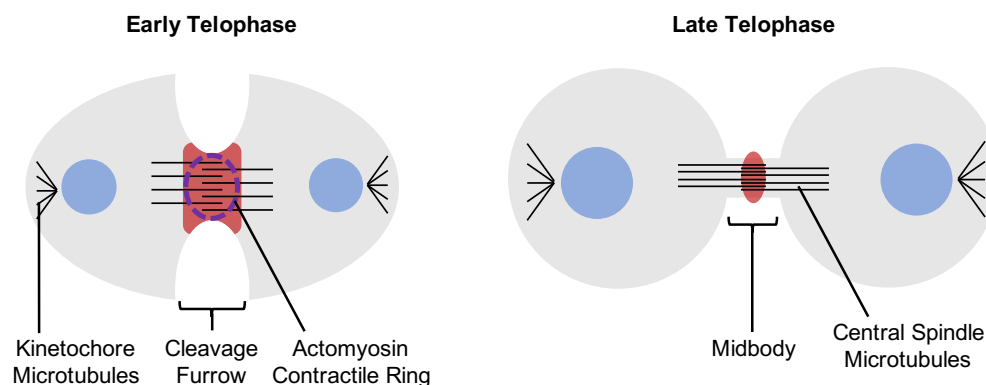


Figure 4.1 *Localization of CITK in early and late telophase.* DNA/chromosomes are depicted in blue, the actomyosin contractile ring in purple, and microtubules in black. The CITK protein accumulates at the cleavage furrow in early telophase and forms a ring-like distribution surrounding the midbody in late telophase, depicted here in red. Adapted from D’Avino 2017.

Little research has been conducted to determine the roles of CIT in skeletal muscle. One study reported decreased CITK expression in adult rat *tibialis anterior* muscle samples transfected *in vivo* with the rhoGEF (rho-guanine nucleotide exchange factor) domain of Obscurin, a protein involved in sarcomere development and function. This decrease was proposed to be a result of increased rhoA activity, which dysregulated CITK expression and altered its localization in the sarcomere (Ford-Speelman et al. 2009). Other studies defining the functions of *CIT* were conducted *in vitro* in various cell lines (e. g. HeLa, HEK, COS7) (Di Cunto et al. 1998; Gai et al. 2011; Horton et al. 2015; Dema et al. 2018).

Our goal is to investigate the role of *CIT* in skeletal muscle proliferation and differentiation, using C2C12 mouse myoblasts as a model. We hypothesize that disruption of *CIT* expression would inhibit the proliferation of skeletal muscle cells, possibly promoting their differentiation. Abscission failure caused by disrupted *CIT* expression and subsequent disorganization of midbody proteins could lead to the production of multinucleated cells, potentially resulting in withdrawal from the cell cycle and the initiation of differentiation. To test this hypothesis, we will conduct transfection experiments inhibiting *CIT* expression in C2C12 mouse myoblasts and assess its impacts on cellular proliferation, differentiation, and the expression of *CIT* and other genes of interest.

Overall, this dissertation research has improved our understanding of the genetic architecture of miRNA expression in skeletal muscle and created opportunities for further study to elucidate downstream impacts of miRNA regulation on complex traits in pigs. We conducted the first integrated miR-eQTL analysis in a livestock species concerning traits relevant to food production, identifying genomic regions associated with variation in the expression of 17 unique miRNAs. Many of these miRNAs had not previously been characterized for their roles in pig

skeletal muscle. Through target prediction, correlation, and co-localization analyses, we generated putative regulatory relationships between miRNAs, mRNAs, and phenotypes of interest comprising growth, carcass composition, and meat quality traits in pigs (Daza et al. 2021). These candidate miRNAs and genes serve as potential subjects of further study, with the overarching goals of improving our understanding of skeletal muscle tissue at the molecular level and promoting improvement in the quality and consistency of pork products.

REFERENCES

REFERENCES

1. D'Avino, P. P. (2017). Citron kinase – renaissance of a neglected mitotic kinase. *J. Cell Sci.* 130, 1701–1708.
2. Daza, K. R., Steibel, J. P., Velez-Irizarry, D., Raney, N. E., Bates, R. O., and Ernst, C. W. (2017). Profiling and characterization of a longissimus dorsi muscle microRNA dataset from an F 2 Duroc × Pietrain pig resource population. *Genomics Data* 13, 50–53.
3. Daza, K. R., Velez-Irizarry, D., Casiró, S., Steibel, J. P., Raney, N. E., Bates, R. O., et al. (2021). Integrated Genome-Wide Analysis of MicroRNA Expression Quantitative Trait Loci in Pig Longissimus Dorsi Muscle. *Front. Genet.* 12, 1–15.
4. Dema, A., Macaluso, F., Sgrò, F., Berto, G. E., Bianchi, F. T., Chiotto, A. A., et al. (2018). Citron kinase-dependent F-actin maintenance at midbody secondary ingression sites mediates abscission. *J. Cell Sci.* 131, 1–8.
5. Di Cunto, F., Calautti, E., Hsiao, J., Ong, L., Topley, G., Turco, E., et al. (1998). Citron Rho-interacting kinase, a novel tissue-specific Ser/Thr kinase encompassing the Rho-Rac-binding protein citron. *J. Biol. Chem.* 273, 29706–29711.
6. Ford-Speelman, D. L., Roche, J. A., Bowman, A., and Bloch, R. J. (2009). The Rho-Guanine Nucleotide Exchange Factor Domain of Obscurin Activates RhoA Signaling in Skeletal Muscle. *Mol. Biol. Cell* 20, 3905–3917.
7. Gai, M., Camera, P., Dema, A., Bianchi, F., Berto, G., Scarpa, E., et al. (2011). Citron kinase controls abscission through RhoA and anillin. *Mol. Biol. Cell* 22, 3768–3778.
8. Horton, J. S., Wakano, C. T., Speck, M., and Stokes, A. J. (2015). Two-pore channel 1 interacts with citron kinase, regulating completion of cytokinesis. *Channels* 9, 21–29.
9. Velez-Irizarry, D., Casiro, S., Daza, K. R., Bates, R. O., Raney, N. E., Steibel, J. P., et al. (2019). Genetic control of longissimus dorsi muscle gene expression variation and joint analysis with phenotypic quantitative trait loci in pigs. *BMC Genomics* 20, 1–19.